[Table of Contents](#)[Index](#)[Reviews](#)[Reader Reviews](#)[Errata](#)[Academic](#)

## Apache Cookbook

By [Rich Bowen](#), [Ken Coar](#)

[Start Reading](#) ▶

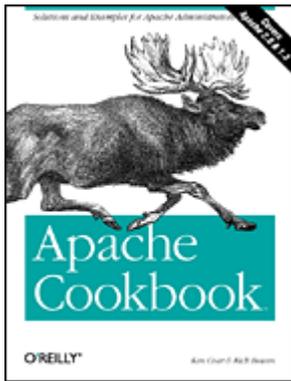
Publisher: O'Reilly

Pub Date: November 2003

ISBN: 0-596-00191-6

Pages: 254

The Apache Cookbook is a collection of problems, solutions, and practical examples for webmasters, web administrators, programmers, and everyone else who works with Apache. Instead of poking around mailing lists, online documentation, and other sources, you can rely on the Apache Cookbook for quick solutions to common problems, and then you can spend your time and energy where it matters most.



[Table of Contents](#)

[Index](#)

[Reviews](#)

[Reader Reviews](#)

[Errata](#)

## Apache Cookbook

By [Rich Bowen](#), [Ken Coar](#)

[Start Reading](#) ▶

Publisher: O'Reilly

Pub Date: November 2003

ISBN: 0-596-00191-6

Pages: 254

[Copyright](#)

[Preface](#)

[What's in This Book](#)

[Platform Notes](#)

[Other Books](#)

[Other Sources](#)

[How This Book Is Organized](#)

[Conventions Used in This Book](#)

[We'd Like to Hear from You](#)

[Acknowledgments](#)

[Chapter 1. Installation](#)

[Recipe 1.1. Installing from Red Hat Linux's Packages](#)

[Recipe 1.2. Installing Apache on Windows](#)

[Recipe 1.3. Downloading the Apache Sources](#)

[Recipe 1.4. Building Apache from the Sources](#)

[Recipe 1.5. Installing with ApacheToolbox](#)

[Recipe 1.6. Starting, Stopping, and Restarting Apache](#)

[Recipe 1.7. Uninstalling Apache](#)

[Chapter 2. Adding Common Modules](#)

[Recipe 2.1. Installing a Generic Third-Party Module](#)

[Recipe 2.2. Installing mod\\_dav on a Unixish System](#)

[Recipe 2.3. Installing mod\\_dav on Windows](#)

[Recipe 2.4. Installing mod\\_perl on a Unixish System](#)

[Recipe 2.5. Installing mod\\_php on a Unixish System](#)

[Recipe 2.6. Installing mod\\_php on Windows](#)

[Recipe 2.7. Installing the mod\\_snake Python Module](#)

[Recipe 2.8. Installing mod\\_ssl](#)

## [Chapter 3. Logging](#)

[Recipe 3.1. Getting More Details in Your Log Entries](#)

[Recipe 3.2. Getting More Detailed Errors](#)

[Recipe 3.3. Logging POST Contents](#)

[Recipe 3.4. Logging a Proxied Client's IP Address](#)

[Recipe 3.5. Logging Client MAC Addresses](#)

[Recipe 3.6. Logging Cookies](#)

[Recipe 3.7. Not Logging Image Requests from Local Pages](#)

[Recipe 3.8. Logging Requests by Day or Hour](#)

[Recipe 3.9. Rotating Logs on the First of the Month](#)

[Recipe 3.10. Logging Hostnames Instead of IP Addresses](#)

[Recipe 3.11. Maintaining Separate Logs for Each Virtual Host](#)

[Recipe 3.12. Logging Proxy Requests](#)

[Recipe 3.13. Logging Errors for Virtual Hosts to Multiple Files](#)

[Recipe 3.14. Logging Server IP Addresses](#)

[Recipe 3.15. Logging the Referring Page](#)

[Recipe 3.16. Logging the Name of the Browser Software](#)

[Recipe 3.17. Logging Arbitrary Request Header Fields](#)

[Recipe 3.18. Logging Arbitrary Response Header Fields](#)

[Recipe 3.19. Logging Activity to a MySQL Database](#)

[Recipe 3.20. Logging to syslog](#)

[Recipe 3.21. Logging User Directories](#)

## [Chapter 4. Virtual Hosts](#)

[Recipe 4.1. Setting Up Name-Based Virtual Hosts](#)

[Recipe 4.2. Designating One Name-Based Virtual Host as the Default](#)

[Recipe 4.3. Setting Up Address-Based Virtual Hosts](#)

[Recipe 4.4. Creating a Default Address-Based Virtual Host](#)

[Recipe 4.5. Mixing Address-Based and Name-Based Virtual Hosts](#)

[Recipe 4.6. Mass Virtual Hosting with mod\\_vhost\\_alias](#)

[Recipe 4.7. Mass Virtual Hosting Using Rewrite Rules](#)

[Recipe 4.8. SSL and Name-Based Virtual Hosts](#)

[Recipe 4.9. Logging for Each Virtual Host](#)

[Recipe 4.10. Splitting Up a LogFile](#)

[Recipe 4.11. Port-Based Virtual Hosts](#)

[Recipe 4.12. Displaying the Same Content on Several Addresses](#)

## [Chapter 5. Aliases, Redirecting, and Rewriting](#)

[Recipe 5.1. Showing Highlighted PHP Source Without Symlinking](#)

[Recipe 5.2. Mapping a URL to a Directory](#)

[Recipe 5.3. Creating a New URL for Existing Content](#)

[Recipe 5.4. Giving Users Their Own URL](#)

[Recipe 5.5. Aliasing Several URLs with a Single Directive](#)

[Recipe 5.6. Mapping Several URLs to the Same CGI Directory](#)

[Recipe 5.7. Creating a CGI Directory for Each User](#)

[Recipe 5.8. Redirecting to Another Location](#)

[Recipe 5.9. Redirecting Several URLs to the Same Destination](#)

[Recipe 5.10. Permitting Case-Insensitive URLs](#)

[Recipe 5.11. Replacing Text in Requested URLs](#)

[Recipe 5.12. Rewriting Path Information to CGI Arguments](#)

[Recipe 5.13. Denying Access to Unreferred Requests](#)

[Recipe 5.14. Rewriting Based on the Query String](#)

[Recipe 5.15. Redirecting All—or Part—of Your Server to SSL](#)

[Recipe 5.16. Turning Directories into Hostnames](#)

[Recipe 5.17. Redirecting All Requests to a Single Host](#)

[Recipe 5.18. Turning Document Names into Arguments](#)

## [Chapter 6. Security](#)

[Recipe 6.1. Using System Account Information for Web Authentication](#)

[Recipe 6.2. Setting Up Single-Use Passwords](#)

[Recipe 6.3. Expiring Passwords](#)

[Recipe 6.4. Limiting Upload Size](#)

[Recipe 6.5. Restricting Images from Being Used Off-Site](#)

[Recipe 6.6. Requiring Both Weak and Strong Authentication](#)

[Recipe 6.7. Managing .htpasswd Files](#)

[Recipe 6.8. Making Password Files for Digest Authentication](#)

[Recipe 6.9. Relaxing Security in a Subdirectory](#)

[Recipe 6.10. Lifting Restrictions Selectively](#)

[Recipe 6.11. Authorizing Using File Ownership](#)

[Recipe 6.12. Storing User Credentials in a MySQL Database](#)

[Recipe 6.13. Accessing the Authenticated Username](#)

[Recipe 6.14. Obtaining the Password Used to Authenticate](#)

[Recipe 6.15. Preventing Brute-Force Password Attacks](#)

[Recipe 6.16. Using Digest Versus Basic Authentication](#)

[Recipe 6.17. Accessing Credentials Embedded in URLs](#)

[Recipe 6.18. Securing WebDAV](#)

[Recipe 6.19. Enabling WebDAV Without Making Files Writable by the Web User](#)

[Recipe 6.20. Restricting Proxy Access to Certain URLs](#)

[Recipe 6.21. Protecting Files with a Wrapper](#)

[Recipe 6.22. Protecting All Files Except a Subset](#)

[Recipe 6.23. Protecting Server Files from Malicious Scripts](#)

[Recipe 6.24. Setting Correct File Permissions](#)

[Recipe 6.25. Running a Minimal Module Set](#)

[Recipe 6.26. Restricting Access to Files Outside Your Web Root](#)

[Recipe 6.27. Limiting Methods by User](#)

[Recipe 6.28. Restricting Range Requests](#)

## [Chapter 7. SSL](#)

[Recipe 7.1. Installing SSL](#)

[Recipe 7.2. Generating SSL Certificates](#)

[Recipe 7.3. Generating a Trusted CA](#)

[Recipe 7.4. Serving a Portion of Your Site via SSL](#)

[Recipe 7.5. Authenticating with Client Certificates](#)

## [Chapter 8. Dynamic Content](#)

[Recipe 8.1. Enabling a CGI Directory](#)

[Recipe 8.2. Enabling CGI Scripts in Non-ScriptAliased Directories](#)

[Recipe 8.3. Using Windows File Extensions to Launch CGI Programs](#)

[Recipe 8.4. Using Extensions to Identify CGI Scripts](#)

[Recipe 8.5. Testing That CGI Is Set Up Correctly](#)

[Recipe 8.6. Reading Form Parameters](#)

[Recipe 8.7. Invoking a CGI Program for Certain Content Types](#)

[Recipe 8.8. Getting SSIs to Work](#)

[Recipe 8.9. Displaying Last Modified Date](#)

[Recipe 8.10. Including a Standard Header](#)

[Recipe 8.11. Including the Output of a CGI Program](#)

[Recipe 8.12. Running CGI Scripts as a Different User with suexec](#)

[Recipe 8.13. Installing a mod\\_perl Handler from CPAN](#)

[Recipe 8.14. Writing a mod\\_perl Handler](#)

[Recipe 8.15. Enabling PHP Script Handling](#)

[Recipe 8.16. Verifying PHP Installation](#)

## [Chapter 9. Error Handling](#)

- [Recipe 9.1. Handling a Missing Host Field](#)
- [Recipe 9.2. Changing the Response Status for CGI Scripts](#)
- [Recipe 9.3. Customized Error Messages](#)
- [Recipe 9.4. Providing Error Documents in Multiple Languages](#)
- [Recipe 9.5. Redirecting Invalid URLs to Some Other Page](#)
- [Recipe 9.6. Making Internet Explorer Display Your Error Page](#)
- [Recipe 9.7. Notification on Error Conditions](#)

## [Chapter 10. Proxies](#)

- [Recipe 10.1. Securing Your Proxy Server](#)
- [Recipe 10.2. Preventing Your Proxy Server from Being Used as an Open Mail Relay](#)
- [Recipe 10.3. Forwarding Requests to Another Server](#)
- [Recipe 10.4. Blocking Proxied Requests to Certain Places](#)
- [Recipe 10.5. Proxying mod\\_perl Content to Another Server](#)
- [Recipe 10.6. Configuring a Caching Proxy Server](#)
- [Recipe 10.7. Filtering Proxied Content](#)
- [Recipe 10.8. Requiring Authentication for a Proxied Server](#)

## [Chapter 11. Performance](#)

- [Recipe 11.1. Determining How Much Memory You Need](#)
- [Recipe 11.2. Benchmarking Apache with ab](#)
- [Recipe 11.3. Tuning Keepalive Settings](#)
- [Recipe 11.4. Getting a Snapshot of Your Site's Activity](#)
- [Recipe 11.5. Avoiding DNS Lookups](#)
- [Recipe 11.6. Optimizing Symbolic Links](#)
- [Recipe 11.7. Minimizing the Performance Impact of .htaccess Files](#)
- [Recipe 11.8. Disabling Content Negotiation](#)
- [Recipe 11.9. Optimizing Process Creation](#)
- [Recipe 11.10. Tuning Thread Creation](#)
- [Recipe 11.11. Caching Frequently Viewed Files](#)
- [Recipe 11.12. Sharing Load Between Servers Using mod\\_proxy](#)
- [Recipe 11.13. Distributing Load Evenly Between Several Servers](#)
- [Recipe 11.14. Caching Directory Listings](#)
- [Recipe 11.15. Speeding Up Perl CGI Programs with mod\\_perl](#)

## [Chapter 12. Miscellaneous Topics](#)

- [Recipe 12.1. Placing Directives Properly](#)
- [Recipe 12.2. Renaming .htaccess Files](#)
- [Recipe 12.3. Generating Directory/Folder Listings](#)
- [Recipe 12.4. Solving the "Trailing Slash" Problem](#)
- [Recipe 12.5. Setting the Content-Type According to Browser Capability](#)
- [Recipe 12.6. Handling Missing Host: Header Fields](#)
- [Recipe 12.7. Alternate Default Document](#)
- [Recipe 12.8. Setting Up a Default "Favicon"](#)

## [Appendix A. Using Regular Expressions in Apache](#)

- [Section A.1. What Directives Use Regular Expressions?](#)

## [Appendix B. Troubleshooting](#)

- [Section B.1. Troubleshooting Methodology](#)
- [Section B.2. Debugging the Configuration](#)
- [Section B.3. Debugging Premature End of Script Headers](#)
- [Section B.4. Common Problems on Windows](#)
- [Section B.5. Fixing Build-Time Error Messages](#)
- [Section B.6. Getting Server-Side Includes to Work](#)
- [Section B.7. Debugging Rewrites That Result in "Not Found" Errors](#)
- [Section B.8. .htaccess Files Having No Effect](#)
- [Section B.9. Address Already in Use](#)

## [Colophon](#)



# Copyright

Copyright 2004 Ken Coar and Rich Bowen.

Printed in the United States of America.

Published by O'Reilly & Associates, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly & Associates books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://safari.oreilly.com>). For more information, contact our corporate/institutional sales department: (800) 998-9938 or [corporate@oreilly.com](mailto:corporate@oreilly.com).

Nutshell Handbook, the Nutshell Handbook logo, and the O'Reilly logo are registered trademarks of O'Reilly & Associates, Inc. Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and O'Reilly & Associates, Inc. was aware of a trademark claim, the designations have been printed in caps or initial caps. The association between the image of a moose and the topic of Apache is a trademark of O'Reilly & Associates, Inc.

While every precaution has been taken in the preparation of this book, the publisher and authors assume no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

# Preface

The Apache web server is a remarkable piece of software. The basic package distributed by the Apache Software Foundation is quite complete and very powerful, and a lot of effort has gone into keeping it from suffering software bloat. One facet of the package makes it especially remarkable: it includes extensibility by design. In short, if the Apache package right out of the box does not do what you want, you can generally extend it so that it does. Dozens of extensions (called modules) are included as part of the package distributed by the Apache Software Foundation. And if one of these doesn't meet your needs, with several million users out there, there is an excellent chance someone else has already done your work for you, who has concocted a recipe of changes or enhancements to the server that will satisfy your requirements.

This book is a collection of these recipes. Its sources include tips from the firehose of the USENET newsgroups, the Apache FAQ, Apache-related mailing lists, mail containing "how-to" questions, questions and problems posed on IRC chat channels, and volunteered submissions.

All of the items in this book come from real-life situations, encountered either by us or by other people who have asked for our help. The topics range from basic compilation of the source code to complex problems involving the treatment of URLs that require SSL encryption.

We've collected more than a hundred different problems and their solutions, largely based on how often they occurred, and have grouped them roughly by subject as shown in [What's in This Book](#).

Primarily, these recipes are useful to webmasters who are responsible for the entire server; however, many are equally applicable to users who want to customize the behavior in their own web directories through the use of *.htaccess* files.

We've written Apache Cookbook to be a practical reference, rather than a theoretical discourse: reading it recipe by recipe, chapter by chapter, isn't going to reveal a plot ("Roy Fielding in the Library with an RFC!" [\[1\]](#)). It's intended to provide point solutions to specific problems, located through the table of contents or the index.

[1] An obscure reference to a board game called Clue and an obscure developer of HTTP.

## What's in This Book

Because much of the material in this book is drawn from question-and-answer discussions and consultations, we have tried to make it as complete as possible. Of course, this means that we have included "recipes" for some questions to which there are currently no satisfactory answers (or at least to our knowledge). This has not been done to tease, annoy, or frustrate you; such recipes are included to provide completeness, so that you will know those problems have been considered rather than ignored.

Very few problems remain insoluble forever, and these incomplete recipes are the ones that will receive immediate attention on the book's web site and in revisions of the book. If a reader has figured out a way to do something the book mentions but doesn't explain, or omits mentioning entirely, our research team can be notified, and that solution will go on the web site and in the next revision.

Who knows, you may be the one to provide such a solution!

## Platform Notes

The recipes in this book are geared toward two major platforms: Unixish ones (such as Linux, FreeBSD, and Solaris) and Windows. There are many that have no platform-specific aspects, and for those any mention of the underlying operating system or hardware is gratefully omitted. Due to the authors' personal preferences and experiences, Unixish coverage is more complete than that for the Windows platforms. However, contributions, suggestions, and corrections for Windows-specific recipes will be gladly considered for future revisions and inclusion on the web site.

## Other Books

There are a number of books currently in print that deal with the Apache web server and its operation. Among them are:

- Apache: The Definitive Guide, Third Edition (O'Reilly)
- Apache Unleashed (Macmillan)
- Apache Administrator's Handbook (Macmillan)

You can also keep an eye on a couple of web pages that track Apache titles:

- <http://Apache-Server.Com/store.html>
- [http://httpd.apache.org/info/apache\\_books.html](http://httpd.apache.org/info/apache_books.html)

## Other Sources

In addition to books, there is a wealth of information available online. There are web sites, mailing lists, and USENET newsgroups devoted to the use and management of the Apache web server. The web sites are limitless, but here are some active and useful sources of information.

- The `comp.infosystems.www.servers.unix` and `comp.infosystems.www.servers.ms-dos` USENET newsgroups. Although these aren't dedicated to Apache specifically, there is a lot of traffic concerning it, and experienced Apache users hang out here. If you don't have access to news, or know how to reach USENET, check out <http://groups.google.com/>.
- The Apache Today web site, run by Internet.Com. This site regularly lists articles about the web server and making the most of it. The URL is <http://ApacheToday.Com/>.
- The [users@httpd.apache.org](mailto:users@httpd.apache.org) mailing list is populated with people who have varying degrees of experience with the Apache software, and some of the Apache developers can be found there, too. Posting is only permitted to subscribed participants. To join the list, visit <http://httpd.apache.org/userslist.html>.
- The `#apache` IRC channel on the `irc.freenode.net` network—or on many other IRC networks, for that matter. However, your chances of encountering us are most likely on the freenode network.

We must point out that none of these is an "official" support medium for the web server. In fact, there is no "official" support path, since the software is largely developed by volunteers and is free. However, these informal support forums successfully answer many questions.

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# How This Book Is Organized

This book is broken up into twelve chapters and two appendixes, as follows:

[Chapter 1](#) covers the basics of installing the vanilla Apache software, from source on Unixish systems, and on Windows from the Microsoft Software Installer (MSI) package built by the Apache developers.

[Chapter 2](#) describes the details of installing some of the most common third-party modules, and includes generic instructions that apply to many others that have less complex installation needs.

[Chapter 3](#) includes recipes about recording the visits to your web site(s), and Apache's error logging mechanism.

[Chapter 4](#) tells you how to run multiple web sites using a single Apache server and set of configuration files.

[Chapter 5](#) describes how to manipulate URLs, how to control which files they refer to, how to change them from one thing to another, and how to make them point to other web sites.

[Chapter 6](#) covers some of the basic issues of securing your Apache server against penetration and exposure by the nefarious elements on the Internet.

[Chapter 7](#) addresses the issues of making your Apache web server capable of handling secure transactions with SSL-capable browsers—a must if you're going to be handling sensitive data such as money transfers or medical records.

[Chapter 8](#) tells you how to enhance your server with runtime scripts and make them operate as a particular user.

[Chapter 9](#) describes how to customize the web server's error messages to give your site its own unique flavor.

[Chapter 10](#) describes how to configure your Apache server to act as a proxy between users and web pages and make the processes as transparent and seamless as possible.

[Chapter 11](#) includes a number of recipes for addressing performance bottlenecks and improving the overall function of your Apache server.

[Chapter 12](#) covers a variety of miscellaneous topics that didn't seem to fit into any of the other chapters.

[Appendix A](#) explains how regular expressions are used for pattern-matching in Apache directives.

[Appendix B](#) covers some basic troubleshooting techniques, where to look for messages, common configuration problems, etc.





# Conventions Used in This Book

Throughout this book certain stylistic conventions are followed. Once you are accustomed to them, you can easily distinguish between comments, commands you need to type, values you need to supply, and so forth.

In some cases, the typeface of terms in the main text will be different and likewise in code examples. The details of what the different styles (italic, boldface, etc.) mean are described in the following sections.

## Programming Conventions

In this book, most case examples of code will be in the form of excerpts from scripts, rather than actual application code. When commands need to be issued at a command-line prompt (such as an xterm for a Unixish system or a DOS command prompt for Windows), they will look something like this:

```
% find /usr/local -name apachectl -print  
# /usr/local/apache/bin/apachectl graceful  
C:>cd "\Program Files\Apache Group\Apache\bin"  
C:\Program Files\Apache Group\Apache\bin>apache -k stop
```

On Unixish systems, command prompts that begin with # indicate that you need to be logged in as the superuser (root username); if the prompt begins with %, then the command can be used by any user.

## Typesetting Conventions

The following typographic conventions are used in this book:

### Italic

Used for commands, filenames, abbreviations, citations of books and articles, email addresses, URLs, and USENET group names.

### Bold

Used for labeling menu choices in a graphical interface.

### Constant Width

Used for function names, command options, computer output, environment variable names, literal strings, and code examples.

### Constant Width Bold

Used for user input in computer dialogues and examples.



## We'd Like to Hear from You

We have tested and verified the information in this book to the best of our ability, but you may find that features have changed (which may in fact resemble bugs). Please let us know about any errors you find, as well as your suggestions for future editions, by writing to:

Please address comments and questions concerning this book to the publisher:

O'Reilly & Associates, Inc. 1005 Gravenstein Highway North Sebastopol, CA 95472 (800) 998-9938 (in the United States or Canada) (707) 829-0515 (international or local) (707) 829-0104 (fax)

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at:

<http://www.oreilly.com/catalog/apacheckbk>

To comment or ask technical questions about this book, send email to:

[bookquestions@oreilly.com](mailto:bookquestions@oreilly.com)

For more information about our books, conferences, Resource Centers, and the O'Reilly Network, see our web site at:

<http://www.oreilly.com>

We have a web site for the book, where we'll list errata and plans for future editions. Here you'll also find the source code from the book available for download so you don't have to type it all in:

<http://Apache-Cookbook.Com/>



# Acknowledgments

Originally, each recipe was going to be individually attributed, but that turned out to be logistically impossible.

Many people have helped us during the writing of this book, by posing a problem, providing a solution, proofreading, reviewing, editing, or just (!) providing moral support. This multitude, to each of whom we are profoundly grateful, includes Nat Torkington (our project editor and demonstrator of Herculean feats of patience), Sharco and Guy- from #apache on irc.freenode.net, Mads Toftum, Morbus Iff (known to the FBI under the alias Kevin Hemenway), and Andy Holman.

## Ken Coar

I dedicate this book to my father, Richard J. Coar, for all the various kinds of support he has provided over the years.

My sincere thanks go out to the crew at O'Reilly, including Paula and Nat, who apparently never gave up hope regardless of how many deadlines whooshed past unmet. Thanks also to the WriterBase Authors Support Group and Cabal mailing lists, whereon much balm and advice was offered. Our technical reviewers Morbus and Quasi provided much excellent feedback and helped make this a better book. The VMware software package was very useful, and I'm grateful to its makers for their help and support.

The people who have worked on the Apache web server documentation, and the people who develop the software itself, get a big note of thanks, too; without the former, collating a lot of the information in this book would have been a whole lot more difficult, and without the latter, the book wouldn't have happened at all.

The users of the software, whose frequently challenging questions populate the mailing lists, the IRC channels, and our inboxes, deserve thanks for all the inspiration they unwittingly provided for the recipes in this book.

But foremost among those to whom I owe gratitude is my significantly better half, Cathy, without whose patience, support, and constructive criticism I would never have achieved what I have.

## Rich Bowen

I dedicate this book to Sarah, who is the motivation for all the important things I do.

While it would be impossible to adequately thank all the folks who made this book possible, there are some people, in addition to those already mentioned, that I'd like to especially thank.

A huge thank you to our technical reviewers. You created a huge quantity of additional work for us, made our lives miserable, and made the book late. I don't think I've ever had such a thorough tech review, and the comments that you made served to make this a much better book than it otherwise would have been. Without quasi, large parts of the book would have been significantly less complete—and less correct. And Morbus gave the entire project the sort of surreal, other-worldly aura that you expect from...well, from Morbus. Thank you both and to the many other people who made various remarks about what we'd written.



# Chapter 1. Installation

For this cookbook to be useful, you need to install the Apache web server software. So what better way to start than with a set of recipes that deal with the installation?

There are many ways of installing this package; one of the features of open software like Apache is that anyone may make an installation kit. This allows vendors (such as Debian, FreeBSD, Red Hat, Mandrake, Hewlett-Packard, and so on) to customize the Apache file locations and default configuration settings so that these settings fit with the rest of their software. Unfortunately, one of the consequences of customization is that the various prepackaged installation kits are almost all different from one another.

In addition to installing it from a prepackaged kit, of which the variations are legion, there's always the option of building and installing it from the source yourself. This has both advantages and disadvantages; on the one hand you know exactly what you installed and where you put it, but on the other hand, it's likely that binary add-on packages will expect files to be in locations different than those you have chosen.

If setting up the web server is something you're going to do once and never again, using a packaged solution prepared by your system vendor is probably the way to go. However, if you anticipate applying source patches, adding or removing modules, or just fiddling with the server in general, building it yourself from the ground up is probably the preferred method. (The authors of this book, being confirmed bit-twiddlers, always build from source.)

This chapter covers some of the more common prepackaged installation varieties and also how to build them from the source yourself.

Throughout the chapter, we assume that you will be using *dynamic shared objects* (DSOs) rather than building modules statically into the server. The DSO approach is highly recommended; it not only makes it easy to update individual modules without having to rebuild the entire server, but it also makes adding or removing modules from the server's configuration a simple matter of editing the configuration file.

DSOs on Unixish systems typically have a `.so` extension; on Windows they end with a `.dll` suffix.



# Recipe 1.1 Installing from Red Hat Linux's Packages

## Problem

You have a Red Hat Linux server and want to install or upgrade the Apache web server on it using the packages that Red Hat prepares and maintains.

## Solution

If you are a member of the Red Hat Network (RHN), Red Hat's subscription service, you can use Red Hat's *up2date* tool to maintain your Apache package:

```
# up2date -ui apache apache-devel apache-manual
```

If you're running a more recent version:

```
# up2date -ui httpd httpd-devel httpd-manual
```

If you aren't a member of RHN, you can still download the packages from one of Red Hat's servers (either <ftp://ftp.redhat.com/> or <ftp://updates.redhat.com/>), and install it with the following command:

```
# rpm -Uvh apache
```

## Discussion

The *-Uvh* option to the *rpm* command tells it to:

- Upgrade any existing version of the package already on the system or install it for the first time if it isn't.
- Explain the process, so you can receive positive feedback that the installation is proceeding smoothly.
- Display a pretty line of octothorpes (#) across the screen, marking the progress of the installation.

If you use the packages Red Hat maintains for its own platform, you will benefit from a simple and relatively standard installation. However, you can only update versions that Red Hat has put together an RPM package for, which typically means that you may be lagging weeks to months behind the latest stable version.

There is also the issue of platform compatibility; for instance, at some point the version of Apache provided for Red Hat Linux changed from 1.3 to 2.0, and newer versions of the operating system will probably only have the 2.0 packages available. Similarly, if you run an older version of Red Hat Linux, the newer packages will probably not install properly on your system.

It's a good idea to install the *apache-devel* package as well. It's quite small, so it won't have much impact on your disk usage; however, it includes files and features that a lot of third-party modules will need to install properly.

## See Also

-





# Recipe 1.2 Installing Apache on Windows

## Problem

You want to install the Apache web server software on a Windows platform.



If you already have Apache installed on your Windows system, remove it before installing a new version. Failure to do this results in unpredictable behavior. See [Recipe 1.7](#).

## Solution

Primarily, Windows is a graphically oriented environment, so the Apache install for Windows is correspondingly graphical in nature.

The simplest way to install Apache is to download and execute the Microsoft Software Installer (MSI) package from the Apache web site at <http://httpd.apache.org/download>. The following screenshots come from an actual installation made using this method.

Each step of the installation process is distinct in the process and you can revise earlier decisions, until the files are installed. The first screen ([Figure 1-1](#)) simply confirms what you're about to do and the version of the package you're installing.

**Figure 1-1. First screen of Apache MSI install**



The second screen ([Figure 1-2](#)) presents the Apache license. Its basic tenets boil down to the following: do what you want with the software, don't use the Apache marks (trademarks like the feather or the name Apache) without permission, and provide proper attribution for anything you build based on Apache software. (This only applies if you plan to distribute your package; if you use it strictly on an internal network, this isn't required.) You can't proceed past this screen until you agree to the license terms.





# Recipe 1.3 Downloading the Apache Sources

## Problem

You want to build the Apache web server yourself from the sources directly (see [Recipe 1.4](#)), but don't know how to obtain them.

## Solution

There are a number of ways to obtain the sources. You can access the latest version in close to real-time by using CVS, the tool used by the Apache developers for source control, you can download a release tarball, or you can install a source package prepared by a distributor, among others.

From a prepackaged tarball, download the tarball from <http://httpd.apache.org/dist/>, and then:

```
% tar xzvf apache_1.3.27.tar.gz
```

If your version of tar doesn't support the z option for processing zipped archives, use this command instead:

```
% gunzip -c < apache_1.3.27.tar.gz | tar xvf -
```

From the very latest up-to-the-minute Apache 1.3 source repository (not guaranteed to be completely functional), use:

```
% cvs -d :pserver:anoncvs@CVS.Apache.Org:/home/cvspublic login
Password: anoncvs
% cvs -d :pserver:anoncvs@CVS.Apache.Org:/home/cvspublic checkout apache-1.3
```

You can fetch a particular release version instead of the bleeding edge code if you know the name the developers gave it. For example, this will pull the sources of the 1.3.27 release, which is expected to be stable, unlike the up-to-the-minute version:

```
% cvs -d :pserver:anoncvs@CVS.Apache.Org:/home/cvspublic checkout -r APACHE_1_3_27
apache-1.3
```

From the very latest up-to-the-minute Apache 2.0 source repository (not guaranteed to be completely functional), use:

```
% cvs -d :pserver:anoncvs@CVS.Apache.Org:/home/cvspublic login
Password: anoncvs
% cvs -d :pserver:anoncvs@CVS.Apache.Org:/home/cvspublic checkout -r
APACHE_2_0_BRANCH httpd-2.0
% cd httpd-2.0/src/lib
% cvs checkout apr apr-util
```

As with the method for the 1.3 version of the server, you can fetch a particular release of the 2.0 code if you know the name assigned to it in CVS.

You can find the names of the tags used in the source tree by visiting either <http://cvs.apache.org/viewcvs.cgi/apache-1.3/> or <http://cvs.apache.org/viewcvs.cgi/httpd-2.0/> and pulling down the Show files using tag list at the bottom of the page.



All sorts of tags are used by the developers for various purposes. The tags used to label versions of files used for a release are always of the form APACHE\_*n\_m\_e*, so use these





# Recipe 1.4 Building Apache from the Sources

## Problem

You want to build your Apache web server from the sources directly rather than installing it from a prepackaged kit.

## Solution

Assuming that you already have the Apache source tree, whether you installed it from a tarball, CVS, or some distribution package, the following commands, executed in the top directory of the tree, builds the server package with most of the standard modules as DSOs:

Apache 1.3:

```
% ./configure --prefix= /usr/local/apache --with-layout=Apache
    --enable-shared=max--enable-module=most
% make
# make install
```

For Apache 2.0:

```
% ./buildconf
% ./configure --prefix= /usr/local/apache --with-layout=Apache --enable-
    modules=most --enable-mods-shared=all
% make
# make install
```

If you want more detailed information about the various options and their meanings, you can use the following command:

```
% ./configure --help
```

## Discussion

Building the server from the sources can be complex and time-consuming, but it's required if you intend to make any changes to the source code. It gives you much more control over things, such as the use of shareable object libraries and the database routines available to modules. Building from source is also de rigeur if you're developing your own Apache modules.

The options to the *configure* script are many and varied; if you haven't used it before to build Apache, locate some online tutorials (such as those at <http://Apache-Server.Com/tutorials/> or <http://httpd.apache.org/docs-2.0/install.html>) when you want to change the defaults. The default options generally produce a working server, although the filesystem locations and module choices may not be what you'd like; they may include modules you don't want or omit some you do. (See [Chapter 2](#) for some examples.)

## See Also

- 

- [Recipe 1.3](#)

- 

- <http://Apache-Server.Com/tutorials/>

-





# Recipe 1.5 Installing with ApacheToolbox

## Problem

You have a complicated collection of modules you want to install correctly.

## Solution

Download ApacheToolbox from <http://www.apachetoolbox.com/>. (Note that the version numbers will probably be different than these, which were the latest available when this section was written.) Unpack the file:

```
% bunzip2 Apachetoolbox-1.5.65.tar.bz2
% tar xvf Apachetoolbox-1.5.65.tar
```

(Depending on your version of tar, you may be able to combine these operations into a single `tar xjvf` command.)

Then run the installation script:

```
# cd Apachetoolbox-1.5.65
# ./install.sh
```

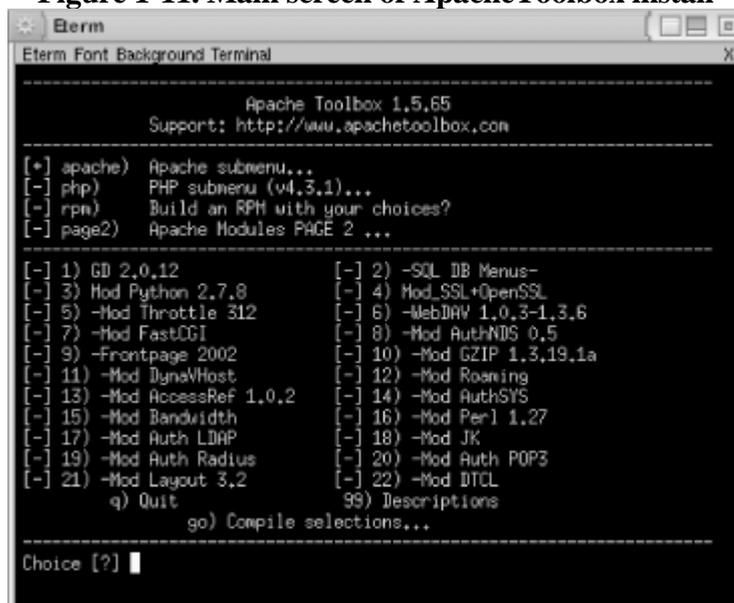
## Discussion

ApacheToolbox is developed and maintained by Bryan Andrews. It is a shell script that assists in the configuration and installation of Apache. It includes support for over 100 commonly used or standard modules.

When you run the script, you select modules from lists appearing on various screens. Once you have decided on your list of modules, ApacheToolbox downloads the third-party modules you have selected and the tools that you don't have installed, and then runs the Apache `configure` script with any arguments needed to create the combination you have requested.

The main screen (see [Figure 1-11](#)) lists the most popular third-party modules that ApacheToolbox can install. Select or deselect a particular module by typing the number next to that module's name.

**Figure 1-11. Main screen of ApacheToolbox install**







# Recipe 1.6 Starting, Stopping, and Restarting Apache

## Problem

You want to be able to start and stop the server at need, using the appropriate tools.

## Solution

On Unixish systems, use the *apachectl* script; on Windows, use the options in the Apache folder of the Start menu.

## Discussion

The basic Apache package includes tools to make it easy to control the server. For Unixish systems, this is usually a script called *apachectl*, but prepackaged distributions may replace or rename it. It can only perform one action at a time, and the action is specified by the argument on the command line. The options of interest are:

```
apachectl start
```

This will start the server if it isn't already running. If it is running, this option has no effect and may produce a warning message.

```
apachectl graceful
```

This option causes the server to reload its configuration files and gracefully restart its operation. Any current connections in progress are allowed to complete. The server will be started if it isn't running.

```
apachectl restart
```

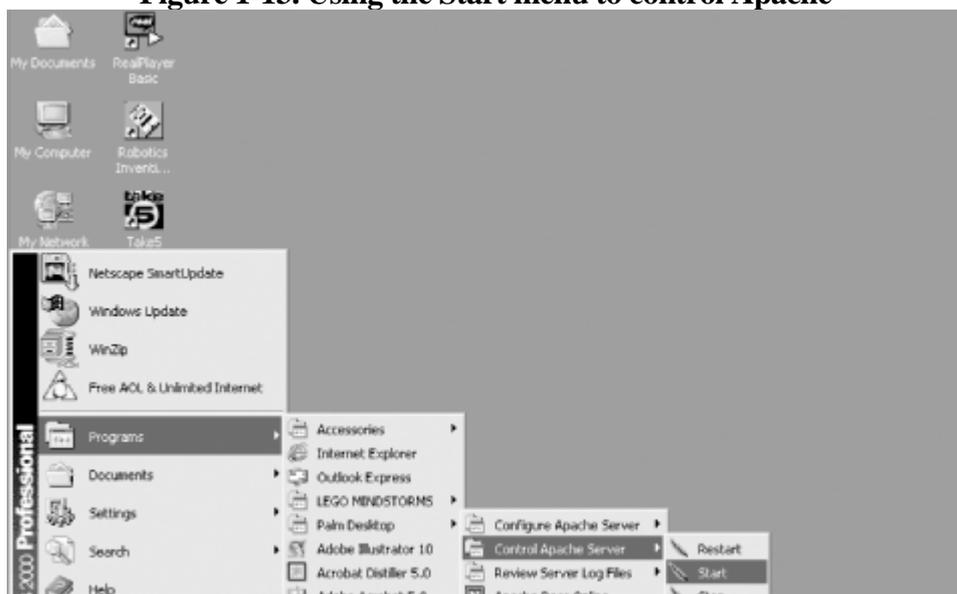
Like the *graceful* option, this one makes the server reload its configuration files. However, existing connections are terminated immediately. If the server isn't running, this command will try to start it.

```
apachectl stop
```

This shuts the server down immediately. Any existing connections are terminated at once.

