Introduction to GACRC Teaching Cluster

Georgia Advanced Computing Resource Center (GACRC)
Enterprise Information Technology Services (EITS)
The University of Georgia
Outline

• GACRC
• Overview
• Working Environment
  ➢ Three Folders
  ➢ Three Computational Partitions
  ➢ Software on Cluster
• Submit a Computational Batch Job
• Run Interactive Jobs
• GACRC Wiki and Support
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
Support: https://wiki.gacrc.uga.edu/wiki/Getting_Help
Web Site: http://gacrc.uga.edu
Kaltura Channel: https://kaltura.uga.edu/channel/GACRC/176125031
Note: You need to connect to the UGA VPN at first when accessing from outside of the UGA main campus.
Working Environment

https://wiki.gacrc.uga.edu/wiki/Systems#Teaching_cluster

- Two nodes, your "username" is your MyID for both of them:
  1. For batch job workflow, the host to log into is teach.gacrc.uga.edu
  2. For file transfers, the host to log into is txfer.gacrc.uga.edu

- Three Directories:
  1. /home/MyID: directory for static data (e.g., scripts, software, etc...)
  2. /scratch/MyID: working space for running computational jobs
  3. /work/bcmb8330: directory for storing data
     a. /work/bcmb8330/MyID: data storage space for individual user in class
     b. /work/bcmb8330/instructor_data: data shared with class by the instructor

- Three Partitions:
  1. batch: for running regular computational jobs
  2. highmem: for running high-memory jobs
  3. gpu: for running GPU jobs
Working Environment (cont.)

- Software
  1. Software names are long and have a Easybuild toolchain name associated to it
  2. Complete module name: Name/Version-toolchain, e.g., Python/3.8.2-GCCcore-8.3.0
  3. Software names are case-sensitive!
     - module spider pattern: Search modules using a name pattern (case-insensitive)
     - module load/unload moduleName: Load/remove a module
     - module avail: List all available modules on the cluster
     - module list: List modules currently loaded
     - module purge: Remove all modules from working environment
Submit a Batch Job

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

1. Log on to Login node using MyID and password, and two-factor authentication with Archpass Duo:
   
   ```bash
   ssh MyID@teach.gacrc.uga.edu
   ```

2. Change directory to /scratch directory:
   
   ```bash
   cd /scratch/MyID
   ```

3. Create a working subdirectory for a job:
   
   ```bash
   mkdir workDir
   ```

4. Change directory to workDir:
   
   ```bash
   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: use `nano` sub.sh

7. Submit a job from workDir: use `sbatch` sub.sh

8. Check job status: use `squeue --me` or Cancel a job: use `scancel` JobID
Step 1: Log on to Login node
https://wiki.gacrc.uga.edu/wiki/Connecting#Connecting_to_the_teaching_cluster

1. Teaching cluster access requires 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.

2. If you are connecting from off-campus, please first connect to the UGA VPN and then connect to teach.gacrc.uga.edu. Information on how to use the VPN is available at https://eits.uga.edu/access_and_security/infosec/tools/vpn/
Step 1: Log on to Login node - Mac/Linux using ssh

1. Open **Terminal** utility

2. Type command line: `ssh MyID@teach.gacrc.uga.edu`

3. You will be prompted for your **UGA MyID password**

4. You will verify your login using **Archpass Duo** authentication
ssh zhuofei@teach.gacrc.uga.edu

1. use ssh to open connection

UGA DUO authentication is required for SSH/SCP access to GACRC systems. For additional help with UGA DUO authentication or to report an issue please visit: https://eits.uga.edu/access_and_security...

Password: 2. Enter your MyID password
When you enter password, no stars or dots will show as you are typing. Please type password carefully!

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 5. SMS passcodes to XXX-XXX-5758 (next code starts with: 1)

Passcode or option (1-5): 1
Success. Logging you in...

5. Verify login using Duo

Last login: Mon Aug 3 11:11:58 2020 from 172.18.114.119
zhuofei@teach-sub1 ~$
Step1 (Cont.) - Windows using PuTTY

1. Download and install PuTTY: [https://www.putty.org/](https://www.putty.org/)

2. Detailed downloading and installation instructions:
   [https://wiki.gacrc.uga.edu/wiki/How_to_Install_andConfigure_PuTTY](https://wiki.gacrc.uga.edu/wiki/How_to_Install_andConfigure_PuTTY)

3. Detailed configuring and usage instructions:
   [https://wiki.gacrc.uga.edu/wiki/How_to_Install_andConfigure_PuTTY#Configuring_PuTTY](https://wiki.gacrc.uga.edu/wiki/How_to_Install_andConfigure_PuTTY#Configuring_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:

- UGA MyID password
- Select DUO option
- Logged on!
Step2 - 4: cd to /scratch dir, make and cd into workDir

zhuofei@teach-sub1 ~$ cd /scratch/zhuofei
⇒ cd command to change directory
zhuofei@teach-sub1 zhuofei$ mkdir workDir
⇒ mkdir command to create a subdirectory
zhuofei@teach-sub1 zhuofei$ cd workDir/
⇒ cd command to change directory
zhuofei@teach-sub1 workDir$ ls
⇒ ls command to list contents of directory
zhuofei@teach-sub1 workDir$
⇒ it is empty in workDir!
Step 5: Transfer data from local computer to workDir - Mac/Linux

https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_scp_2

1. Connect to Transfer node (txfer.gacrc.uga.edu) in Terminal from your local computer
2. Use scp command: scp (-r) [Source] [Target]
3. Enter your MyID password, then select Duo option to verify connection

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

