Nvidia MIG

Official Documentation

- https://docs.nvidia.com/datacenter/tesla/mig-user-guide/#running-with-mig

Disable MIG

```
$ sudo nvidia-smi -mig 0
Disabled MIG Mode for GPU 00000000:06:00.0

Warning: persistence mode is disabled on device 00000000:06:00.0. See the
Known Issues section of the nvidia-smi(1) man page for more information. Run
with [--help | -h] switch to get more information on how to enable
persistence mode.
All done.

$ nvidia-smi
Thu Mar 10 21:25:12 2022
+-----------------------------------------------------------------------------
--+
| NVIDIA-SMI 510.47.03    Driver Version: 510.47.03    CUDA Version: 11.6
|
|-------------------------------+----------------------+--------------------
--+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr.
ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute
M. |
|                               |                      |              MIG
M. |
|===============================+======================+====================
==|
|   0  NVIDIA A100 80G...  Off  | 00000000:06:00.0 Off |
0 |
| N/A   33C    P0    58W / 300W |      0MiB / 81920MiB |     20%
Default |
|                               |                      |
Disabled |
+-------------------------------+----------------------+--------------------
--+

+-----------------------------------------------------------------------------
--+
| Processes:
|
```

```
|  GPU   GI   CI        PID   Type   Process name                  GPU
Memory |
|        ID   ID                                                   Usage
|
|=============================================================================
==|
|  No running processes found
|
+-----------------------------------------------------------------------------
--+
```

Enable MIG

```
$ sudo nvidia-smi -mig 1
Enabled MIG Mode for GPU 00000000:06:00.0

Warning: persistence mode is disabled on device 00000000:06:00.0. See the
Known Issues section of the nvidia-smi(1) man page for more information. Run
with [--help | -h] switch to get more information on how to enable
persistence mode.
All done.

$ nvidia-smi
Thu Mar 10 21:32:23 2022
+-----------------------------------------------------------------------------
--+
| NVIDIA-SMI 510.47.03    Driver Version: 510.47.03    CUDA Version: 11.6
|
|-------------------------------+----------------------+----------------------
--+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr.
ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute
M. |
|                               |                      |               MIG
M. |
|===============================+======================+======================
==|
|   0  NVIDIA A100 80G...  Off  | 00000000:06:00.0 Off |
On |
| N/A   32C    P0    57W / 300W |      0MiB / 81920MiB |     N/A
Default |
|                               |                      |
Enabled |
+-------------------------------+----------------------+----------------------
--+

+-----------------------------------------------------------------------------
--+
| MIG devices:
```

```
|
+------------------+----------------------+-----------+----------------------
--+
| GPU  GI  CI  MIG |         Memory-Usage |        Vol|         Shared
|
|      ID  ID  Dev |           BAR1-Usage | SM     Unc| CE  ENC  DEC  OFA
JPG|
|                  |                      |        ECC|
|
|==================+======================+===========+======================
==|
|  No MIG devices found
|
+-----------------------------------------------------------------------------
--+

+-----------------------------------------------------------------------------
--+
| Processes:
|
|  GPU   GI   CI        PID   Type   Process name                  GPU
Memory |
|        ID   ID                                                   Usage
|
|=============================================================================
==|
|  No running processes found
|
+-----------------------------------------------------------------------------
--+
```

List GPU Instance Profiles (-lgip)

```
$ sudo nvidia-smi mig -lgip

+-----------------------------------------------------------------------------
--+
| GPU instance profiles:
|
| GPU   Name             ID    Instances   Memory     P2P    SM    DEC   ENC
|
|                              Free/Total   GiB              CE    JPEG  OFA
|
|=============================================================================
==|
|   0  MIG 1g.10gb       19     7/7        9.50       No     14     0     0
|
|                                                           1     0     0
|
```

```
+-----------------------------------------------------------------------------
--+
|   0  MIG 1g.10gb+me    20     1/1        9.50       No     14     1     0
|
|                                                          1     1     1
|
+-----------------------------------------------------------------------------
--+
|   0  MIG 2g.20gb       14     3/3        19.50      No     28     1     0
|
|                                                          2     0     0
|
+-----------------------------------------------------------------------------
--+
|   0  MIG 3g.40gb        9     2/2        39.25      No     42     2     0
|
|                                                          3     0     0
|
+-----------------------------------------------------------------------------
--+
|   0  MIG 4g.40gb        5     1/1        39.25      No     56     2     0
|
|                                                          4     0     0
|
+-----------------------------------------------------------------------------
--+
|   0  MIG 7g.80gb        0     1/1        78.75      No     98     5     0
|
|                                                          7     1     1
|
+-----------------------------------------------------------------------------
--+
```

