Getting Started on the FASRC clusters with Open OnDemand

Where to access the slides

https://docs.rc.fas.harvard.edu/kb/training-materials/

Learning objectives

- What is Open OnDemand (OOD)?
- Accessing OOD
- Launching apps
- RStudio Server
- Jupyter Notebook
  - Create conda environment (i.e., jupyter kernel)
- Files tab
- Jobs tab
- Remote Desktop
- FASSE proxy

Some definitions

- **OOD**: Open On Demand
- **FASRC**: Faculty of Arts and Sciences Research Computing
- **Cluster**: large group of servers with lots of memory and processors
- **Cannon**: cluster that handles level 2 data. Named after the 19th century Harvard astronomer Annie Jump Cannon.
- **FASSE**: cluster that handles level 3 data. FAS Secure Enclave.

Glossary of these terms: docs.rc.fas.harvard.edu/kb/glossary

What is Open OnDemand (OOD)?

- Open-source web portal to access clusters
- Web-based
  - Uses modern browser like Google Chrome, Mozilla Firefox, or Microsoft Edge.
    - Safari does not support all of OOD’s features
  - No software other than a browser needs be installed on your local laptop/desktop
- Easy to learn and use
- Very similar to desktop applications
- The easiest way to run graphical applications remotely on a cluster

OPEN OnDemand

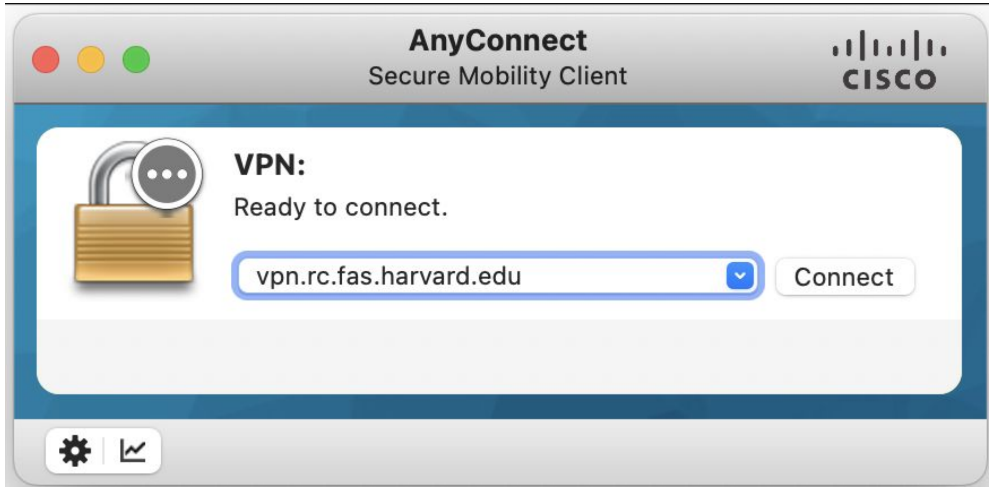
How to access OOD on FASRC Clusters

1. Get an account
    - You can choose a username. The default is first initial, last name. I'll use jharvard as an example
    - This is NOT necessarily the same as your HarvardKey username.
2. Log onto the FASRC VPN. This is NOT the generic Harvard VPN.
    - vpn.rc.fas.harvard.edu
    - username is
        - jharvard@fasrc (Cannon)
        - jharvard@fasse (FASSE)
3. Log into Open On Demand for your cluster
    - username is just jharvard, nothing else

      Cannon link: https://rcood.rc.fas.harvard.edu

      FASSE link: https://fasseood.rc.fas.harvard.edu

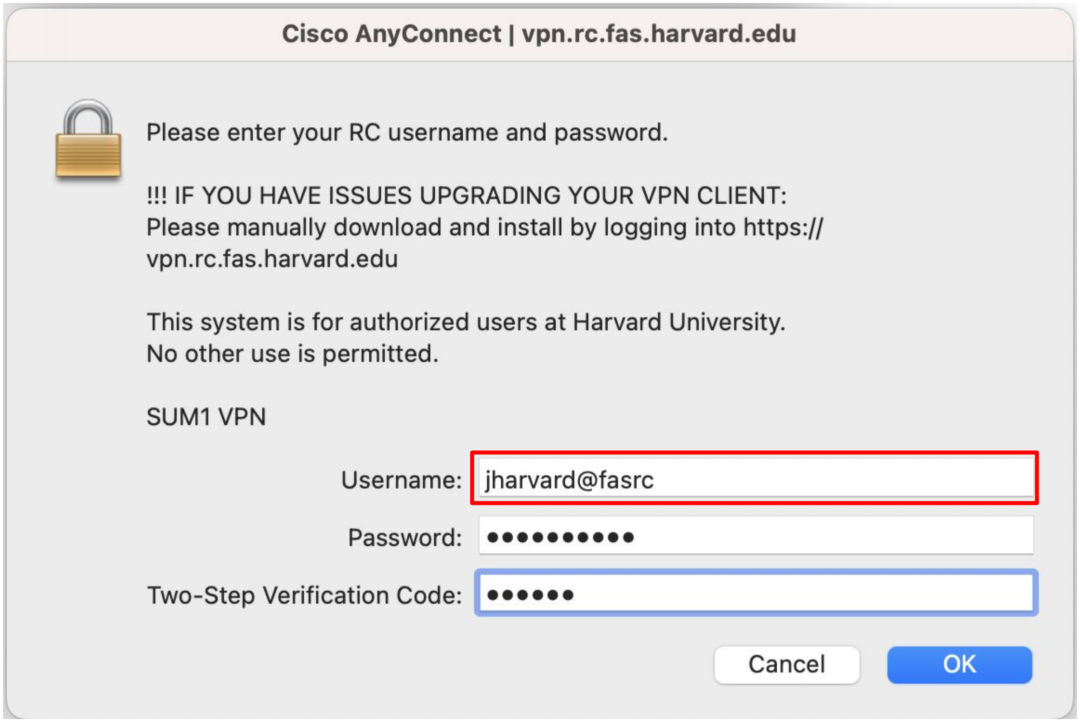
Connecting to VPN



AnyConnect
Secure Mobility Client

VPN:
Ready to connect.

vpn.rc.fas.harvard.edu



Cisco AnyConnect | vpn.rc.fas.harvard.edu

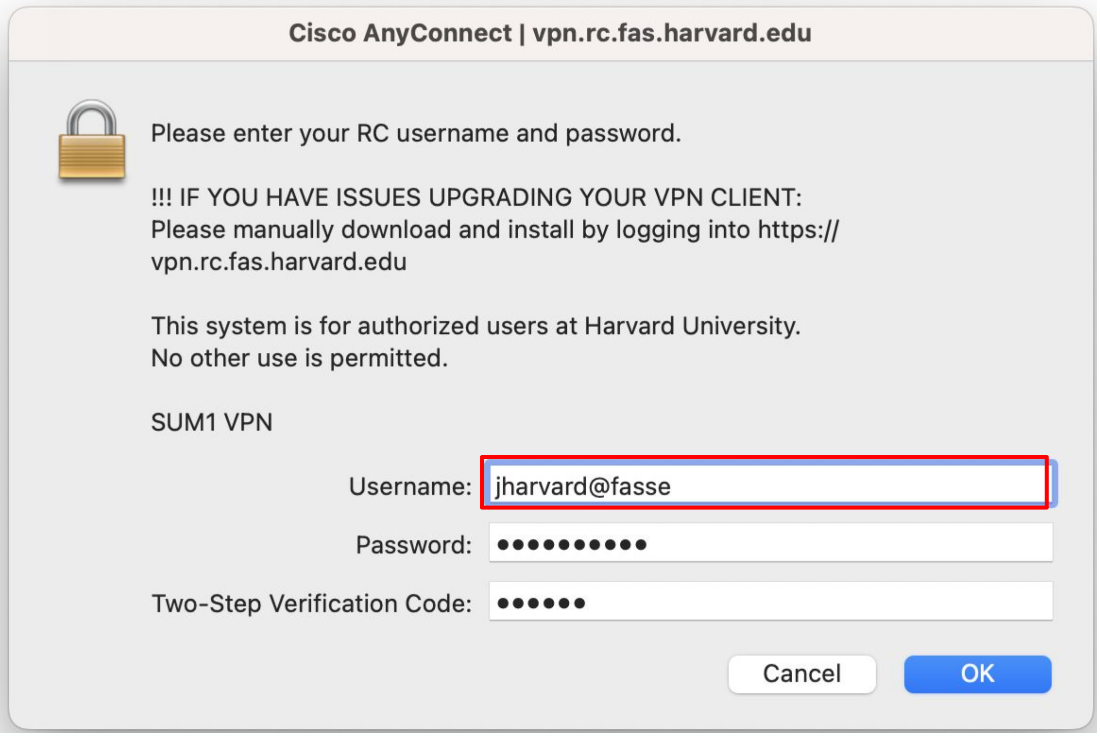
Please enter your RC username and password.

!!! IF YOU HAVE ISSUES UPGRADING YOUR VPN CLIENT:
Please manually download and install by logging into https://vpn.rc.fas.harvard.edu

This system is for authorized users at Harvard University.
No other use is permitted.

SUM1 VPN

Username: jharvard@fasrc
Password: ●●●●●●●●●●
Two-Step Verification Code: ●●●●●●



Cisco AnyConnect | vpn.rc.fas.harvard.edu

Please enter your RC username and password.

