Introduction to the Internet and Web

Internet

- It is the largest network in the world that connects hundreds of thousands of individual networks all over the world.

- The popular term for the Internet is the "information highway".

- Rather than moving through geographical space, it moves your ideas and information through cyberspace – the space of electronic movement of ideas and information.

Internet

- No one owns it

- It has no formal management organization.

- As it was originally developed by the *Department of Defense*, this lack of centralization made it less vulnerable to wartime or terrorist attacks.

- To access the Internet, an existing network need to pay a small registration fee and agree to certain standards based on the TCP/IP (Transmission Control Protocol/Internet Protocol)

Uses of the Internet

- Send e-mail messages.

- Send (upload) or receive (down load) files between computers.

- Participate in discussion groups, such as mailing lists and newsgroups.

- Surfing the web.

World Wide Web

- Invented by Tim Berners-Lee in 1990

- The invention was while working in CERN (the European Organization for Nuclear Research)

- Tim Berners-Lee invented both the HTML and the HTTP protocol used to request and transmit web pages between web servers and web browsers

World Wide Web Consortium (W3C)

- Founded by Tim Berners-Lee in 1994

- devoted to developing nonproprietary, interoperable technologies for the World Wide Web

- also a standardization organization

- Web technologies standardized by the W3C are called *Recommendations (role, syntax, rules, etc..)*

- *(e.g) HTML, CSS, XHTML, XML*

What is WWW?

The World Wide Web (www) is an open source of information space where documents and other web resources are identified by URLs, interlinked by hypertext links, and can be accessed via the Internet

What is Web?

- Consists of information organized into Web pages containing text and graphic images

- Highlighted keywords and images that lead to related information

- A collection of linked Web pages that has a common theme or focus is called a *Web site*

- The main page that all of the pages on a particular Web site are organized around and link back to is called the site's *Home page*

How to access the Internet?

- To access the Internet, an existing network need to pay a small registration fee and agree to certain standards based on the TCP/IP (Transmission Control Protocol/Internet Protocol) reference model.

- Each organization pays for its own networks and its own telephone bills, but those costs usually exist independent of the internet.

- The regional Internet companies route and forward all traffic, and the cost is still only that of a local telephone call.

Internet Service Provider (ISP)

- A commercial organization with permanent connection to the Internet that sells temporary connections to subscribers

- Examples:

 *Prodigy, America Online, Microsoft network, AT&T Networks.*

How to access the Web?

- Once you have your Internet connection, then you need special software called a *browser* to access the Web

- Web browsers are used to connect you to remote computers, open and transfer files, display text and images.

- *Web browsers* are specialized programs.

- Examples of Web browser: Netscape Navigator (Navigator) and Internet Explorer

Client/Server Structure of the Web

- Web is a collection of files that reside on computers, called *Web servers*, that are located all over the world and are connected to each other through the Internet.

- When you use your Internet connection to become part of the Web, your computer becomes a *Web client* in a worldwide client/server network.

