

CSC435: Web Programming

Lecture 1: the internet and the web

Bei Xiao

American University

Jan 13, 2014

Course Info

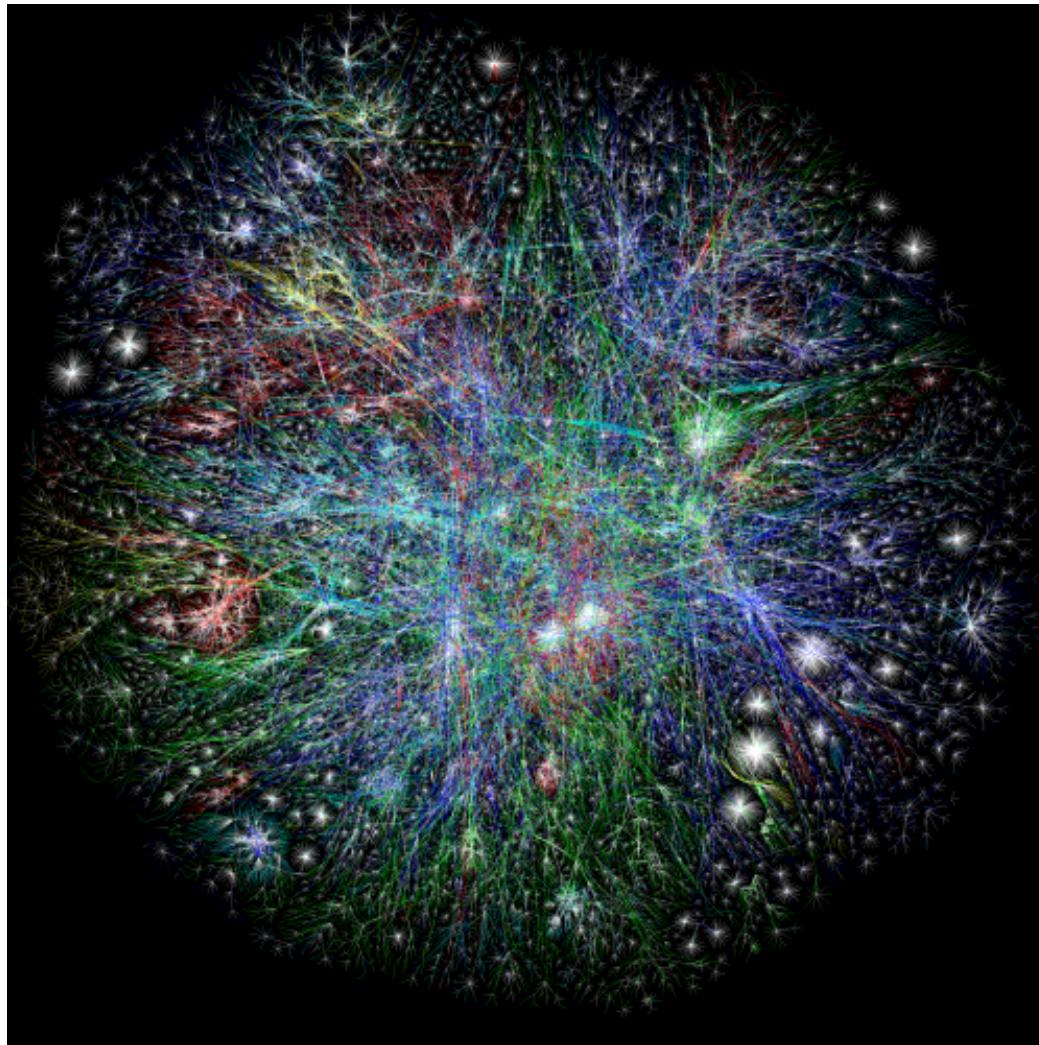
- Instructor: Prof. Xiao, bxiao@american.edu
- Office: SCAN 110
- Office hour: Wed 4-6pm, Thurs 11-Noon.
- Class time: Jan 13- May 5 (Mon/Thurs 4-5:15pm)
- Prerequisite: Intro to Programming (CSC 280)
- Course website:

[http://people.csail.mit.edu/beixiao/Teaching/
CSC435/CSC435_Spring2014.html](http://people.csail.mit.edu/beixiao/Teaching/CSC435/CSC435_Spring2014.html)

Overview

- What is the internet?
- How to access a website?
- What is an IP address
- What is URL?
- How to send request to server?
- What are client side technologies?
- What are server side technologies?

