

GACRC Sapelo2 Cluster New User Training

Georgia Advanced Computing Resource Center (GACRC)

Enterprise Information Technology Services(EITS)

The University of Georgia

Outline

- GACRC
- Sapelo2 Cluster
 - Cluster Diagram and Overview
 - Five Directories
 - Five Computational Partitions
 - Software Environment
- Batch Job Submission Workflow
- Useful Commands: `queue --me`, `sacct`, `sacct-gacrc`, `qlogin`
- GACRC Wiki and User Support
- Appendices

GACRC

- A high-performance-computing (HPC) center at the UGA
- 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

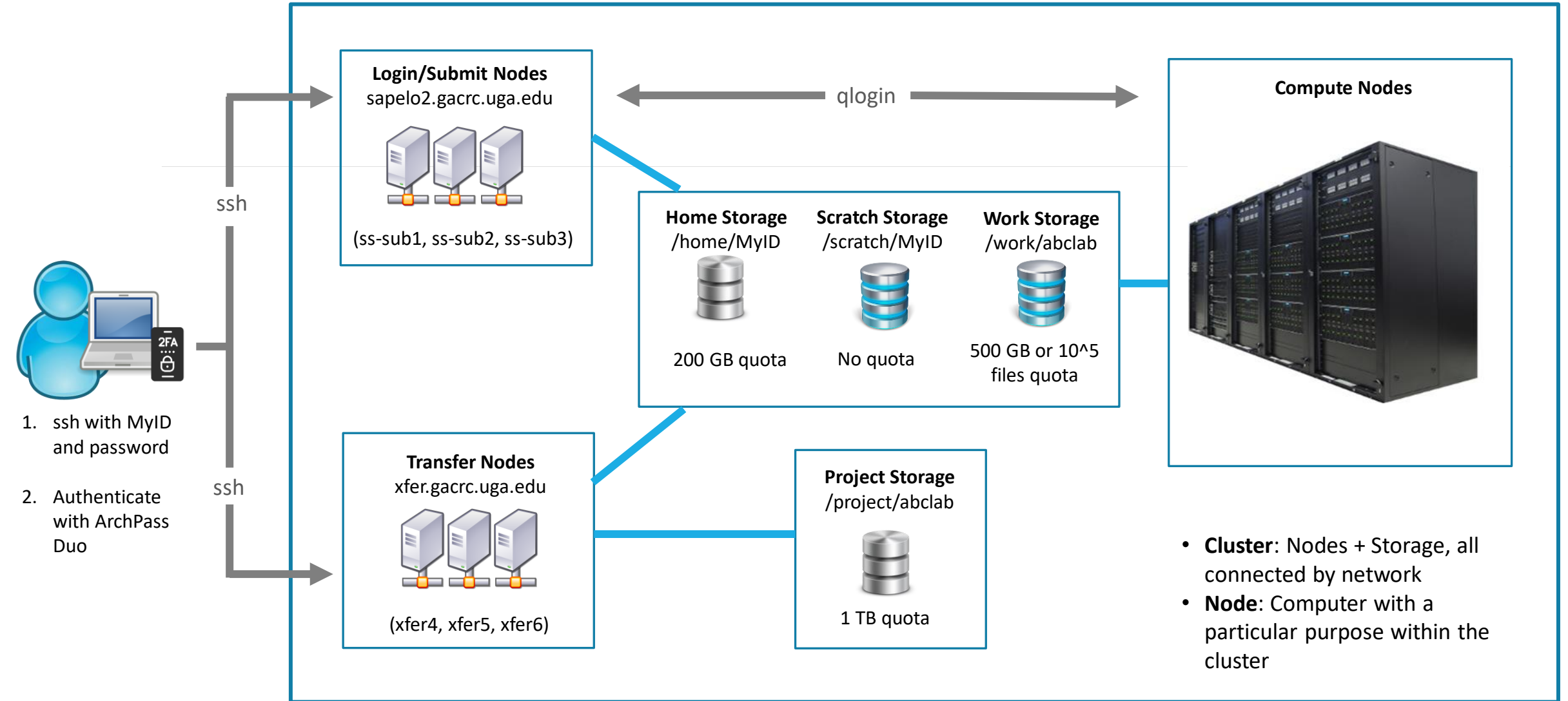
Wiki: <http://wiki.gacrc.uga.edu>

Help and Support: https://wiki.gacrc.uga.edu/wiki/Getting_Help

Web Site: <http://gacrc.uga.edu>

Kaltura Channel: <https://kaltura.uga.edu/channel/GACRC/176125031>

Sapelo2 Cluster



Note: You need to connect to the **UGA network using VPN** when accessing from outside of the **UGA main campus**.

UGA VPN: https://eits.uga.edu/access_and_security/infosec/tools/vpn/

Five Directories https://wiki.gacrc.uga.edu/wiki/Disk_Storage

Directory	Name	Quota	Accessible from	Intended Use	Backed-up	Important Notes
/home/MyID	Home	200GB	Login Transfer Compute	Static data, e.g. 1. Scripts, source codes 2. Local software	Yes	Not for storing data of your jobs!
/scratch/MyID	Scratch	No Limit	Login Transfer Compute	Temporary files needed for currently running jobs	No	Clean up when your job finishes! Subject to “30-day purge” policy
/work/abclab	Work	500GB 10 ⁵ files	Login Transfer Compute	Input files needed for repeated jobs	No	Clean up when your job finishes! Group sharing is possible
/project/abclab	Project	1TB (initial)	Transfer	Temporary data parking	Yes	Group sharing is possible
/lscratch	Local Scratch	200GB - 800GB	Compute	Jobs with heavy disk I/O operations	No	Clean up when job exits from node!

Scratch File System 30-Day Purge Policy

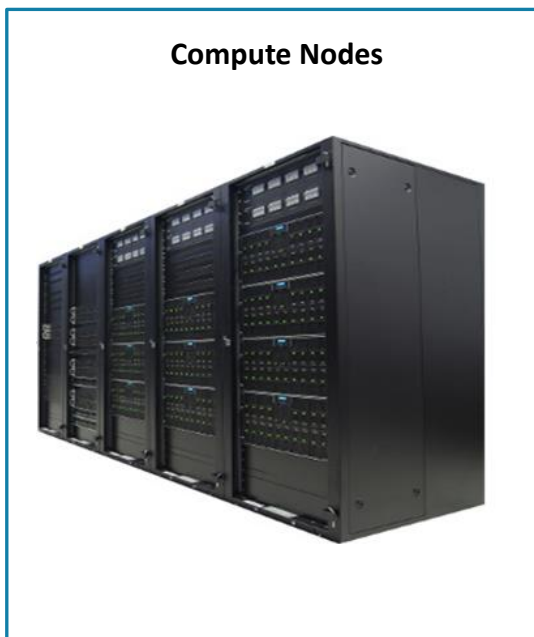
https://wiki.gacrc.uga.edu/wiki/Disk_Storage#Scratch_file_system

Any file that is not accessed or modified by a compute job in a time period **no longer than 30 days** will be automatically deleted off the /scratch file system.

Measures circumventing this policy will be monitored and actively discouraged.

- You have a list of those purgeable files located at **/usr/local/var/lustre_stats/\$USER.over30d.files.lst**
- You are suggested to copy files from /scratch to **/project** or **outside of GACRC**
- You should first move all unnecessary files and folders to **/scratch/trash/\$USER**
- The fastest way to save old files is to copy them to /project area, using the **fpsync** utility on xfer.gacrc.uga.edu
- When you archive data using **tar** on /scratch, please **do not use z option** (compression option). After you achieved data by tar, you can use gzip to compress it.

