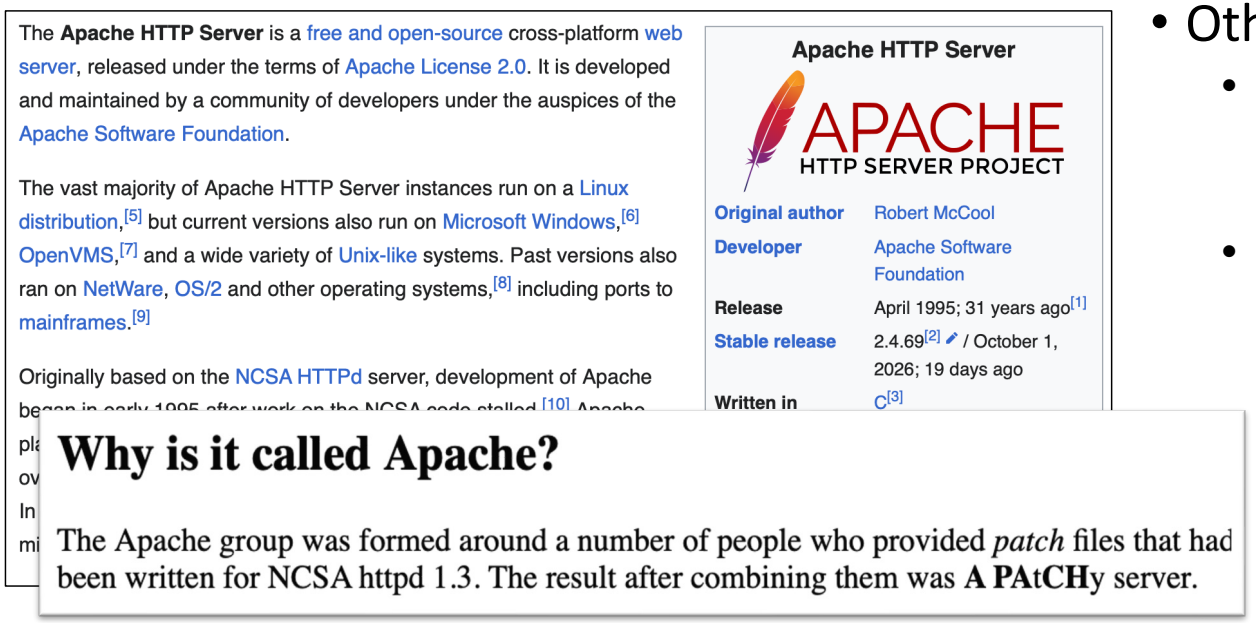
Lecture 18

408/508 *Computational Techniques for Linguists*

Today's Topics

- continue describing Apache2 webserver operation
  - for macOS and Ubuntu Linux
- Homework 7
  - set up Apache2
  - set up two index.html pages

httpd



The **Apache HTTP Server** is a free and open-source cross-platform web server, released under the terms of Apache License 2.0. It is developed and maintained by a community of developers under the auspices of the Apache Software Foundation.

The vast majority of Apache HTTP Server instances run on a Linux distribution,[5] but current versions also run on Microsoft Windows,[6] OpenVMS,[7] and a wide variety of Unix-like systems. Past versions also ran on NetWare, OS/2 and other operating systems,[8] including ports to mainframes.[9]

Originally based on the NCSA HTTPd server, development of Apache began in early 1995 after work on the NCSA code stalled.[10] Apache

Apache HTTP Server

APACHE HTTP SERVER PROJECT

Original author	Robert McCool
Developer	Apache Software Foundation
Release	April 1995; 31 years ago[1]
Stable release	2.4.69[2] / October 1, 2026; 19 days ago
Written in	C[3]

Why is it called Apache?

The Apache group was formed around a number of people who provided *patch* files that had been written for NCSA httpd 1.3. The result after combining them was **A PAtCHy** server.

