

Cloud-native refactor

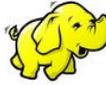
Anu Engineer, Marton Elek

What is Apache Hadoop Ozone?

Generic **Object store** based on Hadoop Storage layer.



S3 protocol



Hadoop FS



CSI

Apache Hadoop Ozone



hadoop.apache.org/ozone



What is Apache Hadoop Ozone?

Generic **Object store** based on Hadoop Storage layer.

SCALABLE

Ozone is designed to scale to tens of billions of files and blocks and, in the future, even more.

SECURE

Ozone integrates with kerberos infrastructure for access control and supports TDE and on-wire encryption.

CONSISTENT

Ozone is a strongly consistent object store. This consistency is achieved by using protocols like RAFT.

MULTI-PROTOCOL SUPPORT

Ozone supports different protocols like S3 and Hadoop File System APIs.

CLOUD-NATIVE

Ozone is designed to work well in containerized environments like YARN and Kubernetes.

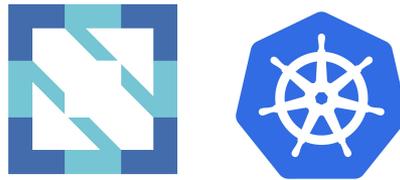
HIGHLY AVAILABLE

Ozone is a fully replicated system that is designed to survive multiple failures.

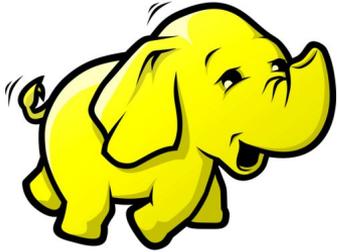
Disclaimer



“All the world's a stage,
And all the men and women merely
players”

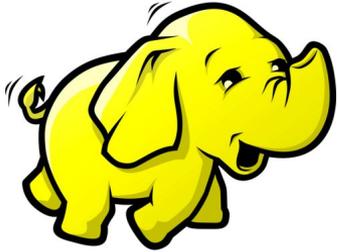


Cloud-Native world



Hadoop Storage

Cloud-Native world



Hadoop Storage



Cloud-Native world

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)

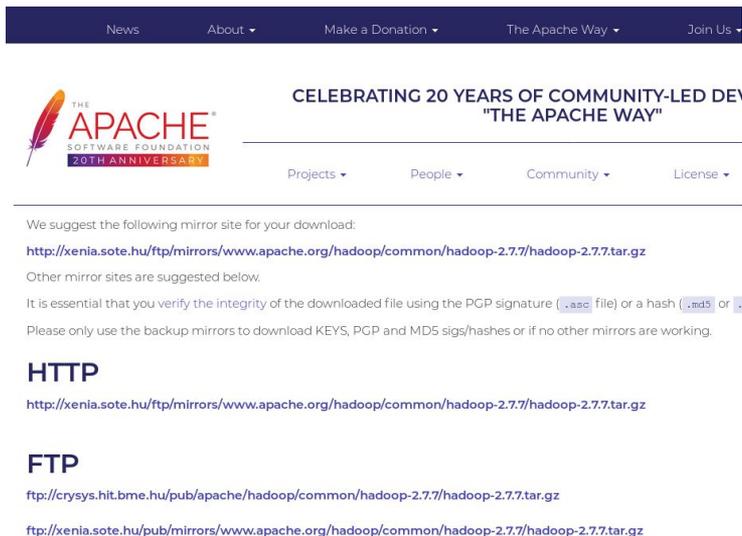
Cloud native computing uses an **open source** software stack to deploy applications as **microservices**, packaging each part into its own **container**, and dynamically orchestrating those containers to **optimize resource** utilization (cncf.io)

Explain Like I'm 5?

What is cloud-native (def1) ?

Hadoop 2.7 release?

- Download one tar



News About ▾ Make a Donation ▾ The Apache Way ▾ Join Us ▾

 CELEBRATING 20 YEARS OF COMMUNITY-LED DEV
"THE APACHE WAY"

Projects ▾ People ▾ Community ▾ License ▾

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 `.sha`).

