Overview of GACRC resources

Georgia Advanced Computing Resource Center (GACRC)
Enterprise Information Technology Services (EITS)

UNIVERSITY OF GEORGIA
Georgia Advanced Computing Resource Center

Outline

- GACRC
- Overview of Sapelo2 Cluster
    - Cluster diagram
    - Computing resources
    - Computational Partitions
    - Storage and file systems
    - CSP compute nodes
    - Software Environment
- Globus – file transfer automation
- Open OnDemand interface
- GACRC wiki and User Support

GACRC

- A high-performance-computing (HPC) center at UGA
- Advanced computing environment for the UGA research and education community
  - HPC infrastructure location in the Boyd Data Center
  - Comprehensive collection of scientific, engineering and business applications
  - Consulting and training services
  - Teaching cluster

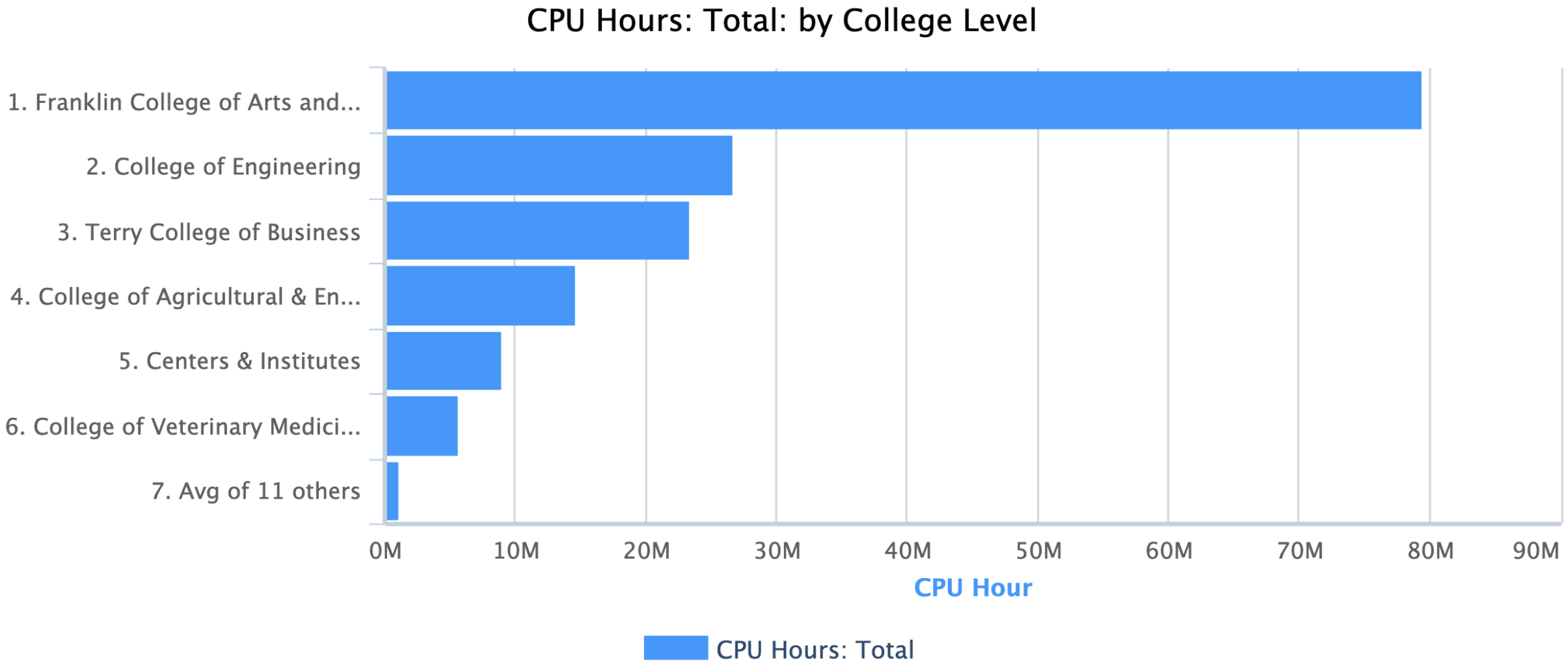
- GACRC Wiki: http://wiki.gacrc.uga.edu
- Help and Support: http://help.gacrc.uga.edu
- GACRC website: http://gacrc.uga.edu
- Kaltura Channel (tutorial videos): https://kaltura.uga.edu/channel/GACRC/176125031

