GACRC AC Meeting Sept 2015

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

Sapelo Update

- 31 groups – 66 users
- 14 buy-in groups, 3 groups in discussion
- (20+4) buy-in nodes, (16+3) match nodes
- 78 software, libraries, languages installed

- Engaged Penguin in streamlining user/group management
- Engaged Adaptive in queuing/scheduling configurations

Ⅱ EITS

ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

User Community

User Community

	Groups	Users
Active (past year)	183	919
Inactive (past year)	53	117
Total	236	1036

EITS
ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

Training

Update on Training

- 5x workshops developed – more to come
  - Intro to zcluster
  - Intro to Sapelo
  - Linux Basics I & Linux Basics II
  - GACRC storage and data management
- 20 sessions delivered, 7 sessions planned for September
- All new users must go through training before getting accounts
- To date, sessions have touched 142 individuals

FY16 Budget

FY16 Budget

- $125k (EITS) & $125k (OVPR)
- $350k (OVPR) (expended with Sapelo)
- $290k (EITS) additional funding
- $100k (VPIT) buy-in match funds

Back-up & Data Migration

GACRC backup services

Goal: Implement restoration and DR services for UGA research data

A software based solution is ideal
- Using vendor supported GACRC hardware
- Backup from disk to disk using resiliently configured InfiniBand connected storage systems

Backup cycles
- Target file systems backed up every (7) days
- Snapshots or daily backups are being considered for /home
- Files deleted by users pruned from backup storage after (30) days
- *Options are being assessed for disk to disk (perhaps to tape)*

Target file systems
- /home, /project, /usr/local, /db

GACRC user data migrations

Goal: Reduce complexity, improve performance and protect research data

- All Sapelo SCRATCH data migrated from Icebreaker storage to the Lustre storage (on-going)

- All OFLOW data migrated from ArchStor and Panasas to PROJECT (on-going)

- All zcluster HOME folders migrated from Panasas to Icebreaker storage (HOME or PROJECT)

- ArchStor servers to be decommissioned or re-assigned to other groups

- Panasas re-purposed as zcluster SCRATCH file system

Staffing

Staffing Proposals

- **Bioinformatics Trainer**
  - Waiting for H1B process to be completed
  - Staff member could start mid-September

- **Sysadmin/Storage Engineer**
  - Increased operational complexity
  - Brings sysadmin staff from 2 to 3

- **Computational Scientist(s)**
  - Short-term embedding within research groups
  - Workflows, dataflows, scripting, serial → parallel, optimizations, etc.

- **Enabler(s) "for-hire" – sysadmins, DBAs, Viz, Data scientists, etc.**
  - %FTE across many research groups, core facilities
  - Linux-based research computing
  - EITS bridge funding
  - Not to replace dept. IT staff

EITS
ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

Other Issues

Other Topics

- GACRC now reports to CTO
- EITS is under Administrative Review in FY16
  - GACRC is subject to this review
  - Organizational self-study already initiated (Bob, Mike, Peter & Guy involved)
  - GACRC-AC might be involved
- Georgia Informatics Institute
  - GACRC seen as resource
  - Jessie on committee

Extra

GACRC user data migrations   (cont.)
Migration Status

- **/scratch** migration to Lustre (in progress)
  - **August 31**: initial communication to users to migrate data by September 11
  - **September 14**: assess user compliance
  - **September 19**: disable access to the relevant icebreakers used for legacy /scratch

- **/oflow** migration (in progress)

- /home migrations (phase 1)
  - Initial technical plan finalised
  - Dependencies on icebreaker storage liberated from /scratch migration effort

EITS
ENTERPRISE INFORMATION
TECHNOLOGY SERVICES

GACRC user data migrations (cont.)
Migration Status

- **/home** migrations (phase 2)
  - Communication instructing users to migrate data to new home directories (100 GB quota), and lab /project folders (1 TB quota)
  - Monitor data migration progress and assist users when required through completion

- *zcluster* **/scratch** migration (pending)
  - Provision the Panasas as a scratch file system for the zcluster
  - Communicate data migration details to users and assist users when required through completion
  - Re-purpose liberated Icebreaker storage to expand the **/project** file system