



KubeCon



CloudNativeCon

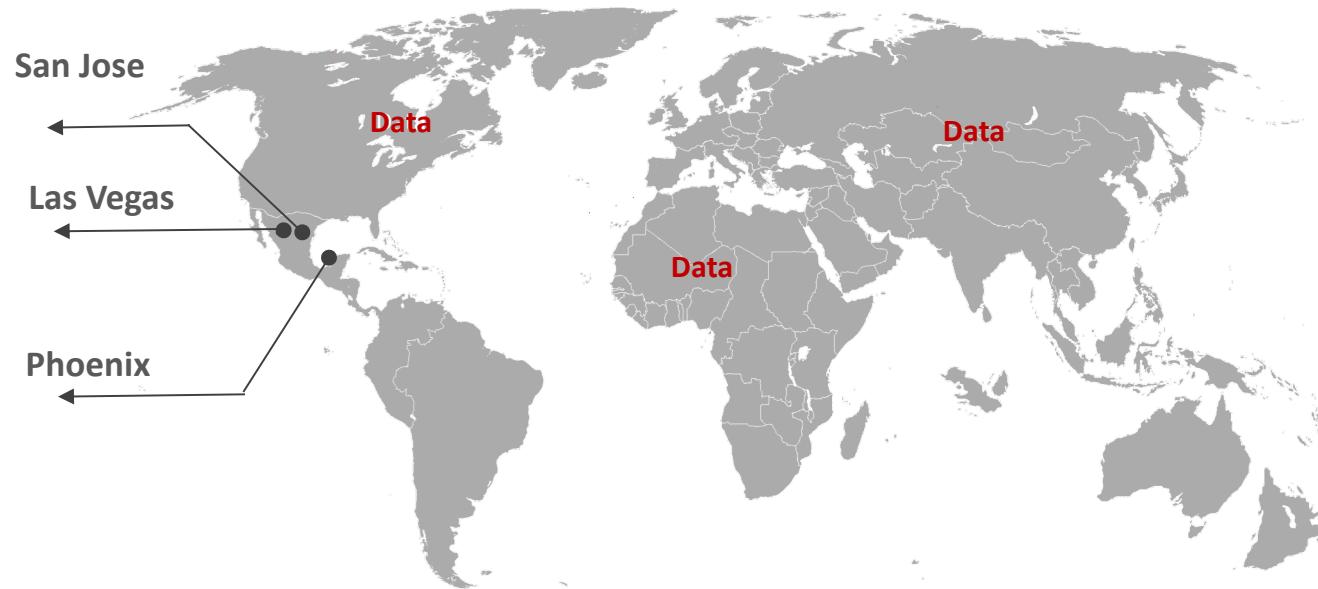
China 2018

Operator Framework to manage Stateful Workloads in eBay

AGENDA

- Background
- Introduction
- Features
- Conclusion

BACKGROUND



Data Platform

eBay has the data platform and shared data ecosystem to provide the off-line and real-time data that power eBay's vision of a buyer experience, seller insights and agile, data driven business to serve the ebay employees all around the world.

