proofpoint.™

Proofpoint on Demand (PoD) Log API

November 2020

Revision C

Contents

PoD Log API

The PoD Log service is a webservice for Proofpoint on Demand customers that offers a real-time email processing log feed for use with Security Information and Event Management (SIEM) solutions. This webservice uses the secure WebSocket (WSS) protocol to stream logs to supporting solutions.

This log feed can be used to identify mail and message filtering events.

Connection Notes

The PoD Log API does not allow use of the same token for more than one session at the same time. If you need to open more than one simultaneous connection to receive the same type of data, additional token(s) must be requested.

When the connection between the client and the service is dropped and restored within one hour, the data will be sent from the moment of time when the previous session had dropped, so there is no need to perform any additional action from the client side.

In the case where the client was connected to the PoD Log service and disconnected for more than one hour, after a new session is established, the client will start receiving the accumulated data starting from the last one hour of the new session.

For example:

The previous session was terminated at 1PM EST on 11/17/2020 and at 3PM EST 11/17/2020 a new connection was established. In this case the client will start receiving "real-time" data from 2PM EST 11/17/2020. To backfill the gap and retrieve the data from 1PM EST to 2PM EST, a separate request to the archive must be made. To do this, the `sinceTime` parameter should be set as 2020-11-17T14:00:00-0005 and the `toTime` as 2020-11-17T14:59:59-0005.

API Features

Endpoint

The PoD Logging Service production endpoint is

```
wss://logstream.proofpoint.com/
```

The API signature is

```
/v1/stream?cid={clusterId}&type=[message|maillog]&sinceTime={sinceTime}
&toTime={toTime}
```

The `sinceTime` and `toTime` parameters must be specified to request the archived data.

The date format is `YYY-MM-DDTHH:SS-0000` where `0000` is the time zone.

Correct zones are shown here:

EST = 0500	EDT = 0400
CST= 0600	CDT = 0500
MST = 0700	MDT = 0600
PST = 0800	PDT = 0700

Part	Required	Type	Example	Default	Description
type	yes	string	`message` `maillog`	N/A	Valid values are “message” or “maillog”

sinceTime	no	string	`2018-01-25T02:37:40-0800`
			`2018-01-25T02:37:40.000-0800`
				N/A	Start time to begin streaming log data, in ISO8601 format, which includes timezone information. Data timestamp is specific to the millisecond. Is used only if the specified timestamp is older than at least one hour from now. Rounds *down* to the nearest hour.

Part	Required	Type	Example	Default	Description

toTime	no	string	`2018-02-25T02:37:40-0800`
			`2018-02-25T02:37:40.000-0800`
				N/A	End time to stop streaming log data, in ISO 8601 format. Data timestamp is specific to the millisecond. Defaults to Now when the `sinceTime` is defined. If specified, must be greater than `sinceTime`. Rounds *up* to the nearest hour.

cid	yes	string	*customer hosted*	N/A	The Cluster ID must be a legal user group string. This is required for server authentication purposes.

Note: If a time is not specified, it means “now.” The limit to data availability is 30 days. You can specify a query going back in time 30 days from the present time. The archived data granularity is one (1) hour, not a minute or second. The service rounds *down* the `sinceTime` and rounds *up* the `toTime` parameter values to the nearest hour.

For example, `sinceTime=2018-01-25T14:12:34-0800` will be rounded *down* to `2018-01-25T14:00:00-0800` and `toTime=2018-01-25T14:31:23-0800` will be rounded *up* to `2018-01-25T15:00:0800`.

Fields

The JSON schema format is used to describe each field. This service supports only JSON.

Authentication

The authorization header must be set as part of the request to authenticate and be authorized to stream log data.

Proofpoint will provide the token and credentials to connect to the webservice.

Required header:

```
Authorization: Bearer <token>
```

The token value is uniquely generated and provided by Proofpoint for a customer cluster to authenticate with the service. The service uses JSON Web Token (JWT) to communicate the client identity to the service.

Signing Key

This is your CLUSTER_ID assigned by Proofpoint. The CLUSTER_ID is displayed in the upper-right corner of the management interface next to the release number.

Examples for Testing Streaming Requests

This section contains examples to test connectivity between your system and the Proofpoint PoD log service. In the examples `sinceTime` is optional to stream historical data. If `sinceTime` is not specified, the server will stream data in real time.

