EECS1012

Net-centric Introduction to Computing

Lecture 1:
The Internet, WWW (and a bit of history)

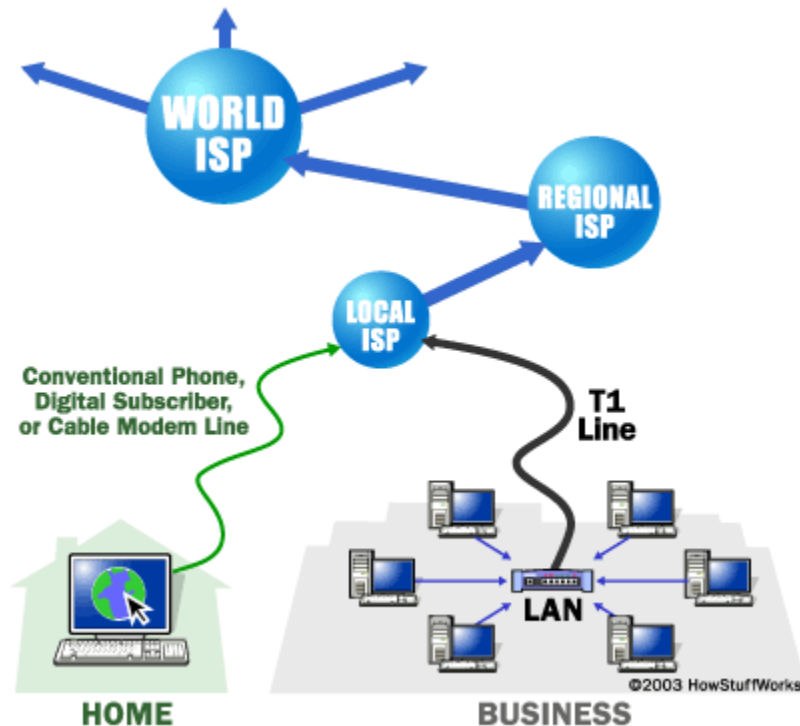# The Internet and the WWW

What is the Internet?

- A "series of tubes"

- How many Internets are out there?
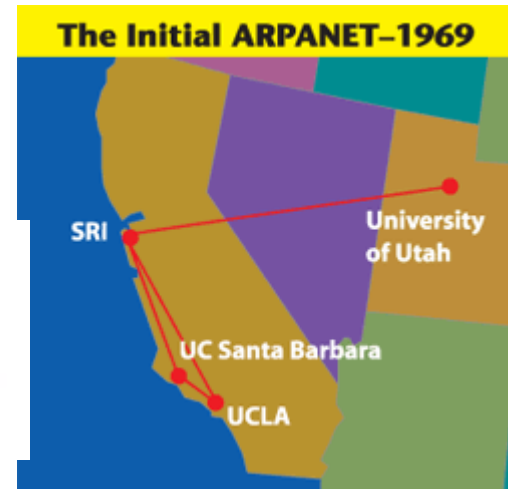
- Is Google one of them?

EECS 1012

What is the Internet?

- A collection of computer networks that use a protocol to exchange data

- Is the World Wide Web (WWW) and the internet the same?

EECS 1012

Brief history

- Began as a US Department of Defense network called ARPANET (1960s-70s)
- Packet switching (in the 60s)
- E-mail is born on 1971
- TCP/IP beginning on 1974 (Vinton Cerf)
- USENET (1979)
- By 1987: Internet includes nearly 30,000 hosts

EECS 1012

The Initial ARPANET-1969

SRI
University of Utah
UC Santa Barbara
UCLA

USENET

Usenet Help
UIC – [Part 1 | Intro | FAQ | History | Menu/Help
| Part 2 | Home]

* What Is Usenet:
* Where and How should I post:
* Alt.Binaries.Pictures Info:
* Getting On Usenet:
* Creating Newsgroups:
* Usenet Software:

Brief history (cont.)

- WWW created in 1989-91 by Tim Berners-Lee
- Popular web browsers released:
  - Netscape 1994
  - IE 1995
- Amazon.com opens in 1995
- Google, January 1996
- Wikipedia launched in 2001
- MySpace opens in 2003
- Facebook February 2004

Berners-Lee

The future of the Internet?