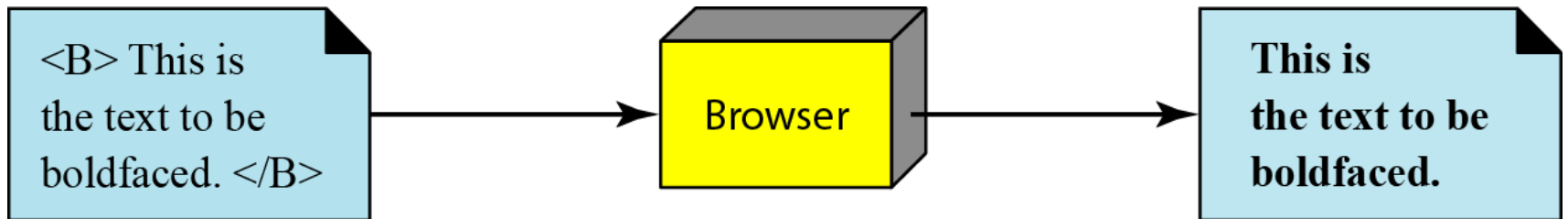- A *Web browser* is the software that you run on your computer to make it work as a web client.
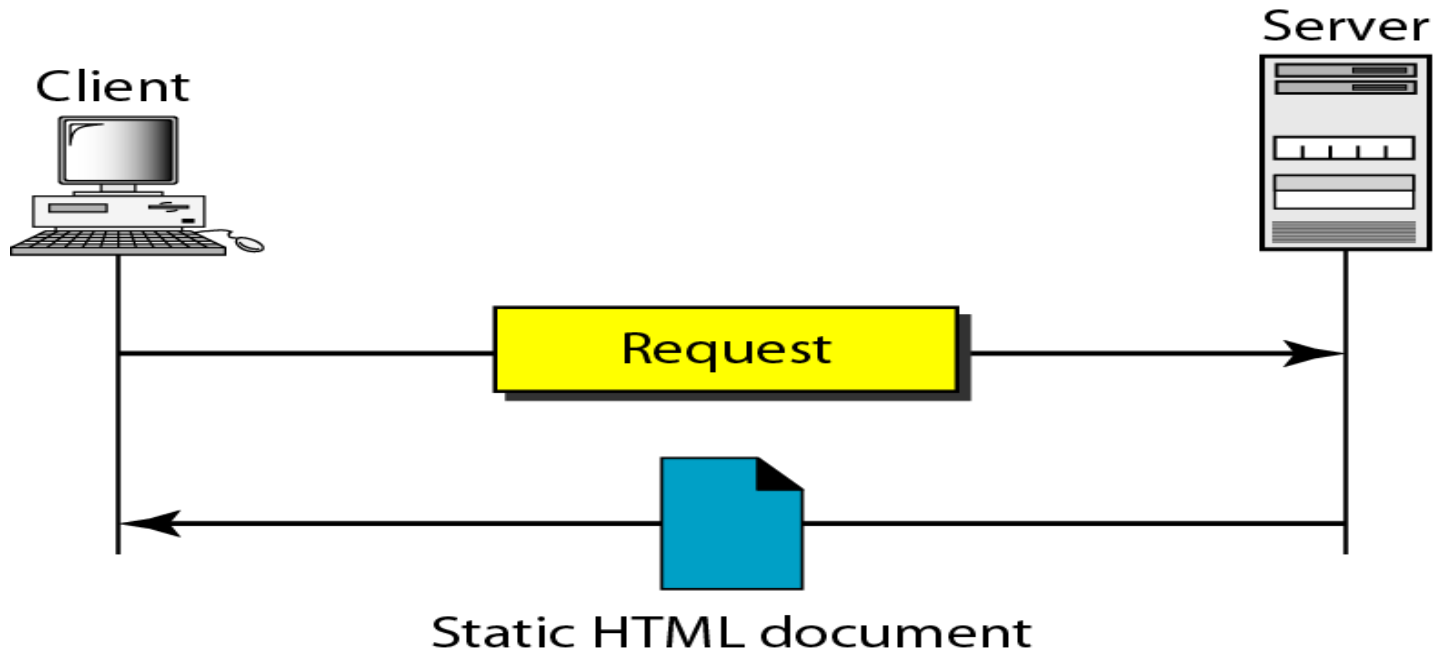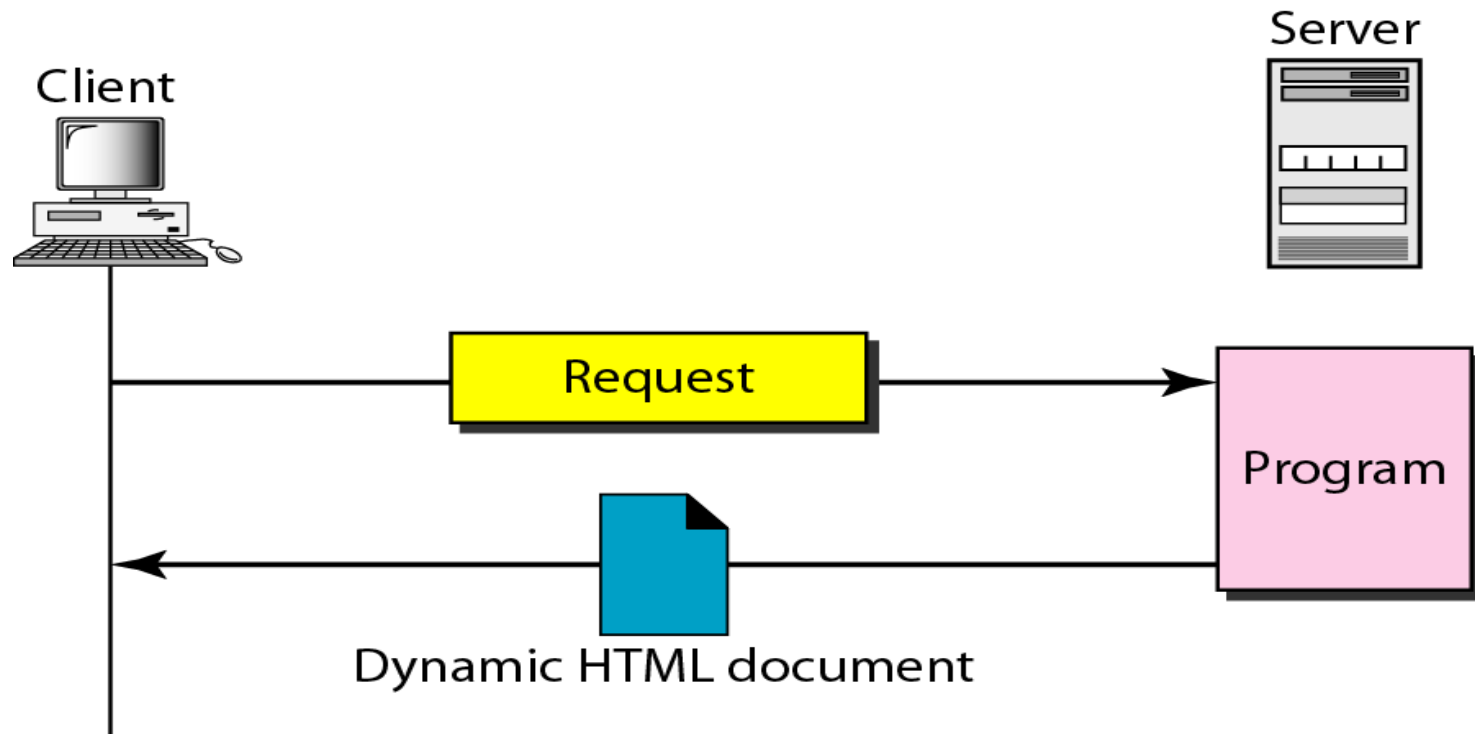
Architecture of WWW
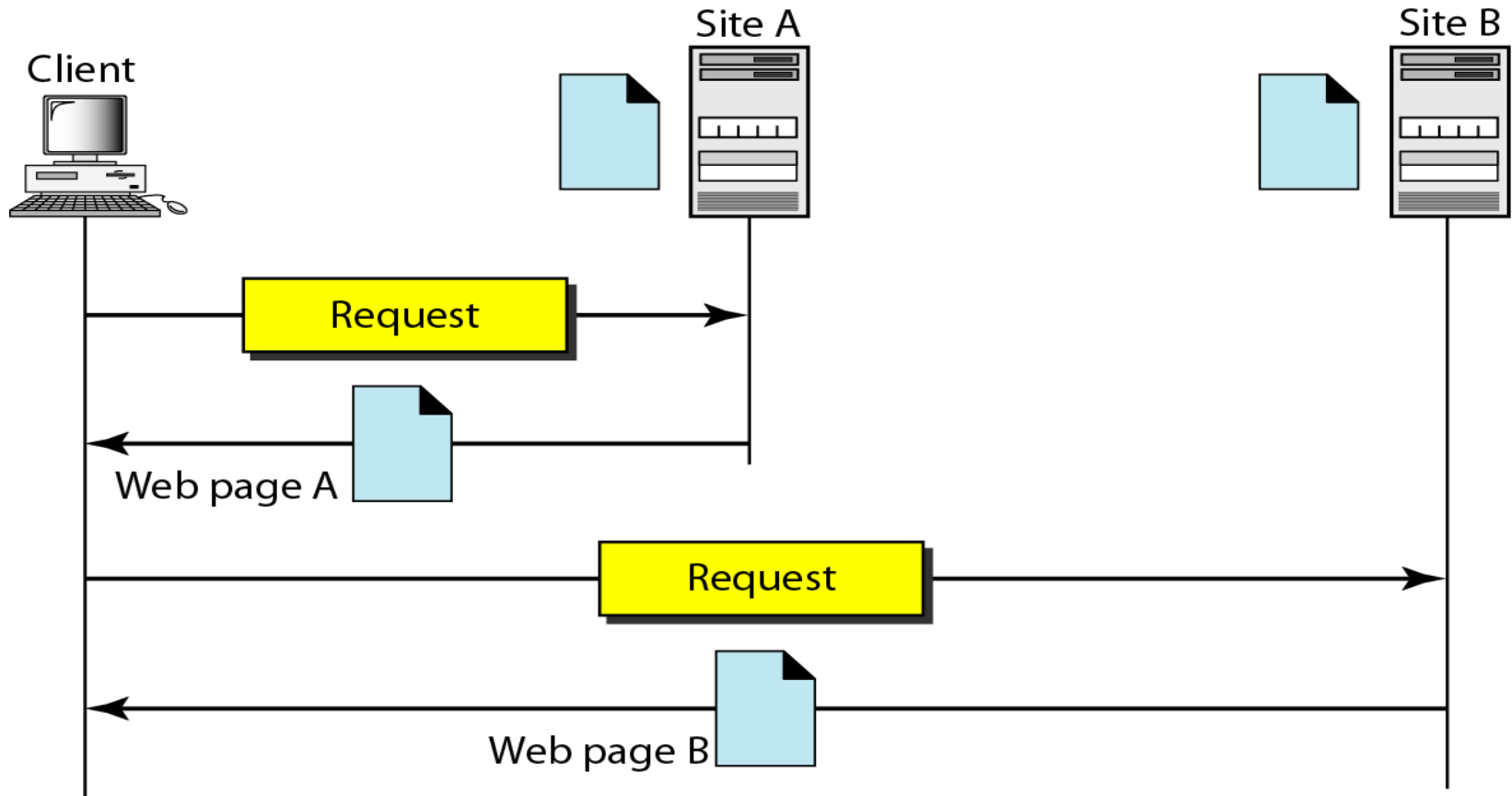
Types of Web Documents

