

# CSC435: Web Programming

## Lecture 1: the internet and the web

Bei Xiao

American University

Jan 12, 2016

# Course Info

- Instructor: Prof. Xiao, [bxiao@american.edu](mailto:bxiao@american.edu)
- Office: SCAN 110
- Office hour: Wed 4-6pm.
- Class time: Jan 13- May 5 (Mon/Thurs 5:15-6:30pm)
- Prerequisite: Intro to Programming (CSC 280)
- Course website:

# Course objectives

- Fundamental technologies behind web applications, **not just building websites.**
- Technology for designing, developing, and evaluating web apps.
- Learn a set of tools to develop user-friendly dynamic web apps.

# Specific tools

- HyperText Markup Language (HTML) for authoring web pages
- Cascading Style Sheets (CSS) for supplying stylistic information to web pages
- JavaScript for creating interactive web pages
- PHP Hypertext Processor for generating dynamic pages on a web server
- Python CGI for server-side scripting (if time allows).
- Asynchronous JavaScript and XML (Ajax) for enhanced web interaction and applications
- Structured Query Language (MySQL) for interacting with databases

# Take-home work

Please read the course syllabus

Please download Ammps:

Code academy tutorial on HTML

# Grading

- 50% Assignments (five projects)
- 15% Mid-term exam
- 15% Final Project (Open topic)
- 10% Class attendance (in-class performances and attendance)
- 10% in-class quiz.

# Late Policy

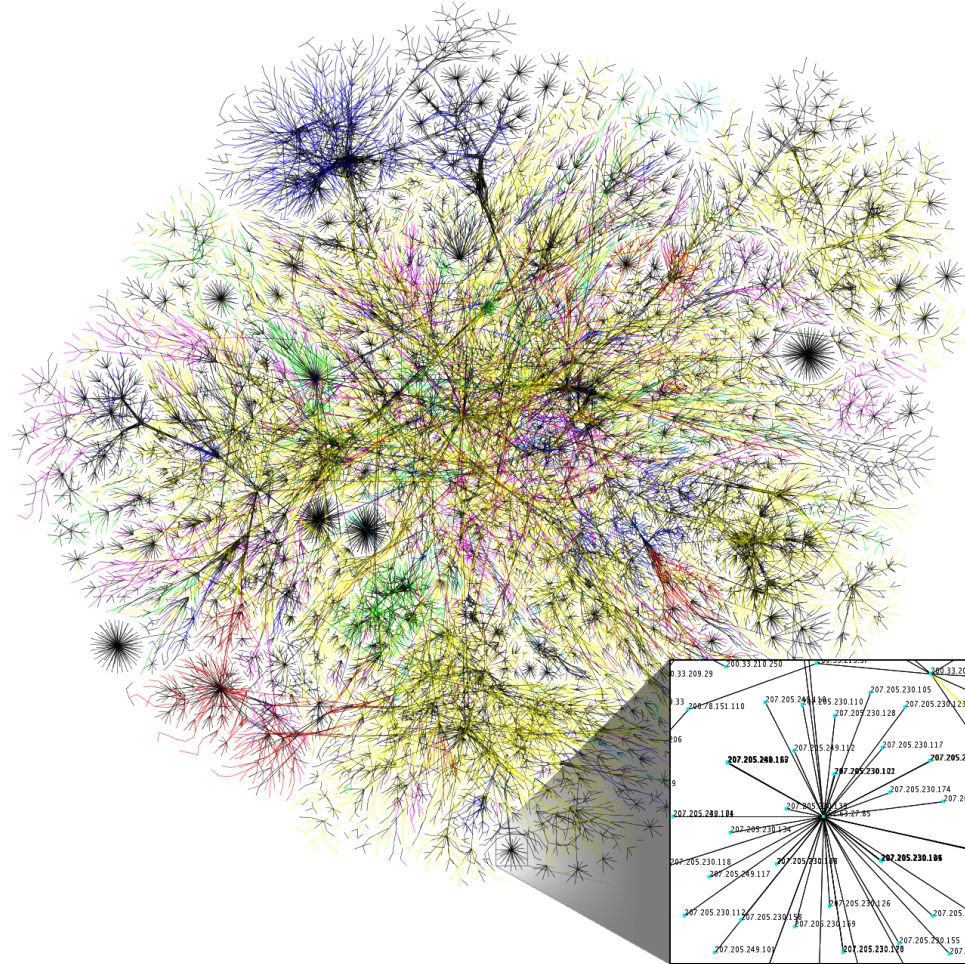
- No late work is accepted. All homework have to be turned in by due date.

