Chapter – 1

Internet and WWW

Introduction:

The Internet is the global system of interconnected computer networks that use the Internet protocol suite (TCP/IP) to link devices worldwide. It is a network of networks that consists of private, public, academic, business, and government networks of local to global scope, linked by a broad array of electronic, wireless, and optical networking technologies. The Internet carries an extensive range of information resources and services, such as the inter-linked hypertext documents and applications of the World Wide Web (WWW), electronic mail, telephony, and peer-to-peer networks for file sharing.

History:

Internet was originally conceived by the Department of Defense as a way to protect government communications systems in the event of a military strike. The original network, dubbed ARPANet (for the Advanced Research Projects Agency that developed it) evolved into a communications channel among contractors, military personnel, and university researchers who were contributing to ARPA projects.

The network employed a set of standard protocols to create an effective way for these people to communicate and share data with each other.

ARPAnet's popularity continued to spread among researchers, and in the 1980's the National Science Foundation, whose NSFNet, linked several high speed computers, and took charge of what had come to be known as the Internet.

By the late 1980's, thousands of cooperating networks were participating in the Internet.

In 1991, the U.S. High Performance Computing Act established the NREN (National Research & Education Network). NREN's goal was to develop and maintain high-speed networks for research and education, and to investigate commercial uses for the Internet.

The rest, as they say, is history in the making. The Internet has been improved through the developments of such services as Gopher and the World Wide Web.

Even though the Internet is predominantly thought of as a research oriented network, it continues to grow as an informational, creative, and commercial resource every day and all over the world.

Application areas of Internet:

The Internet has many important applications. Of the various services available via the Internet, the following are the areas where internet use:

1. **Communication:**

It is used for sending and receiving message from one and other through internet by using electronic mail. Some of the web sites providing this service are yahoomail.com Hotmail.com rediffmail.com etc.

2. **Job Searches:**

Getting information regarding availability of job in different sectors and areas. We can publish our resume in online for prospective job. Some of the web sites providing this service are naukri.com, monster.com, summerjob.com, recuritmentindia.com etc.

3. **Finding Books And Study Material :**

Books and other study material stored around the world can be easily located through internet. Latest encyclopedias are available online.

4. **Health And Medicine:**

Internet provide information and knowledge about field of health medicine people can have information about various disease and can receive help .patient can be taken to virtual check room where they can meet doctors. Some of the web sites providing this service are

5. **Travel:**

One can use internet to gather information about various tourist place. It can be used for booking Holiday tours, hotels, train and flights. Some of the web sites providing this service are indiatravelog.com, rajtravel.com, makemytrip.com.

6. **Entertainment:**

One can download jokes, songs movies, and latest sports updates through internet. Some of the web sites providing this service arecricinfo.com, movies.com espn.com. We can also watch online videos through internet with the help of YouTube.

7. **Shopping:**

Internet is also used for online shopping. By just giving accounts details we can perform the transaction. We can even pay our bills and perform bank related transaction.

8. **Stock Market Updates:**

We can sell or buy shares while sitting on computer through internet. Several websites like ndtvprofit.com, moneypore.com, provide information regarding investment.

9. **Research :**

A large number of people are using internet for research purposes we can download any kind information by using internet.

10. Business Use Of Internet:

There are different ways by which internet can be used for business are:

- Information about the product can be provided online to the customer
- Provide market information to the business e.g. Stock Market
- It help business to recruit talented people.
- Help in locating suppliers of the product
- Feedback and reviews about companies product
- Eliminate middle men and have a direct contact with customer
- Providing information to the investor by providing companies back ground and financial information on web site.

INTERNET SERVICE PROVIDER:

An Internet service provider (ISP) is a company that provides customers with Internet access. The ISP may be organized in various form such as: commercial community owned, Non-profit or privately owned. Data may be transmitted using several technologies, including dial-up, DSL, cable modem, wireless or dedicated high-speed interconnects.
Typically, ISPs also provide their customers with the ability to communicate with one another by providing Internet email accounts, usually with numerous email addresses at the customer's discretion. Other services, such as telephone and television services, may be provided as well. The services and service combinations may be unique to each ISP.
An Internet service provider is also known as an Internet access provider (IAP). Some ISP in Nepal are:

- WorldLink Communication Pvt Ltd
- Vianet Communication
- Websurfer Nepal Pvt Ltd
- Mercantile Communication
- Websurfer Nepal Pvt Ltd

DOMAIN NAME SERVER (DNS):

