



NEPAL COLLEGE OF INFORMATION TECHNOLOGY

COURSE CODE :CMP483.3

WEB TECHNOLOGY

Course Manual on Web Technology

Author:

Roshan Kr. Shah

November 13, 2016

Contents

1	Unit 1: Fundamentals	2
1.1	Introduction to Internet	2
1.1.1	What is internet?	2
1.1.2	What is WWW?	3
1.1.3	What is Website?	3
1.1.4	What is Web Browser?	4
1.1.5	What is Web Server?	4
1.1.6	What is URL?	5
1.1.7	Multipurpose Internet Mail Extensions(MIME)	6
1.2	Overview of different protocols	6
1.2.1	What is HTTP?	6
1.2.2	What is SMTP Server?	7
1.2.3	What is FTP?	7
1.2.4	What is POP?	8
1.2.5	What is WAP?	8
1.2.6	Web Architecture	9
1.2.7	Web Standards	9
1.3	Domain name and Hierarchy	10
1.4	Domain name Registration Process	10
1.5	Web Hosting	11
2	Unit 2: Introduction to HTML And XHTML	12

1 Unit 1: Fundamentals

1.1 Introduction to Internet

The Internet is essentially a global network of computing resources. We can think the Internet as a physical collection of routers, circuits and other communicating devices as a set of shared resources. Few definitions provided previously include :—

- A network of networks based on the TCP/IP communications protocol.
- A community of people who use and develop those networks.

some of the basic services available to the internet users are —

Emial —A fast, easy, and inexpensive way to communicate with other Internet users around the world.

Telnet — Allows a user to log into a remote computer as though it were a local system.

FTP — Allows a user to log into a remote computer as though it were a local system.

UseNet news — Allows a user to log into a remote computer as though it were a local system.

World Wide Web WWW — A hypertext interface to Internet information resources.

1.1.1 What is internet?

WWW stands for **W**orld **W**ide **W**eb. A technical definition of the World Wide Web is — All the resources and users on the Internet that are using the Hypertext Transfer Protocol HTTP.

A broader definition comes from the organization that Web inventor Tim Berners-Lee helped found, the World Wide Web Consortium W3C: The World Wide Web is the universe of network accessible information, an embodiment of human knowledge.

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

1.1.2 What is WWW?

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

A broader definition comes from the organization that Web inventor Tim Berners-Lee helped found, the World Wide Web Consortium W3C: The World Wide Web is the universe of network accessible information, an embodiment of human knowledge.

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

1.1.3 What is Website?

A website XYZ.com is a collection of various pages written in HTML markup language. This is a location on the web where people can find tutorials on latest technologies. Similarly, there are millions of websites available on the web.

Each page available on the website is called a web page and first page of any website is called home page for that site. A web page refers to any document on the web. Every web page is identified by a unique URL in the internet and it contains hyper text i.e different types of text such as plain text ,graphics, animation, Audio, video etc.

Web technology relates to the interface between web servers and their clients. It includes markup languages, programming interfaces and languages, and standards for document identification and display.

A web application is any application that is specifically used through a web browser. Some examples of web applications include , ON Line Book Store , Shopping , Citrix Meta frame applications and any number of tools such as package tracking that are found on the Internet.

Accessing information on the Web

Information on the Web is stored in documents, using a language called **HTML** (HyperText Markup Language). Web clients must interpret HTML to be able to display the documents to a user. The protocol that governs the exchange of information between the Web server and Web client is named **HTTP** (HyperText Transfer Protocol). Any Web Application we consider It follows **Client – Server** Technology

