ctIPd™ Integration for WAN Deployment

ctIPd Module and Documentation Usage Restrictions

This program and the accompanied documentation is SECRET AND CONFIDENTIAL, and constitute a proprietary trade secret of CYREN Inc. (herein after referred to as "CYREN").

No person is allowed to copy, decompile, reverse engineer, use, sublicense or otherwise access this program unless the prior express, written consent is received from CYREN. The possession and use of this program shall be governed by the terms of a license agreement between CYREN and each authorized licensee. Unauthorized use of this program is strictly prohibited, and those perpetrating such unauthorized uses shall be prosecuted to the fullest extent of the law. The confidentiality and non-disclosure obligations of licensee shall be strictly maintained at all times by licensee and licensee, in receiving a copy of this program, acknowledges that it shall not be disclosed to third parties; rather, only to employees or consultants having a firm need to know, and provided that they are bound by confidentiality restrictions at least as restrictive as those adopted by licensee within the framework of its relationship with CYREN.

The failure to maintain confidentiality will likely cause severe damages and irreparable harm to CYREN and, therefore, in addition to any other remedies and rights available at law, CYREN shall be entitled to seek injunctive relief without the need for the posting of any bond or other guarantee.

Contacts

Any technical questions you or your developers have about using the ctwsd should be addressed to support@cyren.com.

Trademark and Copyright Statement

ctIPd is a licensed SDK product featuring patented technology. CYREN's patented solution is protected by U.S. patent #6,330,590.

RPD, Zero-Hour Protection, IPRep, ctIPd are trademarks of CYREN Inc. For more information, visit our website: http://www.cyren.com/

Linux is a trademark of Linus Torvalds. FreeBSD is a registered trademark of Wind River Systems, Inc. RBL is a trademark for the proprietary MAPS DNSBL.

All other trademarks and registered trademarks are the property of their respective owners.

About CYREN

CYREN™ provides proven Internet security technology to more than 150 security companies and service providers including 1&1, Check Point, F-Secure, Google, Microsoft, Panda Security, Rackspace, US Internet, and WatchGuard, for integration into their solutions. CYREN's GlobalView™ and patented Recurrent Pattern Detection™ (RPD™) technologies are founded on a unique cloud-based approach, and protect effectively in all languages and formats. CYREN Antivirus utilizes a multi-layered approach to provide award winning malware detection and industry-leading performance.

CYREN technology automatically analyzes billions of Internet transactions in real-time in its global data centers to identify new threats as they are initiated, enabling our partners to protect end-users from spam and malware, and ensure safe, compliant browsing. The company's expertise in building efficient, mass-scale security services mitigate Internet threats for thousands of organizations and hundreds of millions of users in 190 countries.

CYREN, formerly known as Commtouch, was founded in 1991, is headquartered in the US in McLean, Virginia, with offices in Palo Alto, California, Herzliya, Israel, Berlin, Germany, and Reykjavik, Iceland.

For more information about enhancing security offerings with CYREN technology, visit our website at www.cyren.com, see our blog at http://blog.cyren.com or write to info@cyren.com.

Table of Contents

1 Introduction

CYREN GlobalView™ Mail Reputation services are used primarily to weed out spam messages and email-borne malware at the entry point, before these messages enter the customer's messaging network, thereby relieving the need for resource-consuming downstream filtering. By applying Mail reputation services to senders of inbound messages before or during the SMTP connection and before they enter the messaging network, ctIPd delivers cost-effective benefits such as the following:

- Reducing IT resources such as server count, CPU load, storage, etc.
- Eliminating multiple security risks
- Reducing the level of false positives
- Minimizing the cost of downstream filtering
- Lowering the overall bandwidth consumption
- Optimizing IT labor required to manage the overall messaging process

Additionally, CYREN GlobalView Mail Reputation services are used as part of an overall strategy to optimize network accessibility. By weeding out these resource-draining SMTP connections before any actual messages are transmitted, CYREN's solution enables the network's messaging processes to work more efficiently, so that valuable IT resources are correctly focused on allowing legitimate sources full and uninterrupted access while making access to unauthorized sources that are attempting to abuse the network, more difficult to achieve.

