

#### **UPDATE 13 APRIL 2019 :**

Since leela zero 0.17 release, it is needed to use gcc-8 or later version, scripts have been updated to support it now, thanks to [@ozymandias8](#) for testing and feedback in : <https://github.com/leela-zero/leela-zero/issues/2333>

#### **VERSION OF 02 FEBRUARY 2019 :**

See changelogs at the end of this message : <https://github.com/gcp/leela-zero/issues/1905#issue-366048983>

This tutorial may seem long, but it is because a lot of screenshots are included and take too much place in an unoptimal way.

As you can see in the video tutorial (link will be added later), all these instructions will take you less than 5 minutes to get through once you are used to them.

**UPDATE 02 february 2019 :** <https://github.com/gcp/leela-zero> is now <https://github.com/leela-zero/leela-zero>

Current scripts are likely not to work anymore, recreate a new template with the startup script now or you free trial will be wasted on unworking instructions..

**UPDATE 26 november 2018 :** due to a recent change in google cloud policy, november 2018 and newer accounts need to make a quota request for GPU and preemptible CPU or they get the following error message at instance creation : **Quota 'GPUS\_ALL\_REGIONS' exceeded. Limit: 0.0 globally.**

See “Quota requests” part in page 3 for more details

#### **IMPORTANT UPDATE : 13 november 2018 !!!**

**TO ALL THOSE WHO USED THE FREE TRIAL INSTRUCTIONS BEFORE 13 NOVEMBER 2018 :**

**OLD STARTUP-SCRIPT DOESN'T WORK ANYMORE ! UPDATE NEEDED, OR YOUR FREE CREDIT WILL BE CONSUMED WITHOUT DOING ANY GAME PRODUCTION !**

**YOU NEED TO DELETE YOUR INSTANCE GROUP AND INSTANCE TEMPLATE, AND CREATE A NEW TEMPLATE WITH THE UPDATED STARTUP-SCRIPT IN PAGE 10 OF THIS DOC**

Here is a preview of what you can do if you follow this tutorial :

