PhraseApp

Continous Delivery with Kubernetes

Tobias Schwab, phraseapp.com

Container Scheduler?

Well, yes, but...

… there is a little bit more

Framework to build **Distributed Applications**

REST API

CNCF: independent from commercial vendors

Linux Kernel for the Cloud?

Getting Started

- Google Container Engine (GKE)
- minikube
- kubeadm (alpha in v1.4)
- ...
- Kubernetes the Hard Way

Pods

Pods

Set of **tightly coupled** containers running on a **single node**

Pod: Features

- dedicated ip
- shared network interface
- shared file system

Pod: Manifest

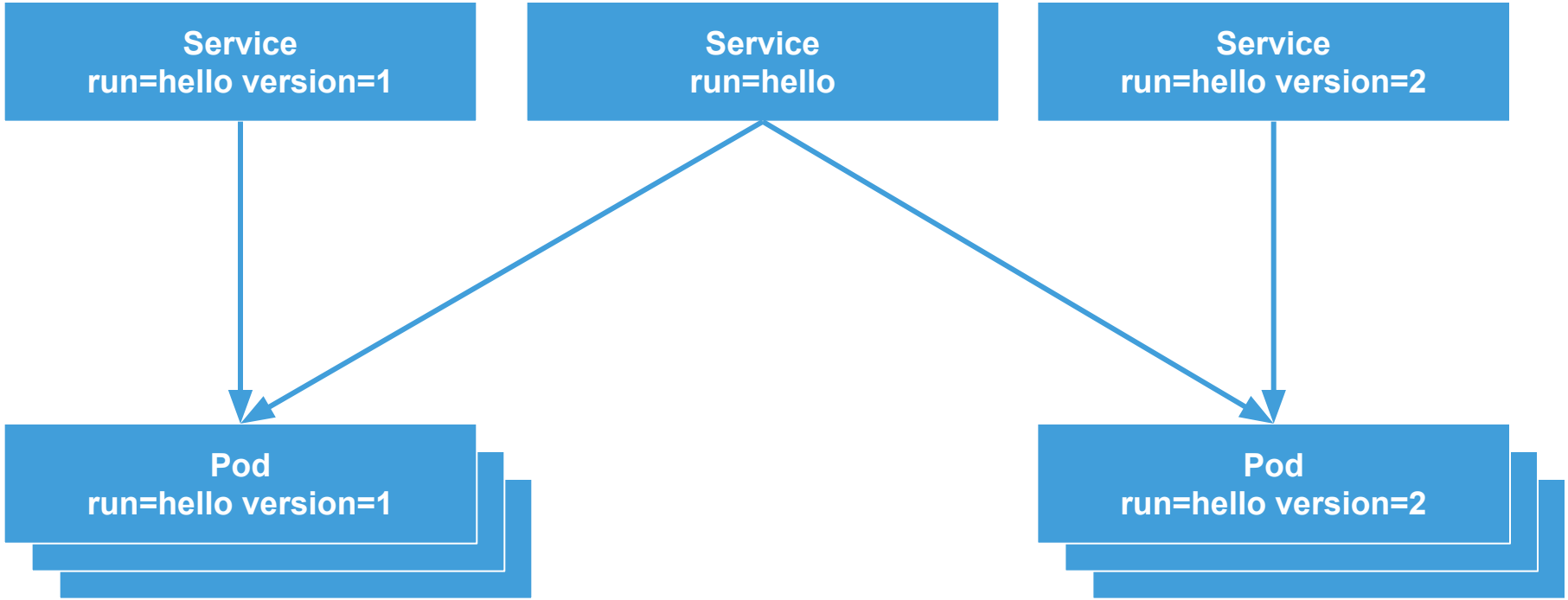
```
kind: Pod
apiVersion: v1
metadata:
  name: hello
  labels:
    run: hello
spec:
  containers:
  - name: hello
    image: quay.io/tobstarr/hello:v1
```

Services

Services

Policy to access a **logical set** of pods

Services



Service: Manifest

```
apiVersion: v1
kind: Service
metadata:
  labels:
    run: hello
  name: hello
spec:
  ports:
  - port: 80
    targetPort: 8080
    protocol: TCP
  selector:
    run: hello
```

