

Overview of GACRC Resources

Georgia Advanced Computing Resource Center (GACRC)

Enterprise Information Technology Services (EITS)

The University of Georgia

Outline

- What is GACRC?
- Overview of Sapelo2 hardware
- Overview of Sapelo2 software
- Transferring Files using Globus
- Open OnDemand
- Consulting and Support
- Documentation

What is GACRC?

- A high-performance computing (HPC) center at UGA
- A partnership between the Offices of the VP for Information Technology and for Research
- Provide to the UGA research and education community an advanced computing environment:
 - Computing hardware, storage, and networking infrastructure
 - Large collection of math libraries, scientific, and engineering applications
 - Consulting and training services
- Sapelo2 (main cluster, for research use)
- Teaching cluster (for instructional use, student accounts)
- Buy-in program (faculty contribute nodes to Sapelo2)

Wiki: <http://wiki.gacrc.uga.edu>

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

Computing clusters at Boyd Data Center



Accounts and Access

- UGA faculty (PIs) can register a lab and request accounts for group members.

Step 1: Lab registration by PI:

<https://uga.teamdynamix.com/TDClient/2060/Portal/Requests/ServiceDet?ID=25846>

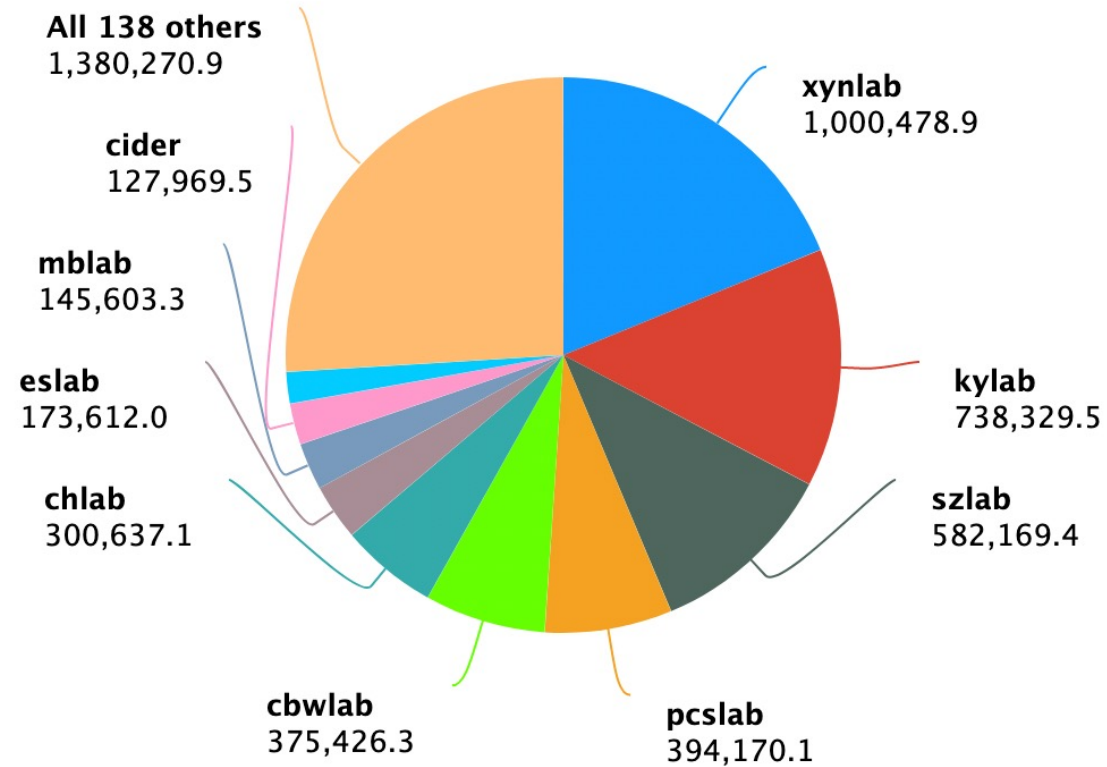
Step 2: PI requests user accounts:

<https://uga.teamdynamix.com/TDClient/2060/Portal/Requests/ServiceDet?ID=25839>

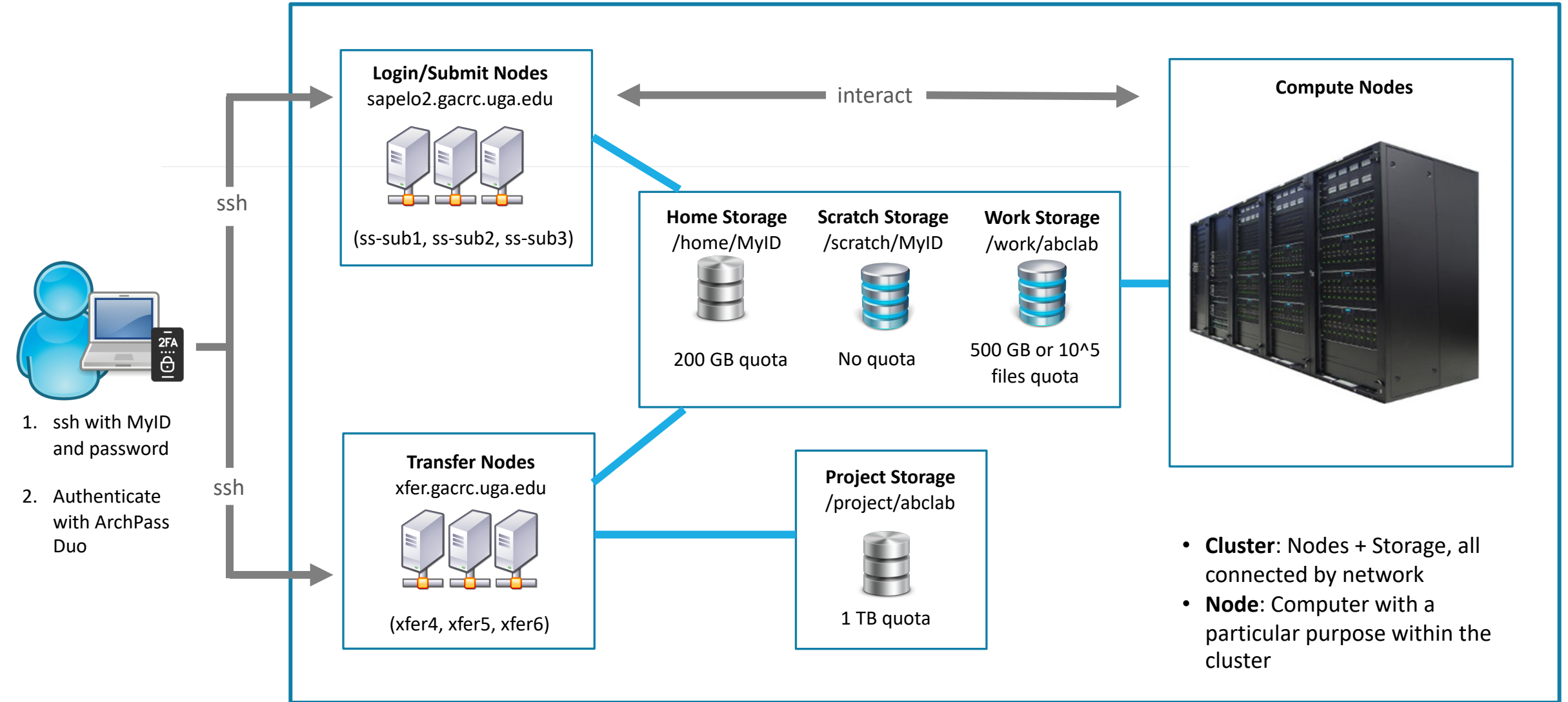
- User accounts provided using UGA MyID.
- Connection to the cluster via SSH on Linux/Mac OS X, PuTTY on Windows.
- Authenticate with UGA MyID password and Archpass Duo.
- From off campus, connect to the UGA VPN first.

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

Sapelo2 usage by group – July 2022



Sapelo2 Cluster Overview



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/

Sapelo2 – Computing Hardware

- Resources available to all users (summary on next slide)
 - Regular memory nodes: 28 to 64 cores/node, 2GB to 6GB/core
 - High memory nodes: 512GB, 1TB or 2TB/node
 - GPU nodes: NVIDIA A100, P100, K40m or K20Xm devices
 - Number of CPU cores: ~16,000
- Buy-in nodes
 - Various configurations, including GPU nodes.
 - Number of CPU cores: ~8,000
- Total number of CPU cores: ~24,000

Processor Type	Cores/Node	RAM/node (GB)	# Nodes	GPU Cards/Node	
AMD EPYC Milan	64	128	6	-	Regular Nodes
AMD EPYC Rome	64	128	106	-	
AMD EPYC Naples	32	128	64	-	
AMD Opteron	48	128	87	-	
Intel Xeon Skylake	32	192	42	-	
Intel Xeon Broadwell	28	64	34	-	
AMD EPYC Rome	32	2000	2	-	High-Memory Nodes
AMD EPYC Naples	64	1000	4	-	
AMD EPYC Naples	32	512	18	-	
AMD Opteron	48	1000	1	-	
AMD Opteron	48	512	6	-	
Intel Xeon Broadwell	28	1000	4	-	
AMD EPYC Milan	64	96	1	4 NVIDIA A100	GPU Nodes
Intel Xeon Skylake	32	192	4	1 NVIDIA P100	
Intel Xeon Ivy Bridge	16	128	2	8 NVIDIA K40m	
Intel Xeon Westmere	12	96	4	7 NVIDIA K20Xm	

