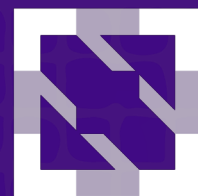




KubeCon

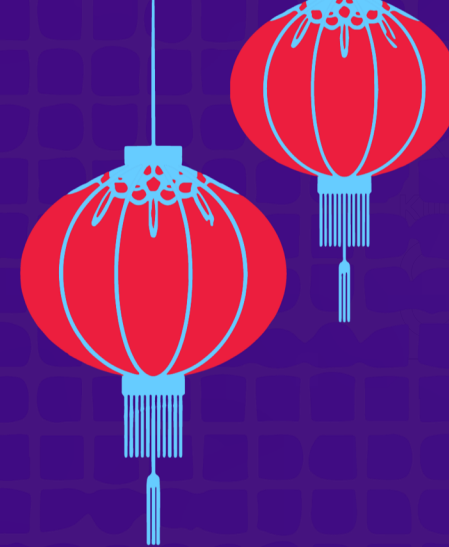
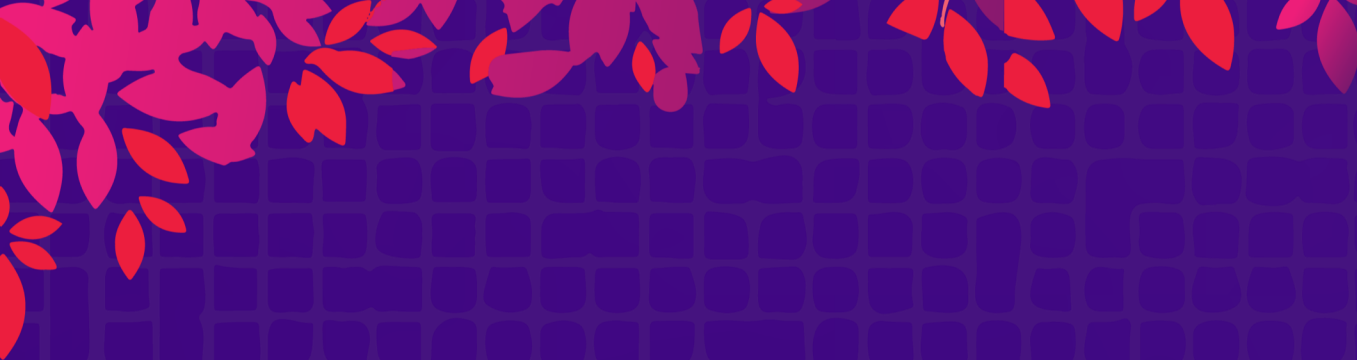


CloudNativeCon

OPEN SOURCE SUMMIT

China 2019





KubeCon



CloudNativeCon

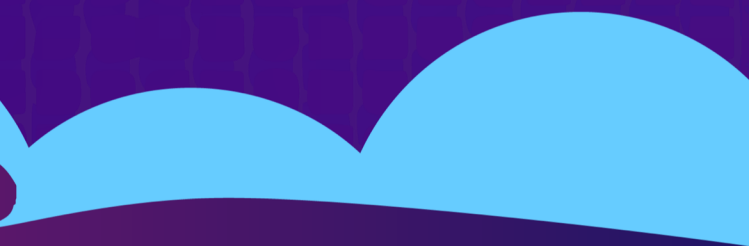


OPEN SOURCE SUMMIT

China 2019

Managing Large-Scale Kubernetes Clusters Effectively and Reliably

Yong Zhang – Ant Financial



About Me



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- **Yong Zhang**
 - **Ant Financial - Infra & Data**
 - **PAAS & Automated Cluster Management System**

Agenda



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

- **Background and Motivation**
- **Design Concept**
- **Cluster Management Operators**
- **Q & A**

Background: Cluster Scale



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

- Tens of clusters
 - Tens of thousands of nodes in one cluster
- Hundreds of thousands of pods
 - Tens of thousands of jobs
- Resource cost is huge



Motivation



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

- **Cluster Life Cycle Management**

- **Create**
- **Delete**
- **Upgrade**
- ...