Sapelo2 Cluster at the Boyd Data Center

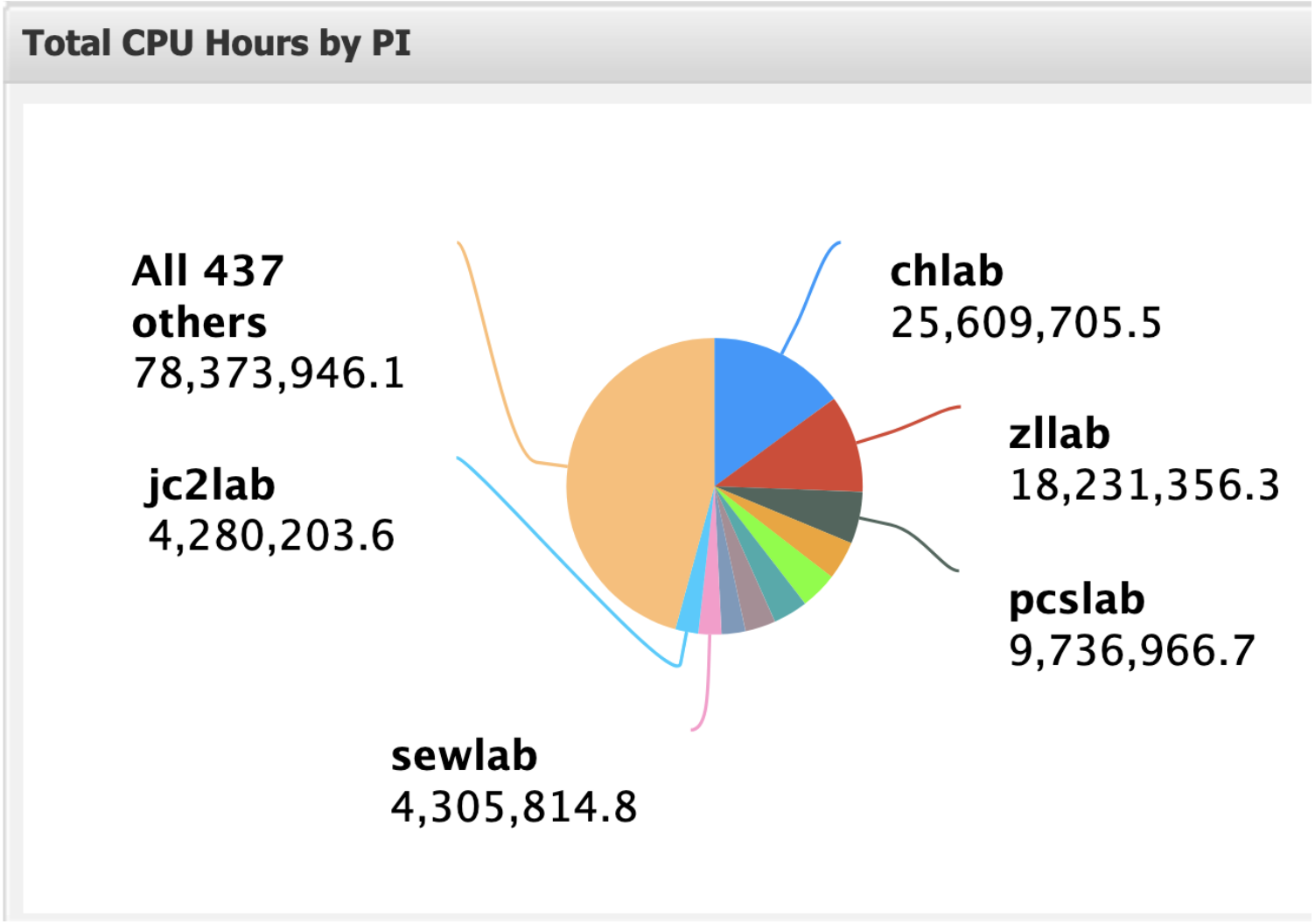
Rows with racks of compute nodes

