**Course Web Programming Unit I**

***Contents***

*What is Internet? Introduction to internet and its applications, E- mail, telnet, FTP, e-commerce, video conferencing, e-business. Internet service providers, domain name server, internet address World Wide Web (WWW) : World Wide Web and its evolution, uniform resource locator (URL), browsers -Search engine, web saver - apache, IIS, proxy server, HTTP protocol.*

**WHAT IS INTERNET?**

Internet is defined as an Information super Highway, to access information over the web. However, It can be defined in many ways as follows:

1. Internet is a world-wide global system of interconnected computer networks.
2. Internet uses the standard Internet Protocol (TCP/IP).
3. Every computer in internet is identified by a unique IP address.
4. IP Address is a unique set of numbers (such as 110.22.33.114) which identifies a computer location.
5. A special computer DNS (Domain Name Server) is used to give name to the IP Address so that user can locate a computer by a name.
6. For example, a DNS server will resolve a name **http://www.tutorialspoint.com** to a particular IP address to uniquely identify the computer on which this website is hosted.
7. Internet is accessible to every user all over the world.

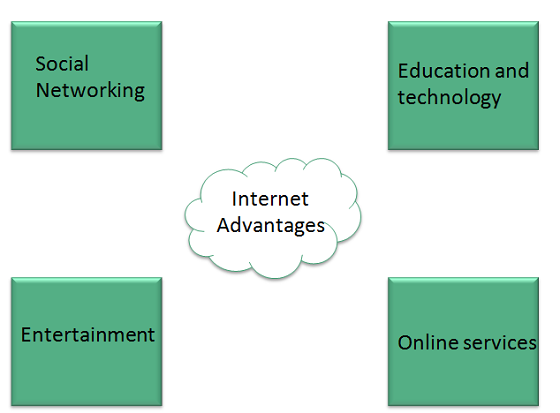
**Evolution of Internet**

The concept of Internet was originated in 1969 and has undergone several technological & Infrastructural changes as discussed below:

* The origin of Internet devised from the concept of Advanced Research Project Agency Network (ARPANET).
* ARPANET was developed by United States Department of Defense.
* Basic purpose of ARPANET was to provide communication among the various bodies of government.
* Initially, there were only four nodes, formally called Hosts.
* In 1972, the ARPANET spread over the globe with 23 nodes located at different countries and thus became known as Internet.
* By the time, with invention of new technologies such as TCP/IP protocols, DNS, WWW, browsers, scripting languages etc.,Internet provided a medium to publish and access information over the web.

**Advantages**

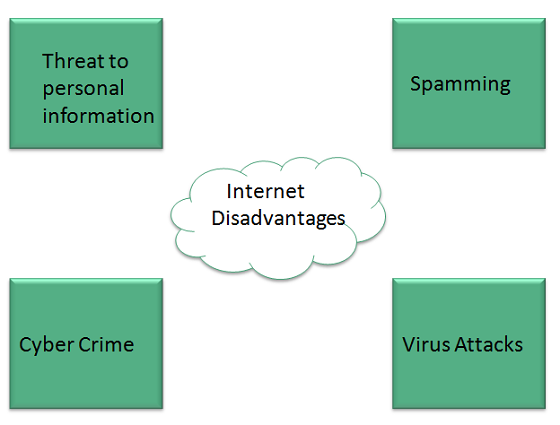
Internet covers almost every aspect of life, one can think of. Here, we will discuss some of the advantages of Internet:



* Internet allows us to communicate with the people sitting at remote locations. There are various apps available on the wed that uses Internet as a medium for communication. One can find various social networking sites such as:
  + Facebook
  + Twitter
  + Yahoo
  + Google+
  + Flickr
  + Orkut
* One can surf for any kind of information over the internet. Information regarding various topics such as Technology, Health & Science, Social Studies, Geographical Information, Information Technology, Products etc can be surfed with help of a search engine.
* Apart from communication and source of information, internet also serves a medium for entertainment. Following are the various modes for entertainment over internet.
  + Online Television
  + Online Games
  + Songs
  + Videos
  + Social Networking Apps
* Internet allows us to use many services like:
  + Internet Banking
  + Matrimonial Services
  + Online Shopping
  + Online Ticket Booking
  + Online Bill Payment
  + Data Sharing
  + E-mail
* Internet provides concept of electronic commerce, that allows the business deals to be conducted on electronic systems

**Disadvantages**

However, Internet has prooved to be a powerful source of information in almost every field, yet there exists many disadvanatges discussed below:



* There are always chances to loose personal information such as name, address, credit card number. Therefore, one should be very careful while sharing such information. One should use credit cards only through authenticated sites.
* Another disadvantage is the **Spamming**. Spamming corresponds to the unwanted e-mails in bulk. These e-mails serve no purpose and lead to obstruction of entire system.
* **Virus** can easily be spread to the computers connected to internet. Such virus attacks may cause your system to crash or your important data may get deleted.
* Also a biggest threat on internet is pornography. There are many pornographic sites that can be found, letting your children to use internet which indirectly affects the children healthy mental life.
* There are various websites that do not provide the authenticated information. This leads to misconception among many people.

**Email**

Email is a service which allows us to send the message in electronic mode over the internet. It offers an efficient, inexpensive and real time mean of distributing information among people.

**E-Mail Address**

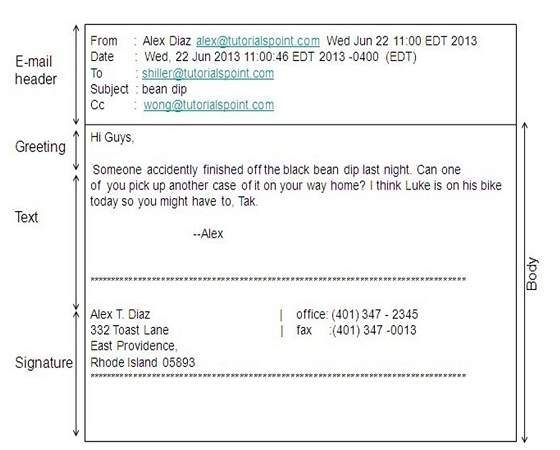
Each user of email is assigned a unique name for his email account. This name is known as E-mail address. Different users can send and receive messages according to the e-mail address.

E-mail is generally of the form username@domainname. For example, webmaster@tutorialspoint.com is an e-mail address where webmaster is username and tutorialspoint.com is domain name.

* The username and the domain name are separated by **& (at)** symbol.
* E-mail addresses are not case sensitive.
* Spaces are not allowed in e-mail address.

