Introduction to High Performance Computing (HPC) Resources at GACRC

Georgia Advanced Computing Resource Center
University of Georgia
Zhuofei Hou, HPC Trainer
zhuofei@uga.edu
Outline

• What is GACRC?

• Concept of High Performance Computing (HPC)

• What is GACRC zcluster?

• What is GACRC New Cluster – Sapelo?
What is GACRC?

Who Are We?

- Georgia Advanced Computing Resource Center
- Collaboration between the Office of Vice President for Research (OVPR) and the Office of the Vice President for Information Technology (OVPIT)
- Guided by a faculty advisory committee (GACRC-AC)

Why Are We Here?

- To provide computing hardware and network infrastructure in support of high-performance computing (HPC) at UGA

Where Are We?

- http://gacrc.uga.edu (Web)
- http://wiki.gacrc.uga.edu (Wiki)
- http://gacrc.uga.edu/help/ (Web Help)
- https://wiki.gacrc.uga.edu/wiki/Getting_Help (Wiki Help)
# GACRC Users September 2015

<table>
<thead>
<tr>
<th>Colleges &amp; Schools</th>
<th>Depts</th>
<th>PIs</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Franklin College of Arts and Sciences</td>
<td>14</td>
<td>117</td>
<td>661</td>
</tr>
<tr>
<td>College of Agricultural &amp; Environmental Sciences</td>
<td>9</td>
<td>29</td>
<td>128</td>
</tr>
<tr>
<td>College of Engineering</td>
<td>1</td>
<td>12</td>
<td>33</td>
</tr>
<tr>
<td>School of Forestry &amp; Natural Resources</td>
<td>1</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td>College of Veterinary Medicine</td>
<td>4</td>
<td>12</td>
<td>29</td>
</tr>
<tr>
<td>College of Public Health</td>
<td>2</td>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>College of Education</td>
<td>2</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Terry College of Business</td>
<td>3</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>School of Ecology</td>
<td>1</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>School of Public and International Affairs</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>College of Pharmacy</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>40</td>
<td>214</td>
<td>970</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Centers &amp; Institutes</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TOTALS:</strong></td>
<td>49</td>
<td>233</td>
<td>1029</td>
</tr>
</tbody>
</table>
## GACRC Users September 2015

<table>
<thead>
<tr>
<th>Centers &amp; Institutes</th>
<th>PIs</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center for Applied Isotope Study</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Center for Computational Quantum Chemistry</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Complex Carbohydrate Research Center</td>
<td>6</td>
<td>28</td>
</tr>
<tr>
<td>Georgia Genomics Facility</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Institute of Bioinformatics</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Savannah River Ecology Laboratory</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Skidaway Institute of Oceanography</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Center for Family Research</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Carl Vinson Institute of Government</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>19</strong></td>
<td><strong>59</strong></td>
</tr>
</tbody>
</table>
Concept of High Performance Computing (HPC)

Serial:

- Serial problem can not be broken
- Discrete instructions executed sequentially
- Only 1 instruction executed at any moment on a single processor

Parallel:

- Problem broken into parallel parts can be solved concurrently
- Instructions executed simultaneously on multiply processors
- Synchronization/communication employed
- Shared-memory multithreaded job or MPI job (Message Passing Interface)
What is GACRC zcluster?

- Cluster Structural Diagram
- General Information
- Computing Resources
- Software Installed
- Submit Jobs
GACRC Linux HPC Zcluster Structural Diagram

Zcluster Connection Speed:
- 10 Gbit Ethernet
- 1 Gbit Ethernet

Aggregator Switch

Storage Switch

Top of Rack Switch

Top of Rack Switch

Login Switch

User

Icebreaker Storage

Panasas Storage

Copy Nodes

Copy Nodes

Compute Nodes on Rack 1-11

Node index: c1-1, c1-2, ...., c11-1, c11-2, ....
CPU: Intel Xeon/AMD Opteron multicore
GPU: NVIDIA Tesla (Fermi/Kepler)

/home
/usr/local
/db

Node index: 1, 2, 3

Login Node

zcluster.rcc.uga.edu

World
zcluster General Information

zcluster is a Linux high performance computing (HPC) cluster:

• Operating System: 64-bit Red Hat Enterprise Linux 5 (RHEL 5)

• User can login to:
  Login node: zcluster.rcc.uga.edu (for login & job submission)
  Copy node: copy.rcc.uga.edu (for data transferring & compression)

• Internodal Communication: 1Gbit network
  compute nodes $\leftrightarrow$ compute nodes
  compute nodes $\leftrightarrow$ storage systems

• Queueing System: Sun Grid Engine (SGE) with qsub, qstat, qdel, etc. commands
# zcluster Computing Resources

