Forcepoint

Web Security Cloud

Configuring Full Traffic Logging

Revision A

Contents

Configuring Full Traffic Logging

Administrators using Forcepoint Web Security Cloud have the option to download full web traffic logs for retention and analysis.

At this time, the full traffic logging feature is a limited availability feature. Contact Forcepoint Technical Support to enable the feature for your organization.

Once you have enabled traffic logging in the Forcepoint Cloud Security Gateway Portal, also known as the cloud portal, you can schedule a regular process to download the logs and save them to a location of your choice. Logs are retained in the cloud service for 14 days.

Important

Full traffic logging is an add-on for Forcepoint Web Security Cloud, and is separate from standard web reporting. Standard reporting data is retained for 90 days and can be accessed through standard and custom reports; full traffic logs, once enabled, are retained for 14 days and are accessible through download only.

Once the feature has been enabled for your account, follow the steps in this paper to set up and use full traffic logging. See:

1) *Setting up full logging* provides step-by-step instructions for setting up full traffic logging in the cloud portal, accessing the log files, and understanding the sample download script.

2) *Downloading log files* describes the issues you must be aware of when downloading the logs, and how to schedule the download process.

3) *File format definition for full traffic logging* describes the contents of a log file, with examples.

If you encounter unexpected issues while setting up full traffic logging, see *Troubleshooting full traffic logging*.

Related concepts

Troubleshooting full traffic logging on page 15

Related tasks

Setting up full logging on page 3

Related reference

Downloading log files on page 7
File format definition for full traffic logging on page 10

Setting up full logging

Before you begin

Even though the correct permission (Log export) may be visible in your Account settings, the feature is not available by default. To make it available in your account, contact Support.

Once the feature is available, to set up full traffic logging in the cloud portal:

Steps

1) *Create a new administrator contact*.
   We strongly recommend that the log download process has its own user name and password to gain access to the Forcepoint Web Security Cloud service. This keeps the process separate from your other administration tasks and enables you to establish longer password expiration policies.

2) *Enable full traffic logging*.
   Full traffic logging can be enabled for your whole account or for specific policies.

3) *Set up a download script*.

Related tasks

Create a new administrator contact on page 3
Enable full traffic logging on page 4
Set up a download script on page 6

Create a new administrator contact

To create the new contact:

Steps

1) In the cloud portal, on the main toolbar, click **Account**, then select **Contacts**.

2) Under the Contacts list, click **Add**.

3) Enter identifying information for the new contact in the **First name** and **Surname** fields. For example, “Traffic” and “Logging.”

4) Click **Submit**.

5) Click the link provided to supply a **User name** for the account.

6) Enter a password for the contact. It must conform to the password policy on the main Contacts page.

7) Enter a password expiration date for the contact. To avoid having to regularly update it, this should be different than the regular account settings; it should span a longer period. The maximum period is 365 days.

8) Under **Account Permissions**, check the **Log Export** box, and any other permissions you want to give this user. You can act as an administrator from this logon.

Note

If you give this contact only the **Log Export** permission and nothing else, the user name and password cannot be used to log on to the cloud portal. The **View Reports** permission is the minimum permission a user needs to be able to log on.

9) Click **Submit**.

Enable full traffic logging

To enable log retention for your account:

Steps

1) In the cloud portal, on the main toolbar, click **Web**, then select **Full Traffic Logging** (under **Settings**).

2) Click **Edit**.

3) Mark the **Enable full Web traffic logging** checkbox.

   The text on this page states the conditions for using full traffic logging—namely, all data is retained for only 14 days, and if you do not download any files for a period of 14 days, full traffic logging is automatically disabled. For more information, see *Troubleshooting full traffic logging*.

   This page also contains a link to a sample script that you can use to download and store your log files. You can edit this script to suit your needs. For more information, see *Set up a download script*.

4) Click **Submit**.

Related concepts

Troubleshooting full traffic logging on page 15

Related tasks

Set up a download script on page 6

Change the log retention

By default, all web policies have the logging setting that you define at the account level. If you want to change the log retention for a particular policy:

Steps

1) On the main toolbar, click **Web**, then select **Policies** (under **Policy Management**).

