

RESILIENCE
REALIZED

Fluid Build Data Orchestration for Kubernetes



KubeCon



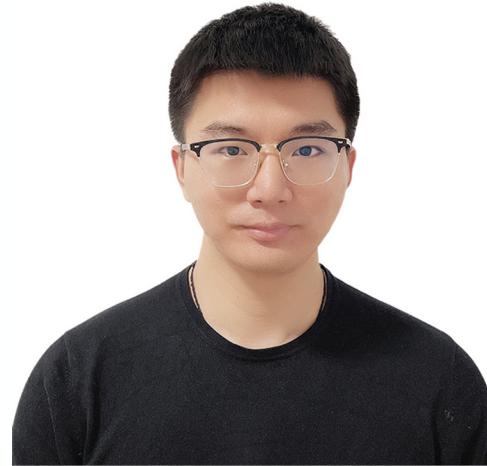
CloudNativeCon

North America 2021

About us



Yang Che
Fluid community co-founder
Staff Engineer of Alibaba Cloud



Yuandong Xie
Fluid core developer
Senior Engineer of Tencent Cloud

Agenda

- **Why Fluid**
- **What's Fluid**
- **Architecture**
- **Demo**
- **Community**

Data-intensive Applications Are Trending on Kubernetes

“By 2023, 70% of AI workloads will use Application container technology like Kubernetes” predicted by Gartner^[1]

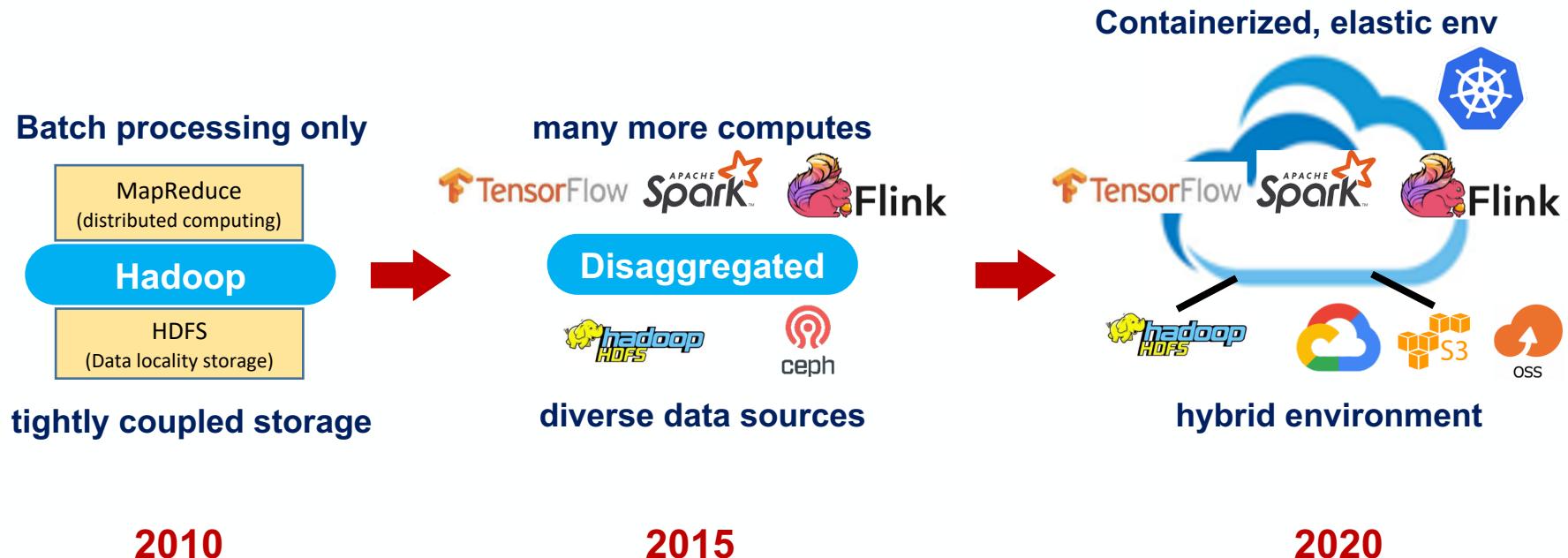
“Google Replaces YARN with Kubernetes to Schedule Apache Spark” at ApacheCon North America^[2]