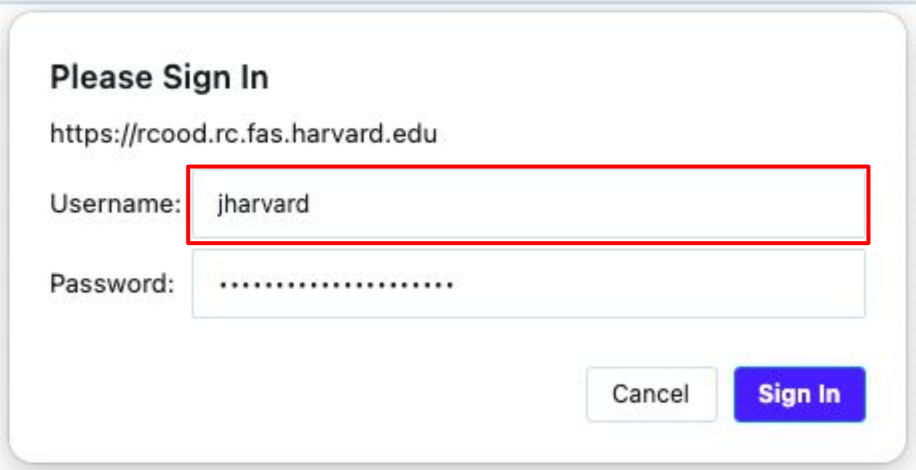
!!! IF YOU HAVE ISSUES UPGRADING YOUR VPN CLIENT:
Please manually download and install by logging into https://vpn.rc.fas.harvard.edu

This system is for authorized users at Harvard University.
No other use is permitted.

SUM1 VPN

Username: jharvard@fasse
Password: ●●●●●●●●●●
Two-Step Verification Code: ●●●●●●

Signing in to the OOD Dashboard



Please Sign In

https://rcood.rc.fas.harvard.edu

Username: jharvard

Password: ••••••••••••••••••••••

When you load the site

https://rcood.rc.fas.harvard.edu

or

https://fasseood.rc.fas.harvard.edu

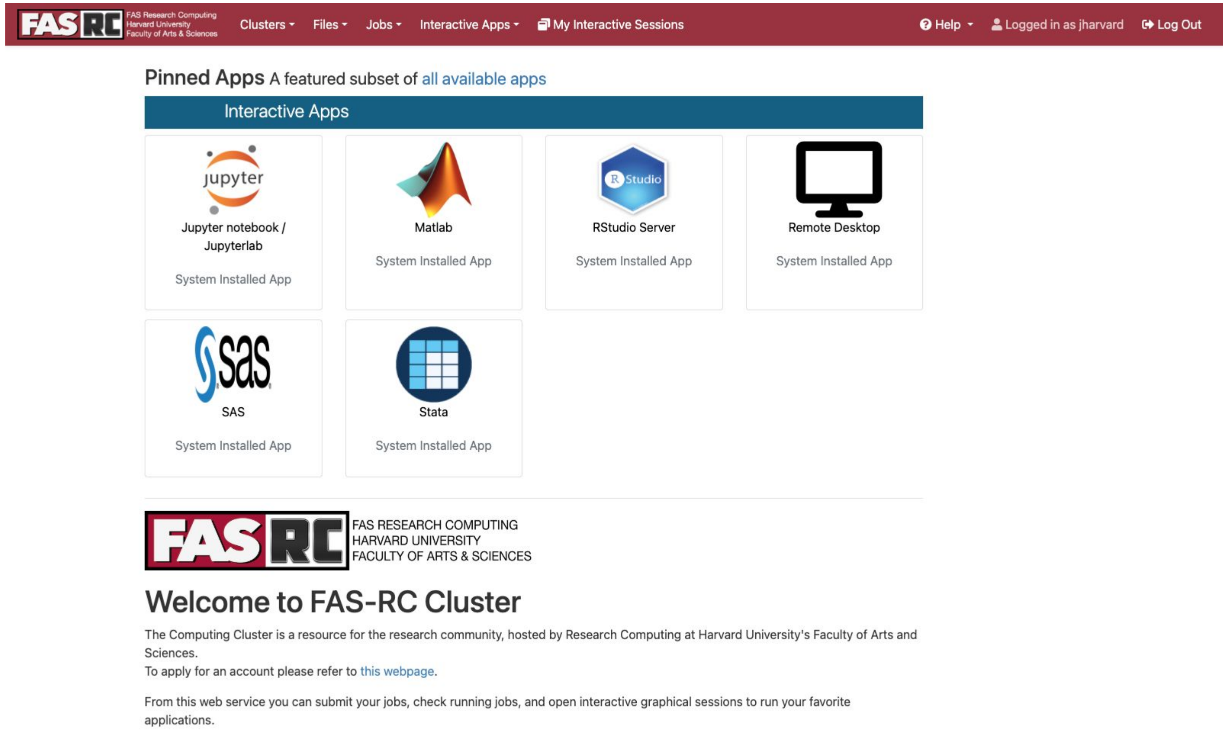
You will be prompted to log in.

Unlike when you log in to the VPN, on the website you need to use

ONLY your username, NOT username@cluster

OOD dashboard on Cannon and FASSE

Cannon



Pinned Apps A featured subset of all available apps

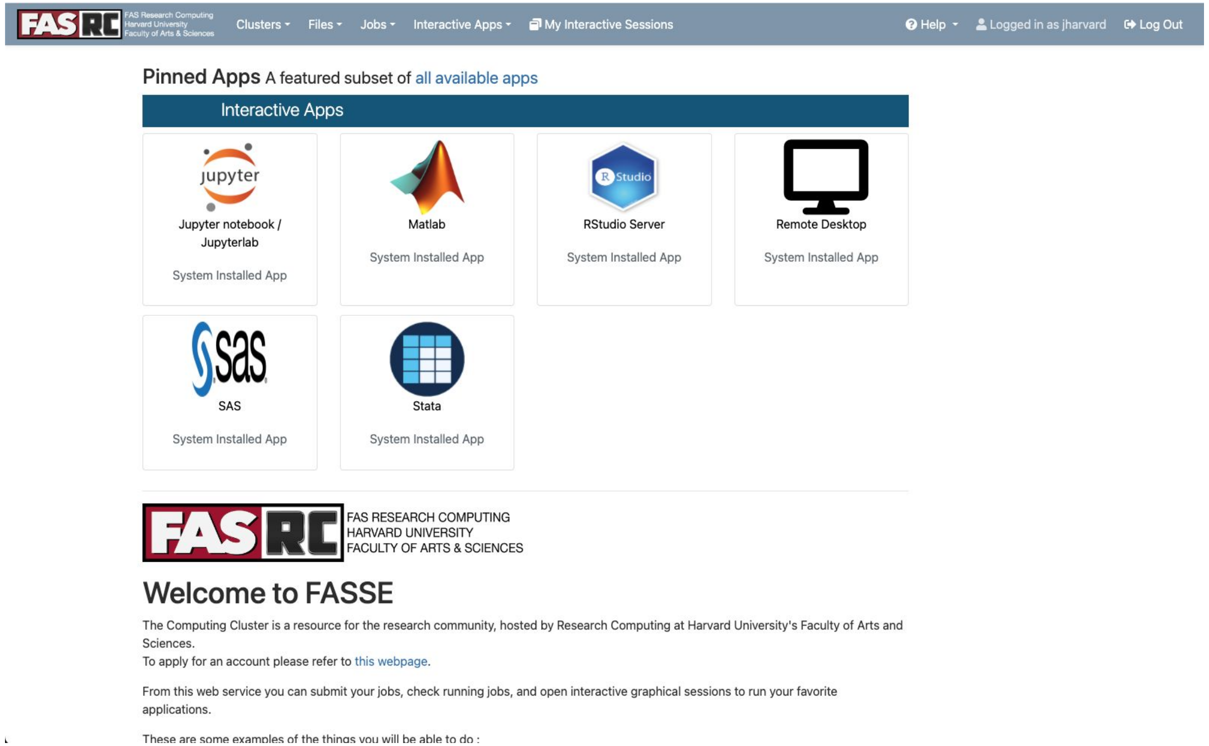
Welcome to FAS-RC Cluster

The Computing Cluster is a resource for the research community, hosted by Research Computing at Harvard University's Faculty of Arts and Sciences.
To apply for an account please refer to this webpage.

From this web service you can submit your jobs, check running jobs, and open interactive graphical sessions to run your favorite applications.