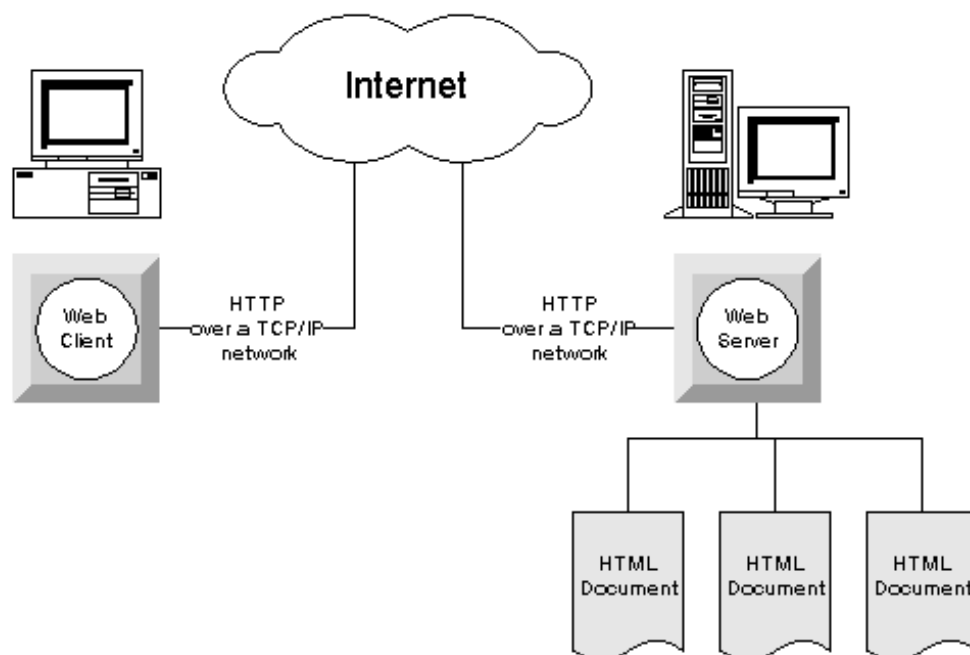


fig1:- Accessing Information on the web

1.1.4 What is Web Browser?

Web Browsers are software installed on your PC. To access the Web you need a web browsers, such as Netscape Navigator, Microsoft Internet Explorer or Mozilla Firefox. Currently you must be using any sort of Web browser while you are navigating through my site tutorialspoint.com. On the Web, when you navigate through pages of information this is commonly known as browsing or surfing.

1.1.5 What is Web Server?

Every Website sits on a computer known as a Web server. This server is always connected to the internet. Every Web server that is connected to the Internet is given a unique address made up of a series of four numbers between 0 and 256 separated by periods. For example, 68.178.157.132 or 68.122.35.127. When you register a Web address, also known as a domain name, such as tutorialspoint.com

you have to specify the IP address of the Web server that will host the site. We will see different type of Web servers in a separate chapter.

1.1.6 What is URL?

URL stands for **Uniform Resource Locator**, and is used to specify addresses on the World Wide Web. A **URL** is the fundamental network identification for any resource connected to the web i.e. g. , hypertextpages, images, andsoundfiles. A URL will have the following format –

```
protocol://hostname/other_information
```

URLs have 3 components: **A Prefix, A Hostname and A Path**

The protocol specifies how information is transferred from a link. The protocol used for webresources is HyperText Transfer Protocol HTTP. Other protocols compatible with most web browsers include FTP, telnet, newsgroups, and Gopher.

The protocol is followed by a colon, two slashes, and then the domain name. The domain name is the computer on which the resource is located.

Links to particular files or subdirectories may be further specified after the domain name. The directory names are separated by single forward slashes.

Example:

