Enterprise Kubernetes Logging & Monitoring

Observability in VMware PKS

Denny Zhang

Staff Engineer

VMware, Cloud Native BU

11/12/2018

About Me

Kubernetes

DevOps

Blogger

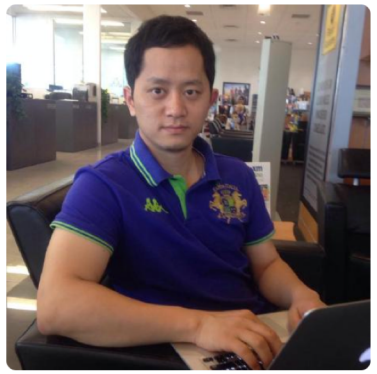
OpenSource ?

Denny Zhang
Cloud Native & Blogger || 23K+ IT Connections
VMware • Sun Yat-Sen University
San Francisco Bay Area • 500+

Slack: https://www.dennyzhang.com/slack
I welcome connections from DevOps/Ops/Agile....

Articles & activity
23,679 followers | Manage followers

See all connections



DennyZhang - DevOps & Cloud Native
dennyzhang

Slack:
https://dennyzhang.com/slack

Edit bio

Developer Program Member

dennyzhang.com (张巍)
MountainView, CA
denny.zhang001@gmail.com
https://www.dennyzhang.com

Organizations

Overview | Repositories 86 | Stars 91 | Followers 354 | Following 78

Pinned repositories — Customize your pinned repositories

challenges-kubernetes
Challenges Your Kubernetes Skills And Knowledge
Shell ★ 161 72

cheatsheet-kubernetes-A4
Kubernetes CheatSheets In A4
★ 55 41

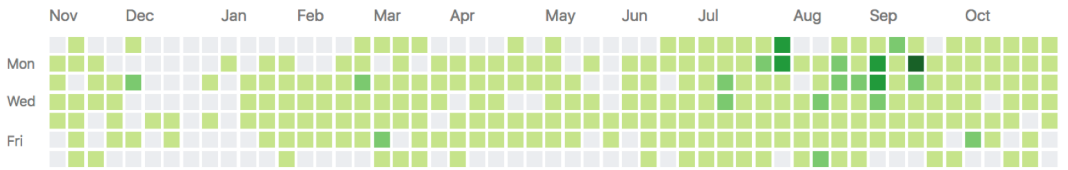
cheatsheet.dennyzhang.com
Apply best practices via CheatSheets
Shell ★ 8 8

challenges-cloudformation-jenkins
Challenges Your AWS And Cloudformation Skills By Solving Real Questions.
Shell ★ 40 44

devops_public
DevOps Scripts
Shell ★ 92 63

Denny-s-emacs-configuration
Emacs shapes me to be a better programmer
Emacs Lisp ★ 37 5

