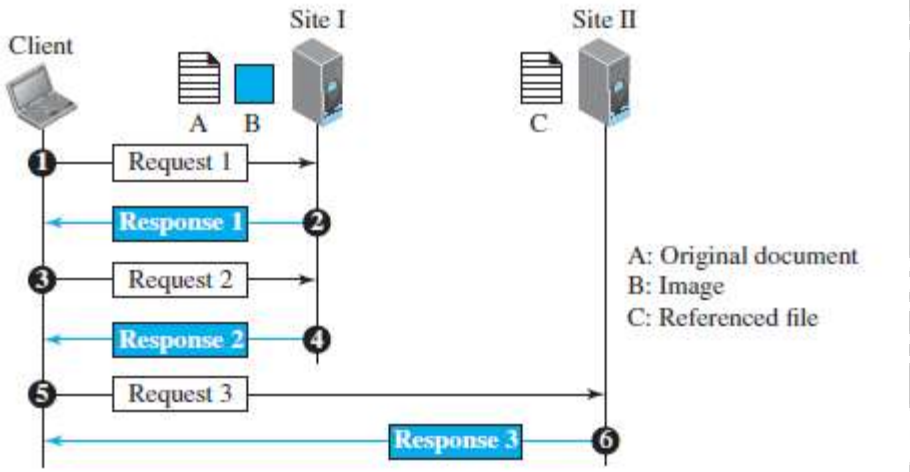
WORLD WIDE WEB

World Wide Web (abbreviated WWW or Web).

Architecture

- The WWW today is a distributed client-server service, in which a client using a browser can access a service using a server. However, the service provided is distributed over many locations called *sites.*
- Each site holds one or more web pages. Each web page, however, can contain some links to other web pages in the same or other sites. A simple web page has no links to other web pages; a composite web page has one or more links to other web pages.
- Each web page is a file with a name and address.



Web Client (Browser)

- A variety of vendors offer commercial **browsers** that interpret and display a webpage, and all of them use nearly the same architecture. Each browser usually consistsof three parts: a controller, client protocols, and interpreters.

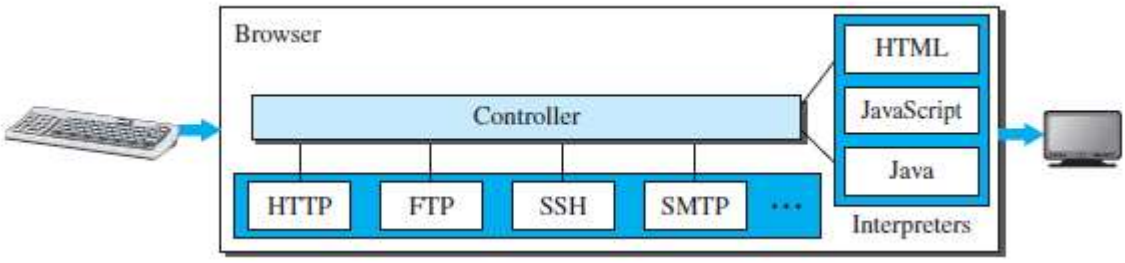


Fig: Browser

- The controller receives input from the keyboard or the mouse and uses the client programs to access the document. After the document has been accessed, the controller uses one of the interpreters to display the document on the screen.

Web Server

- The web page is stored at the server. Each time a request arrives, the corresponding document is sent to the client.
- To improve efficiency, servers normally store requested files in a cache in memory; memory is faster to access than a 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. Some popular web servers include Apache and Microsoft Internet Information Server.

Uniform Resource Locator (URL)

- Define a web page, we need three identifiers: *host, port,* and *path*, *protocol.*

✓ ***Protocol*.** The first identifier is the abbreviation for the client-server program that we need in order to access the web page. Although most of the time the protocol is HTTP (HyperText Transfer Protocol). we can also use other protocols such as FTP (File Transfer Protocol).

✓ ***Host*.** The host identifier can be the IP address of the server or the unique name given to the server.

✓ ***Port*.** The port, a 16-bit integer, is normally predefined for the client-server application.

✓ ***Path*.** The path identifies the location and the name of the file in the underlying operating system. To combine these four pieces together, the **uniform resource locator (URL)** has been designed.

protocol://host/path	**Used most of the time**
protocol://host:port/path	**Used when port number is needed**

Static Documents:

- **Static documents** are fixed-content documents that are created and stored in a server. The client can get a copy of the document only. In other words, the contents of the file are determined when the file is created.
- Static documents are prepared using one of several languages: HyperText Markup Language (HTML), Extensible Markup Language (XML), Extensible Style Language (XSL), and Extensible Hypertext Markup Language (XHTML).

Dynamic Documents:

- A **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.
- *Java Server Pages* (*JSP*),or *Active Server Pages* (*ASP*), Visual Basic language for scripting, or *ColdFusion,* Structured Query Language (SQL).

Active Documents:

- For many applications, we need a program or a script to be run at the client site. These are called ***active documents*.**
- When a browser requests an active document, the server sends a copy of the document ora script. The document is then run at the client (browser) site. One way to create an active document is to use *Java applets,* a program written in Java on the server.
- It is compiled and ready to be run. The document is in byte code (binary) format. Another way is to use *Java Scripts* but download and run the script at the client site.