UNIT 11 INTERNET TECHNOLOGY

Structure

11.0 OBJECTIVES

After reading this unit you will be able to :

- understand the various components of Internet Architecture;
- know various Internet Protocols; and
- make yourself aware of security issues connected with the Internet.

11.1 INTRODUCTION

The Internet is the largest network the world has ever seen. Thousands of millions of people use it everyday. Technically the -Internet can be defined as a Transmission Control Protocol/Internet Protocol (TCP/IP)-bound network of networks using standard protocols for communication. Protocols are the rules that all the networks use to understand each other. The various protocols are sets of technical specifications that let computers exchange information, no matter what kind of computers they are, or what kind of technology hooks them together. Vendors of software and hardware want their products to be useful on the Internet, and so they make sure those products understand the Internet protocols and operate within them. The term *interoperability* has been coined to describe this ability of disparate types of hardware and software to work together under a common set of rules.

11.2 INTERNET ARCHITECTURE

The formal definition of a network is: “a data communications system that interconnects computer systems at various sites”. At its simplest, a network may consist of two computers or devices with a length of wire between them, letting them communicate. At its most complex, as in the Internet, a network is a globe-spanning, heterogeneous mix of computers. The Internet connects million of computers hooked to a number of heterogeneous networks.

Think of a corporate-wise network: each department has a LAN that allows it to share files and maybe a printer or two. Several departments, working together, interconnect their networks so that information may be shared more easily among the departments. These “regional” networks are interconnections based on geography (same city, same state, same group of states) or function (accounts-receivable grouped with accounts-payable into an accounting network, for example). Then the regional networks are connected together on to a corporate network, sometimes called a backbone. So, there is a user connected to a local Net; a local Net connected to a regional Net; and regional nets connected to a backbone (a backbone can be defined as a set of paths that local or regional networks connect to for long-distance interconnection).

The backbones connected to each other at physical network meeting points called gateways (a gateway is a network point that acts as an entrance to another network) illustrates the global Internet. We say “global” because networks from most countries with some sort of infrastructure are connected to it. Practically, this means people can use their computers on their local networks to send messages or access data from other computers located in another state, in another country, or in fact anywhere, that is connected to the Internet.

As mentioned above, the Internet links millions of computers for communicating data from one computer to another. A computer linked on the Internet is known as the host computer. The term “host” means any computer that has full two-way access to other computers on the Internet. The millions of host computer are linked on the Internet for communicating with each other. The connectivity from one computer to another computer is being provided using some standard mode of linkages called Internet Protocols. A protocol can be defined as special set of rules governing connectivity for telecommunication connections. Protocols may exist at several levels and in order to communicate both end points must recognise and observe standard protocols. Peer-to-Peer and client/server are two popular systems of communication.

11.2.1 Peer-to-Peer Communication

Peer-to-peer is a communications model in which each party has the same capabilities and either party can initiate a communication session. Other models with which it might be contrasted include the *client/server* model and the *master/slave* model. In some cases, peer-to-peer communication is implemented by giving each communication node both server and client capabilities.

On the Internet, peer-to-peer (referred to as P2P) is a type of transient Internet network that allows a group of computer users with the same networking programme to connect with each other and directly access files from one another's hard drives. *Napster* and *Gnutella* are examples of this kind of peer-to-peer software. Corporations are looking at the advantages of using P2P as a way for employees to share files without the expense involved in maintaining a centralized server and as a way for businesses to exchange information with each other directly. These are usually operated in small offices. IBM's Advanced Peer-to-Peer Networking (APPN) and Gnutellanet are the example of products that supports the peer-to-peer communication model.

11.2.2 Client-Server Architecture

The Client-Server Architecture is based on the principle where the client computer requests for some data and the data are sent by the server computer through the network. The concept of *client/server* computing has particular importance on the Internet because most of the programmes are built using this design. A server is a programme that "serves" (or delivers) something, usually information, to a client programme. A server usually runs on a computer that is connected to a network. The size of that network is not important in the client/server concept - it could be a small local area network or the global Internet.

The advantage of this type of design is that a server has to store the information in one format: which could be accessed by various clients working on multiple platforms and located at different places. In the client/server model, multiple client programmes share the services of a common server programme. Both client programmes and server programmes are often part of a larger programme or application.

In the case of the Internet, the Web browser is a client programme that requests services from a Web server. The server is designed to interact with client programmes so that people using the system can determine whether the information they want is there, and if so, have it sent.

11.2.3 Accessing the Internet

Let us see the model of centralized or cooperating utilities, such as the telephone or electricity. We can comfortably compare the Internet to one of these utilities. For example, there is a phone service in almost every part of India. A person who wants the telephone facility contacts a local area service provider (MTNL in case of Delhi). The service provider gives a "hook-up" from the residence or office to the service network. This arrangement allows you to connect to telephones almost anywhere in the world. The Internet where data moves among networks of computers works much the same way. In order to access the Internet, one requires the following:

- A computer with necessary client software
- The connectivity through Internet Service Provider (ISP) for flow of data
- Host computer(s) hosting the desired data

11.2.3.1 The Computer (The Client Side)

In order to connect to the Internet one needs a computer with necessary hardware and software devices for connecting. In order to have proper connectivity one needs a right mix of hardware and software. Depending on the need one can select hardware based on Pentium processor (Intel based) or Macintosh. These can further run on DIS, Windows, Unix, Linux OS/2 or such other operating systems. All these popular operating systems now have built-in support for connecting to the Internet. In order to access the data from server computers a large number of client software are available to suits various operating systems.

