

Introduction to TACC Computing Cluster

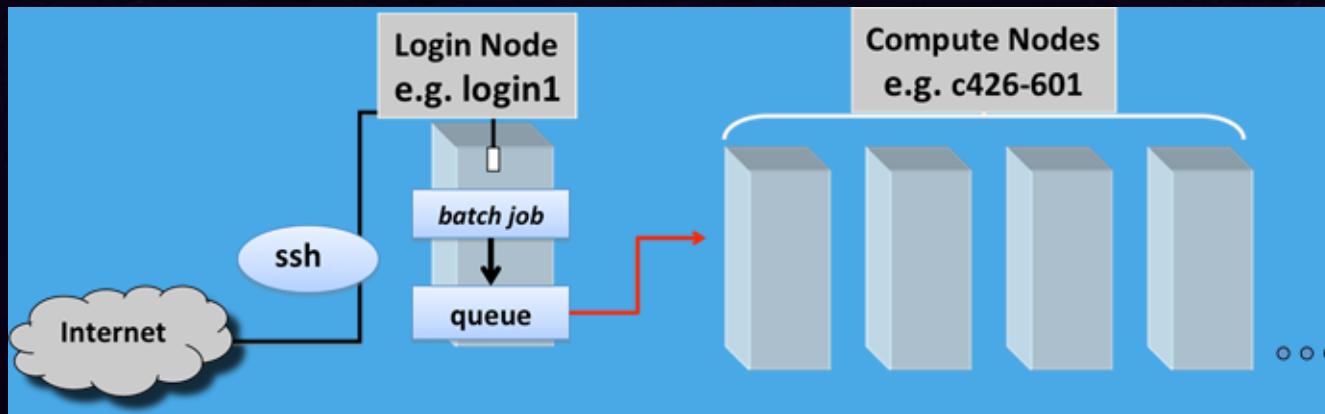
Scalable Computational Intelligence
Texas Advanced Computing Center (TACC)

Overview

Practical introduction to using TACC resources

- Nodes and File system at TACC
- How to Remotely Access TACC
- Navigating the File System on TACC machines
- Hands On: Accessing TACC via TAP

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

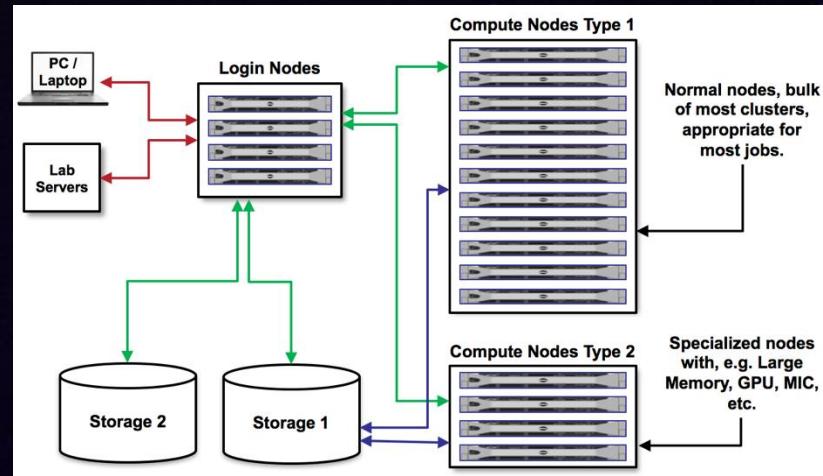
- Where research computation occurs
- Training machine learning models, running simulations, etc.

Types of Compute Nodes: Partitions

- Compute nodes are divided into different pools called queues
 - For example here are a few queues on VISTA
 - Grace/Grace (gg)
 - Grace/Hoppersmall (gh)
 - Development Grace Hopper (gh-dev)
 - Refer to the user guide for the queue and accounting details

File Systems

- Three (3) main file systems
 - \$HOME
 - Quota (25G)
 - Backed up
 - Used for: small scripts, environment settings, and other routine file management task
 - \$WORK
 - Quota (1TB)
 - Not Backed up
 - Used For: Intermediate staging/preparation of files, storing original datasets
 - \$SCRATCH
 - No Quota
 - Not Backed up, subject to purge
 - Used for: High I/O bandwidth jobs
- Refer to VISTA User Guide / TACC tutorials
 - <https://docs.tacc.utexas.edu/hpc/vista/#system-filesystems>
 - <https://docs.tacc.utexas.edu/tutorials/managingio/>



