

Purdue University - ITaP Research Computing

Gladys Andino

Dan Dietz

Jieyu Gao

Lev Gorenstein

Erik Gough

Stephen Harrell

Randy Herban

Steve Kelley

Boyu Zhang

Xiao Zhu

rcac-help@purdue.edu

August 29, 2017

Adopted to Scholar.

Original Radon tutorial slides available:

www.rcac.purdue.edu/training/clusters101/

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Vocabulary

Logging In

- Windows
- Mac
- Linux
- What to do afterwards

Modules

Types of Work

- Serial vs Parallel
- Parallel Computations

Good Citizenship

Jobs, Queues, and PBS

- Jobs
- Queues
- PBS
- Life of a Job

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Running Jobs

- Submission File
- Submitting
- Environment Variables
- Tips
- Example

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Vocabulary

Vocabulary

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

A cluster is a cluster is a cluster?

- Hardware (compute nodes + interconnect + storage)
- Software (OS + compilers + libraries + apps + queue manager)
- Infrastructure (front-ends + power + cooling + data center + staff)



Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

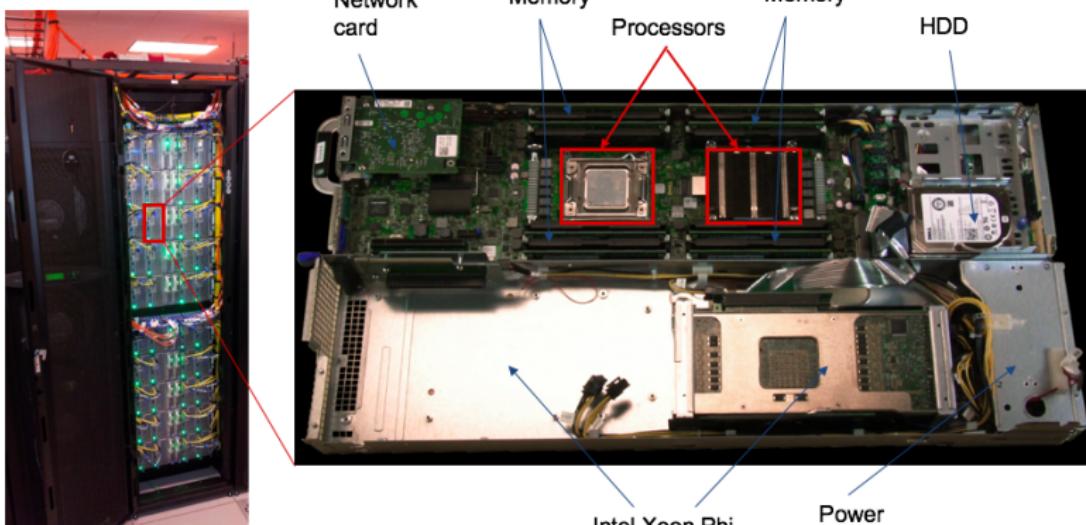
Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Node:



Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

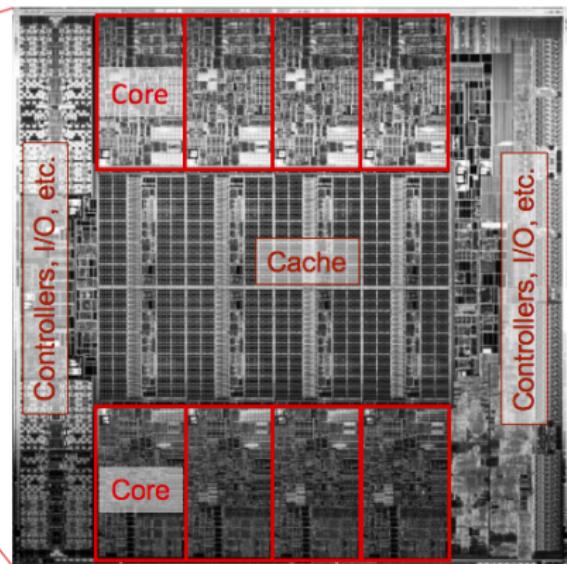
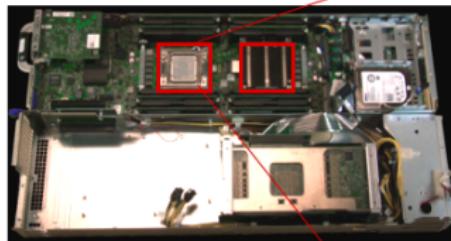
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

CPU:



Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

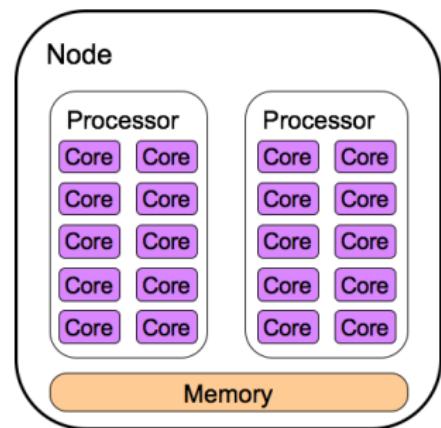
Storage

Globus

Workflows

Processor Core: individual compute unit ("slot") on the chip

- You see:
2 physical processors, 10 cores ea.
- Queuing system sees:
20 logical processors
(hence 'nodes=X:ppn=20'
effectively stands for "processor
cores per node").
- From now on, we will be mostly
concerned with cores (logical
processors), not physical chips
("I ran on 5 CPUs" == "on 5
processors" == "on 5 cores")



Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

IBM SYSTEM 360 (1 Computer, 1 Processor):



Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

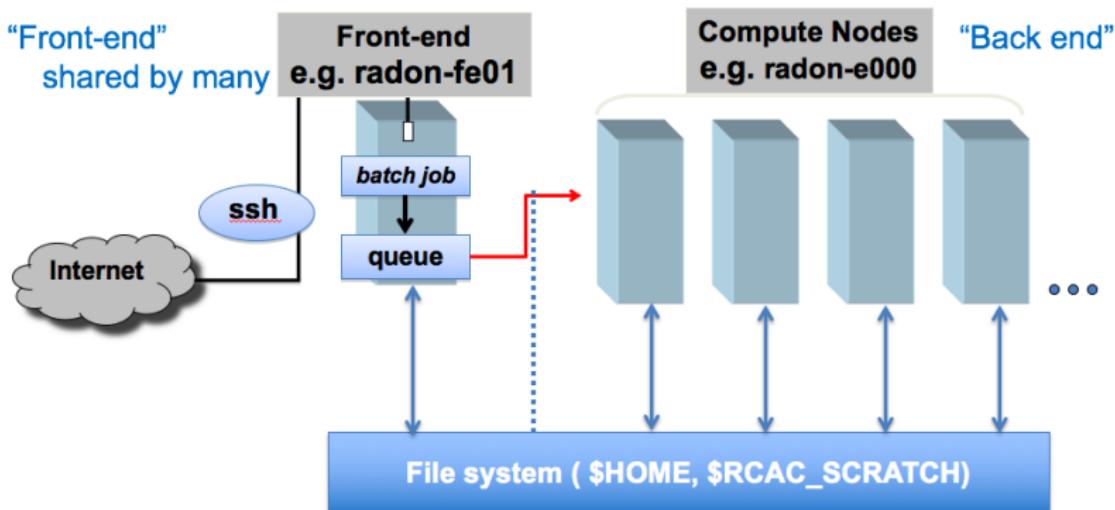
Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Front-end vs compute node:



Running Jobs: The goal is getting to the compute nodes

Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

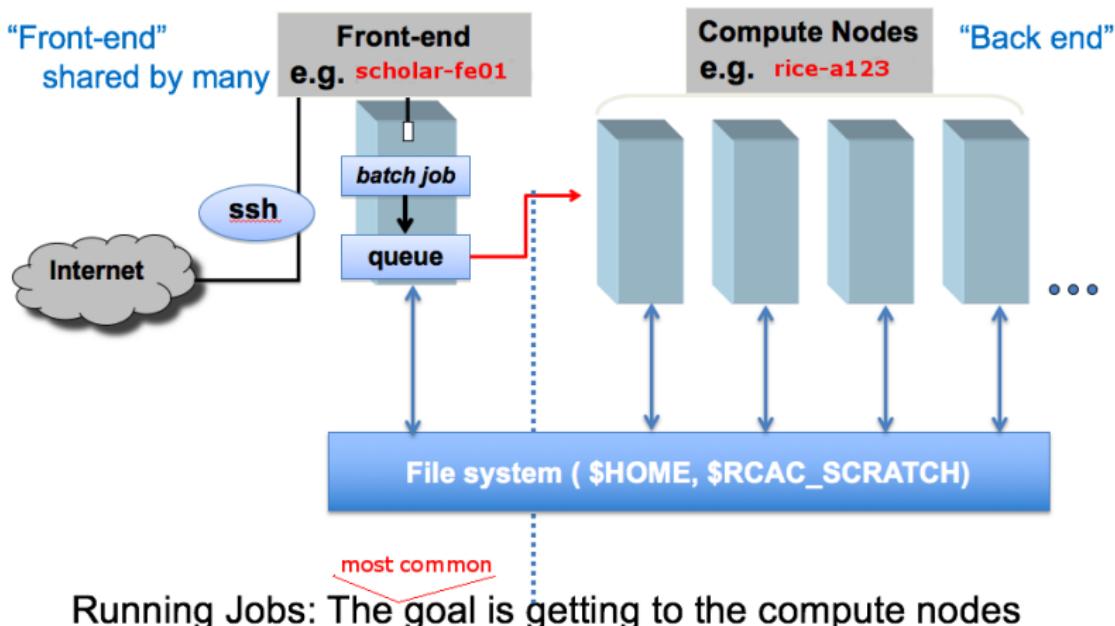
Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Scholar is special (own front-end, Rice compute nodes):



Vocabulary

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Community clusters:

HPC (Conte, Rice, Halstead, Brown): Multiple cores or nodes, probably MPI. Benefit from high-performance network and parallel filesystem. Some have accelerators (Xeon Phi or GPUs). The vast majority of campus - 80% of all work!

Data-Intensive Life Science (Snyder): Use entire node to get large amounts of memory. Less need for high-performance network. Needs large, fast storage.

Scholar: Special case for teaching. Uses Rice nodes, so mostly MPI at first glance. But also tweaked for interactive use (tasks on front-ends, Jupyter notebooks, Rstudio, etc)

Vocabulary

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Need help with ITaP Research Computing resources?

Check our User Guides and documentation:

www.rcac.purdue.edu/knowledge

You can always send us an email:

rcac-help@purdue.edu

Come to our coffee hour consultations:

www.rcac.purdue.edu/coffee

Logging In

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Logging In

- Windows
- Mac
- Linux
- What to do afterwards

Logging In

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thlnlnc

Storage

Globus

Workflows

You will be using the Scholar cluster:

- www.rcac.purdue.edu/compute/scholar/
- Everyone has been given an account on the cluster for the duration of the class
- If you wish to continue using [other] clusters after the class concludes, please make a request under your advisor or PI's name:

<https://www.rcac.purdue.edu/account/request/>

SSH (Secure SHell) is the main connection method.

