

Internet and Intranet

Lecture by:
Jalauddin Mansur

July 2015

Chapter 4: HTTP and Web Services

Topics :

- HTTP, Web Servers and Web Access
- Universal naming with URLs
- WWW Technology: HTML, DHTML, WML, XML
- Tools: WYSIWYG Authoring Tool
- Helper Applications: CGI, PERL, JAVA, JAVA SCRIPTS, PHP, ASP, .Net Applications
- Introduction to AJAX (Programming)
- Browser as a rendering engine: text, HTML, gif and jpeg

Web Servers and Web Access

- A server is the basic part that differentiates a provider from a user
- To provide information on the web you must either have your own server or rent space on a server
- A Web server is a program that generates and transmits responses to client requests for Web resources.
- Handling a client request consists of several key steps:
 - Parsing the request message
 - Checking that the request is authorized
 - Associating the URL in the request with a file name
 - Constructing the response message
 - Transmitting the response message to the requesting client

- The server can generate the response message in a variety of ways:
 - The server simply retrieves the file associated with the URL and returns the contents to the client.
 - The server may invoke a script that communicates with other servers or a back-end database to construct the response message.
- *Web site* and *Web server* are different:
 - A **Web site** consists of a collection of Web pages associated with a particular hostname.
 - A **Web server** is a program to satisfy client requests for Web resources.

- A Web server may limit which users can access certain resources.
- Access control requires a combination of ***authentication*** and ***authorization***.
 - Authentication identifies the user who originated the request.
 - Authorization determines which users have access to a particular resource

AUTHENTICATION

- Most client-server systems authenticate a user by asking for a user name and password.
- Web server must perform authentication for every request for a resource that has access restrictions.
- The server returns an HTTP response that indicates that the request requires authorization.
- The response also identifies what kind of authentication is required.

AUTHORIZATION

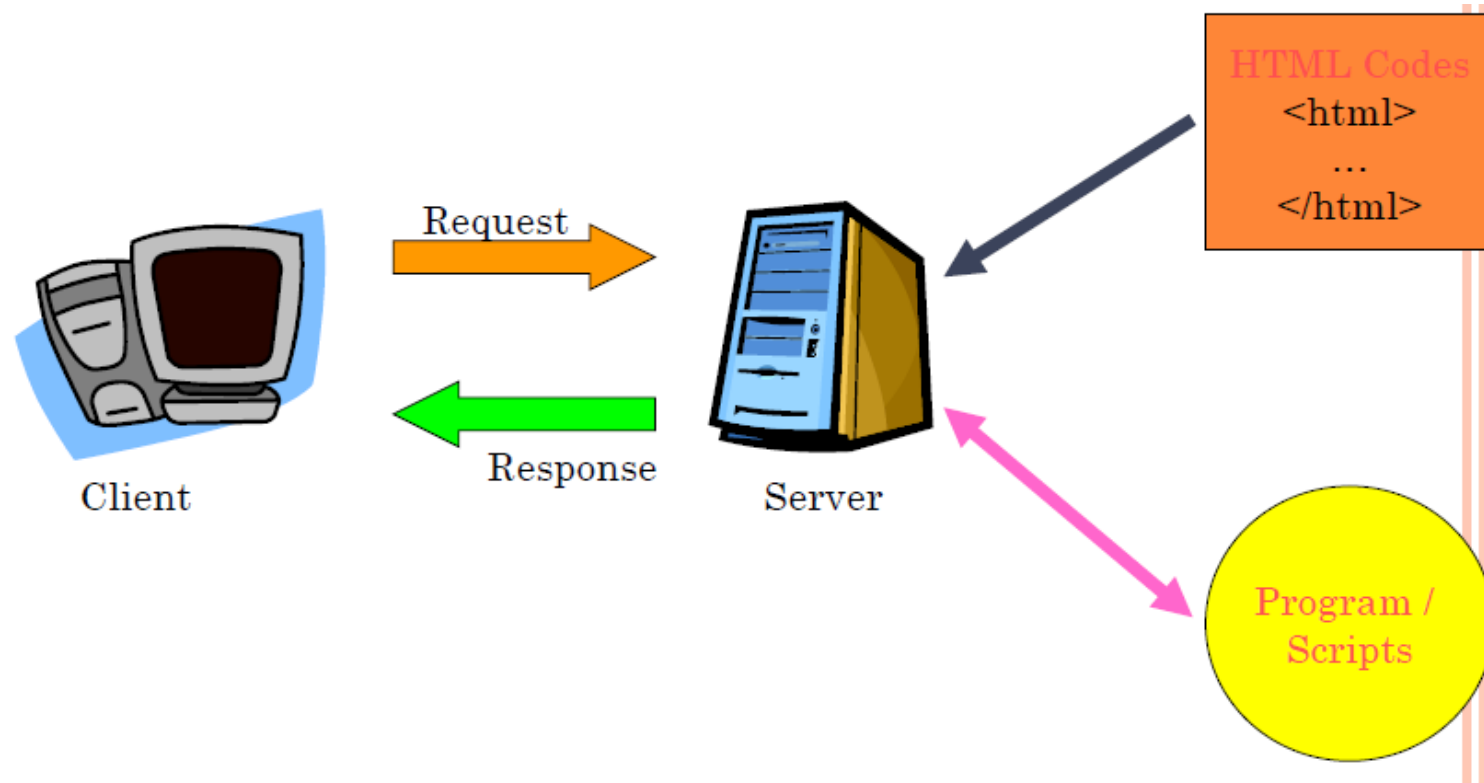
- To control access to Web resources, the server must employ an authorization policy.
- A policy typically expressed in terms of an access control list that enumerates the users who are granted or denied access to the resources.

HOW DOES THE WEB WORK?

- The web information is stored in the Web pages.
 - In HTML format.
- The web pages are stored in the computers called Web servers.
 - In the Web server file system.
- The computer reading the pages is called web clients with specific web browser.
 - Most commonly Internet Explorer or Netscape.
- The web server waits for the request from the web clients over the Internet.
 - Internet Information Server (IIS) or Apache.

HOW DOES THE WEB WORK?

