

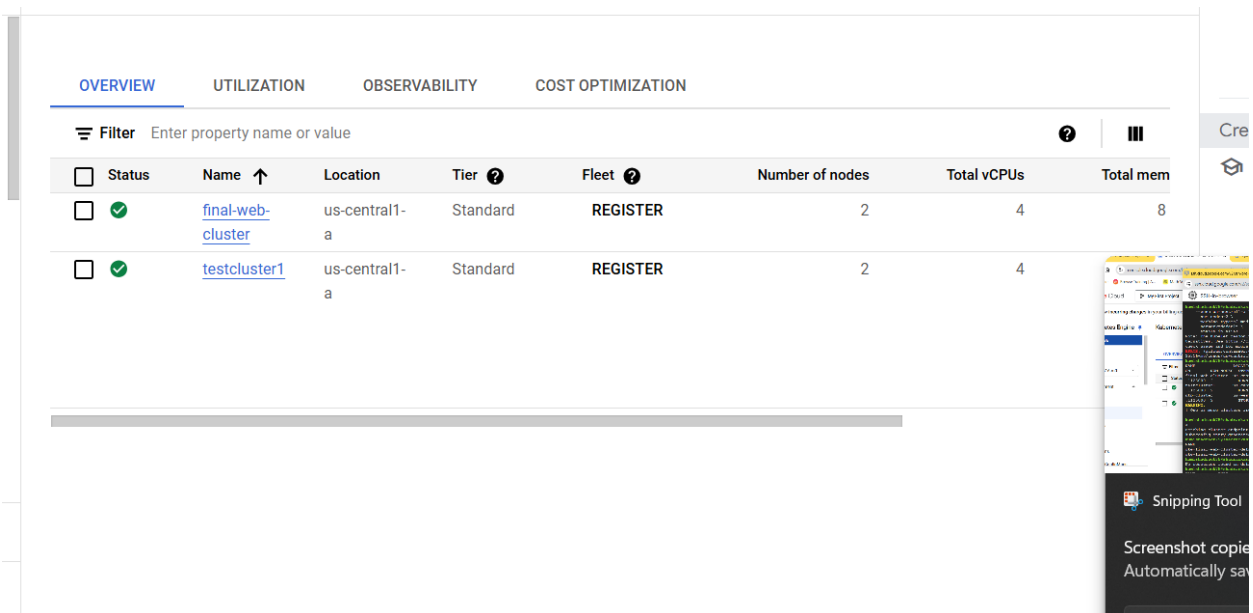
## 1. Create a GKE Cluster

### 1. GCP Project Setup:

- Navigate to the Google Cloud Console.
- Create a new project and enable the **Google Kubernetes Engine API**.

### 2. Cluster Configuration:

- Use the **GKE Console** or **gcloud CLI** to create a cluster:



The screenshot displays the Google Cloud Platform GKE console. At the top, there are tabs for OVERVIEW, UTILIZATION, OBSERVABILITY, and COST OPTIMIZATION. Below these is a filter bar with the text "Filter Enter property name or value". The main content is a table listing GKE clusters. The table has columns for Status, Name, Location, Tier, Fleet, Number of nodes, Total vCPUs, and Total mem. Two clusters are listed: "final-web-cluster" and "testcluster1", both with a status of "REGISTER" and 2 nodes. A Snipping Tool window is visible in the bottom right corner, showing a screenshot of the console and the text "Screenshot copied. Automatically saved."

Status	Name	Location	Tier	Fleet	Number of nodes	Total vCPUs	Total mem
<input checked="" type="checkbox"/>	<a href="#">final-web-cluster</a>	us-central1-a	Standard	REGISTER	2	4	8
<input checked="" type="checkbox"/>	<a href="#">testcluster1</a>	us-central1-a	Standard	REGISTER	2	4	

