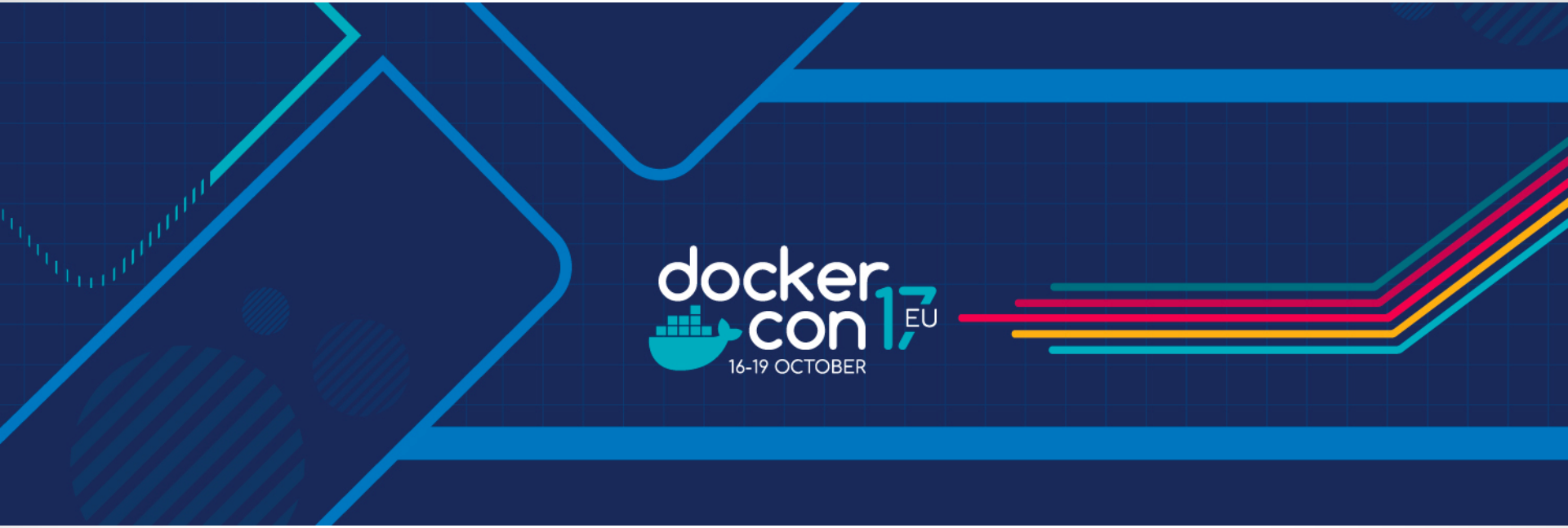
"Swarm & K8S" We are one big community

dockercon17EU

Agenda

- The brief history of Docker and K8S
- Horizontal Pod Autoscaler Concept
- Demo Session
  - Enable kubernetest dashboard/metric server
  - Demo: Deploy simple python webpage
  - Demo: Generate load and HPA monitor

The brief history of Docker & K8S



"Swarm & K8S" We are one big community

docker con 17 EU

The brief history of Docker & K8S

- "We are one big community"

