Kubernetes with Containers and DevOps Workshop

Hands-on lab step-by-step

Aralık 2018

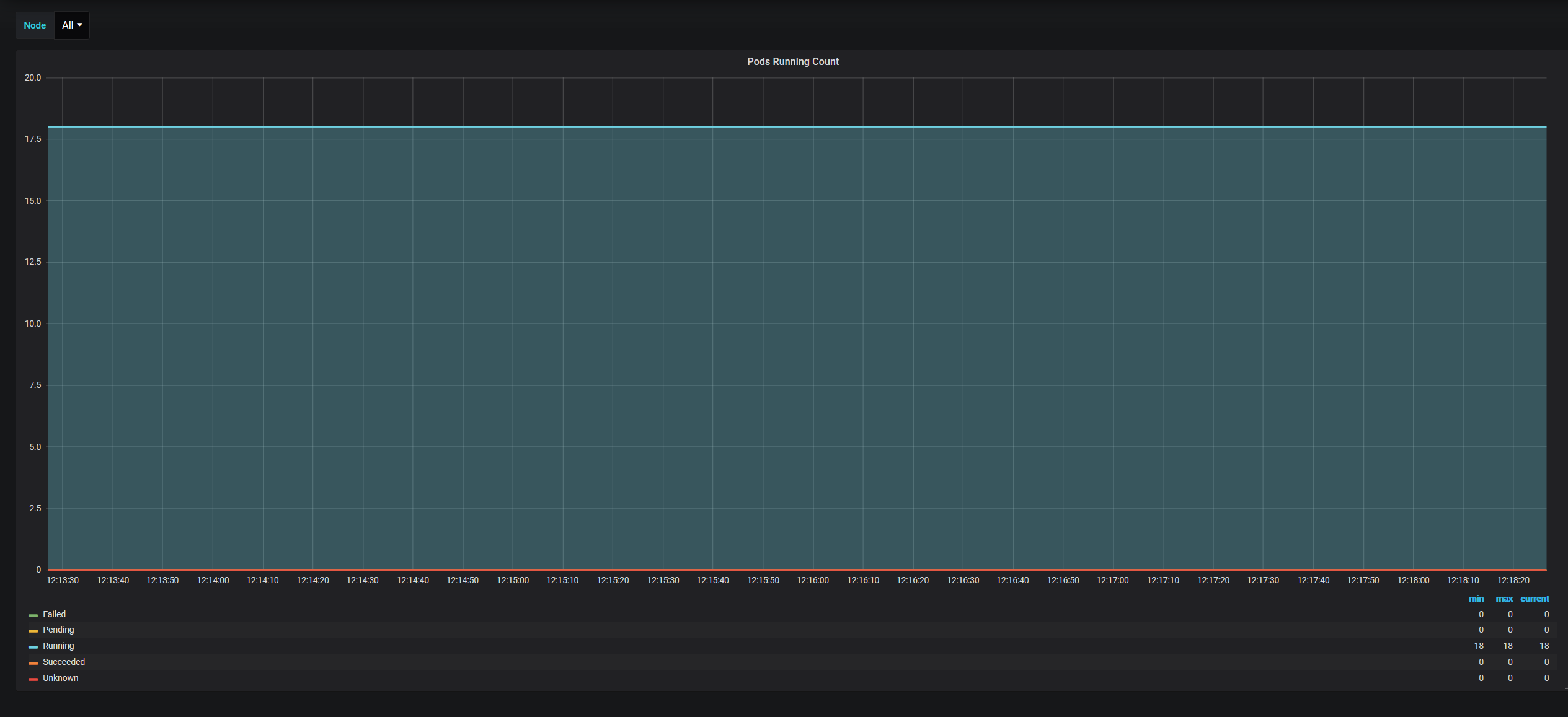
8.Application and Infrastructure Scaling

Imagine a scenario where your realize that your existing cluster is at capacity and you need to scale it out to add more nodes in order to increase capacity and be able to deploy more PODS.