Please only use the backup mirrors to download KEYS, PGP and MDS 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

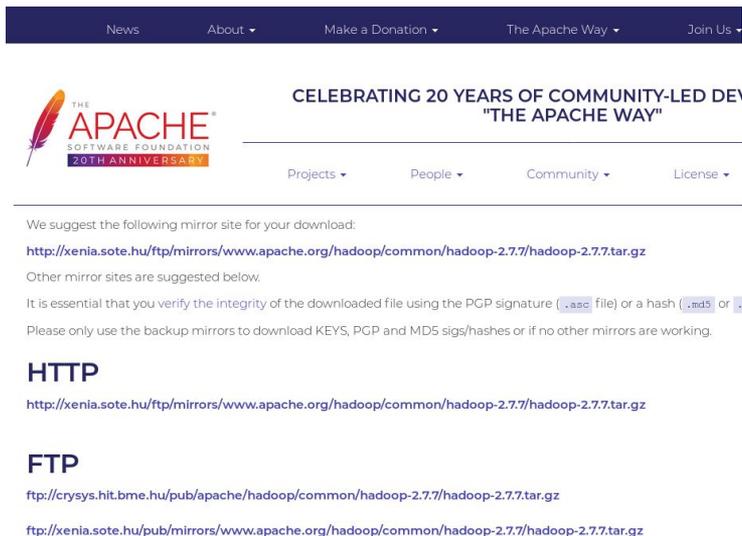
<ftp://crysyst.hit.bme.hu/pub/apache/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

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

What is cloud-native (def1) ?

Hadoop 2.7 release?

- Download one tar



The screenshot shows the Apache Hadoop 2.7.7 release page. At the top, there is a dark blue navigation bar with links for News, About, Make a Donation, The Apache Way, and Join Us. Below this is a banner celebrating the 20th anniversary of the Apache Software Foundation, with the text "CELEBRATING 20 YEARS OF COMMUNITY-LED DEVELOPMENT 'THE APACHE WAY'". The Apache logo is on the left, and navigation links for Projects, People, Community, and License are on the right. The main content area suggests a mirror site for download: <http://xenia.sote.hu/ftp/mirrors/www.apache.org/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>. It also provides instructions on how to verify the integrity of the downloaded file using PGP signatures or hashes, and lists backup mirrors for KEYS, PGP, and MDS signatures.

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` file).

Please only use the backup mirrors to download KEYS, PGP and MDS 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://crysyst.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>

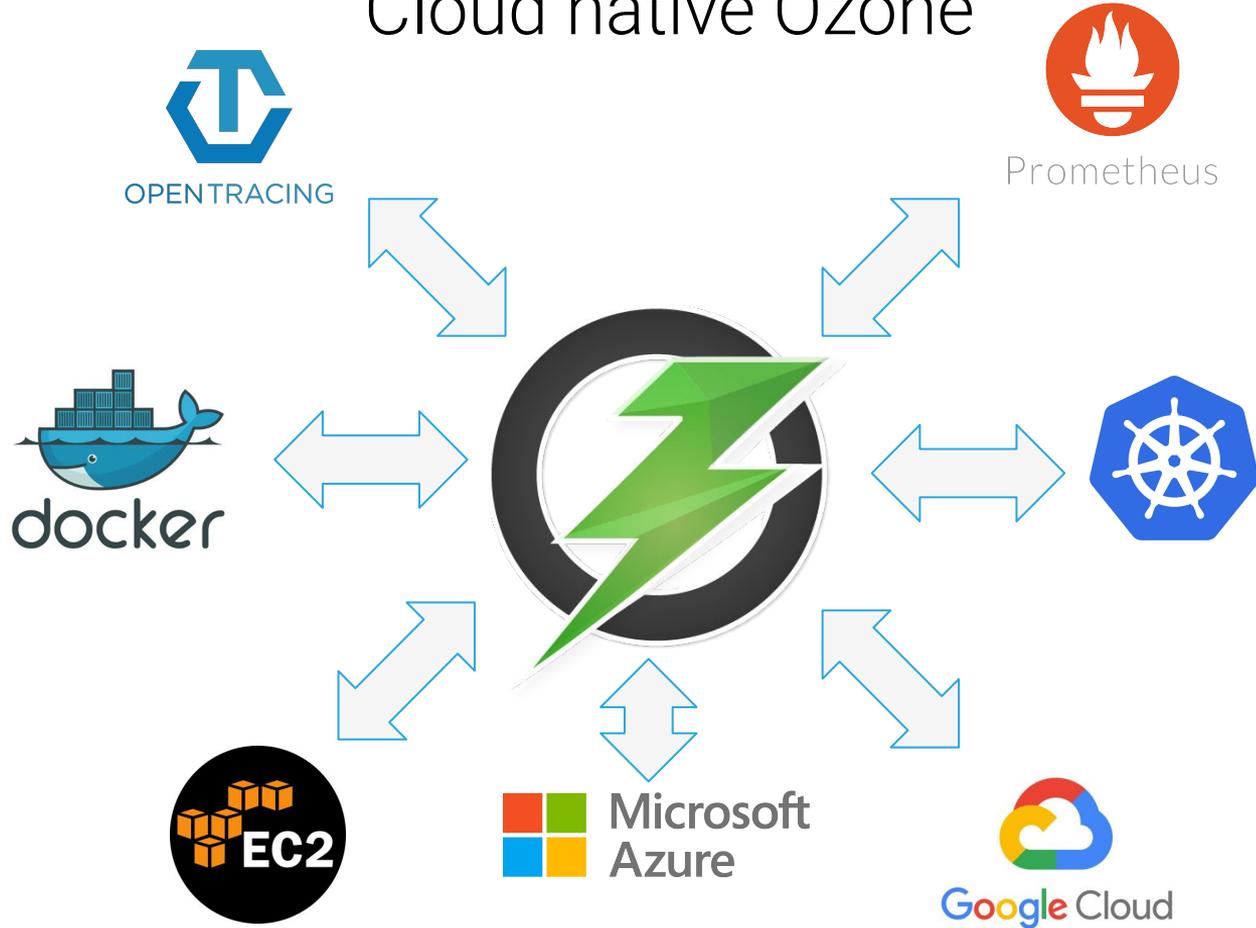
Cloud-native Hadoop release?

