Administering the API Gateway

Intermediate to advanced goals to manage your installation.
Intermediate to advanced goals to manage your installation.
Abstract

This guide provides the information regarding configuration tasks, which can be performed after the basic installation.
# Table of Contents

**PREFACE**  ........................................................................................................ 7

**PART I. THE API GATEWAY**  ......................................................................... 8

**CHAPTER 1. OPERATING APICAST**  ............................................................. 9

1.1. PUBLIC BASE URL  .................................................................................. 9
1.2. MAPPING RULES  .................................................................................... 10
    1.2.1. Mapping rules in products and backends  .................................. 10
    1.2.2. Matching of mapping rules  ......................................................... 11
    1.2.3. Mapping rules workflow  ............................................................. 12
1.3. HOST HEADER  ......................................................................................... 13
1.4. PROTECTING YOUR API BACKEND  .................................................... 13
1.5. USING APICAST WITH PRIVATE APIS  ............................................... 13
1.6. CONFIGURING APICAST WITH OPENTRACING  .................................. 13
    1.6.1. Prerequisites  ............................................................................... 14
    1.6.2. Procedure  ................................................................................... 14
    1.6.3. Additional information  ............................................................... 14
    1.6.4. Installing Jaeger on your OpenShift instance  .......................... 14

**CHAPTER 2. OPERATING DOCKER-CONTAINERIZED ENVIRONMENTS**  ...... 17

2.1. TROUBLESHOOTING APICAST ON THE DOCKER-CONTAINERIZED ENVIRONMENT  ................................................................................................. 17
    2.1.1. Cannot connect to the Docker daemon error  ............................ 18
    2.1.2. Basic Docker command-line interface commands  .................. 18

**CHAPTER 3. ADVANCED APICAST CONFIGURATION**  ................................ 18

3.1. DEFINE A SECRET TOKEN  ....................................................................... 18
3.2. CREDENTIALS  ....................................................................................... 19
3.3. CONFIGURING ERROR MESSAGES  ..................................................... 19
3.4. CONFIGURATION HISTORY  ................................................................. 20
3.5. DEBUGGING  .......................................................................................... 20
3.6. PATH ROUTING  .................................................................................... 21

**CHAPTER 4. APICAST POLICIES**  .............................................................. 22

4.1. APICAST STANDARD POLICIES  ...................................................... 22
    4.1.1. 3scale Auth Caching policy  ....................................................... 23
    4.1.2. 3scale Batcher policy  ................................................................. 24
    4.1.3. Anonymous Access policy  .......................................................... 25
    4.1.4. CORS Request Handling policy  ............................................... 26
    4.1.5. Echo policy  ............................................................................... 28
    4.1.6. Edge Limiting policy  ................................................................. 28
        4.1.6.1. Types of limits  ................................................................. 29
        4.1.6.2. Limit definition  ............................................................... 29
        4.1.6.3. Liquid templating  ......................................................... 30
        4.1.6.4. Applying conditions  ..................................................... 30
        4.1.6.5. Configuring the store  ..................................................... 31
        4.1.6.6. Error handling  ............................................................. 31
    4.1.7. Header Modification policy  ....................................................... 31
    4.1.8. IP Check policy  ........................................................................ 31
    4.1.9. JWT Claim Check policy  ............................................................ 33
        4.1.9.1. About JWT Claim Check policy  .................................. 34
        4.1.9.2. Configuring JWT Claim Check policy in your policy chain 34
            4.1.9.2.1. Prerequisites:  ....................................................... 34
4.1.22.2. Configuring Upstream Connection in your policy chain
4.1.22.2.1. Prerequisites:
4.1.22.2.2. Configuring the policy
4.1.23. URL Rewriting policy
4.1.23.1. Commands for rewriting the path
4.1.23.2. Commands for rewriting the query string
4.1.24. URL Rewriting with Captures policy
4.2. ENABLING A STANDARD POLICY
4.3. CREATING CUSTOM APICAST POLICIES
4.4. ADDING CUSTOM POLICIES TO APICAST
4.4.1. Adding custom policies to the built-in APIcast
4.4.2. Adding custom policies to APIcast on another OpenShift Container Platform
4.5. CREATING A POLICY CHAIN IN 3SCALE
4.6. CREATING A POLICY CHAIN JSON CONFIGURATION FILE

CHAPTER 5. INTEGRATING A POLICY CHAIN WITH APICAST NATIVE DEPLOYMENTS 
5.1. USING VARIABLES AND FILTERS IN POLICIES

CHAPTER 6. TRANSFORMING 3SCALE MESSAGE CONTENT USING POLICY EXTENSIONS IN FUSE
6.1. INTEGRATING APICAST WITH APACHE CAMEL TRANSFORMATIONS IN FUSE
6.2. CONFIGURING AN APICAST POLICY EXTENSION CREATED USING APACHE CAMEL IN FUSE ON OPENSHIFT

CHAPTER 7. APICAST ENVIRONMENT VARIABLES
all_proxy, ALL_PROXY
APICAST_ACCESS_LOG_FILE
APICAST_BACKEND_CACHE_HANDLER
APICAST_CONFIGURATION_CACHE
APICAST_CONFIGURATION_LOADER
APICAST_CUSTOM_CONFIG
APICAST_ENVIRONMENT
APICAST_EXTENDED_METRICS
APICAST_HTTPS_CERTIFICATE
APICAST_HTTPS_CERTIFICATE_KEY
APICAST_HTTPS_PORT
APICAST_HTTPS_VERIFY_DEPTH
APICAST_LOAD_SERVICES_WHEN_NEEDED
APICAST_LOG_FILE
APICAST_LOG_LEVEL
APICAST_MANAGEMENT_API
APICAST_MODULE
APICAST_OIDC_LOG_LEVEL
APICAST_PATH_ROUTING
APICAST_POLICY_LOAD_PATH
APICAST_PROXY_HTTPS_CERTIFICATE
APICAST_PROXY_HTTPS_CERTIFICATE_KEY
APICAST_PROXY_HTTPS_PASSWORD_FILE
APICAST_PROXY_HTTPS_SESSION_REUSE
APICAST_REPORTING_THREADS
APICAST_RESPONSE_CODES
APICAST_SERVICE_${ID}_CONFIGURATION_VERSION
APICAST_SERVICES_LIST
APICAST_SERVICES_FILTER_BY_URL
APICAST_UPSTREAM_RETRY_CASES
CHAPTER 8. CONFIGURING APICAST FOR BETTER PERFORMANCE ............................ 86

8.1. GENERAL GUIDELINES 86
8.2. DEFAULT CACHING 86
8.3. ASYNCHRONOUS REPORTING THREADS 88
8.4. 3SCALE BATCHER POLICY 88

CHAPTER 9. EXPOSING 3SCALE APICAST METRICS TO PROMETHEUS ................. 91

9.1. ABOUT PROMETHEUS 91
9.1.1. Prometheus queries 91
9.2. APICAST INTEGRATION WITH PROMETHEUS 91
9.2.1. Additional options 92
9.3. OPENSSHIFT ENVIRONMENT VARIABLES FOR 3SCALE APICAST 92
9.4. 3SCALE APICAST METRICS EXPOSED TO PROMETHEUS 93

PART II. API VERSIONING ................................................................. 96

CHAPTER 10. API VERSIONING .......................................................... 97

10.1. GOAL 97
10.2. PREREQUISITES 97
10.3. URL VERSIONING 97
10.4. ENDPOINT VERSIONING 100
10.5. CUSTOM HEADER VERSIONING 100

PART III. API AUTHENTICATION ...................................................... 101

CHAPTER 11. AUTHENTICATION PATTERNS .......................................... 102

11.1. SUPPORTED AUTHENTICATION PATTERNS 102
11.2. SETTING UP AUTHENTICATION PATTERNS 102
11.2.1. Select the authentication mode for your service 102
11.2.2. Select the Authentication mode you want to use 103
11.2.3. Ensure your API accepts the correct types of credentials 103
11.2.4. Create an application to test credentials 103
11.3. STANDARD AUTHENTICATION PATTERNS 103
11.3.1. API key 103
11.3.2. App_ID and App_Key pair 104
11.3.3. OpenID Connect 104
11.4. REFERER FILTERING 104

CHAPTER 12. OPENID CONNECT INTEGRATION ...................................... 108

12.1. JWT VERIFICATION AND PARSING BY APICAST 108
12.2. CLIENT CREDENTIALS SYNCHRONIZATION BY ZYNC

12.3. CONFIGURE RED HAT SINGLE SIGN-ON INTEGRATION
   12.3.1. Configure Zync to use custom CA certificates
   12.3.2. Configure Red Hat Single Sign-On
   12.3.3. Configure 3scale

12.4. CONFIGURE HTTP INTEGRATION WITH THIRD-PARTY IDENTITY PROVIDERS
   12.4.1. Pre-requisites
   12.4.2. Procedure
   12.4.3. Zync REST API example
      12.4.3.1. Prerequisites
      12.4.3.2. Creating, updating and deleting clients
      12.4.3.3. Payload
      12.4.3.4. Using OAuth2 authentication

12.5. OAUTH 2.0 SUPPORTED FLOWS
   12.5.1. How OAuth 2.0 supported flows work
   12.5.2. Configuring OAuth 2.0 supported flows

12.6. TEST THE INTEGRATION
   12.6.1. Test the client synchronization
   12.6.2. Test the API authorization flow

12.7. EXAMPLE OF THE INTEGRATION
PREFACE

This guide helps you to apply intermediate to advanced configuration features to your 3scale installation. For basic details regarding installation, refer to Installing 3scale.
PART I. THE API GATEWAY
CHAPTER 1. OPERATING APICAST

This section describes the concepts to consider when working with advanced APIcast configurations.

1.1. PUBLIC BASE URL

The **Public Base URL** is the URL that your developers use to make requests to your API product, which is exposed publicly with 3scale. This will be the URL of your APIcast instance.

If you are using one of the **Self-managed deployment options**, you can choose your own Public Base URL for each one of the environments provided (staging and production) on a domain name you are managing. Note that this URL should be different from the one of your API backend, and could be something like `https://api.yourdomain.com:443`, where `yourdomain.com` is the domain that belongs to you. After setting the Public Base URL make sure you save the changes and, if necessary, promote the changes in staging to production.

APIcast will only accept calls to the hostname specified in the Public Base URL. As an example for the **Echo API**, if you specify `https://echo-api.3scale.net:443` as the Public Base URL, the correct call would be:

```bash
curl "https://echo-api.3scale.net:443/hello?user_key=YOUR_USER_KEY"
```

In case you do not have a public domain for your API, you can use the APIcast IP address in the requests, but you still need to specify a value in the **Public Base URL** field even if the domain is not real. In this case, make sure you provide the host in the Host header. For example:

```bash
curl "http://192.0.2.12:80/hello?user_key=YOUR_USER_KEY" -H "Host: echo-api.3scale.net"
```

If you are deploying on local machine, you can also just use "localhost" as the domain, so the Public Base URL will look like `http://localhost:80`, and then you can make requests like this:

```bash
curl "http://localhost:80/hello?user_key=YOUR_USER_KEY"
```

In case you have multiple API products, set this Public Base URL appropriately for each product. APIcast will route the requests based on the hostname.

1.2. MAPPING RULES

Based on the requests to your API, mapping rules define the metrics or methods that you want to report. The following is an example mapping rule:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Pattern</th>
<th>Metric or Method</th>
<th>Last?</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>GET</td>
<td>/</td>
<td>hits</td>
<td>false</td>
<td>1</td>
</tr>
</tbody>
</table>

This rule means that any **GET** requests that start with `/` will increment the metric **hits** by 1. This rule will match any request to your API. But most likely, you will change this rule because it is too generic, and often leads to double counts when more specific rules are added.

The following rules for the Echo API show more specific examples:
1.2.1. Mapping rules in products and backends

Mapping rules work at the API product and API backend levels. In this section, you will learn about the behavior of mapping rules at each level, as well as see an example to describe how mapping rules operate.

Mapping rules at the product level

- The mapping rule takes precedence. This means that the product mapping rule is the first one to be evaluated.
- The mapping rule is always evaluated, independent of which backend receives the redirected traffic.

Mapping rules at the backend level

- When you add mapping rules to a backend, these are added to all the products bundling said backend.
- The mapping rule is evaluated after the mapping rules defined at the product level.
- The mapping rule is evaluated only if the traffic is redirected to the same backend the mapping rule belongs to.
- The path of the backend for a product is automatically prepended to each mapping rule of the backend bundled to said product.

Example of mapping rules with products and backends

For this example, we start with a backend and a product.

- The Echo API backend:
  - Has the private endpoint: https://echo-api.3scale.net
  - Contains 2 mapping rules with the following patterns:
    - /hello
    - /bye

- The Cool API product:
  - Has this public endpoint: https://cool.api
  - Uses the Echo API backend with this routing path: /echo.
Mapping rules with the following patterns are automatically part of the Cool API product:

/echo/hello
/echo/bye

- This means that a request sent to the public URL https://cool.api/echo/hello is redirected to https://echo-api.3scale.net/hello
- Similarly, a request sent to https://cool.api/echo/bye redirects to https://echo-api.3scale.net/bye

Now, consider an additional product called Tools For Devs using the same Echo API backend.

- The Tools For Devs product:
  - Has public endpoint: https://dev-tools.api
  - Uses the Echo API backend with the following routing path: /tellmeback.
- Mapping rules with the following patterns are automatically part of the Tools For Devs product:

  /tellmeback/hello
  /tellmeback/bye

  - Therefore, a request sent to the public URL https://dev-tools.api/tellmeback/hello is redirected to https://echo-api.3scale.net/hello
  - Similarly, a request sent to https://dev-tools.api/tellmeback/bye redirects to https://echo-api.3scale.net/bye

- If you add a mapping rule with the /ping pattern to the Echo API backend, both products - Cool API and Tools For Devs- are affected:
  - Cool API will have a mapping rule with this pattern: /echo/ping
  - Tools For Devs will have a mapping rule with this pattern: /tellmeback/ping

1.2.2. Matching of mapping rules

The matching of mapping rules is performed by prefix and can be arbitrarily complex. The notation follows OpenAPI and ActiveDocs specifications:

- A mapping rule must start with a forward slash (/).
- You can perform a match on the path over a literal string (for example, /hello).
- Mapping rules can include parameters on the query string or in the body (for example, /{word}?value={value}). APIcast fetches the parameters in the following ways:
  - GET method: From the query string.
  - POST, DELETE, or PUT method: From the body.
- Mapping rules can contain named wildcards (for example, /{word}). This rule will match anything in the placeholder {word}, making requests like /morning match the rule. Wildcards can appear between slashes or between slash and dot. Parameters can also include wildcards.
By default, all mapping rules are evaluated from first to last, according to the sort you specified. If you add a rule /v1, it will be matched for requests whose paths start with /v1 (for example, /v1/word or /v1/sentence).

You can add a dollar sign ($) to the end of a pattern to specify exact matching. For example, /v1/word$ will only match /v1/word requests, and will not match /v1/word/hello requests. For exact matching, you must also ensure that the default mapping rule that matches everything (/) has been disabled.

More than one mapping rule can match the request path, but if none matches, the request is discarded with an HTTP 404 status code.

1.2.3. Mapping rules workflow

Mapping rules have the following workflow:

- You can define a new mapping rule (see Add mapping rules).
- Mapping rules will be grayed out on the next reload to prevent accidental modifications.
- To edit an existing mapping rule, you must enable it first by clicking the pencil icon on the right.
- To delete a rule, click the trash icon.
- All modifications and deletions are saved when you promote the changes in Integration > Configuration.

Add mapping rules

To add a new mapping rule, perform the following steps:

1. Click Add Mapping Rule.
2. Specify the following settings:
   - **Verb**: The HTTP request verb (GET, POST, DELETE, or PUT).
   - **Pattern**: The pattern to match (for example, /hello).
   - **Metric or method to increment**: The metric or method name.
   - **Increment by**: The metric increment number (for example, 1).
   - **Last?**: If this mapping rule should be considered as the last one, to stop processing other mapping rules.
   - **Position**: Number that indicates the position of the execution of the mapping rule, to sort the mapping rules.
3. Click Create Mapping Rule to apply the changes.

Stop other mapping rules

To stop processing other mapping rules, you can select Last? when creating a new mapping rule. For example, if you have the following mapping rules defined in API Integration Settings and you have different metrics associated with each rule:
When calling with `(get) /path/to/example/search`, APIcast stops processing the remaining mapping rules and incrementing their metrics after the rule is matched.

### 1.3. HOST HEADER

This option is only needed for those API products that reject traffic unless the `Host` header matches the expected one. In these cases, having a gateway in front of your API product will cause problems since the `Host` will be the one of the gateway, e.g. `xxx-yyy.staging.apicast.io`

To avoid this issue you can define the host your API product expects in the `Host Header` field in the Authentication Settings, and the hosted APIcast instance will rewrite the host.

![Host Header](authentication_settings.png)

### 1.4. PROTECTING YOUR API BACKEND

Once you have APIcast working in production, you might want to restrict direct access to your API product without credentials. The easiest way to do this is by using the Secret Token set by APIcast. Please refer to the [Advanced APIcast configuration](#) for information on how to set it up.

### 1.5. USING APICAST WITH PRIVATE APIs

With APIcast it is possible to protect the APIs which are not publicly accessible on the Internet. The requirements that must be met are:

- Self-managed APIcast must be used as the deployment option.
- APIcast needs to be accessible from the public internet and be able to make outbound calls to the 3scale Service Management API.
- The API product should be accessible by APIcast.

In this case you can set your internal domain name or the IP address of your API in the `Private Base URL` field and follow the rest of the steps as usual. Note, however, that you will not be able to take advantage of the Staging environment, and the test calls will not be successful, as the Staging APIcast instance is hosted by 3scale and will not have access to your private API backend. But once you deploy APIcast in your production environment, if the configuration is correct, APIcast will work as expected.

### 1.6. CONFIGURING APICAST WITH OPENTRACING

OpenTracing is an API specification and method used to profile and monitor microservices. From version 3.3 onwards, APIcast includes OpenTracing Libraries and the [Jaeger Tracer library](#).
1.6.1. Prerequisites

To add distributed tracing to your APIcast deployment, you need to ensure the following prerequisites:

- Each external request should have a unique request ID attached, usually via a HTTP header.
- Each service should forward the request ID to other services.
- Each service should output the request ID in the logs.
- Each service should record additional information, like start and end time of the request.
- Logs need to be aggregated, and provide a way to parse via HTTP request ID.

1.6.2. Procedure

To configure OpenTracing, use the following environment variables:

- **OPENTRACING_TRACER**: To define which tracer implementation to use. Currently, only Jaeger is available.
- **OPENTRACING_CONFIG**: To specify the default configuration file of your tracer. You can see an example [here](https://github.com/3scale/apicast).
- **OPENTRACING_HEADER_FORWARD**: Optional. You can set this environment variable according to your OpenTracing configuration.

For more information about these variables, refer to APIcast environment variables.

To test if the integration is properly working, you need to check if traces are reported in the Jaeger tracing interface.

1.6.3. Additional information

The OpenTracing and Jaeger integration are available in the upstream project: [https://github.com/3scale/apicast](https://github.com/3scale/apicast)

1.6.4. Installing Jaeger on your OpenShift instance

This section provides information about the installation of Jaeger on the OpenShift instance you are running.

**WARNING**

Jaeger is a third-party component, which 3scale does not provide support for, with the exception of uses with APIcast. The following instructions are provided as a reference example only, and are not suitable for production use.

