KubeCon | CloudNativeCon

China 2018

ebay

Production Cluster Monitoring and Remediation for High Reliability at eBay

钱世俊, Cloud Software Engineer, ebay @danielqsj

刘应科, MTS1, Cloud Software Engineer, ebay @keyingliu

Agenda

Growing Clusters

Monitoring

Remediation

Q&A

Growing Clusters

30+

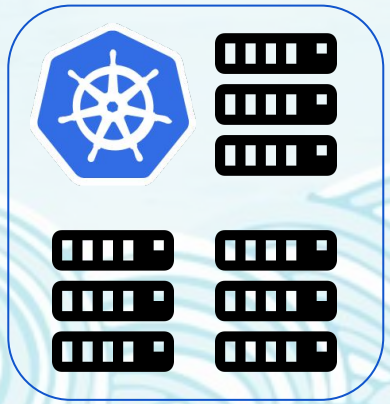
Clusters

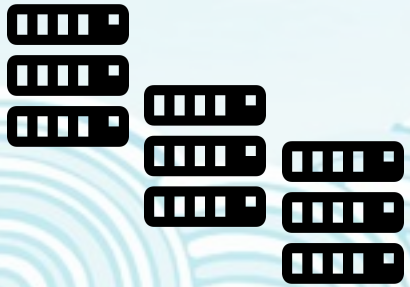
8K+

Nodes（BMs+VMs）

