Getting Started on the FASRC clusters with Open OnDemand

Learning objectives

- What is Open OnDemand (OOD)?
- How to access OOD?
- Understanding the form to launch apps
- RStudio Server
- Jupyter Notebook
  - Create conda environment (i.e., jupyter kernel)
- Remote Desktop
- FASSE proxy
- Files tab
- Jobs tab

What is Open OnDemand (OOD)?

- Open-source web portal to access clusters
- Web-based, no software needs be installed on your local laptop/desktop (except for a modern browser like Google Chrome, Mozilla Firefox)
- Easy to learn and simple to use
- Very similar to desktop applications
- The easiest way to run GUI applications remotely on a cluster
- Safari is not recommended for OOD

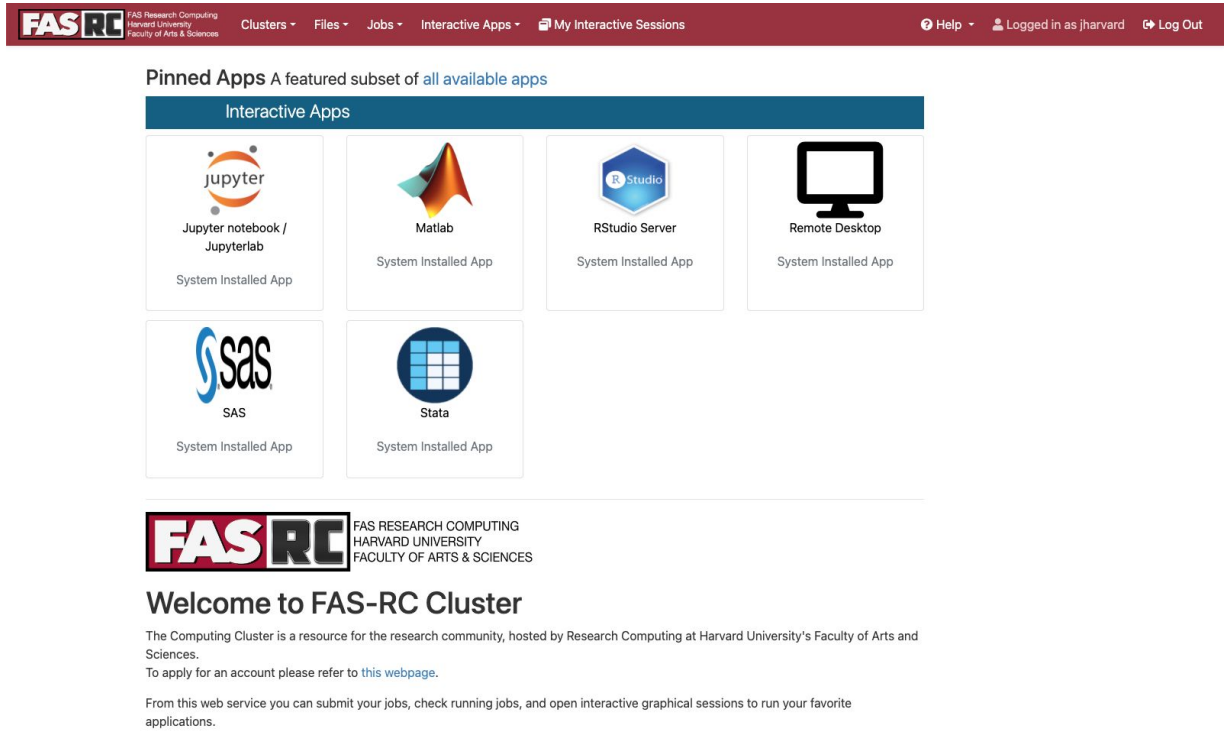
OPEN OnDemand

How to access OOD on FASRC Clusters

- FASRC clusters
  - Cannon & how to get an account - IQSS Cannon Quickstart Guide – FASRC DOCS
  - FASSE & how to get an account - IQSS FASSE Quickstart Guide – FASRC DOCS
- Accessing OOD from Cannon
  - Connect to FASRC VPN - Virtual Desktop (VDI) through Open OnDemand – FASRC DOCS
  - Then go to https://rcood.rc.fas.harvard.edu
- Accessing OOD from FASSE
  - Connect to FASSE VPN - FASSE VDI Apps – FASRC DOCS
  - Then go to https://fasseood.rc.fas.harvard.edu