The hardware devices attached to the client computer also play a role in providing proper Internet connectivity. These can be either a modem or network connection. In order to connect on a Local Area Network, normally used in offices or universities/ colleges, there is a need to have Network Interface Card (NIC). These cards are designed to handle different speeds and network architecture. The details have been given in the network chapter. In order to connect from home or a small office, a modem is connected to a computer. A modem can be an external device or fitted inside the computer, i.e, internal modem. The modems come with different speeds, i.e. 14400 bits per second (bps), 28,800 bps or 58600 bps, etc. A modem provides connectivity to the external world through various types of communication lines.

By using web browsers we can locate servers of the Internet, send a query, process the query results, and display them using the tools familiar to us. A client programme is designed for a particular computing platform (for example, Windows, Macintosh, Unix) to take advantage of the strengths of the platform. The client software is designed to make you comfortable: it uses interface elements just like the ones you use to do word processing or a spreadsheet, or even to play a game. For example, a client programme used on the Internet to view web pages are called a browser such as Netscape, Internet explorer, etc.

11.2.3.2 The Connectivity

Each user can access the Internet through a connection on an existing network or via a modem (a device that allows the computer to use a telephone line to a remote network or ISP) from a remote site such as a private residence. The data and information that can be accessed on the Internet comes in numerous different formats and there is a wide range of applications that interpret the information for the user. The connections to the Internet fall under two basic categories: dial-up and dedicated

11.2.3.2.1 Dial-Up Connections

Dial-up connections to the Internet are not permanent connections. When you want Internet service, you dial-up to your service provider. When you are finished, you hang-up and your connection is broken. Of the dial up variety, there are two categories: *Analogue and Digital.*

a) **Analogue Dial-up Connections**

Analogue dial-up connections are the simplest and least expensive connections to make. The only hardware that's required besides the computer is a modem. The speed of an analogue dial-up connection is determined by the speed of the modem and the condition of the telephone line. Analogue, using normal telephone lines, refers to data transmissions that use a continuous wave form to transmit data. In this case, modems convert analogue signals to digital signals at the transmission end and convert digital signals back to analogue ones at the receiving end.

b) **Digital Dial-up Connections**

Digital transmissions, such as using fibre optic devices, pass data along using discrete, on/off pulses. This type of transmission does not require a modem at each end of the connection.

ISDN, which stands *for Integrated Services Digital Network,* is an example of digital dial-up service. ISDN is a set of protocols defining how data are transmitted over digital networks. Unlike the dial-up analogue service, ISDN offers a higher bandwidth and is capable of transmitting voice and data simultaneously on the same connect. Transfer speeds range from 64Kbps to 128Kbps or higher. ISDN service can be delivered over the same two copper wires that provide telephone service to your library. Therefore, no additional wiring will be needed in most cases.

11.2.3.2 Dedicated Connections

Dedicated connections differ from dial-up connections in that they are up and running 24 hours a day. Whether anyone is using them or not, the connection remains open. This type of connection is appropriate for organizations that transfer large amounts of data and have many users and workstations that must be connected to the Internet. This option requires that dedicated lines be leased through a network provider and special network hardware be installed on site, making this a complicated operation. Dedicated lines can be normal telephone lines, cables or radio frequency links.

11.2.4 Internet Service Providers (ISPs)

Internet Service Providers are companies which provide access to the Internet. This can be via a dial-up connection using a modem, or using a higher speed connection. Various charging levels may exist, but a popular method for home users is flat rate (per month unlimited time and data amount). Traditionally the Internet was purely a text-based global pool of information, and access was either limited or required a certain specialised knowledge. The development of the Internet today has ensured that information now comes in other formats such as graphical, audio and animated images, and the interface for such information is now a lot more dynamic and user friendly.

11.3 ORGANIZATION OF INTERNET

Internet links millions of computers for communicating data from one computer to another. A computer linked on the Internet is known as the host computer. The millions of host computers communicate with each other by using standard Internet protocols described in the last section. The data to be moved from one host to another host is broken into small pieces called packets. Each packet has a header with the address of destination host. The packets of different sizes move on various networks before reaching the destination. Various packets of one file may take different routes to reach a destination. The different networks on the Internet are connected with special purpose computers called routers. These routers look for destination address given on each packet and direct the packet to take the best route to the destination. Routers take their decisions based on information that is constantly reaching them from all over the Internet. They also hear from other routers about the links that are down or congested/slow, or about routers that are no longer accepting packets for certain destinations. Each packet's destination and proposed route is evaluated individually, in the blink of an eye, and sent off along the best route for that particular packet at that particular moment.