EECS 1012

The future of the Internet?

EECS 1012

Key aspects of the Internet

- ☐ Sub-networks are independent

- ☐ Computers can dynamically join and leave the network

- ☐ Built on **open standards**

- ☐ Lack of centralized control (mostly)

- ☐ Everyone can use it with simple, commonly available software

EECS 1012

People and organizations

- Internet Engineering Task Force (IETF): internet protocol standards

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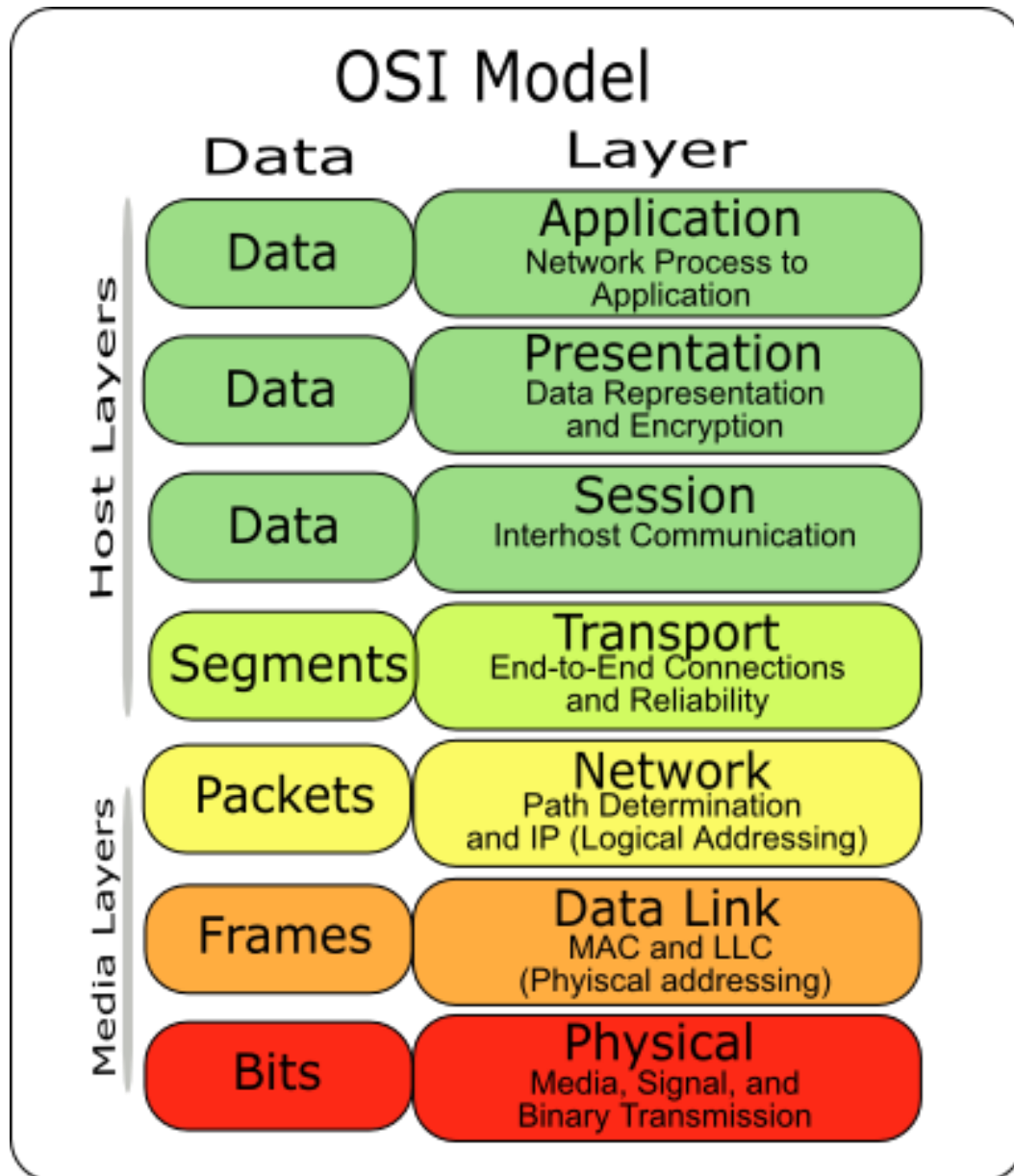
- World Wide Web Consortium (W3C): web standards

