WWW

- The WWW is a distributed client/server service, in which a client using a browser can access a service using a server. However, the service(web pages) provided is distributed over many locations called sites.

- The World Wide Web (WWW), commonly known as the Web, is an information system where documents and other web resources are identified by Uniform Resource Locators (URLs, such as https://www.example.com/), which may be interlinked by hypertext, and are accessible over the Internet.

- A web page is a document commonly written in HTML (Hypertext Markup Language) that is accessible through the Internet or other networks using an Internet browser. A web page is accessed by entering a URL address and may contain text, graphics, and hyperlinks to other web pages and files. The web page is stored at the web Server

- Leading Web servers include Apache (the most widely-installed Web server), Microsoft's Internet Information Server (IIS), Google Web Server (GWS) and IBM's family of Domino servers.

- A website is the group of web pages which are placed in a location on the internet under a domain. For example, a company website can have various web pages such as home, about us, contact us, products, services and other.

- It is accessible through a web address. The website can be designed using static web pages or dynamic web pages. Contents on a website are globally viewed, remains same for the different individuals.

World Wide 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.

The service provided is distributed over many locations called *sites.*

Each site holds one or more web pages. Each web page can contain some links to other web pages in the same or other sites.

- Simple web page has no links to other web pages.

- Composite web page has one or more links to other web pages.

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

Web Server

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

ARCHITECTURE

- Assume we need to retrieve a scientific document that contains one reference to another text file and one reference to a large image. Figure shows the situation.

- The main document and the image are stored in two separate files in the same site (file A and file B); the referenced text file is stored in another site (file C). Since we are dealing with three different files, we need three transactions if we want to see the whole document.

