

# CS5284 : Graph Machine Learning

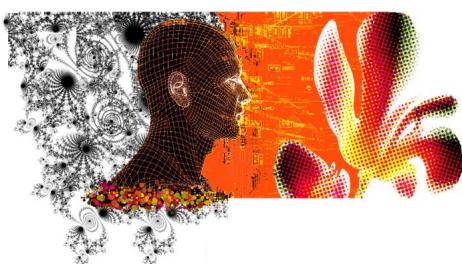
Google Cloud GPU

Semester 1 2025/26

Xavier Bresson

<https://x.com/xbresson>

Department of Computer Science  
National University of Singapore (NUS)



# Outline

1. Redeem coupon
2. Verify the billing information
3. Create a new project
4. Increase the GPUs quotas
5. Create a new VM instance
6. Connect to the VM by SSH



# \$300 in free credits

In addition to the Coupon, Google Cloud offers US\$ 300 credit to spend in the next 90 days:

<https://console.cloud.google.com>

This is likely not needed for the project, but for you to know (another project).

The screenshot shows the Google Cloud homepage with a prominent banner for a \$300 free trial. The banner includes a 'START FREE' button. Below the banner, there's a section for 'Popular getting started resources' with filters for 'General' selected. Two pre-built solution templates are shown: 'Deploy a three-tier web app' and 'Deploy load balanced managed VMs'.

## Popular getting started resources

Filter by

Web, Mobile, Game, Storage · Containers, VMs, Hybrid/Multi, Move Workload · Data, AI/ML, SAP · Maps, APIs · General

