ଓଡ଼ିଶା ରାଜ୍ୟ ମୁକ୍ତ ବିଶ୍ୱବିଦ୍ୟାଳୟ, ସମ୍ବଲପୁର, ଓଡ଼ିଶା
Odisha State Open University, Sambalpur, Odisha
Established by an Act of Government of Odisha.

DIPLOMA IN COMPUTER APPLICATION

DCA-03

NETWORK FUNDAMENTALS

BLOCK

3

NETWORK PROTOCOLS AND SECURITY

Unit-1

Network Layer Protocols

Unit-2

Transport Layer Protocols

Unit-3

Application Layer Protocols

Unit-4

Internet and its Services

Unit-5

Network Security

ଓଡ଼ିଶା ରାଜ୍ୟ ମୁକ୍ତ ବିଶ୍ୱବିଦ୍ୟାଳୟ, ସମ୍ବଲପୁର, ଓଡ଼ିଶା
Odisha State Open University, Sambalpur, Odisha
Established by an Act of Government of Odisha.

EXPERT COMMITTEE

Dr P.K.Behera (Chairman)
Reader in Computer science
Utkal University
Bhubaneswar, Odisha

Dr.J.R.Mohanty (Member)
Professor and HOD
KIIT University
Bhubaneswar, Odisha

Sh PabitranandaPattnaik (Member)
Scientist-E, NIC
Bhubaneswar, Odisha

Sh Malaya Kumar das(Member)
Scientist-E, NIC
Bhubaneswar, Odisha

Dr. Bhagirathi Nayak (Member)
Professor and Head(IT & System)
Sri Sri University
Bhubaneswar,Odisha

Dr.ManoranjanPradhan(Member)
Professor and Head(IT & System)
G.I.T.A
Bhubaneswar, Odisha

Sr V.S Sandhilya(Convener)
Academic Consultant (I.T)
Odisha State Open University
Sambalpur,Odisha

DIPLOMA IN COMPUTER APPLICATION

Course Writer

Chandrakant Mallick
Consultant (Academic)
School of Computer and Information Science
Odisha state Open University

UNIT-1
NETWORK LAYER PROTOCOLS

UNIT STRUCTURE

1.0 INTRODUCTION

At the physical and data link layers, TCP/IP reference architecture does not define any specific protocol. It supports all the standard and proprietary protocols in the underlying network. A network in a TCP/IP internetwork can be a local-area network or a wide-area network.

The network layer, also called the TCP/IP Internet Layer, holds the whole architecture together. Its job is to permit hosts to inject packets into any network and have they travel independently to the destination (potentially on a different network). They may even arrive in a different order than they were sent, in which case it is the job of higher layers to rearrange them, if in-order delivery is desired.

The internet layer defines an official packet format and protocol called IP (Internet Protocol).

The job of the internet layer is to deliver IP packets where they are supposed to go. Packet routing is clearly the major issue here, as is avoiding congestion.

At the network layer/ Internet Layer, TCP/IP supports the Internetworking Protocol. IP, in turn, uses four supporting protocols: ARP, RARP, ICMP, and IGMP. In this unit we will discuss each of these protocols in the network.

1.1 LEARNING OBJECTIVES

After going through this unit, you should be able to:

- Know the basic functions of the network layer
- Understand the concepts of internetworking
- Know the functions internetworking protocols
- Understand the IP addressing concepts
- Know the field format of IP datagram
- Know the address resolution concepts
- Know process of Error reporting protocols at network layer

1.2 NETWORK LAYER

The network layer is responsible for source-to-destination delivery of a packet across multiple networks. It ensures that each packet gets from its point of origin to its final destination. It does not recognize any relationship between those packets. It treats each one independently as though each belong to separate message.

Network Layer Functions

The functions of the network layer include the following.

Internetworking: The logical gluing of heterogeneous physical networks together to look like a single network to the upper layers.

Logical Addressing: If a packet has to cross the network boundary then the header contains information of the logical addresses of the sender and the receiver.

Routing: When independent networks or links are connected to create an internetwork or a large network the connective devices route the packet to the final destination.

Packetizing: Encapsulating packets received from the upper layer protocol. Internet Protocol is used for packetizing in the network layer.

Fragmenting: Router has to process the incoming frame and encapsulate it as per the protocol used by the physical network to which the frame is going.

1.3 INTERNETWORKING

An **internetwork** is a collection of individual networks, connected by intermediate networking devices, that functions as a single large network. Internetworking refers to the industry, products, and procedures that meet the challenge of creating and administering internetworks.

An internetwork is also referred to as a network of networks because it is made up of lots of smaller networks. The nodes that interconnect the networks are called routers. They are also sometimes called gateways.

Internetworking is the way of connecting two or more computer networks via network devices such as routers or gateways using a common routing technology. The result is called an internetwork (often shortened to internet).

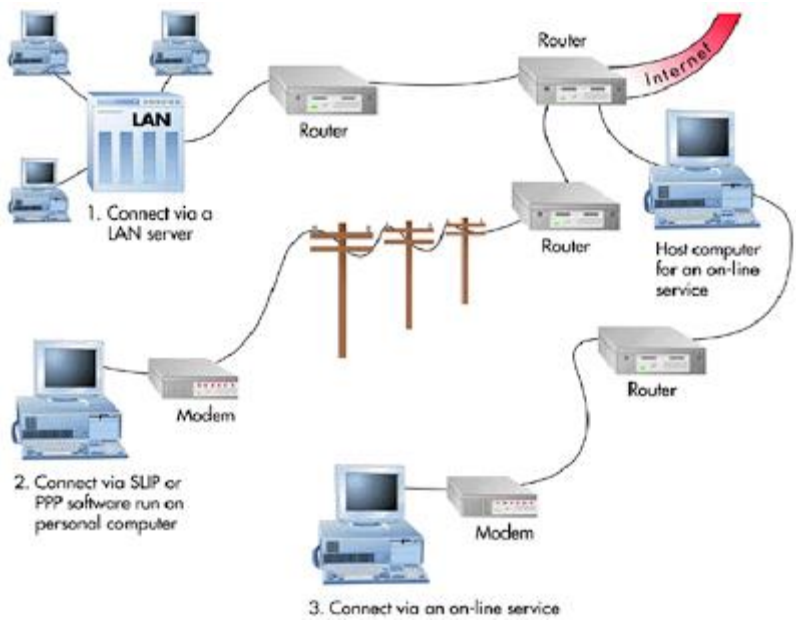


Fig: Internetwork

The figure above illustrates some different kinds of network technologies that can be interconnected by routers and other networking devices to create an internetwork.

The Internet Protocol or the Internetworking Protocol (IP) is the key tool used today to build scalable, heterogeneous internetwork.

1.4INTERNET PROTOCOL (IP)

The Internetworking Protocol or simply Internet Protocol (IP) is the transmission mechanism used by the TCP/IP protocols stack. It is an unreliable and connectionless protocol. It provides a best-effort delivery service. The term *best effort* means that IP provides no error checking or tracking.

IP transports data in packets called *datagrams*, each of which is transported separately. Data grams can travel along different routes and can arrive out of sequence or be duplicated. IP does not keep track of the routes and has no facility for reordering datagram once they arrive at their destination. IP provides bare-bones transmission functions that free the user to add only those facilities necessary for a given application and thereby allows for maximum efficiency.

IP address versions: IP became the official protocol for the internet in 1983.As the internet has evolved, so has the IP. There have been six versions since its inception.

Mainly there are two versions.

1. Version 4(IPv4)
2. Version 6(Ipv6)

1.5 INTERNET PROTOCOL VERSION 4 (IPV4)

Packets in the IPv4 format are called datagram. An IP datagram consists of a header part and a text part (payload). The header has a 20-byte fixed part and a variable length optional part. It is transmitted in big-endian order: from left to right, with the high-order bit of the Version field going first. IPv4 can be explained with the help of following points:

1. IP addresses
2. Address Space
3. Notations used to express IP address
4. Classfull Addressing
5. Subnetting
6. CIDR
7. NAT
8. IPv4 Header Format

1.5.1 IP addresses

Every host and router on the Internet has an IP address, which encodes its network number and host number.

The combination is unique: in principle, no two machines on the Internet have the same IP address.

An IPv4 address is 32 bits long

They are used in the Source address and Destination address fields of IP packets.

An IP address does not refer to a host but it refers to a network interface.

1.5.2 Address Space

An address space is the total number of addresses used by the protocol. If a protocol uses *N* bits to define an address, the address space is ***2N*** because each bit can have two different values (0 or 1) and *N* bits can have **2N** values.

IPv4 uses 32-bit addresses, which means that the address space is 232 or 4,294,967,296 (more than 4 billion).

1.5.3 Notations

There are two notations to show an IPv4 address:

1.5.3.1 Dotted decimal notation

To make the IPv4 address easier to read, Internet addresses are usually written in decimal form with a decimal point (dot) separating the bytes. Each byte (octet) is 8 bits hence each number in dotted-decimal notation is a value ranging from 0 to 255. Ex. 129.11.11.239

1.5.4 Classful addressing: In classful addressing, the address space is divided into five classes: A, B, C, D, and E.

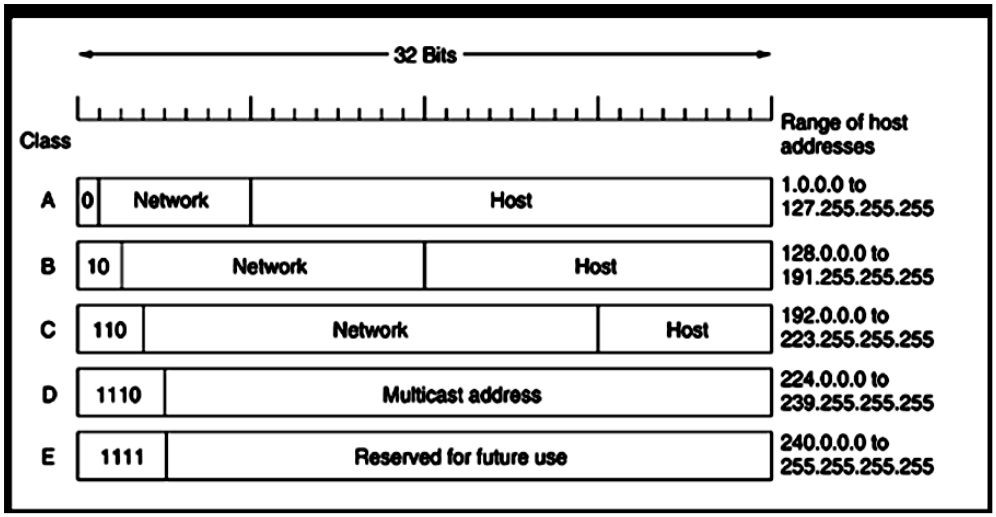


Fig: Classful addressing: IPv4 Netid and Hostid

In classful addressing, an IP address in class A, B, or C is divided into netid and hostid.

These parts are of varying lengths, depending on the class of the address as shown above.

Information on the Number of networks and host in each class is given in the table below:

Table: Number of networks and host in each class

Class	Number of Networks	Number of Hosts	Application
A	128	16,777,216	Unicast
B	16,384	65,536	Unicast
C	2,097,152	256	Unicast
D	1	268,435,456	Multicast
E	1	268,435,456	Reserved

The IP address 0.0.0.0 is used by hosts when they are being booted.

All addresses of the form 127.xx.yy.zz are reserved for loopback testing; they are processed locally and treated as incoming packets.

1.5.5 Sub-netting

It allows a network to be split into several parts for internal use but still act like a single network to the outside world.

To implement subnetting, the router needs a subnet mask that indicates the split between network + subnet number and host.

Ex. 255.255.252.0/22.A‖/22‖ to indicate that the subnet mask is 22 bits long.

Consider a class B address with 14 bits for the network number and 16 bits for the host number where some bits are taken away from the host number to create a subnet number.

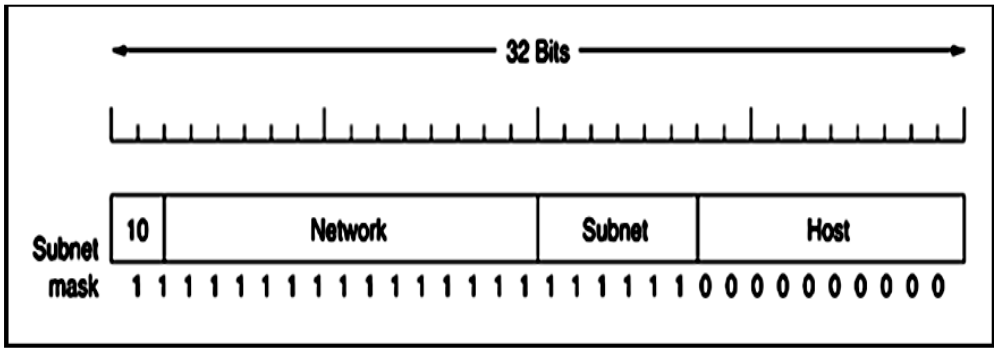


Fig: A Class B network sub netted into 64 subnets

If 6 bits from the host Id are taken for subnet then available bits are:

14 bits for network + 6 bits for subnet + 10 bits for host

With 6 bits for subnet the number of possible subnets is 26 which is 64.

With 10 bits for host the number of possible host are 210 which is 1022 (0 & 1 are not available)

1.5.6 CIDR A class B address is far too large for most organizations and a class C network, with 256 addresses is too small. This leads to granting Class B address to organizations that do not require all the address in the address space wasting most of it. This is resulting in depletion of Address space. A solution is CIDR (Classless Inter-Domain Routing) The basic idea behind CIDR, is to allocate the remaining IP addresses in variable-sized blocks, without regard to the classes.

1.5.7 NAT (Network Address Translation)

The scarcity of network addresses in IPv4 led to the development of IPv6.

IPv6 uses a 128 bit address; hence it has 2128 addresses in its address space which is larger than 232 addresses provided by IPv4.

Transition from IPv4 to IPv6 is slowly occurring, but will take years to complete, because of legacy hardware and its incompatibility to process IPv6 address.

NAT (Network Address Translation) was used to speed up the transition process

The only rule is that no packets containing these addresses may appear on the Internet itself. The three reserved ranges are:

10.0.0.0 – 10.255.255.255/8 (16,777,216 hosts) 172.16.0.0 – 172.31.255.255/12 (1,048,576 hosts) 192.168.0.0 – 192.168.255.255/16 (65,536 hosts)

Operation:

Within the Organization, every computer has a unique address of the form 10.x.y.z. However, when a packet leaves the organization, it passes through a NAT box that converts the internal IP source address, 10.x.y.z, to the organizations true IP address, 198.60.42.12 for example.

1.5.8 IP Header

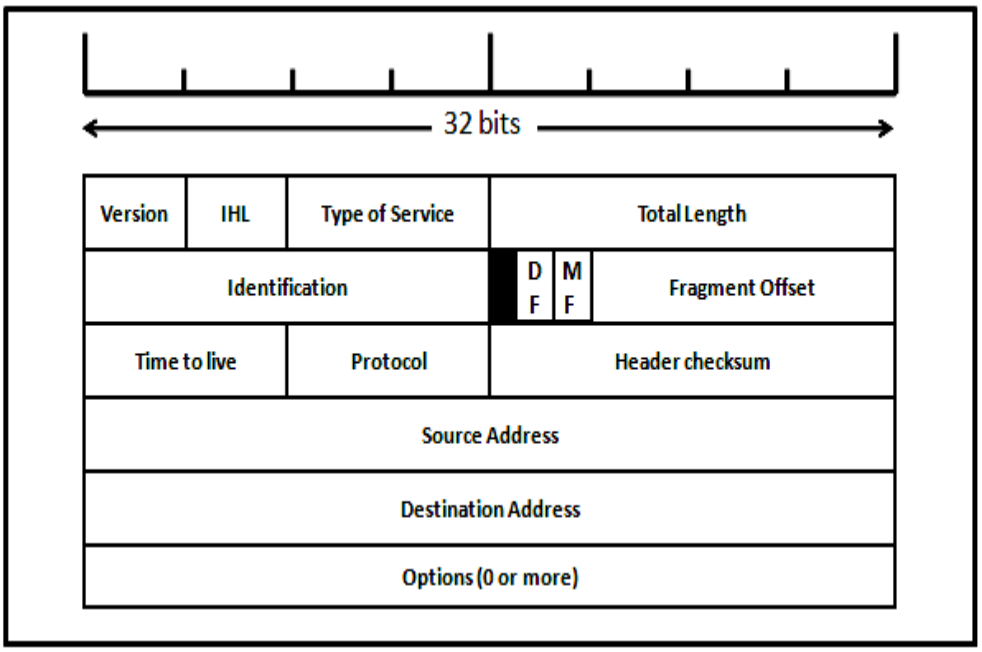


Fig: The IPv4 (Internet Protocol) header

The description of the fields of the IPv4 header is shown in the table given below:

Table: Description of each field of IPv4 PACKET header

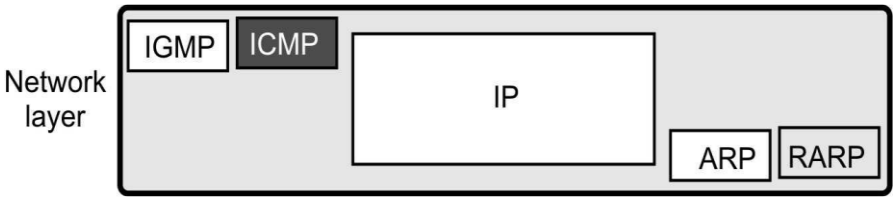
No	Field Name	Description
1	Version	Keeps track of the version of the protocol the datagram belongs to (IPV4 or IPv6)
2	IHL	Used to indicate the length of the Header. Minimum value is 5 Maximum value 15
3	Type of service	Used to distinguish between different classes of service
4	Total length	It includes everything in the datagram—both header and data. The maximum length is 65,535 bytes
5	Identification	Used to allow the destination host to identify which datagram a newly arrived fragment belongs to. All the fragments of a datagram contain the same Identification value
6	DF	1 bit field. It stands for Don't Fragment. Signals the routers not to fragment the datagram because the destination is incapable of putting the pieces back together again
7	MF	MF stands for More Fragments. All fragments except the last one have this bit set. It is needed to know when all fragments of a datagram have arrived.
8	Fragment offset	Used to determine the position of the fragment in the current datagram.
9	Time to live	It is a counter used to limit packet lifetimes. It must be decremented on each hop. When it hits zero, the packet is discarded and a warning packet is sent back to the source host.
10	Header checksum	It verifies Header for errors.