mygooglename@instance-group-5-g0gr: ~ - Google Chrome

https://ssh.cloud.google.com/projects/endless-upgrade-217714/zones/us-central1-b/instances/instance-group-5-g0gr?authuser=0&hl=en\_US&projectNumber=601699849685

```
D14) 158 (W J4) 206 (W E14) 207 (B C13) 159 (B K4) 208 (W D13) 160 (W L4) 209 (B D12) 210 (W D14) 211 (B B13) 212 (W B15) 213 (B B16) 214 (W C12) 161 (B D3) 215 (B B18) 216 (W M15) 217 (B N13) 16
2 (W M5) 218 (W D17) 163 (B C4) 219 (B A15) 164 (W B2) 220 (W D11) 165 (B A2) 166 (W B4) 167 (B A4) 221 (B C17) 222 (W D18) 168 (W A3) 169 (B P17) 223 (B A14) 224 (W E10) 225 (B B19) 226 (W A17) 2
27 (B A18) 170 (W N4) 228 (W M13) 229 (B L14) 171 (B N6) 172 (W O3) 173 (B P4) 230 (W N18) 231 (B M16) 174 (W Q17) 232 (W M19) 175 (B B3) 233 (B L15) 176 (W B5) 177 (B B1) 234 (W A16) 235 (B B17)
236 (W A16) 237 (B H13) 238 (W Q18) 239 (B Q19) 240 (W R19) 178 (W D19) 241 (B J5) 179 (B K5) 242 (W J6) 243 (B R18) 180 (W K3) 181 (B H4) 182 (W J5) 183 (B H5) 184 (W J6) 185 (B H6) 244 (W F8) 18
6 (W H3) 187 (B J7) 188 (W K6) 245 (B B6) 189 (B K7) 190 (W L6) 191 (B C1) 246 (W B7) 192 (W G6) 193 (B H7) 247 (B B5) 194 (W F8) 195 (B O10) 248 (W R19) 196 (W L7) 197 (B O15) 249 (B K16) 198 (W
P18) 250 (W O18) 199 (B O17) 200 (W O14) 251 (B D2) 201 (B F15) 202 (W G14) 252 (W M8) 203 (B D15) 253 (B K9) 204 (W E15) 205 (B E16) 206 (W O16) 207 (B J15) 254 (W E2) 208 (W F16) 255 (B C2) 209
(B G3) 256 (W L17) 257 (B A7) 210 (W J2) 211 (B E16) 212 (W O5) 213 (B P5) 214 (W F16) 215 (B O13) 258 (W P18) 259 (B M15) 216 (W Q12) 260 (W A8) 217 (B E16) 261 (B A6) 262 (W C7) 218 (W N15) 263
(B A17) 219 (B G15) 220 (W F16) 221 (B B17) 222 (W B18) 264 (W C5) 223 (B E16) 224 (W K8) 225 (B G4) 265 (B B4) 226 (W F16) 227 (B G19) 266 (W L4) 228 (W E19) 229 (B E16) 267 (B L3) 230 (W H15) 23
1 (B J16) 232 (W F16) 268 (W K4) 233 (B N14) 269 (B E8) 234 (W E16) 270 (W E9) 235 (B P14) 236 (W O18) 271 (B H19) 237 (B N16) 238 (W H13) 239 (B H12) 240 (W N17) 241 (B K16) 272 (W J18) 242 (W P1
6) 243 (B N16) 244 (W I9) 273 (B K8) 245 (B A18) 274 (W K7) 275 (B L8) 276 (W L7) 246 (W F13) 277 (B J9) 247 (B F12) 278 (W G10) 248 (W B16) 279 (B H10) 249 (B A16) 250 (W A17) 280 (W E6) 251 (B M
8) 252 (W L8) 281 (B F7) 253 (B B17) 254 (W G12) 282 (W G8) 255 (B G11) 283 (B G9) 256 (W A17) 284 (W F9) 257 (B F4) 285 (B J1) 258 (W A15) 259 (B G2) 286 (W G1) 260 (W N3) 287 (B E1) 261 (B P2) 2
88 (W F1) 289 (B D1) 290 (W K3) 262 (W T17) 291 (B L2) 292 (W J19) 293 (B G19) 263 (B H2) 294 (W E5) 264 (W B13) 295 (B H4) 296 (W G3) 265 (B B14) 297 (B N13) 266 (W A12) 267 (B D10) 298 (W B12) 2
68 (W O11) 299 (B M13) 300 (W A13) 301 (B E18) 269 (B N11) 302 (W A16) 303 (B A15) 304 (W A14) 270 (W J10) 305 (B D19) 306 (W A16) 271 (B H10) 307 (B D9) 308 (W C11) 309 (B A15) 272 (W J11) 310 (W
J16) 273 (B H11) 311 (B C19) 274 (W M11) 312 (W A16) 275 (B M10) 313 (B B8) 314 (W A9) 315 (B A15) 276 (W N12) 316 (W K17) 277 (B N10) 317 (B K15) 318 (W A16) 278 (W O12) 279 (B M12) 319 (B A10)
280 (W K10) 281 (B L11) 320 (W B9) 282 (W J1) 321 (B A15) 322 (W E19) 283 (B J3) 323 (B F19) 324 (W A16) 325 (B N7) 326 (W M7) 327 (B A15) 284 (W P10) 328 (W N4) 329 (B O4) 330 (W A16) 331 (B Q10)
285 (B H1) 332 (W N7) 333 (B A15) 286 (W K2) 334 (W M3) 335 (B resign) Game has ended.
Oct 27 04:22:11 instance-group-5-g0gr startup-script[1417]: INFO startup-script: Score: W+Resign
Oct 27 04:22:11 instance-group-5-g0gr startup-script[1417]: INFO startup-script: Winner: white
Oct 27 04:22:12 instance-group-5-g0gr startup-script[1417]: INFO startup-script: Uploading match: 11ba0232743941b180155c90d33151d2.sgf for networks 2da87ea8da0f54e87b70159e6bb82811b61d1c31091b6e01
9fbc62aeaa803b9c and 5c81864b8b9ac195a24e58ceb5fd9ddc9465ca6b6a0d774f578ee16e0f56d919
Oct 27 04:22:12 instance-group-5-g0gr startup-script[1417]: INFO startup-script: Match data 61387dec13662e66be0c5d9f94819043548a2e36770282e035817c9dbedf36f stored in database
Oct 27 04:22:12 instance-group-5-g0gr startup-script[1417]: INFO startup-script: 229 game(s) (193 self played and 36 matches) played in 1188 minutes = 311 seconds/game, 1630 ms/move, last game too
k 794 seconds.
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script: 287 (B E9) {
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "black_hash": "5c81864b8b9ac195a24e58ceb5fd9ddc9465ca6b6a0d774f578ee16e0f56d919",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "black_hash_gzip_hash": "3580befd0253e00332c635e820bdd10817a2c2a03599099ccbc501f58557cff8",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "cmd": "match",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "minimum_autogtp_version": "16",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "minimum_leelaz_version": "0.15",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "options": {
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:     "noise": "false",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:     "playouts": "0",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:     "randomcnt": "0",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:     "resignation_percent": "5",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:     "visits": "1600"
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script: },
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "options_hash": "0ef1f9ca44d9e3b51797cfa45367747010753eaab49a34e9a35d28a9866ca12b7703ec",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "random_seed": "6782596068867827701",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "required_client_version": "16",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "white_hash": "2da87ea8da0f54e87b70159e6bb82811b61d1c31091b6e019fbc62aeaa803b9c",
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script:   "white_hash_gzip_hash": "11f2fdf61a20b3b52e7b4052e2819f2757fc514f45e9ab0f3e43bb69856b65e0"
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script: }
Oct 27 04:22:13 instance-group-5-g0gr startup-script[1417]: INFO startup-script: Got new job: match
Oct 27 04:22:15 instance-group-5-g0gr startup-script[1417]: INFO startup-script: 288 (W F9) first network: 5c81864b8b9ac195a24e58ceb5fd9ddc9465ca6b6a0d774f578ee16e0f56d919.
Oct 27 04:22:15 instance-group-5-g0gr startup-script[1417]: INFO startup-script: second network 2da87ea8da0f54e87b70159e6bb82811b61d1c31091b6e019fbc62aeaa803b9c.
Oct 27 04:22:27 instance-group-5-g0gr startup-script[1417]: INFO startup-script: 289 (B G9) 290 (W C10) 291 (B E10) 292 (W D12) 293 (B E12) 294 (W N13) 295 (B A14) Engine has started.
Oct 27 04:22:27 instance-group-5-g0gr startup-script[1417]: INFO startup-script: time settings 0 1 0
Oct 27 04:22:27 instance-group-5-g0gr startup-script[1417]: INFO startup-script: Thinking time set.
Oct 27 04:22:38 instance-group-5-g0gr startup-script[1417]: INFO startup-script: 296 (W B15) 297 (B P13) 298 (W O17) 299 (B A13) 300 (W B12) 301 (B M7) Engine has started.
Oct 27 04:22:38 instance-group-5-g0gr startup-script[1417]: INFO startup-script: time settings 0 1 0
Oct 27 04:22:38 instance-group-5-g0gr startup-script[1417]: INFO startup-script: Thinking time set.
]
```

If you want the pictures only, here they are (directly added from the drive to this doc) :  
<https://drive.google.com/drive/folders/1kqr5QRRiOSdC2N4CtetUtyOBcBFQ-d6E?usp=sharing>

Quick facts before starting :

- This free trial is entirely and totally free of charge, without any obligatory end of free credit condition (this is the case for Google Cloud, i don't know about the other services)
- With powerful GPU like the Tesla V100, you will be able to produce 16 games per hour (all 5% resign) for leela-zero 40 blocksx256, which is much more than what a public card like a GTX 1080 Ti can do.
- To prevent abuse, spam, robots, etc, an id check will be performed with a valid credit card, but you will not be charged at all during the free trial
- These instructions will create an entirely automated leela-zero autogtp VM instance thanks to a startup-script in metadata : after setting it up correctly, it will not require any operation and will install all needed packages, compile and run leela-zero with autogtp, and will produce games automatically
- The instance uses cloud resource, not your personal machine
- The instance is running on a server : it will stay online independently from you (even if your computer is powered off)
- This instance will be Preemptible : it uses cloud resource that are not always available, causing it to be 60% cheaper (aka. to consume the free trial credit much slower) but the instance is ephemere : after 24 hours max it cannot "live" anymore and will be terminated by preemptible use rules
- The Preemptible terminations will not be a problem though, because our instance will be in a managed instance group
- Our managed instance group will automatically create our first instance, install all needed packages on it (which takes exactly 10 minutes), then automatically reboot it and automatically starting to produce games with autogtp
- Everytime our instance "dies" (max 24 hours because of preemptibility, or if you manually delete it), our managed instance group will automatically delete our "dead" instance and automatically recreate a new "child" preemptible instance (a new one, does not contain old data of the "parent" instance)
- Then, our managed instance group will automatically restart our new "child" instance, install all needed packages including leela-zero (takes exactly 10 minutes), then auto reboot, and then at reboot automatically start to produce games with autogtp, until the "child" instance "dies", giving "birth" by the group to a new "child of the child" instance, etc.
- The exception to this automated recreation+autostart by the instance group is for scheduled maintainance by Google (rare, once every few weeks) which will require you to manual restart the instance (it takes 1 minute), then the auto-start script will handle everything again.



