ICS 434
Advanced Database Systems

Dr. Abdallah Al-Sukairi
sukairi@kfupm.edu.sa

Second Semester 2003 - 2004 (032)

King Fahd University of Petroleum & Minerals
Information & Computer Science Department

Outline

1. The Relational Data Model: Version 2
2. Advanced Data Modeling
3. Client-Server Architecture
4. Client-Server Databases & Tools
5. **Databases on the Web**
6. The System Catalog
7. Query Processing and Optimization
8. Transaction Processing
9. Concurrency Control
10. Recovery
11. Administration & Security
12. Distributed Databases
13. Database Replication
14. Object-Oriented Databases
15. Data Warehousing and Data Mining
16. Other Emerging Database Technologies

5. Databases on the Web

Requirements for Web-DBMS Integration

- The ability to access valuable corporate data in a secure manner
- Support for session and application-based authentication
- The ability to interface to the database independent of any proprietary Web browser or Web server
- An open architecture
- Acceptable performance
- Support for transactions that span multiple HTTP requests
- Minimal administration overhead
- A connectivity solution that takes advantage of all the features of an organization's DBMS
- A set of high-level application development tools

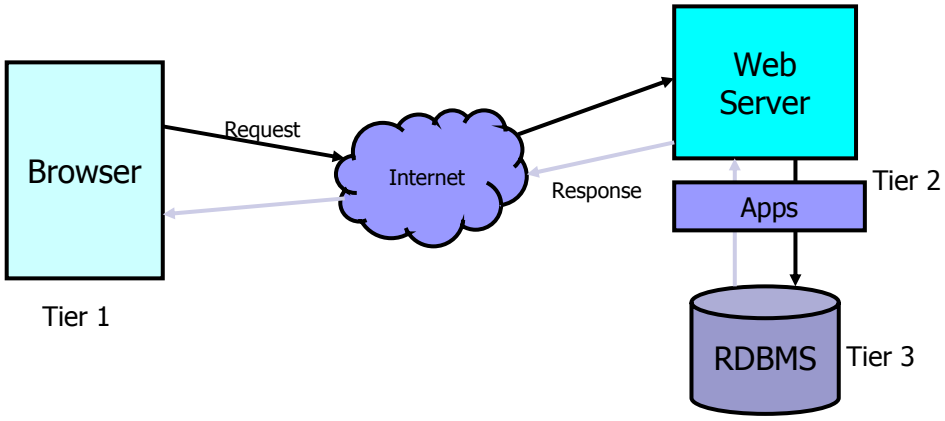
Advantages of the Web-DBMS Approach

- DBMS advantages
- Simplicity
- Platform independence
- Graphical user interface
- Standardization
- Cross-platform support
- Transparent network access
- Scalable deployment
- Innovation

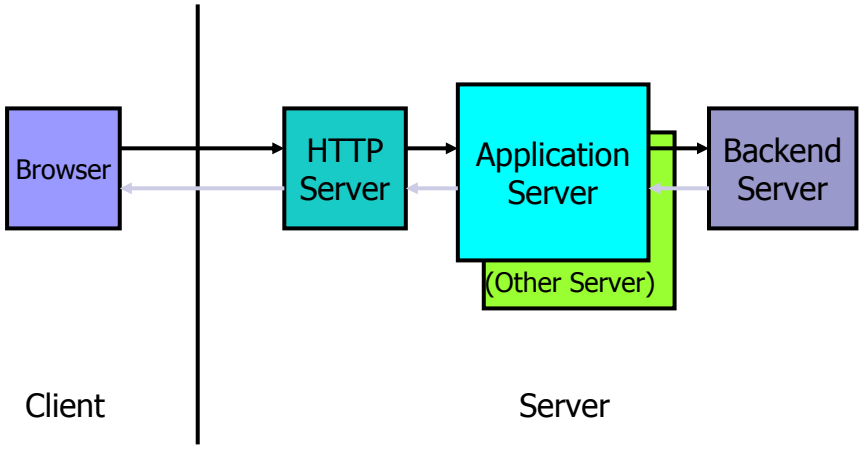
What is a server?

