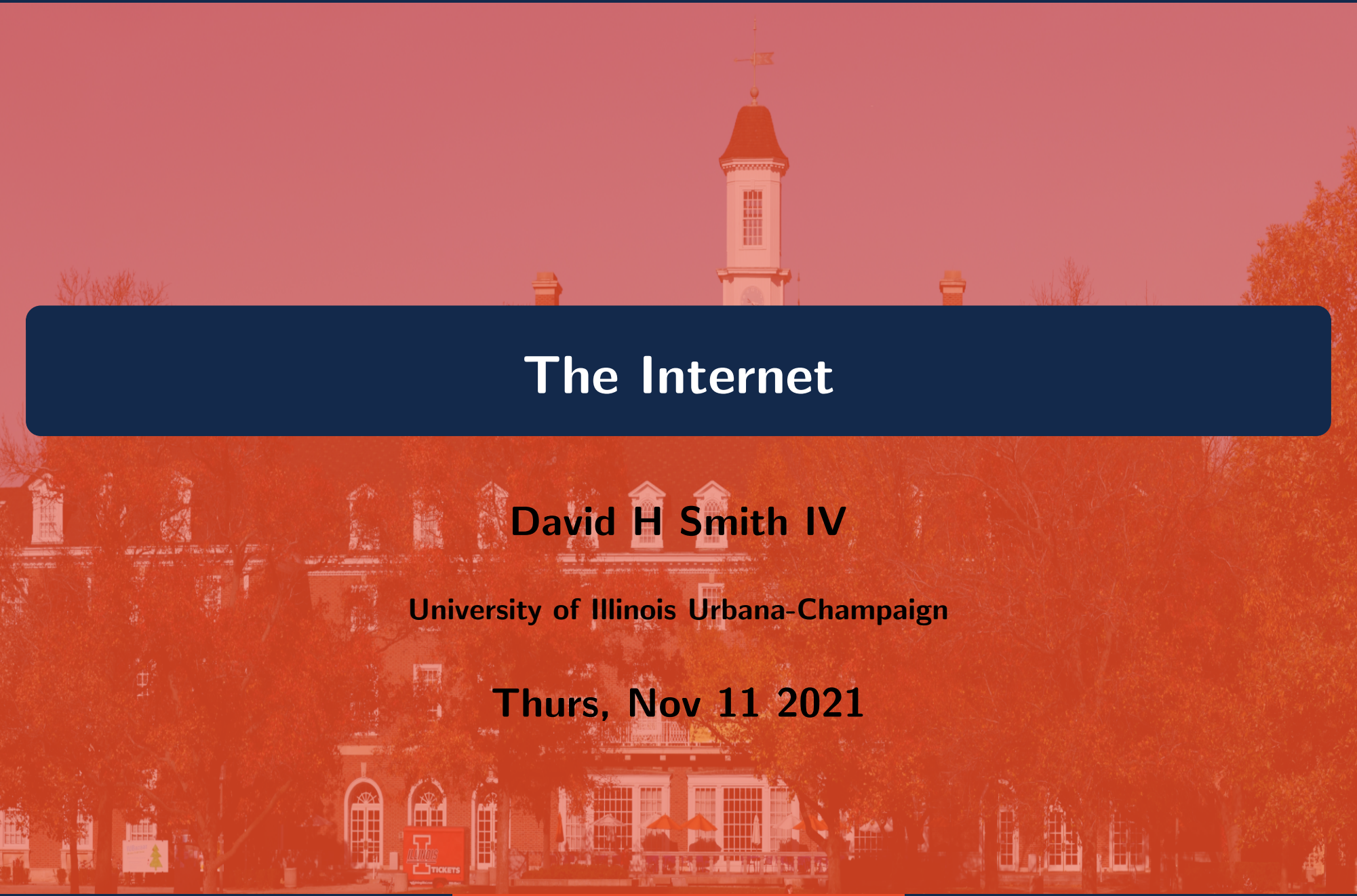


The Internet

David H Smith IV

University of Illinois Urbana-Champaign

Thurs, Nov 11 2021

Reminders

Reminders

Things that are due tomorrow:

- Homework 12
- Game of Life
- Participation 13p1
- Post-reading 13p1

Why we care

Where Data Comes From

1. Started with the data being hard-coded

Where Data Comes From

1. Started with the data being hard-coded
2. Then we got data from the user: `input()`

Where Data Comes From

1. Started with the data being hard-coded
2. Then we got data from the user: `input()`
3. Then we got it from files: `open(filepath)`.