The screenshot shows the Google Cloud Platform console with a terminal window open. The terminal displays the following commands and output:

```
kumishadrack75@shadrackasianvm:~$ gcloud container clusters create final-web-cluster \
--zone us-central1-a \
--num-nodes=2 \
--machine-type=e2-medium \
--network=default \
--enable-ip-alias

Note: The Kubelet readonly port (10255) is now deprecated. Please update your workloads to use the recommended alternatives. See https://cloud.google.com/kubernetes-engine/docs/how-to/disable-kubelet-readonly-port for ways to check usage and for migration instructions.
ERROR: (gcloud.container.clusters.create) ResponseError: code=409, message=Already exists: projects/ace-mission-435516-m0/zones/us-central1-a/clusters/final-web-cluster.

kumishadrack75@shadrackasianvm:~$ gcloud container clusters list

NAME          LOCATION    MASTER_VERSION  MASTER_IP      MACHINE_TYPE  NODE_VERSION
final-web-cluster  us-central1-a  1.30.6-gke.1125000  34.45.155.79   e2-medium     1.30.6-gke.1125000
testcluster1     us-central1-a  1.30.6-gke.1125000  34.123.122.214 e2-medium     1.30.6-gke.1125000
ntp-cluster      us-west1     1.30.6-gke.1125000 (! 13 days left !)  34.169.18.197  e2-medium     1.30.6-gke.1125000
! One or more clusters are approaching expiration and will be deleted.

kumishadrack75@shadrackasianvm:~$ gcloud container clusters get-credentials final-web-cluster --zone us-central1-a
Fetching cluster endpoint and auth data.
kubeconfig entry generated for final-web-cluster.

kumishadrack75@shadrackasianvm:~$ kubectl get nodes

NAME                                STATUS    ROLES    AGE    VERSION
gke-final-web-cluster-default-pool-b80c753b-lqx7  Ready    <none>    7m59s  v1.30.6-gke.1125000
gke-final-web-cluster-default-pool-b80c753b-v29h  Ready    <none>    8m     v1.30.6-gke.1125000

kumishadrack75@shadrackasianvm:~$ kubectl get pods
No resources found in default namespace.

kumishadrack75@shadrackasianvm:~$ kubectl get services

NAME      TYPE        CLUSTER-IP    EXTERNAL-IP    PORT(S)    AGE
kubernetes  ClusterIP   34.118.224.1   <none>         443/TCP    9m50s
```

Note the cluster's name, zone (us-central1-a), machine type, and number of nodes.

## Firewall Rules:

- Ensure the default VPC network has firewall rules to allow external traffic on the web server's port (e.g., HTTP/80).

## 2. Create a Persistent Disk

### 1. Persistent Disk Setup:

- Create a 10GB Persistent Disk:

The screenshot shows a Google Cloud Platform console window with an SSH-in-browser terminal. The terminal displays the following commands and output:

```
kumishadrack75@shadrackasianvm:~$ gcloud compute disks create web-content-pd --size=10GB --zone=us-central1-a
WARNING: You have selected a disk size of under [200GB]. This may result in poor I/O performance. For more information, see: https://developers.google.com/compute/docs/disks#performance.
Created [https://www.googleapis.com/compute/v1/projects/ace-mission-435516-m0/zones/us-central1-a/disks/web-content-pd].
```

NAME	ZONE	SIZE_GB	TYPE	STATUS
web-content-pd	us-central1-a	10	pd-standard	READY

Below the table, a message states: "New disks are unformatted. You must format and mount a disk before it can be used. You can find instructions on how to do this at: https://cloud.google.com/compute/docs/disks/add-persistent-disk#formatting".

On the right side of the console, there are links for "GKE overview", "Help document", "Create a cluster and workload", and "Explore the cluster".

Document its name (web-content-pd), size, and zone.

