Prerequisites

• Network
• Hardware
• Software
• Wetware (people!)
Prerequisites: Network

- Permanent and direct IP access
- Vulnerable periods?
- Support?
- 24hrs/day, 365 days/year?
- Holiday/Illness cover?
Prerequisites: Hardware

- Macs, PCs, Suns, ...
- Hardware support? (24x7?)
- Backups?
- Disc space
- Network speed
- Memory
- Processor power
Prerequisites: Software

- Permanently running daemon
- Software support?
- Service rates?
- DNS lookup rates?
- CGI?
Prerequisites: Wetware

- Checking logfiles
- Changing configuration files
- Software updates & patches
- Data files
- Backups
- Holiday/Illness cover
Support tools

- Editors
- HTML checkers
- Graphics manipulators
- Scanners etc.
- Log file analyser
- CGI programs
Support tools: Text editors

- Plain text editor
- Configuration files
- HTML data files
- emacs, vi, pico
Deprecated support tools: HTML editors

- There exist specialist HTML editors
- Inflexible & incomplete
- Poor quality HTML
- Plain text editors still pretty good
- Avoid MS Word like the plague
Support tools: HTML checkers

- Check HTML syntax
- Check HTML quality
- Check links still work
- weblint
- cron job
Support tools: Graphics manipulators

- Best all-rounder is gimp—the GNU Image Manipulation Program
- Also ee—Electric Eyes
- Bother available as Red Hat packages.
Support tools: Scanners etc.

- Flat bed scanners
- Digital cameras
Support tools: CGI programs

- Common Gateway Interface
- Not covered in this course
- SSI
- SSlexec
- PHP
- perl CGI module
- python CGI module
Support tools: Secure access

- ssh: Replacement for rsh, rlogin, rcp
- Machine daemon: sshd
- Red Hat package
- Unix Support’s CD
Example server

- 3Com 3c905B, 700MHz Athlon, 256MB RAM, 20GB disc
- Red Hat Linux 7.3
- Apache v1.3.23
Apache installation

- As root
- Unix Support’s NFS server
- Mount Red Hat mirror
- Locate Apache package
- Install Apache package
- Unmount Red Hat mirror
Apache installation: Mounting the mirror

- Unix Support mirror: nfs-uxsup.csx.cam.ac.uk
- Red Hat mirror: /linux/redhat

# mount -o ro nfs-uxsup.csx.cam.ac.uk:/linux/redhat /mnt
# cd /mnt/updates/7.3/en/os/i386/
# ls -l apache-*
-rw-r--r-- ... apache-1.3.23-14.i386.rpm
-rw-r--r-- ... apache-devel-1.3.23-14.i386.rpm
-rw-r--r-- ... apache-manual-1.3.23-14.i386.rpm
Apache installation: Examining the package

# rpm --query --info --package apache-1.3.23-14.i386.rpm
Name : apache
Version : 1.3.23
Release : 14
Install date: (not installed)
Group : System Environment/Daemons
Size : 1248999
Build Host: daffy.perf.redhat.com
Summary : The most widely used Web server on the Internet.
Description :
Apache is a powerful, full-featured, efficient, and freely-available Web server. Apache is also the most popular Web server on the
Internet.
Apache installation: Examining the package

# rpm --query --list --package apache-1.3.23-14.i386.rpm
/etc/httpd/conf
/etc/httpd/conf/httpd.conf
...
/etc/rc.d/init.d/httpd.init
...
/var/www
/var/www/html
/var/www/html/index.html
/var/www/icons
/var/www/icons/a.gif
...
/usr/man/man8/httpd.8
...
/usr/sbin/httpd
...
Apache installation: Installing the package

- This has not started the server.
- Please remember to unmount the mirror.

```
# rpm --install apache-1.3.23-14.i386.rpm
# cd
# umount /mnt
```
Apache installation: Configuration file layout

```
+---- conf/ ----+----* .conf
|              +---- access.log
/etc/httpd/ ---+---- logs -> /var/log/httpd/ ----+
|              +---- error.log
+---- modules -> /usr/lib/apache
```
Apache installation: Data file layout

+--- cgi-bin/                     empty
    |
/var/www/---+--- icons/    --- *.gif
    |
+--- html/    --- index.html    default
Apache installation: System file layout

- /usr/sbin: Binaries
- /usr/man: Manual pages
- /etc/rc.d: Startup/Shutdown scripts
- /etc/logrotate.d: Log rotation
Configuring the operating system