Logging In

Windows

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Many clients are available for Windows:

- PuTTY SSH client
 - Download PuTTY, no install required
 - <http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html>
(or Google search *putty*)
 - Download *putty.exe* for Intel x86 to your desktop
- MobaXTerm SSH client
 - Download MobaXTerm (free Home Edition), no install required
 - <http://mobaxterm.mobatek.net/download.html>
 - Has built-in file transfer client, X-windows server, Mosh support

Logging In

Windows

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good

Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

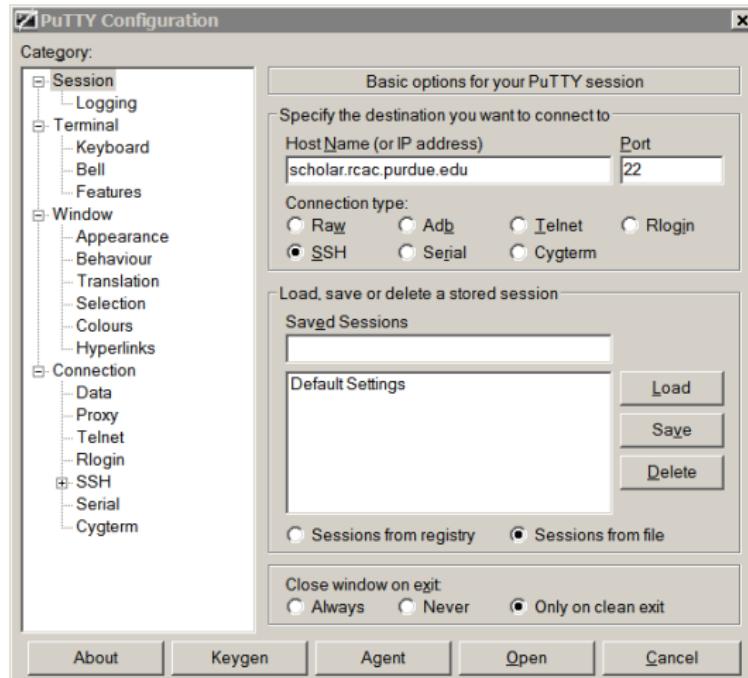
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Host Name for Scholar is scholar.rcac.purdue.edu



Logging In

Mac

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good

Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job

Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Open Terminal app and connect using:

```
ssh myusername@scholar.rcac.purdue.edu
```

Logging In

Linux

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good

Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

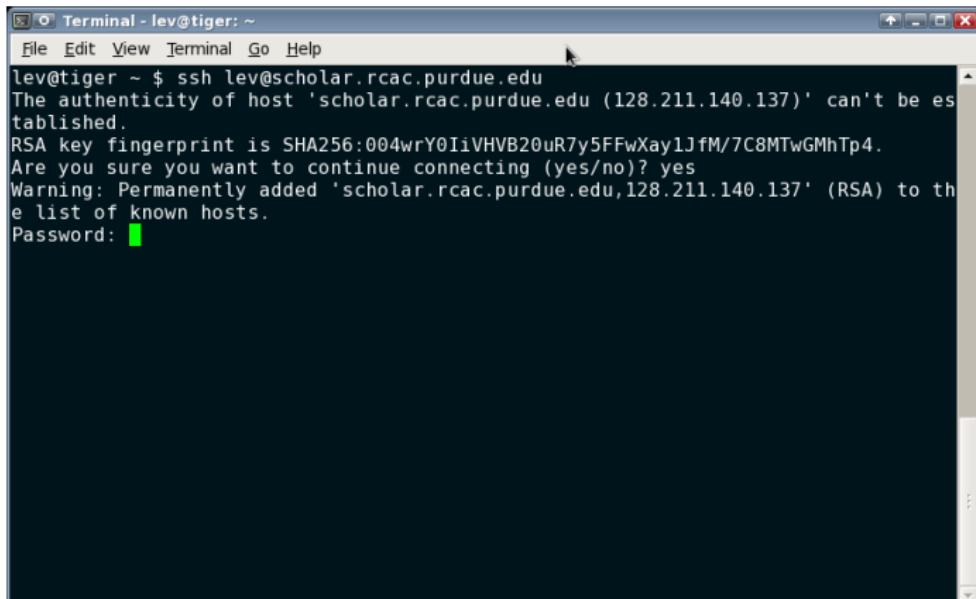
Storage

Globus

Workflows

Linux also has a built in terminal client, similar to Mac:

```
ssh myusername@scholar.rcac.purdue.edu
```



Logging In

What to do afterwards

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good

Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job

Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

- Plenty of general Linux tutorials online.
- We've got a couple, too:
 - <https://www.rcac.purdue.edu/training/unix101/>
 - <https://www.rcac.purdue.edu/training/unix201/>
- Scholar user guide:
<https://www.rcac.purdue.edu/knowledge/scholar>

Logging In

What to do afterwards

Vocabulary

Logging In

Windows

Mac

Linux

What to do afterwards

Modules

Types of Work

Good

Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job

Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Ways to transfer files between cluster and your computer:

- SCP and SFTP (with command line or GUI clients)
- Mapping as network drive
- Globus (great for large transfers)

Relevant user guide section: <https://www.rcac.purdue.edu/knowledge/scholar/storage/transfer>

Modules

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Modules

Modules

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

All ITaP Research Computing clusters use a module system

- Module system provides for the dynamic modification of a user's environment
- Module commands allow you to add applications and libraries to your environment
- This allows us to simultaneously and safely provide several versions of the same software
- You can set up the environment for running applications using simple commands:

```
$ module load intel
$ module load openmpi
```

- All ITaP Research Computing clusters have a default programming environment (compiler and MPI library) loaded for you when you log in.

Modules

Vocabulary

Logging In

Modules

Types of Work

Good

Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Show all modules currently loaded in my environment:

```
$ module list

Currently Loaded Modules:
 1) intel/16.0.1.150  2) openmpi/1.8.1  3) rcac
```

Load a new module:

```
$ module load matlab
$ module list

Currently Loaded Modules:
 1) intel/16.0.1.150  2) openmpi/1.8.1  3) rcac
 4) matlab/R2016a
```

Modules

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Shows you all software modules currently available to you:

```
$ module avail
```

Filters software modules currently available to you by a search term. In this case, show all Python versions available:

```
$ module avail python

----- Applications built with Intel 16.0.1.150 -----
python/2.7.8

----- Core Applications -----
python/anaconda (D)      python/2.7.2      python/3.4.1
```

Modules

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thlnlnc

Storage

Globus

Workflows

All ITaP Research Computing clusters use an hierarchical software configuration. Modules are organized by the programming environment (compiler and libraries) that they were built with and depend on. You will only see (with avail) modules compatible with the programming environment currently loaded.

As you swap in other programming environments (compiler and libraries), you may find additional software available to you. Any libraries or software dependent on the programming environment will be automatically swapped with compatible versions as you change environments.

Modules

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

We can look for software not immediately available to us - that is, things that may not show up in avail output:

```
$ module avail gromacs  
$ module spider gromacs
```

```
gromacs: gromacs/5.0
```

You will need to load one of the set of module(s) below before the "gromacs/5.0" module is available to load.

```
intel/13.1.1.163  openmpi/1.8.1
```

This tells us we must swap into a new programming environment to load gromacs:

```
$ module load intel/13.1.1.163 openmpi/1.8.1 gromacs/5.0
```

Modules

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Unload a currently loaded module:

```
$ module unload matlab
$ module list

Currently Loaded Modules:
 1) intel/13.1.1.163   2) openmpi/1.8.1       3) rcac
 4) gromacs/5.0
```

Unload and purge all currently loaded modules:

```
$ module purge
$ module list
No modules loaded
```

Modules

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

What exactly does module load do?

It will change the environmental variables defined for that module. Typically these are \$PATH and \$LD_LIBRARY_PATH. Also a few extra module-specific environmental variables could be set.

```
$ module show vmd
-----
/opt/modulefiles/core/vmd/1.9.1.lua:
-----
whatis      invoke VMD 1.9.1
setenv      VMD_HOME "/apps/rhel6/vmd-1.9.1"
prepend_path PATH "/apps/rhel6/vmd-1.9.1/bin"
prepend_path LD_LIBRARY_PATH "/apps/rhel6/vmd-1.9.1/lib"
setenv      LIBGL_ALWAYS_INDIRECT "y"
```

Modules

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

A few tips from us:

- Don't load modules in `.bashrc/.profile/.login/.cshrc`
- Don't load more than needed for current task
- Do `module purge` when in doubt
- Do check with `module show` what a given module does
- Do check with `module list` what's actually loaded
- Do script and automate your builds - even if its just a 3-line snippet! Reproducibility is good!

Types of Work

Vocabulary

Logging In

Modules

Types of Work

Serial vs Parallel

Parallel

Computations

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Types of Work

- Serial vs Parallel
- Parallel Computations

Types of Work

Serial vs Parallel

Vocabulary

Logging In

Modules

Types of Work

Serial vs Parallel

Parallel Computations

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinvnc

Storage

Globus

Workflows

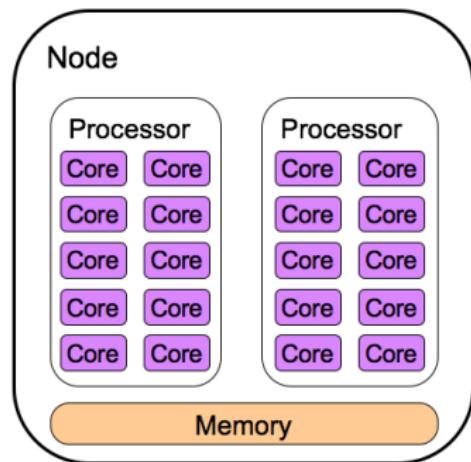


More than one mower!

Mow the lawn faster ...

