UNIT-1

INTERNET BASICS

DATA COMMUNICATIONS : Data communications are the exchange of data between two devices via some form of transmission medium such as a wire cable.

Components of Data Communication:

The different components of Data communication are shown in the following figure.

1. ***Message:*** The message is the information (data) to be communicated. Popular forms of information include text, numbers, pictures, audio, and video.

2. ***Sender****:* The sender is the device that sends the data message. It can be a computer, workstation, telephone handset, video camera, and so on.

3. ***Receiver:*** The receiver is the device that receives the message. It can be a computer, workstation, telephone handset, television, and so on.

4. ***Transmission medium****:* The transmission medium is the physical path by which a message travels from sender to receiver. Some examples of transmission media include twisted-pair wire, coaxial cable, fiber-optic cable, and radio waves.

5. Protocol: A protocol is a set of rules that govern data communications. It represents an agreement between the communicating devices.

Computer Networks :

- A network is a set of nodes connected by communication links.
- A node can be a computer, printer, or any other device capable of sending and/or receiving data generated by other nodes on the network.

Categories of networks

- Network divided in to three primary categories: LAN, MAN, WAN. In to which category a network falls is determined by its Size, Ownership, Distance it covers, and Physical architecture

1. LOCAL-AREA NETWORK(LAN) :

- LAN is usually Privately owned and Links devices in single office, building or campus.
- LAN size is Limited to few kilometres.
- LANs are designed to allow resources (i.e. hardware or software) to be shared between PCs and workstations.
- LAN will use a single transmission media.
- The most common LAN Topologies are Ring, bus, star.

2.METROPOLITAN-AREA NETWORK (MAN):

- A MAN is designed to extend over an entire city.
- It may be single network such as cable television network, or it may be a means of connecting number of LANs in to a larger networks.
- A MAN be wholly Owned and operated by a private company, or it may be a Service provider by Public company such as a local telephone company.

1. **WIDE-AREA NETWORK(WAN):**

WAN provides long-transmission of data, voice, image and video information over large geographic areas that may comprise a country, a continent or even the whole world.

WAN that is wholly owned and used by a single company is often referred to as an enterprise network.

Application of Computer Network:

- Sharing of resources such as printers
- Sharing of expensive software's and database
- Communication from one computer to another computer
- Exchange of data and information among users via network
- Sharing of informationover geographically wide areas.

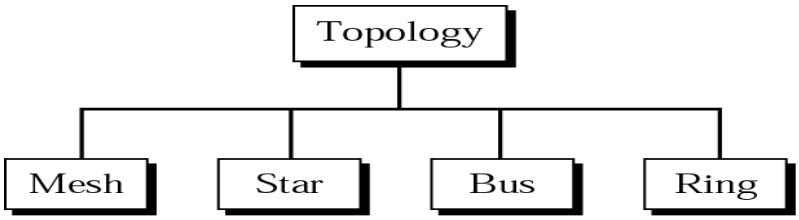
TOPOLOGY:

1.Topology refers to the way in which a network is laid out physically.

2. The topology of a network is the geometric representation of the relationship of all the links and linking devices (usually called nodes) to one another.

Categories of topology:

There are four basic topologies possible.



1.Mesh topology:

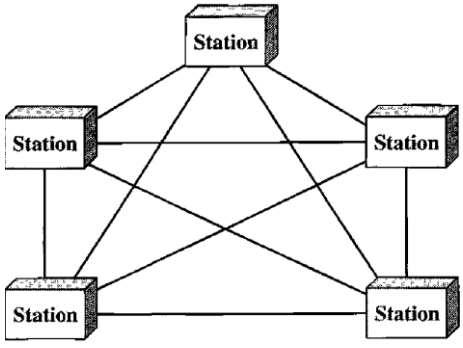
1. In a mesh topology, every device has a dedicated point-to-point link to every other device.

2. The term *dedicated* means that the link carries traffic only between the two devices it connects.

3. A fully connected mesh network therefore has ***n(n -1) /2*** physical channels link n devices.

To accommodate that many links, every device on the network must have *n* - 1

input/output (*I/O)* ports to be connected to the other *n* - 1 stations.



ADVANTAGES:

1. The use of dedicated links guarantees that each connection can carry its own data load, thus eliminating the traffic problems
2. A mesh topology is robust. i.e. If one link becomes unusable, it does not incapacitate the entire system.

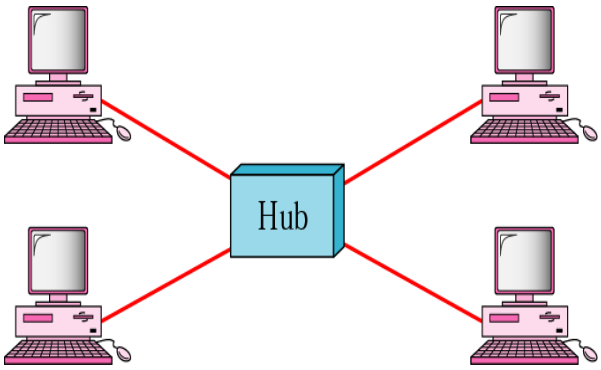
3. There is the advantage of privacy or security.
4. point-to-point links make fault identification and fault isolation easy.

DISADVANTAGES:

1. Because every device must be connected to every other device, installation and reconnection are difficult.
2. The bulk of the wiring can be greater than the available space (in walls, ceilings, or floors) can accommodate.

2.Star Topology:

1. In a star topology, each device has a dedicated point-to-point link only to a central controller, usually called a hub.
2. The devices are not directly linked to one another.
3. A star topology does not allow direct traffic between devices. The controller acts as an exchange: If one device wants to send data to another, it sends the data to the controller, which then relays the data to the other connected devices.



ADVANTAGES:

1. A star topology is less expensive than a mesh topology.
2. It is easy to install and reconfigure.
3. Other advantages include robustness

DISADVANTAGE:

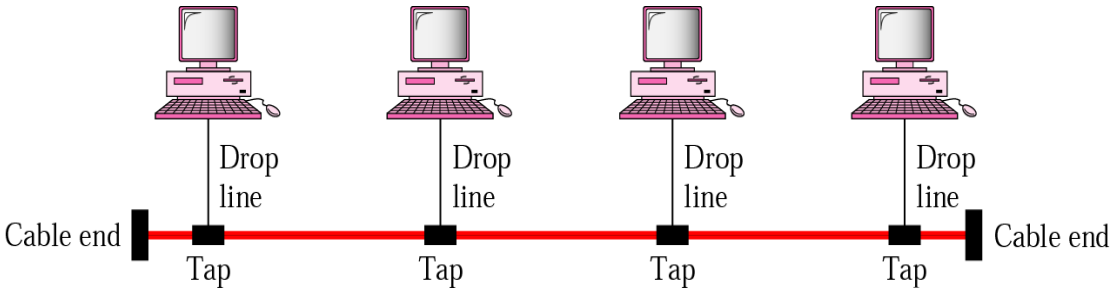
The dependency of the whole topology on one single point, the hub. If the hub goes down, the whole system is dead.

