Technical Guide

URL Filtering

Feature Overview and Configuration Guide

Introduction

URL filtering blocks all HTTP and HTTPS access to a list of websites or portions of web sites. You can specify a short list of websites to block (up to 1000 blacklist and 1000 whitelist rules), and/or subscribe to the blacklist service offered by a third-party provider.

- A **whitelist** is a list of URLs that are known to comply with organisational policies.
- A **blacklist** is a list of URLs that are known to violate organisational policies.
- A **third-party provider** is a subscription-based service that classifies websites among dozens of pre-defined categories of content that will not comply with some organisations' policies.
  If you subscribe to a third-party service, you can create additional blacklists to block extra URLs or whitelists to allow URLs that the service blocks.

URL Filtering provides a fast efficient (stream-based) method of blocking web traffic from locations that are known to be undesirable. It acts on a global basis and should be used when traffic is to be blocked for everyone on the blacklist, or allowed for selective URLs as configured in a whitelist.

This contrasts with Web Control, which has finer grained control as URLs are proxied and categorized and access to websites are controllable on a per-category and per-firewall entity basis. And since the Web Control service is proxy-based there is increased latency compared to the (stream-based) URL Filtering service.

It is possible to use Web Control and URL Filtering at the same time. Connections must be permitted by both URL Filtering and Web Control in order to be allowed through the device. A block action in either feature will cause a failure to load the web page.

Contents

Products and software version that apply to this guide

This guide applies to the following AlliedWare Plus™ products, running version **5.4.6** or later.

- AR3050S
- AR4050S
- AR2050V (user-configured lists only supported)
- AR2010V (user-configured lists only supported)

However, implementation varies between products. To see whether a product supports a feature or command, see the following documents:

- The product's Datasheet
- The AlliedWare Plus Datasheet ??
- The product's Command Reference

These documents are available from the above links on our website at alliedtelesis.com.

Version 5.4.7-1.x and later support:

- Logging of all URL requests
- URL filtering of HTTPS web sites using TLS SNI

Feature support may change in later software versions. For the latest information, see the above documents.

How Does URL Filtering Work?

URL filtering works by sniffing traffic as it traverses the AlliedWare Plus firewall and detecting the HTTP and HTTPS transactions that are taking place. These transactions are then processed, and when an HTTP Request is detected, the URL in question is compared against the whitelists (if any) and blacklists configured.

In AlliedWare Plus version 5.4.7-1 and later, the URL Filtering feature includes the ability to filter SSL-protected websites. For these HTTPS requests, the original URLs are encrypted, therefore they are not visible for processing. Instead the domain name specified in TLS SNI (Transport Layer Security Server Name Indication) for each HTTPS request is used as the URL for matching.

The SNI field is contained within the Client Hello message supplied during the TLS handshake when a client web browser first attempts to access a secure HTTPS server website. The SNI information is supplied in clear-text, and represents the domain part of the URL of the HTTPS request. The SNI field is used by secure web servers hosting multiple secure websites, and allows a secure web server with a single public IP address to host multiple websites. It allows the secure web server to supply the correct digital certificate containing the correct domain name(s) to the requesting web browser client, so that the negotiation of the encrypted connection to the website can proceed.

- If a whitelist match is found, the traffic will not be blocked (it will be logged if configured to do so).
- If a blacklist match is found, the request will be dropped (and logged if configured to do so)—it will not be forwarded to the destination.
- If neither whitelist nor blacklist matches are found, the traffic will not be blocked.
- Pattern checking stops as soon as a match is found. So if traffic matches any configured whitelist, then it will be allowed though the device. Or if traffic matches any configured blacklist then it will immediately be blocked. That same traffic will not be subsequently checked against additional whitelists or blacklists.

How to use URL Filtering

To use URL filtering, you can either use:

- a blacklist provided by a third-party
- custom lists (black/white)
- a combination of custom and third-party lists.

Creating custom lists

A custom list is an ASCII formatted text file containing zero or more single-line pattern matches.

For example, the content of a text file named ***blacklist-example.txt***, consisting of three patterns to match, (listed line-by-line) could look like this:

```
example.net/viruses/*
*/viruses/*
bad_url.com
```

URL pattern matches listed within the text file may take two forms:

- either a base domain, which will match all content of that domain, and all content of sub-domains:

```
example.com
```

- or a wild-card match, where an asterisk will match zero or more characters in a URL:

```
example.net/viruses/*
*/viruses/*
```

Once this list is available to the system (stored in Flash, USB, or on an SD card), the configuration to enable URL filtering is straight forward, as described below in the section "Configuring URL Filtering" on page 7.

