

Data Mobility for Kubernetes Persistent Volumes



Xing Yang, Principal Architect, Huawei

@2000Xyang

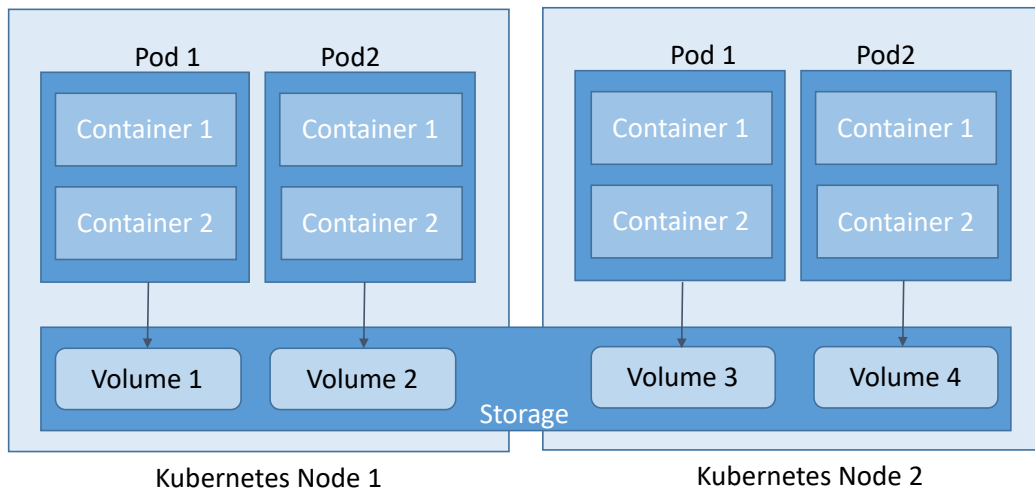
Agenda

- Kubernetes Persistent Volumes and Snapshots
- Container Storage Interface (CSI)
- Why Data Mobility - Use Cases
- Policy Based Data Mobility using OpenSDS
- Multi-Cloud Data Control Architecture
- Demo
- OpenSDS Community
- OpenSDS Roadmap
- Getting Involved

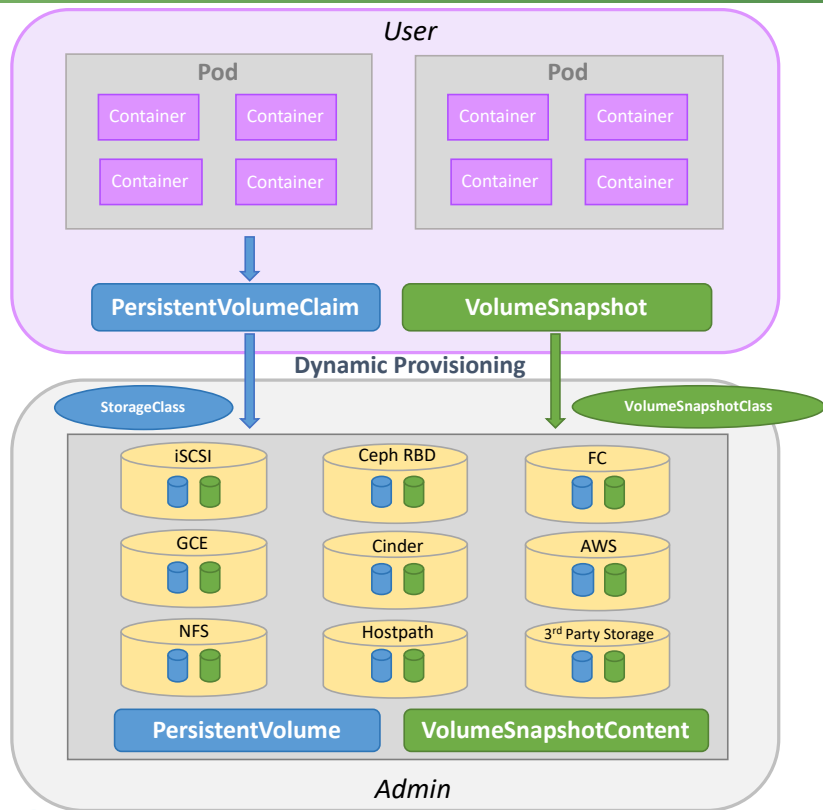
Kubernetes Persistent Volumes and Snapshots

Kubernetes Persistent Volumes

- Persistent Volume
- Persistent Volume Claim
- Storage Class
- Dynamic and Static Provisioning

