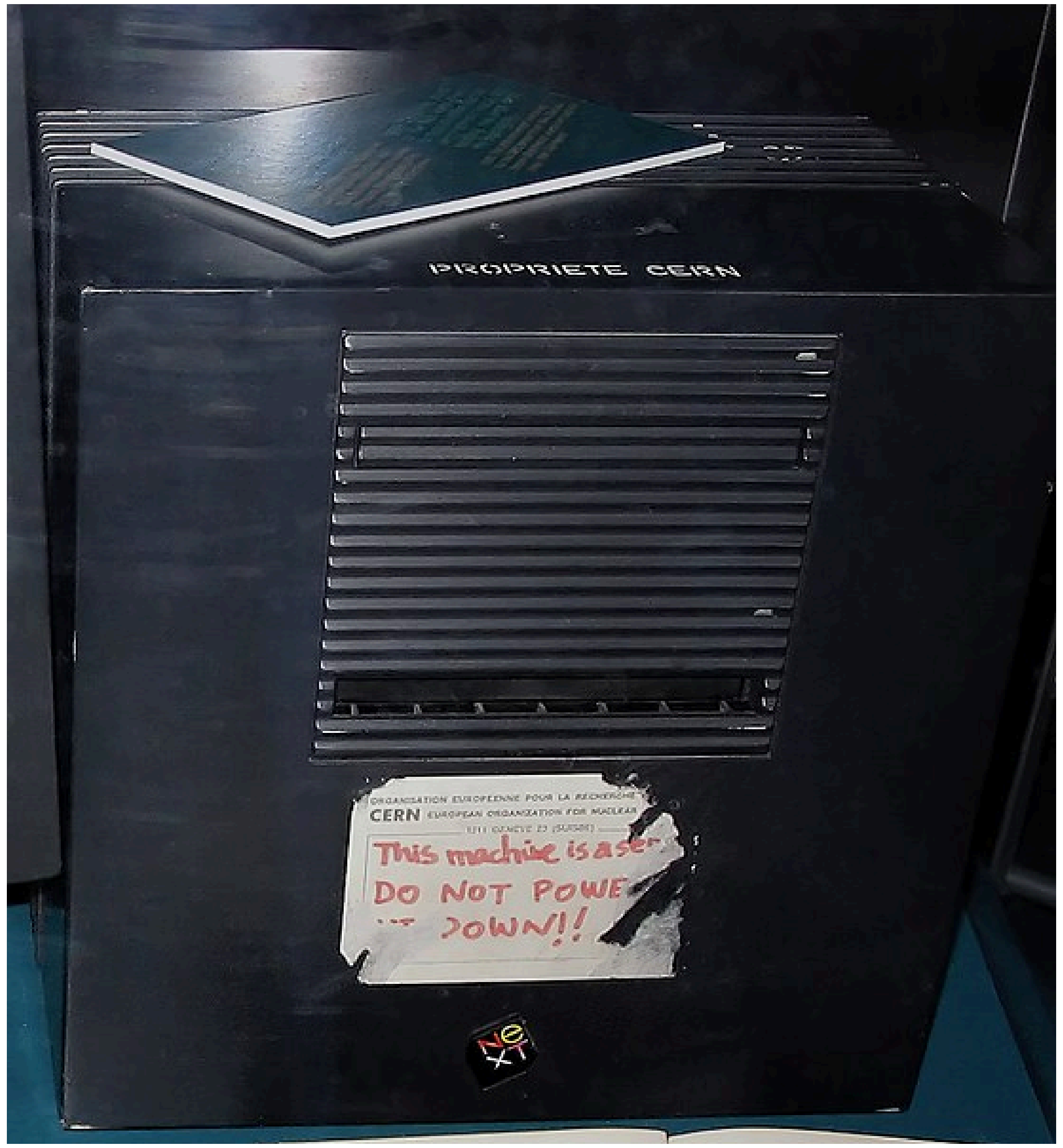
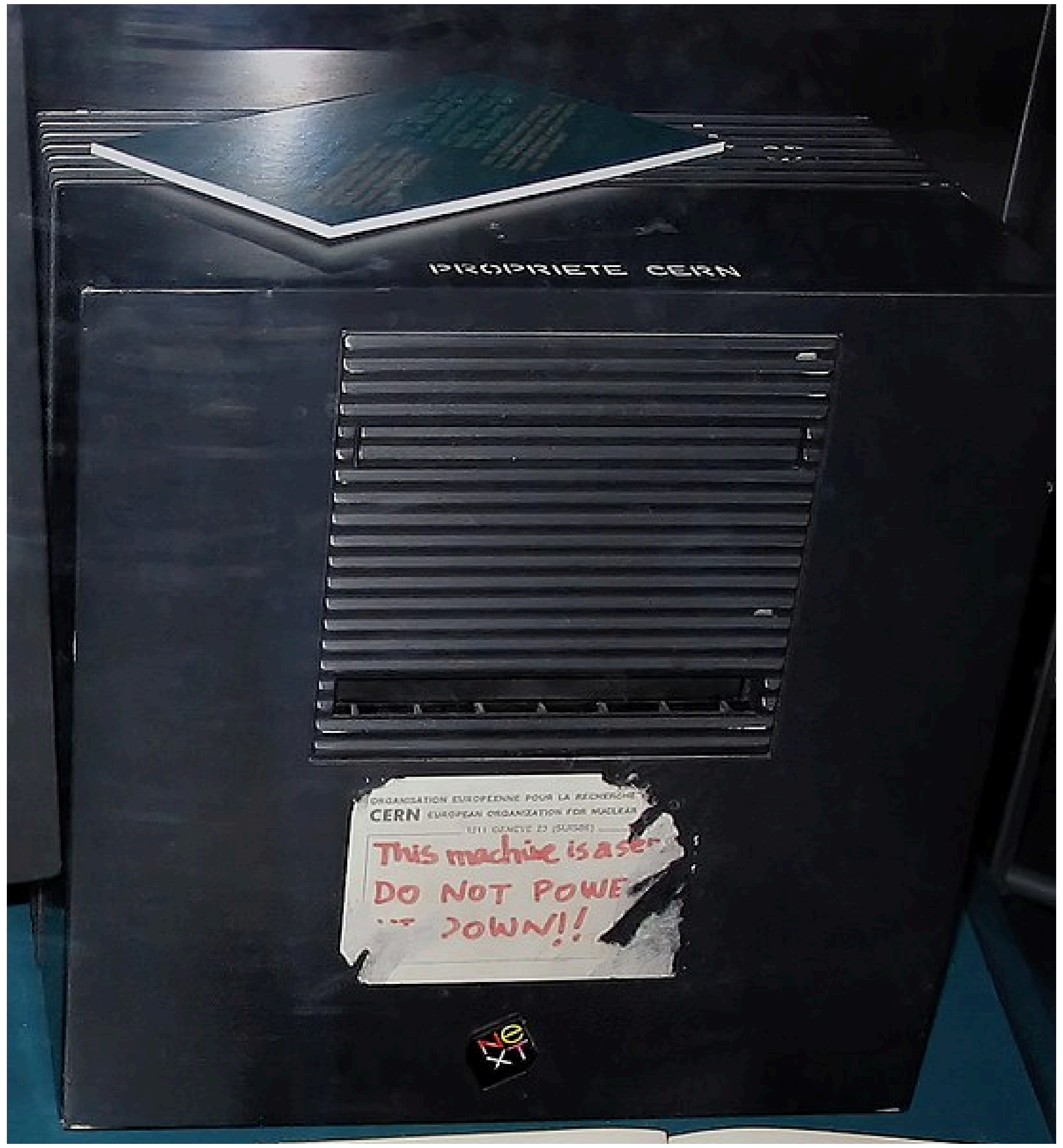
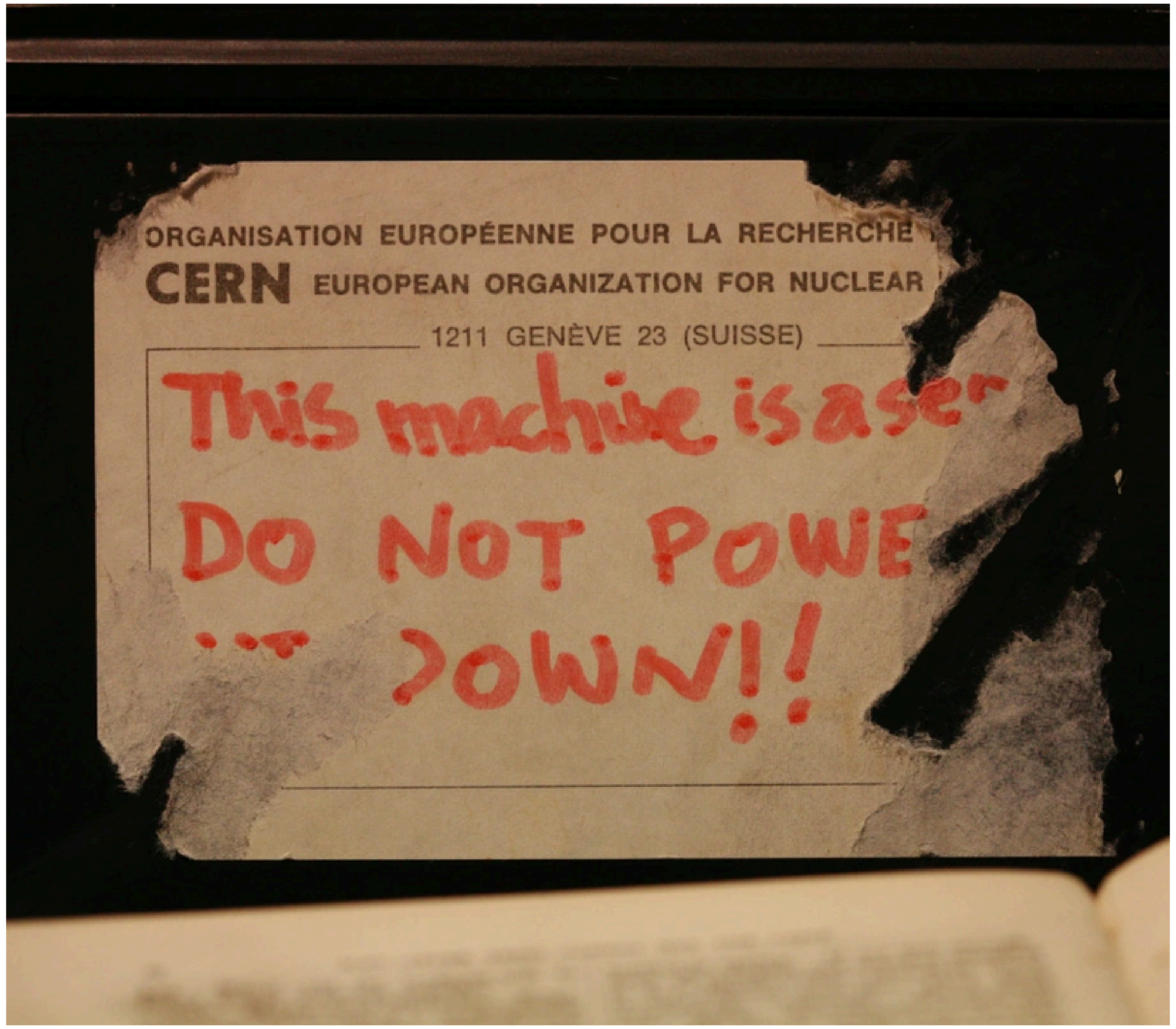
Deploying applications

Alberto Pimpo

```
[ber@apimpo-laptop]~% whoami
Alberto Pimpo
Dev Ops Engineer (IT-PW-AH)
Formerly BE-ICS
[ber@apimpo-laptop]~%
```


PROPRIETE CERN
ORGANISATION EUROPEENNE POUR LA RECHERCHE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR
1211 GENEVE 23 (SUISSE)
This machine is a ser
DO NOT POWE
DOWN!!
NeXT


PROPRIETE CERN
ORGANISATION EUROPÉENNE POUR LA RECHERCHE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR
1211 GENÈVE 23 (SUISSE)
This machine is a ser
DO NOT POWE
DOWN!!
NeXT


ORGANISATION EUROPÉENNE POUR LA RECHERCHE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR
1211 GENÈVE 23 (SUISSE)
This machine is a ser
DO NOT POWE
DOWN!!

What can go wrong?

Deploying correctly is hard,
we need to rely on services.