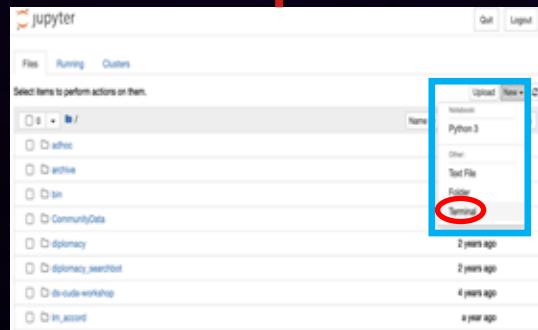
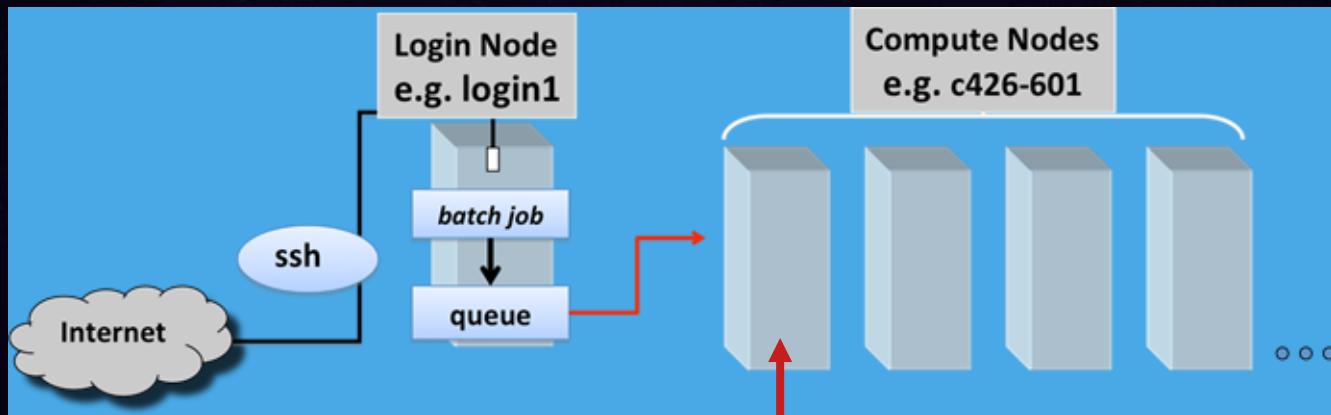
- **/tmp**
 - Each compute node comes with local storage (/tmp)
 - Fast I/O
 - **Temporary:** storage only last the duration of the job and should be copied for permanent storage

How to Remotely Access TACC

We will remotely access Frontera in two ways in this event:

1. SSH: Access login nodes via command line

2. TACC Analysis Portal: UI for accessing jupyter notebook session on a compute nodes



Accessing Compute Nodes

Let's assume we have logged into TACC via ssh and we are on a log in node. How do we connect to a compute node?

- submit a batch job via the sbatch command
 - job waits in queue until resources are available
- 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`



Bash – basic file commands

- TACC machines use the Linux operating system
- Default shell and most popular is bash (Bourne Again Shell)

Command	Function
pwd	Shows absolute path of your present working directory
ls <directory name>	Lists a directory (if mentioned) or present working directory
mkdir <name>	Creates a new directory with mentioned name
cd <name>	Changes working directory to name mentioned (or go to home directory if nothing is specified)
cp <File1> <File2> cp <File1> <Directory>	Makes a copy of File1 and names it File2 or makes a copy of the file with the same name in Directory1
cp -r <directory1> <directory2>	Makes a copy of directory1 and all its contents into directory2
mv <File1> <Directory> mv <File1> <File2>	moves File1 to Directory1 or rename File1 to File2
rm <File1> rm -r <Directory1>	Removes a file or Remove a directory (add –f flag to force remove...be careful)

