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
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

FASSE

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

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
  - **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](https://fasrcdocs.com/rstudio-server-vs-rstudio-desktop-ood-apps)
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
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/](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

```bash
# 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

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>User</th>
<th>Account</th>
<th>Time Used</th>
<th>Queue</th>
<th>Status</th>
<th>Cluster</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2469887</td>
<td>fasrccood/sys/dashboard/sys/RemoteDesktop</td>
<td>jharvard</td>
<td>jharvard_lab</td>
<td>01:35:49</td>
<td>serial_requeue</td>
<td>Completed</td>
<td>Cannon Cluster</td>
<td></td>
</tr>
<tr>
<td>2474168</td>
<td>fasrccood/sys/dashboard/sys/Jupyter</td>
<td>jharvard</td>
<td>jharvard_lab</td>
<td>00:09:37</td>
<td>test</td>
<td>Running</td>
<td>Cannon Cluster</td>
<td></td>
</tr>
</tbody>
</table>

### Running 2474168

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Cannon Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job Id</td>
<td>2474168</td>
</tr>
<tr>
<td>Job Name</td>
<td>.fasrccood/sys/dashboard/sys/Jupyter</td>
</tr>
<tr>
<td>User</td>
<td>jharvard</td>
</tr>
<tr>
<td>Account</td>
<td>jharvard_lab</td>
</tr>
<tr>
<td>Partition</td>
<td>test</td>
</tr>
<tr>
<td>State</td>
<td>RUNNING</td>
</tr>
<tr>
<td>Reason</td>
<td>None</td>
</tr>
<tr>
<td>Total Nodes</td>
<td>1</td>
</tr>
<tr>
<td>Node List</td>
<td>holy?c02412</td>
</tr>
<tr>
<td>Total CPUs</td>
<td>2</td>
</tr>
<tr>
<td>Time Limit</td>
<td>2:00:00</td>
</tr>
<tr>
<td>Time Used</td>
<td>9:39</td>
</tr>
<tr>
<td>Memory</td>
<td>8192M</td>
</tr>
</tbody>
</table>
### Jobs tab (2)

#### Matlab (2474322)
- **Created at:** 2023-09-18 15:28:58 EDT
- **Time Requested:** 1 hour
- **Session ID:** 0846728-1d9f-4a91-8776-582272b74ec0

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

#### Active Jobs

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>User</th>
<th>Account</th>
<th>Time Used</th>
<th>Queue</th>
<th>Status</th>
<th>Cluster</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2469887</td>
<td>.fasrcod/sys/dashboard/sys/RemoteDesktop</td>
<td>jharvard</td>
<td>jharvard_lab</td>
<td>01:35:49</td>
<td>serial_requeue</td>
<td>Completed</td>
<td>Cannon Cluster</td>
<td></td>
</tr>
<tr>
<td>2474322</td>
<td>.fasrcod/sys/dashboard/sys/Matlab</td>
<td>jharvard</td>
<td>jharvard_lab</td>
<td>00:02:27</td>
<td>test</td>
<td>Undetermined</td>
<td>Cannon Cluster</td>
<td></td>
</tr>
<tr>
<td>2474168</td>
<td>.fasrcod/sys/dashboard/sys/Jupyter</td>
<td>jharvard</td>
<td>jharvard_lab</td>
<td>00:15:45</td>
<td>test</td>
<td>Running</td>
<td>Cannon Cluster</td>
<td></td>
</tr>
</tbody>
</table>
Jobs tab (3)

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Cannon Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job Id</td>
<td>2474322</td>
</tr>
<tr>
<td>Job Name</td>
<td>.fasrcood/sys/dashboard/sys/Matlab</td>
</tr>
<tr>
<td>User</td>
<td>jharvard</td>
</tr>
<tr>
<td>Account</td>
<td>jharvard_lab</td>
</tr>
<tr>
<td>Partition</td>
<td>test</td>
</tr>
<tr>
<td>State</td>
<td>OUT_OF_MEMORY</td>
</tr>
<tr>
<td>Reason</td>
<td>OutOfMemory</td>
</tr>
<tr>
<td>Total Nodes</td>
<td>1</td>
</tr>
<tr>
<td>Total CPUs</td>
<td>2</td>
</tr>
<tr>
<td>Time Limit</td>
<td>1:00:00</td>
</tr>
<tr>
<td>Time Used</td>
<td>2:27</td>
</tr>
<tr>
<td>Memory</td>
<td>4096M</td>
</tr>
</tbody>
</table>
If job no longer appears on “Active Jobs”, check job status from command line with slurm job ID

[jharvard@boslogin01 ~]$ sacct -j 2464856

<table>
<thead>
<tr>
<th>JobID</th>
<th>JobName</th>
<th>Partition</th>
<th>Account</th>
<th>AllocCPUS</th>
<th>State</th>
<th>ExitCode</th>
</tr>
</thead>
<tbody>
<tr>
<td>2464856</td>
<td>.fasrcood+</td>
<td>test</td>
<td>jharvard_+</td>
<td>2</td>
<td>TIMEOUT</td>
<td>0:0</td>
</tr>
<tr>
<td>2464856</td>
<td>bat+</td>
<td>batch</td>
<td>jharvard_+</td>
<td>2</td>
<td>CANCELLED</td>
<td>0:15</td>
</tr>
<tr>
<td>2464856</td>
<td>ext+</td>
<td>extern</td>
<td>jharvard_+</td>
<td>2</td>
<td>COMPLETED</td>
<td>0:0</td>
</tr>
</tbody>
</table>

[jharvard@holy7c02111 ~]$ sacct -j 2471535

<table>
<thead>
<tr>
<th>JobID</th>
<th>JobName</th>
<th>Partition</th>
<th>Account</th>
<th>AllocCPUS</th>
<th>State</th>
<th>ExitCode</th>
</tr>
</thead>
<tbody>
<tr>
<td>2471535</td>
<td>.fasrcood+</td>
<td>test</td>
<td>jharvard_+</td>
<td>2</td>
<td>OUT_OF_ME+</td>
<td>0:125</td>
</tr>
<tr>
<td>2471535</td>
<td>bat+</td>
<td>batch</td>
<td>jharvard_+</td>
<td>2</td>
<td>OUT_OF_ME+</td>
<td>0:125</td>
</tr>
<tr>
<td>2471535</td>
<td>ext+</td>
<td>extern</td>
<td>jharvard_+</td>
<td>2</td>
<td>COMPLETED</td>
<td>0:0</td>
</tr>
</tbody>
</table>
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/](https://docs.rc.fas.harvard.edu/)
- GitHub User codes: [https://github.com/fasrc/User_Codes/](https://github.com/fasrc/User_Codes/)

Getting help

- Office hours: [https://www.rc.fas.harvard.edu/training/office-hours/](https://www.rc.fas.harvard.edu/training/office-hours/)
- Ticket
  - Portal: [http://portal.rc.fas.harvard.edu/rcrt/submit_ticket](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