Finacle from Infosys

T101 – Web Fundamentals

PARTNERSHIP PROPELS INNOVATION,
TO SCALE NEW HEIGHTS

Infosys Technologies Limited
Plot No. 44, 3rd Cross
Electronics City, Hosur Road
Bangalore - 560 100
India

Document number:		VersionRev:	1.00
Authorized by:	R N NAGARAJ	Signature/Date:	

Document revision list

Ver.Rev	Date	Author	Description
1.00	10-04-2008	ARUN THOMAS	*Original version*

This document has been written and produced by the USER EDUCATION TEAM of Infosys Technologies Limited.

TABLE OF CONTENTS

1 OVERVIEW

Finacle is a web based application. Since the application is installed in a remote machine and the user will need to access it using a network connection. The network may be internet or intranet (network inside an organisation). So the knowledge of internet and basics of web is basic to understand the operation of Finacle application.

This document discusses the basics of web like

- History of internet
- Basics of web
- Some basic terms related to web
- Web interface to Finacle

2 INTERNET

2.1 COMPUTER NETWORK

A **computer network** is an interconnected group of computers. **Internetwork** is the interconnection between two or more networks. Internet is actually an internetwork which consists of worldwide interconnection of several academic, governmental, public and private networks.

2.2 WHAT IS INTERNET?

Large amount of information can be stored in a computer. A group of computers, connected to each other is called as a network. A network can store an even greater amount of information. Many people can share and work with each other's information by being on a network. If a network like this can exchange information with other networks, based on certain rules (known as protocols) it becomes a part of Internet. In simple words, Internet is a network of many networks. The information available on the Internet changes constantly. As new computers are constantly added to the network, the amount of information is ever growing. An individual or a company does not run the Internet. The content available on the Internet is called World Wide Web (www).WWW is the collection of all the information present in the internet.

The networks connected to each other would form the Internet. Roughly 6 million hosts and over 50 million users are hooked on to the net today.

Using Internet you can:

- Send and receive e-mail messages
- Read Newsgroups
- Visit websites

2.3 HISTORY OF INTERNET

Internet started nearly 25years ago as a project of US department of defence. The Technical group then created was known as Defence Advance Research Projects Administration (DARPA). Its goal was to create a way for widely separated computers to transfer information and data

and to make this data communications as much reliable as possible. DARPA required to make a network that was smart enough to recover on its own from problems such as power failure interruption in communication lines even in nuclear attacks. DARPA called its network as DARPA net.

Eventually the government dropped the idea that its network was only useful for defence related projects and the network became known as **ARPA net**. Around this time the government also started connecting many of the country's universities to the network. Since then, generations of students have studied, used and improved what is today known as Internet. ARPA net is generally considered as the first version of internet.

2.3.1 ARPANET

The **ARPANET** (**Advanced Research Projects Agency Network**) built by DARPA of the United States Department of Defense, was the world's first operational network. It is generally called as the predecessor of internet. The first ARPANET link was set up on November 21, 1969. It was established in UCLA (University Of California, Los Angeles) for the interconnection of four Universities. The first message sent over this network was on October 29, 1969 at 10:30 PM.

There were four nodes in the network each representing one university. Node is actually a machine in the network. The four nodes are:-

- University of UTAH
- UCLA (University of California, Los Angeles)
- UCSB (University of California, Santa Barbara)
- SRI (Stanford Research Institute).

In October 1972 ARPANET went 'public'. In the First International Conference on Computers and Communication which was held in Washington DC, the system in operation was demonstrated by ARPA scientists by connecting computers together from 40 different locations which inspired the Western world to research further in science. Other networks appeared soon after this. The Washington conference also established an Internetworking Working Group (IWG) to organize the research taking place. Meanwhile ARPA scientists were occupied on enhancing the system and expanding its capabilities.

- In 1972, they successfully made use of a new program which is now referred to as e-mail, that would allow to send the messages over the net, allowing direct person-to-person communication.

- The new host-to-host protocols developed in the early 70s allowed access to the hosts' programs. Before then the system only allowed a 'remote terminal' to access the files of each separate host.

- A common language called TCP/IP (Transmission Control Protocol/Internet Protocol) was developed by ARPA scientist in 1974 with help from experts in Stanford which would allow communication between different networks

The development of TCP/IP marked a key stage in networking development. In 1982, ARPANET allowed TCP/IP for its functioning. This is known as the birth of internet.

During the initial stages internet was only used for transferring files between computers. A protocol (rule) known as FTP is used for this purpose. FTP stands for file transfer protocol. Then a language known as GOPHER is introduced by the University of Minnesota. It was helpful in increasing the contents of internet. GOPHER can be used to create web

pages in the internet. But it has some drawbacks. It does not support multimedia content. So a new technology was introduced for writing web pages, known as **Hyper Text Markup Language**. After the introduction of HTML, internet became very powerful. HTML allows the web pages to contain videos, sound clips, animations and so on.

So the history of internet can be divided into three stages

- Initial stage: internet is generally used for file transfer purposes based on FTP protocol.
- Second stage: GOPHER was introduced. Web pages contained text only.
- Third stage: HTML was introduced. Multimedia features were introduced in web pages.

From there on, internet has grown very rapidly. Internet is now a global network of networks. It actually consists of many networks. The number of computers in such a network can range from two or three in a small intranet to several thousands in large organisations.

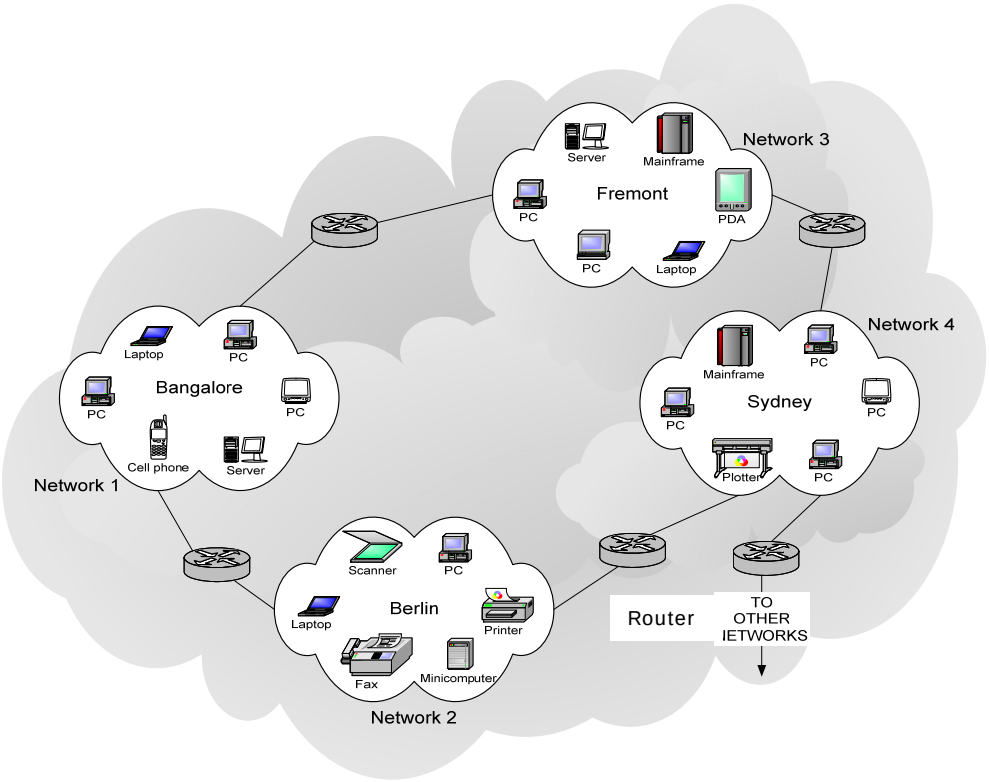
2.4 WORKING OF INTERNET

Internet is a network of networks. Each network will contain several computers. Each machine in the network is known as a **host**. It will have a unique address known as internet address (generally known as IP address).**IP address** is a 32 bit address. It will be in the form of aaa.bbb.ccc.ddd.

One network is connected to another network using certain devices known as **routers**. Information is sent from one machine to other in the form of **packets.** Each packet will have a header(which contains address of the machine to which the packets are sent) and a body. So router

checks the header of the packet and sends the information to the specified machine.

The following diagram shows the structure of a portion of internet which contains four networks.



3 NETWORK TOPOLOGY

Network topology refers to the arrangement of various network elements (nodes) in a network. There are three types of topologies.

- Physical topology
- Signal topology
- Logical topology

3.1 PHYSICAL TOPOLOGY

Physical topology refers to the physical connections and the mapping between the network nodes – i.e., cabling or wiring system, the layout of cables and wiring the locations of nodes and interconnections between them. There are various types of physical topologies. The commonly known physical topologies are:-

3.1.1 POINT- POINT TOPOLOGY

The two end points with permanent link between them would form the simplest topology called point-to-point topology. The value of a permanent point-to-point network is the value of guaranteed communications between the two endpoints.

The above diagram shows a point-point communication channel between a printer and a computer.

The advantages of point-point topology are

- **Permanent connection: -**

  The communication channel is permanently associated with two points. It is also known as **dedicated** connection.