For Windows, the MSI installation of Apache includes menu items for controlling the server, as shown in [Figure 1-13](#).

**Figure 1-13. Using the Start menu to control Apache**







# Recipe 1.7 Uninstalling Apache

## Problem

You have the Apache software installed on your system, and you want to remove it.

## Solution

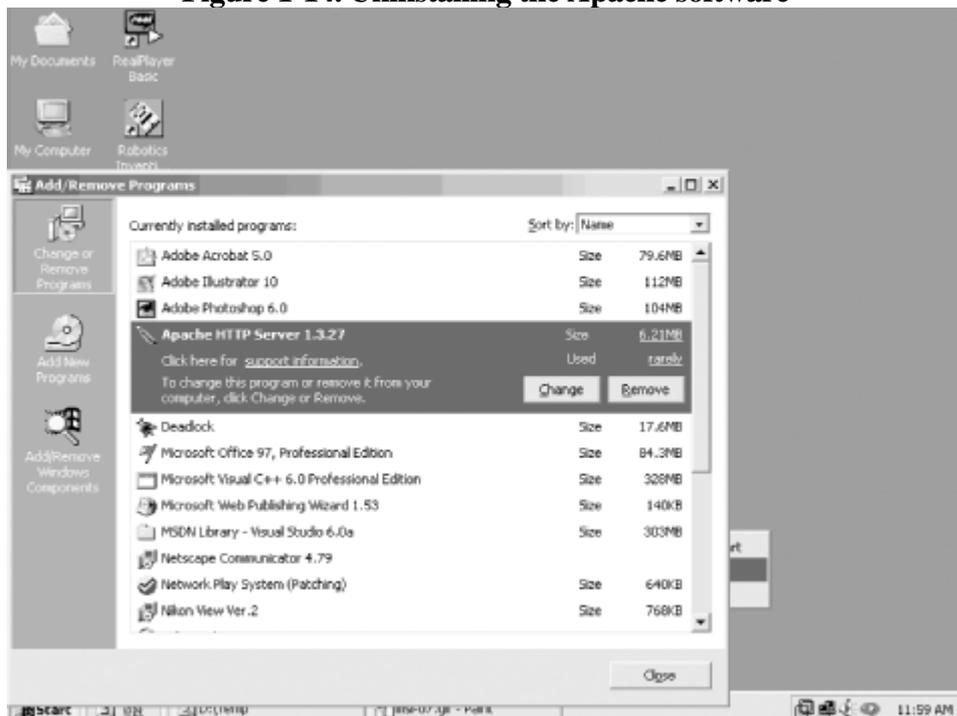
On Red Hat Linux, to remove an Apache version installed with the RPM tool, use:

```
# rpm -ev apache
```

Other packaging systems may provide some similar mechanism.

On Windows, Apache can be typically removed like any other MSI-installed software (see [Figure 1-14](#)).

**Figure 1-14. Uninstalling the Apache software**



## Discussion

Unfortunately, there is no generic works-for-all removal method for Apache installations on Unixish systems. Some packages, such as Red Hat's RPM, do remember what they installed so they can remove all the pieces, as shown in the solution. However, if the software was installed by building from the sources (see [Recipe 1.4](#)), the burden of knowing where files were put rests with the person who did the build and install. The same applies if the software was installed from source on a Windows system; it's only the MSI or InstallShield packages that make the appropriate connections to allow the use of the Add/Remove Software control panel.

## See Also

-



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Chapter 2. Adding Common Modules

There are a number of extremely popular modules for the Apache web server that are not included in the basic distribution. Most of these are separate because of licensing or support reasons; some are not distributed by the Apache Software Foundation because of a decision by the Apache developers; and some are integral parts of other projects. For instance, *mod\_ssl* for Apache 1.3 is developed and maintained separately not only because of the U.S. export control laws (which were more restrictive when the package was originally developed), but because it requires changes to the core software that the Apache developers chose not to integrate.

This chapter provides recipes for installing some of the most popular of these third-party modules; when available, there are separate recipes for installation on Unixish systems and on Windows.

The most comprehensive list of third-party modules can be found in the Apache Module Registry at <http://modules.apache.org/>. Some modules are so popular—or complex—that they have entire sites devoted to them, as do the ones listed in this chapter.

Although hundreds of third-party modules are available, many module developers are only concerned with their single module. This means that there are potentially as many different sets of installation instructions as there are modules. The first recipe in this chapter describes an installation process that should work with many Apache 1.3 modules, but you should check with the individual packages' instructions to see if they have a different or more detailed process.

Many of the modules are available from organizations that prepackage or distribute Apache software, such as in the form of an RPM from Mandrake or Red Hat, but such prebuilt module packages include the assumptions of the packager. In other words, if you build the server from source and use custom locations for the files, don't be surprised if the installation of a packaged module fails.

All of the modules described in this chapter are supported with Apache 1.3 on Unixish systems. Status of support with Apache 2.0 on Windows is shown in [Table 2-1](#).

Table 2-1. Module support status

Module name	Windows	Support on Apache 2.0
<i>mod_dav</i>	Yes	Included; no installation necessary
<i>mod_perl</i>	Yes	Yes
<i>mod_php</i>	Yes	Yes
<i>mod_snake</i>	No	Unknown



## Recipe 2.1 Installing a Generic Third-Party Module

### Problem

You have downloaded a third-party module that isn't listed in this chapter, and you want to install it.

### Solution

Move to the directory where the module's source file was unpacked, and then:

```
% /path/to/apache/ bin/apxs -cia module.c
```

### Discussion

In the case of a third-party module that consists of a single `.c` file, there is a good chance that it can be built and installed using the Solution. Modules that involve multiple source files should provide their own installation instructions.

The `-cia` options mean to compile, install, and activate. The first is pretty straightforward; install means put the `.so` file in the place Apache expects to find it, and activate means to add the module to the `httpd.conf` file.

### See Also

- 

The `apxs` manpage, typically `ServerRoot/man/man8/apxs.8`

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 2.2 Installing mod\_dav on a Unixish System

### Problem

You want to add or enable WebDAV capabilities to your server. WebDAV permits specific documents to be reliably and securely manipulated by remote users without the need for FTP, to perform such tasks as adding, deleting, or updating files.

### Solution

If you're using Apache 2.0, *mod\_dav* is automatically available, although you may need to enable it at compile time with `—enable-dav`.

If you are using Apache 1.3, download and unpack the *mod\_dav* source package from [http://webdav.org/mod\\_dav/](http://webdav.org/mod_dav/), and then:

```
% cd mod_dav-1.0.3-1.3.6
% ./configure --with-apxs=/usr/local/apache/bin/apxs
% make
# make install
```

Restart the server, and be sure to read [Recipe 6.18](#).

### Discussion

*mod\_dav* is an encapsulated and well-behaved module that is easily built and added to an existing server. To test that it has been properly installed, you need to enable some location on the server for WebDAV management and verify access to that location with some WebDAV-capable tool. We recommend *cadaver*, which is an open source command-line WebDAV tool. (The URL for the *cadaver* tool is found at the end of this recipe.)

To enable your server for WebDAV operations, you need to add at least two directives to your *httpd.conf* file. The first identifies the location of the locking database used by *mod\_dav* to keep WebDAV operations from interfering with each other; it needs to be in a directory that is writable by the server. For example:

```
# cd /usr/local/apache
# mkdir var
# chgrp nobody var
# chmod g+w var
```

Now add the following line to your *httpd.conf* file, outside any containers:

```
<IfModule mod_dav.c>
    DAVLockDB var/DAVlock
</IfModule>
```



The *DAVLockDB* location must not be on an NFS-mounted filesystem, because NFS doesn't support the sort of locking *mod\_dav* requires. Putting the lock database on an NFS filesystem may result in unpredictable results.

Next, create a temporary directory for testing WebDAV functionality:

```
# cd /usr/local/apache
```



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 2.3 Installing mod\_dav on Windows

### Problem

You want to enable WebDAV capabilities on your existing Apache 1.3 server with *mod\_dav*.

### Solution

Apache 2.0 includes *mod\_dav* as a standard module, so you do not need to download and build it.

Download and unpack the *mod\_dav* Windows package from [http://webdav.org/mod\\_dav/win32/](http://webdav.org/mod_dav/win32/). Verify that your Apache installation already has the *xmlparse.dll* and *xmltok.dll* files in the *ServerRoot* directory; if they aren't there, check through the Apache directories to locate and copy them to the *ServerRoot*. *mod\_dav* requires the Expat package, which is included with versions of the Apache web server after 1.3.9; these files hook into Expat, which *mod\_dav* will use.

Put the *mod\_dav* DLL file into the directory where Apache keeps its modules:

```
C:\>cd mod_dav-1.0.3-dev
C:\mod_dav-1.0.3-dev>copy mod_dav.dll C:\Apache\modules
C:\mod_dav-1.0.3-dev>cd \Apache
```

Add the following lines to your *httpd.conf* file:

```
LoadModule dav_module modules/mod_dav.dll
```

You may also need to add an *AddModule* line if your *httpd.conf* file includes a *ClearModuleList* directive and re-adds the other modules. Alternatively, you can insert the *LoadModule* for *mod\_dav* after the *ClearModuleList* directive.

### Discussion

*mod\_dav* is an encapsulated and well-behaved module that is easily built and added to an existing server. To test that it has been properly installed, you need to enable some location on the server for WebDAV management and verify access to that location with some WebDAV-capable tool, or browse to it in Windows Explorer, which knows how to access WebDAV locations (as of Windows 2000), or access it from a different system where cadaver or another WebDAV tool is available.

To enable your server for WebDAV operations, you need to add at least two directives to your *ServerRoot/conf/httpd.conf* file. The first identifies the location of the locking database used by *mod\_dav* to keep WebDAV operations from interfering with each other; it needs to be in a directory that is writable by the server. For example:

```
C:\Apache-1.3>mkdir var
```

Now add the following lines to your *httpd.conf* file to enable WebDAV:

```
<IfModule mod_dav.c>
    DAVLockDB "C:/Apache-1.3/var/dav-lock"
</IfModule>
```

Create a temporary directory for testing *mod\_dav*'s ability to function:

```
C:\Apache-1.3>mkdir htdocs\dav-test
```



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 2.4 Installing mod\_perl on a Unixish System

### Problem

You want to install the *mod\_perl* scripting module to allow better Perl script performance and easy integration with the web server.

### Solution

Download and unpack the *mod\_perl* source package from <http://perl.apache.org/>. Then use the following command:

```
% perl Makefile.PL \  
> USE_APXS=1 \  
> WITH_APXS= /usr/local/apache/bin/apxs \  
> EVERYTHING=1 \  
> PERL_USELARGEFILES=0  
% make  
% make install
```

Restart your server.

### Discussion

*mod\_perl* is quite a complex module, and there are several different ways to add it to your server. This recipe is the fastest and lowest-impact one; if it doesn't suit your needs, check the various *README.\** files in the package directory after unpacking. Because its primary language is Perl rather than C, the installation instructions are significantly different from those for most other modules.

Once you have restarted your server successfully, *mod\_perl* should be available and configured as part of it. You can test it by making some changes to the *httpd.conf* file, adding a few scripts, and seeing whether the server processes them correctly. Here is a sample set of steps to test *mod\_perl*'s operation.

1.

Create a directory where your *mod\_perl* scripts can live:

```
# cd ServerRoot  
# mkdir lib lib/perl lib/perl/Apache
```

2.

Create a file named *startup.pl* in your server's *conf/* directory that will give *mod\_perl* some startup instructions:

```
#!/usr/bin/perl  
BEGIN {  
    use Apache ( );  
    use lib Apache->server_root_relative('lib/perl');  
}  
use Apache::Registry ( );  
use Apache::Constants ( );  
use CGI qw(-compile :all);  
use CGI::Carp ( );  
1;
```

3.

Next, create the *lib/perl/Apache/HelloWorld.pm* file that will be used for our test:





## Recipe 2.5 Installing mod\_php on a Unixish System

### Problem

You want to add the *mod\_php* scripting module to your existing Apache web server.

### Solution

Download the *mod\_php* package source from the web site at <http://php.net/> (follow the links for downloading) and unpack it. Then:

```
% cd php-4.3.2
% ./configure \
>     --with-apxs= /usr/local/apache/bin/ apxs
% make
# make install
```

Restart the server.

### Discussion

To test that your installation was successful, create a file named *info.php* in your server's *DocumentRoot*; the file should contain the single line:

```
<?php phpinfo( ) ; ?>
```

Add the following lines to your server's *httpd.conf* file:

```
<IfModule mod_php4.c>
    AddType application/x-httpd-php .php
    AddHandler application/x-httpd-php .php
</IfModule>
```

After restarting your server, try fetching the document *info.php* using a browser. You should see a detailed description of the PHP options that are active.

There are numerous additional options and extensions available for PHP; the recipe given here is only for the most basic installation.

### See Also

- [Recipe 8.15](#)
- [Recipe 8.16](#)
- <http://php.net/>





## Recipe 2.6 Installing mod\_php on Windows

### Problem

You want to add the *mod\_php* scripting module to your existing Apache server on Windows.

### Solution

This recipe needs to be described largely in terms of actions rather than explicit commands to be issued.

1.

Download the PHP Windows binary *.zip* file with API extensions (not the *.exe* file) from <http://php.net/>.

2.

Unpack the *.zip* file into a directory where you can keep its contents indefinitely (such as *C:\PHP4*). If you use WinZip, be sure to select the Use folder names checkbox to preserve the directory structure inside the *.zip* file.

3.

Copy the *PHP4\SAPI\php4apache.dll* to the *\modules\* directory under your Apache installation's *ServerRoot*.

4.

In a command-prompt window, change to the *PHP4* directory where you unpacked the *.zip* file, and type:

```
... \PHP4>copy php.ini-dist %SYSTEMROOT%\php.ini  
... \PHP4>copy php4ts.dll %SYSTEMROOT%
```

(If installing on Windows 95 or Windows 98, use *%WINDOWS%* instead of *%SYSTEMROOT%*.)

5.

Edit the *%SYSTEMROOT%\php.ini* file, locate the line that starts with *extensions\_dir*, and change the value to point to the *PHP4\extensions* directory. For instance, if you unpacked the *.zip* file into *C:\PHP4*, this line should look like:

```
extensions_dir = C:\PHP4\extensions
```

6.

Edit the *conf\httpd.conf* file under the Apache *ServerRoot* and add the following lines near the other *LoadModule* lines:

```
LoadModule php4_module modules/php4apache.dll
```

Add the following lines in some scope where they will apply to your *.php* files:

```
<IfModule mod_php4.c>  
AddType application/x-httpd-php .php  
</IfModule>
```

7.

Restart the Apache server, and the PHP module should be active.

### Discussion

The PHP module installation on Windows requires a lot of nit-picky manual steps. To test that your installation was successful, create a file named *info.php* in your server's *DocumentRoot*; the file should contain the single line:



## Recipe 2.7 Installing the `mod_snake` Python Module

### Problem

You want to add the `mod_snake` Python scripting module to your existing Apache server.

### Solution

To install `mod_snake` on a Unixish system, download the source from the <http://modsnake.sourceforge.net/> web site (follow the link for downloading), unpack it, and then:

```
% cd mod_snake-0.5.0
% ./configure --with-apxs= /usr/local/apache/bin/apxs
% make
# make install
```

Restart the server.



At the time of this writing, `mod_snake` cannot be installed on Windows.

### Discussion

`mod_snake` is a fairly standard Apache module; it can be added to an existing Apache installation without any source-level changes. It requires that you have Python installed, and that the `apxs` script has been properly configured and installed as part of the web server package.

See the `README` and `INSTALL` files in the unpacked package directory, and the HTML documentation in the `examples/tut/` directory, for additional information, examples, and ways to test that your installation was successful.

### See Also

- 

[Recipe 2.1](#)

- 

<http://modsnake.sourceforge.net/>



## Recipe 2.8 Installing mod\_ssl

### Problem

You want to add SSL support to your Apache server with the *mod\_ssl* secure HTTP module.

### Solution

*Windows*

At the time of this writing, there is no supported means of installing *mod\_ssl* on Windows.

*Apache 2.0*

*mod\_ssl* is included with 2.0, although it is not automatically compiled nor installed when you build from source. You need to include the `--enable-ssl` option on your `./configure` line, and enable it with `LoadModule` and `AddModule` directives.

*Apache 1.3*

To install *mod\_ssl* on a Unixish system, download the tarball package from the <http://www.modssl.org/> web site and unpack it. Then:

```
% cd mod_ssl-2.8.14-1.3.27
% ./configure \
>   --with-apache=.. /apache_1.3.27 \
>   --with-ssl=SYSTEM \
>   --prefix= /usr/local/apache
% cd ../apache_1.3.27
% make
% make certificate
```

### Discussion

The *mod\_ssl* package requires source-level changes to the base Apache code, and so the version of the *mod\_ssl* package you install must match the version of the Apache distribution you have. If your Apache installation doesn't include the source (such as if you installed a binary-only RPM or other vendor distribution), you won't be able to add *mod\_ssl* to it.

In addition to the Apache source, *mod\_ssl* requires that you have Perl and the OpenSSL libraries installed. The `--with-ssl` option on the build configuration statement indicates where this is located; if it is in a vendor distributed directory, the special keyword `SYSTEM` tells the build to look for it, and you don't have to find it yourself.

Unlike most other Apache modules, when adding *mod\_ssl* you run the `./configure` script that's in *mod\_ssl*'s directory rather than the one in the Apache source directory; the module's script makes changes to Apache's and then invokes it directly.

This recipe is the bare basics; there are many optional components and features that *mod\_ssl* allows you to specify at configuration time. For more information, consult the `README` and `INSTALL` files in the *mod\_ssl* source directory, or the *mod\_ssl* web site at <http://www.modssl.org/>.

### See Also

-





# Chapter 3. Logging

Apache can, and usually does, record information about every request it processes. Controlling how this is done and extracting useful information out of these logs after the fact is at least as important as gathering the information in the first place.

The logfiles may record two types of data: information about the request itself, and possibly one or more messages about abnormal conditions encountered during processing (such as file permissions). You, as the webmaster, have a limited amount of control over the logging of error conditions, but a great deal of control over the format and amount of information logged about request processing (activity logging). The server may log activity information about a request in multiple formats in multiple log files, but it will only record a single copy of an error message.

One aspect of activity logging you should be aware of is that the log entry is formatted and written after the request has been completely processed. This means that the interval between the time a request begins and when it finishes may be long enough to make a difference.

For example, if your logfiles are rotated while a particularly large file is being downloaded, the log entry for the request will appear in the new logfile when the request completes, rather than in the old logfile when the request was started. In contrast, an error message is written to the error log as soon as it is encountered.

The web server will continue to record information in its logfiles as long as it's running. This can result in extremely large logfiles for a busy site and uncomfortably large ones even for a modest site. To keep the file sizes from growing ever larger, most sites rotate or roll over their logfiles on a semi-regular basis. Rolling over a logfile simply means persuading the server to stop writing to the current file and start recording to a new one. Due to Apache's determination to see that no records are lost, cajoling it to do this according to a specific timetable may require a bit of effort; some of the recipes in this chapter cover how to accomplish the task successfully and reliably (see [Recipe 3.8](#) and [Recipe 3.9](#)).

The log declaration directives, *CustomLog* and *ErrorLog*, can appear inside *<VirtualHost>* containers, outside them (in what's called the *main* or *global server*, or sometimes the *global scope*), or both. Entries will only be logged in one set or the other; if a *<VirtualHost>* container applies to the request or error and has an applicable log directive, the message will be written only there and won't appear in any globally declared files. On the other hand, if no *<VirtualHost>* log directive applies, the server will fall back on logging the entry according to the global directives.

However, whichever scope is used for determining what logging directives to use, all *CustomLog* directives in that scope are processed and treated independently. That is, if you have a *CustomLog* directive in the global scope and two inside a *<VirtualHost>* container, both of these will be used. Similarly, if a *CustomLog* directive uses the *env=* option, it has no effect on what requests will be logged by other *CustomLog* directives in the same scope.

Activity logging has been around since the Web first appeared, and it didn't take long for the original users to decide what items of information they wanted logged. The result is called the *common log format* (CLF). In Apache terms, this format is:

```
"%h %l %u %t \"%r\" %>s %b"
```

That is, it logs the client's hostname or IP address, the name of the user on the client (as defined by RFC 1413 and if



## Recipe 3.1 Getting More Details in Your Log Entries

### Problem

You want to add a little more detail to your access log entries.

### Solution

Use the *combined* log format, rather than the *common* log format:

```
CustomLog logs/access_log combined
```

### Discussion

The default Apache logfile enables logging with the *common* log format, but it also provides the *combined* log format as a predefined *LogFormat* directive.

The *combined* log format offers two additional pieces of information not included in the *common* log format: the Referer (where the client linked from) and the User-agent (what browser they are using).

Every major logfile parsing software package is able to handle the *combined* format as well as the *common* format, and many of them give additional statistics based on these added fields. So you lose nothing by using this format and potentially gain some additional information.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_log\\_config.html](http://httpd.apache.org/docs/mod/mod_log_config.html)

- 

[http://httpd.apache.org/docs-2.0/mod/mod\\_log\\_config.html](http://httpd.apache.org/docs-2.0/mod/mod_log_config.html)



## Recipe 3.2 Getting More Detailed Errors

### Problem

You want more information in the error log in order to debug a problem.

### Solution

Change (or add) the *LogLevel* line in your *httpd.conf* file. There are several possible arguments, which are enumerated below:

For example:

```
LogLevel Debug
```

### Discussion

There are several hierarchical levels of error logging available, each identified by its own keyword. The default value of *LogLevel* is *warn*. Listed in descending order of importance, the possible values are:

emerg

Emergencies; web server is unusable  
alert

Action must be taken immediately  
crit

Critical conditions  
error

Error conditions  
warn

Warning conditions  
notice

Normal but significant condition  
info

Informational  
debug

Debug-level messages

*emerg* results in the least information being recorded and *debug* in the most. However, at debug level a lot of information will probably be recorded that is unrelated to the issue you're investigating, so it's a good idea to revert to the previous setting when the problem is solved.

Even though the various logging levels are hierarchical in nature, one oddity is that *notice* level messages are always logged regardless of the setting of the *LogLevel* directive.



## Recipe 3.3 Logging POST Contents

### Problem

You want to record data submitted with the POST method, such as from a web form.

### Solution

Generally not possible in Apache 1.3 unless the POST-handling module explicitly records the data; possible in Apache 2.0 via input filter functionality, but no such filters are available at the time of this writing.

### Discussion

In Version 1.3 of Apache, only one module receives the chance to process the message body of a request, which contains the POST variable settings, and it's up to this module to record them. The request information is read from the network exactly once by the module chosen by the server to handle the response, and therefore, the information isn't available to the logging phase, which follows the content handling phase.

For example, if you're using *mod\_perl*, you may be able to record the information if the content handler that delivers the response is a Perl script being handled by *mod\_perl*.

### See Also

- 

Watch this book's web site for updates about the availability of an input filter to do this:

<http://Apache-Cookbook.Com/>

- 

See various articles on the Web about Apache 2.0 filters, such as:

<http://OnLAMP.Com/apache/http://ApacheToday.Com/>

## Recipe 3.4 Logging a Proxied Client's IP Address

### Problem

You want to log the IP address of the actual client requesting your pages, even if they're being requested through a proxy.

### Solution

None.

### Discussion

Unfortunately, the HTTP protocol itself prevents this from being possible. From the client side, proxies are intended to be completely transparent; from the side of the origin server, where the content actually resides, they are meant to be almost utterly opaque, concealing the identity of a request.

Your best option is to log the IP address from which the request came. If it came directly from a browser, it will be the client's address; if it came through one or more proxy servers, it will be the address of the one that actually contacts your server.

Both the *combined* and *common* log formats include the *%h* format effector, which represents the (remote) client's identity. However, this may be a hostname rather than an address, depending on the setting of your *HostNameLookups* directive, among other things. If you always want the client's IP address to be included in your logfile, use the *%a* effector instead.

### See Also

- 

The HTTP protocol specification at <ftp://ftp.isi.edu/in-notes/rfc2616.txt>

## Recipe 3.5 Logging Client MAC Addresses

### Problem

You want to record the MAC (hardware) address of clients that access your server.

### Solution

This cannot be logged reliably in most network situations and not by Apache at all.

### Discussion

The MAC address is not meaningful except on local area networks (LANs) and is not available in wide-area network transactions. When a network packet goes through a router, such as when leaving a LAN, the router will typically rewrite the MAC address field with the router's hardware address.

### See Also

- The TCP/IP protocol specifications (see <http://www.rfc-editor.org/cgi-bin/rfcsearch.pl> and search for "TCP" in the title field)



## Recipe 3.6 Logging Cookies

### Problem

You want to record all the cookies sent to your server by clients and all the cookies your server asks clients to set in their databases; this can be useful when debugging web applications that use cookies.

### Solution

To log cookies received from the client:

```
CustomLog logs/cookies_in.log "%{UNIQUE_ID}e %{Cookie}i"  
CustomLog logs/cookies2_in.log "%{UNIQUE_ID}e %{Cookie2}i"
```

To log cookie values set and sent by the server to the client:

```
CustomLog logs/cookies_out.log "%{UNIQUE_ID}e %{Set-Cookie}o"  
CustomLog logs/cookies2_out.log "%{UNIQUE_ID}e %{Set-Cookie2}o"
```

Using the `%(Set-Cookie)o` format effector for debugging is not recommended if multiple cookies are (or may be) involved. Only the first one will be recorded in the logfile. See the Discussion text for an example.



At the time of this writing, the Apache package includes no way to record all cookie values, but one of the authors of this book is working on one. When it's available, it should be mentioned on this book's web site, <http://Apache-Cookbook.Com/>.

### Discussion

Cookie fields tend to be very long and complex, so the previous statements will create separate files for logging them. The cookie log entries can be correlated against the client request access log using the server-set `UNIQUE_ID` environment variable (assuming that `mod_unique_id` is active in the server and that the activity log format includes the environment variable with a `%(UNIQUE_ID)e` format effector).

At the time of this writing, the `Cookie` and `Set-Cookie` header fields are most commonly used. The `Cookie2` and corresponding `Set-Cookie2` fields are newer and have been designed to correct some of the shortcomings in the original specifications, but they haven't yet achieved much penetration.

Because of the manner in which the syntax of the cookie header fields has changed over time, these logging instructions may or may not capture the complete details of the cookies.

Bear in mind that these logging directives will record all cookies, and not just the ones in which you may be particularly interested. For example, here is the log entry for a client request that included two cookies, one named `RFC2109-1` and one named `RFC2109-2`:

```
PNCsUsCoF2UAACI3CZs RFC2109-1="This is an old-style cookie, with space characters  
embedded"; RFC2109-2=This_is_a_normal_old-style_cookie
```

Even though there's only one log entry, it contains information about two cookies.





# Recipe 3.7 Not Logging Image Requests from Local Pages

## Problem

You want to log requests for images on your site, except when they're requests from one of your own pages. You might want to do this to keep your logfile size down, or possibly to track down sites that are hijacking your artwork and using it to adorn their pages.

## Solution

Use *SetEnvIfNoCase* to restrict logging to only those requests from outside of your site:

```
<FilesMatch \.(jpg|gif|png)$>  
    SetEnvIfNoCase Referer "^http://www.example.com/" local_referrer=1  
</FilesMatch>  
CustomLog logs/access_log combined env=!local_referrer
```

## Discussion

In many cases, documents on a web server include references to images also kept on the server, but the only item of real interest for log analysis is the referencing page itself. How can you keep the server from logging all the requests for the images that happen when such a local page is accessed?

The *SetEnvIfNoCase* will set an environment variable if the page that linked to the image is from the *www.example.com* site (obviously you should replace that site name with your own) and the request is for a GIF, PNG, or JPEG image.



*SetEnvIfNoCase* is the same as *SetEnvIf* except that variable comparisons are done in a case-insensitive manner.

The *CustomLog* directive will log all requests that do not have that environment variable set, i.e., everything except requests for images that come from links on your own pages.

This recipe only works for clients that actually report the referring page. Some people regard the URL of the referring page to be no business of anyone but themselves, and some clients permit the user to select whether to include this information or not. There are also "anonymising" sites on the Internet that act as proxies and conceal this information.

## See Also

- 

[Recipe 6.5](#)



# Recipe 3.8 Logging Requests by Day or Hour

## Problem

You want to automatically roll over the Apache logs at specific times without having to shut down and restart the server.

## Solution

Use CustomLog and the *rotatelog*s program:

```
CustomLog "| /path/to/rotatelog /path/to/logs/access_log.%Y-%m-%d 86400" combined
```

## Discussion

The *rotatelog*s script is designed to use an Apache feature called *pipelined logging*, which is just a fancy name for sending log output to another program rather than to a file. By inserting the *rotatelog*s script between the web server and the actual logfiles on disk, you can avoid having to restart the server to create new files; the script automatically opens a new file at the designated time and starts writing to it.

The first argument to the *rotatelog*s script is the base name of the file to which records should be logged. If it contains one or more % characters, it will be treated as a *strftime(3)* format string; otherwise, the rollover time (in seconds since 1 January 1970), in the form of a 10-digit number, will be appended to the base name. For example, a base name of *foo* would result in logfile names like *foo.1020297600*, whereas a base name of *foo.%Y-%m-%d* would cause the logfiles to be named something like *foo.2002-04-29*.

The second argument is the interval (in seconds) between rollovers. Rollovers will occur whenever the system time is a multiple of this value. For instance, a 24-hour day contains 86,400 seconds; if you specify a rollover interval of 86400, a new logfile will be created every night at midnight—when the system time, which is based at representing midnight on 1 January 1970, is a multiple of 24 hours.



Note that the rollover interval is in actual clock seconds elapsed, so when time changes because of daylight savings, this does not in any way affect the interval between rollovers.

## See Also

- 

The *rotatelog*s manpage; try

```
% man -M /path/to/ServerRoot /man rotatelog.s.8
```

replacing the */path/to/ServerRoot* with the actual value of your installation's *ServerRoot* directive in *httpd.conf*

## Recipe 3.9 Rotating Logs on the First of the Month

### Problem

You want to close the previous month's logs and open new ones on the first of each month.

### Solution

The Apache distribution doesn't include a canned script to perform this duty, but you can find one on this book's companion site at <http://Apache-Cookbook.Com/sources/Chapter04/rotate-log-monthly.pl>.

You use it by piping the desired activity log through the script, such as with:

```
CustomLog "| rotate-log-monthly.pl logs/access_log-%Y-%M" CLF
```

The argument to the script is the name of the logfile; the special sequences beginning with `%` are passed to `strftime(3)` to form the new logfile name.

### Discussion

As with the other logging solutions in this chapter, the solution given here addresses a single specific need. If you want to combine functions, such as splitting logfiles according to virtual host and rotating them on the first of each month, you are going to need to use custom scripts.

The *rotate-log-monthly.pl* script is a rather brute-force approach and may not be appropriate for extremely high traffic servers due to its sampling of the system time at each log entry. However, it illustrates the technique.

### See Also

- 

<http://httpd.apache.org/docs/logs.html#piped>

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 3.10 Logging Hostnames Instead of IP Addresses

### Problem

You want to see hostnames in your activity log instead of IP addresses.

### Solution

You can let the web server resolve the hostname when it processes the request by enabling runtime lookups with the Apache directive:

```
HostnameLookups On
```

Or, you can let Apache use the IP address during normal processing and let a piped logging process resolve them as part of recording the entry:

```
HostnameLookups Off  
CustomLog "| /path/to/logresolve -c >> /path/to/logs/access_log.resolved" combined
```

Or, you can let Apache use and log the IP addresses, and resolve them later when analyzing the logfile. Add this to *http.conf*:

```
CustomLog /path/to/logs/access_log.raw combined
```

And analyze the log with:

```
% /path/to/logresolve -c < access_log.raw > access_log.resolved
```

### Discussion

The Apache activity logging mechanism can record either the client's IP address or its hostname (or both). Logging the hostname directly requires that the server spend some time to perform a DNS lookup to turn the IP address (which it already has) into a hostname. This can have some serious impact on the server's performance, however, because it needs to consult the name service in order to turn the address into a name; and while a server child or thread is busy waiting for that, it isn't handling client requests. One alternative is to have the server record only the client's IP address and resolve the address to a name during logfile postprocessing and analysis. At the very least, defer it to a separate process that won't directly tie up the web server with the resolution overhead.

In theory this is an excellent choice; in practice, however, there are some pitfalls. For one thing, the *logresolve* application included with Apache (usually installed in the *bin/* subdirectory under the *ServerRoot*) will only resolve IP addresses that appear at the very beginning of the log entry, and so it's not very flexible if you want to use a nonstandard format for your logfile. For another, if too much time passes between the collection and resolution of the IP addresses, the DNS may have changed sufficiently so that misleading or incorrect results may be obtained. This is especially a problem with dynamically allocated IP addresses such as those issued by ISPs.

An additional shortcoming becomes apparent if you feed your log records directly to *logresolve* through a pipe: as of Apache 1.3.24 at least, *logresolve* doesn't flush its output buffers immediately, so there's the possibility of lost data if the logging process or the system should crash.

### See Also

- 

The *logresolve* manpage:



## Recipe 3.11 Maintaining Separate Logs for Each Virtual Host

### Problem

You want to have separate activity logs for each of your virtual hosts, but you don't want to have all the open files that multiple *CustomLog* directives would use.

### Solution

Use the *split-logfile* program that comes with Apache. To split logfiles after they've been rolled over (replace */path/to/ServerRoot* with the correct path):

```
# cd /path/to /ServerRoot
# mv logs/access_log logs/access_log.old
# bin/apachectl graceful
  [wait for old logfile to be completely closed]
# cd logs
# ../bin/split-logfile < access_log.old
```

To split records to the appropriate files as they're written, add this line to your *httpd.conf* file:

```
CustomLog "| /path/to/split-logfile /usr/local/Apache/logs" combined
```

### Discussion

In order for *split-logfile* to work, the logging format you're using must begin with "%v " (note the blank after the v). This inserts the name of the virtual host at the beginning of each log entry; *split-logfile* will use this to figure out to which file the entry should be written. The hostname will be removed from the record before it gets written.

There are two ways to split your access logfile: after it's been written, closed, and rolled over, or as the entries are actually being recorded. To split a closed logfile, just feed it into the *split-logfile* script. To split the entries into separate files as they're actually being written, modify your configuration to pipe the log messages directly to the script.

Each method has advantages and disadvantages. The rollover method requires twice as much disk space (for the unsplit log plus the split ones) and that you verify that the logfile is completely closed. (Unfortunately there is no guaranteed, simple way of doing this without actually shutting down the server or doing a graceless restart; it's entirely possible that a slow connection may keep the old logfile open for a considerable amount of time after a graceful restart.) Splitting as the entries are recorded is sensitive to the logging process dying—although Apache will automatically restart it, log messages waiting for it can pile up and constipate the server.

### See Also

- 

[Recipe 3.10](#)



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 3.12 Logging Proxy Requests

### Problem

You want to log requests that go through your proxy to a different file than the requests coming directly to your server.

### Solution

Use the *SetEnv* directive to earmark those requests that came through the proxy server, in order to trigger conditional logging:

```
<Directory proxy:*>
    SetEnv is_proxied 1
</Directory>
CustomLog logs/proxy_log combined env=is_proxied
```

### Discussion

Apache 1.3 has a special syntax for the *<Directory>* directive, which applies specifically to requests passing through the proxy module. Although the *\** makes it appear that wildcards can be used to match documents, it's misleading; it isn't really a wildcard. You may either match explicit paths, such as `proxy:http://example.com/foo.html`, or use *\** to match everything. You can not do something like `proxy:http://example.com/*.html`.

If you want to apply different directives to different proxied paths, you need to take advantage of another module. Because you're dealing with requests that are passing through your server rather than being handled by it directly (i.e., your server is a *proxy* rather than an *origin server*), you can't use *<Files>* or *<FilesMatch>* containers to apply directives to particular proxied documents. Nor can you use *<Location>* or *<LocationMatch>* stanzas, because they can't appear inside a *<Directory>* container. You can, however, use *mod\_rewrite*'s capabilities to make decisions based on the path of the requested document. For instance, you can log proxied requests for images in a separate file with something like this:

```
<Directory proxy:*>
    RewriteEngine On
    RewriteRule "\.(gif|png|jpg)$" "-" [ENV=proxied_image:1]
    RewriteCond "%{ENV:proxied_image}" "!1"
    RewriteRule "^" "-" [ENV=proxied_other:1]
</Directory>
CustomLog logs/proxy_image_log combined env=proxied_image
CustomLog logs/proxy_other_log combined env=proxied_other
```

Directives in the *<Directory proxy:\*>* container will only apply to requests going through your server. The first *RewriteRule* directive sets an environment variable if the requested document ends in *.gif*, *.png*, or *.jpg*. The *RewriteCond* directive tests to see if that environment variable isn't set, and the following *RewriteRule* will set a different environment variable if so. The two *CustomLog* directives send the different types of requests to different logfiles according to the environment variables.