Many public services nowadays already deliver free information about blacklisted sources; typically, these lists are static with limited commitment for accuracy by the owners. Other commercial services also enable setting central as well as personal policies for whitelisted sources during the SMTP connection. However, these solutions are not designed to handle the vast 'grey' area of sources for which little or no information is readily available. Included in this group, for example, are millions of machines hijacked on a daily basis and forced into botnets and armies of zombies discrediting their reputation, dynamically. Many of these bots are detected at a later time and are removed and thus the reputation of the once-infected machines is raised.

CYREN analyzes hundreds of millions of messages every day. Included in this analysis is a process whereby CYREN dynamically and in real-time classifies and reclassifies the reputation of each source IP with a series of meaningful and measurable values. These values are constantly updated to include the most up-to-date information and reflect even the slightest change in credibility.

The CYREN GlobalView Mail Reputation daemon (**ctIPd**™) is an embedded reputation engine with a small footprint responsible for maintaining communication with the CYREN Datacenter. ctIPd delivers reputation data in real-time to messaging, security and networking devices. These devices are responsible for posting queries to ctIPd over HTTP or UDP protocol requesting reputation data on source IP addresses attempting to establish SMTP sessions for sending messages to recipients. ctIPd analyzes all available data and provides up-to-date reputation data back to the querying device. In response to ctIPd's returned data the querying device will implement a mechanism to accept, tempfail, or permfail the connection from a sending SMTP source. This

is normally done via a non-CYREN flow control or connection management application on the client-side that is configured to receive reputation data from ctIPd and then to weigh the responses and make its own determination for the appropriate actions to implement.

2 Deploying ctIPd over WAN

When ctIPd is deployed over WAN, either on a CYREN Datacenter or on a CYREN OEM partner's Datacenter the following considerations must be applied.

- Make sure that each request to ctIPd (except in the case of `GetStatus` request) contains the header `X-CTCH-Key`. Use the same license key that is used in ctipd.conf.
- It is recommended that you provision a failover mechanism to the remote ctIPd units as described later in this section.

2.1 Implementing a Failover Mechanism

When deploying ctIPd over WAN, the connectivity between the querying devices and ctIPd units may be interrupted from time to time and the responsibility for enabling connectivity continuity to the ctIPd unit is with the CYREN OEM partner deploying ctIPd.

If you plan a failover mechanism to ctIPd units that are deployed over WAN follow these guidelines:

- Check the IP addresses of all ctIPd units periodically, (i.e., `gethostbyname`) to find out if one or more addresses were changed (a good mechanism will check every 30-50 seconds).
- Try to maintain connectivity with the first responding ctIPd unit rather than switching back and forth frequently between ctIPd units.
- If the last-used ctIPd unit is not responding, you may implement a retry mechanism for several times (i.e., 3 times) and only then attempt to connect the next ctIPd unit.
- Switch back to the first ctIPd unit if connectivity with it was resumed after a failure and if it is available for several consecutive connection attempts.

3 ctIPd Protocol

CYREN has developed a simple protocol to enable communication between the querying device and ctIPd. Using a pre-determined syntax for the envelope and data contained in a query, the querying device sends requests to ctIPd over HTTP or UDP. Once processed, the ctIPd responds to the request (or an error message, if the request could not be processed). The protocol structure for HTTP/UDP is extensible, meaning that the order of the headers is not mandatory and not all headers are required to be included in all responses.

The following sections detail the protocol options that can be used to initiate these requests by the querying device, and the response that will be sent by ctIPd with the most up-to-date information available on the sender IP.

3.1 Protocol and Syntax Options

ctIPd uses a simple protocol to control the transfer of data to and from a querying device. Communication between ctIPd and the querying device can be accomplished over:

- **HTTP:** standard HTTP/1.0 headers to POST requests are used in the following URL: http://<host>/ctipd/iprep.
- **UDP:** either a numbered port or UNIX socket over UDP is used.

When using HTTP or UDP, the same reputation data is sent to ctIPd as part of the request and the same data type values are returned in the response. When using RBL the zone contains the following data: queried IP, one or more data elements identifying the zone and the domain name of the ctIPd host machine.