Example for `curl` command to receive uncompressed data:

```
curl -i --no-buffer -H "Connection: Upgrade" -H "Upgrade: websocket" -H
"Host: logstream.proofpoint.com:443" -H "Authorization: Bearer
<ACCESS_TOKEN>" -H "Sec-WebSocket-Key: SGVsbG8sIHdvcmxkIQ==" -H "Sec-
WebSocket-Version: 13"
"https://logstream.proofpoint.com:443/v1/stream?cid=<CLUSTER_ID>&type=m
essage&sinceTime=2018-08-31T00:00:00-0800"
```

Example for `curl` command to request a data stream compressed by the Deflate algorithm:

```
curl -i --no-buffer -H "Connection: Upgrade" -H "Upgrade: websocket" -H
"Host: logstream.proofpoint.com:443" -H "Sec-WebSocket-Extensions:
permessage-deflate; client_no_context_takeover;
server_no_context_takeover" -H "Authorization: Bearer <ACCESS_TOKEN>" -
H "Sec-WebSocket-Key: SGVsbG8sIHdvcmxkIQ==" -H "Sec-WebSocket-Version:
13"
"https://logstream.proofpoint.com:443/v1/stream?cid=<CLUSTER_ID>&type=m
essage&sinceTime=2018-08-31T00:00:00-0800"
```

> **Note:** The PoD Log API service supports only the `permessage-deflate`, `client_no_context_takeover`, and `server_no_context_takeover` extensions. Refer to RFC7692 for a description of Compression Extensions for WebSocket.

Error Codes

The following table describes error handling codes.

Code	Protocol	Message	Scenarios
400	HTTP	Bad Request	Malformed URL query: - missing or empty *clusterID* - missing or empty message type - invalid *sinceTime* or *toTime* (if present)
401	HTTP	Unauthorized	- Missing or empty Authorization Header - Invalid type of access token - Missing or empty access token - Invalid or expired access token - Invalid *clusterID* - Missing or expired *remote syslog* license for the given *clusterID*
404	HTTP	Not Found	- Invalid URL - Invalid protocol (for example, *http/https* are not supported
405	HTTP	Method not allowed	- Client is sending non GET requests
409	HTTP	Exceeded maximum number of sessions per token	The access token is being used by another session

Message Schema

The following tables describe the message data fields. Fields that are required are indicated as such in the Description column.

Top Level Elements

Name	Required?	Description	Data Type
guid	Required	Globally unique identifier for the message object.	string
connection	Required	Connection-related data.	object

Name	Required?	Description	Data Type
envelope	Required	Envelope-related data.	object
msg	Required	Message-related data.	object
msgParts	Required	Message Parts-related data (includes attachment data).	array
filter	Required	Email filtering data.	object
pps	Required	PPS-specific data.	object

Connection/Session Object Data

Name *(Connection/Session Object Data)*	Required?	Description	Data Type
sid	Required	The ID of the connection/session object; this is otherwise known as the "sid" in filter.log	string
country		The country code of the sender IP.	string
helo	Required	The FQDN or IP reported via the HELO or EHLO command.	string
host	Required	The host name of the reverse lookup of the sender IP.	string (hostname)
ip	Required	The sender IP in IPv4 or IPv6 format.	string (ipv4/ipv6)
protocol	Required	The connection protocol info.	string
resolveStatus		Can the sender IP be resolved with a reverse lookup.	string
tls.inbound.cipher		Inbound TLS cipher algorithm detected.	string
tls.inbound.cipherBits		Inbound TLS cipher algorithm strength (in #bits).	integer

Name *(Connection/Session Object Data)*	Required?	Description	Data Type
tls.inbound.policy		Inbound TLS policy.	string
tls.inbound.version	Required	Inbound TLS protocol version.	string

Envelope Object Data

Name *(Envelope Object Data)*	Required?	Description	Data Type
rcpts	Required	The envelope recipients.	array
from	Required	The envelope sender.	string (email)

Message Object Data

Note: None of these fields is required.

