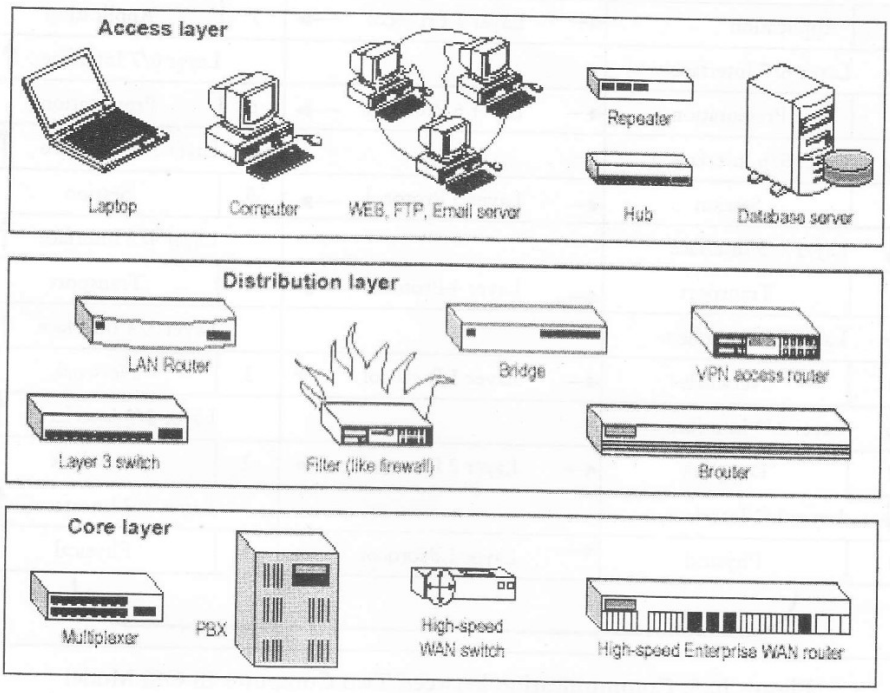


Source: http://www.mcmcse.com/cisco/guides/hierarchical_model.shtml

Figure 10.6: Displays the Three Layer Model

On the implementation of these layers, each layer can work with multiple layers.

10.3.3 TCP/IP Reference Model

The TCP/IP model is considered the oldest protocol of all computer networks like the ARPANET and its successor Internet. TCP/IP stands for Transmission Control Protocol/Internet Protocol. It was developed with the objective to specify a suite of protocols capable of providing transparent communications interoperability services between computers of all sizes, regardless of the hardware or operating system platforms supporting them. Over the years, TCP/IP has become the most widespread of today's protocols. One reason for TCP/IP's popularity is the public availability of its protocols' specifications. In this sense, TCP/IP can justifiably be considered an open system. Most users rely on TCP/IP for the purpose of file transfers, electronic mail (e-mail), and remote login services.

The TCP/IP model was aimed to connect multiple networks together in a seamless way even in case of breakdown of the subnet hardware. Not only providing seamless communication, but also providing a flexible architecture that should support applications with divergent requirements, ranging from transferring files to real-time speech transmission. These objectives could be achieved because of the inclusion of the research work on packet-switching network to the ARPnet.

TCP corresponds to the fourth layer of OSI reference model. IP corresponds to the third layer of the same model. TCP provides a connection type service. That is, a logical connection must be established prior to communication to continuously transmit large amount of data with acknowledgement. IP is a connectionless type service and prior to transmission of data, no logical connection is needed.

TCP/IP defines a suite of communications and applications protocols in layer structure, with each layer handling distinct communication services. TCP/IP defines a four-layer model as shown in Figure 10.7 consisting of the internet layer, the transport layer, the application layer and the host-to-network layer. This architecture is based on three sets of interdependent processes, namely, application-specific processes, host-specific processes, and network-specific processes.

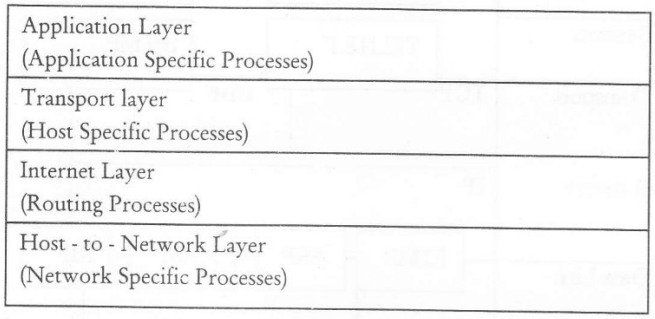


Figure 10.7: TCP/IP Communication Architecture

Internet layer

The packet format and protocol at this layer is called Internet Protocol (IP). IP is a connectionless type service that introduces IP packets into any network. The packets travel independently to the destination. Prior to transmission of data, no logical connection is needed. The TCP/IP Internet layer corresponds to the network layer of the OSI reference model in functionality, as shown in Figure 2.6.

Transport layer

The transport layer of TCP/IP model corresponds to the transport layer of the OSI reference model as shown in Figure 10.8. It is represented by two end-to-end protocols namely, TCP (Transmission Control Protocol) and UDP (User Datagram Protocol). TCP is a reliable connection-oriented protocol and UDP is an unreliable connectionless protocol.

Application layer

The TCP/IP model was the first of its kind model and therefore did not contain session or presentation layers because of its little use to most of the applications. This layer has all the higher-level protocols, as shown in Figure 10.8.

Host-to-network layer

The layer below the Internet layer is not defined and varies from host and network to network. The TCP/IP model suggests that the host has to connect to the network using some protocol so it can send IP packets over it.

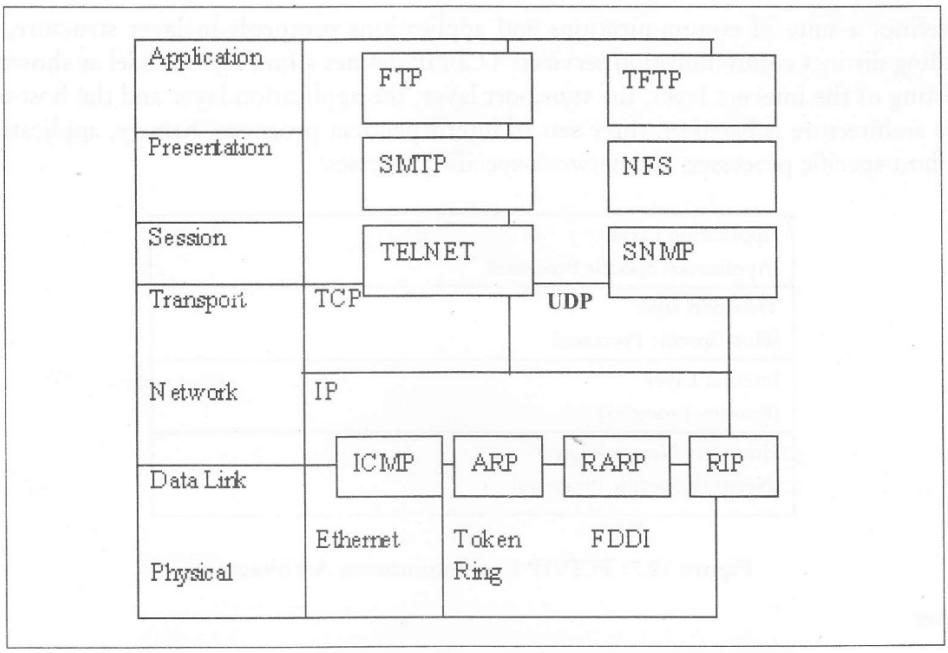


Figure 10.8: Correspondence: TCP/IP Model and the OSI Model

Comparison of the OSI and TCP/IP Reference Models

Figure 10.8 shows the similarity between the TCP/IP and OSI reference model. Both the models were developed based on the concept a stack of independent protocols with similar functionality of the layers.

In spite of similarity between the two models they also contrast in functionalities provided by services, interfaces and protocols. OSI reference model clearly distinguish them while the TCP/IP model did not explicitly distinguish them. Other differences are:

- The OSI model has seven layers and the TCP/IP model has only four layers.
- The OSI model was developed before the protocols were devised. The TCP/IP model was developed after the development of the protocols.
- The OSI model has both connection-oriented and connectionless communication in the network layer and connection-oriented communication in the transport layer. The TCP/IP model supports connectionless mode in the Internet layer and both modes in the transport layer.

10.4 INTERNETWORKING PROTOCOLS

Computers are connected by many different technologies. A network is a system of two or more interconnected computers either in a peer to peer or client to server fashion, most often over a shared and virtual connection. In other words, networks provide the connections between computer resources in order to accommodate the flow of information. This is in direct contrast to the old terminal to host hardwired connection. A network can support terminal to host connections via terminal emulators or terminal servers, and provides greater flexibility in switching connections. The

downside of this exploding information sharing is rather annoying situation when one computer wants to extend its information system to another computer but the latter happened to have a different network technology with different network protocols. As a result, even if they could agree on a type of network technology to interconnect the two computers at different locations, their applications still would not be able to communicate with each other because of the different protocols.

The following points can justify the need for networks:

- Networking allows sharing of resources.
- Reliability - It is also an offshoot of sharing, when one computer breaks down, you can use other computer available on the network. This could be possible because there is no central computer as in the case of mainframe.
- Networks allow us to be mobile.

The term networking applies to:

- the exchange of information among computers of individuals, groups, or institutions
- the process of electronic voice or data communications

Following three basic components are mandatory to implement a network:

- Hardware
- Protocols (software)
- Applications (useful software)

Over the years, TCP/IP has become the most widespread of today's protocols. One reason for its popularity is the public availability of its protocols' specifications. In this sense, TCP/IP can justifiably be considered an open system. Most users rely on TCP/IP for the purpose of file transfers, electronic mail (e-mail), and remote login services. This chapter also discussed how TCP/IP, a common protocol method is used to interconnect computers together, and also serve as the default protocol for accessing information over the Internet. TCP/IP is the protocol used by computers on the Internet and may be considered as two separate protocols such as TCP and IP. Each computer has an IP address. A protocol is a set of rules that govern how computers talk to each other. With TCP/IP, different computer systems can reliably exchange data on an interconnected network. It also provides a consistent set of application programming interfaces (API's) to support application development. This means that software program can use TCP/IP to exchange data. An example of this is web browsers, software applications that use TCP/IP to exchange data.

10.4.1 TCP/IP Protocols

TCP/IP stands for Transmission Control Protocol/Internet Protocol. It was developed with the objective to specify a suite of protocols capable of providing transparent communications interoperability services between computers of all sizes, regardless of the hardware or operating system platforms supporting them. Over the years, TCP/IP has become the most widespread of today's protocols. One reason for TCP/IP's popularity is the public availability of its protocols' specifications. In this sense, TCP/IP can justifiably be considered an open system. Most users rely on TCP/IP for the purpose of file transfers, electronic mail (e-mail), and remote login services.

10.4.2 IP Protocol

In contrast to TCP, it is a connectionless type service and operates at third layer of OSI reference model. That is, prior to transmission of data, no logical connection is needed. This type of protocol is suitable for the sporadic transmission of data to a number of destinations. It does not have such functions as sequence control, error recovery and control, flow control but it identifies the connection with port number. The IP datagram has a header of 20-byte fixed size and a text of variable length optional parts. The header format of IP datagram is depicted in Figure 10.9. The header format is transmitted from left to right, with the high order bit of Version field is transmitted first.

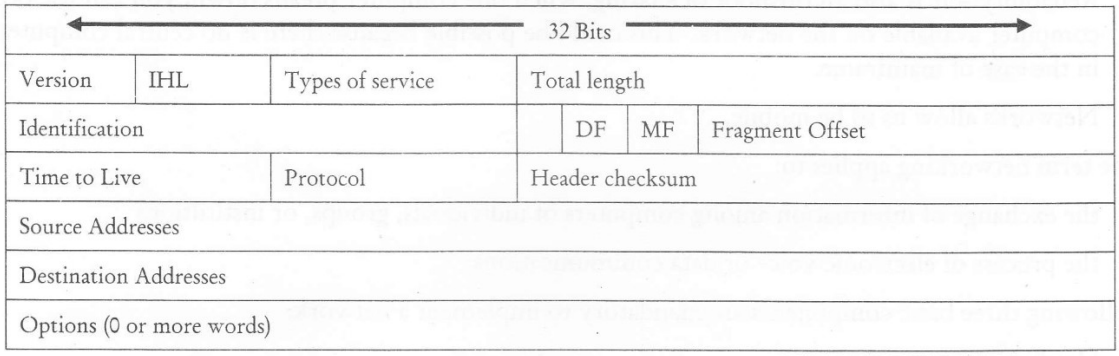


Figure 10.9: IP (Internet Protocol) Header

Data encapsulation adds the IP header to the data. The IP header consists of five or six 32-bit words; the sixth word is attributed to the IP options field. The different fields of the IP header are given as below:

- Version refers to the version of the IP protocol in use and keeps track of the version of the protocol to which the datagram belongs to. The current version of IP is 4.
- Internet Header Length (IHL) indicates the length of the header field in 32-bit words. The minimum value of the header field is 5 that apply when no option is present. The maximum value of this 4 bit filed is 15 that restricts the header to 60 bytes and thus Option field to 40 byte.
- Type of service enables the host to indicate the subnet what kind of service (e.g., reliability and speed) it wants. It refers to any of the type of services that IP supports. Desired service type is normally specified by user level applications. Examples of service type include minimum and maximum throughput, requested by applications such as the File Transfer Protocol (FTP) and Simple Mail Transfer Protocol (SMTP).
- Total length has everything in the datagram (max. 64 KB). If it is subtracted from the IHL field, it indicates to IP the actual length of the data field.
- Identification enables the destination host to determine which datagram a newly arrived fragment belongs to.
- DF means Do not Fragment.
- MF is for More Fragments.
- Fragment offset indicates the source location of the current datagram. The elementary fragment unit size is 8 bytes.

- Time to live that counts hops is expressed in seconds. A zero count indicates that the packet is discarded. TTL is employed by IP to prevent a lost datagram from endlessly looping around the network. IP achieves this objective by initializing the TTL field to the maximum number of routers that the packet can traverse on the network. Every time the datagram traverses a router, IP decrements the TTL field by 1.
- Protocol indicates the destination which transports process to give the datagram to (TCP, UDP, or others).
- Header checksum verifies the header only. The algorithm is to add up all the 16-bit halfwords as they arrive, using one's complement arithmetic.
- Source/Destination address tells the network number and host number.

Options provides an escape to allow subsequent versions of the protocol to have information not present in the original design, to allow experimenters to try out new ideas, and to avoid allocating header bits to information that is rarely needed. On its presence, it includes optional control information. An example of optional information includes the route record, which includes a record of every router that the datagram traversed during its trip around the network.

IP Addresses

Using Internet has become common. We will now understand how Internet interprets the Internet address. The Internet addresses are written as www.hotmail.com, for the instance we write one more address as server.institution.domain. The address www.hotmail.com is not actual address; it is a text version of the Internet address, which is basically a binary representation. Now we compare www.hotmail.com, and server.institution.domain. WWW is the name of the server owned by the institution (in this case, it is hotmail) and this server is connected to the Internet to a domain server namely (com in this case) which maintains a database of the addresses of different servers using the same domain com. The domain name has no geographical relevance and two sites with same domain name may exist at two end of this world.

