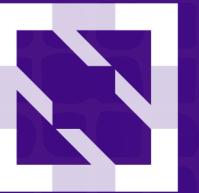




KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019





KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

# No More Chaos - Inspect Kubernetes at Scale

Jie Chen - Alibaba Cloud  
Jim Ma - Ant Financial



# About Us



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

## Jie Chen

- Senior Engineer at Alibaba Cloud

WeChat: anole2010

Email: hantang.cj@alibaba-inc.com

## Jim Ma

- Engineer II at Ant Financial

WeChat: majinjing3

Email: chuxian.mjj@alipay.com



# Agenda

- **Why We Need Inspection on Kubernetes Clusters?**
- **Considerations on an Inspection Service**
- **Build A Self-Healing Inspection System**
- **The Future**

# We have A Large Scale Kubernetes Cluster

- **Nodes (>10k)**
  - Types of Hardware
  - Types of OS
  - Types of Kubelet
- **Pods (>100k)**
  - Jobs (>10k)
- **Controllers (>100)**
  - Custom Resource Types (>100)



**So many Kubernetes resources and custom resources**

**Does the cluster really keep eventual consistency ?**

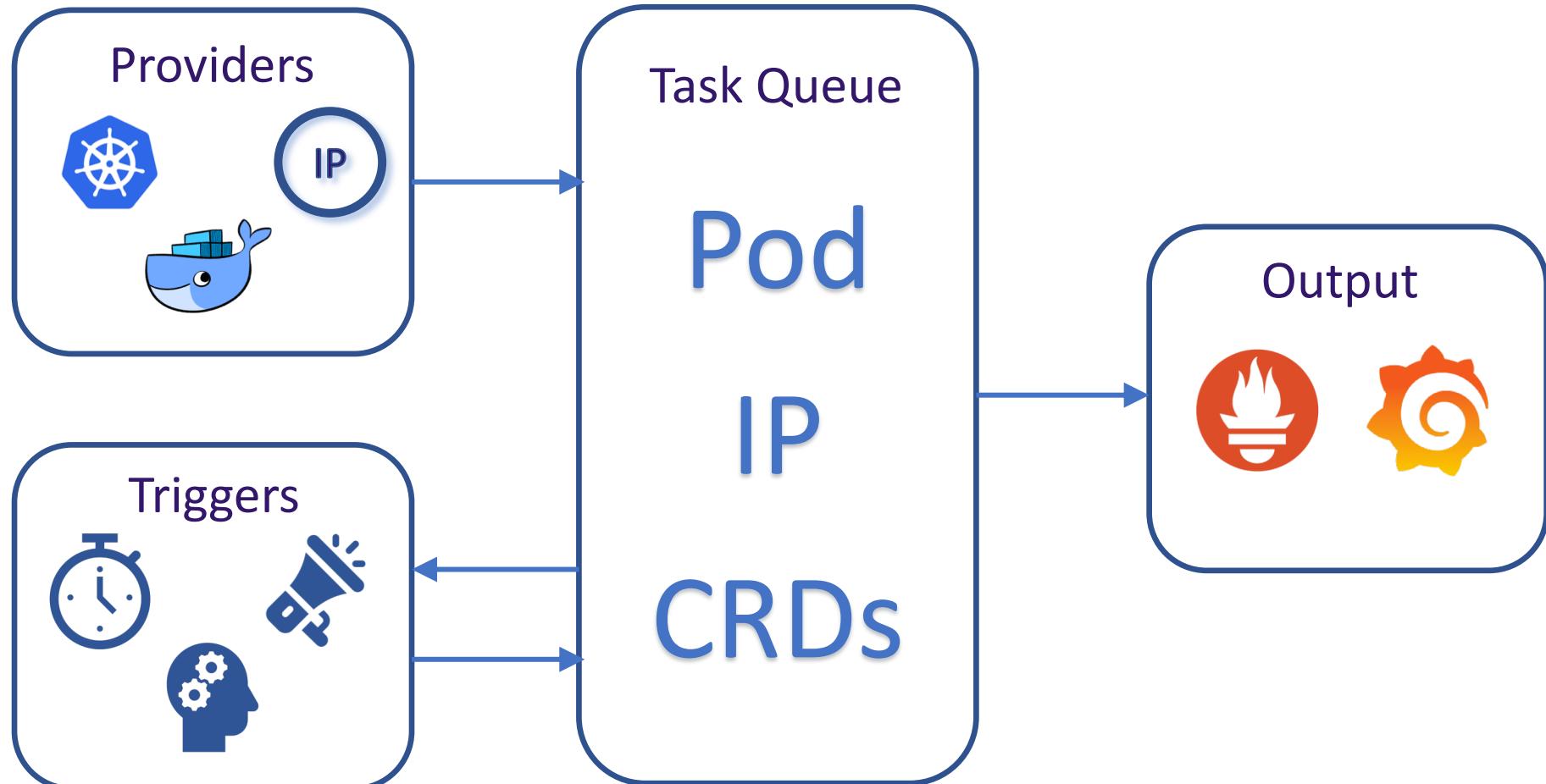
- IP Conflict
  - Service Interrupt
- Dangling Resources
  - Pods
  - PVs
  - CRDs
- Inconsistent with External System
  - CMDB
  - IPAM

