UNIT - 1

1

INTERNET AND WORLD WIDE WEB

Unit Structure :

1.0 OBJECTIVE:

After reading through this chapter, you will be able to –

- Understand concept of Internet and World Wide Web, their applications.
- List the services provided by Internet Service providers with examples.
- Define domain name server and list various domains.
- Understand the concept of Internet address.
- Understand the function of a URL and web browsers.

- Use different web browsers.
- Use the search engines to search for required information over the internet.
- Understand the need and use of a web server and proxy server.

1.1 INTRODUCTION TO INTERNET AND ITS APPLICATIONS

Internet is ---

- A computer network consisting of a worldwide network of computers that use the TCP/IP network protocols to facilitate data transmission and exchange.
- The Internet is a global system of interconnected computer networks that use the standard Internet Protocol Suite (TCP/IP) to serve billions of users worldwide.
- The term Internet actually refers to the combined collection of academic, commercial, and government networks connected over international telecommunication backbones and routed using IP addressing.

The internet has gained popularity rapidly as it is used for various purposes. Few of the main applications of internet are listed below –

Applications of Internet ---

1.1.1 E mail (Electronic mail)

- Electronic mail (also known as email or e-mail) is one of the most commonly used services on the Internet, allowing people to send messages to one or more recipients.
- Email was invented by Ray Tomlinson in1972.
- Electronic mail is a method of exchanging digital messages from an author to one or more recipients.
- Modern email operates across the Internet or other computer networks. Today's email systems are based on a store-and-forward model.
- Email servers accept, forward, deliver and store messages. Neither the users nor their computers are required to be online simultaneously; they need connect only briefly, typically to an email server, for as long as it takes to send or receive messages.
- An email message consists of three components, the message *envelope*, the message *header*, and the message *body*.
- Header contains information about who sent the message, the recipient(s) and the route.
- Header also usually contains descriptive information, such as a subject header field and a message submission date/timestamp.

- Email message body contains text (7bit ASCII) as well as multimedia messages. These processes are declared in Multipurpose Internet Mail Extensions (MIME). MIME is set of RFCs (Request for Comment)
- Network based emails are exchanged over the internet using the SMTP (Simple Mail Transfer protocol).
- In the process of transporting email messages between systems, SMTP communicates delivery parameters using a message *envelope* separate from the message (header and body)itself.
- Email addresses (both for senders and recipients) are two strings separated by the character "@" (the "at sign"): such as user@domain
- The right-hand part describes the domain name involved, and the left-hand part refers to the user who belongs to that domain.
- An email address can be up to 255 characters long and can include the following characters:
  - Lowercase letters from a toz;
  - Digits
  - The characters ".","_" and "-" (full stop, underscore, and hyphen)

In practice, an email address often looks something like this:
fname.lname@provider.domain

1.1.2 Telnet

- Telnet is a network protocol used in any network (internet or LAN) for bidirectional text oriented communication.
- telnet standard was defined in 1973, before which it was considered as a dhocprotocol.
- Original purpose of the telnet protocol was to login to the remote computers on the network.
- Telnet protocol uses 'virtual terminal' to connect to the remote hosts.
- Virtual terminal is a application service that allows host in a multi terminal network to communicate with other hosts irrespective of terminal type or characteristics.
- Telnet uses the TCP protocol for transmission of 8 byte data.
- Most network equipment and operating systems with a TCP/IP stack support a Telnet service for remote configuration (including systems based on Windows NT)
- The term telnet may also refer to the software that implements the client part of the protocol.
- telnet is a client server protocol, which is based upon reliable connection oriented communication transport and basic use of telnet is to make a connection to the TCP protocol.

- Data transferred over telnet is vulnerable as telnet does not use any encryption technique to mask or protect the data.
- Most implementations of Telnet have no authentication that would ensure communication is carried out between the two desired hosts and not intercepted in the middle.
- Commonly used Telnet daemons have several vulnerabilities discovered over the years.
- Extensions to the Telnet protocol provide Transport Layer Security (TLS) security and Simple Authentication and Security Layer (SASL) authentication that address the above issues.
- Few applications of telnet include the 'putty' TCP client which can access a linux server using windows operating system, Absolute telnet (windows client) and RUMBA (terminal emulator).

