Santa Clara University

COEN 233 - Computer Networks

Project

# WEB SERVER

By
Chetan Pangam
(School of Engineering)

**Prof. Dr. Kevyan Moataghed**

## Audience-

This paper has collective information regarding web servers with respect to computer networks. It describes the basic concepts of webserver, requirements and maintenance of webservers and the operation of webservers through the networks.

The paper is perfectly suitable for the person who is new to networking and keen to know about webservers as well as the networking person who wants to know about a webserver and its working as well as the fresh developer wanted to understand the client server model. This document will give overall idea of webserver to every new beginner.

As mentioned any beginner can read this document, there is no such prerequisite required to know before start reading. Having said that, having knowledge of basic reference model and layers of them is plus for the reader

## Acknowledgement-

I would like to express my sincere gratitude to prof. Dr. Kevyan Moataghed for his constant guidance and valuable support. Working on this project “Web Servers” was a source of immense knowledge to me.

I would also like to thanks our grader Priya Maheshwari for providing regular inputs as well as support. Her valuable guidance was recommendable. Thank you to all who directly or indirectly help me out throughout the project.

## Table of contents-

## Table of Figures-



## Table of Tables-

## 1. Introduction

A web server is software which process and responds to the HTTP requests from clients and provide HTML data. A web server is key of client server programming model. A web server plays important role in web hosting. Server side scripting is an additional feature of Web server which can deliver dynamic web pages. After client initiates request, the transformation of domain name to IP address takes place and reaches to actual hosting server which provides appropriate response. All important topics are distributed in 5 chapters.

Chapter two will tell about the concepts of a webservers, history related to webserver, their types and explanations, features of webservers.

Chapter third describes about the requirements of webservers. Basically, it gives idea of Hardware and software requirements with respect to operating systems.

Chapter four explains the detail working of webserver. It also tells about the key terminologies used in the working process of webservers.

Chapter five describes about the maintenance of the webservers. It is important and key point which need to take care during maintenance.

## 2. Concept

A web server is a software which deals with HTTP requests. It is responsible for accepting request from web browsers that is client and serves them HTTP response with related data. The response is usually web page like HTML document and linked objects.

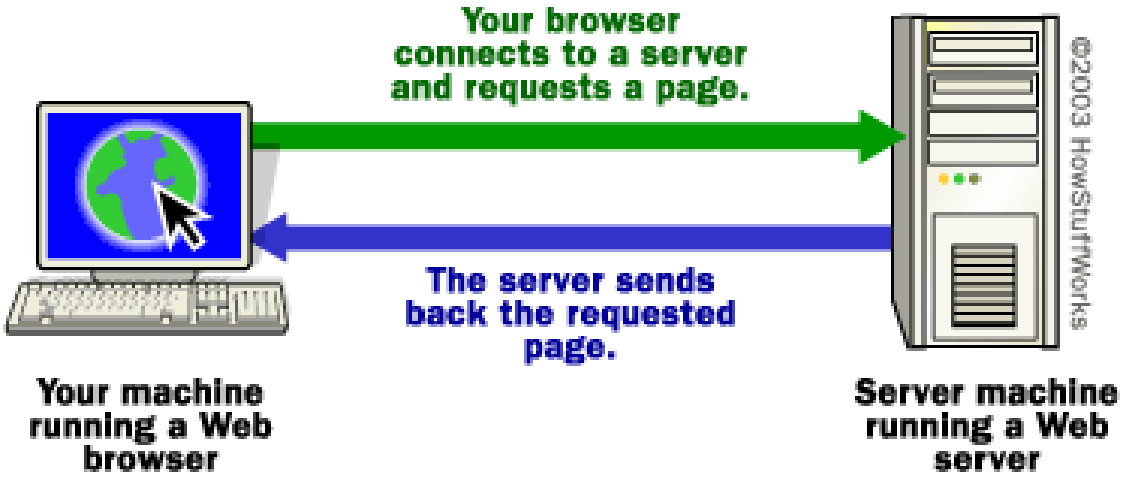


Fig. 1- Basic concept of Webserver

### 2.1 Overview

A web server has primary function to store, process and deliver web pages to clients. These web pages are most commonly HTML documents having images, CSS and scripts along with the text content. A client and server communicate with each other using Hyper Text Transfer Protocol (HTTP). It is application layer protocol for world wide web(www).

