Horizontal Pod Autoscaling

( you need to download metric server )

* Wget -o metricserver.yml <https://github.com/kubernetes-sigs/metricserver/releases/latest/download/components.yml>
* Ls ( you will get see metric file )
* Vi metricserver.yml ( edit this file eg:- --kubelet-insecure-tls )
* Kubelet apply -f metricserver.yml
* Kubectl get pods -n kube-system ( now you can see metric-server pod )
* Kubectl logs -f metrics-server-4834837fdf -n kube-system ( to see detail of metric server )
* Ctrl+c
* Vi deployhpa.yml

( eg:- kind: Deployment

apiVersion: apps/v1

metadata:

name: mydeploy

spec:

replicas: 1

selector:

matchLabels:

name: deployment

template:

metadata:

name: testpod8

labels:

name: deployment

spec:

containers:

- name: c00

image: httpd

ports:

- containerPort: 80

resources:

limits:

cpu: 500m

requests:

cpu: 200m )

* Kubectl apply -f deployhpa.yml
* Kubectl get all ( you will see only pod and container which you define in yml file )
* kubectl autoscale deployment mydeploy --cpu-percent=20 --min=1 --max=10 (configuring auto scaling )
* kubectl get all ( now you can see horizontal auto-scaling is added )

(open one duplicate terminal with same public ip)

* kubectl get pods
* kubectl exec mydeploy-dcdjfkw82r24 -it -- /bin/bash
* apt update -y

( after updating in previous screen when load increase pod will be automatically increase )