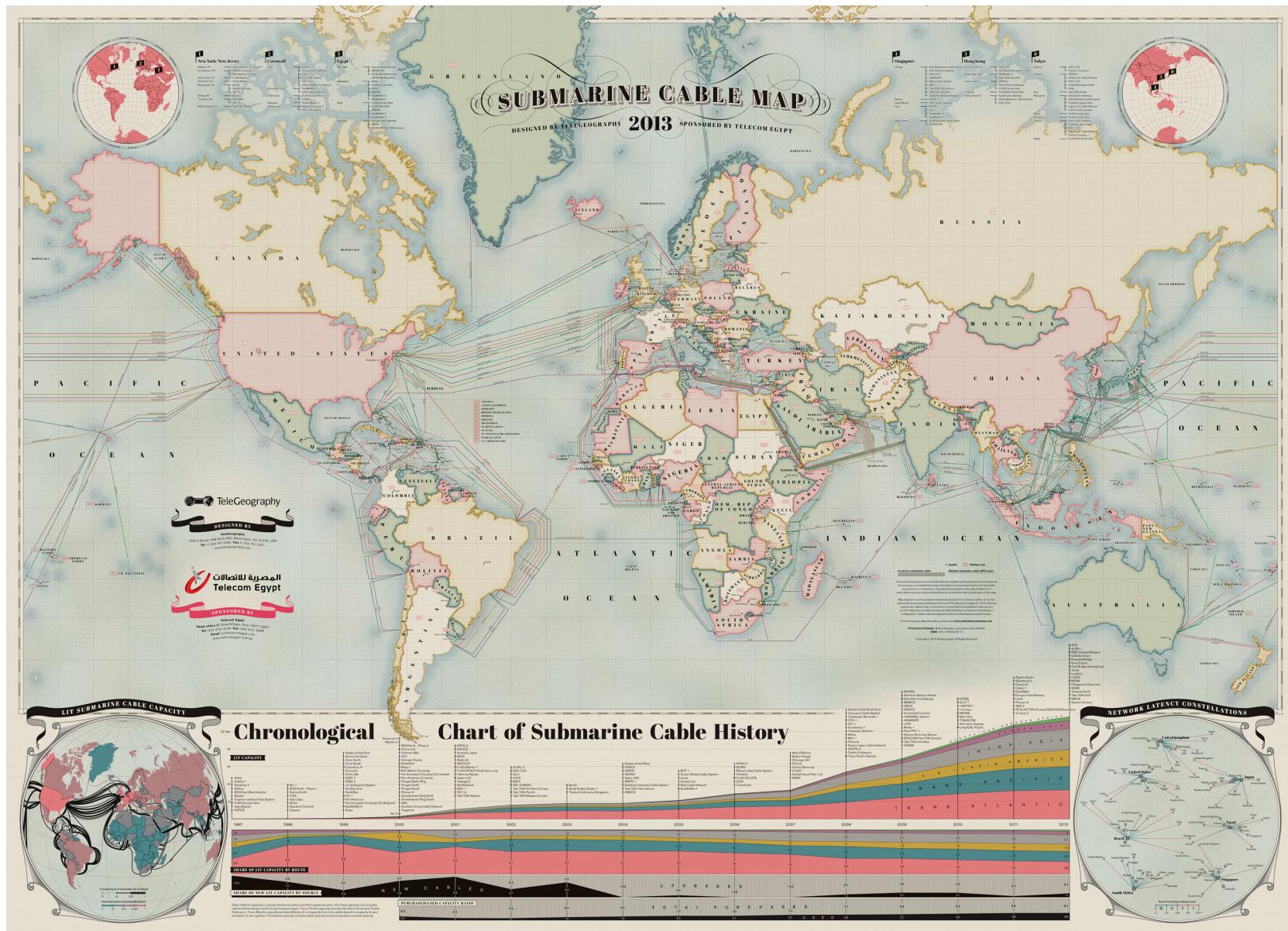
What is the internet?



http://en.wikipedia.org/wiki/Opt_e_Project

What is the internet?

- The internet is a physical network connecting millions of computers using the same protocols (TCP/IP) of sharing transmitting information. In reality, internet is network of small networks.



[http://www.telegeography.com/telecom-resources/
map-gallery/submarine-cable-map-2013/](http://www.telegeography.com/telecom-resources/map-gallery/submarine-cable-map-2013/)

Fun facts about the internet

- 1 out of 4 people have internet access.
- The internet reaches 50 million users in 5 years.
- More videos are uploaded to YouTube in the last 2 months than had ABC, NBC, and CBS all airing new content 24/7 since 1948.
- There are 121,013,927 unique domains.
- Most domains and pages are saved on servers and servers are saved on massive data centers.
- Average connection speed (mb/sec): 11.7 (south korea), 8.6 (Hongkong), 5.3 Netherland, 3.8 (US, 22nd place in the world !!)

