



Introduction to HPC Using Sapelo Cluster at GACRC

Level 1/2

Georgia Advanced Computing Resource Center (GACRC)

EITS/University of Georgia

Zhuofei Hou zhuofei@uga.edu



Outline

- GACRC
- Sapelo Cluster
 - Overview
 - Computing Resources
 - Storage Environment
 - Software Environment
- Job Submission Workflow
- Work on Sapelo
- Guideline and Practical Tips



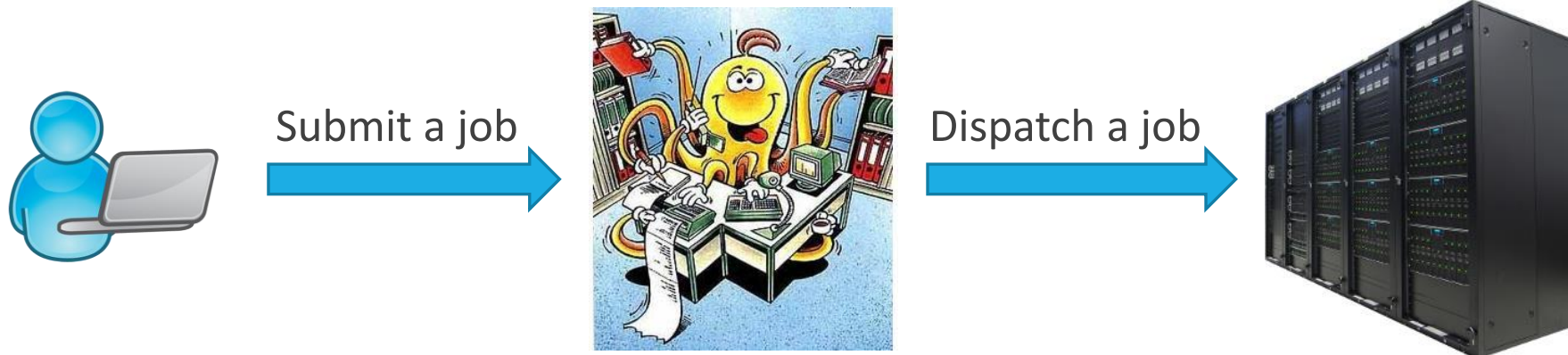
GACRC

- We are a high-performance-computing (HPC) center at UGA
- We provide to the UGA research and education community an advanced computing environment:
 - HPC computing and networking infrastructure located at the Boyd Data Center
 - Comprehensive collection of scientific, engineering and business applications
 - Consulting and training services
- <http://wiki.gacrc.uga.edu> (GACRC Wiki)
- https://wiki.gacrc.uga.edu/wiki/Getting_Help (GACRC Support)
- <http://gacrc.uga.edu> (GACRC Web)

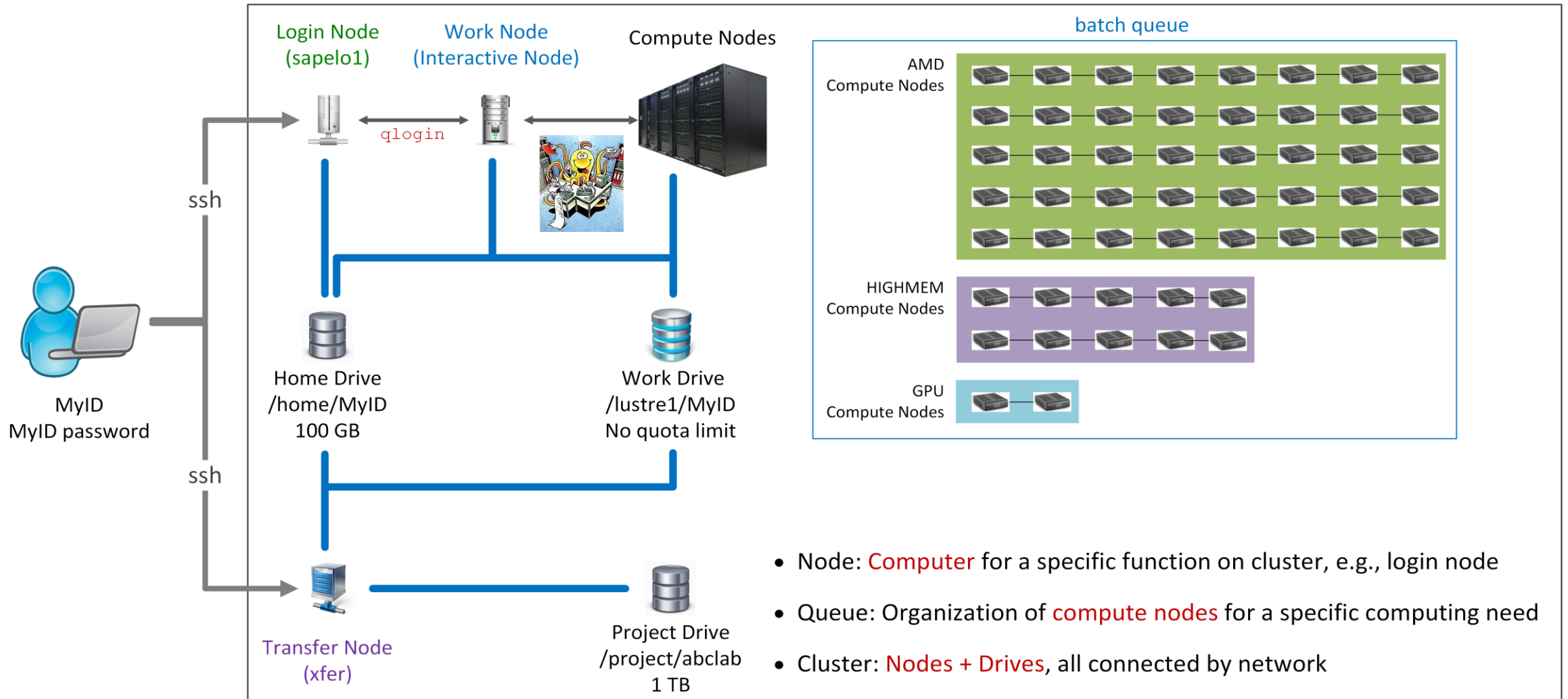
What is a Cluster?

