Copyright Notice

©2013-2024 Tencent Cloud. All rights reserved.

Copyright in this document is exclusively owned by Tencent Cloud. You must not reproduce, modify, copy or distribute in any way, in whole or in part, the contents of this document without Tencent Cloud's the prior written consent.

Trademark Notice

All trademarks associated with Tencent Cloud and its services are owned by Tencent Cloud Computing (Beijing) Company Limited and its affiliated companies. Trademarks of third parties referred to in this document are owned by their respective proprietors.

Service Statement

This document is intended to provide users with general information about Tencent Cloud's products and services only and does not form part of Tencent Cloud's terms and conditions. Tencent Cloud's products or services are subject to change. Specific products and services and the standards applicable to them are exclusively provided for in Tencent Cloud's applicable terms and conditions.
Contents

Origin Configuration
Load Balancing
  Overview
  Quickly Create Load Balancers
  Health Check Policies
  Viewing the Health Status of Origin Server
Related References
  Load Balancing-Related Concepts
  Introduction to Request Retry Strategy
Origin Group Configuration
Id Version Origin Group Compatible Related Issues
Collect EdgeOne origin-pull node IP
Host Header Rewrite
Range GETs
Configuring Origin-Pull HTTPS
HTTP/2 Origin-Pull
Redirect Following During Origin-Pull
Controlling Origin-pull Requests
Origin Configuration
Load Balancing
Overview

Last updated: 2024-05-29 10:33:37

EdgeOne Load Balancing is ideal for scenarios where high availability of origins is crucial. It supports the configuration of multi-level secondary sources for disaster recovery switching. It can proactively probe the health status of origins. This proactive measure blocks failed origins and directs business traffic to healthy origins.

Note:
EdgeOne Load Balancing feature is in beta testing. If you want to use it, you can Contact Us.

Use Cases

Hardware failures, network failures, configuration errors, security attacks, natural disasters, human errors, and other unforeseen circumstances, can affect the availability of the origin. For businesses that require high availability, such as finance, gaming, audio and video, and e-commerce, even brief failures of the origin can result in significant losses. Therefore, it is necessary to implement primary/secondary disaster recovery and health checks for the origin.

Primary/Secondary Disaster Recovery: When the primary source becomes unavailable, it is automatically switched to the secondary source to ensure business continuity.

Proactively Checking origin Health Status: Preemptively disables failed origins, and redirects business traffic to healthy origins. Prevents a situation where a significant number of legitimate service requests are still directed to malfunctioning origins in the event of a failure.

Supported Capabilities

1. Supports the configuration of multi-level secondary sources for multi-source disaster recovery.
2. Supports the configuration of health check policies such as ICMP Ping, HTTP/HTTPS, TCP, and UDP to preemptively disable failed origins and redirect business traffic to healthy origins.
3. Provides a fallback retry policy. It retries directing the traffic to alternative healthy origins when real business traffic requests fail.
More Information

Quickly Create Load Balancers
Load Balancing-Related Concepts
Health Check Policies
Quickly Create Load Balancers

Last updated: 2024-05-29 10:33:37

This document guides you on how to create a Cloud Load Balancer instance.

**Note:**
EdgeOne Load Balancing feature is in beta testing. If you want to use it, you can Contact Us.

Sample Scenario

For example, you currently have an acceleration domain `www.example.com`, with origins `1.2.3.4`, `2.3.4.5`, and `3.4.5.6`. Under normal circumstances, both `1.2.3.4` and `2.3.4.5` are used as primary origins. You have already configured them as the origin group named `primary_origins` following the Origin Group Operation Guide. The server `3.4.5.6` is used as a standby origin in a group called `backup_origins`, which is only used when the primary origins fail. In cases where a real business request fails, retries are attempted with other healthy servers within the same group. Additionally, there is a requirement for proactive probing to actively identify and disable unhealthy origins.

Directions

1. Log in to the Tencent Cloud EdgeOne console. In the left menu bar, click the Site List. Within this list, click the Site that need to be configured to go to the details page.
2. On the site details page, click Origin Settings > Load Balancing.
3. On the Load Balancing page, click Create Instance.

4. Proceed to step 1 of choosing the origin. You need to fill in the instance name, choose the instance type, and add an origin group.
   Taking this scenario as an example, add the origin group `primary_origins` as a priority 1 origin group, add the origin group `backup_origins` as a priority 2 origin group, and click Next.
### Parameter Description

**Instance name**
- Limit to 1-200 characters in length. Allowed characters are a-z, A-Z, 0-9, _, -.  

**Instance type**
- HTTP-specific Type: Supports adding both HTTP-specific and general origin groups. It is only applicable for reference by site acceleration-related services, such as domain services and rule engines.  
- General Type: Only supports adding general origin groups. It is applicable for site acceleration services including domain services and rule engines, and reference by L4 proxy service.  