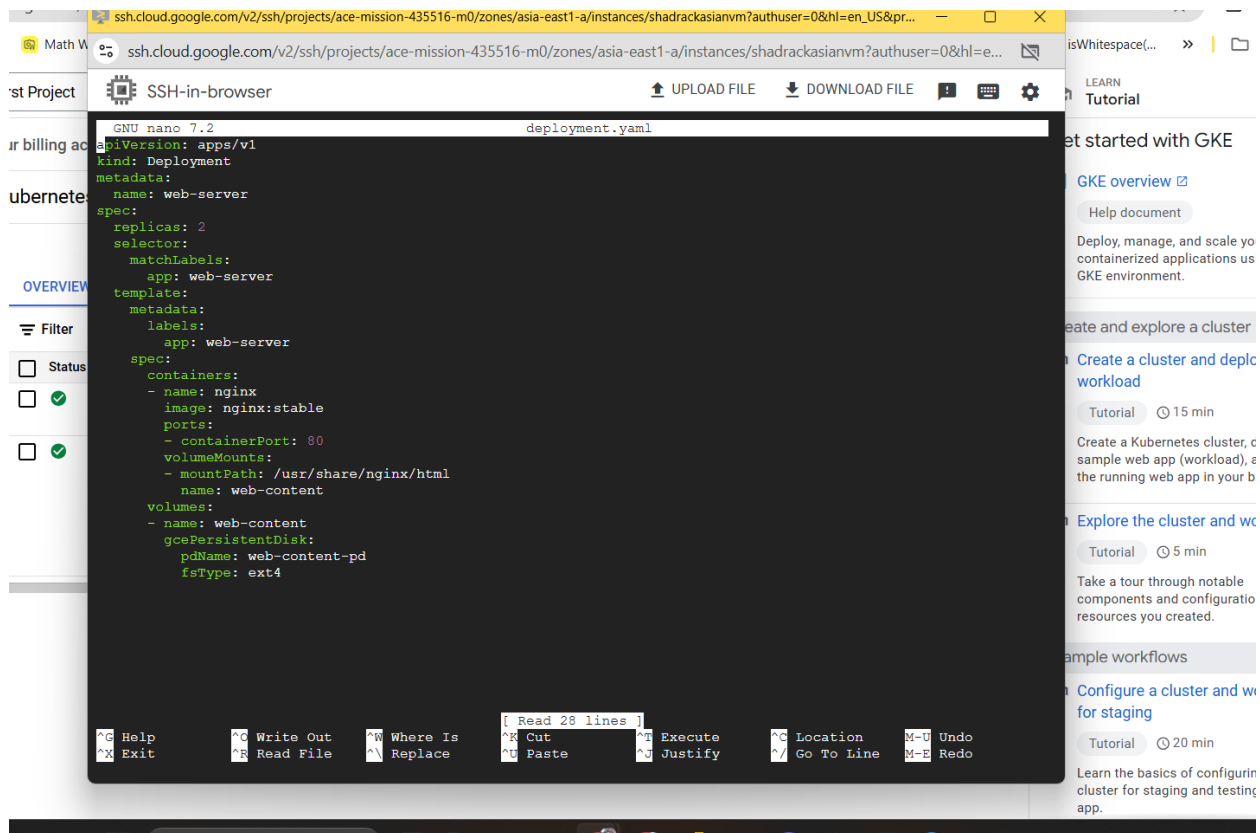
### 3. Deploy the Web Server

#### 1. Choose Container Image:

- Use a pre-built Nginx image: nginx:stable.