1.1.3 FTP

- File transfer protocol is a simple and standard network protocols that transfers a file from one host to the other over a TCP network.
- Based on client server architecture.
- Utilizes separate control and data connection for client and server to transmit and receive file(s) over the network.
- It is an application protocol that uses the internet's TCP/IP suite.
- Mainly used to transfer the web pages or related data from the source or creator to a host that acts as a server to make the page or file available to other hosts (uploading) or downloading programs and other files from server to a host.
- FTP protocol can perform over a active or passive connection.
- When a connection is made from the client to server, it is called as control connection and it remains open for duration of session. This connection is responsible for establishing connectivity between client and server.
- Other connection opened by client (passive) or server (active) is called data connection and is used to transfer the data.
- As separate ports are used by client and server for these connections, FTP becomes an put of band protocol.
- Data transfer can take place in following three modes
    - Stream mode : data is sent in a continuous stream where FTP does not do any formatting.
    - Block mode: FTP breaks the data into several blocks (block header, byte count, and data field) and then passes it on to TCP.
    - Compressed mode: Data is compressed using a single algorithm.
- FTP is a old protocol and is basically low in security aspect. Data transferred over FTP is not encrypted and is in clear text format.

Hence the data like usernames, passwords can be read by anyone who can capture the FTPed package. Newer versions of the protocol, however, apply secure shell protocol (SSH) and avoid all the problems faced by FTP.

- Following are few types of FTP protocol with additional features
- Anonymous FTP : Users login using an 'anonymous' account to protect their confidential data.
- Remote FTP: FTP commands executed on a remote FTP server
- FTP with web browser and firewall support.
- Secure FTP (SFTP,FTPS)

1.1.4 ECommerce

- Electronic commerce can be defined as use of electronic communications, particularly via the internet, to facilitate the purchase/ sale of goods and services. E-commerce includes all forms of electronic trading including electronic data interchange (EDI), electronic banking, electronic mail and other online services.
- E transactions are of two categories. – virtual products like policies and actual retail products.
- Most of e transactions of actual products involve physical transportation of goods which are purchased over the electronic media.
- Online retailing has gained a name of Etailing.
- Electronic commerce is generally considered to be the sales aspect of e-business. It also consists of the exchange of data to facilitate the financing and payment aspects of the business transactions.
- Originally, electronic commerce was identified as the facilitation of commercial transactions electronically, using technology such as Electronic Data Interchange (EDI) and Electronic Funds Transfer (EFT).
- Other forms of e commerce were established with the growth and use of credit cards, and air line reservation system going online.
- Electronic commerce of the modern era (post 1990) includes technologies like enterprise resource planning (ERP), data warehousing and data mining.
- The electronic transactions between two businesses like dealer and wholesaler or wholesaler and retailer come in the B2B (business to business) E commerce category.
- Other popular E commerce categories would be business to consumer (B2C) and business to government(B2G)
- Volume of B2B transaction is much higher as compared to the volume of B2C transactions. Reason for this is, many transactions at B2B level lead to finished good and this leads to just one B2C transaction.

- In an example, if a customer buys a product, say a pen, that would be a B2C transaction. But the transaction leading this one, including purchase of plastic, ink, refill, moulds etc would be B2B transaction. Also the sale of the pen to the retailer by the manufacturing company, like cello, is B2Btransaction.
- OtherformofB2Ctransactionsarebusinesstoindividual,where the record of an individual's transaction is maintained.
- C2C is consumer to consumer, or citizen to citizen E commerce. Here customers can perform transaction via a third party. Like a product can be posted on amazon.com and will be sold to another consumer through Amazon.
- C2B E commerce model is reverse of traditional business to consumer approach. This can be explained by a internet blog or a social networking site where author can have a link in his blog article to online sale of a product (promoting the business). This has become possible due to advancements in technology and reduced costs of technology.
- Unique attribute of e commerce is negotiation facility and its immediate results. Also, in E commerce transactions, integration of transactions is automated.

1.1.5 Video Conferencing

- Video conferencing or video teleconference is a set of telecommunication technologies which allow one or more locations to transmit and receive video and audio signals simultaneously.
- This is known as visual collaboration.
- Simple analog video conferencing is achieved by two closed circuit television systems connected with coaxial cables or radio waves.
- This type of communication was established from1968.
- Modern video conferencing is IP based and through more efficient video compression technologies, allowing desktop or PC based video conferencing.
- Video telephony is now popular due to free internet services.
- Core technology used for this is compression of audio and video signals. Hardware and software used for this task is called as codec (coder/ decoder). Compression rate achieved is almost 1:500. The resultant stream of binary data is sent in packet form through digital networks.
- The components required for a videoconferencing system include:
  - Video input : video camera or webcam
  - Video output: computer monitor , television or projector
  - Audio input: microphones, CD/DVD player, cassette player, or any other source of Pre Amp audio outlet.
  - Audio output: usually loudspeakers associated with the display device or telephone
  - Data transfer: analog or digital telephone network, LAN or Internet