A typical communication session over HTTP or UDP would begin with the querying device initiating a request to ctIPd asking for reputation data. The communication would include some or all of the following:

1. The querying device sends a `classifyip` request for reputation data about a single IP address (hereafter, 'Sender IP' or 'Source IP'). The request contains envelope and data. The contents and syntax of the request is the same for interfacing over HTTP or UDP.

2. ctIPd receives the request and sends back a response to the `classifyip` request. This response contains various values about the IP range to which the queried IP belongs. If the Decision Manager is enabled, ctIPd will also return a recommended action (such as accept, permfail or tempfail). If an error occurred, ctIPd will respond with an error.

3. Once the response to `classifyip` request is received, the querying device applies an appropriate flow control decision to the session, based on the information contained in the response. Some OEM partners use only the IP class to determine which decisions and/or actions their flow control application will apply to incoming traffic from a sender IP. Other OEM partners incorporate additional response

data types in the configuration or policy options. The flow control application may also weight responses from multiple processes before applying its final decision.

The following sections contain details on the protocol syntax used for all types of requests and responses, as well as explanations for the response data types contained in ctIPd's responses.

3.2 HTTP and UDP Requests and Responses

In the following sections, the request and response syntax are detailed. In the first section, a sample for all HTTP requests and responses is included. The second section provides information about each header.

3.2.1 Sample Requests and Responses over HTTP

classifyip Request over HTTP

```
URL: http://<host>/ctipd/iprep method: POST


POST /ctIPd/iprep HTTP/1.0
Accept-Language: en-us
Accept: */*
Content-Length: 3867
Host: 176.211.45.4
User-Agent: CYREN HTTP Client

x-ctch-request-id: 12345678
x-ctch-request-type: classifyip
x-ctch-pver: 1.0

x-ctch-ip: 216.163.176.213
```

Response to classifyip Request over HTTP

```
x-ctch-request-id: 81263
x-ctch-request-status: 0
x-ctch-pver: 1.0

x-ctch-iprange: 216.163.176.213-216.163.176.213
```

```
x-ctch-refid: 0001.0A090303.4693AF75.0114
x-ctch-ipclass: T3
x-ctch-volume: 0,3,3
x-ctch-risk-level: 73
x-ctch-spam-ratio: 100,42
x-ctch-valid-bulk-ratio: 57
x-ctch-maxhits: 3
x-ctch-ttl: 7126
x-dm-action:tmpfail
```

Response with Error over HTTP

```
HTTP/1.0 200 OK
Date: Sat, 12 May 2006 22:25:21 GMT
Server: CYREN IP Reputation daemon /1.0
Content-Length: 5421
Connection: close
Content-Type: text/plain

x-ctch-request-id: 12345678
x-ctch-pver: 1.0
x-ctch-request-status: 201

Your license key is invalid or blocked in the CYREN Datacenter.
Contact CYREN technical support.
```

3.2.2 Sample Requests and Responses over UPD

classifyip Request over UDP

```
x-ctch-request-id: 12345678
x-ctch-request-type: classifyip
x-ctch-pver: 1.0

x-ctch-ip: 216.163.176.213
```

Response to classifyip Request over UDP

```
x-ctch-iprange: 216.163.176.213-216.163.176.213
x-ctch-refid: 0001.0A090303.46D402B5.0016
x-ctch-ipclass: R7
x-ctch-volume: 0,10,10
x-ctch-risk-level: 78
x-ctch-spam-ratio: 87,87
x-ctch-valid-bulk-ratio: 0
x-ctch-24h: 468452,53
x-ctch-ttl: 1800
x-ctch-from-cache: no
x-ctch-dm-action:tmpfail
```

Response with Error

```
x-ctch-request-id: 12345678
x-ctch-pver: 1.0
x-ctch-request-status: 201


Your license key is invalid or blocked in the CYREN Datacenter.
Contact CYREN technical support.
```

3.3 classifyip Request for HTTP and UDP

Each classifyip request contains the request envelope as well as the request data. This enables you to define the IP address for each query.

