

Apache Bigdata



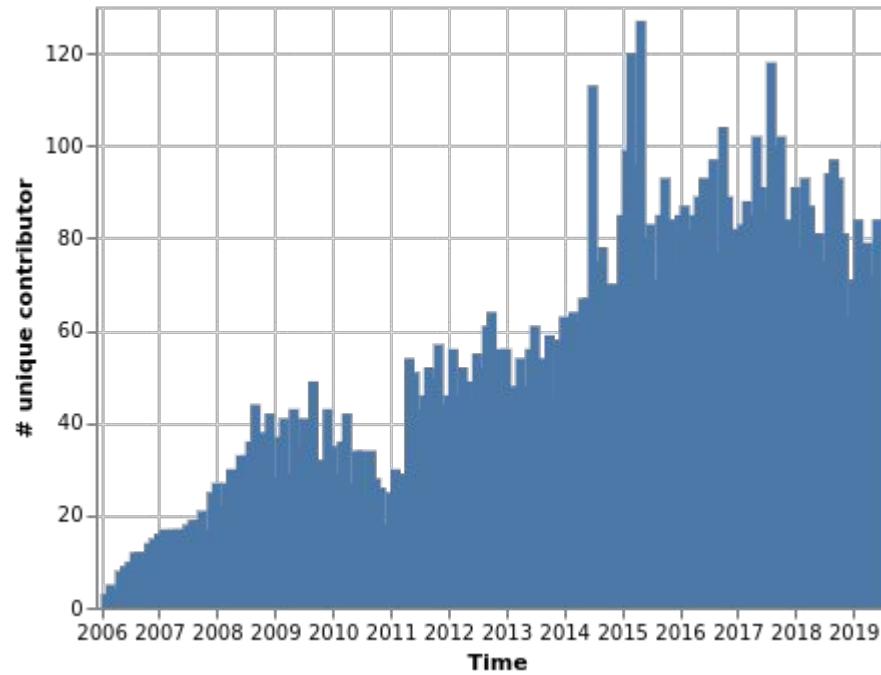
Cloud-native

Márton Elek

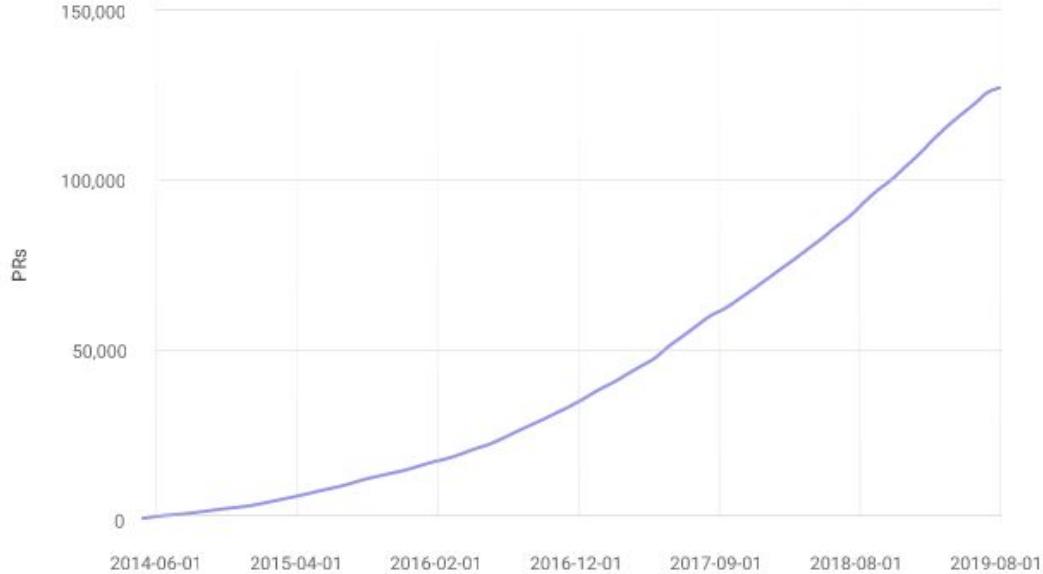
project resolved

1	HADOOP	1930
2	ARROW	1693
3	FLINK	1600
4	SPARK	1364
5	BEAM	1074
6	AIRFLOW	1073
7	HBASE	982
8	IGNITE	817
9	CAMEL	779
10	GEODE	708
11	IMPALA	627
12	HIVE	579
13	OAK	449
14	DLAB	445
15	KAFKA	441

Apache Hadoop

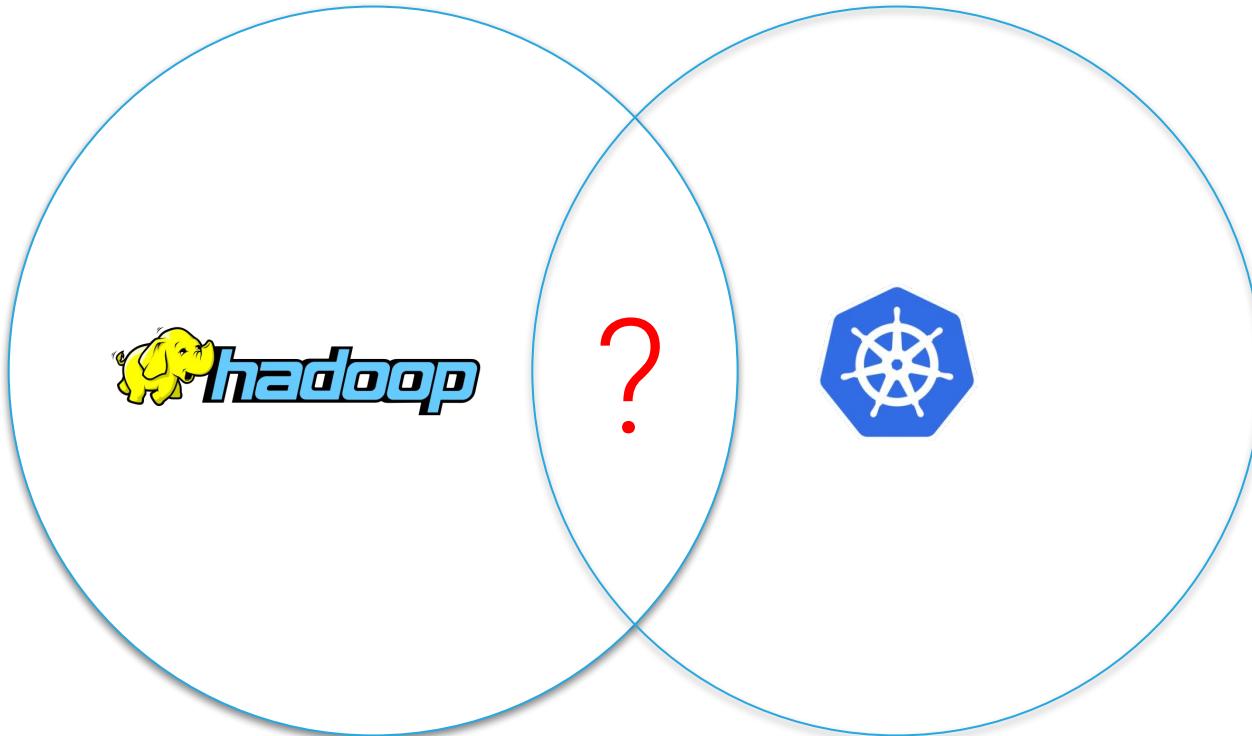


Kubernetes Pull Requests



Cumulative growth of Kubernetes pull requests over time

Source: [CNCF Kubernetes Project Journey Report](#)