https://rcood.rc.fas.harvard.edu

FASSE



Pinned Apps A featured subset of all available apps

Welcome to FASSE

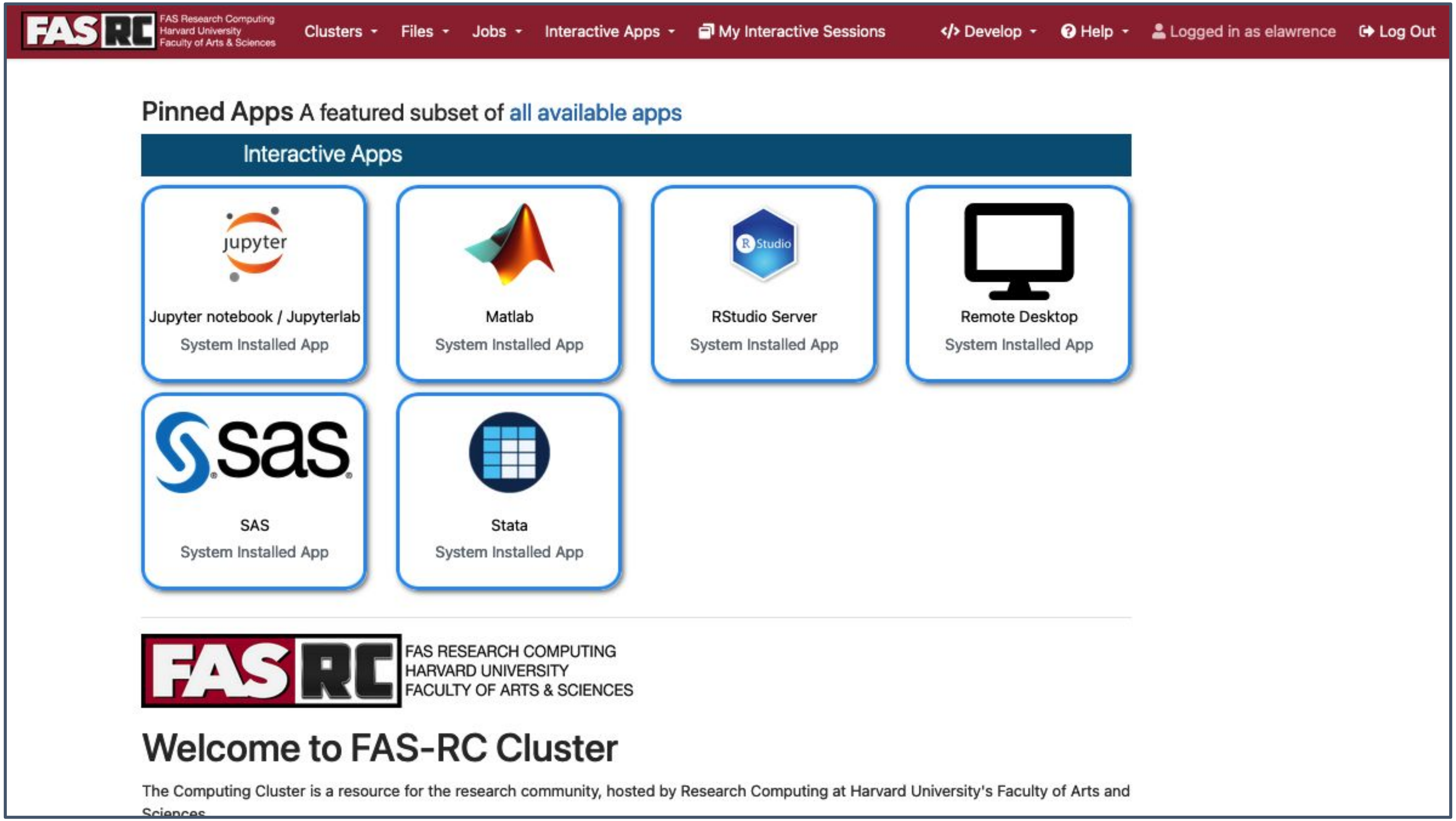
The Computing Cluster is a resource for the research community, hosted by Research Computing at Harvard University's Faculty of Arts and Sciences.
To apply for an account please refer to this webpage.

From this web service you can submit your jobs, check running jobs, and open interactive graphical sessions to run your favorite applications.

These are some examples of the things you will be able to do :

https://fasseood.rc.fas.harvard.edu

Launching an app from the Dashboard



Pinned Apps A featured subset of all available apps

Interactive Apps

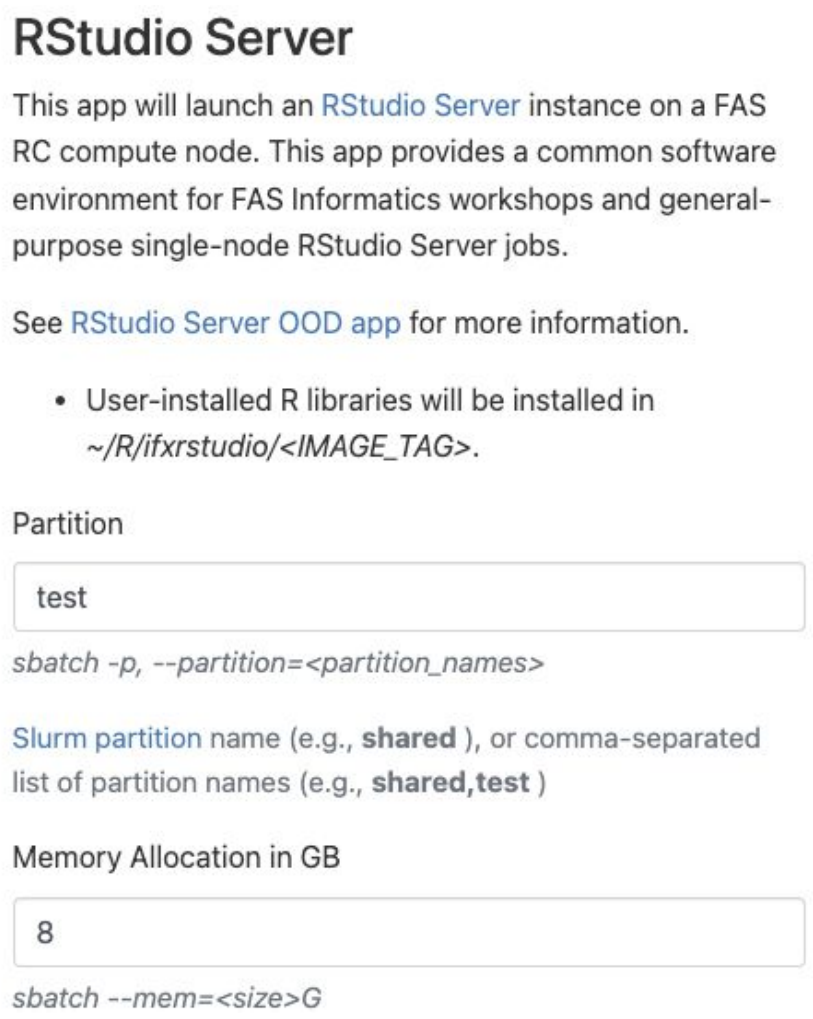
Jupyter notebook / Jupyterlab System Installed App	Matlab System Installed App	RStudio Server System Installed App	Remote Desktop System Installed App
SAS System Installed App	Stata System Installed App		

FAS RESEARCH COMPUTING
HARVARD UNIVERSITY
FACULTY OF ARTS & SCIENCES

Welcome to FAS-RC Cluster

The Computing Cluster is a resource for the research community, hosted by Research Computing at Harvard University's Faculty of Arts and Sciences

Beginning of a form



RStudio Server

This app will launch an RStudio Server instance on a FAS RC compute node. This app provides a common software environment for FAS Informatics workshops and general-purpose single-node RStudio Server jobs.

See RStudio Server OOD app for more information.

- User-installed R libraries will be installed in *~/R/ifxrstudio/<IMAGE_TAG>*.

Partition

test

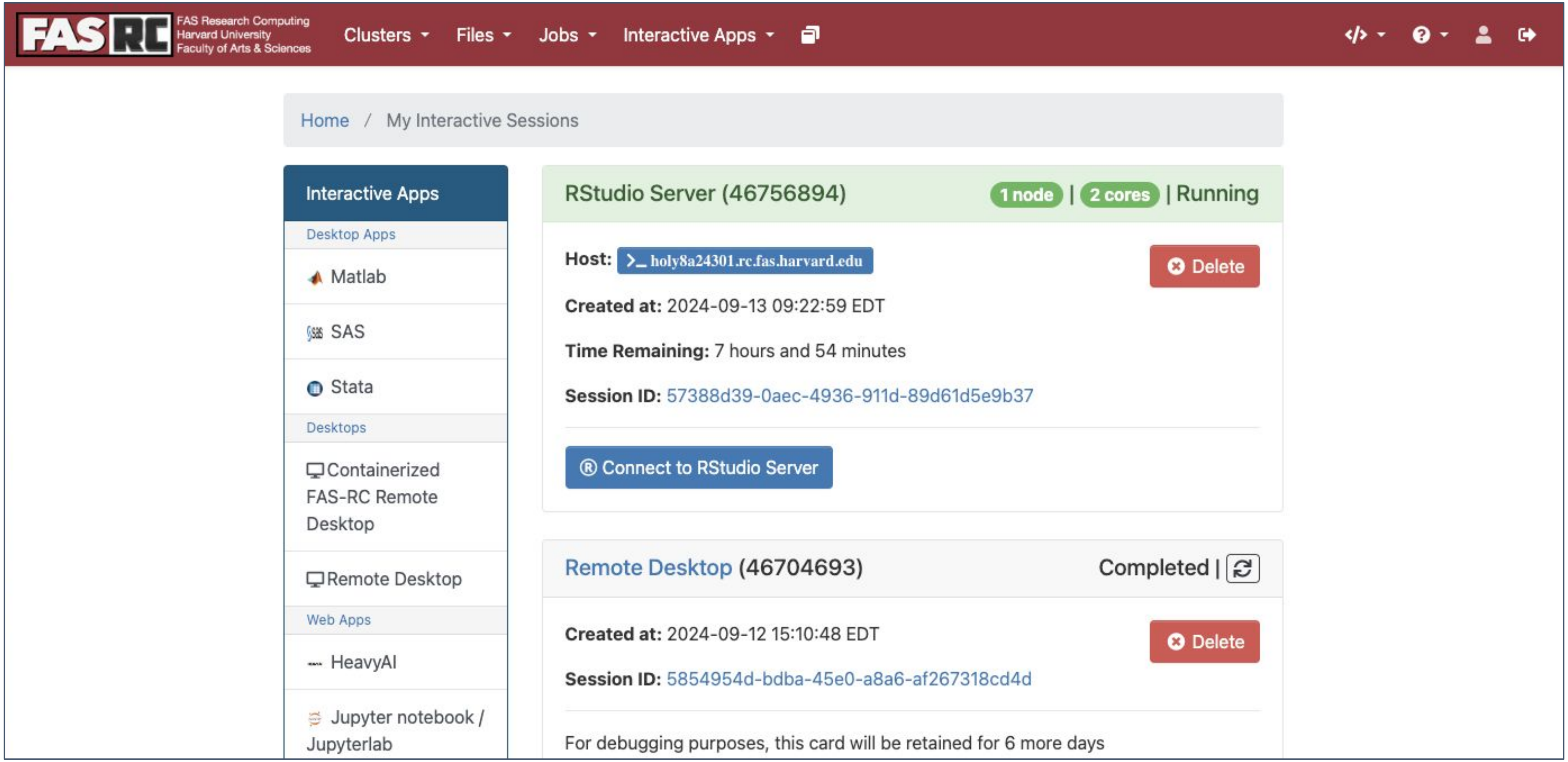
sbatch -p, --partition=<partition_names>

