Web Proxy

Document revision 1.5 (December 12, 2007, 11:44 GMT)
This document applies to V3.0

Table of Contents

Table of Contents
- Summary
- Quick Setup Guide
- Specifications
- Description

Setup
- Property Description
- Notes
- Example

Proxy Monitoring
- Property Description

Access List
- Description
- Property Description
- Notes

Direct Access List
- Description
- Property Description
- Notes

Cache Management
- Description
- Property Description

Connection List
- Description
- Property Description

Cache Contents
- Description
- Property Description

Cache inserts
- Description
- Property Description

Cache Lookups
- Description
- Property Description

Complementary Tools
- Description
- Command Description

Transparent Mode
- Description
- Notes
- Example

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

HTTP Methods
Description

General Information

Summary

The MikroTik RouterOS implements the following proxy server features:

- Regular HTTP proxy
- Transparent proxy. Can be transparent and regular at the same time
- Access list by source, destination, URL and requested method
- Cache access list (specifies which objects to cache, and which not)
- Direct Access List (specifies which resources should be accessed directly, and which - through another proxy server)
- Logging facility

Quick Setup Guide

To set up a 1 GiB large web cache, which will listen on port 8000, do the following:

```
[admin@MikroTik] ip proxy> set enabled=yes port=8000 max-cache-size=1048576
[admin@MikroTik] ip proxy> print
                   enabled: yes
               src-address: 0.0.0.0
                      port: 8000
              parent-proxy: 0.0.0.0
         parent-proxy-port: 0
               cache-drive: system
      cache-administrator: "webmaster"
           max-cache-size: 1048576KiB
            cache-on-disk: no
  max-client-connections: 600
  max-server-connections: 600
           max-fresh-time: 3d
   serialize-connections: no
       always-from-cache: no
            cache-hit-dscp: 4
[admin@MikroTik] ip proxy>
```

Remember to secure your proxy by preventing unauthorized access to it, otherwise it may be used as an open proxy. Also you need to setup destination NAT in order to utilize transparent proxying facility:

```
[admin@MikroTik] ip firewall nat> add chain=dstnat protocol=tcp dst-port=80
action=redirect to-ports=8000
[admin@MikroTik] ip firewall nat>
```

Specifications

Packages required: ***web-proxy***
License required: ***level3***
Home menu level: ***/ip web-proxy***
Standards and Technologies: ***HTTP/1.0***, ***HTTP/1.1***, ***FTP***
Hardware usage: *uses memory and disk space, if available (see description below)*

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

Description

Web proxy performs Internet object cache function by storing requested Internet objects, i.e., data available via HTTP and FTP protocols on a system positioned closer to the recipient than the site the data is originated from. Here 'closer' means increased path reliability, speed or both. Web browsers can then use the local proxy cache to speed up access and external reduce bandwidth consumption.

When setting up Web proxy, make sure it serves only your clients, and is not misused as relay. Please read the security notice in the Access List Section!

Note that it may be useful to have Web proxy running even with no cache when you want to use it as something like HTTP and FTP firewall (for example, denying access to mp3 files) or to redirect requests to external proxy with large cache drives transparently.

Setup

Home menu level: ***/ip proxy***

Property Description

always-from-cache (yes | no; default: **no**) - ignore client refresh requests if the content is considered fresh

cache-administrator (*text*; default: **webmaster**) - administrator's e-mail displayed on proxy error page

cache-drive (*systemname*; default: **system**) - specifies the target disk drive to be used for storing cached objects. You can use console completion to see the list of available drives

cache-hit-dscp (*integer*: 0..63) - automatically mark cache hit with the provided DSCP value

cache-on-disk (yes | no; default: **no**) - whether to store cache files on disk or in RAM filesystem

enabled (yes | no; default: **no**) - specifies whether the web proxy is enabled

max-cache-size (*none* | *unlimitedinteger*: 0..4294967295; default: **none**) - specifies the maximal disk cache size, measured in kibibytes