Domain Name Servers (DNS) are the Internet's equivalent of a phone book. They maintain a directory of domain names and translate them to Internet Protocol (IP) addresses.
This is necessary because, although domain names are easy for people to remember, computers or machines, access websites based on IP addresses.
Information from all the domain name servers across the Internet are gathered together and housed at the Central Registry. Host companies and Internet Service Providers interact with the Central Registry on a regular schedule to get updated DNS information.
When we type in a web address, e.g., www.jimsbikes.com, our Internet Service Provider views the DNS associated with the domain name, translates it into a machine friendly IP address (for example

216.168.224.70 is the IP for jimsbikes.com) and directs our Internet connection to the correct website.
After we register a new domain name or when we update the DNS servers on our domain name, it usually takes about 12-36 hours for the domain name servers world-wide to be updated and able to access the information. This 36-hour period is referred to as propagation.

NETWORK PROTOCOL:

A network protocol defines rules and conventions for communication between network devices. Network protocols include mechanisms for devices to identify and make connections with each other, as well as formatting rules that specify how data is packaged into messages sent and received. Some protocols also support message acknowledgement and data compression designed for reliable and/or high-performance network communication.
Modern protocols for computer networking all generally use packet switching techniques to send and receive messages in the form of packets - messages subdivided into pieces that are collected and re-assembled at their destination. Hundreds of different computer network protocols have been developed each designed for specific purposes and environments.

PROTOCOLS USED IN INTERNET:

TCP/IP (Transmission Control Protocol/Internet Protocol):

TCP (Transmission Control Protocol) and IP (Internet Protocol) are two different procedures that are often linked together. The linking of several protocols is common since the functions of different protocols can be complementary so that together they carry out some complete task. The combination of several protocols to carry out a particular task is often called a "stack" because it has layers of operations. In fact, the term "TCP/IP" is normally used to refer to a whole suite of protocols, each with different functions. This suite of protocols is what carries out the basic operations of the Web. TCP/IP is also used on many local area networks.
TCP/IP defines how electronic device (like computer) should be connected over the internet and how data should be transmitted between them. It is a set of rule that an application can use to package its information for sending across the network of networks.

TCP (Transmission Control Protocol):

When information is sent over the Internet, it is generally broken up into smaller pieces or "packets". The use of packets facilitates speedy transmission since different parts of a message can be sent by different routes and then reassembled at the destination. It is also a safety measure to minimize the chances of losing information in the transmission process. TCP is the means for creating the packets, putting them back together in the correct order at the end, and checking to make sure that no packets got lost in transmission. If necessary, TCP will request that a packet be resent.

❖ IP (Internet Protocol):

Internet Protocol (IP) is the method used to route information to the proper address. Every computer on the Internet has to have its own unique address known as the IP address. Every packet sent will contain an IP address showing where it is supposed to go. A packet may go through a number of computer routers before arriving at its final destination and IP controls the process of getting everything to the designated computer. Note that IP does not make physical connections between computers but relies on TCP for this function. IP is also used in conjunction with other protocols that create connections.

Common types of TCP/IP are:

a. HTTP (Hyper Text Transfer Protocol):

HTTP is a communication protocol. It defines mechanism for communication between browser and the web server. It is also called request and response protocol because the communication between browser and server takes place in request and response pairs.

b. HTTPS(Hypertext Transfer Protocol over Secure Socket Layer):

HTTPS was first introduced by Netscape. HTTPS takes care of secure communication between a web server and web browser. HTTPS typically handles credit card transaction and other sensitive data. A Web page using this protocol will have https: at the front of its URL.

c. FTP (File Transfer Protocol):

File Transport Protocol is a program that allows users and computers to send and receive large portions of data through a private or public network. It can also be used to send configuration files and software updates to network devices, such as switches and routers. FTP uses ports for communications and also uses encryption to protect the information being received and sent.

d. SMTP (Simple Mail Transfer Protocol):

SMTP is a TCP/IP protocol used in sending and receiving e-mail. However, since it is limited in its ability to queue massages at the receiving end, it is usually used with one of two other protocols POP3 or IMAP that let the user save the messages in a server mailbox and download them periodically from the server. In other words, users typically use a program that uses SMTP for sending e-mail and either POP3 or IMAP for receiving e-mail. On UNIX based systems, sendmail is the most widely used SMTP server for e-mail.

e. UDP (User Datagram Protocol):

The User Datagram Protocol is one of the core members of the internet protocol suite. UDP uses a simple connectionless transmission model with a minimum of protocol mechanism. UDP provides checksums for data integrity, and port numbers for addressing different functions at the source and destination of the datagram. It has no handshaking dialogues, and thus exposes the user's program to any unreliability of the underlying network and so there is no guarantee of delivery, ordering, or duplicate protection. If error correction facilities are needed at the network interface level, an

application may use the TCP or SCTP (Stream Control Transmission Protocol) which are design for this purpose.