- There are basically two types of videoconferencing systems.
  - Dedicated systems: all required components (i.e. software and hardware based codec, control computer and video camera, electrical interfaces) packed in a single console application. They include large group, small group, portable and non portable video conferencing systems.
  - Desktop Systems: add ons to normal computing systems transforming these systems to videoconferencing devices.
- There are following layers in the videoconferencing technology–
  - User interface
  - Conference control
  - Control or signal plane
  - Media plane.
- Videoconferencing has following modes
  - Voice activated switch
  - Continuous presense
- Problems faced by videoconferencing
  - Echo: echo is defined as reflected source wave interference with new wave created by the source. i.e. signal coming out from the source interferes with newly coming source and generating unwanted input signal. This may result into remote party receiving their own sounds again. This can be avoided by using an algorithm called as AEC (Acoustic Echo cancellation).
  - Eye contact : in videoconferencing, due to time delays and parallax, communicators have a feel of the other party avoiding eye contact and can result into issues in professional communication. This can be avoided by using special stereo reconstruction technique.
  - Signal latency: The information transport of digital signals in many steps needs time. In a telecommunicated conversation, an increased latency larger than about 150-300ms becomes noticeable and is soon observed as unnatural and distracting. Therefore, next to a stable large bandwidth, a small total round-trip time is another major technical requirement for the communication channel for interactive videoconferencing.

1.1.6 E business

- E business is conduct of business over the internet, which includes buying and selling of goods and even services.
- In other words it is application of information and communication technologies in support of all activities in business.
- Applications of E business are divided into following categories–

- Internal business systems--
- Customer Relationship Management(CRM)
- Enterprise Resource Planning(ERP)
- Human resource management system(HRMS)
- Enterprise Communication and collaboration
- Content management
- E- mails
- Voicemails
- Web conferencing
- Electronic commerce
- B2B (business to business)
- B2C (business to customer)
- B2E business-to-employee
- B2Gbusiness-to-government
- G2Bgovernment-to-business
- G2G (government-to-government)
- G2C (government-to-citizen)
- C2C (consumer-to-consumer)
- C2B(consumer-to-business)

- A business model is defined as the organization of product, service and information flows, and the source of revenues and benefits for suppliers and customers. The concept of e-business model is the same but used in the online presence.
- Few e business models are–
  - E-shops
  - E-commerce
  - E-procurement
  - E-malls
  - E-auctions
  - Virtual Communities
- E business has more security risks as compared to a regular business, as E business has many more users at a time. Keeping the large information confidential is a difficult task. Also, data integrity, authenticity and storage of data are some challenges faced by Ebusiness.
- Some methods to provide security are physical security as well as encryption in data storage, transmission, antivirus software and firewalls. Digital signature is another way to confirm the ownership of a document.

1.2 INTERNET SERVICE PROVIDER

- An Internet service provider (ISP) is a company that provides access to the Internet, hosts data, or does both. ISP is also known as IAP (internet access provider) Access ISPs connect customers to the Internet using copper, wireless or fibre connections. Hosting ISPs lease server space for smaller businesses and host other people servers (colocation). Transit ISPs provide large tubes for connecting hosting ISPs to access ISPs.

- As internet gained popularity, it was essential to provide internet access to many people or many hosts. Due to the increasing demand to access internet, commercial ISP came into existence in1990.

Technologies used –

For users and small business applications -
- Dial up connection
- DSL (digital subscriber line)
- Broadband wireless connection
- Cable modem
- Fibre optical connection

For medium to large businesses or for other ISPs,
- DSL
- Ethernet
- Metro Ethernet
- Gigabyte Ethernet
- Frame relay
- Satellite Ethernet

- ***ISP connections–***

ISPs which provide connections through phone lines like dial ups, do not seek any information about the caller's (user of the internet) physical location or address. So, caller from any location which is in reach of the ISP, can use the services provided.

Other way of getting connected through ISP is using cable or any other lines. Here, fixed registration of the user at the ISP side is essential.

- ***Services provided –***

ISP host usually provide e mail, FTP and web hosting services. Other services can be like virtual machines, clouds or entire physical servers where clients can run their own softwares.

ISPs often take services from their upstream ISPs. i.e. they work in hierarchy. The ISPs are divided into three categories –

- **Peering :** ISPs taking services from upstream and getting connected to each other to exchange data and to control network traffic through peering points, orinternetex change points. These points help save one

more upstream ISP and cut down on the cost. The ISPs which do not need upstream ISP are called Tier 1 ISP.

