

# UAB Research Computing

Resources and Activities

Research Computing Day  
September 13, 2012

UAB IT Research Computing

# UAB IT Research Computing Team



**Bob Cloud**  
Executive Director  
Infrastructure Services  
UAB IT



**David Shealy**  
Director  
Research Computing  
Physics / UAB IT



**Puri Bangalore**  
HPC Consultant to  
Research Computing  
CIS / UAB IT



**John-Paul Robinson**  
System Architect  
Research Computing  
UAB IT



**Mike Hanby**  
Information Systems  
Specialist  
Engineering / UAB IT



**Thomas Anthony**  
Intern  
Research Computing  
UAB IT



**Shantanu Pavgi**  
Programmer/Analyst  
Research Computing  
UAB IT



**Bill Bradley**  
System Programmer  
Infrastructure Services  
UAB IT

Engaged professional staff collaborate with investigators to maximize the research impact

# Research Computing System

The Research Computing System is designed to provide services to researchers in three core areas

- **Data Analysis and Simulation**

Using the High Performance Computing (HPC) fabric we call [Cheaha](#) for analyzing data and running simulations. Many [applications are already available](#) or you can install your own

- **Data Storage and Sharing**

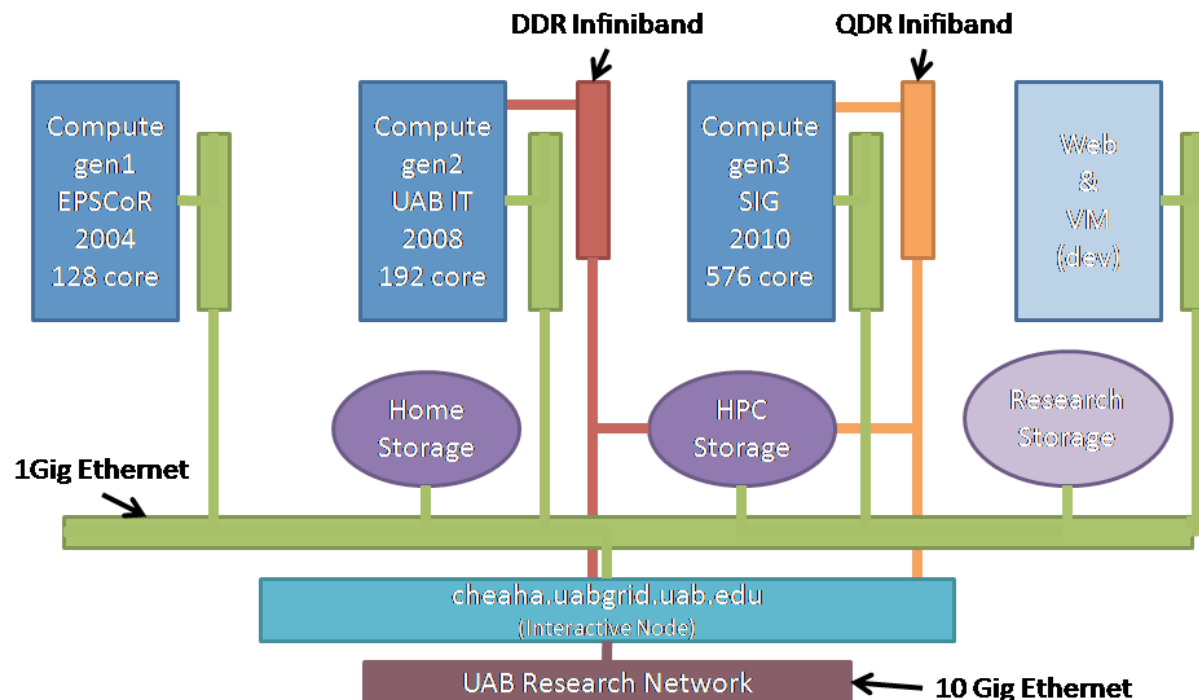
Supporting storage solutions and the trusted exchange of information using virtual data containers to spark new ideas (available beginning October 2012)

- **Application Development**

Providing virtual machines and web-hosted development tools empowering you to serve others with your research