When you work on cluster, **3** roles are working:

- User: logon and submit a job
- Queueing System: dispatch a job to run on cluster
- Cluster: run a job



Sapelo Cluster Diagram





Overview

Sapelo is a Linux high performance computing (HPC) cluster:

- Operating System: 64-bit CentOS Linux 6.5
- You can log on (open connection) to:

Login (username@sapelo1.gacrc.uga.edu) : edit script, submit batch job (not suggested!)

Transfer (username@xfer.gacrc.uga.edu) : transfer, compress, package data

- Login $\xrightarrow{\text{qlogin}}$ Work Node : edit script, submit batch job, run interactive job



Computing Resources

Queue	Node Feature	Total Nodes	RAM(GB) /Node	Max RAM(GB) /Single-node Job	Cores /Node	Processor	GPU	GPU Cards /Node	InfiniBand
batch (regular)	AMD	112	128	126	48	AMD Opteron	N/A		Yes
		4	256	252					
	HIGHME M	6	512	504					
		4	1024 (1)	997	28	Intel Xeon	N/A		
	1024 (3)								
	GPU	2	128	126	16		NVIDIA K40m	8	
abcnode (buy-in)	85	variable	variable	variable	variable	variable	variable		
training_q	training	1	128	126	48	AMD	N/A		

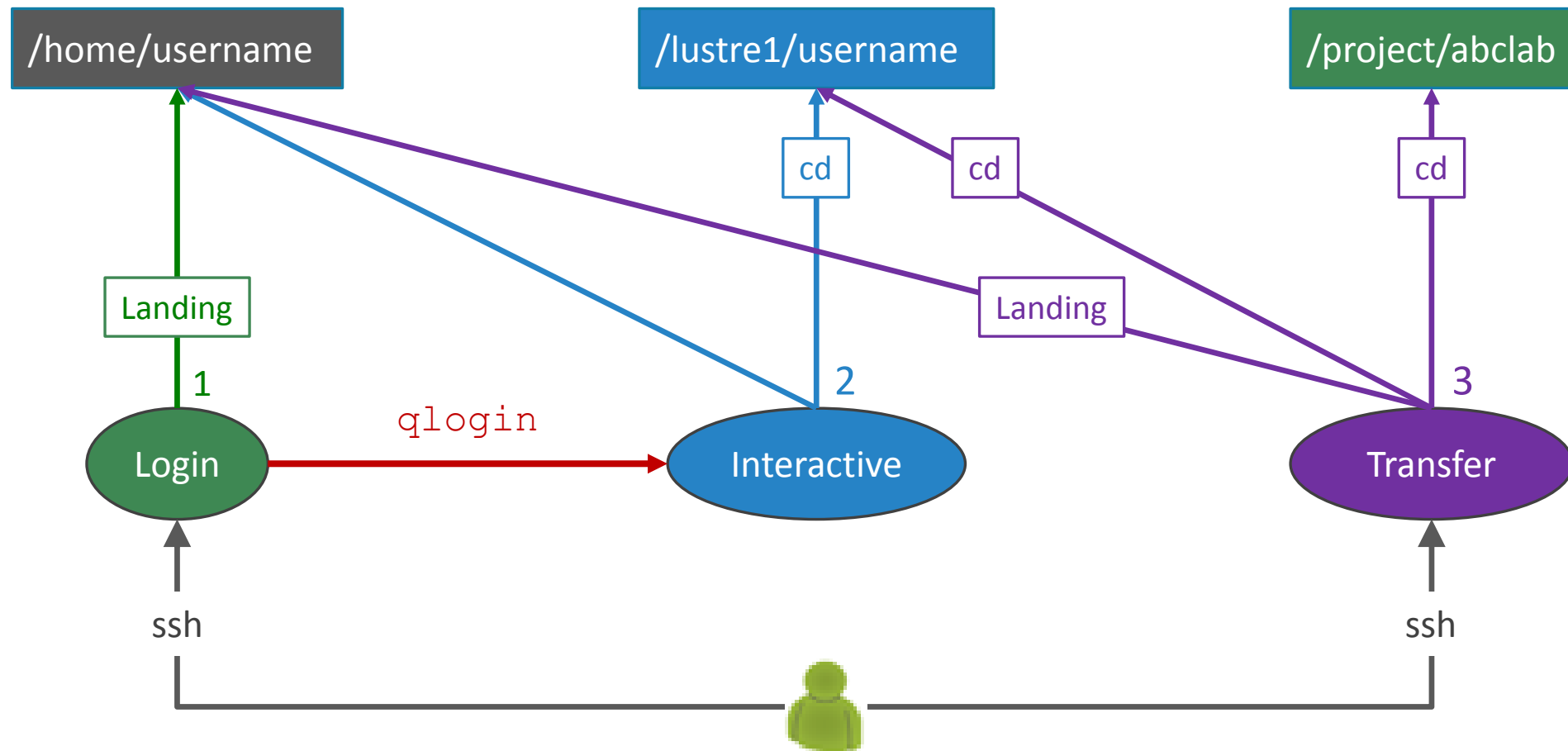


Storage Environment

4 Directories	Role	Quota	Accessible from	Intended Use	Notes
/home/MyID	Home	100GB	Login node Work node Transfer node	Static data: 1. Scripts, source codes 2. Local software	Snapshots is turned on
/lustre1/MyID	Global Scratch	No Limit	Work node Transfer node	Current job data: data being read/written by running jobs	User to clean up! Subject to deletion in 30 days*
/project/abclab	Storage	1TB (Initial)	Transfer node	Temporary data parking: non-current active data	Group sharing possible
/tmp/lscratch	Local Scratch	N/A	Compute node	Jobs with heavy disk I/O	User to clean up! When job exists from the node!

1. /usr/local/apps : Software installation directory ; /db : Bioinformatics database installation directory
2. * denotes component or policy to be implemented in the future

Storage Environment (Cont.) – Accessing Directories from Nodes





Storage Environment (Cont.) - Intended Use of Directories

I/O speed

Fast



Slow

`/lustre1/username/`

← **Current Job Data** being used by **current running jobs** on cluster

`/home/username/`

← **Static Data** being used frequently and **not being modified often**, e.g., scripts or local software

`/project/abclab/`

← **Non-current Active Data** to be analyzed in the future, e.g., 2 months

User Local Computer

← **Final Data**, e.g., final outputs and results



Storage Environment (Cont.)