11	Source address	IP address of the source
12	Destination address	IP address of the destination
13	Options	The options are variable length. Originally, five options were defined: 1. Security : specifies how secret the datagram is 2. Strict source routing : Gives complete path to be followed 3. Loose source routing : Gives a list of routers not to be missed 4. Record route: Makes each router append its IP address 5. Timestamp: Makes each router append its IP address and timestamp

1.6 INTERNET CONTROL PROTOCOLS

Apart from Internet protocol (IP), other prevalent protocols in the internet layer include mainly:

- Address Resolution Protocol(ARP)
- Reverse Address Resolution Protocol(RARP)
- Internet Group Message Protocol(IGMP)
- Internet Control Message Protocol (ICMP)



1.6.1 Address Resolution Protocol

The Address Resolution Protocol (ARP) is used to associate a logical address with a physical address. On a typical physical network, such as a LAN, each device on a link is identified by a physical or station address, usually imprinted on the network interface card (NIC).

ARP is used to find the physical address of a node when its Internet address (IP address) is known.

If a machine talks to another machine in the same network, it requires its physical or MAC address. But, since the application has given the destination's IP address it requires some mechanism to bind the IP address with its MAC address. This is done through Address Resolution protocol (ARP).IP address of the destination node is broadcast and the destination node informs the source of its MAC address.

- Assume broadcast nature of LAN
- Broadcast IP address of the destination
- Destination replies it with its MAC address.
- Source maintains a cache of IP and MAC address bindings

But this means that every time machine A wants to send packets to machine B, A has to send an ARP packet to resolve the MAC address of B and hence this will increase the traffic load too much, so to reduce the communication cost computers that use ARP maintains a cache of recently acquired IP_to_MAC address bindings, i.e. they dont have to use ARP repeatedly. Several refinements of ARP are possible: When machine A wants to send packets to macine B, it is possible that machine B is going to send packets to machine A in the near future. So to avoid ARP for machine B, A should put its IP_to_MAC address binding in the special packet while requesting for the MAC address of B. Since A broadcasts its initial request for the MAC address of B, every machine on the network should extract and store in its cache the IP_to_MAC address binding of A When a new machine appears on the network (e.g. when an operating system reboots) it can broadcast its IP_to_MAC address binding so that all other machines can store it in their caches. This will eliminate a lot of ARP packets by all other machines, when they want to communicate with this new machine.

1.6.2 Reverse Address Resolution Protocol

The Reverse Address Resolution Protocol (RARP) allows a host to discover its Internet address when it knows only its physical address. It is used when a computer is connected to a network for the first time or when a diskless computer is booted.

RARP is a protocol by which a physical machine in a local area network can request to learn its IP address from a gateway server's Address Resolution Protocol table or cache. This is needed since the machine may not have permanently attached disk where it can store its IP address permanently. A network administrator creates a table in a local area network's gateway router that maps the physical machine (or Medium Access Control - MAC) addresses to corresponding Internet Protocol addresses. When a new machine is set up, its RARP client program requests from the RARP server on the router to be sent its IP address. Assuming that an entry has been set up in the router table, the RARP server will return the IP address to the machine which can store it for future use.

Drawbacks of RARP

- Since it operates at low level, it requires direct addresses to the network which makes it difficult for an application programmer to build a server.
- It doesn't fully utilizes the capability of a network like Ethernet which is enforced to send a minimum packet size since the reply from the server contains only one small piece of information, the 32-bit internet address.

1.6.3 Internet Control Message Protocol (ICMP)

The Internet Control Message Protocol (ICMP) is a mechanism used by hosts and gateways to send notification of datagram problems back to the sender. ICMP sends query and error reporting messages.

This protocol discusses a mechanism that gateways and hosts use to communicate control or error information. The Internet protocol provides unreliable, connectionless datagram service, and that a datagram travels from gateway to gateway until it reaches one that can deliver it directly to its final destination. If a gateway cannot route or deliver a datagram, or if the gateway detects an unusual condition, like network congestion, that affects its ability to forward the datagram, it needs to instruct the original source to take action to avoid or correct the problem. The Internet Control Message Protocol allows gateways to send error or control messages to other gateways or hosts; ICMP provides communication between the Internet Protocol software on one machine and the Internet Protocol

software on another. This is a special purpose message mechanism added by the designers to the TCP/IP protocols. This is to allow gateways in an internet to report errors or provide information about unexpected circumstances. The IP protocol itself contains nothing to help the sender test connectivity or learn about failures.

Error reporting vs. Error Correction

ICMP only reports error conditions to the original source; the source must relate errors to individual application programs and take action to correct problems. It provides a way for gateway to report the error It does not fully specify the action to be taken for each possible error. ICMP is restricted to communicate with the original source but not intermediate sources.

ICMP Message Delivery

ICMP messages travel across the internet in the data portion of an IP datagram, which itself travels across the internet in the data portion of an IP datagram, which itself travels across each physical network in the data portion of a frame. Datagram carrying ICMP messages are routed exactly like datagram carrying information for users; there is no additional reliability or priority. An exception is made to the error handling procedures if an IP datagram carrying an ICMP messages are not generated for errors that result from datagram carrying ICMP error messages.

ICMP Message Format

It has three fields; an 8-bit integer message TYPE field that identifies the message, an 8-bit CODE field that provides further information about the message type, and a 16-bit CHECKSUM field (ICMP uses the same additive checksum algorithm as IP, but the ICMP checksum only covers the ICMP message).In addition, ICMP messages that report errors always include the header and first 64 data bits of the datagram causing the problem. The ICMP TYPE field defines the meaning of the message as well as its format.

The Types include:

TYPE FIELD	ICMP MESSAGE TYPE
0	ECHO REPLY
3	DESTINATION UNREACHABLE
4	SOURCE QUENCH

5	REDIRECT(CHANGE A ROUTE)
8	ECHO REQUEST
11	TIME EXCEEDED FOR A DATAGRAM
12	PARAMETER PROBLEM ON A DATAGRAM
13	IMESTAMP REQUEST
14	TIMESTAMP REPLY
15	INFORMATION REQUEST(OBSOLETE)
16	INFORMATION REPLY(OBSOLETE)
17	ADDRESS MASK REQUEST
18	ADDRESS MASK REPLY TESTING DESTINATION

1.6.4 Internet Group Message Protocol (IGMP)

The Internet Group Message Protocol (IGMP) is used to facilitate the simultaneous transmission of a message to a group of recipients.

It is a communications protocol used by hosts and adjacent routers on IPv4 networks to establish multicast group memberships. IGMP is an integral part of IP multicast.

IGMP can be used for one-to-many networking applications such as online streaming video and gaming, and allows more efficient use of resources when supporting these types of applications.

IGMP is used on IPv4 networks. Multicast management on IPv6 networks is handled by Multicast Listener Discovery (MLD) which uses ICMPv6 messaging in contrast to IGMP's bare IP encapsulation.

Types of messages in IGMP

IGMP has three types of messages: the query, the membership report, and the leave report. There are two types of query messages as shown in the figure below: general and special.

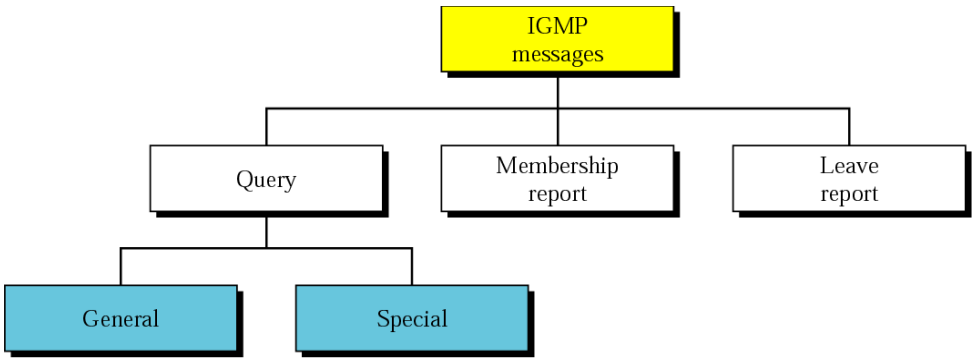


Fig: Types of messages in IGMP

1.7 KEY WORDS & CONCEPTS

- **An internet is a c**ollection of communications networks interconnected by bridges and/or routers/gateways.
- **Internet Protocol (IP)** is TCP / IP's primary network protocol, which provides addressing and routing information.
- TCP/IP has 4 layers: host to network, IP Layer, Transport & Application
- It uses 4 levels of address: physical, logical, port & specific
- IP address uniquely identifies a device on the Internet.
- IT address is a 32 bit global internet address consisting two parts Network ID and host ID.
- Some of the Internet**control protocols are:** ICMP, ARP, RARP, BOOTP
- **Address Resolution Protocol (ARP)** – A protocol in TCP/IP that accepts the IP address and returns the corresponding physical address of a computer or router.
- **Internet Control Message Protocol** (ICMP) provides error-reporting mechanisms, transfer control messages from routers and hosts, to other hosts, ie Feedback about problems.
- **Internet Group Management Protocol (IGMP**) is the protocol used to support multicasting.

1.8 SELF ASSESMENT QUESTIONS

1. What are the functions of network layer?

……

……

……

……

2. Write about IP address and its classifications?

.....

.....

.....

.....

3. What is ICMP? What is its purpose?

……

……

……

……..

4. What is subnet mask? Write with an example?

……

……

……

……

5. What is InternetProtocol? What are its services?

……

……

……

……....

1.9 REFERENCES AND SUGGESTEDREADING

1. Behrouz A. Forouzan, "*Introduction to Data Communications and Networking*", McGraw-Hill Education (India), New Delhi.
2. Andrew S. Tanenbaum, "*Computer Networks*", PHI Learning Pvt. Ltd
3. James F. Kurose, Keith W. Ross, "*Computer Networking: A Top-Down Approach Featuring the Internet*", Pearson Education Inc., New Delhi.
4. Wayne Tomasi, "*Introduction to Data Communications and Networking*", Pearson Education Inc., New Delhi.
5. L. L. Peterson & B. S. Davie," *Computer Networks*", Elsevier Inc,

UNIT-2
TRANSPORT LAYER PROTOCOLS

UNIT STRUCTURE

2.0 INTRODUCTION

The transport layer deals with two main protocols in TCP/IP protocol suite. One is the Transmission Control Protocol (TCP) and another is User Datagram Protocol (UDP). TCP is a connection oriented protocol that provides reliable end-to-end delivery of data. Whereas, UDP is a connection-less protocol. It provides well sequenced transport function when reliability and serving are less important than size and speed. Transport layer services are implemented by transport protocols used between two transport entities. Transport layer services are similar to the data link services. But Data link layer is designed to provide its services within a single network, while the transport layer provides services across an internetwork made up of many networks.

In this unit we will discuss in detail about these two transport protocols TCP and UDP with their services as well.

2.1 LEARNING OBJECTIVE

After going through this unit, you should be able to:

- Know the Functions and Services of transport layer
- The functions of transport protocols TCP and UDP
- Understand the difference in approaches of both transport layer protocols
- Understand the header of both TCP and UDP segments.

2.2 TRANSPORT SERVICES

Protocols at the transport layer responsible for the delivery of data from an application program on one device to an application program on another device. It is the first end-to-end layer in the OSI model. It provides its services to the upper layer to use the services of the network layer and other lower layer protocols.

There are seven categories of services provided by the transport layer. The main transport services are End-to-end-delivery, Addressing, Reliable delivery, Flow control, Connection management, Multiplexing, Congestion Control and Quality of Services (QoS).

2.2.1 End-to-end-delivery

A logical address at network layer facilitates the transmission of data from source to destination device. But the source and the destination both may be having multiple processes communicating with each other. Hence it is important to deliver the data not only from the sender to the receiver but from the correct process on the sender to the correct process on the receiver. The transport layer takes care of process to process delivery of

data and makes sure that it is intact and in order. This is also called end-to-end delivery.

2.2.2Addressing

To ensure process to process delivery the transport layer makes use of **port address** to identify the data from the sending and receiving process. A Port Address is the name or label given to a process. It is a 16 bit address. Ex. TELNET uses port address 23, HTTP uses port address 80. Port address is also called as Service Point Address.

2.2.3 Segmentation and Reassembly

The data can be transported in a connection oriented or connectionless manner. If the connection is connection oriented then all segments are received in order else they are independent of each other and are received out of order and have to be rearranged.

The Transport layer is responsible for segmentation and reassembly of the message into segments which bear sequence numbers. This numbering enables the receiving transport layer to rearrange the segments in proper order.

2.2.4 Flow Control & Error control: The transport layer also carries out flow control and error control functions; but unlike data link layer these are end to end rather than node to node.

2.2.5Connection Management

End to end delivery can be accomplished in two ways: connection oriented and connectionless. A Connection-OrientedInternet service is accomplished using TCP (Transmission Control Protocol). Reliable transport, flow control and congestion control are types of services that are provided to an application by TCP. These services are acknowledged. The TCP's congestion control mechanism is used to maintain throughput.

The connection oriented mode is most commonly used from both two modes. A connection oriented protocol establishes a virtual circuited or pathway through the internal between sender and receiver. All of the packets belonging to a message are then sent over this same path. Using a single pathway for the entire message facilitates the acknowledgement process and retransmission of damaged and lost frames connection oriented services is generally considered reliable.

The reliable services includes features such as flow control, error control and sequence delivery. Flow control makes sure that neither side of a connection overwhelms the other side by sending too many pockets too quickly.

Connection Oriented transmission has three stages:

1. Connection establishment.
2. Data transfer
3. Connection termination.

Connection Establishment:

Before communicating device can send data to the other, the initializing device must first determine the availability of the other to exchange data and a pathway must be found through the network by which the data can be sent.

This step is called connection establishment. Connection establishment requires three Transport Layer actions called three way handshakes as shown in figure given below.

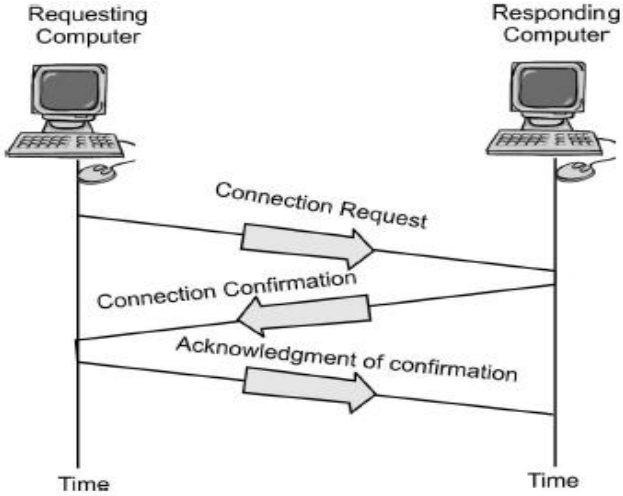


Fig: Connection Establishment

The computer requesting the connection sends a connection request packet to the intended receiver.

• The responding computer returns a confirmation packet to the requesting computer.

• The requesting computing returns a packet acknowledging the confirmation

Connection Termination:

Once all of the data have been transferred, the connection must be terminated.

Connection termination also requires three way handshakes as shown in figure below.

- Requesting computer sends a disconnection request packet.
- Responding computers confirm the disconnection request.
- The requesting computer acknowledges the confirmation.

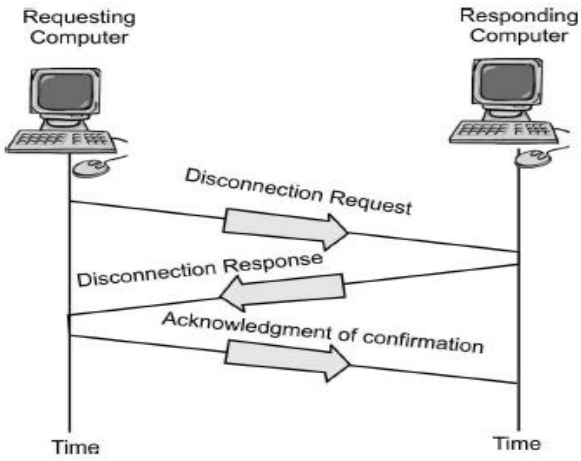


Fig: Connection Termination

2.2.6 Multiplexing

To improve transmission efficiency, the transport layer has the option of multiplexing. Multiplexing at this layer occurs in two ways:

1. **Upward Multiplexing** means that multiple transport layer connections use the same network connection.

In Upward Multiplexing, Transport layer uses virtual circuits based on the services of the lower three layers. The underlying network charge for each virtual circuit connection. To make well cost-effective use of an established circuit, the transport layer sends several transmissions based for the same destination along the same path by upward multiplexing.

2. **Downward Multiplexing** means that one transport layer connection uses multiple network connections.

