Tutorial Google Cloud Platform

Arthur Pesah, Sebastian Bujwid, Yonk Shi

March 25, 2019

Contents

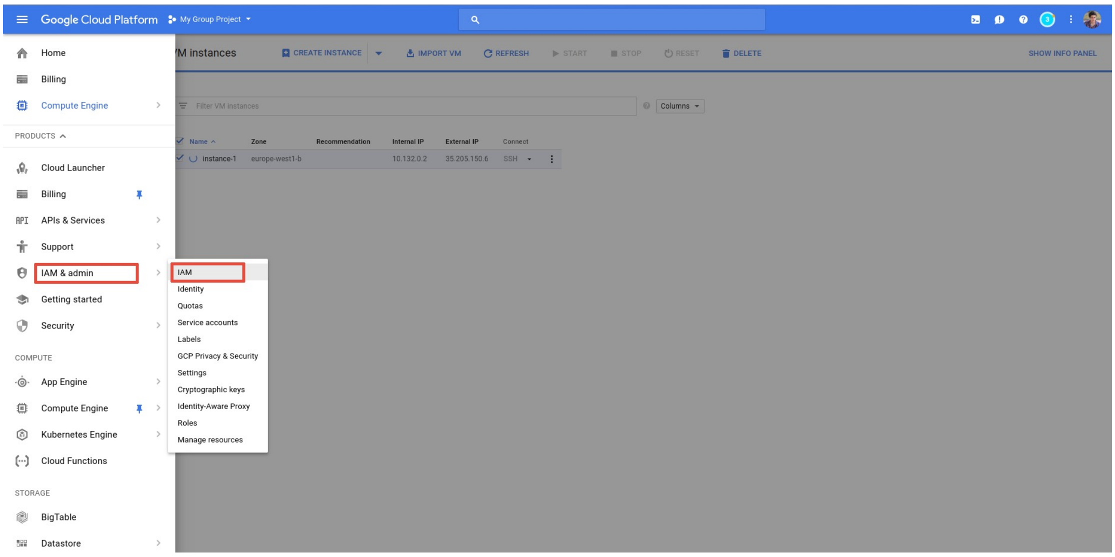
1 Introduction

The main option to have access to a GPU for your deep learning project is to use Google Cloud Platform (GCP). We've obtained some GCP educational grant for this course, each student's email is entitled $50 for the course (If your team has three members, you will have a combined $150 credits - which can correspond to around 360 hours of using NVIDIA K80!). To claim your credit, simply fill out this request form with your KTH email address (since we have a limited number of coupons, please do not redistribute this link!). Follow the instructions in the email, once you have setup the coupon, you will see a billing account named **DD2424**. Make sure to use this DD2424 billing account when you create a new project / applying billing to an existing project.

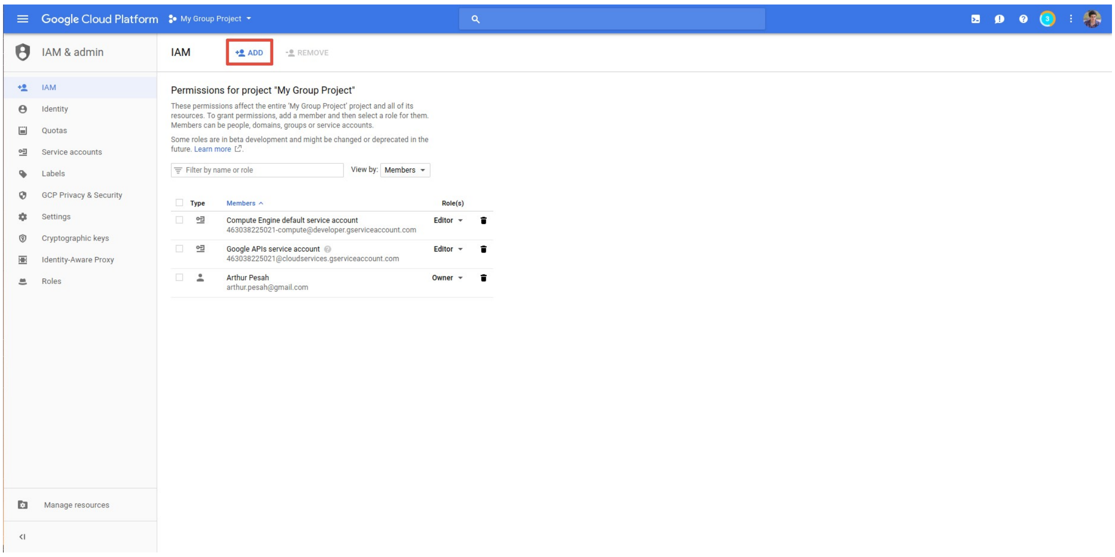
Google also allows everyone to get a free trial account with $300 of credits. If you run out of credits with us, don't hesitate to start using this trial.