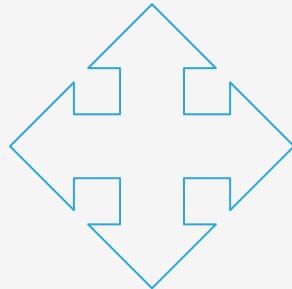


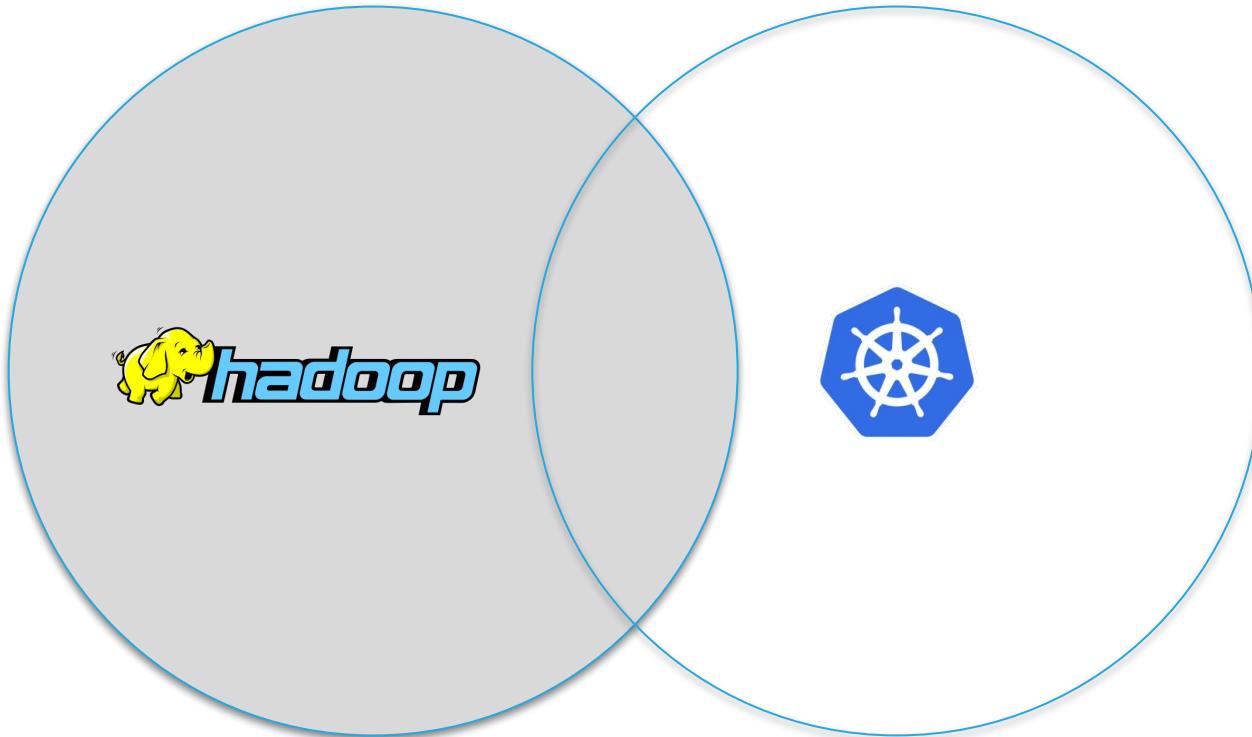


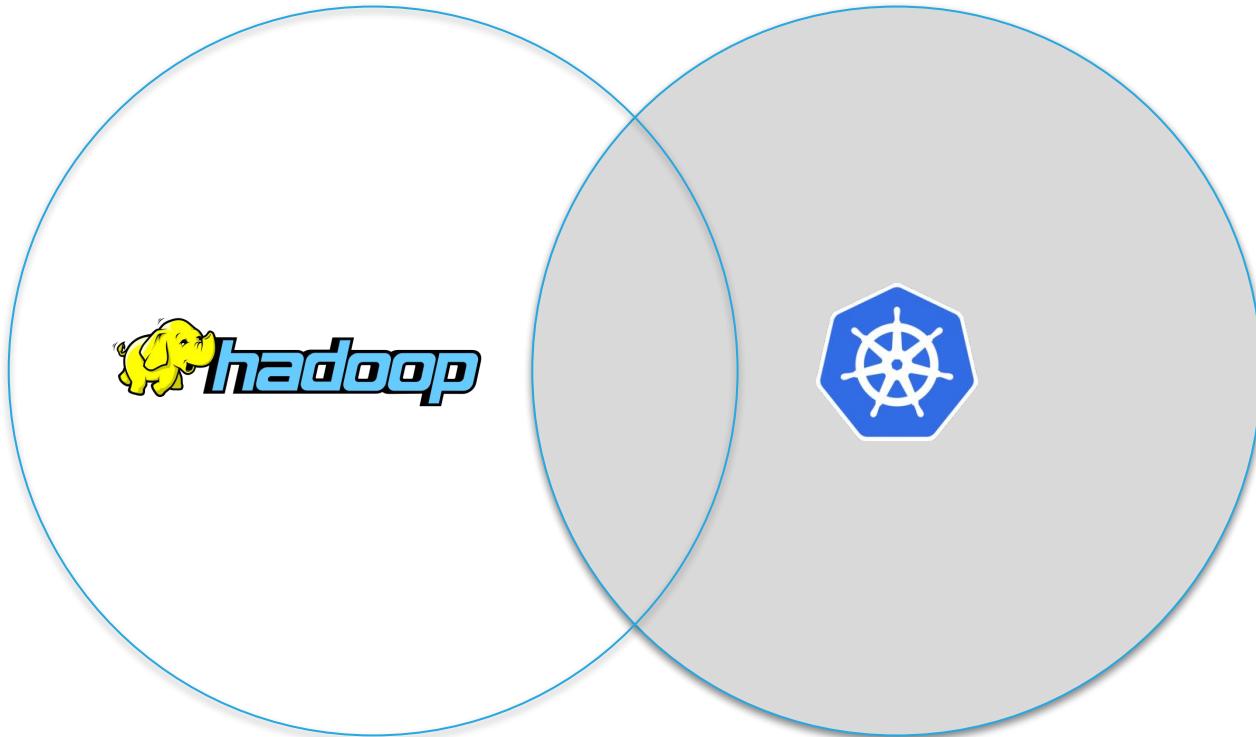












Cloud-native



?

Cloud-native



“Cloud native technologies empower organizations to build and run scalable applications in modern, dynamic environments such as public, private, and hybrid clouds.” ([CNCF charter](#))

WHAT?



Cloud-Native for me?

Márton Elek



Apache Hadoop contributor / PMC

→ Working on **Apache Hadoop Ozone**

Apache Ratis (Incubating) contributor / PMC

Principal Software Engineer **@Cloudera**

Márton Elek



Apache Hadoop contributor / PMC

→ Working on **Apache Hadoop Ozone**

Apache Ratis (Incubating) contributor / PMC

Principal Software Engineer **@Cloudera**



aws  S3 protocol

 Hadoop FS

 CSI

Apache Hadoop Ozone



hadoop.apache.org/ozone

Márton Elek



Apache Hadoop contributor / PMC

→ Working on **Apache Hadoop Ozone**

Apache Ratis (Incubating) contributor / PMC

elek@apache.org

@anzix

Flokkr → <https://github.com/flokkrr> (Like Bigtop but for Kubernetes)

Flekszible → <https://github.com/elek/flekszible> (Helm + Kustomize = Flekszible)



Cloud-Native for me?

What is cloud-native (def1) ?

Hadoop 2.7 release?

- Download one tar

The screenshot shows the Apache Software Foundation homepage. At the top, there is a navigation bar with links for News, About, Make a Donation, The Apache Way, and Join Us. Below the navigation bar, the Apache logo is displayed with the text "THE APACHE SOFTWARE FOUNDATION 20TH ANNIVERSARY". To the right of the logo, the text "CELEBRATING 20 YEARS OF COMMUNITY-LED DEV 'THE APACHE WAY'" is visible. Below this, there is another navigation bar with links for Projects, People, Community, and License. The main content area contains text and links related to downloading Hadoop 2.7.7. It includes a link to a mirror site (<http://xenia.sote.hu/ftp/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>) and other mirror sites. It also provides instructions for verifying file integrity using PGP signatures or hashes.

We suggest the following mirror site for your download:

<http://xenia.sote.hu/ftp/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

Other mirror sites are suggested below.

It is essential that you verify the integrity of the downloaded file using the PGP signature ([.asc](#) file) or a hash ([.md5](#) or [.](#)).

