

OPEN ON DEMAND 101

Rosen Center for Advanced Computing

August 2022



Agenda

- Background on Clusters and Open On Demand
- Introduction to Open on Demand Interface
- Launching Interactive Applications
- Overview of Installing Packages in R and Python
- Questions

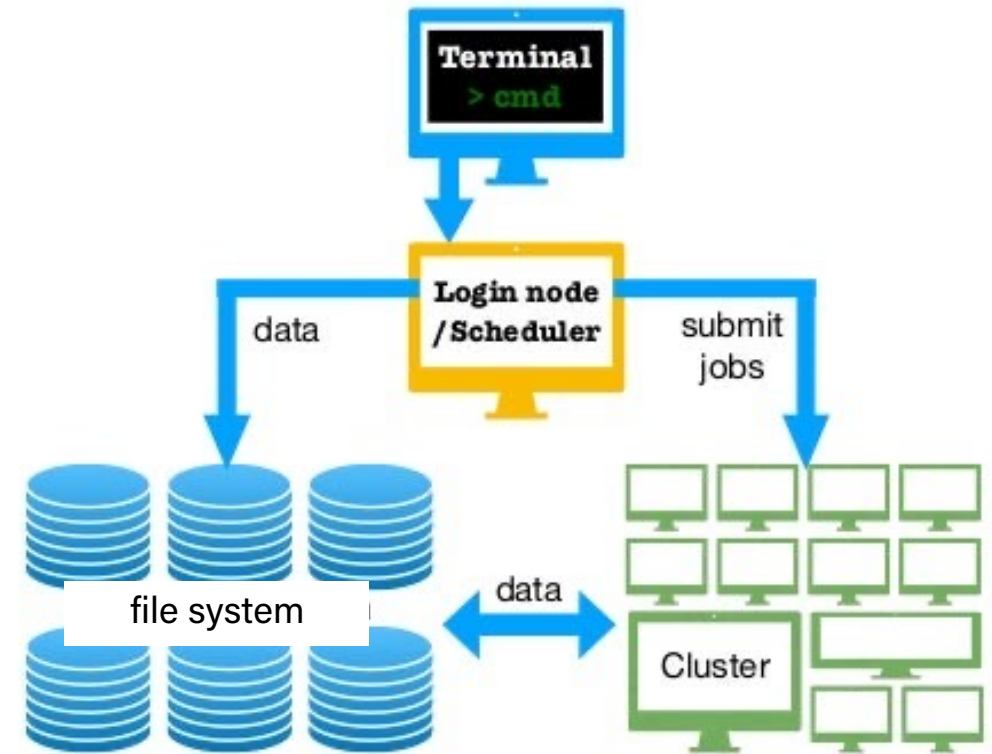
Why use a Cluster?

- Simplified environment set-up and standardization across group (no local installation of software needed)
- Easy collaboration and data sharing
- Built-in redundancies protect against data loss in the case of hardware failure
- RCAC/IT Support for Technical Issues
- Scalable and discounted storage and computing power compared to purchasing individually



HPC Paradigm

- Compute resources (back-end) requested by submitting jobs (via front-end nodes)
 - Jobs specify the resources needed and what should be run
- SLURM scheduler allocates the requested resources to each job
 - Once allocated, resources are exclusive to your job for the duration of your job
- Standard way to use the cluster is via the command line
- **Problem: steep learning curve and little support for software with graphical interfaces**



Bell Cluster

- 8 front-end nodes for submitting jobs
- 450+ compute nodes
 - 2 processors
 - 128 cores
 - 256 GB of memory each
- Shared resource for work with heavy compute requirements



Scholar Cluster

- Designed for classroom use and for teaching about HPC
- More limited compute resources available than on research clusters
- More powerful front-end nodes than typical clusters
- Both CPU and GPU nodes
- Primarily used for interactive jobs



Requesting a Class Allocation on Scholar

- Go to <https://www.rcac.purdue.edu/compute/scholar>
- Click on 'Class Account Request' and fill out the form
 - Need to manually add TAs and auditors
- Student accounts are created the week prior to the start of the semester and removed one week after the grade deadline. Instructor and TA accounts are created overnight.

Overview of Scholar

Scholar is a small computer cluster, suitable for classroom learning about high performance computing (HPC). It consists of 7 interactive login servers and 28 batch worker nodes.

It can be accessed as a typical cluster, with a job scheduler distributing batch jobs onto its worker nodes, or as an interactive resource, with software packages available through a desktop-like environment on its login servers.

If you have a class that you think will benefit from the use of Scholar, you can schedule it for your class through our web page at: **Class Account Request**. You only need to register your class itself. All students who register for the class will automatically get login privileges to the Scholar cluster.

What is Open on Demand (OOD)?

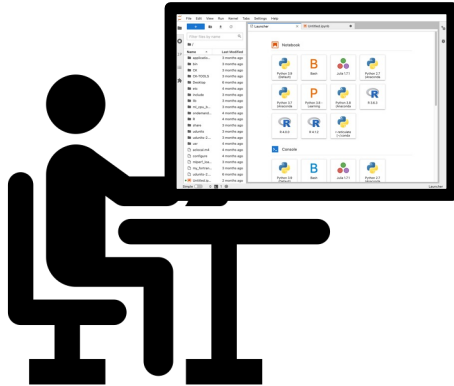
- Open-source portal to interact with resources
 - Web browser portal
 - No installations/plugins
 - User-friendly graphical interface
 - Allows file management, job monitoring/management, and desktop environments/applications
 - Jupyter, R, Matlab, Virtual Desktop
 - Soon: Stata, SAS
- Developed by Ohio Supercomputer Center
 - Widely used for academic HPC clusters today



Why OOD?

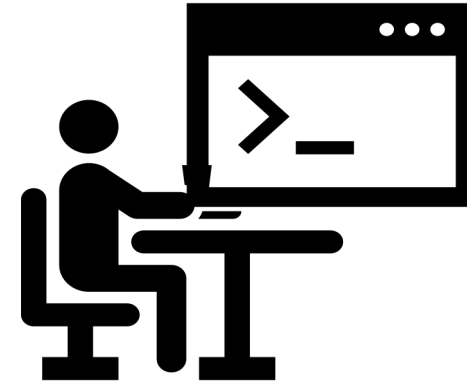
PROS

- Simple to use/learn – abstracts the complexity of submitting jobs
- GUI-Application friendly
- Does not require local software installation



CONS

- Not as fast as command line
- Can be difficult to submit complex, multi-task jobs



OOD Is Ideal For Most Krannert User Cases

Getting Started

- Navigate to gateway.bell.rcac.purdue.edu for Bell or gateway.scholar.rcac.purdue.edu for Scholar and log in with your BoilerKey
- OOD can also be launched via the RCAC website (<https://www.rcac.purdue.edu/compute/bell> or <https://www.rcac.purdue.edu/compute/scholar>)
- Everyone should have an account set up and be able to log in



[Outages & Maintenance](#)

[Bell User Guide](#)

Gateway

Launch

Remote
Desktop



Launch

Overview of Bell

Bell is a Community Cluster optimized for communities running engineering applications. Bell is being built through a partnership of 2020. Bell consists of Dell compute nodes with two 64-core nodes (64 cores per node) and 256 GB of memory. All nodes have 100 Gb/s network and 6-year warranty.

New with Bell is that access is being offered on the basis of a node share. To purchase access to Bell today, go to the [Cluster](#) page and subscribe to our Community Cluster Program Mailing List to stay updated.

Navigation

Bell - Gateway Files ▾ Jobs ▾ Interactive Apps ▾ Documentation ▾ Cluster ▾  </> ▾ ? ▾  Log Out

OnDemand provides an integrated, single access point for all of your HPC resources.

See documentation for more help in using OnDemand.

Current Quota Usage

[/home/srodenb](#)

Updated 1 week, 5 days, 4 hours, and 34 minutes ago

Using 11.7 GB of quota 25 GB

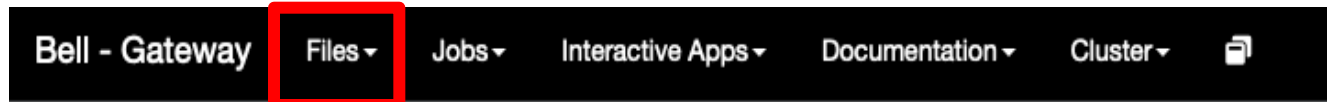
46%

No file count limit.

Navigation



Likely to be commonly used



OnDemand provides an integrated, single access point for all of your HPC res

See documentation for more help in using OnDemand.