Replica Sets

Fixed number of pod **replicas** running on **multiple nodes**

Deployments

declarative Pod updates

- image
- ENV
- configuration files
- labels

Configuration

ConfigMap: **plain text** configuration

- e.g. redis.conf, nginx.conf

Secret: **sensitive** information

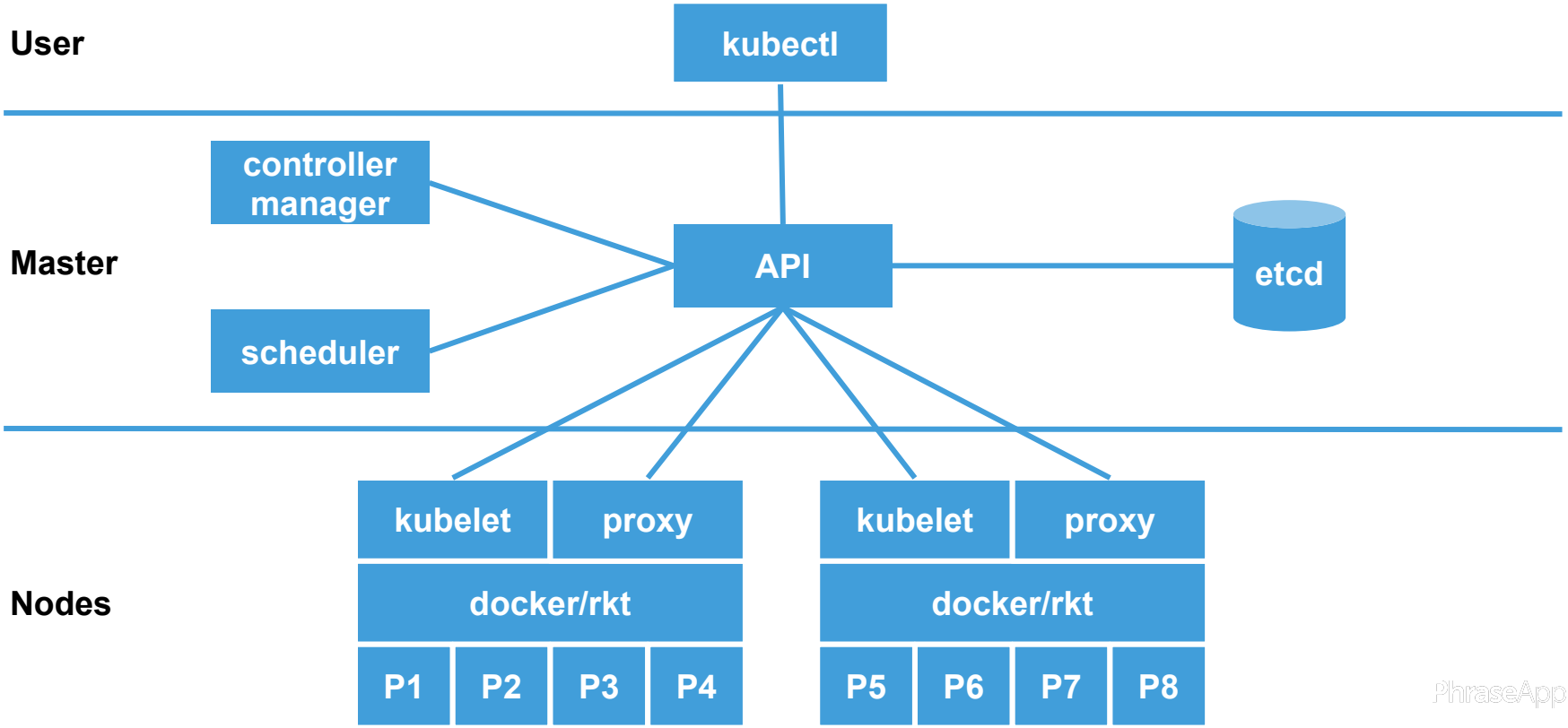
- e.g. TLS certs, registry credentials, etc.

Volumes

Persistent storage beyond container lifetime

- configMap
- secrets
- hostPath
- gcePersistentDisk
- awsElasticBlockStore
- nfs

Architecture



Demo

Continuous Delivery

Continuous Delivery

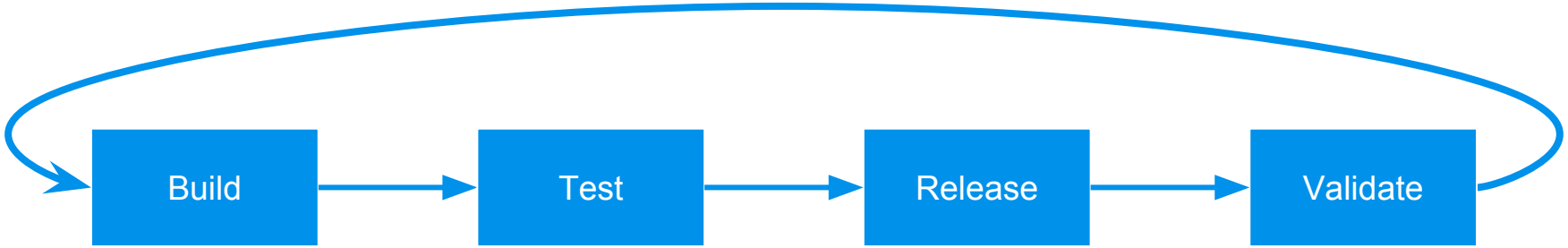
build, **test**, and **release** software **faster** and **more frequently**