7 Main Functions	Related Filesystem	Related Node
Login Landing	/home/username (Home) (Always!)	Login or Transfer
Batch Job Submitting	/home/username (Home)	Login
	/lustre1/username (Global Scratch) (Suggested!) /home/username (Home)	Work
Interactive Job Running	/lustre1/username (Global Scratch) /home/username (Home)	
Data Transferring, Archiving , Compressing	/lustre1/username (Global Scratch) /home/username (Home)	Transfer
Non-current Active Data Temporary Parking	/project/abclab (Project Storage)	
Job Data Temporarily Storing	/lustre1/username (Global Scratch) /tmp/lscratch (Local Scratch)	Compute



Software Environment

The cluster uses **environment modules** to define the various paths for software packages (more than 600 and expanding daily!)


- `module avail` : List all available modules installed on cluster
- `module load` : Load a module into your working environment
- `module list` : List modules currently loaded
- `module unload` : Remove a module
- `module spider` : Search a module



Software Environment (Cont.)

```
File Edit View Search Terminal Help
zhuofei@uga-2f0f976:~ $
zhuofei@uga-2f0f976:~ $ module avail

----- /usr/local/modulefiles -----
Core/StdEnv
Data/cache/moduleT.new
Data/cache/moduleT (D)
Data/system.txt
R/3.0.3
R/3.1.2
R/3.2.1
R/3.2.3
R/3.3.1
R/3.3.2 (D)
StdEnv (D)
abyss/1.5.2
abyss/1.9.0 (D)
aeneas/1.7.3
age/0.4
albacore/1.2.5-cp35
albacore/2.0.2-cp35 (D)
allmaps/0.6.6
amber/14-at15
amber/14
amber/14-20160315 (D)
amos/3.1.0
anaconda/2-4.1.1
gobject-introspection/1.50.0 (D)
grace/5.1.25
grads/2.1a3/gcc/4.7.4
graftm/0.9.5
graftm/0.10.1 (D)
graph-tool/2.19
graph-tool/2.22 (D)
graphviz/2.40.1
gromacs/5.0.2
gromacs/5.1.3-mpi (D)
gromacs/5.1.3 (D)
gsl/1.16/gcc/4.4.7-src
gsl/1.16/gcc/4.4.7
gsl/1.16/gcc/4.7.4
gsl/1.16/gcc/5.3.0 (D)
gsl/2.1/gcc447
gtk+/3.22.4
gtk-doc/1.25
hapcut2/20170126
haploview/4.2
harp/130306_105556
hdf5/1.10.0-patch1/gcc/5.3.0
hdf5/1.8.14/gcc/4.4.7
perl/5.24.0-thread (D)
petsc/3.7.6/intel140-mpi183-fc
petsc/3.7.6/intel140-mpi183 (D)
pgap/1.2.1
pgi/14.9
pgi/14.10
pgi/17.3 (D)
phase/2.1.1
phaser/1.2.0
phasetank/1.0
phenix/1.12
photoscan/1.3.4-openssl
photoscan/1.3.4 (D)
phrap/0.990329
phred/071220
phylip/3.69
phylobayes-mpi/1.7
phylogenie/11082017
phylonet/3.5.6
phylonet/3.6.0 (D)
phylosift/1.0.1
phyluce/1.5.0
phym1/3.0
```

 type key of q to quit from module avail



Software Environment (Cont.)

```
zhuofei@uga-2f0f976:~  
File Edit View Search Terminal Help  
zhuofei@uga-2f0f976:~ $  
zhuofei@uga-2f0f976:~ $ module list  
  
Currently Loaded Modules:  
 1) moab/8.1.1  2) StdEnv  ← once you log on, you will have two default modules  
  
zhuofei@uga-2f0f976:~ $ module load R/3.2.3  ← You can copy a complete name of a module from module avail, then paste here!  
zhuofei@uga-2f0f976:~ $ module list  
  
Currently Loaded Modules:  
 1) moab/8.1.1  2) StdEnv  3) R/3.2.3  
  
zhuofei@uga-2f0f976:~ $ module load R/3.3.1  
  
The following have been reloaded with a version change:  ← module load is smart to replace modules for you!  
 1) R/3.2.3 => R/3.3.1  
  
zhuofei@uga-2f0f976:~ $ module list  
  
Currently Loaded Modules:  
 1) moab/8.1.1      4) bzip2/1.0.6/gcc530  7) pcre/8.39/gcc530    10) proj.4/4.9.3/gcc/5.3.0  13) gcc/5.3.0          16) R/3.3.1  
 2) StdEnv         5) xz/5.2.2-gcc530    8) curl/7.47.1/gcc530  11) gsl/1.16/gcc/5.3.0    14) root/5.34.36  
 3) zlib/1.2.8/gcc530  6) java/jdk1.8.0_20  9) gdal/2.1.0/gcc/5.3.0  12) jags/4.2.0/gcc/5.3.0  15) geos/3.5.0/gcc447
```



Job Submission Workflow

1. Log on to **Login** node: `ssh MyID@sapelol.gacrc.uga.edu`
2. From **Login** node, transfer to **Work** node: `qlogin`
3. On **Work** node, change directory to global scratch: `cd /lustrel/username`
4. Create a working subdirectory for a job: `mkdir ./workDir`
5. Change directory to workDir: `cd ./workDir`
6. Transfer data from local computer to workDir: use `scp` or **SSH File Transfer** to connect **Transfer** node
Transfer data on cluster to workDir: log on to **Transfer** node and then use `cp` or `mv`
7. Make a job submission script in workDir: `nano ./sub.sh`
8. Submit a job from workDir: `qsub ./sub.sh`
9. Check job status: `qstat_me` or Cancel a job: `qdel JobID`