 Virtual ISP (VISP) : this is an ISP which purchases services from other ISP and gives them to the end user, without any set up of its own.

 Free ISP: these are ISPs which provide services free of cost to the users and display advertisements till the users are connected. These are called as freenets. These are normally run on no profit basis.

1.3 DOMAIN NAME SERVERS

- ***Domain Name System is–***

DNS is part of a domain name system. It is hierarchical naming system built on a distributed database for resource connected to the internet or a private network. The main purpose of this system is to translate domain names meaningful to humans into names or rather numeric streams which help the corresponding network devices to identify the resource or domain. Domain name system makes it possible to give or allot names to domains or group of networks irrespective of their physical locations.

- ***Domain Name Server is–***

 Domain name system assigns domain name servers for allotting names and mapping these names to IP addresses. In other words domain name servers are nodes of the domain name system which acts like a client server system. Each domain has at least one authoritative DNS server that publishes information about that domain and the name servers of any domains subordinate to it. The top of the hierarchy is served by the root name servers, the servers to query when looking up (resolving) a TLD (Top level domain). Authoritative DNS can either be a master or a slave. Master DNS keeps record of all zone records. Slave DNS uses a automatic update mechanism to maintain copies of records of its master. Every top level domain requires a primary DNS and at least one secondary DNS. Every DNS query must start with recursive queries at the root zone, for authoritative DNS.

 To improve efficiency, reduce DNS traffic across the Internet, and increase performance in end-user applications, the Domain Name System supports DNS cache servers which store DNS query results for a period of time determined in the configuration (time-to-live) of the domain name record in question.

 The client-side of the DNS is called a DNS resolver. It is responsible for initiating and sequencing the queries that ultimately lead to a full resolution (translation) of the resource sought, e.g., translation of a domain name into an IP address.

 A DNS query may be either a non-recursive query or a recursive query

- The resolver, or another DNS server acting recursively on behalf of the resolver, egotiates use of recursive service using bits in the query headers.
- Resolving usually entails iterating through several name servers to find the needed information. However, some resolvers function simplistically and can communicate only with a single name server. These simple resolvers (called "stub resolvers") rely on a recursive name server to perform the work of finding information for them.
- ***Operation of DNS–***
- Domain name resolvers determine the appropriate domain name servers responsible for the domain name in question by a sequence of queries starting with the right-most (top-level) domain label.
- DNS recorsor consults three name servers to resolve one address. The process is as follows–
  - A network host is configured with an initial cache (so called *hints*) of the known addresses of the root name servers. Such a *hint file* is updated periodically by an administrator from a reliable source.
  - A query to one of the root servers to find the server authoritative for the top-level domain.
  - A query to the obtained TLD server for the address of a DNS server authoritative for the second-level domain.
  - Repetition of the previous step to process each domain name label in sequence, until the final step which returns the IP address of the host sought.

1.4 INTERNET ADDRESS

- Internet address follows the TCP/IP suite hence, it is also known as the IP address.
- Internet address has a job of identifying a node on the network. In other words, it is a numeric lable attached to every system (computer or any other device). The basic function of IP address are two–
- Identification of computer or node or device and location addressing.
- The designers of the Internet Protocol defined an IP address as a 32-bit number[1] and this system, known as Internet Protocol Version 4 (IPv4), is still in use today. However, due to the enormous growth of the Internet and the predicted depletion of available addresses, a new addressing system (IPv6), using 128 bits for the address, was developed in 1995,[3] standardized as RFC 2460 in 1998,[4] and is being deployed worldwide since themid-2000s.
- IP addresses are binary numbers, but they are usually stored in text files and displayed in human-readable notations, such as 172.16.254.1 (forIPv4)

- IPV4 address is a 32 bit number, which uses the decimal doted notation consisting of 4 decimal numbers each ranging from 0 to 255 separated by dots. Network administration divides the IP address into two parts. – the most significant 8 bits are called network address portion the remaining bits are known as rest bits or host bits or identifiers and they are used for host numbering in anetwork.
- Although IPV4 provides 4.3 billion addresses, they are exhausted due to high demand and as a result, insufficient addresses available with IANA (Internet assigned numbers authority). The primary address pool of IANA is expected to get exhausted by mid 2011. To permanently address the problem, new version of IP i.e. IPV6 was brought forward, this version moved the size of IP address from 32 bit to 128 bits.
- Both IPV4 as well as IPV6 have reserved addresses for private or internal networks. This is termed as private addressing.
- Both IPV4 and IPV6 have sub netting effect. That mean, IP networks can be divided into smaller groups or subnets. IP addresses two constituents that is network address and host identifier or interface identifier are used for this purpose.
- 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*

1.5 WORLD WIDE WEB AND ITS EVOLUTION

- The World Wide Web, abbreviated as WWW or W3 and commonly known as the 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.
- The World-Wide Web was developed to be a pool of human knowledge, and human culture, which would allow collaborators in remote sites to share their ideas and all aspects of a common project.