Back of the compute node racks

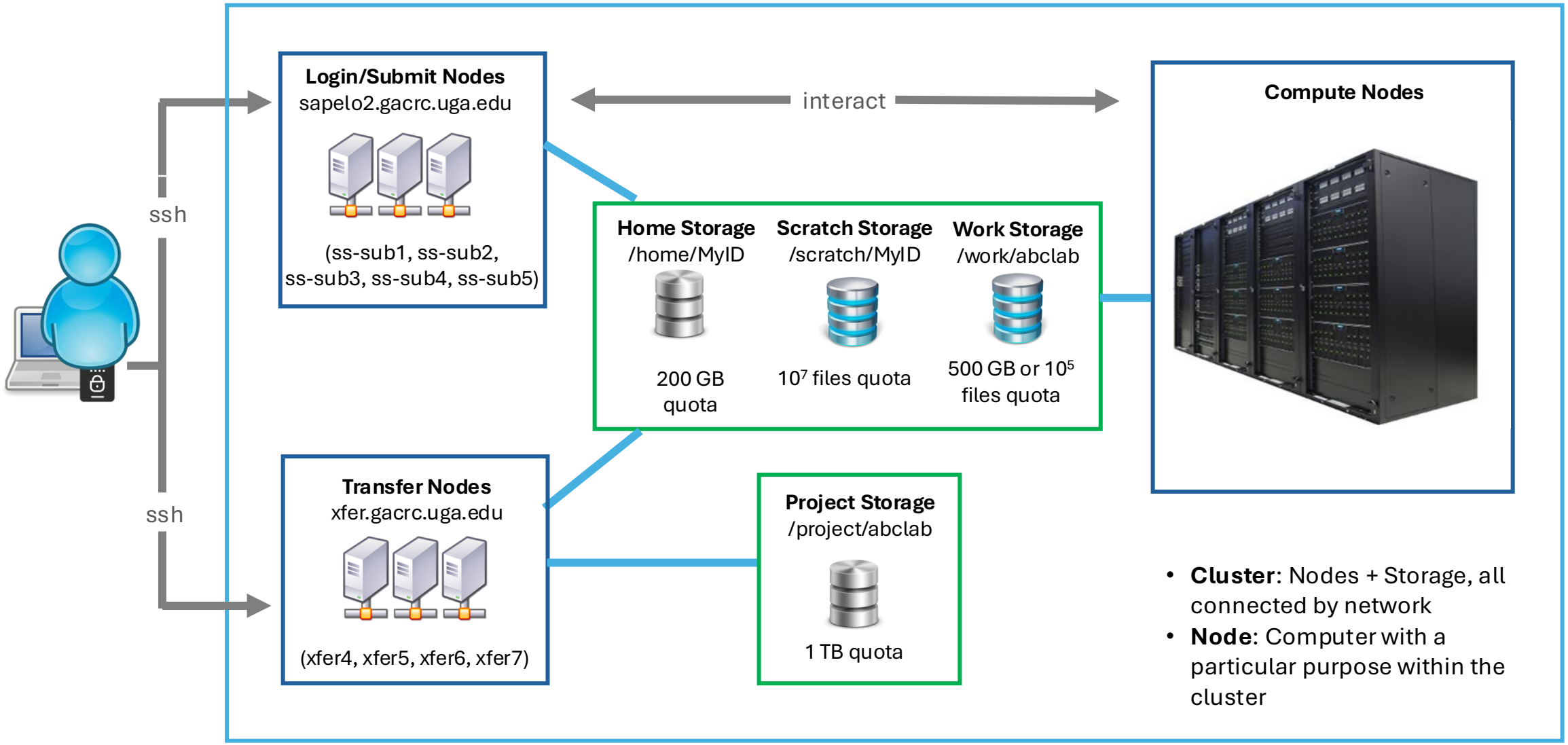
Sapelo2 usage by College (7/31/25-7/31/26)