**E-mail Message Components**

E-mail message comprises of different components: E-mail Header, Greeting, Text, and Signature. These components are described in the following diagram:



**E-mail Header**

The first five lines of an E-mail message is called E-mail header. The header part comprises of following fields:

***From***

The From field indicates the sender’s address i.e. who sent the e-mail.

***Date***

The Date field indicates the date when the e-mail was sent.

***To***

The To field indicates the recipient’s address i.e. to whom the e-mail is sent.

***Subject***

The Subject field indicates the purpose of e-mail. It should be precise and to the point.

***CC***

CC stands for Carbon copy. It includes those recipient addresses whom we want to keep informed but not exactly the intended recipient.

***BCC***

BCC stands for Black Carbon Copy. It is used when we do not want one or more of the recipients to know that someone else was copied on the message.

***Greeting***

Greeting is the opening of the actual message. Eg. Hi Sir or Hi Guys etc.

***Text***

It represents the actual content of the message.

***Signature***

This is the final part of an e-mail message. It includes Name of Sender, Address, and Contact Number.

**ADVANTAGES**

E-mail has proved to be powerful and reliable medium of communication. Here are the benefits of e-mail:

***Reliable***

Many of the mail systems notify the sender if e-mail message was undeliverable.

***Convenience***

There is no requirement of stationary and stamps. One does not have to go to post office. But all these things are not required for sending or receiving an mail.

***Speed***

E-mail is very fast. However, the speed also depends upon the underlying network.

***Inexpensive***

The cost of sending e-mail is very low.

***Printable***

It is easy to obtain a hardcopy of an e-mail. Also an electronic copy of an e-mail can also be saved for records.

***Global***

E-mail can be sent and received by a person sitting across the globe.

***Generality***

It is also possible to send graphics, programs and sounds with an e-mail.

**Disadvantages**

Apart from several benefits of E-mail, there also exists some disadvantages as discussed below:

***Forgery***

E-mail doesn’t prevent from forgery, that is, someone impersonating the sender, since sender is usually not authenticated in any way.

***Overload***

Convenience of E-mail may result in a flood of mail.

***Misdirection***

It is possible that you may send e-mail to an unintended recipient.

***Junk***

Junk emails are undesirable and inappropriate emails. Junk emails are sometimes referred to as spam.

***No Response***

It may be frustrating when the recipient does not read the e-mail and respond on a regular basis.

**TELNET**

Telnet is a Network Protocol and the tool which uses that protocol (i.e Telnet) is also known as Telnet. Telnet is used to connect to an External or a Remote Computer via Port 23. You’ll need to enter the Login and Password in order to connect to that remote PC. After connecting to your remote PC you can carry out commands to preform your task on your Remote PC for example you want to create a text file or you want to check how your system is working or you want to check the directory of your remote PC.

Telnet offers a bunch of commands, here are some:

**close** ---- close current connection

**logout** ---- forcibly logout remote user and close the connection

**display** ---- display operating parameters

**mode** ---- try to enter line or character mode ('mode ?' for more)

**open** ---- connect to a site

**quit** ---- exit telnet

**send** ---- transmit special characters ('send ?' for more)

**set** ---- set operating parameters ('set ?' for more)

**unset** ---- unset operating parameters ('unset ?' for more)

**status** ---- print status information

**toggle** ---- toggle operating parameters ('toggle ?' for more)

**slc** ---- set treatment of special characters

**z** ---- suspend telnet

**Connecting using Telnet**

There are two ways to connect using telnet, one is the direct way and the other is using Telnet commands.

***Direct Way:***

telnet 192.168.1.1 23

***Telnet Way:***

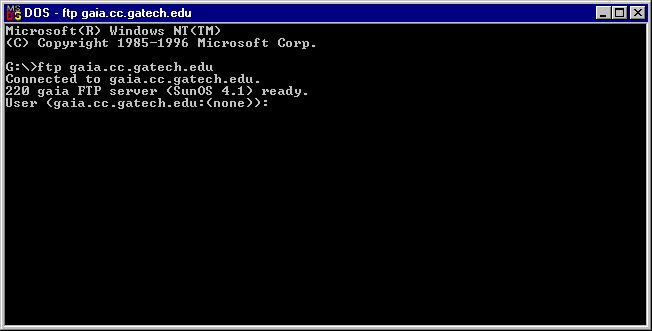
First type telnet in console then you’ll enter the Telnet session, type

open 192.168.1.1

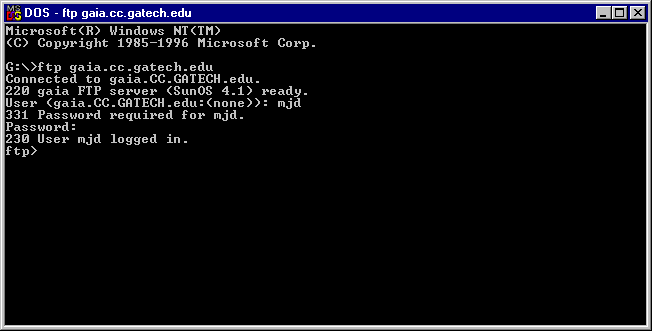
**FTP (File Transfer Protocol)**

FTP ("File Transfer Program") allows you to "put" and "get" files to and from remote machines so you can edit your programs on your local machine and upload them when you are ready to compile. To make FTP work, you need a client (your machine) and a server (the machine to/from which you are putting/getting files). Most UNIX and file server machines operate an FTP server for your file transferring pleasure so all you need to worry about is operating the client software, and that's easy!

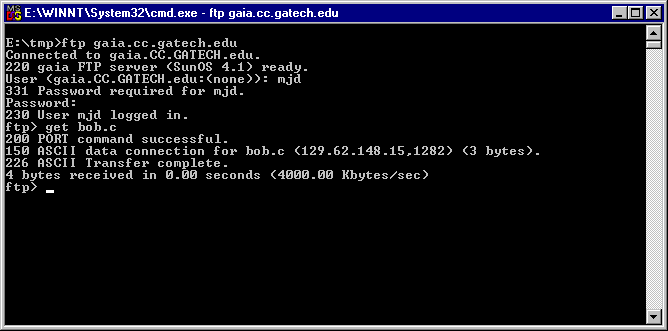
The basic FTP client is a command-line program. Most operating systems with a command-line environment (e.g. DOS, UNIX, VMS) have a standard FTP client program. To use the client, type "ftp <FTP Server>" (no quotes) where <FTP Server> is the name (e.g. ftp.netscape.com) or IP address (e.g. 193.4.5.6) of the FTP Server.