- The documents in the WWW can be grouped into three broad categories: *static, dynamic,* and *active*

- The category is based on the time at which the contents of the document are determined
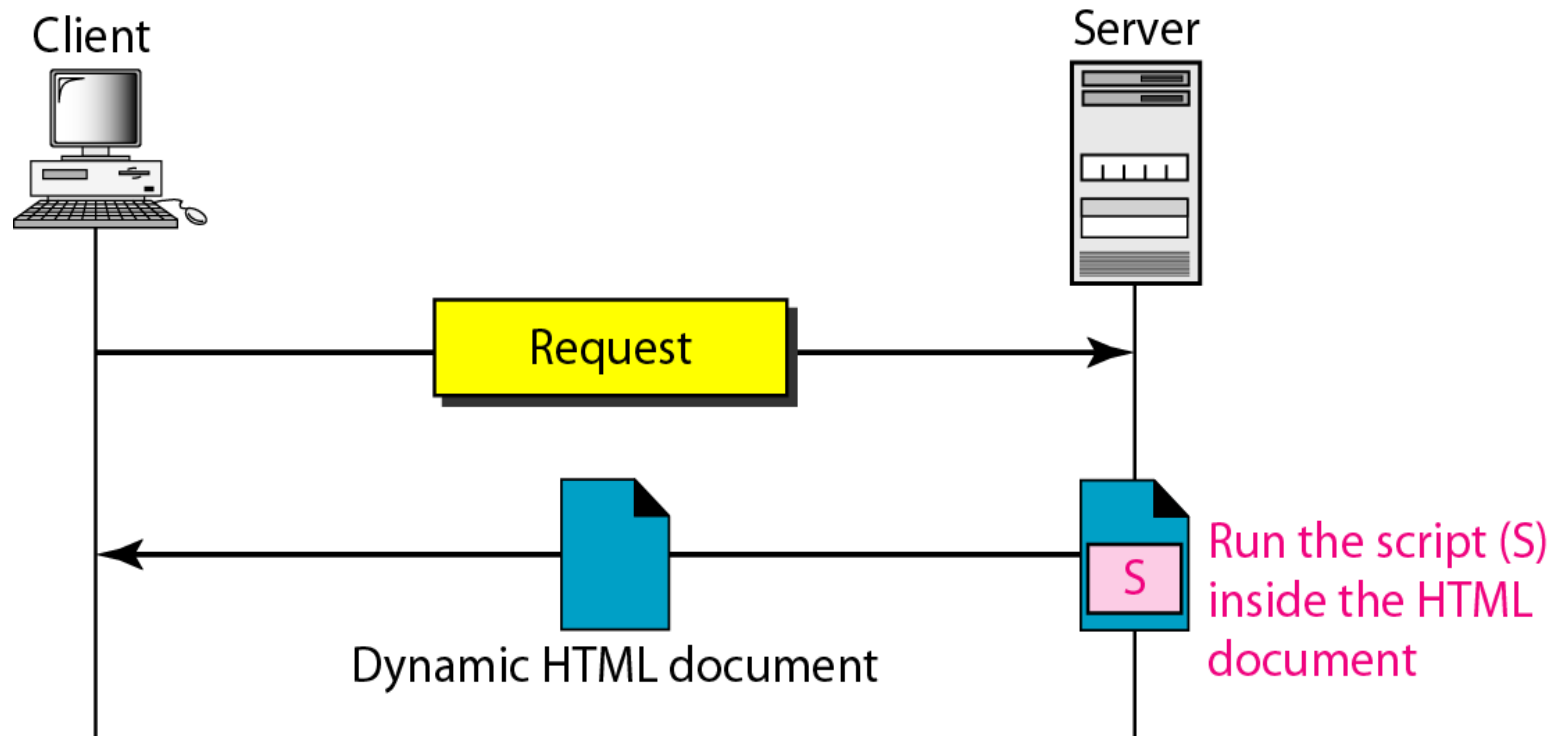
Static document