The above case is the simplest case. In another instance an organization may be large enough and have several other servers for different purposes such as web server, email server, print server etc. Suppose we now take an example www.sun.planet.universe.in. This address has five parts separated by three dots. If we try to understand this address, this address will indicate that a group Planets (planet) comes under a Universe sub domain which is a part of India domain and maintaining one server sun out of many servers, which is linked to Internet through its web server. Likewise any organization with several departments may create addresses for its sub domain with different servers being maintained there.

Internet is the collection of several independent networks, which are interconnected with one another. Now each independent network may have several hosts. Keeping this in mind, you can now think of address of your house. Your house has a unique house number, which is not assigned, to any other house in your locality. In this case, your house can be considered as a host. Your locality can be considered as network and your city as domain. You can write your address in Internet addressing notation as houseno.locality.city. If suppose you want to tell your address to a foreigner, then you will have to add your country name in your address. In this case it will become houseno.locality.city.country. Now if anybody desires to send you a letter or visit your house, he will

first has to come to your country and then to your city. After that he will reach to your locality and then your house by your house number. The same analogy applies in case of Internt addressing.

We have already noted that a host on Internet has two parts. These are identification of the network and identification of the host on the network. In this manner, the address of a host is therefore comprised of two parts namely network address and host address. These two parts together make 32 bit long IP address for a particular host on the Internet. The IP address, which will see in the subsequent discussion, is written in four octets each separated by a dot. It may have a form like 197.23.207.10. Presently, we are using IP address version 4 (IPv4). However, IP address version 6 (IPv6) is gradually under implementation stage.

IPv4 Addressing

IPv4 addresses are uniquely used as identifiers, which work at network layer to identify the source or destination of IP packets. Presently, the version of IP, which is in use, is called as IPv4. In this version, every node on Internet may have one or more interfaces, and we are required to identify each of these devices with a unique address assigned to each of them. It means that each node is assigned one or more IP addresses to invoke TCP/IP. These are logical addresses and have 32 bits.

Technically, IP addresses are expressed using binary notation with 32 bit long string. In order to make these strings to remember easily, dotted decimal notations are used, in which periods or dots separate four decimal numbers from 0 to 255 representing 32 bits. As there are 32 bits therefore each decimal number contains 8 bits and called octet.

For example, the IPv4 address 11000000101010000000101000011001 is expressed as 192.168.10.25 in dotted decimal notation. Below are the steps to convert an IPv4 address from binary notation to dotted decimal notation:

1. Break 32 bit long address into segments of 8-bit blocks: 11000000 10101000 00001010 00011001
2. Write decimal equivalent of each segment: 192 168 10 25
3. Separate the blocks with periods: 192.168.10.25

Figure 10.10 shows the IP address structure.

11000000	10101000	00001010	00011001
192	168	10	25

Figure 10.10: IP address in Dotted Decimal Notation

Dotted Decimal Notation

We have seen that IPv4 address is expressed as a 32-bit number in dotted decimal notation. IP addresses may have fixed part and variable part depending upon the allocation of total addresses to you or your organization. Fixed part of the address may be from one octet to three octets and remaining octets will then be available for variable part. An IPv4 address is assigned using these parts. All bits in the fixed octet (s) are set to 1 while variable octet(s) are set to 0 bits. Thereafter, convert the result into dotted decimal notation. For example, you may take an IP address as 192.168.10.25. Now set all fixed bits to 1 and set all variable bits to 0. This gives 11111111 11111111 00000000 00000000. On converting it in dotted decimal notation, the result is 255.255.0.0. This dotted decimal notation with fixed and variable parts is used as address prefix to 192.168.10.25 and is expressed as 192.168.10.25, 255.255.0.0. This way

of expressing the prefix length as a dotted decimal number is known as network mask or subnet mask notation.

Classification of IPv4 Addresses

Internet standards allow the following addresses:

- *Unicast:* It is assigned to a single network interface located on a specific subnet ad facilitates one-to-one communication. This is unique address globally for the identification of a device on the network. It may be understood as the house number on a particular locality. It includes a subnet prefix and a host ID portion.
- *Subnet prefix:* The subnet prefix is basically network identifier or network address portion of an IP unicast address. It should be noted that all nodes on the same physical or logical subnet must use the same subnet prefix, which eventually becomes unique within the entire TCP/IP network.
- *Host ID:* The host ID, which is a host address portion of an IP unicast address, identifies a network node to which some devices are interfaced. It is also unique within the network segment.
- *Multicast:* It is used for one or more network interfaces located on various subnets. It allows one-to-many communication. It delivers single packets from one source to many destinations. These addresses are part of Class D addressing scheme.
- *Broadcast:* It is allocated to all network interfaces located on a subnet and is used for one-to-everyone on a subnet communication. It delivers packets from one source to all interfaces on the subnet. Broadcast addresses may be further classified as network broadcast, subnet broadcast, all subnets directed broadcast and limited broadcast.

Internet Addresses are further classified into different classes. It is based on the number bits are used for the address prefix of a single subnet and the number of bits are used for the host ID. It therefore allocates the number of networks and the number of hosts per network.

Subnetting for IP Addresses

Over the past several years, the Internet has scaled enormous volume in terms of hosts connected to it and therefore IPv4 addresses yet available are becoming scare. You may have confusion here that 32 bits give 2^{32} unique addresses which comes around 4.3 billion different addresses. But this not the condition because of the different classes of the IPv4 addresses. Suppose a medium sized organization gets Class B address based on its current user population of say 1000. It uses 1000 different addresses. But the organization management has the ability to assign 2^{16} = 65,536 different identifiers. It means that there is 64,536 addresses wastage. Since they all belong to the same class B network number, they cannot be reclaimed by any other organization. A network administrator may suggest using Class C network address, which may require at least four class C networks. Later on, suppose, the number of users increase and the organization applies for another class C network, it might not get the same or if it gets, it has to pass through a hell of paper works and delays. In addition there is another angle of this problem with regard to additional routing. With many Class C networks, you need to have more network number for routers to track. Consequently performance of the network deteriorates. The solution of these problems lie either in increasing the number of bits in IP address or Classless Inter Domain Routing (CIDR).

We may also use a technique called subnetting to efficiently divide the address space allocated to an organization to the different users divided among different subnets of an organization network.

Therefore subnetting is a process through which the address space of a unicast address prefix is efficiently divided for allocation among the subnets of an organization network. As we know that a unicast address have fixed and variable portions. The fixed portion of a unicast address prefix has a defined value. The variable portion of a unicast address prefix has the bits beyond the prefix length, which needs to set to 0. Subnetting uses the variable portion of a unicast address prefix for assignment to the subnets of an organization network.

In order to implement subnetting, you need to follow the some guidelines:

- Assess the number of subnets requirement.
- Assess the number of host IDs for each subnet.
- After this, a set of subnetted address prefixes with a range of valid IP addresses may be defined. Following steps are followed for Subnetting:
- Estimate the number host bits for the subnetting.
- Determine the new subnetted address prefixes.
- Determine the range of IP addresses for each new subnetted address prefix.

We may now learn as to how the subnet prefix of an IP address is determined. Following steps give you a way to determine the same without the use of binary numbers:

1. Write the number n (the prefix length) as the sum of 4 numbers by successively subtracting 8 from n. For example, 22 is 8+8+6+0.
2. In a table with four columns and three rows, place the decimal octets of the IP address in the first row. The second row will then contain the four digits of the sum as has been determined in step 1.
3. The columns having 8 in the second row, write the corresponding octet from the first row to the third row. In case of 0 in a column in the second row, place 0 in the third row.
4. The column in the second row having a number between 0 and 8, convert the decimal number in the first row to binary. Now select the high-order bits for the number of bits indicated in the second row and put zero for the remaining bit and then convert back to decimal number. This will be the entry in that column. For our example the entry in third column of first row is 10. Therefore the binary equivalent is 00001010. Again the third column of second row is having 6. It means we have to take 6 bits as such from high bi side and converting the remaining two bits as 00. This will give us a binary number as 00001000 which is decimal equivalent to 8. Therefore, the entry 8 will go in that column.

192	168	10	25
8	8	6	0
192	168	8	0

This gives the subnet prefix for the IPv4 address configuration 192.168.10.25/22 as 192.168.204.0/22.

Now, we have to extract the subnet prefix from an arbitrary IPv4 address using an arbitrary subnet mask. For this purpose a mathematical operation logical AND is used. A logical comparison between the 32-bit IP address and the 32-bit subnet mask is performed. It gives the subnet prefix. For example, we may consider the following possible addresses for Class C.

Class C Network	Bit Representation	Address Range
210.195.8.0	11010010-11000011-00001000-xxxxxxxx	210.195.8.0-211.195.8.255
210.195.9.0	11010010-11000011-00001001-xxxxxxxx	210.195.9.0-211.195.9.255
210.195.10.0	11010010-11000011-00001010-xxxxxxxx	210.195.10.0-211.195.10.255
210.195.11.0	11010010-11000011-00001011-xxxxxxxx	210.195.11.0-211.195.11.255

These Class C networks define the contiguous set of addresses from 210.195.8.0 to 210.195.11.255. On examining these addresses, it is observed that the first 22 bits are same for each address. It means that any of these Class C networks has 22 bit network number followed by a 10 bit local identifier for hosts. A router then can extract the network number using a logical AND operation between a 22-bit subnet mask and an IP address. For this example, we can say that a router can represent the four networks using the single entry 210.195.8.0/22, where /22 indicates the network number is 22 bits long. Likewise, 210.195.8.0/20 address would first 20 bits and so on. This indicates that we are grouping different smaller networks together and they are being treated same for the routing purposes.

Let us know take an example. Our IPv4 address is 210.195.8.0 and a 22 bit subnet mask is 255.255.252.0.

11010010 - 11000011 - 000010xx - xxxxxxxx (IP Address)

AND

11111111 - 11111111 - 11111100 - 00000000 (22 bit subnet mask)

11010010 - 11010011 - 00001000 - 00000000 (network number)

(210) (195) (8) (0)

The result of the bit-wise logical AND of the 32 bits of the IPv4 address and the subnet mask is the subnet prefix 210.195.8.0. It may therefore be noted that the bits in the fixed portion of the address (in which the bits in the subnet mask are set to 1), the subnet prefix bits are copied from the IPv4 address, essentially extracting the subnet prefix of the IPv4 address. On the other side the bits in the variable portion of the address where these are set to zero, the subnet prefix bits are also set to 0 and thus discarding the host ID portion of the IPv4 address.

10.4.3 Internet Transmission Protocol

The key protocols of the Transport Layer are **Transmission Control Protocol** (TCP) and **User Datagram Protocol** (UDP). TCP enables reliable data delivery service with end-to-end error detection and correction. UDP facilitates low-overhead, connectionless datagram delivery service. Both protocols are responsible for delivering data between the session layer and the network layer.

User Datagram Protocol (UDP)

The User Datagram Protocol enables application programs to have direct access to a datagram delivery service like the delivery service that IP provides. This enables applications to exchange messages over the network with a minimum of protocol overhead. UDP is connectionless unreliable datagram protocol in which the sending terminal does not check whether data has been received by receiving terminal. The unreliable service indicates that there is no guarantee that the data reaches at the receiving end of the network correctly. It can be understood more clearly by Figure 10.11

However, this protocol makes it possible to omit a variety of processes thus reducing the load on the CPU. UDP has 16-bit Source Port and Destination Port numbers. Figure 4.4 shows the data structure of the UDP header. The simplicity of the UDP header stems from the unsophisticated nature of the services it provides.

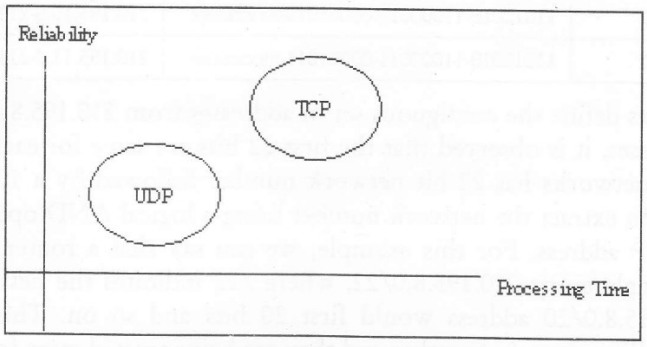


Figure 10.11: UDP vs TCP

0	16	31
Source Port	Destination Port	
Length	UDP Checksum	
Data		

Figure 10.12: UDP Format

Following is a brief description of each field:

- *Source Port:* Source port specifies port number of the application relating to the user data.
- *Destination Port:* As it name indicates, this pertains to the destination application.
- *Length:* It describes the total length of the UDP datagram, including both data and header information.
- *UDP Checksum:* It gives an option of integrity checking.

At this point, it is important to understand the layering concept along with the need for headers. The relationship between the IP and UDP has been depicted in Figure 10.13.

There are a number of good reasons for choosing UDP as a data transport service. When the amount of data being transmitted is small, UDP is considered the most efficient choice for a transport layer protocol because of the overhead for establishing connections and ensuring reliable delivery may be greater than the work of re-transmitting the entire data. Applications for a query-response model also works excellent for using UDP. The response is used as a positive acknowledgment to the query. When a response is not received within a certain time period, the application initiates another query.

Some examples of the usage of UDP are Remote file server (NFS), name translation (DNS), intra-domain routing (RIP), network management (SNMP), multimedia applications and telephony.

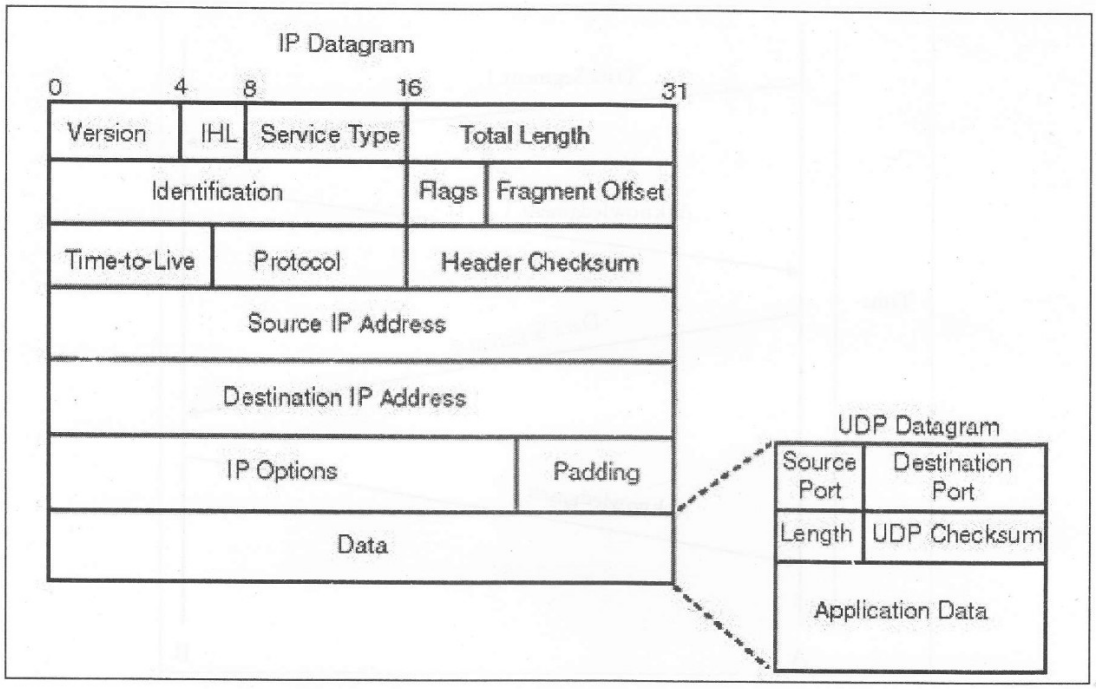


Figure 10.13: Relationship between the IP and UDP

Transmission Control Protocol

It provides a connection type service. That is, a logical connection must be established prior to communication. Because of this a continuous transmission of large amount of data is possible. It ensures a highly reliable data transmission for upper layers using IP protocol. This is possible because TCP uses positive acknowledgement to confirm the sender about the proper reception of data as shown in Figure 10.14. The sender keeps on send data at constant intervals until it receives a positive acknowledgement.