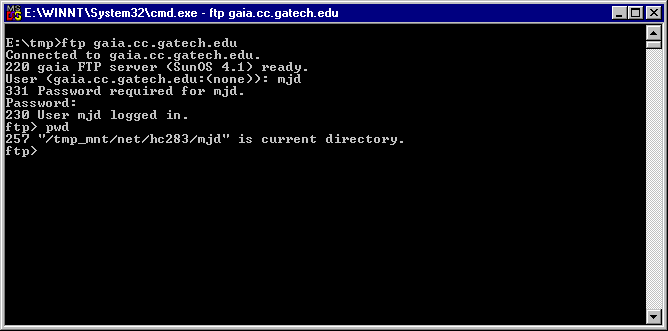
Upon connection, the FTP Server prompts the user for a user id and password. This is usually the same user id and password needed to log onto the machine. After entering the user id and password, the FTP client is logged on as indicated by the FTP prompt



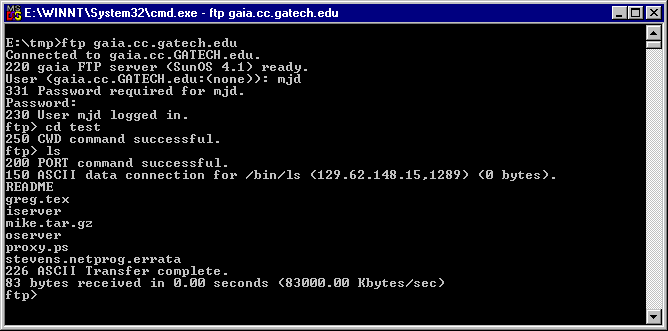
The user can either "put" files from the client to the server or "get" file from the server to the client. If you have been preparing a program to upload to the server, you must "put" it onto the server. To put a file, you simply type "put <filename>" at the FTP prompt where <filename> is the name of the file you want to upload (e.g. put bob.c uploads the file bob.c). To download a file, you simply type "get <filename>" at the FTP prompt. If you want to use wildcards (e.g. to upload all ".c" files), use the mget or mput command ("m" stands for "multiple"). For example, "mget \*.c" downloads all files with the ".c" extension.



An obvious question is "From what directory does the file I **put** come, and to which directory does it go?" (or vice-versa for **get**). In FTP, you must keep track of both the client and server directory. The client directory is the directory from which you ftp'd. In the example above, I am in directory E:\tmp when I ftp to the server. Any puts or gets (without an absolute pathname) will be put from or gotten to this directory. When you FTP to a server, you are initially placed in a default starting directory, usually your "home" directory on the server. To determine where you are on the server, try the "pwd" (path of working directory) command at the FTP prompt as in



You may traverse the directories on the server with the "cd" (change directory) command and list the files in the current server directory with the "ls" (list) command as in



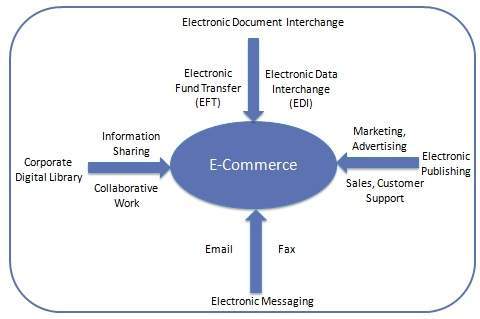
You may change the current working directory of the **client** with the "lcd" (local change directory) command. To see a list of commands, type "help" at the ftp prompt. You may get help with a specific command by typing "help <command>" (e.g. help lcd).

Graphical programs exist that hide the command-line complexities of FTP for the Macintosh (e.g. Fetch) and PC (e.g. CuteFTP and WS-FTP) so you may want to try these.

**E-COMMERCE**

E-Commerce or Electronics Commerce is a methodology of modern business which addresses the need of business organizations, vendors and customers to reduce cost and improve the quality of goods and services while increasing the speed of delivery. E-commerce refers to paperless exchange of business information using following ways.

* Electronic Data Exchange (EDI)
* Electronic Mail (e-mail)
* Electronic Bulletin Boards
* Electronic Fund Transfer (EFT)
* Other Network-based technologies



**Features**

E-Commerce provides following features

* **Non-Cash Payment** − E-Commerce enables use of credit cards, debit cards, smart cards, electronic fund transfer via bank's website and other modes of electronics payment.
* **24x7 Service availability** − E-commerce automates business of enterprises and services provided by them to customers are available anytime, anywhere. Here 24x7 refers to 24 hours of each seven days of a week.
* **Advertising / Marketing** − E-commerce increases the reach of advertising of products and services of businesses. It helps in better marketing management of products / services.
* **Improved Sales** − Using E-Commerce, orders for the products can be generated any time, any where without any human intervention. By this way, dependencies to buy a product reduce at large and sales increases.
* **Support** − E-Commerce provides various ways to provide pre sales and post sales assistance to provide better services to customers.
* **Inventory Management** − Using E-Commerce, inventory management of products becomes automated. Reports get generated instantly when required. Product inventory management becomes very efficient and easy to maintain.

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Traditional Commerce** | **E-Commerce** |
| 1 | Heavy dependency on information exchange from person to person. | Information sharing is made easy via electronic communication channels making little dependency on person to person information exchange. |
| 2 | Communication/ transaction are done in synchronous way. Manual intervention is required for each communication or transaction. | Communication or transaction can be done in asynchronous way. Electronics system automatically handles when to pass communication to required person or do the transactions. |
| 3 | It is difficult to establish and maintain standard practices in traditional commerce. | A uniform strategy can be easily established and maintain in e-commerce. |
| 4 | Communications of business depends upon individual skills. | In e-Commerce or Electronic Market, there is no human intervention. |
| 5 | Unavailability of a uniform platform as traditional commerce depends heavily on personal communication. | E-Commerce website provides user a platform where al l information is available at one place. |
| 6 | No uniform platform for information sharing as it depends heavily on personal communication. | E-Commerce provides a universal platform to support commercial / business activities across the globe. |

**E-BUSINESS**

E-business (electronic business) is the conduct of business processes on the Internet. These electronic business processes include buying and selling products, supplies and services; servicing customers; processing payments; managing production control; collaborating with business partners; sharing information; running automated employee services; recruiting; and more.