The same sort of decision-making is made for all packets that traverse the Internet. Each time a packet reaches a router, its address is examined and the packet is forwarded either to another router nearer its ultimate destination or to that destination if the router is the final router on the path. The destination computer is the one that unpacks and merges all the packets, throws away the "envelopes," and hands off the data.

11.3.1 Internet Protocol

To exchange information, computers must understand what each other computer is saying. They use a common language. We use a common language in class, called English. That is so because we can understand what is being said. A protocol is simply a set of conventions that determines how data will be transmitted from one point to another. This also determines how to move messages and handle errors; using them allows the creation of standards separate from a particular hardware system. The data on the Internet are transmitted from one computer to another using some standard protocols. A protocol can be defined as a formal description of formats and rules two or more computers must follow to exchange that data. These can be low-level details of computer-to-computer interfaces (for example the order in which bits from a byte are sent across a wire) or high-level exchange between application programmes (for example the way in which two programmes transfer a file across a network). In simple terms, protocols are a set of technical specifications that let computers exchange information.

11.3.1.1 Transmission Control Protocol, Internet Protocol (TCP/IP)

The two protocols predominantly used by the Internet are Transmission Control Protocol (TCP) and Internet Protocol (IP) and are popularly referred to as TCP/IP. These protocols are so common for the Internet that the definition of the Internet given by many experts says "Internet is a TCP/IP bound network of networks to access resources from one computer to another." These protocols were developed in 1974 by Robert Kahn of ARPANET and computer scientist Vinton G. Gerf. TCP/IP's great strength is that it easily enabled computers of different architectures and operating systems to communicate with each other.

TCP/IP is a two-layer program. The higher **layer, Transmission Control Protocol,** manages the assembling of a message or file into smaller packets that are transmitted over the Internet and received by a TCP layer that reassembles the packets into the original message. The lower layer, **Internet Protocol,** handles the **address** part of each packet so that it gets to the right destination. Each **gateway** computer on the network checks this address to see where to forward the message. Even though some packets from the same message are routed differently than others, they'll be reassembled at the common destination.

TCP/IP uses the **client/server** model of communication in which a computer user (a client) requests and is provided a service (such as sending a web page) by another computer (a server) in the network. TCP/IP communication is primarily point-to-point, meaning each communication is from one point (or **host** computer) in the network to another point or host computer. TCP/IP and the higher-level applications that use it are collectively said to be "stateless" because each client request is considered a new request unrelated to any previous one (unlike ordinary phone conversations that require a dedicated connection for the call duration). Being stateless frees network paths so that everyone can use them continuously. (Note that the TCP layer itself is not stateless as far as any one message is concerned. Its connection remains in place until all packets in a message have been received.)

The Internet Protocol (IP) is the method or **protocol** by which **data** are sent from one computer to another on the **Internet.** Each computer (known as a **host)** on the Internet has at least one **IP address** that uniquely identifies it from all other computers on the Internet. When you send or receive data (for example, an e-mail note or a Web page), the message gets divided into little chunks called packets. Each of these packets contains both the sender's Internet address and the receiver's address. Any **packet** is sent first to a gateway computer that understands a small part of the Internet. The gateway computer reads the destination address and forwards the packet to an adjacent gateway that in turn reads the destination address and so forth across the Internet until one gateway recognizes the packet as belonging to a computer within its immediate neighborhood or **domain.** That gateway then forwards the packet directly to the computer whose address is specified.

Since a message is divided into a number of packets, each packet can, if necessary, be sent by a different route across the Internet. Packets can arrive in a different order than the order they were sent in. The Internet protocol just delivers them. It's up to another protocol, the Transmission Control Protocol **(TCP),** to put them back in the right order.

IP is a connectionless protocol, which means that there is no continuing connection between the end points that are communicating. Each packet that travels through the Internet is treated as an independent unit of data without any relation to any other unit of data. (The reason the packets do get put in the right order is because of TCP, the connection-oriented protocol that keeps track of the packet sequence in a message.) In the Open Systems Interconnection **(OSI)** communication model, IP is in **layer 3,** the Networking Layer.

Another TCP/IP advantage is that it's not bound in any way to the physical medium. Whether it is wireless, token-ring, ordinary phone lines, LAN or other network, one can transmit data using TCP/IP.

11.3.1.2 Hypertext Transfer Protocol (HTTP)

The Hypertext Transfer Protocol (HTTP) is the set of rules for exchanging files (text, graphic images, sound, video, and other multimedia files) on the **World Wide Web.** Relative to the **TCP/IP** suite of protocols (which are the basis for information exchange on the Internet), HTTP is an application **protocol.**

Essential concepts that are part of HTTP include (as its name implies) the idea that files can contain references to other files whose selection will elicit additional transfer requests. Any **Web server** machine contains, in addition to the HTML and other files it can serve, an HTTP **daemon,** a programme that is designed to wait for HTTP requests and handle them when they arrive. Your web **browser** is an HTTP

client, sending requests to server machines. When the browser user enters file requests by either "opening" a web file (typing in a **Uniform Resource Locator)** or clicking on a **hypertext link,** the browser builds an HTTP request and sends it to the **Internet Protocol address** indicated by the URL. The HTTP daemon in the destination server machine receives the request and, after any necessary processing, the requested file is returned.