## Cloud Companies :

We are not affiliated with any cloud company, and we provide these instructions as they are a free of charge way to help public contributing to leela-zero project.

We are thankful to these cloud companies for giving us these free trial opportunities.

The instructions below are for Google Cloud Free Trial as Google is a widespread company, but much of the documentation here can be used if slightly modified for other cloud companies offering similar cloud free trials that include a GPU (Microsoft Azure, Oracle Cloud, etc.)

## Video Tutorial

---

As an interactive help to text instructions, a video tutorial is provided here : youtube....(link will be added later)

## Start the Google Cloud Free Trial (id check)

---

You'll first need to start your google cloud free trial here :

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

The cloud public resource available being limited, especially when it comes to powerful GPU like the Tesla V100, in order to prevent abuse, spam, robots, multi accounts, etc, Google will ask to check your id with a valid credit card, but you will not be charged anything at all, even when your free trial credit ends.

You can check by yourself here :

<https://cloud.google.com/free/docs/frequently-asked-questions>

And this free trial does not force you to subscribe to anything at all. It is indeed entirely free of charge.

Also, just like for Google Colab, please do NOT try to use multiple accounts in any way, as Google won't hesitate to ban you, as it happened for google Colab users. If you want to help leela-zero, please rather try to spread these instructions so that more people join us in this contributed effort.

## Getting Started

---

If you are not redirected, after your id is successfully checked, you will be redirected here : <https://cloud.google.com/getting-started>

To navigate in Google Cloud menus, click on the sandwich bar (in the top left) in the top left.

## Free credit consumption estimation :

---

You can check your remaining free credit, in the billing menu of the sandwich bar (in the top left).

Or, you can alternately use this link :

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

At start, a free 300\$ (or 257€) free trial credit is given to you, free of charge. With the settings later explained, our instance will consume 0,773\$/hour with a Tesla V100

This will allow us to use it for around 390 hours, which would produce around 6000 games leela-zero 40 blocks per free trial (5% resign only, as 0% resign games take much more time), or around 4000 games per free trial if you include a mix of -r5 and -r0 games, all this entirely free of charge !

These 300\$ can be used for around 16,1 days for a h24 7/7 use of a Tesla V100, entirely free of charge !

# Quota Request for GPU (all regions) and Preemptible CPUs (all regions)

---

Starting from november 2018, google cloud policy changed :

Free trial accounts now have a default quota of GPU set to 0, which means if we try to create an instance with a GPU in any region we'll get the following error :

**Quota 'GPUS\_ALL\_REGIONS' exceeded. Limit: 0.0 globally.**

To solve this, we now need to send a quota increase request to google cloud support, and they'll answer you by email within 48 hours

Note : preemptible CPUs are also set to 0 so you need to increase them to 24 or more as we will see below

In google cloud console, go to the sandwich bar on top left, and click on : IAM & Admin -> quotas

Alternatively, you can use this link : <https://console.cloud.google.com/iam-admin/quotas>

Then in the quota page :

- in services select only "Compute Engine API"
- and on "Limit" click on the arrow to sort by lowest limit first (starting from 0 it will be easier to see the quota increases we need)

Scroll down and turn pages until you find GPU (all regions) - Global

As you can see here GPU (all regions) - Global quota is set to 0 :

And preemptible CPUs are also all set to 0 in every region

Google Cloud Platform

My First Project

IAM & admin

IAM

Identity & Organization

Organization policies

Quotas

Service accounts

Labels

Privacy & Security

Settings

Cryptographic keys

Identity-Aware Proxy

Roles

Audit Logs

Manage resources

Quotas

EDIT QUOTAS

Quota type

All quotas

Service

Compute Engine API

Metric

All metrics

Location

All locations

Clear

	Service	Location	Current Usage	7 Day Peak Usage	Limit
<input type="checkbox"/>	Compute Engine API Preemptible CPUs	us-central1		0	0
<input type="checkbox"/>	Compute Engine API Preemptible CPUs	us-east1		0	0
<input type="checkbox"/>	Compute Engine API Preemptible CPUs	us-east4		0	0
<input type="checkbox"/>	Compute Engine API Preemptible CPUs	us-west1		0	0
<input type="checkbox"/>	Compute Engine API Preemptible CPUs	us-west2		0	0
<input type="checkbox"/>	Compute Engine API GPUs (all regions)	Global		0	0
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	asia-east1	0	1	1
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	asia-east2	0	1	1
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	asia-northeast1	0	1	1
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	asia-south1	0	1	1
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	asia-southeast1	0	1	1
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	australia-southeast1	0	1	1
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	eu-west-1	0	1	1
<input type="checkbox"/>	Compute Engine API NVIDIA K80 GPUs	eu-west-2	0	1	1

The steps are now what you will see in the screenshots below :

- 1) go back to page 1 and select all preemptible CPUs
- 2) go again to page 2 and select the remaining preemptible CPUs, as well as GPU (all regions) - Global
- 3) click on “edit quotas”
- 4) Write your personal information (name, mail, etc.)
- 5) input the new quota limit :
  - For GPU all regions - Global : set to 1
  - For Preemptible CPUs : set them to any number that is higher than 8 (i chose 24 in this example because the non preemptible CPUs have this limit)
- 6) write your personal information (name, mail, etc.)
- 7) You can write a smalll description like i did if you want, then click on “Done”
- 8) Click on “Submit Request”

As you can see below in the screenshots :

Google Cloud Platform

My First Project

IAM & admin

IAM

Identity & Organization

Organization policies

Quotas

Service accounts

Labels

Privacy & Security

Settings

Cryptographic keys

Identity-Aware Proxy

Roles

Audit Logs

Manage resources

Quotas

EDIT QUOTAS

<input type="checkbox"/>	Compute Engine API Preemptible Local SSD (GB)	us-east4	—	0
<input type="checkbox"/>	Compute Engine API Preemptible Local SSD (GB)	us-west1	—	0
<input type="checkbox"/>	Compute Engine API Preemptible Local SSD (GB)	us-west2	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	asia-east1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	asia-east2	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	asia-northeast1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	asia-south1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	asia-southeast1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	australia-southeast1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	europa-north1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	europa-west1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	europa-west2	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	europa-west3	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	europa-west4	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	northamerica-northeast1	—	0
<input checked="" type="checkbox"/>	Compute Engine API Preemptible CPUs	southamerica-east1	—	0