3.Bus Topology:

1. A bus topology, is multipoint connected . One long cable acts as a backbone to link all the devices in a network.

2. Nodes are connected to the bus cable by drop lines and taps. A drop line is a connection

Running between the device and the main cable. A tap is a connector that either splices into the main cable or punctures the sheathing of a cable to create a contact with the metallic core.



ADVANTAGES

1. Advantages of a bus topology include ease of installation

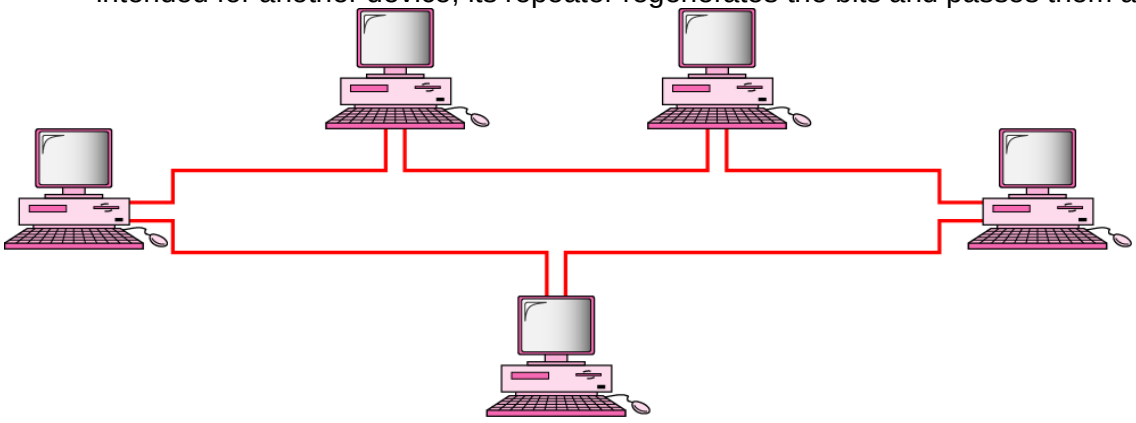
2. Bus uses less cabling.

DISADVANTAGES:

1. Difficult reconnection and fault isolation is also difficult.
2. Signal reflection at the taps can cause degradation in quality.

4.Ring topology:

- In a ring topology, each device has a dedicated point-to-point connection with only the two devices on either side of it.
- A signal is passed along the ring in one direction, from device to device, until it reaches its destination.
- Each device in the ring incorporates a repeater. When a device receives a signal intended for another device, its repeater regenerates the bits and passes them along.



ADVANTAGE:

1. A ring is relatively easy to install and reconfigure.
2. Fault isolation is simplified.

DISADVANTAGE: Unidirectional traffic can be a disadvantage.

INTERNET

Internet is a worldwide network of networks that uses the standard Internet protocol suite (TCP/IP) to link several billion devices worldwide.

Hardware and software Requirements

To connect to the Internet we need the following four things:

1. A computer
2. A modem and telephone line (if you are using dial up access)
   A data line of some sort (if you are not using dial up access)
3. An Internet browser (software) and software to connect you to the ISP
4. An account with an Internet Service Provider (ISP) Souvenirs

Applications of Internet:

- Sending and receiving email
- Searching and browsing information archives
- Copying files between computers
- Conducting financial transactions
- Navigating (in your car, smart scooter, smart bike, or other)
- Playing interactive games.
- Video and music streaming
- Chat or voice communication (direct messaging, video conferencing) etc.

Intranet

An intranet is a computer network for sharing information, collaboration tools, operational systems, and other computing services within an organization, usually to the exclusion of access by outsiders.

Uses of the intranet:

- Streamlining everyday activities by making repeated tasks more feasible.
- Centralizing and managing important information and company data in a single database.
- Making collaboration easier since information can be shared across the entire network.
- Providing personalized content to employees based on their role within the company.
- Improving internal communication by making employee directories, company news and organization charts readily available.
- Providing fast and easy access to information about company policies, benefits and updates.

Extranet:

The extranet is a private network that uses the internet that allows people outside a business partners, vendors or authorized customers to access business information.

Parameter	Internet	Intranet
Usage	Public	Private
User Types	Any user having dial up of Internet access line.	Organization employees and Internal company departments
Usage	Access all kind of information	Internal employee communication , telephone directories etc.
Security	Low security. Configured under 0 security level in firewall	High security. Configured under 100 security level in firewall
Regulated by	Internet Architecture Board (IAB): Oversees the technical and engineering development of the IETF and IRTF. Internet Corporation for Assigned Names and Numbers (ICANN).	It is regulated by an organization.
Coverage	Wide Area	Within an organization
Access	Large number of users	Limited number of users
System failure	Unpredictable	System availability is high since system is monitored by authority

MODEM

Modem, ("*mo*dulator/*dem*odulator") is electronic device that convert digital data signals into modulated analog signals suitable for transmission over analog telecommunications circuits.

Working principle of MODEM

Modulator:

• This unit is used to convert the digital data from computer into analog data This process is called modulation

• This is done by adding a carrier signal to the digital signal.

Demodulator:

• This unit is used to convert the analog data from telephone system into digital data

• This is done by eliminating the carrier signal from analog signal

Types of modems

Two types of modem

- **Internal modem:** Internal Modem is modem that plugged directly into the CPU. Physically internal modem in the form of a card that is plugged into one of the expansion slots on the mainboard, usually on the ISA or PCI slot.
- **External modem:** External Modem is modem that installed outside of the CPU. External modem connected to the CPU via the COM port or USB. This type of modem typically uses separate voltage source in the form of an adapter.

Difference between Internal and External Modem

Internal Modem	External Modem
An internal modem is a modem that fits inside of a computer. Internal modems typically ship with the computer and come pre-installed.	he external modem sits outside the computer. The external modem can be used when a computer is unable to fit an internal modem inside of it.
Low in Price.	Comparatively high in price.
No external accessory has to buy.	In an external modem, RS232 interface cable has to buy.
It is difficult to move the internal modem to another computer.	The external modem can be moved easily.
The internal modem is powered by PC.	The external modem needs plugs into the wall to power on.