Current Quota Usage

[/home/srodenb](#)

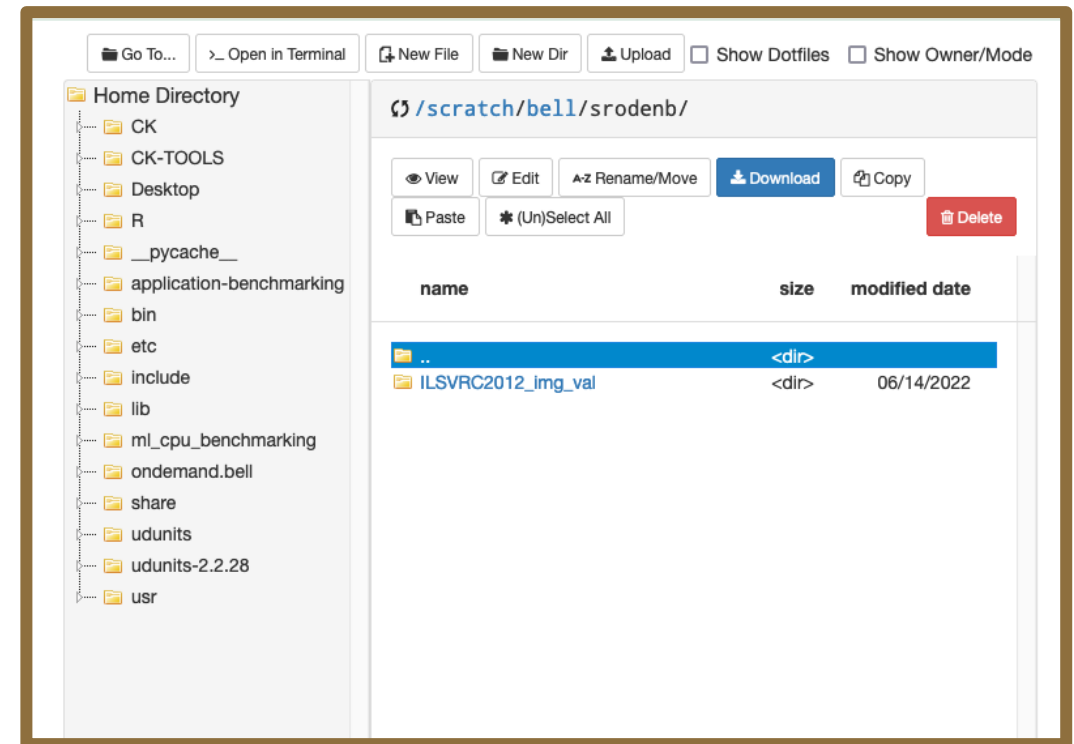
Updated 1 week

Using 11.7 GB of quota 25 GB

46%

No file count limit.

- Links to file explorer interface for your storage spaces
 - Everyone should have a /depot/mylab space allocated

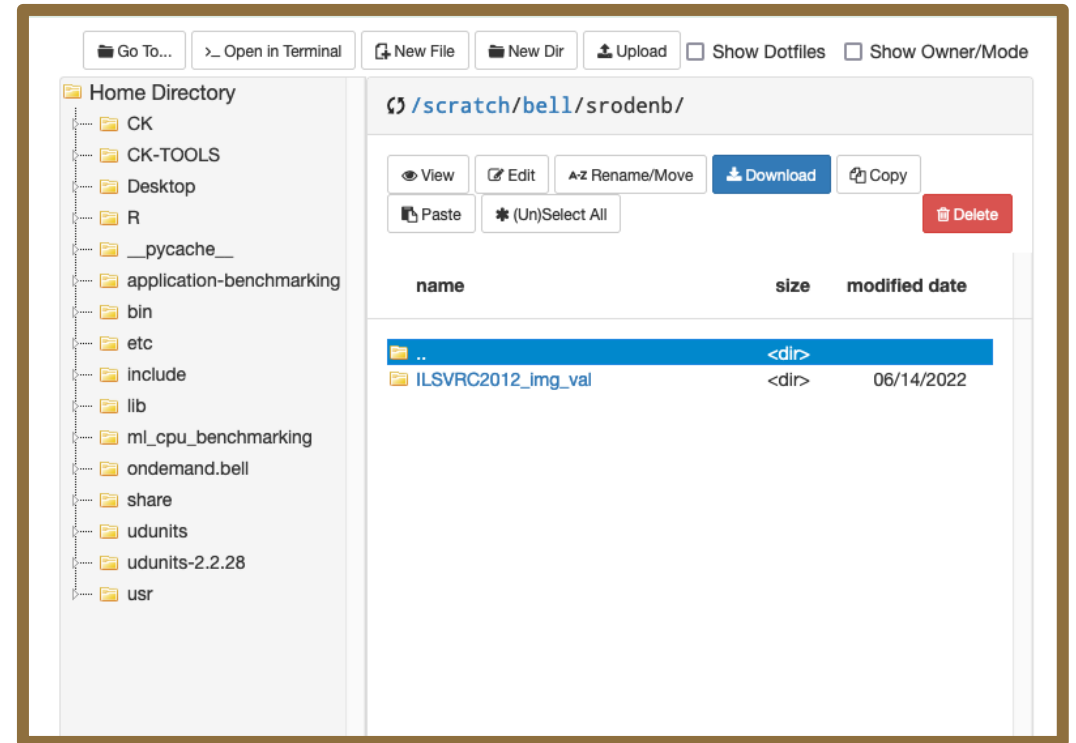


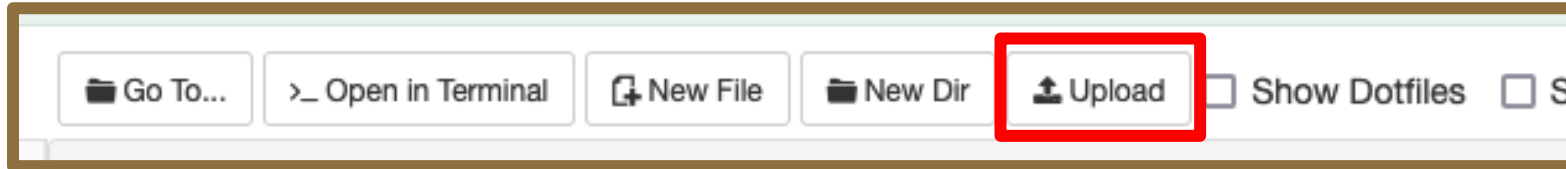
File Storage Systems



Likely to be commonly used

- /home/myusername
 - Relatively small quota (25 GB)
 - Not purged
 - Data not easily sharable with other users
 - Easy access
- /depot/mylabname
 - Lab-wide storage and data sharing
 - Not purged
 - Initial allocation 100 GB; can purchase more for \$70/TB
- /scratch/clustername/myusername
 - Temporary storage – large quota (200 TB on Bell/1 TB on Scholar) but unused files purged after 30 days
 - Easy, fast access





- For smaller files, the upload button can be used to transfer files from a local machine
 - Pro: Native integration with OOD
 - Con: Not ideal for very large file transfers (gigabyte+ scale)
- **Globus** is a good option for large transfers
 - Fast transfers over good networks
 - Robust transfers over flaky networks





- “Fire and forget”
 - Globus monitors the transfer, auto-resumes on errors, and sends an email at the end
- Your computer is only used for the command channel
 - If the transfer is not from your computer, your computer does not have to stay on





- Login to Globus
 - transfer.rcac.purdue.edu or globus.org
 - Select "Purdue University Main Campus" as Organization.
 - Will be taken to the BoilerKey 2FA page

Docs: docs.globus.org/how-to/get-started/

Log in to use Purdue Globus Web App


Use your existing organizational login

e.g., university, national lab, facility, project

Purdue University Main Campus ▼

By selecting Continue, you agree to Globus [terms of service](#) and [privacy policy](#).

