The layers below the application layer are there to provide reliable transport, but they do not real work for users.

Even in the application layer there is a need for support protocols, to allow the applications to function.

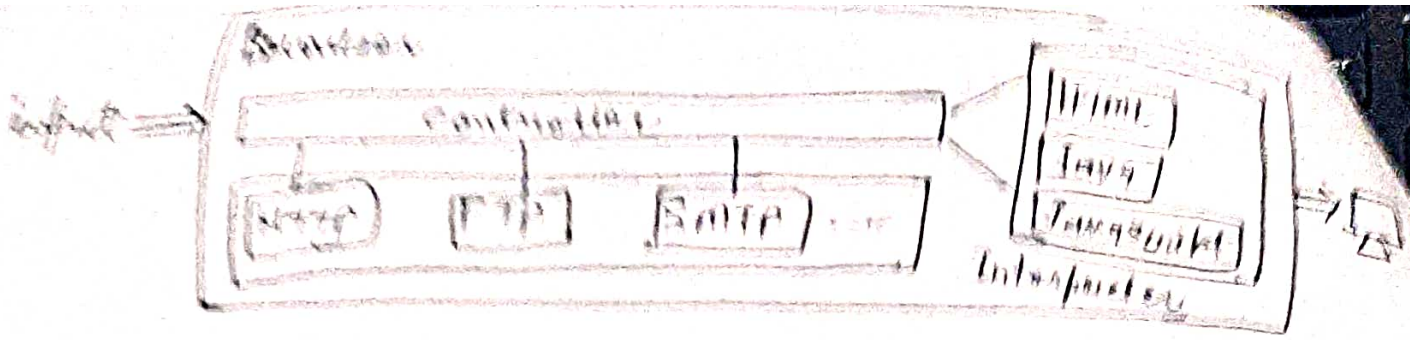
WWW (The World Wide Web):

WWW is a distributed client-server service, in which a client using a browser can access a service using a server. The service provided is distributed over many locations called sites. Each sites holds one or more documents, referred to as web pages. Each web page can contain some links to other links to other web pages in the same or other sites.

Each web page is a file with a name and address.

WWW: All the resources and users on the internet that are using the HTTP.

Web client (Browser): A variety of vendors offers commercial browsers that interpret and display a web page, and all of them use nearly the same architecture. Each browser usually consists of three three parts: a controller client protocol and interpreters.



When an item is selected, the browser follows the hyperlink and fetches the page selected. Pages are named using URL (Uniform Resource Locators).

—: Steps that occur when link is selected:

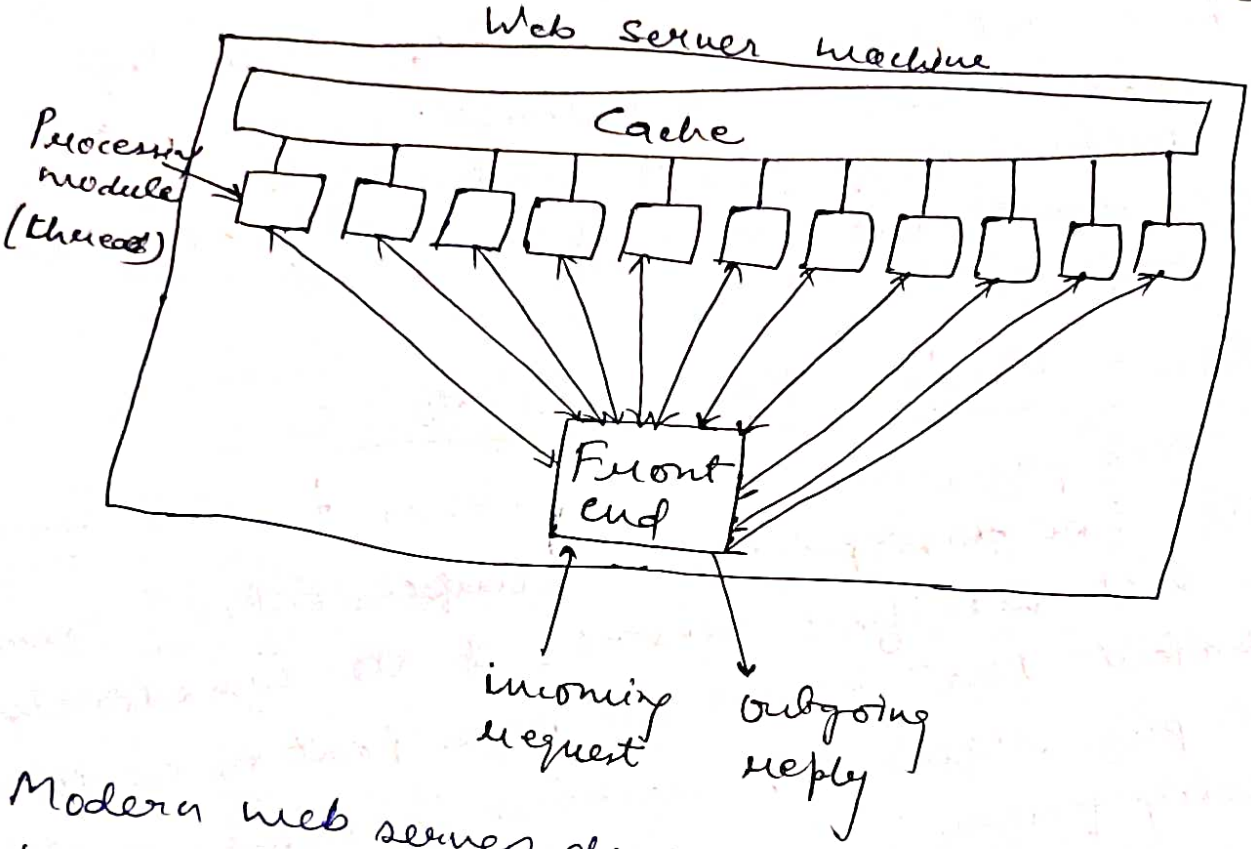
1. The browser determine the URL.
2. The browser asks DNS for the IP address.
3. DNS replies IP add^{n}.
4. The browser makes a TCP connection to port which is required on IP add^{n}.
5. It then sends over a request asking for file.
6. The server sends the file.
7. The TCP connection is released.
8. The browser display all the text.
9. The browser fetches and display all images in this file.

