Marwadi University

01CE0306-Web Technology

Unit-1 Introduction to Web

Prof. Vaibhav K Matalia
Computer Engineering Department

Introduction to WWW

Marwadi University

- The World Wide Web is abbreviated as WWW and is commonly known as the web.
- World Wide Web is a collection of websites or web pages stored in web servers and connected to local computers through the internet.
- These websites contain text pages, digital images, audios, videos, etc. Users can access the content of these sites from any part of the world over the internet using their devices such as computers, laptops, cell phones, etc.
- The WWW, along with internet, enables the retrieval and display of text and media to your device.
- The World Wide Web is based on several different technologies: Web browsers, Hypertext Markup Language (HTML) and Hypertext Transfer Protocol (HTTP).
- The World Wide Web was originally designed in 1991 by Tim Berners-Lee.

Internet

- It connects millions of computers together globally, forming a network in which any computer can communicate with any other computer as long as they are both connected to the Internet.
- The information travelling over the Internet uses a variety of languages known as protocols. So, we can say that Internet is network of computers which connected together and any computer communicate with any other computer.
- Use the standard Internet Protocol suite(TCP/IP) for communication.

Difference between Internet and WWW

Marwadi University

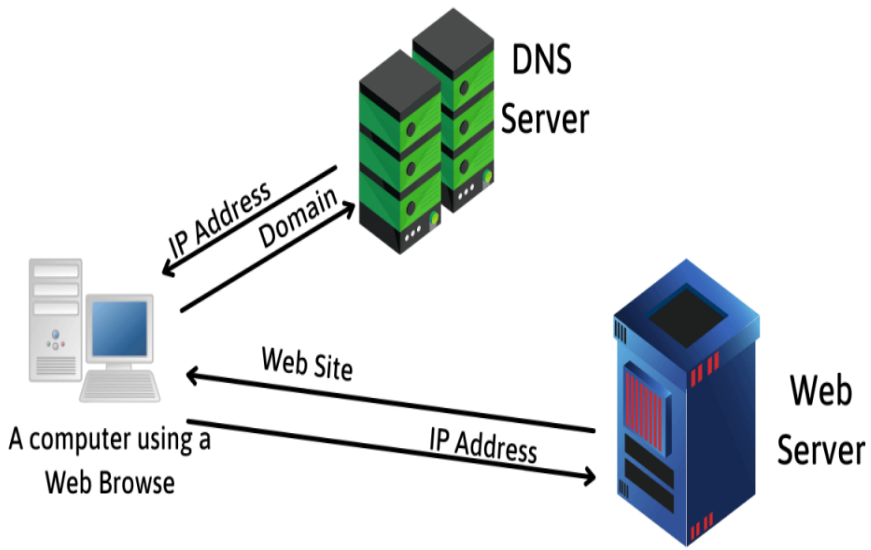
Internet	WWW
Internet is a means of connecting a computer to any other computer anywhere in the world.	World Wide Web which is a collection of information which is accessed via the Internet.
Internet is infrastructure.	Web is a service on top of that infrastructure.
Internet can be viewed as a big book-store.	Web can be viewed as collection of books on that store.
Internet is hardware-based.	Web is software oriented.
Internet uses IP address.	WWW uses HTTP

