Introduction to NCAR HPC Resources

2024 CESM Tutorial

Rory Kelly
HPC Consultant

August 5, 2024

Getting Help

HPC Documentation -   https://ncar-hpc-docs.readthedocs.io

Getting Help

HPC Documentation -   https://ncar-hpc-docs.readthedocs.io

Getting Help

HPC Documentation -   https://ncar-hpc-docs.readthedocs.io

Getting Help

HPC Documentation -   https://ncar-hpc-docs.readthedocs.io

- Available systems and their uses
- Signing in and managing data
- Accessing software
- Managing jobs using Batch schedulers
- Customizing your user environment

This is only an introduction; for full documentation, see:

https://ncar -hpc-docs.readthedocs.io/en/latest/

- HPC Cray EX, 19.87 petaflops
- 2488 CPU Nodes
  - Dual socket, 64-core AMD 7763 "Milan" CPUs
  - 256 GB DDR4 Memory
  - 1 Cassini Slingshot-11 NIC
- 82 GPU Nodes
  - Single socket, 64-core AMD 7763 "Milan" CPUs
  - 512 GB DDR4 Memory
  - 4 NVIDIA A100 GPUs
  - NVLink GPU interconnect
 - 4 Cassini Slingshot-11 NICs
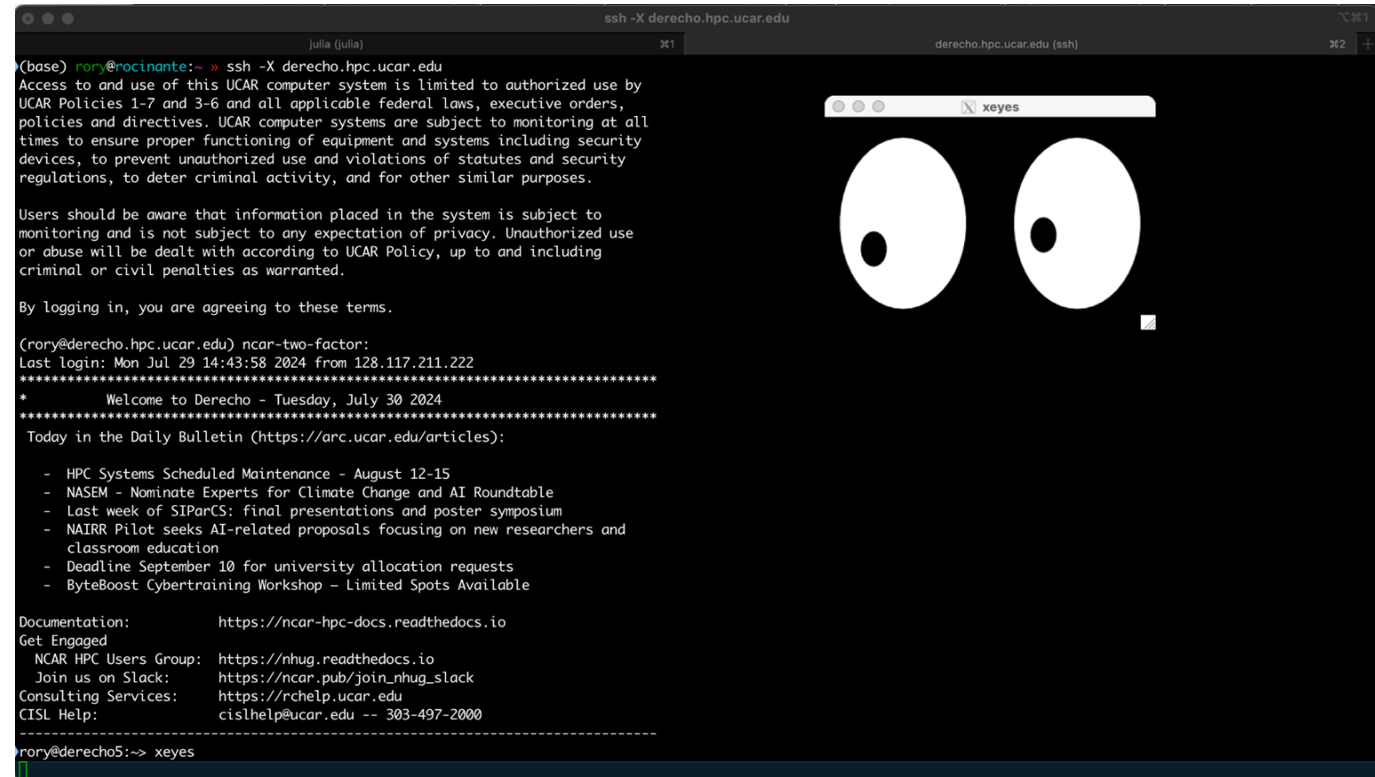
- 22 nodes for data analysis and visualization jobs. Each node has 36 cores and up to 384 GB memory.
  - 9 of these nodes also feature an NVIDIA Quadro GP100 GPU.
- 18 nodes feature large-memory, dense GPU configurations to support explorations in machine learning (ML) and GPU computing
  - 4 of these nodes feature 4 NVIDIA Tesla V100 GPUs
  - 6 of these nodes feature 8 NVIDIA Tesla V100 GPUs
  - 8 of these nodes feature 4 NVIDIA Ampere A100 GPUs
- 64 high-throughput computing (HTC) nodes for small computing tasks using 1 or 2 CPUs.
  - 62 HTC nodes have 384 GB of available memory
  - 2 HTC nodes have 1.5 TB of available memory
- 4 nodes are reserved for Research Data Archive workflows

- Use ssh along with your username to log in

  ssh -XY -l username derecho.hpc.ucar.edu

  ssh -XY -l username casper.hpc.ucar.edu

- Use Duo for authentication
- Derecho - 8 login nodes
- Casper - 2 login nodes

VNC can be used to run a remote GNOME/KDE desktop

Need to install a VNC client first - We recommend TigerVNC, but other VNC clients such as TurboVNC will also work

Usage:

`vncmgr create -A <project code>`

`vncmgr     (interactive)`
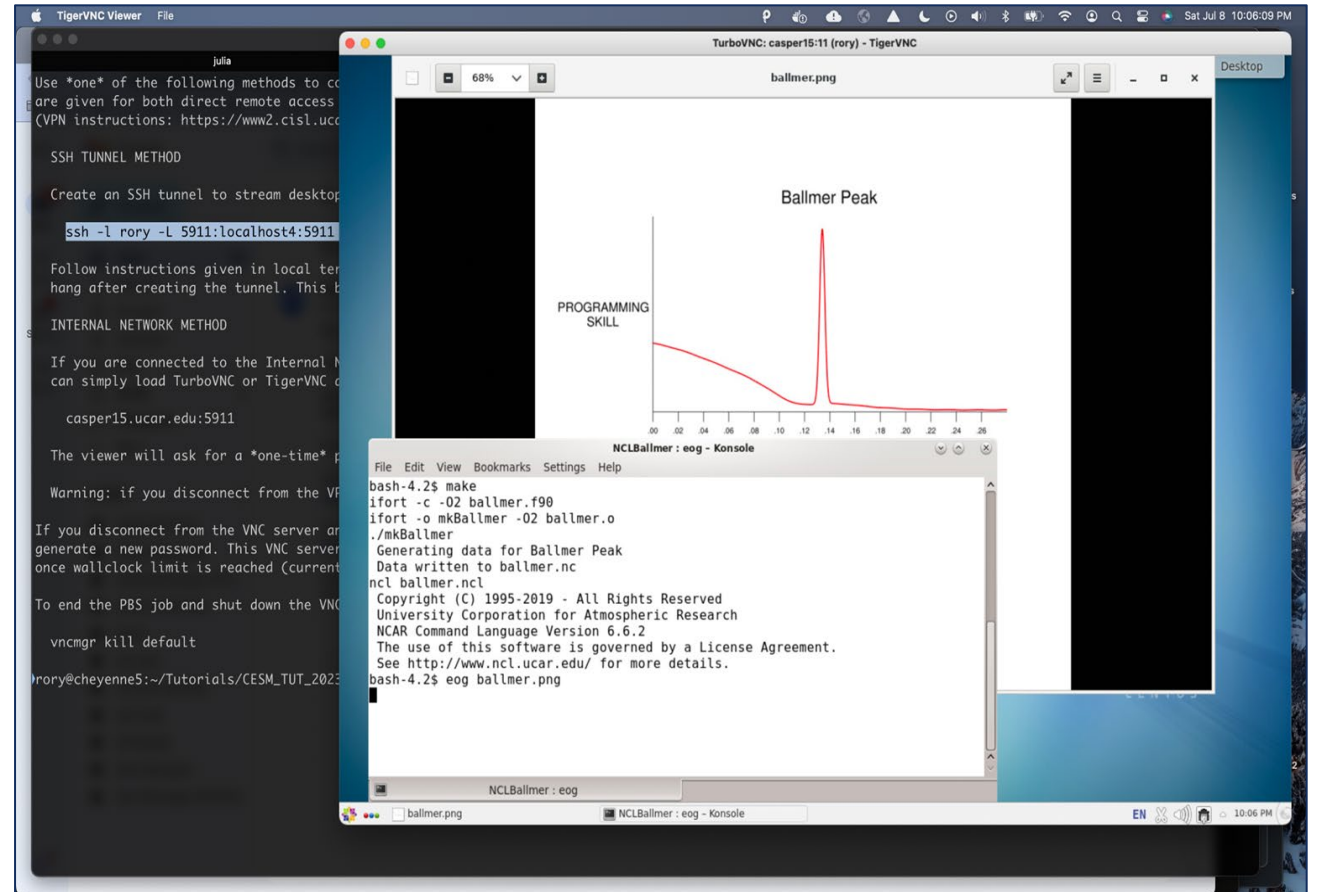
FastX can be used to run a remote desktop or terminal session in a browser without a client.