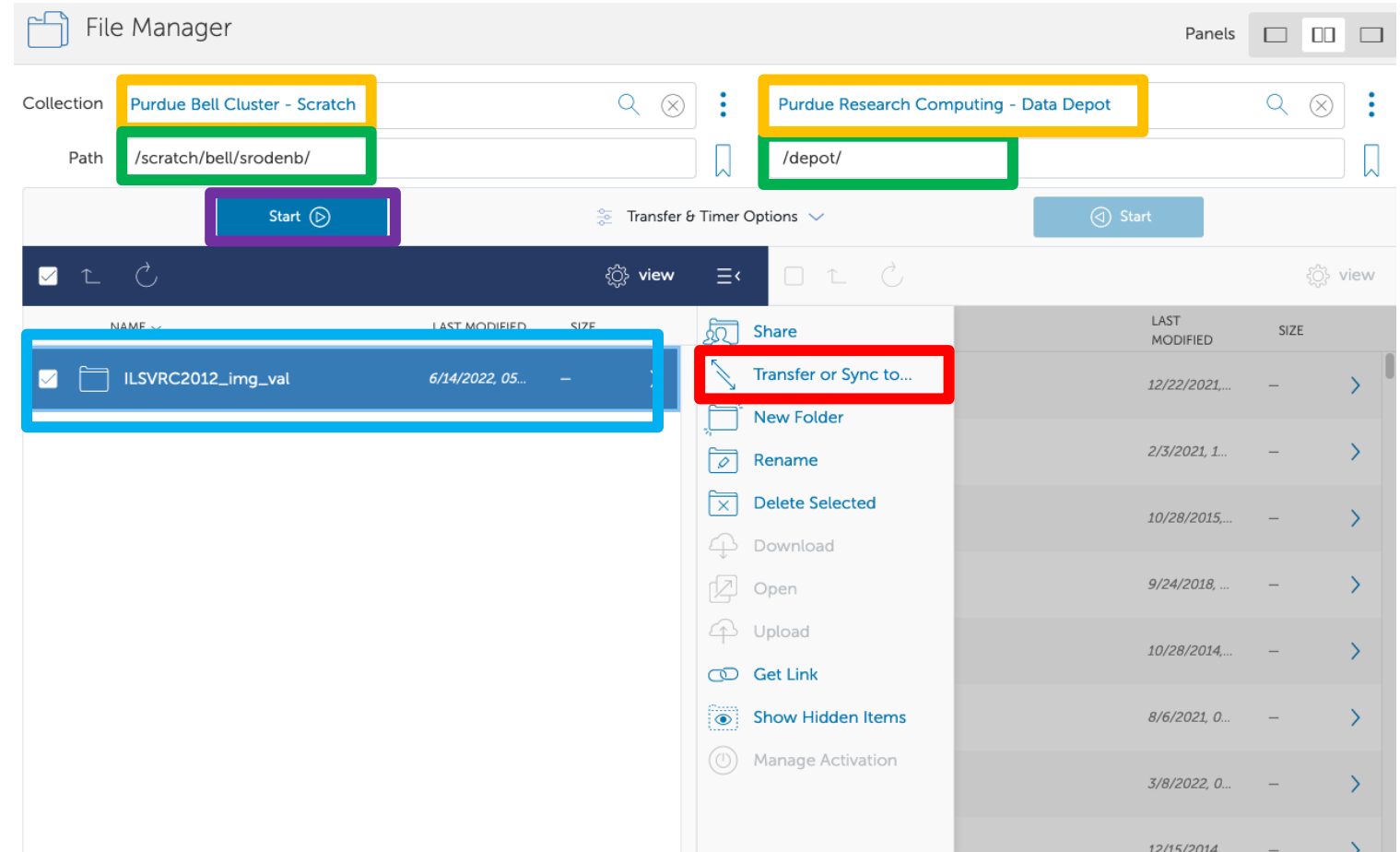
Continue



Globus uses CI Logon to enable you to Log In from this organization. By clicking Continue, you agree to the [CI Logon privacy policy](#) and you agree to share your username, email address, and affiliation with CI Logon and Globus. You also agree for CI Logon to issue a certificate that allows Globus to act on your behalf.



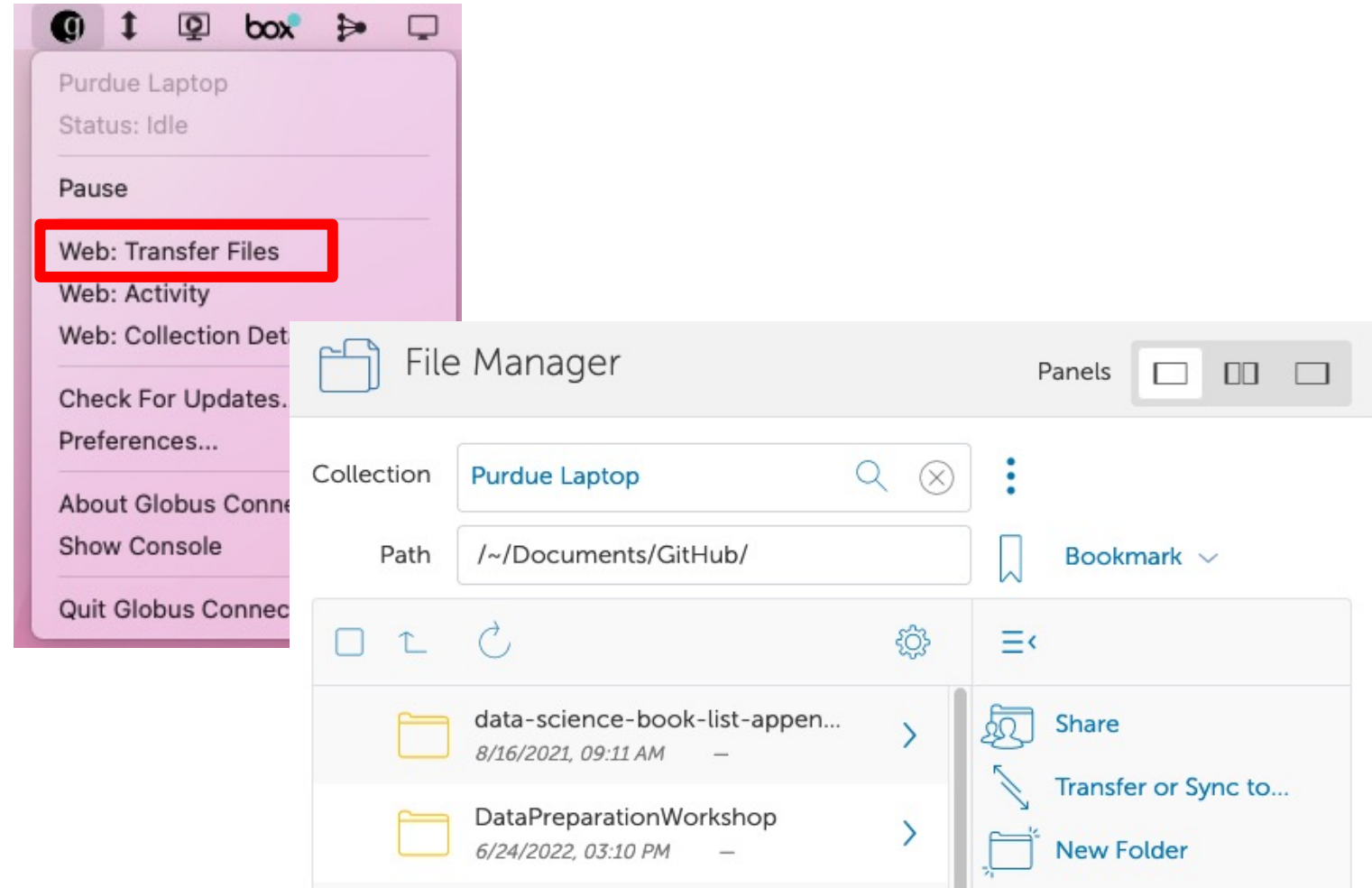
1. Click "Transfer or Sync to..."
2. Search for and select target and destination endpoints
3. Specify paths within the endpoints
4. Select the files/directories to be transferred
5. Click "Start"



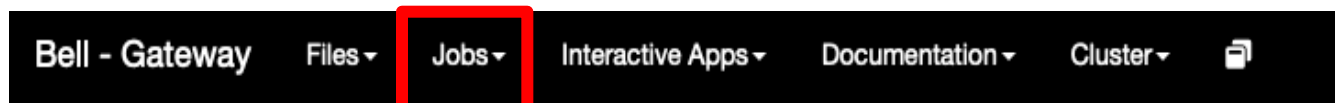


To transfer to/from your personal computer:

- Download Globus Connect Personal:
<https://app.globus.org/file-manager/gcp>
- Add new endpoint for your machine
- Once added, this runs in the taskbar
 - Click on “Web: Transfer Files” to open transfer interface



Navigation



OnDemand provides an integrated, single access point for all of your HPC res

See documentation for more help in using OnDemand.