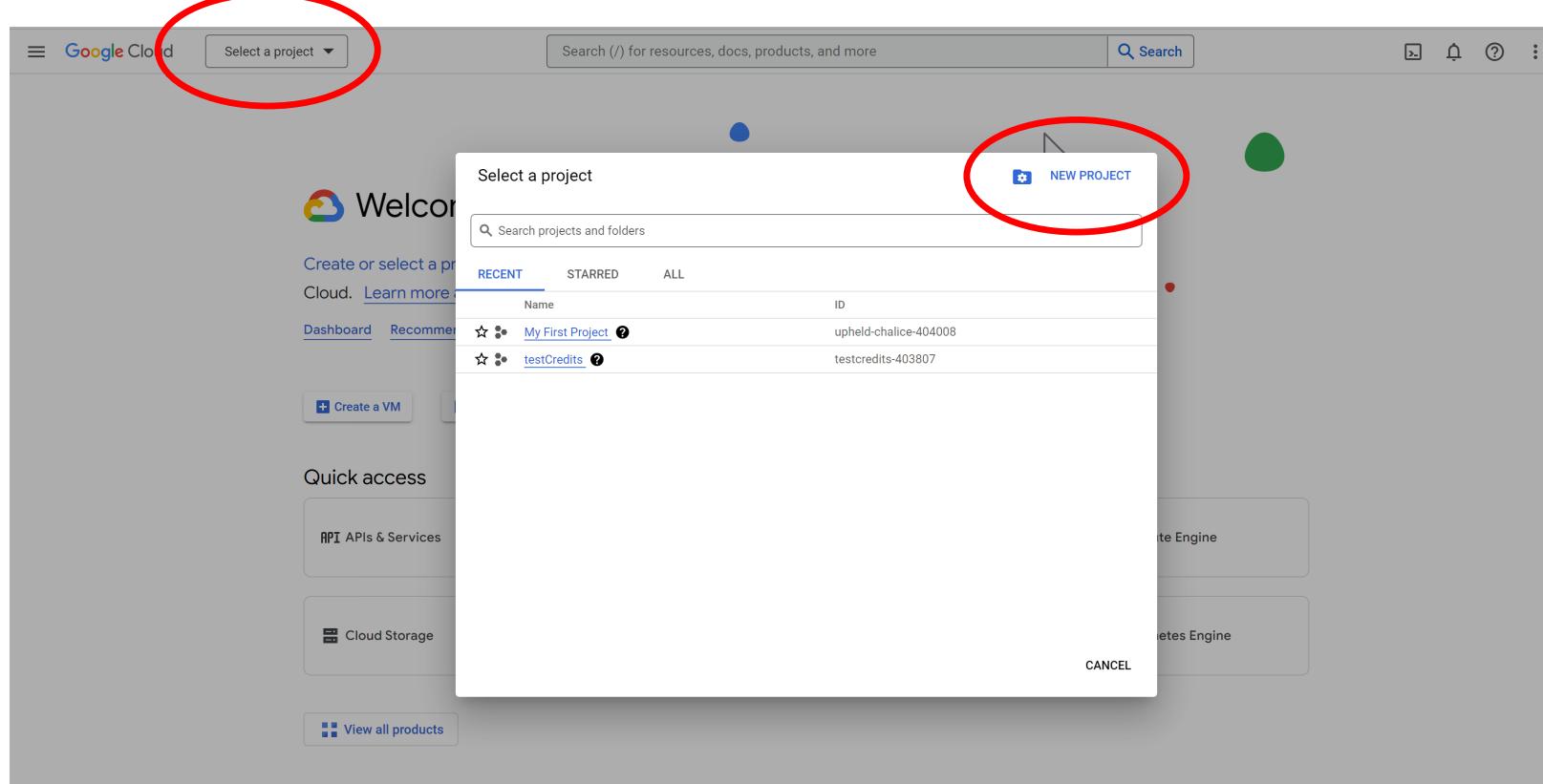
### Pre-built solution templates

Deploy a three-tier web app  
Web app, rich media site, ecommerce website, database-backed website

Deploy load balanced managed VMs  
Data analysis, data pipeline, application logs, log management, log intelligence

# Create a Project

- In the main page: <https://console.cloud.google.com>
- Select a Project > Create a Project.



# Billing Information

- You can verify that your coupon worked in the main page of the project > Billing >

The screenshot shows the Google Cloud Billing interface. On the left, the main Google Cloud Welcome screen is visible, featuring a 'Billing' button highlighted with a red oval. A red arrow points from this button to the central 'Billing' interface. The central interface is titled 'Billing' and shows 'Billing account: Billing Account for Education'. It includes sections for 'Overview', 'BILLING ACCOUNT OVERVIEW', 'Billing health checks', and a prominent 'Credits' section. The 'Credits' section displays '\$50.00' in a large green circle, indicating the remaining credit balance. A green box highlights this 'Credits' section.

Google Cloud Welcome

You're working in testCredits

Project number: 459513599115 | Project ID: testcredits-403807

Dashboard Recommendations

Create a VM Run a query in BigQuery Create a GKE cluster Create a storage bucket

Quick access

RPC APIs & Services IAM & Admin Billing

Cloud Storage BigQuery VPC network

View all products

Google Cloud Search (/) for resources, docs, products, and more

Billing Overview

Billing account: Billing Account for Education

Overview

BILLING ACCOUNT OVERVIEW

Billing health checks

Check out your account health results to avoid common billing-related issues and adopt our best practice recommendations. [Learn more](#)

0 1 1

Credits

\$50.00 Remaining credits Out of \$50.00

Remaining credits CS3244 Machine Learning pt11

Credit details

# Change GPU Quotas

- In the free version of Google Cloud, you do not have access to VM with GPUs, you need to ask for an increase of your GPU Quotas. In IAM&Admin>Quotas, look for “GPUS\_ALL\_REGIONS” using the filter. It should be set at 0. We will ask for an increase.

The screenshot shows the Google Cloud IAM & Admin Quotas interface. On the left, the navigation menu is open, with 'IAM & Admin' selected. Under 'Quotas', 'Quotas' is also selected, highlighted with a red circle. On the right, the main page displays 'Quotas for project "testCredits"'. A modal window titled 'Set up quota alerts' is open, providing instructions on how to manage quota alerts. Below the modal, a table shows current usage statistics: 'Current usage > 90%' is 0, and 'All quotas' total is 11,463. At the bottom, a filtered table lists quotas for the 'compute.googleapis.com/gpus\_all\_regions' metric. One entry for 'Compute Engine API' has a 'Limit' of 1, which is circled in green. The 'Current usage percentage' is 0%, and 'Current usage' is 0.