List GPU Instance Profiles possible placements (-lgipp)

```
$ sudo nvidia-smi mig -lgipp
GPU  0 Profile ID 19 Placements: {0,1,2,3,4,5,6}:1
GPU  0 Profile ID 20 Placements: {0,1,2,3,4,5,6}:1
GPU  0 Profile ID 14 Placements: {0,2,4}:2
GPU  0 Profile ID  9 Placements: {0,4}:4
GPU  0 Profile ID  5 Placement : {0}:4
GPU  0 Profile ID  0 Placement : {0}:8
```

Creating GPU Instances

Create only GPU instances

```
$ sudo nvidia-smi mig -cgi 14,14,19,19,20
```

```
Successfully created GPU instance ID  5 on GPU  0 using profile MIG 2g.20gb
(ID 14)
Successfully created GPU instance ID  3 on GPU  0 using profile MIG 2g.20gb
(ID 14)
Successfully created GPU instance ID 13 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created GPU instance ID  9 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created GPU instance ID 10 on GPU  0 using profile MIG
1g.10gb+me (ID 20)
```

Create GPU instances and apply default compute instance profile

```
$ sudo nvidia-smi mig -cgi 14,14,19,19,20 -C
Successfully created GPU instance ID  5 on GPU  0 using profile MIG 2g.20gb
(ID 14)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  5
using profile MIG 2g.20gb (ID  1)
Successfully created GPU instance ID  3 on GPU  0 using profile MIG 2g.20gb
(ID 14)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  3
using profile MIG 2g.20gb (ID  1)
Successfully created GPU instance ID 13 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID 13
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID  9 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  9
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID 10 on GPU  0 using profile MIG
1g.10gb+me (ID 20)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID 10
using profile MIG 1g.10gb (ID  0)
```

One big GPU

```
$ sudo nvidia-smi mig -cgi 0 -C
Successfully created GPU instance ID  0 on GPU  0 using profile MIG 7g.80gb
(ID  0)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  0
using profile MIG 7g.80gb (ID  4)

$ nvidia-smi -L
GPU 0: NVIDIA A100 80GB PCIe (UUID: GPU-
a166d193-64ba-8969-91b0-7729983c680b)
  MIG 7g.80gb     Device  0: (UUID: MIG-c4f17314-2b46-5569-ac07-
bb0781b44ba7)
$ nvidia-smi
Thu Mar 10 21:34:25 2022
```

```
+-----------------------------------------------------------------------------
--+
| NVIDIA-SMI 510.47.03    Driver Version: 510.47.03    CUDA Version: 11.6
|
|-------------------------------+----------------------+----------------------
--+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr.
ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute
M. |
|                               |                      |               MIG
M. |
|===============================+======================+======================
==|
|   0  NVIDIA A100 80G...  Off  | 00000000:06:00.0 Off |
On |
| N/A   32C    P0    67W / 300W |      0MiB / 81920MiB |     N/A
Default |
|                               |                      |
Enabled |
+-------------------------------+----------------------+----------------------
--+

+-----------------------------------------------------------------------------
--+
| MIG devices:
|
+------------------+----------------------+-----------+----------------------
--+
| GPU  GI  CI  MIG |         Memory-Usage |        Vol|         Shared
|
|      ID  ID  Dev |           BAR1-Usage | SM     Unc| CE  ENC  DEC  OFA
JPG|
|                  |                      |        ECC|
|
|==================+======================+===========+======================
==|
|  0    0   0   0  |      0MiB / 81069MiB | 98      0 |  7   0    5    1
1 |
|                  |      1MiB / 13107... |           |
|
+------------------+----------------------+-----------+----------------------
--+

+-----------------------------------------------------------------------------
--+
| Processes:
|
|  GPU   GI   CI        PID   Type   Process name                  GPU
Memory |
|        ID   ID                                                   Usage
```

```
|
|=============================================================================
==|
|  No running processes found
|
+-----------------------------------------------------------------------------
--+
```