These are remote computers or programs which delivers web pages through internet to end user as per request from web browser. Every web server get identified through its domain name and IP address. The domain name is a string which identifies web pages within the internet [1]. This name is registered in DNS and it is used for addressing purpose. IP address is numerical identification to each device in computer network which uses Internet protocol for communication. There are two versions of IP, IPv4 and IPv6. DNS translates the domain name into the IP address.

The client is commonly web browser or web crawler. A web browser is a software application to retrieve information to user, display that information and navigate on other information. Web crawler, also called as spider, is an internet bot for web indexing purpose which browses world wide web systematically. The client

make request using HTTP for resources and web server provides response which is typically a real file on primary storage or secondary storage depending on implementation of web server. The response has the required content or error message if not able to get the resource.

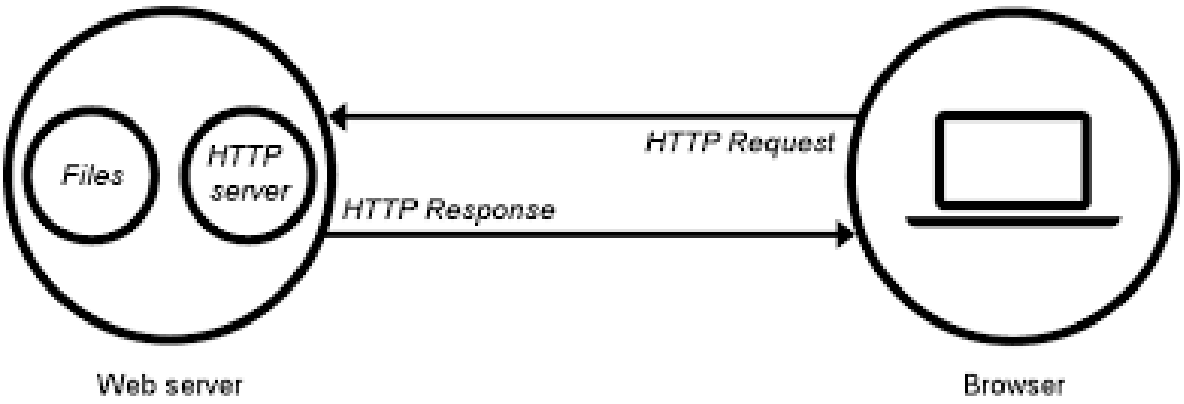


Fig. 2- Overview of Web server request-response

Even though the primary function of web server is to server HTTP response, HTTP implementation also has way to receive content from client which is used in file uploading and web form submission. Along with serving the world wide web, a web server also embedded in devices like printers, routers, webcams and serve only a local network. It used as a part of system for device monitoring.

Many web servers are reinforcing server side scripting using scripting languages like PHP, Active server pages (ASP). It gives advantage of without changing actual sever software, we can script the behavior of the web server in separate scripted file. It helps to generate dynamic data in response to client's request instead of static HTML content. The below diagram shows the process of handling request for a PHP file.