2 How to add new members to your project

You might want other members of your team to be able to access to GCP. In that case, it's pretty straight-forward to add new members. Go to *IAM & admin / IAM*:



Then click on *Add*:



Permissions for project "My Group Project"

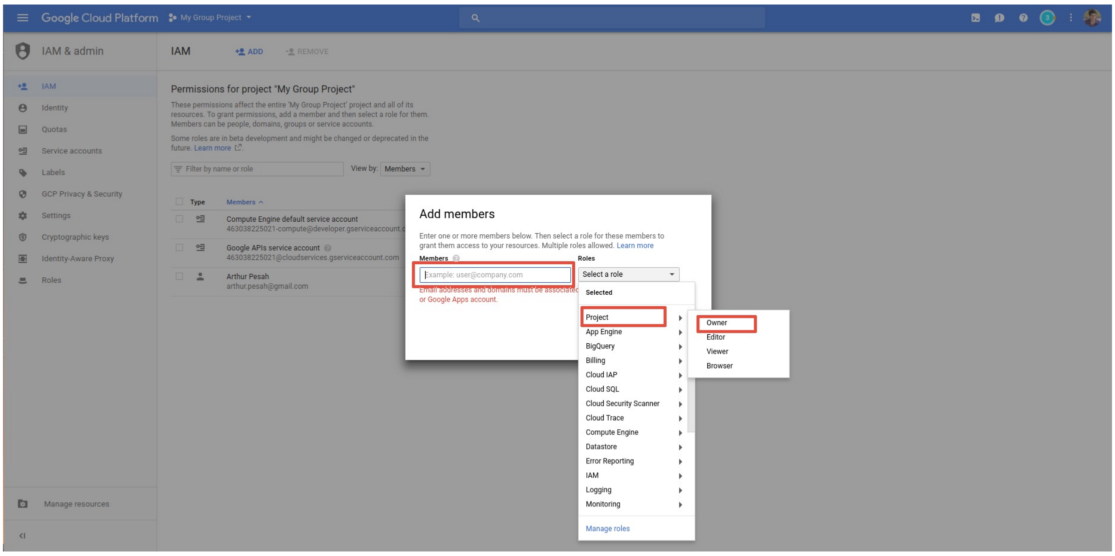
These permissions affect the entire 'My Group Project' project and all of its resources. To grant permissions, add a member and then select a role for them. Members can be people, domains, groups or service accounts.

Some roles are in beta development and might be changed or deprecated in the future. Learn more

Filter by name or role   View by: Members

□	Type	Members	Role(s)
□		Compute Engine default service account 463038225021-compute@developer.gserviceaccount.com	Editor
□		Google APIs service account 463038225021@cloudservices.gserviceaccount.com	Editor
□		Arthur Pesah arthur.pesah@gmail.com	Owner

Add the google address of your other team members and select *Project/Owner* as a role:



Add members

Enter one or more members below. Then select a role for these members to grant them access to your resources. Multiple roles allowed. Learn more

Members: Example: user@company.com

Email addresses and domains must be associated ... or Google Apps account.