- Many types of server
  - File server — file: networked file space
  - FTP server — ftp: remote file space, often read-only
  - Web server — http: web pages and more
  - Mail server — mail: email system
  - News server — news: newsgroups messages

Dynamic, Interactive Web Server (3-tier Architecture)



General Multi-tier Architecture



Web Content

- HTML documents are static
- Dynamic content
  - user-side programming
  - Middleware
    - ColdFusion, PHP etc.
  - server-side programming
    - scripting languages
    - programming languages

Server-Side Programming

- Database
- Searches
- Form processing
- Counters
- Mailing lists
- Customized pages etc.

Why Server-Side Programming?

- Accessibility
  - You can reach the Internet from any browser, any device, any time, anywhere
- Manageability
  - Does not require distribution of application code
  - Easy to change code
- Security
  - Source code is not exposed
  - Once user is authenticated, can only allow certain actions
- Scalability
  - Web-based 3-tier architecture can scale out

History of Dynamic Web Content

- Common Gateway Interface (CGI) was the first generation approach to providing dynamic web content
  - Used scripts
  - A process dispatched for each web page generated.
  - Hence inefficient and did not scale well.
- Numerous second generation alternatives were invented:
  - FastCGI
  - mod_perl
  - NSAPI
  - ISAPI
  - Java Servlets
- These embedded HTML in programming code so that programmers had to develop them. Costly.

Scripting - the Third Generation Approach

- Embed simple code in HTML pages
- The HTML pages use the code to decide what elements to display and what data should be displayed
- Classes and/or subroutines can be called to compute information for inclusion in the web page. Existing APIs can be invoked.
- This is known as ‘scripting’

Scripting Language or Compiled?

- Scripting Languages:
  - Server Side Includes (SSI)
  - Perl
  - PHP
  - ASP (VBScript)
  - Python
- Compiled Languages:
  - C
  - C++
  - C#
  - ASP .Net
  - Java Servlets
  - Java Server Pages (JSP)
    - Looks like a scripting language, but is actually compiled into a Java Servlet
- Either portable byte code (such as a Java .class file) or a true executable (native to the microprocessor) is produced
- Common to all scripting languages is some sort of real time interpreter that parses text and turns it into executable instructions for the server

Some Approaches to Scripting

- JavaServer Pages (JSP) by Sun Microsystems
- Hypertext Preprocessor (PHP) [open-source]
- ColdFusion (CFML) by Macromedia
- Active Server Pages (ASP and ASP.NET) by Microsoft

Criteria Affecting Decisions

- Web server availability
- Knowledge of language
- Scalability and efficiency
- Personal preference

What is JSP?

- JavaServer Pages
- Java-based technology that simplifies the development of dynamic websites
- Designed around the Java philosophy
- Packaged with J2EE
- As all XML-based technologies – it separates presentation aspects from programming logic contained in the code

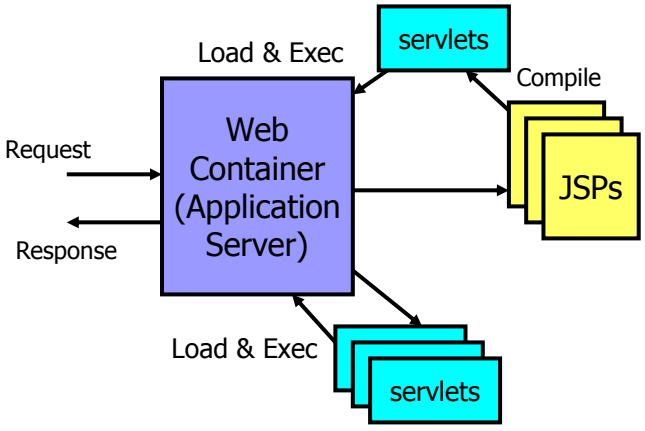
J2EE Technology

- Web container (Application Server)
- EJB container
- Servlets
- JSP (JavaServer Pages), Tag Library
- Applets, XML, JDBC, JNDI, JMS, RMI, IIOP, JavaMail, …