11.3.1.3 File Transfer Protocol (FTP)

FTP is a standard Internet **protocol.** It is the simplest way to exchange files between computers on the Internet. Like the Hypertext Transfer Protocol **(HTTP),** which transfers displayable web pages and related files, and the Simple Mail Transfer Protocol **(SMTP),** which transfers e-mail, FTP is an application protocol that uses the Internet's **TCP/IP** protocols. FTP is commonly used to transfer web page files from their creator to the computer that acts as their **server** for everyone on the Internet. It's also commonly used to **download** programmes and other files to your computer from other servers.

As a user, you can use FTP with a simple command line interface (for example, from the windows MS-DOS Prompt window) or with a commercial programme that offers a graphical user interface. Your web browser can also make FTP requests to download the programmes you select from a web page. Using FTP, you can also update (delete, rename, move, and copy) files at a server. You need to **log on** to an FTP server. However, publicly available files are easily accessed using **anonymous FTP.**

Basic FTP support is usually provided as part of a suite of programmes that come with TCP/IP. However, any FTP client programme with a graphical user interface usually must be downloaded from the company that makes it.

11.3.1.4 Serial Line Internet Protocol (SLIP) & Point-to-Point Protocol (PPP)

Personal computer users usually get to the Internet through the Serial Line Internet Protocol **(SLIP)** or the Point-to-Point Protocol **(PPP).** These protocols encapsulate the IP packets so that they can be sent over a dial-up phone connection to an access provider's modem.

SLIP is a **TCP/IP** protocol used for communication between two machines that are previously configured for communication with each other. For example, your Internet server provider may provide you with a SLIP connection so that the provider's server can respond to your requests, pass them on to the Internet, and forward your requested Internet responses back to you. Your dial-up connection to the server is typically on a slower serial line rather than on the **parallel** or multiplex lines such as a line of the network you are hooking up to.

PPP (Point-to-Point Protocol) is a **protocol** for communication between two computers using a **serial** interface, typically a personal computer connected by a phone line to a server. For example, your Internet server provider may provide you with a PPP connection so that the provider's server can respond to your requests, pass them on to the Internet, and forward your requested Internet responses back to you. PPP uses the Internet protocol **(IP)** (and is designed to handle others). It is sometimes considered a member of the TCP/IP suite of protocols. Relative to the Open Systems Interconnection **(OSI)** reference model, PPP provides layer 2 (data-link layer) service. Essentially, it packages your computer's **TCP/IP** packets and forwards them to the server where they can actually be put on the Internet. PPP is a **full-duplex** protocol that can be used on various physical media, including twisted pair or fiber optic lines or satellite transmission. It uses a variation of High Speed Data Link Control **(HDLC)** for packet encapsulation.

PPP is usually preferred over the earlier de facto standard Serial Line Internet Protocol **(SLIP)** because it can handle **synchronous** as well as **asynchronous** communication. PPP can share a line with other users and it has error detection that SLIP lacks. Where a choice is possible, PPP is preferred.

11.3.1.5 Z39.50

Z39.50 is an American National Standard for information retrieval (IR). Prepared by the National Information Standards Organisation (NISO), Z39.50 defines how one system can cooperate with other systems for the purpose of searching databases and receiving records. ANSI/NISO Z39.50-1995 (ISO 23950) is one of a set of standards produced to facilitate the interconnection of computer systems.. As a network protocol, the Z39.50 standard provides a set of rules that govern the formats and procedure used by computers to interact with one another. The standard establishes the permissible sequences of events at each of the two computer systems and specifies the content and structure of information parcels that are exchanged between systems.

The standard specifies formats and procedures governing the exchange of messages between a client and server, enabling the user to search remote databases, identify records which meet specified criteria, and to retrieve some or all of the identified records. It is concerned, in particular, with the search and retrieval of information in databases.

One of the major advantages of using Z39.50 is that it enables uniform access to a large number of diverse and heterogeneous information sources. Z39.50 does offer one TRUE interface to a variety of databases. Changing the interface is not frequent, but the beauty of this one is that it still remains the same of all servers! Some of these products are starting to offer UNICODE support, which will become increasingly important as we move into multi lingual record displays. They are functionally rich because this kind of product can support simultaneous searching of multiple databases. This is a very valuable feature in that it greatly compresses the amount of time required to sequentially query multiple databases.

11.3.2 Internet Addressing

Each host computer on the Internet has its own unique address. To identify a host on the Internet, three addressing systems have been evolved: A numerical system called IP addressing, a hierarchical naming system called the Domain Name System, and an addressing system called URLs, which are used for identifying sites on the web.

IP address : Each computer has a unique numerical address, such as 194.170.32.23

Domain name : Each computer must have a unique name, such as www.hct.ac.ae

Uniform Resource Locator : Address of file(s) to be accessible from a host computer

i) IP Addresses

Every host on the Internet is assigned a unique identifier called an *IP address or Internet protocol address.* The IP address is a numerical address consisting of four numbers separated by periods. An IP address looks like this: 128.86.8.7 and is read as, "128 dot 86 dot 8 dot 7."