Sapelo2 usage by Group (7/31/25-7/31/26)



Sapelo2 Cluster



- **Cluster**: Nodes + Storage, all connected by network
- **Node**: Computer with a particular purpose within the cluster

Sapelo2 – Computing Hardware

➢ **Resources available to all users**

- Regular memory nodes: 32 to 384 cores/node, 2GB to 6GB/core
- High memory nodes: 512GB, 1TB, 2TB, or 3TB/node
- GPU nodes: NVIDIA quad A100, quad H100, quad L4, P100 devices
- Number of CPU cores: ~ 39,000

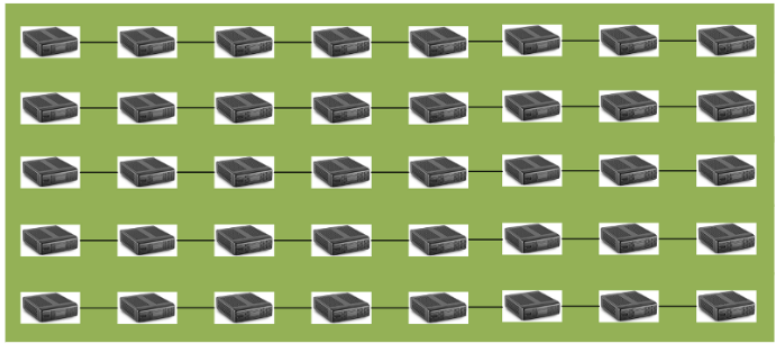
➢ **Buy-in nodes**

- Various configurations, including GPU nodes (NVIDIA L40S, A100, H100, H200, B200).
- Number of CPU cores: ~ 11,500

➢ Total number of CPU cores: ~ 50,500

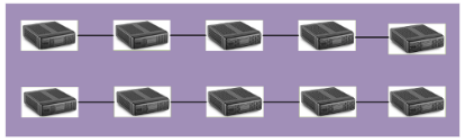
Computational Partitions

Compute Nodes



batch for up to 7 days
batch_30d for up to 30 days

→ For regular jobs up to 740GB per node

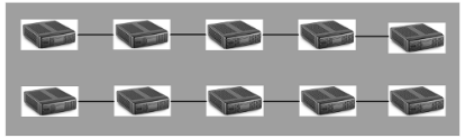


highmem_p for up to 7 days
highmem_30d_p for up to 30 days

→ For high memory jobs up to 990GB per node

hugemem_p for up to 7 days
hugemem_30d_p for up to 30 days

→ For huge memory jobs up to 3000GB per node



gpu_p for up to 7 days
gpu_30d_p for up to 30 days

→ For GPU enabled jobs

inter_p for interactive jobs

→ For interactive jobs

Compute nodes are divided into groups called **partitions**. A **partition** is a collection of compute nodes for a particular computing need.

Computational Partitions - continued

Partition	Total Nodes	Max Mem(GB) /Single-node job	Cores /Node	Processor Type	GPU Cards /Node
batch batch_30d	1	740	384	AMD EPYC Turin	N/A
	64	740	128	AMD EPYC Genoa	
	120	500	128	AMD EPYC Milan	
	4	250	64	AMD EPYC Milan	
	114	120	64	AMD EPYC Rome	
	25	120	32	AMD EPYC Naples	
	40	180	32	Intel Xeon Skylake	
highmem_p highmem_30d_p	10	500	32	AMD EPYC Naples	N/A
	12	990	32	AMD EPYC Milan	
	2	990	128		
hugemem_p, hugemem_30d_p	3	3000	48	AMD EPYC Genoa	N/A
	2	2000	32	AMD EPYC Rome	
gpu_p gpu_30d_p	14	1000	64	AMD EPYC Milan	4 NVDIA A100
	14	1000	64	Intel Xeon SapphireRapids	4 NVIDIA H100
	12	740	128	AMD EPYC Genoa	4 NVIDIA L4