```bash
scp ./file zhuofei@txfer.gacrc.uga.edu:/home/zhuofei/workDir
scp -r ./folder/ zhuofei@txfer.gacrc.uga.edu:/home/zhuofei/workDir
```

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

```bash
scp zhuofei@txfer.gacrc.uga.edu:/home/zhuofei/workDir/file .
scp -r zhuofei@txfer.gacrc.uga.edu:/home/zhuofei/workDir/folder/ .
```
Step 5 (Cont.) - Windows using WinSCP
https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_WinSCP_2

1. You need to connect to cluster’s Transfer node (txfer.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

Step 5 (Cont.) - Windows using WinSCP

![WinSCP Screenshot]

- Login to the GACRC Teaching Cluster using WinSCP.
  - Host name: ufer.gacrce.uga.edu
  - Port number: 22
  - Username: aduser
  - Password: **********
  - Click on the 'Login' button.
Step 5 (Cont.) - Windows using WinSCP

Select DUO option
Step 5 (Cont.) - Windows using WinSCP

Change paths on your local computer and transfer node

Drag to transfer files or folders
Step 5 (Cont.): Transfer data on cluster to workDir

- Log on to Transfer node (txfer.gacrc.uga.edu)
  - Mac/Linux: ssh MyID@txfer.gacrc.uga.edu (page 9-10)
  - Windows: use PuTTY to log in MyID@txfer.gacrc.uga.edu (page 11-13)

- Directories you can access on transfer node:
  1. /home/MyID
  2. /scratch/MyID
  3. /work/bcmb8330/

- Transfer data between two folders on cluster using cp or mv, e.g.:

  mv /work/bcmb8330/MyID/datafile /scratch/MyID/workDir
Step 6: Make a job submission script in workDir using nano
https://wiki.gacrc.uga.edu/wiki/Sample_batch_job_submission_scripts_on_the_teaching_cluster

$ 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
Step 6 (Cont.)

1. Copy sample job to workDir:
   
cp -r /usr/local/training/bcmb8330/* .

2. Job submission script:
   
sub.sh

3. Amber Wiki:
   
https://wiki.gacrc.uga.edu/wiki/AMBER-R-Sapelo2

#!/bin/bash
#SBATCH --job-name=Amber_test
#SBATCH --partition=batch
#SBATCH --nodes=1
#SBATCH --ntasks=10
#SBATCH --cpus-per-task=1
#SBATCH --mem-per-cpu=2gb
#SBATCH --time=24:00:00
#SBATCH --output=minimization.%j.out
#SBATCH --error=minimization.%j.err
#SBATCH --mail-user=MyID@uga.edu
#SBATCH --mail-type=ALL

ml Amber/18-foss cuda-2018b-AmberTools-18-patchlevel-10-8
source ${AMBERHOME}/amber.sh

cd $SLURM_SUBMIT_DIR

# PMEMD: Job1: minimization, solvent
mpiexec --mca mpi_cuda_support 0 ${AMBERHOME}/bin/pmemd.MPI -O 
   -i min_solvent.in
   -o min_solvent.out
   -p gfp.parm7
   -c gfp.rst7
   -ref gfp.rst7
   -r gfp_min_solvent.rst7

# Job name
# Computational partition
# Number of compute nodes
# MPI processes/rank number
# Number of CPU cores per task/MPI process
# Memory per CPU core
# Time limit hrs:mins:secs
# Standard output log
# Standard error log
# Where to send mail
# Mail events (BEGIN, END, FAIL, ALL)
$ sbatch sub.sh
Submitted batch job 33780

Tips: sub.sh is a job submission script for

1. specifying computing resources
2. loading software using module load
3. running any Linux commands you want to run
4. running the Amber MPI command
Step 8: Check job status using squeue

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