# 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?



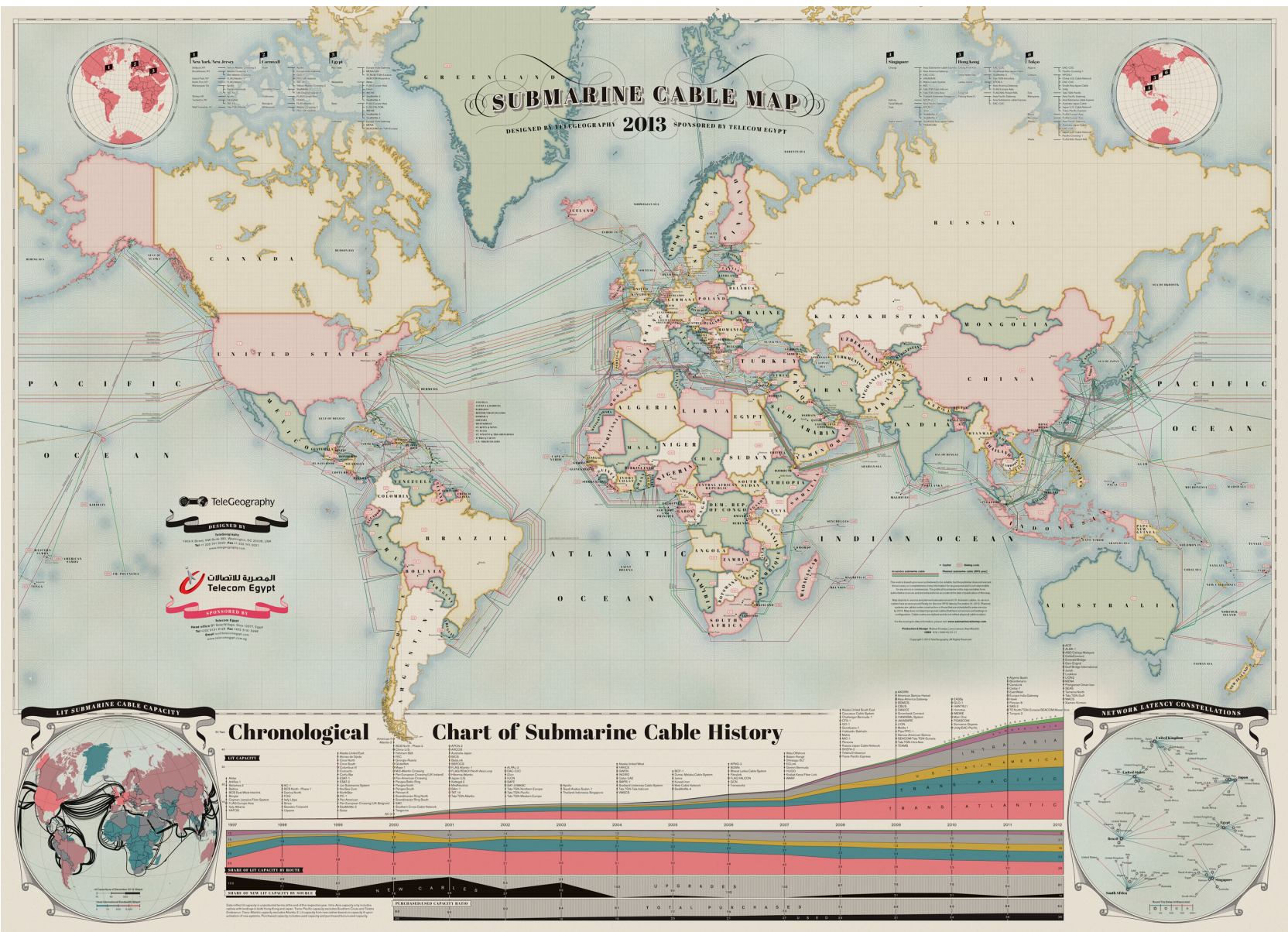
[http://en.wikipedia.org/wiki/Opte\\_Project](http://en.wikipedia.org/wiki/Opte_Project)

# 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.

# The internet?

- Wikipedia:  
<https://en.wikipedia.org/wiki/Internet>
- A connection of computer networks the internet protocol (IP)
- Layers of communication protocols: IP-> TCP/UDP-> HTTP/POP/FTP
- What is the difference between Internet and the web?



