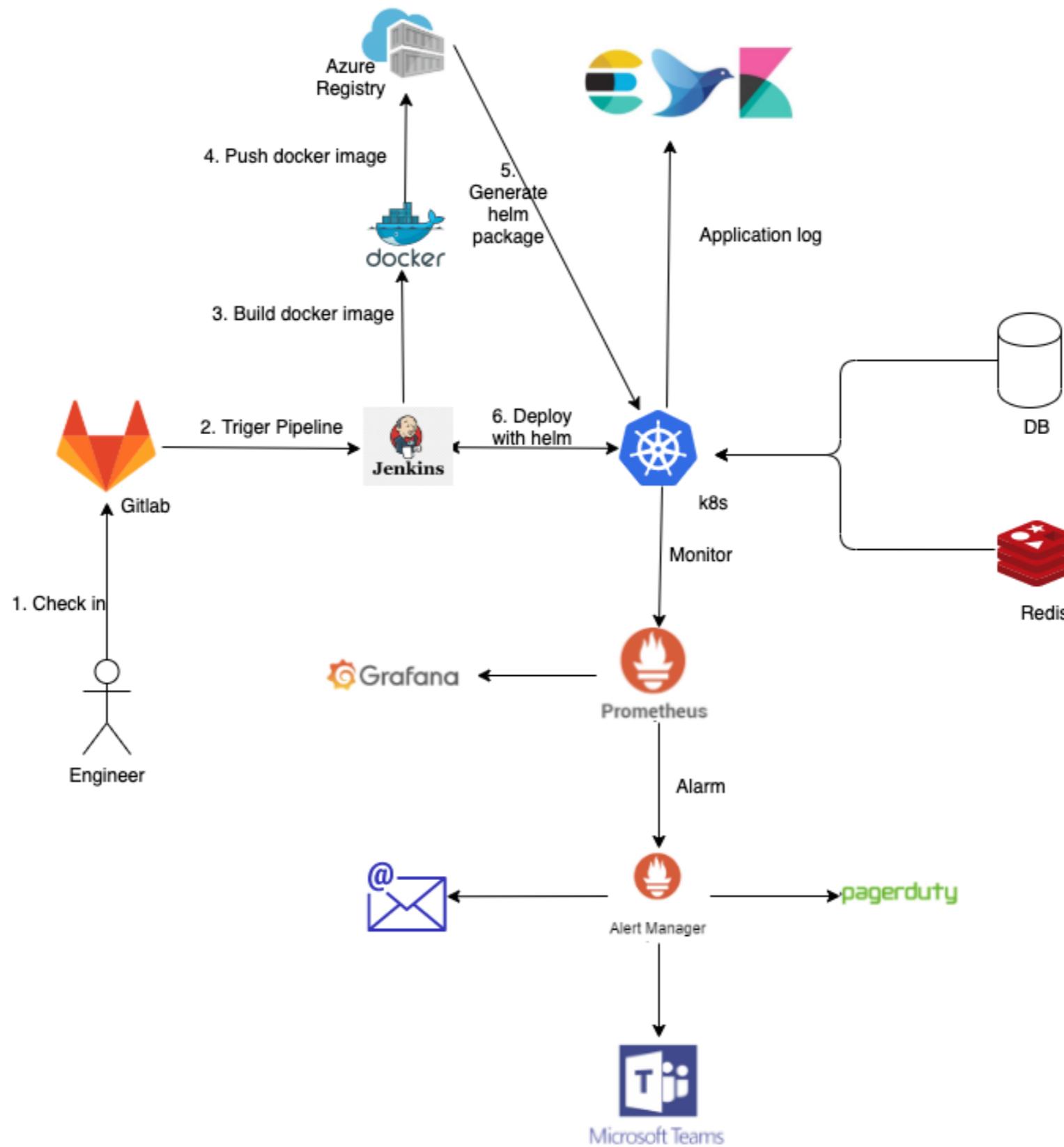


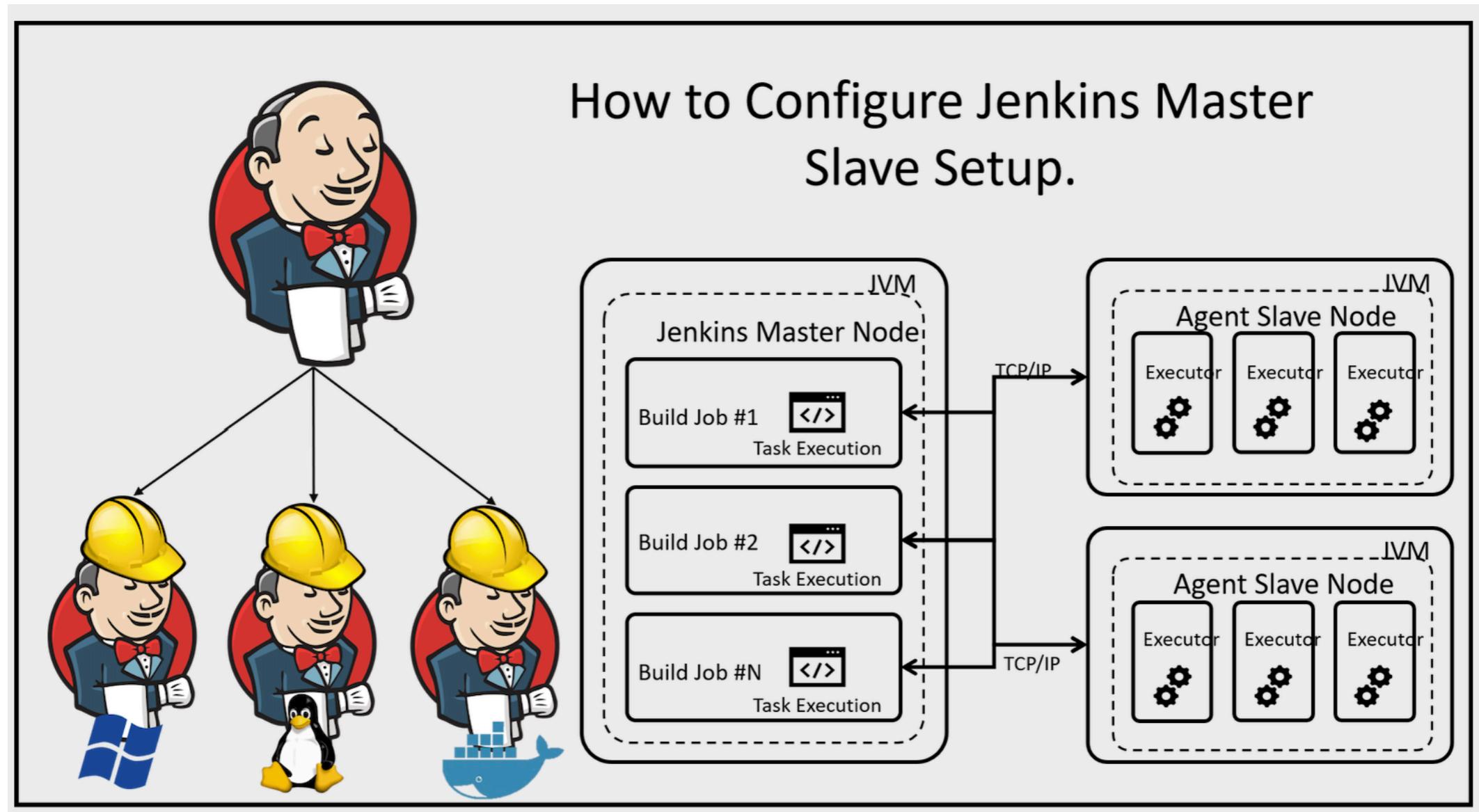
Kubernetes + Cloud

Qingjie Zhao
03/31/2020

System Architecture Diagram with Jenkins

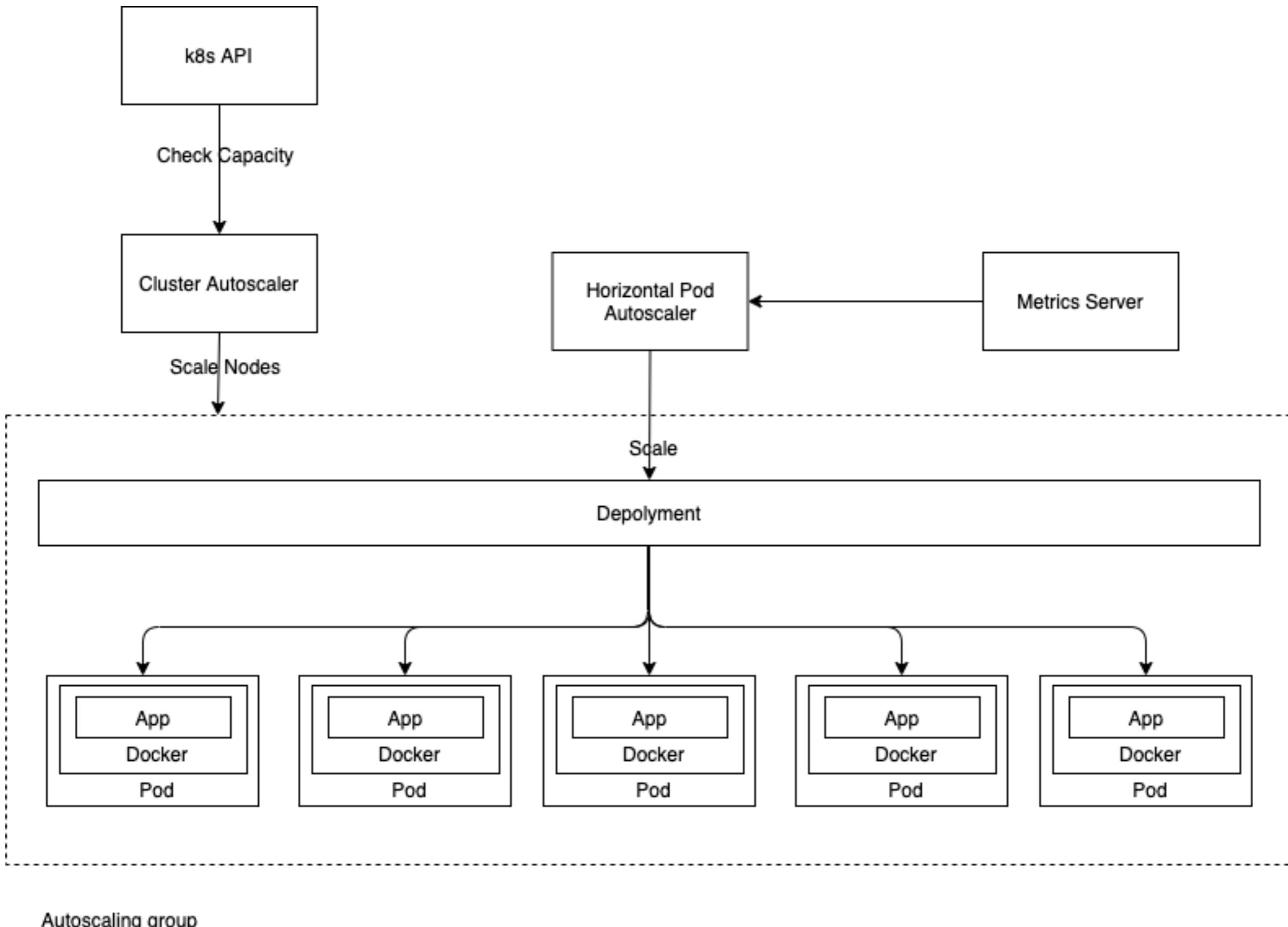


Jenkins - master - slave

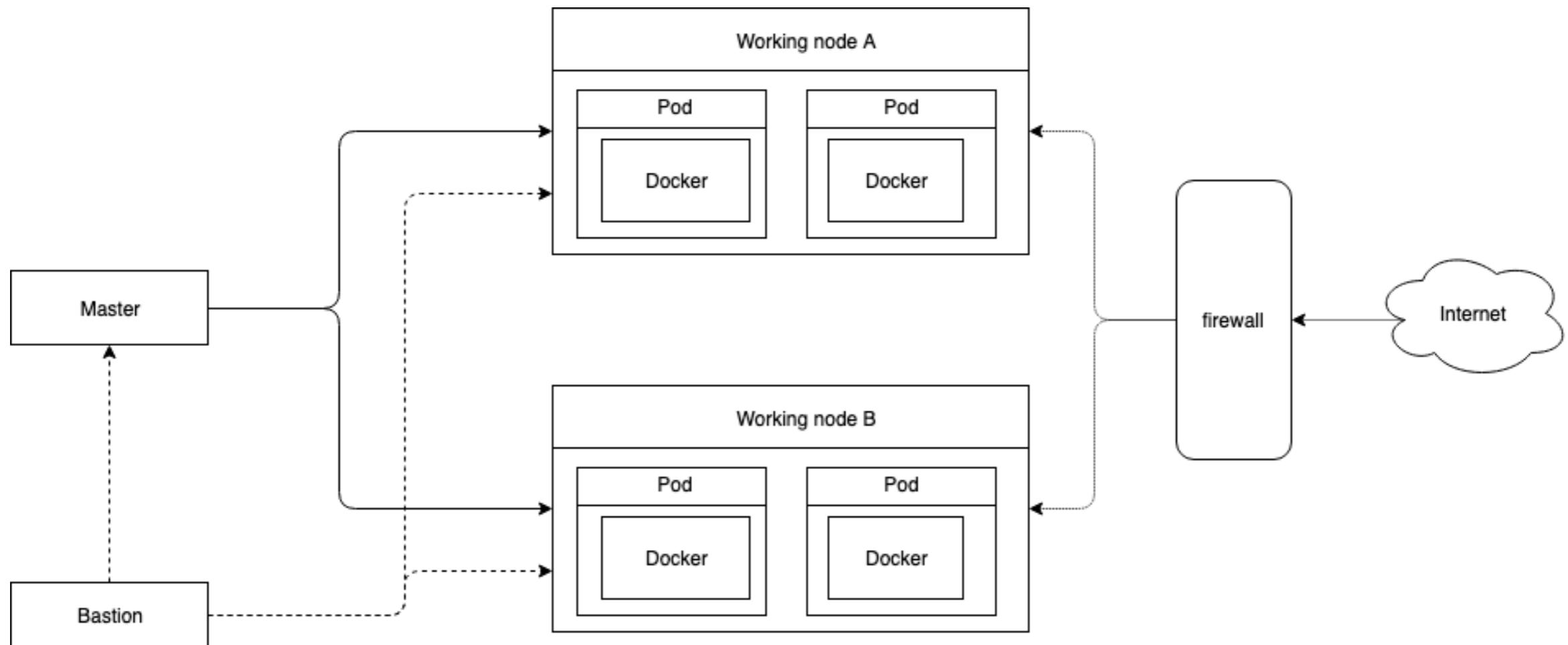


- Application will be start/run/destroy on slave

Cluster Autoscaler + Horizontal Pod Autoscaler

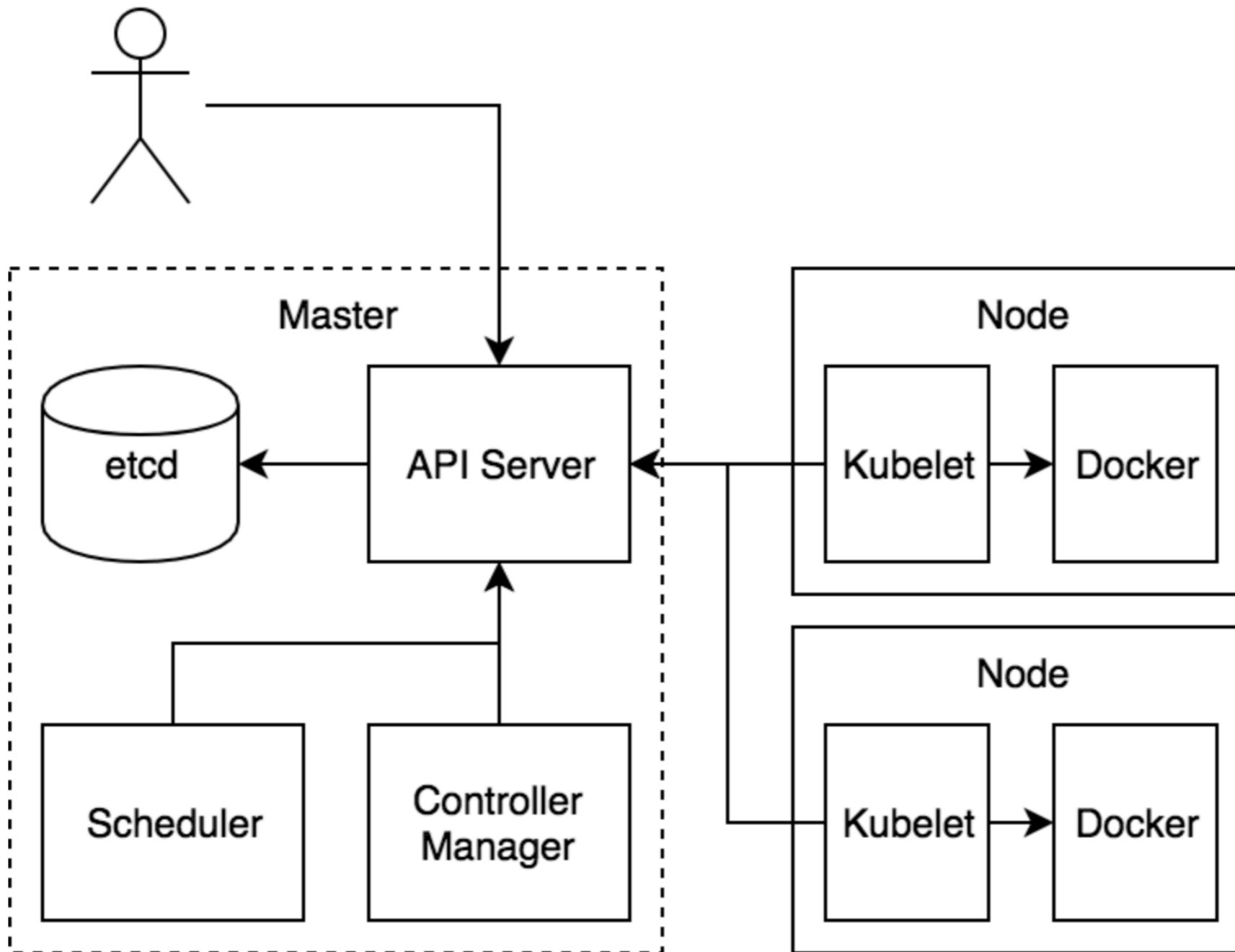


Kubernetes real architecture



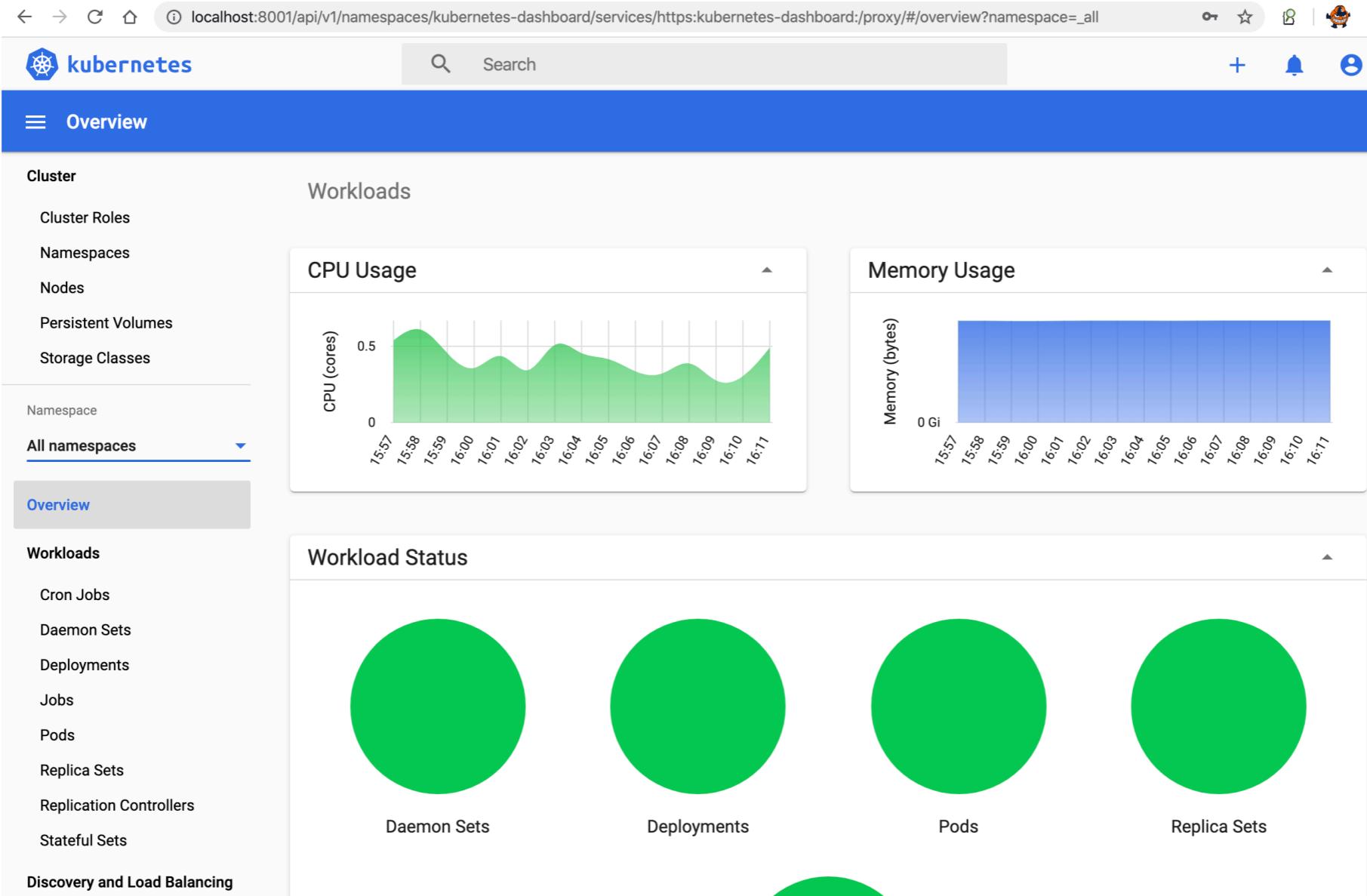
- It doesn't show master and bastion on AKS/EKS

Master Node



- It doesn't show Master Node on AKS and EKS

Kubernetes-dashboard + matrix server



- kubectl proxy

namespace/deployment/service/replicaset/pod/hpa

namespace								
NAME	READY	STATUS	RESTARTS	AGE				
pod/jsreport-server-896dd9d64-qwtsj	1/1	Running	0	2d				
pod								
NAME	TYPE	CLUSTER-IP	EXTERNAL-IP	PORT(S)	AGE			
service/jsreport-server	ClusterIP	10.100.164.80	<none>	5488/TCP	2d			
service								
NAME	READY	UP-TO-DATE	AVAILABLE	AGE				
deployment.apps/jsreport-server	1/1	1	1	2d				
