

Using the Cluster

Quick guide to using the cluster by Pulkit M. Requires account on Dvorak.

Assumes knowledge of navigating system directories and the following terminal commands: `cd`, `ls`, `pwd`, `mkdir`, `mv`, `cat`.

GitHub Gnina Repository: <https://github.com/gnina>

1) Logging onto cluster:

`ssh username@gpu.csb.pitt.edu`

To exit cluster/server, type `exit`

* If not in lab, SSH into dvorak (`username@dvorak.csb.pitt.edu`) first, and from there to the gpu cluster. VPN tool required (e.g. Pulse Secure).

2) Read through the README file located on the cluster.

3) Place all necessary files (model, data, python scripts, pbs script) into a single directory on local machine.

- ❖ Model: <https://github.com/gnina/models>
 - File normally ends in `.model` or `.prototxt`
- ❖ Scripts: <https://github.com/gnina/scripts>
 - Ensure that all files ending in `.py` are in your directory.
- ❖ Data: Refer to Dr. Koes
<https://github.com/gnina/models/tree/master/data>

4) In the model file, in all layers of type “MolGridData” change the root folder to:

`"/net/pulsar/home/koes/dkoes/PDBbind/refined-set/"`

```
name: "residual-net-v0.1"
layer {
  name: "data"
  type: "MolGridData"
  top: "data"
  top: "label"
  top: "affinity"
  include {
    phase: TEST
  }
  molgrid_data_param {
    source: "TESTFILE"
    batch_size: 1
    dimension: 23.5
    resolution: 0.5
    shuffle: false
    balanced: false
    has_affinity: true
    root_folder: "/net/pulsar/home/koes/dkoes/PDBbind/refined-set/"
  }
}
```

5) Include the following exports in your PBS script (before the python command):

Refer to provided pbs script for a complete template.

```
export PATH=/usr/bin:$PATH
export
LD_LIBRARY_PATH=/net/pulsar/home/koes/dkoes/local/lib:/usr/lib64:/usr/lib/x86_64-linux-gnu:/usr/local/cuda-8.0/lib64
export PYTHONPATH=/net/pulsar/home/koes/dkoes/local/python:$PYTHONPATH
```

6) Copy working directory onto server (scp command):

```
scp -r ~/Desktop/test_folder username@gpu.csb.pitt.edu:~
scp -r test_folder username@perigee/apogee.csb.pitt.edu:~/Desktop
```

7) *Test on cluster nodes, not head node:**

Launch job with **qsub script.pbs** from directory with required files. Use **qstat -au username** to check job status.

Do not run python directly in terminal after ssh.

Troubleshooting

Use **cat** to read output file (located in folder **qsub** was run in) and **pip** to manually install missing python packages (e.g. numpy):

```
pip install -I --user package
```

Quick Tips

Viewing files in terminal:

cat /path/to/file

Editing files in terminal:

vi /path/to/file

Vi Basics

Default is command mode.

- ❖ **x** to delete character under cursor
- ❖ **v** to start selection (for copy/cut operation)
 - Move cursor to select, then **y** to copy or **x** to cut
- ❖ Position cursor, then **p** to paste
- ❖ Save & exit **:wq** (MUST BE IN COMMAND MODE)
- ❖ Exit **:q!**

To enter insert mode press **i** (to type normally), and **Esc** to go back to command mode.