### See Also

- 

The *mod\_rewrite* and *mod\_log\_config* documentation



## Recipe 3.13 Logging Errors for Virtual Hosts to Multiple Files

### Problem

Unlike access logs, Apache only logs errors to a single location. You want Apache to log errors that refer to a particular virtual host to the host's error log, as well as to the global error log.

### Solution

There are at least two possible ways of doing this:

1.

Use piped logging to send entries to a custom script that will copy and direct error messages to the appropriate files.

2.

Use piped logging to duplicate log entries:

```
ErrorLog "| tee logfile1 | tee logfile2 > logfile3"
```

### Discussion

Unlike activity logs, Apache will log error messages only to a single location. If the error is related to a particular virtual host and this host's `<VirtualHost>` container includes an `ErrorLog` entry, the error will be logged only in this file, and it won't appear in any global error log. If the `<VirtualHost>` does not specify an `ErrorLog` directive, the error will be logged only to the global error log. (The global error log is the last `ErrorLog` directive encountered that isn't in a `<VirtualHost>` container.)

Currently, the only workaround to this is to have the necessary duplication performed by a separate process (i.e., by using piped logging to send the error messages to the process as they occur). Of the two solutions given above, the first, which involves a custom script you develop yourself, has the most flexibility. If all you want is simply duplication of entries, the second solution is simpler but requires that your platform have a `tee` program (Windows does not). It may also be subject to lagging messages if your `tee` program doesn't flush its buffers after each record it receives. This could also lead to lost messages if the pipe breaks or the system crashes.

### See Also

- 

<http://httpd.apache.org/docs/logs.html#piped>

## Recipe 3.14 Logging Server IP Addresses

### Problem

You want to log the IP address of the server that responds to a request, possibly because you have virtual hosts with multiple addresses each.

### Solution

Use the `%A` format effector in a *LogFormat* or *CustomLog* directive:

```
CustomLog logs/served-by.log "%{UNIQUE_ID}e %A"
```

### Discussion

The `%A` effector signals the activity logging system to insert the local IP address—that is, the address of the server—into the log record at the specified point. This can be useful when your server handles multiple IP addresses. For example, you might have a configuration that includes elements such as the following:

```
Listen 10.0.0.42
Listen 192.168.19.243
Listen 263.41.0.80
<VirtualHost 192.168.19.243>
    ServerName Private.Example.Com
</VirtualHost>
<VirtualHost 10.0.0.42 263.41.0.80>
    ServerName Foo.Example.Com
    ServerAlias Bar.Example.Com
</VirtualHost>
```

This might be meaningful if you want internal users to access `Foo.Example.Com` using the `10.0.0.42` address rather than the one published to the rest of the network (such as to segregate internal from external traffic over the network cards). The second virtual host is going to receive requests aimed at both addresses even though it has only one *ServerName*; using the `%A` effector in your log format can help you determine how many hits on the site are coming in over each network interface.

### See Also

- 

The *mod\_log\_config* documentation

## Recipe 3.15 Logging the Referring Page

### Problem

You want to record the URL of pages that refer clients to yours, perhaps to find out how people are reaching your site.

### Solution

Add the following effector to your activity log format:

```
%{Referer}i
```

### Discussion

One of the fields that a request header may include is called the Referer. Referer is the URL of the page that linked to the current request. For example, if file *a.html* contains a link such as:

```
<a href="b.html">another page</a>
```

When the link is followed, the request header for *b.html* will contain a Referer field that has the URL of *a.html* as its value.

The *Referer* field is not required nor reliable; some users prefer software or anonymising tools that ensure that you can't tell where they've been. However, this is usually a fairly small number and may be disregarded for most web sites.

### See Also

- [Recipe 3.17](#)
- [Recipe 6.5](#)

## Recipe 3.16 Logging the Name of the Browser Software

### Problem

You want to know the software visitors use to access your site, for example, so you can optimize its appearance for the browser that most of your audience uses.

### Solution

Add the following effector to your activity log format:

```
%{User-Agent}i
```

### Discussion

Request headers often include a field called the User-agent. This is defined as the name and version of the client software being used to make the request. For instance, a User-agent field value might look like this:

```
User-Agent: Mozilla/4.77 [en] (X11; U; Linux 2.4.4-4GB i686)
```

This tells you that the client is claiming to be Netscape Navigator 4.77, run on a Linux system and using X-windows as its GUI.

The User-agent field is neither required nor reliable; many users prefer software or anonymising tools that ensure that you can't tell what they're using. Some software even lies about itself so it can work around sites that cater specifically to one browser or another; users have this peculiar habit of thinking it's none of the webmaster's business which browser they prefer. It's a good idea to design your site to be as browser-agnostic as possible for this reason, among others. If you're going to make decisions based on the value of the field, you might as well believe it hasn't been faked—because there's no way to tell if it has.

### See Also

- 

[Recipe 3.17](#)

## Recipe 3.17 Logging Arbitrary Request Header Fields

### Problem

You want to record the values of arbitrary fields clients send to their request header, perhaps to tune the types of content you have available to the needs of your visitors.

### Solution

Use the `%{...}i` log format variable in your access log format declaration. For example, to log the Host header, you might use:

```
%{Host}i
```

### Discussion

The HTTP request sent by a web browser can be very complex, and if the client is a specialized application rather than a browser, it may insert additional metadata that's meaningful to the server. For instance, one useful request header field is the Accept field, which tells the server what kinds of content the client is capable of and willing to receive. Given a *CustomLog* line such as this:

```
CustomLog logs/accept_log "%{UNIQUE_ID}e \"%{Accept}i\""
```

a resulting log entry might look like this:

```
PNb6VsCoF2UAAH1dAUo "text/html, image/png, image/jpeg, image/gif,  
image/x-xbitmap, */*"
```

This tells you that the client that made that request is explicitly ready to handle HTML pages and certain types of images, but, in a pinch, will take whatever the server gives it (indicated by the wildcard `/*` entry).

### See Also

- [Recipe 3.15](#)
- [Recipe 3.17](#)

## Recipe 3.18 Logging Arbitrary Response Header Fields

### Problem

You want to record the values of arbitrary fields the server has included in a response header, probably to debug a script or application.

### Solution

Use the `%{...}o` log format variable in your access log format declaration. For example, to log the Last-Modified header, you would do the following:

```
%{Last-Modified}o
```

### Discussion

The HTTP response sent by Apache when answering a request can be very complex, according to the server's configuration. Advanced scripts or application servers may add custom fields to the server's response, and knowing what values were set may be of great help when trying to track down an application problem.

Other than the fact that you're recording fields the server is sending rather than receiving, this recipe is analogous to [Recipe 3.17](#) in this chapter; refer to that recipe for more details. The only difference in the syntax of the logging format effectors is that response fields are logged using an *o* effector, and request fields are logged using *i*.

### See Also

- 

[Recipe 3.17](#)

## Recipe 3.19 Logging Activity to a MySQL Database

### Problem

Rather than logging accesses to your server in flat text files, you want to log the information directly to a database for easier analysis.

### Solution

Install the latest release of *mod\_log\_sql* from [http://www.grubbybaby.com/mod\\_log\\_sql/](http://www.grubbybaby.com/mod_log_sql/) according to the modules directions (see [Recipe 2.1](#)), and then issue the following commands:

```
# mysqladmin create apache_log
# mysql apache_log
# mysql apache_log < access_log.sql
mysql> grant insert,create on apache_log.* to webserver@localhost identified by
      'wwwpw' ;
```

Add the following lines to your *httpd.conf* file:

```
<IfModule mod_log_mysql.c>
  MySQLLoginInfo localhost webserver wwwpw
  MySQLDatabase apache_log
  MySQLTransferLogTable access_log
  MySQLTransferLogFormat huSUsbTvRA
</IfModule>
```

### Discussion

Note that there is currently a difference between the name of the web location, the tarball, and the actual module. The module file and directory retain the "mysql" name, but the directory and the tarball have moved to a more generic "sql" naming convention.

### See Also

- 

[http://www.grubbybaby.com/mod\\_log\\_sql/](http://www.grubbybaby.com/mod_log_sql/)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 3.20 Logging to syslog

### Problem

You want to send your log entries to syslog.

### Solution

To log your error log to syslog, simply tell Apache to log to syslog:

```
ErrorLog syslog:user
```



Some *syslog* reporting class other than *user*, such as *local1*, might be more appropriate in your environment.

Logging your access log to syslog takes a little work. Add the following to your configuration file:

```
CustomLog |/usr/local/apache/bin/apache_syslog combined
```

where *apache\_syslog* is a program that looks like the following:

```
#!/usr/bin/perl
use Sys::Syslog qw( :DEFAULT setlogsock );

setlogsock('unix');
openlog('apache', 'cons', 'pid', 'user');

while ($log = <STDIN>) {
    syslog('notice', $log);
}

closelog;
```

(Note that this script is only a skeleton; an actual production quality version should include code to check for errors, etc.)

### Discussion

There are two main reasons for logging to syslog. The first of these is to have many servers log to a central logging facility. The second is that there are many existing tools for monitoring syslog and sending appropriate notifications on certain events. Allow Apache to take advantage of these tools, and your particular installation may benefit.

Apache supports logging your error log to syslog by default. This is by far the more useful log to handle this way, because syslog is typically used to track error conditions rather than merely informational messages.

The syntax of the *ErrorLog* directive allows you to specify syslog as an argument or to specify a particular syslog facility. In the previous example, the user syslog facility was specified. In your */etc/syslog.conf* file, you can specify where a particular log facility should be sent—whether to a file or to a remote syslog server.

Since Apache does not support logging your access log to syslog by default, you need to accomplish this with a





## Recipe 3.21 Logging User Directories

### Problem

You want each user directory web site (i.e., those accessed via `http://servername/~username`) to have its own logfile.

### Solution

In `httpd.conf`, add the directive:

```
CustomLog "|/usr/local/apache/bin/userdir_log" combined
```

Then, in the file `/usr/local/apache/bin/userdir_log`, place the following code:

```
#!/usr/bin/perl

my $L = '/usr/local/apache/logs'; # Log directory

my %is_open = ( ); # File handle cache
$|=1;
open(F, ">>$L/access_log"); # Default error log

while (my $log = <STDIN>) {
    if ($log =~ m!\s/~(.*?)!\) {
        my $u = $1;
        unless ($is_open{$u}) {
            my $fh;
            open $fh, '>>' . $L . '/' . $u;
            $is_open{$u} = $fh;
        }
        select ($is_open{$u});
        $|=1;
        print $log;
    }
    else {
        select F;
        $|=1;
        print F $log;
    }
}

close F;
foreach my $h (keys %is_open) {
    close $h;
}
```

(Note that this script is only a skeleton; an actual production quality version should include code to check for errors, etc.)

### Discussion

Usually, requests to user directory web sites are logged in the main server log, with no differentiation between one user's site and another. This can make it very hard for a user to locate log messages for his personal web site.

The Solution allows you to break out those requests into one logfile per user, with requests not going to a userdir web site but going to the main logfile. The log handler can, of course, be modified to put all log messages in the main logfile as well as in the individual logfiles.



## Chapter 4. Virtual Hosts

A web server system supports multiple web sites in a way similar to a person who responds to her given name, as well as her nickname. In the Apache configuration file, each alternate identity, and probably the "main" one as well, is known as a virtual host (sometimes written as vhost) identified with a `<VirtualHost>` container directive. Depending on the name used to access the web server, Apache responds appropriately, just as someone might answer differently depending on whether she is addressed as "Miss Jones" or "Hey, Debbie!" If you want to have a single system support multiple web sites, you must configure Apache appropriately.

There are two different types of virtual host supported by Apache. The first type, called address-based or IP-based, is tied to the numeric network address used to reach the system. Bruce Wayne never answered the parlour telephone with "Batman here!" nor did he answer the phone in the Batcave by saying, "Bruce Wayne speaking." However, it's the same person answering the phone, just as it's the same web server receiving the request.

The other type of virtual host is name-based, because the server's response depends on what it is called. To continue the telephone analogy, consider an apartment shared by multiple roommates; you call the same number whether you want to speak to Dave, Joyce, Amaterasu, or George. Just as multiple people may share a single telephone number, multiple web sites can share the same IP address. However, all IP addresses shared by multiple Apache virtual hosts need to be declared with a `NameVirtualHost` directive.

In the most simple of Apache configurations, there are no virtual hosts. Instead, all of the directives in the configuration file apply universally to the operation of the server. The environment defined by the directives outside any `<VirtualHost>` containers is sometimes called the default server or perhaps the global server. There is no official name for it, but it can become a factor when adding virtual hosts to your configuration.

But what happens if you add a `<VirtualHost>` container to such a configuration? How are those directives outside the container interpreted, and what is their effect on the virtual host?

The answer is not a simple one: essentially, the effect is specific to each configuration directive. Some get inherited by the virtual hosts, some get reset to a default value, and some pretend they've never been used before. You'll need to consult the documentation for each directive to know for sure.

There are two primary forms of virtual hosts: IP-based virtual hosts, where each virtual host has its own unique IP address; and name-based virtual hosts, where more than one virtual host runs on the same IP address but with different names. This chapter will show you how to configure each one and how to combine the two on the same server. You'll also learn how to fix common problems that occur with virtual hosts.

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 4.1 Setting Up Name-Based Virtual Hosts

### Problem

You have only one IP address, but you want to support more than one web site on your system.

### Solution

Use the *NameVirtualHost* \* directive in conjunction with *<VirtualHost>* sections:

```
ServerName 127.0.0.1
NameVirtualHost *

<VirtualHost *>
    ServerName TheSmiths.name
    DocumentRoot "C:/Apache/Sites/TheSmiths"
</VirtualHost>

<VirtualHost *>
    ServerName JohnSmith.name
    DocumentRoot "C:/Apache/Sites/JustJohnSmith"
</VirtualHost>
```

### Discussion

With IP addresses increasingly hard to come by, name-based virtual hosting is the most common way to run multiple web sites on the same Apache server. The previous recipe works, for most users, in most virtual hosting situations.

The \* in the previous rules means that the specified hosts run on all addresses. For a machine with only a single address, this means that it runs on that address but will also run on the loopback, or localhost address. Thus if you are sitting at the physical server system, you can view the web site.

The argument to the *<VirtualHost>* container directive is the same as the argument to the *NameVirtualHost* directive. Putting the hostname here may ignore the virtual host on server startup, and requests to this virtual host may unexpectedly go somewhere else. If your name server is down or otherwise unresponsive at the time that your Apache server is starting up, then Apache can't match the particular *<VirtualHost>* section to the *NameVirtualHost* directive to which it belongs.

Requests for which there is not a virtual host listed will go to the first virtual host listed in the configuration file. In the case of the previous example, hostnames that are not explicitly mentioned in one of the virtual hosts will be served by the TheSmiths.name virtual host.

It is particularly instructive to run *httpd -S* and observe the virtual host configuration as Apache understands it, to see if it matches the way that you understand it. *httpd -S* returns the virtual host configuration, showing which hosts are name-based, which are IP-based, and what the defaults are.

Multiple names can be listed for a particular virtual host using the *ServerAlias* directive, as shown here:

```
ServerName TheSmiths.name
ServerAlias www.TheSmiths.name Smith.Family.name
```



## Recipe 4.2 Designating One Name-Based Virtual Host as the Default

### Problem

You want all requests, whether they match a virtual host or use an IP address, to be directed to a default host, possibly with a "host not found" error message.

### Solution

Add the following `<VirtualHost>` section, and list it before all of your other ones:

```
<VirtualHost *>
  ServerName default
  DocumentRoot /www/htdocs
  ErrorDocument 404 /site_list.html
</VirtualHost>
```

### Discussion

Note that this recipe is used in the context of name-based virtual hosts, so it is assumed that you have other virtual hosts that are also using the `<VirtualHost *>` notation, and that there is also an accompanying `NameVirtualHost *` appearing above them. We have used the default name for clarity; you can call it whatever you want.

Setting the `ErrorDocument 404` to a list of the available sites on the server directs the user to useful content, rather than leaving him stranded with an unhelpful 404 error message. You may wish to set `DirectoryIndex` to the site list as well, so that users who go directly to the front page of this site also get useful information.

It's a good idea to list explicitly all valid hostnames either as `ServerNames` or `ServerAliases`, so that nobody ever winds up at the default site. However, if someone accesses the site directly by IP address, or if a hostname is added to the address in question before the appropriate virtual host is created, the user still gets useful content.

### See Also

- 

[Recipe 4.4](#)

## Recipe 4.3 Setting Up Address-Based Virtual Hosts

### Problem

You have multiple IP addresses assigned to your system, and you want to support one web site on each.

### Solution

Create a virtual host section for each IP address you want to list on:

```
ServerName 127.0.0.1

<VirtualHost 10.0.0.1>
    ServerName TheSmiths.name
    DocumentRoot "C:/Apache/Sites/TheSmiths"
</VirtualHost>

<VirtualHost 10.0.0.2>
    ServerName JohnSmith.name
    DocumentRoot "C:/Apache/Sites/JustJohnSmith"
</VirtualHost>
```

### Discussion

The virtual hosts defined in this example catch all requests to the specified IP addresses, regardless of what hostname is used to get there. Requests to any other IP address not listed go to the virtual host listed in the main body of the configuration file.

The *ServerName* specified is used as the primary name of the virtual host, when needed, but is not used in the process of mapping a request to the correct host. Only the IP address is consulted to figure out which virtual host to serve requests from, not the Host header field.

### See Also

- 

<http://httpd.apache.org/docs/vhosts/>

## Recipe 4.4 Creating a Default Address-Based Virtual Host

### Problem

You want to create a virtual host to catch all requests that don't map to one of your address-based virtual hosts.

### Solution

Use the `_default_` keyword to designate a default host:

```
<VirtualHost _default_>
    DocumentRoot /www/htdocs
</VirtualHost>
```

### Discussion

The `_default_` keyword creates a virtual host that catches all requests for any address:port combinations for which there is not a virtual host configured.

The `_default_` directive may be used in conjunction with a particular port number, such as:

```
<VirtualHost _default_:443>
```

Using this syntax means that the specified virtual host catches all requests to port 443, on all addresses for which there is not an explicit virtual host configured. This is the usual way to set up SSL, which you see in the default SSL configuration file.

`_default_` typically does not work as people expect in the case of name-based virtual hosts. It does not match names for which there are no virtual host sections, only address:port combinations for which there are no virtual hosts configured. If you wish to create a default name-based host, see [Recipe 4.2](#).

### See Also

- 

[Recipe 4.2](#)

## Recipe 4.5 Mixing Address-Based and Name-Based Virtual Hosts

### Problem

You have multiple IP addresses assigned to your system, and you want to support more than one web site on each address.

### Solution

Provide a *NameVirtualHost* directive for each IP address, and proceed as you did with a single IP address:

```
ServerName 127.0.0.1
NameVirtualHost 10.0.0.1
NameVirtualHost 10.0.0.2

<VirtualHost 10.0.0.1>
    ServerName TheSmiths.name
    DocumentRoot "C:/Apache/Sites/TheSmiths"
</VirtualHost>

<VirtualHost 10.0.0.1>
    ServerName JohnSmith.name
    DocumentRoot "C:/Apache/Sites/JustJohnSmith"
</VirtualHost>

<VirtualHost 10.0.0.2>
    ServerName TheFergusons.name
    DocumentRoot "C:/Apache/Sites/TheFergusons"
</VirtualHost>

<VirtualHost 10.0.0.2>
    ServerName DoriFergusons.name
    DocumentRoot "C:/Apache/Sites/JustDoriFerguson"
</VirtualHost>
```

### Discussion

Using the address of the server, rather than the wildcard \* argument, makes the virtual hosts listen only to that IP address. However, you should notice that the argument to *<VirtualHost>* still must match the argument to the *NameVirtualHost* with which they are connected.

### See Also

- 

<http://httpd.apache.org/docs/vhosts/>



# Recipe 4.6 Mass Virtual Hosting with `mod_vhost_alias`

## Problem

You want to host many virtual hosts, all of which have exactly the same configuration.

## Solution

Use `VirtualDocumentRoot` and `VirtualScriptAlias` provided by `mod_vhost_alias`.

```
VirtualDocumentRoot /www/vhosts/%-1/%-2.1/%-2/htdocs
VirtualScriptAlias  /www/vhosts/%-1/%-2.1/%-2/cgi-bin
```

## Discussion

This recipe uses directives from `mod_vhost_alias`, which you may not have installed when you built Apache, as it is not one of the modules that is enabled by default.

These directives map requests to a directory built up from pieces of the hostname that was requested. Each of the variables represents one part of the hostname, so that each hostname is mapped to a different directory.

In this particular example, requests for content from `www.example.com` is served from the directory `/www/vhosts/com/e/example/htdocs`, or from `/www/vhosts/com/e/example/cgi-bin` (for CGI requests). The full range of available variables is shown in [Table 4-1](#).

Table 4-1. `mod_vhost_alias` variables

Variable	Meaning
%%	insert a %
%p	insert the port number of the virtual host
%M.N	insert (part of) the name

M and N may have positive or negative integer values, which have the following meanings (see [Table 4-2](#)).

Table 4-2. Meanings of variable values

Value	Meaning
0	The whole name
1	The first part of the name
2	The second part of the name
...	...
N	The Nth part of the name
-1	The last part of the name
-2	The second to last part of the name
...	...
-N	The Nth to last part of the name



## Recipe 4.7 Mass Virtual Hosting Using Rewrite Rules

### Problem

Although there is a module—*mod\_vhost\_alias*—which is explicitly for the purpose of supporting large numbers of virtual hosts, it is very limiting and requires that every virtual host be configured exactly the same way. You want to support a large number of vhosts, configured dynamically, but, at the same time, you want to avoid *mod\_vhost\_alias*.

### Solution

Use directives from *mod\_rewrite* to map to a directory based on the hostname:

```
RewriteEngine on
RewriteCond    %{HTTP_HOST}      ^(www\.)?([^.]+\)\.com$
RewriteRule    ^(.*)$           /home/%2$1
```

### Discussion

*mod\_vhost\_alias* is useful, but it is best for settings where each virtual host is identical in every way but hostname. Using *mod\_vhost\_alias* precludes the use of other URL-mapping modules, such as *mod\_userdir*, *mod\_rewrite*, and *mod\_alias*, and it can be very restrictive. Using *mod\_rewrite* is less efficient, but it is more flexible.

For example, when using *mod\_vhost\_alias*, you must do all of your hosts with *mod\_vhost\_alias*; whereas with this alternate approach, you can do some of your hosts using the rewrite rules and others using conventional virtual host configuration techniques.

The directives in the Solution map requests for *www.something.com* (or without the *www*) to the directory */home/something*.

### See Also

- [Recipe 5.16](#)
- <http://httpd.apache.org/docs/vhosts/>
- <http://httpd.apache.org/docs/vhosts/>
- [http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)



# Recipe 4.8 SSL and Name-Based Virtual Hosts

## Problem

You want to have multiple SSL web sites on the same server.

## Solution

In most common implementations of SSL, you are limited to one SSL host per address and port number. Thus, either you need to have a unique IP address for each SSL host or run them on alternate ports to get more than one on a particular IP address:

```
Listen 443
Listen 444

<VirtualHost 10.0.1.2:443>
    ServerName secure1.example.com
    DocumentRoot /www/vhosts/secure1

    SSLEngine On
    SSLCertificateFile /www/conf/ssl/secure1.crt
    SSLCertificateKeyFile /www/conf/ssl/secure1.key
</VirtualHost>

<VirtualHost 10.0.1.3:443>
    ServerName secure2.example.com
    DocumentRoot /www/vhosts/secure2

    SSLEngineOn
    SSLCertificateFile /www/conf/ssl/secure2.crt
    SSLCertificateKeyFile /www/conf/ssl/secure2.key
</VirtualHost>

<VirtualHost 10.0.1.3:444>
    ServerName secure3.example.com
    DocumentRoot /www/vhosts/secure3

    SSLEngineOn
    SSLCertificateFile /www/conf/ssl/secure3.crt
    SSLCertificateKeyFile /www/conf/ssl/secure3.key
</VirtualHost>
```

## Discussion

The limitation that restricts you to one SSL host per IP address is not a limitation imposed by Apache but by the way that SSL works. When the browser connects to the server, the first thing that it does is negotiate for a secure connection. During this process, the server sends its certificate to the client, which indicates that the rest of the transaction will be encrypted.

Because this happens before the browser tells the server what resource it wants, this part of the conversation can be based only on the IP address on which the client connected. By the time the server receives the Host header field, it is too late—the certificate has already been sent.

It is possible to run SSL hosts on ports other than port 443, if the port number is explicitly specified in the URL. This would allow you to get around this limitation, but it would put an additional burden on the end user to type the correct URL with the port number.



## Recipe 4.9 Logging for Each Virtual Host

### Problem

You want each virtual host to have its own logfiles.

### Solution

Specify `ErrorLog` and `CustomLog` within each virtual host declaration:

```
<VirtualHost *>
    ServerName    waldo.example.com
    DocumentRoot  /home/waldo/www/htdocs

    ErrorLog      /home/waldo/www/logs/error_log
    CustomLog     /home/waldo/www/logs/access_log combined
</VirtualHost>
```

### Discussion

The various logging directives can be placed either in the main body of your configuration file or within a `<VirtualHost>` section. When it is placed within a virtual host, log entries for that virtual host go in the specified logfile, rather than into the log file(s) defined in the main server configuration.



Each logfile counts against the total number of files and network connections your server is allowed to have. If you have 100 virtual hosts, each with its own error and activity log, that's 200 open channels—and if the server's quota is 256, you can only handle 56 concurrent requests at any one time.

In the recipe given here, the logfiles are placed within the home directory of a particular user, rather than in the main log directory. This gives you easier access to those files, but you still need to take adequate precautions to set the permissions on the directory in question. Consult [Chapter 6](#) for a discussion on file permissions.

### See Also

- [Chapter 3](#)
- [Chapter 6](#)

## Recipe 4.10 Splitting Up a LogFile

### Problem

Due to a large number of virtual hosts, you want to have a single logfile and split it up afterwards.

### Solution

```
LogFormat "%v %h %l %u %t \"%r\" %>s %b" vhost  
CustomLog logs/vhost_log vhost
```

Then, after rotating your logfile:

```
split-logfile < logs/vhost_log
```

### Discussion

The *LogFormat* directive in this recipe creates a logfile that is similar to the common log file format but additionally contains the name of the virtual host being accessed. The *split-logfile* utility splits up this logfile into its constituent virtual hosts.

### See Also

- 

[Recipe 3.11](#)

## Recipe 4.11 Port-Based Virtual Hosts

### Problem

You want to present different content for HTTP connections on different ports.

### Solution

Explicitly list the port number in the `<VirtualHost>` declaration:

```
Listen 8080

<VirtualHost 10.0.1.2:8080>
    DocumentRoot /www/vhosts/port8080
</VirtualHost>

Listen 9090

<VirtualHost 10.0.1.2:9090>
    DocumentRoot /www/vhosts/port9090
</VirtualHost>
```

### Discussion

Port-based virtual hosting is somewhat less common than other techniques shown in this chapter. However, there are a variety of situations in which it can be useful. If you have only one IP address, have no ability to add hostnames to DNS, or if your ISP blocks in-bound traffic on port 80, it may be useful to run virtual hosts on other ports.

Visitors to your web site must list the port number in the URL that they use. For example, to load content from the second virtual host previously listed, the following URL might be used:

```
http://server.example.com:9090/
```

### See Also

- 

<http://httpd.apache.org/docs/vhosts/>

## Recipe 4.12 Displaying the Same Content on Several Addresses

### Problem

You want to have the same content displayed on two of your addresses.

### Solution

Specify both addresses in the `<VirtualHost>` directive:

```
NameVirtualHost 192.168.1.1
NameVirtualHost 172.20.30.40

<VirtualHost 192.168.1.1 172.20.30.40>
    DocumentRoot /www/vhosts/server
    ServerName server.example.com
    ServerAlias server
</VirtualHost>
```

### Discussion

This setup is most useful on a machine that has addresses that are internal to your network, as well as those that are accessible only from outside your network. If these are the only addresses, you could use the `*` notation introduced in [Recipe 4.1](#). However, if there are more addresses, this allows you to specify what content appears on what address.

### See Also

- 

<http://httpd.apache.org/docs/vhosts/>

## Chapter 5. Aliases, Redirecting, and Rewriting

When Apache receives a request, it is assumed that the client will be served a file out of the *DocumentRoot* directory. However, there will be times when you want these resources to be served from some other location. For example, if you wanted to place a set of documents on your web site, it may be more convenient to leave them where they are, rather than to move them to a new location.

In this chapter, we deal with three general categories of these sort of cases. *Aliasing* refers to mapping a URL to a particular directory. *Redirecting* refers to mapping a URL to another URL. And *Rewriting* refers to using *mod\_rewrite* to alter the URL in some way.

Other recipes in this chapter are related because they map URLs to resources that are at unexpected places in the filesystem.

These topics are particularly interesting to webmasters who want to avoid link-rot or have sites that are periodically subject to upheaval (files or directories are moved around, or even moved from server to server). The redirection and rewriting capabilities of the Apache web server allow you to conceal such ugly behind-the-scenes disturbances from the eyes of your Internet visitors.



# Recipe 5.1 Showing Highlighted PHP Source Without Symlinking

## Problem

You want to be able to see the syntax-enhanced source to your PHP scripts without having to set up symbolic links for all of them.

## Solution

Add a line such as the following to your *httpd.conf* file:

```
RewriteRule "^(\.*\.\php)s$" "/cgi-bin/show.php?file=$1" [PT,L]
```

Create a file named *show.php* as shown below, and put it in your server's */cgi-bin/* directory:

```
<?php
/*
 * Show the highlighted source of a PHP script without a symlink or copy.
 */
if ((! isset($_GET))
    || (! isset($_GET['file']))
    || (! ($file = $_GET['file']))) {
    /*
     * Missing required arguments, so bail.
     */
    return status('400 Bad Request',
                 "Data insufficient or invalid.\r\n");
}

$file = preg_replace('/\.\phps$/', '.php', $file);
if (! preg_match('/\.\php$/', $file)) {
    return status('403 Forbidden',
                 "Invalid document.\r\n");
}
$docroot = $_SERVER['DOCUMENT_ROOT'];
if ((! preg_match(";$docroot;", $file))
    || (! preg_match("^/home/[^/]+/public_html;", $file))) {
    return status('403 Forbidden',
                 "Invalid document requested.\r\n");
}
Header('Content-type: text/html; charset=iso-8859-1');
print highlight_file($file);
return;

function status($msg, $text) {
    Header("Status: $msg");
    Header('Content-type: text/plain; charset=iso-8859-1');
    Header('Content-length: ' . strlen($text));
    print $text;
}
?>
```

## Discussion

The script in the solution uses a built-in PHP function to display the script's source in highlighted form. The *preg\_match* against *\$docroot* verifies the requested file is under the server's DocumentRoot. The next *preg\_match* also permits files in users' *public\_html* directories.

## See Also





# Recipe 5.2 Mapping a URL to a Directory

## Problem

You want to serve content out of a directory other than the *DocumentRoot* directory. For example, you may have an existing directory of documents, which you want to have on your web site that you do not want to move into the Apache document root.

## Solution

```
Alias /desired-URL-prefix /path/to/other/directory
```

## Discussion

The example given maps URLs starting with */desired-URL-prefix* to files in the */path/to/other/directory* directory. For example, a request for the URL:

```
http://example.com/desired/something.html
```

results in the file */path/to/other/directory/something.html* being sent to the client.

This same effect could be achieved by simply creating a symbolic link from the main document directory to the target directory and turning on the *Options +FollowSymLinks* directive.<sup>[1]</sup> However, using *Alias* explicitly allows you to keep track of these directories more easily. Creating symlinks to directories makes it hard to keep track of the location of all of your content. Additionally, a stray symlink may cause you to expose a portion of your filesystem that you did not intend to.

[1] See the documentation for the *Option* directive at <http://httpd.apache.org/docs/mod/core.html#options>.

You may also need to add a few configuration directives to permit access to the directory that you are mapping to. An error message (in your *error\_log* file) saying that the request was "denied by server configuration" usually indicates this condition. It is fairly common—and recommended in the documentation ([http://httpd.apache.org/docs/misc/security\\_tips.html#protectserverfiles](http://httpd.apache.org/docs/misc/security_tips.html#protectserverfiles))—to configure Apache to deny all access, by default, outside of the *DocumentRoot* directory. Thus, you must override this for the directory in question, with a configuration block as shown below:

```
<Directory /path/to/other/directory>
    Order allow,deny
    Allow from all
</Directory>
```

This permits access to the specified directory.

Note that the *Alias* is very strict with respect to slashes. For example, consider an *Alias* directive as follows:

```
Alias /puppies/ /www/docs/puppies/
```

This directive aliases URLs starting with */puppies/* but does not alias the URL */puppies*. This may result in a trailing slash problem. That is, if a user attempts to go to the URL <http://example.com/puppies> he gets a 404 error, whereas if he goes to the URL <http://example.com/puppies/> with the trailing slash, he receives content from the desired directory. To avoid this problem, create *Aliases* without the trailing slash on each argument.

Finally, make sure that if you have a trailing slash on the first argument to *Alias*, you also have one on the second argument. Consider the following example:



## Recipe 5.3 Creating a New URL for Existing Content

### Problem

You have an existing directory which you want to access using a different name.

### Solution

Use an *Alias* directive in *httpd.conf*:

```
Alias /newurl /www/htdocs/oldurl
```

### Discussion

While *Alias* is usually used to map URLs to a directory outside of the *DocumentRoot* directory tree, this is not necessarily required. There are many times when it is desirable to have the same content accessible via a number of different names. This is typically the case when a directory has its name changed, and you wish to have the old URLs continue to work, or when different people refer to the same content is by different names.

Remember that *Alias* only affects the mapping of a local URI (the */foo/bar.txt* part of *http://example.com/foo/bar.txt*); it doesn't affect or change the hostname part of the URL (the *http://example.com/* part). To alter that portion of the URL, use the *Redirect* or *RewriteRule* directives.

### See Also

- [Recipe 5.2](#)
- [http://httpd.apache.org/docs/mod/mod\\_alias.html](http://httpd.apache.org/docs/mod/mod_alias.html)
- [http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)



# Recipe 5.4 Giving Users Their Own URL

## Problem

You want to give each user on your system his own web space.

## Solution

If you want users' web locations to be under their home directories, add this to your *httpd.conf* file:

```
UserDir public_html
```

To put all users' web directories under a central location:

```
UserDir /www/users/*/htdocs
```

If you have *mod\_perl* installed, you can do something more advanced like this (again, added to your *httpd.conf* file):

```
<Perl>
# Folks you don't want to have this privilege
my %forbid = map { $_ => 1 } qw(root postgres bob);
opendir H, '/home/';
my @dir = readdir(H);
closedir H;
foreach my $u (@dir) {
    next if $u =~ m/^\./;
    next if $forbid{$u};
    if (-e "/home/$u/public_html") {
        push @Alias, "$u/", "/home/$u/public_html/";
    }
}
</Perl>
```

## Discussion

The first solution is the simplest and most widely used of the possible recipes we present here. With this directive in place, all users on your system are able to create a directory called *public\_html* in their home directories and put web content there. Their web space is accessible via a URL starting with a tilde (~), followed by their usernames. So, a user named bacchus accesses his personal web space via the URL:

```
http://www.example.com/~bacchus/
```

If you installed Apache from the standard source distribution, your default configuration file includes an example of this configuration. It also contains a *<Directory>* section referring to the directory */home/\*/public\_html*, with various options and permissions turned on. You need to uncomment that section in order for anyone to have access to these user web sites. This section should look something like the following:

```
<Directory /home/*/public_html>
    AllowOverride FileInfo AuthConfig Limit
    Options MultiViews Indexes SymLinksIfOwnerMatch IncludesNoExec
    <Limit GET POST OPTIONS PROPFIND>
        Order allow,deny
        Allow from all
    </Limit>
    <LimitExcept GET POST OPTIONS PROPFIND>
        Order deny,allow
        Deny from all
    </LimitExcept>
</Directory>
```

Make sure you understand what each of these directives is enabling before you uncomment this section in your



## Recipe 5.5 Aliasing Several URLs with a Single Directive

### Problem

You want to have more than one URL map to the same directory but don't want multiple *Alias* directives.

### Solution

Use *AliasMatch* in *http.conf* to match against a regular expression:

```
AliasMatch ^/pupp(y|ies) /www/docs/small_dogs
```

### Discussion

The *AliasMatch* directive allows you to use regular expressions to match arbitrary patterns in URLs and map anything matching the pattern to the desired URL. Think of it as *Alias* with a little more flexibility.

This example causes URLs starting with */puppy*, as well as URLs starting with */puppies*, to be mapped to the directory */www/docs/small\_dogs*.

Apache's regular expression syntax is discussed in much greater detail in [Appendix A](#).

### See Also

- 

- [Appendix A](#)

- 

Mastering Regular Expressions by Jeffrey Friedl (O'Reilly)

## Recipe 5.6 Mapping Several URLs to the Same CGI Directory

### Problem

You want to have a number of URLs map to the same CGI directory but don't want to have multiple *ScriptAlias* directives.

### Solution

Use `ScriptAliasMatch` in *httpd.conf* to match against a regular expression:

```
ScriptAliasMatch ^/([s]cripts?|cgi(-bin)?)/ /www/cgi-bin/
```

### Discussion

This is a more complicated recipe than the previous one and may require that you read [Appendix A](#). This directive maps requests starting with `/script/`, `/scripts/`, `/Script/`, `/Scripts/`, `/cgi/`, and `/cgi-bin/` to the directory `/www/cgi-bin/`, and it causes all files in that directory to be treated as CGI programs.

This kind of directive is generally used to clean up a mess that you have made. If you design your web site well from the start, this sort of thing is never necessary, but the first time you redesign, or otherwise rearrange your web site, you'll find the necessity for these sorts of contortions.