1. Install the Jaeger all-in-one in the current namespace:
2. Create a Jaeger configuration file `jaeger_config.json` and add the following:

```json
{
    "service_name": "apicast",
    "disabled": false,
    "sampler": {
        "type": "const",
        "param": 1
    },
    "reporter": {
        "queueSize": 100,
        "bufferFlushInterval": 10,
        "logSpans": false,
        "localAgentHostPort": "jaeger-agent:6831"
    },
    "headers": {
        "jaegerDebugHeader": "debug-id",
        "jaegerBaggageHeader": "baggage",
        "TraceContextHeaderName": "uber-trace-id",
        "traceBaggageHeaderPrefix": "testctx-"
    },
    "baggage_restrictions": {
        "denyBaggageOnInitializationFailure": false,
        "hostPort": "127.0.0.1:5778",
        "refreshInterval": 60
    }
}
```

- set a `sampler` constant of 1 to sample all requests
- set the location and queue size of the `reporter`
- set `headers`, including `TraceContextHeaderName` which we will use to track requests

3. Create a ConfigMap from our Jaeger configuration file and mount it into APIcast:

```bash
oc create configmap jaeger-config --from-file=jaeger_config.json
oc volume dc/apicast --add -m /tmp/jaeger/ --configmap-name jaeger-config
```

4. Enable OpenTracing and Jaeger with the configuration we have just added:

```bash
oc set env deploymentConfig/apicast OPENTRACING_TRACER=jaeger
OPENTRACING_CONFIG=/tmp/jaeger/jaeger_config.json
```

5. Find the URL the Jaeger interface is running on:

```bash
oc get route
(...)
jaeger-query-myproject.127.0.0.1.nip.io
```

6. Open the Jaeger interface from the previous step, which shows data being populated from Openshift Health checks.
7. The final step is to add OpenTracing and Jaeger support to your backend APIs so that you can see the complete request trace. This varies in each back end, depending on the frameworks and languages used. As a reference example, you can see Using OpenTracing with Jaeger to collect Application Metrics in Kubernetes.

For more information on configuring Jaeger, see:

- Jaeger on OpenShift development setup
- Jaeger on OpenShift production setup
- Distributed tracing on OpenShift Service Mesh
CHAPTER 2. OPERATING DOCKER-CONTAINERIZED ENVIRONMENTS

2.1. TROUBLESHOOTING APICAST ON THE DOCKER-CONTAINERIZED ENVIRONMENT

This section describes the most common issues that you can find when working with APIcast on a Docker-containerized environment.

2.1.1. Cannot connect to the Docker daemon error

The docker: Cannot connect to the Docker daemon. Is the docker daemon running on this host? error message may be because the Docker service hasn’t started. You can check the status of the Docker daemon by running the `sudo systemctl status docker.service` command.

Ensure that you are run this command as the root user because the Docker containerized environment requires root permissions in RHEL by default. For more information, see [here](#).

2.1.2. Basic Docker command-line interface commands

If you started the container in the detached mode (`-d` option) and want to check the logs for the running APIcast instance, you can use the log command: `sudo docker logs <container>`. Where, `<container>` is the container name ("apicast" in the example above) or the container ID. You can get a list of the running containers and their IDs and names by using the `sudo docker ps` command.

To stop the container, run the `sudo docker stop <container>` command. You can also remove the container by running the `sudo docker rm <container>` command.

For more information on available commands, see [Docker commands reference](#).
CHAPTER 3. ADVANCED APICAST CONFIGURATION

This section covers the advanced settings option of 3scale's API gateway in the staging environment.

3.1. DEFINE A SECRET TOKEN

For security reasons, any request from the 3scale gateway to your API backend contains a header called X-3scale-proxy-secret-token. You can set the value of this header in Authentication Settings on the Integration page.

Setting the secret token acts as a shared secret between the proxy and your API so that you can block all API requests that do not come from the gateway if you do not want them to. This adds an extra layer of security to protect your public endpoint while you are in the process of setting up your traffic management policies with the sandbox gateway.

Your API backend must have a public resolvable domain for the gateway to work, so anyone who knows your API backend can bypass the credentials checking. This should not be a problem because the API gateway in the staging environment is not meant for production use, but it is always better to have a fence available.

3.2. CREDENTIALS

The API credentials within 3scale are either user_key or app_id/app_key depending on the authentication mode that you are using. OpenID Connect is valid for the API gateway in the staging environment, but it cannot be tested in the Integration page.

However, you might want to use different credential names in your API. In this case, you need to set custom names for the user_key if you are using the API key mode:

Alternatively, for the app_id and app_key:
For instance, you could rename `app_id` to `key` if that fits your API better. The gateway will take the name `key` and convert it to `app_id` before doing the authorization call to the 3scale backend. Note that the new credential name has to be alphanumeric.

You can decide if your API passes credentials in the query string (or body if not a GET) or in the headers.

![Credentials Location]

**NOTE**

APIcast normalizes header names when extracting credentials. This means they are case insensitive, and underscores and hyphens are treated equally. For example, if you set the App Key parameter as `App_Key`, other values such as `app-key` are also accepted as valid app key headers.

### 3.3. CONFIGURING ERROR MESSAGES

This section describes how to configure APIcast error messages.

As a proxy, 3scale APIcast manages requests in the following ways:

- If there are no errors, APIcast passes the request from the client to the API back end server, and returns the API response to the client without modifications. In case you want to modify the responses, you can use the **Header Modification policy**.

- If the API responds with an error message, such as `404 Not Found` or `400 Bad Request`, APIcast returns the message to the client. However, if APIcast detects other errors such as **Authentication missing**, APIcast sends an error message and terminates the request.

Hence, you can configure these error messages to be returned by APIcast:

- **Authentication failed**: This error means that the API request does not contain the valid credentials, whether due to fake credentials or because the application is temporarily suspended. Additionally, this error is generated when the metric is disabled, meaning its value is 0.

- **Authentication missing**: This error is generated whenever an API request does not contain any credentials. It occurs when users do not add their credentials to an API request.

- **No match**: This error means that the request did not match any mapping rule and therefore no metric is updated. This is not necessarily an error, but it means that either the user is trying random paths or that your mapping rules do not cover legitimate cases.
• Usage limit exceeded: This error means that the client reached its rate limits for the requested endpoint. A client may reach more than one rate limit if the request matches multiple mapping rules.

To configure errors, follow these steps:

1. Navigate from [Your_API_name] > Integration > edit APIcast configuration > Gateway response.

2. Choose the error type you want to configure.

3. Specify values for these fields:
   - Response Code: the three-digit HTTP response code.
   - Content-type: the value of the Content-Type header.
   - Response Body: the value of the response message body.

4. To save your changes, click **Update & test in Staging Environment**

### 3.4. CONFIGURATION HISTORY

Every time you click the **Update & Test Staging Configuration** button, the current configuration is saved in a JSON file. The staging gateway will pull the latest configuration with each new request. For each environment, staging or production, you can see a history of all the previous configuration files.

Note that it is not possible to automatically roll back to previous versions. Instead a history of all your configuration versions with their associated JSON files is provided. Use these files to check what configuration you had deployed at any point of time. If you want to, you can recreate any deployments manually.

### 3.5. DEBUGGING

Setting up the gateway configuration is easy, but you may still encounter errors. In such cases, the gateway can return useful debug information to track the error.

To get the debugging information from APIcast, you must add the following header to the API request: **X-3scale-debug: {SERVICE_TOKEN}** with the service token corresponding to the API service that you are reaching to.

When the header is found and the service token is valid, the gateway will add the following information to the response headers:

- **X-3scale-matched-rules**: /v1/word/{word}.json, /v1
- **X-3scale-credentials**: app_key=APP_KEY&app_id=APP_ID
- **X-3scale-usage**: usage%5Bversion_1%5D=1&usage%5Bword%5D=1

**X-3scale-matched-rules** indicates which mapping rules have been matched for the request in a comma-separated list.

The header **X-3scale-credentials** returns the credentials that were passed to 3scale backend.

**X-3scale-usage** indicates the usage that was reported to 3scale backend. Usage in the form `usage%5Bversion_1%5D=1&usage%5Bword%5D=1` is a URL-encoded string and shows that the API request incremented the methods
3.6. PATH ROUTING

APIcast handles all the API services configured on a 3scale account (or a subset of services, if the APICAST_SERVICES_LIST environment variable is configured). Normally, APIcast routes the API requests to the appropriate API service based on the hostname of the request, by matching it with the Public Base URL. The first service where the match is found is used for the authorization.

The Path routing feature allows using the same Public Base URL on multiple services and routes the requests using the path of the request. To enable the feature, set the APICAST_PATH_ROUTING environment variable to true or 1. When enabled, APIcast will map the incoming requests to the services based on both hostname and path.

This feature can be used if you want to expose multiple backend services hosted on different domains through one gateway using the same Public Base URL. To achieve this you can configure several API services for each API backend (i.e. Private Base URL) and enable the path routing feature.

For example, you have 3 services configured in the following way:

- Service A Public Base URL: api.example.com Mapping rule: /a
- Service B Public Base URL: api2.example.com Mapping rule: /b
- Service C Public Base URL: api.example.com Mapping rule: /c

If path routing is disabled (APICAST_PATH_ROUTING=false), all calls to api.example.com will try to match Service A. So, the calls api.example.com/c and api.example.com/b will fail with a “No Mapping Rule matched” error.

If path routing is enabled (APICAST_PATH_ROUTING=true), the calls will be matched by both host and path. So:

- api.example.com/a will be routed to Service A
- api.example.com/c will be routed to Service C
- api.example.com/b will fail with “No Mapping Rule matched” error, i.e. it will NOT match Service B, as the Public Base URL does not match.

If path routing is used, you must ensure there is no conflict between the mapping rules in different services that use the same Public Base URL, i.e. each combination of method + path pattern is only used in one service.
CHAPTER 4. APICAST POLICIES

APICAST policies are units of functionality that modify how APICAST operates. Policies can be enabled, disabled, and configured to control how they modify APICAST. Use policies to add functionality that is not available in a default APICAST deployment. You can create your own policies, or use standard policies provided by Red Hat 3scale.

The following topics provide information about the standard APICAST policies, creating your own custom APICAST policies, and creating a policy chain.

- APICAST Standard Policies
- Creating custom APICAST policies
- Creating a policy chain in the AMP

Control policies for a service with a policy chain. Policy chains do the following:

- specify what policies APICAST uses
- provide configuration information for policies 3scale uses
- specify the order in which 3scale loads policies

**NOTE**

Red Hat 3scale provides a method for adding custom policies, but does not support custom policies.

In order to modify APICAST behavior with custom policies, you must do the following:

- Add custom policies to APICAST
- Define a policy chain that configures APICAST policies
- Add the policy chain to APICAST

4.1. APICAST STANDARD POLICIES

3scale provides the following standard policies:

- Section 4.1.1, “3scale Auth Caching policy”
- Section 4.1.2, “3scale Batcher policy”
- Section 4.1.3, “Anonymous Access policy”
- Section 4.1.4, “CORS Request Handling policy”
- Section 4.1.5, “Echo policy”
- Section 4.1.6, “Edge Limiting policy”
- Section 4.1.7, “Header Modification policy”
- Section 4.1.8, “IP Check policy”
4.1.1. 3scale Auth Caching policy

The 3scale Auth Caching policy caches authentication calls made to APIcast. You can select an operating mode to configure the cache operations.

3scale Auth Caching is available in the following modes:

1. **Strict - Cache only authorized calls.**

   "Strict" mode only caches authorized calls. If a policy is running under the "strict" mode and if a call fails or is denied, the policy invalidates the cache entry. If the backend becomes unreachable, all cached calls are rejected, regardless of their cached status.

2. **Resilient – Authorize according to last request when backend is down.**

   The "Resilient" mode caches both authorized and denied calls. If the policy is running under the "resilient" mode, failed calls do not invalidate an existing cache entry. If the backend becomes unreachable, calls hitting the cache continue to be authorized or denied based on their cached status.

3. **Allow – When backend is down, allow everything unless seen before and denied.**

   The "Allow" mode caches both authorized and denied calls. If the policy is running under the "allow" mode, cached calls continue to be denied or allowed based on the cached status. However, any new calls are cached as authorized.
IMPORTANT

Operating in the "allow" mode has security implications. Consider these implications and exercise caution when using the "allow" mode.

4. None - Disable caching.

The "None" mode disables caching. This mode is useful if you want the policy to remain active, but do not want to use caching.

Configuration properties

<table>
<thead>
<tr>
<th>property</th>
<th>description</th>
<th>values</th>
<th>required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>caching_type</td>
<td>The <code>caching_type</code> property allows you to define which mode the cache will operate in.</td>
<td>data type: enumerated string [resilient, strict, allow, none]</td>
<td>yes</td>
</tr>
</tbody>
</table>

Policy object example

```
{
  "name": "caching",
  "version": "builtin",
  "configuration": {
    "caching_type": "allow"
  }
}
```

For information on how to configure policies, see the Creating a policy chain section of the documentation.

4.1.2. 3scale Batcher policy

The 3scale Batcher policy provides an alternative to the standard APIcast authorization mechanism, in which one call to the 3scale backend (Service Management API) is made for each API request that APIcast receives.

The 3scale Batcher policy caches authorization statuses and batches usage reports, thereby significantly reducing the number of requests to the 3scale backend. With the 3scale Batcher policy you can improve APIcast performance by reducing latency and increasing throughput.

When the 3scale Batcher policy is enabled, APIcast uses the following authorization flow:

1. On each request, the policy checks whether the credentials are cached:
   - If the credentials are cached, the policy uses the cached authorization status instead of calling the 3scale backend.
   - If the credentials are not cached, the policy calls the backend and caches the authorization status with a configurable Time to Live (TTL).

2. Instead of reporting the usage corresponding to the request to the 3scale backend immediately, the policy accumulates their usage counters to report them to the backend in batches.
separate thread reports the accumulated usage counters to the 3scale backend in a single call, with a configurable frequency.

The 3scale Batcher policy improves the throughput, but with reduced accuracy. The usage limits and the current utilization are stored in 3scale, and APIcast can only get the correct authorization status when making calls to the 3scale backend. When the 3scale Batcher policy is enabled, there is a period of time in which APIcast is not sending calls to 3scale. During this time window, applications making calls might go over the defined limits.

Use this policy for high-load APIs if the throughput is more important than the accuracy of the rate limiting. The 3scale Batcher policy gives better results in terms of accuracy when the reporting frequency and authorization TTL are much less than the rate limiting period. For example, if the limits are per day and the reporting frequency and authorization TTL are configured to be several minutes.

The 3scale Batcher policy supports the following configuration settings:

- **auths_ttl**: Sets the TTL in seconds when the authorization cache expires. When the authorization for the current call is cached, APIcast uses the cached value. After the time set in the `auths_ttl` parameter, APIcast removes the cache and calls the 3scale backend to retrieve the authorization status.

- **batch_report_seconds**: Sets the frequency of batch reports APIcast sends to the 3scale backend. The default value is 10 seconds.

**IMPORTANT**

To use this policy, enable both the 3scale APIcast and 3scale Batcher policy in the policy chain.

### 4.1.3. Anonymous Access policy

The Anonymous Access policy exposes a service without authentication. It can be useful, for example, for legacy applications that cannot be adapted to send the authentication parameters. The Anonymous policy only supports services with API Key and App Id / App Key authentication options. When the policy is enabled for API requests that do not have any credentials provided, APIcast will authorize the calls using the default credentials configured in the policy. For the API calls to be authorized, the application with the configured credentials must exist and be active.

Using the Application Plans, you can configure the rate limits on the application used for the default credentials.

**NOTE**

You need to place the Anonymous Access policy before the APIcast Policy, when using these two policies together in the policy chain.

Following are the required configuration properties for the policy:

- **auth_type**: Select a value from one of the alternatives below and make sure the property corresponds to the authentication option configured for the API:
  - **app_id_and_app_key**: For App ID / App Key authentication option.
  - **user_key**: For API key authentication option.
- **app_id** (only for **app_id_and_app_key** auth type): The App Id of the application that will be used for authorization if no credentials are provided with the API call.

- **app_key** (only for **app_id_and_app_key** auth type): The App Key of the application that will be used for authorization if no credentials are provided with the API call.

- **user_key** (only for the **user_key** auth_type): The API Key of the application that will be used for authorization if no credentials are provided with the API call.

**Figure 4.1. Anonymous Access Policy**

Anonymous access
builtin – Provides default credentials for unauthenticated requests

This policy allows to expose a service without authentication. It can be useful, for example, for legacy apps that cannot be adapted to send the auth params. When the credentials are not provided in the request, this policy provides the default ones configured. An app_id + app_key or a user_key should be configured.

- Enabled

  auth_type*

  app_id_and_app_key

  app_key*

    myappid

  app_id*

    secret-app-key-123

**4.1.4. CORS Request Handling policy**

The Cross Origin Resource Sharing (CORS) Request Handling policy allows you to control CORS behavior by allowing you to specify:

- Allowed headers
- Allowed methods
- Allowed credentials
- Allowed origin headers

The CORS request handling policy will block all unspecified CORS requests.
NOTE
You need to place the CORS Request Handling policy before the APIcast Policy, when using these two policies together in the policy chain.

Configuration properties

<table>
<thead>
<tr>
<th>property</th>
<th>description</th>
<th>values</th>
<th>required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>allow_headers</td>
<td>The <strong>allow_headers</strong> property is an array in which you can specify which CORS headers APIcast will allow.</td>
<td>data type: array of strings, must be a CORS header</td>
<td>no</td>
</tr>
<tr>
<td>allow_methods</td>
<td>The <strong>allow_methods</strong> property is an array in which you can specify which CORS methods APIcast will allow.</td>
<td>data type: array of enumerated strings [GET, HEAD, POST, PUT, DELETE, PATCH, OPTIONS, TRACE, CONNECT]</td>
<td>no</td>
</tr>
<tr>
<td>allow_origin</td>
<td>The <strong>allow_origin</strong> property allows you to specify an origin domain APIcast will allow</td>
<td>data type: string</td>
<td>no</td>
</tr>
<tr>
<td>allow_credentials</td>
<td>The <strong>allow_credentials</strong> property allows you to specify whether APIcast will allow a CORS request with credentials</td>
<td>data type: boolean</td>
<td>no</td>
</tr>
</tbody>
</table>

Policy object example

```json
{
    "name": "cors",
    "version": "builtin",
    "configuration": {
        "allow_headers": [
            "App-Id", "App-Key",
            "Content-Type", "Accept"
        ],
        "allow_credentials": true,
        "allow_methods": [
            "GET", "POST"
        ],
        "allow_origin": "https://example.com"
    }
}
```
For information on how to configure policies, refer to the Creating a policy chain section of the documentation.

### 4.1.5. Echo policy

The Echo policy prints an incoming request back to the client, along with an optional HTTP status code.

**Configuration properties**

<table>
<thead>
<tr>
<th>property</th>
<th>description</th>
<th>values</th>
<th>required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>status</td>
<td>The HTTP status code the echo policy will return to the client</td>
<td>data type: integer</td>
<td>no</td>
</tr>
<tr>
<td>exit</td>
<td>Specifies which exit mode the echo policy will use. The request exit mode stops the incoming request from being processed. The set exit mode skips the rewrite phase.</td>
<td>data type: enumerated string [request, set]</td>
<td>yes</td>
</tr>
</tbody>
</table>

**Policy object example**

```json
{
   "name": "echo",
   "version": "builtin",
   "configuration": {
      "status": 404,
      "exit": "request"
   }
}
```

For information on how to configure policies, refer to the Creating a policy chain section of the documentation.

### 4.1.6. Edge Limiting policy

The Edge Limiting policy aims to provide flexible rate limiting for the traffic sent to the backend API and can be used with the default 3scale authorization. Some examples of the use cases supported by the policy include:

- End-user rate limiting: Rate limit by the value of the "sub" (subject) claim of a JWT token passed in the Authorization header of the request (configured as `{{ jwt.sub }}`).
- Requests Per Second (RPS) rate limiting.
- Global rate limits per service: Apply limits per service rather than per application.
- Concurrent connection limit: Set the number of concurrent connections allowed.
4.1.6.1. Types of limits

The policy supports the following types of limits that are provided by the `lua-resty-limit-traffic` library:

- **leaky_bucket_limiters**: Based on the "leaky_bucket" algorithm that is based on the average number of requests plus a maximum burst size.