**Add origin group**
- In the CLB instance, the smallest configuration dimension for an origin is the origin group. You need to configure the origin into an origin group and add it here. For more details, see Origin Group Configuration.  
- You can set priorities for the added origin groups. Traffic will not be directed to origins in lower-priority origin groups if there are healthy origins in higher-priority origin groups. Up to 10 origin groups can be configured, with lower numbers indicating higher priorities.

5. Proceed to step 2 of health check policy. It supports four types of probes: ICMP Ping, HTTPS/HTTP, TCP, and UDP. Tencent Cloud EdgeOne will actively send probe requests to your origin to check its latency and health status. You can choose the appropriate probe frequency based on the load condition of your origin. Here, choose ICMP Ping as the probe policy. For a detailed introduction to probe policy configuration, see Introduction to Health Check Policy. After configuration is completed, click **Next**.
Note:

If you do not want EdgeOne’s nodes to initiate any probe requests to origins, you can choose Not Enabled. In this case, the Load Balancing instance will default to traffic scheduling based on the priority order of the origin groups from step 1. If a request to a particular origin fails 5 times within 60 seconds, the corresponding origin will be disabled for 10 minutes according to the default policy.

Using this policy will not be able to disable the origin of the failure in advance, and it will not be able to automatically and quickly recover the traffic scheduling after the origin returns to normal. Compared with enabling active probe, using this policy may cause you to encounter more failed requests during the origin failure period. Therefore, if you want your business to have higher availability, it is recommended that you enable active probe.

6. Proceed to step 3 of traffic scheduling policy. The current traffic scheduling policy defaults to failover based on the priority order according to the results of active probes. When real business requests fail to retrieve content from the origin during the backsource process, support for request retry is available. There are two request retry policies available. For details, see Introduction to Request Retry Policy.

Policy 1: When a real business request fails to access a certain origin, it directly retries with another origin within the next lower priority origin group. This is suitable for scenarios where the performance of both origin group 1 and origin group 2 is similar.

Policy 2: When a real business request fails to access a certain origin, it directly retries with another origin within the same priority origin group. This is suitable for scenarios where the performance of origin group 1 is significantly better than that of origin group 2.
7. Taking this sample scenario as an example, policy 2 can be chosen. Click **Complete** to finish creating the instance.
Health Check Policies

Last updated: 2024-05-29 10:33:37

This document introduces the probe methods and their principles within health checks, the origin health determination criteria and the calculation methods.

Note:
EdgeOne Load Balancing feature is in beta testing. If you want to use it, you can Contact Us.

Principle of Health Checks

After configuring health check policies, EdgeOne’s probe nodes in different regions will send probe requests to your origin and determine the health status of the origin based on the response results. Health check policies consist of probe methods and origin health determination criteria. The probe method determines the type of probe request, while the origin health determination criteria determine how the response results are processed.

Probe Method

Currently, supports for ICMP Ping, HTTP/HTTPS, TCP, and UDP as the four methods of probe. For more details, see The Principle Introduction of Probe Methods. The following are the explanations for the corresponding configuration items:

<table>
<thead>
<tr>
<th>Probe Method</th>
<th>Applicable Scenario</th>
<th>Configuration Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICMP Ping</td>
<td>Only probes network connectivity, and host reachability.</td>
<td>Probe Frequency</td>
<td>Required, with optional intervals of every 30 seconds, every 60 seconds, every 3 minutes, every 5 minutes, or every 10 minutes.</td>
</tr>
<tr>
<td>HTTP/HTTPS</td>
<td>Applicable for applications that require content recognition in requests, such as web applications and app services.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probe Frequency</td>
<td>Required, with optional intervals of every 30 seconds, every 60 seconds, every 3 minutes, every 5 minutes, or every 10 minutes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td>Required, the full URL for health checks, for example: <a href="http://www.example.com/test">www.example.com/test</a>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probe Port</td>
<td>Required, defaulting to port 80. It is recommended not to modify this unless a specific port needs to be designated.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HTTP Method</td>
<td>Required, the HTTP method for health checks is by default HEAD, with options including GET or HEAD. If the HEAD method is used, the server returns only HTTP header information, which can reduce backend overhead and enhance request efficiency. The corresponding origin service must support HEAD. If the GET method is used, the origin service simply needs to support GET.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HTTP Status Code</td>
<td>Required, the origin is considered healthy when the status code matches the selected status codes. By default, this includes 2XX, with options to select: 1XX, 2XX, 3XX, 4XX, 5XX.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow Redirects</td>
<td>Disabled by default. When enabled, the probe node will initiate another probe based on the 301/302 redirect address responded by the origin. It uses the status code of the final redirection response as the determination result for the health status. Up to 3 redirects are supported.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Custom Request Headers</td>
<td>Optional, custom request headers can be configured to be sent with the health check requests to the origin, with a maximum of 8 configurations allowed, for example: host: <a href="http://www.example.com">www.example.com</a>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCP</td>
<td>Suitable for scenarios where high reliability and data accuracy are essential, but</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probe Frequency</td>
<td>Required, with optional intervals of every 30 seconds, every 60 seconds, every 3 minutes, every 5 minutes, or every 10 minutes.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Probe Port | Required, defaulting to port 80. It is
<table>
<thead>
<tr>
<th>Transmission Speed</th>
<th>UDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suitable for scenarios where high transmission efficiency is crucial and a relatively lower level of accuracy is acceptable, such as instant messaging and online video streaming.</td>
<td>transmission speed is of lesser importance, such as file transfers and remote log-ins.</td>
</tr>
</tbody>
</table>

