CS101 Lecture 04:
How the Web Works
Publishing Web pages

Aaron Stevens
14 September 2010

What You’ll Learn Today

– How does the WWW work?
– What are web servers, anyway?
– So I got some HTML pages and stuff.
  How do I make them show up on the web?
– Who can see my pages?
The World Wide Web

A system of interlinked hypertext documents and other resources accessed via the Internet.

http://www.w3.org/People/Berners-Lee/

How do we find stuff on the web?

Uniform Resource Locator
A standard way of specifying the location of a resource, it’s name, and how to get it.
Protocols

A protocol is a set of rules governing the exchange or transmission of data between devices.

How do you visit a website?

You don’t visit a website. The website gets delivered to you.
HTTP

HyperText Transfer Protocol
http://www.ltg.ed.ac.uk/~ht/WhatAreURIs/

What’s a Web Server?
Web server: a software application which waits for/responds to HTTP requests.

Tim Berners-Lee wrote 2 applications to make the web:
– A web browser called WorldWideWeb
– A server called HTTPd

The first WWW Server at CERN.
What’s a Web Server?

Today’s web servers use high-performance hardware like this: (fast network and disk access)

Pictured: IBM Blade Servers hosting files.myopera.com, photo from Wikipedia

LAMP Model

The most common structure for web applications uses this configuration:

**Linux** operating system for a server
**Apache** web server software
**MySQL** database software
**PHP/Perl/Python** scripting language to create dynamic HTML
How a Web Server Works

• Receives HTTP Request
• Search for resource (file) on disk
• Send HTTP Response (status code + data)
  – If not found: status 404 (NOT FOUND)
  – If not permitted: status 403 (FORBIDDEN)
  – ...
  – Else: status 200 (OK) + send data

How to Publish a Web Page

1. Create HTML document, locate ancillary files (e.g. images).
2. Transfer files to web server
3. Set permissions for read access
4. Test the URL in your browser
cs-people.bu.edu

Our WWW server is cs-people.bu.edu.

Main page is http://cs-people.bu.edu/

- Subdirectories for individual users:
  - http://cs-people.bu.edu/<username>
- Example:
  - http://cs-people.bu.edu/azs

Your UNIX Home Directory

Each user has a UNIX “home” directory:

General form:
/cs/course/<section>/<username>/

Example:
/cs/course/cs101a2/azs/

This has a UNIX pseudonym of ~.
Also mapped to your Windows Z:/ drive.
Your WWW Directory

The web server will map your URL:
http://cs-people.bu.edu/<username>/
to your UNIX’s account
~/public_html/ directory

(or Z:/public_html/ on Windows).

Locate files in Finder/Windows Explorer

Find your files on your local computer
How to Transfer Files to csa2.bu.edu

Use a file transfer client-program:
- Fetch (Mac)
  http://fetchsoftworks.com/
  (a free academic license is available)
- WinSCP (Windows)
  http://winscp.net/eng/index.php
  (also free)

Connecting by WinSCP:
- Hostname: csa2.bu.edu
- Be sure to use the “SFTP” protocol
WinSCP to \texttt{csa2.bu.edu}

After you connect, transfer files by drag 'n drop.

Then right-click to set permissions.

File Protection

Users and Groups

Many operating systems (for example, Unix) have a concept of users (each with unique username/password). Users are organized into groups.

Example: I’m in all of these groups:
\texttt{faculty3 cs108grades cs101stuff cs108stuff cs101grades}

Example: you are in one of these group:
\texttt{cs101a2, cs101a3, cs101a4, cs101a5, cs101b2, cs101b3 or cs101b4}
File Permissions

Each file has its own set of permissions for:

– Reading, writing or executing
– Owner, group, or others
– This leads to a 3x3 matrix of permissions:

<table>
<thead>
<tr>
<th>Owner</th>
<th>Group</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>Write</td>
<td>Execute</td>
</tr>
<tr>
<td>Write</td>
<td>Read</td>
<td>Execute</td>
</tr>
<tr>
<td>Execute</td>
<td>Read</td>
<td>Write</td>
</tr>
</tbody>
</table>

UNIX equivalent: 644

Setting Permissions by WinSCP

• Right-click to open this dialog

• Set the permissions to 644
Transferring File by Fetch

• Use drag’n’drop interface to transfer files...

Then use the Get Info button to set permissions.

Setting Permissions by Fetch

• Set permissions to 644
Testing the Webpage:

After uploading the files, test in your browser:

26

index.html

There are two special filenames that the web server looks for automatically:

- home.html
- index.html

Create the file index.html as the “main landing” page for your website, which corresponds to the URL:

http://cs-people.bu.edu/<username>/

Example:
What You Learned Today

- HTTP revisited, Web server
- UNIX home directory
- File Transfer Protocol
- UNIX File Permissions
- index.html

Announcements & To Do List

- Readings:
  - HTML Tutorial http://www.w3schools.com/HTML/
  - Reed ch 3, pp. 53-57 (today)
  - http://www.webpagesthatsuck.com

- HW 1 due WED 9/15
- Quiz 1 on TUE 9/21
  - Covers Lectures 1-5
  - See section on syllabus about how to prepare