Service	Quota	Dimensions (e.g. location)	Limit	Current usage percentage	Current usage
<input type="checkbox"/> Compute Engine API	GPUs (all regions)		1	0%	0

# Increase GPU Quota

- Select the quota, click on “Edit Quota”, request a new quota value at 1, and provide a short justification on the fact that you are a student and you use it to complete a homework. This can take several business days to be accepted.

The screenshot shows the Google Cloud Quotas interface for the project "testCredits".

**Left Panel (Quotas for project "testCredits"):**

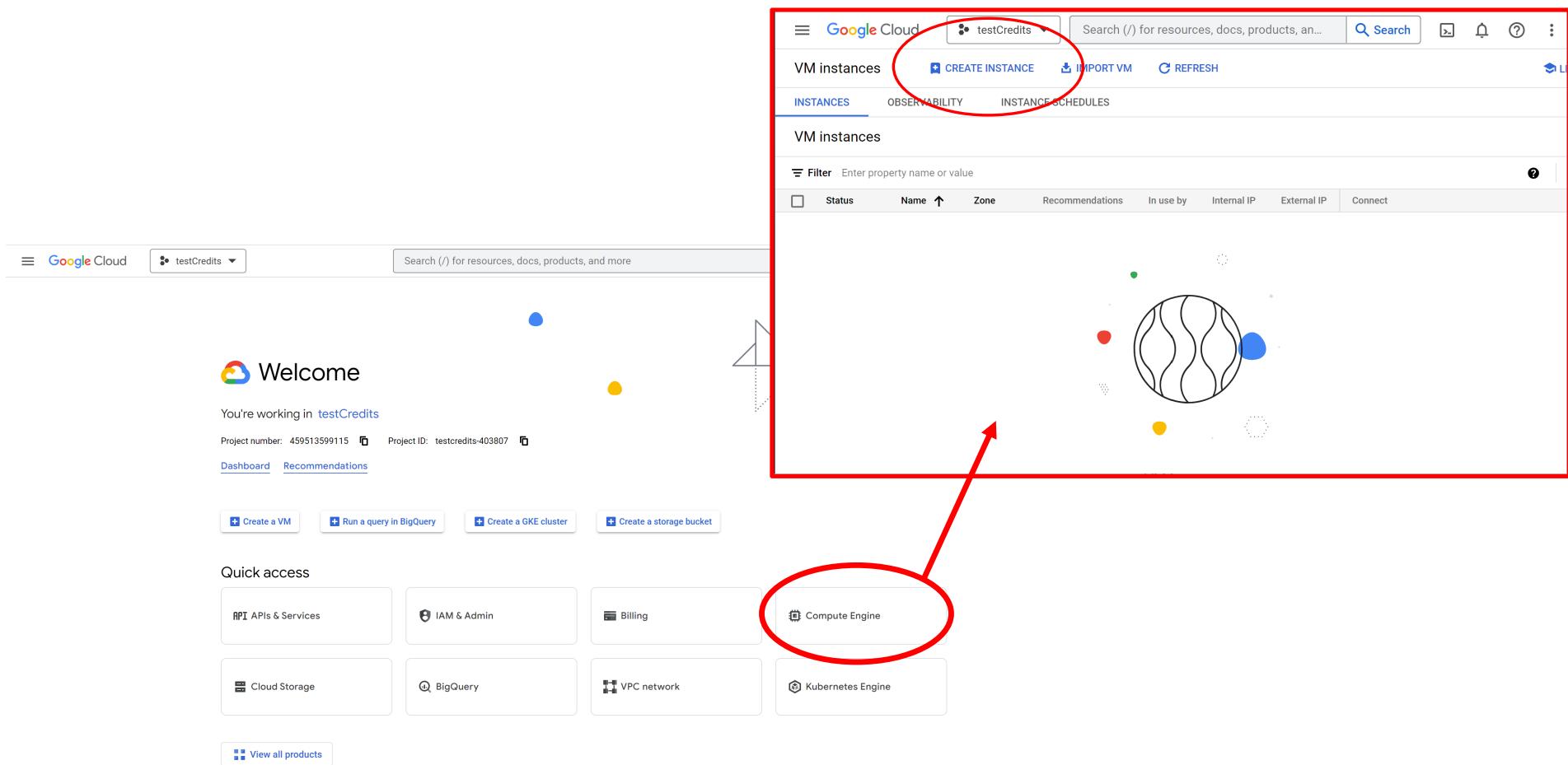
- Header: Google Cloud, testCredits, Search for resources, docs, products, and more.
- Tabs: QUOTAS (selected) and INCREASE REQUESTS.
- Section: Set up quota alerts (with a red circle around the alert icon).
- Section: Current usage > 90% (0) and All quotas (11,463).
- Table: A filtered view of quotas. The filter is set to Metric: compute.googleapis.com/gpus\_all\_regions. It shows two rows:
  - Service: Compute Engine API, Quota: GPUs (all regions), Dimensions: (e.g. location), Limit: 1, Current usage percentage: 0%, Current usage: 0.

