Cryo-EM Research Center

Intro to the Cryo-EM HPC Cluster Resource

Jennifer Scheuren and Matt Larson

Introduction to HPC

1. About the Cryo-EM Cluster
2. Getting started:
    1. How to request an account
    2. Connecting with Open OnDemand
    3. Basics with SBatch scripts or submit_to_<queue> commands.
3. Getting data to/from the HPC Cluster with Globus
4. Examples
5. Policies
6. Help and more information

Where to get HPC help

- Jennifer (jfscheuren@wisc.edu)
  - HPC Administrator for CEMRC
- Matt (mrlarson2@wisc.edu)
  - Sys admin for the CEMRC/MCCET facilities

Email to jfscheuren@wisc.edu and
CC: cryoem@biochem.wisc.edu

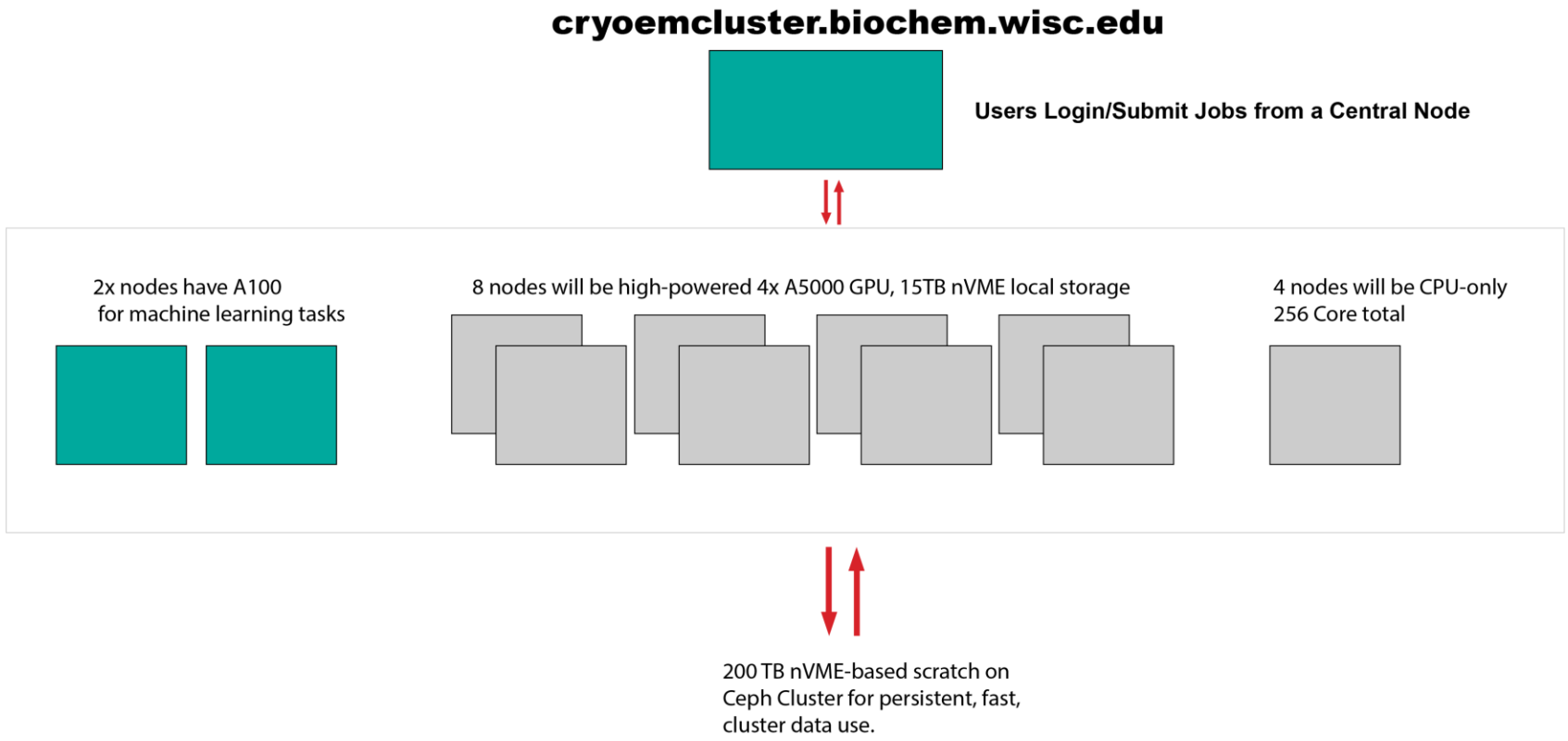
HPC Administrator

Computer Scientist/System Administrator

Slurm

- “Simple Linux Utility for Resource Management”
- Open-source job scheduling and cluster management tool
- Highly customizable – Control resources, jobs and users
- **Cluster** – set of computers for high-performance computing
- **Head node** – computer from which users can submit jobs
- **Compute nodes** – computers that jobs will be assigned by the scheduler. Users do not log in directly.
- **Partition (or Queue)** – a pool of nodes from some part of the cluster
  - Ex. “GPU”, ”A5000”, “A100” and “CPU” (what hardware is available in each queue in the cryoemcluster.biochem.wisc.edu)

slurm
workload manager



Cryo-EM HTC Cluster

Comparing different compute nodes

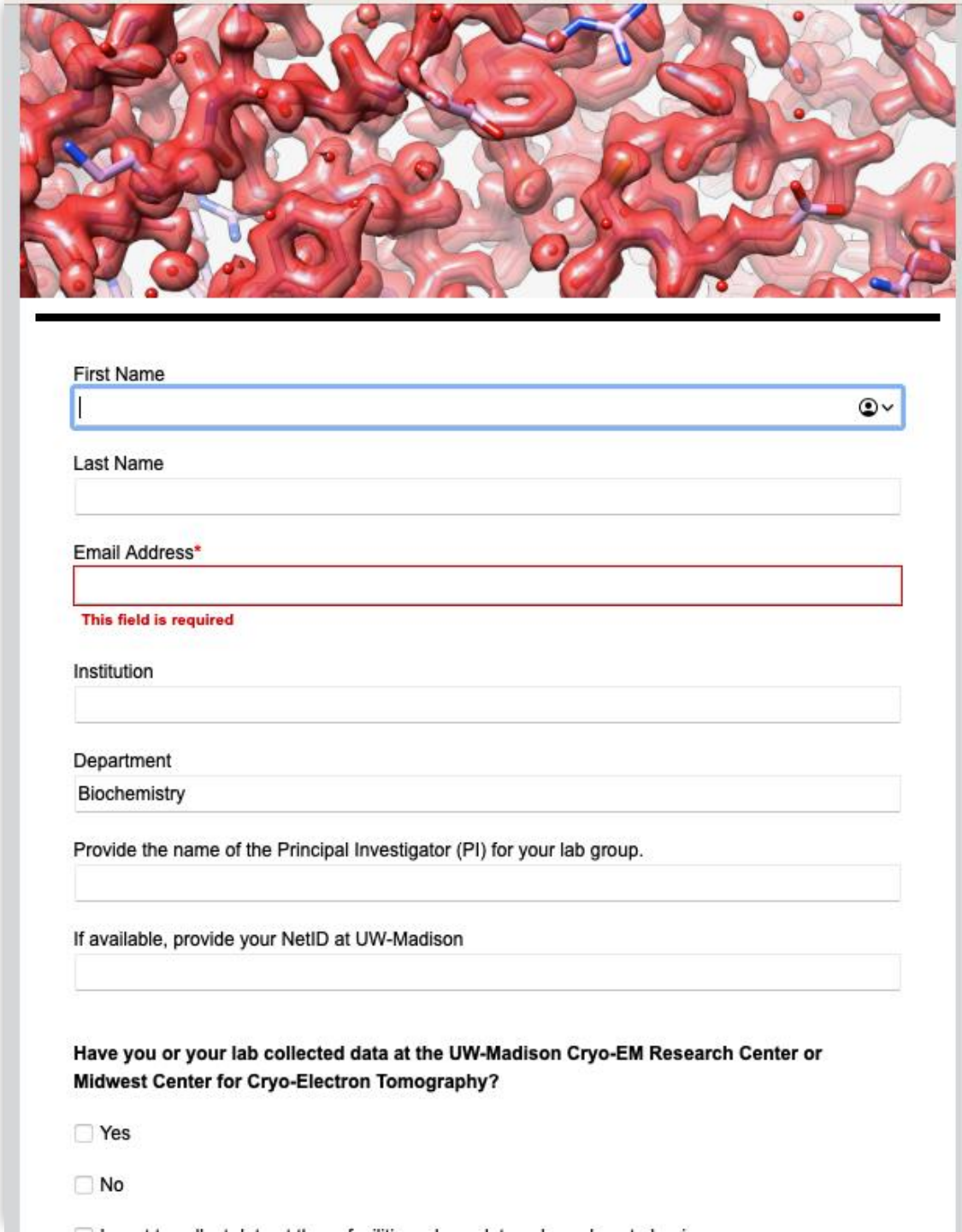
Compute node	GPU	Memory	CPU
a100	4x Nvidia a100s	512 GB	2x AMD 7443 48 cores
a5000	4x Nvidia a5000s	512GB	1x AMD EPYC 64 cores
cpu	none	256 GB	1x AMD EPYC 64 cores

- Use gpu nodes for jobs that can support gpu processing.
- Use cpu nodes for jobs that cannot support gpu processing.