Current Quota Usage

[/home/srodenb](#)

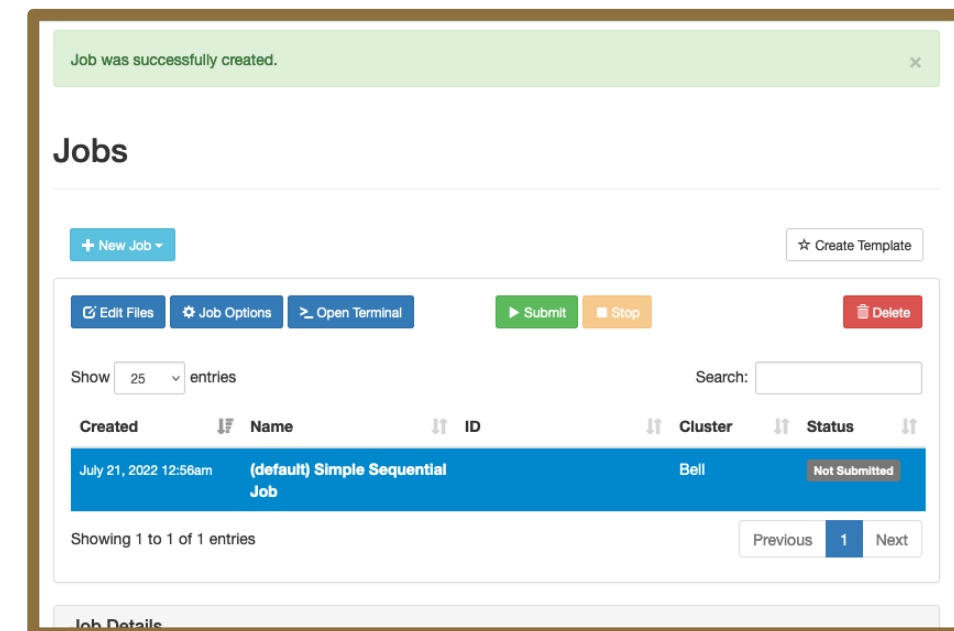
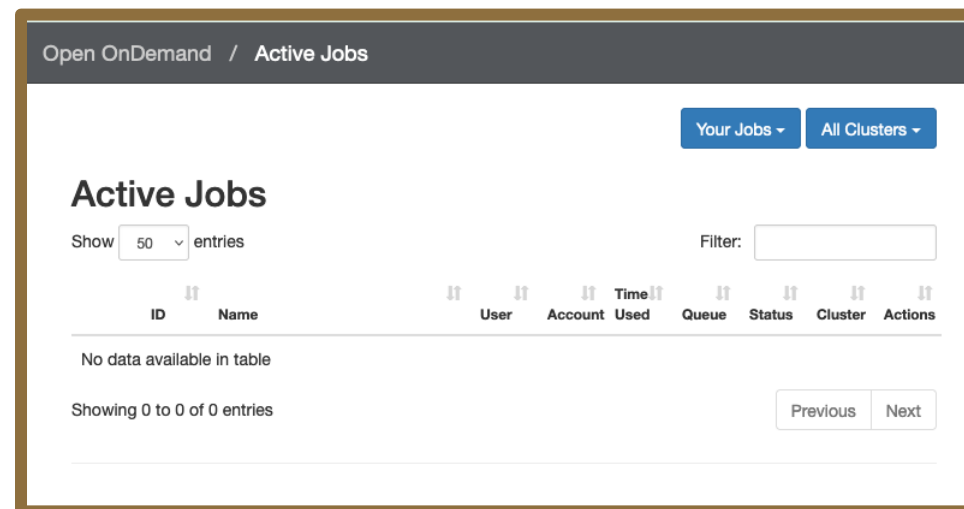
Updated 1 week

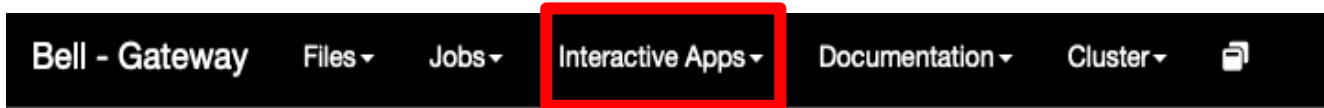
Using 11.7 GB of quota 25 GB

46%

No file count limit.

- Active jobs: Allows you to monitor any active jobs
- Job Composer: Walks you through creating and submitting a new job





OnDemand provides an integrated, single access point for all of your HPC res

See documentation for more help in using OnDemand.

Current Quota Usage

[/home/srodenb](#)

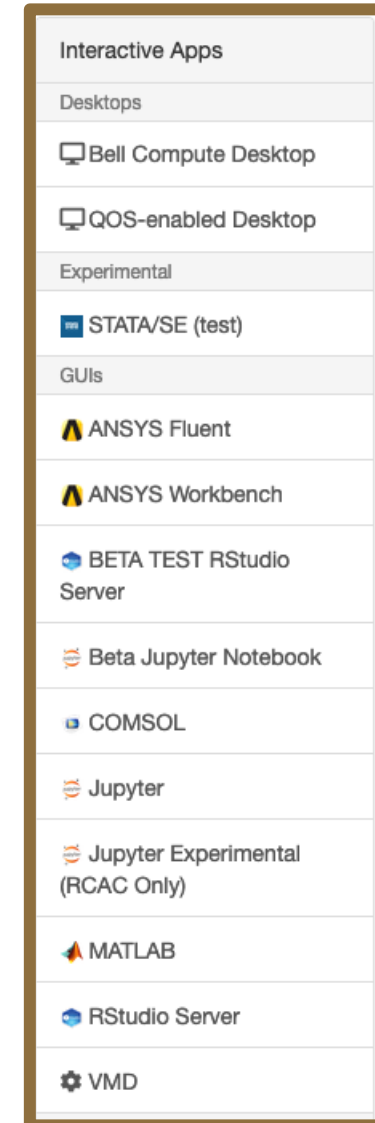
Updated 1 week

Using 11.7 GB of quota 25 GB


No file count limit.

46%

- Allows you to launch interactive apps with graphical user interfaces



Navigation

Bell - Gateway Files ▾ Jobs ▾ Interactive Apps ▾ Documentation ▾ Cluster ▾ 

OnDemand provides an integrated, single access point for all of your HPC res

See documentation for more help in using OnDemand.

Current Quota Usage

[/home/srodenb](#)

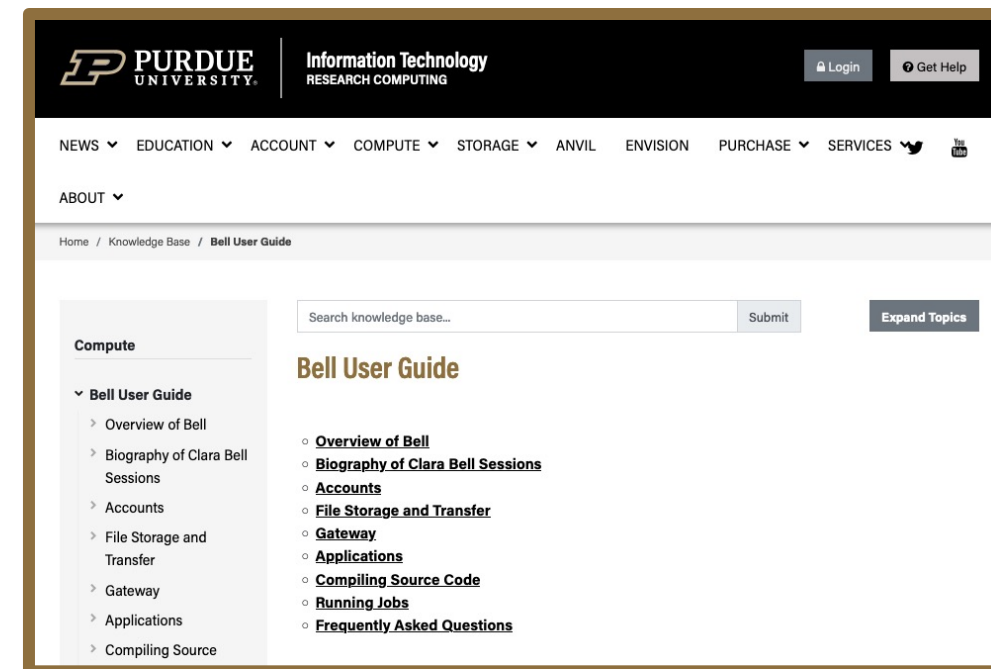
Updated 1 week

Using 11.7 GB of quota 25 GB

46%

No file count limit.

- Links to the User Guide



The screenshot shows the Purdue University Information Technology Research Computing website. The header includes the Purdue University logo, the text "Information Technology RESEARCH COMPUTING", and links for "Login" and "Get Help". A navigation bar contains links for "NEWS", "EDUCATION", "ACCOUNT", "COMPUTE", "STORAGE", "ANVIL", "ENVISION", "PURCHASE", and "SERVICES". Below this is an "ABOUT" link. The main content area is titled "Bell User Guide" and features a search bar with the text "Search knowledge base..." and a "Submit" button. A sidebar on the left lists the "Compute" section, which includes a "Bell User Guide" subsection with links to "Overview of Bell", "Biography of Clara Bell Sessions", "Accounts", "File Storage and Transfer", "Gateway", "Applications", and "Compiling Source". The main content area lists the "Bell User Guide" subsection with links to "Overview of Bell", "Biography of Clara Bell Sessions", "Accounts", "File Storage and Transfer", "Gateway", "Applications", "Compiling Source Code", "Running Jobs", and "Frequently Asked Questions".