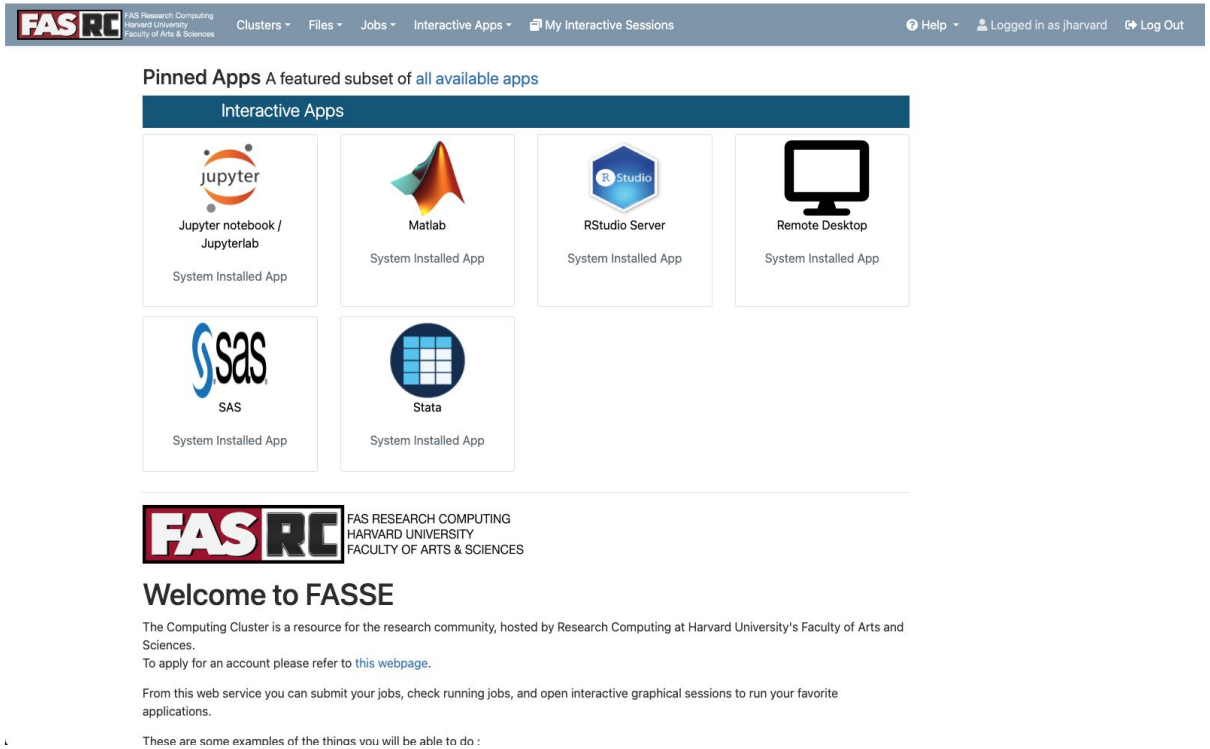
OOD dashboard on Cannon and FASSE

Cannon



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

FASSE



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

Filling 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

RStudio Server vs. RStudio Desktop

RStudio Server

- Go-to RStudio application with easy-to-install packages (pre-compiled)
- Cannot set `R_LIBS_USER` (location where packages are saved)
- `R_LIBS_USER` is set to `~/R/ifxrstudio/\<IMAGE_TAG\>`
- Cannot use `module load`
- Cannot use slurm commands (e.g. `sbatch`)