A negative acknowledgment implies that the failed data segment needs to be retransmitted.

What happens when a packet is lost on the network and fails to reach its ultimate destination? When host A sends data, it starts a time down counter. If the timer expires without receiving an acknowledgment, host A assumes that the data segment was lost. Consequently, the sending computer retransmits a duplicate of the failing segment.

Its other functions include sequence control, error recovery and control, flow control and identification of port number.

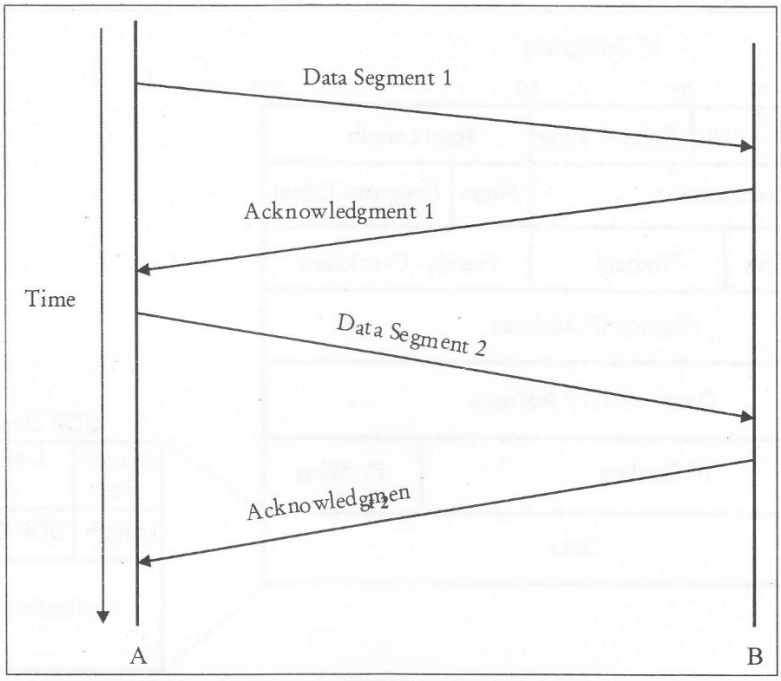


Figure 10.14: TCP Establishes Virtual Circuits

Figure 10.15 shows the format of the TCP data segment. The TCP header includes both source and destination port fields for identifying the applications for which the connection is established. The sequence and acknowledgment number fields underlie the positive acknowledgment and retransmission technique. Integrity checks are accommodated using the checksum field.

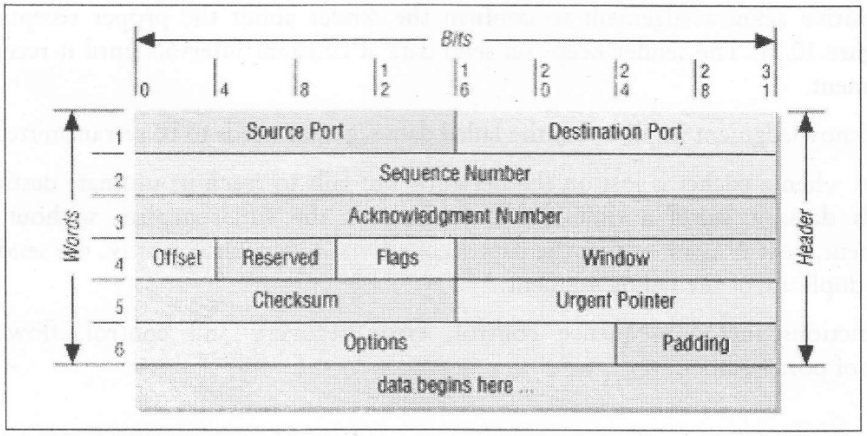


Figure 10.15: Data Segment Format of the TCP Protocol

TCP, therefore, unlike to UDP, TCP is a reliable connection-oriented byte-stream protocol.

Reliable: TCP provides reliable delivery of data using Positive Acknowledgment with Re-transmission (PAR) mechanism. PAR is a mechanism where the data is transmitted again and again until it hears from the remote system that the data arrived correctly. The unit of data exchanged between source and destination host is called a segment as shown in the Figure 10.15.Each segment has a checksum to

verify that the data arrives at the destination end undamaged. When the data segment is received undamaged, the receiver sends a positive acknowledgment back to the source end. When the data segment is damaged, the destination machine discards it. When the source machine does not receive any positive acknowledgement with in a specified time out period, it re-transmits the data segment.

Connection-oriented: TCP creates a logical end-to-end connection between the source and destination hosts. Handshake that is control information is exchanged between the source and destination hosts to set a dialogue before data is sent. TCP indicates the control function in a segment by setting the flag in a Flags field in the segment header. TCP uses a three-way handshake that indicates that three segments are exchanged. Figure 10.16 depicts the simplest form of the three-way handshake. Host A initiates the connection by transmitting host B a segment with the "Synchronize sequence numbers" (SYN) bit set. This segment indicates to host B that host A requests to create a connection. The segment also indicates to host B the sequence number host A will use as a starting number for its segments so that data can be put in the proper order. Host B replies to host A with a segment that has the "Acknowledgment" (ACK) and SYN bits set. Host B's segment acknowledges the receipt of A's segment and tells host A the Sequence Number host B will begin with. Finally, host A transmits a segment that acknowledges receipt of host B's segment. Thus, host A transfers the first actual data.

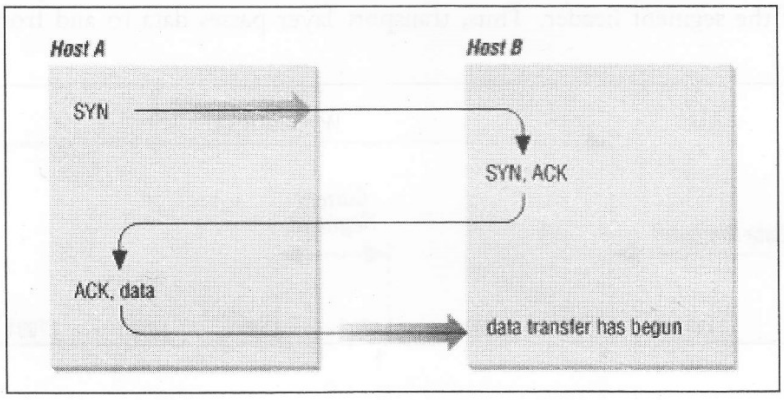


Figure 10.16: Three-way Handshake

This exchange of data also indicates to the TCP of host A has indication that the remote TCP is active and ready to receive data. When the connection is created, data can be exchanged. As soon as the source and destination machines have completed the data exchange, they initiate a three-way handshake with segments containing the "No more data from sender" bit (called the FIN bit) to release the connection. Thus, end-to-end exchange of data using the logical connection between the source and host machines is accomplished.

Continuous stream of bytes: TCP considers the data it transmits as a continuous stream of bytes, not as independent packets. This necessitates TCP to take care to maintain the sequence in which bytes are sent and received. The sequence number and acknowledgment number fields in the TCP segment header keep track of the bytes. In order to keep track of the data stream correctly, each end of the processes are required to know the other end's initial number. The source and destination ends of the processes synchronize byte-numbering systems by exchanging SYN segments during the handshake. The sequence number field in the SYN segment has the initial sequence Number (ISN). This is considered the starting point for the byte-numbering system. Thereafter, each byte of data is numbered sequentially from the ISN to start with ISN+1 for the first real byte of data to be transmitted.

The acknowledgment segment (ACK) has positive acknowledgment and flow control functions. The acknowledgment indicates to the sender the amount of data received and the data which can be received further. The acknowledgment number is the sequence number of the next byte the receiver is about to receive.

Figure 10.17 illustrates a TCP data stream that begins with an ISN of 0. The destination machine has received and acknowledged 2000 bytes. Therefore, the current acknowledgment number is 2001. The destination machine has enough buffer space for another 6000 bytes. The source machine is currently transmitting a segment of 1000 bytes starting with sequence number 4001. The source machine has received no acknowledgment for the bytes from 2001 onwards, but continues transmitting data as long as it is within the window. When the source machine fills the window and receives no acknowledgment of the data previously sent, it will, after time-out, transmit the data again beginning from the first unacknowledged byte. In Figure10.17 re-transmission begins from byte 2001 when no further acknowledgments are received. This makes source machine to believe that data is reliably received at the remote locations of the network.

TCP also ensures for delivering data received from IP to the correct application. The application is identified by 16-bit port number. The source machine and destination machine ports are included in the first word of the segment header. Thus, transport layer passes data to and from the application layer correctly.

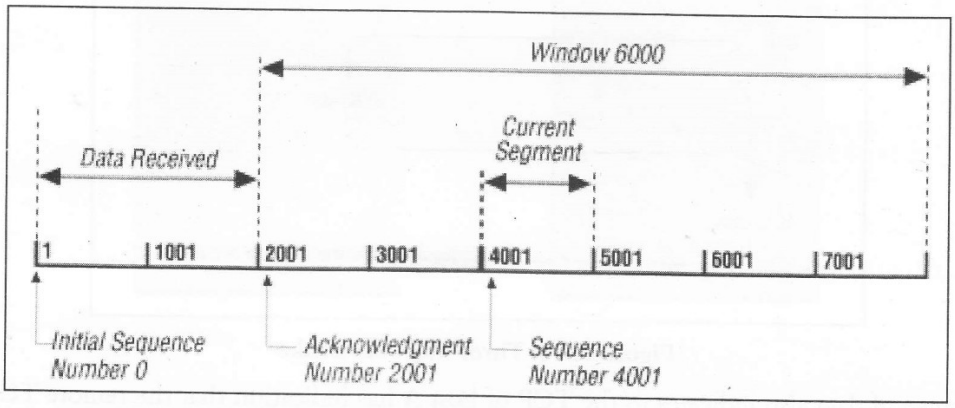


Figure 10.17: TCP Data Stream

Some of the applications of TCP are Electronic mail (SMTP), file transfer (FTP), remote login (Telnet), web (HTTP), etc.

10.4.4 Address Resolution Protocol (ARP)

The data link layer does not understand IP addresses. The hosts are attached to a LAN by an interface board that only understands LAN addresses. Each Ethernet contains a unique 48-bit Ethernet address. This address is entirely different from 32-bit IP addresses. It, therefore, becomes necessary to map 32-bit IP addresses into hardware addresses, such as 48-bit Ethernet addresses. The IP addresses are software addresses, which therefore are required to be resolved into equivalent destination's hardware address before transmitting datagram across a physical network. Therefore, the IP software must translate the IP address of the destination host into an equivalent hardware address. This process of translation of IP address into an equivalent MAC address or hardware address is called as address

resolution. This could be possible only when the destination host and the receiving hosts are connected on the same network.

ARP (Address Resolution Protocol) Table

The problem encountered while performing local routing may be summarized as follows:

1. How does host A know whether host B is in the same subnet?
2. How does host A obtain the MAC address of host B?
3. How does host A know that it can send the packet to router R so that the packet can reach host C that is in other subnet?
4. How do hosts determine which router to use if two or more routers are connected to the same LAN?

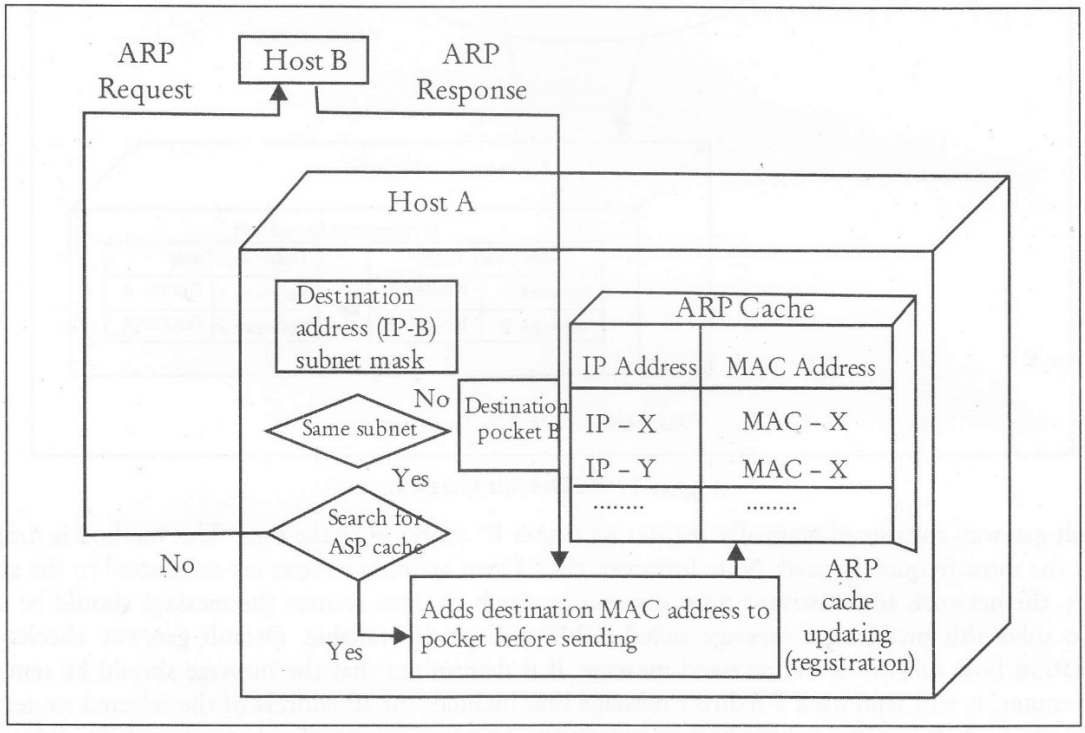


Figure 10.18: ARP Request and ARP Table Cache

The solution to problems 1 and 2 is shown in Figure 10.18.

For problem 1, because the source host knows its own subnet mask, it multiplies its own IP address by the subnet mask. It also multiplies the IP address of the destination host by the same mask. If the two results are the same, the source host knows that the destination host is in the same subnet.

