



Georgia Advanced Computing
Resource Center
UNIVERSITY OF GEORGIA

GACRC User Quick Guide: Migrating from Zcluster to Sapelo

The GACRC Staff
Version 1.0

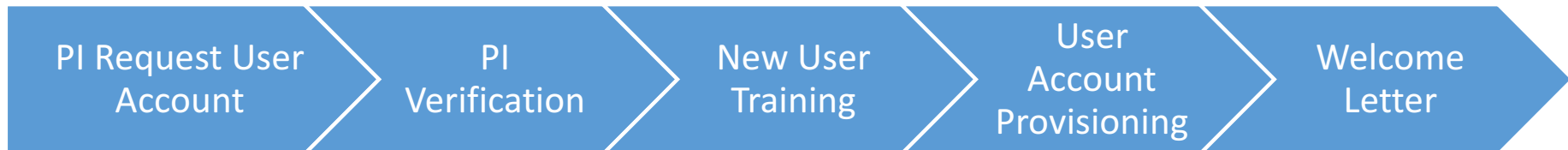
Discussion Points

- I. Request Sapelo User Account
- II. Systems
- III. Transfer Files
- IV. Configure Software Environment
- V. Job Submission Workflow
- VI. Job Submission Script
- VII. Appendices

Request Sapelo User Account

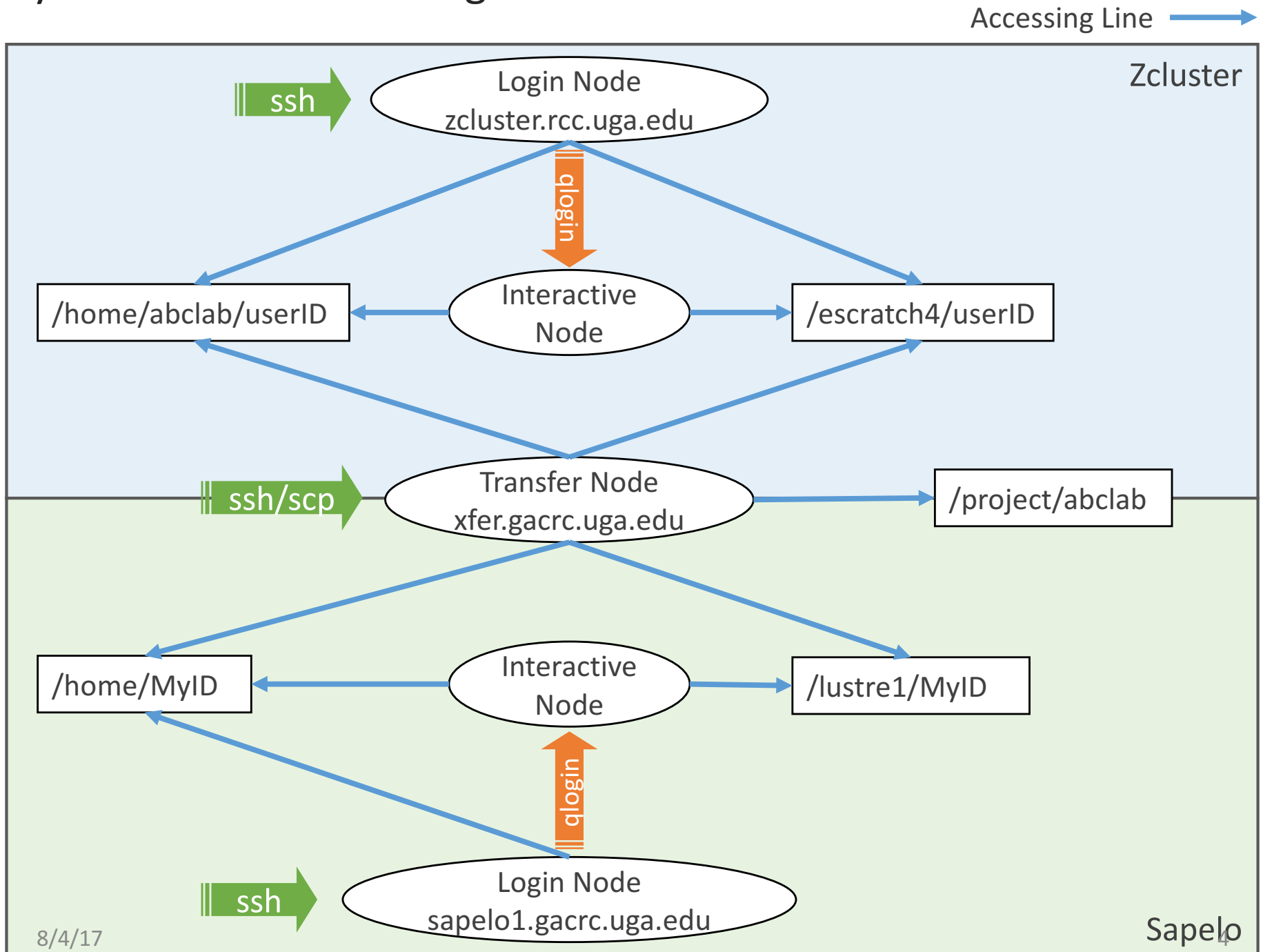
Sapelo cluster user account: MyID@sapelo1.gacrc.uga.edu