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`

Check out the User guide to learn more:

https://lmod.readthedocs.io/en/latest/010_user.html

Hands on: Accessing TACC via the TACC Analysis Portal and Environment Setup



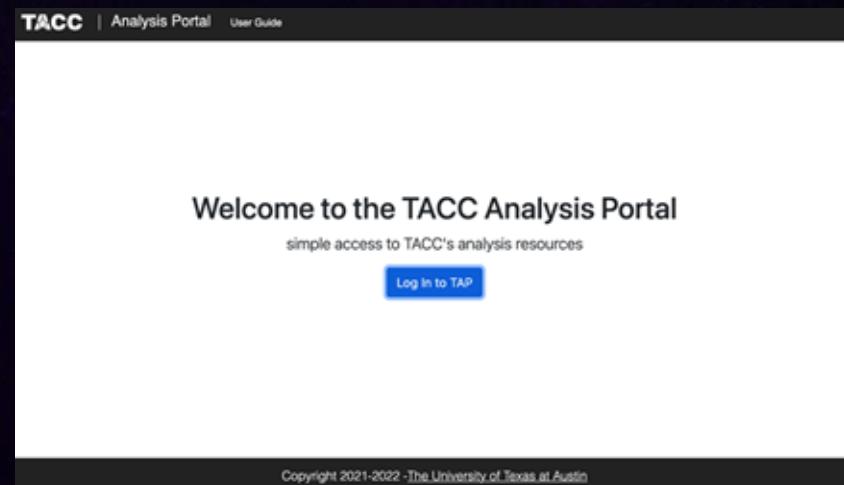
Development Environment

- ▶ Jupyter Notebook
 - ▶ <http://jupyter.org/>
 - ▶ Web application for creating and sharing computational documents.
 - ▶ It offers a simple, streamlined, document-centric experience.
 - ▶ Provides an interactive interface in the browser
 - ▶ Can view output and plots inline
 - ▶ Makes Iterative/Interactive Development easy
 - ▶ Can be turned into a pure script for non-interactive work



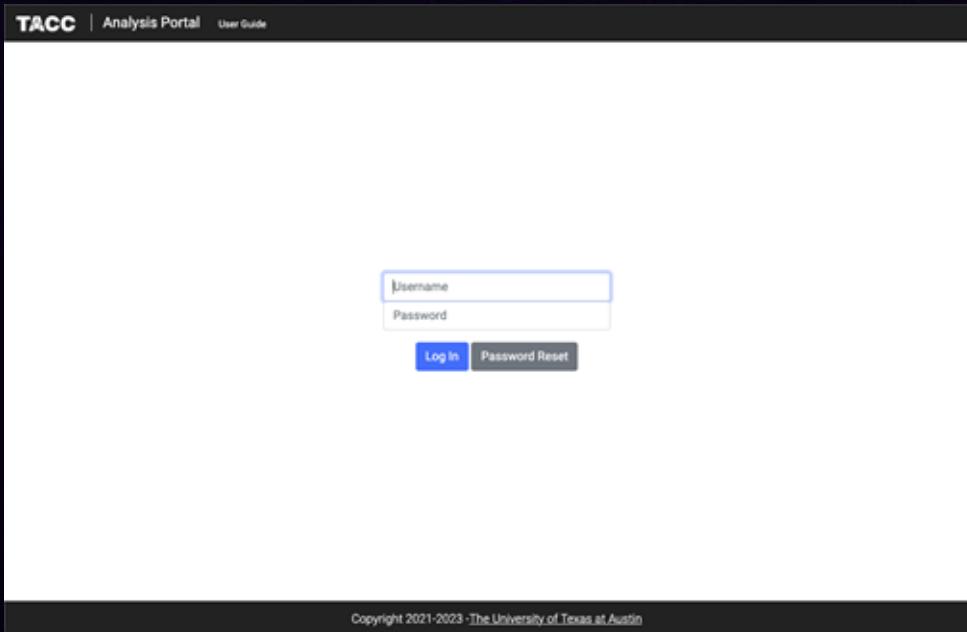
TACC Analysis Portal

- 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, remote linux desktop
- <https://tap.tacc.utexas.edu>



Login with your TACC account

- TACC Analysis Portal:
login with your TACC account credentials



Jupyter notebook: Accessing shell

TACC | Analysis Portal [User Guide](#)

jrduan 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	<input checked="" type="checkbox"/> Open	99%	Running: 312 Queued: 1251
Lonestar6	<input checked="" type="checkbox"/> Open	69%	Running: 135 Queued: 98
Longhorn	<input checked="" type="checkbox"/> Open	74%	Running: 30 Queued: 40
Maverick2	<input checked="" type="checkbox"/> Open	16%	Running: 4 Queued: 7
Stampede2	<input checked="" type="checkbox"/> 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	02/21/2022	Details	Resubmit