- THE HTTP REQUEST/RESPONSE MODEL Client



Web Server Example

Microsoft Internet Information Services (IIS)

- An enterprise-level Web server that is included with Windows
- FTP Site
 - Used for transferring large files across the Internet
- HTTP Site
 - Used most frequently to request documents from Web servers
- SMTP Virtual Server
 - Sends and receives electronic mail
- Web Site Content Directory
 - Directory containing the documents that clients will view

Apache Web Server

- Currently the most popular Web server
 - Stability
 - Efficiency
 - Portability
 - Open-source

Universal Naming with URLs

URI (Uniform Resource Identifiers)

- An extensible scheme for identifying resources
 - Uniform-common method for naming, locating things
 - Resource-any entity
 - Identifier-character, string that identifies the entity
- URI is simply formatted, case-insensitive string containing name, location etc to identify a resource
 - for example a website, a web service etc.
- URL are a subset of URI that identifies resources by their primary access mechanism (HTTP, FTP, etc)

- URN are a subset of URI that identifies resources that are globally unique and persist even when resource disappears.
- A general syntax of URI used for HTTP is as follows:
 - `URI = "http:" "://" host [":" port] [abs_path ["?" query]]`
- Here if the port is empty or not given,
 - port 80 is assumed for HTTP and
 - an empty `abs_path` is equivalent to an `abs_path` of `"/"`
- For example, the following three URIs are equivalent:
 - `http://abc.com:80/~smith/home.html`
 - `http://ABC.com/%7Esmith/home.html`
 - `http://ABC.com:/%7esmith/home.html`

URL Example

<http://example.edu/fac/lpress/shortbio.htm>

- This request is for a Web (http) server
- The server program is running on a computer with the domain name *example.edu*
- Look in a subdirectory called *fac/lpress*
- If there is a file called *shortbio.htm*, send it back to the client to be displayed; if not, send an error message back to the client

Domain name = the unique name of a computer on the Internet

WWW

- The world wide web is a sophisticated system for universal information capture and delivery
- The world wide web consortium (W3C) definition of the web:
 - *"The World Wide Web is the universe of network-accessible information, an embodiment of human knowledge"*

- Provides information access in ways not previously possible
 - Hyperlinked (Hypertext)
 - Graphical user interface
 - Pictorial and non-text information
 - Information that changes rapidly
 - Immediate access
 - Anyone can author a web site
 - Multi-user access to the same information
 - Searchable information
- The functionality of the WWW is based on 3 standards:
 - URL (Universal Resource Locator)
 - HTML (Hypertext Markup Language)
 - HTTP (Hypertext transfer Protocol)

HTML

- Markup language: Full set of instructions used to comprehensively describe the structural format of a piece of text or other media element
- Standard Generalized Markup Language (SGML):
 - SGML itself is not a markup language
 - the standard on which the most common markup languages are based
- HTML stands for **H**yper **T**ext **M**arkup **L**anguage
- HTML (Hypertext Markup Language) based on SGML
- An HTML file is a text file containing small **markup tags**
- The markup tags tell the Web browser **how to display** the page

- An HTML file must have an **htm** or **html** file extension
- An HTML file can be created using a **simple text editor**
- Basic HTML Tags:

`<html>`

`<head>`

`<title>` Title of page `</title>`

`</head>`

`<body>`

This is my first homepage.

`` This text is bold ``

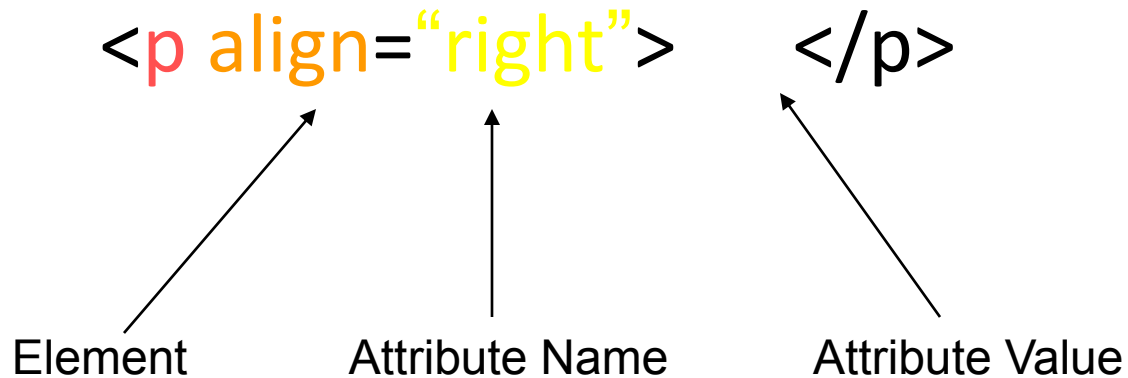
`</body>`

`</html>`

Basic Tags

- Text display:
 - ``, ``, ``
- Structure:
 - `<h1>`, `<h2>`, `<h3>`
 - `<p>`
 - ``, ``
- Attributes:
 - Align, text, bgcolor, etc
- Links:
 - `...`
- Images:
 - `` an empty tag

HTML Element / Tag



Not case-sensitive.

What is XHTML?

- XHTML stands for **EX**tensible **HyperText Markup Language**.
- XHTML is aimed to **replace** HTML.
- XHTML is almost **identical** to HTML 4.01.
- XHTML is a **stricter and cleaner** version of HTML.
- XHTML is HTML defined as an **XML application**.
- XHTML is a W3C Recommendation.

Why XHTML?

- XML was designed to describe data and HTML was designed to display data.
- Therefore by combining HTML and XML, we got a markup language that is useful now and in the future
- XHTML pages can be read by all XML enabled devices AND XHTML gives you the opportunity to write "well-formed" documents now, that work in all browsers and that are backward browser compatible

Advantages of XHTML

- XHTML is easier to maintain.
- XHTML is XSL ready.
- XHTML is easier to teach and to learn.
- XHTML is ready for the future.

DHTML

- DHTML stands for Dynamic HTML.
- DHTML is the art of making HTML pages dynamic.
- DHTML is a combination of technologies used to create dynamic and interactive Web sites.
 - Combination of HTML 4.0, Style Sheets, DOM and JavaScript
- DHTML is not a standard defined by W3C.
- DHTML is a "marketing term" - used by Netscape and Microsoft to describe the new technologies the 4.x generation browsers would support.
- All major Web browsers support DHTML, but differ in terms of what each considers acceptable syntax and scripting code

DHTML Technologies

- HTML 4.0
- Cascading Style Sheets (CSS)
- The Document Object Model (DOM)
- JavaScript