- Others:
  - Nginx
    - (open sourced 2004)
    - **C**
  - Pingora
    - (Cloudfare: open sourced 2024)
    - **Rust**

Recap: Apache Webserver on macOS

Default static webpage storage location:

- http://localhost/
- **`/Library/WebServer/Documents/index.html`**
- `sudo apachectl start`
- `sudo apachectl stop`
- `sudo apachectl restart`

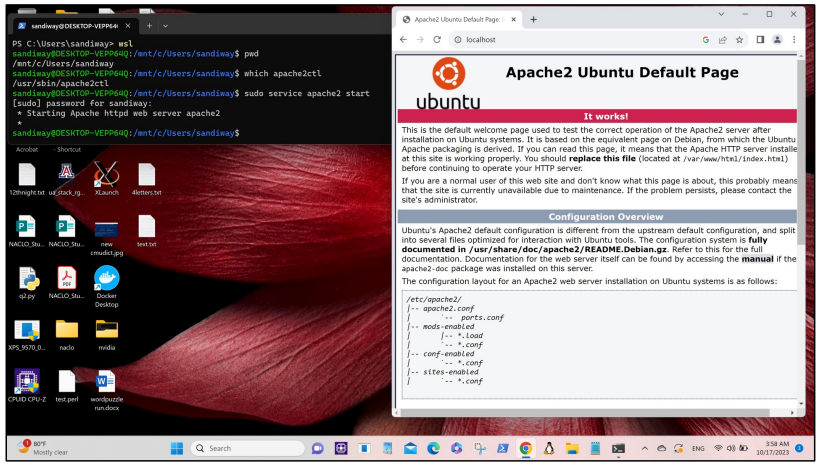


`nano` is a simple text editor
`^` means use the Control key
save file as
`/Library/Webserver/Documents/index.html`

Recap: Apache Webserver on WSL2 (Ubuntu)

- From Windows Powershell:
  - `wsl`
- Default page (Document Root):
  - `/var/www/html/index.html`
  - owner of file is `root`, so use the `sudo` prefix to edit it

    e.g. `sudo nano /var/www/html/index.html`,
- Commands (or `apache2ctl`):
  - `sudo service apache2 start`
  - `sudo service apache2 stop`
  - `sudo service apache2 restart`



DocumentRoot

DocumentRoot Directive

Description:	Directory that forms the main document tree visible from the web
Syntax:	`DocumentRoot directory-path`
Default:	`DocumentRoot "/usr/local/apache/htdocs"`
Context:	server config, virtual host
Status:	Core
Module:	core

This directive sets the directory from which `httpd` will serve files. Unless matched by a directive like `Alias`, the server appends the path from the requested URL to the document root to make the path to the document. Example:

```
DocumentRoot "/usr/web"
```

Apache2 on Ubuntu

- Master configuration file:
  - **`/etc/apache2/httpd.conf`**

```
sandiway@sandiway-VirtualBox:~$ cd /etc/apache2/
sandiway@sandiway-VirtualBox:/etc/apache2$ ls
apache2.conf    conf-enabled  magic           mods-enabled  sites-available
conf-available  envvars       mods-available  ports.conf    sites-enabled
sandiway@sandiway-VirtualBox:/etc/apache2$ ls -l
total 80
-rw-r--r-- 1 root root  7224 Oct  3 07:41 apache2.conf
drwxr-xr-x 2 root root  4096 Oct 24 20:43 conf-available
drwxr-xr-x 2 root root  4096 Oct 24 20:43 conf-enabled
-rw-r--r-- 1 root root  1782 Jun 27 10:05 envvars
-rw-r--r-- 1 root root 31063 Jun 27 10:05 magic
drwxr-xr-x 2 root root 12288 Oct 24 20:43 mods-available
drwxr-xr-x 2 root root  4096 Oct 24 20:43 mods-enabled
-rw-r--r-- 1 root root   320 Jun 27 10:05 ports.conf
drwxr-xr-x 2 root root  4096 Oct 24 20:43 sites-available
drwxr-xr-x 2 root root  4096 Oct 24 20:43 sites-enabled
sandiway@sandiway-VirtualBox:/etc/apache2$
```