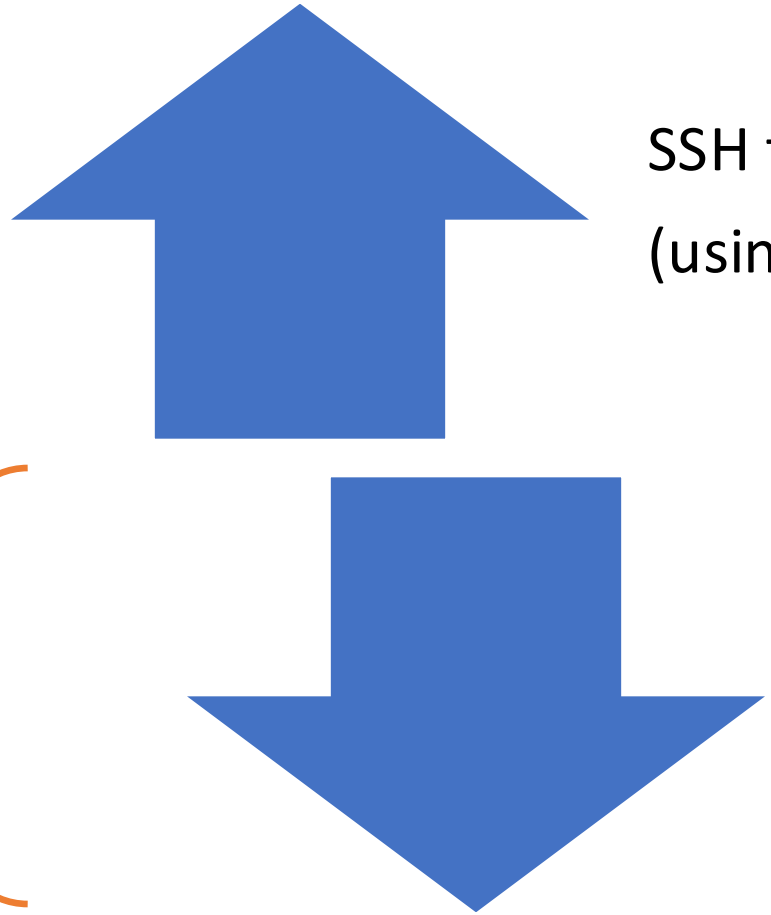
Getting started: request an account

https://explore.wisc.edu/cryoem_hpc_access_request

- You will be added to an allowed access group, and a network directory with quota will be created for use on the HPC Cluster.
- Afterwards you be contacted to you let you know you can now login.



Two ways to connect!



SSH to cryoemcluster.biochem.wisc.edu

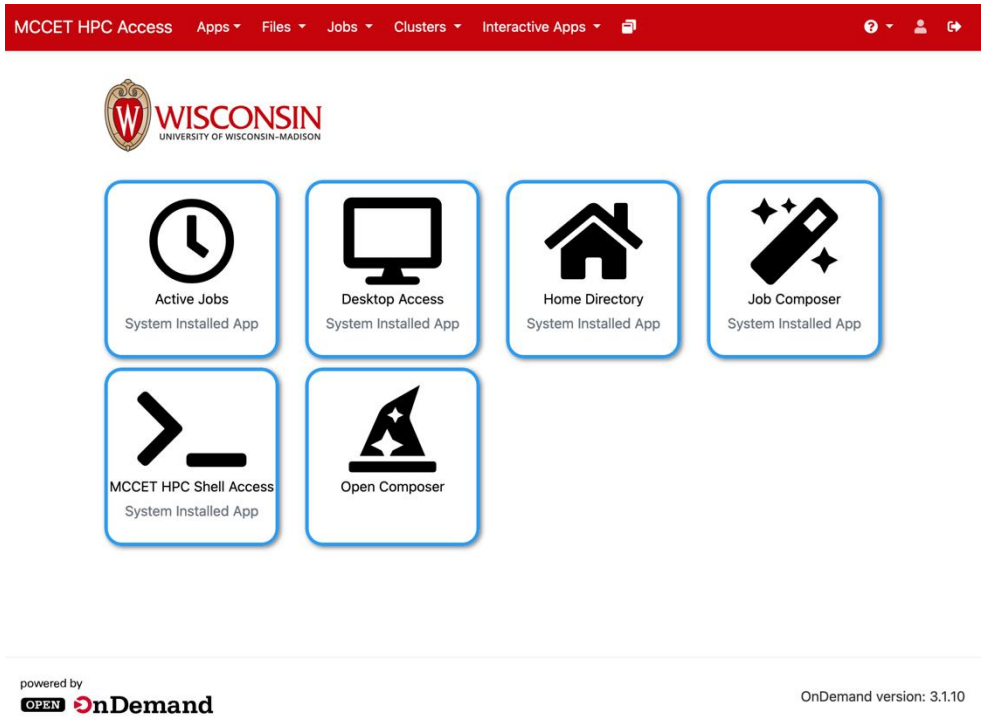
(using X11 forwarding for graphical apps)

Open OnDemand web portal at
https://ondemandcryo.biochem.wisc.edu

OPEN OnDemand

Provides cluster access via web browser.

- Home directory access
- Active jobs
- Shell access
- Interactive Desktop
- Open composer



Go to https://ondemandcryo.biochem.wisc.edu
Login with your NetID *(Let's Try this Now!)*

Alternatively, login via ssh

- Login via SSH:

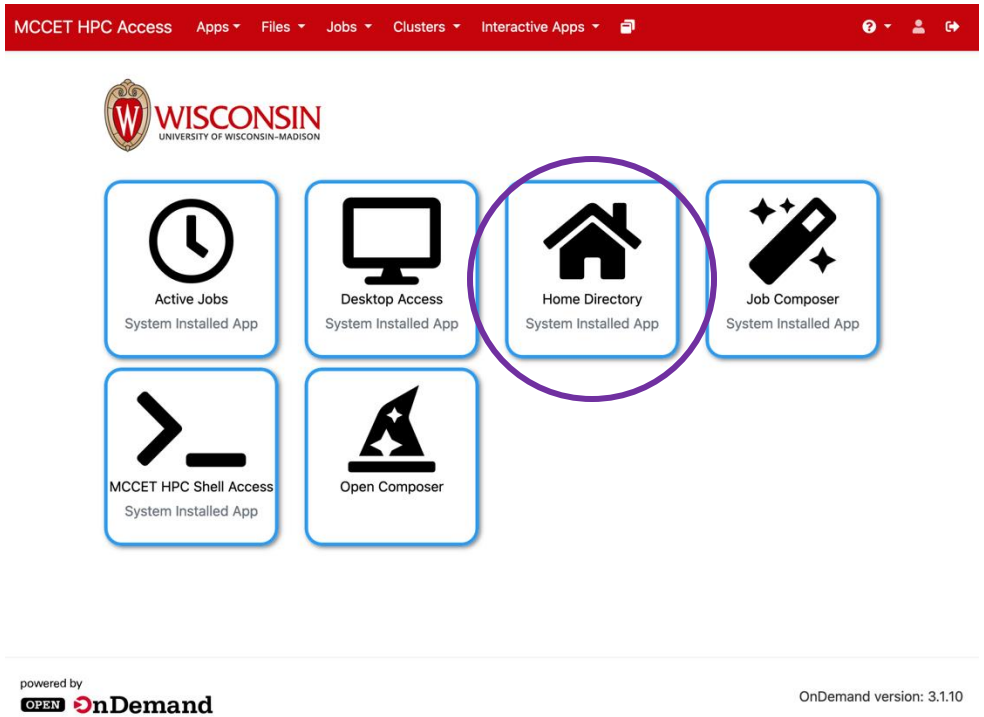
  `ssh -YC netid@wisc.edu@cryoemcluster.biochem.wisc.edu`

- The (-Y) and (-C) options with SSH enable X11 forwarding, with compression for best performance.
- You need to be on the Biochemistry VPN first.

OPEN OnDemand

Provides cluster access via web browser.

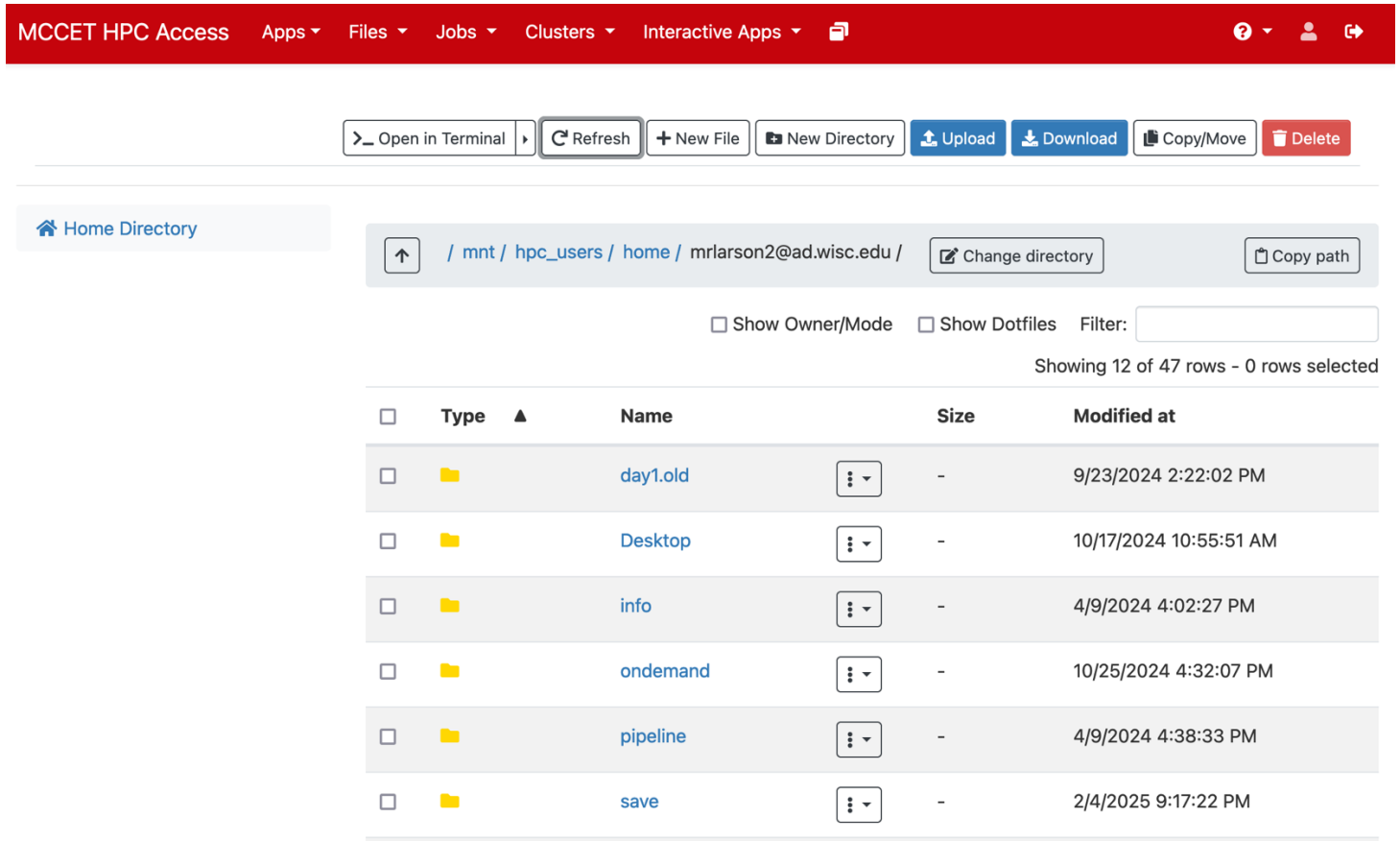
- Home directory access
- Active jobs
- Shell access
- Interactive Desktop
- Open composer