2) On the Policies page, click the name of the policy you want to configure.

3) On the General tab for the selected policy, click **Edit**.

4) Under Full Traffic Logging, change the selection in the drop-down list from **Use policy-wide default** to either **Enabled** or **Disabled**. This overrides the account- level setting.

5) Click **Submit**.

Next steps

You can view the logs available for your account by going to https://sync-web.mailcontrol.com/hosted/logs and logging on with the user name and password that you set up with the Log Export permission. If you access this site immediately after you have set up full traffic logging, you will see only an empty XML script, but once Forcepoint Web Security Cloud has started to retain your logs, the page will show all available log files for download.

Each file name has the following format:

hosted_<SourceID>_<AccountID>_<ClusterIP>_<Version>_<Epoch>_<SequenceNo>.gz

Elements of the string

The elements of the string are defined as follows:

Element	Description
SourceID	Internal string to identify where in the cloud service the log data was created.
AccountID	The cloud service internal identifier for your account.
ClusterIP	The IP address for the cluster that processed your web requests.
Version	Current version number for the log file format.
Epoch	The UNIX time for each log file, representing the time intervals between each generated log.
SequenceNo	If your logs exceed a certain size, the cloud service will deliver multiple log files for the configured time interval. This number identifies each log in the given period.

For example, for log files from a cluster with the IP address 10.12.14.16 generated every 10 minutes, you might see the following:

hosted_xxxx1_1234_10.12.14.16_1_1236779400_1.gz

hosted_xxxx2_1234_10.12.14.16_1_1236780000_1.gz

hosted_xxxx1_1234_10.12.14.16_1_1236780000_2.gz

hosted_xxxx3_1234_10.12.14.16_1_1236780000_3.gz

hosted_xxxx1_1234_10.12.14.16_1_1236780600_1.gz

Set up a download script

To download the log files and save them to a location of your choice, you can either use the sample Perl script or create a script of your own. To save the sample script to your network:

Steps

1) On the main cloud portal toolbar, click **Web**, then select **Full Traffic Logging** (under **Settings**).

2) Click **Edit**.

3) Click the sample script link, and save the file to a location of your choice. By default the file is named full_traffic_log_download.pl.

Warning

Forcepoint provides the sample log download script as a convenience to its customers, but does not provide support for customization and will not be responsible for any problems that may arise from editing the script.

After downloading a script

The script can be run on Windows or Linux, and does the following:

- Connects to the cloud service using the URL specified in the script
- Optionally reports the log files available for download
- Downloads the available log files to a location of your choice, or by default to the directory where the script is located
- Optionally checks the MD5 hash of each downloaded file to verify the file's integrity before deletion from the server
- Uses the HTTP DELETE method to request that the cloud service delete the downloaded files

Note

Running the script on Windows requires ActivePerl, which you can download from https://www.activestate.com/activeperl/downloads .

To see the ActivePerl modules required for running the script, open the script in a text editor. The modules are listed at the beginning of the file.

If you customize the sample script or choose to write your own script, you must always include the DELETE method to remove the downloaded files from the server. This is because files are only retained for 14 days, and any files that have not been deleted after 7 days will trigger a warning email. For more information, see *Troubleshooting full traffic logging*.

Related concepts

Troubleshooting full traffic logging on page 15

Downloading log files

To download log data when it is available, run the script that you have set up. If you are using the provided sample script, the available parameters to use with the script are described below.

Some parameters have a short form (for example, **-v**) and a long form (for example, **--verbose**). For these parameters, both options are listed.