<http://www.telegeography.com/telecom-resources/map-gallery/submarine-cable-map-2013/>

# Brief history of the internet

- Began as US Department of Defense network called APPANET (1960s-70s).
- Initial services: electronic mail, file transfer
- Opened to commercial interests in late 80s.
- WWW created in 1989-91 by [Tim Berners-Lee](#)
- Popular web browsers released: Netscape 1994, IE 1995, Mozilla 1998, Safari, 2003, Firefox, 2004, Chrome 2011.
- Amazon.com opens in 1995; Google January 1999.
- Facebook opened in 2004, Twitter in 2006.

# 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, 22<sup>nd</sup> place in the world !!)

# Question

- Who “runs” the internet? Who is responsible for overseeing it?

# People and organizations

- Internet Engineering Task Force (IETF): internet protocol standards
- Internet Corporation for Assigned Names and Numbers (ICANN): decides top-level [domain name](#).
- World Wide Web consortium (W3C): web standards.



# 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
- News Media
- Games
- Be creative to create your own

# 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.

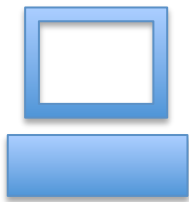
# Accessing a website

- Type in a URL in an address bar in your browser: [www.xkcd.com](http://www.xkcd.com)