Servlet Container/Engine

- Servlets/JSP require a Container
- Apache Tomcat is the reference implementation of the Servlet/JSP Specs
- It is open source, small, install quickly,and is FREE
- Web Site: jakarta.apache.org/tomcat
- It include a simple HTTP 1.1 server, good enough for development and small intranets

Servlets and JSP



JSP Basics

- Individual JSP pages are text files stored on the web server
- When a page is first requested, the JSP engine uses the page to generate a Servlet
- The compiled Servlet is saved and used to service additional requests
- When a page is modified, the Servlet is regenerated
- Precompilation of pages is also feasible

A Simple JSP

```
<HTML>
<BODY>
Hello, visitor, It is now <%= new java.util.Date().toString() %>
</BODY>
</HTML>
```

- "new java.util.Date().toString()"
  - This creates a date object and converts it to a String that can be displayed.
- The <% %> element can be used to insert regular Java code to control the flow of a page

JSP Pros & Cons

- Not widely supported by web hosting companies
- Not a JavaScript, includes all Java technology for use
- Real Java, not scripting language
- Many deployment choices

What is PHP?

- Open Source server-side scripting language designed specifically for the web
- Conceived in 1994, now used on +10 million web sites
- Supports a wide-range of databases (20 + ODBC).
  - Tightly integrated with MySQL
- Perl- and C-like syntax
- It is cross platform, unlike ASP which generally works only on Microsoft web servers (IIS)
- Website: http://www.php.net/

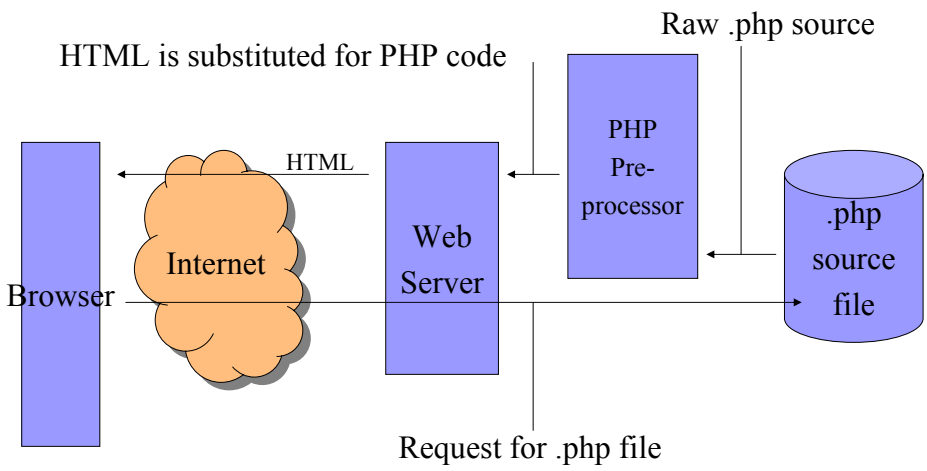
What is PHP?

- Designed similar to Active Server Pages
  - You embed PHP commands into your web pages
- Commands are preprocessed by the PHP processor and appropriate HTML is sent to the web server
- Lots of free or low cost software:
  - http://www.hotscripts.com/PHP/Scripts_and_Programs/ lists over 4000 programs
  - PHP Nuke is an excellent example: free portal software

What is MySQL?

- Relational database management system (RDBMS)
- Free
- Can run on UNIX, Windows and Mac OS
- Website: http://www.mysql.com/

PHP Process Flow



PHP Example

- Begin and end script tags: <?php ?>
  can also use <? and ?>
- PHP statements must be terminated with a semicolon (";")

```
<html>
<body>
<?php echo "Hello, World! I'm using PHP!"; ?>
</body>
</html>
```

PHP Form Data

- Access to the HTTP POST and GET data is simple in PHP
- The global variables $_POST[] and $_GET[] contain the request data

```
<?php
    if ($_POST["submit"])
        echo "<h2>You clicked Submit!</h2>";
    else if ($_POST["cancel"])
        echo "<h2>You clicked Cancel!</h2>";
?>
<form action="post.php" method="post">
    <input type="submit" name="submit" value="Submit">
    <input type="submit" name="cancel" value="Cancel">
</form>
```