Client

Server

Request

Static HTML document

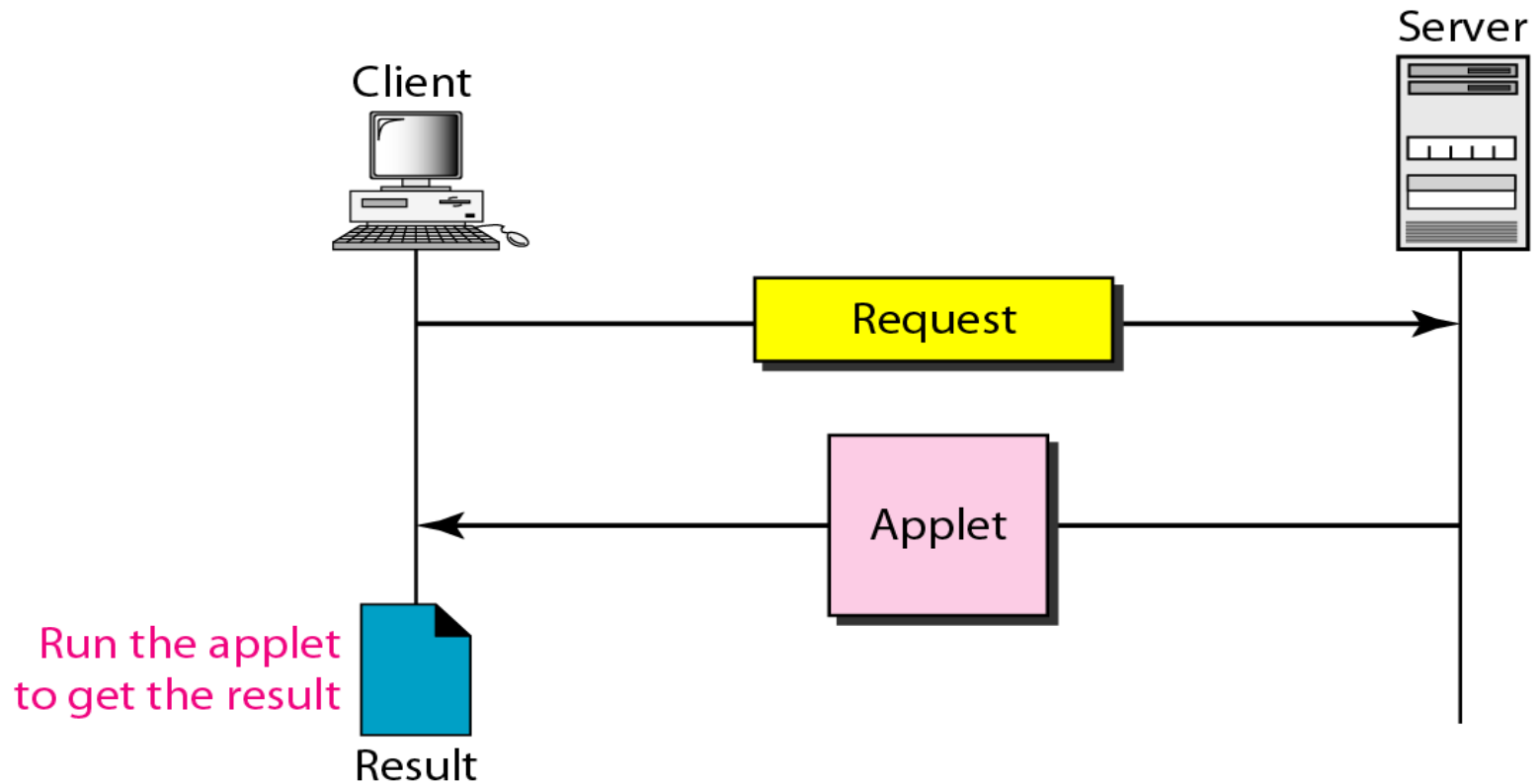| \<B\> This is the text to be boldfaced. \</B\> | → | Browser | → | **This is the text to be boldfaced.** |

Dynamic document using PHP

Dynamic document using Server-side Script
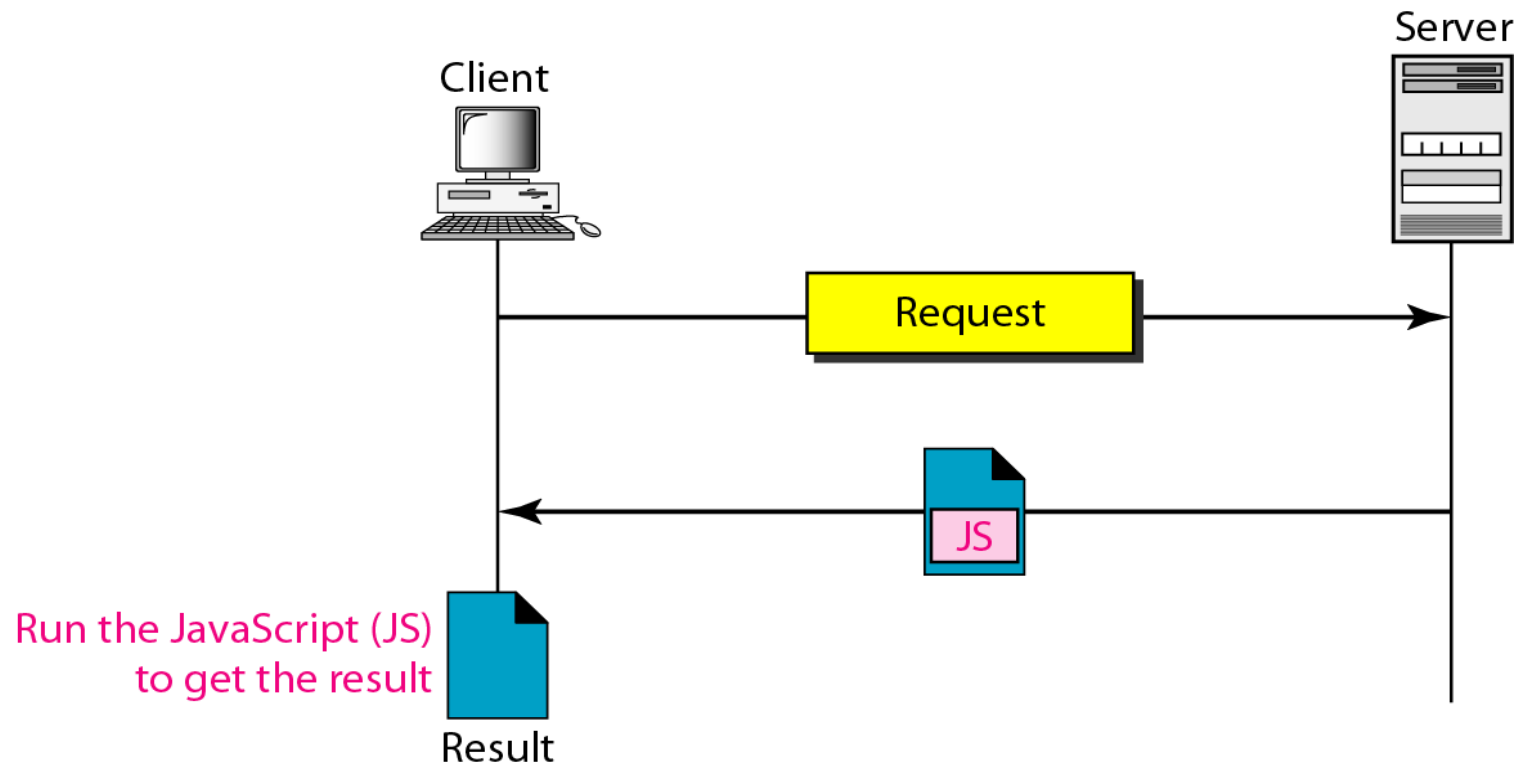