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



1. The UGA PI uses the GACRC online form <http://help.gacrc.uga.edu/userAcct.php> to request a user account for a group member.
2. Once we received the request, we will verify it with the PI.
3. After verification by the PI, the new user will be required to attend a training session.
4. After the user attended training, we will provision a Sapelo account for the user.
5. A welcome letter is sent to the user once user account is ready.

Systems – Overview Diagram



Systems – Overview

Log on to Zcluster Login Node:	<code>'ssh userID@zcluster.rcc.uga.edu'</code> (using zcluster password)
Open Zcluster Interactive Node:	type <code>'qlogin'</code> command from login node
Zcluster Home Directory:	/home/abclab/userID
Zcluster Global Scratch:	/escratch4/userID (4TB quota)
Create Working Directory on Zcluster Global Scratch:	type <code>'make_escratch'</code> command from login node
Log on to Sapelo Login Node:	<code>'ssh MyID@sapelo1.gacrc.uga.edu'</code> (using UGA MyID password)
Open Sapelo Interactive Node:	type <code>'qlogin'</code> command from login node
Sapelo Home Directory:	/home/MyID (100GB quota)
Sapelo Global Scratch:	/lustre1/MyID
Create Working Directory on Sapelo Global Scratch:	use <code>'mkdir'</code> command in /lustre1/MyID from interactive node
Log on to GACRC Transfer Node:	<code>'ssh MyID@xfer.gacrc.uga.edu'</code> (using UGA MyID password)
Group Project Space:	/project/abclab (“abc” are initials of the PI of the group)

Systems – Storage Environment

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

Name	Role	Purpose	Zcluster	Sapelo
Home	ssh landing spot	static data, e.g., scripts, software	/home/abclab/userID	/home/MyID (100GB quota)
Global Scratch	Job working space	Current Job data being used by running jobs	/escratch4/userID (4TB quota)	/lustre1/MyID (no quota – limited by capacity)
Project	Active data storage	Non-current active data for future analysis	/project/abclab (1TB initial quota)	

Systems – Computing Resources Overview

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

Zcluster	150x 8-core 16GB RAM nodes with Intel Xeon processors (total 1200 cores) 45x 12-core 48GB RAM nodes with Intel Xeon processors (total 540 cores)
	4x 8-core 192GB RAM nodes with Intel Xeon processors (total 32 cores) 10x 12-core 256GB RAM nodes with Intel Xeon processors (total 120 cores) 2x 32-core 512GB RAM nodes with Intel Xeon processors (total 64 cores)
	6x 32-core 64GB RAM nodes with AMD Opteron processors (total 192 cores)
	4x 12-core 96GB RAM nodes with Intel Xeon processors, 8 NVIDIA K20Xm GPU cards each
Sapelo	112x 48-core 128GB RAM nodes with AMD Opteron processors (total 5376 cores) 4x 48-core 256GB RAM nodes with AMD Opteron processors (total 192 cores)
	6x 48-core 512GB RAM nodes with AMD Opteron processors (total 288 cores) 1x 48-core 1TB RAM node with AMD Opteron processors (total 48 cores) 3x 28-core 1TB RAM node with Intel Xeon processors (total 84 cores)
	2x 16-core 128GB RAM nodes with Intel Xeon processors, 8 NVIDIA K40m GPU cards each
	85+ buy-in nodes with variable configurations

Transfer of Files

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

All Sapelo and zcluster users, as well as all GACRC PIs, can access the GACRC transfer nodes (xfer.gacrc.uga.edu) and using UGA **MyID** and **MyID password** to authenticate.

Note that a user's home directory on transfer node is the same as the user's Sapelo home directory, which is not the same as the user's Zcluster home directory.