- **Fault self-recovery:**

- **hardware failure**
- **component service exception**

- **Change Controllable**



Design Concept: Imperative



KubeCon

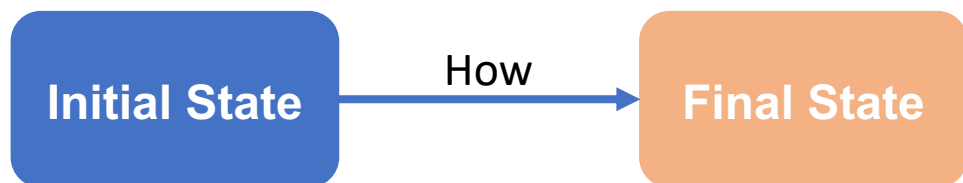


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- Imperative

- Tell What

- Tell How

Design Concept: Declarative



KubeCon

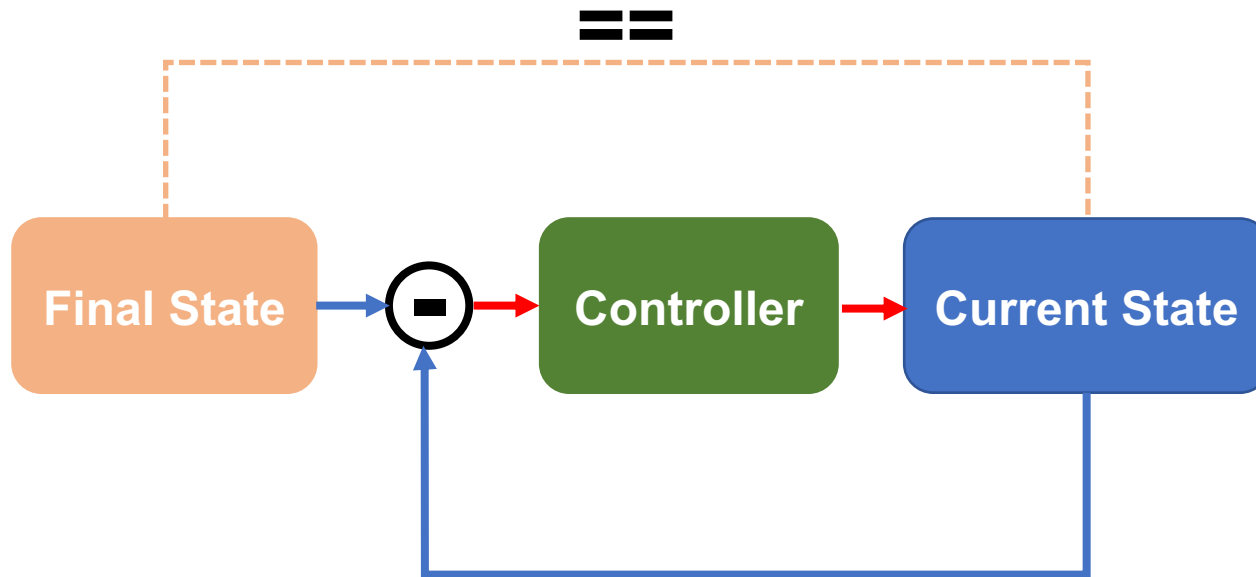


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- **Declarative**
 - **Tell What**
 - **Not How**