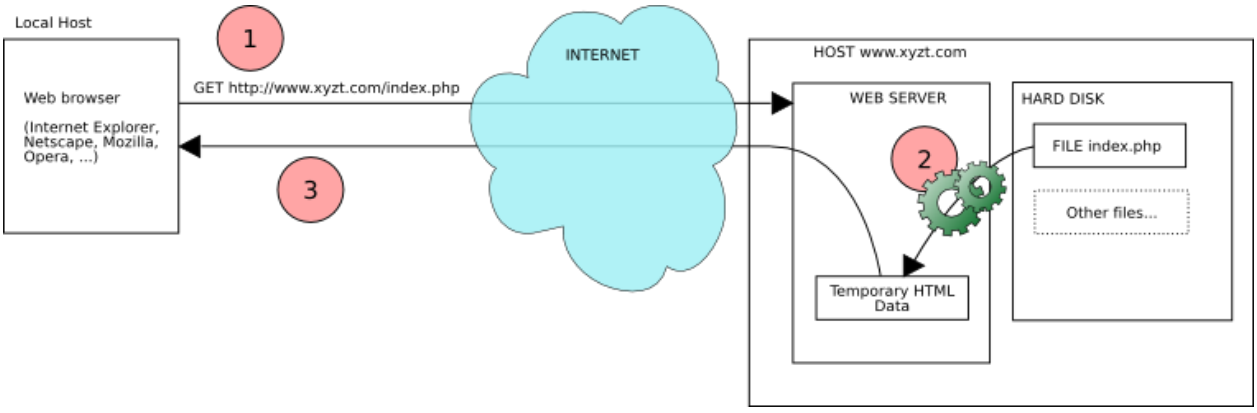


Fig. 3 – Web server handling PHP request

First, the browser sends the request to web server for PHP file through URL for example http://www.wxyt.com/index.php. The web server identify the .php extension, and get the file from storage. Then it processes the php file to generate temporary HTML data. And this temporary generated data sends as a response by web server to client [2]. As the processing done in server itself, it's called as server side programming.

### 2.2 History

In 1989, Tim Berners-Lee kept new project proposal to his CERN employers with the objective to exchange an information using a hypertext system. The results of project were showed in 1990 where he wrote two programs-

1. A browser which then called as World Wide Web
2. First web server of the world, known as CERN httpd


Fig 4 – A first web server of the world with label "This machine is a server. DO NOT POWER IT DOWN!!"

This easy and effective technology used to surf and exchange data through WWW support them to move on different operating systems between 1991 and 1994. By that time, the use of this technology was increased in scientific centers and universities as well as in industry. To regulate the future development through standardized process, Berners-Lee decided to form the World Wide Web Consortium (W3C).

| Web Servers | 2016 | 2016 | 2016 | 2016 | 2016 | 2016 | 2017 | 2017 | 2017 |
|---|---|---|---|---|---|---|---|---|---|
| | 1-Feb | 1-Apr | 1-May | 1-Jul | 1-Sep | 1-Nov | 1-Jan | 1-Feb | 21-Feb |
| Apache | 55.20% | 53.90% | 52.80% | 52.20% | 51.80% | 51.50% | 50.90% | 50.70% | 50.50% |
| Nginx | 27.10% | 28.70% | 29.90% | 30.50% | 30.80% | 31.30% | 32.10% | 32.40% | 32.60% |
| Microsoft-IIS | 12.30% | 12.00% | 12.00% | 12.00% | 12.00% | 11.90% | 11.60% | 11.60% | 11.50% |
| LiteSpeed | 2.30% | 2.30% | 2.30% | 2.30% | 2.30% | 2.30% | 2.30% | 2.40% | 2.40% |
| Google Servers | 1.40% | 1.40% | 1.40% | 1.30% | 1.30% | 1.30% | 1.30% | 1.30% | 1.30% |
| Tomcat | 0.50% | 0.50% | 0.50% | 0.60% | 0.60% | 0.60% | 0.60% | 0.60% | 0.60% |
| IdeaWebServer | 0.30% | 0.30% | 0.30% | 0.30% | 0.30% | 0.30% | 0.30% | 0.30% | 0.30% |
| Node.js | 0.20% | 0.20% | 0.20% | 0.20% | 0.20% | 0.20% | 0.20% | 0.30% | 0.30% |
| Apache Traffic Server | 0.20% | 0.20% | 0.20% | 0.20% | 0.20% | 0.20% | 0.30% | 0.30% | 0.30% |
| Tengine | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.20% | 0.20% |
| Cowboy | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% |
| Lighttpd | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% |
| IBM Servers | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% |
| Oracle Servers | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% | 0.10% |

Table 1 - trends in usage of web servers since February 2016 till now [3].

The table gives the idea of recent trends in usage of web servers. The figures are important in case of Apache and Nginx. In last one year, the percentage of Nginx users has been increasing on the other hand the percentage of Apache users decreasing. Other web servers are quite stable that is not much increment or decrement.

### 2.3 Types of Web servers

There are two basic type of web server depending how they are implemented, Shared web server and Dedicated web server. It is really challenging to choose in between these two. We need to do research about them and we need to choose accordingly. Here we will see some points of each of them.

#### 2.3.1 Shared Web Servers-

In shared web servers, the resources are shared with other users. So, a single web server allocated to many clients and they will have their website along with our web site on that web server. As others are sharing the server, we are allotted with limited disk space and bandwidth. If we exceed the allocated limit, then there will be a charge for it. It becomes cheaper as the cost of server gets divided into number of users sharing that server.