Where Data Comes From

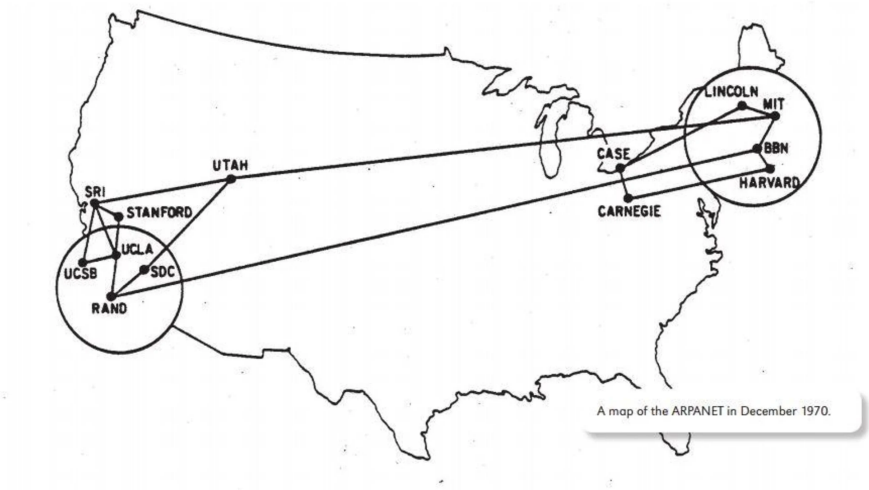
1. Started with the data being hard-coded
2. Then we got data from the user: `input()`
3. Then we got it from files: `open(filepath)`.
4. Now, we can get it from the internet.
   - Recall, websites are just collections of files and software on another person's computer (server):

```
dev_website
|---index.html
|___elements
    |--- about.html
    |--- cv.html
    |___ projects.html

```

History of the Internet

ARPANET



A map of the ARPANET in December 1970.

- Only a few users at government research facilities and universities
- Used for file transfer, not websites
- File Transfer Protocol (FTP, 1971) is still available and used today.

Tim Berners-Lee and the Internet

- TBL was a physicist at CERN (another scientist interested in sharing work).

Tim Berners-Lee and the Internet

- TBL was a physicist at CERN (another scientist interested in sharing work).
- Invented in the 1990s.

Tim Berners-Lee and the Internet

- TBL was a physicist at CERN (another scientist interested in sharing work).
- Invented in the 1990s.
- Composed of the following components:

Tim Berners-Lee and the Internet

- TBL was a physicist at CERN (another scientist interested in sharing work).
- Invented in the 1990s.
- Composed of the following components:
  1. HTML files containing links to other HTML files

Tim Berners-Lee and the Internet

- TBL was a physicist at CERN (another scientist interested in sharing work).
- Invented in the 1990s.
- Composed of the following components:
  1. HTML files containing links to other HTML files
     - Link: First proposed by Vannevar Bush in "As we may think".
     - Proposed a hypothetical machine that could let a user jump to different locations across multiple documents. Sound familiar?

