Cloud-native refactor

Anu Engineer, Marton Elek

cloudera

What is Apache Hadoop Ozone?

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

aws S3 protocol

Hadoop FS

CSI

Apache Hadoop Ozone

hadoop.apache.org/ozone

THE APACHE SOFTWARE FOUNDATION

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.

CONSISTENT

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

CLOUD-NATIVE

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

SECURE

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

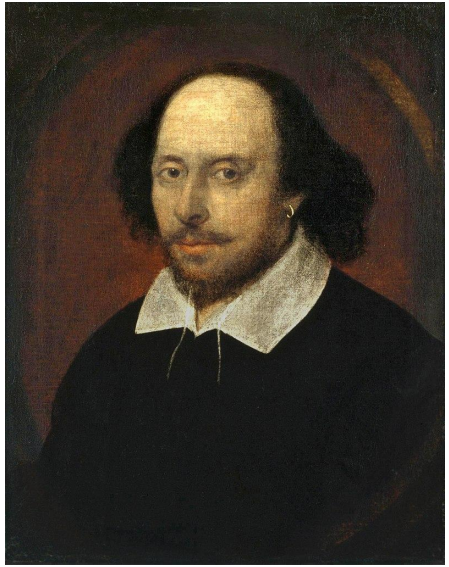
MULTI-PROTOCOL SUPPORT

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

HIGHLY AVAILABLE

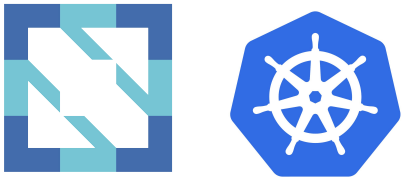
Ozone is a fully replicated system that is designed to survive multiple failures.

Disclaimer

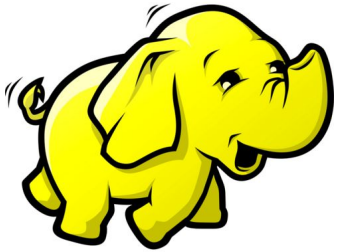


“All the world's a stage,

And all the men and women merely players”

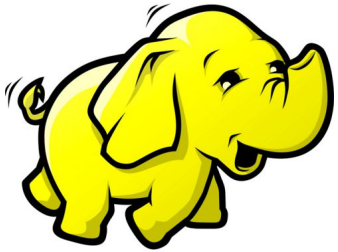


Cloud-Native world



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)

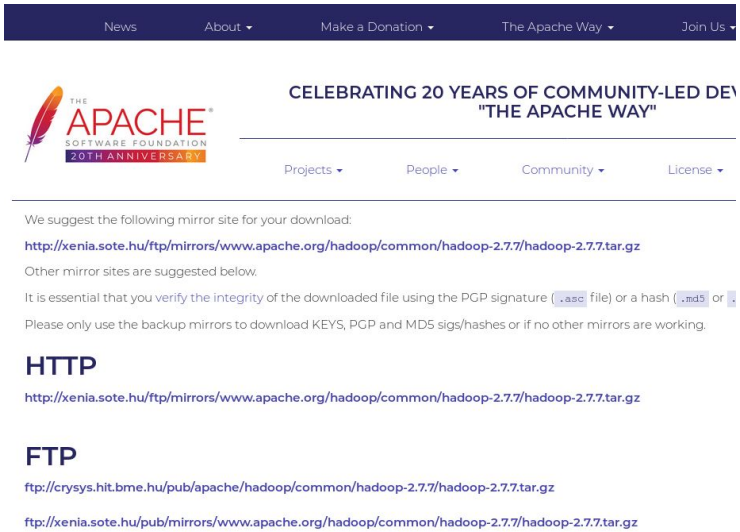
Cloud native computing uses an **open source** software stack to deploy applications as **microservices**, packaging each part into its own **container**, and dynamically orchestrating those containers to **optimize resource** utilization (cncf.io)

Explain Like I'm 5?

What is cloud-native (def1) ?

Hadoop 2.7 release?