Transfer nodes can access the following file systems using their **full paths**:

Sapelo Home Directory:	<code>cd /home/MyID</code>
------------------------	----------------------------

Zcluster Home Directory:	<code>cd /panfs/pstor.storage/home/abclab/userID</code>
--------------------------	---

Zcluster Global Scratch:	<code>cd /escratch4/userID</code>
--------------------------	-----------------------------------

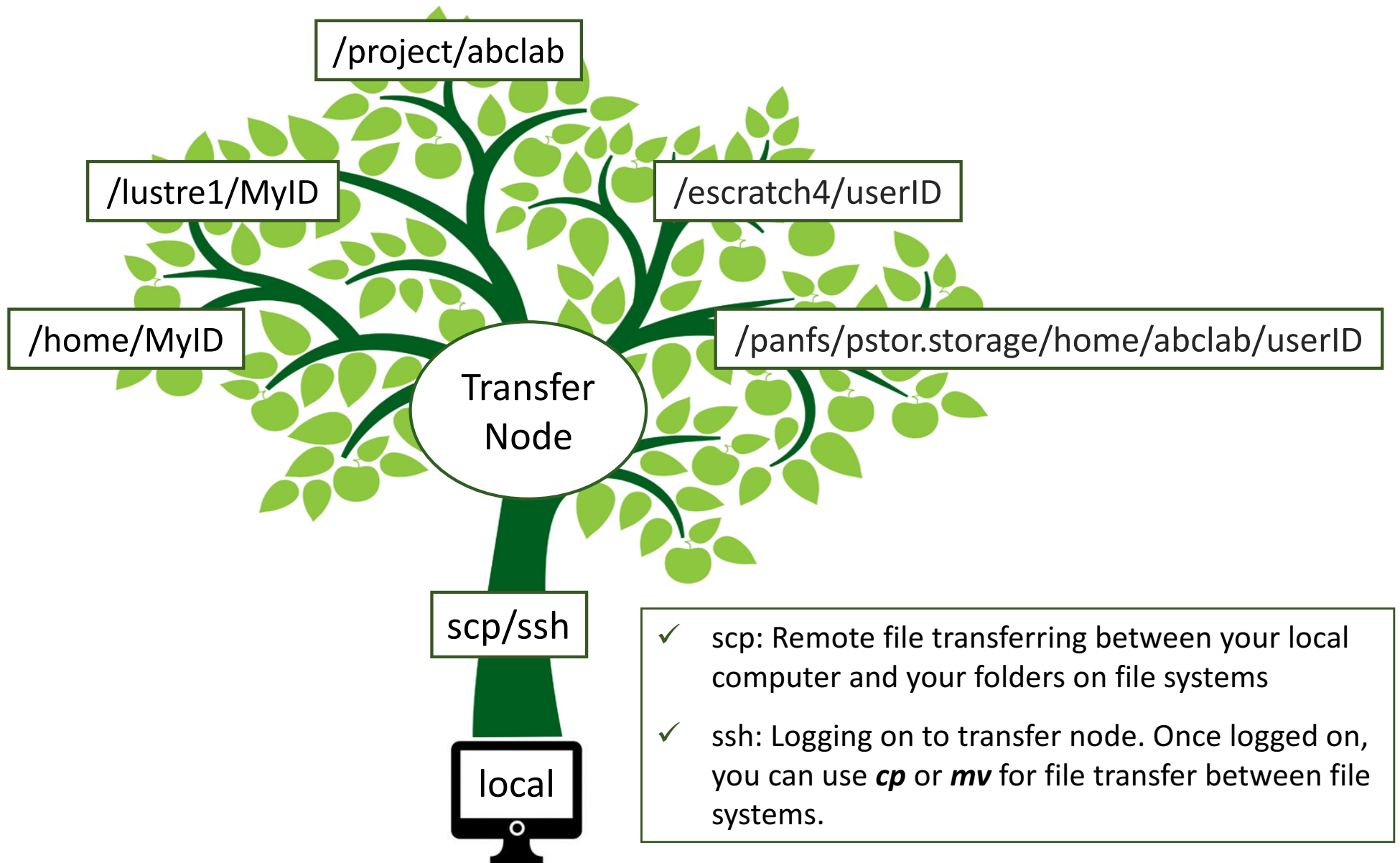
Sapelo Global Scratch:	<code>cd /lustre1/MyID</code>
------------------------	-------------------------------

Group Project Folder	<code>cd /project/abclab</code>
----------------------	---------------------------------

Detailed instructions can be found at the link above

File Transferring Tree

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



Configure Software Environment

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

Zcluster

User “manually” exports various paths for software packages (in a job submission script or on the command line of an interactive session), for example:

```
export PATH=.....:$PATH
```

Sapelo

User issues module commands to configure various paths for software packages
(in a job submission script or on the command line of an interactive session):

