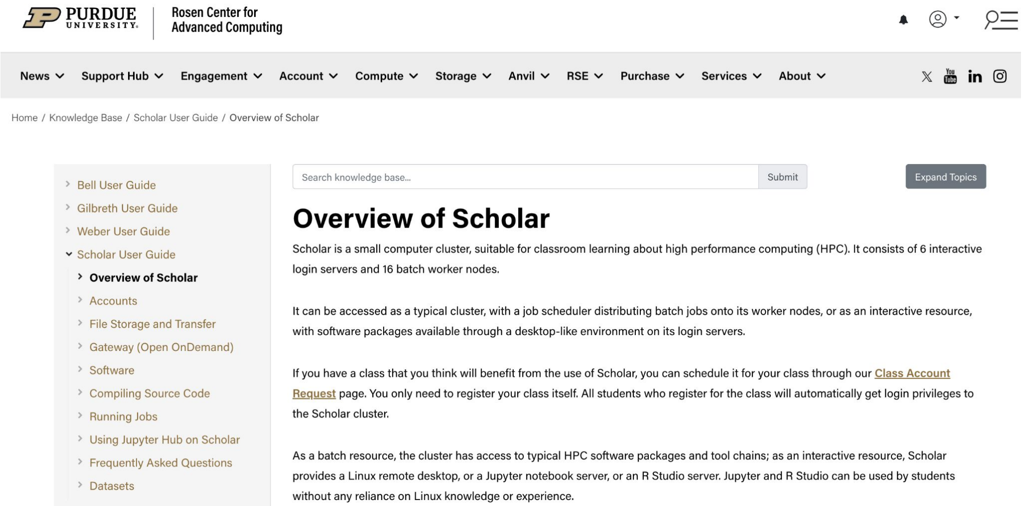
RCAC Tutorial

Guven Gergerli

RCAC Scholar

- Groups of computer clusters for research at Purdue
- Job scheduling or remote desktop options
- There are different Community Clusters but we will use Scholar
    - classroom learning about high performance computing (HPC).
- For anything that is not mentioned, there is a detailed guide on how to use Scholar.



Overview of Scholar

Scholar is a small computer cluster, suitable for classroom learning about high performance computing (HPC). It consists of 6 interactive login servers and 16 batch worker nodes.

It can be accessed as a typical cluster, with a job scheduler distributing batch jobs onto its worker nodes, or as an interactive resource, with software packages available through a desktop-like environment on its login servers.

If you have a class that you think will benefit from the use of Scholar, you can schedule it for your class through our Class Account Request page. You only need to register your class itself. All students who register for the class will automatically get login privileges to the Scholar cluster.

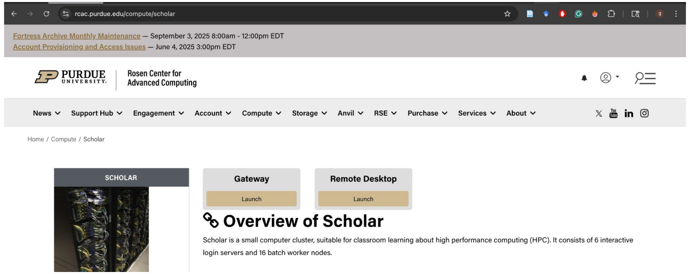
As a batch resource, the cluster has access to typical HPC software packages and tool chains; as an interactive resource, Scholar provides a Linux remote desktop, or a Jupyter notebook server, or an R Studio server. Jupyter and R Studio can be used by students without any reliance on Linux knowledge or experience.

How to use Scholar for GPU training

1. Gateway

rcac.purdue.edu/compute/scholar

- Many functionalities with browser access
- This is what you will see:



rcac.purdue.edu/compute/scholar

Fortress Archive Monthly Maintenance — September 3, 2025 8:00am - 12:00pm EDT
Account Provisioning and Access Issues — June 4, 2025 3:00pm EDT