- Download one tar

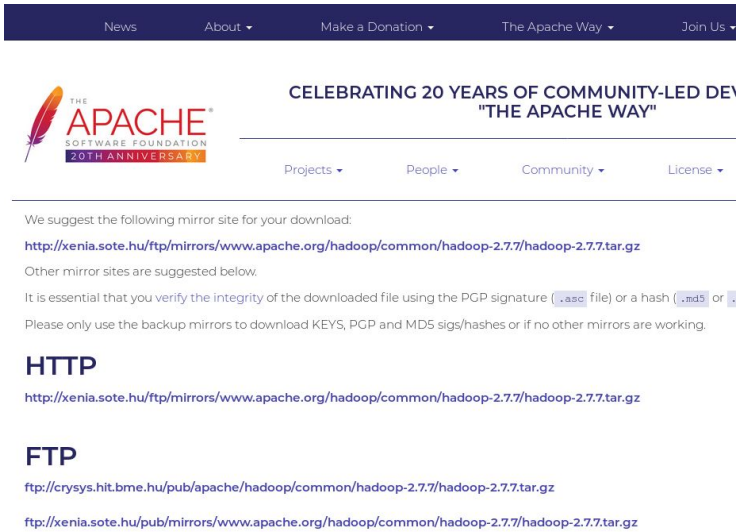


cloudera

What is cloud-native (def1) ?

Hadoop 2.7 release?

- Download one tar



Cloud-native Hadoop release?

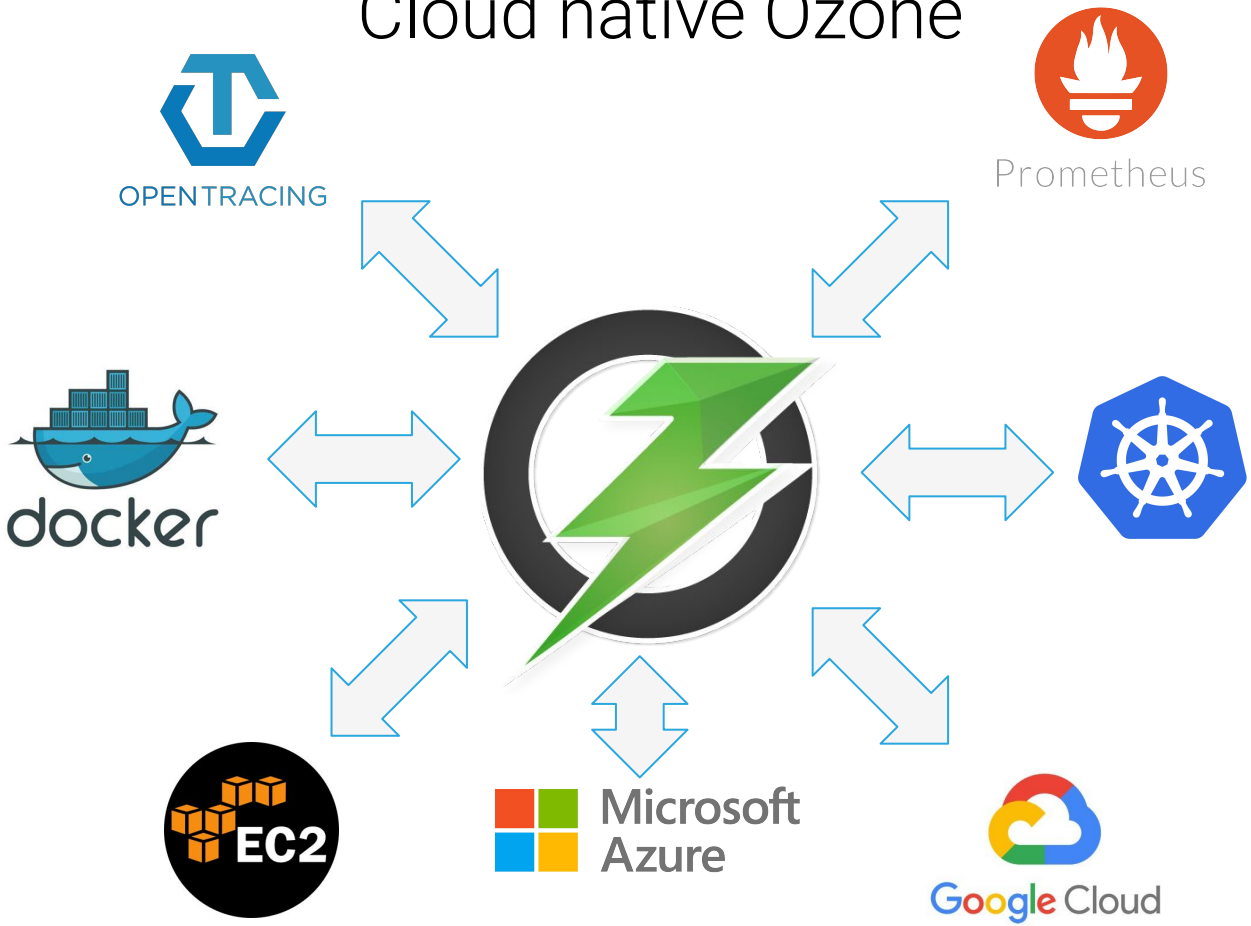
- tar xvzf + ./bin/hdfs start
- docker run apache/hadoop
- docker-compose up -d
- kubectl apply -f