Many (7) smallest GPUs

```
~$ sudo nvidia-smi mig -cgi 19,19,19,19,19,19,19 -C
Successfully created GPU instance ID 13 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID 13
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID 11 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID 11
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID 12 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID 12
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID  7 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  7
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID  8 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  8
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID  9 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  9
using profile MIG 1g.10gb (ID  0)
Successfully created GPU instance ID 10 on GPU  0 using profile MIG 1g.10gb
(ID 19)
Successfully created compute instance ID  0 on GPU  0 GPU instance ID 10
using profile MIG 1g.10gb (ID  0)

$ nvidia-smi -L
GPU 0: NVIDIA A100 80GB PCIe (UUID: GPU-
a166d193-64ba-8969-91b0-7729983c680b)
  MIG 1g.10gb     Device  0: (UUID: MIG-6655c8fd-3878-5030-
a9b0-65b40839f65e)
  MIG 1g.10gb     Device  1: (UUID: MIG-52c663bb-eef5-580c-
b98f-2c2d93f4d668)
  MIG 1g.10gb     Device  2: (UUID: MIG-
efafa9b6-09cc-5b9f-9a43-73dc927794f0)
  MIG 1g.10gb     Device  3: (UUID:
```

```
MIG-3f2ece72-3e45-560d-8698-79d951b1312b)
  MIG 1g.10gb     Device  4: (UUID: MIG-09f3602a-747a-5357-9d60-
fd883da6fae2)
  MIG 1g.10gb     Device  5: (UUID:
MIG-38d561b9-0ddf-5033-9fc0-1016588a3854)
  MIG 1g.10gb     Device  6: (UUID: MIG-33709a1d-5135-5f0d-a6ab-
b936857ed5e4)

$ nvidia-smi
Thu Mar 10 21:36:19 2022
+-----------------------------------------------------------------------------
--+
| NVIDIA-SMI 510.47.03    Driver Version: 510.47.03    CUDA Version: 11.6
|
|-------------------------------+----------------------+----------------------
--+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr.
ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute
M. |
|                               |                      |               MIG
M. |
|===============================+======================+======================
==|
|   0  NVIDIA A100 80G...  Off  | 00000000:06:00.0 Off |
On |
| N/A   32C    P0    68W / 300W |     45MiB / 81920MiB |     N/A
Default |
|                               |                      |
Enabled |
+-------------------------------+----------------------+----------------------
--+

+-----------------------------------------------------------------------------
--+
| MIG devices:
|
+------------------+----------------------+-----------+----------------------
--+
| GPU  GI  CI  MIG |         Memory-Usage |        Vol|         Shared
|
|      ID  ID  Dev |           BAR1-Usage | SM     Unc| CE  ENC  DEC  OFA
JPG|
|                  |                      |        ECC|
|
|==================+======================+===========+======================
==|
|  0    7   0   0  |      6MiB /  9728MiB | 14      0 |  1   0    0    0
0 |
|                  |      0MiB / 16383MiB |           |
|
```

```
+------------------+----------------------+-----------+----------------------
--+
|  0    8   0   1  |      6MiB /  9728MiB | 14      0 |  1   0    0    0
0 |
|                  |      0MiB / 16383MiB |           |
|
+------------------+----------------------+-----------+----------------------
--+
|  0    9   0   2  |      6MiB /  9728MiB | 14      0 |  1   0    0    0
0 |
|                  |      0MiB / 16383MiB |           |
|
+------------------+----------------------+-----------+----------------------
--+
|  0   10   0   3  |      6MiB /  9728MiB | 14      0 |  1   0    0    0
0 |
|                  |      0MiB / 16383MiB |           |
|
+------------------+----------------------+-----------+----------------------
--+
|  0   11   0   4  |      6MiB /  9728MiB | 14      0 |  1   0    0    0
0 |
|                  |      0MiB / 16383MiB |           |
|
+------------------+----------------------+-----------+----------------------
--+
|  0   12   0   5  |      6MiB /  9728MiB | 14      0 |  1   0    0    0
0 |
|                  |      0MiB / 16383MiB |           |
|
+------------------+----------------------+-----------+----------------------
--+
|  0   13   0   6  |      6MiB /  9728MiB | 14      0 |  1   0    0    0
0 |
|                  |      0MiB / 16383MiB |           |
|
+------------------+----------------------+-----------+----------------------
--+

+-----------------------------------------------------------------------------
--+
| Processes:
|
|  GPU   GI   CI        PID   Type   Process name                  GPU
Memory |
|        ID   ID                                                   Usage
|
|=============================================================================
==|
|  No running processes found
```

```
|
+-----------------------------------------------------------------------------
--+
```