| **Probe Frequency** | Required, with optional intervals of every 30 seconds, every 60 seconds, every 3 minutes, every 5 minutes, or every 10 minutes. |
| **Probe Port** | Required, defaulting to port 80. It is recommended not to modify this unless a specific port needs to be designated. |
| **Probe Request** | Required, customize the content of the health check request, with a limit of 500 characters. |
| **Probe Response Result** | Required, customize the content of the health check request, with a limit of 500 characters. |

**Origin Health Determination Criteria**

Choose any of the probe policies: ICMP Ping, HTTP/HTTPS, TCP, and UDP. Click **Show Advanced Configuration** to configure origin health determination criteria. The following are the descriptions for each configuration item:
### Configuration Item

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timeout</strong> The allowed timeout duration for a single health check request to the origin. If no response is received within this period, the origin is considered Unhealthy. The default is 5 seconds, with a configurable range of ([1, 30]) seconds.</td>
</tr>
<tr>
<td><strong>Unhealthy Threshold</strong> The number of consecutive probe failures required to determine an origin Unhealthy. Once this threshold is reached, the origin is considered Unhealthy. The default is 2 times, with a configurable range of ([1, 5]). For example, if this value is set to 2, and an origin is initially Healthy, upon receiving two consecutive Unhealthy probe results, the origin will be considered Unhealthy.</td>
</tr>
<tr>
<td><strong>Healthy Threshold</strong> The number of consecutive successful probes required to restore an origin to a Healthy state, making it available again. The default is 3 times, with a configurable range of ([1, 5]). For example, if this value is set to 3, and an origin is Unhealthy, after three consecutive Healthy probe results, the origin will be restored to a Healthy status.</td>
</tr>
</tbody>
</table>

**Active Probing Cycle for Origin Health Status Change**
For example, suppose the health determination conditions for the origin are set as follows: timeout of 5 seconds, unhealthy threshold of 3 times, healthy threshold of 3 times, and a probe interval of every 30 seconds. The time required to consider an origin Unhealthy would then be: 5+30+5+30+5=75 seconds. The time required to restore the origin to a Healthy state (assuming a successful active probe response takes 1 second) would be: 1+30+1+30+1=63 seconds.

More Information

Introduction to Principle of Probe Method

ICMP Ping
HTTP/HTTPS
TCP
UDP
The ICMP Ping health check mechanism is as follows:

1. EdgeOne probe node sends a Ping command to your origin.

2. If the Ping is successful, and within the backsource timeout period, the origin receives an ICMP reply, the service is considered normal, and the result of this check is considered healthy.

3. If the Ping fails, and within the backsource timeout period, the probe node does not receive an ICMP reply from the origin, the service is considered abnormal, and the result of this check is considered unhealthy.

**Note:**

ICMP Ping requires your origin to support Ping.
The HTTP/HTTPS health check mechanism is as follows:

1. EdgeOne probe node sends an HTTP request to your origin. It requires configuration of the corresponding URL and port, with the option to include a custom HOST header.

2. If, within the backsource timeout period, the EO probe node receives an HTTP status code from the origin that successfully matches the configured HTTP status codes, the result of this check is considered healthy.