Details of the content of custom lists

A custom list is an ASCII formatted text file containing zero or more single-line pattern matches. So far, we have looked at the general syntax of the entries in these files. Here we look in more detail at the rules governing the content of these files:

- There is no ordering or precedence for patterns in the file.
- Spaces in the pattern are not allowed.
- The wildcard, asterisk '*' can be used in the pattern to indicate a match on zero or more characters.
- If there are no '/' or '*' characters present, then all content of the domain is blocked.
- "Match everything" patterns are not allowed (e.g. '*' or '*/*').
- Empty or comment lines (starting with '#' or ';') are ignored.

- The ‘www.’ prefix should not be included in the pattern. However patterns and URLs are normalized before matching. More specifically:
  - The ‘www.’ prefix and authentication prefix ‘login:<password>@’ that may pre-pend a URL are automatically stripped from the URL before pattern matching.
  - Patterns are converted to lower case.
  - Only the domain name should be specified for blocking HTTPS traffic because TLS SNI contains only the domain name for the HTTPS request.

The table below describes how the pattern ***mysite.com/** is matched (Blocked URLs) or not matched (Non-blocked URLs) for a blacklist.

Table 1: A pattern matching example with explanations.

THIS PATTERN	BLOCKS THE URLS	NON-BLOCKED URLS
*mysite.com/	mysub.mysite.com www.mysite.com	mysub.mysite.com/mypage
Pattern matching explanations	**mysub.mysite.com** is a match (and is therefore blocked) because: ■ The wildcard, asterix ‘*’ matches the prepended text ‘mysub’ in the URL, and the remaining text in the URL matches the pattern. **www.mysite.com** is a match because: ■ The “www.” prefix is stripped off prior to matching, and the remaining text in the URL matches the pattern.	**mysub.mysite.com/mypage** is not a match (and is therefore non-blocked) because: ■ The text ‘mypage’ in the URL is not part of the pattern.

The following table lists a series of blacklisted ‘domain and string pattern’ match criteria, and examples of URLs that would or would not be matched by these criteria.

Table 2: Blacklisted domain and string pattern match criteria

PATTERN	BLOCKED URLS	NON-BLOCKED URLS
com	www.mydotcomurl.com	myausurl.com.au
com.au	www.myausurl.com.au:8080/file.txt	mydotcomurl.com
ru	myrussian.pp.ru	myfakerussian.ru.org
z	faz.com auzi.id.au zulu.com me.kiwi.nz fish.com/folder1/file.gz www.google.co.nz/search?client=ubuntu&channel=fs&q=ziare&ie=utf-8&oe=utf-8&gfe_rd=cr&ei=ZfKWVqgtk5PABN6YtqgD	
*mysite.com/	mysub.mysite.com www.mysite.com	mysub.mysite.com/mypage
mysite.com/*	www.mysite.com/mypage.html www.mysite.com/ www.mysite.com	mysub.mysite.com/mypage www.mysite.com.au
mysite.com	mypage.mysite.com.au mysite.com.au www.mysite.com mypage.mysite.com/folder/file.txt somescript.sc?mysite.com.au	
mysite.com/*/filename*.exe	www.mysite.com/folder/filename.exe mysite.com/folder/filename-bad.exe mysite.com/subdomain/folder2/folder3.html/abcd/filename.exe mysite.com/folder/filename-bad/file.exe	www.myurl.mysite.com/subdomain/filename*.exe search-engine.com/search?q=mysite.com/folder/filename.exe mysite.com/filename.exe mysite.com/subdomain/file.exe mysite.com/subdomain/filename.exe1 mysite.com/subdomain/filename.html
192.168.1*/abcd-efgh/subdomain/*	192.168.10.com/abcd-efgh/subdomain/filename.exe 192.168.1.10/abcd-efgh/subdomain/filename.exe	192.168.2.10/abcd-efgh/subdomain/filename.exe 192.168.1.10/abcd-efgh/filename.exe 192.168.1.10/abcd-efgh/subdomain2/filename.exe

Limits

URL filtering is limited to 1000 custom whitelist and 1000 custom blacklist rules, spread over any number of list files.

Configuring URL Filtering

URL filtering is turned on by configuring a whitelist that uses a custom file, a blacklist that uses a custom file, or blacklisting that uses a third-party service.