deployment								
NAME	DESIRED	CURRENT	READY	AGE				
replicaset.apps/jsreport-server-896dd9d64	1	1	1	2d				
rs								
NAME	REFERENCE			REPLICAS				
horizontalpodautoscaler.autoscaling/jsreport-server	Deployment/jsreport-server			TARGETS 2%/10%	MINPODS 1			
hpa								
hpa								
If CPU is 10%, pod will be auto increased								
TARGETS 2%/10%								
MINPODS 1								
MAXPODS 10								
REPLICAS 1								
AGE 68m								

namespace

```
qzhaos-MBP:dev-k8s qzhao$ kubectl get ns
NAME          STATUS  AGE
cert-manager   Active  7d21h
default        Active  22d
echo1          Active  5d9h
gitlab-managed-apps  Active  19d
ingress-nginx  Active  7d21h
jenkins        Active  2d20h
jsreport-server Active  26h
kube-node-lease Active  22d
kube-public    Active  22d
kube-system    Active  22d
kubernetes-dashboard Active  6d7h
locust          Active  5d6h
monitoring     Active  7d5h
sentry          Active  5d9h
test1           Active  30h
weave           Active  5d10h
```

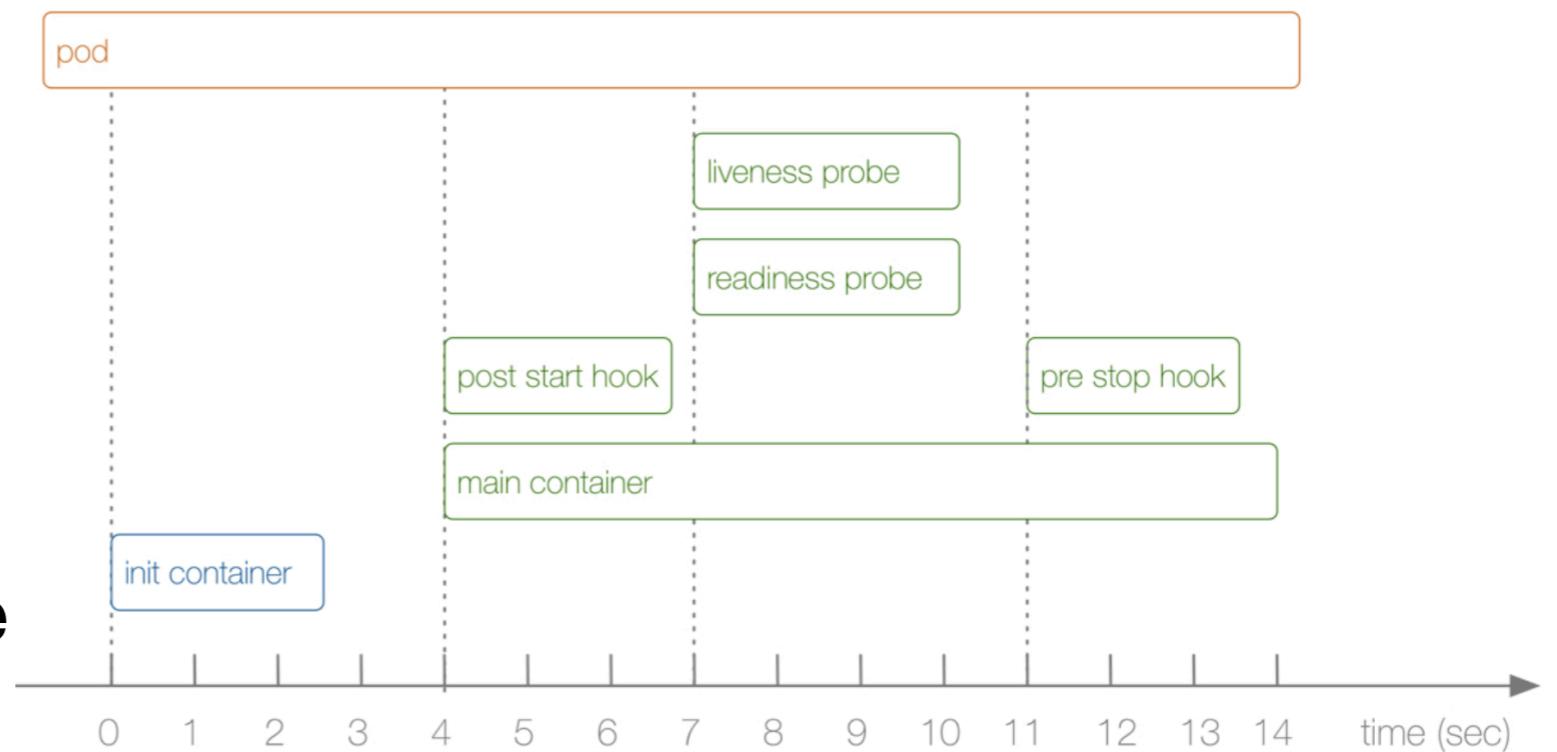
- one application should be run in one namespace