- **fixed_window_limiters**: Based on a fixed window of time (last X seconds).

- **connection_limiters**: Based on the concurrent number of connections.

You can scope any limit by service or globally.

4.1.6.2. Limit definition

The limits have a key that encodes the entities that are used to define the limit (an IP, a service, an endpoint, an ID, the value for a specific header, etc.). The Key is specified in the `key` parameter of the limiter.

**key** is an object that is defined by the following properties:

- **name**: It is the name of the key. It must be unique in the scope.

- **scope**: It defines the scope of the key. The supported scopes are:
  - Per service scope that affects one service (**service**).
  - Global scope that affects all the services (**global**).

- **name_type**: It defines how the "name" value will be evaluated:
  - As plain text (**plain**)
  - As Liquid (**liquid**)

Each limit also has some parameters that vary depending on their types:

- **leaky_bucket_limiters**: **rate, burst**
  - **rate**: It defines how many requests can be made per second without a delay.
  - **burst**: It defines the amount of requests per second that can exceed the allowed rate. An artificial delay is introduced for requests above the allowed rate (specified by **rate**). After exceeding the rate by more requests per second than defined in **burst**, the requests get rejected.

- **fixed_window_limiters**: **count, window. count** defines how many requests can be made per number of seconds defined in **window**.

- **connection_limiters**: **conn, burst, delay**
  - **conn**: Defines the maximum number of the concurrent connections allowed. It allows exceeding that number by **burst** connections per second.
  - **delay**: It is the number of seconds to delay the connections that exceed the limit.

**Examples**
1. Allow 10 requests per minute to service_A:

```json
{
  "key": { "name": "service_A" },
  "count": 10,
  "window": 60
}
```

2. Allow 100 connections with bursts of 10 with a delay of 1s:

```json
{
  "key": { "name": "service_A" },
  "conn": 100,
  "burst": 10,
  "delay": 1
}
```

You can define several limits for each service. In case multiple limits are defined, the request can be rejected or delayed if at least one limit is reached.

### 4.1.6.3. Liquid templating

The Edge Limiting policy allows specifying the limits for the dynamic keys by supporting Liquid variables in the keys. For this, the `name_type` parameter of the key must be set to "liquid" and the `name` parameter can then use Liquid variables. Example: `{ { remote_addr } }` for the client IP address or `{ { jwt.sub } }` for the 'sub' claim of the JWT token.

**Example:**

```json
{
  "key": { "name": "{ { jwt.sub } }", "name_type": "liquid" },
  "count": 10,
  "window": 60
}
```

For more information about Liquid support, see Section 5.1, “Using variables and filters in policies”.

### 4.1.6.4. Applying conditions

Each limiter can have an optional condition that defines when the limiter must be applied. The condition is specified in the `condition` property of the limiter.

**condition** is defined by the following properties:

- **combine_op**: It is the boolean operator applied to the list of operations. The following two values are supported: **or** and **and**.

- **operations**: It is a list of conditions that need to be evaluated. Each operation is represented by an object with the following properties:
  - **left**: The left part of the operation.
  - **left_type**: How the **left** property is evaluated (plain or liquid).
  - **right**: The right part of the operation.
○ **right_type**: How the right property is evaluated (plain or liquid).

○ **op**: Operator applied between the left and the right parts. The following two values are supported: == (equals) and != (not equals).

Example:

```json
"condition": {
  "combine_op": "and",
  "operations": [
    {
      "op": "==",
      "right": "GET",
      "left_type": "liquid",
      "left": "\{ http_method \}",
      "right_type": "plain"
    }
  ]
}
```

### 4.1.6.5. Configuring the store

By default, Edge Limiting policy uses OpenResty shared dictionary for the rate limiting counters. However, an external Redis server can be used instead of the shared dictionary. This can be useful when multiple APIcast instances are used. Redis server can be configured using the `redis_url` parameter.

### 4.1.6.6. Error handling

The limiters support the following parameters to configure how the errors are handled:

- **limits_exceeded_error**: Allows to configure the error status code and message that will be returned to the client when the configured limits are exceeded. The following parameters should be configured:

  ○ **status_code**: The status code of the request when the limits are exceeded. Default: 429.

  ○ **error_handling**: Specifies how to handle the error, with following options:

    - **exit**: Stops processing request and returns an error message.

    - **log**: Completes processing request and returns output logs.

- **configuration_error**: Allows to configure the error status code and message that will be returned to the client in case of incorrect configuration. The following parameters should be configured:

  ○ **status_code**: The status code when there is a configuration issue. Default: 500.

  ○ **error_handling**: Specifies how to handle the error, with following options:

    - **exit**: Stops processing request and returns an error message.

    - **log**: Completes processing request and returns output logs.

### 4.1.7. Header Modification policy
The Header Modification policy allows you to modify the existing headers or define additional headers to add to or remove from an incoming request or response. You can modify both response and request headers.

The Header Modification policy supports the following configuration parameters:

- **request**: List of operations to apply to the request headers
- **response**: List of operations to apply to the response headers

Each operation consists of the following parameters:

- **op**: Specifies the operation to be applied. The **add** operation adds a value to an existing header. The **set** operation creates a header and value, and will overwrite an existing header’s value if one already exists. The **push** operation creates a header and value, but will not overwrite an existing header’s value if one already exists. Instead, **push** will add the value to the existing header. The **delete** operation removes the header.

- **header**: Specifies the header to be created or modified and can be any string that can be used as a header name (e.g. Custom-Header).

- **value_type**: Defines how the header value will be evaluated and can either be **plain** for plain text or **liquid** for evaluation as a Liquid template. For more information, see Section 5.1, “Using variables and filters in policies”.

- **value**: Specifies the value that will be used for the header. For value type "liquid" the value should be in the format **{{ variable_from_context }}**. Not needed when deleting.

**Policy object example**

```json
{
  "name": "headers",
  "version": "builtin",
  "configuration": {
    "response": [
      {
        "op": "add",
        "header": "Custom-Header",
        "value_type": "plain",
        "value": "any-value"
      }
    ],
    "request": [
      {
        "op": "set",
        "header": "Authorization",
        "value_type": "plain",
        "value": "Basic dXNlcm5hbWU6cGFzc3dvcmQ="
      },
      {
        "op": "set",
        "header": "Service-ID",
        "value_type": "liquid",
        "value": "{{service.id}}"
      }
    ]
  }
}
```
For information on how to configure policies, see the Creating a policy chain section of the documentation.

4.1.8. IP Check policy

The IP Check policy is used to deny or allow requests based on a list of IPs.

Configuration properties

<table>
<thead>
<tr>
<th>property</th>
<th>description</th>
<th>data type</th>
<th>required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>check_type</td>
<td>The check_type property has two possible values, whitelist or blacklist.</td>
<td>string, must be either whitelist or blacklist</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>blacklist will deny all requests from IPs on the list.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>whitelist will deny all requests from IPs not on the list.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ips</td>
<td>The ips property allows you to specify a list of IP addresses to whitelist or blacklist. Both single IPs and CIDR ranges can be used.</td>
<td>array of strings, must be valid IP addresses</td>
<td>yes</td>
</tr>
<tr>
<td>error_msg</td>
<td>The error_msg property allows you to configure the error message returned when a request is denied.</td>
<td>string</td>
<td>no</td>
</tr>
<tr>
<td>client_ip_sources</td>
<td>The client_ip_sources property allows you to configure how to retrieve the client IP. By default, the last caller IP is used. The other options are X-Forwarded-For, X-Real-IP, last_caller.</td>
<td>array of strings, valid options are one or more of X-Forwarded-For, X-Real-IP, last_caller.</td>
<td>no</td>
</tr>
</tbody>
</table>

Policy object example

```yaml
{}
```
For information on how to configure policies, refer to the Creating a policy chain section of the documentation.

4.1.9. JWT Claim Check policy

4.1.9.1. About JWT Claim Check policy

Based on JSON Web Token (JWT) claims, the JWT Claim Check policy allows you to define new rules to block resource targets and methods.

In order to route based on the value of a JWT claim, you need a policy in the chain that validates the JWT and stores the claim in the context that the policies share.

If the JWT Claim Check policy is blocking a resource and a method, the policy also validates the JWT operations. Alternatively, in case that the method resource does not match, the request continues to the backend API.

Example: In case of a GET request, the JWT needs to have the role claim as admin, if not the request will be denied. On the other hand, any non GET request will not validate the JWT operations, so POST resource is allowed without JWT constraint.

```
{
  "name": "apicast.policy.jwt_claim_check",
  "configuration": {
    "error_message": "Invalid JWT check",
    "rules": [
      {
        "operations": [
          {"op": "==", "jwt_claim": "role", "jwt_claim_type": "plain", "value": "admin"}
        ],
        "combine_op": "and",
        "methods": ["GET"],
        "resource": "/resource",
        "resource_type": "plain"
      }
    ]
  }
}
```

4.1.9.2. Configuring JWT Claim Check policy in your policy chain

4.1.9.2.1. Prerequisites:

- You need to have access to a 3scale installation.
• You need to wait for all the deployments to finish.

4.1.9.2.2. Configuring the policy

1. To add the JWT Claim Check policy to your API, follow the steps described in Enabling a standard Policy and choose JWT Claim Check.

2. Click the JWT Claim Check link.

3. To enable the policy, select the Enabled checkbox.

4. To add rules, click the plus + icon.

5. Specify the resource_type.

6. Choose the operator.

7. Indicate the resource controlled by the rule.

8. To add the allowed methods, click the plus + icon.

9. Type the error message to show to the user when traffic is blocked.

10. When you have finished setting up your API with JWT Claim Check, click Update Policy.

Additionally:

• You can add more resource types and allowed methods by clicking the plus + icon in the corresponding section.

To save your changes, click Update & test in Staging Environment

4.1.10. Liquid Context Debug policy

NOTE

The Liquid Context Debug policy is meant only for debugging purposes in the development environment and not in production.

This policy responds to the API request with a JSON, containing the objects and values that are available in the context and can be used for evaluating Liquid templates. When combined with the 3scale APIcast or Upstream policy, Liquid Context Debug must be placed before them in the policy chain in order to work correctly. To avoid circular references, the policy only includes duplicated objects once and replaces them with a stub value.

An example of the value returned by APIcast when the policy is enabled:

```json
{
    "jwt": {
        "azp": "972f7b4f",
        "iat": 1537538097,
        ...
        "exp": 1537574096,
        "typ": "Bearer"
    }
}
```
4.1.11. Logging policy

The Logging policy has two purposes:

- To enable and disable access log output.
- To create a custom access log format for each service and be able to set conditions to write custom access log.

You can combine the Logging policy with the global setting for the location of access logs. Set the `APICAST_ACCESS_LOG_FILE` environment variable to configure the location of APIcast access logs. By default, this variable is set to `/dev/stdout`, which is the standard output device. For further details about global APIcast parameters, see Chapter 7, APIcast environment variables.

Additionally, the Logging policy has these features:

- This policy only supports the `enable_access_logs` configuration parameter.
- To enable the access logs, select the `enable_access_logs` parameter or disable the Logging policy.
- To disable access logging for an API:
  1. Enable the policy.
  2. Clear the `enable_access_logs` parameter
  3. Click the Submit button.
- By default, this policy is not enabled in policy chains.

4.1.11.1. Global configuration for all APIs

Logging options help to avoid issues with logs that are not correctly formatted in APIs. A custom APIcast environment variable can be set and all APIs implement the Logging policy. Here is an example of a policy that is loaded in all services: `customEnv.lua`

```lua
local cjson = require('cjson')
```
local PolicyChain = require('apicast.policy_chain')
local policy_chain = context.policy_chain

local logging_policy_config = cjson.decode([{
  "enable_access_logs": false,
  "custom_logging": "{{request}}" to service {{service.id}} and {{service.name}}"
}])
]

policy_chain:insert( PolicyChain.load_policy('logging', 'builtin', logging_policy_config), 1)

return {
  policy_chain = policy_chain,
  port = { metrics = 9421 },
}

To run APIcast with this specific environment:

docker run --name apicast --rm -p 8080:8080 \
-v $(pwd):/config \
-e APICAST_ENVIRONMENT=/config/custom_env.lua \
-e THREESCALE_PORTAL_ENDPOINT=https://ACCESS_TOKEN@ADMIN_PORTAL_DOMAIN \
quay.io/3scale/apicast:master

These are key concepts of the Docker command to consider:

- Current Lua file must be shared to the container -v $(pwd):/config.
- APICAST_ENVIRONMENT variable must be set to the Lua file that is stored in the /config
directory.

4.1.11.2. Examples

This section describes some examples when working with the Logging policy. These examples consider
the following caveats:

- If custom_logging or enable_json_logs property is enabled, default access log will be
disabled.

- If enable_json_logs is enabled, the custom_logging field will be omitted.

Disabling access log

```lua
{
  "name": "apicast.policy.logging",
  "configuration": {
    "enable_access_logs": false
  }
}
```

Enabling custom access log

```lua
{
  "name": "apicast.policy.logging",
```
Enabling custom access log with the service identifier

```json
{
  "name": "apicast.policy.logging",
  "configuration": {
    "enable_access_logs": false,
    "custom_logging": "{{request}}" to service {{service.id}} and {{service.name}}",
  }
}
```

Configuring access logs in JSON format

```json
{
  "name": "apicast.policy.logging",
  "configuration": {
    "enable_access_logs": false,
    "enable_json_logs": true,
    "json_object_config": [
      {
        "key": "host",
        "value": "{{host}}",
        "value_type": "liquid"
      },
      {
        "key": "time",
        "value": "{{time_local}}",
        "value_type": "liquid"
      },
      {
        "key": "custom",
        "value": "custom_method",
        "value_type": "plain"
      }
    ]
  }
}
```

Configuring a custom access log only for a successful request

```json
{
  "name": "apicast.policy.logging",
  "configuration": {
    "enable_access_logs": false,
    "custom_logging": "{{request}}" to service {{service.id}} and {{service.name}}",
    "condition": {
      "operations": [
        {"op": "==", "match": "{{status}}", "match_type": "liquid", "value": "200"}
      ]
    }
  }
}
```
Customizing access logs where the response status matches either 200 or 500

```json
{
  "name": "apicast.policy.logging",
  "configuration": {
    "enable_access_logs": false,
    "custom_logging": "{{request}}" to service {{service.id}} and {{service.name}}",
    "condition": {
      "operations": [
        {"op": "==", "match": "{{status}}", "match_type": "liquid", "value": "200"},
        {"op": "==", "match": "{{status}}", "match_type": "liquid", "value": "500"}
      ],
      "combine_op": "or"
    }
  }
}
```

4.1.11.3. Additional information: custom logging

For custom logging, you can use Liquid templates with exported variables. These variables include:

- NGINX default directive variable: log_format. For example: {{remote_addr}}.
- Response and request headers:
  - {{req.headers.FOO}}: To get the FOO header in the request.
  - {{res.headers.FOO}}: To retrieve the FOO header on response.
- Service information, such as {{service.id}}, and all the service properties provided by these parameters:
  - THREESCALE_CONFIG_FILE
  - THREESCALE_PORTAL_ENDPOINT

4.1.12. Maintenance Mode policy

The Maintenance Mode policy allows you reject incoming requests with a specified status code and message. It is useful for maintenance periods or to temporarily block an API.

Configuration properties

The following is a list of possible properties and default values.

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
<th>default</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>status</td>
<td>integer, optional</td>
<td>503</td>
<td>Response code</td>
</tr>
</tbody>
</table>
**Maintenance mode policy example**

```json
{
  "policy_chain": [
    {
      "name": "maintenance-mode",
      "version": "1.0.0",
      "configuration": {
        "message": "Be back soon..",
        "status": 503
      }
    }
  ]
}
```

For information on how to configure the Maintenance Mode policy in 3scale, see the Section 4.5, “Creating a policy chain in 3scale” section of the documentation.

### 4.1.13. OAuth 2.0 Token Introspection policy

The OAuth 2.0 Token Introspection policy allows validating the JSON Web Token (JWT) token used for services with the OpenID Connect (OIDC) authentication option using the Token Introspection Endpoint of the token issuer (Red Hat Single Sign-On).

APIcast supports the following authentication types in the `auth_type` field to determine the Token Introspection Endpoint and the credentials APIcast uses when calling this endpoint:

- **use_3scale_oidcIssuer_endpoint**: APIcast uses the client credentials, *Client ID* and *Client Secret*, as well as the Token Introspection Endpoint from the OIDC Issuer setting configured on the Service Integration page. APIcast discovers the Token Introspection endpoint from the `token_introspection_endpoint` field. This field is located in the `.well-known/openid-configuration` endpoint that is returned by the OIDC issuer.

**Example 4.1. Authentication type set to use_3scale_oidcIssuer_endpoint**

```json
"policy_chain": [
  ...
  {
    "name": "apicast.policy.token_introspection",
    "configuration": {
      "auth_type": "use_3scale_oidcIssuer_endpoint"
    }
  }
  ...
],
```

- **client_id+client_secret**: This option enables you to specify a different Token Introspection Endpoint, as well as the *Client ID* and *Client Secret* APIcast uses to request token information. When using this option, set the following configuration parameters:
  - **client_id**: Sets the Client ID for the Token Introspection Endpoint.
  - **client_secret**: Sets the Client Secret for the Token Introspection Endpoint.
- **introspection_url**: Sets the Introspection Endpoint URL.

Example 4.2. Authentication type set to `client_id+client_secret`

```json
"policy_chain": [
...
{
"name": "apicast.policy.token_introspection",
"configuration": {
"auth_type": "client_id+client_secret",
"client_id": "myclient",
"client_secret": "mysecret",
"introspection_url": "http://red_hat_single_sign-on/token/introspection"
}
...
].
```

Regardless of the setting in the **auth_type** field, APIcast uses Basic Authentication to authorize the Token Introspection call (**Authorization: Basic <token>** header, where `<token>` is Base64-encoded `<client_id>:<client_secret>` setting).
The response of the Token Introspection Endpoint contains the **active** attribute. APIcast checks the value of this attribute. Depending on the value of the attribute, APIcast authorizes or rejects the call:
- **true**: The call is authorized
- **false**: The call is rejected with the **Authentication Failed** error

The policy allows enabling caching of the tokens to avoid calling the Token Introspection Endpoint on every call for the same JWT token. To enable token caching for the Token Introspection Policy, set the `max_cached_tokens` field to a value from 0, which disables the feature, and 10000. Additionally, you can set a Time to Live (TTL) value from 1 to 3600 seconds for tokens in the `max_ttl_tokens` field.

### 4.1.14. Referrer policy

The Referrer policy enables the Referrer Filtering feature. When the policy is enabled in the service policy chain, APICAST sends the value of the Referrer policy to the Service Management API as an upwards AuthRep call. The value of the Referrer policy is sent in the `referrer` parameter in the call.

For more information on how Referrer Filtering works, see the Referrer Filtering section under Authentication Patterns.

### 4.1.15. Retry policy

The Retry policy sets the number of retry requests to the upstream API. The retry policy is configured per service, so users can enable retries for as few or as many of their services as desired, as well as configure different retry values for different services.

**IMPORTANT**

As of 3scale 2.8, it is not possible to configure which cases to retry from the policy. This is controlled with the environment variable `APICAST_UPSTREAM_RETRY_CASES`, which applies retry requests to all services. For more on this, check out `APICAST_UPSTREAM_RETRY_CASES`.

An example of the retry policy JSON is shown below:

```json
{
  "$schema": "http://apicast.io/policy-v1/schema#manifest#",
  "name": "Retry",
  "summary": "Allows retry requests to the upstream",
  "description": "Allows retry requests to the upstream",
  "version": "builtin",
  "configuration": {
    "type": "object",
    "properties": {
      "retries": {
        "description": "Number of retries",
        "type": "integer",
        "minimum": 1,
        "maximum": 10
      }
    }
  }
}
```

### 4.1.16. RH-SSO/Keycloak Role Check policy
This policy adds role check when used with the OpenID Connect authentication option. This policy verifies realm roles and client roles in the access token issued by Red Hat Single Sign-On (RH-SSO). The realm roles are specified when you want to add role check to every client resource of 3scale.