We need to dive into our Kubernetes cluster to  
inspect all the resources

# Considerations on An Inspection Service

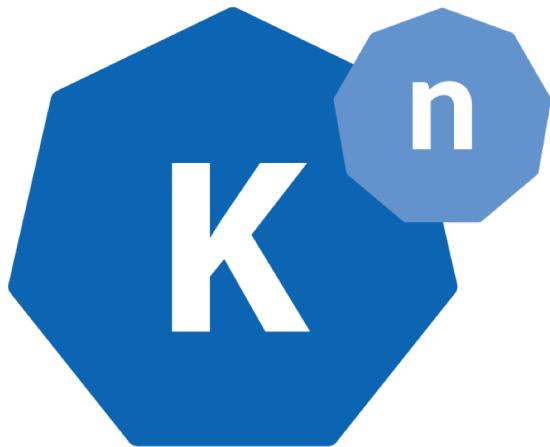
- **Efficiency**
  - Capable to inspect ten of thousands of nodes
  - Capable to inspect hundreds of thousands of pods
  - Capable to inspect hundreds of custom resource (\* pods)
- **Scalability**
  - Support more teams/groups
  - Setup hundreds of inspecting points
- **Observation**
  - Auditing
  - Monitoring

# What We Build in Version 1



# Problems in V1

- **One pod per cluster**
  - Limited performance per node
  - Resource conflicts among tasks
  - Unable to scale horizontally
- **Lack of generality**
  - Only an internal closed-loop