PHP Pros & Cons

- Platform independent
- Origins in the C programming language
- Open source
- Many deployment choices
- Easy to learn
- Widely supported

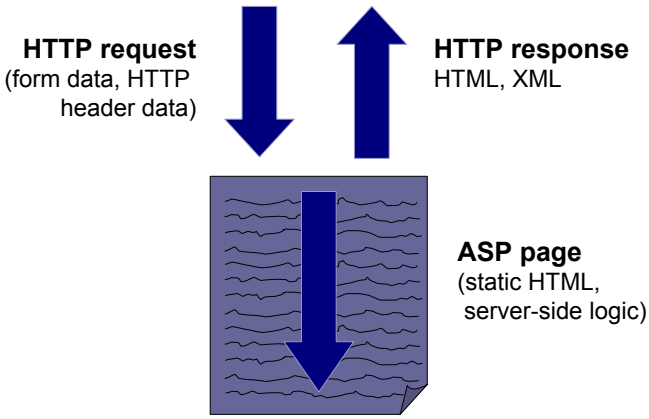
ASP.NET

- Based on .NET framework and the Common Language Runtime (CLR)
- Compiled language
- New languages: Visual Basic .NET, C#, …
- Improved component model
- Web Forms
- ADO.NET
- Web Services

What is ASP?

- Server-side programming technology
- Consists of static HTML interspersed with script
- ASP intrinsic objects (Request, Response, Server, Application, Session) provide services
- Commonly uses ADO to interact with databases
- Application and session variables
- Application and session begin/end events
- ASP manages threads, database connections, ...

What is ASP?



HelloWorld.asp

```
<html>
<head><title>HelloWorld.asp</title></head>
<body>
<form method="post">
<input type="submit" id=button1 name=button1
value="Push Me" />
<%
if (Request.Form("button1") <> "") then
  Response.Write "<p>Hello, the time is " & Now()
end if
%>
</form>
</body>
</html>
```

ASP Successes

- Simple procedural programming model
- Access to COM components
  - ActiveX Data Objects (ADO)
  - File System Object
  - Custom components
- Script-based: no compiling, just edit, save & run
  - VBScript, JScript – leverages existing skills
- Support for multiple scripting languages

ASP Challenges

- Coding overhead (too much code)
  - Everything requires writing code!
- Code readability (too complex; code and UI intermingled)
- Maintaining page state requires more code
- Reuse is difficult
- Supporting many types of browsers is difficult
- Deployment issues (e.g. DLL locking)
- Session state scalability and availability
- Limited support for caching, tracing, debugging, etc.
- Performance and safety limitations of script

ASP.NET Overview

- ASP.NET provides services to allow the creation, deployment, and execution of Web Applications and Web Services
- Like ASP, ASP.NET is a server-side technology
- Web Applications are built using Web Forms
- Web Forms are designed to make building web-based applications as easy as building Visual Basic applications

Goals

- Keep the good parts of ASP and improve the rest
- Simplify: less code, easier to create and maintain
- Multiple, compiled languages
- Fast
- Scalable
- Manageable
- Available
- Customizable and extensible
- Secure
- Tool support

Key Features

- Web Forms
- Web Services
- Built on .NET Framework
- Simple programming model
- Maintains page state
- Multibrowser support
- XCOPY deployment
- XML configuration
- Complete object model
- Session management
- Caching
- Debugging
- Extensibility
- Separation of code and UI
- Security
- ASPX, ASP side by side
- Simplified form validation
- Cookieless sessions

HelloWorld.aspx

```
<%@ Page language="c#" %>
<html>
<head></head>
<script runat="server">
public void B_Click (object sender, System.EventArgs e) {
  Label1.Text = "Hello, the time is " + DateTime.Now;
}
</script>
<body>
  <form method="post" runat="server">
    <asp:Button onclick="B_Click" Text="Push Me“
      runat="server“ /> <p>
    <asp:Label id=Label1 runat="server" />
  </form>
</body>
</html>
```