Downward Multiplexing allows the transport layer to split a single connection among several different paths to improve throughput (speed of delivery). This option is useful when the underlying networks have low or slow capacity. For example, some network layer protocols have restriction on the sequence number that can be handled .X.25 uses a three bit numbering code, so sequence number are restricted to the range of 0 to 7. In this case throughput can be unacceptably low. To counteract this problem the transport layer can use more than one virtual circuit at the network layer to improve throughput by sending several data segment at once delivery is faster.

2.2.7 Congestion Control

Congestion control refers to techniques and mechanisms that can either prevent congestion, before it happens, or remove congestion, after it has happened.

2.2.8 Quality of Services (QoS)

A stream of packet, from a source to a destination is called a flow. In connection oriented network, all the packets belonging to a flow the same

route. But in connection less network they may follow different routes. The need of each flow can be characterized by four primary parameters:

1. **Reliability**: Reliability is the ability of a system to perform and maintain its functions in normal conditions as well as under unexpected conditions.
2. **Delay** is defined as the time interval elapsed between the departures of data from the source to its arrival at the destination.
3. **Jitter:** Jitter refers to the variation in time between packets arriving at the destination.
4. **Bandwidth** refers to the data rate supported by a network connection or interface.

2.3 INTRODUCTION TO TCP

TCP is designed to provide reliable communication between pairs of processes (TCP users) across a variety of reliable and unreliable networks and Internets. TCP is streaming oriented (Connection Oriented). Stream means that every connection is treated as stream of bytes. The user application does not need to package data in individual datagram as it is done in UDP. In TCP a connection must be established between the sender and the receiver. By creating this connection, TCP requires a VC at IP layers that will be active for the entire duration of a transmission.
Both IP and UDP treat multiple datagram's belongings to a single transmission as entirely separate units, un-related to each other. TCP on the other hand is responsible for the reliable delivery of entire segments. Every segment must be received and acknowledged before the VC is terminated.

2.4 TCP SEGMENT HEADER

In this section, we will discuss about TCP Segment Header, different fields in TCP Header and the use of these fields.

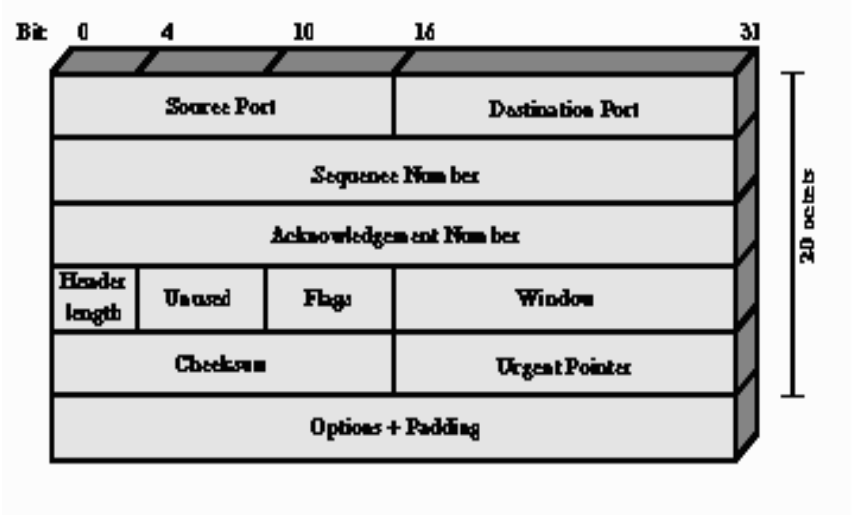


Fig: TCP header format

The functions of fields shown in the above figure are as follows:
Source port: 16 Bit number which identifies the Source Port number (Sending Computer's TCP Port).

Destination port: 16 Bit number which identifies the Destination Port number (Receiving Port).

Sequence number: 32 Bit number used for byte level numbering of TCP segments. If you are using TCP, each byte of data is assigned a sequence number. If SYN flag is set (during the initial three way handshake connection initiation), then this is the initial sequence number. The sequence number of the actual first data byte will then be this sequence number plus 1.

Acknowledgment Number: 32 Bit number field which indicates the next sequence number that the sending device is expecting from the other device.

Header Length: 4 Bit field which shows the number of 32 Bit words in the header. Also known as the Data Offset field. The minimum size header is 5 words (binary pattern is 0101).

Reserved: Reserved for future use. Always set to 0 (Size 6 bits).

Control Bit Flags: We have seen before that TCP is a Connection Oriented Protocol. The meaning of Connection Oriented Protocol is that, before any data can be transmitted, a reliable connection must be obtained and acknowledged. Control Bits govern the entire process of connection establishment, data transmissions and connection termination. The control bits are listed as follows: They are:

URG: Urgent Pointer.

ACK: Acknowledgement.

PSH: This flag means Push function. Using this flag, TCP allows a sending application to specify that the data must be pushed immediately. When an application requests the TCP to push data, the TCP should send the data that has accumulated without waiting to fill the segment.

RST: Reset the connection. The RST bit is used to RESET the TCP connection due to unrecoverable errors. When an RST is received in a TCP segment, the receiver must respond by immediately terminating the connection. A RESET causes both sides immediately to release the connection and all its resources. As a result, transfer of data ceases in both directions, which can result in loss of data that is in transit. A TCP RST indicates an abnormal termination of the connection.

SYN: This flag means synchronize sequence numbers. Source is beginning a new counting sequence. In other words, the TCP segment contains the sequence number of the first sent byte (ISN).

FIN: No more data from the sender. Receiving a TCP segment with the FIN flag does not mean that transferring data in the opposite direction is not possible. Because TCP is a fully duplex connection, the FIN flag will cause the closing of connection only in one direction. To close a TCP connection gracefully, applications use the FIN flag.

Window: indicates the size of the receive window, which specifies the number of bytes beyond the sequence number in the acknowledgment field that the receiver is currently willing to receive.

Checksum: The 16-bit checksum field is used for error-checking of the header and data.

Urgent Pointer: Shows the end of the urgent data so that interrupted data streams can continue. When the URG bit is set, the data is given priority over other data streams (Size 16 bits).

Options (Variable): At present, only one option is defined, which specifies the maximum segment size that will be accepted.

The source port and destination port specify the sending and receiving users of TCP. As with IP, there are a number of common users of TCP that have been assigned numbers; these numbers should be reserved for that purpose in any implementation. Other port numbers must be arranged by agreement between communicating parties.

2.5 TCP SERVICES

When an application invokes TCP for its transport protocol; the application receives the following services from it. Connection oriented service, Reliable transport service, Congestion-control mechanism.

Now let us describe these services in brief:

Connection-oriented service: As you know that the connection oriented service is comprised of the handshake procedure which is a full duplex connection in that, two processes can send messages to each other over the connection at the same time. When the application has finished sending the message, it must remove the connection. The service is referred to as a "connection oriented" service. We are also aware from the discussion on the network layer that this service is implemented through the virtual circuit mechanism.

Reliable transport service: Communicating processes can rely on TCP to deliver all the data that is sent, without error and in the proper order. When one side of the application passes a stream of bytes into a socket, it can count on TCP to deliver the same stream of data to the receiving socket, with no missing or duplicate bytes. Reliability in the Internet is achieved with the use of acknowledgement and retransmissions.

Congestion-control mechanism: TCP also includes a congestion-control mechanism for the smooth functioning of Internet processes. When the packet load offered to the network exceeds its handling capacity congestion builds up. One point to be noted is that, although, the Internet connection oriented service is bundled with suitable data transfer, flow control and congestion control mechanisms; they are not the essential components of the connection oriented service. A connection oriented service can be provided with bundling these services through a different type of a network.

Services that is not provided by TCP

Some of the services that are not provided by TCP are as follows:

i) It does not guarantee a minimum transmission rate,

ii) It does not provide any delay guarantee. But it guarantees delivery of all data.

However, it provides no guarantee to the rate of data delivery.

2.6 INTRODUCTION TO UDP

UDP is connectionless and unreliable transport protocol. Apart from multiplexing /de-multiplexing and some error correction UDP adds little to the IP protocol. Here, we will discuss about the basic functions of UDP and how it works.

In case, we choose UDP instead of TCP for application development, then the application is almost directly talking to the IP layer. UDP takes messages from the application process, attaches the source and destination port number fields for the multiplexing/de-multiplexing service, adds two other small fields, and passes the resulting segment to the network layer. Without establishing handshaking between sending and receiving transport-layer entities, the network layer encapsulates the segment into an IP datagram and then makes a **best-effort** attempt to deliver the segment to the receiving host. If the segment arrives at the receiving host, UDP uses the destination port number to deliver the segment's data to the desired application process. So you might ask a question then why is UDP required? TCP should be the choice for all types of application layer/protocol. DNS stands for domain name system. It provides a directory service for the internet. It is commonly used by other application protocols (HTTP, FTP etc.) to translate user given host names to IP address. But before we answer your question let-we look at another application called the DNS, which runs exclusively in UDP only but unlike other protocols DNS is not an application with which users interact directly. Instead DNS is a core Internet function that translates the host name to IP addresses. Also unlike other protocols, DNS typically uses UDP. When the DNS application in a host wants to make a query, it constructs a DNS query message and passes the message to the UDP. Without performing any handshaking with the UDP entity running on the destination end system, UDP adds header fields to the message and passes the resulting segment to the network layer. The network layer encapsulates the UDP segment into a datagram and sends the datagram to a name server. The DNS application at the querying host then waits for a reply to its query. If it doesn't receive a reply (possibly because the underlying network lost the query or the reply), either it tries sending the query to another name server, or it informs the invoking application that it can't get a reply. Like DNS, there are many applications, which are better suited for UDP for the following reasons.

No connection establishment: Since UDP does not cause any delay in establishing a connection, this is a probably the principal reason why DNS runs TCP. HTTP uses TCP rather than UDP, since reliability is critical for Web pages with text.

More Client support: TCP maintains connection state in the end systems. This connection state includes receiving and sending buffers, congestion-control parameters, and sequence and acknowledgement number parameters. UDP, on the other hand, does not maintain connection state and does not track any of these parameters. A server devoted to a particular application can typically support many more active clients when the application runs over UDP rather than TCP. Because UDP, (unlike TCP) does not provide reliable data service and congestion control mechanism, therefore, it does not need to maintain and track the receiving and sending of buffers (connection states), congestion control parameters, and sequence and acknowledgement number parameters. Small packet header overhead. The TCP segment has 20 bytes of header over-head in every segment, whereas UDP has only eight bytes of overhead. Now let us examine the application services that are currently using the UDP protocol. For example, remote file server, streaming media, internet telephony, network management, routing protocol such as RIP and, of course, DNS use UDP. Other applications like e-mail, remote terminal access web surfing and file transfer use TCP.

2.7 UDP SEGMENT STRUCTURE

UDP is an end to end transport level protocol that adds only port addresses, checksum error control and length information to the data from the upper layer. The application data occupies the data field of the UDP segment. For example, for DNS, the data field contains either a query message or a response message. For a streaming audio application, audio samples fill the data field. The packet produced by UDP is called a user datagram.

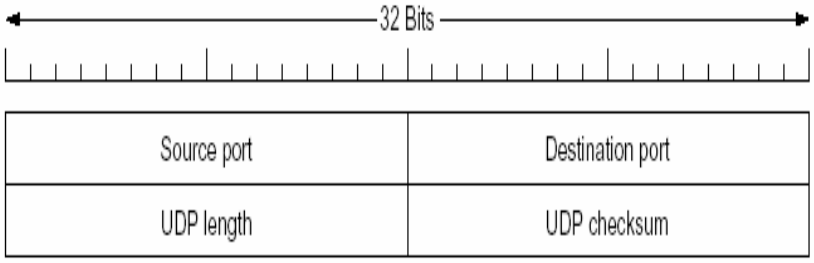


Fig: UDP segment structure

The UDP header has only four fields, each consisting of two bytes as in the figure above. Let us discuss each field separately:

Source port address: It is the address of the application program that has created the message.

Destination port address: It is the address of the application program that will receive the message.

Total length: It specifies the length of UDP segment including the header in bytes.

Checksum: The checksum is used by the receiving host to check whether errors have been introduced into the segment. In truth, the checksum is also calculated over a few of the fields in the IP header in addition to the UDP segment.

2.8UDP SERVICES

Some of the important features of UDP are as follows:

i) UDP is connectionless, so there is no hand shaking before the two processes start communications.

ii) It provides unreliable data transfer service therefore, there is no guarantee that the message will reach. Due to this, the message may arrive as the receiving process at a random time.

UDP does not provide a congestion-control service. Therefore, the sending process can pump data into a UDP socket at any rate it pleases. Then why do we require such a protocol at the transport layer? There are certain types of applications such as real time. Real-time applications are applications that can tolerate some loss but require a minimum rate. Developers of real-time applications often choose to run their applications over the UDP because their applications cannot wait for acknowledgements for data input. Now coming back to TCP, since it guarantees that all the packets will be delivered to the host, many application protocols, layer protocols are used for these services. For example, SMTP (E-mail), Telnet, HTTP, FTP exclusively uses TCP whereas NFS and Streaming Multimedia may use TCP or UDP. But Internet telephony uses UDP exclusively.

2.9 SELFASSESMENT QUESTIONS

1. What is the need for transport layer?

……
……
……
……..

2. Write down all the transport layer services.

……
……
……
……

3. What are the advantages of UDP?

……
……
……
……

4. How do transport services differ from the data link layer services?

……
……
……
……

5. Which is reliable - TCP or UDP? Why?

……
……
……
……

6. What is multiplexing? Differentiate between upward and downward multiplexing?

……
……
……
…….....

7. Define Congestion Control?

……
……
……
……....

8. What is meant by quality of service?

……
……
……
……

9. What are the three events involved in the connection?

……
……
……

10. Differentiate between TCP and UDP.

……
……
……
…….....
.....

2.10 KEYWORDS AND CONCEPTS

1. The transport layer takes care of process to process delivery of data and makes sure that it is intact and in order. This is also called **end-to-end delivery.**
2. To ensure process to process delivery the transport layer makes use of **port address** to identify the data from the sending and receiving process.
3. A **connection oriented service** has three steps for transmission:
    - Connection establishment.
    - Data transfer
    - Connection termination.
4. The **Quality of Service** (QoS) of a network is characterized by four primary parameters: Reliability, Delay Jitter, and Bandwidth**.**
5. UDP is connectionless protocol and it provides unreliable data transfer service therefore, there is no guarantee that the message will reach. Due to this, the message may arrive as the receiving process at a random time.
6. TCP is a reliable transport protocol, because, it provides connection oriented, reliable transport services between sending and receiving processes. It also provides congestion control mechanism for the smooth functioning of Internet processes.

2.11 REFERENCES AND SUGGESTED READINGS

1. Behrouz A. Forouzan, "*Introduction to Data Communications and Networking*", McGraw-Hill Education (India), New Delhi.
2. Andrew S. Tanenbaum, "*Computer Networks*", PHI Learning Pvt. Ltd
3. James F. Kurose, Keith W. Ross, "*Computer Networking: A Top-Down Approach Featuring the Internet*", Pearson Education Inc., New Delhi.
4. Wayne Tomasi, "*Introduction to Data Communications and Networking*", Pearson Education Inc., New Delhi.
5. Indira Gandhi National Open University, Study Material, *Data Communication and Computer Networks*, Block-4, unit-1 Transport Services and Mechanism Concepts of Communication and Networking.

UNIT-3
APPLICATION LAYER PROTOCOLS

UNIT STRUCTURE

3.0 INTRODUCTION

An application layer protocol defines how an application processes (clients and servers), running on different end systems, pass messages to each other.

In particular, an application layer protocol defines:

- The types of messages, e.g., request messages and response messages.
- The syntax of the various message types, i.e., the fields in the message and how the fields are delineated.
- The semantics of the fields, i.e., the meaning of the information that the field is supposed to contain;
- Rules for determining when and how a process sends messages and responds to messages.

3.1 LEARNING OBJECTIVE

After going through this unit you will be able to

- Understand the concept of client server model
- Know the working of HTTP protocol
- Know the working to e-mail with SMTP.
- Know the functions of POP
- Explain the working of TELNET and FTP

3.2 CLIENT SERVER MODEL

In order to avail the services provided by the internet two or more computers communicate with each other through application programs. The application programs are the entities that communication with each other, not the computers or users .The application programs using the internet follow these client server model strategies.

The client-server model is a distributed communication framework of network processes among service requestors, clients and service providers. The client-server connection is established through a network or the Internet.

The client-server model is a core network computing concept also building functionality for email exchange and Web/database access. Web technologies and protocols built around the client-server model are:

- Hypertext Transfer Protocol (HTTP)
- Domain Name System (DNS)
- Simple Mail Transfer Protocol (SMTP)
- TELNET

Clients include Web browsers, chat applications, and email software, among others. Servers include Web, database, application, chat and email, etc.

Clients: A clients is a program running on a local machine requesting service from a server. A client program is finite which means a client program is run by the user or any other application when it is needed and terminates when the service is complete. Services needed frequently and by many users have specific client server application programs.

Server: A server is an application program running on the remote machine providing services to the client. A server program is an infinite program. When it starts, it runs infinitely unless a problem arises. It waits for incoming requests from client .When a request comes it responds and provides services to the request.