... or mow a bigger lawn

Lawn mowers = cores



Types of Work

Serial vs Parallel

Vocabulary

Logging In

Modules

Types of Work

Serial vs Parallel

Parallel Computations

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

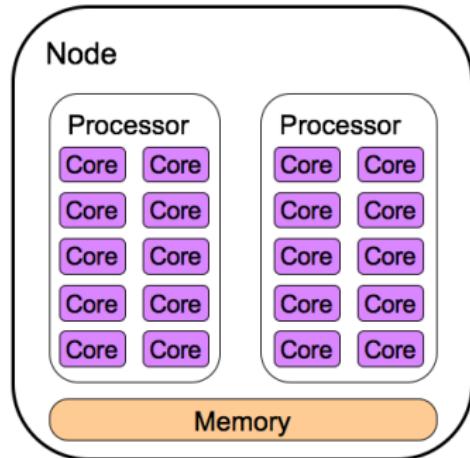
Storage

Globus

Workflows



More than one mower!
Mow the lawn faster ...
... or mow a bigger lawn
Lawn mowers = cores



Types of Work

Parallel Computations

Vocabulary

Logging In

Modules

Types of Work

Serial vs Parallel

Parallel
Computations

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Parallel computations come in a few flavors:

- MPI - Message Passing Interface, multiple nodes
- OpenMP - confine yourself to one node
- Hybrid - mix MPI and OpenMP
- Accelerators/coprocessors - GPU, Intel Xeon Phi
- Parallel R, Matlab, Hadoop and Spark

Types of Work

Parallel Computations

Vocabulary

Logging In

Modules

Types of Work

Serial vs Parallel

Parallel
Computations

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

A few tips from us:

- Know your computational needs (high performance? high throughput?)
- Know the nature of your program (single-core? single node? multi-node? CPU-bound? memory-bound? communications-bound?)
- Before running massive production simulations, consider investing a little time in small scaling studies (performance vs # of cores/nodes)

Good Citizenship

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Good Citizenship

Good Citizenship

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

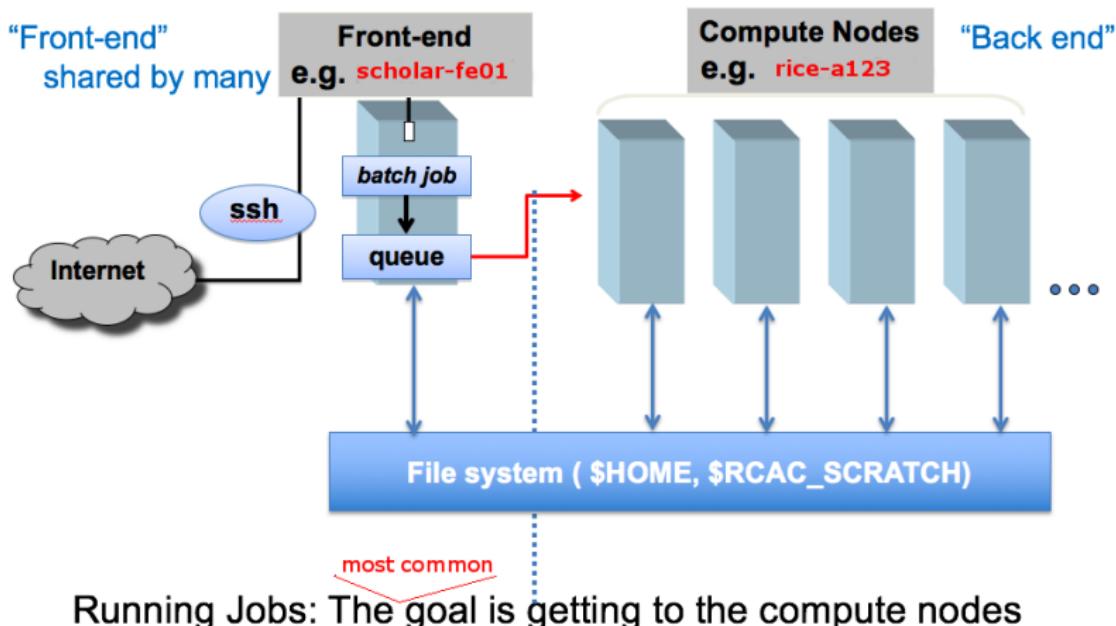
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Front-end vs compute node:



Good Citizenship

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

What does this mean?

- Do not do science on the front end
- Front-ends are limited shared resources and you'll annoy the system administrators and other users
- Watchdog will monitor and quickly terminate (or at the very least, severely throttle) any inappropriate resource-intensive process
- Instead: submit a job. We'll get to that

Scholar is an exception to the "kill" rule, and you could run some things on its front-ends... but for anything heavy you'd still be much better off submitting a job.

Good Citizenship

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

What can I do on the front-ends?

- Building/compiling software
- Managing files: editing, transferring, tar, gzip, hsi
- Submitting, monitoring, and managing batch jobs
- Launching interactive jobs
- Modest post-processing and analysis

Jobs, Queues, and PBS

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Jobs, Queues, and PBS

- Jobs
- Queues
- PBS
- Life of a Job

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

What is a job?

- A job is simply a set of tasks to be performed by a cluster
- A script to instruct the cluster precisely what to do to complete your work
- Self contained to be executed without any interaction
- Submit and forget!

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Remember, cluster front-end nodes are shared resources for

- Creating, submitting, and monitoring jobs
- File transfers
- Preparing inputs
- Editing and compiling code
- Small-scale testing

May be used by 50+ people simultaneously

- Check out the `who` command
- Jobs should have no interference by other people
- Want jobs carefully arranged on compute nodes

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Jobs are submitted to the cluster:

- Cluster executes jobs on back-end compute nodes
- Jobs are carefully scheduled and arranged on the compute nodes

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Jobs need to specify the resources they require:

- Three basic units:
 - Number of nodes
 - Number of cores
 - Time
- Queue
- Memory
- Node access policy
- Other resources

Cluster will allocate requested resources once they are available. Job starts once resources are allocated.

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

The most basic unit you can request is nodes:

- Serial applications probably only need one node
- Parallel applications may need 10s or 100s of nodes

Job should also request how many cores per node it needs:

- Cluster allows your job to use only the cores you request
- Often your job should request all of the cores in a node
- Remember, we want to avoid other people!

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

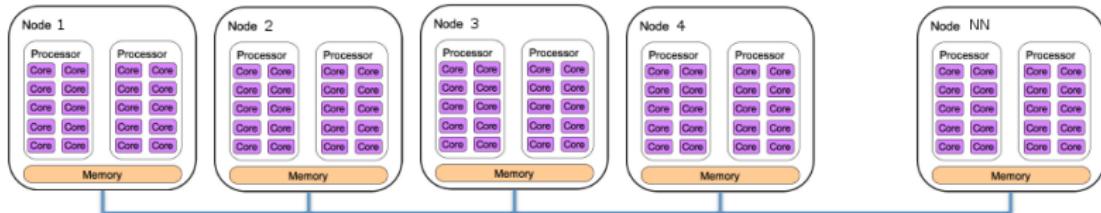
Globus

Workflows

Nodes are physically different pieces of hardware.

A job with 20 nodes and 20 cores each, is **not** 400 threads

- 20 individual computers!
- Program needs to be written with MPI to tie all the nodes together



Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

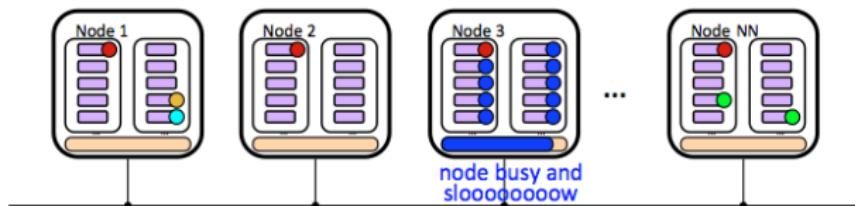
Storage

Globus

Workflows

Nature of the job: how much inter-process communications?
1 core on 20 nodes, or for 20 cores on 1 node?

- Communication inside a node is faster than between nodes
- Collisions with other people's jobs ($20 \text{ nodes} \times 19 \text{ other jobs} = 380 \text{ jobs}$)
- Most performant, most predictable, least vulnerable, least interfering



Tight packing preferred. Most ITaP clusters enforce exclusive nodes unless explicitly specified for single node jobs.

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

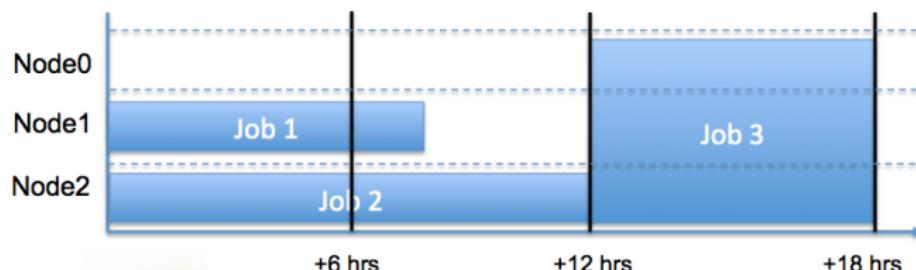
Globus

Workflows

The third basic unit jobs can request is walltime. The cluster won't let your job run forever!

Cluster needs to have a time limit for your job:

- Helps the cluster efficiently schedule you on the nodes
- Make request accurate with some safety buffer
- Tune your walltime request through testing and experimenting



Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Jobs need memory, but...

- Clusters can **not** limit your job to an amount memory
- All of a compute node's memory is available to your job
- All of that memory is also available to any other job on the node!
- 1 core job needs $16\text{ GB} \times 20\text{ jobs} = 320\text{ GB}$

But wait, my Scholar (Rice) node only has 64 GB!!

Jobs crash and burn

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Memory across multiple compute nodes is distributed. Memory is discrete chunks of memory in different address spaces.

Can't just "add more nodes" to get more memory:

- 20 nodes with 16 GB memory each is **not** 320 GB of memory
- 20 discrete chunks of 16 GB memory
- Can't allocate 320 GB of memory to a single process
- Need MPI to tie memory together



Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Best bet is to always request all cores of a node or request exclusive access with the node access policy. You get full control of the node and all of its memory. You can try requesting proportionate numbers of cores but not a guarantee!!

Another option is to use the node access policy to request that only your own jobs can share nodes. You don't have to worry about somebody else getting in the way, but you do have to be mindful of what resources your own job is using.

The Tao of Defensive Computing: *Protect yourself from others, protect others from yourself*

Jobs, Queues, and PBS

Jobs

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

There are a number of other resources you can request in a job:

- Accelerators (GPUs, Xeon Phis)
- Specific types of nodes
- Certain licenses: MATLAB and its Toolboxes

Jobs, Queues, and PBS

Queues

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Cluster can't run all jobs at once so sometimes you must wait

- Jobs are submitted into queues
- Jobs wait until cluster and queue has free resources
- Queues set constraints on jobs that can be submitted
 - Max nodes
 - Max walltime
- Sets initial priorities
- Our clusters have several queue types

Jobs, Queues, and PBS

Queues

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Owner queues:

- Named for your PI, lab, or group
- Number of nodes set by what your PI buys into
- Allows for longest jobs - 336 hours
- Jobs should start in 4 hours or less (assuming your queue isn't full of other jobs in the queue)

Queues are **not** tied to specific nodes in the cluster. They allow you to use the number of nodes somewhere in the cluster - whichever nodes are free, functioning, and of the type the queue can access.

Jobs, Queues, and PBS

Queues

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Standby queue:

- Uses idle nodes from owner queues
- Everyone on cluster gets access
- Limited to 4 hours (owners must be able get quick access from their queue)
- No limit on number of nodes (probably won't get whole cluster though)
- Lowest priority jobs, no promises on turnaround time (can be minutes or days)
- If you can run under 4 hours, go for it!

