# Kubernetes

It manages the cluster

It’s a cluster of servers which can manage ur application always by increasing and decreasing the no of servers

Cluster== group of servers



It’s a container orchestration tool, used to manage the cluster of servers

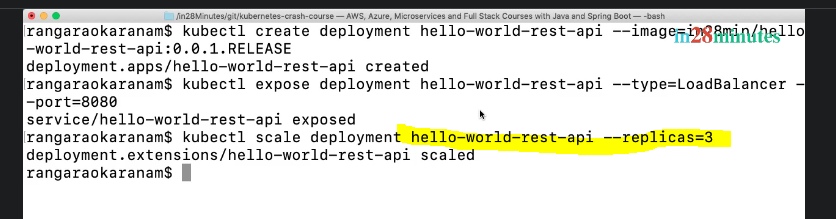
If u want to scale up the application with more servers / if u want to dynamically increase the number of servers its possible with Kubernetes.

With Kubernetes, the load will be distributed equally amongst all the servers.

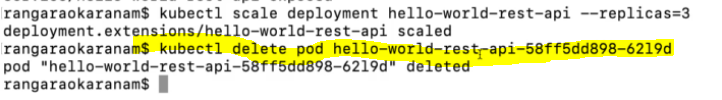
With Kubernetes

* Enable scaling -U can create multiple instances of an application
* U can distribute the load equally.
* If 1 instance goes down , within 5-10 seconds, a new instance will be created automatically.
* Zero down time deployment -means Without app going down , old will be un-installed and new will be installed parallelly

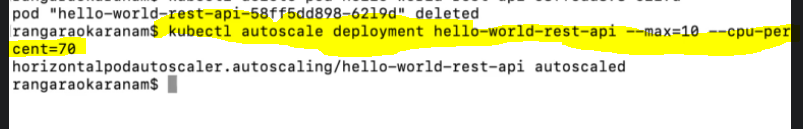
### Manual scaling -- Command to create 3 instances



### How to delete an instance



### Auto-scaling upto 10 instances

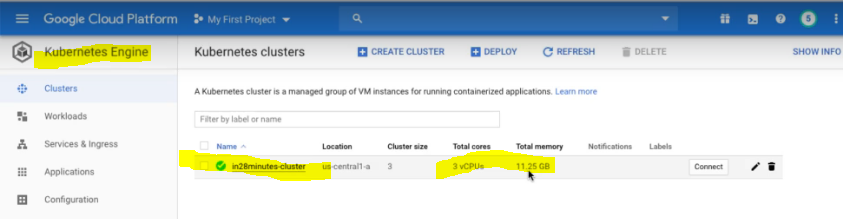


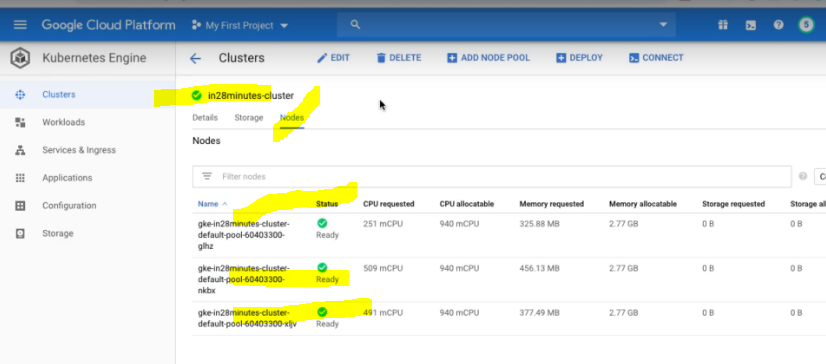
# GCP

Go to cloud.google.com

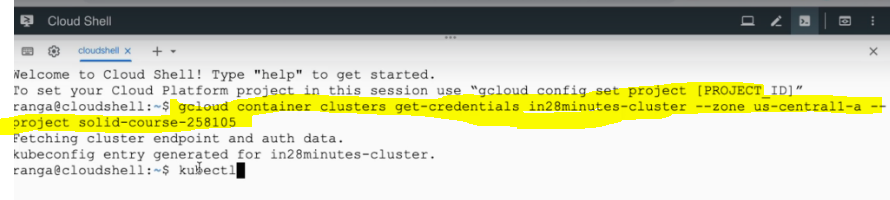
In GCP we created a cluster- this cluster is having 3 nodes / servers

### Cluster creation





### Connect to the Cluster using cmd



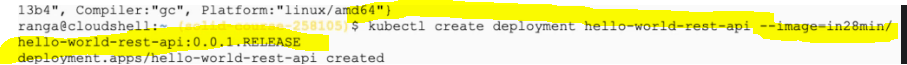
Kubectl means -kube controller

Sample command

Kubectl version

### Deploying our appn to GCP

Before that, make the docker image ready, because we will give that docker image to that Google Kubernetes engine.



### Kubernetes commands

Kubectl get events

Means while deploying 1st pod will be created, and replica set will be created, later image will be pulled .. these are the list of events will be occurring