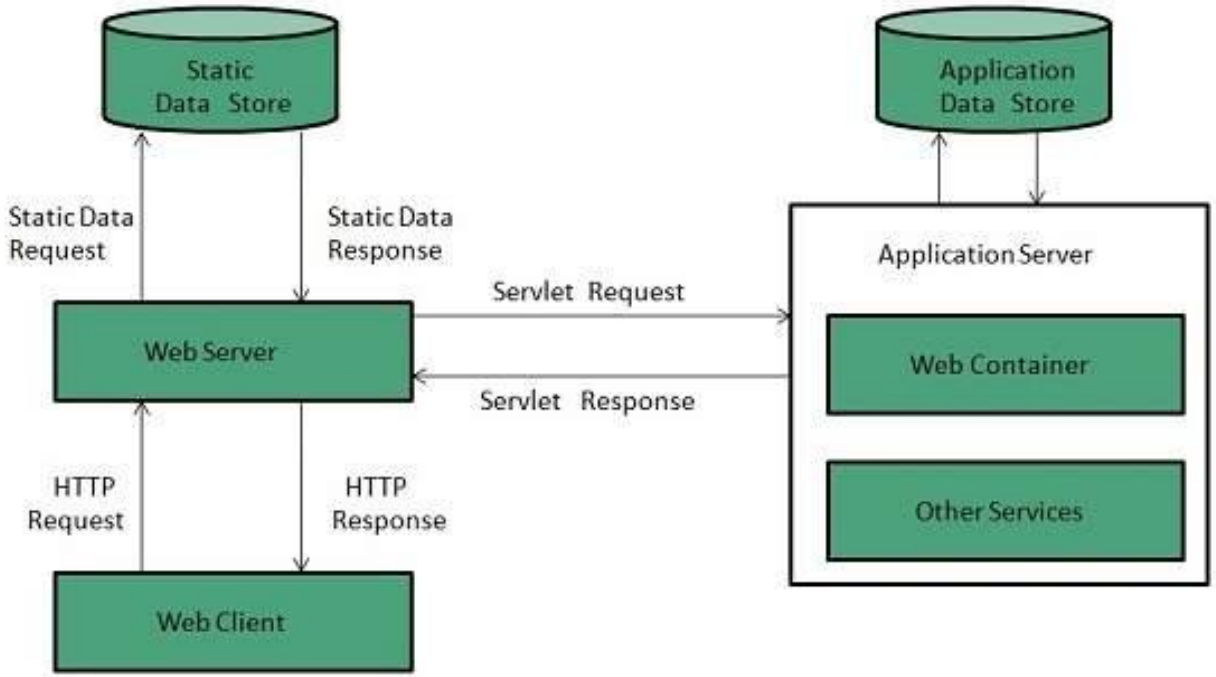
Web Server

- A web server is a computer that hosts web pages, making them accessible online. When a user loads a site, the web server will retrieve the relevant files and send them to the browser so the user can interact with them.
- Web server consists of the hardware and software that is responsible for responding to user requests. Let's break down the main components of a web server:

Software:-The web server software controls how users access hosted files. It consists of http server that is responsible to process and respond to incoming requests.

Hardware:-It stores web server software and another website's data, such as HTML files, images, text documents, and JavaScript files. The web server hardware also connects to the internet, enabling data exchange with other physical devices.


DNS
Server
IP Address
Domain
Web Site
IP Address
A computer using a
Web Browse
Web
Server


Static
Data Store
Application
Data Store
Static Data
Request
Static Data
Response
Application Server
Servlet Request
Web Server
Web Container
Servlet Response
HTTP
Request
HTTP
Response
Other Services
Web Client

Marwadi University

How Does a Web Server Work?

- Let's discover the step-by-step process of what happens whenever a web browser approaches the web server and requests a file. Follow the below steps:

1. A user specifies the URL they want to access in the address bar.
2. The browser fetches the IP address of the domain name. This would take the web browser to the web server.
3. After finding the corresponding web server, the browser sends an HTTP request to retrieve site content.
4. The web server receives and processes the HTTP request through its HTTP server. Once the HTTP server accepts the request, it will search through the database to obtain the relevant data.
5. Finally, the server returns the files to the web browser and delivers them to users.

- When the HTTP server fails to find or process the requested files, it will send an HTTP error status code to the browser.
- The most common error message is a 404 error, which means the requested page is missing.

Types of Web server

- Some of the most common web servers include:

Apache:-Apache Web Server is one of the most popular web servers in the world today. It is compatible with platforms such as Linux, Windows, Mac, and more.