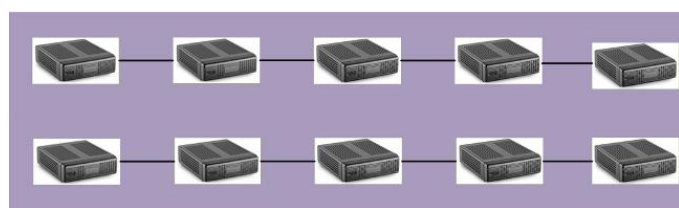
Computational Partitions



batch/batch_30d
For regular jobs



highmem_p/
highmem_30d_p
For high memory jobs



gpu_p/gpu_30d_p
For GPU-enabled jobs



inter_p
For interactive jobs



_p = "partition"
_30d = 30 day (time limit)

Compute nodes are divided into groups called **partitions**. A **partition** is a collection of compute nodes for a particular computing need.

Computational Partitions [https://wiki.gacrc.uga.edu/wiki/Job Submission partitions on Sapelo2](https://wiki.gacrc.uga.edu/wiki/Job_Submission_partitions_on_Sapelo2) <https://wiki.gacrc.uga.edu/wiki/Systems#Sapelo2>

Type	Partition	Time limit	Max jobs Running	Max jobs Submit	Notes
Regular	batch	7 days	250	10,000	Regular nodes
	highmem_p		15	100	For running high memory jobs
	gpu_p		18	20	For running GPU-enabled jobs
Long-term	batch_30d	30 days	1	2	30-day partition counterparts
	highmem_30d_p				
	gpu_30d_p		2		
Interactive	inter_p	2 days	3	20	Regular nodes, for interactive jobs.
Buy-in	name_p	variable			Partitions that target different groups' buy-in nodes. The name string is specific to each group.

Partition	Total Nodes	Max Mem(GB) /Single-node job	Cores /Node	Processor Type	GPU Cards /Node
batch batch_30d	~275	120	64	AMD EPYC	N/A
			32		
		48	AMD Opteron		
		180	Intel Xeon Skylake		
		58	Intel Xeon Broadwell		
highmem_p highmem_30d_p	~30	500	32	AMD EPYC	N/A
			48	AMD Opteron	
		990	64	AMD EPYC	
			28	Intel Xeon Broadwell	
			48	AMD Opteron	
gpu_p gpu_30d_p	~10	180	32	Intel Xeon Skylake	1 NVIDIA P100
		120	16	Intel Xeon	8 NVIDIA K40m
		90	12		7 NVIDIA K20Xm
name_p	variable				

Software Environment

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

- Approximately 900 software modules are installed (as of March 2021)
- Most modules are compiled EasyBuild (EB) toolchains GCC-8.3.0 or foss-2019b.
- Name format: **Name/Version-Toolchain**, e.g., **Python/3.8.2-GCCcore-8.3.0** (case-sensitive)
- Module commands:
 - **ml spider *pattern*** : Search module names matching a *pattern*
 - **ml *moduleName*** : Load a module into your working environment
 - **DO NOT LOAD/USE MODULES ON THE LOGIN/SUBMIT NODES! (ss-sub1, ss-sub2, ss-sub3, etc...)**
 - **ml av** : List all available software modules installed on cluster
 - **ml** : List modules currently loaded
 - **ml -*moduleName*** : Remove a module from working environment
 - **ml purge** : Remove all modules from working environment

Important Tip using Software

[https://wiki.gacrc.uga.edu/wiki/Available Toolchains and Toolchain Compatibility](https://wiki.gacrc.uga.edu/wiki/Available_Toolchains_and_Toolchain_Compatibility)

- When you load more than one software modules, **toolchain compatibility** is the most important thing you need to pay attention to
- If you load more than one module and some toolchains are incompatible, your job will end up with failing dependencies or Lmod errors, such as:

Lmod has detected the following error:

These module(s) exist but cannot be loaded as requested



ml Python/3.7.4-GCCcore-8.3.0
ml Perl/5.28.0-GCCcore-7.3.0



ml Python/3.7.4-GCCcore-8.3.0
ml Perl/5.30.0-GCCcore-8.3.0



ml Beast/2.6.3-foss-2019b
ml Perl/5.28.0-GCCcore-7.3.0



ml Beast/2.6.3-foss-2019b
ml Perl/5.30.0-GCCcore-8.3.0

Job Submission Workflow

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo2

1. Log on to Login node using MyID and password, and two-factor authentication with Archpass Duo:
`ssh MyID@sapelo2.gacrc.uga.edu`
2. On Login node, change directory to your scratch space: `cd /scratch/MyID`
3. Create a working subdirectory for a job : `mkdir ./workDir`
4. Change directory to workDir : `cd ./workDir`
5. Transfer data from local computer to workDir : use `scp` or **WinSCP** to connect Transfer node
Transfer data on cluster to workDir : log on to Transfer node and then use `cp` or `mv`
6. Make a job submission script in workDir : `nano ./sub.sh`
7. Submit a job from workDir : `sbatch ./sub.sh`
8. Check job status : `squeue --me` or Cancel a job : `scancel jobID`

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

<https://wiki.gacrc.uga.edu/wiki/Connecting>

1. Open **Terminal** utility
2. Type command line: `ssh MyID@sapelo2.gacrc.uga.edu`
3. You will be prompted for your **MyID password**
4. Sapelo2 access requires ID verification using two-factor authentication with Archpass

Duo. If you are not enrolled in Archpass Duo, please refer to

https://eits.uga.edu/access_and_security/infosec/tools/archpass_duo/ on how to enroll

Step1 (Cont.) - Mac/Linux using ssh

Use Terminal utility on
Mac or Linux!



```
[zhuofei@localhost ~]$  
[zhuofei@localhost ~]$ ssh zhuofei@sapelo2.gacrc.uga.edu ← Log on  
Password: ← Input MyID password!
```

.....
Enter a passcode or select one of the following options:

1. Duo Push to XXX-XXX-5758
2. Phone call to XXX-XXX-5758
3. Phone call to XXX-XXX-1925
4. SMS passcodes to XXX-XXX-5758

Passcode or option (1-5): 1 ← Select Duo authentication option!

Success. Logging you in...

Last login: Tue Sep 15 11:22:42 2020 from 128.192.75.65

```
zhuofei@ss-sub1 ~$ ← I am on login node ss-sub1!
```