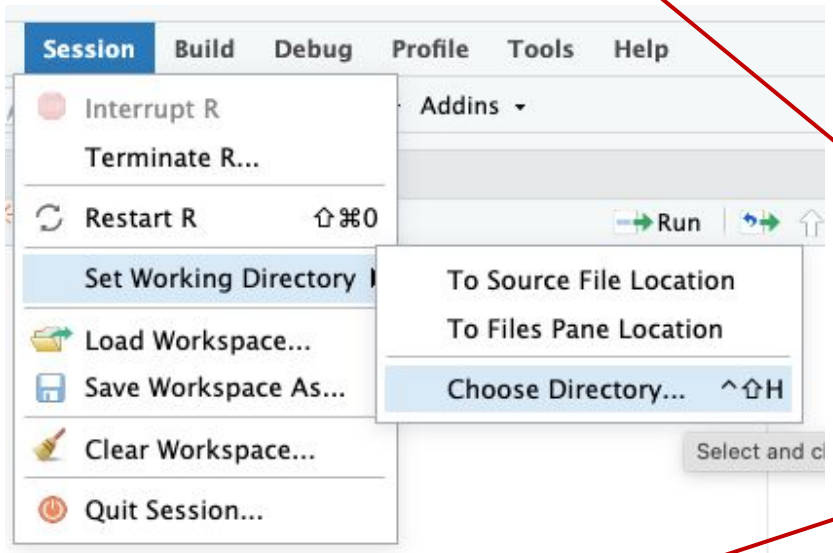
R via Remote Desktop

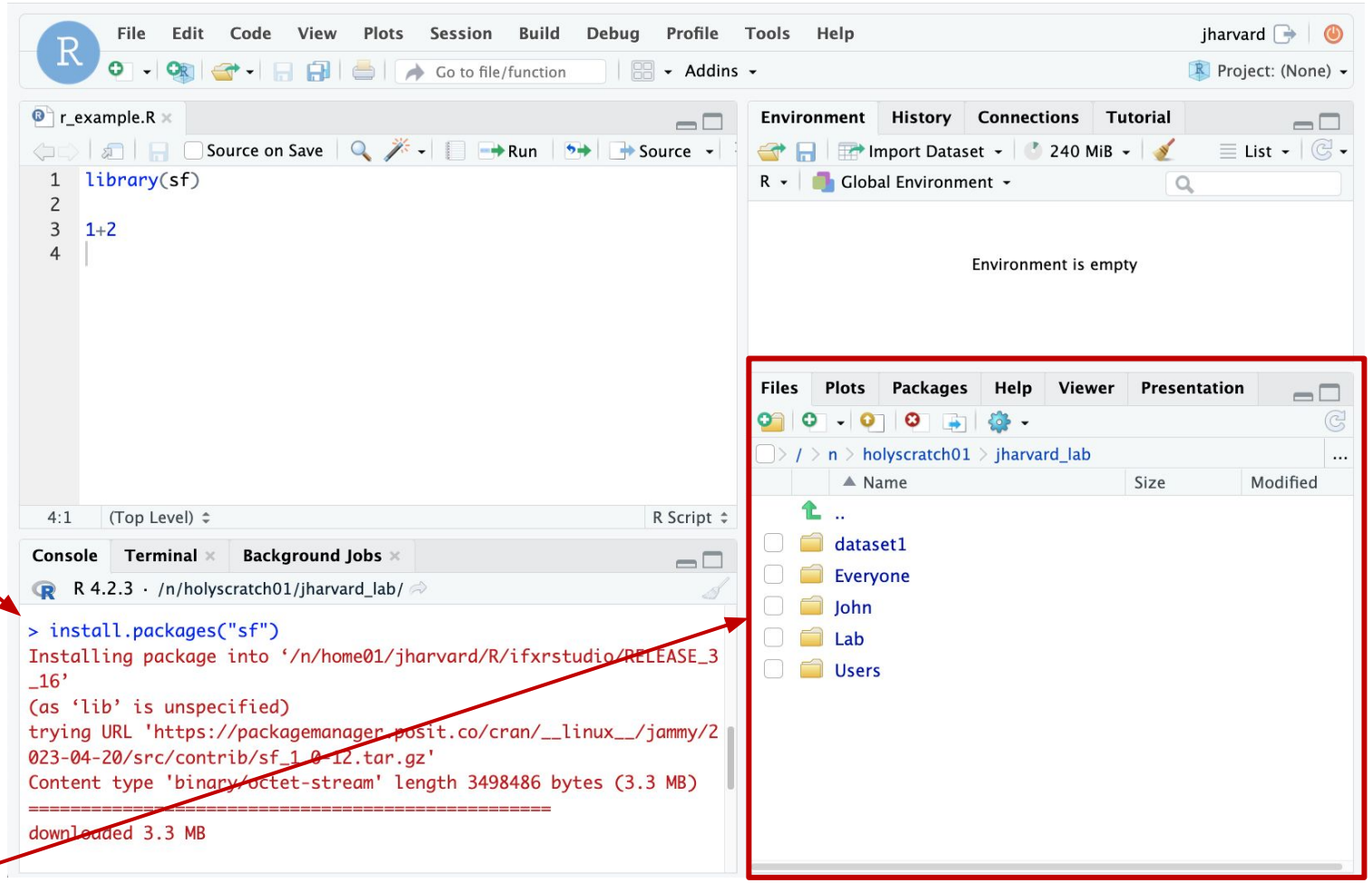
- Highly customized environment
- Can set `R_LIBS_USER`
- Can use `module load` → you can set specific compilers (e.g. `openmpi`, `gcc`)
- Can use slurm commands (e.g. `sbatch`)