Overview of Scholar

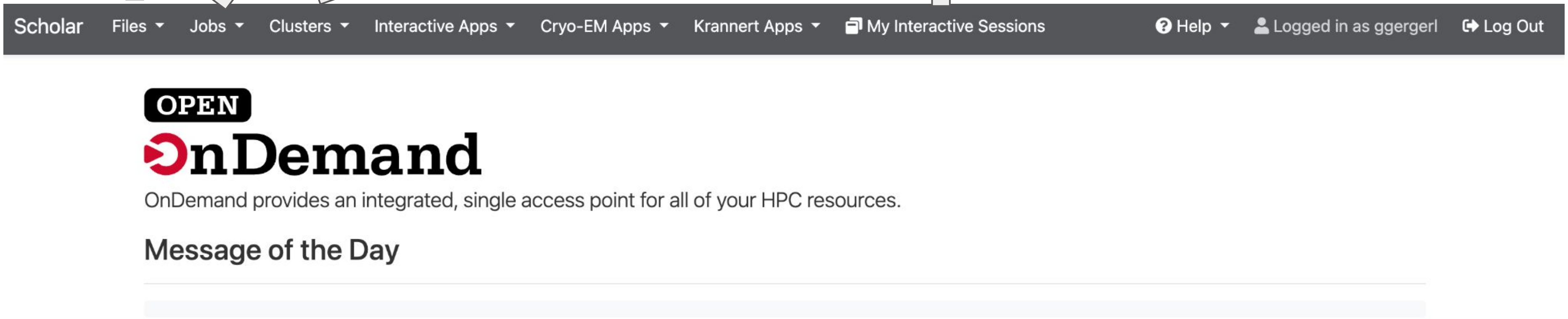
Scholar is a small computer cluster, suitable for classroom learning about high performance computing (HPC). It consists of 6 interactive login servers and 16 batch worker nodes.

To check your jobs

Directories

Remote Desktop

Shell

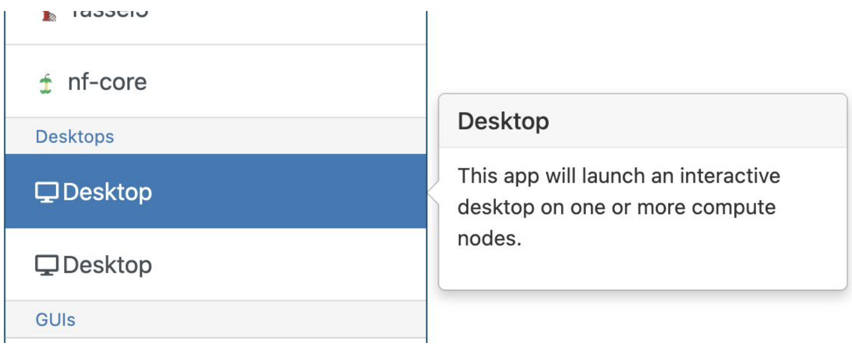


OnDemand provides an integrated, single access point for all of your HPC resources.

Message of the Day

How to use Scholar for GPU training
2. Remote Desktop

- Best to initially see what is going on inside the cluster



nf-core

Desktops

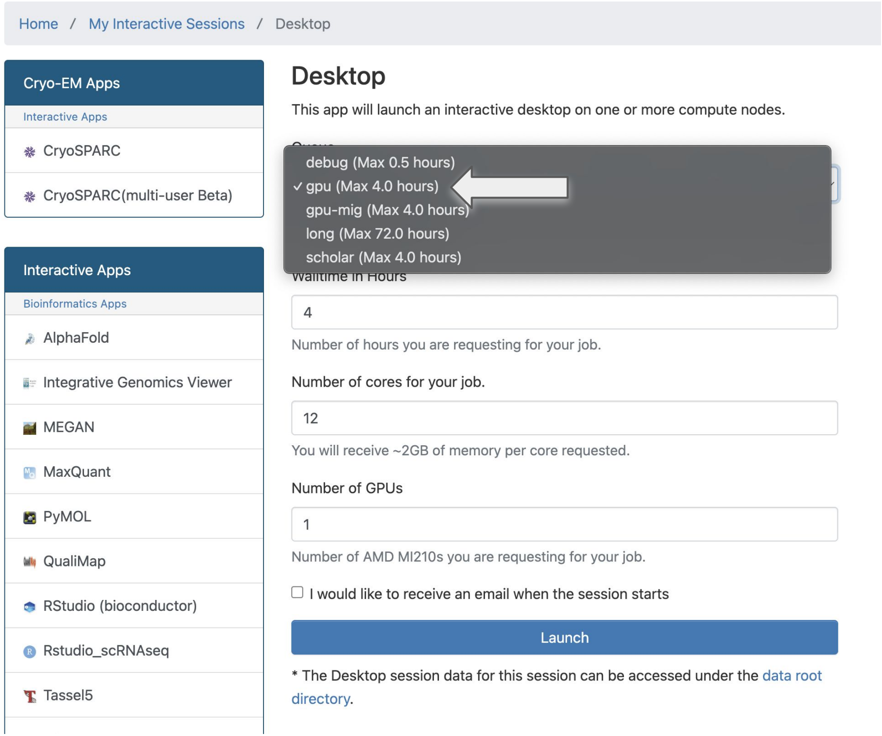
Desktop

Desktop

GUIs

Desktop

This app will launch an interactive desktop on one or more compute nodes.