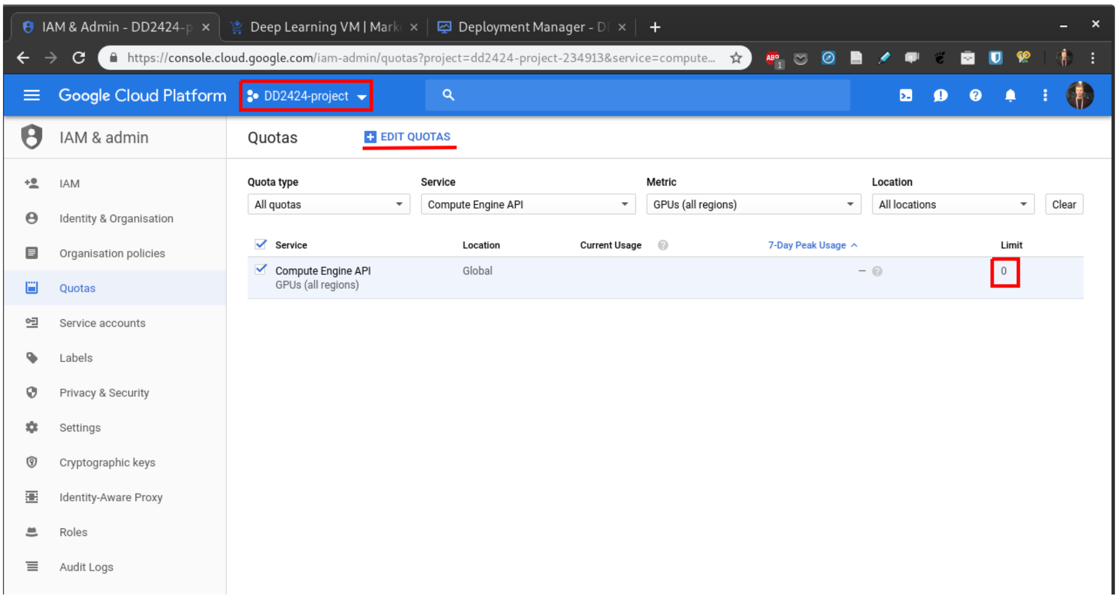
Roles: Select a role

3 GPU Quota & Setup

If you are planning on using GPU resources, you will have to set GPU quota for your project. It is likely that the GPU quota for your project is 0 by default. You can check that in *IAM & admin* → *Quotas*, under service called *Compute Engine API - GPUs (all regions)*. If it says 0 you need to increase it to be able to deploy GPU machines. You can do that by selecting the service and clicking *EDIT QUOTAS*.

When creating the virtual machine/choosing quota, we recommend **K80 GPU** over expensive GPUs (V100, P100 etc). K80s are a fraction of the cost comparing to other GPUs, and their performances are comparable. In our experience, two K80s will outperform a single V100, yet still cheaper.

Note: do not apply for higher quota than you really need. The higher the quota you apply for, the more time it can take for Google to process the request. Requesting the limit of 1 should be approved very quickly.



Quotas

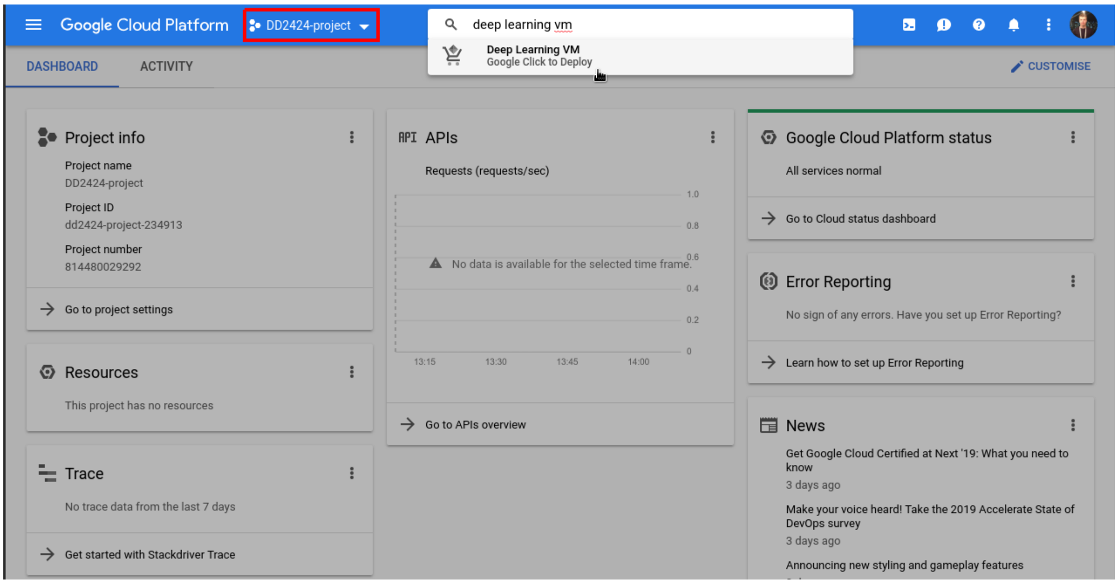
Quota type: All quotas
Service: Compute Engine API
Metric: GPUs (all regions)
Location: All locations

☑ Service	Location	Current Usage	7-Day Peak Usage	Limit
☑ Compute Engine API GPUs (all regions)	Global		–	0

4 How to create a Deep Learning virtual machine

The most convenient way to do use GCP resources for Deep Learning is to use existing Deep Learning image templates. They are easy and quick to setup and come with software packages such as Tensorflow or PyTorch and have pre-installed NVIDIA drivers.

To deploy a new *Deep Learning VM* go to the GCP Console page and search for *Deep Learning VM* in the search bar at the top of the page.



Google Cloud Platform DD2424-project

deep learning vm

Deep Learning VM
Google Click to Deploy

DASHBOARD ACTIVITY CUSTOMISE

Project info

Project name
DD2424-project

Project ID
dd2424-project-234913

Project number
814480029292

Go to project settings

Resources

This project has no resources

Trace

No trace data from the last 7 days

Get started with Stackdriver Trace

APIs

Requests (requests/sec)

No data is available for the selected time frame.

Go to APIs overview

Google Cloud Platform status

All services normal

Go to Cloud status dashboard

Error Reporting

No sign of any errors. Have you set up Error Reporting?