# UAB Research Computing System

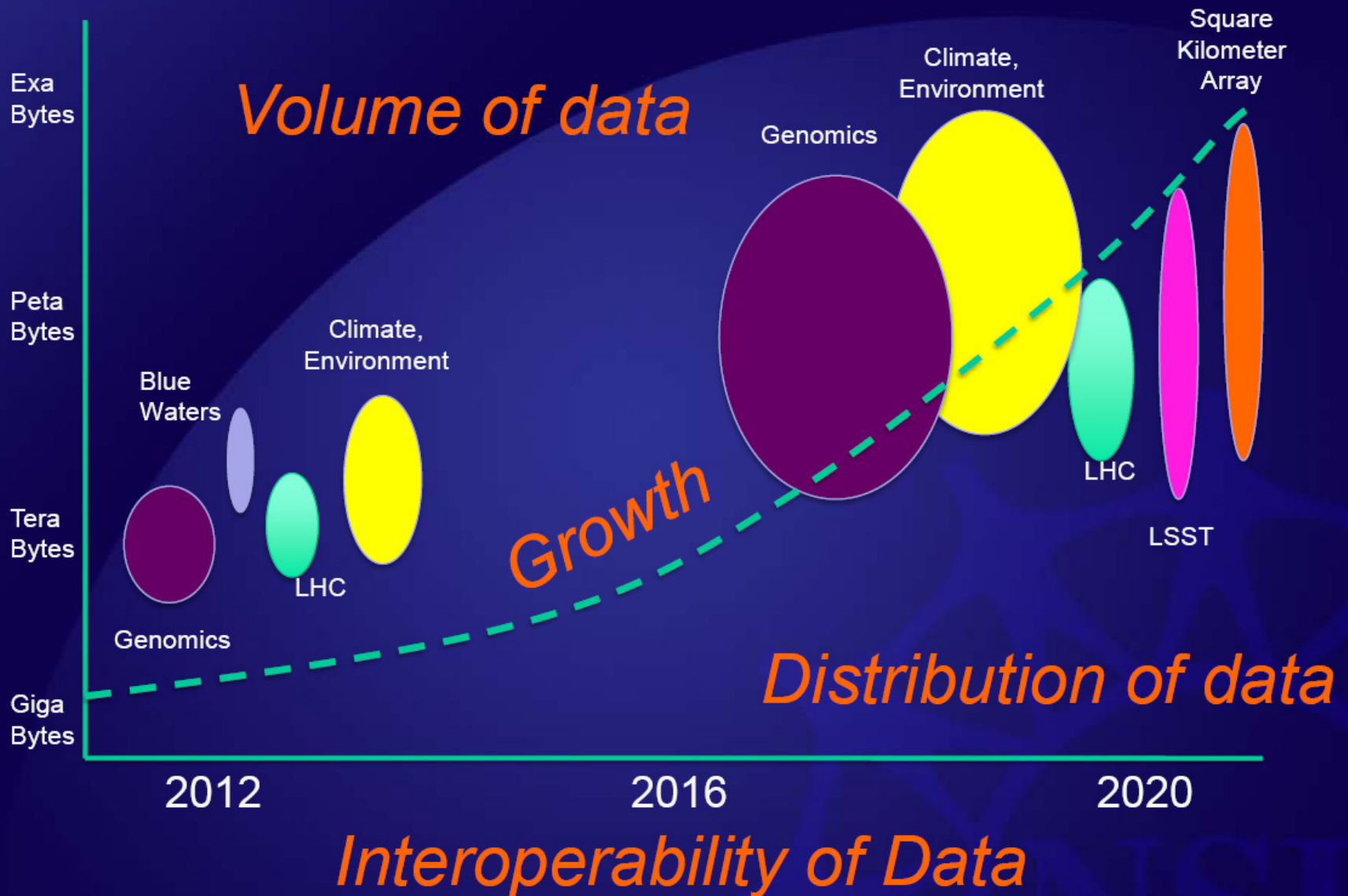
- Operated by UAB IT Research Computing Group
- 868 cores, three hardware generations, one SGE scheduler
- ~9Tflops compute power, 180TB High performance storage
- **Research Storage – bringing ½ Petabyte on-line over next few months**
- Virtual Machine fabric with OpenNebula



<https://docs.uabgrid.uab.edu/wiki/Cheaha>



# Data Challenges



# Storage Cloud-scape



Other Clouds

## Other Clouds

- Commercial price schedules
- Not Close to UAB HPC
- Features dictated by service provider
- Eg.
  - Dropbox, SugarSync, Box.net, AmazonS3, Rackspace, \*Cloud



Campus Cloud

## Campus Cloud

- < Commercial Cost
- Close to UAB HPC
- Fully customizable

# Research Storage Project

- Fall 2011:
  - Announce intent to expand storage capacity for research use, including a goal of the “first Terabyte for free”
- Spring 2012 and Summer 2012:
  - Pilot a storage architecture with Dell Inc.
    - Includes storage and capacity to build storage-oriented applications
  - Explore use cases for research storage
    - Desktop integration – same view of files on cluster from desktop
    - Virtual containers – storage is virtualized for use in applications
    - Backup to storage – eg. build a Timecapsule for Macs
    - Group storage – dedicate storage for use storage for
    - Secure storage – TrueCrypt containers across platforms
    - Mobile access – iFiles on the iPhone
  - Acquire 420TB storage from Dell Inc.
- Fall 2012
  - Begin roll out of new storage (beginning in October)
  - First TeraByte for Free
  - Offer storage cost model both incremental and condo
  - Initial features
    - Cluster access
    - Campus access (desktop integration via CIFS)
    - Mobile access (via standard apps)