**Right Panel (Edit Quota dialog):**

- Header: 1 quota selected.
- Section: Quota changes (Expand each service card to change individual quotas).
- Card: Edit quota for Compute Engine API.
  - Quota: GPUs (all regions)
  - Current limit: 1
  - New limit: 1 (highlighted with a red border)
  - Error message: The requested quota limit cannot be equal to current limit.
- Card: Request description \*
  - I am a student in National University of Singapore and I need to create a VM with GPU for a ML coursework class.
  - Your description will be sent to your service provider and is used to evaluate your request. It's useful to include the intent of the quota usage, future growth plans, region or zone spread, and any additional requirements or dependencies.
- Buttons: DONE and NEXT.

# Creating VM Instance

- From main screen : Go to Compute Engine > Create VM Instance



# Create VM Instance

The screenshot shows the Google Cloud 'Create an instance' page. On the left, a sidebar lists options: 'New VM instance' (selected), 'New VM instance from template', 'New VM instance from machine image', and 'Marketplace'. The main area shows fields for 'Name' (instance-1), 'Region' (asia-southeast1 (Singapore)), 'Zone' (asia-southeast1-b), and 'Machine configuration' (GPU type NVIDIA T4, Number of GPUs 1). A red box highlights the 'Region' field with the note 'Select Singapore server (faster), or Taiwan'. A blue box highlights the 'GPUs' tab in the 'Machine configuration' section with the note 'Select GPUs'. An orange box highlights the 'GPU type' field with the note 'Choose GPU model and number. Choose one that is available. K80 in Taiwan works, and T4 in SG sometimes too.' To the right, a sidebar shows a monthly estimate of US\$220.11 and a note: 'You have 50USD, so around 150h in this case'. Another sidebar shows compute engine pricing.

To create a VM instance, select one of the options:

- New VM instance
- New VM instance from template
- New VM instance from machine image
- Marketplace

**Create an instance**

Name \* instance-1

Region \* asia-southeast1 (Singapore)

Zone is permanent

asia-southeast1-b Zone is permanent

Machine configuration

General purpose Compute-optimised Memory-optimised  GPUs

Graphics processing units (GPUs) accelerate specific workloads on your instances such as machine learning and data processing. [Learn More](#)

GPU type NVIDIA T4

Number of GPUs 1

Enable Virtual Workstation (NVIDIA GRID)

Series	Description	vCPUs	Memory	Platform
N1	Balanced price and performance	1 - 96	1.8 - 624 GB	Intel Skylake

Machine type

Choose a machine type with preset amounts of vCPUs and memory that suit most workloads. Or, you can create a custom machine for your workload's particular needs. [Learn more](#)

PRESET CUSTOM

n1-standard-1 (1 vCPU, 3.75 GB memory)

vCPU Memory

Monthly estimate **US\$220.11**  
That's about US\$0.30 hourly  
Pay for what you use: No upfront costs and per-second billing

You have 50USD, so around 150h in this case

Compute Engine pricing [Learn more](#)

Monthly estimate US\$42.78

1 NVIDIA T4 US\$270.10

10 GB balanced persistent disk US\$1.10

Use discount -US\$93.86

Total US\$220.11

[Compute Engine pricing](#)

[LESS](#)

# Create a VM Instance

[Create an instance](#)

To create a VM instance, select one of the options:

**New VM instance**  
Create a single VM instance from scratch

**New VM instance from template**  
Create a single VM instance from an existing template

**New VM instance from machine image**  
Create a single VM instance from an existing machine image

**Marketplace**  
Deploy a ready-to-go solution onto a VM instance

## Boot disk ?

Name: instance-1  
Type: New balanced persistent disk  
Size: 50 GB  
Licence type: Free  
Image: Deep Learning VM M112

[CHANGE](#)

## Monthly estimate

**US\$224.51**

That's about US\$0.31 hourly

Pay for what you use: No upfront costs and per-second billing

Item	Monthly estimate
1 vCPU + 3.75 GB memory	US\$42.78
1 NVIDIA T4	US\$270.10
50 GB balanced persistent disk	US\$5.50
Use discount	-US\$93.86
Total	US\$224.51

[Compute Engine pricing](#)

[LESS](#)

## Boot disk

Select an image or snapshot to create a boot disk, or attach an existing disk. Can't find what you're looking for? Explore hundreds of VM solutions in [Marketplace](#)

[PUBLIC IMAGES](#) [CUSTOM IMAGES](#) [SNAPSHOTS](#) [ARCHIVE SNAPSHOTS](#) [EXISTING DISKS](#)

Operating system: Deep Learning on Linux

Version: Deep Learning VM M112

Debian 10, Python 3.7, Intel MKL. With Intel optimized NumPy, SciPy, and scikit-learn preinstalled.

Boot disk type: Balanced persistent disk

[COMPARE DISK TYPES](#)

Size (GB): 50

Provision between 50 and 65536 GB

[SHOW ADVANCED CONFIGURATION](#)

[SELECT](#)

[CANCEL](#)

Change the OS from Debian to the one adapted for Deep Learning (CUDA and python already installed)

Allow HTTP and HTTPS traffic. This could be useful to install a jupyter server

## Firewall ?

Add tags and firewall rules to allow specific network traffic from the Internet

- Allow HTTP traffic
- Allow HTTPS traffic
- Allow load balancer health checks

## Advanced options

Networking, disks, security, management, sole-tenancy

[CREATE](#)

[CANCEL](#)

[EQUIVALENT CODE](#)

# Stopping the VM

- **/!\ STOP YOUR VM IF YOU ARE NOT USING IT /!\\**
- You are paying as long as the green icon is present.

The screenshot shows the Google Cloud VM instances page. At the top, there is a navigation bar with the Google Cloud logo, a dropdown for 'testCredits', a search bar containing 'firewa', and a 'Search' button. Below the navigation bar, there are tabs for 'VM instances', '+ CREATE INSTANCE', 'IMPORT VM', and 'REFRESH'. Underneath these tabs, there are three sub-tabs: 'INSTANCES' (which is selected), 'OBSERVABILITY', and 'INSTANCE SCHEDULES'. A message '1 instance selected' is displayed, followed by buttons for 'START / RESUME', 'STOP', 'SUSPEND', 'RESET', 'DELETE', 'LABELS', and 'PERMISSIONS'. A red circle highlights the 'STOP' button. Below this, a table lists the selected instance: 'instance-2' in 'asia-east1-b' zone, with internal IP 10.140.0.2 (nic0) and external IP 34.81.122.198 (nic0). A red circle highlights the green status icon next to the instance name. In the 'Related actions' section, there are several cards: 'Explore Backup and DR' (NEW), 'View billing report', 'Monitor VMs', 'Set up firewall rules', 'Patch management', and 'Load balance between V'. To the right of the table, a vertical menu is open, listing actions: 'Start / Resume', 'Stop' (highlighted with a red circle), 'Suspend', 'Reset', 'Delete', 'View network details', 'Create new machine image', 'View logs', and 'View monitoring'.

Status	Name	Zone	Recommendations	In use by	Internal IP	External IP	Connect
	instance-2	asia-east1-b			10.140.0.2 (nic0)	34.81.122.198 (nic0)	SSH

**Related actions**

- Explore Backup and DR **NEW**  
Back up your VMs and set up disaster recovery
- View billing report  
View and manage your Compute Engine billing
- Monitor VMs  
View outlier VMs across metrics like CPU and network
- Set up firewall rules  
Control traffic to and from a VM instance
- Patch management  
Schedule patch updates and view patch compliance on VM instances
- Load balance between V  
Set up Load Balancing for your application as your traffic and users grow

Start / Resume  
Stop  
Suspend  
Reset  
Delete  
View network details  
Create new machine image  
View logs  
View monitoring

# Connect to the VM instance

- You can directly connect by SSH to the instance. You can then create your python environment, install your packages and launch .py scripts from this console.

