Using GPUs on the patas cluster

Shane Steinert-Threlkeld

January 24 2020

The GPU Nodes on patas

Patas currently has two GPU nodes (though a third is coming soon!). Configured as follows:

node	num GPUs	GPU type	memory per GPU
gn1	2	Tesla P40	8GB
gn2	8 (2x4)	Tesla M10	8GB

Requesting GPUs

Requesting a GPU is simple: simply add `request_GPUs = 1` to your condor job file. See this page for more information. *Note:* Please only request one GPU. These are a limited resource, with a fairly heavy demand. If you have a job that genuinely requires multiple GPUs (e.g. for memory), try to wait until a time when usage is very low.

Setting up your local environment

You should be using a virtual environment (either `pyenv` or `conda`) to manage the dependencies of your project. I recommend `conda`. As a minimal example of setting up an environment with `pytorch` and `AllenNLP` installed, you can execute the following steps:

1. `wget https://repo.anaconda.com/archive/Anaconda3-2019.10-Linux-x86_64.sh`

   NB: replace the URL with whatever the newest version is at https://www.anaconda.com/distribution/#linux
2. `bash Anaconda3-2019.10-Linux-x86_64.sh`

   Say "yes" when prompted
3. `conda create -n my-project`
4. `conda activate my-project`
5. `conda install cudatoolkit=10.0 pytorch -c pytorch -c anaconda`

   NB: the `=10.0` is very important!
6. `pip install allennlp`

Now, in the future, you can just add `conda activate my-project` to the start of your job scripts.