However, there is high security risk in the shared web servers as others using same web server. Also, the response time and performance degradation happens when the resources are shared and there is a high network traffic. There is risk of getting blacklisted by search engine as someone else might do bad practices which may blacklist entire IP address [4].

### 2.3.2 Dedicated Web Servers

As name suggests, it is available for only one user. It is ideal for the website having heavy web traffic. There will be only one website hosted on this server that's ours. It will be costly as nobody will share cost and we should have to pay for whole server. There is no limit on disk space and bandwidth so there will be no upper limit for extra charges.

A dedicated web server will be more secure, until we do some unethical thing as we are the only one on the web server. So, the chances of getting backlisted is depends on us and the activity we performed on server. As we have dedicated bandwidth, there will be no problem of poor response time or bad performance. As we are the owner of web server, we must have some skilled person to manage and control thing related to web server.

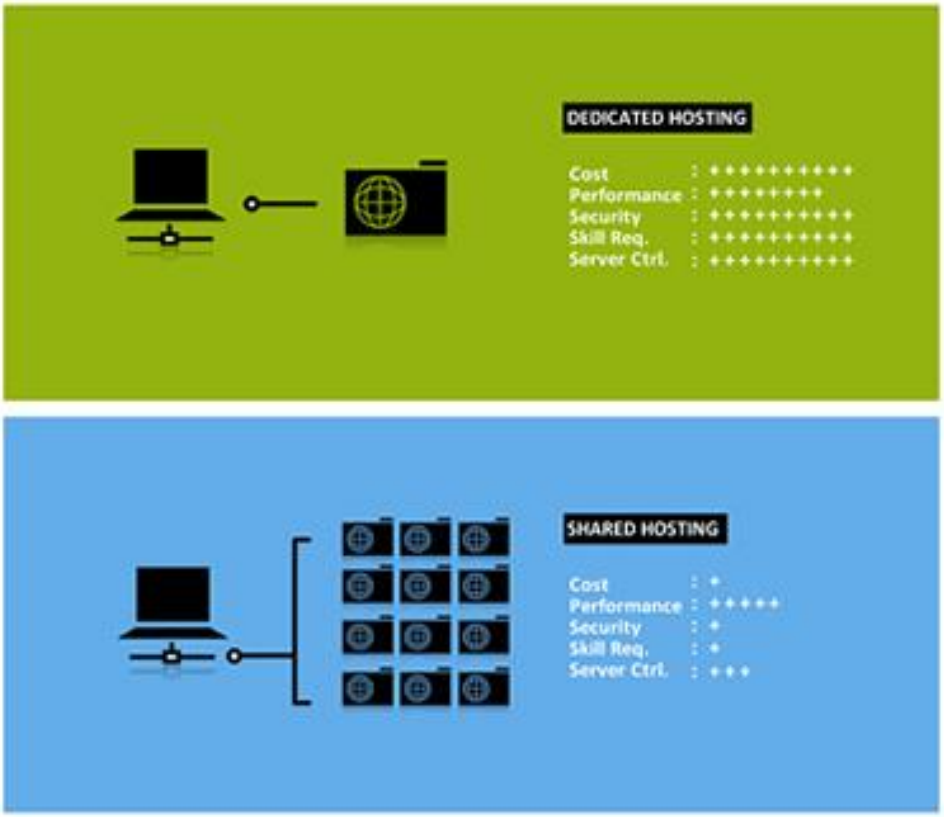


Fig 5 – Dedicated Vs Shared web hosting server.

### 2.4 Features of Web server

Most of the web browsers have a common set of features. As web servers are designed specifically for website hosting, all features are related to creating and maintaining the hosting environment. Some the key features are as below-

- Large data storage support-

  It is an important and key feature of web server. A web server supports large data storage space to store data from multiple websites.

- Bandwidth controlling-

  This feature is used to minimize the over network traffic. The host environment can set the bandwidth values to regulate the traffic over internet which avoid the downtime causes due to heavy network traffic.

- Server side scripting-

  This feature provides dynamic web pages as a response. This is done by scripting languages at server side.