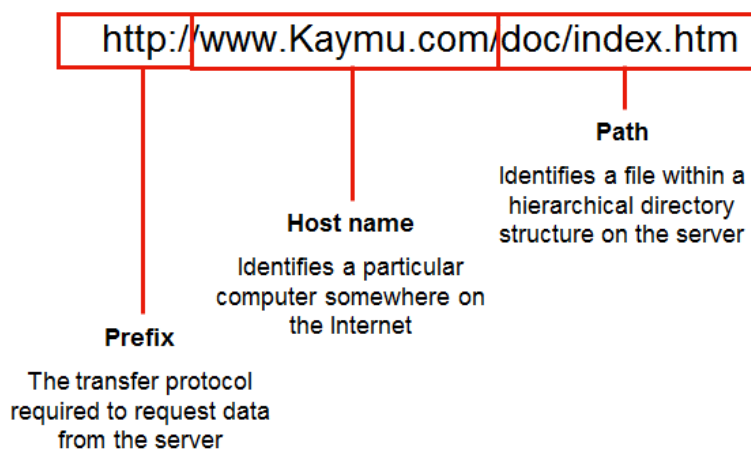


fig3:-Components of URLs

1.1.7 Multipurpose Internet Mail Extensions(MIME)

MIME (multi - Purpose Internet Mail Extension) is an extension of the original internet e - mail protocol that lets people use the protocol to exchange different kinds of data files on the internet: audio,video, images, application programs, and other kinds, as well as the ASCII text handled in the original protocol. It is a standard which was proposed by Bell Communicatoins in 1991 in order to extend upon the limited capabilities of email, and particular to allow documents (such as images, sound, and text) to be inserted in a message.

MIME adds the following features to the email service:

- Be able to send multiple attachments with a single message;
- Unlimited message length;
- Use of character sets other than ASCII code;
- Use of rich text (layouts, fonts,colors,etc)
- Binary attachments (executables,images, audio or viedo files) which may be divided if needed.

1.2 Overview of different protocols

For the data transfer in communication between client and server in web technology , we have to follow the various protocols. few of these are:-

1.2.1 What is HTTP?

HTTP stands for **H**ypertext **T**ransfer **P**rotocol. This is the protocol being used to transfer hypertext documents that makes the World Wide Web possible.

HTTP makes use of TCP to open connections between clients and servers and to pass the requests and responses between them.

When a user clicks a link or types a web address: A TCP connection is opened between browser and server,then requests and responses are sent back and forth using HTTP

HTTP is a stateless protocol, meaning each request for data is dealt with in isolation; once the server sends the response, it forgets everything about the original request. A standard web address such as [ABC.com](#) is called a URL and here the prefix http indicates its protocol.

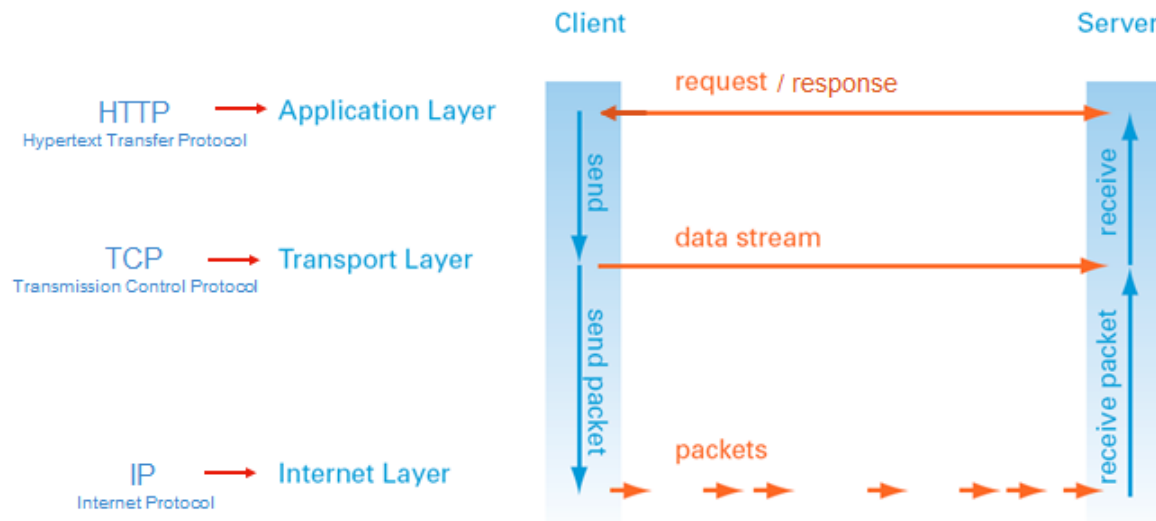


fig3:- Fetching Pages over the Internet: HTTP

1.2.2 What is SMTP Server?

SMTP stands for Simple Mail Transfer Protocol Server. This server takes care of delivering emails from one server to another server. When you send an email to an email address, it is delivered to its recipient by a SMTP Server.