Evolution of WWW

- In March 1989, Tim Berners-Lee wrote a proposal that referenced ENQUIRE, a database and software project he had built in 1980, and described a more elaborate information management system.
- on November 12, 1990, with help from Robert Cailliau, Tim Berners- Lee published a more formal proposal to build a "Hypertext project" called "World Wide Web" (one word, also "W3") as a "web" of "hypertext documents" to be viewed by "browsers" using a client– server architecture.

- This proposal estimated that a read-only web would be developed within three months and that it would take six months to achieve "the creation of new links and new material by readers, to achieve universal authorshipl" as well as "the automatic notification of a reader when new material of interest to him/her has become available."
- A NeXT Computer was used by Berners-Lee as the world's first web server and also to write the first web browser, World Wide Web, in 1990.
- Tools needed were a working Web the first web browser (which was a web editor as well); the first web server; and the first web pages, which described the project itself.
- On August 6, 1991, Tim Berners-Lee posted a short summary of the World Wide Web project on the alt. hypertext newsgroup.
- This date also marked the debut of the Web as a publicly available service on the Internet. The first photo on the web was uploaded by Berners-Lee in 1992, an image of the CERN house band Les Horribles Cernettes.
- The first server outside Europe was set up at SLAC to host the SPIRES-HEP database in 91 –92.
- The concept of hypertext originated with older projects from the 1960s, such as the Hypertext Editing System (HES) at Brown University by Ted Nelson and Douglas Engelbart.
- Tim Berners Lee introduced the concept of the Universal Document Identifier (UDI), later known as Uniform Resource Locator (URL) and Uniform Resource Identifier (URI); the publishing language Hyper Text Markup Language (HTML); and the Hypertext Transfer Protocol(HTTP).
- In 1993, a graphical browser was developed by a team at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign (NCSA-UIUC), led by Marc Andreessen. This was the first web browserever.

1.6 URL

- Uniform Resource Locator (URL) is a Uniform Resource Identifier (URI) that specifies where an identified resource is available and the mechanism for retrieving it.
- An example of the use of URLs is the addresses of web pages on the World Wide Web, such as*http://www.example.com/.*
- The format is based on Unix file path syntax, where forward slashes are used to separate directory or folder and file or resource names.
- Conventions already existed where server names could be prepended to complete file paths, preceded by adouble-slash

- Every URL consists of some of the following:
  - The scheme name (commonly called protocol), followed by a colon. The scheme name defines the namespace, purpose, and the syntax of the remaining part of the URL.
  - Domain Name depending upon scheme(alternatively, IP address). The domain name or IP address gives the destination location for the URL.
  - An optional port number; if omitted, the default for the scheme is used
  - Path of the resource to be fetched or the program to be run. The path is used to specify and perhaps find the resource requested. It may be case-sensitive for non window based servers. Eg: *http://www.mudlle.ac.in/news.html*
  - A query string for scripts The query string contains data to be passed to software running on the server. It may contain name/value pairs separated by ampersands, for example *? first_ name= John & last_name=Doe.*
  - Optional fragment identifier that specifies a part or a position within the overall resource or document. When used with HTTP, it usually specifies a section or location within the page, and the browser may scroll to display that part of the page.
- When resources contain references to other resources, they can use relative links to define the location of the second resource.
- relative URLs are dependent on the original URL containing a hierarchical structure against which the relative link is based.
- the ftp, http, and file URL schemes are examples of some that can be considered hierarchical, with the components of the hierarchy being separated by"/"
- A URL is a URI that, "in addition to identifying a resource, provides a means of locating the resource by describing its primary access mechanism.

1.7 BROWSERS

- A web browser or Internet browser is a software application for retrieving, presenting, and traversing information resources on the World Wide Web.
- Web browsers can also be used to access information provided by Web servers in private networks or files in file systems. Some browsers can also be used to save information resources to file systems.
- Primary function of a browser is to identify the URI and brings the information resource to user.