Rows per page: 50 1 - 50 of 680

Google Cloud Platform

My First Project

IAM & admin

IAM

Identity & Organization

Organization policies

Quotas

Service accounts

Labels

Privacy & Security

Settings

Cryptographic keys

Identity-Aware Proxy

Roles

Audit Logs

Manage resources

Quotas

EDIT QUOTAS

Quota type

Service

Metric

Location

Clear

Service	Location	Current Usage	7 Day Peak Usage	Limit
<input checked="" type="checkbox"/> Compute Engine API Preemptible CPUs	us-central1	—	0	0
<input checked="" type="checkbox"/> Compute Engine API Preemptible CPUs	us-east1	—	0	0
<input checked="" type="checkbox"/> Compute Engine API Preemptible CPUs	us-east4	—	0	0
<input checked="" type="checkbox"/> Compute Engine API Preemptible CPUs	us-west1	—	0	0
<input checked="" type="checkbox"/> Compute Engine API Preemptible CPUs	us-west2	—	0	0
<input checked="" type="checkbox"/> Compute Engine API GPUs (all regions)	Global	—	0	0
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	asia-east1	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	asia-east2	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	asia-northeast1	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	asia-south1	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	asia-southeast1	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	australia-southeast1	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	europa-north1	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	europa-west1	0	—	1
<input type="checkbox"/> Compute Engine API NVIDIA K80 GPUs	europa-west2	0	—	1

19 quotas selected

Edit quotas

Name

Email

Phone

Next



Google Cloud Platform

My First Project

IAM & admin

IAM

Identity & Organization

Organization policies

Quotas

Service accounts

Labels

Privacy & Security

Settings

Cryptographic keys

Identity-Aware Proxy

Roles

Audit Logs

Manage resources

Quotas

EDIT QUOTAS

<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-central1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east4	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west2	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API CPUs	asia-east1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-east2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-northeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-south1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	australia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-central-1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-3	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-4	<div></div>	0	– ?	24

Rows per page: 50 451 - 500 of 680

19 quotas selected

Compute Engine API

Quota: GPUs (all regions)  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
1

Quota: Preemptible CPUs - us-west2  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - us-west1  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - us-east4  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - us-east1  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - us-central1  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.

Google Cloud Platform

My First Project

IAM & admin

IAM

Identity & Organization

Organization policies

Quotas

Service accounts

Labels

Privacy & Security

Settings

Cryptographic keys

Identity-Aware Proxy

Roles

Audit Logs

Manage resources

Quotas

EDIT QUOTAS

<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-central1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east4	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west2	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API CPUs	asia-east1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-east2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-northeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-south1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	australia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-central-1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-3	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	eu-west-4	<div></div>	0	– ?	24

Rows per page: 50 451 - 500 of 680

New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - europe-west3  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - europe-west4  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - northamerica-northeast1  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Quota: Preemptible CPUs - southamerica-east1  
New quota limit  
Enter a new quota limit. Your request will be sent to your service provider for approval.  
24

Request description  
Required  
i want to use tesla p100 and tesla v100 in preemptible VM instances with the free trial (pay as you go) thanks

Done Cancel

Submit request Back

Google Cloud PlatformMy First Project

Navigation menuIAMIdentity & OrganizationOrganization policiesQuotasService accountsLabelsPrivacy & SecuritySettingsCryptographic keysIdentity-Aware ProxyRolesAudit LogsManage resources

QuotasEDIT QUOTAS

<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-central1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east4	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west2	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API CPUs	asia-east1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-east2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-northeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-south1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	australia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-north1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west3	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west4	<div></div>	0	– ?	24

Rows per page: 50451 - 500 of 680

19 quotas selected

Compute Engine API

Submit requestBack

Google Cloud PlatformMy First Project

IAM & adminIAMIdentity & OrganizationOrganization policiesQuotasService accountsLabelsPrivacy & SecuritySettingsCryptographic keysIdentity-Aware ProxyRolesAudit LogsManage resources

QuotasEDIT QUOTAS

<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-central1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-east4	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west1	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API Regional managed instance groups	us-west2	<div></div>	0	– ?	20
<input type="checkbox"/>	Compute Engine API CPUs	asia-east1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-east2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-northeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-south1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	asia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	australia-southeast1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-north1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west1	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west2	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west3	<div></div>	0	– ?	24
<input type="checkbox"/>	Compute Engine API CPUs	europa-west4	<div></div>	0	– ?	24

Rows per page: 50451 - 500 of 680

Edit quotas

Compute Engine API

Thank you for submitting Case # (ID ██████████) to Google Cloud Platform support for the following quotas:

- Change GPUs (all regions) from 0 to 1
- Change Preemptible CPUs - us-west2 from 0 to 24
- Change Preemptible CPUs - us-west1 from 0 to 24
- Change Preemptible CPUs - us-east4 from 0 to 24
- Change Preemptible CPUs - us-east1 from 0 to 24
- Change Preemptible CPUs - us-central1 from 0 to 24
- Change Preemptible CPUs - asia-east1 from 0 to 24
- Change Preemptible CPUs - asia-east2 from 0 to 24
- Change Preemptible CPUs - asia-northeast1 from 0 to 24
- Change Preemptible CPUs - asia-south1 from 0 to 24
- Change Preemptible CPUs - asia-southeast1 from 0 to 24
- Change Preemptible CPUs - australia-southeast1 from 0 to 24
- Change Preemptible CPUs - europa-north1 from 0 to 24
- Change Preemptible CPUs - europa-west1 from 0 to 24
- Change Preemptible CPUs - europa-west2 from 0 to 24
- Change Preemptible CPUs - europa-west3 from 0 to 24
- Change Preemptible CPUs - europa-west4 from 0 to 24
- Change Preemptible CPUs - northamerica-northeast1 from 0 to 24
- Change Preemptible CPUs - southamerica-east1 from 0 to 24

Your request is being processed and you should receive an email confirmation for your request. Should you need further assistance, you can respond to that email.

Wait for google cloud support answer within 48 hours  
It should be enough to solve the issue

But if the issue still remains, email google support again and please also tell us on github so that we update these instructions again

Note : you may find extra information on this page if you want :  
<https://cloud.google.com/compute/quotas>

# Create an instance template

In google cloud console, go to the sandwich bar on top left, and click on : Compute Engine -> Instance templates



Alternatively, you can use this link : <https://console.cloud.google.com/compute/instanceTemplates/>

An instance template is a recording of the settings we want every new instance to be automatically recreated with. Click on blue button "Create instance template" :

Then, choose these settings :

As you can see in the 4 screenshots below :

Google Cloud Platform

My First Project

Compute Engine

Committed use discounts

VM instances

Instance templates

Sole tenant nodes

Disks

Snapshots

Images

TPUs

Instance groups

Metadata

Health checks

Zones

Network endpoint groups

Operations

Quotas

Security scans

Settings

Marketplace

Create an instance template

Describe a VM instance once and then use that template to create groups of identical instances

Name

instance-template-try3

Machine type

Customize to select cores, memory and GPUs.