Please only use the backup mirrors to download KEYS, PGP and MD5 sigs/_hashes or if no other mirrors are working.

HTTP

<http://xenia.sote.hu/ftp/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

FTP

<ftp://crysystech.bme.hu/pub/apache/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

<ftp://xenia.sote.hu/pub/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

What is cloud-native (def1) ?

Hadoop 2.7 release?

- Download one tar

The screenshot shows the Apache Software Foundation homepage with a banner celebrating 20 years of community-led development "THE APACHE WAY". Below the banner, there's a navigation bar with links for News, About, Make a Donation, The Apache Way, and Join Us. Underneath the banner, there are links for Projects, People, Community, and License. A note suggests using xenia.sote.hu as a mirror site for the download. It provides a direct link to the Hadoop 2.7 tar.gz file: <http://xenia.sote.hu/ftp/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>. It also lists other mirror sites and instructions for verifying file integrity using PGP signatures or MD5 hashes.

We suggest the following mirror site for your download:

<http://xenia.sote.hu/ftp/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

Other mirror sites are suggested below.

It is essential that you verify the integrity of the downloaded file using the PGP signature ([.asc](#) file) or a hash ([.md5](#) or [.sha1](#)).

Please only use the backup mirrors to download KEYS, PGP and MD5 sigs/_hashes or if no other mirrors are working.

HTTP

<http://xenia.sote.hu/ftp/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

FTP

<ftp://crysystech.hit.bme.hu/pub/apache/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

<ftp://xenia.sote.hu/pub/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

Cloud-native Hadoop Ozone release?

- tar xvzf + ./bin/hdfs start
- docker run apache/ozone
- docker-compose up -d
- kubectl apply -f

...

Make it easy to start anywhere!

[Overview](#)[Getting Started](#)[Command Line Interface](#)[Programming Interfaces](#)[Security](#)[Concepts](#)[Beyond Basics](#)[Recipes](#)[Easy Start](#)[Running Ozone from Docker Hub](#)

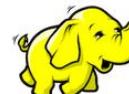
You can try out Ozone using docker hub without downloading the official release. This makes it easy to explore Ozone.

[Starting ozone inside a single container](#)

The simplest and easiest way to start an ozone cluster to explore what it can do is to start ozone via docker.

[Ozone in Docker](#)[Recommended](#)[Running Ozone from an Official Release.](#)

Apache Ozone can also be run from the official release packages. Along with the official source releases, we also release a set of convenience binary packages. It is easy to run these binaries in different configurations.

[Deploying Ozone on a physical cluster.](#)

Ozone is designed to work concurrently with HDFS. The physical cluster instructions explain each component of Ozone and how to deploy with maximum control.

[On-Prem Ozone Cluster](#)

The word "minikube" in a bold, blue, sans-serif font.

[Deploy Ozone using MiniKube.](#)

Ozone comes with a standard set of K8s resources. You can deploy them to MiniKube and experiment with the K8s based deployments.

[MiniKube Cluster](#)[Deploying Ozone on K8s](#)

Ozone is designed to work well under kubernetes. These are instructions on how to deploy Ozone on K8s platform. Ozone provides a replicated storage solution for K8s based Applications.

[Ozone on Kubernetes](#)[An Ozone cluster in Local Node.](#)

We also ship standard docker files with official release, if you want to use them. These are part of official release and not depend upon Docker Hub.

[Multi-Container Cluster](#)

Cloud native Ozone



AWS Services

Services ▾ Resource Groups ▾

History

EC2

S3

Console Home

IAM

Elastic Beanstalk

Find a service by name or feature (for example, EC2, S3 or VM, storage).

Group A-Z

Compute

- EC2
- Lightsail
- ECR
- ECS
- EKS
- Lambda
- Batch
- Elastic Beanstalk
- Serverless Application Repository

Robotics

- AWS RoboMaker

Blockchain

- Amazon Managed Blockchain

Analytics

- Athena
- EMR
- CloudSearch
- Elasticsearch Service
- Kinesis
- QuickSight
- Data Pipeline
- AWS Glue
- MSK

Business Applications

- Alexa for Business
- Amazon Chime
- WorkMail

End User Computing

- WorkSpaces
- AppStream 2.0
- WorkDocs
- WorkLink

Storage

- S3
- EFS
- FSx
- S3 Glacier
- Storage Gateway
- AWS Backup

Management & Governance

- AWS Organizations
- CloudWatch
- AWS Auto Scaling
- CloudFormation
- CloudTrail
- Config
- OpsWorks
- Service Catalog
- Systems Manager
- Trusted Advisor
- Managed Services
- Control Tower
- AWS License Manager
- AWS Well-Architected Tool
- Personal Health Dashboard

Security, Identity, & Compliance

- IAM
- Resource Access Manager
- Cognito
- Secrets Manager
- GuardDuty
- Inspector
- Amazon Macie
- AWS Single Sign-On
- Certificate Manager
- Key Management Service
- CloudHSM
- Directory Service
- WAF & Shield
- Artifact
- Security Hub

Internet Of Things

- IoT Core
- Amazon FreeRTOS
- IoT 1-Click
- IoT Analytics
- IoT Device Defender
- IoT Device Management
- IoT Events
- IoT Greengrass
- IoT SiteWise
- IoT Things Graph

Game Development

- Amazon GameLift

Migration & Transfer

- AWS Migration Hub
- Application Discovery Service
- Database Migration Service
- Server Migration Service
- AWS Transfer for SFTP

Media Services

- Elastic Transcoder
- Kinesis Video Streams
- MediaConnect
- MediaConvert
- MediaLive

Mobile

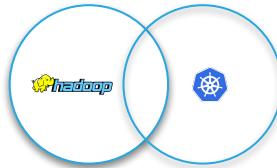
- AWS Amplify
- Mobile Hub
- AWS AppSync
- Device Farm

▲ close

Cloud-native is a
User Experience!

Improve
User Experience

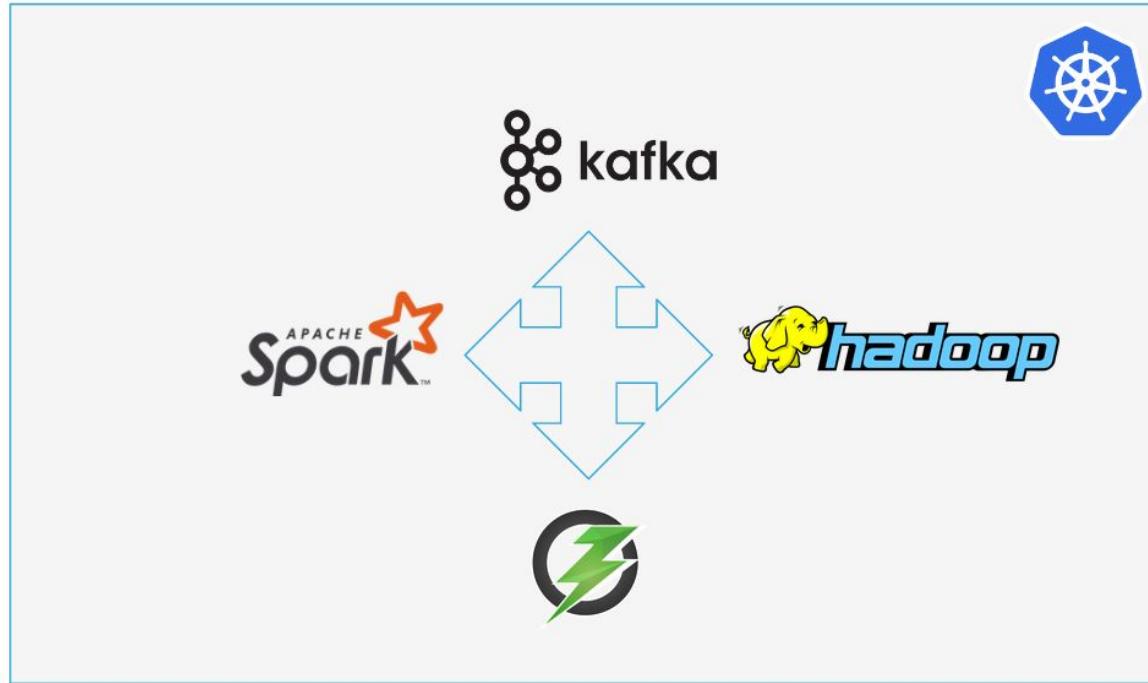
with the help of
collaboration



Requirements



Mange the whole platform





Fully customizable

Organize the complexity
Don't **hide** it



Sharable/Reusable part (Packaging)

