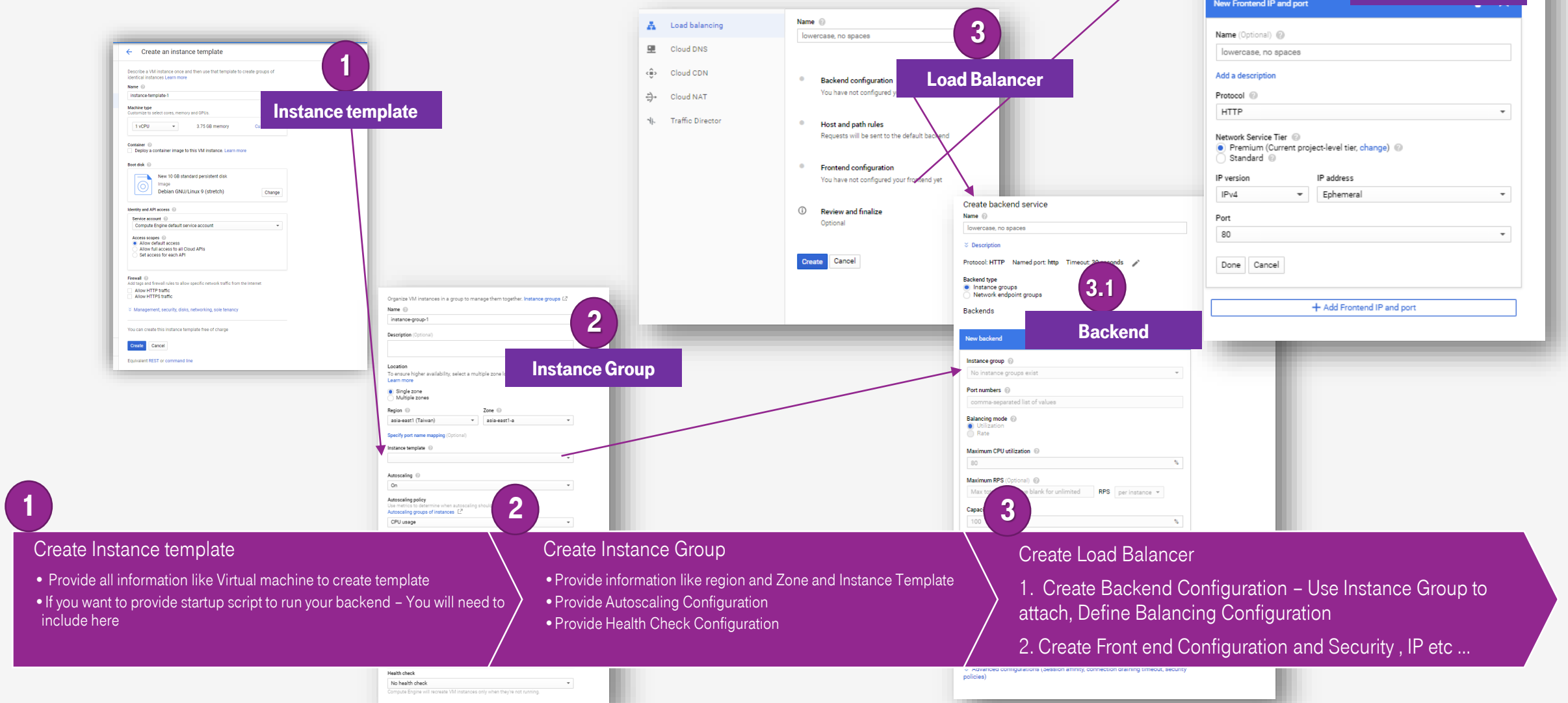


Assignment and Exam Content

HTTP Load balancer

HTTP Load Balancer Configuration - 3 Major Steps



1

Create Instance template

Create Instance template

- Provide all information like Virtual machine to create template
- If you want to provide startup script to run your backend – You will need to include here

Create Instance Group

- Provide information like region and Zone and Instance Template
- Provide Autoscaling Configuration
- Provide Health Check Configuration

Create Load Balancer

1. Create Backend Configuration – Use Instance Group to attach, Define Balancing Configuration
2. Create Front end Configuration and

1

Go To -> COMPUTE -> Compute Engine -> Instance Template, Click on Create

2

Information is equivalent to that of VM with only difference being Instance Template is **Global Resource** and you can not attach Zone and Region.