# Accessing a website

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



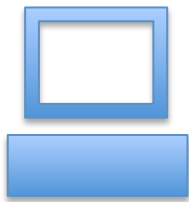
Client  
Your  
Browser



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



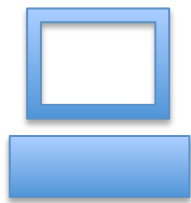
Client  
Your  
Browser



Web server

# Accessing a website

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

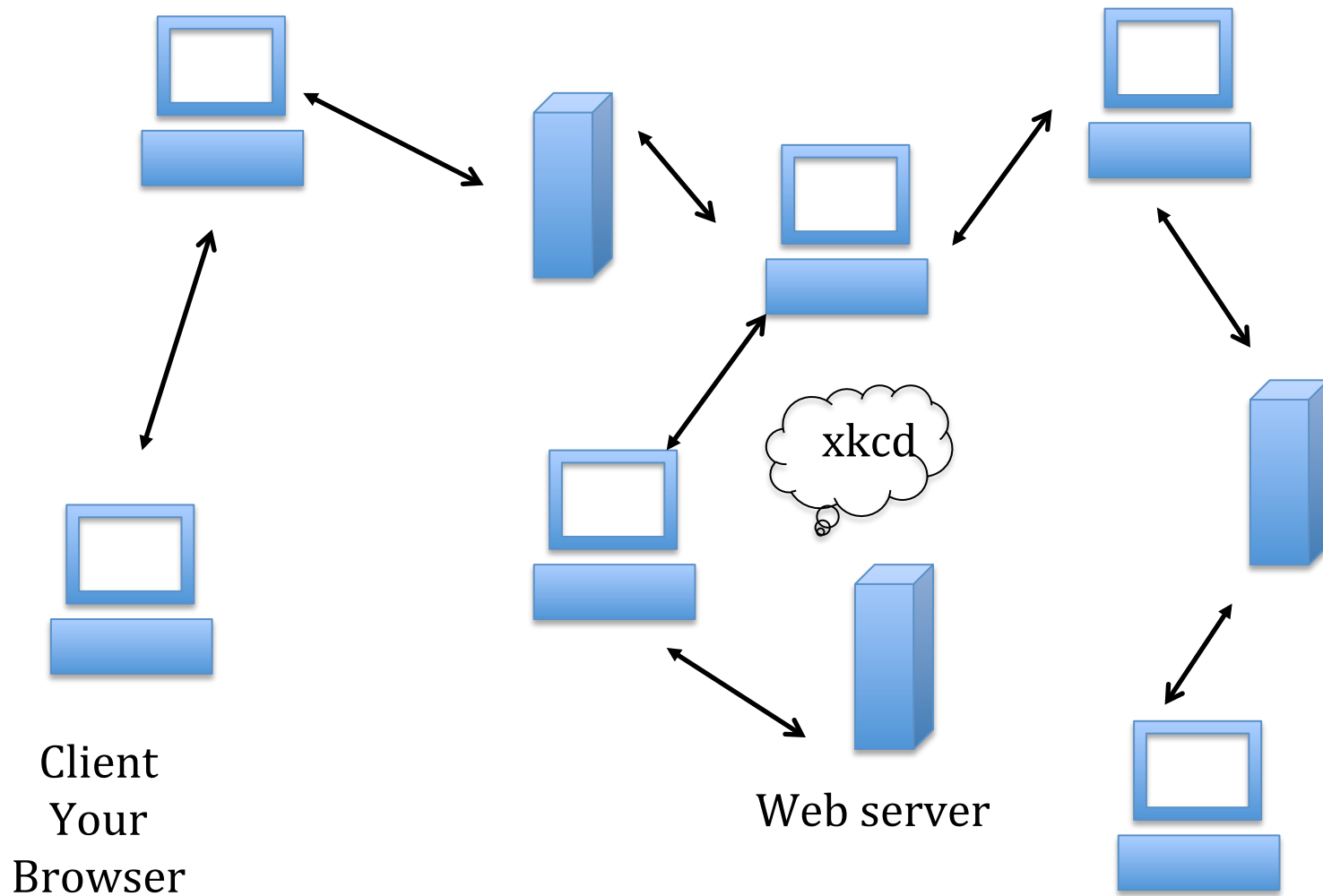


Client  
Your  
Browser

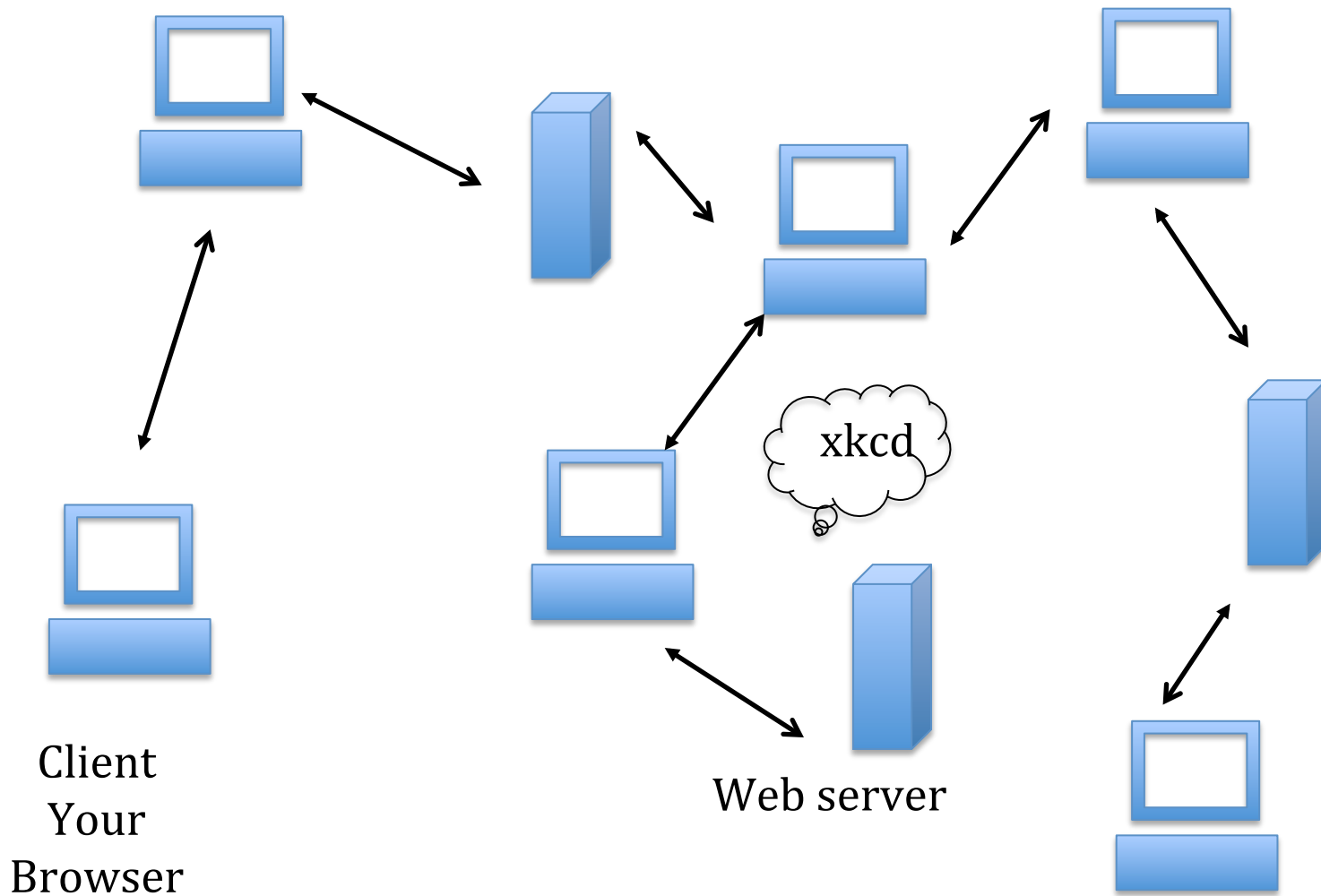


Web server

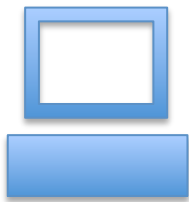
# The Internet



# Which server has the content of the requested website?



# IP Address



Client  
Your  
Browser



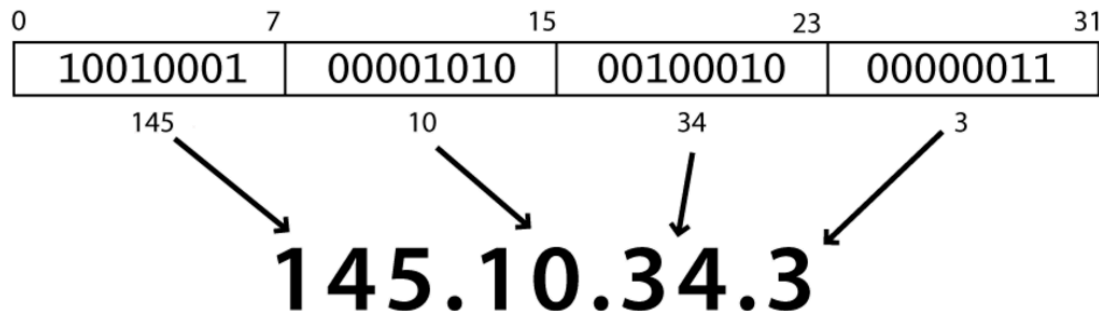
Web server  
IP: 72.26.203.99

# IP (Internet Protocol) 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 (Internet Protocol)

- A simple protocol for attempting to send data between two computers.
- Each device has a 32-bit IP address written for four 8-bit numbers (0-255)
- Find out your local IP address
  - In a terminal, type ipconfig(Windows) and ifconfig (Mac)
- Find out your internet IP address: [whatismyip.com](http://whatismyip.com)





# 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):

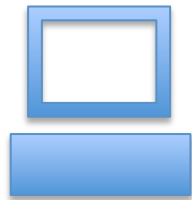
[http://en.wikipedia.org/wiki/  
Uniform\\_resource\\_locator](http://en.wikipedia.org/wiki/Uniform_resource_locator)

# Question

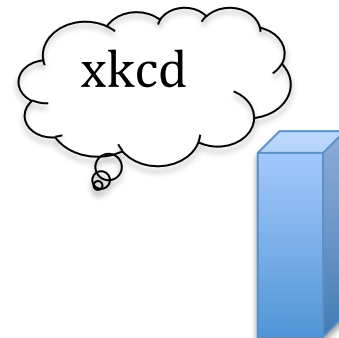
- You give the browser a URL, how does the browser know the IP address of the URL?