Sapelo2 – CSP compute nodes

➢ Available to all CSP members (partition name: csp_p)

- ra1-8 : 52 cores, 256GB RAM, Intel Icelake, local drive with 890GB of space
- ra1-9 : 40 cores, 192GB RAM, Intel CascadeLake, local drive with 890GB of space
- ra1-10 : 36 cores, 256GB RAM, Intel CascadeLake, local drive with 440GB of space
- ra1-11 : 24 cores, 256GB RAM, Intel Skylake, local drive with 930GB of space
- ra1-12 : 36 cores, 256GB RAM, Intel Skylake, local drive with 930GB of space

➢ Available to Dr. Hall's group (partition name: hall_p)

- a1-[14-20]: 32 cores, 128GB RAM, AMD EPYC Rome, local drive with 890GB of space

➢ Maximum walltime limit of 30 days, can be extended

Sapelo2 - Storage

➢Storage to support projects that use the computing cluster only.

•ZFS storage appliance (500TB):
/home (200GB/user quota, no purge, snapshots, backup)

•Aeon hybrid Lustre appliance (2.4PB flash & 9.6PB spinning drives)
/scratch (no capacity quota, 10^7 file/user, 30-day purge, no snapshots, no backup)
/work (500GB and 500k file/group quota, no purge, no snapshots, no backup)

•Local hard drive (SSD, NVMe) on compute nodes: /lscratch (3.5TB of GPU nodes, up to 890GB on other nodes)

•ZFS storage appliance (8PB): /project (1TB/group, no purge, snapshots, backup)

•ZFS storage application (8PB): backups for /home and /project

Storage Spaces (Directories)

Directory	Name	Quota	Accessible from	Intended Use	Backed-up	Important Notes
/home/MyID	Home	200GB	**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	10^7 files	**Login** **Transfer** **Compute**	Temporary files needed for currently running jobs	No	Clean up when your job finishes! Subject to "30-day purge" policy
/work/abclab	Work	500GB 10^5 files	**Login** **Transfer** **Compute**	Input files needed for repeated jobs	No	Clean up when your job finishes! Group sharing is possible
/project/abclab	Project	1TB (initial)	**Transfer**	Temporary data parking	Yes	Group sharing is possible
/lscratch	Local Scratch	3.5 TB	**Compute**	Jobs with heavy disk I/O operations	No	Clean up when job exits from node!

Scratch 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 within **30 days** will be automatically deleted off the /scratch file system.
- Measures circumventing this policy will be monitored and actively discouraged.

- Copy files in /scratch that you want to keep to /project when the job is done
  - Can use Globus or fpsync on xfer nodes to do this quickly
- Move all unnecessary files and folders to /scratch/trash/$USER
- When you archive data using tar on /scratch, please do not use z option (compression option).
  - After you archive data with tar, you can use gzip to compress it.

Sapelo2 - Software

- 64-bit Linux OS (Rocky 9.7).
- Slurm queueing system to submit jobs to the compute nodes (e.g. sbatch, squeue, sacct).
- Over 1500 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 (e.g. MATLAB, Amber, Gaussian), some limited to a few groups that purchased licenses (e.g. SAS, Stata, Ansys, Schrodinger).

Software Environment

- Software is installed as “software modules”
- Naming format of software modules: **Name/Version-Toolchain**
  - e.g., **Python/3.12.3-GCCcore-13.3.0**
- Module commands:
  - ml spider *pattern* : Search module names matching a *pattern*
  - ml *ModuleName* : Load a module into your submission script
  - ml : List modules currently loaded
  - ml -*ModuleName* : Remove a module
  - ml purge : Remove all currently loaded modules

NOTE: DO NOT LOAD/USE MODULES ON THE LOGIN/SUBMIT NODES! (ss-sub1, ss-sub2, ss-sub3, etc…)

Software Environment - toolchains

- When you load more than one software module, you must pay attention to their toolchain compatibility
- Having modules loaded with incompatible toolchains can cause errors in your job, such as:

```
Lmod has detected the following error:
These module(s) exist but cannot be loaded as requested
```

- The toolchain compatibility table can be found at https://wiki.gacrc.uga.edu/wiki/Available_Toolchains_and_Toolchain_Compatibility

Globus – Transferring and sharing 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.
- Automate data backup.

