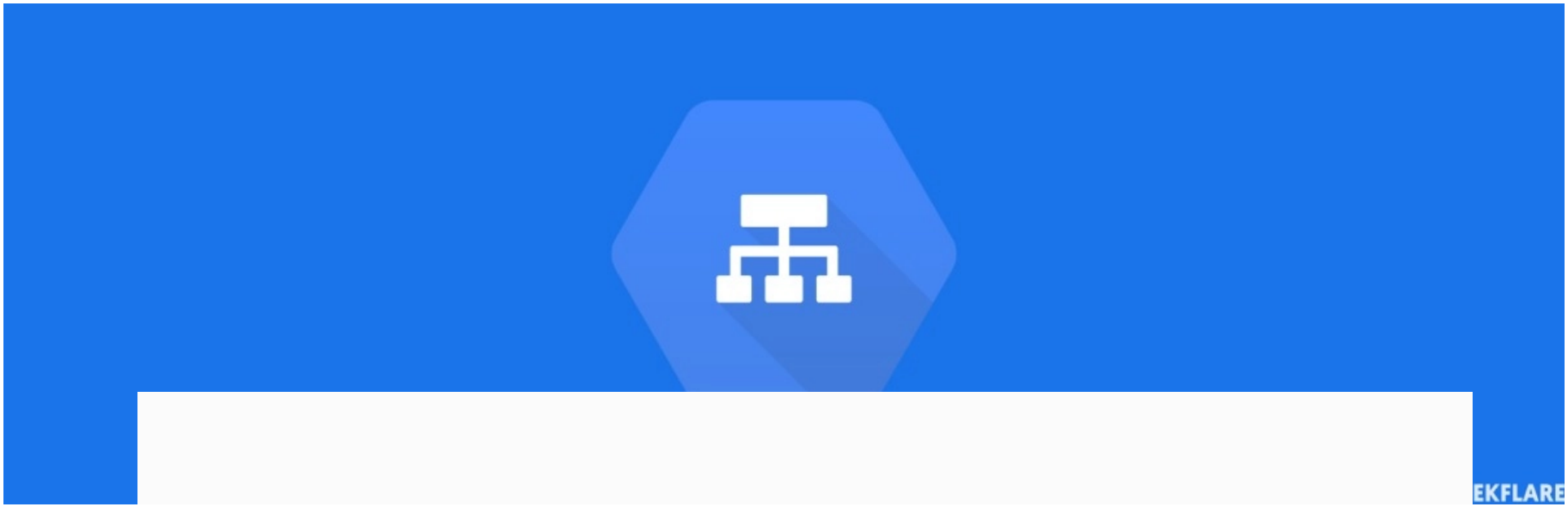


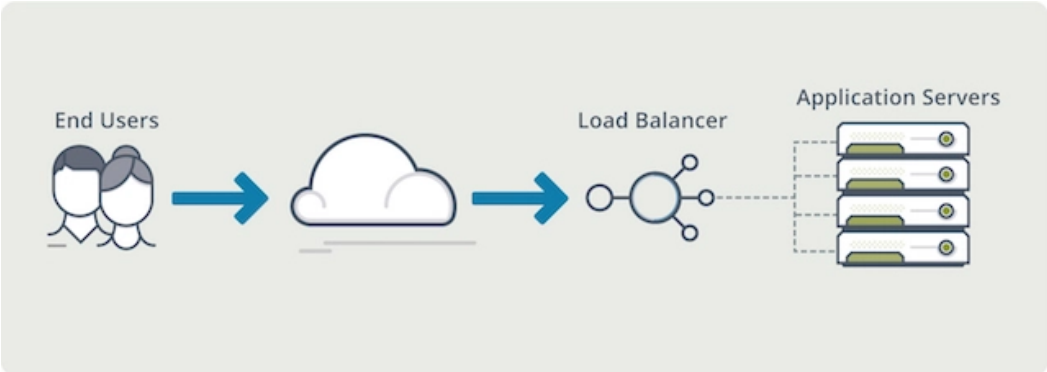
# How to Create a Load Balancer on Google Cloud?

By Chandan Kumar on March 14, 2019

Posted in Cloud Computing

> Creating an HTTP(s) cloud load balancer on the Google Cloud Platform (GCP)

If you are hosting your applications on Google Cloud and looking for better high-availability, then you should try implementing a load balancer (LB).




Load Balancer Concept

Google Cloud LB is smart. It offers more than a traditional one.

- HTTP/2 enabled

- Terminate SSL handshake
- Custom SSL/TLS policies
- Route traffic to the nearest server
- Path-based routing
- Auto-scaling

and a lot more...

