

ECE 435 – Network Engineering

Lecture 4

Vince Weaver

<https://web.eece.maine.edu/~vweaver>

vincent.weaver@maine.edu

29 January 2025

Announcements

- Homework #1 due Friday



Traditional Internet Servers

- Often Client/Server
- Server “daemon”
- Listens on port
 - IANA (Internet Assigned Name Authority) “well-known” ports 0-1023
 - Registered (reserved) ports: 1024-49151
 - Dynamic/Private 49152-65535
- Start at boot time? On demand? Old days inetd, these days systemd



Server Types

- Concurrent – handle multiple connections at time (forks or threads)
 - Concurrent Connectionless – when need fast turnaround, low latency DNS, NFS
 - Concurrent Connection – widely used. WWW.
- Iterative – handles one connection at a time, rest wait on queue
 - Iterative Connectionless – common+trivial, short lived
 - Iterative Connection – high latency



Protocols

- What type of protocol should talk?
- Fixed-length binary?
- Free-form ASCII text?
- 7-bit ASCII vs Unicode?
- Encrypted or compressed?
(security issues from compression?)



Internet, Pre-Web

- These days if someone says internet, often mean WWW
- What was the world like before the web?
 - Mostly Text based (simple, some networks not handle binary)
 - Low-speed, high-latency connections
- Linux machines in 90s came pre-installed with all kinds of servers running
 - e-mail, ftp, usenet, telnet
 - chargen / time / motd / fortune



○ finger / talk / write



What changed?

- Commercial companies allowed on
- Eternal September
- Endless security issues



The World Wide Web – precursors

- Before: getting files via cd-rom or ftp (or e-mail/ftp gateways!), search with archie (archive w/o the V, not comic related)
- gopher: university of Minnesota, 1991. search with jughead/veronica
Why fail? UMN tried to charge license fee, much more restricted file format than html.



The World Wide Web – history

- World-Wide-Web: Tim Berners-Lee, CERN, (NeXT machine) Initial Proposal 1989, first text-based prototype 1991
- Marc Anderson UIUC worked on graphical browser, Mosaic, 1993
- Anderson went on to form Netscape Communications 1994. Webserver software, made Navigator (“mozilla”) relatively cheap/free to drive uptake of web servers.
- Microsoft Internet Explorer. Licensed version of Mosaic.



1995 (as add-on to Win95). MS paid percentage royalties to Spyglass Mosaic, so what happened when they gave it away for free?

- Browser wars.
- Netscape bought by AOL in 1998
- By 2000, IE had over 80% due to bundling with windows, famous lawsuit
- Gap between IE6 and IE7 of 5 years (2001 to 2006)
- Netscape released firefox as open source in 2004
- Safari/Webkit browsers based off of KDE browser
- Google Chrome took over the lead around 2012 or so



- Standards fight. ACID test.
- IE cancelled, with Edge Microsoft joins everyone besides firefox by using webkit-based engine



Top Browsers

1996	Mosaic 1.2%	Netscape 77.3%	IE 19%				
2003	IE 94%	Firefox 2%	Safari —	Opera 1%	Navigator 1%		
2010	IE 42%	Firefox 29%	Chrome 11%	Safari 6%			
2017	Chrome 46.5%	Safari 21.5%	IE 10.1%	Firefox 6.3%	Edge 1.9%	Opera 1.3%	Android 1.2%
2021	Chrome 52.5%	Safari 23.9%	Edge 3.0%	Firefox 4.4%	Samsung 2.2%	Opera 1.0%	

Stats can vary for desktop vs mobile. Stats from Wikipedia. EWS for 1996. TheCounter.com for 2003, wikimedia 2010,2017,2021



Other browsers: midori, lynx, links, w3m



HTML

- HTML – hyper text markup language
- Based on SGML (Standard Generalized Markup Language)
- Hypertext (documents that can link to each other) actually proposed by Vannevar Bush in 1945
- Simplest form, just a text file with some extra commands specified in angle brackets, and usually a closing tag with a / in it. Case insensitive (though supposed to use lowercase these days).