Lets Create Instance Template for USA

Fill in information as you like (Keep Machine Type to Low configuration – default 1 CPU is good.)

Download and Copy Startup Script from

<https://github.com/dhanajimusale/GCPTTrain>

Open Firewall – for HTTP Traffic (Click on Check box and make it on)

Click on **Management, Security Disks..**

Create an instance template

Describe a VM instance once and then use that template to create groups of identical instances [Learn more](#)

Name

Machine type
Customize to select cores, memory and GPUs.
1 vCPU **3.75 GB memory** [Customize](#)

Container
☐ Deploy a container image to this VM instance. [Learn more](#)

Boot disk
 New 10 GB standard persistent disk
Image: **Debian GNU/Linux 9 (stretch)** [Change](#)

Identity and API access
Service account

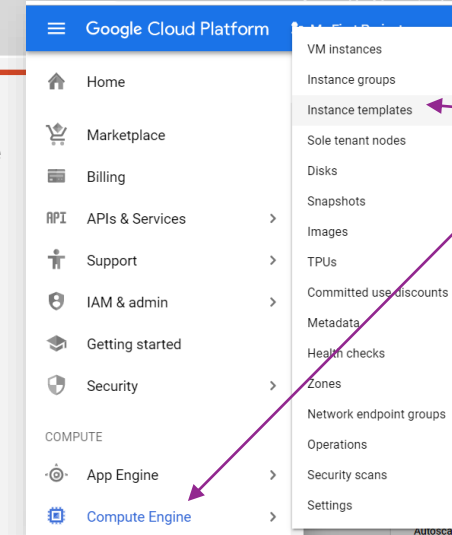
Access scopes
☒ Allow default access
☐ Allow full access to all Cloud APIs
☐ Set access for each API

Firewall
Add tags and firewall rules to allow specific network traffic from the Internet
☐ Allow HTTP traffic
☐ Allow HTTPS traffic
☒ Management, security, disks, networking, sole tenancy

You can create this instance template free of charge

Create **Cancel**

Equivalent REST or command line



instance template

We should Create

1 Instance Template per region for at least 2 regions – Please name it accordingly.
with Different Startup Script as per Demo for Auto Scaling demo as well.

1

Create Instance template

3

Enter Startup Script attached (please check demo video)

Create 2nd Instance template as well for Asia with Startup script in it and name it appropriately.

Again Code is Located at

<https://github.com/dhanajimusale/GCPTrain>

Create Instance template

- Provide all information like Virtual machine to create template
- If you want to provide startup script to run your backend – You will need to include here

Create Instance Group

- Provide information like region and Zone and Instance Template
- Provide Autoscaling Configuration
- Provide Health Check Configuration

Create Load Balancer

1. Create Backend Configuration – Use Instance Group to attach, Define Balancing Configuration
2. Create Front end Configuration and Security , IP etc ...

← Create an instance template

Management Security Disks Networking Sole Tenancy

Description (Optional)

Labels (Optional)

+ Add label

Automation

Startup script (Optional)

You can choose to specify a startup script that will run when your instance boots up or restarts. Startup scripts can be used to install software and updates, and to ensure that services are running within the virtual machine. [Learn more](#)

Startup Script

Metadata (Optional)

You can set custom metadata for an instance or project outside of the server-defined metadata. This is useful for passing in arbitrary values to your project or instance that can be queried by your code on the instance. [Learn more](#)

Key	Value

+ Add item

Availability policy

Preemptibility

A preemptible VM costs much less, but lasts only 24 hours. It can be terminated sooner due to system demands. [Learn more](#)

Off (recommended)

On host maintenance

When Compute Engine performs periodic infrastructure maintenance it can migrate your VM instances to other hardware without downtime

Migrate VM instance (recommended)

Automatic restart

Compute Engine can automatically restart VM instances if they are terminated for non-user-initiated reasons (maintenance event, hardware failure, software failure and so on)

On (recommended)

[Less](#)

Things to remember

1. Instance template is *Global Resource*, but you can attach *Zonal resources* to make it *Zonal*.

2. Instance template is just template definitions to create instance. Creating Instance Template will not create actual instance.