Cores

4

vCPU

1 - 8

Memory

5.25

GB

3.6 - 26

Extend memory

CPU platform

Automatic

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.

Number of GPUs

GPU type

1

NVIDIA Tesla V100

Machines with GPUs can't migrate on host maintenance

Choosing a machine type

Upgrade your account

to create instances with up to 96 cores

Container

Deploy a container image to this VM instance.

Boot disk

New 10 GB standard persistent disk

These are estimated costs for a VM instance created using this template

You have €62.598489 free trial credits remaining

\$564.58 monthly estimate

That's about \$0.773 hourly

Show costs for location

US

Details

☰

Google Cloud Platform

My First Project ▾

🔍

🖨️

Compute Engine

📊

Committed use discounts

👤

VM instances

📄

Instance templates

👤

Sole tenant nodes

💾

Disks

📷

Snapshots

🖼️

Images

🔧

TPUs

🏠

Instance groups

☰

Metadata

🏠

Health checks

📶

Zones

🌐

Network endpoint groups

🕒

Operations

📄

Quotas

🛡️

Security scans

⚙️

Settings

🛒

Marketplace

⏪

⬅️

Create an instance template

🔗

Boot disk

📄

New 10 GB standard persistent disk

🖼️

Image

Ubuntu 18.04 LTS

Change

🔗

Identity and API access

🔗

Service account

Compute Engine default service account

🔗

Access scopes

☒ Allow default access

☐ Allow full access to all Cloud APIs

☐ Set access for each API

🔗

Firewall

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

☒ Allow HTTP traffic

☒ Allow HTTPS traffic

Management

Security

Disks

Networking

Sole Tenancy

Description (Optional)

🔗

Labels (Optional)

+ Add label

Automation

🔗

Startup script (Optional)

You can choose to specify a startup script that will run when your instance boots up or restarts. Startup scripts can be used to install software and updates, and to ensure that services are running within the virtual machine. [Learn more](#)

```
#!/bin/bash
PKG_OK=$(dpkg-query -W --showformat='${Status}\n' glances|grep "install ok installed")
echo Checking for glanceslib: $PKG_OK
if [ "" == "$PKG_OK" ]; then
```



Google Cloud Platform

My First Project

Compute Engine

Committed use discounts

VM instances

Instance templates

Sole tenant nodes

Disks

Snapshots

Images

TPUs

Instance groups

Metadata

Health checks

Zones

Network endpoint groups

Operations

Quotas

Security scans

Settings

Marketplace

Create an instance template

Automation

Startup script (Optional)

You can choose to specify a startup script that will run when your instance boots up or restarts. Startup scripts can be used to install software and updates, and to ensure that services are running within the virtual machine. [Learn more](#)

```
#!/bin/bash
PKG_OK=$(dpkg-query -W --showformat='${Status}\n' glances|grep "install ok installed")
echo Checking for glanceslib: $PKG_OK
if [ "" == "$PKG_OK" ]; then
  echo "No glanceslib. Setting up glanceslib and all other leela-zero packages."
  sudo apt-get update && sudo apt-get -y upgrade && sudo apt-get -y dist-upgrade && sudo add-apt-repository -y ppa:graphics-drivers/ppa && sudo apt-get update && sudo apt-get -y install nvidia-driver-410 linux-headers-generic nvidia-opengl-dev && sudo apt-get -y install clinfo cmake git libboost-all-dev libopenblas-dev zlib1g-dev build-essential qtbase5-dev qttools5-dev qttools5-dev-tools libboost-dev libboost-program-options-dev opengl-headers ocl-icd-libopengl1 ocl-icd-opengl-dev qt5-default qt5-qmake curl && git clone https://github.com/gcp/leela-zero && cd leela-zero && git checkout next && git pull && git clone https://github.com/gcp/leela-zero && git submodule update --init --recursive && mkdir build && cd build && cmake .. && cmake --build . && cd ../autogtp && cp ../build/autogtp/autogtp . && cp ../build/leelaz . && sudo apt-get -y install glances zip && sudo apt-get clean && sudo reboot
else
  sudo -i && cd /leela-zero/autogtp && ./autogtp -g 2
fi
```

Metadata (Optional)

You can set custom metadata for an instance or project outside of the server-defined metadata. This is useful for passing in arbitrary values to your project or instance that can be queried by your code on the instance. [Learn more](#)

Key

Value

X

+ Add item

Availability policy

Preemptibility

A preemptible VM costs much less, but lasts only 24 hours. It can be terminated sooner due to system demands. [Learn more](#)

On

An instance group created from a template with a preemptible VM will not behave like other instance groups. [Learn more](#)

← Create an instance template

```
else
    sudo -i && cd /leela-zero/autogtp && ./autogtp -g 2
fi
```

### Metadata (Optional)

You can set custom metadata for an instance or project outside of the server-defined metadata. This is useful for passing in arbitrary values to your project or instance that can be queried by your code on the instance. [Learn more](#)


Key	Value
<a href="#">+ Add item</a>	

### Availability policy

## Preemptibility

A preemptible VM costs much less, but lasts only 24 hours. It can be terminated sooner due to system demands. [Learn more](#)

On ▼

 An instance group created from a template with a preemptible VM will not behave like other instance groups. [Learn more](#)

### Automatic restart

Compute Engine can automatically restart VM instances if they are terminated for non-user-initiated reasons (maintenance event, hardware failure, software failure, etc.)