The IP address is a set of numbers that expresses the exact physical connection between a computer and the network on the Internet. In some senses you can think of them in the same way you think about telephone numbers: a phone number uniquely describes your connection to the telephone network. IP addresses work somewhat similarly but are more complex than phone numbers because there are literally millions of network connections possible and because IP addresses are intended for use by computers rather than people. An IP address that looks like this:

154.135.186.235

An IP address consists of a 32-bit integer that's represented by four 8-bit numbers, written in base 10, separated by periods. IP addresses are organized from left to right with the left-hand octet describing the largest network organisation and the rightmost octet describing the actual network connection. IP addresses are unique on the network and allow it to know specifically which computer is to receive which electronic packet as well as from which specific computer the electronic packet came. The IP address of the Parliament of India computer is 164.100.24.3. As it may be quite difficult to remember such long IP addresses, a system to translate it in domain names has been developed. The Domain Name System (DNS) serves as a directory programme for IP addresses and it takes care of translating IP addresses to simpler English names. The DNS for the above mentioned IP address is parliamentofindia.nic.in

ii) Domain Names

A **domain** name locates an organization or other entity on the Internet. For example, the domain name *www.nic.in* or www.hindustantimes.com

The domains names have been designed in such a way that they broadly describe organizational or geographic realities. They indicate what country the network connection is in, what kind of organization owns it, and sometimes further details.

Servers or host computers have special names for each country. All countries in the world have a country suffix, except the USA. The UAE uses ae, New Zealand uses .nz, while Canada's is ca.

The domain name of a host computer looks like:

i) Host computer name

ii) organization name

iii) type of organization

iv) country name

IETF who designed the addressing system, have planned a system which looks like words. These words roughly map to a parallel system of numerical addresses called IP addresses. Every computer on the Internet has both a domain name and an IP address, and when you use a domain name, the computers translate that name to the corresponding IP address.

The server www.hct.ac.ae means

i) a host computer called www

ii) an organization called HCT

iii) an academic institution (ac stands for academic)

iv) located in the UAE (ae means Arab Emirates)

Similarly, the server **www.yahoo.com** means a host called www, belonging to an organization called yahoo, which is a commercial organization (com means commercial) located in the United States (if there is no country code then it is in the United States).

In actuality, a host computer uses only numbers, turning all domain name addresses into numbers. This translation process is taken care of behind the scenes by software. The reason domain names exist in the first place is because names are more convenient for people to use and easier to remember than numbers. For this reason, you are more apt to use domain names for addressing hosts than IP addresses.

A third level can be defined to identify a particular host server at the Internet address. In our example, “www” is the name of the server that handles Internet requests. (A second server might be called “www2”.) A third level of domain name is not required. For example, the fully-qualified domain name could have been “totalbaseball.com” and the server assumed.

Second-level domain names must be unique on the Internet and registered with one of the **ICANN-**accredited registrars for the COM, NET, and ORG top-level domains. Where appropriate, a top-level domain name can be geographic (currently, most non-U.S. domain names use a top-level domain name based on the country the server is in.). To register a U. S. geographic domain name or a domain name under a country code, see an appropriate registrar.

More than one domain name can be mapped to the same Internet address. This allows multiple individuals, businesses, and organizations to have separate Internet identities while sharing the same Internet server.

Top-level Domain names: On the Internet, a top-level domain (TLD) identifies the most general part of the **domain name** in an Internet address. A TLD is either a generic top-level domain **(gTLD),** such as **“com”** for “commercial,” **“edu”** for “educational,” and so forth, or a country code top-level domain (ccTLD), such as “fr” for France or "is" for Iceland.

A second-level domain (SLD) is the portion of a Uniform Resource Locator **(URL)** that identifies the specific and unique administrative owner associated with an **IP address.** The second-level domain name includes the top-level domain name. For example, in: whatis.com “whatis” is a second-level domain. “whatis.com” is a second-level domain name (and includes the top-level domain name of “com”). Second-level domains can be divided into further domain levels. These sub-domains sometimes represent different computer servers within different departments. More than one second-level domain name can be used for the same IP address

The top level domain names include country names known as Geographic Domains and type of organizations known as Non-Geographic Domains. The geographically based top-level domains use two-letter country designations. For example, .us is used for the United States,.ca for Canada (not California),.uk or .gb for the United Kingdom or Great Britain, and .il for Israel. Each domain has a number of hosts. A few more examples are given in the following table.

ABBREVIATION	MEANING
Au	Australia
Be	Belgium
Ge	Germany
Jp	Japan
Mx	Mexico
Nz	New Zealand
Uk	United kingdom

Non-Geographic Domains

There are six common top-level domain types that are non-geographical:

.com for commercial organizations such as netcom.com, apple.com, sun.com, etc.

.net for network organizations, such as internic.net

.gov for parts of governments within the United States, such as nasa.gov, Oklahoma.gov, etc.

.edu for organizations of higher education, such as sjsu.edu, ucsc.edu, mit,edu, etc.