Learn how to set up Error Reporting

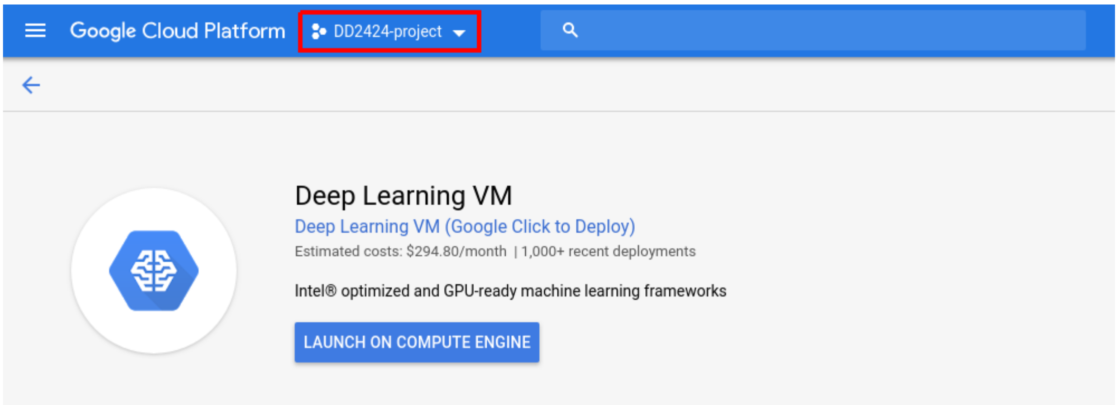
News

Get Google Cloud Certified at Next '19: What you need to know
3 days ago

Make your voice heard! Take the 2019 Accelerate State of DevOps survey
3 days ago

Announcing new styling and gameplay features

When selecting the image **make sure that you have the right project active** - you can change the project next to GCP logo on top left corner.



Google Cloud Platform DD2424-project

Deep Learning VM

Deep Learning VM (Google Click to Deploy)

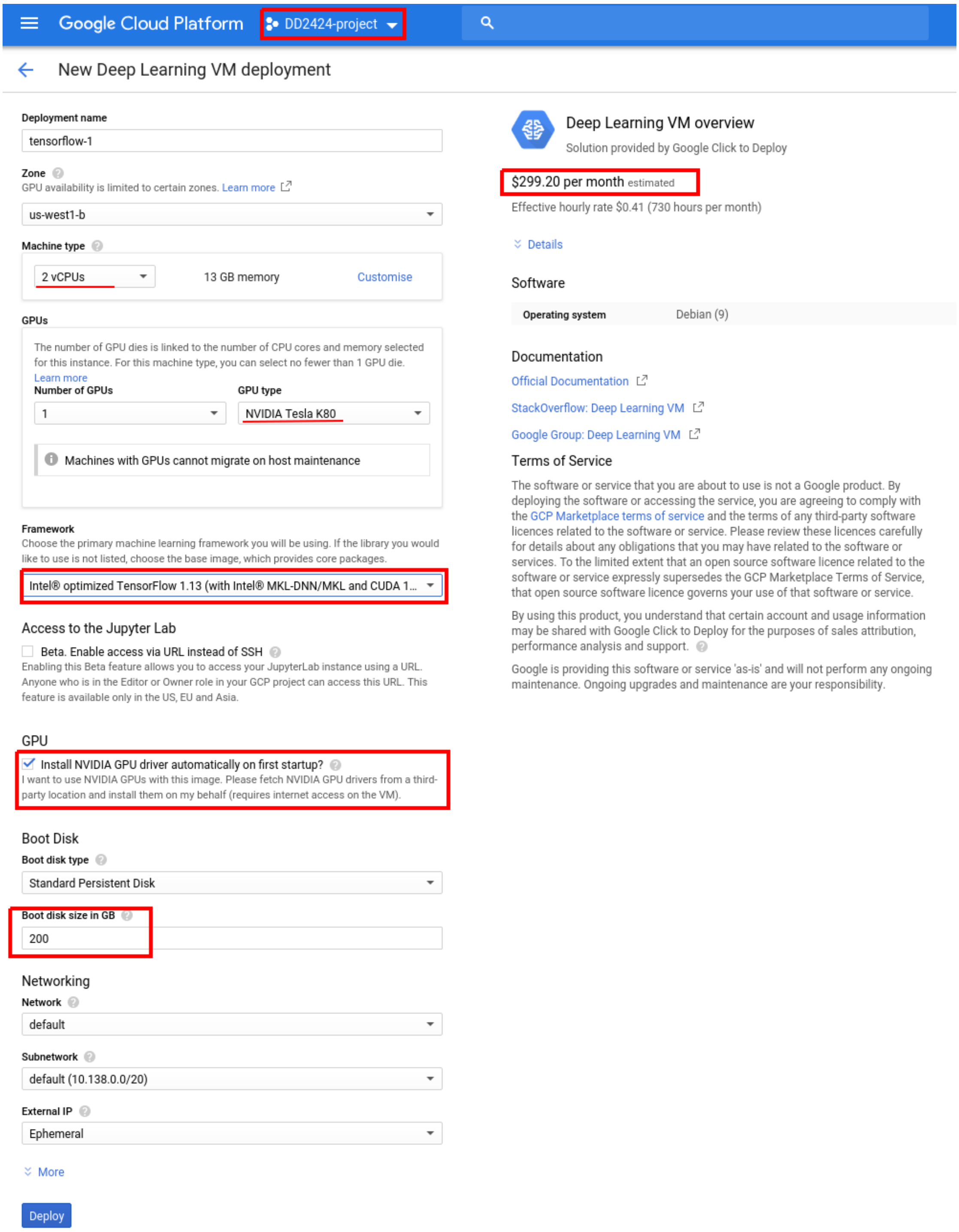
Estimated costs: $294.80/month | 1,000+ recent deployments