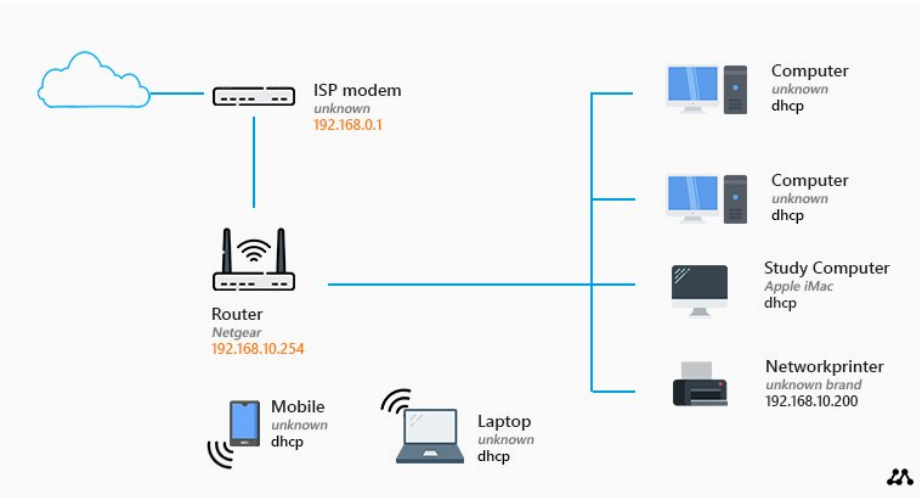
Tim Berners-Lee and the Internet

- TBL was a physicist at CERN (another scientist interested in sharing work).
- Invented in the 1990s.
- Composed of the following components:
  1. HTML files containing links to other HTML files
     - Link: First proposed by Vannevar Bush in "As we may think".
     - Proposed a hypothetical machine that could let a user jump to different locations across multiple documents. Sound familiar?
  2. A browser, which could read and render the files.

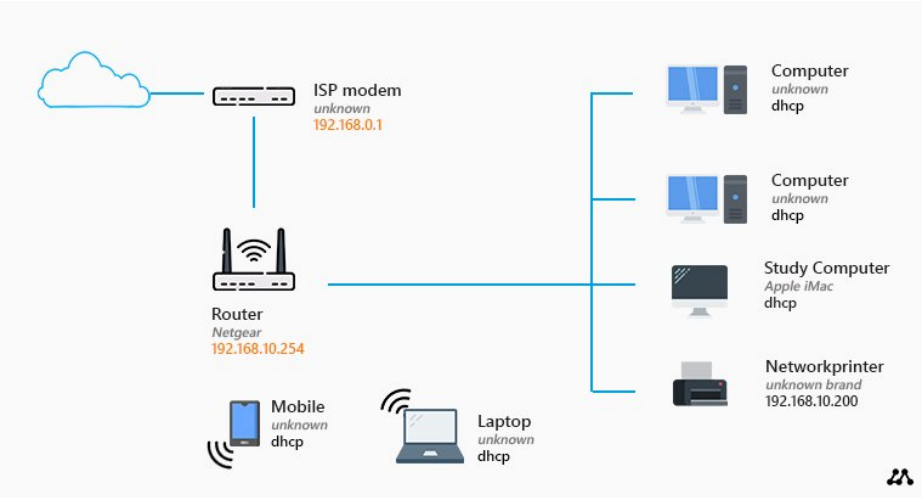
Tim Berners-Lee and the Internet

- TBL was a physicist at CERN (another scientist interested in sharing work).
- Invented in the 1990s.
- Composed of the following components:
  1. HTML files containing links to other HTML files
     - Link: First proposed by Vannevar Bush in "As we may think".
     - Proposed a hypothetical machine that could let a user jump to different locations across multiple documents. Sound familiar?
  2. A browser, which could read and render the files.
  3. A set of rules (protocol), for transferring these files (HTTP).