2,719 contributions in the last year — Contribution settings



GitHub LinkedIn Blog

Agenda

- Problem Space
- Existing Solutions
- VMware Solution
- Q & A

Problem Space

Business values for logging and monitoring

Trouble Shooting

Monitoring

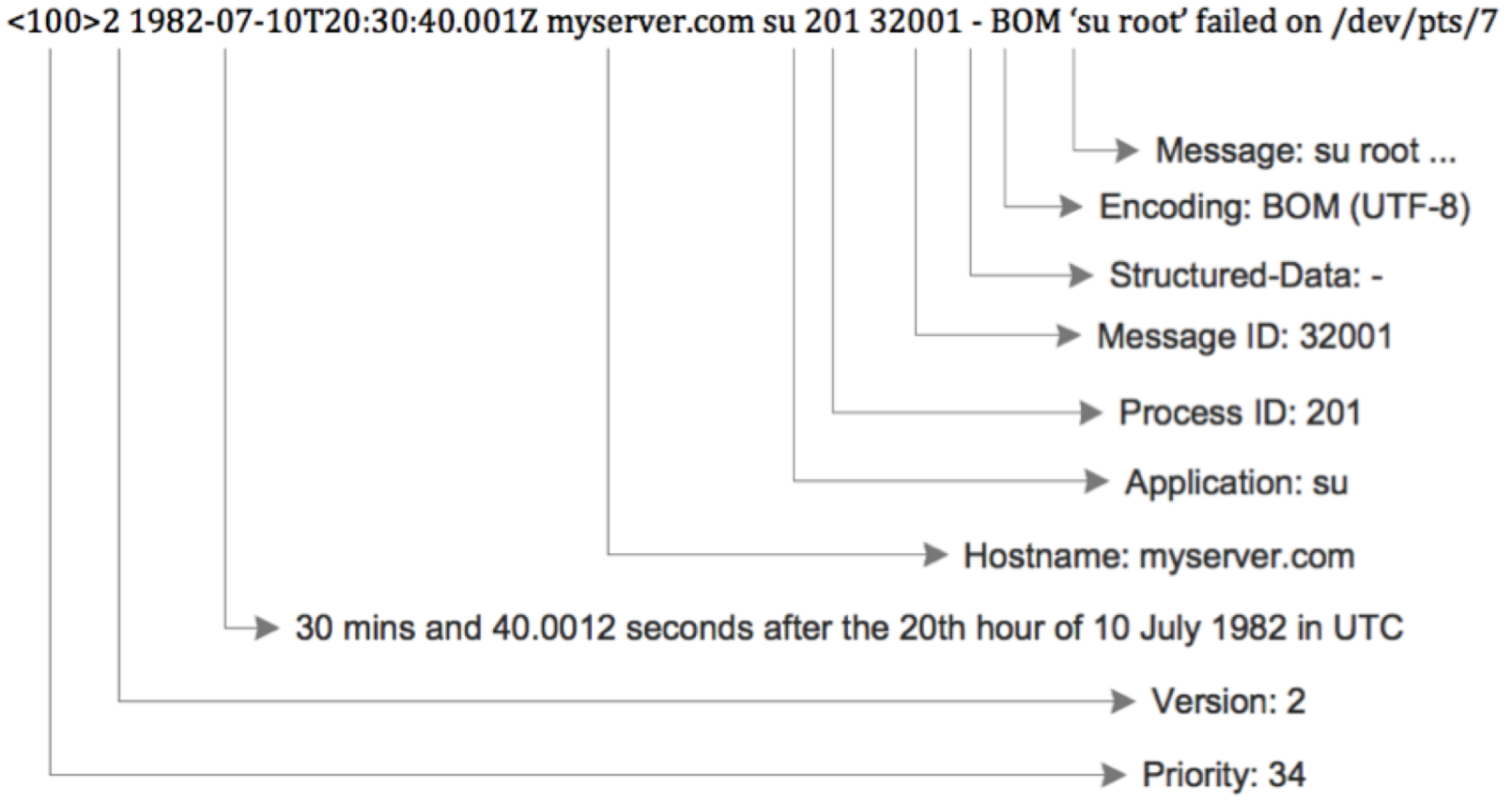
Resource Utilization

Audit & Compliance

Data Models In Logging & Monitoring

Data Models Convention

Log format: RFC-5424



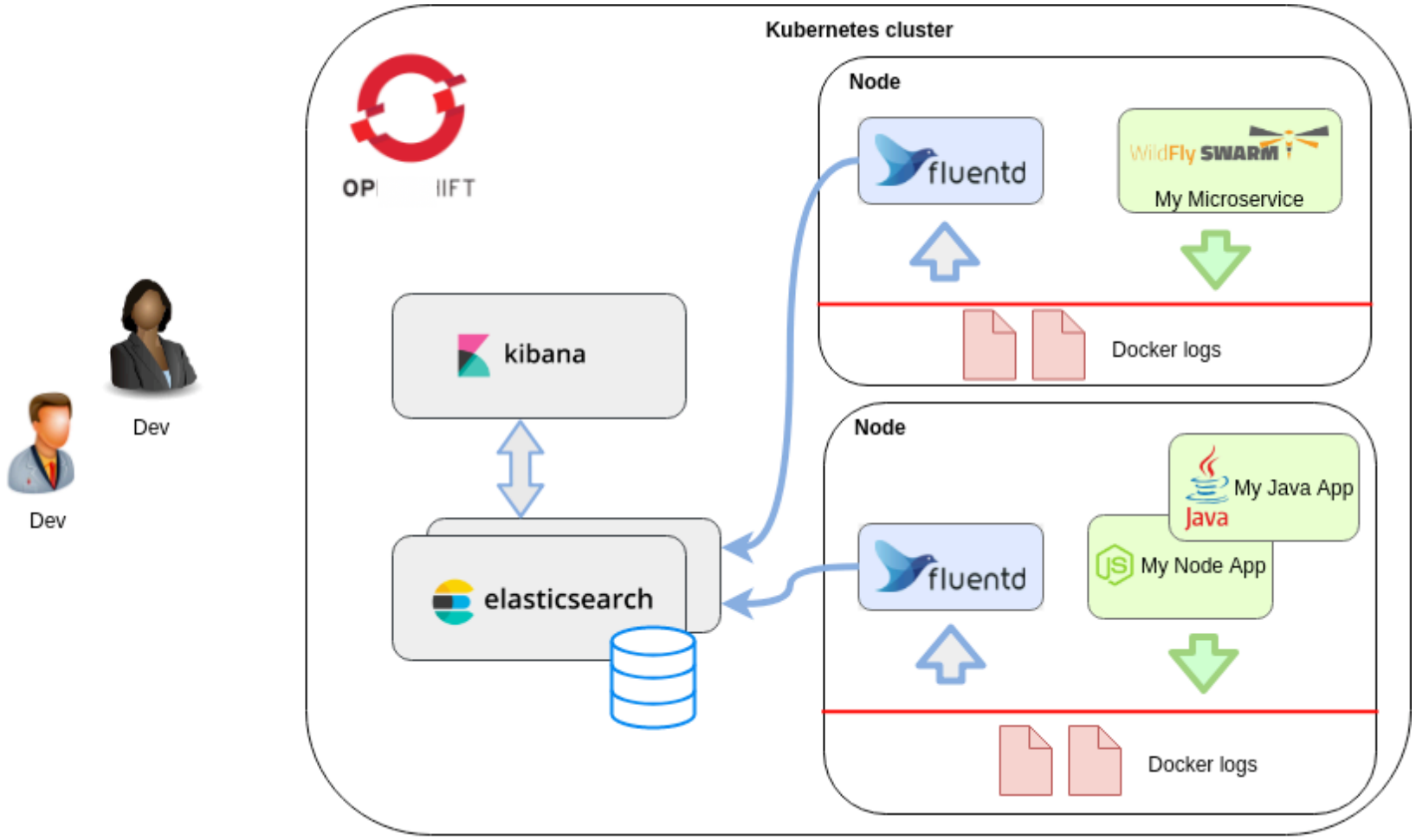
Data Models Convention

Metric format: Prometheus metrics? Kubernetes Open Metrics?

```
# HELP http_requests_total The total number of HTTP requests.
# TYPE http_requests_total counter
http_requests_total{method="post",code="200"} 1027 1395066363000
http_requests_total{method="post",code="400"}    3 1395066363000

# Escaping in label values:
msdos_file_access_time_seconds{path="C:\\DIR\\FILE.TXT",error="Cannot find file:\n\"FILE.TXT\""} 1.4582559

# Minimalistic line:
metric_without_timestamp_and_labels 12.47

# A weird metric from before the epoch:
something_weird{problem="division by zero"} +Inf -3982045

# A histogram, which has a pretty complex representation in the text format:
# HELP http_request_duration_seconds A histogram of the request duration.
# TYPE http_request_duration_seconds histogram
http_request_duration_seconds_bucket{le="0.05"} 24054
http_request_duration_seconds_bucket{le="0.1"} 33444
http_request_duration_seconds_bucket{le="0.2"} 100392
http_request_duration_seconds_bucket{le="0.5"} 129389
http_request_duration_seconds_bucket{le="1"} 133988
http_request_duration_seconds_bucket{le="+Inf"} 144320
http_request_duration_seconds_sum 53423
http_request_duration_seconds_count 144320
```