Navigation

Bell - Gateway

Files ▾

Jobs ▾

Interactive Apps ▾

Documentation ▾

Cluster ▾



OnDemand provides an integrated, single access point for all of your HPC resources.

See documentation for more help in using OnDemand.

Current Quota Usage

[/home/srodenb](#)

Updated 1 week

Using 11.7 GB of quota 25 GB

46%

No file count limit.

- Allows you to launch a command line terminal

```
***** Use of Purdue BoilerKey or SSH keys is Required *****
*****
Last login: Mon Jul 18 15:15:27 2022 from vpnclient-10-195-24-110.itap.purdue.edu
=====
Welcome to the Bell Cluster
=====
Bell consists of:
=====
Nodes:
Bell-A   ppn=128   256 GB memory
Bell-B   ppn=128   1 TB   memory
Bell-G   ppn=128   256 GB memory 2x AMD MI50 32GB
Bell-X   ppn=48    384 GB memory 6x AMD MI50 32GB
=====
Scratch:
Quota: 200 TB / 2 million files
Path: $CLUSTER_SCRATCH
Type command: "myquota"
=====
Queues:
Type command: "qlist"
=====
Software:
Type command: "module avail" or "module spider"
=====
User guide:
www.rcac.purdue.edu/knowledge/bell
=====
Help:
www.rcac.purdue.edu/help
=====
News:
www.rcac.purdue.edu/news/Bell
=====
srodenb@bell-fe06:~ $
```

Interactive Apps

- Interactive apps allow you to use programs that have graphical interfaces
 - Very little difference from how you would use the program on a laptop!
- Upon selecting the interactive app you want it will walk you through launching a job
- Go ahead and select the application of your choice
- You can also launch a virtual desktop from which you can launch additional applications.



Jupyter

This app will launch a Jupyter Notebook or a Jupyter Lab session on the [Bell cluster](#) using the new jupyterhub module.

Queue

debug (Max 0.5 hours) ▾

Please select a queue from the drop-down and enter the number of hours below (up to the max listed above).

Number of hours

.5

Processor Cores requested

1 ▴ ▾

Please select the number of processor cores to dedicate to your job. Note: You will be allocated 2 GB of memory per requested core. If the session exceeds this threshold, the scheduler will terminate it.

☐ High-Priority

You may designate one job as High Priority for scheduling

☐ I would like to receive an email when the session starts

Launch Jupyter Lab or Jupyter Notebook

Jupyter Lab ▾

Launch

Launching a GUI App

- Applies to Jupyter, R, Stata
- Typical components on composer:
 - Queue
 - "Line" the job goes into to be scheduled
 - Duration of reservation
 - Limited by the maximum for the selected queue
 - Cores
 - Typically 1 unless you have extensive compute/memory requirements
 - Software version (if multiple installed)



Jupyter

This app will launch a Jupyter Notebook or a Jupyter Lab session on the [Bell cluster](#) using the new jupyterhub module.

Queue

debug (Max 0.5 hours) ▾

Please select a queue from the drop-down and enter the number of hours below (up to the max listed above).

Number of hours

.5

Processor Cores requested

1 ▾

Please select the number of processor cores to dedicate to your job. Note: You will be allocated 2 GB of memory per requested core. If the session exceeds this threshold, the scheduler will terminate it.

☐ High-Priority

You may designate one job as High Priority for scheduling

☐ I would like to receive an email when the session starts

Launch Jupyter Lab or Jupyter Notebook

Jupyter Lab ▾

Launch

Launching a GUI App

- It can take a few minutes for resources to be allocated and for your app to launch
 - Timing depends on how busy the cluster is and the resources you request

Jupyter (18840349) Queued

Created at: 2022-07-21 01:25:41 EDT Delete

Time Requested: 30 minutes

Session ID: [ec971964-1584-46ef-9cf8-64bb766e54bb](#)

Please be patient as your job currently sits in queue. The wait time depends on the number of cores as well as time requested.



Jupyter (18840345) 1 node | 1 core | Running

Host: [>_bell-a009.rcac.purdue.edu](#) Delete

Created at: 2022-07-21 01:23:53 EDT

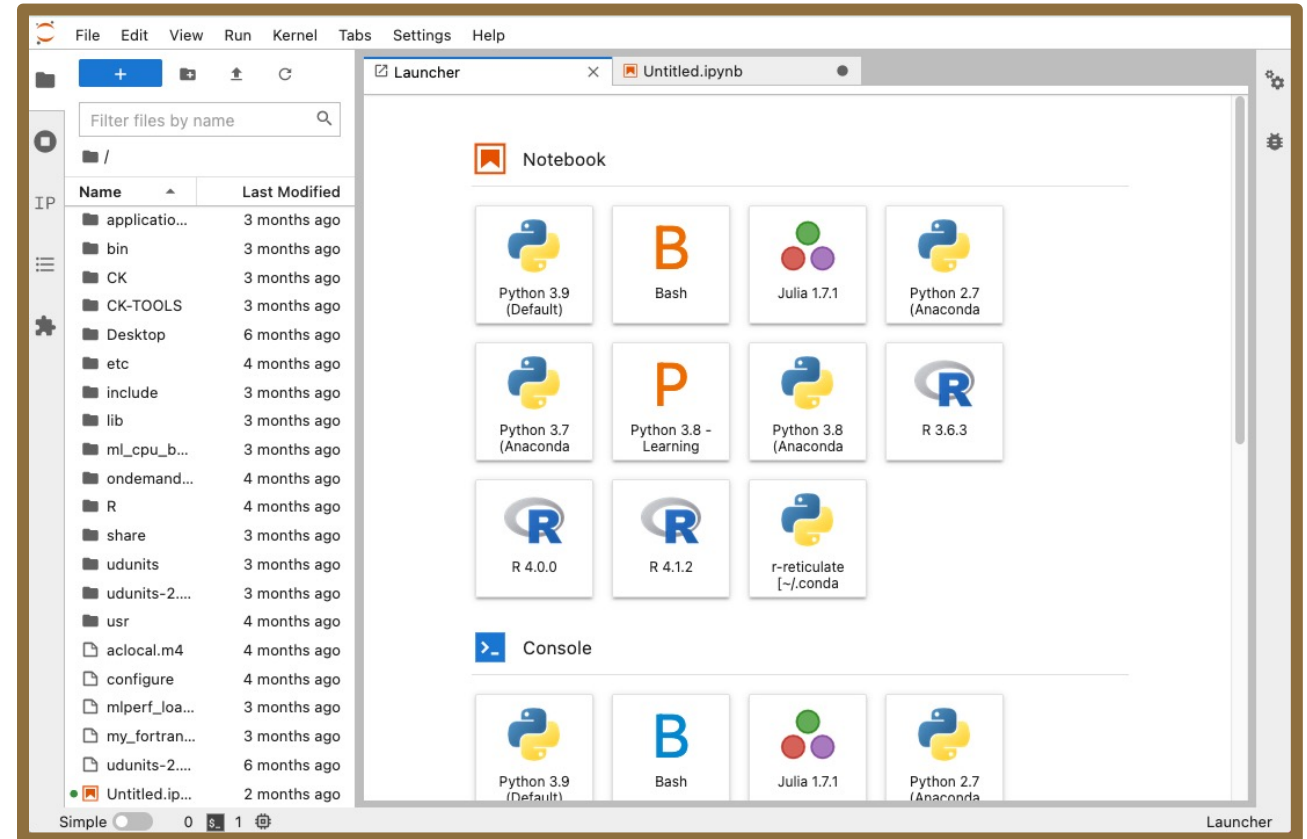
Time Remaining: 29 minutes

Session ID: [6e556107-39bf-46a4-87b5-6e712c0dd39e](#)

[⚙️ Connect to Jupyter](#)

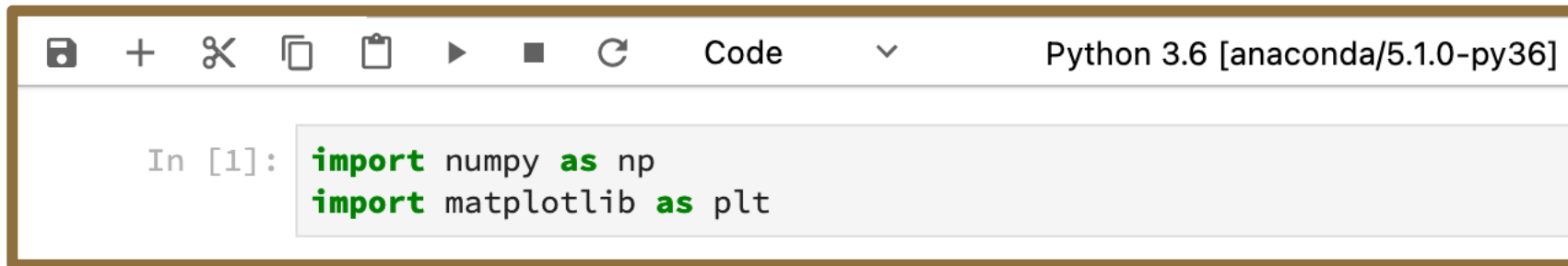
Launching a GUI App

- Once the application launches you should be able to interact with the application like you normally would!