IIS:-It stands for Internet Information Services. It is a web server that is owned by Microsoft. IIS comes with the Windows Server Operating System.

NGINX:-Its primary job is to handle hundreds of concurrent connections. It is suitable for high-traffic websites.

Apache Tomcat:-Apache Tomcat is specializes in Java Servlets and popularly known as a Java container. It supports PHP, ASP.net, Perl, Python, and more.

lighttpd:-lighttpd was developed in 2003. This web server requires low memory and CPU and disk space. Web cameras, internet routers, and other things of a similar nature use lighttpd as their web server.

Web Browser

- A Web browser is a software program that allows us to access web pages.
- A browser is a client program as it runs on a user computer or mobile device and contacts the webserver for the information requested by the user.
- The web server sends the data back to the browser that displays the results on internet supported devices.
- Browser sends requests to web servers using HTTP.
- Browsers translate web pages and websites into human-readable content.

Features of Web Browser:-

Most Web browsers offer common features such as:

Refresh button:- Refresh button allows the website to reload the contents of the web pages. Most of the web browsers store local copies of visited pages to enhance the performance by using a caching mechanism. Sometimes, it stops you from seeing the updated information; in this case, by clicking on the refresh button, you can see the updated information.

Stop button:- It is used to cancel the communication of the web browser with the server and stops loading the page content. For example, if any malicious site enters the browser accidentally, it helps to save from it by clicking on the stop button.

Home button:- It provides users the option to bring up the predefined home page of the website.

Web address bar:- It allows the users to enter a web address in the address bar and visit the website.

Tabbed browsing:-It provides users the option to open multiple websites on a single window. It helps users to read different websites at the same time. For example, when you search for anything on the browser, it provides you a list of search results for your query. You can open all the results by right-clicking on each link, staying on the same page.

Bookmarks:- It allows the users to select particular website to save it for the later retrieval of information, which is predefined by the users.

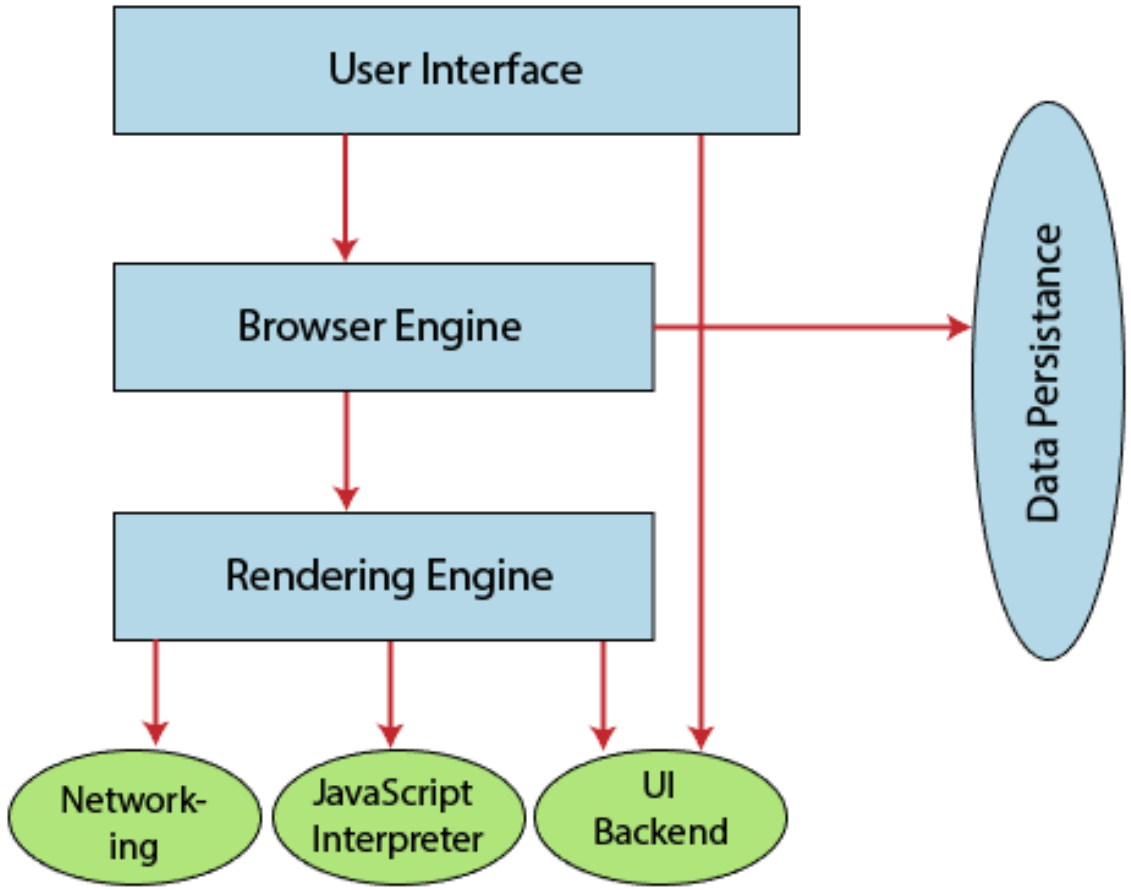
What is the URL (Uniform Resource Locator)?