Home / My Interactive Sessions / Desktop

Cryo-EM Apps

Interactive Apps

CryoSPARC

CryoSPARC(multi-user Beta)

Interactive Apps

Bioinformatics Apps

AlphaFold

Integrative Genomics Viewer

MEGAN

MaxQuant

PyMOL

QualiMap

RStudio (bioconductor)

Rstudio_scRNAseq

Tassel5

Desktop

This app will launch an interactive desktop on one or more compute nodes.

debug (Max 0.5 hours)
✓ gpu (Max 4.0 hours)
gpu-mig (Max 4.0 hours)
long (Max 72.0 hours)
scholar (Max 4.0 hours)

Walltime in Hours

4

Number of hours you are requesting for your job.

Number of cores for your job.

12

You will receive ~2GB of memory per core requested.

Number of GPUs

1

Number of AMD MI210s you are requesting for your job.

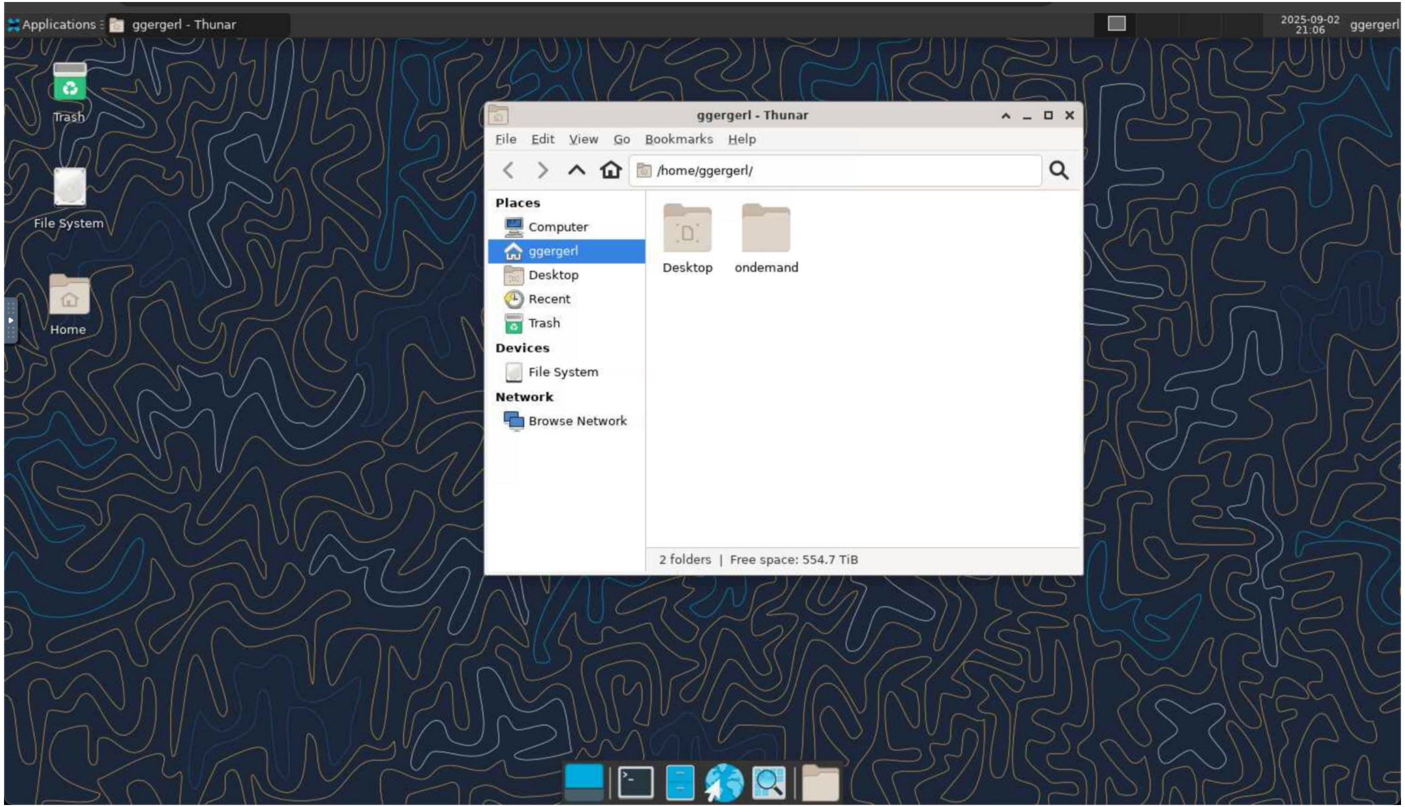
□ I would like to receive an email when the session starts