Slurm partition name (e.g., **shared**), or comma-separated list of partition names (e.g., **shared,test**)

Memory Allocation in GB

8

sbatch --mem=<size>G



RStudio Server (46756894) | 1 node | 2 cores | Running

Host: holy8a24301.rc.fas.harvard.edu

Created at: 2024-09-13 09:22:59 EDT

Time Remaining: 7 hours and 54 minutes

Session ID: 57388d39-0aec-4936-911d-89d61d5e9b37

Remote Desktop (46704693) Completed |

Created at: 2024-09-12 15:10:48 EDT

Session ID: 5854954d-bdba-45e0-a8a6-af267318cd4d

For debugging purposes, this card will be retained for 6 more days

Jupyter Notebook (1)

- You can create your own kernels, but some command line needed
- Note: a kernel is the same as a conda, python, or mamba environment

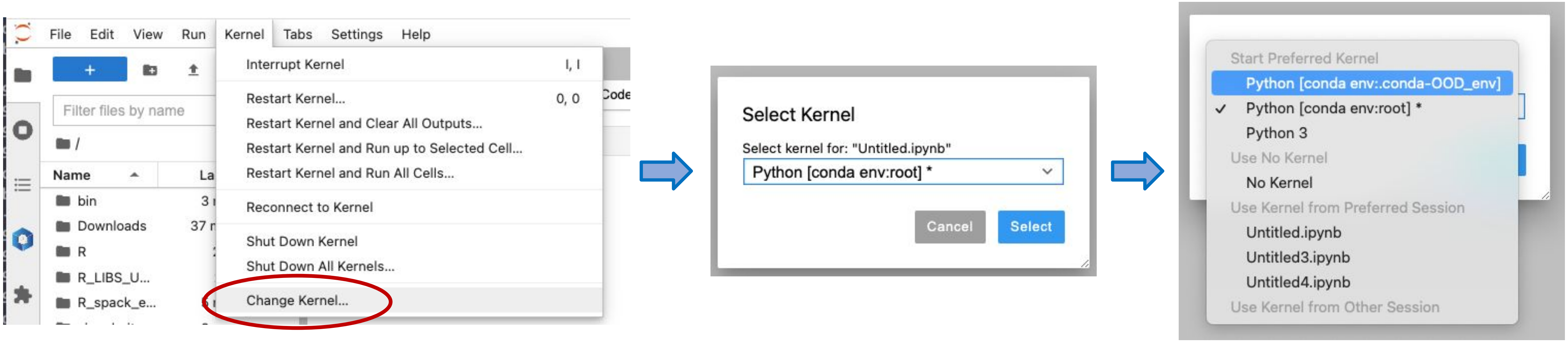
1. Open a terminal in the ”Remote Desktop” app
   ⇒Don’t create mamba environments inside Jupyter Notebook/Lab!
2. Create mamba environment and install package `ipykernel` and `nb_conda_kernels`
3.

```
[jharvard@holy7c02111 ~]$ module load python
[jharvard@holy7c02111 ~]$ mamba create –n OOD_env python=3.11 pip wheel numpy
[jharvard@holy7c02111 ~]$ mamba activate OOD_env
(OOD_env)[jharvard@holy7c02111 ~]$ mamba install ipykernel nb_conda_kernels
```

https://docs.rc.fas.harvard.edu/kb/python-package-installation/#Use_mamba_environment_in_Jupyter_Notebooks

Jupyter Notebook (2)

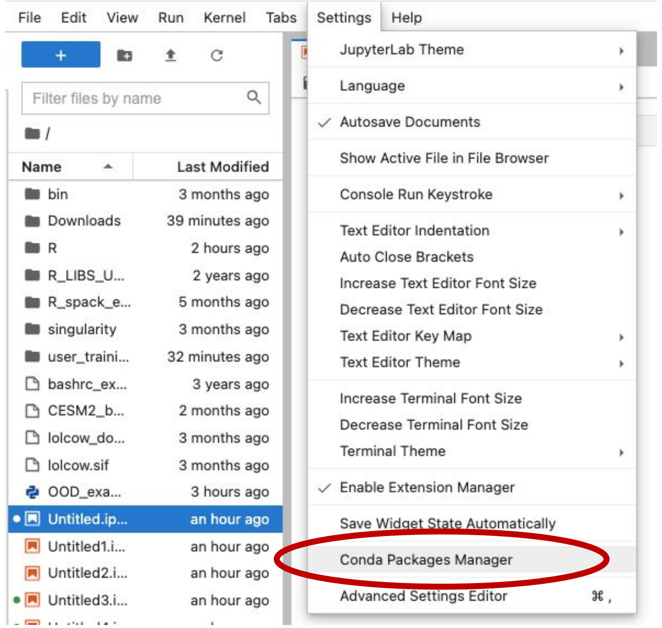
3. Launch **new** Jupyter Notebook session (existing session will not work!)
4. Select newly created mamba environment as the kernel
    a. Open a notebook
    b. On the top menu, click Kernel -> Select Kernel -> Click on OOD_env



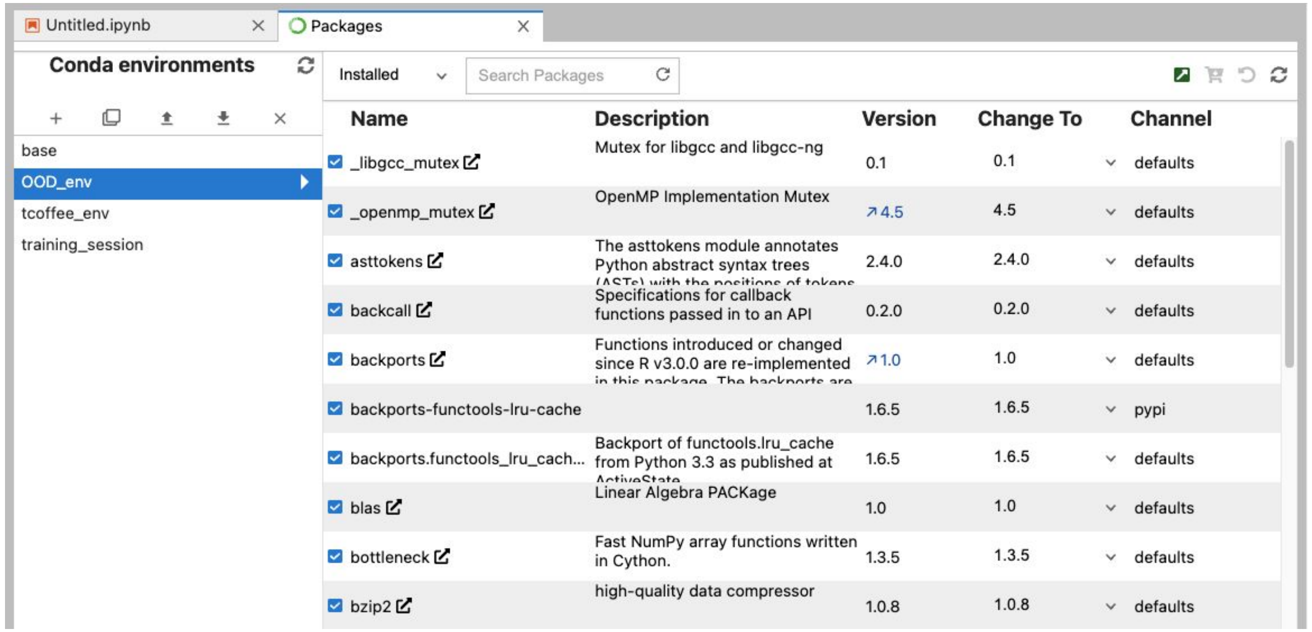
Select kernel for: "Untitled.ipynb"
Python [conda env:root] *

Jupyter Notebook (3)

5. Managing (install, uninstall, update) packages
    a. We recommend using the command line
    b. You can also use the conda package manager, which is the same thing as a mamba package manager: On the top menu, click Settings -> Conda Package Manager -> OOD_env



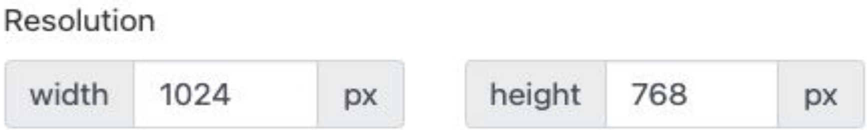
Name	Last Modified
bin	3 months ago
Downloads	39 minutes ago
R	2 hours ago
R_LIBS_U...	2 years ago
R_spack_e...	5 months ago
singularity	3 months ago
user_traini...	32 minutes ago
bashrc_ex...	3 years ago
CESM2_b...	2 months ago
lolcow_do...	3 months ago
lolcow.sif	3 months ago
OOD_exa...	3 hours ago
Untitled.ip...	an hour ago
Untitled1.i...	an hour ago
Untitled2.i...	an hour ago
Untitled3.i...	an hour ago



	Name	Description	Version	Change To	Channel
☑	_libgcc_mutex	Mutex for libgcc and libgcc-ng	0.1	0.1	defaults
☑	_openmp_mutex	OpenMP Implementation Mutex	↗4.5	4.5	defaults
☑	asttokens	The asttokens module annotates Python abstract syntax trees	2.4.0	2.4.0	defaults
☑	backcall	Specifications for callback functions passed in to an API	0.2.0	0.2.0	defaults
☑	backports	Functions introduced or changed since R v3.0.0 are re-implemented	↗1.0	1.0	defaults
☑	backports-functools-lru-cache		1.6.5	1.6.5	pypi
☑	backports.functools_lru_cach...	Backport of functools.lru_cache from Python 3.3 as published at	1.6.5	1.6.5	defaults
☑	blas	Linear Algebra PACKage	1.0	1.0	defaults
☑	bottleneck	Fast NumPy array functions written in Cython.	1.3.5	1.3.5	defaults
☑	bzip2	high-quality data compressor	1.0.8	1.0.8	defaults