- Package provides a user and group for the daemon
- We need to add a group for the apache administrators
- And at least one group for the web authors
- Avoid use of root
- Log rotation
Configuring the O/S: User & groups

# groupadd -r webadmins
# groupadd -r webeditor
# vi /etc/group
Configuring the O/S: File permissions as installed

- Only root can make modifications.

```bash
# ls -ld /var/www /etc/httpd /var/log/httpd
drwxr-xr-x  3 root root  1024 Jun 27 12:09 /etc/httpd
drwxr-xr-x  5 root root  1024 Jun 27 12:09 /var/www
drwxr-xr-x  2 root root  1024 Jun 27 16:36 /var/log/httpd
```
Configuring the O/S: File permissions

- Change the group to webadmins:

  ```
  # chgrp -R webadmins /etc/httpd /var/log/httpd /etc/logrotate.d/apache
  # chgrp -R webeditor /var/www
  ```

- Let the group write to the directories:

  ```
  # chmod -R g+w /var/www /etc/httpd /var/log/httpd /etc/logrotate.d/apache
  ```
• Make the group ownership “setgid”:

```
# find /var/www /etc/httpd /var/log/httpd -type d -exec chmod g+s {} \;
```
Configuring the O/S: File permissions—as changed

- The daemon will run as user apache.
- How can the daemon write its log files?
- It *starts* life and opens the log files as user root.

```
# ls -ld /var/www /etc/httpd /var/log/httpd /etc/logrotate.d/apache
drwxrwsr-x 3 root webadmins 1024 Jun 27 12:09 /etc/httpd
-rw-rw-r-- 1 root webadmins 172 Jun 27 12:09 /etc/logrotate.d/apache
drwxrwsr-x 5 root webeditor 1024 Jun 27 12:09 /var/www
drwxrwsr-x 2 root webadmins 1024 Jun 27 12:09 /var/log/httpd
```
Being a webadmin

• A fresh login will pick up membership of group webadmins.

• This gives access to existing webadmins-writable files.

• Files created in *setgid* directories will be owned by group webadmins

• Check your permissions mask
Starting the server

```
# /etc/rc.d/init.d/httpd start
Starting httpd: [ OK ]
```
Restarting or stopping the server

# /etc/rc.d/init.d/httpd restart
Shutting down http: [ OK ]
Starting httpd: [ OK ]

# /etc/rc.d/init.d/httpd stop
Shutting down http: [ OK ]
Configuring the service

• As a webadmin, not as root!

• Directory: /etc/httpd/conf/

• Directory and contents are group-writable by webadmins

• httpd.conf: Configuration file

• srm.conf & access.conf: Obsolete & empty

• Directory: /etc/logrotate.d/
• **apache**: Controls the rotation of the log files.

• File is writable by members of group webadmins.
httpd.conf: Running the daemon

ServerType standalone
ServerRoot /etc/httpd
DocumentRoot /var/www/html
Port 80
User apache
Group apache
ServerAdmin rjd4@cam.ac.uk
ServerName www.inst.cam.ac.uk
ErrorLog /var/log/httpd/error_log
LogLevel info
Options None
Syntax: Running the daemon

- ServerType standalone
  The daemon will not rely on inetd to launch it on demand but will run permanently.

- ServerRoot /etc/httpd
  Any files referred to in this configuration file will either be fully qualified or resolved relative to this directory.

- DocumentRoot /var/www/html
  The documents to be served are found in this directory.
• **Port 80**
This is the standard port of WWW services. It is *privileged* on a Unix system so must be opened by root. Once opened, the port can be passed to unprivileged services (e.g. running user apache). Ports 8000 and 8080 are commonly used ports for completely unprivileged servers.

• **User apache**
  
  Group apache

We created a user and group specifically for the webserver. These two lines tell the server to use them. The server can change its user and group ids only if it is started as root.

• **ServerAdmin rjd4@cam.ac.uk**

Some error messages displayed to the client can contain a contact email address. This is where it is defined.
• **ServerName www.inst.cam.ac.uk**
You may not need this line. If your machine’s real name is boring.\textit{inst}.cam.ac.uk but there is a DNS record pointing www.\textit{inst}.cam.ac.uk to it as well then you want the server to identify itself as www.\textit{inst}.cam.ac.uk. This is how you override the machine’s host name.

• **ErrorLog /var/log/httpd/error_log**
Any error messages will be logged to the file /var/log/httpd/error_log.