EECS 1012

Layered architecture

- ☐ The internet uses a layered hardware/software architecture (also called the "OSI model"): OSI model

- ☐ **physical layer** : devices such as ethernet, coaxial cables, fiber-optic lines, modems

- ☐ **data link layer** : basic hardware protocols (ethernet, wifi, DSL PPP)

- ☐ **network / internet layer** : basic software protocol (IP)

- ☐ **transport layer** : adds reliability to network layer (TCP, UDP)

- ☐ **application layer** : implements specific communication for each kind of program (HTTP, POP3/IMAP, SSH, FTP)

Layered architecture

OSI Model

Data	Layer
Host Layers	
Data	**Application** — Network Process to Application
Data	**Presentation** — Data Representation and Encryption
Data	**Session** — Interhost Communication
Segments	**Transport** — End-to-End Connections and Reliability
Media Layers	
Packets	**Network** — Path Determination and IP (Logical Addressing)
Frames	**Data Link** — MAC and LLC (Phyiscal addressing)
Bits	**Physical** — Media, Signal, and Binary Transmission

Internet Protocol (IP)

- Simple protocol for data exchange between computers
- IP Addresses:
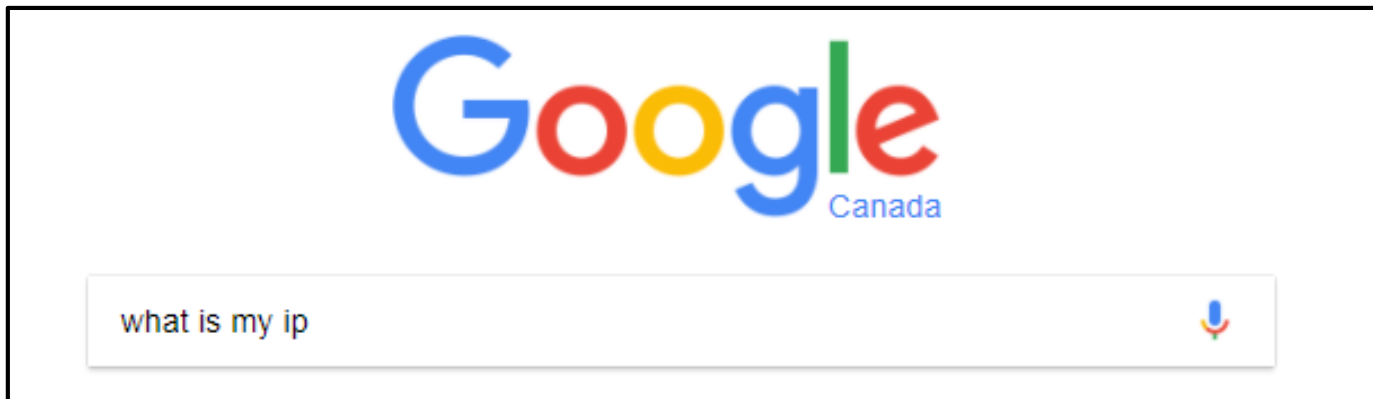 - 32-bit for IPv5 - (Four numbers between 0-255)
  - 128-bit for IPv6