Connect to the NCAR VPN, then go to

https://fastx.ucar.edu:3300

Can also be setup using an SSH tunnel
- ssh -L 3300:fastx.ucar.edu:3300 username@fastx.ucar.edu (duo auth)
- Go to https://localhost:3300 in a browser
    ignore unsafe warnings
    duo auth again
- Open a KDE Desktop
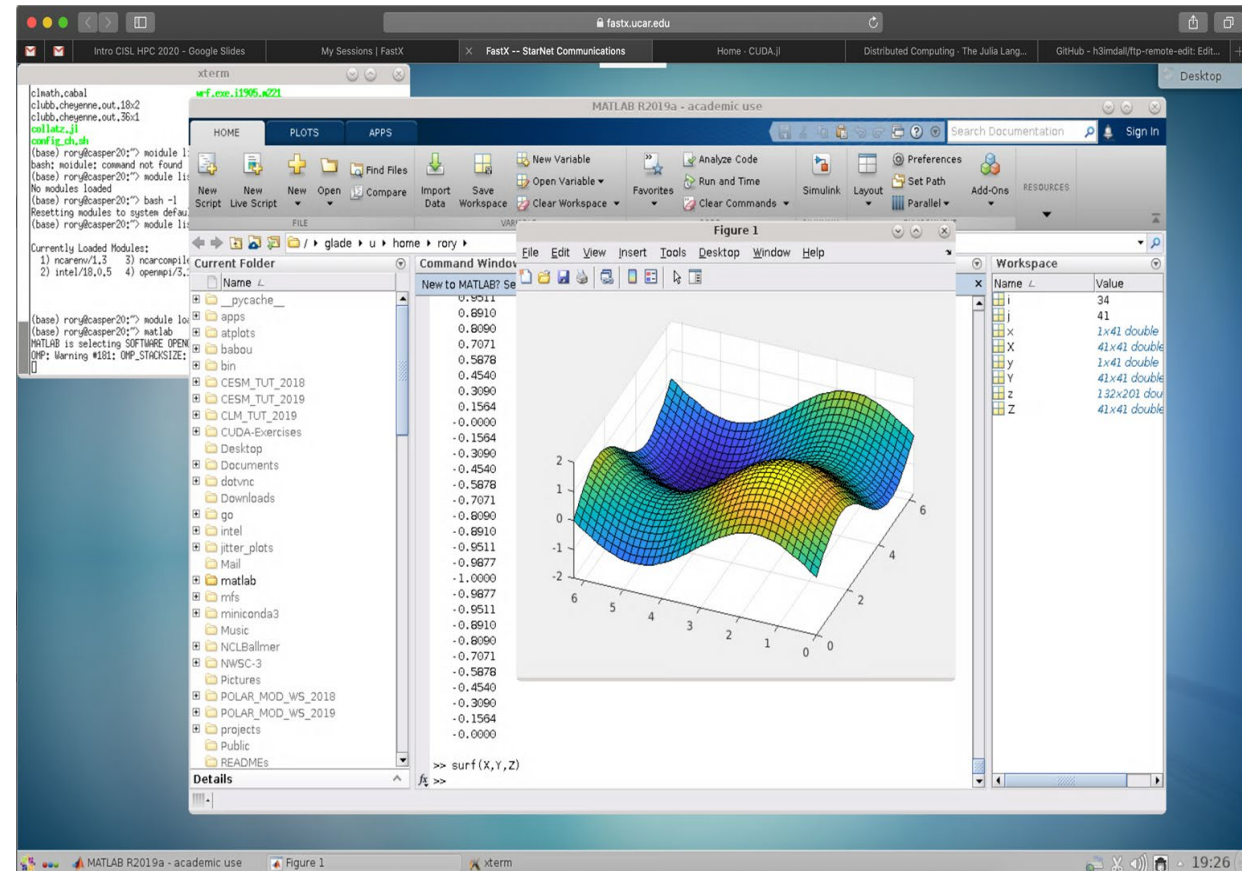
Derecho and Casper are also accessible via the JupyterHub service at
jupyterhub.hpc.ucar.edu