Request Envelope

Header	Explanation
x-ctch-request-id	Optional header. This value is highly recommended in the case of UDP due to the connectionless nature of UDP.
x-ctch-request-type	Defines the request type
x-ctch-pver	The current client version of the CYREN protocol

Note: *The Request Envelope must be separated from the Request Data with an empty line.*

Request Data

Header	Explanation
`x-ctch-ip`	The IP address for which the query was sent. Each request can contain only one IP address.
`x-ctch-cache-only`	This optional header indicates if zero-latency is enabled. The default value is zero (0). If set to one (1), then zero latency response will be used.

3.4 Response to classifyip Request for HTTP and UDP

The response to `classifyip` request contains the IP reputation for IP range to which the queried IP is belonged as calculated by CYREN. Within the reply, various data types are used to determine the IP reputation of the calculated IP range and a reference ID is added to help track the transactions for diagnostics.

Response Envelope

Header	Explanation
`x-ctch-request-id`	The same as in the `x-ctch-request-id` headers used in the request
`x-ctch-request-status`	0 OK
	101 Unknown Command
	102 Bad Request
	201 Auth. Error
	301 Center Not Operational
	302 Connection to Datacenter failed
	303 Connection to Datacenter timed out
	304 Request to Datacenter timed out
	305 General Communication Error
	401 Internal Error
`x-ctch-pver`	The current server version of the CYREN protocol

This is really weird. I

Note: *The Response Envelope must be separated from the Request Data with an empty line.*

Response Data

Header	Explanation
`x-ctch-iprange`	Represents the first and last addresses of the IP range to which the queried IP belongs
`x-ctch-refid`	RefID of the transaction
`x-ctch-ipclass`	Calculated IP class of the Sender IP
`x-ctch-volume`	Calculated IP volume for the Sender IP
`x-ctch-risk-level`	Calculated risk level of the Sender IP
`x-ctch-spam-ratio`	Calculated Spam ratio for the Sender IP
`x-ctch-valid-bulk-ratio`	Calculated valid Bulk ratio for the Sender IP
`x-ctch-maxhits`	Maximum hits: Represents how long the reputation data for a specific IP is saved, based on the number of attempts to establish an SMTP connection. When implementing a local cache on the querying device, it is advised not to query ctIPd until maxhits is exceeded.
`x-ctch-ttl`	Time-to-live (ttl): Represents how long the reputation data for a specific IP is saved, based on time. When implementing a local cache on the querying device, it is advised not to query ctIPd until the ttl is exceeded.
`x-ctch-dm`	Recommended call-to action. Options include: `accept`, `tempfail` and `permfail`.

Notes: *Each of the Data Types is explained in the section Response Data Types.*

When a value for both x-ctch-maxhits and x-ctch-ttl are specified, it is recommended that the record be deleted from the local cache as soon as either value is reached.

3.5 Response with Error for HTTP and UDP

The response with Error from ctIPd to the querying device contains the status and an explanation of why the error occurred. The same format is used to respond with errors for `classifyip` requests.

Response Envelope

Header	Explanation
`x-ctch-request-id`	Optional header
`x-ctch-pver`	The current server version of the CYREN protocol
`x-ctch-request-status`	Number representing a status. For example, for an invalid license key, you would receive: `201`

Response Data

Header	Explanation
Free text	Full text with description of error. For example: Our license key is invalid or blocked in the CYREN Datacenter. Contact CYREN technical support.

3.6 Response Data Types

To offer the maximum flexibility to the OEM partner, CYREN answers each query with a detailed response containing several reputation-related data types. Included in this detailed response is the IP class, which represents a composite reputation value.

- Some OEM partners use only the IP class to determine which decisions and/or actions their flow control application will apply to incoming traffic from a sender IP.
- Others use only the Action contained in the ctIPd response without consulting additional reputation sources.
- Still other OEM partners choose to incorporate additional response data types in the configuration or policy options.

ctIPd's comprehensive response syntax offers maximum flexibility and ease of integration by enabling each OEM partner to make use of any part of the IP reputation response to determine an appropriate action regarding sender IP reputations.