Cluster 1:

- Spark, Ozone, Kafka
- Secure + HA

Cluster2:

- HDFS
- Secure + Router based federation



Composable feature set

Collaboration between components:

- Turn on HA for ALL the components
- Setup storage for all the components

Configuration in the Hadoop

- **core-site.xml**
 - **Default storage type / parameters**
 - **Security**
- hdfs-site.xml
 - HDFS specific configuration (ports, settings)
- yarn-site.xml
 - Execution specific configuration

YARN + HDFS

- HDFS containers
 - core-site.xml
 - hdfs-site.xml
- YARN containers
 - core-site.xml
 - hdfs-site.xml
 - yarn-site.xml

YARN + Ozone

- **Ozone containers**
 - **core-site.xml**
 - **ozone-site.xml**
- YARN containers
 - **core-site.xml**
 - **ozone-site.xml**
 - **yarn-site.xml**



Collaboration, collaboration, collaboration



Versioning / Migration (+GitOps)

- How to migrate the settings between multiple environments?
- How check previous version of the configs?
- How to rollback to the previous version?



Mange the whole platform



Fully customizable



Sharable/Reusable part



Composable, collaborative feature set



Versioning / Migration (+GitOps)

Traditional Hadoop cluster managers
(outside Kubernetes)



Mange the whole platform



Search

Clusters

Hosts

Diagnostics

Audits

Charts

Replication

Administration

Home

Status

All Health Issues

Configuration 13

All Recent Commands

Cluster 1

Cloudera Runtime 7.0.2 (Parcels)

4 Hosts

4

HBASE-1

1

HDFS-1

1

HIVE-1

1

HIVE_ON_TEZ-1

1

HUE-1

1

IMPALA-1

1

KAFKA-1

1

KUDU-1

1

LIVY-1

1

OOZIE-1

1

OZONE-1

3

SOLR-1

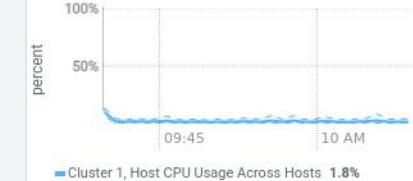
1

SPARK_ON_YARN-1

1

Charts

Cluster CPU



HDFS IO

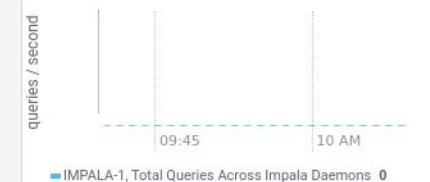


Home

Cluster Disk IO



Completed Impala Queries





Mange the whole platform



Fully customizable



Filters

[Clear All](#)[SCOPE](#)

HDFS-1 (Service-Wide)	20
Balancer	3
DataNode	13
Gateway	5
Https	13
JournalNode	10
NFS Gateway	9
NameNode	34
SecondaryNameNode	11
Failover Controller	10

[CATEGORY](#)

Advanced	96
Checkpointing	2
Cloudera Navigator	4
Erasure Coding	4
High Availability	5
Logs	37
Main	46
Monitoring	101
Performance	20
Ports and Addresses	26
Proxy	26
Replication	5
Resource Management	14
Security	61
Stacks Collection	5

[STATUS](#)

<input checked="" type="checkbox"/> Error	0
<input type="triangle-up"/> Warning	0
<input type="checkbox"/> Edited	1
<input type="checkbox"/> Pending	0

[Show All Descriptions](#)

Block Replica Placement Policy

dfs.block.replicator.classname

HDFS-1 (Service-Wide)

 HDFS Default Upgrade Domains

Enable HDFS Block Metadata API

dfs.datanode.hdfs-blocks-metadata.enabled

 HDFS-1 (Service-Wide)

System User

HDFS-1 (Service-Wide)

hdfs



System Group

HDFS-1 (Service-Wide)

hdfs



Shared Hadoop Group Name

HDFS-1 (Service-Wide)

hadoop



System User's Home Directory

HDFS-1 (Service-Wide)

/var/lib/hadoop-hdfs



HDFS Service Advanced Configuration Snippet (Safety Valve) for hdfs-site.xml

HDFS-1 (Service-Wide) [Undo](#)[View as XML](#)

Name



Value



Description

 Final



Mange the whole platform



Fully customizable



Sharable/Reusable part



Composable, collaborative feature set



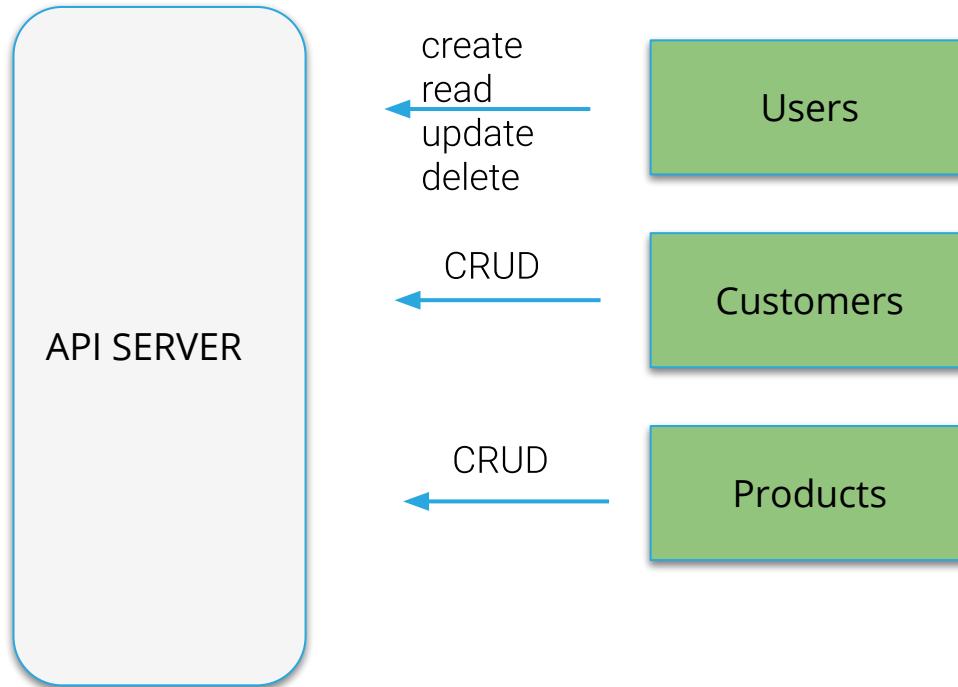
Versioning / Migration (+GitOps)