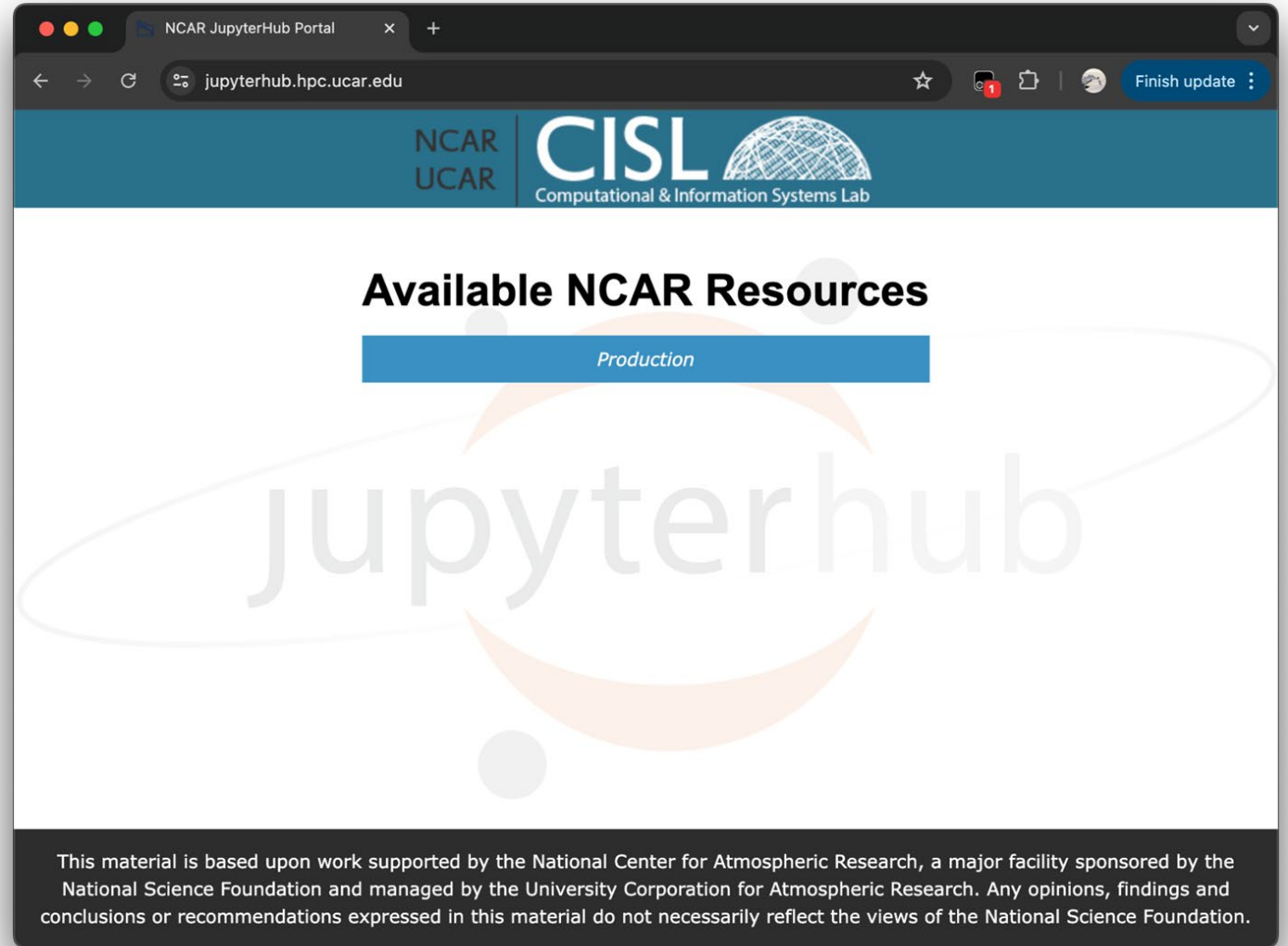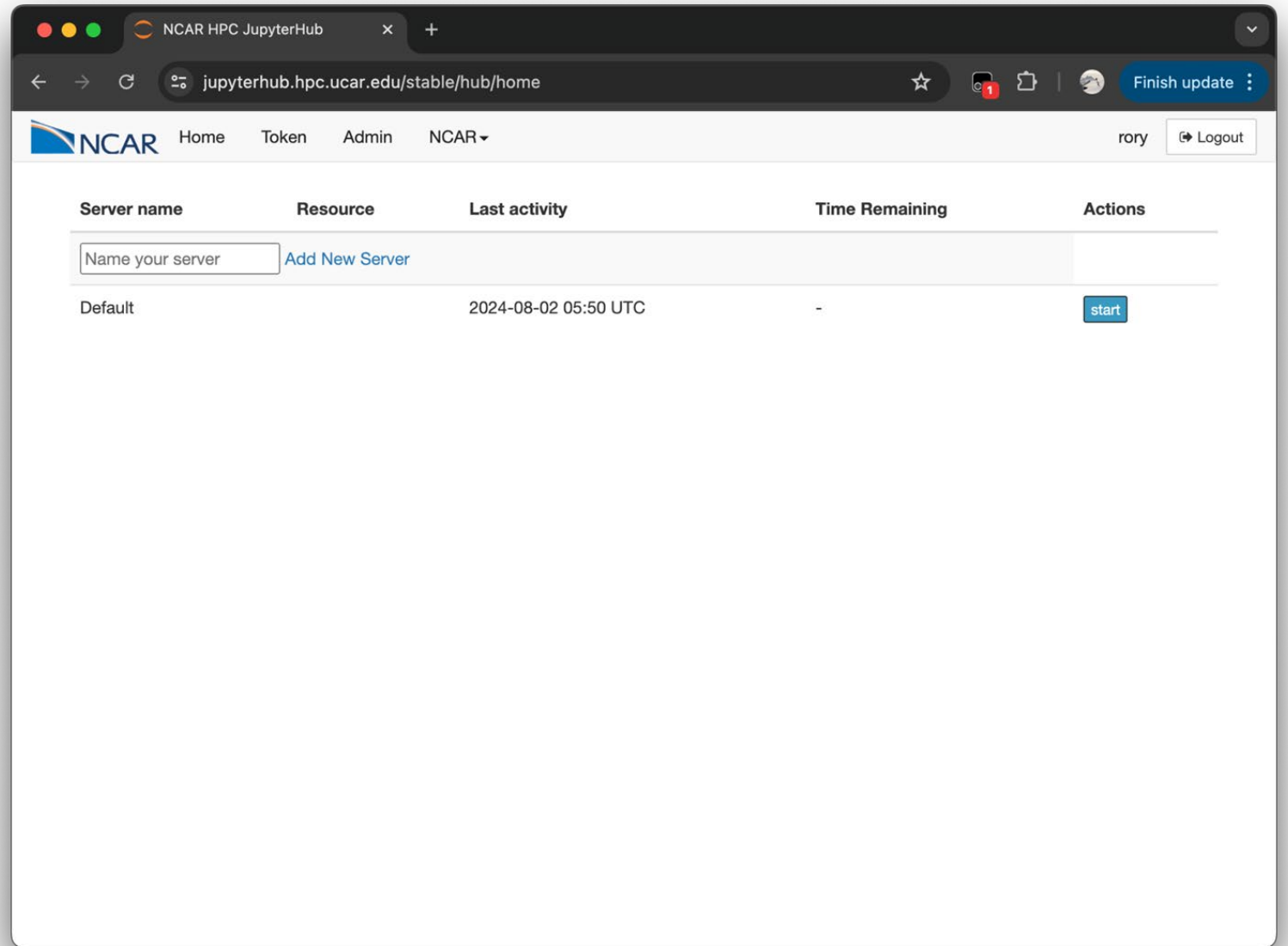
Log in with your standard Duo credentials and respond to the push notification
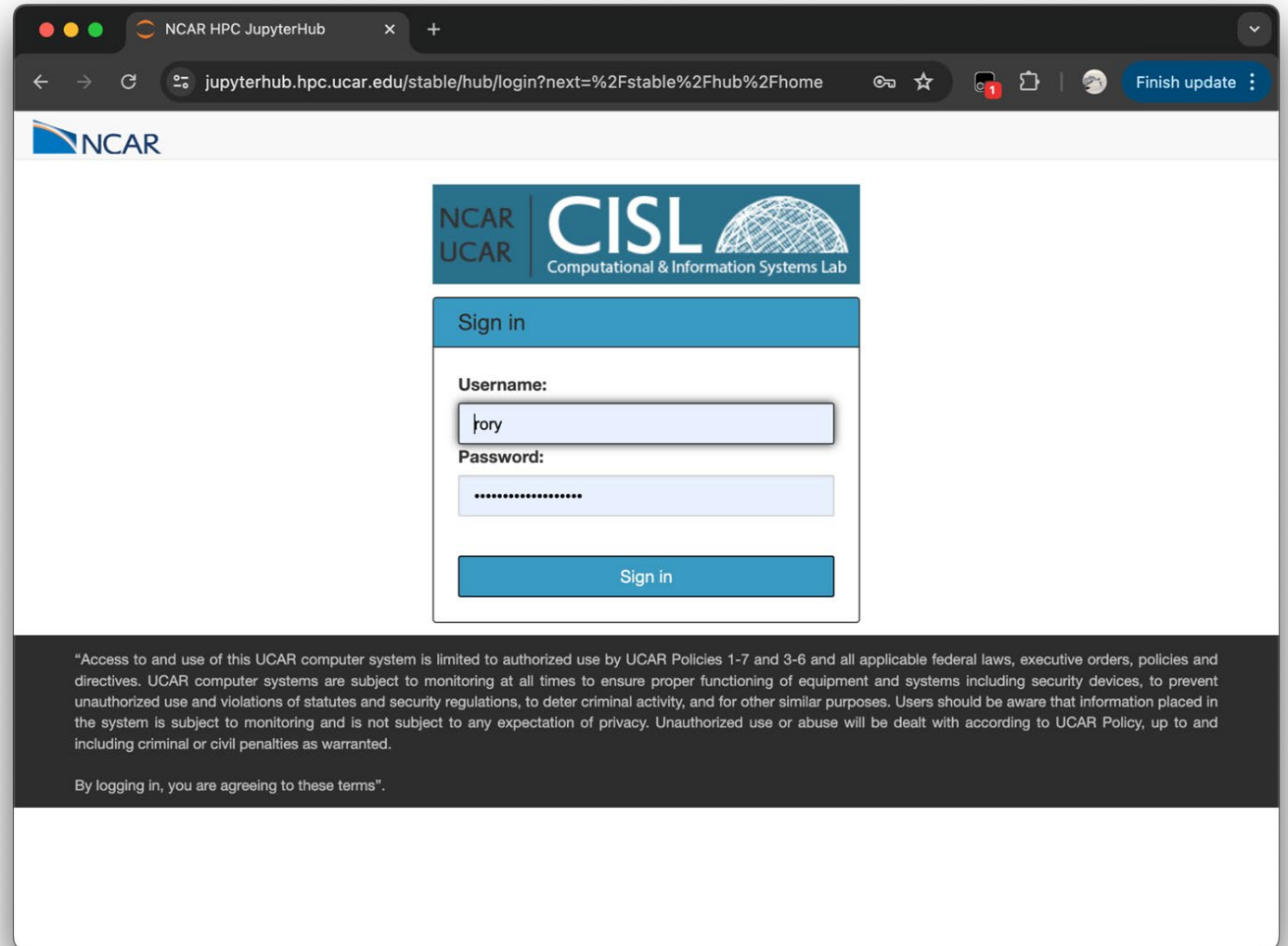
Start a server if you do not already have one running
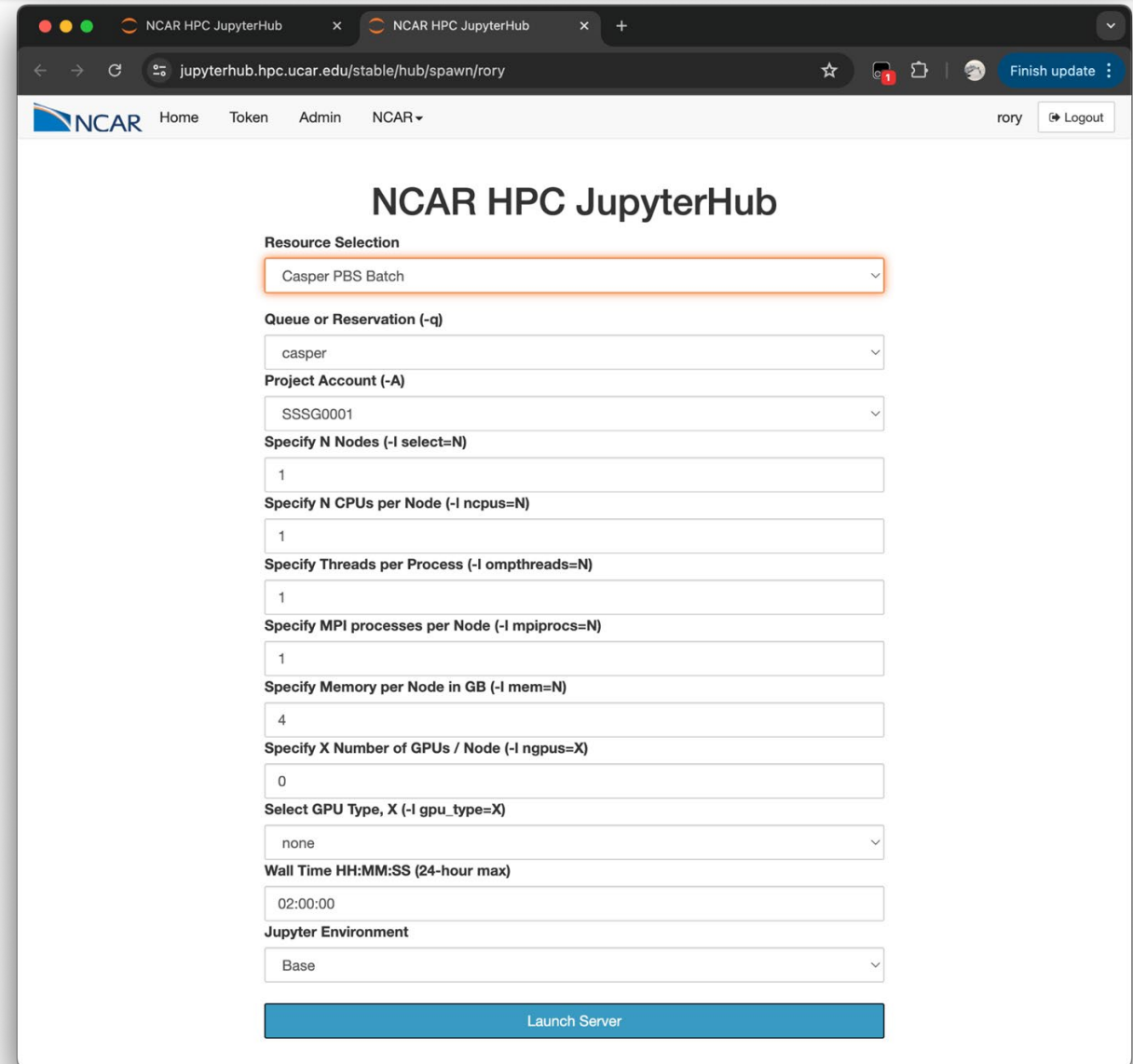
Select a Resource and fill in batch request details, including number of cores and project account