What is the web?

- World wide web: a collection of interlinked multimedia documents that are stored on the Internet and accessed using a common protocol HTTP.

Data Center at CERN



What is the use of the web?

- Homepage
- Wikipedia
- Photo sharing websites
- Shopping
- Social networks
- Services
- Games

What are other internet-based applications?

- Email
- Telnet
- FTP
- Instant messaging
- File sharing
- Data storage

Quiz

- Is the internet the same as the web?

Quiz

- Is the internet the same as the web?
- If not, what is the key distinction?

Quiz

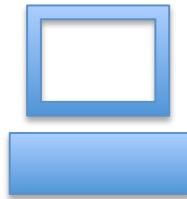
- Is the internet the same as the web?
- If not, what is the key distinction
- Answer: Internet is hardware; web is software along with data.

Accessing a website

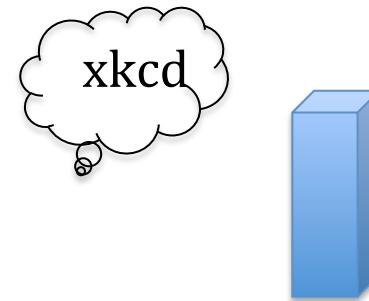
- Type in a URL in an address bar in your browser: www.xkcd.com

Accessing a website

- Type in a URL in an address bar in your browser: www.xkcd.com
- Web server: a machine that stores all the content such as images and videos



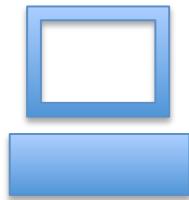
Client
Your
Browser



Web server

Accessing a website

- Computer communicate the server to tell which webpage you want to view
- Server process the data and return the web content the browser needs



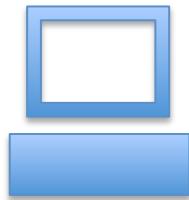
Client
Your
Browser



Web server

Accessing a website

- How does the browser know which server contains contents of the website xkcd?