- **Switched Connection: -**

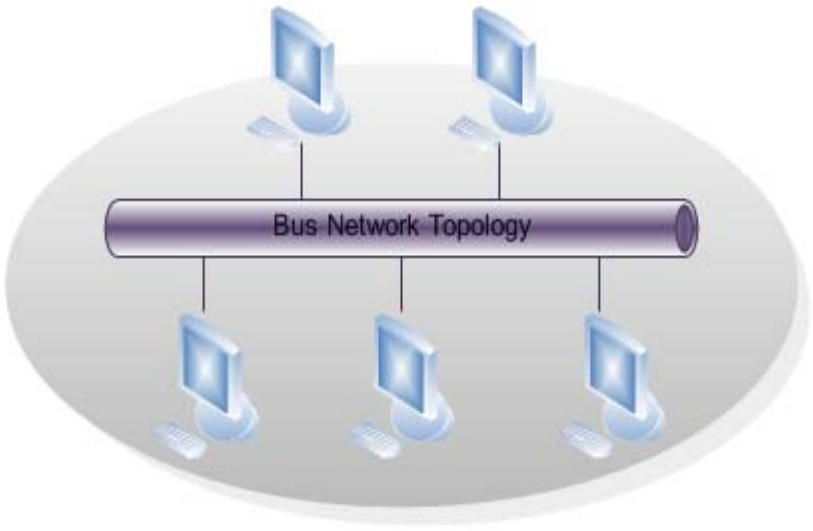
  A point-point connection can be switched-on or switched-off according to the requirements. It is easy to operate a point-point communication channel.

3.1.2 BUS TOPOLOGY

Bus topology is of two types – **linear bus topology** and **distributed bus topology.**

- **Linear BUS Topology: -**

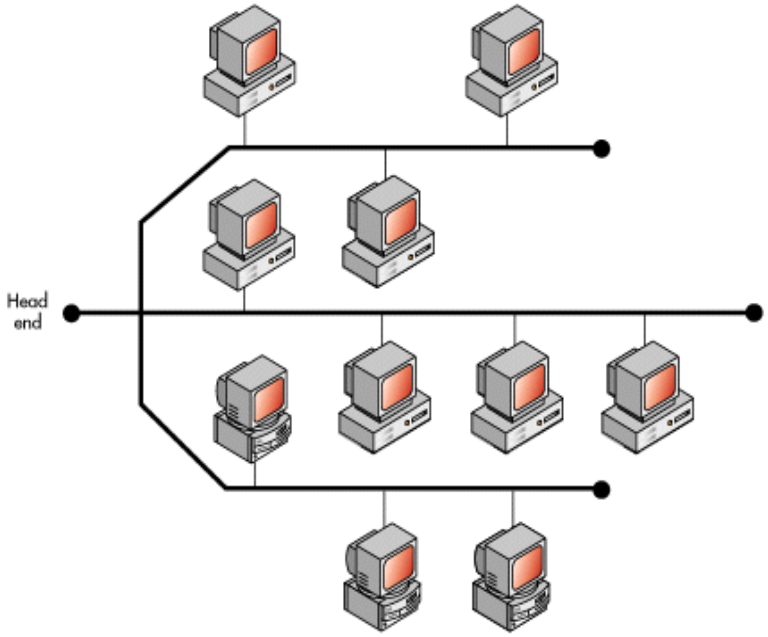
  The nodes in this network are connected to a general transmission medium called as 'bus' which has just two endpoints which acts as the backbone or the trunk in this topology. All information transmitted in between nodes in the network is sent over this common transmission medium and is accessible by all nodes in the network.



Linear Bus Topology

- **Distributed BUS Topology: -**

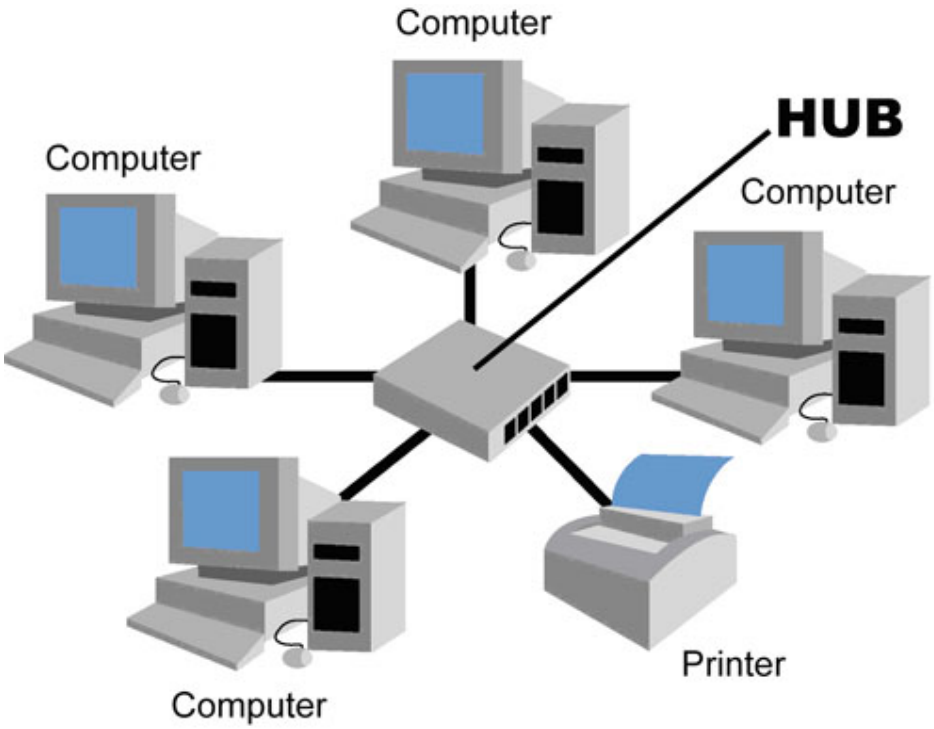
  A topology where in the nodes are connected to a common bus which has more than two endpoints which are formed by adding branches to the main bus is called distributed bus topology.



Distributed BUS topology

3.1.3 STAR TOPOLOGY

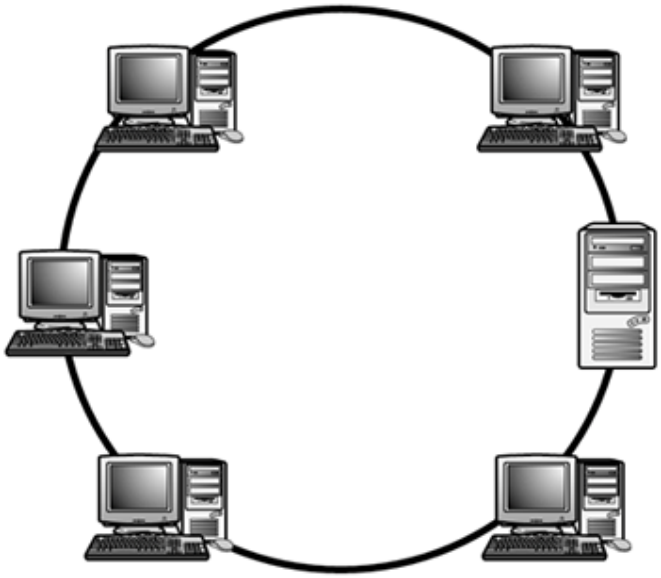
Star topology is one in which the cables are connected to a hub and the data flow from the nodes happens through this hub. The disadvantage in this topology is, if the central hub or switch fails the network would also fail.



Star topology

3.1.4 RING TOPOLOGY

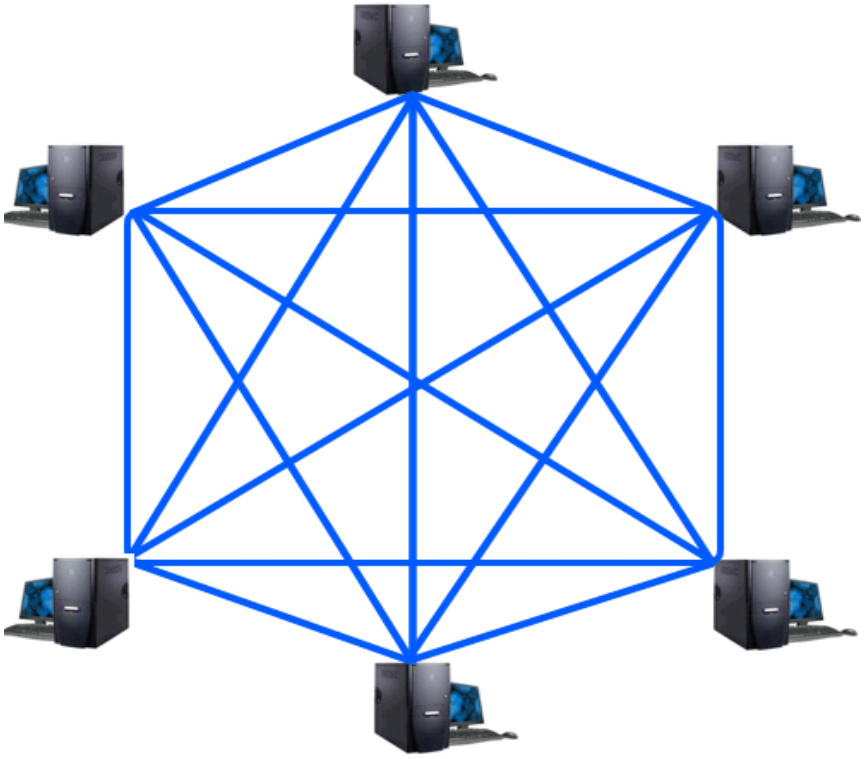
A topology in which a node is connected to two other nodes, with last and the first node connected to form a ring is called Ring topology. The data flow from one node to another in a circular manner and is unidirectional.



Ring topology

3.1.5 MESH TOPOLOGY

In mesh topology, every node is connected to each other such that there is a dedicated point-point connection between every node. The number of communication channels is very high in mesh topology. So it is costlier compared to other topologies.