[1] <https://www.gartner.com/en/conferences/emea/data-analytics-switzerland/featured-topics/topic-ai-machine-learning>

[2] <https://thenewstack.io/big-data-google-replaces-yarn-with-kubernetes-to-schedule-apache-spark/>

Journey of A Data-driven Company



Technical Challenges Blocking Faster Adoption



KubeCon



CloudNativeCon

North America 2021

Heterogeneous Data Source



Complexity

Disaggregated Compute & Storage



I/O Bottleneck

Locality-ignoring Scheduling

Computing Resource
(MR slots, Spark executors)

Memory Resource
(cache unit, data buffer)

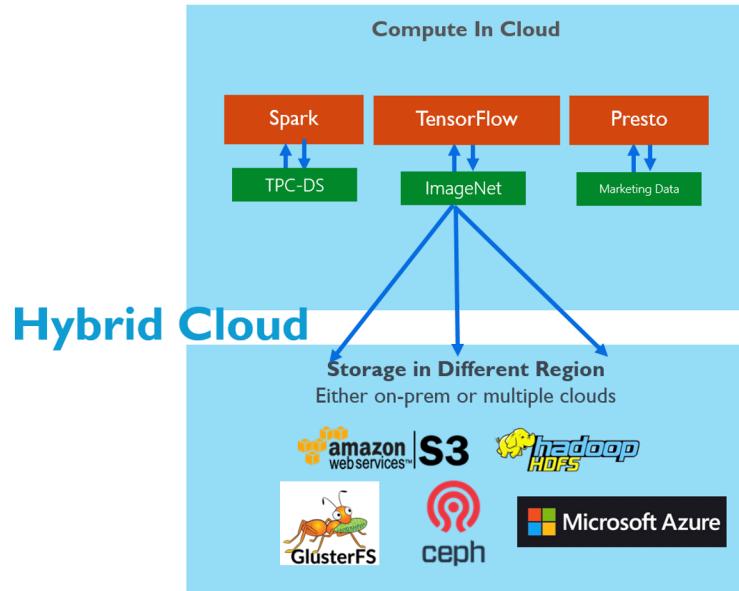
Static Resource Allocation



Inefficiency

1st Challenge: Complicated Data Storage

- **Heterogeneous storage:** HDFS, Ceph, S3 ...
- **Various data semantics:** files, objects, ...
- **Low-level data access APIs:** DFS, S3, POSIX, and complicated settings.



Fluid introduces “**DataSet**”, a high-level abstraction to applications, hiding details of heterogeneous data sources

2nd Challenge: I/O Bottleneck



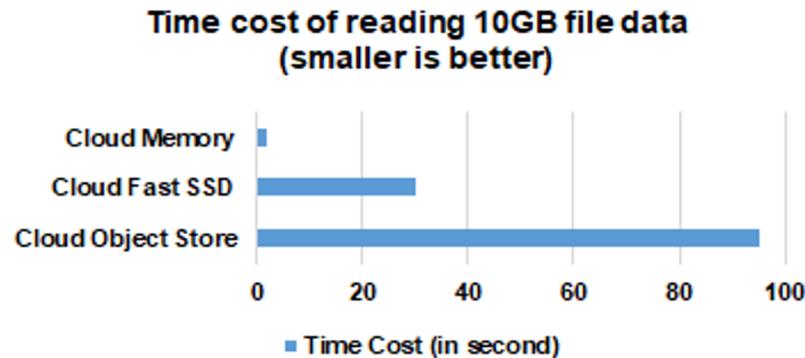
KubeCon



CloudNativeCon

North America 2021

- Disaggregated compute and storage: more I/O overhead
- Elastic infrastructure leads to more remote data access
- Concurrent data applications competing for limited data access bandwidth



Fluid accelerates data access with distributed cache runtimes

3rd Challenge: Locality-ignoring Scheduling



KubeCon



CloudNativeCon

North America 2021

- **Workload has no aware on scheduler:**

move data or move compute?