The screenshot shows the Google Cloud Platform interface for managing VM instances. The top navigation bar includes the Google Cloud logo, a dropdown for 'testCredits', a search bar, and various user icons. Below the header, the 'VM instances' section is active, with options to 'CREATE INSTANCE', 'IMPORT VM', and 'REFRESH'. A 'LEARN' button is also present. The main area displays a table of VM instances, with one row selected for 'instance-2'. The table columns include Status, Name (sorted), Zone, Recommendations, In use by, Internal IP, External IP, and a 'Connect' button. The 'Connect' button for 'instance-2' is set to 'SSH'. To the right of the table, a terminal window titled 'ssh.cloud.google.com/v2/ssh/projects/testcredits-403807/zones/asia-east1-b/instances/instance-2?aut...' is open. The terminal session shows a welcome message for the Google Deep Learning VM, details about the operating system (Debian 10), and a list of resources available. It also provides instructions for reinstalling the Nvidia driver and notes about the software's license and warranty.

Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP	Connect
<input type="checkbox"/>	instance-2	asia-east1-b			10.140.0.2 (nic0)	34.81.122.198 (nic0)	SSH

**Related actions**

- Explore Backup and DR NEW  
Back up your VMs and set up disaster recovery
- View billing report  
View and manage your Compute billing
- Set up firewall rules  
Control traffic to and from a VM instance
- Patch management  
Schedule patch updates and view compliance on VM instances

**SSH-in-browser**

```
=====
Welcome to the Google Deep Learning VM
=====

Version: common-cpu.m12
Based on: Debian GNU/Linux 10 (buster) (GNU/Linux 4.19.0-25-cloud-amd64 x86_64)

Resources:
* Google Deep Learning Platform StackOverflow: https://stackoverflow.com/questions/tagged/google-dl-platform
* Google Cloud Documentation: https://cloud.google.com/deep-learning-vm
* Google Group: https://groups.google.com/forum/#!forum/google-dl-platform

To reinstall Nvidia driver (if needed) run:
sudo /opt/deeplearning/install-driver.sh
Linux instance-2 4.19.0-25-cloud-amd64 #1 SMP Debian 4.19.289-2 (2023-08-08) x86_64

The programs included with the Debian GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*copyright.

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
(base) mmichelezza@instance-2:~$ 
```

# Use Notebooks : Configure Jupyter server

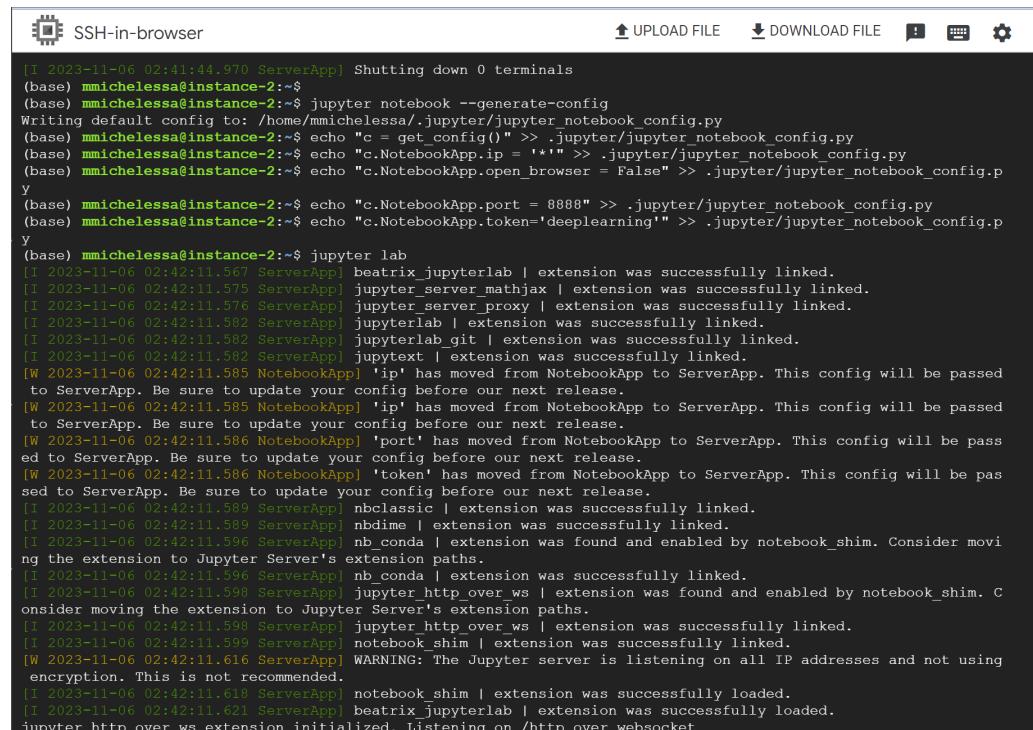
Jupyter lab is already pre-installed.

