Web Infrastructure

CS 130:
Tools and Technologies of the World Wide Web
Spring, 2022

Announcements

- Evening office hours have been moved back this week (now 7-9PM)
- Nice job on Lab 1 – keep writing that code!
- During this week's lab, we will be getting an introduction to GitHub, which is how we will be publishing your HTML, CSS, and JavaScript files to the server (so that the rest of the world can look at your stuff).
- There will be a (masked) photographer in class on Wednesday to take photos.

Let’s find out what you already know about the Internet...

Outline

1. **What is the internet?**
2. What is the web?

Some questions about the Internet...

1. What is the cloud?
2. Who invented the Internet?
3. Where is the Internet?
4. Who controls the Internet?
5. What is Net Neutrality?
6. Who can access your data over the internet?

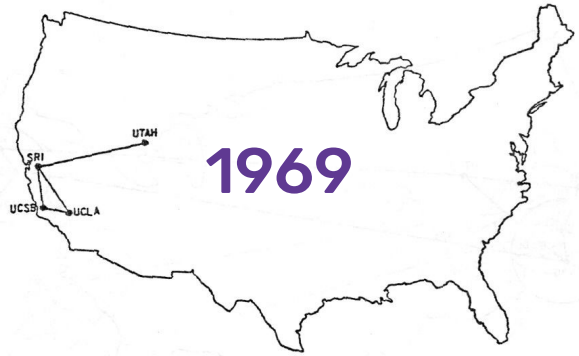
What is the cloud? What are some examples of cloud services?

- The cloud is basically having access to someone else's computer(s)
  - Refers to software and information and computing services (including storage) that run on someone else's computer instead of on yours
- Cloud services can be accessed through a Web browser like Firefox or Google Chrome, through mobile devices, IoT services, etc.
  - What this means is that your computer is **constantly** interacting with the cloud.
- You can also access the cloud via command line tools (e.g. Heroku client, Git, cURL, wget, etc.)

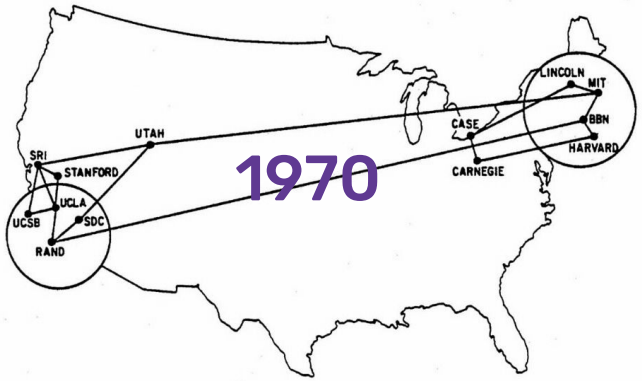
Who Invented the Internet?

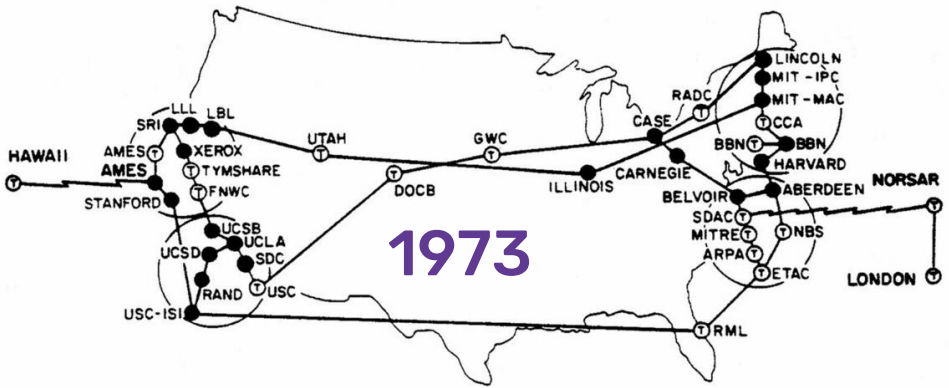
- The internet began as ARPANET, an academic research network that was funded by the military (DARPA), beginning in 1969
- In 1973, Vint Cerf and Bob Kahn began work on TCP/IP, the next networking standards that became the foundation of the modern internet
- In 1981 funding for the internet shifted to the NSF, which funded the long-distance networks that served as the internet's backbone until 1994
- In 1994, the Clinton Administration privatized the internet backbone