Go to https://ondemandcryo.biochem.wisc.edu

Home Directory

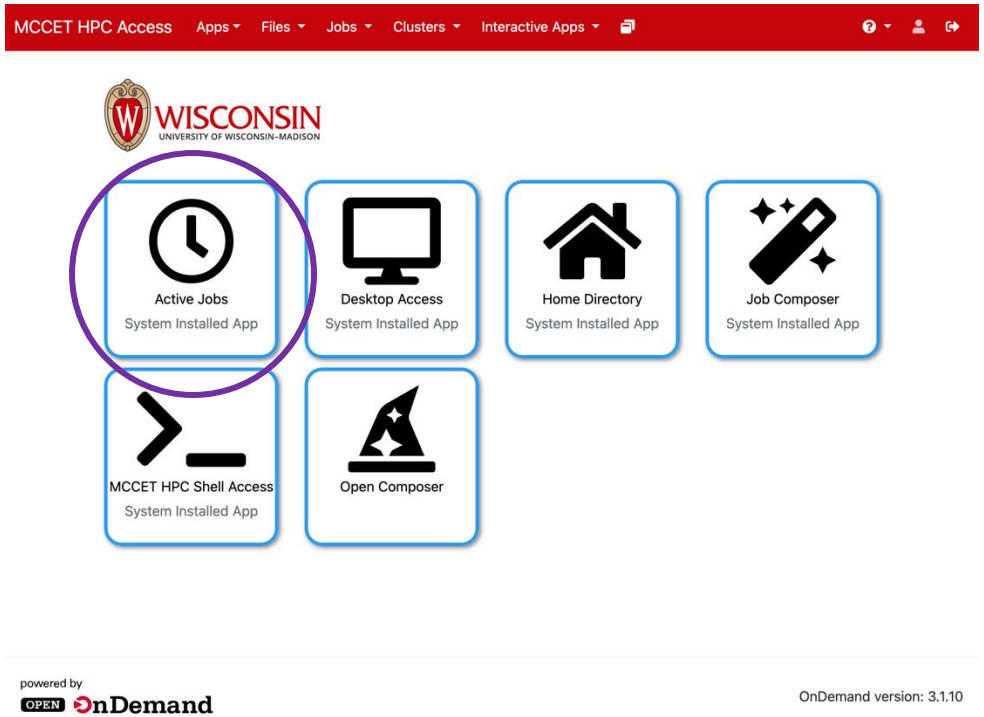
- View contents of your home directory, and transfer files.
- *But.. use Globus for larger data transfers.*



OPEN OnDemand

Provides cluster access via web browser.

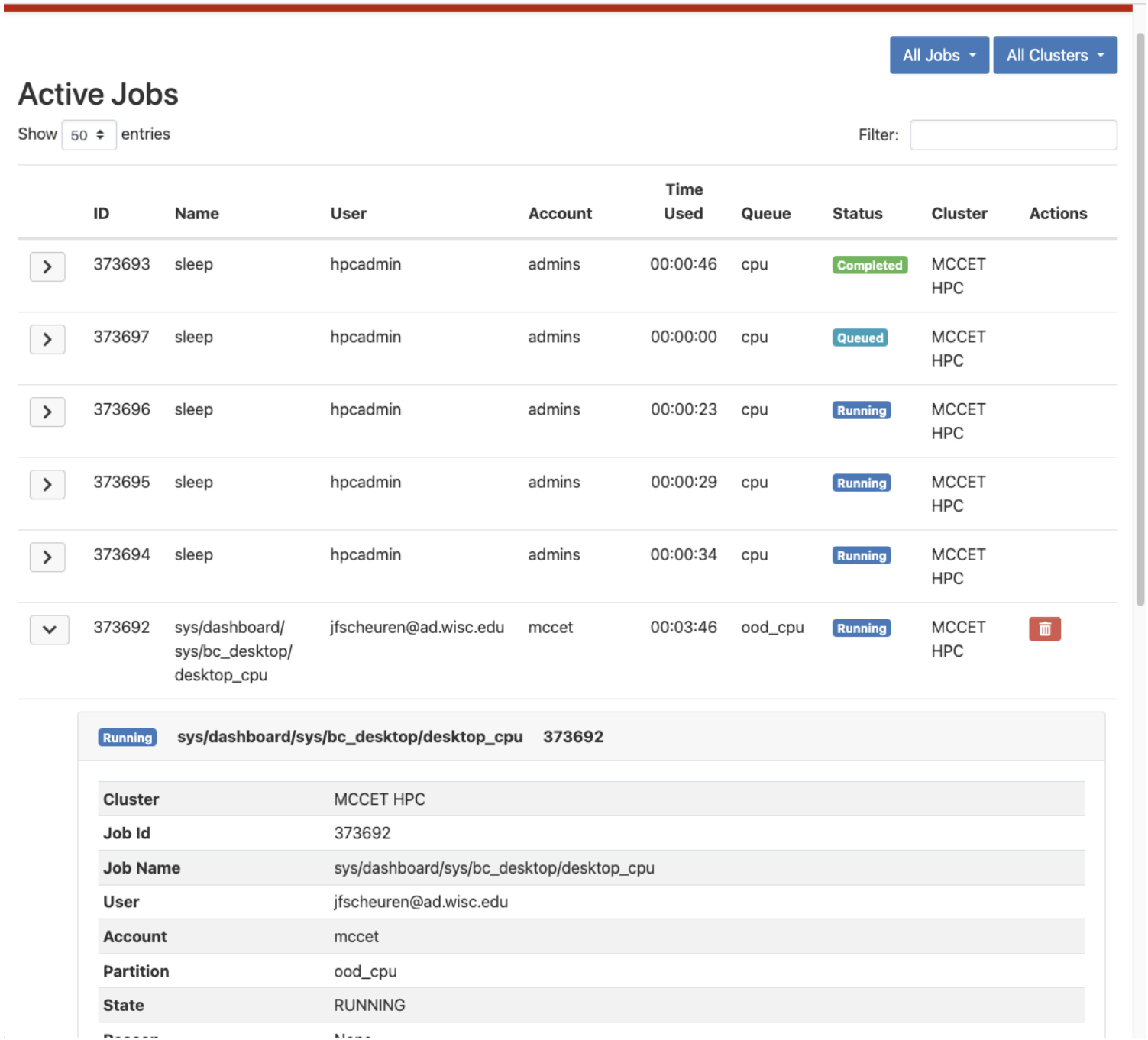
- Home directory access
- Active jobs
- Shell access
- Interactive Desktop
- Open composer



Go to https://ondemandcryo.biochem.wisc.edu

Active Jobs

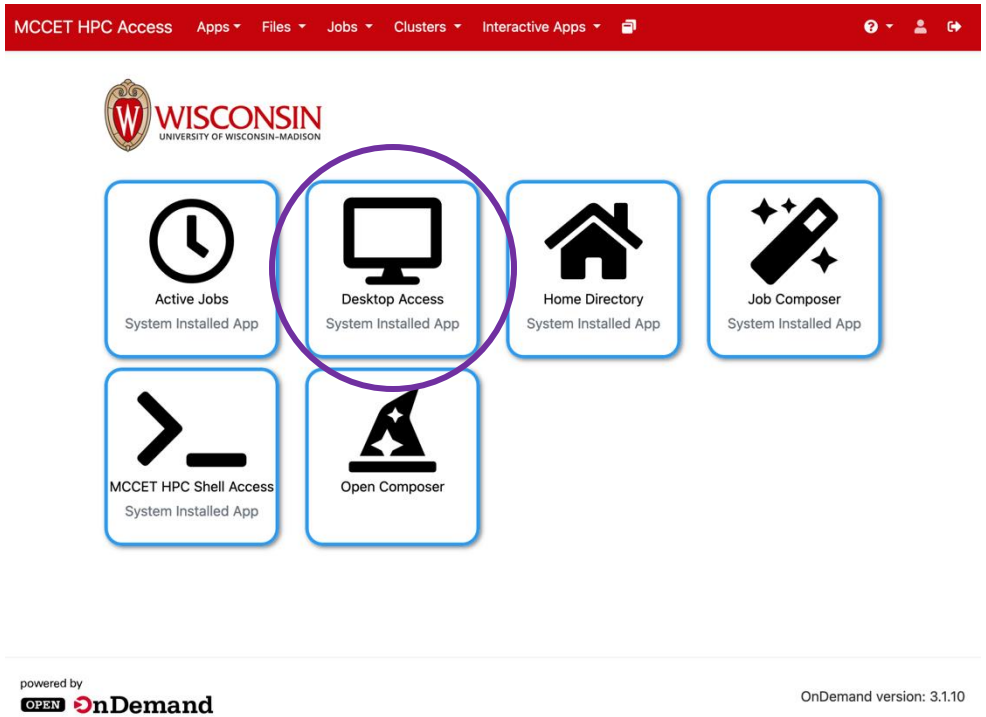
- Show running, queued, completed, and failed jobs.
- Manage running jobs.



