Cloud Computing Assignment Report
GCP Setup with Auto-Scaling & Security

Harshika Arya, B22CS023

Introduction

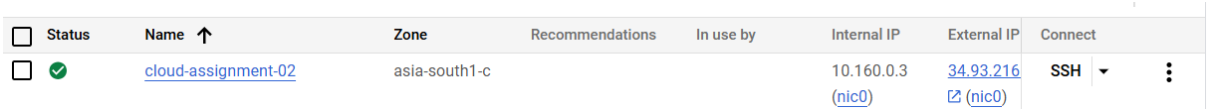
This report documents the process of setting up a Virtual Machine (VM) in Google Cloud Platform (GCP), implementing auto-scaling policies, and applying security measures such as IAM roles and firewall rules. I also tested auto-scaling by generating CPU load and observed how instances were created and deleted dynamically.

Creating a Virtual Machine (VM)

To start, I created a virtual machine in GCP by navigating to Compute Engine > VM Instances and clicking “Create Instance.”

Configuration Details:

- Name: cloud-assignment-02
- Region: asia-south1-c
- Machine Type: E2
- Boot Disk: Ubuntu 20.04 LTS
- Networking: Allowed HTTP & HTTPS traffic

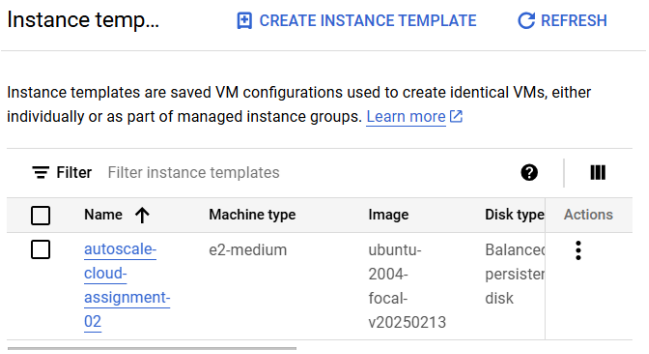


Configuring Auto-Scaling

To enable auto-scaling, I created a Managed Instance Group (MIG) using an Instance Template.

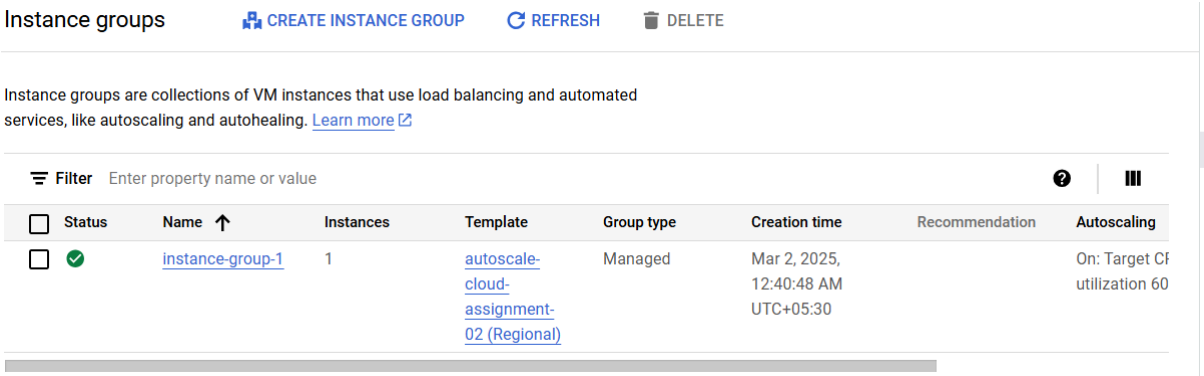
1. Creating an Instance Template

1. Compute Engine > Instance Templates.
2. Clicked "Create Instance Template" and configured it similarly to my VM.
3. Saved the template.



2. Creating a Managed Instance Group

1. Compute Engine > Instance Groups.
2. Clicked "Create Instance Group", selected Managed Instance Group, and attached my instance template.
3. Set the auto-scaling policy:
    - Scale based on CPU utilization.
    - Target CPU usage: 60%.
    - Minimum instances: 1, Maximum instances: 5.