```
$ squeue --me -l
Tue Jan 17 09:25:20 2023
JOBID PARTITION NAME USER    STATE TIME TIME_LIMI NODES NODELIST(REASON)
33780 batch Amber_te zhuofei RUNNING 0:23 1-00:00:00 1 tcn18

$ squeue -me
JOBID PARTITION NAME USER ST TIME NODES NODELIST(REASON)
33780 batch Amber_te zhuofei R 2:15 1 tcn18
```

Common STATE: R for Running; PD for PenDing; TO for TimedOut; S for Suspended; F for FAILED
TIME: the elapsed time used by the job, not remaining time, not CPU time.
Step 8 (Cont.): Cancel job using `scancel`

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_the_teaching_cluster#How_to_delete_a_running_or_pending_job

```bash
$ squeue --me -l
Tue Jan 17 09:29:34 2023
JOBID PARTITION NAME      USER    STATE       TIME TIME_LIMI  NODES NODELIST(REASON)
33780 batch               Amber_te zhuofei RUNNING  4:37 1-00:00:00 1  tcn18

$ scancel 33780

$ squeue --me -l
Tue Jan 17 09:29:42 2023
JOBID PARTITION NAME      USER    STATE       TIME TIME_LIMI  NODES NODELIST(REASON)
33780 batch               Amber_te zhuofei COMPLETED 4:44 1-00:00:00 1  tcn18

$ squeue --me -l
Tue Jan 17 09:29:45 2023
JOBID PARTITION NAME      USER    STATE       TIME TIME_LIMI  NODES NODELIST(REASON)
```

Step 8 (Cont.): Check job details using `sacct-gacrc -X` and `seff`

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

```bash
$ sacct-gacrc -X

JobID  JobName  User  Partition  NNode  NCPUS  ReqMem  CPURusage  Elapsed  Timelimit  State  ExitCode  NodeList
-------  --------  -----  ---------  ------  ------  -------  ----------  -------  ----------  ------  --------  -------
33780   Amber_test  zhuofei  batch  1  10  2Gc  02:31:28  00:04:44  1-00:00:00  CANCELLED+  0:0  tcn18
33798   Amber_test  zhuofei  batch  1  10  2Gc  00:34:40  00:03:28  1-00:00:00  COMPLETED  0:0  tcn18
```

$ seff 33798