Features of Modems:

- **Speed**: The speed at which the modem can send data in bps (bits per second). Typically modem speeds are: 300, 600, 1200, 2400, 4800, 9600, 14.4K, 19.2K, 28.8K bps
- **Auto Dial /Redial**: Smart Modems can dial the phone number and & auto redial if a busy signal is received.
- **Auto Answer**: Most modems can automatically answer the phone when an incoming call comes in. They have Ring Detect capability.
- **Self-Testing:** New modems have self-testing features. They can test the digital connection to the terminal /computer and the analog connection to a remote modem. They can also check the modem's internal electronics.
- **Voice over Data:** Voice over Data modems allow a voice conversation to take place while data is being transmitted. This requires both the source and destination modems to have this feature.
- **Synchronous or Asynchronous Transmission:** Newer modems allow a choice of synchronous or asynchronous transmission of data. Normally, modem transmission is asynchronous. We send individual characters with just start and stop bits. Synchronous transmission or packet transmission is used in specific applications.

Functions of Modems:

- **Data Compression:** Data compression is the ability of the modem to take data in from the computer, reduce it in volume, and then send it out via the modem.

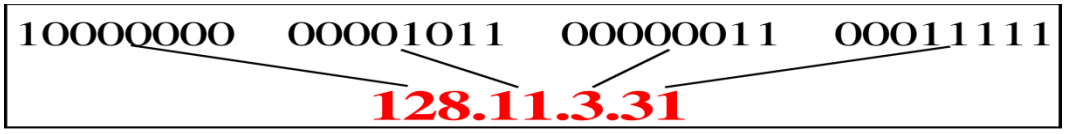
- **Error correction:** Error correction standards provide a way of correcting errors that result from outside interference, such as noise on the phone line. Error correction ensures that data coming out of the receiving modem is exactly the same as data going into the sending modem.
- **Flow Control:** Individual modems send information at different speeds. It's necessary for faster modems to slow down so that slower modems can catch up, otherwise the slower modem will receive more data than it can process.

IP ADDRESS :

- An IP (Internet Protocol) address is a 32 bit binary number which uniquely identify a node or host connection on an IP network.
- Each Internet address consists of 4 bytes (32-bits) defining 3 fields: class type, network identification (netid) and host identification (hosted).
- These parts are of varying lengths, depending on the class of the address

Dotted decimal notation:

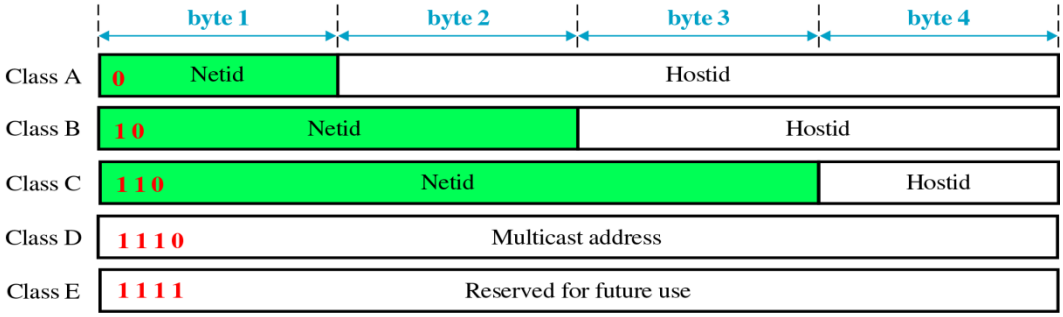
- Machines read the IP address as a stream of 32 bits.
- However, for human consumption, the IP address is written in dotted decimal notation.
  – The 32-bit address is divided into 4 groups of 8 bits (an octet or a byte).
  – Each octet is written as a decimal number ranging from 0 to 255.
  – The decimal numbers are separated by periods, or dots.



ADDRESS CLASSES

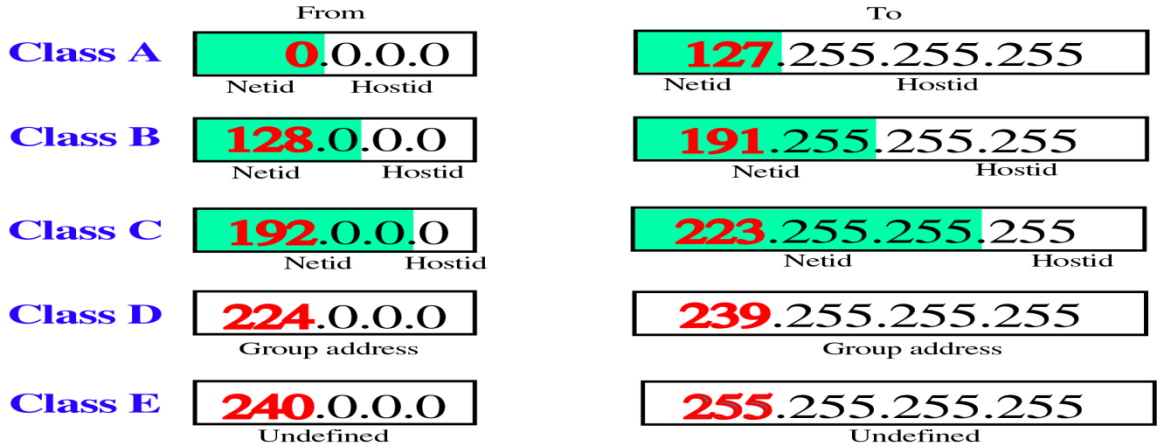
The designers of the Internet decided to create classes of networks based on network size. For the small number of networks possessing a very large number of nodes, they created the rank *Class A network*. At the other extreme is the *Class C network*, reserved for the numerous networks with a small number of nodes. The class distinction for networks in between very large and very small is predictably called a *Class B network*. How one would subdivide an IP address into a network and node address is determined by the class designation of one's network.

- There are 5 different address classes (A, B,C,D and E). We can determine which class any IP address is in by examining the first 4 bits of the IP address.



- Class A – C widely used. Class D for multicasting (allows copies of a datagram to be passed to select group of hosts rather than to an individual host.) and class E for future use

Class Ranges of Internet Addresses



Class A addresses were designed for large organizations with a large number of attached hosts or routers. Class B addresses were designed for midsize organizations with tens of thousands of attached hosts or routers. Class C addresses were designed for small organizations with a small number of attached hosts or routers.

Flaw in Classful Address

- A block in class A address is too large for almost any organization
- A block in class B is also very large, probably too large for any of the organizations that received a class B block.
- A block in class C is probably too small.
- A and B always wasted. But C is always not enough
- In Overall, classful addressing, a large part of the available addresses were wasted.