- Virtual hosting-

  This serves multiple websites using same IP address. It hosts multiple domain name on single server. The resources of server like memory, CPU get shared

- Log file configuration-

  We can configure which kind of data can get stored in log file and where it will be saved.

- Error message configuration-

  It will help to show user-friendly error message to user so that we can decide which page to display for specific error e.g. 404 error

- Creation of virtual directories and mapping them with physical directories.

## 3. Requirements-

If we are running business which depends on web, we must want to keep our website up for 24*7. It should have good performance throughout. A hardware and software combination give the desired output. So, the requirements of web server are important while setting up the server. It depends upon various factors. There are two kind of requirements known as Hardware requirements and Software requirements.

### 3.1 Hardware requirements

As we are discussing about the hardware requirements, the processor, memory and storage are come in the picture. What will be the best combination will give me best output? Mostly its depend upon the size of business and number of instances that visiting our web site.

The minimum hardware requirements are difficult to infer for small instances. Here are some figures which are absolutely minimum requirement and we will require better hardware than this.

For 5 Concurrent Users-

- **CPU:** 2 x Intel Core 2 (2.66 Ghz, 128K cache)
- **RAM:** 4GB
- **Minimum database space:** 10GB

For 25 Concurrent Users

- **CPU:** Quad 2GHz+ CPU
- **RAM:** 6GB
- **Minimum database space:** 10GB.

Here is some example for the hardware specification. The RAM mentioned below is not confirmed records as it is either allocated to JVM or else it is server total memory. The table have number of CPUs, speed of CPU and RAM as per the accounts active. Active account is from the time index page accessed till the connection get closed.

| Accounts | CPUs | CPU (GHz) | RAM (MB) |
|---|---|---|---|
| 150 | 1 | 2.6 | 1,024 |
| 350 | 2 | 2.8 | 1,536 |
| 5,000 | 4 | 3 | 2,048 |
| 10,000 | 2 | 3.8 | 2,048 |
| 10,000 | 2 | 3.6 | 4,096 |
| 21,000 | 2 | 3.6 | 4,096 |
| 85,000 | 4 | 2.6 | 4,096 |

Table 2 - Example Hardware Specification [5].

### 3.2 Software Requirements

The software requirements vary depending upon the operating system under use. The software requirement involves supportive operating system, supportive browsers, scripting languages, database, scheduler etc.

We will see the software requirement for Windows server and Linux like server. We will compare them by software differences they have.

| | Windows Server | Linux like Server |
|---|---|---|
| Web server Software | • Apache v1.3+<br>• Apache v2.2+<br>• IIS6<br>• IIS7<br>• IIS8 | • Apache v1.3+<br>• Apache v2.2+<br>• LigHTTPD v1.4.12+<br>• Ngnix |
| Scripting Language | • PHP v5.3.*, v5.4.*, v5.5.*, v5.6.* (VC9 x86 Non Thread Safe is recommended for IIS).<br>• PHP v5.2.12+, v5.3.*, v5.4.*, v5.5.*, v5.6.* (for Apache) | • PHP v5.2.12+, v5.3. *, v5.4.*, v5.5.*, v5.6.* |
| | | |

| | | |
|---|---|---|
| Database | • MySQL Community Server v5.1+<br>• Microsoft SQL 2005 or higher | • MySQL Community Server v5.1+ |
| Scheduler | Windows Scheduler (always available as standard Windows tool) | Cron (commonly available) |

Table 3 – Software requirement for Windows and Linux servers [5]

## 4. Operation

If we want any information, we just search it on google and go through the information retrieved. But have we ever think that what will be the logic which display us the related information. Whenever we clicked on any link or enter the web address in URL, what process takes place behind the scene to get the data on our screen.

Considering that I am sitting at my laptop, surfing through web, and I get an email from professor having important links for projects. I check it out and it is http://www.networkcomputing.com/home.html. So, I paste it in the URL of my browser and just press enter. I get useful information within a minute. but from where the URL becomes live.

### 4.1 Behind the scenes

To understand what actually happens behind the scene, I get into more detail of this process and found some interesting things. Below are some basic steps happens once we press the enter.