Step1 (Cont.) - Windows using PuTTY

1. Download and install PuTTY: <https://www.putty.org/>

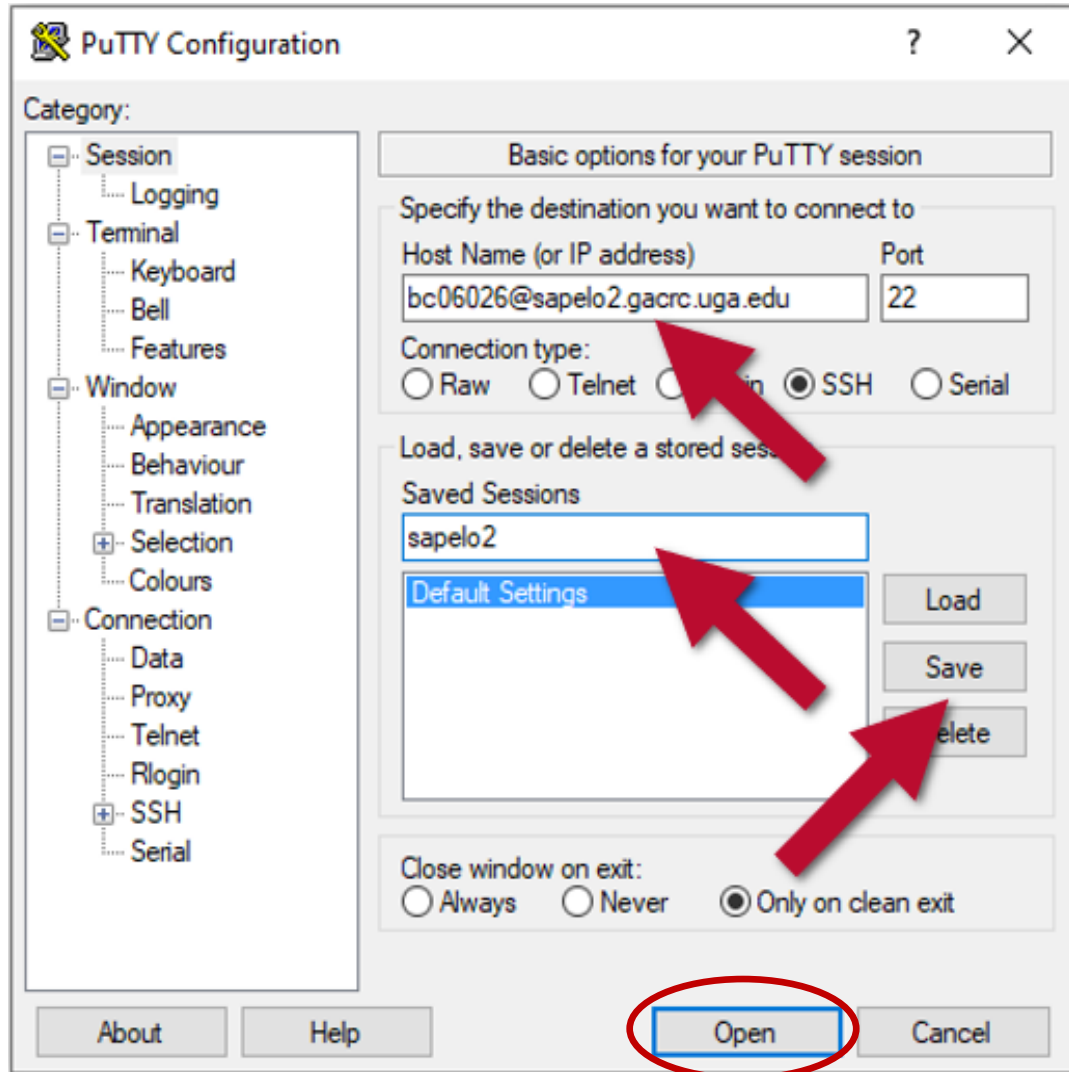
2. Detailed downloading and installation instructions:

[https://wiki.gacrc.uga.edu/wiki/How to Install and Configure PuTTY](https://wiki.gacrc.uga.edu/wiki/How_to_Install_and_Configure_PuTTY)

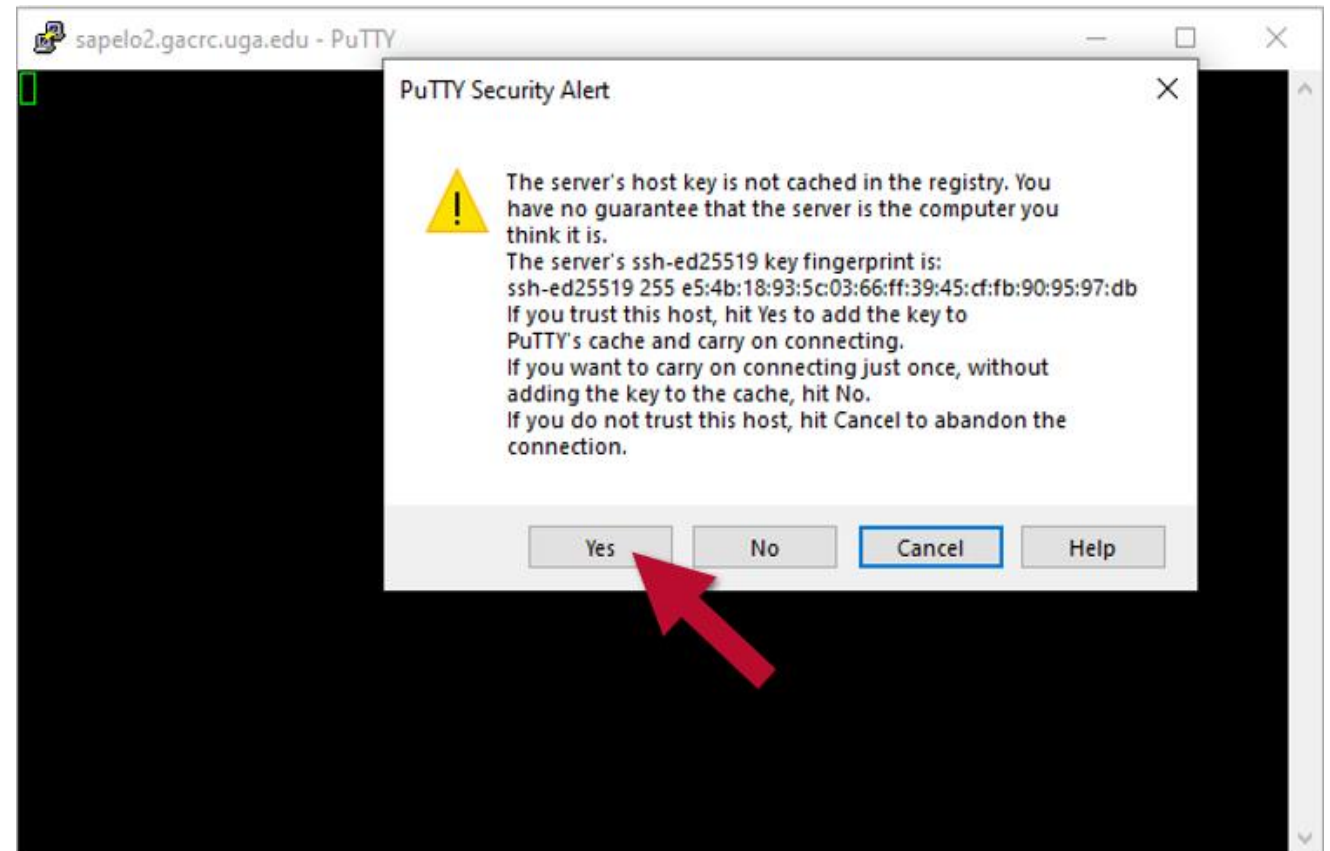
3. Detailed configuring and usage instructions:

[https://wiki.gacrc.uga.edu/wiki/How to Install and Configure PuTTY#Configuring PuTTY](https://wiki.gacrc.uga.edu/wiki/How_to_Install_and_Configure_PuTTY#Configuring_PuTTY)

Step1 (Cont.) - Windows using PuTTY



The first time you connect to login node, PuTTY will give you this security alert window. Please click "Yes"