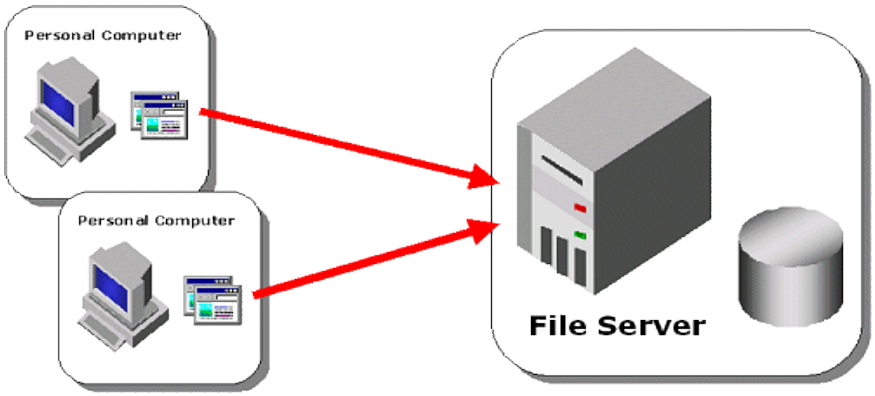


Fig: A client – server Model

A client – server relationship is many – to –one.Many clients can use the services of one server.

Advantages of Client Server Model:

(i) Client server can be used in distributed system
(ii) The likely hood of system break down is less.
(iii) Resources of other computer can be trapped.
(iv) Distance is no barrier for availing information

3.3 BOOTSTRAP PROTOCOL (BOOTP)

Each computer that is attached to a TCP/IP internet must know the following information.

- Its IP address
- Its subnet mask
- The IP address of a router
- The IP address of a name server.

This information is usually stored in a configuration file and accessed by the computer during the boot strap process.

Bootstrap protocol(**BOOTP)** is a client – server protocol designed to provide above four information for a disk text computer that is booted the

first time RARP used to provide only IP address of a diskless computer but a BOOTP all the four information and we do not need RARP.

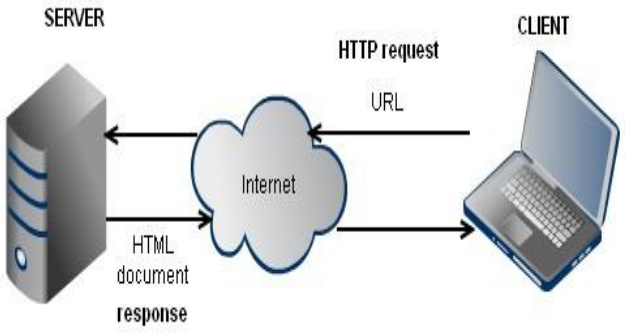
The primary motivation for replacing RARP with BOOTP is that RARP was a link layer protocol. This made implementation difficult on many server platforms, and required that a server be present on each individual IP subnet. BOOTP introduced the innovation of relay agents, which forwarded BOOTP packets from the local network using standard IP routing, so that one central BOOTP server could serve hosts on many subnets.

3.4 HYPERTEXT TRANSFER PROTOCOL (HTTP)

The hypertext transfer protocol (HTTP) is a protocol used mainly to access data on the World Wide Web (www). The protocol is used to transfer plain text, hyper text, audio, video and soon. The Hypertext Transfer Protocol (HTTP) the Web's main application-layer protocol although current browsers can access other types of servers.
A repository of information spread all over the world and linked together. HTTP allows the user to switch from one hypertext to another or one document to another. It allows transferring file and used the services of TCP. HTTP utilizes TCP connections to send client requests and server replies.

3.4.1 HTTP transaction:

The client initializes the transaction by sending the request message. The server replies by sending a response. There are two types of HTTP messages; request message and response message.



A Request message consists of a request line, header and the body.

A **response message consists** of a status line, header and sometimes body.

HTTP functions as a request–response protocol in the client–server computing model. A web browser, for example, may be the *client* and an application running on a computer hosting a web site may

be the *server*. The client submits an HTTP *request* message to the server. The server, which provides *resources* such as HTML files and other content, or performs other functions on behalf of the client, returns a *response* message to the client. The response contains completion status information about the request and may also contain requested content in its message body.

A web browser is an example of a *user agent* (UA). Other types of user agent include the indexing software used by search providers (web crawlers), voice browsers, mobile apps, and other software that accesses, consumes, or displays web content.

HTTP is designed to permit intermediate network elements to improve or enable communications between clients and servers. High-traffic websites often benefit from web cache servers that deliver content on behalf of upstream servers to improve response time. Web browsers cache previously accessed web resources and reuses them when possible to reduce network traffic. HTTP proxy servers at private network boundaries can facilitate communication for clients without a globally routable address, by relaying messages with external servers.

HTTP is an application layer protocol designed within the framework of the Internet Protocol Suite. Its definition presumes an underlying and reliable transport layer protocol,[2]and Transmission Control Protocol (TCP) is commonly used. However HTTP can be adapted to use unreliable protocols such as the User Datagram Protocol (UDP), for example in HTTPU and Simple Service Discovery Protocol (SSDP).

HTTP resources are identified and located on the network by uniform resource locators (URLs), using the uniform resource identifier (URI) schemes *http* and *https*. URIs and hyperlinks in Hypertext Mark-up Language (HTML) documents form inter-linked hypertext documents.

3.4.2 HTTPS

HTTPs is also called **HTTP Secure,** as it is used asa protocol for secure communication over a computer network which is widely used on the Internet. HTTPS consists of communication over Hypertext Transfer Protocol (HTTP) within a connection encrypted by Transport Layer Security or its predecessor, Secure Sockets Layer. The main motivation for HTTPS is authentication of the visited website and protection of the privacy and integrity of the exchanged data.

In its popular deployment on the internet, HTTPS provides authentication of the website and associated web server with which one is communicating, which protects against man-in-the-middle attacks. Additionally, it provides bidirectional encryption of communications between a client and server, which protects against eavesdropping and tampering with or forging the

contents of the communication.[6] In practice, this provides a reasonable guarantee that one is communicating with precisely the website that one intended to communicate with (as opposed to an impostor), as well as ensuring that the contents of communications between the user and site cannot be read or forged by any third party.

Historically, HTTPS connections were primarily used for payment transactions on the World Wide Web, e-mail and for sensitive transactions in corporate information systems. In the late 2000s and early 2010s, HTTPS began to see widespread use for protecting page authenticity on all types of websites, securing accounts and keeping user communications, identity and web browsing private.

3.5 SMTP (SIMPLE MAIL TRANSFER PROTOCOL)

One of the most popular network services on the Internet is the electronic mail (e-mail).The TCP/IP protocol that supports electronic mail on the Internet is called Simple Mail Transfer Protocol (SMTP).SMTP transfers messages from senders' mail servers to the recipients' mail servers using TCP connections. Users are identified based on e-mail addresses.

SMTP provides services for mail exchange between users on the same or different computers.

SMTP works based on the client/server model as follows:

SMTP has two sides: a client side which executes on a sender's mail server, and server side which executes on recipient's mail server. Both the client and server sides of SMTP run on every mail server. When a mail server sends mail (to other mail servers), it acts as an SMTP client.

When a mail server receives mail (from other mail servers) it acts as an SMTP server.

3.5.1 How E-mail works?

Email is based around the use of electronic mailboxes. When an email is sent, the message is routed from server to server, all the way to the recipient's email server. More precisely, the message is sent to the mail server tasked with transporting emails (called the **MTA**, for *Mail Transport Agent*) to the recipient's MTA. On the Internet, MTAs communicate with one another using the protocol SMTP, and so are logically called **SMTP servers** (or sometimes *outgoing mail servers*).

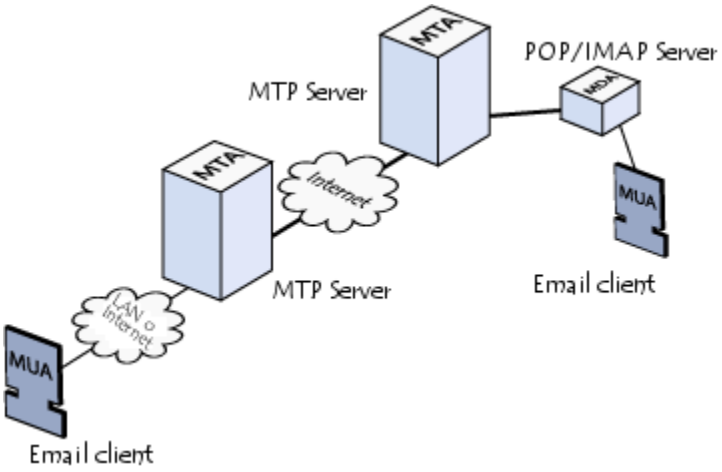


Fig: Working of E-Mail

The recipient's MTA then delivers the email to the incoming mail server (called the **MDA**, for *Mail Delivery Agent*), which stores the email as it waits for the user to accept it. There are two main protocols used for retrieving email on an MDA:

- POP3 (*Post Office Protocol*), the older of the two, which is used for retrieving email and, in certain cases, leaving a copy of it on the server.
- IMAP (*Internet Message Access Protocol*), which is used for coordinating the status of emails (read, deleted, moved) across multiple email clients. With IMAP, a copy of every message is saved on the server, so that this synchronization task can be completed.

For this reason, incoming mail servers are called **POP servers** or **IMAP servers**, depending on which protocol is used.

To use a real-world analogy, MTAs act as the post office (the sorting area and mail carrier, which handle message transportation), while MDAs act as mailboxes, which store messages (as much as their volume will allow) until the recipients check the box. This means that it is not necessary for recipients to be connected in order for them to be sent email.

To keep everyone from checking other users' emails, MDA is protected by a user name called a**login** and by a **password**.

Retrieving mail is done using a software program called an **MUA** (*Mail User Agent*).

When the MUA is a program installed on the user's system, it is called an **email client** (such as Mozilla Thunderbird, Microsoft Outlook, Eudora Mail, Incredimail or Lotus Notes).

When it is a web interface used for interacting with the incoming mail server, it is called **webmail**.

3.6 MULTIPURPOSE INTERNET MAIL EXTENSIONS (MIME)

Multipurpose Internet Mail Extensions (MIME) is an Internet standard that extends the format of email to support: Text in character sets other than ASCII. Non-text attachments: audio, video, images, application programs etc. Message bodies with multiple parts.

It is an extension of SMTP that allows the transfer of multimedia messages. If binary data is included in a message MIME headers are used to inform the receiving mail agent:

- Content-Transfer-Encoding: Header alerts the receiving user agent that the message body has been ASCII encoded and the type of encoding used.
- Content-Type: Header informs the receiving mail agent about the type of data included in the message.

3.7 POST OFFICE PROTOCOL (POP)

The **Post Office Protocol** (**POP**) is an application-layer Internet standard **protocol** used by local e-**mail** clients to retrieve e-**mail** from a remote server over a TCP/IP connection.POP is also called as POP3 protocol.

This is a protocol used by a mail server in conjunction with SMTP to receive and holds mail for hosts.

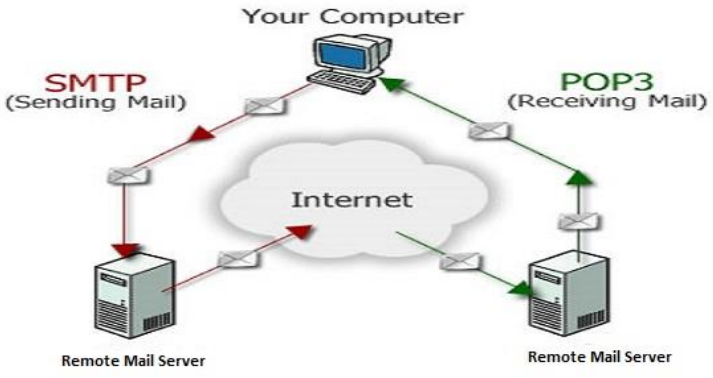


Fig: Function of POP

POP3 mail server receives e-mails and filters them into the appropriate user folders. When a user connects to the mail server to retrieve his mail, the messages are downloaded from mail server to the user's hard disk.

3.8 DOMAIN NAME SYSTEM (DNS)

To identify an entity, TCP/IP protocol uses the IP address which uniquely identifies the connection of a host to the Internet. The Domain Name System (DNS) is the system used to translate alphanumeric domain names into Internet Protocol numbers.

The DNS is made up of many servers and databases which, through a series of lookups in various caches, configure Domain Names into IP Addresses. DNS is a hierarchical system, based on a distributed database that uses a hierarchy of Name Servers to resolve Internet host names into the corresponding IP addresses required for packet routing by issuing a DNS query to a name server.Its purpose is to provide a layer of abstraction between Internet services (web, email, etc.) and the numeric addresses (IP

addresses) used to uniquely identify any given machine on the Internet. The DNS associates a variety of information with the domain names assigned and, most importantly, translates the domain names meaningful to humans into the numerical identifiers that locate the desired destination.

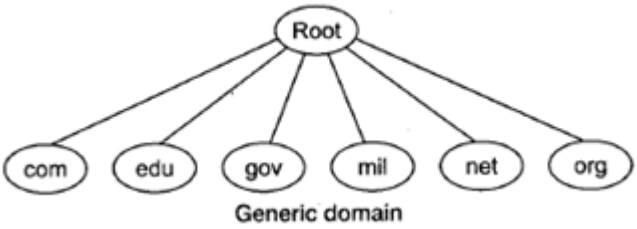
However, people refer to use names instead of address. Therefore, we need a system that can map a name to an address and conversely an address to name. In TCP/IP, this is the domain name system.

DNS in the Internet: DNS is protocol that can be used in different platforms.

3.8.1 Domain Name Space

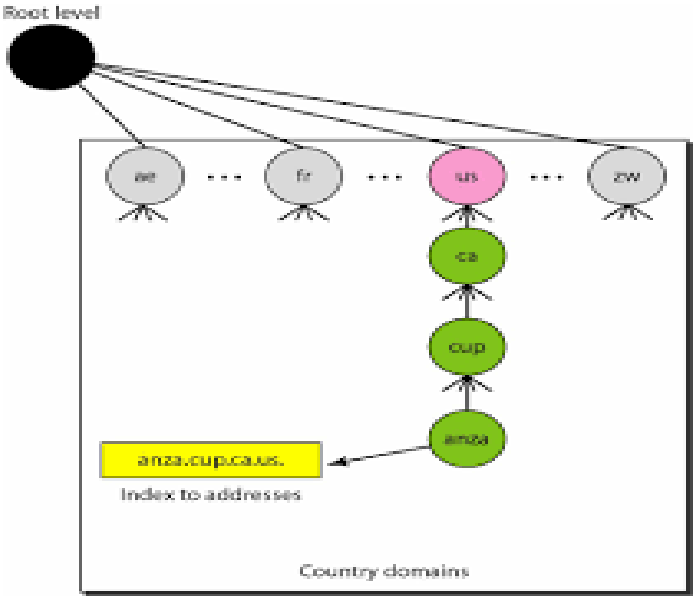
Domain name space is divided into three categories.

- **Generic Domain:** The generic domain defines registered hosts according, to their generic behaviour. Each node in the tree defines a domain which is an index to the domain name space database



.

- **Country Domain:** The country domain section follows the same format as the generic domain but uses 2 characters country abbreviations (e.g., US for United States) in place of 3 characters.



- **Inverse Domain:** The inverse domain is used to map an address to a name.

  Table: Domain names with description

Label	Description
aero	Airlines and aerospace companies
biz	Businesses or firms (similar to "com")
com	Commercial organizations
coop	Cooperative business organizations
edu	Educational institutions
gov	Government institutions
info	Information service providers
int	International organizations
mil	Military groups
museum	Museums and other nonprofit organizations
name	Personal names (individuals)
net	Network support centers
org	Nonprofit organizations
pro	Professional individual organizations

3.8.2 How does DNS work?

The DNS makes it possible to assign domain names in a meaningful way to Internet resources as well as to users, regardless of the entity's location. As a result, the WWW hyperlinks remain consistent, even for mobile devices. A domain name is an easy way to remember an address, but that needs to be converted to its numerical, IP format.

Coordination across the Internet is maintained by means of a complex authoritative root system known as the Top Level Domain (TLD), as well as the DNS and other smaller name servers responsible for hosting individual domain information.

DNS includes three types of top-level domains: generic (gTLD), country code (ccTLD), and sponsored (sTLD). gTLDs include domains that could be obtained by anyone (.com, .info, .net, and .org). Since 2014 many other gTLDs have been added like .pub, .ngo, .sucks. sTLDs are limited to a specific group e.g .aero (for air-transport industry).

For each domain, the DNS spreads the responsibility by mapping the domain names and assigning them into IP addresses, and vice-versa. This is accomplished through authoritative name servers which have been designated for each domain. Each authoritative name server is responsible for its own particular domain, but it has the authority to assign new authoritative name servers to any of its sub-domains. The DNS is able to store many types of information, even the mail server lists for a specific domain. The DNS is a core element which ensures the functionality of the Internet through its distributed keyword-based redirection service.

However, the DNS does not include security extensions, which was instead developed as DNSSEC.