See RStudio Server vs. RStudio Desktop OOD apps – FASRC DOCS

RStudio Server

- Package install
- Change directory





- Open file

Jupyter Notebook (1)

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

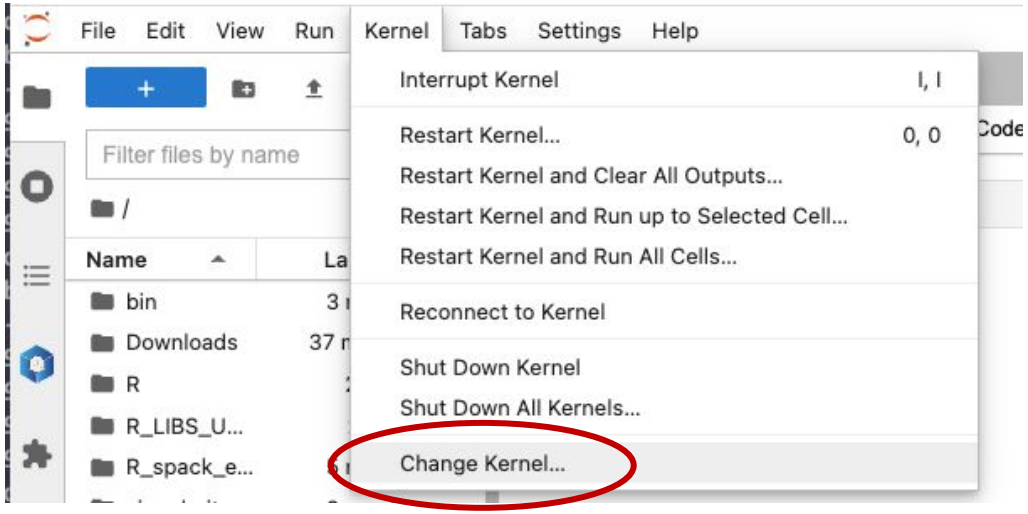
1. Launch ”Remote Desktop” app
   ⇒Don’t create conda environments inside Jupyter Notebook/Lab!!
2. Create conda environment and install package `ipykernel`

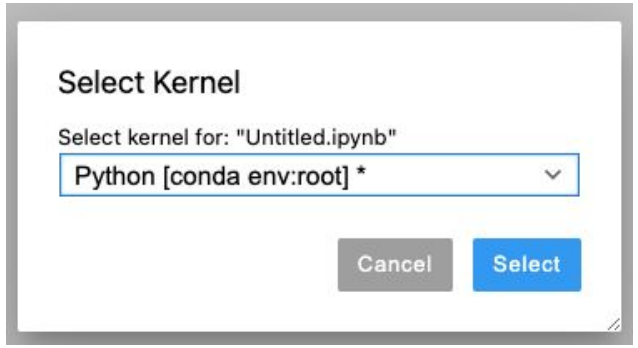
```
[jharvard@holy7c02111 ~]$ module load python
[jharvard@holy7c02111 ~]$ mamba create -n OOD_env python=3.11 pip wheel numpy
[jharvard@holy7c02111 ~]$ source activate OOD_env
(OOD_env)[jharvard@holy7c02111 ~]$ mamba install ipykernel
```