# Question

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



Client  
Your  
Browser

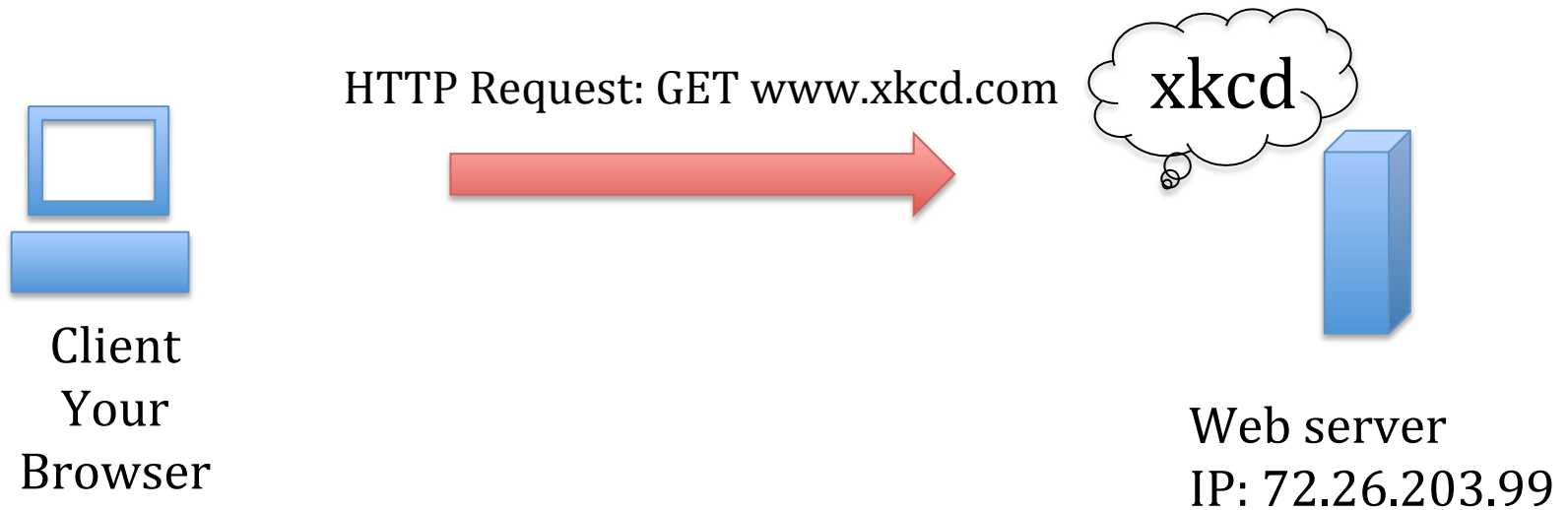


Web server  
IP: 72.26.203.99

# DNS (Domain Name System)



# HTTP Request



# 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

Read more here:

[http://www.jmarshall.com/easy/http/  
#whatis](http://www.jmarshall.com/easy/http/#whatis)



# HTTP Response



# Static versus Dynamic server-side processing

- Static website: [www.xkdc.com](http://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](http://www.facebook.com)

The specific content that a user sees the 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   | <b>Status Line</b>      |
| Date: Thu, 20 May 2004 21:12:58 GMT<br>Connection: close  | <b>General Headers</b>  |
| Server: Apache/1.3.27<br>Accept-Ranges: bytes   | <b>Response Headers</b> |
| Content-Type: text/html<br>Content-Length: 170<br>Last-Modified: Tue, 18 May 2004 10:14:49 GMT  | <b>Entity Headers</b>   |
| <html><br><head><br><title>Welcome to the Amazing Site!</title><br></head><br><body><br><p>This site is under construction. Please come<br>back later. Sorry!</p><br></body><br></html> | <b>Message Body</b>     |

# 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 webservice 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: e.g. Python CGI, 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