C 190+
COUNTRY

D 600PB
DATA



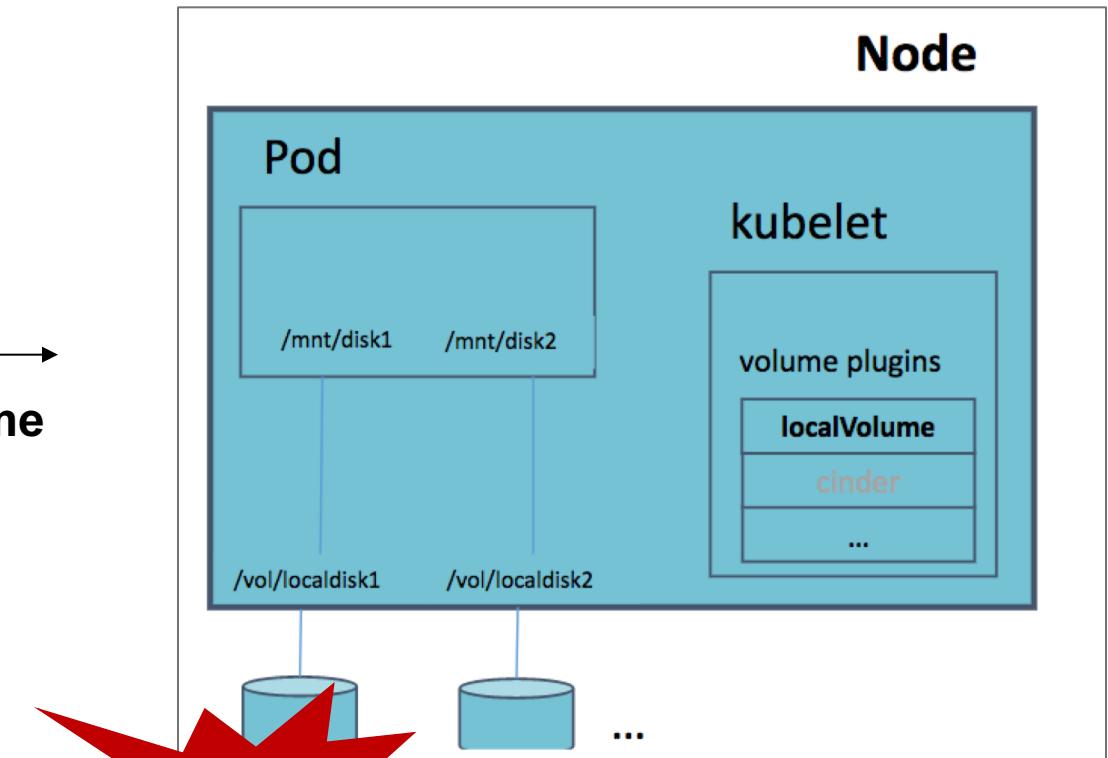
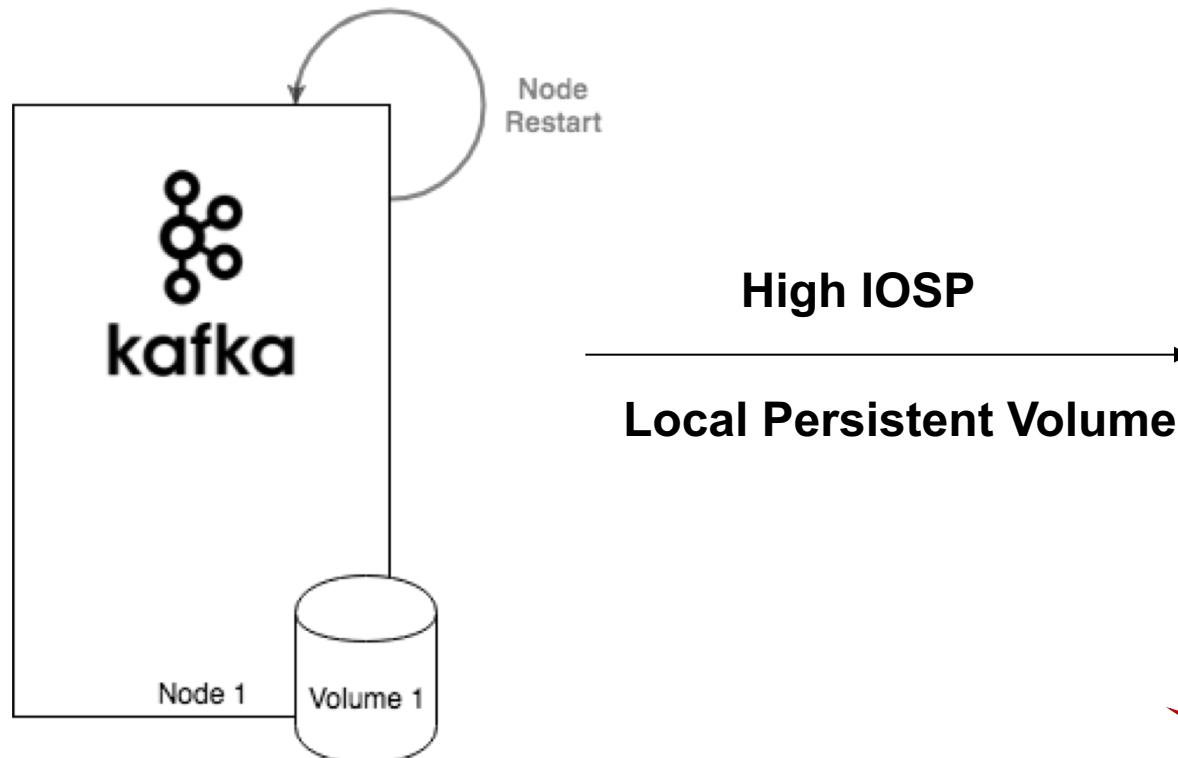
BACKGROUND