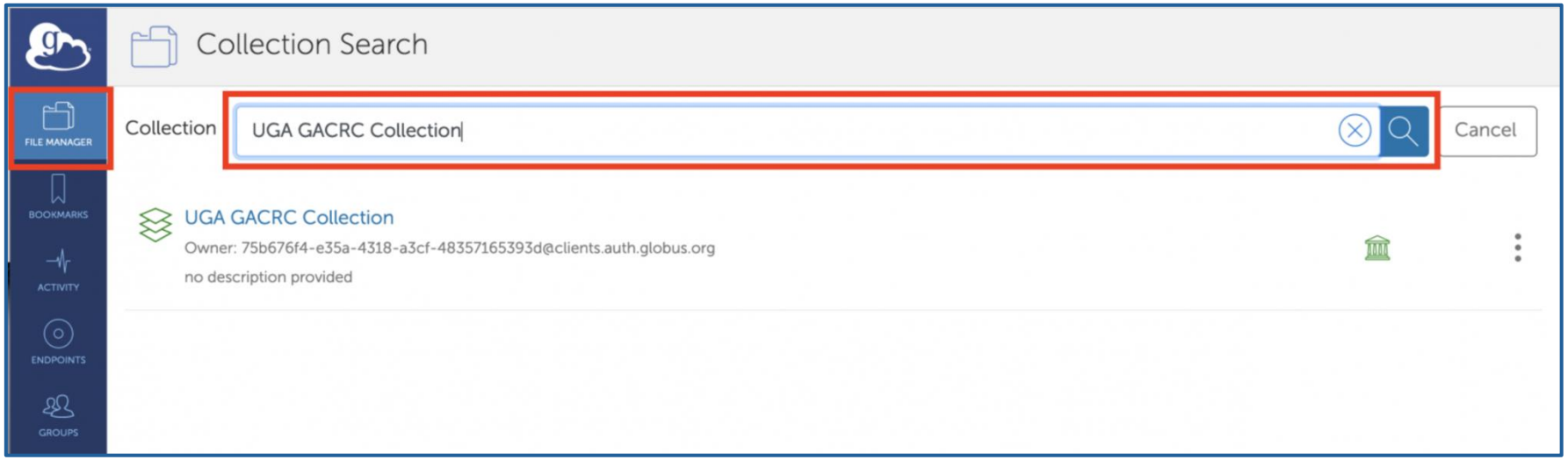
https://wiki.gacrc.uga.edu/wiki/Globus

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

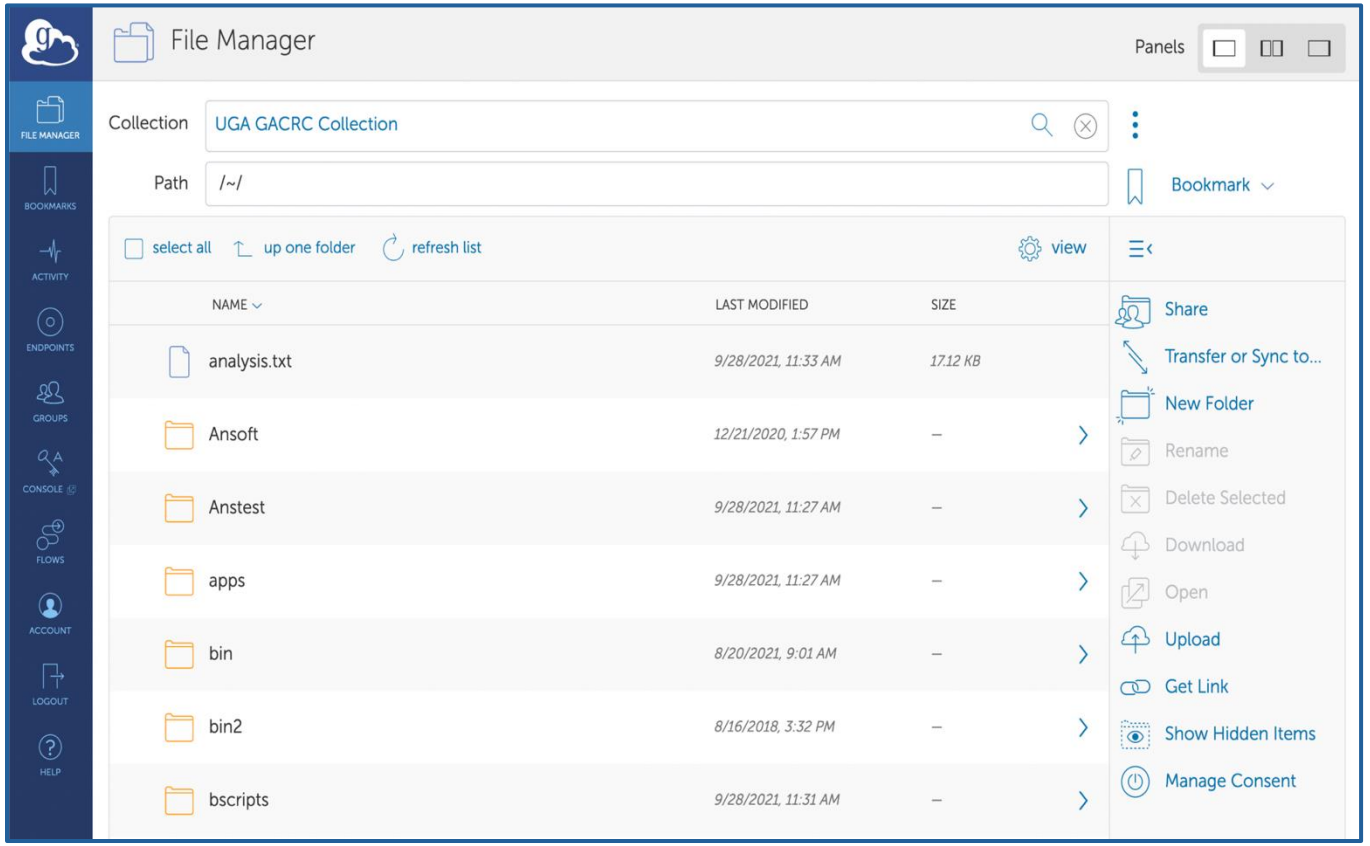
Globus – Transferring files

- Globus is a website you can use to transfer files between your local computer and the Cluster and is free for UGA members
- Create a Globus Identity Account at globus.org following the instructions at https://wiki.gacrc.uga.edu/wiki/Globus#Getting_Started
- From the File Manager tab at globus.org, search for the “UGA GACRC Collection”



Globus – Transferring files and automation

- When you open the UGA GACRC Collection in Globus, by default you will be in your home directory. You can change to any of your or your lab's other directories by typing a path in the Path bar.
- Use the **Upload/Download** options on the right for transferring **small** or a **few** files.
- Use the **Transfer or Sync** option for **many** or **large** files (requires Globus Connect Personal on your computer: https://wiki.gacrc.uga.edu/wiki/Globus_Connect_Personal)



Open OnDemand interface

➢Alternative way to access Sapelo2, connecting from a browser.
URL: https://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, Jupyter notebook, VSCode, 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

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, virtual 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

- GACRC Wiki http://wiki.gacrc.uga.edu
- GACRC Website http://gacrc.uga.edu
- GACRC training videos https://kaltura.uga.edu/channel/GACRC/176125031
- Globus https://wiki.gacrc.uga.edu/wiki/Globus
- Open OnDemand https://wiki.gacrc.uga.edu/wiki/OnDemand
- Trainings Offered https://wiki.gacrc.uga.edu/wiki/Training
- Getting Support http://help.gacrc.uga.edu
- GACRC Offices: Computing Services Building, rooms 101 to 108

Thank you