Step1 (Cont.) - Windows using PuTTY

Next you will enter your UGA MyID password and initiate DUO authentication procedure:

```
sapelo2.gacrc.uga.edu - PuTTY
Using username "bc06026".
Keyboard-interactive authentication prompts from server:
| UGA DUO authentication is required for SSH/SCP access to
| GACRC systems.
|
| UGA DUO is a two-factor authentication service which
| requires a password (one factor) and a code, phone,
| or device (second factor) to successfully authenticate.
|
| If you are not enrolled in the UGA DUO service please
| visit the UGA DUO service self-service portal to enroll
| and configure or manage your DUO enabled devices.
|
| https://eits.uga.edu/access_and_security/infosec/tools/duo/portal/
|
| For additional help with UGA DUO authentication or to
| report an issue please visit:
|
| https://eits.uga.edu/access_and_security/infosec/tools/archpass/
|
| Password: █ ← UGA MyID password
```

```
sapelo2.gacrc.uga.edu - PuTTY
| requires a password (one factor) and a code, phone,
| or device (second factor) to successfully authenticate.
|
| If you are not enrolled in the UGA DUO service please
| visit the UGA DUO service self-service portal to enroll
| and configure or manage your DUO enabled devices.
|
| https://eits.uga.edu/access_and_security/infosec/tools/duo/portal/
|
| For additional help with UGA DUO authentication or to
| report an issue please visit:
|
| https://eits.uga.edu/access_and_security/infosec/tools/archpass/
| Duo two-factor login for zhuofei
|
| Enter a passcode or select one of the following options:
|
| 1. Duo Push to XXX-XXX-5758
| 2. Phone call to XXX-XXX-5758
| 3. Phone call to XXX-XXX-1925
| 4. Phone call to XXX-XXX-3535
| 5. SMS passcodes to XXX-XXX-5758
|
| Passcode or option (1-5): 1 █ ← DUO option
```

Step2: On Login node change directory to global scratch

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

```
zhuofei@ss-sub1 ~$ pwd  
/home/zhuofei
```

← this is my home directory!

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

```
zhuofei@ss-sub1 ~$ cd /scratch/zhuofei/  
zhuofei@ss-sub1 zhuofei$ pwd  
/scratch/zhuofei
```

← this is my scratch space!

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

```
zhuofei@ss-sub1 zhuofei$ ls  
user_test
```

Step3 - 4: Create and cd to a working subdirectory

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

```
zhuofei@ss-sub1 zhuofei$ mkdir workDir
zhuofei@ss-sub1 zhuofei$ ls
user_test  workDir
```

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

```
zhuofei@ss-sub1 zhuofei$ cd workDir
zhuofei@ss-sub1 workDir$ pwd
/scratch/zhuofei/workDir
zhuofei@ss-sub1 workDir$ ls
```

← it is empty!

Step5: Transfer data from local computer to workDir - Mac/Linux

https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_scp

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:/scratch/zhuofei/workDir/
```

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

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

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

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

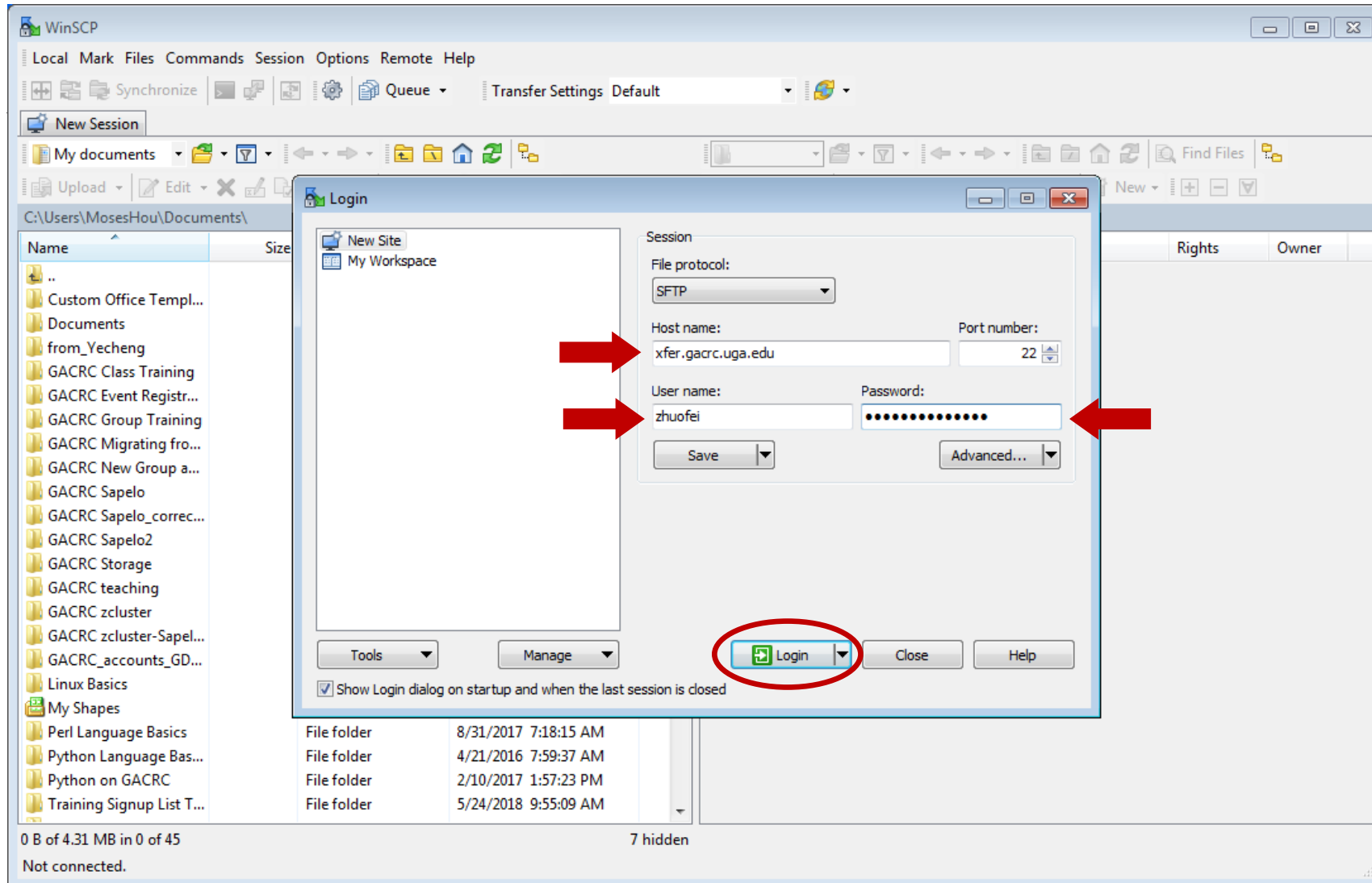
Step5 (Cont.) - Windows using WinSCP

https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_WinSCP

1. You need to connect to cluster's Transfer node (xfer.gacrc.uga.edu)
2. Use **WinSCP** on local computer
 - WinSCP can be downloaded from <https://winscp.net/eng/index.php>
 - Default installation procedure is simple
3. Alternative FileZilla https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_FileZilla

Step5 (Cont.) - Windows using WinSCP

https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_WinSCP



Step5 (Cont.) - Windows using WinSCP

https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_WinSCP

The screenshot shows the WinSCP application window titled "zhuofei@xfer.gacrc.uga.edu - WinSCP". The interface includes a menu bar (Local, Mark, Files, Commands, Session, Options, Remote, Help), a toolbar with icons for Synchronize, Queue, and Transfer Settings, and a file explorer view of the local drive "C:\Users\MosesHou\Documents\". A modal dialog box titled "Server prompt - zhuofei@xfer.gacrc.uga.edu" is overlaid on the interface. The dialog contains the following text:

UGA DUO authentication is required for SSH/SCP access to GACRC systems.

UGA DUO is a two-factor authentication service which requires a password (one factor) and a code, phone, or device (second factor) to successfully authenticate.

If you are not enrolled in the UGA DUO service please visit the UGA DUO service self-service portal to enroll and configure or manage your DUO enabled devices.

https://eits.uga.edu/access_and_security/infosec/tools/duo/portal/

For additional help with UGA DUO authentication or to report an issue please visit:

https://eits.uga.edu/access_and_security/infosec/tools/archpass/

Duo two-factor login for zhuofei

Enter a passcode or select one of the following options:

1. Duo Push to XXX-XXX-5758
2. Phone call to XXX-XXX-5758
3. Phone call to XXX-XXX-1925
4. Phone call to XXX-XXX-3535
5. SMS passcodes to XXX-XXX-5758

Passcode or option (1-5):

Buttons: OK, Cancel, Help

A red arrow points to the input field with the text "DUO option".

Step5 (Cont.) - Windows using WinSCP

https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_WinSCP

Local Mark Files Commands Session Options Remote Help

Synchronize Queue Transfer Settings Default

zhuofei@xfer.gacrc.uga.edu X New Session

C: Local Disk zhuofei

C:\Users\MosesHou\ /home/zhuofei/

Change paths on your local computer and transfer node

Drag to transfer files or folders

Name	Size	Type	Changed	Name	Size	Changed	Rights	Owner
..		Parent directory	7/10/2020 6:29:42 AM	..		1/4/2021 1:29:10 PM	rw-r-xr-x	root
Contacts		File folder	6/24/2020 5:30:50 AM	apps		1/4/2021 1:59:38 PM	rw-r-xr-x	zhuofei
Desktop		File	1/4/2021 8:46:03 AM	Backups_20200630		6/30/2020 7:47:15 AM	rw-r-xr-x	zhuofei
Documents		File folder	11/18/2020 1:04:09 PM	dask_conda		6/18/2020 6:52:03 PM	rw-r-xr-x	zhuofei
Downloads		File folder	1/4/2021 9:26:48 AM	documents		8/24/2020 9:30:05 AM	rw-r-xr-x	zhuofei
Favorites		File folder	6/24/2020 5:30:50 AM	globusconnectperson...		8/14/2020 3:18:22 PM	rw-r-xr-x	zhuofei
Links		File folder		language		11/25/2020 1:49:50 PM	rw-r-xr-x	zhuofei
Music		File folder		b		11/1/2018 10:36:45 AM	rw-r-xr-x	zhuofei
Pictures		File folder		mailmerge		6/16/2020 1:46:37 PM	rw-r-xr-x	zhuofei
Saved Ga		File folder		miniconda3		6/16/2020 9:35:35 AM	rw-r-xr-x	zhuofei
Searches		File folder	6/24/2020 5:30:50 AM	orthomcl.config		4/27/2020 8:09:04 AM	rw-r-xr-x	zhuofei
Tracing		File folder	7/7/2015 10:45:05 AM	research		10/21/2019 10:34:03 AM	rw-r-xr-x	zhuofei
Videos		File folder	6/24/2020 5:30:50 AM	sap2test		7/20/2020 10:21:13 AM	rw-r-xr-x	zhuofei
				sapelo2_account_pro...		12/15/2020 11:35:00 AM	rw-r-xr-x	zhuofei
						10/8/2020 11:29:50 AM	rw-r-xr-x	zhuofei
						6/22/2018 8:26:17 AM	rw-r-xr-x	zhuofei
						12/22/2020 12:39:04 PM	rw-r-xr-x	zhuofei
						9/9/2020 10:04:27 AM	rw-r-xr-x	zhuofei
						10/26/2020 2:07:38 PM	rw-r-xr-x	zhuofei
						7/16/2020 7:07:19 AM	rw-r-xr-x	zhuofei
						2/14/2019 2:02:52 PM	rw-r-xr-x	zhuofei
						5/11/2020 6:53:35 AM	rw-r-xr-x	zhuofei
						7/6/2020 11:09:19 AM	rw-r-xr-x	zhuofei

0 B of 0 B in 0 of 12 27 hidden 0 B of 77.5 MB in 0 of 40 310 hidden

SFTP-3 0:00:40

Step5 (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 15-16)
 - ✓ Windows: use PuTTY to log in `MyID@xfer.gacrc.uga.edu` (page 17-19)
- Landing folder: `/home/MyID` (Home)
- You can transfer data between following directories on cluster using `cp` or `mv`:
 1. `/home/MyID` (Home)
 2. `/scratch/MyID` (Scratch)
 3. `/work/abclab` (Work)
 4. `/project/abclab` (Project)
- Most file systems on Transfer are *auto-mounted* upon the first time *full-path* access, e.g.,
`cd /project/abclab/`

Step6: Make a job submission script in workDir

https://wiki.gacrc.uga.edu/wiki/Sample_batch_job_submission_scripts_on_Sapelo2

