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