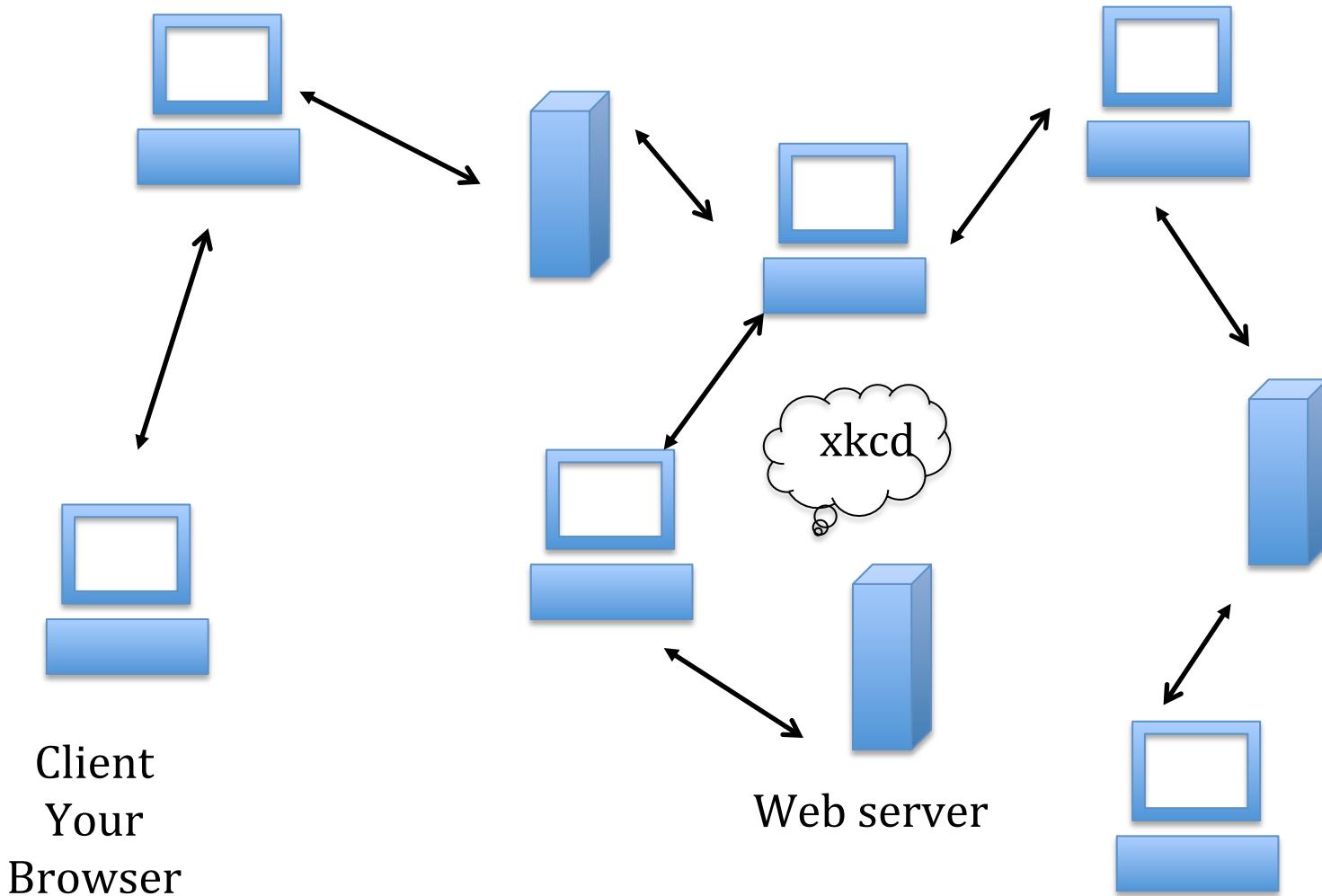


Client
Your
Browser

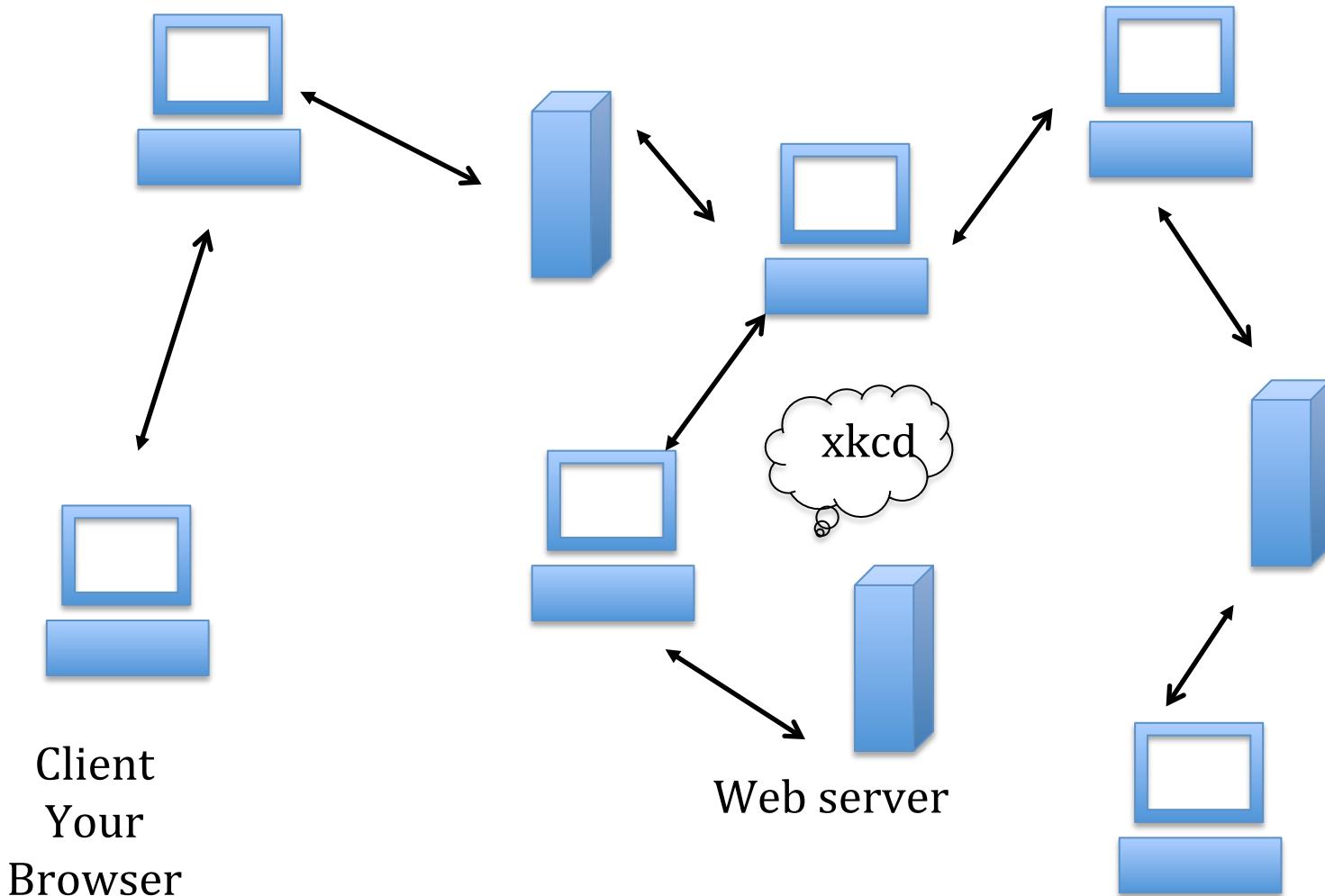


Web server

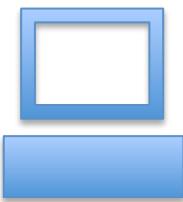
The Internet



Which server has the content of the requested website?



IP Address



Client
Your
Browser



Web server
IP: 72.26.203.99

IP Address

- The IP address give the location of the computer so that your computer can reach the server. Every computer at AU has the IP address from 147.
- It contains 4 numbers, ranging from 1 to 255.
- For more information, look at:
[Internet Protocol at wikipedia](#)

IP address versus URL

- IP address: the actual address of a targeted web server.
- URL refers to Uniform Resource Locator
- URL: human-readable pseudonym for a web-server.
- URL is mapped into an IP by a DNS server.

Question

- Why don't we directly type IP address into the browser?

Answer

- URLs allow users to change their web server without remembering the IP address
- Further reading:

Uniform Resource Locator (URL):

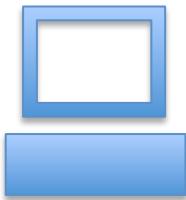
[http://en.wikipedia.org/wiki/
Uniform_resource_locator](http://en.wikipedia.org/wiki/Uniform_resource_locator)

Question