Apache2 on Ubuntu

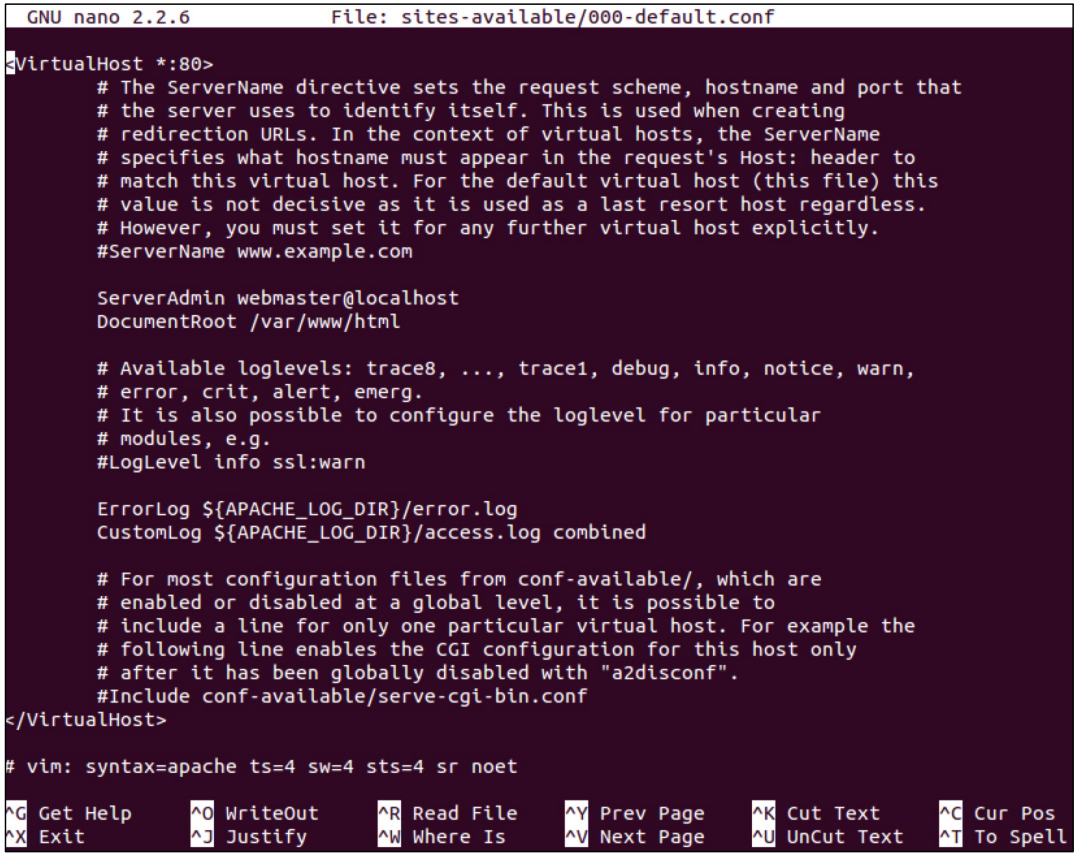
- `cd /etc/apache2/`
- `grep –r DocumentRoot` (`grep –r` searches into subdirectories for string)

```
sandiway@sandiway-VirtualBox:/etc/apache2$ grep -r DocumentRoot
sites-available/000-default.conf:       DocumentRoot /var/www/html
sites-available/default-ssl.conf:               DocumentRoot /var/www/html
sandiway@sandiway-VirtualBox:/etc/apache2$
```

- For httpd, relevant file is:
`/etc/apache2/sites-enabled/000-default.conf`

Apache2 on Ubuntu

`nano /etc/apache2/sites-enabled/000-default.conf`



Apache Webserver on macOS

- Configuration file: `/etc/apache2/httpd.conf`
  - DocumentRoot

```
232 # DocumentRoot: The directory out of which you will serve your
233 # documents. By default, all requests are taken from this directory, but
234 # symbolic links and aliases may be used to point to other locations.
235 #
236 DocumentRoot "/Library/WebServer/Documents"
237 <Directory "/Library/WebServer/Documents">
238     #
239     # Possible values for the Options directive are "None", "All",
240     # or any combination of:
241     #   Indexes Includes FollowSymLinks SymLinksifOwnerMatch ExecCGI MultiViews
242     #
243     # Note that "MultiViews" must be named *explicitly* --- "Options All"
244     # doesn't give it to you.
245     #
246     # The Options directive is both complicated and important.  Please see
247     # http://httpd.apache.org/docs/2.4/mod/core.html#options
248     # for more information.
249     #
250     Options FollowSymLinks Multiviews
251     MultiviewsMatch Any
252
253     #
254     # AllowOverride controls what directives may be placed in .htaccess files.
255     # It can be "All", "None", or any combination of the keywords:
256     #   AllowOverride FileInfo AuthConfig Limit
257     #
258     AllowOverride None
259
260     #
261     # Controls who can get stuff from this server.
262     #
263     Require all granted
264 </Directory>
```

User web files

Per-user web directories

Available Languages: en | es | fr | ja | ko | tr

On systems with multiple users, each user can be permitted to have a web site in their home directory using the `UserDir` directive. Visitors to a URL `http://example.com/~username/` will get content out of the home directory of the user "`username`", out of the subdirectory specified by the `UserDir` directive.