1. **module avail** : show all modules available on cluster
2. **module load moduleName/version** : load the needed module “moduleName” with a version
3. **module list** : list all modules you already loaded

For example:

```
$ module load python/2.7.8
```

```
$ module load R/3.2.3
```

```
$ module list
```

Currently Loaded Modules:

```
1) moab/8.1.1 2) StdEnv 3) python/2.7.8 4) R/3.2.3
```

Job Submission Workflow (general)

Zcluster	Sapelo
<ol style="list-style-type: none">1. Log on to zcluster login node: <i>ssh userID@zcluster.rcc.uga.edu</i>2. Create a working dir on global scratch: <i>make_escalch</i>3. Change to working dir on global scratch: <i>cd /escalch4/userID/userID_mmm_dd</i> (mmm, dd: date abbreviations)4. Create a zcluster job submission script: <i>nano ./sub.sh</i>5. Submit job: <i>qsub -q rcc-30d ./sub.sh</i>	<ol style="list-style-type: none">1. Log on to Sapelo login node: <i>ssh MyID@sapelo1.gacrc.uga.edu</i>2. Go to Sapelo Interactive node: <i>qlogin</i>3. Change to Sapelo global scratch: <i>cd /lustre1/MyID</i>4. Create a working dir on global scratch: <i>mkdir ./workDir</i>5. Change to working dir on global scratch: <i>cd ./workDir</i>6. Create a Sapelo job submission script: <i>nano ./sub.sh</i>7. Submit job: <i>qsub ./sub.sh</i>

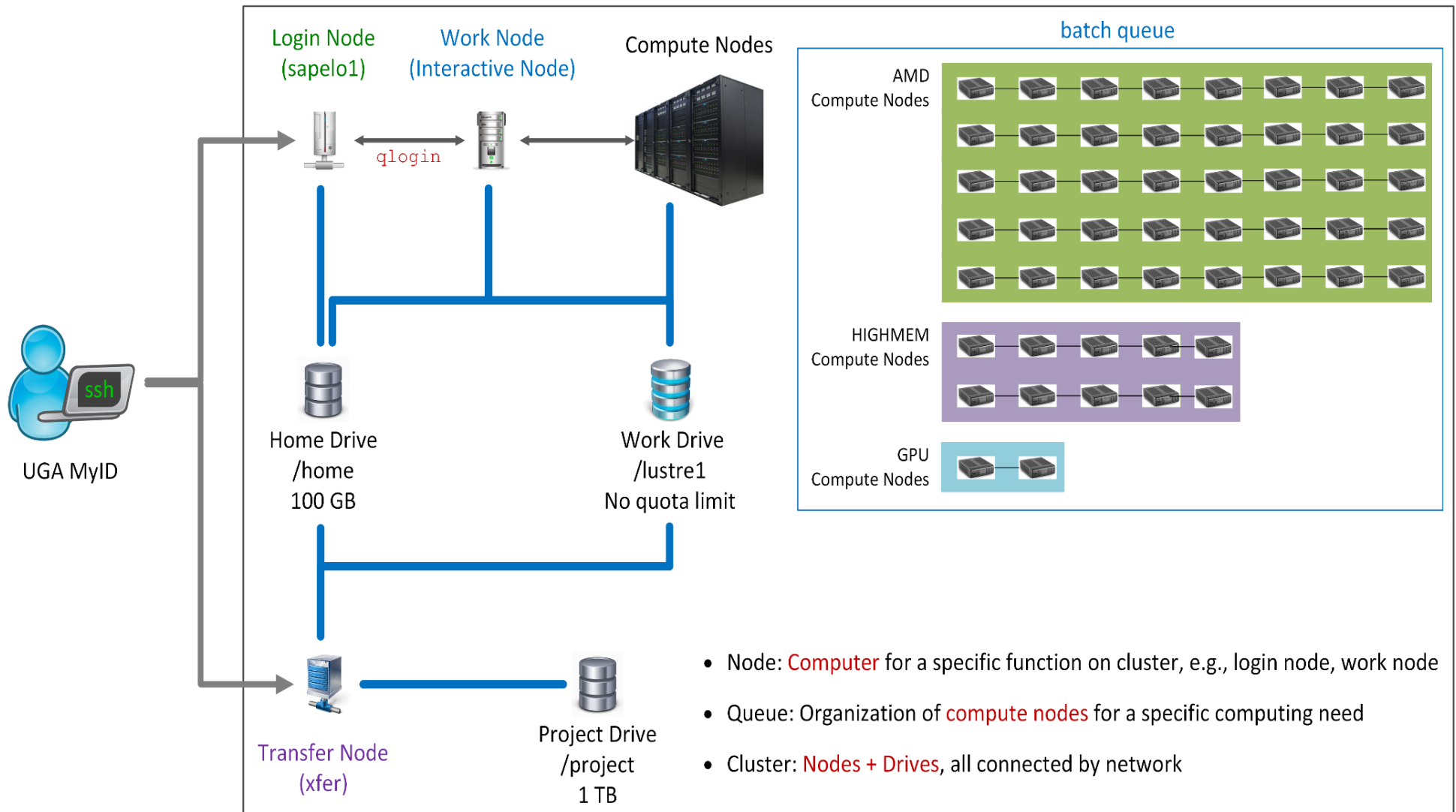
Job Submission Script (NCBI Blast as example)

