Introduction to GACRC Teaching Cluster

PHYS8602

Georgia Advanced Computing Resource Center (GACRC)

Enterprise Information Technology Services (EITS)

The University of Georgia
Outline

- GACRC
- Overview
- Working Environment
  - Two Nodes and Three Folders
  - Computational Partitions
  - Software
- Submit a Computational Batch Job
- 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](http://wiki.gacrc.uga.edu)
Support: [https://wiki.gacrc.uga.edu/wiki/Getting_Help](https://wiki.gacrc.uga.edu/wiki/Getting_Help)
Web Site: [http://gacrc.uga.edu](http://gacrc.uga.edu)
Kaltura Channel: [https://kaltura.uga.edu/channel/GACRC/176125031](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.

1. ssh with MyID and password
2. Verify with Archpass Duo two-factor authentication

- **Node**: Computer for a specific function on cluster, e.g., login node
- **Partition**: Collection of **compute nodes** for specific computing need
- **Cluster**: **Nodes + Drives**, all connected by network
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 folders:
  1. /home/MyID: working space for running computational jobs
  2. /work/phys8602/MyID: data storing space for individual user in a class
  3. /work/phys8602/instructor_data: data shared with class by the instructors

- Partitions for your class: fsr8602
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 module names matching a pattern (case-insensitive)
   - module load moduleName: Load a module into your working environment
   - module avail: List all available software modules installed on cluster
   - module list: List modules currently loaded
   - module unload moduleName: Remove a module from working environment
Submit a Computational Batch Job

1. Log on to Login node using MyID and password, and two-factor authentication with Archpass Duo:
   \[ \text{ssh MyID@teach.gacrc.uga.edu} \]

2. Change directory to your scratch space:
   \[ \text{cd /scratch/MyID} \]

3. Create a working subdirectory for a job:
   \[ \text{mkdir ./workDir} \]

4. Change directory to workDir:
   \[ \text{cd ./workDir} \]

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. Compile C code \textit{mult.c} into a binary code

7. Make a job submission script in workDir:
   \[ \text{nano ./sub.sh} \]

8. Submit a job from workDir:
   \[ \text{sbatch ./sub.sh} \]

9. Check job status: \texttt{squeue --me (-l)} or Cancel a job:
   \[ \text{scancel JobID} \]

---

GACRC TEACHING CLUSTER NEW USER TRAINING WORKSHOP

1/9/2023
Step1: 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. SMS passcodes to XXX-XXX-5758 (next code starts with: 1)

Passcode or option (1-5): 1  ➕  3. Select Duo option
Success. Logging you in...
Last login: Mon Aug 3 11:11:58 2020 from 172.18.114.119
zhuofei@teach-sub1 ~$  ➕  4. Logged on!

5. Verify login using Duo
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

3. Detailed configuring and usage instructions:
   
   https://wiki.gacrc.uga.edu/wiki/How_to_Install_and_Configure_PuTTY#Configuring_PuTTY
The first time you connect to login node, PuTTY will give you this security alert window. Please click "Yes"
Next you will enter your UGA MyID password and initiate DUO authentication procedure:

- **UGA MyID password**
- **Select DUO option**
- **Logged on!**
Step 2 - 4: cd to /scratch dir, make and cd into workDir

```
zhuofei@teach-sub1 ~$ cd /scratch/zhuofei
zhuofei@teach-sub1 zhuofei$ mkdir workDir
zhuofei@teach-sub1 zhuofei$ cd workDir/
zhuofei@teach-sub1 workDir$ ls
zhuofei@teach-sub1 workDir$
```

- cd command to change directory
- mkdir command to create a subdirectory
- cd command to change directory
- ls command to list contents of directory
- it is empty in workDir!
Step5: 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

```
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