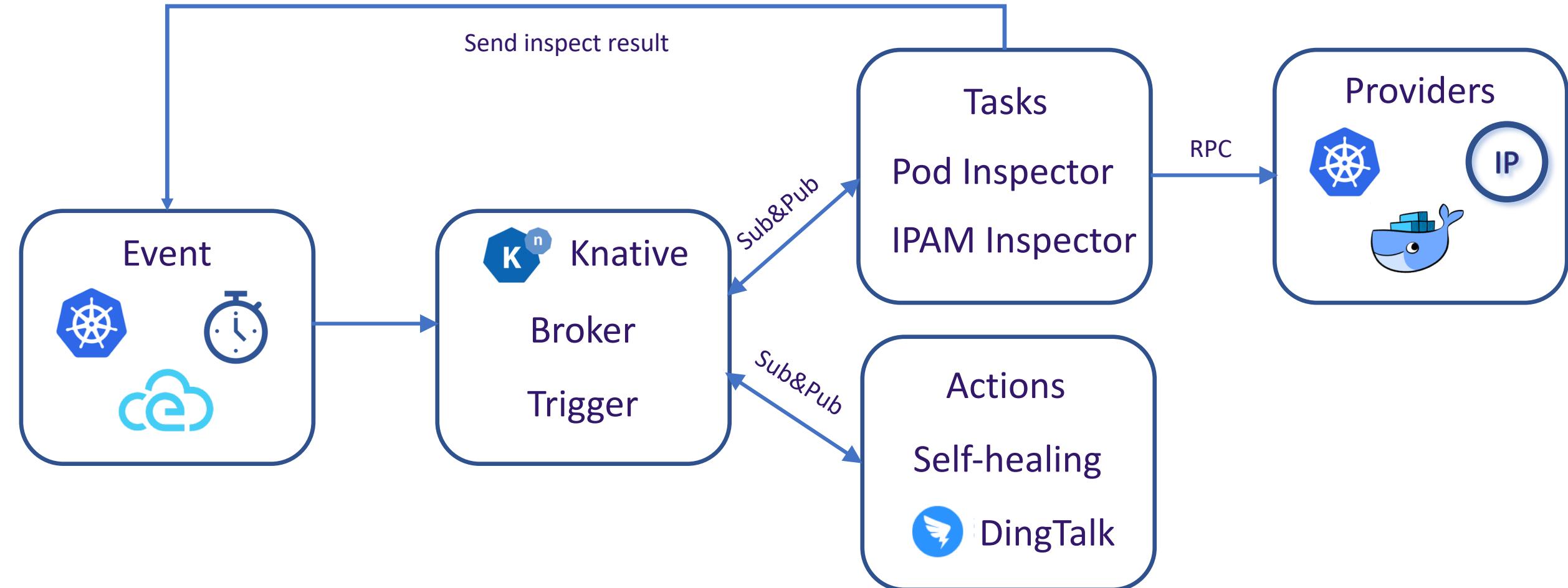


# Knative

# Knative : The Killer Component

- Use Serverless to horizontally scale inspections
- Use CloudEvent to be more scalable
- To be a Framework not a internal service any more

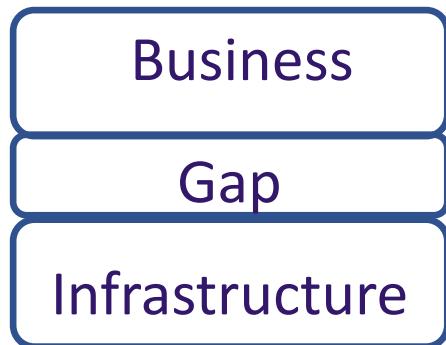
# We Build V2 in Serverless with Knative



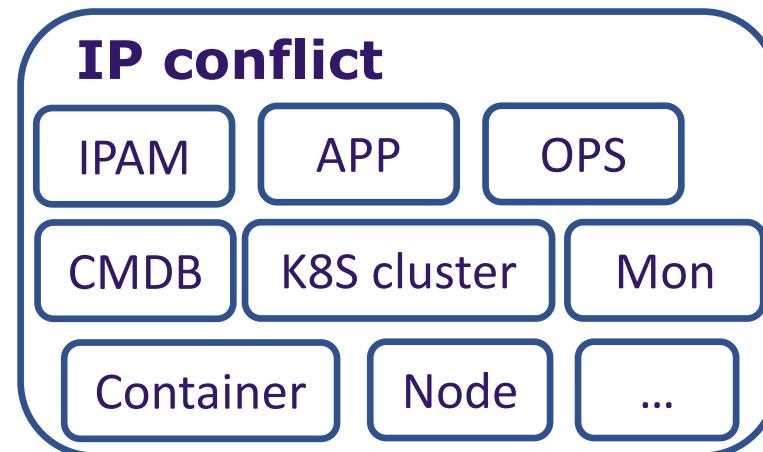
# How to Build A Complete Inspection system

# Some Problems in Business Stability

1) A huge gap between infrastructure and business



2) Lack of global perspective of status inspection



# Changes When We Moving to Cloud Native and Kubernetes?



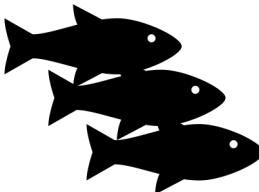
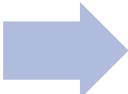
**Imperative**