3. If, within the backsource timeout period, the EO probe node does not receive a response from the origin or receives a status code that does not match the configured codes, the result of this check is considered unhealthy.
The TCP health check mechanism is as follows:

1. EdgeOne probe node sends a SYN connection request packet to a specific port (configurable) on your origin.
2. Upon receiving the SYN request packet, if the corresponding port on the origin is in a normal listening state, it will respond with a SYN+ACK packet.
3. If, within the backsource timeout period, the probe node receives a SYN+ACK response packet from the origin, it indicates that the service is running normally. The result of this check is considered healthy. The probe node then replies with an ACK packet to the origin and sends an RST reset packet to terminate the TCP connection.
4. If, within the backsource timeout period, the probe node does not receive a SYN+ACK response packet from the origin, it indicates that the service is running abnormally. The result of this check is considered unhealthy. The probe node sends an RST reset packet to the origin to terminate the TCP connection.
The UDP health check mechanism is as follows:

1. EdgeOne probe node sends a customized probe packet to a specific port (configurable) on your origin.
2. If, within the backsource timeout period, the probe node receives a customized response packet from the origin, it indicates that the service is running normally. The result of this check is considered healthy.
3. If, within the backsource timeout period, the probe node does not receive a customized response packet from the origin or receives a response packet that does not conform to the defined content, it indicates that the service is running abnormally. The result of this check is considered unhealthy.

**Note:**
Both the request content and response content are customized, and you need to configure the corresponding request-response content on your origin.

**Probe Request Identification**

Active probes do not carry special request identifiers. When you choose ICMP Ping or TCP probes, there are no related features. When you choose UDP probes, customized content can be configured to serve as identifiers. For HTTP/HTTPS probes, separate customized request headers can be configured to serve as identifiers.
Viewing the Health Status of Origin Server

The node probe results will display the outcomes of probes initiated by EdgeOne from various nodes and regions within the global availability zones towards the current origin group. Users can view these probe results to find whether the origin is healthy across different zones.

Note:
EdgeOne Load Balancing feature is in beta testing. If you want to use it, you can Contact Us.

1. Log in to the EdgeOne console. In the left menu bar, click the Site List. Within this list, click the Site need to be configured to go to the details page.
2. On the site details page, click Origin Settings > Load Balancing.
3. On the Load Balancing page, click the desired Load balancer.

4. In the instance details page, click View details.

5. In the node probe results, nodes are differentiated by the following three colors:

**Green Node:** Indicates that the probe node in the region has considered all origins in the origin group to be healthy.

**Red Node:** Indicates that the probe node in the region has considered one or more origins in the origin group to be unhealthy.
**Gray Node:** Indicates that the probe node in the region cannot probe any origins. Probing is done at the IP level, meaning if the origin is a domain, the domain will be resolved into an IP before it probes. This situation usually occurs if you have entered an incorrect domain origin which cannot be resolved into an IP. In this case, it is recommended to check for potential spelling mistakes in the origin domain or whether the corresponding domain has expired.

<table>
<thead>
<tr>
<th>Detection results</th>
<th>Number of total nodes</th>
<th>Number of nodes with all origins healthy</th>
<th>Number of nodes with unhealthy origin</th>
<th>Number of nodes where no origin is detected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17</td>
<td>16</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Note:**

Probe nodes in different regions make independent decisions. Edge nodes will route requests back to the origin based on the probe results from the nearest probe nodes in each region.

For example: If your origin is in Hong Kong (China), and the probe node in Singapore considers the origin to be unhealthy whereas the probe node in Germany considers it to be healthy, traffic from the Singapore region will not be routed to that origin, while traffic from the Germany region will continue to be directed to that origin.

In the scenario described above, you can refer to the probe results from other regions for a comprehensive view. If only a few nodes consider the origin to be unhealthy, it might be due to network fluctuations in certain areas. If the majority of nodes consider the origin to be unhealthy, it is recommended to check whether the origin has malfunctioned.
This document introduces the relevant concepts involved in Load Balancing.

**Note:**
Tencent Cloud EdgeOne Load Balancing feature is in beta testing. If you want to use it, you can [Contact Us](#).

### Load Balancer

A load balancer is a virtual concept, comprising Origin groups and health check policies. Within a single load balancer, up to ten origin groups can be configured in priority order, accompanied by one health check policy. The load balancer intelligently directs business traffic based on probe results and the configured traffic scheduling policy.
Origin Group

An origin group is the smallest unit of origin configuration within the Load Balancing. You can add one or more origins. When you add multiple Origin Servers, you can configure weights to adjust traffic load. For more details, see Origin Group Configuration.

Health Check Policy

The health check policy consists of probe methods and health assessment criteria. Currently, four probe methods are supported: ICMP Ping, HTTP/HTTPS, TCP, and UDP. For more details, see Detailed Health Checks.

Traffic Scheduling Policy

The traffic scheduling policy only takes effect when the health check policy is enabled. Currently, it supports a Failover-by-Priority-Order policy, that is, based on probe results, it disables failed origins and routes traffic to healthy ones according to the priority order of origin groups.