OPEN OnDemand

Provides cluster access via web browser.

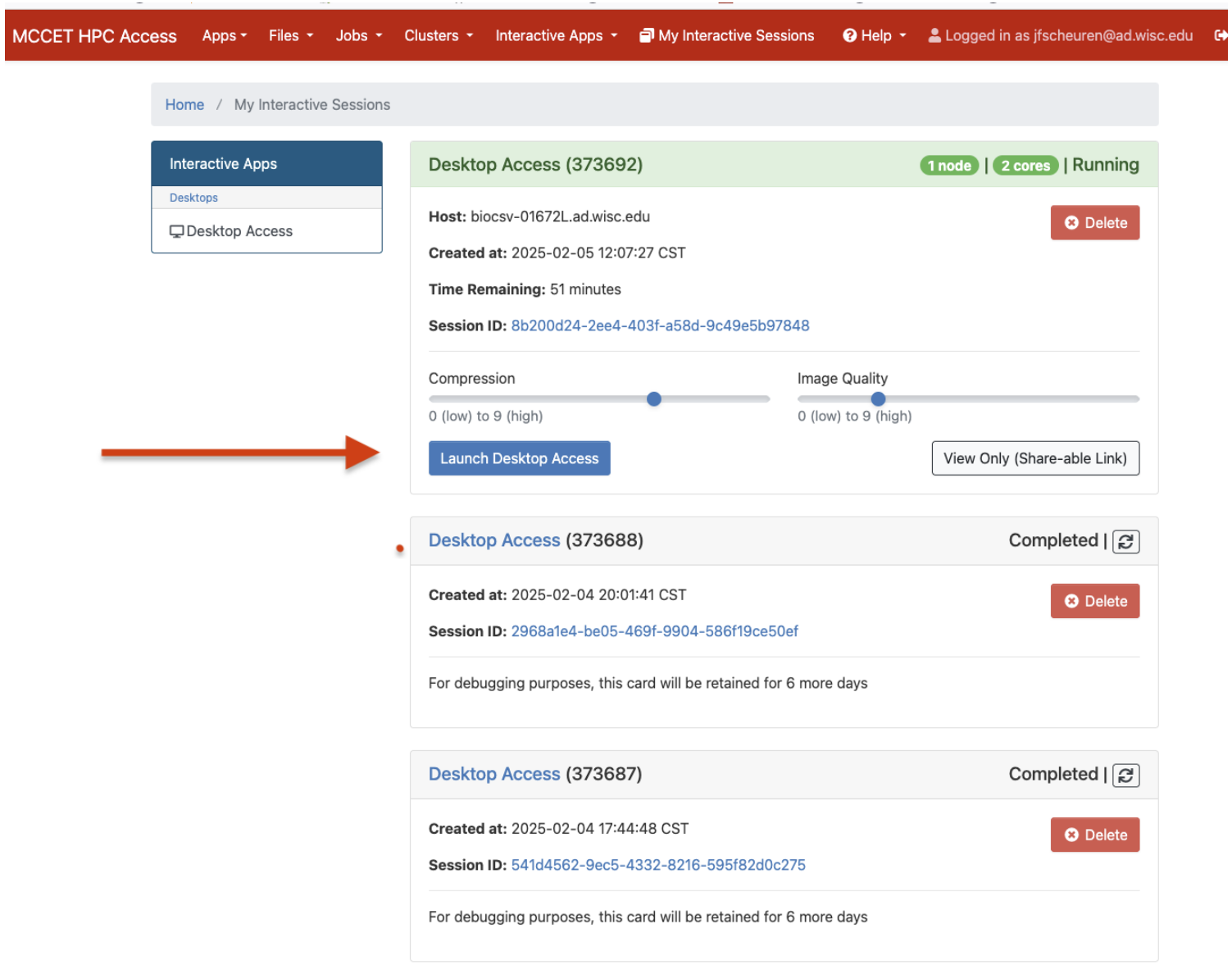
- Home directory access
- Active jobs
- Shell access
- Interactive Desktop
- Open composer



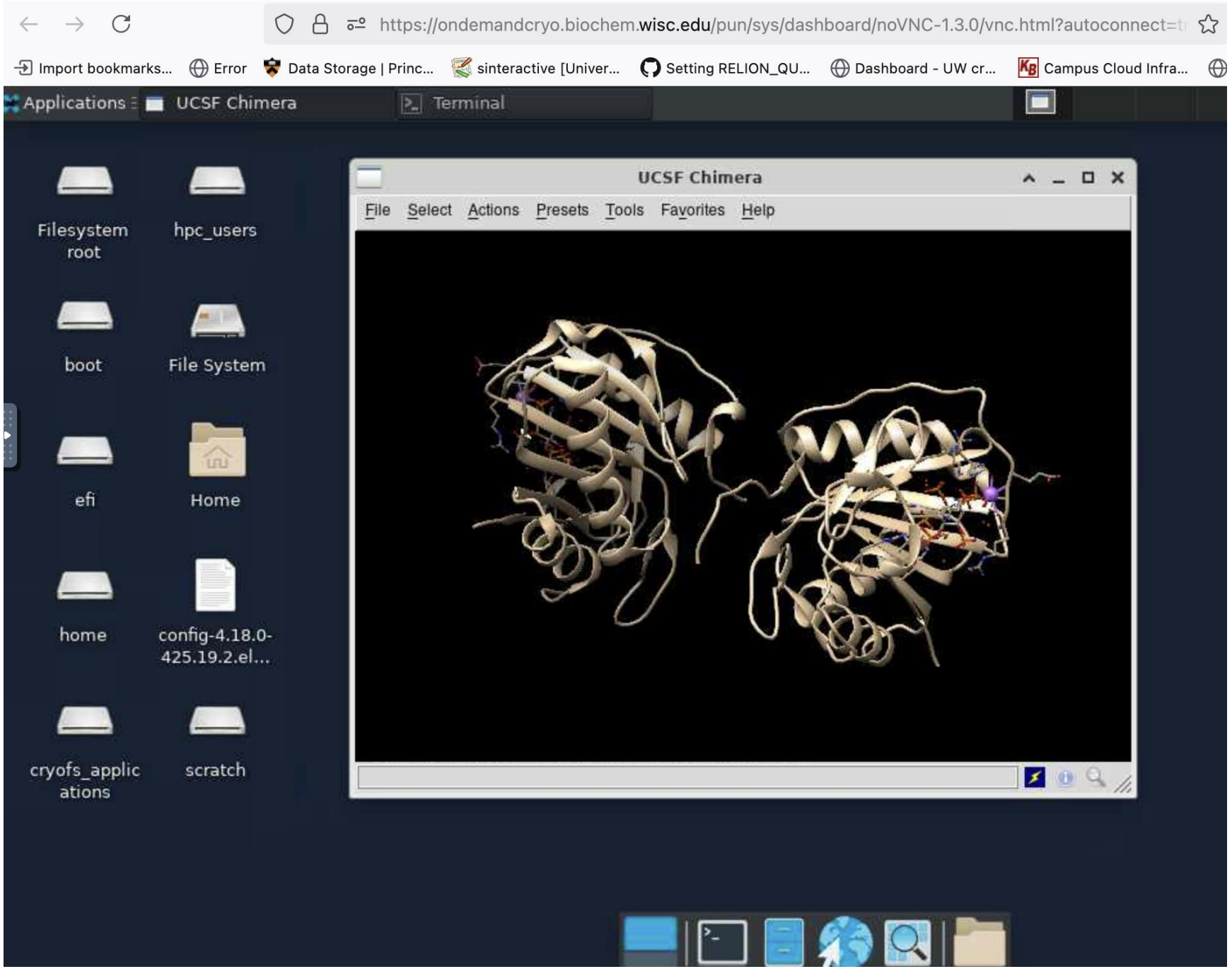
Go to https://ondemandcryo.biochem.wisc.edu
Login with your NetID *(Let's Try this Now!)*