Step1: Log on to Login node - Mac/Linux using ssh

1. Open **Terminal** utility
2. Type command line: `ssh MyID@sapelo1.gacrc.uga.edu`
3. You will be prompted for your **MyID password** (when you type in the password, the prompt blinks and does not move)

```
zhuofei@s172-20-34-h20:~$  
zhuofei@s172-20-34-h20:~$  
zhuofei@s172-20-34-h20:~$ssh zhuofei@sapelo1.gacrc.uga.edu  
zhuofei@sapelo1.gacrc.uga.edu's password: █
```




Step1 (Cont.) - Windows

Download to install:

http://eits.uga.edu/hardware_and_software/software/

The screenshot shows the SSH Secure Shell 3.2.9 (Build 283) application window. A 'Connect to Remote Host' dialog box is open in the foreground. The dialog box contains the following fields and buttons:

- Host Name: **2** sapelo1.gacrc.uga.edu
- User Name: **3** zhuofei
- Port Number: 22
- Authentication Method: <Profile Settings>
- Buttons: **4** Connect, Cancel

At the bottom of the dialog box, there is a text box with the following text:

You will be prompted for your **MyID password** after step 4



Step2: From Login node go to Work node using qlogin

- Once you logged on Login node, your current directory will be your home directory:

```
zhuofei@uga-2f0f976:~ $ pwd  
/home/zhuofei
```

← that's my home dir on cluster!

- Type `qlogin` command on your command line:

```
zhuofei@uga-2f0f976:~ $ qlogin  
qsub: waiting for job 2983350.pbs.scm to start  
qsub: job 2983350.pbs.scm ready
```

```
zhuofei@n148:~ $
```

← now I am on a Work node with index of n148



Step2 (Cont.)

- Once you are on Work node, your current directory will be your home directory:

```
zhuofei@n148:~ $ pwd  
/home/zhuofei
```

← that's my home dir on cluster!

- Type `exit` command on your command line to go back to Login node:

```
zhuofei@n148:~ $ exit  
logout  
  
qsub: job 2983575.pbs.scm completed  
zhuofei@uga-2f0f976:~ $
```

← now I am back on Login node



Step3: On Work node, change directory to global scratch

- Use `cd` command to change your current directory to `/lustre1/MyID`

```
zhuofei@n148:~ $ cd /lustre1/zhuofei
zhuofei@n148:/lustre1/zhuofei $ pwd
/lustre1/zhuofei
```

← that's my working space on cluster!

- Use `ls` command to take a look in `/lustre1/MyID`

```
zhuofei@n148:/lustre1/zhuofei $ ls
user_test  workDir_Alex
```



Step4: Create a working subdirectory for a job

- Use `mkdir` command to make a subdirectory in `/lustre1/MyID`

```
zhuofei@n148:/lustre1/zhuofei $ mkdir workDir
zhuofei@n148:/lustre1/zhuofei $
```

- Use `ls` command to take a look in `/lustre1/MyID`

```
zhuofei@n148:/lustre1/zhuofei $ ls
user_test  workDir  workDir_Alex  ← a new folder workDir was created!
```



Step5: Change directory to workDir

- Use `cd` command to change your current directory to `/lustre1/MyID/workDir`

```
zhuofei@n148:/lustre1/zhuofei $ cd workDir
zhuofei@n148:/lustre1/zhuofei/workDir $
```

- Use `ls` command to take a look in `/lustre1/MyID/workDir`

```
zhuofei@n148:/lustre1/zhuofei/workDir $ ls
zhuofei@n148:/lustre1/zhuofei/workDir $
```

← It is empty!



Step6: Transfer data from local computer to workDir - Mac/Linux

1. You need to connect to cluster's Transfer node (xfer.gacrc.uga.edu)
2. Open Terminal utility on local computer to use `scp (-r) [Source] [Target]`

E.g. 1: working on local computer, from Local → workDir on cluster

```
scp ./file zhuofei@xfer.gacrc.uga.edu:/lustrel1/zhuofei/workDir/
```

```
scp -r ./folder/ zhuofei@xfer.gacrc.uga.edu:/lustrel1/zhuofei/workDir/
```

