Tutorial Google Cloud Platform

DD2424 Deep Learning in Data Science

Arthur Pesah

Contents

1 Introduction

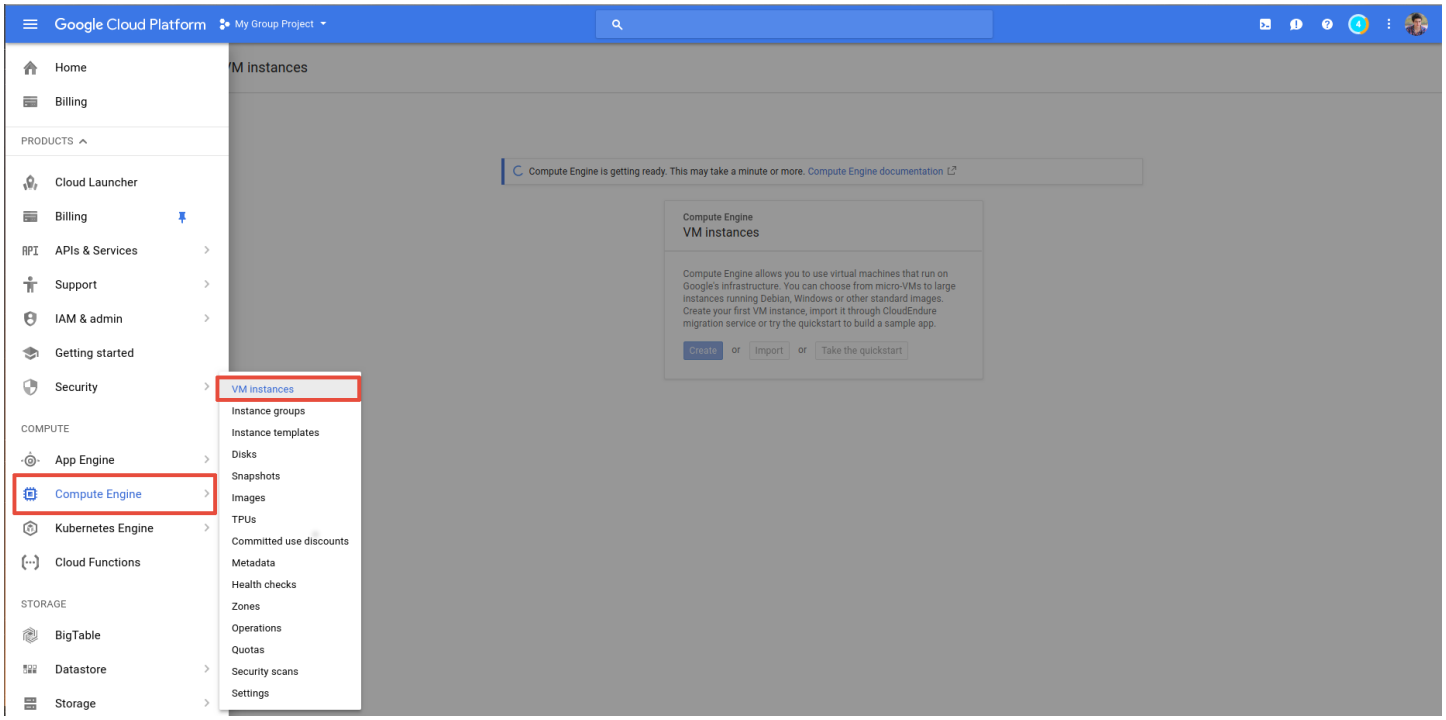
Apart from PDC and Colab, another option to have access to a GPU for your deep learning project is to use Google Cloud Platform (GCP). Indeed, we've obtained 5000$ of GCP educational grant for this course, so your group will be able to use up to around 80$ (which can correspond to around 200 hours!). In order to get access to it, just fill the request form with your group name and the Google address of one of your team members and we'll take care of creating a GCP project for you. You will then receive an email with the project and can start using GCP.

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.