• **LogLevel info**
An error in Apache comes with a severity rating. This directive specifies what the minimum level to log is.

• **Options None**
Apache has various options, almost all of which default to “on”. We will turn them off so we are forced to meet them explicitly in this course.
Syntax: Suboptions to LogLevel

- **emerg**
  Emergencies—system is unusable. e.g. “Child cannot open lock file. Exiting.”

- **alert**
  Alert—Action must be taken immediately. e.g. “getpwuid: couldn’t determine user name from uid.”

- **crit**
  Critical condition—Any different from alert? e.g. “socket: Failed to get a socket, exiting child”

- **error**
  Error condition—effects a single transfer, not the system as a whole. e.g. “Premature end of script headers”
• **warn**
  Warning e.g “child process 1234 did not exit, sending another SIGHUP”

• **notice**
  Notice—Normal but significant condition. e.g “caught SIGTERM, shutting down”

• **info**
  Informational messages e.g “Server seems busy, (you may need to increase StartServers, or Min/Max SpareServers).”

• **debug**
  Debugging messages e.g “Opening config file /etc/httpd/conf/httpd.conf”
Pool of daemons
httpd.conf: Parameters for daemon pool

PidFile /var/run/httpd.pid
LockFile /var/lock/httpd.lock
ScoreBoardFile /var/run/httpd.scoreboard
Timeout 300
KeepAlive On
MaxKeepAliveRequests 100
KeepAliveTimeout 15
MinSpareServers 5
MaxSpareServers 20
StartServers 8
MaxClients 150
MaxRequestsPerChild 100
Apache’s functionality

• Our server has very little functionality.

• It serves all documents as “text/plain”.

• It can only log errors.

• We can add functionality as we need it.

• “Modules”
httpd.conf: Initialising the modules

# Start with an empty module list

ClearModuleList
AddModule mod_so.c
Syntax: Starting up the module system

• **ClearModuleList**
  Lose all information about modules in use.

• **AddModule mod_so.c**
  Use the *mod_so.c* module. Because it is built in to the binary we don’t need to specify the external file the module lives in.
httpd.conf: Following symbolic links

Options +FollowSymLinks
Syntax: Option suboptions for symbolic links

- Options +FollowSymLinks
  The web server will follow symbolic links.

- Options +SymLinksIfOwnerMatch
  The web server will follow symbolic links if the owner of the link (typically its creator) and the owner of the target of the link are the same.
httpd.conf: Adding support for MIME types

LoadModule mime_module modules/mod_mime.so
AddModule mod_mime.c

TypesConfig /etc/mime.types
DefaultType text/plain

AddEncoding x-compress Z
AddEncoding x-gzip gz tgz
Syntax: Loading an external module

- **LoadModule mime_module modules/mod_mime.so**
  This line says that the file `modules/mod_mime.so` (resolved relative to the ServerRoot definition at the start of the configuration file) contains a module called `mime_module`. This module is added to the list of modules that the server knows about. As yet the server won’t use the module; it just knows where to get it should it be called upon to use it.

- **AddModule mod_mime.c**
  This line tells the server to look through all the modules it knows about (either built-in or located with `LoadModule` directives) looking for a module whose original source file was called `mod_mime.c` (stupid, but that’s how they chose to do it) and activate it.
mod_mime: Directives

- **TypesConfig /etc/mime.types**
  Red Hat ships with a file called `/etc/mime.types` (part of the mailcap package) which identifies the file name extensions used for various MIME content types on the system. This line instructs the web server to use that file to identify MIME content types of files.

- **DefaultType text/plain**
  This says that if the server cannot determine the MIME content type of the file it is about to send then it should presume text/plain.

- **AddEncoding x-compress Z**
  This declares that any file whose name ends in “.Z” should be declared as having MIME encoding type “x-compress” (i.e. it is compressed) and the file name without the .Z suffix should be used to determine the underlying MIME content type.
<table>
<thead>
<tr>
<th>MIME Type</th>
<th>Extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>application/activemessage</td>
<td></td>
</tr>
<tr>
<td>application/andrew-inset</td>
<td>ez</td>
</tr>
<tr>
<td>application/applefile</td>
<td></td>
</tr>
<tr>
<td>application/mac-binhex40</td>
<td>hqx</td>
</tr>
<tr>
<td>application/octet-stream</td>
<td>bin dms lha lzh exe class</td>
</tr>
<tr>
<td>application/postscript</td>
<td>ai eps ps</td>
</tr>
<tr>
<td>application/x-dvi</td>
<td>dvi</td>
</tr>
<tr>
<td>application/x-javascript</td>
<td>js</td>
</tr>
<tr>
<td>image/gif</td>
<td>gif</td>
</tr>
<tr>
<td>image/jpeg</td>
<td>jpeg jpg</td>
</tr>
<tr>
<td>image/x-xwindowdump</td>
<td>xwd</td>
</tr>
<tr>
<td>message/partial</td>
<td></td>
</tr>
<tr>
<td>message/rfc822</td>
<td></td>
</tr>
<tr>
<td>model/vrml</td>
<td>wrl vrml</td>
</tr>
<tr>
<td>text/plain</td>
<td>asc txt</td>
</tr>
</tbody>
</table>
httpd.conf: Logging transfers