Pod description

```
Image ID: docker-pullable://gitlab.americanportfolios.com:4567/infrastructure/jsreport-server/tst@sha256:c9577ec1e421d5746b8a971b7a544e9d0bb2d1554042232be3e18991c8f
Port: 5488/TCP
Host Port: 0/TCP
State: Running
Started: Tue, 31 Mar 2020 09:30:35 -0400
Ready: True
Restart Count: 0
Limits:
cpu: 200m
memory: 250Mi
Requests:
cpu: 200m
memory: 250Mi
Liveness: http-get http://:5488/ delay=10s timeout=5s period=10s #success=1 #failure=3
Readiness: http-get http://:5488/ delay=10s timeout=5s period=10s #success=1 #failure=3
Environment:
APP: server.js
CLUSTERS: 1
NAME: jsreport-server
Mounts:
/app/data from jsreport-server-home-data (rw)
/var/run/secrets/kubernetes.io/serviceaccount from jsreport-server-token-ts7wk (ro)
Conditions:
Type Status
Initialized True
Ready True
ContainersReady True
PodScheduled True
Volumes:
jsreport-server-home-data:
Type: PersistentVolumeClaim (a reference to a PersistentVolumeClaim in the same namespace)
ClaimName: jsreport-server-pvc-data
ReadOnly: false
jsreport-server-token-ts7wk:
Type: Secret (a volume populated by a Secret)
SecretName: jsreport-server-token-ts7wk
Optional: false
QoS Class: Guaranteed
Node-Selectors: <none>
```

QoS and Health check

- QoS
 - Guaranteed
 - Burstable
 - BestEffect
- Health check
 - livenessProbe
 - readinessProbe



Deployment description

```
qzhaos-MBP:dev-k8s qzhao$ kubectl describe deployment.apps/jsreport-server -n jsreport-server
Name:           jsreport-server
Namespace:      jsreport-server
CreationTimestamp:  Tue, 31 Mar 2020 09:30:24 -0400
Labels:         app.kubernetes.io/instance=jsreport-server
                app.kubernetes.io/managed-by=Tiller
                app.kubernetes.io/name=jsreport-server
                app.kubernetes.io/version=1.0
                helm.sh/chart=jsreport-server-0.1.0
Annotations:    deployment.kubernetes.io/revision: 1
Selector:       app.kubernetes.io/instance=jsreport-server,app.kubernetes.io/name=jsreport-server
Replicas:       2 desired | 2 updated | 2 total | 2 available | 0 unavailable
StrategyType:   RollingUpdate
MinReadySeconds: 0
RollingUpdateStrategy: 25% max unavailable, 25% max surge
Pod Template:
  Labels:      app.kubernetes.io/instance=jsreport-server
                app.kubernetes.io/name=jsreport-server
  Service Account: jsreport-server
  Containers:
    jsreport-server:
      Image:      gitlab.americanportfolios.com:4567/infrastructure/jsreport-server/tst:latest
      Port:       5488/TCP
      Host Port: 0/TCP
      Limits:
        cpu:      200m
        memory:  250Mi
      Requests:
        cpu:      200m
        memory:  250Mi
      Liveness:   http-get http://:5488/ delay=10s timeout=5s period=10s #success=1 #failure=3
      Readiness:  http-get http://:5488/ delay=10s timeout=5s period=10s #success=1 #failure=3
      Environment:
        APP:      server.js
        CLUSTERS: 1
        NAME:     jsreport-server
      Mounts:
        /app/data from jsreport-server-home-data (rw)
  Volumes:
    jsreport-server-home-data:
```

Service description

```
qzhaos-MBP:dev-k8s qzhao$ kubectl describe service/jsreport-server -n jsreport-server
Name:           jsreport-server
Namespace:      jsreport-server
Labels:         app.kubernetes.io/instance=jsreport-server
                app.kubernetes.io/managed-by=Tiller
                app.kubernetes.io/name=jsreport-server
                app.kubernetes.io/version=1.0
                helm.sh/chart=jsreport-server-0.1.0
Annotations:    <none>
Selector:       app.kubernetes.io/instance=jsreport-server,app.kubernetes.io/name=jsreport-server
Type:          ClusterIP
IP:            10.100.164.80
Port:          http  5488/TCP
TargetPort:     http/TCP
Endpoints:     192.168.37.195:5488,192.168.37.202:5488
Session Affinity: None
Events:        <none>
```

PersistentVolume

The screenshot shows the AWS EC2 Dashboard with the 'Volumes' section selected under 'ELASTIC BLOCK STORE'. A red box highlights the 'Volumes' link in the sidebar. Another red box highlights the volume name 'kubernetes-dynamic-pvc-1164a9fa-5e64-11ea-853d-160a68a82d21' in the list table. A third red box highlights the 'Size' column, showing '7 GiB'. A fourth red box highlights the 'Volume Type' column, showing 'io1'.

Name	Volume ID	Size	Volume Type	IOPS	Snap
kubernetes-dynamic-pvc-1164a9fa-5e64-11ea-853d-160a68a82d21	vol-05c8494...	7 GiB	io1	100	
	vol-0049e21...	20 GiB	gp2	100	snap-

Volumes: vol-05c8494eb5eb5812c (kubernetes-dynamic-pvc-1164a9fa-5e64-11ea-853d-160a68a82d21)

Description	Status Checks	Monitoring	Tags
Volume ID: vol-05c8494eb5eb5812c	Size: 7 GiB	Created: March 4, 2020 at 5:04:12 PM UTC-5	Alarm status: None
Lifecycle Manager	State: in-use	Attachment information: i-01d0273ec5cc1e26:/dev/xvdbd (attached)	Snapshot: -
Network & Security	Volume type: io1	Product codes: -	Availability Zone: us-east-1b
Security Groups			Encryption: Not Encrypted
Elastic IPs			KMS Key ID:
Placement Groups			KMS Key Aliases:
Key Pairs			KMS Key ARN:
Network Interfaces			

How to use pvc on deployment.yaml

```
1 ---  
2 # What do want?  
3 apiVersion: v1  
4 kind: PersistentVolumeClaim  
5 metadata:  
6   name: jsreport-server-pvc-data  
7   namespace: jsreport-server  
8 spec:  
9   storageClassName: es-cloud-ssd-pv-jsreport-server-data  
10  accessModes:  
11    - ReadWriteOnce  
12  resources:  
13    requests:  
14      storage: 16Gi  
15 ---  
16 # How do we want it implemented  
17 apiVersion: storage.k8s.io/v1  
18 kind: StorageClass  
19 metadata:  
20   name: es-cloud-ssd-pv-jsreport-server-data  
21   namespace: jsreport-server  
22 reclaimPolicy: Retain  
23 provisioner: kubernetes.io/aws-ebs  
24 parameters:  
25   type: io1  
26  
27 ---  
28 # What do want?  
29 apiVersion: apps/v1  
30 kind: Deployment  
31 metadata:  
32   name: jsreport-server  
33 spec:  
34   selector:  
35     matchLabels:  
36       app: jsreport-server  
37   template:  
38     metadata:  
39       labels:  
40         app: jsreport-server  
41     spec:  
42       containers:  
43         - name: jsreport-server  
44           image: "quay.io/jsreport/js-report-server:2.4.1"  
45           ports:  
46             - containerPort: 8080  
47             - containerPort: 8081  
48           volumeMounts:  
49             - name: jsreport-server-home-data  
50               mountPath: /app/data  
51       volumes:  
52         - name: jsreport-server-home-data  
53           persistentVolumeClaim:  
54             claimName: jsreport-server-pvc-data # must match the claim name  
55       securityContext:  
56         allowPrivilegeEscalation: false  
57         capabilities:  
58           drop:  
59             - CAP_NET_RAW  
60             - CAP_SYS_NICE  
61             - CAP_SYS_PTRACE  
62             - CAP_SYS_RESOURCE  
63             - CAP_SYS_TIME  
64             - CAP_SYS_WAIT  
65           add:  
66             - CAP_CHOWN  
67             - CAP_FOWNER  
68             - CAP_FSETID  
69             - CAP_MKNOD  
70             - CAP_SETGID  
71             - CAP_SETUID  
72             - CAP_SETPCAP  
73             - CAP_MAC_OVERRIDE  
74             - CAP_MAC_ADMIN  
75             - CAP_IPC_OWNER  
76             - CAP_NET_BIND_SERVICE  
77             - CAP_NET_RAW  
78             - CAP_SYS_NICE  
79             - CAP_SYS_PTRACE  
80             - CAP_SYS_RESOURCE  
81             - CAP_SYS_TIME  
82             - CAP_SYS_WAIT  
83           drop:  
84             - CAP_CHOWN  
85             - CAP_FOWNER  
86             - CAP_FSETID  
87             - CAP_MKNOD  
88             - CAP_SETGID  
89             - CAP_SETUID  
90             - CAP_SETPCAP  
91             - CAP_MAC_OVERRIDE  
92             - CAP_MAC_ADMIN  
93             - CAP_IPC_OWNER  
94             - CAP_NET_BIND_SERVICE  
95             - CAP_NET_RAW  
96             - CAP_SYS_NICE  
97             - CAP_SYS_PTRACE  
98             - CAP_SYS_RESOURCE  
99             - CAP_SYS_TIME  
100            - CAP_SYS_WAIT
```

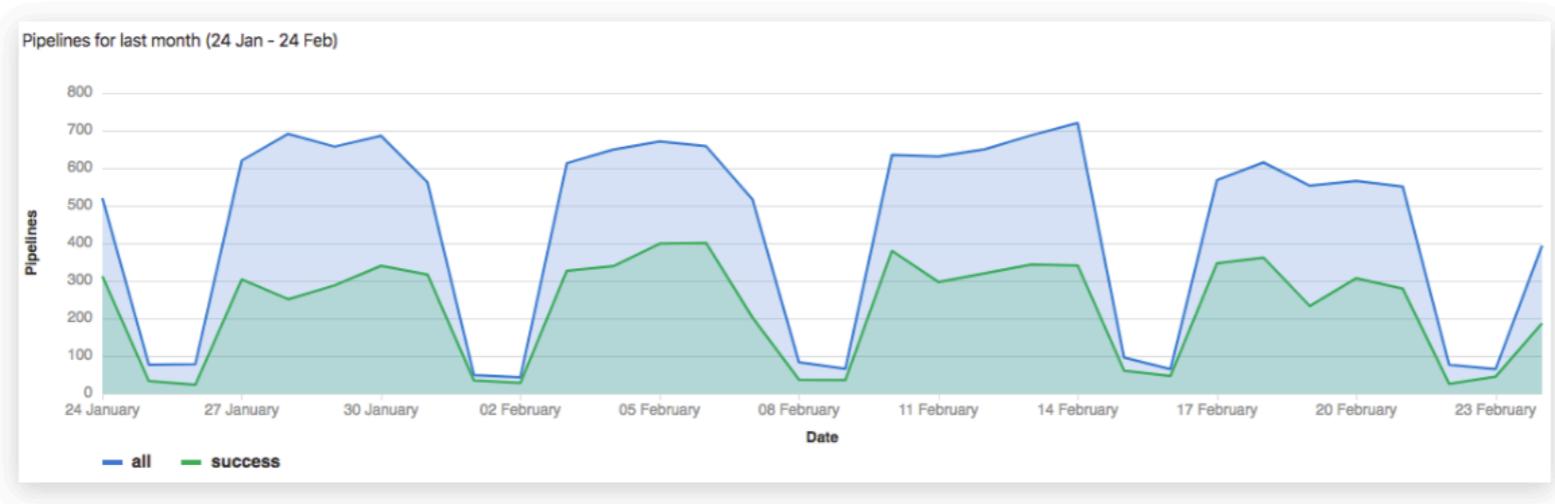
Application install/upgrade/delete with helm chart

Pipeline success and duration charts

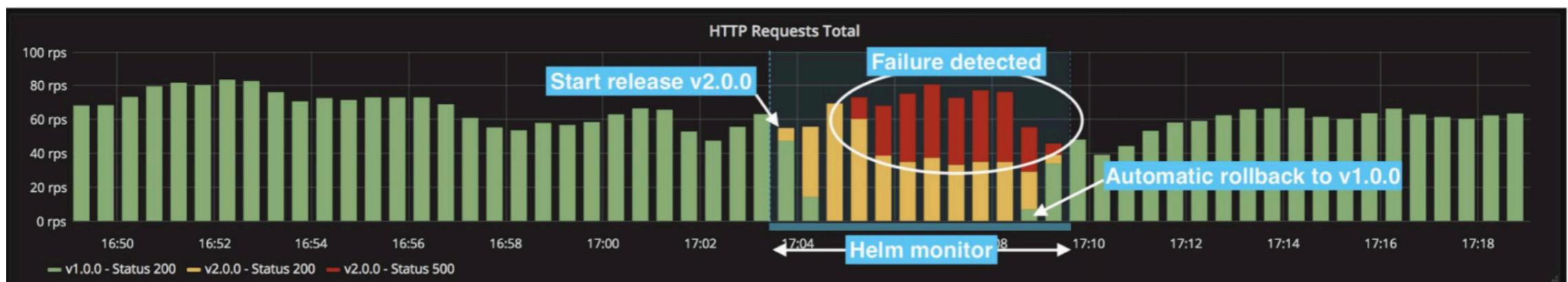
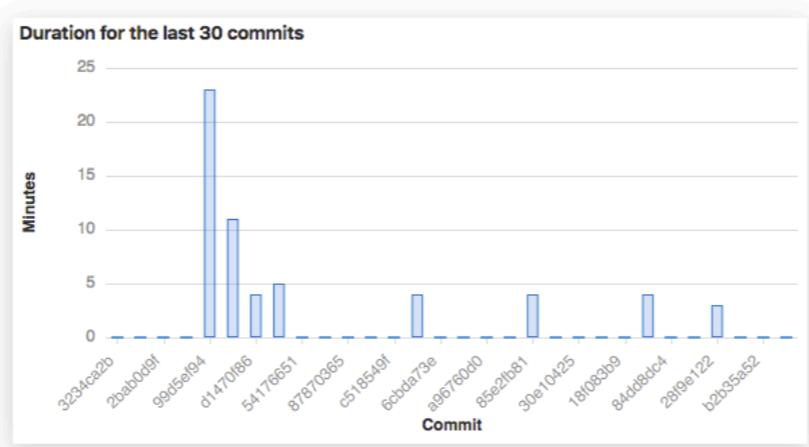
Version history [...](#)

GitLab tracks the history of your pipeline successes and failures, as well as how long each pipeline ran. To view this information, go to [Analytics > CI / CD Analytics](#).