E-business can comprise a range of functions and services, ranging from the development of [intranets](http://searchwindevelopment.techtarget.com/definition/intranet) and [extranets](http://searchenterprisewan.techtarget.com/definition/extranet) to [e-service](http://searchcio.techtarget.com/definition/e-services), the provision of services and tasks over the Internet by [application service providers](http://searchsoa.techtarget.com/definition/application-service-provider). Today, as major corporations continuously rethink their businesses in terms of the Internet, specifically its availability, wide reach and ever-changing capabilities, they are conducting e-business to buy parts and supplies from other companies, collaborate on sales promotions, and conduct joint research. With the security built into today's [browsers](http://searchwindevelopment.techtarget.com/definition/browser), and with [digital certificates](http://searchsecurity.techtarget.com/definition/digital-certificate) now available for individuals and companies from Verisign, a certificate issuer, much of the early concern about the security of business transaction on the Web has abated, and e-business by whatever name is accelerating

**VIDEO CONFERENCING**

A video conference uses a computer, camera, speakers, microphone, coder/decoder and network (such as the Internet) to conduct a live conference between two or more people. It uses audio and video telecommunication to bring people at different sites together for a meeting. This can be as simple as a conversation between two people in private offices (point-to-point) or involve several sites (multi-point) with more than one person in large rooms at different sites. Besides the audio and visual transmission of meeting activities, videoconferencing can be used to share documents, computer-displayed information and even whiteboards. (This information was sourced and adapted from: <http://en.wikipedia.org/wiki/Videoconferencing>)

Video conferencing is an extremely useful method of communication.

* **In education...**
  + Students from diverse communities and backgrounds can explore, communicate, analyze and share information and ideas through two-way communication forums.
  + Teachers and lecturers can communicate with one another from remote or otherwise isolated educational facilities.
  + Faculty members can keep in touch with classes while attending conferences.
  + Faculty members can participate in thesis defences at other institutions.
  + Administrators on tight schedules can collaborate on budget preparations from different parts of campus.
  + Faculty committee members can audition scholarship candidates.
  + Guest lecturers can be brought into the classroom from other institutions.
  + Researchers can collaborate with colleagues at other institutions on a regular basis without loss of time due to travel.
  + Schools with multiple campuses can collaborate and share professors.
  + Researchers can answer questions about grant proposals from agencies or review committees.
* **In business...**
  + People in distant locations can participate in meetings on short notice; saving them time and money.
  + Job interviews can be conducted from different cities, removing the need to fly candidates around the country.
* **In law...**testimony can be given by an individual who is not able to attend the physical legal settings.
* **In media...**journalists can participate in international press conferences from any location, without having to leave their offices.

(Information sourced and adapted from: <http://en.wikipedia.org/wiki/Videoconferencing>)

**Equipment needed**

The following video conferencing equipment will need to be set up according to the type of video conference you're going to have.

* ***Camera****:*A camera is necessary to send images of your venue to recipients. It is optional when you join a hosted video conferencing session and you only need to view and listen to the broadcast. A basic webcam (including those build-in to notebooks) is quite sufficient for broadcasting the image of a one or two persons. For larger groups it is recommended that you use a mid-range webcam, similar to the Microsoft LifeCam Studio. Specialized video conferencing rooms require a specialized camera with a specific lens suited for the venue and the type of video conference to be set up.
* ***Display device:***A display device is required to view the image or video broadcast from your remote venue(s). The basic notebook or computer monitor would be sufficient for one to two people to view a particular broadcast. If your audience increases in size, the display device should be substituted with a digital projector or [LCD](http://www.icts.uct.ac.za/modules.php?name=News&file=article&sid=4988#LCD) Monitor. In cases where it is necessary to view multiple groups at the same time or where applications are shared (as part of the video conferencing session) a second display device would be advisable.
* ***Speakers:***Speaker technology has advanced so much so that basic build-in notebook or computer speakers are often sufficient for a one-to-one person broadcast. A USB microphone headset would greatly enhance personal sound and voice quality. Depending on the requirements for a specific session, additional speakers might be required. In most cases this could be accomplished by connecting the computer to the venue's sound system.
* ***Microphone:*** A microphone is only required when sound needs to be contributed. When only viewing a hosted session, a microphone is not required. Specialised microphones are available for larger venues and are selected based on the requirements and venue acoustics. In most cases only one person at each venue would be contributing to the conversation and a USB microphone headset would be sufficient.
* ***Venue:*** When viewing a broadcast conference using your computer or if you're having a one-to-one conference, the venue itself might not be that important. However, there are some universal environmental factors to consider. When equipping a room with video conferencing facilities it is very important to consider the acoustics, layout, furniture, lighting, wall colour and decorations of the room. See the detailed section on [venue considerations](http://www.icts.uct.ac.za/modules.php?name=News&file=article&sid=4988#venue_considerations) for more information.
* ***Coder or Decoder:***Non-computer based video conferencing requires a coder or decoder to perform the communication function that the computer would perform. The devices support [ISDN](http://www.icts.uct.ac.za/modules.php?name=News&file=article&sid=4988#ISDN), [IP](http://www.icts.uct.ac.za/modules.php?name=News&file=article&sid=4988#IP) or both. ISDN make use of a connection via the telephone network and also require the necessary [PABX](http://www.icts.uct.ac.za/modules.php?name=News&file=article&sid=4988#PABX) infrastructure to be in place. ISDN preceded IP technology and therefore most of the older units in use only support ISDN. IP units use an Internet connection to connect with one another. When setting up a multi-location video conferencing session you need to consider the audience and the equipment they will use to connect. Should a combination of ISDN and IP equipment be used then a bridging function is required. It is possible to purchase a video conferencing bridge, however it is more economical to rent time on a hosted bridge. There are a number of service providers available. Reservations on a hosted bridge need to be made in advance and require the opening an account with the specific vendor.

**Venue considerations**

* ***Table:*** The shape of the table will determine the seating position, which in turn will determine if all participants will be visible to the camera. Reflected light is coloured by the surface from which it is reflected. The table surface will not only affect the light reflection, but will also affect the acoustics. For example, a leather-covered table or a table with a table cloth will absorb noise created by someone writing or clicking on the table, whereas a metal or glass table will amplify those sounds.
* ***Room:***Colours and decor will determine the amount of reflected light, which will affect the quality of the image transmitted. The shape and layout of your room will greatly influence the quality of the sound and image transmitted from the venue. Non-uniform rooms are best suited to reduce reflection and noise vibration. The ideal shape of a conference venue is a rectangle, as opposed to a square room.
* ***Walls:***These should be painted in matte finish as gloss creates unwanted reflections. The most suitable colours are mid-grey and greyish tones of blue and green. Sharp or dark colours should be avoided. Should noise levels be an issue, sound dampening paint or materials could be applied to the walls.
* ***Floors:*** These play a similar role to that of the table surface depending on the floor covering used. Carpeted floors are ideal as they dampen noise and reduce reflection.  
  ***Windows:***They provide light, normally from one side of the room, which can cause the faces of participants to be either too light or too dark (depending which side of the room they are located). Try to install either curtains, blinds or some mechanism that will allow control over the amount of light that enters the room. Another alternative is to tint or frost the windows.
* ***Lighting:***This will determine the image quality of the viewing parties. Soft diffused lighting is ideal. Fluorescent lighting with diffusers are best suited for video conferencing. Down lighters should be avoided as they cause uneven lighting and cast shadows.
* ***Furniture:***The colour of the furniture and its finishes are very important as they affect both the acoustics in the room and the amount of reflected light. Padded chairs will help dampen sound while matte finished covering (without bold patterns) will assist in creating a clear video image. Try to arrange your participants to sit in a horseshoe arrangement. A rectangular or oval table is strongly recommended. Mid-brown tables with limited reflection will help reduce shadows on faces that would be created by an upward reflection. In large venues, table inputs for power, microphones and digital displays should also be considered.

