

Machine Learning on HPC: Introduction to ML Support at TACC

Sikan Li

Research Engineer

Scalable Computational Intelligence Group

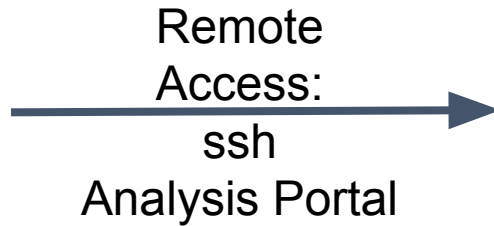
Texas Advanced Computing Center

University of Texas at Austin

Accessing TACC Machines Remotely



Personal Laptop



Secure Shell (ssh): Accessing Shell

```
localhost$ ssh username@frontera.tacc.utexas.edu
```

- The standard way to connect TACC machines
- SSH is available within Linux and from the terminal app in the Mac OS. If you are using Windows, you will need an SSH client that supports the SSH-2 protocol: e.g. [Bitvise](#), [OpenSSH](#), [PuTTY](#), or [SecureCRT](#)
- Read [User Guide](#) for more information on how to access TACC machine with ssh

TACC Analysis Portal

- TACC Analysis Portal (TAP) provides access to interactive sessions on TACC machines
- Allows user to utilize web-based interactions on TACC computers
 - e.g. Jupyter notebook, Rstudio

TACC Analysis Portal

- TACC Analysis Portal (TAP) provides access to interactive sessions on TACC machines
- Allows user to utilize web-based interactions on TACC computers
 - e.g. Jupyter notebook, Rstudio
- tap.tacc.utexas.edu

TACC | Analysis Portal User Guide

Welcome to the TACC Analysis Portal

simple access to TACC's analysis resources

Log in to TAP

Jupyter notebook: Accessing shell

- TACC Analysis Portal:

TACC | Analysis Portal [User Guide](#) jrduncan [Log Out](#)

Submit New Job

System:

Application:

Project:

Queue:

Nodes: Tasks:

Options

Job Name:

Time Limit:

Reservation:

VNC Desktop Resolution:

[Submit](#) [Utilities](#)

System Status

System	Status	Utilization	Job Count
Frontera	✓ Open	99%	Running: 312 Queued: 1251
Lonestar6	✓ Open	69%	Running: 135 Queued: 98
Longhorn	✓ Open	74%	Running: 30 Queued: 40
Maverick2	✓ Open	16%	Running: 4 Queued: 7
Stampede2	✓ Open	96%	Running: 830 Queued: 736

Past Jobs

JNB-Frontera	03/18/2022	Details	Resubmit
JNB-Frontera	03/18/2022	Details	Resubmit
JNB-Frontera	03/18/2022	Details	Resubmit
JNB-Frontera	03/18/2022	Details	Resubmit
JNB-Frontera	02/21/2022	Details	Resubmit

Jupyter notebook: Accessing shell

TAP Job Status

Job: Jupyter notebook on Frontera (4175197, 2022-03-21T17:28-05:00)

Status: RUNNING

Start: March 21, 2022, 5:28 p.m.

End: March 21, 2022, 5:33 p.m.

Refresh: in 873 seconds

Message:

```
TAP: Your session is running at https://frontera.tacc.utexas.edu:60752  
/?token=9cbad0f26752e7dd14fcf090d6a30b6ec5c15c63ed7d9e2b626f214712fb8b4d
```

Connect

End Job

Show Output

Back to Jobs

Jupyter notebook: Accessing shell



Quit Logout

Files Running Clusters

Select items to perform actions on them.

The screenshot shows the Jupyter file browser interface. At the top, there are tabs for 'Files', 'Running', and 'Clusters'. Below the tabs, there are buttons for 'Upload', 'New', and a refresh icon. A dropdown menu is open under the 'New' button, showing options: 'Notebook: Python 3', 'Other: Text File', 'Folder', and 'Terminal'. The 'Terminal' option is circled in red. Below the menu, a list of files and folders is visible, including 'design-safe-tutorial', 'python_for_ML_training', 'tutorial-0716', 'UTSA_DL_Tutorial', and 'Untitled.ipynb'.