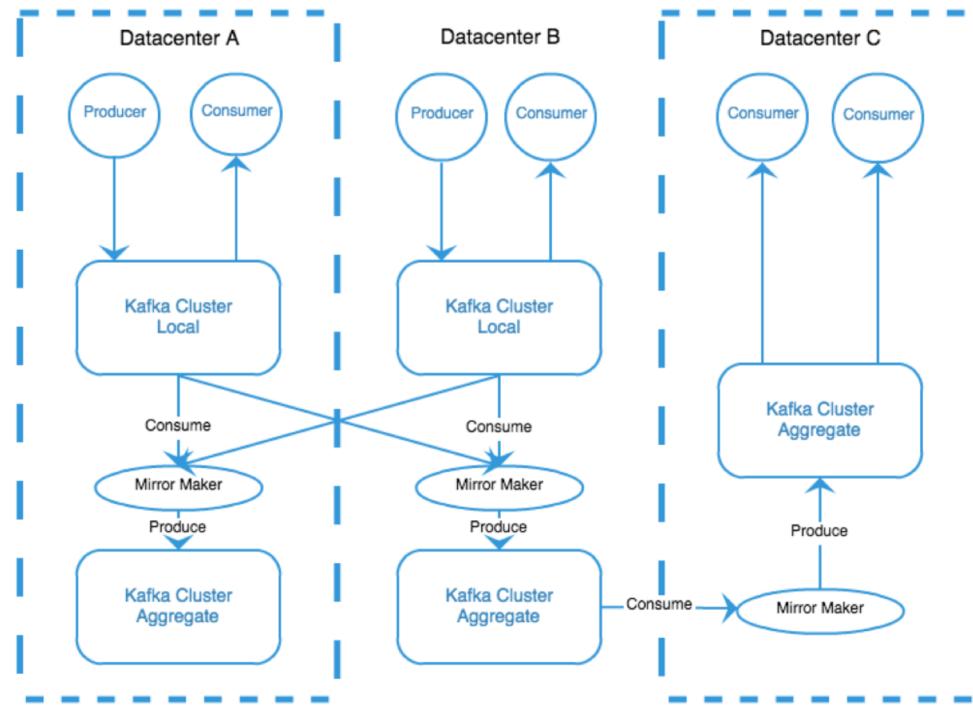
KubeCon

CloudNativeCon

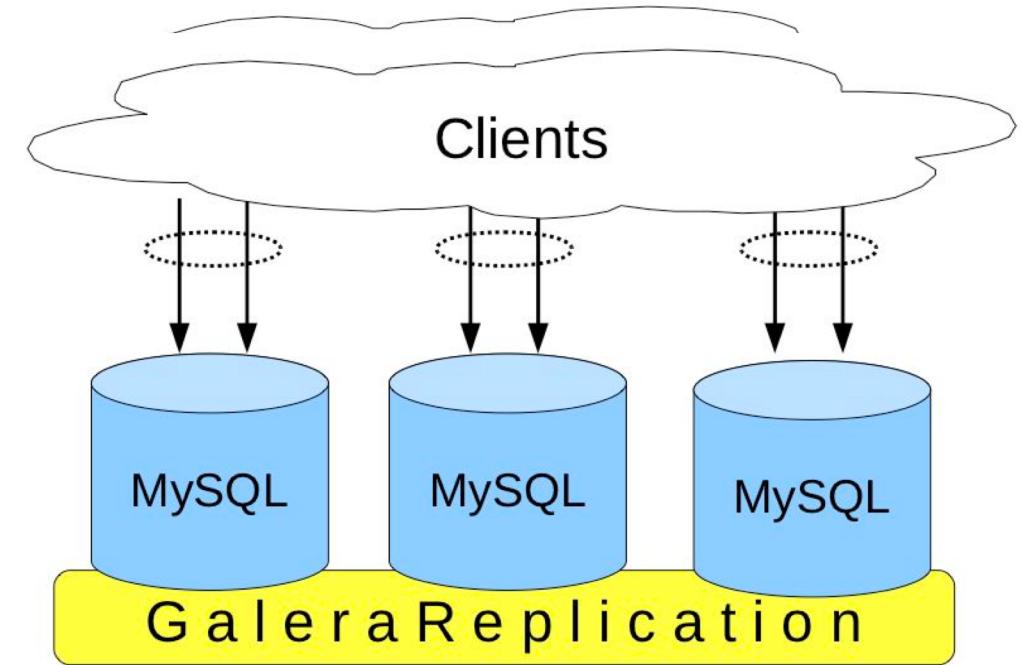
China 2018



BACKGROUND



Example: Tiered Kafka Architecture



Example: Active-Active Mysql Cluster

BACKGROUND

Simplify Management

- Multiple Kubernetes Clusters
- Internal Dependency
- Complex Data Operation
 - Scaling
 - Rolling Restart

Improve Reliability

- Always Available
- Resiliency
- Highly Scalable
- High Performance

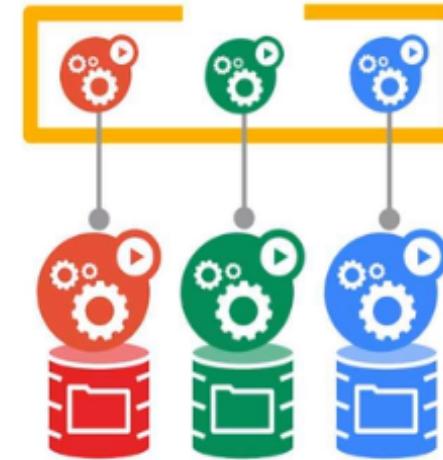
Security

- Integrate with existing enterprise security policies

BACKGROUND

StatefulSet:

1. Cross K8s clusters deployment
2. Auto Recover for Disk Failure
3. Defined order for rolling restart



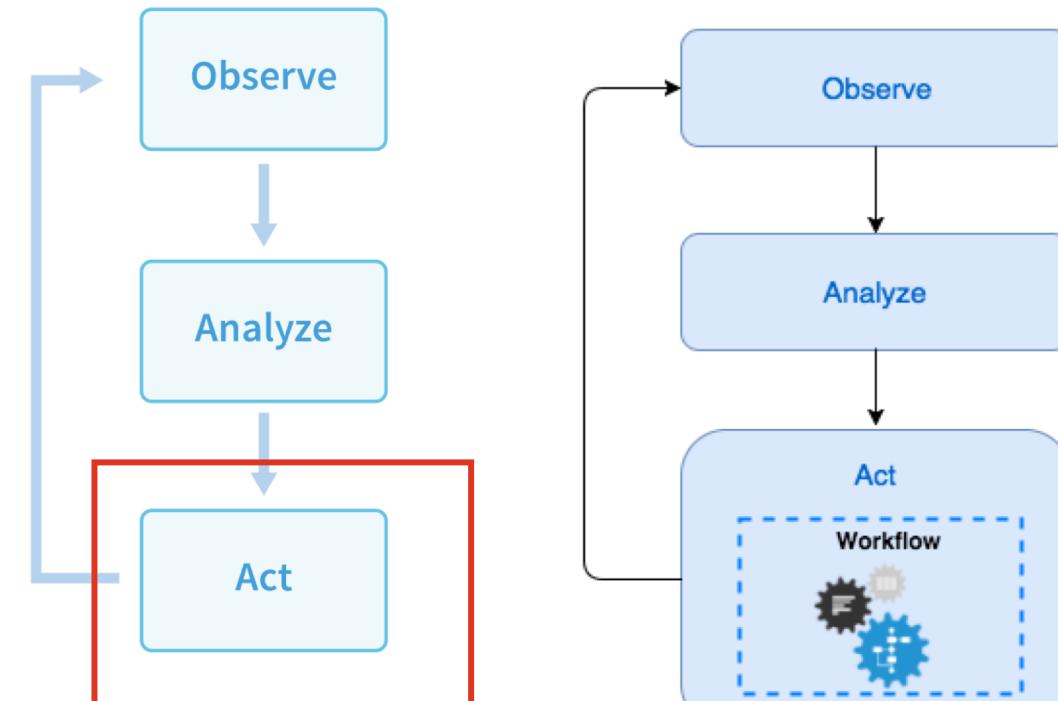
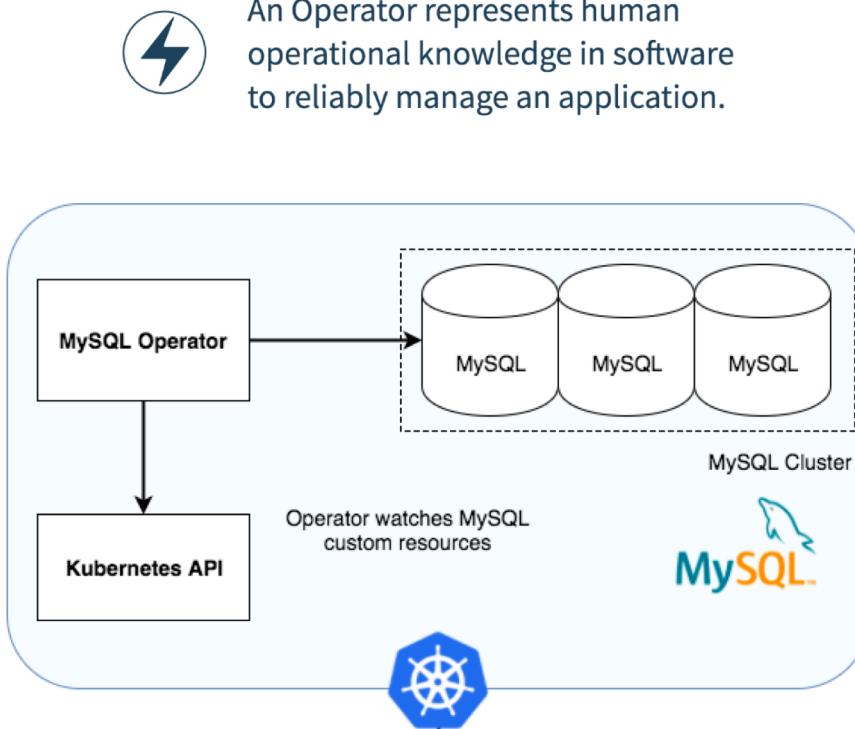
Helm Charts:

1. On-demand configuration changes for related components
2. Customize the docker images



INTRODUCTION

Operator Pattern + Workflow



INTRODUCTION

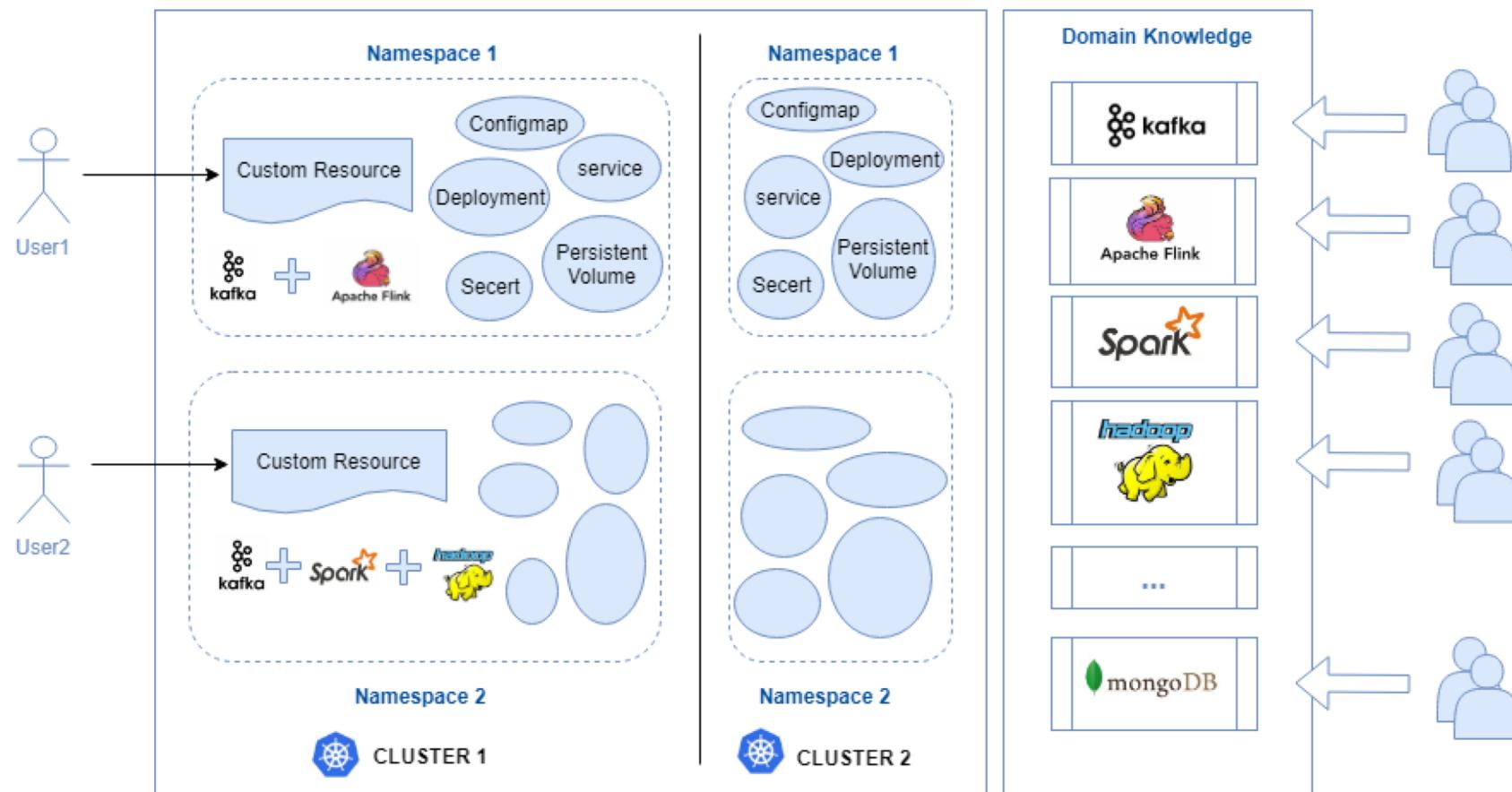


KubeCon

CloudNativeCon

China 2018

Operator Framework: One Stack Management for Stateful Workloads



FEATURES – Deploy Pattern



KubeCon

CloudNativeCon

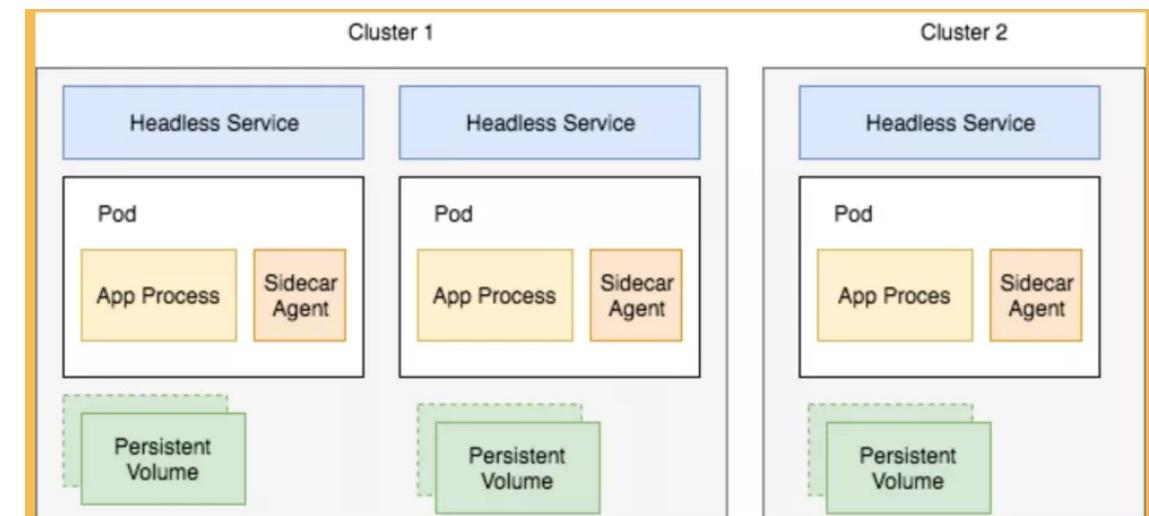
China 2018

```
1 ---
2 apiVersion: milkyway.ebay.com/v1
3 kind: Application
4 metadata:
5   name: stream1
6   namespace: milkyway
7 spec:
8   components:
9     - name: zookeeper1
10    config: default
11    regions:
12      - name: NA
13        zones:
14          - name: phx
15            size: 3
16        version: 3.4.7
17      - name: kafka1
18        config: default
19        regions:
20          - name: NA
21            zones:
22              - name: phx
23                size: 3
24            version: 1.1.0
25          - name: zookeeper2
26            config: default
27            regions:
28              - name: NA
29                zones:
30                  - name: phx
31                    size: 3
32            version: 3.4.7
33          - name: flink-jobmanager
34            config: default
35            regions:
36              - name: NA
37                zones:
38                  - name: phx
39                    size: 1
40            version: 1.4.2
41          - name: flink-taskmanager
42            config: default
43            regions:
44              - name: NA
45                zones:
46                  - name: phx
47                    size: 3
48            version: 1.4.2
```