100K+

Pods

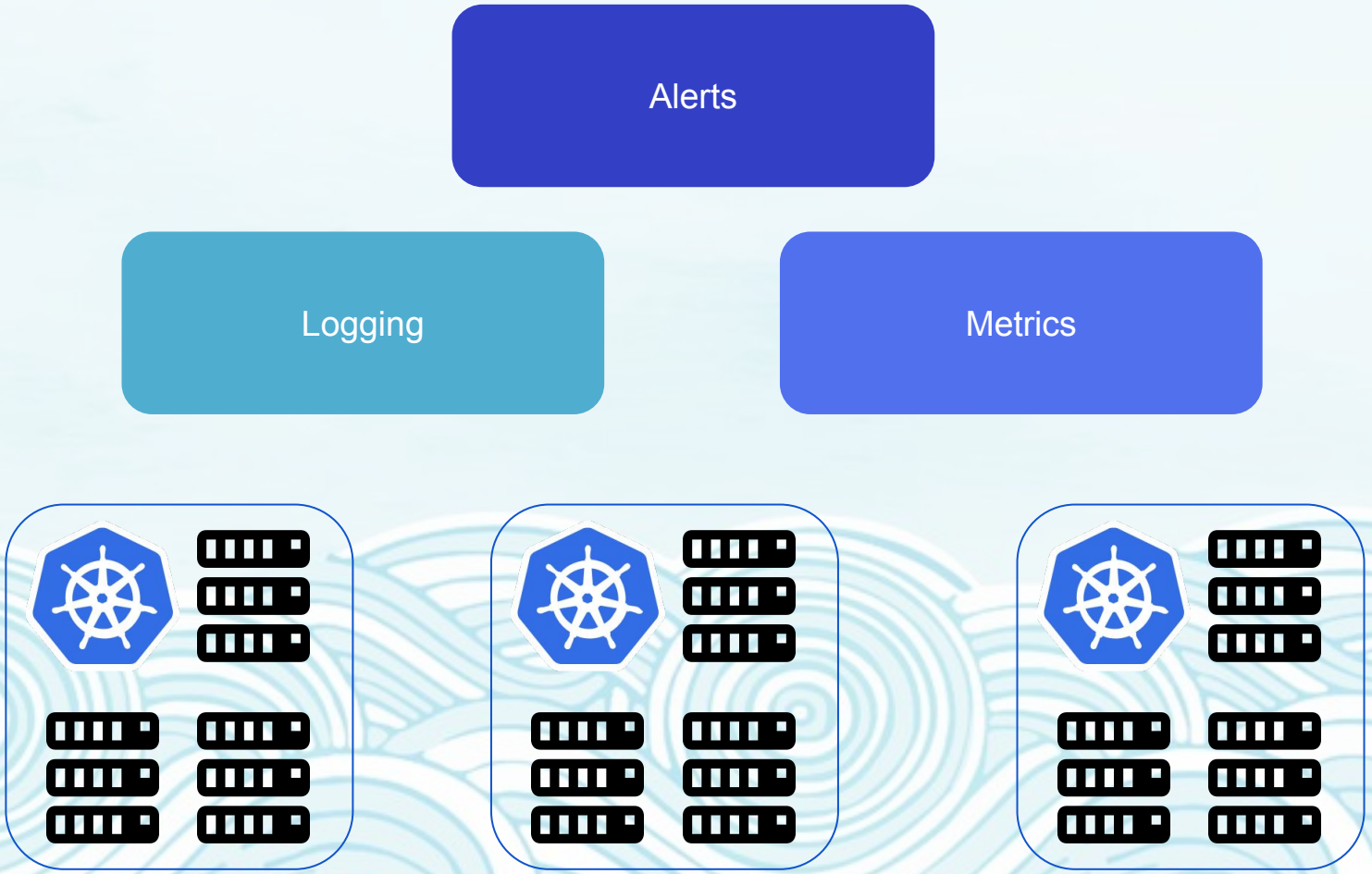




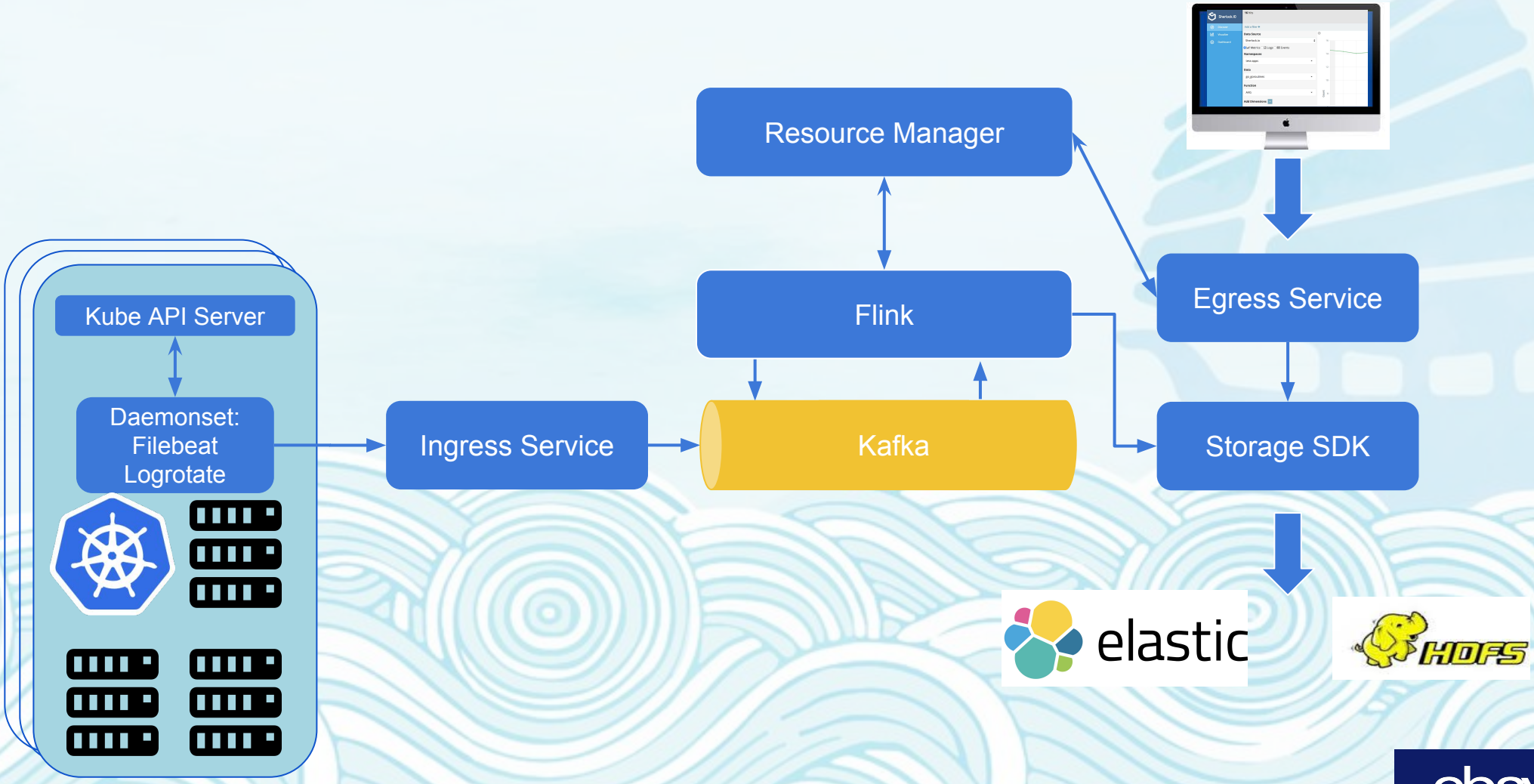
Monitoring Goals

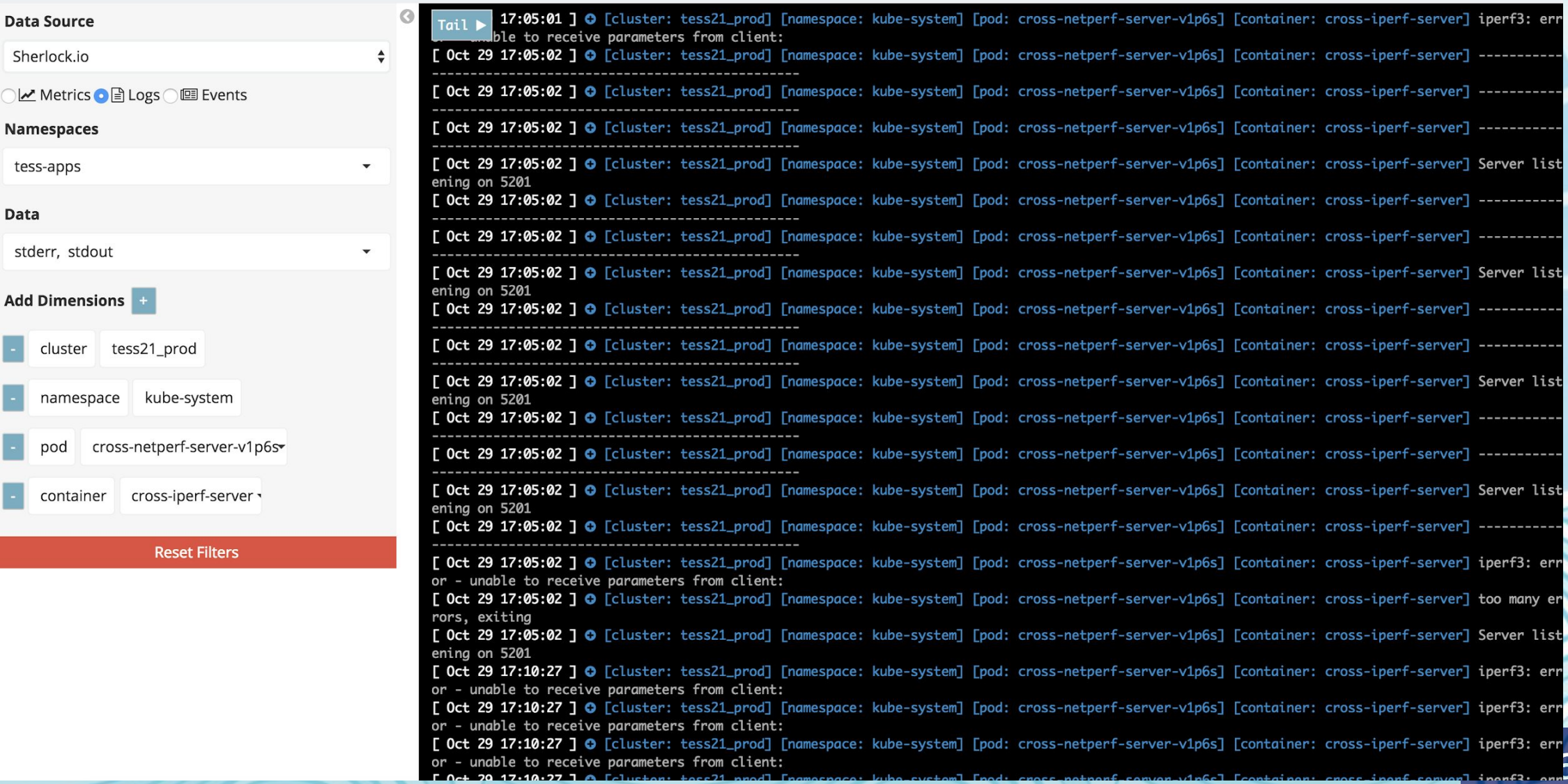
- Control Plane Management
  - Apiserver
  - ETCD
  - Scheduler
  - Controller
- Data Plane Management
  - Node Lifecycle Management
  - Pod Lifecycle Management
  - Daemonset / Deployment / Service / Ingress …
- Alert Management
- AIOps