3.6.1 Response Data Types

Each time ctIPd receives a `classifyip` request from the querying device over HTTP or UDP, it returns a response with values for each of the following data types:

- IP range
- IP class
- IP volume

- Risk level
- IP Spam ratio
- IP valid Bulk ratio
- RefID
- TTL
- Call-to-Action

3.6.2 IP Class Groups

The IP class data type represents a composite reputation value assigned by CYREN. There are three groups, each represented by a letter (R, T, or G) and a number. The number represents the level of risk.

- Groups R and T represent the volume received in a 24-hour period.
- Group G, the **Known Reputation** group, represents sources with fixed decisions, not based on their monitored volume. These include blacklisted, whitelisted and private IP sources.

A variety of elements are used to determine the risk value. By combining the volume and risk, CYREN has determined set IP classes that fall into three groups:

- High Volume
- Transitory or Low Volume
- Known Reputation

Group	Explanation
Rx: R1..R9	Represents sources with substantial volume history. For example, R1 being a source that was monitored over time with high volume and low risk and R9 a source with very high volume and high risk.
Tx: T1..T5	Represents sources with transitory volume or no volume history. For example, T1 being a source with low transitory volume and low risk and T5 a source with low transitory volume and high risk.
Gx: G1..G3	Represents sources for which there are fixed decisions and regardless of their monitored volume. In this group, G1 are all whitelisted sources, G2 are all blacklisted sources, and G3 are all private IP sources.

3.6.3 IP Range

While the querying device queries about a specific IP address, ctIPd will respond with a calculated range of IP addresses to which the queried IP was calculated to belong. The range is typically made of one or more successive IP addresses found to have the same behavior as the queried IP.

When the querying device generates a new request about a different IP address within the same range, then ctIPd will respond immediately using data from the local cache and will not forward the request to a CYREN datacenter unless the record in the local cache was already expired for this IP range.

The IP range includes two sets of IP addresses:

- First IP in the range
- Last IP in the range
- Example: 1.2.3.0 – 1.2.3.255

Example of Private IP Addresses

From	To	Representation
10.0.0.0	10.255.255.255	10/8
172.16.0.0	172.31.255.255	172.16/12
192.168.0.0	192.168.255.255	192.168/16
127.0.0.0	127.0.0.255	127/8

3.6.4 Volume

The IP volume is a 3-value array as detailed below:

- **Recent Peak**, indicates the deviation of the recent volume from the approximated hourly peak calculated from the daily average over the last 30 days [signed integer].
- **Calculated daily average of the last 30 days**, ranging from 0 to 10 with 0 being a source IP having no monitored traffic [unsigned integer].
- **Sdtdev**. of the last 24 hours from the daily average of last 30 days, ranging from 0 to 10. [unsigned integer].

3.6.5 Risk Level

This value represents the risk level, a composite value that weights a host of variables on the Datacenter-side. A value of '0' in the Risk Level represents a whitelisted IP and a value '100' represents a blacklisted IP.

3.6.6 Spam Ratio

The Spam ratio is calculated by dividing the monitored spam volume by the total monitored volume. The Spam ratio is a 2-value array as detailed below:

- **Recent Peak**, indicates the deviation of the recent Spam ratio from the approximated hourly peak calculated from the daily average over the last 30 days, ranging from 0 to 100 [percentage].

- **Calculated daily average of the last 30 days**, ranging from 0 to 100 with '0' being a source IP having no monitored Spam traffic [percentage].

3.6.7 Valid Bulk Ratio

The valid Bulk ratio is calculated by dividing the monitored valid Bulk volume by the total monitored volume. The valid Bulk ratio has a single value as detailed below:

- **Calculated daily average of the last 30 days**, ranging from 0 to 100 with '0' being a source IP having no monitored valid Bulk traffic [percentage].

3.6.8 RefID

The refID is a reference ID record that will be used by CYREN to diagnose various support issues, per-transaction. It is recommended that you keep this IP RefID in case you need to trace back to determine why ctIPd returned any particular response.

Note: *Without this key, CYREN is unable to retroactively determine the reason why a query was returned with any specific set of data values.*

3.6.9 TTL

The TTL represents how long to reuse the reputation data before sending a new query for the same source IP. The TTL can be determined by ctIPd either as time unit (e.g., 2 hrs.) or max hits (e.g., 8 times).