Package Installation

- Many common packages may already be installed!
- Try to load/import the packages you need using the Interactive App you launched
- For any packages that need to be installed please reach out to rcac-help@purdue.edu
- Package installation is a one-time process!
 - Once packages installed all students can use the same installation from any node on the cluster

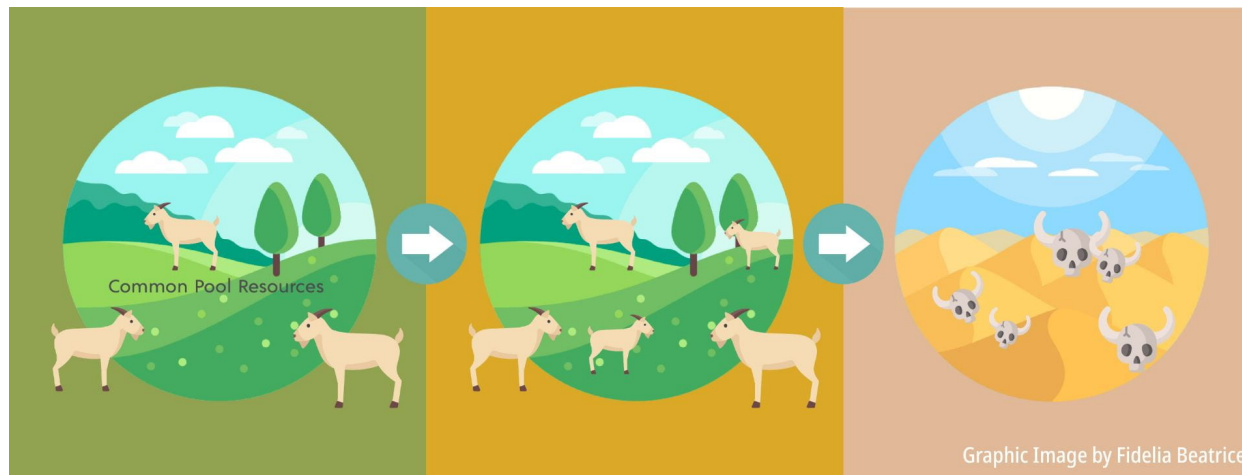


The screenshot shows a Jupyter Notebook interface. The top toolbar includes icons for saving, adding, deleting, copying, pasting, running, and refreshing, along with a dropdown menu currently set to 'Code'. The environment is identified as 'Python 3.6 [anaconda/5.1.0-py36]'. The active code cell, labeled 'In [1]:', contains the following Python code:

```
import numpy as np
import matplotlib as plt
```

Best Practices and Being a Good Steward

- OOD interactive apps are designed to be used for developing and running code/visualizations
 - Not designed for long-running production jobs
 - When requesting resources keep in mind that the cluster is a shared resource
 - Don't request 2+ cores unless your code needs the memory or is designed for that
 - Limit requested wall time to how long you will actively be working
- Shut down applications/delete the job if you finish before time runs out
 - This frees up resources for other users



QUESTIONS?

Contact rcac-help@purdue.edu with any technical problems

APPENDIX

Streamlined Instructions for OOD

- Navigate to gateway.bell.rcac.purdue.edu for Bell or gateway.scholar.rcac.purdue.edu for Scholar and log in with your BoilerKey
- OOD can also be launched via the RCAC website (<https://www.rcac.purdue.edu/compute/bell> or <https://www.rcac.purdue.edu/compute/scholar>)
- Everyone should have an account set up and be able to log in



[Outages & Maintenance](#)

[Bell User Guide](#)

Gateway

Launch

Remote
Desktop

Launch

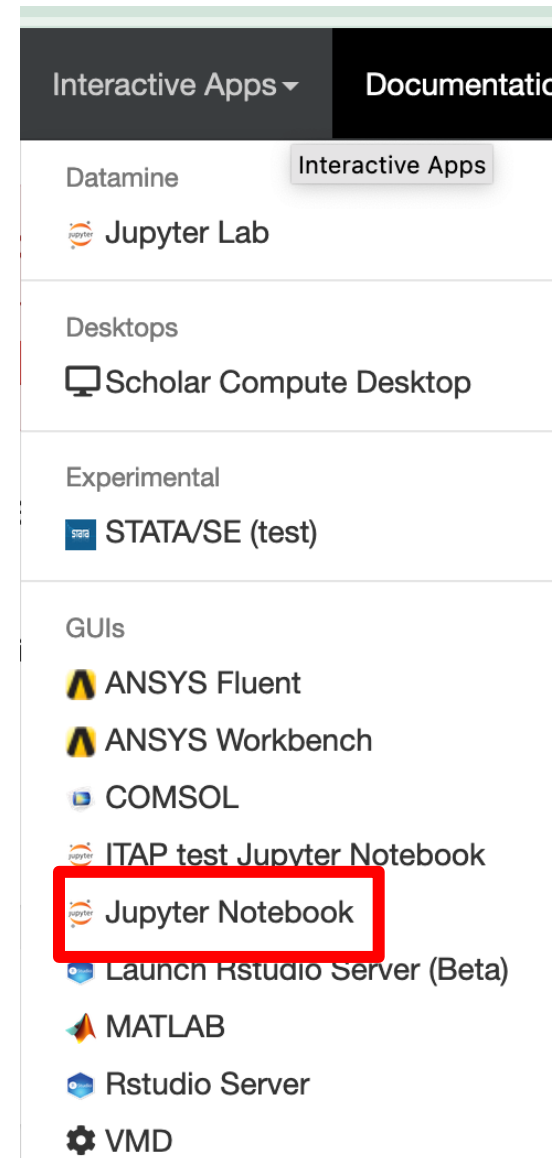
Overview of Bell

Bell is a Community Cluster optimized for communities running engineering applications. Bell is being built through a partnership of 2020. Bell consists of Dell compute nodes with two 64-core nodes (64 cores per node) and 256 GB of memory. All nodes have 100 Gb/s network and 6-year warranty.

New with Bell is that access is being offered on the basis of a node share. To purchase access to Bell today, go to the [Cluster](#) page and subscribe to our Community Cluster Program Mailing List to stay updated.

Streamlined Instructions for OOD

- Select the "Interactive Apps" Tab and GUI that you wish to use



Streamlined Instructions for OOD

- Most users should select the `scholar` queue
- Enter how many hours you would like the reservation to last (0-4 hours)
 - Do not select a time longer than you will reasonably be continuously working on the cluster
- If prompted for number of processor cores, select 1
- For jobs 30 minutes or less, select `debug` queue and use .5 as the number of hours
- Click `Launch`



Jupyter

This app will launch a Jupyter Notebook or a Jupyter Lab session on the [Bell cluster](#) using the new jupyterhub module.

Queue

debug (Max 0.5 hours) ▾

Please select a queue from the drop-down and enter the number of hours below (up to the max listed above).

Number of hours

.5

Processor Cores requested

1 ▾

Please select the number of processor cores to dedicate to your job. Note: You will be allocated 2 GB of memory per requested core. If the session exceeds this threshold, the scheduler will terminate it.

☐ High-Priority

You may designate one job as High Priority for scheduling

☐ I would like to receive an email when the session starts

Launch Jupyter Lab or Jupyter Notebook

Jupyter Lab ▾

Launch

Streamlined Instructions for OOD

- It can take a few minutes for resources to be allocated and for your app to launch
- Once your reservation is ready click `Connect`

Jupyter (18840349) Queued

Created at: 2022-07-21 01:25:41 EDT Delete

Time Requested: 30 minutes

Session ID: [ec971964-1584-46ef-9cf8-64bb766e54bb](#)

Please be patient as your job currently sits in queue. The wait time depends on the number of cores as well as time requested.