Request Retry Policy

In the event of a request failure to a particular origin during normal business operations, the Load Balancing feature, guided by its request retry policy, can schedule the request to another origin for a retry. This helps reduce business request failures due to network issues or origin malfunctions. For more details, see the Introduction to Request Retry Policy.
Introduction to Request Retry Strategy

Last updated : 2024-05-29 10:33:37

Load Balancing is capable of redirecting a request to an alternative origin for retrial when a request to an initially designated server fails, in accordance with the retry policy. This reduces business request failures caused by network issues or origin fails.

**Note:**

EdgeOne Load Balancing feature is in beta testing. If you want to use it, you can Contact Us.

Actual business requests may fail due to the following reasons:

1. **Origin Failure and Have Not Actively Probe to Disable:** After health check policies are configured, active probe is conducted periodically. Traffic is directed based on the results of the previous probe until new results are available. If an origin becomes unhealthy between two probes, business traffic might still be directed to the unhealthy origin. This leads to business request failure.

2. **Network Jitter:** The origin is healthy, but network issues occur during the access. This leads to business request failure.

**Note:**

Request failures include origin connection establishment failures and origin response reception failures.

For the situations mentioned above, EdgeOne provides the following two fallback request retry policies:

**Policy 1:** When a real business request fails to access a certain origin, it directly retries with another origin within the next lower priority origin group. This is suitable for scenarios where the performance of both higher and lower priority origin groups is similar.

**Policy 2:** When a real business request fails to access a certain origin, it directly retries with another origin within the same priority origin group. This is suitable for scenarios where the performance of the higher priority origin group is
significantly better than that of the lower priority origin group.
Origin Group Configuration

Last updated: 2023-12-15 09:51:57

Overview

Manage business origins in the form of origin groups. The origin groups configured here can be used in functions such as adding acceleration domain names and L4 proxy.

Create Origin Group

1. Log in to the EdgeOne console and click Site List in the left sidebar. In the site list, click the target site to enter the site details page.
2. On the site details page, click origin configuration > origin group.
3. Click Create origin group.
4. Fill in the origin group name and select the origin type. The specific type descriptions are as follows:
   - **HTTP Dedicated**: Supports adding IP/domain name origins and object storage origins, and can only be used for site acceleration-related services (e.g., Domain Name Service and rule engine - Modify origin).
   - **Universal**: Only supports adding IP/domain name as origin, does not support adding object storage origin, and can be used for site acceleration services (such as Domain Name Service and rule engine) and L4 proxy.

**Note:**
After the configuration is complete, the origin group type cannot be modified.
5. Click the **Add Origin** button to configure the origin. The supported origin types are as follows, with up to 100 origins supported.

**Object storage origin:** Tencent Cloud COS or other object storage buckets compatible with **AWS S3**.

**IP/domain name origin:** Supports IPv4 addresses, IPv6 addresses, and domain names as origins.

**Note:**

Explanation of weight-related configurations in the origin group:

1. If a weight is set for an origin in the origin group, all origins in the group must also set corresponding weights. Weights can be integers between 0 and 100. If the weight of an origin is set to 0, no origin-pull requests will be allocated to that origin. Other non-zero weight origins will be allocated origin-pull requests based on their respective weight ratios.

2. If you do not set a weight, all origins in the origin group should not set weights at the same time. In this case, if “smart acceleration” is not enabled, EdgeOne will distribute origin-pull requests equally to each origin. If “smart acceleration” is enabled, EdgeOne will select the best quality origin for each origin-pull request.
6. Click **Create** to complete the origin group creation.
Id Version Origin Group Compatible Related Issues

Last updated: 2023-10-24 15:45:49

The origin group has carried out a product capability upgrade since October 24, 2023. After the upgrade, the old version of the origin group will be processed for compatibility in the following ways. At the same time, we also suggest you switch to the usage of the new version of the origin group.

Origin type & Configuration method compatibility

The new version of the origin group will no longer distinguish between self-owned origin, object storage origin, and Tencent Cloud COS type origin. The original origin groups with origin type of object storage and Tencent Cloud COS will be automatically updated to the new version of dedicated HTTP origin group, and the original origin groups with self-owned origin type will be automatically updated to the universal origin group.