Networks

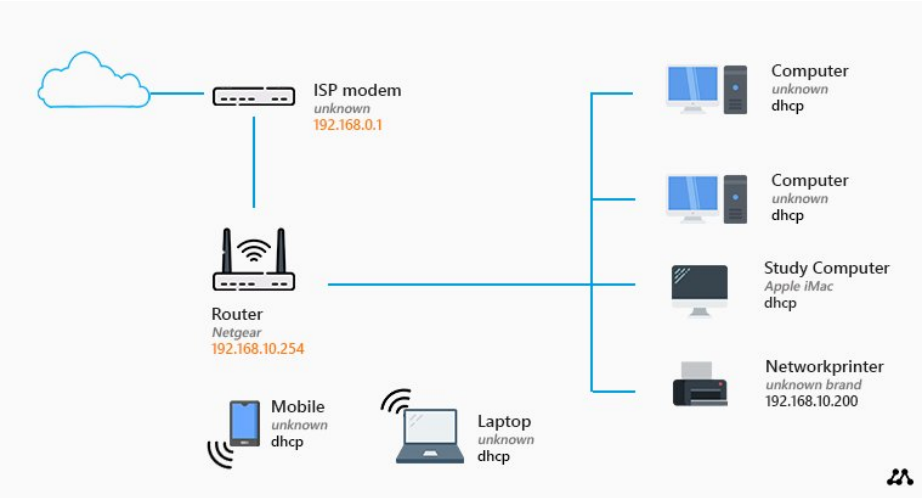

ISP modem
unknown
192.168.0.1
Router
Netgear
192.168.10.254
Mobile
unknown
dhcp
Laptop
unknown
dhcp
Computer
unknown
dhcp
Computer
unknown
dhcp
Study Computer
Apple iMac
dhcp
Networkprinter
unknown brand
192.168.10.200

Networks



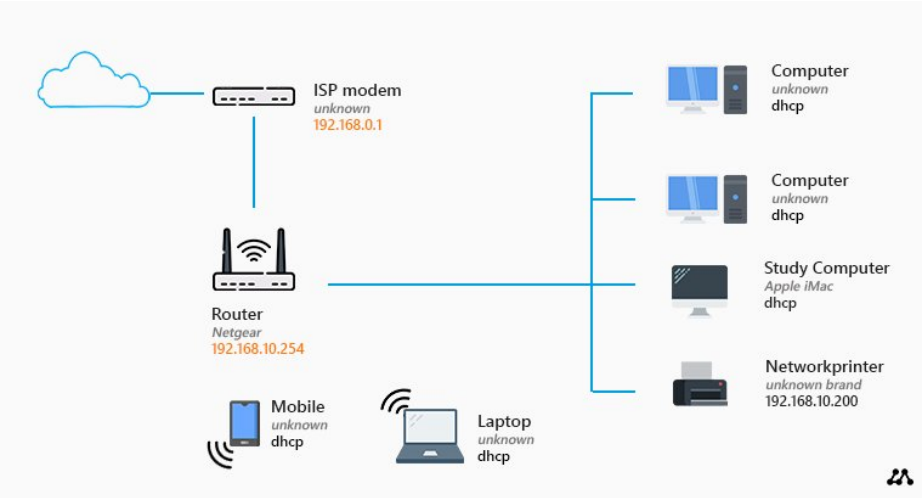
- Local Devices $\rightarrow$ Printers, TVs, Thermostats, Laptops,

Networks



- Local Devices $\rightarrow$ Printers, TVs, Thermostats, Laptops,
- Router $\rightarrow$ Route data/traffic between networks.

Networks



- Local Devices $\rightarrow$ Printers, TVs, Thermostats, Laptops,
- Router $\rightarrow$ Route data/traffic between networks.
- Modem $\rightarrow$ (1) Stands for "modulation-demodulation" (2) Converts between analog (waves) and digital (10101011101).

The early internet

- Dial-up Connections $\rightarrow$ An internet connection that uses telephone lines as the method of connecting to the ISP.