Many Notebooks and applications can be run from the main jupyterhub launcher screen.
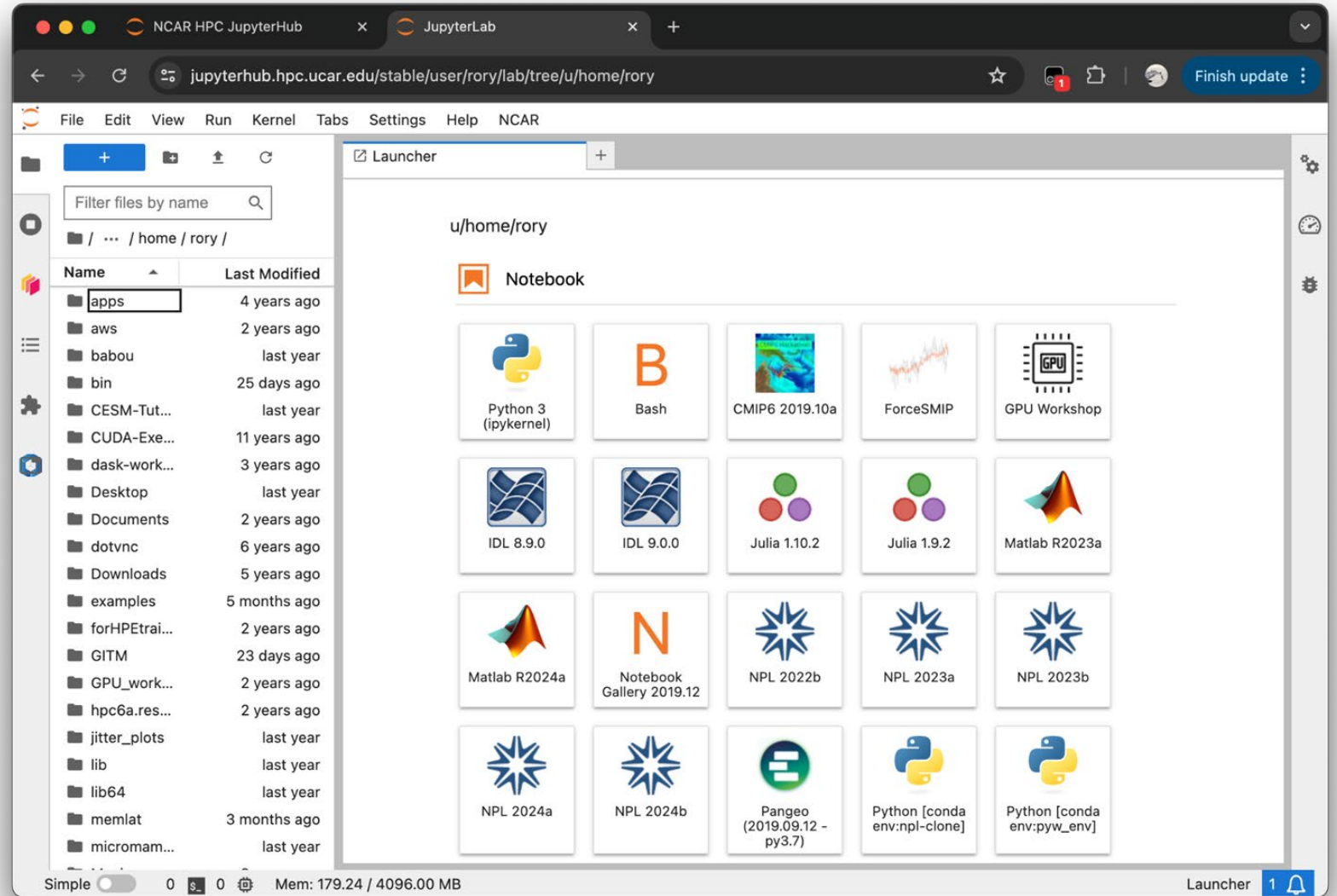
Jupyterhub is used to create sophisticated interactive computational notebooks for analysis, education, etc.