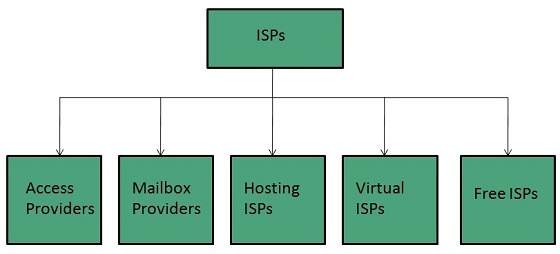
**INTERNET SERVICE PROVIDERS**

**Internet Service Provider (ISP)** is a company offering access to internet. They offer various services:

* Internet Access
* Domain name registration
* Dial-up access
* Leased line access

**ISP Types**

ISPs can broadly be classified into six categories as shown in the following diagram:



**Access providers**

They provide access to internet through telephone lines, cable wi-fi or fiber optics.

**Mailbox Provider**

Such providers offer mailbox hosting services.

**Hosting ISPs**

Hosting ISPs offers e-mail, and other web hosting services such as virtual machines, clouds etc.

**Virtual ISPs**

Such ISPs offer internet access via other ISP services.

**Free ISPs**

Free ISPs do not charge for internet services.

**Connection Types**

There exist several ways to connect to the internet. Following are these connection types available:

1. Dial-up Connection
2. ISDN
3. DSL
4. Cable TV Internet connections
5. Satellite Internet connections
6. Wireless Internet Connections

**Dial-up Connection**

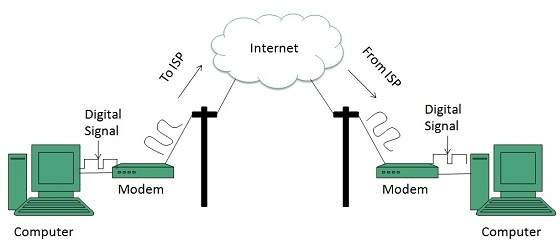
**Dial-up** connection uses telephone line to connect PC to the internet. It requires a modem to setup dial-up connection. This modem works as an interface between PC and the telephone line.

There is also a communication program that instructs the modem to make a call to specific number provided by an ISP.

Dial-up connection uses either of the following protocols:

1. Serial Line Internet Protocol (SLIP)
2. Point to Point Protocol (PPP)

The following diagram shows the accessing internet using modem:



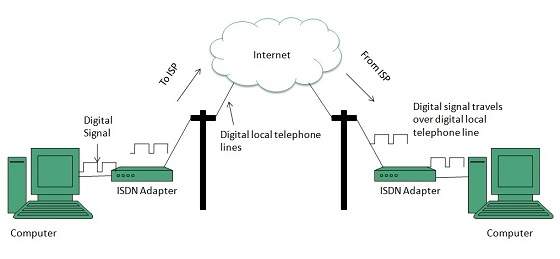
**ISDN**

**ISDN** is acronym of **Integrated Services Digital Network.** It establishes the connection using the phone lines which carry digital signals instead of analog signals.

There are two techniques to deliver ISDN services:

1. Basic Rate Interface (BRI)
2. Primary Rate Interface (PRI)

The following diagram shows accessing internet using ISDN connection:



**DSL**

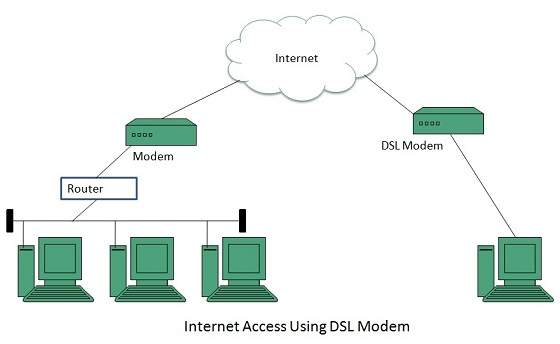
**DSL** is acronym of **Digital Subscriber Line.** It is a form of broadband connection as it provides connection over ordinary telephone lines.

Following are the several versions of DSL technique available today:

1. Asymmetric DSL (ADSL)
2. Symmetric DSL (SDSL)
3. High bit-rate DSL (HDSL)
4. Rate adaptive DSL (RDSL)
5. Very high bit-rate DSL (VDSL)
6. ISDN DSL (IDSL)

All of the above mentioned technologies differ in their upload and download speed, bit transfer rate and level of service.

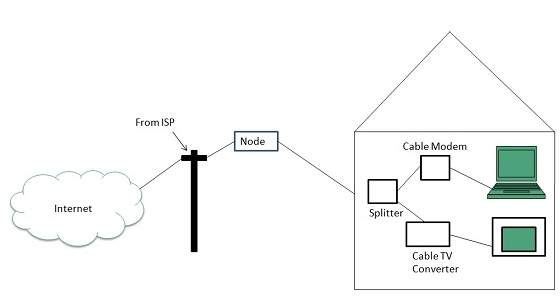
The following diagram shows that how we can connect to internet using DSL technology:



Cable TV Internet Connection

Cable TV Internet connection is provided through Cable TV lines. It uses coaxial cable which is capable of transferring data at much higher speed than common telephone line.

The following diagram shows that how internet is accessed using Cable TV connection:



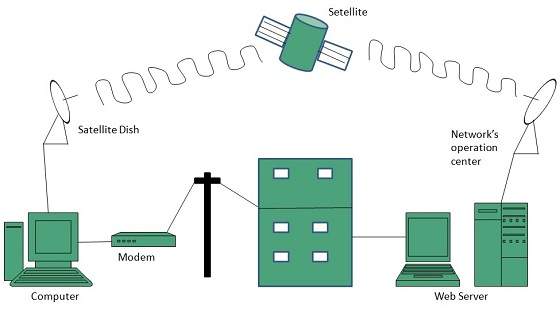
**Satellite Internet Connection**

Satellite Internet connection offers high speed connection to the internet. There are two types of satellite internet connection: one way connection or two way connection.

In one way connection, we can only download data but if we want to upload, we need a dialup access through ISP over telephone line.