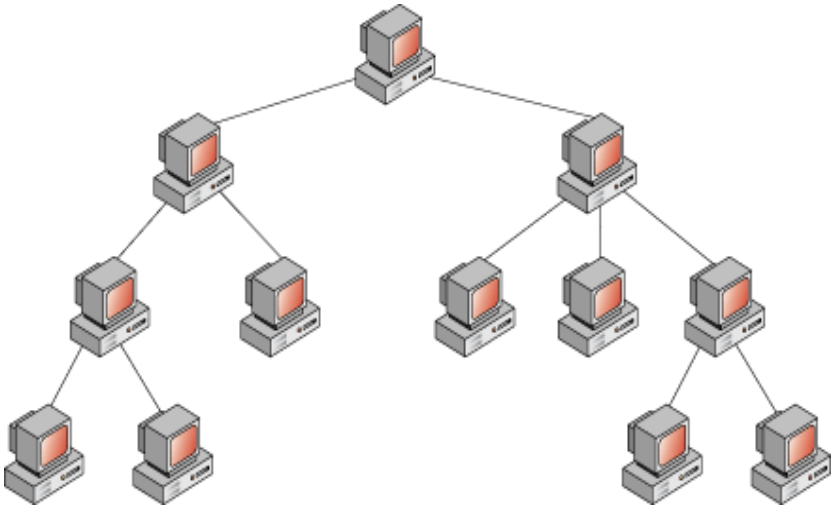


Mesh topology.

3.1.6 TREE TOPOLOGY

A topology in which a central root node is connected to more than one node in the form of a inverted tree structure is known as tree topology. This is also called hierarchical topology.

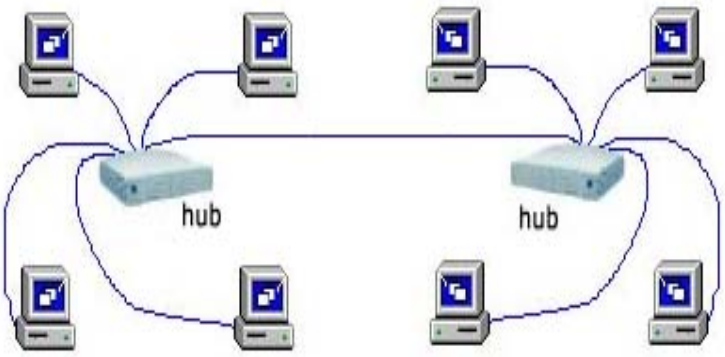
A topology in which a central root node is connected to more than one node



Tree topology

3.1.7 HYBRID TOPOLOGIES

Networks based on different physical topologies which are connected to each other will form the hybrid topology..It can also consists of two or more similar topologies, provided the resulting topology must not be same as the individual topology. There are various hybrid topologies like star-bus, star of stars, hybrid mesh etc.



Star-bus hybrid topology

The above diagram shows a hybrid network. It is actually a star-bus network. Here two star topologies are connected as a bus topology.

3.1.8 COMPARISON OF DIFFERENT PHYSICAL TOPOLOGIES (BUS, STAR AND RING)

Features	Bus topology	Star topology	Ring topology
Expense	Low	Medium	High
Reliability	Good	Excellent	Good

Geographically coverage ability	Poor	Good	Excellent
Ease of troubleshooting	Poor	Excellent	Good

3.2 SIGNAL TOPOLOGY

Signal topology is the mapping of different nodes in a network based on the paths by which the signals pass through. Signal topology generally refers to the actual path that the signals (e.g., electromagnetic, electrical, optical, etc.) take when propagating between nodes.

3.3 LOGICAL TOPOLOGY

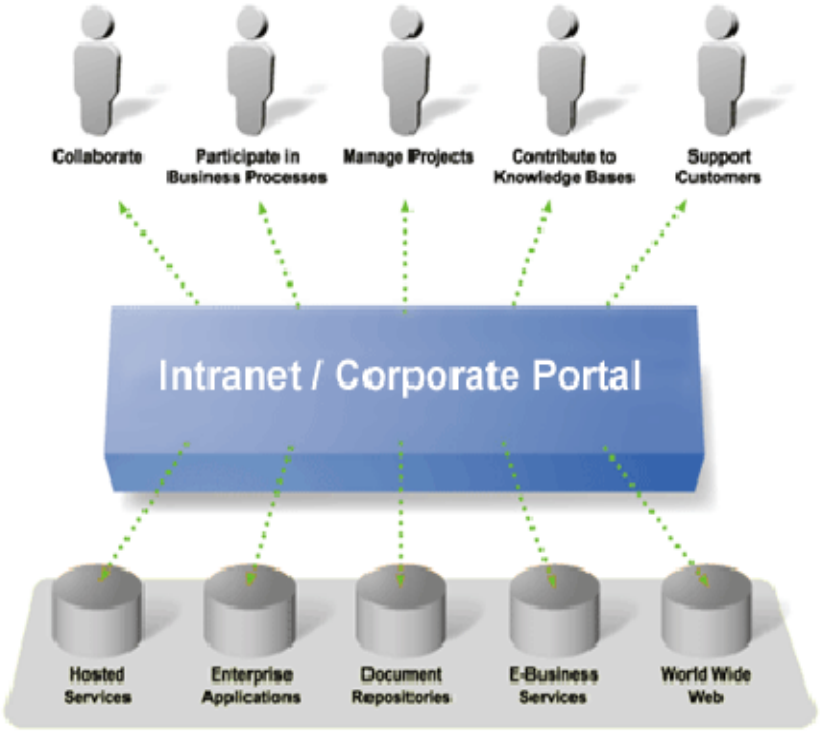
Logical topology refers to the mapping of nodes based on the paths taken by actual data in the network. It is similar to signal topology, but here the determining factor is path of data, not signals. In many cases logical topology and signal topology are used interchangeably.

4 INTRANET

Intranet is the internal network of an organisation. It is also known as corporate portal. It contains all the web pages, email services, product details, telephone directory etc. of an organisation. An intranet of an organisation is accessible to employees of that particular organisation only. A software known as **firewall** is used to ensure this. Firewall checks

whether the authorised person only can access the intranet. It checks all network traffic through the company's intranet.

Intranet also known as corporate portal or private business network utilizes standard network software and hardware technologies like Ethernet, TCP/IP, WiFi, Web browsers and Web servers. Internet access included in the organization's intranet cannot be directly accessed from outside as it will be firewalled. Intranets use the same HTTP server (Web server) technology, HTML hypertext links and communications protocols as the public Web.



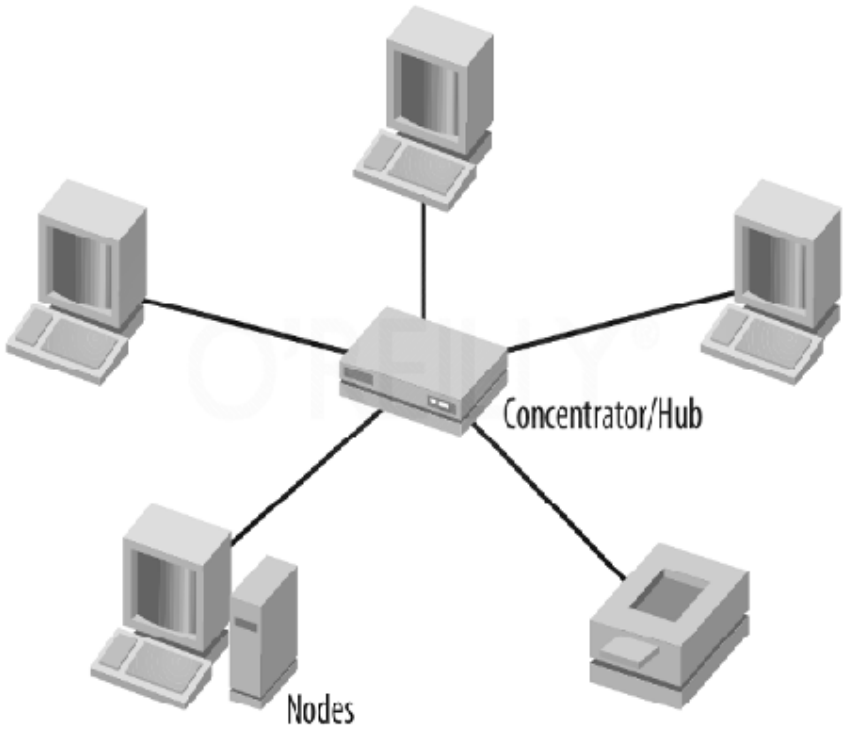
The above diagram shows a schematic representation of an intranet of an organisation. It may contain different ingredients like management of projects, document repositories, customer supports, e-business activities etc.

An intranet can be called as " a private version of an Internet" or " a version of the Internet confined to an organisation" because Intranet uses the same technologies and concepts used to build Intranet like clients and servers running based on a protocol. When there is need for internet from Organization's Intranet, it is provided through a firewalled gateway, along with encryption of messages, user authentication and often makes use of **virtual private networks (VPNs)**. Through such systems and devices off-site employees can access company information, computing resources and internal communications.

4.1 LOCAL AREA NETWORK (LAN)

A computer network which covers small geographic area like office, home or group of buildings and which is used often to achieve Intranet is called a LAN. The major differences between a LAN and internet are: -

- The data is transfer in LAN is much faster when compared to that of internet.
- LAN is limited to a small geographical region, but internet can be accessed from anywhere in the world.
- A dedicated telecommunication line is required for internet but it is not needed in LAN.
- Telecommunication line is required for internet.



