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 Queues
  ➢ Software

• Submit Batch Job

• Transfer Data

• 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
Help and Support: https://wiki.gacrc.uga.edu/wiki/Getting_Help
Web Site: http://gacrc.uga.edu
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/
Working Environment

- 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 : working space for computational jobs
  2. /work/CourseID/MyID : data parking for individual user in the class (/work/gene4220/MyID)
  3. /work/CourseID/instructor_data : data shared with class by the instructors

- Three Queues:
  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 an 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 Batch Job

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

2. Open an interactive session to compile ens code: `qlogin`

3. Change directory to ens folder: `cd ens`

4. Load PGI/19.10 compiler module: `ml PGI`

5. Compile source code ens.f into a binary executable: `pgfortran -O ens.f -o ens`

6. Make a job submission script using nano: `nano ens.sh`

7. Submit a job: `sbatch ens.sh`

8. Check job status: `squeue --me -l` or Cancel a job: `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
https://wiki.gacrc.uga.edu/wiki/Connecting#Connecting_to_the_teaching_cluster

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. Log on

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 login option 1
Success. Logging you in...
Last login: Mon Aug 3 11:11:58 2020 from 172.18.114.119
zhuofei@teach-sub1 ~$ ➔ 5. Logged on!
Step1 (Cont.) - Windows

1. Download and install SSH Secure Utilities: http://eits.uga.edu/hardware_and_software/software/
2. You can use PuTTY as an alternative: https://www.putty.org/
Step1 (Cont.) - Windows using SSH Secure Utilities

https://wiki.gacrc.uga.edu/wiki/Connecting#Connecting_to_the_teaching_cluster

Please Note:
Authentication Method needs to be set as
Keyboard Interactive in default <profile Setting>
Step1 (Cont.) - Windows using SSH Secure Utilities

4. Connect to Remote Host
5. Host Name: teach.gacrc.uga.edu
6. User Name: MyID
7. Port Number: 22
8. Enter your UGA MyID password and click OK

Host Name: teach.gacrc.uga.edu
User Name: MyID
Port Number: 22
Step1 (Cont.) - Windows using SSH Secure Utilities

9. Enter “push” and click OK

10. Verify login using Duo
Step1 (Cont.) - Windows using SSH Secure Utilities

11. Click OK

12. Logged on!
Step2 - 5:

zhuofei@teach-sub1 ~$ qlogin ➔ 2. Open an interactive session to compile ens code
zhuofei@tcn26 ~$

zhuofei@tcn26 ~$ cd ens ➔ 3. Change directory to ens folder
zhuofei@tcn26 ens$

zhuofei@tcn26 ens$ ls
ens.def ens.f ens.i01 ens.i02 ens.i12 ens.notes ens.notes~ ens.sh
zhuofei@tcn26 ens$

zhuofei@tcn26 ens$ ml PGI ➔ 4. Load PGI/19.10 compiler module
zhuofei@tcn26 ens$

zhuofei@tcn26 ens$ pgfortran -O ens.f -o ens ➔ 5. Compile binary
zhuofei@tcn26 ens$

zhuofei@tcn26 ens$
Step 6: Make a job submission script using nano

```bash
$ nano ens.sh
```

nano is a small and friendly text editor on Linux.

Ctrl-x to save file and quit from nano.
#!/bin/bash

#SBATCH --job-name=j_ens # Job name (j_ens)
#SBATCH --partition=batch # Partition name (batch)
#SBATCH --ntasks=1 # Run job in single task, by default using 1 CPU core on a single node
#SBATCH --cpus-per-task=1 # CPU core count per task, by default 1 CPU core per task
#SBATCH --mem=5G # Memory per node (5GB); by default using M as unit
#SBATCH --time=96:00:00 # Time limit hrs:min:sec or days-hours:minutes:seconds
#SBATCH --export=NONE # Do not export any user’s explicit environment variables to compute node
#SBATCH --output=log.%j # Standard output log, e.g., log.12345
#SBATCH --mail-user=yourMyID@uga.edu # Where to send mail
#SBATCH --mail-type=END,FAIL # Mail events (BEGIN, END, FAIL, ALL)

cd $SLURM_SUBMIT_DIR # Change directory to job submission directory (Optional!!)
ml PGI/19.10-GCC-8.3.0-2.32 # Load PGI/19.10 compiler module
time ./ens # Run binary ens in this job

More Information: https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_the_teaching_cluster
Step7: Submit a job using `sbatch`

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

```bash
$ sbatch ens.sh
Submitted batch job 3809
```

**Tips:** `ens.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 ens binary
**Step 7: Check job status using squeue --me -l**

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

```
$ squeue --me -l
Wed Aug  8 13:40:02 2018

<table>
<thead>
<tr>
<th>JOBID</th>
<th>PARTITION</th>
<th>NAME</th>
<th>USER</th>
<th>STATE</th>
<th>TIME</th>
<th>TIME_LIMIT</th>
<th>NODES</th>
<th>NODELIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>3809</td>
<td>batch</td>
<td>j_ens</td>
<td>zhuofei</td>
<td>RUNNING</td>
<td>0:50</td>
<td>4:00:00</td>
<td>1</td>
<td>tcn18</td>
</tr>
<tr>
<td>3807</td>
<td>interacti</td>
<td>qlogin</td>
<td>zhuofei</td>
<td>RUNNING</td>
<td>16:13</td>
<td>8:00:00</td>
<td>1</td>
<td>tcn26</td>
</tr>
</tbody>
</table>
```