Starting Jupyter notebook: Accessing shell

- TACC Analysis Portal:
Select the following options

Submit New Job

System	Vista
Application	Jupyter notebook
Project	Frontera-Training
Queue	gh
Nodes	1
Tasks	1
Options	
Job Name	20 characters max
Time Limit	H:M:S (default 2:0:0)
Reservation	ML-Institute-Mon
VNC Desktop Resolution	WIDTHxHEIGHT
<input type="button" value="Submit"/> <input type="button" value="Utilities"/>	

Waiting for Job to start (can take a minute)

TACC

| Analysis Portal

User Guide

ngaffney

TAP Job Status

Job: Jupyter notebook on Frontera (6314485, 2024-05-13T13:42-05:00)

Status: PENDING

Refresh: in 41 seconds

Message:

TAP: Your script has been submitted to Frontera but it is not yet running.

'sinfo -p development' output is:

PARTITION	AVAIL NODES(A/I)
development*	up 390/4

Check Status

End Job

Show Output

Back to Jobs

Jupyter notebook: Accessing shell

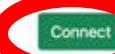
TACC | Analysis Portal User Guide  jrduncan Log Out

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

Setting up for the tutorial

- ▶ Note, the following setup only has to be done once.
- ▶ This is copying files into the right places so you can run these notebooks with us.



Jupyter notebook: Accessing shell

The screenshot shows the Jupyter Notebook interface. At the top, there is a navigation bar with tabs for "Files", "Running", and "Clusters". On the right side of the header, there are "Quit" and "Logout" buttons. Below the header, a message says "Select items to perform actions on them." To the left, a file tree shows a root directory "/" containing several subfolders: "adhoc", "archive", "bin", "CommunityData", "diplomacy", "diplomacy_searchbot", "ds-cuda-workshop", and "lm_accord". To the right of the file tree, a modal dialog box is open. This dialog has fields for "Name" (set to "Python 3"), "Notebook:" (set to "Python 3"), and "Other:" (set to "Text File"). A dropdown menu lists "Folder" and "Terminal". The "Terminal" option is highlighted with a red oval and is also enclosed in a blue rectangular box. Below the dropdown, the text "2 years ago" is visible. The footer of the interface features the TACC logo and the word "TEXAS".

Select items to perform actions on them.