```
$ nano sub.sh
```

nano is a simple text editor on Linux. You are welcome to use other editors like vim or emacs.

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 ^Y Prev Page ^C Cur Pos
^X Exit ^J Justify ^W Where Is ^V Next Page ^U UnCut Text ^T To Spell
```



```
#!/bin/bash
#SBATCH --job-name=testBowtie2      # Job name (testBowtie2)
#SBATCH --partition=batch           # Queue name (batch)
#SBATCH --ntasks=1                 # Run single task using one CPU core on a single node
#SBATCH --mem=4G                   # Job memory limit (4 GB)
#SBATCH --time=1:00:00             # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --export=NONE              # Do not load any users' explicit environment variables
#SBATCH --output=%x_%j.out         # Standard output, testBowtie2_1234.out
#SBATCH --error=%x_%j.err          # Standard error log, testBowtie2_1234.err
#SBATCH --mail-type=END,FAIL       # Mail events (BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu # Where to send mail

cd $SLURM_SUBMIT_DIR                # Change directory to job submission directory
ml Bowtie2/2.4.1-GCC-8.3.0          # Load software module and run bowtie2 below

bowtie2 -x ./index/lambda_virus -U ./myreads.fq -S output.sam
```

```
^G Get Help      ^O WriteOut     ^R Read File    ^Y Prev Page    ^K Cut Text
^X Exit          ^J Justify      ^W Where Is     ^V Next Page    ^U UnCut Text
```

To run the demo example, please copy these files into your working dir:

```
cp -r /usr/local/training/Sapelo2/* .
```

Step7: Submit a job from workDir using **sbatch**

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo2#How_to_submit_a_job_to_the_batch_partition

```
zhuofei@ss-sub1 workDir$ pwd
/scratch/zhuofei/workDir
zhuofei@ss-sub1 workDir$ ls
index  myreads.fq  sub.sh
zhuofei@ss-sub1 workDir$ sbatch sub.sh
Submitted batch job 32860
```

sub.sh is a 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

Step8: Check job status using `squeue --me`

https://wiki.gacrc.uga.edu/wiki/Monitoring_Jobs_on_Sapelo2

```
zhuofei@ss-sub1 workDir$ squeue --me
JOBID PARTITION      NAME      USER  ST      TIME  NODES  NODELIST (REASON)
32862   batch  testBowt  zhuofei PD       0:00     1      (None)
32861   batch  testBowt  zhuofei R        0:05     1      c5-19
32860   batch  testBowt  zhuofei R        4:37     1      c5-19
32859   batch  testBowt  zhuofei CD       9:29     1      b1-2
```

R : job is **R**unning

PD : job is **P**en**D**ing, waiting for resources to become available

CD : job is **C**omplete**D** and is not longer running

CA : job is **C**Ancel**e**d and is not longer running

F : job is **F**ailed (crashed) on the node and is not longer running

Note: "TIME" is the **elapsed wall-clock time** of your job running on cluster, instead of the CPU time.