Interactive Desktop Access

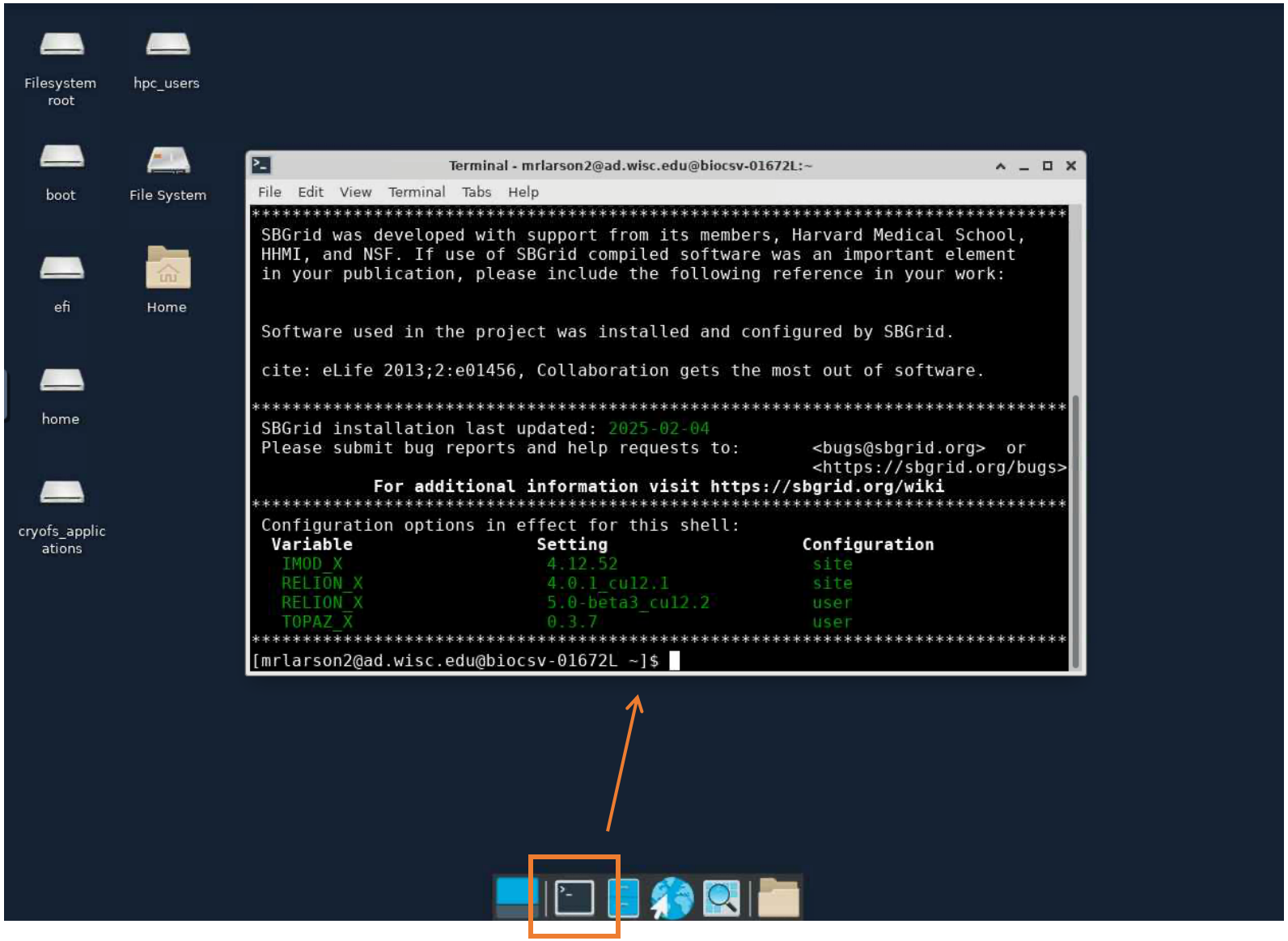
- Access the desktop of one of the compute nodes. You can currently select between a machine in the cpu or a5000 queues.



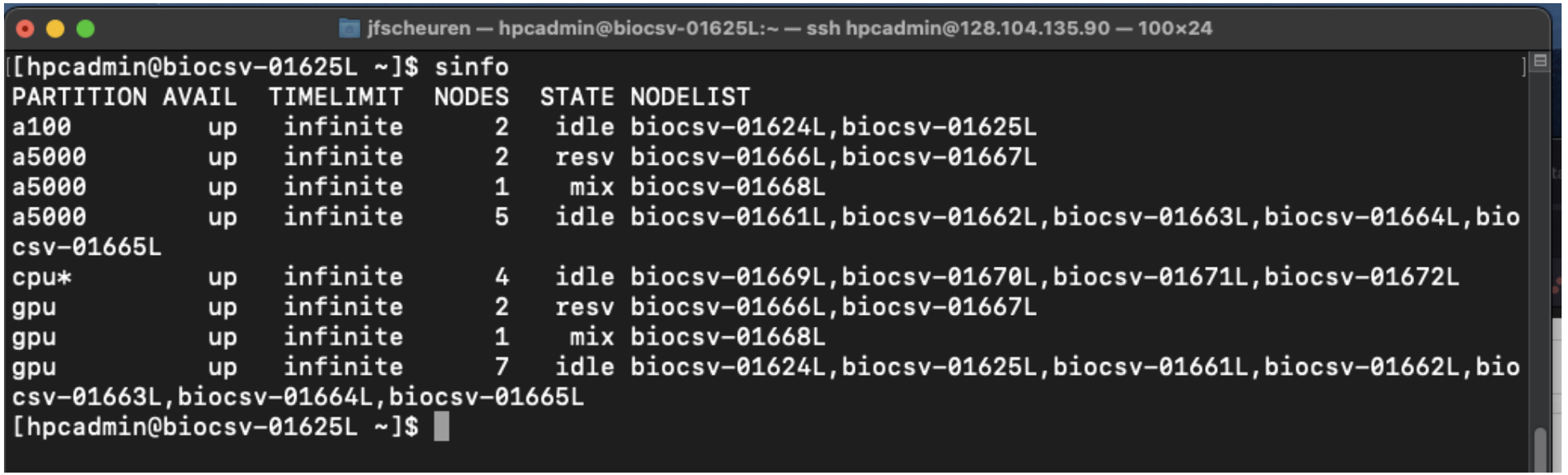
Example Desktop: Chimera



Open a terminal in the OOD desktop



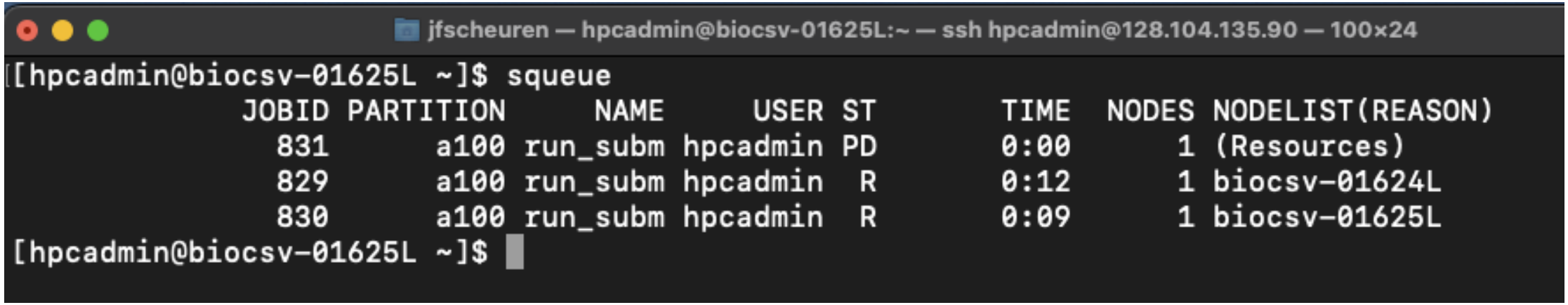
Getting started commands: sinfo



Displays partitions and state of the nodes

- See which nodes are busy – “STATE” column

Getting started commands: squeue



Displays pending and currently running jobs

Sbatch and Srun

- Sbatch – submits a batch script with configuration and task commands to Slurm
- Srun – runs job interactively from command line (if you close terminal session the job will be cancelled)
- Srun – in sbatch scripts creates ’job steps’ - srun jobs can be run in parallel using multiple srun commands with an “&” after the command

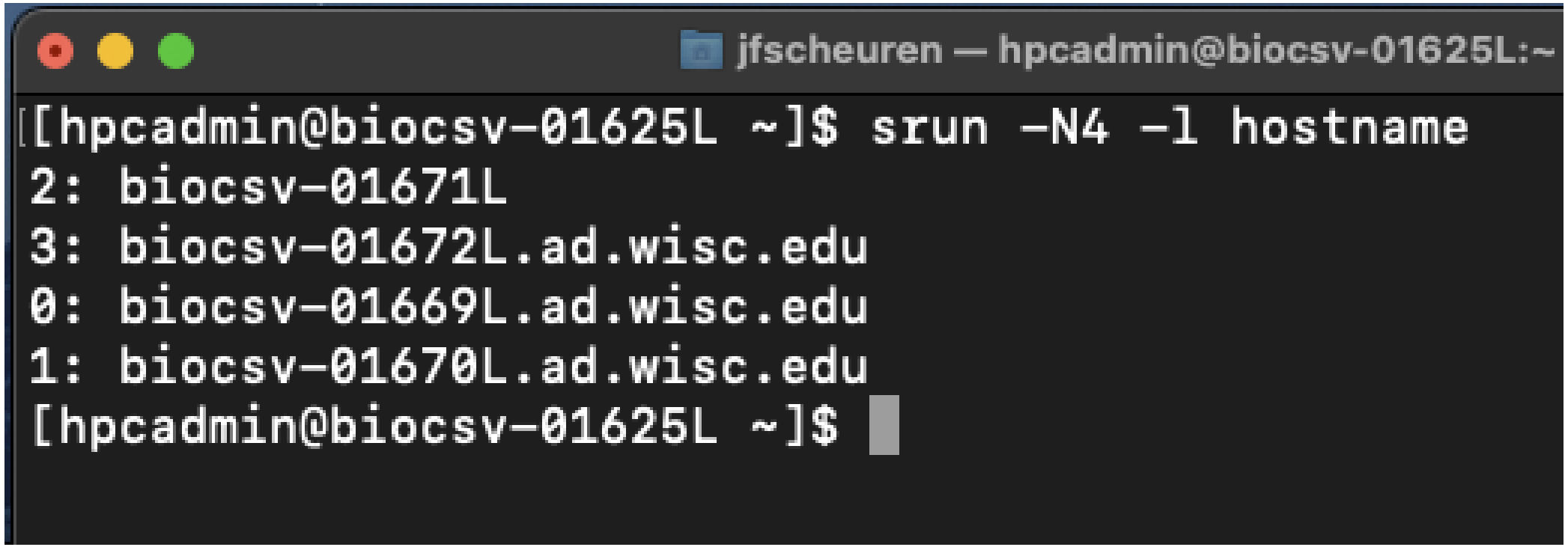
Job Profiles

- Each job can specify the resources it will require:
  - Partition to run on
  - Number of compute nodes to reserve
  - CPU count per node
  - CPU features
  - GPU count
  - GPU features (minimum size, etc)
  - Minimum free local space
  - Runtime required