View successful pipelines:



View pipeline duration history:



Representation of a failed release using Prometheus and Grafana

Helm Chart category

```
qzhaos-MBP:dev-k8s qzhao$ tree k8s-helm-chart/
k8s-helm-chart/
├── server_configmap.yaml
├── README.md
├── server_ingress.yaml
├── img
│   └── 53.png
├── jsreport-server
│   ├── ingress
│   │   ├── README.md
│   │   ├── jsreport-server_configmap.yaml
│   │   └── jsreport-server_nginx.yaml
│   ├── jsreport-server
│   │   ├── Chart.yaml
│   │   ├── charts
│   │   ├── templates
│   │   │   ├── NOTES.txt
│   │   │   ├── _helpers.tpl
│   │   │   ├── deployment.yaml
│   │   │   ├── ingress.yaml
│   │   │   ├── service.yaml
│   │   │   └── serviceaccount.yaml
│   │   ├── tests
│   │   │   └── test-connection.yaml
│   │   └── values.yaml
│   └── pv
│       └── jsreport-server-storage-pv-pvc-aws.yaml
└── storageclass.yaml

7 directories, 16 files
```

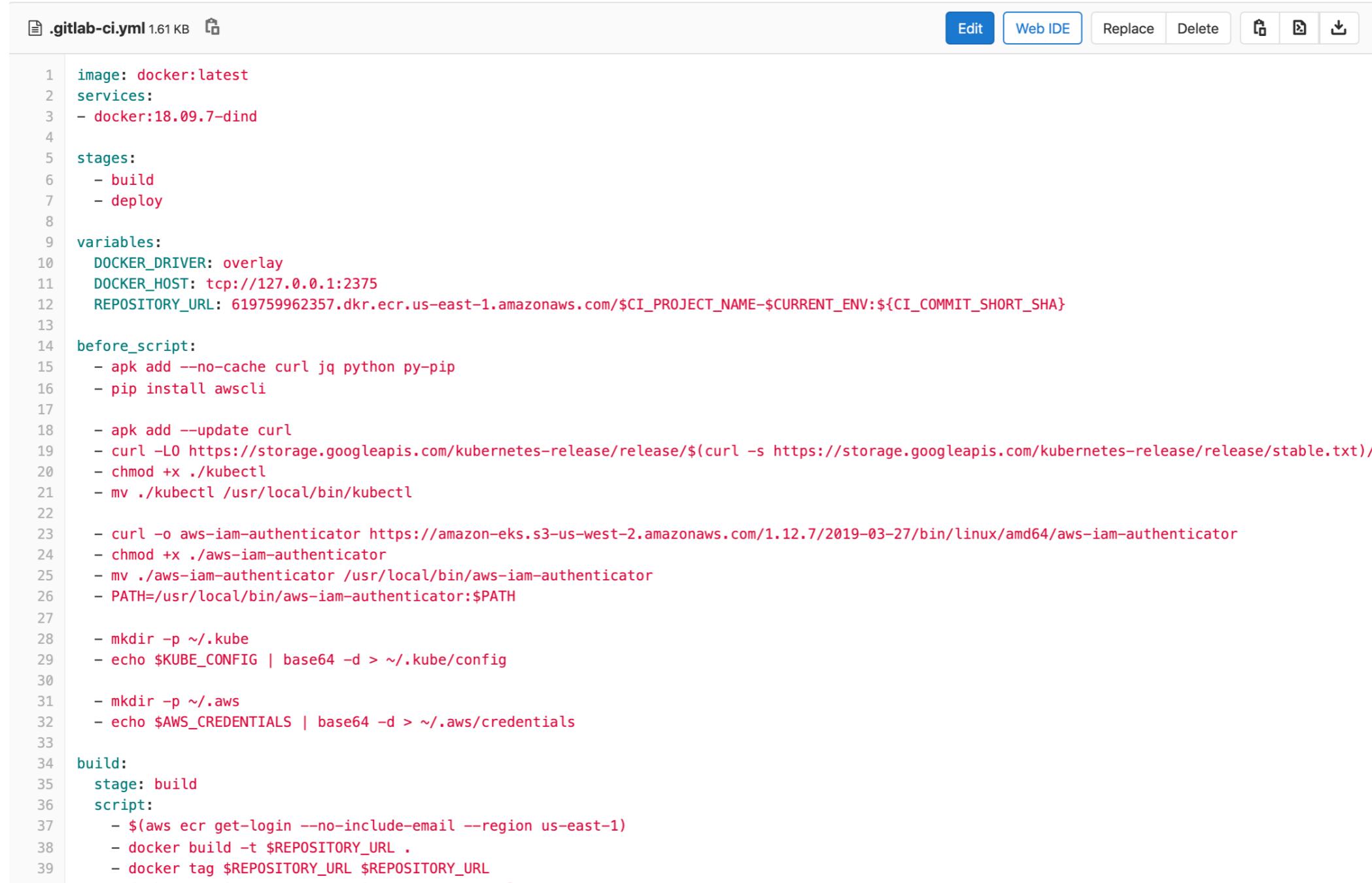
- <https://gitlab.americanportfolios.com/infrastructure/k8s-helm-chart>

Gitlab + k8s

The screenshot shows the GitLab interface for managing Kubernetes clusters. The URL in the address bar is `gitlab.americanportfolios.com/qzhao/go-hello-world-k8s/-/clusters/13`. The page title is "k8stst-aws". The left sidebar has a "G" icon and several other navigation items. The main content area has tabs for "Details", "Applications", and "Advanced Settings", with "Details" selected. A "GitLab Integration" toggle switch is turned on. Under "Environment scope", there is a field containing an asterisk (*) and a note explaining it is the default environment scope for the cluster. A "Base domain" field is present with a note about using it for Auto Review Apps and Auto Deploy stages. A "Save changes" button is at the bottom of this section. Below this is a "Provider details" section, which is collapsed. It contains fields for "Kubernetes cluster name" (set to "k8stst-aws"), "API URL" (set to `https://F4D970FCD66FFCD9CC38D7F5DF7AE9B5.gr7.us-east-1.eks.amazonaws.com`), and "CA Certificate" (containing a long string of certificate data).

- Gitlab community version only allow to set 1 kubernetes

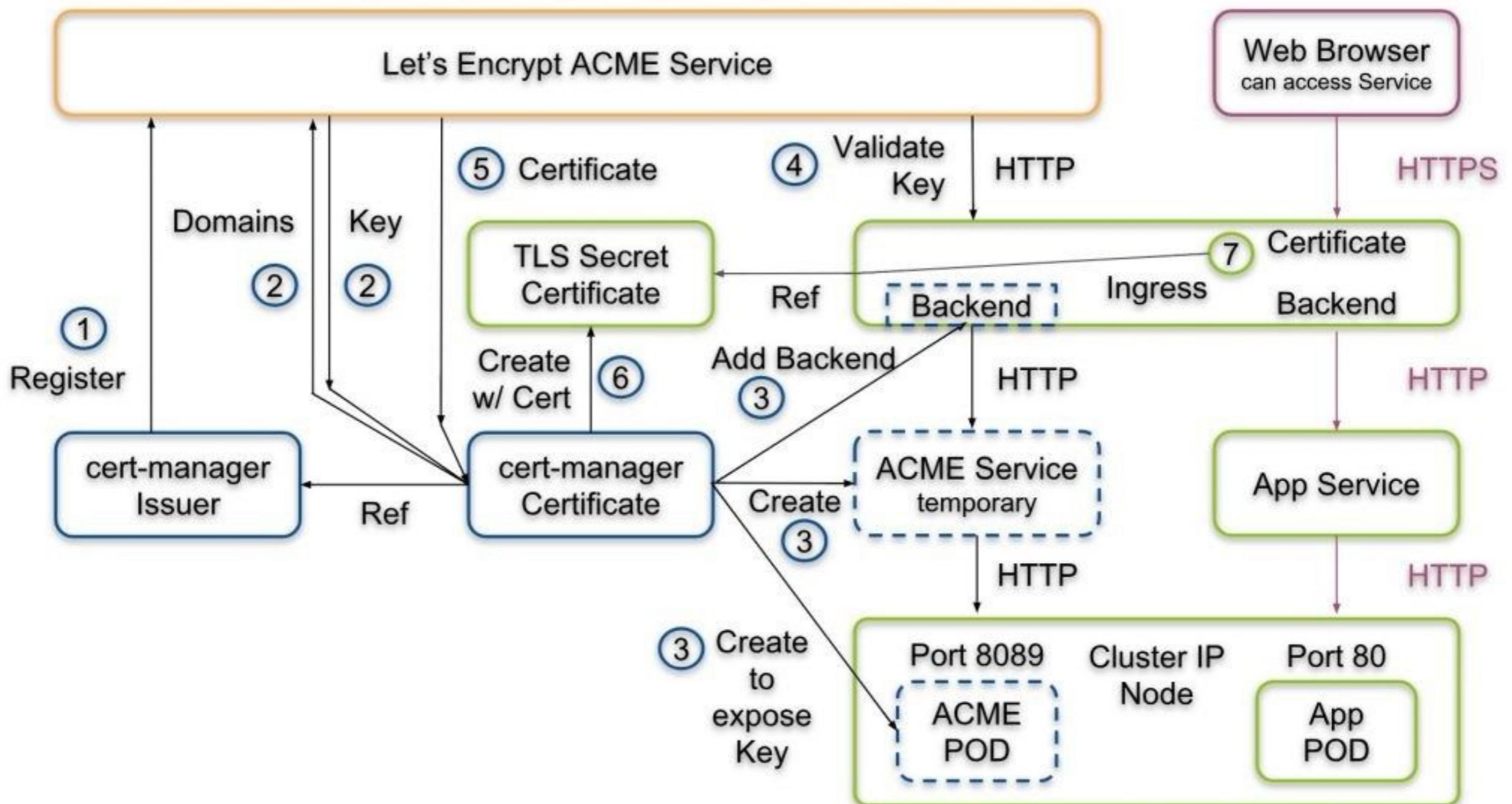
helm is called in .gitlab-ci.yaml(pipeline script)



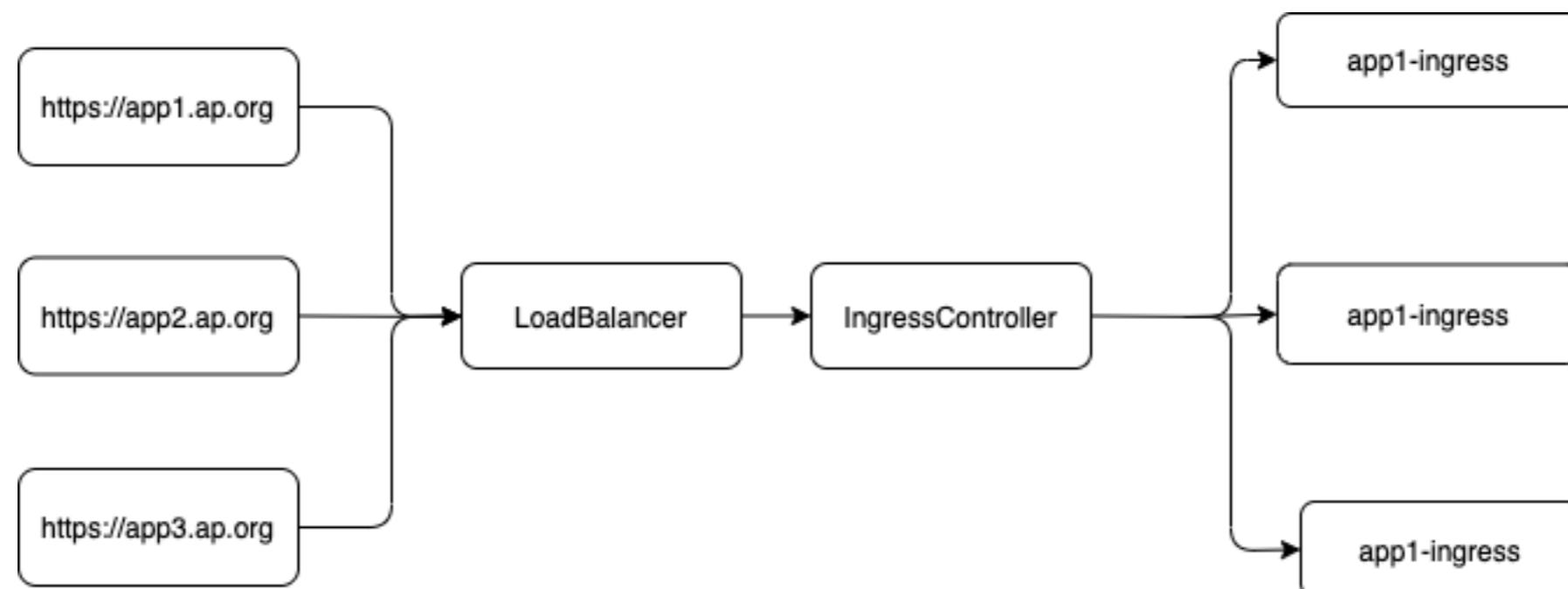
```
1  image: docker:latest
2  services:
3    - docker:18.09.7-dind
4
5  stages:
6    - build
7    - deploy
8
9  variables:
10   DOCKER_DRIVER: overlay
11   DOCKER_HOST: tcp://127.0.0.1:2375
12   REPOSITORY_URL: 619759962357.dkr.ecr.us-east-1.amazonaws.com/$CI_PROJECT_NAME-$CURRENT_ENV:${CI_COMMIT_SHORT_SHA}
13
14 before_script:
15   - apk add --no-cache curl jq python py-pip
16   - pip install awscli
17
18   - apk add --update curl
19   - curl -L0 https://storage.googleapis.com/kubernetes-release/release/$(curl -s https://storage.googleapis.com/kubernetes-release/release/stable.txt)/
20   - chmod +x ./kubectl
21   - mv ./kubectl /usr/local/bin/kubectl
22
23   - curl -o aws-iam-authenticator https://amazon-eks.s3-us-west-2.amazonaws.com/1.12.7/2019-03-27/bin/linux/amd64/aws-iam-authenticator
24   - chmod +x ./aws-iam-authenticator
25   - mv ./aws-iam-authenticator /usr/local/bin/aws-iam-authenticator
26   - PATH=/usr/local/bin/aws-iam-authenticator:$PATH
27
28   - mkdir -p ~/.kube
29   - echo $KUBE_CONFIG | base64 -d > ~/.kube/config
30
31   - mkdir -p ~/.aws
32   - echo $AWS_CREDENTIALS | base64 -d > ~/.aws/credentials
33
34 build:
35   stage: build
36   script:
37     - $(aws ecr get-login --no-include-email --region us-east-1)
38     - docker build -t $REPOSITORY_URL .
39     - docker tag $REPOSITORY_URL $REPOSITORY_URL
        . . . . .
```

- <https://gitlab.americanportfolios.com/qzhao/go-hello-world-k8s/-/blob/master/.gitlab-ci.yaml>