- **No intelligent data warmup:**

ETL is costly & error-prone

- **No data/app affinity-aware scheduling**

for apps like Spark, Tensorflow

Computing Resource
(MR slots, Spark executors)

Memory Resource
(cache unit, data buffer)

Static Resource Allocation

**Fluid builds data-aware scheduling strategies on
Kubernetes to coordinate applications with cache
runtime automatically**

What is Fluid, indeed

An Elastic Data Abstraction and Acceleration Platform in Cloud Native Environment.

Data Abstraction

for heterogeneous
data sources

Data Acceleration

with autoscaling and
portable cache runtimes

Data-aware Scheduling

to improve data affinity
for apps intelligently

We are a [Cloud Native Computing Foundation](#) sandbox project.



**CLOUD NATIVE
COMPUTING FOUNDATION**

Fluid Architecture

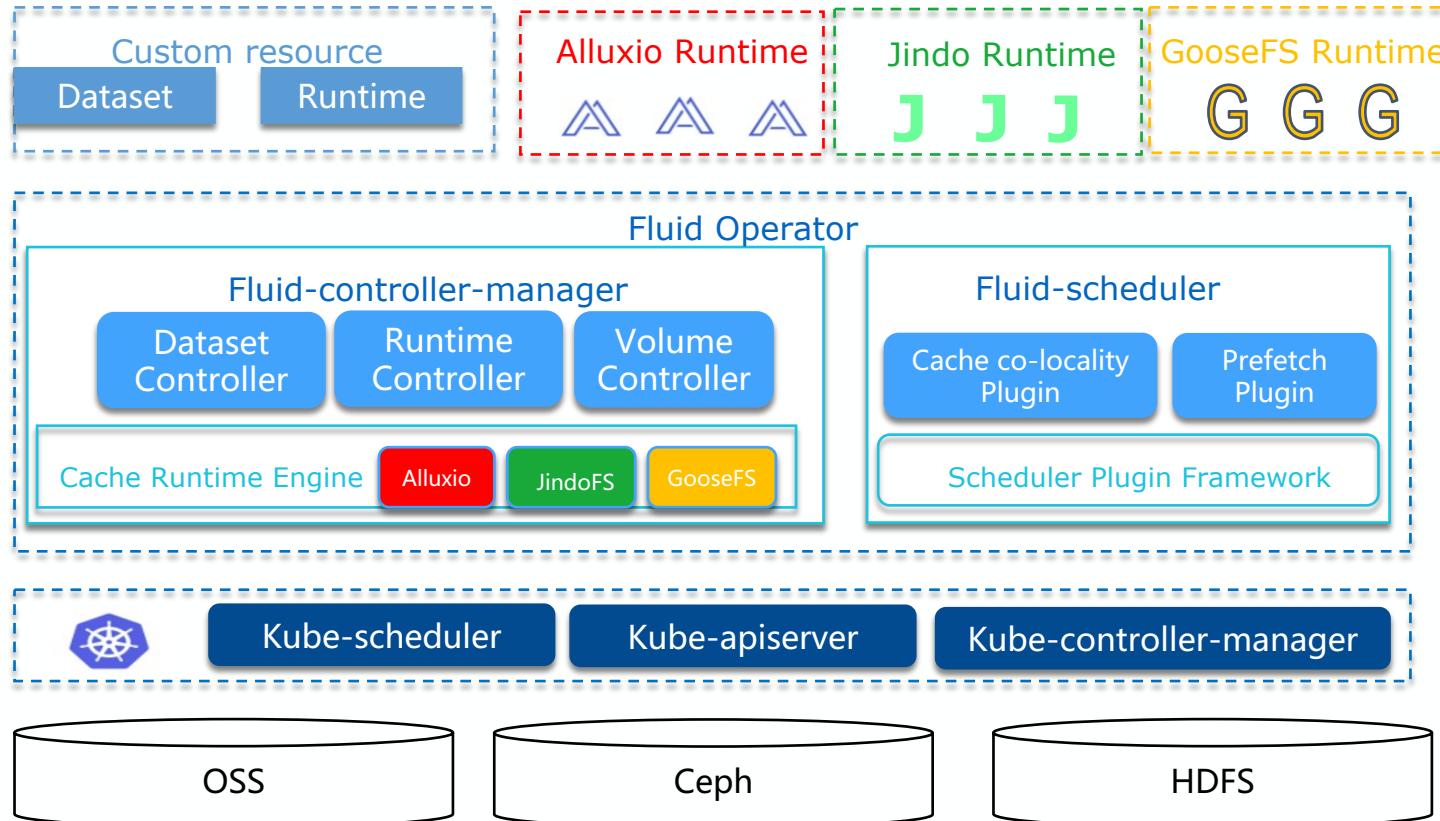


KubeCon



CloudNativeCon