Let’s make a pizza

Buy all the ingredients

Let’s make a pizza

Buy all the ingredients

Buy a premade dough

Let’s make a pizza

Buy all the ingredients

Buy a premade dough

Just order it

Paradigms of cloud computing

- **IaaS**
  - virtualized components that replicates the physical architectures (VMs, Block Storage, etc...)
- **PaaS**
  - a platform for building and deploying applications without managing the infrastructure
- **SaaS**
  - just use the software

Some examples

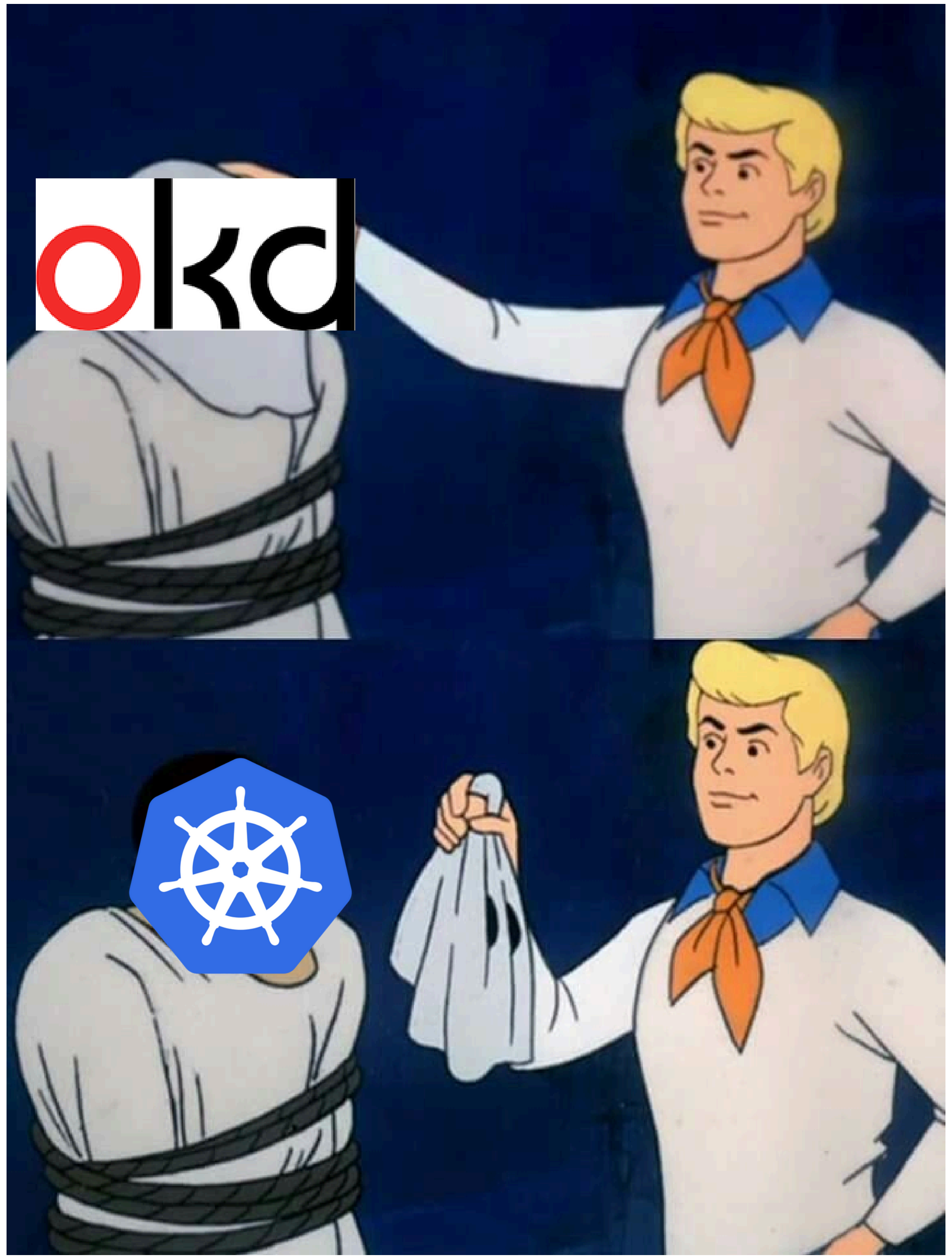
SaaS	Grafana GitLab
PaaS	DBoD okd Kubeflow reana
IaaS	openstack®

Our PaaS platform

Let’s see together how it works

https://paas.cern.ch/

It’s Kubernetes under the hood



A package manager for kubernetes

Let’s see together an example

Thanks!

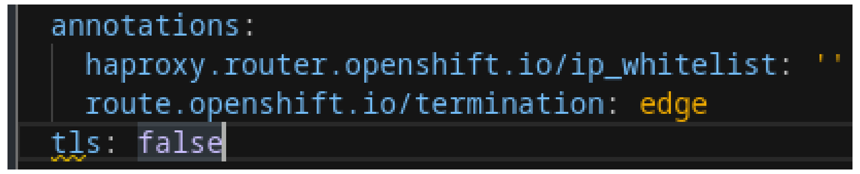
Exercise 1

- Create a project on paas-stg (use webservices-portal) with *Test* category
- Deploy the application from *Modern application and deployment* lecture
    - using S2I (Import from git)
    - deploying the container image directly
- Look at the logs of your app on logs.paas.cern.ch

Exercise 2

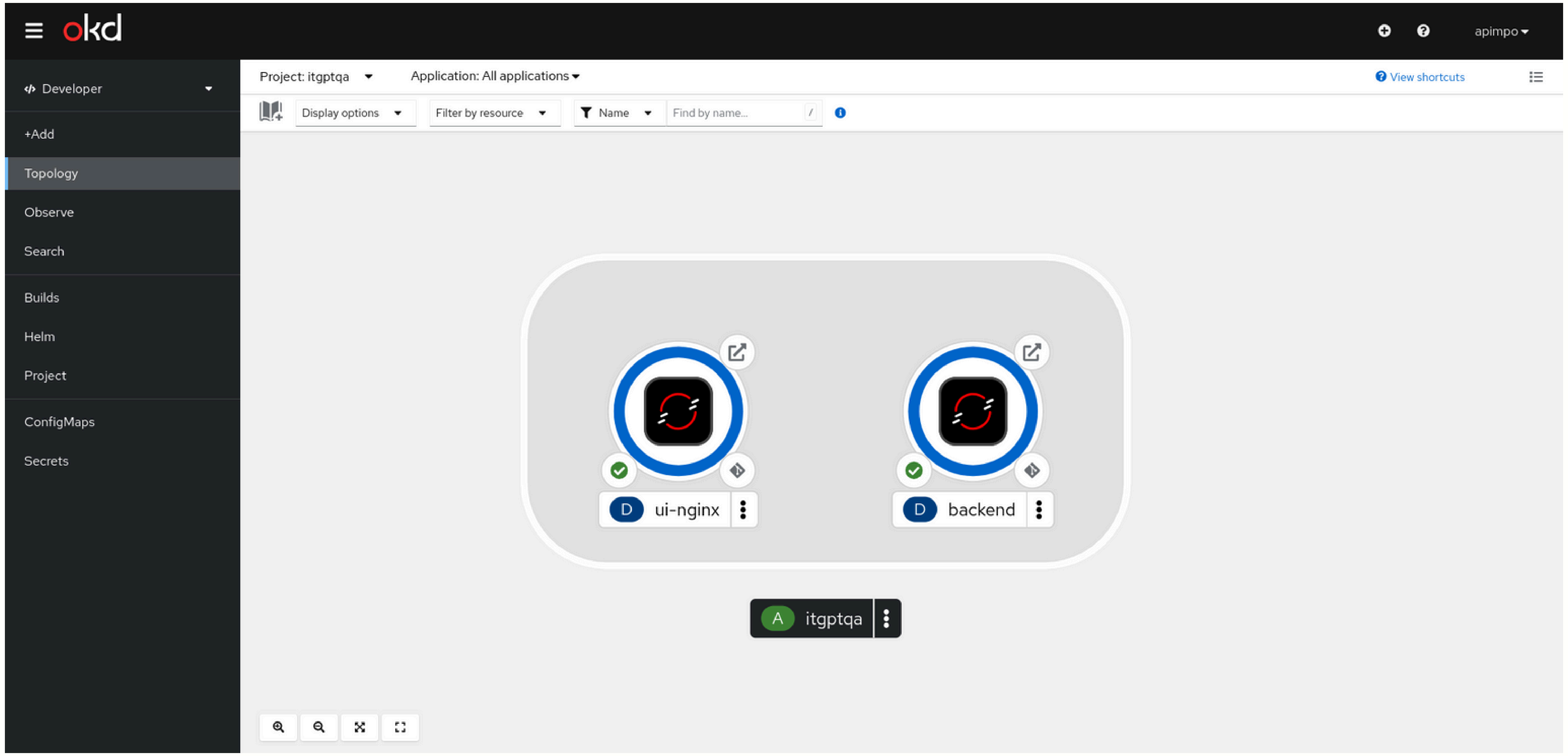
- Deploy openwebui helm chart with a custom values file
  - disable ollama and pipelines components
  - set the hostname to *myname-test.app.cern.ch*
    - *add [1] to the ingress config to make it visible outside CERN*
  - increase the replicas to 3
  - use image gitlab-registry.cern.ch/apimpo/open-webui/myimage with tag latest-x86_64
- Extra:
  - connect it with a DBoD Postgres instance (extra envvar DATABASE_URL)

[1]

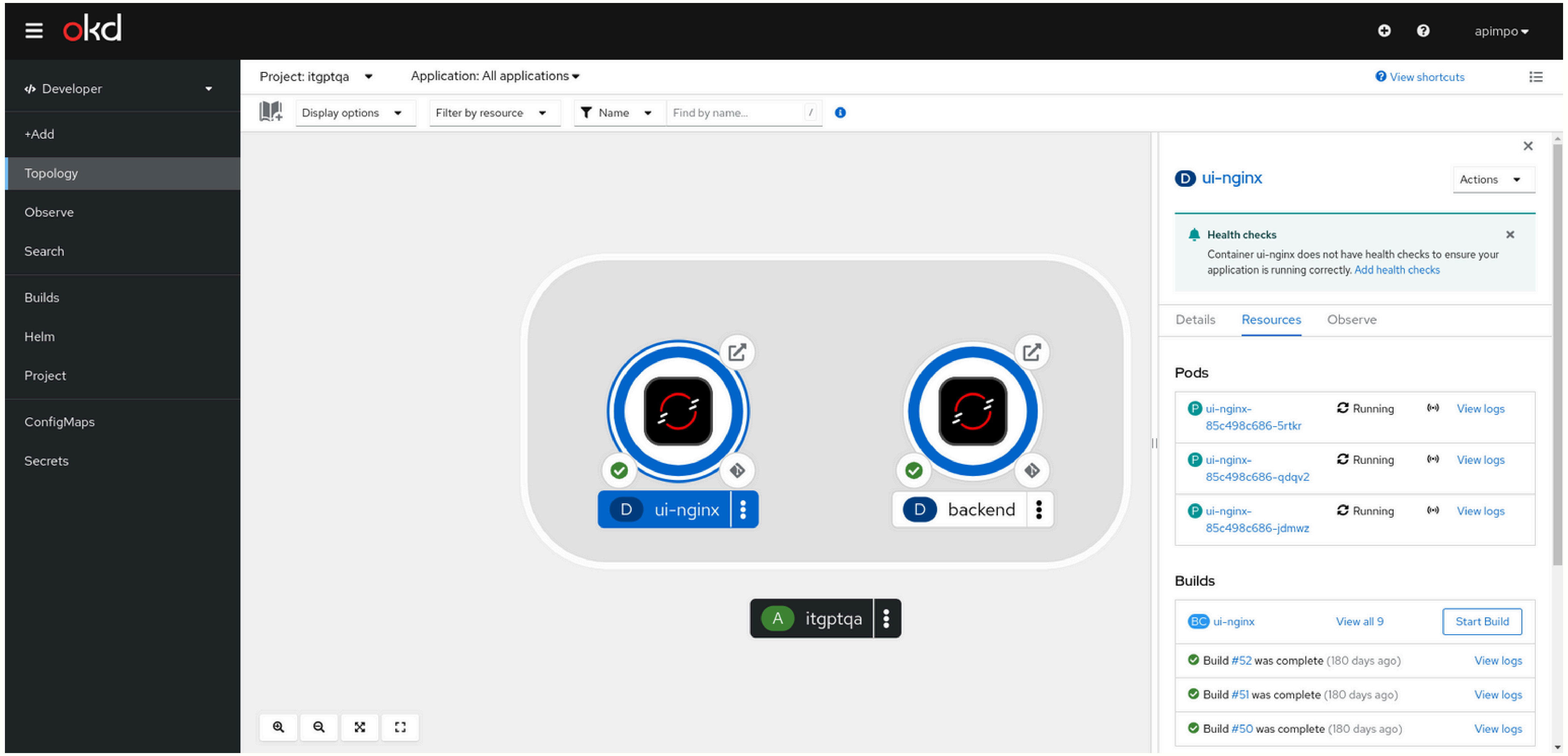


```yaml
annotations:
  haproxy.router.openshift.io/ip_whitelist: ''
  route.openshift.io/termination: edge
tls: false
```

Screenshots



Screenshots



Health checks

Container ui-nginx does not have health checks to ensure your application is running correctly. Add health checks

