Introduction to GACRC Sapelo2 Cluster

CSP Lunch Seminar

Georgia Advanced Computing Resource Center (GACRC)

EITS/University of Georgia

Zhuofei Hou zhuofei@uga.edu

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://uga.teamdynamix.com/TDClient/Requests/ServiceCatalog?CategoryID=11593

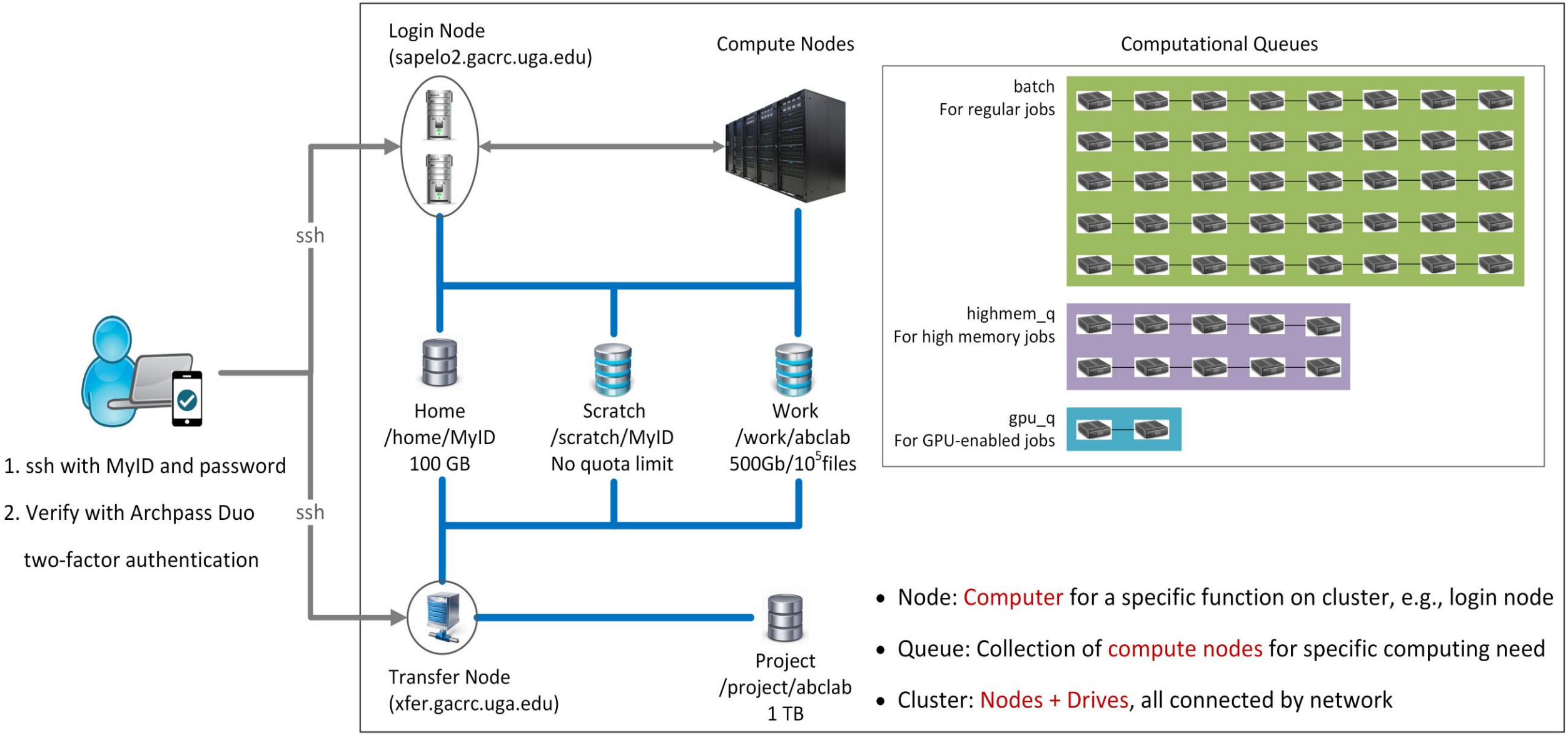
Web Site: http://gacrc.uga.edu

Outline

- What is Sapelo2 Cluster
- Work on Sapelo2 Cluster
- Obtain Sapelo2 User Account
- GACRC Links

What is Sapelo2 Cluster

1. Diagram and Overview
2. Storage Environment
3. Computing Resources
4. Software Environment



Please 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/

Overview https://wiki.gacrc.uga.edu/wiki/Systems#Sapelo2

- Two Nodes:
  1. Login node for batch job workflow: MyID@sapelo2.gacrc.uga.edu
  2. Transfer node for data transferring: MyID@xfer.gacrc.uga.edu
- Five Directories:
  1. Home: Login landing spot; 100GB quota; Backed-up
  2. Scratch: High-speed storage for temp files needed for current jobs; NO quota; NOT backed-up
  3. Work: High-speed storage for input files needed for repeated jobs; per group quota of 500GB and max 100,000 single files; NOT backed-up
  4. Project: Temporary data parking; per group quota of 1TB; Backed-up (ONLY accessible from Transfer node!)
  5. Local Scratch: Local storage on each individual compute node; 200GB quota; NOT backed-up
- Four Computational Queues: batch, highmem_q, gpu_q, groupBuyin_q