- This process begins when user inputs the URI in the browser. Prefix of the URI describes how to interpret the URI. Most URIs have resource retrieved over Hyper text Transfer Protocol. Some web browsers also support prefixes like FTP.
- Once this is done, the HTML script is passed to the browser's layout engine. To make the script interactive java script support is needed. With this, browser can interpret text, images, video and interactive scripts.
- All major browsers allow users to access multiple information resources at the same time in different windows or in tabs. Major browsers include pop up blockers to prevent windows to open without users consent.
- Most major web browsers have these user interface elements in common: Back and forward buttons to go back to the previous resource and forward again.
    - A history list, showing resources previously visited in a list (typically, the list is not visible all the time and has to be summoned)
    - A refresh or reload button to reload the current resource.
    - A stop button to cancel loading the resource. In some browsers, the stop button is merged with the reload button.
    - A home button to return to the user's homepage
    - An address bar to input the Uniform Resource Identifier (URI) of the desired resource and display it.
    - A search bar to input terms into a search engine
    - A status bar to display progress in loading the resource and also the URI of links when the cursor hovers over them, and page zooming capability.

- The usage share of web browsers is as shown below. (Source: Median values)

- Internet Explorer(43.55%)
- Mozilla Firefox (29.0%; Usage by version number)
- Google Chrome (13.89%)
- Safari(6.18%)
- Opera(2.74%)
- Mobile browsers (4.45%)
- Some special web browsers are listed below–

1.7.1 INTERNET EXPLORER

- Windows Internet Explorer (formerly Microsoft Internet Explorer, commonly abbreviated IE or MSIE) is a series of graphical web browsers developed by Microsoft and included as part of the Microsoft Windows line of operating systems starting in1995.

- It was first released as part of the add-on package Plus! for Windows 95 that year. Later versions were available as free downloads, or in service packs. It was part of later versions of windows operating systems.
- The latest stable release is Internet Explorer 9, which is available as a free update for Windows 7, Windows Vista, Windows Server 2008 and Windows Server 2008R2.
- Internet Explorer uses a componentized architecture built on the Component Object Model (COM) technology. It consists of several major components, each of which is contained in a separate Dynamic- link library (DLL) and exposes a set of COM programming interfaces hosted by the Internet Explorer main executable, 'iexplore.exe'
- Internet Explorer uses a zone-based security framework that groups sites based on certain conditions, including whether it is an Internet- or intranet-based site as well as a user-editable white list. Security restrictions are applied per zone; all the sites in a zone are subject to the restrictions.

1.7.2 NETSCAPE NAVIGATOR

- Netscape Navigator is a proprietary web browser that was popular in the 1990s. It was the most popular web browser till 2002, after which competitor browsers have taken over the market of Netscape.
- Netscape Navigator was based on the Mosaic web browser.
- Netscape announced in its first press release (October 13, 1994) that it would make Navigator available without charge to all non-commercial users, and Beta versions of version 1.0 and 1.1 were indeed freely downloadable in November 1994 and March 1995, with the full version 1.0 available in December1994.
- The first few releases of the product were made available in "commercial" and "evaluation" versions.
- During development, the Netscape browser was known by the code name Mozilla. Mozilla is now a generic name for matters related to the open source successor to Netscape Communicator.

1.7.3 OPERA

- Opera is a web browser and Internet suite developed by Opera Software. The browser handles common Internet-related tasks such as displaying web sites, sending and receiving e-mail messages, managing contacts, chatting on IRC, downloading files via Bit Torrent, and reading webfeeds.
- Opera is offered free of charge for personal computers and mobile phones. This is the most popular mobile phone browser and is not packages in desktop operating system.
- Features include tabbed browsing, page zooming, mouse gestures, and an integrated download manager. Its security features include built-in

phishing and malware protection, strong encryption when browsing secure websites, and the ability to easily delete private data such as HTTP cookies.

- Opera runs on a variety of personal computer operating systems, including Microsoft Windows, Mac OS X, Linux, and FreeBSD
- Opera includes built-in tabbed browsing, ad blocking, fraud protection, a download manager and Bit Torrent client, a search bar, and a web feed aggregator. Opera also comes with an e-mail client called Opera Mail and an IRC chat client builtin.
- Opera has several security features visible to the end user. One is the option to delete private data, such as HTTP cookies, the browsing history, and the cache, with the click of a button. This lets users erase personal data after browsing from a shared computer.
- Opera Mobile is an edition of Opera designed for smart phones and personal digital assistants(PDAs)

1.7.4 MOZILLA FIREFOX

- Mozilla Firefox is a free and open source web browser descended from the Mozilla Application Suite and managed by Mozilla Corporation. As of February 2011[update], Firefox is the second most widely used browser with approximately 30% of worldwide usage share of web browsers.
- To display web pages, Firefox uses the Gecko layout engine, which implements most current web standards.
- The latest Firefox features[15] include tabbed browsing, spell checking, incremental find, live book marking, a download manager, private browsing, location-aware browsing (also known as "geolocation") based exclusively on a Google service.
- Firefox runs on various operating systems including Microsoft Windows, Linux, Mac OS X, FreeBSD, and many other platforms.