Design Concept: Self-recovery



KubeCon

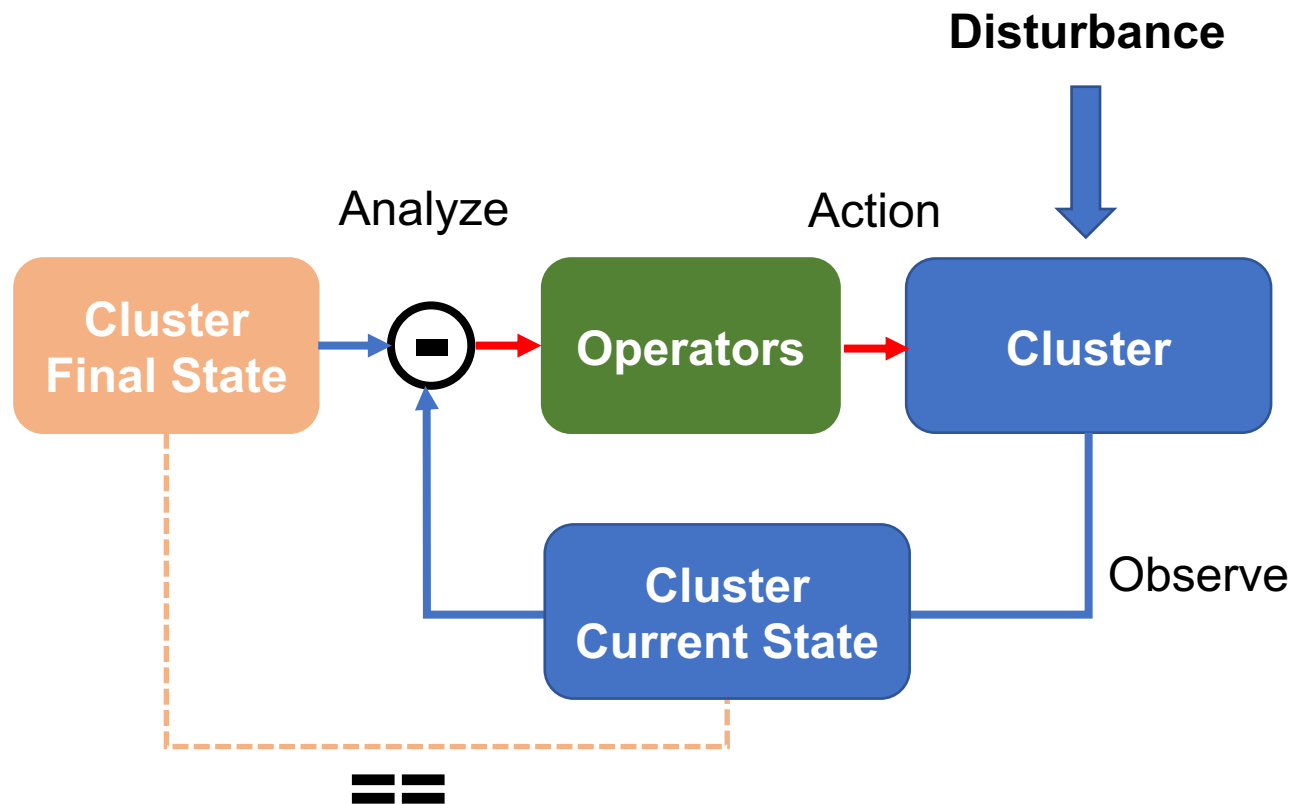


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- **Observe**

watch actual state

- **Analyze**

difference from desired and actual state

- **Action**

change to desired state

Cluster-Operator