&

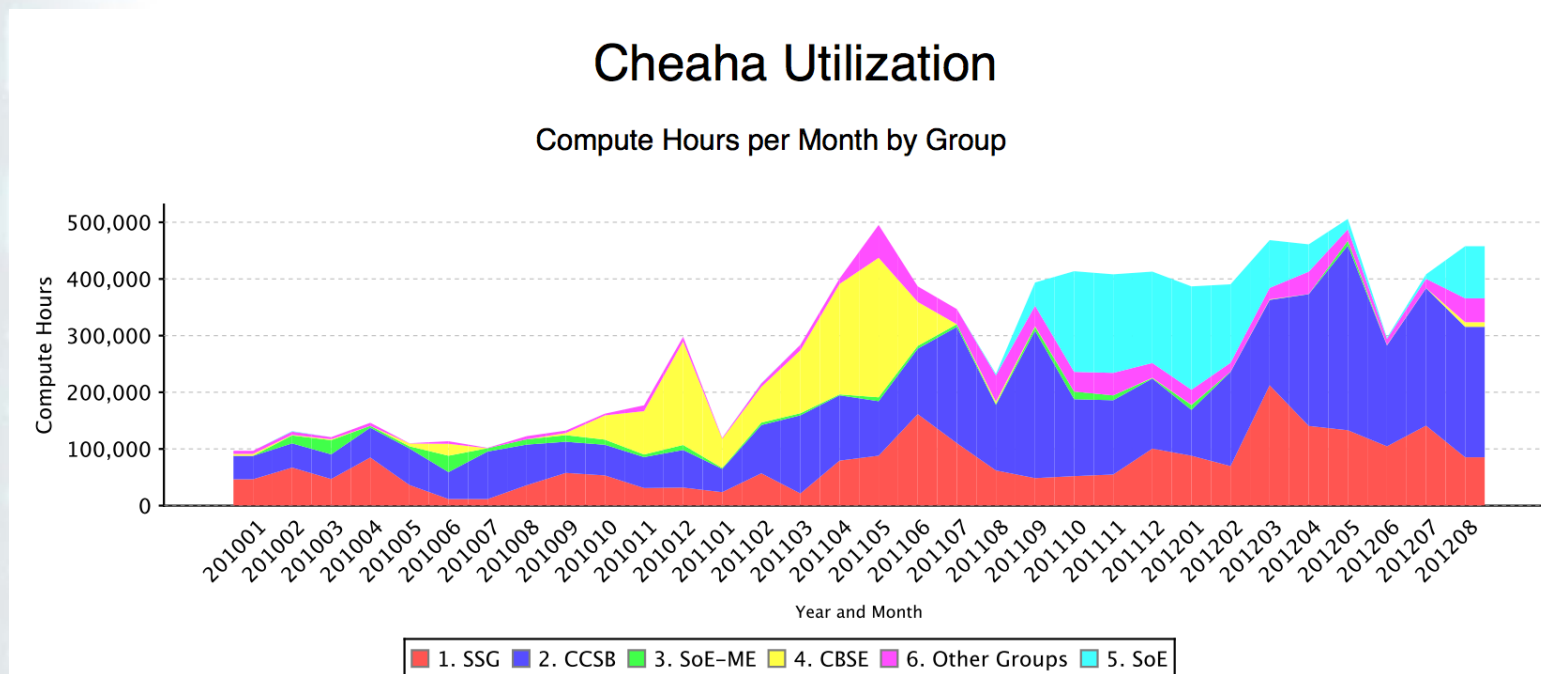


# Research Storage Cost Models

- Campus Cloud Storage (raw storage):
  - 1 TB – free for researchers as announced on Sept 15, 2011 (no backup)
  - >1 TB- Cost: \$150/TB/Yr
  - Condo model: \$12,000 for 36TB Raw for 3Yrs
- Hosted Internet2 Net+ services with box.net
  - Data synchronization solution for all platforms
  - Special discussion at end of presentation