Jupyter (18840345) 1 node | 1 core | Running

Host: [>_bell-a009.rcac.purdue.edu](#) Delete

Created at: 2022-07-21 01:23:53 EDT

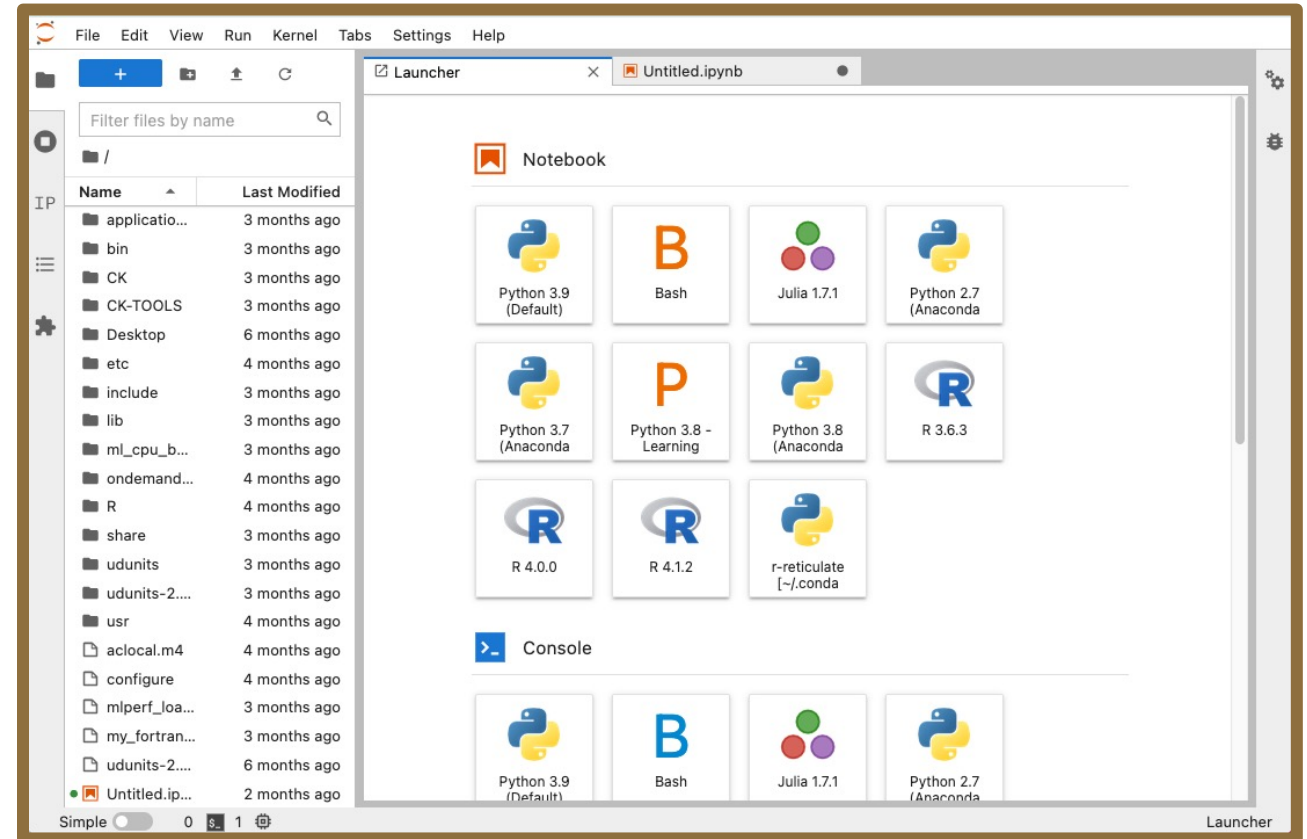
Time Remaining: 29 minutes

Session ID: [6e556107-39bf-46a4-87b5-6e712c0dd39e](#)

[⚙️ Connect to Jupyter](#)

Streamlined Instructions for OOD

- Once the application launches you should be able to interact with the application like you normally would!



Streamlined Instructions for OOD

- When finished, shut down your application and delete the job if your allocation has not yet expired

Jupyter Notebook (152710)1 node | 1 core | Running

Host: >_scholar-a000.rcac.purdue.edu

Created at: 2022-08-15 13:09:54 EDT

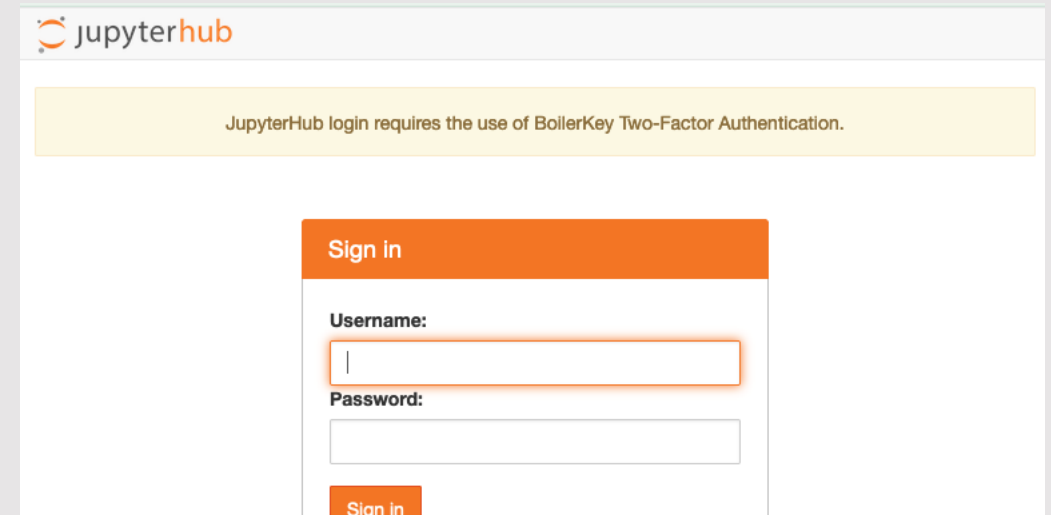
Time Remaining: 29 minutes

Session ID: ae5b4de4-d2b7-4670-84d8-13392ff7ce94

Delete

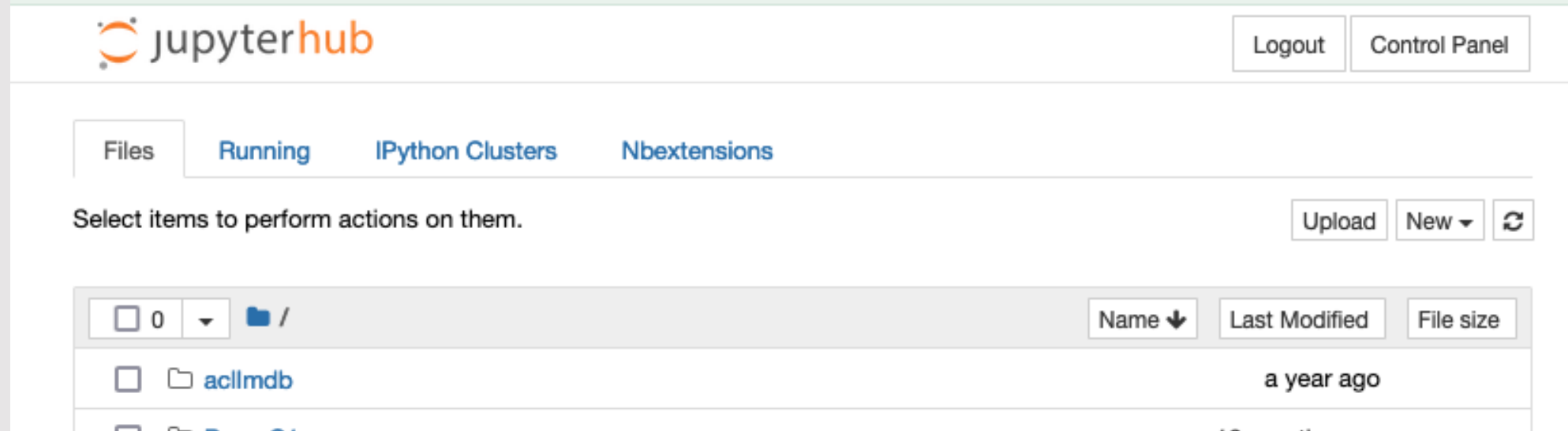
Notebooks

- Alternative method to access Scholar resources that does not require job submissions
 - Immediately able to run notebooks once accessing the URL
- **However, in the long term this option will be going away to better manage compute resources**
 - Using this option should be seen as a temporary solution in conjunction with further discussions with RCAC if OOD does not work for your purpose



Go to <https://notebook.scholar.rcac.purdue.edu/> and log in with your BoilerKey credentials