Name *(Message Object Data)*	Description	Data Type
header.cc	Carbon copy of email addresses.	array of strings
header.from	The header sender.	array of strings
header.message-id	The header message-id.	array of strings
header.reply-to	The header *Reply to* address.	array of strings
header.return-path	The header return path address.	array of strings
header.subject	The header subject.	array of strings

Name ***(Message Object Data)***	**Description**	**Data Type**
header.to	The header recipients.	array of strings
lang	The detected language of the message.	string
normalizedHeader	The "normalized" counterpart to the "header" object.	object
parsedAddresses.cc		array of strings
parsedAddresses.from		array of strings
parsedAddresses.to		array of strings
sizeBytes	The original, raw message size in bytes.	integer

Message Parts Object Data

Multiple message parts, in-line or attached, can be associated to an email message and this table lists the allowed fields for each attachment object.

Name ***(Message Parts Object Data)***	**Required?**	**Description**	**Data Type**
detectedCharset	Required	The detected charset of the message part.	string
detectedExt	Required	The detected extension of the message part.	string
detectedMime	Required	The detected MIME type of the message part.	string
detectedName	Required	The detected file name of the message part.	string
detectedSizeBytes	Required	The detected file size of the message part in bytes.	integer
disposition	Required	The content disposition value.	string
md5	Required	The ID of the message part in MD5.	string

Name ***(Message Parts Object Data)***	**Required?**	**Description**	**Data Type**
sha256	Required	The ID of the message part in SHA256.	string
isArchive	Required	Is the message part an archive type?	boolean
isCorrupted	Required	Is the message part corrupted?	boolean
isDeleted	Required	Is the message part deleted?	boolean
isProtected	Required	Is the message part password protected?	boolean
isTimedOut	Required	Did the message part analysis or text extraction time out?	boolean
isVirtual	Required	Is the message part virtual (a file member in an archive type of attachment)?	boolean
labeledCharset	Required	The charset of the message part as given.	string
labeledExt	Required	The extension of the attachment as given.	string
labeledMime	Required	The detected MIME type of the message part as given.	string
labeledName	Required	The name of the message part as given.	string
metadata		The metadata of the message part as reported by cvtd (interface to the document extraction engine).	object
sandboxStatus		The sandbox module status for the message part.	string
sizeDecodedBytes		The size of the decoded message part in bytes.	integer
structureId		The Structural ID of the message part with respect to container type attachments.	string
urls		The URLs that were detected.	array
urls.[].url	Required	The URL found in the corresponding message part.	string

Name ***(Message Parts Object Data)***	Required?	Description	Data Type
urls.[].isRewritten		Whether the URL was rewritten by URL Defense.	boolean
urls.[].notRewrittenReason	Required	The reason why the corresponding URL was not rewritten by URL Defense. The value is an empty string if it was rewritten.	string
urls.[].src	Required	The PPS sources that detected the URL.	array of strings

Filter Object Data

Name ***(Filter Object Data)***	Required?	Description	Data Type
actions	Required	The actions triggered; each array element is an object consisting of the action, module, and rule. The final disposition/action is marked with *isFinal*.	array
disposition	Required	The message disposition string as determined by *filterd* (the filtering engine daemon).	string
pe.rcpts		Recipients encrypted via Proofpoint Encryption.	array
quarantine.folder	Required	Quarantine folder containing a copy of the message.	string
quarantine.rule	Required	Rule that causes the message to be quarantined.	string
durationSecs	Required	Time spent processing the message.	number
currentFolder	Required	The folder to which the message is currently assigned.	string
isMsgEncrypted	Required	Is the message encrypted?	boolean
isMsgReinjected	Required	Was the message reinjected?	boolean
mid	Required	The message id.	integer
modules.av.virusNames	Required	The virus names reported by the AV module.	array