There are the two types of role checks that the **type** property specifies in the policy configuration:

- **whitelist** (default): When `whitelist` is used, APIcast will check if the specified scopes are present in the JWT token and will reject the call if the JWT doesn’t have the scopes.

- **blacklist**: When `blacklist` is used, APIcast will reject the calls if the JWT token contains the blacklisted scopes.

It is not possible to configure both checks – `blacklist` and `whitelist` in the same policy, but you can add more than one instance of the **RH-SSO/Keycloak role check** policy to the APIcast policy chain.

You can configure a list of scopes via the `scopes` property of the policy configuration.

Each `scope` object has the following properties:

- **resource**: Resource (endpoint) controlled by the role. This is the same format as Mapping Rules. The pattern matches from the beginning of the string and to make an exact match you must append `$` at the end.

- **resource_type**: This defines how the `resource` value is evaluated.
  - As plain text (**plain**): Evaluates the `resource` value as plain text. Example: `/api/v1/products$`.
  - As Liquid text (**liquid**): Allows using Liquid in the `resource` value. Example: `/resource_{jwt.aud}` manages access to the resource containing the Client ID.

- **methods**: Use this parameter to list the allowed HTTP methods in APIcast, based on the user roles in RH-SSO. As examples, you can allow methods that have:
  - The `role1` realm role to access `/resource1`. For those methods that do not have this realm role, you need to specify the `blacklist`.
  - The `client1` role called `role1` to access `/resource1`.
  - The `role1` and `role2` realm roles to access `/resource1`. Specify the roles in `realm_roles`. You can also indicate the scope for each role.
  - The client role called `role1` of the application client, which is the recipient of the access token, to access `/resource1`. Use `liquid` client type to specify the JSON Web Token (JWT) information to the client.
  - The client role including the client ID of the application client, the recipient of the access token, to access `/resource1`. Use `liquid` client type to specify the JWT information to the `name` of the client role.
  - The client role called `role1` to access the resource including the application client ID. Use `liquid` client type to specify the JWT information to the `resource`.

- **realm_roles**: Use it to check the realm role (see the **Realm Roles in Red Hat Single Sign-On** documentation). The realm roles are present in the JWT issued by Red Hat Single Sign-On.
"realm_access": {
  "roles": [
    "<realm_role_A>", "<realm_role_B>"
  ]
}

The real roles must be specified in the policy.

"realm_roles": [
  { "name": "<realm_role_A>" }, { "name": "<realm_role_B>" }
]

Following are the available properties of each object in the `realm_roles` array:

- **name**: Specifies the name of the role.
- **name_type**: Defines how the name must be evaluated; it can be **plain** or **liquid** (works the same way as for the `resource_type`).
- **client_roles**: Use `client_roles` to check for the particular access roles in the client namespace (see the Client Roles in Red Hat Single Sign-On documentation). The client roles are present in the JWT under the `resource_access` claim.

"resource_access": {
  "<client_A>": {
    "roles": [
      "<client_role_A>", "<client_role_B>"
    ]
  },
  "<client_B>": {
    "roles": [
      "<client_role_A>", "<client_role_B>"
    ]
  }
}

Specify the client roles in the policy.

"client_roles": [
  { "name": "<client_role_A>", "client": "<client_A>" },
  { "name": "<client_role_B>", "client": "<client_A>" },
  { "name": "<client_role_A>", "client": "<client_B>" },
  { "name": "<client_role_B>", "client": "<client_B>" }
]

Following are the available properties of each object in the `client_roles` array:

- **name**: Specifies the name of the role.
- **name_type**: Defines how the `name` value must be evaluated; it can be **plain** or **liquid** (works the same way as for the `resource_type`).
- **client**: Specifies the client of the role. When it is not defined, this policy uses the **aud** claim as the client.
• **client_type**: Defines how the **client** value must be evaluated; it can be **plain** or **liquid** (works the same way as for the **resource_type**).

### 4.1.17. Routing policy

The Routing policy allows you to route requests to different target endpoints. You can define target endpoints and then you will be able to route incoming requests from the UI to those using regular expressions.

Routing is based on the following rules:

- Request path rule
- Header rule
- Query argument rule
- JSON Web Token (JWT) claim rule

When combined with the APIcast policy, the Routing policy should be placed before the APIcast one in the chain, as the two policies that comes first will output content to the response. When the second gets a change to run its content phase, the request will already be sent to the client, so it will not output anything to the response.

#### 4.1.17.1. Routing rules

- If multiple rules exist, the Routing policy applies the first match. You can sort these rules.

- If no rules match, the policy will not change the upstream and will use the defined Private Base URL defined in the service configuration.

#### 4.1.17.2. Request path rule

This is a configuration that routes to **http://example.com** when the path is **/accounts**:

```json
{
    "name": "routing",
    "version": "builtin",
    "configuration": {
        "rules": [
            {
                "url": "http://example.com",
                "condition": {
                    "operations": [
                        {
                            "match": "path",
                            "op": "==",
                            "value": "/accounts"
                        }
                    ]
                }
            }
        ]
    }
}
```
4.1.17.3. Header rule

This is a configuration that routes to http://example.com when the value of the header Test-Header is 123:

```
{
  "name": "routing",
  "version": "builtin",
  "configuration": {
    "rules": [
      {
        "url": "http://example.com",
        "condition": {
          "operations": [
            {
              "match": "header",
              "header_name": "Test-Header",
              "op": "==",
              "value": "123"
            }
          ]
        }
      }
    ]
  }
}
```

4.1.17.4. Query argument rule

This is a configuration that routes to http://example.com when the value of the query argument test_query_arg is 123:

```
{
  "name": "routing",
  "version": "builtin",
  "configuration": {
    "rules": [
      {
        "url": "http://example.com",
        "condition": {
          "operations": [
            {
              "match": "query_arg",
              "query_arg_name": "test_query_arg",
              "op": "==",
              "value": "123"
            }
          ]
        }
      }
    ]
  }
}
```
4.1.17.5. JWT claim rule

To route based on the value of a JWT claim, there needs to be a policy in the chain that validates the JWT and stores it in the context that the policies share.

This is a configuration that routes to `http://example.com` when the value of the JWT claim `test_claim` is 123:

```json
{
  "name": "routing",
  "version": "builtin",
  "configuration": {
    "rules": [
      {
        "url": "http://example.com",
        "condition": {
          "operations": [
            {
              "match": "jwt_claim",
              "jwt_claim_name": "test_claim",
              "op": "==",
              "value": "123"
            }
          ]
        }
      }
    ]
  }
}
```

4.1.17.6. Multiple operations rule

Rules can have multiple operations and route to the given upstream only when all of them evaluate to true (using the 'and' `combine_op`), or when at least one of them evaluates to true (using the 'or' `combine_op`). The default value of `combine_op` is 'and'.

This is a configuration that routes to `http://example.com` when the path of the request is `/accounts` and when the value of the header `Test-Header` is 123:

```json
{
  "name": "routing",
  "version": "builtin",
  "configuration": {
    "rules": [
      {
        "url": "http://example.com",
        "condition": {
          "combine_op": "and",
          "operations": [
            {
              "match": "path",
              "op": "==",
              "value": "/accounts"
            }
          ]
        }
      }
    ]
  }
}
```
This is a configuration that routes to http://example.com when the path of the request is /accounts or when the value of the header Test-Header is 123:

```json
{
  "name": "routing",
  "version": "builtin",
  "configuration": {
    "rules": [
      {
        "url": "http://example.com",
        "condition": {
          "combine_op": "or",
          "operations": [
            {
              "match": "path",
              "op": "==",
              "value": "/accounts"
            },
            {
              "match": "header",
              "header_name": "Test-Header",
              "op": "==",
              "value": "123"
            }
          ]
        }
      }
    ]
  }
}
```

4.1.17.7. Combining rules

Rules can be combined. When there are several rules, the upstream selected is one of the first rules that evaluates to true.

This is a configuration with several rules:

```json
{
  "name": "routing",
  "version": "builtin",
  "configuration": {
    "rules": [
```

CHAPTER 4. APICAST POLICIES
4.1.17.8. Catch-all rules

A rule without operations always matches. This can be useful to define catch-all rules.

This configuration routes the request to `http://some_upstream.com` if the path is `/abc`, routes the request to `http://another_upstream.com` if the path is `/def`, and finally, routes the request to `http://default_upstream.com` if none of the previous rules evaluated to true:
4.1.17.9. Supported operations

The supported operations are `==`, `!=`, and `matches`. The latter matches a string with a regular expression and it is implemented using `ngx.re.match`.

This is a configuration that uses `!=`. It routes to `http://example.com` when the path is not `/accounts`:

```json
{
   "name": "routing",
   "version": "builtin",
   "configuration": {
      "rules": [
         {
            "url": "http://example.com",
            "condition": {
               "operations": [
                  {
                     "match": "path",
                     "op": "!=",
                     "value": "/accounts"
                  }
               ]
            }
         }
      ]
   }
}
```

4.1.17.10. Liquid templating

It is possible to use liquid templating for the values of the configuration. This allows you to define rules with dynamic values if a policy in the chain stores the key `my_var` in the context.

This is a configuration that uses that value to route the request:

```json
{
   "name": "routing",
   "version": "builtin",
   "configuration": {
      "rules": [
         {
            "url": "http://example.com",
            "condition": {
               "operations": [
                  {
                     "match": "path",
                     "op": "=",
                     "value": "{{my_var}}"
                  }
               ]
            }
         }
      ]
   }
}
```
4.1.17.11. Set the host used in the Host header

By default, when a request is routed, the policy sets the Host header using the host of the URL of the rule that matched. It is possible to specify a different host with the `host_header` attribute.

This is a configuration that specifies `some_host.com` as the host of the Host header:

```
{
  "name": "routing",
  "version": "builtin",
  "configuration": {
    "rules": [
      {
        "url": "http://example.com",
        "host_header": "some_host.com",
        "condition": {
          "operations": [
            {
              "match": "path",
              "op": "==",
              "value": "/"
            }
          ]
        }
      }
    ]
  }
}
```

4.1.18. SOAP policy

The SOAP policy matches SOAP action URIs provided in the `SOAPAction` or `Content-Type` header of ...
The SOAP policy matches SOAP action URIs provided in the `SOAPAction` or `Content-Type` header of an HTTP request with mapping rules specified in the policy.

### Configuration properties

<table>
<thead>
<tr>
<th>property</th>
<th>description</th>
<th>values</th>
<th>required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>pattern</td>
<td>The <em>pattern</em> property allows you to specify a string that APIcast will seek matches for in the SOAPAction URI.</td>
<td>data type: string</td>
<td>yes</td>
</tr>
<tr>
<td>metric_system_name</td>
<td>The <em>metric_system_name</em> property allows you to specify the 3scale backend metric with which your matched pattern will register a hit.</td>
<td>data type: string, must be a valid metric</td>
<td>yes</td>
</tr>
</tbody>
</table>

### Policy object example

```json
{
    "name": "soap",
    "version": "builtin",
    "configuration": {
        "mapping_rules": [
            {
                "pattern": "http://example.com/soap#request",
                "metric_system_name": "soap",
                "delta": 1
            }
        ]
    }
}
```

For information on how to configure policies, refer to the Creating a policy chain section of the documentation.

### 4.1.19. TLS Client Certificate Validation policy

#### 4.1.19.1. About TLS Client Certificate Validation policy

With the TLS Client Certificate Validation policy, APIcast implements a TLS handshake and validates the client certificate against a whitelist. A whitelist contains certificates signed by the Certified Authority (CA) or just plain client certificates. In case of an expired or invalid certificate, the request is rejected and no other policies will be processed.

The client connects to APIcast to send a request and provides a Client Certificate. APIcast verifies the authenticity of the provided certificate in the incoming request according to the policy configuration. APIcast can also be configured to use a client certificate of its own to use it when connecting to the upstream.
4.1.19.2. Setting up APIcast to work with TLS Client Certificate Validation

APIcast needs to be configured to terminate TLS. Follow the steps below to configure the validation of client certificates provided by users on APIcast with the Client Certificate Validation policy.

4.1.19.2.1. Prerequisites:

- You need to have access to a 3scale installation.
- You need to wait for all the deployments to finish.

4.1.19.2.2. Setting up APIcast to work with the policy

To set up APIcast and configure it to terminate TLS, follow these steps:

1. You need to get the access token and deploy APIcast self-managed, as indicated in Deploying APIcast using the OpenShift template.

   **NOTE**

   APIcast self-managed deployment is required as the APIcast instance needs to be reconfigured to use some certificates for the whole gateway.

2. For testing purposes only, you can use the lazy loader with no cache and staging environment and --param flags for the ease of testing:

   ```
   oc new-app -f https://raw.githubusercontent.com/3scale/3scale-amp-openshift-templates/master/apicast-gateway/apicast.yml --param CONFIGURATION_LOADER=lazy --param DEPLOYMENT_ENVIRONMENT=staging --param CONFIGURATION_CACHE=0
   ```

3. Generate certificates for testing purposes. Alternatively, for production deployment, you can use the certificates provided by a Certificate Authority.

4. Create a Secret with TLS certificates:

   ```
   oc create secret tls apicast-tls --cert=ca/certs/server.crt --key=ca/keys/server.key
   ```

5. Mount the Secret inside the APIcast deployment:

   ```
   oc set volume dc/apicast --add --name=certificates --mount-path=/var/run/secrets/apicast --secret-name=apicast-tls
   ```

6. Configure APIcast to start listening on port 8443 for HTTPS:

   ```
   oc set env dc/apicast APICAST_HTTPS_PORT=8443
   APICAST_HTTPS_CERTIFICATE=/var/run/secrets/apicast/tls.crt
   APICAST_HTTPS_CERTIFICATE_KEY=/var/run/secrets/apicast/tls.key
   ```

7. Expose 8443 on the Service:

   ```
   oc patch service apicast -p '{"spec":{"ports":[{"name":"https","port":8443,"protocol":"TCP"}]}}'
   ```
8. Delete the default route
   
   `oc delete route api-apicast-staging`

9. Expose the apicast service as a route
   
   `oc create route passthrough --service=apicast --port=https --hostname=api-3scale-apicast-staging.$WILDCARD_DOMAIN`

   **NOTE**
   
   This step is needed for every API you are going to use and the domain changes for every API.

10. Verify that the previously deployed gateway works and the configuration was saved, by specifying [Your_user_key] in the placeholder.
   
   `curl https://api-3scale-apicast-staging.$WILDCARD_DOMAIN?user_key=[Your_user_key] -v --cacert ca/certs/ca.crt`

### 4.1.19.3. Configuring TLS Client Certificate Validation in your policy chain

#### 4.1.19.3.1. Prerequisites

- You need 3scale login credentials.
- You need to have configured APIcast with the TLS Client Certificate Validation policy.

#### 4.1.19.3.2. Configuring the policy

1. To add the TLS Client Certificate Validation policy to your API, follow the steps described in Enabling a standard Policy and choose TLS Client Certificate Validation.

2. Click the **TLS Client Certificate Validation** link.

3. To enable the policy, select the **Enabled** checkbox.

4. To add certificates to the whitelist, click the plus + icon.

5. Specify the certificate including `-----BEGIN CERTIFICATE-----` and `-----END CERTIFICATE-----`.

6. When you have finished setting up your API with TLS Client Certificate Validation, click **Update Policy**.

   Additionally:

   - You can add more certificates by clicking the plus + icon.
   - You can also reorganize the certificates by clicking the up and down arrows.

   To save your changes, click **Update & test in Staging Environment**
4.1.19.4. Verifying functionality of the TLS Client Certificate Validation policy

4.1.19.4.1. Prerequisites:

- You need 3scale login credentials.
- You need to have configured APIcast with the TLS Client Certificate Validation policy.

4.1.19.4.2. Verifying policy functionality

You can verify the applied policy by specifying [Your_user_key] in the placeholder.

```bash
curl https://api-3scale-apicast-staging.$WILDCARD_DOMAIN/?user_key=[Your_user_key] -v --cacert ca/certs/ca.crt --cert ca/certs/client.crt --key ca/keys/client.key

curl https://api-3scale-apicast-staging.$WILDCARD_DOMAIN/?user_key=[Your_user_key] -v --cacert ca/certs/ca.crt --cert ca/certs/server.crt --key ca/keys/server.key

curl https://api-3scale-apicast-staging.$WILDCARD_DOMAIN/?user_key=[Your_user_key] -v --cacert ca/certs/ca.crt
```

4.1.19.5. Removing a certificate from the whitelist

4.1.19.5.1. Prerequisites

- You need 3scale login credentials.
- You need to have set up APIcast with the TLS Client Certificate Validation policy.
- You need to have added the certificate to the whitelist, by configuring TLS Client Certificate Validation in your policy chain.

4.1.19.5.2. Removing a certificate

1. Click the TLS Client Certificate Validation link.
2. To remove certificates from the whitelist, click the x icon.
3. When you have finished removing the certificates, click Update Policy.

To save your changes, click Update & test in Staging Environment

4.1.19.6. Reference material

For more information about working with certificates, you can refer to Red Hat Certificate System.

4.1.20. TLS Termination policy

This section provides information about the Transport Layer Security (TLS) Termination policy: concepts, configuration, verification and file removal from the policy.

4.1.20.1. About TLS Termination policy

With the TLS Termination policy, you can configure APIcast to finish TLS requests for each API without
using a single certificate for all APIs. APIcast pulls the configuration setting before establishing a connection to the client; in this way, APIcast uses the certificates from the policy and makes the TLS terminate. This policy works with these sources:

- Stored in the policy configuration.
- Stored on the file system.

By default, this policy is not enabled in policy chains.

4.1.20.2. Configuring TLS Termination in your policy chain

This section describes the prerequisites and steps to configure the TLS Termination in your policy chain, with Privacy Enhanced Mail (PEM) formatted certificates.

4.1.20.2.1. Prerequisites

- Certificate issued by user
- A PEM-formatted server certificate
- A PEM-formatted certificate private key

4.1.20.2.2. Configuring the policy

1. To add the TLS Termination policy to your API, follow the steps described in Enabling a standard Policy and choose TLS Termination.

2. Click the **TLS Termination** link.

3. To enable the policy, select the **Enabled** checkbox.

4. To add TLS certificates to the policy, click the plus + icon.

5. Choose the source of your certificates:
   - Embedded certificate: Specify the path to the server certificate, and the path to the certificate private key.
   - Certificate from local file system: Browse the files for the certificate private key, and the server certificate.

6. When you have finished setting up your API with TLS Termination, click **Update Policy**.

Additionally:

- You can add more certificates by clicking the plus + icon.
- You can also reorganize the certificates by clicking the up and down arrows.

To save your changes, click **Update Policy Chain**.

4.1.20.3. Verifying functionality of the TLS Termination policy

4.1.20.3.1. Prerequisites:
• You need 3scale login credentials.
• You need to have configured APIcast with the TLS Termination policy.

4.1.20.3.2. Verifying policy functionality
You can test in the command line if the policy works with the following command:

curl "${public_URL}:${port}/?user_key=${user_key}" --cacert ${path_to_certificate}/ca.pem -v

where:
• public_URL = The staging public base URL
• port = The port number
• user_key = The user key you want to authenticate with
• path_to_certificate = The path to the CA certificate in your local file system

4.1.20.4. Removing files from TLS Termination
This section describes the steps to remove the certificate and key files from the TLS Termination policy.

4.1.20.4.1. Prerequisites
You need 3scale login credentials.
You need to have added the certificate to the policy, by configuring APIcast with the TLS Termination policy.

4.1.20.4.2. Removing a certificate
1. Click the TLS Termination link.
2. To remove certificates and keys, click the x icon.
3. When you have finished removing the certificates, click Update Policy.

