

Kubernetes

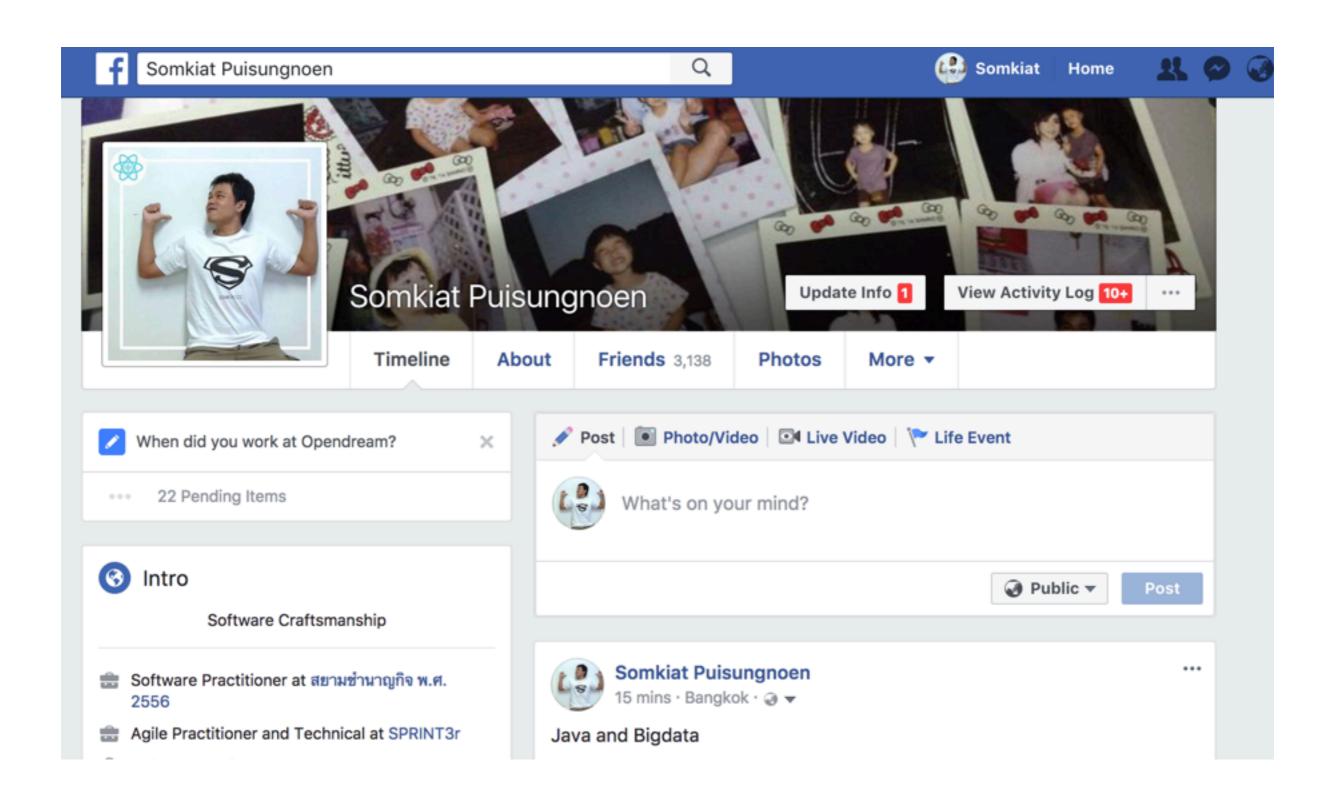
Workshop with Java



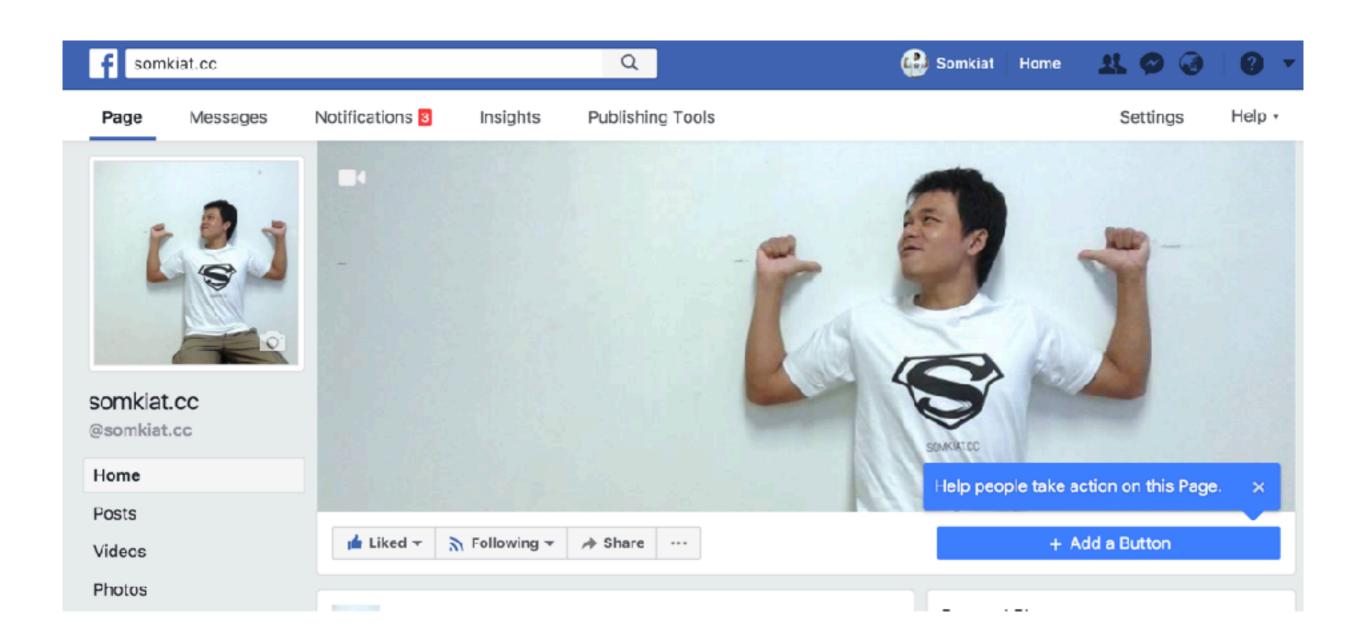














Topics day 1

Kubernetes concepts

- 1. Pods
- 2. Replication Controllers
- 3. Replica Sets
- 4. Deployments
- 5. Services
- 6. Jobs
- 7. Volumes

Deploy the Java web app to Kubernetes



Topics day 2

Advance concepts

- 1. Persistent volume
- 2. Stateful Sets
- 3. Horizontal Pod Autoscaling (HPA)
- 4. Deamon Sets
- 5. Health of Pods
- 6. Namespaces
- 7. Rolling update