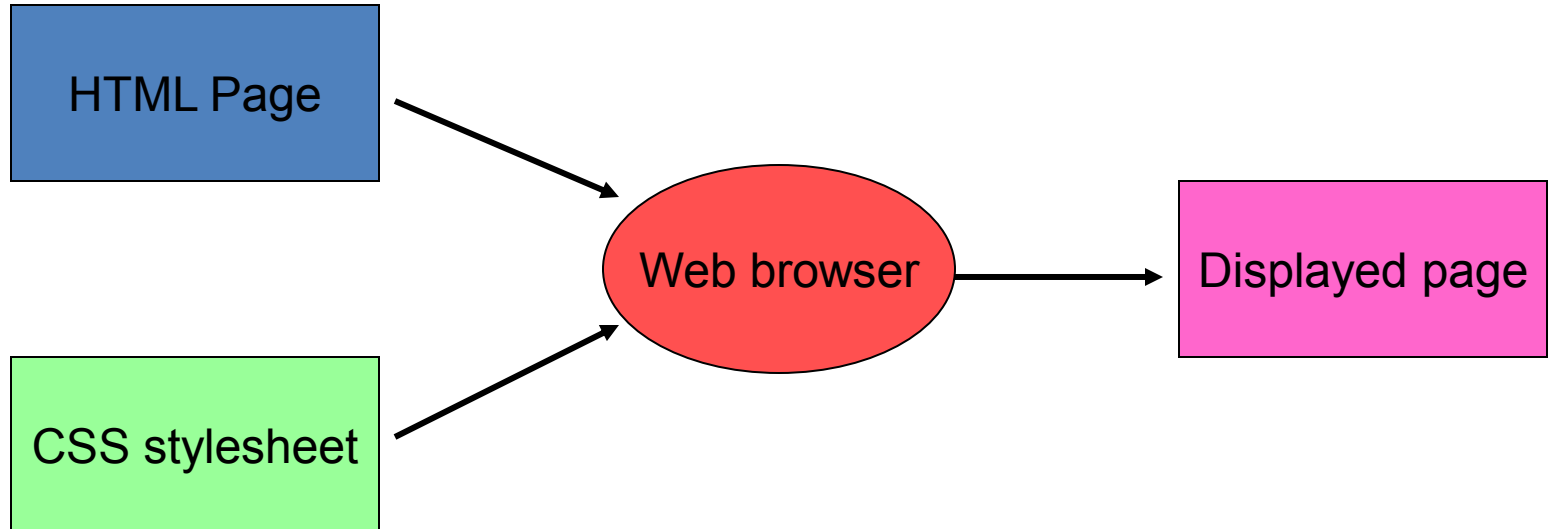
HTML 4.01 Features

- Non Linearity- It is the function of the architecture of web pages to make sense to visitors
- Supports XHTML -that enables you to use the features of XML while creating web pages
- Supports multiple languages- It provides language identifiers that describe the language in which the text on the web is to be displayed
- Supports embedded controls such as ActiveX controls and plug-ins
- Supports scripting-enables you to create dynamic web pages
- Supports enhanced printing options- eg., you can print all web pages that are linked to a specific web pages

Cascading Style Sheets (CSS)

- **CSS** stands for **Cascading Style Sheets**.
- Styles define **how to display** HTML elements.
- CSS is used to impose a particular style on the document.
- Styles are normally stored in **Style Sheets**.
- Styles were added to HTML 4.0
- **External Style Sheets** can save you a lot of work.
- External Style Sheets are stored in **CSS files**.
- Multiple style definitions will **cascade** into one priority:
 - Inline Style (inside HTML element)
 - Internal Style Sheet (inside the <head> tag)
 - External Style Sheet
 - Browser default style

Using Stylesheets to add presentation



DOM

- The document Object Model is an Application programming interface (API) that was developed by W3C
- The HTML Document Object Model (HTML DOM) defines a standard way for accessing and manipulating HTML documents.
- The DOM is platform independent and language neutral
- The DOM is a proposed specification for how objects on a web pages are represented
- The DOM presents an HTML document as a tree-structure (a node tree), with elements, attributes, and text.

- The access, along with methods and properties to add, move, change, or remove HTML elements, is given through the Document Object Model (DOM).
- All browsers have implemented this recommendation, and therefore, incompatibility problems in the DOM have almost disappeared.
- The DOM can be used by JavaScript to read and change HTML, XHTML, and XML documents.

CLIENT-SIDE & SERVER-SIDE TECHNOLOGIES

Client-Side	Server-Side
<p>HTML, XML</p> <p>Cascading Style Sheets (CSS)</p> <p>Scripting languages</p> <ul style="list-style-type: none">- JavaScript, VBScript <p>Java Applets</p> <p>ActiveX controls</p> <p>Plug-ins and Helpers application</p>	<p>CGI/Perl</p> <p>PHP</p> <p>ColdFusion</p> <p>Scripting Languages</p> <ul style="list-style-type: none">- Server-side JavaScript- ASP, JSP, Java Servlets <p>ISAPI/NSAPI programs</p>

JavaScript

- JavaScript was designed to add interactivity to HTML pages.
- JavaScript is a scripting language.
- A JavaScript consists of lines of executable computer code.
- A JavaScript is usually embedded directly into HTML pages.
- JavaScript is an interpreted language (means that scripts execute without preliminary compilation).
- Everyone can use JavaScript without purchasing a license.
- JavaScript is seldom used to write complete “programs”
 - Instead, small bits of JavaScript are used to add functionality to HTML pages
- JavaScript is *reasonably* platform-independent

What can a JavaScript Do?

- JavaScript gives HTML designers a programming tool.
- JavaScript can put dynamic text into an HTML page.
- JavaScript can react to event.
- JavaScript can read and write HTML elements.
- JavaScript can be used to validate data.
- JavaScript can be used to detect the visitor's browser.
- JavaScript can be used to create cookies.

Where to Put the JavaScript

- Head section
- Body section
- External scripts

JavaScript in the head section

```
<html>
<head>
  <script type="text/javascript">
    function message()
    {
      alert("This alert box was called with
the onload event")
    }
  </script>
</head>

<body onload="message()" >

</body>
</html>
```

Scripts in the body section

<html>

<head>

</head>

<body>

<script type="text/javascript">

document.write("Hello World!")

</script>

</body>

</html>

Using External JavaScript

```
<html>
```

```
<head>
```

```
    <script src="abc.js">
```

```
    </script>
```

```
</head>
```

```
<body>
```

```
</body>
```

```
</html>
```

WML

- Wireless Markup Language
 - Formerly called HDML (Handheld Devices Markup Languages)
 - Allows the text portions of web pages to be displayed on cell phones or PDAs via wireless media.
 - It is part of the Wireless Application Protocol (WAP).
 - WML is read and interpreted by a micro browser that is built in to a wireless handheld device

WML Features

- Small display
- Limited input capacity
- Narrowband network connection
- Limited memory
- Limited CPU power
- It provides WML users with games, email services and instant messaging

XML

- XML stands for **EX**tensible **M**arkup **L**anguage.
- XML is a **markup language** much like HTML.
- XML was designed to **describe data**.
- XML tags are not predefined. You must **define your own tags**.
- XML uses a **Document Type Definition (DTD)** or an **XML Schema** to describe the data.
- XML with a DTD or XML Schema is designed to be **self-descriptive**.
- XML is a W3C Recommendation.

XML Features

- Written in plain text and is readable
 - Can view and edit XML document with simple text editor
- Cross platform software and hardware independent tool that is used for transmitting information
- Used primarily for web pages
- XML is meant to emphasize intelligent and logical formatting within technical documents
- XML provides a way for databases from different vendors to exchange information across the Internet

Differences between XML and HTML

- XML was designed to carry data.
- XML is not a replacement for HTML.
- XML and HTML were designed with different goals:
 - XML was designed to describe data and to focus on what data is.
 - HTML was designed to display data and to focus on how data looks.
- HTML is about displaying information, while XML is about describing information.

