Introducing the web

CS 640, Lecture 5

The web in a nutshell

- HTML document
- Conversation using http

Web client ➔ Web server

Inventor of the web: Tim Berners-Lee

Lecture outline

- The client
- The protocol
- The server
- HTML
The browser
- Complex program running on client computer
- Main goal: get web pages from servers, display them and handle interaction with user
- Standards prescribe browser behavior
- Various implementations
  - Internet Explorer, Mozilla/Firefox/Netscape, Safari
  - Have inconsistent interpretation of standards, non-standard extensions
- Executes programs embedded in web pages

Client side programming
<table>
<thead>
<tr>
<th>Feature</th>
<th>Scripts</th>
<th>Applets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>JavaScript</td>
<td>Java</td>
</tr>
<tr>
<td>Access local files</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Network transfers</td>
<td>Limited</td>
<td>Limited</td>
</tr>
<tr>
<td>Graphics</td>
<td>Limited (HTML)</td>
<td>Strong</td>
</tr>
<tr>
<td>User interaction</td>
<td>Extensive</td>
<td>Limited</td>
</tr>
<tr>
<td>Typical uses</td>
<td>Interactive pages</td>
<td>Web games</td>
</tr>
<tr>
<td>Covered in class</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Lecture outline
- The client
- The protocol
- The server
- HTML
HyperText Transfer Protocol

- RFCs: 1945 HTTP/1.0; 2068, 2616 HTTP/1.1
- Well known port number: 80
- Fundamental http operations
  - Client opens TCP connection, requests document
  - Server sends reply to client (includes document)
  - Text-based protocol similar to SMTP, uses MIME
- http conversations also include
  - Context information (e.g. browser type)
  - Negotiation of various options

Uniform Resource Locators

- URLs are used to refer to web documents
  - Browser displays URL of current page
  - URLs are part of http conversations
  - Links between documents also use URLs
  - URLs have a well-defined structure
    - Protocol used to access document (also ftp, https …)
    - DNS name of server holding document
    - Location of file (path + file name)
    - May also have port number – http://server:8080/dir/

Lecture outline

- The client
- The protocol
- The server
- HTML
The web server
- Multithreaded program handling hundreds of requests in parallel
  - IIS (Microsoft), apache (open source)
- When the server receives a request
  - Static pages are read from disk and sent to client
  - For dynamic pages the server also executes some programs that build (parts of) the page
- Other functions
  - Logging accesses
  - Authenticating users

Server side programming
- Short history
  - CGI – separate programs launched by web server
    - They produce an HTML document as output
    - They receive arguments as input
    - Strong isolation, bad performance
  - Programs embedded inside web page (php, ASP, JSP)
    - Program executed inside web server process
- What are dynamic pages used for?
  - Personalizing based on user identity
  - Interacting with databases (e.g. on-line banking)
  - Web applications (e.g. web based email)
  - Separate database keeps persistent data

Web applications gain ground
<table>
<thead>
<tr>
<th>Feature</th>
<th>Web app.</th>
<th>Desktop app.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphics</td>
<td>Strong</td>
<td>Unlimited</td>
</tr>
<tr>
<td>User interaction</td>
<td>Strong</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Network usage</td>
<td>High</td>
<td>Varies</td>
</tr>
<tr>
<td>Accessible from</td>
<td>Any computer</td>
<td>Where installed</td>
</tr>
<tr>
<td>Upgrade cost</td>
<td>Update servers</td>
<td>Update desktop</td>
</tr>
<tr>
<td>Data backup cost</td>
<td>Backup servers</td>
<td>Backup desktop</td>
</tr>
<tr>
<td>Popularity</td>
<td>Increasing</td>
<td>Dominant</td>
</tr>
</tbody>
</table>
Lecture outline

- The client
- The protocol
- The server
- HTML

HyperText Markup Language

- Standard defining how the structure and format of web documents is encoded
- Most browsers render valid HTML consistently
- HTML documents contain text – can include images, forms and interactive elements
- HTML uses tags (e.g. `<b>bold</b>`)  
- HTML documents can be rendered by clients with different constraints (e.g. window size)
- HTML allows links to other documents

An example of a web page

```html
<TITLE>Bucky Badger's web page</TITLE>
<BODY>
<H1>Welcome to Bucky's web page</H1>
<img SRC="bucky.gif">
<p>I am Bucky, the mascot for University of Wisconsin athletics. Please visit </p>
<A HREF="http://www.uwbadgers.com/football/index.html"> the web page of our football team</A> and  
<A HREF="http://www.uwbadgers.com/basketball/index.html"> the web page of our basketball team</A>.
</BODY>
```
Welcome to Bucky's web page

I am Bucky, the mascot for University of Wisconsin athletics. Please visit the web page of our football team and the web page of our basketball team.

Web-related classes

- Goal: learn how to develop web applications using both client and server side technologies
- Hands-on introduction to languages and tools
- "Big picture" issues
- Things covered elsewhere
  - Full mastery of technologies – years of experience
  - Building system software – CS 537
  - Using databases – CS 364
  - Building database systems – CS 564

Schedule of web classes

- HTML
- Javascript
- DHTML
- http
- C#, Visual Studio
- SQL interaction, ADO.Net
- ASP.NET big picture
- ASP.Net controls
- Web services