Dynamic documents are sometimes referred to as server-side dynamic documents.

Active document using Java Applet

Dynamic document using Client-side Script

Active documents are sometimes referred to as client-side dynamic documents.

Addresses on the Web:IP Addressing

- Each computer on the internet does have a unique identification number, called an *IP (Internet Protocol) Address*

- The IP addressing system currently in use on the Internet uses a four-part number

- Each part of the address is a number ranging from 0 to 255, and each part is separated by period

- For example, **106.29.242.17**

IP Addressing

- The combination of the four IP address parts provides 4.2 billion possible addresses (256 x 256 x 256 x 256)

- This number seemed adequate until 1998

- Members of various Internet task forces are working to develop an alternate addressing system that will accommodate the projected growth

- However, all of their working solutions require extensive hardware and software changes throughout the Internet
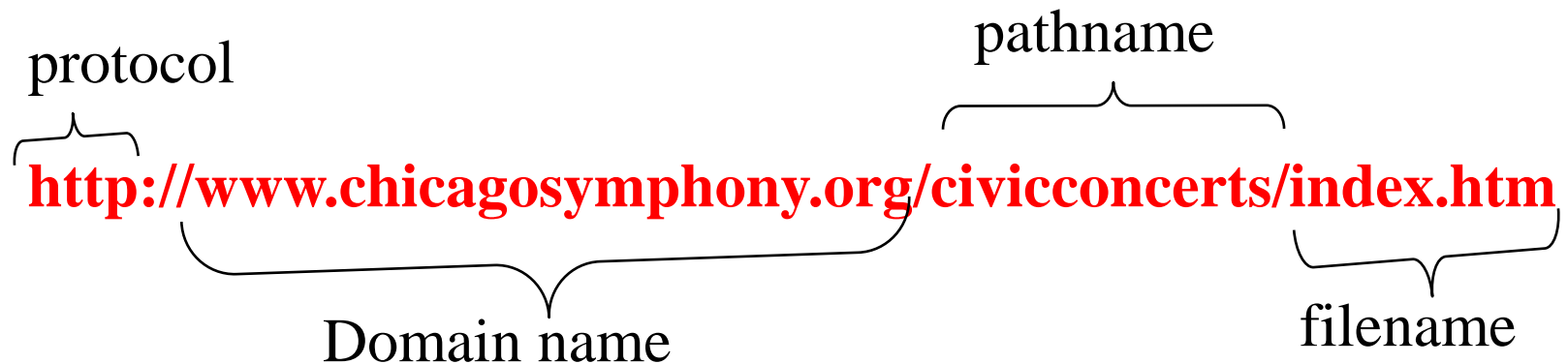
Domain Name Addressing

- Most web browsers do not use the IP address to locate Web sites and individual pages

- They use domain name addressing

- A **domain name** is a unique name associated with a specific IP address by a program that runs on an Internet host computer

- This program, which coordinates the IP addresses and domain names for all computers attached to it, is called **DNS (Domain Name System) software**

- The host computer that runs this software is called a **Domain Name Server**

Uniform Resource Locators

- The IP address and the domain name each identify a particular computer on the Internet.

- However, they do not indicate where a Web page's HTML document resides on that computer.