.mil for non-classified military networks, such as army.mil, etc. (The classified networks are not connected to the wider Internet.)

.org for organizations that do not otherwise fit the commercial or educational designations, such as eff.org, farnet.org, etc.

The lowest level in the domain name system is the host name. Home names identify a computer on the Internet. In our example, the host's name is uafsysb, which is short for the University of Arkansas at Fayetteville, System B computer.

Host Name: In some instances, there are second-level domains delegated to organizations such as K-12 schools, community colleges, private schools, libraries, museums, as well as city and country governments. Examples of second-level domains are shown here:

CC - Community colleges
TEC - Technical colleges
LIB - Libraries
K12 - Kindergarten through 12th grade schools and districts
STATE-State Government
MUS- Museums

iii) **Uniform Resource Locator (URL)**

A URL (Uniform Resource Locator) (pronounced YU-AHR-EHL or, in some quarters, UHRL) is the address of a file (resource) accessible on the Internet. The type of resource depends on the Internet application protocol. Using the World Wide Web's protocol, the Hypertext Transfer Protocol **(HTTP),** the resource can be an **HTML** page (like the one you're reading), an image file, a programme such as a **common gateway interface** application or Java **applet,** or any other file supported by HTTP. The URL contains the name of the protocol required to access the resource, a **domain name** that identifies a specific computer on the Internet, and a hierarchical description of a file location on the computer.

On the Web (which uses the Hypertext Transfer Protocol), an example of a URL is:

http://www.mhrcc.org/kingston

Which describes a web page to be accessed with an HTTP (web browser) application that is located on a computer named www.mhrcc.org. The specific file is in the directory named /kingston and is the default page in that directory (which, on this computer, happens to be named index.html).

An HTTP URL can be for any web page (not just a home page) or any individual file.

11.4 INTERNET SECURITY

11.4.1 Introduction

We are in the midst of the Information Era! There has been an enormous information explosion, and the mushrooming popularity of the Internet puts huge amounts of information right at our fingertips. How safe is transmission of data on the Internet is a big question. Threats to privacy are being widely publicized from time to time. A major security hole was observed in Microsoft's Hotmail service that allowed anyone to view any hotmail member's mailbox by using a correctly configured URL that included the username, but not the password.

If one is running a business, one would probably want information to be easily accessible about products or services but certainly not other information like filling of tenders, prices quoted to various customers, etc. On the other side we have to ensure that data kept on our servers is not attempted by intruders or hackers and wrong messages are sent to the people visiting our site. Let us see how we can save our personal information and secure our system from hackers, intruders or virus attacks.

11.4.2 Password Protection

The storage space on a host computer is used as a temporary holding place for files before downloading them to the personal computer. There is a need to protect such files as these can be accessed easily through the Internet. This calls for protecting the system as a whole.

In this respect, the first level of security mechanism is your password. Passwords are an important first line of defence against intruders and it is in everyone's best interest to use this principal mode of protection. You don't want to be the weak link in a chain because of a poorly chosen password. If an intruder can break into your host's system using your password, he or she may then be able to find other security holes in that system or use it as a means of entry into other systems on the Internet.

When you establish an Internet account, you should follow some basic rules when choosing a password. You make it easy for a cracker to decipher your password if you choose one that is easy to remember, like your first name, your spouse's first name or a pet's name. To help make your account more secure, consider the following principles when creating and using your password:

- Don’t use your name or a modification of your name for your password.
- Don’t use a word or modification of a word that occurs in any dictionary.
- Don’t use an acronym.
- Once you’ve created an account password, don't share it with anyone.
- Change your password often, at least every three months.
- Don’t leave your terminal unattended when you're logged in.
- Don’t write your ID and password on a piece of paper or send it to friends via e-mail.

11.4.3 Computer Viruses

A computer virus is a computer programme that infects your computer applications or system files. When the virus becomes active, it can destroy data on any computer. The virus does this by getting into your computer's memory and from there it can copy itself to its hard disk or floppy disks.

The computer virus is a piece of programming code usually disguised as something else that causes some unexpected and usually undesirable event. A virus is often designed so that it is automatically spread to other computer users. Viruses can be transmitted as attachments to an e-mail note, as downloads, or be present on a diskette or CD. The source of the e-mail note, downloaded file, or diskette you've received is often unaware of the virus. Some viruses wreak their effect as soon as their code is executed; other viruses lie dormant until circumstances cause their code to be executed by the computer. Some viruses are playful in intent and effect (“Happy Birthday, Ludwig!”) and some can be quite harmful, erasing data or causing your hard disk to require reformatting.

Viruses are inactive until you execute an application that's infected. You can also activate a virus by starting up your computer with a floppy disk that’s infected by a boot sector virus. Going online and downloading an infected programme isn't harmful. Only when you execute the infected programme can the virus infect your PC. Generally, there are three main classes of viruses:

File infectors: Some file infector viruses attach themselves to programme files, usually selected .COM or .EXE files. Some can infect any programme for which execution is requested, including .SYS, .OVL, .PRG, and .MNU files. When the programme is loaded, the virus is loaded as well. Other file infector viruses arrive as wholly-contained programmes or scripts sent as an attachment to an e-mail note.