LoadModule config_log_module modules/mod_log_config.so
AddModule mod_log_config.c

HostnameLookups On
IdentityCheck Off

CustomLog /var/log/httpd/access_log "%t %h \"%r\" %>s %B"
mod_log_config: Directives

- CustomLog filename "format"
  Log to the file with the given format. Multiple log files may be defined.

- HostnameLookups On
  Convert IP addresses to hostnames.

- IdentityCheck On
  Do an ident lookup for each incoming request.
mod_log_config: Logging escape sequences

- `%t`: Time of the request
- `%h`: Remote hostname
- `%r`: First line of the request
- `%s`: Status code
- `%B`: Data bytes sent
<table>
<thead>
<tr>
<th>Status Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>OK</td>
</tr>
<tr>
<td>301</td>
<td>Moved Permanently</td>
</tr>
<tr>
<td>307</td>
<td>Temporary Redirect</td>
</tr>
<tr>
<td>400</td>
<td>Bad Request</td>
</tr>
<tr>
<td>401</td>
<td>Unauthorized</td>
</tr>
<tr>
<td>403</td>
<td>Forbidden</td>
</tr>
<tr>
<td>404</td>
<td>Not Found</td>
</tr>
<tr>
<td>500</td>
<td>Internal Server Error</td>
</tr>
<tr>
<td>505</td>
<td>HTTP Version Not Supported</td>
</tr>
</tbody>
</table>
mod_log_config: Common logging escape sequences

- `%a`: Client’s IP address
- `%B`: Bytes sent, excluding HTTP headers.
- `%f`: The name of the file served.
- `%h`: Remote hostname, or IP address is hostname lookups are off.
- `%l`: Remote logname from identd if IdentityCheck is on.
- `%r`: The first (typically only) line of the request.
- `%s`: Status code of the request.
- `%T`: Number of seconds taken to service the request.
• %t: Time of the request.

• %u: The URL requested.

• %u: The userid used if this is a page that requires userid/password.

• %{header}i: Argument of header in the incoming request

• %{header}o: Argument of header in the outgoing response
HTTP request headers

• Authorization: Access rights to restricted pages.

• From: E-mail address of the user making the request. (Often blank.)

• If-Modified-Since: Only send the data if necessary.

• Referer: The URL of the referring page.

• User-Agent: The web client. Many lie.
Some example log lines

[17/Apr/2000:10:10:25 +0100] hostname "GET /index.html HTTP/1.0" 200 1316
[17/Apr/2000:10:11:00 +0100] hostname "GET /bogus.html HTTP/1.0" 404 0
[17/Apr/2000:10:12:00 +0100] hostname "GET http://elsewhere/index.html HTTP/1.0" 200 1316
[17/Apr/2000:10:30:23 +0100] hostname "GET /cgi-bin/phf?Qalias=x%0a/bin/cat/%20/etc/passwd HTTP/1.0" 404 0
/etc/logrotate.conf

# rotate log files weekly
weekly

# keep 4 weeks worth of backlogs
rotate 4

# send errors to root
errors root

# create new (empty) log files after rotating old ones
create

# RPM packages drop log rotation information into this directory
include /etc/logrotate.d
/etc/logrotate.d/apache—as shipped

/var/log/httpd/access_log /var/log/httpd/error_log {
    missingok
    sharedscripts
    postrotate
        /bin/kill -
        HUP `cat /var/run/httpd.pid 2>/dev/null` 2> /dev/null || true
    endscript
}