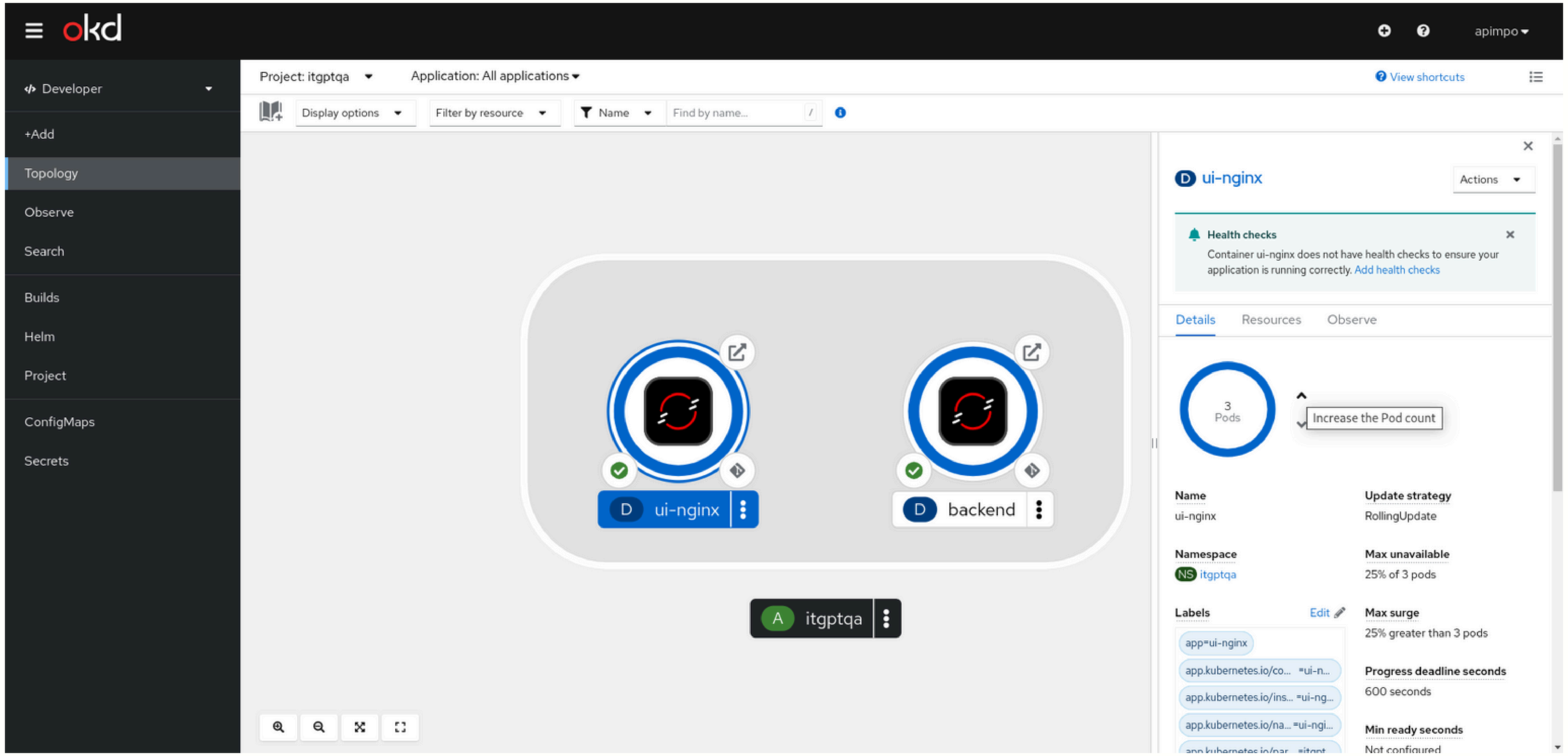
Pods

```
P ui-nginx-85c498c686-5rtkr    Running   View logs
P ui-nginx-85c498c686-qdqv2    Running   View logs
P ui-nginx-85c498c686-jdmwz    Running   View logs
```

Builds

```
BC ui-nginx          View all 9          Start Build
Build #52 was complete (180 days ago)    View logs
Build #51 was complete (180 days ago)    View logs
Build #50 was complete (180 days ago)    View logs
```

Screenshots



Health checks

Container ui-nginx does not have health checks to ensure your application is running correctly. Add health checks

3 Pods

Name: ui-nginx

Namespace: NS itgptqa

Labels:
- app=ui-nginx
- app.kubernetes.io/co... =ui-n...
- app.kubernetes.io/ins... =ui-ng...
- app.kubernetes.io/na... =ui-ngi...