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

...

Make it easy to start anywhere!

Cloud native Ozone



Cloud-Native (def2)

The screenshot displays the AWS Services console interface. At the top, the AWS logo is on the left, and 'Services' and 'Resource Groups' are on the right. Below the logo is a 'History' sidebar with links to EC2, S3, Console Home, IAM, and Elastic Beanstalk. A search bar is positioned above the service grid, containing the text 'Find a service by name or feature (for example, EC2, S3 or VM, storage)'. To the right of the search bar are 'Group' and 'A-Z' buttons. The main area is a grid of service cards, each with an icon and a list of services. The categories and services shown are:

- Compute:** EC2, Lightsail, ECR, ECS, EKS, Lambda, Batch, Elastic Beanstalk, Serverless Application Repository
- Storage:** S3, EFS, FSx, S3 Glacier, Storage Gateway, AWS Backup
- Database:** RDS, DynamoDB, ElastiCache, Neptune, Amazon Redshift, Amazon DocumentDB
- Migration & Transfer:** AWS Migration Hub, Application Discovery Service, Database Migration Service, Server Migration Service, AWS Transfer for SFTP
- Robotics:** AWS RoboMaker
- Blockchain:** Amazon Managed Blockchain
- Satellite:** Ground Station
- 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
- Media Services:** Elastic Transcoder, Kinesis Video Streams, MediaConnect, MediaConvert, MediaLive
- 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
- 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
- Mobile:** AWS Amplify, Mobile Hub, AWS AppSync, Device Farm

At the bottom center of the grid, there is a 'close' button.

Cloud-native is an
User Experience!