For problem 2, the ARP request packet that includes the IP address of the destination host is broadcast on the subnet. Then, the IP address and MAC address of the destination host, included in the ARP reply packet returned from the destination host, are recorded on the table. Note, however, that contents of the table are deleted if they are not updated within a specified period of time. This is intended to accommodate changes made to MAC due to transfer of hosts or replacement of hardware.

Subnet mask refers to dividing an IP address into the network address and host address. For example, if 24 bits are assigned to the network address while 8 bits are assigned to the host address, then the subnet mask should be set to 255.255.255.0 (11111111111111111111111100000000).

Default Gateway

The solution to problems 3 and 4 from above is shown in Figure 10.19.

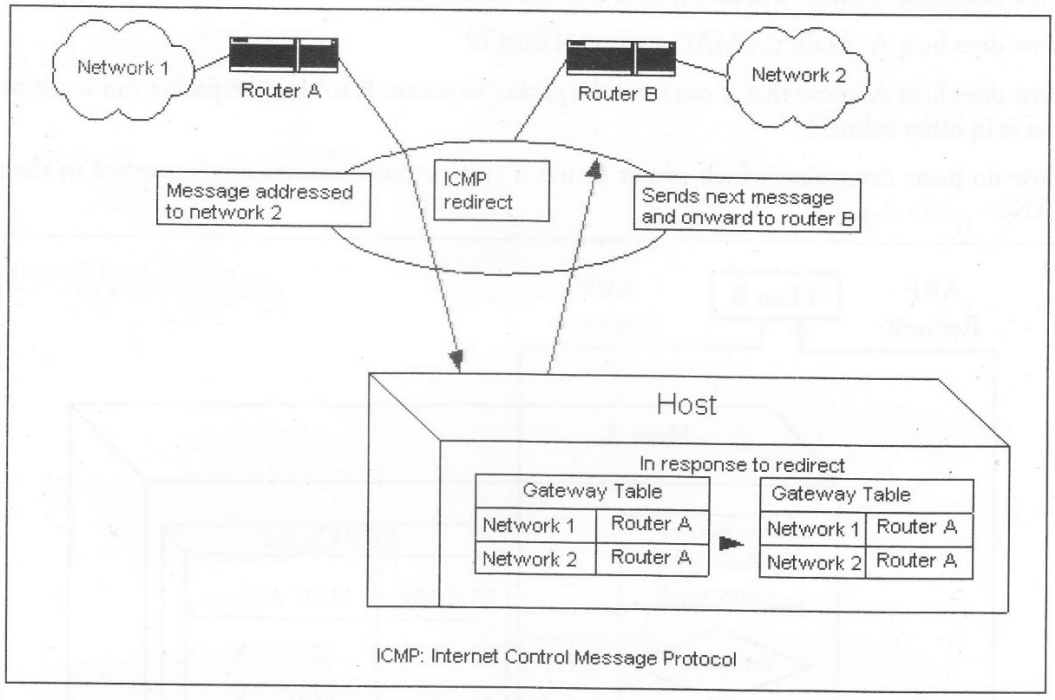


Figure 10.19: Default Gateway

Default gateway consists of manually registering router IP addresses in the host. This method is simple, and is the most frequently used. Note however, that if two or more routers are connected to the same subnet, the network administrator must determine which of these routers the message should be sent to. To solve this problem, a message called ICMP redirect is available. Default gateway checks the destination host address in the received message. If it determines that the message should be sent via other router, it will send back a redirect message that includes the IP address of the selected router, to the source host. When the source host receives this message, it stores the destination host address as well as the second router address so that it can send messages addressed to this host via the second router.

10.4.5 Reverse Address Resolution Protocol (RARP)

Like ARP, it also performs the same function but in reverse order. It determines the IP address with the help of a given MAC address. The RARP enables a host to discover its IP address when it knows only its MAC address. The host wishing to know its IP address broadcasts an RARP query containing its MAC address to all nodes on its physical network. A server on the network recognizes the RARP packet and returns the host's IP address. This usually occurs when booting a diskless workstation. Normally, such a machine gets the binary image of its operating system from a remote file server and

learns its IP address. A newly-booted workstation broadcasts (using the all 1s address) its Ethernet address and say: "My 48-bit Ethernet address is 14.04.05.18.01.25. Does anyone out there know my IP address?" The RARP server at the local network looks this request, looks up the Ethernet address in its configuration files and forwards the corresponding IP address.

10.4.6 ICMP: Future IP: Error Reporting Mechanism

The Internet Control Message Protocol (ICMP), an error reporting protocol that is an integral part of the IP protocol. ICMP communicate control data, information data, and error recovery data across the network. Problems that is less severe than transmission errors result in error conditions that can be reported. For example, suppose some of the physical paths in Internet fail, causing the Internet to be partitioned into two sets of networks with no path between the sets, a datagram sent from a host in one set to a host in other cannot be delivered.

The TCP/IP suite includes a protocol called ICMP that IP uses to send error messages when condition such as the one described above arises. The protocol is required for a standard implementation of IP. We will see that the two protocols are co-dependent. IP uses ICMP when it sends an error message, and ICMP uses IP to transport messages.

Following is a brief description of some of the error messages defined by ICMP protocol:

- *Source Quench:* A router or host whose receive communication buffers are nearly full normally triggers this message. A source quench message is sent to the sending host, the receiver is simply requesting the sending host to reduce the rate at which it is transmitting until advised otherwise.
- *Time Exceeded:* A time-exceeded message is sent in two cases. Whenever a router reduces the TTL field in a datagram to zero, the router discards the datagram and sends a time-exceeded message. In addition, a time-exceeded message is sent by a host if the reassembly timer expires before all fragments from a given datagram arrive.
- *Route Redirect:* A router sends this message to a host that is requesting its routing services. When a host creates a datagram destined for a network, it sends the datagram to a router, which forwards the datagram to its destination. If a router determines that a host has incorrectly sent a datagram that should be sent to a different router, the router uses route redirect message to cause the host to change its route. In this manner, a route redirect message improves the efficiency of the routing process by informing the requesting host of a shorter path to the desired destination.
- *Host Unreachable:* Whenever a gateway or a router determines that a datagram cannot be delivered to its final destination (due to link failure or bandwidth congestion), an ICMP host unreachable message is sent to the originating node on the network. Normally the message includes the reason the host cannot be reached.
- *Fragmentation and Reassembly:* The largest datagram the IP protocol can handle is 64 Kbytes. The maximum datagram size is dictated by the width of the total length field in the IP header. Realistically, most underlying data link technologies cannot accommodate this data size. For example, the maximum size of the data frame supported by Ethernet is 1,514 bytes. Unless rectified, something is done about situations like this. IP has to discard data that is delivered to it from upper-layer protocols with sizes exceeding the maximum tolerable size by the data link layer. To circumvent this difficulty, IP is built to provide data fragmentation and reassembly.

- Whenever an upper-layer protocol delivers data segments whose sizes exceed the limit allowed by the underlying network, IP breaks the data into smaller pieces that are manageable within the allowed limit. The small datagrams are then sent to the target host, which reassembles them for subsequent delivery to an upper-layer protocol.
- Data fragments, however, takes the same route but there is instances when they may adopt alternate route too. Fragments traversing different routes may reach their destination out of the order in which they were sent. To allow for recovery from such a behavior, IP employs the fragmentation-offset field in its header. The fragmentation-offset field includes sequencing information that the remote IP host uses to recover the sequence in which the datagrams were sent. The fragmentation-offset field also contains information for detecting missing fragments, which is used by IP. Data is passed to the protocol described in the protocol field only when all related fragments are duly received and reordered, it is known as data reassembly.
- Fragments belonging to two or more independent large data can be differentiated by IP using identification field. Fragments of the same datagram are uniquely assigned in the identification field. The receiving end uses this number to recover the IP fragments to their respective datagrams.
- A host that creates a datagram can set a bit in the flag field to specify the fragmentation. This bit is set to 1 in all fragments belonging to a datagram except for the final fragment. This ensures that all fragments of a datagram are received.
- *Echo request/Echo reply:* These two ICMP messages are exchanged between ICMP software on any two hosts in a bid to check connectivity between them. The ping command is an example of a diagnostic command commonly used by network users to check for the reachability of a certain host. On invoking this command, ICMP echo request message is sent to the target host. The target host responds with an echo as proof of reachability. It should however be operational and connected to the network. In other words, the reply carries the same data as the request.
- *Address Mask Request/Reply:* A host broadcasts an address mask request when it boots, and routers that receive the request send an address mask reply that contains the correct 32-bit subnet mask being used on the network.

ICMP uses IP to transport each error message. When a router has an ICMP message to send, it creates an IP datagram and encapsulates the ICMP message in the datagram. It means that the ICMP message is placed in the data area of the IP datagram. The datagram is forwarded as usual with the complete datagram being encapsulated in a frame for transmission. Figure 10.20 illustrates the two level of encapsulation. Figure 10.20 illustrates two levels of data encapsulation.

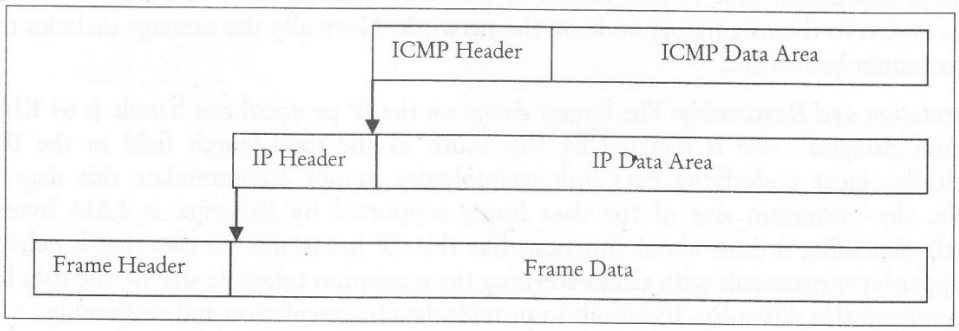


Figure 10.20: Two Levels of Encapsulation in case of ICMP Datagram Transmission

Check Your Progress

1. What are virtual circuits?
2. Define transport layer.

10.5 LET US SUM UP

The computer with internetworking has become a potent tool for education, productivity and enlightenment. The Internet can improve the quality of life at a relatively low cost. The Internet services like e-mail, surfing Internet, FTP, Telnet, and database access, gopher; Archie, have become popular and useful. Thus, we have learnt and understood about the necessity of computer networks in today's ever expanding technology, the network hardware and software, reference models to widely distributed computers and networks working on different operating systems and protocols, etc. We have also learnt as to how the evolution of Internet from ARPAnet could be possible under example of networks. The three basic components namely; hardware, protocols (software) and applications (useful software) are mandatory to implement a computer network. It is also explained that the concept of layers is important in networking. Each layer with two layers work as the interface and protects the upper layer that each one layer can change with minimum impact on the upper layers. In some cases, this protection is so proficient that an application may not know that it is running on different hardware. The OSI network model has seven layers.

Unlike bridge, routers are made to interconnect dissimilar as well as similar LANs/networks together. This operates at network layer of OSI model. There also exist many routers such as inter router protocols, serial line protocols, etc. There is also one category known as gateway. The installed base of ever growing different networks with variation in technologies will always be around. The advancement in hardware and software is also pushing the cost to be always down. This is leading to a situation in which different sections of a big company install different LANs. This initiates circumstances where it is desirable to connect different networks together. The network layer ensures the delivery of the packet from source to destination. There are many devices such as fibre modem, repeaters, bridges, routers, gateways and switching hubs, which provide an extension of the network globally. Repeater is a device, which compensates for the attenuation of the signal in the media and therefore increases the length of the LAN by connecting similar LANs together. This corresponds to the physical layer of OSI model.

The IPv4 address, which is a 32- bit integer, is difficult to remember. Therefore easily remembered host names were devised. Due to Internet explosion, it is not practical to keep an exhaustive hosts file for every host because of the sheer volume of listing as well as addition, deletion and updating of new, old and current hosts. TCP/IP, a common protocol method is used to interconnect computers together, and also serve as the default protocol for accessing information over the Internet. TCP/IP is the protocol used by computers on the Internet and may be considered as two separate protocols such as TCP and IP. Each computer has an IP address. A protocol is a set of rules that govern how computers talk to each other. With TCP/IP, different computer systems can reliably exchange data on an interconnected network. ICMP packets contain information about failures on the network or error control such as inoperative nodes and gateways, packet congestion etc. The IP software interprets ICMP message. ICMP messages often travel across many networks to reach their destination so they are encapsulated in the data portion often IP datagram.

10.6 KEYWORDS

Internetwork: A scheme for interconnecting multiple networks of dissimilar technologies.

Internet Control Message Protocol (ICMP): An error reporting protocol that is an integral part of the IP protocol. ICMP communicate control data, information data, and error recovery data across the network.

IP protocol: A connectionless type service and operates at third layer of OSI reference model

Virtual Circuits: Connecting a number of dissimilar computer networks presents a seamless communication channel to which many systems are attached.

Address Resolution Protocol (ARP): The hosts are attached to a LAN by an interface board that only understands LAN addresses.

Fragmentation: Each autonomous system places limits on the maximum size of a packet.

Reverse Address Resolution Protocol (RARP): Like ARP, it also performs the same function but in reverse order. It determines the IP address with the help of a given MAC address.

10.7 QUESTIONS FOR DISCUSSION

1. Why there are only 16,777,214 host IDs available in Class A of IP addressing scheme?
2. What is ICMP and ARP?
3. How did CIDR ease the problem of IPv4 addressing?
4. What is fragmentation?
5. Compare OSI and TCP/IP models.

Check Your Progress: Model Answers

1. Connecting a number of dissimilar computer networks presents a seamless communication channel to which many systems are attached.
2. The transport layer of TCP/IP model corresponds to the transport layer of the OSI reference model. It is represented by two end-to-end protocols namely, and UDP.

10.8 SUGGESTED READINGS

Rajneesh Agrawal and Bhata Bhushan Tiwari, *Computer Networks and Internet*, Vikas Publication

Burton, Bill, *Remote Access for Cisco Networks*, McGraw-Hill Osborne Media

Behrouz A. Forouzan, Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies

Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall

LESSON

11

NETWORK APPLICATIONS

CONTENTS

11.0 AIMS AND OBJECTIVES

After studying this lesson, you will be able to:

- Discuss the use of different network applications like ftp, http, etc.
- Understand WWW concept
- Discuss web browsers
- Know how to design a webpage
- Discuss how search engines are used to retrieve information from www

11.1 INTRODUCTION

Application layer functions are to identify source and destination machine wished to communicate with one another, find out resource availability and synchronize exchange of data between source and destination machines. However, each application is different but some applications are so useful that they have become standardized. To identify source and destination machine wishing for exchanging data, the application layer finds out the identity and availability of source and destination machines for

an application with data to send. To determine the resource availability, the application layer decides if sufficient network resources for the requested exchange of data are available. In order to synchronize exchange of data, the application layer provides the necessary support.

11.2 NETWORK APPLICATIONS

The Internet has defined standards for File Transfer Protocol (FTP) that connects to a remote machine and sends or fetches an arbitrary file. FTP addresses the issues of authentication, listing a directory contents, ASCII or binary files, etc.

Another aspect of the application layer is to login remotely. This application is called Telnet. It is a remote terminal protocol that enables a user at one location to establish a TCP connection to another location and then pass keystrokes from the local host to the remote host.

