

Lecture Notes 1:

The Internet and World Wide Web

CSE 190 M (Web Programming), Spring 2007
University of Washington

Reading: Sebasta Ch. 1 sections 1.1 - 1.5.2, 1.7 - 1.8.5, 1.8.8, 1.9



What is the Internet?

- A "series of tubes" ([1](#), [2](#))

Stevens internet "series of tubes" Search

powered by

Video



[Series of Tubes](#)
Senator Ted Stevens talking about the Net Neutrality Bill
Jul 17, 2006
www.youtube.com

- How many internets are there, anyway?

bush rumors internets Search

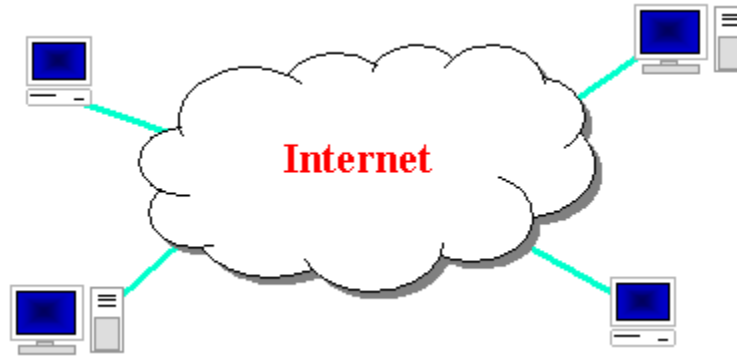
powered by

Video



[Rumors on the internets](#)
George W. Bush clip with the famous: "I hear there're rumors on
Feb 10, 2006
www.youtube.com

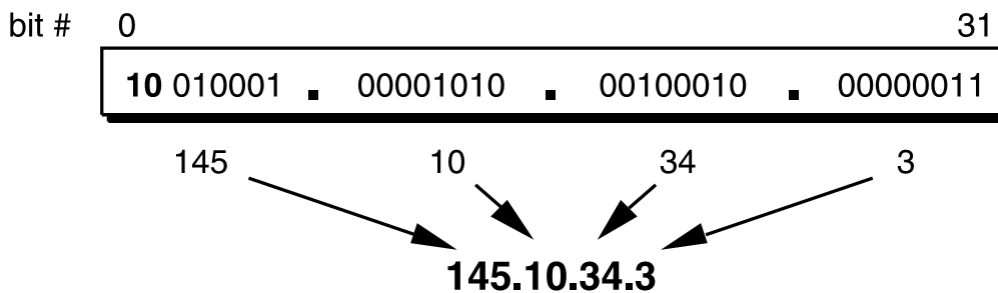
The Internet



- Wikipedia: <http://en.wikipedia.org/wiki/Internet>
- a connection of computer networks using the Internet Protocol (IP)
- What's the difference between the Internet and the World Wide Web (WWW)?
- the Web is the collection of web sites and pages around the world; the Internet is larger and also includes other services such as email, chat, online games, etc.

Internet Protocol (IP)

- a simple protocol for attempting to send data between two computers
- each device has a 32-bit IP address
- addresses written as four 8-bit numbers (between 0 and 255)
e.g. 145.10.34.3



- find out your internet IP address: whatismyip.com
- find out your local IP address:
in a terminal, type: `ipconfig` (Windows) or `ifconfig` (Mac/Linux)

Transmission Control Protocol (TCP)

- adds multiplexing, guaranteed message delivery on top of IP
- **multiplexing**: multiple programs using the same IP address
 - **port**: a number given to each program or service
 - port 80: web browser
 - port 25: email
 - port 22: ssh
 - port 5190: AOL Instant Messenger
 - more common ports
- some programs (games, streaming media programs) use simpler UDP protocol instead of TCP

Domain Name System (DNS)

- a set of servers that map written names to IP addresses
 - Example: `www.cs.washington.edu` → `128.208.3.88`
- many systems maintain a local cache called a hosts file
 - Windows: `C:\Windows\system32\drivers\etc\hosts`
 - Mac: `/private/etc/hosts`
 - Linux: `/etc/hosts`

People and organizations

- World Wide Web Consortium (W3C): web standards
- Internet Engineering Task Force (IETF): internet protocol standards
- Internet Corporation for Assigned Names and Numbers (ICANN): decides top-level domain names



Web Server

- a computer running web server software that listens for web page requests on TCP port 80
- popular web server software:
 - Apache: www.apache.org
 - Microsoft Internet Information Server (IIS)
 - available as part of Windows XP (directions)



Web Browser

- a software application that displays web pages
- popular web browser software:
 - Mozilla Firefox: getfirefox.com
 - Microsoft Internet Explorer (IE): part of Windows
 - Apple Safari: part of Mac OS X
 - Opera: opera.com



Hypertext Transport Protocol (HTTP)

- the set of commands understood by a web server and sent from a browser
- some HTTP commands (your browser sends these internally):
 - GET *filename* : download
 - POST *filename* : send a web form response
 - PUT *filename* : upload
- simulating a browser with a terminal window:

```
$ telnet www.cs.washington.edu 80
Trying 128.208.3.88...
Connected to 128.208.3.88 (128.208.3.88).
Escape character is '^]'.
GET /index.html
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 ...">
<html>
...

```

HTTP error codes

- the web server returns a special "error code" number to the browser, possibly followed by an HTML document
- common error codes:

Number Meaning

200 OK

301-303 page has moved (permanently or temporarily)

403 you are forbidden to access this page

- 404 page not found
- 500 internal server error

- complete list of error codes

Uniform Resource Locator (URL)

- an identifier for the location of a document on a web site
- a basic URL:

```

http://www.aw-bc.com/info/regesstepp/index.html
~~~~~
protocol      host                path

```

- upon entering this URL into the browser, it would:
 - ask the DNS server for the IP address of `www.aw-bc.com`
 - connect to that IP address at port 80
 - ask the server to GET `/info/regesstepp/index.html`
 - display the resulting page on the screen

More advanced URLs

- **anchor:** jumps to a given section of a web page

```

http://www.textpad.com/download/index.html#downloads

```

- the above URL fetches `index.html` and then jumps downward to a part of the page labeled `downloads`
- **port:** for web servers on ports other than the default 80

```

http://www.cs.washington.edu:8080/secret/money.txt

```

- **query string:** a set of parameters passed to a web program

```

http://www.google.com/search?q=miserable+failure&start=10

```

- the above URL asks the server at `www.google.com` to run the program named `search` and pass it two parameters:
 - `q` (set to "miserable+failure")
 - `start` (set to 10)

Web programming technologies

- Hypertext Markup Language (HTML): used for writing web pages
 - XHTML: A newer, more rigidly standardized specification of HTML
- Cascading Style Sheets (CSS): supplies stylistic info to web pages
- Javascript: allows interactive and programmable web pages
- Asynchronous Javascript and XML (AJAX): allows fetching of web documents in the background for enhanced web interaction
- Google AJAX Search API: embeds Google services in web pages
- PHP Hypertext Processor (PHP): allows the web server to create pages dynamically
- Structured Query Language (SQL): interaction with databases