Name ***(Filter Object Data)***	**Required?**	**Description**	**Data Type**
modules.dkimv		The DKIM module data.	array
modules.dkimv.[].domain	Required	The DKIM d= value in the signature line.	string
modules.dkimv.[].selector	Required	The DKIM s= value in the signature line.	string
modules.dkimv.[].result	Required	The DKIM result.	string
modules.dmarc.filterdResult		The rollup DMARC result (generated by *filterd* for the rules, i.e. *$dmarcresult*).	string
modules.dmarc.authResults		The detailed authentication results.	array
modules.dmarc.authResults.[].emailIdentities		The email identities for a DMARC authorization result object.	object
modules.dmarc.authResults.[].emailIdentities.header.from		The *header.from* email identity for a DMARC authorization result object.	string
modules.dmarc.authResults.[].emailIdentities.smtp.helo		The *smtp.helo* email identity for a DMARC authorization result object.	string
modules.dmarc.authResults.[].emailIdentities.smtp.mailfrom		The *smtp.mailfrom* email identity for a DMARC authorization result object	string
modules.dmarc.authResults.[].method		The authorization result method.	string
modules.dmarc.authResults.[].propspec		The property specification for the authorization result per DMARC spec.	object
modules.dmarc.authResults.[].propspec.header.s		The *header.s* value for the property specification for the authorization result per DMARC spec.	string
modules.dmarc.authResults.[].reason		The reason string for the authorization result.	string
modules.dmarc.authResults.[].result		The result value for the authorization result.	string
modules.dmarc.records		The actual raw DMARC TXT record.	array
modules.dmarc.srvid		DMARC Auth Service ID as defined in *filter.cfg*.	string

Name ***(Filter Object Data)***	**Required?**	**Description**	**Data Type**
modules.dmarc.alignment		DMARC alignment report data.	array
modules.dmarc.alignment.[] .fromDomain		The DMARC TLD from the MAIL FROM data.	string
modules.dmarc.alignment.[] .results		The DMARC results array object; there can be multiple of these per method-identity combinations.	array
modules.dmarc.alignment.[] .results.[].identity		The DMARC domain identity as reported in the signature.	string
modules.dmarc.alignment.[] .results.[].identityOrg		The DMARC identifying organization as a Top Level Domain.	string
modules.dmarc.alignment.[] .results.[].method		The DMARC method involved for an alignment result object.	string
modules.dmarc.alignment.[] .results.[].result		The DMARC result involved for the alignment result object.	string
modules.pdr.v1.rscore		The PDR (Proofpoint Dynamic Reputation) v1 *rscore* value.	integer
modules.pdr.v1.spamscore		The PDR v1 *spamscore* value.	integer
modules.pdr.v1.virusscore		The PDR v1 *virusscore* value.	integer
modules.pdr.v2.response		The PDR v2 response status.	string
modules.pdr.v2.rscore		The PDR v2 *rscore* value.	integer
modules.sandbox.errorStatus	Required	The Attachment Defense error status string.	string
modules.spam	Required	The spam engine analysis on the message.	object
modules.spam.triggeredClassifier		The one spam classifier as defined by policy rules that determined the spam disposition.	string
modules.spf.result		The SPF (Sender Policy Framework) result.	string
modules.urldefense.rewrittenUrls	Required	The URLs rewritten by URL Defense.	array

Name ***(Filter Object Data)***	**Required?**	**Description**	**Data Type**
modules.urldefense	Required	Metadata reported by URL Defense.	object
modules.urldefense.version	Required	Version info for URL Defense.	object
modules.urldefense.version.engine	Required	Engine version for the URL Defense Module.	string
modules.urldefense.counts	Required	Metrics about the URLs evaluated by the URL Defense Module.	object
modules.urldefense.counts.maxLimit		The configured defined maximum number of **unique** URLs the URL Defense Module can process.	integer
modules.urldefense.counts.total	Required	The total number of URLs the URL Defense processed.	integer
modules.urldefense.counts.unique	Required	The total **unique** number of URLs the URL Defense Module processed.	integer
modules.urldefense.counts.rewritten	Required	The total number of URLs the URL Defense Module rewrote.	integer
modules.urldefense.counts.noRewriteIsEmail		The total number of URLs the URL Defense Module did not rewrite due to "is email".	integer
modules.urldefense.counts.noRewriteIsLargeMsgPartSize		The total number of URLs the URL Defense Module did not rewrite due to "is large message part size".	integer
modules.urldefense.counts.noRewriteIsExcludedDomain		The total number of URLs the URL Defense Module did not rewrite due to "is excluded domain".	integer
modules.urldefense.counts.noRewriteIsUnsupportedScheme		The total number of URLs the URL Defense Module did not rewrite due to "is unsupported scheme".	integer
modules.urldefense.counts.noRewriteIsSchemeless		The total number of URLs the URL Defense Module did not rewrite due to "is schemeless".	integer
modules.urldefense.counts.noRewriteIsMaxLengthExceeded		The total number of URLs the URL Defense Module did not rewrite due to "is max length exceeded".	integer
modules.urldefense.counts.noRewriteIsContentTypeText		The total number of URLs that the URL Defense did not rewrite due to "is content type text".	integer
modules.zerohour.score	Required	The ZeroHour threat score.	string
msgSizeBytes	Required	The size of the email in bytes.	integer