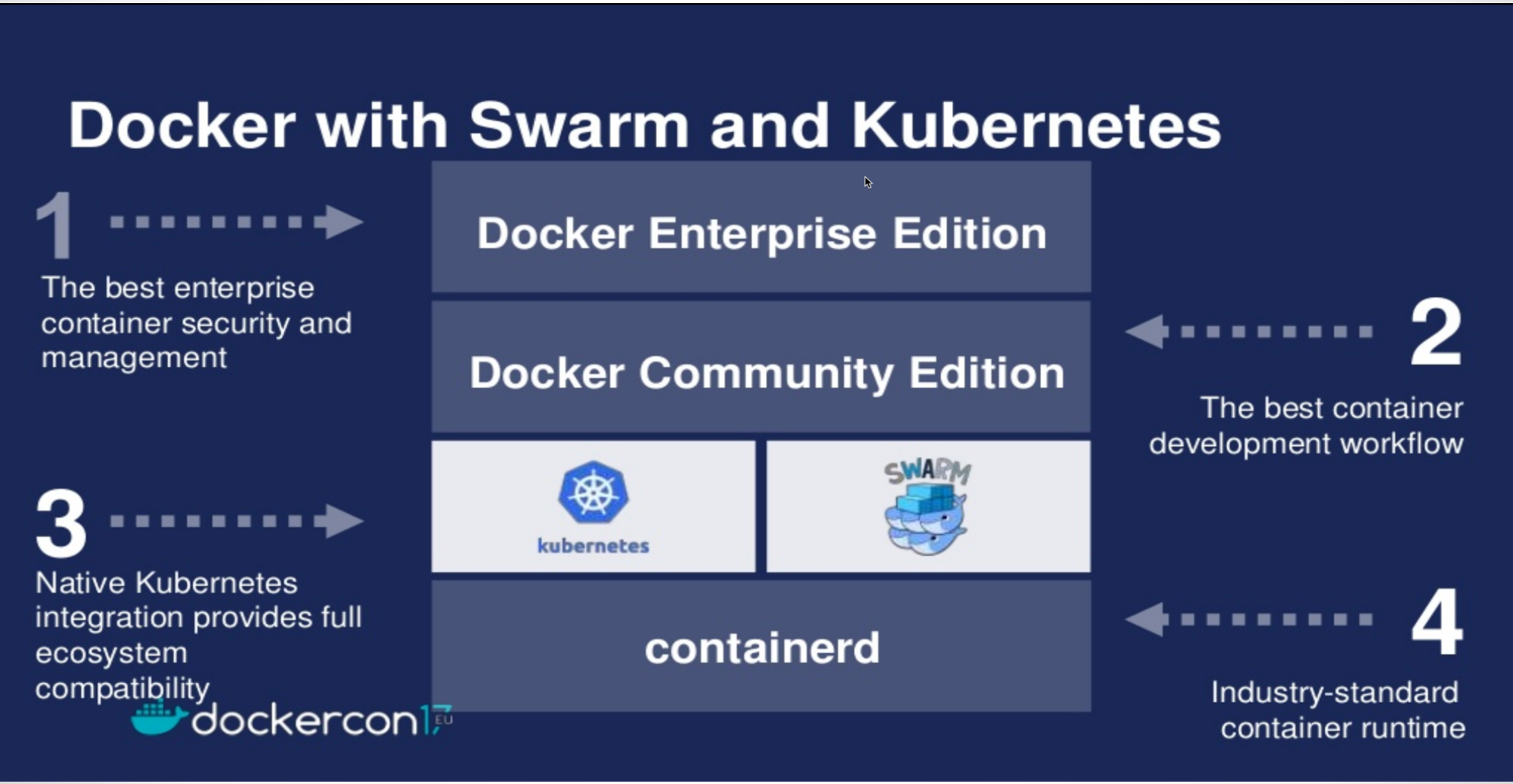
docker con 17 EU

"Swarm & K8S" We are one big community

docker con 17 EU

The brief history of Docker & K8S

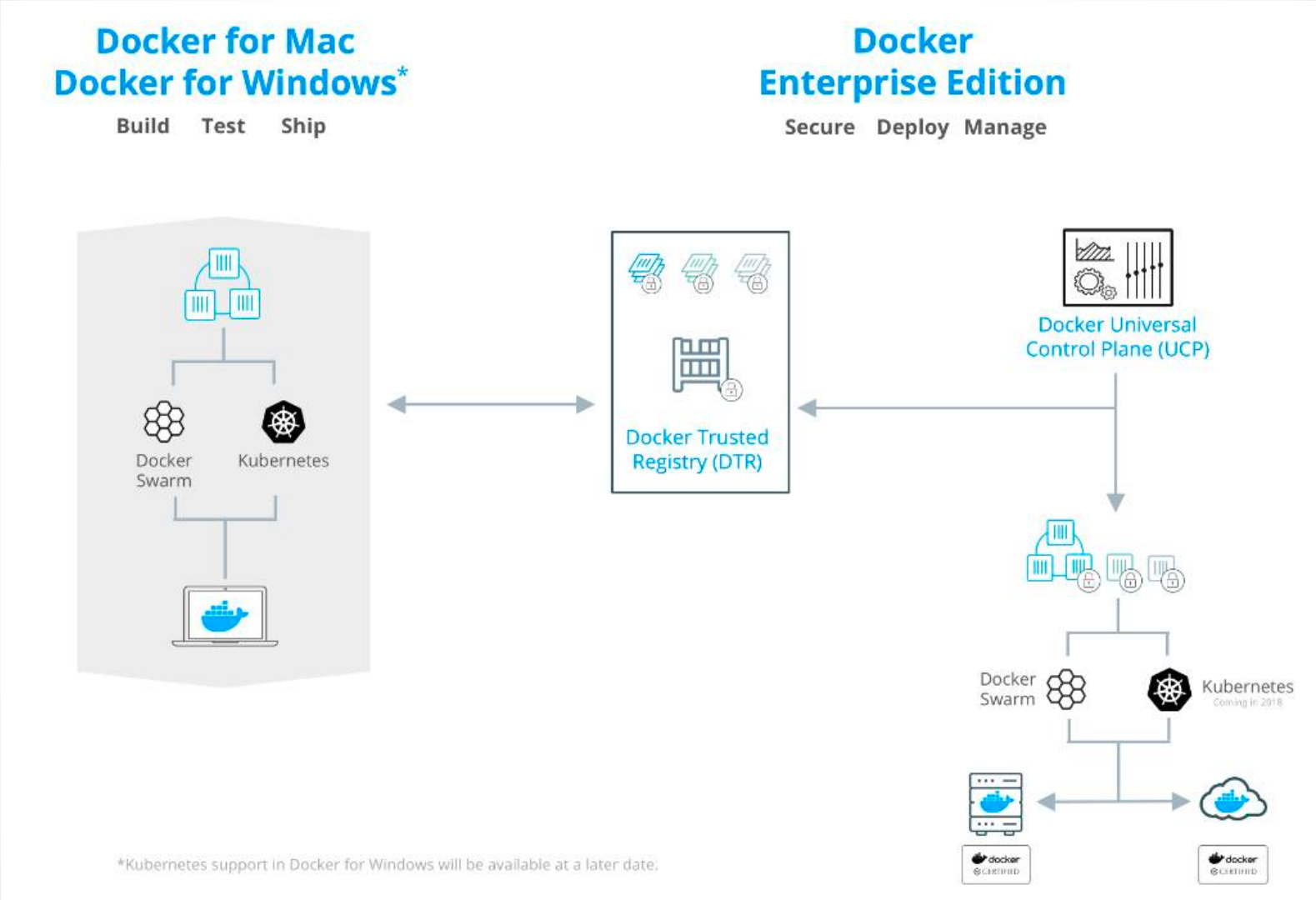
- Official support orchestrator both K8S and Swarm