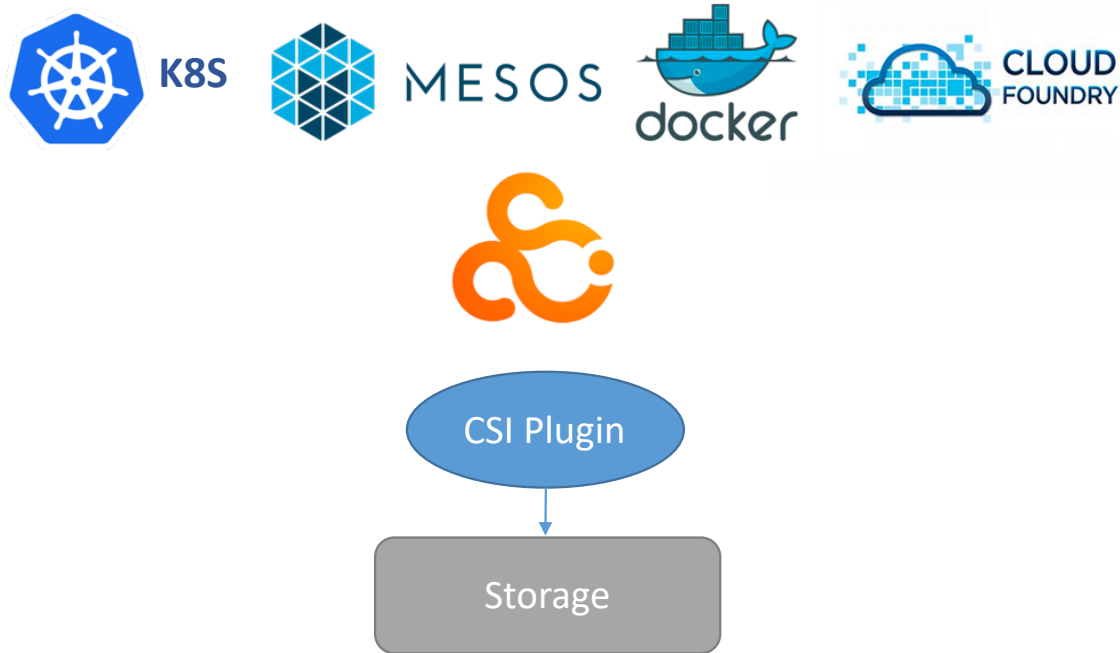


Kubernetes Volume Snapshots

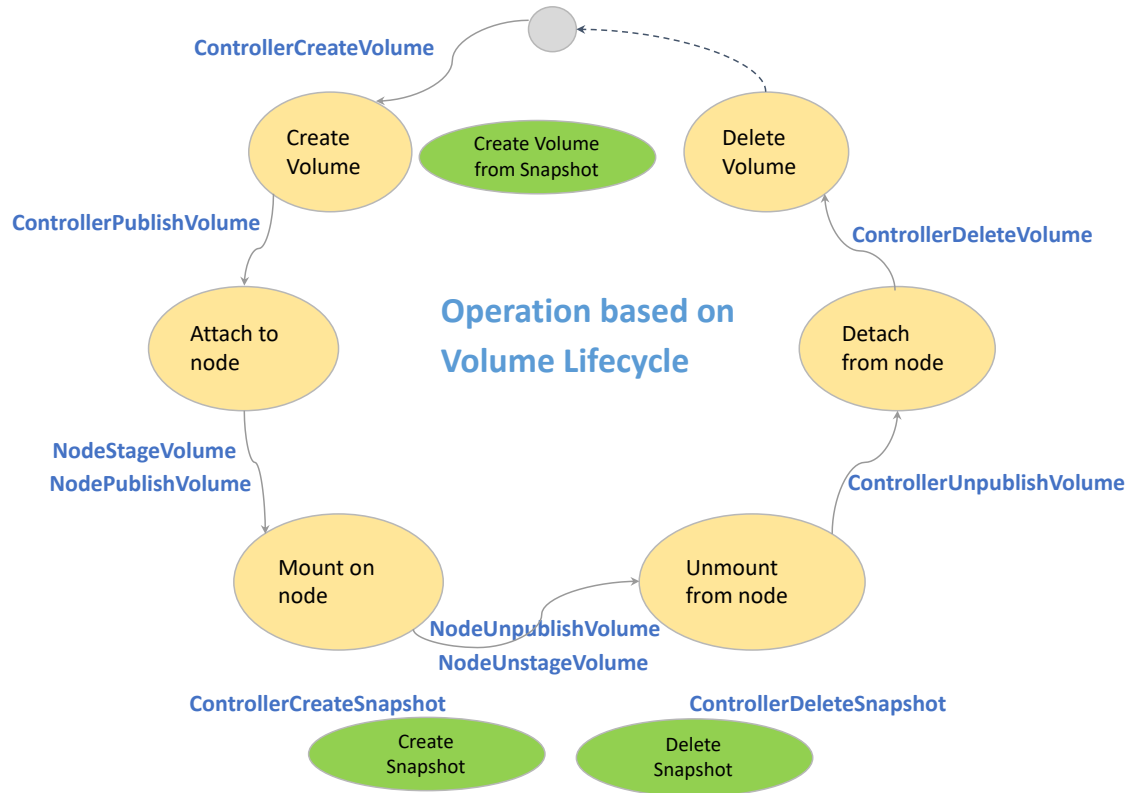


- Volume Snapshot
- Volume Snapshot Content
- Volume Snapshot Class
- Dynamic and Static Provisioning
- Supports CSI Drivers