1. To add a **whitelist** that uses a custom file (that is stored on USB, for example) and then enable URL filtering, use the commands:

```
awplus#configure terminal
awplus(config)#url-filter
awplus(config-url-filter)#whitelist usb:/my_whitelist.txt
awplus(config-url-filter)#protect
```

2. To add a **blacklist** that uses a custom file (that is stored on Flash, for example) and then enable URL filtering, use the commands:

```
awplus#configure terminal
awplus(config)#url-filter
awplus(config-url-filter)#blacklist flash:/blacklist-example.txt
awplus(config-url-filter)#protect
```

3. To add a blacklist provided by **Kaspersky** and then enable URL filtering, use the commands:

```
awplus#configure terminal
awplus(config)#url-filter
awplus(config-url-filter)#provider kaspersky
awplus(config-url-filter)#protect
```

- To check that Kaspersky is active, enter the command **show url-filter:**

```
awplus#show url-filter
Status:       Enabled (Loading)
Provider:     Kaspersky
 Status:               Enabled
 Resource version:     not set
 Update interval:      1 hour
 Blacklist entries:    -
Custom blacklists   Entries
 blacklist-example.txt  3
Custom whitelists   Entries
```

- Invalid entries in URL filter lists are ignored (not loaded).
- Expiry of the Kaspersky URL Filtering Subscription License will cause URL filtering to reload without a Kaspersky blacklist.

Using multiple whitelists and blacklists

The AlliedWare Plus firewalls support pattern checking against multiple whitelists and multiple blacklists.

Multiple custom whitelists or blacklists can be configured and checked as follows:

```
awplus(config)#url-filter
awplus(config-url-filter)#blacklist blacklist1.txt
awplus(config-url-filter)#blacklist blacklist2.txt
awplus(config-url-filter)#blacklist blacklist3.txt
awplus(config-url-filter)#whitelist whitelist1.txt
awplus(config-url-filter)#whitelist whitelist2.txt
awplus(config-url-filter)#whitelist whitelist3.txt
awplus(config-url-filter)#protect
```

You can check the configuration using the **show url-filter**, **show running-config url-filter** and **dir** commands:

```
awplus#show url-filter
Status:      Enabled (Active)
Provider:    not set
Custom blacklists   Entries
 blacklist1.txt       18
 blacklist2.txt       23
 blacklist3.txt       39
Custom whitelists   Entries
 whitelist1.txt       11
 whitelist2.txt       26
 whitelist3.txt       33
```

```
awplus#show running-config url-filter
url-filter
 blacklist blacklist1.txt
 blacklist blacklist2.txt
 blacklist blacklist3.txt
 whitelist whitelist1.txt
 whitelist whitelist2.txt
 whitelist whitelist3.txt
 protect
!
```

```
awplus#dir
       107 -rw- May 11 2016 04:52:44  whitelist1.txt
       229 -rw- May 11 2016 04:52:39  whitelist2.txt
       318 -rw- May 11 2016 04:52:32  whitelist3.txt
       372 -rw- May 11 2016 04:51:50  blacklist3.txt
       202 -rw- May 11 2016 04:51:38  blacklist2.txt
       170 -rw- May 11 2016 04:51:31  blacklist1.txt
```

Rules for processing lists

The order of processing of lists is:

- First—whitelists
- Second—custom blacklists
- Third—third-party-provided blacklists

The matching logic is that as soon as a URL matches an entry in a list that it is being compared against, then comparing stops and the relevant action (allow, if the match occurs in a whitelist, or deny if the match occurs in a blacklist) is taken.

Because whitelist matching precedes blacklist matching, you can use custom whitelists to override any corresponding blacklist entries. An HTTP or HTTPS request that has a URL matching an entry in a whitelist will be permitted immediately, and the URL will not be matched against any blacklists.

So, if websites you actually want to access are being blocked by the third-party blacklist, or some subsection of an otherwise dangerous site is desirable, a whitelist may be created.

Example For this example, the ***example.net/viruses/research*** folder contains information that is needed within the otherwise completely blocked site.

This can be allowed by creating a whitelist file named ‘whitelist-example.txt’ in Flash memory, with the contents:

```
example.net/viruses/research/*
```

And configuring it as follows:

```
awplus#configure terminal
awplus(config)#url-filter
awplus(config-url-filter)#whitelist whitelist-example.txt
awplus(config-url-filter)#protect
```

This whitelist will be processed prior to the blacklist, and will allow matching traffic through.