WWW (WORLD WIDE WEB):

The World Wide Web (abbreviated WWW or the Web) is an information space where documents and other web resources are identified by Uniform Resource Locators (URLs), interlinked by hypertext links, and can be accessed via the Internet. English scientist Tim Berners-Lee invented the World Wide Web in 1989. He wrote the first web browser computer program in 1990 while employed at CERN in Switzerland. The Web browser was released outside of CERN in 1991, first to other research institutions starting in January 1991 and to the general public on the Internet in August 1991.

The World Wide Web has been central to the development of the Information Age and is the primary tool billions of people use to interact on the Internet. Web pages are primarily text documents formatted and annotated with Hypertext Markup Language (HTML). In addition to formatted text, web pages may contain images, video, audio, and software components that are rendered in the user's web browser as coherent pages of multimedia content. Embedded hyperlinks permit users to navigate between web pages. Multiple web pages with a common theme, a common domain name, or both, make up a website. Website content can largely be provided by the publisher, or interactive where users contribute content or the content depends upon the user or their actions. Websites may be mostly informative, primarily for entertainment, or largely for commercial, governmental, or non-governmental organizational purposes.

WEB SERVER:

A Web server is a program that uses HTTP (Hypertext Transfer Protocol) to serve the files that form Web pages to users, in response to their requests, which are forwarded by their computers' HTTP clients. Dedicated computers and appliances may be referred to as Web servers as well.

The process is an example of the client/server model. All computers that host Web sites must have Web server programs. Leading Web servers include Apache (the most widely-installed Web server), Microsoft's Internet Information Server (IIS) and nginx (pronounced engine X) from NGNIX. Other Web servers include Novell's NetWare server, Google Web Server (GWS) and IBM's family of Domino servers.

Web servers often come as part of a larger package of Internet- and intranet-related programs for serving email, downloading requests for File Transfer Protocol (FTP) files, and building and publishing Web pages. Considerations in choosing a Web server include how well it works with the operating system and other servers, its ability to handle server-side programming, security characteristics, and the particular publishing, search engine and site building tools that come with it.

WEB BROWSER:

A browser is an application program that provides a way to look at and interact with all the information on the World Wide Web. The word "browser" seems to have originated prior to the Web as a generic term for user interfaces that let you browse (navigate through and read) text files online.

Technically, a Web browser is a client program that uses HTTP (Hypertext Transfer Protocol) to make requests of Web servers throughout the Internet on behalf of the browser user. Most browsers support e-mail and the File Transfer Protocol (FTP) but a Web browser is not required for those Internet protocols and more specialized client programs are more popular.

The first Web browser, called WorldWideWeb, was created in 1990. That browser's name was changed to Nexus to avoid confusion with the developing information space known as the World Wide Web. The first Web browser with a graphical user interface was Mosaic, which appeared in 1993.

The most popular web browsers are Google Chrome, Microsoft Edge (preceded by Internet Explorer), Safari, Opera and Firefox.

SEARCH ENGINE:

A search engine is a software program or script available through the Internet that searches documents and files for keywords and returns the results of any files containing those keywords. Today, there are thousands of different search engines available on the Internet, each with their own abilities and features. The first search engine ever developed is considered Archie, which was used to search for FTP files and the first text-based search engine is considered Veronica. Today, the most popular and well-known search engine is Google.

Large search engines contain millions and sometimes billions of pages, many search engines not only just search the pages but also display the results depending upon their importance. This importance is commonly determined by using various algorithms.

For users, a search engine is accessed through a browser on their computer, smartphone, tablet, or another device.