To save your changes, click Update Policy Chain.

4.1.21. Upstream policy
The Upstream policy allows you to parse the Host request header using regular expressions and replace the upstream URL defined in the Private Base URL with a different URL.

For Example:

A policy with a regex /foo, and URL field newexample.com would replace the URL https://www.example.com/foo/123/ with newexample.com

Policy chain reference:
<table>
<thead>
<tr>
<th>property</th>
<th>description</th>
<th>values</th>
<th>required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>regex</td>
<td>The regex property allows you to specify the regular expression that the Upstream policy will use when searching for a match with the request path.</td>
<td>data type: string, Must be a valid regular expression syntax</td>
<td>yes</td>
</tr>
<tr>
<td>url</td>
<td>Using the url property, you can specify the replacement URL in the event of a match. Note that the upstream policy does not check whether or not this URL is valid.</td>
<td>data type: string, ensure this is a valid URL</td>
<td>yes</td>
</tr>
</tbody>
</table>

Policy object example

```json
{
    "name": "upstream",
    "version": "builtin",
    "configuration": {
        "rules": [
            {
                "regex": "^/v1/.*",
                "url": "https://api-v1.example.com",
            }
        ]
    }
}
```

For information on how to configure policies, refer to the Creating a policy chain section of the documentation.

**4.1.22. Upstream Connection policy**

**4.1.22.1. About Upstream Connection policy**

The Upstream Connection policy allows you to change the default values of the following directives, for each API, depending on how you have configured the API back end server in your 3scale installation:

- `proxy_connect_timeout`
- `proxy_send_timeout`
- `proxy_read_timeout`

**4.1.22.2. Configuring Upstream Connection in your policy chain**
4.1.22.2.1. Prerequisites:
- You need to have access to a 3scale installation.
- You need to wait for all the deployments to finish.

4.1.22.2.2. Configuring the policy

1. To add the Upstream Connection policy to your API, follow the steps described in Enabling a standard Policy and choose Upstream connection.
2. Click the Upstream connection link.
3. To enable the policy, select the Enabled checkbox.
4. Configure the options for the connections to the upstream:
   - send_timeout
   - connect_timeout
   - read_timeout
5. When you have finished setting up your API with Upstream connection, click Update Policy.

To save your changes, click Update & test in Staging Environment

4.1.23. URL Rewriting policy

The URL Rewriting policy allows you to modify the path of a request and the query string.

When combined with the 3scale APIcast policy, if the URL rewriting policy is placed before the 3scale APIcast policy in the policy chain, the APIcast mapping rules will apply to the modified path. If the URL rewriting policy is placed after APIcast in the policy chain, then the mapping rules will apply to the original path.

The policy supports the following two sets of operations:

- **commands**: List of commands to be applied to rewrite the path of the request.
- **query_args_commands**: List of commands to be applied to rewrite the query string of the request.

4.1.23.1. Commands for rewriting the path

Following are the configuration parameters that each command in the commands list consists of:

- **op**: Operation to be applied. The options available are: **sub** and **gsub**. The **sub** operation replaces only the first occurrence of a match with your specified regular expression. The **gsub** operation replaces all occurrences of a match with your specified regular expression. See the documentation for the **sub** and **gsub** operations.

- **regex**: Perl-compatible regular expression to be matched.

- **replace**: Replacement string that is used in the event of a match.
• **options** (optional): Options that define how the regex matching is performed. For information on available options, see the `ngx.re.match` section of the OpenResty Lua module project documentation.

• **break** (optional): When set to true (checkbox enabled), if the command rewrote the URL, it will be the last one applied (all posterior commands in the list will be discarded).

### 4.1.23.2. Commands for rewriting the query string

Following are configuration parameters that each command in the `query_args_commands` list consists of:

- **op**: Operation to be applied to the query arguments. The following options are available:
  - **add**: Add a value to an existing argument.
  - **set**: Create the arg when not set and replace its value when set.
  - **push**: Create the arg when not set and add the value when set.
  - **delete**: Delete an arg.

- **arg**: The query argument name that the operation is applied on.

- **value**: Specifies the value that is used for the query argument. For value type "liquid" the value should be in the format `{{ variable_from_context }}`. For the **delete** operation the value is not taken into account.

- **value_type** (optional): Defines how the query argument value is evaluated and can either be **plain** for plain text or **liquid** for evaluation as a Liquid template. For more information, see Section 5.1, “Using variables and filters in policies”. If not specified, the type "plain" is used by default.

**Example**

The URL Rewriting policy is configured as follows:

```json
{
    "name": "url_rewriting",
    "version": "builtin",
    "configuration": {
        "query_args_commands": [
            {
                "op": "add",
                "arg": "addarg",
                "value_type": "plain",
                "value": "addvalue"
            },
            {
                "op": "delete",
                "arg": "user_key",
                "value_type": "plain",
                "value": "any"
            },
            {
                "op": "push",
                "arg": "pusharg",
```
The original request URI that is sent to the APIcast:

https://api.example.com/api/v1/products/123/details?
user_key=abc123secret&pusharg=first&setarg=original

The URI that APIcast sends to the API backend after applying the URL rewriting:

https://api-backend.example.com/internal/products/123/details?
pusharg=first&pusharg=pushvalue&setarg=setvalue

The following transformations are applied:

1. The substring /api/v1/ matches the only path rewriting command and it is replaced by /internal/.

2. user_key query argument is deleted.

3. The value pushvalue is added as an additional value to the pusharg query argument.

4. The value original of the query argument setarg is replaced with the configured value setvalue.

5. The command add was not applied because the query argument addarg is not present in the original URL.

For information on how to configure policies, see the Creating a policy chain section of the documentation.

4.1.24. URL Rewriting with Captures policy

The URL Rewriting with Captures policy is an alternative to the Section 4.1.23, “URL Rewriting policy” policy and allows rewriting the URL of the API request before passing it to the API backend.

The URL Rewriting with Captures policy retrieves arguments in the URL and uses their values in the rewritten URL.
The policy supports the `transformations` configuration parameter. It is a list of objects that describe which transformations are applied to the request URL. Each transformation object consists of two properties:

- **match_rule**: This rule is matched to the incoming request URL. It can contain named arguments in the `{nameOfArgument}` format; these arguments can be used in the rewritten URL. The URL is compared to `match_rule` as a regular expression. The value that matches named arguments must contain only the following characters (in PCRE regex notation): `[w-~%!$&'()*,;=@:]`. Other regex tokens can be used in the `match_rule` expression, such as `^` for the beginning of the string and `$` for the end of the string.

- **template**: The template for the URL that the original URL is rewritten with; it can use named arguments from the `match_rule`.

The query parameters of the original URL are merged with the query parameters specified in the `template`.

**Example**

The URL Rewriting with Captures is configured as follows:

```json
{
  "name": "rewrite_url_captures",
  "version": "builtin",
  "configuration": {
    "transformations": [
      {
        "match_rule": "/api/v1/products/{productId}/details",
        "template": "/internal/products/details?id={productId}&extraparam=anyvalue"
      }
    ]
  }
}
```

The original request URI that is sent to the APIcast:

```
https://api.example.com/api/v1/products/123/details?user_key=abc123secret
```

The URI that APIcast sends to the API backend after applying the URL rewriting:

```
https://api-backend.example.com/internal/products/details?
user_key=abc123secret&extraparam=anyvalue&id=123
```

**4.2. ENABLING A STANDARD POLICY**

Perform the following steps to enable policies in the Admin Portal:

1. Log in to 3scale.
2. Navigate to the API service.
3. From [your_API_name] > Integration > Configuration select edit APIcast configuration.
4. Under the POLICIES section, click add policy.
5. Select the policy you want to add and fill out the required fields.

6. Click the **Update and test in Staging Environment** button to save the policy chain.

### 4.3. Creating Custom APIcast Policies

You can create custom APIcast policies entirely or modify the standard policies. In order to create custom policies, you must understand the following:

- Policies are written in Lua.
- Policies must adhere to and be placed in the proper file directory.
- Policy behavior is affected by how they are placed in a policy chain.
- The interface to add custom policies is fully supported, but not the custom policies themselves.

### 4.4. Adding Custom Policies to APIcast

If you have created custom policies, you must add them to APIcast. How you do this depends on where APIcast is deployed.

You can add custom policies to the following APIcast self-managed deployments:

- APIcast built-in gateways as part of a 3scale on-premises deployment on OpenShift
- APIcast on OpenShift and the Docker containerized environment

You cannot add custom policies to APIcast hosted.

**WARNING**

Never make policy changes directly onto a production gateway. Always test your changes.

### 4.4.1. Adding custom policies to the built-in APIcast

To add custom APIcast policies to an on-premises deployment, you must build an OpenShift image containing your custom policies and add it to your deployment. Red Hat 3scale provides a sample repository you can use as a framework to create and add custom policies to an on-premises deployment.

This sample repository contains the correct directory structure for a custom policy, as well as a template which creates an image stream and BuildConfigs for building a new APIcast OpenShift image containing any custom policies you create.
WARNING

When you build **apicast-custom-policies**, the build process “pushes” a new image to the **amp-apicast:latest** tag. When there is an image change on this image stream tag (``:latest``), both the **apicast-staging** and the **apicast-production** tags, by default, are configured to automatically start new deployment. To avoid any disruptions to your production service (or staging, if you prefer) it is recommended to disable automatic deployment ("Automatically start a new deployment when the image changes" checkbox), or configure a different image stream tag for production (e.g. **amp-apicast:production**).

Follow these steps to add a custom policy to an on-premises deployment:

1. Fork the [https://github.com/3scale/apicast-example-policy](https://github.com/3scale/apicast-example-policy) [public repository with the policy example] or create a private repository with its content. You need to have the code of your custom policy available in a Git repository for OpenShift to build the image. Note that in order to use a private Git repository, you must set up the secrets in OpenShift.

2. Clone the repository locally, add the implementation for your policy, and push the changes to your Git repository.

3. Update the **openshift.yml** template. Specifically, change the following parameters:
   a. **spec.source.git.uri**: `https://github.com/3scale/apicast-example-policy.git` in the policy BuildConfig – change it to your Git repository location.
   b. **spec.source.images[0].paths.sourcePath**: `/opt/app-root/policies/example` in the custom policies BuildConfig - change `example` to the name of the custom policy that you have added under the **policies** directory in the repository.
   c. Optionally, update the OpenShift object names and image tags. However, you must ensure that the changes are coherent (example: **apicast-example-policy** BuildConfig builds and pushes the **apicast-policy:example** image that is then used as a source by the **apicast-custom-policies** BuildConfig. So, the tag should be the same).

4. Create the OpenShift objects by running the command:
   ```bash
   oc new-app -f openshift.yml --param AMP_RELEASE=2.8.0
   ```

5. In case the builds do not start automatically, run the following two commands. In case you changed it, replace **apicast-example-policy** with your own BuildConfig name (e.g. **apicast-</name>-policy**). Wait for the first command to complete before you execute the second one.
   ```bash
   oc start-build apicast-example-policy
   oc start-build apicast-custom-policies
   ```

If the build-in APIcast images have a trigger on them tracking the changes in the **amp-apicast:latest** image stream, the new deployment for APIcast will start. After **apicast-staging** has restarted, go to the Integration page on the admin portal, and click the **Add Policy** button to see your custom policy listed. After selecting and configuring it, click **Update & test in Staging Environment** to make your custom policy work in the staging APIcast.
4.4.2. Adding custom policies to APIcast on another OpenShift Container Platform

You can add custom policies to APIcast on OpenShift Container Platform (OCP) by fetching APIcast images containing your custom policies from the Integrated OpenShift Container Platform registry.

Add custom policies to APIcast on another OpenShift Container Platform

1. Add policies to APIcast built-in
2. If you are not deploying your APIcast gateway on your primary OpenShift cluster, establish access to the internal registry on your primary OpenShift cluster.
3. Download the 3scale 2.8 APIcast OpenShift template.
4. To modify the template, replace the default image directory with the full image name in your internal registry.
   
   ```
   image: <registry>/<project>/amp-apicast:latest
   ```
5. Deploying APIcast using the OpenShift template, specifying your customized image:
   
   ```
   oc new-app -f customizedApicast.yml
   ```

**NOTE**

When custom policies are added to APIcast and a new image is built, those policies are automatically displayed as available in the Admin Portal when APIcast is deployed with the image. Existing services can see this new policy in the list of available policies, so it can be used in any policy chain.

When a custom policy is removed from an image and APIcast is restarted, the policy will no longer be available in the list, so you can no longer add it to a policy chain.

4.5. CREATING A POLICY CHAIN IN 3SCALE

Create a policy chain in 3scale as part of your APIcast gateway configuration. Follow these steps to modify the policy chain in the UI:

1. Log in to your AMP
2. Navigate to the API service
3. From [your_API_name] > Integration > Configuration select edit APIcast configuration

4. Under the POLICIES section, use the arrow icons to reorder policies in the policy chain. Always place the APIcast policy last in the policy chain.

5. Click the Update and test in Staging Environment button to save the policy chain

4.6. CREATING A POLICY CHAIN JSON CONFIGURATION FILE
If you are using a native deployment of APIcast, you can create a JSON configuration file to control your policy chain outside of the AMP.

A JSON configuration file policy chain contains a JSON array composed of the following information:

- the **services** object with an **id** value that specifies which service the policy chain applies to by number
- the **proxy** object, which contains the policy_chain and subsequent objects
- the **policy_chain** object, which contains the values that define the policy chain
- individual **policy** objects which specify both **name** and **configuration** data necessary to identify the policy and configure policy behavior

The following is an example policy chain for a custom policy `sample_policy_1` and the API introspection standard policy `token_introspection`:

```json
{
  "services": [ {
    "id": 1,
    "proxy": {
      "policy_chain": [
        {
          "name": "sample_policy_1", "version": "1.0",
          "configuration": {
            "sample_config_param_1": ["value_1"],
            "sample_config_param_2": ["value_2"]
          }
        },
        {
          "name": "token_introspection", "version": "builtin",
          "configuration": {
            "introspection_url": ["https://tokenauthorityexample.com"],
            "client_id": ["exampleName"],
            "client_secret": ["secretexamplekey123"]
          }
        },
        {
          "name": "apicast", "version": "builtin",
        }
      ]
    }
  }
}
```

All policy chains must include the built-in policy **apicast**. Where you place APIcast in the policy chain will affect policy behavior.
CHAPTER 5. INTEGRATING A POLICY CHAIN WITH APICAST NATIVE DEPLOYMENTS

For native APIcast deployments, you can integrate a custom policy chain by specifying a configuration file using the `THREESCALE_CONFIG_FILE` environment variable. The following example specifies the config file `example.json`:

```
THREESCALE_CONFIG_FILE=example.json bin/apicast
```

5.1. USING VARIABLES AND FILTERS IN POLICIES

Some Section 4.1, “APIcast standard policies” support Liquid templating that allows using not only plain string values, but also variables that are present in the context of the request.

To use a context variable, wrap its name in `{{ and }}`, example: `{{ uri }}`. If the variable is an object, you can also access its attributes, for example: `{{ somevar.attr }}`.

Following are the standard variables that are available in all the policies:

- **uri**: The path of the request with query parameters excluded from this path. The value of the embedded NGINX variable `$uri`.
- **host**: The host of the request (the value of the embedded NGINX variable `$host`).
- **remote_addr**: The IP address of the client (the value of the embedded NGINX variable `$remote_addr`).
- **headers**: The object containing the request headers. Use `{{headers['Some-Header']}}` to get a specific header value.
- **http_method**: The request method: GET, POST, etc.

These standard variables are used in the context of the request, but policies can add more variables to the context. A phase refers to all the execution steps that APIcast has. Variables can be used by all the policies in the policy chain, provided these cases:

- Within the same phase, if the variable is added in the policy and then used in the following policy after the addition.
- If a variable is added in a phase, this variable can be used in the next phases.

Following are some examples of variables that the standard 3scale APIcast policy adds to the context:

- **jwt**: A parsed JSON payload of the JWT token (for OpenID Connect authentication).
- **credentials**: An object that holds the application credentials. Example: "app_id": "972f7b4f", "user_key": "13b668c4d1e10eaebaa5144b4749713f".
- **service**: An object that holds the configuration for the service that the current request is handled by. Example: the service ID would be available as `{{ service.id }}`.

For a full list of objects and values available in the context, see the Section 4.110, “Liquid Context Debug policy”.

The variables are used with the help of Liquid templates. Example: `{{ remote_addr }}`, `{{`
headers['Some-Header'], {{ jwt.aud }}. The policies that support variables for the values have a special parameter, usually with the _type suffix (example: value_type, name_type, etc.) that accepts two values: "plain" for plain text and "liquid" for liquid template.

APIcast also supports Liquid filters that can be applied to the variables’ values. The filters apply NGINX functions to the value of the Liquid variable.

The filters are placed within the variable output tag {{ }}, following the name of the variable or the literal value by a pipe character | and the name of the filter. Examples:

- {{ 'username:password' | encode_base64 }}, where username:password is a variable.
- {{ uri | escape_uri }}.

Some filters do not require parameters, so you can use an empty string instead of the variable. Example: {{ " | utctime }} will return the current time in UTC time zone.

Filters can be chained as follows: {{ variable | function1 | function2 }}. Example: {{ " | utctime | escape_uri }}.

Following is the list of the available functions:

- escape_uri
- unescape_uri
- encode_base64
- decode_base64
- crc32_short
- crc32_long
- hmac_sha1
- md5
- md5_bin
- sha1_bin
- quote_sql_str
- today
- time
- now
- localtime
- utctime
- cookie_time
- http_time
• parse_http_time
CHAPTER 6. TRANSFORMING 3SCALE MESSAGE CONTENT USING POLICY EXTENSIONS IN FUSE

You can create highly flexible policy extensions for Red Hat 3scale API Management using Red Hat Fuse. You can do this by creating policy extensions in Fuse on OpenShift and configuring them in the 3scale Admin Portal. This enables you to use an APIcast proxy policy to perform complex transformations to request/response message content (for example, XML to JSON) that are implemented in the Apache Camel integration framework.

This also enables you to add or modify custom policy extensions dynamically in Camel, instead of rebuilding and redeploying the static APIcast container image. You can use any Camel Enterprise Integration Pattern (EIP) available using Camel Domain Specific Language (DSL) to implement an APIcast policy extension. This enables you to write policy extensions using a familiar programming language such as Java or XML. The example in this topic uses the Camel Netty4 HTTP component to implement the HTTP proxy in Java.

Required software components

You must have the following Red Hat Integration components deployed on the same OpenShift cluster:

- Fuse on OpenShift 7.6
- 3scale On-premises
- APIcast self-managed or embedded

**NOTE**

You should install these components on the same OpenShift cluster, but not necessarily in the same namespace or project. There is no direct communication between Fuse and 3scale. All communication occurs through OpenShift only.

Architecturally, it is best to have Fuse custom code running in a different project than 3scale. However, you must ensure that communication between both projects is possible. For details, see Configuring network policy with OpenShift SDN.

Additional resources

- Fuse on OpenShift Guide
- Apache Camel Component Reference - Netty4 HTTP component

6.1. INTEGRATING APICAST WITH APACHE CAMEL TRANSFORMATIONS IN FUSE

You can integrate APIcast with a transformation written as an Apache Camel application in Fuse on OpenShift. When the policy extension transformation is configured and deployed in 3scale, the 3scale traffic goes through the Camel policy extension, which transforms the message content. In this case, Camel works as a reverse HTTP proxy, where APIcast sends the 3scale traffic to Camel, and Camel then sends the traffic on to the API backend.

The example in this topic creates the HTTP proxy using the Camel Netty4 HTTP component:

- The request received over the HTTP proxy protocol is forwarded to the target service with the HTTP body converted to uppercase.
The response from the target service is processed by converting it to uppercase and then returned to the client.