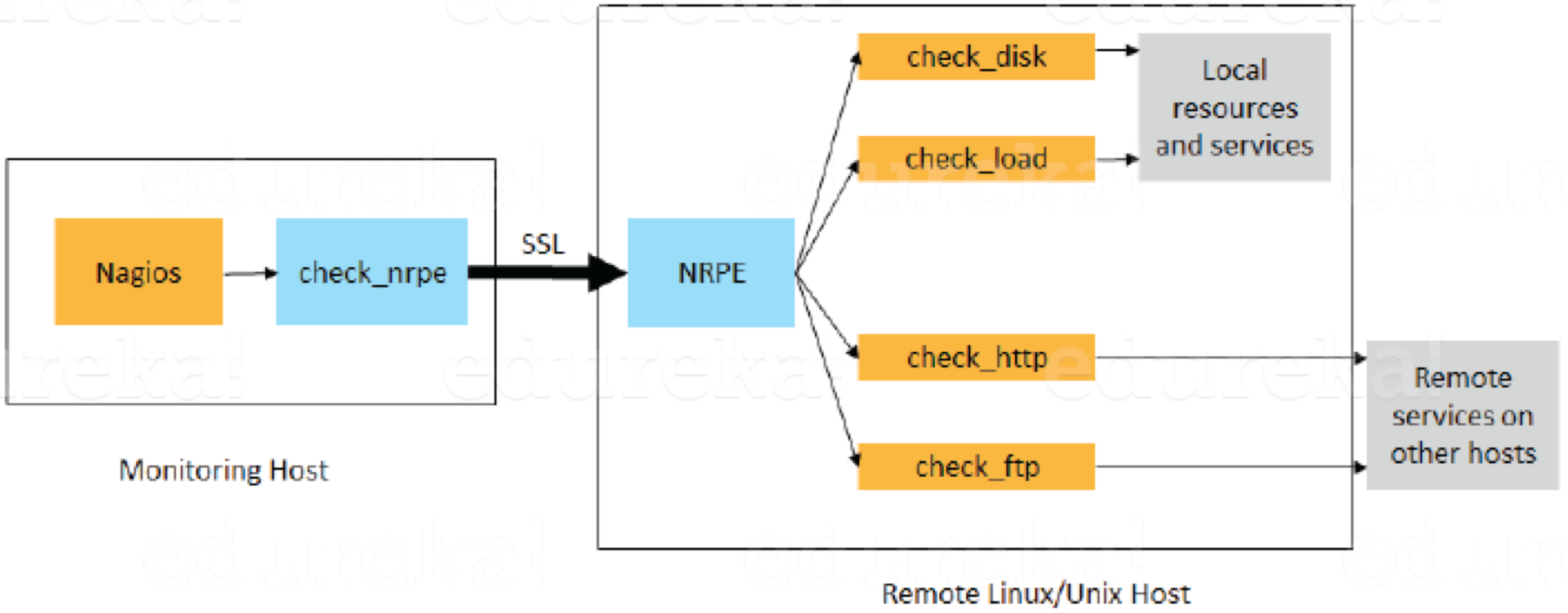
Existing Container Logging Solution

Fluentd: tail log file



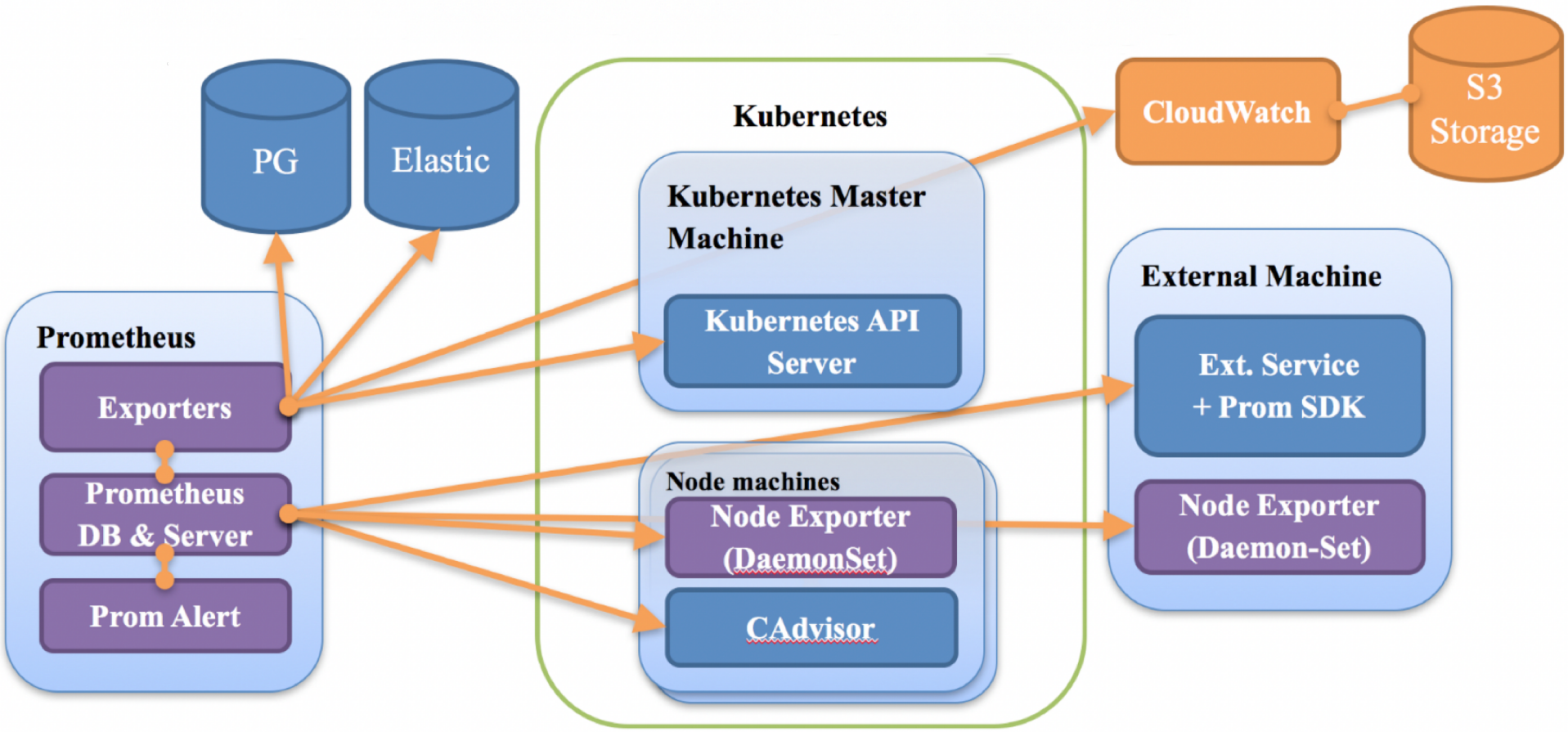
Existing Container Monitoring Solution

Nagios/Zabbix: agent base



Existing Container Monitoring Solution

Prometheus: exporter base + pull model



Problem Solved?

New Problems When Moving To Container World

The foundation has changed!

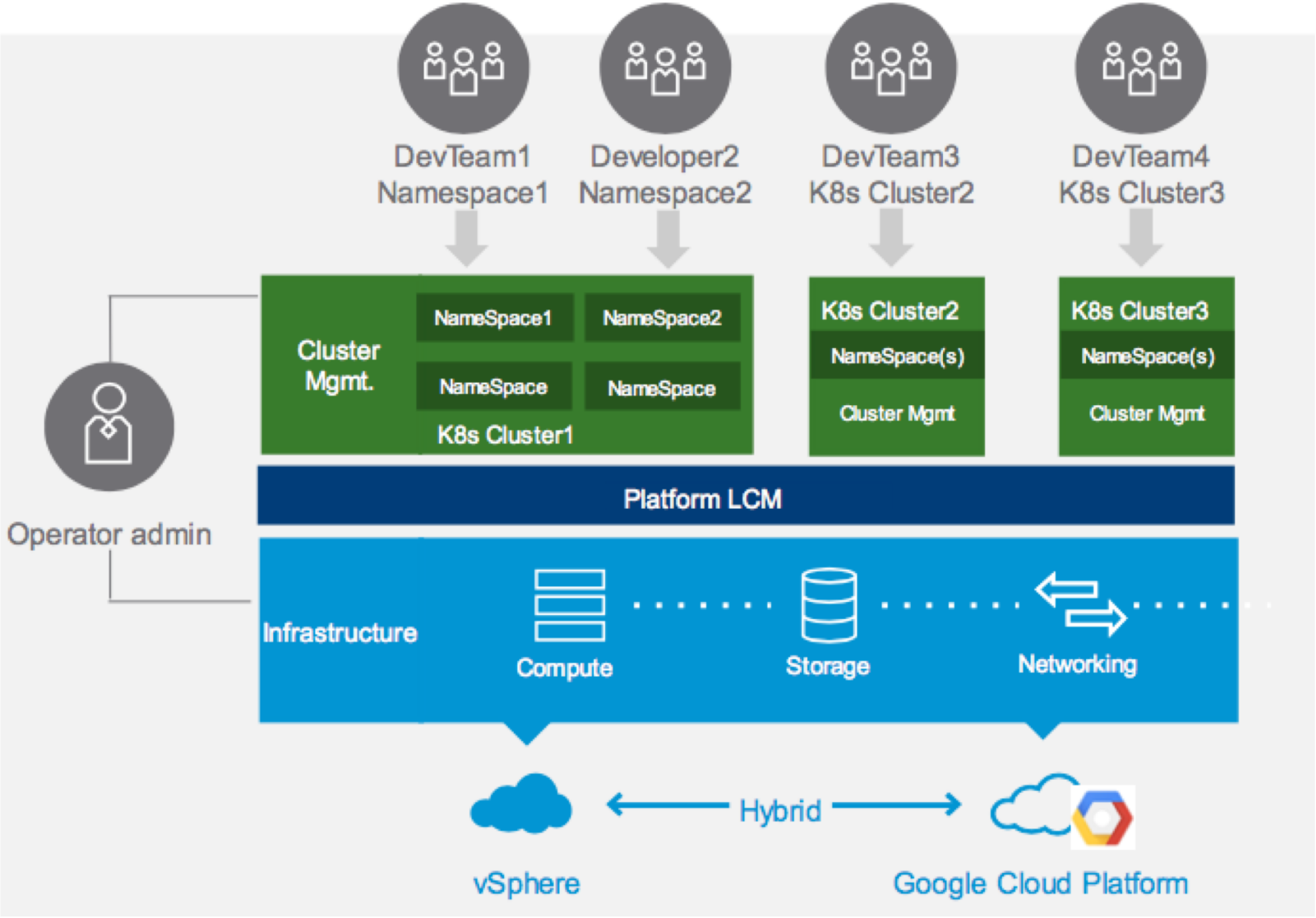
	VM World	Container World
When process die/restart	Log file still there	Pod deleted, log file deleted
When process run	Write log to file	Log prints to stdout/stderr
When process deploy	One service for one VM	Co-located, shared, elastic