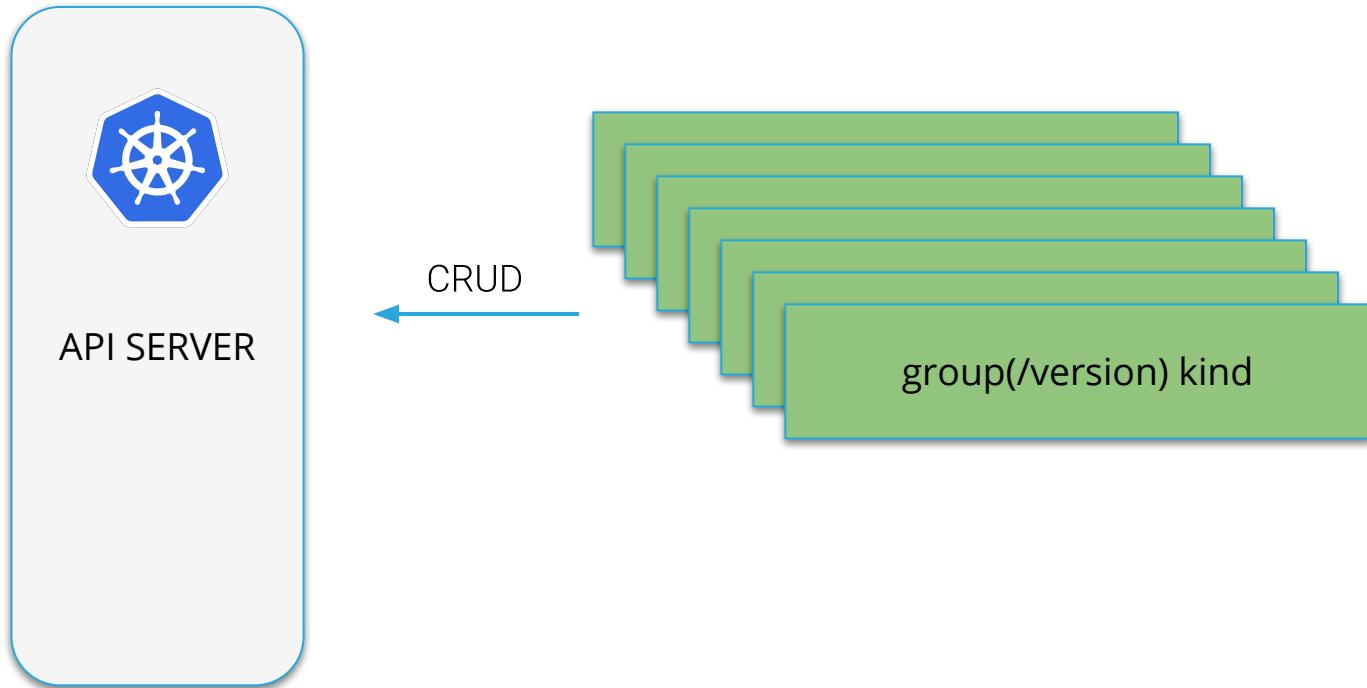
Same in  ?

How does  work?