Jobs, Queues, and PBS

Queues

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs

Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Debug queue:

- Allows for running small jobs for testing and debugging
- Up to 2 nodes for 30 minutes at a time
- Highest priority
- Starting time <1 minute

Jobs, Queues, and PBS

Queues

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues

PBS
Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

In this class you are using Scholar, which is a special dedicated queue on Rice cluster.

scholar queue:

- Sort of like owner queues, but with a twist
- Has higher priority
- Tighter limits on # of simultaneous jobs from one user
- Shorter jobs (but longer than standby)

RCAC has `qlist` command to see all queues you have access to, their state and limitations.

Jobs, Queues, and PBS

PBS

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

What is PBS?

- Portable Batch System
- Originally developed for NASA in 1991
- Open source fork and commercial fork remain
- We use the open source fork, TORQUE

en.wikipedia.org/wiki/Portable_Batch_System

Jobs, Queues, and PBS

PBS

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

What is PBS, really? It is the resource manager for the cluster:

- Job submission
- Job status
- Executes jobs
- Manages queues
- Manages nodes

Jobs, Queues, and PBS

PBS

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues

PBS

Life of a Job

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

The scheduler is actually separate software. We use a scheduler called Moab. It decides which jobs to run and where.

"PBS" == "Scheduler" == "Batch system" == "Server"

- We often interchange them in conversation
- All refer to the whole system of servers, schedulers, and managers

Jobs, Queues, and PBS

Life of a Job

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

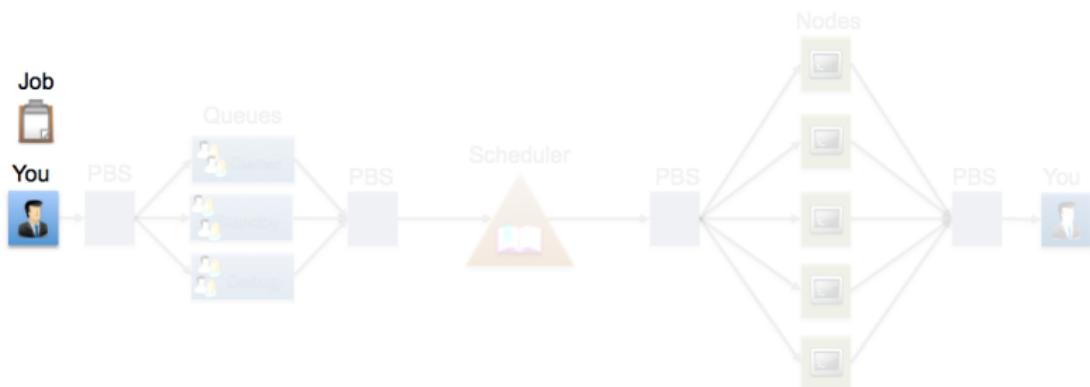
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

You prepare a job submission script



Jobs, Queues, and PBS

Life of a Job

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

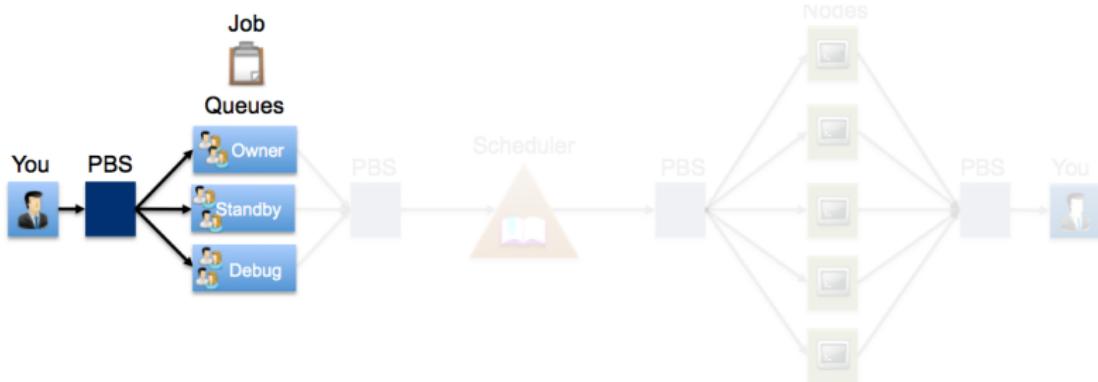
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

You submit job script into PBS. The job is placed into a queue.



Jobs, Queues, and PBS

Life of a Job

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

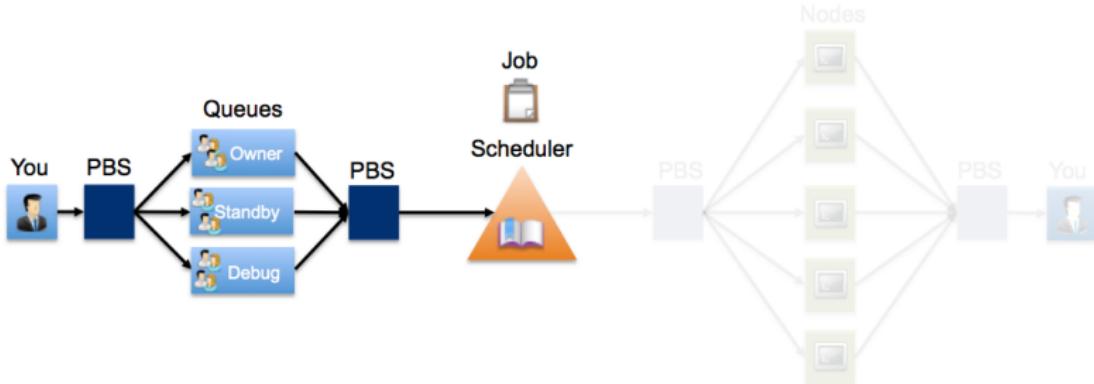
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Scheduler iteratively asks PBS for new jobs, status of old jobs, and status of nodes. Each cycle can last several minutes.



Jobs, Queues, and PBS

Life of a Job

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

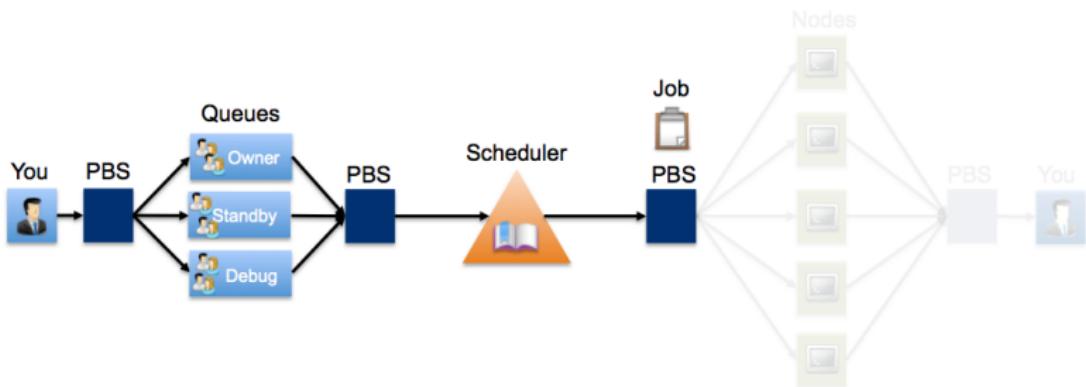
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Scheduler considers jobs, acts as a handler. Tells PBS to start jobs one at a time. Goes back and does this over and over.



Jobs, Queues, and PBS

Life of a Job

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

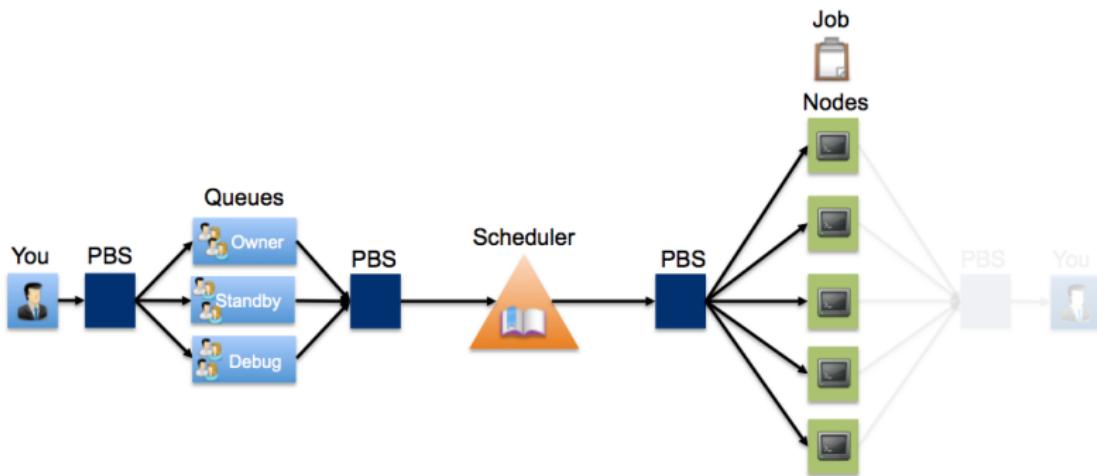
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

PBS checks health of the node. Send job script to node(s) and executes it.



Jobs, Queues, and PBS

Life of a Job

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Jobs
Queues
PBS
Life of a Job

Running Jobs

Job
Monitoring

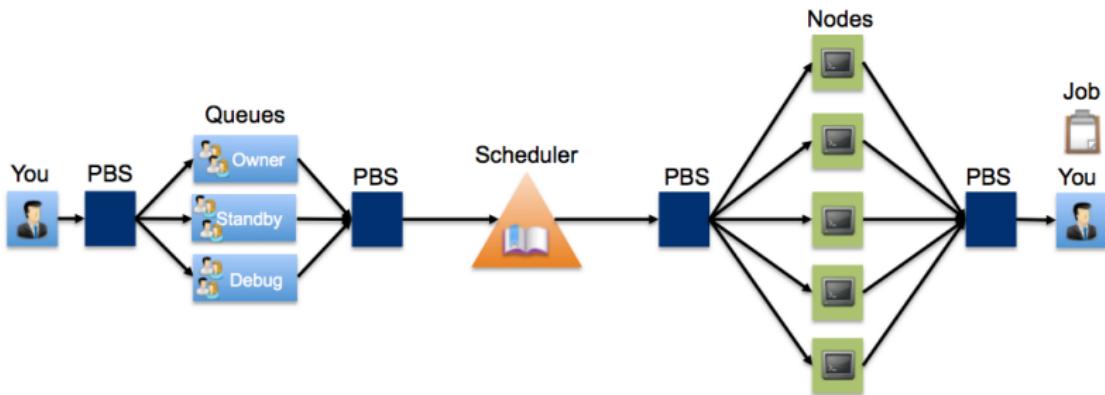
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

PBS watches for the job script to complete on nodes. Collects job output files from nodes and sends files to you.



Running Jobs

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs
Submission File
Submitting
Environment Variables
Tips
Example

Job Monitoring

Interactive Jobs and Thinlinec

Storage

Globus

Workflows

Running Jobs

- Submission File
- Submitting
- Environment Variables
- Tips
- Example

Running Jobs

Submission File

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

To run a job you must first create a submission file. This file instructs the cluster everything it needs to know to run your job.