KubeCon

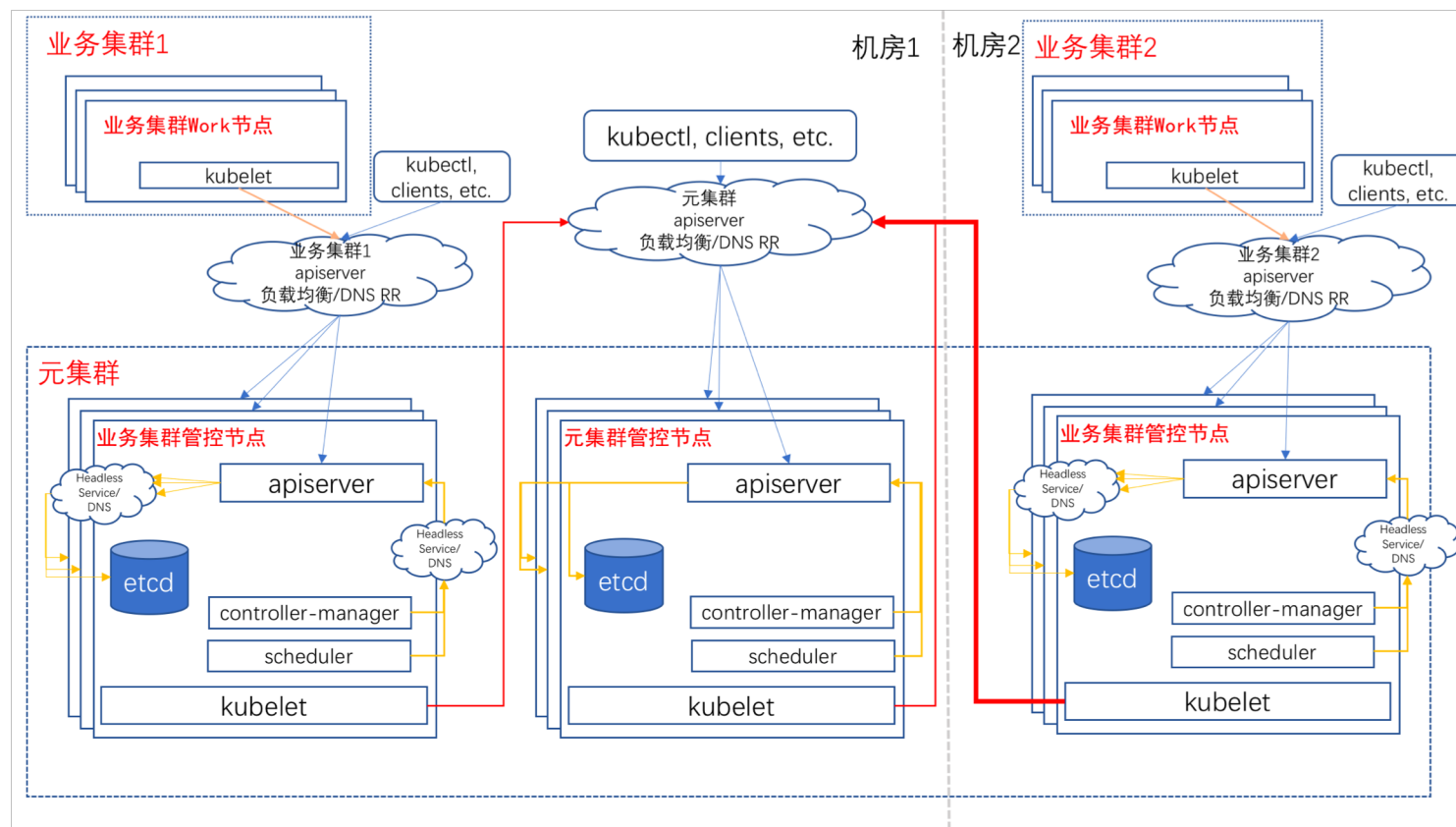


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- Cluster CRD
- ClusterPackageVersion
- Cluster-Operator

Machine-Operator



KubeCon



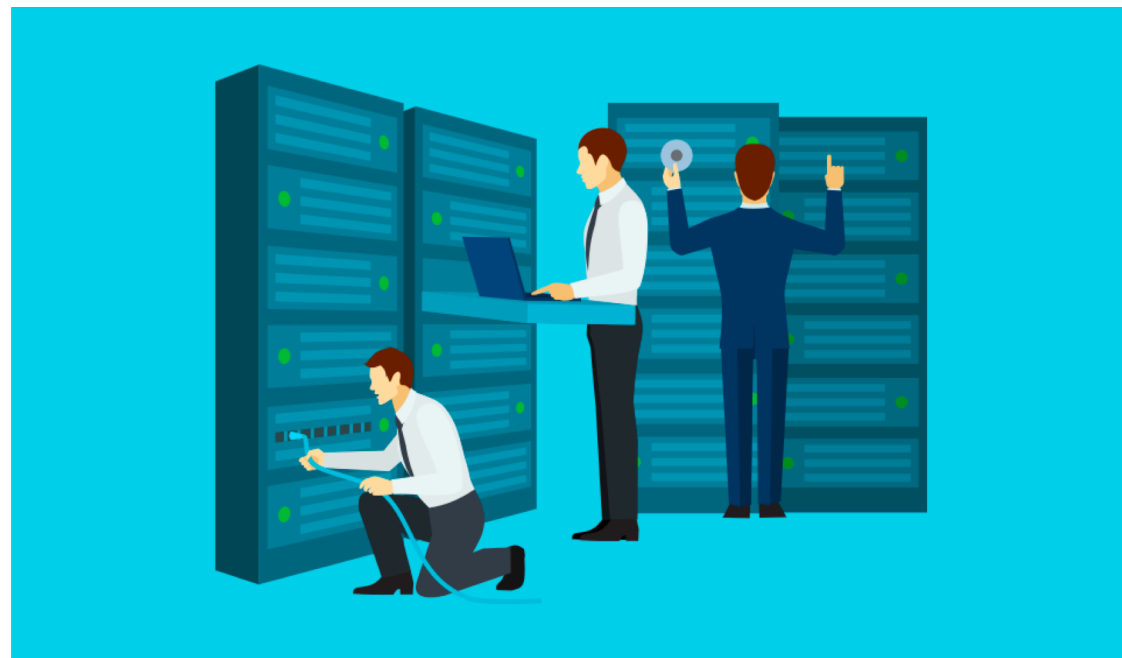
CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