1. Declarative interface and Free combination

2. Model data applications with K8s resources

- Pure Pods + Local Volume
- Deployment
- DeploymentSets
- StatefulSet



FEATURES – Deploy Pattern

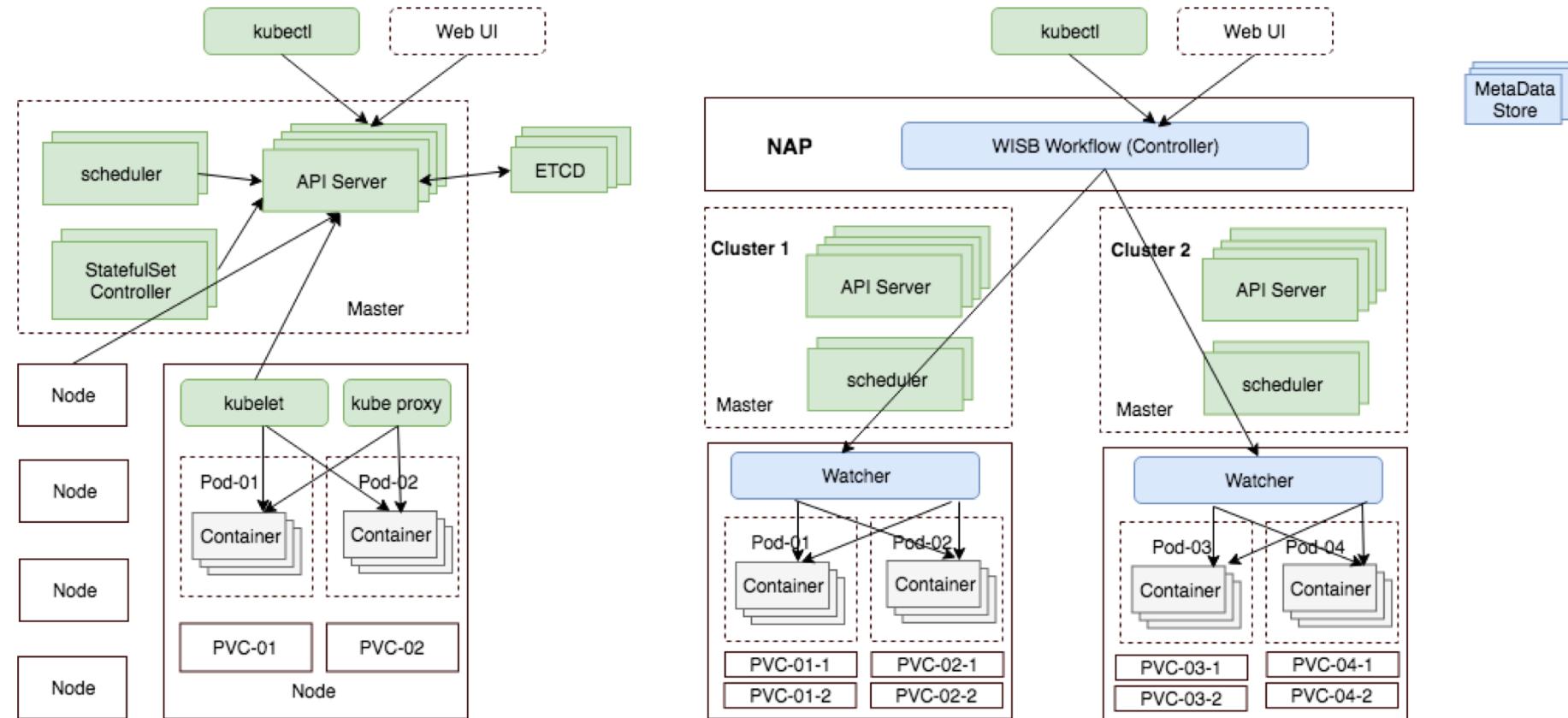


KubeCon

CloudNativeCon

China 2018

3. Cross Kubernetes Clusters Deployment and Management



FEATURES – Lifecycle Management

1. On-demand Management functions to reduce the maintain efforts

Provision/ Decommission

Rest API and Kubernetes native API for one step create or delete a cluster

Scaling

The cluster could on-demand flexup and flexdown

Configuration Management

Modify the application configuration parameters and update to the cluster

Auto Remediation

If one node is down, the WISB flow will automatically triggered to bring back the missed node

Rolling Restart / Upgrade

Upgrade the cluster to new version

Replacement

Replace the bad node or low performance node

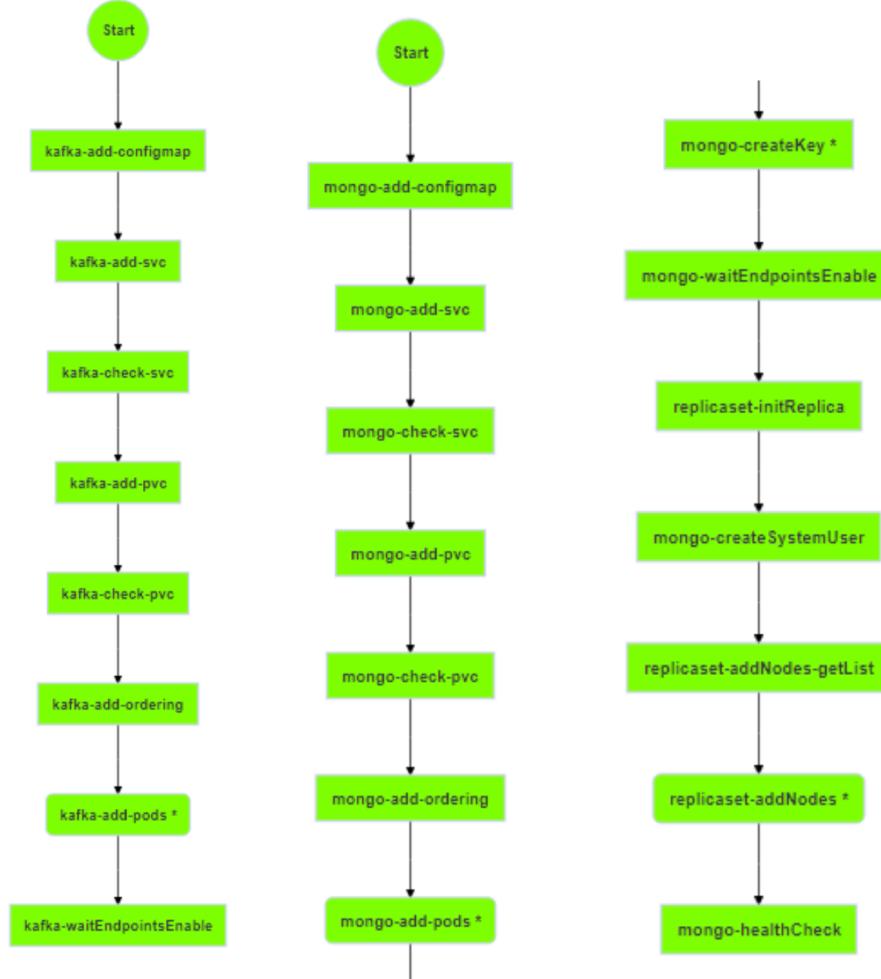
WISB Management Workflow

- Declarative with simple syntax via yaml, easy to modify and maintain
- Decompose complex logic with idempotent reusable and retryable tasks
- Design for failure
- Reusable and parallelism subflows
- Special design for group operation