<table>
<thead>
<tr>
<th>Queue Type</th>
<th>Queue Name</th>
<th>Nodes</th>
<th>Processor</th>
<th>Cores/Node</th>
<th>RAM(GB)/Node</th>
<th>Cores</th>
<th>NVIDIA GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular</td>
<td>rcc-30d</td>
<td>45</td>
<td>Intel Xeon</td>
<td>12</td>
<td>48</td>
<td>540</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>150</td>
<td>Intel Xeon</td>
<td>8</td>
<td>16</td>
<td>1200</td>
<td></td>
</tr>
<tr>
<td>High Memory</td>
<td>rcc-m128-30d</td>
<td>4</td>
<td>Intel Xeon</td>
<td>8</td>
<td>192</td>
<td>32</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>Intel Xeon</td>
<td>12</td>
<td>256</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rcc-m512-30d</td>
<td>2</td>
<td>Intel Xeon</td>
<td>32</td>
<td>512</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Multi Core</td>
<td>rcc-mc-30d</td>
<td>6</td>
<td>AMD Opteron</td>
<td>32</td>
<td>64</td>
<td>192</td>
<td>N/A</td>
</tr>
<tr>
<td>Interactive</td>
<td>interq</td>
<td>2</td>
<td>AMD Opteron</td>
<td>48</td>
<td>132</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>GPU</td>
<td>rcc-sgpu-30d</td>
<td>2</td>
<td>Intel Xeon</td>
<td>8</td>
<td>48</td>
<td>16</td>
<td>4 Tesla S1070 cards</td>
</tr>
<tr>
<td></td>
<td>rcc-mgpu-30d</td>
<td>2</td>
<td>Intel Xeon</td>
<td>12</td>
<td>48</td>
<td>24</td>
<td>9 Tesla (Fermi) M2070 cards</td>
</tr>
<tr>
<td></td>
<td>rcc-kgpu-30d</td>
<td>4</td>
<td>Intel Xeon</td>
<td>12</td>
<td>96</td>
<td>24</td>
<td>32 Tesla (Kepler) K20Xm cards</td>
</tr>
</tbody>
</table>

**Total peak performance:** 23 Tflops
Layer 1: Node
- A standalone “computer in a box”
- Multiple processors, e.g. 2, sharing memory
- Local disk storage, network interface, etc.
- Networked into a cluster

Layer 2: Processor
- A single computing component
- Multicore processor, e.g. 4 cores

Layer 3: Core
- Central processing unit (CPU) reading and executing instructions independently
- Each core is assigned to a software thread
Software Installed on zcluster

- Perl, Python, Java, awk, sed, C/C++ and Fortran compilers
- Matlab, Maple, R
- Many Bioinformatics applications: NCBI Blast+, Velvet, Trinity, TopHat, MrBayes, SoapDeNovo, Samtools, RaxML, etc.
- RCCBatchBlast (RCCBatchBlastPlus) to distribute NCBI Blast (NCBI Blast+) searches to multiple nodes.
- Many Bioinformatics Databases: NCBI Blast, Pfam, uniprot, etc.
- For a complete list of applications installed: https://wiki.gacrc.uga.edu/wiki/Software
Submit Jobs on zcluster

• To submit a batch job, you need:
  ➢ **Software** installed
  ➢ **Job submission script** to run the software,
    ✓ Specifying working directory
    ✓ Exporting environment variables, e.g.,
      `OMP_NUM_THREADS` (OpenMP threads number)
      `LD_LIBRARY_PATH` (searching paths for shared libraries)

• Job queueing commands:
  ➢ `qsub` with specifying **queue name**
  ➢ `qstat`, `qdel`
  ➢ `qacct`, `qsj`, etc.
Submit Jobs on zcluster

- **Step 1**: Create a job submission script `sub.sh` running *Samtools*:

  ```bash
  #!/bin/bash
  cd ${HOME}/testdir
  time /usr/local/samtools/latest/samtools <command> [options]
  ```

  - Linux shell (`bash`)
  - Specify and enter (`cd`) the working directory (`${HOME}/testdir`)
  - Run `samtools` with `time` command to measure amount of time it takes to run the application

- **Step 2**: Submit it to the queue:

  `$ qsub -q rcc-30d sub.sh`

  - Submit a job
  - Your job submission script
  - to the queue rcc-30d with **16GB** RAM/Node
What is GACRC Sapelo

- Cluster Structural Diagram
- General Information
- Computing Resources
- Software installed
- Submit Jobs
The New GACRC Linux HPC Cluster Structural Diagram

Network Connection Speed:
- InfiniBand
- 10 Gbit Ethernet
- 1 Gbit Ethernet

User

World

Aggregator Switch

Storage Switch

Top of Rack Switch

Top of Rack Switch

Login Switch

InfiniBand Switch

User

Icebreaker Storage
Node index: sn3
/project

Icebreaker Storage
Node index: sn[0-2]
/home
/usr/local
/db

Seagate Storage
Lustre Parallel File System
/lustre1

Transfer Node
xfer2.gacrcc.uga.edu
Build Node
build1.gacrcc.uga.edu
AMO Opteron 16-core processor, 64 GB RAM

Compute Nodes
Node index: n0, n1, n2, .... n165, n166, n167
CPU: AMD Opteron / Intel Xeon multicore processors
GPU: NVIDIA Tesla (Kepler) K40m GPU cards
/lscratch