Launch

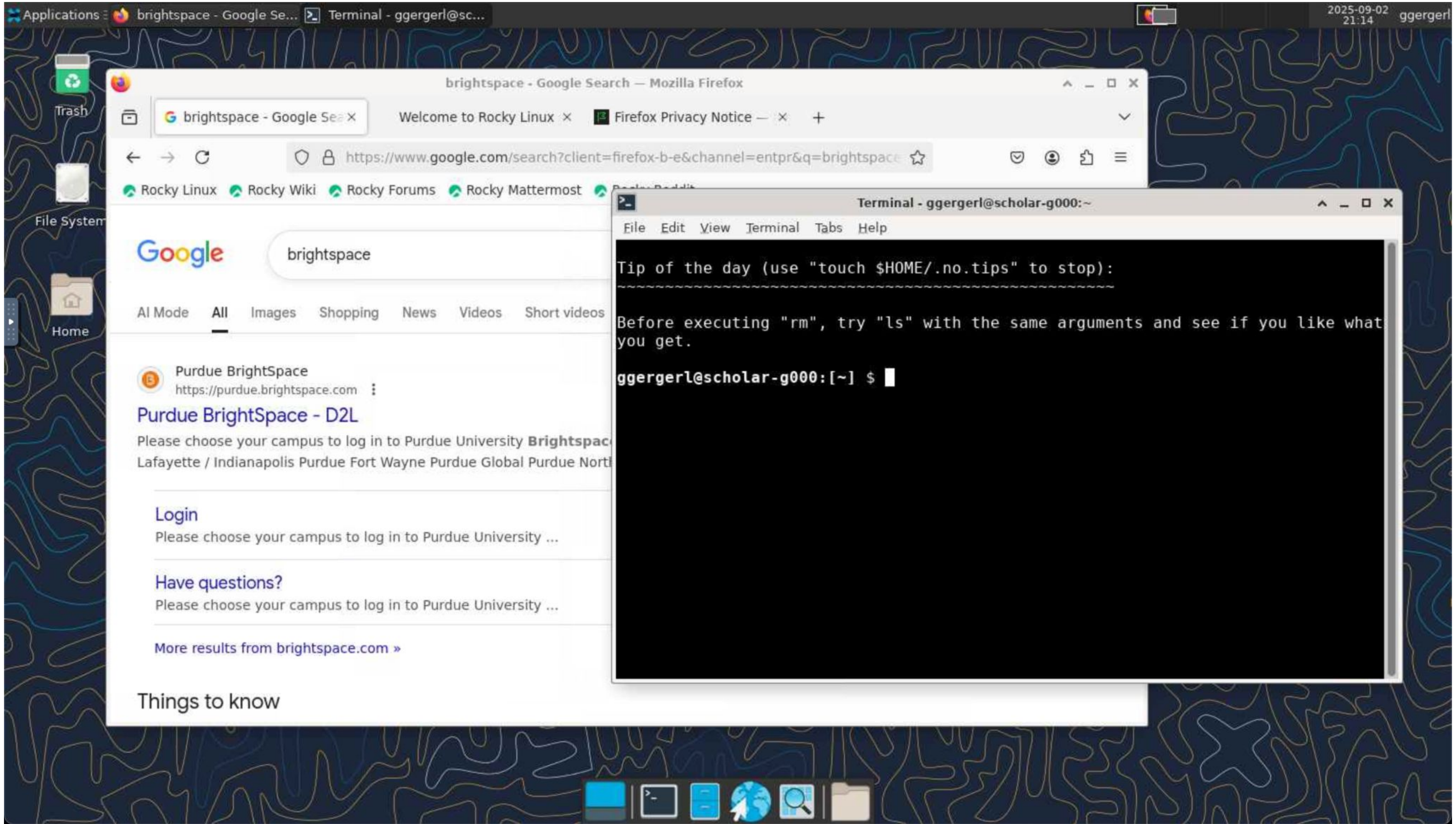
* The Desktop session data for this session can be accessed under the data root directory.

