The Internet and WWW

Class 1

Administrative

- class web site `https://borax.truman.edu/315`
- syllabus
- office hours
- Blackboard for grades
- coding style guide

Sources

- online sources
  - developer.mozilla.org
  - w3.org
  - php.net
- make sure you look at HTML5, CSS3, PHP7 stuff
- JavaScript (ECMAScript) ES6 minimum
- online references are for low-level things, not for solutions

Software Environment

- some of the work in this course can be done on your own computer
- most will have to be done on, or at least uploaded to, the department's Linux system sand which has a public webserver
- you will need to use a code editor such as Emacs, vim, bluefish, or VSCode

Written Assignments

- homework assignments will mostly be source code in text files
  - HTML
  - CSS
  - JavaScript
  - PHP
  - SQL
- occasionally submitted as standalone files
- usually installed on *sand* where I will access them

Tests

- tests will be answers written in ASCII text files, using a text editor
- your choice of text editor is very important

Code Editor

- Linux
  - Emacs (this is what I use)
  - vim
  - geany
  - bluefish
  - VScode
- Mac
  - TextMate
  - atom
  - Emacs
  - vim
  - VScode
- Windows
  - atom
  - Emacs
  - vim
  - notepad++
  - VScode

Keys to Success

1. attend class
2. by 8:30, you need to be seated ready to take notes

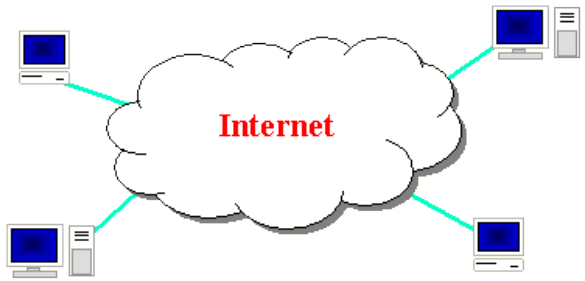
3. I post the slides of each unit on the course calendar after the unit is complete
4. participate! don't sit there confused — interrupt me and ask your question
5. interaction is part of your participation grade

Keys to Success

2. program!
   - computer science is so much more than programming
   - but you cannot be a computer scientist without programming
   - this course requires good programming skills and adherence to style standards
   - you must have completed CS181 with C or better
   - my recommendation is that you take CS260 before this course

   - program every day
   - type in, run, and experiment with
     - my code examples
     - code you find online
     - programs you write yourself

The Internet

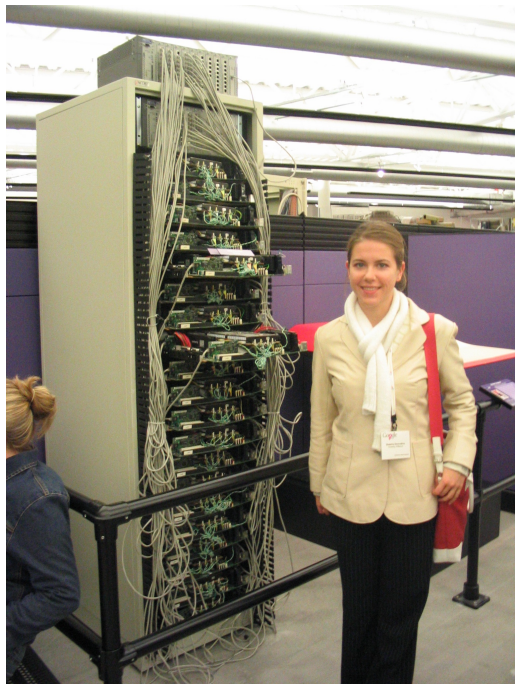


A connection of networks using the internet protocol (IP)

Layer	Name	Protocol
7	Application	HTTP
6	Presentation	MIME
5	Session	full duplex
4	Transport	TCP
3	Network	IP
2	Data Link	connections, errors
1	Physical	pins, voltages, cables

A Few Milestones of the Internet

- 1960s–1970s US DoD network ARPANET
  - initial services: electronic mail, FTP file transfer
- late 1980s opened for educational and then commercial use
- 1991 gopher by Mark McCahill at UMinn
- 1991 WWW by Tim Berners-Lee at CERN
- web browsers
  - 1993 Mac Samba text-only browser
  - 1993 NCSA's Mosaic first graphical browser
  - 1994 Netscape free product, but not open source
  - 1995 IE bundled with Windows
  - 2004 Firefox free open source software
- 1995 Amazon
- 1996 Google



Key Aspects of the Internet

- subnetworks can stand on their own
- computers can dynamically join and leave the network
- built on open standards
  - anyone can create a new internet device
  - accessible with open, simple, commonly available software
- little centralized control