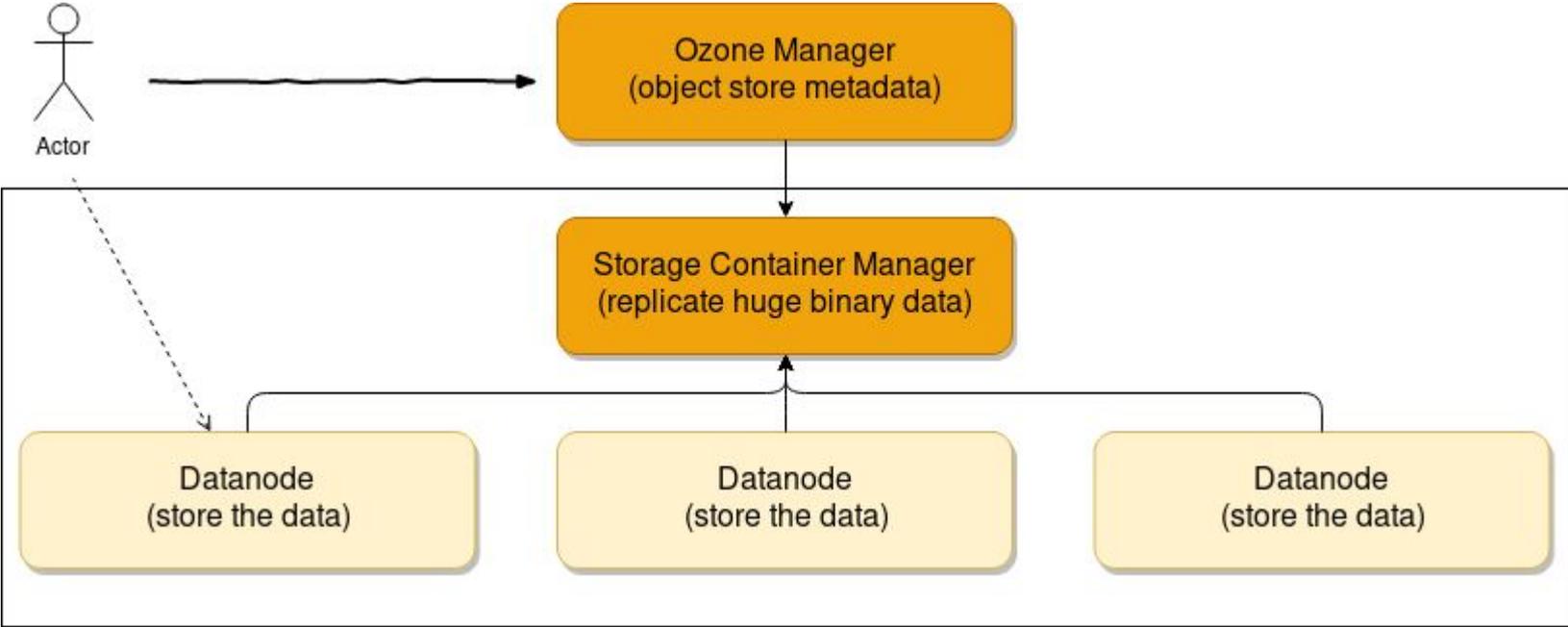


Results

- Cloud-native user experience
- **Support multiple runtime environments**
 - **Kubernetes?** kubectl apply -f .
 - **Local environment?** docker run -P 9878 apache/ozone
 - **Pseudo cluster?** cd compose/ozonesecure && docker-compose up -d
 - **Standalone cluster?** ./bin/ozone scm start
- awscli s3 cp s3://ozonebucket/file /tmp/localfile