Components of a submission file:

- PBS directives: specify resources needed such as number of nodes, cores, and time
- Module load to set up paths, libraries
- PBS environment variables
- Customized commands to make your job run

Running Jobs

Submission File

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs
Submission File
Submitting Environment Variables
Tips
Example

Job Monitoring

Interactive Jobs and Thinvnc

Storage

Globus

Workflows

PBS directives set PBS job attributes. They appear at the top of your submission file and are prefaced with #PBS in your submission file or could be specified at submit time as arguments.

Common PBS job attributes include the following. None are absolutely required, but you should always set the following three. Without them, PBS will assume defaults which are likely not useful to you.

PBS Directive	Description
-l nodes=2:ppn=20	Number of nodes and cores
-l walltime=00:10:00	Maximum walltime allowed
-q scholar	Destination queue

Running Jobs

Submission File

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

There are also several other directives you can optionally set. Some of the more common ones follow. As usual `man qsub` will document all available options.

PBS Directive	Description
<code>-N jobname</code>	Give your job a custom name
<code>-o /path/to/output.out</code>	Specify the job output file
<code>-e /path/to/error.err</code>	Can use \$PBS_JOBID to make unique Specify the job error file
<code>-l naccesspolicy=policy</code>	Can use \$PBS_JOBID to make unique Node access policy
<code>-n</code>	Request exclusive access to the node

Running Jobs

Submission File

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

`naccesspolicy` can be set to one of the following:

- shared: allowing sharing of a node between jobs (default on Radon)
- singlejob: only allow one job per node (default on most other ITaP clusters)
- singleuser: only allow one user's jobs on a node

Relevant user guide section: [www.rcac.purdue.edu/
knowledge/scholar/run/pbs/naccesspolicy](http://www.rcac.purdue.edu/knowledge/scholar/run/pbs/naccesspolicy)

Running Jobs

Submission File

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Submission File

Submitting Environment Variables

Tips

Example

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

After requesting resources, your submission script should load required modules to set up the environment for your application or code. Pretend you have just logged in and have nothing set up yet.

It may be a good idea to purge any currently loaded modules (whether from your login scripts or system default) so you get a predictable and reproducible environment.

```
#PBS -l nodes=1:ppn=1
#PBS -q scholar

module purge
module load r
```

Note that by default this will still reserve you a whole node (20 cores) despite ppn=1 (because naccesspolicy).

Running Jobs

Submission File

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Now that you have resources requested and your environment set up, cd to where your files are, then run the commands necessary to do your work.

A job to run an R script:

```
#PBS -l nodes=1:ppn=1
#PBS -q scholar

module purge
module load r

cd ~/myproject/ # PBS starts job in home directory
                 # so need to move where project files are

Rscript my_script.r arg1 arg2
```

Running Jobs

Submitting

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Once your job submission script is written you can submit it with the qsub command:

```
$ qsub myjob.sub  
1234567.rice-adm.rcac.purdue.edu
```

If submission was successful, you should get back the Job ID that PBS has assigned to your job. Only the numerical part is strictly required - the rest of the ID identifies which cluster the job is on but is not required for further operations on the job.

Job scripts do not require the .sub extension, however, it is convenient to make it easy to identify submission files.

Running Jobs

Submitting

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

You can also specify PBS directives directly to qsub rather than in the job file. Any directives that will never change should be put in the file, however, by specifying directly to qsub you can use the same submission file with slight variations on one or more of the directives.

Here we can specify different job names for the same job script. We'll see in one of the following examples how we can take advantage of this.

```
$ qsub -N input_1  
1234567.rice-adm.rcac.purdue.edu  
$ qsub -N input_2  
1234568.rice-adm.rcac.purdue.edu  
$ qsub -N input_3  
1234569.rice-adm.rcac.purdue.edu
```

Running Jobs

Environment Variables

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Submission File
Submitting

Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

PBS defines several variables automatically for you. They can be used within your job submission file. Some common variables include:

Variable	Description
PBS_O_WORKDIR	Absolute path of the current working directory when you submitted this job
PBS_JOBID	Job ID assigned to this job by PBS
PBS_JOBNAME	Job name (defined by you or PBS)
PBS_NODEFILE	Path to a file containing the list of nodes assigned to this job

Running Jobs

Environment Variables

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

\$PBS_O_WORKDIR contains the path to the directory from which you submitted the job. \$PBS_O_WORKDIR can be used to cd into that directory within the job.

When a job starts it will default to your home directory,
so you likely need to cd somewhere first.

```
#PBS -l nodes=1:ppn=20
#PBS -q scholar

# Change to the directory from which you submitted this job
cd $PBS_O_WORKDIR

# Print out the current working directory path
pwd
```

Running Jobs

Environment Variables

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Submission File

Submitting

Environment Variables

Tips

Example

Job

Monitoring

Interactive Jobs and Thinvnc

Storage

Globus

Workflows

\$PBS_JOBID can be used to store output from a program in a uniquely named file. In this way you can reuse the same job file across multiple inputs and run several instances simultaneously without overwriting output from the other jobs.

```
#PBS -l nodes=1:ppn=20
#PBS -q scholar

# Change to the directory from which you submitted this job
cd $PBS_O_WORKDIR

# Make a directory for each job to store the job output and error
file
mkdir $PBS_JOBID
./myprogram 1> $PBS_JOBID/myprogram.out 2> $PBS_JOBID/myprogram.err
```

Running Jobs

Environment Variables

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

PBS_JOBNAME allows you to use the same script for multiple inputs/arguments. You can submit the same submission file with different job names (specify directly to qsub rather than hard code in the file) and then feed the job name into your program as an argument.

```
#PBS -l nodes=1:ppn=20
#PBS -q scholar

# Change to the directory from which you submitted this job
cd $PBS_O_WORKDIR

# Use PBS_JOBNAME as the input argument to my program
./myprogram $PBS_JOBNAME
```

Running Jobs

Environment Variables

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

PBS_NODEFILE provides a path to a file containing the list of nodes assigned to your job (one entry for each core). This is typically passed into an MPI program telling it where to run.

MPI programs run inside a PBS job will automatically detect the nodefile. However, you may want to change the nodefile and run your MPI program with some complex geometry.

```
#PBS -l nodes=2:ppn=20
#PBS -q scholar
#PBS -N input_800

module load intel openmpi

# Change to the directory from which you submitted this job
cd $PBS_O_WORKDIR

# And run the program in parallel
mpiexec -n $PBS_NP -machinefile $PBS_NODEFILE ./mpi_hello
```

Running Jobs

Tips

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables

Tips

Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

If you have lots of small single core tasks, you have two options:

Option 1: one process per PBS job. Drawback: overhead!

Option 2: pack processes into a single PBS job

Drawbacks: need to be careful not to overwhelm the node at once, slightly more complicated submission script.

```
#PBS ...
command1 &
command2 &
...
command20 &
wait
```

& will let the command run in the background and will move onto the next command. If you do this **you must have a wait at the end**. Otherwise PBS will think your job has completed the instant all commands are started and exit prematurely.

Running Jobs

Tips

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips

Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

When scripting PBS submissions, use some sleep between qsubs. Rapid submissions may cause server to trip and drop some submissions.

```
$ for i in {1..10}; do qsub -N input_$i; sleep 0.5; done
```

```
#PBS -l nodes=1:ppn=20
#PBS -q scholar

# Change to the directory from which you submitted this job
cd $PBS_O_WORKDIR

# Use PBS_JOBNAME (-N) as the input argument to my program
./myprogram $PBS_JOBNAME
```

Running Jobs

Example

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs
Submission File
Submitting
Environment
Variables
Tips
Example

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Real life example:

```
#!/bin/bash -l
#PBS -N blastx_sp
#PBS -q scholar
#PBS -l nodes=1:ppn=20
#PBS -l walltime=03:00:00

# set up modules
module purge
module load bioinfo
module load blast
module list

# move to our working directory
cd $PBS_O_WORKDIR
pwd

# log the start time
date +"%d %B %Y %H:%M:%S"
```

Running Jobs

Example

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Submission File
Submitting
Environment
Variables

Tips

Example

Job

Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Real life example (continued):

```
echo " "  
  
# run blast with our inputs  
blastx -query sequences.fasta -out seq.blastx_sp.fmt6 \  
-db swissprot -num_threads 8 -outfmt 6 -evalue 1E-06 \  
-max_target_seqs 5  
  
# log completion time  
echo " "  
date +"%d %B %Y %H:%M:%S"
```

```
$ qsub subjob4cluster101.sub  
787233.rice-adm.rcac.purdue.edu
```

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

**Job
Monitoring**

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Job Monitoring

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

qlist can be used to list all queues you can use and their current status and limits:

\$ qlist					
Queue	Current Number of Cores				
	Total	Queue	Run	Free	Max Walltime
=====	=====	=====	=====	=====	=====
debug	80	20	0	80	0:30:00
scholar	320	0	20	300	48:00:00
standby	11,260	54,380	5,340	609	4:00:00

This is a custom command unique to ITaP clusters, so this may not work on other systems.

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

`qstat -u myusername` can be used to list all current jobs from the user `myusername`. Typically this would be your own username and your own jobs:

```
$ qstat -u gandino
rice-adm.rcac.purdue.edu:

          Req'd   Req'd   Elap
          -----   -----   -----
Job ID  Username Queue    Jobname   SessID NDS TSK Memory Time S Time
-----  -----   -----   -----   -----  --  --  --  --
787233  gandino  scholar  blastx_sp  --      1  20   --  03:00:00 Q  --
```

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Some of the qstat fields explained in more detail:

Field	Description
SessID	Process ID of running job. Will show up shortly after job starts.
NDS	Number of nodes assigned to job
TSK	Total processors assigned to job
Req'd Memory	Requested memory (if specified)
Req'd Time	Requested maximum walltime
S	Status: Queued (Q), Running (R), Held (H), Complete (C)
Elap Time	Elapsed walltime of running job

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

**Job
Monitoring**

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

qstat -a myqueue can be used to check jobs in your lab's queue and see how busy your queue is:

```
$ qstat -a scholar
```

```
...
```

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

checkjob is useful to see why your job hasn't started:

```
$ checkjob 787233
job 787233
Aname: blastx_sp
State: Idle
...
BecameEligible: Mon Oct 19 17:15:04
    # If a queue is running capacity, your job will not
    # be immediately eligible for scheduling
SubmitTime: Mon Oct 19 17:13:34
...
NOTE: job req cannot run in partition rice-adm (available procs do
      not meet requirements : 0 of 20 procs found)
idle procs: 27 feasible procs: 0
    # if there are no nodes available to run your job you will
    # get a message similar to this. There may be free (idle)
    # processors scattered across the cluster, but may not be
    # on contiguous nodes to fit your job
```

checkjob queries the scheduler directly, so may not respond while it is busy

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

showstart will give you an estimate of when your job should start:

```
$ showstart 787233

job 787233 requires 8 procs for 3:00:00
*Estimated Rsv based start in 17:14:47 on Tue Oct 20 10:39:31
Estimated Rsv based completion in 20:14:47 on Tue Oct 20 13:39:31
```

This is a very rough estimate and is only accurate in the current state of the cluster (assuming no more jobs will be submitted). It will not be accurate at all for low-priority queues (such as standby), as jobs there are constantly pushed back in line as owner queue jobs come are submitted.