In two way connection, we can download and upload the data by the satellite. It does not require any dialup connection.

The following diagram shows how internet is accessed using satellite internet connection:



**Wireless Internet Connection**

Wireless Internet Connection makes use of radio frequency bands to connect to the internet and offers a very high speed. The wireless internet connection can be obtained by either WiFi or Bluetooth.

**DOMAIN NAME SERVER**

When **DNS** was not into existence, one had to download a **Host file** containing host names and their corresponding IP address. But with increase in number of hosts of internet, the size of host file also increased. This resulted in increased traffic on downloading this file. To solve this problem the DNS system was introduced.

**Domain Name System** helps to resolve the host name to an address. It uses a hierarchical naming scheme and distributed database of IP addresses and associated names

**IP Address**

IP address is a unique logical address assigned to a machine over the network. An IP address exhibits the following properties:

* IP address is the unique address assigned to each host present on Internet.
* IP address is 32 bits (4 bytes) long.
* IP address consists of two components: **network component** and **host component**.
* Each of the 4 bytes is represented by a number from 0 to 255, separated with dots. For example 137.170.4.124

IP address is 32-bit number while on the other hand domain names are easy to remember names. For example, when we enter an email address we always enter a symbolic string such as webmaster@tutorialspoint.com.

**Uniform Resource Locator (URL)**

**Uniform Resource Locator (URL)** refers to a web address which uniquely identifies a document over the internet.

This document can be a web page, image, audio, video or anything else present on the web.

For example, **www.tutorialspoint.com/internet\_technology/index.html** is an URL to the index.html which is stored on tutorialspoint web server under internet\_technology directory.

**URL Types**

There are two forms of URL as listed below:

1. Absolute URL
2. Relative URL

**Absolute URL**

Absolute URL is a complete address of a resource on the web. This completed address comprises of protocol used, server name, path name and file name.

* **http** is the protocol.
* **tutorialspoint.com** is the server name.
* **index.htm** is the file name.

The protocol part tells the web browser how to handle the file. Similarly we have some other protocols also that can be used to create URL are:

* FTP
* https
* Gopher
* mailto
* news

**Relative URL**

Relative URL is a partial address of a webpage. Unlike absolute URL, the protocol and server part are omitted from relative URL.

Relative URLs are used for internal links i.e. to create links to file that are part of same website as the WebPages on which you are placing the link.

Difference between Absolute and Relative URL

|  |  |
| --- | --- |
| **Absolute URL** | **Relative URL** |
| Used to link web pages on different websites | Used to link web pages within the same website. |
| Difficult to manage. | Easy to Manage |
| Changes when the server name or directory name changes | Remains same even of we change the server name or directory name. |
| Take time to access | Comparatively faster to access. |

**Domain Name System Architecture**

The Domain name system comprises of **Domain Names, Domain Name Space, Name Server** that have been described below:

**Domain Names**

Domain Name is a symbolic string associated with an IP address. There are several domain names available; some of them are generic such as **com, edu, gov, net** etc, while some country level domain names such as **au, in, za, us** etc.

The following table shows the **Generic** Top-Level Domain names:

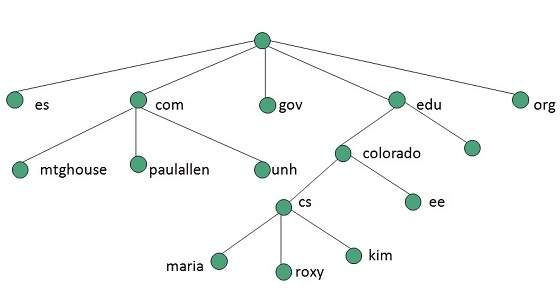
|  |
| --- |
|  |
| **Domain Name** | **Meaning** |
| Com | Commercial business |
| Edu | Education |
| Gov | U.S. government agency |
| Int | International entity |
| Mil | U.S. military |
| Net | Networking organization |
| Org | Non profit organization |

The following table shows the **Country top-level** domain names:

|  |
| --- |
|  |
| **Domain Name** | **Meaning** |
| Au | Australia |
| In | India |
| Cl | Chile |
| fr | France |
| us | United States |
| za | South Africa |
| uk | United Kingdom |
| jp | Japan |
| es | Spain |
| de | Germany |
| ca | Canada |
| ee | Estonia |
| hk | Hong Kong |

**Domain Name Space**

The domain name space refers a hierarchy in the internet naming structure. This hierarchy has multiple levels (from 0 to 127), with a root at the top. The following diagram shows the domain name space hierarchy:



In the above diagram each subtree represents a domain. Each domain can be partitioned into sub domains and these can be further partitioned and so on.

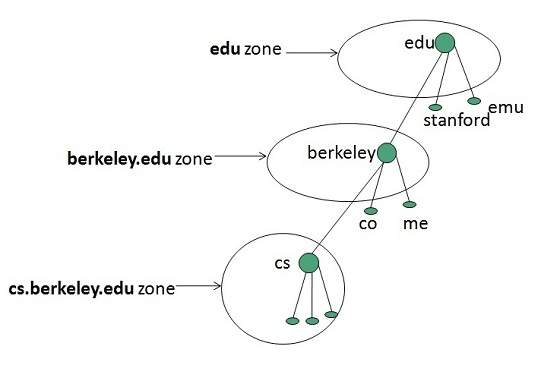
**Name Server**

Name server contains the DNS database. This database comprises of various names and their corresponding IP addresses. Since it is not possible for a single server to maintain entire DNS database, therefore, the information is distributed among many DNS servers.

* Hierarchy of server is same as hierarchy of names.
* The entire name space is divided into the zones

**Zones**

Zone is collection of nodes (sub domains) under the main domain. The server maintains a database called zone file for every zone.



If the domain is not further divided into sub domains then domain and zone refers to the same thing.

The information about the nodes in the sub domain is stored in the servers at the lower levels however; the original server keeps reference to these lower levels of servers.

**WORLD WIDE WEB (WWW)**

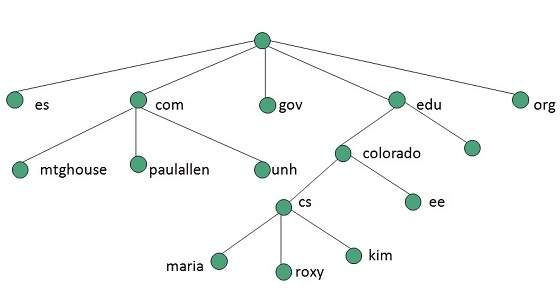
**WWW** stands for **World Wide Web.** A technical definition of the World Wide Web is : all the resources and users on the Internet that are using the Hypertext Transfer Protocol (HTTP).