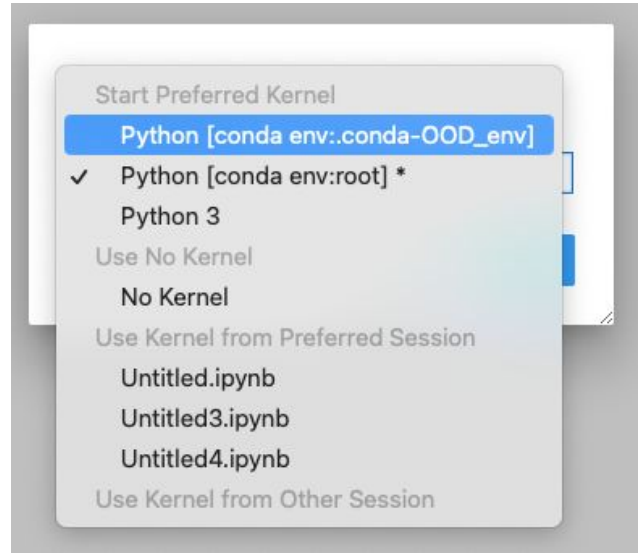
https://docs.rc.fas.harvard.edu/kb/python/#Mamba

Jupyter Notebook (2)

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

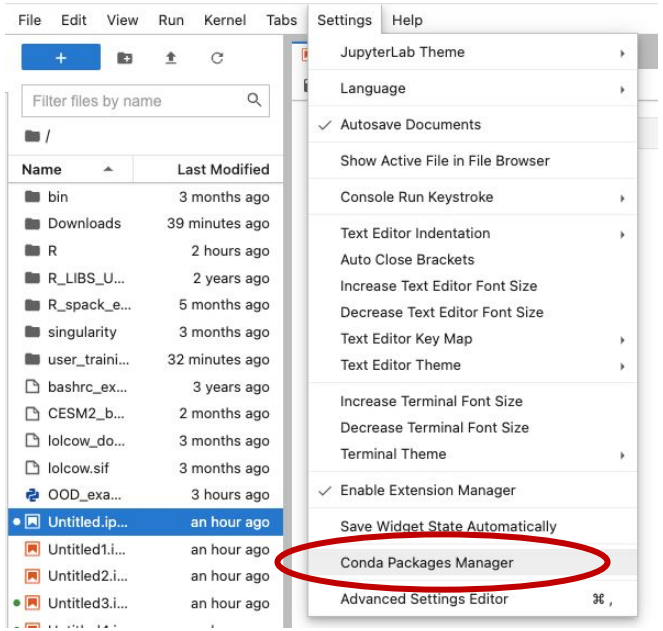


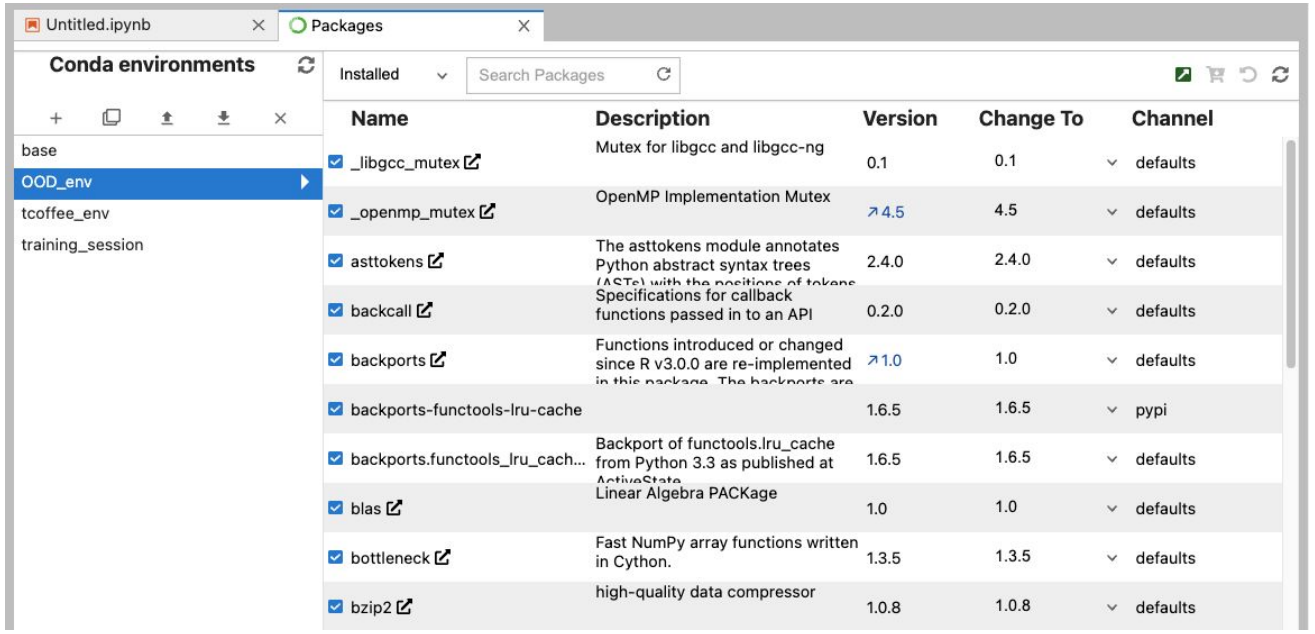




Jupyter Notebook (3)

5. Managing (install, uninstall, update) packages
   a. We recommend using the command line https://docs.rc.fas.harvard.edu/kb/python/#Mamba
   b. You can also use the conda package manager: On the top menu, click Settings -> Conda Package Manager -> OOD_env





