

HPC Clusters at ISU

Best practices

October 5, 2021



Agenda

Introduction

Logins

Storage

Fast data transfer

Using Job scripts

Software environment

Tips for improving code efficiency

Q&A



Introduction

<https://www.hpc.iastate.edu/guides/introduction-to-hpc-clusters>

- Education Cluster hpc-class

- <https://www.hpc.iastate.edu/guides/classroom-hpc-cluster>
- Used for classwork
- Instructors request access for students taking the class

- Free Research Cluster Condo

- <https://www.hpc.iastate.edu/guides/condo>
- Researchers request free access for their groups and manage user access for their groups

- Research Cluster Nova

- <https://www.hpc.iastate.edu/guides/nova>
- PIs purchase nodes and storage on the cluster and manage user access for their groups
- ResearchIT manages user access for LAS users



Logins

- Need ssh client to login
 - Linux and MacOS - open terminal
 - Windows – open Windows PowerShell
 - Older Windows – install putty from <https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>
- Must be
 - on the iastate.edu network or
 - using ISU VPN installed from <http://vpn.iastate.edu>

Logins

- `ssh <ISU NetID>@<clustername>.its.iastate.edu`
 - `ssh kraeva@hpc-class.its.iastate.edu`
- On first login from any machine, you will first be asked if you want to store the cluster's host key. Answer YES
- Type ISU password when prompted for Password
- Condo and Nova use Google Authenticator (GA) as second login factor
 - At the first attempt to ssh to Condo or Nova, a new GA Account is created, and an email with the instructions is sent to your ISU address
 - The GA Account is shared across Condo and Nova clusters



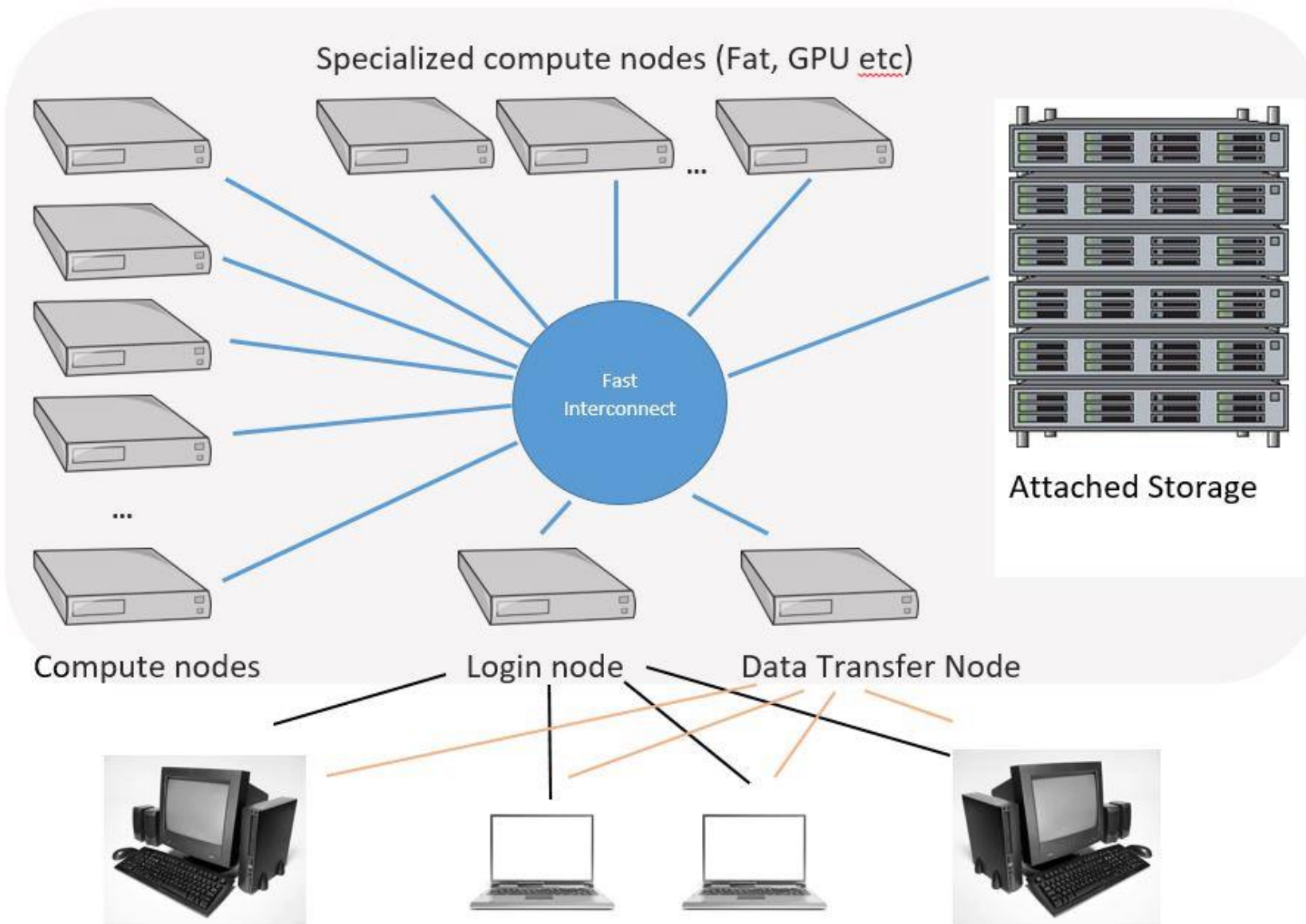
Logins (GA)