...

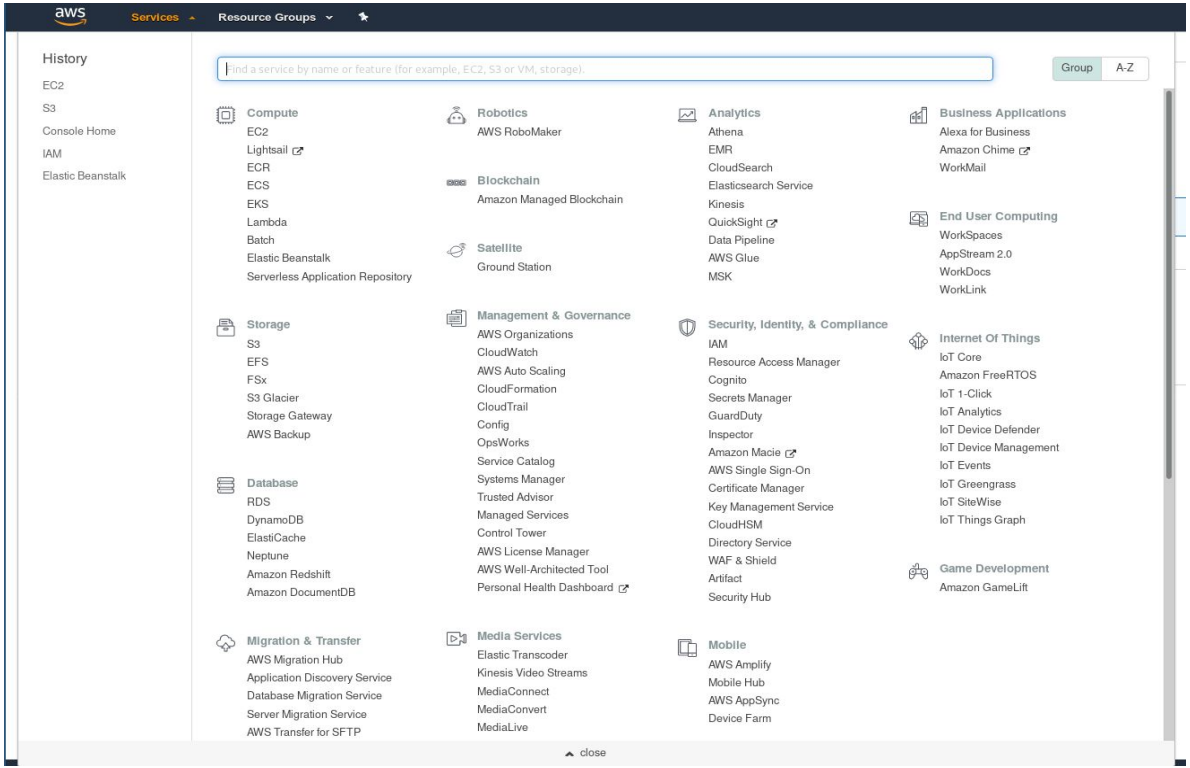
Make it easy to start anywhere!

cloudera

Cloud native Ozone



Cloud-Native (def2)



cloudera

Cloud-native is an User Experience!