Update strategy: RollingUpdate

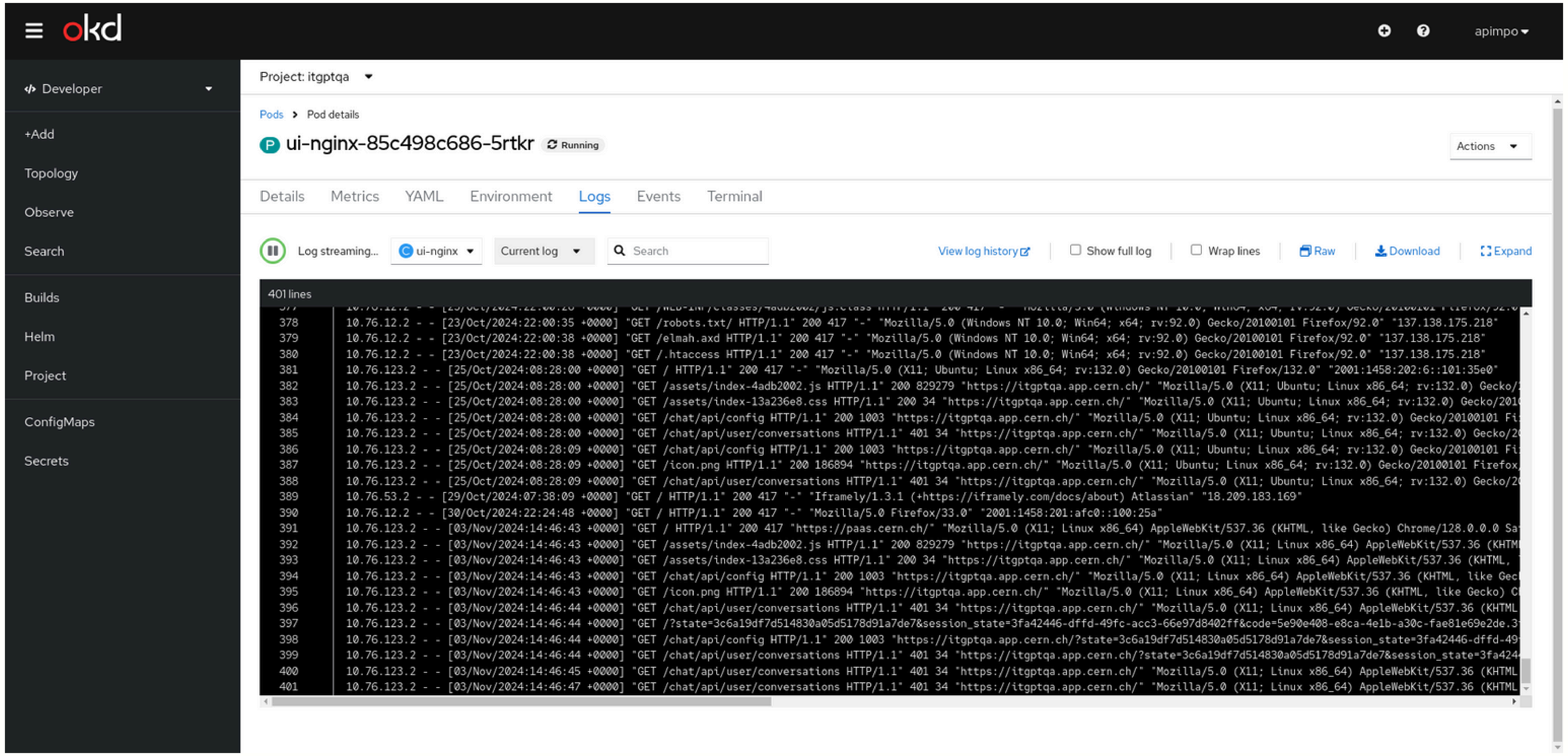
Max unavailable: 25% of 3 pods

Max surge: 25% greater than 3 pods

Progress deadline seconds: 600 seconds

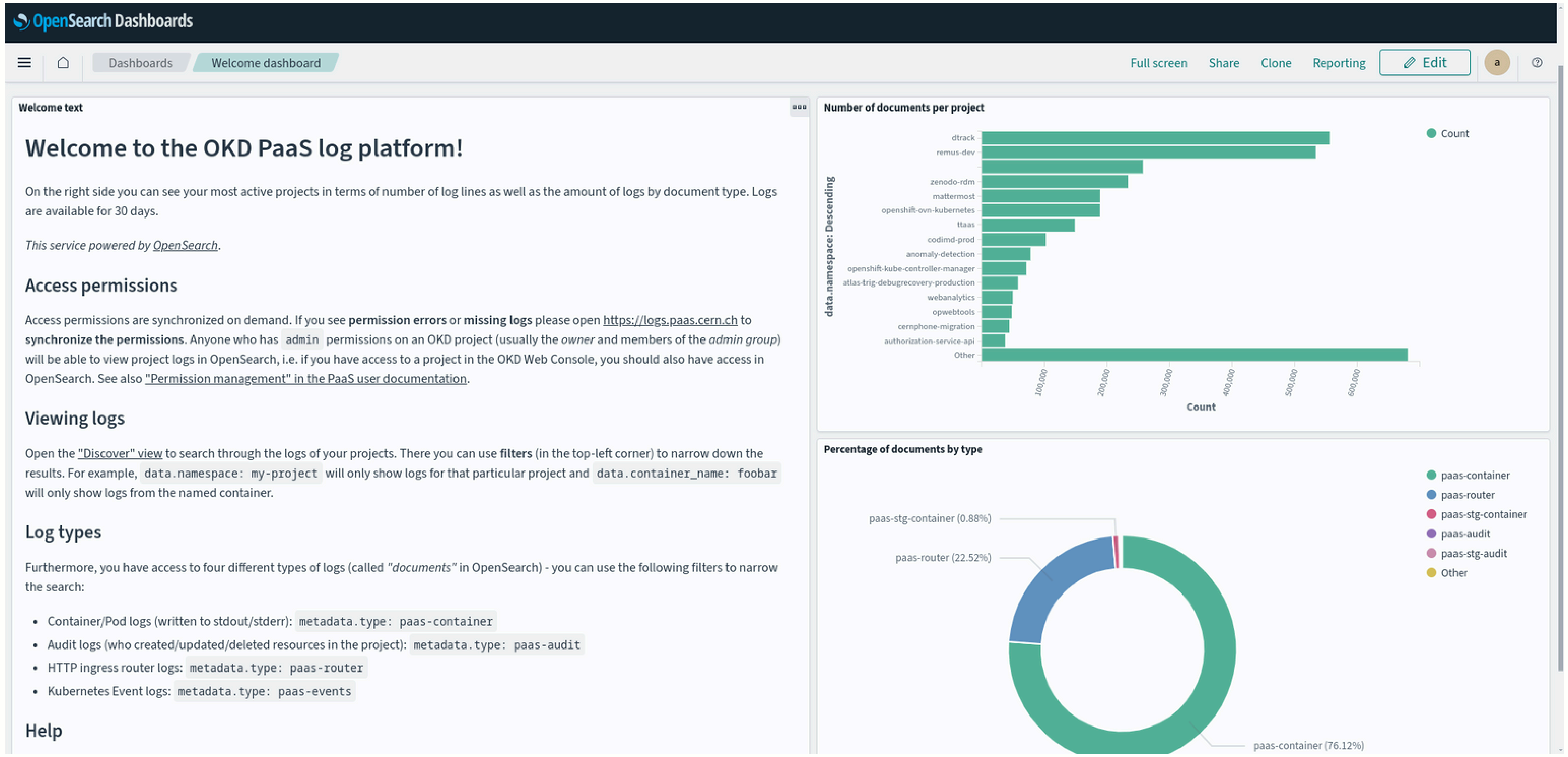
Min ready seconds: Not configured

Screenshots



```
10.76.12.2 - - [23/Oct/2024:22:00:35 +0000] "GET /robots.txt/ HTTP/1.1" 200 417 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:92.0) Gecko/20100101 Firefox/92.0" "137.138.175.218"
10.76.12.2 - - [23/Oct/2024:22:00:38 +0000] "GET /elmah.axd HTTP/1.1" 200 417 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:92.0) Gecko/20100101 Firefox/92.0" "137.138.175.218"
10.76.12.2 - - [23/Oct/2024:22:00:38 +0000] "GET /.htaccess HTTP/1.1" 200 417 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:92.0) Gecko/20100101 Firefox/92.0" "137.138.175.218"
10.76.123.2 - - [25/Oct/2024:08:28:00 +0000] "GET / HTTP/1.1" 200 417 "-" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/20100101 Firefox/132.0" "2001:1458:202:6::101:35e0"
10.76.123.2 - - [25/Oct/2024:08:28:00 +0000] "GET /assets/index-4adb2002.js HTTP/1.1" 200 829279 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/2
10.76.123.2 - - [25/Oct/2024:08:28:00 +0000] "GET /assets/index-13a236e8.css HTTP/1.1" 200 34 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/201
10.76.123.2 - - [25/Oct/2024:08:28:00 +0000] "GET /chat/api/config HTTP/1.1" 200 1003 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/20100101 Fi
10.76.123.2 - - [25/Oct/2024:08:28:00 +0000] "GET /chat/api/user/conversations HTTP/1.1" 401 34 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/2
10.76.123.2 - - [25/Oct/2024:08:28:09 +0000] "GET /chat/api/config HTTP/1.1" 200 1003 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/20100101 Fi
10.76.123.2 - - [25/Oct/2024:08:28:09 +0000] "GET /icon.png HTTP/1.1" 200 186894 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/20100101 Firefox
10.76.123.2 - - [25/Oct/2024:08:28:09 +0000] "GET /chat/api/user/conversations HTTP/1.1" 401 34 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:132.0) Gecko/2
10.76.53.2 - - [29/Oct/2024:07:38:09 +0000] "GET / HTTP/1.1" 200 417 "-" "Iframely/1.3.1 (+https://iframely.com/docs/about) Atlassian" "18.209.183.169"
10.76.12.2 - - [30/Oct/2024:22:24:48 +0000] "GET / HTTP/1.1" 200 417 "-" "Mozilla/5.0 Firefox/33.0" "2001:1458:201:afc0::100:25a"
10.76.123.2 - - [03/Nov/2024:14:46:43 +0000] "GET / HTTP/1.1" 200 417 "https://paas.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/128.0.0.0 Sa
10.76.123.2 - - [03/Nov/2024:14:46:43 +0000] "GET /assets/index-4adb2002.js HTTP/1.1" 200 829279 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTM
10.76.123.2 - - [03/Nov/2024:14:46:43 +0000] "GET /assets/index-13a236e8.css HTTP/1.1" 200 34 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML,
10.76.123.2 - - [03/Nov/2024:14:46:43 +0000] "GET /chat/api/config HTTP/1.1" 200 1003 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gec
10.76.123.2 - - [03/Nov/2024:14:46:43 +0000] "GET /icon.png HTTP/1.1" 200 186894 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) C
10.76.123.2 - - [03/Nov/2024:14:46:44 +0000] "GET /chat/api/user/conversations HTTP/1.1" 401 34 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML
10.76.123.2 - - [03/Nov/2024:14:46:44 +0000] "GET /?state=3c6a19df7d514830a05d5178d91a7de7&session_state=3fa42446-dffd-49fc-acc3-66e97d8402ff&code=5e90e408-e8ca-4e1b-a30c-fae81e69e2de.3
10.76.123.2 - - [03/Nov/2024:14:46:44 +0000] "GET /chat/api/config HTTP/1.1" 200 1003 "https://itgptqa.app.cern.ch/?state=3c6a19df7d514830a05d5178d91a7de7&session_state=3fa42446-dffd-49
10.76.123.2 - - [03/Nov/2024:14:46:44 +0000] "GET /chat/api/user/conversations HTTP/1.1" 401 34 "https://itgptqa.app.cern.ch/?state=3c6a19df7d514830a05d5178d91a7de7&session_state=3fa424
10.76.123.2 - - [03/Nov/2024:14:46:45 +0000] "GET /chat/api/user/conversations HTTP/1.1" 401 34 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML
10.76.123.2 - - [03/Nov/2024:14:46:47 +0000] "GET /chat/api/user/conversations HTTP/1.1" 401 34 "https://itgptqa.app.cern.ch/" "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML
```

Screenshots



Welcome to the OKD PaaS log platform!

On the right side you can see your most active projects in terms of number of log lines as well as the amount of logs by document type. Logs are available for 30 days.

This service powered by OpenSearch.

Access permissions

Access permissions are synchronized on demand. If you see **permission errors** or **missing logs** please open https://logs.paas.cern.ch to **synchronize the permissions**. Anyone who has `admin` permissions on an OKD project (usually the *owner* and members of the *admin group*) will be able to view project logs in OpenSearch, i.e. if you have access to a project in the OKD Web Console, you should also have access in OpenSearch. See also "Permission management" in the PaaS user documentation.

Viewing logs

Open the "Discover" view to search through the logs of your projects. There you can use **filters** (in the top-left corner) to narrow down the results. For example, `data.namespace: my-project` will only show logs for that particular project and `data.container_name: foobar` will only show logs from the named container.

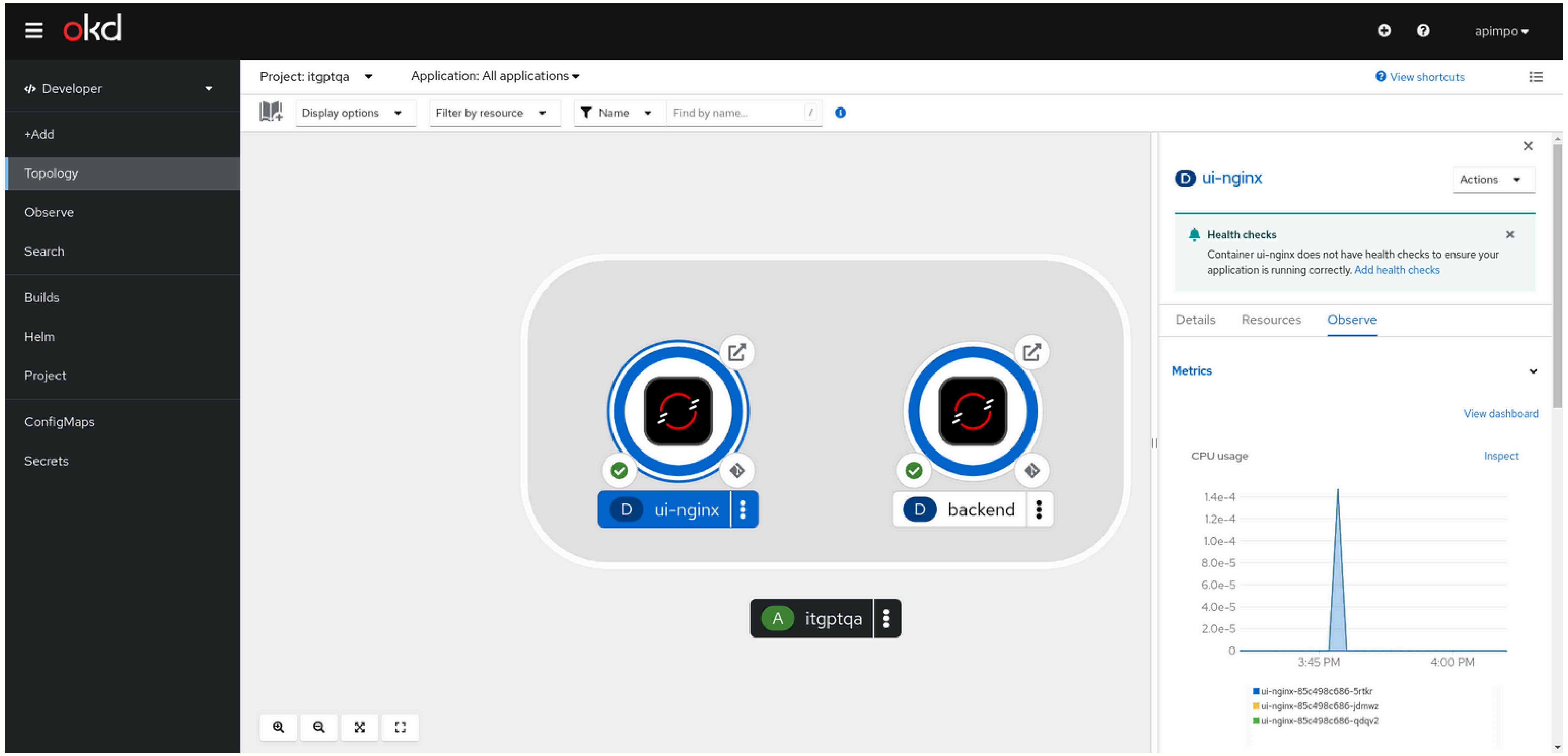
Log types

Furthermore, you have access to four different types of logs (called *"documents"* in OpenSearch) - you can use the following filters to narrow the search:

- Container/Pod logs (written to stdout/stderr): `metadata.type: paas-container`
- Audit logs (who created/updated/deleted resources in the project): `metadata.type: paas-audit`
- HTTP ingress router logs: `metadata.type: paas-router`
- Kubernetes Event logs: `metadata.type: paas-events`