Mail using Simple Mail Transport Protocol (SMTP) enables a mail delivery agent on a local machine to connect to a mail delivery agent on a remote machine and deliver mail. Similarly, there are many other applications such as News (NNTP) enabling a communication between a news server and a news client and Web (HTTP) based protocol for communication on the World Wide Web. These applications can be standardized for making common applications for users.

Client Server Interaction

In the client server architecture, a machine (client) requests to another machine (server) to create a connection for providing some service. The services running on the server run on ports. The ports are application identifiers. The client machine should know the address of the server machine for getting the desired services from this port and to connect to the server machine. However, the server machine should not know the address or the port of the client machine at the time of connection initiation. The first packet transmitted by the client machine as a request to the server machine contains details about the client which are further used by the server to send any information. Client machine acts as the active device which makes the first move to establish the connection whereas the server machine passively waits for such requests from some client.

Various network applications are discussed below.

11.2.1 File Transfer Protocol (FTP)

The File Transfer protocol is among the oldest protocols still used in the Internet. FTP is widely available on almost all-browsers indicating that all computing platforms, including DOS, OS/2, UNIX, and up to the mainframe level have this service available. You can very well understand from its name that it facilitates the majority of file transfers across the Internet. In other word, FTP is a file server access protocol that enables a user to transfer files between two hosts across the network or Internet using TCP. You may see the versatility of this application layer protocol, it accomplishes its job even intended hosts at separate locations could potentially be running different operating systems, using different file storage systems, and using different character sets. Accessing FTP sites over the Internet requires that the user must have the knowledge of the location and the name of the desired files.

Unlike Telnet, FTP does not require any familiarity with the remote operating system. The user is still required, however, to be familiar with the FTP command set built into the protocol itself so that he or she can productively manage the session.

Modern FTP servers known as ftpd support two different TCP connections, namely control and data connections. First control connection is invoked for the entire duration of transfer of file or FTP session. It facilitates the exchange of commands issued by the client, and replies originating from server. Data connection is established as and when it is required. Its main function is to facilitate transfer of files and directory listings to and from the client at the client's request.

Whenever you wish to do FTP, you need to invoke a few commands. These commands basically are related to transfer a file from remote computer to your computer or from your computer to the remote computer. There are anonymous as well as authorized privileges with regard to transfer of a file from a server. In case of anonymous FTP servers, you can do FTP without authorization. You need to login with a username, which is anonymous, and a password that is your e-mail address. Apart from this, there are authorized servers for which you need to register before you are permitted to do FTP. After registration, you will get a password.

Trivial File Transfer Protocol (TFTP)

TFTP, like FTP, is also an Internet service intended for the transfer of files between from one computer to another over a network. It does not provide password protection or user directory capability. Unlike FTP, however, TFTP does not rely on TCP for transport services. Instead, TFTP uses UDP to shuttle the requested file to the TFTP client. Furthermore, diskless devices that keep software in ROM to use it to boot themselves can use it. It is simpler than the File Transfer Protocol (FTP) but less capable. FTFTP facilitates to quickly send files across the network with fewer security features than FTP.

11.2.2 Electronic Mail

Electronic mail is one of the most popular network services. The use of Electronic mail or e-mail has probably may be cited as the foremost reason for the popularity of Internet. The proliferation of cyber café can be credited to e-mail or World Wide Web. E-mail provides an efficient and fast means of communication with relatives, friends or colleagues throughout the world. You cannot only communicate with one person at a time or thousands but also you can receive and send files and other information in a short time. In e-mail communication, the intended receiver or receivers of the message are not required to be present at their desktop at the time of receiving of the message by their computer. It works like a postal mail. In postal mail postman puts the sender's message in your mailbox and when you come back from your work, you access your mailbox to retrieve the message. Therefore we may consider it in a way a substitute of postal mail. However, it has many-many more superior features than postal mail. E-mail has two parts:

1. *User agent:* It is the user interface to the mail system. The user agent system enables to provide ways to view, edit, and reply to messages, etc. It also accesses messages stored in a system mailbox. The user agent enables the user to use a text editor to create a file that the user agent hands over to the message transfer agent.
2. *Message Transfer Agent (MTA):* It is a software package that transports messages created by a user to destination mailboxes possibly on remote machines. The MTA has to perform more complex jobs than other applications:
   (a) MTA handles temporary failures when a destination machine is temporarily unavailable; it stores the message on the local machine for later delivery. Thus, the User Agent typically just stores messages into a storage area.

(b) MTA distinguishes between local and remote recipients.

(c) MTA needs to deliver copies of a message to several machines.

(d) MTA has to allow mixing text, voice, and video in a message and appending documents and files to a message.

As discussed above, the email addresses consist of the following components:

- *Mailbox names:* A mailbox is associated with one login id within a mail server to store the emails of the user. Therefore, a specific name is provided to the mailbox associated with each IDs.
- *Symbolic names:* It refers to the name of a service rather than a specific user. For example, postmaster is universally recognized as an address for post mail problems. In email system, the symbolic names are aliases for specific mailboxes.
- *Group names (mail exploders):* It refers to an alias for a set of recipients. MTA consults an internal database to specify the mail addresses.

There are number of e-mail packages available. Some of them are free like Goggle's mail, Yahoo mail, hotmail etc while some are paid. All of them are also not alike but most of the e-mail software has some basic functionality common. These are:

- Send and receive mail messages
- Save your messages in a file
- Print mail messages
- Forward a mail message to other recipient
- Reply to mail messages
- Attach a file to a mail message

In order to send a message we need to first type the address of the intended recipient. E-mail addresses have some sort of similarity with phone numbers with regard to the identification of person, organization, or a geographic location. E-mail addresses likewise, telephone numbers, which have usually area code, have rules for use. Usually, the e-mail address has three parts:

- A user identity or name
- An "at" sign (@)
- The domain name, which basically specifies the address of the user's mail server. It is the right most part of the address and follows a particular naming conventions. You can now understand the email address by the help of the following example:-

Example: services@jalandhar.in

The left most part before the @ (at sign) is the identity or name if the user and the right most part after the @ is the server which indicates India. There are some naming conventions like edu, com, org etc used for education, commercial, organization respectively.

Simple Mail Transfer Protocol (SMTP)

Electronic mail (E-mail) is considered the most widely used TCP/IP application. The Internet mail protocols enable a client machine to exchange mail and message between TCP/IP hosts. Three

standard protocols are applied to provide such mail application. The SMTP is one of them. The three standards are given below:

1. *SMTP:* It is a standard for exchange of mail between two computers (STD 10/RFC 821), which specifies the protocol used to send mail between TCP/IP hosts.
2. *Mail:* It is a standard (STD 11) defining the format of the mail messages, syntax of mail header fields, a set of header fields and their interpretation and about a set of document types other than plain text ASCII to be used in the mail body.
3. *DNS-MX:* It is a standard for the routing of mail using the Domain Name System (RFC 974).

SMTP, an application layer protocol, is used to send e-mail messages across the Internet. It utilizes TCP as the transport protocol to send email to a destination mail exchanger, referred as mail server. A client machine sends email to a mail exchanger or an email is sent from mail exchanger to another mail exchanger. E-mail transmitted using SMTP is normally transmitted from one mail exchanger to another directly. E-mail was never designed to be instantaneous but it appears so often.

Mail Exchangers are nothing but the software application programs to support the SMTP protocol. Mail Exchangers such as send mail or Microsoft Exchange wait for IP datagrams that arrive on the network interface with a TCP port number of 25. When a message is arrived, the mail exchanger checks to find out if it is for one of its users and accordingly move the mail to the user's mailbox. The data sent using SMTP is 7-bit ASCII data, with the high-order bit cleared to zero is found adequate in most instances for the transmission of English text messages but is inadequate for non-English text or non-textual data. To overcome these limitations, Multipurpose Internet Mail Extensions (MIME) defines a mechanism for encoding text and binary data as 7-bit ASCII within the mail envelope and SMTP Service Extensions specifies a mechanism to extend the capabilities of SMTP beyond the limitations.

How SMTP Works

SMTP is end-to-end delivery in which an SMTP client machine contacts the destination host's SMTP server directly to deliver the mail. Unlike the store-and-forward principle that delivers the mail content to the destination host through a number of intermediary nodes in the same network, SMTP continues the mail content being transmitted until it has been successfully copied to the host's SMTP. In case of store and forward mechanism, the successful transmission from the sender only indicates that the mail content has reached the first intermediate hop. There are instances when mail is exchanged between the TCP/IP SMTP mailing system and the locally used mailing systems. Such applications are referred as mail gateways or mail bridges. However, SMTP guarantees only delivery to the mail-gateway host, not to the real destination host, which is located beyond the TCP/IP network. In case of a mail gateway, the SMTP end-to-end transmission is host-to-gateway, gateway-to-host or gateway-to-gateway. SMTP does not specify the format of mail beyond the gateway.

Each message of SMTP contains the following fields:

- A header or envelope that is terminated by a null line.
- Contents – Everything after the null or blank line is the message body with sequence of lines containing ASCII characters.
- Simple Mail Transfer Protocol defines a client/server protocol. The client SMTP machine initiates the session by sending SMTP message and the mail server responds by receiving SMTP message to the session request.

Mail Exchange

The SMTP design is based on the model of communication illustrated Figure 11.1. After the client machine mail request, the sender-SMTP sets a two-way connection with a receiver-SMTP. The receiver-SMTP may be the destination machine or an intermediate machine (mail gateway). The sender-SMTP will initiate commands which are replied to by the receiver-SMTP.

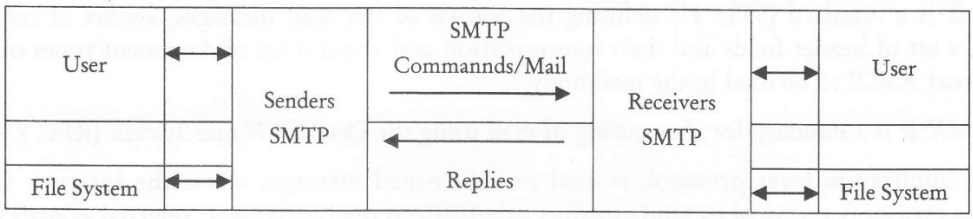


Figure 11.1: SMTP Communication

1. The client machine SMTP sets a TCP connection with the destination machine SMTP and then waits for the server to send a service ready message or a service not available message.
2. HELO (HELO is an abbreviation for hello) is sent and the receiver machine will identify itself by sending back its domain name. The client machine SMTP uses this to verify if it reached the right destination SMTP. If the client machine SMTP supports SMTP Service Extensions, it substitutes an EHLO command in place of the HELO command. A destination machine SMTP which does not support service extensions responds with a 500 Syntax error, command unrecognized message. The client machine SMTP then retries with HELO, or if it cannot transmit the message without one or more service extensions, it should send a QUIT message. If a receiver-SMTP supports service extensions, it responds with a multi-line 250 OK message which includes a list of service extensions which it supports.
3. The client machine now initiates the start of a mail transaction by sending a MAIL command to the destination machine. This command has the reverse-path that is used to report errors. It should be noted that a path is more than just the user mailbox@host domain name pair. Besides, it has a list of routing hosts.
4. The next step of the actual mail exchange provides the server SMTP with the destinations for the message, the message may go to more than one recipient. This is accomplished by sending one or more RCPT TO:<forward-path> commands. Each of them will receive a reply 250 OK when the destination is known to the server or a 550 No such user here when it is not known to the server.
5. When all RCPT commands are sent, the sender forwards a DATA command to notify the destination machine that the message contents are following. The server replies with 354 Start mail input, end with <CRLF>.<CRLF>. It should be noted that the ending sequence that the client machine uses to terminate the message data.
6. The client machine now sends the data line by line ending with the 5-character sequence <CRLF>.<CRLF> line upon which the destination machine acknowledges with a 250 OK or an appropriate error message when anything went wrong.

Now, there are several possible actions:

- The destination machine has no more messages to transmit, it will end the connection with a QUIT command. This command is answered with a 221 Service closing transmission channel reply.

- The destination machine has no more messages to transmit, but it is ready to receive messages (if any) from the other side. It will issue the TURN command. The two SMTPs now switch their role of sender/receiver and the client machine that was previously the destination machine now transmits messages by starting with step 3 above.
- The client machine wants to transmit another message and simply follows step 3 to transmit a new MAIL command.

11.2.3 Remote Login - Telnet

A remote login facility enables a user to create a login session to a remote machine and then execute commands. Telnet is an Internet standard remote login protocol to connect a local terminal with a remote login session. It copies keystrokes to the remote machine and copies output from the remote machine to the source machine. Telnet is a program that allows a user with remote login capabilities to use the computing resources and services available on the host. It emulates the remote terminal on your desktop and therefore referred as terminal emulation protocol of TCP/IP. Telnet can also be used to connect other ports serving user defined as well as well-known services. It works as client server model where it establishes a virtual connection using the TCP transport protocol. The telnet program requires two arguments that is the name of a computer on which the server runs and the protocol port number of the server. After establishing connection, the Telnet server and client enter a phase of option negotiation to determine the options that each side will like to support for the connection. They are always free to change their options even after establishing the connection. This provides it a versatile terminal emulation due to the many options. It can transfer binary data, support byte macros, emulate graphics terminals, and convey information to support centralized terminal management.

Telnet service is unique in that it is not platform-specific like other TCP/IP services. A DOS user running Telnet, for example, can connect to a UNIX host or a mainframe computer. The down side of using Telnet, however, is that unless the user is familiar with the operating system running on the remote platform, he or she cannot use the desired resources easily.Telnet aims to provide three services:

1. Telnet defines a Network Virtual Terminal (NVT) standard to describe a standard terminal. Client programs then interact with the NVT. The server translates NVT operations into specific commands to the actual hardware/operating system.
2. Telnet enables the remote machines connecting together to negotiate options with one another. Option negotiation makes agree both the remote machines on a common level of service.
3. Telnet treats connections of the remote machines symmetrical and enabling them to use programs. Telnet also defines data and command sequences to deal with heterogeneity. The client machine translates keystrokes into NVT format and sends them to the server machine at remote location. The server machine translates NVT operations into the appropriate local representation.

Some of the Telnet commands are given below:

- Interrupt process (IP) - It terminates the running program.
- Abort output (AO) - It refers to discard any buffered output.
- Are you there (AYT) - This command allows client to send and out-of-band query to verify the remote end is still there.
- Erase character (EC) - It refers to erase the previous character.

- Erase line (EL) - It deletes the entire current line.
- Synchronize - It clears data path to remote party.
- Break - It is equivalent of the BREAK or ATTENTION key.

11.2.4 Domain Name System (DNS)

Now we have two types IP address in the form of decimal numbers and text for the same host. You know that list of all IP addresses are maintained centrally by ICANN in the form of distributed database directory. There are several distributed servers, which maintain this list of IP addresses. The reasons behind the distributed server are very logical and simple. It helps in disaster management and in diverting the load of the traffics in the form of requests from clients to other DNS servers located at different sites. DNS server maintains database in both the form that is textual as well as decimal notations. For example, DNS server maintains the address of google site as www.google.com and 216.23.9.53.99. In this manner, DNS is used to provide host-to-IP address mapping of remote hosts to the local hosts and vice versa. It is now amply clear that the DNS maintains a distributed database to map between hostnames and IP addresses. Whenever a client requests a service from a site, then both the site runs DNS protocol to access the distributed database which is nothing but Domain Name Systems. Therefore, the DNS provides the protocol, which allows clients and servers to communicate with each other. DNS enables a system to use a resolver, which resolves the host name to IP address understandable by server.