max-client-connections (*integer*; default: **600**) - maximum number of concurrent client connections accepted by the proxy. All further connections will be rejected

max-fresh-time (*time*; default: **3d**) - an upper limit on how long objects without an explicit expiry time will be considered fresh

max-server-connections (*integer*; default: **600**) - maximum number of concurrent proxy connections to external servers. All further connections will be put on hold until some of the existing server connections will terminate

parent-proxy (*IP addressport*; default: **0.0.0.0**) - IP address of the upper-level (parent) proxy

parent-proxy-port (*port*) - TCP port the parent proxy is active on

port (*port*; default: **3128**) - specifies the port(s) the web proxy will be listening on

serialize-connections (yes | no; default: **no**) - Do not make multiple connections to server for multiple client connections, if possible (i.e. server supports persistent HTTP connections). Clients will be served on FIFO principle; next client is processed when response transfer to the previous one is completed. If a client is idle for too long (max 5 seconds by default), it will give up waiting

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

and open another connection to the server

src-address (*IP address*; default: **0.0.0.0**) - the web-proxy will use this address connecting to the parent proxy or web site.

- **0.0.0.0** - appropriate src-address will be automatically taken from the routing table (preferred source of the respective route)

Notes

The web proxy listens to all IP addresses that the router has in its IP address list.

Example

To enable proxy on port 8080 with maximal available cache size:

```
[admin@MikroTik] ip proxy> set enabled=yes port=8080 \
\...  max-cache-size=unlimited
[admin@MikroTik] ip proxy> print
                  enabled: yes
              src-address: 0.0.0.0
                     port: 8000
             parent-proxy: 0.0.0.0
       parent-proxy-port: 0
              cache-drive: system
     cache-administrator: "webmaster"
          max-cache-size: 21000KiB
           cache-on-disk: no
  max-client-connections: 600
  max-server-connections: 600
          max-fresh-time: 3d
   serialize-connections: no
       always-from-cache: no
          cache-hit-dscp: 4
[admin@MikroTik] ip proxy>
```

Note how the **max-cache-size** value has been calculated from the **unlimited** to an accurate value in kibibytes

Proxy Monitoring

Command name: ***/ip proxy monitor***

Property Description

cache-used (*read-only: integer*) - the amount of disk (or RAM if the cache is stored only in RAM) used by the cache

free-disk-space (*read-only: integer*) - the amount of free space on the cache drive

hits (*read-only: integer*) - number of client requests resolved from the cache

hits-sent-to-clients (*read-only: integer*) - the amount of cache hits sent to client

received-from-servers (*read-only: integer*) - total amount of data received from the external servers

requests (*read-only: integer*) - total number of client requests to the proxy

sent-to-clients (*read-only: integer*) - total amount of data sent to the clients

status (*read-only: text*; default: **stopped**) - display status information of the proxy server

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

- **stopped** - proxy is disabled and is not running
- **running** - proxy is enabled and running
- **formatting-disk** - the cache drive is being formatted
- **checking-disk** - the cache drive is being checked for errors and cache inconsistencies
- **invalid-address** - proxy is enabled, but not running because of invalid address (you should change address or port)

total-disk-size (*read-only: integer*) - size of the cache drive

total-ram-used (*read-only: integer*) - the amount of memory used by the proxy (excluding RAM cache size)

uptime (*read-only: time*) - the time since the proxy has been started last time

Access List

Home menu level: ***/ip proxy access***

Description

Access list is configured in the same way as MikroTik RouterOS firewall rules. Rules are processed from the top to the bottom. First matching rule specifies decision of what to do with this connection. There is a total of 6 classifiers that specify matching constraints. If none of these classifiers is specified, the particular rule will match any connection.

If connection is matched by a rule, **action** property of this rule specifies whether connection will be allowed or not. If some connection does not match any rule, it will be allowed.