### See Also

- 

- [Recipe 5.5](#)

- 

- [Appendix A](#)

## Recipe 5.7 Creating a CGI Directory for Each User

### Problem

You want each user to have their own *cgi-bin* directory rather than giving them all access to the main server CGI directory.

### Solution

Put this in your *httpd.conf*:

```
<Directory /home/*/public_html/cgi-bin/>
    Options ExecCGI
    SetHandler cgi-script
</Directory>
```

### Discussion

You can't use *ScriptAlias* in this case, because for each user, the first argument to *ScriptAlias* would be different. Using *ScriptAliasMatch* would also be impossible. The second argument to *ScriptAliasMatch* must be a constant string.

This recipe lets each user put CGI scripts in her own personal web space. Files accessed via URLs starting with:  
`http://www.example.com/~username/cgi-bin/`

are treated as CGI scripts.

If you have *suexec* enabled, CGI programs run from this target directory will be run with the *userid* of the user specified in the URL. For example, a CGI program accessed via the URL  
`http://www.example.com/~rbowen/cgi-bin/example.cgi` would be run as the user *rbowen*.

### See Also

- 

[Recipe 8.1](#)



# Recipe 5.8 Redirecting to Another Location

## Problem

You want requests to a particular URL to be redirected to another server.

## Solution

Use a `Redirect` directive in `httpd.conf`, and give an absolute URL on the second argument:

```
Redirect /example http://www.other.server/new/location
```

## Discussion

Whereas *Alias* maps a URL to something in the local filesystem, *Redirect* maps a URL to another URL, usually on another server. The second argument is a full URL and is sent back to the client (browser), which makes a second request for the new URL.

It is also important to know that the *Redirect* directive preserves path information, if there is any. Therefore, this recipe redirects a request for `http://original.server/example/something.html` to `http://www.other.server/new/location/something.html`.

Redirections come in several different flavors, too; you can specify which particular type of redirect you want to use by inserting the appropriate keyword between the *Redirect* directive and the first URL argument. All redirects instruct the client where the requested document is now; the different types of redirection inform where the client should look for the document in the future.

temp

A temporary redirection is used when the document is not in the originally requested location at the moment, but is expected to be there again some time in the future. So the client remembers the URL it used on the original request and will use it on future requests for the same document.

permanent

A permanent redirection indicates that not only is the requested document not in the location specified by the client, but that the client should never look for it there again. In other words, the client should remember the new location indicated in the redirect response and look there in all subsequent requests for the resource.

gone

This keyword means that the document doesn't exist in this location, and it shouldn't bother asking any more. This differs from the 404 Not Found error response in that the gone redirection admits that the document was once here, even though it isn't any more. It's not considered an error, unlike the 404 status.



## Recipe 5.9 Redirecting Several URLs to the Same Destination

### Problem

You want to redirect a number of URLs to the same place. For example, you want to redirect requests for */fish* and */Fishing* to `http://fish.example.com/`.

### Solution

Use `RedirectMatch` in `httpd.conf` to match against a regular expression:

```
RedirectMatch ^/[fF]ish(ing)? http://fish.example.com
```

### Discussion

This recipe redirects requests on one server for URLs starting with fish or fishing, with either an upper-case or lower-case f, to a URL on another server, *fish.example.com*. As with *Redirect*, the path information, if any, is preserved. That is, a request for `http://original.server/Fishing/tackle.html` is redirected to `http://fish.example.com/tackle.html` so that existing relative links continue to work.

As with several of the earlier examples, *RedirectMatch* uses regular expressions to provide arbitrary text pattern matching.

### See Also

- 

[Appendix A](#)

## Recipe 5.10 Permitting Case-Insensitive URLs

### Problem

You want requested URLs to be valid whether uppercase or lowercase letters are used.

### Solution

Use *mod\_speling* to make URLs case-insensitive:

```
CheckSpelling On
```

### Discussion

The *mod\_speling* module is part of the standard Apache distribution but is not enabled by default, so you need to explicitly enable it.

In addition to making URLs case-insensitive, *mod\_speling*, as the name implies, provides simple spellchecking capability. In particular, in the case of a "not found" error, *mod\_speling* attempts to find files that may have been intended, based on similar spelling, transposed letters, or perhaps letters swapped with similar-looking numbers, like O for o and i for 1.

When *mod\_speling* is installed, it may be turned on for a particular scope (such as a directory, virtual host, or the entire server) by setting the *CheckSpelling* directive to *On*.

And, yes, that is the correct spelling of the module name.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_speling.html](http://httpd.apache.org/docs/mod/mod_speling.html)

## Recipe 5.11 Replacing Text in Requested URLs

### Problem

You want to change all occurrences of *string1* to *string2* in a request's URI.

### Solution

```
RewriteCond %{REQUEST_URI} "string1"  
RewriteRule "(.*)string1(.*)" "$1string2$2" [N,PT]
```

### Discussion

The [N] flag tells Apache to rerun the rewrite rule. This rule will get run repeatedly until the *RewriteCond* fails. Thus, it will get rerun as long as the URL contains the string that you want to replace. As soon as all occurrences of this string have been replaced, the *RewriteCond* will fail, and the rule will stop. The [PT] tells *mod\_rewrite* to pass the rewritten URL on to the rest of Apache for any additional processing once the rewriting is done.

### See Also

- 

[Appendix A](#)

## Recipe 5.12 Rewriting Path Information to CGI Arguments

### Problem

You want to pass arguments as part of the URL but have these components of the URL rewritten as CGI QUERY\_STRING arguments.

### Solution

This is just an example, of course; make appropriate changes to the RewriteRule line to fit your own environment and needs:

```
RewriteEngine on
RewriteRule ^/book/([^/]+)/([^/]+)/cgi-bin/book.cgi?author=$1&subject=$2 [PT]
```

### Discussion

One reason you might want or need to do this is if you're gluing together two legacy systems that do things in different ways, such as a client application and a vendor script.

For example, the RewriteRule in the Solution will cause:

```
http://www.example.com/book/apache/bowen
```

to be rewritten as:

```
http://www.example.com/cgi-bin/book.cgi?subject=apache&author=bowen
```

The [PT] flag on the RewriteRule directive instructs Apache to keep processing the URL even after it has been modified; without the flag, the server would directly try to treat the rewritten URL as a filename, instead of continuing to the step at which it determines it's a CGI script. It also allows multiple RewriteRule directives to make additional refinements to the URL.

If the URL being rewritten already has a query string, or might, change the [PT] to [QSA,PT]. The QSA means "query string add" and will cause the query string generated by the rewrite to be added to the query string in the original URL. Without QSA, the original query string will be replaced.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

## Recipe 5.13 Denying Access to Unreferred Requests

### Problem

You want to prevent other web sites from using your images (or other types of documents) in their pages and allow your images to be accessed only if they were referred from your own site.

### Solution

Put this in your *httpd.conf*:

```
RewriteEngine On
RewriteCond %{HTTP_REFERER} !=""
RewriteCond %{HTTP_REFERER} "!^http://mysite.com/.*$" [NC]
RewriteCond %{REQUEST_URI} "\.(jpg|gif|png)$"
RewriteRule .* - [F]
```

### Discussion

This recipe is a series of RewriteCond directives, designed to determine whether an image file is requested from within a document on your site or if it is embedded in a page from another server. If the the latter, then the other site is stealing your images and needs to be stopped.

The first rule checks to see if the referer is even set. Some clients don't send a referer, and some browsers can be configured not to send referers. If we deny requests from all clients that don't send a referer, we'll deny a lot of valid requests; so we let these ones in.

Next, we check to see if the referer appears to be from some site other than our own. If so, we keep going through the rules. Otherwise, we'll stop processing the rewrite.

Finally, we check to see if this is a request for an image file. If the file is a nonimage file, such as an HTML file, then we want to allow people to link to these files from somewhere offsite.

If we've reached this point in the ruleset, we know that we have a request for an image file from within a page on another web site. The RewriteRule matches a request and returns Forbidden to the client.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

## Recipe 5.14 Rewriting Based on the Query String

### Problem

You want to translate one URI into another based on the value of the query string.

### Solution

Put this in your *httpd.conf*:

```
RewriteCond "%{QUERY_STRING}"    "^user=([^\=]*)"  
RewriteRule "/userdirs"         "http://%1.users.example.com/" [R]
```

### Discussion

*mod\_rewrite* does not consider the query string as part of the URI for matching and rewriting purposes, so you need to treat it separately. The given example translates requests of the form:

`http://example.com/people?user=jones` `http://jones.users.example.com/`

The [R] tells *mod\_rewrite* to direct the browser to the URL constructed by the RewriteRule directive.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_alias.html](http://httpd.apache.org/docs/mod/mod_alias.html)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 5.15 Redirecting All—or Part—of Your Server to SSL

## Problem

You want certain parts of your non-SSL web space to be redirected to a secured area.

## Solution

You can redirect everything that is attached to port 80 with the following *RewriteRule*:

```
RewriteCond "%{SERVER_PORT}"          "^80$"
RewriteRule "^(.*)$"                  "https://%{SERVER_NAME}$1" [R,L]
```

You can redirect particular URLs to a secure version:

```
RewriteRule "^/normal/secure(/.*)" "https://%{HTTP_HOST}$1" [R,L]
```

You can check to see whether the HTTPS environment variable is set:

```
RewriteCond %{HTTPS} !=on
RewriteRule "^(/secure/.*)" "https://%{HTTP_HOST}$1" [R,L]
```

Or, you can simply use the *Redirect* directive in the http section of *httpd.conf* file to cause a URL to be served as HTTPS:

```
Redirect / https://secure.example.com/
```

Make sure that this appears only in in the http scope and not in the https scope, or all https requests will loop.

## Discussion

The first solution causes all requests that come in on port 80 (normally the unencrypted HTTP port) to be redirected to the same locations on the current server but accessed through SSL. Note the use of `SERVER_NAME`; because this is a complete site redirection, it's simplest to use the server's official name for itself.

The directive shown in the second solution causes all portions of the server's web space under `http://myhost/normal/secure` to be redirected to the SSL location rooted at `https://myhost/`. The use of `HTTP_HOST` rather than `SERVER_NAME` means that only the location and the scheme in the visitor's browser, not the server name.

Note that the paths to the SSL and non-SSL locations differ; if you want the paths to be the same except for the security, you can use something like the directives given in the third solution.

## See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

## Recipe 5.16 Turning Directories into Hostnames

### Problem

You want to migrate pathnames under a single hostname to distinct hostnames.

### Solution

Use RewriteRule in *httpd.conf*:

```
RewriteRule "^/(patha|pathb|pathc)(.*)" "http://$1.example.com$2" [R]
RewriteRule "^/([^./]*) (.*)" "http://$1.example.com$2" [R]
RewriteRule "^/~([^./]*) (.*)" "http://$1.example.com$2" [R]
```

### Discussion

The first recipe redirects requests of the form `http://example.com/pathseg/some/file.html` to a different host, such as `http://pathseg.example.com/some/file.html`, but only for those requests in which *pathseg* is *patha*, *pathb*, or *pathc*.

The second recipe does the same thing, except that any top-level path segment is redirected in this manner.

The third recipe splits the difference, redirecting all "user" requests to distinct hosts with the same name as the user.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

## Recipe 5.17 Redirecting All Requests to a Single Host

### Problem

You want all requests made of your system to be redirected to a specific host.

### Solution

Put this in your *httpd.conf*:

```
RewriteCond "%{HTTP_HOST}"    "!^www.example.com$"    [NC,OR]
RewriteCond "%{SERVER_NAME}"  "!^www.example.com$"    [NC]
RewriteRule "(.*)"            "http://www.example.com$1" [R]
```

### Discussion

Any request handled by your server within the scope of the directives in the Solution (which aren't directed to the `www.example.com` host) is redirected there.

The two different *RewriteCond* directives are used to catch all requests made by some host other than `www.example.com`, regardless of the redirection method.

The NC (No Case) flag makes the regular expression case-insensitive. That is, it makes it match regardless of whether letters are upper- or lowercase.

The OR flag is a logical "or," allowing the two conditions to be strung together so that either one being true is a sufficient condition for the rule to be applied.

Finally, the R flag causes an actual Redirect to be issued, so that the browser will make another request for the generated URL.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

## Recipe 5.18 Turning Document Names into Arguments

### Problem

You want to redirect requests for documents to a CGI script, or other handler, that gets the document names as an argument.

### Solution

Use RewriteRule in *httpd.conf*:

```
RewriteRule "^/dir/([^./]*)\.html" "/dir/script.cgi?doc=$1" [PT]
```

### Discussion

This solution causes all requests for HTML documents in the specified location to be turned into requests for a handler script that receives the document name as an argument in the QUERY\_STRING environment variable.

The *PT* flag should be included to allow any appropriate subsequent URL rewriting or manipulation to be performed.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Chapter 6. Security

In this chapter, security means allowing people to see what you want them to see and preventing them from seeing what you don't want them to see. Additionally, there are the issues of what measures you need to take on your server in order to restrict access via non-Web means. This chapter illustrates the precautions you need to take to protect your server from malicious access and modification of your web site.

The most common questions ask how to protect documents and restrict access. Unfortunately, due to the complexity of the subject and the nature of the web architecture, these questions tend to also have the most complex answers or often no convenient answers at all.

Normal security nomenclature and methodology separate the process of applying access controls into two discrete steps; in the case of the Web, they may be thought of as the server asking itself these questions:

- Are you really who you claim to be?
- Are you allowed to be here?

These steps are called *authentication* and *authorization*, respectively. Here's a real-world example: a flight attendant checks your photo identification (authentication) and your ticket (authorization) before permitting you to board an airplane.

Authentication can be broken down into what might be called *weak* and *strong*. Weak authentication is based on the correctness of credentials that the end user supplies (which therefore may have been stolen from the real owner—hence the name "weak"), whereas strong authentication is based on attributes of the request over which the end user has little or no control, and it cannot change from request to request—such as the IP address of his system.

Although checking authentication and authorization are clearly separate activities, their application gets a bit blurred in the context of the Apache web server modules. Even though the main difference between the many security modules is how they store the credentials (in a file, a database, an LDAP directory, etc.), they nevertheless have to provide the code to retrieve the credentials from the store, validate those supplied by the client, and check to see if the authenticated user is authorized to access the resource. In other words, there's a lot of functionality duplicated from module to module, and although there are frequently similarities between their behavior and directives, the lack of shared code means that sometimes they're not quite as similar as you'd hope. This overloading of functionality has been somewhat addressed in the next version of the web server after 2.0 (still in development at the time of this writing).

In addition to the matter of requiring a password to access certain content from the web server, there is the larger issue of securing your server from attacks. As with any software, Apache has, at various times in its history, been susceptible to conditions that would allow an attacker to gain inappropriate control of the hosting server. For example, they may have been able to access, or modify, files that the site administrator had not intended to give access to, or they may have been able to execute commands on the target server. Thus, it is important that you know what measures need to be taken to ensure that your server is not susceptible to these attacks.





# Recipe 6.1 Using System Account Information for Web Authentication

## Problem

You want all the users on your Unixish system to be able to authenticate themselves over the Web using their already-assigned usernames and passwords.

## Solution

Set up a realm using *mod\_auth* and name */etc/passwd* as the *AuthUserFile*:

```
<Directory "/home">
  AuthType Basic
  AuthName HomeDir
  AuthUserFile /etc/passwd
  Require valid-user
  Satisfy All
</Directory>
```

## Discussion

We must stress that using system account information for web authentication is a very bad idea, unless your site is also secured using SSL. For one thing, any intruder who happens to obtain one of your users' credentials not only can access the protected files over the Web, but can actually log onto your system where it's possible to do significant damage. For another, web logins don't have the same security controls as most operating systems; over the Web, an intruder can keep hammering away at a username with password after password without the system taking any defensive measures; all *mod\_auth* will do is record a message in the Apache error log. However, most operating systems will enter a paranoid mode and at least ignore login attempts for a while after some number of failures.

If you still want to do this, either because you consider the risk acceptable or because it doesn't apply in your situation, the *httpd.conf* directives in the Solution will do the trick. The syntax and order of the fields in a credential record used by *mod\_auth* happens (and not by accident) to match the standard layout of the */etc/passwd* lines. *mod\_auth* uses a simple text file format in which each line starts with a username and password and may optionally contain additional fields, with the fields delimited by colons. For example:

```
smith:$apr1$GLWeF/..$8hOXRFUpHhBJHpOUdNFe51
```

*mod\_auth* ignores any additional fields after the password, which is what allows the */etc/passwd* file to be used. Note that the password in the example is encrypted.

You can manage Apache *mod\_auth* credential files with the *htpasswd* utility, but don't use this utility on the */etc/passwd* file! Use the normal account maintenance tools for that.

Note that this technique will not work if shadow passwords are in use, because the password field of */etc/passwd* contains nothing useful in that situation. Instead, the passwords are stored in the file */etc/shadow*, which is readable only by root, while Apache runs as an unprivileged user. Furthermore, most modern Unixish operating systems use the */etc/shadow* means of user authentication by default.

## See Also



## Recipe 6.2 Setting Up Single-Use Passwords

### Problem

You want to be able to provide credentials that will allow visitors into your site only once.

### Solution

No solution is available with standard Apache features.

### Discussion

As described in [HTTP, Browsers, and Credentials](#), the concept of being "logged in" to a site is an illusion. In order to achieve the desired one-time-only effect, the server needs to complete the following steps:

1.

Note the first time the user successfully presents valid credentials.

2.

Somehow, associate that fact with the user's "session."

3.

Never allow those credentials to succeed again if the session information is different from the first time they succeeded.

The last step is not a simple task, and it isn't a capability provided in the standard Apache distribution. To complicate matters, there is the desire to start a timeout once the credentials have succeeded, so that the user doesn't authenticate once and then leave his browser session open for days and retain access.

Fulfilling this need would require a custom solution. Unfortunately, we are not aware of any open or public modules that provide this capability; however, search and watch the module registry for possible third-party implementations.

### See Also

- 

[Recipe 6.3](#)

- 

<http://modules.apache.org/>

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 6.3 Expiring Passwords

## Problem

You want a user's username and password to expire at a particular time or after some specific interval.

## Solution

No solution is available with standard Apache features, but a few third-party solutions exist.

## Discussion

Refer to [HTTP, Browsers, and Credentials](#). In order for Apache to provide this functionality, it would need to store more than just the valid username and password; it would also have to maintain information about the credentials' expiration time. No module provided as part of the standard Apache distribution does this.

There are several third-party solutions to this problem, including the Perl module *Apache::Htpasswd::Perishable* and the *mod\_perl* handler *Apache::AuthExpire*.

There are two slightly different ways to look at this problem, which will influence your choice of a solution. You may want a user's authentication to be timed out after a certain amount of time, or perhaps after a certain period of inactivity, forcing them to log in again. Or you may want a particular username/password pair to be completely expired after a certain amount of time, so that it no longer works. The latter might be used instead of a single-use password, which is impractical to implement in HTTP.

*Apache::Htpasswd::Perishable* partially implements the latter interpretation of the problem by adding expiration information to the password file. Inheriting from the *Apache::Htpasswd* module, it adds two additional methods, *expire* and *extend*, which set an expiration date on the password and extend the expiration time, respectively.

For example, the following code will open a password file and set an expiration date on a particular user entry in that file:

```
use Apache::Htpasswd::Perishable;

my $pass = Apache::Htpasswd::Perishable->new("/usr/local/apache/passwords/user.pass")
    or die "Could not open password file.";
$pass->expire('waldo',5); # Set the expiration date 5 days in the future
```

Such a mechanism is only useful if expired passwords are removed from the password file periodically. This can be accomplished by running the following *cron* script every day. This will delete those users for whom the expiration date has passed:

```
#!/usr/bin/perl
use Apache::Htpasswd::Perishable;

my $password_file = '/usr/local/apache/passwords/user.pass';

open (F,$password_file) or die "Could not open password file.";
my @users;
while (my $user = <F>) {
    $user =~ s/^([:])+:.*$/\1/;
```





# Recipe 6.4 Limiting Upload Size

## Problem

With more and more web hosting services allowing customers to upload documents, uploads may become too large. With a little creativity, you can put a limit on uploads by using the security capabilities of the server.

## Solution

Assume you want to put a limit on uploads of ten thousand (10,000) bytes. Here's how you could do that for your */upload* location:

```
SetEnvIf Content-Length "^[1-9][0-9]{4,}" upload_too_large=1
<Location /upload>
    Order Deny,Allow
    Deny from env=upload_too_large
    ErrorDocument 403 /cgi-bin/remap-403-to-413
</Location>
```

You can tailor the response by making the */cgi-bin/remap-403-to-413* script look something like this:

```
#!/usr/local/bin/perl
#
# Perl script to turn a 403 error into a 413 IFF
# the forbidden status is because the upload was
# too large.
#
if ($ENV{'upload_too_large'}) {
    #
    # Constipation!
    #
    print <<EOHT
Status: 413 Request Entity Too Large
Content-type: text/plain; charset=iso-8859-1
Content-length: 84
```

Sorry, but your upload file exceeds the limits set forth in our terms and conditions.

```
EOHT
}
else {
    #
    # This is a legitimate "forbidden" error.
    #
    my $uri = $ENV{'REDIRECT_REQUEST_URI'};
    my $length = 165 + length($uri);
    print <<EOHT
Status: 403 Forbidden
Content-type: text/html; charset=iso-8859-1
Content-length: $length
```

```
<html>
<head>
  <title>Forbidden</title>
</head>
<body>
  <h1>Forbidden</h1>
  <p>
    You don't have permission to access $uri
    on this server.
  </p>
</body>
```



## Recipe 6.5 Restricting Images from Being Used Off-Site

### Problem

Other sites are linking to images on your system, stealing bandwidth from you and incidentally making it appear as though the images belong to them. You want to ensure that all access to your images is from documents that are on your server.

### Solution

Add the following lines to the *.htaccess* file in the directory where the images are, or to the appropriate *<Directory>* container in the *httpd.conf* file. Replace the *myserver.com* with your domain name:

```
<FilesMatch "\.(jpg|jpeg|gif|png)$">
    SetEnvIfNoCase Referer "^http://([^\/*]*\.)?myserver.com/" local_referrer=1
    Order Allow,Deny
    Allow from env=local_referrer
</FilesMatch>
```

In fact, by using the following recipe, you can even go one step further, and return a different image to users accessing your images via an off-site reference:

```
SetEnvIfNoCase Referer "^http://([^\/*]*\.)?myserver.com/" local_referrer=1
RewriteRule %ENV{local_referrer} !1 /Stolen-100x100.png [L]
```

### Discussion

The first solution will cause all requests for image files to be refused with a 403 Forbidden status unless the link leading to the request was in one of your own documents. This means that anyone linking to your images from a different web site system will get the error instead of the image, because the referer does not match the approved server name.

Note that this technique can cause problems for requests that do not include a Referer request header field, such as people who visit your site through an anonymising service or who have their browser configured not to send this information.

The second solution is similar to the first, except that it substitutes an image of your choice for the one requested, rather than denying the request. Using the values in the Solution, you can construct a *Stolen-100x100.png* that has whatever admonitory message or perhaps just some picture that will deter the visitor from "stealing" your images.

### See Also

- 

[Recipe 6.21](#)

## Recipe 6.6 Requiring Both Weak and Strong Authentication

### Problem

You want to require both weak and strong authentication for a particular resource. For example, you wish to ensure that the user accesses the site from a particular location and to require that he provides a password.

### Solution

Use the *Satisfy* directive to require both types of authentication:

```
<Directory /www/htdocs/sensitive>

# Enforce all restrictions
Satisfy All

# Require a password
AuthType Basic
AuthName Sensitive
AuthUserFile /www/passwords/users
AuthGroupFile /www/passwords/groups
Require group salesmen

# Require access from a certain network
Order deny,allow
Deny from all
Allow from 192.168.1
</Directory>
```

### Discussion

In this example, a user must provide a login, identifying him as a member of the salesmen group, and he must also use a machine on the 192.168.1 network.

The *Satisfy All* directive requires that all access control measures be enforced for the specified scope. A user accessing the resource from a nonmatching IP address will immediately receive a Forbidden error message in his browser, while, in the logfile, the following error message is logged:

```
[Sun May 25 15:31:53 2003] [error] [client 208.32.53.7] client denied by server
configuration: /usr/local/apache/htdocs/index.html
```

Users who are in the required set of IP addresses, however, receive a password dialog box and are required to provide a valid username and password.

### See Also

- 

[Recipe 6.9](#)



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 6.7 Managing .htpasswd Files

### Problem

You wish to create password files for use with Basic HTTP authentication.

### Solution

Use the *htpasswd* utility to create your password file, as in [Table 6-1](#).

Table 6-1. Managing password files with htpasswd

Command	Action
% <b>htpasswd -c</b> user.pass waldo	Create a new password file called <i>user.pass</i> with this one new entry for user waldo. Will prompt for password.
% <b>htpasswd</b> user.pass ralph	Add an entry for user ralph in password file <i>user.pass</i> . Will prompt for password.
% <b>htpasswd -b</b> user.pass ralph mydogspot	Add a user ralph to password file <i>user.pass</i> with password mydogspot.

Or, use the Perl module *Apache::Htpasswd* to manage the file programmatically:

```
use Apache::Htpasswd;
$pass = new Apache::Htpasswd("/usr/local/apache/passwords/user.pass") or
die "Couldn't open password file.";

# Add an entry
$pass->htpasswd("waldo", "emerson");

# Delete entry
$pass->htDelete("waldo");
```

### Discussion

The *htpasswd* utility, which comes with Apache, is located in the *bin* subdirectory.



On some third-party distributions of Apache, the *htpasswd* program has been copied into a directory in your path, but ordinarily it will not be in your path; you will either have to put it there, or provide the full path to the program in order to run it, such as */usr/local/apache/bin/htpasswd*.



## Recipe 6.8 Making Password Files for Digest Authentication

### Problem

You need to create a password file to be used for Digest authentication.

### Solution

Use the following command forms to set up a credential file for a realm to be protected by Digest authentication:

```
% htdigest -c "By invitation only" rbowen
% htdigest "By invitation only" krietz
```

### Discussion

Digest authorization, implemented by *mod\_auth\_digest*, uses an MD5 hash of the username, password, and authentication realm to check the credentials of the client. The *htdigest* utility, which comes with Apache, creates these files for you.

The syntax for the command is very similar to the syntax for the *htpasswd* utility, except that you must also specify the authentication realm that the password will be used for. The resulting file contains one line per user, looking something like the following:

```
rbowen:By invitation only:23bc21f78273f49650d4b8c2e26141a6
```

Note that, unlike entries in the password files created by *htpasswd*, which can be used anywhere, these passwords can be used only in the specified authentication realm, because the encrypted hash includes the realm.

As with *htpasswd*, the *-c* flag creates a new file, possibly overwriting an existing file. You will be prompted for the password and then asked to type it again to verify it.

*htdigest* does not have any of the additional options that *htpasswd* does.

### See Also

- 

[Recipe 6.7](#)



## Recipe 6.9 Relaxing Security in a Subdirectory

### Problem

There are times when you might want to apply a tight security blanket over portions of your site, such as with something like:

```
<Directory /usr/local/apache/htdocs/BoD>
  Satisfy All
  AuthUserFile /usr/local/apache/access/bod.htpasswd
  Require valid-user
</Directory>
```

Due to Apache's scoping rules, this blanket applies to all documents in that directory and in any subordinate subdirectories underneath it. But suppose you want to make a subdirectory, such as *BoD/minutes*, available without restriction?

### Solution

The *Satisfy* directive is the answer. Add the following to either the *.htaccess* file in the subdirectory or in an appropriate *<Directory>* container:

```
Satisfy Any
Order Deny,Allow
Allow from all
```

## HTTP, Browsers, and Credentials

It is easy to draw incorrect conclusions about the behavior of the Web; when you have a page displayed in your browser, it is natural to think that you are still connected to that site. In actuality, however, that's not the case—once your browser fetches the page from the server, both disconnect and forget about each other. If you follow a link, or ask for another page from the same server, a completely new exchange has begun.

When you think about it, this is fairly obvious. It would make no sense for your browser to stay connected to the server while you went off to lunch or home for the day.

Each transaction that is unique and unrelated to others is called *stateless*, and it has a bearing on how HTTP access control works.

When it comes to password-protected pages, the web server doesn't remember whether you've accessed them before or not. Down at the HTTP level where the client (browser) and server talk to each other, the client has to prove who it is every time; it's the client that remembers your information.

When accessing a protected area for the first time in a session, here's what actually gets exchanged between the client and the server:





## Recipe 6.10 Lifting Restrictions Selectively

### Problem

You want most documents to be restricted, such as requiring a username and password, but want a few to be available to the public. For example, you may want *index.html* to be publicly accessible, while the rest of the files in the directory require password authentication.

### Solution

Use the *Satisfy Any* directive in the appropriate place in your *.htaccess* or *httpd.conf* file:

```
<Files index.html>
  Order Deny,Allow
  Allow from all
  Satisfy Any
</Files>
```

### Discussion

Regardless of what sorts of restrictions you may have on other files, or on the directory as a whole, the *<Files>* container in the solution makes the *index.html* file accessible to everyone without limitation. *Satisfy Any* tells Apache that any of the restrictions in place may be satisfied, rather than having to enforce any particular one. In this case, the restriction in force will be *Allow from all*, which permits access for all clients.

## Weak and Strong Authentication

The basic Apache security model for HTTP is based upon the concepts of *weak* and *strong* authentication mechanisms. Weak mechanisms are those that rely on information volunteered by the user; strong ones use credentials obtained without asking him. For instance, a username and password constitute a set of weak credentials, while the IP address of the user's client system is regarded as a strong one.

One difference between the two types lies in how Apache handles an authentication failure. If invalid weak credentials are presented, the server will respond with a 401 Unauthorized status, which allows the user to try again. In contrast, a failure to authenticate when strong credentials are required will result in a 403 Forbidden status—for which there is no opportunity to retry.

In addition, strong and weak credentials can be required in combination; this is controlled by the *Satisfy* directive. The five possible requirements are:

- None. No authentication required.
- Only strong credentials are needed.



## Recipe 6.11 Authorizing Using File Ownership

### Problem

You wish to require user authentication based on system file ownership. That is, you want to require that the user that owns the file matches the username that authenticated.

### Solution

Use the *Require file-owner* directive:

```
<Directory /home/*/public_html/private>
  AuthType Basic
  AuthName "MyOwnFiles"
  AuthUserFile /some/master/authdb
  Require file-owner
</Directory>
```

### Discussion

The goal here is to require that username jones must authenticate in order to access the */home/jones/public\_html/private* directory.

The user does not authenticate against the system password file but against the *AuthUserFile* specified in the example. Apache just requires that the name used for authentication matches the name of the owner of the file or directory in question. Note also that this is a feature of *mod\_auth* and is not available in other authentication modules.



This feature was added in Apache 1.3.22.

### See Also

- 

The *Require file-group* keyword at [http://httpd.apache.org/docs/mod/mod\\_auth.html#require](http://httpd.apache.org/docs/mod/mod_auth.html#require)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 6.12 Storing User Credentials in a MySQL Database

## Problem

You wish to use user and password information in your MySQL database for authenticating users.

## Solution

For Apache 1.3, use *mod\_auth\_mysql*:

```
AuthMySQLInfo db_host.example.com db_user my_password
AuthMySQLGeneralDB auth_database_name
```

```
<Directory /www/htdocs/private>
    AuthName "Protected directory"
    AuthType Basic
    require valid-user
</Directory>
```

For Apache 2.1 and later, use *mod\_authn\_dbi*:

```
AuthnDbiDriver Config1 mysql
AuthnDbiHost Config1 db.example.com
AuthnDbiUsername Config1 db_username
AuthnDbiPassword Config1 db_password
AuthnDbiName Config1 auth_database_name
AuthnDbiTable Config1 auth_database_table
AuthnDbiUsernameField Config1 user_field
AuthnDbiPasswordField Config1 password_field
AuthnDbiIsActiveField Config1 is_active_field

AuthnDbiConnMin Config1 3
AuthnDbiConnSoftMax Config1 12
AuthnDbiConnHardMax Config1 20
AuthnDbiConnTTL Config1 600

<Directory "/www/htdocs/private">
    AuthType Digest
    AuthName "Protected directory"
    AuthBasicProvider dbi
    AuthnDbiServerConfig Config1
    Require valid-user
</Directory>
```

## Discussion

There are a number of modules called *mod\_auth\_mysql*. The module used in the previous example is the *mod\_auth\_mysql* from [http://www.diegonet.com/support/mod\\_auth\\_mysql.shtml](http://www.diegonet.com/support/mod_auth_mysql.shtml). For the full explanation of the database fields that you will need to create, and the additional options that the module affords, you should consult the documentation on the web site.

If you are running Apache 2.1 or later, you will want to take advantage of the new authentication framework, and use the module *mod\_authn\_dbi*, available from [http://open.cyanworlds.com/mod\\_authn\\_dbi/](http://open.cyanworlds.com/mod_authn_dbi/). Due to the new authentication API in Apache 2.1, a number of things are possible that were not possible in earlier versions. For example, a single module, such as *mod\_authn\_dbi*, can be used for either Basic or Digest authentication, by simply changing the *AuthType* directive from *Basic* to *Digest*. (*AuthBasicProvider* would also become *AuthDigestProvider* in the previous example.)



## Recipe 6.13 Accessing the Authenticated Username

### Problem

You want to know the name of the user who has authenticated.

### Solution

Consult the environment variable `REMOTE_USER`.

In a Server-Side Include (SSI) directive, this may look like:

```
Hello, user <!--#echo var="REMOTE_USER" -->. Thanks for visiting.
```

In your CGI code, it might look like:

```
my $username = $ENV{REMOTE_USER};
```

### Discussion

When a user has authenticated, the environment variable `REMOTE_USER` is set to the name with which she authenticated. You can access this variable in CGI programs, SSI directives, PHP files, and a variety of other methods. The value will also appear in your *access\_log* file.

Note that, while it is the convention for an authentication module to set this variable, there are reportedly some third-party authentication modules that do not set it, but provide other methods for accessing that information.

### See Also

- 

[Recipe 6.14](#)

## Recipe 6.14 Obtaining the Password Used to Authenticate

### Problem

You want to get the password that the user authenticated with.

### Solution

Standard Apache modules do not make this value available. It is, however, available from the Apache API if you wish to write your own authentication methods.

In the Apache 1.3 API, you need to investigate the *ap\_get\_basic\_auth\_pw* function. In the 2.0 API, look at the *get\_basic\_auth* function.

If you write an authentication handler with *mod\_perl*, you can retrieve the username and password with the *get\_username* function:

```
my ($username, $password) = get_username($r);
```

### Discussion

For security reasons, although the username is available as an environment variable, the password used to authenticate is not available in any simple manner. The rationale behind this is that it would be a simple matter for unscrupulous individuals to capture passwords so that they could then use them for their own purposes. Thus the decision was made to make passwords near to impossible to attain.

The only way to change this is to rebuild the server from the sources with a particular (strongly discouraged) compilation flag. Alternately, if you write your own authentication module, you would of course have access to this value, as you would need to verify it in your code.

### See Also

- 

[Recipe 6.13](#)

## Recipe 6.15 Preventing Brute-Force Password Attacks

### Problem

You want to disable a username when there are repeated failed attempts to authenticate using it, as if it is being attacked by a password-cracker.

### Solution

There is no way to do this with standard Apache authentication modules. The usual approach is to watch your logfile carefully. Or you can use something like *Apache::BruteWatch* to tell you when a user is being attacked:

```
PerlLogHandler Apache::BruteWatch
PerlSetVar BruteDatabase      DBI:mysql:brutelog
PerlSetVar BruteDataUser      username
PerlSetVar BruteDataPassword password

PerlSetVar BruteMaxTries      5
PerlSetVar BruteMaxTime       120
PerlSetVar BruteNotify        rbowen@example.com
```

### Discussion

Due to the stateless nature of HTTP and the fact that users are not, technically, "logged in" at all (see [HTTP, Browsers, and Credentials](#)), there is no connection between one authentication attempt and another. This makes it possible to repeatedly attempt to log in with a particular username.

*Apache::BruteWatch* is one way to watch the logfile and send notification when a particular account is being targeted for a brute-force password attack. With the configuration shown previously, if a given account fails authentication 5 times in 2 minutes, the server administrator will be notified of the situation, so that she can take appropriate measures, such as blocking the offending address from the site.

### See Also

- 

[HTTP, Browsers, and Credentials](#)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 6.16 Using Digest Versus Basic Authentication

### Problem

You want to understand the distinction between the Basic and Digest authentication methods.

### Solution

Use *AuthType Basic* and the *htpasswd* tool to control access using Basic authentication. Use *AuthType Digest* and the *htdigest* tool for the Digest method.

### Discussion

Basic web authentication is exactly that: primitive and insecure. It works by encoding the user credentials with a reversible algorithm (essentially base-64 encoding) and transmitting the result in plaintext as part of the request header. Anyone (or anything) that intercepts the transmission can easily crack the encoding of the credentials and use them later. As a consequence, Basic authentication should only be used in environments where the protected documents aren't truly sensitive or when there is no alternative.

In contrast, Digest authentication uses a more secure method that is much less susceptible to credential theft, spoofing, and replay attacks. The exact details don't matter; the essential ingredient is that no username or password traverses the network in plaintext.

Preparing a realm to use Basic authentication consists of simply storing the username/password pair and telling the server where to find them. The password may or may not be encrypted. The same credentials may be applied to any realm on the server, or even copied to a completely different server and used there. They may be stored in a variety of databases; multiple modules exist for storing Basic credentials in flat text files, GDBM files, MySQL databases, LDAP directories, and so on.

Setting up Digest authentication is a little more involved. For one thing, the credentials are not transportable to other realms; when you generate them, you specify the realm to which they apply. For another, the only storage mechanism currently supported directly by the Apache package is flat text files; if you want to keep your Digest credentials in an LDAP directory or Oracle database, you're going to have to look for third-party modules to do it or else write one yourself.