Monitoring Overview



How we logging



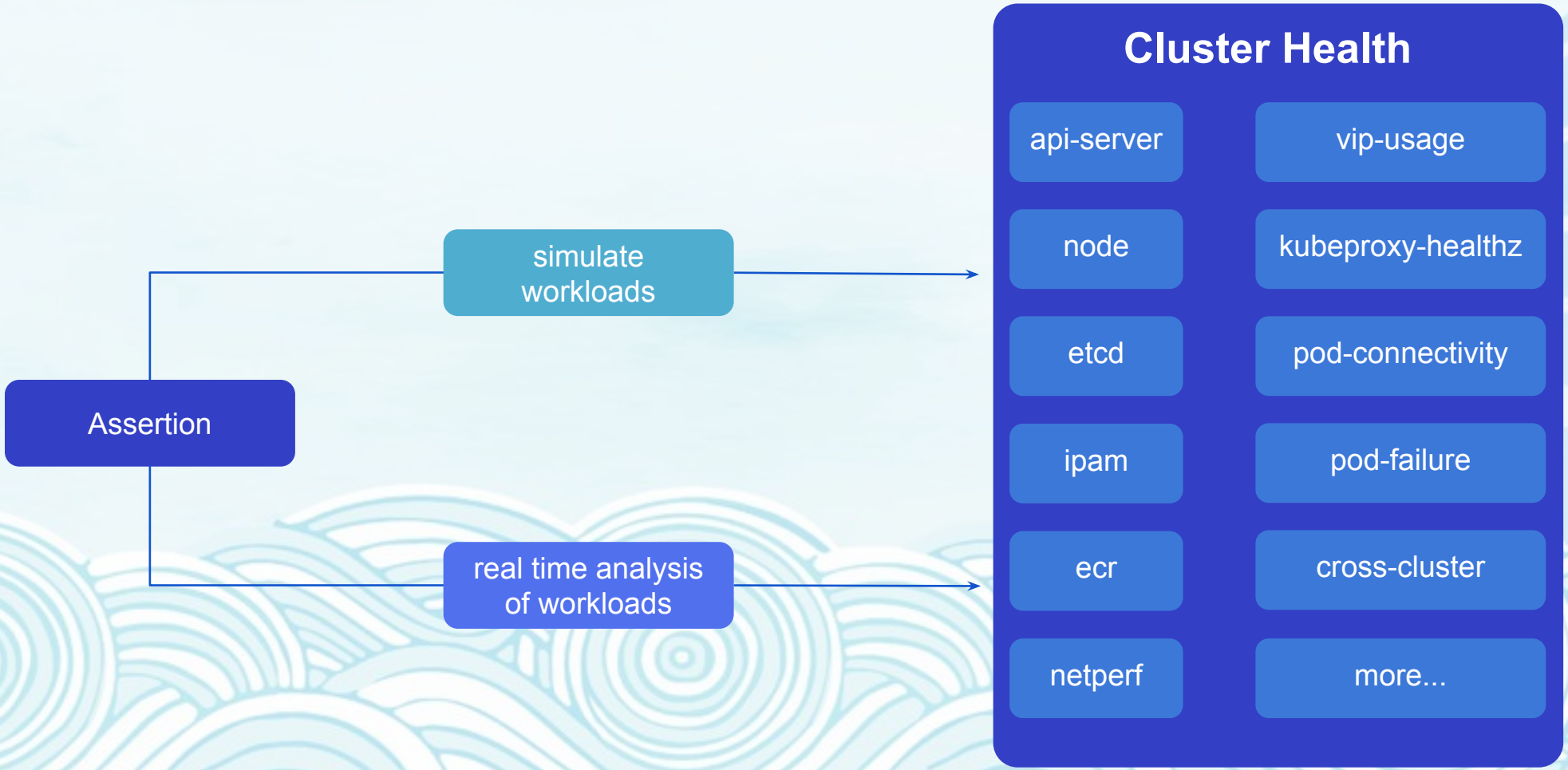

Data Source
Sherlock.io
Metrics Logs Events
Namespaces
tess-apps
Data
stderr, stdout
Add Dimensions
cluster tess21_prod
namespace kube-system
pod cross-netperf-server-v1p6s
container cross-iperf-server
Reset Filters
Tail
[Oct 29 17:05:01] [cluster: tess21_prod] [namespace: kube-system] [pod: cross-netperf-server-v1p6s] [container: cross-iperf-server] iperf3: err
unable to receive parameters from client:
[Oct 29 17:05:02] [cluster: tess21_prod] [namespace: kube-system] [pod: cross-netperf-server-v1p6s] [container: cross-iperf-server] ----------
[Oct 29 17:05:02] [cluster: tess21_prod] [namespace: kube-system] [pod: cross-netperf-server-v1p6s] [container: cross-iperf-server] Server list
ening on 5201
[Oct 29 17:05:02] [cluster: tess21_prod] [namespace: kube-system] [pod: cross-netperf-server-v1p6s] [container: cross-iperf-server] iperf3: err
or - unable to receive parameters from client:
[Oct 29 17:05:02] [cluster: tess21_prod] [namespace: kube-system] [pod: cross-netperf-server-v1p6s] [container: cross-iperf-server] too many er
rors, exiting
[Oct 29 17:10:27] [cluster: tess21_prod] [namespace: kube-system] [pod: cross-netperf-server-v1p6s] [container: cross-iperf-server] iperf3: err
or - unable to receive parameters from client:

How we collect metrics

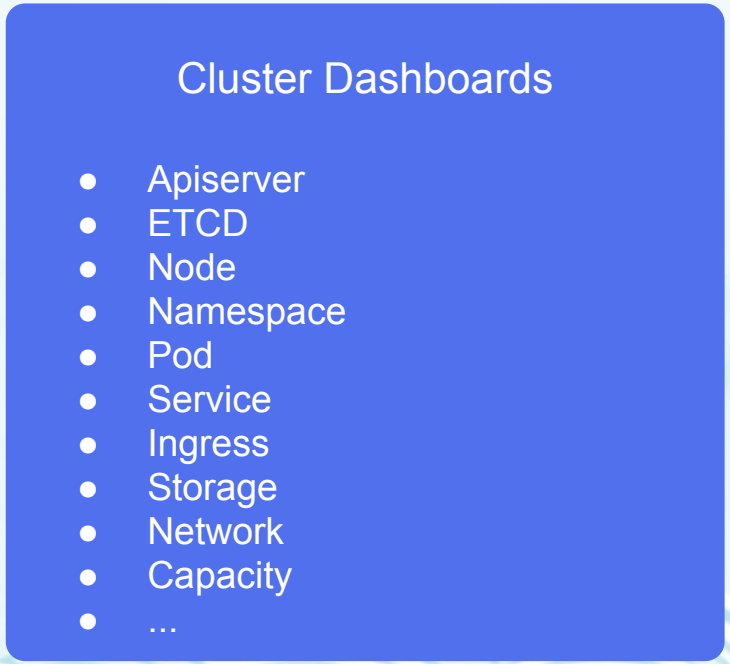
Federated Prometheus

Cluster Prometheus A	Cluster Prometheus B	Cluster Prometheus C
Control Plane Management • Apiserver Latency • Scheduling Latency • IP Allocation Latency • ETCD Latency • ETCD disk usage • Namespace Resource Usage ...	**Node Lifecycle Management** • NotReady Nodes: Amount and Timestamp • SchedulingDisabled Nodes: Amount, Timestamp and Reason • Cpu, Memory, Disk usage • Network Status • PID, FD status ...	**Pod Lifecycle Management** • Pod Creation Latency • Pod Terminating Latency • Pod Restart Times • Pod Resource Usage • Container Creation Latency • Container Terminating Latency • Container Exit Status ...
K8S Key Components	Exporter	Node Problem Detector	Kube State Metrics	Assertion	...

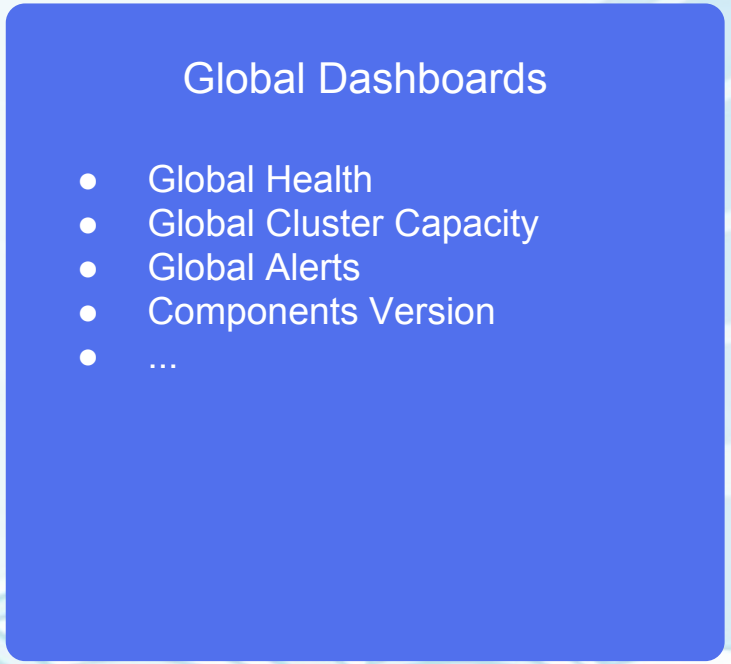
Assertion


Assertion
simulate workloads
real time analysis of workloads
Cluster Health
api-server
vip-usage
node
kubeproxy-healthz
etcd
pod-connectivity
ipam
pod-failure
ecr
cross-cluster
netperf
more...