- This configuration if you would like to train inside the remote desktop

How to use Scholar for GPU training
2. Remote Desktop



How to use Scholar for GPU training
2. Remote Desktop



Purdue BrightSpace
https://purdue.brightspace.com
Purdue BrightSpace - D2L
Please choose your campus to log in to Purdue University **Brightspac**
Lafayette / Indianapolis Purdue Fort Wayne Purdue Global Purdue Nort

Login
Please choose your campus to log in to Purdue University ...

Have questions?
Please choose your campus to log in to Purdue University ...

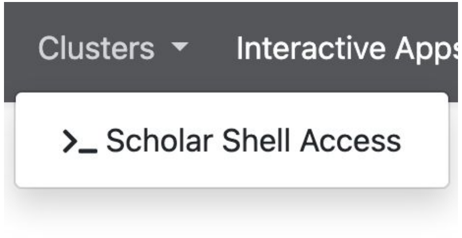
More results from brightspace.com »

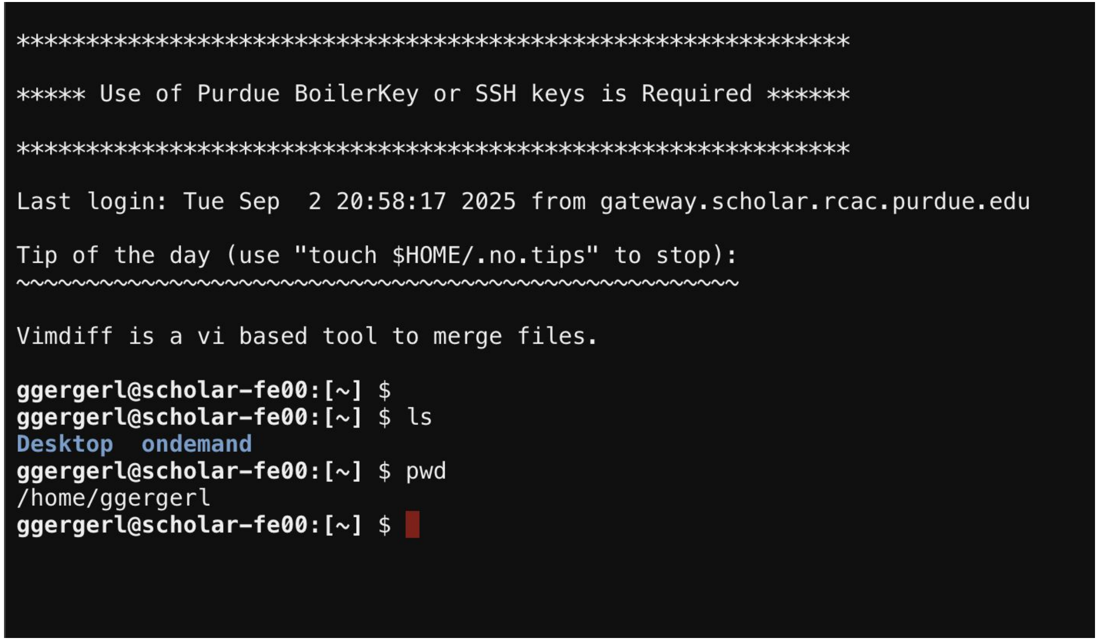
Things to know

```
Tip of the day (use "touch $HOME/.no.tips" to stop):
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Before executing "rm", try "ls" with the same arguments and see if you like what
you get.

ggergerl@scholar-g000:[~] $
```

How to use Scholar for GPU training
3. Shell

- When you need to access the terminal through the browser (without having to start remote desktop)