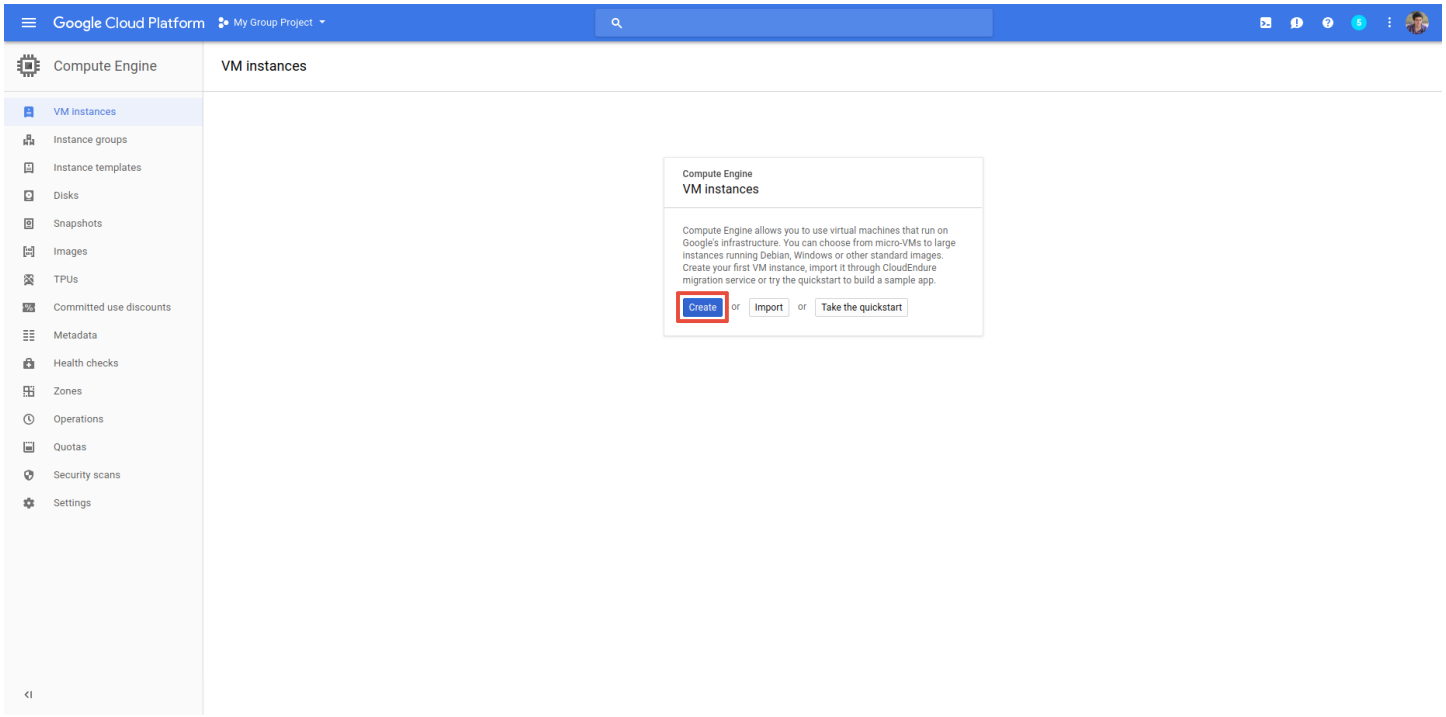
2 How to use Google Cloud Platform

Once you've obtained access to your GCP project, you're free to use your 80$ of credits however you want on the platform. Nevertheless, we have some advice on how to optimize this budget in the context of the course.

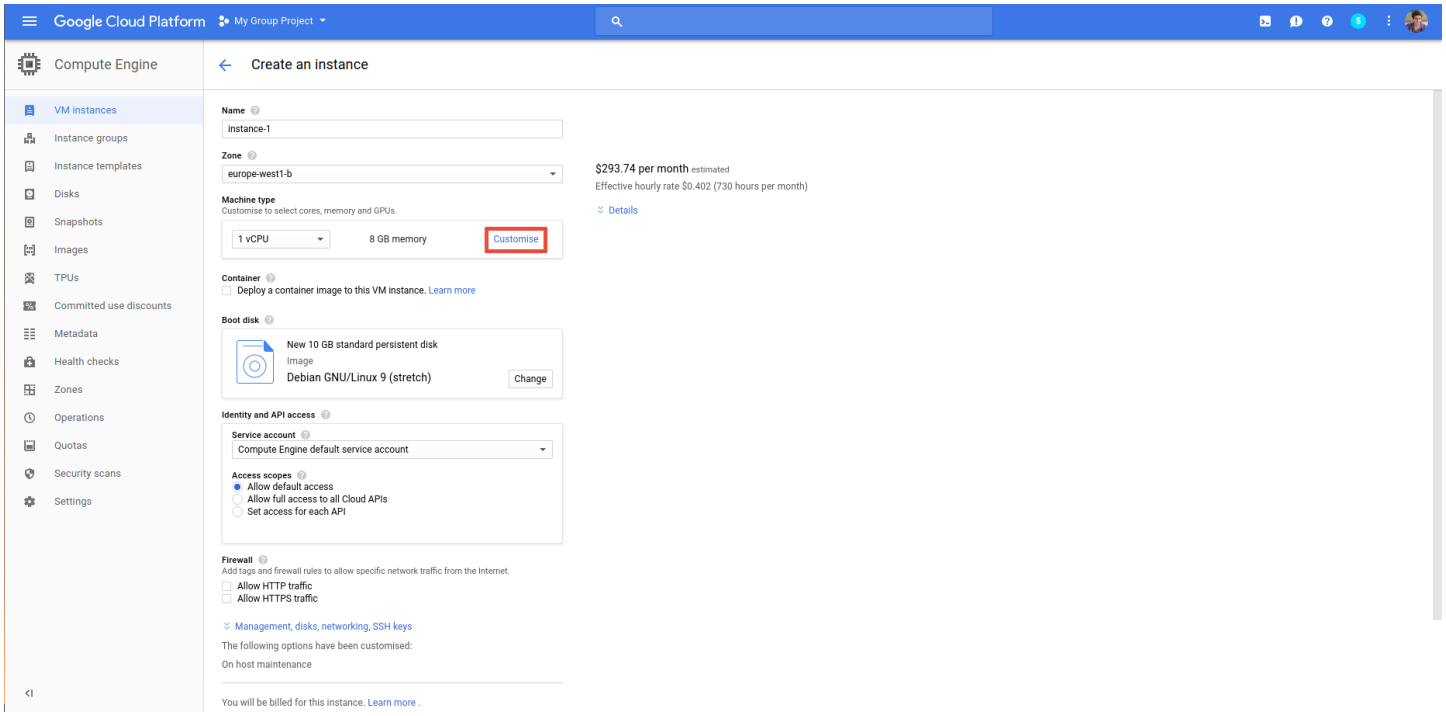
We propose you to create a Virtual Machine (VM) instance and to run your code on it (through ssh). For that, go to the GCP website and select your project on the top bar. Then, click on *Compute Engine* and *VM instances*:



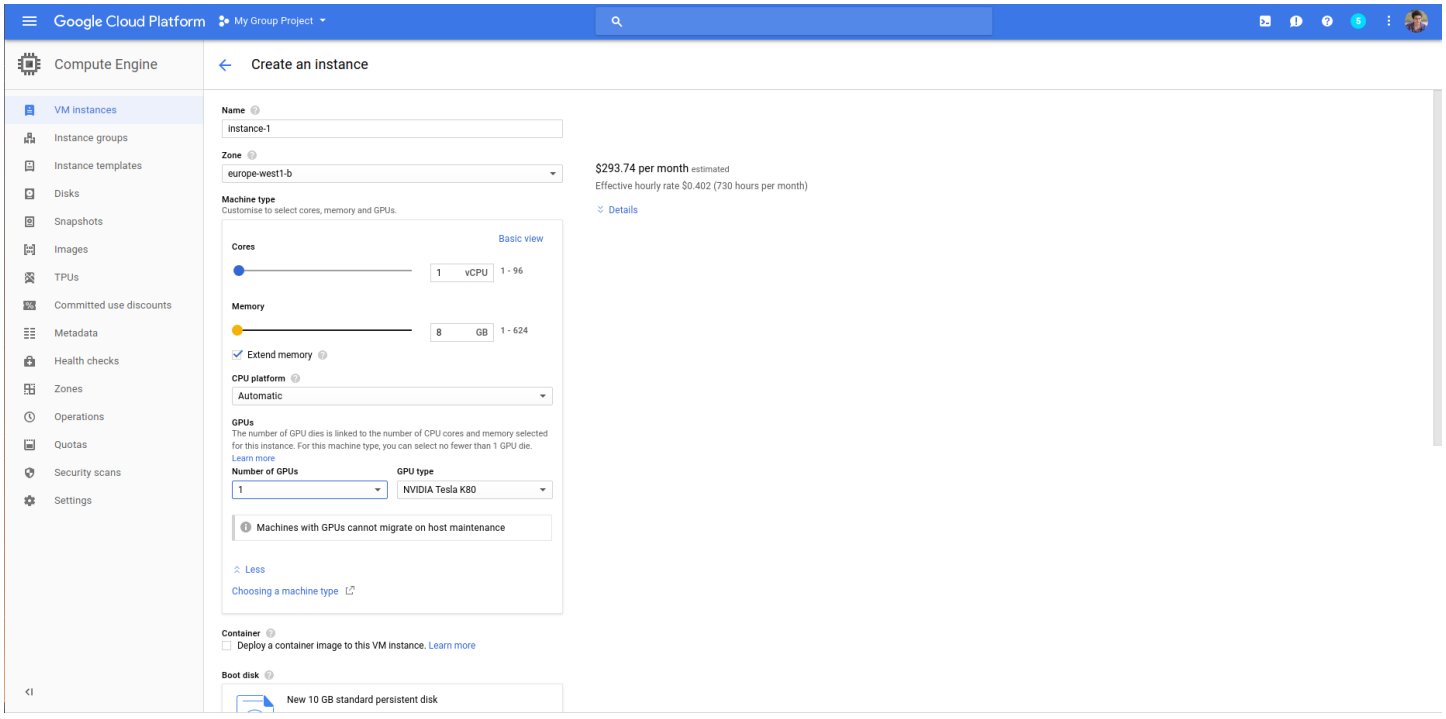
It might take a while to configure the Compute Engine. Once it's done, create a VM instance:



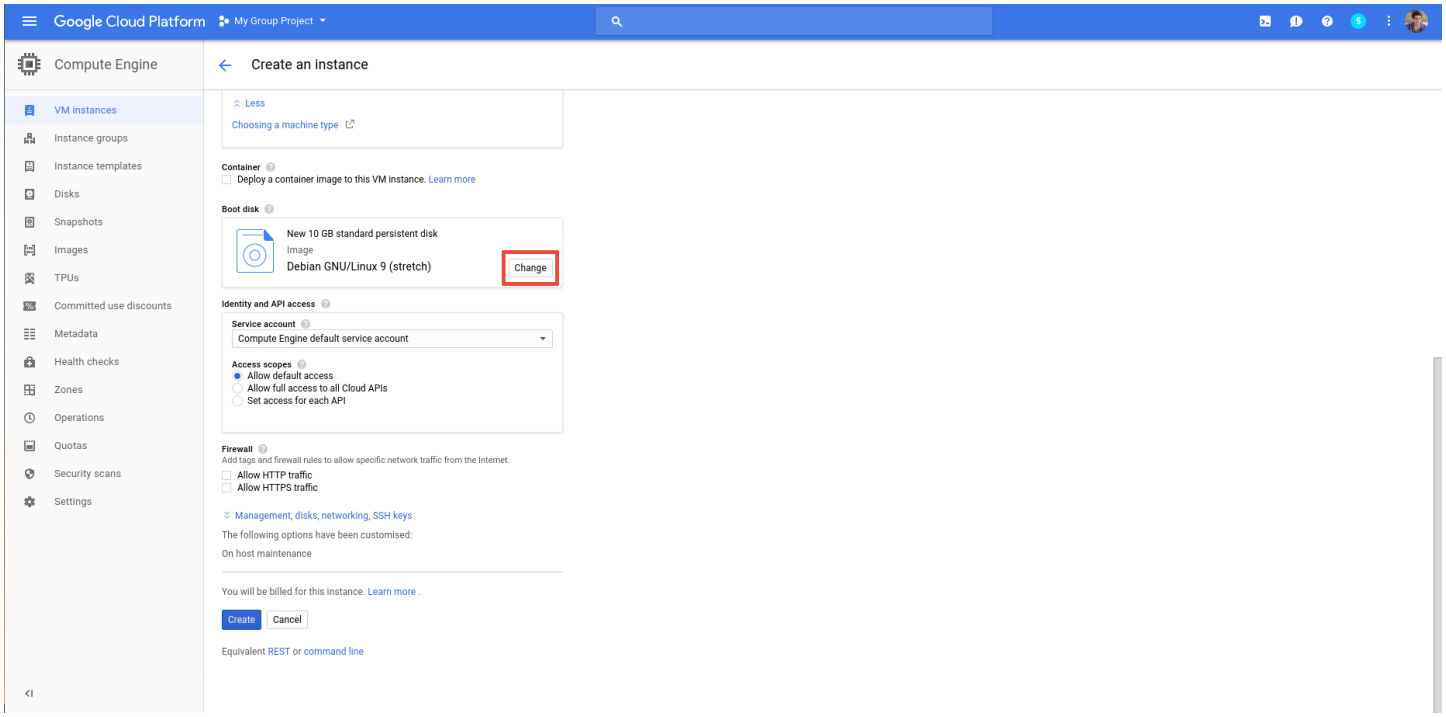
You can now choose the parameters you want for your VM. Below is a good set of parameters that we've been using:

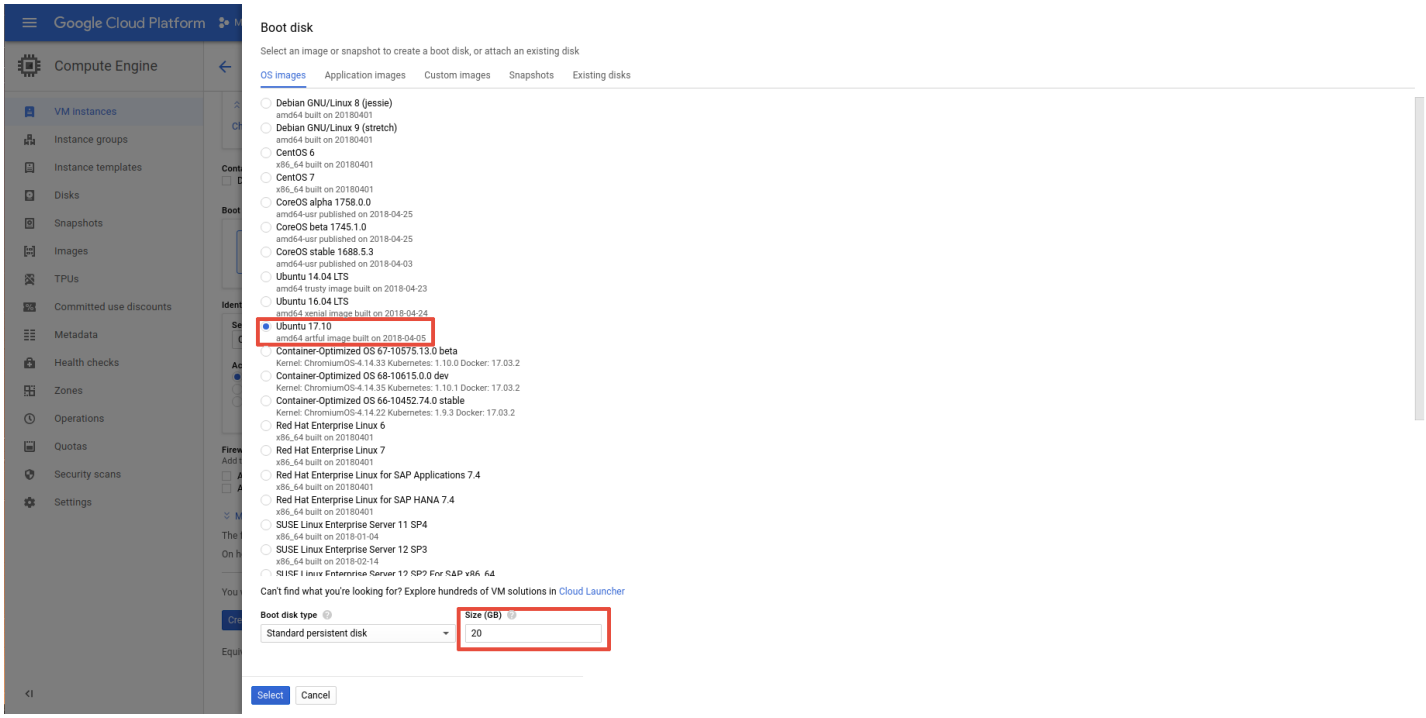


The zone *europe-west1-b* contains a K80 and is not too far from Sweden, hence can be a good choice. Select carefully the GPU (we advise you to take the K80) and the RAM (depending on the size of your dataset). If you're using large datasets that requires a parallel reader and fast pre-fetching, you might also want to choose a better CPU (with typically more cores).



Then choose your OS (Ubuntu is easier for installing the Nvidia drivers) and your hard disk size (we advise you to take around 20GB, since Python libraries for deep learning can be really heavy).





Finally, you can enable http(s) in the hidden options if you plan to use Tensorboard. Your VM is now created! From now, 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)

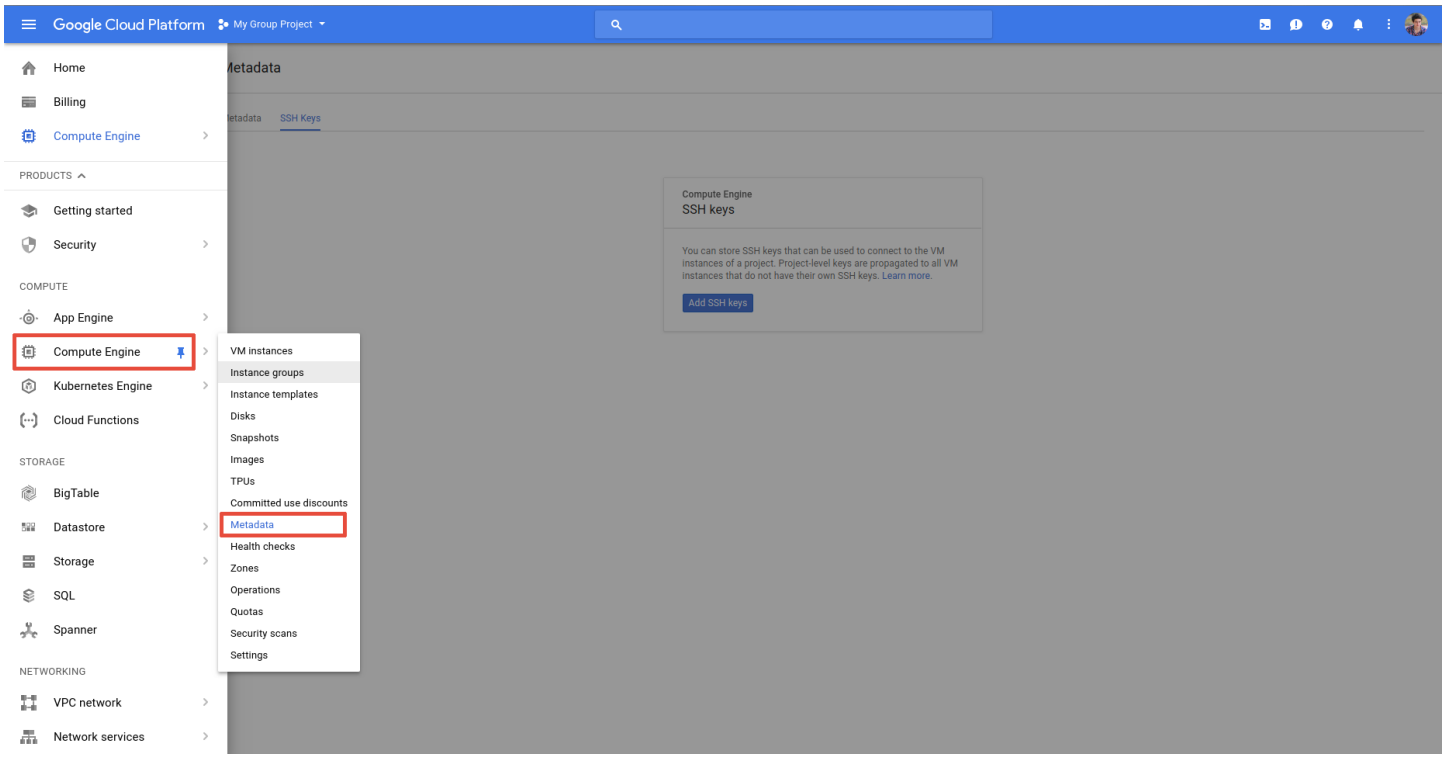
The next step to be able to connect is to handle your ssh keys. The instructions depend on your OS and are well explained on the GCP documentation. You can create as many ssh keys as you want, so don't hesitate to repeat the instructions for everyone in your group connecting from his own computer. Here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