3.8.2.1 Top-Level Domain (TLD)

Whenever you use a domain name, in a web address (URL), email address, or wherever, it *ends* in a "top-level domain" or "TLD". This is the last part of the name. We often thing of .COM, .ORG, .NET, etc., as in:

- www.disruptiveconversations.**COM**
- www.forimmediaterelease.**BIZ**
- internetsociety.**ORG**

TLDs are broadly classified into two categories:
a. generic top-level domains (gTLDs)

b. country code top-level domains (ccTLDs)
The entity responsible for the administration of these TLDs in the "root" of the Domain Name System (DNS) is the Internet Assigned Numbers Authority (IANA) that is currently operated by the Internet Corporation for Assigned Names and Numbers (ICANN). You can see the full list of current TLDs at: https://www.iana.org/domains/root/db

3.8.2.2 Second-Level Domain

The next part of the domain name to the left of the TLD (and separated by a dot) is the "second-level domain". These are the domains that you are typically able to register with a registrar. Examples include:

- **www.disruptiveconversations.com**
- **www.forimmediaterelease.biz**
- **internetsociety.org**

The next part of the domain name to the left ("www" in the first two examples above) would be called the "third-level domain", and so on.

a) **gTLD (Generic Top-Level Domain):** Generic top-level domains (gTLDs) are TLDs that are not tied to any specific country and are "generic" in terms of being able to be used (in theory, anyway) by anyone on the Internet anywhere in the world. The "original" TLDs such as .COM, .ORG, .NET, .GOV, .MIL are all classified as "generic TLDs". There were a couple of rounds of "expansion" of the gTLDs that brought the total to 22 gTLDs prior to the "newgTLD" expansion currently underway

b) **ccTLD (Country Code Top-Level Domain):** Country code top-level domains(ccTLDs) are *two letter TLDs* that are assigned to *countries* based mostly on the ISOC 3166 list of country codes. Some countries have chosen to operate their ccTLD exclusively for domains within their country or geographic territory. Some do not allow people to register "second-level domains" under the TLD and instead require

people to register third-level domains under one of several different second-level domains. For example, the .UK domain as to date required registrations to be under domains such as ".co.uk" and ".org.uk", basically duplicating part of the original gTLD scheme inside their ccTLD.

Many ccTLDs have chosen **NOT** to restrict their ccTLD to people in their country and have in fact marketed their domains very widely encouraging everyone to use them. Some prominent examples of this include Columbia (.CO), Montenegro (.ME), Tuvulu(.TV), Federated States of Micronesia(.FM) and many more. Essentially, any time you are using a two-letter TLD, it is a ccTLD for some country.

c) **newgTLD Top-Level Domain:** After many years of discussion, ICANN's board voted in 2011 to allow the creation of new generic TLDs using almost any text string (and in multiple character sets) and began the "newgTLD" program. This resulted in 1,930 applications by various companies to operate new gTLDs. These newgTLDs are now being rolled out in phases and people are able to register second-level domains under many of these domains. More newgTLDs are being made available pretty much every week - and the expansion will continue for many months and years ahead of us.

At a technical level, "new gTLDs" are effectively the same as "gTLDs"... the designation is just really that these new gTLDs are coming out in this current round of expansion.

d) **IDN = Internationalized Domain Name:** The original TLDs were all in the ASCII character set, but over time ICANN decided to allow the creation of "internationalized domain names"(IDNs) that use other character sets such as Cyrillic, Arabic, Chinese, Korean, etc. The first IDN for a country code TLD appeared in 2010 and the newgTLDs contain many IDNs. (In fact, the very first of the "newgTLDs" were four IDNs.)

3.9 TELNET (TERMINAL NETWORK)

TELNET is a general purpose client –server program that lets a used to access any application program on a remote computer. It allows the user to log on to a remote computer.

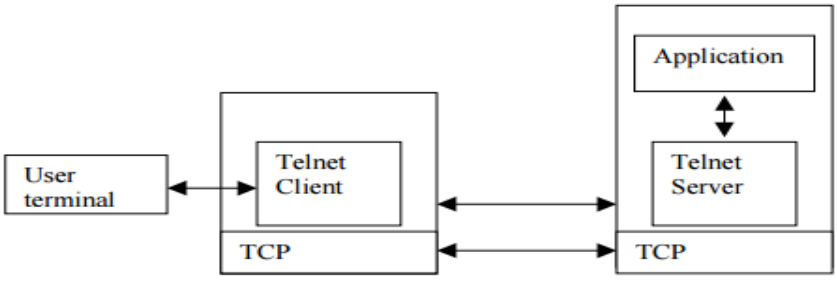


Fig. TELNET protocol model

After logging on a user can use the service available on the remote computer and transfer the results back to the local computer.

TELNET uses the NVT (Network Virtual Terminal) system to encode characters on the local system. On the server (remote) machine; NVT decodes the characters to a form acceptable to the remote machine.

TELNET is a protocol that provides a general, bi-directional, eight-bit byte oriented communications facility. Many application protocols are built upon the TELNET protocol.
So the TELNET is an abbreviation for TErminaL NET work.

TELNET enable the establishment of a connection to a remote system in such a way that the local terminal appears to be a terminal art the remote system.

Virtual Terminal in TELNET

To ensure compatibility between the user and the remote server, Telnet used a translation services between two machines.

- The TELNET client translates the output from the actual terminal to standard code.
- The information in the standard code is sent to the TELNET server in the remote host.
- The TELNET server translates the information in to characters accepted by the remote host .A virtual terminal is connected to the remote host.
- The concept of virtual terminal solves the problems of universal connectivity is connecting multiple incompatible terminals to a single remote server.

Local Login: When a user logs in to a local time – sharing system, it is called local login.

Remote Login: When a user wants to access an application program or utility located on a remote machine, he/she performs remote login. Here the TELNET client and server programs come in to use.

3.10 FILE TRANSFER PROTOCOL (FTP)

FTP is the standard mechanism provided by TCP/IP for copying a file from one host to another.FTP differs from other client server application in that it establishes two connections between hosts. One connection is used for data transfer, the other for control information

Separation of commands and responses /data transfer makes FTP more efficient.The control connection used very simple rules of communication.

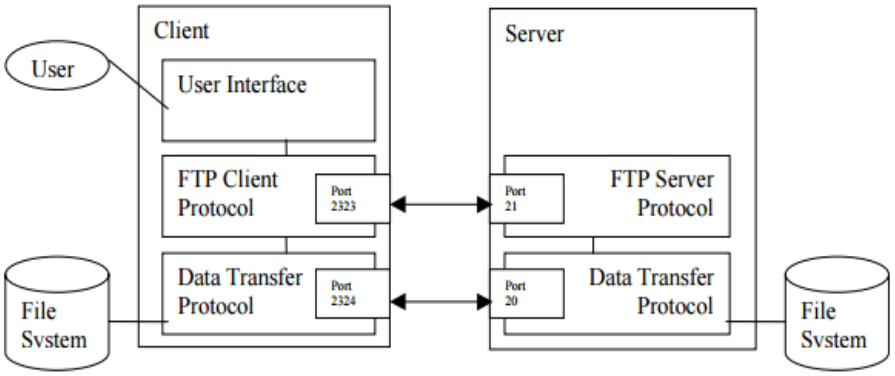


Fig: FTP protocol model

The control connection (used) remains connected during the entire interactive (Progresses) FTP session.

The data connection is opened and then closed for each file transferred. It opens each time commands that involve transferring files are used and closes when the file is transferred .The two FTP connections, control data use different strategies and port numbers.

Uses of FTP:

(i) It is very useful to transfer files from one network to another in an organisation.
(ii) It is an effective way to get a geographically – dispersed group to cooperate on a project.
(iii) It is a potent and popular way to share – information over the Internet.
(iv) It is both a program and method used to transfer files from one network.
(v) File transfer is quite rapid.
(vi) It does not need any special .Only the web- browser will suffice.

Working: FTP is not just name of the protocol; it is also name of the program. Issue the command by typing FTP followed by address of another site and presses enter.

FTP works as a client /server process you give the command FTP using a remote address such as FTP newday. horizon.com

The FTP running on your machine is client to an FTP process that acts as server on new day, horizon.com

Main Objectives of FTP are:

- Promote sharing of files
- To encourage and impolite use of remote computes
- To shield user from variations in file storage systems among hosts.
- To transfer remote data reliably and efficiently.

3.10 KEY TERMS AND CONCEPTS

Server - systems that store information shown on the web; stores web pages and other information used both on the internet and intranet.

Clients: A clients is a program running on a local machine requesting service from a server.

http:// - Hypertext Transfer Protocol; the protocol or standard that defines how all information is sent over the internet; usually precedes a URL.

ftp:// (File Transfer Protocol) moves a file between 2 computers.

telnet:// (Telnet client) allows remote login to another computer.

Domain – an identification string located at the end of a web address; examples - .net, .org, .com

The **Domain Name System (DNS)** is the system used to translate alphanumeric domain names into Internet Protocol numbers.

SMTP provides services for mail exchange between users on the same or different computers.

3.11 SELF-ASSESMENT QUESTIONS

1. Differentiate between HTTP and SMTP
   .....
   .....
   .....
   .....
2. Compare the protocols HTTP and FTP.
   .....
   .....
   .....
   .....
3. What are the two agents in the architecture of e-mail?
   .....
   .....
   .....
   .....
4. Write the basic functions of e –mail systems
   .....
   .....
   .....
   .....
5. What are the additional services provided by e-mail?
   .....
   .....
   .....
   .....
   ……..

3.13 RFERENCES AND SUGGESTED READINGS

1. Behrouz A. Forouzan, "*Introduction to Data Communications and Networking*", McGraw-Hill Education (India), New Delhi.
2. Andrew S. Tanenbaum, "*Computer Networks*", PHI Learning Pvt. Ltd
3. James F. Kurose, Keith W. Ross, "*Computer Networking: A Top-Down Approach Featuring the Internet*", Pearson Education Inc., New Delhi.
4. Wayne Tomasi, "*Introduction to Data Communications and Networking*", Pearson Education Inc., New Delhi.
5. Fundamentals of Information Security (PGDCS-01), Study Materials of Uttarakhand Open University, Haldwani, for Certificate in e e-Governance and Cyber Security.

UNIT-4
INTERNET AND ITS SERVICES

UNIT STRUCTURE

4.0 INTRODUCTION

The term "Internet" refers to a global network of routers and hosts that run the TCP/IP protocol suite, and use the same global IP address space - a network of networks in which users at any one computer can get information from any other computer in any part of the world. It was conceived by the Advanced Research Projects Agency (ARPA) of the U.S. government in 1969 and was first known as the ARPANet.

Today, the Internet is accessible to hundreds of millions of people worldwide. Physically, the Internet uses a portion of the total resources of the currently existing public telecommunication networks. Technically, what distinguishes the Internet is its use of a set of protocols called TCP/IP (for Transmission Control Protocol/Internet Protocol).

Two recent adaptations of Internet technology, the intranet and the extranet, also make use of the TCP/IP protocol.

For most Internet users, electronic mail (email) practically replaced the postal service for short written transactions. People communicate over the Internet in a number of other ways including Internet Relay Chat (IRC), Internet telephony, instant messaging, video chat or social media.

In this unit we will discuss in detail about the Internet, its working and its applications.

4.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Know the evolution of Internet.
- Understand the working of the Internet.
- Understand Domain Name System.
- Know the working of an ISP.
- Differentiate www with the Internet.

4.2 HISTORY OF THE INTERNET

1. J.C.R. Licklider of MIT first proposed a global network of computers in 1962, and moved over to the Defence Advanced Research Projects Agency (DARPA) in late 1962 to head the work to develop it.
2. Leonard Kleinrock of MIT and later UCLA developed the theory of packet switching, which was to form the basis of Internet connections.
3. The Internet, also known as ARPANET, was brought online in 1969 under a contract let by the renamed Advanced Research Projects Agency (ARPA) which initially connected four major computers at universities in the south-western US (UCLA, Stanford Research Institute, UCSB, and the University of Utah).

4. The Internet was designed to provide a communications network that would work even if some of the major sites were down.
5. E-mail was adapted for ARPANET by Ray Tomlinson of BBN in 1972.
6. In 1986, the National Science Foundation funded NSFNet as a cross country 56 Kbps backbone for the Internet. They maintained their sponsorship for nearly a decade, setting rules for its non-commercial government and research uses.
7. The first effort, other than library catalogs, to index the Internet was created in 1989, as Peter Deutsch and Alan Emtage, students at McGill University in Montreal, created an archiver for ftp sites, which they named Archie.
8. Since the Internet was initially funded by the government, it was originally limited to research, education, and government uses.

The early Internet was used by computer experts, engineers, scientists, and librarians. There was nothing friendly about it. There were no home or office personal computers in those days, and anyone who used it, whether a computer professional or an engineer or scientist or librarian, had to learn to use a very complex system.

4.3 HOW INTERNET WORKS?

The Internet is a network of networks—millions of them, actually. If the network at your university, your employer, or in your home has Internet access, it connects to an Internet service provider (ISP).

The Internet has no centre and no one owns it. The Internet was designed to be redundant and fault-tolerant—meaning that if one network, connecting wire, or server stops working, everything else should keep on running.

Now let‘s see how the Internet works! If you want to communicate with another computer on the Internet then your computer needs to know the answer to three questions: What are you looking for? Where is it? And how do we get there? The computers and software that make up Internet infrastructure can help provide the answers. Let‘s look at how it all comes together.

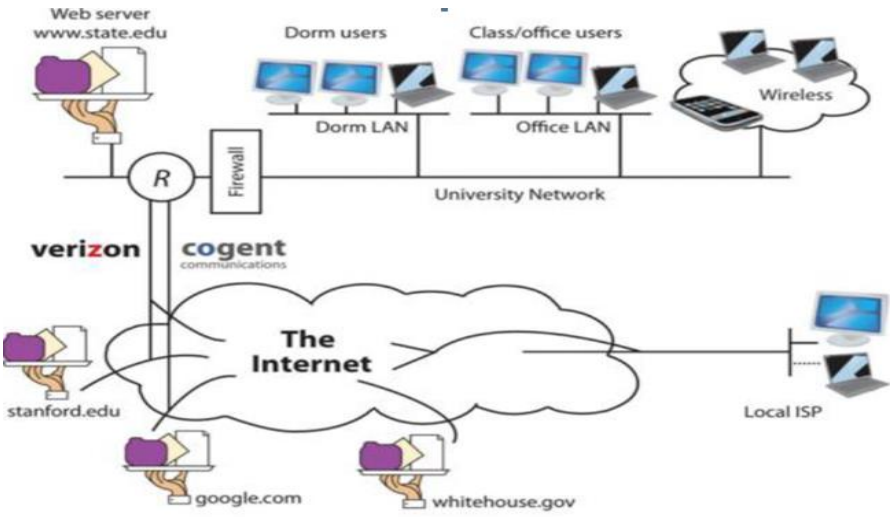


Fig: Working of the Internet

When you type an address into a Web browser (sometimes called a URL for uniform resource locator), you're telling your browser what you're looking for, Figure below describes how to read a typical URL.

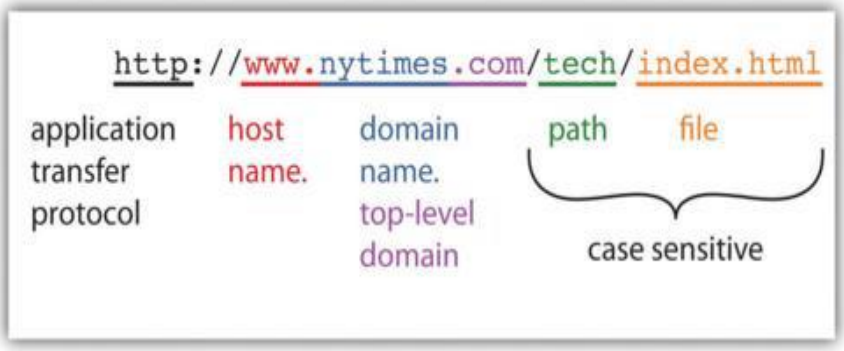


Fig: Anatomy of a Web Address

The http:// you see at the start of most Web addresses stands for hypertext transfer protocol. The http protocol defines how Web browser and Web servers communicate and is designed to be independent from the computer's hardware and operating system.

4.3.1 Hosts and Domain Names

The next part of the URL in our diagram holds the host and domain name. Think of the domain name as the name of the network you're trying to connect to, and think of the host as the computer you're looking for on that network.

Many domains have lots of different hosts. For example, Yahoo!'s main Web site is served from the host named ―www‖ (at the address http://www.yahoo.com), but Yahoo! also runs other hosts including those named ―finance‖ (finance.yahoo.com), ―sports‖ (sports.yahoo.com), and ―games‖ (games.yahoo.com).

Host and domain names are not case sensitive, so you can use a combination of upper and lower case letters and you'll still get to your destination.

4.4 ADDRESSING SCHEME IN THE INTERNET

An Internet Protocol address (IP address) is a numerical label assigned to each device (e.g., computer, printer) participating in a computer network that uses the Internet Protocol for communication. An IP address serves two principal functions: host or network interface identification and location addressing. Its role has been characterized as follows: "A name indicates what we seek. An address indicates where it is. A route indicates how to get there.

4.4.1 IP versions

Two versions of the Internet Protocol (IP) are in use: IP Version 4 and IP Version 6. Each version defines an IP address differently. Because of its prevalence, the generic term IP address typically still refers to the addresses defined by IPv4. The gap in version sequence between IPv4 and IPv6 resulted from the assignment of number 5 to the experimental Internet Stream Protocol in 1979, which however was never referred to as IPv5.

4.4.2 IPv4 Addresses

In IPv4 an address consists of 32 bits which limits the address space to 4294967296 (232) possible unique addresses. IPv4 reserves some addresses for special purposes such as private networks (~18 million addresses) or multicast addresses (~270 million addresses). IPv4 addresses are canonically represented in dot-decimal notation, which consists of four decimal numbers, each ranging from 0 to 255, separated by dots, e.g., 172.16.254.1. Each part represents a group of 8 bits (octet) of the address. In some cases of technical writing, IPv4 addresses may be presented in various hexadecimal, octal, or binary representations.