Challenges

Logging – Horizontal partition

Can’t mix:

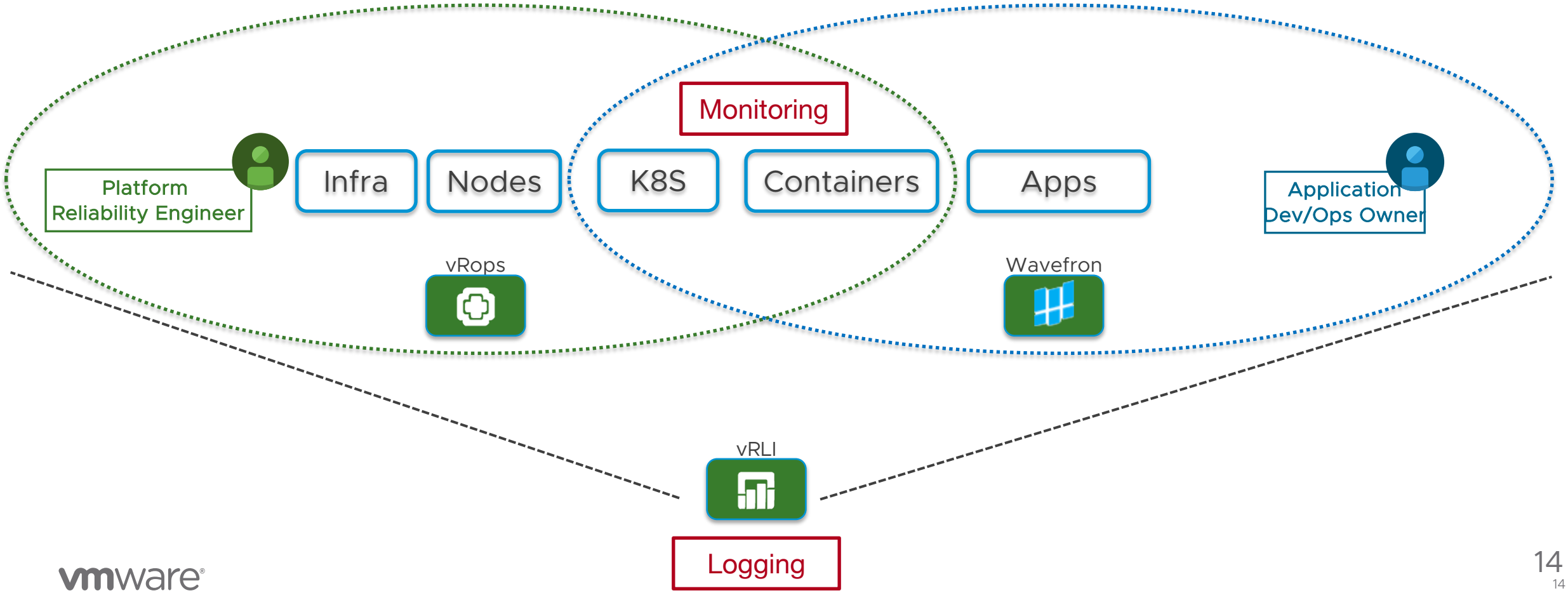
- Different k8s clusters
- Different namespaces



Challenges

Logging - Vertical partition

- Data sources from different layers
- Different layers have different domain knowledgebase



Challenges

Fail to get all unnecessary logs

Log-agent is too slow for pod events

Tree: b593e6f4ad ▾ fluentd / example / in_tail.conf

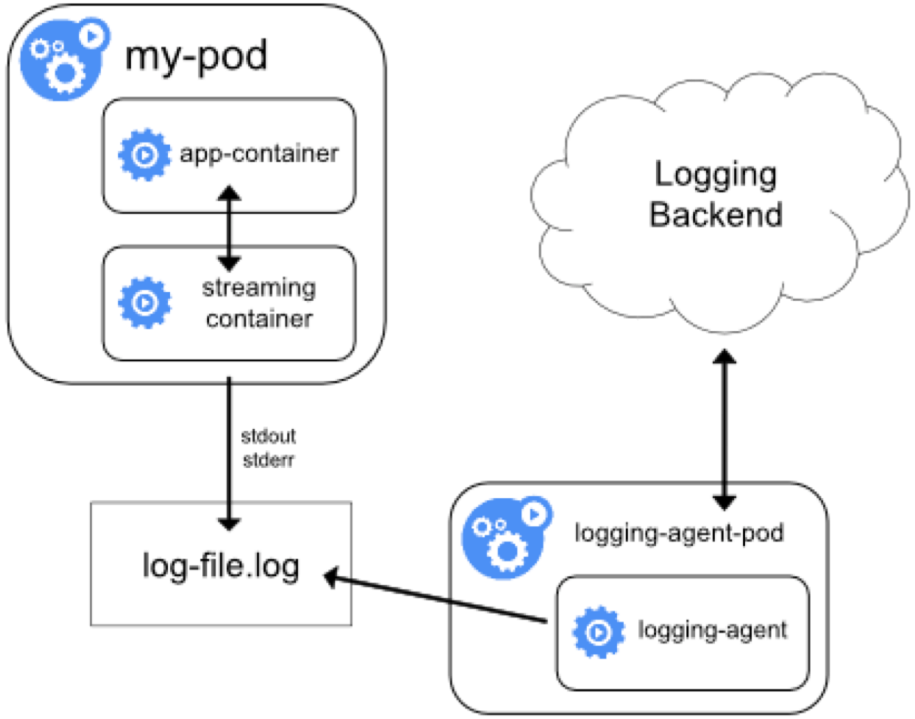
cosmo0920 example: Use `@type` in example confs

2 contributors

15 lines (13 sloc) | 225 Bytes

```
1   <source>
2     @type tail
3     format none
4     path /var/log/fluentd_test.log
5     pos_file /var/log/fluentd_test.pos
6     tag test
7     rotate_wait 5
8     read_from_head true
9     refresh_interval 60
10  </source>
11
12  <match test>
13    @type stdout
14  </match>
```

