Lab Overview

Building highly available applications on Compute Engine.

Task 1 Prepare

1. Select or create a Google Cloud Platform project. GO TO THE MANAGE RESOURCES PAGE
2. Make sure that billing is enabled for your Google Cloud Platform project.
3. Enable the Compute Engine API. ENABLE THE API

Task 2 Create the health check

Create a health check.

1. Go to the Health checks page in the GCP Console. GO TO THE HEALTH CHECKS PAGE
2. Click Create health check.
3. Set Name to autohealer-check
4. For Protocol select HTTP
5. Set Request path to /health. This indicates what HTTP path the health check uses. For this tutorial, the demo web server defines the path /health to return either a HTTP 200 (OK) response when healthy or a HTTP 500 (Internal Server Error) response when unhealthy.
6. Set the Health criteria:
    1. Set Check interval to 10.
    2. Set Timeout to 5.
    3. Set Healthy threshold to 2.
    4. Set Unhealthy threshold to 3.
7. Click Create at the bottom.

Create a firewall rule to allow health check probes to make HTTP requests.

1. Go to the Create firewall rule page in the GCP Console. GO TO THE CREATE FIREWALL RULE PAGE
2. For Name, enter default-allow-http-health-check
3. For Network, select default
4. For Targets, select All instances in the network
5. For Source filter, select IP ranges
6. For Source IP ranges, enter 130.211.0.0/22 and 35.191.0.0/16
7. In Protocols and ports, select tcp and enter 80
8. Click Create.

Task 2 Set up the web service

Create an instance template. Include a startup script that starts up the demo web server.

1. Go to the Instance templates page in the GCP Console. GO TO THE INSTANCE TEMPLATES PAGE
2. Click Create instance template.
3. Set the Name to webserver-template
4. For Machine type select micro (f1-micro).
5. Under Firewall, select the Allow HTTP traffic checkbox.
6. Click Management, security, disks, networking, sole tenancy to to reveal advanced settings. You should see a number of tabs.
7. Under the Management tab, find Automation and enter the following Startup script:

```
sudo apt-get update && sudo apt-get install git gunicorn3 python3-pip -y
git clone https://github.com/GoogleCloudPlatform/python-docs-samples.git
cd python-docs-samples/compute/managed-instances/demo
sudo pip3 install -r requirements.txt
sudo gunicorn3 --bind 0.0.0.0:80 app:app --daemon
```

8. Click Create at the bottom of the page.

Deploy the web server as a managed instance group.

1. Go to the Instance groups page in the GCP Console. GO TO THE INSTANCE GROUPS PAGE
2. Click Create instance group.
3. Set the Name to webserver-group
4. For Region select *europe-west1 *
5. For Zone select europe-west1-b
6. For Instance template select webserver-template
7. For Autoscaling select Off.
8. Set Number of instances to 3
9. For Health check select autohealer-check
10. Set Initial delay to 90
11. Click Create.

Create a firewall rule that will allow HTTP requests to the web servers.

1. Go to the Create firewall rule page in the GCP Console. GO TO THE CREATE FIREWALL RULE PAGE
2. For Name, enter default-allow-http
3. For Network, select default
4. For Targets, select Specified target tags
5. For Target Tags, enter http-server
6. For Source filter, select IP ranges
7. For Source IP ranges, enter 0.0.0.0/0
8. In Protocols and ports, select tcp and enter 80
9. Click Create.

Task 3 Simulate health check failures

Navigate to a web server instance.

1. Go to the VM instances page in the GCP Console. GO TO THE VM INSTANCES PAGE
2. Under the External IP column, click the ip address for any webserver-group instance. A new tab should open in your web browser. If the request times out or web page is not available, wait a minute to let the server finish setting up and try again. The demo web server displays a page similar to the following:

3. On the demo web page, click Make unhealthy.

This causes the web server to fail the health check. Specfically, the web server makes the /health path return a HTTP 500 (Internal Server Error). You can verify this yourself by quickly clicking the Check health button (this will stop working after the autohealer has started rebooting the instance).

4. Wait for the autohealer to take action.

    1. Go to the VM instances page in the GCP Console. GO TO THE VM INSTANCES PAGE
    2. Wait for the status of the web server instance to change. The green checkmark next to the instance name should change to a grey square, indicating the autohealer has started rebooting the unhealthy instance.
    3. Click Refresh at the top of the page periodically to get the most recent status.
    4. The autohealing process is finished when the grey square changes back to a green checkmark, indicating the instance is healthy again.