How we build dashboards



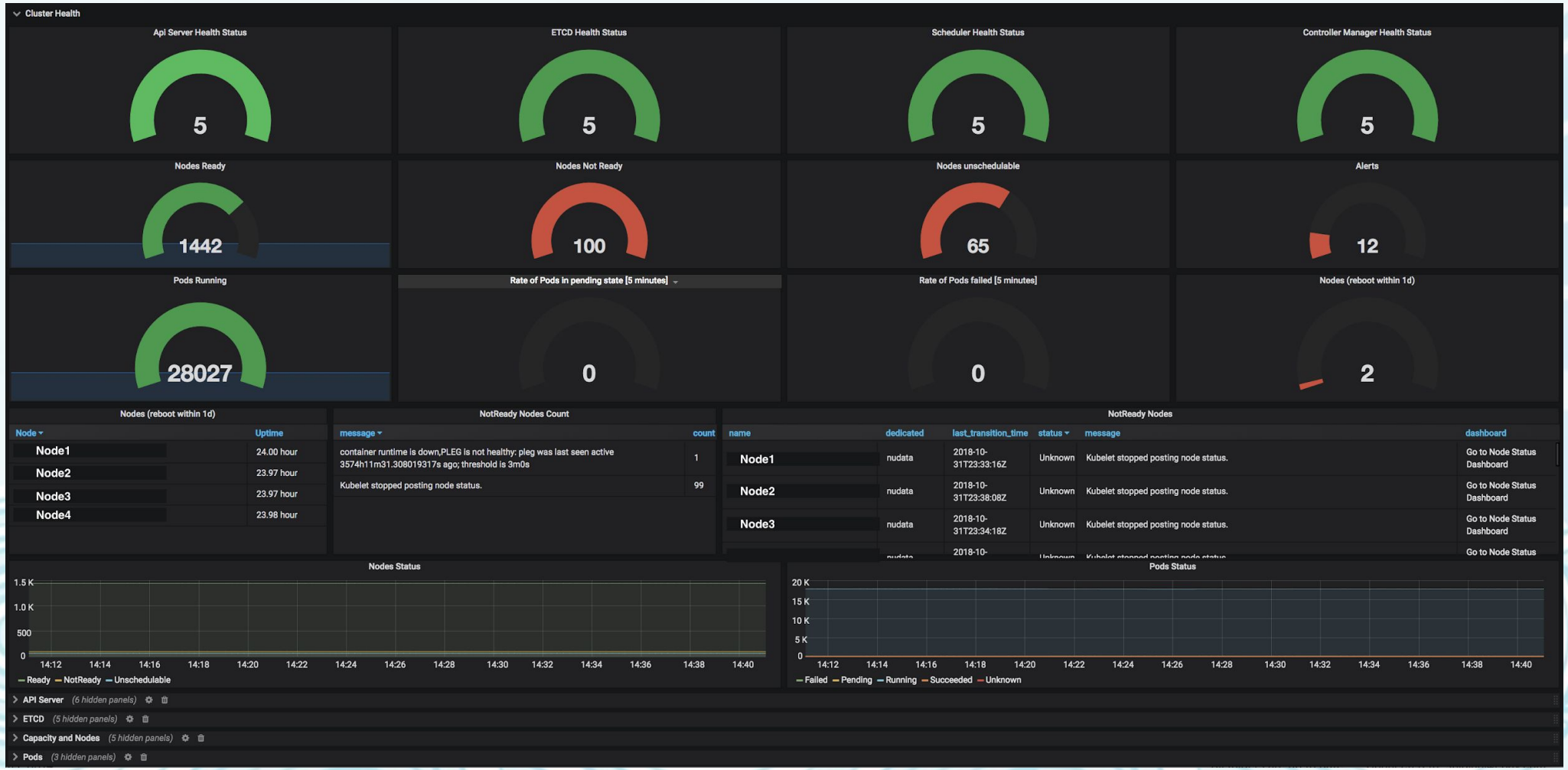
Cluster Prometheus



Federated Prometheus

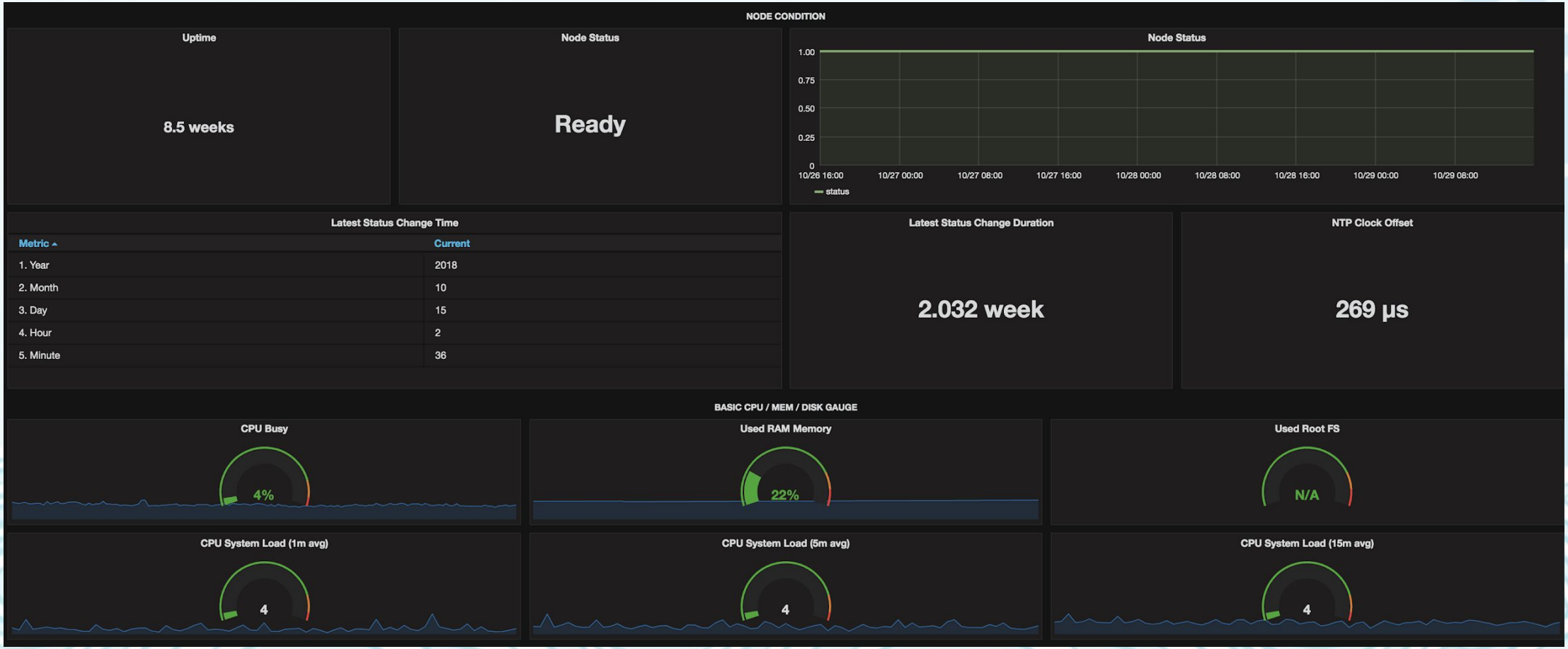
Global Health Dashboard

Cluster Health Dashboard



ebay

Node Health Dashboard