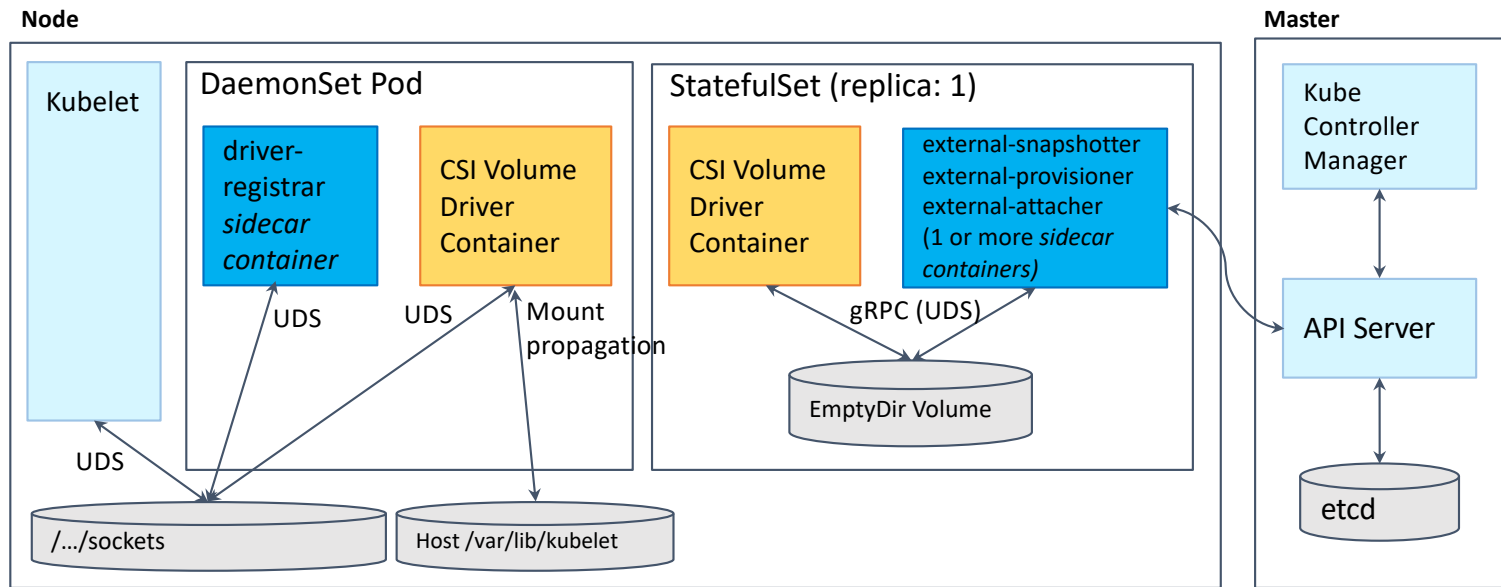
Container Storage Interface (CSI)



CSI Functions



Deploy CSI Plugin on Kubernetes



Source: <https://github.com/kubernetes/community/blob/master/contributors/design-proposals/storage/container-storage-interface.md>

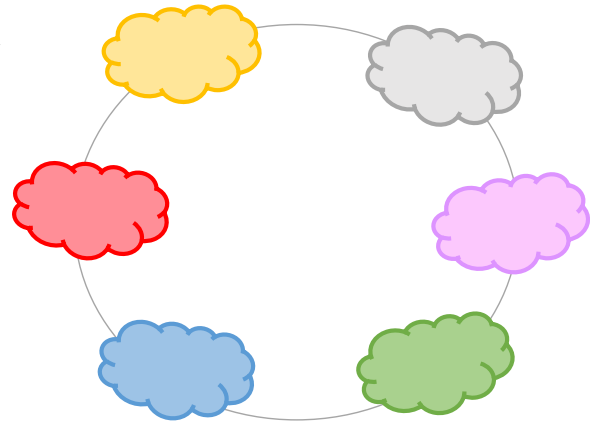
Why Data Mobility

Why do we need Data Mobility

- Data mobility allows data to be accessible any time, any where, on premise or in the cloud.
- Use cases for data mobility by OpenSDS multi-cloud data control
 - High availability
 - Data analytics
 - Data lifecycle management
 -

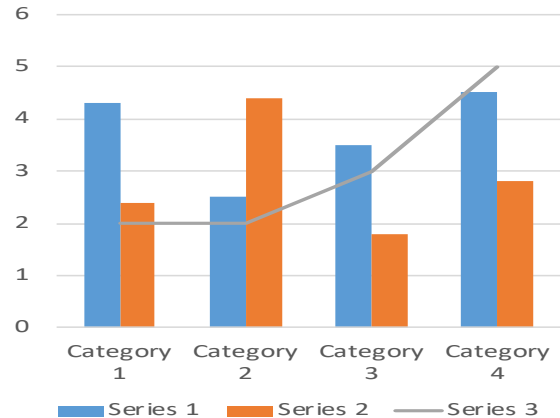
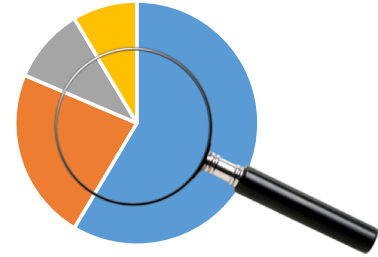
Use Case: High Availability

- Make data available at different geographic location.
- Make data available and accessible all the time without downtime.



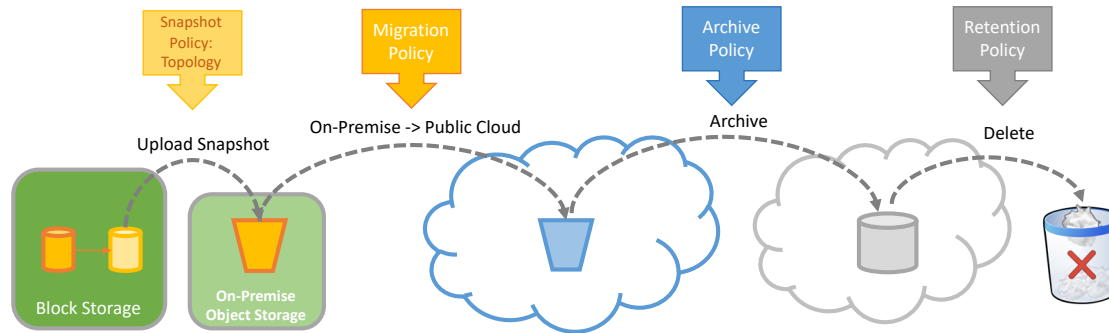
Use Case: Data Analytics

- Upload datasets to public cloud for data analytics.
- Utilize compute resources in public cloud to do CPU intensive work.