**Declarative**



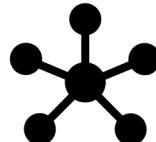
**Giant Container like VM**



**Flexible Containers in Pod**



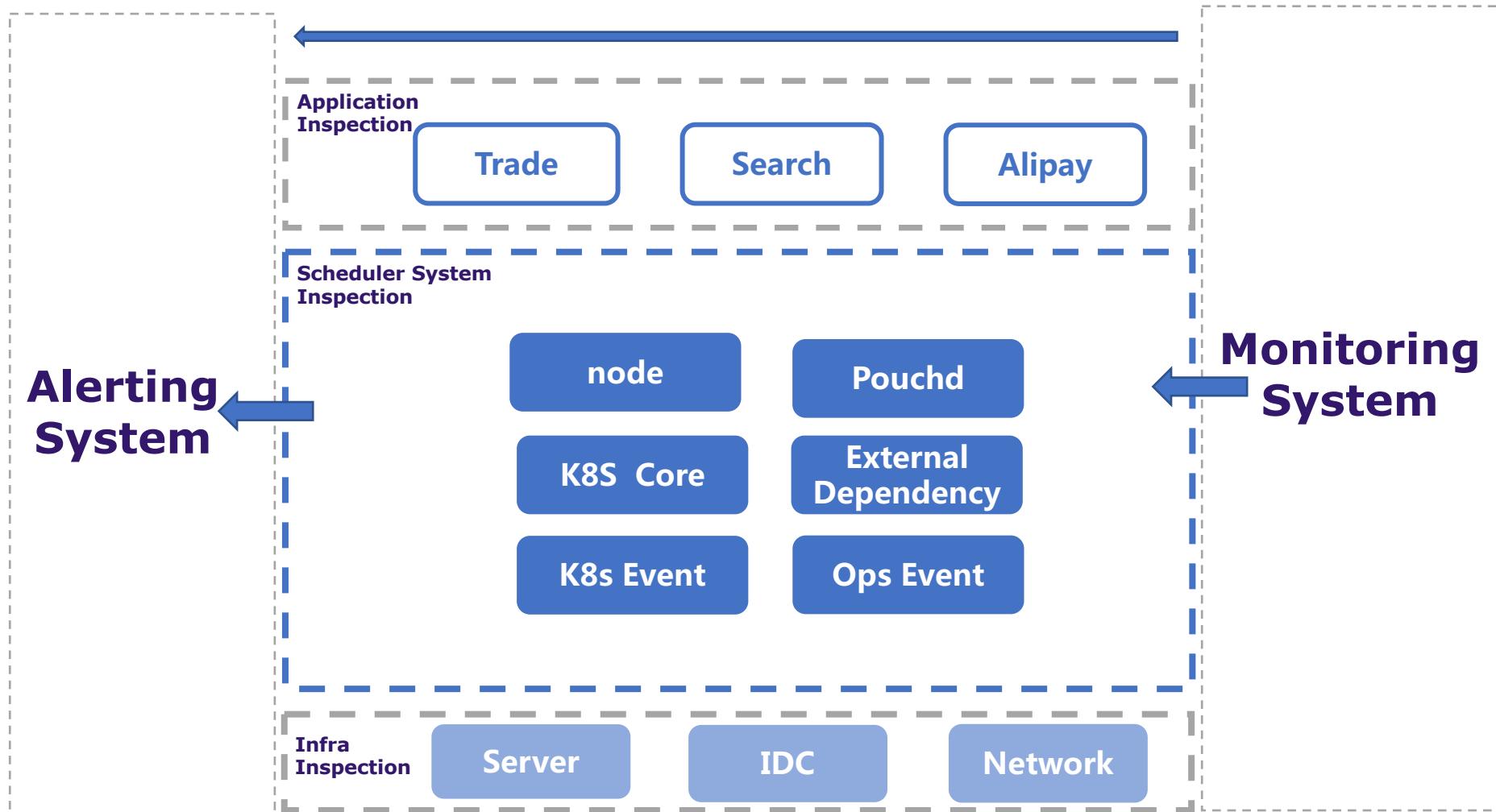
**Just Resource Allocation**



**Works more**

