Creating a Virtual Machine

Overview

Google Compute Engine lets you create virtual machines running different operating systems, including multiple flavors of Linux (Debian, Ubuntu, Suse, Red Hat, CoreOS) and Windows Server, on Google infrastructure. You can run thousands of virtual CPUs on a system that has been designed to be fast and to offer strong consistency of performance.

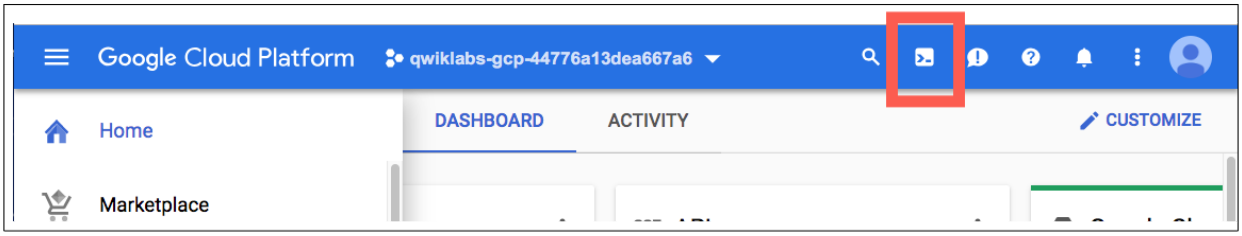
What you'll do

- Create a virtual machine with the GCP Console
- Create a virtual machine with `gcloud` command line
- Deploy a web server and connect it to a virtual machine