A broader definition comes from the organization that Web inventor **Tim Berners-Lee** helped found, the **World Wide Web Consortium (W3C).**

The World Wide Web is the universe of network-accessible information, an embodiment of human knowledge.

In simple terms, The World Wide Web is a way of exchanging information between computers on the Internet, tying them together into a vast collection of interactive multimedia resources.

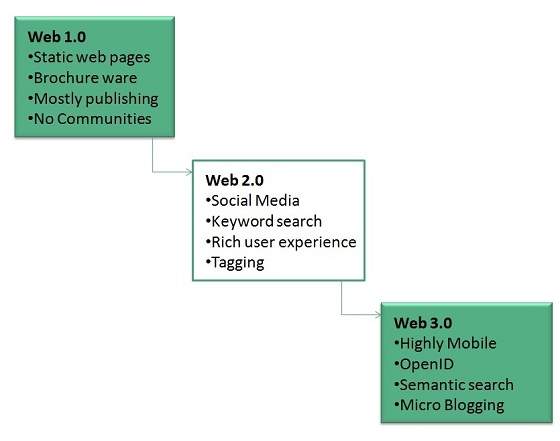
**Internet** and **Web** is not the same thing: Web uses internet to pass over the information.



**EVOLUTION**

**World Wide Web** was created by **Timothy Berners Lee** in 1989 at **CERN** in **Geneva.** World Wide Web came into existence as a proposal by him, to allow researchers to work together effectively and efficiently at **CERN.** Eventually it became **World Wide Web.**

The following diagram briefly defines evolution of World Wide Web:



**UNIFORM RESOURCE LOCATOR (URL)**

Every document on the Web has a unique address. This address is known as **U**niform **R**esource **L**ocator (URL).

Several HTML/XHTML tags include a URL attribute value, including hyperlinks, inline images, and forms. All of them use the same syntax to specify the location of a web resource, regardless of the type or content of that resource. That's why it is known a Uniform Resource Locator.

**BROWSERS**

**web Browser** is an application software that allows us to view and explore information on the web. User can request for any web page by just entering a URL into address bar.

Web browser can show text, audio, video, animation and more. It is the responsibility of a web browser to interpret text and commands contained in the web page.

Earlier the web browsers were text-based while now a days graphical-based or voice-based web browsers are also available. Following are the most common web browser available today:

|  |  |
| --- | --- |
| **Browser** | **Vendor** |
| Internet Explorer | Microsoft |
| Google Chrome | Google |
| Mozilla Firefox | Mozilla |
| Netscape Navigator | Netscape Communications Corp. |
| Opera | Opera Software |
| Safari | Apple |
| Sea Monkey | Mozilla Foundation |
| K-meleon | K-meleon |

Architecture

There are a lot of web browser available in the market. All of them interpret and display information on the screen however their capabilities and structure varies depending upon implementation. But the most basic component that all web browser must exhibit are listed below:

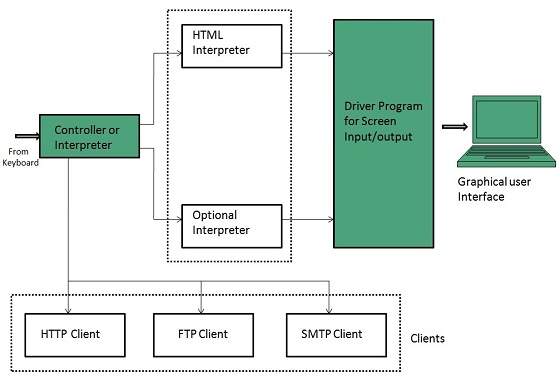
* Controller/Dispatcher
* Interpreter
* Client Programs

**Controller** works as a control unit in CPU. It takes input from the keyboard or mouse, interpret it and make other services to work on the basis of input it receives.

**Interpreter** receives the information from the controller and execute the instruction line by line. Some interpreter are mandatory while some are optional For example, HTML interpreter program is mandatory and java interpreter is optional.

**Client Program** describes the specific protocol that will be used to access a particular service. Following are the client programs tat are commonly used:

* HTTP
* SMTP
* FTP
* NNTP
* POP



**SEARCH ENGINE**

**Search Engine** refers to a huge database of internet resources such as web pages, newsgroups, programs, images etc. It helps to locate information on World Wide Web.

User can search for any information by passing query in form of keywords or phrase. It then searches for relevant information in its database and return to the user.



**Search Engine Components**

Generally there are three basic components of a search engine as listed below:

1. Web Crawler
2. Database
3. Search Interfaces

**Web crawler**

It is also known as **spider** or **bots.** It is a software component that traverses the web to gather information.

**Database**

All the information on the web is stored in database. It consists of huge web resources.

**Search Interfaces**

This component is an interface between user and the database. It helps the user to search through the database.

**Search Engine Working**

Web crawler, database and the search interface are the major component of a search engine that actually makes search engine to work. Search engines make use of Boolean expression AND, OR, NOT to restrict and widen the results of a search. Following are the steps that are performed by the search engine:

1. The search engine looks for the keyword in the index for predefined database instead of going directly to the web to search for the keyword.
2. It then uses software to search for the information in the database. This software component is known as web crawler.
3. Once web crawler finds the pages, the search engine then shows the relevant web pages as a result. These retrieved web pages generally include title of page, size of text portion, first several sentences etc.
4. These search criteria may vary from one search engine to the other. The retrieved information is ranked according to various factors such as frequency of keywords, relevancy of information, links etc.
5. User can click on any of the search results to open it.

**Examples**

Following are the several search engines available today:

|  |  |
| --- | --- |
| **Search Engine** | **Description** |
| Google | It was originally called **BackRub.** It is the most popular search engine globally. |
| Bing | It was launched in 2009 by **Microsoft.** It is the latest web-based search engine that also delivers Yahoo’s results. |
| Ask | It was launched in 1996 and was originally known as **Ask Jeeves.** It includes support for match, dictionary, and conversation question. |
| AltaVista | It was launched by **Digital Equipment Corporation** in 1995. Since 2003, it is powered by Yahoo technology. |
| AOL.Search | It is powered by Google. |
| LYCOS | It is top 5 internet portal and 13th largest online property according to Media Matrix. |
| Alexa | It is subsidiary of Amazon and used for providing website traffic information. |

**WEB SERVER**

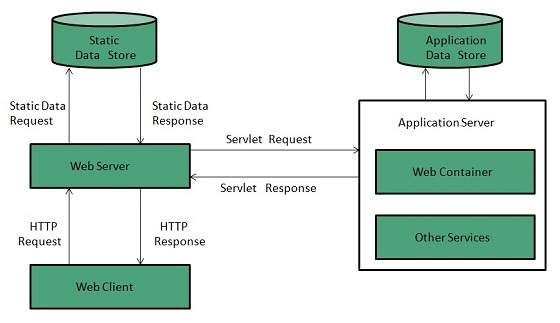
**Web server** is a computer where the web content is stored. Basically web server is used to host the web sites but there exists other web servers also such as gaming, storage, FTP, email etc.

