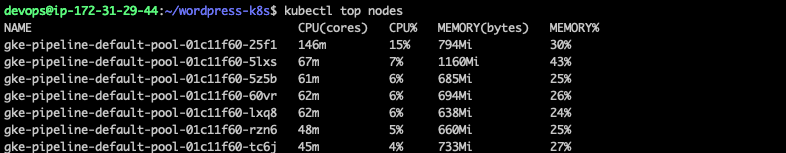
Replace **<your-name>** with your **name** throughout the Lab.

**1. SSH to the AWS Workstation**

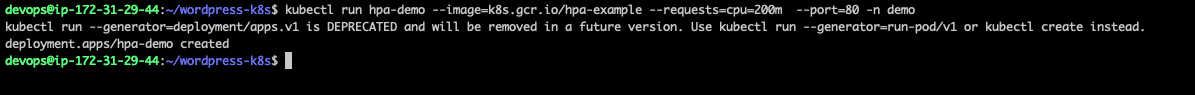
|  |
| --- |
| $ mkdir /home/devops/hpa  $ cd /home/devops/hpa $ kubectl top nodes |

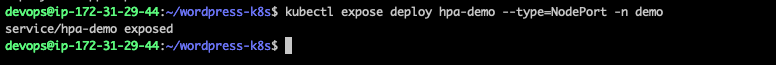


**2. Run & expose the Application**

Create a new deployment with the below command.

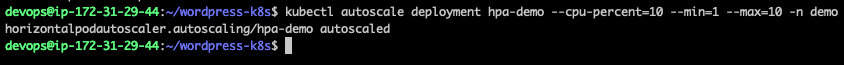
|  |
| --- |
| $ kubectl run hpa-<your-name> --image=k8s.gcr.io/hpa-example --requests=cpu=200m --port=80 -n <your-name>  $ kubectl expose deploy hpa-<your-name> --type=NodePort -n <your-name> |





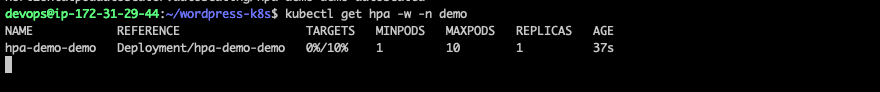
## **3. Create Horizontal Pod Autoscaler**

|  |
| --- |
| $ kubectl autoscale deployment hpa-<your-name> --cpu-percent=10 --min=1 --max=5 -n <your-name> |



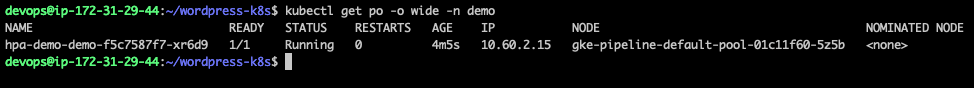
**4. Please wait for 2-3 minutes before running the below command**

|  |
| --- |
| $ kubectl get hpa -w -n <your-name> |

****Press **Ctrl+C** to exit.

**5. Check the NODE where the HPA App has been deployed.**

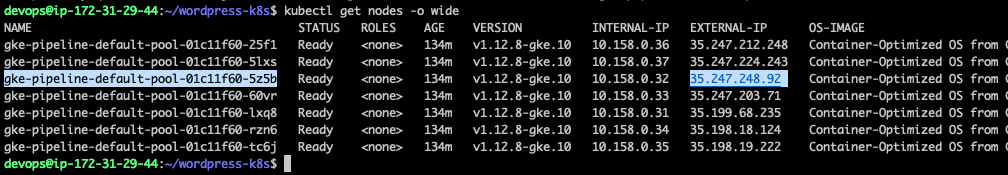
|  |
| --- |
| $ kubectl get po -o wide -n <you-name> |



In this demo the POD is running on **gke-pipeline-default-pool-01c11f60-5z5b NODE.**

**6. Run the below command to get the PUBLIC IP of the NODE where the POD is running.**

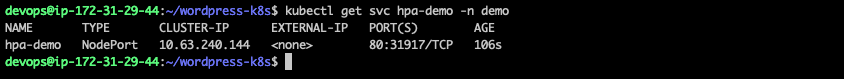
|  |
| --- |
| $ kubectl get nodes -o wide |



Public IP for **is 35.247.248.92**

**7. Check the NODEPORT on which the service is exposed.**

|  |
| --- |
| $ kubectl get svc hpa-<your-name> -n <your-name> |

****

**In this Example the hpa app is exposed is exposed on NODEPORT 31917**

**8. Run the below command to increase the Load.**

|  |
| --- |
| nohup bash -c "while true; do wget -q -O- http://<NODE-Public-IP>:<NodePort>/ ; done &" |

**Where, NODE-Public-IP is the Public IP of the Node where the hpa-demo POD is deployed and NodePort is the Port on which the hpa-demo POD is exposed to, in this Example 35.247.248.92 is the Public IP of the NODE and 31917 is the NodePort**

Example

|  |
| --- |
| nohup bash -c "while : ; do wget -q -O- http://35.247.248.92:31917/ ; done &" |

**9. Run the below command on the second terminal.**

|  |
| --- |
| $ watch -n 1 kubectl get hpa hpa-<your-name> -n <your-name> |

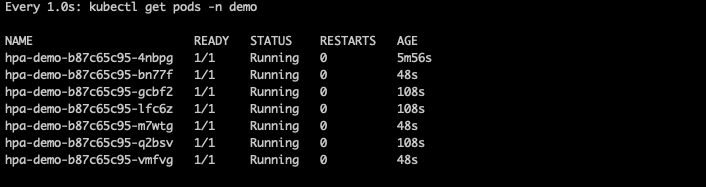
Press CTRL+C to exit the watch.



**10. Run the below command to check the PODs status**

|  |
| --- |
| $ watch -n 1 kubectl get pods -n <your-name> |

We can observe that the PODS have been horizontally scaled up due to the increased loadl, as shown in the below screenshot.



Press CTRL+C to exit the watch.

**Cleanup the HPA deployments after completing LAB-9**

11. Run the below commands to delete the Deployments and HPA

|  |
| --- |
| $ kubectl delete deploy hpa-demo-<your-name> -n <your-name> $ kubectl delete hpa hpa-demo-<your-name> -n <your-name> |