Results

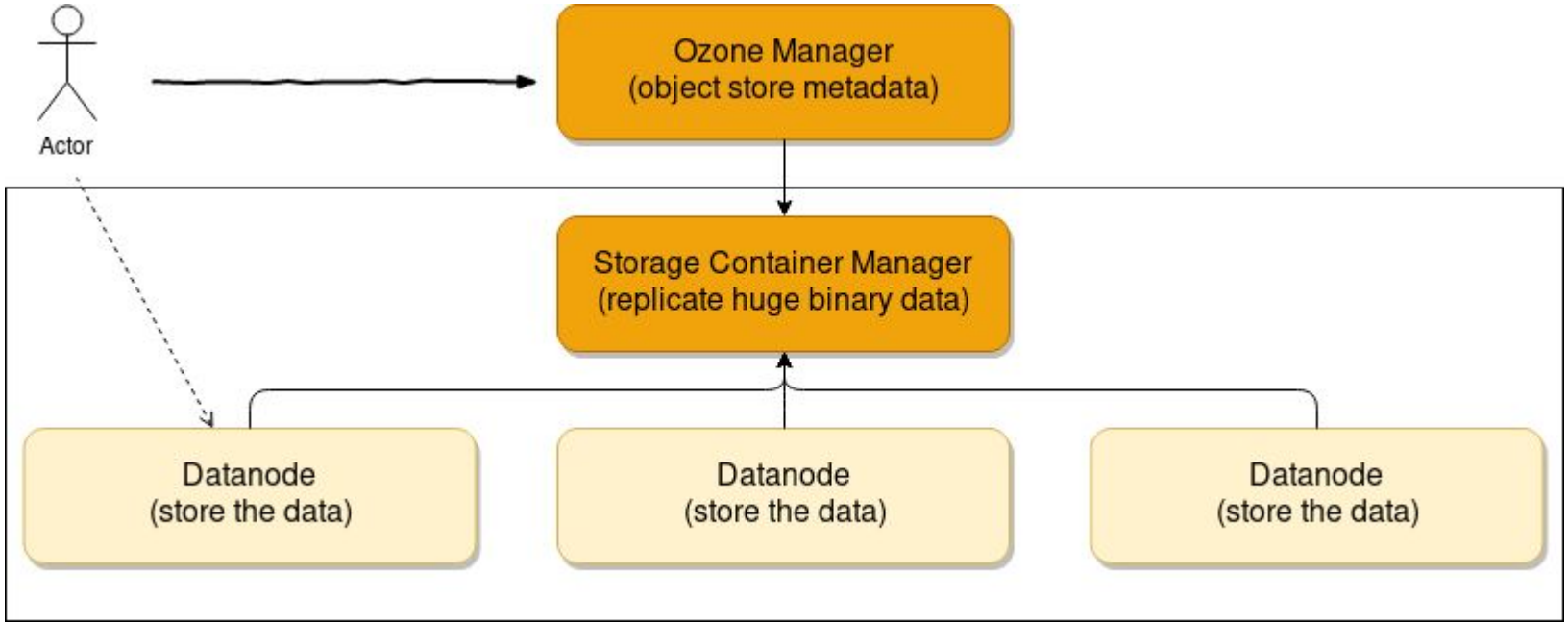
- Cloud-native user experience

- **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

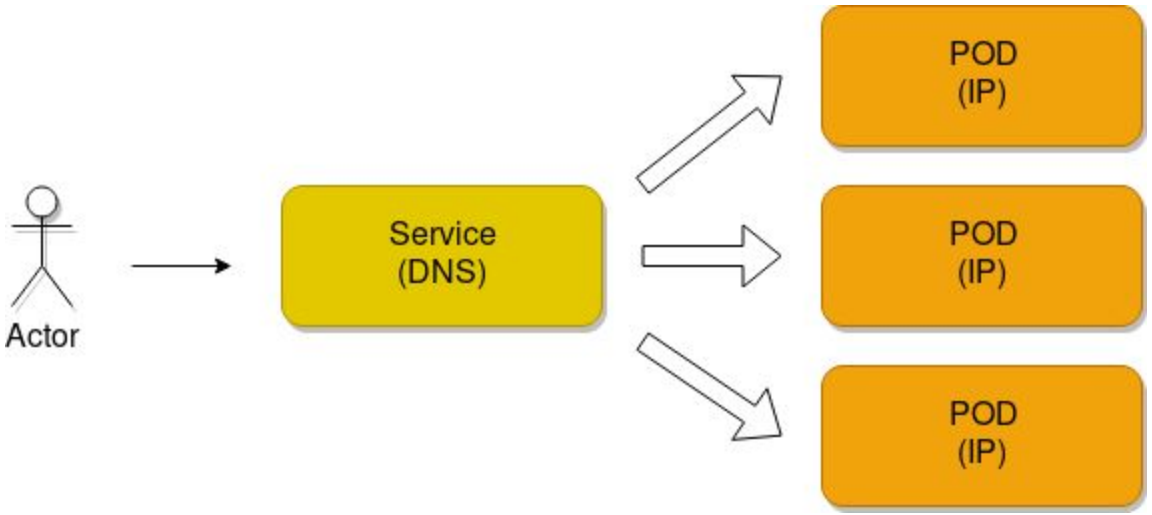
Network

Apache Hadoop Ozone

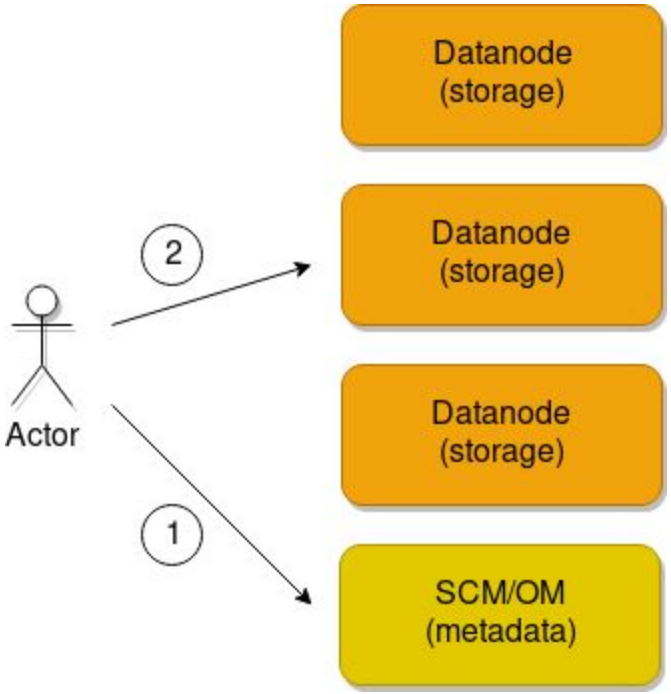