Type the following commands to configure the jupyter server.

```
jupyter notebook --generate-config  
echo "c = get_config()" >> .jupyter/jupyter_notebook_config.py  
echo "c.NotebookApp.ip = '*' >> .jupyter/jupyter_notebook_config.py  
echo "c.NotebookApp.open_browser = False" >> .jupyter/jupyter_notebook_config.py  
echo "c.NotebookApp.port = 8888" >> .jupyter/jupyter_notebook_config.py  
echo "c.NotebookApp.token='deeplearning'" >> .jupyter/jupyter_notebook_config.py
```

To start the server, type:

```
jupyter lab
```



The screenshot shows an SSH-in-browser interface. At the top, there are buttons for UPLOAD FILE, DOWNLOAD FILE, and various system icons. The main area is a terminal window with the following text:

```
[I 2023-11-06 02:41:44.970 ServerApp] Shutting down 0 terminals  
(base) mmichelessa@instance-2:~$ jupyter notebook --generate-config  
Writing default config to: /home/mmichelessa/.jupyter/jupyter_notebook_config.py  
(base) mmichelessa@instance-2:~$ echo "c = get_config()" >> .jupyter/jupyter_notebook_config.py  
(base) mmichelessa@instance-2:~$ echo "c.NotebookApp.ip = '*' >> .jupyter/jupyter_notebook_config.py  
(base) mmichelessa@instance-2:~$ echo "c.NotebookApp.open_browser = False" >> .jupyter/jupyter_notebook_config.py  
(base) mmichelessa@instance-2:~$ echo "c.NotebookApp.port = 8888" >> .jupyter/jupyter_notebook_config.py  
(base) mmichelessa@instance-2:~$ echo "c.NotebookApp.token='deeplearning'" >> .jupyter/jupyter_notebook_config.py  
(base) mmichelessa@instance-2:~$ jupyter lab  
[I 2023-11-06 02:42:11.567 ServerApp] beatrix_jupyterlab | extension was successfully linked.  
[I 2023-11-06 02:42:11.575 ServerApp] jupyter_server_mathjax | extension was successfully linked.  
[I 2023-11-06 02:42:11.576 ServerApp] jupyter_server_proxy | extension was successfully linked.  
[I 2023-11-06 02:42:11.582 ServerApp] jupyterlab | extension was successfully linked.  
[I 2023-11-06 02:42:11.582 ServerApp] jupyterlab_git | extension was successfully linked.  
[I 2023-11-06 02:42:11.582 ServerApp] jupytertext | extension was successfully linked.  
[W 2023-11-06 02:42:11.585 NotebookApp] 'ip' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.  
[W 2023-11-06 02:42:11.585 NotebookApp] 'ip' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.  
[W 2023-11-06 02:42:11.586 NotebookApp] 'port' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.  
[W 2023-11-06 02:42:11.586 NotebookApp] 'token' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.  
[I 2023-11-06 02:42:11.589 ServerApp] nbclassic | extension was successfully linked.  
[I 2023-11-06 02:42:11.589 ServerApp] nbdimx | extension was successfully linked.  
[I 2023-11-06 02:42:11.596 ServerApp] nb_conda | extension was found and enabled by notebook_shim. Consider moving the extension to Jupyter Server's extension paths.  
[I 2023-11-06 02:42:11.596 ServerApp] nb_conda | extension was successfully linked.  
[I 2023-11-06 02:42:11.598 ServerApp] jupyter_http_over_ws | extension was found and enabled by notebook_shim. Consider moving the extension to Jupyter Server's extension paths.  
[I 2023-11-06 02:42:11.598 ServerApp] jupyter_http_over_ws | extension was successfully linked.  
[I 2023-11-06 02:42:11.599 ServerApp] notebook_shim | extension was successfully linked.  
[W 2023-11-06 02:42:11.616 ServerApp] WARNING: The Jupyter server is listening on all IP addresses and not using encryption. This is not recommended.  
[I 2023-11-06 02:42:11.618 ServerApp] notebook_shim | extension was successfully loaded.  
[I 2023-11-06 02:42:11.621 ServerApp] beatrix_jupyterlab | extension was successfully loaded.  
[I 2023-11-06 02:42:11.621 ServerApp] jupyter_http_over_ws extension initialized. Listening on /http_over_websocket
```