- In SLURM, these are defined by the “Sbatch” script. (#SBATCH lines)

Run a simple command

- srun -N4 -l hostname. ← runs the hostname command on 4 different nodes

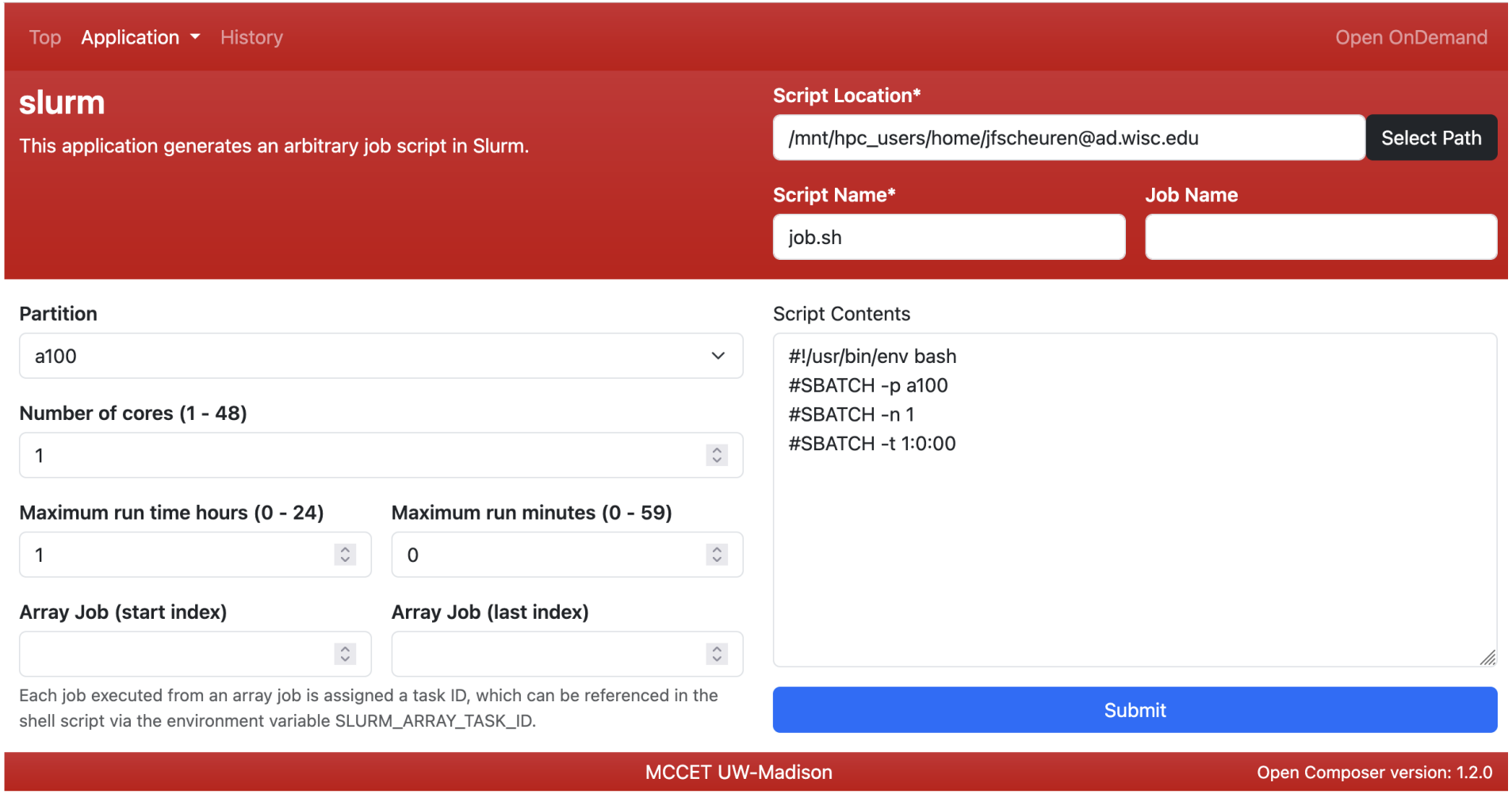


Sbatch example

```
#!/bin/bash

#SBATCH --partition=cpu     ## specify which partition (group of nodes) to run on
#SBATCH --nodes=1           ## specify how many nodes to use
#SBATCH --cpus-per-task=1   ## number of cpus per task
#SBATCH --mem=100M          ## amount of memory to be allocated from node
#SBATCH --time=1:00:00      ## allowed runtime


srun hostname               ## command to be run

~
```

Building jobs with Open OnDemand: Open Composer

- View and submit jobs
- Works better than built-in Job Composer



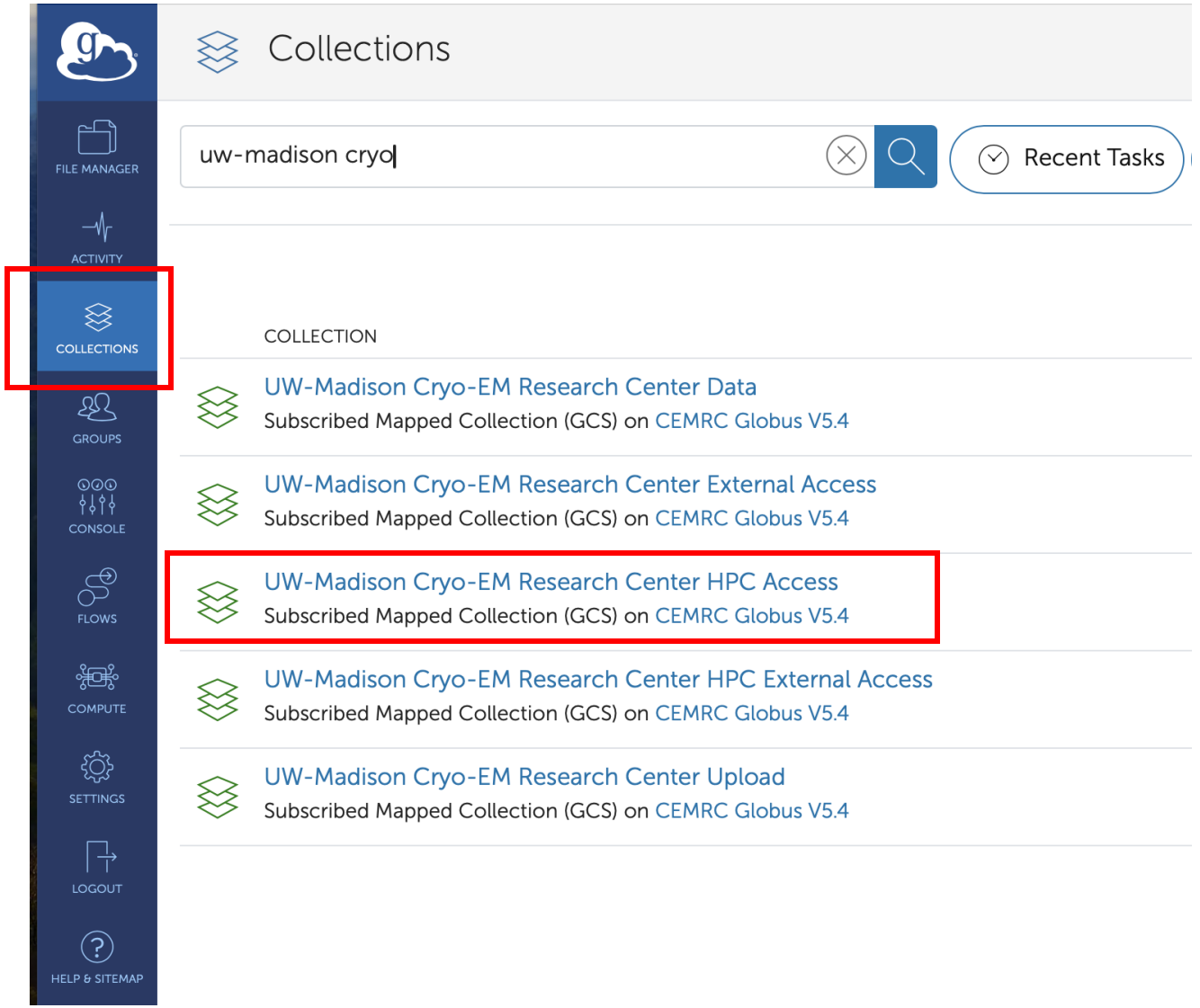
Finding on Globus

Go to https://app.globus.org

to use the Globus tool for moving data between facility and research group collections