A: Original document
B: Image
C: Referenced file
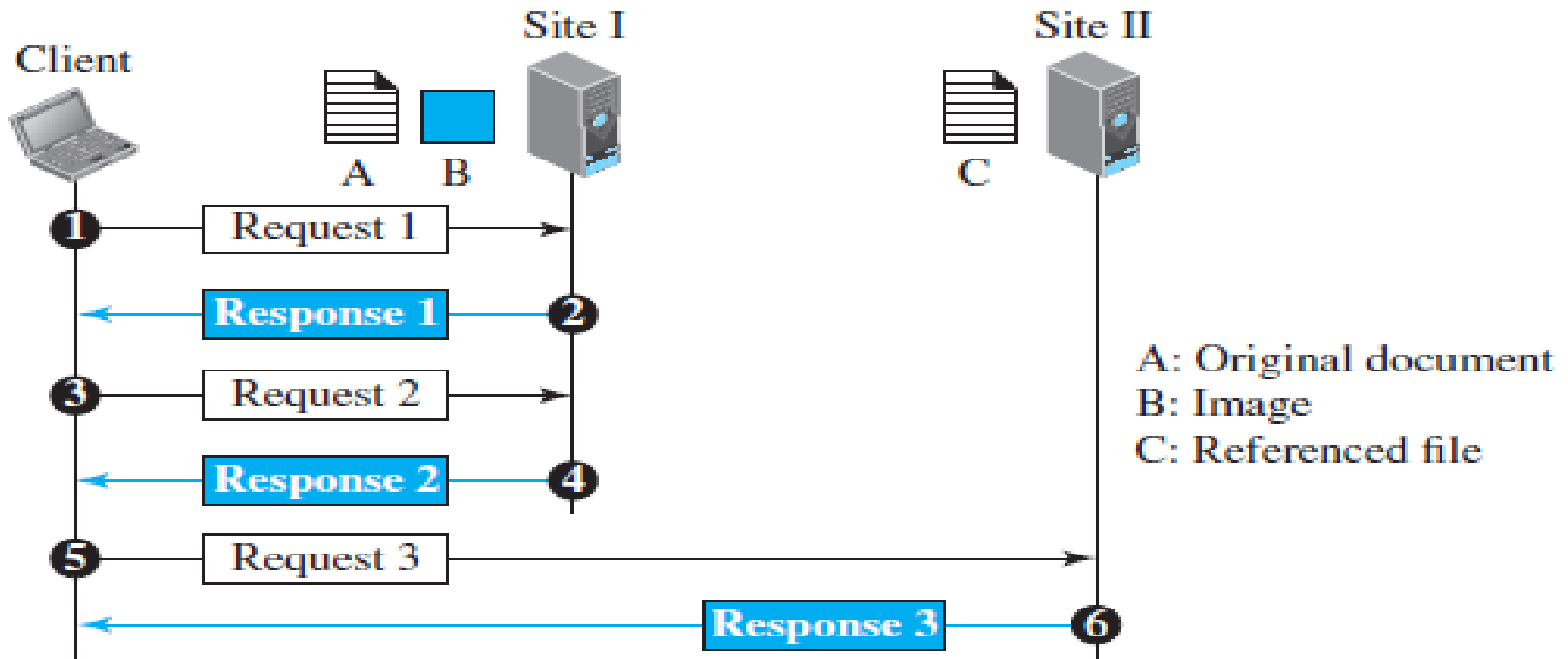
browsers interpret and display a web page.
browsers usually consists of three parts: a controller, client protocols, and interpreters.
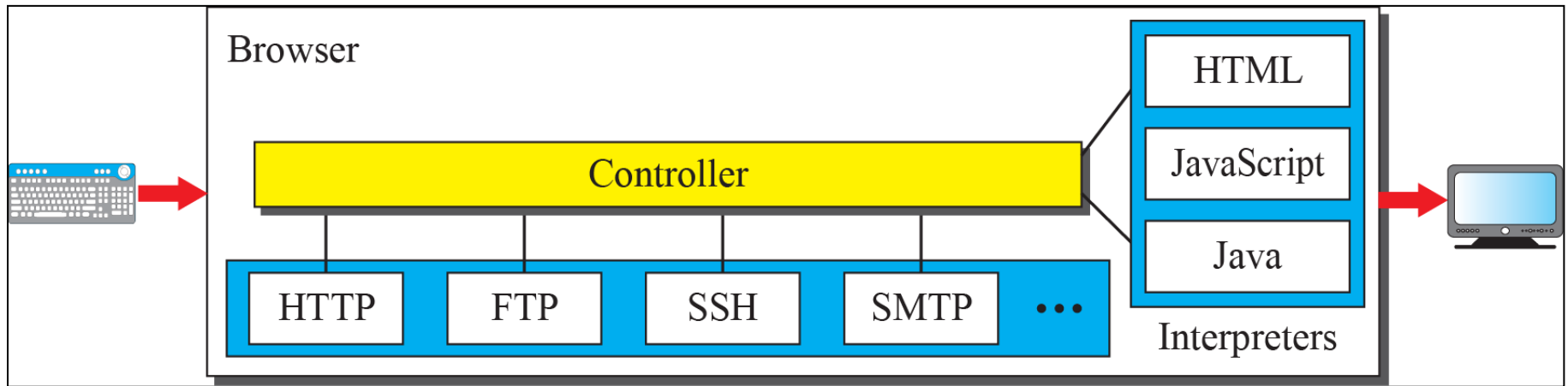
- The controller receives input from the keyboard or the mouse
- The controller uses the client programs to access the document.
- The controller uses one of the interpreters to display the document on the screen.

The client protocol: HTTP or FTP.
The interpreter: HTML, Java, or JavaScript, depending on the type of document.

Uniform Resource Locator (URL)

URL: the address of a World Wide Web page.
The uniform resource locator (URL) combine four identifiers to define the web page to distinguish it from other web pages:

❏ **Protocol:** the client-server program that we need in order to access the web page (HTTP or FTP).

❏ **Host:** IP address of the server or the unique name given to the server such as *https://snscourseware.org/*

❏ **Port:** If the HTTP protocol is used for accessing the web page, the well-known port number is 80. if a different port is used, the number canbe explicitly given.

❏ **Path:** The path identifies the location and the name of the file in the underlying operating system. For example, /snsce/cse/computernetworks is a path that uniquely defines a file named computernetworks.

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

Web Documents

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

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, not when it is used.

- 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)*.

- 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.

- Because a fresh document is created for each request, the contents of a dynamic document may vary from one request to another.

- A very simple example of a dynamic document is the retrieval of the time and date from a server.

Active Documents

- For many applications, we need a program or a script to be run at the client site. These are called *active documents.*

- *For example, suppose we want to run a program that creates* animated graphics on the screen or a program that interacts with the user

- 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 bytecode (binary) format.

- Another way is to use *JavaScripts but download and run the script at the client site.*

Static HTML document

Run the script (S) inside the HTML document

Dynamic HTML document

Applet

Run the applet to get the result

Result