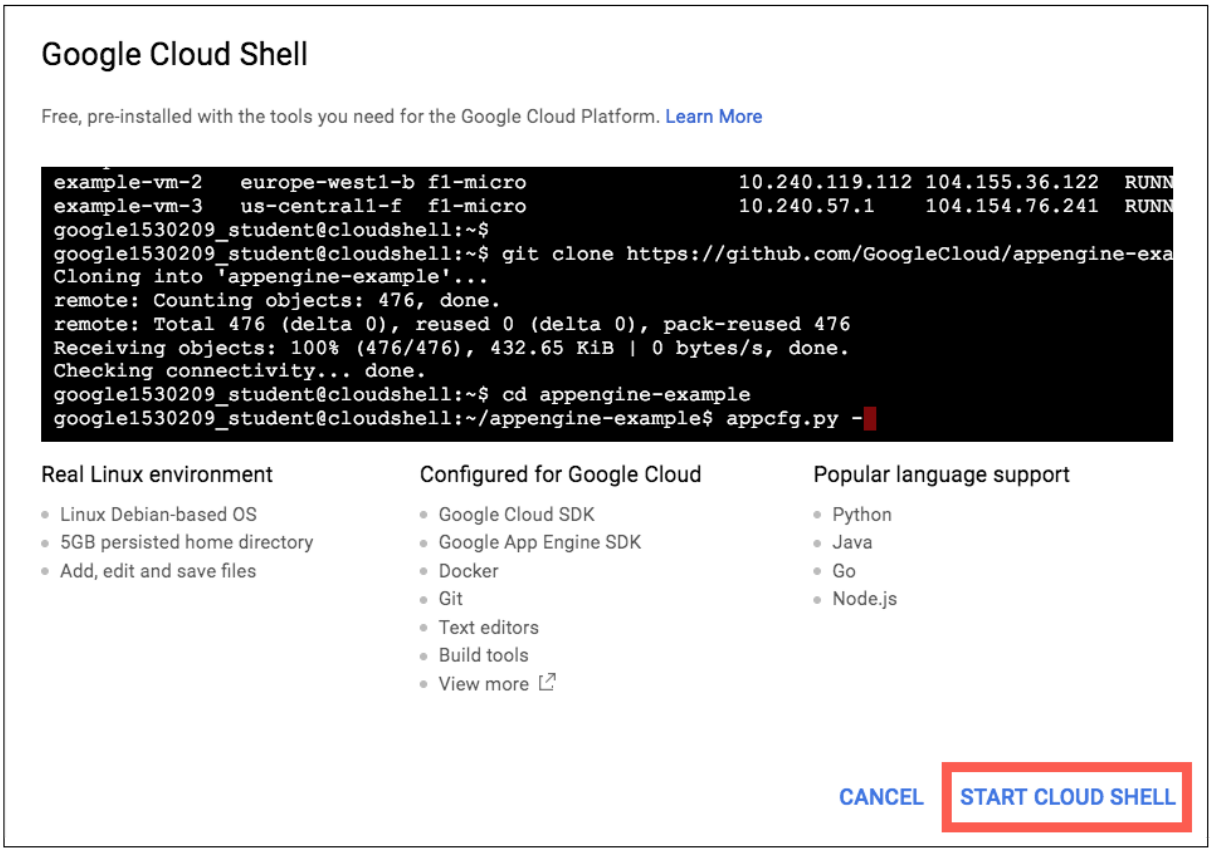
Activate Google Cloud Shell

Google Cloud Shell is a virtual machine that is loaded with development tools. It offers a persistent 5GB home directory and runs on the Google Cloud. Google Cloud Shell provides command-line access to your GCP resources.

1. In GCP console, on the top right toolbar, click the Open Cloud Shell button.

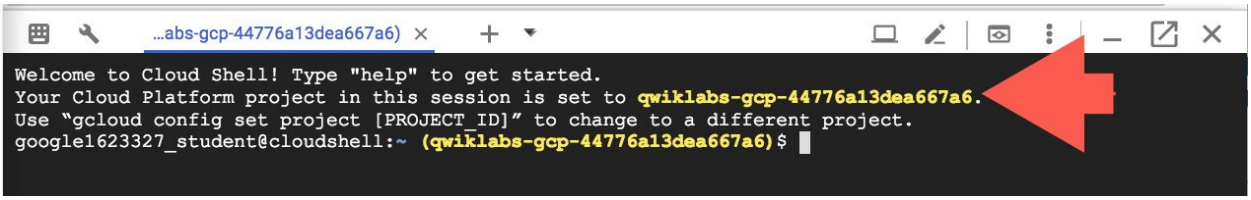


2. In the dialog box that opens, click **START CLOUD SHELL**:



You can click "START CLOUD SHELL" immediately when the dialog box opens.

It takes a few moments to provision and connect to the environment. When you are connected, you are already authenticated, and the project is set to your *PROJECT_ID*. For example:



gcloud is the command-line tool for Google Cloud Platform. It comes pre-installed on Cloud Shell and supports tab-completion.

You can list the active account name with this command:

```
gcloud auth list
```

Output:

```
Credentialed accounts:
 - <myaccount>@<mydomain>.com (active)
```

Example output:

```
Credentialed accounts:
 - google1623327_student@qwiklabs.net
```

You can list the project ID with this command:

```
gcloud config list project
```

Output:

```
[core]
project = <project_ID>
```

Example output:

```
[core]
project = qwiklabs-gcp-44776a13dea667a6
```

Full documentation of **gcloud** is available on Google Cloud gcloud Overview.

Understanding Regions and Zones

Certain Compute Engine resources live in regions or zones. A region is a specific geographical location where you can run your resources. Each region has one or more zones. For example, the us-central1 region denotes a region in the Central United States that has zones `us-central1-a`, `us-central1-b`, `us-central1-c`, and `us-central1-f`.