Domain Name System (DNS)

- DNS Stands for ***D**omain **N**ame **S**ystem* (or ***S**ervice* or ***S**erver*).
- It is an Internet service that translates *domain names* into IP addresses.
- Because domain names are alphabetic, they're easier to remember. The Internet however, is really based on IP addresses. Every time you use a domain name, therefore, a DNS service must translate the name into the corresponding IP address. For example, the domain name *www.example.com* might translate to *198.105.232.4*.

Three main components of DNS

1. Name resolver
2. Name server
3. database of resource records(RRs)

DNS resolver

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

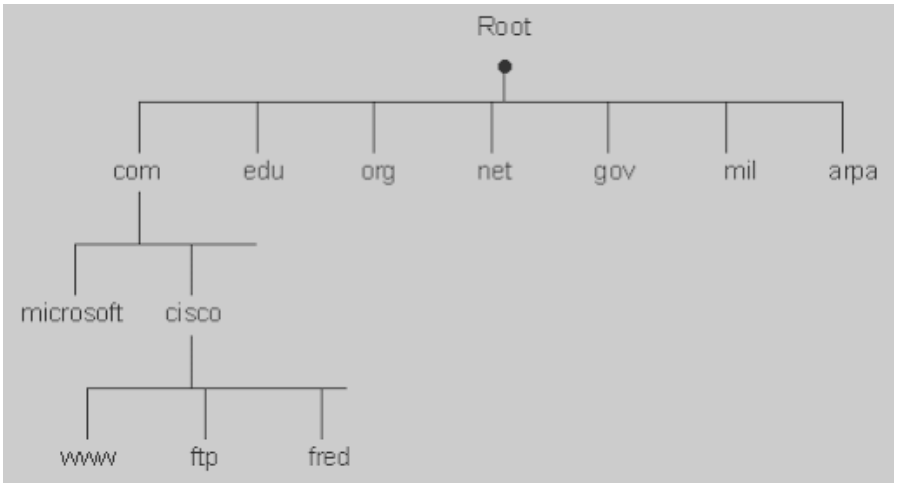
Name Sever

Server responsible for answering DNS queries

- Exists at all levels of hierarchy
- Authoritative name servers hold part of the DNS database
- One name server can serve more then one zone
- Many name servers “should” serve the same zone
- Some name servers are authoritative for certain zones

STRUCTURE

The structure of DNS is hierarchical or tree structure. At the top node is called the root and it is the start of all other branches in the DNS tree. It is designated with a period. Each branch moves down from level to level. When referring to DNS addresses, they are referred to from the bottom up with the root designator (period) at the far right. Example: "myhost.mycompany.com.".



DNS is hierarchical in structure. A domain is a subtree of the domain name space. From the root, the assigned top-level domains in the U.S. are:

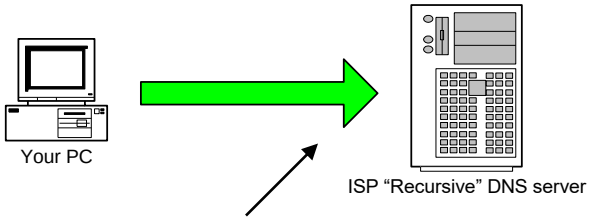
- GOV - Government body.
- EDU - Educational body.
- INT - International organization
- NET - Networks
- COM - Commercial entity.
- MIL - U. S. Military.
- ORG - Any other organization not previously listed.

MAPPING OF DOMAIN NAME TO IP ADDRESS

The domain mechanism for making name to addresses consists of a mane server. In name server there is a server application does the work of mapping of domain name to IP address. When the name is translated from client to name server , the task is performed by a client software called name resolver.

FOR EXAMPLE:-

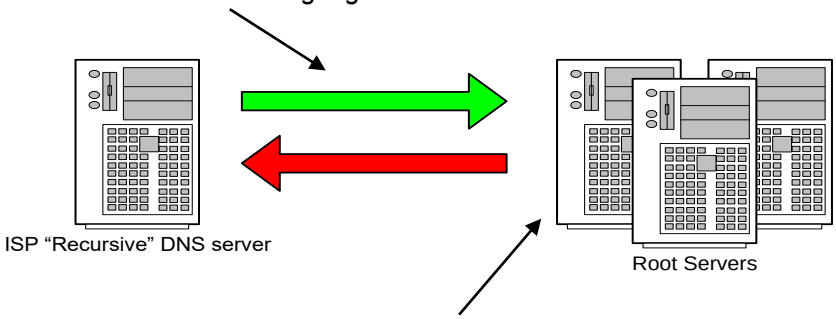
Step 1: Your PC sends a resolution request to its configured DNS Server, typically at your ISP.



Tell me the Address of “www.google.com

Step 2: Your ISPs recursive name server starts by asking one of the root servers predefined in its "hints" file

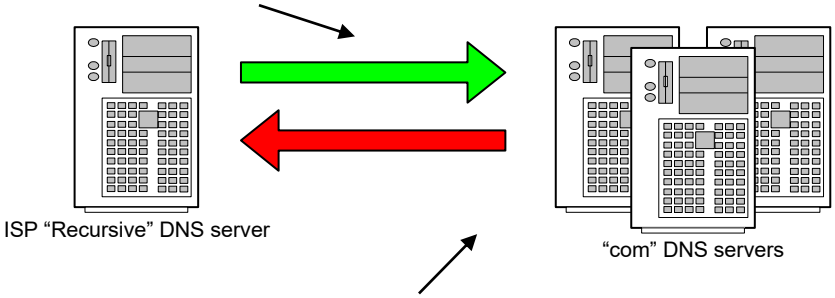
Tell me the Address of "www.google.com"



I don't know the address but I know who's authoritative for the"com" domain ask them

Step 3: Your ISPs recursive name server then asks one of the "com" name servers as directed.

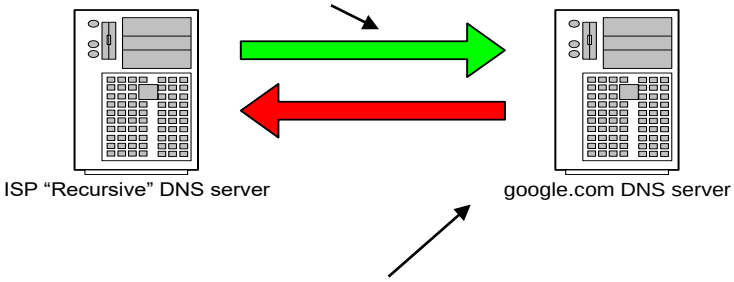
Tell me the Address of" www.google.com "



I don't know the address but I know who's authoritative for the "google.com" domain ask them

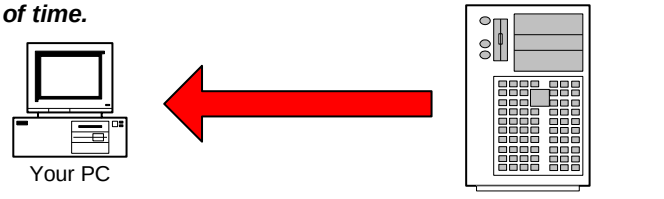
Step 4: Your ISPs recursive name server then asks one of the "google.com" name servers as directed.