What about DNS?

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 certificats
  - Can be initialized with kerberos
  - Or with pre-generated certificate

Kubernetes security

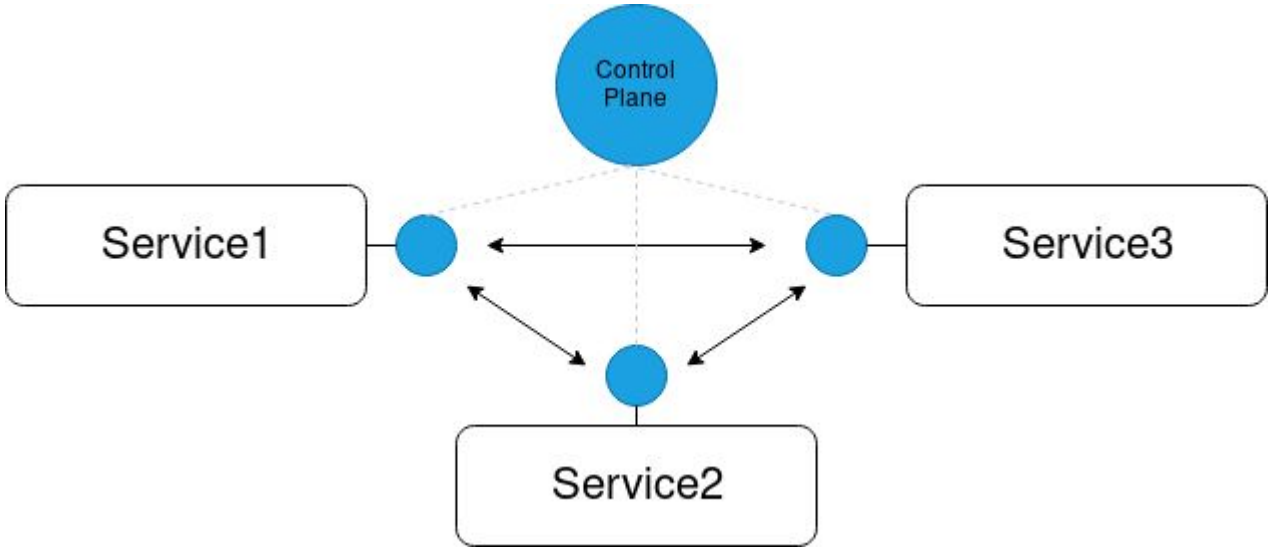
- RBAC for K8s API usage
- Secret management