Help

Screenshots



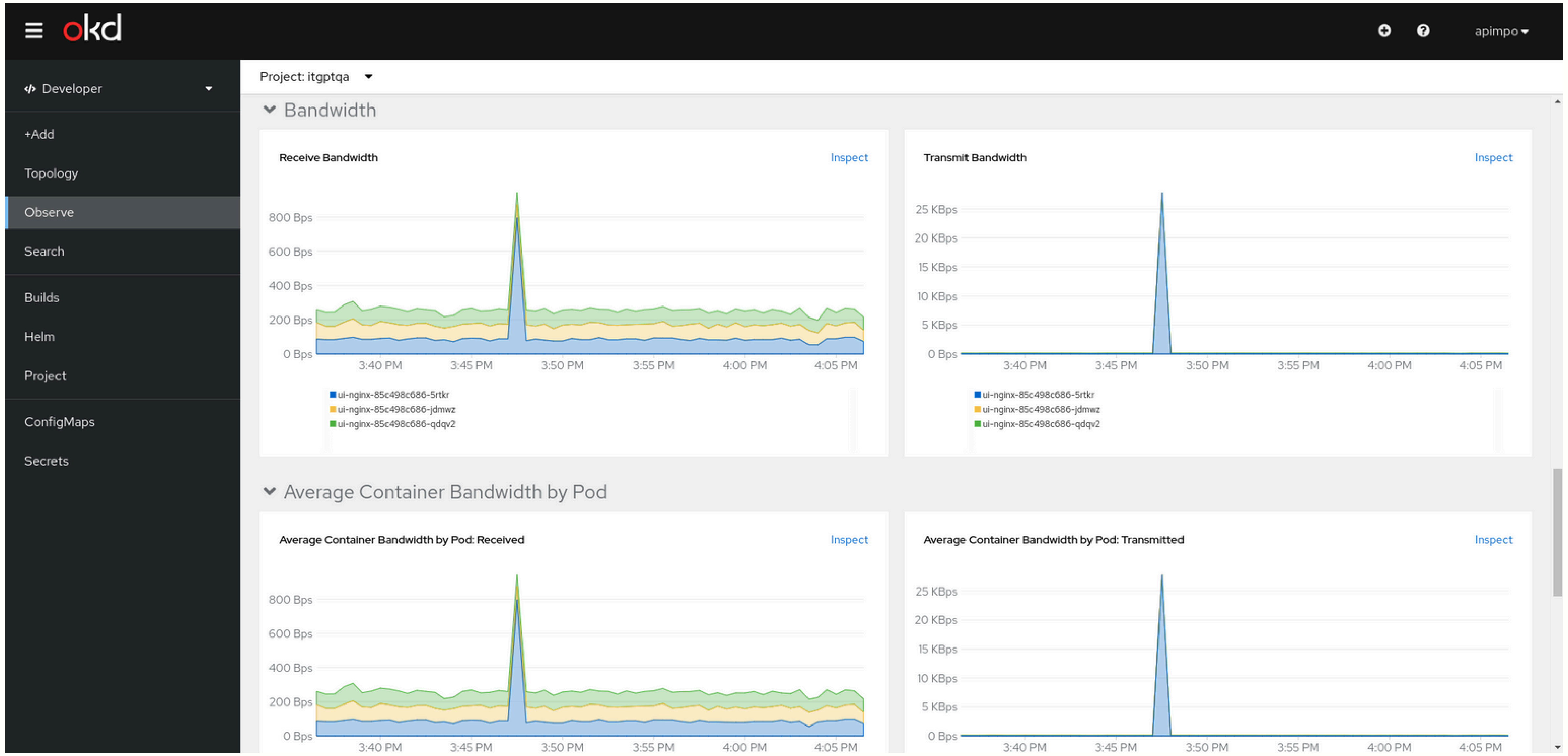
Project: itgptqa

Application: All applications

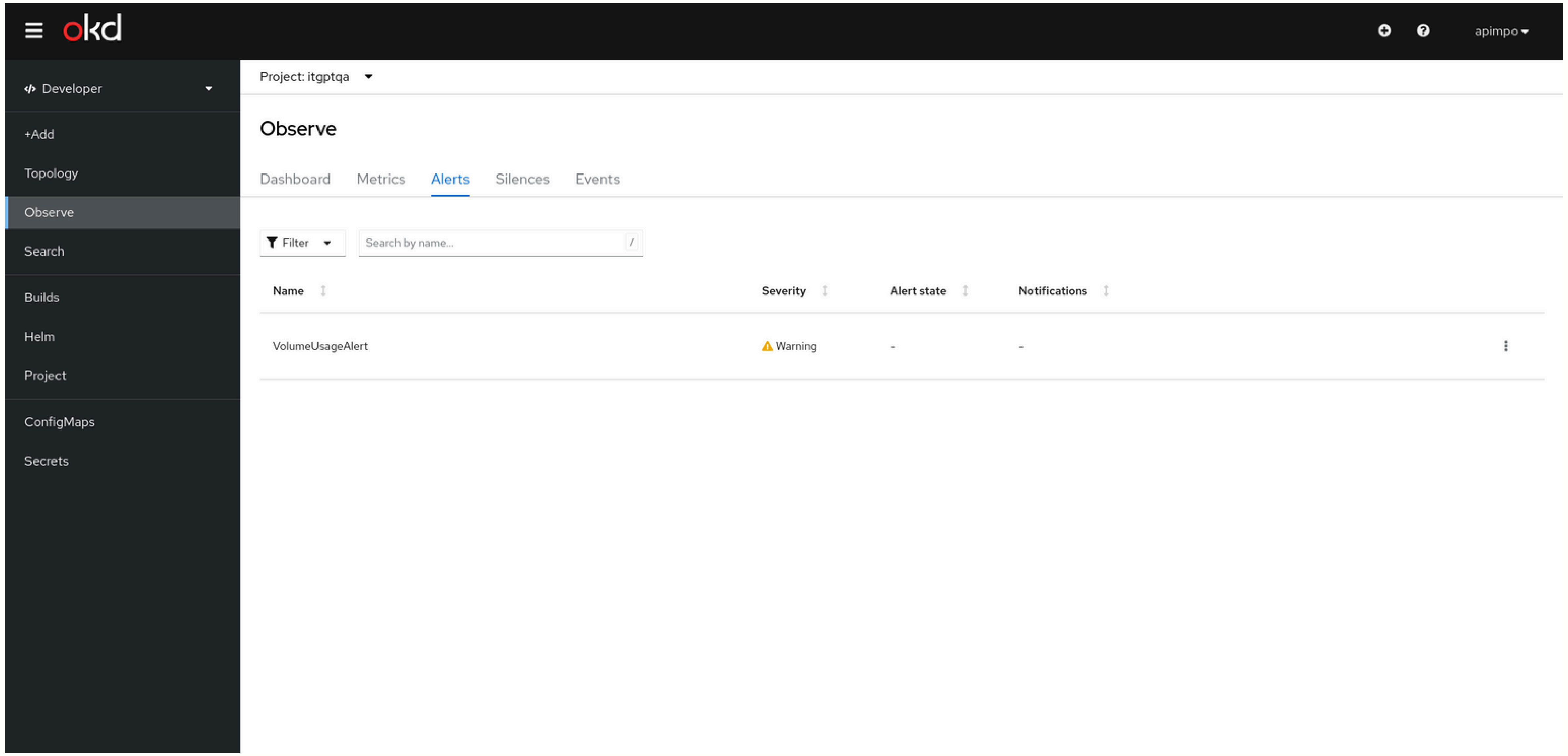
Health checks

Container ui-nginx does not have health checks to ensure your application is running correctly. Add health checks

Screenshots



Screenshots

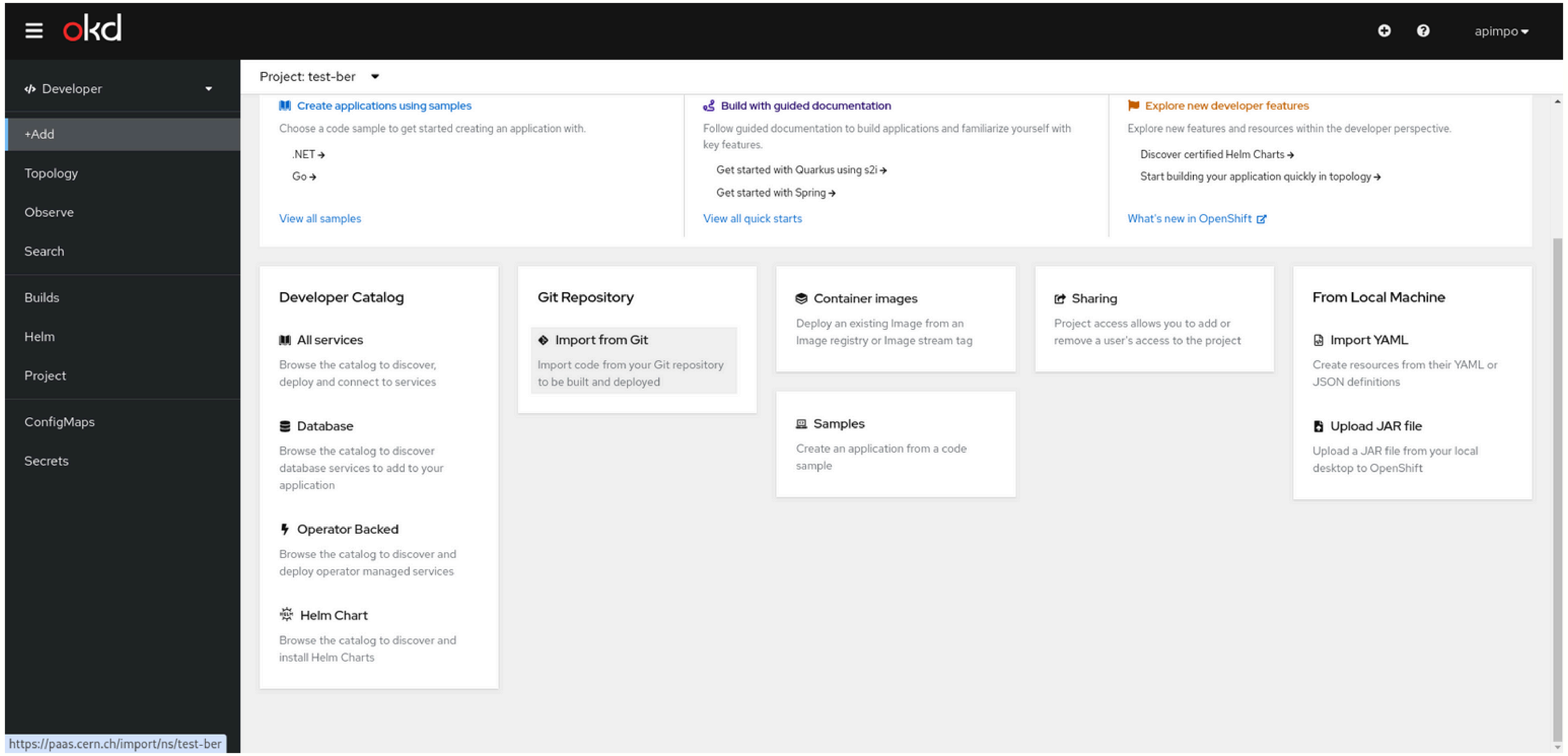


Project: itgptqa

Observe

Name	Severity	Alert state	Notifications
VolumeUsageAlert	⚠ Warning	-	-

Screenshots



Project: test-ber

Create applications using samples

Choose a code sample to get started creating an application with.

.NET →

Go →

View all samples

Build with guided documentation

Follow guided documentation to build applications and familiarize yourself with key features.

Get started with Quarkus using s2i →

Get started with Spring →

View all quick starts

Explore new developer features

Explore new features and resources within the developer perspective.

Discover certified Helm Charts →

Start building your application quickly in topology →

What's new in OpenShift

Developer Catalog

All services

Browse the catalog to discover, deploy and connect to services

Database

Browse the catalog to discover database services to add to your application

Operator Backed

Browse the catalog to discover and deploy operator managed services

Helm Chart

Browse the catalog to discover and install Helm Charts

Git Repository

Import from Git

Import code from your Git repository to be built and deployed

Container images

Deploy an existing Image from an Image registry or Image stream tag

Samples

Create an application from a code sample

Sharing

Project access allows you to add or remove a user's access to the project

From Local Machine

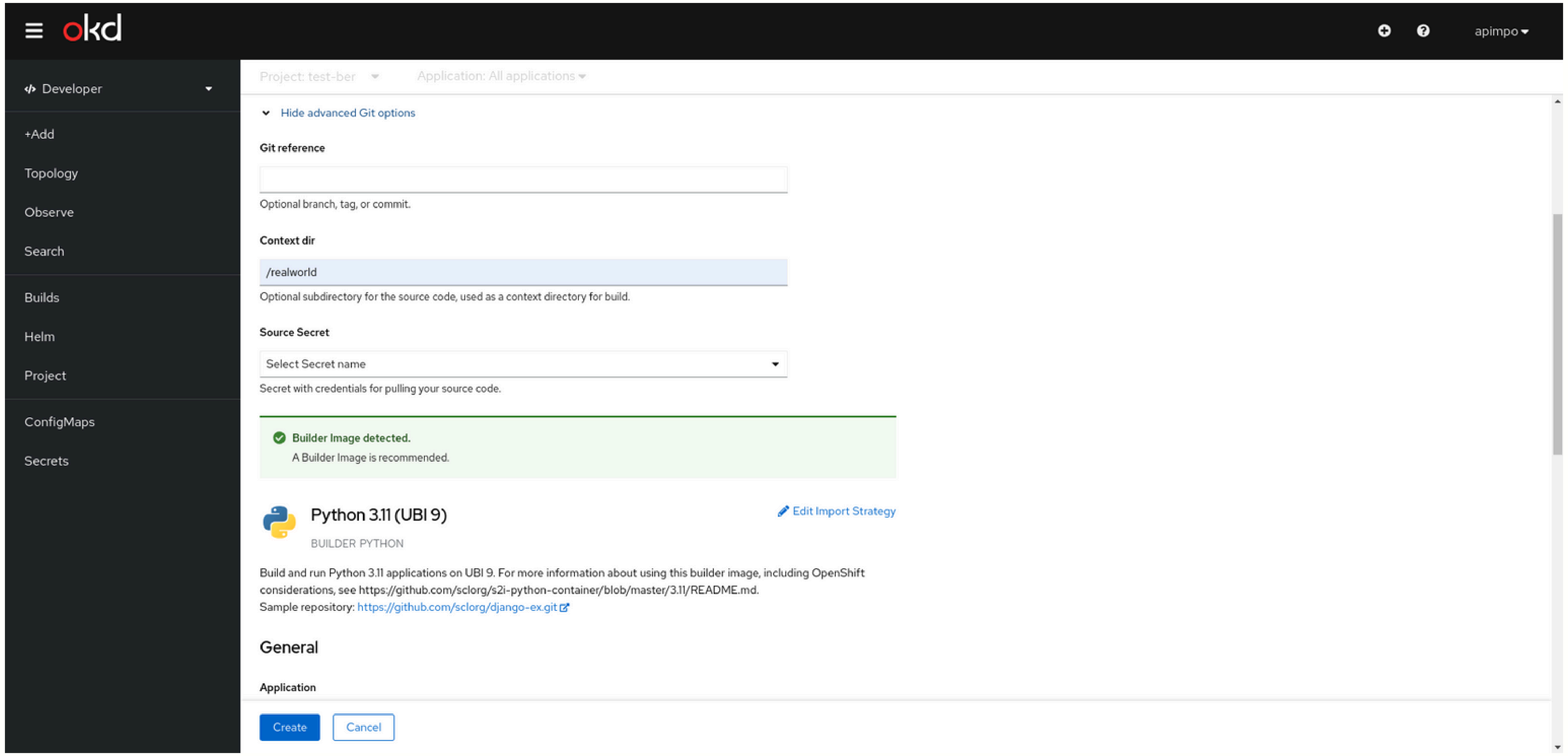
Import YAML

Create resources from their YAML or JSON definitions

Upload JAR file

Upload a JAR file from your local desktop to OpenShift

Screenshots



Git reference

Optional branch, tag, or commit.

Context dir

/realworld

Optional subdirectory for the source code, used as a context directory for build.

Source Secret

Select Secret name

Secret with credentials for pulling your source code.

Builder Image detected.

A Builder Image is recommended.