### On host maintenance

When Compute Engine performs periodic infrastructure maintenance it can migrate your VM instances to other hardware without downtime

Terminate VM instance

⌵ Less

You can create this instance template free of charge

Create Cancel

Equivalent REST or command line

- template name : any name you want
- on machine type, click "customize"
- 4 vcpu / 5.25 GB ram
- 1 GPU : Tesla V100
- untick "Extend memory"
- Boot Disk : click on "change" and choose Ubuntu 18.04 LTS with 10GB HDD (standard persistent disk)
- Firewall : allow http/https Then click on "Management, security, disks, networking, sole tenancy" : These new options will appear :
- Automation : in startup script, copy paste all the script below :

CHOOSE one of the 2 scripts below :

(updated scripts : 7 december 2018 versions now work again with google cloud, shrinked) :



- -> OPTION A : if you want to use MASTER (more stable) branch, choose this script :

```
#!/bin/bash
PKG_OK=$(dpkg-query -W --showformat='${Status}\n' glances|grep "install ok installed")
echo Checking for glanceslib: $PKG_OK
if [ "" == "$PKG_OK" ]; then
    echo "No glanceslib. Setting up glanceslib and all other leela-zero packages."
    sudo -i && sudo add-apt-repository -y ppa:graphics-drivers/ppa && sudo apt-get update && sudo apt-get -y -f install nvidia-driver-410 libboost-dev
libboost-program-options-dev libboost-filesystem-dev opencv4-headers ocl-icd-libopencv4 ocl-icd-opencv4-dev zlib1g-dev clinfo cmake qt5-default qt5-qmake
curl git zip glances && sudo add-apt-repository -y ppa:ubuntu-toolchain-r/test && sudo apt-get update && sudo apt-get -y install gcc-8 g++-8 && sudo update-alternatives --install /usr/bin/gcc gcc
/usr/bin/gcc-8 40 --slave /usr/bin/g++ g++ /usr/bin/g++-8 && git clone https://github.com/leela-zero/leela-zero && cd leela-zero && git submodule update --init --recursive
&& mkdir build && cd build && cmake .. && cmake --build . && cp leelaz autogtp && sudo reboot
else
    sudo -i && cd /leela-zero/build/autogtp && ./autogtp -g 2
fi
```

- -> OPTION B : else, if you want to use NEXT (latest improvements) branch, choose this script :

```
#!/bin/bash
PKG_OK=$(dpkg-query -W --showformat='${Status}\n' glances|grep "install ok installed")
echo Checking for glanceslib: $PKG_OK
if [ "" == "$PKG_OK" ]; then
    echo "No glanceslib. Setting up glanceslib and all other leela-zero packages."
    sudo -i && sudo add-apt-repository -y ppa:graphics-drivers/ppa && sudo apt-get update && sudo apt-get -y -f install nvidia-driver-410 libboost-dev
libboost-program-options-dev libboost-filesystem-dev opencv4-headers ocl-icd-libopencv4 ocl-icd-opencv4-dev zlib1g-dev clinfo cmake qt5-default qt5-qmake
curl git zip glances && sudo add-apt-repository -y ppa:ubuntu-toolchain-r/test && sudo apt-get update && sudo apt-get -y install gcc-8 g++-8 && sudo update-alternatives --install /usr/bin/gcc gcc
/usr/bin/gcc-8 40 --slave /usr/bin/g++ g++ /usr/bin/g++-8 && git clone https://github.com/leela-zero/leela-zero -b next && cd leela-zero && git submodule update --init
--recursive && mkdir build && cd build && cmake .. && cmake --build . && cp leelaz autogtp && sudo reboot
else
    sudo -i && cd /leela-zero/build/autogtp && ./autogtp -g 2
fi
```

What this script makes is basically :

- if glances package is not installed (i.e: if it is first boot of this new instance) : then install all needed packages (system,driver,leela-zero), then reboot
- if glances package is installed (i.e: if it is 2nd or more boot of this new instance) : then run autogtp

Note : please make sure you dont add a CAPS at the “fi” while editing this google doc.

- Availability Policy : Preemptibility ON

This is important as it will reduce cost (free credit consumption) by 60%

(0.773\$/hour instead of 1.8\$/hour for a Tesla V100 !)

note : in nvidia-driver-410 , replace 410 version number by whatever latest version number you find here

:<https://launchpad.net/~graphics-drivers/+archive/ubuntu/ppa>

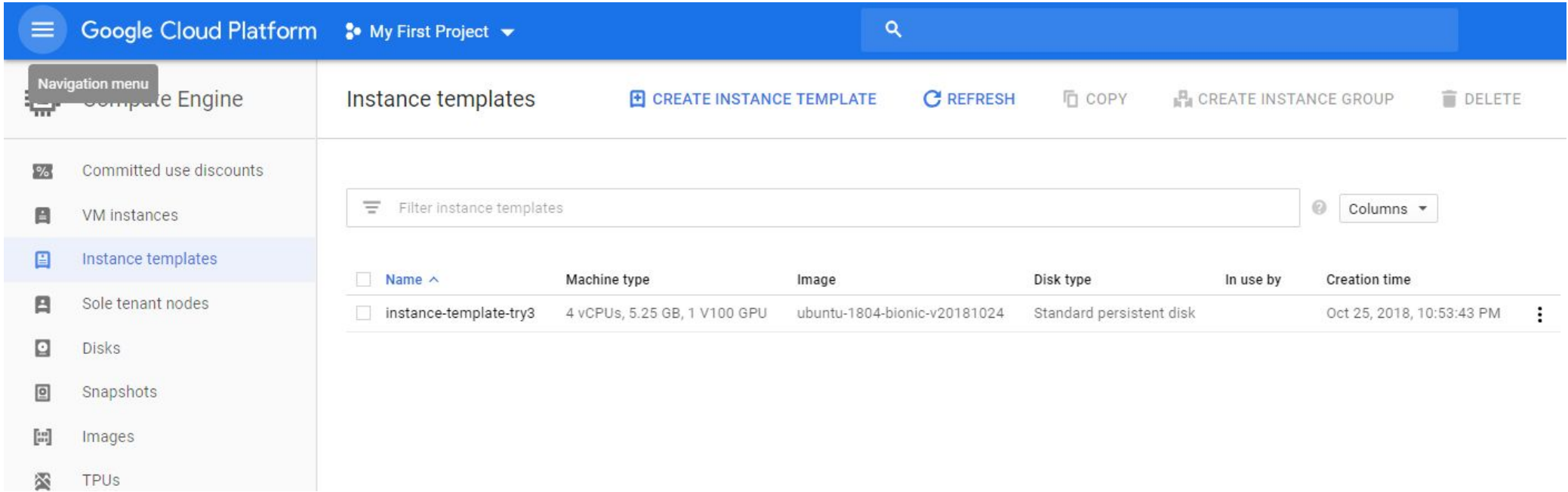
note 2 : an instance template cannot be modified/updated, if you want for example to change nvidia-driver-410 in the script, change username, or modify anything in the instance template, you will have to create a new instance template, then delete your current instance group, and create a new instance group with your latest instance template.

note 3 : i chose ./autogtp -g 2 in the script as it is 25% faster on a Tesla V100 from my extensive long time tests



note 4 : credits for the script go to (i modified it) : <https://stackoverflow.com/a/10439058>

After instance template finishes creating, you will get a screen like that :



## Preparations : Check the regions and subregions that can provide a Tesla V100 :

In this part, we will only be checking which regions have a Tesla V100 available, but we will NOT create any instance. When we are finished checking all regions, we will click on “Cancel” and exit the VM instance page without creating anything.