Service Mesh

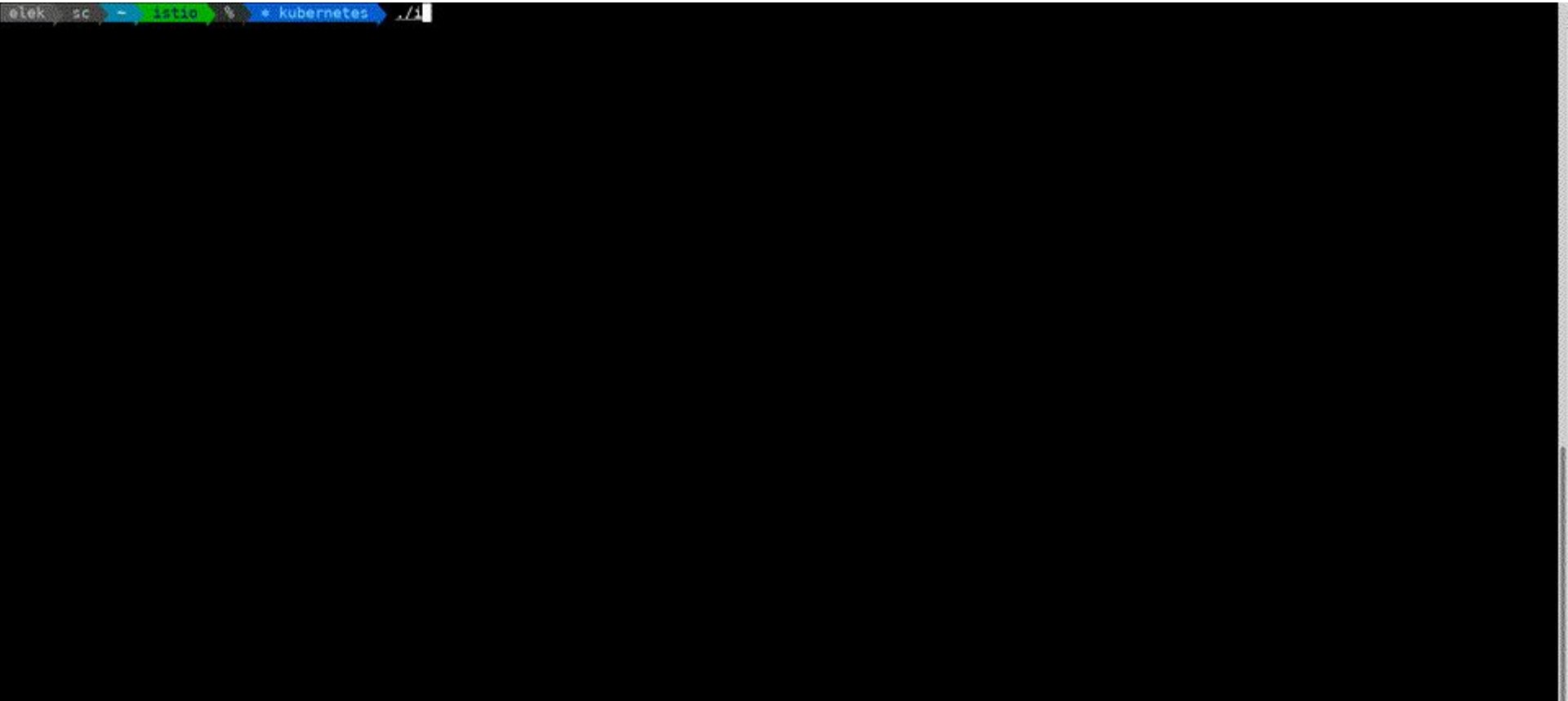
“The term service mesh is used to describe the network of microservices that make up such applications and the interactions between them.” (istio.io)

Explain Like I'm 5?

Centrally managed, component side proxies/reverse proxies



Nothing is free



To be or not to be?