Continuous Delivery

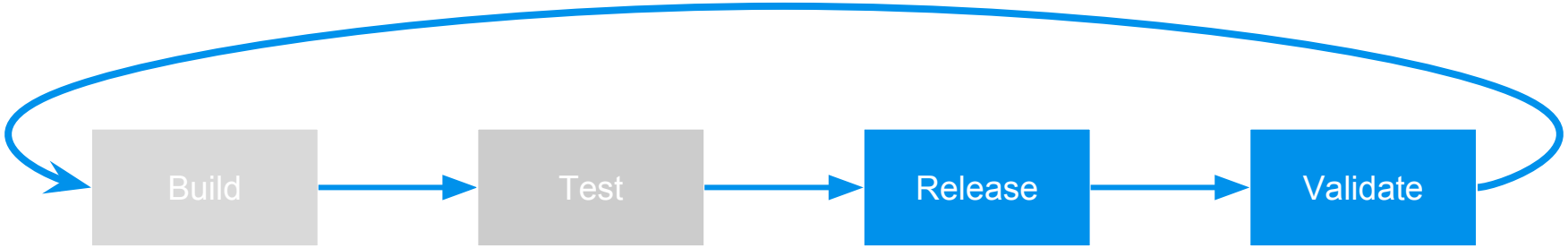
Our Philosophy:

- everyone **can** deploy any time
- you build it, you run it!
- **automation!**

Continuous Delivery Pipeline



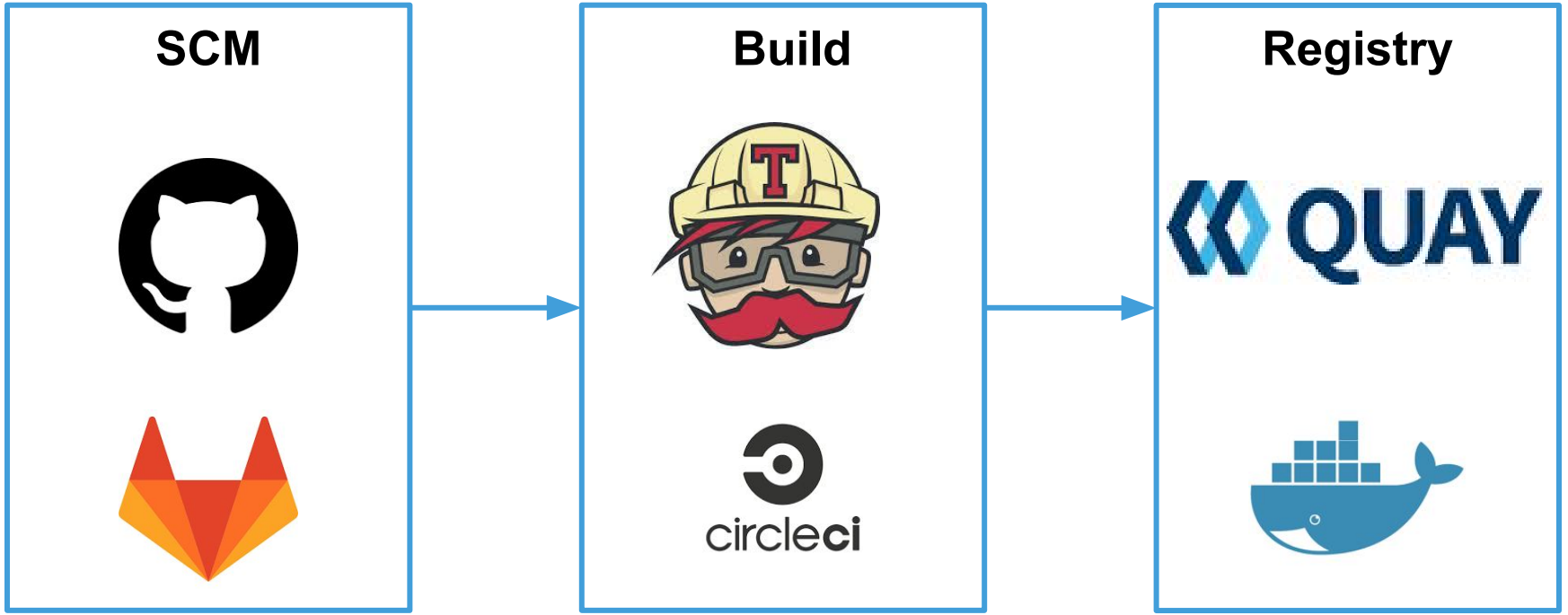
Build & Test



Build & Test

- compile
- run unit and functional tests
- build container image
- push container image to registry