"Swarm & K8S" We are one big community

dockercon17 EU

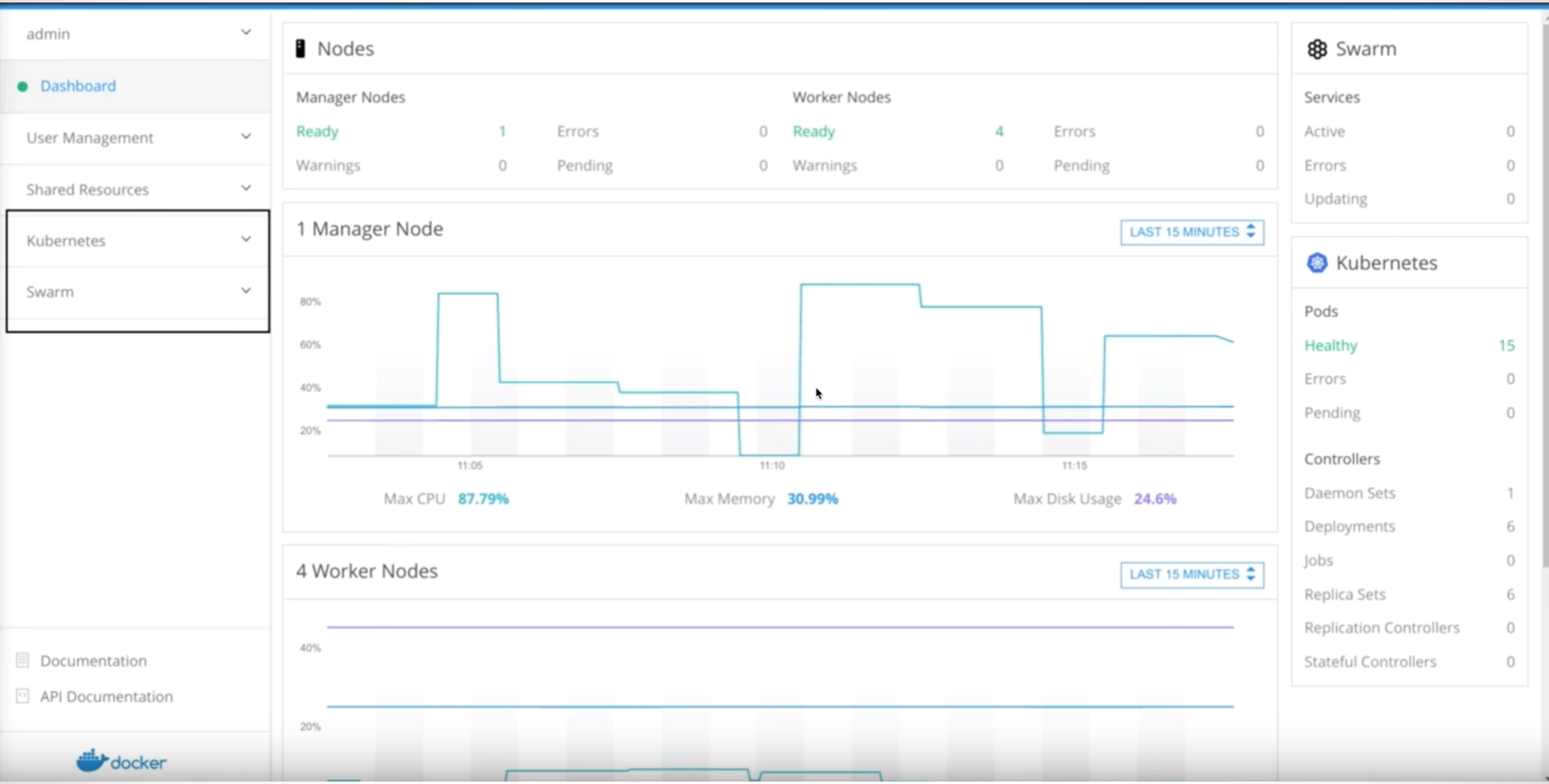
The brief history of Docker & K8S



"Swarm & K8S" We are one big community

dockercon17EU

The brief history of Docker & K8S

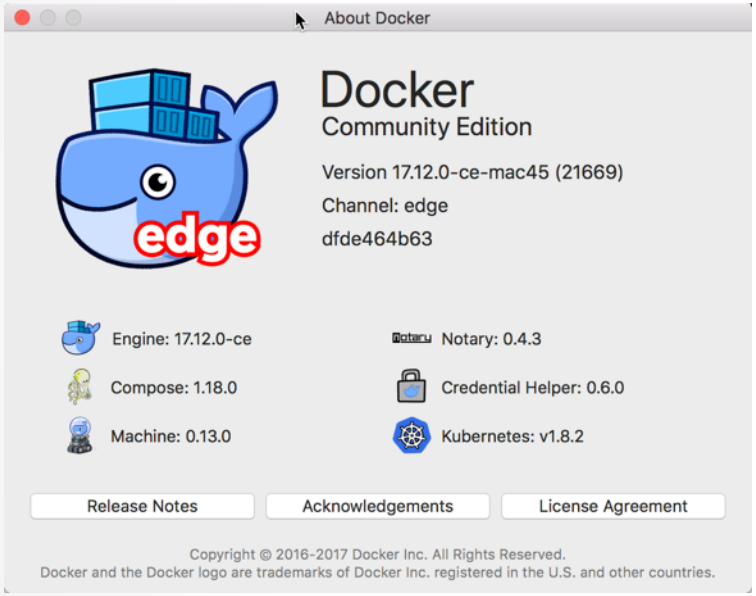


"Swarm & K8S" We are one big community

dockercon17EU

The brief history of Docker & K8S

- Docker Community Edition (CE)
  - Docker for Windows (**Available Now with Edge Edition !!!**)
  - Docker for MAC (**Available Now with Edge Edition !!!**)



"Swarm & K8S" We are one big community

The brief history of Docker & K8S

What are the best Docker orchestration tools?

11	77	Mar 13, 2018
OPTIONS CONSIDERED	RECOMMENDATIONS	LAST UPDATED

THE BEST 1 OF 11 OPTIONS WHY?

11 Options Considered	Price	Last Updated
THE BEST Kubernetes	GET IT HERE	Mar 13, 2018
Docker Swarm	GET IT HERE	Dec 11, 2017
Docker Compose	GET IT HERE	Sep 23, 2017
Nomad	GET IT HERE	Sep 23, 2017
OpenShift	GET IT HERE	Jan 22, 2018

See Full List

RELATED QUESTIONS

- What are the best power user tools for macOS?
- What are the best continuous integration tools?
- What are the best developer tools for Mac OSX?
- What are the best software tools for live streaming?
- What are the best log aggregation & monitoring tools?
- What are the best mockup and wireframing tools for websites?
- What are the best diff tools for Git?
- What are the best mind mapping tools?
- What are the best 2D animation tools for game development?
- What are the best power user tools for Windows?

Reference: https://www.slant.co/topics/3929/~docker-orchestration-tools

"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

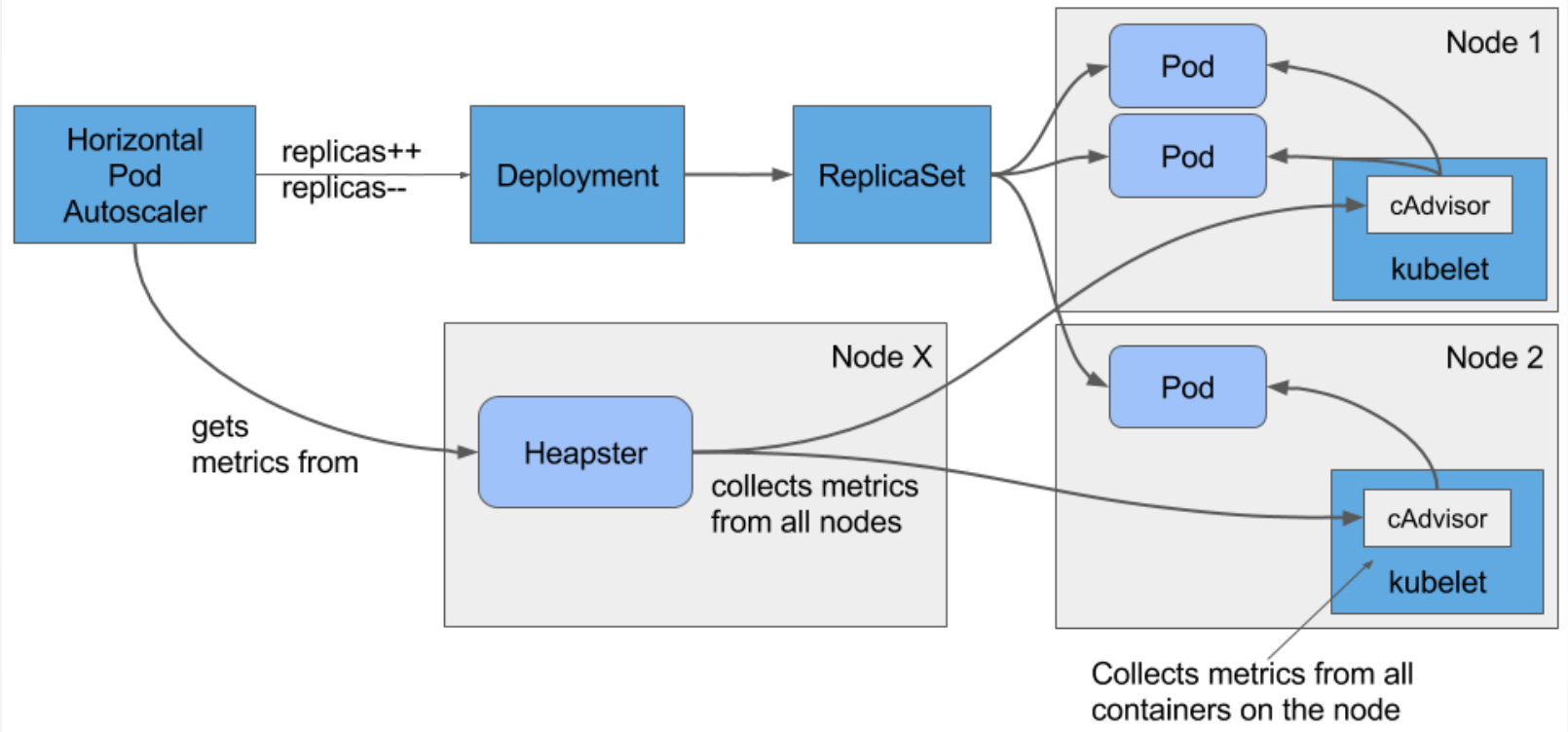
- Normally application (Container inside) was designed base on requirement and performance test result
  - Ex:
    - Spec: 1 Container: 200 Concurrent API Call/Sec
    - Traffic Average: 5000 Concurrent API Call/Sec
    - Total: ~ 25-30 Containers (+ overcap 15%)