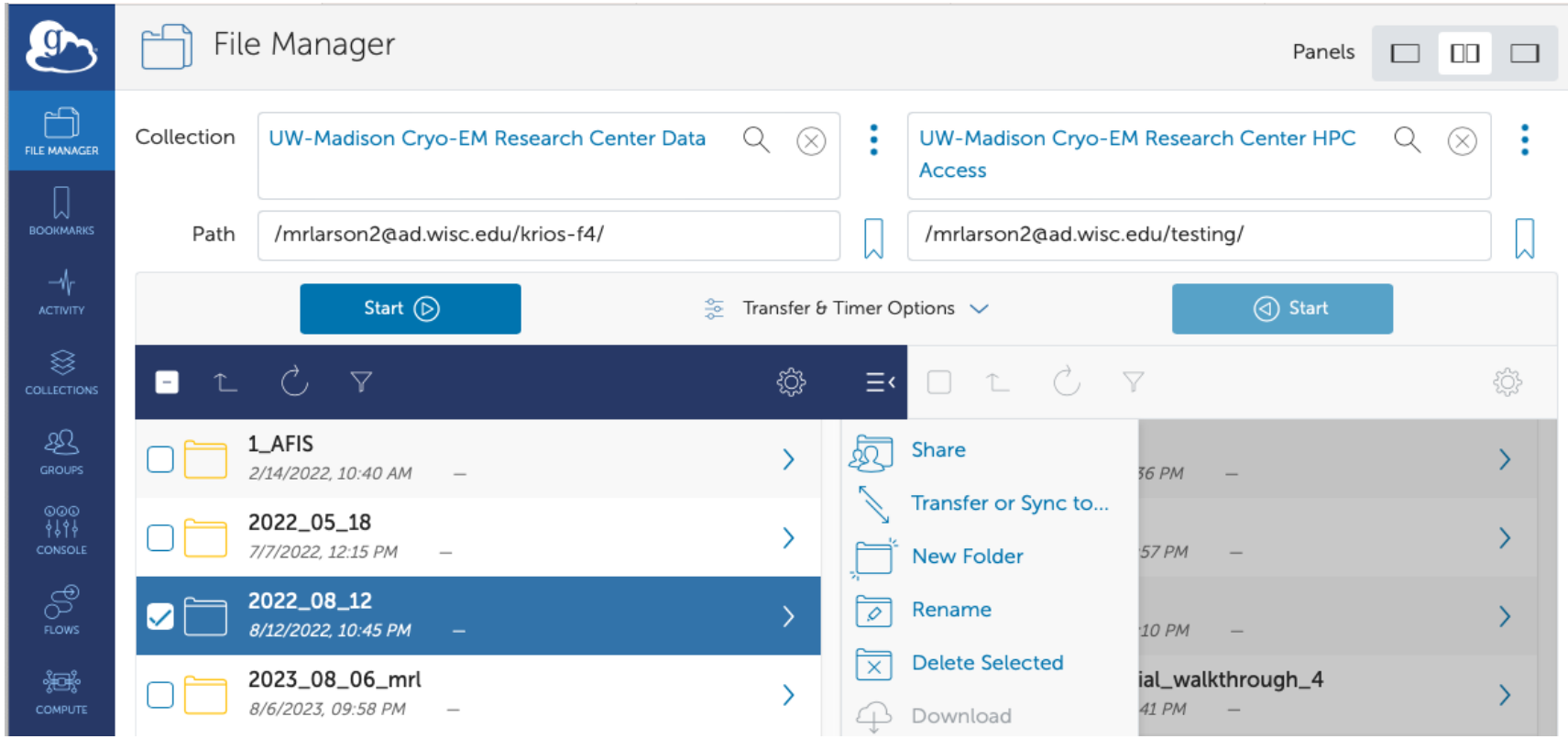
This will go directly to your home directory on the cluster.

You can setup a second collection with Cryo-EM Center Data or Wright lab collections to copy data.



Examples: transfer via Globus

- Choose a dataset and set destination to the HPC cluster, Start!



See also “UW-Madison Wright” for transfers to/from cryofs_wright data.

Getting started: submitting an AreTomo job

submit_to_a5000.sh AreTomo -InMrc 1138_G1__L4_TS_001_aligned.st -OutMrc test/tomogram.mrc -VolZ 1350 -AlignZ 1200 -OutBin 6 -DarkTol 0.1 -FlipVol 1 -Kv 300 -PixSize 1.4 -Wbp 1 -AngFile angles.txt -Patch 4 4 -TiltAxis 89.9

Submits a job via the pre-built Sbatch script from **/mnt/hpc_users/share/sbatch/a5000_gpu_sbatch.sh**

Outputs are “Submitted batch job 833” after job starts, and then you can can follow output of the jobs via “slurm.out” and “slurm.err” and the results in the “test/” output folder.

tail –f slurm_<id>.out to follow the standard output.

ssh –YC netid@wisc.edu@cryoemcluster.biochem.wisc.edu

cd $HOME

tar -xvf /mnt/hpc_users/share/resources/HPC_sample_data.tar

cd HPC_sample_data/AreTomo_2023_07_13/Evaluation/1138_G1_L4_TS_001/

> Examine the command for AreTomo

cat Example_Command.sh

> Submit your job!

submit_to_gpu.sh ./Example_Command.sh

> Check what is running!

squeue

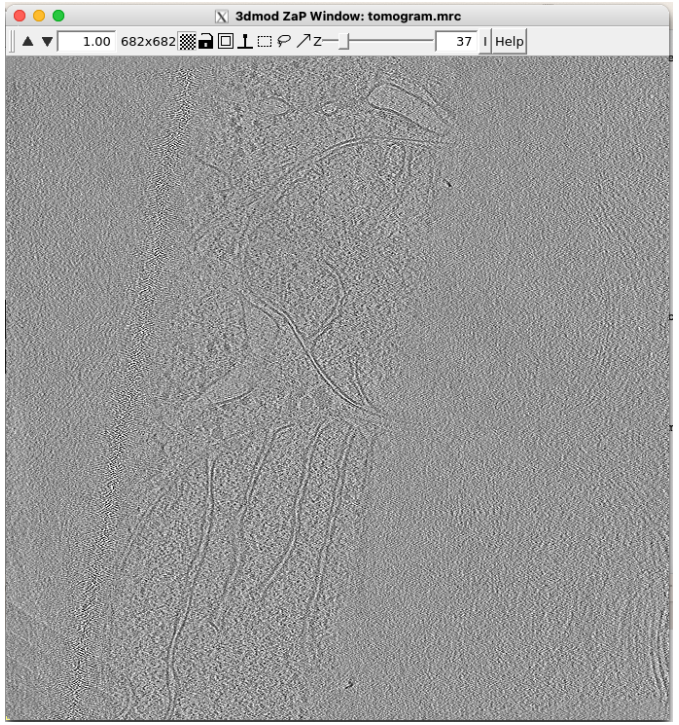
```
JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
  845       gpu gpu_sbat mrlarson  R       9:18      1 biocsv-01668L
```