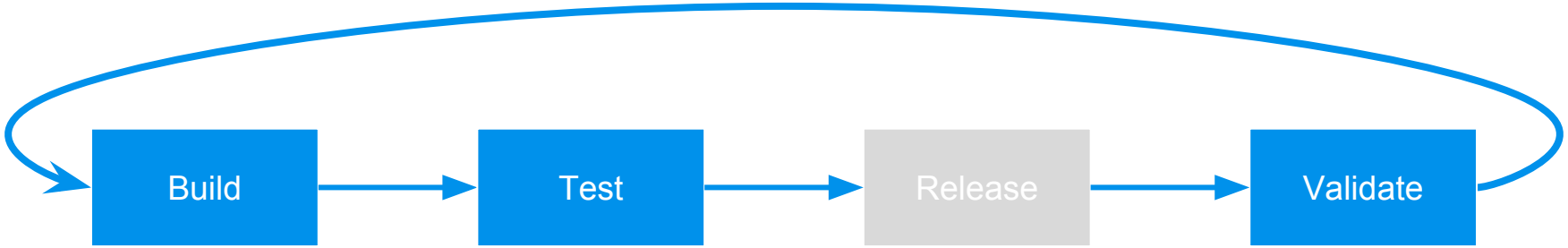
Build & Test - v1



Build & Test: Kubernetes

- jenkins container
- private registry container
- Pipeline plugin
- docker and kubectl in jenkins image
- /var/run/docker.sock via *VolumeMount*
- scale out builds via kubectl