Name ***(Filter Object Data)***	**Required?**	**Description**	**Data Type**
origGuid		The parent GUID for the message from which the current message was split.	string
qid	Required	The *sendmail* queue ID.	string
routes	Required	The policy routes triggered by the message.	array
routeDirection		inbound outbound internal external	string
smime.rcpts		Recipients encrypted via S/MIME.	array
smime.signedRcpts		Recipients signed and encrypted via S/MIME.	array
startTime	Required	Timestamp for when message processing begins.	date-time
suborgs.sender	Required		string
suborgs.rcpts	Required		array
throttleIp		The IP address being rate-controlled.	string (ipv4/ipv6)
verified.rcpts		Verified recipients.	array

PPS Object Data

Name ***(PPS Object Data)***	**Required?**	**Description**	**Data Type**
agent	Required	The source/MFA host from which the email was received.	string (hostname)
cid	Required	The cluster ID license for the PPS deployment.	string
version	Required	The release PPS version.	string

Mail Schema

These fields represent the data in the mail logs. Each record or object matches a log line in the *maillog* given a particular *qid* (queue ID).

Field Properties

Name	Required?	Description	Data Type
data	Required	The raw data that corresponds to one log line from maillog.	string
id	Required	A unique ID for the object.	string
pps.agent	Required	The FQDN of the source agent on which the mail log line is produced.	string
pps.cid	Required	The cluster ID from which the data log line originated.	string
sm.auth			string
sm.class		The class (i.e., numeric precedence) of the message.	string
sm.ctladdr		The "controlling user", that is, the name of the user whose credentials are used for delivery.	string
sm.daemon		The daemon name from the **DaemonPortOptions** setting.	string
sm.delay		The total message delay: the time difference between reception and final delivery or bounce). Format is delay=HH:MM::SS for a delay of less than one day and delay=days+HH:MM::SS otherwise.	string
sm.dsn		The enhanced error code (RFC2034) if available.	string
sm.from		The envelope sender address.	string
sm.mailer		The name of the mailer used to deliver to this recipient.	string
sm.msgid	Required	The message id of the message (from the header).	string

Name	Required?	Description	Data Type
sm.nrcpts		The number of envelope recipients for this message (after aliasing and forwarding).	number
sm.pri		The initial message priority (used for queue sorting).	string
sm.proto		The protocol used to receive this message (e.g., ESMTP or UUCP).	string
sm.qid	Required	The corresponding sendmail queue ID for the log line.	string
sm.relay		Shows which user or system sent / received the message; the format is one of relay=user(a)domain [IP], relay=user(a)localhost, or relay=fqdn host.	string
sm.sizeBytes		The size of the incoming message in bytes during the DATA phase, including end-of-line characters.	number
sm.stat		The delivery status of the message. For successful delivery, stat=Sent (text) is printed, where text is the actual text that the other host printed when it accepted the message, transmitted via SMTP. For local delivery, stat=Sent is printed. Other possibilities are stat=Deferred: reason, stat=queued, or stat=User unknown.	string