How we manage alerts

Alert Rules

Labels

Component

Severity

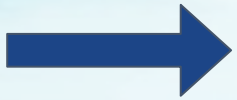
Annotations

Description

Summary

Runbook

Cluster



Grafana

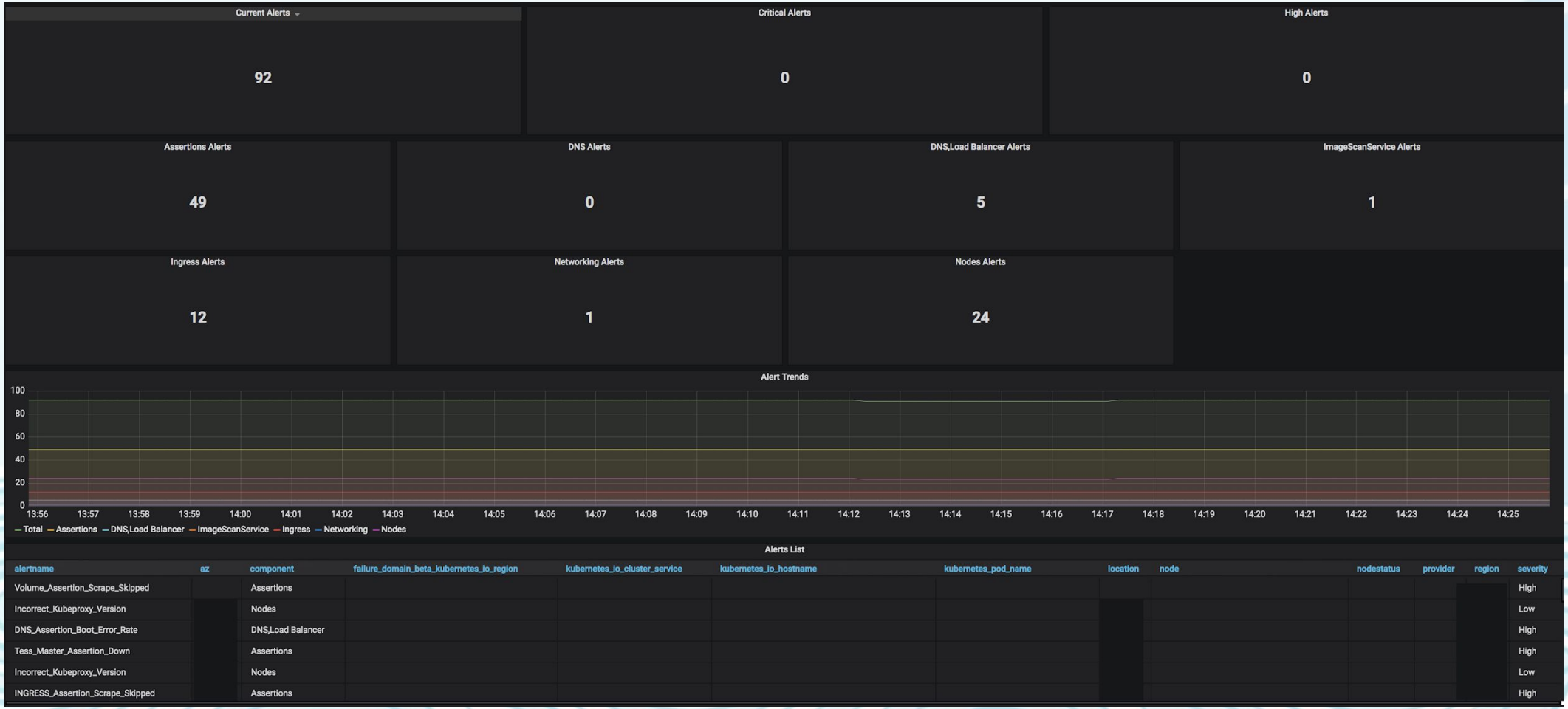
Cluster Alert Dashboard
Global Alert Dashboard