An example of an XML document

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<bookstore>
```

```
  <book category="COOKING">
```

```
    <title lang="en">Everyday Italian</title>
```

```
    <author>Giada De Laurentiis</author>
```

```
    <year>2005</year>
```

```
    <price>30.00</price>
```

```
  </book>
```

```
  <book category="CHILDREN">
```

```
    <title lang="en">Harry Potter</title>
```

```
    <author>J K. Rowling</author>
```

```
    <year>2005</year>
```

```
    <price>29.99</price>
```

```
  </book>
```

```
  .....
```

```
</bookstore>
```

Introduction to AJAX

- AJAX is an acronym for **A**synchronous **J**avaScript **A**nd **X**ML.
- AJAX is not a programming language but simply a development technique for creating interactive web applications.
- The technology uses JavaScript to send and receive data between a web browser and a web server.
- The AJAX technique makes web pages more responsive by exchanging data with a server behind the scenes, instead of reloading an entire web page each time a user makes a change.
- With AJAX, web applications can be faster, more interactive, and more user friendly.
- AJAX uses an XMLHttpRequest object to send data to a web server, and XML is commonly used as the format for receiving server data.

Why use Ajax?

- Users demand
 - Very user-visible effect
 - Rich UI experience in a Web page
- Portable across browsers
- Geeky reasons:
 - Multithreaded data retrieval from Web servers
 - User can continue to interact with a page while the request is in progress
 - Progress indicators
 - Appearance of speed
 - Avoids need for `setTimeout()`
 - Less bandwidth required; less server load
 - Reload partial page, not entire page
 - Load data only, not even partial page

Why use Ajax?

- Client/Server Apps:
 - Dynamic data
 - Static forms, controls, code, etc.
 - Efficient, but not flexible
- Traditional Web Apps:
 - Dynamic data
 - Dynamic forms, controls, code, etc.
 - Flexible, but inefficient, and noticeably slow
- Ajax Apps:
 - Dynamic data
 - Static or dynamic forms, controls, code, etc.
 - Best of both worlds

Pros and Cons of AJAX

Pros

- Allows web applications to interact with data on the server
- Avoid clunky GET/POST send/receive interfaces – web apps
- Look more and more like real applications
- Some applications can only be realized this way
- Eg: Google Suggest offers interactive access to one of the largest data collections in the world
- For office style applications, user's data is stored on a reliable server, accessible from any web browser

Cons

- Tough to make compatible across all browsers
- Should have a low-latency connection to the server
- Can be server intensive
- Eg: Google Suggest generates a search for every keystroke entered

AJAX Application Example

- Online test
 - Many multiple choice questions
 - All questions are displayed on one page
 - After the user answers one question, the correct answer and explanation about why the user answer is wrong is shown on the page
- Pure sever-side solution using conventional web application
 - For each question answer submission, the whole page with most of repeated data sent to the browser
- Pure client-side solution using conventional JavaScript