/etc/logrotate.d/apache—as modified

/var/log/httpd/access_log /var/log/httpd/error_log {
    missingok
    sharedscripts
    create 0640 root webadmins
    postrotate
        /bin/kill -
        HUP `cat /var/run/httpd.pid 2>/dev/null` 2> /dev/null || true
    endscript
}
Resolving a URL to a file via an alias

http://server/wombat/index.html

/var/www/htdocs/wombat/index.html

http://server/icons/new.png

/var/www/htdocs/

/var/www/icons/new.png
httpd.conf: Aliases in Apache configuration

# Aliases

LoadModule alias_module modules/mod_alias.so
AddModule mod_alias.c

Alias /icons/ /var/www/icons/
Access log: Failing to read a directory

[27/Apr/2000:15:47:11 +0100] hostname "GET /index.html HTTP/1.0" 200 2537
[27/Apr/2000:15:48:09 +0100] hostname "GET / HTTP/1.0" 404 0
httpd.conf: Module for automatic indexing

# Automatic indexing of directory URLs

LoadModule autoindex_module modules/mod_autoindex.so
AddModule mod_autoindex.c

Options +Indexes
Browser’s view of automatic indexing

Index of /
* Parent Directory
* index.html
* poweredby.png
httpd.conf: Fancy indexing

IndexOptions +FancyIndexing
Browser’s view of fancy indexing

Index of /

<table>
<thead>
<tr>
<th>Name</th>
<th>Last modified</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent Directory</td>
<td>25-Apr-2000 14:00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>index.html</td>
<td>25-Apr-2000 18:08</td>
<td>2k</td>
<td></td>
</tr>
<tr>
<td>poweredby.png</td>
<td>01-Mar-2000 18:37</td>
<td>1k</td>
<td></td>
</tr>
</tbody>
</table>
httpd.conf: Fancy indexing options

IndexOptions +SuppressLastModified +ScanHTMLTitles
# Browser’s view of fancy indexing options

## Index of /

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent Directory</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>index.html</td>
<td>2k</td>
<td>Test Page for the Apache Web Server on Re&gt;poweredby.png</td>
</tr>
<tr>
<td>poweredby.png</td>
<td>1k</td>
<td></td>
</tr>
</tbody>
</table>
httpd.conf: Adding icons to the fancy listing

IndexOptions IconWidth IconHeight

AddIconByType (HTM,/icons/layout.gif) text/html
AddIconByType (TXT,/icons/text.gif) text/*
AddIconByType (IMG,/icons/image2.gif) image/*
AddIconByType (MOD,/icons/world2.gif) model/*
AddIconByType (SND,/icons/sound2.gif) audio/*
AddIconByType (VID,/icons/movie.gif) video/*
httpd.conf: Application subtypes

AddIconByType (_PS,/icons/a.gif) application/postscript
AddIconByType (PDF,/icons/a.gif) application/pdf
AddIconByType (HQX,/icons/binhex.gif) application/mac-binhex40
AddIconByType (DVI,/icons/dvi.gif) application/x-dvi
AddIconByType (TEX,/icons/tex.gif) application/x-tex
AddIconByType (TAR,/icons/tar.gif) application/x-tar
AddIconByType (BIN,/icons/binary.gif) application/octet-stream
AddIconByType (XXX,/icons/unknown.gif) application/*
httpd.conf: Directories

AddIcon (_UP,/icons/back.gif)  ..
AddIcon (DIR,/icons/folder.gif)  ^^DIRECTORY^^
AddIcon (---,/icons/blank.gif)  ^^BLANKICON^^
# Browser’s view of a fully labelled web page

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[_UP] Parent Directory</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>[HTM] index.html</td>
<td>2k</td>
<td>Test Page for the Apache Web Server on Re&gt;</td>
</tr>
<tr>
<td>[DIR] manual/</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>[IMG] poweredby.png</td>
<td>1k</td>
<td></td>
</tr>
</tbody>
</table>
mod_autoindex: IndexOptions suboptions

- **FancyIndexing**: Turns on the “long” format.

- **ScanHTMLTitles**: Display the HTML title or web pages as their description. This can be intensive on the disc.

- **SuppressDescription**: Turn off the description column altogether.

- **SuppressLastModified**: Turn off the column for the last modification date and time.

- **SuppressSize**: Turn off the column for the size of documents.

- **IconWidth [=X]**: Specify the width of all the icons in pixels (defaults to 20).
• **IconHeight [=Y]**: Specify the height of all the icons in pixels (defaults to 22).

• **NameWidth=X**: Width in characters of the file name column. An asterisk means “as wide as the widest element”.