- Orchestrator was support scaling as request (Swarm/K8S/Mesos etc)

- But…
  - Why we need to scale bigger farm when empty workload ?
  - Are we confidence that all workload was handled ?
  - How can we handle workload spike ?

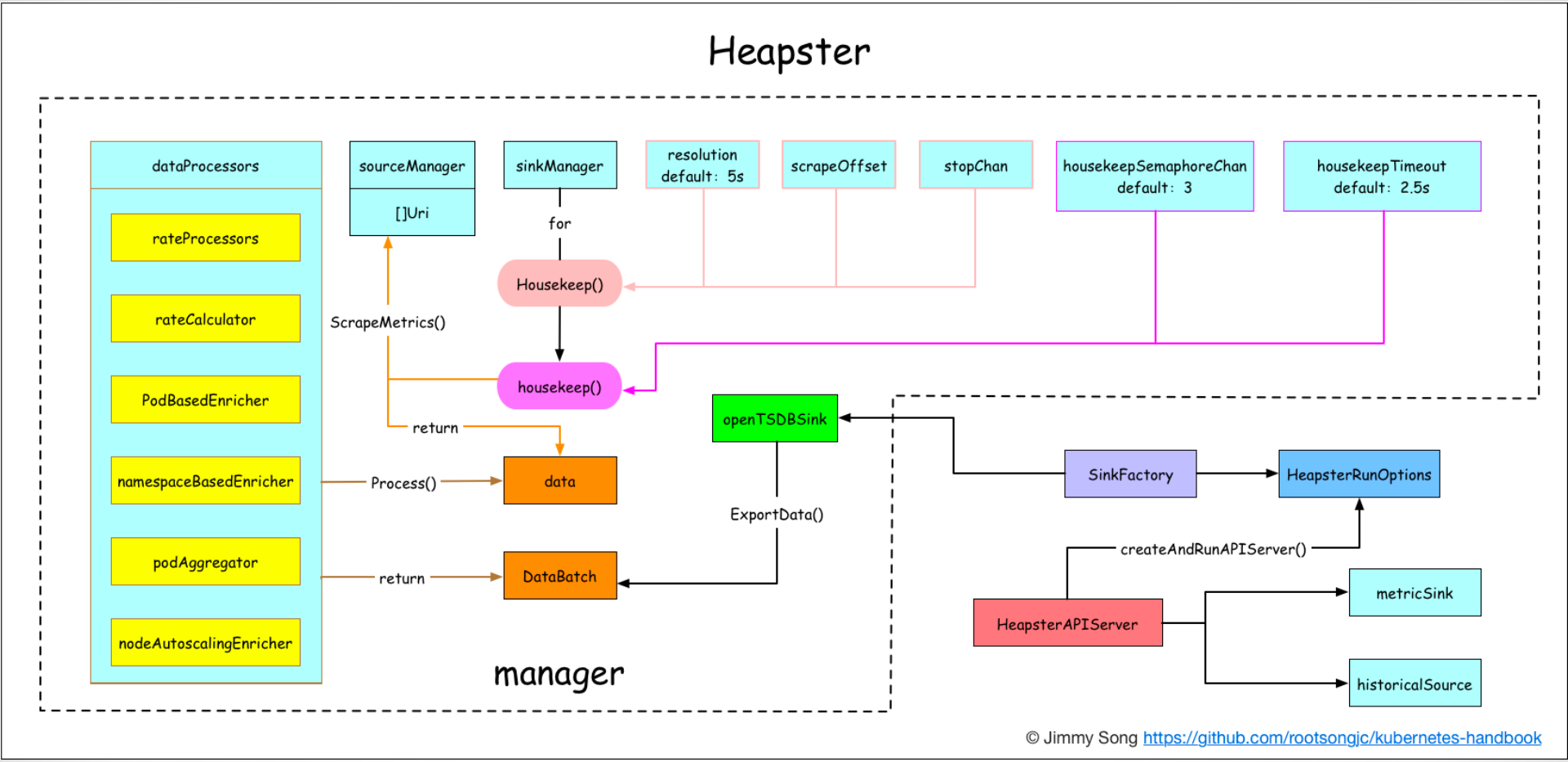
- HPA will response for monitor workload on Pods (Now base on CPU) and automatic trigger deployment to scale-up application