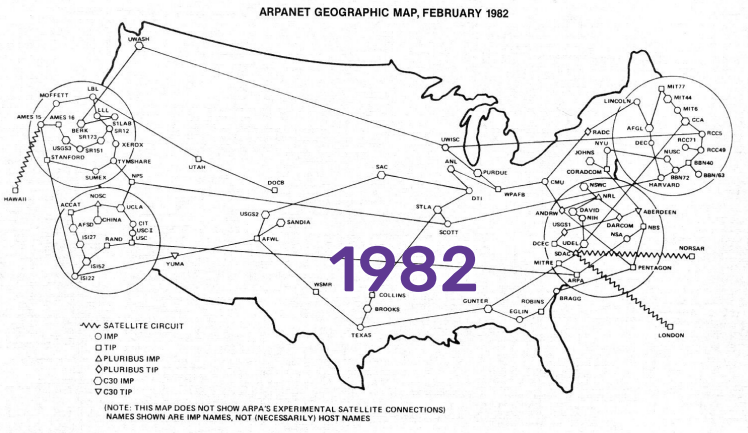
Source: https://www.vox.com/2014/6/16/18076282/the-internet



The ARPANET in December 1969







Source: https://www.vox.com/a/internet-maps

Where is the Internet?

Three primary components:

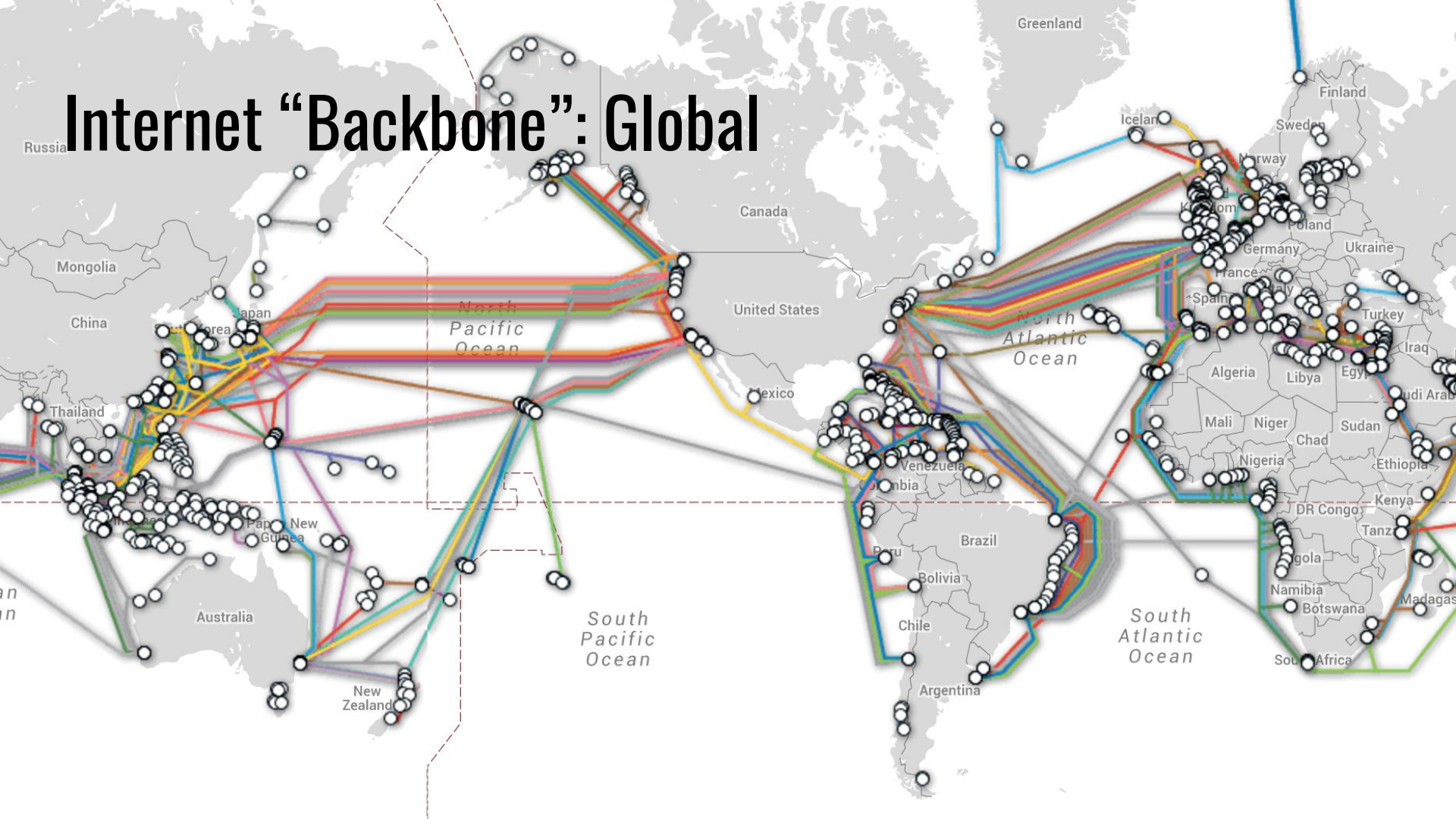
1. **The Internet Backbone**
   Long-distance networks — mostly on fiber optic cables — that carry data between data centers and consumers