Tell me the Address of "www.google.com"



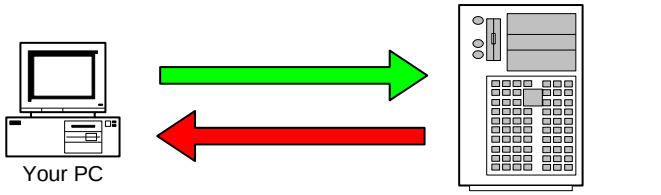
The Address of www.google.com is 216.239.53.99

Step 5: ISP DNS server then send the answer back to your PC. The DNS server will "remember" the answer for a period of time.



The Address of www.google.com is 216.239.53.99

Step 6: Your PC can then make the actual HTTP request to the web server.



Here it is!

INTERNET SERVICE PROVIDER

An Internet service provider (ISP) is an organization that provides a myriad of services for accessing, using, or participating in the Internet.

Factors to consider when choosing ISP

- **Bandwidth :** Data transferring speed provided by ISP Company**.**
- **Availability :** Availiability of Network & performance to its users
- **COST** : Refers pricing of the connection as well as services
- **Network security :** It is an important issue related to the network over the Internet. Everyone has its own private information being safe on their servers & nodes.
- **Customer Services** : Better Customer service is highly required an ISP.
- **Location & they need for speed** : It is an also a important factor, when we looking an Internet provider is the location where we live or work. A better location refers a good level of customers support.

CIDR Notation

CIDR, which stands for Classless Inter-Domain Routing, is an IP addressing scheme that improves the allocation of IP addresses. It replaces the old system based on classes A, B, and C.

How does CIDR work?

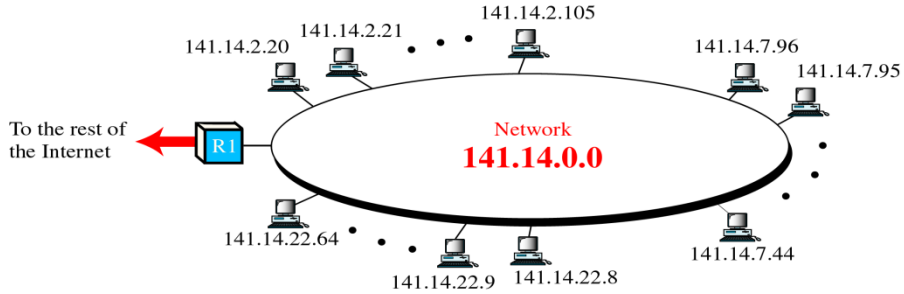
CIDR is based on variable-length subnet masking (VLSM). This allows it to define prefixes of arbitrary lengths making it much more efficient than the old system. CIDR IP addresses are composed of two sets of numbers. The network address is written as a prefix, like you would see a normal IP address (e.g. 192.255.255.255). The second part is the suffix which indicates how many bits are in the entire address (e.g. /12). Putting it together, a CIDR IP address would look like the following:

192.255.255.255/12

The network prefix is also specified as part of the IP address. This varies depending upon the number of bits required. Therefore, taking the example above, we can say that the first 12 bits are the network part of the address while the last 20 bits are for host addresses.

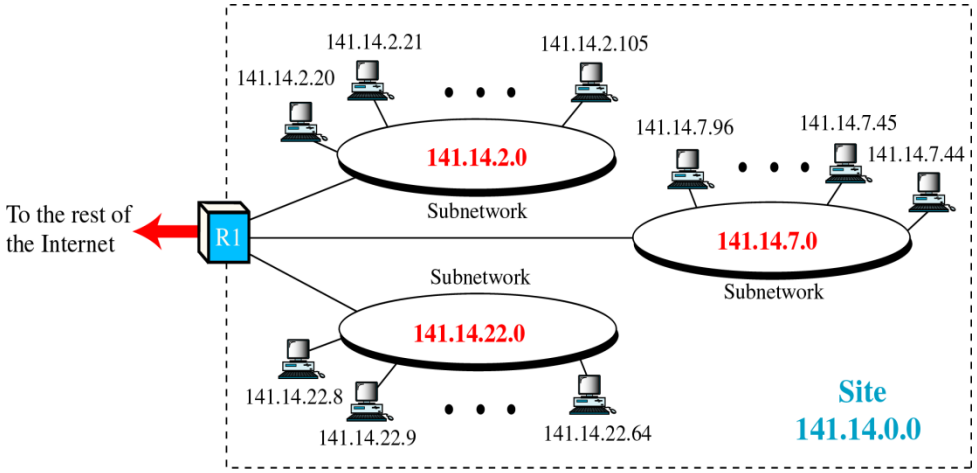
Subnetting

- Basically without subnetting, most of organization is limited to two levels of hierarchy

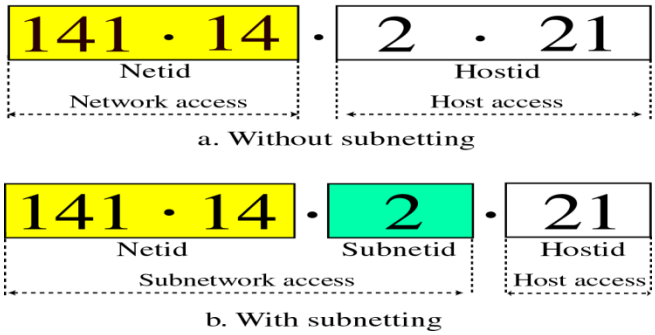


(**A Network with Two Levels of Hierarchy**)

- In this case, the hosts cannot be organized into groups, and all of the hosts are at the same level.
- As a result the organization has one network with many many hosts .
- To make a network more organize, three levels of hierarchy is implemented.
- Subnetting creates an intermediate level of hierarchy in the IP addressing system. Now we have 3 levels: netid, subnetid, and hostid.



(**A Network with Three Levels of Hierarchy**)



(**Addresses with and without Subnetting**)

Masking

- A process that extracts the address of the physical network from an IP address.
- Masking can be done whether we have subnetting or not.
- If we have not subnetted the network, masking extracts the network address from an IP address.
- If we have subnetted, masking extracts the subnetwork address from an IP address.

7) **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.