One big GPU with many compute instances

```
$sudo nvidia-smi mig -cci 0
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)

$ sudo nvidia-smi mig -cci 0,0,0,0,0,0,0
Successfully created compute instance ID  0 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)
Successfully created compute instance ID  1 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)
Successfully created compute instance ID  2 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)
Successfully created compute instance ID  3 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)
Successfully created compute instance ID  4 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)
Successfully created compute instance ID  5 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)
Successfully created compute instance ID  6 on GPU  0 GPU instance ID  0
using profile MIG 1c.7g.80gb (ID  0)

$ nvidia-smi -L
GPU 0: NVIDIA A100 80GB PCIe (UUID: GPU-
a166d193-64ba-8969-91b0-7729983c680b)
  MIG 1c.7g.80gb  Device  0: (UUID: MIG-c4f17314-2b46-5569-ac07-
bb0781b44ba7)
  MIG 1c.7g.80gb  Device  1: (UUID: MIG-2b5e73e1-66c1-5801-8084-
ef0d0ec08b6c)
  MIG 1c.7g.80gb  Device  2: (UUID: MIG-000028a5-5c26-532f-
b903-38ffc3ed1422)
  MIG 1c.7g.80gb  Device  3: (UUID: MIG-3046d0e8-aaf9-546e-
a9a7-5673d31eeea8)
  MIG 1c.7g.80gb  Device  4: (UUID: MIG-80c0ecbd-29bb-5944-8576-
a0c04b437a27)
  MIG 1c.7g.80gb  Device  5: (UUID: MIG-e5a655a5-9430-51ef-
abca-3f8c1527023c)
  MIG 1c.7g.80gb  Device  6: (UUID: MIG-
b6de6961-47b2-56c1-88f9-26775c22e1a7)
$ nvidia-smi
Thu Mar 10 21:21:23 2022
+-----------------------------------------------------------------------------
--+
| NVIDIA-SMI 510.47.03    Driver Version: 510.47.03    CUDA Version: 11.6
|
|-------------------------------+----------------------+----------------------
```

```
--+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr.
ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute
M. |
|                               |                      |              MIG
M. |
|===============================+======================+==================
==|
|   0  NVIDIA A100 80G...  Off  | 00000000:06:00.0 Off |
On |
| N/A   32C    P0    61W / 300W |      0MiB / 81920MiB |     N/A
Default |
|                               |                      |
Enabled |
+-------------------------------+----------------------+------------------
--+

+-----------------------------------------------------------------------------
--+
| MIG devices:
|
+------------------+----------------------+-----------+---------------------
--+
| GPU  GI  CI  MIG |         Memory-Usage |        Vol|         Shared
|
|      ID  ID  Dev |           BAR1-Usage | SM     Unc| CE  ENC  DEC  OFA
JPG|
|                  |                      |        ECC|
|
|==================+======================+===========+=====================
==|
|  0    0   0   0  |      0MiB / 81069MiB | 14      0 |  7   0    5    1
1 |
|                  |      1MiB / 13107... |           |
|
+------------------+                      +-----------+---------------------
--+
|  0    0   1   1  |                      | 14      0 |  7   0    5    1
1 |
|                  |                      |           |
|
+------------------+                      +-----------+---------------------
--+
|  0    0   2   2  |                      | 14      0 |  7   0    5    1
1 |
|                  |                      |           |
|
+------------------+                      +-----------+---------------------
--+
```

```
|  0    0   3   3  |                      | 14      0 |  7   0    5    1
1 |
|                  |                      |           |
|
+------------------+                      +-----------+--------------------
--+
|  0    0   4   4  |                      | 14      0 |  7   0    5    1
1 |
|                  |                      |           |
|
+------------------+                      +-----------+--------------------
--+
|  0    0   5   5  |                      | 14      0 |  7   0    5    1
1 |
|                  |                      |           |
|
+------------------+                      +-----------+--------------------
--+
|  0    0   6   6  |                      | 14      0 |  7   0    5    1
1 |
|                  |                      |           |
|
+------------------+----------------------+-----------+--------------------
--+

+-----------------------------------------------------------------------------
--+
| Processes:
|
|  GPU   GI   CI        PID   Type   Process name                  GPU
Memory |
|        ID   ID                                                   Usage
|
|=============================================================================
==|
|  No running processes found
|
+-----------------------------------------------------------------------------
--+
```

Delete all

```
$ sudo nvidia-smi mig -dci && sudo nvidia-smi mig -dgi
Successfully destroyed compute instance ID  0 from GPU  0 GPU instance ID  9
Successfully destroyed compute instance ID  0 from GPU  0 GPU instance ID 13
Successfully destroyed compute instance ID  0 from GPU  0 GPU instance ID  3
Successfully destroyed compute instance ID  0 from GPU  0 GPU instance ID  5
Successfully destroyed compute instance ID  0 from GPU  0 GPU instance ID 10
```

```
Successfully destroyed GPU instance ID  9 from GPU  0
Successfully destroyed GPU instance ID 13 from GPU  0
Successfully destroyed GPU instance ID  3 from GPU  0
Successfully destroyed GPU instance ID  5 from GPU  0
Successfully destroyed GPU instance ID 10 from GPU  0
```

From:
https://servizi.di.unimi.it/ - **Servizi Tecnologici Dipartimento di Informatica**

Permanent link:
https://servizi.di.unimi.it/hpc:start

Last update: **2022/03/10 21:31**