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

Directory	Name	Quota	Accessible from	Intended Use	Backed-up	Important Notes
/home/MyID	Home	100GB	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		Temporary files needed for current running jobs	No	Clean up when your job finished! Subject to "30-day purge" policy
/work/abclab	Work	500GB 10^5 files		Input files needed for repeated jobs	No	Clean up any old data! Group sharing is possible
/project/abclab	Project	1TB (initial)	Transfer	Temporary data parking	Yes	Group sharing is possible
/lscratch	Local Scratch	200GB	Compute	Jobs with heavy disk I/O operations	No	Clean up when job exits from node! Data are persistent

More about 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 your old files is to copy them to /project area, using the fpsync utility on xfer.gacrc.uga.edu
- If you want to first create a tar archive of your /scratch area, DO NOT compress the archive when creating the archive

Queue	Total Nodes	RAM(GB) /Node	Max Mem(GB) /Single-node job	Cores /Node	Processor Type	GPU Cards /Node	InfiniBand
batch	42	192	184	32	Intel Xeon Skylake	N/A	Yes
	32	64	58	28	Intel Xeon Broadwell		
	106	128	120	48	AMD Opteron		
	16			32	AMD EPYC		
highmem_q	5	1024	990	28|48	Intel Xeon Broadwell (4) AMD Opteron (1)		
	15	512	502	32|48	Intel Xeon Nehalem (1) AMD EPYC (8)|Opteron (6)		
gpu_q	4	192	184	32	Intel Xeon Skylake	1 NVDIA P100	
	2	128	120	16	Intel Xeon	8 NVIDIA K40m	
	4	96	90	12		7 NVIDIA K20Xm	
groupBuyin_q	variable						

Software Environment https://wiki.gacrc.uga.edu/wiki/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/3.6.4-foss-2018a
3. Software names are case-sensitive!
   - `module avail` : List all available software modules installed on cluster
   - `module load` *`moduleName`* : Load a module into your working environment
   - `module list` : List modules currently loaded
   - `module unload` *`moduleName`* : Remove a module from working environment
   - `ml spider` *`pattern`* : Search module names matching a pattern (case-insensitive)

Work on Sapelo2 Cluster

1. Job Submission Workflow
2. How to Know Job Details
3. How to Know Node Details
4. qlogin Commands: Open Interactive Node for Running Interactive Tasks
5. Code Compilation

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 SSH File Transfer 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 : `qsub ./sub.sh`
8. Check job status : `qstat_me` or Cancel a job : `qdel JobID`

How to Know Job Details

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

Option 1: **`qstat -f JobID`** for ***running jobs*** or ***finished jobs in 24 hours***

Option 2: Email notification from ***finished jobs (completed, canceled, or crashed)***,

if using:

```
#PBS -M MyID@uga.edu
#PBS -m ae
```

Option 1: qstat -f JobID (running jobs or finished jobs in 24 hour)

```
$ qstat -f 12222
Job Id: 12222.sapelo2
    Job_Name = testBlast
    Job_Owner = zhuofei@10.56.200.51
    resources_used.cput = 00:00:00
    resources_used.vmem = 316864kb
    resources_used.walltime = 00:15:01
    resources_used.mem = 26780kb
    resources_used.energy_used = 0
    job_state = C
    queue = batch
    .
    Error_Path = sapelo2-sub2.ecompute:/lustre1/zhuofei/examples/testBlast.e12222
    exec_host = n236/0-3
    Output_Path = sapelo2-sub2.ecompute:/lustre1/zhuofei/examples/testBlast.o12222
    .
    Resource_List.nodes = 1:ppn=4:Intel
    Resource_List.mem = 20gb
    Resource_List.walltime = 02:00:00
    Resource_List.nodect = 1
    .
    Variable_List = PBS_O_QUEUE=batch,PBS_O_HOME=/home/zhuofei,……
                    PBS_O_WORKDIR=/lustre1/zhuofei/workDir,
```

Option 2: Email notification from finished jobs

```
BS Job Id: 12332.sapelo2        Job Name: bowtie2_test
Queue: batch
Exechost: n232/0

Message: Execution terminated

Details:
Exit_status=0
resources_used.cput=00:09:26
resources_used.vmem=755024kb
resources_used.walltime=00:09:51
resources_used.mem=1468676kb
resources_used.energy_used=0
.
Short reason:
Execution terminated
```

```
PBS Job Id: 12331.sapelo2        Job Name: bowtie2_test
Queue: batch
Exechost: n235/0

Message: Execution terminated

Details:
Exit_status=271
resources_used.cput=00:02:58
resources_used.vmem=755024kb
resources_used.walltime=00:03:24
resources_used.mem=420712kb
resources_used.energy_used=0
.
Short reason:
Execution terminated
```