```
scp zhuofei@txfer.gacrc.uga.edu:/home/zhuofei/workDir/file.
scp -r zhuofei@txfer.gacrc.uga.edu:/home/zhuofei/workDir/folder/.
```
Step5 (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

Step5 (Cont.) - Windows using WinSCP
https://wiki.gacrc.uga.edu/wiki/Transferring_Files#Using_WinSCP_2
Step 5 (Cont.) - Windows using WinSCP

Select DUO option
Step5 (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 (Landing home)
  2. /work/phys8602/MyID
  3. /work/phys8602/instructor_data

• Transfer data between two folders on cluster using `cp` or `mv`, e.g.:
  ```
  mv /work/phys8602/MyID/datafile /home/MyID/workDir
  ```
Step 6: Compile C code `mult.c` into a binary

```
zhuofei@teach-sub1 workDir$ interact
zhuofei@tcn26 workDir$ cp /usr/local/training/phys8602/mult.c .
zhuofei@tcn26 workDir$ cat mult.c
/* Program mult
 * Multiple two integer numbers */
#include <stdio.h>
int main(void)
{
    int i=3, j=4, iprod;
    FILE *fp;
    fp = fopen("output.txt","w");
    iprod=i*j;
    fprintf(fp, "The product of %d and %d is %d\n", i,j,iprod);
    fclose(fp);
    return 0;
}
zhuofei@tcn26 workDir$ module load GCC/8.3.0
zhuofei@tcn26 workDir$ gcc mult.c -o mult.x
zhuofei@tcn26 workDir$ ls
mult.c mult.x
zhuofei@tcn26 workDir$ exit
```

- Start an interactive session
- Copy source code to your working dir
- Load GCC compiler module
- Compile source code into a binary
- Binary is generated in your working dir
- Exit from interactive session
Step7: Make a job submission script `sub.sh` using `nano`.

- Copy `sub.sh` to your working dir:
  ```bash
cp /usr/local/training/phys8602/sub.sh
```
- Show contents of `sub.sh`:
  ```bash
cat sub.sh
```
- Make a job submission script `sub.sh` using `nano`:
  ```bash
#!/bin/bash
#SBATCH --job-name=test
#SBATCH --partition=fsr8602
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=1
#SBATCH --mem=2gb
#SBATCH --time=00:10:00
#SBATCH --output=log.%j
#SBATCH --mail-user=MyID@uga.edu
#SBATCH --mail-type=ALL

cd $SLURM_SUBMIT_DIR
module load GCC/8.3.0
time ./mult.x
```
- Use `nano` to make modifications to `sub.sh`, e.g., email address:
  ```bash
  nano sub.sh
  ```
Step 8: Submit a job from workDir using sbatch

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

$ sbatch sub.sh
Submitted batch job 12109

**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 your binary code
Step 9: Check job status using squeue

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

zhuofei@teach-sub1 workDir$ squeue --me

<table>
<thead>
<tr>
<th>JOBID</th>
<th>PARTITION</th>
<th>NAME</th>
<th>USER</th>
<th>ST</th>
<th>TIME</th>
<th>NODES</th>
<th>NODELIST (REASON)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12109</td>
<td>fsr8602</td>
<td>test</td>
<td>zhuofei</td>
<td>R</td>
<td>0:05</td>
<td>1</td>
<td>tcn18</td>
</tr>
</tbody>
</table>

zhuofei@teach-sub1 workDir$ squeue --me -l

Mon Jan 11 12:03:14 2021

<table>
<thead>
<tr>
<th>JOBID</th>
<th>PARTITION</th>
<th>NAME</th>
<th>USER</th>
<th>STATE</th>
<th>TIME</th>
<th>TIME_LIMI</th>
<th>NODES</th>
<th>NODELIST (REASON)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12109</td>
<td>fsr8602</td>
<td>test</td>
<td>zhuofei</td>
<td>RUNNING</td>
<td>0:11</td>
<td>10:00</td>
<td>1</td>
<td>tcn18</td>
</tr>
</tbody>
</table>

Job State: R for Running; PD for PenDing; F for Failed
TIME: the elapsed time used by the job, not remaining time, not CPU time
Step9 (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