Job Monitoring

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

`qdel myjobid` will stop and delete the specified job:

```
$ qdel 787233
```

Occasionally the node running your job will become unresponsive (might be why you're trying to delete the job!) and you will receive an error message from `qdel`. In this case, give it a few minutes to clear on its own, and if it still doesn't clear email us at rcac-help@purdue.edu and we will manually purge the job.

Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

**Interactive
Jobs and
Thinlinc**

Storage

Globus

Workflows

Interactive Jobs and Thinlinc

Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Jobs are typically run without user interaction, but some do need interaction. An interactive job will give you an interactive session on a (dedicated to you if so requested) compute node where you can do heavy computations (vs a shared front-end).

- Test code without impacting others
- Quicker develop / test / debug cycle
- Run GUI apps on a dedicated compute node (Matlab, IGV, Fluent, etc)

Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

You can submit an interactive job:

```
$ qsub -I  
qsub: waiting for job 7867033.rice-adm.rcac.purdue.edu to start
```

The job will be queued like any other job, and may take some time to start. Eventually the job will start and you will be placed on a compute node:

```
qsub: job 7867033.rice-adm.rcac.purdue.edu ready  
$ hostname  
rice-a123
```

Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

To have graphics support in an interactive job, you need to a) have an X11 server installed on your local machine, b) enable X11 forwarding when connecting to cluster, and then c) enabled X11 forwarding in your interactive job:

```
$ ssh -Y scholar.rcac.purdue.edu  
$ qsub -I -X  
qsub: waiting for job 7867033.rice-adm.rcac.purdue.edu to start
```

If you don't enable X11 forwarding you might encounter an error like this:

```
Exception in thread "main" java.awt.HeadlessException:  
No X11 DISPLAY variable was set, but this program performed an  
operation which requires it.  
        at java.awt.GraphicsEnvironment.checkHeadless(  
GraphicsEnvironment.java:204)  
        at java.awt.Window.<init>(Window.java:536)  
        at java.awt.Frame.<init>(Frame.java:420)  
        ...
```

Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

We recommend using Thinlinc:

- Can be accessed from web browser or via stand-alone client
- Gives you a remote desktop environment
- Thinlinc keeps session alive after disconnect
- Faster than X11 forwarding directly to your laptop,
especially if you are off-campus

Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Start a browser and navigate to thinlinc.rcac.purdue.edu (general use, lands on Thinlinc server farm) or better yet, desktop.scholar.rcac.purdue.edu (dedicated to Scholar, will land on Scholar front-end):



Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

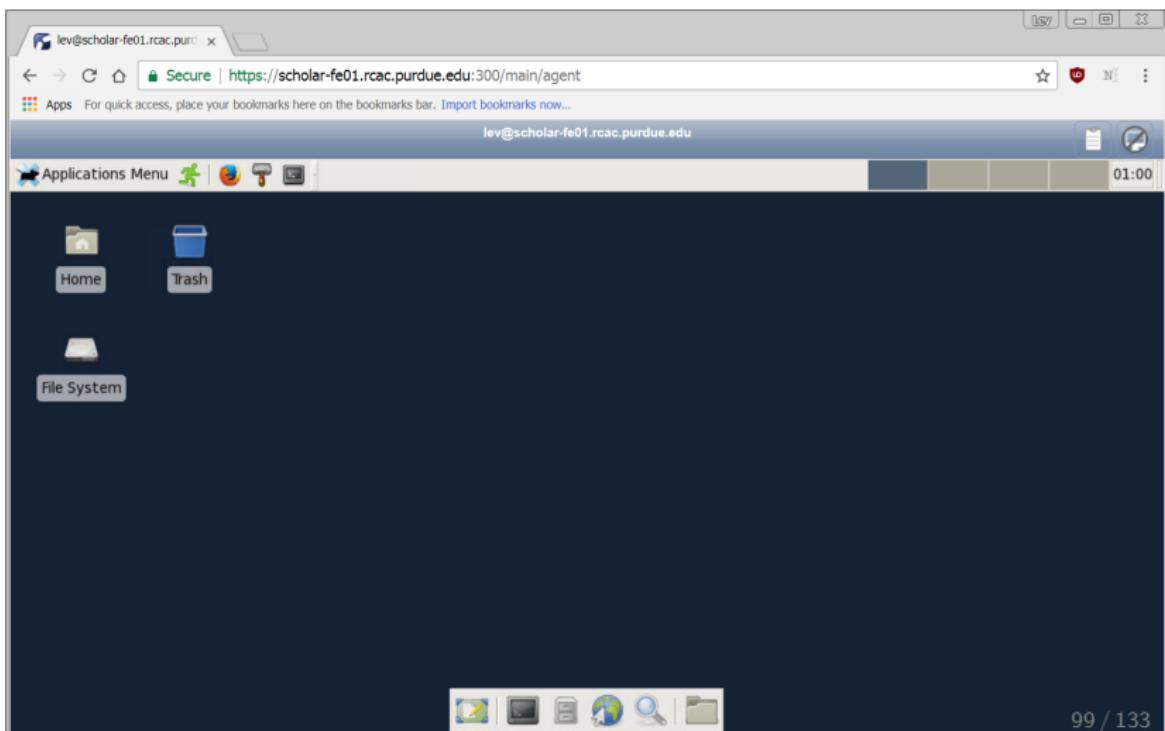
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Log in and wait for remote desktop to start:



Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

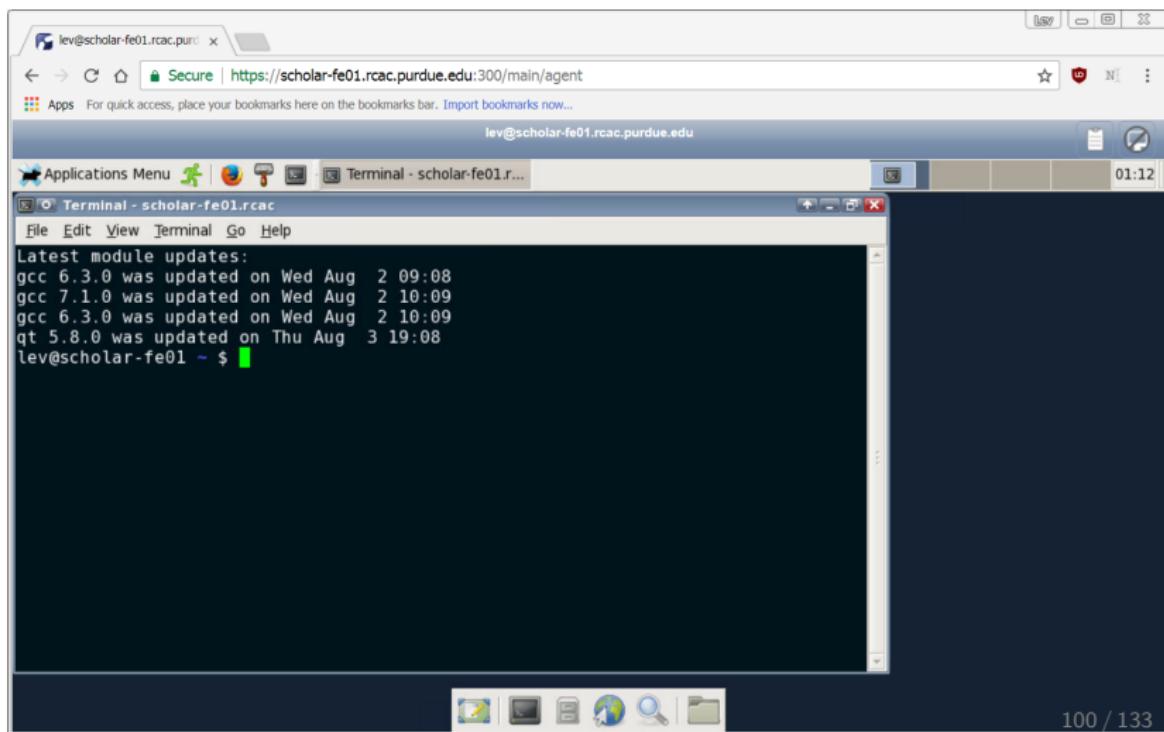
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Open terminal from Applications Menu:



Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

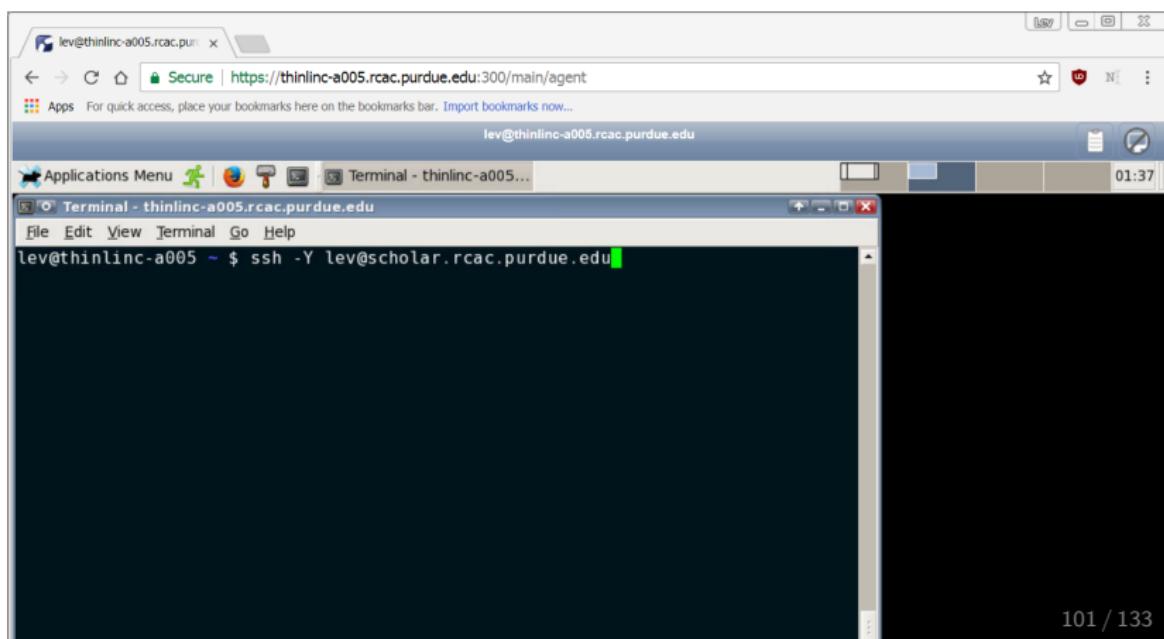
Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