Name	Required?	Description	Data Type
sm.tls.verify		The tls_verify data is included in two log lines. When the data appears in the from= log line, it describes TLS results when the message was received by the Proofpoint Protection Server. When the data appears in the to= log line, it describes TLS results when the message was sent from the Proofpoint Protection Server. Results for tls_verify from = lines: NONE - Client did not use STARTTLS or it was disabled. NOT - Client used STARTTLS; PPS was configured to not request a client certificate. NO - Client used STARTTLS and PPS requested a client certificate, but the client did not send one. FAIL - Client used STARTTLS, PPS requested a client certificate, and the client sent one, but certificate validation failed. OK - Client used STARTTLS, PPS requested a client certificate, the client sent one, and certificate validation succeeded. Results for tls_verify to= lines TEMP - Non-TLS temporary error occurred. PROTOCOL - Non-TLS protocol error occurred. SOFTWARE - TLS handshake error occurred. NONE - STARTTLS was not offered by the remote server or PPS was configured to not use it (with this server). NO - PPS used STARTTLS and managed to negotiate an anonymous cipher suite. FAIL - PPS used STARTTLS, but validation of the remote server certificate failed. OK - PPS used STARTTLS and validation of the remote server certificate succeeded.	object
sm.to		Recipients to this mailer.	string array
sm.xdelay		The total time the message took to be transmitted during final delivery. This differs from the delay= equate, in that the xdelay= equate only counts the time in the actual final delivery.	string
ts	Required	Timestamp of logging time in ISO8601 format.	string

Mail Schema

```
{
    "$schema": "http://json-schema.org/draft-04/schema#",
    "id": "https://www.proofpoint.com/v2/schemas/maillog.json",
    "properties": {
        "data": {
            "id": "/properties/data",
            "type": "string"
        },
        "id": {
            "id": "/properties/id",
            "type": "string"
        },
        "pps": {
            "id": "/properties/pps",
            "properties": {
                "agent": {
                    "id": "/properties/pps/properties/agent",
                    "type": "string"
                },
                "cid": {
                    "id": "/properties/pps/properties/cid",
                    "type": "string"
                }
            },
            "required": [
                "agent",
                "cid"
            ],
            "type": "object"
        },
        "sm": {
            "id": "/properties/sm",
            "properties": {
                "ctladdr": {
                    "id": "/properties/sm/properties/ctladdr",
                    "type": "string"
                },
                "delay": {
                    "id": "/properties/sm/properties/delay",
                    "type": "string"
                },
                "dsn": {
                    "id": "/properties/sm/properties/dsn",
                    "type": "string"
                },
                "mailer": {
                    "id": "/properties/sm/properties/mailer",
                    "type": "string"
                },
                "pri": {
                    "id": "/properties/sm/properties/pri",
                    "type": "integer"
                },
                "qid": {
                    "id": "/properties/sm/properties/qid",
                    "type": "string"
                },
                "stat": {
                    "id": "/properties/sm/properties/stat",
                    "type": "string"
                },
                "tls": {
                    "id": "/properties/sm/properties/tls",
                    "properties": {
                        "verify": {
                            "id":
"/properties/sm/properties/tls/properties/verify",
                            "type": "string"
                        }
                    },
                    "required": [
                        "verify"
                    ],
                    "type": "object"
                },
                "to": {
                    "id": "/properties/sm/properties/to",
                    "items": {
                        "id": "/properties/sm/properties/to/items",
                        "type": "string"
                    },
                    "type": "array"
                },
                "xdelay": {
                    "id": "/properties/sm/properties/xdelay",
                    "type": "string"
                }
            },
            "required": [
                "qid"
            ],
            "type": "object"
        },
        "ts": {
            "id": "/properties/ts",
            "type": "string"
        }
    },
    "required": [
        "pps",
        "data",
        "ts",
        "sm",
        "id"
    ],
    "type": "object"
}
```

Example

```
{
  "pps": {
    "agent": "example.proofpoint.com",
    "cid": "mmeng_uivm071"
  },
  "ts": "2017-08-17T14:54:12.949180-07:00",
  "data": "2017-08-17T14:54:12.949180-07:00 example sendmail[30641]:
v7HLqYbx029423: to=/dev/null, ctladdr=<user1@example.com> (8/0),
delay=00:00:00, xdelay=00:00:00, mailer=*file*, tls_verify=NONE, pri=35342,
dsn=2.0.0, stat=Sent",
  "sm": {
    "tls": { "verify": "NONE" },
    "stat": "Sent",
    "qid": "v7HLqYbx029423",
    "dsn": "2.0.0",
    "mailer": "*file*",
    "to": ["/dev/null"],
    "ctladdr": "<user1@example.com> (8/0)",
    "delay": "00:00:00",
    "xdelay": "00:00:00",
    "pri": 35342
  },
  "id": "ZeYGULpZmL5N0151HN1OyA"
}
```