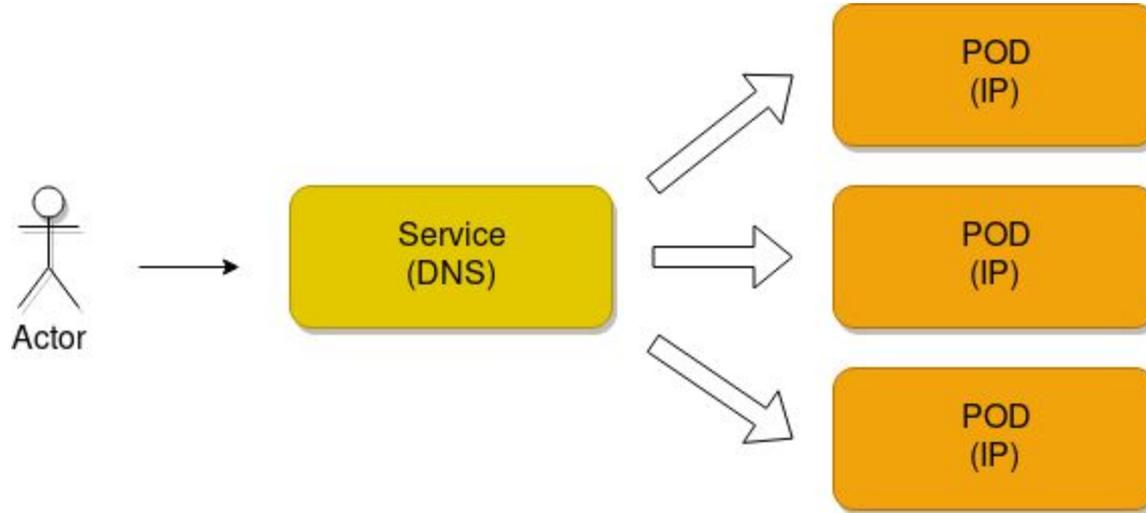
Network

Apache Hadoop Ozone

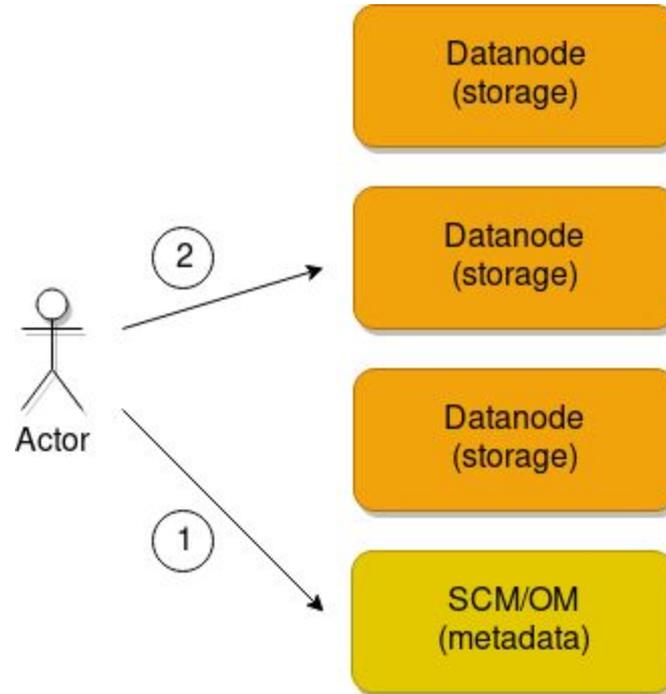


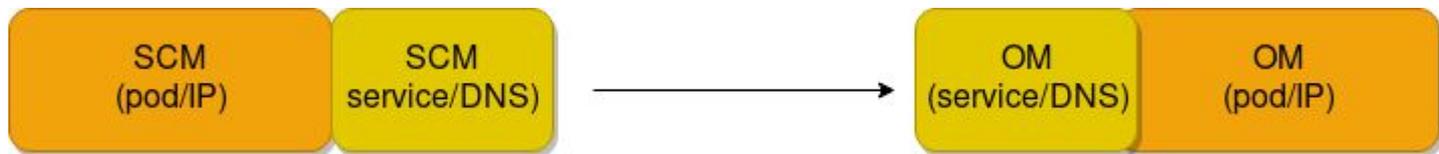
What about DNS?

Kubernetes for Stateless apps



Architecture of a storage app





DNS is dynamic

Please be prepared

- DNS can be changed (and may not be available at the time of start)
- DNS can be separated from pod (pod IP may have no reverse DNS)

Security

Security support

Hadoop Security:

- Kerberos based security
- Ozone: based on certificates
 - Can be initialized with kerberos
 - Or with pre-generated certificate

Kubernetes security

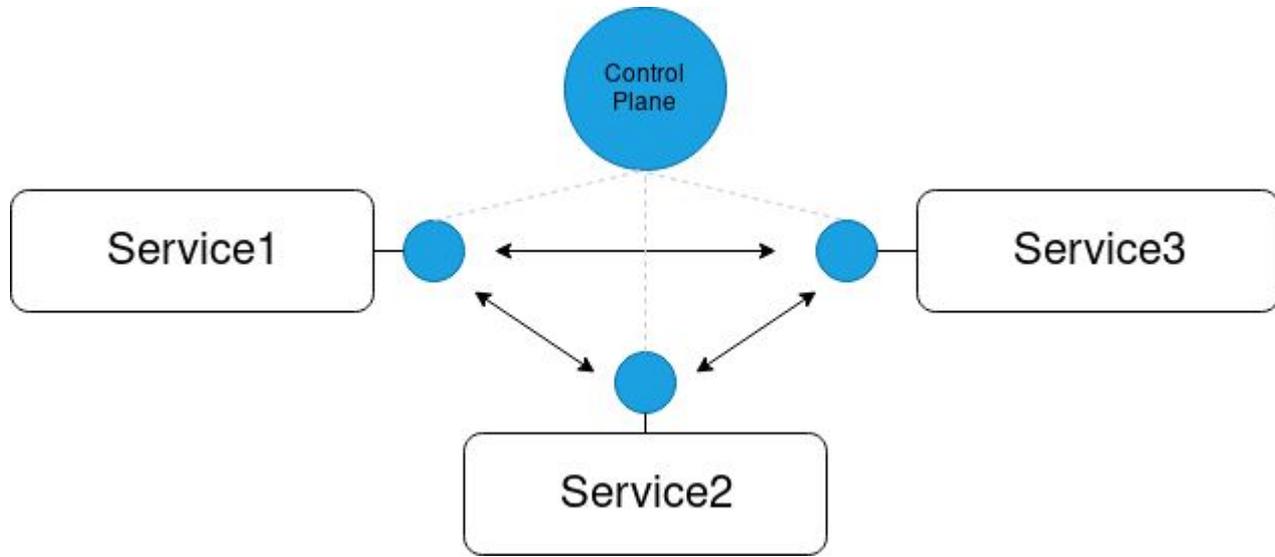
- RBAC for K8s API usage
- Secret management