 - The user can read JavaScript source code to view what is correct answer
 - Large amount of explanation data will be carried by the JavaScript code
- AJAX solution
 - After the user answers a question, use XMLHttpRequest to ask the server to send the correct answer and explanation.
 - Display the correct answer and explanation received from the server

CGI

- CGI = Common Gateway Interface
- Provides a standardized way for web browsers to:
 - Call programs on a server.
 - Pass data to programs on a server.
 - Receive responses from programs on a server.
- CGI is the interface between server programs and other software.
- CGI is not a Perl specific concept.
- Any language can produce CGI programs.
- Why Perl?
 - Perl provides a nice interface for creating CGI scripts.

How Does CGI Work?

- Phase 1: Create.
- Phase 2: Request/Execute.
- Phase 3: Respond/Display.

Phase 1

- A CGI script is created.
 - e.g. a Perl program to do your taxes.
- The script is placed on a server.
 - Made executable.
 - Given appropriate permissions.
- A webpage is created and linked to the CGI script.
 - Webpage is the script's interface to the world.

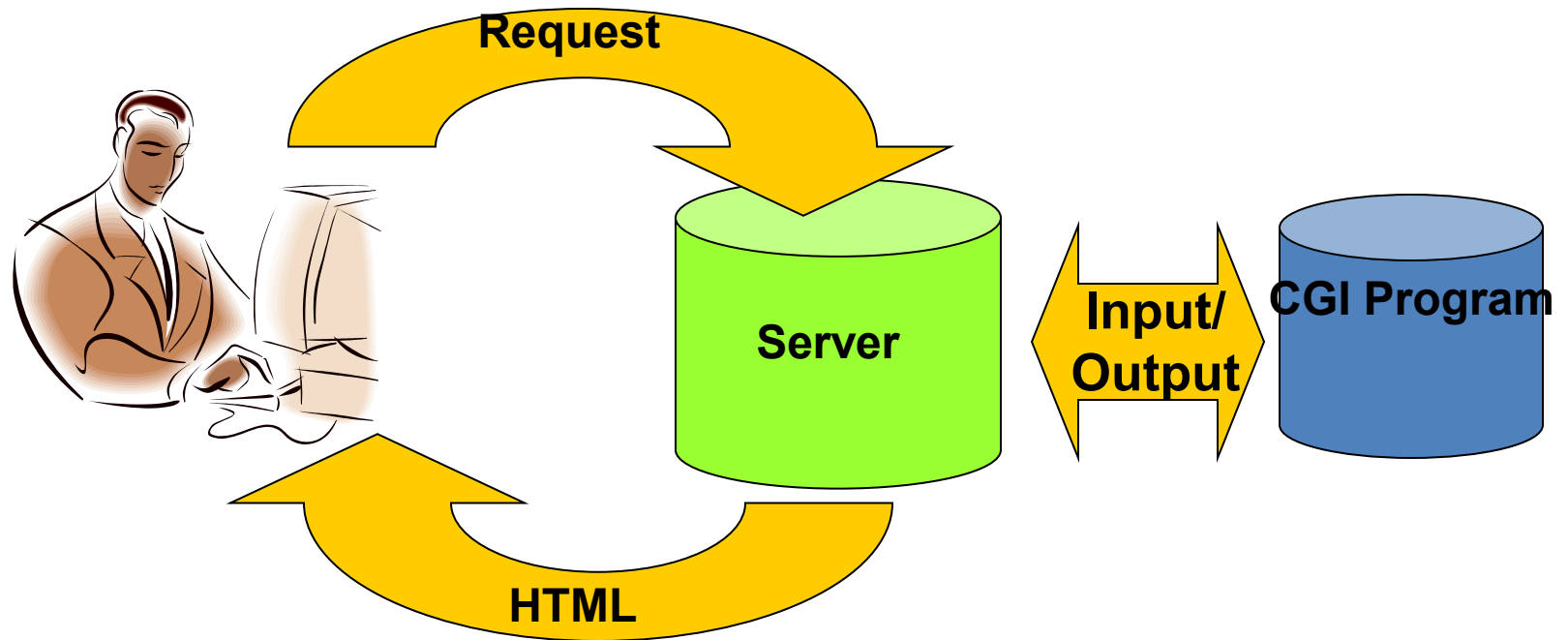
Phase 2

- A person visits the webpage and submits a request to run the script.
- Browser contacts the server with the CGI script:
 - Asks to run the script.
 - Passes input parameters to the script.
- Server runs the script on the input parameters.

Phase 3

- Script produces output in the form of HTML.
- Server takes the output and returns it to the web browser.
- Browser displays the output as an HTML page.
 - Page may reference other CGI scripts.

An Overview of the Process



Perl

- Perl is a general-purpose programming language,
- Can be used for practically any programming task that any other high-level language can be used for.
- However, Perl is usually thought of as a “glue” language, so called because it binds things together (such as tying databases to Web pages, converting files from one format to another, and so on).
- Perl is very flexible and is currently available on over two dozen operating system platforms
- The name Perl comes from “Practical Extraction and Report Language”.
- Perl has many features borrowed from other programming languages.

Where to get Perl

- Perl is available free of charge from many public sites.
- For Linux or UNIX systems, visit perl.com for the latest releases
- For Windows systems, you can compile the Perl source code yourself (a hassle) or download a preconfigured Windows release at activestate.com
- For Macintosh, visit macperl.com for MacPerl

Java

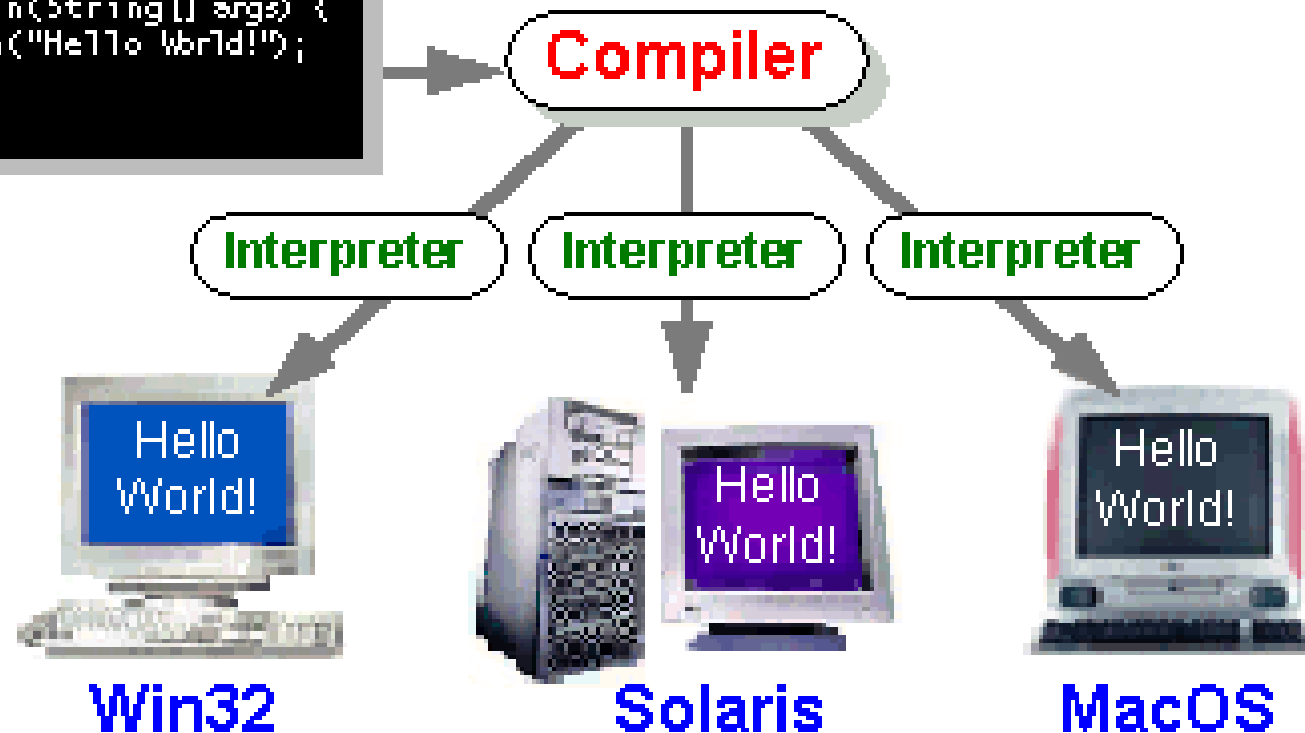
- Java is a general-purpose computer programming language that is concurrent, class-based, object-oriented
- Specifically designed to have as few implementation dependencies as possible.
- Intended to let application developers "write once, run anywhere" (WORA), meaning that compiled Java code can run on all platforms that support Java without the need for recompilation.
- Java applications are typically compiled to byte-code that can run on any Java virtual machine (JVM) regardless of computer architecture.

Java – compiled and interpreted

Java Program