- Node Configuration and Kernel Patch Management
- Docker / Kubelet Install、Uninstall、Upgrade
- Node Final-state Management
- Node Fault Self-recovery



Machine-Operator



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

```
kind: Machine
spec:
  idc: xxx
  ip: 10.10.10.1
  versions:
    pouch: 1.0
    kubelet: 1.1
status:
  phase: Running
  readinessGates:
    - conditionType: PouchOK
    - conditionType: OSOK
  versions:
    pouch: 1.0
    kubelet: 1.0
```

Machine-Operator

Node 10.10.10.1

NPD Pod

docker, CNI, kubelet,
Configure Files

```
kind: MachinePackageVersion
metadata:
  name: pouch-1.0
spec:
  packageName: pouch
  config:
    rpm: http://pouch-1.0.rpm
  configMaps:
    - name: pouch-config
      value: v1
```

Node Final-state Management



KubeCon

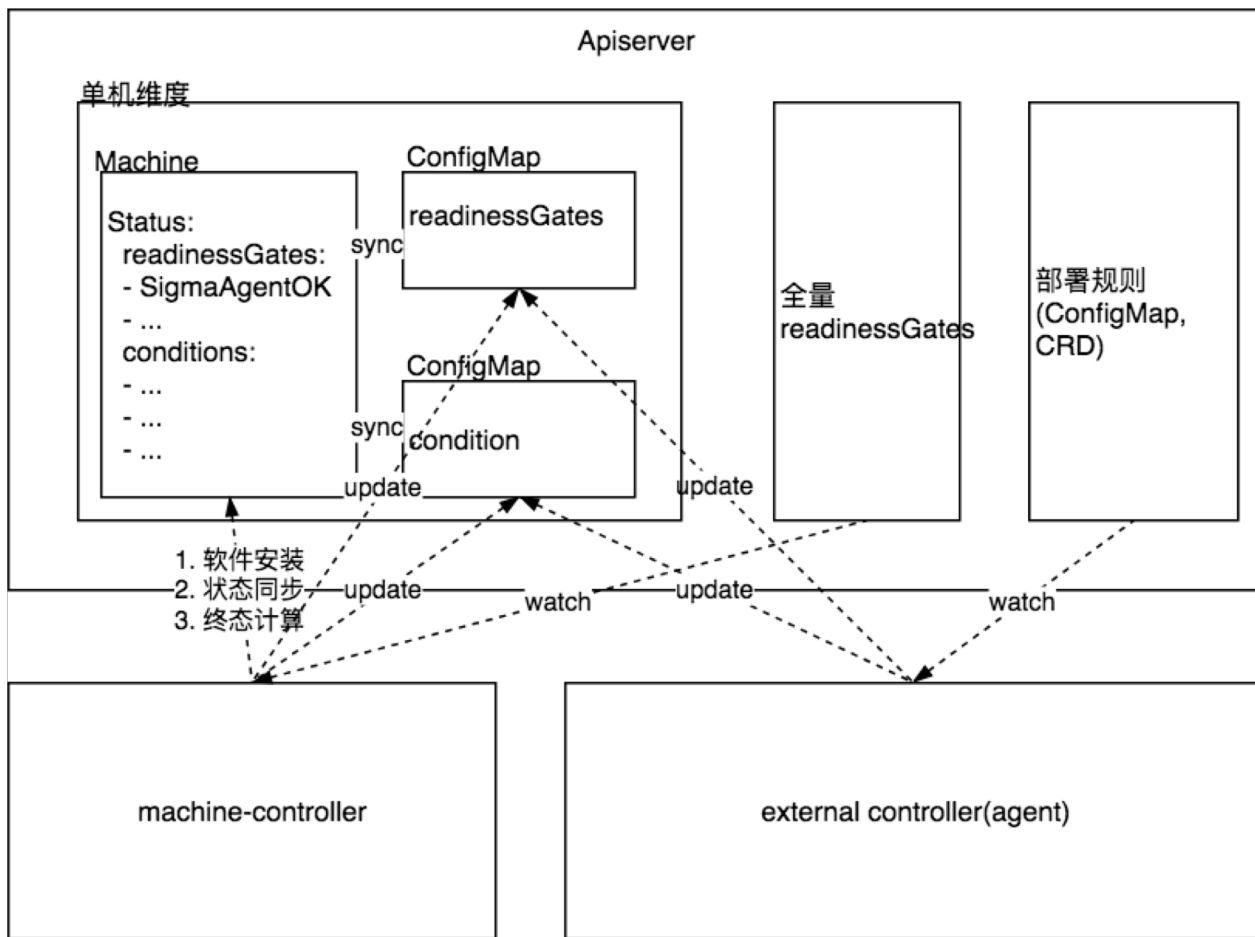


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- **ReadinessGates**
 - Node Schedulable Conditions
- **External controller**
 - DaemonSet
 -
- **Condition ConfigMap**
 - External Conditions