You may be now thinking of how DNS is able to provide the quick translation of text of the IP addresses within fraction of seconds from a directory of billions of such addresses. This could be made possible by using Domain concepts, which uses hierarchical arrangements of text addresses translation.

You can see from the Figure 11.2 that at the top level is the root server, which has null label. Below this is another level domain or domain as com, edu, int and so on which are grouped together. Below this different sub domains or groups have been created. Table 11.2 below corresponds to some commonly appearing domain names with their respective sites. The DNS can accommodate almost all kinds of organizations by allowing each group to choose between geographical or organizational naming hierarchies.

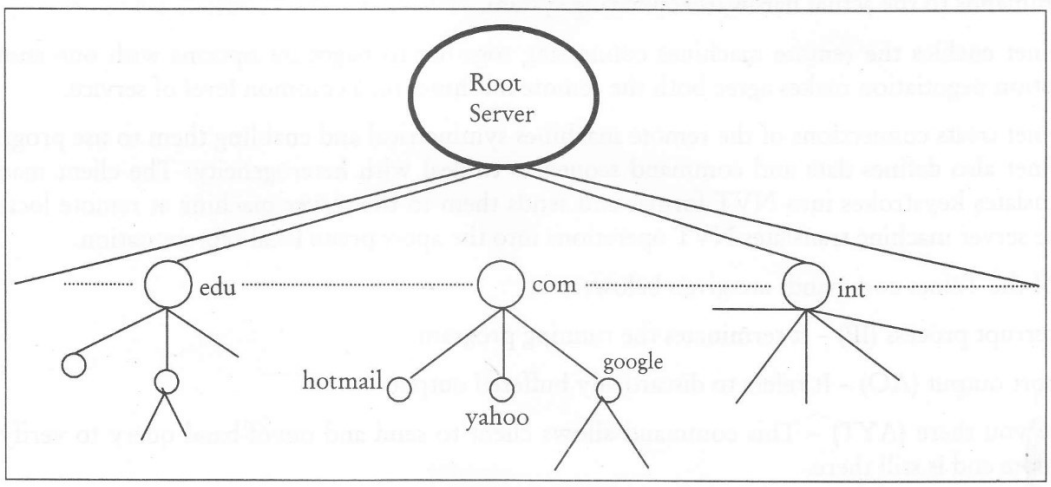


Figure 11.2: DNS Hierarchy

Table 11.1: Internet Domains

Domain	Indicative Site
Com	Commercial institute
Edu	Educational institute
Org	Nonprofit organization
Net	Network service provider
Gov	Government department
Mil	Military
Biz	Business
Country code	For example, in for India, us for USA, au for Australia, jp for Japan and so on

As we know that the servers maintaining addresses are distributed and have locations throughout the world. Then this question arises as to how text addresses are organized in hierarchical arrangement. You may refer Figure 11.2 above and Table 11.1. The hierarchy is represented into zones and each zones is a hierarchy of one or more nodes without any overlapping. Each zone is represented by a server and undoubtedly with one backup server. Root server as shown in Figure 11.2 is only one, which is just indicative; there may be several root servers at several locations in the world. Each root is aware of the location of each DNS server of specific domains.

The process is now very simple to understand. When you need to connect with a particular site, you first send your request to your local host. If your local host can provide the translation, your request is completed. If not, your local host then sends your request one level above in the hierarchy. If the server at one level above is able to handle the same, you get your intended website at your desktop through your local server. If not, then the server at one level above from your local server either sends your request again to another server or informs your local server that your request is failed and gives the address of another server to process your request. This process continues till a server is found who knows the address, otherwise, the request is filtered up to the root server. Depending upon the domain address, root server forwards the request to the one of the domain servers represented at the next level of hierarchy. This process continues and the information of text address is returned to Root server and then back to your local server.

11.1.5 Hyper Text Transfer Protocol

In order to access any website, the web browsers are used which are assisted by the URL that uses the http scheme. It is the URL or the port number that assists the browser to link with a Web site. The server indicates a computer connected to the Internet while the port number indicates a type of socket to which the browser plugs in to link with the Web server. The web server not only provides the requisite web pages but also describes a computer program that runs on a computer to provide web pages. When a browser receives an URL will attempt to connect with the server computer having the required web pages by connecting to the specified port number. The URL can be provided to the browser either by typing it at its specified location or by clicking on the link available on some already displayed web page or document.

It is the role of the browser to connect with the server where the requisite requests from client or user is stored or available. When the web server receives the request from browser it replies back to the

browser, which is client in this case. The information basically contains the HTTP protocol version, name of the server, the media type of the document and date etc. The media type of the document is quite important information because the browser is required to know what kind of document this is before it can process it. HTML is the most common media type transferred over the Web. Other media types are GIF image and JPEG image. Several times when a response like "HTTP 404 Not Found" is displayed which means that the request document is not available at the link. There are different responses defined in HTTP. Briefly, in order to access a web page, HTTP involves browser that issues a request followed by a few headers. In response, the server replies back with a few headers and a document.

The web server basically maps the URLs to files on its hard disks. The web server interprets the path in any URL to map it with a filename on its hard disk. In order to make it work to map with the requisite file, the web server is configured to contain a "document root" directory relative to which all URLs are resolved as filenames. Let us take an example, suppose the URL is http://myspace.tutorial.in, and the document root is D:\WWWFiles\. When a user types the URL http://myspace.tutorial.in/lesson1/networking.htm into browser, the browser requests the server for the document/lesson1/networking.htm. The web server begins searching in the directory D:\WWWFiles\lesson1 for a file called networking.htm. If the requisite file is available it responds with a header followed by the document. If it is not available, it responds a 404 Not Found followed by a helpful error message telling the user to search elsewhere.

11.3 WORLD WIDE WEB

The World Wide Web and Internet have impacted the world including business, social, political life in the last few years. It is expected that this trend will certainly continue well into the future too. You must be very much familiar with the term World Wide Web, which is also known as web or WWW or W3 and has established itself as the most popular part of the Internet by far. It is an incredible mines of information, once you start searching anything ranging from documents to pictures to software, it almost appears limitless. It provides you documents, sound files, view images, animation, and video, speak and hear voice, and view programs that run on practically on any software in the world. Therefore, it facilitates the rich and diverse communication by enabling you to access and interact with text, graphics, animation, photos, audio and video. It has now become so simple for you to understand how the web works and what it is. Its implementation is based on client server system and employs your personal computer as client, web browser software, a connection to an Internet service provider, servers, routers and switches to direct the flow of information. You may be aware of all terms used in the formation of a web except web browser.

A browser is software, which your computer uses to view WWW documents and access the Internet. The browser program residing in your computer facilitates you with the advantages of text formatting, hypertext links, images, sounds, motion, and other features. Internet Explorer and Netscape are some of the widely used browsers. Browsers have sub programs called plug-ins to handle the documents you find on the Web. It may also other plug-ins stored elsewhere in your computer.

Web is very simple to use. Whenever you wish to visit any website, say your institute's website, you simply enter the address or URL of the website in your web browser to forward your request to the web server of the institute to provide you the intended web page. The institute's web server then sends your request on the Internet to find the intended website. Once it is obtained, the web server returns

the same to your computer where the browser loaded with different plug-ins interprets the data, displaying it on your computer screen. The intended web page, which is now available on your desktop, may have click able links. On clicking on the same, you may visit other pages. In this manner, the information scattered across the globe can be linked together.

It now becomes essential to explain as to how the different web pages with different text format and standards could be linked to a particular web page. The binding forces that hold the Web together are the hypertext and the hyperlinks. The hyperlink allows electronic files on the Web to be linked so you can jump easily between them using hypertext protocol. As you have learnt that web browsers that enables you to access the Web also distinguish between web pages and other types of data on the Internet because web pages are written in a computer language called Hypertext Markup Language or HTML.

World Wide Web (WWW): The World Wide Web is an information space that provides resources, which are identified by global identifiers called Uniform Resource Identifiers (URI). There is software, which in conjunction with servers, proxies, spiders, browsers and multimedia applications enables an user to identify and explore resources on web space. In networking language, www is a client-server information system that utilizes the Internet to access computers containing millions of hypertext documents.

11.3.1 Advantages

Many companies have understood the advantages of having a presence on the World Wide Web and have successfully addressed their corporate objectives by integrating their web site as part of their business strategy. As we are aware that a website can generate awareness of the products and services of the company and provide a global storefront for the company 24 hours a day with automating many business procedures. It is relatively inexpensive and versatile for establishing and maintaining a website. Its interactive feature make it superior than other advertising mediums. Below are the some advantages offered by WWW:

1. ***Presence on the web:*** It enables businesses to be in touch with several million people who have access to the World Wide Web with more and more added every day. No business can afford to ignore this many potential customers.

2. ***Networking:*** It helps in developing lines of communication to enhance contact with potential clients and organizations. It helps in speedy and reliable communication and advertisement.

3. ***Provide business information:*** It facilitates the websites to publish business services, hours, location, phone and e-mail for the public to view like any printed form of advertising. Unlike the conventional advertisement, the website provides instant communication with information about the business that may change regularly.

4. ***Service to customers:*** A website provides access to business information and services to their customers online that may not be available any other way. The customers can be from anywhere in the world and shop in online stores like never before and from the comfort of their homes. They can easily and quickly search the database to locate the exact item that they were looking for and purchase it online.

5. ***Conduct business:*** A website may provide means of doing business.

6. ***Provide files to download:*** Details and information of products and services in the form of pamphlets, brochures, advertisements, and even a demonstration video can be down loaded from the company's website.

7. ***Remote office access:*** It facilitates offices and employees to be in touch with one another from remote places to accomplish their tasks effectively.

11.2.2 Designing a Web Page

HTML Basics

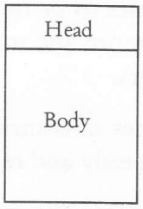
HTML is a simple scripting language that is used to write WebPages. It is an abbreviation and stands for Hyper Text Markup Language that is predominantly markup language for the creation of web pages. Hypertext in HTML is simply a piece of text that works as a link. It is basically a text file containing small markup tags as headings, paragraphs, lists, and so on. It also supplements the text with interactive forms, embedded images, and other objects. Its file extension is htm or html. It can be created using a simple text editor. These markup tags guide the Web browser as to how display the page.

When an HTML file is opened in a web browser, the browser looks for HTML codes in the text and use them to change the layout, insert images, or create links to other pages. As HTML documents are only text files therefore they can be written in even the simplest text editor. FrontPage and Dreamweaver are some of the most popular HTML editors.

In general, a computer takes an "A" as simply an "A" whether it is bold, italic, big or small. In order to tell the computer that "A" should be italic or bold, a software program namely Browser is used which is specified with a markup in front of the A. Such a markup is known as a Tag. It is customary that all HTML tags should be enclosed in < and >. For example: bold. It makes the text bold.

HTML Page Structure

A web page is consisted of a head and a body as shown in the figure below:

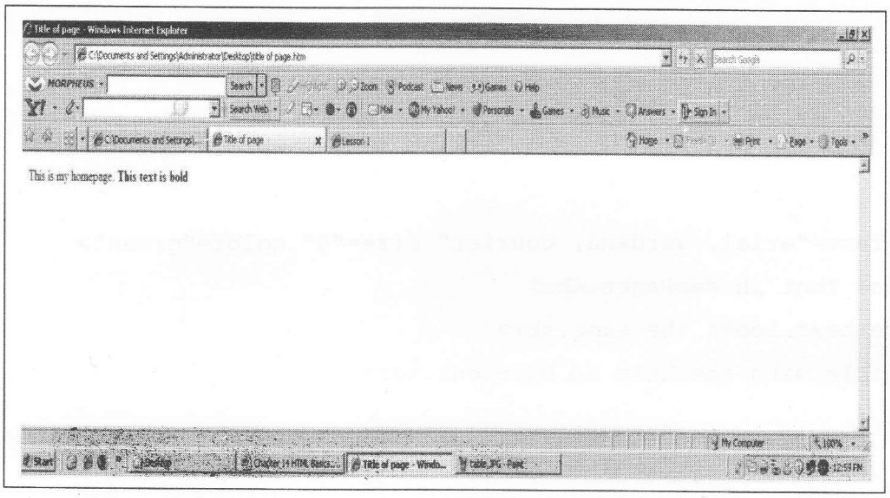


The purpose of using head is not to show the text and tags directly on the page while in case of body they appear directly on the web page. All web pages starts with an <html> tag at the beginning and the end, indicating the browser where the document starts and where it stops. Metatags are used in the head section for, among other things, to improve the rankings in search engines. Often, the head section also includes javascript, which is a programming language for more complex HTML pages.

HTML uses tags that are used to mark-up HTML elements and are surrounded by the two characters < and > and are called angle brackets that indicate start and end of the tag. These tags are not case sensitive.

Take an example:

```
<html>
<head>
<title>Title of page</title>
</head>
<body>
This is my homepage. <b>This text is bold</b>
</body>
</html>
```



The above example is an element. Let us take <title>Title of page</title>. As it has been indicated earlier that the HTML element starts with a start tag: <title> and the content of the HTML element is: Title of Page. It ends with an end tag: </title>.

Tags can have attributes to provide additional information about the HTML elements on the web page. Attributes always written as name/value pairs like this: name="value". They are always added to the start tag of an HTML element.

For example

Attribute tag defines the body element of an HTML page: <body>. Using bgcolor attribute, browser can be enabled to define the background color of web page. For example if background color is blue, it should be like this: <body bgcolor="blue">.

Attribute also defines an HTML table: <table>. Using an added border attribute, the browser could be enabled to define the border. If the table has no border, it should like this: <table border="0">.

Other important tags in HTML are headings, paragraphs and line breaks. It also enables to define a lot of elements for formatting output, like bold or italic text.

HTML Text

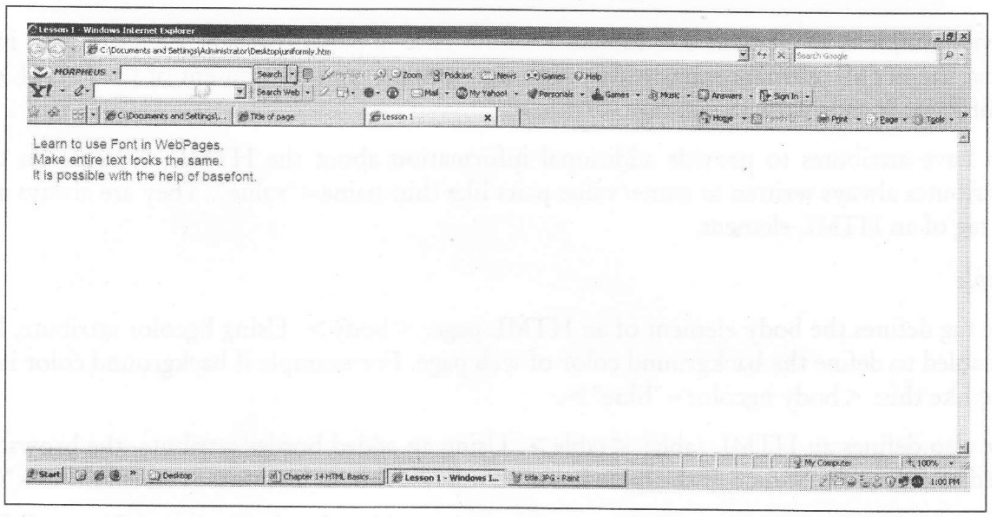
HTML text is entered in the same manner as any text is entered. It requires attributes to be defined for indicating size, font etc. The web browsers show the fonts available on the user's system.