Step8: Check job status using `squeue --me -l`

https://wiki.gacrc.uga.edu/wiki/Monitoring_Jobs_on_Sapelo2

```
zhuofei@ss-sub1 workDir$ squeue --me -l
```

```
Tue Sep 15 15:00:51 2020
```

JOBID	PARTITION	NAME	USER	STATE	TIME	TIME_LIMI	NODES	NODELIST (REASON)
32866	batch	testBowt	zhuofei	RUNNING	0:14	1:00:00	1	rc6-10
32865	batch	testBowt	zhuofei	RUNNING	0:30	1:00:00	1	rc6-10
32864	batch	testBowt	zhuofei	RUNNING	0:33	1:00:00	1	rc6-10

Step8 (Cont.): Cancel job using **scancel**

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo2#How_to_cancel_.28delete.29_a_running_or_pending_job

```
zhuofei@ss-sub1 workDir$ queue --me -l
Tue Sep 15 15:08:27 2020
JOBID PARTITION      NAME      USER      STATE      TIME  TIME_LIMI  NODES  NODELIST (REASON)
32869      batch  testBowt  zhuofei   RUNNING   0:01    1:00:00     1  c5-19
32868      batch  testBowt  zhuofei   RUNNING   0:05    1:00:00     1  c5-19
32867      batch  testBowt  zhuofei   RUNNING   0:10    1:00:00     1  c5-19
zhuofei@ss-sub1 workDir$
zhuofei@ss-sub1 workDir$ scancel 32867
zhuofei@ss-sub1 workDir$ queue --me -l
Tue Sep 15 15:08:45 2020
JOBID PARTITION      NAME      USER      STATE      TIME  TIME_LIMI  NODES  NODELIST (REASON)
32867      batch  testBowt  zhuofei  COMPLETI  0:26    1:00:00     1  c5-19
32869      batch  testBowt  zhuofei   RUNNING   0:19    1:00:00     1  c5-19
32868      batch  testBowt  zhuofei   RUNNING   0:23    1:00:00     1  c5-19
zhuofei@ss-sub1 workDir$ queue --me -l
Tue Sep 15 15:08:50 2020
JOBID PARTITION      NAME      USER      STATE      TIME  TIME_LIMI  NODES  NODELIST (REASON)
32869      batch  testBowt  zhuofei   RUNNING   0:19    1:00:00     1  c5-19
32868      batch  testBowt  zhuofei   RUNNING   0:23    1:00:00     1  c5-19
```

```
bc06026@b1-24 workdir$ sacct
```

JobID	JobName	Partition	Account	AllocCPUS	State	ExitCode
3326893	testBowti+	batch	gacrc-ins+	1	RUNNING	0:0
3326893.bat+	batch		gacrc-ins+	1	RUNNING	0:0
3326893.ext+	extern		gacrc-ins+	1	RUNNING	0:0

```
bc06026@b1-24 workdir$ sacct -X
```

JobID	JobName	Partition	Account	AllocCPUS	State	ExitCode
3326893	testBowti+	batch	gacrc-ins+	1	RUNNING	0:0

```
bc06026@b1-24 workdir$ sacct -X --format jobid,state
```

JobID	State
-------	-------

3326893	RUNNING
---------	---------

```
bc06026@b1-24 workdir$
```


zhuofei@ss-sub1 workDir\$ sacct-gacrc

JobID	JobName	User	Partition	NodeList	AllocNodes	NTasks	NCPUS	ReqMem	MaxVMSize	State	CPUTime	Elapsed	Timelimit	ExitCode	WorkDir
275	bowti+	zhuofei	batch	rc6-10	1	1	4Gn			COMPLETED	00:09:13	00:09:13	01:00:00	0:0	/scratch/zhu..
275.batch	batch			rc6-10	1	1	4Gn	396868K		COMPLETED	00:09:13	00:09:13		0:0	
275.extern	extern			rc6-10	1	1	4Gn	142616K		COMPLETED	00:09:13	00:09:13		0:0	
276	amberjob	shtsai	gpu_p	c4-23	1	1	10Gn			CANCELLED+	00:03:19	00:03:19	02:00:00	0:0	/scratch/sht..
276.batch	batch			c4-23	1	1	10Gn	221140K		CANCELLED	00:03:20	00:03:20		0:15	
276.extern	extern			c4-23	1	1	10Gn	169800K		COMPLETED	00:03:19	00:03:19		0:0	
277	mpitest	shtsai	batch	c2-[11-12]	2	24	600Mc			COMPLETED	04:01:12	00:10:03	02:00:00	0:0	/scratch/sht..
277.batch	batch			c2-11	1	1	12	600Mc	221268K	COMPLETED	02:00:36	00:10:03		0:0	
277.extern	extern			c2-[11-12]	2	2	24	600Mc	169800K	COMPLETED	04:01:12	00:10:03		0:0	
277.0	orted			c2-12	1	1	1	600Mc	265640K	COMPLETED	00:00:01	00:00:01		0:0	
278	bash	shtsai	inter_p	c2-4	1	1	2Gn			RUNNING	00:13:37	00:13:37	12:00:00	0:0	/scratch/sht..
278.extern	extern			c2-4	1	1	1	2Gn		RUNNING	00:13:37	00:13:37		0:0	
278.0	bash			c2-4	1	1	1	2Gn		RUNNING	00:13:37	00:13:37		0:0	

zhuofei@ss-sub1 workDir\$ sacct-gacrc-v 47939

JobID	47939
JobName	testBowti+
User	zhuofei
Partition	batch
NodeList	c1-3
AllocNodes	1
State	RUNNING
CPUTime	00:00:28
Elapsed	00:00:28
Timelimit	01:00:00
ExitCode	0:0
WorkDir	/scratch/zhuofei/workDir_sapelo2
NTasks	1
NCPUS	1
ReqMem	4Gn

zhuofei@ss-sub1 workDir\$ seff 37259

Job ID: 37259

Cluster: tc2

User/Group: zhuofei/gacrc-instruction

State: COMPLETED (exit code 0)

Cores: 1

CPU Utilized: 00:09:45

CPU Efficiency: 99.66% of 00:09:47 core-walltime

Job Wall-clock time: 00:09:47

Memory Utilized: 197.34 MB

Memory Efficiency: 4.82% of 4.00 GB

Obtain Job Details

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo2#How_to_check_resource_utilization_of_a_running_or_finished_job

Option 1: `squeue --me` for details of a pending or running jobs

Option 2: `sacct` or `sacct-gacrc (-v)` for details of computing resource usage of a running or finished job

Option 3: `seff` for details of computing resource usage of a finished job

Option 4: Email notification from finished jobs (completed, canceled, or crashed), if using:

```
#SBATCH --mail-user=username@uga.edu
```

```
#SBATCH --mail-type=ALL
```

Interactive jobs

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo2#How_to_open_an_interactive_session

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo2#How_to_run_an_interactive_job_with_Graphical_User_Interface_capabilities

Description	Command
Start an interactive session	qlogin
Start an interactive session with X forwarding	xqlogin

qlogin	srun --pty -p inter_p --mem=2G --nodes=1 --ntasks-per-node=1 --time=12:00:00 --job-name=qlogin bash -l
xqlogin	srun --pty --x11 -p inter_p --mem=2G --nodes=1 --ntasks-per-node=1 --time=12:00:00 --job-name=xqlogin bash -l

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

Kaltura channel <https://kaltura.uga.edu/channel/GACRC/176125031>

System: <https://wiki.gacrc.uga.edu/wiki/Systems#Sapelo2>

Connection: https://wiki.gacrc.uga.edu/wiki/Connecting#Connecting_to_Sapelo2

Software: https://wiki.gacrc.uga.edu/wiki/Software_on_Sapelo2

Running Jobs: https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo2

Monitoring Jobs: https://wiki.gacrc.uga.edu/wiki/Monitoring_Jobs_on_Sapelo2

Sample scripts : https://wiki.gacrc.uga.edu/wiki/Sample_batch_job_submission_scripts_on_Sapelo2

Transfer File: https://wiki.gacrc.uga.edu/wiki/Transferring_Files

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

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

GACRC Help and Support

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

➤ **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

When you ask GACRC to test or troubleshoot your jobs, Please make sure of the correctness of your datasets being used!