pagerduty

Jira

- RCA
- Execution Plan
- Time Consumption

Global Alert Dashboard

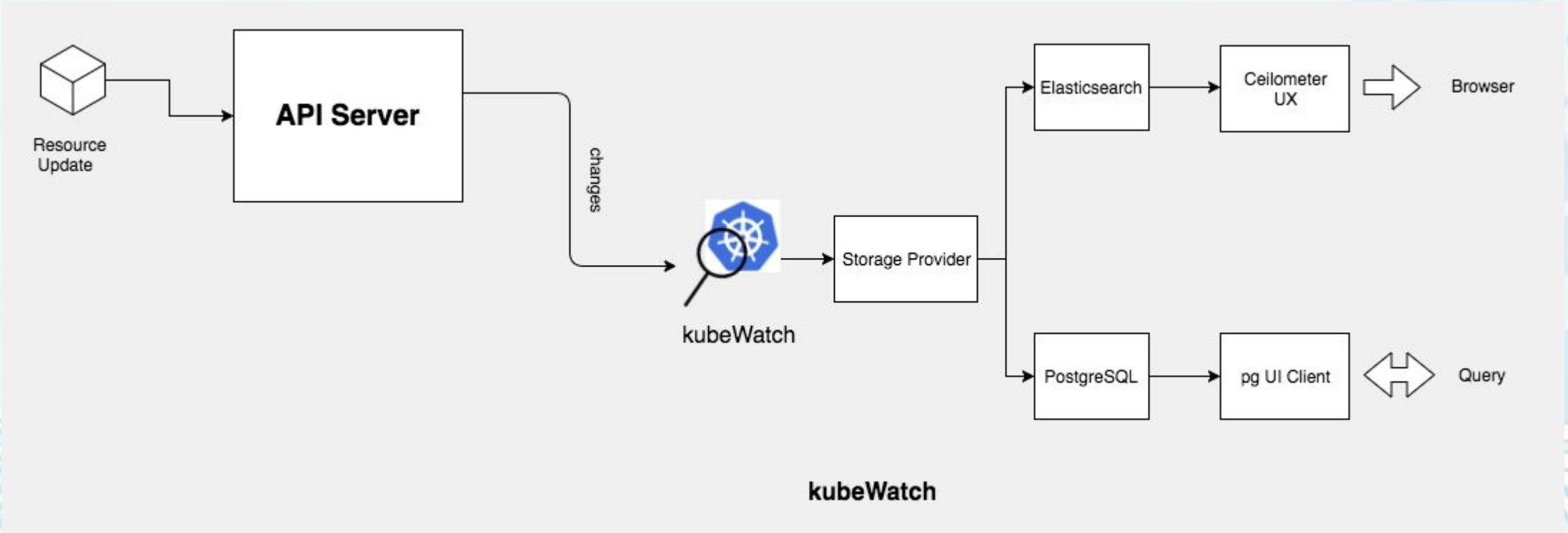


KubeWatch

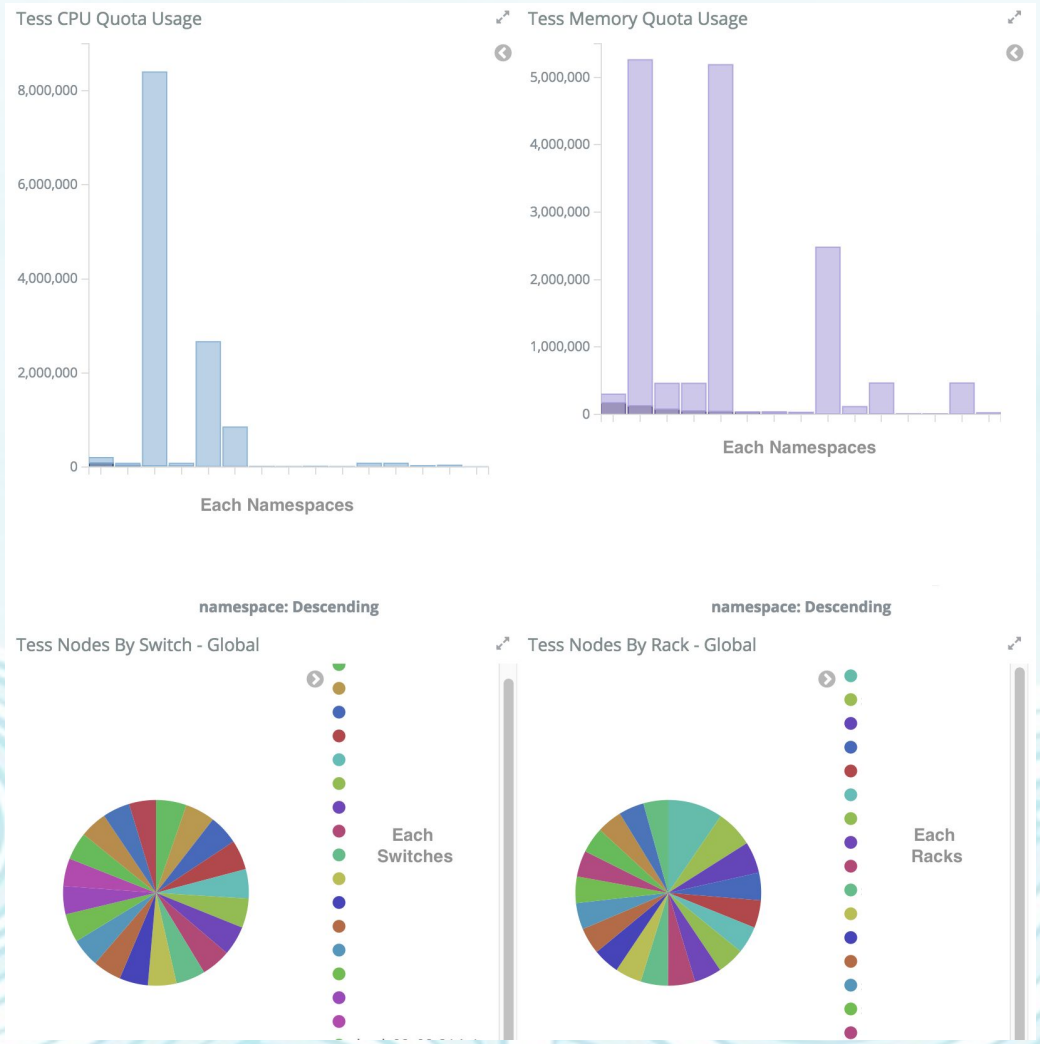
How to audit across clusters?

How to execute complex queries quickly?

KubeWatch Architecture



KubeWatch



Query Example:

1. **Get all pods for namespace kube-system**

```
SELECT
  name AS podName,
  data->'metadata'->>'namespace' AS namespace ,
  data->'spec'->>'nodeName' AS nodeName ,
  meta,
  data
FROM pods
WHERE data->'metadata'->>'namespace'
      LIKE 'kube-system'
      AND deleted = FALSE;
```

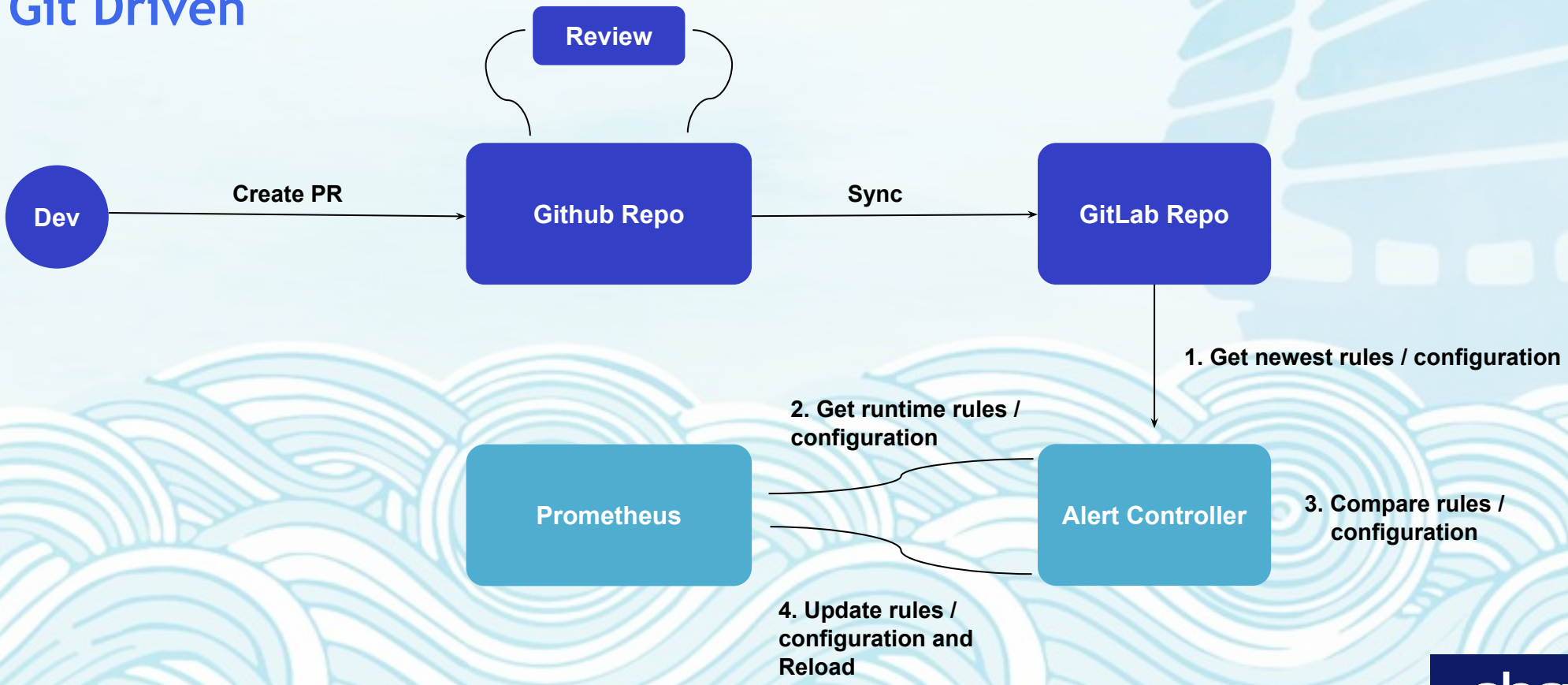
2. **Get all services of type load balancer**

```
SELECT
  name AS serviceName,
  data->'spec'->>'type' AS type,
  data->'metadata'->>'namespace' AS nameSpace
FROM svcs
WHERE data->'spec'->>'type'
      LIKE 'LoadBalancer';
```