https://github.com/up1/course-kubernetes-in-practice



Kubernetes

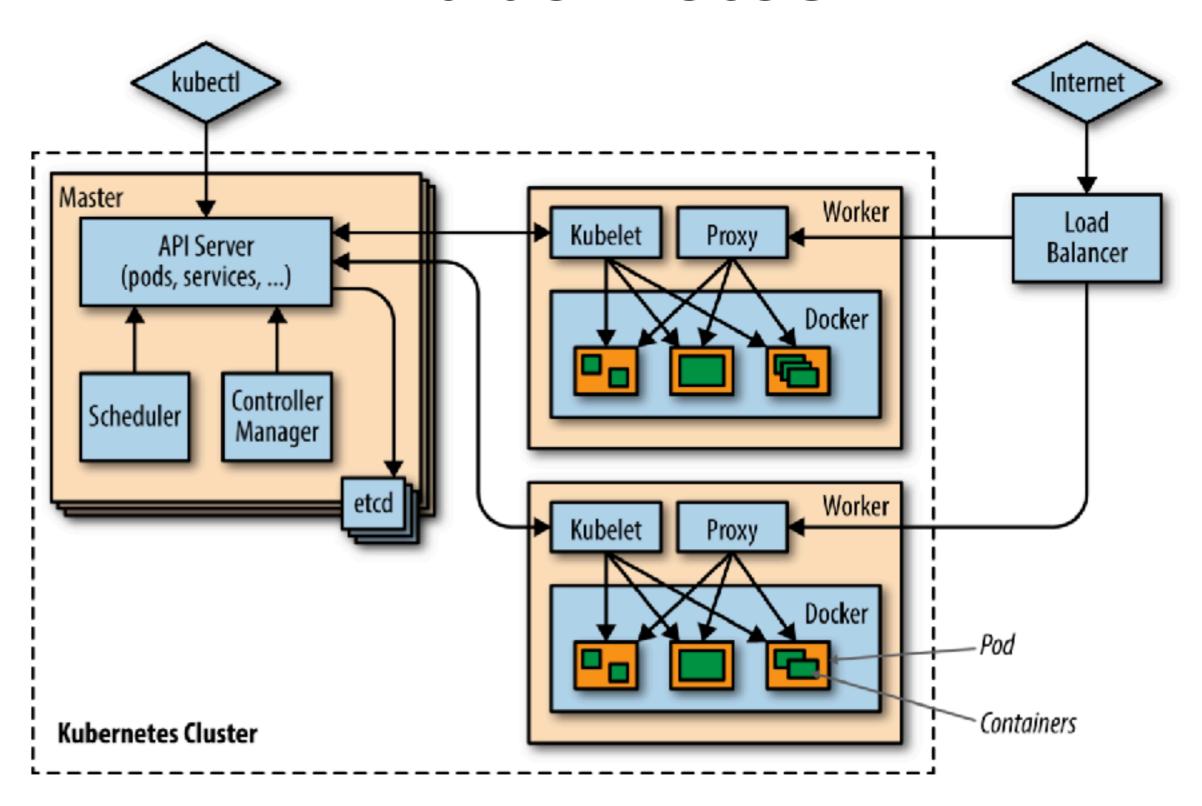
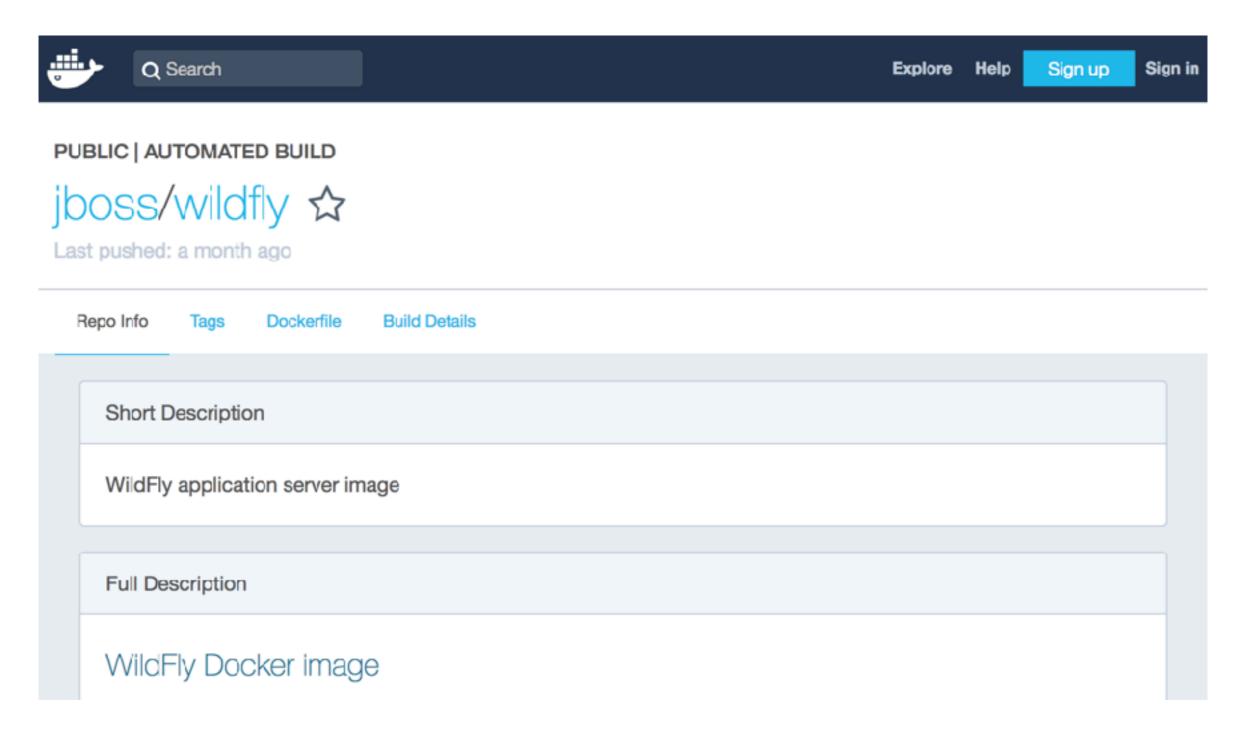




Image jboss/wildfly



https://hub.docker.com/r/jboss/wildfly/



Let's start



1. Pods

Smallest deployable unit
Logical collection of containers
Pods are created in namespaces(ns)
All containers in a pod share the ns, volume and networking