3. Instance Template can be used to create VM directly or used by Instance Group/

2

Create Instance Group

1

Go To -> COMPUTE -> Compute Engine -> Instance Groups, Click on Create

2

There are three main components of Instance Groups Information -> 1. Instance template and remaining information (for Compute engine to start) Region and Zone , 2. Autoscaling configuration, 3. Health Check

Try to Choose Multiple Zones (Click radio button) to have configurations from failure of one Zone.

Create two instance Groups one for USA and other for Asia.

1. Choose Instance Template and corresponding Region (USA and Asia)
2. Choose Autoscaling Policy – CPU Utilization and Other param in Screen.
3. Configure Health Check – Use Port 80.

Once you create Instance Group – Compute Engine will start Virtual machines specified minimum number of instances.

Create Instance template

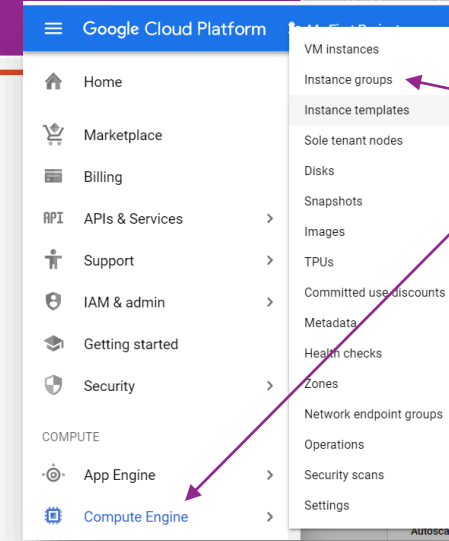
- Provide all information like Virtual machine to create template
- If you want to provide startup script to run your backend – You will need to include here

Create Instance Group

- Provide information like region and Zone and Instance Template
- Provide Autoscaling Configuration
- Provide Health Check Configuration

Create Load Balancer

1. Create Backend Configuration – Use Instance Group to attach, Define Balancing Configuration
2. Create Front end Configuration and



Instance Groups

Organize VM instances in a group to manage them together. [Instance groups](#) [?]

Name [?]
instance-group-1

Description (Optional)

Location
To ensure higher availability, select a multiple zone location for an instance group. [Learn more](#)

☒ Single zone
☐ Multiple zones

Region [?]
asia-east1 (Taiwan)

Zone [?]
asia-east1-a

Specify port name mapping (Optional)

Instance template [?]

Autoscaling [?]
On

Autoscaling policy
Use metrics to determine when autoscaling should resize the group. [Autoscaling groups of instances](#) [?]
CPU usage

Target CPU usage [?]
60 %

Minimum number of instances [?]
1

Maximum number of instances [?]
10

Cool down period [?]
60 seconds

Autohealing [?]
To use autohealing, configure firewall rules. This will allow health checks to connect to VM instances in a group. [How to configure firewall rules?](#) [?]

Health check
No health check

Compute Engine will recreate VM instances only when they're not running.

Things to remember

1. Instance Group used for configuration to mention how many instance should run.
2. HTTP Load Balancer use Instance groups to configure backend for Compute Service.

3

Configure HTTP Load Balancer

Create Instance template

- Provide all information like Virtual machine to create template
- If you want to provide startup script to run your backend – You will need to include here

Create Instance Group

- Provide information like region and Zone and Instance Template
- Provide Autoscaling Configuration
- Provide Health Check Configuration

Create Load Balancer

1. Create Backend Configuration – Use Instance Group to attach, Define Balancing Configuration
2. Create Front end Configuration and Security , IP etc ...

1

Go To -> NETWORKING -> Network Services -> Load balancing , Click on Create Load Balancer and Select HTTP Load Balancer.

2

Load Balancer has 3 Main Configurations

Backend, Host & Path Rules, Frontend configurations

1. Backend Configuration : backend let your configure backend compute service or Cloud Storage (for Images , Videos etc). Backend configure also let you configure balancing mode which will let HTTP Load Balancer how to choose backend instances.

Select

Backend Type : Instance Group

- Configure New Backend - Select Instance Group you created for USA
- Select **Balancing Mode** as Utilization.
- **Max CPU – 80%** Leave all other things as default

Add Another Backend for **Asia** instance group similarly.