Not all log streamed to pod stdout/stderr



Challenges

Fail to get log fast enough

- Co-located: Lots of Pods in one k8s worker VM
- Noisy neighbors: Back-pressure

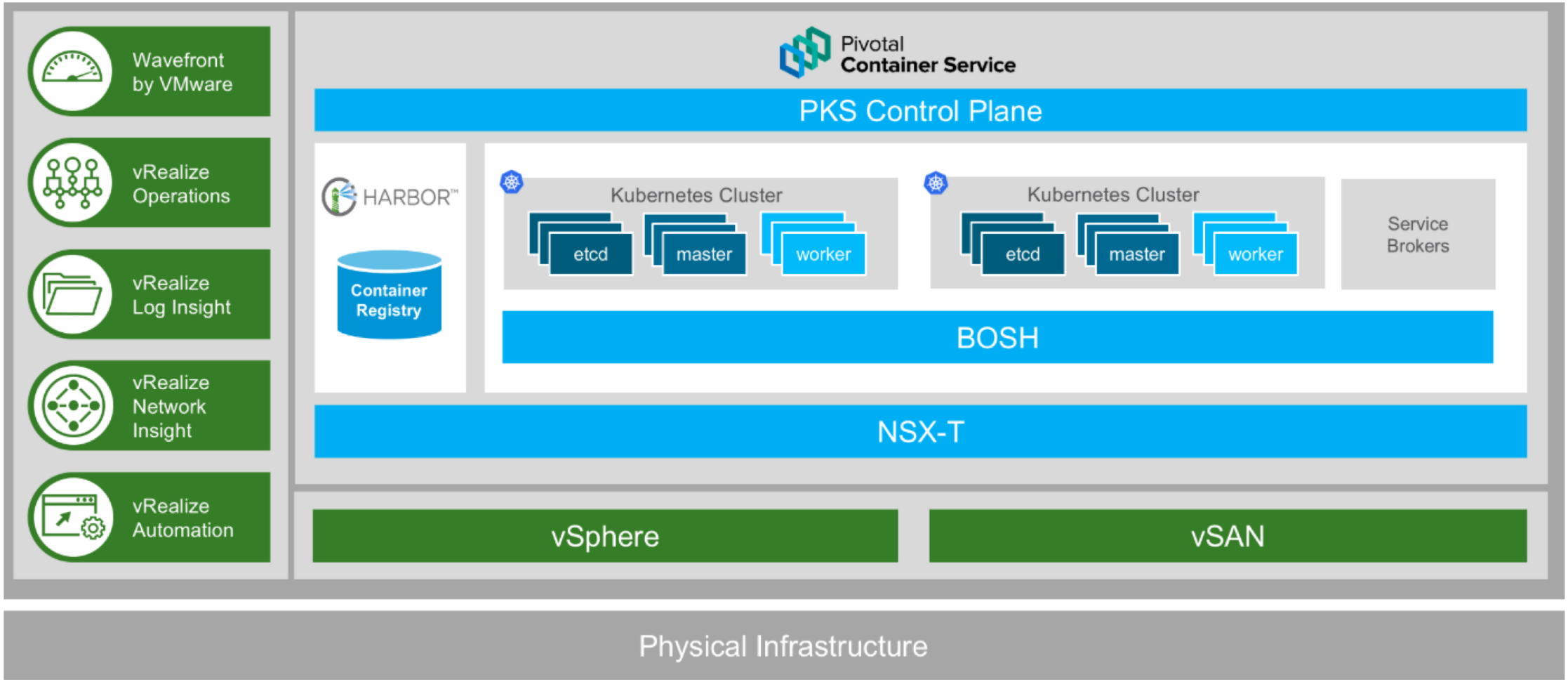
VMware Solution

Benefits we want to achieve

- Built-in VMware Solution for administrators
- Namespace multi-tenancy for developers
- Bring your own solutions