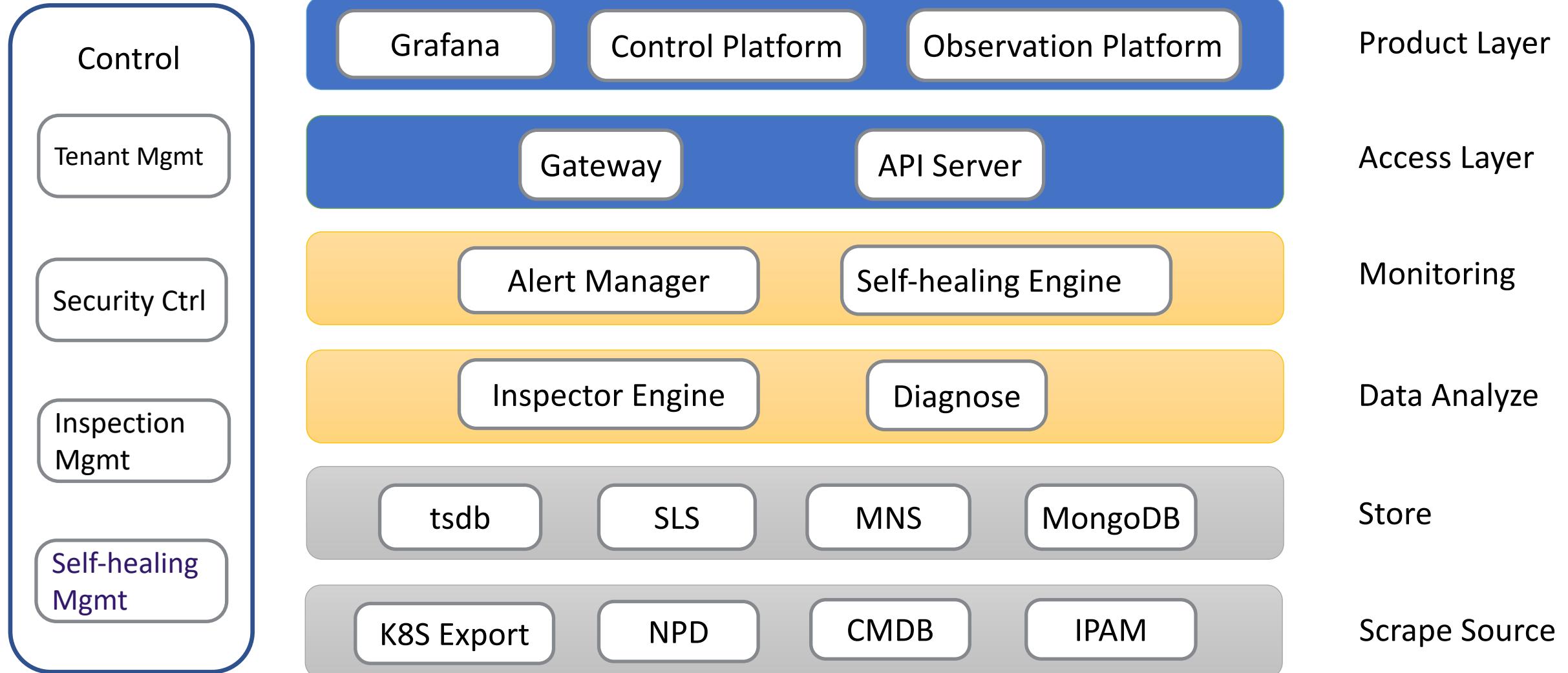
# How to Handle These Problems and Changes?

# Boundary of Inspection System



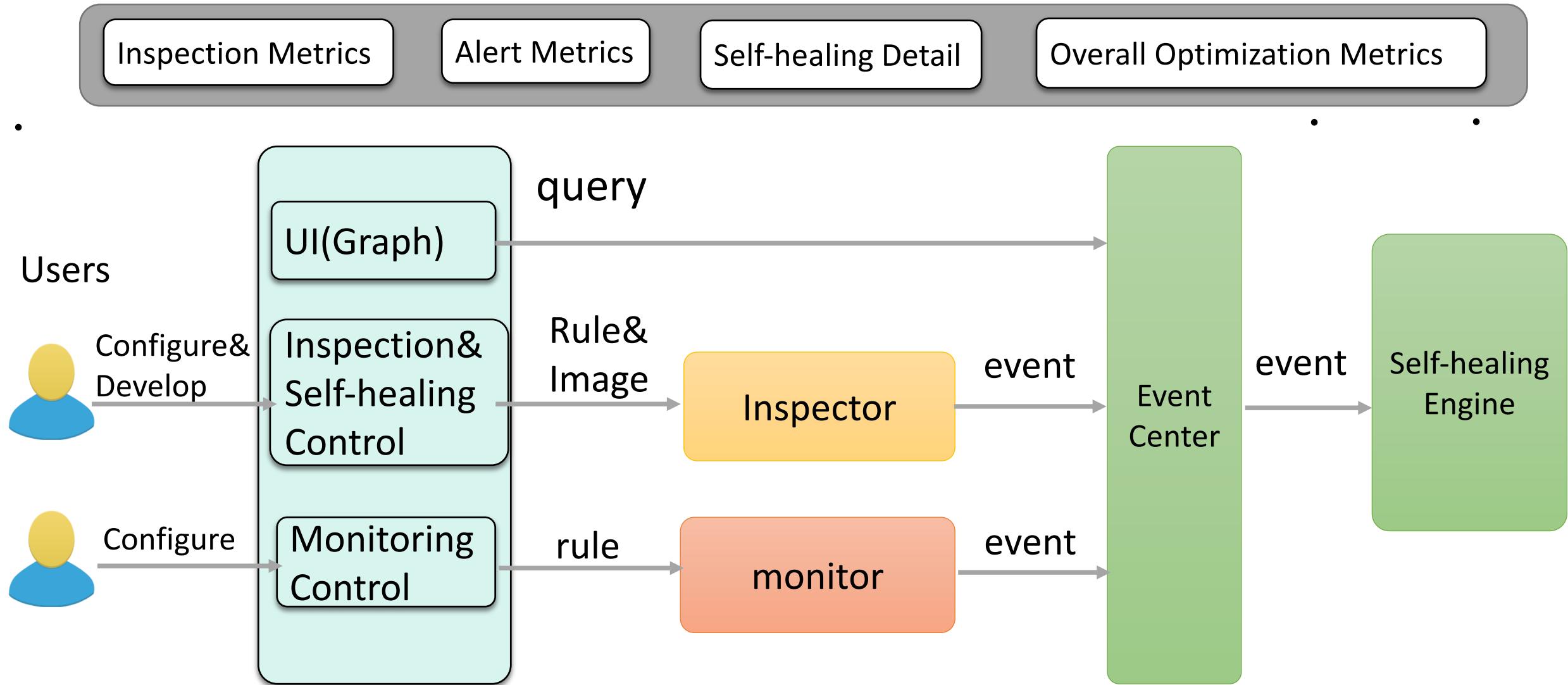
# We Can Do More about Analysis and Make Decision