You don't have to choose **Cloud CDN**

The screenshot illustrates the Google Cloud Platform console interface for configuring an HTTP Load Balancer. It shows the navigation path from the 'Load Balancing' menu to the 'Create backend service' page. The 'Backend' configuration screen is highlighted, showing the 'Instance groups' backend type and the 'Utilization' balancing mode with a maximum CPU utilization of 80%.

3.1 Backend

Create backend service

Name: lowercase, no spaces

Description:

Protocol: HTTP Named port: http Timeout: 30 seconds

Backend type: ☒ Instance groups ☐ Network endpoint groups

Backends

New backend

Instance group: No instance groups exist

Port numbers: comma-separated list of values

Balancing mode: ☒ Utilization ☐ Rate

Maximum CPU utilization: 80%

Maximum RPS (Optional): Max total RPS. Leave blank for unlimited RPS per instance

Capacity: 100%

Cloud CDN: ☐ Enable Cloud CDN

Health check:

Advanced configurations (Session affinity, connection draining timeout, security policies)

3

Configure HTTP Load Balancer

Create Instance template

- Provide all information like Virtual machine to create template
- If you want to provide startup script to run your backend – You will need to include here

Create Instance Group

- Provide information like region and Zone and Instance Template
- Provide Autoscaling Configuration
- Provide Health Check Configuration

Create Load Balancer

1. Create Backend Configuration – Use Instance Group to attach, Define Balancing Configuration
2. Create Front end Configuration and Security , IP etc ...

2

2. **Host and Path Rules** -> If you want to route traffic to backend based on type of request e.g. php , Images or videos etc

We don't have specific configurations now. Leave it blank.

3

3. **Frontend Configurations** : frontend let you configure Load balancer IP , Network Tier, Protocol etc

Leave HTTP protocol and leave all others as default.

Frontend configuration

Specify an IP address, port and protocol. This IP address is the frontend IP for your clients requests. For SSL, a certificate must also be assigned.

New Frontend IP and port

Name (Optional) ?
lowercase, no spaces

Add a description

Protocol ?
HTTP

Network Service Tier ?
☒ Premium (Current project-level tier, change) ?
☐ Standard ?

IP version
IPv4

IP address
Ephemeral

Port
80

Done Cancel

+ Add Frontend IP and port

Test HTTP Load Balancer

Create Instance template

- Provide all information like Virtual machine to create template
- If you want to provide startup script to run your backend – You will need to include here

Create Instance Group

- Provide information like region and Zone and Instance Template
- Provide Autoscaling Configuration
- Provide Health Check Configuration

Create Load Balancer

1. Create Backend Configuration – Use Instance Group to attach, Define Balancing Configuration
2. Create Front end Configuration and Security , IP etc ...

You will clients in different Zone – Which you can do it using VM again – You can use AWS or GCP VM. We are using GCP VMs, Refer Demo for any doubts

1. Client Instances

Launch VM in Regions/Zones you created Instance Template/Groups (at least two)

2. SSH instance using console and Use following commands to access Service

Curl Command Slow

```
$ while true ; do (curl http://[Load Balancer ExternalIP]/service ) ; sleep 2 ; done
```

Curl Command Fast

```
$ while true ; do (curl http://[Load Balancer ExternalIP]/service ) ; sleep 0.1 ; done
```


HTTP Load Balancer : Try Yourself

1

Install Apache , and Host different type of content like PHP and Images or Videos

Use Host Rule to configure different end point or backend. Host images on Cloud Storage and Enable CDN.

2

Create Instance Group without Auto Scaling and see behavior

3

HTTP Load Balancer – Backend Service configure Balancing Mode to RPS (Request per Second)

4

Exam Tips

Try different techniques/Configurations on Balancing Modes , Auto Scaling – **Very Important for Exam.**

Test HTTP Load Balancer

Create Instance template

- Provide all information like Virtual machine to create template
- If you want to provide startup script to run your backend – You will need to include here

Create Instance Group

- Provide information like region and Zone and Instance Template
- Provide Autoscaling Configuration
- Provide Health Check Configuration

Create Load Balancer

1. Create Backend Configuration – Use Instance Group to attach, Define Balancing Configuration
2. Create Front end Configuration and Security , IP etc ...

End of HTTP LB Assignment