ebay

Monitoring Automation

- Rollout alert rules
- Rollout monitoring configurations
- Git Driven



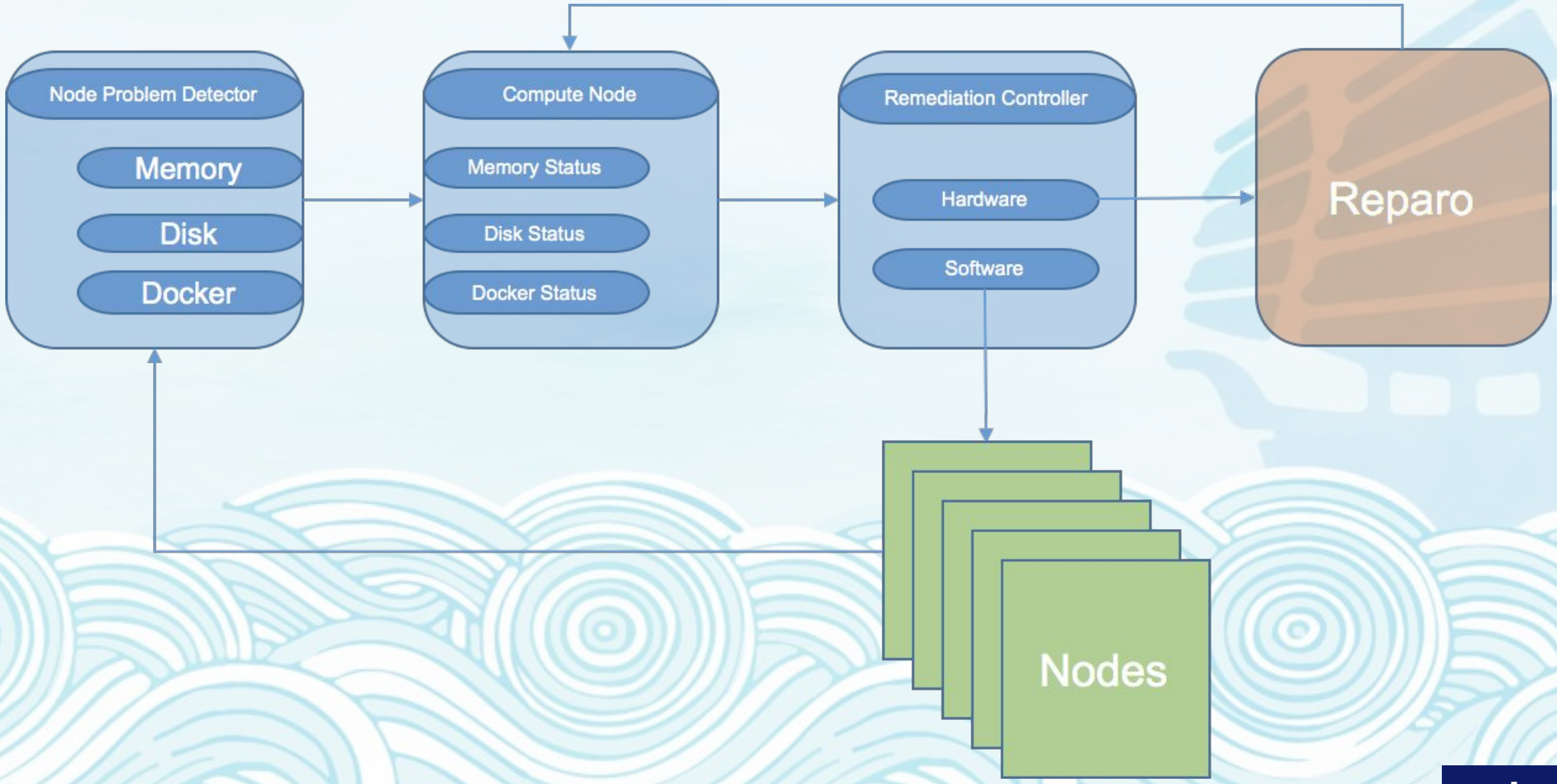
AIOps

- Real-time analysis and alerts
- Reducing MTTD and MTTR

We have

- More than 30 clusters
- Thousands of nodes
- Including both BMs and VMs

Overall


Node Problem Detector
Memory
Disk
Docker
Compute Node
Memory Status
Disk Status
Docker Status
Remediation Controller
Hardware
Software
Reparo
Nodes

Hardware Failures

Sensors (Non intrusive detection)

- TEMP
- CPU
- PSU
- MEMORY
- VOLT
- HDD
- FAN

In OS (intrusive detection)

- Kernel message
- MCE message
- Disk check

- Define each pattern to correspond with each known failure
- More patterns can be defined if new failure found
- Check if the failure can be tolerant
- Mark the hardware as failures
- Get notified if hardware issues have been fixed

Software Failures

Health check for key components:

Kube*

Configurations

Container Runtime

Key Services

Kernel soft lockup

etc.

KubeCon

CloudNativeCon

China 2018

Thank You !

KubeCon

CloudNativeCon

China 2018

Q&A