**Exercise 1 - Scale Application**

1. Check to see current number of pods running via Grafana Dashboard.

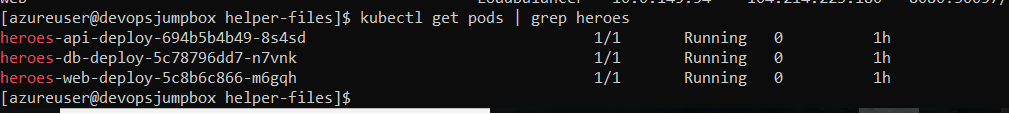
* Go to the same Grafana Dashboard from exercise 8 and look at the **Pods Running Count** section. You will see the total count of Pods and the various phases they are in.



1. Check to see current number of heroes pods running via K8s CLI.

kubectl get pods | grep heroes

# You should see something like the following as output (one replica of each pod):

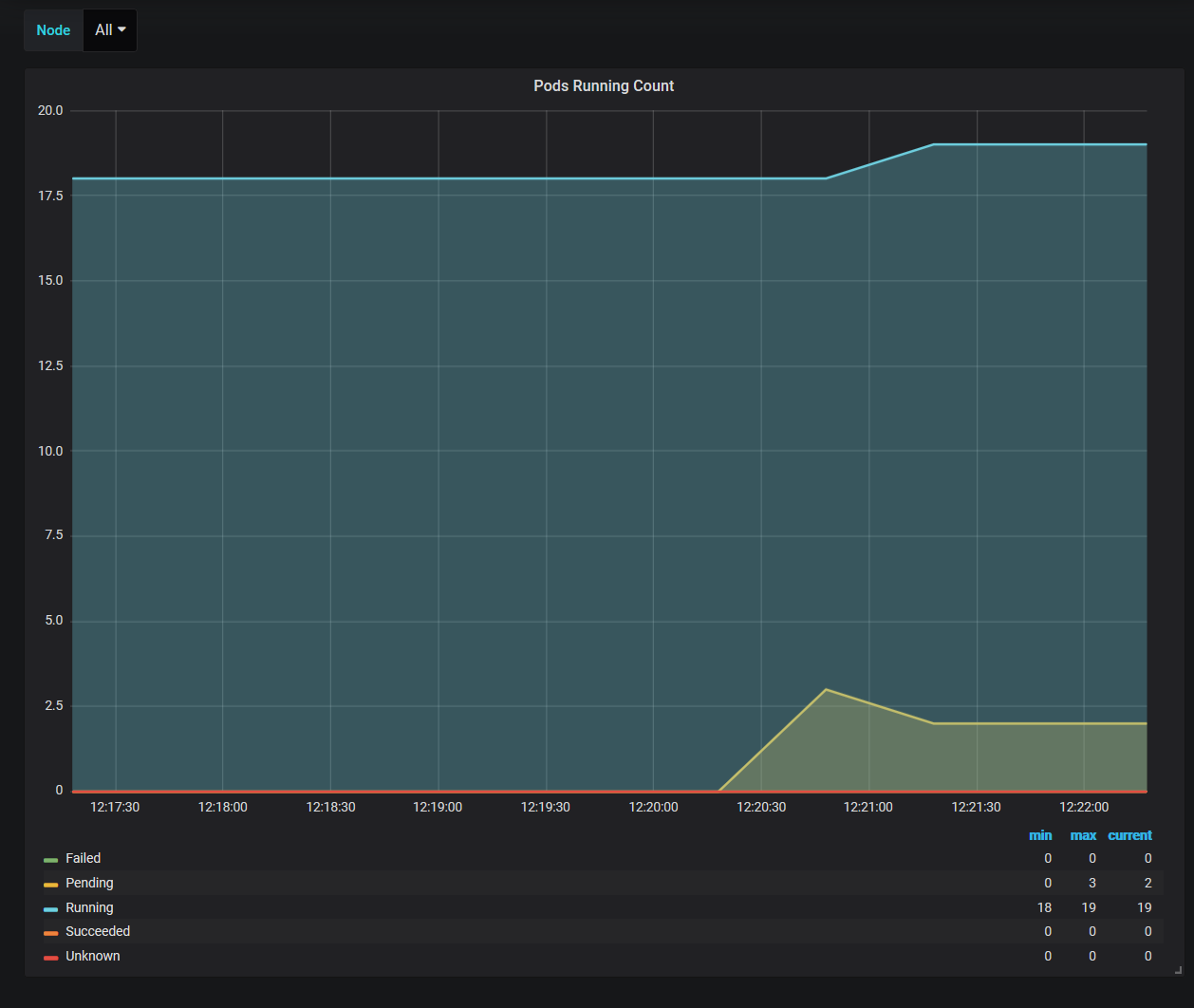
3. Scale out the Web application

* To simulate a real-world scenario we are going to scale the web app to handle increased load.