sub.sh on Zcluster	sub.sh on Sapelo
<pre>#!/bin/bash cd `pwd` export PATH=/usr/local/ncbiblast/latest/bin:\$PATH blastall -p blastn -d /db/ncbiblast/latest/nt \ -a 4 -i QUERY \ -o out.QUERY</pre>	<pre>#PBS -S /bin/bash #PBS -q batch #PBS -N testBlast #PBS -l nodes=1:ppn=4:AMD #PBS -l mem=20gb #PBS -l walltime=480:00:00 cd \$PBS_O_WORKDIR module load ncbiblast+/2.2.26 blastall -p blastn -d /db/ncbiblast/nrte/latest/nt \ -a 4 -i QUERY \ -o out.QUERY > outputfile 2>&1</pre>
<pre>qsub -q rcc-30d -pe thread 4 ./sub.sh</pre>	<pre>qsub ./sub.sh</pre>

Appendix 1:



Sapelo Cluster Diagram



- Node: **Computer** for a specific function on cluster, e.g., login node, work node
- Queue: Organization of **compute nodes** for a specific computing need
- Cluster: **Nodes + Drives**, all connected by network

Appendix 2: Common Command Summary

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

Tasks	Zcluster	Sapelo
Common file/dir operations	ls, cd, pwd, file, cat, less, more, cp mv, mkdir, rm, rmdir, dos2unix, mac2unix, etc.	
Transfer remote files	scp, sftp or SSH Secure Client File Transfer, FileZilla, WinSCP	
Open interactive session	qlogin	
Create working directory on global scratch	make_escratch	mkdir subdirectoryName
Edit script files	nano sub.sh or vim sub.sh	
Submit a batch job	qsub -q queueName -l ... -pe ... sub.sh	qsub sub.sh
check job status	qstat	
Check detailed job info	qsj jobID, qstat -j jobID, qacct -j jobID	qstat -f jobID, showjobs jobID
Cancel a job from queue	qdel jobID, qdel -u userID	qdel jobID

Appendix 3: Current process for a UGA PI to request (re)activation or creation of a MyID for an external collaborator

To keep a user's MyID active **after** he/she leaves UGA, or to create a MyID for an external research collaborator:

1. The PI accesses the EITS helpdesk portal at <https://uga.teamdynamix.com/TDClient/Home/>
2. Click on "MyID Account Request" on the right side of the page in "Popular Services" section.
3. Use MyID credentials to log on the service page.
4. In the MyID Request From, enter his/her information and select "GACRC Remote Collaborator" for the Account Type field.
5. After verification by the GACRC Director, EITS will generate collaborator's MyID
6. If collaborator is visiting UGA campus, MyID can be obtained through OVPR's Visiting Researcher & Scholar program – <https://research.uga.edu/visiting/>

Appendix 4: Useful Links



Georgia Advanced Computing
Resource Center
UNIVERSITY OF GEORGIA

- **GACRC Website:** <http://gacrc.uga.edu>
- **GACRC Wiki:** https://wiki.gacrc.uga.edu/wiki/Main_Page
- **Systems:** <https://wiki.gacrc.uga.edu/wiki/Systems>
- **How to Transfer Files:** https://wiki.gacrc.uga.edu/wiki/Transferring_Files
- **Current Software List:** <https://wiki.gacrc.uga.edu/wiki/Software>
- **How to Submit a Job:** https://wiki.gacrc.uga.edu/wiki/Running_Jobs_on_Sapelo
- **How to Get Help:** <http://help.gacrc.uga.edu/>
- **GACRC Training Opportunities:** <https://wiki.gacrc.uga.edu/wiki/Training>
- **How to register a new GACRC New Lab:** <http://help.gacrc.uga.edu/labAcct.php>
- **How to register a new GACRC User:** <http://help.gacrc.uga.edu/userAcct.php>