

EECS 485 (Web Databases & Information Systems)

Discussion
Jan 8th, 2010

Administrative

- Me: Vahed Qazvinian
- email: vahed@umich.edu
- Discussion: Fridays 10:30-11:30 (DOW 1010)
- Office Hours: Fridays 12:00-1:00 (CAEN lab)

PA1

- Please send me an email
 - with “485group” in the title
 - A group name, and your unique names.
- You’ll receive an email about your login info to “eecs485-2.eecs.umich.edu”
- PA1 on ctools today; due next week.
- After class you can stay and find a group if you haven’t already got one.

Web Applications

Web

- In 1990 and 1991, Tim Berners-Lee created the World Wide Web at the European Laboratory for Particle Physics (CERN) in Geneva, Switzerland
- The original purpose of the **World Wide Web (WWW)** was to provide easy access to cross-referenced documents that existed on the CERN computer network
- **Hypertext linking** allows you to quickly open other Web pages

HTTP

- A Web page is identified by a unique address called the **URL**
- Each URL consists of two basic parts:
 - A **protocol** (usually HTTP) and
 - Either the **domain name** for a Web server or a Web server’s **Internet Protocol address**
- Hypertext Transfer Protocol (**HTTP**) manages the hypertext links that are used to navigate the Web

Sample URL

http://www.google.com/help/index.html

Protocol Domain name Directory Filename

HTML

- Web pages are created using Hypertext Markup Language (**HTML**)
- Web pages are commonly referred to as **HTML pages** or **documents**
- A **markup language** is a set of characters or symbols that define a document's logical structure

HTML

```
<html>
  <head>
    <title>Michael J. Cafarella</title>
    <meta http-equiv="Content-Type" content="text/html" />
  </head>
  <body>
    <h1>Michael J. Cafarella</h1>
    <img alt="Michael J. Cafarella" data-bbox="238 521 288 571" />
    <p>Email: mjca@cs.washington.edu  

    Physical mail: mjca@cs.washington.edu  

    Department of Computer Science and Engineering  

    University of Washington  

    Box 357350  

    Seattle, WA 98195-3570  

    Office: 402 Allen Center
  </body>
</html>
```

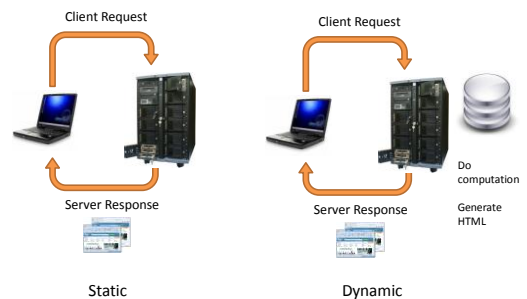
Michael J. Cafarella



Email: mjca@cs.washington.edu
 Physical mail: mjca@cs.washington.edu
 Department of Computer Science and Engineering
 University of Washington
 Box 357350
 Seattle, WA 98195-3570
 Office: 402 Allen Center

I recently graduated from the PhD program of the Department of Computer Science and Engineering at the University of Washington. I am currently working on my dissertation, and have been working on it for the last two years. I am currently working on my dissertation, and have been working on it for the last two years. I am currently working on my dissertation, and have been working on it for the last two years.

Web server



Web Applications

- What do we use the Web for?
 - amazon.com; cnn.com; [Facebook.com](http://facebook.com); [Blogger.com](http://blogger.com)
- What do they all have in common?
 - Web server
 - Apache; IBM HTTP Server; Microsoft Internet Information Services
 - Application layer
 - PHP; WebSphere Application Server
 - Database
 - MySQL; DB2 Universal Database; MS SQL Server

Web Applications

- You can mix technologies
 - Apache / WAS / DB2
 - IHS / PHP / MySQL
 - Apache / PHP / DB2
- You can run them side by side
 - Apache or IHS can take a look at a request and send it to either WAS or PHP
 - PHP can connect to both MySQL and DB2
 - Can use MySQL for logging and DB2 for customer information

Architecture



LAMP

- LAMP stands for Linux-Apache-MySQL-PHP.
- Instead of PHP, Perl and Python are also used.
- This is a free alternative to “WISA,” Windows-IIS-SQL Server-ASP.Net

Apache HTTP Server

- Apache is the most widely used web server software package in the world.
- In 2009 it became the first web server software to surpass the 100 million web site milestone
- In April 1996, Apache stood at 29% (and IIS at 1.6%) of the web server market:
 - <http://survey.netcraft.com/Reports/9604/ALL/>
- You can pull the distribution down from the Apache Web site
 - <http://www.apache.org/>



Apache 2 Configuration

- Apache 2 configuration: **/etc/apache2/***:
 - apache2.conf** general config file
 - httpd.conf** not used in Apache 2
 - sites-available/** default, and other sites (Apache 2 can serve multiple sites with their own IP addresses with one executable)
 - sites-enabled/** soft-links to enabled sites
 - ... (see **/etc/apache2/README**)
- **sites-available/default** may have:
 - DocumentRoot /var/www/

Modules

- Most tasks in Apache are handled by modules
- Some popular Apache modules:
 - PHP
 - mod_ssl (SSL: Secure Sockets Layer)
 - OpenSSL
 - perl

MySQL

- MySQL is a popular, scalable, free, open-source database management system (DBMS).
- It runs on Windows, Linux, and many other operating systems.
- that MySQL is used in 11 million installations in the world.

MySQL

- SQL: Structured Query Language
- Database Management System (DBMS): A system that manages relational databases; A collection of programs that enabling the storage, modification, and extraction of information from a database
- MySQL available online
<http://dev.mysql.com/>



PA1

- This is a group assignment. After forming your group, you will receive an email with the following information:
 - The hostname of the machine which will run your website (e.g. eecs485-2.eecs.umich.edu)
 - A port number (e.g. 12345)
 - A string of random characters (e.g. abcdefghij)
 - A user name (e.g. group1)

PA1

- In your home directory
 - **conf**: configuration files for Apache
 - **log**: log files for your Apache
 - **html**: location of code for your group's website
- Start Apache by
 - `/usr/sbin/apache2ctl -f ~/conf/apache2.conf -k start`
- You'll have access to a database on the MySQL server running on the machine

Resources

- Primary sources
 - Apache <http://apache.org/>
 - PHP <http://php.net/>
 - MySQL <http://mysql.com/>
- Secondary sources
 - Textbooks
 - New York PHP <http://nyphp.org/>
 - O'Reilly's OnLAMP <http://onlamp.com/>
 - Tons of other resources on Web