If you were coming through thinlinc.rcac.purdue.edu,
don't forget to go to Scholar (not needed if coming through
desktop.scholar.purdue.edu):



Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

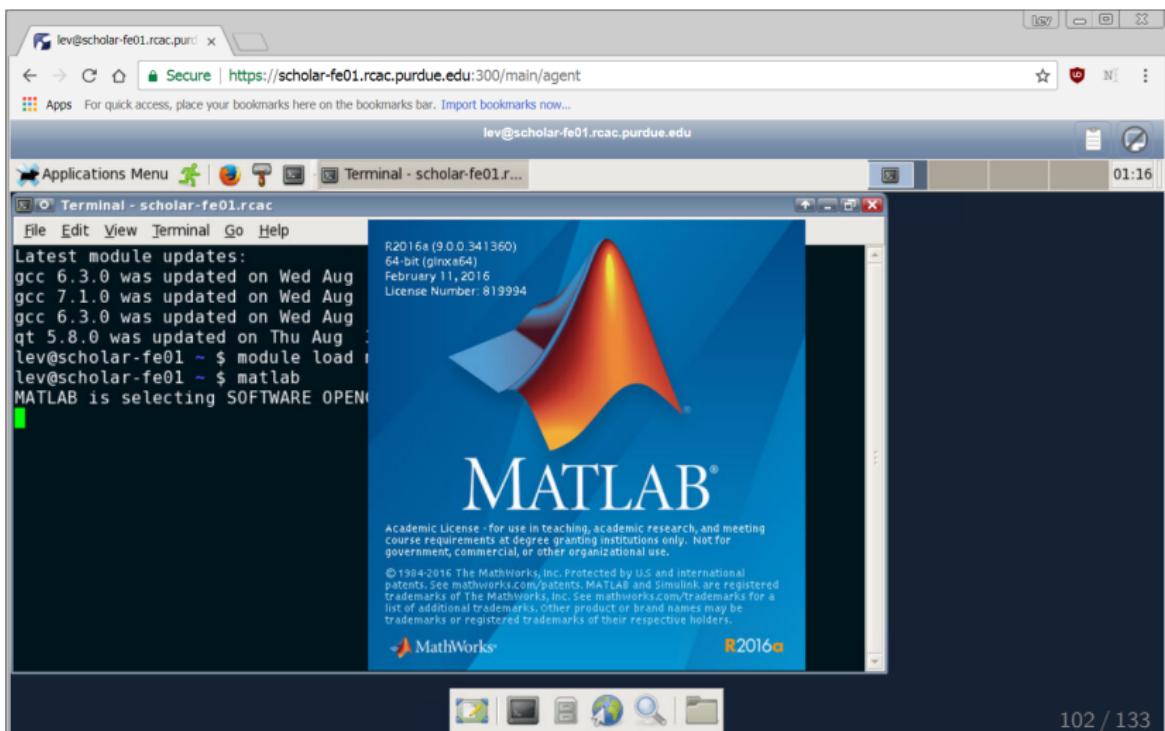
Interactive Jobs and Thinlinc

Storage

Globus

Workflows

module load matlab and start Matlab program:



Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

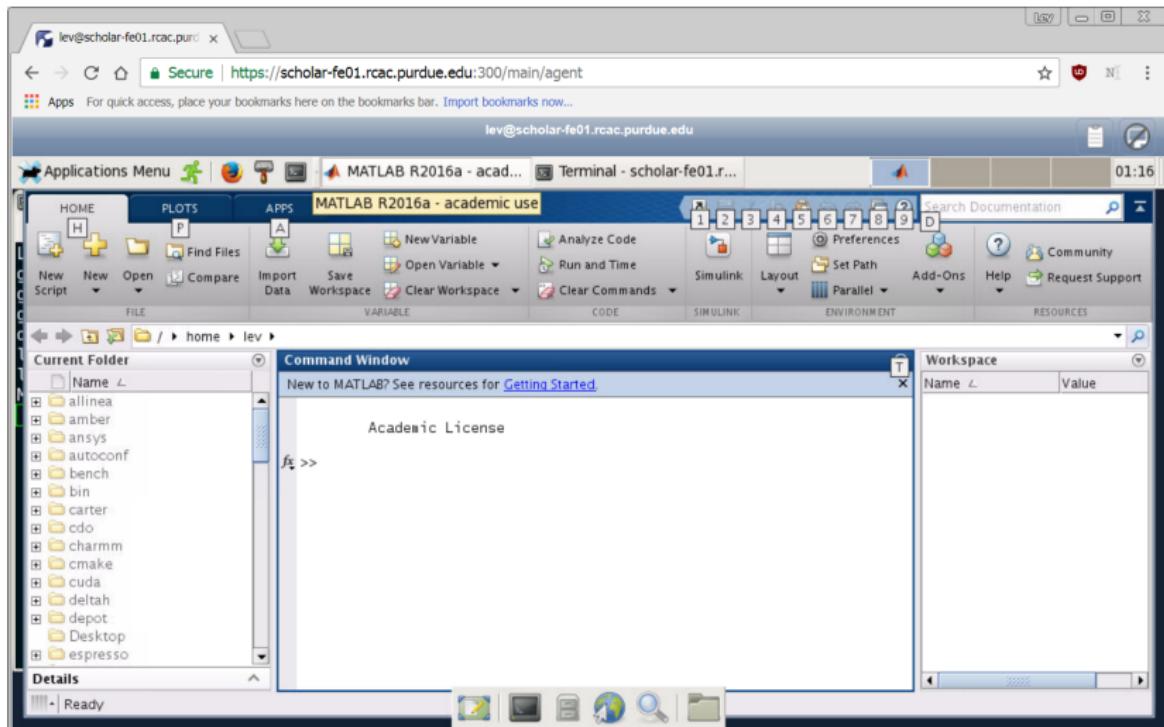
Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows



Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

But wait, this just ran on the front-end! Let's get onto a node and try again:

```
$ qsub -I -X -q scholar -l nodes=1:ppn=20,walltime=00:10:00
qsub: waiting for job 4218241.rice-adm.rcac.purdue.edu to start
qsub: job 4218241.rice-adm.rcac.purdue.edu ready

$ echo $DISPLAY
localhost:50.0
$ hostname
rice-a260.rcac.purdue.edu
$ module load matlab
$ matlab
```

Storage

Globus

Workflows

Interactive Jobs and Thinlinc

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

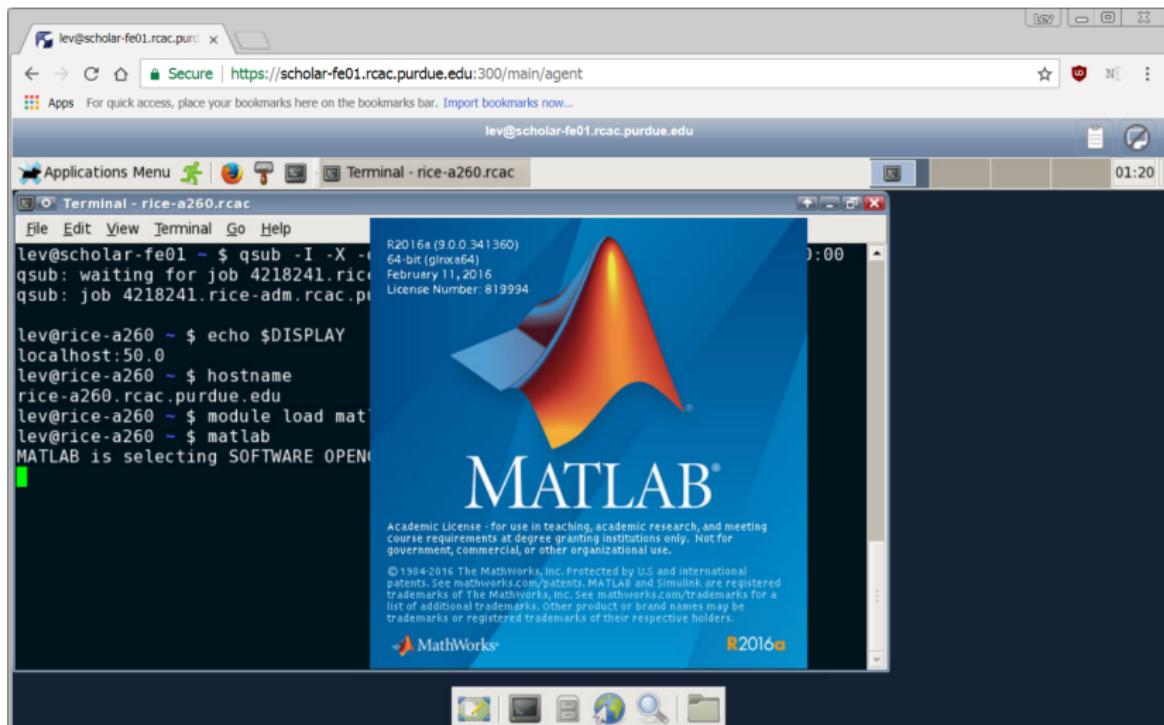
Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows



Storage

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Storage

Storage

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

ITaP Research clusters have several types of storage available - each has its own purpose and use case.

- \$HOME vs \$RCAC_SCRATCH
- Local disk (/tmp) on each individual node
- Research Data Depot
- Fortress archive

	\$HOME	Scratch	/tmp	/depot	Fortress
Capacity	25 GB	100 TB	~250 GB	100 GB+	unlimited
Resilient to hardware failures	yes	yes	no	yes	yes
Resilient to human errors	yes (snapshots)	no	no	yes (snapshots)	no
Subject to purging	no	yes	yes	no	no
Performance	medium	high	medium	medium	very slow
Designed for running jobs	no	yes	no	in moderation	no
Common within cluster	yes	yes	no	yes	yes (hsf/htar)
Common across clusters	yes	no	no	yes	yes (hsf/htar)
Globus endpoint	yes	yes	no	yes	yes

Storage

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Data Depot - a high-capacity, fast, reliable and secure data storage service.

- Basically, a "home directory for labs"
- Redundant, fast, flexible, accessible from anywhere
- Don't have to own nodes
- Data Depot space is mounted and accessible from all nodes and clusters

To learn more: www.rcac.purdue.edu/storage/depot/

Storage

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Fortress - a large, long-term storage system

- Robotic tape library - offers vast amounts of storage for cheap
- HSI - FTP-style interface without requiring any special user knowledge.
- HTAR - a utility to aggregate a set of files into a single tar archive directly into Fortress.
- Does not support direct FTP, SFTP or SCP transfers

To learn more: www.rcac.purdue.edu/storage/fortress/

Storage

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Globus – a powerful and easy to use file transfer service.
Transfer large amounts of research data to and from campus,
or share data with collaborators around the country or around
the globe.

Purdue Globus Portal: transfer.rcac.purdue.edu

Storage

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Types of storage and when to use them: capacity, speed, longevity – pick any two

ITaP Research Computing offerings are designed around these scenarios:

- Codes, executables, scripts – develop and store in \$HOME
- Lab-wide codes, executables, scripts, settings in /depot
- Produce/analyze data in \$RCAC_SCRATCH
- Permanently store results in Fortress and/or in /depot

`myquota` command

Use variables instead of explicit paths

Snapshots are not a substitute for backups!