Unified transport library (hadoop-rpc)

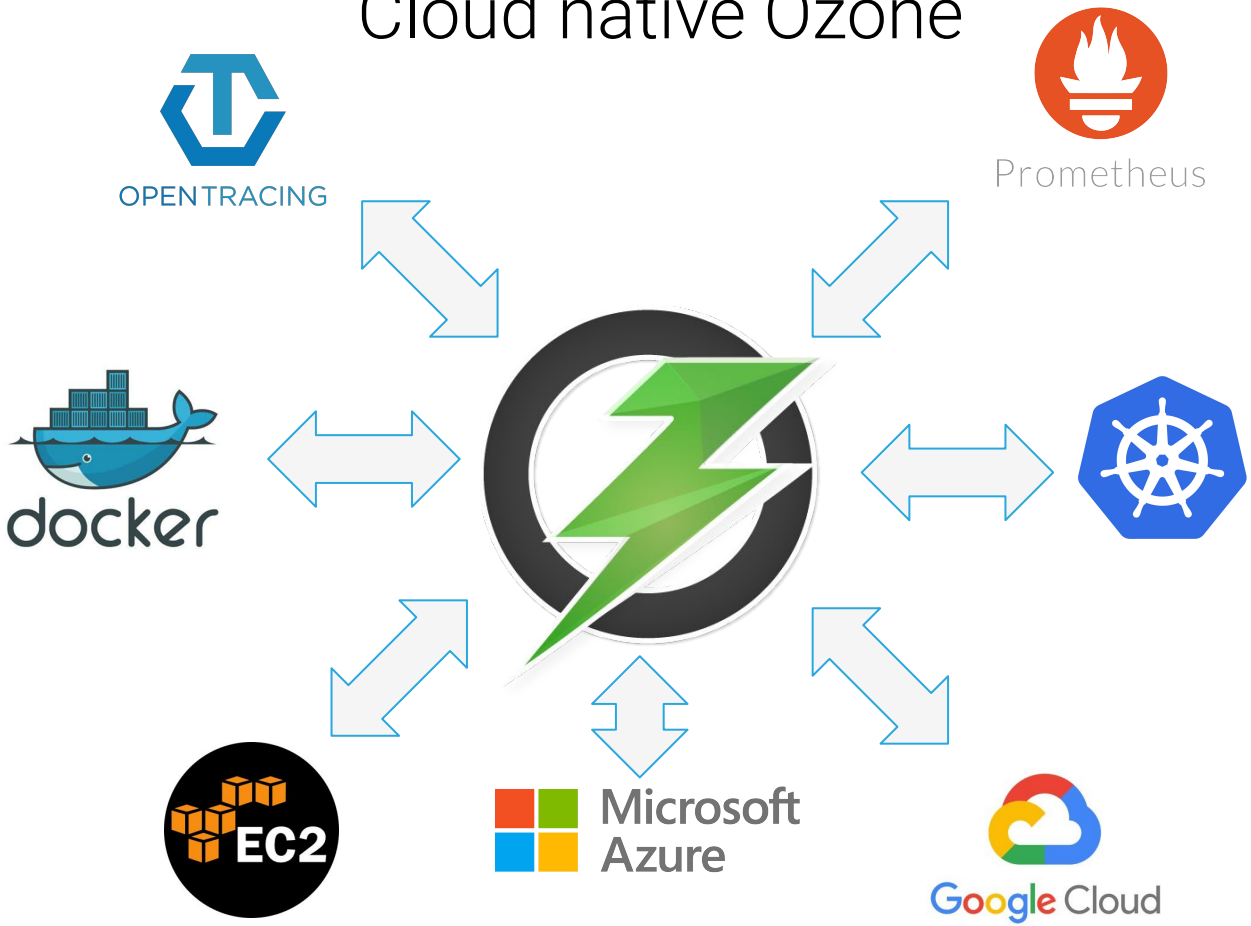
- Predefined network rules
- Advanced visibility
- Fine grained protocol support
- Security is included

