

The World Wide Web

www.w3schools.com

www.webdesign.com

...



Really everybody can design an own website

Overview

- Internet / Web Concepts
- Three-tier architectures
- Presentation layer
- Middle tier

History: The Internet and the Web

- 12xx Incas use Quipu
- 1945 idea of linking together microfiche published by Vannevar Bush
- 1960s Internet as (D)ARPA project:
fault-tolerant, heterogeneous WAN (cold war!)
term "Hypertext" coined by Ted Nelson
at ACM 20th National Conference
- 1976 Queen Elizabeth sends her first email. First state leader.
- 1980 Berners-Lee at CERN writes notebook program to link arbitrary nodes
- 1989 Berners-Lee makes a proposal on information management at CERN
- 1990 Berners-Lee's boss approves purchase of a NeXT cube
Berners-Lee begins hypertext GUI browser+editor and dubs it "WorldWideWeb"
First web server developed
- 1991 May 17 – general release of WWW on central CERN machines
- 1992 more browsers: Viola & Erwise released
- 1994 > 200 web servers by start of year
Mosaic: first inline images; later "Netscape"

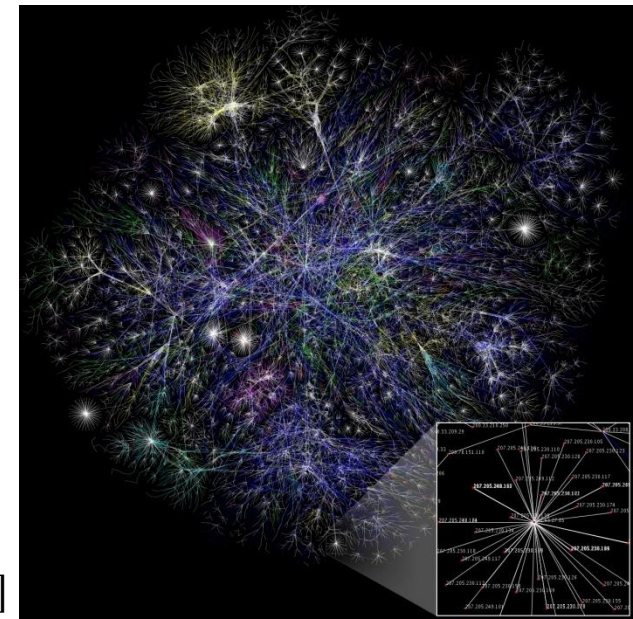
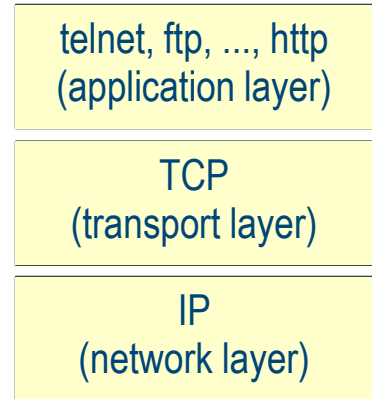


Internet & WWW

- Internet originally **4 basic services**, based on TCP & IP:
 - telnet, ftp, mail, news
 - Later many more: IRC, SSL, NTP, ...

- Each computer has worldwide unique id
 - **IP address**: n.n.n.n (32 bit IPv4, 128 bit IPv6)
 - **Domain name**: subdomain.host.top-level-domain
 - **DNS** to resolve

- **World-Wide Web** just another Internet service
 - HTTP: Hypertext Transfer Protocol
 - HTML: Hypertext Markup Language
 - **URIs** (Uniform Resource Identifiers)



Hypertext Transfer Protocol

- What is a **communication protocol**?
 - Set of rules that defines the structure of messages & communication process
 - Examples: TCP, IP, **HTTP**
- What happens if you click on www.cs.wisc.edu/~dbbook/index.html?
 - Client **connects** to server, **transmits** HTTP request to server
 - Server **generates** response, **transmits** to client
 - Both **disconnect**
- HTTP **header** describes content/action (text = ISO-8859-1), **content** for data
 - RFC 2616

HTTP Request Structure

- Request line

GET ~/index.html HTTP/1.1

- Http **method** field (GET and POST, more later)
- local **resource** field
- HTTP **version** field

- Type of client

User-agent: Mozilla/4.0

- What types of files (MIME types) the client will accept

Accept: text/*, image/gif, image/jpeg

- **MIME** = Multipurpose Internet Mail (!) Extensions = file type naming system
- MIME types other than text/*, image/jpeg, image/gif, image/png need **browser plug-in** or **helper application**

HTTP Response Structure

- **Status line**

HTTP/1.1 200 OK

- HTTP version: HTTP/1.1
- Status code
- Server message, textual

- *200 OK: Request succeeded*
- *400 Bad Request: Request could not be fulfilled by the server*
- *404 Not Found: Requested object does not exist on the server*
- *505 HTTP Version not supported*