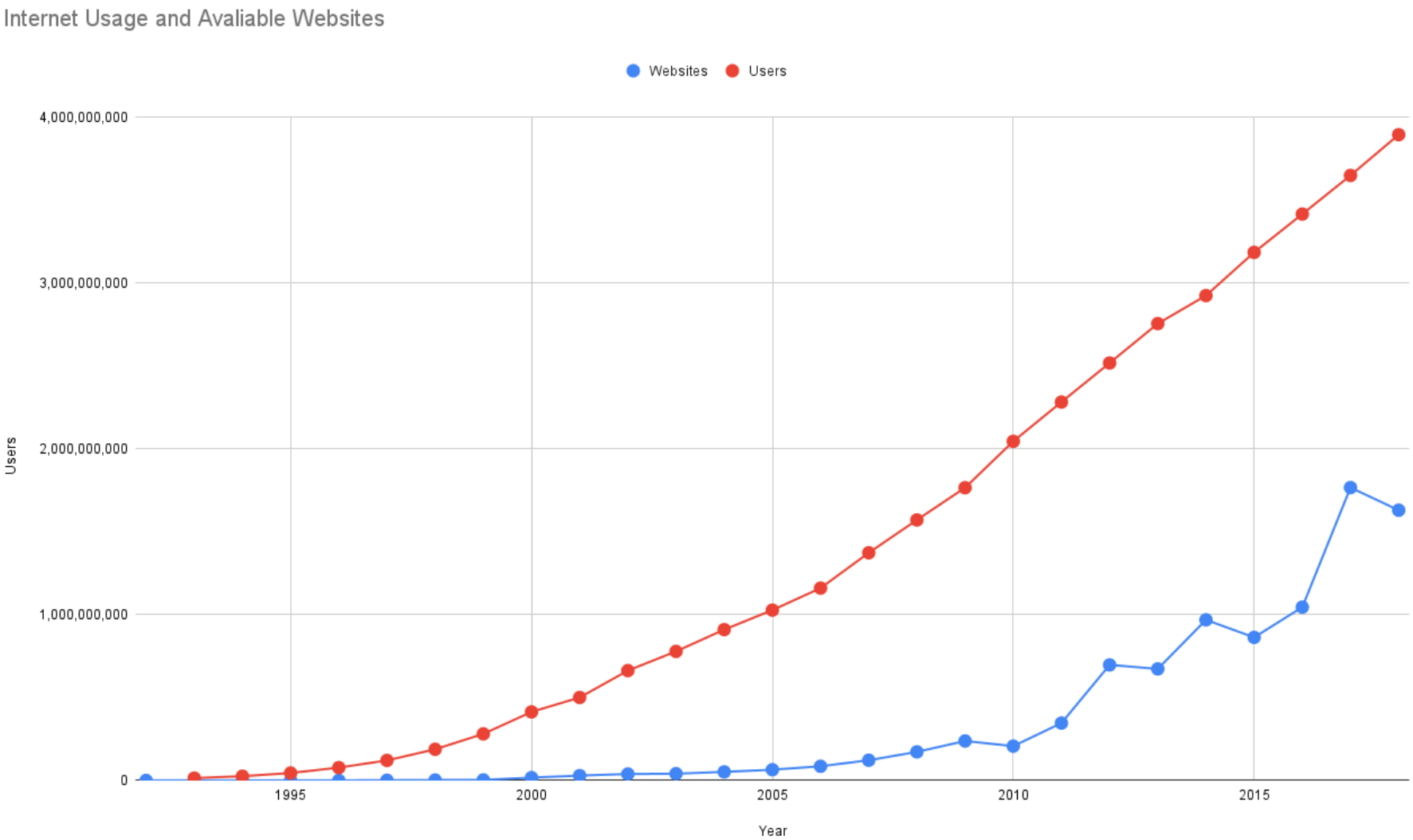
The early internet

- Dial-up Connections $\rightarrow$ An internet connection that uses telephone lines as the method of connecting to the ISP.
- Bulletin Board Systems (BBS) $\rightarrow$ An early example of a server that you could run on you computer in order to allow others to use software on that server.
  - These still exist and can be accessed using the "teletype network" (telnet) protocol.
  - Telnet Example: Telnet Star Wars

The early internet

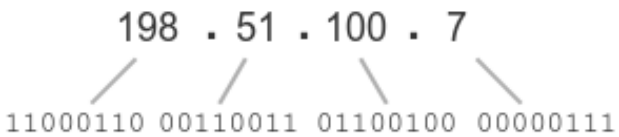
- Dial-up Connections $\rightarrow$ An internet connection that uses telephone lines as the method of connecting to the ISP.
- Bulletin Board Systems (BBS) $\rightarrow$ An early example of a server that you could run on you computer in order to allow others to use software on that server.
  - These still exist and can be accessed using the "teletype network" (telnet) protocol.
  - Telnet Example: Telnet Star Wars
- Internet Relay Chat (IRC) $\rightarrow$ An early instant messaging system that allowed someone to use a IRC client program to connect to a server and chat.