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)*

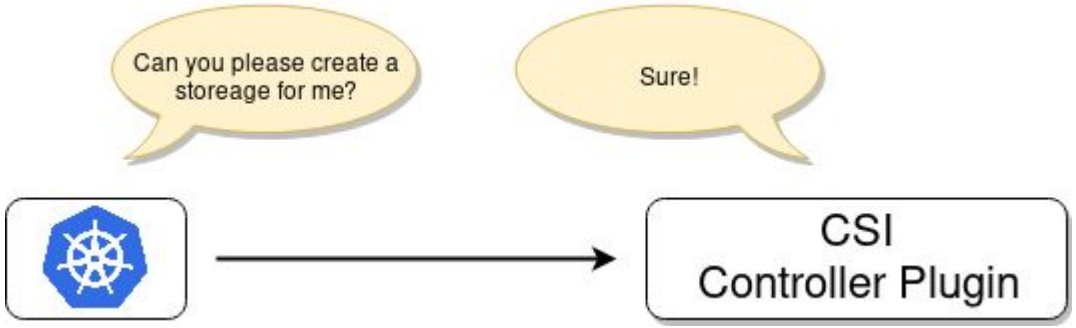
Cloud-native

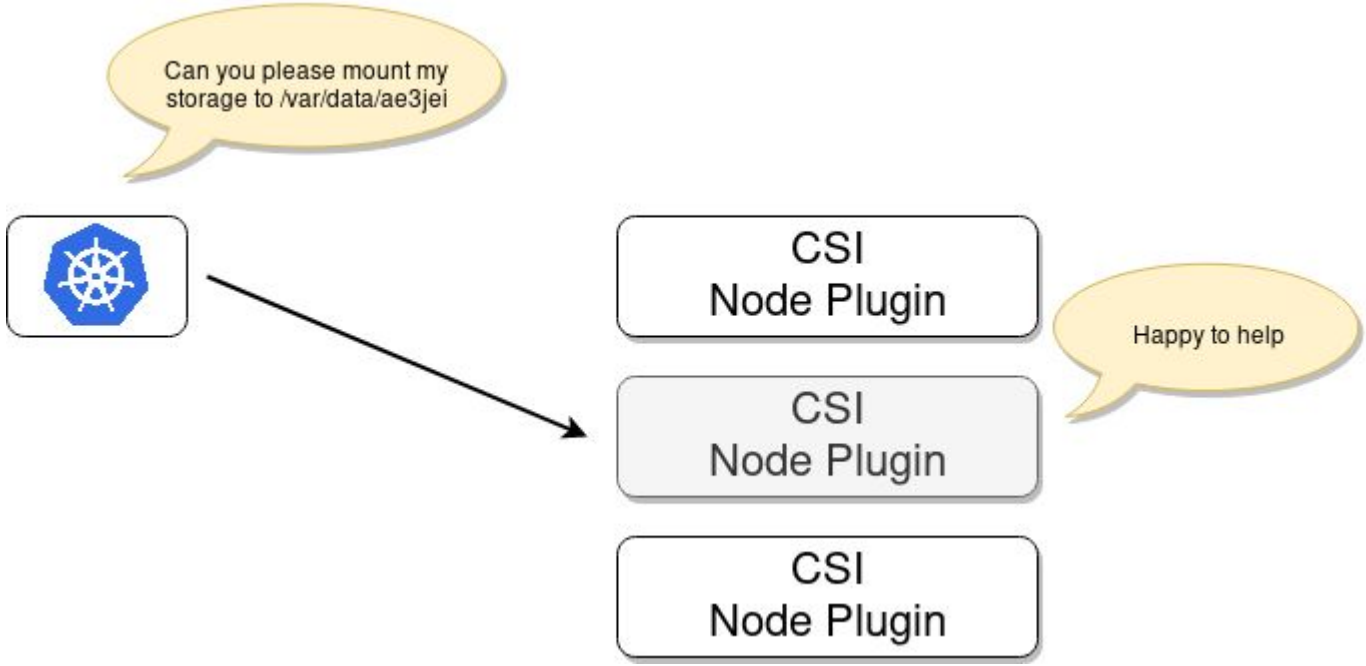
- Metrics
  - **Prometheus**
  - Visualization with other tools (***embedded Grafana***)
- Tracing
  - **OpenTracing**
  - OpenCensus
- ***Log collection***
  - Fluentd, ...

CSI

cloudera

Container Storage Interface: Vendor-neutral interface for volume management





cloudera

Results

CSI server is included (easy part)

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

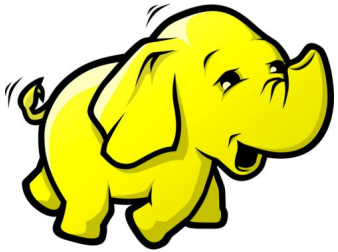
Hard part is to mount the file system (data path):

- 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

Cloud-Native world

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

aws S3 protocol

Hadoop FS

CSI

Apache Hadoop Ozone

hadoop.apache.org/ozone

THE APACHE SOFTWARE FOUNDATION

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