CertManager+Let's encrypt

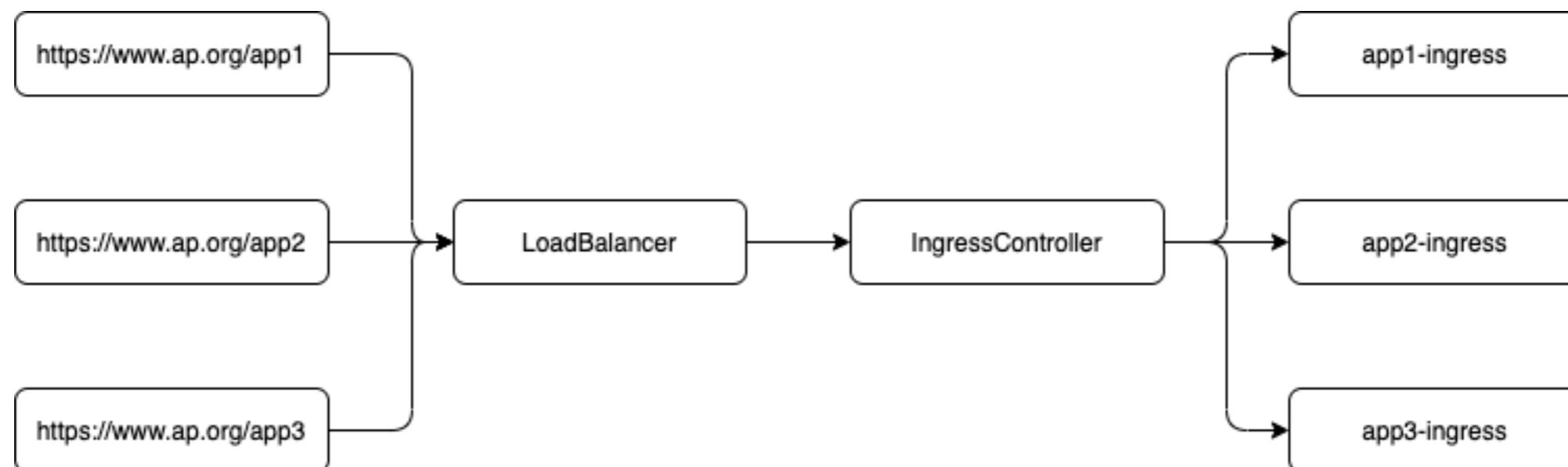


IngressController(k8s-official)



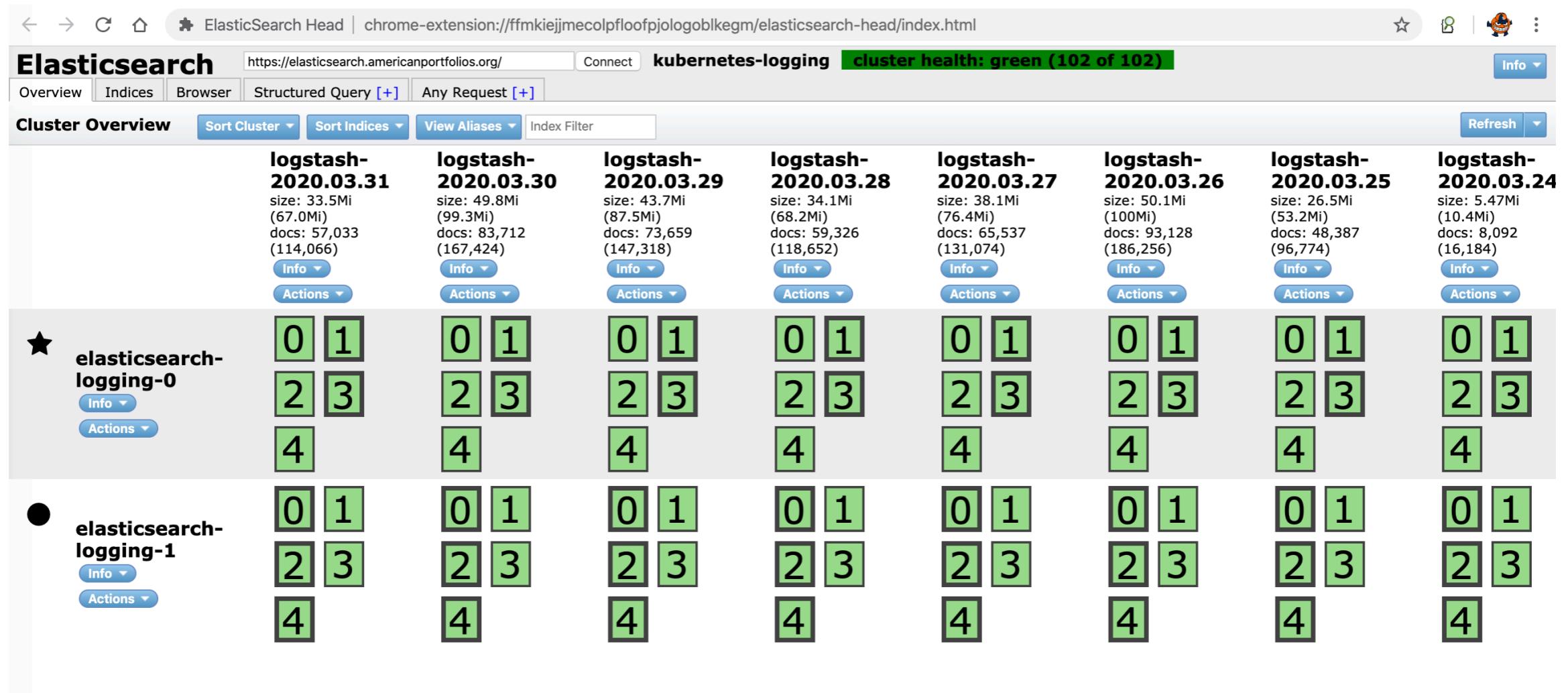
- https://gitlab.americanportfolios.com/infrastructure/k8s-helm-chart/-/blob/master/ingress/jsreport-server_ingress.yaml

IngressController(nginx-official)



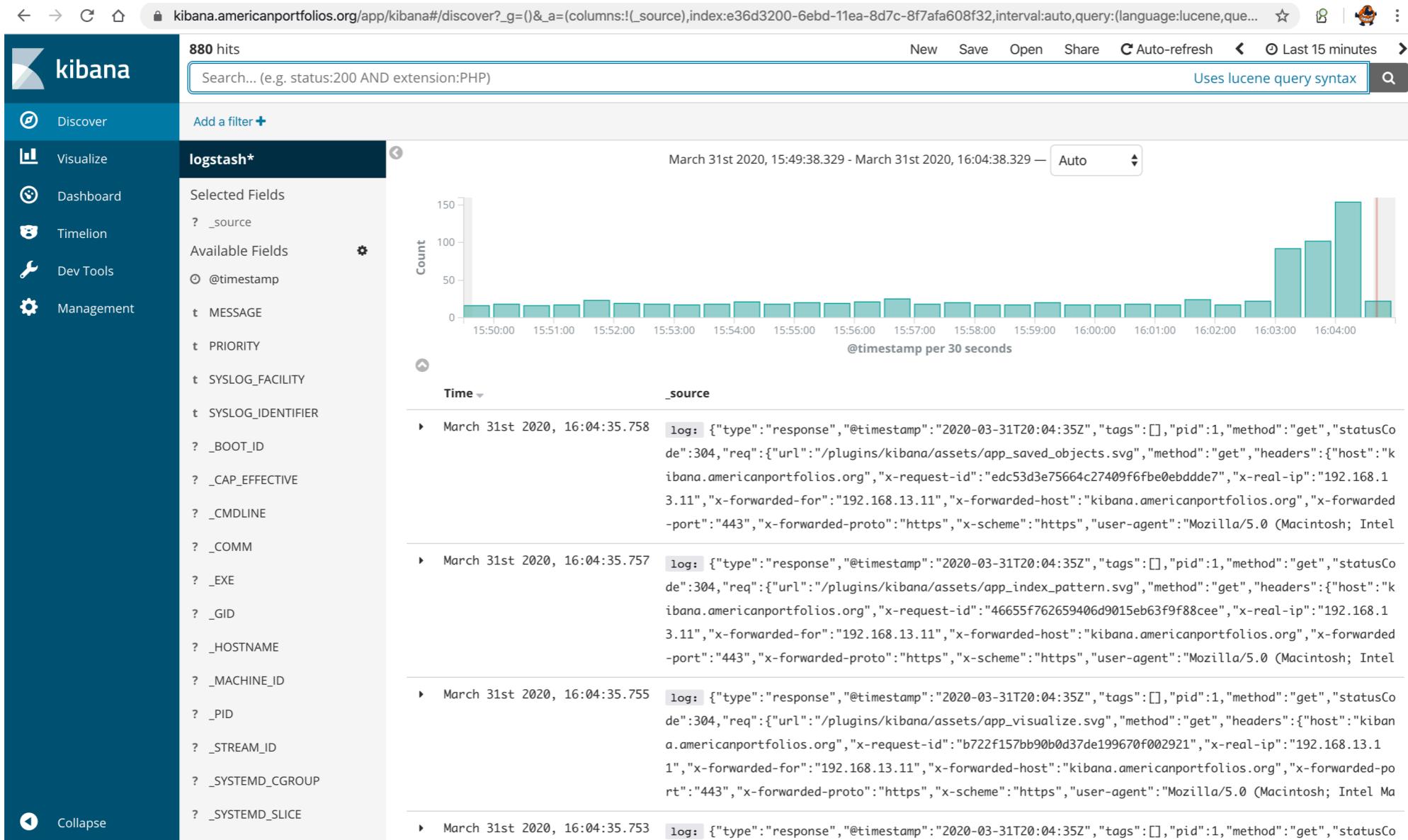
- <https://americanportfolios.atlassian.net/wiki/spaces/PD/pages/1153302541/Set+ingress-nginx+nginx-official+as+a+loadbalancer>

ElasticSearch



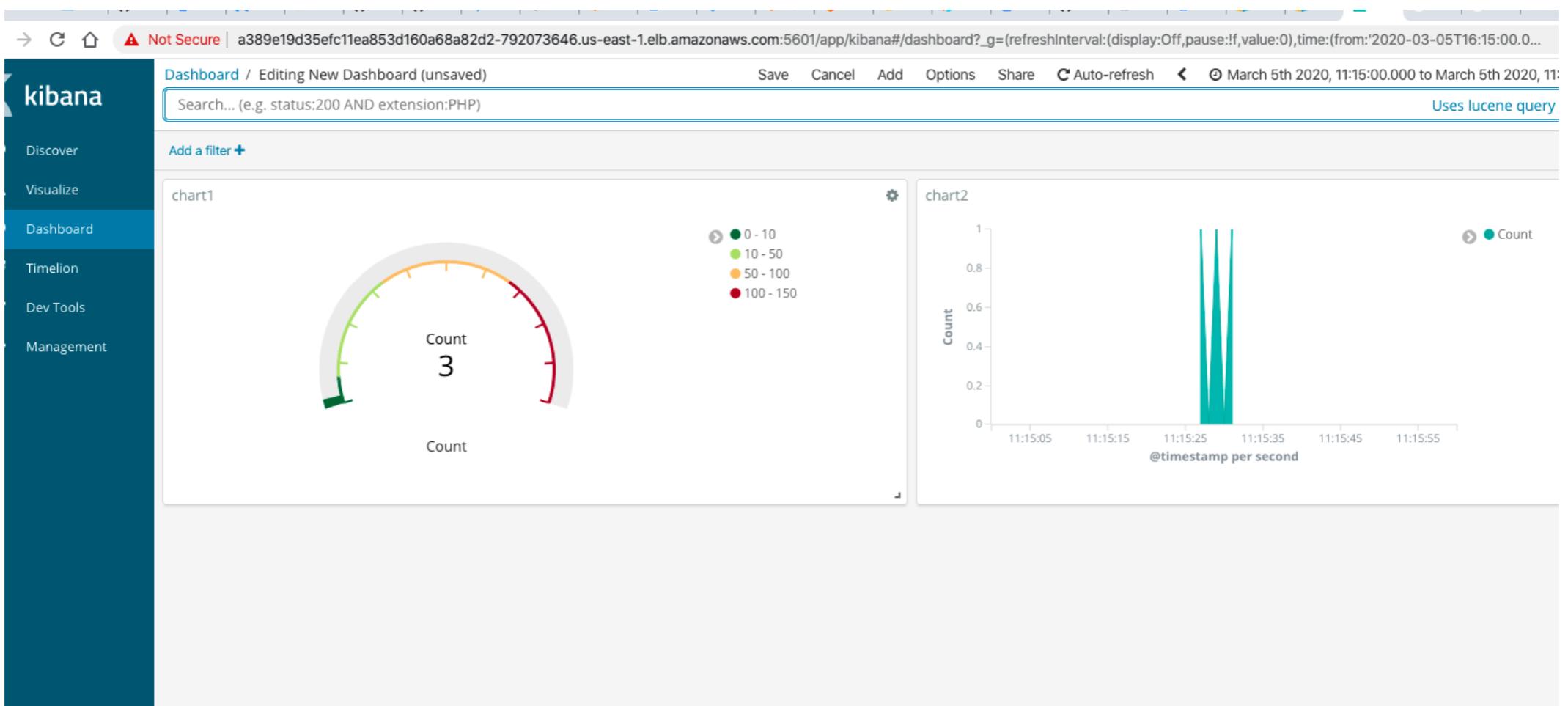
- <https://elasticsearch.americanportfolios.org/>

Fluentd+Kibana



- <https://kibana.americanportfolios.org>

Kibana dashboard



Weave

The screenshot shows the WeaveScope interface, a monitoring tool for a Weave-based network. The main dashboard displays a network graph with 40 nodes, where each node is represented by a hexagon icon. Nodes include locust, nginx-ingress-c, sentry-workers, sentry-sentry-r..., sentry-cron, sentry-sentry-r..., sentry-web, prometheus, node-exporter, grafana, kube-state-met..., alertmanag, echo1, rules-configma..., config-reloader, prometheus-co..., cert, echo1, jenkins, cert, jsreport-server, nginx-ingress-c, and various dashboards like grafana-sc-das... and tiller.