In addition to the more complex setup process, Digest authentication currently suffers from a lack of market penetration. That is, even though Apache supports it, not all browsers and other web clients do; so you may end up having to use Basic authentication simply, because there's nothing else available to your users.

### See Also

- 

[Recipe 6.18](#)



## Recipe 6.17 Accessing Credentials Embedded in URLs

### Problem

You know people access your site using URLs with embedded credentials, such as `http://user:password@host/`, and you want to extract them from the URL for validation or other purposes.

### Solution

None; this is a nonissue that is often misunderstood.

### Discussion

For nonproxy requests, this doesn't even exist; the browser dissects the URL and turns it into the appropriate request header fields (i.e., `WWW-Authenticate`). For proxy requests, who knows?

## Recipe 6.18 Securing WebDAV

### Problem

You want to allow your users to upload and otherwise manage their web documents with WebDAV, but without exposing your server to any additional security risks.

### Solution

Require authentication to use WebDAV:

```
<Directory "/www/htdocs/dav-test">
  Order Allow,Deny
  Deny from all
  AuthDigestFile /www/acl/.htpasswd-dav-test
  AuthDigestDomain /dav-test/
  AuthName "DAV access"
  Require
  Satisfy any
</Directory>
```

### Discussion

Because WebDAV operations can modify your server's resources and *mod\_dav* runs as part of the server, locations that are WebDAV-enabled need to be writable by the user specified in the server's *User* directive. This means that the same location is writable by any CGI scripts or other modules that run as part of the Apache server. To keep remote modification operations under control, you should enable access controls for WebDAV-enabled locations. If you use weak controls, such as user-level authentication, you should use Digest authentication rather than Basic, as shown in the Solution.

The contents of the *<Directory>* container could be put into a *dav-test/.htaccess* file, as well. Note that the authentication database (specified with the *AuthDigestFile* directive) is not within the server's URI space, and so it cannot be fetched with a browser nor with any WebDAV tools.

Your authentication database and *.htaccess* files should not be modifiable by the server user; you don't want them getting changed by your WebDAV users!

### See Also

- 

[Recipe 6.16](#)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 6.19 Enabling WebDAV Without Making Files Writable by the Web User

### Problem

You want to run WebDAV but don't want to make your document files writable by the Apache server user.

### Solution

Run two web servers as different users. The DAV-enabled server, for example, might run as *User dav*, *Group dav*, while the other server, which is responsible for serving your content, might run as *User nobody*, *Group nobody*. Make the web content writable by the dav user, or the dav group.



Remember that only a single web server can be handling a particular port/IP address combination. This means that your WebDAV-enabled server will have to be using either a different address, a different port, or both than the non-WebDAV server.

### Discussion

A big security concern with DAV is that the content must be modifiable by the web server user for DAV to be able to update that content. This means that any content can also be edited by CGI programs, SSI directives, or other programs running under the web server. While the Apache security guidelines caution against having any files writable by the web server user, DAV requires it.

By running two Apache servers, you can move around this limitation. The DAV-enabled web server, running on an alternate port, has the *User* and *Group* directives set to an alternate user and group, such as:

```
User dav
Group dav
```

which is the owner of the web content in question. The other web server, which will be responsible for serving content to users, runs as a user who does not have permission to write to any of the documents.

The DAV-enabled web server should be well authenticated, so that only those who are permitted to edit the site can access that portion of the server. You should probably also set up this server to be very lightweight, both in the modules that you install as well as in the number of child processes (or threads) that you run.

Finally, it should be noted that the *perchild* MPM, under Apache 2.0, supports the idea of running different virtual hosts with different user ids, so that this recipe could be accomplished by enabling DAV just for the one particular vhost. However, as of this writing, the *perchild* MPM is not working yet.

### See Also



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 6.20 Restricting Proxy Access to Certain URLs

## Problem

You don't want people using your proxy server to access particular URLs or patterns of URLs (such as MP3 or streaming video files).

## Solution

You can block by keyword:

```
ProxyBlock .rm .ra .mp3
```

You can block by specific backend URLs:

```
<Directory proxy:http://other-host.org/path>
    Order Allow,Deny
    Deny from all
    Satisfy All
</Directory>
```

Or you can block according to regular expression pattern matching:

```
<Directory proxy:*>
    RewriteEngine On
    #
    # Disable proxy access to Real movie and audio files
    #
    RewriteRule "\.(rm|ra)$" "-" [F,NC]
    #
    # Don't allow anyone to access .mil sites through us
    #
    RewriteRule "[a-z]+://[-.a-z0-9]*\.mil($|/)" "-" [F,NC]
</Directory>
```

## Discussion

All of these solutions will result in a client that attempts to access a blocked URL receiving a 403 Forbidden status from the server.

The first solution uses a feature built into the proxy module itself: the *ProxyBlock* directive. It's simple and efficient, and it catches the results so that future accesses to the same URL are blocked with less effort; however, the pattern matching it can perform is extremely limited and prone to confusion. For instance, if you specify:

```
ProxyBlock .mil
```

the server denies access to both <http://www.navy.mil/> and <http://example.com/spec.mil/list.html>. This is probably not what was intended!

The second method allows you to impose limitations based on the URL being fetched (or gateway, in the case of a *ProxyPass* directive).

The third method, which allows more complex what-to-block patterns to be constructed, is both more flexible and more powerful, and somewhat less efficient. Use it only when the other methods prove insufficient.



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 6.21 Protecting Files with a Wrapper

## Problem

You have files to which you want to limit access using some method other than standard web authentication (such as a members-only area).

## Solution

In *httpd.conf*, add the following lines to a *<Directory>* container whose contents should be accessed only through a script:

```
RewriteEngine On
RewriteRule "\.(dll|zip|exe)$" protect.php [NC]
RewriteCond %{REMOTE_ADDR} "!^my.servers.ip"
RewriteRule "\.cgi$" protect.php [NC]
```

And an example *protect.php* that just displays the local URI of the document that was requested:

```
<?php
/*
 * The URL of the document actually requested is in
 * $_SERVER['REQUEST_URI']. Appropriate decisions
 * can be made about what to do from that.
 */
Header('Content-type: text/plain');
$body = sprintf("Document requested was: %s\n", $_SERVER['REQUEST_URI']);
Header('Content-length: ' . strlen($body));
print $body;
?>
```

## Discussion

In the situation that prompted this recipe, authentication and authorization were completed using a cookie rather than the standard mechanisms built into the web protocols. Any request for a document on the site was checked for the cookie and redirected to the login page if it wasn't found, was expired, or had some other problem causing its validity to be questioned.

This is fairly common and straightforward. What is needed in addition to this is a way to limit access to files according to the cookie and ensure that no URL-only request could reach them.

To this end, a wrapper is created (called *protect.php* in the Solution), which is invoked any time one of the protected document types is requested. After validating the cookie, the *protect.php* script figures out the name of the file from the environment variables, determines the content-type from the extension, and opens the file and sends the contents.

This is illustrated in the Solution. Any time a document ending in one of the extensions *.dll*, *.zip*, *.exe*, or *.cgi* is requested from the scope covered by the *mod\_rewrite* directives, and the request comes from some system other than the web server system itself (i.e., from a client system), the *protect.php* script will be invoked instead. In the Solution, the script simply displays the local URI of the document that is requested; applying additional access control or other functionality is easily developed from the example.

If access control is the main purpose of the wrapper and the access is granted, the wrapper needs to send the



## Recipe 6.22 Protecting All Files Except a Subset

### Problem

You want to deny all web access to files in a directory, except for those with a particular extension (i.e., a directory with HTML files in it, where you don't want other files to be accessible).

### Solution

Use a Files container in a Directory container to limit where authentication is required:

```
<Directory "/usr/local/apache/htdocs">
  Satisfy All
  Order allow,deny
  Deny from all
  <Files *.html>
    Order deny,allow
    Allow from all
    Satisfy Any
  </Files>
</Directory>
```

### Discussion

This method can be easily extended to apply to arbitrary filename patterns using shell global characters. To extend it to use regular expressions for the filename, use the *<FilesMatch>* directive instead.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_access.html](http://httpd.apache.org/docs/mod/mod_access.html)

## Recipe 6.23 Protecting Server Files from Malicious Scripts

### Problem

Scripts running on your web server may access, modify, or destroy files located on your web server if they are not adequately protected. You want to ensure that this cannot happen.

### Solution

Ensure that none of your files are writable by the nobody user or the nobody group, and that sensitive files are not readable by that user and group:

```
# find / -user nobody
# find / -group nobody
```

### Discussion

The *User* and *Group* directives specify a user and group under whose privileges the web server will run. These are often set to the values of nobody and nobody, respectively, but they can vary in different setups. It is often advisable to create a completely new user and group for this purpose, so that there is no chance that the user has been given additional privileges of which you are not aware.

Because everything runs with these privileges, any files or directories that are accessible by this user and/or group will be accessible from any script running on the server. This means that a script running under one virtual host may possibly modify or delete files contained within another virtual host, either intentionally or accidentally, if those files have permissions making this possible.

Ideally, no files anywhere on your server should be owned by, or writable by, the server user, unless for the explicit purpose of being used as a datafile by a script. And, even for this purpose, it is recommended that a real database be used, so that the file itself cannot be modified by the server user. And if files simply must be writable by the server, they should definitely not be in some web-accessible location, such as */cgi-bin/*.

### See Also

- [Recipe 8.12](#)
- [Recipe 6.24](#)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 6.24 Setting Correct File Permissions

## Problem

You want to set file permissions to provide the maximum level of security.

## Solution

The *bin* directory under the *ServerRoot* should be owned by user *root*, group *root*, and have file permissions of *755* (*rwxr-xr-x*). Files contained therein should also be owned by *root.root* and be mode *755*.

Document directories, such as *htdocs*, *cgi-bin*, and *icons*, will have to have permissions set in a way that makes the most sense for the development model of your particular web site, but under no circumstances should any of these directories or files contained in them be writable by the web server user.



The solution provided here is specific to Unixish systems. Users of other operating systems should adhere to the principles laid out here, although the actual implementation will vary.

The *conf* directory should be readable and writable only by *root*, as should all the files contained therein.

The *include* and *libexec* directories should be readable by everyone, writable by no one.

The *logs* directory should be owned and writable by *root*. You may, if you like, permit other users to read files in this directory, as it is often useful for users to be able to access their logfiles, particularly for troubleshooting purposes.

The *man* directory should be readable by all users.

Finally, the *proxy* directory should be owned by and writable by the server user.



On most Unixish file systems, a directory must have the *x* bit set in order for the files therein to be visible.

## Discussion

You should be aware that if you ask 12 people for the correct ways to set file permissions on your Apache server, you will get a dozen different answers. The recommendations here are intended to be as paranoid as possible. You should feel free to relax these recommendations, based on your particular view of the world and how much you trust





# Recipe 6.25 Running a Minimal Module Set

## Problem

You want to eliminate all modules that you don't need in order to reduce the potential exposure to security holes. What modules do you really need?

## Solution

For Apache 1.3, you can run a bare-bones server with just three modules. (Actually, you can get away with not running any modules at all, but it is not recommended.)

```
% ./configure --disable-module=all --enable-module=dir \  
> --enable-module=mime --enable-module=log_config \  

```

For Apache 2.0, this is slightly more complicated, as you must individually disable modules you don't want:

```
% ./configure --disable-access \  
> --disable-auth --disable-charset-lite \  
> --disable-include --disable-log-config --disable-env --disable-setenvif \  
> --disable-mime --disable-status --disable-autoindex --disable-asis \  
> --disable-cgid --disable-cgi --disable-negotiation --disable-dir \  
> --disable-imap --disable-actions --disable-alias --disable-userdir
```

Note that with 2.0, as with 1.3, you may wish to enable *mod\_dir*, *mod\_mime*, and *mod\_log\_config*, by simply leaving them off of this listing.

## Discussion

A frequent security recommendation is that you eliminate everything that you don't need; if you don't need something and don't use it, then you are likely to overlook security announcements about it or forget to configure it securely. The question that is less frequently answered is exactly what you do and don't need.

A number of Apache package distributions come with everything enabled, and people end up running modules that they don't really need—or perhaps are not even aware that they are running.

This recipe is an attempt to get to the very smallest Apache server possible, reducing it to the minimum set of modules that Apache will run. That is, if you take any of these out, Apache will not even start up, let alone serve a functional web site.

### Apache 1.3

With Apache 1.3, this question is fairly easy to answer. We've reduced it to a set of three modules, and, actually, you can eliminate all of the modules if you really want to, as long as you're aware of the implications of doing so.

*mod\_dir* is the module that takes a request for / and turns it into a request for */index.html*, or whatever other file you have indicated with the *DirectoryIndex* directive as the default document for a directory. Without this module, users typing just your hostname into their browser will immediately get a 404 error, rather than a default document. Granted, you could require that users specify a hostname and filename in their URL, in which case you could dispense with this module requirement. This would, however, make your web site fairly hard to use.



## Recipe 6.26 Restricting Access to Files Outside Your Web Root

### Problem

You want to make sure that files outside of your web directory are not accessible.

### Solution

For Unixish systems:

```
<Directory />
  Order deny,allow
  Deny from all
  AllowOverride None
  Options None
</Directory>
```

For Windows systems:

```
<Directory C:/>
  Order deny,allow
  Deny from all
  AllowOverride None
  Options None
</Directory>
```

Repeat for each drive letter on the system.

### Discussion

Good security technique is to deny access to everything, and then selectively permit access where it is needed. By placing a *Deny from all* directive on the entire filesystem, you ensure that files cannot be loaded from any part of your filesystem unless you explicitly permit it, using a *Allow from all* directive applied to some other *<Directory>* section in your configuration.

If you wanted to create an *Alias* to some other section of your filesystem, you would need to explicitly permit this with the following:

```
Alias /example /var/example
<Directory /var/example>
  Order allow,deny
  Allow from all
</Directory>
```

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_access.html](http://httpd.apache.org/docs/mod/mod_access.html)



## Recipe 6.27 Limiting Methods by User

### Problem

You want to allow some users to use certain methods but prevent their use by others. For instance, you might want users in group A to be able to use both GET and POST but allow everyone else to use only GET.

### Solution

Apply user authentication per method using the *Limit* directive:

```
AuthName "Restricted Access"
AuthType Basic
AuthUserFile filename
Order Deny,Allow
Allow from all
<Limit GET>
    Satisfy Any
</Limit>
<LimitExcept GET>
    Satisfy All
    Require valid-user
</Limit>
```

### Discussion

It is often desirable to give general access to one or more HTTP methods, while restricting others. For example, while you may wish any user to be able to GET certain documents, you may wish for only site administrators to POST data back to those documents.

It is important to use the *LimitExcept* directive, rather than attempting to enumerate all possible methods, as you're likely to miss one.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_auth.html](http://httpd.apache.org/docs/mod/mod_auth.html)

- 

[http://httpd.apache.org/docs/mod/mod\\_access.html](http://httpd.apache.org/docs/mod/mod_access.html)

- 

<http://httpd.apache.org/docs/mod/core.html#limit>

- 

<http://httpd.apache.org/docs/mod/core.html#limitexcept>





## Recipe 6.28 Restricting Range Requests

### Problem

You want to prevent clients from requesting partial downloads of documents within a particular scope, forcing them to request the entire document instead.

### Solution

You can overload `ErrorDocument 403` to make it handle range requests. To do this, put the following into the appropriate `<Directory>` container in your `httpd.conf` file or in the directory's `.htaccess` file:

```
SetEnvIf "Range" "." partial_requests
Order Allow,Deny
Allow from all
Deny from env=partial_requests
ErrorDocument 403 /forbidden.cgi
```

Then put the following into a file named `forbidden.cgi` in your server's `DocumentRoot`:

```
#!/usr/bin/perl -w
use strict;
my $message;
my $status_line;
my $body;
my $uri = $ENV{'REDIRECT_REQUEST_URI'} || $ENV{'REQUEST_URI'};
my $range = $ENV{'REDIRECT_HTTP_RANGE'} || $ENV{'HTTP_RANGE'};
if (defined($range)) {
    $body = "You don't have permission to access "
        . $ENV{'REQUEST_URI'}
        . " on this server.\r\n";
    $status_line = '403 Forbidden';
}
else {
    $body = "Range requests disallowed for document '"
        . $ENV{'REQUEST_URI'}
        . "'.\r\n";
    $status_line = '416 Range request not permitted';
}
print "Status: $status_line\r\n"
    . "Content-type: text/plain;charset=iso-8859-1\r\n"
    . "Content-length: " . length($body) . "\r\n"
    . "\r\n"
    . $body;
exit(0);
```

Or use `mod_rewrite` to catch requests with a `Range` header. To do this, put the following into the appropriate `<Directory>` container in your `httpd.conf` file or in the directory's `.htaccess` file:

```
RewriteEngine On
RewriteCond "%{HTTP:RANGE}" "."
RewriteRule "(.*)" "/range-disallowed.cgi" [L,PT]
```

Then put the following into a file named `range-disallowed.cgi` in your server's `DocumentRoot`:

```
#!/usr/bin/perl -w
use strict;
my $message = "Range requests disallowed for document '"
    . $ENV{'REQUEST_URI'}
    . "'.\r\n";
print "Status: 416 Range request not permitted\r\n"
    . "Content-type: text/plain;charset=iso-8859-1\r\n"
```



# Chapter 7. SSL

Secure Socket Layers (SSL) is the standard way to implement secure web sites. By encrypting the traffic between the server and the client, which is what SSL does, that content is protected from a third party listening to the traffic going past.

The exact mechanism by which this encryption is accomplished is discussed extensively in the SSL specification, which you can read at <http://wp.netscape.com/eng/ssl3/>. For a more user-friendly discussion of SSL, we recommend looking through the *mod\_ssl* manual, which you can find at <http://www.modssl.org/docs/2.8/index.html>. This document discusses not only the specific details of setting up *mod\_ssl*, but also covers the general theory behind SSL it and has pictures illustrating the concepts.

In this chapter, we talk about some of the common things that you might want to do with your secure server, including how to install it.



# Recipe 7.1 Installing SSL

## Problem

You want to install SSL on your Apache server.

## Solution

The solutions to this problem fall into several categories, depending on how you installed Apache in the first place (or whether you are willing to rebuild Apache to get SSL).

If you installed a binary distribution of Apache, your best bet is to return to the place from which you acquired that binary distribution, and try to find the necessary files for adding SSL to it.

If you built Apache yourself from source, then the solution will depend on whether you are running Apache 1.3 or Apache 2.0.

In Apache 1.3, SSL is an add-on module, which you must acquire and install from a different location than that from where you obtained Apache. There are two main choices available: *mod\_ssl* (<http://www.modssl.org/>) and Apache-SSL (<http://www.apache-ssl.org/>); the installation procedure will vary somewhat depending on which one of these you choose.

If you are building Apache 2.0 from source, the situation is somewhat simpler; just add `—enable-ssl` to the `./configure` arguments when you build Apache to include SSL as one of the built-in modules.

Consult [Chapter 1](#) and [Chapter 2](#) for more information on installing third-party modules, particularly if you have installed a binary distribution of Apache rather than building it yourself from the source code.

If you are attempting to install SSL on Apache for Windows, there is a discussion of this in the Compiling on Windows document, which you can find at [http://httpd.apache.org/docs-2.0/platform/win\\_compiling.html](http://httpd.apache.org/docs-2.0/platform/win_compiling.html) for Apache 2.0. Or, if you are using Apache 1.3 on Windows and wish to install SSL, you should consult the file *INSTALL.Win32*, which comes with the SSL distribution, or look at the HowTo at <http://tud.at/programm/apache-ssl-win32-howto.php3>.

Finally, note that the Apache SSL modules are an interface between Apache and the OpenSSL libraries, which you must install before any of this can work. You can obtain the OpenSSL libraries from <http://www.openssl.org/>. Although you may already have these libraries installed on your server, it is recommended that you obtain the latest version of the libraries to have the most recent security patches and to protect yourself from exploits.

## Discussion

So, why is this so complicated? Well, there are a variety of reasons, most of which revolve around the legality of encryption. For a long time, encryption has been a restricted technology in the U.S.. Since Apache is primarily based out of the U.S., there is a great deal of caution regarding distributing encryption technology with the package. Even though the U.S. has relaxed its restrictions on exporting SSL technology, it is still a sensitive area.



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 7.2 Generating SSL Certificates

### Problem

You want to generate certificates to use on your SSL server.

### Solution

Use the *openssl* command-line program that comes with OpenSSL:

```
% openssl genrsa -out hostname.key 1024
% openssl req -new -key hostname.key -out hostname.csr
```

At this point, you can either send your Certificate Signing Request (CSR) off to one of the certificate authority companies, such as Thawte or Entrust, for them to sign, or, if you prefer, you can sign the key yourself:

```
% openssl x509 -req -days 365 -in hostname.csr -signkey hostname.key
    -out hostname.crt
```

Then move these files to your Apache server's configuration directory, such as */www/conf/*, and then add the following lines in your *httpd.conf* configuration file:

```
SSLCertificateFile /www/conf/hostname.crt
SSLCertificateKeyFile /www/conf/hostname.key
```

### Discussion

The SSL certificate is a central part of the SSL conversation and is required before you can run a secure server. Thus, generating the certificate is a necessary first step to configuring your secure server.

Generating the key is a multistep process, but it is fairly simple.

#### Generating the private key

In the first step, we generate the private key. SSL is a private/public key encryption system, with the private key residing on the server and the public key going out with each connection to the server and encrypting data sent back to the server.

The first argument passed to the *openssl* program tells *openssl* that we want to generate an RSA key (*genrsa*), which is an encryption algorithm that all major browsers support.

The next argument gives *openssl* something to use as the source of randomness. The *-rand* flag will accept one or more filenames, which will be used as a key for the random number generator. If no *-rand* argument is provided, OpenSSL will attempt to use */dev/urandom* by default if that exists, and it will try */dev/random* if */dev/urandom* does not exist. It is important to have a good source of randomness in order for the encryption to be secure. If your system has neither */dev/urandom* nor */dev/random*, you should consider installing a random number generator, such as *egd*. You can find out more information about this on the OpenSSL web site at [http://www.openssl.org/docs/crypto/RAND\\_egd.html](http://www.openssl.org/docs/crypto/RAND_egd.html).

The *-out* argument specifies the name of the key file that we will generate. This file will be created in the directory in



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 7.3 Generating a Trusted CA

## Problem

You want to generate SSL keys that browsers will accept without a warning message.

## Solution

Issue the following commands:

```
% CA.pl -newca
% CA.pl -newreq
% CA.pl -signreq
% CA.pl -pkcs12
```

## Discussion

[Recipe 7.2](#) discusses the lengthy steps that are required to create keys and sign them. Fortunately, OpenSSL comes with a script to automate much of this process, so that you don't have to remember all of those arguments. This script, called *CA.pl*, is located where your SSL libraries are installed, for example, */usr/share/ssl/misc/CA.pl*.

The lines in the Solution hide a certain amount of detail, as you will be asked a number of questions in the process of creating the key and the certificate. Note also that you will probably need to be in the directory where this script lives to get successful results from this recipe.

If you want to omit the passphrase on the certificate so that you don't have to provide the passphrase each time you start up the server, use *-newreq-nodes* rather than *-newreq* when generating the certificate request.

After running this sequence of commands, you can generate more certificates by repeating the *-newreq* and *-signreq* commands.

Having run these commands, you will have generated a number of files. The file *newcert.pem* is the file you specify in your *SSLCertificateFile* directive, the file *newreq.pem* is your *SSLCertificateKeyFile*, and the file *demoCA/cacert.pem* is the CA certificate file, which will need to be imported into your users' browsers (for some browsers) so that they can automatically trust certificates signed by this CA. And, finally, *newcert.p12* serves the same purpose as *demoCA/cacert.pem* for certain other browsers.

## Importing the CA

If your users are using Internet Explorer, you need to create a special file for them to import. Use the following command:

```
openssl X509 -demoCA/cacert.pem -out cacert.crt -outform DER
```

Then you can send them the *cacert.crt* file.

Clicking on that file will launch the SSL certificate wizard and guide the user through installing the CA certificate into their browser.



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 7.4 Serving a Portion of Your Site via SSL

## Problem

You want to have a certain portion of your site available via SSL exclusively.

## Solution

This is done by making changes to your *httpd.conf* file.

For Apache 1.3, add a line such as the following:

```
Redirect /secure/ https://secure.domain.com/secure/
```

For Apache 2.0:

```
<Directory /www/secure>
    SSLRequireSSL
</Directory>
```

Or, with *mod\_rewrite*:

```
RewriteEngine On
RewriteCond %{HTTPS} !=on
RewriteRule ^/(.*) https://%{SERVER_NAME}/$1 [R,L]
```

## Discussion

It is perhaps best to think of your site's normal pages and its SSL-protected pages as being handled by two separate servers, rather than one. While they may point to the same content, they run on different ports, are configured differently, and, most importantly, the browser considers them to be completely separate servers. So you should too.

Don't think of enabling SSL for a particular directory; rather, you should think of it as redirecting requests for one directory to another.

Note that the *Redirect* directive preserves path information, which means that if a request is made for */secure/something.html*, then the redirect will be to <https://secure.domain.com/secure/something.html>.

Be careful where you put this directive. Make sure that you only put it in the HTTP (non-SSL) virtual host declaration. Putting it in the global section of the *config* file may cause looping, as the new URL will match the *Redirect* requirement and get redirected itself.

Finally, note that if you want the entire site to be available only via SSL, you can accomplish this by simply redirecting all URLs, rather than a particular directory:

```
Redirect / https://secure.domain.com/
```

Again, be sure to put that inside the non-SSL virtual host declaration.

You will see various solutions proposed for this situation using *RedirectMatch* or various *RewriteRule* directives. There are special cases where this is necessary, but in most cases, the simple solution offered here works just fine.



## Recipe 7.5 Authenticating with Client Certificates

### Problem

You want to use client certificates to authenticate access to your site.

### Solution

Add the following *mod\_ssl* directives to your *httpd.conf* file:

```
SSLVerifyClient require
SSLVerifyDepth 1
SSLCACertificateFile conf/ssl.crt/ca.crt
```

### Discussion

If you happen to be lucky enough to have a small, closed user community, such as an intranet, or a web site for a group of friends or family, it is possible to distribute client certificates so that each user can identify himself.

Create client certificates, signing them with your CA certificate file, and then specify the location of this CA certificate file using the *CACertificateFile* directive, as shown above.

### See Also

- 

- [Recipe 7.2](#)

- 

[http://httpd.apache.org/docs-2.0/mod/mod\\_ssl.html](http://httpd.apache.org/docs-2.0/mod/mod_ssl.html)

# Chapter 8. Dynamic Content

CGI programs are one of the simplest ways to provide dynamic content for your web site. They tend to be easy to write, because you can write them in any language. Thus, you don't have to learn a new language to write CGI programs.

Other dynamic content providers, such as PHP and mod\_perl, also enjoy a great deal of popularity, because they provide many of the same functions as CGI programs but typically execute faster.

Very few web sites can survive without some mechanism for providing dynamic content—content that is generated in response to the needs of the user. The recipes in this chapter guide you through enabling various mechanisms for producing this dynamic content and help you troubleshoot possible problems that may occur.

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 8.1 Enabling a CGI Directory

## Problem

You want to designate a directory that contains only CGI scripts.

## Solution

Add the following to your *httpd.conf* file:

```
ScriptAlias /cgi-bin/ /www/cgi-bin/
```

## Discussion

A CGI directory will usually be designated and enabled in your default configuration file when you install Apache. However, if you want to add additional directories where CGI programs are permitted, the *ScriptAlias* directive does this for you. You may have as many *ScriptAlias*'ed directories as you want.

The one line previously introduced is equivalent to these directive lines:

```
Alias /cgi-bin/ /www/cgi-bin/
```

```
<Location /cgi-bin/>  
    Options ExecCGI  
    SetHandler cgi-script  
</Location>
```



Note that URLs that map to the directory in question via some other mechanism, such as another *Alias* or a *RewriteRule*, will not benefit from the *ScriptAlias* setting, as this mapping is by URL (<Location>), not by directory. As a result, accessing the scripts in this directory through some other URL path may result in their code being displayed rather than the script being executed.

## See Also

- 

- [Chapter 5](#)

- 

- [Recipe 8.2](#)

- 

[http://httpd.apache.org/docs/mod/mod\\_alias.html](http://httpd.apache.org/docs/mod/mod_alias.html)

## Recipe 8.2 Enabling CGI Scripts in Non-ScriptAliased Directories

### Problem

You want to put a CGI program in a directory that contains non-CGI documents.

### Solution

Use *AddHandler* to map the CGI handler to the particular files that you want to be executed:

```
<Directory "/foo">
  Options +ExecCGI
  AddHandler cgi-script .cgi .py .pl
</Directory>
```

### Discussion

Enabling CGI execution via the *ScriptAlias* directive is preferred, for a number of reasons, over permitting CGI execution in arbitrary document directories. The primary reason is security auditing. It is much easier to audit your CGI programs if you know where they are, and storing them all in a single directory ensures that.

However, there are cases where it is desirable to have this functionality. For example, you may want to keep several files together in one directory—some of them static documents, and some of them scripts—because they are part of a single application.

Using the *AddHandler* directive maps certain file extensions to the *cgi-script* handler so they can be executed as CGI programs. In the case of the aforementioned example, programs with a *.cgi*, *.py*, or *.pl* file extension will be treated as CGI programs, while all other documents in the directory will be served up with their usual MIME type.

Note that the *+ExecCGI* argument is provided to the *Options* directive, rather than the *ExecCGI* argument—that is, with the *+* sign rather than without. Using the *+* sign adds this option to any others already in place, whereas using the option without the *+* sign will replace the existing list of options. You should use the argument without the *+* sign if you intend to have only CGI programs in the directory, and with the *+* sign if you intend to also serve non-CGI documents out of the same directory.

### See Also

- 

[Recipe 8.1](#)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 8.3 Using Windows File Extensions to Launch CGI Programs

## Problem

You want to have CGI programs on Windows executed by the program associated with the file extension. For example, you want *.pl* files to be executed by *perl.exe* without having to change the *#!* line to point at the right location.

## Solution

Add the following line to your *httpd.conf* file:

```
ScriptInterpreterSource registry
```

## Discussion

Since Apache has its roots in the Unixish world, there are a number of things that are done the Unixish way, even on Microsoft Windows. CGI execution is one of these things, but the *ScriptInterpreterSource* directive allows you to have Apache behave more in the way that Windows users are accustomed to.

Usually, on Windows, a file type is indicated by the file extension. For example, a file named *example.pl* is associated with the Perl executable; when a user clicks on this file in the file explorer, Perl is invoked to execute this script. This association is created when you install a particular program, such as Perl or MS Word, and the association is stored in the Windows registry.

On Unixish systems, on the other hand, most scripts contain the location of their interpreter in the first line of the file, which starts with the characters *#!*. This line is often called the shebang line (short for sharp bang, which are the shorthand names for the two characters).

For example, a Perl program might start with the line:

```
#!/usr/bin/perl
```

The shell running the script looks in this first line and uses the program at the indicated path to interpret and execute the script. In this way, files with arbitrary file extensions (or no extension at all) may be invoked with any interpreter desired. In the case of Perl, for example, one might have several versions of Perl installed, and the particular version desired may be invoked by using the appropriate *#!* line.

However, you may be accustomed to the operating system's innate way of executing a program, and this can be somewhat nonintuitive. Thus, in the early days of Apache on Windows, the *ScriptInterpreterSource* directive was added to make Apache behave the way that Windows users expected.

*ScriptInterpreterSource* may have one of two values. When set to the default value, *script*, Apache will look in the script itself for the location of the interpreter that it is to use. When it is set to *registry*, it will look in the Windows registry for the mapping that is associated with the file's extension and use this to execute the script.

This feature can be very useful for users who are running multiple servers, some on Unixish operating systems and others on Windows, but who want the same CGI programs to run both places. Because Perl is unlikely to be located



## Recipe 8.4 Using Extensions to Identify CGI Scripts

### Problem

You want Apache to know that all files with a particular extension should be treated as CGI scripts.

### Solution

Add the following to your *httpd.conf* file in a scope covering the areas where it should apply, or in an *.htaccess* file for the appropriate directory:

```
AddHandler cgi-script .cgi
```

### Discussion

The *AddHandler* directive shown in this solution tells Apache that any files that have a *.cgi* extension should be treated as CGI scripts, and it should try to execute them rather than treat them as content to be sent.

The directive only affects files with that extension in the same scope as the directive itself. You may replace the common *.cgi* extension with another, or even with a list of space-separated extensions.

Note the use of the term extension rather than suffix; a file named *foo.cgi.en* is treated as a CGI script unless a handler with the *.en* extension overrides it.

### See Also

- 

[Recipe 8.2](#)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 8.5 Testing That CGI Is Set Up Correctly

## Problem

You want to test that you have CGI enabled correctly. Alternatively, you are receiving an error message when you try to run your CGI script and you want to ensure the problem doesn't lie in the web server before you try to find a problem in the script.

## Solution

```
#!/usr/bin/perl
print "Content-type: text/plain\r\n\r\n";
print "It's working.\n";
```

And then, if things are still not working, look in the error log.

## Discussion

Because Perl is likely to be installed on any Unixish system, this CGI program should be a pretty safe way to test that CGI is configured correctly. In the event that you do not have Perl installed, an equivalent shell program may be substituted:

```
#!/bin/sh
echo Content-type: text/plain
echo
echo It's working.
```

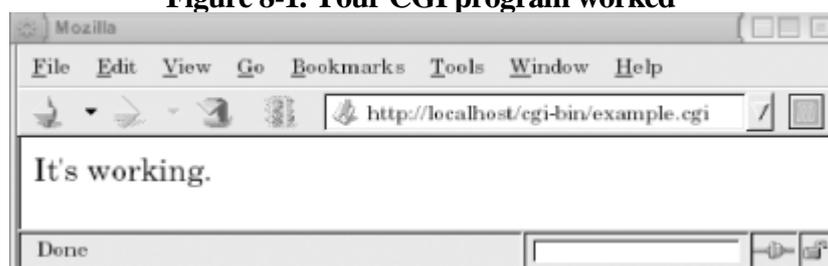
And, if you are running Apache on Windows, so that neither of the above options works for you, you could also try this with a batch file:

```
echo off
echo Content-type: text/plain
echo.
echo It's working.
```

Make sure that you copy the program code exactly, with all the right punctuation, slashes, etc., so that you don't introduce additional complexity by having to troubleshoot the program itself.

In either case, once the program is working, you should see something like the following screen capture (see [Figure 8-1](#)).

**Figure 8-1. Your CGI program worked**



The idea here is to start with the simplest possible CGI program to ensure that problems are not caused by other complexities in your code. We want to ensure that CGI is configured properly, not to verify the correctness of a particular CGI program.





# Recipe 8.6 Reading Form Parameters

## Problem

You want your CGI program to read values from forms for use in your program.

## Solution

First, look at an example in Perl, which uses the popular *CGI.pm* module:

```
#!/usr/bin/perl
use CGI;
use strict;
use warnings;

my $query = CGI->new;

# Load the various form parameters
my $name = $form->param("name");

# Multi-value select lists will return a list
my @foods = $form->param("favorite_foods");

# Output useful stuff
print "Content-type: text/html\n\n";
print "Name: " . $form->{name} . "\n";
print "Favorite foods: <ul>";
foreach my $food (@foods) {
    print "<li>$food</li>";
}
print "</ul>\n";
```

Next, look at the same program in C, which uses the *cgic* C library:

```
#include "cgic.h"
/* Boutell.com's cgic library */

int cgiMain( ) {
    char name[100];

    /* Send content type */
    cgiHeaderContentType("text/html");

    /* Load a particular variable */
    cgiFormStringNoNewlines("name", name, 100);
    fprintf(cgiOut, "Name: ");
    cgiHtmlEscape(name);

    return 0;
}
```

For this example, you will also need a *Makefile*, which looks something like this:

```
CFLAGS=-g -Wall
CC=gcc
AR=ar
LIBS=-L./ -lcgic

libcgic.a: cgic.o cgic.h
    rm -f libcgic.a
    $(AR) rc libcgic.a cgic.o
```





# Recipe 8.7 Invoking a CGI Program for Certain Content Types

## Problem

You want to invoke a CGI program to act as a sort of content filter for certain document types. For example, a photographer may wish to create a custom handler to add a watermark to photographs served from his web site.

## Solution

Use the *Action* directive to create a custom handler, which will be implemented by a CGI program. Then use the *AddHandler* directive to associate a particular file extension with this handler:

```
Action watermark /cgi-bin/watermark.cgi
AddHandler watermark .gif .jpg
```

## Discussion

This recipe creates a watermark handler that is called whenever a *.gif* or *.jpg* file is requested.

A CGI program, *watermark.cgi*, takes the image file as input and attaches the watermark image on top of the photograph. The path to the image file that was originally requested in the URL is available in the `PATH_TRANSLATED` environment variable, and the program needs to load that file, make the necessary modifications, and send the resulting content to the client, along with the appropriate HTTP headers.

Note that there is no way to circumvent this measure, as the CGI program will be called for any *.gif* or *.jpg* file that is requested.

This same technique may be used to attach a header or footer to HTML pages in an automatic way, without having to add any kind of SSI directive to the files. This can be extremely inefficient, as it requires that a CGI program be launched, which can be a very slow process. It is, however, constructive to see how it is done. What follows is a very simple implementation of such a footer script:

```
#!/usr/bin/perl

print "Content-type: text/html\r\n\r\n";

my $file = $ENV{PATH_TRANSLATED};

open FILE, "$file";
print while <FILE>;
close FILE;
print qq~

<p>
FOOTER GOES HERE
~;
```

The requested file, located at `PATH_TRANSLATED`, is read in and printed out, unmodified. Then, at the end of it, a few additional lines of footer are output. A similar technique might be used to filter the contents of the page itself. With Apache 2.0, this may be better accomplished with *mod\_ext\_filter*.