Fault Self-recovery



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

Hardware

- Power
- Disk
- Memory
- Motherboard
-



Resource

- Load
- Memory Pressure
- Disk Pressure

Component

- Component Crash
- Configuration Error

Fault Self-recovery



KubeCon

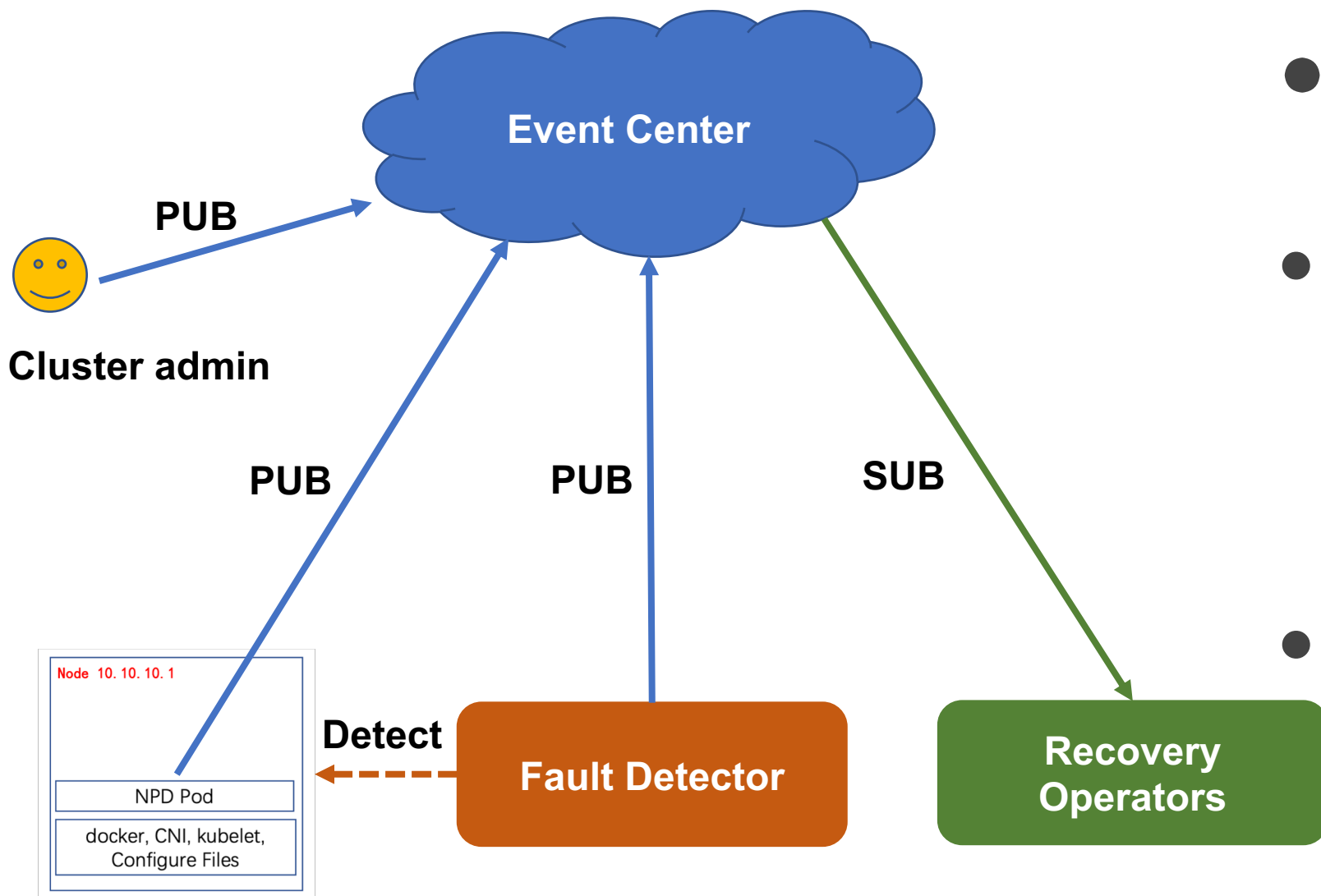


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- Event Center
 - Publish & Subscribe Event
- Fault Detection
 - Cluster admin
 - NPD
 - Fault Detector
- Fault Recovery
 - Recovery Operators