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
- How?
  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
```

Remote Desktop

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

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

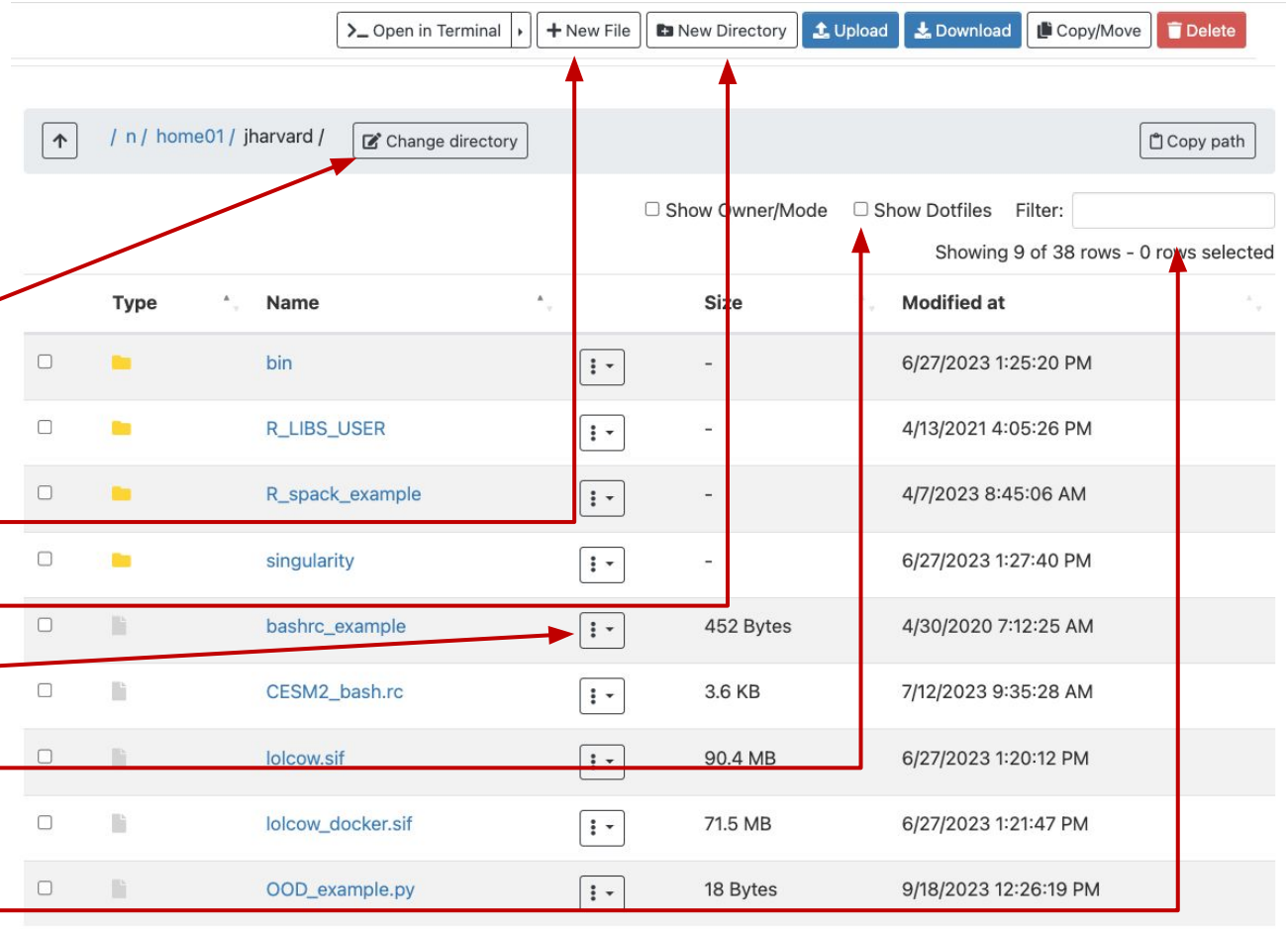
FASSE proxy

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

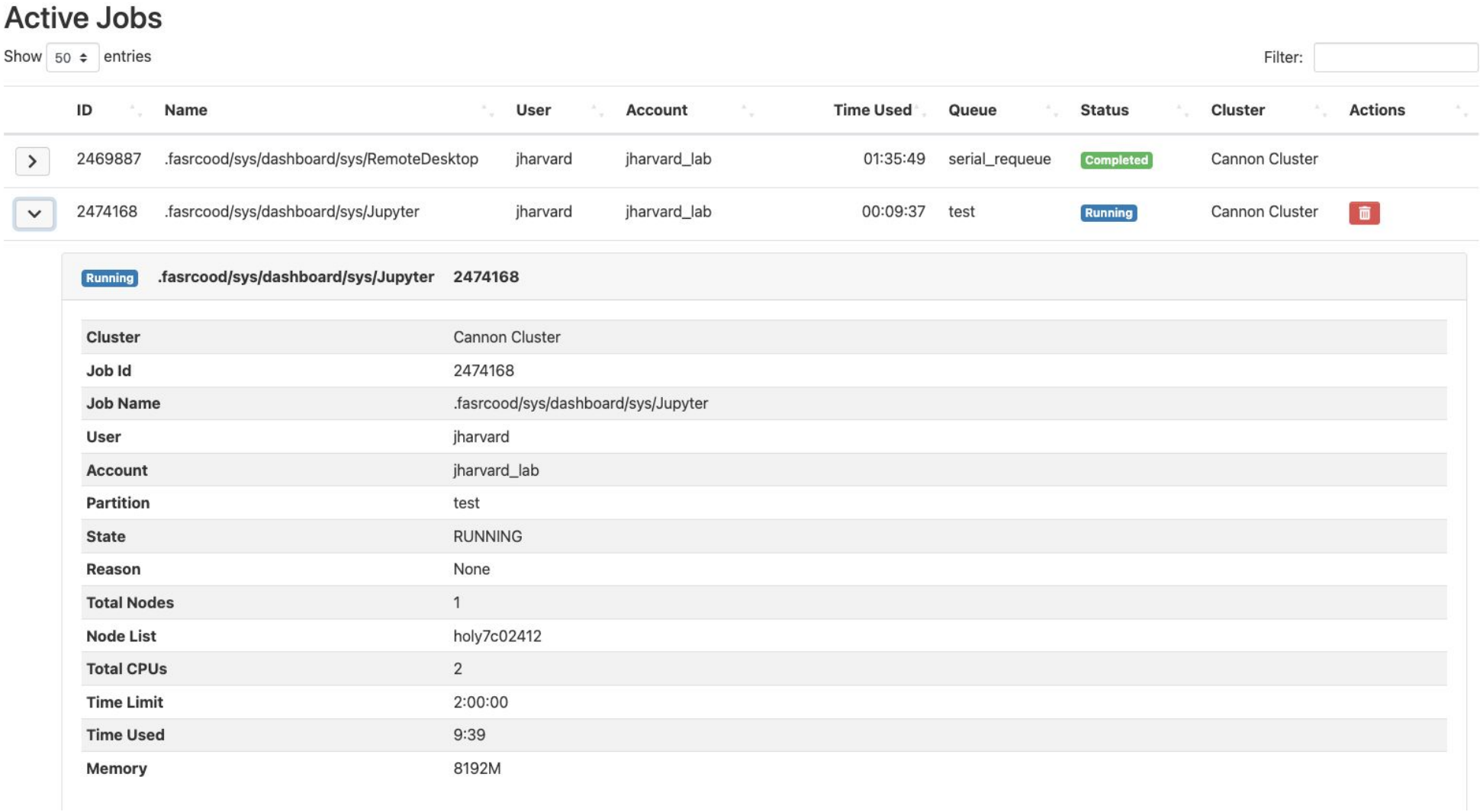
- You may need to set FASSE proxy on
  - RStudio server if you are unable to reach cran and download R packages