3.6.10 Action

The recommended action after weighting the source reputation data and counting how many times it attempted to establish SMTP connection. Optional actions are: *accept*, *permfail* and *tempfail* based on predefined throttling logics.

3.6.11 Zero-Latency Response Time

Although the average response time is extremely short, you may wish for a minimal latency at certain occasions. In this case, the querying device can request that the response be generated from the ctIPd's local cache only, rather than waiting for a response from the CYREN Datacenter.

When this option is enabled, if the requested data is not immediately available in the local cache, ctIPd will respond with the status '0' [OK] and an empty response. At the same time, however, ctIPd will still generate a request to the CYREN Datacenter in relation to the queried sender IP.

When the response comes back from the Datacenter, the response is stored in the local cache for future queries. By default, this functionality is disabled.

The querying device may request to benefit zero-latency response regardless if the request was made over HTTP or UDP. To utilize this mode of operation a special MIME header, `x-ctch-cache-only`, is added to the `classifyip` request as described above in the protocol.

4 ctIPd Testing and Verification

Once you have successfully deployed ctIPd, it is recommended that you perform the following tests to verify that ctIPd is receiving and responding to queries sent by your querying device as expected. To properly test ctIPd, CYREN has provided several Perl scripts which receive a list of IP addresses in a command line and forward them to ctIPd for classification. This process includes the following:

- You can either enter a parameter in the command line, or read the IP addresses from STDIN until EOF.
- You must specify the hostname and port where ctIPd runs. By default, this is local host and port 8080 for HTTP and 5678 for UDP.

For effective testing, CYREN predefined the following IP addresses with fixed values to be returned in responses to `classifyip` requests:

Test IP	IPClass	Risk	Volume	SpamR	BulkR	TTL	MaxHits
216.163.176.201	G1	0	0,5,4	0,0	3	1h	Null
216.163.176.202	G2	100	0,5,4	92,90	0	2h	Null
Private IP Ranges	G3	0	1,0,0	Null	Null	Null	Null
	For more information, see Example of Private IP Addresses.						
216.163.176.211	T1	38	0,5,1	0,35	5	Null	16
216.163.176.212	T2	56	0,1,2	55,59	0	Null	0
216.163.176.213	T3	73	0,3,3	100,42	57	Null	4
216.163.176.214	T4	87	300,2,1	89,84	0	Null	8
216.163.176.215	T5	95	-31,4,5	93,97	1	Null	16
216.163.176.221	R1	3	0,7,0	0,3	0	10min	Null
216.163.176.222	R2	7	0,6,2	0,6	0	30min	Null
216.163.176.223	R3	15	-35,6,6	0,12	67	30min	Null
216.163.176.224	R4	29	-91,6,6	56,0	99	30min	Null

Test IP	IPClass	Risk	Volume	SpamR	BulkR	TTL	MaxHits
216.163.176.225	R5	51	0,7,6	63,1	94	30min	Null
216.163.176.226	R6	68	62,6,5	90,49	20	30min	Null
216.163.176.227	R7	72	0,7,8	59,74	9	30min	Null
216.163.176.228	R8	88	40,8,5	94,81	3	30min	Null
216.163.176.229	R9	95	0,7,5	0,97	0	1h	Null

After running these tests, you may also add `x-ctch-cache-only` to the next `classifyip` request to receive zero-latency response. Additionally, you may want to simulate one or more cases of returned errors by ctIPd. To do this you can review the list of returned errors in the protocol and generate queries that will result in having `x-ctch-request-status` return value different than '0'. For example, you may disrupt the query with bad format or missing data or use an invalid IP in the request, etc.

4.1.1 Evaluation and Testing Sample Scripts

The following sample scripts are available:

For HTTP: scanip_http.pl

For UDP: scanip_udp.pl

Usage is the same for both sample scripts:

```
./<script's name> [OPTIONS][IP(s)]
```

Command line switches are as follows:

OPTIONS

`--host`

Define the server host. Default is 127.0.0.1

--port

Define the server port. Default is '8080' for HTTP and '5678' for UDP.

--cache

Send a cache-only request. Default '0'.

--help

Display the help message.

For input, you may feed the sample scripts with one or more IP(s) from `stdin` or a file.