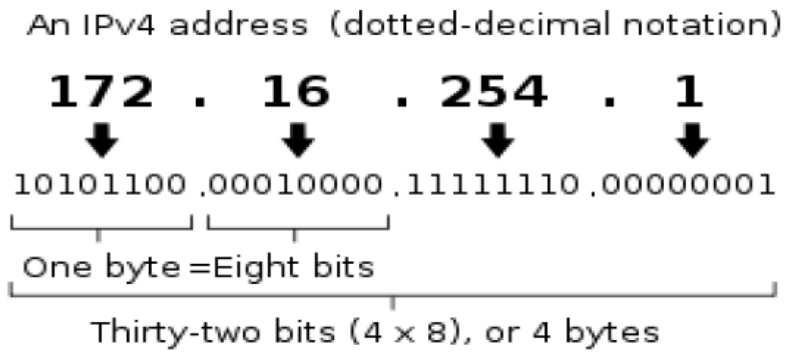


Figure 3: IPv4 address

4.4.2.3 IPv4 address exhaustion

High levels of demand have decreased the supply of unallocated Internet Protocol Version 4 (IPv4) addresses available for assignment to Internet service providers and end user organizations since the 1980s. This development is referred to as IPv4 address exhaustion. IANA's primary address pool was exhausted on 3 February 2011, when the last five blocks were allocated to the five RIRs. APNIC was the first RIR to exhaust its regional pool on 15 April 2011, except for a small amount of address space reserved for the transition to IPv6, intended to be allocated in a restricted process.

4.4.3 IPv6 Addresses

The rapid exhaustion of IPv4 address space prompted the Internet Engineering Task Force (IETF) to explore new technologies to expand the addressing capability in the Internet. The permanent solution was deemed to be a redesign of the Internet Protocol itself. This new generation of the Internet Protocol was eventually named Internet Protocol Version 6 (IPv6) in 1995. The address size was increased from 32 to 128 bits (16 octets), thus providing up to 2128 (approximately 3.403×1038) addresses. This is deemed sufficient for the foreseeable future.

The intent of the new design was not to provide just a sufficient quantity of addresses, but also redesign routing in the Internet by more efficient aggregation of subnetwork routing prefixes. This resulted in slower growth of routing tables in routers. The smallest possible individual allocation is a subnet for 264 hosts, which is the square of the size of the entire IPv4 Internet. At these levels, actual address utilization rates will be small on any IPv6 network segment. The new design also provides the opportunity to separate the addressing infrastructure of a network segment, i.e. the local administration of the segment's available space, from the addressing prefix used to route traffic to and from external networks. IPv6 has facilities that automatically change the routing prefix of entire networks, should the global connectivity or the routing policy change, without requiring internal redesign or manual renumbering. The large number of IPv6 addresses allows large blocks to be assigned for specific purposes and, where appropriate, to be aggregated for efficient routing. With a large address space, there is no need to have complex address conservation methods as used in CIDR.

All modern desktop and enterprise server operating systems include native support for the IPv6 protocol, but it is not yet widely deployed in other devices, such as residential networking routers, voice over IP (VoIP) and multimedia equipment, and network peripherals.

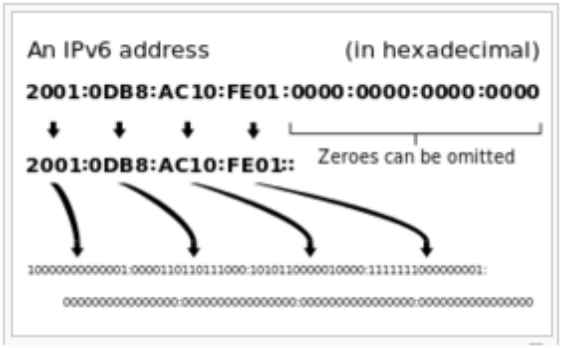


Fig: Decomposition of an IPv6 address from hexadecimal representation to its binary value

4.4.5 IP address assignment

Internet Protocol addresses are assigned to a host either anew at the time of booting, or permanently by fixed configuration of its hardware or software. Persistent configuration is also known as using a static IP address. In contrast, in situations when the computer's IP address is assigned newly each time, this is known as using a dynamic IP address.

4.4.5.1 Methods

Static IP addresses are manually assigned to a computer by an administrator. The exact procedure varies according to platform. This contrasts with dynamic IP addresses, which are assigned either by the computer interface or host software itself, as in Zeroconf, or assigned by a server using Dynamic Host Configuration Protocol (DHCP). Even though IP addresses assigned using DHCP may stay the same for long periods of time, they can generally change. In some cases, a network administrator may implement dynamically assigned static IP addresses. In this case, a DHCP server is used, but it is specifically configured to always assign the same IP address to a particular computer. This allows static IP addresses to be configured centrally, without having to specifically configure each computer on the network in a manual procedure.

4.4.5.2 Uses of dynamic address assignment

IP addresses are most frequently assigned dynamically on LANs and broadband networks by the Dynamic Host Configuration Protocol (DHCP). They are used because it avoids the administrative burden of assigning specific static addresses to each device on a network. It also allows many devices to share limited address space on a network if only some of them will be online at a particular time. In most current desktop operating systems, dynamic IP configuration is enabled by default so that a user does not need to manually enter any settings to connect to a network with a DHCP server. DHCP is not the only technology used to assign IP addresses dynamically. Dialup and some broadband networks use dynamic address features of the Point-to-Point Protocol.

4.4.5.3 Address auto configuration

RFC 3330 defines an address block, 169.254.0.0/16, for the special use in link-local addressing for IPv4 networks. In IPv6, every interface, whether using static or dynamic address assignments, also receives a local-link address automatically in the block fe80::/10.

These addresses are only valid on the link, such as a local network segment or point-to-point connection, that a host is connected to. These addresses are not routable and like private addresses cannot be the source or destination of packets traversing the Internet.

When the link-local IPv4 address block was reserved, no standards existed for mechanisms of address auto configuration. Filling the void, Microsoft created an implementation that is called Automatic Private IP Addressing (APIPA). APIPA has been deployed on millions of machines and has, thus, become a de facto standard in the industry. In RFC 3927, the IETF defined a formal standard for this functionality, entitled Dynamic Configuration of IPv4 Link-Local Addresses.

4.4.5.4 Uses of static addressing

Some infrastructure situations have to use static addressing, such as when finding the Domain Name System (DNS) host that will translate domain names to IP addresses. Static addresses are also convenient, but not absolutely necessary, to locate servers inside an enterprise. An address obtained from a DNS server comes with a time to live, or caching time, after which it should be looked up to confirm that it has not changed. Even static IP addresses do change as a result of network administration (RFC 2072).

4.4.6 Routing

IP addresses are classified into several classes of operational characteristics: unicast, multicast, anycast and broadcast addressing.

4.4.6.1 Unicast addressing

The most common concept of an IP address is in unicast addressing, available in both IPv4 and IPv6. It normally refers to a single sender or a single receiver, and can be used for both sending and receiving. Usually, a unicast address is associated with a single device or host, but a device or host may have more than one unicast address. Some individual PCs have several distinct unicast addresses, each for its own distinct purpose. Sending the same data to multiple unicast addresses requires the sender to send all the data many times over, once for each recipient.

4.4.6.2 Broadcast addressing

In IPv4 it is possible to send data to all possible destinations ("all-hosts broadcast"), which permits the sender to send the data only once, and all

receivers receive a copy of it. In the IPv4 protocol, the address 255.255.255.255 is used for local broadcast. In addition, a directed (limited) broadcast can be made by combining the network prefix with a host suffix composed entirely of binary 1s. For example, the destination address used for a directed broadcast to devices on the 192.0.2.0/24 network is 192.0.2.255. IPv6 does not implement broadcast addressing and replaces it with multicast to the specially-defined all-nodes multicast address.

4.4.6.3 Multicast addressing

A multicast address is associated with a group of interested receivers. In IPv4, addresses 224.0.0.0 through 239.255.255.255 (the former Class D addresses) are designated as multicast addresses. IPv6 uses the address block with the prefix ff00::/8 for multicast applications. In either case, the sender sends a single datagram from its unicast address to the multicast group address and the intermediary routers take care of making copies and sending them to all receivers that have joined the corresponding multicast group.

4.4.6.4 Anycast addressing

Like broadcast and multicast, anycast is a one-to-many routing topology. However, the data stream is not transmitted to all receivers; just the one which the router decides is logically closest in the network. Anycast address is an inherent feature of only IPv6. In IPv4, anycast addressing implementations typically operate using the shortest-path metric of BGP routing and do not take into account congestion or other attributes of the path. Anycast methods are useful for global load balancing and are commonly used in distributed DNS systems.

4.4.7 Public addresses

A public IP address, in common parlance, is synonymous with a globally routable unicast IP address. Both IPv4 and IPv6 define address ranges that are reserved for private networks and link-local addressing. The term public IP address often used excludes these types of addresses.

4.4.8 Modifications to IP addressing

4.4.8.1 IP blocking and firewalls

Firewalls perform Internet Protocol blocking to protect networks from unauthorized access. They are common on today's Internet. They control access to networks based on the IP address of a client computer. Whether using a blacklist or a white list, the IP address that is blocked is the perceived IP address of the client, meaning that if the client is using a proxy server or network address translation, blocking one IP address may block many individual computers.

4.4.8.2 IP address translation

Multiple client devices can appear to share IP addresses: either because they are part of a shared hosting web server environment or because an IPv4 network address translator (NAT) or proxy server acts as an intermediary agent on behalf of its customers, in which case the real originating IP addresses might be hidden from the server receiving a request. A common practice is to have a NAT hide a large number of IP addresses in a private network. Only the "outside" interface(s) of the NAT need to have Internet-routable addresses.

Most commonly, the NAT device maps TCP or UDP port numbers on the side of the larger, public network to individual private addresses on the masqueraded network. In small home networks, NAT functions are usually implemented in a residential gateway device, typically one marketed as a "router". In this scenario, the computers connected to the router would have private IP addresses and the router would have a public address to communicate on the Internet. This type of router allows several computers to share one public IP address.

4.5 INTERNET SERVICE PROVIDER

An Internet service provider (ISP) is an organization that provides services for accessing, using, or participating in the Internet5. Internet service providers may be organized in various forms, such as commercial, community-owned, non-profit, or otherwise privately owned. Internet services typically provided by ISPs include Internet access, Internet transit, domain name registration, web hosting, Usenet service, and collocation.

4.5.1 Classification of ISP6

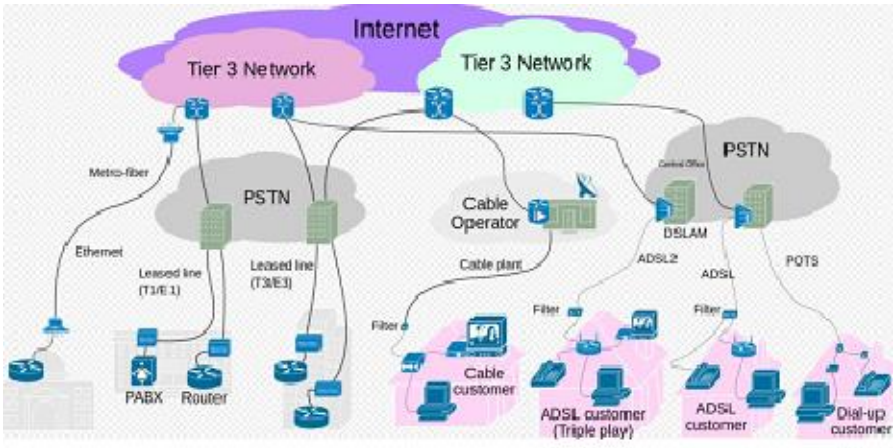


Fig: Internet connectivity options from end-user to tier 3/2 ISPs

4.5.1.1 Access providers ISP

ISPs provide Internet access, employing a range of technologies to connect users to their network. Available technologies have ranged from computer

modems with acoustic couplers to telephone lines, to television cable (CATV), wireless Ethernet (wi-fi), and fiber optics.
For users and small businesses, traditional options include copper wires to provide dial-up, DSL (typically asymmetric digital subscriber line (ADSL), cable modem or Integrated Services Digital Network (ISDN) (typically basic rate interface). Using fiber-optics to end users is called Fiber to the Home or similar names. For customers with more demanding requirements (such as medium-to-large businesses, or other ISPs) can use higher-speed DSL (such as single-pair high-speed digital subscriber line), Ethernet, metropolitan Ethernet, gigabit Ethernet, Frame Relay, ISDN Primary Rate Interface, ATM (Asynchronous Transfer Mode) and synchronous optical networking (SONET). Wireless access is another option, including satellite Internet access.

4.5.1.2 Mailbox providers

A mailbox provider is an organization that provides services for hosting electronic mail domains with access to storage for mail boxes. It provides email servers to send, receive, accept, and store email for end users or other organizations. Many mailbox providers are also access providers, while others are not (e.g., Yahoo! Mail, Outlook.com, Gmail, AOL Mail, PO box). The definition given in RFC 6650 covers email hosting services, as well as the relevant department of companies, universities, organizations, groups, and individuals that manage their mail servers themselves. The task is typically accomplished by implementing Simple Mail Transfer Protocol (SMTP) and possibly providing access to messages through Internet Message Access Protocol (IMAP), the Post Office Protocol, Webmail, or a proprietary protocol.

4.5.1.3 Hosting ISPs

Internet hosting services provide email, web-hosting or online storage services. Other services include virtual server, cloud services, or physical server operation.

4.5.1.4 Transit ISP

Just as their customers pay them for Internet access, ISPs themselves pay upstream ISPs for Internet access. An upstream ISP usually has a larger network than the contracting ISP or is able to provide the contracting ISP with access to parts of the Internet the contracting ISP by itself has no access to. In the simplest case, a single connection is established to an upstream ISP and is used to transmit data to or from areas of the Internet beyond the home network; this mode of interconnection is often cascaded multiple times until reaching a tier 1 carrier. In reality, the situation is often more complex. ISPs with more than one point of presence (PoP) may have separate connections to an upstream ISP at multiple PoPs, or they may be

customers of multiple upstream ISPs and may have connections to each one of them at one or more point of presence. Transit ISPs provide large amounts of bandwidth for connecting hosting ISPs and access ISPs.

4.5.1.5 Virtual ISPs

A virtual ISP (VISP) is an operation that purchases services from another ISP, sometimes called a wholesale ISP in this context, which allow the VISP's customers to access the Internet using services and infrastructure owned and operated by the wholesale ISP. VISPs resemble mobile virtual network operators and competitive local exchange carriers for voice communications.

4.5.1.6 Free ISPs

Free ISPs are Internet service providers that provide service free of charge. Many free ISPs display advertisements while the user is connected; like commercial television, in a sense they are selling the user's attention to the advertiser. Other free ISPs, sometimes called freenets, are run on a nonprofit basis, usually with volunteer staff.

4.5.1.7 Wireless ISP

A wireless internet service provider (WISP) is an Internet service provider with a network based on wireless networking. Technology may include commonplace Wi-Fi wireless mesh networking, or proprietary equipment designed to operate over open 900 MHz, 2.4 GHz, 4.9, 5.2, 5.4, 5.7, and 5.8 GHz bands or licensed frequencies such as 2.5 GHz (EBS/BRS), 3.65 GHz (NN) and in the UHF band (including the MMDS frequency band) and LMDS.

4.6 WORLD WIDE WEB (WWW*)*

The World Wide Web is a system of interlinked hypertext documents accessed via the Internet. With a web browser, one can view web pages that may contain text, images, videos, and other multimedia, and navigate between them via hyperlinks. Using concepts from his earlier hypertext systems like ENQUIRE, British engineer, computer scientist and at that time employee of CERN, Sir Tim Berners-Lee, now Director of the World Wide Web Consortium (W3C), wrote a proposal in March 1989 for what would eventually become the World Wide Web.

4.6.1 Internet vs. WWW?

The World Wide Web is just one way to access information through the internet10. While it does represent a considerable portion of the internet, and is unquestionably the most popular part, the two concepts must not be treated as synonyms because they are not the same. We tend to become used to calling things by the simplest possible name but we also tend to

muddle concepts and mix up one thing with another when the distinction between them isn‘t very clear. One very common case of this is the fact that most people tend to refer to ―the web‖ and ―internet‖ as if they were exactly the same thing, when in fact they‘re not. It can be rather confusing, and even a surprise for many, but the internet and the web are two different things, and one is above the other. Let‘s see what this means.

4.6.1.1 The www

The three Ws that are in the addresses of the websites we access. The World Wide Web or simply the ―web‖, is a way to access information through the internet. The web is a model for sharing information that is built on the internet. The protocol used by the web is HTTP, just one of the many ways that information can be sent through the internet.

If a service uses HTTP to enable applications to communicate with each other, this is a web service. Web browsers, such as Chrome or Firefox, enable us to access web documents that we mainly know as web pages or websites. These sites are connected to each other through hyperlinks as if they were on a spider‘s web (hence the name), and all this thanks to the transfer protocol: HTTP.

Therefore, the web is only one of the ways that information can flow through the internet: it is just a portion, and although it is very large and the most popular part, it does not include the whole of the internet.

4.6.1.2 The Internet

The internet is a massive network, the network of networks. The internet connects millions of computers across the globe through a network that enables any computer to be able to communicate with another, no matter where on the planet they are, provided they are both connected to the internet. A network is any connection between two or more clients. For example, you can access a local network in your home that only the computers of the members of your family can access and which is connected through a switcher or router, or a work network that only people working at the same firm can access. The internet is a global, large-scale network that enables millions and millions of devices to connect at the same time, and is completely free and open.

All the information that travels through the internet does so through a protocol; there are several of these. As we have already explained, the HTTP protocol is the one used by the web to share information. Therefore, web pages such as Twitter, Google, Facebook and even this blog are part of the web and this information travels to us all through the internet.

When it comes down to it, the world won‘t end if we continue using the terms interchangeably – after all, habits are hard to break – but it is a good thing to be clear about the concepts, at least.