GACRC Service Catalog

Georgia Advanced Computing Resource Center (GACRC) service catalog

Services (11)

[Account Creation](#)

For a research group's PI to request user accounts for group members on the GACRC computing systems.

[Class Account Creation](#)

For an instructor to request user accounts for students attending a course that will need to use GACRC computing systems.

[Class Account Modification](#)

For instructors to request changes to be made in previously requested class account.

[Computing Lab Modification/Deletion](#)

[General Internal](#)

[General Support](#)

Report issues and request help with GACRC systems, except for software installation requests and account/lab creation requests.

[Lab Creation](#)

For a research group's PI to register a computing lab on the GACRC computing systems

[Modify/Delete Account](#)

For PIs to request changes in or deletion of user accounts on GACRC computing systems.

[Software Installation/Update](#)

Request software and common application database (e.g. NCBI blast databases) installation and upgrade.

My Recent Requests

[home directory is not fully provisioned: ss57215](#)

[GACRC Sapelo2 New Lab/Use Account Request 2018-11-14_preTraining](#)

[GACRC Sapelo2 Cluster New Lab/Use Account Request 2018-11-05_preTraining](#)

[provision 5 user accounts for ugahelpdesk group](#)

[GACRC Sapelo2 New Lab/Use Account Request 2018-10-22_preTraining](#)

[View All Recent Requests >](#)

Popular Services

[EITS Help Desk Support Request](#)

[MyID Account Request](#)

[Change Request](#)

[02 Restricted VPN Access](#)

[Terry Classroom & Meeting Room Support](#)

[View All Popular Services >](#)

My Recently Visited Services

[Modify/Delete Account](#)

[Class Account Creation](#)

Service - General Support - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Mail - zhuofei@uga.edu x Service - General Support x

https://uga.teamdynamix.com/TDClient/Requests/ServiceDet?ID=25844

90% ...

UNIVERSITY OF GEORGIA

Search the client portal

Zhuofei Hou

Home IT Help Desks Projects/Workspaces **Services** Knowledge Base News

Project Requests Ticket Requests My Favorite My Recent My Approvals Services A-Z Search

[Service Catalog](#) / [Academics, Learning & Research](#) / [GACRC Service Catalog](#) / General Support

General Support

If you do not have a myid, please mail gacrc-help@uga.edu, and we will respond promptly.

The purpose of this form is to provide a method to report issues and to request help with GACRC systems.

Please use this form for all questions and support needs (e.g. to report issues, to troubleshoot jobs, to request resources or grant writing help, etc). Please do not use this form for software installation requests or lab/user account management, which all have separate forms.

Please refer to the GACRC documentation for information on GACRC resources, how to connect and transfer files, how to run jobs, installed software list, training schedule, and a FAQ.

The link to this documentation is <https://wiki.gacrc.uga.edu>

Request Service

Share

Add to Favorites

← Click to request

This site is operated by [Enterprise Information Technology Services \(EITS\)](#) at the University of Georgia.

[Privacy](#) | [Accessibility](#) | [Website Feedback](#)

<https://uga.teamdynamix.com/TDClient/Requests/ServiceCatalogSearch>



🔍 Search the client portal

👤 Zhuofei Hou

General Support

+ Show Help - Hide Help

Report issues and request help with GACRC systems, except for software installation requests and account/lab creation requests.

Short Description *

Email *

MyID *

Phone Number *

Support Needed For

- Galaxy
- Sapelo2
- Teaching Cluster
- Work Filesystem
- Home Filesystem
- Scratch Filesystem
- Project Filesystem
- Xfer Nodes
- Other

Lab *

Slurm job states

Code	State	Meaning
R	Running	Job is running on compute node(s)
PD	Pending	Job is waiting for compute node(s)
CD	Completed	Job completed
CG	Completing	Job is completing
CA	Canceled	Job was canceled
F	Failed	Job terminated with non-zero exit code
NF	Node Fail	Job terminated due to failure of node(s)

Commands for submitting and canceling jobs

Description	Slurm Command
Submit a batch job to queue	<code>sbatch sub.sh</code>
Delete a job from queue	<code>scancel <jobID></code>
Cancel all your job(s)	<code>scancel -u <username></code>
Cancel all your pending job(s)	<code>scancel -t PENDING -u <username></code>
Cancel your job(s) by job name	<code>scancel --name <jobname></code>
Cancel an element (index) of an array job (jobID)	<code>scancel <jobID>_<index></code>

Commands for monitoring jobs

Description	Slurm Command
Command	<code>squeue</code> or <code>squeue -l</code>
Job status of all your jobs	<code>squeue --me</code> or <code>squeue --me -l</code>
Job status of a job	<code>squeue -j <jobID></code>
Job status of job(s) from a user	<code>squeue -u <username></code>
Job status with details	<code>scontrol show job <job ID></code> <code>scontrol show job -dd <job ID></code>
Job's resource usage	<code>sacct-gacrc</code>
View job batch script	<code>scontrol write batch_script <job ID> [filename]</code>

Slurm headers for running a **Serial (single-core)** job

https://wiki.gacrc.uga.edu/wiki/Sample_batch_job_submission_scripts_on_Sapelo2

```
#!/bin/bash
#SBATCH --job-name=testBowtie2          # Job name (testBowtie2)
#SBATCH --partition=batch              # Queue name (batch)
#SBATCH --ntasks=1                    # Run in a single task using one CPU core on a single node
#SBATCH --mem=4G                       # Job memory limit (4 GB)
#SBATCH --time=1:00:00                 # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --export=NONE                  # Do not load any users' explicit environment variables
#SBATCH --output=%x_%j.out             # Standard output log, e.g., testBowtie2_1234.out
#SBATCH --error=%x_%j.err              # Standard error log, e.g., testBowtie2_1234.err
#SBATCH --mail-type=END,FAIL           # Mail events (BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu   # Where to send mail

cd $SLURM_SUBMIT_DIR                   # Change directory to job submission directory
ml Bowtie2/2.4.1-GCC-8.3.0             # Load software module and run bowtie2 below
bowtie2 -x ./index/lambda_virus -U ./myreads.fq -S output.sam
```

Slurm headers for running a Threaded job

```
#!/bin/bash
#SBATCH --job-name=testBowtie2           # Job name (testBowtie2)
#SBATCH --partition=batch                # Queue name (batch)
#SBATCH --nodes=1                       # Run all processes on a single node
#SBATCH --ntasks=1                      # Run in a single task on a single node
#SBATCH --cpus-per-task=8                # Number of CPU cores per task (8)
#SBATCH --mem=10G                       # Job memory limit (10 GB)
#SBATCH --time=1:00:00                   # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --export=NONE                    # Do not load any users' explicit environment variables
#SBATCH --output=%x_%j.out               # Standard output log, e.g., testBowtie2_1234.out
#SBATCH --error=%x_%j.err                # Standard error log, e.g., testBowtie2_1234.err
#SBATCH --mail-type=END,FAIL              # Mail events (BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu     # Where to send mail

cd $SLURM_SUBMIT_DIR
ml Bowtie2/2.4.1-GCC-8.3.0
bowtie2 -p 8 -x ./index/lambda_virus -U ./myreads.fq -S output.sam
```