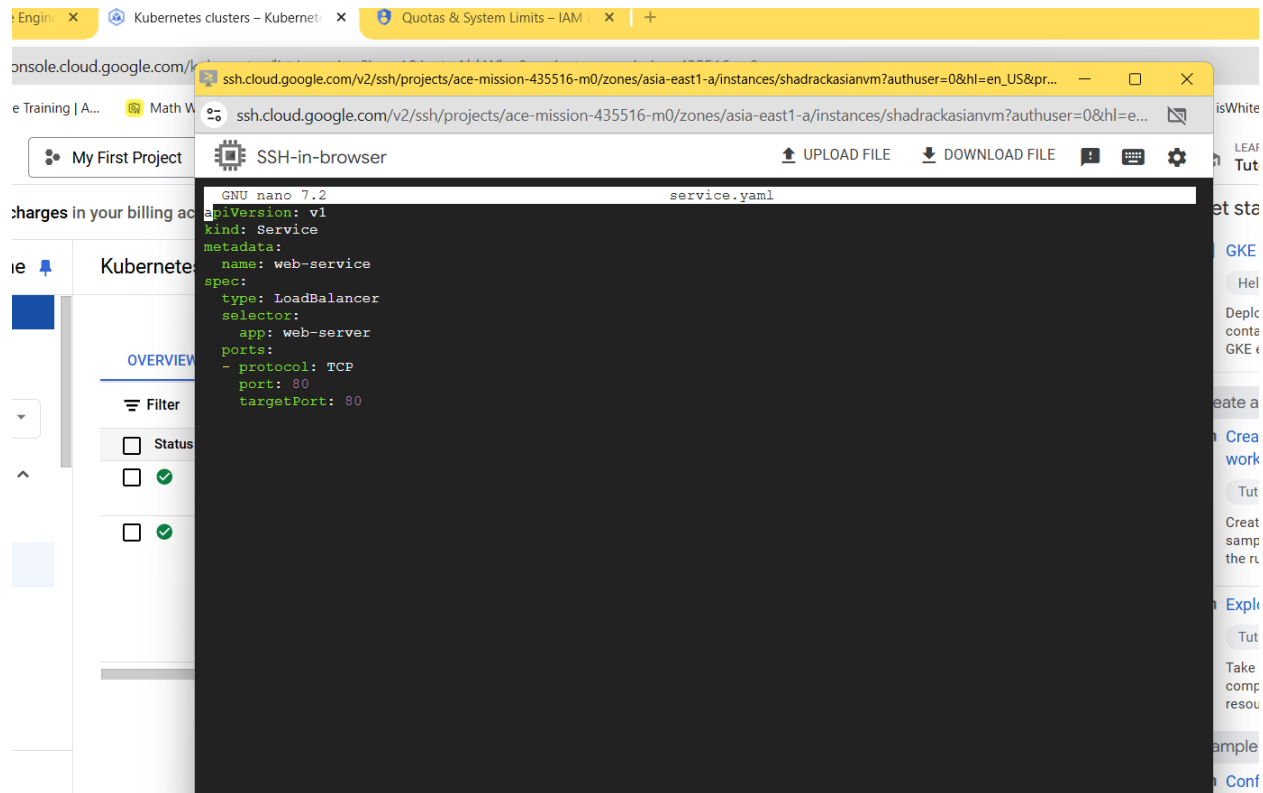
#### 2. Create Deployment Manifest:

- Save the following to deployment.yaml



## Create Service Manifest:

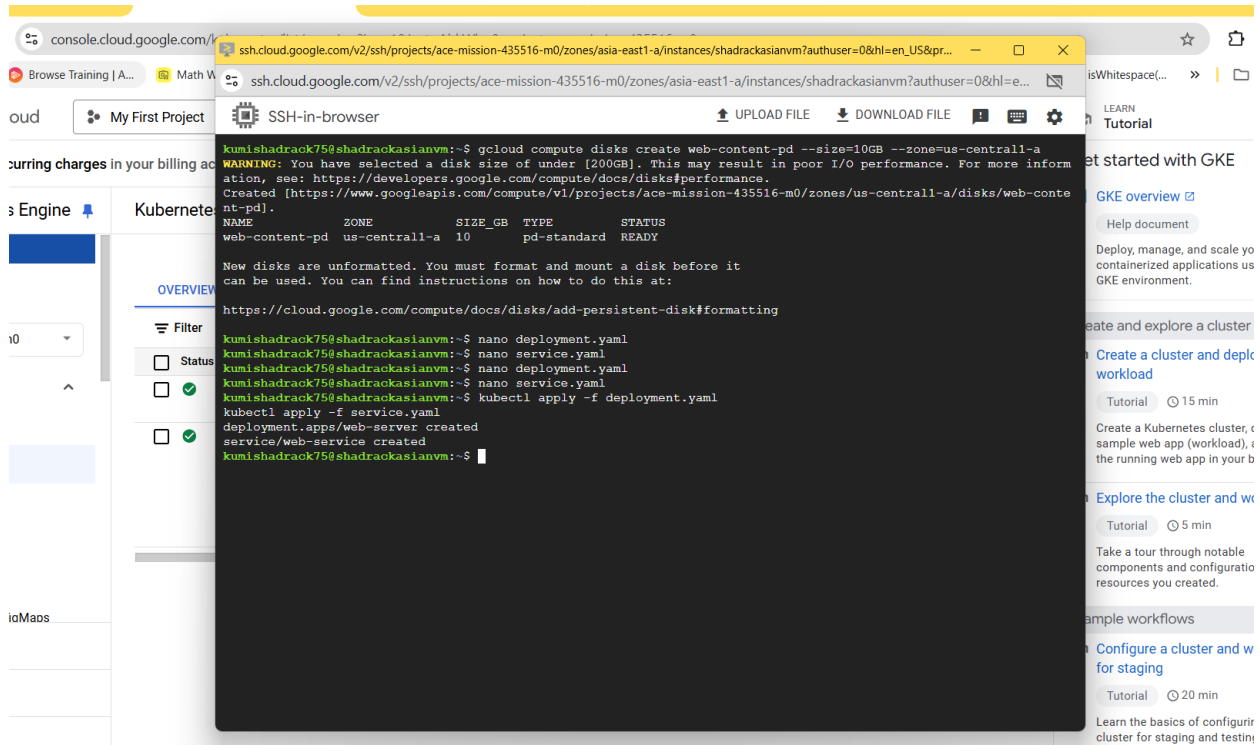
- Save the following to service.yaml



### Deploy Resources:

kubectl apply -f deployment.yaml

kubectl apply -f service.yaml



## Verify Deployment:

- Get the external IP address:

```
kubectl get service web-service
```

Test access using a browser or curl.

SSH-in-browser

UPLOAD FILE DOWNLOAD FILE

LEARN Tutorial

et started

GKE overvi

Help docu

Deploy, man: containerize GKE environi

create and exp

Create a cl workload

Tutorial

Create a Kut sample web the running v

Explore the

Tutorial

Take a tour t components resources yc

ample workf

Configure : for staging

Tutorial

Learn the ba cluster for st app.

```
kumishdrack75@shadrackasianvm:~$ gcloud compute disks create web-content-pd --size=10GB --zone=us-central1-a
WARNING: You have selected a disk size of under [200GB]. This may result in poor I/O performance. For more inform
ation, see: https://developers.google.com/compute/docs/disks#performance.
Created [https://www.googleapis.com/compute/v1/projects/ace-mission-435516-m0/zones/us-central1-a/disks/web-conte
nt-pd].
NAME                ZONE            SIZE_GB  TYPE            STATUS
web-content-pd      us-central1-a   10       pd-standard     READY

New disks are unformatted. You must format and mount a disk before it
can be used. You can find instructions on how to do this at:
https://cloud.google.com/compute/docs/disks/add-persistent-disk#formatting

kumishdrack75@shadrackasianvm:~$ nano deployment.yaml
kumishdrack75@shadrackasianvm:~$ nano service.yaml
kumishdrack75@shadrackasianvm:~$ nano deployment.yaml
kumishdrack75@shadrackasianvm:~$ nano service.yaml
kumishdrack75@shadrackasianvm:~$ kubectl apply -f deployment.yaml
deployment.apps/web-server created
kumishdrack75@shadrackasianvm:~$ kubectl apply -f service.yaml
service/web-service created
kumishdrack75@shadrackasianvm:~$ kubectl get service web-service
NAME                TYPE              CLUSTER-IP    EXTERNAL-IP    PORT(S)          AGE
web-service         LoadBalancer     34.118.230.10  34.136.211.247  80:31087/TCP     107s
kumishdrack75@shadrackasianvm:~$
```

## 4. Configure Monitoring

### 1. Enable GKE Monitoring:

- Use **Metrics Explorer** in the Cloud Console to monitor:
  - CPU and memory usage.
  - HTTP request metrics.
  - Disk I/O performance.

### 2. Set Alerts:

- Configure alerts for critical metrics:
  - Navigate to **Monitoring > Alerting** in Cloud Console.
  - Add policies for thresholds (e.g., CPU > 80%, HTTP errors).

### 3. Security Monitoring:

- Enable **Cloud Security Command Center** for threat detection.

## 5. Test the Web Server

### 1. Verify Content Serving:

- Check the web server by accessing its external IP

```
curl http://<EXTERNAL_IP>
```

### Simulate Failures:

- Kill a pod to test high availability

```
kubectl delete pod <pod-name>
```

Verify traffic is routed to healthy pods

### Trigger Alerts:

- Stress-test resources to ensure alerts work:

```
kubectl exec <pod-name> -- stress --cpu 4 --timeout 60
```