The diagram represents a small local area network (LAN) with five nodes. A device known as concentrator or **hub** is used to connect the nodes. LAN is designed based upon **Ethernet**.

Ethernet:-

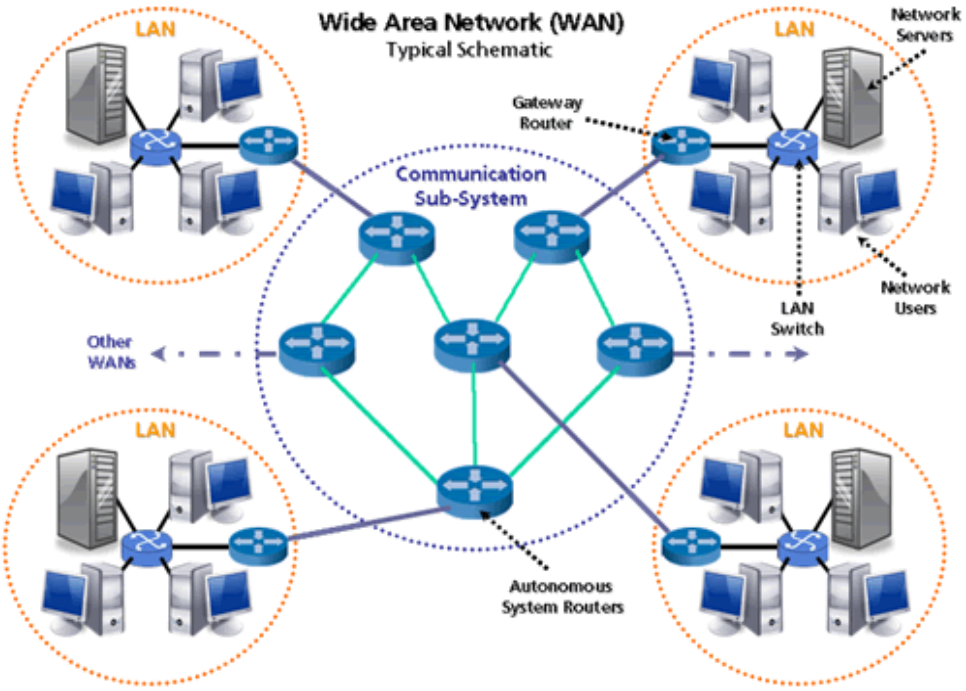
Ethernet as the name suggests has been evolved from the ancient term “ether”, meaning the medium through which light propagates. Basically Ethernet defines the wiring and signalling standards used for the network which is usual a Local Area Network (LAN). It uses the idea of a shared coaxial cable which acts as the media through which the transmission takes place and hence the communication between computers. The cable used for transmission purposes can be referred to as

the "ether" and hence the name Ethernet. This technology has been incorporated from computer networking technologies.

Over a period of time, cost effective and highly reliable Ethernet hubs and switches have replaced the crude but simple concept of coaxial cables, and form the complex networking technology that underlies most LANs.

4.2 WIDE AREA NETWORK (WAN)

Network over a large area is termed as Wide Area Network (WAN). Usually this type of network would be spread across geographical boundaries using routers and public communication links. Internet would be a perfect example for a WAN. Other examples include telephone lines, microwave links and satellite channels.



The above diagram is a schematic representation of a typical wide area network (WAN). It is actually a network of different LANs which are located in

fairly remote areas. These LANs are connected using certain devices known as **routers.**

Usually the transmission rate across this network would be in the range of 1200 bits/sec to 6 Mbit/sec. ATM and leased lines can even have transmission speeds greater than 156 Mbit/sec.

5 EXTRANET

A private network which uses Internet protocols, network connectivity and public telecommunication system to share a part of organization's data to its stakeholders securely is called an extranet. It can also be termed as an extended Intranet shared over the internet to registered users outside the company, enabled by authentication mechanisms on a login page.

Since that extranet is over internet, security and privacy is a must. For these reasons firewalls, digital certificates and VPN (virtual private network) are a must.

For example some parts of sparsh like the leave system can also be accessed from the internet, but we need to give our credentials for authentication purpose.

Extranet has some disadvantages also. They are: -

- Very expensive with hardware, software and training costs coming into picture.
- To secure the data over extranet would be a major concern.
- Extranet may reduce personal contact (face to face meetings) with stakeholders, which can cause a lack of connection between the company and its people thus hurting the loyalty of its business partners.

6 WORLD WIDE WEB (WWW)

World Wide Web (www), generally shortened to **web**, and refers to the complete content of internet. The web contains all the web pages which are written using a markup language like HTML which can be accessed via internet.

6.1 HISTORY OF WWW

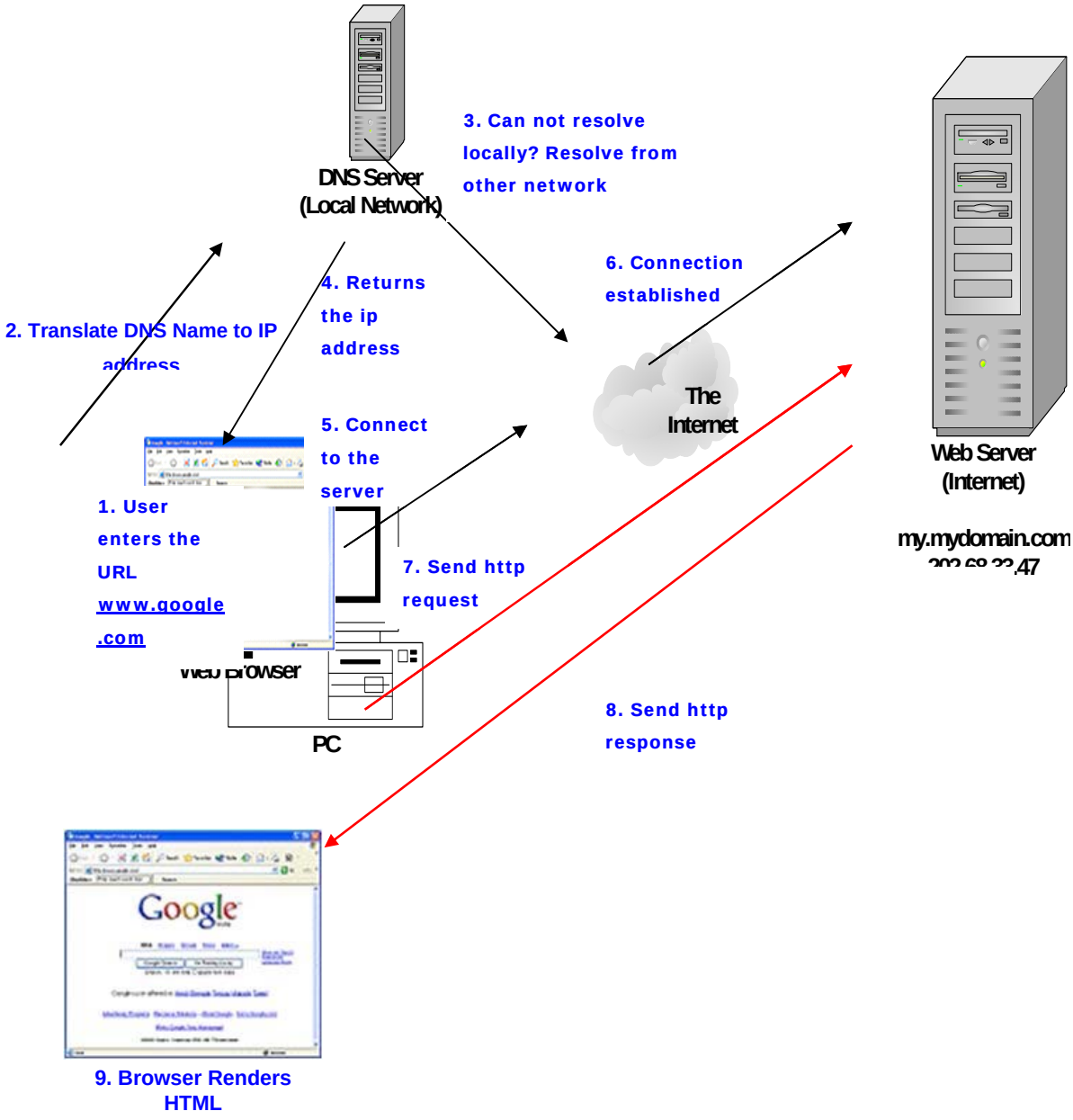
The whole idea of www was introduced in 1989 by **Sir Tim Berners-Lee**, who is considered as the father of internet. Since then the growth of internet was rapid. Tim Berners-Lee himself introduced the terms like URL, http. He is currently the chairman of w3c (World Wide Web consortium) which is the governing body of the web.

6.2 WORKING OF WEB

The working of World Wide Web is based on client-server technology. The client here is a user program known as **browser** which sends the request to server, which is a machine which contains all the web pages. The server machine returns the requested information as the response. So the web generally uses a request-response model.

The transfer of data between client (browser) and the server need to be according to some rules known as protocols. The protocol used in web is known as **HTTP (hyper text transfer protocol).** HTTP specifies the type of communication and information that is sent between a web site and a browser.

The basic working of the web can be explained using the following flowchart



The steps involved in working of a web server are:-

1. User enters the **URL** in browser. (**URL** stands for uniform resource locator, which is used to uniquely identify a resource in the

internet).This URL will contain the host's address, which can be either IP address or a particular name mapped to the IP address. This name is known as **domain name**.