0	7	15	23	31
10010001	00001010	00001010	00000011
145	10	34	3

145.10.34.3
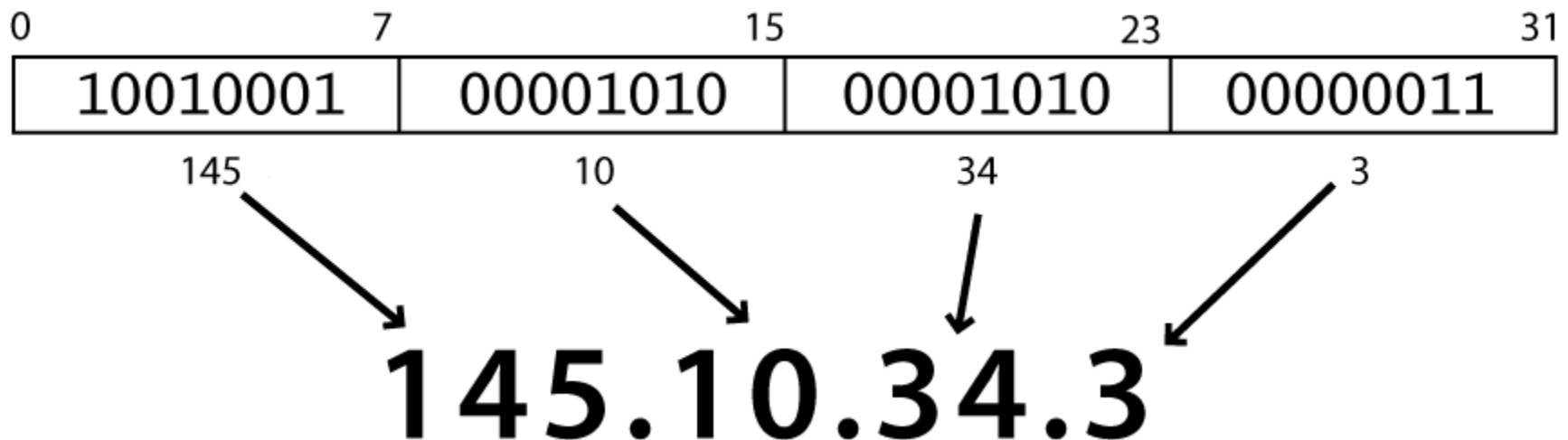
Find your IP address

☐ go to "whatismyip.com"

Your IP Address Is:

184.147.44.55

☐ Or google it:

Google
Canada

what is my ip

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 443 for secure browsing)
  - port 25: email
  - port 22: ssh

- Some programs (games, streaming media programs) use simpler UDP protocol instead of TCP

EECS 1012

The World Wide Web

Web servers and browsers

- **Web servers:** software that listen for web page requests and returns web page content

- **Web browsers:** fetches/displays documents from web servers

Web Servers

- **Web server:** software that listens for web page requests
    - Apache
    - Microsoft Internet Information Server (IIS)

EECS 1012

Web Browser

- **Web browser**: fetches/displays documents from web servers
  - Mozilla Firefox
  - Microsoft Internet Explorer (IE)
  - Apple Safari
  - Google Chrome
  - Opera

EECS 1012

Domain Name Server (DNS)

- Set of servers that map written names to IP addresses
  - Example: yorku.ca → **130.63.236.137**
- Many systems maintain a local cache called a hosts file
  - Windows: C:\Windows\system32\drivers\etc\hosts
  - Mac: /private/etc/hosts
  - Linux: /etc/hosts

Uniform Resource Locator (URL)

- An identifier for the location of a document on the web

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

- URL lets your browser know what type of protocol
- What is the _host_ (DNS server will find the IP)
- What webpage to get (_path_)

More advanced URLs

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

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

  - fetches index.html then jumps down to part of the page labeled "downloads"