1.7.5 CHROME

- Chrome, the web browser by Google, is rapidly becoming popular due to following features-
  - SPEED: Chrome is designed to be fast in every possible way: It's quick in staring up from the desktop, loading web pages and running complex web applications.
  - SIMPLICITY: Chrome's browser window is streamlined, clean and simple. Chrome also includes features that are designed for efficiency and ease of use. For example, you can search and navigate from the same box, and arrange tabs however you wish.
  - SECURITY: Chrome is designed to keep you safer and more secure on the web with built-in malware and phishing protection, auto updates to make sure the browser is up-to-date with the latest

security updates, and more. Learn more about Chrome's security features.

- Chrome is the first browser to incorporate machine translation in the browser itself, without requiring additional plugins orextensions.

1.8 SEARCH ENGINE

- A web search engine is designed to search for information on the World Wide Web and FTP servers. The search results are generally presented in a list of results and are often called hits. The information may consist of web pages, images, information and other types of files. Some search engines also mine data available in databases or open directories.
- The very first tool used for searching on the Internet was Archie.
- The first web robot, the Perl-based World Wide Web Wanderer was built and used by it to generate an index called 'Wandex'. The purpose of the Wanderer was to measure the size of the World Wide Web.
- Around 2000, Google's search engine rose to prominence. The company achieved better results for many searches with an innovation called Page Rank. This iterative algorithm ranks web pages based on the number and Page Rank of other web sites and pages that link there, on the premise that good or desirable pages are linked to more than others.
- Web search engines work by storing information about many web pages, which they retrieve from the html itself. These pages are retrieved by a Web crawler (sometimes also known as a spider) — an automated Web browser which follows every link on the site.
- This information is then analyzed and indexed The contents of each page are then analyzed to determine how it should be indexed . The purpose of an index is to allow information to be found as quickly as possible.

1.9 WEB SERVER

1.9.1 APACHE

- Apache HTTP Server is a free and open-source web server that delivers web content through the internet. It is commonly referred to as Apache and after development; it quickly became the most popular HTTP client on the web.
- The word, Apache, has been taken from the name of the Native American tribe ‘Apache’, famous for its skills in warfare and strategy making.
- Apache is the most widely used Web Server application in Unix-like operating systems but can be used on almost all platforms such as Windows, OS X, OS/2, etc

- It is a modular, process-based web server application that creates a new thread with each simultaneous connection. It supports a number of features; many of them are compiled as separate modules and extend its core functionality, and can provide everything from server side programming language support to authentication mechanism.

Features of Apache:

- Handling of static files
- Loadable dynamic modules
- Auto-indexing
- Compatible with IPv6
- Supports HTTP/2
- FTP connections
- Bandwidth throttling
- Load balancing
- Session tracking
- URL rewriting
- Geolocation based on IP address and many more

1.9.2 IIS

- Internet Information Services (IIS) is a flexible, general-purpose web server from Microsoft that runs on Windows systems to serve requested HTML pages or files. IIS works through a variety of standard languages and protocols
- An IIS web server accepts requests from remote client computers and returns the appropriate response. This basic functionality allows web servers to share and deliver information across local area networks (LAN), such as corporate intranets, and wide area networks (WAN), such as the internet.
- A web server can deliver information to users in several forms, such as static webpages coded in HTML; through file exchanges as downloads and uploads; and text documents, image files and more.
- Features of IIS:
  - IIS is used to host ASP.NET web applications and static websites
  - It can also be used as an FTP server, host WCF services, and be extended to host web applications. built on other platforms such as PHP
  - An invaluable feature is remote management. IIS can also be managed via the CLI or using Power Shell
  - One of the key feature of IIS is application pool.