sinfo

> See output of the job and view the results…

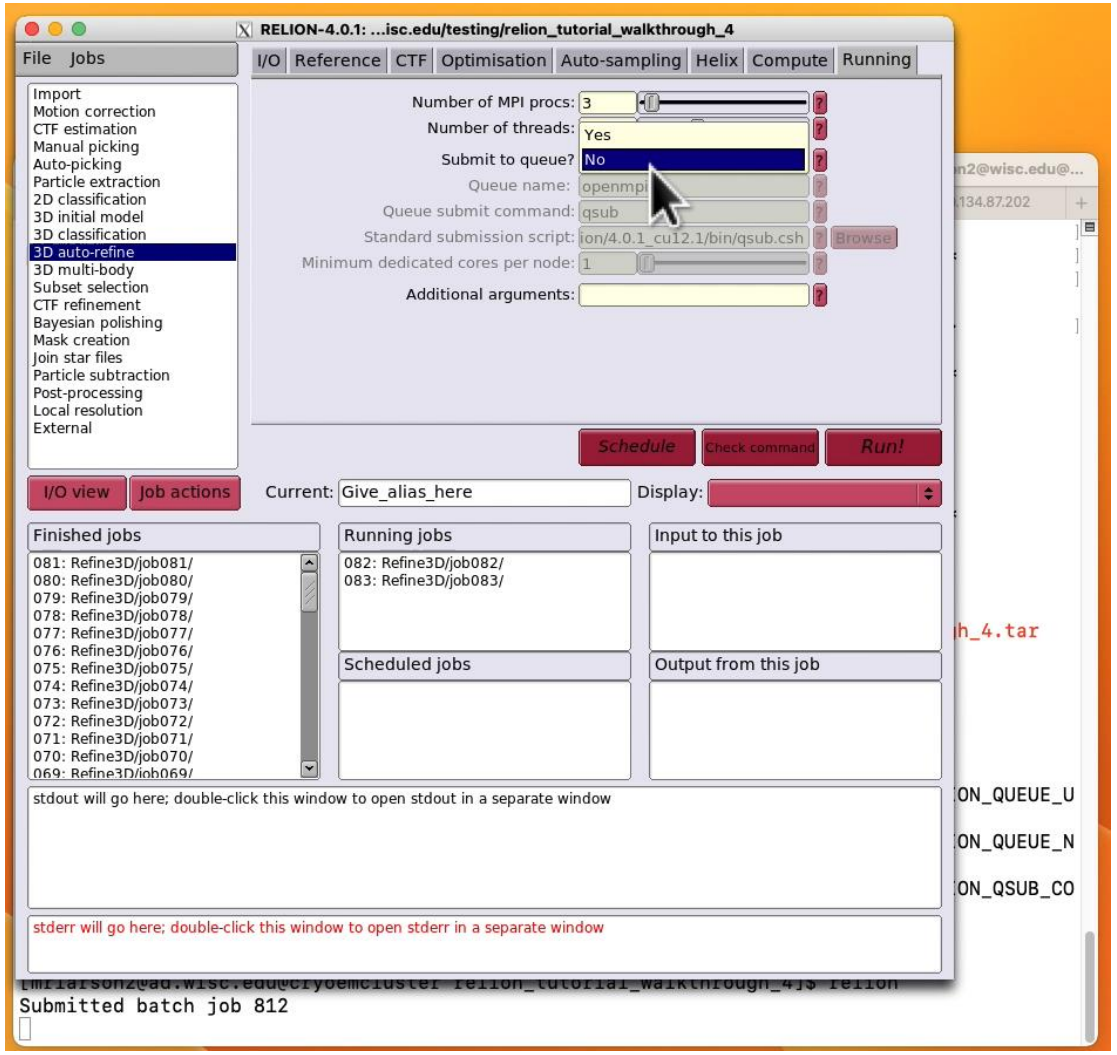
tail –f slurm_<jobid>.out

3dmod output/tomogram.mrc



Demo /w AreTomo

RELION -> SLURM



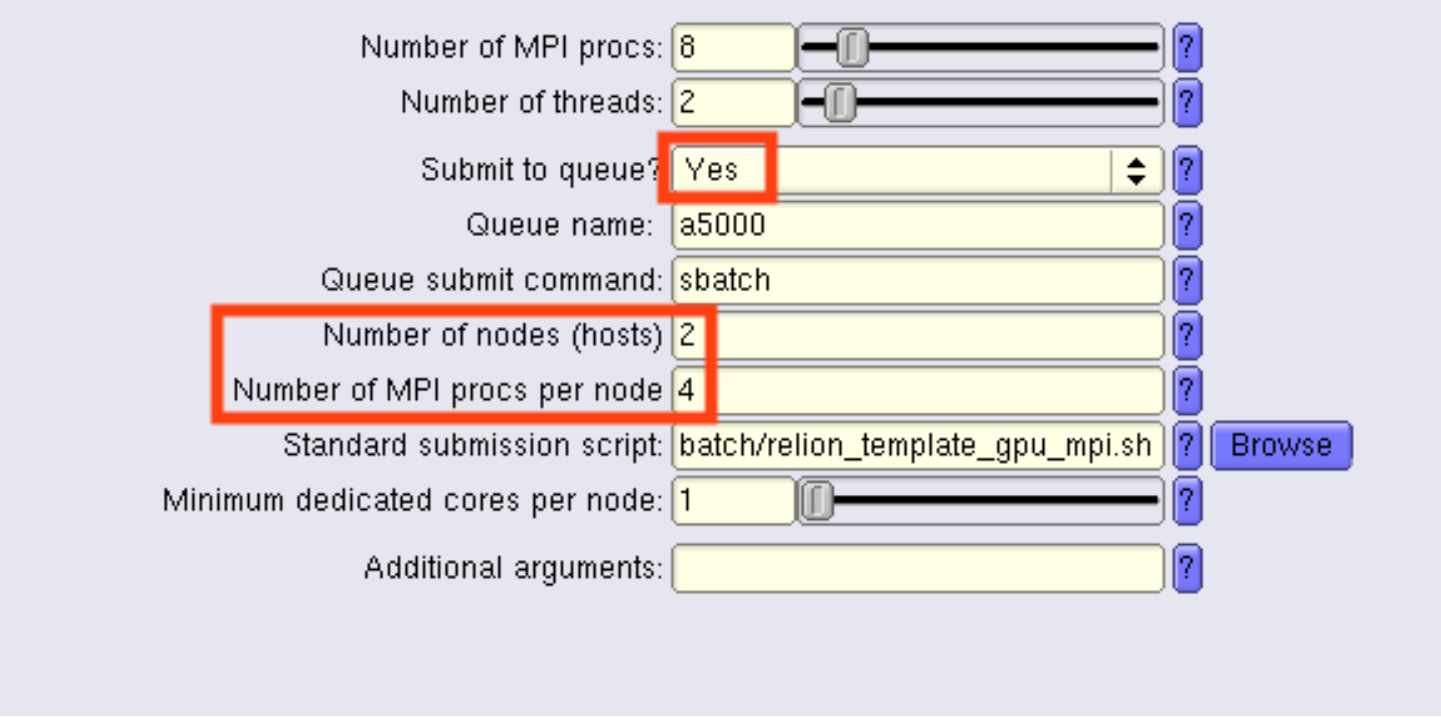
Enable "**Submit to queue**" to provide SLURM specific options command + queue.

Needs a "Standard submission script" :

/mnt/hpc_users/share/sbatch/relion_template_**gpu**.sh
or
/mnt/hpc_users/share/sbatch/relion_template_**cpu**.sh

RELION – using multiple nodes (MPI)

- Some jobs can be split among multiple nodes for faster processing
- Use the relion_template_gpu_mpi.sh or the relion_template_cpu_mpi.sh submission script
- Set number of nodes to < 8 for a5000 jobs. Set number of nodes to 2 for a100 jobs. Set number of nodes to < 4 for cpu jobs.



Command-line example: AlphaFold job

AlphaFold requires a FASTA input (T1083.fasta):

>T1083

GAMGSEIEHIEEAIANAKTKADHERLVAHYEEEAKRLEKKSEEYQELAKVYKKITDVYPNIRSYMVLHYQNLTRRYKE
AAEENRALAKLHHELAIVED

Submit via an Sbatch wrapper script:

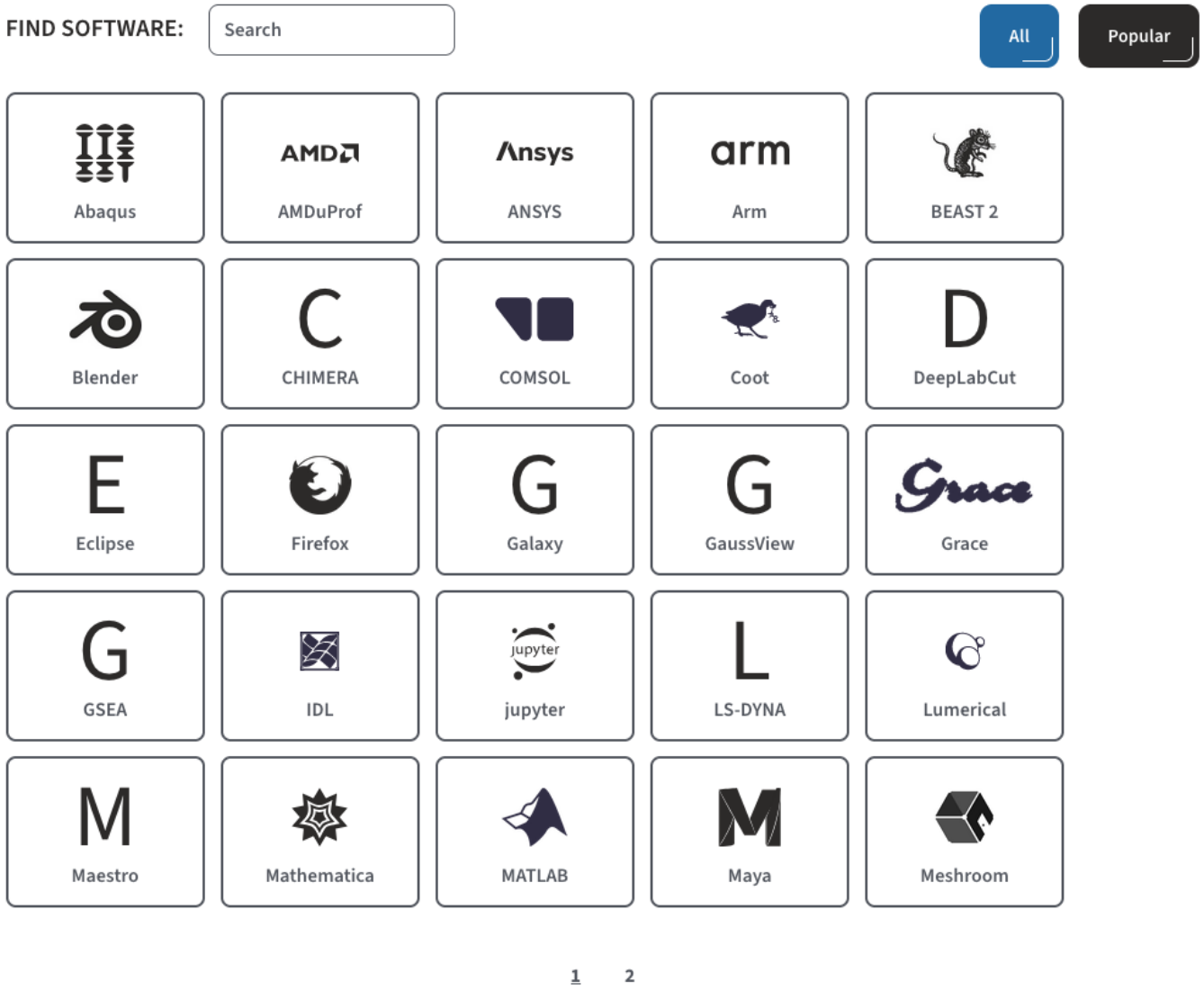
submit_to_a5000.sh /mnt/hpc_users/share/bin/alphafold-monomer.sh T1083.fasta

Policies

- Start jobs from the login node, but do not run significant processing on login node. Using significant resources on the login node will result in a warning and repeat offense will result in a locked account. Jobs running on the login node will be stopped.
- Large datasets should be removed from the cluster storage when jobs are complete. Cluster space will be audited periodically and users with old data will be contacted for transferring off storage.
- Each account is limited to 5TB quota in the shared network storage.

Other integrated OOD applications to consider

- https://openondemand.org/run-open-ondemand#enabled-applications
- Some of these other integrated applications might be useful to test on the HPC cluster.
- Applications available via SBGrid can always be used directly whether connecting via SSH or Open OnDemand.
- Some graphical applications may work better via the Open OnDemand Desktop (ex. Chimera, IMOD, and Napari).



Thank you!

Cryo Electron Microscopy Research Center
DEPARTMENT OF BIOCHEMISTRY
UNIVERSITY OF WISCONSIN–MADISON

Supported by the Department of Biochemistry at UW-Madison and Morgridge Institute for Research

- Director Elizabeth Wright
- CEMRC Staff and Wright Lab members
- NIH support for the National Centers for cryo-ET U24 GM139168-01
- Robert Landick, UW2020: WARF Discovery Initiative Award
- VCRGE Office for Research
- UW SMPH and CALS
- DoIT (networking, Globus, and hosting)

CryoEM

CEMRC Contacts: https://cryoem.wisc.edu cryoem@biochem.wisc.edu

- Elizabeth Wright (erwright2@wisc.edu)
- Jennifer Scheuren (jfscheuren@wisc.edu) - HPC Administrator
- Matt Larson (mrlarson2@wisc.edu) – CS and SysAdmin

Questions and Discussion

- What additional software should we support on the cluster? (Cryosparc, Matlab, Jupyter)
- What additional training would be useful? https://bcrf.biochem.wisc.edu/all-tutorials/
- Visit the UW-Madison Cryo-EM documentation for additional information and guidance (https://uwcryoem.github.io)