- A uniform resource locator is the address of a resource on the internet or the World Wide Web. It is also known as a web address or uniform resource identifier (URI).
- For example,https://www.marwadiuniversity.ac.in/ which is the URL or web address for the marwadi university website. A URL represents the address of a resource, including the protocol used to access it.

Component of Web Browser:-

User Interface: The user interface is an area where the user can use several options like address bar, back and forward button, menu, bookmarking, and many other options to interact with the browser.

Browser Engine: It connects the UI (User Interface) and the rendering engine as a bridge. It queries and manipulates the rendering engine based on inputs from several user interfaces.



Rendering Engine: It is responsible for displaying the requested content on the browser screen. It translates the HTML. It generates the layout of the content and displays it on the browser screen.

Networking: It retrieves the URLs by using internet protocols like HTTP or FTP.It is responsible for maintaining all aspects of Internet communication and security. Furthermore, it may be used to cache a retrieved document to reduce network traffic.

JavaScript Interpreter: As the name suggests, JavaScript Interpreter translates and executes the JavaScript code, which is included in a website. The translated results are sent to the rendering engine to display results on the device screen.

UI Backend: It is used to draw basic combo boxes and Windows (widgets). It specifies a generic interface, which is not platform-specific.

Data storage:-The data storage is a persistence layer that is used by the browser to store all sorts of information locally, like cookies. A browser also supports different storage mechanisms such as IndexedDB, WebSQL, localStorage and FileSystem. It is a database stored on the local drive of your computer where the browser is installed. It handles user data like cache, bookmarks, cookies, and preferences.

HTTP

- HTTP stands for HyperText Transfer Protocol. It is used by the World Wide Web.
- It defines how messages are formatted and transmitted, and what actions Web servers and browsers should take in response to various commands.
- When you enter a URL in your browser, this actually sends an HTTP command to the Web server directing it to fetch and transmit the requested Web page.

Features of HTTP

HTTP is connectionless:-The HTTP client, i.e., a browser initiates an HTTP request and after a request is made, the client waits for the response. The server processes the request and sends a response back after which client disconnect the connection. So client and server knows about each other during current request and response only. Further requests are made on new connection like client and server are new to each other.