Internet Usage



IP Addresses, Domain Names, and URLS

IP Addresses



0 = 00000000

255 = 11111111

- Internet Protocol (IP) Address $\rightarrow$ The 32 bit (four 8-bit groups) address that specifies.
- Goes from 0.0.0.0-255.255.255.255
- Each domain name (e.g., www.google.com) has an IP address
- We use a Domain Name System (DNS) server to lookup the IP attached to a domain name

Domain Names

- Levels of Domains:
  1. Top-Level Domain (TLD): .com, .edu, .net, .org

Domain Names

- Levels of Domains:
  1. Top-Level Domain (TLD): .com, .edu, .net, .org
  2. Second-Level Domain: google, wikipedia, youtube

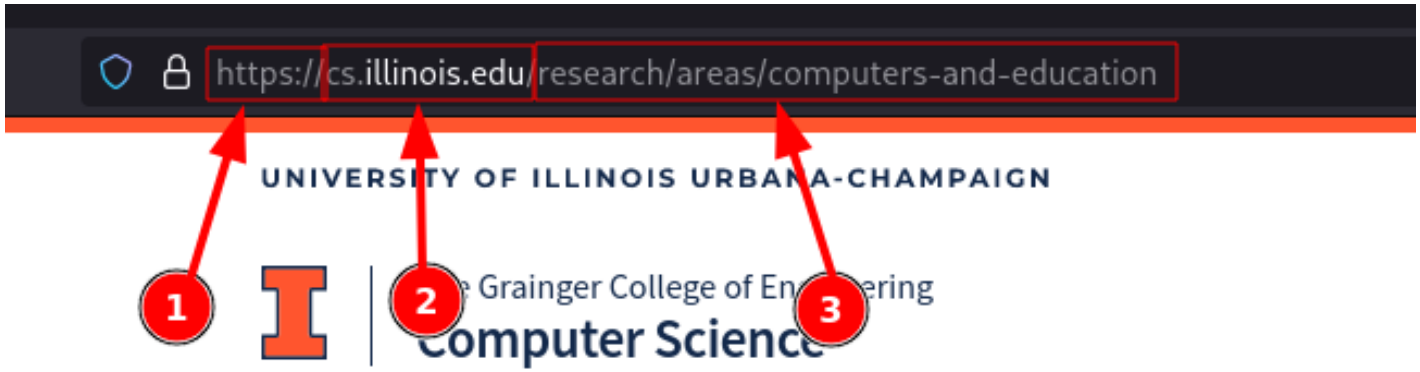
Domain Names

- Levels of Domains:
  1. Top-Level Domain (TLD): .com, .edu, .net, .org
  2. Second-Level Domain: google, wikipedia, youtube
  3. Sub-Domain: sub-domain(s) that fall under the second level domain.

Domain Names

- Levels of Domains:
  1. Top-Level Domain (TLD): .com, .edu, .net, .org
  2. Second-Level Domain: google, wikipedia, youtube
  3. Sub-Domain: sub-domain(s) that fall under the second level domain.
  4. For example, if the top and second-level domain is illinois.edu:
     - cs.illinois.edu
     - law.illinois.edu