Property Description

action (*allow* | *deny*; default: **allow**) - specifies whether to pass or deny matched packets

dst-address (*IP addressnetmask*) - destination address of the IP packet

dst-host (*wildcard*) - IP address or DNS name used to make connection the target server (this is the string user wrote in his/her browser before specifying port and path to a particular web page)

dst-port (*port*) - a list or range of ports the packet is destined to

hits (*read-only: integer*) - the number of requests that were policed by this rule

local-port (*port*) - specifies the port of the web proxy via which the packet was received. This value should match one of the ports web proxy is listening on.

method (*any* | *connect* | *delete* | *get* | *head* | *options* | *post* | *put* | *trace*) - HTTP method used in the request (see HTTP Methods section at the end of this document)

path (*wildcard*) - name of the requested page within the target server (i.e. the name of a particular web page or document without the name of the server it resides on)

redirect-to (*text*) - in case access is denied by this rule, the user shall be redirected to the URL specified here

src-address (*IP addressnetmask*) - source address of the IP packet

Notes

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

It is strongly recommended to deny all IP addresses except those behind the router as the proxy still may be used to access your internal-use-only (intranet) web servers. Also, consult examples in Firewall Manual on how to protect your router.

Wildcard property **url** matches a complete string (i.e., they will not match "example.com" if they are set to "example"). Available wildcards are '*' (match any number of any characters) and '?' (match any one character). Regular expressions are also accepted here, but if the property should be treated as a regular expression, it should start with a colon (':').

Small hits in using regular expressions:

- **** symbol sequence is used to enter **** character in console
- **\.** pattern means **.** only (in regular expressions single dot in pattern means any symbol)
- to show that no symbols are allowed before the given pattern, we use **^** symbol at the beginning of the pattern
- to specify that no symbols are allowed after the given pattern, we use **$** symbol at the end of the pattern
- to enter **[** or **]** symbols, you should escape them with backslash ****.

Direct Access List

Home menu level: ***/ip proxy direct***

Description

If **parent-proxy** property is specified, it is possible to tell the proxy server whether to try to pass the request to the parent proxy or to resolve it connecting to the requested server directly. Direct Access List is managed just like Proxy Access List described in the previous chapter except the **action** argument.

Property Description

action (*allow* | *deny*; default: **allow**) - specifies the action to perform on matched packets

- **allow** - always resolve matched requests directly bypassing the parent router
- **deny** - resolve matched requests through the parent proxy. If no one is specified this has the same effect as allow

dst-address (*IP addressnetmask*) - destination address of the IP packet

dst-host (*wildcard*) - IP address or DNS name used to make connection the target server (this is the string user wrote in his/her browser before specifying port and path to a particular web page)

dst-port (*port*) - a list or range of ports the packet is destined to

local-port (*port*) - specifies the port of the web proxy via which the packet was received. This value should match one of the ports web proxy is listening on.

method (*any* | *connect* | *delete* | *get* | *head* | *options* | *post* | *put* | *trace*) - HTTP method used in the request (see HTTP Methods section in the end of this document)

path (*wildcard*) - name of the requested page within the target server (i.e. the name of a particular web page or document without the name of the server it resides on)

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

src-address (*IP addressnetmask*) - source address of the IP packet

Notes

Unlike the access list, the direct proxy access list has default action equal to **deny**. It takes place when no rules are specified or a particular request did not match any rule.

Cache Management

Home menu level: ***/ip proxy cache***

Description

Cache access list specifies, which requests (domains, servers, pages) have to be cached locally by web proxy, and which not. This list is implemented exactly the same way as web proxy access list. Default action is to cache object (if no matching rule is found).