This step will be fast, but important in order to create a group in the correct regions. In another browser tab, if you're still in Compute engine (else use the top left sandwich bar), on the left panel you will see "instance groups", click on it.

Alternatively, you can use this link : <https://console.cloud.google.com/compute/instances>

To know which regions and subregions have a Tesla V100, click on "create instance" (we will not actually create an instance, but just scroll through regions and subregions, then exit it) Then click on "custom" machine.

And for every region, if you can add a GPU, add 1 GPU, and check if among these GPU a Tesla V100 is providable, as shown in the screenshot below :

Google Cloud Platform

My First Project

← 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

+

Marketplace

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

Name

instance-1

Region

us-central1 (Iowa)

Zone

us-central1-a

Machine type

Customize to select cores, memory and GPUs.

Cores

1

vCPU

1 - 8

Memory

3,75

GB

1 - 6.5

☐ Extend memory

CPU platform

Automatic

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

NVIDIA Tesla P4

NVIDIA Tesla P4 Virtual Workstation

NVIDIA Tesla V100

Machines with GPUs can't migrate

[Choosing a machine type](#)

[Upgrade your account](#) to create instances with up to 96 cores

Repeat this step for every subregion inside a region and similarly for every other regions : At the time of these instructions, these are the regions and subregions that can provide a Tesla V100 :

- us-central1 (iowa) (has Tesla V100 in 3 subregions : a f b)
- eu-west-4 (netherlands) (has Tesla V100 in 3 subregions : a b c)
- us-west-1 (oregon) (has Tesla V100 in 2 subregions : b a)
- asia-east-1 (Taiwan) (has Tesla V100 in 1 subregion : c)

But please remember that these can change in the future for the better or the worse, this is why all the methodology is described here.

Then, when you saw all regions and subregions, go back at the bottom of the page and click on “Cancel” (we will not create a VM instance but an instance group)

We will now use the information we gathered to actually create our instances with an instance group.

## Create a managed instance group :

This managed instance group will take care of automatically recreating a new instance every time preemptible use shuts it down (frequent), or if it is deleted by you (but NOT for scheduled maintenance which will require a manual restart of the instance, much more rare).

If you're still in Compute engine (else use the top left sandwich bar), on the left panel you will see "instance groups", click on it. Alternatively, you can use this link : <https://console.cloud.google.com/compute/instanceGroups/>

Click on "Create instance group"

Then, choose these settings :

As you can see in the screenshots below :



Google Cloud Platform

My First Project

Compute Engine

Committed use discounts

VM instances

Instance templates

Sole tenant nodes

Disks

Snapshots

Images

TPUs

Instance groups

Metadata

Health checks

Zones

Network endpoint groups

Operations

Quotas

Security scans

Settings

Marketplace

Create a new instance group

Use an instance group when configuring a load-balancing backend service or to group VM instances. [Learn more](#)

Name

instance-group-1

Description

Location

Multi-zone groups span multiple zones which assures higher availability [Learn more](#)

Single-zone

Multi-zone

Region

us-central1

Zones

us-central1-a

us-central1-b

us-central1-c

us-central1-f

Less

Specify port name mapping

Instance template

instance-template-try3

An instance group created from a template with a preemptible VM will not behave like other instance groups. [Learn more](#)

Autoscaling

Off

Number of instances

Google Cloud Platform

My First Project

Compute Engine

Committed use discounts

VM instances

Instance templates

Sole tenant nodes

Disks

Snapshots

Images

TPUs

Instance groups

Metadata

Health checks

Zones

Network endpoint groups

Operations

Quotas

Security scans

Settings

Marketplace

Create a new instance group

us-central1-c

us-central1-f

Less

Specify port name mapping

(Optional)

Instance template

instance-template-try3

An instance group created from a template with a preemptible VM will not behave like other instance groups.

Learn more

Autoscaling

Off

Number of instances

1

To maximize availability, the number of instances should be at least equal to the number of zones. Additional instances will be placed in different zones.

Learn more

Autohealing

VMs in the group are recreated as needed. You can use a health check to recreate a VM if the health check finds the VM unresponsive. If you do not select a health check, VMs are recreated only when stopped.

Learn more

Health check

No health check

Initial delay

300

seconds

You will be billed for VM instances in this group.

Compute Engine pricing

Create

Cancel

Equivalent REST or command line

- group name : any name you want
- location : multi-zone, then click on more As explained earlier choose only one of the regions that can provide a Tesla V100 in at least one of its subregions (zones). The instance group that we are creating will continuously be in charge of trying to recreate and restart the instance in any of the subregions (zones) that we will choose until it finds one where hardware is providable on the cloud. This is why it is preferred to choose a region that can provide a Tesla V100 in many of its subregions. In the video tutorial, i chose us-central1 (iowa) in this example. Then click on "configure zones"
- zone (subregion) : tick only the subregions (zones) that can provide a Tesla V100 In this example i ticked subregions a f and b, and unticked all other subregions.
- instance template : choose the template that we created earlier. This template will tell what the group what hardware to use, and also include a startup-script that will automatically install all packages needed, then install leela-zero, and then at every boot will start autogtp. It is entirely automated
- autoscaling : off
- number of instances : 1 As free trial accounts have a GPU quota of 1, we will not be able to create anymore simultaneously running instances with a GPU.
- Health check : no

Click on blue button "create". After some time, we get a screen like this :

Google Cloud Platform

My First Project

Navigation menu

Compute Engine

Committed use discounts

VM instances

Instance templates

Sole tenant nodes

Disks

Snapshots

Images

TPUs

Instance groups

Instance groups

CREATE INSTANCE GROUP

REFRESH

EDIT

DELETE

Filter instance groups

Columns

<input type="checkbox"/>	Name ^	Zone	Instances	Template	Creation time	Recommendation	Autoscaling	In use by
<input type="checkbox"/>	<input checked="" type="checkbox"/> instance-group-1	us-central1 (3/4 zones)	1	instance-template-try3	Oct 25, 2018, 10:55:26 PM		Off	

if you go back to "VM instances" (<https://console.cloud.google.com/compute/instances>) :

You can see that a new instance has been created by the managed instance group This instance will automatically update, upgrade and install all system packages then compile leela-zero and reboot, and then automatically run autogtp.