Note that, by default, access to these directories is **not** enabled. You can enable access when using `UserDir` by uncommenting the line:

```
#Include conf/extra/httpd-userdir.conf
```

in the default config file `conf/httpd.conf`, and adapting the `httpd-userdir.conf` file as necessary, or by including the appropriate directives in a `<Directory>` block within the main config file.

Apache2 on Ubuntu vs macOS

- Logs are in directory:
  `/var/log/apache2/`
  - `access.log`
  - `error.log`
- User web files in:
  - `~/public_html`
- In your home directory do:
  - `mkdir public_html`
  - `nano public_html/index.html`

- Logs are in directory:
  `/var/log/apache2/`
  - `access_log`
  - `error_log`
- User web files in:
  - `~/Sites`
- In your home directory do:
  - `mkdir Sites`
  - `nano Sites/index.html`

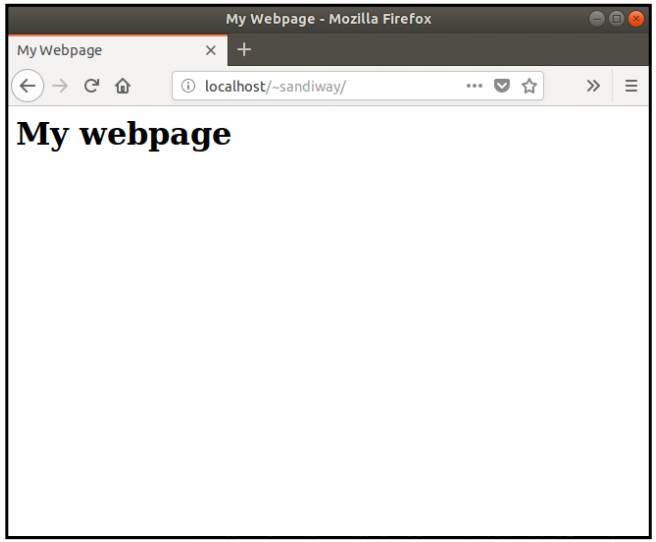
Apache2 on Ubuntu

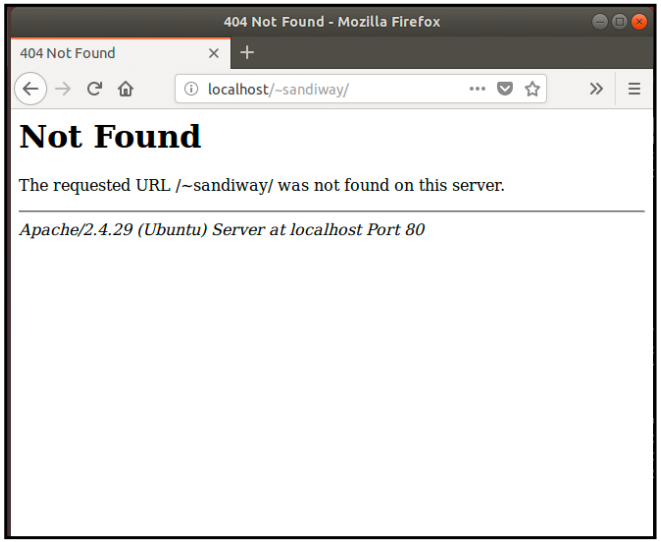
- To enable user web files in **~/public_html**
  - **sudo a2enmod userdir**
  - **sudo systemctl restart apache2**
  - http://localhost/~sandiway/

```
sandiway@sandiway-VirtualBox:~$ mkdir public_html
sandiway@sandiway-VirtualBox:~$ cd public_html
sandiway@sandiway-VirtualBox:~/public_html$ nano index.html
sandiway@sandiway-VirtualBox:~/public_html$ ls
index.html
sandiway@sandiway-VirtualBox:~/public_html$ ls -l
total 4
-rw-r--r-- 1 sandiway sandiway 92 Oct 24 21:12 index.html
sandiway@sandiway-VirtualBox:~/public_html$ sudo a2enmod userdir
[sudo] password for sandiway:
Enabling module userdir.
To activate the new configuration, you need to run:
  systemctl restart apache2
sandiway@sandiway-VirtualBox:~/public_html$ systemctl restart apache2
```

Apache2 on Ubuntu

- To enable user web files in **`~/public_html`**
  - `sudo a2enmod userdir` **`(a2dismod)`**
  - `sudo service apache2 restart`
  - `http://localhost/~sandiway/`