Slurm headers for running an **Array** job

```
#!/bin/bash
#SBATCH --job-name=testBowtie2Array      # Job name (testBowtie2Array)
#SBATCH --partition=batch                # Queue name (batch)
#SBATCH --ntasks=1                      # Run in a single task using one CPU core on a single node
#SBATCH --mem=4G                        # Job memory limit (4 GB)
#SBATCH --time=1:00:00                  # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --export=NONE                    # Do not load any users' explicit environment variables
#SBATCH --output=%x_%j.out               # Standard output log, e.g., testBowtie2Array_1234.out
#SBATCH --error=%x_%j.err                # Standard error log, e.g., testBowtie2Array_1234.err
#SBATCH --array=0-9                      # Array element range from 0 to 9, i.e. 10 element jobs

cd $SLURM_SUBMIT_DIR
ml Bowtie2/2.4.1-GCC-8.3.0                # Original data is split into 10 pieces and run in each element job
bowtie2 -x ./index/lambda_virus -U ./myreads_$$SLURM_ARRAY_TASK_ID.fq \
-S output_$$SLURM_ARRAY_TASK_ID.sam
```




Slurm headers for running a Threaded (OpenMP) job

```
#!/bin/bash
#SBATCH --job-name=testOpenMP           # Job name (testOpenMP)
#SBATCH --partition=batch                # Queue name (batch)
#SBATCH --nodes=1                       # Run all processes on a single node
#SBATCH --ntasks=1                      # Run in a single task on a single node
#SBATCH --cpus-per-task=12              # Number of CPU cores per task (12)
#SBATCH --mem=10G                       # Job memory limit (10 GB)
#SBATCH --export=NONE                   # Do not load any users' explicit environment variables
#SBATCH --time=24:00:00                 # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --output=%x_%j.log             # Standard output and error log, e.g., testOpenMP_1234.log
#SBATCH --mail-type=END,FAIL           # Mail events (BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu    # Where to send mail

cd $SLURM_SUBMIT_DIR
export OMP_NUM_THREADS=$SLURM_CPUS_PER_TASK # Sets the number of threads to use for OpenMP parallel regions
ml foss/2019b                               # Load toolchain module
./myProgram.x                               # Run your program binary compiled with OpenMP
```

Slurm headers for running an MPI job

```
#!/bin/bash
#SBATCH --job-name=testMPI           # Job name (tesMPI)
#SBATCH --partition=batch           # Queue name (batch)
#SBATCH --nodes=2                   # Run on two nodes
#SBATCH --ntasks-per-node=16       # How many tasks on each node; Number of tasks=32=MPI ranks
#SBATCH --cpus-per-task=1           # Number of CPU cores per task; 16 CPU cores per node
#SBATCH --mem-per-cpu=500M          # Memory per allocated CPU; 8GB (500MB*16) memory per node
#SBATCH --time=24:00:00             # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --export=NONE               # Do not load any users' explicit environment variables
#SBATCH --output=%x_%j.log          # Standard output and error log, e.g., testMPI_1234.log
#SBATCH --mail-type=END,FAIL        # Mail events (BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu # Where to send mail

cd $SLURM_SUBMIT_DIR
ml foss/2019b                       # Load toolchain module
mpirun -n 32 ./myProgram.x          # Run your program binary compiled with OpenMPI with 32 ranks
```



Slurm headers for running a Hybrid MPI/OpenMP job

```
#!/bin/bash
#SBATCH --job-name=testHybrid          # Job name (testHybrid)
#SBATCH --partition=batch              # Queue name (batch)
#SBATCH --nodes=2                     # Run on two nodes
#SBATCH --ntasks-per-node=8           # How many tasks on each node; Number of tasks=16=MPI ranks
#SBATCH --cpus-per-task=4             # Number of CPU cores per task; 32 CPU cores per node
#SBATCH --mem-per-cpu=500M            # Memory per allocated CPU; 16GB (500MB*32) memory per node
#SBATCH --time=24:00:00               # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --export=NONE                 # Do not load any users' explicit environment variables
#SBATCH --output=%x_%j.log            # Standard output and error log
#SBATCH --mail-type=END,FAIL          # Mail events (BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu  # Where to send mail

cd $SLURM_SUBMIT_DIR                  # Change directory to job submission directory
export OMP_NUM_THREADS=$SLURM_CPUS_PER_TASK # Sets the number of threads to use for OpenMP parallel regions
ml foss/2019b                          # Load toolchain module
mpirun -n 16 ./myProgram.x            # Run your program binary compiled with OpenMPI with 16 ranks
```

Slurm headers for running a GPU job

```
#!/bin/bash
#SBATCH --job-name=amber                # Job name
#SBATCH --partition=gpu_p                # Partition (queue) name, i.e., gpu_p
#SBATCH --gres=gpu:1                     # Requests one GPU device; --gres=gpu:P100:1, --gres=gpu:K40:1
#SBATCH --ntasks=1                       # Run a single task
#SBATCH --cpus-per-task=2                # Number of CPU cores per task
#SBATCH --mem=40gb                       # Job memory request
#SBATCH --time=10:00:00                  # Time limit hrs:min:sec
#SBATCH --export=NONE                    # Do not load any users' explicit environment variables
#SBATCH --output=%x_%j.out               # Standard output and error log, e.g., amber_1234.out
#SBATCH --mail-type=END,FAIL             # Mail events (BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu     # Where to send mail

cd $SLURM_SUBMIT_DIR
ml Amber/18-fosscuda-2018b-AmberTools-18-patchlevel-10-8
mpirun $AMBERHOME/bin/pmemd.cuda -O -i ./prod.in -o prod_c4-23.out -p ./dimerFBP_GOL.prmtop -c ./restart.rst \
-r prod.rst -x prod.mdcrd
```



Slurm headers for running a Singularity container

```
#!/bin/bash
#SBATCH --job-name=test_sortmerna           # Job name
#SBATCH --partition=batch                   # Partition (queue) name
#SBATCH --ntasks=1                          # Run on a single CPU
#SBATCH --mem=8gb                           # Job memory request
#SBATCH --time=02:00:00                     # Time limit hrs:min:sec
#SBATCH --export=NONE                       # Do not load any users' explicit environment variables
#SBATCH --output=sortmerna.%j.out           # Standard output log, e.g., sortmerna.1234.out
#SBATCH --error=sortmerna.%j.err           # Standard error log, e.g., sortmerna.1234.err
#SBATCH --cpus-per-task=4                   # Number of CPU cores per task
#SBATCH --mail-type=END,FAIL                # Mail events (NONE, BEGIN, END, FAIL, ALL)
#SBATCH --mail-user=username@uga.edu       # Where to send mail

cd $SLURM_SUBMIT_DIR
singularity exec /apps/singularity-images/sortmerna-3.0.3.simg sortmerna --threads 4 --ref db.fasta,db.idx --reads file.fa \
--aligned base_name_output
```

Guideline Tips

- Do NOT use Login node to run CPU/memory intensive tasks directly → submit jobs to compute nodes!
- Do NOT use Login Node to transfer data between your local computer and cluster → use Transfer node!
- Do NOT use Home for storing job data → use /scratch/MyID
- Do NOT park data in Scratch or Local Scratch → clean up when job finishes or exits from node
- Do NOT park data permanently in Project → download data to your local drive
- NO large memory job running on batch partition → use highmem_p
- NO small memory job running on highmem_p partition → use batch
- In general, number of threads you want to run with a parallel job = number of cores requested
- When you archive data using tar on /scratch, please do not use z option (compression option). After you achieved data by tar, you can use gzip to compress it.

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! ✅

Thank You!

Telephone Support

EITS Help Desk: 706-542-3106

Monday – Thursday: 7:30 a.m. – 7:30 p.m.

Friday: 7:30 a.m. – 6 p.m.

Saturday – Sunday: 1 p.m. – 7 p.m.

Georgia Advanced Computing Resource Center

101-108 Computing Services building

University of Georgia

Athens, GA 30602

<https://gacrc.uga.edu/>