Web Server: The web page is stored at the server. Each time a request arrives, the corresponding document is sent to to the client.

—: Steps that the server perform

1. Accept a TCP connection from a client (a browser).
2. Get the name of the file requested.
3. Get the file (from disk).
4. Return the file to the client.
5. Release the TCP connection.

To improve efficiency, servers normally store requested files in a cache in memory; memory is faster to access than disk. A server can also become more efficient through multithreading or multiprocessing. In this case, a server can answer more than one request at a time.



Modern web servers do more than just accept file names and return files. The front end passes each incoming request to the first available module, which then carries it out using some subset of the following steps, depending on which ones are needed for that particular request.

1. Resolve the name of the web page requested.
2. Authenticate the client.
3. Perform access control on the client.
4. Perform access control on the web page.

5. Check the cache.
6. Fetch the requested page from disk.
7. Determine the MIME type to include in the response.
8. Take care of miscellaneous odds and ends.
9. Return the reply to the client.
10. Make an entry in the server log. Such logs can later be mined for valuable info^n about user behaviour, for eg, the order in which people access the pages. This is used for administrative purpose.

URLs – Uniform Resource locators :

Web pages may contain pointers to other web pages. how these pointers are implemented. When the web was first created, it was immediately apparent that having one page point to another web page required mechanism for naming and locating pages.

1. What is the page called ?
2. Where is the page located ?
3. How can the page be accessed ?

These 3 questions had to be answered before a selected page could be displayed.

A web page, as a file, needs to have a unique identifier to distinguish it from other web pages. To define a page we need three identifier : host, Port, and Path.

Before defining the web page, we need to tell

the browser what client-server application we want to use, which is called the pr protocol.

→ Protocol: The first identifier is the abbreviation for the client-server program that we need in order to access the web page. Most of the time the protocol is HTTP.

→ Host: The host identifier can be the IP address of the server or the unique name given to the server. The name is normally the domain name that uniquely defines the host, such as forouzan.com.

→ Port: The port, a 16-bit integer, is normally predefined for the client-server application. For eg, if the HTTP protocol is used for accessing the web page, the well-known port is 80.

→ Path: The path identifies the location and the name of the file in the underlying OS.

eg: http://www.mhhe.com/compsci/forouzan/

defines the web page related to one of the author

www.mhhe.com → host name

compsci/forouzan → path

Name	Used for	Example
http	Hypertext	http://www.cs.vu.nl/~ast/
ftp	FTP	ftp://ftp.cs.vu.nl/pub/minix/README
telnet	Remote login	telnet://www.w3.org:80

[Some common URLs]

Web Documents :

The documents in the WWW can be grouped into ~~three~~ two broad categories : static, dynamic.

Static Documents : These are fixed content documents that are created and stored in a server. When a client accesses the document, a copy of the document is sent. The user can then use a browser to see the document.

Dynamic documents : Dynamic document is created by a web server whenever a browser requests the document. When a request arrives, the web server runs an application program or a script that creates the dynamic document. The server returns the result of the program or script as a response to the browser that requested the document.

JSP (Java Server Pages), ASP (Active Server Pages) are used to retrieve a dynamic document.