**TIME:** the elapsed time used by the job, not remaining time, not CPU time.
Step7 (Cont.): Cancel job using scancel

```
$ squeue --me -l
Wed Aug  8 13:40:02 2018
JOBID PARTITION   NAME      USER       STATE   TIME  TIME_LIMI  NODES NODELIST
 3809     batch     j_ens  zhuofei   RUNNING  0:50  4-00:00:00      1  tcn18
 3807 interacti  qlogin  zhuofei   RUNNING  16:13   8:00:00      1  tcn26

$ scancel 3809

$ squeue --me -l
Wed Aug  8 13:40:02 2018
JOBID PARTITION   NAME      USER       STATE   TIME  TIME_LIMI  NODES NODELIST
 3807 interacti  qlogin  zhuofei   RUNNING  16:13   8:00:00      1  tcn26
```
Step7 (Cont.): Check job details using scontrol show job

```
$ scontrol show job 174
JobId=3810  JobName=j_ens
  UserId=zhuofei(1772) GroupId=gacrc-instruction(21004) MCS_label=N/A
  JobState=RUNNING  Reason=None  Dependency=(null)
  Requeue=1  Restarts=0  BatchFlag=1  Reboot=0  ExitCode=0:0
  RunTime=00:10:28  TimeLimit=4:00:00:00  TimeMin=N/A
  SubmitTime=2020-10-12T08:32:23  EligibleTime=2020-10-12T08:32:23
  StartTime=2020-10-12T08:32:23  EndTime=2020-10-16T08:32:23  Deadline=N/A
  ...  
  Partition=batch  AllocNode:Sid=tcn26:18042
  NodeList=tcn18
  NumNodes=1  NumCPUs=1  NumTasks=1  CPUs/Task=1  ReqB:S:C:T=0:0:*:*:*
  
  Command=/home/zhuofei/ens/ens.sh
  WorkDir=/home/zhuofei/ens
  StdErr=/home/zhuofei/ens/log.3810
  StdOut=/home/zhuofei/ens/log.3810
```
Step7 (Cont.): Check node info using sinfo

```bash
$ sinfo
PARTITION    AVAIL   TIMELIMIT    NODES  STATE    NODELIST
batch*       up 7-00:00:00 1  drain  tcn1
batch*       up 7-00:00:00 1  mix     tcn18
batch*       up 7-00:00:00 22 idle    tcn[2-16,19-25]
batch*       up 7-00:00:00 1  down    tcn17
interactive  up 7-00:00:00 1  mix     tcn26
interactive  up 7-00:00:00 4  idle    tcn[27-30]
gpu          up 7-00:00:00 1  idle    tcgn1
highmem      up 7-00:00:00 2  idle    tchmn[1-2]
```

idle = no cores in use; mix = some cores are still free; alloc = all cores are allocated
Transfer Data: from local computer to workDir - Mac/Linux

1. Connect to Transfer node (MyID@txfer.gacrc.uga.edu) in Terminal on local computer
2. Type scp command: scp (-r) [Source] [Target]
3. Once you input MyID password, scp command will send “push” to your Duo Enrolled mobile device for verification

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/ .
```
Please Note:
Authentication Method needs to be set as Keyboard Interactive in default <profile Setting>
Transfer Data: Windows using SSH Secure Utilities

4. Host Name: txfer.gacrc.uga.edu
   User Name: MyID
   Port Number: 22

5. Connect to Remote Host
6. Click Connect

7. Enter your UGA MyID password and click OK

Steps 9 - 11 are the same as listed on page 13 - 14!
Transfer Data: Windows using SSH Secure Utilities

13. Click yellow button

12. Logged on!

14. Change local and remote paths
Transfer Data: Windows using SSH Secure Utilities

15. Drag data between local computer and remote cluster
Transfer data on cluster to workDir

• Log on to Transfer node (MyID@txfer.gacrc.uga.edu)
  ✔ Mac/Linux: ssh MyID@txfer.gacrc.uga.edu (page 8-9)
  ✔ Windows: use SSH Secure Client app (page 14-16)

• Directories you can access on txfer:
  1. /home/MyID (Landing home)
  2. /work/CourseID/MyID
  3. /work/CourseID/instructor_data

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

  `mv /work/binf8940/MyID/datafile /home/MyID/workDir`
GACRC Wiki [http://wiki.gacrc.uga.edu](http://wiki.gacrc.uga.edu)

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)

Software: [https://wiki.gacrc.uga.edu/wiki/Software](https://wiki.gacrc.uga.edu/wiki/Software)

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)

Code Compilation: [https://wiki.gacrc.uga.edu/wiki/Code_Compilation_on_the_teaching_cluster](https://wiki.gacrc.uga.edu/wiki/Code_Compilation_on_the_teaching_cluster)

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

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.
General Support

If you do not have a myUGA, 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 granting 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.georgia.edu](https://wiki.georgia.edu)

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

Privacy | Accessibility | Website Feedback

[https://uga.teamdynamic.com/TCClient/Requests/ServiceCatalogSearch](https://uga.teamdynamic.com/TCClient/Requests/ServiceCatalogSearch)
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/