- port: for web servers on ports other than the default 80

http://portquiz.net:8080/ ← port 8080

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

http://www.google.com/search?q=poutine

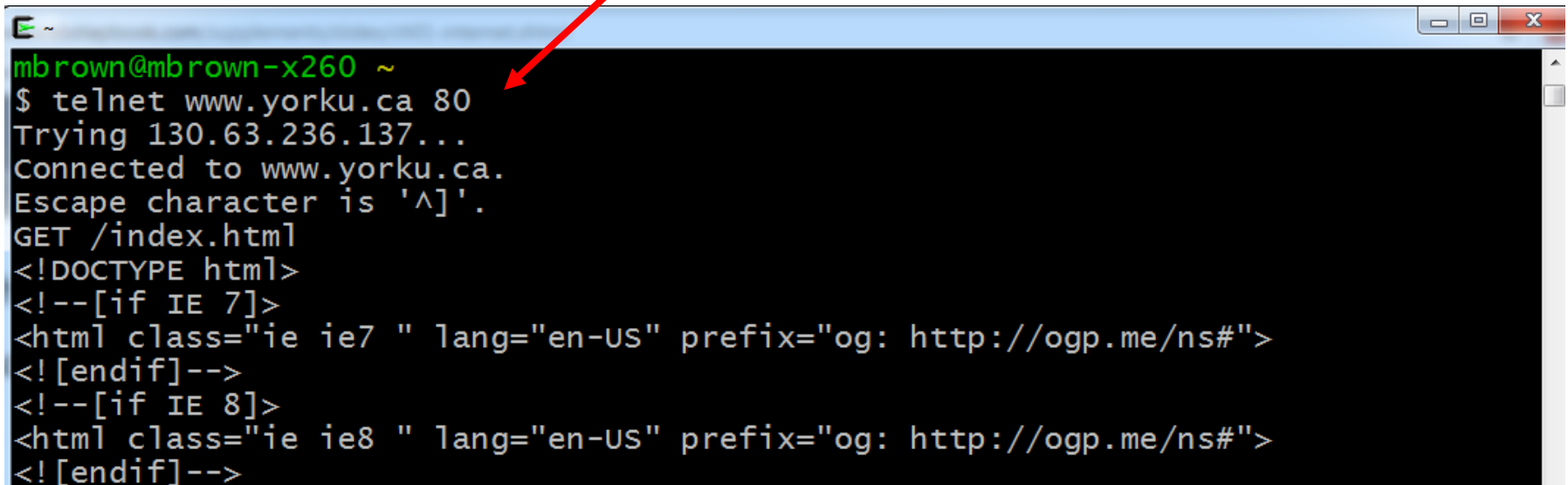 - parameter q is set to "poutine"

Hypertext Transport Protocol (HTTP)

- 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

- *Exercise*: simulate a browser with a terminal window (see next slide)

EECS 1012

HTTP protocol example

I contacted yorku.ca (using telnet)
I typed "GET /index.html"
The webpage contents were returned.

```
mbrown@mbrown-x260 ~
$ telnet www.yorku.ca 80
Trying 130.63.236.137...
Connected to www.yorku.ca.
Escape character is '^]'.
GET /index.html
<!DOCTYPE html>
<!--[if IE 7]>
<html class="ie ie7 " lang="en-US" prefix="og: http://ogp.me/ns#">
<![endif]-->
<!--[if IE 8]>
<html class="ie ie8 " lang="en-US" prefix="og: http://ogp.me/ns#">
<![endif]-->
```

HTTP Error Codes

- When something goes wrong, the web server returns a special "error code" number

- 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

EECS 1012

https://en.wikipedia.org/wiki/List_of_HTTP_status_codes

Error Codes (plain English meaning)

Code	Meaning	Plain English (From user perspective)
1xx	Informational; indicates a provisional response, e.g. 100	OK so far and client should continue with the request
2xx	Successful	All good
3xx	Redirection	Something moved
4xx	Client Error	You messed up
5xx	Server Error	We messed up

Internet Media ("MIME") types

MIME type	file extension
text/html	.html
text/plain	.txt
image/gif	.gif
image/jpeg	.jpg
video/quicktime	.mov
application/octet-stream	.exe

EECS 1012

Web Languages

- Hypertext Markup Language (HTML): used for writing web pages

- Cascading Style Sheets (CSS): stylistic info for web pages

- PHP Hypertext Processor (PHP): dynamically create pages on a web

- server

- JavaScript: interactive and programmable web

Summary

- The internet is developed in the 1960s and 1970s
  - Many of the pioneers who help create it are still alive
- Established protocols (IP, TPC, HTTP, etc) allow machines to "talk" to each other
- HTML and CSS is a language used to describe webpages
- Programming languages such as JavaScript and PHP are used to make webpages more interactive