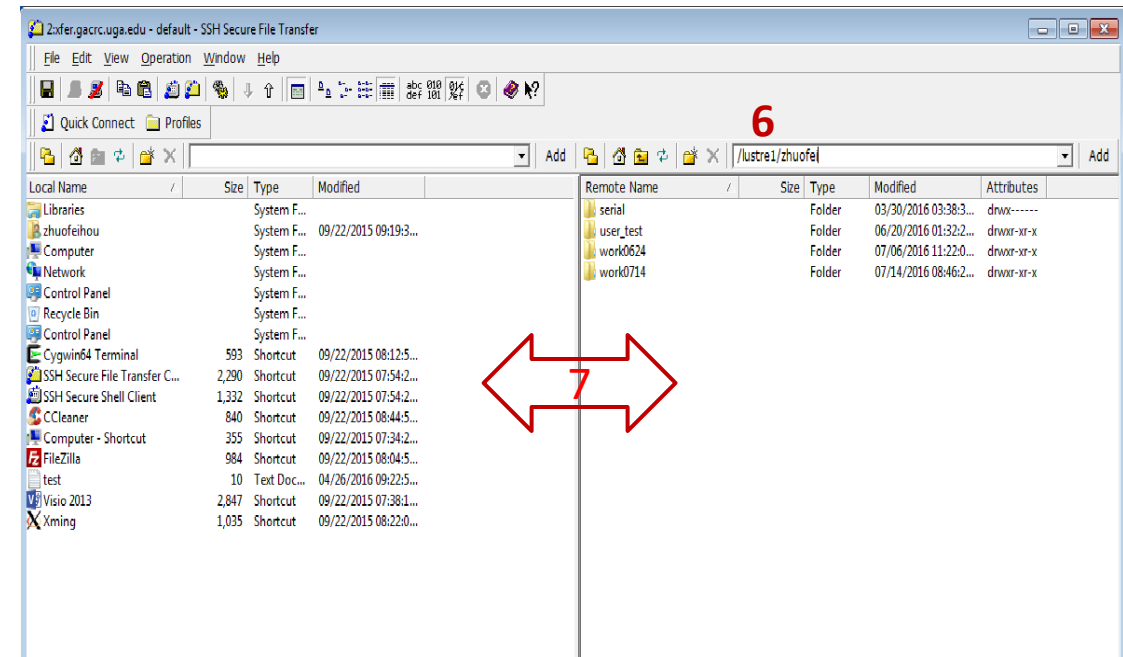
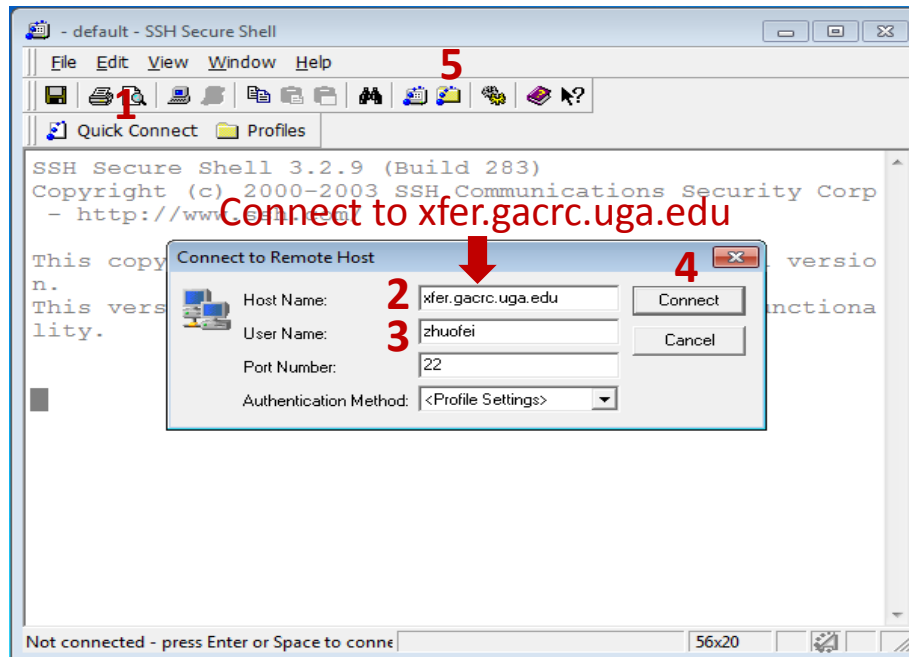
E.g. 2: working on local computer, from workDir on cluster → Local

```
scp zhuofei@xfer.gacrc.uga.edu:/lustrel1/zhuofei/workDir/file ./
```

```
scp -r zhuofei@xfer.gacrc.uga.edu:/lustrel1/zhuofei/workDir/folder/ ./
```

Step6 (Cont.) - Windows

1. You need to connect to cluster's Transfer node (xfer.gacrc.uga.edu)
2. Use SSH File Transfer or FileZilla or WinSCP on local computer





Step6 (Cont.): Transfer data on cluster to workDir

- Log on to Transfer node (xfer.gacrc.uga.edu)
 - ✓ Mac/Linux: `ssh MyID@xfer.gacrc.uga.edu` (page 8)
 - ✓ Windows: use SSH Secure Client app (page 9)
- Landing directory: `/home/MyID` (Sapelo home)
- Transfer data between folders on cluster using `cp`, `mv`
- Filesystems you can access using full path:
 1. `/home/MyID/` : Sapelo home directory (landing directory)
 2. `/lustre1/MyID/` : Sapelo global scratch directory
 3. `/project/abclab/` : Non-current active data temporary parking
- Most file systems on Transfer are *auto-mounted* upon *the first time full-path access*, e.g.,
`cd /project/abclab/`



Step7: Make a job submission script in workDir using nano

```
$ nano sub.sh
```

nano is a small and friendly text editor on Linux.

Ctrl-x to save file and quit from nano



```
zhuofei@n124:/lustre1/zhuofei/workDir
GNU nano 2.0.9 File: sub.sh Modified
hello nano! I am Zhuofei!

^G Get Help      ^O WriteOut     ^R Read File    [ New File ]
^X Exit          ^J Justify      ^W Where Is    ^Y Prev Page
^_              ^L Next Page    ^K Cut Text     ^C Cur Pos
^_              ^U UnCut Text   ^T To Spell
```

Step7 (Cont.)

1. Copy a template from GACRC Wiki Software:
<https://wiki.gacrc.uga.edu/wiki/Bowtie-Sapelo>
2. Modify it as needed for your computing

To run this example, you need to copy input data to your current working folder:

```
cp /usr/local/training/e_coli_data.fq .
```

<http://bowtie-bio.sourceforge.net/tutorial.shtml>

```
zhuofei@n146:/lustre1/zhuofei/workDir
File Edit View Search Terminal Help
GNU nano 2.0.9 File: sub.sh
#!/bin/bash
#PBS -q training_q
#PBS -N bowieJob_1
#PBS -l nodes=1:ppn=1:training
#PBS -l walltime=1:00:00
#PBS -l mem=2gb

#PBS -M zhuofei@uga.edu
#PBS -m ae

cd $PBS_O_WORKDIR

module load bowtie/1.1.1

time bowtie /usr/local/apps/bowtie/1.1.1/indexes/e_coli \
./e_coli_data.fq
█

[ Read 16 lines ]
^G Get Help ^O WriteOut ^R Read File ^Y Prev Page ^K Cut Text ^C Cur Pos
^X Exit ^J Justify ^W Where Is ^V Next Page ^U UnCut Text ^T To Spell
```



Step8: Submit a job from workDir using qsub

```
zhuofei@n148:/lustre1/zhuofei/workDir $ pwd
/lustre1/zhuofei/workDir
zhuofei@n148:/lustre1/zhuofei/workDir $ ls
my_e_coli_10000snp_500mb.fq  sub.sh
zhuofei@n148:/lustre1/zhuofei/workDir $ qsub sub.sh
2987391.pbs.scm
```

sub.sh is job submission script to

1. specify computing resources:
2. load software using **module load**
3. run any Linux commands you want to run
4. run the software



Step9: Check job status using qstat_me

```
zhuofei@n148:/lustrel1/zhuofei/workDir $ qstat_me
Job ID          Name          User          Time Use  S Queue
-----
2987177.pbs     STDIN         zhuofei       00:05:01 R s_interq
2987391.pbs     bowtieJob_1   zhuofei       00:00:35 C training_q
2987401.pbs     bowtieJob_1   zhuofei              0 R training_q
2987402.pbs     bowtieJob_1   zhuofei              0 Q training_q
```

R : job is running

C : job completed (or canceled or crashed) and is not longer running. Jobs stay in this state for 1 hour

Q : job is pending, waiting for resources to become available

Note: "Time Use" is the CPU time, instead of the wall-clock time of your job staying on cluster!