# Self-healing Inspection System - Overview



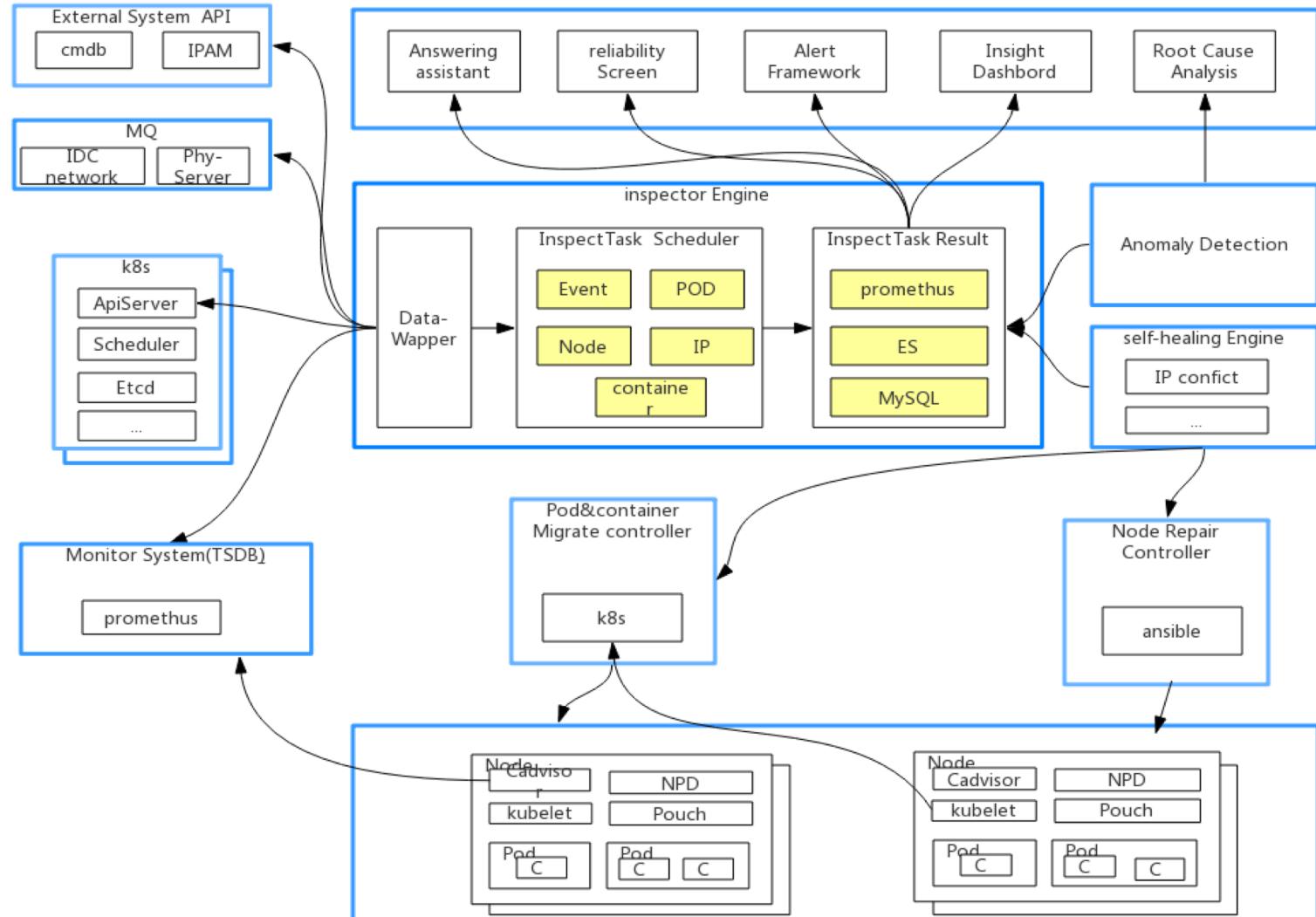
# How to Enable Users do Inspect?