System or boot-record infectors: These viruses infect executable code found in certain system areas on a disk. They attach to the DOS **boot sector** on diskettes or the **Master Boot Record** on hard disks. A typical scenario is to receive a diskette from an innocent source that contains a boot disk virus. When your operating system is running, files on the diskette can be read without triggering the boot disk virus. However, if you leave the diskette in the drive, and then turn the computer off or reload the operating system, the computer will look first in your A drive, find the diskette with its boot disk virus, load it, and make it temporarily impossible to use your hard disk.

Macro viruses: These are among the most common viruses, and they tend to do the least damage. Macro viruses infect your Microsoft Word application and typically insert unwanted words or phrases.

The best protection against a virus is to know the origin of each programme or file you load into your computer or open from your e-mail programme. Since this is difficult, you can buy **anti-virus software** that can screen e-mail attachments and also check all of your files periodically and remove any viruses that are found.

Anti-virus software greatly reduces the chances of experiencing a viral infection in your computer. Most of these programmes will scan all of the files currently on the hard disk to see whether they contain any viruses. They also install memory-resident programmes that keep a continuous look out for any suspicious activity that would indicate a virus.

Today, there are thousands of known viruses and new viruses are being developed daily. To provide yourself with the best protection possible, you should update your virus protection software periodically.

11.4.3.1 How to avoid Viruses

If you share floppies with others or download files from online services, there's no 100% guarantee that you'll always be protected against viruses. To help reduce the risk of your computer being infected, follow these tips:

- Run an anti-virus programme and keep it updated often.
- Scan floppies that you suspect might be infected.
- Don't copy programmes from one computer to another. Use the original distribution diskettes to install programmes.
- Scan your system regularly with the full scanning engine.
- Avoid using floppies from unknown sources.
- Write-project your floppies by covering the notch on 5.25-inch disks or by sliding the little tab to expose the hole on 3.5-inch disks.
- Never boot your computers from unknown diskettes. If you do and you suspect there may be a virus on the diskette, shut the computer down. Boot up from a clean system diskette and check the system with an anti-virus programme.

Utilize your anti-virus programme's memory resident scanners to check all files as they are accessed, even from the Internet.

11.4.4 Spyware

Spyware is software planted on your computer to retrieve and forward information about you to outside systems. The information collected can range from a survey of your surfing habits passed along to advertisers and marketers to the passwords and credit card numbers you type passed along to crackers who may exploit it. The software can be planted through e-mail or even included in software you obtain and install for other purposes, such as Gozilla and CuteFTP. It can also record information about what you do while you are there, and perhaps collect further information from you. Many sites want you to register in order to use their services. You may be asked your real name and e-mail address, your home or business address, your telephone number, your income level, your interests, and so on. This can be valuable information for running a business. The information is voluntary, of course, and they have no way to tell if you are faking the information.

Think twice before giving out such information freely. Think about whom you are giving it to and what uses they could put the information to.

To defend yourself against spyware, always maintain current anti-virus software on your system, and be very careful about installing software from unknown sources.

11.4.5 Firewalls

Is your computer at home or office safe from crackers and hackers? If you use a dial-up connection, you probably are reasonably safe. If you use a full time network connected to the Internet or use a full time connection like cable modem or Digital Subscriber Line (DSL), you may be quite vulnerable. At any time, thousands of automated programmes are running on the Internet just looking for vulnerable computer systems. As a result, your computer is probably being probed repeatedly during the day.

In order to secure your computer from receiving such unwanted programmes and prevent intrusions, firewalls may be installed. To understand the concept, let us take an example from old cars. In early motor cars, steel plates where placed between the engine and the driver seat. If there was a fire or explosion, the steel plates stopped damage to the driver. The plates helped stop the spread of flames or prevent engine parts from flying in to the driver. The steel plates are a wall between the engine and the driver. They protect the driver from damage. This is a protecting wall in cars and is called a firewall. Computers are like the drivers, and the Internet is like the engine. Computers can be damaged by people using the Internet. A firewall protects the computers from damage.

A firewall is a set of related programmes, located at a network **gateway server,** that protects the resources of a private network from users from other networks. (The term also implies the security policy that is used with the programmes.) An enterprise with an **intranet** that allows its workers access to the wider Internet installs a firewall to prevent outsiders from accessing its own private data resources and for controlling what outside resources its own users have access to.

There are a number of firewall screening methods. A simple one is to screen requests to make sure they come from acceptable (previously identified) **domain name and Internet Protocol** addresses. For mobile users, firewalls allow **remote access** into the private network by the use of secure logon procedures and authentication certificates.

