SETTING UP LOAD BALANCER IN GCP

1. Activating cloud shell

2. Setting up default region and zone

To set up the default region in the cloud shell, use the following command:

To set up the default zone –

gcloud config set compute/zone us-central1-a

3. Creating the load balancer template

gcloud compute instance-templates create lb-backend-template \

--network=default \

--subnet=default \

--tags=allow-health-check \

--machine-type=e2-medium \

--image-family=debian-11 \

--image-project=debian-cloud \

--metadata=startup-script='#!/bin/bash \

apt-get update \

apt-get install apache2 -y a2ensite default-ssl a2enmod ssl \

vm\_hostname="$(curl -H "Metadata-Flavor:Google" \

http://169.254.169.254/computeMetadata/v1/instance/name)" echo "Page served from: $vm\_hostname" | tee /var/www/html/index.html systemctl restart apache2'

4. Creating a managed instance group based on the template

Managed instance groups (MIGs) let you operate apps on multiple identical VMs. You can make your workloads scalable and highly available by taking advantage of automated MIG services, including autoscaling, auto-healing, regional (multiple zones) deployment, and automatic updating.

gcloud compute instance-groups managed create lb-backend-group --template=lb-backend-template --size=2

5. Creating the firewall rule

Here, we’re creating a firewall rule named fw-allow-health-check

gcloud compute firewall-rules create fw-allow-health-check \

--network=default \

--action=allow \

--direction=ingress \

--source-ranges=130.211.0.0/22,35.191.0.0/16 \

--target-tags=allow-health-check \

--rules=tcp:80

6. Setting up a global static external IP address

Now that the instances are up and running, set up a global static external IP address that your customers use to reach your load balancer.

gcloud compute addresses create lb-ipv4-1 \

--ip-version=IPV4 \

--global

Note the IPv4 address that was reserved:

gcloud compute addresses describe lb-ipv4-1 \

--format="get(address)" \

--global

7. Creating a health check

Google Cloud provides health checking mechanisms that determine whether backend instances respond properly to traffic.

gcloud compute health-checks create http http-basic-check --port 80

8. **Creating a backend service**

gcloud compute backend-services create web-backend-service \

--protocol=HTTP \

--port-name=http \

--health-checks=http-basic-check \

--global

9. Adding instances to the backend service

gcloud compute backend-services add-backend web-backend-service \

--instance-group=lb-backend-group \

--instance-group-zone=us-central1-a \

--global

10. Creating a URL map

URL map is a Google Cloud configuration resource used to route requests to backend services or backend buckets. For example, with an external HTTP(S) load balancer, you can use a single URL map to route requests to different destinations based on the rules configured in the URL map:

gcloud compute url-maps create web-map-http \

--default-service=web-backend-service

11. Creating a target HTTP proxy

gcloud compute target-http-proxies create http-lb-proxy \

--url-map web-map-http

12. Creating a global forwarding rule

gcloud compute forwarding-rules create http-content-rule \

--address=lb-ipv4-1\

--global \

--target-http-proxy=http-lb-proxy \

--ports=80