4.6.1.3 Intranet

An intranet is a privately owned network that is contained within a company or organization. Each computer in Intranet is also identified by an IP Address which is unique among the computers in that Intranet. It may consist of many interlinked local area networks and also use leased lines in the wide area network. Typically, an intranet includes connections through one or more gateways computers to the outside Internet. The main purpose of an intranet is to share company information and computing resources among employees. An intranet can also be used to facilitate working in groups and for teleconferences. PCs in intranet are not available to the world outside the intranet. Usually each company or organization has their own Intranet network and members/employees of that company can access the computers in their intranet.

4.6.1.4 Similarities between Internet and Intranet

- Intranet uses the internet protocols such as TCP/IP and FTP.
- Intranet sites are accessible via web browser in similar way as websites in internet. But only members of Intranet network can access intranet hosted sites.
- In Intranet, own instant messengers can be used as similar to yahoo messenger/ gtalk over the internet.

4.6.1.5 Differences between Internet and Intranet

- Internet is general to PCs all over the world whereas Intranet is specific to few PCs.
- Internet has wider access and provides a better access to websites to large population whereas Intranet is restricted.
- Internet is not as safe as Intranet as Intranet can be safely privatized as per the need.

4.7 SERVICES PROVIDED BY THE INTERNET

Some of the important services provided by the Internet are:

World Wide Web (WWW): It is a subset of the Internet and it presents text, images, animation, video, sound, and other multimedia in a single interface. The operation of the Web relies primarily on hypertext, as it is a means of information retrieval.

Electronic Mail (E-Mail): It is the process of exchanging messages electronically, via a communications network, using the computer.

File Transfer Protocol (FTP): It is a system of rules and a software program that enables a user to log on to another computer and transfer information between it and his/her computer.

Telnet: It connects one machine to another in such a way that a person may interact with another machine as if it is being used locally.

Internet Relay Chat (IRC): This service allows people to communicate in real time and carry on conversations via the computer with one or more people. It provides the user with the facility to engage in simultaneous synchronous) online 'conversations' with other users from anywhere in the world.

Chatting and Instant Messaging: Chat programs allow users on the Internet to communicate with each other by typing in real time. Instant messaging allows a user on the Web to contact another user currently logged in and type a conversation.

Internet Telephony: It refers to the use of the Internet rather than the traditional telephone company infrastructure, to exchange spoken or other telephonic information.

Video Conferencing: It uses the same technology as IRC, but also provides sound and video pictures. It enables direct face-to-face communication across networks via web cameras, microphones, and other communication tools.

Commerce through Internet: It refers to buying and selling goods and services online.

Newsgroups (Usenet): It is an international discussion group that focuses on a particular topic and helps in gathering information about that topic.

Mailing Lists (List server): It refers to a large community of individuals who carry out active discussions, organized around topic-oriented forums that are distributed via e-mail and this method is known as mailing list.

4.8 APPLICATION OF INTERNET

We can roughly separate internet applications into the following types: online media, online information search, online communications, online communities, online entertainment, e-business, online finance and other applications.

The internet is treated as one of the biggest invention. It has a large number of uses.

1. Communication
2. Job searches
3. Finding books and study material
4. Health and medicine
5. Travel
6. Entertainment
7. Shopping

8. Stock market updates
9. Research
10. Business use of internet: different ways by which internet can be used for business are:
    i. Information about the product can be provided can be provided online to the customer.
    ii. Provide market information to the business
    iii. It helps business to recruit talented people
    iv. Help in locating suppliers of the product
    v. Fast information regarding customers view about companies product

4.9. KEY WORDS AND CONCEPTS

An internet is a collection of communications networks interconnected by bridges and/or routers

The Internet – (note upper case I) is the *global collection* of thousands of individual machines and networks. The Internet has no centre and no one owns it.

Intranet is the corporate internet operating within the organization

Uses Internet (TCP/IP and http) technology to deliver documents and resources

Search engine – a computer program that searches for specific words or phrases on the web

Intranet - an internal network of data and information that is used in many companies; typically password protected, accessible only from within the company's confines, and housed on a separate server. Contains the same features as the internet

Internet service provider (ISP) - a company that provides customers access to the internet

URL – Uniform Resource Locator; a unique address that sends a request to the server which houses the information you are looking.

The **World Wide Web (WWW)** is a system of interlinked hypertext documents accessed via the Internet. The www is a way to access information through the internet.

4.10 SELF ASSESMENT QUESTIONS

1. What is the Internet?

   .....
   .....
   .....
   .....

2. Write a brief notes on WWW

   .....
   .....
   .....
   .....

3. What are Domains?

   .....
   .....
   .....
   .....

4. What is a Web Browser?

   .....
   .....
   .....
   .....

5. What is a Search engine?

   .....
   .....
   .....
   .....

4.11 RFERENCES AND SUGGESTED READINGS

1. Behrouz A. Forouzan, "*Introduction to Data Communications and Networking*", McGraw-Hill Education (India), New Delhi.
2. Andrew S. Tanenbaum, "*Computer Networks*", PHI Learning Pvt. Ltd
3. James F. Kurose, Keith W. Ross, "*Computer Networking: A Top-Down Approach Featuring the Internet*", Pearson Education Inc., New Delhi.
4. Wayne Tomasi, "*Introduction to Data Communications and Networking*", Pearson Education Inc., New Delhi.
5. *Fundamentals of Information Security (PGDCS-01),* Study Materials of Uttarakhand Open University, Haldwani, for Certificate in e e-Governance and Cyber Security.

UNIT-5
NETWORK SECURITY

UNIT STRUCTURE

5.0 INTRODUCTION

Development of Information and Communication Technology has lead to increase in use of electronic devices and technologies such as Computers, Internet (or Network), Mobile phone, Laptops, Tablets, Hard-disk etc among the people day by day.

These devices have internal and external memory which contains electronic data. This data can be confidential, public or private. Now, the security of such data becomes mandatory for all users so as to prevent it from any form of attack which can make this data corrupted.

Therefore, Security is a very essential part of day-to-day activities. Now, we will start with defining the term "Security". Security refers to the state of being secure or the precautions taken to ensure against theft, misuse, etc or protection of assets or free from danger or attack or threat.

Overall, Security ensures that all processes work as expected. We often refer this security mechanism as cyber security or information technology security that focuses on protecting computers, networks, programs and data from unintended or unauthorized access, change or destruction.

It is the most critical factor and has minimal standard which should be maintained by an individual or organization. This brings reliability, safety and assurance of being protected. In this unit, we will introduce to the types of security and its services like Confidentiality, Availability, Integrity, Authentication, and Non-Repudiation etc.

5.1 LEARNING OBJECTIVES

After going through this unit you should be able to

- Identify different types of security
- Know different types of security goals and services
- Explain the functions of firewalls
- Understand the functions of antivirus

5.2 TYPES OF SECURITY

Information Technology (IT) Security – consists of following types:

1. Application security,
2. Computer security,
3. Data security,
4. Information security
5. Network security

5.2.1 Application Security

Application security prevents attack and vulnerabilities on an application. This application can be a mobile application or any other application such as web application etc. The security of an application remains throughout its lifecycle from initial phase to its running phase (or application phase) and on maintenance phase too.

5.2.2 Computer Security

Computer security is about securing a computer system (Desktop or Laptop etc) or a host. This type of security ensures a computer virus free with the help of anti-virus software. Moreover, a computer should use genuine and updated software and hardware. Also it should be protected with a password. This type of security is a form of computer security.

5.2.3 Data Security

Data Security involves security of electronic data which is present on any hard-disks / secondary storage either of computer system or on network, on server, etc. Such security can be implemented by using passwords, cryptography (through encryption and decryption), biometric authentication, or through access control list etc.

5.2.4 Information Security

Information Security is defined as protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, perusal, inspection, recording or destruction. This involves security of electronic data which is present on any database or file in any electronic memory. "Data Security" and "Information Security" are used interchangeably and are almost similar.

5.2.5 Network Security

Network Security takes care of a network, its associated processes and aims to secure it. This network can be an organizational/company internal network or any external network. All data which is coming inside the network and going outside the network is analyzed and monitored to keep the network danger free. Moreover, every process which is part of the network is also monitored.

Thus network security mechanisms need to protect the network and the network-accessible resources from unauthorized access by consistent and continuous monitoring and measurement of its effectiveness.

5.3 NEED OF SECURITY

The following vulnerabilities demand the need for security.

- To isolate data from getting hacked, corrupted, unauthorized access, unauthorized modified, unauthorized deleted etc
- To maintain confidentiality, availability and integrity of data
- To prevent electronic mail from getting hacked and unauthorized access.
- To protect easy passwords and pins being cracked.
- To eradicate vulnerabilities (weakness) in the system or data.

5.4 SECURITY SERVICES

In order to overcome the above mentioned vulnerabilities of a system or data or network and other assets, there are 5 major security services such as

Confidentiality, Integrity, Availability, Non-Repudiation and Authentication which are as follows:

5.4.1 Confidentiality

Confidentiality means keeping information secret from unauthorized access and is probably the most common aspect of information security. It is important to protect confidential information. An organization needs to guard against those malicious actions that endanger the confidentiality of its information. For example, an account user is authorized to see his account transaction online and no other account user can access this data as it is confidential.

5.4.2 Integrity

Information needs to be changed constantly. In a bank, when a customer deposits or withdraws money, the balance of their account needs to be changed. Integrity means that changes should be done only by authorized users and through authorized mechanisms. Moreover, the changes should get reflected at all the ends on which the changed information is accessed.

5.4.3 Availability

The third component of information security services is availability. The information created and stored by an organization needs to be available to authorized users and applications. Information is useless if it is not available o authorized users.

Information needs to be changed constantly, which means that it must be accessible to those authorized to access it. Unavailability of information is just as harmful to an organization as a lack of confidentiality or integrity. Imagine what would happen to a bank if the customers could not access their accounts for transactions. Therefore, information should be accessible and useable upon appropriate demand by an authorized user and availability is the prevention of unauthorized withholding of information.

5.4.4 Authentication

Authentication is the process by which a person or other entity proves that it is who (or what) it says it is. For example, a bank authenticates a person or entity that deal before transferring something valuable, such as information or money, to or from, it.

Authentication is achieved by presenting some unique identifying entity to the endpoint that is undertaking the process. An example of this process is the way you authenticate yourself with an ATM - here you insert your bank card (something you have) and enter your personal identification number (PIN –Personal Identification Number, something you know). Another example can be the authentication process for email account. In this case, you have the email address and you know the corresponding account password to access the account.

5.4.5 Non-Repudiation

Non-repudiation is the prevention of either the sender or the receiver denying a transmitted message. A system must be able to prove that certain

messages were sent and received. Non-repudiation is often implemented by using digital signatures. For example, a user A sent a message to user B. At later stage, user A should not deny of having sent the message to user B.

5.4.6 Access Control

Another additional security service can be access control. Access control means control of access through identification and authentication. A system needs to be able to identify and authenticate users for access to data, applications and hardware. In a large system there may be a complex structure determining which users and applications have access to which objects. This is done through Access Control List (ACL). For example, an account holder while checking his data online can only view data but cannot modify it. This is because of the reason of access given to the user on the basis of his role and identity.

5.5 FIREWALLS

Internet connectivity is no longer an option for most organizations. However, while internet access provides benefits to the organization, it enables the outside world to reach and interact with local network assets. This creates the threat to the organization. While it is possible to equip each workstation and server on the premises network with strong security features, such as intrusion protection, this is not a practical approach. The alternative, increasingly accepted, is the firewall.

The firewall is inserted between the premise network and internet to establish a controlled link and to erect an outer security wall or perimeter. The aim of this perimeter is to protect the premises network from internet based attacks and to provide a single choke point where security and audit can be imposed. The firewall can be a single computer system or a set of two or more systems that cooperate to perform the firewall function.

5.5.1 Firewall characteristics:

All traffic from inside to outside, and vice versa, must pass through the firewall. This is achieved by physically blocking all access to the local network except via the firewall.

Various configurations are possible. Only authorized traffic, as defined by the local security policy, will be allowed to pass. Various types of firewalls are used, which implement various types of security policies. The firewall itself is immune to penetration. This implies that use of a trusted system with a secure operating system. This implies that use of a trusted system with a secure operating system.

Four techniques that firewall use to control access and enforce the site's security policy is as follows:

Service control – determines the type of internet services that can be accessed, inbound or outbound. The firewall may filter traffic on this basis of IP address and TCP port number; may provide proxy software that

receives and interprets each service request before passing it on; or may host the server software itself, such as web or mail service.

Direction control – determines the direction in which particular service request may be initiated and allowed to flow through the firewall.

User control – controls access to a service according to which user is attempting to access it.

Behaviour control – controls how particular services are used.

5.5.2 Capabilities of Firewall

A firewall defines a single choke point that keeps unauthorized users out of the protected network, prohibits potentially vulnerable services from entering or leaving the network, and provides protection from various kinds of IP spoofing and routing attacks.

A firewall provides a location for monitoring security related events. Audits and alarms can be implemented on the firewall system.

A firewall is a convenient platform for several internet functions that are not security related. A firewall can serve as the platform for IPsec.

5.5.3 Limitations of Firewall

The firewall cannot protect against attacks that bypass the firewall. Internal systems may have dial-out capability to connect to an ISP. An internal LAN may support a modem pool that provides dial-in capability for travelling employees and telecommuters. The firewall does not protect against internal threats. The firewall does not protect against internal threats, such as a disgruntled employee or an employee who unwittingly cooperates with an external attacker.

The firewall cannot protect against the transfer of virus-infected programs or files. Because of the variety of operating systems and applications supported inside the perimeter, it would be impractical and perhaps impossible for the firewall to scan all incoming files, e-mail, and messages for viruses.

5.6 MALWARE

Malware stands for ―*Malicious Software* and it is designed to gain access or installed into the computer without the consent of the user. They perform unwanted tasks in the host computer for the benefit of a third party. There is a full range of malwares which can seriously degrade the performance of the host machine. There is a full range of malwares which are simply written to distract/annoy the user, to the complex ones which captures the sensitive data from the host machine and send it to remote servers.

5.7 TYPES OF MALWARE

There are various types of malwares present in the Internet. Some of the popular ones are:

5.7.1 Adware

It is a special type of malware which is used for forced advertising. They either redirect the page to some advertising page or pop-up an additional page which promotes some product or event. These adware are financially supported by the organizations whose products are advertised.

5.7.2 Spyware

It is a special type of which is installed in the target computer with or without the user permission and is designed to steal sensitive information from the target machine. Mostly it gathers the browsing habits of the user and the send it to the remote server without the knowledge of the owner of the computer. Most of the time they are downloaded in to the host computer while downloading freeware i.e. free application programmes from the internet. Spywares may be of various types; it can keep track of the cookies of the host computer, it can act as key loggers to sniff the banking passwords and sensitive information, etc.

5.7.3 Browser hijacking software

There are some malicious software which are downloaded along with the free software offered over the internet and installed in the host computer without the knowledge of the user. This software modifies the browsers setting and redirect links to other unintentional sites.

5.7.4 Virus

A virus is a malicious code written to damage/harm the host computer by deleting or appending a file, occupy memory space of the computer by replicating the copy of the code, slow down the performance of the computer, format the host machine, etc. It can be spread via email attachment, pen drives, digital images, e-greeting, audio or video clips, etc. A virus may be present in a computer but it cannot activate itself without the human intervention. Until and unless the executable file (.exe) is executed, a virus cannot be activated in the host machine.

5.7.5 Worms

They are a class of virus which can replicate themselves. They are different from the virus by the fact that they does not require human intervention to travel over the network and spread from the infected machine to the whole network. Worms can spread either through network, using the loopholes of the Operating System or via email. The replication and spreading of the worm over the network consumes the network resources like space and bandwidth and force the network to choke. It is a program that views the infection point as another computer rather than as other executable files on an already infected computer.

5.7.6 Trojan Horse

A Trojan virus is a piece of software designed to look like a useful file or software program but performs a possibly nefarious function once installed on a client computer.

A Trojan virus will normally consist of a server and client component. The client component is the portion of the malware that infects the end-user‘s computer. Once established or executed, the virus can be designed to establish a certain level of control over the infected computer.
It not only damages the host computer by manipulating the data but also it creates a backdoor in the host computer so that it could be controlled by a remote computer.

5.7.7 Scare ware

Internet has changed how we talk, shop, play etc. It has even changed the way how the criminal target the people for ransom. While surfing the Internet, suddenly a pop-up alert appears in the screen which warns the presence of dangerous virus, spywares, etc. in the user‘s computer.

As a remedial measure, the message suggests the used download the full paid version of the software. As the user proceeds to download, a malicious code, known as scare ware is downloaded into the host computer. It holds the host computer hostage until the ransom is paid. The malicious code can neither be uninstalled nor can the computer be used till the ransom is paid.

5.8. SECURITY COUNTERMEASURES

Some of the security countermeasures include the following.

Intrusion detection system (**IDS**) are devices or software applications used to monitor a network or systems for malicious activity or policy violations.
Intrusion prevention systems (**IPS**) are network security appliances that are used to monitor network or system activities for malicious activity.
Antivirus Approaches
The ideal solution to the threat of viruses is prevention: Do not allow a virus to get into the system in the first place, or block the ability of a virus to modify any files containing executable code or macros. This goal is, in general, impossible to achieve, although prevention can reduce the number of successful viral attacks. The next best approach is to be able to do the following:
• **Detection:** Once the infection has occurred, determine that it has occurred and locate the virus.
• **Identification:** Once detection has been achieved, identify the specific virus that has infected a program.
• **Removal:** Once the specific virus has been identified, remove all traces of the virus from the infected program and restore it to its original state. Remove the virus from all infected systems so that the virus cannot spread

further. If detection succeeds but either identification or removal is not possible, then the alternative is to discard the infected file and reload a clean backup version. Advances in virus and antivirus technology go hand in hand. Early viruses were relatively simple code fragments and could be identified and purged with relatively simple antivirus software packages. As the virus arms race has evolved, both viruses and, necessarily, antivirus software have grown more complex and sophisticated.