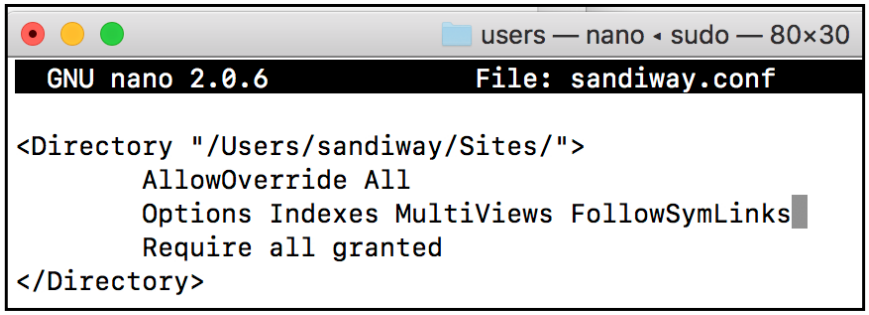


Apache Webserver on macOS

Static webpages

- storage locations:
  - http://localhost/~sandiway/ (**no need to be superuser**)
  - **`mkdir ~sandiway/Sites`** (**`/Users/`*`username`*`/Sites`**)
  - **`~/Sites/index.html`** (create this file!)
  - **`sudo nano /etc/apache2/users/sandiway.conf`**