Fault Self-recovery



KubeCon

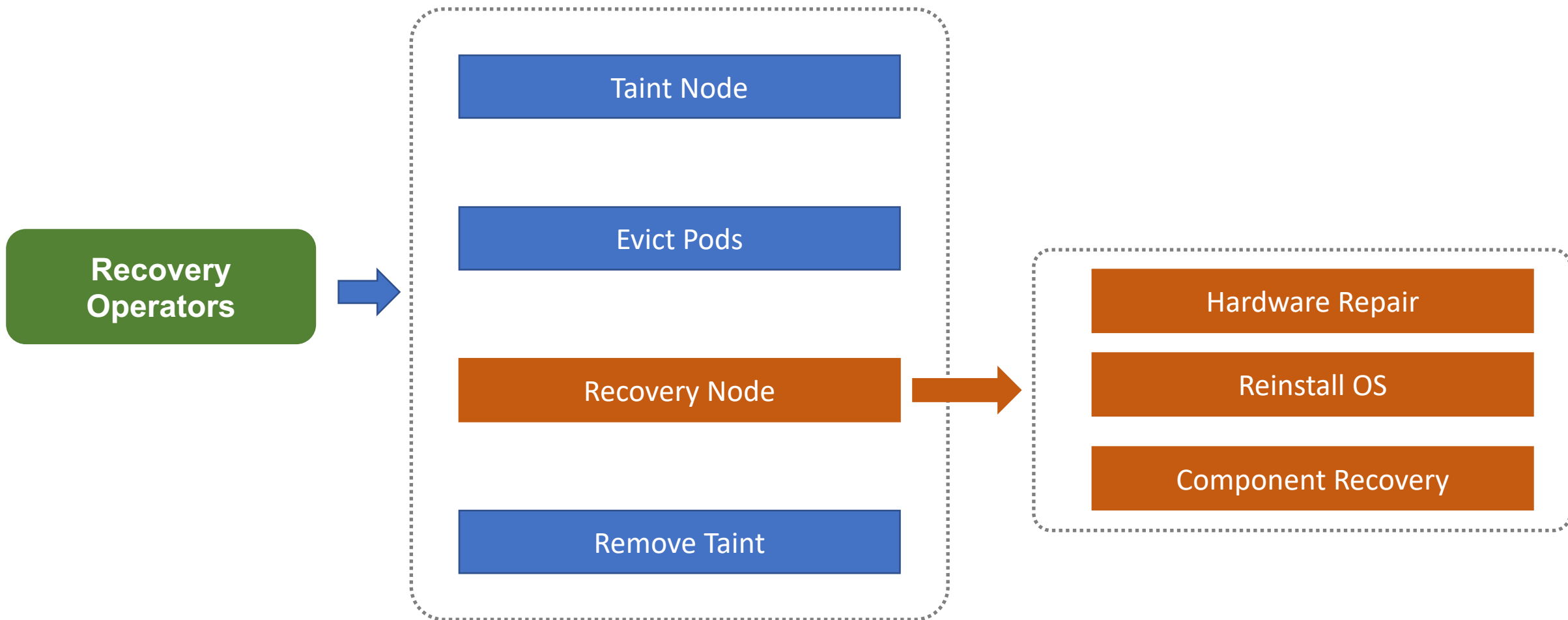


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



Risk Prevention



KubeCon

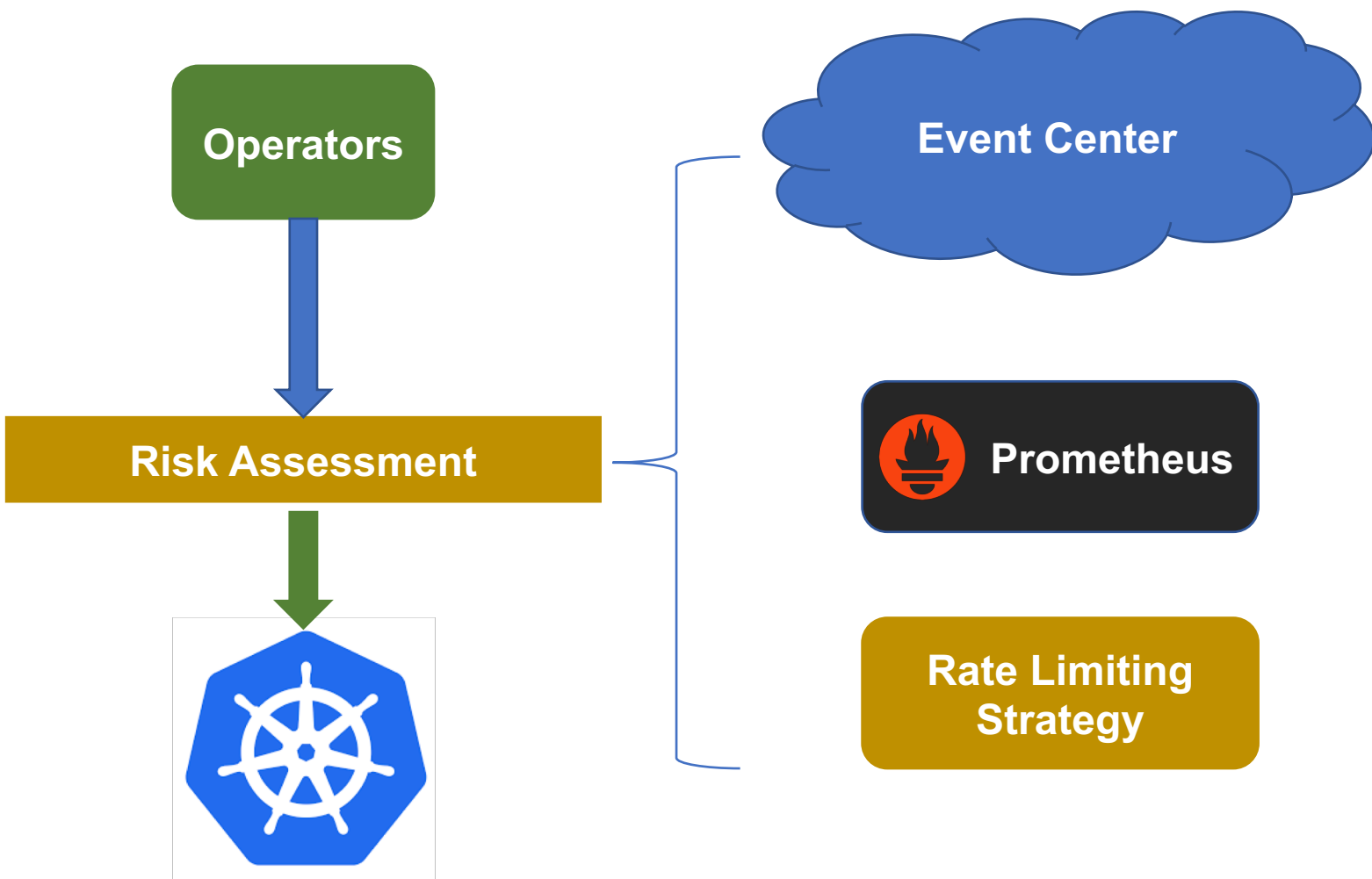


CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



- Circuit-Breaker
- Rate Limit

Q & A



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

Thanks !



扫一扫上面的二维码图案，加我微信