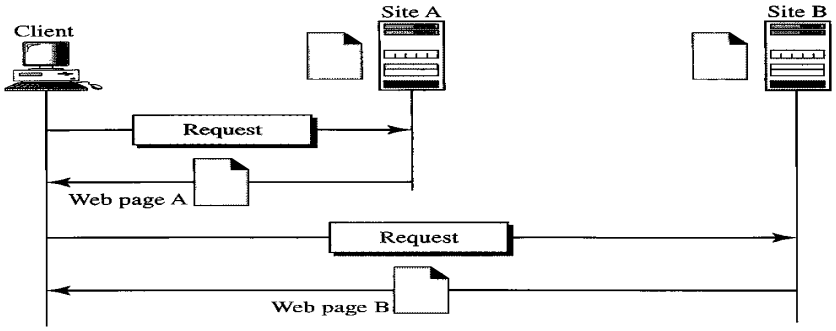
WORLD WIDE WEB (WWW)

- World Wide Web (WWW) is global, seamless environment in which all information (text, image, audio, video, computational service) are accessible in a consistent and simple way by using a standard set of naming and access convention from internet.
- WWW was initially developed by Tim Berners-Lee and others at CERN, Switzerland.
- It is also called WWW or W3 or Web.

ARCHITECTURE:

- The WWW today is a distributed client/server service, in which a client using a browser

  can access a service using a server. However, the service provided is distributed over many locations called *sites.*



- Each site holds one or more documents, referred to as *Web pages.* Each Web page can contain a link to other pages in the same site or at other sites. The pages can be retrieved

  and viewed by using browsers.

Principles of WWW:

- There would be no central control. The web works because people work within the agreed-to guidelines. As part of this the web ethic is that any one publish, and anyone(authorized user) can read information.
- All web servers would use the same protocols/mechanism etc.
- *http* is a fast, stateless, extensible transport mechanism would be used to communicate within the Web.
- httpd or http daemons, would be the base server which receive messages and providing data as requested.
- URLs would be used for network-wide addressing.
- All web browsers would use the same basic language- HTML.

Features of WWW:

The Web has unique features:

1. It is grand globally distributed internet service.
2. It is platform independent.
3. It has GUI features with easy navigability through hypertext and multimedia links.
4. It is interactive, because of search engine has push technology features like Web casting etc.
5. It is continuously expandable and each updated almost by a minute.
6. By processing the Web contents using CGI, Servlets, JavaScript, ASP, and JSP. The services for e-mail, e-com, and m-com are also vi

APPLICATION LAYER PROTOCOLS

There are several protocols which work for users in Application Layer. Application layer protocols can be broadly divided into two categories:

• Protocols which are used by users. For email for example, E-Mail

• Protocols which help and support protocols used by users. For example DNS

Few of Application layer protocols are described below:

1. Hyper Text Transfer Protocol (HTTP) :

The HyperText Transfer Protocol (HTTP) is used to define how the client server programs can be written to retrieve web pages from the Web.

• An HTTP client sends a request; an HTTP server returns a response. The server uses the port number 80

• HTTP uses the services of TCP. This means that, before any transaction between the client and the server can take place, a connection needs to be established between them. After the transaction, the connection should be terminated.

2. Domain Name System:

- DNS Stands for Domain Name System (or Service or Server).
- It is an Internet service that translates domain names into IP addresses.
- Because domain names are alphabetic, they're easier to remember. The Internet however, is really based on IP addresses. Every time you use a domain name, therefore, a DNS service must translate the name into the corresponding IP address. For example, the domain name www.example.com might translate to 198.105.232.4.

3. **File Transfer Protocol:**

- File Transfer Protocol (FTP) is the standard mechanism provided by TCP/IP for copying a file from one host to another or transferring files from one system to another.
- It is not only a protocol but also a service as well as application.
- Some problems in transferring files are two systems may use different file name conventions, two systems may have different ways to represent text and data, two systems may have different directory structures. All these problems have been solved by FTP in a very simple and elegant approach.

4**. Simple Mail Transfer Protocol :**

- This is the method protocol computer use to send messages by message transfer agents (MTA) on the internet.
- MTA are client & server programs that perform e-mail services such as sending or receiving mail for a host computer.

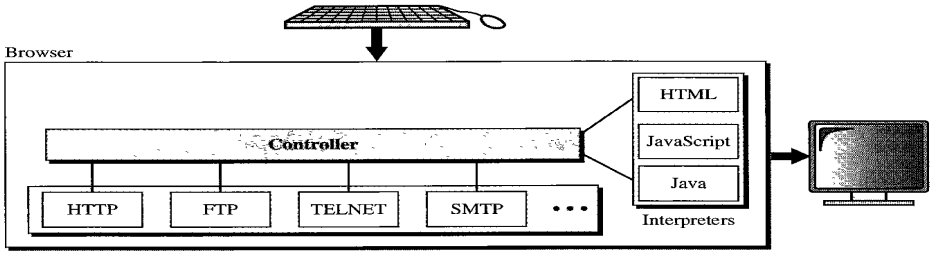
5. **TELNET:**

- Telnet is a protocol or set of rules that enables one computer to connect another computer. This process is also known as remote login.
- The user computer which initiates connection is referred as local computer and the machine being connected to which accepts the connection is referred as the remote or host computer.

- To start telnet session you must login to the server by entering a valid user name and password.
- Telnet is both TCP/IP application and a protocol for connecting a local computer to a remote computer.
- Telnet operates on client server principle. The local computer uses a telnet client program to establish the connection and display data on local computer monitor.

Web Browser

- Browsers are world Wide Web client side software that enables the user to access resources on the web.
- Each browser usually consists of three parts: a controller, client protocol, and interpreters.
- The controller receives input from the keyboard or the mouse and uses the client programs to access the document.



- After the document has been accessed, the controller uses one of the interpreters to display the document on the screen.
- The interpreter can be HTML, Java, or JavaScript, depending on the type of document.

Features of Web Browser

- The most popular browsers have a number of features in common. They allow users to set bookmarks and browse in a private mode. They also can be customized with extensions, and some of them provide a sync service.
- Most browsers have these user interface features:
- Allow the user to open multiple pages at the same time, either in different browser windows or in different tabs of the same window
- Back and forward buttons to go back to the previous page visited or forward to the next one.
- A refresh or reload button to reload the current page.
- As top button to cancel loading the page. (In some browsers, the stop button is merged with the reload button.)
- A home button to return to the user's home page.
- An address bar to input the URL of a page and display it.
- A search bar to input terms into a search engine. (In some browsers, the search bar is merged with the address bar.)
-
- One of the main feature of a browser is to search the information on the current page as well as search the WWW itself.
- Browser give you the facility to save a Web page in a file on your com pull print a Web page on your computer, and send the contents of a Web page e-Mail to others on the Internet.
- Few Web browser are complete Internet package, means they come with components like e-Mail client, newsgroup client an HTML composer, telnet client, ftp client, etc.
- Web browser should be able to handle text, images of the World Wide Web, as well as the hyperlinks to digital video, or other types of information.