• **DescriptionWidth=Y**: Width in characters of the “description” or “title scan” column. An asterisk means that the whole row should be 79 characters wide.
httpd.conf: Headers and footers

HeaderName HEADER.html
ReadmeName README.html
<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[_UP] Parent Directory</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>[HTM] HEADER.html</td>
<td>1k</td>
<td></td>
</tr>
<tr>
<td>[HTM] README.html</td>
<td>1k</td>
<td></td>
</tr>
<tr>
<td>[HTM] index.html</td>
<td>2k</td>
<td>Test Page for the Apache Web Server on R</td>
</tr>
<tr>
<td>[DIR] manual/</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>[IMG] poweredby.png</td>
<td>1k</td>
<td></td>
</tr>
</tbody>
</table>
httpd.conf: Suppressing files from the listing

IndexIgnore .??* *~ *# HEADER* README* SCCS RCS CVS
httpd.conf: Default files

# Default files in directory URLs

LoadModule dir_module modules/mod_dir.so
AddModule mod_dir.c

DirectoryIndex index.html index.htm
httpd.conf: Setting the 404 error document

ErrorDocument 404 /errors/404.html
ErrorDocument 500 "Oops, server goof."
Syntax: Specifying error messages

- ErrorDocument nnn "text": If the server generates status code nnn then a text/plain page will be returned with that status code and text as the text.

- ErrorDocument nnn URL: If the server generates status code nnn then the local web page at URL will be returned along with status code nnn.
Faking a browser with telnet

$ telnet draig.csi.cam.ac.uk 80
Trying 131.111.10.224...
Connected to draig.csi.cam.ac.uk.
Escape character is '^]'.
GET / HTTP/1.0

HTTP/1.1 200 OK
Date: Tue, 16 May 2000 08:54:29 GMT
Server: Apache/1.3.12 (Unix) (Red Hat/Linux)
Last-Modified: Tue, 25 Apr 2000 17:08:10 GMT
ETag: "f242-9e9-3905d0fa"
Content-Length: 2537
Connection: close
Content-Type: text/html

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Final//EN">
<HTML>
  <HEAD>
    ...
  </HEAD>
  </BODY>
</HTML>
HTTP response headers

- **HTTP/1.1 200 OK**: The HTTP protocol version number (our query was version 1.0 but the server is entitled to reply with version 1.1), followed by the status code and a text explanation of the status code.

- **Date**: The timestamp of the response.

- **Server**: A description of the responding server.

- **Last-Modified**: When the page was last modified.

- **ETag**: “Entity tag”: a key used to uniquely identify this version of the page for caches etc.

- **Content-Length**: Number of bytes in the body of the response. (i.e. the HTML page, but not the HTTP headers.)
• **Connection**: Whether the TCP connection should be kept open after this transfer to allow further requests.

• **Content-Type**: The MIME content type of the following document

• **Blank line**: The separator between the headers and the body of the web page.
Adding the mod_asis module

# Send .asis files "as is"

AddType httpd/send-as-is asis

LoadModule asis_module modules/mod_asis.so
AddModule mod_asis.c
A plausible index.asis file

Status: 403 Directory searching is prohibited
Content-Type: text/html

<!DOCTYPE HTML PUBLIC
  "-//W3C//DTD HTML 4.0 Transitional//EN"
  "http://www.w3.org/TR/REC-html40/strict.dtd">

<html>
  <head>
    <title>Security policy violation</title>
  </head>
  <body>
    <h1>Security policy violation</h1>
    <p>This web site’s security policy prohibits the autoindexing of this directory. Your request has been logged.</p>
  </body>
</html>
Faking a browser with telnet again

$ telnet draig.csi.cam.ac.uk 80
GET /two/ HTTP/1.0

Trying 131.111.10.224...
Connected to draig.csi.cam.ac.uk.
Escape character is '^]'.
Connection closed by foreign host.
HTTP/1.1 403 Directory searching is prohibited
Date: Tue, 16 May 2000 11:30:40 GMT
Server: Apache/1.3.12 (Unix) (Red Hat/Linux)
Connection: close
Content-Type: text/html

<!DOCTYPE HTML PUBLIC
    "-//W3C//DTD HTML 4.0 Transitional//EN"
    "http://www.w3.org/TR/REC-html40/strict.dtd">
<HTML><HEAD>
<TITLE>Security policy violation</TITLE>
</HEAD><BODY>
<H1>Security policy violation</H1>
<P>This web site's security policy prohibits the autoindexing of this directory. Your request has been logged.</P>
</BODY></HTML>
httpd.conf: User directories

# Users’ web pages

LoadModule userdir_module modules/mod_userdir.so
AddModule mod_userdir.c

UserDir public_html
user_dir: Remapping http://server/~user/index.html

- UserDir public_html
  Maps URL to ~/user/public_html/index.html.