Login Node (VMs)
sapelo1.gacrcc.uga.edu
Intel Xeon processor
Sapelo General Information

Sapelo is a Linux high performance computing (HPC) cluster:

- Operating System: 64-bit CentOS Linux 6.5
- User can login to:
  - Login node: sapelo1.gacrc.uga.edu (for login & job submission)
  - Transfer mode: xfer2.gacrc.uga.edu (for data transferring & compression)
  - Build node: build1.gacrc.uga.edu (for code compilation)
- Internodal communication: InfiniBand network
  - compute nodes ⇔ compute nodes
  - compute nodes ⇔ storage systems, e.g., /home and /scratch
- Queueing System: Torque + Moab with qsub, qstat, qdel, etc. commands
# Sapelo Computing Resources

## Peak Performance per Node:

- 500 Gflops/Node

## Home directory:

- 100 GB

## Scratch directory on /lustre1:

- No quota limit, auto-moved to /project, if no modification in 30 days!

<table>
<thead>
<tr>
<th>Queue</th>
<th>Node Type</th>
<th>Total Nodes</th>
<th>Processor</th>
<th>Cores / Node</th>
<th>RAM (GB) / Node</th>
<th>GPU</th>
<th>GPU Cards / Node</th>
<th>InfiniBand</th>
</tr>
</thead>
<tbody>
<tr>
<td>batch</td>
<td>AMD</td>
<td>120</td>
<td>AMD Opteron</td>
<td>48</td>
<td>128</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>HIGHMEM</td>
<td>3</td>
<td>AMD Opteron</td>
<td>48</td>
<td>512 (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1024 (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPU</td>
<td></td>
<td>2</td>
<td>Intel Xeon</td>
<td>16</td>
<td>128</td>
<td>NVIDIA K40m</td>
<td>8</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Node 167

Layer 1: Node
- A standalone “computer in a box”
- Multiple processors, e.g. 4, sharing memory
- Local disk storage, network interface, etc.
- Networked into a cluster

Layer 2: Processor
- A single computing component
- Multicore processor, e.g. 12 cores

Layer 3: Core
- Central processing unit (CPU) reading and executing instructions independently
- Each core is assigned to a software thread

48 cores and 128 GB RAM per node
Software Installed on Sapelo

• Sapelo uses **environment modules** to define paths for software

• Current number of modules installed is ~90 and expanding daily!

• `module avail` ➔ List all modules available on Sapelo

• `module list` ➔ List modules currently being loaded for use

• `module load` ➔ Load modules needed

• `module unload` ➔ Unload modules not needed
Currently Loaded Modules:
  1) StdEnv   2) moab/7.2.10

Currently Loaded Modules:
  1) StdEnv   2) moab/7.2.10   3) python/2.7.8

Connection to sailo1.gacrc.uga.edu closed.

The following have been reloaded with a version change:
  1) moab/8.1.1 => moab/7.2.10
Submit Batch Jobs on Sapelo

• To submit a batch job, you need:
  - Software loaded. If not, used `module load`
  - Job submission script to run the software, specifying working directory and computing resources:
    - Number of nodes and cores
    - Amount of memory
    - Type of nodes
    - Maximum wallclock time, etc.

• Job queueing commands:
  - `qsub`, `qstat`, `qdel`
  - `showq`, `checkjob`, etc.
Submit Batch Jobs on Sapelo

• How to submit a job? **Easy!**

```
[zhuofei@75-104 MPIs]$ qsub sub.sh
```

qsub is to submit a job

sub.sh is your job submission script specifying:

- Number of nodes and cores
- Amount of memory
- Type of nodes
- Maximum wallclock time, etc.

• How to make a job submission script? **Next Page!**
Submit Batch Jobs on Sapelo

• Example: Serial job submission script sub.sh running NCBI Blast +

```bash
#PBS -S /bin/bash
#PBS -q batch
#PBS -N testBlast
#PBS -l nodes=1:ppn=1:AMD
#PBS -l mem=20gb
#PBS -l walltime=48:00:00

cd $PBS_O_WORKDIR
module load ncbiblast+/2.2.29
time blastn [options] > outputfile
```

- Linux shell (bash)
- Queue name (batch)
- Name of the job (testBlast)
- Number of nodes (1), number of cores/node (1), node type (AMD)
- Maximum amount of physical memory (20 GB) used by the job
- Maximum wall clock time (48 hours) for the job, default 6 minutes

- Use the directory from which the job is submitted as the working directory
- Load the module of ncbiblast+, version 2.2.29
- Run blastn with ‘time’ command to measure the amount of time it takes to run the application
Where to Find Useful Information?

- GACRC Web: http://gacrc.uga.edu/
- GACRC Wiki: https://wiki.gacrc.uga.edu/wiki/Main_Page
- GACRC Help: http://gacrc.uga.edu/help/
- GACRC Training: https://wiki.gacrc.uga.edu/wiki/Training
- GACRC User Account: https://wiki.gacrc.uga.edu/wiki/User_Accounts
- GACRC Software: https://wiki.gacrc.uga.edu/wiki/Software
Thank You!