Intel® optimized and GPU-ready machine learning frameworks

LAUNCH ON COMPUTE ENGINE

After clicking *Launch on compute engine* you will be moved to the settings of the VM. You can choose the size of the machine and the disk space according to your needs (keep in mind that it affects the price a little bit). If you're using large datasets that require a parallel reader and fast pre-fetching, you might also want to choose a better CPU (with typically more cores). Select the GPU you need (we advise to use K80) and the framework you intend to use, e.g *Intel optimized TensorFlow* or *PyTorch*. **Make sure that you select the option to install NVIDIA GPU driver**.

Optionally, use can change the zone to *europe-west1-b* or *europe-west1-d* - since the servers are closer the connection should have smaller delays.



New Deep Learning VM deployment

Deployment name
tensorflow-1

Zone
GPU availability is limited to certain zones. Learn more
us-west1-b

Machine type
2 vCPUs 13 GB memory Customise

GPUs
The number of GPU dies is linked to the number of CPU cores and memory selected for this instance. For this machine type, you can select no fewer than 1 GPU die.
Learn more

Number of GPUs: 1
GPU type: NVIDIA Tesla K80

Machines with GPUs cannot migrate on host maintenance

Framework
Choose the primary machine learning framework you will be using. If the library you would like to use is not listed, choose the base image, which provides core packages.
Intel® optimized TensorFlow 1.13 (with Intel® MKL-DNN/MKL and CUDA 1...

Access to the Jupyter Lab
□ Beta. Enable access via URL instead of SSH
Enabling this Beta feature allows you to access your JupyterLab instance using a URL. Anyone who is in the Editor or Owner role in your GCP project can access this URL. This feature is available only in the US, EU and Asia.

GPU
☑ Install NVIDIA GPU driver automatically on first startup?
I want to use NVIDIA GPUs with this image. Please fetch NVIDIA GPU drivers from a third-party location and install them on my behalf (requires internet access on the VM).

Boot Disk
Boot disk type
Standard Persistent Disk

Boot disk size in GB
200

Networking
Network
default

Subnetwork
default (10.138.0.0/20)

External IP
Ephemeral

More

Deploy

Deep Learning VM overview
Solution provided by Google Click to Deploy

$299.20 per month estimated
Effective hourly rate $0.41 (730 hours per month)

Details

Software

Operating system	Debian (9)

Documentation
Official Documentation
StackOverflow: Deep Learning VM
Google Group: Deep Learning VM

Terms of Service
The software or service that you are about to use is not a Google product. By deploying the software or accessing the service, you are agreeing to comply with the GCP Marketplace terms of service and the terms of any third-party software licences related to the software or service. Please review these licences carefully for details about any obligations that you may have related to the software or services. To the limited extent that an open source software licence related to the software or service expressly supersedes the GCP Marketplace Terms of Service, that open source software licence governs your use of that software or service.

By using this product, you understand that certain account and usage information may be shared with Google Click to Deploy for the purposes of sales attribution, performance analysis and support.

Google is providing this software or service 'as-is' and will not perform any ongoing maintenance. Ongoing upgrades and maintenance are your responsibility.

After you click *Deploy*, each hour is counted on your budget, so don't forget to **STOP THE VM WHEN YOU'RE NOT USING IT** (instructions at the end).