- UserDir /home/userpages
  Maps URL to /home/userpages/user/index.html.

- UserDir /home/*/webstuff
  Maps URL to /home/user/webstuff/index.html.

- UserDir http://other/home/userpages
  Maps URL to http://other/home/userpages/user/index.html

- UserDir http://other/*/webstuff
  Maps URL to http://other/user/webstuff/index.html
A simple restriction example

- By default:
  - `index.html` files to be respected.
  - Automatic indexing permitted.

- Under `/var/www/html/fubar/`:
  - `index.html` files to be respected.
  - Automatic indexing forbidden.
httpd.conf: Restricting options to subdirectories

# Default
Options +Indexes

# Subdirectory restriction
<Directory /var/www/html/fubar/>
Options -Indexes
</Directory>
httpd.conf: Delegation of (some) control

AccessFileName .config

<Directory /var/www/html>
AllowOverride AuthConfig FileInfo Indexes
</Directory>
Core functionality: Delegation of (some) control

- **AccessFileName** `fname`
  Within the document tree the a file `fname` will override the default behaviour with the behaviour specified within (insofar as is permitted).

- **AllowOverride** `suboptions`
  This directive specifies exactly what aspects of the configuration may and may not be overridden in the files named by the AccessFileName directive.
Core functionality: AllowOverride suboptions

- AuthConfig
  Control the mechanisms used for authenticating users for access to restricted documents. See the section on access control for more on this option.

- FileInfo
  This permits the use of the directives found in the MIME module to change or add MIME types.

- Indexes
  This permits the use of the directives found in the two directory modules.
- **Options**
  Allow the use of the Options directive in the delegated control files.

- **All**
  Permit all overrides.

- **None**
  Permit no overrides. Ignore the delegated control files.
httpd.conf: Restricting options to subdirectories

# Default
Options +Indexes
AccessFileName .config
<Directory /var/www/html>
AllowOverride Options
</Directory>
/var/www/html/fubar/.config contents

Options -Indexes
httpd.conf: Access restrictions

# Access control by IP address

LoadModule access_module modules/mod_access.so
AddModule mod_access.c

order deny,allow
allow from .csi.cam.ac.uk
deny from all
allow from .csx.cam.ac.uk
Request from randompc.example.com

1. Initial state: Access allowed
2. deny from all: Access denied
3. allow from .csi.cam.ac.uk: Inapplicable—No change
4. allow from .csx.cam.ac.uk: Inapplicable—No change
5. Final state: Access denied
Request from ghoul.csi.cam.ac.uk

1. Initial state: Access allowed
2. deny from all: Access denied
3. allow from .csi.cam.ac.uk: Applicable—Access allowed
4. allow from .csx.cam.ac.uk: Inapplicable—No change
5. Final state: Access allowed
mod_access: allow directives

• order deny,allow

1. Initially all access allowed,
2. then apply all deny lines,
3. then apply all allow lines.

• order allow,deny

1. Initially all access denied,
2. then apply all allow lines,

3. then apply all deny lines.

- allow from all

- All requests are allowed.

- allow from host.inst.cam.ac.uk

- Requests from the host are allowed. Requires HostnameLookups On.

- allow from .inst.cam.ac.uk
• requests from hosts within the domain are allowed. Requires Host-nameLookups On.

• allow from 131.111.11.84

• Requests from the host are permitted.

• allow from 131.111.11.0/255.255.255.0

• Requests from any IP address starting 131.111.11. are allowed.

• allow from 131.111.11.0/24
• Requests from any IP address starting 131.111.11. are allowed. (The first three numbers correspond to the first 24 bits of the IP address quoted.)
mod_access: deny directives

• deny from ...