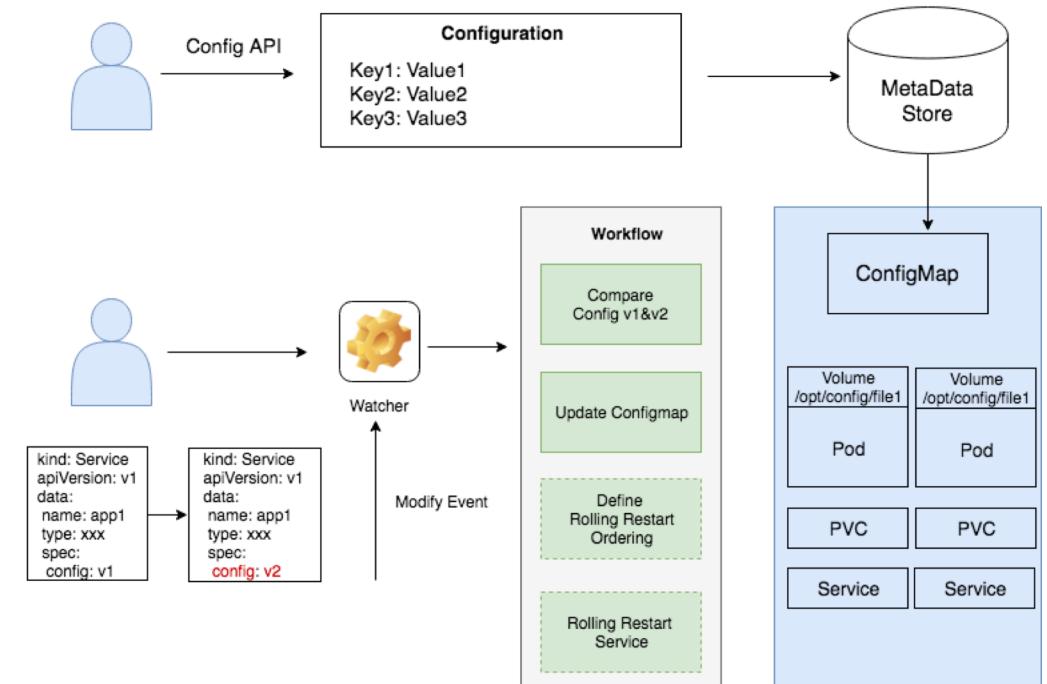


FEATURES – Lifecycle Management

2. Reusable common flow and customize WISB flow



Configuration Management



FEATURES – Reliability

Health Check

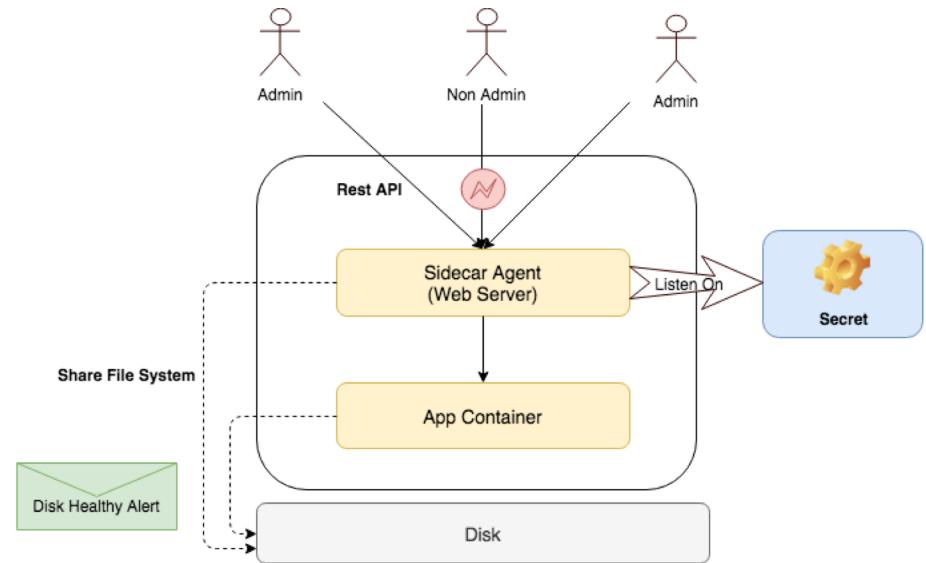
- Scheduling triggered **HealthCheck Flow** to check app status
- Send notification by Email or Stack if any unhealthy detected

The screenshot shows a monitoring interface with three sections:

- Task Log:** [TASK_FAILED:create-change-ticket] Task create-change-ticket:52025 in Flow zookeeper-wisb:50765 failed. Details: Task create-change-ticket (52025) State Scheduled Duration less than a second (443 ms). Message: Exception [Wed Nov 07 23:56:44 MST 2018] Error creating ticket for form <changeimplementation>:ERROR (552): Failure during SQL operation to the database; ORA-16456: switchover to standby in progress or completed. Time: 2018-11-07 23:56:45-07 | https://milkyway.vip.ebay.com
- Email Notifications:** Three separate email notifications for NAP-Production-NoReply@ebay.com. Each notification includes the subject [milkyway] [Heartbeat] NAP Health Report for [mirrormaker, kafka, flink] and the message "No alerts".