META SEARCH ENGINE:

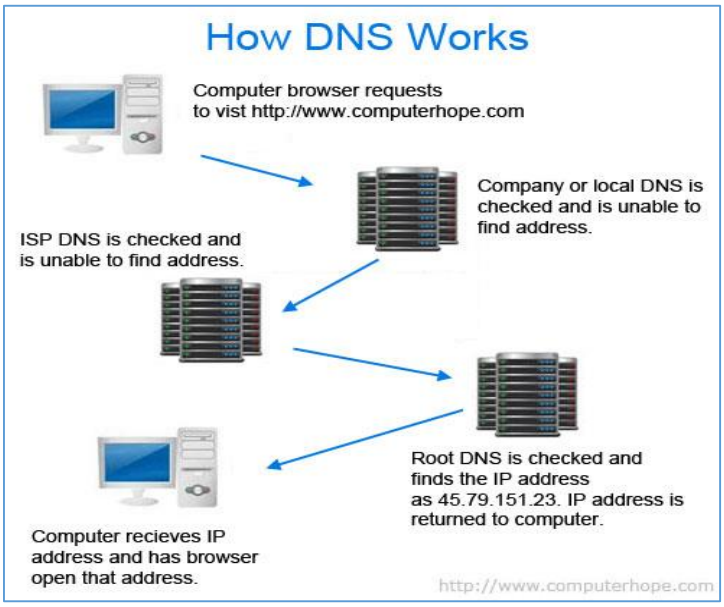
A search engine that queries other search engines and then combines the results that are received from all. In erect, the user is not using just one search engine but a combination of many search engine at once to optimize web searching. For example: Dogpile, Mamma are metasearch engine. A metasearch engine (or aggregator) is a search tool that uses another search engine’s data to produce their own results from the internet. Metasearch engines take input from a user and simultaneously send out queries to third party search engines for results.

DOMAIN NAME:

When referring to an Internet address or name a **domain** or **domain name** is the location of a website. For example, the domain name "computerhope.com" points to the IP address "45.79.151.23", but it is generally easier to remember a name rather than a long string of numbers. A domain name can be a maximum of sixty-three characters with one character minimum, and is entered after the protocol in the URL, as we can see in the following example.

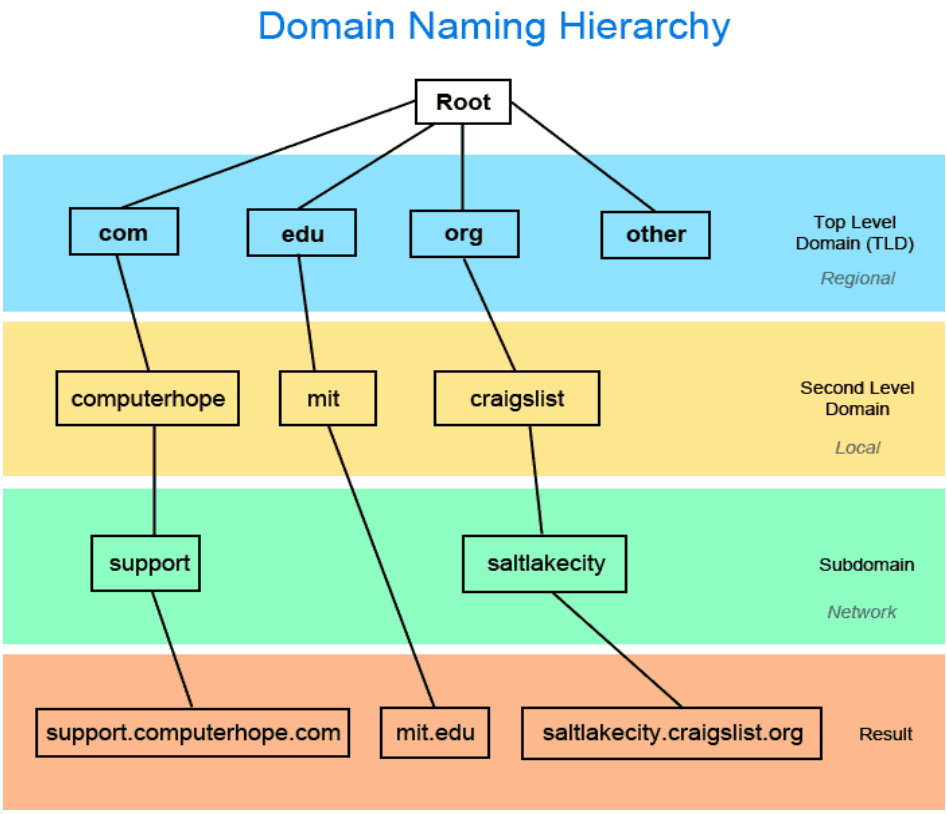
When deciding on a domain name, it's a good idea to keep it simple, something that is easy to remember. Additional promoting tips for websites is on our promotion page. To register or lookup a domain name, we recommend visiting GoDaddy or Network solutions; both companies are domain name registrars.