Horizontal Pods Autoscaler



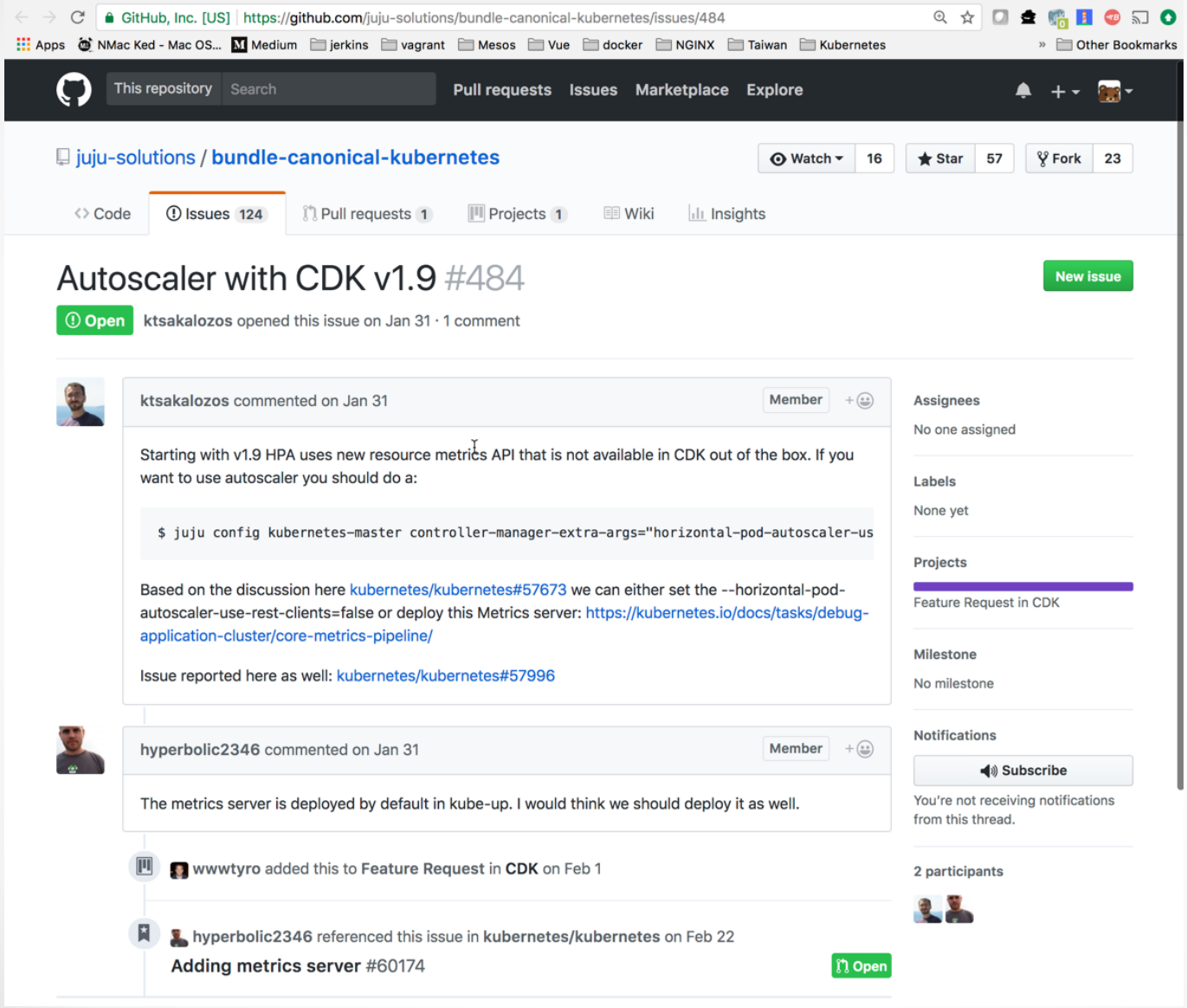
"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler



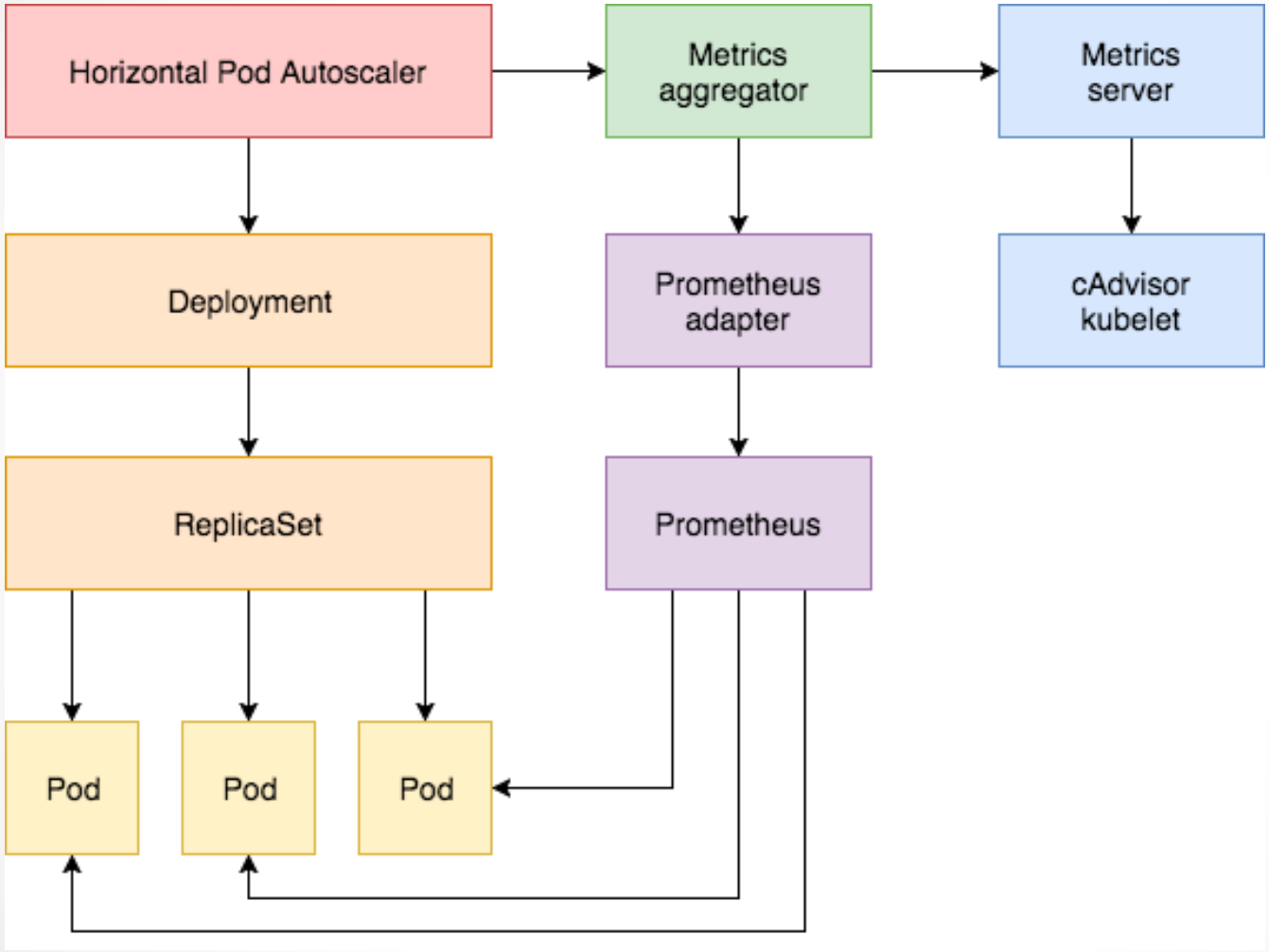
"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler



"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

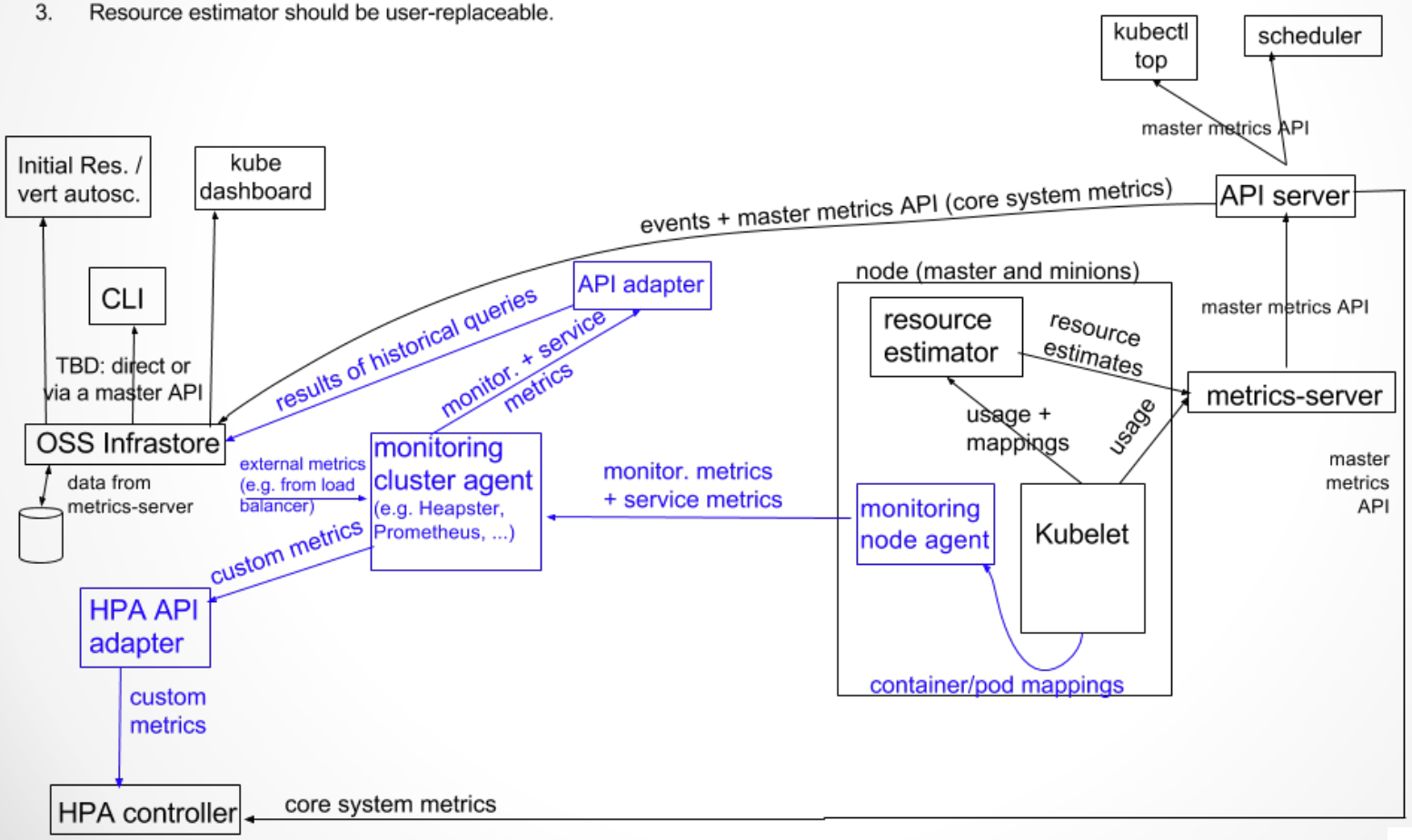


"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

Monitoring architecture proposal: OSS
(arrows show direction of metrics flow)

Notes
1. Arrows show direction of metrics flow.
2. Monitoring pipeline is in blue. It is user-supplied and optional.
3. Resource estimator should be user-replaceable.



"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

- Example: Deployment for python

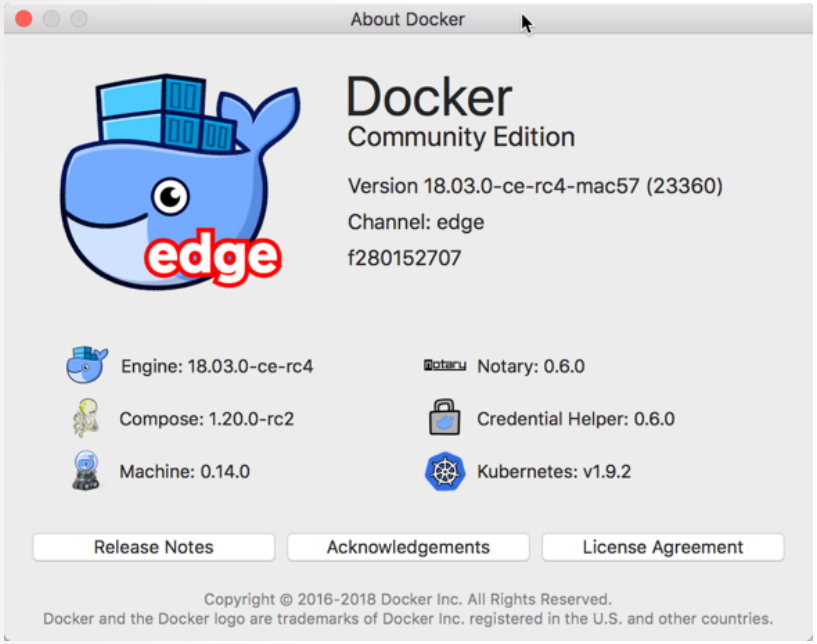
```yaml
apiVersion: v1
kind: Service
metadata:
  name: webtest
  labels:
    name: web
    owner: Praparn_L
    version: "1.0"
    module: WebServer
    environment: development
spec:
  selector:
    name: web
    owner: Praparn_L
    version: "1.0"
    module: WebServer
    environment: development

  type: NodePort
  ports:
   - port: 5000
     name: http
     targetPort: 5000
     protocol: TCP
     nodePort: 32500
```

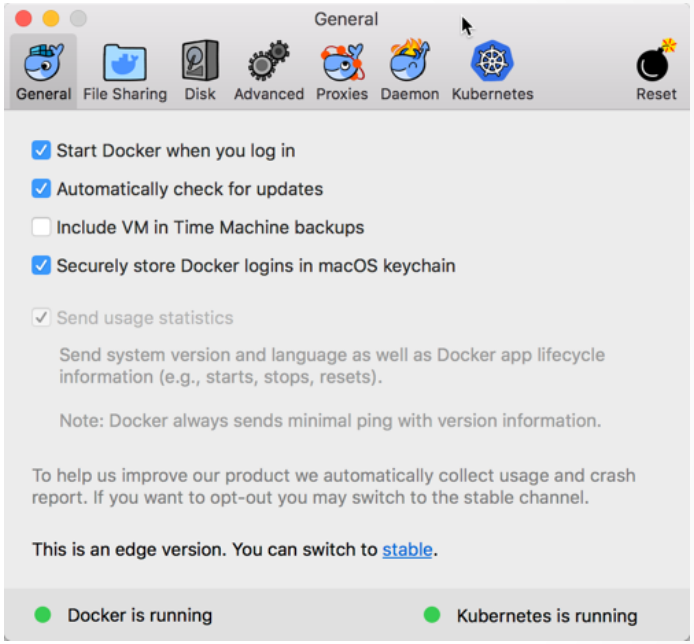
```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: webtest
  labels:
    name: web
    owner: Praparn_L
    version: "1.0"
    module: WebServer
    environment: development
spec:
  replicas: 1
  selector:
    matchLabels:
      name: web
      owner: Praparn_L
      version: "1.0"
      module: WebServer
      environment: development
  template:
    metadata:
      labels:
        name: web
        owner: Praparn_L
        version: "1.0"
        module: WebServer
        environment: development
    spec:
      containers:
        - name: webtest
          image: labdocker/cluster:webservicelite_v1
          resources:
            requests:
              cpu: "200m"
          ports:
          - containerPort: 5000
            protocol: TCP
```