```
class HelloWorldApp {  
    public static void main(String[] args) {  
        System.out.println("Hello World!");  
    }  
}
```

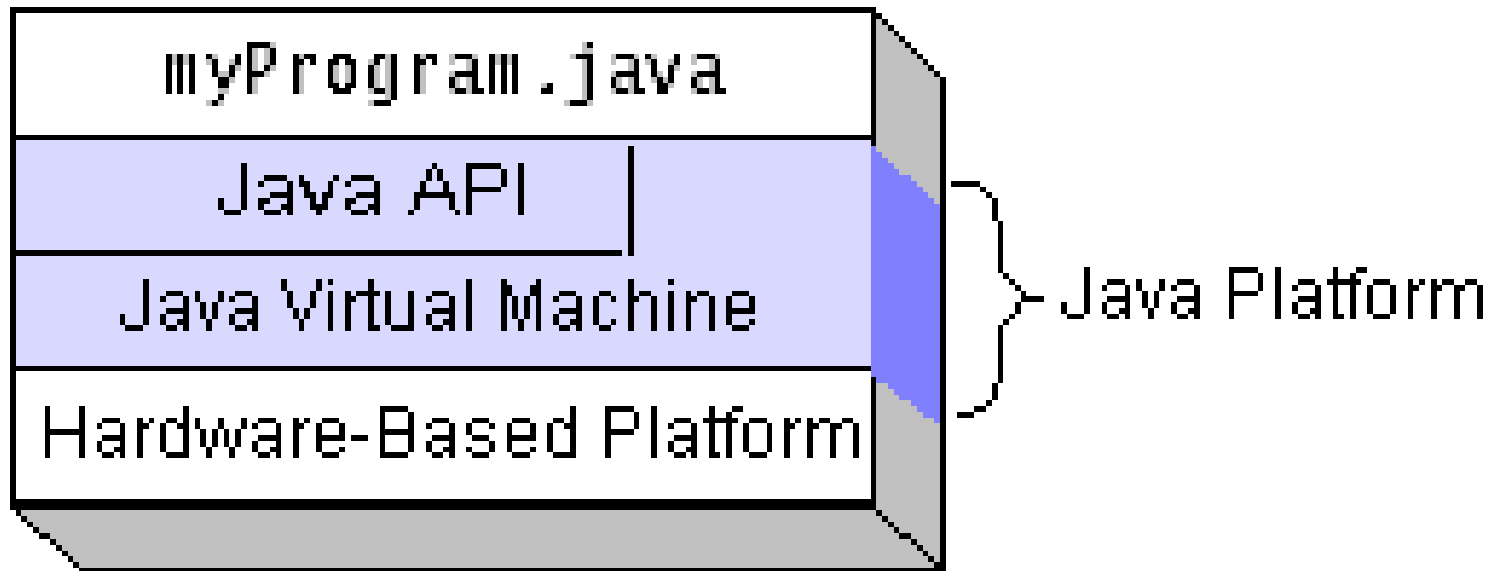
HelloWorldApp.java



Java platform

The Java platform has two components:

- The *Java Virtual Machine* (**Java VM**)
- The *Java Application Programming Interface* (**Java API**)



Java Applet

- It is used to solve the problem that HTML is not a programming language.
- Instead of running a program on the web server, a special kind of Java program (applet) is downloaded to the browser.
- JavaScript is less powerful than Java.
- JavaScript code is physically part of an HTML document, but applets are stored separately from the HTML files.

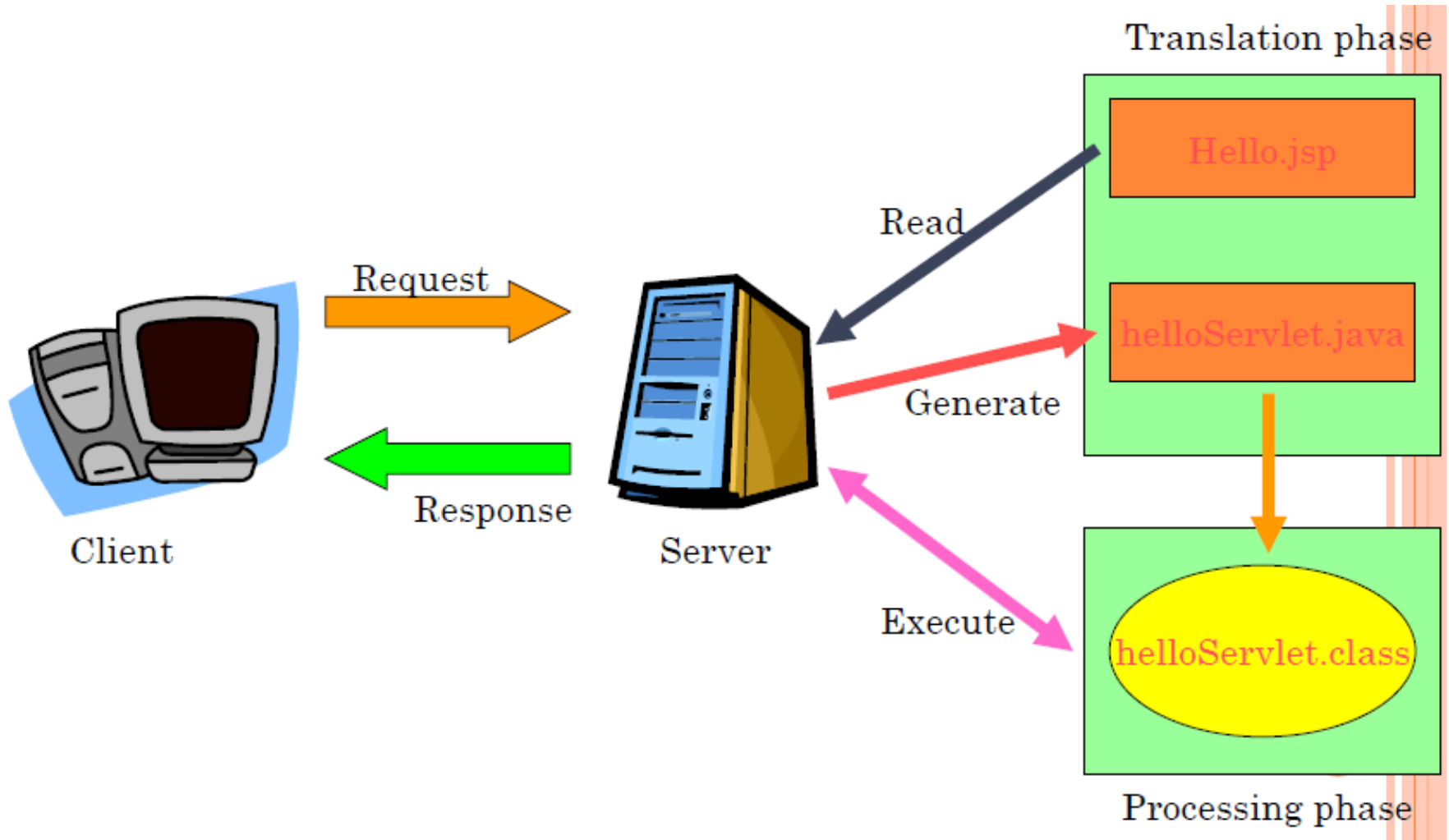
WHAT ARE JAVA SERVLETS?

- An extremely popular Java substitute for CGI scripts.
- They are programs to be run on a web server.
- The web page is based on the data submitted by the user.
- More efficient, easier to use, Powerful and Portable.
- They are Java application programs that are resident on the server and are alternatives to CGI programs.
- Java Servlets allow you to build
 - Web page based on the user's input data
 - Web page that changes frequently
- More efficient, easier to use, more powerful and portable.

WHAT IS JSP?