Use Case: Data Lifecycle Management

- Frequently accessed data stored on-premise.
- Data not accessed for a long time can be stored in archival storage in public cloud.
- Data can be moved to different location depending on usage, age, access frequency.
- Data can be deleted automatically based on policy.



Policy-based Data Mobility using OpenSDS

OpenSDS Projects – Sushi and Hotpot

SUSHI

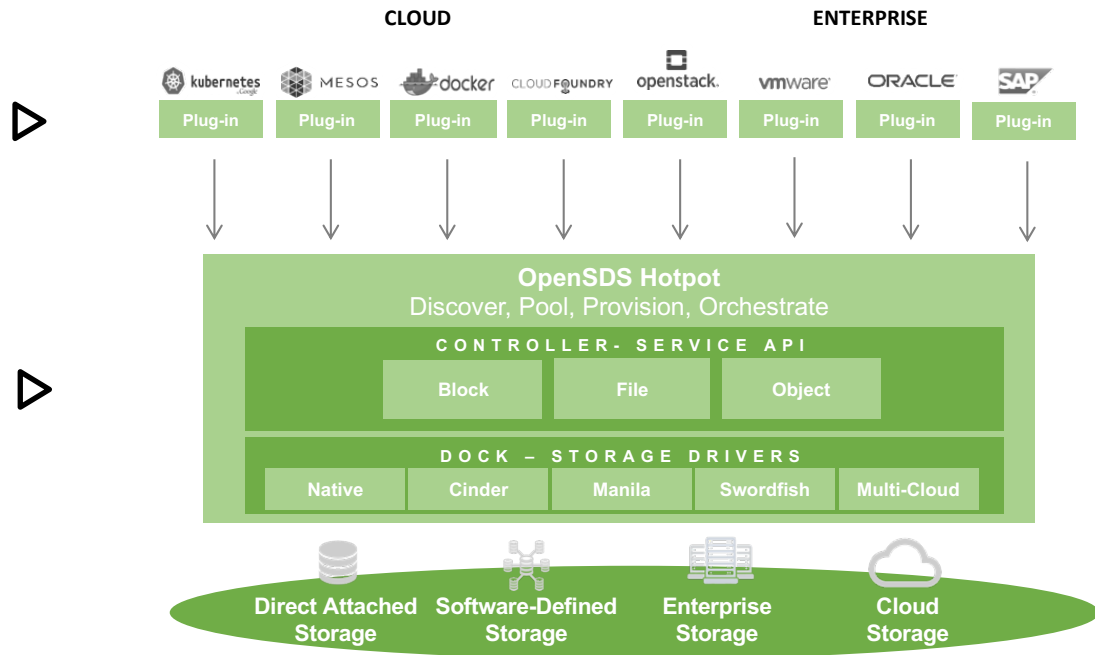
The Northbound Plug-ins Project

Common plug-ins to enable OpenSDS storage services for cloud and application frameworks

HOTPOT

The Storage Controller Project

Single control for block, file, and object services across storage on premise and in clouds