2. There will be a **DNS server** local to the client's network. **DNS server** will map the domain name to the particular IP address. Local DNS Server will check whether that IP address is present in the network.
3. If the IP address is not present in the local network, check will be done against other DNS servers in the internet in order to get the IP address corresponding to the domain name provided in the URL.
4. The IP address corresponding to the domain name will be returned back to the browser.

   (If the IP address of the server itself is entered in the URL, the steps 2, 3, 4 can be avoided).
5. The browser will connect to the particular IP address through the internet.
6. A connection will be established between client (browser) and the server.
7. The browser will send a request (since we are using http, the request is known as **http request**) for the web page.

8. Server will return http response and that response contains contents of the requested page**.**

   (The terms **URL, DNS server, http request, http response** will be explained in detail later).

6.3 WEB BROWSER

Hard disk contains this browser as software that connects the client to internet. Browser is used to display the web pages. A web page can contain graphics, audio and movies as well as text. Browser will be equipped to display all these contents. Browser is generally known as client's window to the World Wide Web. Since documents on the web take long time to download, browsers have a capability to store images temporarily in a cache on the local hard disk to avoid multiple downloads of same images.

6.4 WEB SERVER

A web server is the machine that serves the documents that the client requests. It is capable of reading the request and processing the request sent by the client. The web server responds to the client by sending a web page that contains requested data as a response in HTTP format. All the web pages will be residing in the web server.

6.5 WEBSITE

A website is a collection of Web pages, images, videos or other digital assets that is hosted on one or more web servers, usually accessible via the Internet. HTML(Hyper Text Markup Language) is used to write typical web page. This page is accessible via Hyper Text Transfer protocol (HTTP). These web sites which can be accessed publicly constitute "World Wide Web". The first on-line website was created by Tim Berners-Lee in 1991.

6.6 INTERNET SERVICE PROVIDER (ISP)

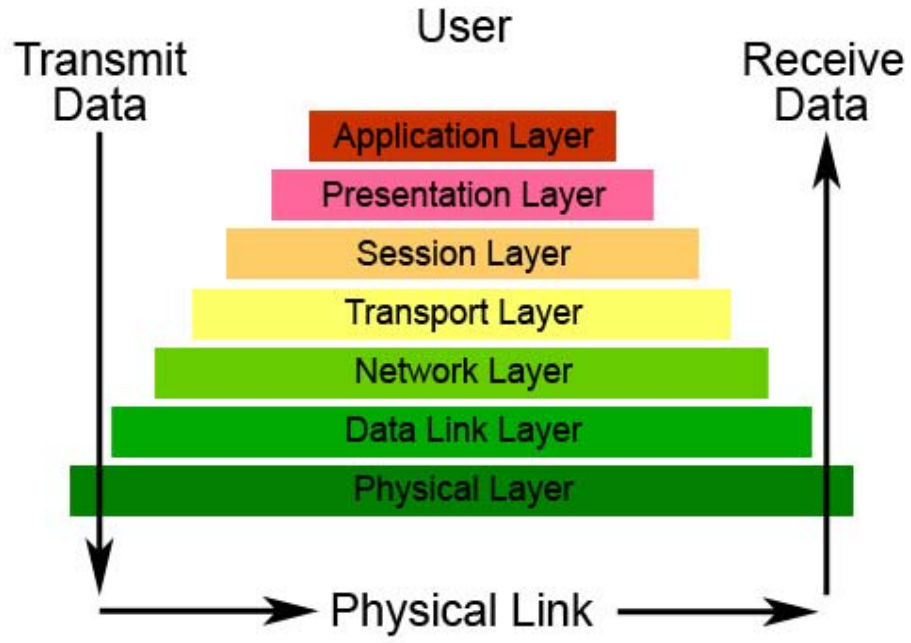
A company or business that provides access to the Internet and other related services is known as ISP. ISP, also called Internet access provider or IAP. In the past, most ISPs or IAPs were run by the phone companies. For example, the major ISPs in India are BSNL and MTNL. Now, any group or individual with sufficient expertise and money is able to act as ISP.

7 NETWORK MODELS

Network is the interconnection of different computers. Network model is the representation of a network as different layers. It illustrates the transfer of data in a network, what are the different protocols used for data transfer etc. There are two network models.

- Open Systems Interconnection (OSI) model
- Transmission Control Protocol/Internet Protocol (TCP/IP) model

7.1 OPEN SYSTEMS INTERCONNECTION (OSI) MODEL



OSI model is a newer network model compared to TCP/IP model. It is also known as the **OSI seven layer model** since it contains seven layers. The OSI Model consists of the Application Layer, Presentation Layer, Session Layer, Transport Layer, Network Layer, Data Link, and Physical layers from top to bottom. Each layer is capable of servicing the layer above it and issues service requests to the layer below it. A layer is a collection of related functions.

From the diagram the data we can see that the data is transmitted from the application layer to the physical layer and data is received from physical layer to application layer.

7.1.1 APPLICATION LAYER

The seventh level of the OSI model is Application Layer. It interfaces directly to the application. Sometimes presentation layer services application layer .HTTP, FTP are examples of protocols in the application layer.

7.1.2 PRESENTATION LAYER

It is the sixth level of the OSI model. It accepts service requests from the application layer and responds to the request .It also issues service requests to the session layer. Different services like encryption (for security purposes), structuring of data using XML (extended markup language) are done in this layer.

7.1.3 SESSION LAYER

It is the fifth level in this model. It responds to service requests from the presentation layer. Transport layer which is below it receives requests from this layer. The session layer helps in opening, closing and managing a session. The session would be between application processes which are at end-user side, i.e. a semi-permanent dialogue. Applications communicate in the form of requests and responses; session layer is responsible of maintaining session for such communications.

This layer provides synchronization. Synchronization is the process where transmitter sends the data and at the same time receiver receives the data. It is very important in cases like web conferences, online TV applications where video and audio should be transmitted simultaneously.

7.1.4 TRANSPORT LAYER

It is the fourth level of OSI model where it responds to service requests from the application layer and will issue service requests to the network layer. Transfer of date will be in the form of packets in this layer which will have a header and a body. Header contains source IP and the destination IP. The body contains the data to be transferred.

TCP is an example of a protocol in this layer. Transport layer provides some services. They are

- Transport layer is connection oriented.
- Data reliability is present in this layer. That means the errors in data transfer is minimised. This is done by re-transmitting the erroneous packages.
- Network congestion can be controlled.
- The concept of ports in introduced in this layer. Ports can be used to operate different data transfers at the same time using a single machine.

7.1.5 NETWORK LAYER

The third layer in OSI model is **network layer**. It is sometimes called as internet layer. It acts as an interface by responding to the service requests from transport layer and issuing them to the data link layer. The services like packet forwarding and packet routing are done in this layer. Forwarding is the process in which a packet is passed from one node to another whereas routing is the process in which the most appropriate path of packet transfer is selected.

Internet protocol (IP) is a protocol in this layer.

7.1.6 DATA LINK LAYER

The second layer in the seven-layered OSI model is **data link layer**. It responds to the service requests from network layer and issues service requests to the physical layer as well. It transmits data between the adjacent network nodes in a wide area network and also between nodes on the same local area network segment. The data transfer between network entities in both functional and procedural means takes place in this layer and also the means to detect and correct the errors that may crop up in the Physical layer.

Following are the services provided by data layer.

- Error detection
- Flow control.

Example of data link layer protocol is Ethernet.

7.1.7 PHYSICAL LAYER

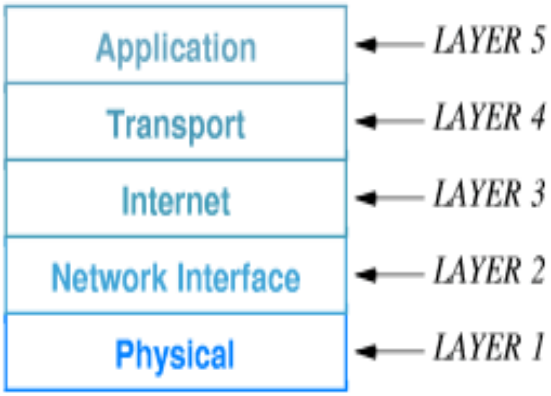
The first layer in the OSI model of computer networking is **physical layer**. It carries out the services requested by data link layer. It is the basic network layer which provides the only means to transmit raw bits rather than sending packets over a physical data link connecting network nodes. The torrent of raw bits are grouped into code words or symbols and then converted into a physical signal which is then transmitted through a physical transmission medium. The type of network topology (which could be LAN, bus, star) is specified in this layer.

7.2 TRANSMISSION CONTROL PROTOCOL/INTERNET PROTOCOL (TCP/IP) MODEL

TCP/IP model was created in 1970 by DARPA (Defence Advanced Research Projects Agency), an agency of United States Department of Defence. **Arpanet**, which is the predecessor of internet is created based on this model. TCP/IP model has five layers. They are –

TCP/IP Model has five layers, namely Physical layer, Data link layer, Transport layer, Network layer and Application layer.

Schematic representation of TCP/IP model is shown below



7.2.1 APPLICATION LAYER

The top layer of the TCP/IP model is **Application layer**. Here, the data in the application specific format is transferred from program to transport layer. It does the work of session layer, application layer and presentation layer in the seven layer OSI model.

Examples of Application layer protocols are TCP, FTP (file transfer protocol) and telnet (which uses a thin client to get connected to the server).