1.2.3 What is FTP?

FTP stands for File Transfer Protocol. It is a protocol used to transfer files between an FTP host/server and an FTP client computer on the Internet. FTP is most commonly used to download files from the World Wide Web. It is an alternative choice to HTTP protocol for downloading and uploading files to FTP servers. FTP is used for:

- Uploading webpages to web servers for publishing on the Internet;
- Browsing and downloading files from public software sites;
- Transferring large files among two parties that are too large for email attachments;

- Distributing the latest revisions of programs by software developers.

1.2.4 What is POP?

The Post Office Protocol (POP3) is an Internet standard protocol used by local email software clients to retrieve emails from a remote mail server over a TCP/IP connection. Since the first version was created in 1984, the Post Office Protocol (currently at Version 3) has since become one of the most popular protocols and is used by virtually every email client to date. Its popularity lies in the protocol's simplicity to configure, operate and maintain.

Email servers hosted by Internet service providers also use POP3 to receive and hold emails intended for their subscribers. Periodically, these subscribers will use email client software to check their mailbox on the remote server and download any emails addressed to them.

Once the email client has downloaded the emails, they are usually deleted from the server, although some email clients allow users to specify that mails be copied or saved on the server for a period of time.

Email clients generally use the well-known TCP port 110 to connect to a POP3 server. If encrypted communication is supported on the POP3 server, users can optionally choose to connect either by using the STLS command after the protocol initiation stage or by using POP3S, which can use the Transport Layer Security (TLS) or Secure Sockets Layer (SSL) on TCP port 995 to connect to the server.

1.2.5 What is WAP?

Wireless application protocol(WAP) is an application environment and set of communication protocols for wireless devices designed to enable manufacturer- vendor, and technology - independent access to the internet and advanced telephony services. WAP is chosen due to following reason:

- Internet standards such as HTML,HTTP,TLS and TCP are inefficient over mobile networks;
- WAP utilizes binary transmission for greater compression of data and is optimized for long latency and low bandwidth.
- As WAP is based on a scalable layered architecture, each layer can develop independently of other. This makes it possible to introduce new transport protocols without changes in other layers.

1.2.6 Web Architecture

Web architecture is an approach to the design and planning of websites that involves technical, and functional criteria. As in traditional architecture, the focus is on the user and on user requirements. This requires particular attention to web contents, the business plan and usability. The web is a two-tiered architecture consisting of a **web browser** which displays information content and a **web server** that transfer information to the client.

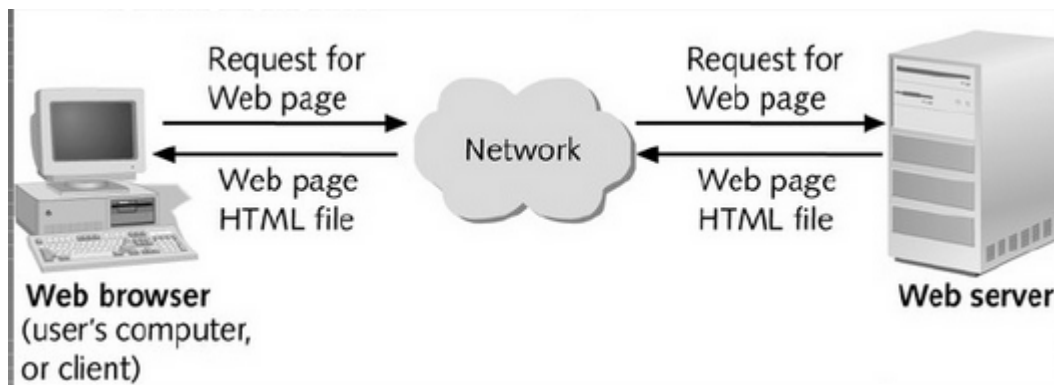


fig3:-Web Architecture

1.2.7 Web Standards

Web standards are recommended by W3C committee designed to deliver greatest benefit to largest number of web users. Using these standards any web document can work across multiple browsers, platforms, and Internet devices. Also this standard will continue to work as new devices are invented. For example:

structural languages: HTML, XML, XHTML;

presentation languages: CSS

object models: W3C DOM etc.