Remote Desktop

- Not as necessary as it used to be
- **For running long Jupyter Notebook sessions**
- Terminal
- Can also be used to open multiple applications in a single window
- Choose the defaults for resolution



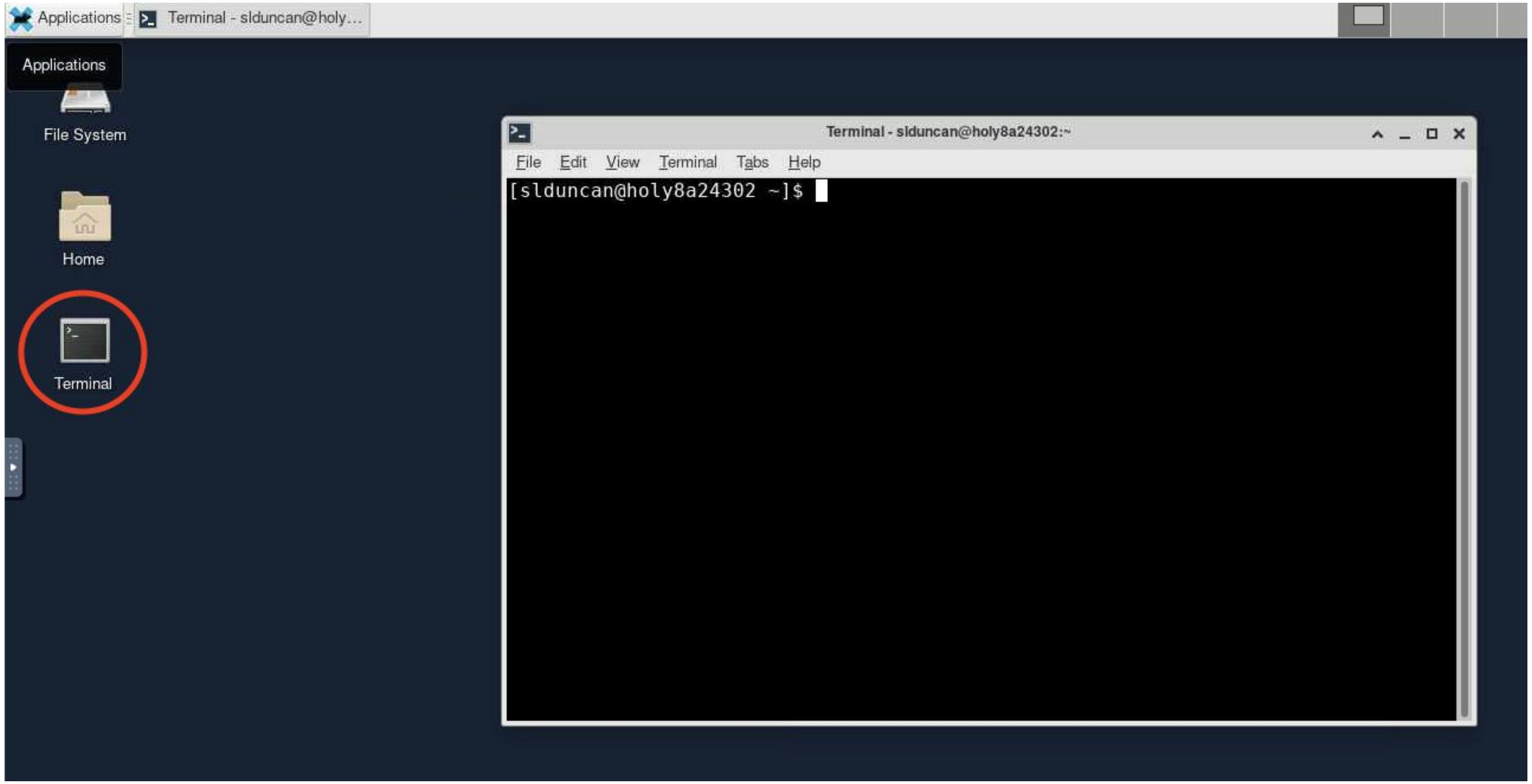
Resolution

width 1024 px

height 768 px

Starting a terminal

- Double click on the icon that looks like a computer screen
- From there you can type in any commands you want to



```
[slduncan@holy8a24302 ~]$ 
```

Running Jupyter Notebook in Remote Desktop

- Very important: Jupyter Notebook will not continue to run if you close the Jupyter notebook page! The cell that is running will lose the data and output files will not be written
    1. Solution: run Remote Desktop app and launch Jupyter Notebook from within Remote Desktop
    2. Documentation: https://docs.rc.fas.harvard.edu/kb/ood-remote-desktop-how-to-open-software/#Jupyter_Notebook

```
#Jupyter Notebook
[jharvard@holy7c02111 ~]$ module load python
[jharvard@holy7c02111 ~]$ mamba activate OOD_env
[jharvard@holy7c02111 ~]$ jupyter notebook
```

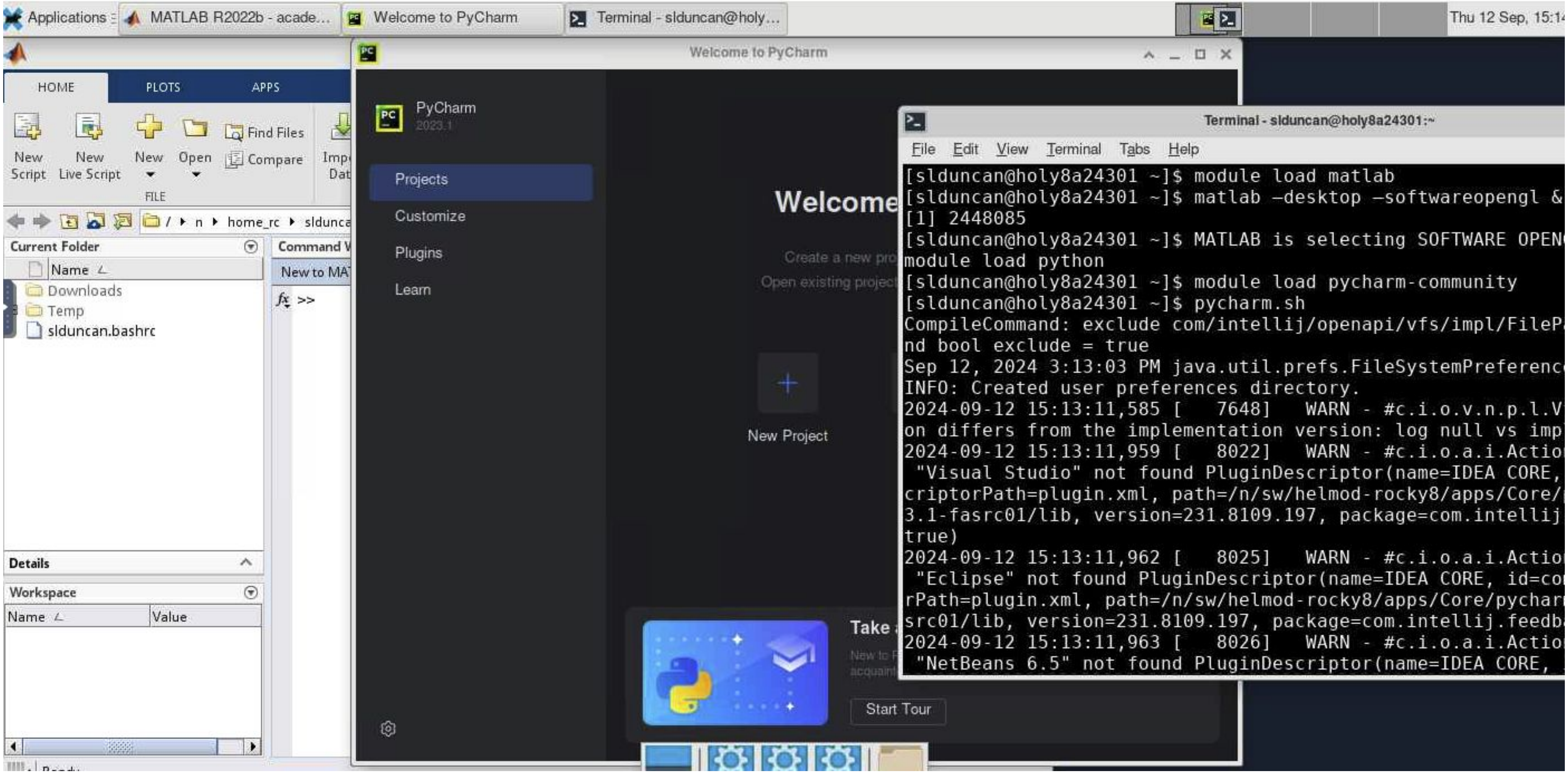
Opening Multiple Applications in Remote Desktop

Documentation: https://docs.rc.fas.harvard.edu/kb/ood-remote-desktop-how-to-open-software/