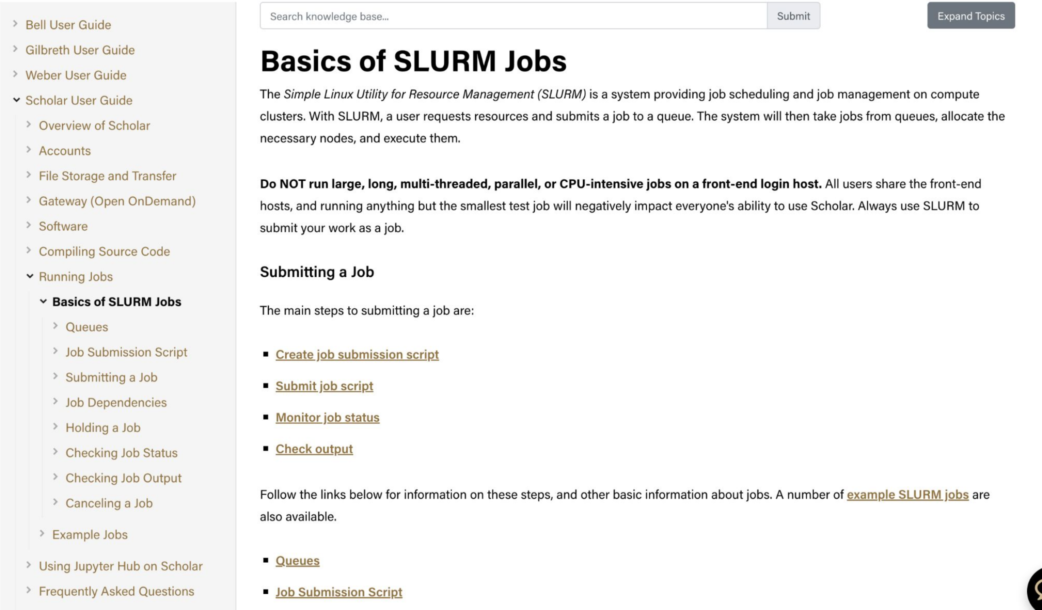




```
************************************************************

***** Use of Purdue BoilerKey or SSH keys is Required ******

************************************************************

Last login: Tue Sep  2 20:58:17 2025 from gateway.scholar.rcac.purdue.edu

Tip of the day (use "touch $HOME/.no.tips" to stop):
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Vimdiff is a vi based tool to merge files.

ggergerl@scholar-fe00:[~] $
ggergerl@scholar-fe00:[~] $ ls
Desktop  ondemand
ggergerl@scholar-fe00:[~] $ pwd
/home/ggergerl
ggergerl@scholar-fe00:[~] $
```

How to use Scholar for GPU training
4. Running jobs

- Clusters can be initiated with a SLURM job
  - SLURM job: a configuration for clusters to run a script with dedicated hardware
- Remote Desktop is also a job but with a desktop application
- For more detail please check RCAC tutorial as well ->



Basics of SLURM Jobs

The *Simple Linux Utility for Resource Management (SLURM)* is a system providing job scheduling and job management on compute clusters. With SLURM, a user requests resources and submits a job to a queue. The system will then take jobs from queues, allocate the necessary nodes, and execute them.

Do NOT run large, long, multi-threaded, parallel, or CPU-intensive jobs on a front-end login host. All users share the front-end hosts, and running anything but the smallest test job will negatively impact everyone's ability to use Scholar. Always use SLURM to submit your work as a job.

Submitting a Job

The main steps to submitting a job are:

- Create job submission script
- Submit job script
- Monitor job status
- Check output

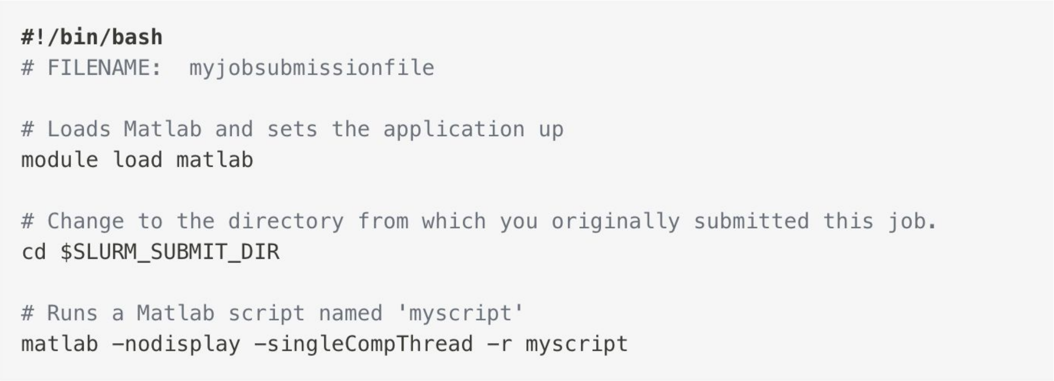
Follow the links below for information on these steps, and other basic information about jobs. A number of example SLURM jobs are also available.