Resources that live in a zone are referred to as zonal resources. Virtual machine Instances and persistent disks live in a zone. To attach a persistent disk to a

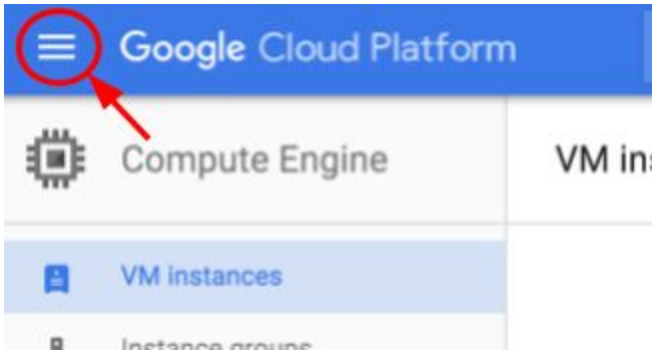
virtual machine instance, both resources must be in the same zone. Similarly, if you want to assign a static IP address to an instance, the instance must be in the same region as the static IP.

Learn more about regions and zones and see a complete list in Regions & Zones documentation.

Create a new instance from the Cloud Console

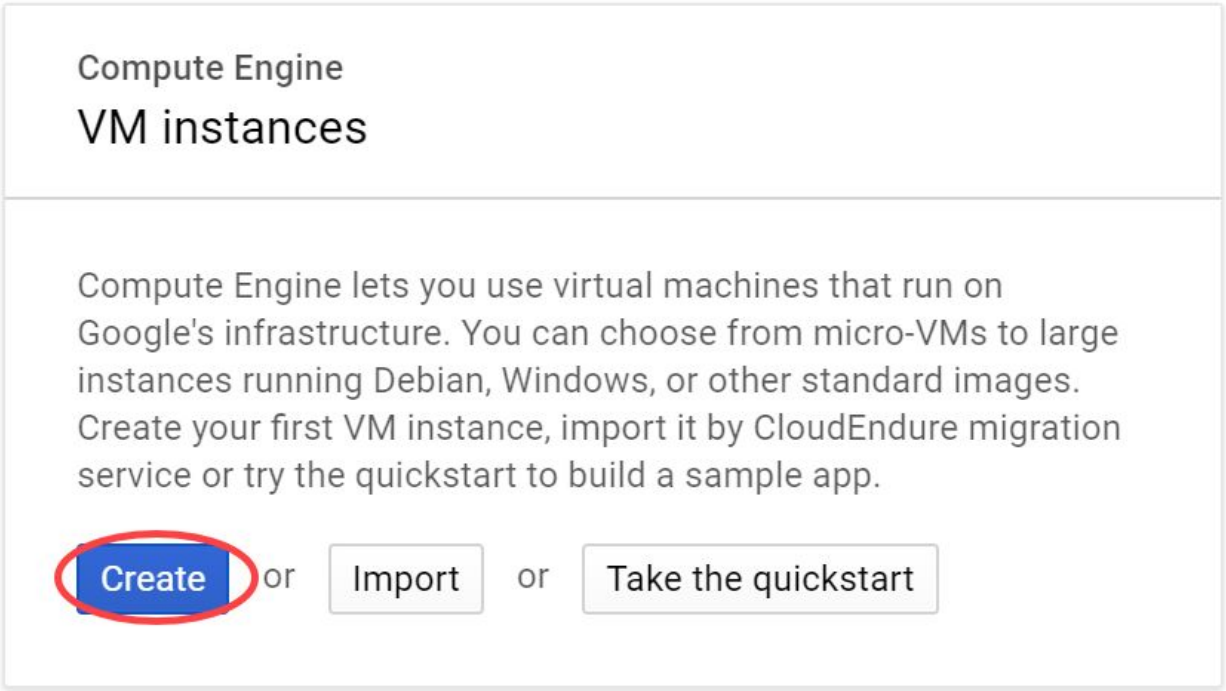
In this section, you'll learn how to create new pre-defined machine types with Google Compute Engine from the Cloud Console.

In the GCP Console, on the top left of the screen, select Navigation menu > **Compute Engine** > **VM Instances**:



This may take a minute to initialize for the first time.

To create a new instance, click **Create**.