2. **The “Last Mile”**
   The part of the internet that connects homes and small businesses to the internet.

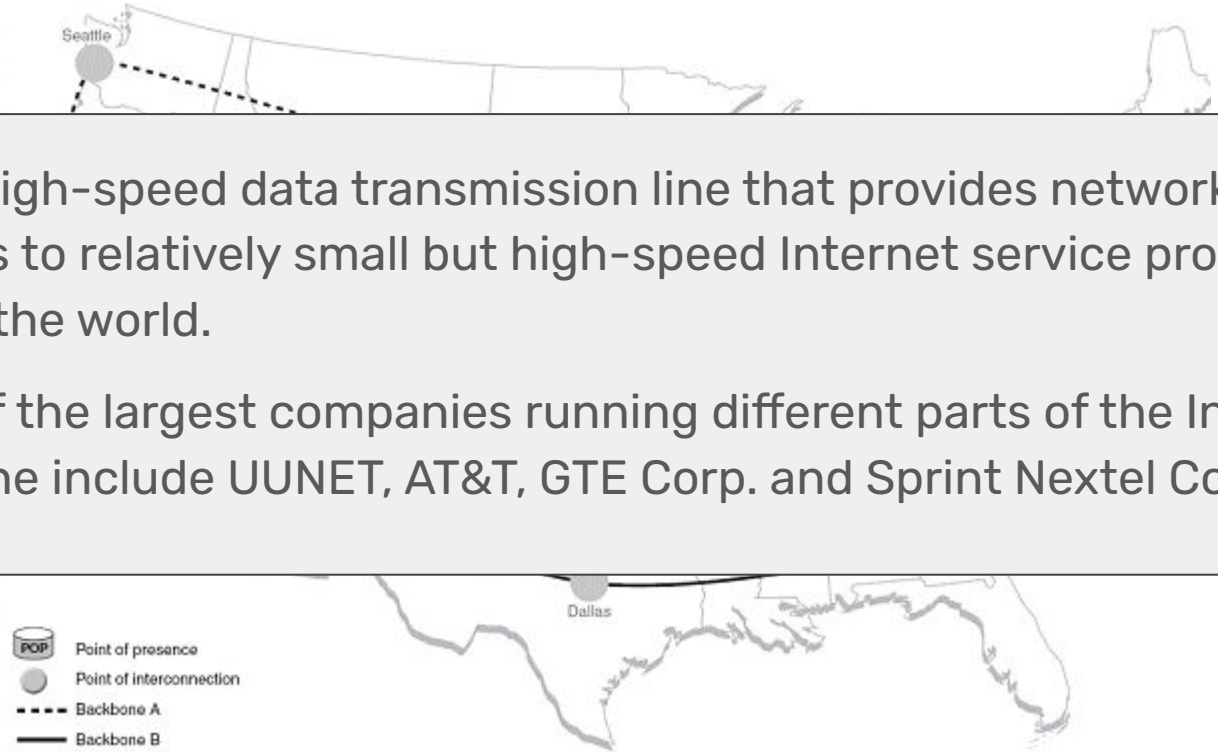
3. **Data Centers**
   Can be located anywhere in the world, but they are often located in remote areas where land and electricity are cheap.

