GACRC AC Meeting Jan 2016

Dr. Guy Cormier

Director of Research Computing

THE UNIVERSITY OF GEORGIA
Office of the Vice President for Information Technology
Enterprise Information Technology Services

Sapelo Update

- 56 (42) groups – 121 (89) users
- 22 (19) buy-in groups, 2 groups in discussion
- 28 buy-in nodes, 24 match nodes
- Cores: 5,936 + 2,080 = 8,106   (26% buy-in)

- Deployed add. 240TB on Lustre scratch
  - 150TB for EuPathDB
  - 90TB for GACRC usage (now at 330TB)
- Deployed 2 additional data transfer nodes

EITS
ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

Update on Training

- New User Training
  - Introduction to zcluster at GACRC
  - Introduction to Sapelo at GACRC

- Linux Basics and Hands-On Sessions
  - Introduction to Linux Basics – Part 1
  - Linux Hands-On Practice Session – Part 1
  - Introduction to Linux Basics – Part 2
  - Linux Hands-On Practice Session – Part 2

- Topical sessions:
  - How to run NCBI Blast on zcluster at GACRC
  - How to install software on zcluster at GACRC
  - Introduction to GACRC Storage Environment

EITS
ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

NSF MRI

- SHARC – Southeast Hub for Advanced Research Computing
- Clemson, UGA, Georgia State & College of Charleston
- The investment will provide UGA with five years of –
  - reserved 10% portion of the instrument, including HPC and Informatics resources, high-speed storage and a backup environment.
  - shared access to a non-reserved 50% portion of the instrument which would be made accessible to institutions across the Southeast.
  - one voting seat on the Center's Executive/Steering Committee and two seats on the Science Advisory and User Committee. These seats are only available to the four founding institutions of the Center.
  - access to computational support experts from participating institutions.

EITS
ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

NSF MRI (cont.)

- $5.6M project
- UGA investment is $225k over the intial 3 years of the project and $50k for the last 2 years of the project
- Clemson provides
  - colocation in a data center with power, cooling, security etc.
  - Systems administration, NOC, network engineering, etc.
  - Access to a 9PB backup environment not part of the MRI request
- 1.6 Pflop/s (theoretical)
- heavy on GPUs (384x nVIDIA Pascal-class)
- 4x 2TB compute nodes
- 6.5PB in disk storage
- 454TB in SSD-based storage for Informatics (Hadoop/Spark)

EITS
ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

Projects Underway

- Backup environment
- Development cluster
- Expansion of IB fabric on Sapelo
- New cluster management stack on Sapelo
- Discuss the future of zcluster
  - Dismantle and repurpose parts in Sapelo?
  - Decommission completely & increase Sapelo resources?
- Research IFS

Request from faculty in Marine Sciences & Engineering

- 9 faculty involved in Marine-related research wanting to pool their buy-in resources

- Effectively create a Marine cluster

- Feasible through queueing/scheduling system

- Do we require specific MOU?

- Any other concerns?

Request from faculty in Mathematics, CPH & Pharmacy

REDCap instance for UGA

- widely used software for research data capture
- http://project-redcap.org/
- 1,749 institutional partners across 97 countries
- Request was made for GACRC to host through a VMware instance.
- Would require EITS support. Cost-recovery service?
- MOU/SLA? With who?
- David Lee aware of project

Starting January 2016, the GACRC will institute monthly maintenance windows in order to perform maintenance operations requiring system operations to be reduced or interrupted.

The schedule will be as follows:
- The last Wednesday of each month from 10AM to 4PM will be reserved for partial cluster maintenance.
- Twice a year, a two-day shut-down of GACRC services will be scheduled for more complex maintenance operations. These will occur on the last Tuesday and Wednesday of the months of January and July.

These maintenance windows represent periods when the GACRC may choose to drain the queues of running jobs and suspend access to either or both clusters, as well as storage devices for maintenance purposes. Interruptions will be kept as brief as possible.

The GACRC will notify all users at least 10 days in advance that a maintenance window will be in effect. The notification will describe the nature and extent (partial or full) of the interruptions of cluster and or storage services. In case a maintenance window has to be extended due to unavoidable technical reasons, adequate communications will be made to all users.

The impact of the outages will vary, and the GACRC will do its best to preserve pending and running jobs, which is often very doable. Nevertheless, users will need to plan their job submissions around the maintenance windows.