Sapelo2 – Storage and Cluster Network

- Storage to support projects that use the computing cluster only.
- ZFS storage: home directory `/home` (200GB/user quota, no purge, backup)
- DDN Lustre storage device:
 - `/scratch` (no quota, 30-day purge, no backup)
 - `/work` (500GB and 500k file/group quota, no purge, no backup)
- Panasas and ZFS JBOD: `/project` (1TB/group, no purge, backup)
- Local hard drive (SSD, NVMe) on compute nodes: `/lscratch` (~890GB)
- Internodal and compute node to Lustre network:
 - AMD Opterons – QDR Infiniband (32 Gb/s bandwidth)
 - All other nodes – EDR Infiniband (100 Gb/s bandwidth)

Sapelo2 – Software

- 64-bit Linux OS (CentOS 7.9).
- Slurm queueing system to submit jobs to the compute nodes (e.g. sbatch, squeue, sacct).
- Over 1100 environment modules, conda environments, Singularity containers.
- Cannot run applications that need Windows OS.
- Cannot run Docker containers, convert to Singularity (Apptainer) containers.
- Very limited number of commercial software, some limited to a few groups that purchased licenses.

Software Environment

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

- Software loaded via **environment modules** (Lmod).
- Compilers (C/C++, Fortran): GCC, Intel , PGI
- MPI Libraries: OpenMPI, Intel MPI, MVAPICH
- Math Libraries: Intel MKL, GSL, OpenBLAS, LAPACK, FFTW, etc.
- GPU Tools: NVIDIA CUDA Toolkit, cuDNN, NVHPC
- High-Level Languages: Python, Perl, R, Julia, Java
- Plotting: gnuplot, xmgrace, matplotlib, ParaView, etc.
- Other Programming/Visualization Tools: MATLAB, Mathematica
- Machine Learning and Deep Learning Tools: scikit-learn, Tensorflow, Keras, PyTorch, etc.
- Applications for Bioinformatics, Chemistry, Physics, Engineering, etc.

Transferring Files

➤ UGA has a subscription for **Globus**, a high-performance data-transfer platform that allows you to perform and/or automate:

- Transfer data from your local machine to GACRC resources.
- Transfer data between file systems on Sapelo2.
- Data transfers between servers in your group or a server and your laptop.
- Sharing data with researchers at UGA and at other institutions.
- Sharing data with the world.

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

https://kaltura.uga.edu/media/t/1_vlprwoc7/176125031

➤ Other file transfer programs are available: scp, FileZilla, WinSCP

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

Open OnDemand

- Alternative way to access Sapelo2, **connecting from a browser.**

URL: ondemand.gacrc.uga.edu

- Provides a graphical desktop on the cluster

Graphical applications can be utilized on the cluster far more easily and smoothly than using X11 forwarding.

- Interactive applications (MATLAB, Rstudio, ParaView, etc)

- Sessions are not lost when network connection is interrupted or local computer is shutdown.

Can start a session from a browser on one computer, and connect to it from a browser on another computer.

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

https://kaltura.uga.edu/media/t/1_u9d1xrpp/176125031

Some applications that interface with the cluster

➤ MATLAB

Run client with GUI on local machine and configure it to offload work to the cluster.

<https://wiki.gacrc.uga.edu/wiki/MATLAB-Sapelo2>

➤ Jupyter Notebook

Run the notebook on Sapelo2 and connect to it from a browser on your local machine

<https://wiki.gacrc.uga.edu/wiki/Jupyter-Sapelo2>

Consulting and Support

- Install and update software per user request.
- Troubleshoot programs, jobs, and workflows.
- Help users optimize their code.
- Help users implement or automate pipelines.
- Provide user training (Linux, Sapelo2 new user, Python, R, conda environment, etc.).
- Assistance with grant proposal preparation, where GACRC resources are used.

https://gacrc.uga.edu/about/how_we_help_researchers.php

Documentation and Contact

Wiki: <http://wiki.gacrc.uga.edu>

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

Help: <https://help.gacrc.uga.edu>

Videos: <https://kaltura.uga.edu/channel/GACRC/176125031>

Offices: GACRC staff are located in the Computing Services Building (formerly called Statistics Building), rooms 101 to 108.

Office Hours via Zoom (preferred).

Thank you!