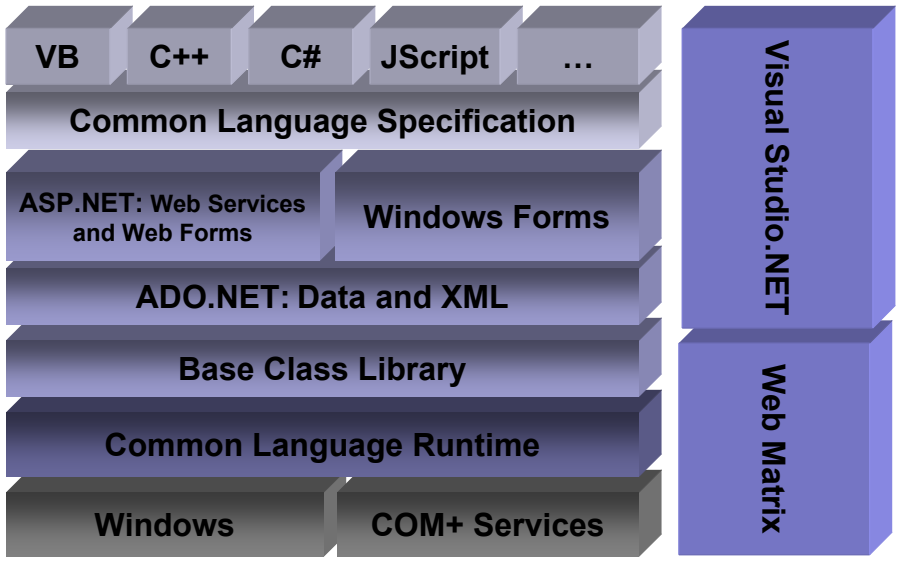
Architecture

- ASP.NET is built upon
  - .NET Framework
  - Internet Information Server (IIS)

Architecture

- Internet Information Server (IIS)
  - IIS MMC Snap-In (Internet Services Manager)
    - Tool to manage IIS
  - Virtual Directories
    - Provides a mapping between URL and file path

Architecture



ASP .NET Web Matrix Project

- Lightweight, simple, community-oriented tool for building ASP.NET apps
- Full WYSIWYG support
- Small (~ 1.4 Mb)
- Community features
  - IM integration, code sharing, chat features
- Available free-of-charge at www.asp.net

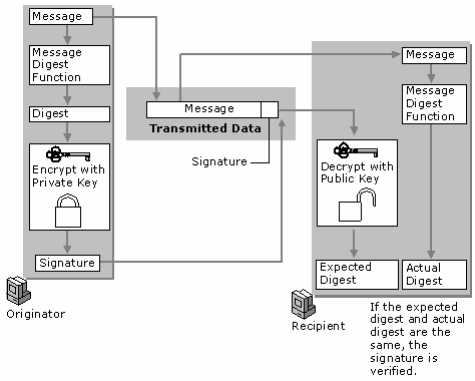
ASP .NET Pros & Cons

- Microsoft only platforms
- Closed source
- Widely supported
- Not only limited to one language (C++, C#, VB, J#)
- One deployment choice -> “Wintel”

Network and Information Security

- Cryptography is the science of providing security for information
- Good cryptography makes attempts to violate security cost-prohibitive
- cryptography can offer the following basic functions:
  - Confidentiality
  - Authentication
  - Integrity
- Basic Components of Modern Cryptography
  - Symmetric Key Encryption
  - Public Key Encryption (e.g. RSA)
  - Secret Key Exchange
  - Message Digest Functions
  - Digital Signatures

RSA Digital Signature Process



Internet Security

- Proxy Servers
  - Improve performance
  - Filter requests
- Firewalls
  - A system designed to prevent unauthorized access to or from a private network
- Encryption
- Secure Sockets Layer (SSL)
- Kerberos
  - A server of secured user names and passwords