create this file …

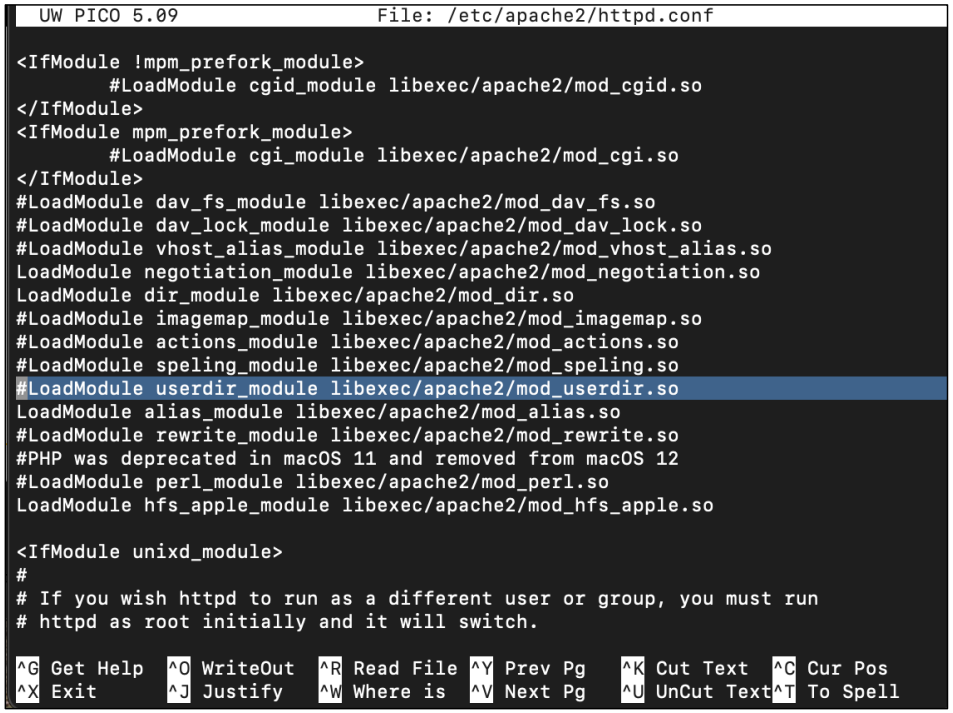


Apache Webserver on macOS

Static webpages

- http://localhost/~sandiway/

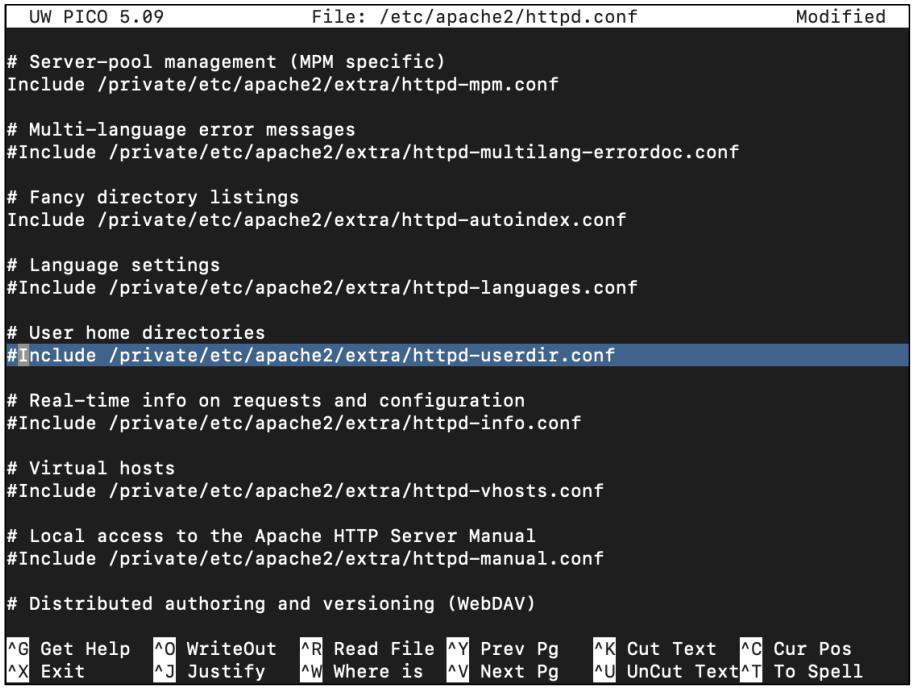
- storage locations:
  - `sudo nano /etc/apache2/httpd.conf`
- uncomment `mod_userdir.so` line
  - (*remove the comment char #*)



Apache Webserver on macOS

Static webpages

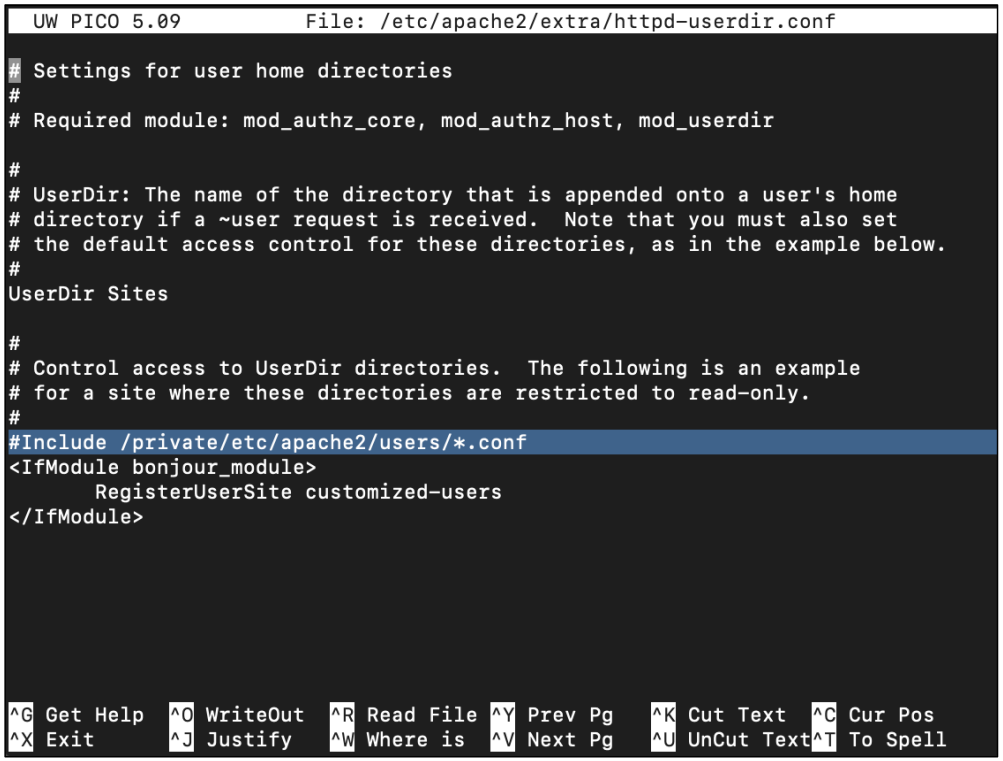
- also need to uncomment another line in `/etc/apache2/httpd.conf`
- uncomment the `httpd_userdir.conf` line
- (*remove the #*)