The central part of the interface features a large hexagonal icon for the **jsreport-server** container, which is highlighted. A modal window provides detailed information about this container:

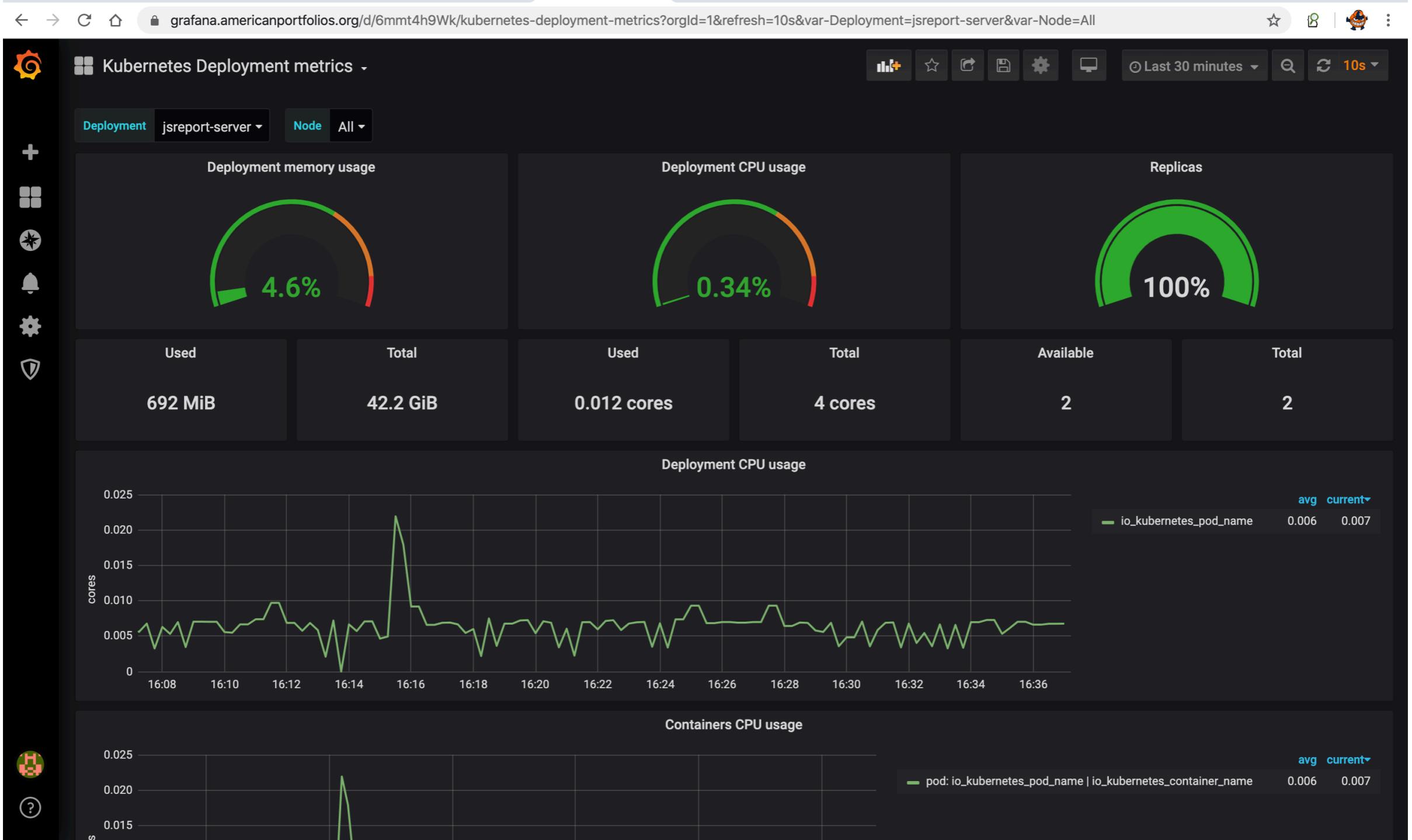
- jsreport-server**
jsreport-server/tst jsreport-server-896dd9d...
ip-192-168-57-201
- Status**
CPU: 0.11 % | Memory: 172.4 MB
- Info**
 - Image tag: latest
 - Image name: infrastructure/jsreport-server/tst
 - Command: /bin/sh /start
 - State: Up 7 hours
 - Uptime: 7 hours
 - Restart #: 0
 - IPs: 192.168.37.202
- Processes**

Process	PID	CPU	Memory
node	12362	0.00 %	60.4 MB
/bin/sh	12337	0.00 %	688 KB

At the bottom of the interface, there are tabs for **All**, **System containers**, and **Application containers**, with **Running containers** selected. Other buttons include **Show uncontained** and **Hide uncontained**. The footer shows the version **Version 1.13.0 on weave-scope-app-7b679dfcd5f** and some icons.

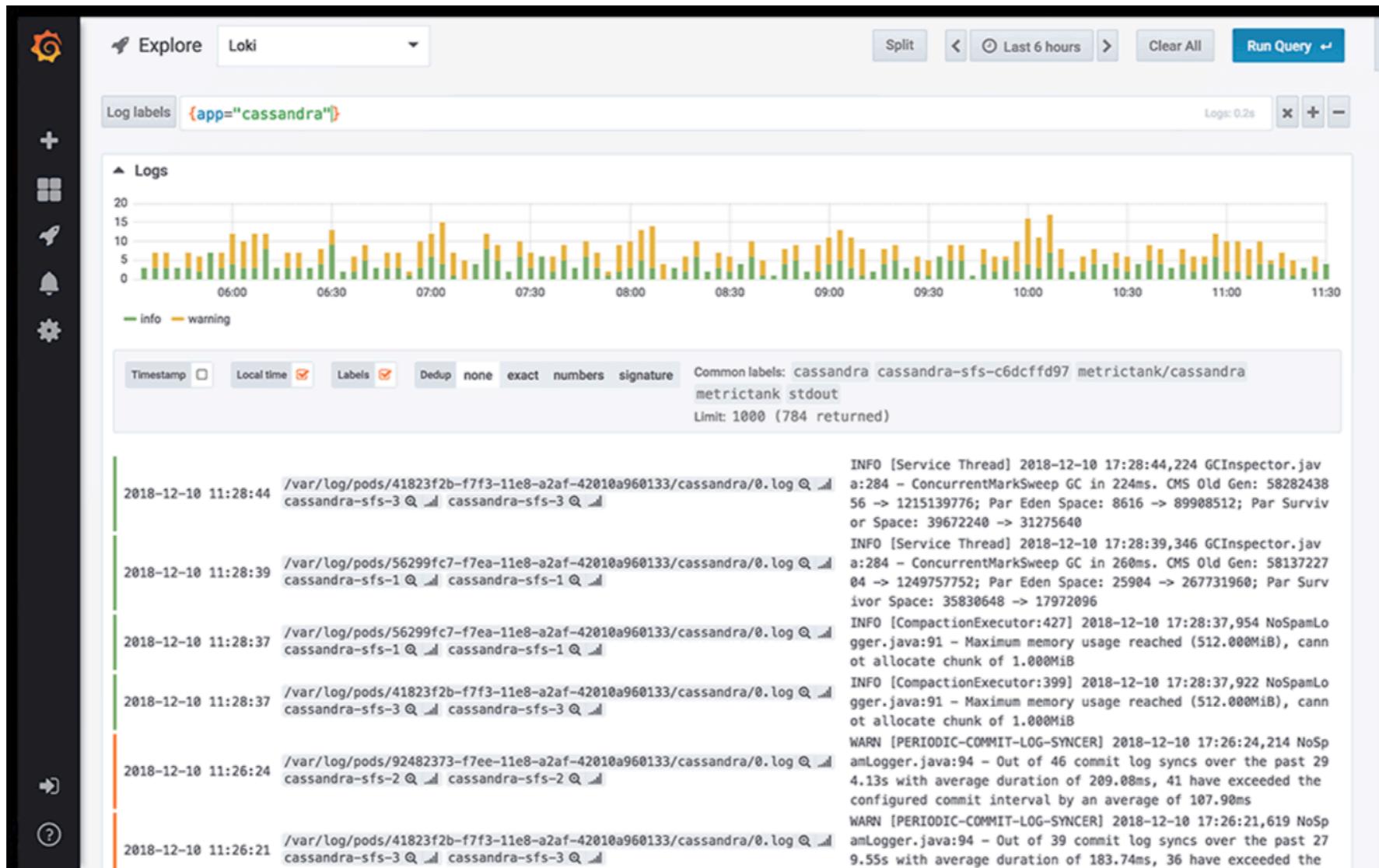
- <https://weave.americanportfolios.org/>

Grafana



- <https://grafana.americanportfolios.org/>

Loki



- Loki is a horizontally-scalable, highly-available, multi-tenant log aggregation system.

Prometheus

The screenshot shows the Prometheus web interface at <http://prometheus.americanportfolios.org/graph>. The top navigation bar includes links for Prometheus, Alerts, Graph, Status, and Help. Below the navigation is a toolbar with a checkbox for "Enable query history", a "Try experimental React UI" link, and a text input field labeled "Expression (press Shift+Enter for newlines)". A blue "Execute" button is next to a dropdown menu that says "- insert metric at cursor -". Below this is a tab bar with "Graph" selected and "Console" as an option. A search bar contains the word "Moment". On the right, there's a table with two columns: "Element" and "Value", which currently displays "no data". At the bottom right is a "Remove Graph" button, and at the bottom left is a blue "Add Graph" button.

- <http://prometheus.americanportfolios.org/>

Alert Manager

The screenshot shows the Alert Manager interface at alertmanager.americanportfolios.org/#/alerts. The top navigation bar includes links for Alertmanager, Alerts, Silences, Status, Help, and a prominent blue button labeled "New Silence". Below the navigation is a search bar with "Filter" and "Group" tabs, and a "Custom matcher, e.g. `env=production`" input field with a "+" button and a "Silence" icon.

On the right, there are filter options: "Receiver: All" with three radio buttons for "All", "Silenced", and "Inhibited". Below these are three alert groups:

- + Expand all groups
- + Not grouped 4 alerts
- + Not grouped 1 alert
- + job="kube-proxy" + 1 alert

- <https://alertmanager.americanportfolios.org/>

Alert Manager + MS Team for tst/uat/sit/prod



- This is in progressing

Sentry

The screenshot shows the Sentry interface for an internal issue. The URL in the browser is <https://sentry.americanportfolios.org/sentry/internal/issues/1/>.

Left Sidebar:

- Sentry v admin@sentry.local
- Projects
- Assigned to me
- Bookmarked issues
- Recently viewed
- Activity
- Stats
- Settings

Top Bar:

- internal
- Environment: All environments

Issue Details:

TypeError poll(..../sentry/scripts/views.js)

Object [object Object] has no method 'updateFrom'

Action Buttons: Resolve, Ignore, Share

Issue Summary: ISSUE # INTERNAL-1 | EVENTS 1 | USERS 1 | ASSIGNEE

Event Details: Event ac5c7eacf7c8485fb1dfc10112f921e9 | Mar 26, 2020 3:34:07 PM UTC | JSON (5.9 KB)

User & Device: mail@example.... ID: 1 | Chrome Version: 65.0.3325 | Mac OS X Version: 10.13.4

Tags: browser: Chrome 65.0.3325, browser.name: Chrome, level: error, os: Mac OS X 10.13.4, os.name: Mac OS X, url: http://example.com/foo, user: id:1

Message: This is an example JavaScript exception

Exception: EXCEPTION (most recent call first) | Get a tour of the issue page | Only | Full | Raw

Ownership Rules: Create Ownership Rule

Environment Metrics: All Environments (LAST 24 HOURS), LAST 30 DAYS

First Seen: When: 5 days ago (Mar 26, 2020 3:34:07 PM UTC), Release: not configured

Last Seen: When: 5 days ago (Mar 26, 2020 3:34:07 PM UTC), Release: not configured

- Sentry VS. segment
- <https://sentry.americanportfolios.org>

Locust(Performance testing)

The screenshot shows the Locust web interface at locust.americanportfolios.org. The top navigation bar includes icons for back, forward, refresh, and home, along with a lock icon indicating a secure connection. The main header features the Locust logo and the text "LOCUST". Below the header, the "HOST" is listed as <https://echo1.americanportfolios.org>. The "STATUS" is shown as "HATCHING 0 users", with an "Edit" link. The "SLAVES" count is 2, "RPS" is 0, and "FAILURES" are at 0%. A red "STOP" button and a grey "Reset Stats" button are visible.

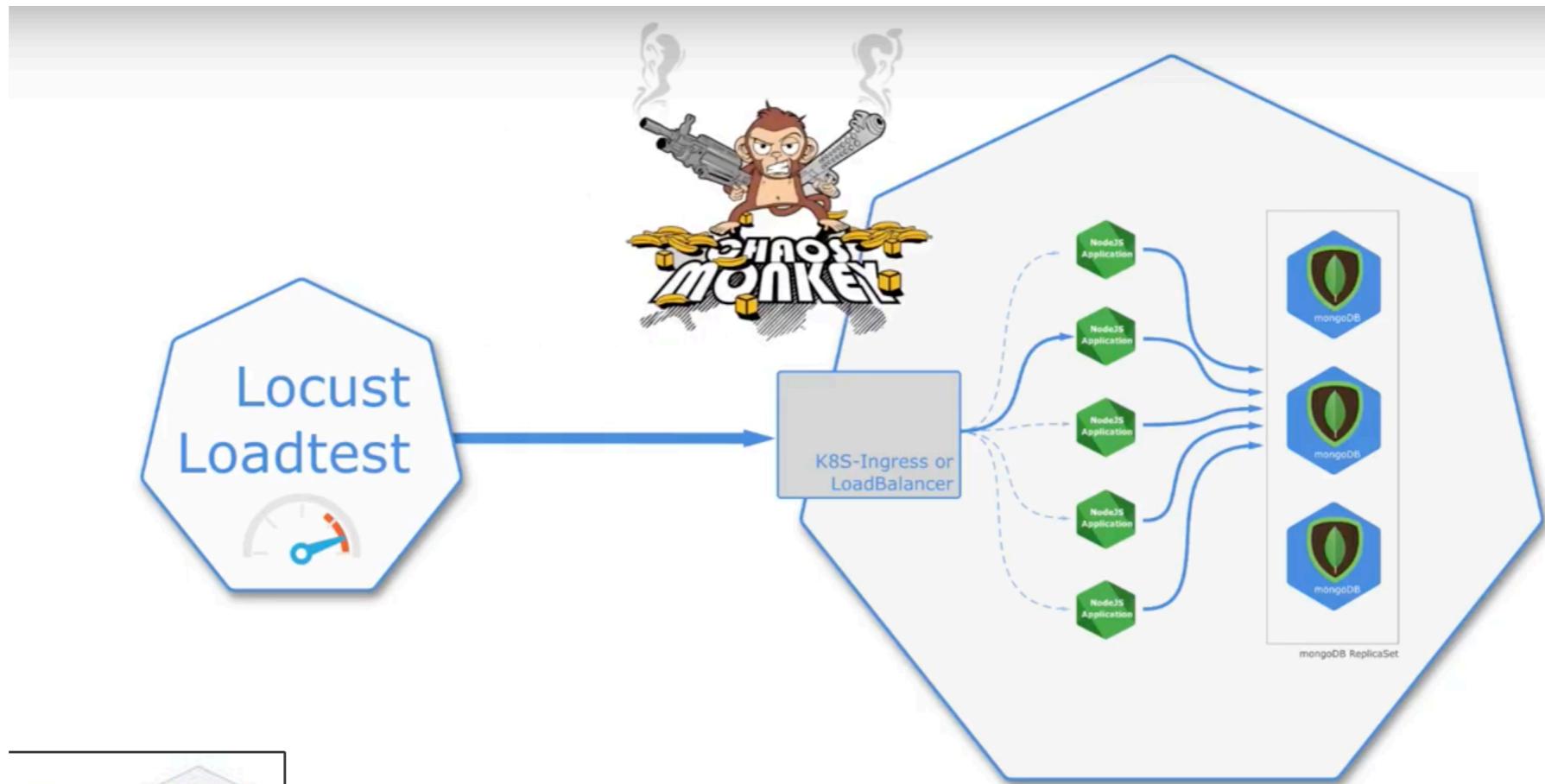
The main content area has tabs for "Statistics", "Charts", "Failures", "Exceptions", "Download Data", and "Slaves". The "Statistics" tab is selected. A table displays performance metrics:

Type	Name	# requests	# fails	Median (ms)	Average (ms)	Min (ms)	Max (ms)	Content Size (bytes)	# reqs/sec
Total		0	0				0	0	0

A modal dialog box is open, titled "Change the locust count". It contains two input fields: "Number of users to simulate" with the value "10" and "Hatch rate (users spawned/second)" with the value "5". A "Close" button is located in the top right corner of the modal. At the bottom of the modal is a green "Start swarming" button.

- <http://locust.americanportfolios.org/>
- <http://locust.io>

Locust + Chao's Monkey (Disaster recovery testing)



- It should randomly break anything in QA env.

Spinnaker

The screenshot shows the Spinnaker web interface at 127.0.0.1:9000/#/applications/jsreport/clusters/serverGroupDetails/kubernetes/default/jsreport-server/replicaSet%20jsreport-server-896dd9d64. The top navigation bar includes links for SPINNAKER, Search, Projects, Applications, a search bar, Help, and user icons.

The main header shows the application name **jsreport** and tabs for PIPELINES, INFRASTRUCTURE, TASKS, and CONFIG. Below this, the INFRASTRUCTURE tab is selected, showing CLUSTERS, LOAD BALANCERS, and FIREWALLS.

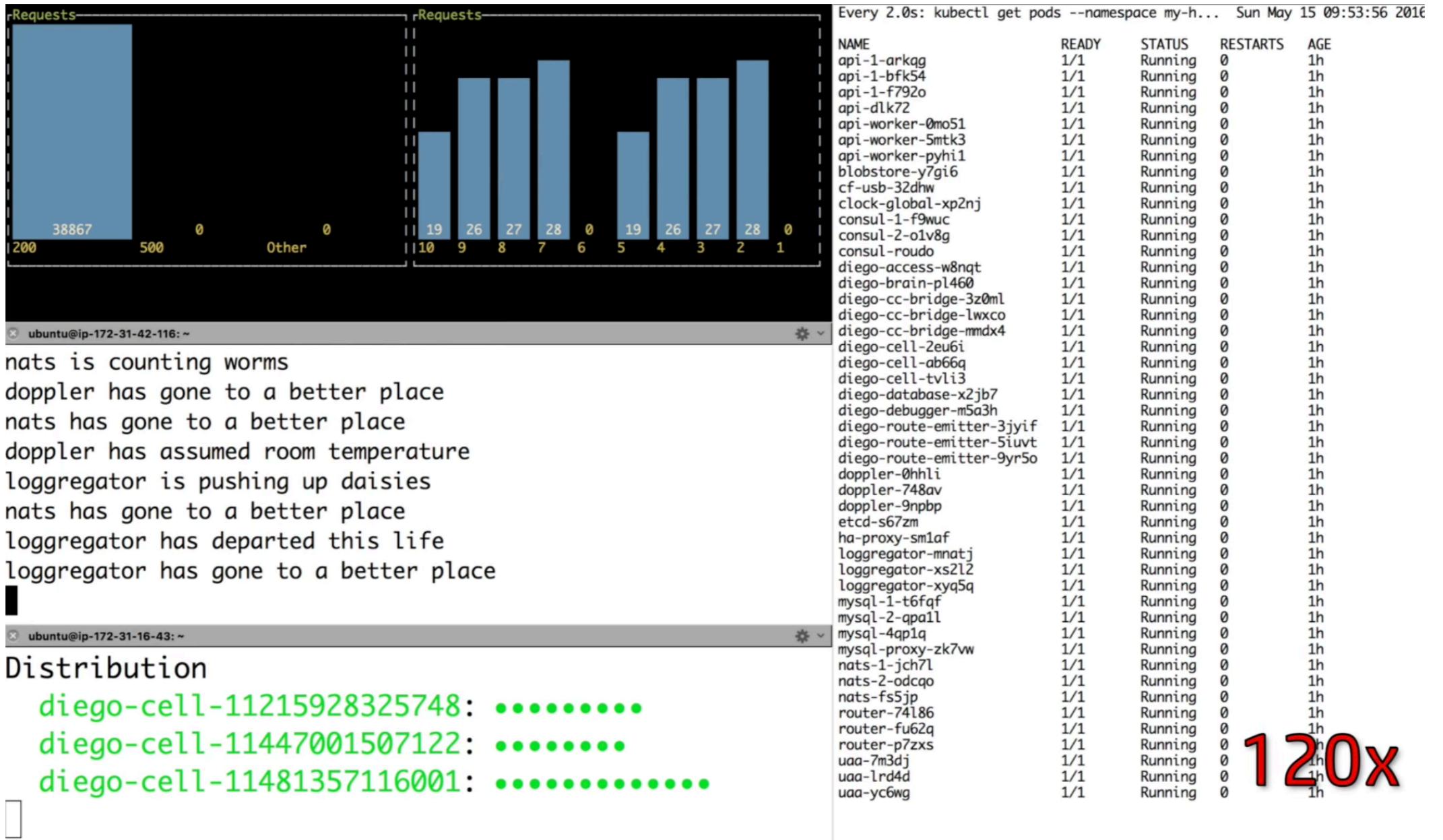
The left sidebar contains filters for SEARCH, ACCOUNT (default), REGION (jsreport-server), STACK ((none)), DETAIL ((none)), STATUS (Healthy, Unhealthy, Disabled, Starting, Out of Service, Unknown), and AVAILABILITY ZONES.

The central cluster view shows a list of clusters under the DEFAULT account. The first cluster is **jsreport-server-896dd9d64**, which is currently active (2 instances at 100% health). Below it is a deployment named **jsreport-server**, and further down is a task named **V001**.

A detailed sidebar for the **jsreport-server-896dd9d64** cluster provides information on Replica Set Actions, Events (no recent events), Labels (app.kubernetes.io/instance: jsreport-server, app.kubernetes.io/name: jsreport-server, pod-template-hash: 896dd9d64), Size (Current 2), and Health.

At the bottom, a message states: "The application **jsreport** has not been configured."

Chao's testing



- <https://www.youtube.com/watch?v=QNOtE-199EQ&t=92s>

Kubernetes deployment

- Recreate
- Rolling-update
- Blue-green
- Canary
- A/B testing

Recreate

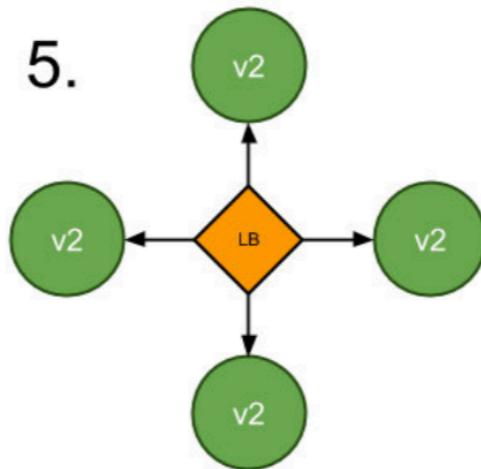
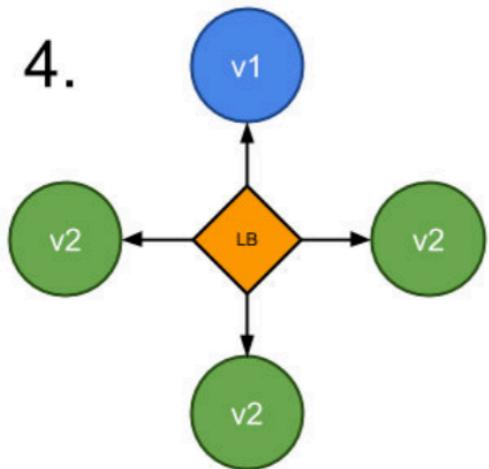
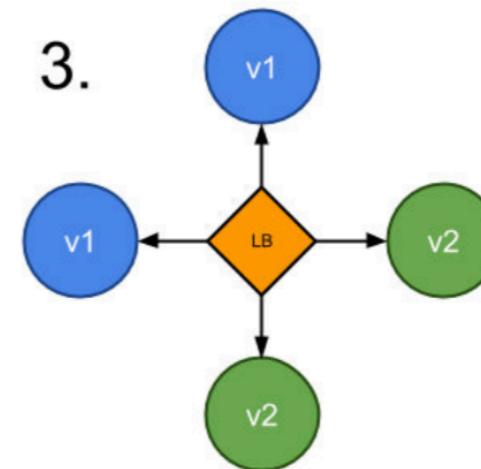
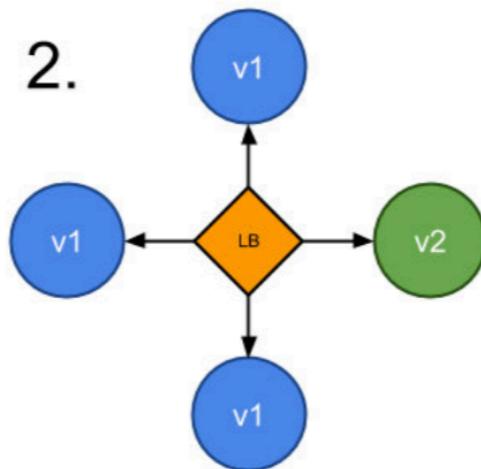
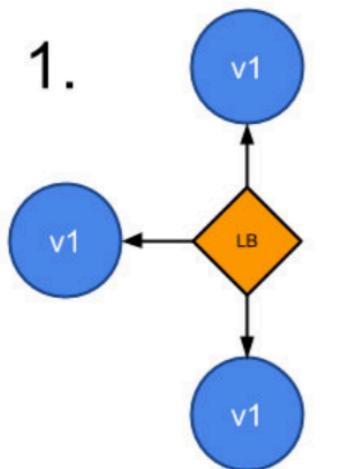
```
spec:  
  replicas: 3  
  strategy:  
    type: Recreate
```



Recreate

- Only for dev env

Rolling-update-1



ramped

- A lot of companies are using rolling update

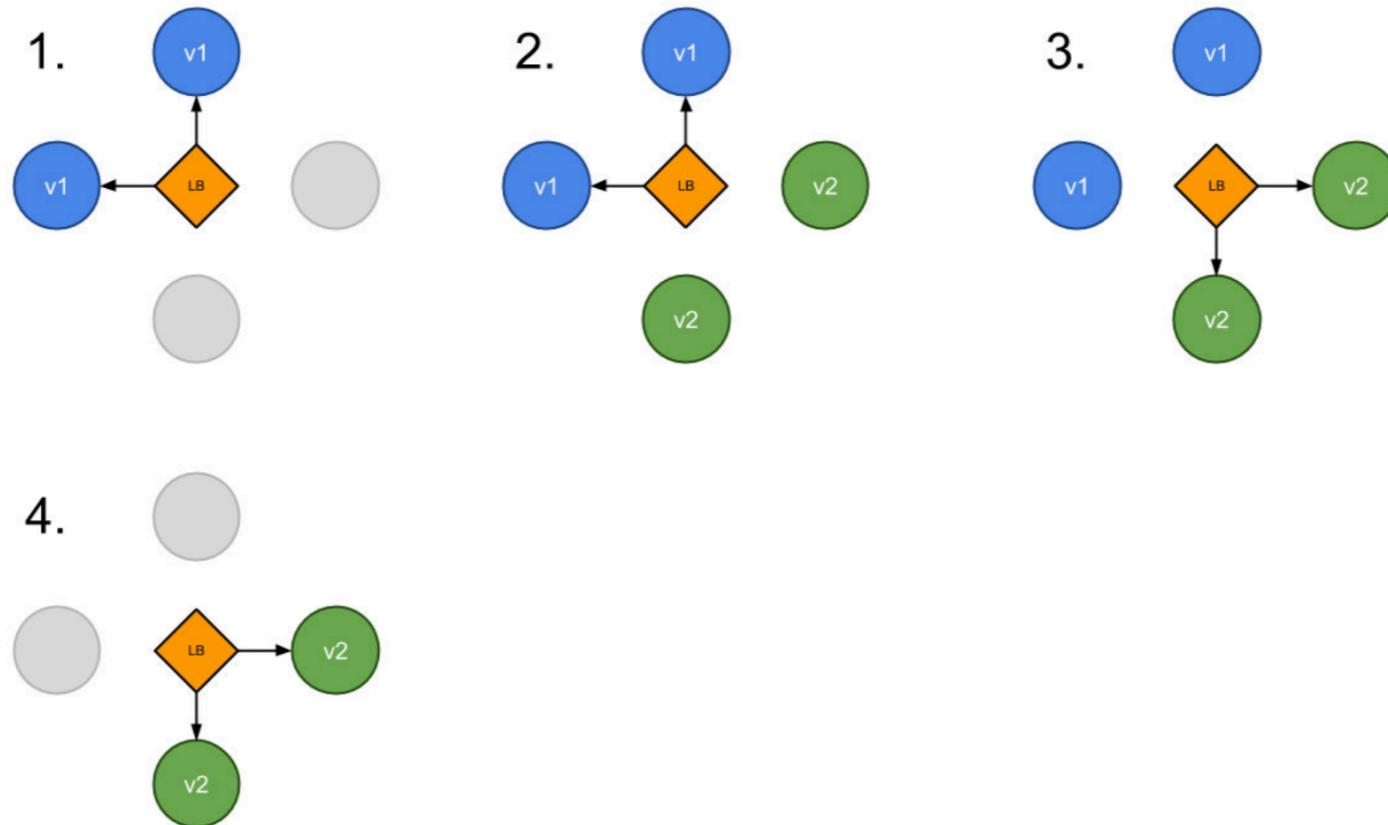
Rolling-update-2

```
spec:  
  replicas: 3  
  strategy:  
    type: RollingUpdate  
    rollingUpdate:  
      maxSurge: 2          # 一次可以添加多少个Pod  
      maxUnavailable: 1    # 滚动更新期间最大多少个Pod不可用
```