Sidecar Agent

- Proxy admin actions and permission control
- View Log and collect metrics
- Alerting (Disk healthy/ Disk usage)
- Secret rotation



FEATURES – Security



Authentication

Keystone authentication and integration with LDAP



RBAC

Grant the CRU permission to specific namespace, only the user have the CRU permission could manage the cluster



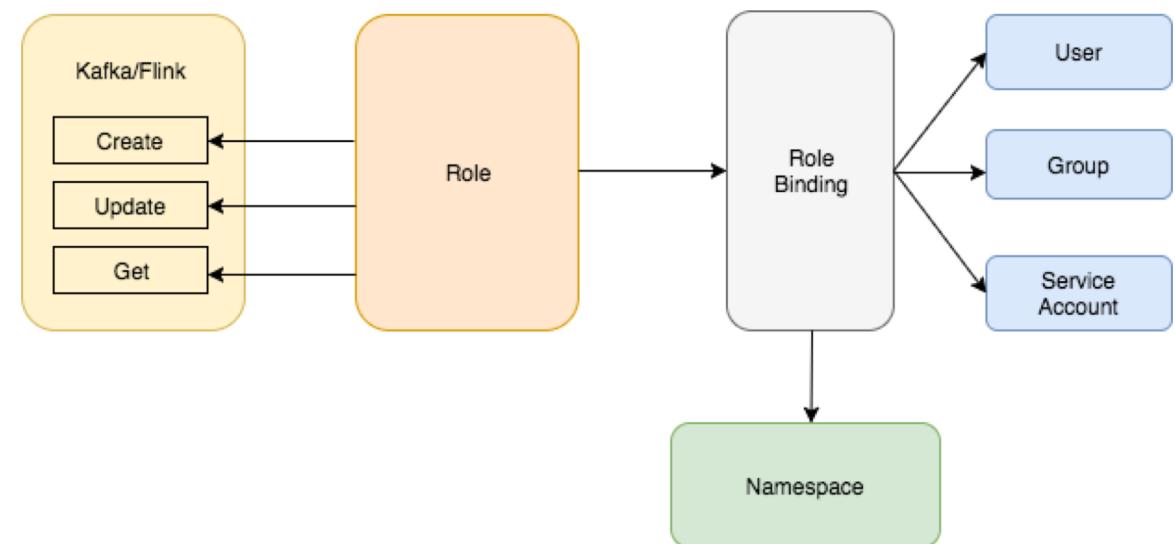
Non-root user

security context / setcap



Standard approve process

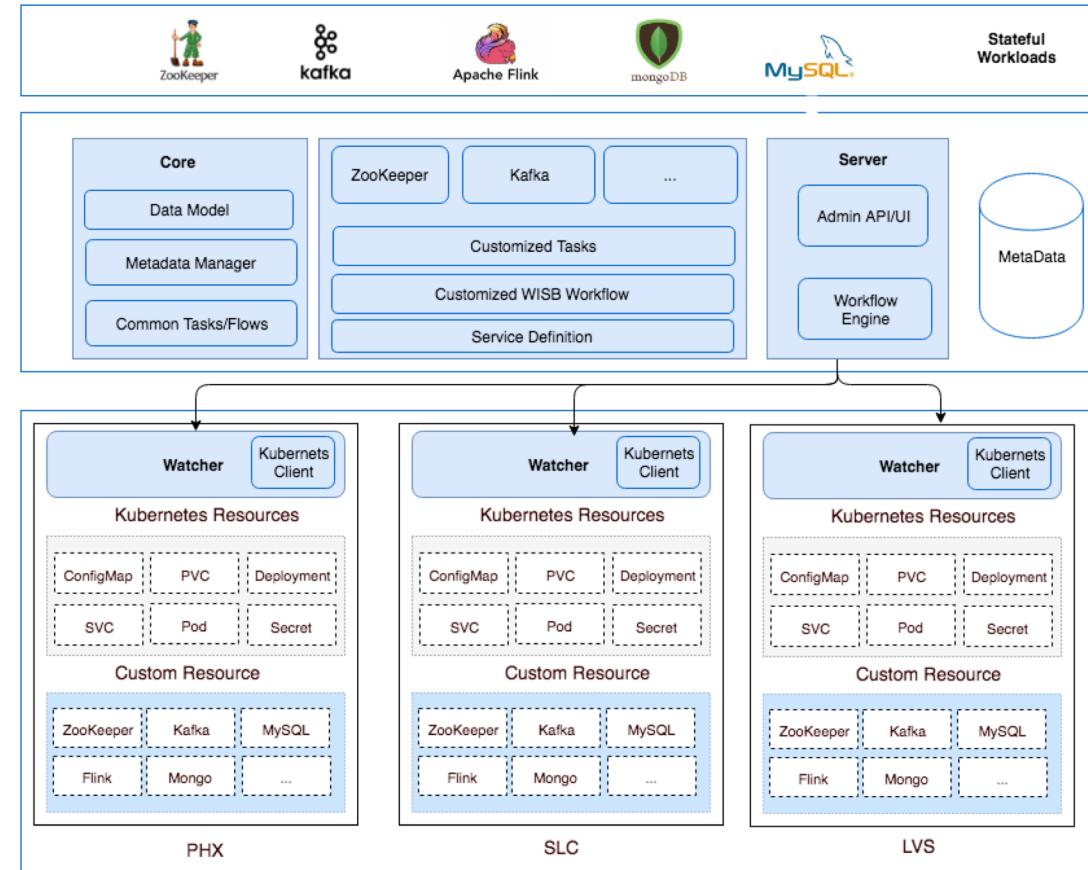
sudo action / human approve / Trace system



Conclusion

Operator Framework

We utilize the operator pattern, and benefit from the workflow engine to build a framework to makes it easier and more agile to manage all the multi-component, GEO-distributed stateful workloads.





KubeCon



CloudNativeCon

China 2018

Thanks !



KubeCon



CloudNativeCon

China 2018