Google Cloud Platform

My First Project

Compute Engine

Committed use discounts

VM instances

Instance templates

Sole tenant nodes

Disks

Snapshots

Images

TPUs

VM instances

CREATE INSTANCE

IMPORT VM

REFRESH

START

STOP

RESET

DELETE

Filter VM instances

Columns

<input type="checkbox"/>	Name ^	Zone	Recommendation	Internal IP	External IP	Connect
<input type="checkbox"/>	<input checked="" type="checkbox"/> instance-group-1-00tv	us-central1-b		10.128.0.2 (nic0)	35.239.125.175	SSH

## The last obligatory thing you need to know :

How to manually restart an instance after it has been stopped due to scheduled maintainance :

As explained earlier, Preemptible instances are much cheaper (60% less) but can't last more than 24 hours, and can be stopped sooner (frequent).

This case is not at all a problem though, because, every time the instance is stopped by preemptible use or deleted by yourself, the managed instance group will immediately and continuously automatically keep trying to recreate and restart our instance in any of the subregions(zones) we selected earlier that have a Tesla V100 available, until it succeeds. You don't have to do anything at all for that.

However, sometimes both preemptible and non preemptible instances will get stopped by Google Cloud due to scheduled maintainance events.

note : when your instance is stopped, the free credit stops being consumed, so you don't have to worry about failing to restart a stopped instance, you can do it anytime later.

This case is rare (maybe once every few weeks), but Preemptible instances can't be automatically restarted after that, unlike non Preemptible instances.

If that rare case happens, the user (you) will have to manually recreate the instance group.



To do that, go on "instance groups" page in Compute engine : <https://console.cloud.google.com/compute/instanceGroups/>

Then click on the 3 dots at the right of your instance group (or tick the box at the left of your group name), and choose “Delete”.

Deleting the group will also automatically delete all VM instances managed by this group (in our case only 1), but will not delete other instances that are not related to this group (if you have other projects, etc).

Then just recreate our instance group as we saw earlier (it takes 2 minutes) and you’ll be good to go for a long time again !

note : recreating the instance will start consuming the free trial credit again, as explained in the pre-start message.

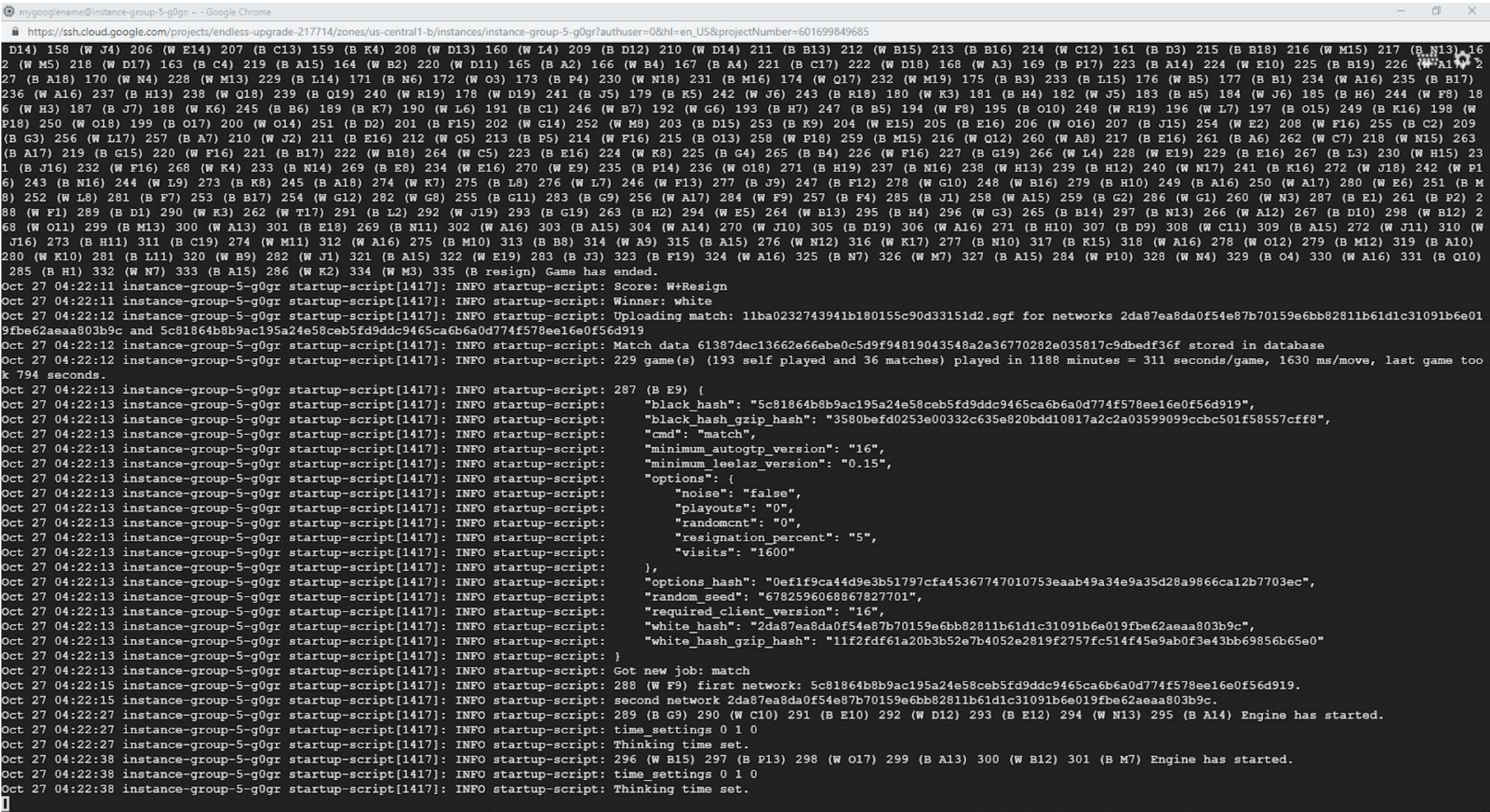
Finally, whenever you want, you can willingly choose to stop contributing for some time by deleting the instance group, then whenever you want start contributing again.

This concludes the obligatory steps ! You are now a Tesla V100 leela-zero contributor ! You can exit chrome, power off your computer, you don't have to do anything ! It is entirely automated (except for scheduled maintainance events requiring a manual restart as explained earlier).

## Optional extra steps

The following instructions are optional, but they may interest you.

It includes for example how to easily access your instance and see games being produced in live mode



You can find them here : <https://docs.google.com/document/d/1SobPo-D9Ry1jBlt4CChvsNIYfaFk9WJDI9IEAgeAMJI/edit?usp=sharing>