## See Also





# Recipe 8.8 Getting SSIs to Work

## Problem

You want to enable Server-Side Includes (SSIs) to make your HTML documents more dynamic.

## Solution

There are at least two different ways of doing this.

Specify which files are to be parsed by using a filename extension such as *.shtml*. For Apache 1.3, add the following directives to your *httpd.conf* in the appropriate scope:

```
<Directory /www/html/example>
  Options +Includes
  AddHandler server-parsed .shtml
  AddType "text/html; charset=ISO-8859-1" .shtml
</Directory>
```

Or, for Apache 2.0:

```
<Directory /www/html/example>
  Options +Includes
  AddType text/html .shtml
  AddFilter INCLUDES .shtml
</Directory>
```

Add the *XBitHack* directive to the appropriate scope in your *httpd.conf* file and allow the file permissions to indicate which files are to be parsed for SSI directives:

```
XBitHack On
```

## Discussion

SSIs provide a way to add dynamic content to an HTML page via a variety of simple tags. This functionality is implemented by the *mod\_include* module, which is documented at

[http://httpd.apache.org/docs/mod/mod\\_include.html](http://httpd.apache.org/docs/mod/mod_include.html). There is also a howto-style document available at <http://httpd.apache.org/docs/howto/ssi.html>.

The first solution provided here tells Apache to parse all *.shtml* files for SSI directives. So, to test that the solution has been effective, create a file called *something.shtml*, and put the following line in it:

```
File last modified at '<!--#echo "LAST_MODIFIED" -->'.
```



Note the space between the last argument and the closing "-->". This space is surprisingly important; many SSI failures can be traced to its omission.

Accessing this document via your server should result in the page displaying the date and time when you modified (or created) the file.

If you wish to enable SSIs, but do not wish to permit execution of CGI scripts, or other commands using the *#exec*



## Recipe 8.9 Displaying Last Modified Date

### Problem

You want your web page to indicate when it was last modified but not have to update the date every time.

### Solution

Use SSI processing by putting a line in the HTML file for which you want the information displayed:

```
<!--#config timefmt="%B %e, %Y" -->
```

```
This document was last modified on <!--#echo var="LAST_MODIFIED" -->
```

### Discussion

The *config* SSI directive allows you to configure a few settings governing SSI output formats. In this case, we're using it to configure the format in which date/time information is output. The default format for date output is 04-Dec-2037 19:58:15 EST which is not the most user-friendly message. The recipe provided changes this to the slightly more readable format December 4, 2002. If you want another output format, the *timefmt* attribute can take any argument accepted by the C *strftime(3)* function.

### See Also

- 

[Recipe 8.8](#)

- 

The *strftime(3)* documentation

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 8.10 Including a Standard Header

## Problem

You want to include a header (or footer) in each of your HTML documents.

## Solution

Use SSI by inserting a line in all your parsed files:

```
<--#include virtual="/include/headers.html" -->
```

## Discussion

By using the SSI *include* directive, you can have a single header file that can be used throughout your web site. When your header needs to be modified, you can make this change in one place and have it go into effect immediately across your whole site.

The argument to the *virtual* attribute is a local URI and subject to all normal *Alias*, *ScriptAlias*, *RewriteRule*, and other commands, which means that:

```
<--#include virtual="/index.html" -->
```

will include the file from your *DocumentRoot*, and:

```
<--#include virtual="/cgi-bin/foo" -->
```

will include the output from the *foo* script in your server's *ScriptAlias* directory.

If the argument doesn't begin with a / character, it's treated as being relative to the location of the document using the *#include* directive.



Be aware that URIs passed to *#include virtual* may not begin with *../*, nor may they refer to full URLs such as `http://example.com/foo.html`. Documents included using relative syntax (i.e., those not beginning with */*) may only be in the same location as the including file, or in some sublocation underneath it. Server processing of the URI may result in the actual included document being located somewhere else, but the restrictions on the *#include virtual* SSI command syntax permit only same-location or descendent-location URIs.

## See Also

- 

- [Recipe 8.7](#)

- 

- [Recipe 8.8](#)



## Recipe 8.11 Including the Output of a CGI Program

### Problem

You want to have the output of a CGI program appear within the body of an existing HTML document.

### Solution

Use SSIs by adding a line such as the following to the document (which must be enabled for SSI parsing):

```
<--#include virtual="/cgi-bin/content.cgi" -->
```

### Discussion

The SSI *#include* directive, in addition to being able to include a plain file, can also include other dynamic content, such as CGI programs, other SSI documents, or content generated by any other method.

The *#exec* SSI directive may also be used to produce this effect, but for a variety of historical and security-related reasons, its use is deprecated. The *#include* directive is the preferred way to produce this effect.

Note that *Options IncludeNoExec*, in addition to disabling the *#exec* SSI directive, also forbids the inclusion of CGI programs using the *#include virtual* syntax. Thus, this method cannot be used to circumvent the restriction.

### See Also

- 

[Recipe 8.8](#)

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 8.12 Running CGI Scripts as a Different User with suexec

### Problem

You want to have CGI programs executed by some user other than nobody. For example, you may have a database that is not accessible to anyone except a particular user, so the server needs to temporarily assume that user's identity to access it.

### Solution

When building Apache, enable *suexec* by passing the `—enable-suexec` argument to *configure*.

Then, in a virtual host section, specify which user and group you'd like to use to run CGI programs:

```
User rbowen  
Group users
```

Also, *suexec* will be invoked for any CGI programs run out of username-type URLs for the affected virtual host.

### Discussion

The *suexec* wrapper is a *suid* (runs as the user ID of the user that owns the file) program that allows you to run CGI programs as any user you specify, rather than as the nobody user which Apache runs as. *suexec* is a standard part of Apache and is enabled by default.



The *suexec* concept does not fit well into the Windows environment, and so *suexec* is not available under Windows.

When *suexec* is installed, there are two different ways that it can be invoked, as shown in the Solution.

A *User* and *Group* directive may be specified in a *VirtualHost* container, and all CGI programs executed within the context of that virtual host are executed as that user and group. Note that this only applies to CGI programs. Documents are still accessed as the user and group specified in the *User* and *Group* directives in the main server configuration, not those in the virtual host, and need to be readable by that user and group.

Second, any CGI program run out of a *UserDir* directory is run with the permissions of the owner of that directory. That is, if a CGI program is accessed via the URL `http://example.com/~rbowen/cgi-bin/test.cgi`, then that program will be executed, via *suexec*, with a *userid* of *rbowen*, and a *groupid* of *rbowen*'s primary group.



If *UserDir* points to a nonstandard location, you must tell *suexec* about this when you build it. In a default configuration, *suexec* is invoked when CGI programs are invoked in a directory such as `/home/username/public_html/` for some *username*. If, however, you





## Recipe 8.13 Installing a `mod_perl` Handler from CPAN

### Problem

You want to install one of the many `mod_perl` handler modules available on CPAN. For example, you want to install the `Apache::Perldoc` module, which generates HTML documentation for any Perl module you happen to have installed.

### Solution

Assuming you already have `mod_perl` installed, you'll just need to install the module from CPAN, and then add a few lines to your Apache configuration file.

To install the module, run the following command from the shell as root:

```
# perl -MCPAN -e 'install Apache::Perldoc'
```

Then, in your Apache configuration file, add:

```
<Location /perldoc>
    SetHandler perl-script
    PerlHandler Apache::Perldoc
</Location>
```

After restarting Apache, you can access the handler by going to a URL such as `http://example.com/perldoc/Apache/Perldoc`.

### Discussion

The CPAN shell, which is installed when Perl is installed, gives you an easy way to install Perl modules from CPAN. CPAN, if you're not familiar with it, is the Comprehensive Perl Archive Network, at <http://cpan.org>, a comprehensive archive of Perl stuff, including Perl modules for every purpose you can imagine and several you can't. This includes a substantial number of `mod_perl` handlers.

The module specified in this recipe is a very simple one that gives you HTML documentation for any Perl module you have installed, accessible via your Apache server. Other ones provide photo albums, weblog handlers, and DNS zone management, among other things.

The first time you run the CPAN shell, you will need to answer a series of questions about your configuration, what CPAN server you want to get modules from, where it should find your FTP clients, and so on. This only happens once, then it just works.

The specific way that you need to configure Apache to use your newly-installed module will vary from one module to another, but many of them will look like the example given. The `SetHandler perl-script` directive tells Apache that the content will be handled by `mod_perl`, while the `PerlHandler` directive specifies what Perl module contains the actual handler code.

### See Also



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 8.14 Writing a `mod_perl` Handler

### Problem

You want to write your own `mod_perl` handler.

### Solution

Here's a simple handler:

```
package Apache::Cookbook::Example;

sub handler {
    my $r = shift;
    $r->send_http_header( 'text/plain' );
    $r->print( "Hello, World." );
}

1;
```

Place this code in a file called `Example.pm`, in a directory `Apache/Cookbook/`, somewhere that Perl knows to look for it.

### Discussion

The example handler given is fairly trivial and does not do anything useful. More useful examples may be obtained from the `mod_perl` web site (<http://perl.apache.org/>) and from Geoffrey Young's (et al.) excellent book `mod_perl Developer's Cookbook`. Also, although it is somewhat dated, the "Eagle book" (Writing Apache modules with Perl and C) by Lincoln Stein and Doug MacEachern (O'Reilly) is an excellent introduction to `mod_perl` and the Apache API.

The real question here, however, is how and where you should install the file that you've created. There are two answers to this question, and which one you choose will be largely personal preference.

When Perl looks for a module, it looks through the list called `@INC` for directories where that module might be. You can either put your module in one of those directories, or you can add a directory to the list.

To find out where Perl is looking, you can examine the values stored in `@INC` with the following:

```
perl -le 'print join "\n", @INC;'
```

This will give you a listing that will look something like:

```
/usr/local/lib/perl5/5.8.0/i686-linux
/usr/local/lib/perl5/5.8.0
/usr/local/lib/perl5/site_perl/5.8.0/i686-linux
/usr/local/lib/perl5/site_perl/5.8.0
/usr/local/lib/perl5/site_perl
.
```

This will of course vary from one system to another, from one version of Perl to another, but will bear some resemblance to that listing.



## Recipe 8.15 Enabling PHP Script Handling

### Problem

You want to enable PHP scripts on your server.

### Solution

If you have *mod\_php* installed, use *AddHandler* to map *.php* and *.phtml* files to the PHP handler:

```
AddHandler application/x-httpd-php .phtml .php
```

### Discussion

This recipe maps all files with *.phtml* or *.php* to the PHP handler. You must ensure that the *mod\_php* module is installed.

### See Also

- - [Recipe 2.5](#)
  -
- Installation instructions on the *mod\_php* web site at <http://www.php.net/manual/en/install.apache.php> for Apache 1.3 or <http://www.php.net/manual/en/install.apache2.php> for Apache 2.0

## Recipe 8.16 Verifying PHP Installation

### Problem

You want to verify that you have PHP correctly installed and configured.

### Solution

Put the following in your test PHP file:

```
<?php phpinfo( ); ?>
```

### Discussion

Place the above text in a file called *something.php* in a directory where you believe you have enabled PHP script execution. Accessing that file should give you a list of all configured PHP system variables. The first screen of the output should look something like [Figure 8-2](#).

**Figure 8-2. Sample phpinfo() output**



PHP Version 4.4.0-dev	
System	Linux [redacted] 2.4.20-19.7 #1 Tue Jul 15 13:44:14 EDT 2003 i686
Build Date	Dec 30 2002 08:11:57
Configure Command	./configure '--with-mysql=/usr' '--with-zlib=/usr' '--with-xml' '--with-gd' '--with-png' '--with-jpeg-dir=/usr' '--with-zlib-dir=/usr' '--enable-xml' '--enable-calendar' '--with-apxs=[redacted]/bin/apxs'
Server API	Apache
Virtual Directory Support	disabled
Configuration File (php.ini) Path	/usr/local/lib/php.ini
PHP API	20020910
PHP Extension	20020429
Zend Extension	20021010
Debug Build	no
Thread Safety	disabled
Registered PHP Streams	php, http, ftp, compress.zlib

This program makes use of the Zend Scripting Language Engine:  
Zend Engine v1.4.0, Copyright (c) 1998-2002 Zend Technologies

### See Also

- 

[Recipe 8.15](#)

# Chapter 9. Error Handling

When you're running a web site, things go wrong. And when they do, it's important that they are handled gracefully, so that the user experience is not too greatly diminished. In this chapter, you'll learn how to handle error conditions, return useful messages to the user, and capture information that will help you fix the problem so that it does not happen again.



# Recipe 9.1 Handling a Missing Host Field

## Problem

You have multiple virtual hosts in your configuration, and at least one of them is name-based. For name-based virtual hosts to work properly, the client must send a valid Host field in the request header. This recipe describes how you can deal with situations in which the field is not included.

## Solution

Add the following lines to your *httpd.conf* file:

```
Alias /NoHost.cgi /usr/local/apache/cgi-bin/NoHost.cgi
RewriteEngine On
RewriteCond "%{HTTP_HOST}" "^[^$]"
RewriteRule "(.*)" "/NoHost.cgi$1" [PT]
```

The file *NoHost.cgi* can contain something like the following:

```
#!/usr/bin/perl -Tw

my $msg = "To properly direct your request, this server requires that\n"
        . "your Web client include the HTTP 'Host' request header field.\n"
        . "The request which caused this response did not include such\n"
        . "a field, so we cannot determine the correct document for you.\n";
print "Status: 400 Bad Request\r\n"
    . "Content-type: text/plain\r\n"
    . "Content-length: " . length($msg) . "\r\n"
    . "\r\n"
    . $msg;
exit(0);
```

## Discussion

Once the directives in the solution are in place, all requests made of the server that do not include a Host: field in the request header are redirected to the specified CGI script, which can take appropriate action.

The solution uses a CGI script so that the response text can be tailored according to the attributes of the request and the server's environment. For instance, the script might respond with a list of links to valid sites on the server, determined by the script at runtime by examining the server's own configuration files. If all you need is a "please try again, this time with a Host: field" sort of message, a static HTML file would suffice:

```
RewriteRule .* /nohost.html [PT]
```

A more advanced version of the script approach could possibly scan the *httpd.conf* file for *ServerName* directives, construct a list of possibilities from them, and present links in a 300 Multiple Choices response. Of course, there's an excellent chance they wouldn't work, because the client would still not be including the Host: field.

## See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)



## Recipe 9.2 Changing the Response Status for CGI Scripts

### Problem

There may be times when you want to change the status for a response—for example, you want 404 Not Found errors to be sent back to the client as 403 Forbidden instead.

### Solution

Point your *ErrorDocument* to a CGI script instead of a static file. The CGI specification permits scripts to specify the response status code.

In addition to the other header fields the script emits, like the `Content-type:` field, include one named `Status:` with the value and text of the status you want to return:

```
#!/bin/perl -w
print "Content-type: text/html;charset=iso-8859-1\r\n";
print "Status: 403 Access denied\r\n";
:
```

### Discussion

If Apache encounters an error processing a document, such as not being able to locate a file, by default it will return a canned error response to the client. You can customize this error response with the *ErrorDocument* directive, and Apache will generally maintain the error status when it sends your custom error text to the client.

However, if you want to change the status to something else, such as hiding the fact that a file doesn't exist by returning a Forbidden status, you need to tell Apache about the change.

This requires that the *ErrorDocument* be a dynamic page, such as a CGI script. The CGI specification provides a very simple means of specifying the status code for a response: the `Status:` CGI header field. The Solution shows how it can be used.

### See Also

- 

- [Chapter 8](#)

- 

- <http://httpd.apache.org/docs/mod/core.html#errordocument>

- 

- <http://CGI-Spec.Golux.Com/>



## Recipe 9.3 Customized Error Messages

### Problem

You want to display a customized error message, rather than the default Apache error page.

### Solution

Use the `ErrorDocument` directive in *httpd.conf*:

```
ErrorDocument 405 /errors/notallowed.html
```

### Discussion

The *ErrorDocument* directive allows you to create your own error pages to be displayed when particular error conditions occur. In the previous example, in the event of a 405 status code (Method Not Allowed), the specified URL is displayed for the user, rather than the default Apache error page.

The page can be customized to look like the rest of your web site. When an error document looks significantly different from the rest of the site, this can leave the user feeling disoriented, or she may feel as if she has left the site that she is currently on.

### See Also

- 

<http://httpd.apache.org/docs/mod/core.html#errordocument>

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## Recipe 9.4 Providing Error Documents in Multiple Languages

### Problem

On a multilingual (content negotiated) web site, you want your error documents to be content negotiated as well.

### Solution

The Apache 2.0 default configuration file contains a configuration section, initially commented out, that allows you to provide error documents in multiple languages customized to the look of your web site, with very little additional work.

Uncomment those lines. You can identify the lines by looking for the following comment in your default configuration file:

```
# The internationalized error documents require mod_alias, mod_include  
# and mod_negotiation. To activate them, uncomment the following 30 lines.
```

In Apache 1.3 this is harder, but there's a solution in the works, as of this writing, that will make it similar to the 2.0 implementation. Check the Apache Cookbook web site for more information.

### Discussion

The custom error documents provided with Apache 2.0 combine a variety of techniques to provide internationalized error messages. As of this writing, these error messages are available in German, English, Spanish, French, Dutch, Swedish, Italian and Portuguese. Based on the language preference set in the client browser, the error message is delivered in the preferred language of the end-user.

Using content negotiation, the correct variant of the document (i.e., the right language) is selected for the user, based on their browser preference settings. For more information about content negotiation, see the content negotiation documentation at <http://httpd.apache.org/docs-2.0/content-negotiation.html> (for Apache 2.0) or <http://httpd.apache.org/docs/content-negotiation.html> (for Apache 1.3).

In addition to delivering the error message in the correct language, this functionality also lets you customize the look of these error pages so that they resemble the rest of your web site. To facilitate this, the files *top.html* and *bottom.html*, located in the *include* subdirectory of the *error* directory, should be modified to look like the standard header and footer content that appears on your web site. The body of the error message documents is placed between the header and the footer to create a page that is less jarring to users when they transition from your main site to the error pages that are generated.

You will also note that the error documents themselves contain SSI directives, which are used to further customize the error documents for the user. For example, in the case of the 404 (file not found) error document, the page will provide a link back to the page that the user came from, if the environment variable `HTTP_REFERER` is defined, and if that variable is not found, the page will merely notify the user that the URL was not found. Other SSI directives may be put in these documents, if you wish, to further customize them.

### See Also



## Recipe 9.5 Redirecting Invalid URLs to Some Other Page

### Problem

You want all "not found" pages to go to some other page instead, such as the front page of the site, so that there is no loss of continuity on bad URLs.

### Solution

Use the `ErrorDocument` to catch 404 (Not Found) errors:

```
ErrorDocument 404 /index.html
DirectoryIndex index.html /path/to/notfound.html
```

### Discussion

The recipe given here will cause all 404 errors—every time someone requests an invalid URL—to return the URL `/index.html`, providing the user with the front page of your web site, so that even invalid URLs still get valid content. Presumably, users accessing an invalid URL on your web site will get a page that helps them find the information that they were looking for.

On the other hand, this behavior may confuse the user who believes she knows exactly where the URL should take her. Make sure that the page that you provide as the global error document does in fact help people find things on your site, and does not merely confuse or disorient them. You may, as shown in the example, return them to the front page of the site. From there they should be able to find what they were looking for.

When users get good content from bad URLs, they will never fix their bookmarks and will continue to use a bogus URL long after it has become invalid. You will continue to get 404 errors in your log file for these URLs, and the user will never be aware that they are using an invalid URL. If, on the other hand, you actually return an error document, they will immediately be aware that the URL they are using is invalid and will update their bookmarks to the new URL when they find it.

Note that, even though a valid document is being returned, a status code of 404 is still returned to the client. This means that if you are using some variety of tool to validate the links on your web site, you will still get good results, if the tool is checking the status code, rather than looking for error messages in the content.

### See Also

- <http://httpd.apache.org/docs/mod/core.html#errordocument>
- [http://httpd.apache.org/docs/mod/mod\\_dir.html](http://httpd.apache.org/docs/mod/mod_dir.html)



## Recipe 9.6 Making Internet Explorer Display Your Error Page

### Problem

You have an *ErrorDocument* correctly configured, but IE is displaying its own error page, rather than yours.

### Solution

Make the error document bigger—at least 512 bytes.

### Discussion

Yes, this seems a little bizarre, and it is. In this case, Internet Explorer thinks it knows better than the web site administrator. If the error document is smaller than 512 bytes, it will display its internal error message page, rather than your custom error page, whenever it receives a 400 or 500 series status code. This size is actually configurable in the browser, so this number may in fact vary from one client to another. "Friendly error messages" can also be turned off entirely in the browser preferences.

This can be extremely frustrating the first time you see it happen, because you just know you have it configured correctly and it seems to work in your other browsers. Furthermore, when some helpful person tells you that your error document just needs to be a little larger, it's natural to think that he is playing a little prank on you, because this seems a little too far-fetched.

But it's true. Make the page bigger. It needs to be at least 512 bytes, or IE will ignore it and gleefully display its own "friendly" error message instead.

Exactly what you fill this space with is unimportant. You can, for example, just bulk it up with comments. For example, repeating the following comment 6 times would be sufficient to push you over that minimum file size:

```
<!-- message-obscuring clients are an abomination  
and an insult to the user's intelligence -->
```

### See Also

- 

<http://httpd.apache.org/docs/mod/core.html#errordocument>



# Recipe 9.7 Notification on Error Conditions

## Problem

You want to receive email notification when there's an error condition on your server.

## Solution

Point the *ErrorDocument* directive to a CGI program that sends mail, rather than to a static document:

```
ErrorDocument 404 /cgi-bin/404.cgi
```

*404.cgi* looks like the following:

```
#!/usr/bin/perl
use Mail::Sendmail;
use strict;

my $message = qq~
Document not found: $ENV{REQUEST_URI}
Link was from: $ENV{HTTP_REFERER}
~;

my %mail = (
    To => 'admin@server.com',
    From => 'website@server.com',
    Subject => 'Broken link',
    Message => $message,
);
sendmail(%mail);

print "Content-type: text/plain\n\n";
print "Document not found. Admin has been notified";
```

## Discussion

This recipe is provided as an example, rather than as a recommendation. On a web site of any significant size or traffic level, actually putting this into practice generates a substantial quantity of email, even on a site that is very well maintained. This is because people mistype URLs, and other sites, over which you have no control, will contain incorrect links to your site. It may be educational, however, to put something like this in place, at least briefly, to gain an appreciation for the scale of your own web site.

The *ErrorDocument* directive will cause all 404 (Document Not Found) requests to be handled by the specified URL, and so your CGI program gets run and is passed environment variables that will be used in the script itself to figure out what link is bad and where the request came from.

The script used the *Mail::Sendmail* Perl module to deliver the email message, and this module should work fine on any operating system. The module is not a standard part of Perl, so you may have to install it from CPAN (<http://www.cpan.org/>). A similar effect can, of course, also be achieved in PHP or any other programming language.

The last two lines of the program display a very terse page for the user, telling him that there was an error condition. You may wish, instead, to have the script redirect the user to some more informative and attractive page elsewhere on your web site. This could be accomplished by replacing those last two lines with something like the following:

```
print "Location: http://server.name/errorpage.html\n\n";
```



# Chapter 10. Proxies

Proxy means to act on behalf of another. In the context of a web server, this means one server fetching content from another server, then returning it to the client. For example, you may have several web servers that hide behind a proxy server. The proxy server is responsible for having requests end up going to the right backend server.

*mod\_proxy*, which comes with Apache, handles proxying behavior. The recipes in this chapter cover various techniques that can be used to take advantage of this capability. We discuss securing your proxy server, caching content proxied through your server, and ways to use *mod\_proxy* to map requests to services running on alternate ports.

Additional information about *mod\_proxy* can be found at [http://httpd.apache.org/docs/mod/mod\\_proxy.html](http://httpd.apache.org/docs/mod/mod_proxy.html) for Apache 1.3, or [http://httpd.apache.org/docs-2.0/mod/mod\\_proxy.html](http://httpd.apache.org/docs-2.0/mod/mod_proxy.html) for Apache 2.0.

Please make sure you don't enable proxying until you understand the security concerns involved and have taken steps to secure your proxy server. (See [Recipe 6.20](#) for details.)

You may also wish to consider a dedicated proxy server, such as Squid (<http://www.squid-cache.org/>), which is focused entirely on one task, and thus has more options related to this task.



# Recipe 10.1 Securing Your Proxy Server

## Problem

You want to enable proxying, but you don't want an open proxy that can be used by just anyone at all.

## Solution

For Apache 1.3:

```
<Directory proxy:*>
  Order deny,allow
  Deny from all
  Allow from .yourdomain.com
</Directory>
```

For Apache 2.0:

```
<Proxy *>
  Order Deny,Allow
  Deny from all
  Allow from .yourdomain.com
</Proxy>
```

## Discussion

Running an open proxy is a concern because it permits users from the Internet to use your proxy server to cover their tracks as they visit web sites. This can be a problem for a variety of reasons. The user is effectively stealing your bandwidth and is certainly part of the problem. However, perhaps more concerning is the fact that you are probably enabling people to circumvent restrictions that have been put in place by their network administrators, or perhaps you are providing users with anonymity while they visit a web site, and as a consequence, these visits appear to come from your network.

In these recipes, *.yourdomain.com* should be replaced by the name of your particular domain, or, better yet, the network address(es) that are on your network. (IP addresses are harder to fake than host and domain names.) For example, you might use, rather than the line appearing in the recipe, a line such as:

```
Allow from 192.168.1
```

Note that every request for resources that goes through your proxy server generates a logfile entry, containing the address of the client and the resource that they requested through your proxy server. For example, one such request might look like:

```
192.168.1.5 - - [26/Feb/2003:21:26:13 -0500] "GET http://httpd.apache.org/docs/mod/
mod_proxy.html HTTP/1.1" 200 49890
```

Your users, if made aware of this fact, will no doubt find it invasive, because this will show all HTTP traffic through the proxy server.

It is possible to configure your server not to log these requests. The technique for doing this is to set an environment variable for proxied requests:

```
<Directory proxy:*>
  SetEnv PROXIED 1
</Directory>
```



## Recipe 10.2 Preventing Your Proxy Server from Being Used as an Open Mail Relay

### Problem

If your Apache server is set up to operate as a proxy, it is possible for it to be used as a mail relay unless precautions are taken. This means that your system may be functioning as an "open relay" even though your mail server software is actually securely configured.

### Solution

Use *mod\_rewrite* to forbid proxy requests to port 25 (SMTP):

```
<Directory proxy:*>
  RewriteEngine On
  RewriteRule "^proxy:[a-z]*://[^/]*:25(/|$)" "-" [F,NC,L]
</Directory>
```

### Discussion

To use the Apache proxy as an SMTP relay is fairly trivial, but then so is preventing it. The solution simply tells the server to respond with a 403 Forbidden to any attempts to use it to proxy to a remote mail server (port 25). Other ports, such as HTTP (port 80), HTTPS (port 443), and FTP (ports 20 and 21), which are commonly permitted proxy access, will not be affected.

### See Also

- [http://httpd.apache.org/docs/mod/mod\\_proxy.html](http://httpd.apache.org/docs/mod/mod_proxy.html)
- <http://httpd.apache.org/docs/mod/core.html#directory>
- [http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

## Recipe 10.3 Forwarding Requests to Another Server

### Problem

You want requests for particular URLs to be transparently forwarded to another server.

### Solution

Use `ProxyPass` and `ProxyPassReverse` directives in your `httpd.conf`:

```
ProxyPass /other/ http://other.server.com/  
ProxyPassReverse /other/ http://other.server.com/
```

### Discussion

Use this recipe when you have a frontend server and one or more backend servers, inaccessible from the Internet, and you wish to serve content from them. In the example given, when a request is made for a URL starting with `/other/`, Apache makes a request for the URL `http://other.server.com/`, and returns the content obtained by the client. For example, a request for the URL `/other/example.html` results in a request for the URL `http://other.server.com/example.html`.

The `ProxyPassReverse` directive ensures that any header fields returned by the secondary server (which contain the name of the server, such as Location headers) will be rewritten to contain the URL that the end user will actually be using, ensuring that the redirect actually functions as desired.

Note that links within HTML documents on the secondary site should all be relative, rather than absolute, so that these links work for users using the content via the proxy server. In the recipe given, for example, a link to `/index.html` removes the `/other/` portion of the URL, causing the request to no longer hit the proxied portion of the server.

Using this technique, you can have content for one web site actually served by multiple web server machines. This can be used as a means to traverse the border of your network, or it can be used as a load-sharing technique to lessen the burden on your primary web server.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_proxy.html](http://httpd.apache.org/docs/mod/mod_proxy.html)

## Recipe 10.4 Blocking Proxied Requests to Certain Places

### Problem

You want to use your proxy server as a content filter, forbidding requests to certain places.

### Solution

Use ProxyBlock in the *httpd.conf* to deny access to particular sites:

```
ProxyBlock forbiddensite.com www.competitor.com monster.com
```

### Discussion

This example forbids proxied requests to the sites listed. These arguments are substring matches; `example.com` will also match `www.example.com`, and an argument of `example` would match both.

If you want more fine-grained control of what content is requested through your proxy server, you may want to use something more sophisticated, such as Squid, which is more full-featured in that area.

### See Also

- 

The Squid proxy server, found at <http://www.squid-cache.org/>

## Recipe 10.5 Proxying `mod_perl` Content to Another Server

### Problem

You want to run a second HTTPD server for dynamically generated content and have Apache transparently map requests for this content to the other server.

### Solution

First, install Apache, running on an alternate port, such as port 90, on which you will generate this dynamic content. Then, on your main server:

```
ProxyPass /dynamic/ http://localhost:90/  
ProxyPassReverse /dynamic/ http://localhost:90/
```

### Discussion

Most dynamic content generation techniques use a great deal more system resources than serving static content. This can slow down the process of serving static content from the same server, because child processes will be consumed with producing this dynamic content, and thus unable to serve the static files.

By giving the dynamic content its own dedicated server, you allow the static content to be served much more rapidly, and the dynamic content has a dedicated server. Each of the servers can have a smaller set of modules installed than they would otherwise require, because they will be performing a smaller subset of the functionality needed to do both tasks.

This technique can be used for a `mod_perl` server, a PHP server, or any other dynamic content method. Or you could reverse the technique and have, for example, a dedicated machine for serving image files using `mod_mmap_static` to serve the files very rapidly out of an in-memory cache.

In the example given, all URLs starting with `/dynamic/` will be forwarded on to the other server, which will, presumably, handle only requests for dynamic content. URLs that do not match this URL, however, will fall through and be handled by the frontend server.

## Recipe 10.6 Configuring a Caching Proxy Server

### Problem

You want to run a caching proxy server.

### Solution

Configure your server to proxy requests, and provide a location for the cached files to be placed:

```
ProxyRequests on
CacheRoot /var/spool/httpd/proxy
```

### Discussion

Running a caching proxy server allows users on your network to have more rapid access to content that others have already requested. They will, perhaps, not be getting the most recent version of the document in question, but, because they are retrieving the content from a local copy rather than from the remote web server, they will get it much more quickly.

With the contents of the WWW growing ever more dynamic, running a caching proxy server perhaps makes less sense than it once did, when most of the Web was composed of static content. However, since *mod\_proxy* is fairly smart about what it caches and what it does not cache, this sort of setup will still speed things up by caching the static portions of documents, such as the image files, while retrieving the freshest version of those documents that change over time.

The directory specified in the *CacheRoot* directive specifies where cached content will be stored. This directory must be writable by the user that Apache is running as (typically nobody), so that it is able to store these files there.

Finally, note that, while in Apache 1.3, the functions discussed here are provided by *mod\_proxy*; in Apache 2.0, the proxying and caching functionality have been split into the modules *mod\_proxy* and *mod\_cache*, respectively. In either case, these modules are not enabled by default.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_proxy.html](http://httpd.apache.org/docs/mod/mod_proxy.html)

## Recipe 10.7 Filtering Proxied Content

### Problem

You want to apply some filter to proxied content, such as altering certain words.

### Solution

In Apache 2.0 and later, you can use *mod\_ext\_filter* to create output filters to apply to content before it is sent to the user:

```
ExtFilterDefine naughtywords mode=output intype=text/html  
    cmd="/bin/sed s/darned/blasted/g"
```

```
<Proxy *>  
    SetOutputFilter naughtywords  
</Proxy>
```

### Discussion

The recipe offered is a very simple-minded "naughty word" filter, replacing the naughty word "darned" with the sanitized alternate "blasted." This could be expanded to a variety of more sophisticated content modification, because the *cmd* argument can be any command line, such as a Perl script, or arbitrary program, which can filter the content in any way you want. All proxied content will be passed through this filter before it is delivered to the client.

Note that this recipe will work only in Apache 2.0, as the module *mod\_ext\_filter*, the *SetOutputFilter* directive, and the *<Proxy>* directive are available only in Apache 2.0.

Note also that there are ethical and legal issues surrounding techniques like this, which you may need to deal with. We don't presume to take a position on any of them. In particular, modifying proxied content that does not belong to you may be a violation of the owner's copyright and may be considered by some to be unethical. Thankfully, this is just a technical book, not a philosophical one. We can tell you how to do it, but whether you should is left to your conscience and your lawyers.

### See Also

- 

[http://httpd.apache.org/docs-2.0/mod/mod\\_proxy.html](http://httpd.apache.org/docs-2.0/mod/mod_proxy.html)

- 

[http://httpd.apache.org/docs-2.0/mod/mod\\_ext\\_filter.html](http://httpd.apache.org/docs-2.0/mod/mod_ext_filter.html)

## Recipe 10.8 Requiring Authentication for a Proxied Server

### Problem

You wish to proxy content from a server, but it requires a login and password before content may be served from this proxied site.

### Solution

Use standard authentication techniques to require logins for proxied content:

```
ProxyPass "/secretserver/" "http://127.0.0.1:8080 "  
<Directory "proxy:http://127.0.0.1:8080/">  
    AuthName SecretServer  
    AuthType Basic  
    AuthUserFile /path/to/secretserver.htpasswd  
    Require valid-user  
</Directory>
```

### Discussion

This technique can be useful if you are running some sort of special-purpose or limited-function web server on your system, but you want to apply Apache's rich set of access control and its other features to access it. This is done by using the *ProxyPass* directive to make the special-purpose server's URI space part of your main server, and using the special *proxy:path* *<Directory>* container syntax to apply Apache settings only to the mapped URIs.

### See Also

- 

[Recipe 6.7](#)

# Chapter 11. Performance

Your web site can probably be made to run faster, if you are willing to make a few tradeoffs, and spend a little time benchmarking your site to see what is really slowing it down.

There are a number of things that you can configure differently to get a performance boost. Although, there are other things to which you may have to make more substantial changes. It all depends on what you can afford to give up and what you are willing to trade off. For example, in many cases, you may need to trade performance for security, or vice versa.

In this chapter, we make some recommendations of things that you can change, and we warn against things that can cause substantial slow-downs. Be aware that web sites are very individual, and what may speed up one web site may not necessarily speed up another web site.

Topics covered include hardware considerations, configuration file changes, and dynamic content generation, which can all be factors in getting every ounce of performance out of your web site.



Very frequently, application developers develop programs in conditions that don't nearly enough reflect the conditions under which they will be run in production. Consequently, the application that seemed to run adequately fast with the test database of 100 records, runs painfully slowly with the production database of 200,000 records.

By ensuring that your test environment is at least as demanding as your production environment, you greatly reduce the chances that your application will perform unexpectedly slowly when you roll it out.



# Recipe 11.1 Determining How Much Memory You Need

## Problem

You want to ensure that you have sufficient RAM in your server.

## Solution

Find the instances of Apache in your process list, and determine an average memory footprint for an Apache process. Multiply this number by your peak load (maximum number of concurrent web clients you'll be serving).

## Discussion

Because there is very little else that you can do at the hardware level to make your server faster, short of purchasing faster hardware, it is important to make sure that you have as much RAM as you need.

Determining how much memory you need is an inexact science, to say the least. In order to take an educated guess, you need to observe your server under load, and see how much memory it is using.

The amount of memory used by one Apache process will vary greatly from one server to another, based on what modules you have installed and what the server is being called upon to do. Only by looking at your own server can you get an accurate estimate of what this quantity is for your particular situation.

Tools such as *top* and *ps* may be used to examine your process list and determine the size of processes. The *server-status* handler, provided by *mod\_status*, may be used to determine the total number of Apache processes running at a given time.

If, for example, you determine that your Apache processes are using 4 MB of memory each, and under peak load, you find that you are running 125 Apache processes, then you will need, at a bare minimum, 500 MB of RAM in the server to handle this peak load. Remember that memory is also needed for the operating system, and any other applications and services that are running on the system, in addition to Apache, and so in reality you will need more than this amount.

If, on the other hand, you are unable to add more memory to the server, for whatever reason, you can use the same technique to figure out the maximum number of child processes that you are capable of serving at any one time, and use the *MaxClients* directive to limit Apache to that many processes:

```
MaxClients 125
```

## See Also

- 

<http://httpd.apache.org/docs/misc/perf-tuning.html>



# Recipe 11.2 Benchmarking Apache with ab

## Problem

You want to benchmark changes that you are making to verify that they are in fact making a difference in performance.

## Solution

Use *ab* (Apache bench), which you will find in the *bin* directory of your Apache installation:

```
ab -n 1000 -c 10 http://www.example.com/test.html
```

## Discussion

*ab* is a command-line utility that comes with Apache and lets you do very basic performance testing of your server. It is especially useful for making small changes to your configuration and testing server performance before and after the change.