# User-driven System Inspecting



# Practices

# Core Engine for Self-Healing Inspection System





KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019

# Examples & Results

## 1) Observation

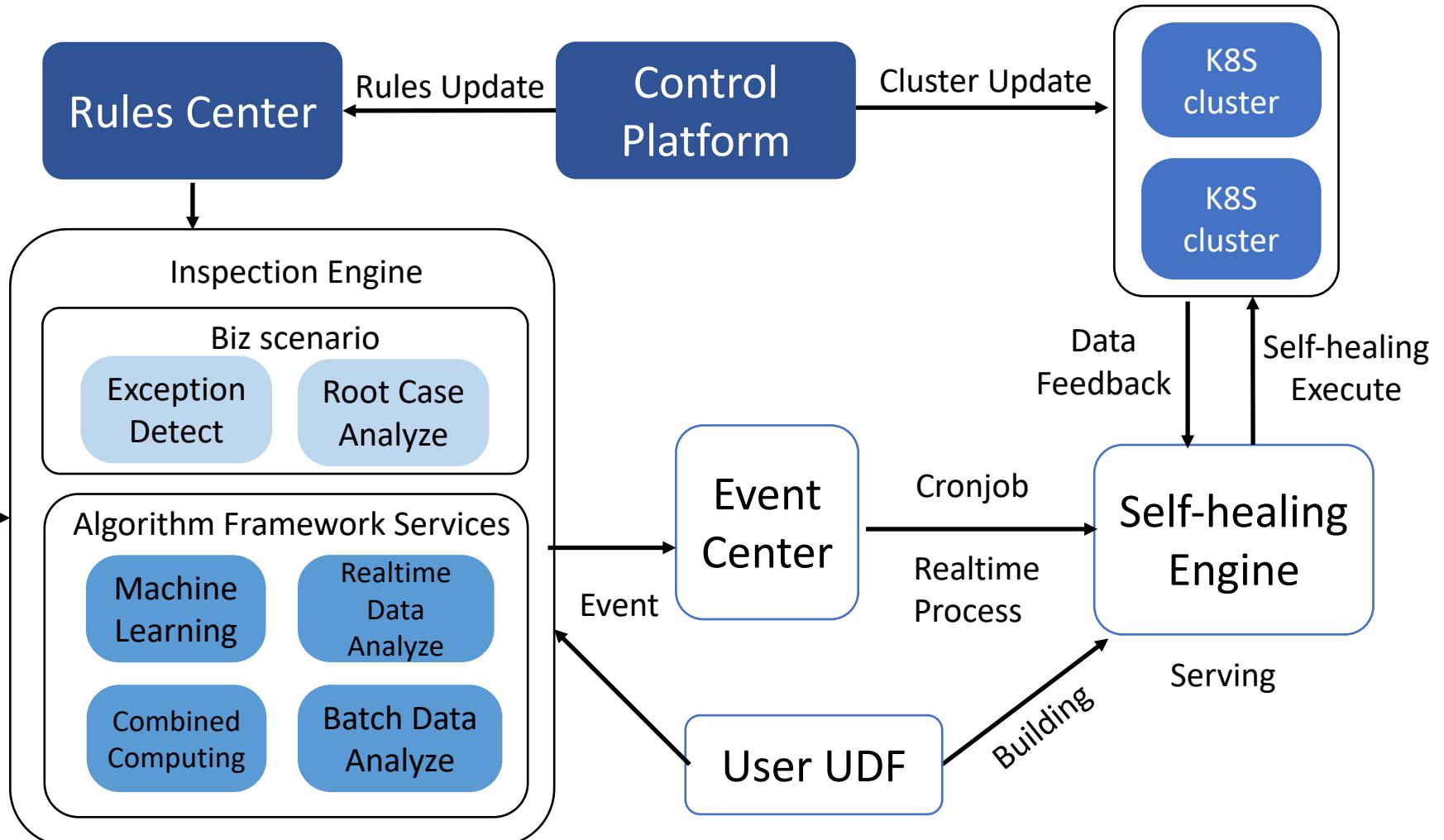
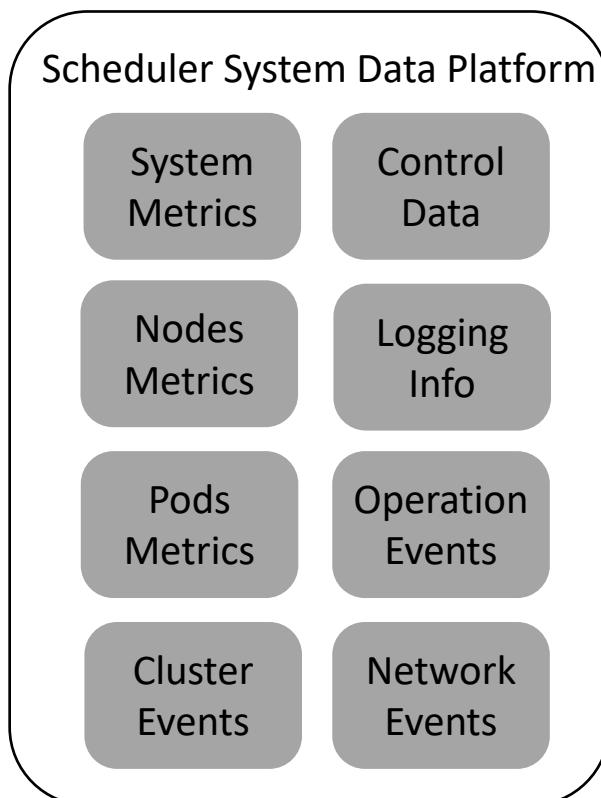