Prerequisites

- You must have Fuse on OpenShift 7.6 and 3scale 2.8 deployed on the same OpenShift cluster. For installation details, see:
  - Fuse on OpenShift Guide
  - Installing 3scale

- You will need cluster administrator privileges to install Fuse on OpenShift and 3scale and to create projects. However, you can create deployment configurations, deploy pods, or create services with edit access privileges per project.

Procedure

1. Write an Apache Camel application in Java using the Camel netty4-http component to implement the HTTP proxy. You can then use any Camel component to transform the message. The following simple example performs an uppercase transformation of the request and the response from the service:

```java
import java.util.Locale;
import org.apache.camel.Exchange;
import org.apache.camel.Message;
import org.apache.camel.builder.RouteBuilder;

public class ProxyRoute extends RouteBuilder {

    @Override
    public void configure() throws Exception {
        from("netty4-http:proxy://0.0.0.0:8080")
            .process(ProxyRoute::uppercase)
            .toD("netty4-http:"
                + "${headers." + Exchange.HTTP_SCHEME + "}://
                + "${headers." + Exchange.HTTP_HOST + "}:
                + "${headers." + Exchange.HTTP_PORT + "}"
                + "${headers." + Exchange.HTTP_PATH + "}")
            .process(ProxyRoute::uppercase);
    }

    public static void uppercase(final Exchange exchange) {
        final Message message = exchange.getIn();
        final String body = message.getBody(String.class);
        message.setBody(body.toUpperCase(Locale.US));
    }
}
```

2. Deploy your Camel application as described in Creating and Deploying Applications on Fuse on OpenShift.

NOTE

The HTTP over TLS (HTTPS) protocol is not supported in this example.
6.2. CONFIGURING AN APICAST POLICY EXTENSION CREATED USING APACHE CAMEL IN FUSE ON OPENSIFT

After you have created the Apache Camel transformation, you can configure it as a policy extension in the policy chain in the 3scale Admin Portal.

The policy extension enables you to configure a 3scale service to use an HTTP proxy. This service is used to send the 3scale traffic over the HTTP proxy to perform request/response modifications in a third-party proxy. In this case, the third-party proxy is Apache Camel implemented in Fuse on OpenShift.

NOTE
The policy extension code is created in an Apache Camel application in Fuse on OpenShift and cannot be modified or deleted from 3scale.

Prerequisites

- You must have Fuse on OpenShift 7.6 and 3scale 2.8 deployed on the same OpenShift cluster. For installation details, see:
  - Fuse on OpenShift Guide
  - Installing 3scale
- You must have implemented a 3scale policy extension using an Apache Camel application in Fuse on OpenShift.
- You must have deployed the Apache Camel application in a pod and exposed it as a service on OpenShift. For details, see link:https://access.redhat.com/documentation/en-us/red_hat_fuse/7.6/html/fuse_on_openshift_guide/get-started-developers#create-and-deploy-applications[Creating and Deploying Applications on Fuse on OpenShift].

Procedure

1. Select Integration > Configuration > edit APIcast configuration in the 3scale Admin Portal.
2. Select POLICIES > Add Policy > Proxy service.
3. Enter the route for the Camel proxy service in the appropriate field:
   - **https_proxy**: Connects to an HTTPS proxy service.
   - **http_proxy**: Connects to an HTTP proxy service.
   - **all_proxy**: Connects to services when the protocol is not specified. For example, the following shows the route for an HTTP proxy service:
     ```
     http://camel-proxy.my-3scale-management-project.svc
     ```
4. Click Update & test in staging environment to apply the new proxy policy.
5. Test the HTTP proxy using `curl`. For example:
   ```
   curl "https://testapi-3scale-apicast-staging.myuser.app.dev.3sca.net:443/?user_key=MY_USER_KEY" -k
   ```
6. Confirm that the message content has been transformed, which in this example means converted to uppercase.

Additional resources

- 3scale HTTP proxy policy and example in GitHub
CHAPTER 7. APICAST ENVIRONMENT VARIABLES

APICAST environment variables allow you to modify behavior for APICAST. The following values are supported environment variables:

- **all_proxy** ALL_PROXY
- **APICAST_ACCESS_LOG_FILE**
- **APICAST_BACKEND_CACHE_HANDLER**
- **APICAST_CONFIGURATION_CACHE**
- **APICAST_CONFIGURATION_LOADER**
- **APICAST_CUSTOM_CONFIG**
- **APICAST_ENVIRONMENT**
- **APICAST_EXTENDED_METRICS**
- **APICAST_HTTPS_CERTIFICATE**
- **APICAST_HTTPS_CERTIFICATE_KEY**
- **APICAST_HTTPS_PORT**
- **APICAST_HTTPS_VERIFY_DEPTH**
- **APICAST_LOAD_SERVICES_WHEN_NEEDED**
- **APICAST_LOG_FILE**
- **APICAST_LOG_LEVEL**
- **APICAST_MANAGEMENT_API**
- **APICAST_MODULE**
- **APICAST_OIDC_LOG_LEVEL**
- **APICAST_PATH_ROUTING**
- **APICAST_POLICY_LOAD_PATH**
- **APICAST_PROXY_HTTPS_CERTIFICATE**
- **APICAST_PROXY_HTTPS_CERTIFICATE_KEY**
- **APICAST_PROXY_HTTPS_PASSWORD_FILE**
CHAPTER 7. APICAST ENVIRONMENT VARIABLES

- APICAST_PROXY_HTTPS_SESSION_REUSE
- APICAST_REPORTING_THREADS
- APICAST_RESPONSE_CODES
- APICAST_SERVICE_${ID}_CONFIGURATION_VERSION
- APICAST_SERVICES_LIST
- APICAST_SERVICES_FILTER_BY_URL
- APICAST_UPSTREAM_RETRY_CASES
- APICAST_WORKERS
- BACKEND_ENDPOINT_OVERRIDE
- HTTP_KEEPALIVE_TIMEOUT
- http_proxy HTTP_PROXY
- https_proxy HTTPS_PROXY
- no_proxy NO_PROXY
- OPENSSL_VERIFY
- OPENTRACING_CONFIG
- OPENTRACING_HEADER_FORWARD
- OPENTRACING_TRACER
- RESOLVER
- THREESCALE_CONFIG_FILE
- THREESCALE_DEPLOYMENT_ENV
- THREESCALE_PORTAL_ENDPOINT

**all_proxy, ALL_PROXY**
Default: no value  
Value: string  
Example: http://forward-proxy:80

Defines a HTTP proxy to be used for connecting to services if a protocol-specific proxy is not specified. Authentication is not supported.

**APICAST_ACCESS_LOG_FILE**
Default: stdout

Defines the file that will store the access logs.

**APICAST_BACKEND_CACHE_HANDLER**
Values: strict | resilient

Default: strict
Deprecated: Use the Caching policy instead.

Defines how the authorization cache behaves when backend is unavailable. Strict will remove cached application when backend is unavailable. Resilient will do so only on getting authorization denied from backend.

**APICAST_CONFIGURATION_CACHE**
Values: a number

Default: 0

Specifies the interval (in seconds) that the configuration will be stored for. The value should be set to 0 (not compatible with boot value of **APICAST_CONFIGURATION_LOADER**) or more than 60. For example, if **APICAST_CONFIGURATION_CACHE** is set to 120, the gateway will reload the configuration from the API manager every 2 minutes (120 seconds). A value < 0 disables reloading.

**APICAST_CONFIGURATION_LOADER**
Values: boot | lazy

Default: lazy

Defines how to load the configuration. Boot will request the configuration to the API manager when the gateway starts. Lazy will load it on demand for each incoming request (to guarantee a complete refresh on each request **APICAST_CONFIGURATION_CACHE** should be 0).

**APICAST_CUSTOM_CONFIG**
Deprecated: Use policies instead.

Defines the name of the Lua module that implements custom logic overriding the existing APIcast logic.

**APICAST_ENVIRONMENT**
Default:

Value: string[:]

Example: production:cloud-hosted

Double colon (:) separated list of environments (or paths) APIcast should load. It can be used instead of -e or --environment parameter on the CLI and for example stored in the container image as default environment. Any value passed on the CLI overrides this variable.

**APICAST_EXTENDED_METRICS**
Default: false

Value: boolean

Example: "true"

Enables additional information on Prometheus metrics. The following metrics have the service_id and service_system_name labels which provide more in-depth details about APIcast:

- total_response_time_seconds
- upstream_response_time_seconds
- upstream_status
APICAST_HTTPS_CERTIFICATE
Default: no value
Path to a file with X.509 certificate in the PEM format for HTTPS.

APICAST_HTTPS_CERTIFICATE_KEY
Default: no value
Path to a file with the X.509 certificate secret key in the PEM format.

APICAST_HTTPS_PORT
Default: no value
Controls on which port APIcast should start listening for HTTPS connections. If this clashes with HTTP port it will be used only for HTTPS.

APICAST_HTTPS_VERIFYDEPTH
Default: 1
Values: positive integers.
Defines the maximum length of the client certificate chain. If this parameter has 1 as its value, it implies that this length might include one additional certificate, for example intermediate CA.

APICAST_LOAD_SERVICES_WHEN_NEEDED
Values:
- true or 1 for true
- false, 0 or empty for false
Default: false
This option can be used when there are many services configured. However, its performance depends on additional factors, for example, the number of services, the latency between APIcast and the 3scale Admin Portal, and the Time To Live (TTL) of the configuration.

By default, APIcast loads all the services each time it downloads its configuration from the Admin Portal. When this option is enabled, the configurations uses lazy loading. APIcast will only load the ones configured for the host specified in the host header of the request.

NOTE
- The caching defined by APICASTCONFIGURATION_CACHE applies.
- This option will be disabled when APICASTCONFIGURATION_LOADER is boot.
- Not compatible with APICAST_PATH_ROUTING.

APICAST_LOG_FILE
Default: stderr
Defines the file that will store the OpenResty error log. It is used by bin/apicast in the error_log directive. Refer to NGINX documentation for more information. The file path can be either absolute, or relative to the prefix directory (apicast by default).
APICAST_LOG_LEVEL
Values: debug | info | notice | warn | error | crit | alert | emerg
Default: warn

Specifies the log level for the OpenResty logs.

APICAST_MANAGEMENT_API
Values:
  - disabled: completely disabled, just listens on the port
  - status: only the /status/ endpoints enabled for health checks
  - debug: full API is open

The Management API is powerful and can control the APIcast configuration. You should enable the debug level only for debugging.

APICAST_MODULE
Default: apicast

Deprecated: Use policies instead.

Specifies the name of the main Lua module that implements the API gateway logic. Custom modules can override the functionality of the default apicast.lua module. See an example of how to use modules.

APICAST_OIDC_LOG_LEVEL
Values: debug | info | notice | warn | error | crit | alert | emerg
Default: err

Allows to set the log level for the logs related to OpenID Connect integration.

APICAST_PATH_ROUTING
Values:
  - true or 1 for true
  - false, 0 or empty for false

When this parameter is set to true, the gateway will use path-based routing in addition to the default host-based routing. The API request will be routed to the first service that has a matching mapping rule, from the list of services for which the value of the Host header of the request matches the Public Base URL.

APICAST_POLICY_LOAD_PATH
Default: APICAST_DIR/policies

Value: string[:]

Example: ~/.apicast/policies:$PWD/policies

Double colon (:) separated list of paths where APIcast should look for policies. It can be used to first load policies from a development directory or to load examples.
CHAPTER 7. APICAST ENVIRONMENT VARIABLES

**APICAST_PROXY_HTTPS_CERTIFICATE**
Default:

Value: string

Example: /home/apicast/my_certificate.crt

The path to the client SSL certificate that APIcast will use when connecting with the upstream. Notice that this certificate will be used for all the services in the configuration.

**APICAST_PROXY_HTTPS_CERTIFICATE_KEY**
Default:

Value: string

Example: /home/apicast/my_certificate.key

The path to the key of the client SSL certificate.

**APICAST_PROXY_HTTPS_PASSWORD_FILE**
Default:

Value: string

Example: /home/apicast/passwords.txt

Path to a file with passphrases for the SSL cert keys specified with **APICAST_PROXY_HTTPS_CERTIFICATE_KEY**.

**APICAST_PROXY_HTTPS_SESSION_REUSE**
Default: on

Values:

- **on**: reuses SSL sessions.
- **off**: does not reuse SSL sessions.

**APICAST_REPORTING_THREADS**
Default: 0

Value: integer >= 0

**Experimental**: Under extreme load might have unpredictable performance and lose reports.

Value greater than 0 is going to enable out-of-band reporting to backend. This is a new experimental feature for increasing performance. Client won’t see the backend latency and everything will be processed asynchronously. This value determines how many asynchronous reports can be running simultaneously before the client is throttled by adding latency.

**APICAST_RESPONSE_CODES**

Values:

- **true**, **1** for true
- **false**, **0** or empty for false
When set to true, APIcast will log the response code of the response returned by the API backend in 3scale. In some plans this information can later be consulted from the 3scale Admin Portal. For more information, see Response Code Tracking.

**APICAST_SERVICE_${ID}_CONFIGURATION_VERSION**
Replace ${ID} with the actual Service ID. The value should be the configuration version you can see in the configuration history on the Admin Portal. Setting it to a particular version will prevent it from auto-updating and will always use that version.

**APICAST_SERVICES_LIST**
*Value:* a comma-separated list of service IDs

Used to filter the services configured in the 3scale API Manager, and only use the configuration for specific services in the gateway, discarding those service IDs that are not specified in the list. You can find the service IDs in the Admin Portal: Dashboard > [Your_API_name] tagged as ID for API calls.

**APICAST_SERVICES_FILTER_BY_URL**
*Value:* a PCRE (Perl Compatible Regular Expression) such as `.*.example.com`

Filters the services configured in the 3scale API Manager. This filter matches with the public base **URL**. Services that do not match the filter are discarded. If the regular expression cannot be compiled, no services are loaded.

**Example 7.1. example**

The Regexp filter `http://.*.foo.dev` is applied to the following backend endpoints:

1. http://staging.foo.dev
2. http://staging.bar.dev

In this case, 1 and 3 are configured in the embedded APIcast and 2 and 4 are discarded.

**APICAST_UPSTREAM_RETRY_CASES**
*Values:* error | timeout | invalid_header | http_500 | http_502 | http_503 | http_504 | http_403 | http_404 | http_429 | non_idempotent | off

**NOTE**

This is only used when the retry policy is configured and specifies when a request to the upstream API should be retried. It accepts the same values as Nginx’s **PROXY_NEXT_UPSTREAM Module**.
APICAST_WORKERS
Default: auto

Values: number | auto

This is the value that will be used in the nginx worker_processes directive. By default, APIcast uses auto, except for the development environment where 1 is used.

BACKEND_ENDPOINT_OVERRIDE
URI that overrides backend endpoint from the configuration. Useful when deploying outside OpenShift deployed AMP. Example: https://backend.example.com.

HTTP_KEEPALIVE_TIMEOUT
Default: 75 Value: positive integers Example: 1

This parameter sets a timeout during which a keep-alive client connection will stay open on the server side. The zero value disables keep-alive client connections.

By default, the gateway keeps HTTP_KEEPALIVE_TIMEOUT disabled. This configuration allows the use of the keepalive timeout from NGINX whose default value is 75 seconds.

http_proxy, HTTP_PROXY
Default: no value Value: string Example: http://forward-proxy:80

Defines a HTTP proxy to be used for connecting to HTTP services. Authentication is not supported.

https_proxy, HTTPS_PROXY
Default: no value Value: string Example: https://forward-proxy:443

Defines a HTTP proxy to be used for connecting to HTTPS services. Authentication is not supported.

no_proxy, NO_PROXY
Default: no value Value: string\[,<string>\]; *Example: foo,bar.com,.extra.dot.com

Defines a comma-separated list of hostnames and domain names for which the requests should not be proxied. Setting to a single * character, which matches all hosts, effectively disables the proxy.

OPENSSL_VERIFY
Values:

- 0, false: disable peer verification
- 1, true: enable peer verification

Controls the OpenSSL Peer Verification. It is off by default, because OpenSSL can’t use system certificate store. It requires custom certificate bundle and adding it to trusted certificates.

It is recommended to use https://github.com/openresty/lua-nginx-module#lua_ssl_trusted_certificate and point to to certificate bundle generated by export-builtin-trusted-certs.

OPENTRACING_CONFIG
This environment variable is used to determine the config file for the opentracing tracer, if OPENTRACING_TRACER is not set, this variable will be ignored.

Each tracer has a default configuration file: * jaeger: conf.d/opentracing/jaeger.example.json
You can choose to mount a different configuration than the provided by default by setting the file path using this variable.

Example: /tmp/jaeger/jaeger.json

OPENTRACING_HEADER_FORWARD
Default: uber-trace-id

This environment variable controls the HTTP header used for forwarding opentracing information, this HTTP header will be forwarded to upstream servers.

OPENTRACING_TRACER
Example: jaeger

This environment variable controls which tracing library will be loaded, right now, there’s only one opentracing tracer available, jaeger.

If empty, opentracing support will be disabled.

RESOLVER
Allows to specify a custom DNS resolver that will be used by OpenResty. If the RESOLVER parameter is empty, the DNS resolver will be autodiscovered.

THREESCALE_CONFIG_FILE
Path to the JSON file with the configuration for the gateway. You can download the configuration from the Admin Portal using the URL: <schema>://<admin-portal-domain>/admin/api/nginx/spec.json (Example: https://account-admin.3scale.net/admin/api/nginx/spec.json).

When the gateway is deployed using Docker, the file has to be injected to the docker image as a read only volume, and the path should indicate where the volume is mounted, i.e. path local to the docker container.

You can find sample configuration files in examples folder.

It is required to provide either THREESCALE_PORTAL_ENDPOINT or THREESCALE_CONFIG_FILE (takes precedence) for the gateway to run successfully.

THREESCALE_DEPLOYMENT_ENV
Values: staging | production

Default: production

The value of the environment variable defines the environment for a particular configuration. The configuration will be downloaded for that environment from 3scale (Staging or Production) when using new APIcast.

The value of THREESCALE_DEPLOYMENT_ENV is also used in the X-3scale-User-Agent header for the authorize and report requests made to 3scale Service Management API. The header value is used by 3scale for statistics.

THREESCALE_PORTAL_ENDPOINT
URI that includes your password and portal endpoint in the following format: <schema>://<password>@<admin-portal-domain>. The <password> can be either the provider key or an access token for the 3scale Account Management API. <admin-portal-domain> is the URL to log in to the Admin Portal.
Example: https://access-token@account-admin.3scale.net.

When **THREESCALE_PORTAL_ENDPOINT** environment variable is provided, the gateway downloads the configuration from 3scale on initializing. The configuration includes all the settings provided on the Integration page of the API(s).

You can also use this environment variable to create a single gateway with the Master Admin Portal.

It is **required** to provide either **THREESCALE_PORTAL_ENDPOINT** or **THREESCALE_CONFIG_FILE** (takes precedence) for the gateway to run successfully.
CHAPTER 8. CONFIGURING APICAST FOR BETTER PERFORMANCE

This document provides general guidelines to debug performance issues in APIcast. It also introduces the available caching modes and explains how they can help in increasing performance, as well as details about profiling modes. The content is structured in the following sections:

- Section 8.1, “General guidelines”
- Section 8.2, “Default caching”
- Section 8.3, “Asynchronous reporting threads”
- Section 8.4, “3scale Batcher policy”

8.1. GENERAL GUIDELINES

In a typical APIcast deployment, there are three components to consider:

- APIcast
- The 3scale back-end server that authorizes requests and keeps track of the usage
- The upstream API

When experiencing performance issues in APIcast:

- Identify the component that is responsible for the issues.
- Measure the latency of the upstream API, to determine the latency that APIcast plus the 3scale back-end server introduce.
- With the same tool you are using to run the benchmark, perform a new measurement but pointing to APIcast instead of pointing to the upstream API directly.