Step9 (Cont.): Cancel job using qdel

```
zhuofei@n148:/lustrel1/zhuofei/workDir $ qdel 2987402
zhuofei@n148:/lustrel1/zhuofei/workDir $ qstat_me
```

Job ID	Name	User	Time Use	S	Queue
2987177.pbs	STDIN	zhuofei	00:05:01	R	s_interq
2987391.pbs	bowtieJob_1	zhuofei	00:00:35	C	training_q
2987401.pbs	bowtieJob_1	zhuofei	0	R	training_q
2987402.pbs	bowtieJob_1	zhuofei	0	C	training_q

C status will stay in list for 1 hour



Step9 (Cont.): Check Job using qstat -n -u MyID

```
zhuofei@n14:/lustre1/zhuofei/workDir $ qstat -n -u zhuofei
```

```
pbs.scm:
```

Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Req'd Memory	Req'd Time	Elap S	Time
2991949.pbs.scm n146/31	zhuofei	s_interq	STDIN	819642	1	1	--	12:00:00	R	00:00:34
2991951.pbs.scm → n6/0	zhuofei	training	bowtieJob_1	842001	1	1	2gb	01:00:00	R	00:00:06
2991952.pbs.scm n6/1	zhuofei	training	bowtieJob_1	844269	1	1	2gb	01:00:00	R	00:00:04
2991953.pbs.scm n6/2	zhuofei	training	bowtieJob_1	844885	1	1	2gb	01:00:00	R	00:00:04

Note: "Elap Time" is the wall-clock time, instead of the CPU time, which qstat_me can give you!



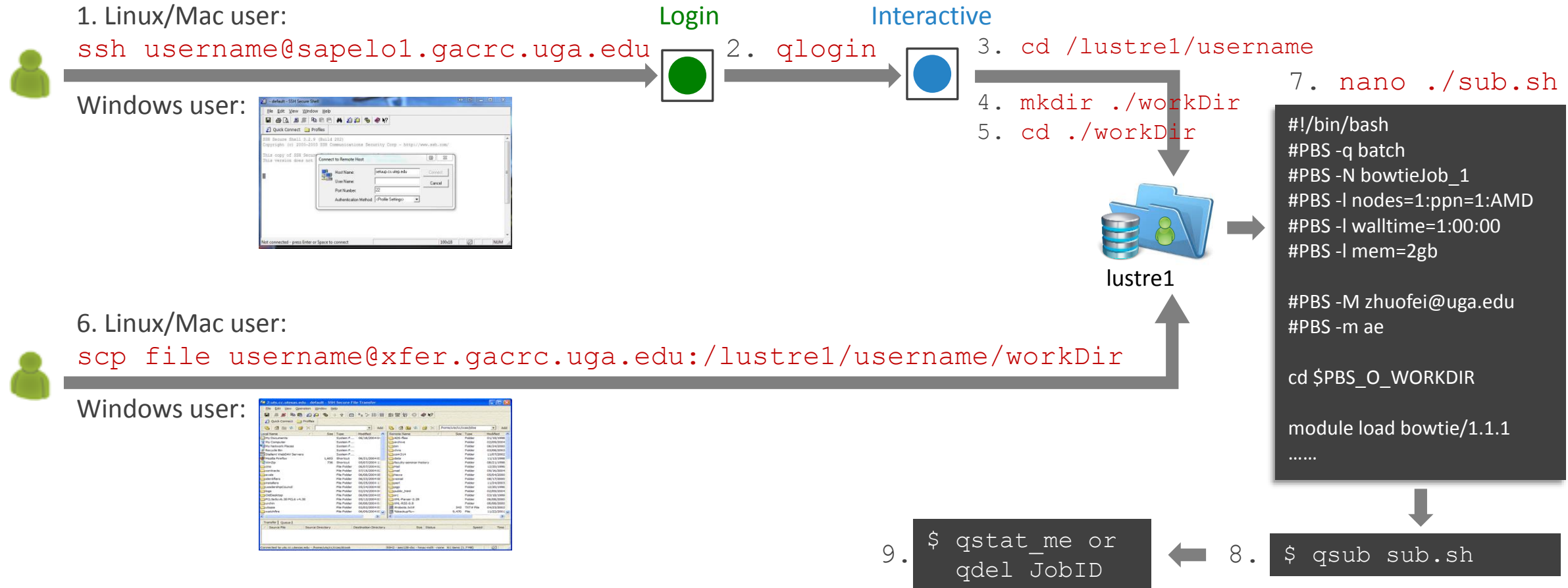
Step9 (Cont.): Check all Jobs on cluster using qstat

```
$ qstat
Job ID          Name                User                Time Use S Queue
-----
2392787.pbs    G_s-Butyl_0_opt     gtp72866            101:34:5 R batch
2392792.pbs    G_s-Butyl_20_opt    gtp72866            56:14:32 R batch
2393599.pbs    ...80-oz0.6-K220    rbchan              89:99:59 R batch
2395891.pbs    NoRiver_jul_2008    hluo76              0 Q batch
2395893.pbs    NoRiver_jul_2012    hluo76              0 Q batch
2396261.pbs    ...opyl_freq_005    gtp72866            31:55:56 R batch
2396297.pbs    ...opyl_freq_041    gtp72866            202:53:1 R batch
2400998.pbs    ...rrect_lab4.sh    xwwang              844:06:3 R batch
2406268.pbs    abundTrinity3       mpodio              0 Q batch
2423833.pbs    L80-307             jx57780             393:37:1 C batch
2423838.pbs    L80-312             jx57780             373:57:4 C batch
2513741.pbs    ...3012_ah100_00    byang               00:00:00 R batch
2513743.pbs    ...4012_ah070_01    byang               138:22:3 R batch
.
.
.
```

Note: qstat command will give you a long list of all jobs from all users on cluster!