Parameter	Description
-u *<username>* --username	Mandatory. Defines the logon user name for connecting to the cloud service. This must be an administrator contact with Log Export permissions. For example: *-u FTL_user@example.com*
-p *<password>* --password	Mandatory. This is the password for the specified user name. For example: *-p Ft2016Logs*
-v *--verbose*	Optional. Runs the script in verbose mode, which displays progress messages. Verbose mode provides feedback on the script's progress, for example: ■ Downloading filelist from <host name> as <user name> ■ No files available to download ■ Downloading <file> to <file name location>
-h *<hostname>* --host	Optional. Defines the host name to connect to. This is specified in the script by default, so you would only need this option if you have edited the script to remove it, or if you have been given a different URL to connect to. For example: *-h https://sync-web.mailcontrol.com*
-d *<file path>* --destination	Optional. Defines the destination directory for the downloaded log files. If not specified, the files are downloaded into your current working directory. For example: `-d /cloudweb/logs`
-m *--md5sum*	Optional. Checks the md5sum of each downloaded file. The MD5 hash is commonly used to verify the integrity of files (i.e. to verify that a file has not changed as a result of file transfer or disk error), and can therefore be used to check the files before they are deleted from the server.

Parameter	Description
-l *--list-only*	Optional. Displays a list of available log files without downloading them.
--proxy *<proxy details>*	Optional. Specifies an HTTP proxy to use if you are having difficulty connecting to the cloud service. The proxy must be in the form http://*username:password@host:port* For example: --proxy http:// jsmith:Abc123@proxy_server:80
--format= *<format>*	Optional. Creates a new data file containing the original downloaded data rewritten in the desired format. The new file's name has the relevant data format as a suffix. Note that when this parameter is used, by default the original *.gz file from the source server is not saved to the destination directory. Valid data formats are: csv: Comma Separated Values cef: ArcSight Common Event Format wc3: WC3 Extended Log file Format (http://www.w3.org/TR/WD-logfile.html) For example: --format=csv
--keepgz	Optional. Use in conjunction with the format parameter to download and keep a copy of the *.gz data file in the destination directory. This overrides the default behavior of the **format** parameter. For example: --format=csv --keepgz
--delete	Optional. Deletes the original data file from the source server following download. The default option is to delete the file from the server.
--nodelete	Optional.Keeps the original data file on the source server after download. This parameter is provided for testing purposes while configuring the **format** parameter described above, enabling you to download a file in different formats. Note that files are still only retained for 14 days, and you will still receive a warning after 7 days if a downloaded data file remains on the server.
--max_batch_size	Optional. Specifies the maximum number of files to download. When set, each time the script is run, the configured number of files are downloaded, starting with the newest files.
--man	Optional. Displays the list of parameters with their descriptions.

Parameter	Description
--help	Optional. Displays a brief description of the program's purpose.

Due to the volume of data, we recommend importing the information into a database to analyze the downloaded log files. For more information about the downloaded data, see *File format definition for full traffic logging*.

Related reference
File format definition for full traffic logging on page 10

Scheduling log file download

Once you have run an initial download and determined the parameters you want to use in your script, set up a scheduled service to run automatic downloads.

We recommend that you download the log files at least once a day. To avoid periods of high network traffic, select a random time for the download (for example, somewhere between 10 and 50 minutes past the hour).

Scheduling on Windows

Before scheduling downloads from the cloud service, make sure that the Windows Task Scheduler service is started. To check this:

Steps

1) Open the Windows **Services** tool (**Start** > **Control Panel** > **Administrative Tools** > **Services or Server Manager** > **Tools** > **Services**).

2) Scroll down to **Task Scheduler**.
   - If the status is Started, you need do nothing.
   - Otherwise, click **Start** or **Resume** to start the service.

Procedure to schedule the log file download

To schedule the log file download:

Steps

1) Open the Windows **Scheduled Tasks** tool (**Start** > **Control Panel** > **Scheduled Tasks or Server Manager** > **Tools** > **Scheduled Tasks**).

2) Double-click **Add Scheduled Task**.

3) Work through the Scheduled Task Wizard as follows:

   a) Browse to the location where you have stored your script.

   b) Select how often to perform the task (daily is advisable).

   c) Select a time to start the task, and the start date.

   d) Enter your network user name and password (**not** the user name and password you set up in the Cloud TRITON Manager).

   e) Mark the **Open advanced properties for this task** checkbox, then click **Finish**.

4) On the Task tab, add the **-u**, **-p**, and **-d** parameters to the end of the **Run** field, before the closing quotes, as well as any other parameters you want to use.