Source: https://www.vox.com/2014/6/16/18076282/the-internet

Internet “Backbone”: Global



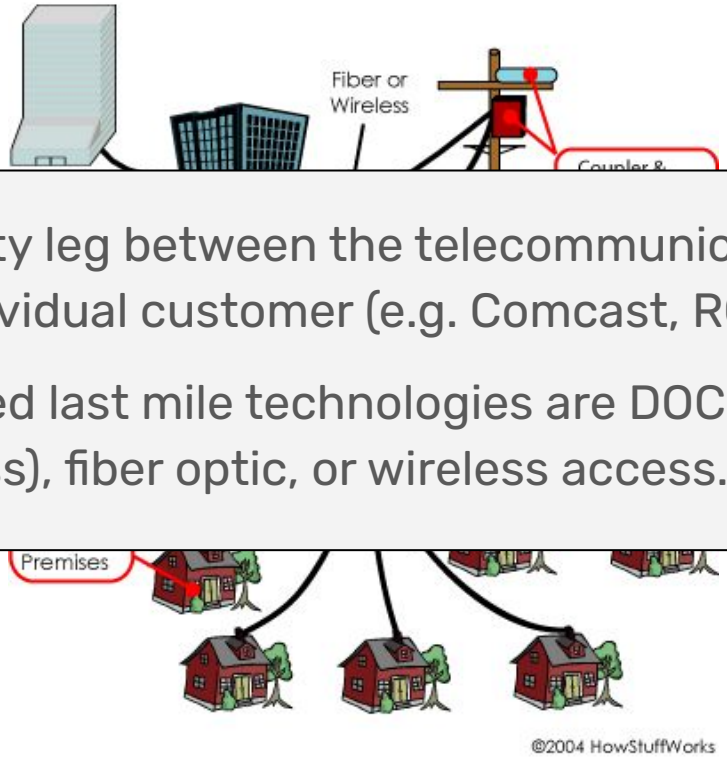
Internet “Backbone”: National



- A very high-speed data transmission line that provides networking facilities to relatively small but high-speed Internet service providers all around the world.
- Some of the largest companies running different parts of the Internet backbone include UUNET, AT&T, GTE Corp. and Sprint Nextel Corp.

Source: http://computer.howstuffworks.com/bpl4.htm

The “Last Mile”: Local



- The final connectivity leg between the telecommunication service provider and an individual customer (e.g. Comcast, RCN, etc.).
- The most widely used last mile technologies are DOCSIS (xDSL Cable and Cable Modem Access), fiber optic, or wireless access.

Source: http://computer.howstuffworks.com/bpl4.htm


Pictured: A Google Data Center

Who controls the Internet?

- No one runs the internet. It's organized as a decentralized network of networks.
- Thousands of companies, universities, governments, and other entities operate their own networks and exchange traffic with each other based on voluntary interconnection agreements
- There are technical standards committees (ICANN, IETF, etc.) that do meet and agree on rules of traffic exchange
- Governments can filter and block traffic (as can companies, schools, etc.)

Source: https://www.vox.com/2014/6/16/18076282/the-internet

What is Net Neutrality?

The principle that Internet service providers and governments should treat all data on the Internet the same, not discriminating or charging differentially by user, content, site, platform, application, type of attached equipment, or mode of communication.

Source: https://en.wikipedia.org/wiki/Net_neutrality

What is Net Neutrality?

The principle that Internet service providers and governments should treat all data on the Internet the same, not discriminating or charging differentially by user, content, site, platform, application, type of attached equipment, or mode of communication. In the US, this is controlled by the FCC.

- Repealed in 2015, championed by Ajit Pai
- Biden administration revisiting this determination

Questions to think about:

- What are the arguments FOR certain kinds of traffic getting prioritized?
- What are the arguments against it?

Article that provides some context:
https://www.cnet.com/tech/services-and-software/fcc-net-neutrality-repeal-ajit-pai-what-you-need-to-know/

Arguments for NN

Arguments against NN

Who can access your data (transmission & storage)?

What did the 2013 Snowden leaks reveal?