The origin group will no longer support the configuration of origin-pull by region/protocol. If you have previously configured related origin-pull rules by region/protocol, the rules will be migrated to the rule engine as shown below:
### modify origin-http/https

<table>
<thead>
<tr>
<th>IF</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HOST</strong></td>
<td>Is</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IF</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Request protocol</strong></td>
<td>Is</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>IF</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Request protocol</strong></td>
<td>Is</td>
</tr>
</tbody>
</table>

Origin group port migration description

The new version of the origin group will no longer support port configuration. All port configurations will be migrated to the service configuration entry, such as L4 proxy or Domain Management.
Primary and Standby Origin Configuration Instructions

In the **Domain Management** and **Rule Engine - Modify Origin**, directly configuring primary and standby origins is no longer supported. Existing configurations will not be affected, but modifications are no longer supported. If you currently have a demand for primary and standby origin configurations, please contact us for support.
Collect EdgeOne origin-pull node IP

Last updated: 2023-12-15 09:54:21

To obtain the EdgeOne’s IPs for requesting origin, which can be used to set these EdgeOne’s IPs as an allowlist in the origin firewall, only allowing fixed source (IP) requests to the origin, thus implementing protection for the origin.

Use Guide

1. Directly access [https://api.edgeone.ai/ips](https://api.edgeone.ai/ips) through a browser or curl command. This will collect all IPv4 and IPv6 origin-pull node IP addresses of EdgeOne in the Global availability zone. The responded result is UTF-8 encoded plain text, with one IP segment per line.

2. If you only need to obtain the origin-pull node IP of a specified region or a specified IP Type, you can also filter the origin-pull node IP by carrying a specified query string. The supported query strings are as follows:

<table>
<thead>
<tr>
<th>Query string</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>version</td>
<td>Specifies the Type of origin-pull node IP address to be collected, with the following values: &lt;br&gt; v4: All IPv4 origin-pull node IP addresses &lt;br&gt; v6: All IPv6 origin-pull node IP addresses &lt;br&gt; By default, all IPv4 and IPv6 IP addresses are returned when this parameter is not included.</td>
</tr>
<tr>
<td>area</td>
<td>Specifies the region of the origin-pull node IP to be collected, with the following values: &lt;br&gt; global: All origin-pull node IP addresses in the Global availability zone &lt;br&gt; mainland-china: All origin-pull node IP addresses in the Chinese mainland availability zone &lt;br&gt; overseas: All origin-pull node IP addresses in the Global availability zone (excluding Chinese mainland) &lt;br&gt; By default, all origin-pull node IP addresses in the Global availability zone are returned when this parameter is not included.</td>
</tr>
</tbody>
</table>

Note:

1. In general, it is suggested to select and obtain the origin-pull node IP address of the corresponding region based on your site’s service area. For example, if the site’s service area is in the Chinese mainland availability zone, you can obtain the origin-pull node IP address within the Chinese mainland availability zone, and the same applies to other service areas.

2. This function is mutually exclusive with the origin protection function. If you need to collect the latest origin-pull IP through the method in this Document, please confirm that you have turned the business origin protection function Off.
Request example

If you want to obtain all IPv6 origin-pull node IPs in the Global availability zone (excluding Chinese mainland), you can query the origin-pull IP address under this condition by carrying `version=v6&area=overseas` in the request URL. The specific URL is:  

https://api.edgeone.ai/ips?version=v6&area=overseas

The response result example is as follows (this result is for reference only, please refer to the real-time request result for the actual origin-pull IP):
240d:c010::/28
2001:ee0:324b:100::/64
2405:3200:101:63::/64
2405:4800:a601::/64
2602:ffe4:c02:1001::/64
2602:ffe4:c12:101::/64
2602:ffe4:c12:105::/64
2602:ffe4:c15:124::/64
2602:ffe4:c18:c003::/64
2602:ffe4:c18:c201::/64
2602:ffe4:c18:c203::/64
2602:ffe4:c27:1003::/64
2604:980:4002:2::/64
2604:980:5003:2::/64
2604:980:7002:6::/64
2a02:b60:2001::/64
Host Header Rewrite

Last updated: 2023-10-11 10:28:05

Overview

Host header rewriting enables you to rewrite the host header to the actual origin domain when the origin domain is different from the acceleration domain in the load balancing task.

Directions

1. Log in to the EdgeOne console. Select Origin Configuration > Rule Engine on the left sidebar.
2. On the rule engine page, select the target site and click

   ![+] to configure host header rules as needed.
3. On the rule engine page, select Host for the match type and Rewrite host header for the action, and configure other parameters as needed. Click Save and publish or Save only.

Note

Supported match types: "Host".
Range GETs

Overview

Range GETs can be enabled to reduce the consumption of large file origin-pulls and response time.

Why can Range GETs improve the efficiency of large file delivery?

When caching large files, nodes will split them into smaller parts in order to improve cache efficiency. All parts cached expire at the same time and follow the node cache TTL configuration. Range requests are also supported. For example, if a client request carries the HTTP header `Range: bytes = 0-999`, only the first 1000 bytes of the file will be returned to the user.

If Range GETs is enabled: When parts of the file are requested and their caches expire, nodes only pull and cache the requested parts and return them to the user, so that origin-pull consumption and response time are greatly reduced.