Configuration of Pod

```
apiVersion: v1
kind: Pod
metadata:
  name: wildfly-pod
  labels:
    name: wildfly-pod
spec:
  containers:
  - name: wildfly
    image: jboss/wildfly:10.1.0.Final
    ports:
    - containerPort: 8080
```



2. Replication Controllers (RC)

Ensure that a specified # of pod replicas are running at any time

Automatically replaces if pods fail Support scale-up and scale-down



Configuration of RC

```
apiVersion: v1
kind: ReplicationController
metadata:
  name: wildfly-rc
spec:
  replicas: 2
  selector:
    app: wildfly-rc-pod
  template:
    metadata:
      labels:
        app: wildfly-rc-pod
    spec:
      containers:
      - name: wildfly
        image: jboss/wildfly:10.1.0.Final
        ports:
        - containerPort: 8080
```



3. Replica Sets

Next-generation of RC Selector support



Configuration of Replica Sets

```
apiVersion: extensions/v1beta1
kind: ReplicaSet
metadata:
  name: wildfly-rs
spec:
  replicas: 2
  selector:
    matchLabels:
      app: wildfly-rs-pod
    matchExpressions:
      - {key: tier, operator: In, values: ["backend"]}
      - {key: environment, operator: NotIn, values: ["prod"]}
  template:
    metadata:
      labels:
        app: wildfly-rs-pod
        tier: backend
        environment: dev
    spec:
      containers:
      - name: wildfly
```



4. Deployments

Provide declarative updates for Pods and Replica Sets

Easy to start a RC or Replica set
Easy to check status of deployment
Update deployment to use a new image
Rollback deployment to previous version



Configuration of Deployments

```
apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  name: wildfly-deployment
spec:
  replicas: 3
  template:
    metadata:
      labels:
        app: wildfly
    spec:
      containers:
      - name: wildfly
        image: jboss/wildfly:10.1.0.Final
        ports:
        - containerPort: 8080
```



5. Services

Abstraction to define a logical set of Pods and policy to access them

IP of service doesn't change over time

Pods belong to a service are defined by a label selector



Configuration of Services

```
apiVersion: v1
kind: Service
metadata:
  name: wildfly-service
spec:
  selector:
    app: wildfly-rc-pod
  ports:
    - name: web
      port: 8080
```



6. Volumes

Pods are ephemeral and work well with stateless

Stateful containers require data to be persisted outside

Directory that is accessible to the containers in a Pod



Configuration of Volumes

```
spec:
  containers:
  - name: couchbase
    image: somkiat/couchbase:latest
                      Use Volume on Host Path
    ports:
    - containerPort: 8091
    volumeMounts:
    - mountPath: /var/couchbase/lib
      name: couchbase-data
  volumes:
   name: couchbase-data
    hostPath:
      path: /opt/data
```



Workshop#0

File /working-with-java/01-basic



Start Wildfly server

\$kubectl run hello-wildfly --image=jboss/wildfly: 12.0.0.Final —port=8080

\$kubectl get pod, deployment

NAME hello-wildfly-7bd4cfc58-jw9xp			READY 0/1		STATUS ContainerCreating			RESTARTS 0	AGE 3m
NAME hello-wildfly	DESIRED 1	CURREN 1	⊤ U 1	JP-TO-DAT L	Έ	AVAILAE 0	BLE	AGE 3m	
NAME hello-wildfly-7	bd4cfc58	DESIRE	D C	CURRENT	RE/ Ø	ADY	AGE 3m		



Start Wildfly server

\$kubectl proxy



Welcome to WildFly 11

Your WildFly 11 is running.

Documentation | Quickstarts | Administration Console

WildFly Project | User Forum | Report an Issue

JBess Community

To replace this page simply deploy your own war with / as its context path.

To disable it, remove the "welcome-content" handler for location / in the undertow subsystem.