- The URL gets break into three parts by the browser:
    1. Protocol – it is nothing but “http”
    2. Server name – that will be www.networkcomputing.com
    3. File name – that will be "home.html"
- Then browser ask for IP address for the server name it has from Domain mane server(DNS). Using that IP address, it establishes connection with server machine of networkcomputing.com.
- As per the HTTP protocol, GET request has sent to server by browser, to get the file “home.html” from network computing server.

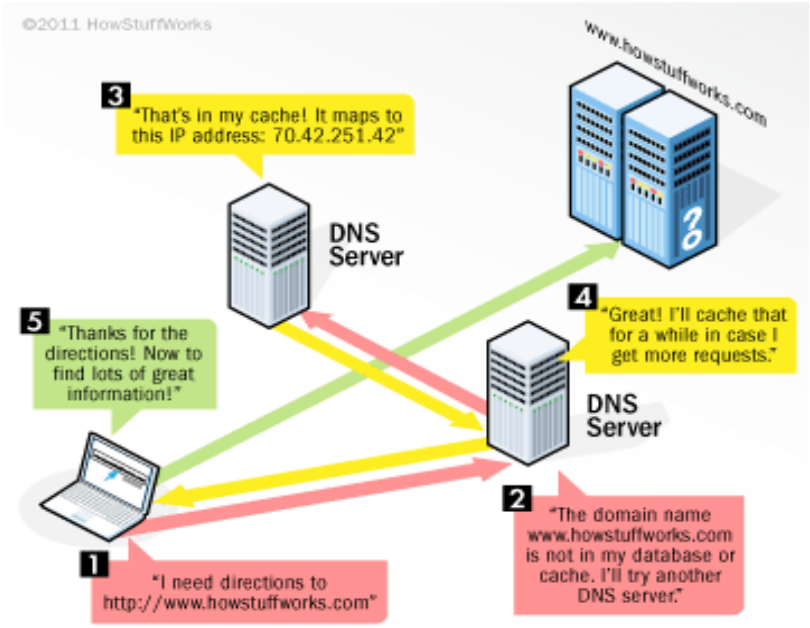


Fig 6- Actual working happens behind the scene

- Server respond with HTML text to the browser and then browser display the information in formatted manner by reading HTML tags.

This is how the we can able to see the related information within few seconds. To understand this process in more detail we will understand the concept appears in the whole process one by one.

### 4.2 The Internet

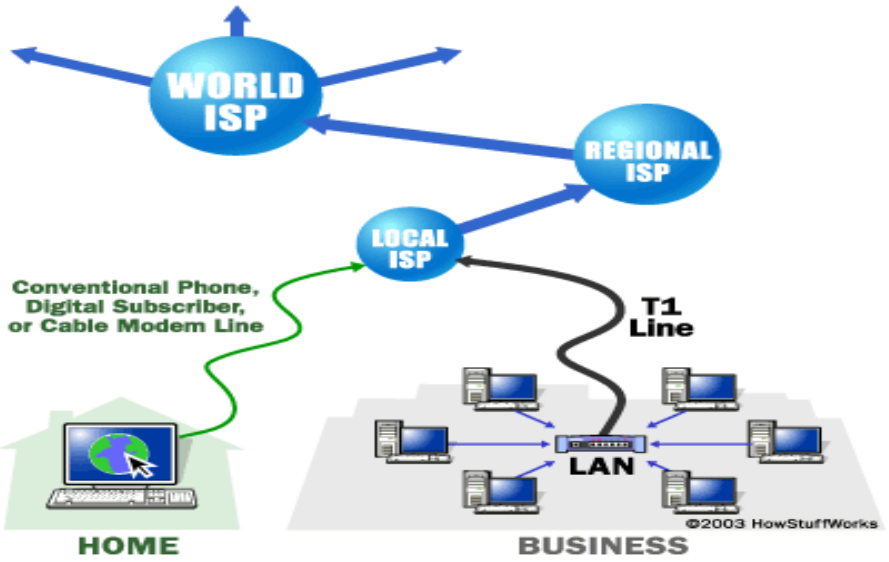


Fig 7 – The Internet overview

The internet is collection of number of computer systems, interconnected together through a computer network. Due to interconnection, all the systems can communicate with any other computer in the network. Our home computer system or laptop is connected to an internet service provider (ISP) via wired or wireless connection. A computer systems in organization or University are connected to LAN using Network Interface card (NIC) and LAN is connected to ISP using high speed telephone line which can handle approximately 1.5 million bits per second. These ISPs are interconnected and connected to largest ISPs through the fiber-optics backbones. All Backbones in the world are connected to through fiber-optics lines which are under sea or satellite links.