Key Organizations

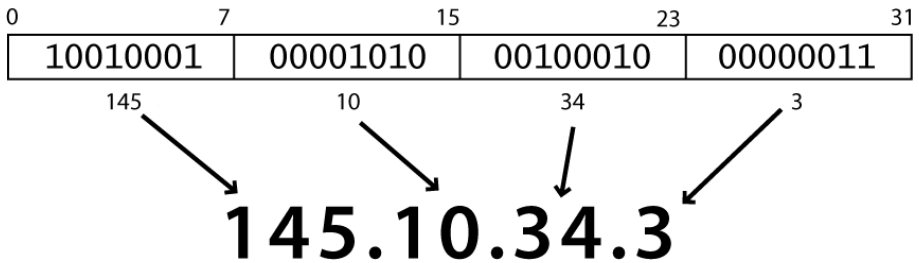
IETF Internet Engineering Task Force
internet protocol standards

ICANN Internet Corporation for Assigned Names and Numbers
addresses and domain names

W3C World Wide Web Consortium
web document standards

Internet Protocol (IP)

- a simple protocol to send data packets between two devices
- each device has a 32-bit IP address written as four 8-bit unsigned numbers (each 0–255)[1]



- find out your outward-facing IP address: whatismyip.com
- find out your local IP address with shell command:
  - `ipconfig` (Windows)
  - `ifconfig` (Mac and Linux)

[1] This is IPv4; IPv6 is similar but has a bigger address space.

Transmission Control Protocol (TCP)

- packet reassembly: if packets arrive out of order, re-order them
- address multiplexing: multiple programs can use the same IP address with different ports

port	purpose
22	incoming ssh
25	incoming email
80	unencrypted web server (http)
110	email client via POP3
118	SQL database access
143	email client via IMAP
389	LDAP directory queries
443	encrypted web server (https)

Domain Name Service (DNS)

- a distributed protocol that maps text names to IP addresses
  - www.google.com $\Rightarrow$ 74.125.225.144
  - www.truman.edu $\Rightarrow$ 150.243.160.15
  - borax.truman.edu $\Rightarrow$ 150.243.160.61
- look up an address (on Linux) with

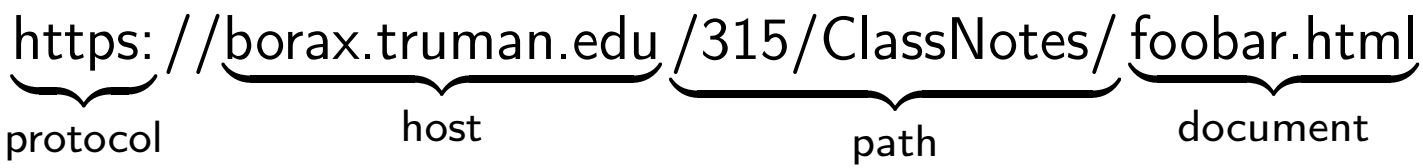
  ```
  $ host www.microsoft.com
  ```

The World Wide Web

- a subset of the internet (note lower case)
- a set of documents that are linked
- a set of web servers that listen for document requests
- a set of clients that send requests for documents

Uniform Resource Locator (URL)

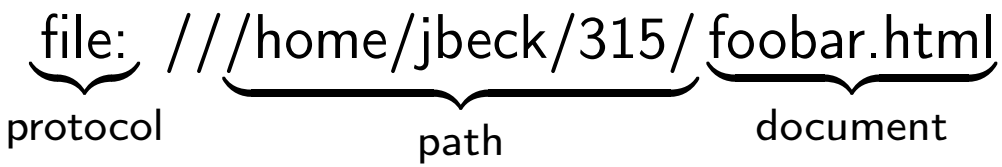
- resources on the web are specified by URLs



- given this URL a browser would:
  1. ask a DNS server for the IP address of borax.truman.edu
  2. connect via TCP to that IP address, port 443
  3. send "GET foobar.html" on the path /315/ClassNotes/ to Apache (which listens on that port)
  4. receive the resulting stream as a document
  5. render and display the resulting document

A Useful URL

- the file: protocol instructs the browser to explore the local filesystem
- slightly different path format between Unix-based and Windows-based
- renders HTML, CSS, and JavaScript without requiring a server



Links

- before WWW, links did not exist
- the user had to know where a document was located and manually retrieve it using the fully qualified document address

- links are what make the WWW a web
- links are designated in HTML with the a (anchor) element

```
<p>
  It is easy to validate your HTML code at
  <a href="https://validator.w3.org">W3C</a>.
</p>
```

It is easy to validate your HTML code at W3C.

Media Types (MIME types)

- the web comprises many document file formats
- listed in `/etc/mime.types` on Linux and Mac
- filename suffixes are quite important

MIME Type	**e**xtensions	**n**otes
audio/mpeg	.mp3, .mpg	audio files
image/jpeg	.jpg, .jpeg	JPEG images
image/png	.png	PNG images
multipart/form-data		web form data
text/css	.css	style sheet files
text/html	.htm, .html, others	web pages
text/javascript	.js	Javascript programs
text/plain	.txt	plain text
text/xml	.xml	XML data
video/mp4	.mp4	video & audio
video/quicktime	.qt, .mov	QuickTime movies

Assignment

- make sure you can access sand
- read the FAQ: https://sand.truman.edu/faq
- make sure your home directory and public_html exist with correct permissions
- play with editors, decide which one you want to use
- look up how to customize it for no tabs, indent size, etc.

official assignment due noon on Thursday, 13 January
linked to the course calendar