HTML Standards

- Internet Engineering Task Force (IETF) HTML 2.0 in 1994
RFC 1866, 1867, 1942, 1980, 2070
- Since 1996 by the World Wide Web Consortium (W3C)
- 2000 HTML (ISO/IEC 15445:2000)
- HTML 4.01 in 1999
- HTML5 by Web Hypertext Application Technology Working Group in 2014 (in response to FLASH)
- Javascript (ECMAScript) ECMA-262 and ISO/IEC 16262



- HTML4 vs HTML5 vs XHTML
- XML (trendy for a bit) eXtensible markup language, can do things like add new tags on fly



Sample ancient HTML

```
<!DOCTYPE html>
```

```
<html>
```

```
<head><title>ECE435 Test</title></head>
```

```
<body>
```

```
<center><h1>ECE435 Test</h1></center>
```

```
<hr>
```

```
This is a test.
```

```

```



```
<br>Line Break
```

```
<!-- Comment -->
```

```
<p> Paragraph
```

```
<b>Bold</b> <i>Italic</i>
```

```
<a href="other.html">A link to another page</a>
```

```
</body>
```

```
</html>
```



Other HTML tags

- Bulleted list
- Tables relatively easy (also abused for formatting)
- Input / Submitting back to the website (HTML forms)
- Early on vendors went crazy with custom tags: Marquee tag, Blink Tag. Frames.
- “view source”



HTML Formatting

- Originally idea was no formatting, web browser should automatically display simple text in a way to best be displayed on your local machine
Publishers/graphics designers got a hold of it and that's where all the pixel perfect positioning stuff came in
- CSS (cascading style sheets), Javascript



Dynamic Content – Server Side

- cgi-bin: Write a program that takes input as environment vars, output as standard out sent to the requesting browser.

Can write in any program. Typically was things like perl, I often did this in C or even Fortran

- Dynamic content – SSI (server side includes)
- Server extensions (such as PHP, modperl, ASP, .NET) more commonly used (with security issues)



Dynamic Content – Client Side

- Mostly Javascript
- Huge blobs of code that run on your computer, rather than the server
- Lots of security implications



Plugins

- Ways to extend browser with external plugins
- Flash player
- JAVA applets
- Ad blockers
- On the way out for security issues but also companies not wild about ad blockers



Other Modern Web Content

- Web Assembly?
- WebGL?
- Infinite scrolling websites?



WWW client – brief overview

- A URL (uniform resource locator) specifies the document you want (sometimes partly hidden these days)
`http://web.eece.maine.edu/~vweaver/`
- Browser parses URL
- It looks up the address of `web.eece.maine.edu` via DNS
- DNS reports `192.168.8.99`
- Browser makes TCP connection to port 80 (default) of `192.168.8.99`
how do you specify other port? `web.eece.maine.edu:8080`



why would you want to?

security

- The client requests the file `vweaver/index.html` (`index.html` is the default). Often there's a server root, and special handling for user dirs (with the tilde) that are often in user home dirs, as in this case
- The server returns this file
- The TCP connection is closed
- The browser displays the text
- The browser might have to fetch any images linked to by the document



URLs

- URI (uniform resource identifier)
- URL (uniform resource locator) subset of URI, includes info on how to find the resource (protocol and server)
- URN (uniform resource name) asks for a document but from anywhere. I.e. give it something like an ISBN and returns the book
- `scheme:[//[user:password@]host[:port]][/]path[?query] [#fragment]`
`: / ? # [] @ reserved, must encode if use %3f`
`query key1=value1&key2=value2 or key1=value1;key2=value2`
- protocol: http, ftp, file, news, gopher, mailto, telnet



Serving Non-HTML Files

- You can serve up any kind of binary file. Often have associated MIME-type like with e-mail
- Browser also often has built in support
- GIF (trouble due to patents), PNG, JPG, SVG
- MP3 music? Movies?
- Plugins. Flash? Java? PDF?