Storage

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Some more tips:

- Avoid running jobs from \$HOME
- Run demanding jobs from \$RCAC_SCRATCH
- Avoid frequent I/O when possible
- Minimize simultaneous I/O from many processes
- Learn to recognize/avoid other stressors (default 1s on big directories)
- Know when its time to learn/use parallel I/O

Globus

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Globus

Globus

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Globus enables file transfers between two sites ("endpoints")

- Like hiring movers to pack your house and get it to new location
- You don't have to be involved
- Get an update when it completes / fails
- Goes directly between the two fast sites
- Doesn't use your laptop's bandwidth!
- Endpoints on all of our clusters, many other Universities/labs and even on your own laptop!
- Sites must set up Globus on their side, so not everyone has it but worth checking!

Globus

Vocabulary

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows



Globus

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Exercise:

- Log into Globus:
- transfer.rcac.purdue.edu

- Vocabulary
- Logging In
- Modules
- Types of Work
- Good Citizenship
- Jobs, Queues, and PBS
- Running Jobs
- Job Monitoring
- Interactive Jobs and Thinlinc
- Storage
- Globus**
- Workflows



PURDUE
UNIVERSITY.

 globus

[Globus Account Log in](#)

Log in to use Purdue Globus Web App

Use your existing organizational login

e.g. university, national lab, facility, project, Google or [Globus ID](#)

(Your Globus username and password used prior to February 13, 2016 is now Globus ID)

Purdue University Main Campus



Continue



Globus uses CILogon to enable you to Log In from this organization. By clicking Continue, you agree to the [CILogon privacy policy](#) and you agree to share your username, email address, and affiliation with CILogon and Globus. You also agree for CILogon to issue a certificate that allows Globus to act on your behalf.

Didn't find your organization? Then use Globus ID to [sign up](#).

Globus

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

The screenshot shows a web browser window for the URL <https://transfer.rcac.purdue.edu/globus-app/transfer>. The page header features the Purdue University logo and the text "Information Technology at Purdue Research Computing (RCAC)". Below the header is a navigation bar with links for "Manage Data", "Groups", "Support", and "Account". A secondary navigation bar below the main one includes links for "Transfer Files", "Activity", "Endpoints", "Bookmarks", "Publish", and "Console".

The main content area is titled "Transfer Files". It contains two large input fields for "Endpoint" and "Path", each with a "Go" button and a star icon. Between these fields are two arrows: a left arrow pointing left and a right arrow pointing right. Below each input field is a message: "Start by selecting an endpoint." on the left and "Start by selecting an endpoint." on the right.

On the right side of the page, there is a "RECENT ACTIVITY" section showing counts for circular and downward arrows, each with a value of 0. There is also a "Go" button next to the second endpoint input field.

Globus

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Endpoint Cancel

Purdue Structural Biology - Saturn
owner: purduestbio@globusid.org
Purdue Cryo-EM Facility

Purdue Research Computing - Home Directories
owner: purdue@globusid.org
This endpoint provides access to home directories storage on all Purdue clusters - /home

Purdue Research Computing - Data Depot
owner: purdue@globusid.org
This endpoint provides access to the Purdue Research Data Depot storage - /depot

Purdue Radon Cluster
owner: purdue@globusid.org
This endpoint provides access to scratch storage on the Purdue Radon cluster - /scratch/radon

Purdue Peregrine1 cluster
owner: purdue@globusid.org
This endpoint provides access to scratch storage on the Purdue Peregrine1 cluster - /scratch/peregrine1

Purdue Hathi Cluster
owner: purdue@globusid.org
This endpoint provides access to scratch storage on the Purdue Hathi cluster - /hadoop/mnt/user

Purdue Hammer Cluster
owner: purdue@globusid.org
This endpoint provides access to scratch storage on the Purdue Hammer cluster - /scratch/hammer

Globus

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Transfer Files | Activity | Endpoints | Bookmarks | Publish | Console

Transfer Files

The screenshot shows the Globus Transfer Files interface. At the top, there are two sets of input fields for endpoints and paths. The left pane displays a file browser for 'Purdue Research Computing - Home Dire' with a path of '/~/' and a single folder named 'unix101-2016'. The right pane is currently empty and contains the text 'Start by selecting an endpoint.'

RECENT ACTIVITY ○ 0 ▽ 0 ○ 0

Endpoint: Purdue Research Computing - Home Dire Path: /~/ Go

Endpoint: Path: Go

select all up one folder refresh list

Folder unix101-2016

Start by selecting an endpoint.

Globus

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Endpoint X Cancel

🔍

Purdue Structural Biology - Saturn owner: purduestbio@globusid.org Purdue Cryo-EM Facility
Purdue Research Computing - Home Directories owner: purdue@globusid.org This endpoint provides access to home directories storage on all Purdue clusters - /home
Purdue Research Computing - Data Depot owner: purdue@globusid.org This endpoint provides access to the Purdue Research Data Depot storage - /depot
Purdue Radon Cluster owner: purdue@globusid.org This endpoint provides access to scratch storage on the Purdue Radon cluster - /scratch/radon
Purdue Peregrine1 cluster owner: purdue@globusid.org This endpoint provides access to scratch storage on the Purdue Peregrine1 cluster - /scratch/peregrine1
Purdue Hathi Cluster owner: purdue@globusid.org This endpoint provides access to scratch storage on the Purdue Hathi cluster - /hadoop/mnt/user
Purdue Hammer Cluster owner: purdue@globusid.org This endpoint provides access to scratch storage on the Purdue Hammer cluster - /scratch/hammer

Globus

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Transfer Files | Activity | Endpoints | Bookmarks | Publish | Console

Transfer Files

The screenshot shows the Globus Transfer Files interface. At the top, there are two endpoint configurations:

- Purdue Research Computing - Home Dir:** Path is set to `/~`. The folder `unix101-2016` is selected.
- Purdue Radon Cluster:** Path is set to `/scratch/radon/`.

Below the endpoints are two file browser panes:

- Left Pane (Source):** Shows the contents of the `unix101-2016` folder. It contains a single item: `Folder`.
- Right Pane (Destination):** Shows a list of items from `a` to `t`, each labeled `Folder`.

At the bottom right of the interface, there is a "RECENT ACTIVITY" section with three circular icons: one orange, one green, and one blue, each with a count of 0.

Globus

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows

Transfer Files | Activity | Endpoints | Bookmarks | Publish | Console

Transfer Files

The screenshot shows the Globus Transfer Files interface. At the top, there are two endpoint configurations:

- Endpoint 1:** Purdue Research Computing - Home Dir. Path: `/~`. Buttons: Go, Back, Forward.
- Endpoint 2:** Purdue Radon Cluster. Path: `/scratch/radon/r/rherban/`. Buttons: Go, Back, Forward.

Below the endpoints are two file browser panes. The left pane shows a single folder named `unix101-2016`. The right pane is currently empty. Both panes have standard file browser controls: select none/all, up one folder, refresh list, and a three-dot menu icon.

Globus

Vocabulary

Transfer Files | Activity | Endpoints | Bookmarks | Publish | Console

Logging In

Modules

Types of Work

Good Citizenship

Jobs, Queues, and PBS

Running Jobs

Job Monitoring

Interactive Jobs and Thinlinc

Storage

Globus

Workflows

Transfer Files

RECENT ACTIVITY 1 0 0

Transfer request submitted successfully. Task id: 2e22df00-961f-11e6-b0a4-22000b92c261

Endpoint Purdue Research Computing - Home Dire



Path /~/

Go

Endpoint Purdue Radon Cluster



Path /scratch/radon/r/rherban/

Go

select none

up one folder

refresh list



unix101-2016

Folder

select all

up one folder

refresh list



Globus

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Can share select folders with outside collaborators: no need to have Purdue ID.

Can set up an endpoint on your personal laptop:
www.globus.org/globus-connect-personal

Workflows

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Workflows

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Unless you are running a one-off simulation, you should stop and think about the workflow:

- What's the optimal resource request?
- Experimentation - scaling studies
- Do I break my workflow into multiple jobs?
- How does my data flow?
 - From original source
 - Working spaces
 - Post-processing/analysis
 - Archives

Workflows

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Sample workflow:

- Grab dataset before batch of jobs:
`rsync -rav /depot/group/data/experiment1
/scratch/radon/u/user/`
- Submit batch of jobs
- Save to `/scratch/radon/u/user/results`
- Copy results back to Depot or Fortress:

```
ssh user@cli.globusonline.org transfer -- purdue#radon:/scratch  
/radon/u/user/results purdue#depot:/depot/group/data/  
experiment1-results
```

- Requires ssh key setup with Globus
- Happens in the background; doesn't waste job time

Workflows

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

This requires SSH keys to script unattended Globus transfers:

- Create GlobusID
- Go to globus.org and create Globus ID
- Log in with it and link your Purdue Credentials
- Get SSH key and pair it with your Globus ID

Workflows

Vocabulary

Logging In

Modules

Types of Work

Good
Citizenship

Jobs, Queues,
and PBS

Running Jobs

Job
Monitoring

Interactive
Jobs and
Thinlinc

Storage

Globus

Workflows

Can use existing key:

```
$ cd ~  
$ ls -las .ssh  
$ cat .ssh/id_rsa.pub
```

Or generate new key:

```
$ ssh-keygen -t rsa -f .ssh/id_globus -N ""  
$ cat .ssh/id_globus.pub
```

Copy the public key.

Workflows

- Vocabulary
- Logging In
- Modules
- Types of Work
- Good Citizenship
- Jobs, Queues, and PBS
- Running Jobs
- Job Monitoring
- Interactive Jobs and Thinlinc
- Storage
- Globus
- Workflows



Randy Herban (rherban2@globusid.org)

[Back to Account Information on the Purdue Globus Web App](#)

[Home](#) | [Log Out](#)

Manage SSH and X.509 Keys

[Add a New Key](#)

You have the following key attached to your Globus ID account:

SSH Public Key

[expand](#) [delete](#) [rename](#)

Workflows

Vocabulary
Logging In
Modules
Types of Work
Good Citizenship
Jobs, Queues, and PBS
Running Jobs
Job Monitoring
Interactive Jobs and Thinlinc
Storage
Globus
Workflows



Randy Herban (rherban2@globusid.org)

[Back to Account Information on the Purdue Globus Web App](#)

[Home](#) | [Log Out](#)

Add a New Key

[Back to Key Management](#)

Alias

My Laptop's Key

Type

SSH Public Key

Allows you to use the [Globus command line interface](#) via any standard `ssh` client. Some Globus-based applications may also use this SSH authorized key to provide access to their resources.

X.509 Credential

Allows you to use the [Globus command line interface](#) via a `gsissh` client, and to activate endpoints with a delegated proxy of this X.509 credential via `gsissh` to the CLI.

Body

[Add Key](#)

[Cancel](#)