Apache Webserver on macOS

Static webpages

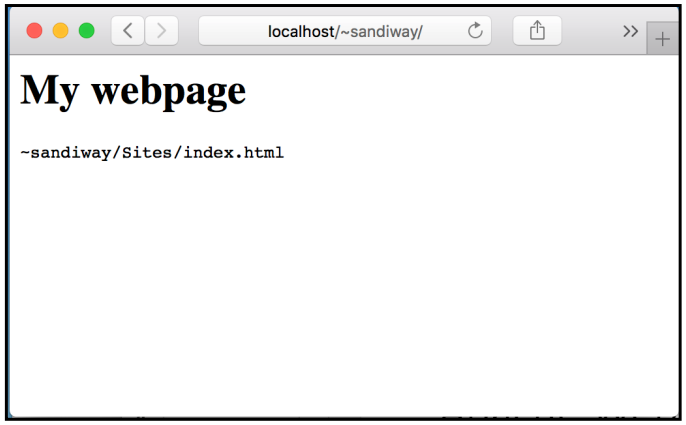
- `sudo nano /etc/apache2/extra/httpd-userdir.conf`
- uncomment this include
- (*remove the #*)

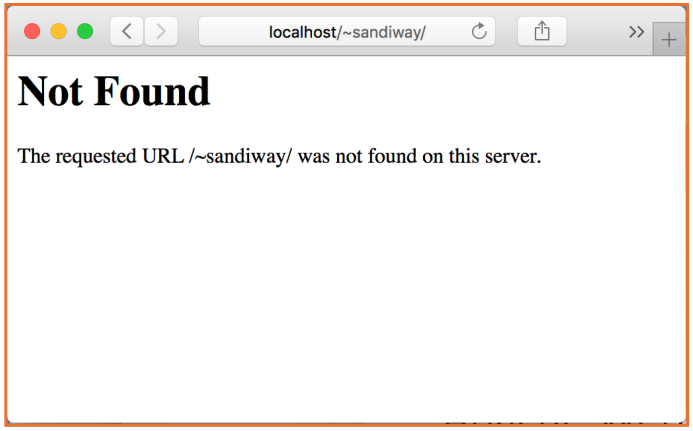


Apache Webserver on macOS

Static webpages

- storage locations:
  - `sudo apachectl –k restart`
  - create a file `~sandiway/Sites/index.html`