There are many parameters you can configure when creating a new instance. Use the following for this lab:

Field

There are a number of images to choose from, including: Debian, Ubuntu, CoreOS as well as premium images such as Red Hat Enterprise Linux and Windows Server. See Operating System documentation for more detail.

Click **Create**.

Wait for it to finish - it shouldn't take more than a minute.

Once finished, you should see the new virtual machine in the **VM Instances** page.

To SSH into the virtual machine, click on **SSH** on the right hand side. This launches a SSH client directly from your browser.

Note: For more information, see the Connect to an instance using ssh documentation.

Install a NGINX web server

Now you'll install NGINX web server, one of the most popular web servers in the world, to connect your virtual machine to something.

Once SSH'ed, get `root` access using `sudo`:

```
sudo su -
```

As the `root` user, update your OS:

```
apt-get update
```

(Output)

```
Get:1 http://security.debian.org stretch/updates InRelease [94.3 kB]
Ign http://deb.debian.org strech InRelease
Get:2 http://deb.debian.org strech-updates InRelease [91.0 kB]
...
```

Install NGINX:

```
apt-get install nginx -y
```

(Output)

```
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
...
```

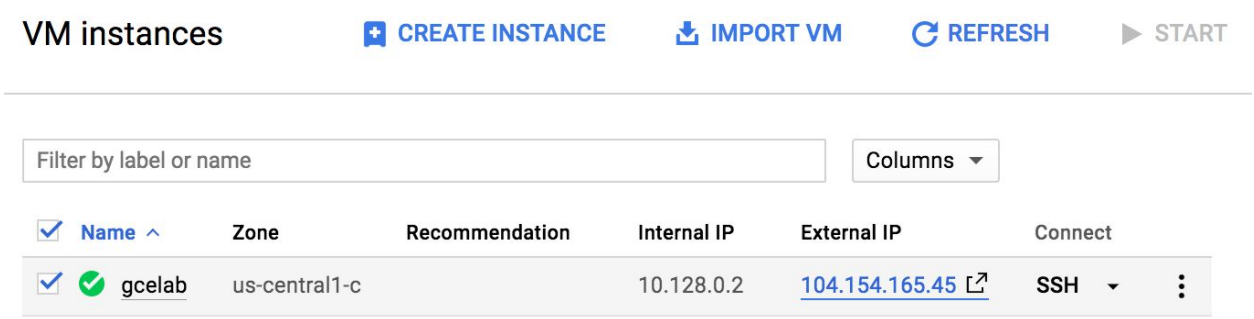
Check that NGINX is running:

```
ps auwx | grep nginx
```

(Output)

```
root      2330  0.0  0.0 159532  1628 ?        Ss   14:06   0:00 nginx: master
process /usr/sbin/nginx -g daemon on; master_process on;
www-data  2331  0.0  0.0 159864  3204 ?        S    14:06   0:00 nginx: worker
process
www-data  2332  0.0  0.0 159864  3204 ?        S    14:06   0:00 nginx: worker
process
root      2342  0.0  0.0  12780   988 pts/0    S+   14:07   0:00 grep nginx
```

Awesome! To see the web page, go to the Cloud Console and click the `External IP`link of the virtual machine instance. You can also see the web page by adding the `External IP` to `http://EXTERNAL_IP/` in a new browser window or tab.



You should see this default web page:

Create a new instance with gcloud

Rather than using the GCP Console to create a virtual machine instance, you can use the command line tool `gcloud`, which is pre-installed in Google Cloud Shell. Cloud Shell is a Debian-based virtual machine loaded with all the development

tools you'll need (`gcloud`, `git`, and others) and offers a persistent 5GB home directory.

If you want to try this on your own machine in the future, read the gcloud command line tool guide.