- It can be used to launch most GUI applications
  1. Load module
  2. Set environmental variables (if needed)
  3. Launch software
- You can have multiple applications open

```
# Matlab
[jharvard@holy7c02111 ~]$ module load matlab
[jharvard@holy7c02111 ~]$ matlab -desktop -softwareopengl &

#PyCharm
[jharvard@holy7c02111 ~]$ module load python
[jharvard@holy7c02111 ~]$ module load pycharm-community
[jharvard@holy7c02111 ~]$ pycharm.sh
```



```
[slduncan@holy8a24301 ~]$ module load matlab
[slduncan@holy8a24301 ~]$ matlab -desktop -softwareopengl &
[1] 2448085
[slduncan@holy8a24301 ~]$ MATLAB is selecting SOFTWARE OPEN
module load python
[slduncan@holy8a24301 ~]$ module load pycharm-community
[slduncan@holy8a24301 ~]$ pycharm.sh
CompileCommand: exclude com/intellij/openapi/vfs/impl/FileP
nd bool exclude = true
Sep 12, 2024 3:13:03 PM java.util.prefs.FileSystemPreferenc
INFO: Created user preferences directory.
2024-09-12 15:13:11,585 [   7648]   WARN - #c.i.o.v.n.p.l.V
on differs from the implementation version: log null vs imp
2024-09-12 15:13:11,959 [   8022]   WARN - #c.i.o.a.i.Actio
 "Visual Studio" not found PluginDescriptor(name=IDEA CORE,
criptorPath=plugin.xml, path=/n/sw/helmod-rocky8/apps/Core/
3.1-fasrc01/lib, version=231.8109.197, package=com.intellij
true)
2024-09-12 15:13:11,962 [   8025]   WARN - #c.i.o.a.i.Actio
 "Eclipse" not found PluginDescriptor(name=IDEA CORE, id=co
rPath=plugin.xml, path=/n/sw/helmod-rocky8/apps/Core/pychar
src01/lib, version=231.8109.197, package=com.intellij.feedb
2024-09-12 15:13:11,963 [   8026]   WARN - #c.i.o.a.i.Actio
 "NetBeans 6.5" not found PluginDescriptor(name=IDEA CORE,
```

Inactivity lock out

- It may lock out due to inactivity
- Use your FASRC password to unlock

Filling out a form to launch an app

- Request the resources that you need
  (If you don't know for a first trial run, use similar resources as your laptop/desktop)
  - Partition (Name): depends on Cannon vs FASSE
  - Memory (RAM): amount of memory in GB
  - Number of cores: recommended at least 2
  - Number of GPUs: if >= 1, make sure you **select** a gpu partition
  - Allocated time: time you would like your session to run.

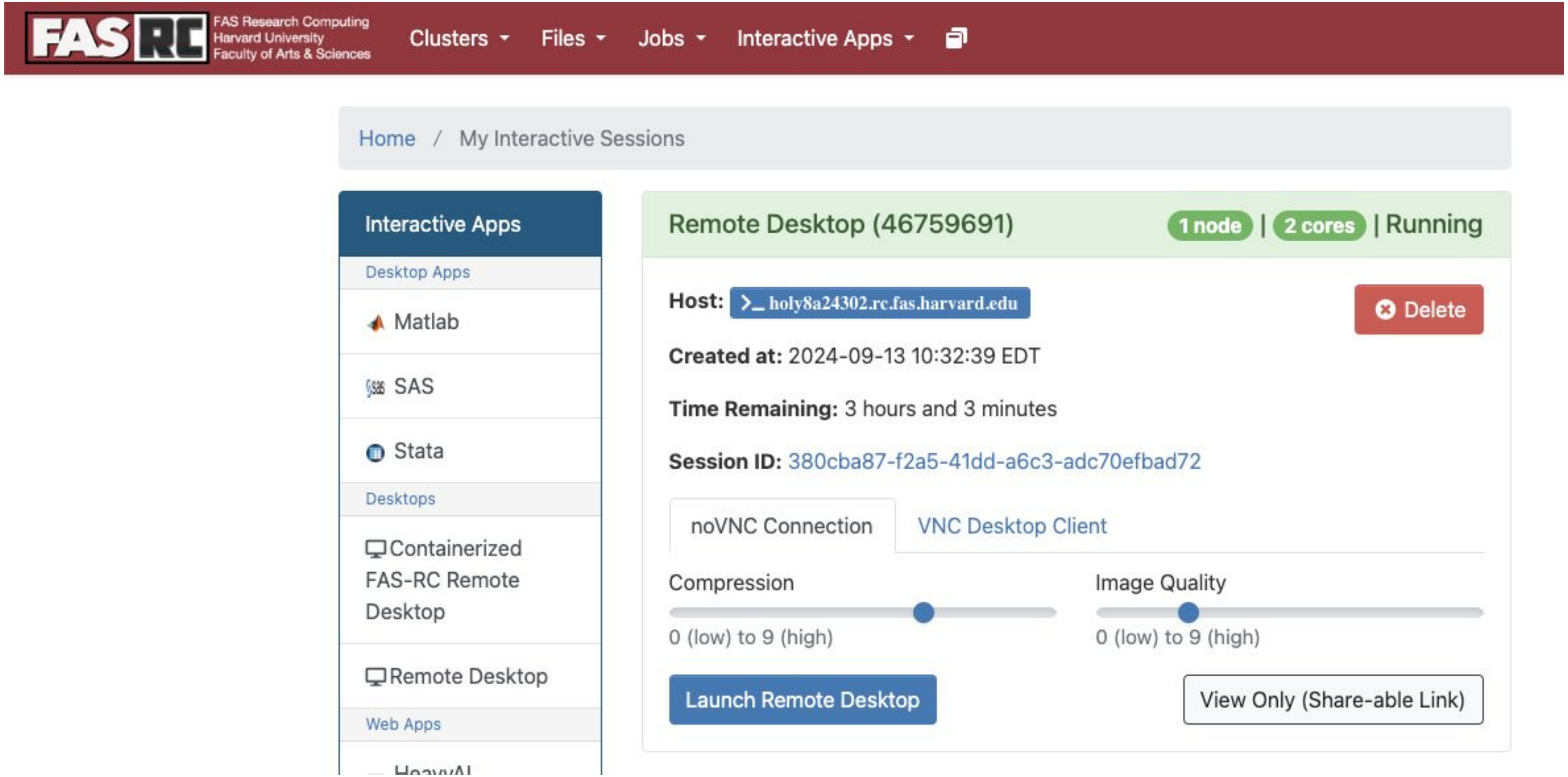
  the minimum and/or maximum values of each field depends on the selected partition

  - Email for status notification: to know when job starts, ends
  - Reservation: if you have a special reservation (this requires approval from FASRC)
  - Account: use this if you have more than one PI_lab affiliation

Menu bar

- My Interactive Sessions
- Files
- Jobs
- Interactive apps

My Interactive Sessions



Remote Desktop (46759691) 1 node | 2 cores | Running

Host: holy8a24302.rc.fas.harvard.edu

Created at: 2024-09-13 10:32:39 EDT

Time Remaining: 3 hours and 3 minutes

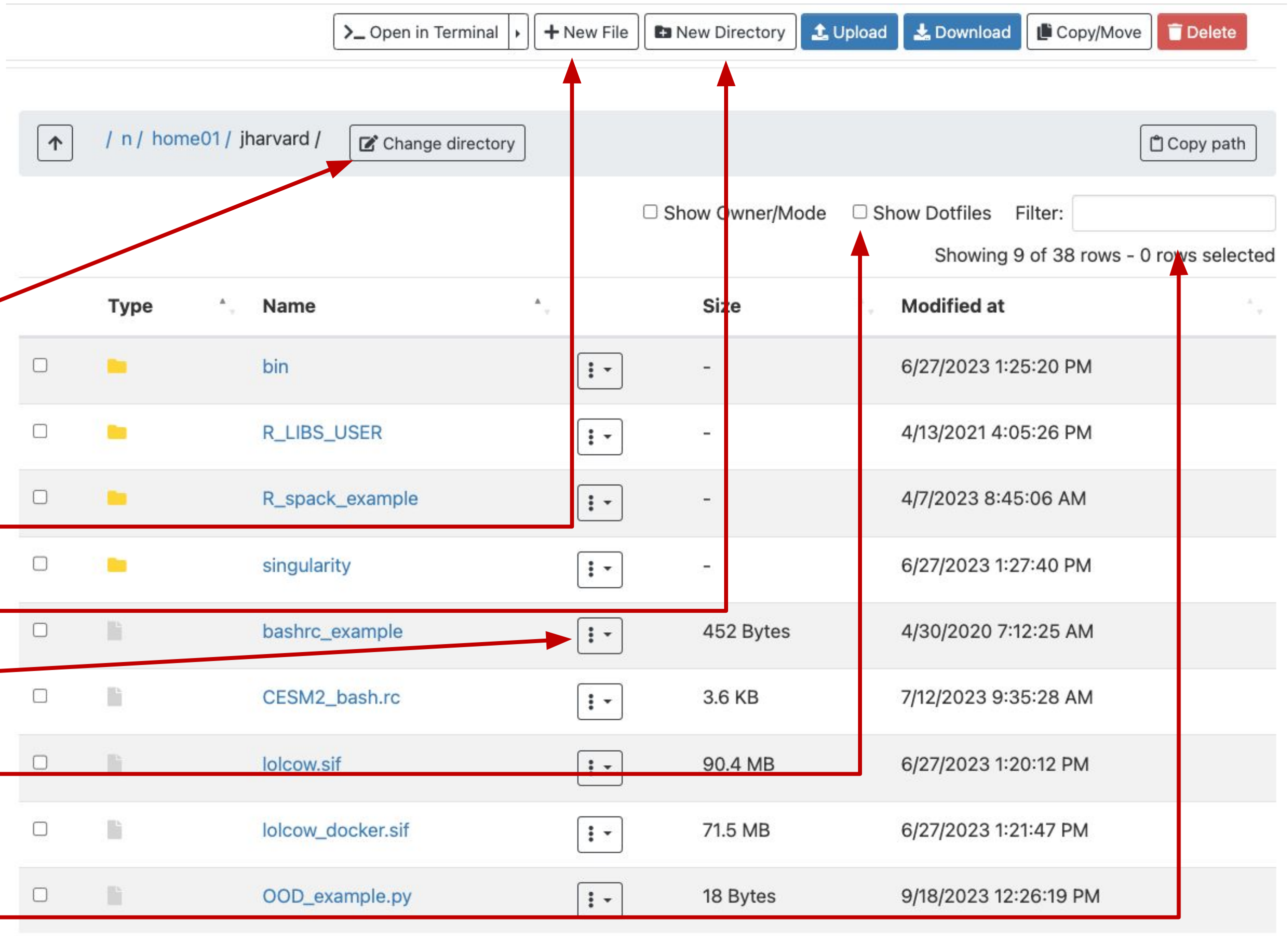
Session ID: 380cba87-f2a5-41dd-a6c3-adc70efbad72