Python 3.11 (UBI 9)

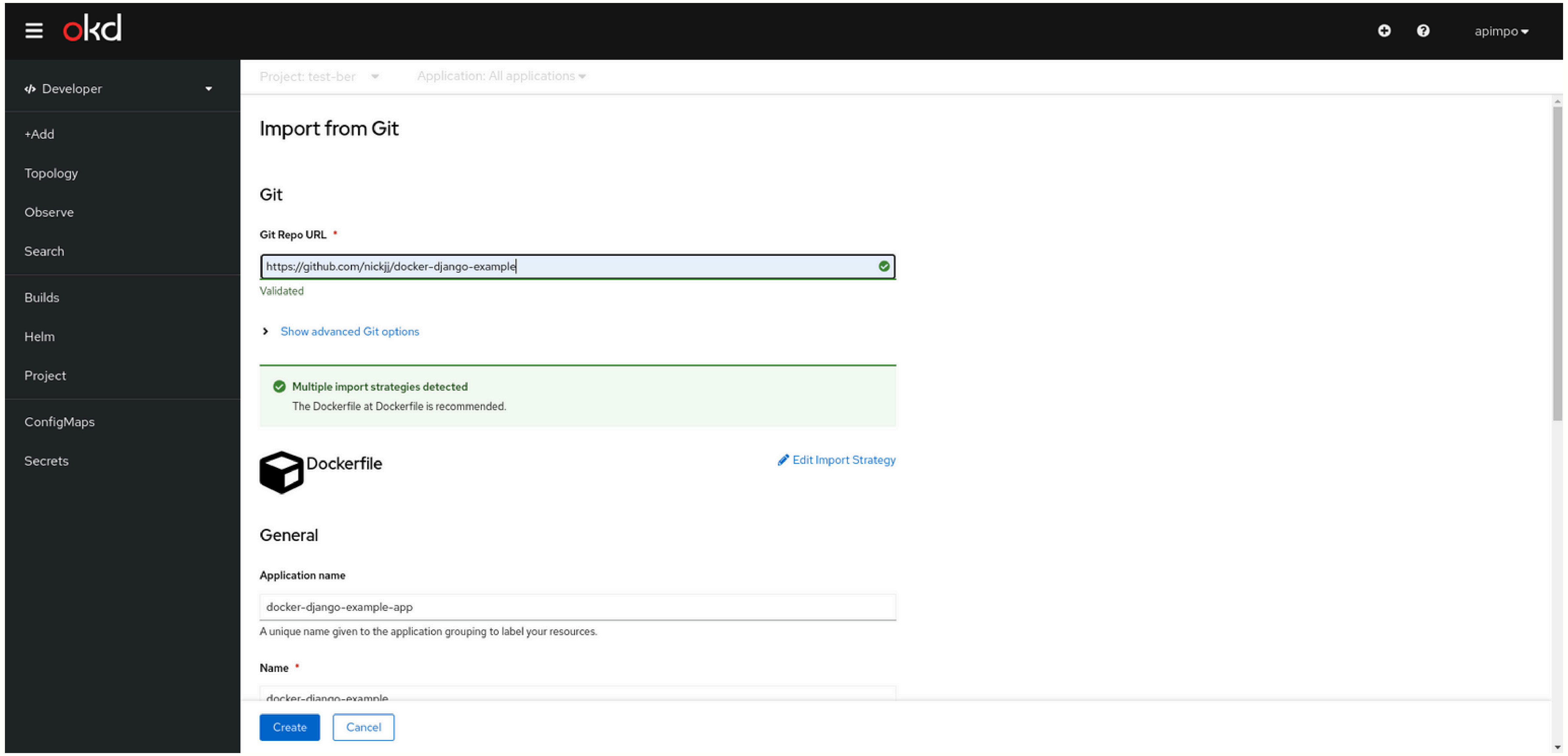
BUILDER PYTHON

Build and run Python 3.11 applications on UBI 9. For more information about using this builder image, including OpenShift considerations, see https://github.com/sclorg/s2i-python-container/blob/master/3.11/README.md.
Sample repository: https://github.com/sclorg/django-ex.git

General

Application

Screenshots



Import from Git

Git

Git Repo URL *

https://github.com/nickjj/docker-django-example

Validated

Multiple import strategies detected

The Dockerfile at Dockerfile is recommended.

Dockerfile

General

Application name

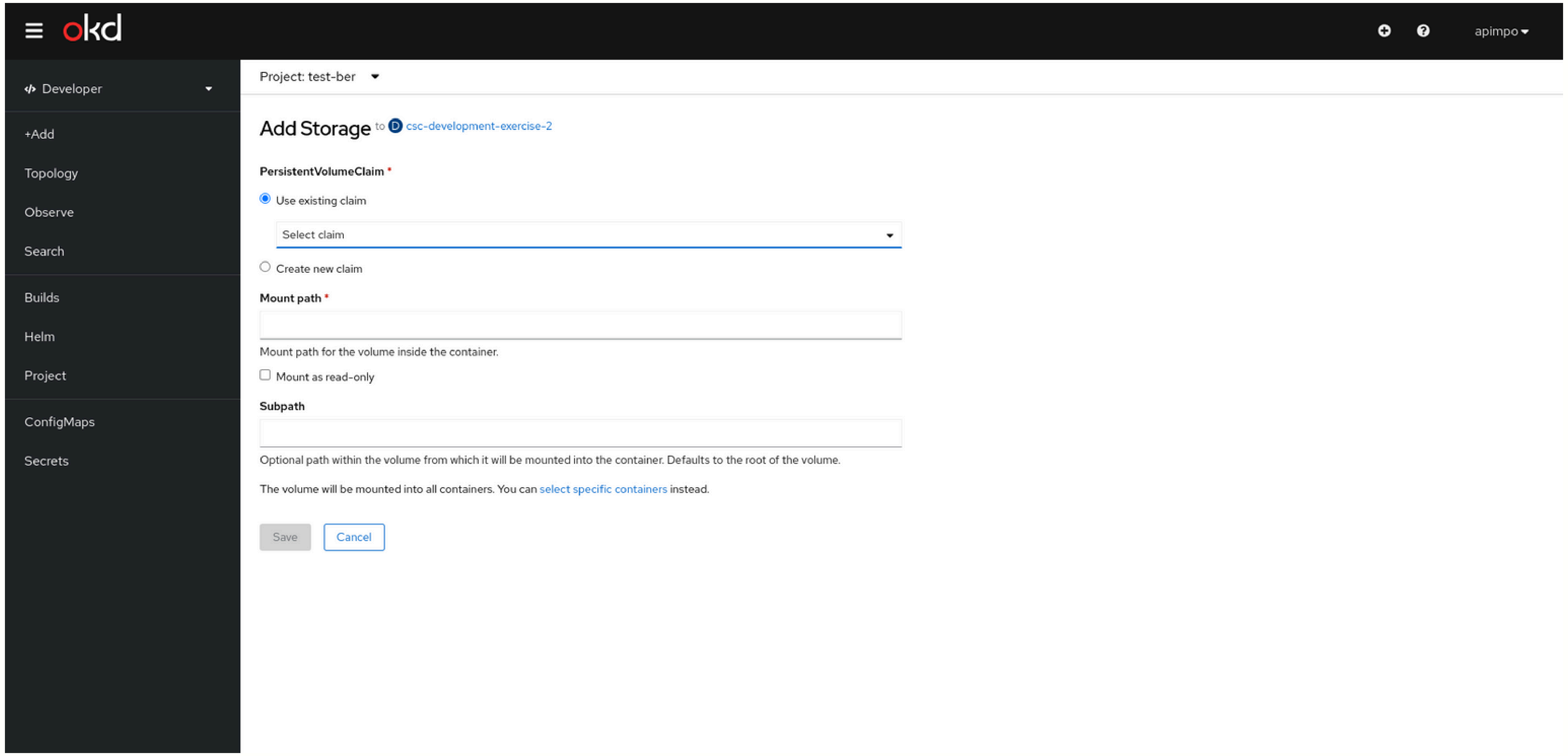
docker-django-example-app

A unique name given to the application grouping to label your resources.

Name *

docker-django-example

Screenshots



Project: test-ber

Add Storage to D csc-development-exercise-2

PersistentVolumeClaim *

● Use existing claim

Select claim

○ Create new claim

Mount path *

Mount path for the volume inside the container.

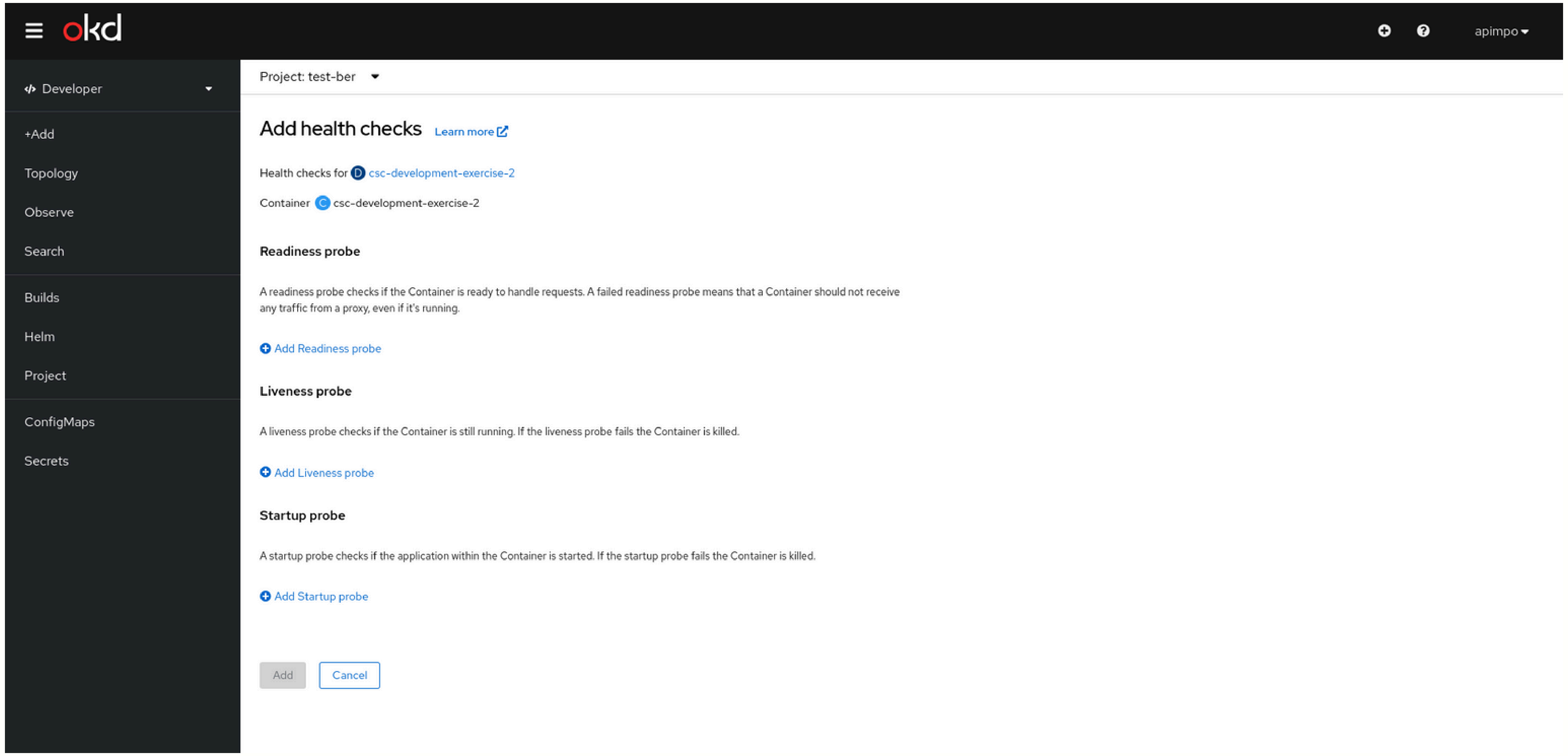
□ Mount as read-only

Subpath

Optional path within the volume from which it will be mounted into the container. Defaults to the root of the volume.

The volume will be mounted into all containers. You can select specific containers instead.

Screenshots



Add health checks Learn more

Health checks for D csc-development-exercise-2

Container C csc-development-exercise-2

Readiness probe

A readiness probe checks if the Container is ready to handle requests. A failed readiness probe means that a Container should not receive any traffic from a proxy, even if it's running.