It can also be used for really unsophisticated things like viewing images
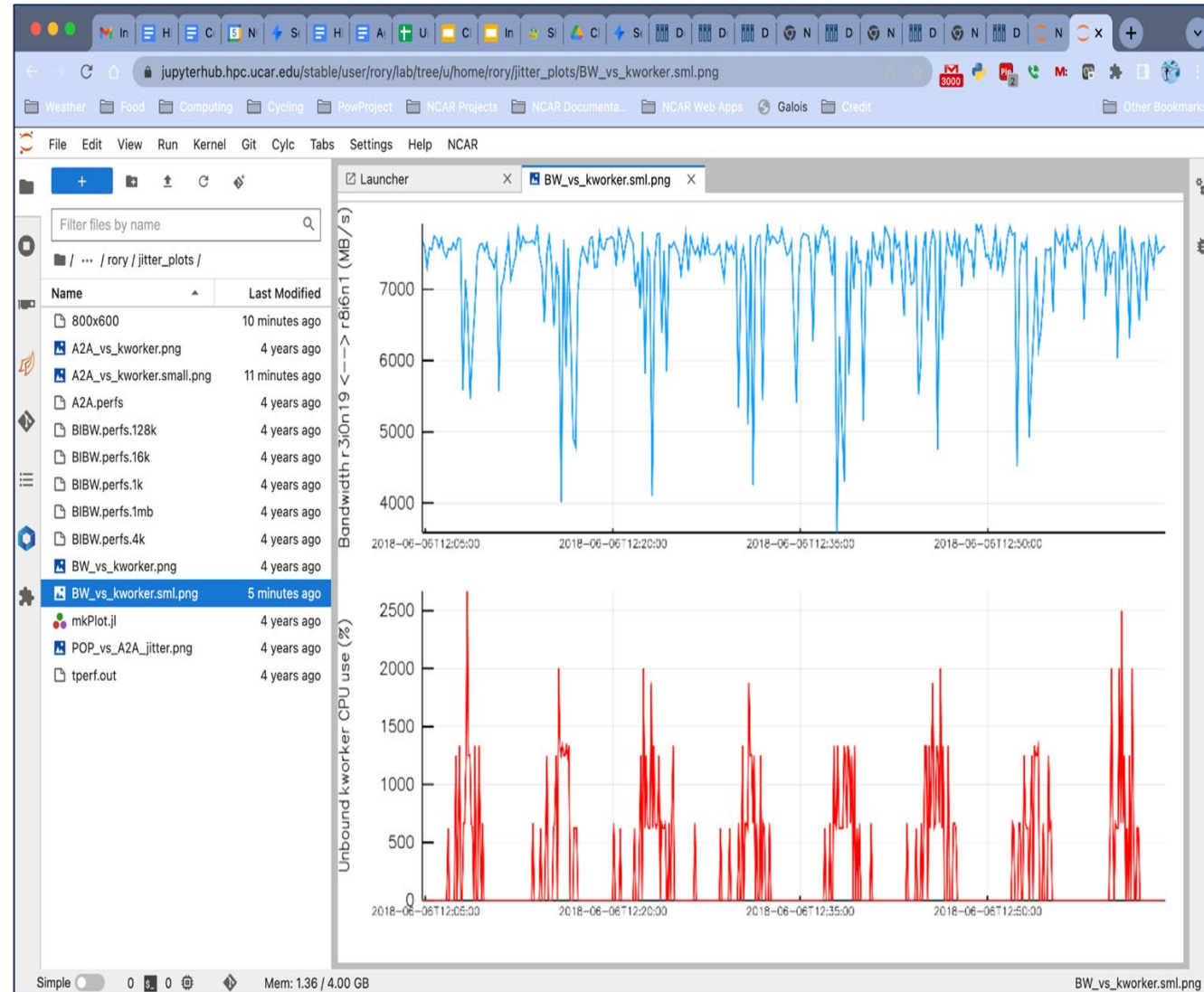
Be a Good Neighbor on Shared Resources