```
zhuofei@teach-sub1 workDir$ scancel 12109
zhuofei@teach-sub1 workDir$ squeue --me
```

<table>
<thead>
<tr>
<th>JOBID</th>
<th>PARTITION</th>
<th>NAME</th>
<th>USER</th>
<th>ST</th>
<th>TIME</th>
<th>NODES</th>
<th>NODELIST(REASON)</th>
</tr>
</thead>
</table>
Step9 (Cont.): Check job details using sacct-gacrc -X and seff

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

$ sacct-gacrc -X

<table>
<thead>
<tr>
<th>JobID</th>
<th>JobName</th>
<th>User</th>
<th>Partition</th>
<th>NodeList</th>
<th>AllocNodes</th>
<th>NTask</th>
<th>NCPUS</th>
<th>ReqMem</th>
<th>MaxVMSize</th>
<th>State</th>
<th>CPUTime</th>
<th>Elapsed</th>
<th>Timelimit</th>
<th>ExitCode</th>
</tr>
</thead>
<tbody>
<tr>
<td>33561</td>
<td>test</td>
<td>zhuofei</td>
<td>fsr4601</td>
<td>tcn18</td>
<td>1</td>
<td>1</td>
<td>2Gn</td>
<td></td>
<td></td>
<td>COMPLETED</td>
<td>00:00:01</td>
<td>00:00:01</td>
<td>00:01:00</td>
<td>0:0</td>
</tr>
</tbody>
</table>

$ seff 33561  # Check computing resources used by a COMPLETED job

Job ID: 33561
Cluster: gacrc-teach
User/Group: zhuofei/gacrc-instruction
State: COMPLETED (exit code 0)
Cores: 1
CPU Utilized: 00:00:00
CPU Efficiency: 0.00% of 00:00:01 core-walltime
Job Wall-clock time: 00:00:01
Memory Utilized: 1.27 MB
Memory Efficiency: 0.06% of 2.00 GB
Step 7 (Cont.): Check node info using `sinfo`

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

<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</td>
<td>7-00:00:00</td>
<td>1</td>
<td>down*</td>
<td>tcn17</td>
</tr>
<tr>
<td>batch*</td>
<td>up</td>
<td>7-00:00:00</td>
<td>24</td>
<td>idle</td>
<td>tcn[1-16,18-25]</td>
</tr>
<tr>
<td>interactive</td>
<td>up</td>
<td>7-00:00:00</td>
<td>5</td>
<td>idle</td>
<td>tcn[26-30]</td>
</tr>
<tr>
<td>gpu</td>
<td>up</td>
<td>7-00:00:00</td>
<td>1</td>
<td>idle</td>
<td>tcgn1</td>
</tr>
<tr>
<td>highmem</td>
<td>up</td>
<td>7-00:00:00</td>
<td>2</td>
<td>idle</td>
<td>tchmn[1-2]</td>
</tr>
<tr>
<td>fsr4601</td>
<td>up</td>
<td>1:00</td>
<td>1</td>
<td>down*</td>
<td>tcn17</td>
</tr>
<tr>
<td>fsr4601</td>
<td>up</td>
<td>1:00</td>
<td>24</td>
<td>idle</td>
<td>tcn[1-16,18-25]</td>
</tr>
<tr>
<td>fsr8602</td>
<td>up</td>
<td>10:00</td>
<td>1</td>
<td>down*</td>
<td>tcn17</td>
</tr>
<tr>
<td>fsr8602</td>
<td>up</td>
<td>10:00</td>
<td>24</td>
<td>idle</td>
<td>tcn[1-16,18-25]</td>
</tr>
</tbody>
</table>

`idle` = no cores in use; `mix` = some cores are still free; `alloc` = all cores are allocated
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: `seff` for details of computing resource usage of a **finished** job

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

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

```
#SBATCH --mail-user=username@uga.edu
#SBATCH --mail-type=END,FAIL
```
Connecting: 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

Monitoring Jobs: 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

Sample Job Scripts:
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
GACRC 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 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!
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.
Click to request