- To identify a Web pages exact location, Web browsers rely on **Uniform Resource Locator (URL)**

- URL is a *four-part addressing scheme* that tells the Web browser:

  - ✓ *Transfer Protocol*

  - ✓ *Domain name of the computer on which the file resides*

  - ✓ *Pathname of the folder or directory on the computer on which the file resides*

  - ✓ *Name of the file*

Structure of a Uniform Resource Locators

protocol

pathname

http://www.chicagosymphony.org/civicconcerts/index.htm
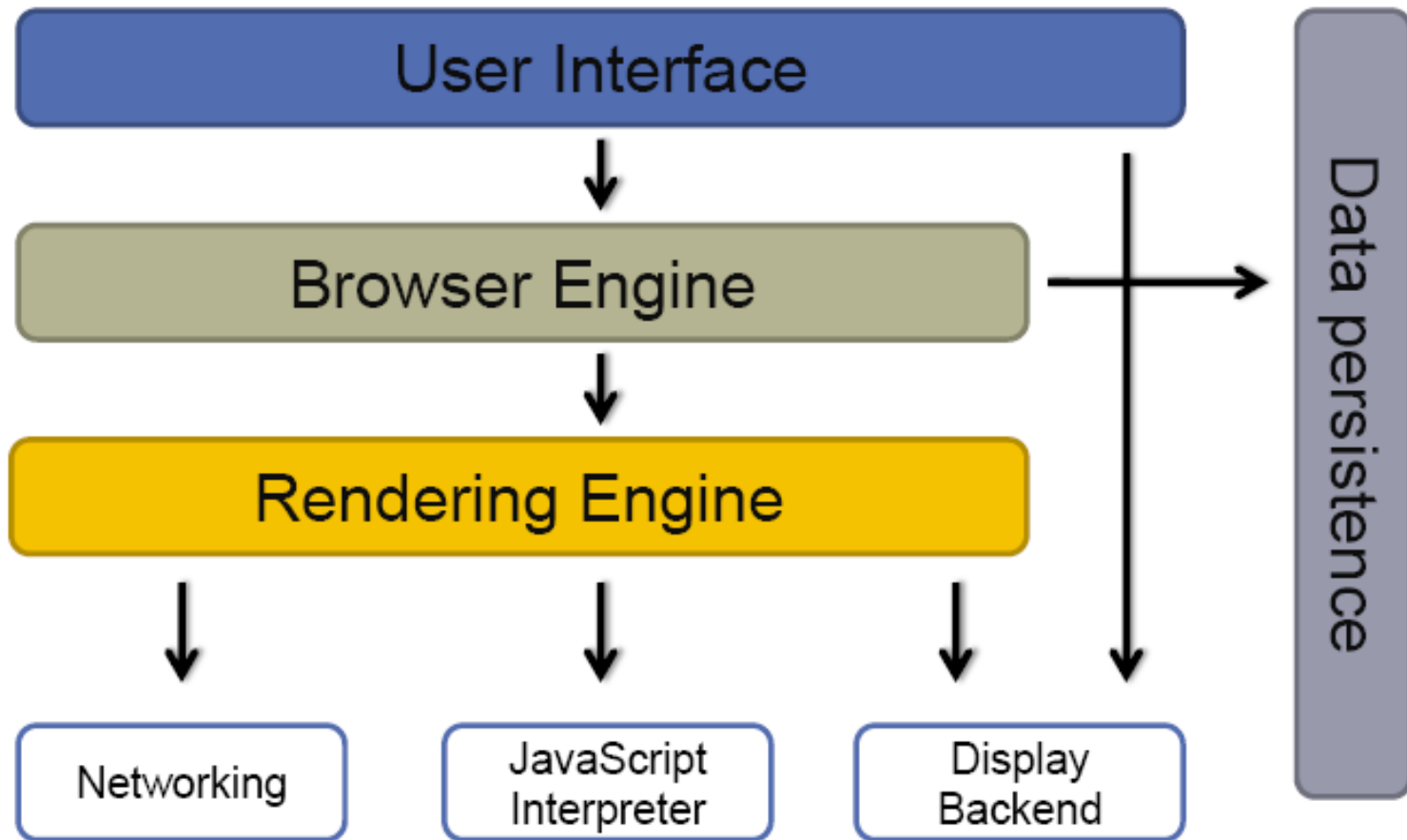
Domain name

filename

http => Hypertext Transfer Protocol

Browser Architecture

What is the Browser?

The browser main functionality is to present the web resource you choose, by requesting it from the server and displaying it on the browser window. The resource format is usually HTML but also PDF, image and more.

Browser Architecture

User Interface

- The space where interaction between users and the browser

- Most of the browsers have common inputs for user interface:

  ✓ Address bar.

  ✓ Next and back buttons.

  ✓ Buttons for home, refresh and stop

  ✓ Bookmark web pages

  ✓ …etc.