7.2.2 TRANSPORT LAYER

The layer next to application layer in TCP/IP model is **transport layer**. It is responsible for end-to-end message transfer capabilities autonomous to the underlying network with error control, flow control and fragmentation. This can be of two types

i. Connection oriented

In a connection oriented type, connection between end points should be established before transferring any data. TCP is an example of connection oriented protocol

ii. Connectionless

In a connectionless type data can be sent between end points without even establishing a connection. UDP (User Datagram Protocol) is an example of connectionless protocol. UDP allows computers to send datagram between each other. Datagram is similar to a packet of information.

There are many advantages by using a connection oriented TCP protocol. They are: -

- data arrives in-order
- data has minimum errors (i.e. correctness)
- data duplicated is discarded
- lost/discarded packets are resent
- includes traffic congestion control

7.2.3 NETWORK LAYER

Network layer forms the third layer from top in the TCP/IP model. It does the functionality of transferring packets between networks. Internet protocol (IP) is an example of network layer protocol. IP performs the basic task of sending packets of data from source to target.

7.2.4 DATA LINK LAYER

The data link layer is the second layerl. This layer has the functionality of adding the header information to the packet and also to transfer packets to the physical layer. The layer where packets are intercepted and sent over a virtual private network.

7.2.5 PHYSICAL LAYER

The first layer in TCP/IP model is **Physical layer**. It is accountable for encoding and spread of data through network communications media. It sends the data in the form of bits from the Physical layer of the source device and is received at the destination device. It also contains many

hardware-related network design issues. For example, LAN, WAN and wireless technology.

8 BASIC TERMINOLOGIES

8.1 HTML (HYPER TEXT MARKUP LANGUAGE)

The basic language for designing the websites is HTML. It is a markup language. **Markup language** is utilized to organize and format text. It consists of text embedded in different tags.

8.2 DOMAIN NAME

Domain name identifies a computer on the Internet belonging to particular network. These names appear as a part of a Website's URL, e.g.infosys.com. These types of domains are also called as hostname.

8.3 DNS (DOMAIN NAME SERVER)

Web server contains all the web pages. Each web server will have a specific IP address. It is difficult to memorize IP addresses of the entire servers. So we can map the IP address to a particular name. This is known as **domain name.** The domain name is mapped to the IP address using DNS(Domain Naming Systems). Generally, there will be a DNS server local to each network in the internet.

8.4 URL (UNIFORM RESOURCE LOCATOR)

Address of the webpage is structured by URL. For example, the URL for Infosys is www.infosys.com . The resource of web pages is uniquely identified by URL. The clients require this URL to access the web pages.

8.5 HOME PAGE

This forms the website's main page. the home page will come up by typing the URL of the website. We can navigate any other web page of that particular web site from the home page.

8.6 HTTP (HYPER TEXT TRANSFER PROTOCOL)

HTTP defines the rules based upon which the data is transferred between web browser and web server. For accessing the required web page, an http request is sent to the server by the browser. After receiving the http request, the server will send required webpage as http **response**.

Detailed explaination of all these terms are made in the coming sections.

9 HTML (HYPER TEXT MARKUP LANGUAGE)

HTML, its a markup language which is used for writing the websites. **Markup language** is used to structure and format text. The compostion of text-based information in a document is described using HTML. This is done by denoting certain text as links, headings, paragraphs, lists and so on and it also adds the text with interactive forms, embedded images and few other objects. HTML is created in the form of tags, surrounded by angle brackets.

First a markup language introduced was called as SGML (standard generalised markup language). All the other markup languages are further developed from SGML. So SGML is called as the mother of all markup languages.

HTML can be understood by all browsers. It is also platform independent (means HTML can be used in any operating system like Windows, Solaris etc.). A simple html page looks like the following.

```
<html>
<head>
<title>A sample HTML page...</title>
</head>
<body>
<center>
<h2>A HTML Page</h2>

<p>This is a sample HTML page which demos <b>Bold</b>,
<i>italic</i> text and a table.
</center>
</body>
</html>
```

The portions enclosed between "<" and ">" are known as **markup tags**. They are used to structure and format the text. <body>, <head>, etc. are examples of markup tags.

9.1 HTML VERSIONS

- The first version of HTML was introduced in July, 1993 by Tim Berners-Lee.
- Second version HTML 2.0 was introduced in April, 1995.
- HTML 3.0 was launched in September, 1995. HTML 3.0 contained some complex mathematical elements, more attributes for tables etc.
- After HTML 3.0, w3c (World Wide Web consortium) launched the next version HTML 3.2 in January, 1997. Interestingly, HTML 3.1 was not officially launched.
- Next version HTML 4.0 was introduced in April, 1998.
- Latest version of HTML is HTML 5 which was introduced on 22 January 2008.

10 DOMAIN NAME

Domain name which are nothing but the hostnames which forms a part of the Web site's URL, e.g.infosys.com helps in identifying a computer or computers on the Internet.

Domain names provide standard names for numeric IP addresses which can be remembered easily. It makes Internet users to simply find and converse with web sites and other server-based services. The domain name system is flexible which will allow assigning multiple IP addresses to single domain or a single IP address can be allotted multiple domain names. Domain names are often referred to as domains.

The domain name is divided into two parts – top level domain and second level domain.

10.1 TOP LEVEL DOMAIN (TLD)

A top-level domain, referred to as a top-level domain name (TLDN) sometimes. It is the part of the domain name which comes after the final dot in it. For instance, in the domain name www.infosys.com, com or COM is the TLDN. The top level domain names are standardized to mainly two divisions.

10.1.1 COUNTRY CODE TLD (CCTLD)

A CCTLD is an Internet top level domain generally reserved for a country. Creation and delegation of ccTLDs is achieved by the Internet Assigned Numbers Authority (IANA).

Examples for ccTLD are .in, .us, .uk etc.

10.1.2 GENERIC TLD (GTLD)

A GTLD is a top-level domain used by a particular organization class. These are to be minimum three letters long, and are named for representing type of organization (for example, .com for commercial organizations).

Examples for gTLD are .com, .edu, .org etc.

10.2 SECOND LEVEL DOMAIN (SLD)

Second-level domain (SLD) is a domain that is directly under a top-level domain (TLD). It is the portion in the domain name that comes before the final dot. Every organization can get their own second level domain name provided they are registered through domain name registrar.

For example, in the domain name www.infosys.com, infosys is the second level domain.

11 URL (UNIFORM RESOURCE LOCATOR)

URL is the address of the web pages. For example, the URL for Infosys is www.infosys.com . URL uniquely identifies a resource of web pages. The clients require this URL to access the web pages.

URL generally has the following structure:

protocol://host:port/path

i. **Protocol:-**

It indicates a rule that is used for the data transfer (it can be http, ftp, etc.).

ii. **Host:-**

It contains the IP address or the domain name of the web server which contains the web pages.

iii. **Port:-**

It specifies the port number using which the http request and http response can be passed. It is optional.

The default port number for http is 80. This port number will be used if no other port is mentioned in the URL.

iv. Path:-

It refers to the complete path of the document in the host.

For example,
Consider a URL http://server1.mydomain.com:8080/index.html

Here,

Protocol: - http

Host: - server1.mydomain.com

Port: - 8080

Path: - /index.html

URL is also known as URI. URI stands for Uniform Resource Identifier.

12 FTP (FILE TRANSFER PROTOCOL)

FTP is a system protocol which will be used for transferring data from one computer to another through a network, like the data that is transferred s over the Internet. The default port for FTP is port no: 21. The two modes of data transfer are ASCII mode and BINARY mode.

12.1 ASCII MODE

When a file is transferred using an ASCII mode, the individual letters, numbers, and characters are transferred using their ASCII character codes. In this mode which is used as default by most FTP clients, the data will be saved in a text file in the proper format at the receiving machine

12.2 BINARY MODE

In "Binary-Type transfer", the sending machine sends each file bit by bit and as the recipient stores the bit stream as it receives it.

FTP has some drawbacks compared to HTTP. They are:-

- **Security Issues**:-

  Passwords sent using FTP is in the form of clear text. Clear text is type of text which can not be encrypted. So passwords can seen by anyone who can access the communication channel. This causes a key security issue.

- **Firewall**:-

  It is very difficult to install firewall for data transfer using FTP. This is a major issue in case of access authentication.

- **Difficulty in programming**:-

FTP programming is difficult compared to that of HTTP. We require a more number of commands to initiate a file transfer.

13 HTTP (HYPER TEXT TRANSFER PROTOCOL)

HTTP refers to the rules based upon which the data is transferred between web browser and web server. It also determines how the data is transferred from the client to the server and vice-versa. The concept of HTTP was introduced by Tim Berners-Lee in 1990. HTTP specifies the functioning of a web based application.

HTTP is a request/response model between a client and a server. The end user will be the client and the web site will be server. The client makes http request using web browser which is called as the user agent. The responding server is know as origin server .It stores or creates resources such as HTML files and images.

13.1 PROPERTIES OF HTTP

13.1.1 CLIENT-SERVER ARCHITECTURE

The HTTP protocol is request/response process. A request is sent in a particular format, once the client establishes the connection with the server and this type of request is known as http request. The web page contents are sent as the part of the response to the client.This is known as **http response.**

13.1.2 CONNECTIONLESS

Even though we mention that connection is established between client and server, the protocol is known as connectionless because once the single request has been processed, the connection is dropped as in the case of the earlier versions of HTTP.

13.1.3 STATELESS

After the client's request receives the response from the server, the connection between client and server is dropped and not recalled. There is no "memory" between client connections. So the server can not store any information which is entered by the client. Server will treat each subsequent request as a new one.

13.1.4 STANDARDISED

HTTP is standardised. That means any vendor's browser (like internet explorer, Mozilla Firefox) can understand the http protocol. It is also platform independent. So HTTP can operate with any operating system.