# This command will create multiple replicas of the heroes-web pod to simulate additional load on the cluster.

kubectl scale deploy/heroes-web-deploy --replicas=4

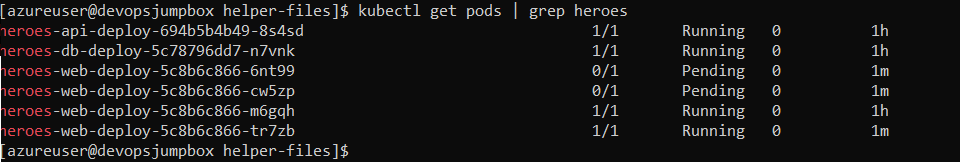
1. Check to see number of pods now running via Grafana Dashboard



1. Check to see number of heroes pods running via kubectl

kubectl get pod | grep heroes

# You should see something like the following as output (more than one heroes-web pod and some of them in different states):



1. Check up on Pods Running in Grafana dashboard

* As you can see we have a number of pods that are in the pending state which means they are trying to be scheduled to run. In this scenario the cluster is out of capacity so they are not able to be scheduled.

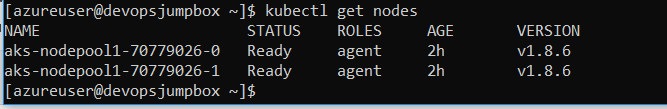


**Exercise 2 - Scale K8s Cluster**

1. Check to see number of current nodes running.

kubectl get nodes

# You should see something like the following as output (there is one node in the cluster):



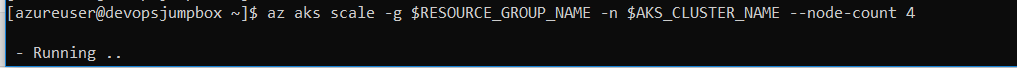
2 Scale out AKS cluster to accomodate the demand