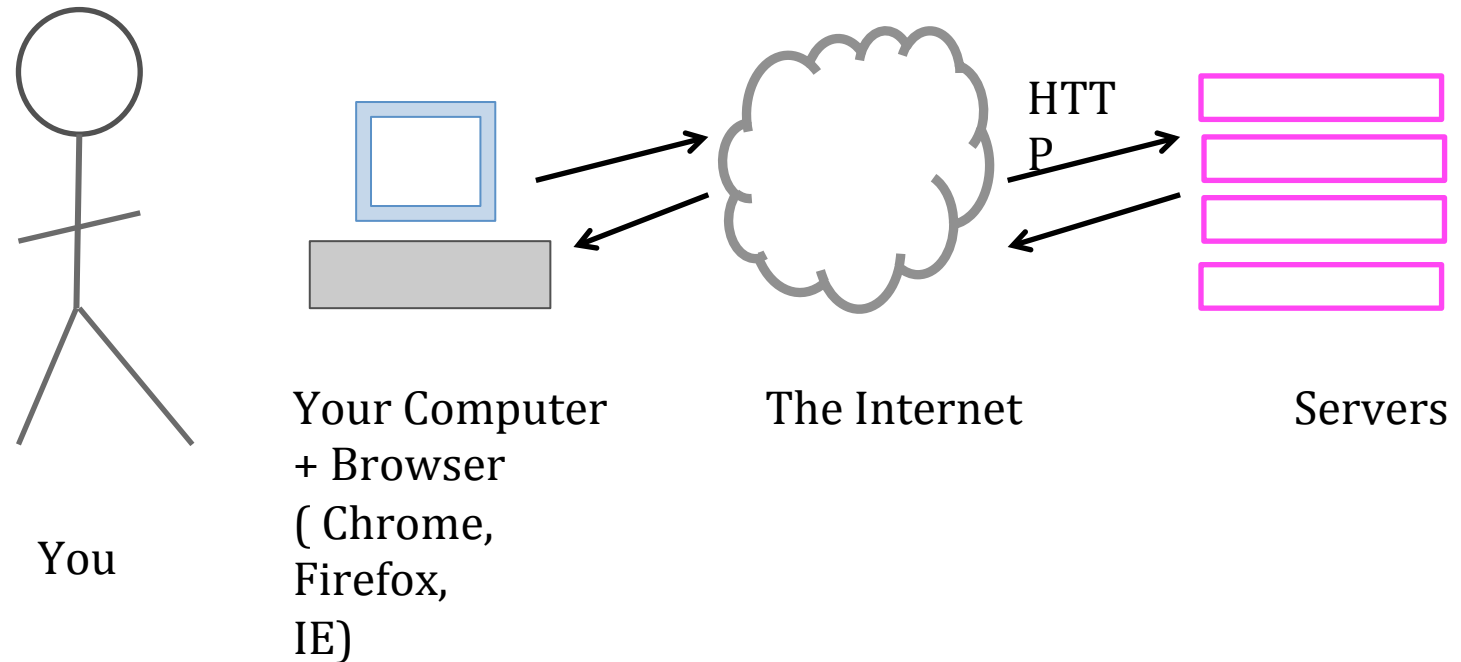
- Often work together
- PHP make queries to MySQL
- MySQL provides dynamic content
- PHP uses the content to generate a dynamic webpage.

# Additional technologies

- Python CGI: The Common Gateway Interface is a standard for external gateway programs to interface with information servers such as HTTP servers.
- 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