North America 2021



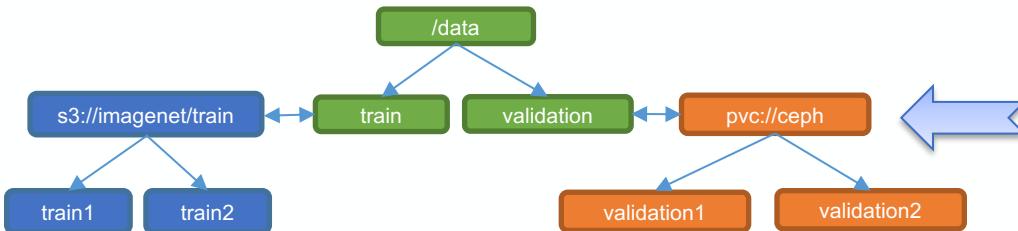
A Quick Start on Fluid: (1) Create a “Dataset”



CloudNativeCon

North America 2021

Edit “dataset.yaml” and run
“kubectl apply -f dataset.yaml”



```
apiVersion: data.fluid.io/v1alpha1
kind: Dataset
metadata:
  name: imagenet
spec:
  mounts:
    - mountPoint: s3://imagenet/train
      name: train
    - mountPoint: pvc://ceph
      name: validation
  nodeAffinity:
    required:
      - nodeSelectorTerms:
          - matchExpressions:
              - key: GPU
                operator: In
                values:
                  - "true"
```

s3

ceph

In GPU
nodes

```
[root@iZuf6afslw5j6hauyzkw21Z dataset]# kubectl get dataset
```

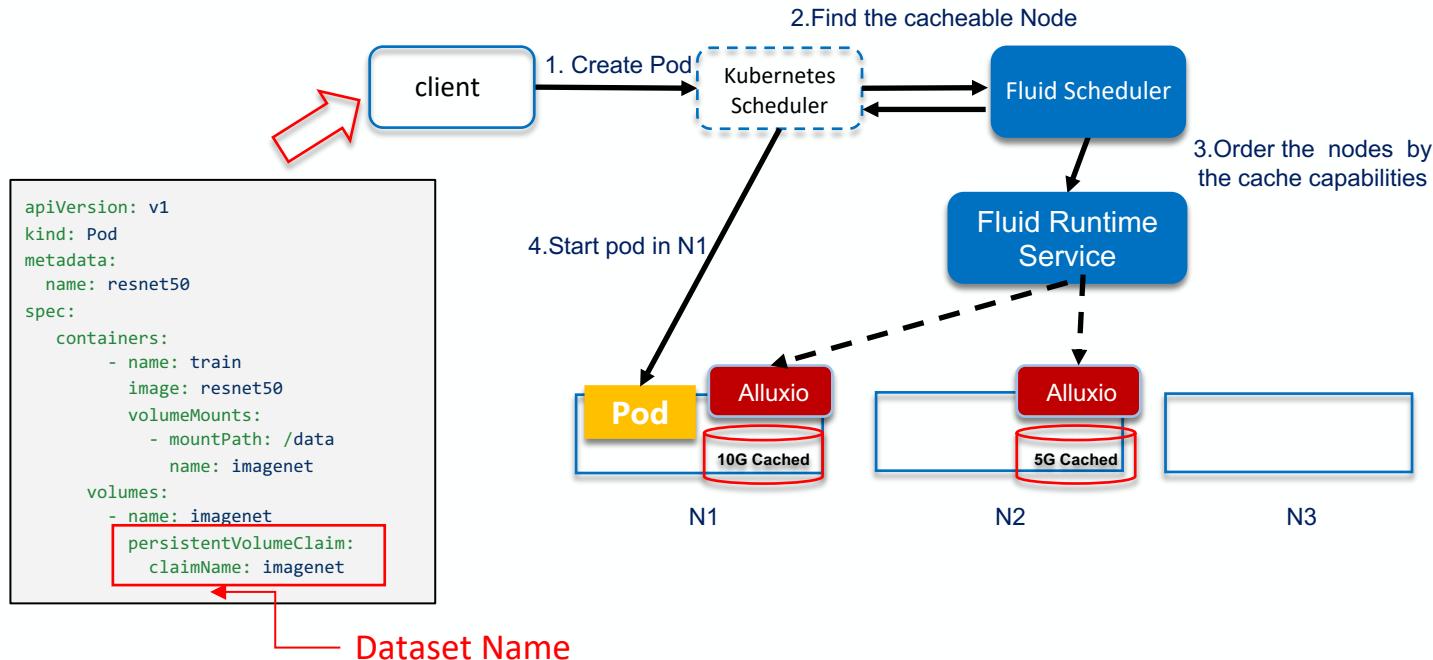
NAME	UFS TOTAL SIZE	CACHED	CACHE CAPACITY	CACHED PERCENTAGE	PHASE	AGE
imagenet	183.08GiB	0.00B	100.00GiB	0.0%	Bound	23m