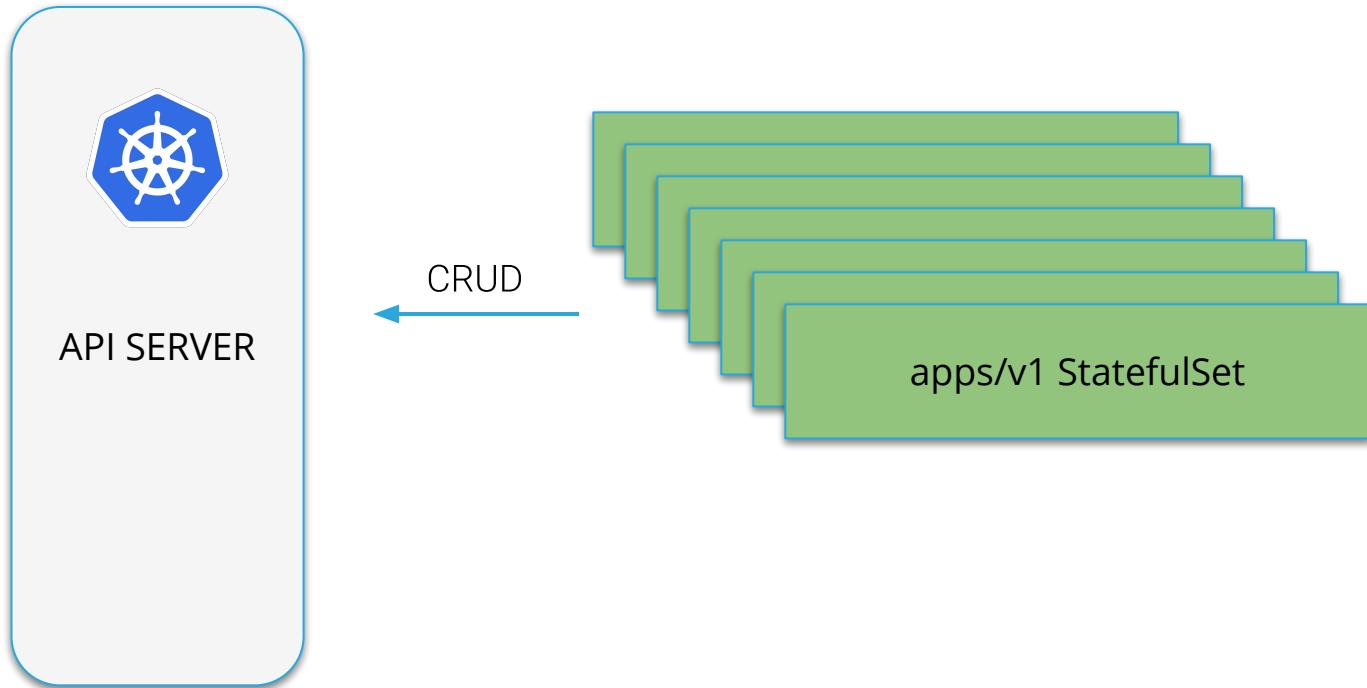
Traditional web service



Kubernetes



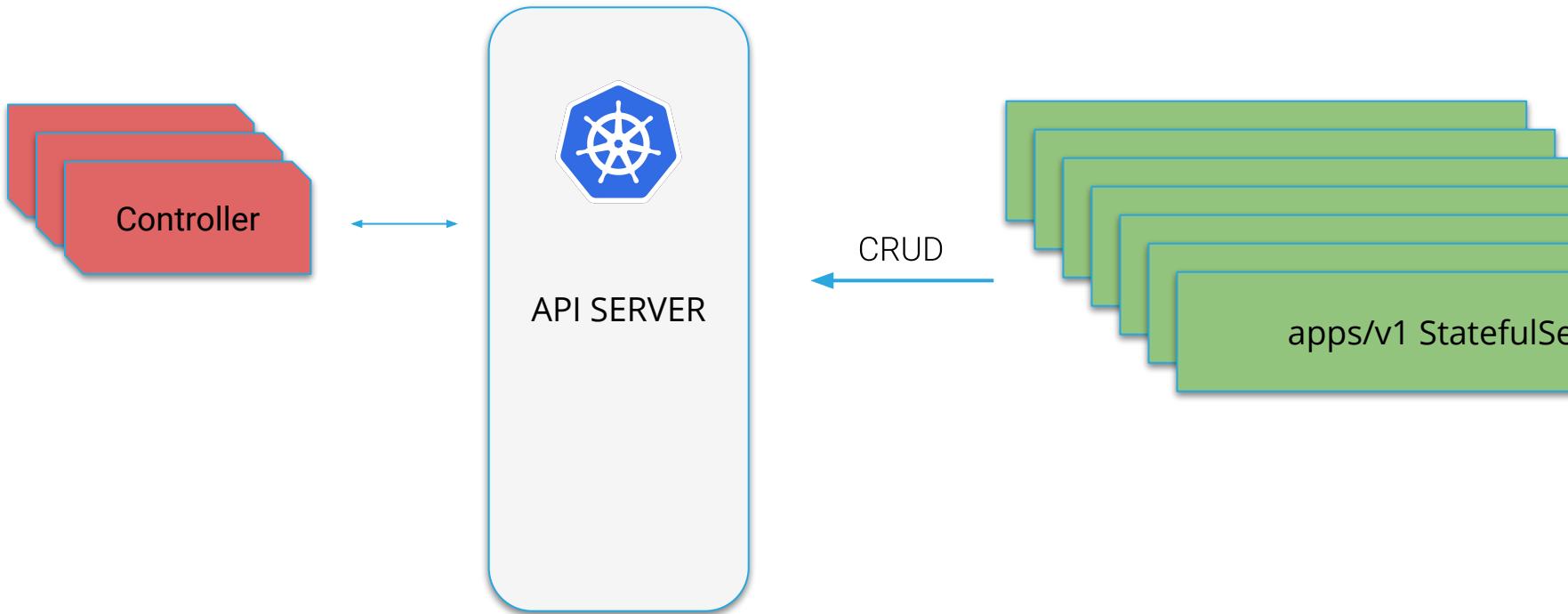
Kubernetes



Game of Yaml

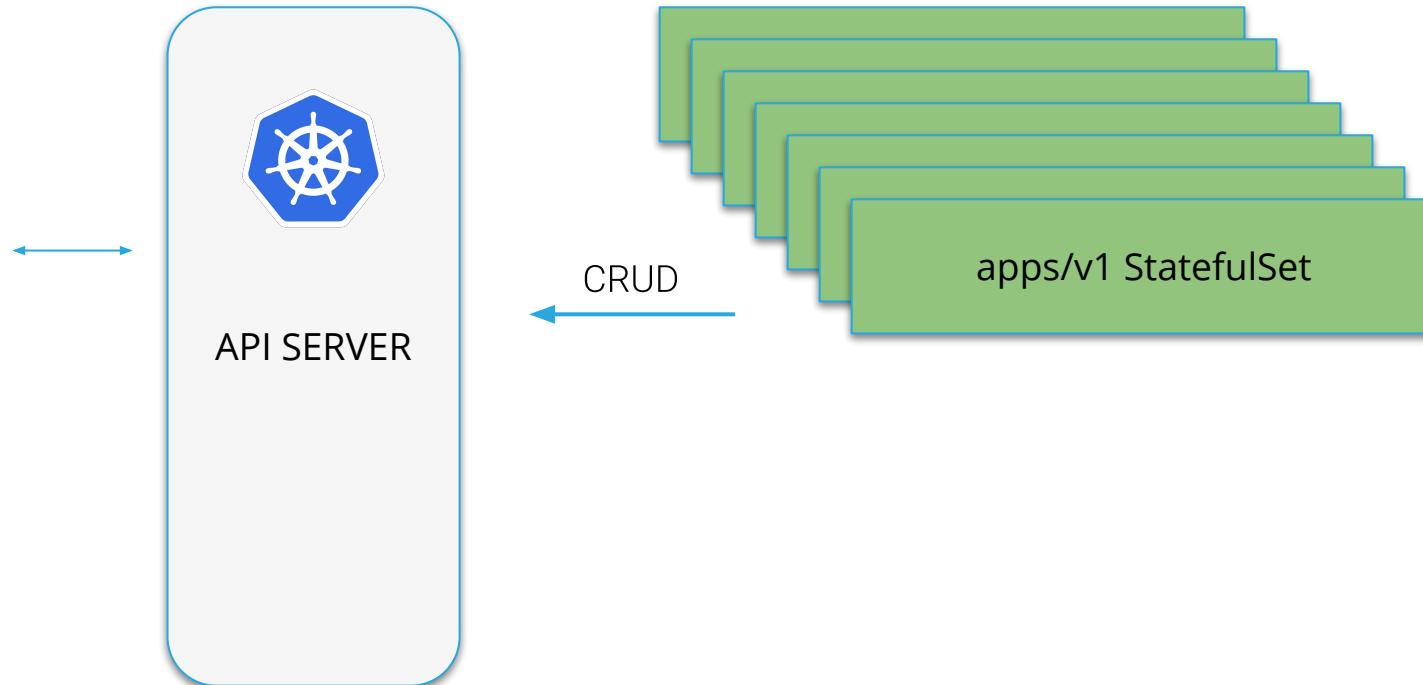
```
17 apiVersion: apps/v1
18 kind: StatefulSet
19 metadata:
20   name: scm
21   labels:
22     app.kubernetes.io/component: ozone
23 spec:
24   selector:
25     matchLabels:
26       app: ozone
27       component: scm
28   serviceName: scm
29   replicas: 1
30   template:
31     metadata:
32       labels:
33         app: ozone
34         component: scm
35       annotations:
```