- Britain's “Tempora” taps fiber optic cables around the world
- NSA could hack into Google and Yahoo data centers w/o their knowledge
- Verizon had been providing the NSA w/all of its phone records
- NSA can request user data from companies, which they are compelled to deliver on by law (PRISM)
- NSA undermines encryption via backdoors and promoting the use of weaker algorithms

Sources:

- https://mashable.com/2014/06/05/edward-snowden-revelations/#u2nj85eX6PqT
- https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/1 (Worth a read!)

Let’s find out what you already know about the Web...

Some questions about the web...

1. What's the difference between the Internet and the Web?
2. Who invented the world wide web?
3. What is a URL?
4. What is HTTP?
5. What is a web server?
6. What is a browser?
7. What is an IP address?
8. What is a domain name?
9. What is a search engine?

What's the difference between the Internet and the Web?

The web is only one of many applications powered by the Internet. Other web applications include:

- Email
- Networked Television (such as cable)
- Peer-to-peer file sharing
- Telephony
- Skype, Google Hangout, GoToMeeting
- Many other networked applications (gaming engines, ITunes, etc.)

Who invented the Worldwide Web?

1989: Tim Berners-Lee, a British computer scientist, conceptualized and built the three foundational web technologies:

- **HTML**: HyperText Markup Language. The markup (formatting) language for the web.
- **URI**: Uniform Resource Identifier. A kind of “address” that is unique and used to identify to each resource on the web. It is also commonly called a URL.
- **HTTP**: Hypertext Transfer Protocol. Allows for the retrieval of linked resources from across the web.

Source: https://webfoundation.org/about/vision/history-of-the-web/

What is a URL?

A URL, or Uniform Resource Locator, is a reference to a web resource that specifies its location on a computer network and a mechanism for retrieving it.

A typical URL could have the form: http://www.example.com/index.html:

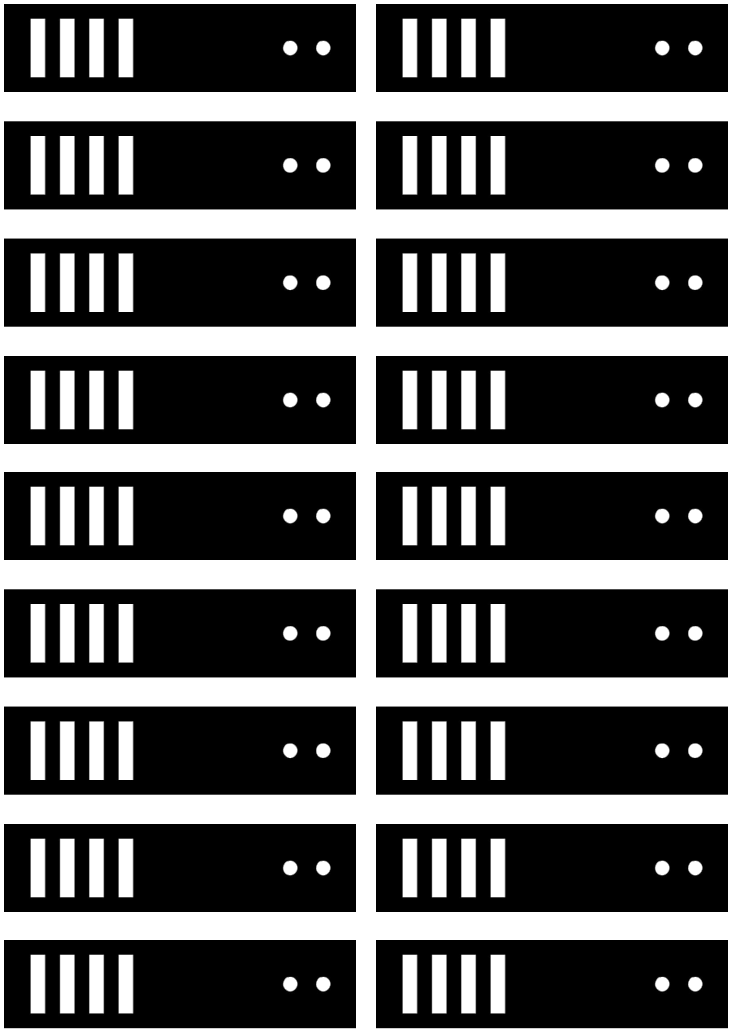
- **http**: indicates the protocol to be used to retrieve the document
- **www.example.com**: the hostname (domain name)
- **index.html**: the file name to be retrieved index.html

What is HTTP?