   The **Run** field might look similar to this:

   ```
   "\\server\users\jsmith\hosted_logs\full_traffic_log_download.pl -u FTL_user@example.com -p Ft2010Logs -d /hostedweb/logs"
   ```

5) Click **OK**.

Scheduling on Linux

Create a cron job to schedule your script to run at the times you want. For more information in Linux, see **man cron** and **man crontab**.

File format definition for full traffic logging

The log files downloaded from the cloud service are in JavaScript Object Notation (JSON) format. For more information about JSON, see http://www.json.org/.

Each log file contains multiple lines, with one request per line. Each line is enclosed in square brackets.

The following table describes the fields that comprise each request.

Field	Description
DateAndTime	The time that a request occurred on the proxy, in seconds in UNIX time.
AccountID	The Forcepoint Web Security Cloud internal identifier for your account.
UserID	The web user's ID, usually their email address.
ClientIP	The client's external Internet IP address, shown in integer format. See *Converting integer IP addresses to dot-decimal IPv4 format*.
RequestCount	The number of requests for a particular site. This will default to 1 per log entry.

Field	Description
RequestSize	Size of the request in bytes.
ResponseSize	Size of the response in bytes.
Disposition	The disposition code of the request. For an explanation of the codes, see *Disposition codes*.
Categories	A comma-separated list of category IDs. To see how the ID numbers relate to category names, go to https://sync-web.mailcontrol.com/hosted/categories?version=2. (Note that this URL is for logs generated with 2015 Release 1 and later. To see ID numbers and category names for logs generated prior to that release, go to https://sync-web.mailcontrol.com/hosted/categories.)
Protocol	The protocol used in the request (for example HTTP, HTTPS, or FTP)
Port	The port number used for the request.
DestinationIP	The IP of the requested address, shown in integer format. See *Converting integer IP addresses to dot-decimal IPv4 format*.
URI	The full URL of the page requested by the user.
AnalyticID	Defines the analytic applied to the request. Can be one of the following: ■ 1, 2 - Real-Time Security Scanning (RTSS) ■ 4, 5, 6 - Advanced Detection (AD) ■ 10 - Antivirus (AE) ■ 11 - Real-Time Classification (RTC) ■ 13 - Malicious iFrame Detection (MIDE) ■ 14 - Malicious PDF Detection (SPIE) ■ 15 - Advanced Secure Hash (ASH) ■ 18 - Meta-analytic Detection (ICE)
ReasonCode	The reason code assigned to the request. For an explanation of the codes, see *Reason codes*.
ContentStripping	This field is blank in this version of the log file.
ReasonString	This is an internal signature ID string.
FileType	One of the following groups: 'unknown', 'text', 'executable', 'image', 'multimedia', 'document', 'suspicious', 'archive', 'ria', 'mime'.
PolicyName	Name of the policy used to filter the request.
ContentType	Content-Type of the response. The default value is an empty string.
RemoteHost	The host name of the origin server.

Field	Description
Method	HTTP method used in the request.
ProxyTime	The total delay, in milliseconds, due to filtering the transaction through the proxy.
OriginTime	The time taken, in milliseconds, to receive the request from the origin server.
ResponseTime	The total response time for the transaction, in milliseconds.

Related reference

Converting integer IP addresses to dot-decimal IPv4 format on page 12
Disposition codes on page 12
Reason codes on page 13

Converting integer IP addresses to dot-decimal IPv4 format

Note

IP addresses in the ClientIP field are exported in integer format. To convert integer IP addresses to the dot-decimal IPv4 address format, you can use the Excel formula shown in the table below. This example assumes an integer IP address is in cell A1 of your Excel spreadsheet.

Example integer IP	Formula to convert integer IP in cell A1	Dot-decimal IPv4 address
1433608197	=HEX2DEC(MID(DEC2HEX(A1,8),1,2))&"."&HEX 2DEC(MID(DEC2HEX(A1,8),3,2))&"."&HEX2DEC(MID(DEC2HEX(A1,8),5,2))&"."&HEX2DEC(MID(D EC2HEX(A1,8),7,2))	85.115.32.5

Disposition codes

The following table explains the meaning of the disposition codes used in the log files.