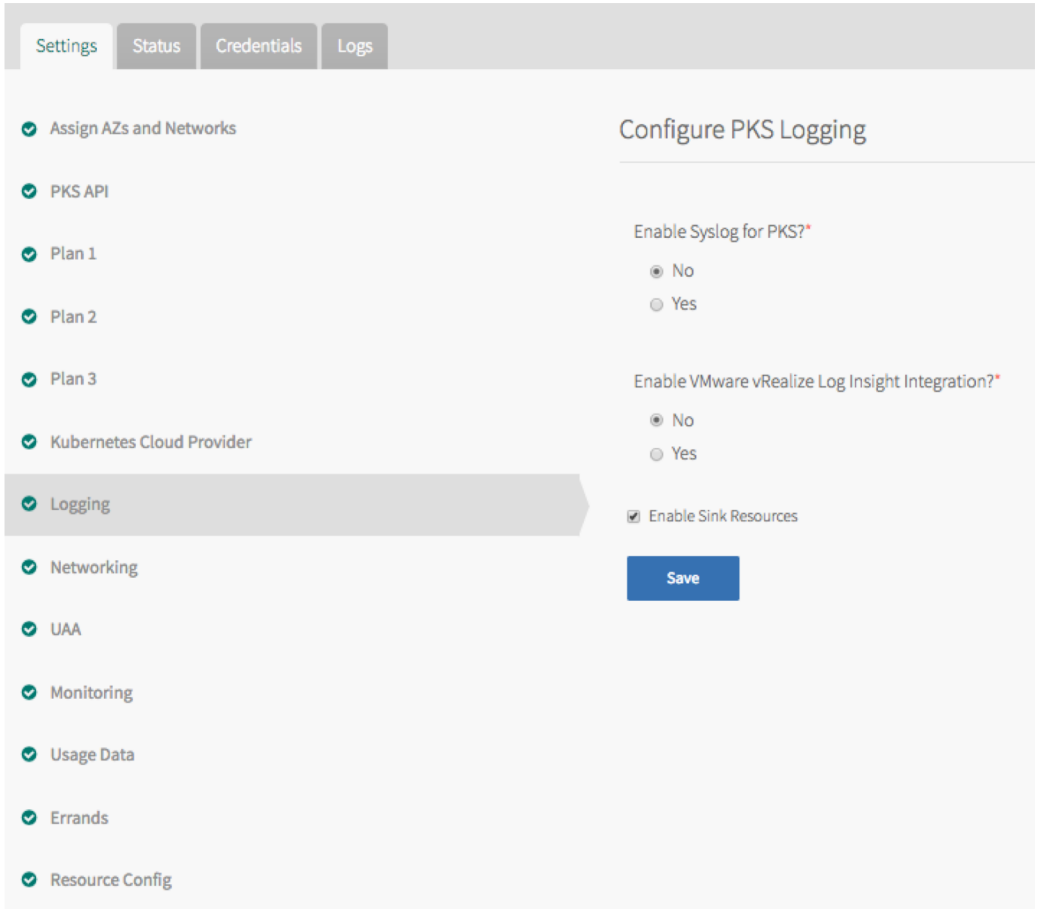
PKS Architecture

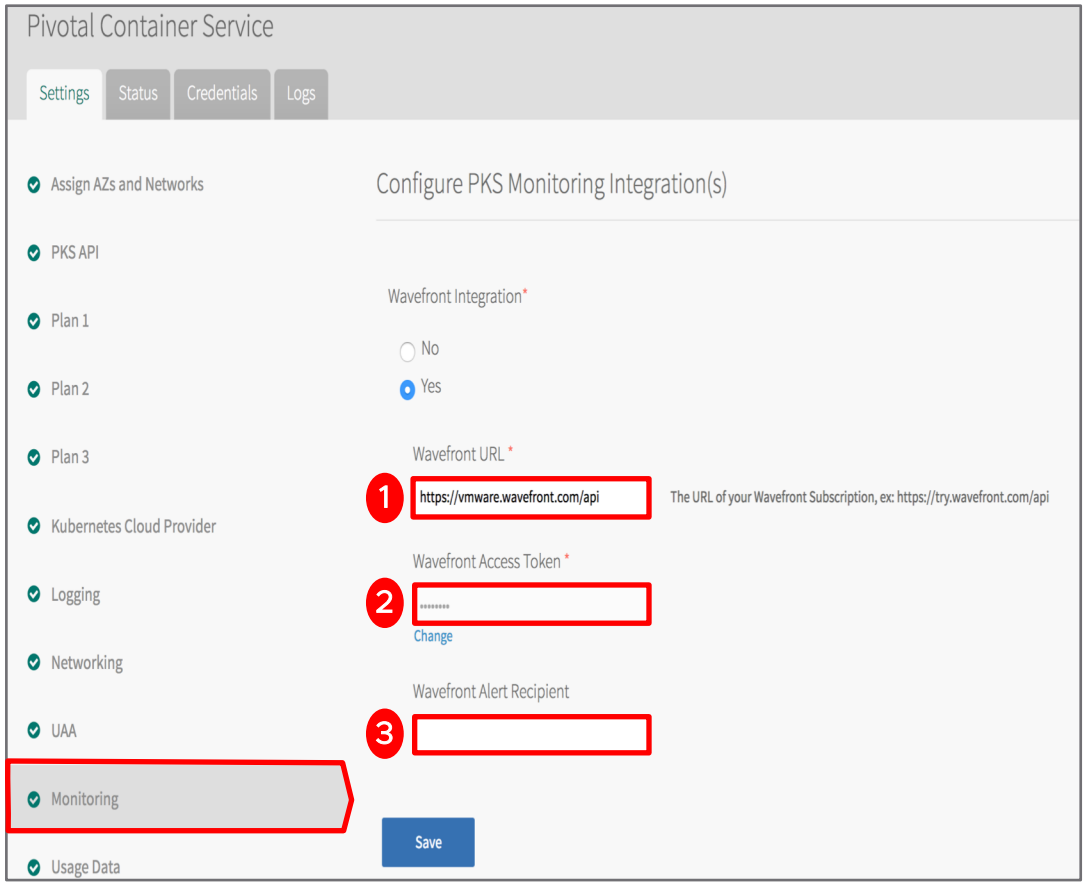
See the big picture



VMware Logging & Monitoring

Easy to configure + work out-of-box





vmware

Workload Logging & Monitoring

Configure in k8s cluster level and namespace level

kubectl apply –f ns-sink.yml

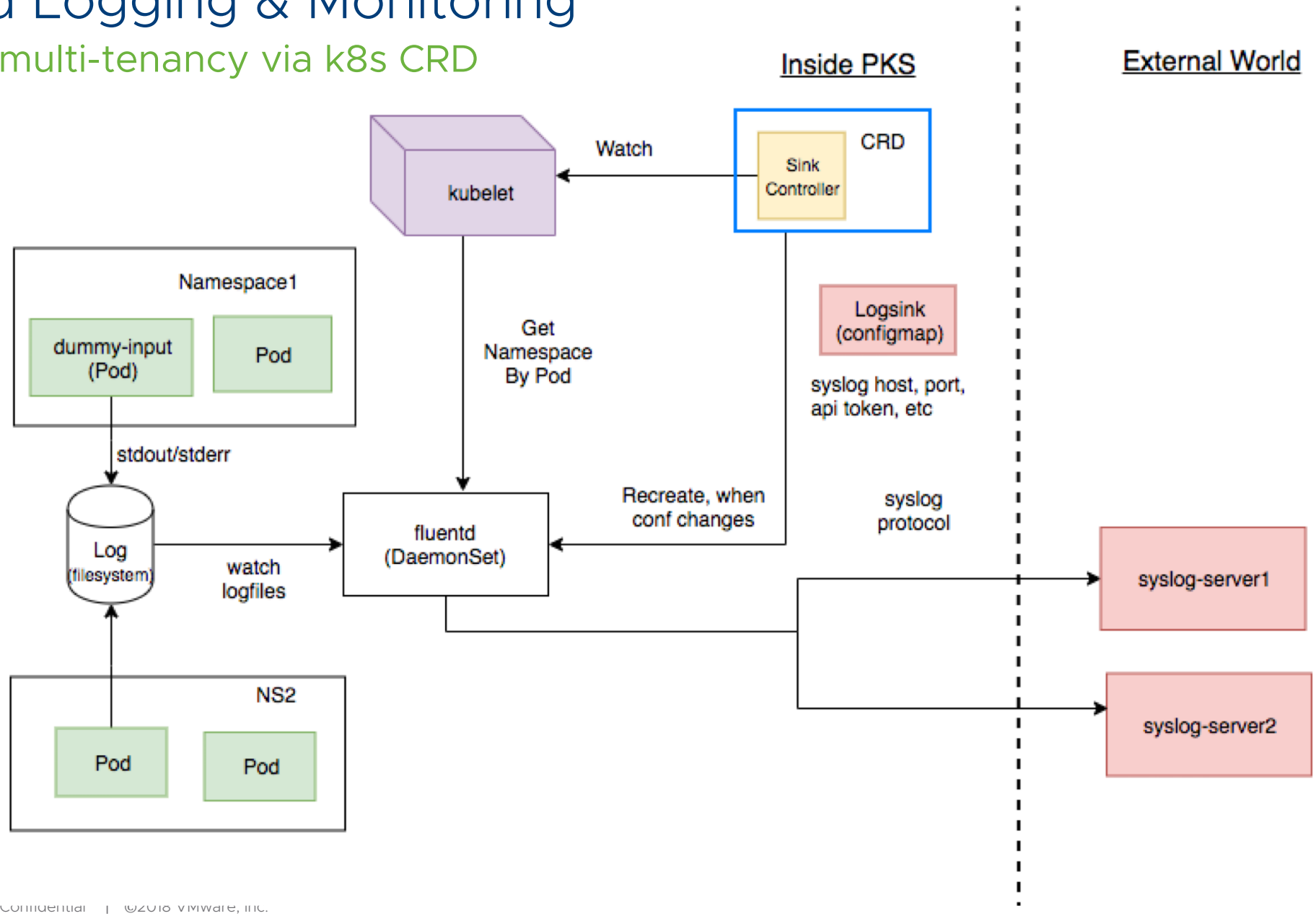
```
apiVersion: apps.pivotal.io/v1beta1
kind: Sink
metadata:
  name: YOUR-SINK
  namespace: YOUR-NAMESPACE
spec:
  type: syslog
  host: YOUR-LOG-DESTINATION
  port: YOUR-LOG-DESTINATION-PORT
  enable_tls: true
```

kubectl apply –f cluster-sink.yml

```
apiVersion: apps.pivotal.io/v1beta1
kind: ClusterSink
metadata:
  name: YOUR-SINK
spec:
  type: syslog
  host: YOUR-LOG-DESTINATION
  port: YOUR-LOG-DESTINATION-PORT
  enable_tls: true
```