- Your activities coexists with those of other users
- CPUs and memory are shared on the login nodes
- Limit your usage to:
  - Reading and editing text/code
  - Compiling small programs
 - Performing data transfers
  - Interacting with the job scheduler
- Programs that use excessive resources on the login nodes will be limited (you will receive automated warnings via email if you exceed usage thresholds)
- In extreme cases, your processes may be terminated

Data Storage - GLADE

File spaces accessible from all HPC systems

File space	Quota	Backup	Uses
Home **/glade/u/home/$USER**	50 GB	Yes	Settings, code, scripts
Work **/glade/work/$USER**	2 TB	No	Compiled codes, models
Scratch **/glade/derecho/scratch/$USER**	30 TB	**Purged!**	Run directories, temp output
Campaign **/glade/campaign**	N/A	No	Project space allocations

*Keep track of usage with "**gladequota**"*

- For short transfers use **scp/sftp** to transfer files
- For large transfers use **Globus**
  - To use Globus, create a Globus ID if you need an account, and search for **NCAR GLADE** or **NCAR Campaign Storage** endpoints
  - CISL endpoints currently can be activated for up to 30 -days
  - Globus has a web interface and a command -line interface
  - **Globus Connect Personal** can manage transfers from your local workstation as well

- CISL installed software is provided as modules
- Modules provide access to runnable applications (compilers, debuggers, ...) as well as libraries (NetCDF, MPI, ...)
- Modules help ensure that all the software in your environment is mutually compatible
- Note that Derecho and Casper each have independent collections of modules due to differences in their base software stacks. We make an effort to keep the stacks similar, but it is not possible in all cases.

Using Modules

- **module load/unload <software>**
- **module avail**
  - show all currently-loadable modules
- **module list**
  - show loaded modules
- **module (--force) purge**
  - remove all loaded modules
- **module save/restore   <name>**
  - create/load a saved software set
- **module spider <software>**
  - search for a particular module

ncarenv/23.09

intel-oneapi
2024.0.2

mkl

netcdf

Compiler

intel-oneapi
2024.0.2

cce 17.0.1

gcc
13.2.0

MPI

cray-mpich
8.1.29

mvapich
3.0

OpenMPI
5.0.0

intel-oneapi
2024.0.2
cray-mpich
8.1.29
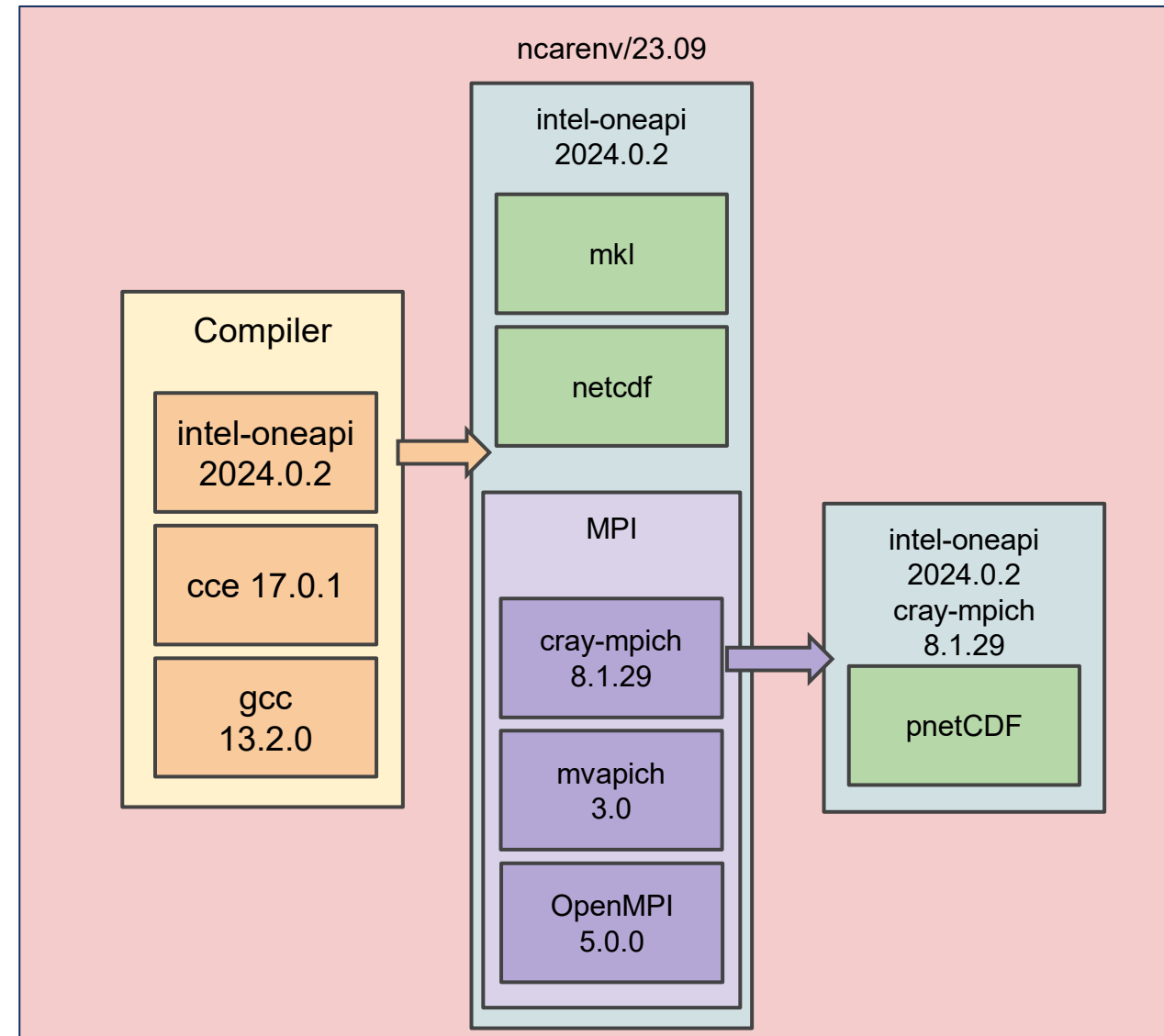
pnetCDF