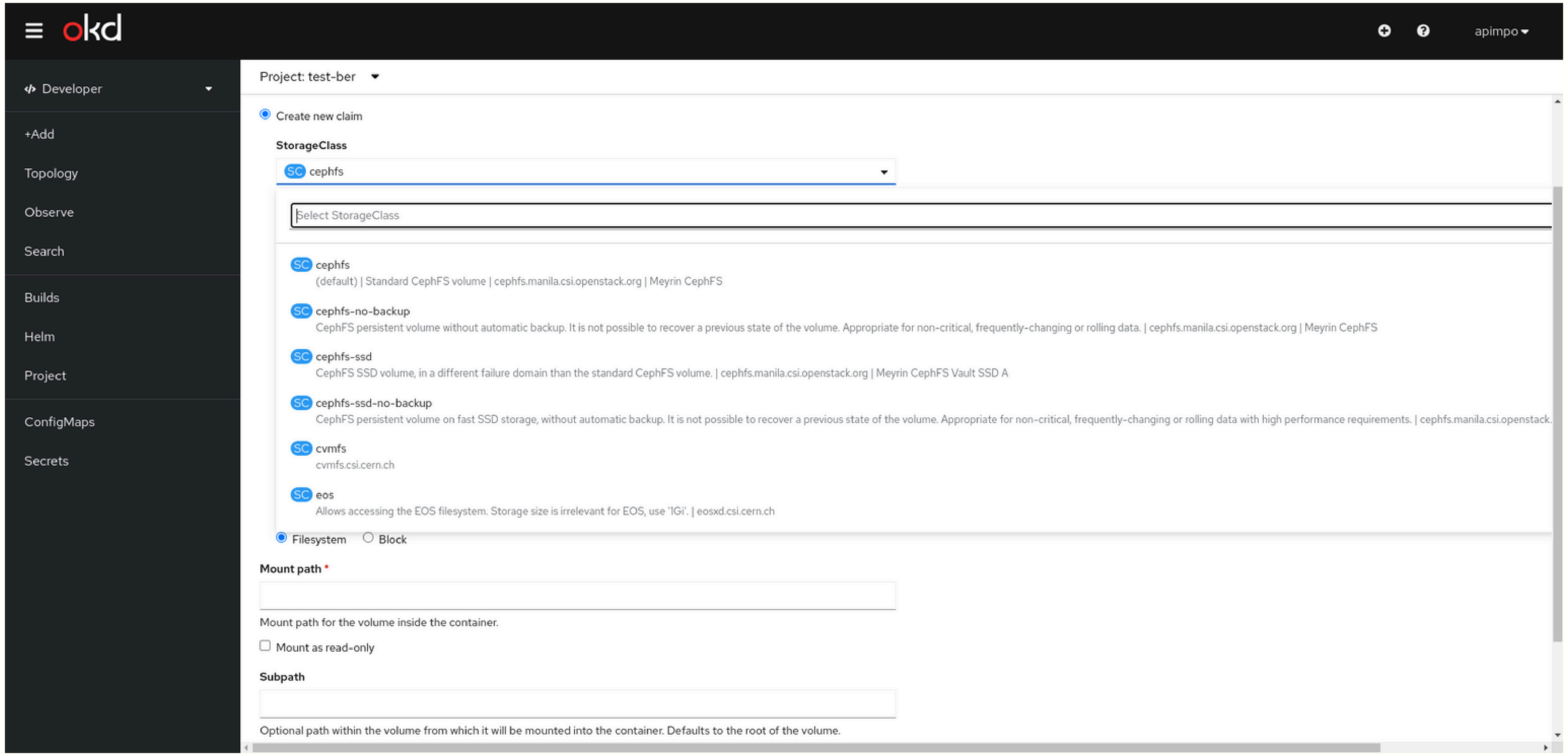
Liveness probe

A liveness probe checks if the Container is still running. If the liveness probe fails the Container is killed.

Startup probe

A startup probe checks if the application within the Container is started. If the startup probe fails the Container is killed.

Screenshots



Project: test-ber

● Create new claim

StorageClass

SC cephfs

Select StorageClass

SC cephfs
(default) | Standard CephFS volume | cephfs.manila.csi.openstack.org | Meyrin CephFS

SC cephfs-no-backup
CephFS persistent volume without automatic backup. It is not possible to recover a previous state of the volume. Appropriate for non-critical, frequently-changing or rolling data. | cephfs.manila.csi.openstack.org | Meyrin CephFS

SC cephfs-ssd
CephFS SSD volume, in a different failure domain than the standard CephFS volume. | cephfs.manila.csi.openstack.org | Meyrin CephFS Vault SSD A

SC cephfs-ssd-no-backup
CephFS persistent volume on fast SSD storage, without automatic backup. It is not possible to recover a previous state of the volume. Appropriate for non-critical, frequently-changing or rolling data with high performance requirements. | cephfs.manila.csi.openstack.

SC cvmfs
cvmfs.csi.cern.ch

SC eos
Allows accessing the EOS filesystem. Storage size is irrelevant for EOS, use '1Gi'. | eosxd.csi.cern.ch

● Filesystem ○ Block

Mount path *

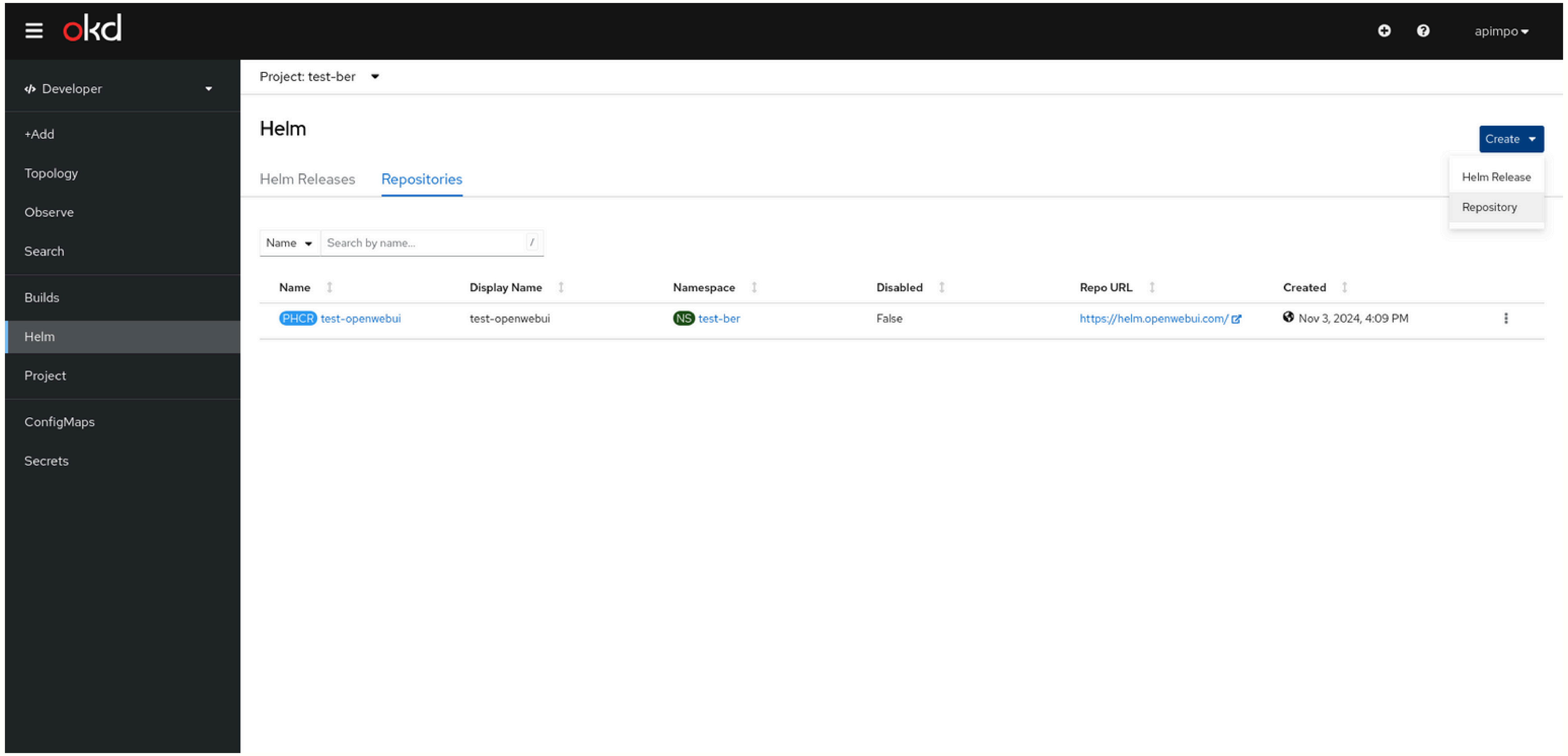
Mount path for the volume inside the container.

□ Mount as read-only

Subpath

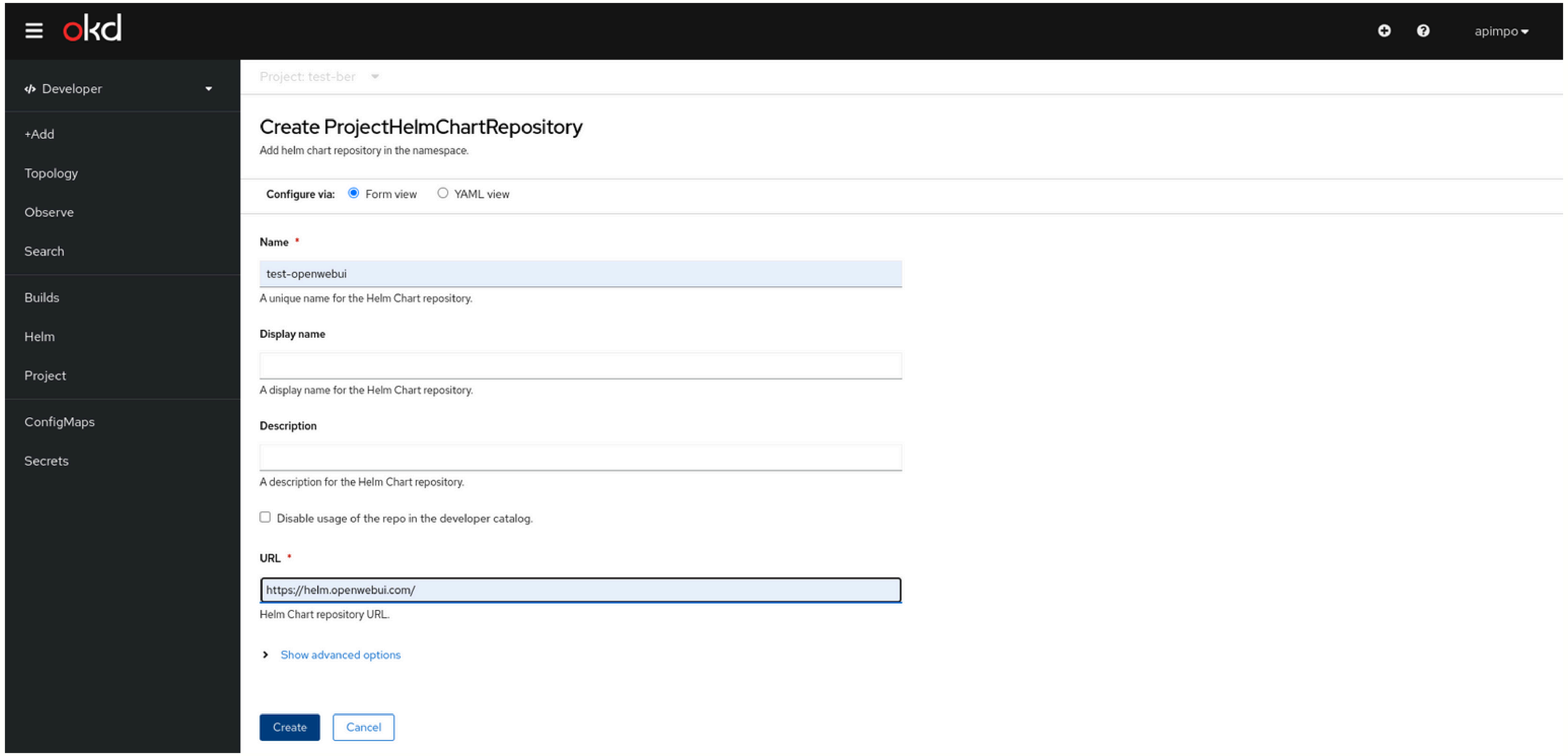
Optional path within the volume from which it will be mounted into the container. Defaults to the root of the volume.

Screenshots



Name	Display Name	Namespace	Disabled	Repo URL	Created
PHCR test-openwebui	test-openwebui	NS test-ber	False	https://helm.openwebui.com/	Nov 3, 2024, 4:09 PM

Screenshots



Create ProjectHelmChartRepository

Add helm chart repository in the namespace.

Configure via: ● Form view ○ YAML view

Name *

test-openwebui

A unique name for the Helm Chart repository.

Display name

A display name for the Helm Chart repository.

Description

A description for the Helm Chart repository.

□ Disable usage of the repo in the developer catalog.