Updating lists

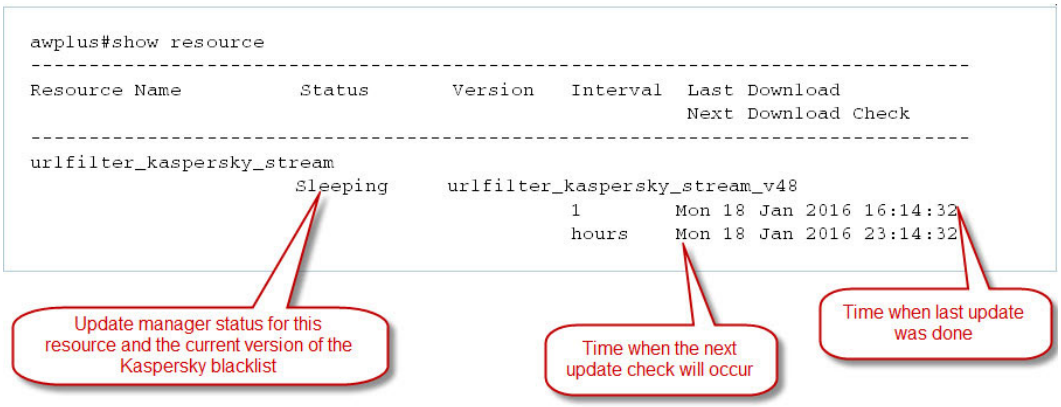
Updating the Kaspersky blacklist

When subscribed to the Kaspersky URL Filter service, updates to the Kaspersky blacklist will be made available. By default URL filtering checks for updates to the Kaspersky blacklist every hour.

You can configure the update interval via the **update-interval** command in **url-filter** configuration mode. The update process is managed by the Update Manager utility.

You can see the update status in two show command outputs: **show url-filter** and **show resource**.

```
awplus#show url-filter
Status:       Enabled (Loading)
Provider:     Kaspersky
 Status:                Enabled
 Resource version:      urlfilter_kaspersky_stream_v48
 Update interval:       1 hour
 Blacklist entries:     63457
...
```



When the Update Manager finds a new version is available, it downloads and instructs URL Filter to start using the new blacklist. An update check can be manually initiated with the **update urlfilter_kaspersky_stream now** or **update all now** commands.

Updating a user-defined blacklist or whitelist

You can modify blacklist and whitelist files that you have created. Once you have completed all the desired changes, use the **url-filter reload custom-lists** command to reload the modified files.

When a new blacklist or whitelist is configured and URL filter is already enabled, it automatically starts using the new file.

Monitoring URL Filtering

The **show url-filter** command displays a summary of the state of URL filtering, including the provider state, and counts of entries in each provided list. Any lists that contain too many entries to load will be noted here.

```
awplus#show url-filter
Status:      Enabled (Active)
Provider:    Kaspersky
 Status:              Enabled
 Resource version:    not set
 Update interval:     1 hour
 Blacklist entries:   -
Custom blacklists   Entries
 blacklist-example.txt  3
Custom whitelists   Entries
 whitelist-example.txt  1
```

URL filtering logging

By default, URL Filtering messages are generated when there are:

- Blacklist and whitelist hits—logged at severity **info (6)** level.
- Invalid match criteria, detected while loading third party and custom blacklist and whitelist files—logged at **err (3)** level.
- Missing configured custom blacklist and/or whitelist files, while starting/restarting the feature—logged at **warning (4)** level.

From AlliedWare Plus version 5.4.7-1.x, you can turn on additional URL request logging to log **all** URL requests, including permitted requests. Use the following commands:

```
awplus(config)# url-filter
awplus(config-url-filter)# log url-requests
```

Log messages for blacklist or whitelist hits include information in the following format:

```
<action> URLFILTER: [URL:<url>] <protocol> <source-ip>:<source-port> ->
<dest-ip>:<dest-port>
```

Table 3: URL Filtering log message elements

Message element	Description
<action>	Which action is applied; [ALERT], [DROP] or [http].
<url>	The requested URL if the flow is HTTP.
<protocol>	The protocol e.g., SMTP, HTTP, TCP, ICMP.
<source-ip>:<source-port>	The source IP address and source port for the packet.
<dest-ip>:<dest-port>	The destination IP address and source port for the packet.

Output 1: Example URL filtering log message for a dropped URL request

```
2016 Nov 17 02:02:21 local5.info awplus IPS[2039]: [Drop] URLFILTER: URL:http:/
kdskspb.ru/ [http] 192.168.1.1:58272 -> 172.16.1.2:80
```

Output 2: Example URL filtering log message for a permitted URL request when **log url-requests** is configured

```
2017 Apr 12 03:47:21 local5.info awplus IPS[3885]: [Http] URL:http://172.16.1.2/
192.168.1.1:53698 -> 172.16.1.2:80
```

For more information about logging, please refer to the Logging Feature Overview and Configuration Guide.

C613-22081-00 REV D

Allied Telesis NETWORK SMARTER

North America Headquarters | 19800 North Creek Parkway | Suite 100 | Bothell | WA 98011 | USA | T: +1 800 424 4284 | F: +1 425 481 3895
Asia-Pacific Headquarters | 11 Tai Seng Link | Singapore | 534182 | T: +65 6383 3832 | F: +65 6383 3830
EMEA & CSA Operations | Incheonweg 7 | 1437 EK Rozenburg | The Netherlands | T: +31 20 7950020 | F: +31 20 7950021

alliedtelesis.com