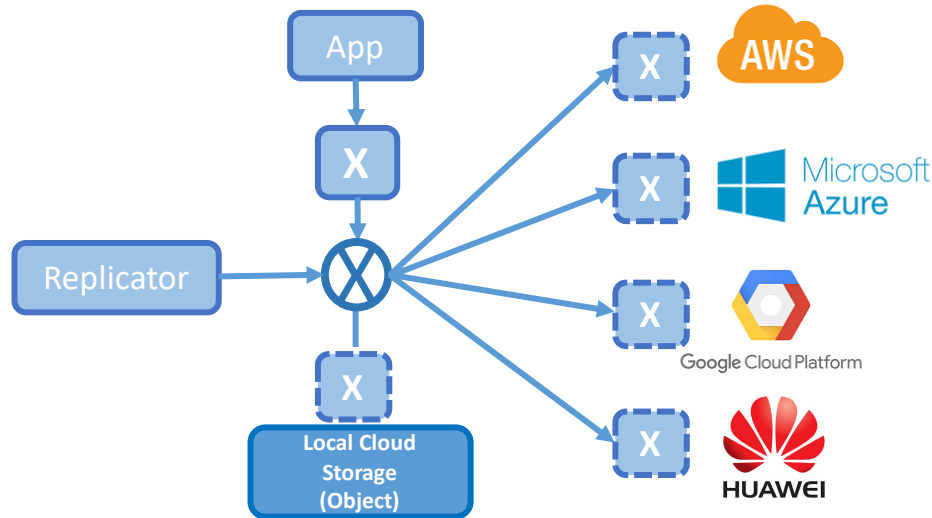


OpenSDS Project - Gelato

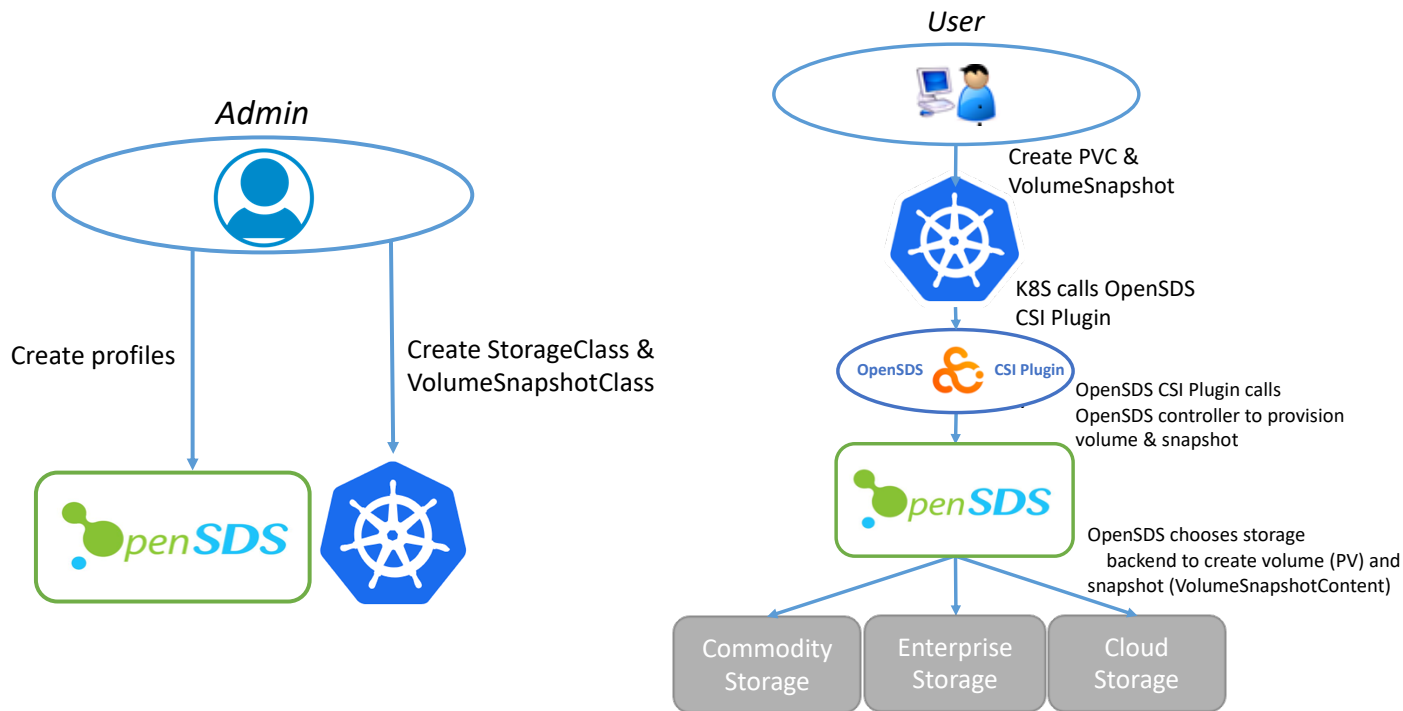
Gelato

The Multi-Cloud Project

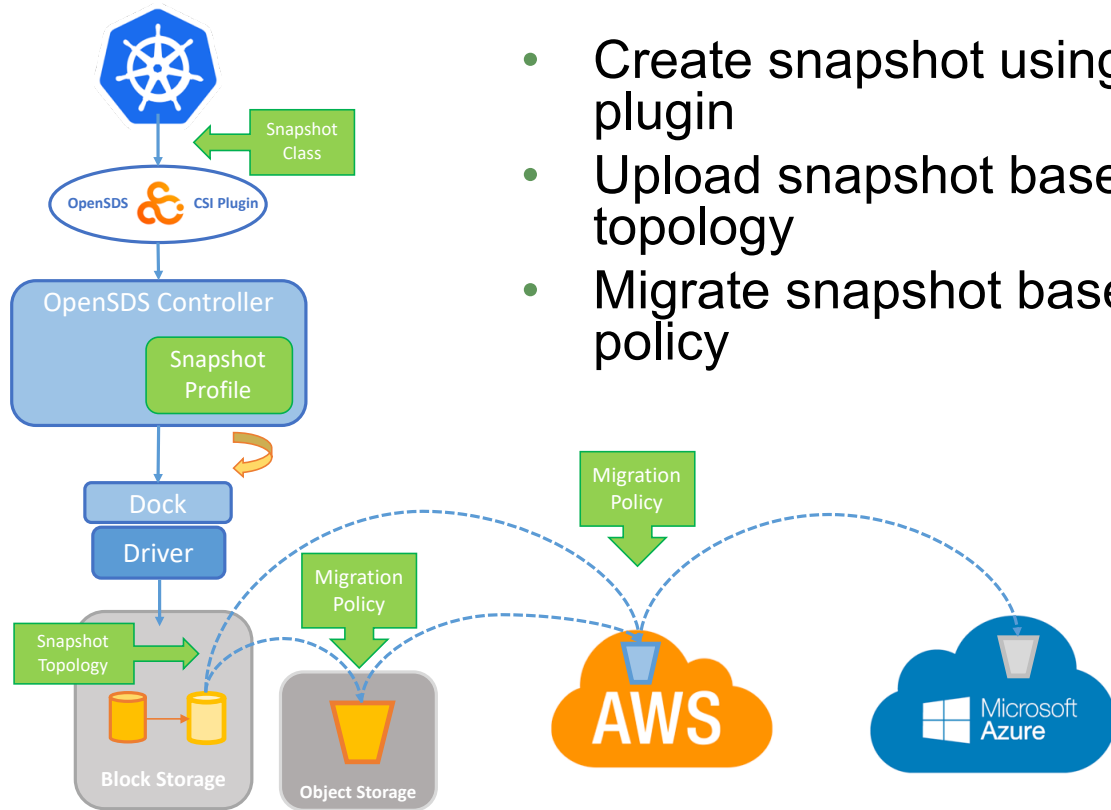
Policy based multi-cloud data control
to enable data mobility across clouds