URL *

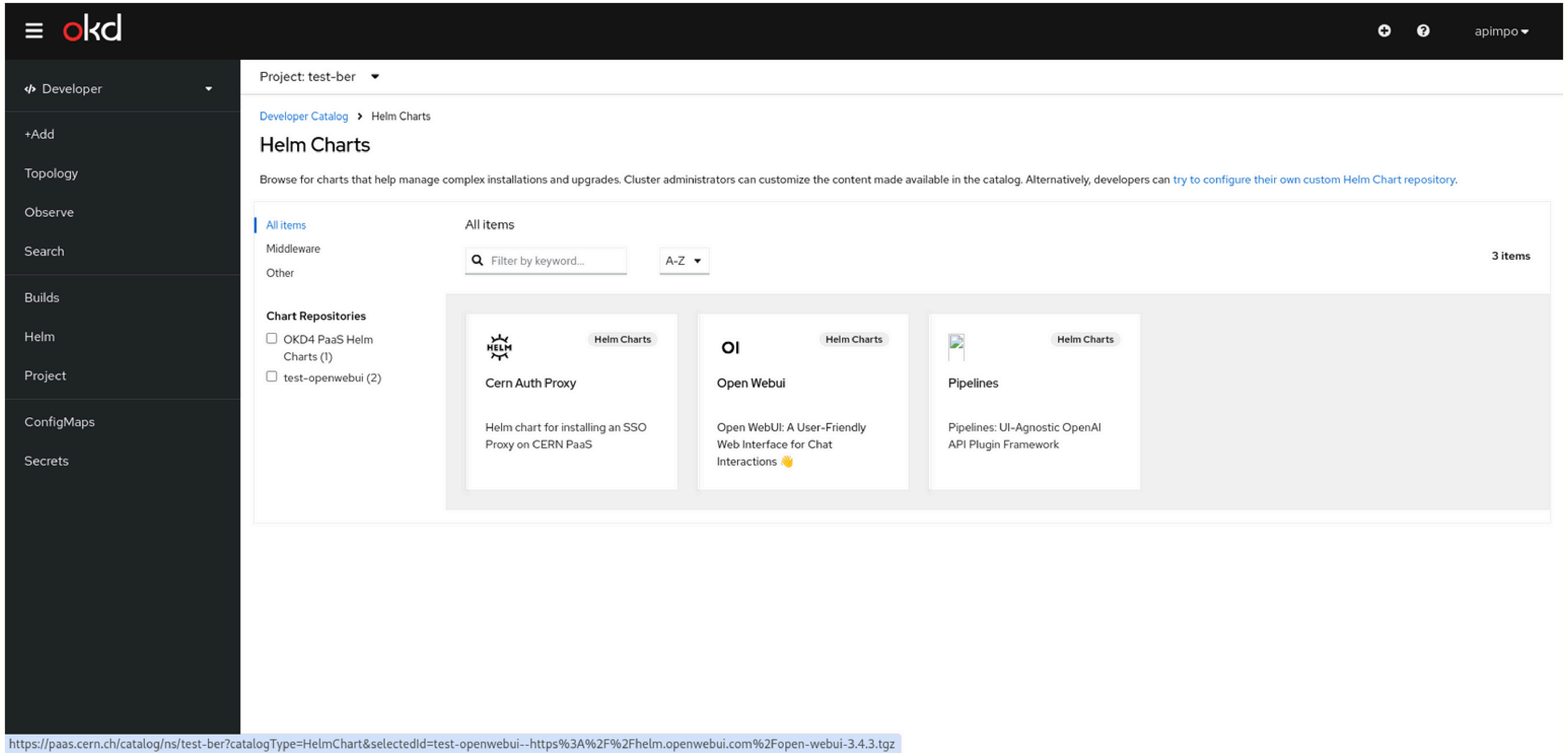
https://helm.openwebui.com/

Helm Chart repository URL.

› Show advanced options

Create | Cancel

Screenshots



Developer Catalog › Helm Charts

Helm Charts

Browse for charts that help manage complex installations and upgrades. Cluster administrators can customize the content made available in the catalog. Alternatively, developers can try to configure their own custom Helm Chart repository.

All items

Chart Repositories

☐ OKD4 PaaS Helm Charts (1)

☐ test-openwebui (2)

Cern Auth Proxy

Helm chart for installing an SSO Proxy on CERN PaaS

Open Webui

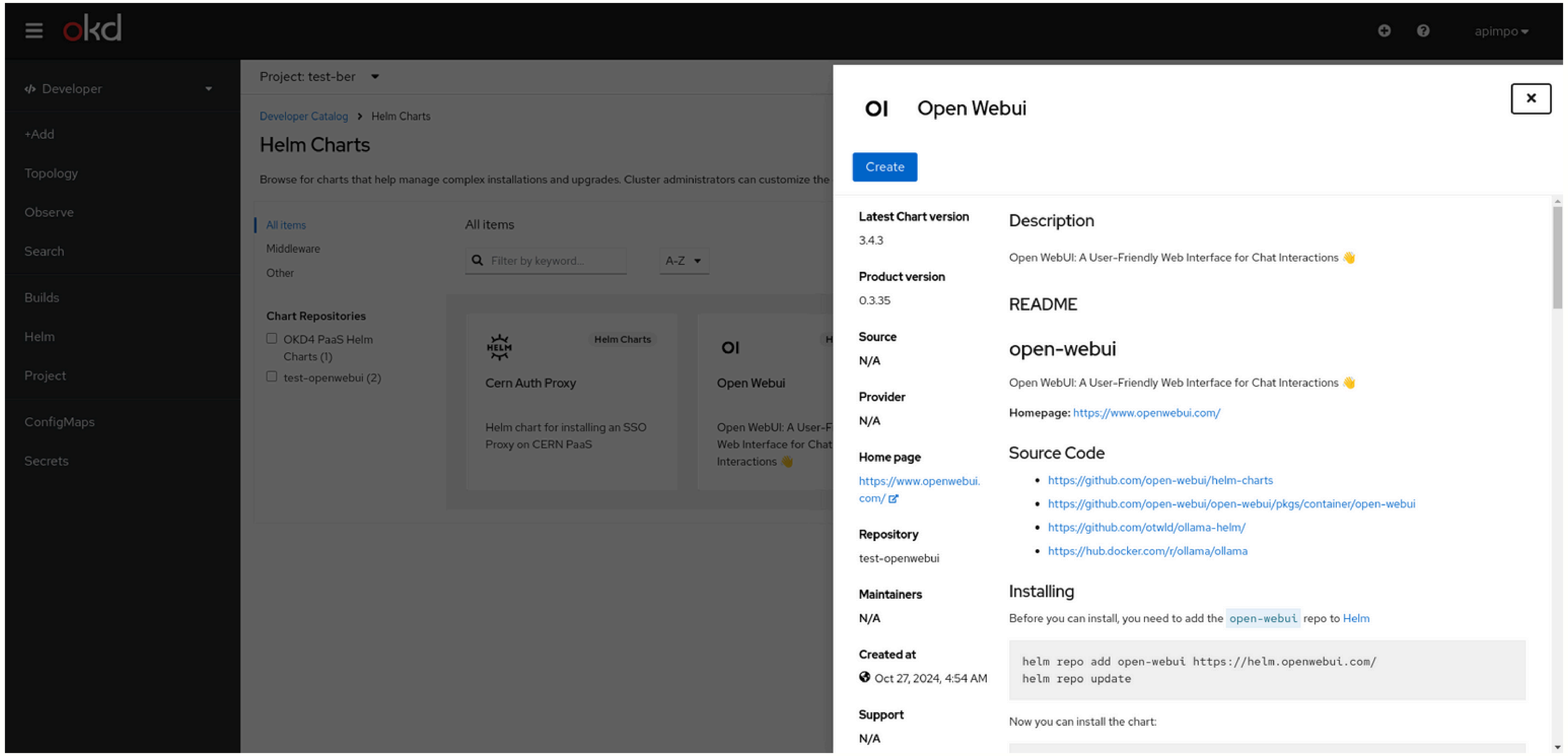
Open WebUI: A User-Friendly Web Interface for Chat Interactions 👋

Pipelines

Pipelines: UI-Agnostic OpenAI API Plugin Framework

https://paas.cern.ch/catalog/ns/test-ber?catalogType=HelmChart&selectedId=test-openwebui--https%3A%2F%2Fhelm.openwebui.com%2Fopen-webui-3.4.3.tgz

Screenshots



Helm Charts

Browse for charts that help manage complex installations and upgrades. Cluster administrators can customize the

Cern Auth Proxy

Helm chart for installing an SSO Proxy on CERN PaaS

Open Webui

Open WebUI: A User-F Web Interface for Chat Interactions 👋

OI Open Webui

Latest Chart version
3.4.3

Product version
0.3.35

Source
N/A

Provider
N/A

Home page
https://www.openwebui.com/

Repository
test-openwebui

Maintainers
N/A

Created at
Oct 27, 2024, 4:54 AM

Support
N/A

Description

Open WebUI: A User-Friendly Web Interface for Chat Interactions 👋

README

open-webui

Open WebUI: A User-Friendly Web Interface for Chat Interactions 👋

Homepage: https://www.openwebui.com/

Source Code

- https://github.com/open-webui/helm-charts
- https://github.com/open-webui/open-webui/pkgs/container/open-webui
- https://github.com/otwld/ollama-helm/
- https://hub.docker.com/r/ollama/ollama

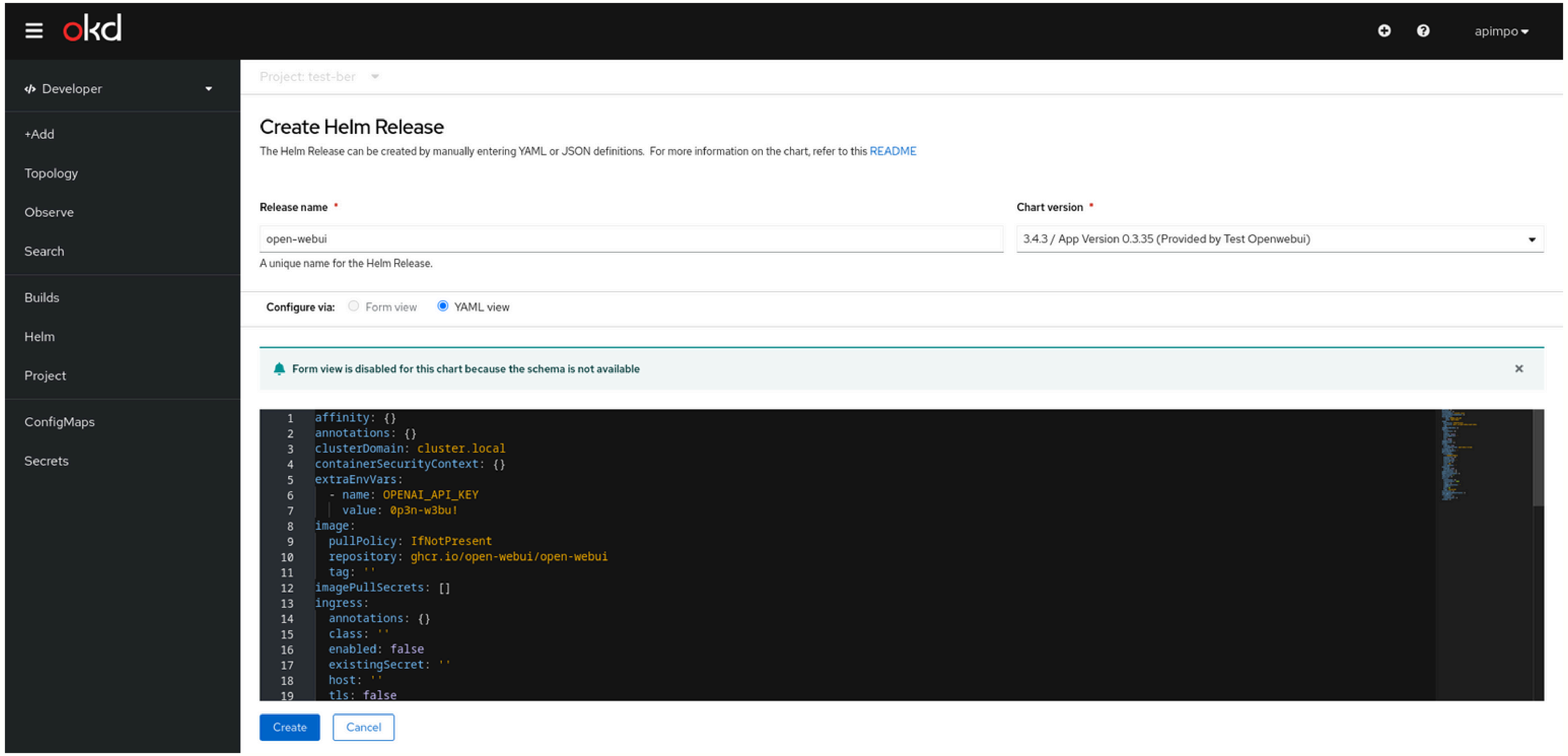
Installing

Before you can install, you need to add the `open-webui` repo to Helm

```
helm repo add open-webui https://helm.openwebui.com/
helm repo update
```

Now you can install the chart:

Screenshots



Create Helm Release

The Helm Release can be created by manually entering YAML or JSON definitions. For more information on the chart, refer to this README

Release name *

open-webui

A unique name for the Helm Release.

Chart version *

3.4.3 / App Version 0.3.35 (Provided by Test Openwebui)

Configure via: ○ Form view ● YAML view

Form view is disabled for this chart because the schema is not available

```yaml
affinity: {}
annotations: {}
clusterDomain: cluster.local
containerSecurityContext: {}
extraEnvVars:
  - name: OPENAI_API_KEY
    value: 0p3n-w3bu!
image:
  pullPolicy: IfNotPresent
  repository: ghcr.io/open-webui/open-webui
  tag: ''
imagePullSecrets: []
ingress:
  annotations: {}
  class: ''
  enabled: false
  existingSecret: ''
  host: ''
  tls: false
```