When referring to a computer network, a domain is a group of resources assigned to a specific group of individuals. It is used to divide global areas or departments of a corporation. A domain may need to be specified when mapping a network computer or drive.



DOMAIN NAME HIERARCHY:

There are millions of websites available on the Internet to make finding an address the domain naming is broken into a hierarchical structure. Below is how DNS is structured on the Internet and an example of three different addresses.



In the above example, all websites are broken into regional sections based on the top level domain (TLD). In the example of http://support.computerhope.com it has a ".com" TLD, with "computerhope" as its second level domain that is local to the .com TLD, and "support" as its subdomain, which is determined by its server.

ISSUES RELATED WITH DOMAIN NAME REGISTRATION:

- **Typosquatting:**

Where a person registers a domain name similar to a real domain name, but with a typo, in hopes that web surfers reach it by accident. These sites are usually filled with paid advertising links that generate revenue for the typosquatter, not to mention the web surfer has been tricked into believing he is on the correct site. This diverts traffic away from the intended site. Sometimes they are routed to a competitor's site or a pornographic site.

- **Cybersquatting:**

When someone registers a domain name, in bad faith, violating the rights of the trademark owner. They usually intend to extort payment from the trademark owner, and they keep the names to sell later to the highest bidder.

- **Pagejacking:**

When the offender copies part of an existing website, and then puts it up on a different website to make it look like the original. Pagejacking is used in phishing schemes, where the fake page gathers account numbers, passwords, and personal information from the unsuspecting user.

- **The Uniform Domain Name Dispute Resolution Policy (UDRP):**

It is a cost-effective and faster alternative to a lawsuit, when there is a domain name dispute that needs to be resolved. This was set up by the Internet Corporation for Assigned Names and Numbers (ICANN), the group responsible for domain name registration.

DOMAIN NAME SYSTEM:

Short for Domain Name System or Domain Name Service, a DNS is an Internet or network server that helps to point domain names or hostnames to their associated Internet Protocol address and was introduced by Paul Mockapetris and Jon Postel in 1983. Without a DNS to resolve a domain name or the proper rights, users would have to know the IP address of each of the web pages or computers you wanted to access.

CLIENT/SERVER ARCHITECTURE:

Client/server architecture is a computing model in which the server hosts, delivers and manages most of the resources and services to be consumed by the client. This type of architecture has one or more client computers connected to a central server over a network or Internet connection. This system shares computing resources.
Client/server architecture may also be referred to as a networking computing model because all the requests and services are delivered over a network.
Client/server architecture works when the client computer sends a resource or process request to the server over the network connection, which is then processed and delivered to the client. A server computer can manage several clients simultaneously, whereas one client can be connected to several servers at a time, each providing a different set of services. In its simplest form, the Internet is also based on client/server architecture where the Web server serves many simultaneous users with Web page and or website data.

CROSS BROWSER COMMUNICATION:

When a software program is developed for multiple computer platforms, it is called a crossplatform program. Similarly, when a website is developed for multiple browsers, it is called a cross-browser website.
The job of a Web developer would be much easier if all browsers were the same. While most browsers are similar in both design and function, they often have several small differences in the way they recognize and display websites. For example, Apple's Safari uses a different HTML rendering engines than Internet Explorer. This means the browsers may display the same Web page with slightly different page and text formatting. Since not all browsers support the same HTML tags,

some formatting may not be recognized at all in an incompatible Web browser. Furthermore, browsers interpret JavaScript code differently, which means a script may work fine in one browser, but not in another.

Because of the differences in the way Web browsers interpret HTML and JavaScript, Web developers must test and adapt their sites to work with multiple browsers. For example, if a certain page looks fine in Firefox, but does not show up correctly in Internet Explorer, the developer may change the formatting so that it works with Internet Explorer. Of course, the page may then appear differently in Firefox. The easiest fix for browser incompatibility problems is to use a more basic coding technique that works in both browsers. However, if this solution is not possible, the developer may need to add code that detects the type of browser, then outputs custom HTML or JavaScript for that browser.

Making a cross-browser site is usually pretty simple for basic websites. However, complex sites with a lot of HTML formatting and JavaScript may require significant extra coding in order to be compatible with multiple browsers. Some developers may even generate completely different pages for each browser. While CSS formatting has helped standardize the appearance of Web pages across multiple browsers, there are still several inconsistencies between Web browsers. Therefore, cross-browser design continues to be a necessary aspect of Web development.