The arguments given in the previous example tell *ab* to request the resource `http://servername.com/test.html` 1000 times (*-n 1000* indicates the number of requests) and to make these requests 10 at a time (*-c 10* indicates the concurrency level).

Other arguments that may be specified can be seen by running *ab* with the *-h* flag. Of particular interest is the *-k* flag, which enables keepalive mode. See the following keepalive recipe for additional details on this matter.

There are a few things to note about *ab* when using it to evaluate performance.

*ab* does not mimic web site usage by real people. It requests the same resource repeatedly to test the performance of that one thing. For example, you may use *ab* to test the performance of a particular CGI program, before and after a performance-related change was made to it. Or you may use it to measure the impact of turning on *.htaccess* files, or content negotiation, for a particular directory. Real users, of course, do not repeatedly load the same page, and so performance measurements made using *ab* may not reflect actual real-world performance of your web site.

You should probably not run the web server and *ab* on the same machine, as this will introduce more uncertainty into the measurement. With both *ab* and the web server itself consuming system resources, you will receive significantly slower performance than if you were to run *ab* on some other machine, accessing the server over the network. However, also be aware that running *ab* on another machine will introduce network latency, which is not present when running it on the same machine as the server.

Finally, there are many factors that can affect performance of the server, and you will not get the same numbers each time you run the test. Network conditions, other processes running on the client or server machine, and a variety of other things, may influence your results slightly one way or another. The best way to reduce the impact of environmental changes is to run a large number of tests and average your results. Also, make sure that you change as few things as possible—ideally, just one—between tests, so that you can be more sure what change has made any differences you can see.



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 11.3 Tuning Keepalive Settings

## Problem

You want to tune the keepalive-related directives to the best possible setting for your web site.

## Solution

Turn on the *KeepAlive* setting, and set the related directives to sensible values:

```
KeepAlive On
MaxKeepAliveRequests 0
KeepAliveTimeout 15
```

## Discussion

The default behavior of HTTP is for each document to be requested over a new connection. This causes a lot of time to be spent opening and closing connections. *KeepAlive* allows multiple requests to be made over a single connection, thus reducing the time spent establishing socket connections. This, in turn, speeds up the load time for clients requesting content from your site.

In addition to turning keepalive on, using the *KeepAlive* directive, there are two directives that allow you to adjust the way that it is done.

The first of these, *MaxKeepAliveRequests*, indicates how many keepalive requests should be permitted over a single connection. There is no reason to have this number set low. The default value for this directive is 100, and this seems to work pretty well for most sites. Setting this value to 0 means that an unlimited number of requests will be permitted over a single connection. This might allow users to load all of their content from your site over a single connection, depending on the value of *KeepAliveTimeout* and how quickly they went through the site.

*KeepAliveTimeout* indicates how long a particular connection will be held open when no further requests are received. The optimal setting for this directive depends entirely on the nature of your web site. You should probably think of this value as the amount of time it takes users to absorb the content of one page of your site before they move on to the next page. If the users move on to the next page before the *KeepAliveTimeout* has expired, when they click on the link for the next page of content, they will get that next document over the same connection. If, however, that time has already expired, they will need to establish a new connection to the server for that next page.

You should also be aware that if users load a resource from your site and then go away, Apache will still maintain that open connection for them for *KeepAliveTimeout* seconds, which makes that child process unable to serve any other requests during that time. Therefore, setting *KeepAliveTimeout* too high is just as undesirable as setting it too low.

In the event that *KeepAliveTimeout* is set too high, you will see (i.e., with the *server-status* handler—see [Recipe 11.4](#)) that a significant number of processes are in keepalive mode, but are inactive. Over time, this number will continue to grow, as more child processes are spawned to take the place of child processes that are in this state.

Conversely, setting *KeepAliveTimeout* too low will result in conditions similar to having *KeepAlive* turned off entirely: when a single client will require many connections over the course of a brief visit. This is less easy to detect



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 11.4 Getting a Snapshot of Your Site's Activity

## Problem

You want to find out exactly what your server is doing.

## Solution

Enable the *server-status* handler to get a snapshot of what child processes are running and what each one is doing. Enable *ExtendedStatus* to get even more detail:

```
<Location /server-status>  
    SetHandler server-status  
</Location>
```

```
ExtendedStatus On
```

Then, view the results at the URL <http://servername/server-status>.

## Discussion

Provided by *mod\_status*, which is enabled by default, the *server-status* handler provides a snapshot of your server's activity. This snapshot includes some basic details, such as when the server was last restarted, how long it has been up, and how much data it has served in that time. Following that, there will be a list of the child processes and what each one is doing. At the bottom of the page is a detailed explanation of the terms used and what each column of the table represents.



The server status display shows activity across the entire server—including virtual hosts. If you are providing hosting services for others, you may not want them to be able to see this level of detail about each other.

It is recommended that, as in the default configuration file that comes with Apache, you restrict access to this handler. Part of the information contained on this page is a list of client addresses and the document that they are requesting. Some users feel that it is a violation of their privacy for you to make this information readily available on your web site. Additionally, it may provide information such as QUERY\_STRING variables, PATH\_INFO variables, or simply URLs, which you wished to not be made public. It is therefore recommended that you add to the above recipe some lines such as:

```
Order deny,allow  
Deny from all  
Allow from 192.168.1
```

This configuration allows access only from the 192.168.0 network, or whatever network you put in there, and denies access from unauthorized Internet users.

## See Also

-



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 11.5 Avoiding DNS Lookups

## Problem

You want to avoid situations where you have to do DNS lookups of client addresses, as this is a very slow process.

## Solution

Always set the *HostNameLookups* directive to *Off*:

```
HostNameLookups Off
```

And make sure that, whenever possible, *Allow from* and/or *Deny from* directives use the IP address, rather than the hostname of the hosts in question.

## Discussion

DNS lookups can take a very long time and should be avoided at all costs. In the event that a client address cannot be looked up at all, it can take up to a minute for the lookup to time out, during which time the child process that is doing the lookup cannot do anything else.

There are a number of cases in which Apache will need to do DNS lookups, and so the goal here is to completely avoid those situations.

### HostNameLookups

Prior to Apache 1.3, *HostNameLookups*, which determines whether Apache logs client IP addresses or hostnames, defaulted to *on*, meaning that each Apache log entry required a DNS lookup to convert the client IP address to a hostname to put in the logfile. Fortunately, that directive now defaults to *off*, and so this is primarily an admonition to leave it alone.

If you need to have these addresses converted to hostnames, then this should be done by another program, preferably running on a machine other than your production web server. That is, you really should copy the file to some other machine for the purpose of processing, so that the effort required to do this processing does not negatively effect your web server's performance.

Apache comes with a utility called *logresolve*, which will process your logfile, replacing IP addresses with hostnames. Additionally, most logfile analysis tools will also do this name resolution as part of the log analysis process.

### Allow and Deny from hostnames

When you do host-based access control, using the *Allow from* and *Deny from* directives, Apache takes additional precautions to make sure that the client is not spoofing its hostname. In particular, it does a DNS lookup on the IP address of the client to obtain the name to compare against the access restriction. It then looks up the name that was obtained, just to make sure that the DNS record is not being faked. [\[1\]](#)

[1] For example, the owner of the IP address could very easily put a PTR record in their reverse-DNS zone,



[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

# Recipe 11.6 Optimizing Symbolic Links

## Problem

You wish to balance the security needs associated with symbolic links with the performance impact of a solution, such as using *Options SymLinksIfOwnerMatch*, which causes a server slowdown.

## Solution

For tightest security, use *Options SymlinksIfOwnerMatch*, or *Options -FollowSymLinks* if you seldom or never use symlinks.

For best performance, use *Options FollowSymlinks*.

## Discussion

Symbolic links are an area in which you need to weigh performance against security and make the decision that makes the most sense in your particular situation.

In the normal everyday operation of a Unixish operating system, symbolic links are considered to be the same as the file to which they link.<sup>[2]</sup> When you *cd* into a directory, you don't need to be aware of whether that was a symlink or not. It just works.

[2] Of course, this is not true at the filesystem level, but we're just talking about the practical user level.

Apache, on the other hand, has to consider whether each file and directory is a symlink or not, if the server is configured not to follow symlinks. And, additionally, if *Option SymlinksIfOwnerMatch* is turned on, Apache not only has to check if the particular file is a symlink, but also has to check the ownership of the link itself and of the target, in the event that it is a symlink. While this enforces a certain security policy, it takes a substantial amount of time and so slows down the operation of your server.

In the tradeoff between security and performance, in the matter of symbolic links, here are the guidelines.

If you are primarily concerned about security, never permit the following of symbolic links. It may permit someone to create a link from a document directory to content that you would not want to be on a public server. Or, if there are cases where you really need symlinks, use *Options SymlinksIfOwnerMatch*, which requires that someone may only link to files that they own and will presumably protect you from having a user link to a portion of the filesystem that is not already under their control.

If you are concerned about performance, then always use *Options FollowSymlinks*, and never use *Options SymlinksIfOwnerMatch*. *Options FollowSymlinks* permits Apache to follow symbolic links in the manner of most Unixish applications—that is, Apache does not even need to check to see if the file in question is a symlink or not.

## See Also





# Recipe 11.7 Minimizing the Performance Impact of .htaccess Files

## Problem

You want per-directory configuration but want to avoid the performance hit of *.htaccess* files.

## Solution

Turn on *AllowOverride* only in directories where it is required, and tell Apache not to waste time looking for *.htaccess* files elsewhere:

```
AllowOverride None
```

Then use *<Directory>* sections to selectively enable *.htaccess* files only where needed.

## Discussion

*.htaccess* files cause a substantial reduction in Apache's performance, because it must check for a *.htaccess* in every directory along the path to the requested file to be assured of getting all of the relevant configuration overrides. This is necessary because Apache configuration directives apply not only to the directory in which they are set, but also to all subdirectories. Thus, we must check for *.htaccess* files in parent directories, as well as in the current directory, to find any directives that would trickle down the current directory.

For example, if, for some reason, you had *AllowOverride All* enabled for all directories and your *DocumentRoot* was */usr/local/apache/htdocs*, then a request for the URL *http://example.com/events/parties/christmas.html* would result in the following files being looked for and, if found, opened and searched for configuration directives:

```
/.htaccess
/usr/.htaccess
/usr/local/.htaccess
/usr/local/apache/.htaccess
/usr/local/apache/htdocs/.htaccess
/usr/local/apache/htdocs/events/.htaccess
/usr/local/apache/htdocs/events/parties/.htaccess
```

Now, hopefully, you would never have *AllowOverride All* enabled for your entire filesystem, so this is a worst-case scenario. However, occasionally, when people do not adequately understand what this configuration directive does, they will enable this option for their entire filesystem and suffer poor performance as a result.

The recommended solution is by far the best way to solve this problem. The *<Directory>* directive is specifically for this situation, and *.htaccess* files should really only be used in the situation where configuration changes are needed and access to the main server configuration file is not readily available.

For example, if you have a *.htaccess* file in */usr/local/apache/htdocs/events* containing the directive:

```
AddEncoding x-gzip tgz
```

You should instead simply replace this with the following in your main configuration file:

```
<Directory /usr/local/apache/htdocs/event>
    AddEncoding x-gzip tgz
</Directory>
```





# Recipe 11.8 Disabling Content Negotiation

## Problem

Content negotiation causes a big reduction in performance.

## Solution

Disable content negotiation where it is not needed. If you do require content negotiation, use the *type-map* handler, rather than the *MultiViews* option:

```
Options -MultiViews
```

```
AddHandler type-map var
```

## Discussion

If at all possible, disable content negotiation. However, if you must do content negotiation—if, for example, you have a multilingual web site—you should use the *type-map* handler, rather than the *MultiViews* method.

When *MultiViews* is used, Apache needs to get a directory listing each time a request is made. The resource requested is compared to the directory listing to see what variants of that resource might exist. For example, if *index.html* is requested, the variants *index.html.en* and *index.html.fr* might exist to satisfy that request. Each matching variant is compared with the user's preferences, expressed in the various Accept headers passed by the client. This information allows Apache to determine which resource is best suited to the user's needs.

However, this process can be very time-consuming, particularly for large directories or resources with large numbers of variants. By putting the information in a *.var* file and allowing the *type-map* handler to be used instead, you eliminate the requirement to get a directory listing, and greatly reduce the amount of work that Apache must do to determine the correct variant to send to the user.

The *.var* file just needs to contain a listing of the variants of a particular resource and describe their important attributes.

If you have, for example, English, French, and Hebrew variants of the resource *index.html*, you may express this in a *.var* file called *index.html.var* containing information about each of the various variants. This file might look like the following:

```
URI: index.html.en
Content-language: en
Content-type: text/html
```

```
URI: index.html.fr
Content-language: en
Content-type: text/html
```

```
URI: index.html.he.iso8859-8
Content-language: he
Content-type: text/html; charset=ISO-8859-8
```

This file should be placed in the same directory as the variants of this resource, which are called *index.html.en*, *index.html.fr*, and *index.html.he.iso8859-8*.





# Recipe 11.9 Optimizing Process Creation

## Problem

You're using Apache 1.3, or Apache 2.0 with the *prefork* MPM, and you want to tune *MinSpareServers* and *MaxSpareServers* to the best settings for your web site.

## Solution

Will vary from one site to another. You'll need to watch traffic on your site and decide accordingly.

## Discussion

The *MinSpareServers* and *MaxSpareServers* directives control the size of the server pool, so that incoming requests will always have a child process waiting to serve them. In particular, if there are fewer than *MinSpareServers* idle processes, Apache will create more processes until that minimum is reached. Similarly, if there are ever more than *MaxSpareServers* processes, Apache will kill off processes until there are fewer than that maximum. These things will happen as the site traffic fluctuates on a normal day.

The best values for these directives for your particular site depends on the amount and the rate at which traffic fluctuates. If your site is prone to large spikes in traffic, *MinSpareServers* needs to be large enough to absorb those spikes. The idea is to never have a situation where requests come in to your site, and there are no idle server processes waiting to handle the request. If traffic patterns on your site are fairly smooth curves with no abrupt spikes, the default values may be sufficient.

The best way to watch exactly how much load there is on your server is by looking at the *server-status* handler output. (See [Recipe 11.4](#).)

You should also set *MaxClients* to a value such that you don't run out of server resources during heavy server loads. For example, if your average Apache process consumes 2 MB of memory and you have a total of 256 MB of RAM available, allowing a little bit of memory for other processes, you probably don't want to set *MaxClients* any higher than about 120. If you run out of RAM and start using swap space, your server performance will abruptly go downhill and will not recover until you are no longer using swap. You can watch memory usage by running a program such as *top*, which shows running processes and how much memory each is using.

## See Also

- [Setting the number of threads on single-child MPMs](#) in [Recipe 11.10](#)
- [Number of threads when using the worker MPM](#) in [Recipe 11.10](#)



# Recipe 11.10 Tuning Thread Creation

## Problem

You're using Apache 2.0 with one of the threaded MPMs, and you want to optimize the settings for the number of threads.

## Solution

Will vary from server to server.

## Discussion

The various threaded MPMs on Apache 2.0 handle thread creation somewhat differently. In Apache 1.3, the Windows and Netware versions are threaded, while the Unixish version is not. Tuning the thread creation values will vary from one of these versions to another.

### Setting the number of threads on single-child MPMs

On MPMs that run Apache with a single threaded child process, such as the Windows MPM (*mpm\_winnt*), and the Windows and Netware versions of Apache 1.3, there are a fixed number of threads in the child process. This number is controlled by the *ThreadsPerChild* directive and must be large enough to handle the peak traffic of the site on any given day. There really is no performance tuning that can be done here, as this number is fixed throughout the lifetime of the Apache process.

### Number of threads when using the worker MPM

The *worker* MPM has a fixed number of threads per child process but has a variable number of child processes, so that increased server load can be absorbed. A typical configuration might look like the following:

```
StartServers 2
MaxClients 150
MinSpareThreads 25
MaxSpareThreads 75
ThreadsPerChild 25
ServerLimit 16
```

The *MinSpareThreads* and *MaxSpareThreads* directives control the size of the idle pool of threads, so that incoming clients will always have an idle thread waiting to serve their request. The *ThreadsPerChild* directive indicates how many threads are in each child process, so when the number of available idle threads drops below *MinSpareThreads*, Apache will launch a new child process, populated with *ThreadsPerChild* threads. Similarly, when server load is reduced and the number of idle threads is greater than *MaxSpareThreads*, Apache will kill off one or more child processes to reduce the idle pool to that number or less.

The goal, when setting these values, is to ensure that there are always idle threads ready to serve any incoming client's request, without having to create a new one. The previous example will work for most sites, as it will ensure that there is at least one completely unused child process, populated with 25 threads, waiting for incoming requests. As soon as threads within this process start to be used, a new child process will be launched for future requests.





# Recipe 11.11 Caching Frequently Viewed Files

## Problem

You want to cache files that are viewed frequently, such as your site's front page, so that they don't have to be loaded from the filesystem every time.

## Solution

Use *mod\_mmap\_static* or *mod\_file\_cache* (for Apache 1.3 and 2.0, respectively) to cache these files in memory:

```
MMapFile /www/htdocs/index.html
MMapFile /www/htdocs/other_page.html
```

For Apache 2.0, you can use either module or the *CacheFile* directive. *MMapFile* caches the file contents in memory, while *CacheFile* caches the file handle instead, which gives slightly poorer performance but uses less memory:

```
CacheFile /www/htdocs/index.html
CacheFile /www/htdocs/other_page.html
```

## Discussion

For files that are frequently accessed, it is desirable to cache that file in some fashion to save disk access time. The *MMapFile* directive loads a file into RAM, and subsequent requests for that file are served directly out of RAM, rather than from the filesystem. The *CacheFile* directive, on the other hand, opens the file and caches the file handle, saving time on subsequent file opens.

In Apache 1.3, this functionality is available with the *mod\_mmap\_static* module, which is labelled as experimental and is not built into Apache by default. To enable this module, you need to specify the `—enable-module=mmap_static` flag to *configure* when building Apache. *mod\_mmap\_static* provides only the *MMapFile* directive.

In Apache 2.0, this functionality is provided by the *mod\_file\_cache* module, which is labelled as experimental, and is not built into Apache by default. To enable this module, you need to specify the `—enable-file_cache` flag to *configure* when building Apache. *mod\_file\_cache* provides both the *MMapFile* and *CacheFile* directives.

These directives take a single file as an argument, and there is not a provision for specifying a directory or set of directories. If you wish to have the entire contents of a directory mapped into memory, the documentation provides the following suggestion. For the directory in question, you would run the following command:

```
% find /www/htdocs -type f -print \  
> | sed -e 's/./mmapfile &/' > /www/conf/mmap.conf
```

In your main server configuration file, you would then load the file created by that command, using the *Include* directive:

```
Include /www/conf/mmap.conf
```

This would cause every file contained in that directory to have the *MMapFile* directive invoked on it.

Note that when files are cached using one of these two directives, any changes to the file will require a server restart



## Recipe 11.12 Sharing Load Between Servers Using `mod_proxy`

### Problem

You want to have a certain subset of your web site served from another machine, in order to share the load of the site.

### Solution

Use *ProxyPass* and *ProxyPassReverse* to have Apache fetch the content from another server:

```
ProxyPass          /other/ http://other.server.com/  
ProxyPassReverse  /other/ http://other.server.com/
```

### Discussion

These directives will cause requests to URLs starting with */other/* to be forwarded to the server `other.server.com`, with the path information preserved. That is to say, a request for `http://www.server.com/other/something.html` will be translated into a request for `http://other.server.com/something.html`. Content obtained from this other server will be returned to the client, which will be unable to determine that any such technique was employed. The *ProxyPassReverse* directive ensures that any redirect headers sent from the backend server (in this case, `other.server.com`) will be modified so that they appear to come from the main server.

This method is often used to have the dynamic portion of the site served by a server running *mod\_perl*—often even on the same machine, but on a different port—while the static portions of the site are served from the main server, which can be lighter weight, and so run faster.

Note that URLs contained within documents are not rewritten as they pass through the proxy, and links within documents should be relative, rather than absolute, so that they work correctly.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_proxy.html](http://httpd.apache.org/docs/mod/mod_proxy.html)

- 

[http://httpd.apache.org/docs-2.0/mod/mod\\_proxy.html](http://httpd.apache.org/docs-2.0/mod/mod_proxy.html)



# Recipe 11.13 Distributing Load Evenly Between Several Servers

## Problem

You want to serve the same content from several servers and have hits distributed evenly among the servers

## Solution

Use DNS round-robin to have requests distributed evenly, or at least fairly evenly, among the servers:

```
www.example.com. 86400 IN A 192.168.10.2
www.example.com. 86400 IN A 192.168.10.3
www.example.com. 86400 IN A 192.168.10.4
www.example.com. 86400 IN A 192.168.10.5
www.example.com. 86400 IN A 192.168.10.6
www.example.com. 86400 IN A 192.168.10.7
```

## Discussion

This example is an excerpt from a BIND zone file. The actual syntax may vary, depending on the particular name server software you are running.

By giving multiple addresses to the same hostname, you cause hits to be evenly distributed among the various servers listed. The name server, when asked for this particular name, will give out the addresses listed in a round-robin fashion, causing requests to be sent to one server after the other. The individual servers need be configured only to answer requests from the specified name.

Running the *host* command on the name in question will result in a list of possible answers, but each time you run the command, you'll get a different answer first:

```
% host www.example.com
www.example.com has address 192.168.10.2
www.example.com has address 192.168.10.3
www.example.com has address 192.168.10.4
www.example.com has address 192.168.10.5
www.example.com has address 192.168.10.6
www.example.com has address 192.168.10.7
% host www.example.com
www.example.com has address 192.168.10.7
www.example.com has address 192.168.10.2
www.example.com has address 192.168.10.3
www.example.com has address 192.168.10.4
www.example.com has address 192.168.10.5
www.example.com has address 192.168.10.6
```



Make sure that when you update your DNS zone file, you also update the serial number, and restart or reload your DNS server.

## See Also

-



## Recipe 11.14 Caching Directory Listings

### Problem

You want to provide a directory listing but want to reduce the performance hit of doing so.

### Solution

Use the *TrackModified* argument to *IndexOptions* to allow browsers to cache the results of an auto-generated directory index:

```
IndexOptions +TrackModified
```

### Discussion

When sending a directory listing to a client, Apache has to open that directory, obtain a directory listing, and determine various attributes of the files contained therein. This is very time consuming, and it would be nice to avoid this when possible.

By default, the Last Modified time sent with a directory listing is the time that the content is being served. Thus, when a client, or proxy server, makes a *HEAD* or conditional *GET* request to determine if it can use the copy that it has in cache, it will always decide to get a fresh copy of the content. The *TrackModified* option to *IndexOptions* cause *mod\_autoindex* to send a Last Modified time corresponding to the file in the directory that was most recently modified. This enables browsers and proxy servers to cache this content, rather than retrieving it from the server each time, and also ensures that the listing that they have cached is in fact the latest version.

Note that clients that don't implement any kind of caching will not benefit from this directive. In particular, testing with *ab* will show no improvement from turning on this setting, as *ab* does not do any kind of content caching.

### See Also

- 

The manpage for the *ab* tool



# Recipe 11.15 Speeding Up Perl CGI Programs with mod\_perl

## Problem

You have existing functional Perl CGI programs and want them to run faster.

## Solution

If you have the *mod\_perl* module installed, you can configure it to run your Perl CGI programs, instead of running *mod\_cgi*. This gives you a big performance boost, without having to modify your CGI code.

There are two slightly different ways to do this.

For Apache 1.3 and *mod\_perl* Version 1:

```
Alias /cgi-perl/ /usr/local/apache/cgi-bin/
<Location /cgi-perl>
    Options ExecCGI
    SetHandler perl-script
    PerlHandler Apache::PerlRun
    PerlSendHeader On
</Location>
```

```
Alias /perl/ /usr/local/apache/cgi-bin/
<Location /perl>
    Options ExecCGI
    SetHandler perl-script
    PerlHandler Apache::Registry
    PerlSendHeader On
</Location>
```

For Apache 2.0 and *mod\_perl* Version 2, the syntax changes slightly:

```
PerlModule ModPerl::PerlRun
Alias /cgi-perl/ /usr/local/apache2/cgi-bin/
<Location /cgi-perl>
    SetHandler perl-script
    PerlResponseHandler ModPerl::PerlRun
    Options +ExecCGI
</Location>
```

```
PerlModule ModPerl::Registry
Alias /perl/ /usr/local/apache2/cgi-bin/
<Location /perl>
    SetHandler perl-script
    PerlResponseHandler ModPerl::Registry
    Options +ExecCGI
</Location>
```

## Discussion

By using *mod\_perl*'s CGI modes, you can improve the performance of existing CGI programs without modifying the CGI code itself in any way. Given the previous configuration sections, a CGI program that was previously accessed via the URL `http://www.example.com/cgi-bin/program.cgi` will now be accessed via the URL `http://www.example.com/cgi-perl/program.cgi` to run it in *PerlRun* mode or via the URL `http://www.example.com/perl/program.cgi` to run it in *Registry* mode.



## Chapter 12. Miscellaneous Topics

With its hundreds of configuration directives, and dozens upon dozens of modules providing additional functionality, the Apache web server can be terrifically complex. So too can the questions about how to use it. We have collected many of the most common questions we have seen and categorized them, putting related topics into their own chapters when there were enough of them.

However, some of the things that come up don't fall readily into one of the categories we have chosen, or perhaps are more fundamental and we've collected them into this catch-all chapter of "things that don't belong anywhere else."



# Recipe 12.1 Placing Directives Properly

## Problem

You know what directive you need but aren't sure where to put it.

## Solution

If you wish the scope of the directive to be global (i.e., you want it to affect all requests to the web server), then it should be put in the main body of the configuration file or it should be put in the section starting with the line `<Directory />` and ending with `</Directory>`.

If you wish the directive to affect only a particular directory, it should be put in a `<Directory>` section that specifies that directory. Be aware that directives specified in this manner also affect subdirectories of the stated directory.

Likewise, if you wish the directive to affect a particular virtual host or a particular set of URLs, then the directive should be put in a `<VirtualHost>` section, `<Location>` section, or perhaps a `<Files>` section, referring to the particular scope in which you want the directive to apply.

In short, the answer to "Where should I put it?" is "Where do you want it to be in effect?"

## Discussion

This question is perhaps the most frequently asked question in every Apache help venue. It is usually answered in a way that is relevant to the specific situation but not in a general all-purpose kind of way.

The situation is further complicated by the fact that the configuration file is frequently split over several files, which are loaded via *Include* directives, and the (usually) mistaken impression that it will make a difference whether a directive is put in one file or another.

Knowing exactly where to put a particular directive comes from understanding how Apache deals with sections (such as `<Directory>` and `<Location>`). There is seldom one magic place that a directive must be placed to make it work. However, there are usually a number of places where you can put a directive and have it produce an undesired effect.

There are two main situations in which a directive, when added to your configuration file, will not have the desired effect. These are when a directive is overridden by a directive appearing in the same scope but later in the configuration, and when there is a directive in a more specific scope.

For the first of these two situations, it is important to understand that the Apache configuration file is parsed from top to bottom. Files that are *Include*'ed are considered to appear in their entirety in the location where the *Include* directive appears. Thus, if you have the same directive appearing twice but with different values, the last one appearing will be the one that is actually in effect.



## Recipe 12.2 Renaming .htaccess Files

### Problem

You want to change the default name of per-directory configuration files to something else, such as on a Windows system, because filenames beginning with a dot can cause problems.

### Solution

Use the `AccessFileName` directive to specify the new name:

```
AccessFileName ht.access
```

### Discussion

In addition to the server-wide configuration files, you can add directives to special files in individual directories. These are called *.htaccess* (aitch tee access) files because that's the default name for them.

However, the Unixish convention of filenames that begin with a dot doesn't play well on all platforms; on Windows in particular it can be difficult to edit files with such names.

Apache allows you to change the name it will use when looking for these per-directory files with the *AccessFileName* directive (which can only appear in the server-wide configuration files). You can use any name that's valid on your platform.

If you use the *AccessFileName* directive, be sure to make any additional appropriate changes to your configuration such as the `<FilesMatch ".ht">` container that keeps the files from being fetchable over the Web:

```
<FilesMatch "^ht\.">
  Order deny,allow
  Deny from all
</FilesMatch>
```

### See Also

- 

- [Recipe 11.7](#)

- 

- <http://httpd.apache.org/docs/howto/htaccess.html>

- 

- <http://httpd.apache.org/docs-2.0/howto/htaccess.html>



## Recipe 12.3 Generating Directory/Folder Listings

### Problem

You want to see a directory listing when a directory is requested.

### Solution

Turn on *Options Indexes* for the directory in question:

```
<Directory /www/htdocs/images>
  Options +Indexes
</Directory>
```

### Discussion

When a URL maps to a directory or folder in the filesystem, Apache will respond to the request in one of three ways:

- If *mod\_dir* is part of the server configuration, and the mapped directory is within the scope of a *DirectoryIndex* directive, and the server can find one of the files identified in that directive, then the file will be used to generate the response.
- If *mod\_autoindex* is part of the server configuration and the mapped directory is within the scope of an *Options* directive that has enabled the *Indexes* keyword, then the server will construct a directory listing at runtime and supply it as the response.
- The server will return a 404 (Resource Not Found) status.

### Enabling directory listings

The real keys to enabling the server's ability to automatically generate a listing of files in a directory are the inclusion of *mod\_autoindex* in the configuration and the *Indexes* keyword to the *Options* directive. This can be done either as an absolute form, as in:

```
Options FollowSymLinks Indexes
```

or in a selective or relative form such as:

```
Options -ExecCGI +Indexes
```

Enabling directory listings should be done with caution. Because of the scope inheritance mechanism, directories farther down the tree will also be affected; and because the server will apply the sequence of rules listed at the beginning of this section in an effort to provide some sort of response, a single missing file can result in the inadvertent exposure of your filesystem's contents.

### Disabling directory indexing below an enabled directory

There are essentially two ways to work around this issue and ensure that the indexing applies only to the single directory:





## Recipe 12.4 Solving the "Trailing Slash" Problem

### Problem

Loading a particular URL works with a trailing slash but does not work without it.

### Solution

Make sure that *ServerName* is set correctly and that none of the *Alias* directives have a trailing slash.

### Discussion

The "trailing slash" problem can be caused by one of two configuration problems: an incorrect or missing value of *ServerName*, or an *Alias* with a trailing slash that doesn't work without it.

#### Incorrect ServerName

An incorrect or missing *ServerName* seems to be the most prevalent cause of the problem, and it works something like this: when you request a URL such as `http://example.com/something`, where *something* is the name of a directory, Apache actually sends a redirect to the client telling it to add the trailing slash.

The way that it does this is to construct the URL using the value of *ServerName* and the requested URL. If *ServerName* is not set correctly, then the resultant URL, which is sent to the client, will generate an error on the client end when it can't find the resulting URL.

If, on the other hand, *ServerName* is not set at all, Apache will attempt to guess a reasonable value when you start it up. This will often lead it to guess incorrectly, using values such as `127.0.0.1` or `localhost`, which will not work for remote clients. Either way, the client will end up getting a URL that it cannot retrieve.

#### Invalid Alias directive

In the second incarnation of this problem, a slightly malformed *Alias* directive may cause a URL with a missing trailing slash to be an invalid URL entirely.

Consider, for example, the following directive:

```
Alias /example/ /home/www/example/
```

The *Alias* directive is very literal, and aliases URLs starting with `/example/`, but it does not alias URLs starting with `/example`. Thus, the URL `http://example.com/example/` will display the default document from the directory `/home/www/example/`, while the URL `http://example.com/example` will generate a "file not found" error message, with an error log entry that will look something like:

```
File does not exist: /usr/local/apache/htdocs/example
```

The solution to this is to create *Alias* directives without the trailing slash, so that they will work whether or not the trailing slash is used:

```
Alias /example /home/www/example
```



## Recipe 12.5 Setting the Content-Type According to Browser Capability

### Problem

You want to set Content-Type headers differently for different browsers, which may render the content incorrectly otherwise.

### Solution

Check the Accept headers with RewriteCond and then set the Content-Type header with a T flag:

```
RewriteCond %{HTTP_ACCEPT} "application/xhtml+xml"  
RewriteCond %{HTTP_ACCEPT} "!application/xhtml+xml\s*;\s*q=0+(?:\.0*[^0-9])"  
RewriteRule . - [T=application/xhtml+xml;charset=iso-8859-1]
```

### Discussion

Different browsers tend to deal with content differently and sometimes need a nudge in the right direction. In this example, for browsers that specify (using the HTTP\_ACCEPT header) that they prefer XHTML content, we want to send a Content-Type header specifying that the content we are sending fulfills that requirement.

The T (Type) flag sets the Content-Type for the response.

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_rewrite.html](http://httpd.apache.org/docs/mod/mod_rewrite.html)

## Recipe 12.6 Handling Missing Host: Header Fields

### Problem

You want to treat differently all requests that are made without a Host: request header field.

### Solution

```
SetEnvIf Host "^$" no_host=1
Order Allow,Deny
Allow from all
Deny from env=no_host
RewriteCond "%{HTTP_HOST}" "^$"
RewriteRule ".*" - [F,L]
```

### Discussion

The Host: request header field is essential to correct handling of name-based virtual hosts (see [Recipe 4.1](#)). If the client doesn't include it, the chances are very good that the request will be directed to the wrong virtual host. All modern browsers automatically include this field, so only custom-written or very old clients are likely to encounter this issue.

The solutions given will cause such requests to be rejected with a 403 Forbidden status; the exact text of the error page can be tailored with an ErrorDocument 403 directive.

The first solution is slightly more efficient.

### See Also

- 

[Recipe 4.1](#)

## Recipe 12.7 Alternate Default Document

### Problem

You want to have some file other than *index.html* appear by default.

### Solution

Use `DirectoryIndex` to specify the new name:

```
DirectoryIndex default.htm
```

### Discussion

When a directory is requested—that is, a URL ending in a `/` rather than in a file name—`mod_dir` will select the index document from that directory and serve that file in response. By default, the index file is assumed to be *index.html*, but this can be configured to something else with the `DirectoryIndex` directive.

Note also that `DirectoryIndex` can be set to several files, which are listed in order of precedence:

```
DirectoryIndex index.html index.htm index.php default.htm
```

Finally, note that you can also provide a relative URL if you want to load content from some other directory, such as a CGI program:

```
DirectoryIndex /cgi-bin/index.pl
```

### See Also

- 

[http://httpd.apache.org/docs/mod/mod\\_dir.html](http://httpd.apache.org/docs/mod/mod_dir.html)

## Recipe 12.8 Setting Up a Default "Favicon"

### Problem

You want to define a default favorite icon, or "favicon," for your site, but allow individual sites or users to override it.

### Solution

Put your default *favicon.ico* file into the */icons/* subdirectory under your *ServerRoot*, and add the following lines to your server configuration file in the scope where you want it to take effect (such as inside a particular *<VirtualHost>* container or outside all of them):

```
AddType image/x-icon .ico
<Files favicon.ico>
    ErrorDocument 404 /icons/favicon.ico
</Files>
```

### Discussion

*favicon.ico* files allow web sites to provide a small (16 x 16 pixels) image to clients for use in labeling pages; for instance, the Mozilla browser will show the favicon in the location bar and in any page tabs. These files are typically located in the site's *DocumentRoot* or in the same directory as the pages that reference them.

What the lines in the solution do is trap any references to *favicon.ico* files that don't exist and supply a default instead. An *ErrorDocument* is used instead of a *RewriteRule*, because we want the default to be supplied only if the file isn't found where expected. A rewrite, unless carefully crafted, would force the specified file to be used regardless of whether a more appropriate one existed.

### See Also

- 

[Chapter 5](#)

# Appendix A. Using Regular Expressions in Apache

A number of the Apache web server's configuration directives permit (or require!) the use of what are called *regular expressions*. Regular expressions are used to determine if a string, such as a URL or a user's name, matches a pattern.

There are numerous resources that cover regular expressions in excruciating detail, so this appendix is not designed to be a tutorial for their use. Instead, it documents the specific features of regular expressions used by Apache—what's available and what isn't. Even though there are quite a number of regular expression packages, with differing feature sets, there are some commonalities among them. The Perl language, for instance, has a particularly rich set of regular expressions but only a small subset of them are available in the Apache regex library, which is different from Perl's.

Regular expressions, as mentioned, are a language that allows you to determine if a particular string or variable looks like some pattern. For example, you may wish to determine if a particular string is all uppercase, or if it contains at least 3 numbers, or perhaps if it contains the word "monkey" or "Monkey." Regular expressions provide a vocabulary for talking about these sort of tests. Most modern programming languages contain some variety of regular expression library, and they tend to have a large number of things in common, although they may differ in small details.

Apache 1.3 uses a regular expression library called *hsregex*, so called because it was developed by Henry Spencer. Note that this is the same regular expression library used in *egrep*, which is the same thing as *grep* on many Unixish platforms.

Apache 2.0 uses a somewhat more full-featured regular expression library called Perl Compatible Regular Expressions (PCRE), so called because it implements many of the features available in the regular expression engine that comes with the Perl programming language. While this appendix does not attempt to communicate all the differences between these two implementations, you should know that *hsregex* is a subset of PCRE, as far as functionality goes, so everything you can do with regular expressions in Apache 1.3, you can do in 2.0, but not necessarily the other way around.

To grossly simplify, regular expressions implement two kinds of characters. Some characters mean exactly what they say (for example, a G appearing in a regular expression will usually mean the literal character G), while some characters have special significance (for example, the period (.) will match any character at all—a wildcard character). Regular expressions can be composed of these characters to represent (almost) any desired pattern appearing in a string.