- It is Java Server Pages that built on top of Java servlets in late 1999.
- In the early days of the Web, the only tool for developing dynamic web content was CGI.
 - For every request, the web server creates a process (not efficient).
- The Java Servlet API was introduced in 1997, however, HTML code has to be embedded inside programs.
- JSP provides a development model for the web authors to experience all the server-side technologies.

JSP PAGE TRANSLATION AND PROCESSING PHASES



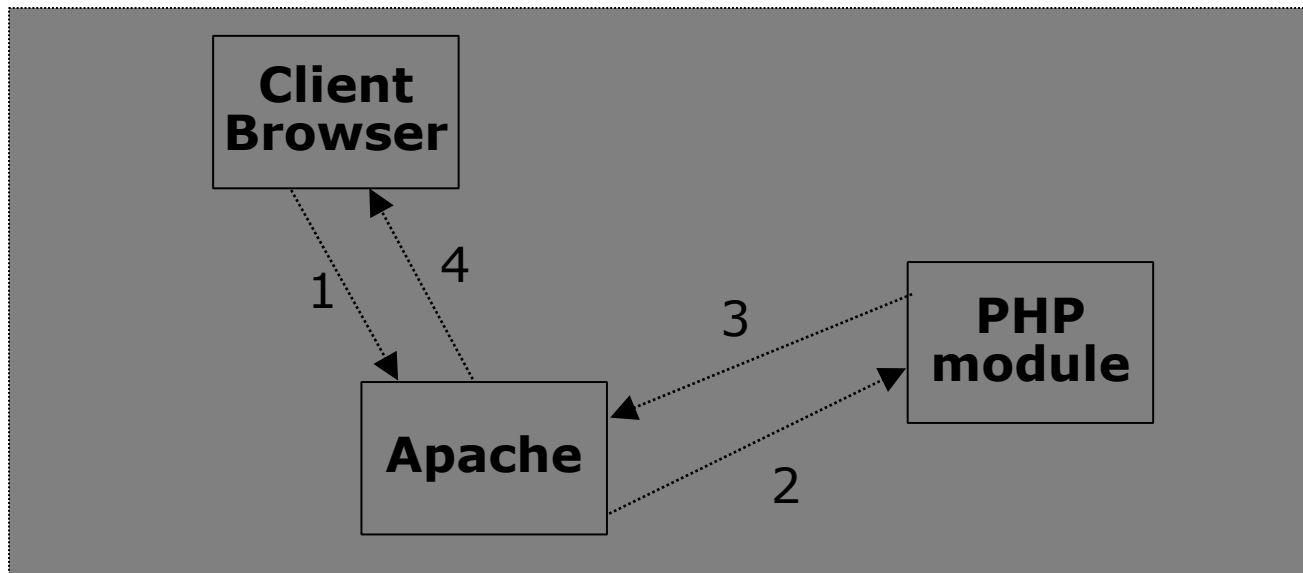
PHP

- PHP stands for "PHP: Hypertext Preprocessor"
- PHP is a server-side scripting language designed specifically for the Web.
- Within an HTML page, you can embed PHP code that will be executed each time the page is visited.

Characteristics

- Very good for creating dynamic content
- Open-source
- Easy to use (C-like and Perl-like syntax)
- Stable and fast
- Multiplatform
- Many databases support
- Many common built-in libraries

How PHP generates HTML/JS Web pages



1. Client from browser send HTTP request (with POST/GET variables)
2. Apache recognizes that a PHP script is requested and sends the request to PHP module
3. PHP interpreter executes PHP script, collects script output and sends it back
4. Apache replies to client using the PHP script output as HTML output

PHP Example