Property Description

action (*allow* | *deny*; default: **allow**) - specifies the action to perform on matched packets

- **allow** - cache objects from matched request
- **deny** - do not cache objects from matched request

dst-address (*IP addressnetmask*) - destination address of the IP packet

dst-port (*port*) - a list or range of ports the packet is destined to

local-port (*port*) - specifies the port of the web proxy via which the packet was received. This value should match one of the ports web proxy is listening on.

method (*any* | *connect* | *delete* | *get* | *head* | *options* | *post* | *put* | *trace*) - HTTP method used in the request (see HTTP Methods section in the end of this document)

path (*wildcard*) - name of the requested page within the target server (i.e. the name of a particular web page or document without the name of the server it resides on)

path (*wildcard*) - name of the requested page within the target server (i.e. the name of a particular web page or document without the name of the server it resides on)

src-address (*IP addressnetmask*) - source address of the IP packet

Connection List

Home menu level: ***/ip proxy connections***

Description

This menu contains the list of current connections the proxy is serving

Property Description

dst-address (*read-only: IP address*) - IP address of to which data are passed via this proxy

protocol (*read-only: text*) - protocol name

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

rx-bytes (*read-only: integer*) - the amount of bytes received from the remote end

src-address (*read-only: IP address*) - IP address of the remote end of the connection

state (*read-only: connecting* | *idle* | *resolving* | *rx-body* | *rx-header* | *tx-body* | *tx-header*) - opened connection state

- **connecting** - establishing connection with server
- **idle** - waiting for next client to serve
- **resolving** - resolving server's DNS name
- **rx-body** - receiving HTTP body
- **rx-header** - receiving HTTP header; or waiting for next request from client
- **tx-body** - transmitting HTTP body
- **tx-header** - transmitting HTTP header

tx-bytes (*read-only: integer*) - the amount of bytes sent to the remote end

Cache Contents

Home menu level: ***/ip proxy cache-contents***

Description

This menu lists all the files stored in the cache

Property Description

file-size (*read-only: integer*) - size of the stored file

last-accessed (*read-only: date*) - date of the last access to the resource

last-accessed-time (*read-only: time*) - time of the last access to the resource

last-modified (*read-only: date*) - modification date

last-modified-time (*read-only: time*) - modification time

uri (*read-only: text*) - full resource name

Cache inserts

Home menu level: ***/ip proxy inserts***

Description

This menu shows statistics on objects stored in cache (cache inserts)

Property Description

denied (*read-only: integer*) - number of inserts denied by the caching list

errors (*read-only: integer*) - number of disk or other system-related errors

no-memory (*read-only: integer*) - number of objects not stored because there was not enough memory

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

successes (*read-only: integer*) - number of successfull cache inserts

too-large (*read-only: integer*) - number of objects too large to store

Cache Lookups

Home menu level: ***/ip proxy lookups***

Description

This menu shows statistics on objects read from cache (cache lookups)

Property Description

denied (*read-only: integer*) - number of requests denied by the access list

expired (*read-only: integer*) - number of requests found in cache, but expired, and, thus, requested from an external server

no-expiration-info (*read-only: integer*) - conditional request received for a page that does not have the information to compare the request with

non-cacheable (*read-only: integer*) - number of requests requested from the external servers unconditionally (as their caching is denied by the cache access list)

not-found (*read-only: integer*) - number of requests not found in the cache, and, thus, requested from an external server (or parent proxy if configured accordingly)

successes (*read-only: integer*) - number of requests found in the cache

Complementary Tools

Description

Web proxy has additional commands to handle non-system drive used for caching purposes and to recover the proxy from severe file system errors.

Command Description

check-drive - checks non-system cache drive for errors

clear-cache - deletes existing cache and creates new cache directories

format-drive - formats non-system cache drive and prepairs it for holding the cache

Transparent Mode

Description

Transparent proxy feature performs request caching invisibly to the end-user. This way the user does not notice that his connection is being processed by the proxy and therefore does not need to perform any additional configuration of the software he is using.

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

This feature may as well be combined with bridge to simplify deployment of web proxy in the existing infrastructure.

To enable the transparent mode, place a firewall rule in destination NAT, specifying which connections, *id est* traffic coming to which ports should be redirected to the proxy.