Comparing these results will give you an idea of the latency introduced by APIcast and the 3scale back-end server.

In a Hosted (SaaS) installation with self-managed APIcast, if the latency introduced by APIcast and the 3scale back-end server is high:

1. Make a request to the 3scale back-end server from the same machine where APIcast is deployed
2. Measure the latency.

The 3scale back-end server exposes an endpoint that returns the version: https://su1.3scale.net/status. In comparison, an authorization call requires more resources because it verifies keys, limits, and queue background jobs. Although the 3scale back-end server performs these tasks in a few milliseconds, it requires more work than checking the version like the /status endpoint does. As an example, if a request to /status takes around 300 ms from your APIcast environment, an authorization is going to take more time for every request that is not cached.

8.2. DEFAULT CACHING
For requests that are not cached, these are the events:

1. APIcast extracts the usage metrics from matching mapping rules.
2. APIcast sends the metrics plus the application credentials to the 3scale back-end server.
3. The 3scale back-end server performs the following:
   a. Checks the application keys, and that the reported usage of metrics is within the defined limits.
   b. Queues a background job to increase the usage of the metrics reported.
   c. Responds to APIcast whether the request should be authorized or not.
4. If the request is authorized, it goes to the upstream.

In this case, the request does not arrive to the upstream until the 3scale back-end server responds.

On the other hand, with the caching mechanism that comes enabled by default:

- APIcast stores in a cache the result of the authorization call to the 3scale back-end server if it was authorized.
- The next request with the same credentials and metrics will use that cached authorization instead of going to the 3scale back-end server.
- If the request was not authorized, or if it is the first time that APIcast receives the credentials, APIcast will call the 3scale back-end server synchronously as explained above.

When the authentication is cached, APIcast first calls the upstream and then, in a phase called post action, it calls the 3scale back-end server and stores the authorization in the cache to have it ready for the next request. Notice that the call to the 3scale back-end server does not introduce any latency because it does not happen in request time. However, requests sent in the same connection will need to wait until the post action phase finishes.

Imagine a scenario where a client is using keep-alive and sends a request every second. If the upstream response time is 100 ms and the latency to the 3scale back-end server is 500 ms, the client will get the response every time in 100 ms. The total of upstream response and the reporting would take 600 ms. That gives extra 400 ms before the next request comes.

The diagram below illustrates the default caching behavior explained. The behavior of the caching mechanism can be changed using the caching policy.
8.3. ASYNCHRONOUS REPORTING THREADS

APIcast has a feature to enable a pool of threads that authorize against the 3scale back-end server. With this feature enabled, APIcast first synchronously calls the 3scale back-end server to verify the application and metrics matched by mapping rules. This is similar to when it uses the caching mechanism enabled by default. The difference is that subsequent calls to the 3scale back-end server are reported fully asynchronously as long as there are free reporting threads in the pool.

Reporting threads are global for the whole gateway and shared between all the services. When a second
TCP connection is made, it will also be fully asynchronous as long as the authorization is already cached. When there are no free reporting threads, the synchronous mode falls back to the standard asynchronous mode and does the reporting in the post action phase.

You can enable this feature using the `APICAST_REPORTING_THREADS` environment variable.

The diagram below illustrates how the asynchronous reporting thread pool works.

---

**8.4. 3SCALE BATCHER POLICY**

By default, APIcast performs one call to the 3scale back-end server for each request that it receives. The goal of the 3scale Batcher policy is to reduce latency and increase throughput by significantly reducing the number of requests made to the 3scale back-end server. In order to achieve that, this policy caches authorization statuses and batches reports.

Section 4.1.2, “3scale Batcher policy” provides details about the 3scale Batcher policy. The diagram below illustrates how the policy works.
CHAPTER 9. EXPOSING 3SCALE APICAST METRICS TO PROMETHEUS

IMPORTANT

For this release of Red Hat 3scale, Prometheus installation and configuration are not supported. Optionally, you can use the community version of Prometheus to visualize metrics and alerts for APIcast-managed API services.

9.1. ABOUT PROMETHEUS

Prometheus is an open-source systems monitoring toolkit that you can use to monitor 3scale APIcast services deployed in the Red Hat OpenShift environment.

If you want to monitor your services with Prometheus, your services must expose a Prometheus endpoint. This endpoint is an HTTP interface that exposes a list of metrics and the current value of the metrics. Prometheus periodically scrapes these target-defined endpoints and writes the collected data into its database.

9.1.1. Prometheus queries

In the Prometheus UI, you can write queries in Prometheus Query Language (PromQL) to extract metric information. With PromQL, you can select and aggregate time series data in real time.

For example, you can use the following query to select all the values that Prometheus has recorded within the last 5 minutes for all time series that have the metric name `http_requests_total`:

```
http_requests_total[5m]
```

You can further define or filter the results of a query by specifying a `label` (a key:value pair) for the metric. For example, you can use the following query to select all the values that Prometheus has recorded within the last 5 minutes for all time series that have the metric name `http_requests_total` and a `job` label set to `integration`:

```
http_requests_total{job="integration"}[5m]
```

The result of a query can either be shown as a graph, viewed as tabular data in Prometheus’s expression browser, or consumed by external systems by using the Prometheus HTTP API. Prometheus provides a graphical view of the data. For a more robust graphical dashboard to view Prometheus metrics, Grafana is a popular choice.

You can also use the the PromQL language to configure alerts in the Prometheus alertmanager tool.

NOTE

Grafana is a community-supported feature. Deploying Grafana to monitor Red Hat 3scale products is not supported with Red Hat production service level agreements (SLAs).

9.2. APICAST INTEGRATION WITH PROMETHEUS

APIcast integration with Prometheus is available for the following deployment options:
- Self-managed APIcast (both with hosted or on-premises API manager)
- Built-in APIcast on-premise

**NOTE**

APIcast integration with Prometheus is not available in hosted API manager and hosted APIcast.

By default, Prometheus can monitor the APIcast metrics listed in Table 9.2, “Prometheus Default Metrics for 3scale APIcast”.

### 9.2.1. Additional options

 Optionally, if you have cluster admin access to the OpenShift cluster, you can extend the `total_response_time_seconds`, `upstream_response_time_seconds`, and `upstream_status` metrics to include `service_id` and `service_system_name` labels. To extend these metrics, set the `APICAST_EXTENDED_METRICS` OpenShift environment variable to `true` with this command:

```
oc set env dc/apicast APICAST_EXTENDED_METRICS=true
```

If you use the APIcast Batch policy (described in Section 4.1.2, “3scale Batcher policy”), Prometheus can also monitor the metrics listed in Table 9.3, “Prometheus Metrics for 3scale APIcast Batch Policy”.

**NOTE**

If a metric has no value, Prometheus hides the metric. For example, if `nginx_error_log` has no errors to report, Prometheus does not display the `nginx_error_log` metric. The `nginx_error_log` metric is only visible if it has a value.

### Additional resources

- For information about how to use Prometheus, go to:

  https://prometheus.io/docs/prometheus/latest/getting_started/

### 9.3. OPENSHIFT ENVIRONMENT VARIABLES FOR 3SCALE APICAST

To configure your Prometheus instance, you can set the OpenShift environment variable described in Table 9.1, “Prometheus Environment Variables for 3scale APIcast”.

**Table 9.1. Prometheus Environment Variables for 3scale APIcast**

<table>
<thead>
<tr>
<th>Environment Variable</th>
<th>Description</th>
<th>Default</th>
</tr>
</thead>
</table>

APICAST_EXTENDED_METRICS

A boolean value that enables additional information on Prometheus metrics. The following metrics have the `service_id` and `service_system_name` labels which provide more in-depth details about APIcast:

- `total_response_time_seconds`
- `upstream_response_time_seconds`
- `upstream_status`

Environment Variable | Description | Default
--- | --- | ---
APICAST_EXTENDED_METRICS | A boolean value that enables additional information on Prometheus metrics. The following metrics have the `service_id` and `service_system_name` labels which provide more in-depth details about APIcast: | false

Additional resources
For information on setting environment variables, see the relevant OpenShift guide:

- The OpenShift 4 Applications guide (https://access.redhat.com/documentation/en-us/openshift_container_platform/4.2/html-single/applications/)
- For more information about supported configurations, see the Red Hat 3scale API Management Supported Configurations page.

9.4. 3SCALE APICAST METRICS EXPOSED TO PROMETHEUS

After you set up Prometheus to monitor 3scale APIcast, by default it can monitor the metrics listed in Table 9.2, “Prometheus Default Metrics for 3scale APIcast”.

The metrics listed in Table 9.3, “Prometheus Metrics for 3scale APIcast Batch Policy” are only available when you use the Section 4.1.2, “3scale Batcher policy”.

Table 9.2. Prometheus Default Metrics for 3scale APIcast

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Type</th>
<th>Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>nginx_http_connections</td>
<td>Number of HTTP connections</td>
<td>gauge</td>
<td>state(accepted,active,handled,reading,total,waiting,writing)</td>
</tr>
<tr>
<td>nginx_error_log</td>
<td>APIcast errors</td>
<td>counter</td>
<td>level(debug,info,notice,warn,error,crit,alert, emerg)</td>
</tr>
<tr>
<td>Metric</td>
<td>Description</td>
<td>Type</td>
<td>Labels</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>openresty_shdict_capacity</td>
<td>Capacity of the dictionaries shared between workers</td>
<td>gauge</td>
<td>dict(one for every dictionary)</td>
</tr>
<tr>
<td>openresty_shdict_free_space</td>
<td>Free space of the dictionaries shared between workers</td>
<td>gauge</td>
<td>dict(one for every dictionary)</td>
</tr>
<tr>
<td>nginx_metric_errors_total</td>
<td>Number of errors of the Lua library that manages the metrics</td>
<td>counter</td>
<td>none</td>
</tr>
<tr>
<td>total_response_time_seconds</td>
<td>Time needed to send a response to the client (in seconds)</td>
<td>histogram</td>
<td>service_id, service_system_name</td>
</tr>
<tr>
<td></td>
<td><strong>Note</strong>: To access the service_id and service_system_name labels, you must set the APICAST_EXTENDED_D_METRICS environment variable to true as described in Section 9.2, “APIcast integration with Prometheus”.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>upstream_response_time_seconds</td>
<td>Response times from upstream servers (in seconds)</td>
<td>histogram</td>
<td>service_id, service_system_name</td>
</tr>
<tr>
<td></td>
<td><strong>Note</strong>: To access the service_id and service_system_name labels, you must set the APICAST_EXTENDED_D_METRICS environment variable to true as described in Section 9.2, “APIcast integration with Prometheus”.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**upstream_status**

HTTP status from upstream servers

*Note:* To access the `service_id` and `service_system_name` labels, you must set the `APICAST_EXTENDED_METRICS` environment variable to `true` as described in Section 9.2, "APIcast integration with Prometheus".

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Type</th>
<th>Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>upstream_status</code></td>
<td>HTTP status from upstream servers</td>
<td>counter</td>
<td><code>status, service_id, service_system_name</code></td>
</tr>
</tbody>
</table>

**threescale_backend_calls**

Authorize and report requests to the 3scale backend (Apisonator)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Type</th>
<th>Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>threescale_backend_calls</code></td>
<td>Authorize and report requests to the 3scale backend (Apisonator)</td>
<td>counter</td>
<td><code>endpoint(authrep, auth, report), status(2xx, 4xx, 5xx)</code></td>
</tr>
</tbody>
</table>

### Table 9.3. Prometheus Metrics for 3scale APIcast Batch Policy

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Type</th>
<th>Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>batching_policy_auths_cache_hits</code></td>
<td>Hits in the auths cache of the 3scale batching policy</td>
<td>counter</td>
<td>none</td>
</tr>
<tr>
<td><code>batching_policy_auths_cache_misses</code></td>
<td>Misses in the auths cache of the 3scale batching policy</td>
<td>counter</td>
<td>none</td>
</tr>
</tbody>
</table>
PART II. API VERSIONING
CHAPTER 10. API VERSIONING

Red Hat 3scale API Management allows API versioning. You have three ways to version your API correctly when you manage your API with 3scale. The following methods are examples of how you could version your API within the 3scale Gateway, which provides extra features due to 3scale architecture.

10.1. GOAL

This guide is designed to give you enough information to implement an API versioning system within 3scale.

Suppose you have an API for finding songs. Users can search for their favorite songs by different keywords: artist, songwriter, song title, album title, and so on. Assume you had an initial version (v1) of the API and now you have developed a new, improved version (v2).

The following sections describe the three most typical ways of implementing an API versioning system using 3scale:

- URL versioning
- Endpoint versioning
- Custom header versioning

10.2. PREREQUISITES

Complete the basics of connecting your API to 3scale before using this quick start guide.

10.3. URL VERSIONING

If you have different endpoints for searching songs (by artist, by song title, and so on), with URL versioning you would include the API version as part of the URI, for example:

1. api.songs.com/v1/songwriter
2. api.songs.com/v2/songwriter
3. api.songs.com/v1/song
4. api.songs.com/v2/song
5. and so on

**NOTE**

When you use this method, you should have planned since v1 that you were going to version your API.

The 3scale Gateway would then extract the endpoint and the version from the URI. This approach allows you to set up application plans for any version/endpoint combination. You can then associate metrics with those plans and endpoints, and you can chart the usage for each endpoint on each version.

The following screen capture shows 3scale’s flexibility.
The only thing left to do is go to [your_API_name] > Integration > Configuration in your 3scale Admin Portal and map your URIs to your metrics, as shown in the following diagram.

You now have two different versions of your API, each with different features enabled. You also have full control and visibility on their usage.
If you want to communicate to all of your users that they should move to the API v2, you can send an internal note asking them to do so. You can monitor who makes the move and see how the activity on v1 decreases while the activity on v2 increases. By adding the metric in your authorization calls to 3scale, you can see how much overall traffic is hitting v1 vs. v2 endpoints and get an idea of when it is safe to deprecate v1.

Figure 10.3. Versioning

If some users continue to use v1, you can filter out only those users to send another internal note about switching to v2.

3scale provides a three-step method for sending deprecation notices.

1. Navigate to **Audience > Applications > Listing** and filter the list by the application plan that you want to send the deprecation note and click **Search**.

2. Click the multiselector to select all of the applications for that particular version. New options display and allow you to perform bulk operations, such as **Send email**, **Change Application Plan**, and **Change State**.

3. Click **Send email** and follow the steps to send a deprecation notice to the owners of the selected applications.

The following image provides a visual reference.

Figure 10.4. Sending deprecation note
For each authrep call that is made to an endpoint, you authenticate only once but report twice: once for the endpoint and once for the API version. There is no double-billing because the call can be authenticated only one time. For each call you make to any endpoint of a specific API version, you aggregate the hits on a convenient metric named after the version number (v1, v2, and so on), which you can use to compare full version traffic with each other.

10.4. ENDPOINT VERSIONING

With endpoint versioning, you can have a different endpoint for each API version such as api.cons.com/author_v1. The gateway extracts the endpoint and the version from the endpoint itself. This method, as well as the previous method, allows the API provider to map external URLs to internal ones.

The endpoint versioning method can only be performed with the on-premise deployment method as it requires a URL rewrite using the LUA scripts that are provided as part of the on-premise configuration.

<table>
<thead>
<tr>
<th>EXTERNAL</th>
<th>INTERNAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>api.songs.com/songwriter_v1</td>
<td>could be rewritten to internal.songs.com/search_by_songwriter</td>
</tr>
<tr>
<td>api.songs.com/songwriter_v2</td>
<td>could be rewritten to internal.songs.com/songwriter</td>
</tr>
</tbody>
</table>

Almost everything (mapping, application plans features, and so on,) works exactly the same as in the previous method.

10.5. CUSTOM HEADER VERSIONING

With custom header versioning, you use a header (that is, "x-api-version") instead of the URI to specify the version.

The gateway then extracts the endpoint from the path and the version from the header. Just as before, you can analyze and visualize any combination of path/version that you want. This approach has several inconveniences, regardless of the API management system you use. See API versioning methods, a brief reference for more information. Here are a few pointers on how 3scale works.

- Just like the previous method, custom header versioning can only be applied to on-premise hosted APIs because it requires some parsing/processing of the request headers to correctly route the authrep calls. This type of custom processing can only be done using Lua scripting.

- With this method, the fine-grained feature separation of the previous methods is much harder to achieve.

- The most important advantage of this method is that the URL and endpoints specified by the developers never change. When a developer wants to switch from one API version to another, they only have to change the header. Everything else works the same.
CHAPTER 11. AUTHENTICATION PATTERNS

By the end of this tutorial you will know how to set the authentication pattern on your API and the effect that this has on applications communicating with your API.

Depending on your API, you may need to use different authentication patterns to issue credentials for access to your API. These can range from API keys to openAuth tokens and custom configurations. This tutorial covers how to select from the available standard Authentication Patterns.

11.1. SUPPORTED AUTHENTICATION PATTERNS

3scale supports the following authentication patterns out of the box:

- **Standard API Keys**: Single randomized strings or hashes acting as an identifier and a secret token.
- **Application Identifier and Key pairs**: Immutable identifier and mutable secret key strings.
- **OpenID Connect**

11.2. SETTING UP AUTHENTICATION PATTERNS

11.2.1. Select the authentication mode for your service

Navigate to the API service you want to work on (there may be only one service named API in which case select this). Go to the **Integration** section.

Each service that you operate can use a different authentication pattern, but only one pattern can be used per service.
11.2.2. Select the Authentication mode you want to use

To select an authentication mode, scroll to the AUTHENTICATION section. Here, you can choose one of the following options:

- API Key (user_key)
- App_ID and App_Key Pair
- OpenID Connect

11.2.3. Ensure your API accepts the correct types of credentials

Depending on the credential type chosen, you may need to accept different parameters in your API calls (key fields, IDs etc.). The names of these parameters may not be the same as those used internally at 3scale. The 3scale authentication will function correctly if the correct parameter names are used in calls to the 3scale backend.

11.2.4. Create an application to test credentials

To ensure that the credential sets are working, you can create a new application to issue credentials to use the API. Navigate to the Accounts area of your Admin Portal Dashboard, click the account you want to use and click new application.

Filling out the form and clicking save will create a new application with credentials to use the API. You can now use these credentials to make calls to your API and records will be checked against the list of applications registered in 3scale.

11.3. STANDARD AUTHENTICATION PATTERNS

3scale supports the authentication patterns detailed in the following sections.

11.3.1. API key

The simplest form of credential supported is the single API model. Here, each application with permissions on the API has a single (unique) long character string; example:

```
API-key = 853a76f7c8d5f4a1ee8bf10a4e0d1f13
```

By default, the name of the key parameter is user_key. You can use this label or choose another, such as API-key. If choosing another label, you need to map the value before you make the authorization calls to 3scale. The string acts as both, an identifier and a secret token, for use of the API. It is recommended that you use such patterns only in environments with low security requirements or with SSL security on API calls. Following are the operations that can be carried out on the token and application:

- Application Suspend: This suspends the applications access to the API and, in effect, all calls to the API with the relevant key will be suspended.
• Application Resume: Undoes the effect of an application suspend action.

• Key Regenerate: This action generates a new random string key for the application and associates it with the application. Immediately after this action is taken, calls with the previous token will cease to be accepted.

The latter action can be triggered from the API Administration in the Admin Portal and (if permitted) from the API Developers User console.

11.3.2. App_ID and App_Key pair

The API Key Pattern combines the identity of the application and the secret usage token in one token; however, this pattern separates the two:

• Each application using the API, issues an immutable initial identifier known as the Application ID (App ID). The App ID is constant and may or may not be secret.

• In addition, each application can have between one and five Application Keys (App_Keys). Each Key is associated directly with the App_ID and should be treated as secret.

```
app_id = 80a4e03 app_key = a1ee8bf10a4e0d1f13853a76f7c8d5f4
```

In the default setting, developers can create up to five keys per application. This allows a developer to create a new key, add it to their code, redeploy their application, and then disable old keys. This does not cause any application downtime the way an API Key Regeneration would.