Workflow Diagram





Work on Sapelo - Useful Links

GACRC Wiki: <http://wiki.gacrc.uga.edu>

GACRC Wiki Running Jobs: https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo

GACRC Wiki Software: <https://wiki.gacrc.uga.edu/wiki/Software>

GACRC Wiki Linux Command: https://wiki.gacrc.uga.edu/wiki/Command_List

GACRC Wiki Training: <https://wiki.gacrc.uga.edu/wiki/Training>

GACRC User Account Request: https://wiki.gacrc.uga.edu/wiki/User_Accounts

GACRC Support: https://wiki.gacrc.uga.edu/wiki/Getting_Help



Work on Sapelo - Get Help

To submit a ticket to us?

➤ **Job Troubleshooting:**

Please tell us details of your question or problem, including but not limited to:

- ✓ Your user name
- ✓ Your job ID
- ✓ Your working directory
- ✓ The queue name and command you used to submit the job

➤ **Software Installation:**

- ✓ Specific name and version of the software
- ✓ Download website
- ✓ Supporting package information if have

Note:

It's **USER's** responsibility to make sure the **correctness of datasets** being used by jobs!





Work on Sapelo – Get User Account

- User Account: **UGAMyID@sapelo1.gacrc.uga.edu**

A **valid official UGA MyID** is a MUST to create a user account!



- To get a user account:

1. Computing Lab Registration: <http://help.gacrc.uga.edu/labAcct.php> (for PI of a new group)
2. User Account Request: <http://help.gacrc.uga.edu/userAcct.php> (for PI of an existing group)
3. New User Training: <http://gacrc.uga.edu/help/training/>
4. **Welcome letter** with whole package of information about your Sapelo user account



Work on Sapelo - Run Batch Jobs

- Components you need to run a job:
 - **Software** already installed (cluster software or the one installed by yourself)
 - **Job submission script** to
 1. specify computing resources:
 - ✓ number of nodes and cores
 - ✓ amount of memory
 - ✓ node's feature
 - ✓ maximum wallclock time
 2. load software using **module load** (for cluster software)
 3. run any Linux commands you want to run, e.g., pwd, mkdir, cd, echo, etc.
 4. run the software
- Common queueing commands you need:
 - **qsub, qstat_me, qstat, qdel**
 - **showjobs, checkjob, showq**



Example 1: **Serial job script** sub.sh running NCBI Blast +

```
#PBS -S /bin/bash
#PBS -q batch
#PBS -N testBlast
#PBS -l nodes=1:ppn=1:AMD
#PBS -l mem=20gb
#PBS -l walltime=48:00:00

cd $PBS_O_WORKDIR

module load ncbiblast+/2.2.29

time blastn [options] >outputfile 2>&1
```

- Linux shell (bash)
- Queue name (batch); **Please do NOT use training_q for regular usage!**
- Name of the job (testBlast)
- Number of nodes (1), number of cores/node (1), node type (**AMD**)
- Maximum amount of physical memory (20 GB) used by the job
- Maximum wall clock time (48 hours) for the job, default 6 minutes
- Use the directory from which the job is submitted as the working directory
- Load the module of ncbiblast+, version 2.2.29
- Run blastn with 'time' command to measure the amount of time it takes to run the application



*Example 2: Threaded job script sub.sh running NCBI Blast+

```
#PBS -S /bin/bash
#PBS -q batch
#PBS -N testBlast
#PBS -l nodes=1:ppn=24:AMD
#PBS -l mem=20gb
#PBS -l walltime=480:00:00

#PBS -M jSmith@uga.edu
#PBS -m ae
#PBS -j oe

cd $PBS_O_WORKDIR

module load ncbiblast+/2.2.29

time blastn -num_threads 24 [options] >outputfile 2>&1
```

→ Number of nodes (1), number of cores/node (24), node type (AMD)
Number of threads (24) = Number of cores requested (24)

→ Email address to receive a notification for computing resources
→ Send email notification when job aborts (a) or terminates (e)
→ Standard error file (testBlast.e1234) will be merged into standard out file (testBlast.o1234)

→ Run blastn with 24 threads (-num_threads 24)



*Example 3: MPI job script sub.sh running RAxML

```
#PBS -S /bin/bash
#PBS -q batch
#PBS -N testRAxML
#PBS -l nodes=2:ppn=48:AMD
#PBS -l walltime=480:00:00
#PBS -l mem=20gb

cd $PBS_O_WORKDIR

module load raxml/8.1.20
module load intel/15.0.2
module load openmpi/1.8.3/intel/15.0.2

mpirun -np 50 raxmlHPC-MPI-AVX [options] >outputfile 2>&1
```

→ Number of nodes (2), number of cores/node (48), node type (AMD)
Total cores requested = $2 \times 48 = 96$
We suggest, Number of MPI Processes (50) \leq Number of cores requested (96)

→ To run raxmlHPC-MPI-AVX, MPI version using OpenMPI 1.8.3/Intel 15.0.2

→ Run raxmlHPC-MPI-AVX with 50 MPI processes (-np 50), default 96


```
#PBS -S /bin/bash
#PBS -q batch
#PBS -N testRAxML
#PBS -l nodes=2:ppn=27:AMD
#PBS -l walltime=480:00:00
#PBS -l mem=20gb
```

→ ppn number (27) fewer than 48 MUST be a multiplier of 3!

```
cd $PBS_O_WORKDIR
```

```
# Context Sharing
CONTEXTS=$(/usr/local/bin/set_contexts.sh $PBS_NUM_PPN)
if [[ "$?" -eq "0" ]] ; then
  export PSM_SHAREDCONTEXTS_MAX=$CONTEXTS
fi
```

} New lines copied from GACRC Wiki

```
module load raxml/8.1.20
module load intel/15.0.2
module load openmpi/1.8.3/intel/15.0.2
```

```
mpirun -np 50 raxmlHPC-MPI-AVX [options] >outputfile 2>&1
```