In order to specify the overall font for web page, use the <basefont> tag at the beginning of the <body> section. It can be comprehended with the following example:

```
Learn to use Font in WebPages.
Make entire text looks the same.
It is possible with the help of basefont.
```

In order to display the above text uniformly, write the following HTML code using <basefont> tag:

```
<html>
<head>
<title>Lesson 1</title>
</head>
<body>
<basefont face="arial, verdana, courier" size="4" color="green">
Learn to use Font in WebPages.<br>
Make entire text looks the same.<br>
It is possible with the help of basefont.<br>
</body>
</html>
```



You may notice the use of color attribute in the above HTML code, which selects the desired color for the text. Similarly the face attribute specifies the desired font. We may now notice that the tag is used to change the font. Following tables shows the tags for text formats:

text	Text as bold
text	Text in italics
<u>text</u>	Text as underline
_{text}	Lowers text and makes it smaller
^{text}	Supcripts text and makes it smaller
<strike>text</strike>	Strikes a line through the text
<big>text</big>	Increases text size by one
text	Decreases the text size by one
text	Text in biggest heading
text	Text in smallest heading
text	Text in smallest fontsize (8 pt)
text	Text in biggest fontsize (36 pt)
text	Adds a paragraph break after the text (2 linebreaks)
text	Left justify text
text	Center text
text	Right justify text
text 	Adds a single linebreak

Home Page

The home page is an index to other pages on that site that you can jump to by clicking an underlined hyperlink or an icon. Links on that site may take you to other related sites.

HTML links

Links that are the most fundamental part of the World Wide Web provides the facility to navigate from one web page to another web page. Links can be classified in three categories. These are links to anchors on the current page, links to other pages within the current site and links to pages outside the current site. Links could be provided for both texts and images.

Links available with HTML are hyperlinks that enables user to link to another document on the Web. To facilitate hyperlinks there are the anchor tag <a> and and the Href attribute, which are used by HTML. An anchor can be used to point to any resource on the Web such as an HTML page, an image, a sound file, a movie, etc. The syntax for creating an anchor is: Text to be displayed. For example - .

The href attribute is used to address the document to link to and the words between the open and close of the anchor tag will be displayed as a hyperlink. Target attribute is used to define where the linked document will be opened. The name attribute is used to create a named anchor. When using named anchors, links can be created that can jump directly into a specific section on a page, instead of letting the user scroll around to find the needed document. Below is the syntax of a named anchor: Text to be displayed.

Defining colors for the HTML links: With the use of a few settings, colors for all links could be defined on the WebPages. The general color of text links which is blue before the click is given in the

<body> tag as: <body link="#C0C0C0" vlink="#800080" alink="#FF0000">. vlink, indicates that the link has been visited by the user and standard color is purple. alink specifies active link which means the color of the link when the mouse is on it. Its standard color is red.

In order to define more links to have different colors than the rest of the page, we can either place the font tags between the <a href> and the tag: here or using a style setting in the <a> tag: here.

Defining link targets: It is evident that a link is usually open in the current window or frame by default. However, in some cases it is required to be opened in another window or frame. It is accomplished by adding a target="" to the <a href>. For example: <a href=http://www.hotmail.com target="_blank">. The _blank loads the page into a new browser window.

Other targets are _self; _parent; and _top that load the web page into the current window, the frame that is superior to the frame the hyperlink is in and cancels all frames, and loads in full browser window respectively.

Defining link within a page: It is required to create a link pointing to the anchor. As it has already been mentioned that an anchor is created using the <a> tag. For example, if an anchor is created for TOYS in an online shopping mall, this word Toys is simply added where it is anchored. The HTML cod is: and then a link pointing to the anchor using the normal <a href> tag, like this: here.

When it is required to create a link to anchors on external WebPages, the syntax is: blabla

Defining links for a frameset: A link in a frameset may provide link to a web page that is loaded in the other frame window. Take an example website having tutorials in a frameset called Contents where different links such as Lesson 1, Lesson 2, Lesson 3 etc are created. The HTML code to go at Lesson 3 will be like: Lesson 2 of the tutorial.

Defining image link: A technique called image mapping is used to link one image to several pages by simply specifying which of the areas of the image should link to where. In other words it explains that an user can go to different websites by simply clicking at different portions of an image. For example:

```
<img src="rainbow.gif" usemap = #example border=0>
<map name=example>
<area shape=circle coords=0,0,29,29 Href="http://www.hotmail.com">
<area shape=circle coords=30,30,59,59 Href="http://www.google.co.in">
</map>
```

In the example above, if mouse is clicked at the upper left corner it links to hotmail website and if it is clicked at the lower right corner, it links to Google website.

For other shapes we can use:

```
<area shape=rect coords= x1,y1,x2,y2 Href="http://www.domain.com"> for Rectangles
<area shape=circle coords= x1,y1,x2,y2 Href="http://www.domain.com"> for Circles
<area shape=polygon coords= x1,y1,x2,y2,..,xn,yn Href="http://www.domain.com">
for Polygons
```

Defining link to the new window: In order to open a page in a new window use the target="_blank" in the <a href> tag. It simply opens a new browser window that will load the linked page. For example linking to the hotmail, the link will be like this: Go to Hotmail.

Defining links to send an email: In order to send email, links are created almost in a similar manner as it is done to link other pages: <a href> tag. The HTML code for the email link is: Email Me.

If a special subject is needed to be added in the email, it can be done using subject= setting: Send Email.

An email link for specific text in the body of the message can be accomplished by simply adding &body=: Send Email

All the above options can be combined in a single email. It will look like: Email Me

Build a simple HTML Document Tables

Tables on websites intend to provide arranged information and create a web page layout with the use of hidden tables. The tables are used to divide the web page into different sections. HTML enables us to create tables. They are defined with the <table> attribute tag. A table has rows (with the <tr> attribute tag, where tr stands for table row). Each row has data cells (with the <td> attribute tag where td stands for table data). A data cell includes text, images, lists, paragraphs, forms, horizontal rules, tables, etc as data. Briefly, each table starts with a table tag and each table row starts with a tr tag. Each table data in a row starts with a td tag. From the following example, it could be understood:

Example:

```
<table border="1">
<tr>
<td>R1, C1</td>
<td>R1, C2</td>
</tr>
<tr>
<td>R2, C1</td>
<td>R2, C2</td>
</tr>
</table>
```

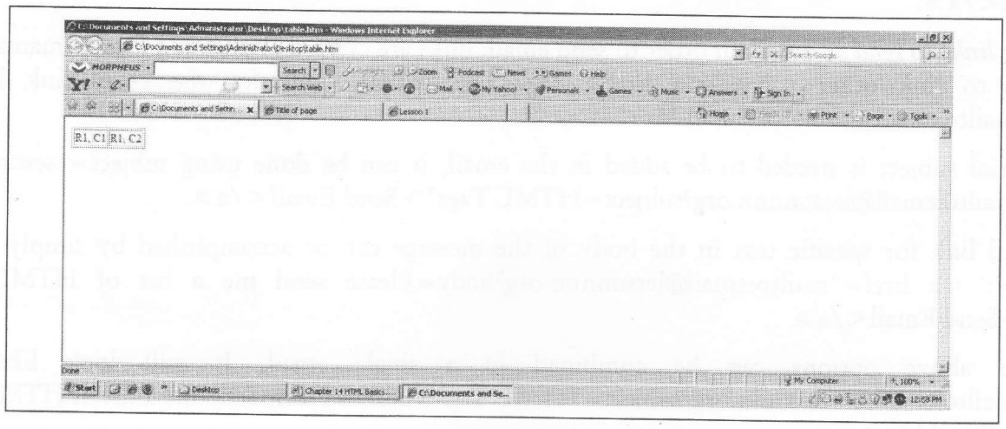
It looks like this:

R1, C1	R1, C2
R2, C1	R2, C2

ables and the Border Attribute: Without border attribute the table will be displayed without any borders. In order to display a table with borders, it is required to use the border attribute:

```
table border="1">
<tr>
<td>R1, C1</td>
```

```
<td>R1, C2</td>
</tr>
</table>
```



Headings in a Table: Headings in a table are defined with the <th> attribute tag.

```
<table border="1">
<tr>
<th>Heading1</th>
<th>Heading2</th>
</tr>
<tr>
<td>R1, C1</td>
<td>R1, C2</td>
</tr>
<tr>
<td>R2, C1</td>
<td>R2, C2</td>
</tr>
</table>
```

Heading1	Heading2
R1, C1	R1, C2
R2, C1	R2, C2

Empty Cells in a Table: Table cells without content are not displayed very well in most browsers.

```
<table border="1">
<tr>
<td>R1, C1</td>
<td>R1, C2</td>
</tr>
<tr>
<td>R2, C1</td>
```

It looks like this:

R1, C1	R1, C2
R2, C1	

```
<td></td>
</tr>
</table>
```

It could be noticed that the border around empty cell is missing. However some browsers also support the feature that in this particular case the border around empty cell should not be missing. This could be avoided by adding a non-breaking space ( ) to empty data cells:

```
<table border="1">
<tr>
<td>R1, C1</td>
<td>R1, C2</td>
</tr>
<tr>
<td>R2, C1</td>
<td> </td>
</tr>
</table>
```

It looks like this:

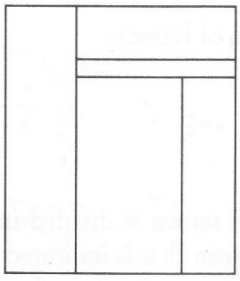
R1, C1	R1, C2
R2, C1	

Briefly the following table defines the table tags:

Table Tags	Tag Description
<table>	A table
<th>	A table header
<tr>	A table row
<td>	A table cell
<caption>	A table caption
<colgroup>	Groups of table columns
<col>	Attribute values for one or more columns in a table
<thead>	A table head
<tbody>	A table body
<tfoot>	A table footer

Frames

It is for displaying more than one HTML document in the same browser window or to divide the screen into separate windows. An example of frame is shown below:



Each of the above windows contains a HTML document. In fact, each HTML document is called a frame and each frame is independent of the others. Using frames in developing web pages have certain disadvantages such as the web developer has to keep track of more HTML documents and it becomes difficult to print the entire page.

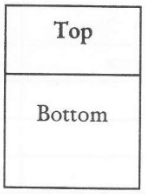
A file, which explains how the screen is sectioned into frames, is called a frameset. When a frameset page is loaded to a web browser, it automatically loads each of the web pages associated with the frames. The frameset tag <frameset> enables a web developer to divide the window into frames wherein it includes a set of rows or columns. The values of the rows/columns indicate the amount of screen area each row/column will occupy.

The frame tag <frame> defines which HTML document can be put into each frame. In the given example, there are a frameset with two columns. The first column is set to 30% of the width of the browser window. The second column is set to 70% of the width of the browser window. The HTML document "frame_1.htm" is put into the first column, and the HTML document "frame_2.htm" is put into the second column.

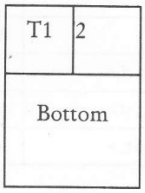
```
<frameset cols="30%,70%">
    <frame src="frame_1.htm">
    <frame src="frame_2.htm">
</frameset>
```

Some examples of different framesets are being given below:

1.
```
<frameset rows="25%,75%">
    <frame src="top.htm" name="top">
    <frame src="bottom.htm" name="bottom">
</frameset>
```



2.
```
<frameset rows="25%,75%">
    <frameset cols="50%,50%">
        <frame src="Tl.htm" name="Tl">
        <frame src="T2.htm" name="T2">
        </frameset>
        <frame src="bottom.htm" name="bottom">
    </frameset>
```



We will now learn in the following steps the designing of frames:

Designing a frame: The HTML code:

```
<frameset cols="120,*">
</frameset>
```

will produce a frame that looks like in a design with screen is divided into two columns. The left is having 120 pixels and the right using the rest of the screen that is indicated by the *.

Adding default pages: It can be added to frame windows with the src setting. When these pages are loaded, the frameset is opened the first time. Names to each frame window using the name setting can also be inserted in order to make a link in one frame window open a page in another frame window. For example:

```
<frameset cols="120,*" >
<frame src="menu.htm" name="menu" >
<frame src="frontf.htm" name="main" >
</frameset>
```

M e n u	Main

Frame borders: In case of invisible border, it is required to simply add the parameters "cols-line" to the frameset:

```
<frameset cols="120,*" frameborder="0" border="0" framespacing="0">
<frame src="menu.htm" name="menu" >
<frame src="frontf.htm" name="main" >
</frameset>
```

Resizable windows: The frame windows can be designed with resizeable with the help of the "noresize" to the frame src lines:

```
<frameset cols="120,*" frameborder="0" border="0" framespacing="0">
<frame src="menu.htm" name="menu" noresize>
<frame src="frontf.htm" name="main" noresize>
</frameset>
```

Adding scroll bar in the menu window: A scrollbar can be added in case of long documents with the help of "noresize scrolling=auto" to the frame src lines:

```
<frameset cols="120,*" frameborder="0" border="0" framespacing="0">
<frame src="menu.htm" name="menu" noresize scrolling=no>
<frame src="frontf.htm" name="main" noresize scrolling=auto>
</frameset>
```

Designing links with HTML frames: A link in the menu window can be used to load a web page in the main window with the help of the following code:

Text. It can be noticed that a parameter target="main" to the <a href> tag accomplishes the task. This enables the link to be opened in the "main" frame window instead of the "menu" frame window where the link itself is located. There are four target names are available:

1. _blank for a new browser window
2. _self for the current window
3. _parent loads the page into the frame that is superior to the frame the hyperlink is in.
4. _top cancels all frames, loads in full browser window.

HTML Images

Images can be added to a web page and they can also be customized with alignments. Mages can also be used for creating links, which has already been explained under links in the same unit earlier. Images in computers are stored in several different ways depending upon the needs. GIF and JPG are some the popular compression image format techniques that help in reducing download times as much as possible. Their characteristics are given in the following table:

Characteristics	JPG	GIF
Colors	Unlimited	256
Transparent images	Can not handle	Can handle
Compression	Excellent	Not good
Uses	Compressing photographs and complex images	Banners, buttons and clipart

Inserting an image: An image could be inserted in an HTML document with the use of tag img. A HTML code to insert the image of Tajmahal on the Web page namely My Spaceeee: <img src="http://www.myspaceeee.in/tajmahal.gif">. The code even become simpler if the image is stored in the same folder as the HTML page: <img src="tajmahal.gif">

Resizing: The size of an image can be altered using the width and height attributes. However it is advisable to reduce the size in a graphics program rather than reducing the size on the web page using the width and height attributes. For example:

<img src="http://www.myspaceeee.in/ tajmahal.gif" width="60" height="60"> can be changed to <img src="http://www.myspaceeee.in/tajmahal.gif" width="120" height="120">. In case of no setting for width and height, at that moment the browser will automatically use the real size of the image. However, it is advisable to always enter the settings for width and height, even when using the real size.