13.2 HTTP REQUEST

The client (browser) requests for a particular web page from the web server. This is known as http request. HTTP request will have a specific structure. The web page cannot be accessed if there is violation of this syntax in the requests sent. The sending of HTTP request happens inside the browser.

HTTP request has the following four parts: -

13.2.1 INITIAL LINE

The initial line of HTTP request contains: -

i. The http method (the method of data transfer between browser and server, which will be explained later).
ii. URL (actually the request will contain the path of the required document).
iii. The protocol version (which can be like HTTP/1.0 or HTTP/1.1)

13.2.2 HEADER INFORMATION

After the initial line the http request can contain some header information. The header information will include details like

- Client information, which tells us which browser, is used.
- Acceptable data formats, which tells us what are the data formats that can be allowed by the browser.

13.2.3 BLANK LINE

A blankline after the header information is used to separate header information from the message body

13.2.4 MESSAGE BODY

Message body contain information. This information will be sent out to the server.This is used only with http **POST** method.

example of an http request is

GET /usr/index.html HTTP/1.0
User_Agent: Mozilla/3.0 Gold
Accept: text/plain

Accept: text/html

Here,

Path of the document: **/usr/index.html**
Browser: **Mozilla/3.0 Gold**
Data formats accepted are **plain text and html text.**

13.3 HTTP RESPONSE

Server receives the request from the client and the contents of the requested web page will be sent as a response. This is known as http response. Http response will have a specific structure.

HTTP response has the following four parts: -

13.3.1 INITIAL LINE

Initial line is also known as status line.

The initial line of HTTP response contains: -

i. The protocol version (which can be like HTTP/1.0 or HTTP/1.1).
ii. Status code, which stand for the status of the http request and the document.
iii. Status description, which describes the status code.

Different HTTP status codes: -

Status code	Description
200	OK
302	Found
400	Bad Request
401	Unauthorised
403	Forbidden
404	Not found
500	Internal Server error

For example, when we request for a page which is not present in the server, the browser will show **"404: page not found"**. This is actually the status code and description.

13.3.2 HEADER INFORMATION

After the initial line the http request can contain some header information. The header information will include details like

- Enter name or numberServer information, which tells us which type of server, is used.
- Date and time of response.
- The requested page contains the type of data, which can be like text/plain or text/html.
- Length of the information being sent along with the response.

13.3.3 BLANK LINE

After the header information the http response should contain a blank line to separate the header information from the message body.

13.3.4 MESSAGE BODY

In HTTP response, the message body has the contents of the documents in an html page or any information that are required by the clients.

Eg:

```
HTTP/1.0 200 OK
Server: NCSA/2.0
Date: Wed, 05 May 2004 16:00:04 GMT
Content-type: text/html
Content-length: 2400
<html>
<head>
<title>Welcome to Infosys Home Page</title>
</head>
<body>
  ...
</body>
</html>
```

13.4 HTTP METHODS

In HTTP methods, the client information is sent from client to server. Some of the examples for methods are GET,POST,HEAD etc.

13.4.1 GET METHOD

In GET method, the client side information is sent to browser as parameters. These parameters will be encoded with the URL as name-value pair.

For example,
http://sparsh/v1/ASPX/Search.aspx?strSearch=bangalore

The above shown URL is an example for using GET method.

The information entered by client is bangalore.'strSearch' is the parameter used in the URL for the value 'bangalore'

- GET method is faster compared to POST method.
- The size of data that can be sent is limited to 2MB. (This is browser dependant).
- Generally critical information like passwords are not passed using GET method.

GET is the default HTTP method.

13.4.2 POST METHOD

In POST method the information from client is sent to server as part of the message body of http request. The data can also be encrypted to ensure the security.

- POST method is slower compared to GET method.
- The pages which uses can not be book marked.
- Large sized data can be sent via POST method.

- Critical information like passwords are passed using POST method. They are generally encrypted.

13.4.3 HEAD METHOD

The HEAD method gets only the information about the document, not the document itself. As a less data is transferred, HEAD is much faster than GET.

For example, consider a situation in which we want to check whether a document is modified after a particular time. HEAD method can be used in such situations, which shows the last modified time.

13.5 HTTP VERSIONS

There are different versions of HTTP which are developed during the last decade.

13.5.1 HTTP/1.0

HTTP/1.0 is the first version of HTTP. It is introduced in May 1996. Before that some deprecated versions of HTTP were used. HTTP/1.0 is a connectionless and stateless protocol.

- In this version, after the response to a particular request has been given, the connection is withdrawn. That means, one network connection can process only one request. So it is known as connectionless protocol.

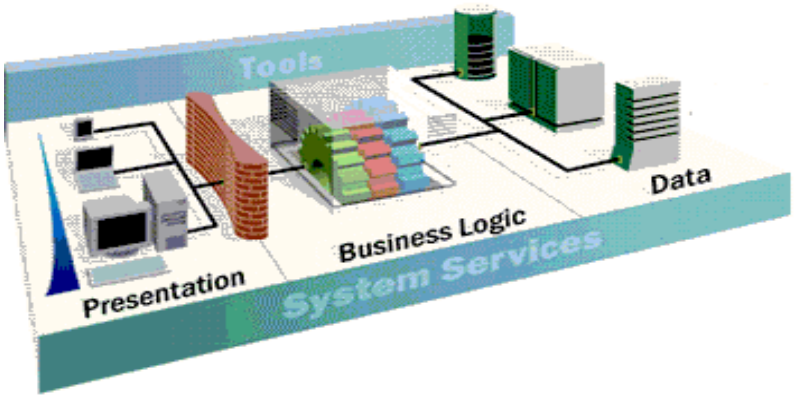
- Also HTTP/1.0 can not store the information entered by the client in memory. So it is known as stateless protocol.

13.5.2 HTTP/1.1

The second version of HTTP, known as HTTP/1.1 is introduced in June 1999. It can not be called connection less because it can process multiple requests with one connection. But it is stateless as it can not store the client information in memory. This is the version which is currently in use.

14 MULTI-TIER APPLICATIONS

Generally business applications are designed in different tiers. Three-tier client-server architecture is generally used in web based applications like Finacle.



The three tiers are: -presentation tier, application tier and data tier.

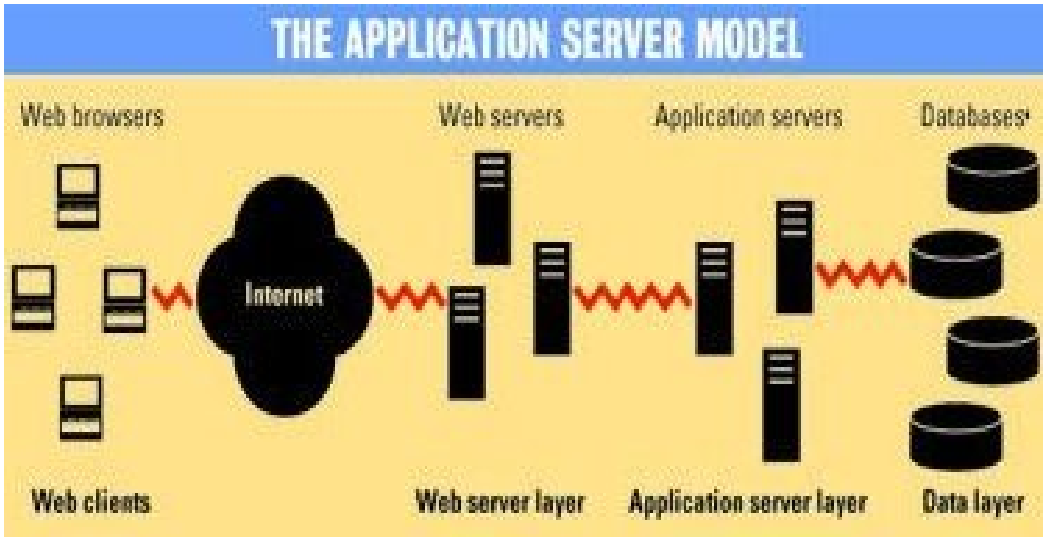
i. **Presentation tier: -** Presentation layer which is the highest layer in the application takes care of the display in the browser. It can communicate with other tiers. **Web servers** are the main components of presentation tier. **Graphical User Interface (GUI**) is provided by the presentation layer.

ii. **Application tier: -** Application tier takes care of the business logic. That means the basic functionality of the application is carried out by the application tier. Application server controls the functionality of the business application.

iii. **Data tier:** - Data is stored in this tier. Data tier mainly consists of Database Servers.

14.1 STRUCTURE

Multi-tier applications will generally have a "**thin client - fat server**" structure. In this kind of structure, the client will have only web browsers to view the web pages. So the client is called as a **thin client**. Web server, Application server and Database server will be the server side. So the server is known as **fat server**. The client-server architecture for a multi-tier application can be represented by the following diagram.



14.1.1 CLIENT SIDE

In the client side we will have only **web browser**. Client can view and enter data in the web pages using web browser. Internet Explorer and Mozilla Firefox are normally used. Here the client is known as a thin client.

14.1.2 SERVER SIDE

The server side in general three tier architecture consists of **Web server, Application server and Database server**. So the server is called a fat server. The Server will process request sent by the client and sends the information back that is requested as an http response.

14.2 WEB SERVER

Web Server is the important component of the presentation tier of the three-tier client server architecture. It is actually a computer program which accepts the http requests and sends the requested data as http response. The data mainly consist of HTML pages. There are two types data: -

a. **Static**: - Static content is the data that is present in the file system in the database server
b. **Dynamic:** - The program that is run by the web server generates Dynamic content. It is slower compared to static content.