Service Mesh

“The term service mesh is used to describe the network of microservices that make up such applications and the interactions between them.” (istio.io)

Explain Like I'm 5?

Centrally managed,
component side proxies/reverse proxies



Nothing is free

```
elek sc ~ istio * kubernetes
```

To be or not to be?

Unified transport library (hadoop-rpc)

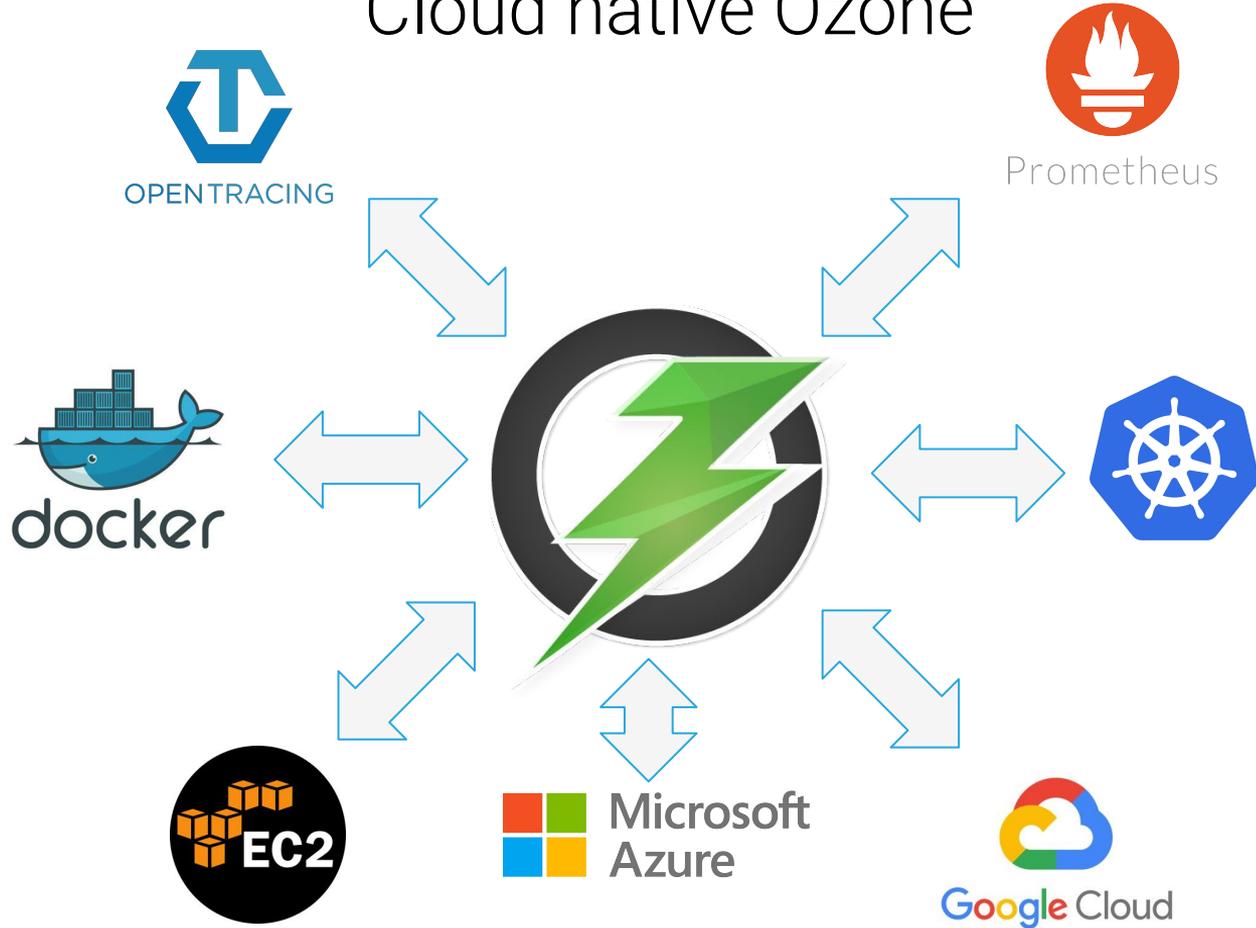
- Predefined network rules
- Advanced visibility
- Fine grained protocol support
- Security is included