## 2) Healing

| 分类      | 指标项           | 是否符合SLO | SLO承诺 | 现状概览 | 异常详情                 |
|---------|---------------|---------|-------|------|----------------------|
| 平台组件可用性 | 超卖            | 不正常     | 10000 | 2000 | <a href="#">查看详情</a> |
| 平台组件可用性 | 资源规则          | 正常      | 10000 | 2000 | <a href="#">查看详情</a> |
| 平台组件可用性 | 脏数据           | 正常      | 10000 | 2000 | <a href="#">查看详情</a> |
| 外部依赖可用性 | 拉镜像成功率        | 正常      | 100   | 90   | <a href="#">查看详情</a> |
| 外部依赖可用性 | 拉镜像成功率        | 正常      | 100   | 90   | <a href="#">查看详情</a> |
| 外部依赖可用性 | 网络Overlay-ENI | 正常      | 100   | 90   | <a href="#">查看详情</a> |
| 外部依赖可用性 | CMDB(Skyline) | 正常      | 100   | 90   | <a href="#">查看详情</a> |
| 调度质量SLO | 组件异常          | 正常      | 200   | 80   | <a href="#">查看详情</a> |

# Future: AIOps-driven Self-Healing Inspection System

Data, Schedule, Control, Ops  
Closed Loop





## Contact Us



KubeCon



CloudNativeCon



OPEN SOURCE SUMMIT

China 2019



Jie Chen 陈杰



Jim Ma 马金晶



免登陆听课

动手实践课后自测

立即听课

CKA课程内容同步

阿里云原生最佳实践

# Thanks