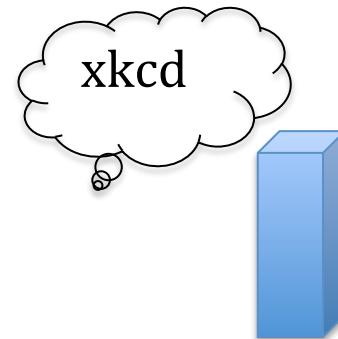
- You give the browser a URL, how does the browser know the IP address of the URL?

Question

- How does your computer know the IP address when you give only the URL?



Client
Your
Browser

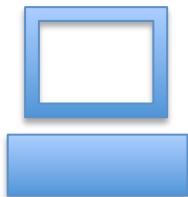


Web server
IP: 72.26.203.99

DNS (Domain Name System)

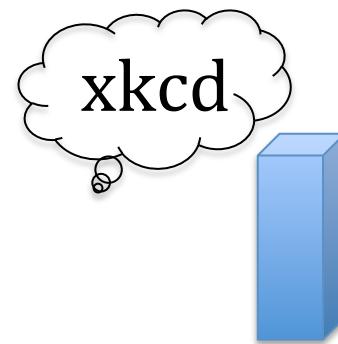
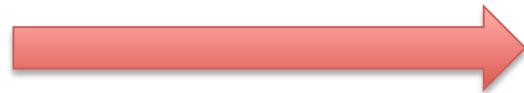


HTTP Request



Client
Your
Browser

HTTP Request: GET www.xkcd.com

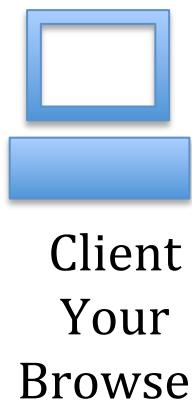


Web server
IP: 72.26.203.99

What is a HTTP Request?

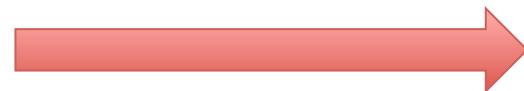
- HTTP: Hyper text Transfer Protocol
<http://djce.org.uk/dumprequest>
- Your browser sends a request to the server:
Get me the file index.html

HTTP Response

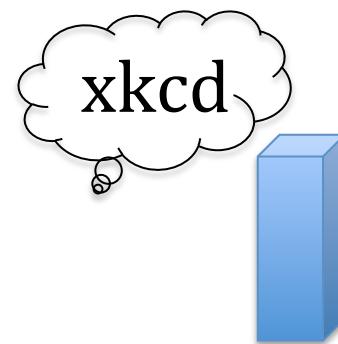


Client
Your
Browser

HTTP Request: GET www.xkcd.com



HTTP Response: web content HTML
file



Web server
IP: 72.26.203.99

Static versus Dynamic server-side processing

- Static website: www.xkdc.com

Every user views the SAME page. So the content can be just saved on the server. In much the same way that your personal computer saves files on its hard drive.