Web site is collection of web pages whileweb server is a software that respond to the request for web resources.

**Web Server Working**

Web server respond to the client request in either of the following two ways:

* Sending the file to the client associated with the requested URL.
* Generating response by invoking a script and communicating with database



**Examples**

Following table describes the most leading web servers available today:

|  |  |
| --- | --- |
| **S.N.** | **Web Server Descriptino** |
| 1 | **Apache HTTP Server** This is the most popular web server in the world developed by the Apache Software Foundation. Apache web server is an open source software and can be installed on almost all operating systems including Linux, UNIX, Windows, FreeBSD, Mac OS X and more. About 60% of the web server machines run the Apache Web Server. |
| 2. | **Internet Information Services (IIS)** The Internet Information Server (IIS) is a high performance Web Server from Microsoft. This web server runs on Windows NT/2000 and 2003 platforms (and may be on upcoming new Windows version also). IIS comes bundled with Windows NT/2000 and 2003; Because IIS is tightly integrated with the operating system so it is relatively easy to administer it. |
| 3. | **Lighttpd** The lighttpd, pronounced lighty is also a free web server that is distributed with the FreeBSD operating system. This open source web server is fast, secure and consumes much less CPU power. Lighttpd can also run on Windows, Mac OS X, Linux and Solaris operating systems. |
| 4. | **Sun Java System Web Server** This web server from Sun Microsystems is suited for medium and large web sites. Though the server is free it is not open source. It however, runs on Windows, Linux and UNIX platforms. The Sun Java System web server supports various languages, scripts and technologies required for Web 2.0 such as JSP, Java Servlets, PHP, Perl, Python, and Ruby on Rails, ASP and Coldfusion etc. |
| 5. | **Jigsaw Server** Jigsaw (W3C's Server) comes from the World Wide Web Consortium. It is open source and free and can run on various platforms like Linux, UNIX, Windows, and Mac OS X Free BSD etc. Jigsaw has been written in Java and can run CGI scripts and PHP programs. |

**PROXY SERVER**

**Proxy server** is an intermediary server between client and the interner. Proxy servers offers the following basic functionalities:

* Firewall and network data filtering.
* Network connection sharing
* Data caching

Proxy servers allow to hide, conceal and make your network id anonymous by hiding your IP address.

**Purpose of Proxy Servers**

Following are the reasons to use proxy servers:

**Monitoring and Filtering**

Proxy servers allow us to do several kind of filtering such as:

* Content Filtering
* Filttering encrypted data
* Bypass filters
* Logging and eavasdropping

**Improving performance**

It fasten the service by process of retrieving content from the cache which was saved when previous request was made by the client.

**Translation**

It helps to customize the source site for local users by excluding source content or substituting source content with original local content. In this the traffic from the global users is routed to the source website through Translation proxy.

**Accessing services anonymously**

In this the destination server receives the request from the anonymzing proxy server and thus does not receive information about the end user.

**Security**

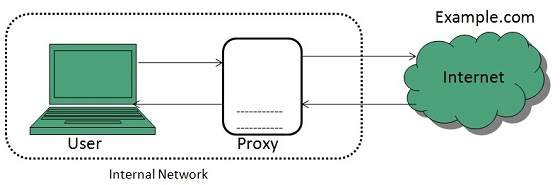
Since the proxy server hides the identity of the user hence it protects from spam and the hacker attacks.

**Type of Proxies**

Following table briefly describes the type of proxies:

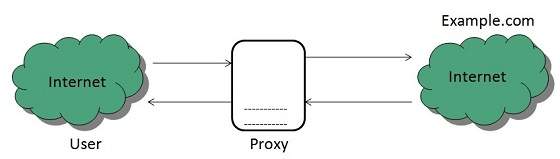
**Forward Proxies**

In this the client requests its internal network server to forward to the internet.



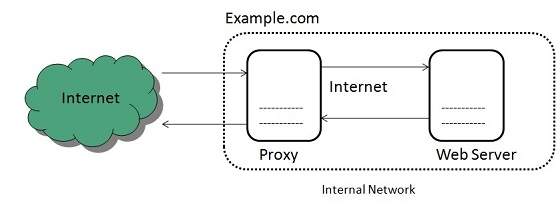
**Open Proxies**

Open Proxies helps the clients to conceal their IP address while browsing the web.



**Reverse Proxies**

In this the requests are forwarded to one or more proxy servers and the response from the proxy server is retrieved as if it came directly from the original Server.



**HTTP PROTOCOL**

The Hypertext Transfer Protocol (HTTP) is an application-level protocol for distributed, collaborative, hypermedia information systems. This is the foundation for data communication for the World Wide Web (i.e. internet) since 1990. HTTP is a generic and stateless protocol which can be used for other purposes as well using extensions of its request methods, error codes, and headers.

Basically, HTTP is a TCP/IP based communication protocol, that is used to deliver data (HTML files, image files, query results, etc.) on the World Wide Web. The default port is TCP 80, but other ports can be used as well. It provides a standardized way for computers to communicate with each other. HTTP specification specifies how clients' request data will be constructed and sent to the server, and how the servers respond to these requests.

Basic Features

There are three basic features that make HTTP a simple but powerful protocol:

* **HTTP is connectionless:** The HTTP client, i.e., a browser initiates an HTTP request and after a request is made, the client disconnects from the server and waits for a response. The server processes the request and re-establishes the connection with the client to send a response back.
* **HTTP is media independent:** It means, any type of data can be sent by HTTP as long as both the client and the server know how to handle the data content. It is required for the client as well as the server to specify the content type using appropriate MIME-type.
* **HTTP is stateless:** As mentioned above, HTTP is connectionless and it is a direct result of HTTP being a stateless protocol. The server and client are aware of each other only during a current request. Afterwards, both of them forget about each other. Due to this nature of the protocol, neither the client nor the browser can retain information between different requests across the web pages.

HTTP/1.0 uses a new connection for each request/response exchange, where as HTTP/1.1 connection may be used for one or more request/response exchanges.

**Basic Architecture**

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.

**Client**

The HTTP client sends a request to the server in the form of a request method, URI, and protocol version, followed by a MIME-like message containing request modifiers, client information, and possible body content over a TCP/IP connection.

**Server**

The HTTP server responds with a status line, including the message's protocol version and a success or error code, followed by a MIME-like message containing server information, entity meta information, and possible entity-body content.

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