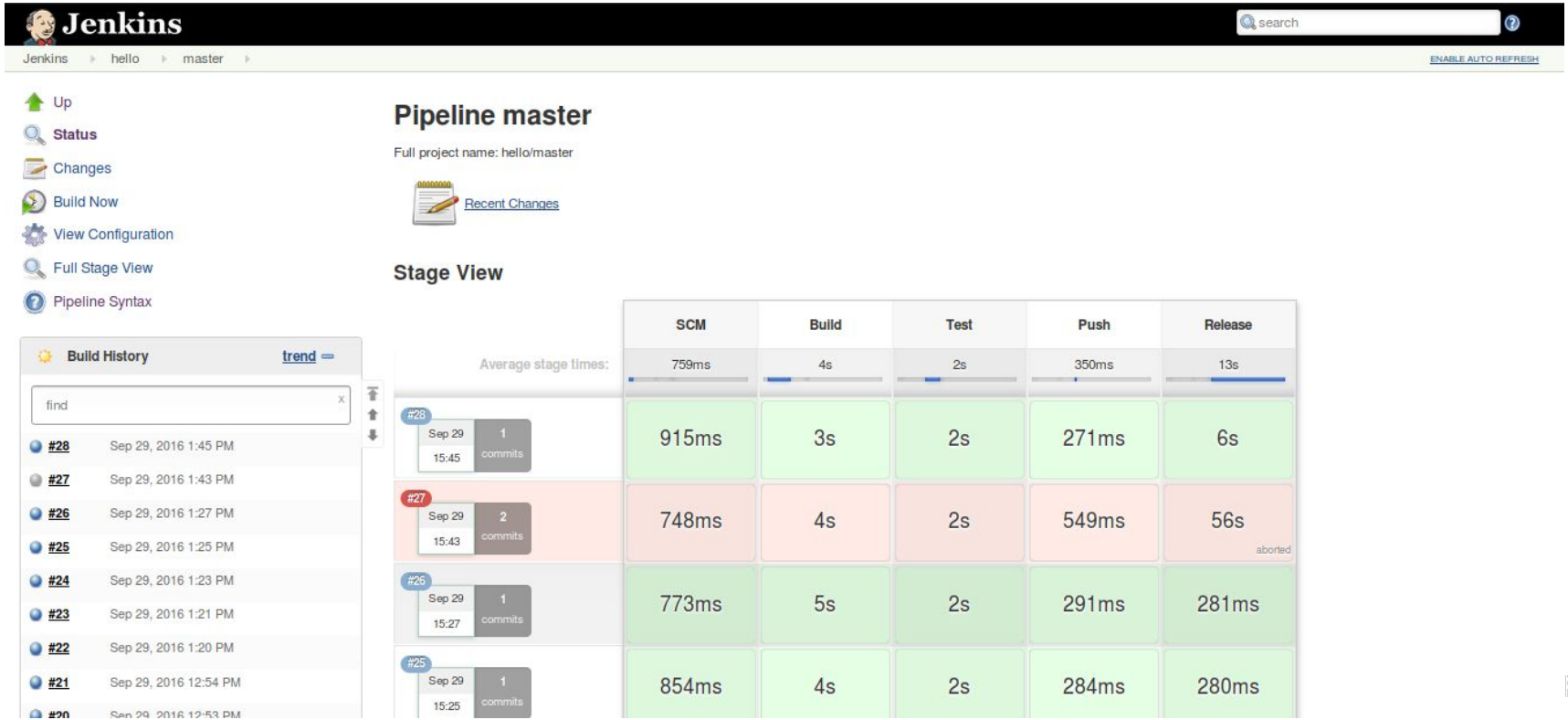
Release



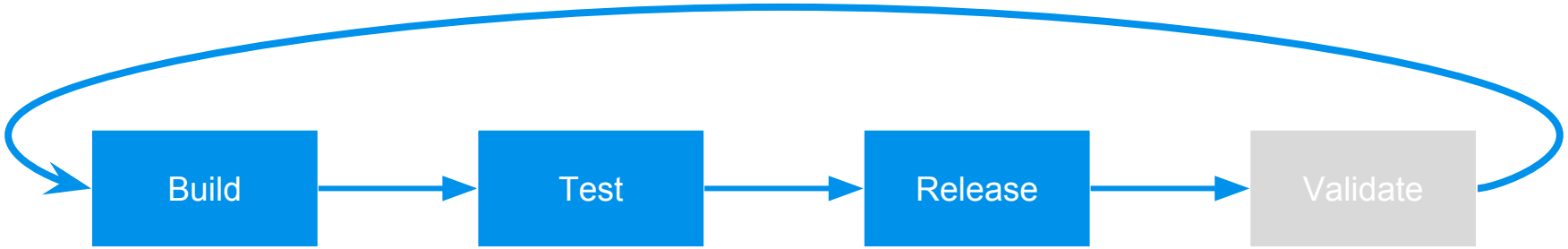
Release

```
kubectl set image ...
```

Kubernetes Build Pipeline



Validate



Monitoring

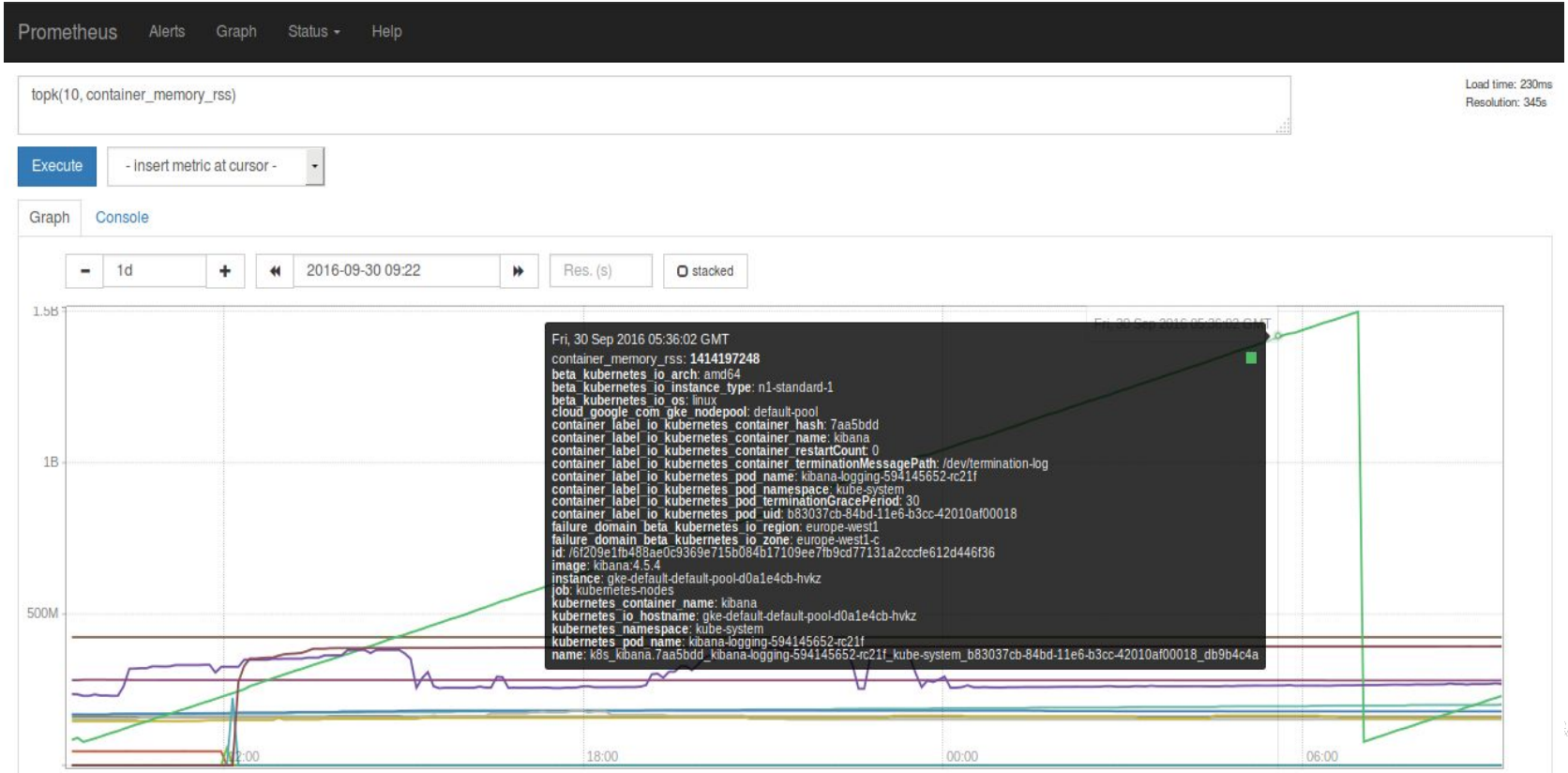
Cadvisor

- expose container metrics
- running on k8s nodes by default

Prometheus

- pull based monitoring & alerting
- https://coreos.com/blog/prometheus-and-kubernetes-up-and-running.html

Prometheus


Prometheus
Alerts
Graph
Status
Help
topk(10, container_memory_rss)
Load time: 230ms
Resolution: 345s
Execute
- insert metric at cursor -
Graph
Console
1d
2016-09-30 09:22
Res. (s)
stacked
1.5B
1B
500M
18:00
00:00
06:00
Fri, 30 Sep 2016 05:36:02 GMT
container_memory_rss: 1414197248
beta_kubernetes_io_arch: amd64
beta_kubernetes_io_instance_type: n1-standard-1
beta_kubernetes_io_os: linux
cloud_google_com_gke_nodepool: default-pool
container_label_io_kubernetes_container_hash: 7aa5bdd
container_label_io_kubernetes_container_name: kibana
container_label_io_kubernetes_container_restartCount: 0
container_label_io_kubernetes_container_terminationMessagePath: /dev/termination-log
container_label_io_kubernetes_pod_name: kibana-logging-594145652-rc21f
container_label_io_kubernetes_pod_namespace: kube-system
container_label_io_kubernetes_pod_terminationGracePeriod: 30
container_label_io_kubernetes_pod_uid: b83037cb-84bd-11e6-b3cc-42010af00018
failure_domain_beta_kubernetes_io_region: europe-west1
failure_domain_beta_kubernetes_io_zone: europe-west1-c
id: /6f209e1fb488ae0c9369e715b084b17109ee7fb9cd77131a2cccfe612d446f36
image: kibana:4.5.4
instance: gke-default-default-pool-d0a1e4cb-hvkz
job: kubernetes-nodes
kubernetes_container_name: kibana
kubernetes_io_hostname: gke-default-default-pool-d0a1e4cb-hvkz
kubernetes_namespace: kube-system
kubernetes_pod_name: kibana-logging-594145652-rc21f
name: k8s_kibana.7aa5bdd_kibana-logging-594145652-rc21f_kube-system_b83037cb-84bd-11e6-b3cc-42010af00018_db9b4c4a