Service Mesh

- Can be controlled from control plane
- Advanced visibility
- Generic protocol support
- Transparent security

Observability

Cloud native Ozone



Visibility

Hadoop:

- Hadoop metrics
 - Custom implementation
 - Supports multiple sink implementation

Cloud-native

- Metrics
 - Prometheus
 - Visualization with other tools (eg. Grafana)

Visibility

Hadoop:

- Hadoop metrics
 - Custom implementation
 - Supports multiple sink implementation
- *HTrace (deprecated)*

Cloud-native

- Metrics
 - Prometheus
 - Visualization with other tools (eg. Grafana)
- Tracing
 - OpenTracing
 - OpenCensus

Visibility

Hadoop:

- Hadoop metrics
 - Custom implementation
 - Supports multiple sink implementation
- *HTrace (deprecated)*
- *Log4j (no collections)*

Cloud-native

- Metrics
 - Prometheus
 - Visualization with other tools (eg. Grafana)
- Tracing
 - OpenTracing
 - OpenCensus
- Log collection
 - Fluentd, ...

Results

Hadoop:

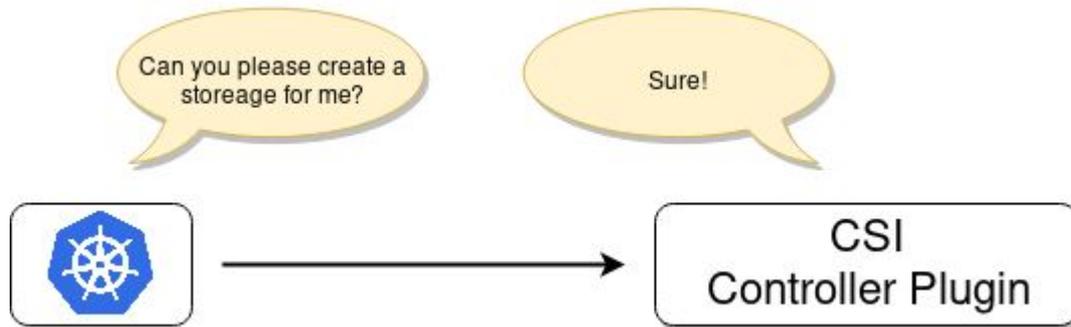
- **Hadoop metrics**
 - Custom implementation
 - Supports multiple sink implementation
 - **Prometheus endpoint**
- *HTrace (deprecated)*
- *Log4j (no collections)*