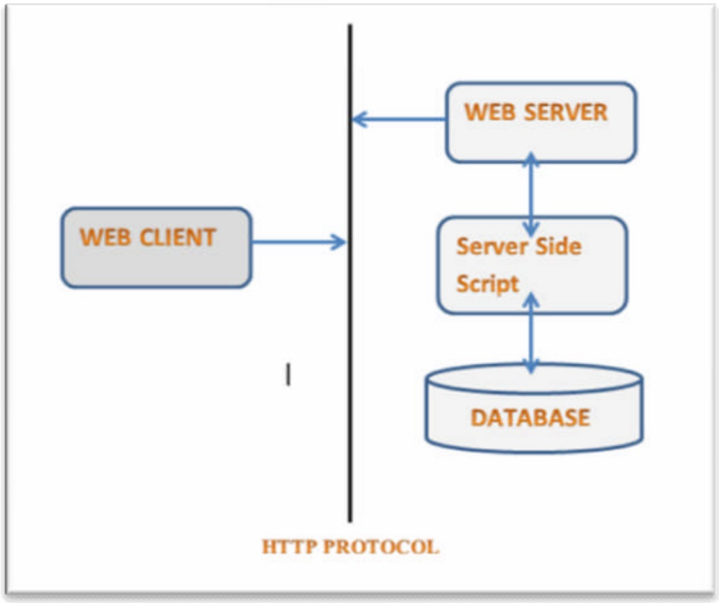
1.9.3 PROXY SERVER

- A proxy server provides a gateway between users and the internet. It is a server, referred to as an "intermediary" because it goes between end-users and the web pages they visit online.
- Proxies provide a valuable layer of security for your computer. They can be set up as web filters or firewalls, protecting your computer from internet threats like malware.
- Benefits of proxy server:
  - Enhanced security
  - Private browsing, watching, listening and shopping
  - Access to location– specific content
  - Prevent employees from browsing inappropriate or distracting sites

1.10 HTTP PROTOCOL

- HTTP stands for Hyper Text Transfer Protocol.
- It is a protocol used to access the data on the World Wide Web (www).
- The HTTP protocol can be used to transfer the data in the form of plain text, hypertext, audio, video, and so on.
- This protocol is known as Hyper Text Transfer Protocol because of its efficiency that allows us to use in a hypertext environment where there are rapid jumps from one document to another document.
- The HTTP protocol is a request/response protocol based on the client/server based architecture where web browsers, robots and search engines, etc. act like HTTP clients and the Web server acts as a server.

The following diagram shows a very basic architecture of a web application and depicts where HTTP :



The HTTP protocol is a request/response protocol based on the client/server based architecture where web browsers, robots and search engines, etc. act like HTTP clients, and the Web server acts as a server.
Basic Features:

- ✓ HTTP is connectionless protocol
- ✓ HTTP is media independent
- ✓ HTTP is stateless

1.11 EXERCISE

1. Define the internet. What protocol suit does it follow?
2. What is email? How is it sent and received?
3. Describe the three components of email.
4. What is meant by email address? What are the required parts of email address?
5. Are following email addresses valid?
   a. 124sir@idol.com
   b. 11 myname@yahoo.org
   c. Seema_Sathye@ server.co.in
   d. Piyush_mishra@myservices.net.in
6. What is telnet used for?
7. What is virtual terminal? What is it used for?
8. Name applications of telnet.
9. Why is FTP protocol used?
10. Explain different connections that can be used by FTP.
11. What are drawbacks of telnet and FTP?
12. Define E commerce. What are the advantages of ecommerce?
13. Give one example of B2B, B2C and C2C ecommerce.
14. Identify following E commerce category–
    a. Sale of online admission form for a college.
    b. Submission of the above form.
    c. Online resale of a second handcar.
    d. Purchase of raw material by an automobile company.
15. What is visual collaboration?
16. What is codec how does it function?
17. What are different types of video conferencing system?
18. List the components of video conferencing system.
19. Discuss the problems faced by video conferencing system.

20. Define E business.
21. List a few applications of E business.
22. What is a E business model? Give three examples of E business model.
23. What are risks for E business? What are the solutions available for these risks?
24. What is ISP? Explain the role of ISP in an internet connection.
25. Classify following technologies of ISP in business or home connections

a. A dial up connection with speed1Mbps

b. A connection to a LAN using leased cable lines

c. Hosting of personal webpage

d. Use of wi-fi for a laptop.

26. What are different services provided by ISP?
27. Explain what is peering? What is its advantage?
28. What is VISP? Give one example to explain its use.
29. Discuss the concept of free net and its importance.
30. What are domain name servers? What is their function?
31. What is internet address? How is it assigned?
32. Write a note on evolution of www
33. What is a web browser? List and compare different available web browsers