- Dynamic website: www.facebook.com

The specific content that a user sees (e.g. profile walls (which will change from time to time and from user to user)) must be generated in real-time processing.

HTTP Response Format

HTTP/1.1 200 OK

Status Line

Date: Thu, 20 May 2004 21:12:58 GMT

General Headers

Connection: close

Server: Apache/1.3.27

Response Headers

Accept-Ranges: bytes

Content-Type: text/html

Entity Headers

Content-Length: 170

Last-Modified: Tue, 18 May 2004 10:14:49 GMT

```
<html>
<head>
<title>Welcome to the Amazing Site!</title>
</head>
```

```
<body>
<p>This site is under construction. Please come
back later. Sorry!</p>
</body>
</html>
```

Message Body

Review

- Browser map an URL into IP through DNS look up
- Send a message to web server to request to view certain file: HTTP request
- The webserver then do some server side processing and then send back the files back to the browser: HTTP responses. Fetch the files from a data server or generate the files on the fly.
- The browser decide how to display the files on the webpages.

Client-side technologies

The following types of files are kind that the server can send to your computer so your browser can render the web-page:

- HTML
- CSS
- JavaScript

Client-side: HTML

- HTML describes the content of a webpage
- Headers
- Text
- Images and Videos
- Hyperlinks

Client-side: HTML

- HTML describes the content of a webpage

```
1  <!DOCTYPE html>
2  <html>
3  <head>
4
5  </head>
6
7  <body>
8
9  <h1>CSS example!</h1>
10 <p>This is a paragraph.</p>
11
12 </body>
13 </html>
```

Client-side: CSS

- Style HTML elements
- Font-color
- Background-color
- Size
- Positioning

Demo: a webpage with and without CSS

JavaScript

- Adds Interactivity to web-pages
- Submit forms
- Button Presses
- React to mouse movement
- Change HTML content
- Example: <http://3d.xkcd.com/?teapot>

Quiz

- What type of files can be found on the web:
 1. Music
 2. Images
 3. HTML
 4. Plain texts
 5. Videos

Server-side technology: PHP

- Used to run scripts on the server, in response to HTTP requests.
- Collect form data
- Generate dynamic page content
- Handle user sessions and credentials
- Interacts with databases

Server-side technology: Database

- Database: an organized collection of data that is maintained on the server
 - Account information and passwords
 - User-generated content (blog posts, comments, clicks on facebook, etc.)
- MySQL is particular database that is commonly used with PHP.
- Queries are requests to the database for information, e.g. last five comments on a blog.
- MySQL process the queries.

PHP and MySQL

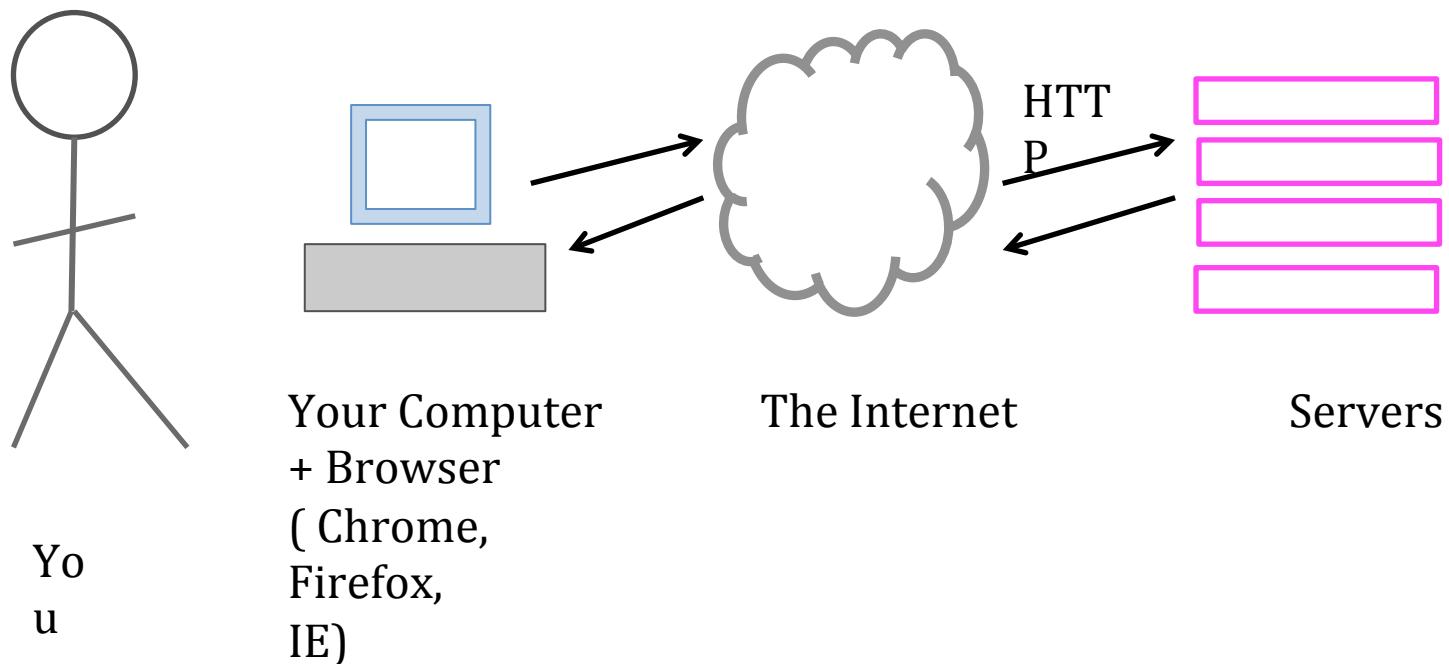
- Often work together
- PHP make queries to MySQL
- MySQL provides dynamic content
- PHP uses the content to generate a dynamic webpage.