Sender: dispatch_root

How to Know Node Details

Option 1: mdiag -v –n | grep [pattern] | …

mdiag -v -n | grep batch | grep AMD
mdiag -v -n | grep batch | grep Intel
mdiag -v -n | grep highmem_q
mdiag -v -n | grep grpBuyin_q

Option 2: from login node, you can ssh to a compute node and run a command there!

ssh n72 'lscpu‘
ssh n222 'free -g‘
ssh n237 "ps aux | grep '^MyID'"

qlogin Commands

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

1. Type qlogin commands from Login node to open Interactive node:
   - `qlogin_intel`: Start an interactive session on an Intel node
   - `qlogin_amd`: Start an interactive session on an AMD node
   - `qlogin`: start an interactive job on either type of nodes
2. Type `exit` command to quit and back to Login node

qlogin Commands

Purpose1: Open interactive node for running interactive tasks of R, Python, Bash scripts, etc.

```
zhuofei@sapelo2-sub1 ~$ qlogin
qsub: waiting for job 12426.sapelo2 to start
qsub: job 12426.sapelo2 ready

zhuofei@n204 ~$ module spider R
-----------------------------------------------------------------------------
  R: R/3.4.1-foss-2016b-X11-20160819-GACRC
-----------------------------------------------------------------------------
…
zhuofei@n204 ~$ ml R/3.4.1-foss-2016b-X11-20160819-GACRC
zhuofei@n204 ~$ R
R version 3.4.1 (2017-06-30) -- "Single Candle"
…
[Previously saved workspace restored]
> a<-1 ; b<-7
> a+b
[1] 8
>
```

qlogin Commands

Purpose2 : Open interactive node for compiling/testing source codes of Fortran, C/C++, Python, etc.

```
zhuofei@sapelo2-sub1 ~$ qlogin_intel
qsub: waiting for job 20912.sapelo2 to start
qsub: job 20912.sapelo2 ready

zhuofei@n206 ~$  module spider iomkl
-----------------------------------------
 iomkl:
-----------------------------------------
   Description:
     Intel Cluster Toolchain Compiler Edition provides Intel C/C++
and Fortran compilers, Intel MKL & OpenMPI.

    Versions:
       iomkl/2018a
...
zhuofei@n206 ~$ module load iomkl/2018a
zhuofei@n206 ~$ icc mysource.c -o myexec.x
zhuofei@n206 ~$
```

Code Compilation

https://wiki.gacrc.uga.edu/wiki/Code_Compilation_on_Sapelo2

GCC/8.3.0-2.32	→ GNU 8.3.0-2.32 compiler suite
PGI/18.10-GCC-6.4.0-2.28	→ PGI 18.10 compiler suite
iccifort/2018.1.163-GCC-6.4.0-2.28	→ Intel 18.0.1.163 compiler suite
foss/2016b	→ GCC 5.4.0, OpenMPI 1.10.3, OpenBLAS 0.2.18, FFTW 3.3.4, ScaLAPACK 2.0.2
foss/2018a	→ GCC 6.4.0, OpenMPI 2.1.2, OpenBLAS 0.2.20, FFTW 3.3.7, ScaLAPACK 2.0.2
foss/2018b	→ GCC/7.3.0, OpenMPI 3.1.1, OpenBLAS 0.3.1, FFTW 3.3.8, ScaLAPACK 2.0.2
fosscuda/2018b	→ foss/2018b with CUDA 9.2.88
gmvolf/2016b	→ GCC 5.4.0, MVAPICH2 2.2, OpenBLAS 0.2.18, FFTW 3.3.4, ScaLAPACK 2.0.2
iomkl/2018a	→ Intel 2018.1.163 compiler suite, OpenMPI 2.1.2, MKL 2018.1.163
imvmkl/2018a	→ Intel 2018.1.163 compiler suite, MVAPICH2 2.2, MKL 2018.1.163

Obtain Sapelo2 User Account

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

Sapelo2 user account: MyID@sapelo2.gacrc.uga.edu

Note: An official activated UGA MyID is a MUST to create a user account!

1. The UGA PI uses the GACRC Account Creation form to request a user account for a group member:

   https://uga.teamdynamix.com/TDClient/Requests/ServiceDet?ID=25839
2. Once we receive the request, the new user will be required to attend a training session
3. During the training session, a Sapelo2 user account is created for the user
4. A welcome letter and a training survey will be sent to the user after the training session

GACRC Wiki and Support

Main Page: http://wiki.gacrc.uga.edu

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

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

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

User Account Request: https://wiki.gacrc.uga.edu/wiki/User_Accounts

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

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/