How Browsers work?

World Wide Web works on the client server model. A user computer works as a client which can receive and send data to the server. When a web page is requested by a user, the browser contacts the requested server (where the website is stored) and by fetching and interpreting the requested files, it displays the web page on the computer screen.

The whole process takes place in these three steps:

1. **Contact to DNS Server**: When a user enters a URL into the address bar and hits enter '', at first browser contacts the DNS server. A DNS server stores the IP addresses of the server associated with the corresponding domain names. The DNS server takes the domain name from the browser and returns the corresponding IP address to the browser.

2. Contact to Server: After getting the IP address of the server for the requested webpage, browser sends a request to that server for the desired files.

3 Rendering :The entire process followed by a browser from fetching the webpage to displaying it on the screen is called Rendering.

Uniform Resource Locator (URL)

- ***U**niform **R**esource **L**ocator is* the global address of documents and other resources on the World Wide Web.
- The URL defines four things: protocol, host computer, port, and path.

- The *protocol* is the client/server program used to retrieve the document. Many different protocols can retrieve a document; among them are FTP or HTTP.
- The host is the computer on which the information is located, although the name of the computer can be an alias. Web pages are usually stored in computers, and computers are given alias names that usually begin with the characters "www".
- The URL can optionally contain the port number of the server. If the *port* is included, it is inserted between the host and the path, and it is separated from the host by a colon.
- Path is the pathname of the file where the information is located. Note that the path can itself contain slashes.

Example:

http://en.wikipedia.org/wiki/File:Raster_to_Vector_Mechanical_Example.jpg

Absolute vs. relative URLs:

An absolute URL is one that completely specifies the desired resource starting from the root of the resource name space. It is unique, meaning that if two absolute URLs are identical, they point to the same resource. An example is:

http://en.wikipedia.org/wiki/File:Raster_to_Vector_Mechanical_Example.jpg

A relative URL points to the location of a resource relative to a base URL. It is preceded by two dots *(../directory_path/file.txt)* for the directory above, one dot *(./directory_path/file.txt)* for the current directory or without the beginning slash (directory_path/file.txt) which is also the current directory. No dots *(/directory_path/file.txt)* for the root directory or domain.

Which results to

http://www.webreference.com/directory_path/file.txt

URI

- Uniform Resource Identifier (URI) is a string of characters used to identify a resource

- Such identification enables interaction with representations of the resource over a network, typically the World Wide Web, using specific protocols.

- Schemes specifying a concrete syntax and associated protocols define each URI

- To guarantee uniformity, all URIs follow a predefined set of syntax rules, but also maintain extensibility through a separately defined hierarchical naming scheme (e.g.http://).

URN

URN stands for Uniform Resource Name. URN is also the subset of URI. One of the best examples of URN is ISBN number which is used to uniquely identify a book. URN is completely different than URL as it doesn't include any protocol.

Difference between URL and URI:

URL	URI
URL is used to describe the identity of an item.	URI provides a technique for defining the identity of an item.
URL links a web page, a component of a web page or a program on a web page with the help of accessing methods like protocols.	URI is used to distinguish one resource from other regardless of the method used.
URL provides the details about what type of protocol is to be used.	URI doesn't contains the protocol specification.
URL is a type of URI.	URI is the superset of URL.

Hypertext

- Hypertext is text which contains links to other texts.
- By clicking on a link in a hypertext document, a user can quickly jump to different content.
- The term was coined by Ted Nelson around 1965.

Characteristics of good hypertext

- **Lots of documents:** Much of the hypertext's power comes from its ability to make large quantities of information accessible. If all the text in your system can be printed on ten pages, it would be just as simple to read through it from beginning to end and forget all this hypertext silliness.

•**Lots of links:**If each document has just one link, then it is little more than normal, sequential text. A hypertext document should present the reader with several links, offering a choice about where to go next. Ideally, a document should present as many relevant links as the reader can easily comprehend and select among.

- **Range of detail:**The great advantage of hypertext is that it permits readers to explore to a breadth and depth that is simply not feasible in print. To make this accessible, available hypertext documents should range from the broadest possible overview of a subject, down to its gritty details**.**
- **Correct links:**This may seem trivial, but it's amazing how many Web links point nowhere. In general, be careful linking to any hypertext document not under your direct control.

Hyperlink

- A hyperlink is a word, phrase, or image that you can click on to jump to a new document or a new section within the current document.
- Hyperlinks are found in nearly all Web pages, allowing users to click their way from page to page.
- When you move the cursor over a hyperlink, whether it is text or an image, the arrow should change to a small hand pointing at the link. When you click it, a new page or place in the current page will open.

Hypermedia

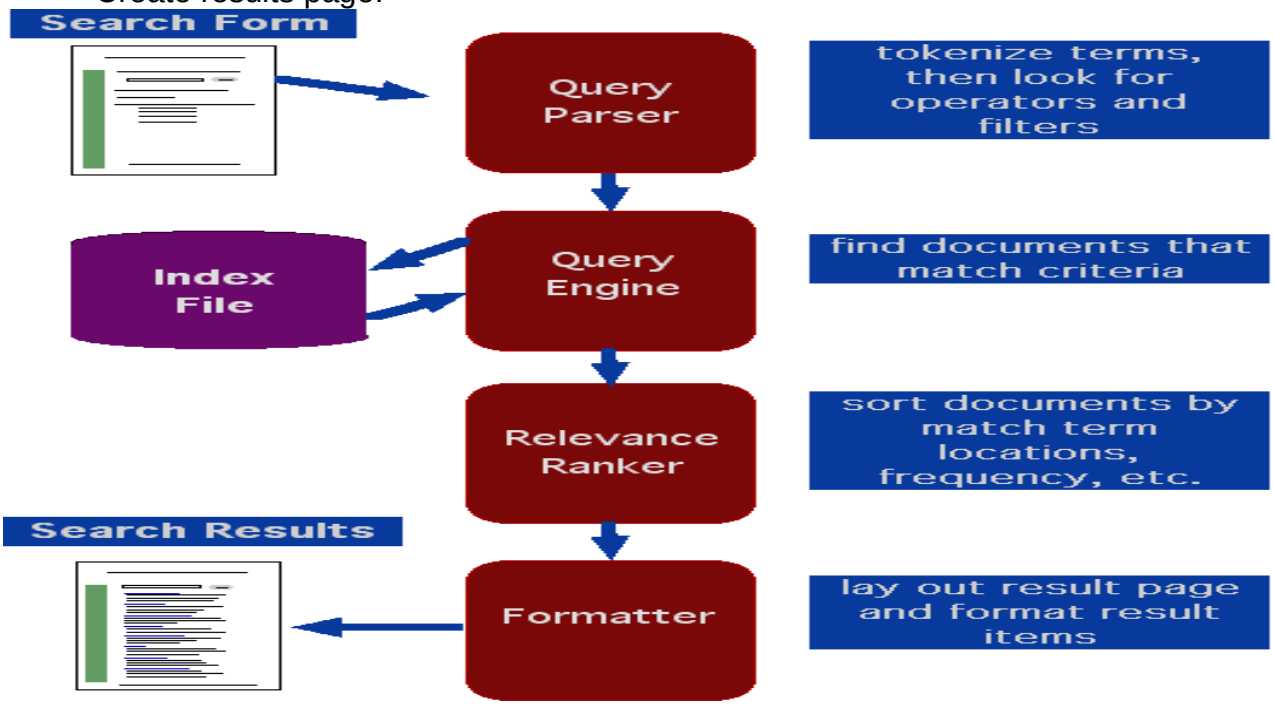
Hypermedia , an extension of the term hypertext, is a nonlinear medium of information that includes graphics, audio, video, plain text and hyperlinks. This designation contrasts with the broader term multimedia , which may include non interactive linear presentations as well as hypermedia