```
# Check computing resources used by a COMPLETED job

Job ID: 33798
Cluster: gacrc-teach
User/Group: zhuofei/ gacrc-instruction
State: COMPLETED (exit code 0)
Nodes: 1
Cores per node: 10
CPU Utilized: 00:34:25
CPU Efficiency: 99.28% of 00:34:40 core-walltime
Job Wall-clock time: 00:03:28
Memory Utilized: 363.66 MB
Memory Efficiency: 1.78% of 20.00 GB
```
Step 8 (Cont.): Check node info using sinfo

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

$ sinfo

<table>
<thead>
<tr>
<th>PARTITION</th>
<th>AVAIL</th>
<th>TIMELIMIT</th>
<th>NODES</th>
<th>STATE</th>
<th>NODELIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch*</td>
<td>up 7-00:00:00</td>
<td>1</td>
<td>drain</td>
<td>tcn17</td>
<td></td>
</tr>
<tr>
<td>batch*</td>
<td>up 7-00:00:00</td>
<td>4</td>
<td>mix</td>
<td>tcn[18-21]</td>
<td></td>
</tr>
<tr>
<td>batch*</td>
<td>up 7-00:00:00</td>
<td>20</td>
<td>idle</td>
<td>tcn[1-16,22-25]</td>
<td></td>
</tr>
<tr>
<td>interactive</td>
<td>up 7-00:00:00</td>
<td>5</td>
<td>idle</td>
<td>tcn[26-30]</td>
<td></td>
</tr>
<tr>
<td>gpu</td>
<td>up 7-00:00:00</td>
<td>1</td>
<td>idle</td>
<td>tcgn1</td>
<td></td>
</tr>
<tr>
<td>highmem</td>
<td>up 7-00:00:00</td>
<td>1</td>
<td>idle</td>
<td>tchmn2</td>
<td></td>
</tr>
<tr>
<td>highmem</td>
<td>up 7-00:00:00</td>
<td>1</td>
<td>down</td>
<td>tchmn1</td>
<td></td>
</tr>
<tr>
<td>fsr4601</td>
<td>up 1:00</td>
<td>1</td>
<td>drain</td>
<td>tcn17</td>
<td></td>
</tr>
<tr>
<td>fsr4601</td>
<td>up 1:00</td>
<td>4</td>
<td>mix</td>
<td>tcn[18-21]</td>
<td></td>
</tr>
<tr>
<td>fsr4601</td>
<td>up 1:00</td>
<td>20</td>
<td>idle</td>
<td>tcn[1-16,22-25]</td>
<td></td>
</tr>
<tr>
<td>fsr8602</td>
<td>up 10:00</td>
<td>1</td>
<td>drain</td>
<td>tcn17</td>
<td></td>
</tr>
<tr>
<td>fsr8602</td>
<td>up 10:00</td>
<td>4</td>
<td>mix</td>
<td>tcn[18-21]</td>
<td></td>
</tr>
<tr>
<td>fsr8602</td>
<td>up 10:00</td>
<td>20</td>
<td>idle</td>
<td>tcn[1-16,22-25]</td>
<td></td>
</tr>
</tbody>
</table>
Obtain Job Details

https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_the_teaching_cluster#How_to_check_resource_utilization_of_a_running_or.finished_job

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

Option 2: `sacct-gacrc -X` 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`
Run 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

<table>
<thead>
<tr>
<th>Description</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start an interactive session</td>
<td>interact</td>
</tr>
<tr>
<td>Start an interactive session with X forwarding</td>
<td>interact --x11</td>
</tr>
</tbody>
</table>

| interact | srun --pty --cpus-per-task=1 --job-name=interact --ntasks=1 --nodes=1 --partition=inter_p --time=12:00:00 --mem=2GB /bin/bash -l |
| interact --x11 | srun --pty --cpus-per-task=1 --job-name=interact --ntasks=1 --nodes=1 --partition=inter_p --time=12:00:00 --mem=2GB --x11 /bin/bash -l |
GACRC Wiki [http://wiki.gacrc.uga.edu](http://wiki.gacrc.uga.edu)
Kaltura Channel [https://kaltura.uga.edu/channel/GACRC/176125031](https://kaltura.uga.edu/channel/GACRC/176125031)

Connecting: [https://wiki.gacrc.uga.edu/wiki/Connecting#Connecting_to_the_teaching_cluster](https://wiki.gacrc.uga.edu/wiki/Connecting#Connecting_to_the_teaching_cluster)

Running Jobs: [https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_the_teaching_cluster](https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_the_teaching_cluster)

Monitoring Jobs: [https://wiki.gacrc.uga.edu/wiki/Monitoring_Jobs_on_the_teaching_cluster](https://wiki.gacrc.uga.edu/wiki/Monitoring_Jobs_on_the_teaching_cluster)

Transfer File: [https://wiki.gacrc.uga.edu/wiki/Transferring_Files#The_File_Transfer_node_for_the_teaching_cluster_28txfer.gacrc.uga.edu_29](https://wiki.gacrc.uga.edu/wiki/Transferring_Files#The_File_Transfer_node_for_the_teaching_cluster_28txfer.gacrc.uga.edu_29)

Sample Job Scripts: [https://wiki.gacrc.uga.edu/wiki/Sample_batch_job_submission_scripts_on_the_teaching_cluster](https://wiki.gacrc.uga.edu/wiki/Sample_batch_job_submission_scripts_on_the_teaching_cluster)

Linux Command: [https://wiki.gacrc.uga.edu/wiki/Command_List](https://wiki.gacrc.uga.edu/wiki/Command_List)
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 partition 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

Please note to make sure the correctness of datasets being used by your jobs!
Georgia Advanced Computing Resource Center (GACRC) service catalog.

If you would like to reach out to GACRC and do not have a UGA MyID, please send an email to gacrc-help@uga.edu, and we will respond promptly.

Categories (3)

Services For Users
General user support, request software installation or update, request training.

Services for PIs
For PIs only: Lab registration, user account creation/modification, class account requests, storage quota modifications.

For GACRC Staff
For GACRC's internal use only.
General Support

If you do not have a myID, please mail gacrclab@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 granting help, etc.). Please do not use this form for software installation requests or let user account management, which all have separate tickets.

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

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

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

Privacy | Accessibility | Website Feedback

Click to request
Need Support?  http://help.gacrc.uga.edu