"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

- Example: Deployment for python






```
praparns-MBP% kubectl config use-context docker-for-desktop
Switched to context "docker-for-desktop".
praparns-MBP% kubectl config get-contexts
CURRENT   NAME                 CLUSTER                      AUTHINFO             NAMESPACE
          minikube             minikube                     minikube
*         docker-for-desktop   docker-for-desktop-cluster   docker-for-desktop
          first                first                        first
          local                default-cluster              default-admin
praparns-MBP%
```

"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

- Example: Deployment for python

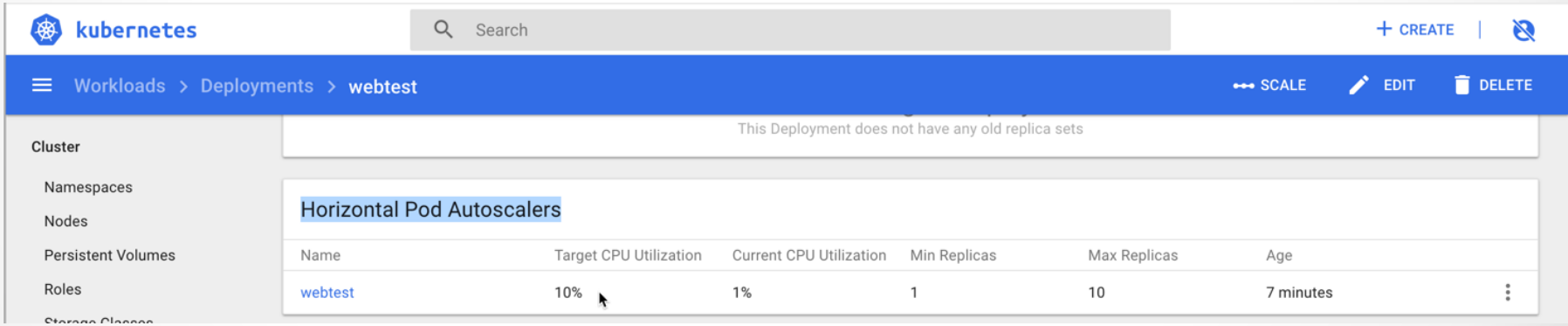
```
praparns-MacBook-Pro% kubectl create -f https://raw.githubusercontent.com/praparn/docker_meetup_20180329/master/webtest_deploy_hpa.yml
service "webtest" created
deployment "webtest" created
praparns-MacBook-Pro% kubectl get deployment/webtest -o wide
        kubectl get svc/webtest -o wide
NAME      DESIRED   CURRENT   UP-TO-DATE   AVAILABLE   AGE         CONTAINERS   IMAGES                                SELECTOR
webtest   1         1         1            1           <invalid>   webtest      labdocker/cluster:webservicelite_v1   environment=developm
ent,module=WebServer,name=web,owner=Praparn_L,version=1.0
NAME      TYPE       CLUSTER-IP      EXTERNAL-IP   PORT(S)          AGE         SELECTOR
webtest   NodePort   10.108.136.22   <none>        5000:32500/TCP   <invalid>   environment=development,module=WebServer,name=web,owner=Pr
aparn_L,version=1.0
praparns-MacBook-Pro%
```

"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

- Example: Deployment for python

```
[praparns-MacBook-Pro% kubectl autoscale deployment/webtest --min=1 --max=10 --cpu-percent=10
deployment "webtest" autoscaled
[praparns-MacBook-Pro% kubectl get hpa
NAME      REFERENCE            TARGETS           MINPODS   MAXPODS   REPLICAS   AGE
webtest   Deployment/webtest   <unknown> / 10%   1         10        0          <invalid>
[praparns-MacBook-Pro% kubectl get hpa
NAME      REFERENCE            TARGETS    MINPODS   MAXPODS   REPLICAS   AGE
webtest   Deployment/webtest   1% / 10%   1         10        1          23s
praparns-MacBook-Pro%
```



"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

- Example: Deployment for python

```
praparns-MacBook-Pro% kubectl run -i --tty load-generator --image=busybox /bin/sh
If you don't see a command prompt, try pressing enter.
/ # wget -q -O- http://webtest.default.svc.cluster.local:5000
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:54:02 2018
/ # while true; sleep 0.01; do wget -q -O- http://webtest.default.svc.cluster.local:5000; done
```

```
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 16:55:30 2018
```

Horizontal Pods Autoscaler

- Example: Deployment for python

```
praparns-MBP:~ praparn$ kubectl get hpa
NAME      REFERENCE            TARGETS    MINPODS   MAXPODS   REPLICAS   AGE
webtest   Deployment/webtest   1% / 10%   1         10        1          10m
praparns-MBP:~ praparn$ kubectl get hpa
NAME      REFERENCE            TARGETS     MINPODS   MAXPODS   REPLICAS   AGE
webtest   Deployment/webtest   27% / 10%   1         10        1          11m
praparns-MBP:~ praparn$ kubectl top pods
NAME                               CPU(cores)   MEMORY(bytes)
load-generator-5c4d59d5dd-m4tsb    43m          1Mi
webtest-7d89786977-tmg76           33m          28Mi
praparns-MBP:~ praparn$ kubectl top nodes
NAME                 CPU(cores)   CPU%      MEMORY(bytes)   MEMORY%
docker-for-desktop   468m         11%       1433Mi          75%
praparns-MBP:~ praparn$
```

```
praparns-MBP:~ praparn$ kubectl top nodes
NAME                 CPU(cores)   CPU%      MEMORY(bytes)   MEMORY%
docker-for-desktop   749m         18%       1508Mi          79%
praparns-MBP:~ praparn$ kubectl get hpa
NAME      REFERENCE            TARGETS     MINPODS   MAXPODS   REPLICAS   AGE
webtest   Deployment/webtest   15% / 10%   1         10        3          14m
praparns-MBP:~ praparn$ kubectl get hpa
NAME      REFERENCE            TARGETS     MINPODS   MAXPODS   REPLICAS   AGE
webtest   Deployment/webtest   15% / 10%   1         10        5          15m
praparns-MBP:~ praparn$ kubectl get hpa
NAME      REFERENCE            TARGETS    MINPODS   MAXPODS   REPLICAS   AGE
webtest   Deployment/webtest   8% / 10%   1         10        5          17m
praparns-MBP:~ praparn$
```