Compression
0 (low) to 9 (high)

Image Quality
0 (low) to 9 (high)

Files tab

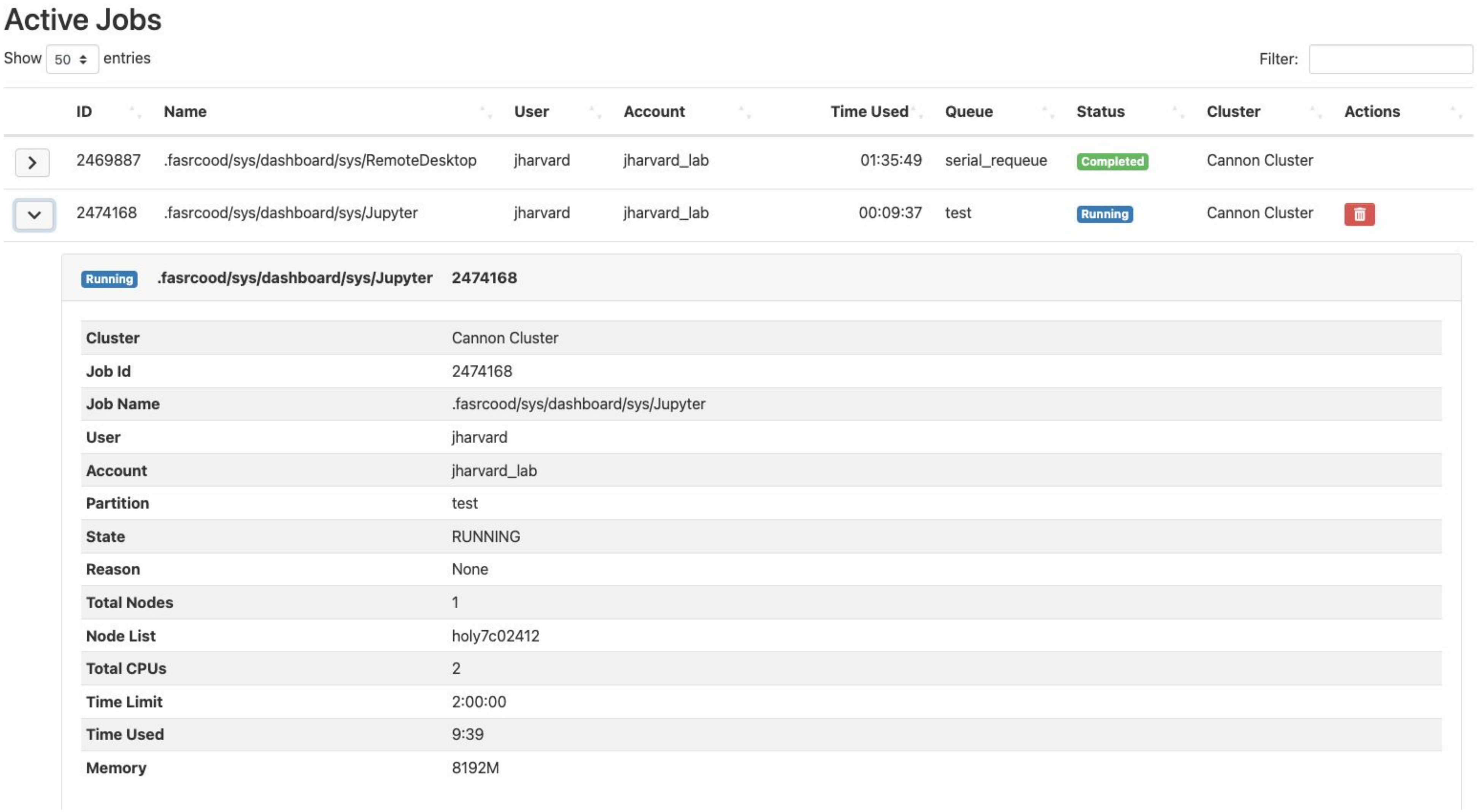
- Default options: home directory and holyscratch
- Click on “Change directory” to go to a lab share at /n/holylabs/LABS
- Create new file
- Create new directory (i.e., folder)
- Click on three dots for options
- Check “Show Dotfiles” to see hidden files
- Filter to find files or directories in current directory



Showing 9 of 38 rows - 0 rows selected

	Type	Name		Size	Modified at
☐	📁	bin	⋮ ▾	-	6/27/2023 1:25:20 PM
☐	📁	R_LIBS_USER	⋮ ▾	-	4/13/2021 4:05:26 PM
☐	📁	R_spack_example	⋮ ▾	-	4/7/2023 8:45:06 AM
☐	📁	singularity	⋮ ▾	-	6/27/2023 1:27:40 PM
☐	📄	bashrc_example	⋮ ▾	452 Bytes	4/30/2020 7:12:25 AM
☐	📄	CESM2_bash.rc	⋮ ▾	3.6 KB	7/12/2023 9:35:28 AM
☐	📄	lolcow.sif	⋮ ▾	90.4 MB	6/27/2023 1:20:12 PM
☐	📄	lolcow_docker.sif	⋮ ▾	71.5 MB	6/27/2023 1:21:47 PM
☐	📄	OOD_example.py	⋮ ▾	18 Bytes	9/18/2023 12:26:19 PM

Jobs tab (1)



Active Jobs

Show 50 ⇕ entries

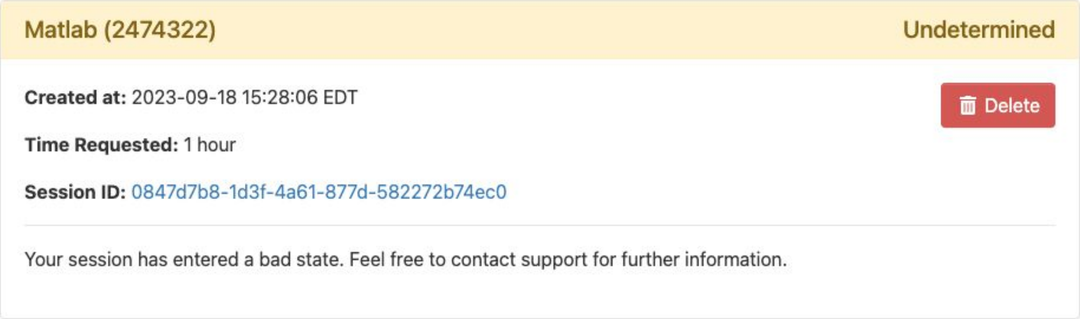
Filter:

	ID	Name	User	Account	Time Used	Queue	Status	Cluster	Actions
>	2469887	.fasrcood/sys/dashboard/sys/RemoteDesktop	jharvard	jharvard_lab	01:35:49	serial_requeue	Completed	Cannon Cluster	
⌄	2474168	.fasrcood/sys/dashboard/sys/Jupyter	jharvard	jharvard_lab	00:09:37	test	Running	Cannon Cluster	🗑

Running **.fasrcood/sys/dashboard/sys/Jupyter** **2474168**

Cluster	Cannon Cluster
Job Id	2474168
Job Name	.fasrcood/sys/dashboard/sys/Jupyter
User	jharvard
Account	jharvard_lab
Partition	test
State	RUNNING
Reason	None
Total Nodes	1
Node List	holy7c02412
Total CPUs	2
Time Limit	2:00:00
Time Used	9:39
Memory	8192M

Jobs tab (2)



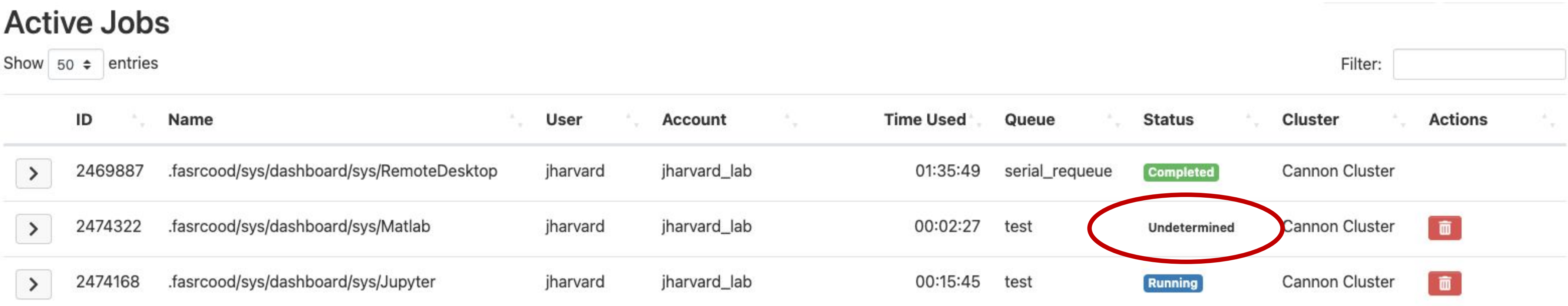
Matlab (2474322) Undetermined

Created at: 2023-09-18 15:28:06 EDT

Time Requested: 1 hour

Session ID: 0847d7b8-1d3f-4a61-877d-582272b74ec0

Your session has entered a bad state. Feel free to contact support for further information.