→ Run raxmlHPC-MPI-AVX with 50 MPI processes (-np 50), default 54



*Work on Sapelo - Check Computing Resources Used by a Job

Option 1: `qstat -f JobID` for *running jobs* or *finished jobs in 1 hour*

Option 2: `showjobs JobID` for *finished jobs over 1 hour, but ≤ 7 days*

Option 3: Email notification from *finished jobs (completed, canceled, or crashed),*

if using:

```
#PBS -M jSmith@uga.edu
```

```
#PBS -m ae
```



*Option 1: qstat -f JobID (running jobs or finished jobs in 1 hour)

```
$ qstat -f 1426778
Job Id: 1426778.pbs.scm
  Job_Name = BS2_GFF_range_methylator5-10.py
  Job_Owner = gent@uga-2f0f976.scm
resources_used.cput = 76:54:46
resources_used.energy_used = 0
resources_used.mem = 44136kb
resources_used.vmem = 90200kb
resources_used.walltime = 76:50:04
job_state = C
queue = batch
.
Error_Path = uga-2f0f976.scm:/home/gent/scripts/BS2_GFF_range_methylator5-10.py.e1426778
exec_host = n2/7
Output_Path = uga-2f0f976.scm:/home/gent/scripts/BS2_GFF_range_methylator5-10.py.o1426778
.
Resource_List.mem = 2gb
Resource_List.nodect = 1
Resource_List.nodes = 1:ppn=1:AMD
Resource_List.walltime = 480:00:00
.
Variable_List = PBS_O_QUEUE=batch,PBS_O_HOME=/home/gent, ..... ,
                PBS_O_WORKDIR=/home/gent/scripts,
```



*Option 2: showjobs JobID (finished jobs over 1 hr, but \leq 7 days)

```
$ showjobs 1426778
Job Id           : 1426778.pbs.scm
Job Name         : BS2_GFF_range_methylator5-10.py
Output File      : uga-2f0f976.scm:/home/gent/scripts/BS2_GFF_range_methylator5-10.py.o1426778
Error File       : uga-2f0f976.scm:/home/gent/scripts/BS2_GFF_range_methylator5-10.py.e1426778
Working Directory : /home/gent/scripts
Home Directory   : /home/gent
Submit Arguments : BS2_GFF_range_methylator5-10.py.sh
User Name        : gent
Group Name       : krdlab
Queue Name       : batch
Wallclock Limit  : 20:00:00:00
Wallclock Duration: 3:04:50:04
CPUTime          : 3:04:54:46
Memory Used      : 43.1Mb
Memory Limit     : 2gb
vmem Used        : 88.1Mb
Submit Time      : Mon Mar 27 20:51:01 2017
Start Time       : Tue Mar 28 03:06:41 2017
End Time         : Fri Mar 31 07:58:32 2017
Exit Code        : 0
Master Host      : n2
```



*Option 3: Email notification from finished jobs

PBS Job Id: 700009.pbs.scm
Job Name: testJob
Exec host: n1/4-27
Execution terminated
Exit_status=0

resources_used.cput=00:05:12
resources_used.energy_used=0
resources_used.mem=410984kb
resources_used.vmem=6548516kb
resources_used.walltime=04:00:59

Error_Path: uga-
2f0f976.scm:/home/zhuofei/MPIs/testJob.o700009
Output_Path: uga-
2f0f976.scm:/home/zhuofei/MPIs/testJob.o700009

PBS Job Id: 700097.pbs.scm
Job Name: testJob
Exec host: n5/4-27
Execution terminated
Exit_status=271

resources_used.cput=00:11:22
resources_used.energy_used=0
resources_used.mem=412304kb
resources_used.vmem=6548524kb
resources_used.walltime=05:00:41

Error_Path: uga-
2f0f976.scm:/home/zhuofei/MPIs/testJob.o700097
Output_Path: uga-
2f0f976.scm:/home/zhuofei/MPIs/testJob.o700097



*Work on Sapelo - Check Queue Status using showq

```
[zhuofei@75-104 MPIs]$ showq
active jobs-----
JOBID                USERNAME           STATE  PROCS   REMAINING          STARTTIME

481914                brant              Running  1      20:46:21  Fri Jun 12 11:32:23
481915                brant              Running  1      20:48:56  Fri Jun 12 11:34:58
481567                becton             Running 288    2:04:15:48 Wed Jun 10 15:01:50
481857                kkim               Running  48     9:18:21:41 Fri Jun 12 09:07:43
481859                kkim               Running  48     9:18:42:21 Fri Jun 12 09:28:23
.
108 active jobs      5141 of 5740 processors in use by local jobs (89.56%)
                    121 of 122 nodes active          (99.18%)

eligible jobs-----
481821                joykai             Idle    48     50:00:00:00 Thu Jun 11 13:41:20
481813                joykai             Idle    48     50:00:00:00 Thu Jun 11 13:41:19
481811                joykai             Idle    48     50:00:00:00 Thu Jun 11 13:41:19
.
50 eligible jobs

blocked jobs-----
JOBID                USERNAME           STATE  PROCS   WCLIMIT          QUEUETIME

0 blocked jobs
Total jobs: 158
```



Guideline Tips

- Do NOT use Login Node to run jobs → Interactive Node or the queue
 - Do NOT use Login Node upload or download data to/from cluster
 - Do NOT use Login Node to transfer data to the home dir
- } Transfer Node
- NO large memory job running on AMD nodes → HIGHMEM nodes
 - NO small memory job running on HIGHMEM nodes → AMD nodes
 - As a general rule, threads # = cores # requested

Practical Tips

- **Each directory should not have too many files inside!** A rule of thumb would be to try to keep no more than a few tens of thousands of files (<10000 would be even better) in any single directory which is accessed frequently



All files are in ONE single dir! ❌



Files are organized in subdirs! ✅



Practical Tips

- Job name should have a specific computational meaning

Good Examples: `#PBS -N blastn_dataSet1_trail2 ; #PBS -N M-10-1121`

Bad Examples: `#PBS -N job1 ; #PBS -N bowtie ; #PBS -N 20160930`

- The stdout .o file and stderr .e file are to be written into files at the finishing time of a job.

Redirect standard output and error of the application to a file, instead of letting it be written in the stdout .o file and stderr .e file of the job, e.g.:

```
time application >file 2>&1
```

- Monitor job progress from time to time, to catch if a job gets stuck



Thank You!