If Range GETs is disabled, when the client requests only parts of a file, the node will pull only the requested parts according to the `Range` header in the client request, cache them, and return them to the client at the same time. However, this may not be able to achieve the optimal performance. In large file scenarios, we recommend you enable Range GETs.

Use Cases

You can use Range GETs to cache large static files in either of the following cases: The origin server supports Range requests, or you use a Tencent Cloud COS origin server and do not apply any data processing methods such as image processing.

Notes

The origin server must support Range requests, or the origin-pull may fail.

The origin-pull may fail if Range GETs is enabled for small static files, or if you enable it while using a Tencent Cloud COS origin server and data processing methods such as image processing.

Directions
For instance, you have a video service website that provides online video watching through `video.example.com`. The videos are mainly long videos with large files. In order to reduce traffic consumption of large files and improve origin-pull speed, you need to support range requests and origin-pull. You can perform the following steps:

1. Log in to the EdgeOne console, click Site List in the left sidebar and click the Site to be configured in the site list.
2. On the site detail page, click on Rule Engine.
3. On the rule engine management page, click on Create rule to enter the new rule's editing page. In this scenario, you can operate as follows:
   3.1 On the rule editing page, select the Matching type as HOST equals `video.example.com`.
   3.2 Click on Action, in the displayed operation list, choose the operation as Range GETs.
   3.3 Click on On/Off to enable Range GETs.
4. Click on Save and publish to complete the configuration of this rule.
Configuring Origin-Pull HTTPS

Last updated : 2023-12-13 11:10:42

Overview

You can specify the protocol that EdgeOne uses in the origin-pull request. In scenarios that require a high level of security, HTTPS can be used to access a website to ensure the security of website data. When HTTPS is specified as the origin-pull protocol, all origin-pull requests from EdgeOne to the origin use HTTPS, which prevents data tampering or theft during transmission.

In scenarios where fast response is required, HTTP can be used for origin-pull requests to speed up website access. When HTTP is specified as the origin-pull protocol, you can avoid complex SSL handshakes and other operations between EdgeOne and the origin, thus improving the website access speed. If your origin does not support HTTPS, please select HTTP.

1. An EdgeOne node initiates an origin-pull request by using the specified origin-pull protocol.
2. The origin responds to the request and establishes a connection by using the same protocol as the request.

Note:
The configuration priority of the rule engine is superior. If the origin protocol rule is configured simultaneously within the domain name service and the rule engine, the final standard is determined by the rule engine.

Scenario 1: Configuring origin-pull HTTPS for multiple domain names in batches in the rule engine

If you need to uniformly change the origin-pull protocol to origin-pull HTTPS for multiple domain names, such as www.example.com, vod.example.com and image.example.com, please refer to the following steps:

1. Log in to the EdgeOne console and click Site List in the left sidebar. In the site list, click the target site.
2. On the site details page, click Rule Engine.
3. On the rule engine management page, click Create rule.
4. On the rule editing page, enter the rule name and select Host matching type to match the request of the specified domain name. In the current scenario, select domain names www.example.com, vod.example.com, and image.example.com.

5. Click on **Action > Select Box**, select **Origin-pull HTTPS** from the dropdown action list that appears.

6. Click on **Save and Publish** to finalize this rule configuration.

**Scenario 2: Configuring origin-pull HTTPS for the specified domain name**

If you need to specify a domain name to modify the origin-pull protocol into origin-pull HTTPS, such as www.example.com, please follow these steps:

1. Log in to the **EdgeOne** console, click **Site List** on the left sidebar, and click the **Site** to be configured in the site list,
2. Choose **Domain Name Service > Domain Name Management** on the **Site Details** page.
3. Select the domain name that needs to be modified and click **Edit** on the **Domain Management** page.

4. In the origin-pull protocol, select **HTTPS** and click **Complete** to finish the modification.
## Edit domain name

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain name</td>
<td></td>
</tr>
<tr>
<td>Origin type</td>
<td>IP/Domain name</td>
</tr>
<tr>
<td>Origin (IP/Domain name)</td>
<td></td>
</tr>
<tr>
<td>IPv6 access</td>
<td>Follow site configuration: Disable</td>
</tr>
<tr>
<td>Origin Protocol</td>
<td>Follow protocol</td>
</tr>
<tr>
<td>Origin Port</td>
<td>HTTP: 80, HTTPS: 443</td>
</tr>
</tbody>
</table>
HTTP/2 Origin-Pull

Last updated : 2023-08-30 15:05:42

Overview

Support EdgeOne nodes to origin-pull using HTTP/2 protocol. HTTP/2 (i.e., HTTP 2.0, Hypertext Transfer Protocol version 2) is the second major version of the HTTP protocol, which can effectively reduce network latency and improve site page loading speed.