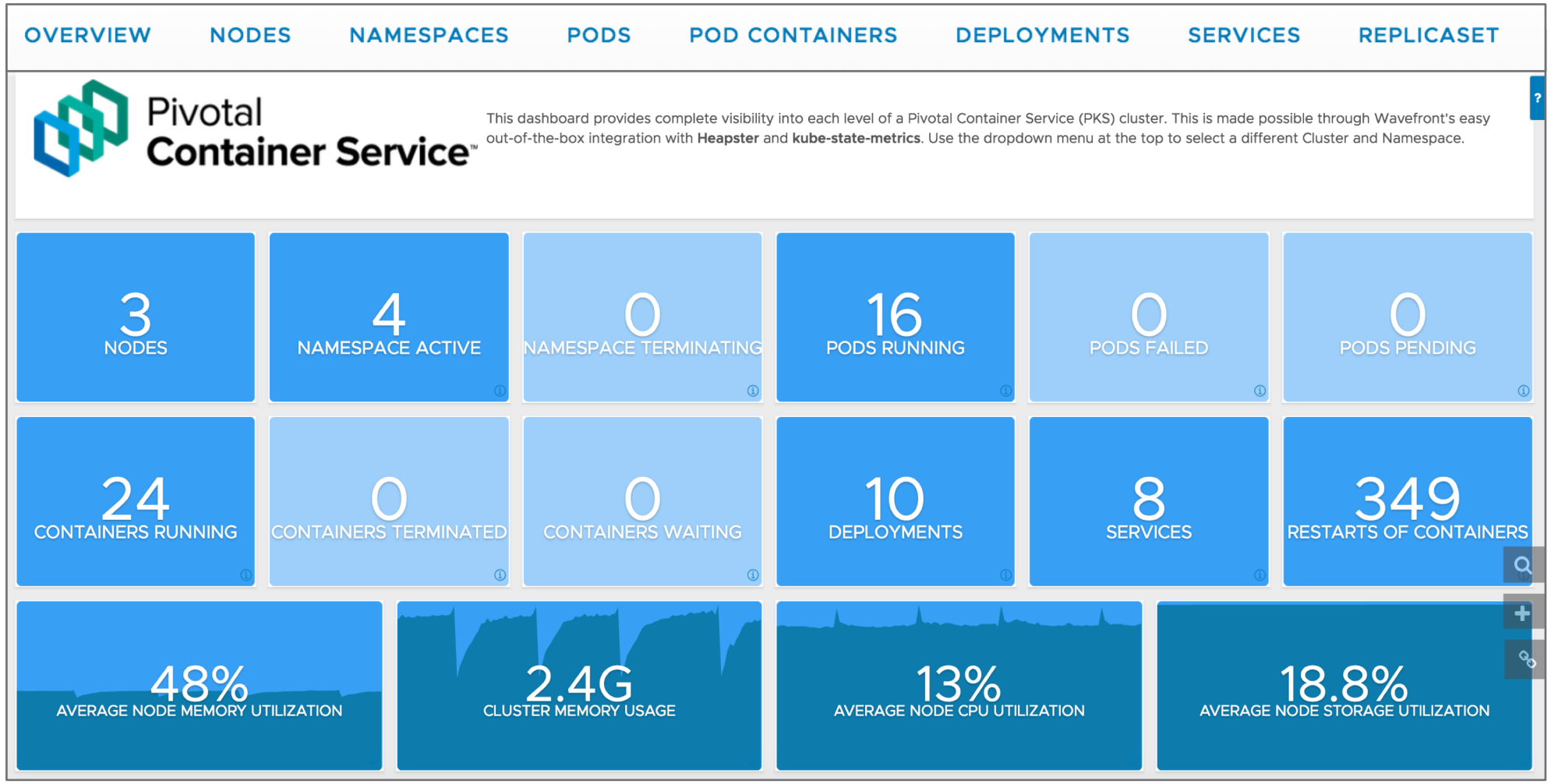
Workload Logging & Monitoring

Namespace multi-tenancy via k8s CRD



VMware Solution – Wavefront

Easy to use



vmware®

VMware Wavefront

Grafana style dashboard, but no maintenance from you

vmware

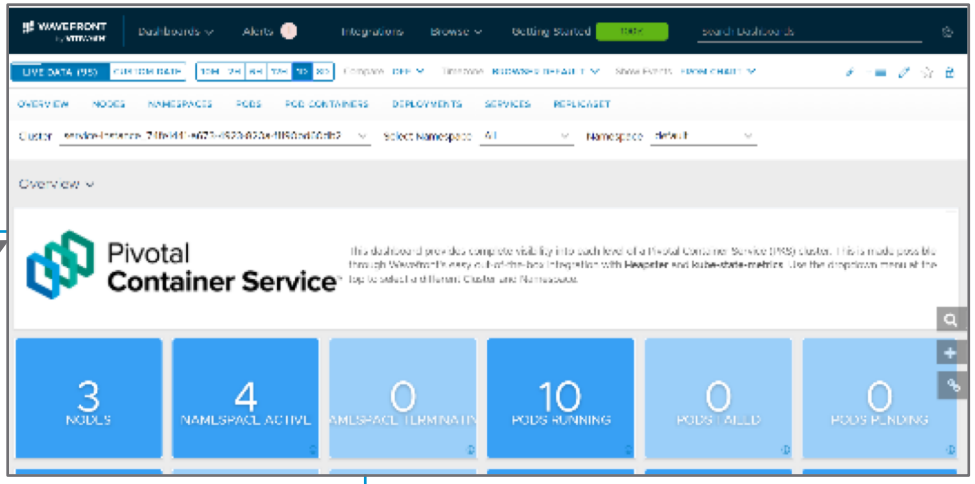
VMware Wavefront

You can create more alerts with wavefront API

Name	Severity	resolveAfterMinutes
Node Memory Usage high	WARN	10
Node Memory Usage too high	SEVERE	10
Node CPU Usage high	WARN	5
Node CPU Usage too high	SEVERE	5
Node Storage Usage high	WARN	10
Node Storage Usage too high	SEVERE	10
Too many Pods crashing	SEVERE	5
Too many Containers not running	SEVERE	5
Node unhealthy	SEVERE	5