Browser Engine

- The bridge between User Interface & Rendering Engine

- Browser Engine provides methods to begin the loading of URL and other high-level browsing actions
  - ✓ Reload, Back, Forward actions
  - ✓ Error messages
  - ✓ Loading progress
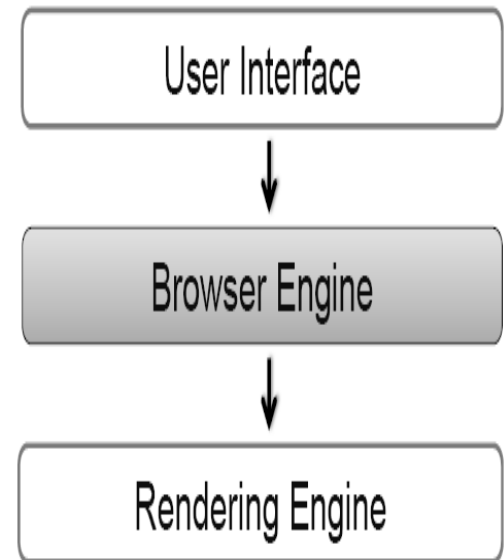
Rendering Engine

- Rendering Engine interprets (render) the HTML, XML, JavaScript and generates the layout that is displayed in the User Interface

Networking

- Access and transfer data on the internet (calls HTTP, HTTPS, FTP)

- The Networking components handles all aspects of internet communication or security

JavaScript Interpreter

- Component parse & executes the JavaScript that is embedded in the website

- Results of the execution a passed to the Rendering Engines for display

Display Backend

- Display common UI components

- Drawing basic widgets like combo boxes, windows

Data Persistence

- Storing the data on the client side.

  ✓Cookies

  ✓Cache

Web Server Architecture

- Web server architecture is the logical layout or design of a web server, based on which a web server is designed, developed and deployed.

- It defines the architectural layout and components of a web server, essential for delivering the required web server-based operations and services.

Web Server Architecture
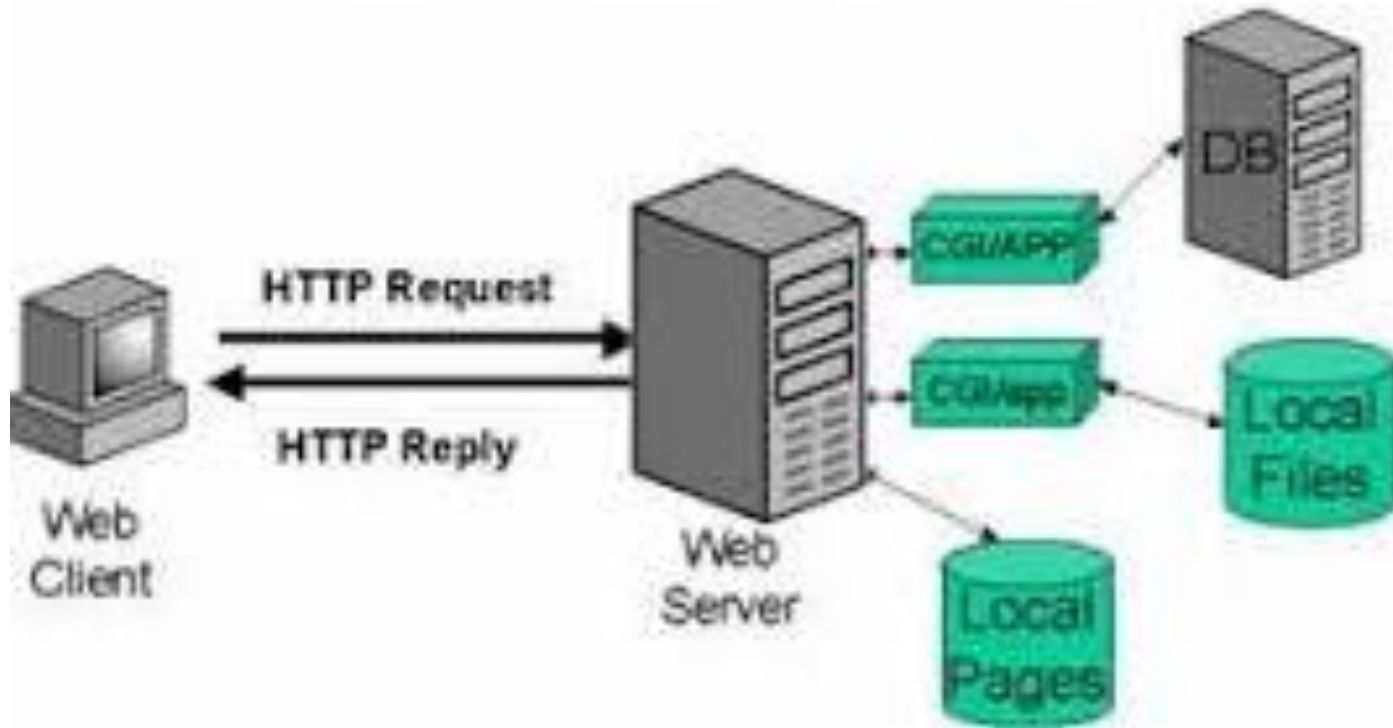
Web Server Architecture

Web server architecture consists of parameters including, but not limited to:

- Physical capacity of the server in terms of computing power, storage and memory

- Performance and quality of service (latency, throughput, low memory utilization)

- Application tiers (type of different applications deployed on the server)

- Platform supported (.Net, LAMP)

- Operating system (Windows, Linux, Solaris)

- Network and/or Internet connectivity (modes of connection and the number of concurrent users it can support)

Exercises

- Comparison of IPv4 & IPv6

- Study about the Brower Architecture

  ✓Firefox

- Study about the Server Architecture

  ✓Apache Web Server