Note :
1. When enabled, the origin must support HTTP/2 protocol access.
2. If you need to configure HTTP/2 access, please refer to HTTP/2.

Directions

If you need to enable or disable HTTP/2 origin-pull for the specified domain www.example.com, you can follow the steps below:

1. Log in to the EdgeOne console, click on the site list in the left sidebar, and click on the site you need to configure within the site list.
2. On the site details page, click on the rule engine.
3. On the rule engine management page, click on create rule to enter the new rule editing page.
   3.1 On the rule editing page, select the matching type as HOST equals www.example.com.
   3.2 Click on the action, and in the pop-up operation list, select the operation as HTTP/2 origin-pull.
3.3 Click on the switch to enable/disable HTTP/2 origin-pull.
4. Click on Save and Publish to complete the rule configuration.
Redirect Following During Origin-Pull

Last updated: 2023-08-30 15:08:04

Overview

Under normal circumstances, when the origin returns a 301/302 request, the node will return the status code to the client by default, and the client will redirect to the corresponding resources for access. EdgeOne supports follow origin redirects. When enabled, if the node receives a 301/302 status code during origin-pull, it will actively follow the redirect (not exceeding the set maximum redirects) to the specified address until the corresponding file is obtained, and then respond to the client with the actual resources, which can improve the user's access response speed.

For example: The client accesses the URL https://a.example.com/test.jpg, the origin A redirects the URL 302 to https://b.example.com/test.jpg, and the domain a.example.com has accessed the EdgeOne Service, while b.example.com has not yet accessed the acceleration service. Then:

**Without enabling origin-pull follow redirect:** After the client initiates the visit, if there is no cache in the EdgeOne node, it will visit the origin A and receive the 302 status code, and then respond to the client with the status code, and the client will directly request the origin B for the corresponding resources. At this time, since the origin B has not accessed the acceleration service, the client's self-initiated access speed is slower, and the obtained file cannot be cached. When other users access the same file, the process needs to be repeated.

**Enable origin-pull follow redirect:** After the client initiates the visit, if there is no cache in the EdgeOne node, it will visit the origin A and receive the 302 status code, and then, according to the status code and the corresponding address, directly request the origin B for the corresponding resources, and cache the resources in the node. This process is carried out by the EdgeOne node for origin-pull requests, the request speed is faster, and the obtained file can be cached in the node. When other users access the same file, there is no need to repeat the origin-pull, and the file can be directly hit and responded to the client.
Directions

For example: If you need to enable origin-pull follow redirect for the specified domain `www.example.com`, with a maximum of 3 redirects. You can refer to the following steps:

1. Log in to the EdgeOne console, click on the site list in the left menu bar, and click on the site to be configured in the site list.
2. On the site details page, click on the rule engine.
3. On the rule engine management page, click Create Rule to enter the editing page of the new rule. In this scenario, you can follow the steps below:
   3.1. On the rule editing page, select the matching type as HOST equals `www.example.com`.
   3.2. Click on the action, and in the pop-up operation list, select the operation as follow origin redirect.
   3.3. Click on the switch, click on the switch to enable, and set the maximum redirects to 3 times. The related configuration instructions are as follows:

Maximum redirects: You can set it between 1-5 times. Within the maximum redirects, the node will follow the redirect address until the corresponding resources are obtained. If the maximum redirects are exceeded, the corresponding status code will be directly responded to the client.
4. Click on **Save and Publish** to complete the rule configuration.
Controlling Origin-pull Requests

Last updated : 2023-08-30 15:09:23

Overview

By default, when origin-pulling, all query strings and Cookies within the request will be retained. If your business origin only allows specified query strings or Cookie information to be carried in the origin-pull request, you can ensure the normal origin-pull request by deleting the specified origin-pull request parameters.

Directions

For example, Client requests Request URL: `http://www.example.com/path/demo.jpg?key1=a&key2=b&key3=c&key4=d`, and only `key1=a` parameter needs to be retained when origin-pulling. You can follow the steps below to configure:

1. Log in to the EdgeOne console, click on the site list in the left menu bar, and click on the site to be configured within the site list.
2. On the site details page, click on the rule engine.
3. On the rule engine Management page, click Create rule to enter the edit page of the new rule.
   3.1. On the rule edit page, select the matching type as HOST equals `www.example.com`.
   3.2. Click on the action, and in the pop-up operation list, select the operation as origin-pull request parameter settings.
   3.3. Select the mode as retaining specified parameters, Enter the parameters `key1` and `key2` to be retained, up to 10 parameters are allowed.
4. Click **Save and Publish** to complete the rule Configuration.