A Quick Start on Fluid: (2) Run Apps with Data Affinity

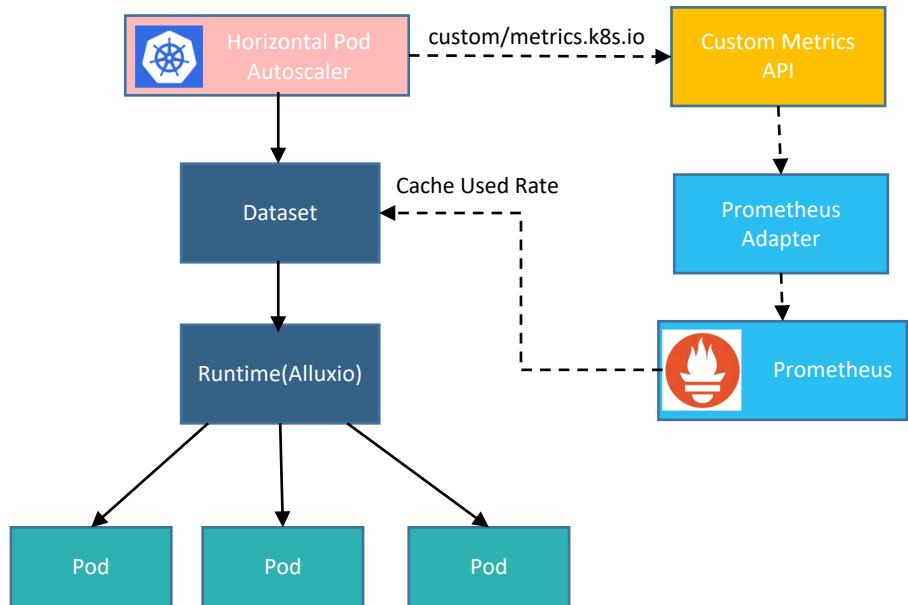


CloudNativeCon
North America 2021

Create “`pod.yaml`” and run “`kubectl create -f pod.yaml`”



Quick Start on Fluid: Auto-scale Dataset



```
apiVersion: autoscaling/v2beta2
kind: HorizontalPodAutoscaler
metadata:
  name: imagenet-hpa
spec:
  scaleTargetRef:
    apiVersion: data.fluid.io/v1alpha1
    kind: AlluxioRuntime
    name: imagenet
  minReplicas: 2
  maxReplicas: 10
  metrics:
  - type: Object
    object:
      metric:
        name: capacity_used_rate
    describedObject:
      apiVersion: data.fluid.io/v1alpha1
      kind: Dataset
      name: imagenet
  target:
    type: Value
    value: "75"
```

Accelerate machine learning training job with Fluid

```
[root@iZuf60ig21389wrykbcmrzz demo]#
```

Demo: Accelerate machine learning training with Fluid

Roadmap

Fluid 0.6

- Accelerate DL on Kubernetes in a simple way
- Scale out/in Dataset
- A small set of datasets

Use case

As a data infrastructure engineer, I want to be able to setup dataset cache for AI immediately that can be easily and efficiently accessed.