Additional technologies

- jQuery: JavaScript library. Very convenient syntax.
- Ajax, a group of jQuery allows you to communicate with the server without loading the pages

Example: Microsoft bing search

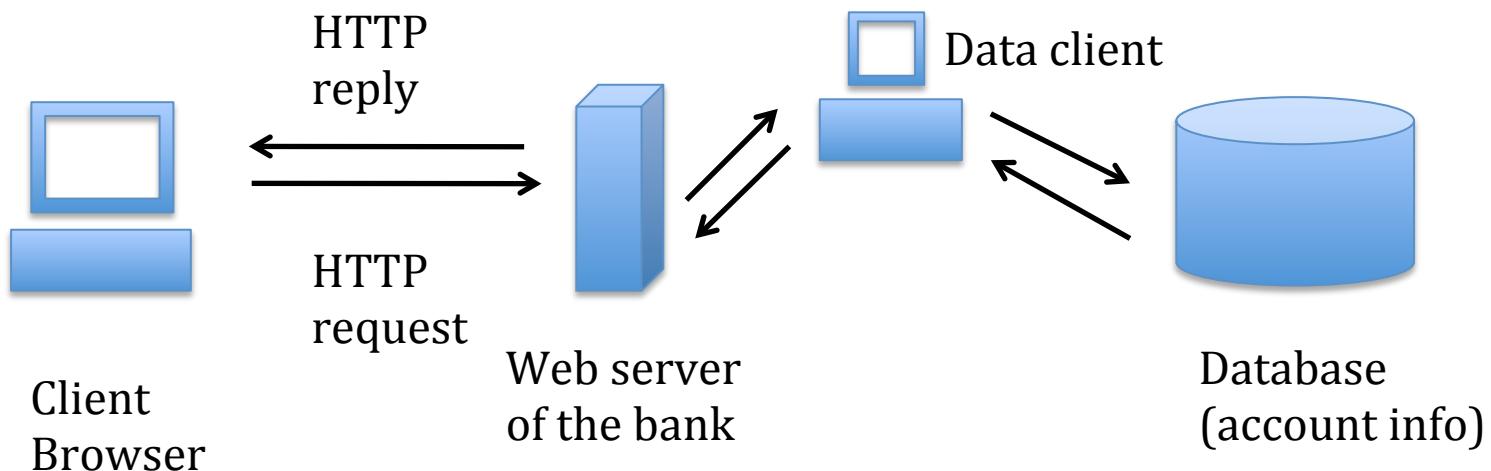
Review: Major pieces of the internet



Example: online banking

- A client access to a web browser (**the client**): www.bankofamerica.com
- The login credentials are stored in the **database**
- The **web server** access the **database server** (e.g. MySQL)
- The database server might fetch financial transaction records from another database server.
- An **application server** interprets the returned data by following the bank's business logic and return to the web server.
- The web server sends the result back to the browser, which interprets the data.