Resources on TACC Machines



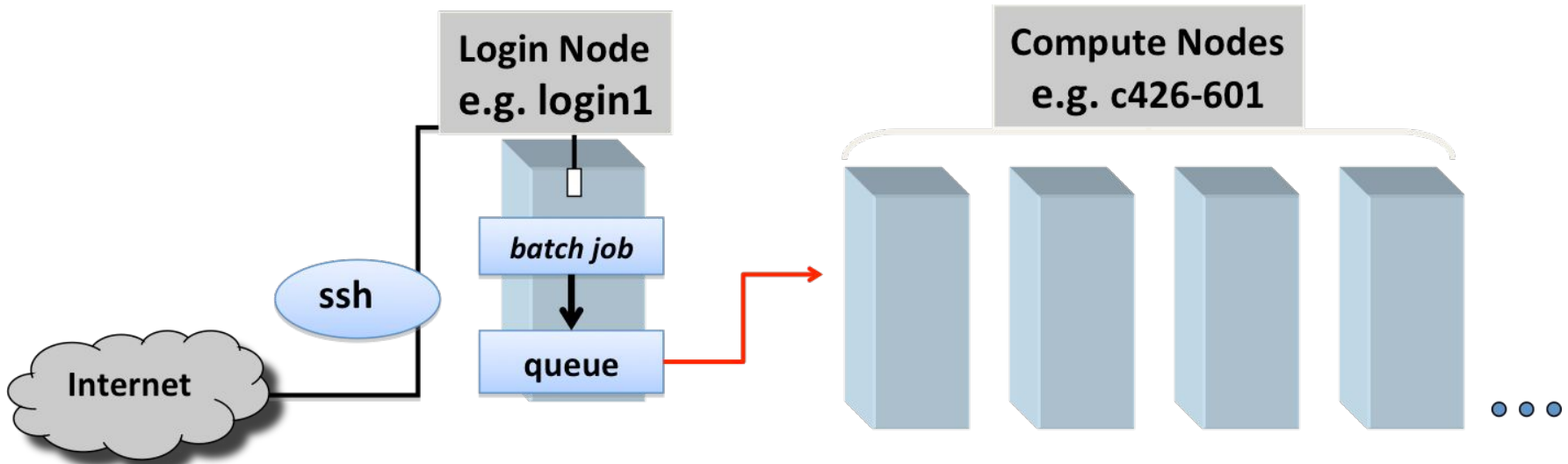
Personal Laptop

Remote
Access:
ssh
Analysis Portal



- Login versus Compute nodes
- File systems

Login vs Compute Nodes



Login Nodes

- Shared by multiple users
- Not meant for CPU or memory intensive jobs
- Meant for managing file and compiling code

Compute Nodes

- Where research computation occurs
- Used for sbatch or interactive session to access compute nodes

Accessing Compute Nodes

Ways to access compute nodes

- submit a batch job via the sbatch command
 - job waits in queue until resources are available
 - [example script](#)
 - sbatch myjobscript
 - queue -u bjones #show all jobs owned by bjones
 - [More job management commands](#)

```
JOBID  PARTITION  NAME      USER ST    TIME  NODES NODELIST(REASON)
25781  debug  idv72397  bjones CG    9:36   2  c190-131,c191-092
25918  debug  ppm_4828  bjones PD    0:00  4828 (Resources)
25915  debug  MV2-test  siliu  PD    0:00  4200 (Priority)
25940  normal  SWMF      xtwang PD    0:00   18 (Nodes required for job are DOWN, DRAINED or reserved for jobs in higher priority partitions)
25589  normal  aatest   slindsey PD    0:00   8 (Dependency)
25949  debug  psdns_la sniffjck PD    0:00  256 (Priority)
25942  normal  WRF2     sniffjck PD    0:00  128 (Nodes required for job are DOWN, DRAINED or reserved for jobs in higher priority partitions)
25618  normal  SP256U   connor PD    0:00   1 (Dependency)
25944  normal  MoTi_hi  wchung R    35:13  1  c112-203
25945  normal  WTi_hi_e wchung R    27:11  1  c113-131
25606  normal  trainA   jackhu R   23:28:28 1  c119-152
```

Accessing Compute Nodes

- [activate an interactive session via idev](#)
 - idev
 - Allows you to get a compute node for testing purposes
 - -m option lets you specify time in minutes
 - `idev -m 30`

File systems

- Three (3) main LUSTRE file systems
 - \$HOME
 - Quota (25G)
 - Backed up
 - Used for: small scripts, environment settings, other routine file management task
 - \$WORK/\$WORK2
 - Quota (1TB)
 - Not Backed up
 - Used For: Intermediate staging/preparation of files
 - Move files in/out of \$SCRATCH
 - Part of one large Global file system, available on all TACC systems
 - \$SCRATCH
 - No Quota
 - Not Backed up, subject to purge
 - Used for: High I/O bandwidth jobs
- Refer to Frontera User Guide
 - <https://frontera-portal.tacc.utexas.edu/user-guide/files/>

module

- On TACC Systems the module system helps setup or teardown useful packages that are supported by TACC, on the fly
- To find more information on a module
 - `module spider package-name`
- To load a module
 - `module load package-name`
 - `module load package-name/version.number`
- [Learn more here](#)

Virtual Environment

- Python virtual environment / Conda virtual environment
 - Find more instructions on python virtual env [here](#)

```
python3 -m venv ~/testp3env
source ~/testp3env/bin/activate
pip install --upgrade pip
# sample pip install
pip install wget
# to exit the virtual env:
deactivate
```

Thanks

- [TACC instruction page](#)

TACC Computing

IHARP → Resources → TACC Computing

Account Setup

Need Help?

Learning Portal

- Questions

- Any Issues on TACC Systems

- Open a ticket a <http://consult.tacc.utexas.edu>