- Queues
- Job Submission Script

How to use Scholar for GPU training
4. Running jobs

- You can run jobs with a job submission file, this is what it looks like:

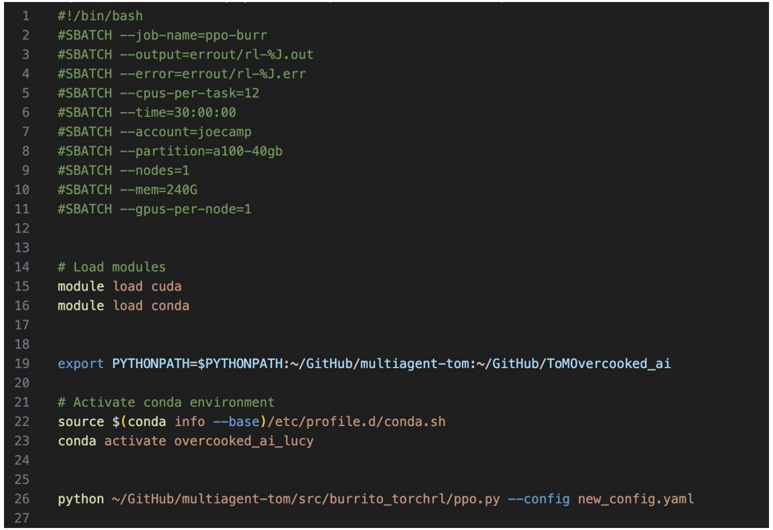
An example from RCAC



```
#!/bin/bash
# FILENAME:  myjobsubmissionfile

# Loads Matlab and sets the application up
module load matlab

# Change to the directory from which you originally submitted this job.
cd $SLURM_SUBMIT_DIR

# Runs a Matlab script named 'myscript'
matlab -nodisplay -singleCompThread -r myscript
```

- You can specify within scholars own limitations.
- The files should be ``.sub`` and ran through ``sbatch file_name.sub``

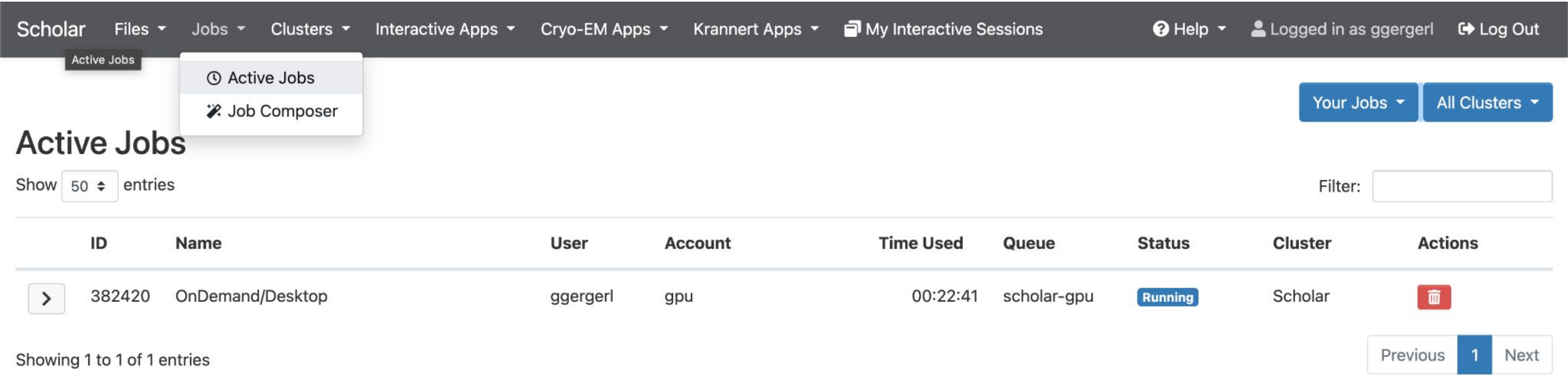
One of my customized job submission file (for another task on a different cluster)



```
#!/bin/bash
#SBATCH --job-name=ppo-burr
#SBATCH --output=errout/rl-%J.out
#SBATCH --error=errout/rl-%J.err
#SBATCH --cpus-per-task=12
#SBATCH --time=30:00:00
#SBATCH --account=joecamp
#SBATCH --partition=a100-40gb
#SBATCH --nodes=1
#SBATCH --mem=240G
#SBATCH --gpus-per-node=1


# Load modules
module load cuda
module load conda


export PYTHONPATH=$PYTHONPATH:~/GitHub/multiagent-tom:~/GitHub/ToMOvercooked_ai

# Activate conda environment
source $(conda info --base)/etc/profile.d/conda.sh
conda activate overcooked_ai_lucy


python ~/GitHub/multiagent-tom/src/burrito_torchrl/ppo.py --config new_config.yaml

```

How to use Scholar for GPU training
4. Running jobs

- A job that is on the process can be seen in the active jobs in gateway



Active Jobs

Show 50 entries

Filter:

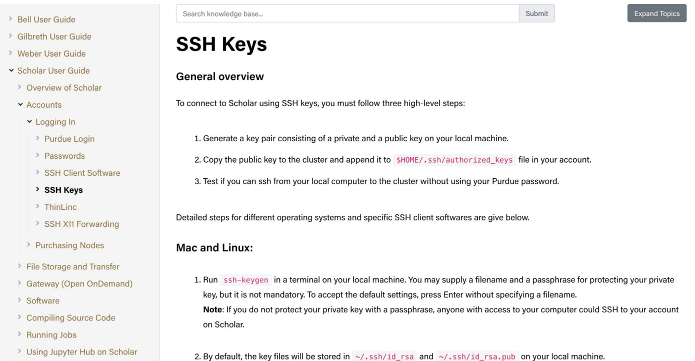
	ID	Name	User	Account	Time Used	Queue	Status	Cluster	Actions
>	382420	OnDemand/Desktop	ggergerl	gpu	00:22:41	scholar-gpu	Running	Scholar	🗑

Showing 1 to 1 of 1 entries

Previous 1 Next

How to use Scholar for GPU training
Bonus: VSCode

- You can also use VSCode for directly playing around with your code inside the cluster
- For this you will need to configure an SSH Host
  - For this you should check the tutorial on RCAC



SSH Keys

General overview

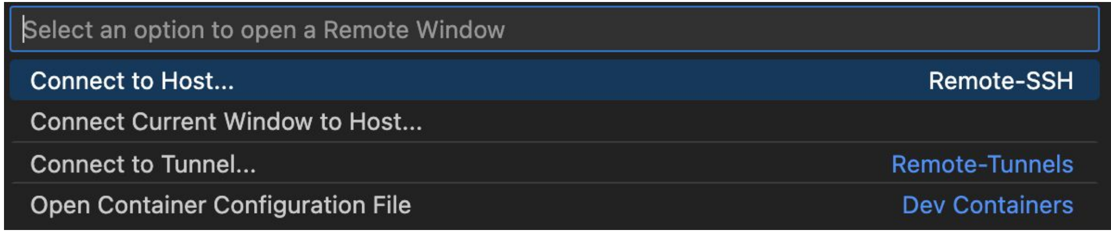
To connect to Scholar using SSH keys, you must follow three high-level steps:

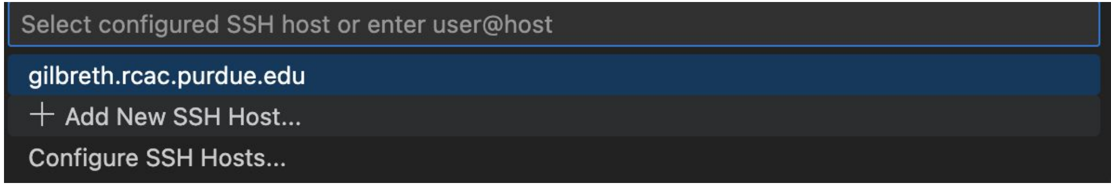
1. Generate a key pair consisting of a private and a public key on your local machine.
2. Copy the public key to the cluster and append it to `$HOME/.ssh/authorized_keys` file in your account.
3. Test if you can ssh from your local computer to the cluster without using your Purdue password.

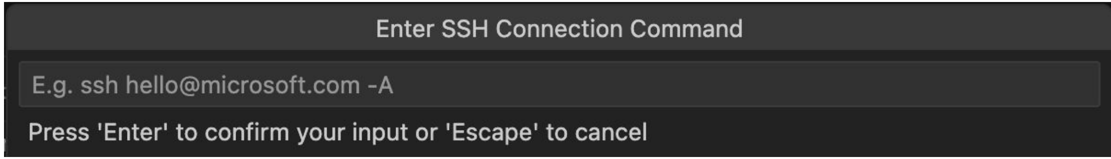
Detailed steps for different operating systems and specific SSH client softwares are give below.

Mac and Linux:

1. Run `ssh-keygen` in a terminal on your local machine. You may supply a filename and a passphrase for protecting your private key, but it is not mandatory. To accept the default settings, press Enter without specifying a filename.
   Note: If you do not protect your private key with a passphrase, anyone with access to your computer could SSH to your account on Scholar.
2. By default, the key files will be stored in `~/.ssh/id_rsa` and `~/.ssh/id_rsa.pub` on your local machine.


Select an option to open a Remote Window
Connect to Host...
Remote-SSH
Connect Current Window to Host...
Connect to Tunnel...
Remote-Tunnels
Open Container Configuration File
Dev Containers





Press 'Enter' to confirm your input or 'Escape' to cancel