```
<h1>Hello from Dr. Chuck's HTML Page</h1>
<p>
<?php
    echo "Hi there.\n";
    $answer = 6 * 7;
    echo "The answer is $answer, what ";
    echo "was the question again?\n";
?>
</p>
<p>Yes another paragraph.</p>
```



.NET

- .NET is a platform that provides a standardized set of services.
 - It exports a common interface so that it's programs can be run on any system that supports .NET.
- A specific software framework
 - Includes a common runtime
- Microsoft .NET is based on the .NET Framework, which consists of two major components:
 - Common Language Runtime (CLR) and
 - Framework Class Libraries (FCL).

.NET Framework

- Framework Class Library (FCL)
 - Provides the core functionality: ASP.NET, Web Services, ADO.NET, Windows Forms, IO, XML, etc.
 - Enforces security and supplies many other programming capabilities
 - Reusable components that programmers can incorporate into their applications
- Common Language Runtime (CLR)
 - The CLR is at the core of the .NET platform - the execution engine
 - Executes programs written in any .NET-compatible programming language

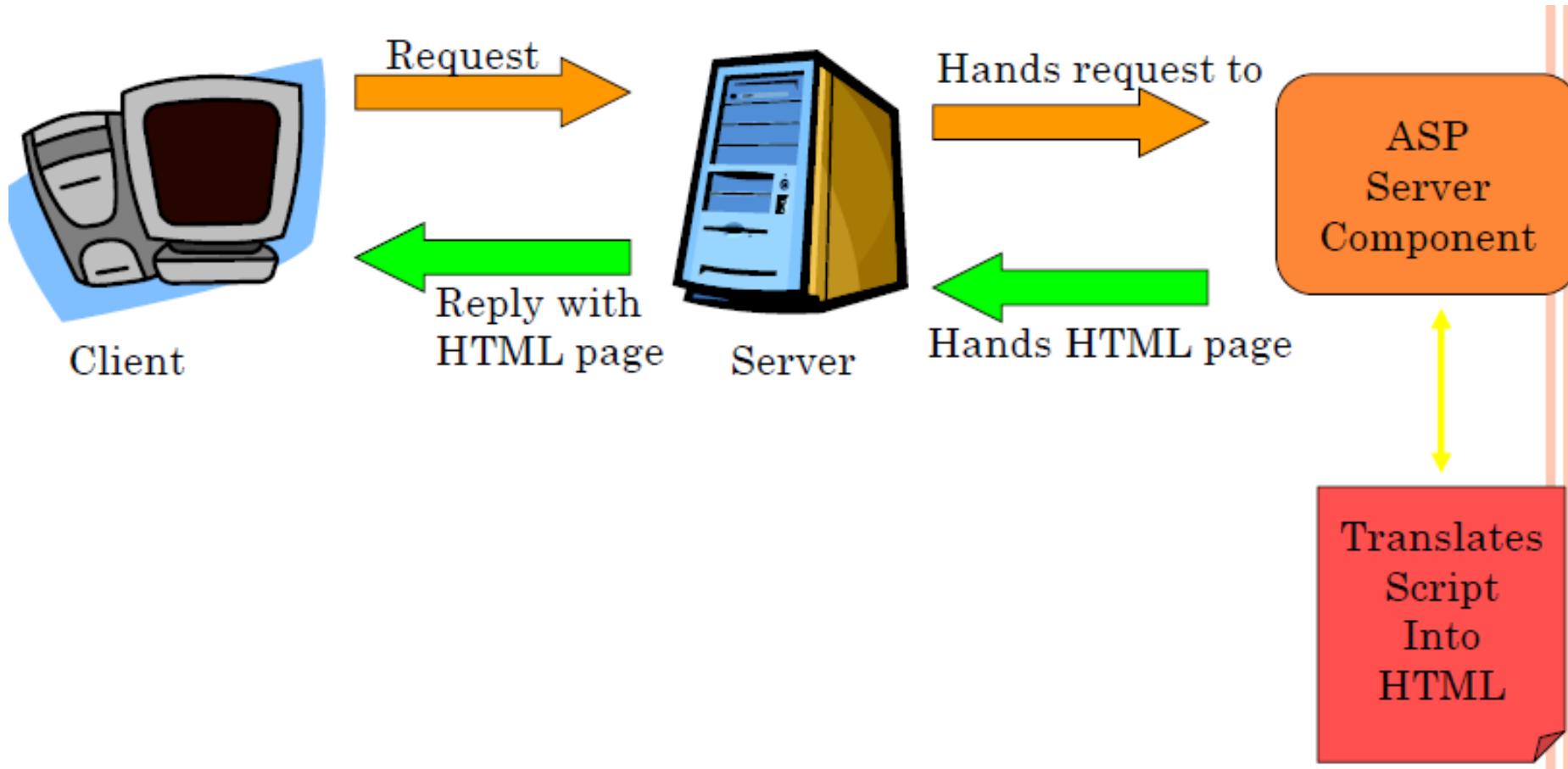
.NET Characteristics

- Independent from a specific programming language
- Promotes software reuse
- Strong emphasis on Web connectivity, using XML web services to connect and share data between smart client devices, servers, and developers/users
- Platform/language independent
- Provides a very good environment to develop networked applications and Web Services
- Provides programming API and unified language-independent development framework

What is ASP?

- ASP (Active Server Pages) is a technology developed by Microsoft.
- Pages using ASP are primarily developed in JavaScript, VBScript , or Perl Script and are integrated into the HTML of your Web pages.
- The ASP code is compiled by the server and the resulting output is standard HTML.
- By using ASP, Web pages can be dynamic, full of ever-changing content, and browser independent
- Active Server Pages enables server-side scripting for IIS with native support for both VBScript and Jscript."

HOW TO LOAD AN ASP PAGE?



JSP vs. ASP

	ASP Technology	JSP Technology
Web Server	IIS or Personal Web Server	Any Web Server
Platforms	Microsoft Windows	Most popular platforms
Reusable components	No	JavaBeans, JSP tags
Security against System crashes	No	Yes
Scripting Language	VBScript, Jscript	Java

- JSP is platform and server independent.
- ASP relies on Microsoft Platforms and Servers.

Web Authoring Tool

- Today's Web authoring tools can provide the power to build an interactive, animated, state-of-the-art Web site suitable for anything from a personal Web page to a midsize business site.
- Example of popular Web Authoring Tools are:
 - NetObjects Fusion
 - Microsoft FrontPage
 - Macromedia Dreamweaver.

WYSIWYG

- What you see is what you get
- In computing, a WYSIWYG editor is a system in which content (text and graphics) displayed onscreen during editing appears in a form closely corresponding to its appearance when printed or displayed as a finished product, which might be a printed document, web page, or slide presentation.
- WYSIWYG implies a user interface that allows the user to view something very similar to the end result while the document is being created
- e.g Dreamweaver 's code view vs Design view

WYSIWYG Contd..

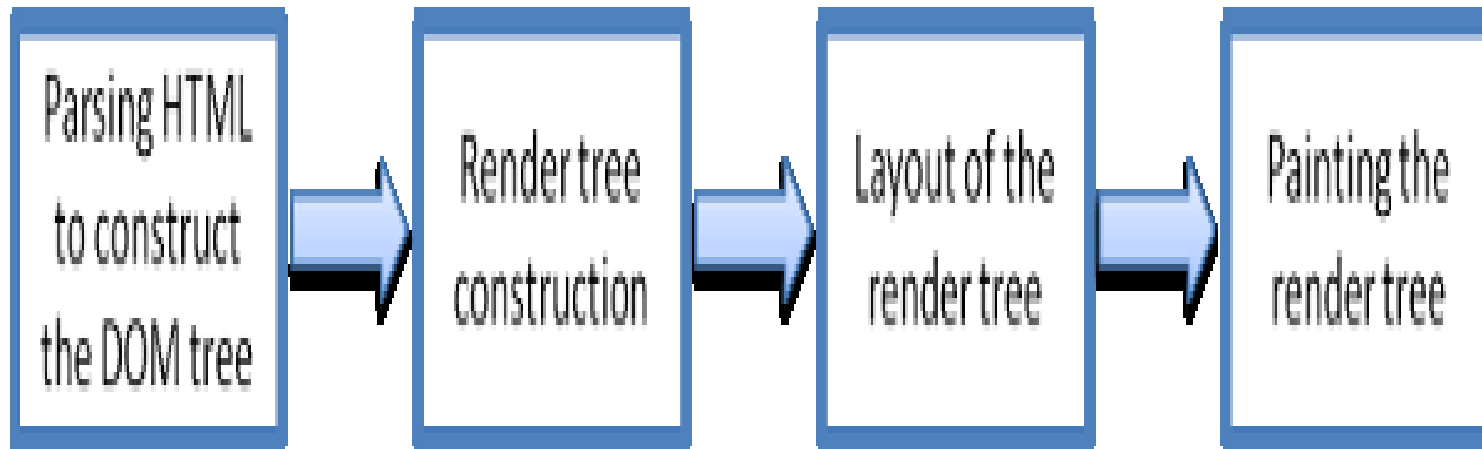
- In general, WYSIWYG implies the ability to directly manipulate the layout of a document without having to type or remember names of layout commands
- Example:
- In presentation programs, compound documents, and web pages,
 - WYSIWYG means the display precisely represents the appearance of the page displayed to the end-user,
 - but does not necessarily reflect how the page will be printed unless the printer is specifically matched to the editing program, as it was with the Xerox Star and early versions of the Apple Macintosh.

BROWSER AS A RENDERING ENGINE

- A web browser engine, (sometimes called layout engine or rendering engine), is a software component that takes marked up content (such as HTML, XML, image files, etc.) and formatting information (such as CSS, XSL, etc.) and displays the formatted content on the screen.
- A layout engine is typically embedded in web browsers, e-mail clients, e-book readers or other applications that require the displaying (and editing) of web content.

- Engines may wait for all data to be received before rendering a page
- May begin rendering before all data are received.
 - This can result in pages changing as more data is received, such as images being filled in if rendering begins before formatting information is received.
- Rendering Engine Example:
 - Gecko: Mozilla
 - WebKit: Safari, Chrome
 - Trident: IE
 - Current versions of Chrome (except iOS version) and Opera are based on Blink, a fork of WebKit.

Basic Flow of Rendering Engine



Thank You

If you have any Queries write to me

@

Jalauddin.mansur@gmail.com

jalawdarling@hotmail.com