Running GPU Jobs

GPUs are available for use in the HPC cluster. To utilize GPU processing in your job, they must be requested. GPUs should only be used with jobs that can make use of them. You must also select the GPU partition to make use of them.

Checking GPU Availability and Versions

To check on the availability of GPUs and their versions, use the following command:

```
nvidia-smi
```

This will output information on current usage, versions of nvidia drivers, GPU memory usage, and quite a bit more:

```
Thu Jul 1 08:19:01 2021


GPU Name Persistence-M| Bus-ID Disp.A Memory-Usage GPU-Util Compute M. MIG M.
Fan Temp Perf Pwr:Usage/Cap
-----------------------------------------------------------------------------------------------
  0 NVIDIA Quadro R... Off 0:00.0 Off 0MiB / 45556MiB 0% Default N/A
  N/A 27C P0 56W / 250W

  1 NVIDIA Quadro R... Off 0:00.0 Off 0MiB / 45556MiB 0% Default N/A
  N/A 26C P0 57W / 250W

  2 NVIDIA Quadro R... Off 0:00.0 Off 0MiB / 45556MiB 0% Default N/A
  N/A 26C P0 55W / 250W

Processes:

  GPU CI CI PID Type Process name GPU Memory Usage
  ID ID

No running processes found
```

Modules

There are several modules available on the HPC cluster, which include CUDA and cuDNN. You can use the module avail command to get further information about the version of these modules installed:

```
module avail cuda
module avail cudnn
```
Your next step will be loading the module you wish to use for GPU jobs. As you can see from running module avail, cuda version 11.0.2 and 11.1.0 are available on the cluster. If we wanted to load 11.0.2, we would use the command:

```
module load CUDA/11.0.2
```