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

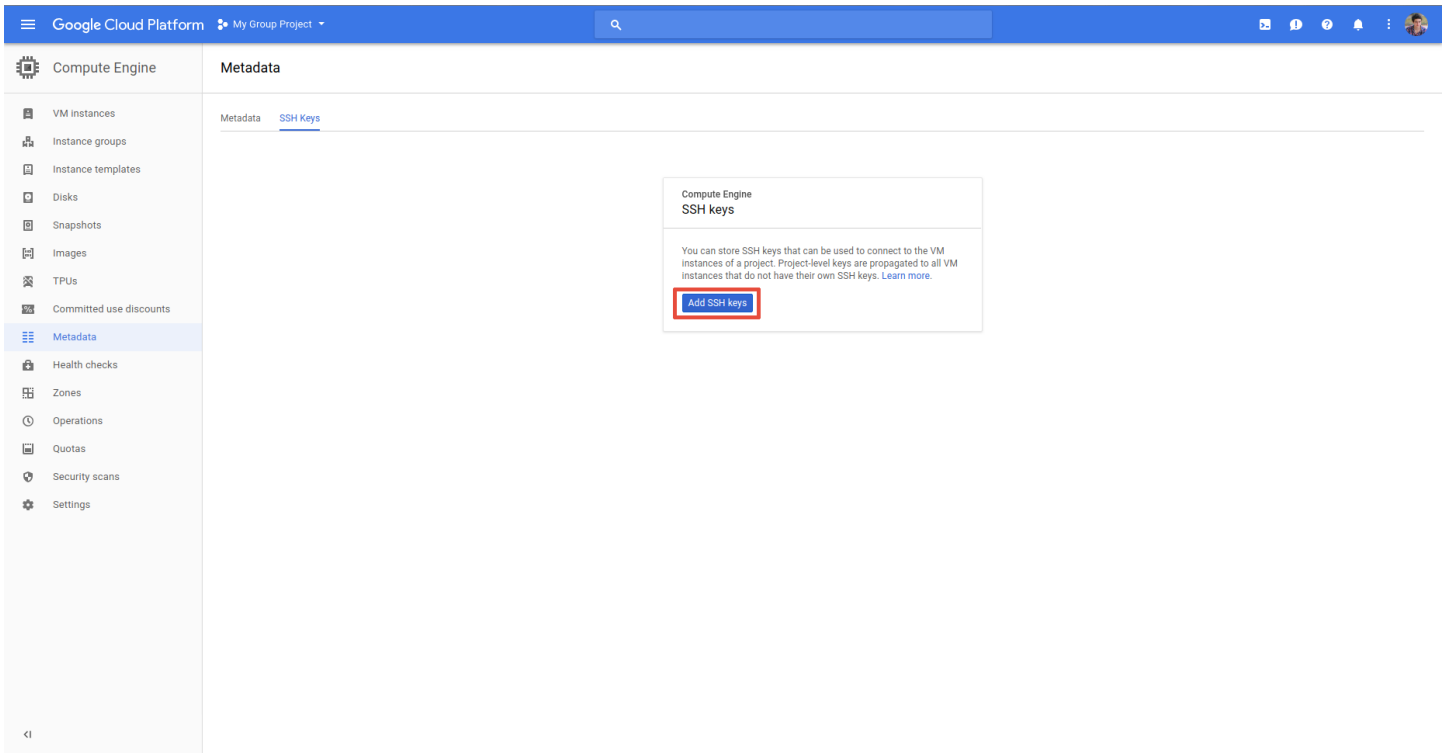
You can use the default parameters for the creation of the key. Then:

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

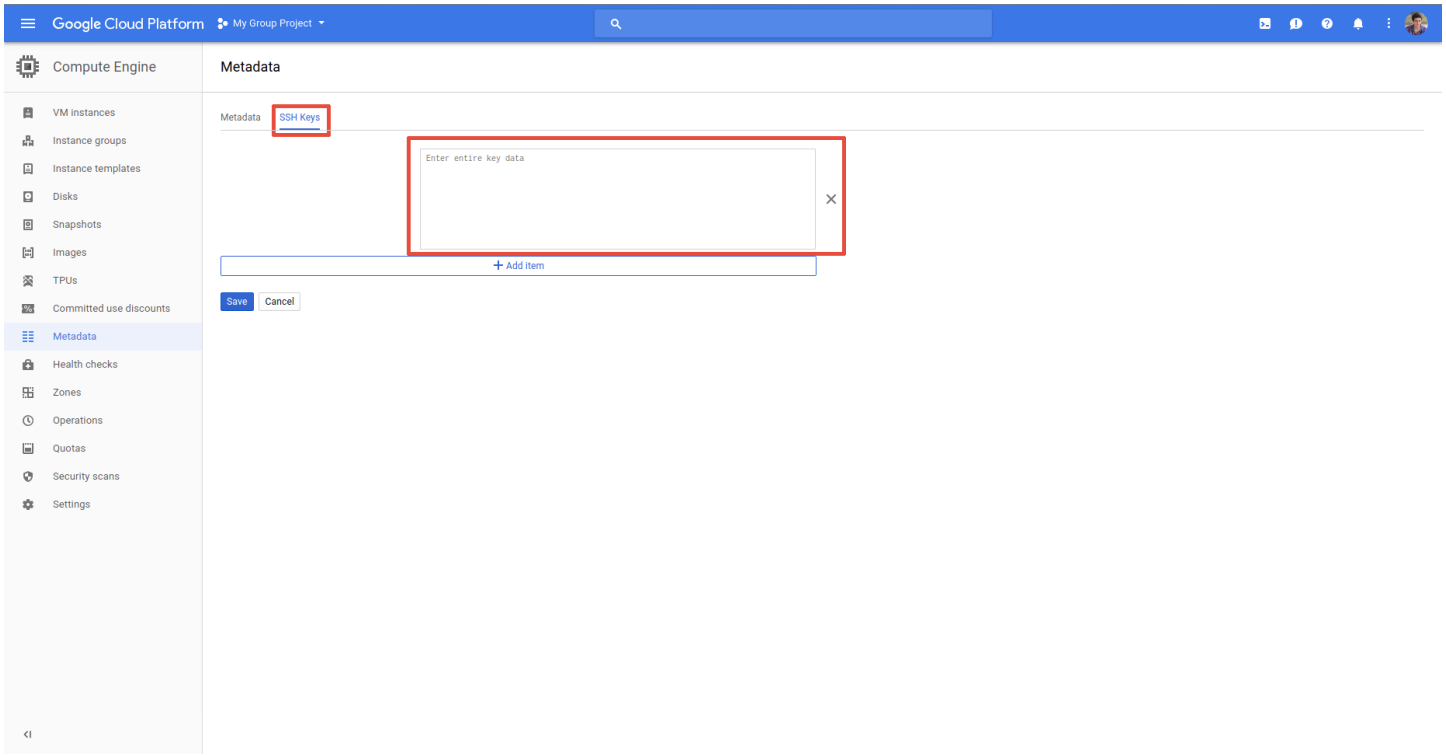
Copy the entire output (including your email address that should appear at the end of the key) and go back to GCP. Click on *Compute Engine* / *Metadata*:



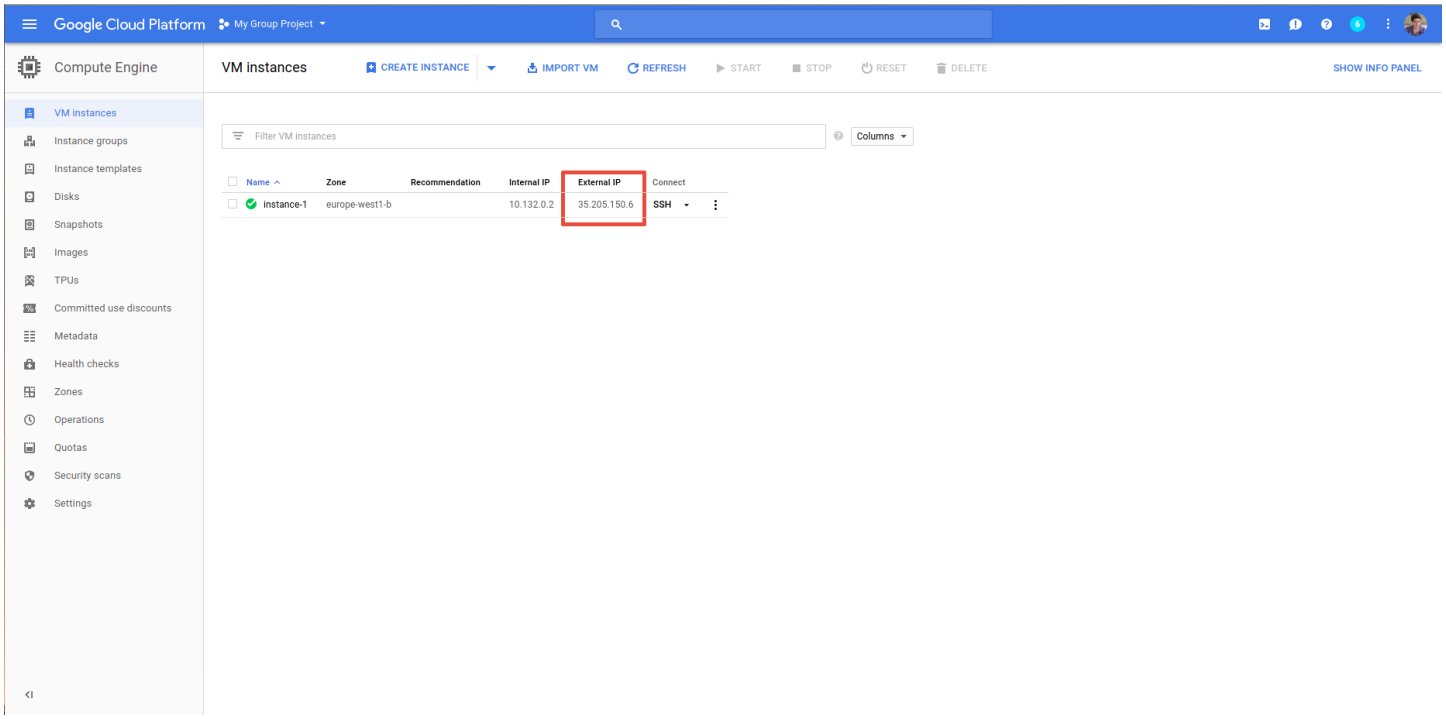
Then click on *SSH Keys* / *Add SSH keys*:



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:



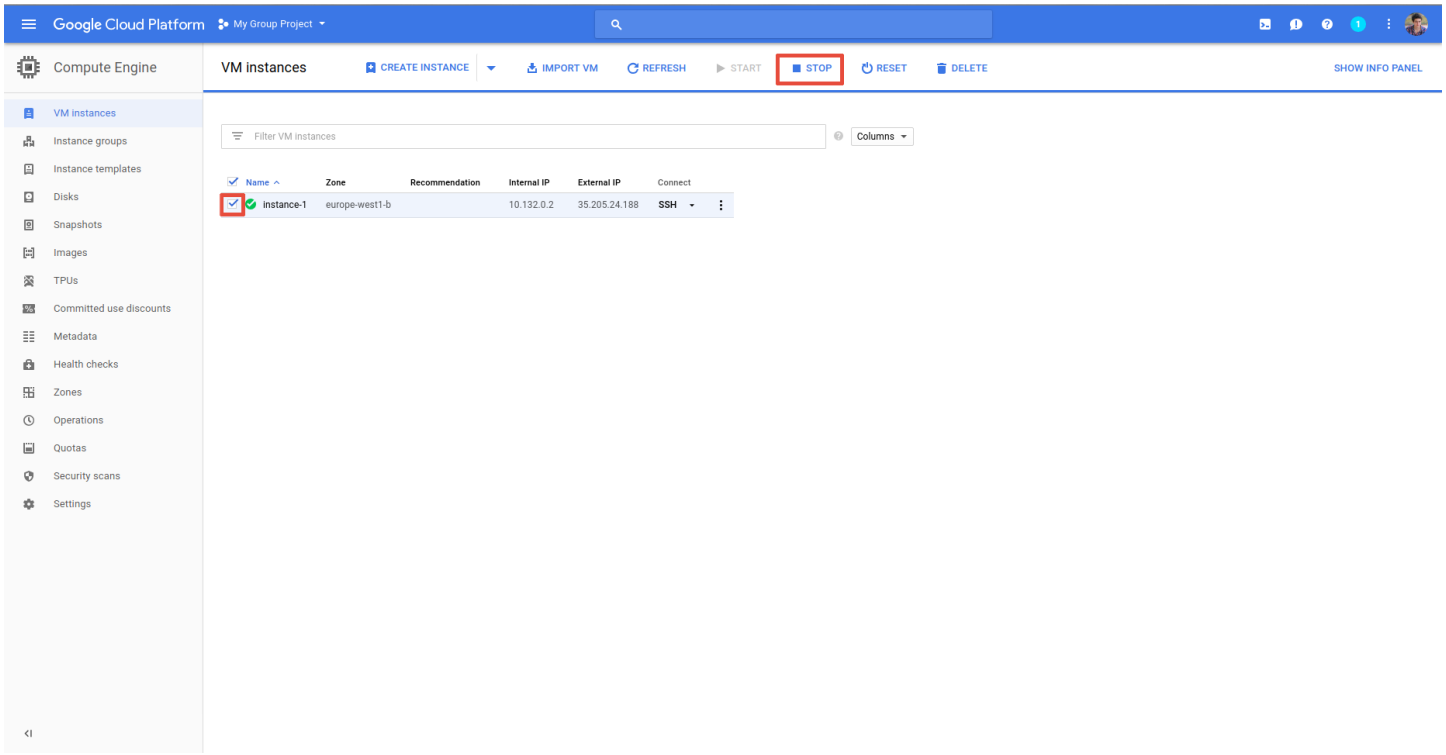
And connect to your machine using a command like that:

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

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

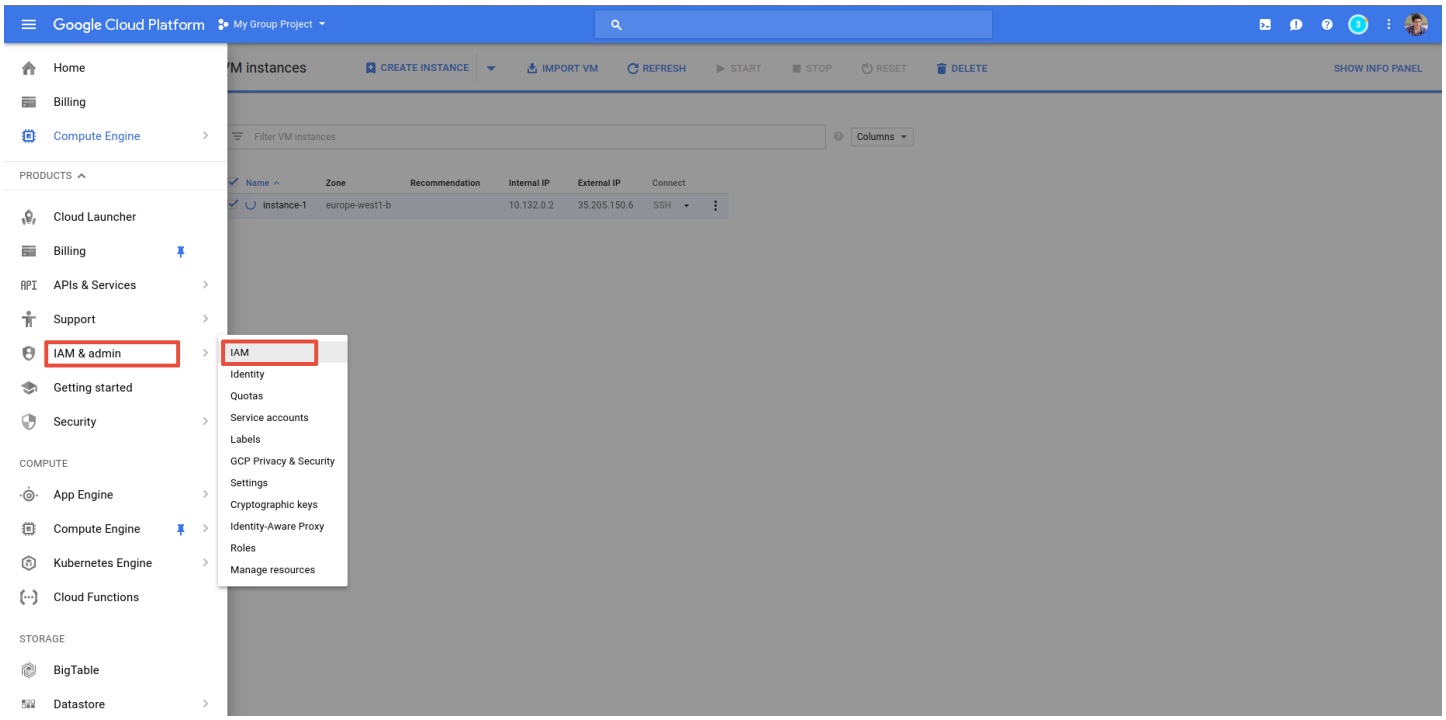
```
scp -i ~/.ssh/gcp-key -r [projectFolder] [google username]@[external ip address]:~
```