- Install Google Authenticator or Twilio Aussy app on your smartphone/table/computer
- Can install on multiple devices
- To add your GA account to the app, scan the QR code or copy in the multi-character key from the email
- ssh to the cluster and enter the 6-digit code from the app when prompted for Verification
- DO NOT reuse the same code, wait for a new one to be generated (a new code is generated every 30 seconds)



Logins (GA Self Reset)

- Do you want to store a phone number to allow you to reset your Google authenticator application (Y/N)?
- Enter Y
- Enter the 10-digit number where you will get a text message when you need to reset your GA account
- To save a different number later, issue `phone-collect.sh`
- To perform self reset, go to <https://hpc-ga1.its.iastate.edu/reset/>



Help!

- <https://www.hpc.iastate.edu/>
- <https://www.hpc.iastate.edu/guides>
- <https://www.hpc.iastate.edu/faq>
- hpc-help@iastate.edu
- researchit@iastate.edu for help with Pronto
- `man <command>`
- `<command> --help`

Storage (Quotas)

- Home directories: 5GB
`/home/<NetID>`
Use for configuration files
- Work directories on Condo/Nova: 160GB for free tier groups or whatever was purchased by the group
`/work/<group_working_directory>/group_storage_usage`
Try to stay below 70% utilization
- `/ptmp`: no quotas, old files regularly deleted

All these directories are available on all cluster nodes



Storage (Work)



`cdw` command (from ISU) will take you to your group working directory

for LAS users: `cd /work/LAS/<PI_NetID>-lab`

Permissions are set to: `drwxrws--T`



to see permissions, issue: `ls -la`

users from other groups can't access your work directory

other users in your group can't delete your files

`/work` uses ZFS file system



- snapshots
- compression (`du --apparent-size` shows the uncompressed usage)



Storage (Temporary)

/ptmp

No backup

files deleted after 90 days from creation on Condo/Nova or
after semester ends on hpc-class



```
mkdir /ptmp/<NetID> or mkdir /ptmp/<groupname>/<NetID>
```

\$TMPDIR and \$PTMPDIR

Job specific space on the node(s) assigned to the job

Can be accessed only while the job is running



Storage (Archive)

Large Scale Storage (LSS)

<https://researchit.las.iastate.edu/large-scale-storage-lss/lss>

Accessible from the data transfer nodes: condodtn and novadtn

Issue `kinit` if can't access your group LSS storage

LSS is backed off-site, \$40/TB/year

Box, Google Drive and other cloud storage

use `rclone` to sync data between HPC clusters (DTN nodes) and Google Drive

<https://www.hpc.iastate.edu/guides/rclone>

File Transfers

Use data transfer nodes: hpc-class-dtn, condodtn and novadtn



scp, sftp

```
scp aa kraeva@condodtn.its.iastate.edu:
```

Can use WinScp, Filezilla ... but we recommend:

Globus Connect

<https://www.hpc.iastate.edu/guides/globus-online>



Login node

Each cluster has its own login node which

- Is shared among multiple users
- Is to be used for editing files, submitting jobs, and viewing output
- Can be used for compiling IF it does not interfere with the usability of the node
- Should NOT be used for file transfers
- Should NOT be used to run memory or compute intensive processes



SLURM workload manager

SLURM Workload Manager

- Receives user's requests for cluster resources (nodes, cores, memory, GPUs)
- Schedules jobs

Jobs can be

- Interactive
- Batch

SLURM workload manager

Use interactive jobs to

- Install programs, software packages
- Debug a program or a process

```
sinfo -p debug
```

```
sinfo
```



```
salloc -N 1 -n 1 -t 1:00:00 -p debug
```

When done, issue `exit`



SLURM workload manager

Otherwise submit a batch job



Use Slurm Script Generator in the cluster user guide at <https://www.hpc.iastate.edu/guides/>

- Unless your application is written to use multiple nodes (e.g. Fluent, OpenFoam or GAMESS), select one node
- If the application does not use threads, request one core
- No need to specify partition unless requesting a GPU or a high memory node
- Can request more memory than default
- Specify long enough runtime – we do not extend runtimes

SLURM workload manager

Copy the generated script into clipboard

Open a new file on a cluster and paste from the clipboard

Add commands at the end of the file

Save the file



```
sbatch <myjob>
```

To see jobs in the queue:

```
squeue
```

```
squeue -j <job_number>
```

```
squeue -u <NetID>
```



SLURM workload manager



To get an estimate on when a job will start, add `--start:`

```
queue --start
```

The real time may be earlier if jobs end before the time they reserved or may be later if higher priority jobs are submitted after your job is submitted.