# set these values to match yours (the cluster and the RG are the same name)

RESOURCE\_GROUP\_NAME=mstrdevopsaks-rg

AKS\_CLUSTER\_NAME=devopsaks

az aks scale -g $RESOURCE\_GROUP\_NAME -n $AKS\_CLUSTER\_NAME --node-count 4

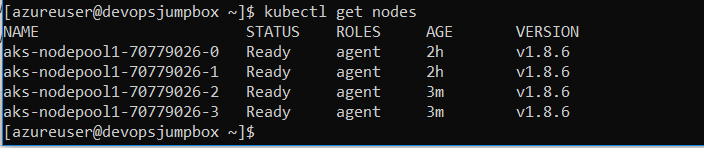


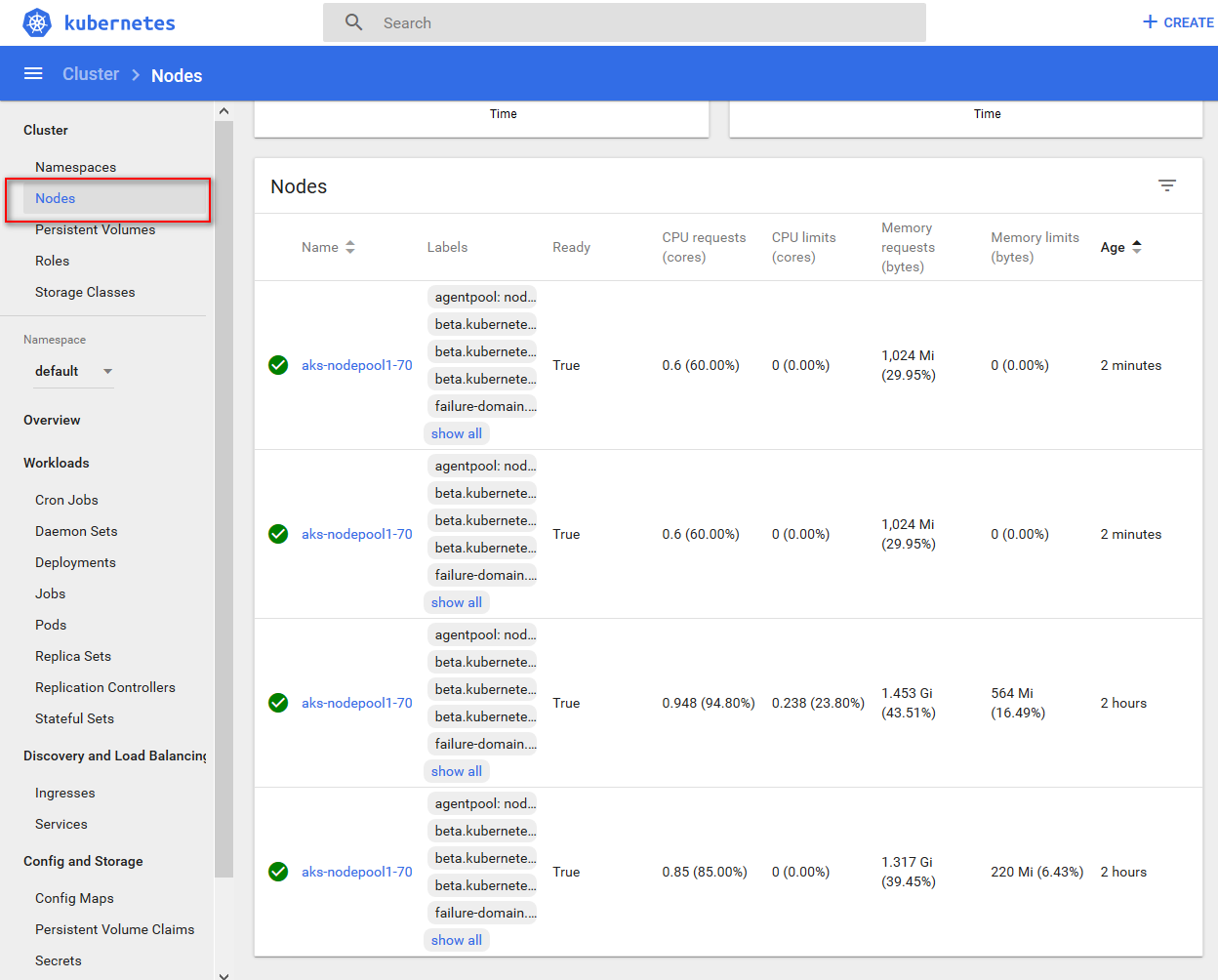
Note this may take some time. Good time to get some coffee.

1. Check to see if the new nodes are deployed and "Ready"

kubectl get nodes

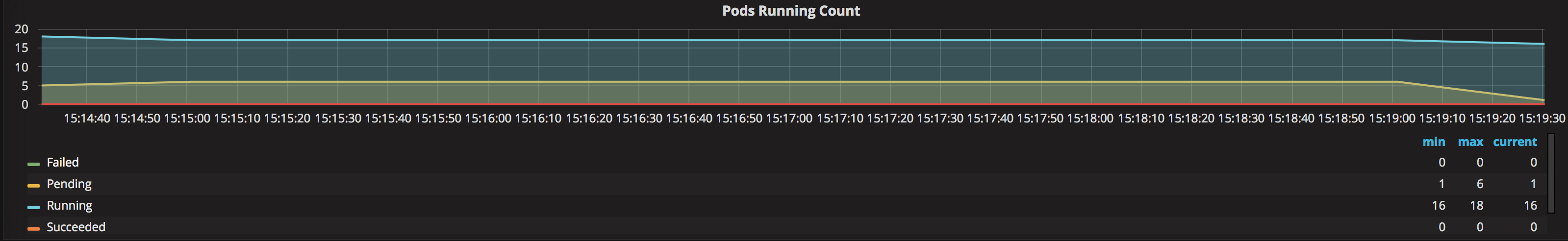
# You should see something like the following as output (there are now 4 nodes in the cluster):





4.Re-visit Grafana Dasboard to validate cluster scale is working.

* Take a look at the **Pods Pending Count** again and you should see that after a few minutes the number of pending pods is going down.

[](https://github.com/eozkurt/blackbelt-aks-hackfest/blob/master/linux-container-workshop/hol-content/img/9-grafana_podsscaling.png)

You now have additional node capacity in your Azure Kubernetes Service cluster to be able to provision more pods.