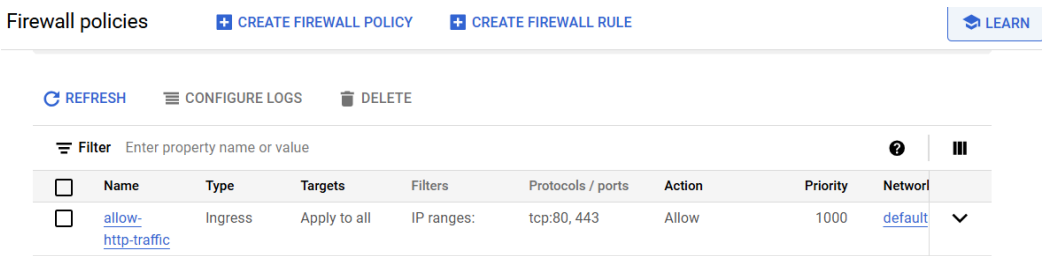


Configuring Security

1. Setting Up Firewall Rules

To control network access, I configured firewall rules:

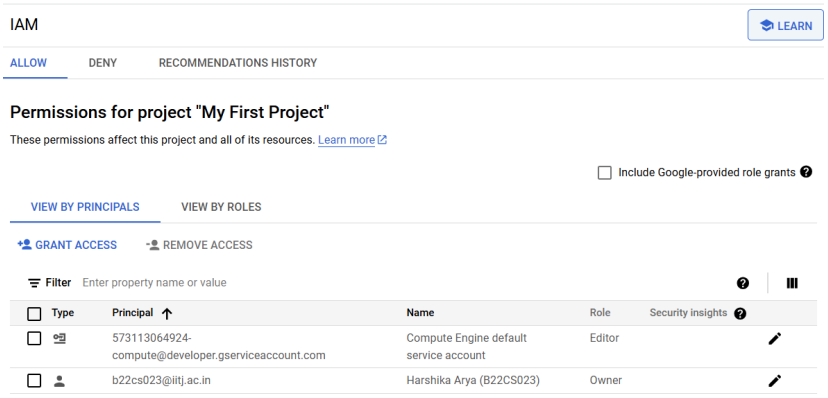
- Created a new firewall rule under VPC Network > Firewall.
- Allowed incoming HTTP & HTTPS traffic on ports 80 and 443.



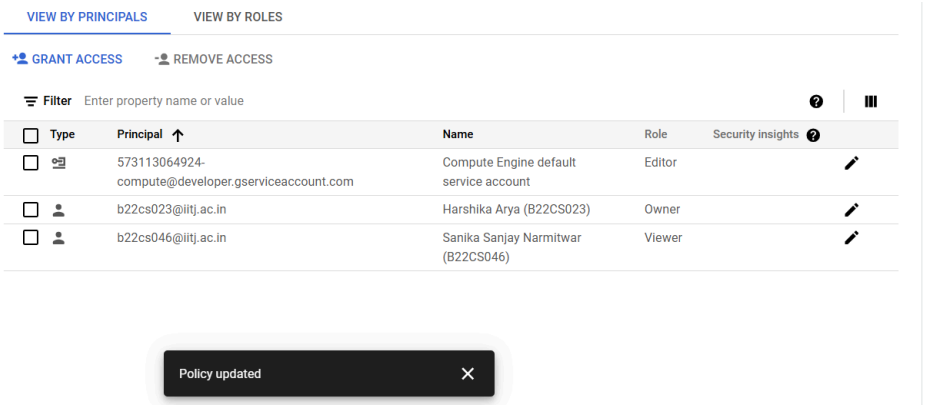
2. Setting IAM Roles for Restricted Access

To manage permissions securely:

- Went to IAM & Admin > IAM.
- There were assigned roles:
    - Compute Viewer (read-only access).
    - Compute Admin (full control).



Example of giving (read) access to another user:



Testing Auto-Scaling

1. Went to Compute Engine > Instance Groups, selected an instance, and clicked SSH.
2. To generate CPU load, I installed the stress tool:

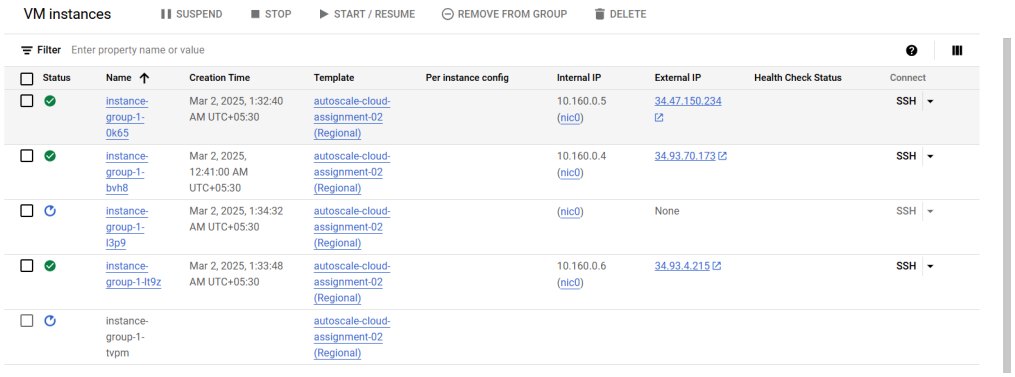
sudo apt update
sudo apt install stress -y