Profiles and StorageClass/VolumeSnapshotClass



CSI Plugin and Policy-based Data Mobility



- Create snapshot using OpenSDS CSI plugin
- Upload snapshot based on snapshot topology
- Migrate snapshot based on migration policy

Volume Snapshot Example

```
apiVersion: snapshot.storage.k8s.io/v1alpha1
kind: VolumeSnapshotClass
metadata:
  name: csi-opensds-snapclass
snapshotter: csi-opensdsplugin
parameters:
  profile: upload-snapshot-to-cloud
```

Note: Profile can be uuid or name of the profile in OpenSDS.

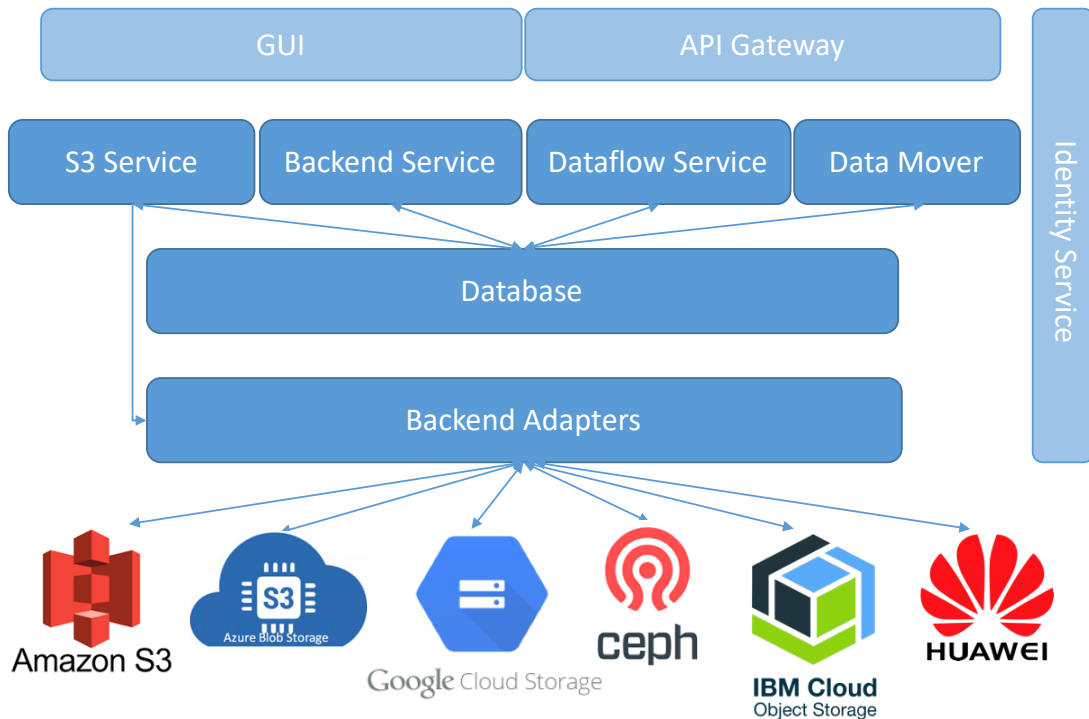
```
apiVersion: snapshot.storage.k8s.io/v1alpha1
kind: VolumeSnapshot
metadata:
  name: new-snapshot-demo
  namespace: demo-namespace
spec:
  snapshotClassName: csi-opensds-snapclass
  source:
    name: opensdspvc
    kind: PersistentVolumeClaim
```