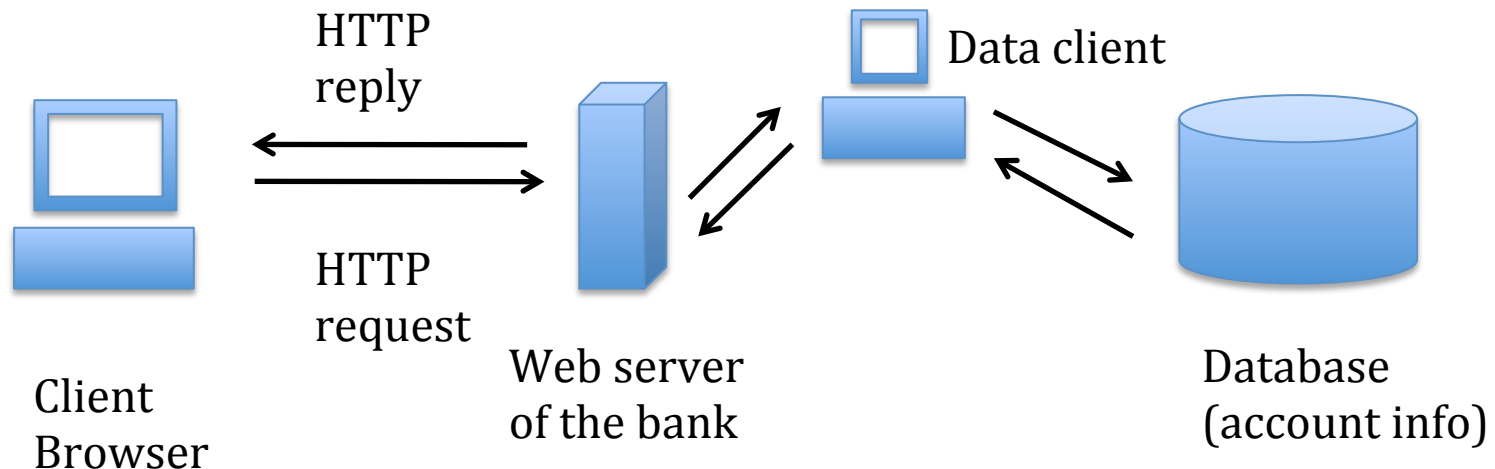




# Example: online banking

- A client access to a web browser (**the client**): [www.bankofamerica.com](http://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

- PythonCGI
- PHP
- MySQL
- Ajax (if time allows)
- Web frameworks

# 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://macromates.com/> (Mac OS)

Create an account at

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

Please read about HTTP:

<http://www.jmarshall.com/easy/http/#whatis>

# Further Readings

URL: [http://en.wikipedia.org/wiki/Uniform\\_resource\\_locator](http://en.wikipedia.org/wiki/Uniform_resource_locator)

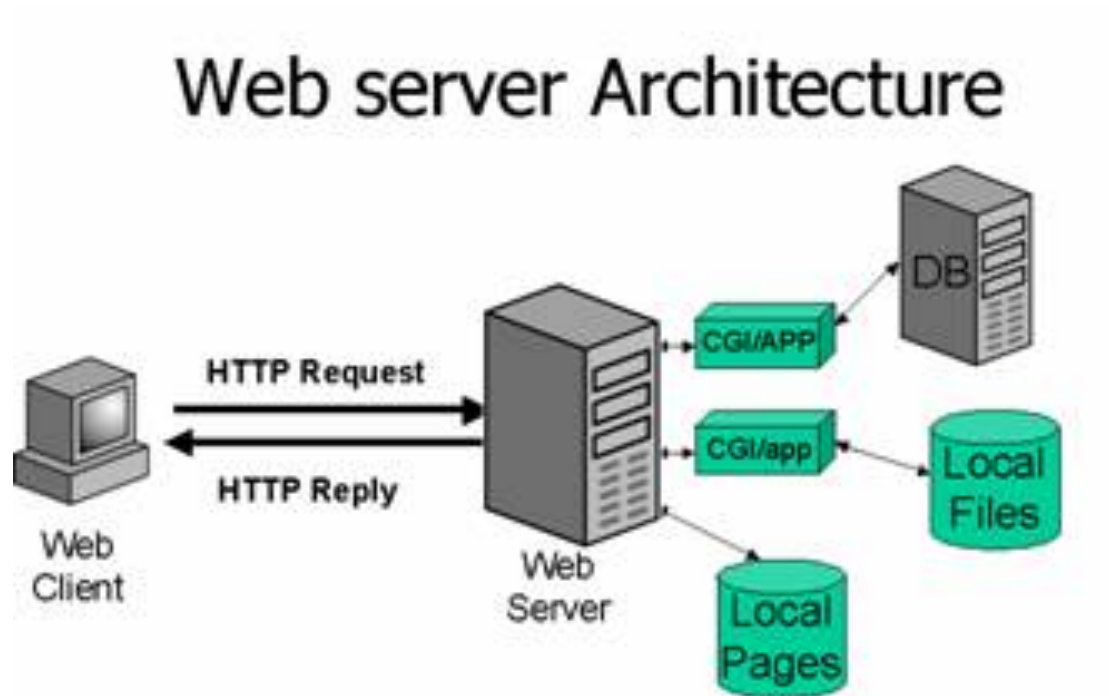
HTTP Request:

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

HTTP Response:

[http://www.tcpipguide.com/free/  
t\\_HTTPResponseMessageFormat.htm](http://www.tcpipguide.com/free/t_HTTPResponseMessageFormat.htm)

# Client-Server Architecture



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

# URL

- URL: Uniform Resource Locator

<http://www.nytimes.com/pages/world/index.html>



Protocol



Host



Path

- Query parameters

[http://www.nytimes.com/pages/world/index.html?](http://www.nytimes.com/pages/world/index.html)

[module=HPMiniNav&contentCollection=World&pgtype=Homepage&region=TopBar&action=click&t=qry69](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](http://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)