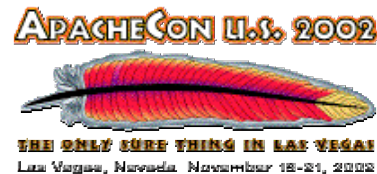


ApacheCon US 2002

Apache mod_rewrite the Swiss Army Knife of URL manipulation

by Mads Toftum



Intro

"The great thing about mod_rewrite is it gives you all the configurability and flexibility of Sendmail.
The downside to mod_rewrite is that it gives you all the configurability and flexibility of Sendmail."

-- Brian Behlendorf
Apache Group

Intro

"Despite the tons of examples and docs, `mod_rewrite` is voodoo.
Damned cool voodoo, but still voodoo. "

-- Brian Moore

Intro

But all this functionality and flexibility has its drawback: complexity.
So don't expect to understand this entire module in just one day.

docs/mod/mod_rewrite.html

Intro – agenda

- _ The main directives
 - RewriteRule
 - RewriteCond
- _ Comparing mod_rewrite to other solutions
- _ The directives
 - Main directives revisited
 - Other directives
- _ mod_rewrite solutions to common problems
- _ Questions

Intro – not covered

- _ Regular Expressions see:
 - perldoc perlre
 - "Mastering Regular Expressions" by Jeffrey Friedl

Directives – RewriteRule

Syntax: RewriteRule *Pattern Substitution*

Default: *None*

Context: server config, virtual host, directory, .htaccess

Override: *FileInfo*

- _ *RewriteRule is the main directive in mod_rewrite*
- _ RewriteRules are checked first
- _ RewriteRules are executed in the *URI translation* or *fixup* phase

Directives – RewriteCond

Syntax: RewriteCond *TestString CondPattern*

Default: *None*

Context: server config, virtual host, directory, .htaccess

Override: *FileInfo*

- _ *RewriteCond is used to add conditions to RewriteRules*
- _ One or more RewriteConds can precede a RewriteRule
- _ Multiple RewriteConds are joined by an implicit and unless [OR] is used

Imitation – Redirect

- _ The mod_alias version:

Redirect /test/ http://example.com/test/

RedirectPermanent /test2/ http://example.com/test2/

- _ The mod_rewrite version:

RewriteEngine On

RewriteRule ^/test/(.+) http://example.com/test/\$1 [R]

RewriteRule ^/test2/(.+) http://example.com/test2/\$1
[R=301]

Imitation – ProxyPass

- _ The mod_proxy version

ProxyPass /example/ http://example.com/

- _ The mod_rewrite version

RewriteRule ^/example/(.*) http://example.com/\$1 [P]

Imitation – UserDir

- _ The mod_userdir version

UserDir /usr/web -> /usr/web/joe/one/two.html

UserDir /home/*/web -> /home/joe/web/one/two.html

- _ The mod_rewrite version

RewriteRule ^/~(.*)/(.*)\$ /usr/web/\$1/\$2

RewriteRule ^/~([a-z][a-z0-9]+)/(.*)\$ /home/\$1/web/\$2

Imitation – mod_vhost_alias

The vhost_alias version:

```
VirtualDocumentRoot /vhosts/%0/docs/
```

_ The mod_rewrite version1:

```
RewriteEngine On
```

```
RewriteRule ^(.+)$ /vhosts/%{HTTP_HOST}/docs$1
```

_ The mod_rewrite version2:

```
RewriteCond %{HTTP_HOST} ^(.*)$
```

```
RewriteRule ^(.*)$ /vhosts/%1/docs$1
```

Imitation – trailing slash

- **Why does accessing directories only work when I include the trailing "/"**

ServerName example.com

- **The mod_rewrite solution:**

RewriteCond	%{REQUEST_FILENAME}	-d
RewriteRule	^(.+[^/])\$	\$1/ [R,L]

Directives – RewriteRule 2

_ RewriteRule *Pattern Substitution [Flags]*

[flags]:

- **'redirect|R [=code]'** (force redirect)
- **'proxy|P'** (force proxy)
- **'forbidden|F'** (force URL to be forbidden)
- **'gone|G'** (force URL to be gone)
- **'type|T=MIME-type'** (force MIME type)
- **'noescape|NE'** (no URI escaping of output)
- **'env|E=VAR:VAL'** (set environment variable)
- **'cookie|CO=NAME:VAL:domain[:lifetime[:path]]'** (set cookie)

Directives – RewriteRule 2b

_ RewriteRule *Pattern Substitution [Flags]*

[flags]:

- **'last|L'** (last rule)
- **'next|N'** (next round)
- **'chain|C'** (chained with next rule)
- **'skip|S=num'** (skip next rule(s))
- **'passthrough|PT'** (pass through to next handler)
- **'nosubreq|NS'** (used only if no internal sub-request)
- **'qsappend|QSA'** (query string append)
- **'nocase|NC'** (no case)

Directives – RewriteCond 2

- _ **Syntax: RewriteCond *TestString CondPattern [Flags]***
- _ *CondPattern is matched like perl regular expression, + these:*
 - *Lexical (string) compares (< > =)*
 - *'-d' (is directory)*
 - *'-f' (is regular file)*
 - *'-s' (is regular file with size > 0)*
 - *'-l' (is symbolic link)*

Directives – RewriteCond 2b

– **Syntax: RewriteCond *TestString CondPattern [Flags]***

- *'-F' (is existing file via subrequest)*
- *'-U' (is existing URL via subrequest)*

All of these can be negated with !

– ***[Flags]***

- *'nocase|NC' (no case)*
- *'ornext|OR' (or next condition)*

Directives – Other 1

- _ **RewriteEngine** *on|off* (*default off*)
- _ **RewriteLog** *file*
- _ **RewriteLogLevel** *Level* (*default 0*)
- _ **RewriteOptions** *Option* (*inherit*)
- _ **RewriteBase** *URL-path* (*context: dir, .htaccess*)

Directives – RewriteMap

Syntax: RewriteMap *MapName MapType:MapSource*

Default: not used per default

Context: server config, virtual host

- _ *Defines a mapping function for key lookups*
- _ *MapType in one of:*
 - *int (toupper,tolower,escape,unescape)*
 - *prg (external program)*
 - *txt*
 - *rnd (like txt but makes random choice from | seperrated list)*
 - *dbm[=type] (one of gdbm, sdbm, ndbm or db – chek valid types with: **bin/apxs -q AP_LIBS**)*

Directives – RewriteMap 2

- *$\$ \{ \text{MapName} : \text{LookupKey} \}$*
- *$\$ \{ \text{MapName} : \text{LookupKey} \mid \text{DefaultValue} \}$*

Example (crude load balancer/service level selection):

RewriteMap service rnd:/some/where/map.txt

*RewriteRule ^/(. *)/(. *)\$ http://\${service:\$1}/\$2 [P]*

/some/where/map.txt:

gold big1|big2|big3|big4|big5|big6 # gold service

silver small1 #
min service

Examples 1

_ No deep links / disallow external referers:

```
RewriteCond    %{HTTP_REFERER} !=""
```

```
RewriteCond %{HTTP_REFERER} !^http://example.com/.*$ [NC]
```

```
RewriteRule .* - [F]
```

_ Disallow UserAgent

```
RewriteCond %{HTTP_USER_AGENT} EvilBot.*
```

```
RewriteRule /nobot/. * http://%{REMOTE_ADDR}/ [R=301]
```


Examples 2

- _ How to restructure homedirs to avoid getting too many directories

RewriteEngine on

RewriteRule ^/~(([a-z])[a-z0-9]+)/(.*)
/data/\$2/\$1/\$3

or

RewriteMap LC int:tolower

RewriteRule ^/~(([a-z])[a-z0-9]+)/(.*) /data/
/data/\${LC:\$2}/\${LC:\$1}/\$3 [NC]

Examples 3

_ How to redirect a logged in user to a specific page:

RewriteEngine on

RewriteMap user-to-page txt:/path/to/map.user-to-page

RewriteCond %{LA-U:REMOTE_USER} ^(.+)\$

RewriteRule ^/\$ \${user-to-page:%1|/notfound.html}

/path/to/map.user-to-page

user1 /some/where/user1.html

user2 /some/where/else/user2.html

...

Examples 4

_ Webmail redirector

RewriteEngine on

RewriteCond %{HTTP_HOST} ^webmail\.(.*)\.dk\$ [NC]

RewriteRule ^/\$ http://mail.dk/login.x?dom=%1 [R]

Examples 5

- _ Moving from index files to autoindex – a regular Redirect
/download/index.htm /dir/ would loop

/download/.htaccess

RewriteEngine on

RewriteBase /download/

RewriteCond %{REQUEST_FILENAME} !-d

RewriteRule ^(.*)/index\.html* \$1/
[R=301]

Questions



Latest version:

<http://www.toftum.org/ApacheCon/>

mads@toftum.dk