Job priority is based on the percentage of your group's allocation that has been used

- All groups on hpc-class and Condo have equal allocation
- Allocation on Nova is based on the # of nodes purchased by the group
- Running on GPU and high-memory nodes costs more



SLURM workload manager

To get information about running job use `qtop` or `sstat`:



```
qtop <job_number>
```

```
sstat --format=MaxRSS%20,AveCPU -j <job_number>
```

See `man sstat` for instructions on how to use `sstat` and examples

After job ends check job efficiency:



```
seff <job_number>
```

For more details, use `sacct` command

See `man sacct` for instructions on how to use `sacct` and examples



Software environment



Software can be accessed through environment modules

```
module avail [<software_name>]
```

```
module spider [<software_name>]
```

```
module load <module_name>
```

```
module list
```

```
module purge
```

Some modules load other modules

Usually more than one version of a software package is available

Software environment

Loading a module will set some environment variables (e.g. `$PATH`) that allow to use software installed as the module.



`which` command shows location of the executable

```
which gcc
```

When using Python, pay attention to the version. If using Python 3, explicitly type `python3` instead of simply `python`, otherwise system `python2` may be used (the behavior depends on the module).

```
module help <module_name>
```



Software environment

How do I Install my own R package?

<https://www.hpc.iastate.edu/faq#Rpack>

How do I install my own Python package?

<https://www.hpc.iastate.edu/faq#python-package>

Be aware of the home directory quota, opt to install software in the group working directory.

If you have software installed in the home directory, you can move it to /work and create a symbolic link:

```
cd
mv R /work/<group_name/
ln -s /work/<group_name/R R
```

Compiling, debugging and profiling

If you need to compile software, you can use GNU or Intel compilers. PGI compilers are also available.

Intel compilers generally provide faster executables.

Use `-O3 -xHost` compiler options to generate faster code.

If need to debug a program, compile it with `-g` option.

To debug a program, use gdb or ddt (<https://www.hpc.iastate.edu/guides/using-ddt-parallel-debugger--map-profiler-and-performance-reports>)

To improve program performance, profile it either with ARM MAP profiler or Intel tools.

Tips for improving code efficiency

Eliminate unnecessary work

With nested if statements, put most likely to fail first

Keep branches outside of innermost loop (if, case)

In matlab or similar programs use array statements

In Perl, associative arrays can be very efficient

In Python, dictionaries and sets can be very efficient

Use libraries where possible

Go parallel within a server (36 cores on nova)

threads, parallel command, OpenMP

Go parallel across servers using MPI



Miscellaneous Tips

Regular users don't have `sudo` privileges

To run GUI software on a cluster, install an X server on your machine and ssh to the cluster with `-X` or `-Y` option

On a Mac computer can install XQuartz (<https://www.xquartz.org/>)

On a Windows computer, look for WSL

It may run slow (performance depends on the network speed)

Machine Learning Singularity Image

Instructions: <https://researchit.las.iastate.edu/machine-learning-singularity-image>



Best Practices

- Use Globus for transfers.
- Use data transfer nodes if you use sftp or scp or other data transfer tools.
- Do not run intensive tasks on the login node, use `salloc`.
- Do not request more cores/nodes than needed for the job. Use `seff` on finished jobs to estimate memory and time requirements for future jobs.
- Keep files that you need in `/work`, not in `/ptmp`, `/tmp`, `/var/tmp` or on compute nodes.
- For I/O intensive tasks, copying files to `$TMPDIR` and using that directory will greatly improve I/O versus accessing `/work` or `/ptmp` over the network, however you will need to copy files back before the job ends.

Upcoming Workshops

- October 14, 11-12:30PM [Building and using containers on the HPC clusters](#)
- October 18, 11-1PM [Using Matlab Parallel Server on the HPC Clusters](#)
- October 29, 11-12:30PM [Using JupyterHub on ISU HPC clusters](#)

You can also find this information at <https://www.hpc.iastate.edu/news>

Survey at https://iastate.qualtrics.com/jfe/form/SV_3dtHKeRLAqOMpQW

The background features a white surface with a network of thin, stylized circuit lines in yellow, orange, and red. These lines meander across the frame, some ending in small circular nodes. A large, solid red rectangle is positioned in the center-left, partially obscuring the circuit lines. In the upper right, a grid of faint, light-blue squares is visible. The overall aesthetic is clean and modern, suggesting a technological or digital theme.

Questions?