VMware Wavefront

Container based agent



Namespace=kube-system

pod=wavefront-proxy

telegraf

wafefront-proxy

kube-state-metrics

heapster

node/pod/container status

resource usage

kube-apiserver

Master

Kubelet

Worker

Kubelet

Worker

Kubelet

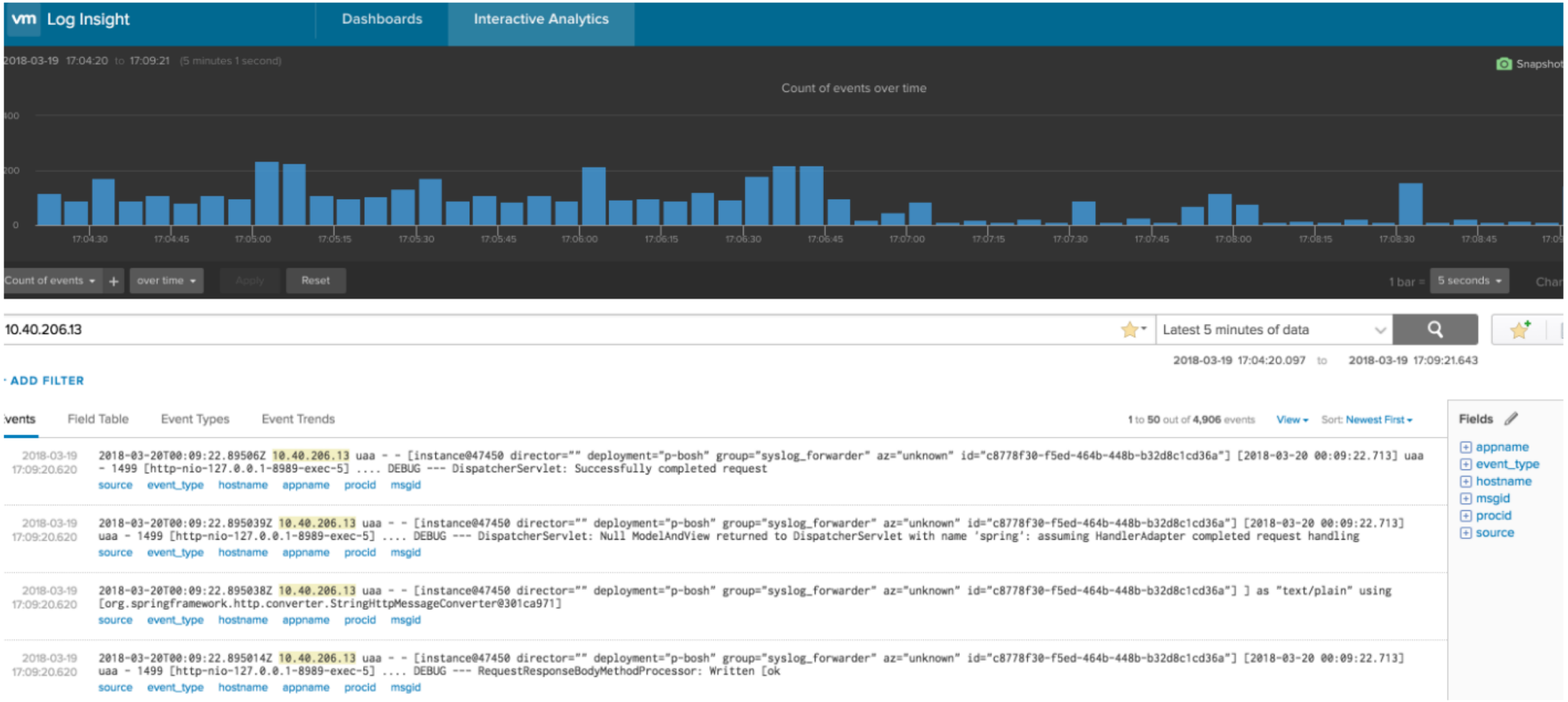
Worker

K8s Cluster

vmware

VMware vRealize Log Insight

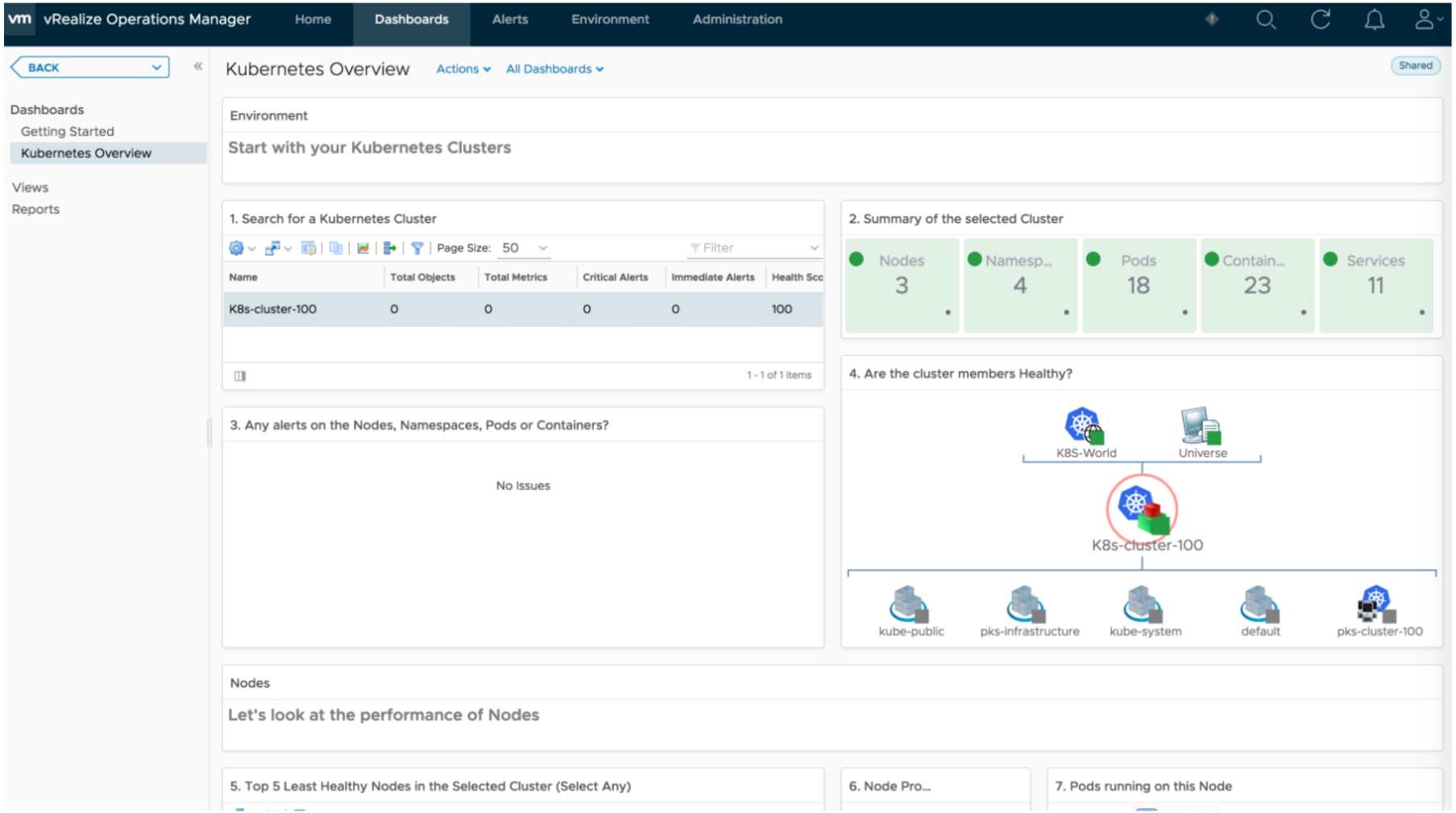
Opinioned log solution



vmware

VMware vRealize Operations

Operation & capacity planning



Built-in VMware Solution – More

More VMware products integration for your scenarios

Bring Your Own Solutions

Choose your favorite recipes, or poisons?

Key Takeaway

Benefits you get with PKS

VMware Realize Suite

Easy To Integrate

Namespace Multi-tenancy

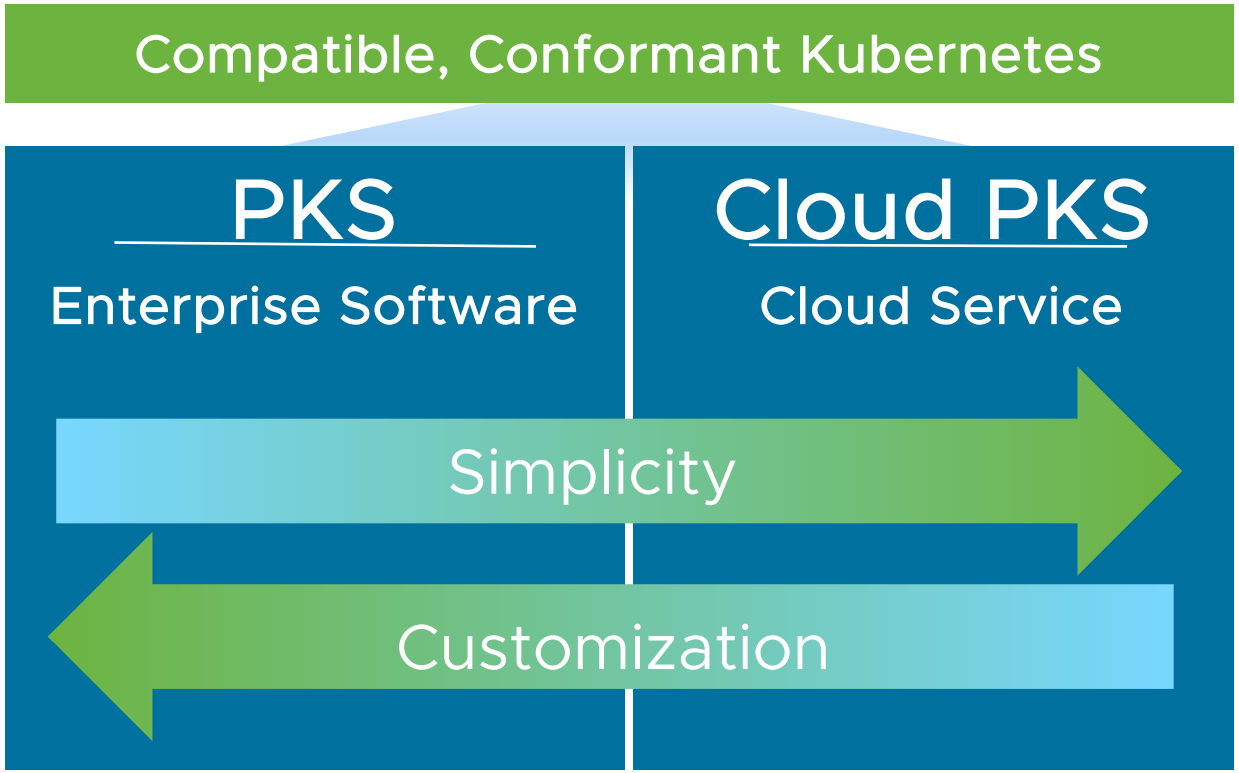
Bring Your Own Solution

Questions?

dennyzhang

Mountain View, United States

Scan the QR code to add me on WeChat



vmware Cloud Foundation™ aws Google Cloud Microsoft Azure

zdenny@vmware.com

vmware