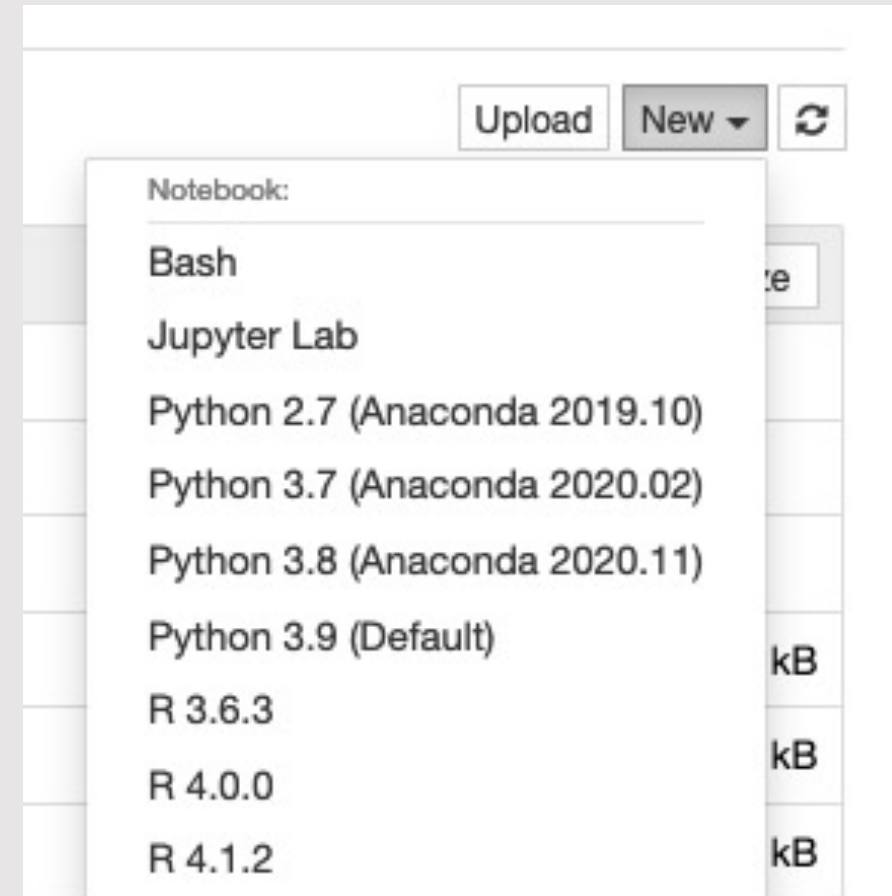
Notebooks



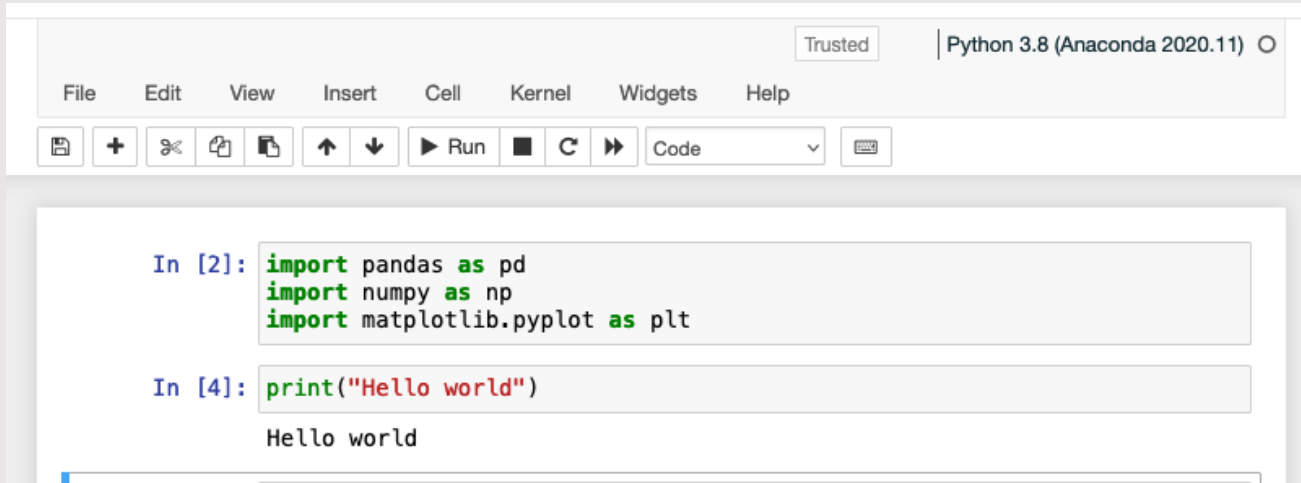
Once you log in you will be directed to a page that looks like this

Notebooks

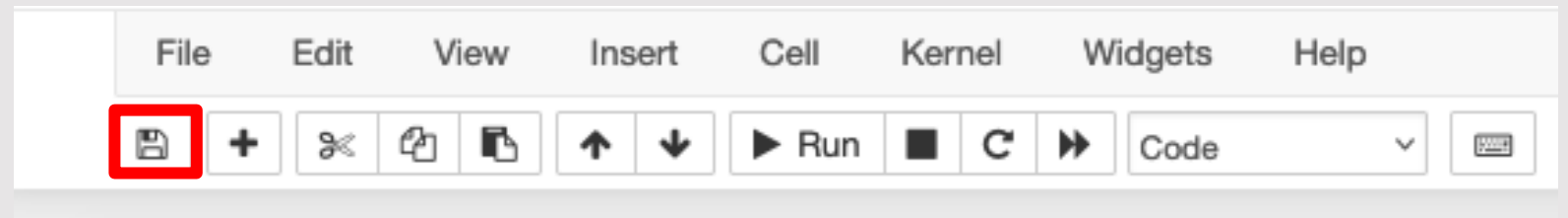
- **To create a new file:** Click “New” in upper right of the screen and select the R or Python version you have been instructed to use. In general, it is recommended to use newer versions of Python/R whenever possible.
- **If you already have a file:** Navigate through folders as needed and click on an existing .ipynb file to open it up.



Notebooks

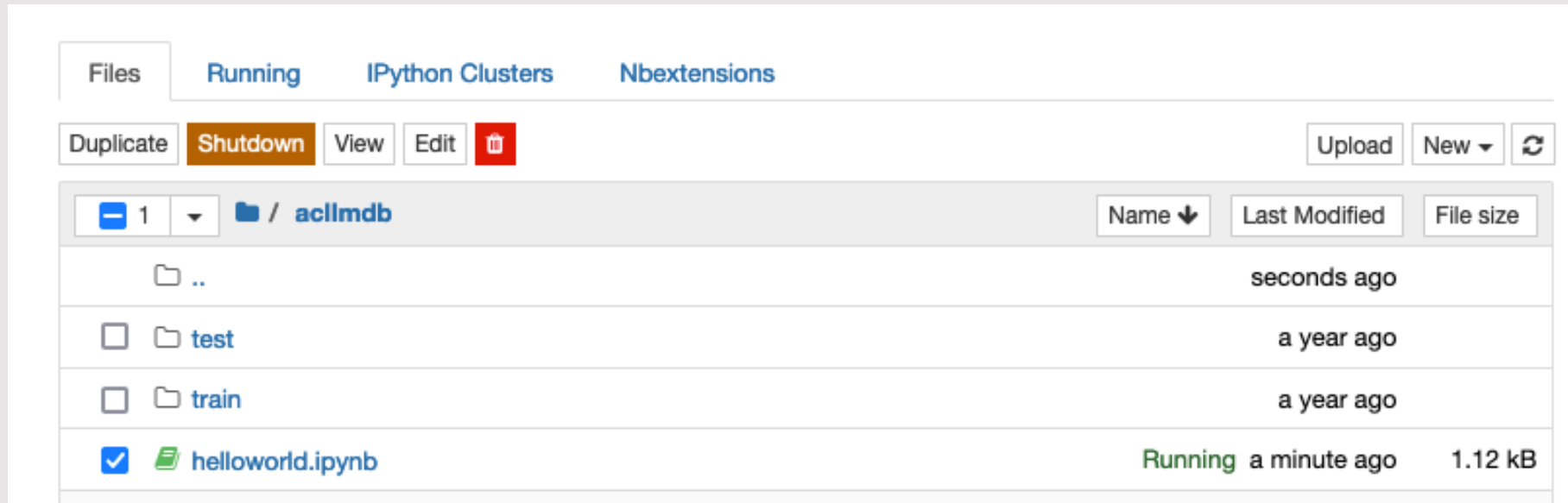


Begin coding!



Once you are done editing save the file and close the tab.

Notebooks



It is best practice to also shut down your kernel when you are done. Select the file name and you will see an option appear to “Shutdown” your kernel.