### 4.3 Clients and Servers

Every machine in the network can be server or client. Systems which serve the request to other systems are called as Servers. And systems those send requests for the services are called as clients. When we connect to google.com and ask for some information, google provides the server which serve our request whereas we ask for the service through request hence, our machine is client or user machine. Machine can be both server and client too.

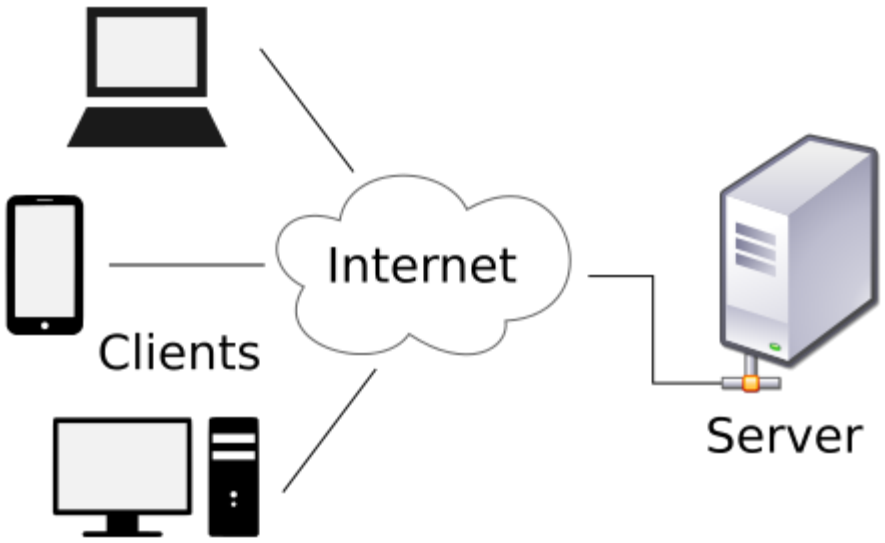


Fig 8 – Client – Server architecture

A server can serve multiple services, and not specific to any one service. For example, a server can provide web service, email service or FTP service. Different clients can request for the service they are interested in. If I opened my web browser, I mostly request for web pages or if I opened my email application, then I must be interested in email server and so on.

### 4.4 IP Address

Every system in the internet assigned with unique id called as IP address. The IP address is of 32 bits which is an Internet Protocol Version 4. As internet growing so fast and we were not having sufficient IP address available, a new IP version (IPv6) was developed in 1995 having 128 bits as address. IPv4 represented in four octets and separated by decimal dot. For example 216.72.16.317. An IP address is sufficient to reach host server. If we type http:// 216.72.16.317 instead of google.com will take us to the right server.

A server machine has fixed IP address that is called Static IP address which does not changed. A system at has IP address allocated by ISP when system connects to internet. It will be unique for that particular session but as we connect again next time, it will have another IP address for that session. To know about internet information, we can use 'IPCONFIG' command in windows operating system and 'nslookup' command use in UNIX operating system.

### 4.5 Domain Names

IP address is the actual address for the host server but it is difficult to remember number string for every host and sometimes IP addresses needs to change so, all servers have readable string called as Domain names. For example, www.google.com. It is readable, easy to remember and permanent name. It is easy to remember google than 216.72.16.317.

The domain name www.google.com has three parts-

1. www is the host name
2. google is the domain name
3. com is the top-level domain name.

These top-level domain names are managed by different registrar. VeriSign is the registrar who manages .com and .net top-level domain. RegistryPro, NeuLevel, public level registry are some other important registries.

### 4.6 Name Servers

The mapping of domain names to IP address is done in domain name server(DNS). These servers are nothing but databases which stores the name with address and its distributed in the network. A central name server takes input from VeriSign and map IP address to domain name.

If we type http://www.google.com/home.html into URL, then browser extract s the domain name www.google.com and send to domain name server. Then DNS maps the proper IP address for the given domain name. In Unix operating system, we can get the IP of any host by using nslookup command. We need to just put domain name along with command, it will respond with the correct IP of server.