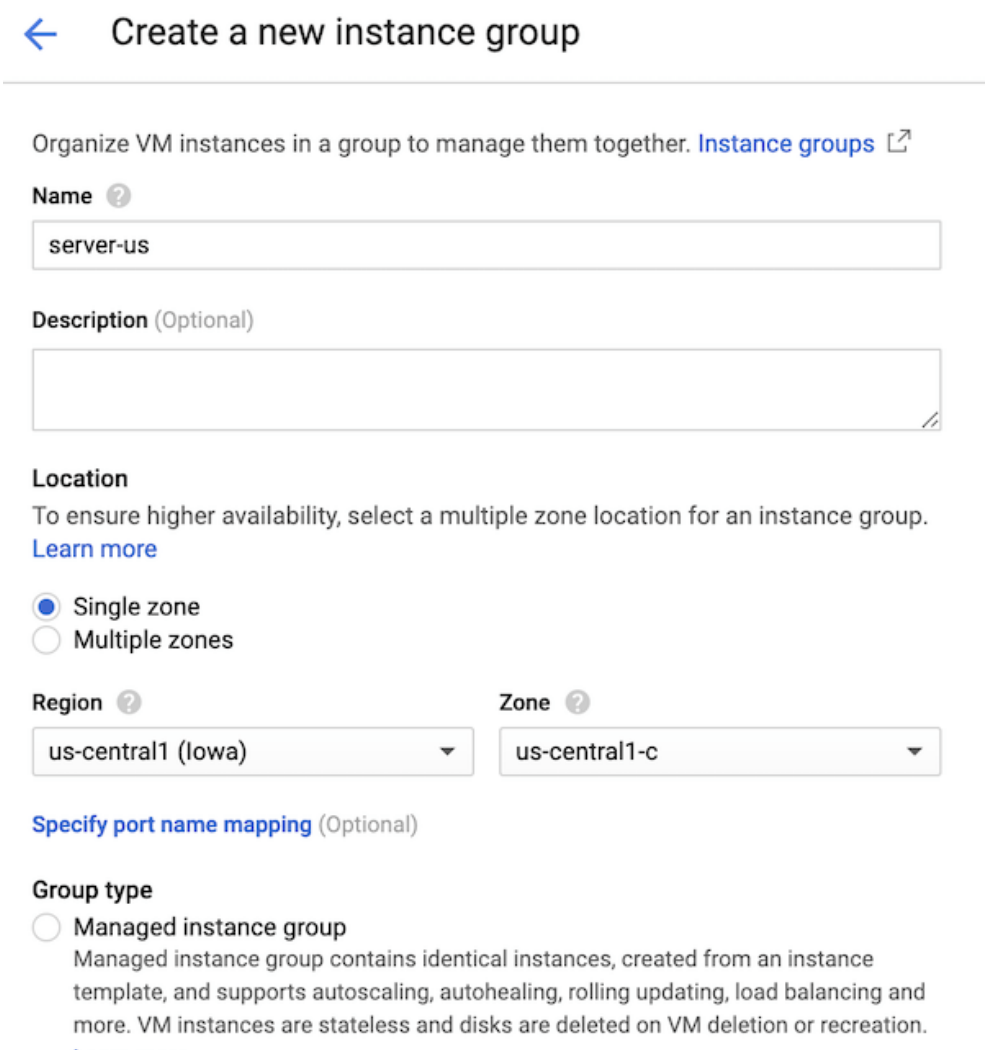
The following, I have two servers (one in the US and another one in the UK). Let's create a load balancer and route traffic to both the servers.