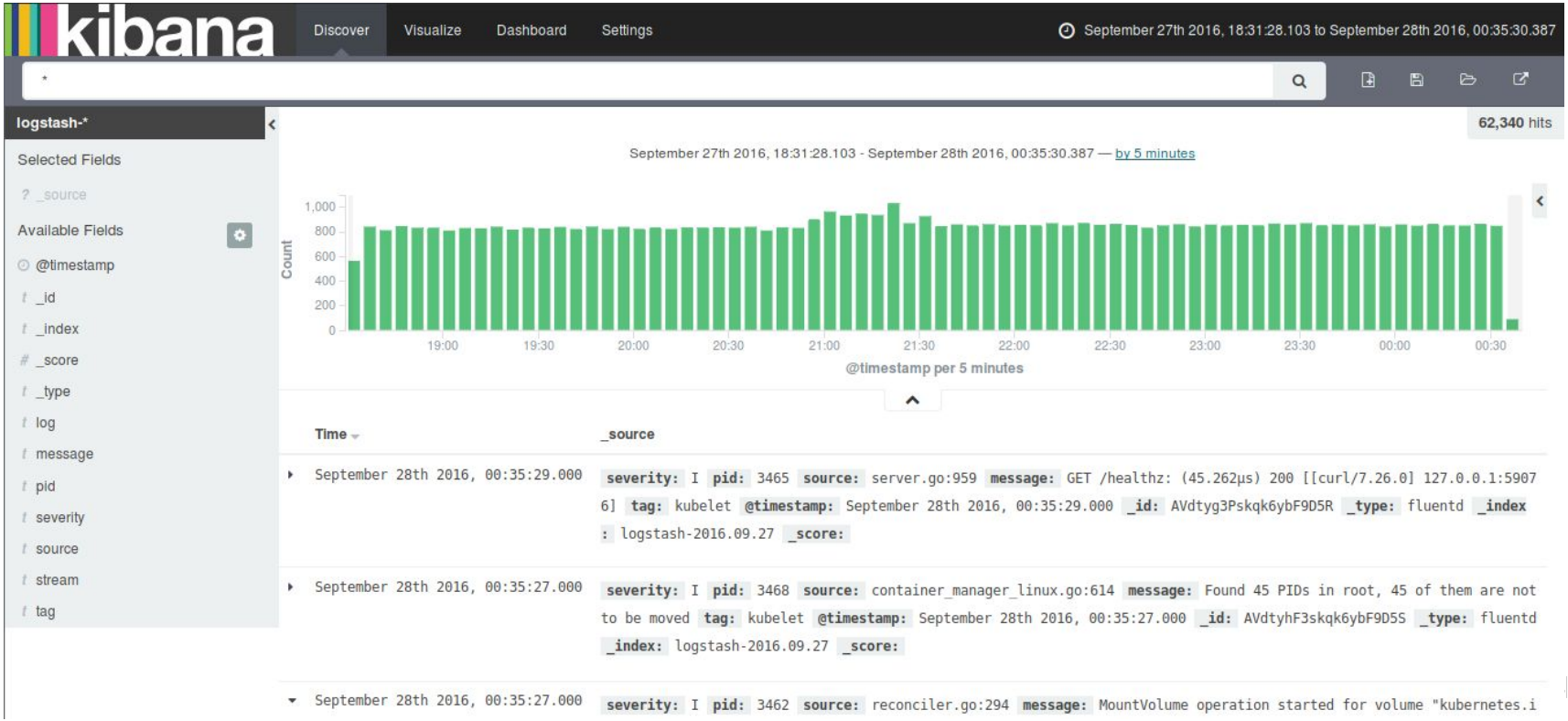
Logging

- apps log to **stdout**
- containers log to local file system
- fluentd
- elasticsearch
- kibana