rolling-update requests

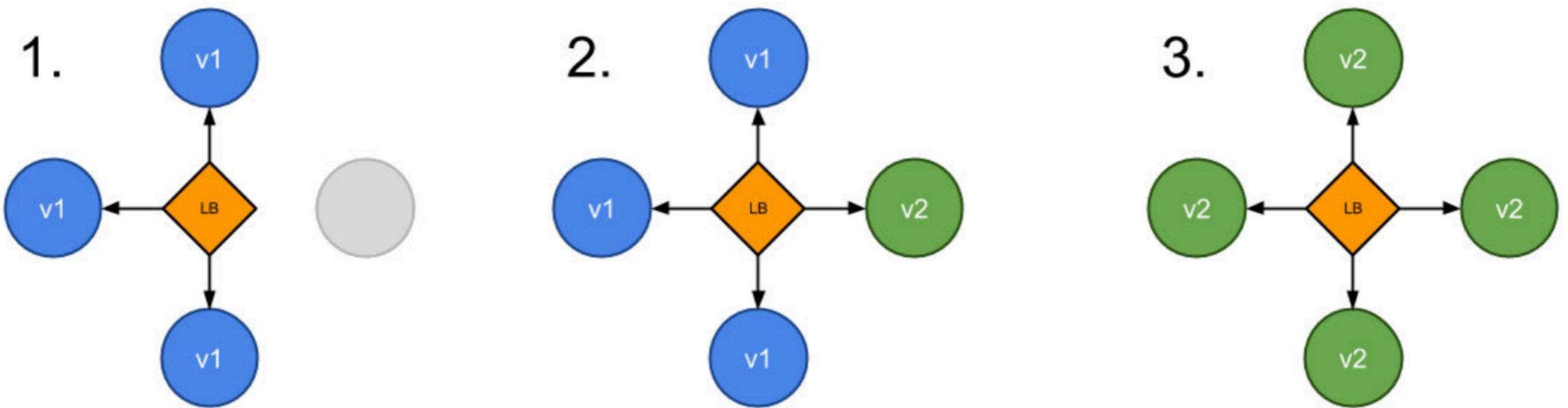
Blue-green



blue/green



Canary

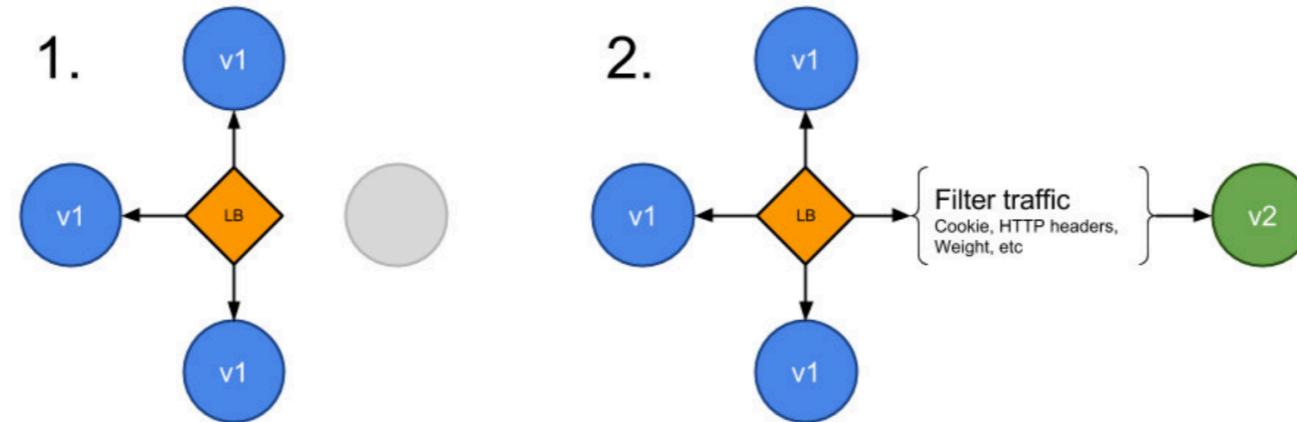


Canary



canary requests

A/B testing



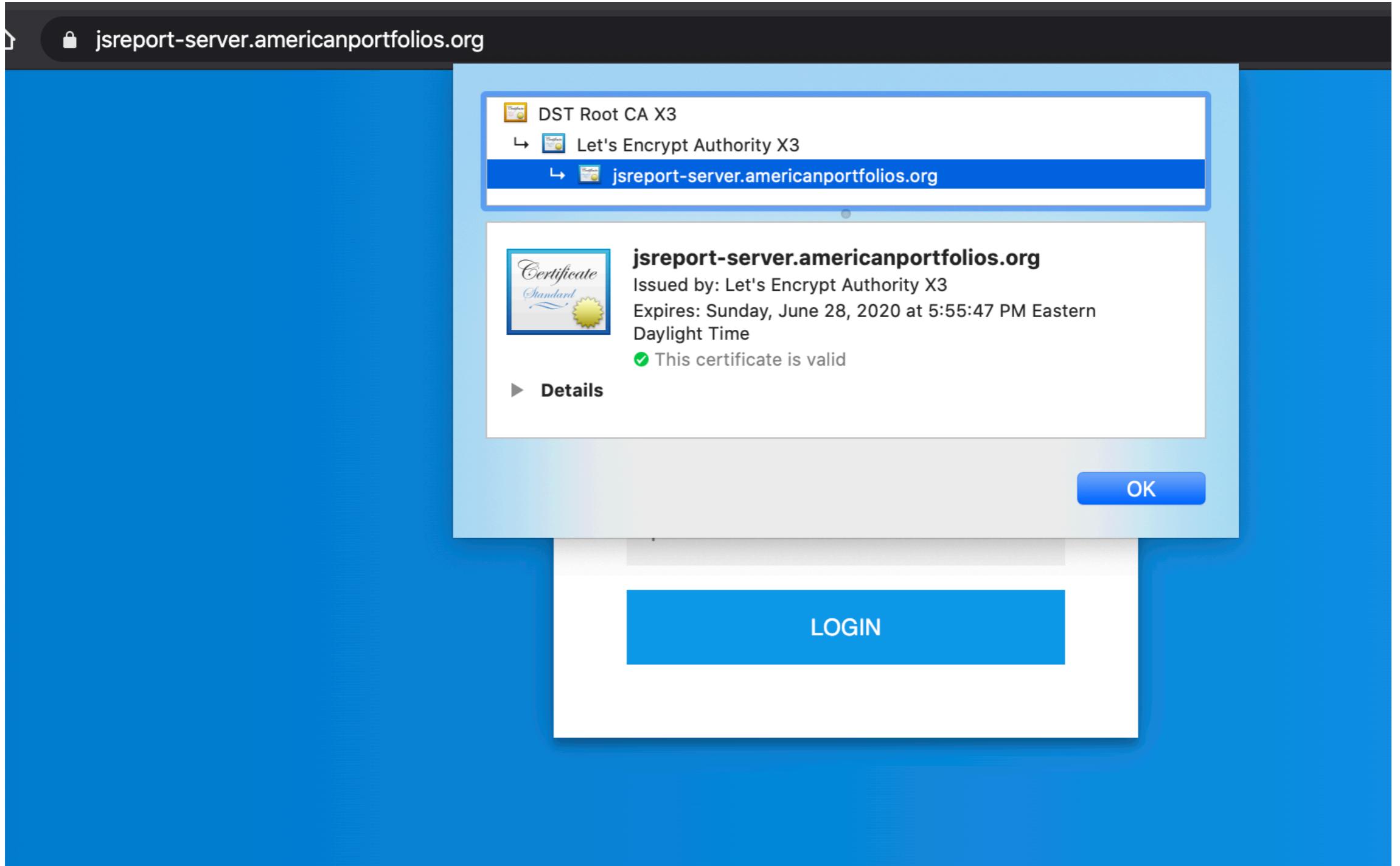
ab test

```
route:  
- tags:  
  version: v1.0.0  
  weight: 90  
- tags:  
  version: v2.0.0  
  weight: 10
```



ab test request

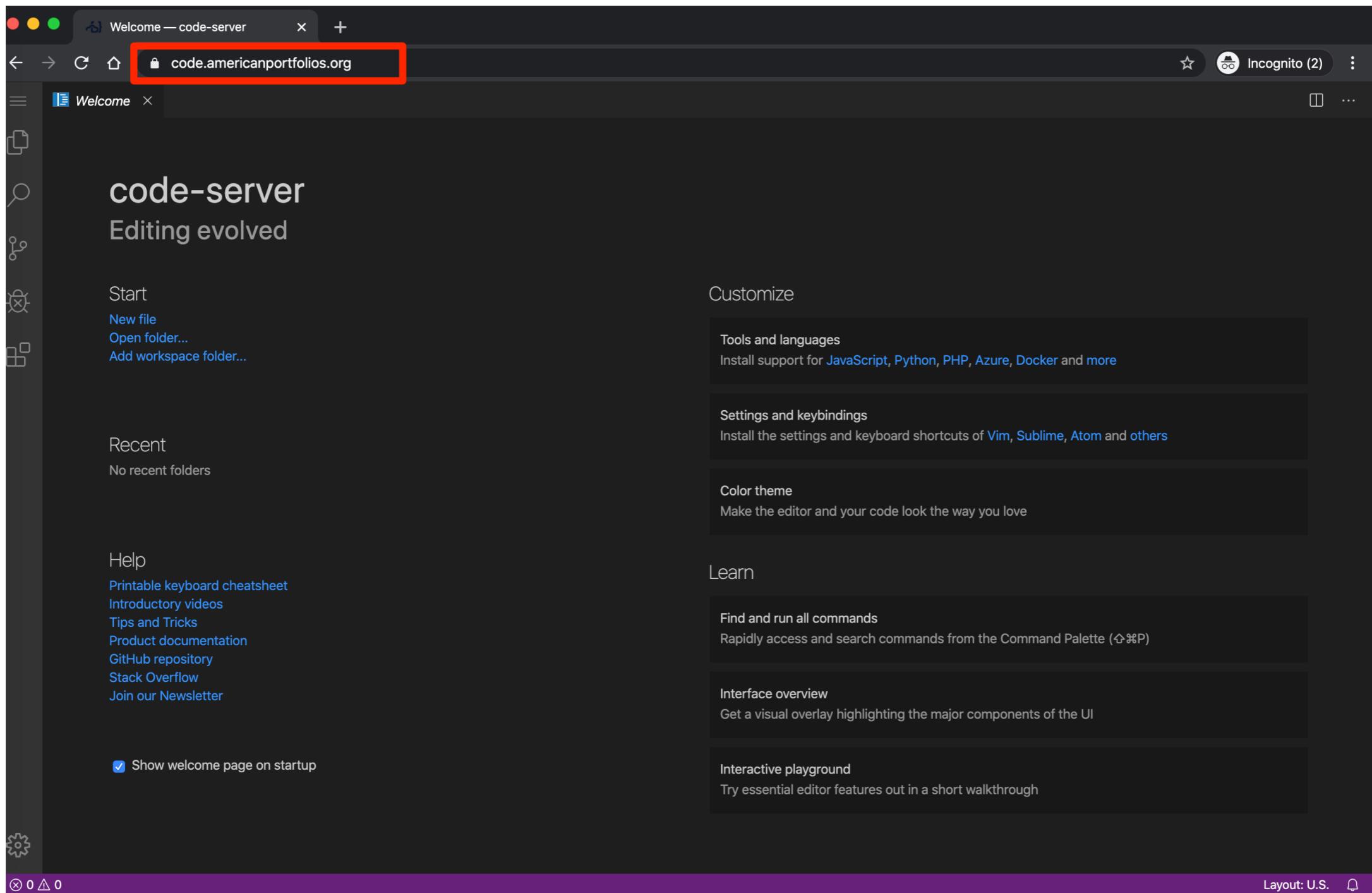
Migration example: jsreport-server



- <https://jsreport-server.americanportfolios.org/>

Migration example: jsreport-server(PersistentVolume testing)

Development tools: VS Code



- <https://code.americanportfolios.org/>
- Password: helloworld

Tool list-1

- Kubernetes <https://kubernetes.io/>
- Elasticsearch <https://www.elastic.co/>
- fluentd <https://www.fluentd.org/>
- kibana <https://www.elastic.co/kibana>
- Prometheus <https://prometheus.io/>
- alertmanager <https://prometheus.io/docs/alerting/alertmanager/>
- grafana <https://grafana.com/>
- Helm2/Tiller <https://v2.helm.sh/>
- Certificate manager <https://cert-manager.io/docs/>
- let's encrypt <https://letsencrypt.org/>
- Ingress from kubernetes <https://kubernetes.github.io/ingress-nginx/deploy/>
- Ingress from Nginx <https://www.nginx.com/products/nginx/kubernetes-ingress-controller/>

Tool list 2

- Weave <https://www.weave.works/docs/net/latest/overview/>
- Loki <https://grafana.com/oss/loki/>
- Sentry/segment <https://sentry.io/>
- Locust <https://locust.io/>
- Spinnaker <https://www.spinnaker.io/>
- Chao's monkey <https://github.com/Netflix/chaosmonkey>
- We will have more later.
 - For example traffic management, istio, kiali, etc...
 - <https://istio.io/>
 - <https://kiali.io/>

Thanks

Q/A