# Use Notebooks: Create firewall rule

To communicate with the jupyter server from your machine, you need to open the port 8888 from your instance

From the VM instance screen, go to “Set up firewall rules” and click on “Create Firewall Rule”

Google Cloud testCredits

Compute Engine

VM instances CREATE INSTANCE IMPORT VM REFRESH

Virtual machines VM instances

Instance templates Sole-tenant nodes Machine images TPUs Committed use discounts Reservations Migrate to Virtual Machin...

INSTANCE OBSERVABILITY INSTANCE SCHEDULES

VM instances

Filter Enter property name or value

Status	Name	Zone	Recommendations	In use by	Connect
<input type="checkbox"/>	instance-2	asia-east1-b			SSH

Related actions

- Explore Backup and DR NEW
- View billing report
- Monitor VMs
- Explore VM logs
- Set up firewall rules** Control traffic to and from a VM instance
- Patch management

Google Cloud testCredits

Firewall policies CREATE FIREWALL POLICY CREATE FIREWALL RULE

Get real-time analytics with Network Intelligence Center

Use Network Intelligence Center for comprehensive monitoring and troubleshooting. [Learn more](#)

- ✓ Visualize your network resources
- ✓ Diagnose and prevent connectivity issues
- ✓ View packet loss and latency metrics
- ✓ Keep your firewall rules strict and efficient

GO TO NETWORK INTELLIGENCE CENTER REMIND ME LATER

# Create Firewall Rule

- In the firewall rule screen, change the following parameters:
  - Target: All instances in the network
  - Source ipv4 range: 0.0.0.0/0
  - Specified Protocols : TCP – Ports: 8888

Google Cloud testCredits firewa

← Create a firewall rule

Targets  
All instances in the network

Source filter  
IPv4 ranges

Source IPv4 ranges \*  
0.0.0.0/0

Second source filter  
None

Destination filter  
None

Protocols and ports ?  
 Allow all  
 Specified protocols and ports

TCP  
Ports  
8888  
E.g. 20, 50-60

UDP  
Ports  
E.g. all

Other  
Protocols

# Use Notebooks : Access Jupyter Server

- You can access Jupyter Lab directly from your browser.

Filter Enter property name or value						
Status	Name ↑	Zone	Recommendations	In use by	Internal IP	External IP
<input checked="" type="checkbox"/>	<a href="#">instance-2</a>	asia-east1-b			10.140.0.2 (nic0)	<a href="#">34.81.122.198 (nic0)</a>

The token is  
“deeplearning”

Token authentication is enabled

If no password has been configured, you need to open the server with its login token in the URL, or paste it above. This requirement will be lifted if you [enable a password](#).

The command:

```
jupyter server list
```

will show you the URLs of running servers with their tokens, which you can copy and paste into your browser. For example:

```
Currently running servers:  
http://localhost:8888/?token=c8de56fa... :: /Users/you/notebook
```

or you can paste just the token value into the password field on this page.

See [the documentation on how to enable a password](#) in place of token authentication, if you would like to avoid dealing with random tokens.

Cookies are required for authenticated access to the Jupyter server.

# Use Notebooks

- You have now access to Jupyter Lab, you can upload your colab notebooks and data and run your codes.