"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

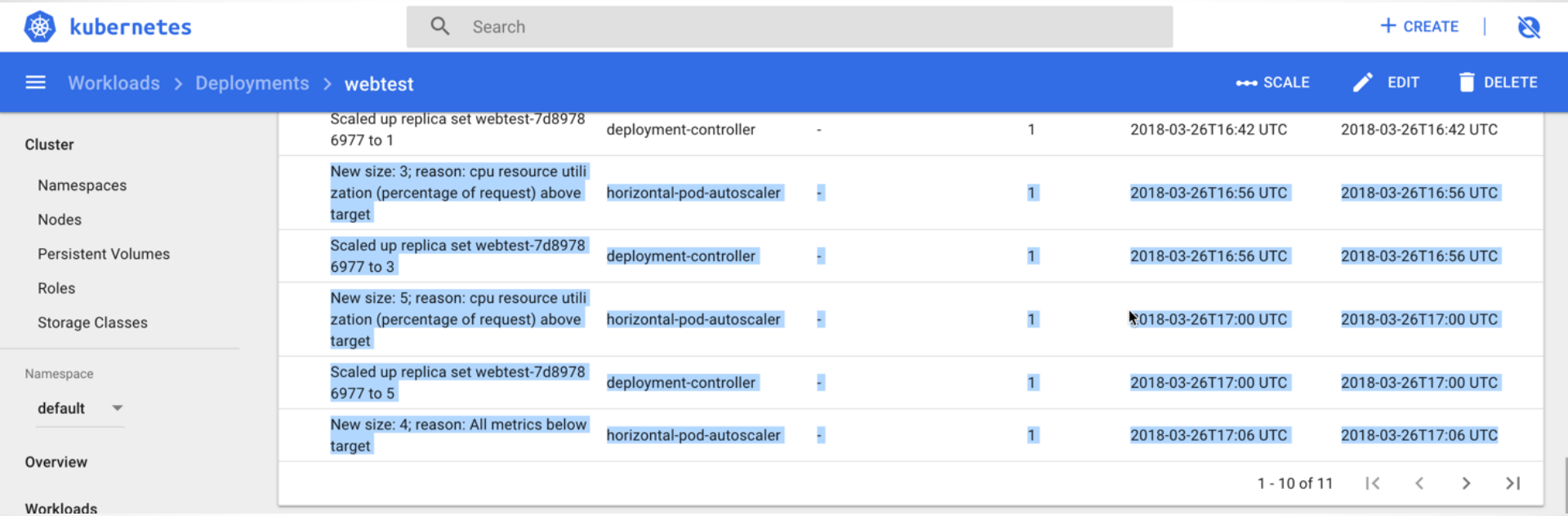
- Example: Deployment for python

```
praparns-MBP:~ praparn$ kubectl describe hpa/webtest
Name:                                                  webtest
Namespace:                                             default
Labels:                                                <none>
Annotations:                                           <none>
CreationTimestamp:                                     Mon, 26 Mar 2018 23:45:03 +0700
Reference:                                             Deployment/webtest
Metrics:                                               ( current / target )
  resource cpu on pods  (as a percentage of request):  8% (17m) / 10%
Min replicas:                                          1
Max replicas:                                          10
Conditions:
  Type            Status  Reason              Message
  ----            ------  ------              -------
  AbleToScale     False   BackoffBoth         the time since the previous scale is still within both the downscale and upscale forbidden w
indows
  ScalingActive   True    ValidMetricFound    the HPA was able to succesfully calculate a replica count from cpu resource utilization (per
centage of request)
  ScalingLimited  False   DesiredWithinRange  the desired count is within the acceptable range
Events:
  Type    Reason             Age   From                       Message
  ----    ------             ----  ----                       -------
  Normal  SuccessfulRescale  6m    horizontal-pod-autoscaler  New size: 3; reason: cpu resource utilization (percentage of request) above
target
  Normal  SuccessfulRescale  2m    horizontal-pod-autoscaler  New size: 5; reason: cpu resource utilization (percentage of request) above
target
praparns-MBP:~ praparn$
```

"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

- Example: Deployment for python

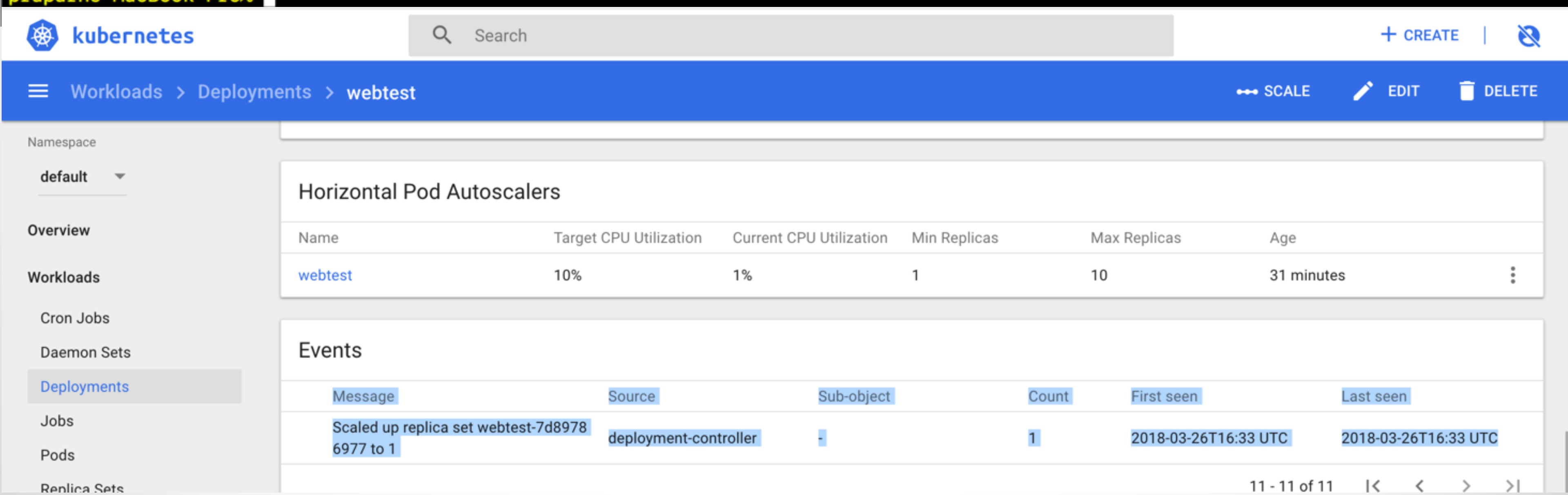


"Swarm & K8S" We are one big community

Horizontal Pods Autoscaler

- Example: Deployment for python

```
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
<H1> Welcome Page from Container Python Lab Web Version 1.00 </H1>Checkpoint Date/Time: Mon Mar 26 17:12:49 2018
^C
/ # exit
Session ended, resume using 'kubectl attach load-generator-5c4d59d5dd-m4tsb -c load-generator -i -t' command when the pod is running
praparns-MacBook-Pro%
```



Demo Session

"Swarm & K8S" We are one big community

dockercon 17 EU

By Praparn Luengphoonlap
Email: eva10409@gmail.com

"Swarm & K8S" We are one big community

dockercon17EU