Create Snapshot

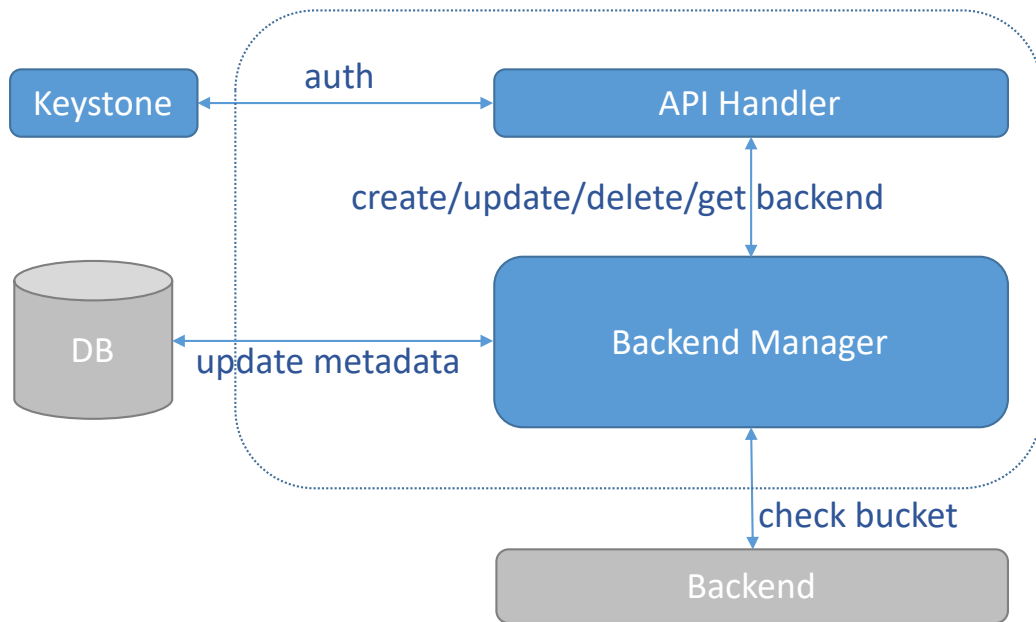


Multi-Cloud Data Control Architecture

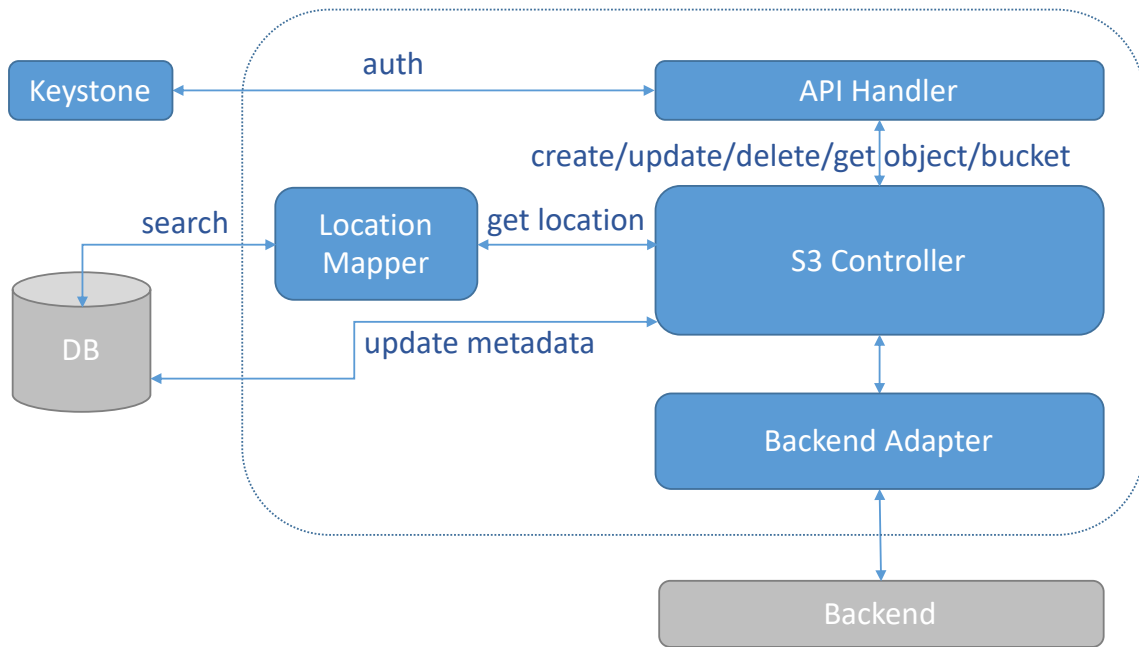
Multi-Cloud Data Control Architecture



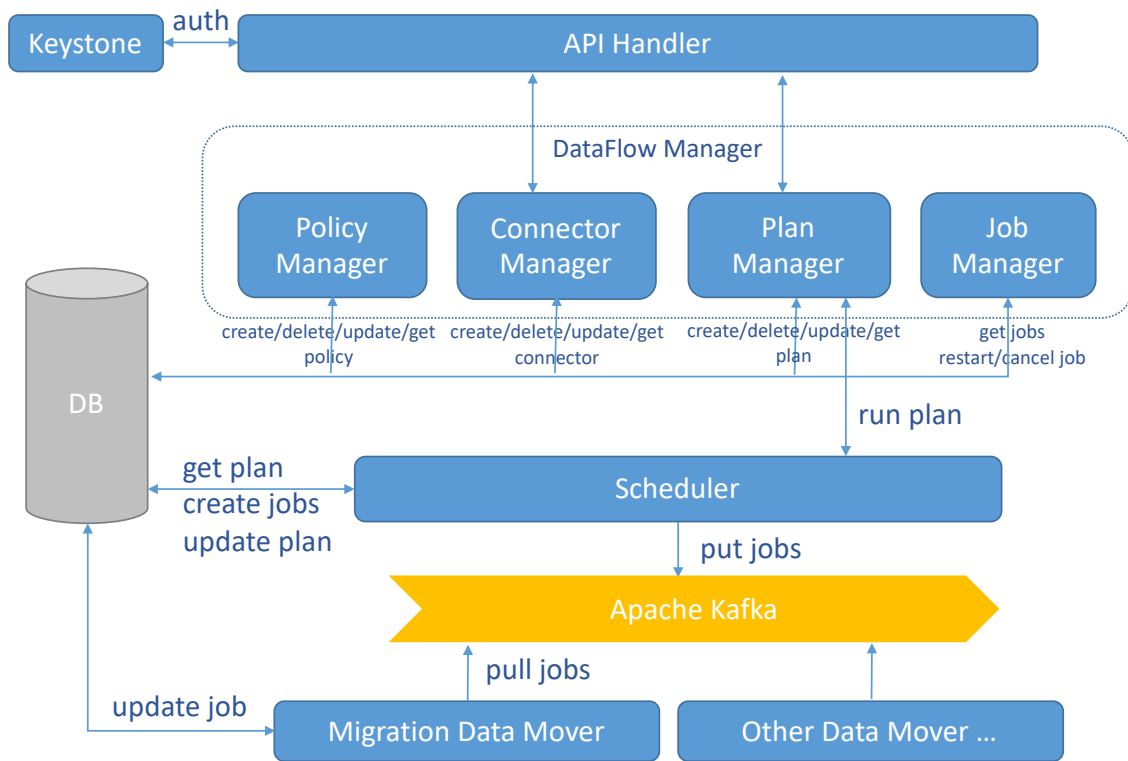
Multi-Cloud Backend Service



Multi-Cloud S3 Service



Multi-Cloud Dataflow Service



Multi-Cloud Dashboard



Multi-Cloud Dashboard – Register Backend

The screenshot displays the Multi-Cloud Dashboard interface. On the left, a sidebar contains navigation links: **admin** (default_region), **Home** (Resource statistics), **Resource** (Volumes / Buckets), **Dataflow** (Through migration / replication capability), **Profile** (Block profiles), **Infrastructure** (Regions, availability zones and storage), and **Identity** (Managing tenants and users). The main dashboard area features a central visualization with clouds labeled 'aws' and 'OpenSDS', and a 'Dataflow Capacity' line graph showing an upward trend from 0 to 25. At the bottom left, a 'Resource' summary shows 5 Volumes, 2 Buckets, and 0 Filesystems. A modal window titled 'Register Storage Backend' is open on the right, allowing the user to select a storage type (Object, File, Block) and enter details: Name, Type (Please select), Region (Please select), Endpoint, Bucket (Bucket name), Access Key, and Secret Key. The OpenSDS logo is visible in the bottom left corner of the dashboard.

Multi-Cloud Dashboard - Migration

The screenshot displays the 'Multi-Cloud Dashboard - Migration' interface. On the left is a dark sidebar with navigation links: 'admin', 'Home', 'Resource', 'Dataflow', 'Profile', 'Infrastructure', and 'Identity'. The main area is titled 'Data flow' and shows 'Migrations(0)' and 'Replications(1)'. A 'Create' button is present. A modal window titled 'Migrate' is open, containing the following fields and options:

- Basic Information**
 - Plan Name :
 - Source Bucket :
 - Destination Bucket :
 - Execution Time : ☒ Immediately
- Rule**
 - Without a rule, all objects in the source bucket will be migrated; otherwise only objects selected based on the rule will be migrated.