A firewall refers to the concept of a security interface or gateway between a closed system or network and the outside Internet that blocks or manages communications in and out of the system. Basically, a firewall, working closely with a **router** programme, examines each network **packet** to determine whether to forward it towards its destination. A firewall also includes or works with a **proxy server** that makes network requests on behalf of workstation users. A firewall is often installed in a specially designated computer separate from the rest of the network so that no incoming request can get directly at private network resources. A firewalls security may be provided by passwords, authentication techniques, software, and hardware, i.e., a firewall is a device that protects a network (group of computers) from outside interference. It can be a hardware or software device A firewall is a router that restricts Internet access by only allowing access to certain computers (and specified services on those host computers) within the organisation. A firewall protects the company computers. It can also stop company workers from accessing the Internet!

A number of companies make firewall products. Features include logging and reporting, automatic alarms at given thresholds of attack, and a graphical user interface for controlling the firewall.

11.4.6 Proxy Servers

In an enterprise that uses the Internet, a proxy server is a **server** that acts as an intermediary between a workstation user and the Internet so that the enterprise can ensure security, administrative control, and caching service. A proxy server is associated with or is part of a **gateway** server that separates the enterprise network from the outside network and a **firewall** server that protects the enterprise network from outside intrusion.

A proxy server receives a request for an Internet service (such as a web page request) from a user. If it passes filtering requirements, the proxy server, assuming it is also a **cache server,** looks in its local **cache** of previously downloaded web pages. If it finds the page, it returns it to the user without needing to forward the request to the Internet. If the page is not in the cache, the proxy server, acting as a client on behalf of the user, uses one of its own IP addresses to request the page from the server out on the Internet. When the page is returned, the proxy server relates it to the original request and forwards it on to the user.

When a firewall is used to stop company workers from accessing the Internet, a proxy server is used to provide access. It also acts as a security device by providing a buffer between inside and outside (on Internet) computers. The functioning of a typical proxy server is given below:

i) When a client wants a file, this request is sent to the proxy server.

ii) The proxy server contacts the web server to get the file.

iii) The proxy server keeps a copy of the file.

iv) The proxy server sends the file back to the client.

If another client wants the same file, the proxy server knows it already has a copy. It does not contact the web server. The proxy server sends the copy of the file to the client. Proxy servers save Internet traffic. If one hundred clients want the same file, the proxy server gets the file once. It keeps a copy of the file in cache. The copy is given to all the clients. If you did not use a proxy server, each client would download their own file from the Internet. This would cost a lot more money. You have to pay for using the Internet. The more you download, the more money you pay.

The functions of proxy, firewall, and caching can be in separate server programmes or combined in a single package. Different server programmes can be in different computers. For example, a proxy server may be in the same machine with a firewall server or it may be on a separate server and forward requests through the firewall. To the user, the proxy server is invisible; all Internet requests and returned responses appear to be directly with the addressed Internet server.

Self Check Exercise

1) Define IP address, Domain name and Host.

2) State the need for Internet Security.

....

....

....

....

....

....

11.5 SUMMARY

This unit discussed specifically the technology associated with the Internet such as Internet architecture, organization of Internet (Internet protocol and Internet addressing) and, more important, Internet security. Internet architecture included peer-to-peer communication, client/server Architecture, accessing Internet and Internet service providers. Organization of Internet encompass Internet Protocol (TCP/IP, HTTP, FTP, SLIP and PPP, Z39.50) and Internet protocol address (IP Address), and Domain names finally the most important aspect is Internet security which includes password protection, computer viruses, spyware, fire walls, proxy servers, etc.

11.6 ANSWERS TO SELF CHECK EXERCISES

1) IP address is the Internet Protocol (IP) address given to every computer connected to the Internet. An IP address is needed to route information much like a street address or PO box is needed to receive regular mail. Example: 66.46.181.116.

   Domain name is a text name which a computer network registers. The domain name is used to give computers text names rather than using the numeric IP addresses. This is like getting a vanity phone number that spells out a word to make it easy to remember. Domain name examples: consumer.net uu.net consumer-info.org.uk.

   Computer (host) name are names given to individual computers. Each host name corresponds to an IP address. Host names and domain names are optional and everything will work fine using just IP addresses. Examples of host names: www.consumer.net mail.consumer.net Cust149.tnt3.sfo3.da.UU.net

2) The need for Internet security is to protect from password computer viruses file infectors, system infectors, spyware, firewalls and proxy servers.

11.7 KEYWORDS

Master Boot-record : A small programme that is executed when a computer boots up.

Radio Frequency : Any frequency within the electronic magnetic spectrum associated with radio wave propagation. Many wireless technologies are based on RF field propagation.

Token Ring : A type of computer network in which all the computers are arranged (schematically) in a circle. A token, which is a special bit pattern, travels around the circle. To send a message, a computer catches a token, attaches a message to it, and then lets it continue to travel around the network.

11.8 REFERENCES AND FURTHER READING

Dawson, A.(1997).s The Internet for Library and Information Professionals, London: Library Association Publishing.

Dern, Daniel. (1994)The Internet Guide for New Users, New York: McGraw Hill.

Mehta, Subhash.(1996) Understanding and using Internet. Delhi, Global Business Press.

Nair, R. Raman.(2002) Accessing Information through Internet. New Delhi, Ess Ess Publications..

Randall, Neil. Teach Yourself the Internet in a Week. New Delhi: Prentice-Hall of India private Limited, New Delhi.