### 4.7 Ports

Server system will provide the service to the request if the request has appropriate Port number. Every service has one port number to mapped to it. For example, web service will be available at port 80 whereas FTP service will be available at port 21. Client has to connect with IP address and respective port number to get specific service.

| Sr No | Service name | Port number |
|---|---|---|
| 1 | Echo | 7 |
| 2 | Daytime | 13 |
| 3 | gotd | 17 |
| 4 | FTP | 21 |
| 5 | telnet | 23 |
| 6 | SMTP (Simple Mail Transfer Protocol) | 25 |
| 7 | Time | 37 |
| 8 | Nameserver | 53 |
| 9 | nickname | 43 |
| 10 | gopher | 70 |
| 11 | Finger | 79 |
| 12 | WWW | 80 |

Table 4 - Well-known services at a well-known port [6].

## 5. Maintenance

The maintenance is key part of the software lifecycle. Maintenance of software is necessary to make them function as they were at start. Doing regular check and maintenance, we can identify the issue before it arises. It also keeps software updated with performance and security upgrades. We generally do maintenance by taking back up and restoring our data, Analyzing Web logs, analyzing the tables and recompiling objects in case of database server, enhancement of the security etc.

Below is some addition task that might be performed to maintain the web server-

- To check server log file regularly, check whether suspicious activity happened
- To verify how much memory every service taking and make sure sufficient memory is available.
- Check the changes happened in the folder permission. It might have chances that most common and required folder get inaccessible.
- Make sure proper and required data redundancy is taking place. It is data which is more important
- Check all the security features are up to the mark or not. It might create loophole to get attacked on system
- Regularly update the security patches to protect the server from malware attacks. Definition should get updated and installed.
- Go through the server logs and check whether any attempt of hacking has been done or not. It is crucial as if ignore such things than it will make bigger impact on server and ultimately on business.
- Take updates of all important and critical services and software.
- Take regular backups of data in case of system failure

### 5.1 Web server vulnerabilities-

Below are some vulnerabilities found in web servers which easily get spread in the server and network which needs to take care while the maintenance of the server-

1. Un-validate input
2. Parameter Tampering
3. Cookie poisoning
4. Cross site scripting
5. SQL Injection
6. Denial of Services
7. Improper error handling

## 6. Conclusion and Future work-

Thus, we have successfully completed this project and studied about web servers and its working through this document. We have gone through the history, types and features of web servers. We also saw the detailed working and related terminologies we came across during the study. We also studied the hardware and software requirements of web server. We understood the maintenance which is key factor of web server.

There is lot of new research going in web server technology. A Courteny Walker, the owner of courserv.com has published new paper [7]. The research paper states that the dynamic web server which is currently in use are slow and expensive. He said that reputed cloud storage providers such as Amazon, google and Microsoft adding static web site features in their cloud storage services which will allow user to create strong and effective web sites using their existing storage account.

This will provide faster web site as its static and using cloud services that too using previous account which make this cheaper. It will impact on traditional web site hosting providers as it is impossible offer same performance as cloud provider will provides.

## 7. Acronyms

ASP – Active Server Pages

CPU – Central Processing Unit

CSS – Cascading Style Sheet

DNS – Domain Name Server

FTP – File Transfer Protocol

GB - Gigabyte

GHz – Giga Hertz

HTML – Hyper Text Markup Language

HTTP – Hyper Text Transfer Protocol

IP – Internet Protocol

IPv4 – Internet Protocol version 4

IPv6 – Internet Protocol version 6

LAN – Local Area Network

MB- Megabyte

NIC – Network Interface Card

PHP- Hypertext Preprocessor

RAM- Random Access Memory

URL - Uniform Resource Locator

W3C – World Wide Web Consortium

WWW – World Wide Web

## 8. References:


1. http://www.webopedia.com/TERM/D/domain_name.html
2. http://fhoerni.free.fr/comp/web_server.html
3. https://w3techs.com/technologies/history_overview/web_server
4. https://www.pueblowebdesign.com/website-hosting
5. https://confluence.atlassian.com/doc/server-hardware-requirements-guide-30736403.html
6. http://computer.howstuffworks.com/web-server.htm
7. https://www.linkedin.com/pulse/future-web-hosting-courtney-walker