0 /

Name: Python 3

Notebook: Python 3

Other: Text File

Folder

Terminal

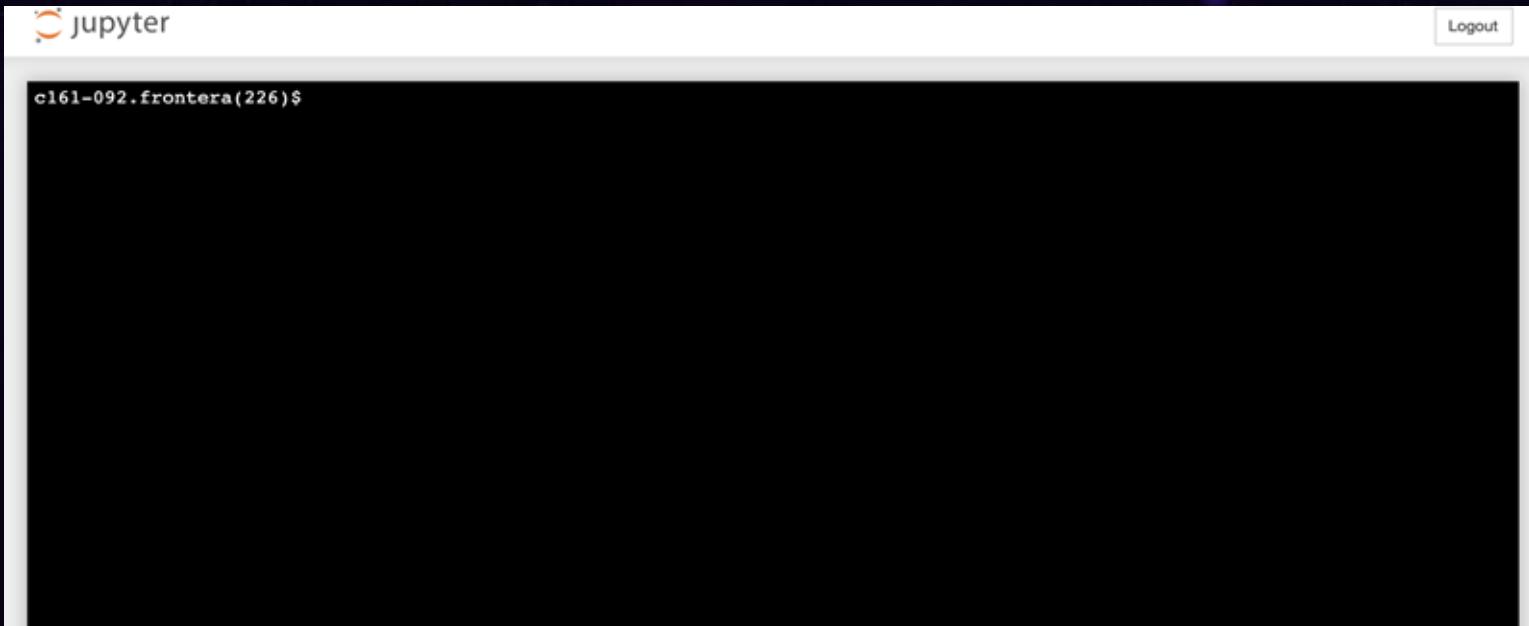
2 years ago

2 years ago

4 years ago

a year ago

Jupyter notebook: Accessing shell (4)



The screenshot shows a Jupyter Notebook interface with a terminal-like code cell. The top bar has the word "jupyter" and a "Logout" button. The code cell displays a command-line prompt: "c161-092.frontera(226)\$".

One time installation

From the command line enter

```
/scratch/install
```

After running this enter **ls** and make sure you see an entry for **institute_summer_25**

Stop the Session

TACC | Analysis Portal User Guide

ngaffney Log Out

TAP Job Status

Job: Jupyter notebook on Frontera (6301729, 2024-05-07T11:49-05:00)
Status: RUNNING
Start: 05/07/2024 11:49
End: 05/07/2024 13:49
Refresh: in 630 seconds
Message:
TAP: Your session is running at <https://frontera.tacc.utexas.edu:60069/>?token=H7Y2lZVbCVPpJRn9-U9tArNvynEX3DOtP29pqnmOYc

Connect End Job Show Output Back to Jobs

Back to the Jobs page

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

TAP Job Status

Job: Jupyter notebook on Frontera (6301729, 2024-05-07T11:49-05:00)
Status: RUNNING
Start: 05/07/2024 11:49
End: 05/07/2024 13:49
Refresh: in 630 seconds
Message:
TAP: Your session is running at <https://frontera.tacc.utexas.edu:60069/>?
token=H7Y2lZVbCVPpJRn9-U9tArNvynEX3DOtP29pqnm0Yc

[Connect](#) [End Job](#) [Show Output](#) [Back to Jobs](#)

- ▶ This is the end of the one time setup. Next we will make sure it worked and then get started with some notebooks about Python



Jupyter notebook: Accessing shell

- TACC Visualization Portal:

Go to <https://tap.tacc.utexas.edu>; Login with your training account
credentials

The screenshot shows the TACC Analysis Portal interface. On the left, there's a 'Submit New Job' form with fields for System, Application, Project, Queue, Nodes, Tasks, Job Name, Time Limit, Reservation, and VNC Desktop Resolution. Below the form are 'Options' like Job Name, Time Limit, Reservation, and VNC Desktop Resolution. At the bottom are 'Submit' and 'Utilities' buttons. On the right, there's a 'System Status' table showing resource utilization for Frontera, Lonestar6, Longhorn, Maverick2, and Stampede2. Underneath is a 'Past Jobs' section listing completed jobs for JNB-Frontera with their submission date, details link, and a blue 'Resubmit' button. The 'Resubmit' button for the first job is circled in red.

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			
	Date	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

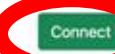
TACC | Analysis Portal User Guide  jrduncan Log Out