Available Software

- Compilers (Intel, CCE, GCC, NVHPC)
- Debuggers / Performance Tools (Linaro Forge, DDT, MAP, CrayPAT)
- MPI Libraries (Cray-MPICH, MVAPICH*, OpenMPI*, IntelMPI*)
- IO Libraries (NetCDF, PNetCDF, HDF5, Parallel IO)
- Analysis Languages (Python, Julia, R, IDL, Matlab)
- Convenience Tools (ncarcompilers, parallel, rclone)
- Many more ...

Using Batch Jobs and Compute Nodes

- Most tasks require too many resources to run on a login node
- Schedule these tasks to run on compute nodes using the **PBS** batch system

ssh derecho.hpc.ucar.edu

Workstation → 8 login nodes → **2570** compute nodes

8 login nodes → **108+** data analysis and visualization, nodes

Workstation → 2 login nodes → **108+** data analysis and visualization, nodes

2570 compute nodes → **108+** data analysis and visualization, nodes

ssh casper.hpc.ucar.edu
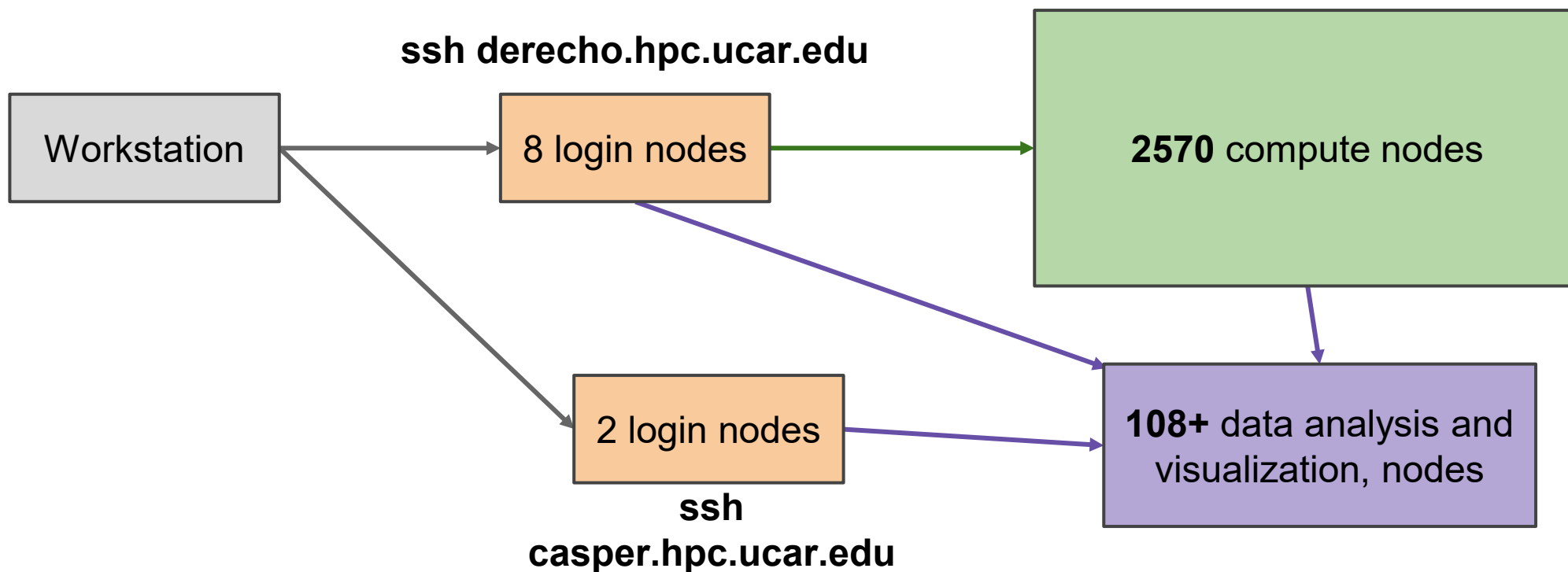
Derecho and Casper use separate allocations!

- Jobs request a given amount of compute resources, for an estimated amount of time, on specific hardware
- Jobs use core-hours, which are charged against your selected project/account
  - Remaining resources are viewable in SAM (sam.ucar.edu)
- Temporary files are often written by programs - set TMPDIR variable to scratch space to avoid job failures

Example PBS Batch Script

```
$ cat basic_mpi.pbs
#!/bin/bash
#PBS -N hello_pbs
#PBS -A <project_code>
#PBS -j oe
#PBS -k eod
#PBS -q main
#PBS -l walltime=00:05:00
#PBS -l select=8:ncpus=128:mpiprocs=128


### Set temp to scratch
export TMPDIR=/glade/derecho/scratch/${USER}/tmp
mkdir -p $TMPDIR


module --force purge
module load ncarenv/23.09 intel-oneapi/2024.0.2 ncarcompilers/1.0.0 cray-mpich/8.1.27

### Run MPT MPI Program
mpibind ./hello_world
```

qsub <script> - submit batch job

qstat <jobid> - query job status

qdel <jobid> - delete/kill a job

qinteractive   -A <project>

    Run an interactive job

qcmd -A <project> -- <command>

    Run command on a compute node

Customize Your Environment with Default Modules

- If you commonly load certain modules, you may wish to have them load automatically when logging onto a cluster
- The right way to do so is with saved module sets:

```
module  --force purge
module load ncarenv/23.09 intel    -oneapi/2024.0.2
module load ncarcompilers/1.0.0 cray    -mpich/8.1.27
module load conda/latest
module save default
```

- Make multiple sets and load them using module restore <set>
- Don't put module load commands in your shell startup files!

Getting Help

HPC Documentation -   https://ncar-hpc-docs.readthedocs.io