Kubernetes

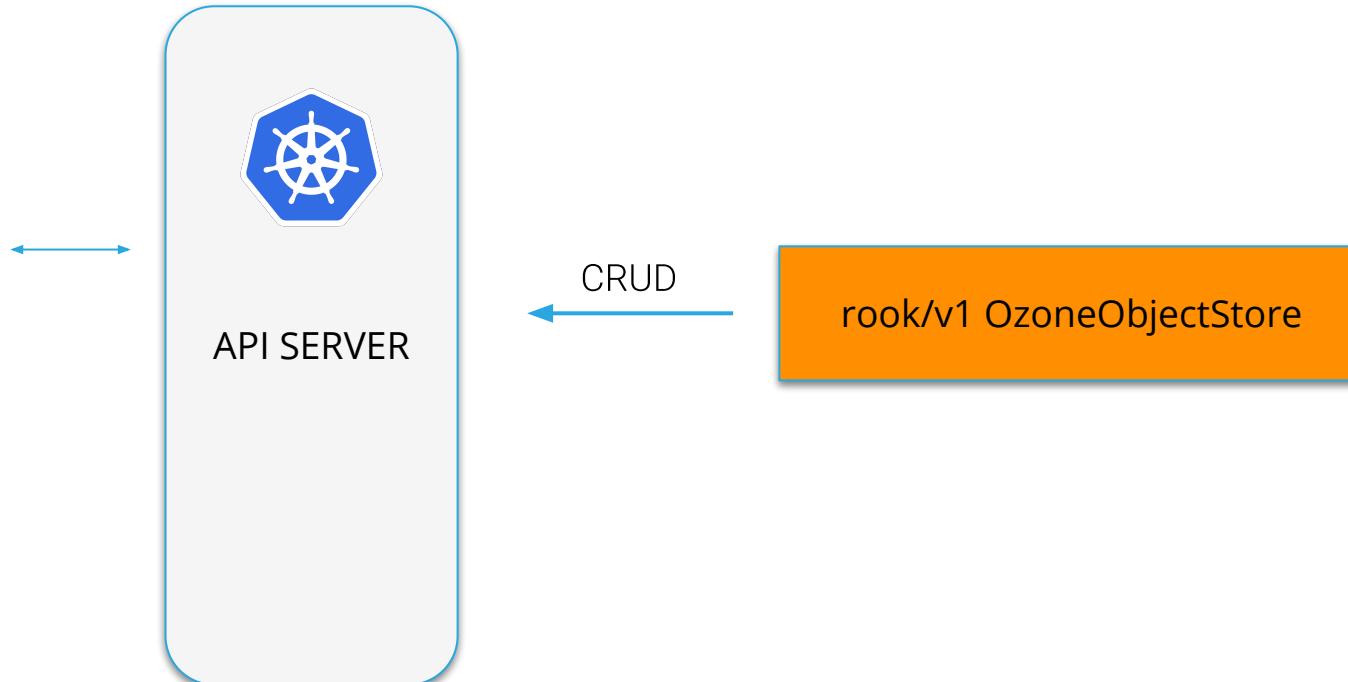


The operator pattern

Kubernetes + built-in resources



Kubernetes + CRD



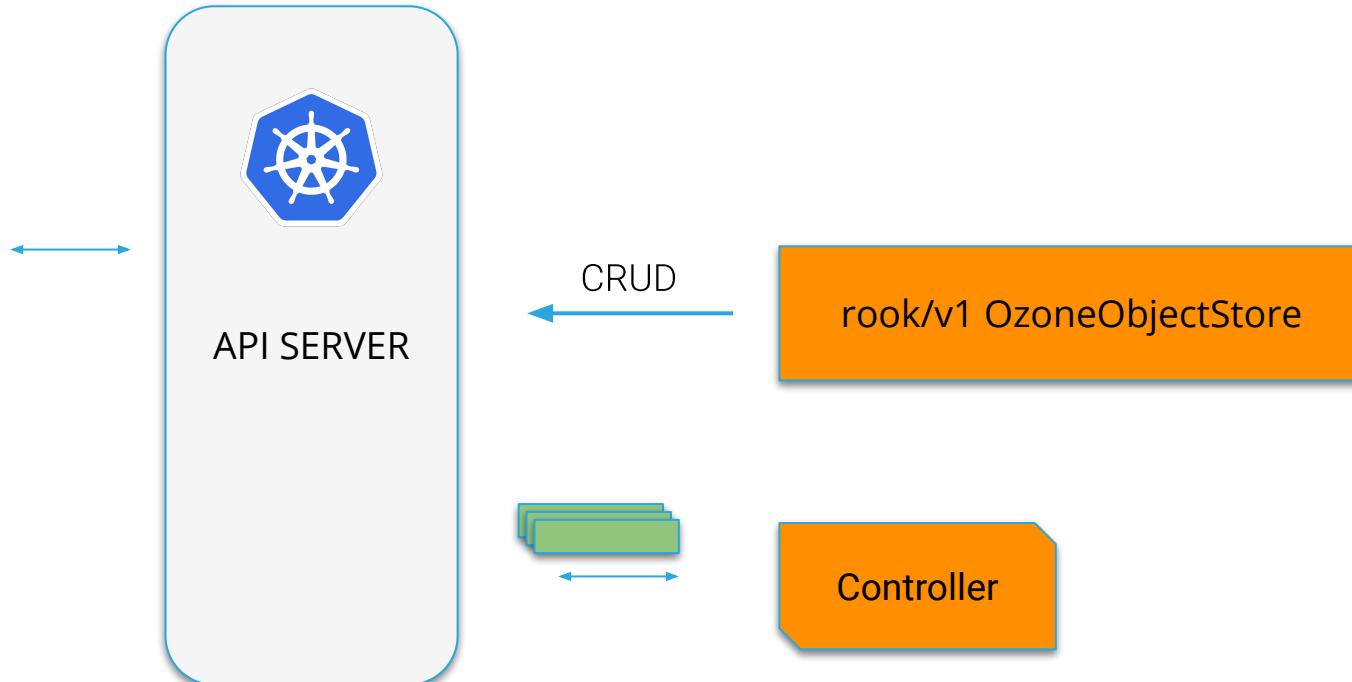
Without operator

```
17 apiVersion: apps/v1
18 kind: StatefulSet
19 metadata:
20   name: scm
21   labels:
22     app.kubernetes.io/component: ozone
23 spec:
24   selector:
25     matchLabels:
26       app: ozone
27       component: scm
28   serviceName: scm
29   replicas: 1
30   template:
31     metadata:
32       labels:
33         app: ozone
34         component: scm
35       annotations:
```

With operator

```
1  apiVersion: ozone.rook.io/v1alpha1
2  kind: OzoneObjectStore
3  metadata:
4    name: my-store
5  spec:
6    ozoneVersion:
7      image: apache/ozone:0.4.1
8    scope:
9      nodeCount: 4
10   storage:
11     hostDir: /data
```

Kubernetes + CRD



Operator pattern



Good to hide complexity from the **user**

```
1  apiVersion: ozone.rook.io/v1alpha1
2  kind: OzoneObjectStore
3  metadata:
4    name: my-store
5  spec:
6    ozoneVersion:
7      image: apache/ozone:0.4.1
8    scope:
9      nodeCount: 4
10   storage:
11     hostDir: /data|
```

Operator pattern



Good to hide complexity from the **user**



Bad to hide complexity from **admin**

OUTPUT = INPUT



+ **current Operator state**



Mange the whole platform



Fully customizable



Sharable/Reusable part



Composable, collaborative feature set



Versioning / Migration (+GitOps)



Tools

K8s Resource managers





[ksz]
fle[ksz]ible

- How to use?
- How does it work?
- Patterns / limitations

Disclaimer



[★ Star](#) 14,274

CNCF incubating



Kustomize

[★ Star](#) 3,923

kubernetes-sig

Disclaimer



★ Star 14,274

CNCF incubating



Kustomize

★ Star 3,923

kubernetes-sig

[ksz]

fle[ksz]ible

★ Star 16

NOT FOR SALE

Patterns are independent from the
implementations



How to use it?

```
elek ~ % cd tmp > deploy > ERROR % * minikube > helm init  
$HELM_HOME has been configured at /home/elek/.helm.
```

Tiller (the Helm server-side component) has been installed into your Kubernetes Cluster.

Please note: by default, Tiller is deployed with an insecure 'allow unauthenticated users' policy.

To prevent this, run `helm init` with the --tiller-tls-verify flag.

For more information on securing your installation see: https://docs.helm.sh/using_helm/#securing-your-helm-installation

```
elek ~ % cd tmp > deploy % * minikube > kc get pod -n kube-system  
NAME          READY   STATUS    RESTARTS   AGE  
coredns-5c98db65d4-85nhf   1/1     Running   1          8h  
coredns-5c98db65d4-shbmx   1/1     Running   1          8h  
etcd-minikube   1/1     Running   0          8h  
kube-addon-manager-minikube 1/1     Running   0          8h  
kube-apiserver-minikube   1/1     Running   0          8h  
kube-controller-manager-minikube 1/1     Running   0          8h  
kube-proxy-q97hg   1/1     Running   0          8h  
kube-scheduler-minikube   1/1     Running   0          8h  
storage-provisioner   1/1     Running   0          8h  
tiller-deploy-75f6c87b87-8frtk 0/1     ContainerCreating   0          6s
```



How to use it?

```
elek@om:~/tmp/deploy% * minikube helm install stable/hadoop --name hadoop1
NAME: hadoop1
LAST DEPLOYED: Thu Oct 24 07:18:01 2019
NAMESPACE: default
STATUS: DEPLOYED

RESOURCES:
==> v1/ConfigMap
NAME          DATA  AGE
hadoop1-hadoop  8    0s

==> v1/Pod(related)
NAME          READY  STATUS           RESTARTS  AGE
hadoop1-hadoop-hdfs-dn-0  0/1   ContainerCreating  0          0s
hadoop1-hadoop-hdfs-nn-0  0/1   ContainerCreating  0          0s
hadoop1-hadoop-yarn-nm-0  0/1   Pending          0          0s
hadoop1-hadoop-yarn-rm-0  0/1   ContainerCreating  0          0s

==> v1/Service
NAME          TYPE      CLUSTER-IP     EXTERNAL-IP   PORT(S)          AGE
hadoop1-hadoop-hdfs-dn  ClusterIP  None          <none>        9000/TCP,50075/TCP  0s
hadoop1-hadoop-hdfs-nn  ClusterIP  None          <none>        9000/TCP,50070/TCP  0s
hadoop1-hadoop-yarn-nm  ClusterIP  None          <none>        8088/TCP,8082/TCP,8042/TCP  0s
hadoop1-hadoop-yarn-rm  ClusterIP  None          <none>        8088/TCP          0s
hadoop1-hadoop-yarn-ui  ClusterIP  10.109.50.96  <none>        8088/TCP          0s
```

How to use it?

```

elek@om:~/tmp$ deploy % * minikube$ helm install stable/hadoop --name hadoop2 --set hdfs.webhdfs.enabled=true
NAME: hadoop2
LAST DEPLOYED: Thu Oct 24 07:21:39 2019
NAMESPACE: default
STATUS: DEPLOYED

RESOURCES:
==> v1/ConfigMap
NAME          DATA   AGE
hadoop2-hadoop  8      1s

==> v1/Pod(related)
NAME           READY  STATUS          RESTARTS  AGE
hadoop2-hadoop-hdfs-dn-0  0/1    ContainerCreating  0          1s
hadoop2-hadoop-hdfs-nn-0  0/1    ContainerCreating  0          0s
hadoop2-hadoop-yarn-nm-0  0/1    Pending          0          0s
hadoop2-hadoop-yarn-rm-0  0/1    Pending          0          0s

==> v1/Service
NAME          TYPE      CLUSTER-IP   EXTERNAL-IP  PORT(S)          AGE
hadoop2-hadoop-hdfs-dn  ClusterIP  None        <none>       9000/TCP,50075/TCP  1s
hadoop2-hadoop-hdfs-nn  ClusterIP  None        <none>       9000/TCP,50070/TCP  1s
hadoop2-hadoop-yarn-nm  ClusterIP  None        <none>       8088/TCP,8082/TCP,8042/TCP  1s
hadoop2-hadoop-yarn-rm  ClusterIP  None        <none>       8088/TCP          1s
hadoop2-hadoop-varn-ui  ClusterIP  10.96.174.124 <none>       8088/TCP          1s

```

```
elek om tmp > deploy % * minikube helm list
NAME      REVISION      UPDATED             STATUS      CHART          APP VERSION      NAMESPACE
hadoop2 1           1        Thu Oct 24 07:21:39 2019  DEPLOYED      hadoop-1.1.1    2.9.0          default
elek om tmp > deploy % * minikube kc get pod
NAME          READY   STATUS    RESTARTS   AGE
hadoop2-hadoop-hdfs-dn-0  1/1     Running   1          36m
hadoop2-hadoop-hdfs-nn-0  1/1     Running   0          36m
hadoop2-hadoop-yarn-nm-0  0/1     Pending   0          36m
hadoop2-hadoop-yarn-rm-0  1/1     Running   0          36m
elek om tmp > deploy % * minikube
```



How does it work?