# A.1 What Directives Use Regular Expressions?

Two main categories of Apache directives use regular expressions. Any directive with a name containing the word *Match*, such as *FileMatch*, can be assumed to use regular expressions in its arguments. And directives supplied by the module *mod\_rewrite* use regular expressions to accomplish their work.

For more about *mod\_rewrite*, see [Chapter 5](#).

*SomethingMatch* directives each implement the same functionality as their counterpart without the *Match*. For example, the *RedirectMatch* directive does essentially the same thing as the *Redirect* directive, except that the first argument, rather than being a literal string, is a regular expression, which will be compared to the incoming request URL.

## A.1.1 Regular Expression Basics

To get started in writing your own regular expressions, you'll need to know a few basic pieces of vocabulary, such as shown in [Table A-1](#) and [Table A-2](#). These constitute the bare minimum that you need to know. Although this will hardly qualify you as an expert, it will enable you to solve many of the regex scenarios you will find yourself faced with.

Table A-1. A basic regex vocabulary

Character	Meaning
.	Matches any character. This is the wildcard character.
+	Matches one or more of the previous character. For example, <i>M+</i> would match one or more Ms; <i>+</i> would match one or more characters of any kind.
*	Matches zero or more of the previous character. For example, <i>M*</i> would match zero or more Ms. This means that it will not only match M, MM, and MMM, but it will also match a string that doesn't have any Ms in it at all.
?	Makes the previous character optional. For example, the regular expression <i>monkeys?</i> will match a string containing either monkey or monkeys. Note that the <i>?</i> applies only to a single character in the absence of any enclosing parentheses.



# Appendix B. Troubleshooting

The Apache web server is a very complex beast. In the vanilla package it includes over 30 functional modules and more than 12 dozen configuration directives. This means that there are significant opportunities for interactions that produce unexpected or undesirable results. This appendix covers some of the more common issues that cause problems, as culled from various support forums.

[\[ Team LiB \]](#)

◀ PREVIOUS

NEXT ▶

## B.1 Troubleshooting Methodology

### B.1.1 In the Error Log

The Apache software does quite a reasonable job of reporting the details when it encounters problems. The reports are recorded in the server's error log, which is usually stored in one of the following places:

- */usr/local/apache/logs/error\_log*
- */var/log/apache/error\_log*
- */var/log/httpd-error.log*
- */var/log/httpd/error\_log*
- *C:\Program Files\Apache Group\error.log*

Where the error log is put depends upon how you installed and configured the server; the wealth of possible locations in the list above is because popular prepackaged installation kits (from Red Hat, SuSE, etc.) each has its own preferred location. Of course, the definitive location can be determined by examining your *httpd.conf* file for the *ErrorLog* directive(s).

So the very first thing you should do when Apache appears to be misbehaving is see if the server has any comments to make.

If the messages in the error log don't make the cause of the problem immediately clear, or if there aren't any messages that seem to relate to the problem, it's a good idea to crank the logging level up by changing the *LogLevel* setting in the *httpd.conf* file:

```
LogLevel debug
```

The debug setting enables all possible error messages and makes the server extremely verbose, so it's a good idea to set it back to warning or error after it has helped you locate the cause of your problem.

### B.1.2 Characterize the Problem

When you're trying to diagnose a problem, here is a question you should ask yourself: "What is the current behavior, and in what ways is it different from the expected or desired behavior?"

If you ask this question, a natural successive question is, "What could cause the current behavior?"

Between the answers to these two questions often lies a "Eureka!" moment. At the very least, they narrow your area of research



## B.2 Debugging the Configuration

When diagnosing a problem by examining your server's configuration, be sure to examine all of the files involved. In particular, look for files identified in Include directives, as well as those in the main *httpd.conf* file and in *.htaccess* files.

If you're editing the server-wide configuration files, be sure to restart the server afterward to make the changes take effect!

If editing a configuration or *.htaccess* file seems to have no effect, test that it's actually being processed by putting a line of gibberish into the file and trying again.

If it seems that an *.htaccess* file is being ignored, even when you insert gibberish, it indicates that it's within the scope of an *AllowOverride None* directive.



## B.3 Debugging Premature End of Script Headers

When you're working with CGI scripts, certain messages can quickly become extremely familiar and tiresome; typically the output in the browser window will be either a blank page or an Internal Server Error page.

This message has several different possible causes. These include, but are not necessarily limited to:

- The CGI script is either not emitting any output at all, or it is emitting content before the required header lines, or it's neglecting to emit the obligatory blank line between the header and the content.
- The script encountered an error and emitted the error message instead of its expected output.
- You're using *suexec* and one or more of the *suexec* constraints has been violated.

To test to see if the problem is an error condition or improper CGI response formatting, run the script interactively from the command line to verify that it is emitting content in compliance with the CGI rules.

If you're using *suexec*, check the *suexec* logfile to see if there are security constraints being violated.

You can tell if you're using *suexec* with the following command:

```
% httpd -l
Compiled-in modules:
  http_core.c
  mod_so.c
suexec: disabled; invalid wrapper /var/www/apache/bin/suexec
```

If you get a message that says that *suexec* is disabled, you can ignore that as a possible cause of the script's execution problems.

If *suexec* is enabled, though, you should look at its logfile to get more details about the problem. You can find the logfile with:

```
# suexec -V
-D DOC_ROOT="/usr/local/apache/htdocs"
-D GID_MIN=100
-D HTTPD_USER="www"
-D LOG_EXEC="/usr/local/apache/logs/suexec.log"
-D SAFE_PATH="/usr/local/bin:/usr/bin:/bin"
-D UID_MIN=100
-D USERDIR_SUFFIX="public_html"
```

The important line is `-D LOG_EXEC="/usr/local/apache/logs/suexec.log"`; it tells you exactly where *suexec* is recording its errors.

You can find out more about CGI and *suexec* here:

-





## B.4 Common Problems on Windows

Windows has its own distinct set of problem areas that don't apply to Unixish environments.

### B.4.1 Cannot Determine Hostname

When trying to start Apache from a DOS window, you receive a message like "Cannot determine hostname. Use ServerName directive to set it manually."

If you don't explicitly supply Apache with a name for your system, it tries to figure it out. This message is the result of that process failing.

The cure for this is really quite simple: edit your *conf/httpd.conf* file, look for the string *ServerName*, and make sure there's an uncommented directive such as:

```
ServerName localhost
```

or:

```
ServerName www.foo.com
```

in the file. Correct it if there is one there with wrong information, or add one if you don't already have one.

Also, make sure that your Windows system has DNS enabled. See the TCP/IP setup component of the Networking or Internet Options control panel.

After verifying that DNS is enabled and that you have a valid hostname in your *ServerName* directive, try to start the server again.

### B.4.2 Finding WS2\_32.DLL on Windows

When trying to start Apache on Windows 95, a message like Unable To Locate WS2\_32.DLL... appears. This file is necessary for Apache to function properly.

Prior to Version 1.3.9, Apache for Windows used Winsock 1.1. Beginning with Version 1.3.9, Apache began using Winsock 2 features (specifically, *WSADuplicateSocket()*). *WS2\_32.DLL* implements the Winsock 2 API. Winsock 2 ships with Windows NT 4.0 and Windows 98. Some of the earlier releases of Windows 95 did not include Winsock 2.

To fix it, install Winsock 2, available at <http://www.microsoft.com/windows95/downloads/>. Then restart your server, and the problem should be gone.

### B.4.3 Fixing WSADuplicateSocket Errors

If, when trying to start Apache on Windows, it fails and the Apache error log contains this message:

```
[crit] (10045) The attempted operation is not supported for the type of object
```



## B.5 Fixing Build-Time Error Messages

### B.5.1 `__inet` Symbols

If you have installed BIND-8, then this is normally due to a conflict between your include files and your libraries. BIND-8 installs its include files and libraries in `/usr/local/include/` and `/usr/local/lib/`, while the resolver that comes with your system is probably installed in `/usr/include/` and `/usr/lib/`.

If your system uses the header files in `/usr/local/include/` before those in `/usr/include/` but you do not use the new resolver library, then the two versions will conflict. To resolve this, you can either make sure you use the include files and libraries that came with your system, or make sure to use the new include files and libraries.

If you're using Apache 2.0 or later, or Apache 1.3 with the APACI build script, you can make changes to the library search lists by defining them on the `./configure` command line:

```
% LIBS=-lbind ./configure ...
```

If you're using Apache 1.3 or earlier and controlling the build process by editing the *Configuration* file directly, just add `-lbind` to the `EXTRA_LDFLAGS` line in the file.

After making the appropriate change to your build configuration process, Apache should build with the correct library.



Apache Versions 1.2 and earlier use `EXTRA_LFLAGS` in the *Configuration* file instead.

As of BIND 8.1.1, the *bind* libraries and files are installed under `/usr/local/bind` by default, so you should not run into this problem. Should you want to use the *bind* resolvers, you'll have to add the following to the respective lines:

- 

For Apache 1.3 with APACI, or 2.0 and later:

```
% CFLAGS=-I/usr/local/bin/include \  
> LDFLAGS=/usr/local/bind/lib LIBS=-lbind \  
> ./configure ...
```

- 

For Apache 1.2 or 1.3 with direct editing of *Configuration*, add/change the following lines in the file:

```
EXTRA_CFLAGS=-I/usr/local/bind/include  
EXTRA_LDFLAGS=-L/usr/local/bind/lib  
EXTRA_LIBS=-lbind
```

## B.6 Getting Server-Side Includes to Work

The solution is to make sure that *Options Includes* is turned on and that either *XBitHack* is turned *On*, or that you have the appropriate *AddHandler* directives set on the file type that you are using.

As discussed in [Recipe 8.8](#), there are a number of ways to enable SSI. If the unparsed SSI directives are appearing in the HTML when the page is loaded, this is a clear indication that SSI execution is not enabled for the document in question.

If the server has difficulty parsing an SSI directive, it will substitute the phrases "An error occurred while processing this directive" in its place in the response. If this happens, the cause of the problem should be listed in the server's error log. See also [Recipe 8.12](#).

## B.7 Debugging Rewrites That Result in "Not Found" Errors

If your *RewriteRule* directives keep resulting in 404 Not Found error pages, add the PT (PassThrough) flag to the *RewriteRule* line. Without this flag, Apache won't process a lot of other factors that might apply, such as *Alias* settings.

You can verify that this is the cause of your problem by cranking the *mod\_rewrite* logging level up to 9 and seeing that the entries relating to the *RewriteRule* mention something about prefixes with *document\_root*:

```
RewriteLog logs/rewrite-log  
RewriteLogLevel 9
```

```
% tail logs/rewrite_log  
ip-address - - [date ] [reqid ] (2) prefixed with document_root to  
/usr/local/apache/htdocs/robots.text  
ip-address - - [date ] [reqid ] (1) go-ahead with /usr/local/apache/htdocs/robots.text  
[OK]
```



Don't forget to turn off the *RewriteLog* directive, or possibly just turn down the logging level, after you've done your checking! Otherwise your disk space may disappear like the snows of yesteryear.

Without the PT flag, *mod\_rewrite* assumes that any rewriting it does will be the last URL manipulation the server needs to do for the request. Since *mod\_rewrite* directives are handled very early in request processing, this can mean that *Alias*, *ScriptAlias*, and other URL manipulations may not get executed. Specifying the flag tells *mod\_rewrite* to not short-circuit processing, but let it continue as usual.

## B.8 .htaccess Files Having No Effect

Make sure that *AllowOverride* is set to an appropriate value. Then, to make sure that the *.htaccess* file is being parsed at all, put the following line in the file and ensure that it causes a server error page to show up in your browser:  
Garbage Goes Here

*.htaccess* files override the settings in the main server configuration file. Because this is frequently an undesired thing, *.htaccess* files are frequently disabled, which will cause your *.htaccess* file to be ignored.

*.htaccess* files are enabled using the *AllowOverride* directive, which lists categories of directives that may appear in an *.htaccess* file. For example, if you wish to put authentication-related directives in an *.htaccess* file, you will need to put the following line in the main server configuration file:

```
AllowOverride AuthConfig
```

*AllowOverride All* permits any directive to be put in the *.htaccess* file, while the directive *AllowOverride None* means, "Please ignore my *.htaccess* files."

Thus, the most common cause of an *.htaccess* file being ignored is simply that your configuration file tells Apache to ignore it.

If you put garbage in your *.htaccess* file, this should generate a Server Error message in the browser, which will verify that Apache is indeed looking at the contents of your file. However, if such a message is not displayed, this is a sure sign that your *.htaccess* file is being completely ignored.

## B.9 Address Already in Use

If, when attempting to start your Apache server, you get the following error message:

```
[Thu May 15 01:23:40 2003] [crit] (98)Address already in use: make_sock: could not bind to port 80
```

One of three things is happening:

- You are attempting to start the server as a nonroot user. Become the root user and try again.
- There is already some process running (perhaps another Apache server) using port 80. Run *netstat*, or perhaps look at the process list and kill any process that seems to fill this role.
- You have more than one *Listen* directive in your configuration file pointing to the same port number. Find the offending duplicate directive and remove it.

In the case of the first condition, you will need to become the root user in order to start Apache. By long tradition, only the root user may bind to any port lower than 1025. Since Apache typically runs on port 80, this requires root privileges.

The second condition can be a little trickier. Sometimes a child process will refuse to die and will remain running after Apache has been shut down. There are numerous reasons this might happen. Most of the time, you can kill this process forcibly using *kill* or *kill -9* while logged in as root. As long as this process is running and has the port occupied, you will be unable to start anything else wanting to bind to that same port.

In the case of the third condition, the second *Listen* directive attempts to bind to port 80, which has already been taken by the first *Listen* directive. Simply removing one of the *Listen* directives will clear up this problem.

## Colophon

Our look is the result of reader comments, our own experimentation, and feedback from distribution channels. Distinctive covers complement our distinctive approach to technical topics, breathing personality and life into potentially dry subjects.

The animal on the cover of *Apache Cookbook* is a moose. The moose roams the forests of North America, Europe, and Russia. It's the largest of the deer family, and the largest moose of all, *Alces alces gigas*, is found throughout Alaska. This particular moose, in fact, is so ubiquitous that it's played an important role in the development of the state-though the relationship between moose and men is often adversarial.

Moose have a high reproductive potential and can quickly fill a range to capacity. And in Alaska, the removal of mature timber through logging and fire has benefited them by providing new stands of young timber-high-quality moose food. Moose get to be a pain when they eat crops, stand on airfields, wander the city streets, and collide with cars and trains.

But in general, these animals are good for the state's economy. Moose are an essential part of the Alaskan landscape, providing tourist photo opportunities when they feed along the highway. Residents and out-of-state hunters harvest 6,000 to 8,000 moose annually-approximately 3.5 million pounds of meat. The future for these animals in Alaska is reasonably bright because humans are learning how to manage moose habitat with wildlife and how to mitigate factors that affect moose populations, such as hunting and predation by wolves and bears.

Sarah Sherman was the production editor and copyeditor, and Marlowe Shaeffer was the proofreader for *Apache Cookbook*. Matt Hutchinson and Claire Cloutier provided quality control. Julie Hawks wrote the index.

Emma Colby designed the cover of this book, based on a series design by Edie Freedman. The cover image is an original engraving from *The Illustrated Natural History: Mammalia*. Emma Colby produced the cover layout with QuarkXPress 4.1 using Adobe's ITC Garamond font.

David Futato designed the interior layout. This book was converted by Julie Hawks to FrameMaker 5.5.6 with a format conversion tool created by Erik Ray, Jason McIntosh, Neil Walls, and Mike Sierra that uses Perl and XML technologies. The text font is Linotype Birka; the heading font is Adobe Myriad Condensed; and the code font is LucasFont's TheSans Mono Condensed. The illustrations that appear in the book were produced by Robert Romano and Jessamyn Read using Macromedia FreeHand 9 and Adobe Photoshop 6. The tip and warning icons were drawn by Christopher Bing. This colophon was written by Ed Stephenson.

The online edition of this book was created by the Safari production group (John Chodacki, Becki Maisch, and Madeleine Newell) using a set of Frame-to-XML conversion and cleanup tools written and maintained by Erik Ray, Benn Salter, John Chodacki, and Jeff Liggett.

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[#apache IRC channel](#)

[#exec SSI directive](#)

[#include directive](#)

[#include virtual, URIs passed to](#)

[%a format effector](#)

[%A format effector](#)

[%d format effector](#)

[%h format effector](#)

[%i format effector](#)

[%m format effector](#)

[%M format effector](#)

[%v logging format](#)

[%Y format effector](#)

[<Directory proxy:\\*> section](#)

[<Directory> directive 2nd](#)

[<Directory> section](#)

[<Files> section](#)

[directives and](#)

[<FilesMatch> section 2nd](#)

[<Location> section](#)

[directives and](#)

[<LocationMatch> section](#)

[<Proxy> directive](#)

[<VirtualHost> section](#)

[directives 2nd](#)

[displaying same content on several addresses](#)

[log](#)

[name-based virtual host](#)

[.cgi files](#)

[.gif files](#)

[.htaccess files \[See also configuration\] 2nd](#)

[having no effect](#)

[minimizing performance impact](#)

[renaming](#)

[restricting access](#)

[Satisfy Any directive](#)

[WebDAV](#)

[.htpasswd files](#)

[.shtml files](#)

[\\_default\\_ <Virtual Host> keyword](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[ab \(Apache bench\)](#)

[Accept header field](#)

[access](#)

[denying to unreferrred requests](#)

[restricted documents](#)

[AccessFileName directive](#)

[account information](#)

[Action directive](#)

[activity logging 2nd](#)

[AddCharset directive](#)

[AddFilter directive](#)

[AddHandler directive 2nd 3rd 4th](#)

[identifying extensions](#)

[AddLanguage directive](#)

[addModule directive](#)

[address already in use error](#)

[address-based virtual hosts](#)

[default](#)

[name-based and](#)

[setting up](#)

[AddType directive](#)

[alert log level](#)

[Alias directive](#)

[invalid](#)

[trailing slash](#)

[aliasing](#)

[several URLs with single directive](#)

[AliasMatch directive](#)

[Allow directive](#)

[Allow from all directive](#)

[Allow from directive](#)

[AllowOverride directive](#)

[Andrews, Bryan](#)

[Apache](#)

[building from sources](#)

[modules](#)

[mod\\_auth](#)

[mod\\_auth\\_digest](#)

[mod\\_auth\\_mysql](#)

[mod\\_authn\\_dbi](#)

[mod\\_autoindex 2nd](#)

[mod\\_cache](#)

[mod\\_cgi](#)

[mod\\_dav](#) [See mod\_dav module]

[mod\\_dir 2nd](#)

[mod\\_ext\\_filter](#)

[mod\\_file\\_cache](#)

[mod\\_log\\_config 2nd](#)

[mod\\_log\\_sql](#)

[mod\\_logio](#)

[mod\\_mime](#)

[mod\\_mmap\\_static 2nd](#)

[mod\\_perl](#) [See mod\_perl modules]

[mod\\_php](#) [See mod\_php module]

[mod\\_proxy](#) [See mod\_proxy module]



[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[bandwidth, stealing](#)

[Basic versus Digest authentication methods](#)

[benchmarking](#)

[bin directory 2nd](#)

[BIND-8 and build-time error messages](#)

[books \(Apache web server\)](#)

[browser software, logging](#)

[browsers](#)

[Content-Type headers](#)

[credentials and](#)

[building Apache from sources](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[CA, importing](#)

[CA.pl script](#)

[CACertificateFile directive](#)

[CacheFile directive](#)

[CacheNegotiatedDocs directive](#)

[CacheRoot directive](#)

[caching files](#)

[cadaver](#)

[cannot determine host name error \(Windows\)](#)

[case-insensitive URLs](#)

[certificate signing request](#)

certificates

[signing](#)

[SSL](#)

[authentication](#)

[keys](#)

CGI programs

[arguments, rewriting path information](#)

directories

[creating one for each user](#)

[mapping several URLs to same](#)

[not-script aliased](#)

[only CGI scripts](#)

[identifying by extension](#)

[making content filters](#)

[outputting to HTML files](#)

[reading from parameters](#)

[redirecting document requests to](#)

[running as different user](#)

[testing](#)

[Windows](#)

[CGI.pm module](#)

[CGI::Carp](#)

[cgic C library](#)

[CheckSpelling directive](#)

[ClearModuleList directive](#)

[combined log format](#) [2nd](#)

[common log format](#) [2nd](#)

[conf directory](#) [2nd](#)

[config SSI directive](#)

[configuration](#) [See also [.htaccess file](#)]

[debugging](#)

[directives](#)

[file, parsing process](#)

content

[creating URL for existing](#)

[dynamic](#) [See [dynamic content](#)]

[filters](#)

[negotiation, disabling](#)

[serving out of directory rather than DocumentRoot](#)

[Content-length field](#)

[Content-Type header field, setting according to browser](#)

[Cookie header field](#)

cookies

[logging](#)



[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[dates, last modified](#)

[DAVLockDB location](#)

[debug log level](#)

[debugging](#) [See troubleshooting]

[DefaultType directive](#)

[Deny directive](#) [2nd](#)

[Deny from all directive](#)

[Deny from directive](#)

[Digest authentication](#)

[versus Basic](#)

directives

[#exec SSI directive](#)

[#include directive](#)

[<Directory>](#) [2nd](#)

[<FilesMatch>](#)

[<Proxy>](#)

[<VirtualHost>](#) [log](#)

[<VirtualHost>](#) [section](#) [See [<VirtualHost>](#) [section directive](#)]

[AccessFileName](#)

[Action](#)

[AddCharset](#)

[AddFilter](#)

[AddHandler](#) [See [AddHandler](#)]

[AddLanguage](#)

[addModule](#)

[AddType](#)

[Alias](#) [See [Alias directive](#)]

[aliasing several URLs](#)

[AliasMatch](#)

[Allow](#)

[Allow from](#)

[Allow from all](#)

[AllowOverride](#)

[AuthDigestFile](#)

[AuthType](#)

[CACertificateFile](#)

[CacheFile](#)

[CacheNegotiatedDocs](#)

[CacheRoot](#)

[CheckSpelling](#)

[ClearModuleList](#)

[config SSI](#)

[configuration](#)

[CustomLog](#) [See [CustomLog directive](#)]

[DefaultType](#)

[Deny](#) [2nd](#)

[Deny from](#)

[Deny from all](#)

[ErrorDocument](#) [See [ErrorDocument directive](#)]

[ErrorDocument](#) [403](#)

[ErrorDocument](#) [404](#)

[ErrorLog](#) [See [ErrorLog directive](#)]

[FollowSymLinks](#)

[in <Directory proxy:\\*>](#) [section](#)

[Include](#) [2nd](#)



[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[email notification, receiving when error occurs](#)

[emerg log level](#)

[encryption technology, distributing](#)

[end of script headers, debugging](#)

error

[conditions, notification](#)

[documents in multiple languages](#)

[handling](#)

[log level](#)

[logging](#)

[logs](#)

messages

[customized](#)

[fixing build-time](#)

[Premature end of script headers](#)

pages

[404 Not Found](#)

[IE displaying its own](#)

ErrorDocument

[as a dynamic page](#)

[ErrorDocument 403 directive](#)

[ErrorDocument 404 directive](#)

[setting to a list of available sites on server](#)

[ErrorDocument directive](#) [2nd](#) [3rd](#)

[notification on error conditions](#)

[ErrorLog directive](#)

[syslog and](#)

[virtual hosts and](#)

extensions

[protecting all files except for certain](#)

[using to identify CGI scripts](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[favicon](#)

[files](#)

[access outside web root](#)

[caching](#)

[permissions](#)

[protecting all except for certain extensions](#)

[protecting server](#)

[filtering proxied content](#)

[folder listings](#)

[disabling indexing](#)

[enabling](#)

[generating](#)

[FollowSymLinks directive](#)

[footers, including in HTML files](#)

[FTP port](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[GDBM files](#)

[global scope](#)

[gone redirection](#)

[Group directive](#) [2nd](#)

[protecting server files](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

header fields

[Accept](#)

[Cookie](#)

[Host:](#)

[logging arbitrary request](#)

[logging arbitrary response](#)

[Set-Cookie](#)

[Status CGI response](#)

[User-agent](#)

headers

[debugging premature end of script](#)

[including in HTML files](#)

[home page, setting something other than index.html as](#)

[Host field, handling missing](#)

[Host: header fields](#)

[HostNameLookups directive](#) [2nd](#)

hostnames

[logging instead of IP addresses](#)

[turning directories into](#)

[hsregex](#)

[htdigest utility](#)

HTML files

[dynamic content](#)

[including CGI output](#)

[including headers and footers](#)

[htpasswd utility](#)

HTTP

[\(secure\) module, installing](#)

[access control](#)

[port](#)

HTTP status codes

[301](#)

[302](#)

[303](#)

[400 series status code](#)

[401 \(unauthorized\) code](#)

[403 Forbidden error](#) [2nd](#) [3rd](#)

[Deny directive and](#)

[open mail relay, preventing](#)

[404 Not Found error](#) [2nd](#) [3rd](#) [4th](#)

[sent as 403 Forbidden instead](#)

[SSI directives and](#)

[using ErrorDocument to catch](#)

[405 status code \(Method Not Allowed\)](#)

[410](#)

[500 series status code](#)

[list of](#)

[various directive formats and](#)

[HTTPS port](#)



[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[icons directory](#)

[IdentityCheck On directive](#)

[IE error pages](#)

[images](#)

[files](#)

[logging requests](#)

[restricting from being used off-site](#)

[Include directive](#) [2nd](#)

[include directory](#) [2nd](#)

[index.html](#)

[language variants](#)

[setting other file as default](#)

[info log level](#)

[installing Apache](#)

[ApacheToolbox](#) [See ApacheToolbox, installing with]

[prepackaged kits](#)

[Red Hat Linux](#)

[Windows](#) [See Windows, installing Apache]

[Internal Server Error page](#)

[Internet.Com](#)

[IP addresses](#) [See also address-based virtual hosts]

[logging hostnames instead of](#)

[logging proxied](#)

[logging server](#)

[IP-based virtual hosts](#) [See address-based virtual hosts]

[IRC channel, #apache](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[keepalive-related directives](#)

[key, private](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[LDAP directories](#)

[libexec directory 2nd](#)

[Limit directive](#)

[LimitExcept directive](#)

[load](#)

[distributing evenly](#)

[sharing between servers](#)

[LoadModule directive](#)

[localhost address](#)

[lock databases on NFS filesystem](#)

[logfiles](#) [\[See also logging\]](#)

[error logs](#)

[maintaining separate logs for each virtual host](#)

[misspellings](#)

[rotating](#)

[first of month](#)

[rotatelog program](#)

[splitting up](#)

[LogFormat directive](#)

[logging server IP addresses](#)

[splitting up a logfile](#)

[logging 2nd](#) [\[See also logfiles\]](#)

[activity](#) [\[See activity logging\]](#)

[arbitrary request header fields](#)

[arbitrary response header fields](#)

[browser software](#)

[combined log format](#)

[common log format](#)

[cookies](#)

[errors](#)

[errors for virtual hosts to multiple files](#)

[hostnames instead of IP addresses](#)

[HTTP status codes](#)

[image requests](#)

[MAC \(hardware\) address](#)

[more detailed errors](#)

[more details in entries](#)

[MySQL database activity](#)

[POST contents](#)

[proxied IP address](#)

[proxy requests](#)

[referring page](#)

[requests by day or hour](#)

[server IP address](#)

[syslog](#)

[types of data](#)

[user directories](#)

[virtual hosts](#)

[LogLevel directive](#)

[documentation](#)

[LogLevel, default value](#)

[logresolve application](#)

[logs directory 2nd](#)

[loopback address](#)



[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[MAC \(hardware\) address, logging](#)

[Mail::Sendmail](#)

[Mail::Sendmail CPAN module](#)

[mailing lists](#)

[man directory](#) [2nd](#)

[MaxClients directive](#)

[MaxSpareServers directive](#)

[MaxSpareThreads directive](#)

[memory, determining how much you need](#)

[methods, limiting by user](#)

[Microsoft Software Installer \(MSI\) package](#)

[MIME type](#)

[MinSpareServers directive](#)

[MinSpareThreads directive](#)

[misspellings in log files](#)

[MMapFile directive](#) [2nd](#)

[mod\\_auth module](#)

[mod\\_auth\\_digest module](#)

[mod\\_auth\\_mysql module](#)

[mod\\_authn\\_dbi module](#)

[mod\\_autoindex module](#) [2nd](#)

[mod\\_cache module](#)

[mod\\_cgi module](#)

[mod\\_dav module](#)

Unix

[enabling](#)

[installing](#)

Windows

[enabling](#)

[installing](#)

[mod\\_dir module](#) [2nd](#)

[mod\\_ext\\_filter module](#)

[mod\\_file\\_cache module](#)

[mod\\_log\\_config module](#) [2nd](#)

[mod\\_log\\_sql module](#)

[mod\\_logio module](#)

[mod\\_mime module](#)

[mod\\_mmap\\_static module](#) [2nd](#)

[mod\\_perl modules](#)

[Apache::AuthExpire](#)

[Apache::Constants](#)

[Apache::Cookbook::Example](#)

[Apache::Htpasswd::Perishable](#)

[Apache::Perldoc](#)

[Apache::PerlRun](#)

[Apache::Registry](#) [2nd](#)

[CGI::Carp](#)

[dynamic content](#)

handler

[modules](#)

[writing](#)

[installing on Unix](#)

[Mail::Sendmail](#)

[ModPerl::PerlRun](#)

[ModPerl::Registry](#)



[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[name-based virtual hosts](#)

[address-based and](#)

[default](#)

[DNS server and](#)

[setting up](#)

[SSL and](#)

[NameVirtualHost directive](#) [2nd](#)

[name-based virtual host](#)

[Netscape cookie proposal](#)

[network MPM](#)

[NFS filesystem, lock databases on](#)

[non-script aliased directories](#)

[notice log level](#)

[NumServers directive](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[online tutorials \(Apache\)](#)

[open mail relay](#)

[Options directive 2nd](#)

[+ExecCGI argument and](#)

[-FollowSymLinks](#)

[SymlinksIfOwnerMatch](#)

[Order directive](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[parsing Apache configuration file](#)

[partial downloads](#)

[passwords](#)

[authentication](#) [2nd](#)

[expiring](#)

[managing .htpasswd files](#)

[password file for Digest authentication](#)

[preventing brute-force attacks](#)

[single use](#)

[PCRE \(Perl Compatible Regular Expressions\)](#)

[perchild MPM](#)

[performance](#)

[minimizing impact of .htaccess files](#)

[Perl CGI programs, speeding up](#)

[Perl Compatible Regular Expressions \(PCRE\)](#)

[PerlHandler directive](#)

[PerlRun mode](#)

[permanent redirection](#)

[PHP](#)

[dynamic content](#)

[module installation on Windows](#)

[script handling, enabling](#)

[seeing syntax-enhanced source without setting up symbolic links](#)

[verifying installation](#)

[port-based virtual hosting](#)

[ports](#)

[POST logging](#)

[prefork MPM](#)

[private key, generating](#)

[process creation, optimizing](#)

[proxies](#) [\[See also proxy servers\]](#) [2nd](#)

[filtering proxied content](#)

[logging IP address](#)

[logging requests](#)

[restricting access to URLs](#)

[proxy directory](#) [2nd](#)

[proxy servers](#)

[authorization](#)

[configuring caching](#)

[content filter, using as](#)

[forwarding requests to](#)

[preventing being used as open relay](#)

[securing](#)

[ProxyBlock directive](#) [2nd](#)

[ProxyPass directive](#) [2nd](#) [3rd](#)

[ProxyPassReverse directive](#) [2nd](#)

[Python module, installing](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[RAM, determining how much you need](#)

[range requests](#)

[Red Hat Linux](#)

[installing Apache](#)

[up2date tool](#)

[Redirect directive](#) [2nd](#)

[preserving path information](#)

[URL served as HTTPS](#)

[redirecting](#) [2nd](#)

[all requests to single host](#)

[document requests to CGI script](#)

[several URLs to same destination](#)

[to secure area](#)

[RedirectMatch directive](#) [2nd](#) [3rd](#)

[Registry mode](#)

[regular expressions](#)

[character classes](#)

[directives](#)

[examples](#)

[library](#)

[resources](#)

[vocabulary](#)

[releases, tags used to label versions](#)

[Require file-owner directive](#)

[response status, changing](#)

[restricted documents](#)

[restrictions, removing selectively](#)

[RewriteCond directive](#) [2nd](#)

[logging proxy requests](#)

[NC \(No Case\) flag](#)

[wrappers](#)

[RewriteRule directive](#) [2nd](#) [3rd](#) [4th](#) [5th](#)

[\[PT\] flag](#)

[404 Not Found error pages and](#)

[logging proxy requests](#)

[multiple](#)

[rewriting](#)

[based on query string](#)

[RFC 2109](#)

[RFC 2165](#)

[RFC 2616](#)

[rotate-log-monthly.pl script](#)

[rotatelog program](#)

[round robin, DNS](#)

[RSA key \(genrsa\)](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[Satisfy directive](#) [2nd](#)

[requiring both weak and strong authentication](#)

[ScriptAlias directive](#)

[ScriptAlias directives](#)

[ScriptAlias'ed directories](#)

[ScriptAliasMatch directive](#)

[ScriptInterpreterSource directive](#) [2nd](#)

[sections and directives](#)

[secure HTTP module, installing](#)

[Secure Socket Layers](#) [\[See SSL\]](#)

[security](#) [\[See also authentication; authorization\]](#) [2nd](#)

[relaxing in subdirectories](#)

[removing restrictions selectively](#)

[seeother redirection](#)

[server-side includes](#)

[server-status handler](#)

[ServerAlias directive](#)

[ServerName](#)

[directives](#)

[incorrect or missing](#)

[ServerRoot directive](#)

[servers](#)

[activity](#)

[protecting files](#)

[proxy](#) [\[See proxy servers\]](#)

[Set-Cookie header field](#)

[SetEnv directive](#)

[SetEnvIf directive](#)

[SetHandler perl-script directive](#)

[SetOutputFilter directive](#)

[size restrictions on documents](#)

[slash, trailing](#)

[SMTP relay](#)

[Spencer, Henry](#)

[split-logfile program](#) [2nd](#)

[Squid proxy server](#)

[SSI #include directive](#)

[SSI directive](#) [2nd](#) [3rd](#) [4th](#)

[parsing for](#)

[SSI include directive](#)

[SSIs, enabling](#)

[SSL \(Secure Socket Layers\)](#)

[authenticating with client certificates](#)

[generating certificates](#)

[keys](#)

[installing](#)

[name-based virtual hosts](#)

[redirecting to](#)

[serving a portion of site via](#)

[support](#)

[SSLCertificateFile directive](#)

[SSLRequireSSL directive](#)

[starting and stopping Apache](#)

[Status CGI response header field](#)

[status for a response, changing](#)



[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[temporary redirection 2nd](#)

[third-party modules, installing](#)

[thread creation, tuning](#)

[ThreadsPerChild directive](#)

[trailing slash](#)

[troubleshooting](#)

[.htaccess files](#)

[address already in use error](#)

[cannot determine host name error](#)

[premature end of script headers](#)

[server-side includes](#)

[system error 1067](#)

[Windows](#)

[WSADuplicateSocket errors](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[uninstalling Apache](#)

[unreferred requests, denying access to](#)

[up2date tool \(Red Hat Linux\)](#)

[upload documents, limiting](#)

[URIs passed to #include virtual](#)

URLs

[accessing credentials embedded in](#)

[aliasing with single directive](#)

[case-insensitive](#)

[creating for existing content](#)

[giving each user own](#)

[logging referring page](#)

mapping

[several to same CGI directory](#)

[to directories](#)

[redirecting](#) [See [redirecting](#)]

[replacing text in](#)

[restricting proxy access](#)

[rewriting path information in CGI arguments](#)

User directive [2nd](#)

[protecting server files](#)

[user directories, logging](#)

[User-agent field](#)

[UserDir directive](#)

usernames

[authentication](#)

[users@httpd.apache.org mailing list](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[versions, tags used to label](#)

[virtual hosts](#)

[address-based](#) [See address-based virtual hosts]

[displaying same content on several addresses](#)

[hosting multiple with same configuration](#)

[using rewrite rules](#)

[logging](#)

[logging errors to multiple files](#)

[maintaining separate logfiles](#)

[name-based](#) [See name-based virtual hosts]

[port-based](#)

[splitting up a logfile](#)

[Virtual Private Networking \(VPN\)](#)

[VirtualDocumentRoot directive](#)

[VirtualScriptAlias directive](#)

[VPN \(Virtual Private Networking\)](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[warn log level](#)

[warnings setting in Perl](#)

[watermark.cgi file](#)

**WebDAV**

[enabling server for](#)

[enabling without making files writable](#)

[securing](#)

[tools](#)

[Unix capabilities](#)

[Windows capabilities](#)

**Windows**

[cannot determine host name error](#)

[CGI programs](#)

[installing Apache](#)

[first time](#)

[module support status](#)

[PHP module installation](#)

[running Apache as service](#)

[troubleshooting](#)

[Winsock 2](#)

[worker MPM](#)

[wrappers](#)

[WS2\\_32.DLL](#)

[WSADuplicateSocket errors](#)

[\[ Team LiB \]](#)

[\[ Team LiB \]](#)

[\[SYMBOL\]](#) [\[A\]](#) [\[B\]](#) [\[C\]](#) [\[D\]](#) [\[E\]](#) [\[F\]](#) [\[G\]](#) [\[H\]](#) [\[I\]](#) [\[K\]](#) [\[L\]](#) [\[M\]](#) [\[N\]](#) [\[O\]](#) [\[P\]](#) [\[R\]](#) [\[S\]](#) [\[T\]](#) [\[U\]](#) [\[V\]](#) [\[W\]](#) [\[X\]](#)

[XBitHack directive](#)

[\[ Team LiB \]](#)