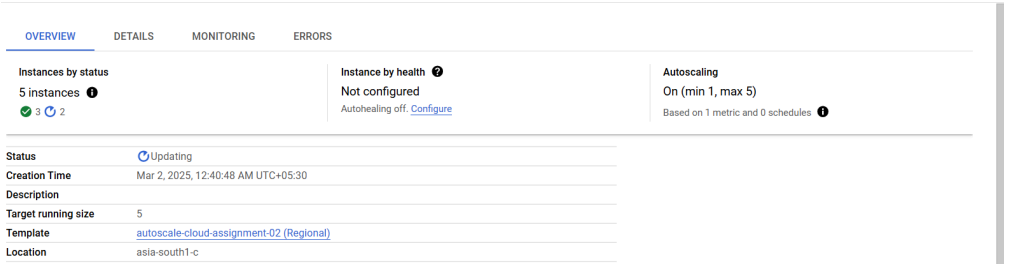
3. For Simulating High CPU Load, ran the following command:

stress --cpu 4 --timeout 300

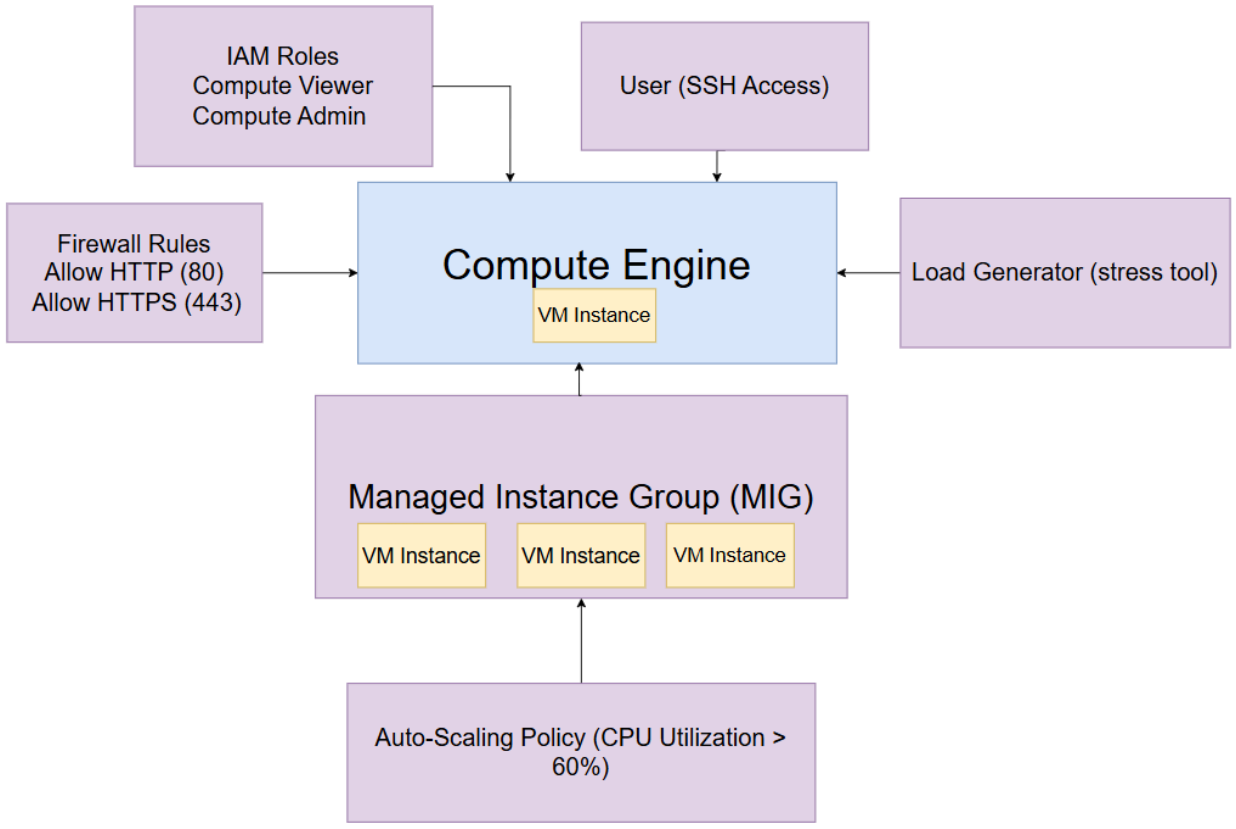
4. Observing Auto-Scaling Behavior

- Checked Compute Engine > Instance Groups and saw new instances were being created.
- After 5 minutes, instances were automatically removed as CPU usage dropped.





Architecture Design



GitHub Repo Link: Repo for Cloud Assignment 02
Video Link: Video Walkthrough of Assignment 02