Code	Description
2	Page blocked
3	Page filtered and permitted
4	Request filtering refused to service the request
6	Could not connect to requested site
9	Blocked, and user access to service disabled
12	Continue/confirm request

Code	Description
13	Quota request
14	Request permitted without filtering

Reason codes

The following table explains the meaning of the reason codes used in the log files.

ID	Analytic	Name
1		Generic
1	Real-time security scanning	Generic
2	Real-time security scanning	Suspicious
3	Real-time security scanning	Exploit
4	Real-time security scanning	Redirection
5	Real-time security scanning	Obfuscation
6	Real-time security scanning	Evasion
7	Real-time security scanning	Counterfeit
8	Real-time security scanning	Spam
9	Real-time security scanning	Hijacked
10	Real-time security scanning	Defaced
11	Real-time security scanning	Tools
12	Real-time security scanning	Infostealer
13	Real-time security scanning	Backchannel traffic
14	Real-time security scanning	Remote control
15	Real-time security scanning	Installer
16	Advanced Detection	Malicious Packed
17	Advanced Detection	Generic Malicious
18	Advanced Detection	Trojan
19	Advanced Detection	Virus
20	Advanced Detection	Worm
21	Advanced Detection	Infected
22	Advanced Detection	Adware
50	Advanced Detection	Zipbomb
300	Malicious PDF Detection	Suspicious document
301	Malicious PDF Detection	Suspicious uncategorized document

ID	Analytic	Name
302	Malicious PDF Detection	Document with active content
400	Malicious iFrame Detection	Malicious iFrame detection
700	Advanced Secure Hash	Generic
800	Meta-Analytic Detection	Generic
900	AppID	Generic
10001	Antivirus	Virus
10002	Antivirus	Adware
10003	Antivirus	Application
10004	Antivirus	Backdoor
10005	Antivirus	Bomb
10006	Antivirus	BootVirus
10007	Antivirus	Denial
10008	Antivirus	Dialer
10009	Antivirus	Downloader
10010	Antivirus	Exploit
10011	Antivirus	Intended
10012	Antivirus	Joke
10013	Antivirus	Macro
10014	Antivirus	MassMailer
10015	Antivirus	MisDisinfection
10016	Antivirus	NetWorm
10017	Antivirus	P2Worm
10018	Antivirus	Proxy
10019	Antivirus	PasswordStealer
10020	Antivirus	Remote
10021	Antivirus	Risk
10022	Antivirus	Spyware
10023	Antivirus	Tool
10024	Antivirus	Trojan
10025	Antivirus	HiddenProcess
10026	Antivirus	Injected Code

Troubleshooting full traffic logging

Your download script attempts to connect to the cloud service to download full traffic logs at an interval that you configure. If your script is unable to make the connection, or if it is unable to retrieve the log files after connecting, the following problems may occur:

- The cloud service stores log files for only 14 days. After that period, the files are deleted, and cannot be recovered. When this occurs, your organization is no longer able to access and analyze web activity recorded in those logs.
- Depending on the volume of Internet activity that your organization sends through the cloud service, log files may grow quickly. If your script is unable to download log files for a day or more, the bandwidth required to download the files and the disk space required to store them may be substantial.

To address this issue:

- Check that your scheduling service (Windows Task Scheduler, or crontab on Linux) is running. If you are using Windows Task Scheduler, check that it is using your most recent network password to run the task.
- Your script may be prevented from accessing the cloud service due to network problems, either affecting Internet or internal network connections. Use a browser or the **ping** utility to verify that the machine running the script can connect to the Internet.
- If the script is connecting to the cloud service but cannot retrieve log records, verify that there is not a problem with the cloud service. Check the administrative email address associated with your full traffic logging account.
- Check that your cloud service password has not expired.

If you do not download traffic logs for a period of 7 days, a notification email is sent to all administrative contacts with Log Export permission enabled, and all policy administrators where full traffic logging is enabled for the policy. The email warns that logging will be disabled if you do not download logs for 14 days. Further notifications are sent after 10 and 13 days, and after 14 days you will be notified that full traffic logging has been deactivated and traffic logs are no longer being generated for your account.

Published 08 August 2024