Apache, Microsoft IIS are examples of web servers which are commonly used.

14.3 APPLICATION SERVER

Application server is the major component of application tier of three tier client-server architecture. Application server helps the thin clients to access the database and do the necessary operations. Business logic and data access is handled by the application server, hence known as centralization. No programming is required for the applications instead; they are assembled in the form of building blocks that are provided by the application server. So application development becomes easier; Application servers are computer programs which help the client to access the database. They can communicate with the web server and also with the database. They are operating system independent. Some of the operating systems like Windows, Solaris etc are used by the application server.

The main advantages of using an application server are: -

i. **Centralised Configuration**: -

The changes to the application, mainly business logic needs to be done at only place. So the application maintenance is easy.

ii. **Security**: -

The application can be accessed only through a single point. So all the database and application accesses can be monitored.

iii. **Performance**: -

The network traffic can be controlled. So the number of users accessing the application can be controlled and hence performance is increased.

14.4 DATABASE SERVER

Database server is a computer program which provides database services to other computers or computer programs. Database servers store database on a single computer system which can be used concurrently. According to the client-server model, an application is divided into two sections by the database server. One is front end and the other is back end. The front end shows requested data which runs on the user's computer. The back end handles tasks like data analysis and storage which runs on the server.

Advantages of using a database server: -

i. Users can store data in one location in the center in the database server.

ii. Complex functions like searching, sorting, and indexing are done on the server itself. So network traffic is reduced because some items are to be transferred between the client and the server.

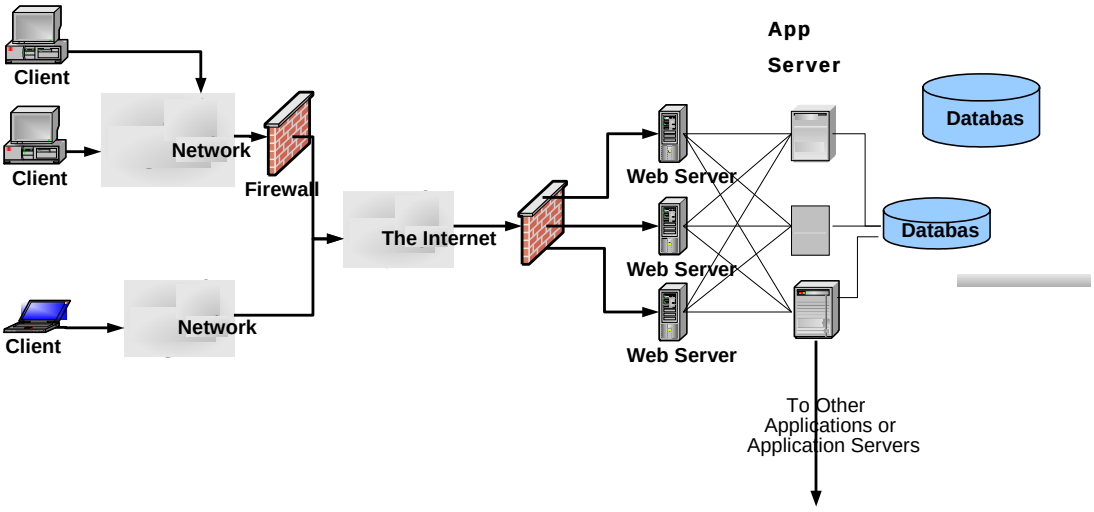
iii. The security of the database server is enhanced because of the centrally stored data.

iv. Using its own processing power the database server searches for the requested data, rather than sending the complete data to the client so that the client searches for the data, as is done in a file server.

v. Concurrent access to data is allowed using the database server

15 REAL WORLD WEB APPLICATIONS

The web applications which are accessible through the internet are known as real world web applications. Real world web applications generally use three-tier client-server architecture. The client is generally known as thin client. A classic web based application is shown in the diagram below,



The above block diagram is divided into two sections: -

a. **Client side**: -

In a common web based application, the client will be connected to a local network, which can be the client's organisation intranet. As we know, each organisation's intranet will have a **firewall**, to check whether the authorised personnel only can access the local network.Therefore, the client can connect to the internet from his intranet through the firewall. As we know client sends the http request from the browser.

b. **Server Side**: -

The server side also have a firewall to monitor the server access. The web server processes the http requests that comes from the client. Web server is responsible for displaying the web pages. Web server handles the presentation part of the application. The web servers generally used are apache and Microsoft IIS. As the web server communicates with the application server webserver acts like an intermediary between client and the application.

Application server gets the requests from the web server. The business logic and the database access is handled by the application server. Business logic is the fundamental functionality of the application. Client can not directly access the database. Client can access the appliace

The data is present in a machine known as database server. This can only be accessed by the application server.

The following is the flow of working of a real time web based application: -

i. Client sends the http request from the browser through the firewall in his intranet.
ii. The client is authorised by the server whether it can access it or not using firewall.
iii. Web server receives the request and web page is shown in the browser.
iv. The client accesses the application and database through the application server. The communication of

Client with the application server is through the web server.

16 FINACLE WEB INTERFACE

Finacle is a web based application. It can be accessed via intranet as well as internet. The key features of Finacle web interface are: -

- Provides a browser-based GUI interface to Finacle

  The Finacle application can be accessed through any browser. User will see an interface which is to be worked on.

- Allows integration with other browser-based applications

  With the help of single sign on (SSO), we can log on to other applications by logging into only one application.

- Designed to work over intranet as well as internet

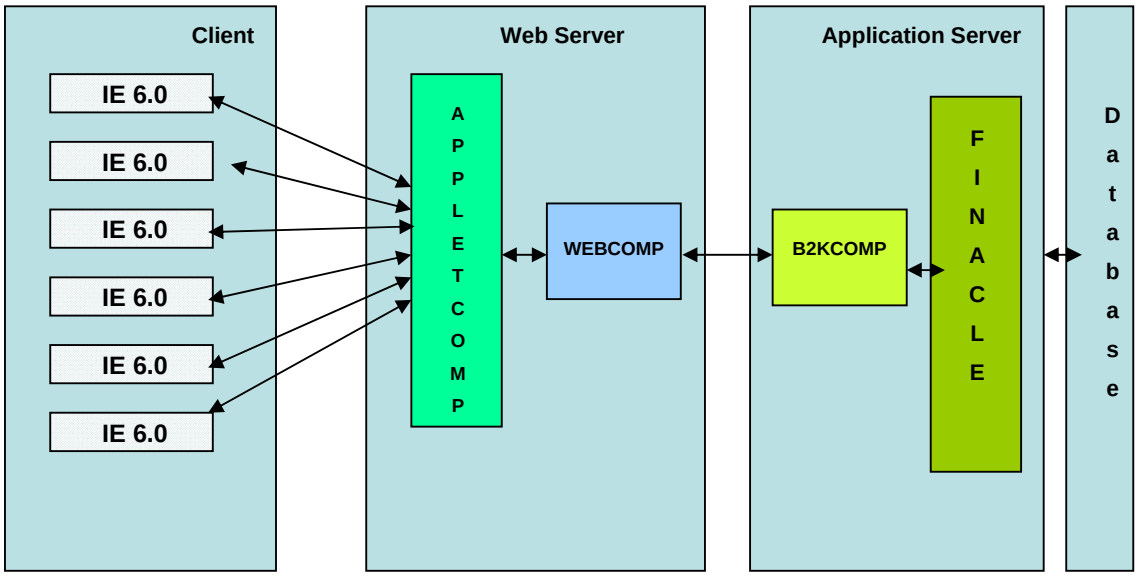
  We can access the Finacle application both from internet and intranet.

- Allows fast data-entry through use of function keys rather than the mouse

  The data entry and navigation in Finacle application happens through function keys. So it is easier for the user to operate.

16.1 STRUCTURE

The structure of Finacle web interface can be explained by the following diagram.



Client Side: -

The client side of Finacle web interface consists of a web browser through which the client can see the web pages and do the required operations. The web browser generally used is Internet Explorer 6.0. But any other browser can also be used.

Server side: -

The server side of Finacle web interface can be divided into two portions: -

i. **Components of web server**

Web Server accepts the http requests and sends the requested data as http response. Web server is responsible for the presentation part of Finacle application. It helps in loading

the web pages. It has mainly two components – Appletcomp and Webcomp.

a. **APPLETCOMP: -**

Appletcomp stands for applet component. It contains the Finacle applet. The Finacle applet is responsible for presenting data received from the server in a graphical user interface. This applet consists of a set of JAVA class files bundled together in a jar (java archive) file. Finacle applet is re-usable, that means the same applet can be used in different servers.

Appletcomp is responsible for the appearance of the web page. Actually there is only one applet. It gets re-painted for every web page. So the layout of every web page is similar.

b. **WEBCOMP: -**

Webcomp stands for web component. It is the place where all the web pages reside. Appletcomp and Webcomp are the components of web server. Upon the client sends the request from the browser the requested page is picked up from the Webcomp and it is displayed on the browser. If client wants to do some operation it can communicate with the application server through the web comp. Appletcomp and Webcomp can be used with any platform like windows or SunOS.

ii. **Components of application server**

Application server helps the clients to access the database and do the necessary operations. It does the business logic for the client. **B2KCOMP** is the main component of Finacle application server.

B2KCOMP: -

Clients can access the application and the database through b2kcomp. It is responsible for the login service. B2kcomp allows the users to login to the application. It also takes care of the session management. B2kcomp is installed in the same server where the Finacle application is present. So it is specific to the platform in which the application server is operating.