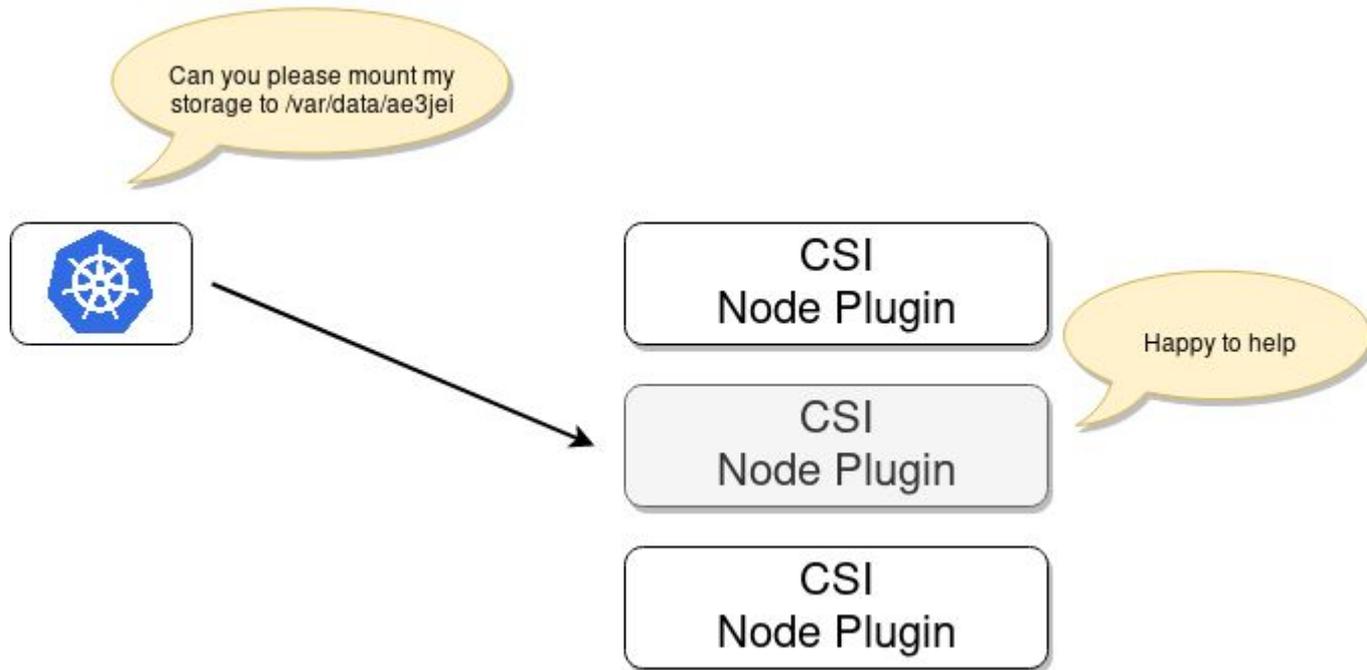
Cloud-native

- Metrics
 - **Prometheus**
 - Visualization with other tools (***embedded Grafana***)
- Tracing
 - **OpenTracing**
 - OpenCensus
- **Log collection**
 - Fluentd, ...

CSI

Container Storage Interface: Vendor-neutral interface for volume management





Results

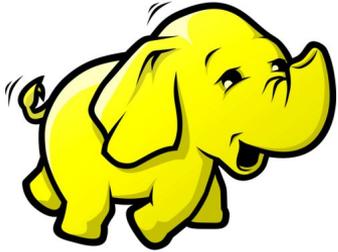
CSI server is included (easy part)

- Create/delete volumes/buckets
- Mount/umount the volume as a real filesystem (on the right node)

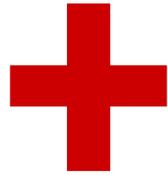
Hard part is to mount the file system (data path):

- Present: mounting via S3 Fuse drivers
 - Multiple Implementation, Multiple caching strategy
- Future/WIP: Native Fuse adapter to mount Ozone buckets/Hadoop fs

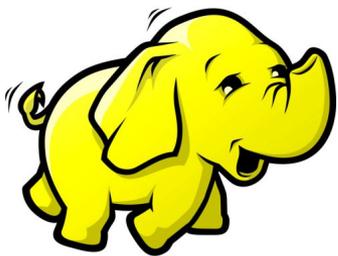
Summary



Hadoop Storage



Cloud-Native world



Hadoop Storage



Hadoop Ozone

Summary

Cloud-Native Ozone

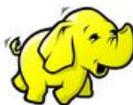
- CSI interface
- Prometheus/Opentracing support
- Easy to use in containers
- First class citizen of Kubernetes

Not the end of the Journey:

- Current version: 0.4.0 (S3, CSI, security)
- GA: 2019Q3 (HA)
- More improvements on the cloud-native side



S3 protocol



Hadoop FS



CSI

Apache Hadoop Ozone



hadoop.apache.org/ozone

Q&A



Apache Hadoop Ozone

<https://hadoop.apache.org/ozone>

Anu Engineer

aengineer@apache.org

Márton Elek

elek@apache.org // @anzix

helm + kustomize =

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