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

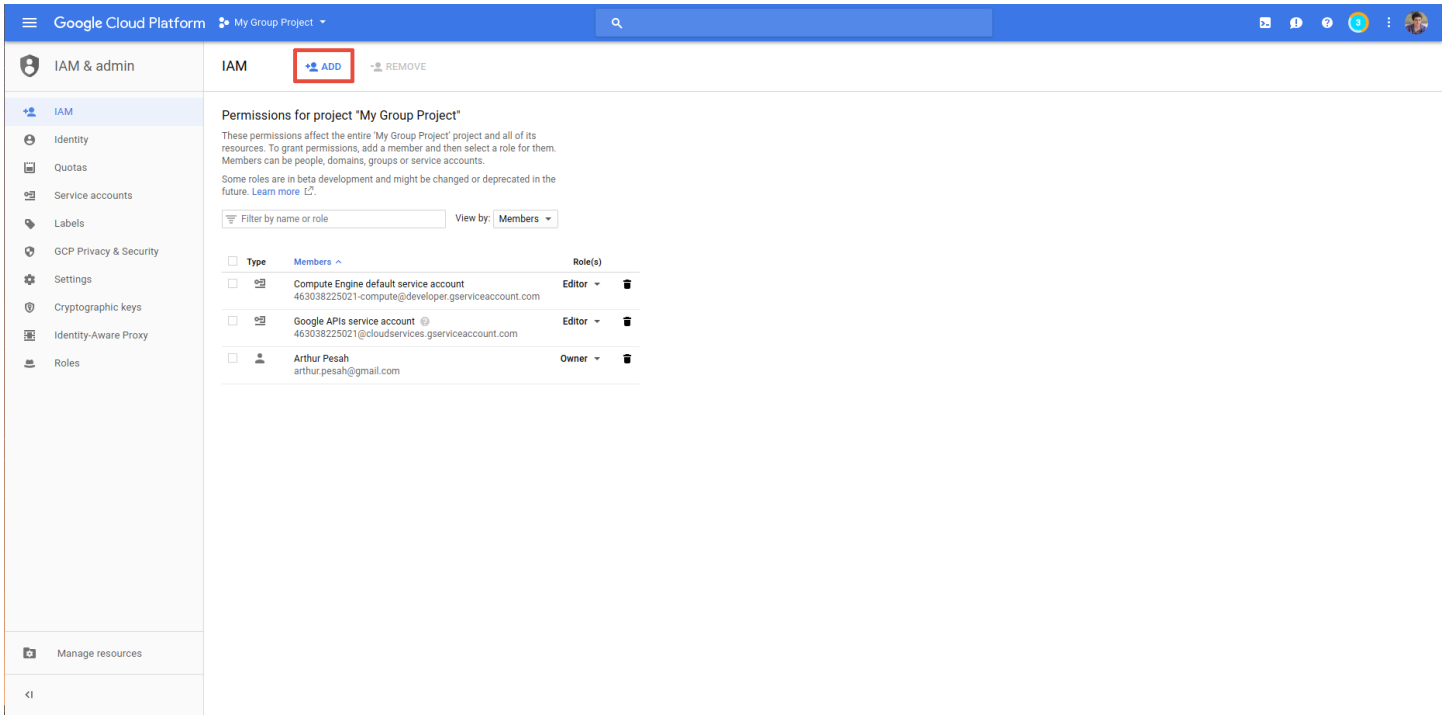


3 How to add new members to your project

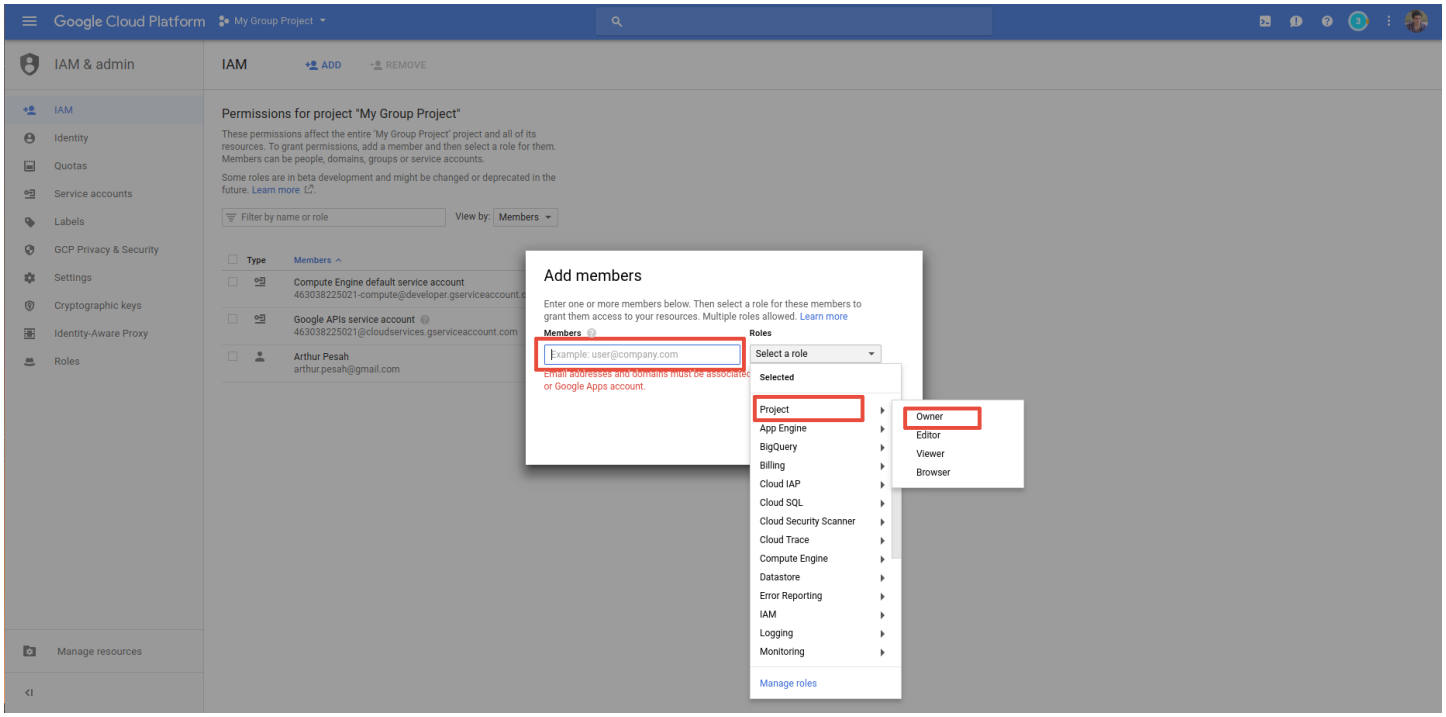
You might want other members of your team to be able to access to GCP. In that case, it's pretty straight-forward to add new members. Go to*IAM & admin / IAM*:



Then click on *Add*:



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



4 How to configure the machine for deep learning

Once you've gained access to the VM, you probably want to be able to use the GPU with deep learning libraries. For that, you first need to install Nvidia drivers and CUDA. To do so, type the following commands on the VM:

```
curl -O
    https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1604/x86_64/cuda-repo-ubuntu1604_8.0.61-1_amd64.deb
sudo dpkg -i ./cuda-repo-ubuntu1604_8.0.61-1_amd64.deb
sudo apt-get update
sudo apt-get install cuda-8-0 -y
```

```
sudo nvidia-smi -pm 1 # enable persistence mode
```

You should now restart the machine (using the web interface for instance, as shown above).

Once your machine has restarted, you can reconnect to the VM and install your Python environment, for instance with Anaconda (package manager used a lot in machine learning):

```
wget https://repo.continuum.io/miniconda/Miniconda3-latest-Linux-x86_64.sh
chmod +x Miniconda3-latest-Linux-x86_64.sh
./Miniconda3-latest-Linux-x86_64.sh
```

Answer *yes* to the last question (*Do you wish the installer to prepend the Miniconda3 install location to PATH in your /home/arthur.pesah/.bashrc ? [yes|no]*. If you've answered no without noticing, you can always add the line *export PATH=/home/[google username]/miniconda3/bin:$PATH* at the end of the file */.bashrc*.

Execute the following line to update your terminal:

```
source ~/.bashrc
```

To install the latest Tensorflow version (for instance), you can now type:

```
conda install -c conda-forge tensorflow tensorflow-gpu
```

Your VM should now be ready! You can test your installation by running for instance some Tensorflow Examples:

```
git clone https://github.com/aymericdamien/TensorFlow-Examples.git
cd TensorFlow-Examples/examples/3_NeuralNetworks/
python convolutional_network.py
```

Good luck for your project and enjoy your 80$ credits!