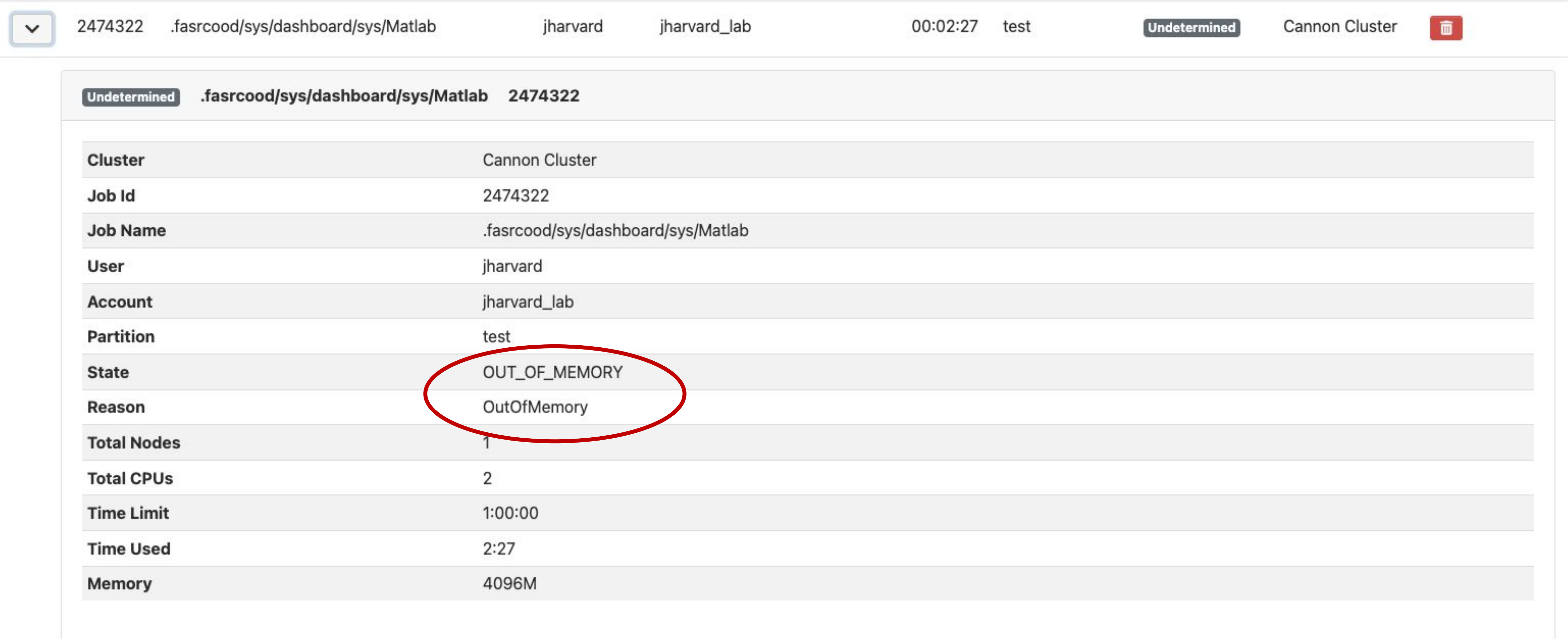
Active Jobs

Show 50 entries

Filter:

	ID	Name	User	Account	Time Used	Queue	Status	Cluster	Actions
>	2469887	.fasrcood/sys/dashboard/sys/RemoteDesktop	jharvard	jharvard_lab	01:35:49	serial_requeue	Completed	Cannon Cluster	
>	2474322	.fasrcood/sys/dashboard/sys/Matlab	jharvard	jharvard_lab	00:02:27	test	Undetermined	Cannon Cluster	🗑
>	2474168	.fasrcood/sys/dashboard/sys/Jupyter	jharvard	jharvard_lab	00:15:45	test	Running	Cannon Cluster	🗑

Jobs tab (3)



2474322	.fasrcood/sys/dashboard/sys/Matlab	jharvard	jharvard_lab	00:02:27	test	Undetermined	Cannon Cluster	🗑

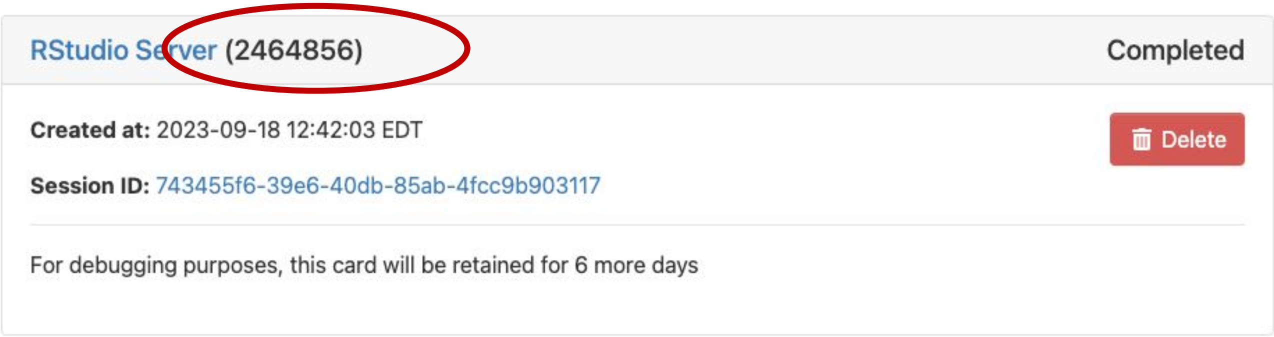
Undetermined **.fasrcood/sys/dashboard/sys/Matlab** **2474322**

Cluster	Cannon Cluster
Job Id	2474322
Job Name	.fasrcood/sys/dashboard/sys/Matlab
User	jharvard
Account	jharvard_lab
Partition	test
State	OUT_OF_MEMORY
Reason	OutOfMemory
Total Nodes	1
Total CPUs	2
Time Limit	1:00:00
Time Used	2:27
Memory	4096M

Job tab (4)

If job no longer appears on “Active Jobs”, check job status from command line with slurm job ID

slurm job ID



RStudio Server (2464856) Completed

Created at: 2023-09-18 12:42:03 EDT

Session ID: 743455f6-39e6-40db-85ab-4fcc9b903117

For debugging purposes, this card will be retained for 6 more days

```
[jharvard@boslogin01 ~]$ sacct -j 2464856
JobID           JobName  Partition    Account  AllocCPUS      State ExitCode
------------ ---------- ---------- ---------- ---------- ---------- --------
2464856      .fasrcood+       test jharvard_+          2    TIMEOUT      0:0
2464856.bat+      batch            jharvard_+          2  CANCELLED     0:15
2464856.ext+     extern            jharvard_+          2  COMPLETED      0:0

[jharvard@holy7c02111 ~]$ sacct -j 2471535
JobID           JobName  Partition    Account  AllocCPUS      State ExitCode
------------ ---------- ---------- ---------- ---------- ---------- --------
2471535      .fasrcood+       test jharvard_+          2 OUT_OF_ME+    0:125
2471535.bat+      batch            jharvard_+          2 OUT_OF_ME+    0:125
2471535.ext+     extern            jharvard_+          2  COMPLETED      0:0
```

Closing running OOD windows/tabs

- In most OOD apps, you can close the browser tab while the code is running, and the code will continue to run on the background
- Jupyter Notebook will not! The cell that is running will lose the data and output files will not be written
  - Solution: run Remote Desktop app and launch Jupyter Notebook from within Remote Desktop
  - Documentation: https://docs.rc.fas.harvard.edu/kb/ood-remote-desktop-how-to-open-software/#Jupyter_Notebook
- Because closing tabs does not end the application, it is important to cancel your job when you are done using it. Otherwise it will be charged to your lab's fairshare.

FASSE proxy

Documentation: https://docs.rc.fas.harvard.edu/kb/proxy-settings/

- You may need to set FASSE proxy on
  - RStudio server if you are unable to reach cran and download R packages
  - Stata if you are unable to load libraries via http
  - Firefox (web browsing)
  - Jupyter Notebook
  - Access Github
  - (Basically, anything outside of FASSE)

Quickstart Guides for using the FASRC Clusters

- Cannon Quickstart Guide
  - https://docs.rc.fas.harvard.edu/kb/iqss-cannon-quickstart-guide
- FASSE Quickstart Guide
  - https://docs.rc.fas.harvard.edu/kb/iqss-fasse-quickstart-guide

- Quickstart guides have more than just information on OOD
  - how to do text based access
  - office hours, training, tickets

FASRC documentation

- FASRC docs: https://docs.rc.fas.harvard.edu/
- GitHub User_codes: https://github.com/fasrc/User_Codes/
- Getting help
  - Office hours: https://www.rc.fas.harvard.edu/training/office-hours/
  - Ticket
    - Email: rchelp@rc.fas.harvard.edu to open a ticket with us

Upcoming trainings

Training calendar: https://www.rc.fas.harvard.edu/upcoming-training/

Containers on the FASRC clusters

This training covers using and managing software/workflow environments using containers. We will touch on the basic differences and how they can be used to help you build/use complex software environments on the cluster.

This training is for those who already have a FASRC account, and want to learn tools and services.

Upcoming trainings

Training calendar: https://www.rc.fas.harvard.edu/upcoming-training/

Getting started on the FASRC clusters with command line interface (CLI)

- Requirement: working FASRC account with cluster access
- Audience
  - Users familiar with command-line interface
  - New to Cannon and FASSE, but familiar with HPC systems
- Content
  - Submit interactive job with `salloc`
  - Submit batch job with `sbatch`
  - Monitor jobs
  - Cluster software overview (modules, spack)

Upcoming trainings

Training calendar: https://www.rc.fas.harvard.edu/upcoming-training/

Advanced Cluster Usage

This training would focus on users who are familiar with the command line interface and would like to improve job submission and management/monitoring.

Objectives:

- Submit interactive and batch jobs
- Request resources appropriate to job requirements
- Monitoring jobs, priority, when jobs will run
- Fairshare
- Scratch vs. home directory performance

Training session evaluation

Please, fill out our training session evaluation. Your feedback is essential for us to improve our trainings!

https://tinyurl.com/FASRC-training

Where to access the slides

https://docs.rc.fas.harvard.edu/kb/training-materials/

Thank you :)

FAS Research Computing