 -
 - ☐ Config Data analysis
 - ☐ Delete the source objects after the migration is completed.

The modal has 'OK' and 'Cancel' buttons at the bottom. The background shows a diagram of multi-cloud architecture with AWS and Huawei Cloud icons.

Demo

- Take snapshot using OpenSDS CSI plugin
- Snapshot uploaded to AWS S3
- Migrate data from AWS S3 to Huawei Cloud object store

OpenSDS Roadmap

2017H2

ZEALAND

Storage For Kubernetes

- Kubernetes FlexVolume
- Vol CRUD
- Standalone Cinder Integration
- CSI Support
- Ceph, LVM

2018H1

ARUBA

Storage Orchestration

- OpenStack
- Replication
Array-Based,
Host-Based
- Dashboard
- Virtual Pools
- Storage Profiles
- NVMeoF preview
- Enumeration
- Block Storage
 - Ceph
 - LVM
 - IBM: XIV, Storwize, SVC
 - Huawei: Dorado

2018H2

BALI

Storage Multi-Cloud

- Data Migration
Offline, Online*
- Monitoring
- Multi-OpenStack
- S3 Object
- Multi-Cloud Control
- NVMeoF
- Storage Groups
Snapshots, Replication
- CSI
Mesos*, Docker*
- Swordfish
Dell-EMC, NetApp

2019H1

CAPRI

Storage Intelligence

- Analytics
- Lifecycle
- Data Protection
- File Share

2019H2++

- Performance
- Optimization
- Tiering
- Security
- Sharing
- Networking
- SCM

OpenSDS Community

Technical Steering Committee



Steven Tan, Chairman
Huawei, VP & CTO Cloud Storage Solution



Rakesh Jain, Vice-Chair
IBM, Research Engineer and Architect



Allen Samuels
Western Digital, R&D Engineering Fellow



Anjaneya "Reddy" Chagam
Intel, Chief SDS Architect



Jay Bryant
Lenovo, Cloud Storage Lead

End-User Advisory Committee



Cosimo Rossetti
Vodafone, Lead Storage Architect



Yusuke Sato
Yahoo Japan, Infrastructure Lead



Kei Kusunoki
NTT Communications, Storage Architect



Yuji Yazawa
Toyota ITC, Group Lead



Wim Jacobs
KPN, Senior Architect



Getting Involved

- Repos: <https://github.com/opensds>
- Slack: <https://opensds.slack.com>
- Mailing list: <https://lists.opensds.io>
- Weekly meetings:
<https://github.com/opensds/design-specs/blob/master/README.md#opensds-technical-meetings>



OPEN SOURCE SUMMIT

EUROPE

THE LINUX FOUNDATION