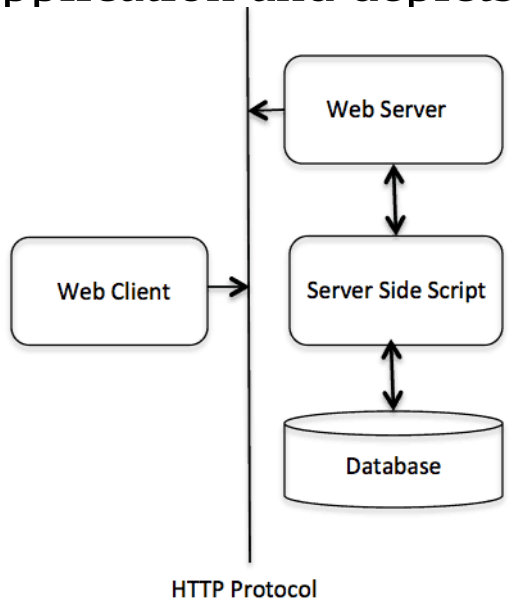
HTTP is media independent:-It means, any type of data(audio,video,image etc) can be sent by HTTP as long as both the client and the server know how to handle the data content.

HTTP is stateless:-The server and client are aware of each other only during a current request. Afterwards, both of them forget about each other. Due to this nature of the protocol, neither the client nor the browser can retain information between different requests across the web pages.

Basic Architecture:-

- The following diagram shows a very basic architecture of a web application and depicts where HTTP sits:
- The HTTP protocol is a request/response protocol

based on the client/server based architecture
where web browsers, robots and search engines etc.
act like HTTP clients, and the Web server acts as a
server.



HTTP Request Method

GET /guide/index.html HTTP/1.1

Host: www.xxxx.com

Accept: image/gif, image/jpeg, */*

Accept-Language: en-us

Accept-Encoding: gzip, deflate

User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)

(blank line)

HTTP Response Message

HTTP/1.1 200 OK
Date: Sun, 18 Oct 2015 08:56:53 GMT
Server: Apache/2.2.14 (Win32)
Last-Modified: Sat, 20 Nov 2015 07:16:26 GMT
Accept-Ranges: bytes
Content-Length: 44
Connection: close
Content-Type: text/html
X-Pad: avoid browser bug
<html><body><h1>Guide</h1> This is guide on HTTP protocol </body></html>

Marwadi University

HTTP Request Methods

- GET
- POST
- PUT
- HEAD
- DELETE

GET Method

- used to request data from a specified resource.
- query string (name/value pairs) is sent in the URL of a GET request
- /test/demo_form.php?name1=kamal&name2=nisha

Notes

- GET requests can be cached
- GET requests remain in the browser history
- GET requests can be bookmarked
- GET requests should never be used when dealing with sensitive data
- GET requests have length restrictions

Post Method

- used to send data to a server to create/update a resource.
- data sent to the server with POST is stored in the request body of the HTTP request:

POST /test/demo_form.php HTTP/1.1

Host: w3schools.com
name1=value1&name2=value2

Notes

- POST requests are never cached
- POST requests do not remain in the browser history
- POST requests cannot be bookmarked
- POST requests have no restrictions on data length based o server configuration.

Put method:

- PUT is used to send data to a server to create/update a resource.

HEAD Method:

- HEAD is almost identical to GET, but without the response body. HEAD requests are useful for checking what a GET request will return before actually making a GET request like before downloading a large file or response body.

DELETE Method:-The DELETE method deletes the specified resource.

OPTIONS Method:-The OPTIONS method describes the communication options for the target resource.

Marwadi University

HTTP Status Code

Status Code	Reason Phrase	Description
200	OK	Standard response for successful request
401	Unauthorized	Resources are password protected
404	Not found	Requested resource is not present currently
509	Bandwidth Limit Exceeded	The server is temporarily unable to respond your request due to the site owner reaching bandwidth limit.
500	Internal server error	Software internal failure

Marwadi University

Difference between http and https

http	https
Transfers data in hypertext (structured text) format	Transfers data in encrypted format
Uses port 80 by default	Uses port 443 by default
Not secure	Secured using SSL technology
Starts with `http://`	Starts with `https://`

Marwadi University

Thank You