Apache, Virtual Hosts, & mod_rewrite

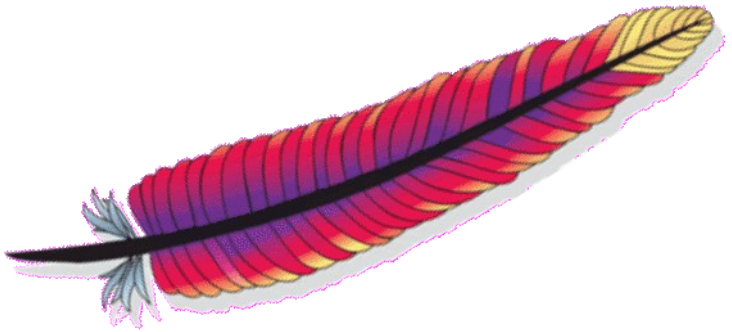
Network Startup Resource Center
www.nsrc.org

Objectives

- Introduce Core Concepts & Terminology
  - Apache HTTPd
  - Modules
  - Virtual Hosts
  - Aliases
  - mod_rewrite
  - Basic Authentication

What is Apache?

- World's Most Widely Used Web Server
  - http://httpd.apache.org/
- A Community of Developers
  - http://www.apache.org/
- An Indigenous People of North America
  - http://en.wikipedia.org/wiki/Apache



Quick Facts on Apache

- Released in 1995
- Evolved from NCSA HTTPd
- Used on 337 million websites
  - per Netcraft's April 2015 Survey
- Runs on all major server operating systems
  - Unix, Linux, Windows, Netware, etc.
- Has its own Software License
  - Compatible with GPL v3

What does Apache do?

- Listens for HTTP(S) requests
  - HTTP on Port 80, HTTPS on Port 443
- Serves documents requested via HTTP(S)
- Processes requests using modules
  - alias, auth, cgi, include, rewrite, userdir
  - many others, both Apache & 3rd party maintained
- Can act as a Proxy Server
  - mod_proxy

What is a Virtual Host?

- Allows one server to host many web sites
- Two kinds of Virtual Hosts
  - IP Based:
    - One IP is used for each site
  - Name Based:
    - One IP is used for many sites

IP Based Virtual Hosting

- Requires a separate IP for every host name
- Is compatible with current SSL
- Can work even if DNS has failed
- Uses up valuable IP address resources

Name Based Virtual Hosts

- Apache reads a Server Name from HTTP headers
  - Different directories are served based on the server name
- SSL fails with name based hosting
  - Hostname is not a part of TLS/SSL handshake
  - Certificate Matching does not work
- SSL can work with name based hosting if:
  - All Virtual Hosts are in the same domain
  - You have a wildcard SSL certificate for the domain
- SSL/TLS will work with virtual hosts some day (RFC4366)

Configuration Considerations

- Directory naming conventions:
  - FreeBSD: /usr/local/www
  - Linux: /var/www
  - Ubuntu 14: /var/www/html
- Directory permissions are important

Aliases

- Our web sites are located in a web root
  - /var/www/html for Ubuntu 14.04 LTS
- We want Apache to serve pages outside the root
  - for example, /usr/local/www/books
- Use an Alias to make this work
  - Add it to your server configuration file
  - Ubuntu: /etc/apache2/sites-available/000-default.conf

mod_rewrite

- Redirect requests for information
  - based on a particular request
  - based on a pattern of pages to a single page
  - based on a pattern of pages to a pattern of pages
- Extremely powerful
- Can use Regular Expressions
- Useful when moving pages or web sites
- Must be enabled: `a2enmod rewrite`

mod_rewrite example

- Some users visit our wiki using HTTP
- We require them use HTTPS
- Configured in /etc/apache2/sites-available/000-default)

```
# turn on the use of the mod_rewrite module
        RewriteEngine on
# trac logins must be secure
        RewriteCond %{SERVER_PORT} !443
        RewriteCond %{REQUEST_URI} ^/trac
        RewriteRule ^(.*)$ https://nsrc.org$1 [R=301]
```

mod_rewrite

- Redirect requests for information
  - based on a particular request
  - based on a pattern of pages to a single page
  - based on a pattern of pages to a pattern of pages
- Extremely powerful
- Can use Regular Expressions
- Useful when moving pages or web sites
- Must be enabled: `a2enmod rewrite`

Basic Access Control: htaccess

- htaccess is a directory-level configuration file
- Put an htaccess file in any directory to restrict
- It restricts directory access to authorized users
- Users/Passwords managed in an htpasswd file
- Use with mod_rewrite + SSL for some security

Review

- Apache HTTPd
- Modules
- Virtual Hosts
- Aliases
- mod_rewrite
- Basic Authentication