  - Stata if you are unable to load libraries via http
  - Firefox (web browsing)
  - Jupyter Notebook
  - Access Github
  - (Basically, anything outside of FASSE)

Files tab

- Default options: home directory and holyscratch
- Click on “Change directory” to go to a lab share
- 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



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

Delete

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

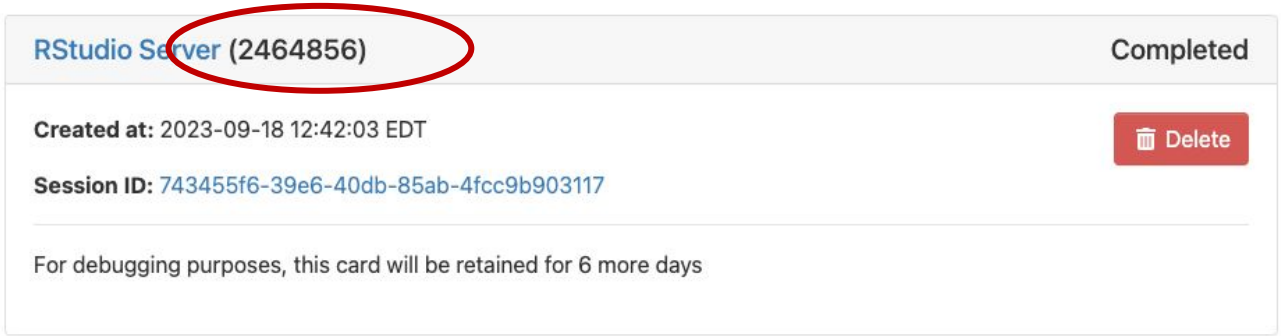
Undetermined .fasrcood/sys/dashboard/sys/Matlab 2474322

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




```
[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
```

Survey

Please, fill out our course survey. Your feedback is essential for us to improve our trainings!!

http://tinyurl.com/FASRCsurvey

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
    - Portal: http://portal.rc.fas.harvard.edu/rcrt/submit_ticket (requires login)
    - Email: rchelp@rc.fas.harvard.edu

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

Thank you :)

FAS Research Computing