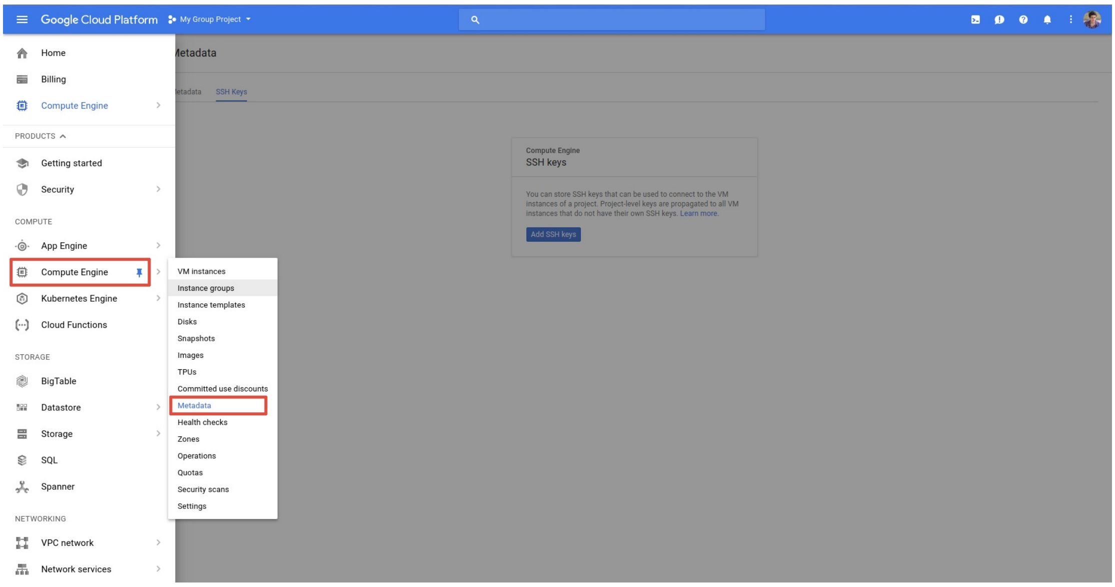
The next step to be able to connect is to handle your ssh keys. The instructions depend on your OS and are well explained in the GCP documentation. connect is to handle your ssh keys. The instructions depend on your OS and are well explained on the GCP documentation. You can create as many ssh keys as you want, so don't hesitate to repeat the instructions for everyone in your group connecting from his own computer. Here is a summary for Linux users:

```
$ ssh-keygen -t rsa -f ~/.ssh/gcp-key -C [GOOGLE EMAIL ADDRESS]
```

You can use the default parameters for the creation of the key. Then:

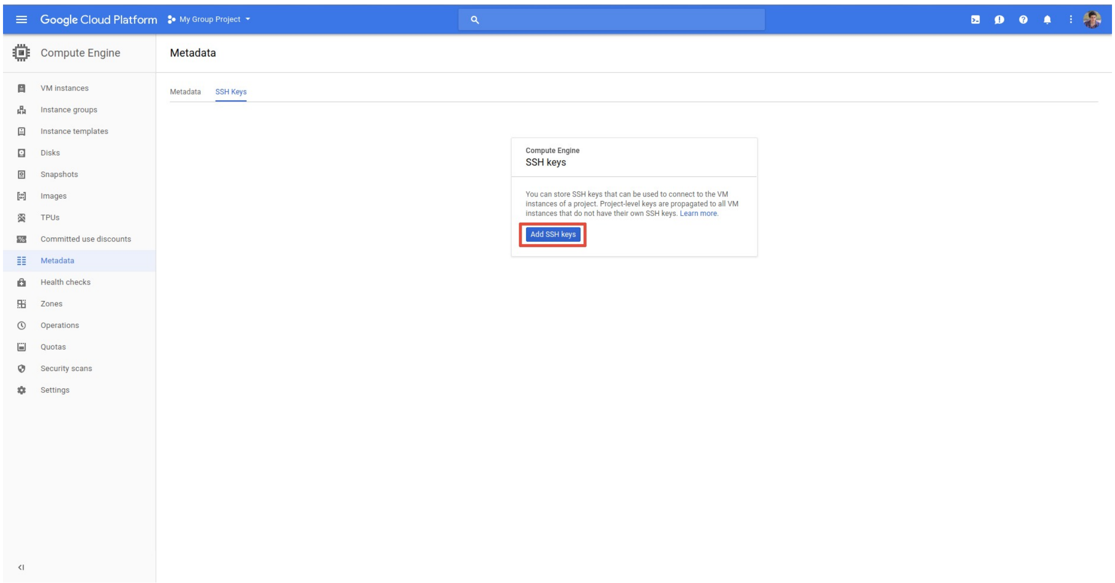
```
$ cat ~/.ssh/gcp-key.pub
```

Copy the entire output (including your email address that should appear at the end of the key) and go back to GCP. Click on *Compute Engine / Metadata*:



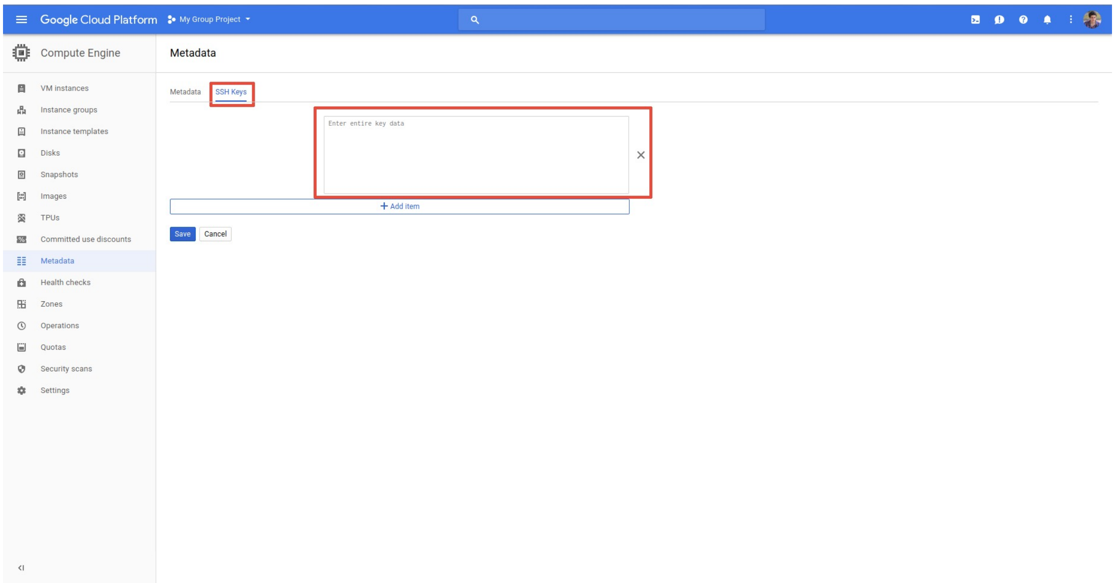
You can store SSH keys that can be used to connect to the VM instances of a project. Project-level keys are propagated to all VM instances that do not have their own SSH keys. Learn more.

Then click on *SSH Keys / Add SSH keys*:

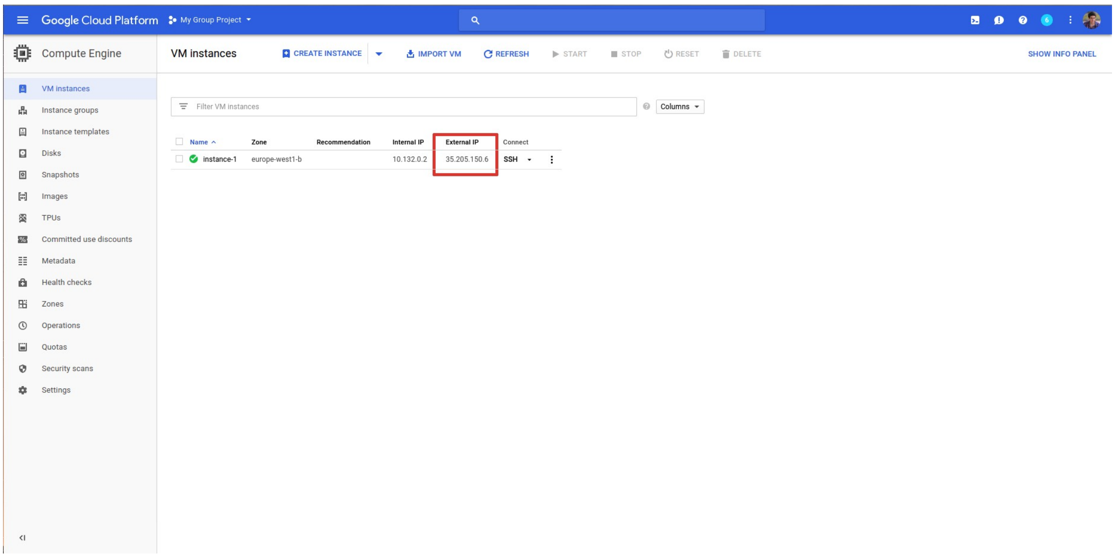


You can store SSH keys that can be used to connect to the VM instances of a project. Project-level keys are propagated to all VM instances that do not have their own SSH keys. Learn more.