Try with configuration file

\$kubectl create -f wildfly-pod.yaml \$kubectl create -f wildfly-rc.yaml \$kubectl create -f wildfly-service.yaml



Workshop #1 with Spring boot service

File /working-with-java/02-web-app



Structure of project





1. Build Spring Boot Service

Create executable JAR file

\$./gradlew bootRepackage

Create Docker image

\$./gradlew buildDocker



2. Push Docker image to Docker Hub

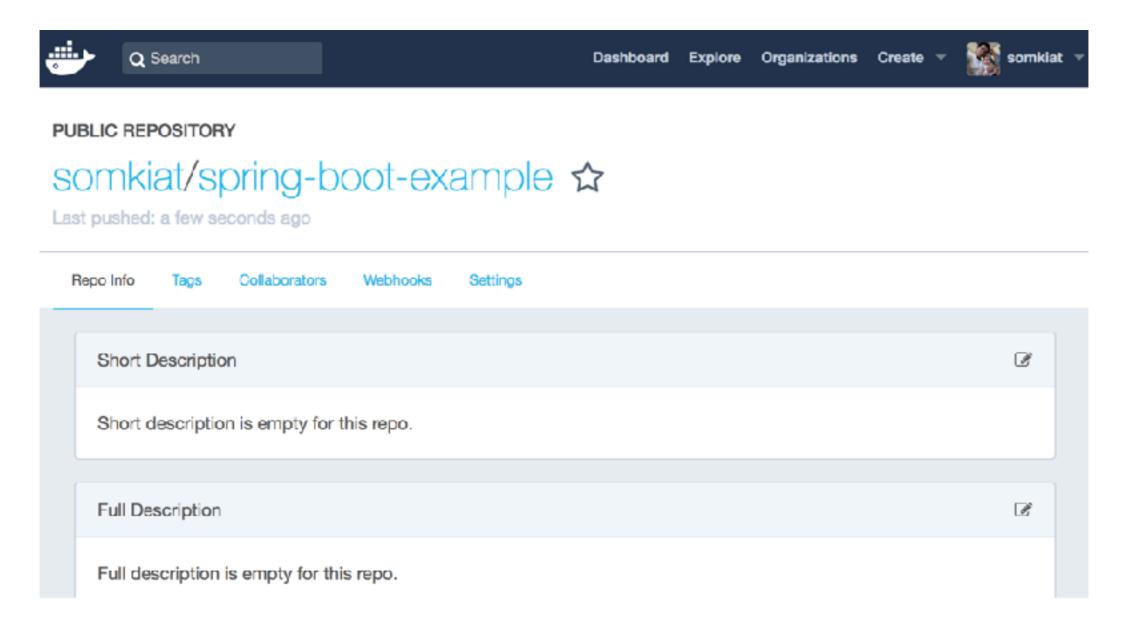
\$docker login \$docker push <username>/spring-boot-example

Required:: register at https://hub.docker.com/



2. Push Docker image to Docker Hub

Check your image



https://hub.docker.com/r/somkiat/spring-boot-example/



3. Deploy application

- 1. MongoDB
- 2. Spring boot service



3.1 Deploy MongoDB

\$kubectl create -f mongo-controller.yaml \$kubectl create -f mongo-service.yaml



3.1 Mongo-controller.yaml

containers:

- image: mongo
name: mongo
ports:

- name: mongo
 containerPort: 27017
 hostPort: 27017

volumeMounts:

Use Volume on Host Path

- name: mongo-persistent-storage
mountPath: /data/db

volumes:

- name: mongo-persistent-storage
hostPath:

path: /data/storage/mongodb



3.1 Mongo-controller.yaml

containers: - image: mongo name: mongo ports:

- name: mongo
containerPort: 27017

hostPort: 27017 Mount volume from container

volumeMounts:

- name: mongo-persistent-storage
mountPath: /data/db

volumes:

- name: mongo-persistent-storage
hostPath:
 path: /data/storage/mongodb



3.2 Deploy Spring boot service

\$kubectl create -f boot-deployment.yaml \$kubectl create -f boot-service.yaml



3.2 Boot-deployment.yaml

spec:

containers:

- name: spring-boot-service

image: somkiat/spring-boot-example

ports: Use environment variables

- containerPort: 8080

env:

name: MONGO_URI

value: mongo-service



4. Testing

Try to add new data and see result in browser

```
① 192.168.99.100:30265/user
 id: "5b043054cff47e0005b9f94e",
 firstName: "Somkiat",
 lastName: "Puisungnoen",
 email: "xxx@gmail.com",
 address: null
```