Border: Border can be added to an image with the help of the following HTML code: <img src="http://www.myspaceeee.in/tajmahal.gif" border="5">

Alternative Text - An alternative text can also be added to an image using the alt setting shown in the example below:

```
<img src="http://www.myspaceeee.in/tajmahal.gif" alt="Taj is for beauty">.
```

Spacing around the image: Space over and under an image can be added with the Vspace attribute. Space to the left and right of the image can be inserted using the Hspace attribute:

```
<img src="tajmahal.gif" Hspace="30" Vspace="10">
```

Spacing on one side of the image is done using a 1x1 pixel transparent gif image. A 10-pixel spacing to the left of the image can be done with the help of pixel.gif:

```
<img src="pixel.gif" width="10" height="1"><img src="tajmahal.gif">
```

Image alignments: Images can be aligned according to the text around it with the help of the alignments parameters. These are for default, left, right, top and texttop aligns etc.

Microsoft Front Page

WebPages have contributed too much in providing the Internet popularity and transforming this world into global village. The WebPages available at different locations are required to be update and

created for which, a programmer needs editor. Earlier Notepad or similar editors were used. Over the years, the WebPages have been becoming more complex, so it is too difficult to program web pages manually using normal word editors. Normal word editors do not allow graphics to be created. Consequently, many companies came out with their own web editors.

Microsoft provides Microsoft Office FrontPage as HTML editor and web site administration tool for the Microsoft Windows compatible operating systems. FrontPage is used to hide the details of pages' HTML code from the user. However, it enables novices to easily create web pages and websites. Advantages of FrontPage are its being simple and functional, scaleable, secure and better performance. The major advantage of using Microsoft FrontPage is that the interface of Microsoft FrontPage is the same as Microsoft Word. Hence, it is easier for persons using Microsoft Word to deal with FrontPage's tools. Moreover, the software software provides amazing themes with which useful websites can be created with the help of lot of web buttons, banners, and images in this software. The beauty of the Microsoft's Front Page is that instead of learning different types of tags HTML, FrontPage enables to work in a WYSI-A-WYG (What You See Is -Almost- What You Get) environment that facilitates a great deal like the Microsoft word processing application, Word.

On the other hand, this software has two disadvantages. It generates lots of redundant codes in the web page programming. The web pages which are created with Microsoft FrontPage could be browsed clearly by Microsoft Internet Explorer, but they couldn't be browsed well by other web browsers such as Mozilla Firefox. It requires a web server in order to run some dynamic activities.

Some features of FrontPage include:

1. It helps navigating through web site and looking the file structure and its contents visually.
2. Front Page provides built-in features for HTML, JavaScript, etc.
3. It contains image editor namely Microsoft Image Composer.
4. It provides point-and-click functionality for common tools suchas mouseovers, e-mail forms, and hit counts.
5. It is very simple to use with previous knowledge of Office products.
6. It provides integrated data display with Office products like Access and Excel.
7. It supports ASP.NET master pages.
8. It provides dynamic changes to all links of a page if its URL is changed.
9. It assigns task for team projects.
10. Content can be edited from anywhere with FrontPage with password.
11. It supports rich clipboard data import (i.e. copy/pasting data from Internet Explorer into FrontPage. It will automatically download media resources such as images and save them locally.
12. It contains built-in support for automated web templates including automatically generated multi-level navigation system.

However, Microsoft in 2007 has stopped offering FrontPage and has been providing two kinds of softwares, Microsoft SharePoint Designer and Microsoft Expression Web.

About Web Servers

The Web page needs to be located on a Web server and therefore a space is needed for Web hosting services. A web server is not simply a server that provides access to shared resources. Support and help from system administrators is sought for specific instructions on how to interact with web server. HTML is used to design a web page from scratch. FrontPage is considered an HTML editor in its simplest form. A FrontPage document is available in Programs in any Microsoft Office suite.

Microsoft Front Page Basics

Microsoft Front Page provides the following facilities for building Web Pages:

- *Testing the Web Page:* It provides a support for previewing of the Webpage being designed in between the design to ensure that it is being created as per the design. It also provides the designer to look into the HTML codes that are creating the Webpage.
- *Local Web Browser File:* Front Page provides testing of a Webpage by viewing it in the local Web browser.

Some of the design features of the Microsoft Front Page software are listed below:

Creating a Web Page Title

1. *Identify the Web Page under design to Search Engines and Browsers:* The search engines and other browsers identify the web page by meta information. For this purpose certain keywords are developed that could be used to search a web page on a website.
2. *Setting the Appearance of the Page:* The colors for background, text, hyperlinks, visited links, wallpaper, theme, active links etc are set to enhance the appearance and usability of the webpage under design.
3. *Adding and Formatting Text:* Like word processor, the text can be types in FrontPage in the same manner. However, formatting text needs different tools.
4. *Using Hyperlinks:* Hyperlink enables to jump to a specific location within a Web file.
5. *Using Images:* Images can be used from many sources in Web pages. They may be purchased images, scanned diagrams and photographs, copied from the Web or created images with a graphics application.
6. *Using Bulleted and Numbered Lists:* They are an effective way to organize the content with scan-ability.
7. *Using Tables:* Tables are frequently used to organize information.
8. *Using FrontPage with Existing Documents:* An existing document created in a word processing application may be inserted to create a Web document easily without re-typing the text.
9. *Advanced and Special Features:* Front Page is also a powerful Web management tool that helps in managing the webpage activities.

11.3.3 Web Browsers

A web browser is the software program used to access the World Wide Web that is the graphical portion of the Internet. The first browser, called NCSA Mosaic, was developed at the National Centre

for Supercomputing Applications lab at Illinois in the early 1990s. The easy-to-use point-and-click interface fuelled the growth of the Web. The web browser software program facilitates a user to display and interact with text, images, videos, music and other information typically located on a web page at a website on the www or a LAN. Text and images on a web page normally designed to provide hyperlinks to other web pages at the same or different website. Thus web browser enables point to point click to reach directly to the targetted web pages.

Some of the popular web browsers are Internet Explorer, Mozilla Firefox, Safari, AOL Explorer, etc. Web browsers belong to HTTP user agent category. Browsers also provide in accessing information provided by web servers in private networks or content in file systems.

Web browsers talk with web servers using HTTP (hypertext transfer protocol) to retrieve web pages. HTTP enables web browsers to provide information to web servers to retrieve web pages from them. Different web pages have a URL address starting with http:// for HTTP access. There may different other URL types and their corresponding protocols and most of the browsers supports them. FTP (file transfer protocol) is one of the examples of such types. Other examples are rtsp: for RTSP (real-time streaming protocol) and https: for HTTPS (an SSL encrypted version of HTTP).

The file format for a Web page is normally HTML (hyper-text markup language) and is identified in the HTTP protocol with a MIME content type. Most of the browsers support different formats such as the JPEG, PNG and GIF image formats including HTML. The combination of HTTP content type and URL protocol specification enables designers to embed images, animations, video, sound and streaming media into a web page or to make them accessible through the web page.

Search Engines

A search engine is an information retrieval system to access and retrieve information stored on WWW or a computer system attached to the Internet. Search engines enable to minimize the time required to find information and the amount of information on a computer system. The computer syatem may be a standalone system or attached to the Internet. The search engines are popular among people as web search engines to explore information on the World Wide Web.

Search engines are interface to a group of contents that allows users to specify criteria about the content by providing some keywords so that the engine can find the several matching contents to the corresponding keywords out of million of webpages. The keywords provided are referred to as a search query. Several styles of search query syntax are used. Search query differs for different types of serch engines, whereas some search engines enable users to enter two or three words separated by space while other search engines may require users to provide entire documents, pictures, sounds, and various forms of languages. Some search engines attempt to enhnce the search queries to provide a quality set of items through a process known as query expansion.

Index-based Search Engine

In such engines, the list of items to meet the criteria specified by the query is typically sorted or indexed. Indexing contents by relevance from highest to lowest minimizes the time needed to explore the desired information. Some search engine uses probabilistic approch to rank contents based on measures of similarity, popularity or authority. Boolean search engines typically provide contents which match exactly without regard to order, although the term boolean search engine may simply refer to the use of boolean-style syntax. Thus, in order to provide a set of matching contents that are sorted based on some criteria quickly, a search engine will typically collect metadata about the group

of contents under consideration through a process called as indexing. The advantge of index is that it needs a smaller amount of computer storage.

Types and Characteristics

Some of the popular search engines with their types and characteristics are given below:

- *Alta Vista:* It is a crawler type which results ranked based on how many times search words appear in the text. It searches full text.
- *Excite:* It is also a crawler type and uses meta tags.
- *Google:* It is also a crawler type that performs based on the number of times other sites are linked to the ranked site.
- *Yahoo:* It is crawler type and performs similar to Google.

Check Your Progress

1. What is FTP?
2. Define Telnet.

11.4 LET US SUM UP

The uppermost layer of OSI models provides a number of services to the users using the TCP/IP protocol. Domain Name System (DNS) provides the quick translation of text of the IP addresses within fraction of seconds from a directory of billions of such addresses. This could be made possible by using Domain concepts, which uses hierarchical arrangements of text addresses translation. The servers maintaining addresses are distributed and have locations throughout the world. Electronic mail is one of the most popular network services and uses user agent and message transfer agent to transport messages created by a user to destination mailboxes possibly on remote machines. Multimedia applications have enthused life in webpages making them interactive. The convergence of different media such as text, pictures, video and sound into a single media has contributed enormously for the growth of Internet and www. Applications of multimedia packages are found in all walks of life. With the advancement and innovation in presentation tools of multimedia, the multimedia applications have been giving impressions of virtual reality to its end users.

Protocols working with TCP/IP like File Transfer Protocol (FTP) and Trivial File Transfer Protocol (TFTP) are file server access protocol that enable a user to transfer files between two hosts across the network or Internet using TCP. Simple Mail Transfer Protocol (SMTP) that is used to transfer mails from one computer system to another computer system attached to the same network or different networks uses end-to-end delivery in which an SMTP client machine contacts the destination host's SMTP server directly to deliver the mail. Telnet is an Internet standard remote login protocol to connect a local terminal with a remote login session. It copies keystrokes to the remote machine and copies output from the remote machine to the source machine.

The World Wide Web that has become de facto standards for any professional belonging to any discipline finds its extensive use and utility in providing information stored in computer system attached to the Internet or www. Different web designing techniques are used to make the information in a presentable form to the users. Some of the web designing techniques such as HTML has become a standard for web pages. In addition to the above, In HTML webpage design, the limited use of colors

often makes the appearance of the colors more powerful. It is also possible to add an image or a plain color as background with the help of its specifications in the <body> tag. Form also gives navigability to a website. Forms are objects that enable to enter information in the form of text boxes, drop-down menus or radio buttons. Front Page provided by Microsoft, however, provides an excellent tool for designing WebPages with very minimal knowledge of HTML has been replaced by more advanced tools. FrontPage enables to work in a WYSI-A-WYG (What You See Is -Almost- What You Get) environment that facilitates a great deal like the Microsoft word processing application, Word. Web browser is software to run the hypertext transmission protocols to retrieve a webpage from its destination. The browser program residing in the end user computer facilitates the advantages of text formatting, hypertext links, images, sounds, motion, and other features. Internet Explorer and Netscape are some of the widely used browsers. Browsers have sub programs called plug-ins to handle the documents you find on the Web. It may also other plug-ins stored elsewhere in your computer. Search engines with different approaches provide an excellent method of searching WebPages on www in minimal possible time. Contribution of search engines is enormous to the growth of www and Internet.

The seventh layer of the OSI model is the application layer providing user related applications such as electronics mail, remote file transfer, remote login, etc. The ITU-T has recognized some of the application for which standardization is possible. Some of them are message handling system, file transfer, access and management, virtual terminal, directory systems and common management protocol.

11.5 KEYWORDS

Frames: It is for displaying more than one HTML document in the same browser window or to divide the screen into separate windows.

Home Page: It provides a point of entry to a Website with help. It also contains all relevant links of that particular website.

HTML: Hypertext Markup Language defines the rules for formatting a web page so that a web browser displays the page properly.

World Wide Web (WWW): It is defined as a client-server information system using the Internet to access computers containing millions of hypertext documents.

Hypertext: It defines the documents containing embedded links (hyperlinks) to other documents or other parts of the same document.

Hyper Text Transport Protocol (HTTP): Hypertext Transfer Protocols are the rules that enable the transmission of web documents from one computer to another via the Internet.

Search Engines: They are software that enables searching of the content available on Internet.

URL - It denotes Uniform Resource Locator. It is the address of a document on the World Wide Web.

Web Browser: It is the client software used to explore and display web pages from a website.

Web Client: It refers to the computer and software used to access a website and web pages.

Web Page: It is a single hypertext document written in Hypertext Markup Language (HTML) and described in HTML basics. This normally contains the basic information and links to navigate in the websites to which it belongs.

Web Site: It is written in HTML and is a collection of linked Web pages on a Web server. Web server is the machine where a website is located or hosted. It may be organization owned or Internet Service Provider (ISP) owned.

Web Server: It is the computer or server which provides a space for hosting a website. The web client access web servers to retrieve information from a website.

11.6 QUESTIONS FOR DISCUSSION

1. Explain the different types of search engines with their characteristics.
2. What is relevance of Front Page in designing web pages when most of the web pages are designed in HTML?
3. How is the HTML document used for making a hyperlink? Explain with example.
4. How a HTML document can be created from normal word document?
5. Define the role of frames in html.
6. How does SMTP work in transferring mails from one computer system to another computer system attached to different networks?
7. Differentiate between HTTP and FTP. Give some advantages of both protocols.

Check Your Progress: Model Answers

1. FTP is a file server access protocol that enables a user to transfer files between two hosts across the network or Internet using TCP.
2. Telnet is an Internet standard remote login protocol to connect a local terminal with a remote login session.

11.7 SUGGESTION READINGS

Rajneesh Agrawal and Bhata Bhushan Tiwari, *Computer Networks and Internet,* Vikas Publication

Behrouz A. Forouzan, Sophia Chung Fegan, *Data Communications and Networking*, McGraw-Hill Companies

Andrew S. Tanenbaum, *Computer Networks*, Prentice Hall

Achyut S Godbole and Atul Kahate, *Web Technologies*, Tata McGraw Hill

J. D. Spragins, *Telecommunications Protocols and Design*, Addison-Wesley, Reading MA

Ferguson P., Huston G., *Quality of Service: Delivering QoS on the Internet and in Corporate Networks*, John Wiley & Sons, Inc., 1998

Spurgeon, Charles E. *Ethernet, The Definitive Guide*. O'Reilly & Associates, 2000.

Nassar, Daniel J. *Ethernet and Token Ring Optimization*. iUniverse.com, 2000. ISBN: 1583482199.

McDysan, David E. and Darren L. Spohn. *ATM Theory and Applications*, McGraw-Hill Osborne Media, 1998.

William A Shay, Understanding Communication and Networks 3rd Edition Thomson Press