Notes

Only HTTP traffic is supported in transparent mode of the web proxy. HTTPS and FTP protocols are not going to work this way.

Example

To configure the router to transparently redirect all connections coming from **ether1** interface to port **80** to the web proxy listening on port **8080**, then add the following destination NAT rule:

```
[admin@MikroTik] > /ip firewall nat add in-interface=ether1 dst-port=80 \
\... protocol=tcp action=redirect to-ports=8080 chain=dstnat
[admin@MikroTik] > /ip firewall nat print
Flags: X - disabled, I - invalid, D - dynamic
 0   chain=dstnat protocol=tcp in-interface=ether1 dst-port=80 action=redirect
     to-ports=8080
[admin@MikroTik] >
```

Be aware, that you will not be able to access the router's web page after addition of the rule above unless you will change the port for the **www** service under **/ip service** submenu to a different value or explicitly exclude router's IP address from those to be matched, like:

```
 /ip firewall nat add in-interface=ether1 dst-port=80 \
\... protocol=tcp action=redirect to-ports=8080 chain=dstnat dst-address=!1.1.1.1/32
```

It is assumed that the router's address is **1.1.1.1/32**.

HTTP Methods

Description

OPTIONS

This method is a request of information about the communication options available on the chain between the client and the server identified by the **Request-URI**. The method allows the client to determine the options and (or) the requirements associated with a resource without initiating any resource retrieval

GET

This method retrieves whatever information identified by the **Request-URI**. If the **Request-URI** refers to a data processing process than the response to the **GET** method should contain data produced by the process, not the source code of the process procedure(-s), unless the source is the result of the process.

The **GET** method can become a *conditional* **GET** if the request message includes an **If-Modified-Since, If-Unmodified-Since, If-Match, If-None-Match**, or **If-Range** header field. The conditional **GET**

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.

method is used to reduce the network traffic specifying that the transfer of the entity should occur only under circumstances described by conditional header field(-s).

The **GET** method can become a *partial* **GET** if the request message includes a **Range** header field. The partial **GET** method intends to reduce unnecessary network usage by requesting only parts of entities without transferring data already held by client.

The response to a **GET** request is cacheable if and only if it meets the requirements for HTTP caching.

HEAD

This method shares all features of **GET** method except that the server must not return a message-body in the response. This retrieves the metainformation of the entity implied by the request which leads to a wide usage of it for testing hypertext links for validity, accessibility, and recent modification.

The response to a **HEAD** request may be cacheable in the way that the information contained in the response may be used to update previously cached entity identified by that **Request-URI**.

POST

This method requests that the origin server accept the entity enclosed in the request as a new subordinate of the resource identified by the **Request-URI**.

The actual action performed by the **POST** method is determined by the origin server and usually is **Request-URI** dependent.

Responses to **POST** method are not cacheable, unless the response includes appropriate **Cache-Control** or **Expires** header fields.

PUT

This method requests that the enclosed entity be stored under the supplied **Request-URI**. If another entity exists under specified **Request-URI**, the enclosed entity should be considered as updated (newer) version of that residing on the origin server. If the **Request-URI** is not pointing to an existing resource, the origin server should create a resource with that URI.

If the request passes through a cache and the **Request-URI** identifies one or more currently cached entities, those entries should be treated as stale. Responses to this method are not cacheable.

TRACE

This method invokes a remote, application-layer loop-back of the request message. The final recipient of the request should reflect the message received back to the client as the entity-body of a 200 (OK) response. The final recipient is either the origin server or the first proxy or gateway to receive a **Max-Forwards** value of **0** in the request. A **TRACE** request must not include an entity.

Responses to this method MUST NOT be cached.

Copyright 1999-2007, MikroTik. All rights reserved. Mikrotik, RouterOS and RouterBOARD are trademarks of Mikrotikls SIA.
Other trademarks and registred trademarks mentioned herein are properties of their respective owners.