- **Date when the object was created**

Last-Modified: Mon, 01 Mar 2002 09:23:24 GMT

- **Number of bytes** being sent

Content-Length: 1024

- What **type** is the object being sent

Content-Type: text/html

- *...plus potentially many more items, such as server type, server time, etc.*

- The **payload!**

<html>...</html>

Conventions

- **index.html** (Windows: **index.htm**), .php, ...
 - If local path ends with directory, this file is assumed
 - *Ex: `http://www.myserver.foo/Downloads`*
 - If not found: **directory listing** is displayed
 - *Put dummy `index.html` if you don't want this, or disable default in server*
- Local path ***~name/path***
 - leads to `~name/public_html/path` where *name* is local user name

HTTP Sample Request/Response

- Client sends:

```
GET ~/dbbook/index.html HTTP/1.1
User-agent: Mozilla/4.0
Accept: text/*, image/gif, image/jpeg
```

- Server responds:

```
HTTP/1.1 200 OK
Date: Mon, 04 Mar 2002 12:00:00 GMT
Server: Apache/1.3.0 (Linux)
Last-Modified: Mon, 01 Mar 2002 09:23:24 GMT
Content-Length: 1024
Content-Type: text/html
```

```
<html> <head></head>
<body>
<h1>Burns and Nobble Internet Bookstore</h1>
Our inventory:
<h3>Science</h3>
<b>The Character of Physical Law</b>
...
</body></html>
```

Try this:

```
$ telnet google.com 80
GET / HTTP/1.1
<3x newline>
```

HTML Primer

- HTML is a data exchange format

- Unformatted ASCII

- Proper indentation increases readability*

- Text interspersed with **tags**, some with **attributes**; usually start and end tag:

```
<h1 align="center">headline</h1>
```

- Opening tags: “<” element name “>”

- Closing tags: “</” element name “>”

- Tags can be **nested**:

```
<h1><em>my</em> text</h1>
```

- Many editors automatically generate HTML directly from your document

- But you **need to know HTML** too, want to generate it later on!

- And tool's code sometimes has bad quality, cf. Microsoft Word “Save as html”

HTML Primer (contd.)

```

<a name="top">
<title>My first HTML document</title>
<h1>An important heading</h1>
<h2>A slightly less important heading</h2>
<p>This is the <em>first</em> paragraph.</p>


```

My link list:

```

<ul>
  <li>This is a link to <a href="http://www.w3.org/">W3C</a>
  <li>This a link to <a href="peter.html">Peter's page</a>
  <li>Go to <a href="#top">top</a>
  <li><a href="/"></a>
</ul>

```

HTML Primer (contd.)

- Text structuring

- Title (for browser title bar)
- Headlines
- Paragraphs, text emphasis

- Links

- External
- Relative
- Internal

- Images

- use **alt**, **width**, **height** attributes!

- Text structuring (contd.)

- Lists

```

<a name="top">
<title>My first HTML document</title>
<h1>An important heading</h1>
<h2>A slightly less important heading</h2>
<p>This is the <em>first</em> paragraph.</p>

My link list:
<ul>
<li>This is a link to <a href="http://www.w3.org/">W3C</a>
<li>This a link to <a href="peter.html">Peter's page</a>
<li>Go to <a href="#top">top</a>
<li><a href="/"></a>
</ul>

```

HTML Primer (contd.)

- Text structuring (contd.)

- tables
- row
- column heading
- regular column

Year	Sales
2000	\$18M
2001	\$25M
2002	\$36M

```

<table>
  <tr>
    <th>Year</th>
    <th>Sales</th>
  </tr>
  <tr>
    <td>2000</td>
    <td>$18M</td>
  </tr>
  <tr>
    <td>2001</td>
    <td>$25M</td>
  </tr>
  <tr>
    <td>2002</td>
    <td>$36M</td>
  </tr>
</table>

```

CSS: Cascading Style Sheets

- Idea: Separate **display style** from **structure & contents**
 - W3C recommendation = standard
- Define appearance of particular items

- HTML element: `body` { font-family: Arial,sans-serif; }
- Self-defined: `a:link` { color: red; }
- Special: `.special` { color: green; font-size: large; }

```
<html>
<body>
  <h1>Title in Arial, but bold</h1>
  <div id="special">I am different</div>
  <a href="#somewhere">link in red</a>
</body>
</html>
```

- All HTML code of site references
common CSS file → Corporate Design

```
<link rel="style sheet" type="text/css" href="books.css">
```

Summary: WWW and HTML

- WWW: another **Internet service**, aimed at easily traversing interconnected documents
- **Protocol**: HTTP, data exchange **format**: HTML
 - captures document structure according to fixed schema
- Browser = program that
 - gets page address; fetches HTML (+ likely additional files); renders page for display
- Separation of concerns:
 - **HTML** for structure and contents
 - **CSS** for layout
 - **JavaScript** for Dynamic HTML