- Stands for Hypertext Transfer Protocol
- Refers to the rules that servers and browsers must follow to in order to transfer web files over the Internet.
- HTTP allows (authorized) users to create, update, or delete resources on a web server
- HTTPS adds encryption and security to the data transmission process

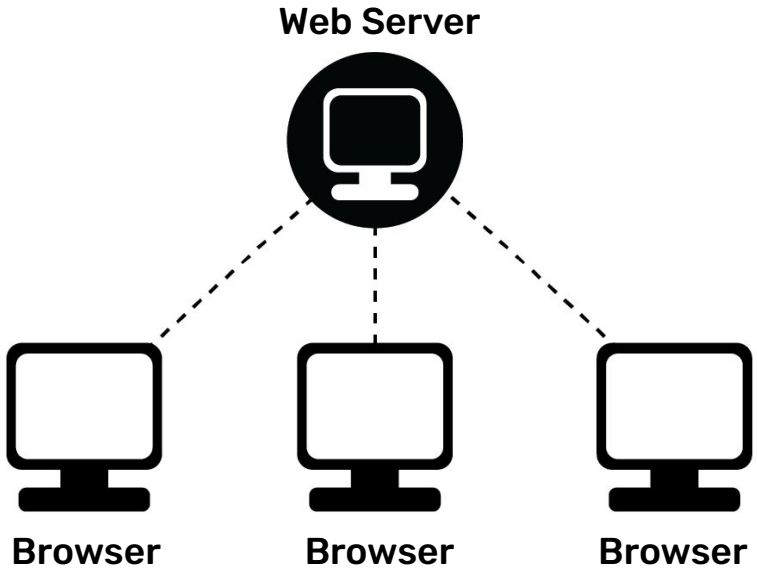
What is a web server?

- A computer or device on a network that other computers can access for information, communication, or computational services
- Any computer can become a web server if...
    a. It has been given a public address (IP address)
    b. It listens to requests over Port 80 (HTTP) or Port 443 (HTTPS)



What is a web browser?

- A software program that allows a user to locate, access, and display Web pages. Examples of web browsers include:
- Safari, Firefox, Chrome, Internet Explorer, and Opera



What is an IP address?

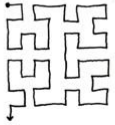
- An IP address is network address for your computer so the Internet knows where to send you emails, data and pictures of cats
- The IANA allocates IP address blocks to regional Internet registries. These are then divided into smaller sub-blocks and assigned to individuals and institutions
- We have already run out of IP addresses (e.g. 66.171.248.170), and so we're now in the process of transitioning to IPv6

MAP OF THE INTERNET

THE IPV4 SPACE, 2006

THIS CHART SHOWS THE IP ADDRESS SPACE ON A PLANE USING A FRACTAL MAPPING WHICH PRESERVES GROUPING -- ANY CONSECUTIVE STRING OF IPS WILL TRANSLATE TO A SINGLE COMPACT, CONTIGUOUS REGION ON THE MAP. EACH OF THE 256 NUMBERED BLOCKS REPRESENTS ONE /8 SUBNET (CONTAINING ALL IPS THAT START WITH THAT NUMBER). THE UPPER LEFT SECTION SHOWS THE BLOCKS SOLD DIRECTLY TO CORPORATIONS AND GOVERNMENTS IN THE 1990'S BEFORE THE RIRS TOOK OVER ALLOCATION.

0 1 14 15 16 19 →
3 2 13 12 17 18
4 7 8 11
5 6 9 10



= UNALLOCATED BLOCK

What is a domain name?

- A way of assigning human-readable names to IP addresses (which are difficult to remember)
- To “Purchase” a domain name (really you’re leasing it), you pay money to an authorized registrar (GoDaddy is the biggest one)
- The domain name system (DNS) is a hierarchical naming system (big dictionary) that keeps track of IP addresses and their associated domain names
- Makes it possible to assign domain names to groups of Internet resources and users, regardless of the entities' physical location

What is a search engine?

A system for organizing and retrieving Web pages. Search engines perform three basic tasks:

1. Crawling – where content is discovered
2. Indexing, where it is analysed and stored in huge databases; and
3. Retrieval, where a user query fetches a list of relevant pages and sorts them.

For Wednesday

On Wednesday, we're going to go over the rules of stylesheets and do some practice exercises!