Statistics and rate limits are always kept at the application ID level of granularity and not per API Key. If a developer wants to track two sets of statistics, they should create two applications rather than two keys.

It is also possible to change the mode in the system and allow applications to be created in the absence of application keys. In this case the 3scale system will authenticate access based on the App ID only (and no key checks are made). This mode is useful for widget type scenarios or where rate limits are applied to users rather than applications. In most cases you will want your API to enforce the presence of at least one application key per application present. This setting is available in [your_API_name] > Integration > Settings.

11.3.3. OpenID Connect

For information on OpenID Connect authentication, see the OpenID Connect integration chapter.

11.4. REFERRER FILTERING

3scale supports the Referrer Filtering feature that can be used to whitelist IP addresses or domain names from where an application can access the API. The API clients specify the referrer value in the Referrer header. The purpose and the usage of the Referrer header are described in the RFC 7231, section 5.5.2: Referrer.

To enable the Referrer Filtering feature go to [your_API_name] > Integration > Settings click the Require referrer filtering checkbox and click Update Service.
The developers with access to your API must configure allowed domain/IP referrers from the developer portal.

In the Admin Portal on the application details page for all applications that belong to this service a new Referrer Filters section displays. Here, the admin can also configure a whitelist of the allowed Referrer header values for this application.
You can set a maximum of five referrer values per application.

The value can only consist of Latin letters, numbers, and special characters *, ., and -. * can be used for wildcard values. If the value is set to *, any referrer value will be allowed, so the referrer check will be bypassed.

For the Referrer Filtering feature to work, you must enable the APIcast Referrer policy in the service policy chain.

When the Require referrer filtering feature and the 3scale Referrer policy are enabled, the authorization works as follows:

1. The applications that do not have Referrer Filters specified are authorized normally only using the provided credentials.

2. For the applications that have Referrer Filters values set, APIcast extracts the referrer value from the Referer header of the request and sends it as referrer param in the AuthRep (authorize and report) request to the Service Management API. The following table shows the AuthRep responses for different combination of the referrer filtering parameters.

<table>
<thead>
<tr>
<th>referrer parameter passed?</th>
<th>Referrer Filters configured for the app?</th>
<th>Referrer parameter value</th>
<th>HTTP Response</th>
<th>Response body</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>matches referrer filter</td>
<td>200 OK</td>
<td>&lt;status&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;authorized&gt;true&lt;/authorized&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;/status&gt;</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>matches referrer filter</td>
<td>200 OK</td>
<td>&lt;status&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;authorized&gt;true&lt;/authorized&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;/status&gt;</td>
</tr>
<tr>
<td>referrer parameter passed?</td>
<td>Referrer Filters configured for the app?</td>
<td>Referrer parameter value</td>
<td>HTTP Response</td>
<td>Response body</td>
</tr>
<tr>
<td>---------------------------</td>
<td>----------------------------------------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>does not match referrer filter</td>
<td>409 Conflict</td>
<td><code>&lt;status&gt; &lt;authorized&gt;false&lt;/authorized&gt; &lt;reason&gt;referrer &quot;test.example.com&quot; is not allowed&lt;/reason&gt; (test.example.com is an example)</code></td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>does not match referrer filter</td>
<td>200 OK</td>
<td><code>&lt;status&gt; &lt;authorized&gt;true&lt;/authorized&gt; &lt;/status&gt;</code></td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>*</td>
<td>200 OK</td>
<td><code>&lt;status&gt; &lt;authorized&gt;true&lt;/authorized&gt; &lt;/status&gt;</code></td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>*</td>
<td>200 OK</td>
<td><code>&lt;status&gt; &lt;authorized&gt;true&lt;/authorized&gt; &lt;/status&gt;</code></td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>–</td>
<td>409 Conflict</td>
<td><code>&lt;status&gt; &lt;authorized&gt;false&lt;/authorized&gt; &lt;reason&gt;referrer is missing&lt;/reason&gt;</code></td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>–</td>
<td>200 OK</td>
<td><code>&lt;status&gt; &lt;authorized&gt;true&lt;/authorized&gt; &lt;/status&gt;</code></td>
</tr>
</tbody>
</table>

The calls that are not authorized by AuthRep are rejected by APIcast with an "Authorization Failed" error. You can configure the exact status code and the error message on the service Integration page.
3scale integrates with the 3rd-party Identity Providers (IdP) for authenticating the API requests using the OpenID Connect specification. OpenID Connect is built on top of OAuth 2.0 that complements the OAuth 2.0 Authorization framework with an authentication mechanism. When OpenID Connect authentication option is used, the API requests are authenticated using the access tokens in the JSON Web Token (JWT) format (RFC 7519).

The integration consists of the following two parts:

- Section 12.1, “JWT verification and parsing by APIcast”
- Section 12.2, “Client credentials synchronization by Zync”

Red Hat 3scale API Management fully supports both integration points with Red Hat Single Sign-On (RH-SSO) acting as the OpenID Provider. See the supported version of RH-SSO on the Supported Configurations page. APIcast integration is also tested with ForgeRock.

In both cases, you can configure the integration by specifying the OpenID Connect Issuer field in the APIcast Configuration on the Integration page of the service using OpenID Connect authentication option. For instructions, see Section 12.3, “Configure Red Hat Single Sign-On integration”.

12.1. JWT VERIFICATION AND PARSGING BY APICAST

The API requests to the service using the OpenID Connect authentication mode should provide the access token in the JWT format, issued by the OpenID Provider, in the Authorization header using Bearer schema. The header should look like the following example:

```
Authorization: Bearer <JWK>
```

Example:

```
Authorization: Bearer: eyJhbGciOiJSUzI1NiIsInR5cCI6IkpXVCJ9.eyJpc3MiOiJSb3IiLCJlbmNyaW Dunkc3MiOiJSb3IiLCJleHAiOjE1Mzg0NjU2NzwiZXhwIjoxNTI0NjE5MzIzLCJpYXQiOjE1Mzg0NjU2Nzi--o0c3uv1R4FTCQGbO1SASHQZqG5Ux5ugaASY-hUHlohy6Pc7dQi0e2NIaegqg4MuZtEwpEPJW-VnGdjrAS0HsXzd6nENM0Z_ofo4ZdTkeVks2KrOvBObcigvYjYongtppR0cw30FwntqfeCkuATeIINN5OKHXOiRA24pQylf18s1C0mnxLm5nVbuv24SFE34aMGRXyZs4icMi8sK65ekxbvvV3PIG3mM0C4iZP26doP0YrLlWfCqEirmENUAcHxz7NuvA
```

The JWT token contains a signature that the token’s receiver can verify and ensure that the token was signed by a known issuer and that its content has not been changed. 3scale supports RSA signature based on the public/private key pair. Here, the issuer signs the JWT token using a private key. APIcast verifies this token using a public key.

APIcast uses OpenID Connect Discovery for getting the JSON Web Keys (JWK) that can be used for verifying the JWT signature.

On each request, APIcast does the following:

1. Verifies the JWT token using the public key.
2. Validates the claims nbf and exp.

3. Verifies that the issuer specified in the claim iss (Issuer) is the same as the one configured in the OpenID Connect Issuer field.

4. Extracts the value of the azp or aud claim and uses it as the Client ID (that identifies the application in 3scale) to authorize the call through the Service Management API.

If any of the JWT validation or the authorization checks fail, APIcast returns an “Authentication failed” error. Otherwise, APIcast proxies the request to the API backend. The Authorization header remains in the request, so the API backend can use the JWT token to check the user and client identity.

12.2. CLIENT CREDENTIALS SYNCHRONIZATION BY ZYNC

Using the Zync component, 3scale synchronizes the client (application) credentials between 3scale and the Red Hat Single Sign-On server configured through the OpenID Connect Issuer setting. Whenever a new application is created, updated, or deleted for a service configured to use OpenID Connect, Zync receives the corresponding event and communicates the change to the Red Hat Single Sign-On instance using RH-SSO API. The Section 12.3, “Configure Red Hat Single Sign-On integration” section provides the steps required to ensure that Zync has the correct credentials to use the RH-SSO API.

12.3. CONFIGURE RED HAT SINGLE SIGN-ON INTEGRATION

12.3.1. Configure Zync to use custom CA certificates

You must establish an SSL connection between Zync and Red Hat Single Sign-On. 3scale 2.2 and above supports custom CA certificates for Red Hat Single Sign-On with the SSL_CERT_FILE environment variable. This variable points to the local path of the certificates bundle. Configure it as follows:

1. Validate the new certificate with the following cURL command. The expected response is a JSON configuration of the realm. If validation fails it is an indicator that your certificate may not be correct.

```
curl -v https://<secure-sso-host>/auth/realms/master --cacert customCA.pem
```

2. Add the certificate bundle to the Zync pod:

   a. Gather the existing content of the /etc/pki/tls/cert.pem file on the Zync pod. Run:

```
   oc exec <zync-pod-id> cat /etc/pki/tls/cert.pem > zync.pem
```

   b. Append the contents of the custom CA certificate file to zync.pem:

```
cat customCA.pem >> zync.pem
```

   c. Attach the new file to the Zync pod as ConfigMap:

```
oc create configmap zync-ca-bundle --from-file=./zync.pem
```

```oc set volume dc/zync --add --name=zync-ca-bundle --mount-path /etc/pki/tls/zync/zync.pem --sub-path zync.pem --source={"configMap":{"name":"zync-ca-bundle","items":[{"key":"zync.pem","path":"zync.pem"}]}}```
3. After deployment, verify that the certificate is attached and the content is correct:

```bash
oc exec <zync-pod-id> cat /etc/pki/tls/zync/zync.pem
```

4. Configure the `SSL_CERT_FILE` environment variable on Zync to point to the new CA certificate bundle:

```bash
oc set env dc/zync SSL_CERT_FILE=/etc/pki/tls/zync/zync.pem
```

### 12.3.2. Configure Red Hat Single Sign-On

To configure Red Hat Single Sign-On, take the following steps:

1. Create a realm (`<REALM_NAME>`).
2. Create a client:
   a. Specify a client ID.
   b. In the `Client Protocol` field, select `openid-connect`.
3. To configure the client permissions, set the following values:
   a. `Access Type` to `confidential`.
   b. `Standard Flow Enabled` to `OFF`.
   c. `Direct Access Grants Enabled` to `OFF`.
   d. `Service Accounts Enabled` to `ON`.
4. Set the service account roles for the client:
   a. Navigate to the `Service Account Roles` tab of the client.
   b. In the `Client Roles` dropdown list, click `realm-management`.
   c. In the `Available Roles` pane, select the `manage-clients` list item and assign the role by clicking `Add selected >>`.
5. Note the client credentials:
   a. Make a note of the client ID (`<CLIENT_ID>`).
   b. Navigate to the `Credentials` tab of the client and make a note of the `Secret` field (`<CLIENT_SECRET>`).
6. Add a user to the realm:
   a. Click the `Users` menu on the left side of the window.
   b. Click `Add user`.
   c. Type the username, set the `Email Verified` switch to `ON`, and click `Save`.
d. On the **Credentials** tab, set the password. Enter the password in both the fields, set the **Temporary** switch to **OFF** to avoid the password reset at the next login, and click **Reset Password**.

e. When the pop-up window displays, click **Change password**.

**12.3.3. Configure 3scale**

After you have created and configured the client in Red Hat Single Sign-On, you must configure 3scale to work with Red Hat Single Sign-On.

To configure 3scale, take the following steps:

1. Enable OpenID Connect.
   a. Select the service on which you want to enable the OpenID Connect authentication, navigate to [your_API_name] > Integration > Configuration
   b. Select **edit integration settings**
   c. Under the **Authentication** deployment options, select **OpenID Connect**
   d. Click **Update Service** to save the settings.

2. Edit the APIcast Configuration:
   a. Navigate to [your_API_name] > Integration > Configuration
   b. Select **edit APIcast configuration**
   c. Under the **Authentication Settings** heading, in the **OpenID Connect Issuer** field, enter the previously noted client credentials with the URL of your Red Hat Single Sign-On server (located at host `<RHSSO_HOST>` and port `<RHSSO_PORT>`).

```
https://<CLIENT_ID>:<CLIENT_SECRET>@<RHSSO_HOST>:<RHSSO_PORT>/auth/realms/<REALM_NAME>
```
   d. To save the configuration, click **Update the Staging Environment**

**12.4. CONFIGURE HTTP INTEGRATION WITH THIRD-PARTY IDENTITY PROVIDERS**

You can configure HTTP integration of OpenID Connect (OIDC) to facilitate syncing credentials with third-party identity providers (IdPs). This means that it is possible to integrate different IdPs other than Red Hat Single Sign-On (RH-SSO), by implementing the OpenAPI specifications we provide.

**12.4.1. Pre-requisites**

- Enable OIDC as authentication mode, as indicated in **Configure 3scale**
- Zync
- Integration with Zync for client synchronization between chosen IdP and 3scale

**12.4.2. Procedure**
To configure HTTP integration of OIDC with third-party identity providers, follow these steps in the Admin Portal:

1. Navigate to [Your_API_name] > Integration > edit APIcast configuration > Authentication Settings.
2. Under OpenID Connect Issuer Type, select REST API.
3. In OpenID Connect Issuer, specify the location of your OpenID Provider.
4. To save your changes, click Update the Staging Environment

12.4.3. Zync REST API example

This example project implements Zync REST API protocol to synchronize OAuth2.0 clients. When a 3scale application is created, updated or deleted Zync tries to replicate that change to http://example.com/api.

12.4.3.1. Prerequisites

3scale must be configured to use:

- OIDC as the authentication mode
- REST API as a OpenID Connect Issuer Type
- http://id:secret@example.com/api as OpenID Connect Issuer

12.4.3.2. Creating, updating and deleting clients

Zync makes the following requests to create, update or delete clients:

- Create and update → PUT /clients/:client_id
- Delete → DELETE /clients/:client_id

All endpoints must reply with a 2xx status code. Otherwise, the request will be retried.

12.4.3.3. Payload

The request payload in case of create and update is application/json:

```json
{
    "client_id": "ee305610",
    "client_secret": "ac0e42db426b4377096c6590e2b06aed",
    "client_name": "oidc-app",
    "redirect_uris": ["http://example.com"],
    "grant_types": ["client_credentials", "password"]
}
```

The request to delete a client has no payload.

12.4.3.4. Using OAuth2 authentication
Zync sends GET requests to the `/.well-known/openid-configuration` endpoint and expects an `application/json` response. The response payload should contain the following:

```json
{
    "token_endpoint": "http://idp.example.com/auth/realm/token"
}
```

With the OAuth2 protocol, Zync uses `token_endpoint` to exchange the `client_id` and `client_secret` provided in the OpenID Connect Issuer address, in order to request an access token. If the API responds with an unsuccessful response, Zync will fallback to HTTP Basic/Digest authentication using the provided credentials.

### 12.5. OAUTH 2.0 SUPPORTED FLOWS

The API clients must get access tokens from the OpenID Connect (OIDC) issuer configured in 3scale, using any OAuth 2.0 flow that is supported by this OpenID provider. In case of Red Hat Single Sign-On (RH-SSO), the following flows are supported (the terms used in RH-SSO clients are specified in parenthesis):

- Authorization Code (Standard Flow)
- Resource Owner Password Credentials (Direct Access Grants Flow)
- Implicit (Implicit Flow)
- Client Credentials (Service Accounts Flow)

In 3scale, if you create a client for which OIDC authentication is enabled, the corresponding clients created by Zync in RH-SSO have only the Authorization Code flow enabled. This flow is recommended as the most secure and suitable for most cases. However, it is possible to enable other flows.

#### 12.5.1. How OAuth 2.0 supported flows work

The client gets the access token using the authorization request, or the token request, or both. The URLs that receive these requests can be discovered using the `/.well-known/openid-configuration` endpoint of the OpenID provider, in the "authorization_endpoint" and "token_endpoint", accordingly. Example: `https://<RHSSO_HOST>:<RHSSO_PORT>/auth/realms/<REALM_NAME>/`.well-known/openid-configuration.

#### 12.5.2. Configuring OAuth 2.0 supported flows

You can configure allowed OAuth 2.0 flows for the 3scale API in the Admin Portal. When you create a new application, the basic integration is finished, including the OpenID Connect (OIDC) configuration.

To configure OAuth 2.0 supported flows, perform these steps:

1. Navigate to the Authentication Settings section: [Your_API_name] > Integration > edit integration settings > Authentication
2. Choose OpenId Connect.
3. The corresponding flows are enabled on the client on RH SSO side. You can view them by navigating through [Your_API_name] > Integration > Edit APIcast configuration > Authentication Settings
   - `standardFlowEnabled` (Authorization Code flow) [selected by default]
Choose one or multiple flows.

To save your changes, click **Update the Staging Environment**

### 12.6. TEST THE INTEGRATION

To test the integration, you must perform the steps listed in the following sections.

#### 12.6.1. Test the client synchronization

To test the client synchronization, take the following steps:

1. Create an application for the service where you configured the OpenID Connect integration.
2. Note the client ID and the client Secret of the generated application.
3. Verify that the client with the same client ID and client secret is now present in the configured
   Red Hat Single Sign-On realm.
4. Update the Redirect URL of the application in the 3scale admin portal. Redirect URLs should be
   as specific as possible.
5. Verify that the Valid Redirect URIs field of the client in Red Hat Single Sign-On has been
   updated accordingly.

#### 12.6.2. Test the API authorization flow

To test the API authorization flow, take the following steps:

1. Get the access token from the Red Hat Single Sign-On server using an OAuth 2.0 flow that is
   enabled on the corresponding RH-SSO client.
2. Use the value of the **access_token** retrieved from RH-SSO in the **Authorization** header as
   follows: `Authorization: Bearer <access_token>`

If the token is correct and the corresponding application in 3scale is authorized, APIcast gateway returns
a response from the API backend.

### 12.7. EXAMPLE OF THE INTEGRATION

The service "API" in 3scale is configured to use the OpenID Connect authentication. The **Public Base URL**
on the service "API" is configured to be **https://api.example.com** and the **Private Base URL** is
configured to be **https://internal-api.example.com**.

The **OpenID Connect Issuer** field is set to **https://zync:41dbb98b-e4e9-4a89-84a3-91d1d19c4207@idp.example.com/auth/realms/myrealm** in the API integration and the client **zync** in
the realm **myrealm** has the correct Service Account roles.

In 3scale, there is an application having the **myclientid** client ID, **myclientsecret** client secret, and a
In Red Hat Single Sign-On, in the myrealm realm, there also exists a client having a myclientid client ID, myclientsecret secret, and https://myapp.example.com Valid Redirect URIs. Standard Flow is enabled on this client. There is a user configured in the myrealm realm having the myuser username and mypassword password.

The flow is as follows:

1. Using the endpoint https://idp.example.com/auth/realms/myrealm/protocol/openid-connect/auth, the application sends an Authorization request to RH-SSO. Within the request, the application provides these parameters: myclientid client ID, and https://myapp.example.com Redirect URL.

2. RH-SSO shows the login window, where the user must provide the user’s credentials: Username myuser and password mypassword.

3. Depending on the configuration, and if it is the first time that the user is authenticating in this specific application, the consent window displays.

4. After the user is authenticated, the application sends a Token request to RH-SSO using the endpoint https://idp.example.com/auth/realms/myrealm/protocol/openid-connect/token and providing the client ID myclientid, client secret myclientsecret and Redirect URL https://myapp.example.com.

5. RH-SSO returns a JSON with an "access_token" field eyJhbGciOiJSUzI1NiIsInR5cCIgOiAiSldUIiwia2lk...xBArNhqF-A.

6. The application sends an API request to https://api.example.com with the header Authorization: Bearer eyJhbGciOiJSUzI1NiIsInR5cCIgOiAiSldUliwia2lk...xBArNhqF-A.

7. The application should receive a successful response from https://internal-api.example.com.