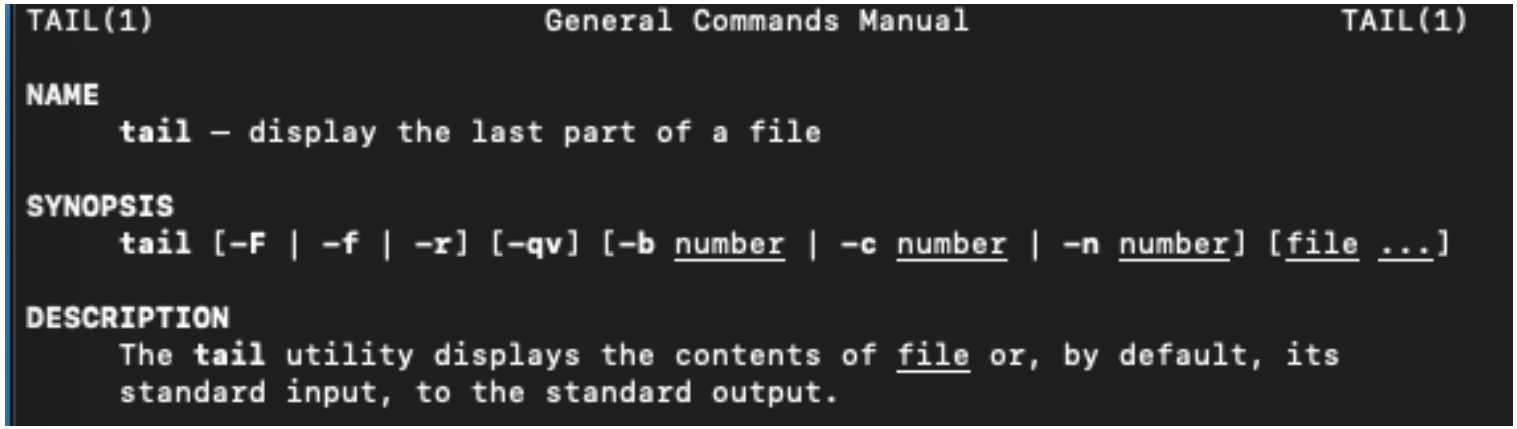


Apache Webserver on macOS

- `tail /var/log/apache2/access_log`

```
(base) apache2$ tail /var/log/apache2/access_log
127.0.0.1 - - [26/Sep/2022:05:56:09 -0700] "GET / HTTP/1.1" 200 555
127.0.0.1 - - [26/Sep/2022:05:56:29 -0700] "-" 408 -
::1 - - [30/Sep/2022:09:03:32 -0700] "GET / HTTP/1.1" 200 555
::1 - - [30/Oct/2022:21:38:44 -0700] "GET / HTTP/1.1" 200 555
::1 - - [30/Oct/2022:21:38:44 -0700] "GET /favicon.ico HTTP/1.1" 404 196
::1 - - [30/Oct/2022:21:59:51 -0700] "-" 408 -
::1 - - [30/Oct/2022:21:59:52 -0700] "GET /~sandiway HTTP/1.1" 404 196
::1 - - [30/Oct/2022:22:00:20 -0700] "GET /~sandiway HTTP/1.1" 404 196
::1 - - [30/Oct/2022:22:17:31 -0700] "GET /~sandiway HTTP/1.1" 301 235
::1 - - [30/Oct/2022:22:17:31 -0700] "GET /~sandiway/ HTTP/1.1" 200 101
(base) apache2$
```

tail command



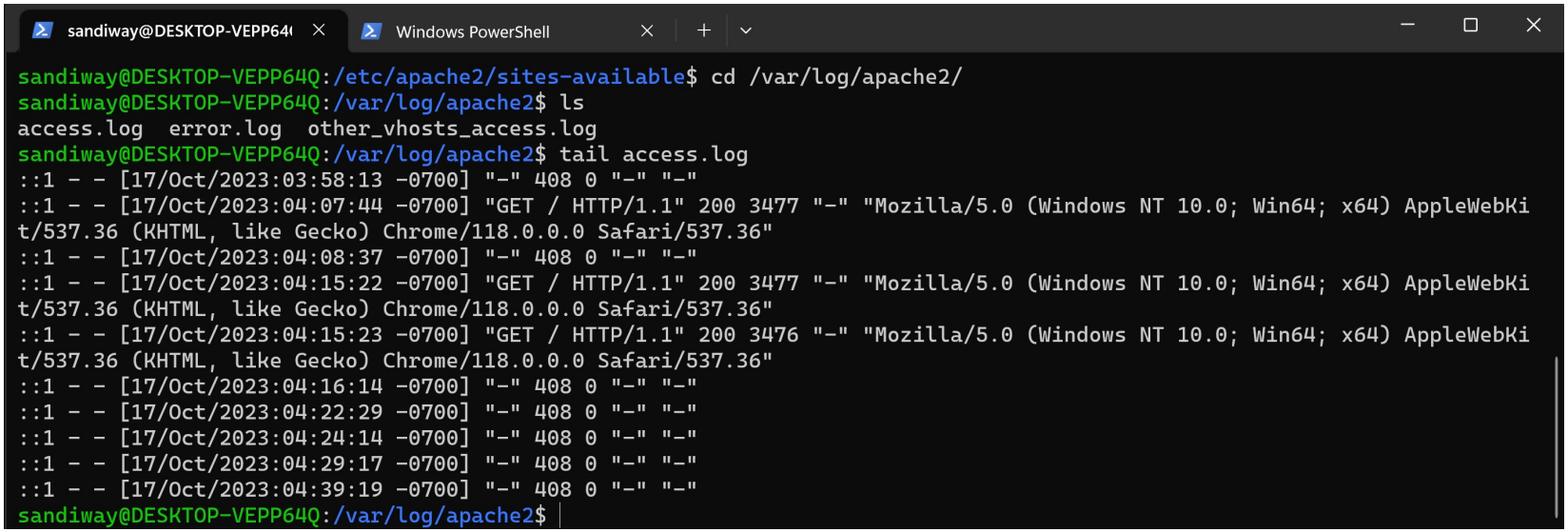
```
TAIL(1)                    General Commands Manual                    TAIL(1)

NAME
     tail – display the last part of a file

SYNOPSIS
     tail [-F | -f | -r] [-qv] [-b number | -c number | -n number] [file ...]

DESCRIPTION
     The tail utility displays the contents of file or, by default, its
     standard input, to the standard output.
```

Apache Webserver on Ubuntu



Homework 7

- For Mac and Ubuntu users
  - set up **Apache2** on your laptop
    - *bear in mind the different file locations!*
- In each case:
  - http://localhost/ (*system-wide homepage*)
  - http://localhost/~yourusername/ (*your personal homepage*)
  - Create two different `index.html` webpages at these locations, e.g. add your photo on your user homepage and your computer's photo on the system-wide homepage
  - Show your system works! (*include screen snapshots please*)
  - Submit one PDF file (by Sunday midnight)