What is the Apache HTTP Server?

• Powerful, flexible, HTTP/1.1 compliant web server
• World's most popular web server since 1996
• Substantially faster, more stable, and more feature-full than many other web servers
• Considered more secure than many other web servers
• Highly configurable and extensible with third-party modules
• Provides full source code and comes with an unrestricted license
• Runs on most operating systems. Runs well on Linux and most Unixes.
Apache Features

- DBM databases for authentication allows you to easily set up password-protected pages with enormous numbers of authorized users, without bogging down the server.

- Customized responses to errors and problems allows you to set up files, or even CGI scripts, which are returned by the server in response to errors and problems.

- Unlimited flexible URL rewriting and aliasing to solve most URL manipulation problems.

- Content negotiation to serve clients of varying sophistication with documents which offer the best representation of information that the client is capable of accepting.

- Virtual Hosts allows the server to distinguish between requests made to different IP addresses or names (mapped to the same machine).

- Configurable Reliable Piped Logs to generate logs in the format that you want. Apache can send log files to a pipe, allowing for log rotation, hit filtering, real-time splitting of multiple vhosts into separate logs, and asynchronous DNS resolving on the fly.
Why “Apache”?

• The original Apache was a set of patch files which you applied against the aging NCSA httpd V1.3 server. Popularly (though incorrectly) accepted, it's a considered cute name which stuck. Apache is "A PAtCHy server".

• Officially Apache says:

  “The name 'Apache' was chosen from respect for the Native American Indian tribe of Apache (Indé), well-known for their superior skills in warfare strategy and their inexhaustible endurance.”
Apache HTTP Server
Market Share

- Number one web server since 1996
- Apache served 96,531,033 web sites in Jan 2009, over 50% of active sites
Origins of the Apache HTTP Server

In February of 1995, the most popular server software on the Web was the public domain HTTP daemon developed at the National Center for Supercomputing Applications, University of Illinois, Urbana-Champaign. However, development of that httpd had stalled, and many webmasters had developed their own extensions and bug fixes that were in need of a common distribution. A small group of webmasters gathered together for the purpose of coordinating their changes (in the form of "patches"). Using NCSA httpd 1.3 as a base, they added all of the published bug fixes and worthwhile enhancements they could find, and made the first official public release (0.6.2) of the Apache server in April 1995.

After extensive beta testing, many ports to obscure platforms, a new set of documentation, and the addition of many features in the form of our standard modules, Apache 1.0 was released on December 1, 1995.

Less than a year after the group was formed, the Apache server passed NCSA's httpd as the #1 server on the Internet.
The Apache Software Foundation (ASF)

• In 1999, members of the Apache Group formed the Apache Software Foundation to provide organizational, legal, and financial support for the Apache HTTP Server. The foundation has placed the software on a solid footing for future development, and greatly expanded the number of Open Source software projects, which fall under this Foundation's umbrella.

• The Apache Software Foundation is a decentralized community of developers. The software they produce is distributed under the terms of the Apache License and is therefore free software / open source software.

• The Apache projects are characterized by a collaborative, consensus-based development process and an open and pragmatic software license.

• Each project is managed by a self-selected team of technical experts who are active contributors to the project.

• The ASF is a meritocracy, implying that membership to the foundation is granted only to volunteers who have actively contributed to Apache projects.

• Among the ASF's objectives are to provide legal protection to volunteers working on Apache projects, and to prevent the Apache brand name from being used by other organizations without permission.
Other ASF Projects

• **Over 85 projects**
  • Ant: Java-based build tool
  • Apache Geronimo: a Java EE server
  • Hadoop: Java software framework that supports data intensive distributed applications.
  • Jakarta: server side Java (including its own set of sub-projects)
  • mod perl: module that integrates the Perl interpreter into Apache server
  • mod python: module that integrates the Python interpreter into Apache server
  • SpamAssassin: email filter used to identify spam.
  • Struts: Java web applications framework
  • Tcl: dynamic websites using Tool Command Language
  • Tomcat: a web container for serving servlets and JSP
  • Xalan: XSLT processors in Java and C++
  • Xerces: A validating XML parser
Apache License (AL)

- Considered a little “different”
- A Free (Libre) software license
  - Requires preservation of copyright notice
  - Requires preservation of disclaimer
- Not a “Copyleft” license
  - Allows use of source code in proprietary products
- Previous versions (1.0, 1.1) not compatible with previous versions of GNU General Public License
- “Not compatible” is bad – you can't distribute GPL code used /combined with AL code
- Latest version ALv2.0 compatible with GPL 3.0
Installing Apache

• On Unix, the usual install methods
  – On Red Hat based Linux distributions
    bash-3.2$ sudo yum install httpd
  – From binaries from ASF
    • More control over final locations
  – From source
    • Let you configure Apache as you need

• On Windows, get the .msi installer package
Starting and Stopping Apache

• Use the supplied apachectl program
  apachectl -k start/stop/restart
• Use rc.d startup script
  /etc/init.d/httpd start/stop/restart
• On Red Hat use the 'service' command
  service httpd start/stop/restart
• On Red Hat use chkconfig to start at boot
  chkconfig --levels 235 httpd on
Configuring Apache

• Apache has a LOT of configuration options
  – Time to buy a “big fat book” at Barnes and Nobel
    • Make sure to get a recent one, a lot are stale (for older Apache versions)
  – Or go to http://httpd.apache.org/docs/2.2/

• Main config file is /etc/httpd/conf/httpd.conf

• The simplest (test) config change may be to uncomment

  #NameVirtualHost *:80
Using Apache

- After a default install, test by:
  - Uncomment NameVirtualHost
  - Start apache
  - On the same machine, point a browser at http://localhost
  - Look at the logs for possible errors
Some Apache Defaults

• On Red Hat installs
  – The ServerRoot is /etc/httpd
  – The logs are under that in /etc/httpd/logs
  – The DocumentRoot is /var/www/html
    • This is where html documents go

• Read the config file for locations
Apache Logs

• Look in the config file for log file locations
  – On Red Hat, the server root is set to /etc/httpd
  – Logs under the server root, in logs

• Two main default logs are access_log and error_log

• Very configurable
  – Can use the system log, syslogd
  – Custom locations, format
Apache Configuration

• Some of the more common things Apache is configured for are:
  – Serving up documents in home directories
  – Rewriting or redirecting URL's
  – Controlling access to web pages
  – Enabling SSL (encrypted) support
  – Setting up virtual hosts

• We'll do a little of each of these in lab