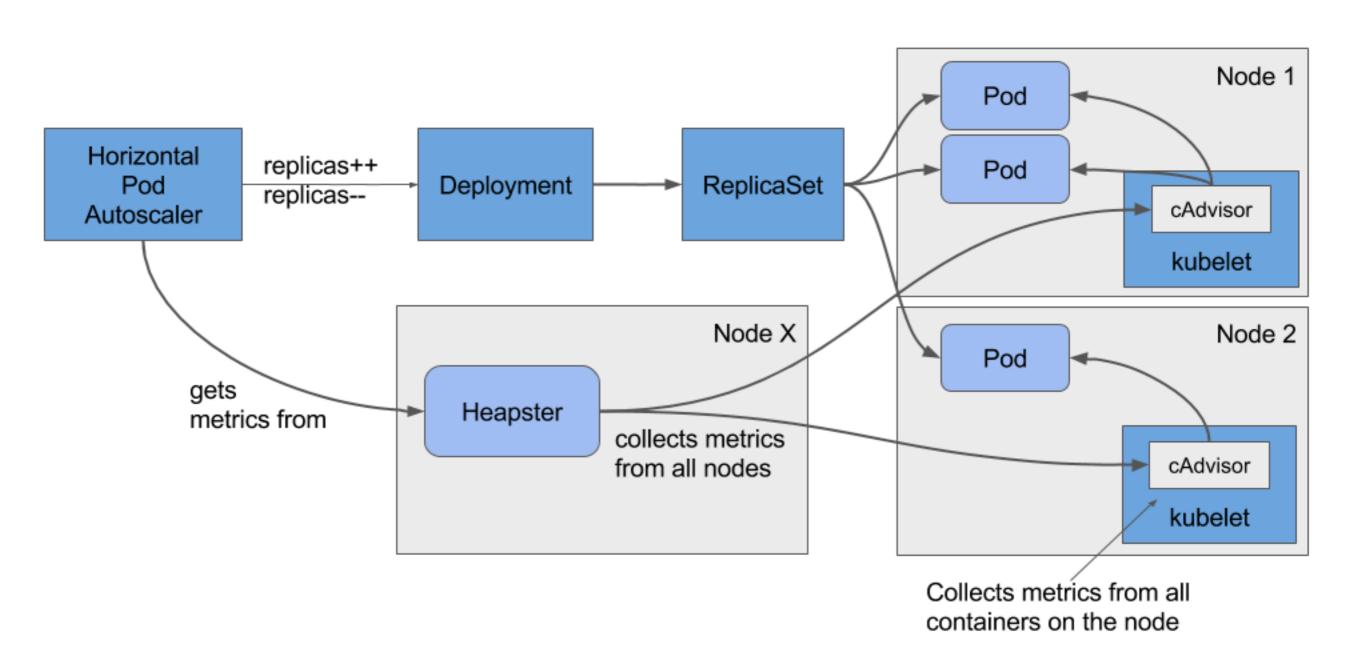


Workshop #2 Horizontal Pods Autoscale (HPA)

File /working-with-java/03-hpa



Components in Autoscaling



https://kubernetes.io/docs/tasks/run-application/horizontal-pod-autoscale-walkthrough/