Helm = package manager + go template

92 lines (91 sloc) | 2.84 KB

Raw Blame History

```
1 apiVersion: apps/v1beta1
2 kind: StatefulSet
3 metadata:
4   name: {{ include "hadoop.fullname" . }}-yarn-nm
5 annotations:
6   checksum/config: {{ include (print $.Template.BasePath "/hadoop-configmap.yaml") . | sha256sum }}
7 labels:
8   app: {{ include "hadoop.name" . }}
9   chart: {{ include "hadoop.chart" . }}
10  release: {{ .Release.Name }}
11  component: yarn-nm
12 spec:
13   serviceName: {{ include "hadoop.fullname" . }}-yarn-nm
14   replicas: {{ .Values.yarn.nodeManager.replicas }}
15 {{- if .Values.yarn.nodeManager.parallelCreate }}
16   podManagementPolicy: Parallel
17 {{- end }}
18   template:
19     metadata:
```

```
1 apiVersion: apps/v1beta1
2 kind: StatefulSet
3 metadata:
4   name: {{ include "hadoop.fullname" . }}-yarn-nm
5 annotations:
6   checksum/config: {{ include (print $.Template.BasePath "/hadoop-configmap.yaml") . | sha256sum }}
7 labels:
8   app: {{ include "hadoop.name" . }}
9   chart: {{ include "hadoop.chart" . }}
10  release: {{ .Release.Name }}
11  component: yarn-nm
12 spec:
13   serviceName: {{ include "hadoop.fullname" . }}-yarn-nm
14   replicas: {{ .Values.yarn.nodeManager.replicas }}
15 {{- if .Values.yarn.nodeManager.parallelCreate }}
16   podManagementPolicy: Parallel
17 {{- end }}
18   template:
19     metadata:
20       labels:
21         app: {{ include "hadoop.name" . }}
22         release: {{ .Release.Name }}
23         component: yarn-nm
24   spec:
25     affinity:
26       podAntiAffinity:
27         {{- if eq .Values.antiAffinity "hard" }}
28           requiredDuringSchedulingIgnoredDuringExecution:
29             - topologyKey: "kubernetes.io/hostname"
30             labelSelector:
```



Fully customizable



Versioning / Migration (+GitOps)





Mange the whole platform



Fully customizable



Sharable/Reusable part



Composable, collaborative feature set



Versioning / Migration (+GitOps)



How to use it?

- create base resource sets
- create transformations
- combine them and generate “kustomize build”



Fully customizable



```
.
├── image-patch-all.yaml
├── image-patch.yaml
└── kustomization.yaml
ozone
├── config.yaml
├── datanode-ss-service.yaml
├── datanode-ss.yaml
├── kustomization.yaml
├── om-ss-service.yaml
├── om-ss.yaml
├── s3g-ss-service.yaml
├── s3g-ss.yaml
├── scm-ss-service.yaml
└── scm-ss.yaml
ozone-freon
└── freon.yaml
└── kustomization.yaml
```

File: **image-patch-all.yaml**

```
1 - op: replace
2   path: /spec/template/spec/containers/0/image
3   value: build/ozone
4 - op: replace
5   path: /spec/template/spec/initContainers/0/image
6   value: build/ozone
```

File: kustomization.yaml

```
1 resources:
2   - ozone
3   - ozone-freon
4   namePrefix: test-
5   patchesJson6902:
6     - target:
7       group: apps
8       version: v1
9       kind: StatefulSet
10      name: om
11      path: image-patch-all.yaml
12     - target:
13       group: apps
14       version: v1
15       kind: StatefulSet
16       name: scm
17       path: image-patch-all.yaml
18     - target:
19       group: apps
```

How does it work?

kustomize build → yaml files are generated

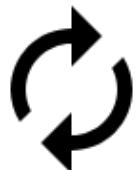
NO MAGIC



Mange the whole platform



Fully customizable



Sharable/Reusable part



Composable, collaborative feature set



Versioning / Migration (+GitOps)



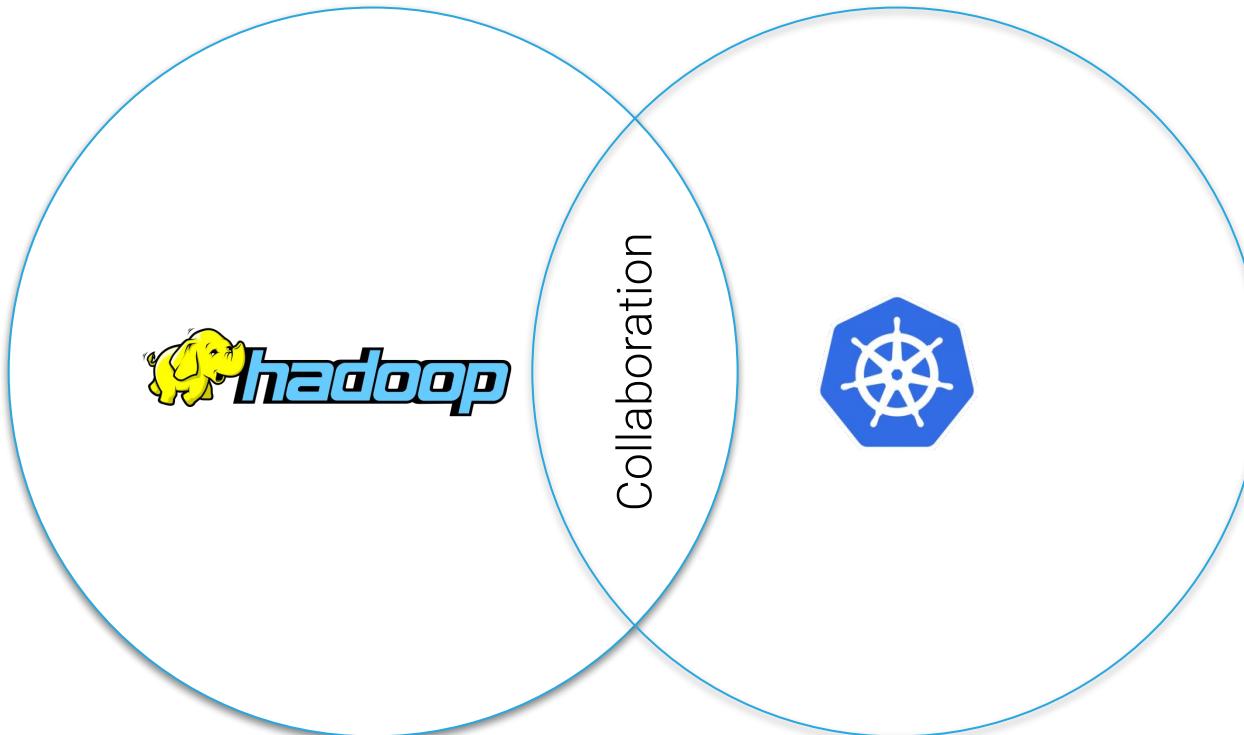
[ksz]

fle[ksz]ible

Flexibility and Collaboration?

Flekszible Demo

Summary



Summary

- **Collaboration** is the key ← *We need to invest more effort*
 - In Cloud-native environment
 - In Hadoop big-data ecosystem
- Collaboration is hard with the existing K8s tool
 - → *We need to invest more effort*

Q&A

Márton Elek

elek@apache.org // @anzix

helm + kustomize =

<https://github.com/elek/flekszible>

Sources

Material icons from:

https://material.io/resources/icons/?icon=format_line_spacing&style=baseline