And paste your key in the form. Delete any extra space or newline, the key should end at the end of the email address. Then save:



You're now able to connect to your VM. One way to do it is to use the SSH button on the right (it will open a new window with a web terminal). But if you prefer connecting from your own terminal, just copy the external IP address like below:



□	Name	Zone	Recommendation	Internal IP	External IP	Connect
□	✓ instance-1	europe-west1-b		10.132.0.2	35.205.150.6	SSH

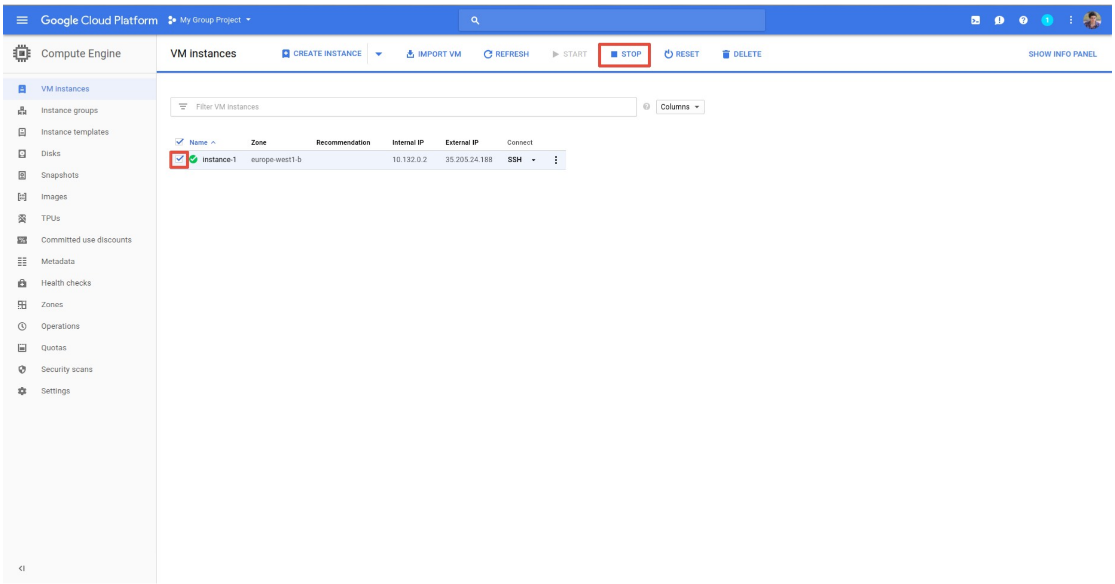
And connect to your machine using a command like that:

```
$ ssh -i ~/.ssh/gcp-key [google username]@[external ip address]
```

You can also copy your project to the machine using scp:

```
$ scp -i ~/.ssh/gcp-key -r [projectFolder] [google username]@[external ip address]:~
```

To stop the machine when you are not using it, just select it and click on STOP:



☑	Name	Zone	Recommendation	Internal IP	External IP	Connect
☑	✓ instance-1	europe-west1-b		10.132.0.2	35.205.24.188	SSH

5 Optional - installing specific versions of packages

5.1 Using conda

In case you need very specific versions of Python deep learning packages the easiest way to get them is through Miniconda. *conda* installs necessary dependencies for you (such as cudnn, cuda-toolkit, etc.).

Note: Packages installed with *conda* can be faster than packages installed with *pip*, because of more compilation flags enabled (support for Intel MKL-DNN, etc.).

If the command *conda* is not available you can install it by executing:

```
$ wget https://repo.anaconda.com/miniconda/Miniconda3-latest-
   Linux-x86_64.sh
$ bash Miniconda3-latest-Linux-x86_64.sh}
```

During the installation process confirm to initialize Miniconda:

```
Do you wish the installer to initialize Miniconda3
in your /home/sebastian/.bashrc ? [yes|no]
[no] >>> yes
```

To be able to access *conda* you need to log out and log in to the machine again through ssh. After that you should be able to use *conda* to create new environments:

```
# e.g. installing TensorFlow 1.10.0
$ conda create -n myenv1 tensorflow-gpu=1.10.0
$ conda activate myenv1
$ python -c ”import␣tensorflow␣as␣tf;␣print(tf.__version__)”
  1.10.0

# e.g. installing Pytorch 0.3.1
$ conda create -n myenv2 pytorch=0.3.1
$ conda activate myenv2
$ python -c ”import␣torch;␣print(torch.__version__)”
  0.3.1

# e.g. installing Caffe
$ conda create -n myenv3
$ conda activate myenv3
$ conda install caffe-gpu
$ caffe --version
caffe version 1.0.0
```

```
    Debug build (NDEBUG not #defined)
```

To install more packages you can use *conda install* or *pip install*.

5.2 Using GCP Marketplace image templates

An alternative approach for obtaining more specific software is to use pre-defined image templates from GCP Marketplace. Use the search bar to find software packages that you are interested in. For example searching for *Caffe* you can find image *Caffe Python 3.6 NVidia GPU Production* by *Jetware*. If you find an image template that meets your particular requirements click *Launch on compute engine* and follow relevant instructions from Section 4.

Note: Using some image templates might not be free, make sure to check the pricing when launching.