Checking account balance:



We will learn

Before Mid-term

- HTML
- CSS
- JavaScript

After Mid-term

- CGI
- PHP
- MySQL
- Ajax (if time allows)
- Design philosophies

Quiz

Which of the following process the request to return web content:

1. A DNS server
2. A Web server
3. Your browser
4. A particular web server
5. All of above

Answer

Which of the following process the request to return web content:

1. A DNS server
2. A Web server
3. Your browser
4. A particular web server
5. All of above

Homework

Install one of the following editors:

<http://www.sublimetext.com/2>

[http://www.barebones.com/products/
textwrangler/](http://www.barebones.com/products/textwrangler/)

<http://macromates.com/> (Mac OS)

Create an account at

<http://www.codecademy.com/>

Further Readings

URL: http://en.wikipedia.org/wiki/Uniform_resource_locator

HTTP Request:

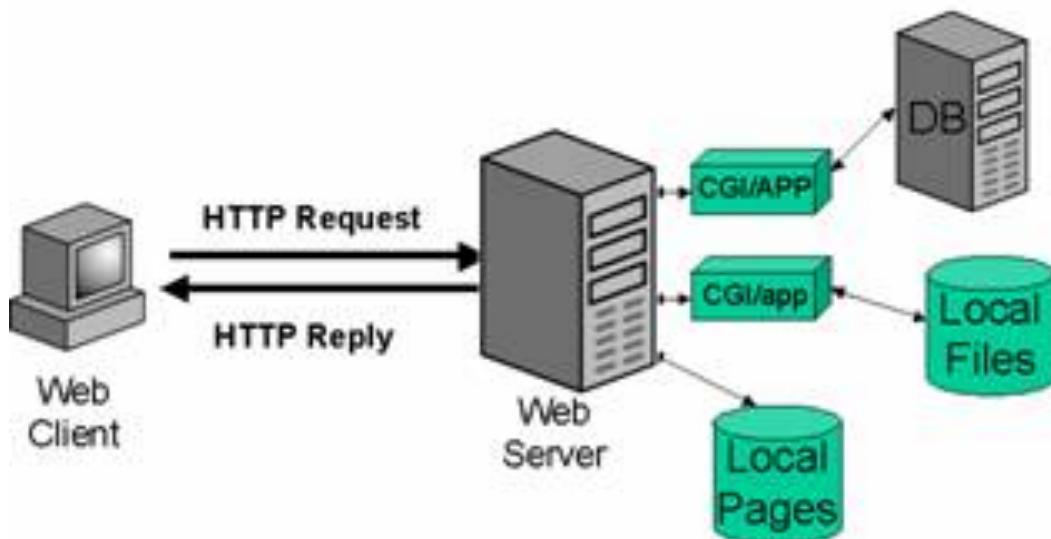
<http://djce.org.uk/dumprequest>

HTTP Response:

<http://www.tcpipguide.com/free/>
[t HTTPResponseMessageFormat.htm](#)

Client-Server Architecture

Web server Architecture



<http://www2.sans.org/security-resources/malwarefaq/guestbook.php>

URL

- URL: Uniform Resource Locator

`http://www.nytimes.com/pages/world/index.html`



- Query parameters

`http://www.nytimes.com/pages/world/index.html?`
`module=HPMiniNav&contentCollection=World&pgtype=Homepage&`
`region=TopBar&action=click&t=qry69`

Quiz

Given the following URL, identify different components of the URL

[http://www.lumosity.com/landing_pages/
431?gclid=CI2Wntii8LsCFUpn0godR1UA5A](http://www.lumosity.com/landing_pages/431?gclid=CI2Wntii8LsCFUpn0godR1UA5A)

1. Protocol
2. Host
3. Path
4. Query

Quiz

Given the protocol https, the host
www.washingtonpost.com, the path /
politics/

What is the URL?

HTTP: how browser talks to the server

- HyperText Transfer Protocol
- HTTP Request methods: Post and Get
- Get: get a document from the server

`http://www.example.com/foo`

Get/ foo HTTP/1.1

Method, Path, Version

HTTP: Get method

- Example: Amazon search for French Press returns the following URL. What is the request line to GET this URL using HTTP/1.1?

[http://www.amazon.com/s/
ref=nb_sb_noss_2?url=search-alias
%3Daps&field-keywords=frensh%20press](http://www.amazon.com/s/ref=nb_sb_noss_2?url=search-alias%3Daps&field-keywords=frensh%20press)

HTTP Responses

- Example: Amazon search for French Press returns the following URL. What is the request line to GET this URL using HTTP/1.1?

[http://www.amazon.com/s/
ref=nb_sb_noss_2?url=search-alias
%3Daps&field-keywords=frensh%20press](http://www.amazon.com/s/ref=nb_sb_noss_2?url=search-alias%3Daps&field-keywords=frensh%20press)