URLS

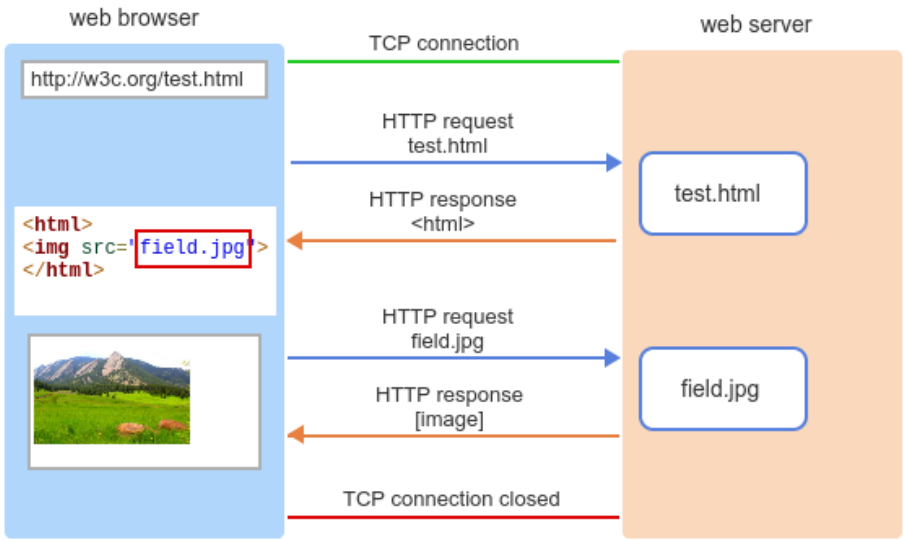


The Uniform Resource Locator (URL) is composed of the following parts:

1. **Protocol Scheme** → The protocol that is being used
2. **Hostname** → The complete domain name of the server you're trying to connect to.
3. **Path** → The path to the resource and (sometimes) query parameters.

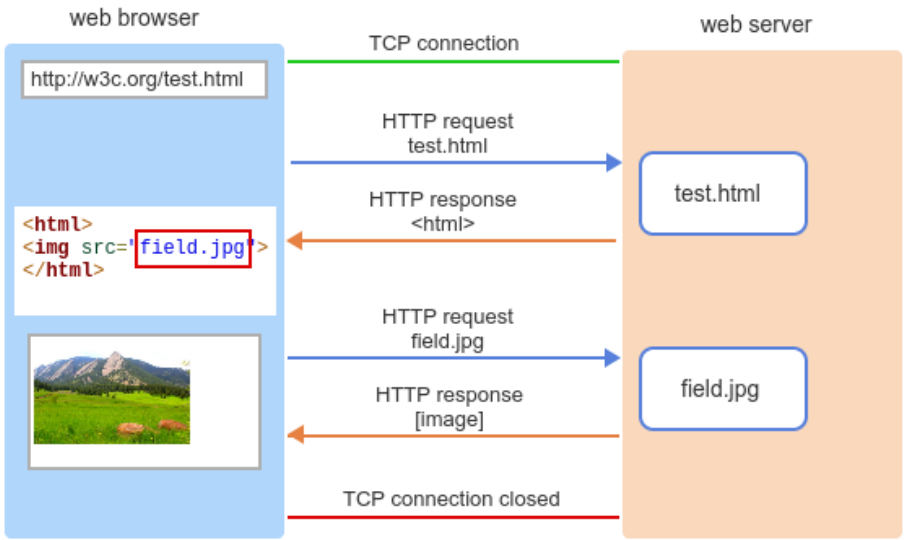
HTTP Protocol

HTTP Basics



- **GET** → Requests a resource from the server.
- **HEAD** → Same as get but without the response body.
- **POST** → Requests the server accept the request body as a new child resource or modify one.
- **PUT** → Requests the server accept the request body as a modification to an existing resource.
- **DELETE** → Requests the deletion of an existing resource.

HTTP Basics



The main ones we care about:

- **GET** $\rightarrow$ Resource requests.
- **POST** $\rightarrow$ Data modifying or creation requests.

Reading data from the internet

```
1 import requests
2
3 response = requests.get("https://www.google.com")
```

A. From a Python program
B. requests module: Given a URL, returns the document at that URL

Reading data from the internet

```
1 import requests
2
3 response = requests.get("https://www.google.com")
```

A. From a Python program
B. requests module: Given a URL, returns the document at that URL

Lets try this on PrairieLearn :D