Fluid 1.0

- Enable big data on Kubernetes
- Comprehensive schedule capability for data and workload
- Dataset lifecycle management in large scale of Kubernetes

Use case

As a data infrastructure engineer, I want to run both AI and Big Data in the same K8s with data acceleration. And I don't want to take care of the data management.

Plan

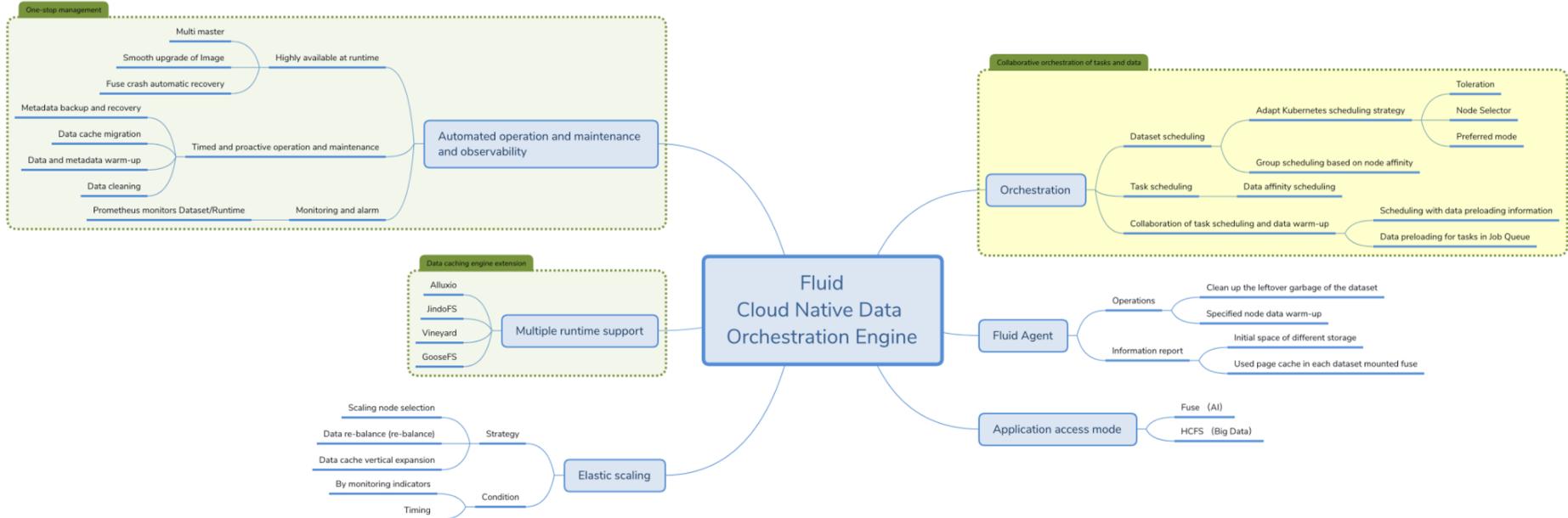


KubeCon



CloudNativeCon

North America 2021



Community & Join Us

- **Contributors:** 80+ contributors, incl.
 - Alibaba Cloud
 - NanJing University
 - Alluxio
 - China Telecom
 - Tencent Cloud
 - 4 Paradigm
 - ...
- **Github:** <https://github.com/fluid-cloudnative/fluid>
- **Slack channel:** <https://cloud-native.slack.com/> #fluid
- **Bi-weekly community meeting:** <https://us05web.zoom.us/j/84070287987?pwd=M0UyNmhlREFxZFZGSFBDWlhHY1N2dz09>
- **Documentation:** <https://fluid-cloudnative.github.io/guide/>
- **Website:** <https://fluid-cloudnative.github.io/>

- Who's adopting Fluid in their own products?





KubeCon



CloudNativeCon

North America 2021

RESILIENCE
REALIZED

Q&A - Thank you !