In the Cloud Shell, create a new virtual machine instance from the command line using `gcloud`, replacing [YOUR_ZONE] with one of the zone choices given earlier:

```
gcloud compute instances create gcelab2 --machine-type n1-standard-2 --zone [your_zone]
```

(Output)

```
Created [...gcelab2].
NAME     ZONE           MACHINE_TYPE  ...    STATUS
gcelab2  us-central1-c  n1-standard-2 ...    RUNNING
```

The instance created has these default values:

- The latest Debian 9 (stretch) image.
- The `n1-standard-2` machine type. In this lab you can select one of these other machine types if you'd like: `n1-highmem-4` or `n1-highcpu-4`. When you're working on a project outside of Qwiklabs, you can also specify a custom machine type.
- A root persistent disk with the same name as the instance; the disk is automatically attached to the instance.

Run `gcloud compute instances create --help` to see all the defaults.

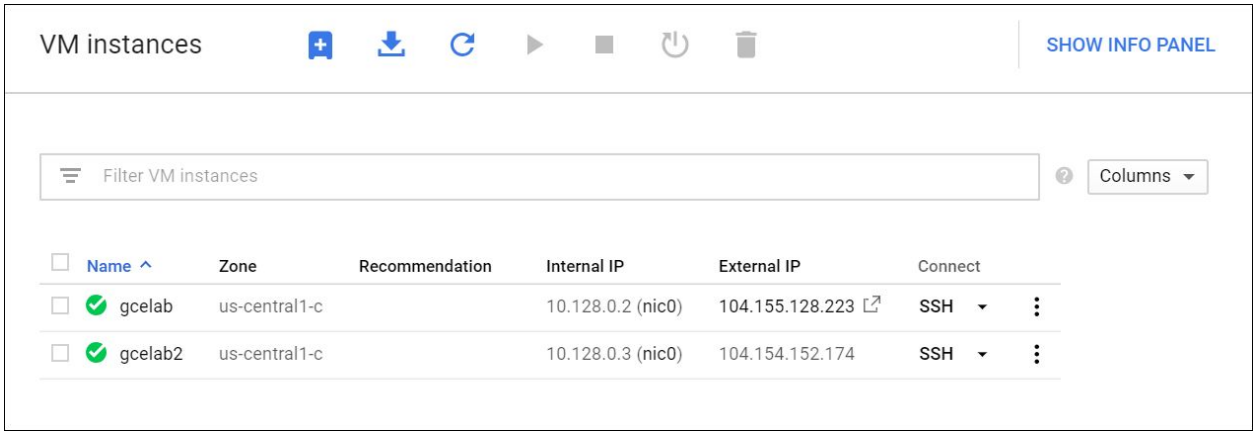
Note: You can set the default region and zones that `gcloud`uses if you are always working within one region/zone and you don't want to append the `--zone`flag every time. Do this by running these commands :

```
gcloud config set compute/zone ...
```

```
gcloud config set compute/region ...
```

To exit `help`, press **Ctrl+c**.

Check out your instances. Select **Navigation menu** > **Compute Engine** > **VM instances**. You should see the 2 instances you created in this lab.



Finally, you can SSH into your instance using `gcloud` as well. Make sure you add your zone, or omit the `--zone` flag if you've set the option globally:

```
gcloud compute ssh gcelab2 --zone [YOUR_ZONE]
```

(Output)

```
WARNING: The public SSH key file for gcloud does not exist.
WARNING: The private SSH key file for gcloud does not exist.
```

```
WARNING: You do not have an SSH key for gcloud.
WARNING: [/usr/bin/ssh-keygen] will be executed to generate a key.
This tool needs to create the directory
[/home/gcpstaging306_student/.ssh] before being able to generate SSH
Keys.
```

Now you'll type **Y** to continue.

```
Do you want to continue? (Y/n)
```

Enter through the passphrase section to leave the passphrase empty.

```
Generating public/private rsa key pair.
Enter passphrase (empty for no passphrase)
```

After connecting, you disconnect from SSH by exiting from the remote shell:

```
exit
```