TAP Job Status

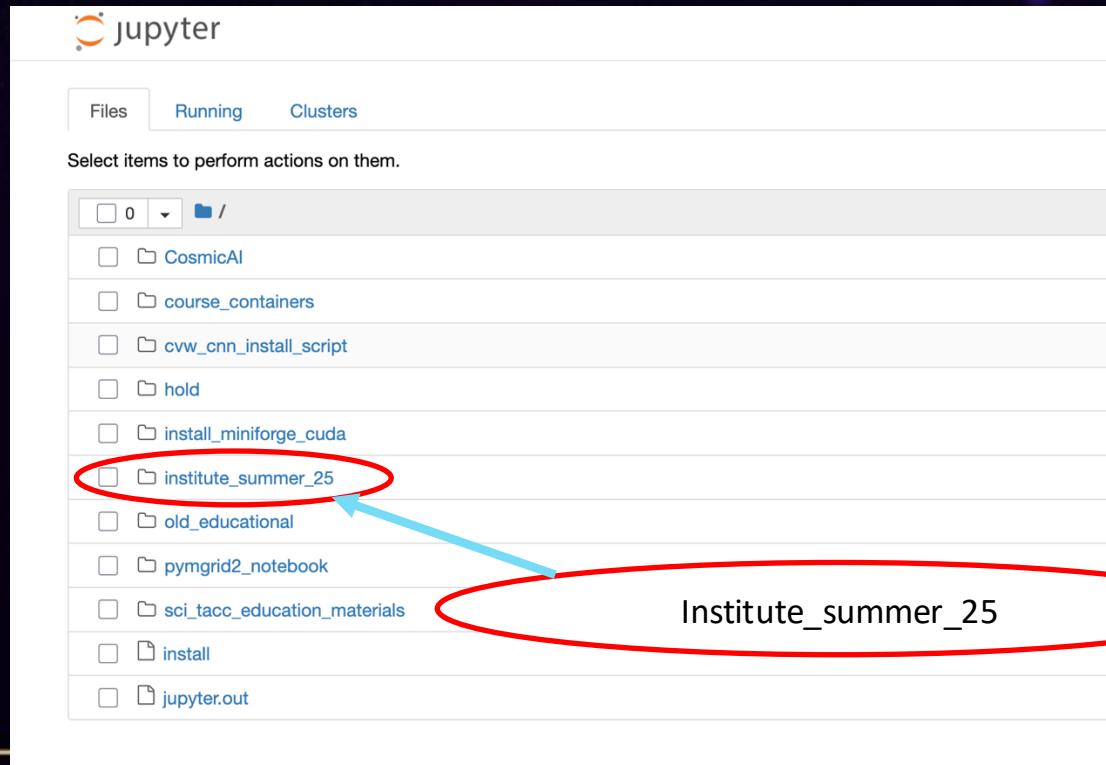
Job: Jupyter notebook on Frontera (4175197, 2022-03-21T17:28-05:00)
Status: RUNNING
Start: March 21, 2022, 5:28 p.m.
End: March 21, 2022, 5:33 p.m.
Refresh: in 873 seconds

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 New Python Notebook



Jupyter Notebook: Open Existing File

- Click on the `institute_summer_25` directory in the list
- Click on **KNNs** directory in the list
- Click on the file notebook file `intro_sklearn.ipynb`
- A Jupyter notebook should launch in a new tab
- Check to make sure you can run this notebook in the containerized kernel

Thanks!!

- Any Issues on TACC Systems that we can not help you with?
 - Open a ticket at <https://tacc.utexas.edu/about/help/>

Supplemental Slides

File Transfer Mechanisms

- Copying files from your local machine to TACC system
- Mac/OSX or Linux
 - Open Terminal window
 - Use scp (Secure Shell Copy)
 - Copies over ssh connection
 - -r option for recursive



```
scp <local file name> user@frontera.tacc.utexas.edu:
```

```
scp -r <local folder name> user@frontera.tacc.utexas.edu:
```

File Transfer Mechanisms (2)

- Copying files from your local machine to TACC system
- Windows
 - Download WinSCP (<https://winscp.net>)
 - Double click installer
 - Choose “Explorer” Option
 - Hostname: frontera.tacc.utexas.edu
 - Username: username
 - TACC Code: <Use Code from your credentials handout>
- Drag and Drop Files/Folders to copy to Stampede2

Command Line Editing

- Often will be required
 - Working with code
 - Editing configuration files
 - Writing Bash Scripts
 - TACC Systems: Writing job submission scripts
- Vim
 - Vi Improved
- Emacs
- Nano
 - Simple, Easy to use



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](#)

Linux/bash

- TACC machines use the Linux operating system
- Default shell and most popular is bash (Bourne Again Shell)
- Bash features auto completion of commands with tab key
- Bash commands are case sensitive

Example:
The “man”
command

man ≠ Man ≠ MAN

```
[jhpowell@isp02 ~]$ man
What manual page do you want?
[jhpowell@isp02 ~]$ Man
-bash: Man: command not found
[jhpowell@isp02 ~]$ MAN
-bash: MAN: command not found
[jhpowell@isp02 ~]$ MaN
-bash: MaN: command not found
[jhpowell@isp02 ~]$ _
```