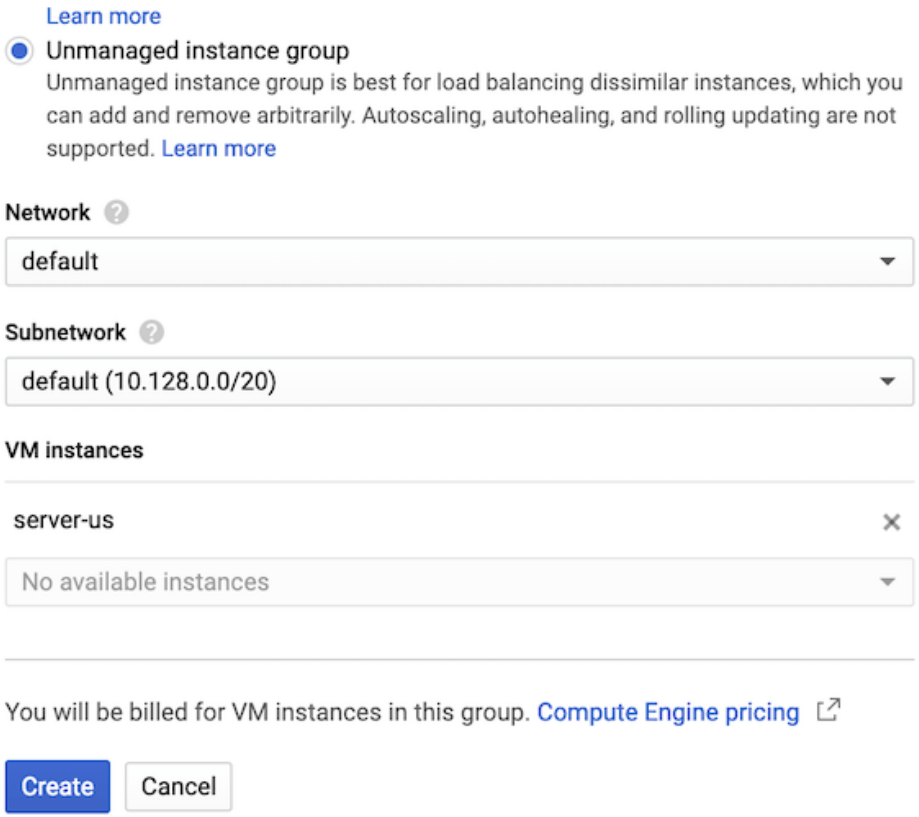

Since Google offers auto-scaling, you have multiple options and choose what your business requires. However, in this article, I will explain how to create a load balancer using unmanaged instance groups which don't support auto-scaling.

# Create Instance Groups

All the servers should be inside the instance groups. So this is a pre-requisite to creating an LB.

- Login to GCP Console
- Navigate to Compute Engine >> Instance groups
- Click create instance group
- Enter the name, select zone as a single, region where your servers are, unmanaged instance group, choose the server from VM instance drop-down and click Create






A single zone can add servers only from that zone. Since I've servers in the UK and US, then I got to create a two instance group. I'll repeat the procedure to add my another server (server-uk).

And, those two groups are created.

# Create an HTTP(s) LB

Google offers three types of LB.

- HTTP (s)
- TCP
- UDP

To manage web application traffic distribution, HTTP(s) is suitable. Let's create that.

- Navigate to Network Services >> load balancing
- Click Create a load balancer

- Start configuration for HTTP(s) load balancing

- Enter the LB name
- On backend configuration tab, select the drop-down and create a backend service

- Enter the name and select backend type as instance groups
- Add both instance groups (server-us and server-uk)
- Adjust the port number – port number of web server or application will be listening on the servers

- Under health check, click create

- Enter the name, select the protocol, port

A health check is essential for LB to know which instance is down, so it stops sending traffic. Below, I am instructing LB to hit the server IP with port 80 every 10 seconds. If a server doesn't response three times consecutively then, LB will mark that instance down.

- Click save and continue to create health
- Click create on backend service wizard
- It will show like below

On host and path rules tab, ensure newly created backend services (geekflarelab in my case) is selected

On frontend configuration tab, enter the name and create new IP address.

- It will popup to reserve a new static IP address. Give the name there and click reserve.
- Click Done

Note: its recommended to get the static IP for your LB so you can use that to point to your domain.

On the review tab, ensure all good and once happy click Create

It will take a few seconds to create, and you should see them on the list.

When you click on the newly created LB, it will show the details like below.

As you can see both instances are healthy and LB is technically operational.

**Next**, you got to update your domain A record to point to the LB frontend IP. Once done, when you hit your domain, it should hit to LB and distribute traffic to the instances.

I did some load test and here is the result.

You can see the traffic from Europe is getting routed to the nearest server located in London, and North America and Asia traffic are to US server. The

good thing is you don't have to configure anything for geo traffic routing, its default feature. The above monitoring is available under backends tab.

## Conclusion

Creating an LB is easy, and I hope this gives you an idea about it. There is a lot of configuration you can do to meet your application requirement like session affinity, CDN integration, SSL cert, etc. If you are exploring an option to have a load balancer for your application, then play around and see how it helps.

Costing is based on usage so there is no monthly or annual locking. I think the minimal usage would cost around $18 per month. If you are curious to learn about Google Cloud administration, then you may consider taking this online course.

TAGS: GCP

## More Great Reading on Geekflare

→

Script to Monitor Google Cloud Unused External IP

→

How to Implement Google Managed Certificate on Cloud Load Balancer?

→

How to Find External IP of Google

→

How to Implement Google Cloud

Cloud VM?

CDN?

→

How to Configure Firewall Rules in Google Cloud Platform?

→

How to Configure Site Uptime Monitoring with Google Stackdriver?

# Power Your Business

Choosing the right product and service is essential to run an online business. Here are some of the tools and services to help your business grow.

## Netsparker

Netsparker uses the Proof-Based Scanning™ to automatically verify the identified vulnerabilities with proof of exploit, thus making it possible to scan thousands of web applications and generate actionable results within just hours.

Try Netsparker

## Kinsta

Probably the best managed WordPress cloud platform to host small to enterprise sites. Kinsta leverages Google's low latency network infrastructure to deliver content faster. Free SSL, CDN, backup and a lot more with outstanding support. You'll love it.

Try Kinsta

## Sucuri

A global CDN and cloud-based web application firewall for your website to supercharge the performance and secure from online threats. SUCURI WAF protects from OWASP top 10 vulnerabilities, brute force, DDoS, malware, and more.

Try Sucuri

+65 More Awesome Resources

Explore site tools

Browse articles

Discover trusted resources