Search Engine

- Search engines are programs that search documents for specified keywords and returns a list of the documents where the keywords were found.

Search Engine Functions

- Accept query
- Look in index for words that match
- Extract documents that match all criteria.
- Sort by relevance (matches, location, link analysis)
- Create results page.



How does a Search Engine work.

Search engine works by following steps:

- Crawling
- Indexing
- Processing Queries
- Ranking

1.Crawling

- Search engines run automated programs, called "robots" or "spiders" that use the hyperlink structure of the web to "crawl" the pages and documents that make up the World Wide Web.
- once a new page is found the spider reads the Content & also checks for Images.

2.Indexing

- Once a page has been crawled by robots, its contents can be "indexed" - stored in a giant database of search engines called an "index". This index allows the Search Engine to sort all the documents in fractions of a second.
- This indexed data is stored in encoded format to save space.

Contents of index

- Basic information for document or record like
  - File name / URL / record ID
  - Title or equivalent
  - Size, date, MIME type
- Full text of item
- More metadata like
  - Product name, picture ID
  - Category, topic, or subject
  - Other attributes, for relevance ranking and display

3.Processing Queries:

- When a request for information comes into the search engine, the search engine retrieves all the documents from the index and tries to match them with the query.
- A match is determined if the term or phrase is found on the page in the manner specified in the query.
- The search engine found millions of matching information, so it uses an algorithm to decide in which order to display the results.

4.Ranking:

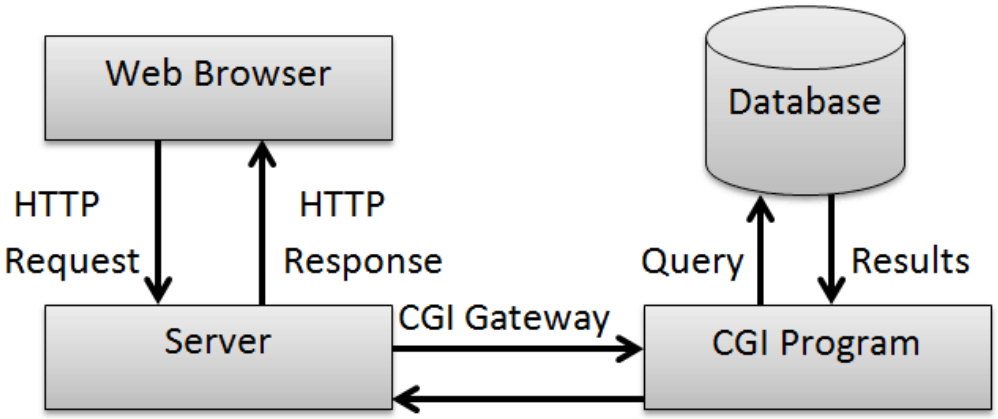
- Once the search engine has determined which results are a match for the query, the engine's algorithm runs calculations on each of the results to determine which is most relevant to the given query.
- They sort these on the results pages in order from most to least relevant so that users can make a choice about which to select.

Common Gateway Interface(CGI)

The Common Gateway Interface CGI) is a specification defining how a program interacts with a Hyper Text Transfer Protocol (HTTP) server.

• The Common Gateway Interface (CGI) provides the middleware between WWW servers and external databases and information sources

Working of CGI (Common Gateway Interface)



The sequence of events for creating a dynamic HTML document on the fly through CGI scripting is as follows:

1. A client makes an HTTP request by means of a URL. This URL could be typed into the 'Location' window of a browser, be a hyperlink or be specified in the 'Action' attribute of an HTML <form> tag.

2. From the URL, the Web server determines that it should activate the gateway program listed in the URL and send any parameters passed via the URL to that program.

3. The gateway program processes the information and returns HTML text to the Web server. The server, in turn, adds a MIME header and returns the HTML text to the Web browser.

4. The Web browser displays the document received from the Web server.

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.

When a computer connects to the internet, it uses an IP address. This is similar to your home’s street address, telling incoming data where to go and marking outgoing data with a return address for other devices to authenticate. A proxy server is essentially a computer on the internet that has an IP address of its own.

Proxy Services

The main purpose of a proxy service is to filter requests to ensure that no dangerous traffic creeps in by applying strict routing rules and to boost the performance of the system. A proxy service works simply – when a proxy service receives a request, for example, to open a Web page, it looks for the already cached pages. If it finds the requested page in the already cached page, it returns it to the user. If the page is not yet cached, proxy service uses its own IP address to fetch the page from the server for the client.

Proxy services are mainly of two types – forward proxy and reverse proxy. Forward proxy is an Internet-facing proxy that is used to retrieve a range of sources. A reverse proxy is particularly used for protection and security of the server. It includes

tasks like caching, authentication and decryption. Other types of proxies include transparent proxies, anonymous proxies, DNS proxies and highly anonymous proxies.

Dreamweaver

Dreamweaver is a web development tool.

Developed by Adobe Systems, Dreamweaver is a proprietary web development tool first created by Macromedia in 1997. Dreamweaver helps developers build web sites for the Internet or an intranet, letting developers design, code and manage websites as well as mobile content. It is currently available for OS X and Windows.

Dreamweaver combines a visual design surface, known as Live View, and a code editor with features such as code collapsing, code completion and syntax highlighting.