There are four generations of antivirus software:

- First generation: simple scanners
- Second generation: heuristic scanners
- Third generation: activity traps
- Fourth generation: full-featured protection

A first-generation scanner requires a virus signature to identify a virus.. Such signature- specific scanners are limited to the detection of known viruses. Another type of first-generation scanner maintains a record of the length of programs and looks for changes in length.

A second-generation scanner does not rely on a specific signature. Rather, the scanner uses heuristic rules to search for probable virus infection. One class of such scanners looks for fragments of code that are often associated with viruses.

Third-generation programs are memory-resident programs that identify a virus by its actions rather than its structure in an infected program. Such programs have the advantage that it is not necessary to develop signatures and heuristics for a wide array of viruses. Rather, it is necessary only to identify the small set of actions that indicate an infection is being attempted and then to intervene.

Fourth-generation products are packages consisting of a variety of antivirus techniques used in conjunction. These include scanning and activity trap components. In addition, such a package includes access control capability, which limits the ability of viruses to penetrate a system and then limits the ability of a virus to update files in order to pass on the infection.

5.9 ANTIVIRUS SOFTWARE

Antivirus or anti-virus software, sometimes known as anti-malware software, is computer software used to prevent, detect and remove malicious software. Antivirus software was originally developed to detect and remove computer viruses, hence the name.

However, with the proliferation of other kinds of malware, antivirus software started to provide protection from other computer threats. In particular, modern antivirus software can protect from: malicious Browser Helper Objects (BHOs), browser hijackers, ransomware, key loggers, backdoors, root kits, Trojan horses, worms, malicious LSPs, dialers, fraud tools, adware and spyware. Some products also include protection from

other computer threats, such as infected and malicious URLs, spam, scam and phishing attacks, online identity (privacy), online banking attacks, social engineering techniques, Advanced Persistent Threat (APT), botnets and DDoS attacks.

5.9.1 Identification methods for viruses

One of the few solid theoretical results in the study of computer viruses is Frederick B. Cohen's 1987 demonstration that there is no algorithm that can perfectly detect all possible viruses. However, using different layer of defense, a good detection rate may be achieved. There are several methods which antivirus engine can use to identify malware:

- **Signature-based detection:** is the most common method. To identify viruses and other malware, the antivirus engine compares the contents of a file to its database of known malware signatures.
- **Heuristic-based detection:** is generally used together with signature-based detection. It detects malware based on characteristics typically used in known malware code.
- **Behavioural-based detection:** is similar to heuristic-based detection and used also in Intrusion Detection System. The main difference is that, instead of characteristics hardcoded in the malware code itself, it is based on the behavioral fingerprint of the malware at run-time. Clearly, this technique is able to detect (known or unknown) malware only after they have starting doing their malicious actions.
- **Sandbox detection:** is a particular behavioural-based detection technique that, instead of detecting the behavioural fingerprint at run time, it executes the programs in a virtual environment; logging what actions the program performs. Depending on the actions logged, the antivirus engine can determine if the program is malicious or not. If not, then, the program is executed in the real environment.
- **Data mining techniques:** are one of the latest approach applied in malware detection. Data mining and machine learning algorithms are used to try to classify the behavior of a file (as either malicious or benign) given a series of file features that are extracted from the file itself.

5.10 KEY TERMS AND CONCEPTS

- **Network Security** is the measures to protect data during their transmission.
- **Internet Security** is the measures to protect data during their transmission over a collection of interconnected networks called Internet.
- **Information Security** is about to prevent attacks or failing that to detect attacks on information based schemes/systems.
- **Security service** means a service that enhances the security of the data processing systems and the information transfers of an organization.

The services are intended to counter security attacks and they make use of one or more security mechanisms to provide the service.

- **Antivirus or anti-virus software**, also known as anti-malware software, is computer software used to prevent, detect and remove malicious software from a computer.

5.11 SELF-ASSESMENT QUESTIONS

1. What is the need for security?

   ……
   ……
   ……
   ……
   ……

2. List and briefly define categories of security services?

   ……
   ……
   ……
   ……
   ……

3. What are characteristics of a firewall?

   ……
   ……
   ……
   ……
   ……
   ……

4. What is a malware? Name different types of malwares?

   ……
   ……
   ……
   ……
   ……

5. What is an antivirus? Why it is necessary?

   ……
   ……
   ……
   ……
   ……
   ……

5.12 REFERENCES AND SUGGESTED READINGS

1. William Stallings, "Cryptography and Network Security-Principles and Practices", Prentice-Hall of India.
2. Fundamentals of Information Security (PGDCS-01), Study Materials of Uttarakhand Open University, Haldwani, for Certificate in e-Governance and Cyber Security.
3. Cyber Security Techniques (PGDCS-02), Study Materials of Uttarakhand Open University, Haldwani, for Certificate in e-Governance and Cyber Security.
4. https://en.wikipedia.org/wiki/Intrusion_detection_system
5. https://en.wikipedia.org/wiki/Antivirus_software

5.13 ANSWERS TO SELF-ASSESMENT QUESTIONS (UNIT-1)

1. What are the functions of network layer?

The main functions of network layer include the following.

- Path determination/routing
- Logical addressing.
- Forwarding
- Datagram encapsulation
- Fragmentation and reassembly
- Error handling and diagnostics

2. Write briefly about IP address and its classifications?

- The source and destination address fields in the IP header each contain a 32-bit global internet address, generally consisting of a network identifier and a host identifier.
- Class A: Few networks, each with many hosts.
- Class B: Medium number of networks, each with a medium number of hosts.
- Class C: Many networks, each with a few hosts.

3. What is ICMP? What is its purpose?

ICMP provides a means for transferring massages from routers and other host. ICMP provides feedback about problems in the communication environment. When the router does not have the buffering capacity to forward a datagram, and when the router can direct the station to send traffic on a shorter route. ICMP message is sent in response to a datagram either by a router along the datagram's path or by the intended destination host.

4. What is subnet mask? Write with an example.

- Subnet is used to identify individual LAN's and MAN's.
- It allows a network to be split into several parts for internal use but still act like a single network to the outside world.
- To implement subnetting, the router needs a subnet mask that indicates the split between network + subnet number and host.
- Subnet mask is used to bit positions containing extended network number are indicated.
- Ex. 255.255.252.0/22.A‖/22‖ to indicate that the subnet mask is 22 bits long.

5. What is Internet Protocol? What are its services?

The internet protocol is part of the TCP/IP suite and is the most widely used as internetworking protocol. . It is an unreliable and connectionless protocol. It provides a best-effort delivery service. The term *best effort* means that IP provides no error checking or tracking.

IP transports data in packet called *datagram*, each of which is transported separately. Data grams can travel along different routes and can arrive out of sequence or be duplicated. IP does not keep track of the routes and has no facility for reordering datagram once they arrive at their destination. IP provides bare-bones transmission functions that free the user to add only those facilities necessary for a given application and thereby allows for maximum efficiency.

5.13 ANSWERS TO SELF ASSESMENT QUESTIONS (UNIT-2)

1. What is the need for transport layer?

The Transport Layer is responsible for delivering data to the appropriate application process on the host computers. This involves statistical multiplexing of data from different application processes, i.e. forming data packets, and adding source and destination port numbers in the header of each Transport Layer data packet. The Transport Layer is responsible for end-to-end data transport. Primary functions include the following:

- Provision of connection oriented or connectionless service.
- Segmentation and reassembling data.
- Setup and release of connections across the network.

2. Write down all the transport layer services.

Transport layer services include the following.

- End-to-end-delivery
- Service point addressing
- Reliable delivery
- Flow control
- Connection management
- Upward and downward Multiplexing
- Congestion Control and
- Quality of Services (QoS).

3. What are the advantages of UDP?

The main advantages of UDP are as follows.

- UDP provides connectionless service.
- No connection establishment delay.
- Provide unreliable, but fast service.

4. How do transport services differ from the data link layer services?

- The data link layer services are at node to node level. But the transport layer services are end to end level. Both the layers are having the flow control and error control mechanisms.
- The data link layer offers at node to node level. But the transport layer offers at end to end level.
- Data link layer is responsible for node to node delivery of the frames while transport layer is responsible for end to end delivery of the entire message.

5. Which is reliable - TCP or UDP? Why?

A reliable protocol ensures that data sent from one machine to another will eventually be communicated correctly. TCP is a reliable one. Because, it provides reliable transport between sending and receiving processes. However it does not guarantee that this data will be transmitted correctly within any particular amount of time -- just that given enough time, it will arrive.

6. What is multiplexing? Differentiate between upward and downward multiplexing?

Transport layer performs multiplexing/ de multiplexing function. Multiple applications employ same transport protocol, but use different port number. According to lower layer n/w protocol, it does upward multiplexing or downward multiplexing.

Downward multiplexing

- Multiple users employ same transport protocol.
- User identified by port number or service access point.

Upward multiplexing

- It may also multiplex with respect to network services used.

7. Define Congestion Control?

It involves preventing too much data from being injected into the network, thereby causing switches or links to become overloaded. Thus flow control is an end to an end issue, while congestion control is concerned with how hosts and networks interact.

8. What is meant by quality of service?

The quality of service defines a set of attributes related to the performance of the connection. The need of each connection can be characterized by four primary parameters:

i) **Reliability**: Reliability is the ability of a system to perform and maintain its functions in normal conditions as well as under unexpected conditions.

ii) **Delay** is defined as the time interval elapsed between the departures of data from the source to its arrival at the destination.

iii) **Jitter:** Jitter refers to the variation in time between packets arriving at the destination.

iv) **Bandwidth** refers to the data rate supported by a network connection or interface.

9. What are the three events involved in the connection?

For security, the transport layer may create a connection between the two end ports. A connection is a single logical path between the source and destination that is associated with all packets in a message.

Creating a connection involves three steps:

- Connection establishment
- Data transfer
- Connection release

10. Differentiate between TCP and UDP.

Major differences between the TCP (Transmission Control) & UDP (User Datagram) transport layer protocols:

TCP	UDP
Stream-Oriented : Data transmitted as a virtual stream of bytes	Message-Oriented : Data transmitted as individual data packets called datagram (the basic unit of data transferred)
Data Reliability: Absolute guarantee that data remains intact during transit.	No guarantee; No data reliability
Two-way connection and reliable	One-way connection reliable
Uses a 20 bytes header	Uses a 8 bytes header
Error checking	No error checking
Flow control	No flow control
Slower speed (due to the above reasons)	Faster than TCP
Used when speed is not critical	Used in games and applications where speed(time) is critical
Examples of protocols those use TCP include SMTP, HTTP, FTP, Telnet, etc	Examples of protocols using UDP include DNS, TFTP, RIP, VoIP, SNMP, etc

5.13 ANSWERS TO SELF-ASSESMENT QUESTIONS (UNIT-3)

1. Differentiate between HTTP and SMTP

- Data transfer between client and server by HTTP look like SMTP messages
- In HTTP message can be sent from the client to the server and server to the client.
- SMTP message are not read by humans but the HTTP messages are read and interpreted by the HTTP server and HTTP client.
- HTTP messages are delivered immediately and SMTP messages are stored and forwarded.

2. Compare the protocols HTTP and FTP

HTTP	FTP
It uses only one TCP connection; data connection.	It uses two connections; data and control connection.
Any short of file can be transferred	Only text file can be transferred
Port no 80 is used	It uses two ports 21 and 20
Two types of messages are used as request and response messages.	It used control information and data exchange between the client and the server.

3. What are the two agents in the architecture of e-mail?

E-mail system consists of two subsystems.

(i) **The user agents:** Allows people to read and send mail. i.e. Local programs that provide user interface

(ii) **The message transfer agent:** moves the messages from the source to the destination .These are system domains.

4. Write the basic functions of e –mail systems.

The main functions of e –mail systems include the following.

- **Composition :** process of creating and sending messages
- **Transfer:**- moves messages from sender to receiver.
- **Reporting: -** Informs the sender whether the message is delivered or not.
- **Displaying:** Incoming messages are needed to be displayed and read by the people.
- **Disposition:** The final step that recipient does with the message after receiving it.

5. What are the additional services provided by e-mail?

Additional services of e-mail are as follows.

- The E-mail system allows mailbox to store the incoming e-mail messages
- The provision of mail list to store e-mail addresses which can be used for sending messages in group
- The provision of carbon copies, blind carbon copies high priority e-mail.
- Widely used in industry for contemporary communication.
- Widely used as a best person –to –person communication.

5.13 ANSWERS TO SELF-ASSESMENT QUESTIONS (UNIT-4)

1. What is the Internet?

The Internet, sometimes called simply "the Net," is a worldwide system of computer networks - a network of networks in which users at any one computer can, if they have permission, get information from any other computer (and sometimes talk directly to users at other computers).

The early Internet was used by computer experts, engineers, scientists, and librarians. There was nothing friendly about it. But now-a day's Internet is one of the basic need for almost every one for many purposes.

2. Write a brief notes on WWW.

The three Ws that are in the addresses of the websites we access. The World Wide Web or simply the web is a way to access information through the internet. The web is a model for sharing information that is built on the internet. The protocol used by the web is HTTP, just one of the many ways that information can be sent through the internet.

If a service uses HTTP to enable applications to communicate with each other, this is a web service. Web browsers, such as Chrome or Firefox, enable us to access web documents that we mainly know as web pages or websites. These sites are connected to each other through hyperlinks as if they were on a spider‘s web (hence the name), and all this thanks to the transfer protocol: HTTP.

Therefore, the web is only one of the ways that information can flow through the internet: it is just a portion, and although it is very large and the most popular part, it does not include the whole of the internet

3. **What are Domains?**

Domains divide World Wide Web sites into categories based on the nature of their owner, and they form part of a site's address, or uniform resource locator (URL).

Common top-level domains are as follows.

.**com** A commercial enterprises

.mil: A military site

.org: An organization site (non-profits, etc.)

.int: A organizations established by international treaty

.**net:** A network

.biz A commercial and personal

.edu: A educational site (universities, schools, etc.)

.info: A commercial and personal

.gov: A government organizations

4. **What is a Web Browser?**

A Web browser contains the basic software you need in order to find, retrieve, view, and send information over the Internet. This includes software that lets you:

- Send and receive electronic-mail (or e-mail) messages worldwide nearly instantaneously.
- Read messages from newsgroups (or forums) about thousands of topics in which users share information and opinions.
- Browse the World Wide Web (or Web) where you can find a rich variety of text, graphics, and interactive information.

The most popular browsers are Microsoft Internet Explorer and Netscape Navigator, Google Chrome, Mozilla Firefox etc.

5. **What is a Search engine?**

A search engine is a searchable database of Internet files which allows the user to enter keywords relating to particular topic and retrieve information about Internet sites containing those keywords. Some of the well-known search engines are www.google.com, www.hotbot.com, www.lycos.com, and www.altavista.com.

5.13 ANSWER TO SELF-ASSESSMENT QUESTIONS (UNIT-5)

1. **What is the need for security?**

   We need security for the following reasons.

   - To isolate data from getting hacked, corrupted, unauthorized access, unauthorized modified, unauthorized deleted etc
   - To maintain confidentiality, availability and integrity of data
   - To prevent electronic mail from getting hacked and unauthorized access
   - To protect easy passwords and pins being cracked
   - To eradicate vulnerabilities (weakness) in the system or data

2. **List and briefly define categories of security services?**

   There are five security services: *data confidentiality*, *data integrity*, *authentication*, *on-repudiation*, and *access control*.

   - *Data confidentiality* is to protect data from disclosure attack.
   - *Data integrity* is to protect data from modification, insertion, deletion, and replaying.
   - *Authentication* means to identify and authenticate the party at the other end of the line.
   - *Non-repudiation* protects against repudiation by either the sender or the receiver of the data.
   - *Access control* provides protection against unauthorized access to data.

3. **What are the characteristics of Firewall?**

   All traffic from inside to outside, and outside to inside of the network must pass through the firewall. This is achieved by physically blocking all access to the local network except via the firewall.

   Various configurations are possible. Only authorized traffic, as defined by the local security policy, will be allowed to pass. Various types of firewalls are used, which implement various types of security policies. The firewall itself is immune to penetration. This implies that use of a trusted system with a secure operating system. This implies that use of a trusted system with a secure operating system.

   Four techniques that firewall use to control access and enforce the site's security policy is as follows:

   (i) **Service control** – determines the type of internet services that can be accessed, inbound or outbound. The firewall may filter traffic on this basis of IP address and TCP port number; may provide proxy software that receives and interprets each service

request before passing it on; or may host the server software itself, such as web or mail service.

(ii) **Direction control** – determines the direction in which particular service request may be initiated and allowed to flow through the firewall.

(iii) **User control** – controls access to a service according to which user is attempting to access it.

(iv) **Behaviour control** – controls how particular services are used.

4. What is a malware? Name different types of malwares?

Malware stands for **—***Malicious Software* and it is designed to gain access or installed into the computer without the consent of the user.

The most common types of malware; adware, bots, bugs, root kits, spyware, Trojan horses, viruses, and worms.

5. What is an antivirus? Why it is necessary?

Antivirus or anti-virus software is used to prevent, detect, and remove malware, including but not limited to computer viruses, computer worms, Trojan horses, spyware and adware.

However, it is possible that a computer may be infected with new malware for which no signature is yet known.

An antivirus is necessary because it protects our computer from malicious attacks and prevents data loss by malicious acts, such as deleting files, accessing personal data, or using your computer to attack other computers.