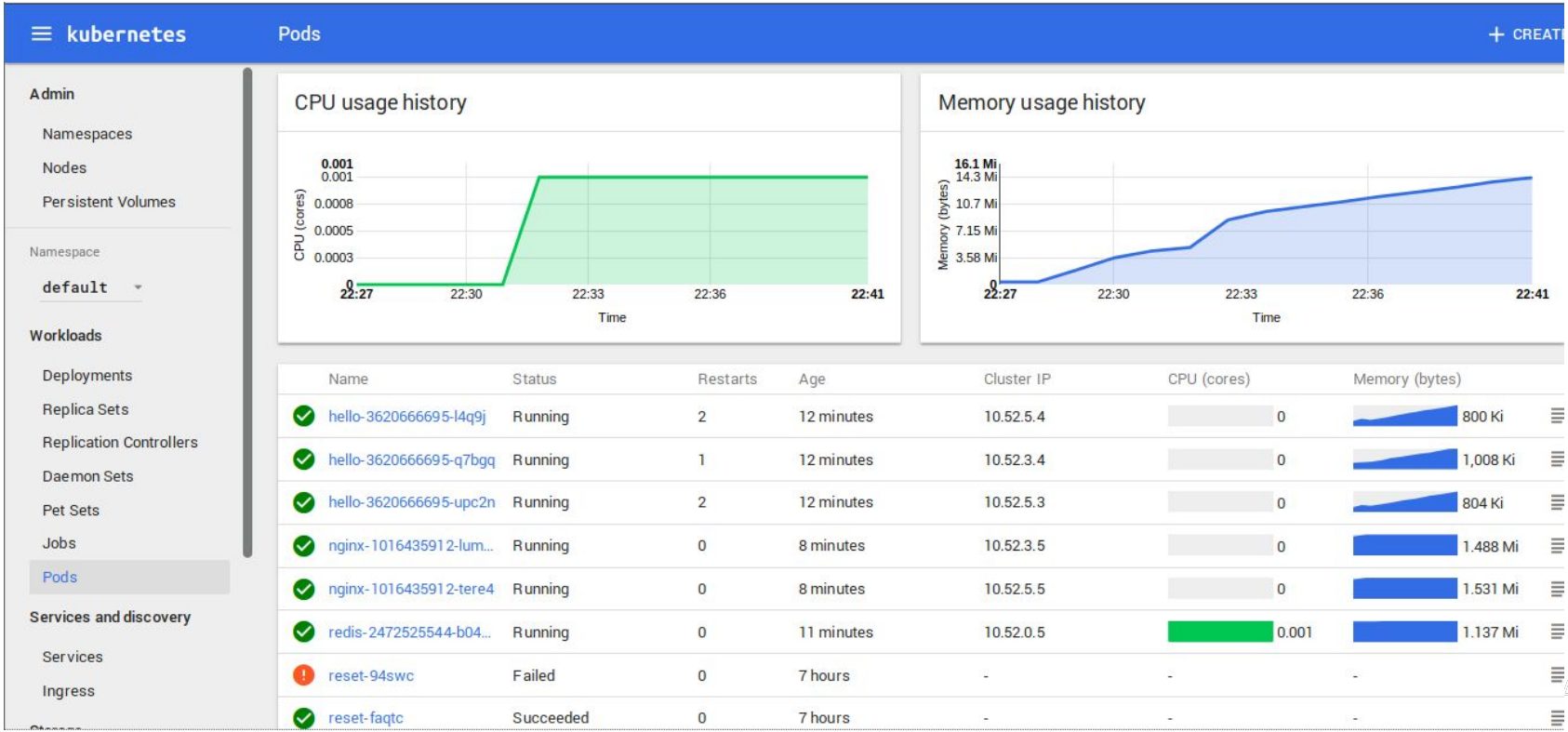
Logging

- *DaemonSet*: Pod running on selected **nodes** in a cluster
- mount node FS via *VolumeMount*
- fluentd sends logs to ElasticSearch
- display/search with Kibana

Kibana


kibana
Discover
Visualize
Dashboard
Settings
September 27th 2016, 18:31:28.103 to September 28th 2016, 00:35:30.387
*
logstash-*
62,340 hits
Selected Fields
? _source
Available Fields
@timestamp
t _id
t _index
_score
t _type
t log
t message
t pid
t severity
t source
t stream
t tag
September 27th 2016, 18:31:28.103 - September 28th 2016, 00:35:30.387 — by 5 minutes
Count
1,000
800
600
400
200
0
19:00
19:30
20:00
20:30
21:00
21:30
22:00
22:30
23:00
23:30
00:00
00:30
@timestamp per 5 minutes
Time
_source
September 28th 2016, 00:35:29.000
severity: I pid: 3465 source: server.go:959 message: GET /healthz: (45.262µs) 200 [[curl/7.26.0] 127.0.0.1:59076] tag: kubelet @timestamp: September 28th 2016, 00:35:29.000 _id: AVdtyg3Pskqk6ybF9D5R _type: fluentd _index: logstash-2016.09.27 _score:
September 28th 2016, 00:35:27.000
severity: I pid: 3468 source: container_manager_linux.go:614 message: Found 45 PIDs in root, 45 of them are not to be moved tag: kubelet @timestamp: September 28th 2016, 00:35:27.000 _id: AVdtyhF3skqk6ybF9D5S _type: fluentd _index: logstash-2016.09.27 _score:
September 28th 2016, 00:35:27.000
severity: I pid: 3462 source: reconciler.go:294 message: MountVolume operation started for volume "kubernetes.i

Dashboard


kubernetes
Pods
+ CREATE
Admin
Namespaces
Nodes
Persistent Volumes
Namespace
default
Workloads
Deployments
Replica Sets
Replication Controllers
Daemon Sets
Pet Sets
Jobs
Pods
Services and discovery
Services
Ingress
CPU usage history
CPU (cores)
0.001
0.001
0.0008
0.0005
0.0003
0
22:27
22:30
22:33
22:36
22:41
Time
Memory usage history
Memory (bytes)
16.1 Mi
14.3 Mi
10.7 Mi
7.15 Mi
3.58 Mi
0
22:27
22:30
22:33
22:36
22:41
Time
Name
Status
Restarts
Age
Cluster IP
CPU (cores)
Memory (bytes)
hello-3620666695-l4q9j
Running
2
12 minutes
10.52.5.4
0
800 Ki
hello-3620666695-q7bgq
Running
1
12 minutes
10.52.3.4
0
1,008 Ki
hello-3620666695-upc2n
Running
2
12 minutes
10.52.5.3
0
804 Ki
nginx-1016435912-lum...
Running
0
8 minutes
10.52.3.5
0
1.488 Mi
nginx-1016435912-tere4
Running
0
8 minutes
10.52.5.5
0
1.531 Mi
redis-2472525544-b04...
Running
0
11 minutes
10.52.0.5
0.001
1.137 Mi
reset-94swc
Failed
0
7 hours
-
-
-
reset-faqtc
Succeeded
0
7 hours
-
-
-

Takeaways

- fully automated CD pipeline
- no **lock-in** or dependency to external services
- failure tolerance towards process **and** node failures
- **scalable!**

Source

- github.com/tobstarr/code-talks-2016
- github.com/tobstarr/hello

PhraseApp

@tobstarr
tobias@phraseapp.com
phraseapp.com