• As per allow from ...
httpd.conf: Restricting access to authenticated users

LoadModule auth_module modules/mod_auth.so
AddModule mod_auth.c

<Directory /var/www/html/restricted>
  AuthType Basic
  AuthName wombat
  AuthUserFile /etc/httpd/conf/passwd
  require valid-user
</Directory>
$ touch /etc/httpd/conf/passwd
$ ls -l /etc/httpd/conf/passwd
-rw-rw-r-- 1 root webadmin 0 Jun 1 10:12 passwd
$ htpasswd /etc/httpd/conf/passwd demouser
New password: demouser
Re-type new password: demouser
Adding password for user demouser
Basic authentication uncovered—1

$ telnet hydra.csi.cam.ac.uk 80
Trying 131.111.11.148...
Connected to hydra.csi.cam.ac.uk.
Escape character is '^['.
GET /restricted/ HTTP/1.0

HTTP/1.1 401 Authorization Required
Date: Thu, 01 Jun 2000 10:29:37 GMT
Server: Apache/1.3.12 (Unix) (Red Hat/Linux)
WWW-Authenticate: Basic realm="wombat"
Connection: close
Content-Type: text/html; charset=iso-8859-1

...
Basic authentication uncovered—2

$ telnet hydra.csi.cam.ac.uk 80
Trying 131.111.11.148...
Connected to hydra.csi.cam.ac.uk.
Escape character is ‘^]’.
GET /restricted/ HTTP/1.0
Authorization: Basic ZGVtb3VzZXI6ZGVtMHVzZXI=

HTTP/1.1 200 OK
Date: Thu, 01 Jun 2000 11:09:15 GMT
Server: Apache/1.3.12 (Unix) (Red Hat/Linux)
Last-Modified: Thu, 01 Jun 2000 10:28:10 GMT
ETag: "6b543-144-39363aba"
Accept-Ranges: bytes
Content-Length: 324
Connection: close
Content-Type: text/html
ID-based access restriction logic

- Authenticate the ID
- Is the ID allowed access?
An example /etc/httpd/conf/passwd file

demouser:RGMhGsfmvLQeE
bob:ylxjJ83Fx7p8E
tom:C6QeAIpNqz9IE
dick:yfPWrksACScys
harry:tXFkoaIYJqbrk
A more refined access control

- /var/www/html/restricted/alpha: Any valid user
- /var/www/html/restricted/beta: tom, dick, harry
- /var/www/html/restricted/gamma: bob, tom
httpd.conf: Finer grained access control

LoadModule auth_module modules/mod_auth.so
AddModule mod_auth.c

DIRECTORY /var/www/html/restricted
AuthType Basic
AuthName wombat
AuthUserFile /etc/httpd/conf/passwd

DIRECTORY /var/www/html/restricted/alpha
require valid-user

DIRECTORY /var/www/html/restricted/beta
require user tom dick harry
<Directory /var/www/html/restricted/gamma>
require user bob tom
</Directory>
httpd.conf: Access control by groups

LoadModule auth_module modules/mod_auth.so
AddModule mod_auth.c

<Directory /var/www/html/restricted>
AuthType Basic
AuthName wombat
AuthUserFile /etc/httpd/conf/passwd
AuthGroupFile /etc/http/conf/group
</Directory>

<Directory /var/www/html/restricted/alpha>
require valid-user
</Directory>

<Directory /var/www/html/restricted/beta>
require group betagrp
</Directory>
</Directory>

DIRECTORY /var/www/html/restricted/gamma
require group gammagrp
</Directory>
An example `/etc/httpd/conf/group` file

betagrp: tom dick harry
gammagrp: bob tom
mod_auth: Directives

- AuthType Basic: Specifies the “basic” authentication mechanism.

- AuthName realm: Specifies the “security realm”.

- AuthUserFile file: Specifies the web ID password file.

- AuthGroupFile file: Specifies the web group file.

- require valid-user: Any authenticated ID may have access.

- require user user1 user2: ID must be authenticated and be one of user1 or user2 to have access.
• require group grp1 grp2: ID must be authenticated and be in group grp1 or grp2 to have access
HTTP request headers

GET / HTTP/1.0
Connection: Keep-Alive
User-Agent: Mozilla/4.72 [en] (X11; U; Linux 2.2.14-6.1.1 i686)
Host: hydra.csi.cam.ac.uk
Accept: image/gif, image/x-xbitmap, image/jpeg, image/pjpeg, image/png, */*
Accept-Encoding: gzip
Accept-Language: es, en
Accept-Charset: iso-8859-1,*,utf-8
DNS entries

www-uxsup.csx.cam.ac.uk. 1D IN CNAME nymph.csi.cam.ac.uk.
nymph.csi.cam.ac.uk. 1D IN A 131.111.10.245
httpd.conf: Setting up a virtual host

# Virtual host example
<VirtualHost cockatrice.csi.cam.ac.uk>
    DocumentRoot /var/www/cock
</VirtualHost>