

# Linux Network Administration

Apache Web Server  
COMP1071 Summer 2020

# Overview

- Apache2 is a software package which provides the infrastructure to deliver web services
- It is flexible, fast, open source, scalable, and modular
- Part of the common LAMP configuration
- The Apache HTTP server software listens for web requests and locates or creates the requested data, then delivers it to the client
- The website for the Apache Project is [apache.org](http://apache.org)

# Hypertext Transfer Protocol

- Hypertext means making text active by allowing the user to follow a logical connection (link) from text in the current document to another location in the document, or to another document
- The transfer protocol refers to the mechanism used by a web server and client to communicate those requests and documents to each other along with control information over TCP
- Modern HTTP has expanded far beyond simply GETting html documents and now handles many different types of documents and other resources which are not like documents
- Links are not restricted to being associated with text

# Document Store

- The document store is the place on a webserver where documents to be provided to clients are stored
- The default for Ubuntu is `/var/www/html` and this is configurable
- Documents can be simple text (e.g. plain text, html, etc.) or more sophisticated files such as multimedia files (e.g. audio, image, video, etc.) and structured data (e.g. xml, json, etc.)
- Apache can run programs found on the server if configured to do so, returning the program output to the client instead of the static content of a file in the document store

# Uniform Resource Locator

- URLs provide the addressing scheme for identifying resources such as documents on the internet
- There are 4 components to a URL

[protocol://host\[:port\]/resource](#)

- protocol is how the client will format communications, host[:port] is where the client will send requests, most protocols have a default port so that you don't need to type it out in the URL
- resource is the part used by the server to decide what to send to the client (e.g. a file or program or module name)

# Resource Content

- Resources requested may be static resources, such as a file containing html which is sent unaltered to the client
- Servers are commonly set up to look for several default index files when a client does not specify a resource
- Resources may also specify a program file which the server will run, or a module to invoke in order to retrieve information or make the web server perform some action, this is known as dynamic content
- Like static content, the server must be configured to permit dynamic content to be requested

# Installation

- Apache 2 is the current major version
- The package and service names for it are just `apache2`
- There are many additional packages to extend the functionality of `apache2` (e.g. to provide scripting capability, make database connections, authenticate users, create graphics)

```
apt-get install apache2
```

# Configuration

- The default installation sets up a default document store, configuration and site
- Apache2 uses a staged approach to configuration files
- Configuration files are created and modified in `/etc/apache2/conf-available`, then deployed and undeployed using `a2enconf` and `a2disconf` commands
- The common configuration elements for the service daemon are specified in `/etc/apache2/apache2.conf` which includes the files ending in `.conf` from `/etc/apache2/conf-enabled`



# apache2.conf

## /etc/apache2/apache2.conf:

```
<Directory />
    Options FollowSymLinks
    AllowOverride None
    Require all denied
</Directory>
<Directory /usr/share>
    AllowOverride None
    Require all granted
</Directory>

AccessFileName .htaccess
<FilesMatch "\.ht">
    Require all denied
</FilesMatch>
```

```
Include ports.conf
```

```
IncludeOptional conf-enabled/*.conf
```

```
IncludeOptional mods-enabled/*.load
```

```
IncludeOptional mods-enabled/*.conf
```

```
IncludeOptional sites-enabled/*.conf
```

## /etc/apache2/ports.conf:

```
Listen 80
```

```
<IfModule ssl_module>
```

```
    Listen 443
```

```
</IfModule>
```

# Modules

- Modules give us a way to extend the functionality of Apache by dynamically loading code into the running service
- Modules are configured in `/etc/apache2/mods-available` and enabled/disabled using `a2enmod` and `a2dismod`
- Many modules have options which can be set up in their `.conf` file found in `/etc/apache2/mods-available`
- Modules can be checked for their enablement state in configuration files using the `<IfModule modname_module>` `</IfModule>` stanza

# Virtual Servers

- A virtual server is an additional configuration file used to provide differentiated service based on the requested server name or IP address
- It compartmentalizes configurations, improves security, abstracts the service from the host platform, and simplifies the server configurations
- In Apache terms, a virtual server is known as a site
- Site files are kept in `/etc/apache/sites-available` and enabled/disabled using `a2ensite` and `a2dissite`, then the server must be reloaded
- Two default site files are created automatically by the package installation, `000-default.conf` (enabled) and `default-ssl.conf` (disabled)

# Sample Site File

- Default site file, enabled by install of apache2, depends on Directory stanza from `/etc/apache2/apache2.conf` file

```
<VirtualHost *:80>
  ServerName www.example.com
  ServerAdmin webmaster@localhost
  DocumentRoot /var/www/html
  ErrorLog ${APACHE_LOG_DIR}/error.log
  CustomLog ${APACHE_LOG_DIR}/access.log combined
<Directory /var/www/html>
  Options Indexes FollowSymLinks
  AllowOverride None
  Require all granted
</Directory>
</VirtualHost>
```

- Only the ServerName and DocumentRoot along with a matching Directory stanza, needs to be set for trivial sites

# Command Line Tools

- The apache service can be controlled using the service command (i.e. `service apache2 start|stop|restart`)
- There is a tool built into apache to instruct the daemon to perform various tasks, `apachectl`
- `apachectl graceful` is the preferred method to reload the configuration files because it includes syntax checks and does not interrupt transactions
- `openssl`, `telnet`, `curl`, and `wget` can be useful diagnostic tools to test a server

# Logs

- Logs are set up in the configuration files and often are specific to the individual sites
- The default log files are text files stored in the directory `/var/log/apache2`, and are automatically aged
- Typical log entries include the timestamp, client IP, request, result code and size, and any errors generated
- A number of packages are available to decipher and summarize the logs (e.g. `loganalyzer`, `analog`)