1.3 Domain name and Hierarchy

Every computer connected to the Internet must have a unique IP address, no matter whether it's a client or a server (or both). An IP address is just a number that identifies a host on the Internet. Example:

212.171.218.17 or 144.214.5.127

The Domain Name System (DNS) is a database that matches IP addresses to host names. And the Domain Name System (DNS) translates host names into IP addresses, which are then used by TCP to establish connections between HTTP clients and servers. Domain names are administered in such a way that they are guaranteed to be unique.

Domain names are organized in a hierarchical structure...

Top Level Domain – Mostly top level domain (TLD) contain country name like .hk, .np, .in, .uk etc or generic top level domain (gTLD) like .com, .org, .net, .info, .biz, .tv etc

Second Level Domain – At this level it contains the actual name of organization or service. It may contain letters (a to z), numbers (0 to 9), dashes (-). like kamyu, apple, google etc.

Third Level Domain – These are strings of characters that designate different services, or host within the second level domain. e.g. "WWW" for the core or main website.

1.4 Domain name Registration Process

Registering a domain name can either be done directly with a registration service, such as HKDNR, or through a website hosting service. Either way, you have to pay a fee for domain registration that is separate from any site hosting fees you may pay.

gTLD domains (.com, .org, .net): US\$12 – 15 *per year*

Country domains in Hong Kong:

.com.hk, .org, .hk, .net.hk: HK\$200 *per year*

.hk : HK\$250 *per year*

1.5 Web Hosting

Web hosting is a service which allocates space for customers to showcase their web-sites on computer servers that are connected to the internet 24 /7. The major type of web hosting are :

- **Free web hosting Service** :Free, (sometimes) advertisement- supported web hosting, and is often limited when compared to paid hosting

Merits –Low cost.It's free, Good for family, hobby or personal sites. Free email is often and option.

Demerits –No domain names..But it has few, limited, or no software options. Limited security options, Limited or no database support with limited technical supports.

- **Shared web hosting services**:One's website is placed on the same server as many other sites, ranging from a few to hundreds or thousands. Typically,all domains may share a common pool of server resources, such as RAM and the CPU.The features available with this type of service can be quite extensive. A shared web site may be hosted with a reseller. Browsing and downloading files from public software sites;

Merits –Low cost.Cost is shared with others. Good for small business and average traffic. Multiple software option.Good support.

Demerits –Reduced security due to many sites on one server. Restriction on traffic volume. Restricted database support. Restricted software support.

- **Virtual Dedicated Server**: Dividing a server into virtual servers, where each user feels like they are on their own dedicated server, but they are actually sharing a server with many other users. the users may have root access to their own virtual space. this is known as virtual private server **VPS**. Customers are sometimes responsible for patching and maintaining the server.

Merits –Low cost.Cost is shared with others. Good for small business and average traffic. Multiple software option. Own domain name. Good support.

Demerits –Reduced security due to many sites on one server. Restriction on traffic volume. Restricted database support. Restricted software support.

- **Dedicated hosting Service**:the user gets his or her own web server and gains full control over it (root access for Linux/administrator access for Windows);however, the user typically does not own the server. Another type of

dedicated hosting is self - managed or unmanaged. This is usually the least expensive for dedicated plans. The user has full administrative access to the box, which means the client is responsible for security and maintenance of his own dedicated box

Merits –Good for large business. Good for high traffic. Multiple domain names. Powerful email solutions. Powerful database support. Strong(Unlimited) software support.

Demerits –expensive. Reuired higer skills.

2 Unit 2: Introduction to HTML And XHTML

References

- [1] Help on BibTeX entry types. <http://nwalsh.com/tex/texhelp/bibtex-7.html>. Accessed: 2015-03-12.
- [2] Help on BibTeX entry types. <http://nwalsh.com/tex/texhelp/bibtex-7.html>. Accessed: 2015-03-12.