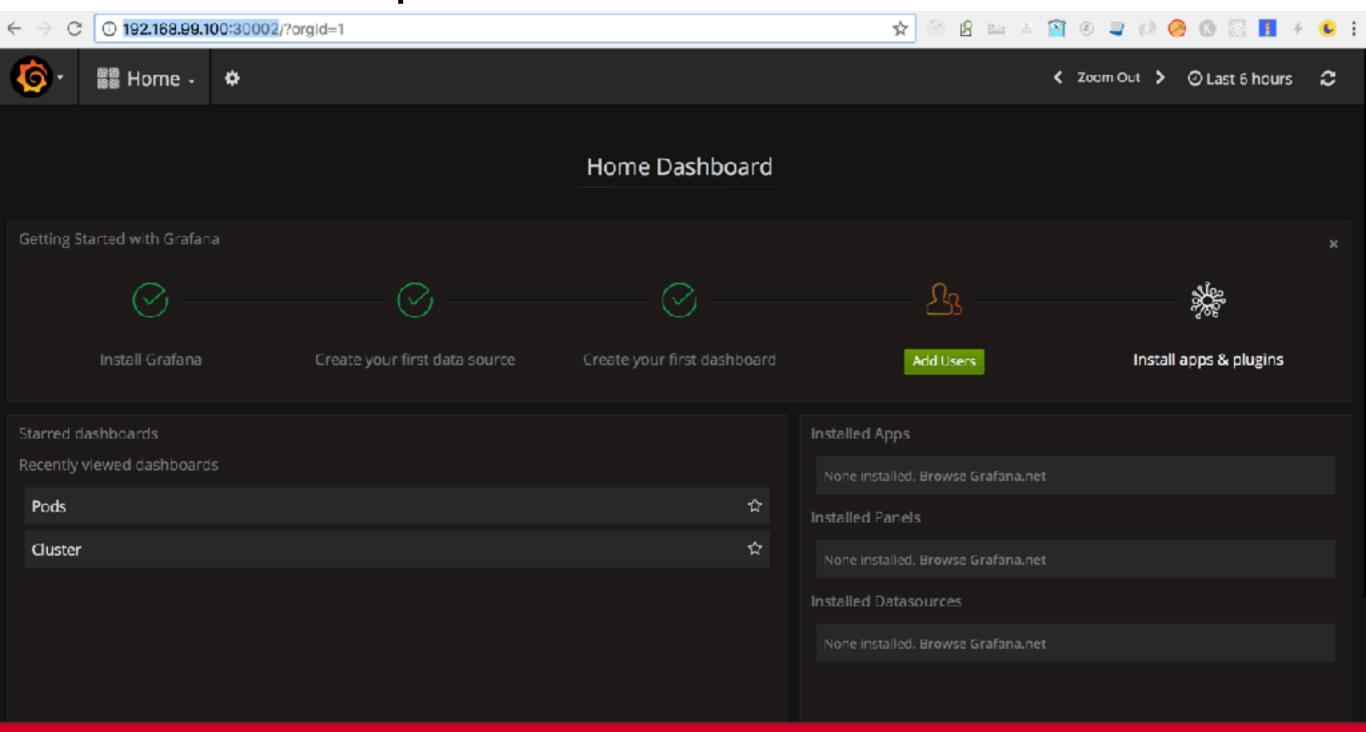
1. Enable hipster in minikube

\$minikube addons enable heapster \$minikube addons start heapster \$minikube addons open heapster



See result in Grafana

http://192.168.99.100:30002





2. Create metric server

Starting from Kubernetes 1.8, resource usage metrics, such as container CPU and memory usage, are available in Kubernetes through the Metrics API. These metrics can be either accessed directly by user, for example by using kubectl top command, or used by a controller in the cluster, e.g. Horizontal Pod Autoscaler, to make decisions.

https://github.com/kubernetes-incubator/metrics-server



3. Deploy Springboot service

\$kubectl create -f boot-deployment.yaml \$kubectl create -f boot-service.yaml



4. Create HPA in command line

\$kubectl autoscale deployment/spring-boot-service-deployment --min=1 --max=5 --cpu-percent=5

```
min = minimum of replica
max = maximum of replica
cpu-percent = average of % of CPU usage
```



4. Create HPA in command line

\$kubectl get hpa

REFERENCE

Deployment/spring-boot-service-deployment

TARGETS M

MINPODS MAXPODS

0%/5%

1



4. Create HPA in command line

Scale-up can only happen if there was no rescaling within the last 3 minutes

Scale-down will wait for 5 minutes from the last rescaling

https://github.com/kubernetes/community/blob/master/contributors/design-proposals/autoscaling/horizontal-pod-autoscaler.md#autoscaling-algorithm



5. Load testing with command line



\$kubectl get pods

NAME	READY	STATUS	RESTARTS	AGE
load-test-6f86656986-9qvdl	0/1	Error	0	4 m
mongo-controller-xdh28	1/1	Running	0	1h
spring-boot-service-deployment-86c7b47b7b-9l9sn	1/1	Running	0	15m
spring-boot-service-deployment-86c7b47b7b-9srkg	1/1	Running	0	1 m
spring-boot-service-deployment-86c7b47b7b-kbtrl	1/1	Running	0	1 m
spring-boot-service-deployment-86c7b47b7b-pqbsm	1/1	Running	0	1 m



