



GACRC AC Meeting 04/29/15

Dr. Guy Cormier

Director of Research Computing



THE UNIVERSITY OF GEORGIA

**Office of the Vice President for Information Technology
Enterprise Information Technology Services**

March 19, 2013

Storage Update

- 240TB Lustre Appliance
 - \$120k investment by EITS, in addition to GACRC budget
 - Installed. Currently being configured/optimized for our workload
 - scratch for Sapelo
 - 30-day policy will be implemented through policy engine (robinhood)
- 1st phase of backup environment
 - 240TB being purchased
 - Located in HSC in EITS data center
 - Will target the Panasas and the new IceBreaker Project space
 - 2nd phase after July 1st with deployment of an additional 480TB
- New data transfer nodes (2x) deployed. Will be used for both Sapelo and zcluster. Significantly higher network performance on these.

Sapelo Update

- Still ramping up to full production
- Currently have 27 users across 16 groups
 - Not just HPC, also HTC
 - Current limiting step is installation of software & dependencies
- Torque/Maui → Moab HPC Suite (\$120k over 3yrs)
- LDAP deployment to use MyID
- Usage is already significant. Jobs waiting in queue

Groups on Sapelo

PI	Dept.	# users
Magdy Alabady	Plant Biology/GGF	1
Michael Bachmann	Physics	3
Renato Castelao	Marine Sciences	3
Daniela Di Iorio	Marine Sciences	2
Natarajan Kannan	Biochemistry & Molecular Biology	1
Jessie Kissinger	Genetics/IOB	1
David Landau	Physics	2
Jim Leebens-Mack	Plant Biology	2
Walt Lorenz	QBCG	1
Bob Schmitz	Genetics	2
Rajagopalbab Srinivasan	Entomology	1
Philip Stancil	Physics	1
CJ Tsai	Forestry	1
XQ Wang	Engineering	4
Shaying Zhao	Biochemistry & Molecular Biology	1
Travis Glenn	Environmental Health Science	1

Past year's usage on zcluster

# Jobs	TOTALS	%
Serial Jobs	3,103,888.0	98.60%
MPI/multicore jobs	44,111.0	1.40%
TOTALS	3,147,999.0	

Wall Time (core.days)	TOTALS	%
Serial Jobs	205,538.6	40.18%
MPI/multicore jobs	306,030.7	59.82%
TOTALS	511,569.3	

Current Buy-In Participants

Robert Schmitz	Genetics	1+1
Xianqiao Wang	Engineering	2+1
Travis Glenn	Environmental Health Science	1+1
Renato Castelao	Marine Sciences	2+1
Jessie Kissinger	Genetics / IOB	1+1
C. Brock Woodson	Engineering	2+1
Catherine Edwards	Skidaway	2+1
Jason Christian	Engineering	1+1
Adrian Burd	Marine Sciences	1+1
		22

in process or under discussion

Kissinger, Patterson, Tsai, Schmitz, Zhao	IOB	240TB JBOD
Jim Leebens-Mack	Plant Sciences	1+1
Devos & Bennetzen	Crop & Soil Sciences/Genetics	1+1
Pej Rohani	Ecology	3+3
Daniela Di Iorio	Marine Sciences	2+1
Merryl Alber	Marine Sciences / Sapelo Island	2+1
Christof Meile	Marine Sciences	2+1

Buy-In Program Issues

- Diversity in fund types
 - State vs. Federal equipment vs. Federal other
 - MOU initially based on service accounting codes
 - Needed to address diversity in accounting codes
- Group participation
 - Diversity in fund types and origins for single MOU leads to issues with audit trail
 - Requires single consolidated account
 - Single administrative point of contact

User Management Issues

- Time for account clean-up
 - Users with no MyID
 - Affiliates will all be identified & verified with PI
 - Groups with no previous year usage. Contact. Remove if requested.
 - **Go-ahead?**
- LDAP to access MyID authentication
 - Currently working on Sapelo
 - Planning to install on zcluster
 - OpenLDAP installation for affiliate accounts
- New account forms
 - Group, User and Affiliate forms
 - **Require actual signatures or webform?**
 - Includes Export Control and strong policy language
 - **Yearly process?**

Training

- Dr. Zhuofei (Moses) Hou on-board
 - In-depth exploration of Sapelo and zcluster
 - running variety of code with Shan-Ho and Yecheng
- Will extend offer to Bioinformatics Trainer candidate
 - H1B will be necessary
 - Once the process is engaged, she will get access to zcluster and Sapelo, for her to explore
- Yecheng is spending much time discussing with IOB students after classes.
- Yecheng and Shan-Ho working one-on-one with Sapelo users

Training Materials

- Presentations being worked on
 - Introduction to Sapelo
 - Introduction to zcluster
 - Comparisons between Sapelo and zcluster
 - GACRC storage environment
 - Intro to Linux CLI and basic scripting
- Sapelo software documentation
- GACRC Cheat Sheet
- Schematics of GACRC infrastructure and workflows
- Updates to the wiki and main website
- Overview flyer for new faculty