\$kubectl describe hpa

```
( current / target )
Metrics:
 resource cpu on pods (as a percentage of request):
                                                       134% (269m) / 5%
Min replicas:
                                                       5
Max replicas:
Conditions:
 Type
                  Status
                          Reason
                                            Message
 AbleToScale
                  False
                          BackoffBoth
                                            the time since the previous scale is still within both the downscale and upscale fo
rbidden windows
                                            the HPA was able to successfully calculate a replica count from cpu resource utiliza
                          ValidMetricFound
 ScalingActive
                  True
tion (percentage of request)
 ScalingLimited True
                         TooManyReplicas
                                            the desired replica count is more than the maximum replica count
           Reason
                                               From
                                                                          Message
 Type
                                         Age
 Normal
           SuccessfulRescale
                                               horizontal-pod-autoscaler New size: 4; reason: cpu resource utilization (percen
                                         1m
tage of request) above target
                                               horizontal-pod-autoscaler unable to get metrics for resource cpu: unable to fet
 Warning FailedGetResourceMetric
                                         19s
ch metrics from API: the server could not find the requested resource (get pods.metrics.k8s.io)
                                               horizontal-pod-autoscaler Operation cannot be fulfilled on horizontalpodautosca
 Warning FailedUpdateStatus
lers.autoscaling "spring-boot-service-deployment": the object has been modified; please apply your changes to the latest versio
n and try again
 Warning FailedComputeMetricsReplicas 19s
                                               horizontal-pod-autoscaler failed to get cpu utilization: unable to get metrics
for resource cpu: unable to fetch metrics from API: the server could not find the requested resource (get pods.metrics.k8s.io)
```

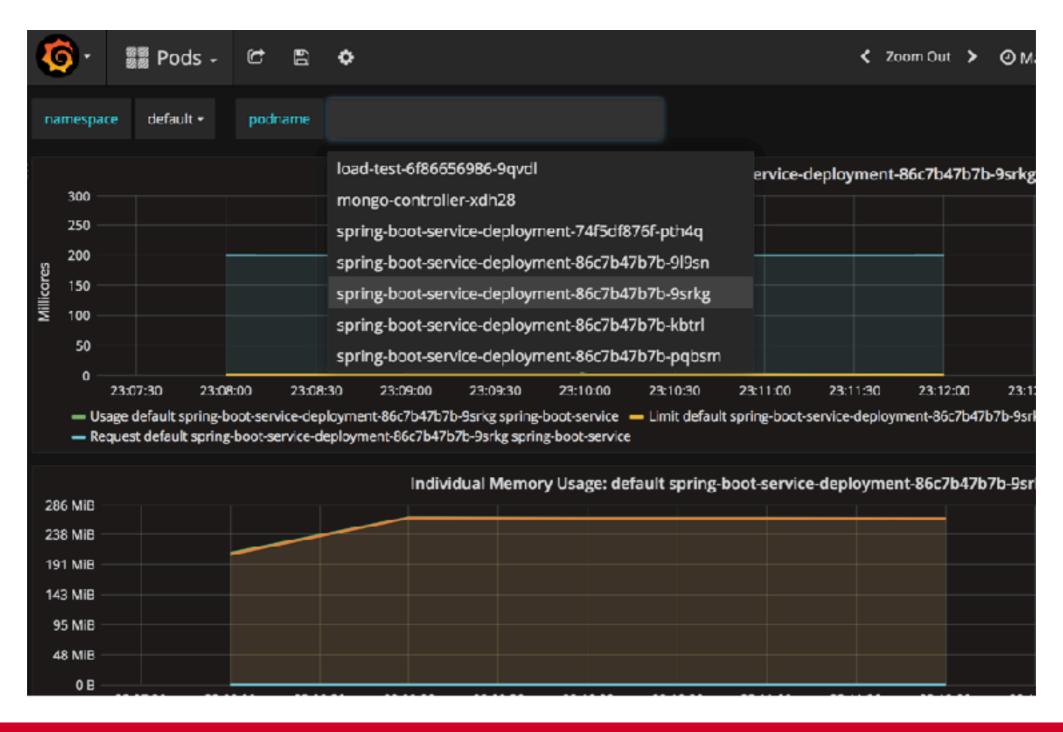


In dashboard of Pods

Pods									÷
	Name \$	Node	Status \$	Restarts	Age 🕏	CPU (cores)	Memory (bytes)		
⊘	spring-boot-service-de	minikube	Running	0	2 minutes	0.249	265.926 Mi	=	:
⊘	spring-boot-service-de	minikube	Running	0	2 minutes	0.03	254.156 Mi	=	:
⊘	spring-boot-service-de	minikube	Running	0	2 minutes	0.001	94.840 Mi	=	:
Ø	load-test-6f86656986-	minikube	Running	1	5 minutes	0	688 Ki	≡	:
⊘	spring-boot-service-de	minikube	Running	0	16 minutes	0.018	33.211 Mi	=	:
Ø	mongo-controller-xdh2	minikube	Running	0	an hour	0.004	24.453 Mi	≡	:



In Grafana dashboard





Workshop #3 Secret and ConfigMap

File /working-with-java/04-secret-configmap