# Computing Use and Scalability



## Application Scaling Focus Areas

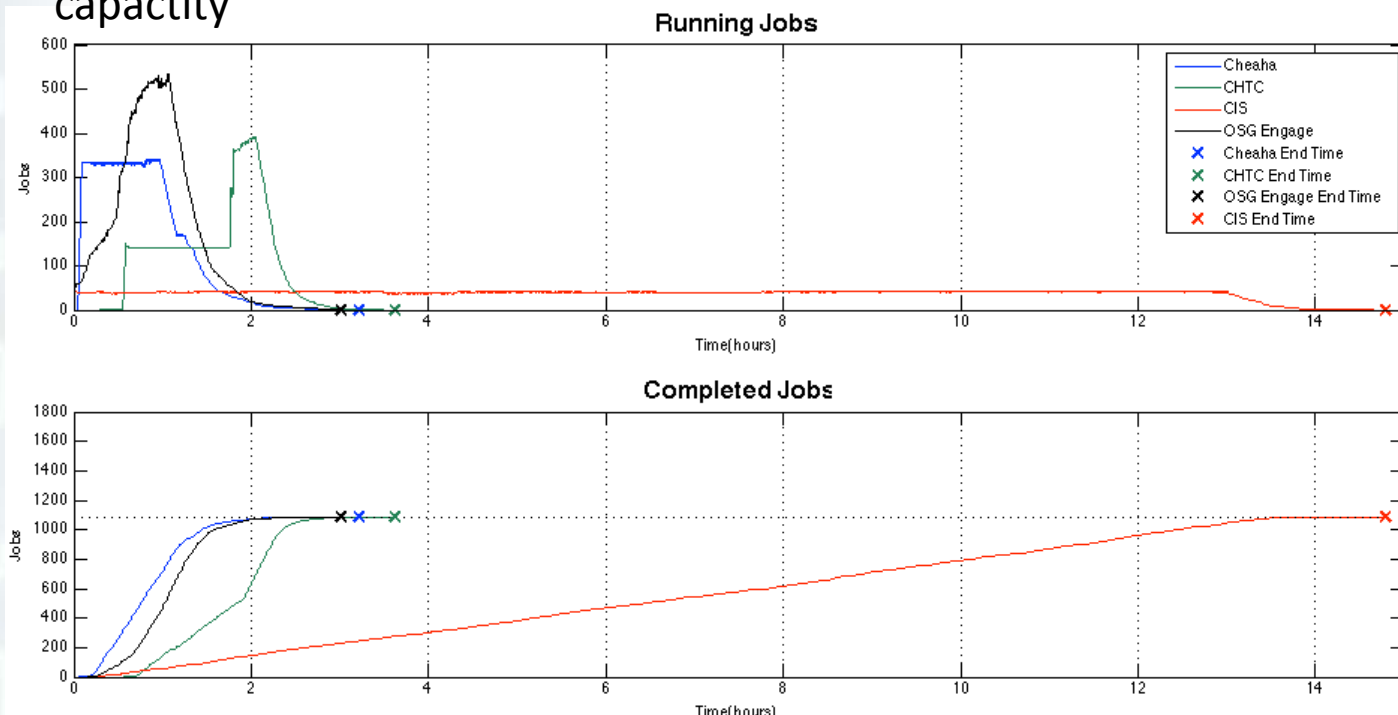
- Structural Biology
  - NAMD 2011
  - Autodock 2012
- Neurology
  - DTI image processing 2012
- Genetics
  - Galaxy 2011 (on going)

# Structural Biology Scaling 2012

## High-Throughput Molecular Docking

- Virtual Screening with Autodock
- UAB Condor Pilot
  - [https://docs.uabgrid.uab.edu/wiki/UAB\\_Condor\\_Pilot](https://docs.uabgrid.uab.edu/wiki/UAB_Condor_Pilot)
- Summary
  - Computing on spare cycle campus condor pool and OSG is comparable to UAB HPC capacity

- Project Partners
  - UAB IT Research Computing
  - Department of Computer and Information Science
  - UAB IT Desktop Computing



# Application Development Preview

## *Virtual Machines!*

The image displays the OpenNebula Sunstone web interface, which is used for managing virtual machines. The dashboard provides a summary of resources, historical monitoring information, and a quickstart guide.

**OpenNebula Sunstone Dashboard Summary:**

- Summary of resources:**
  - VM Templates (total/public): 2 / 0
  - VM Instances (total/running/failed): 2 / 1 / 1
  - Virtual Networks (total/public): 1 / 0
  - Images (total/public): 8 / 0
- Historical monitoring information:**
  - Total VM count: Graph showing total, active, and error counts over time.
  - Total VM CPU: Graph showing CPU usage over time.
  - Total VM Memory: Graph showing memory usage over time.
- Quickstart:**
  - New: ☐ VM Template, ☐ VM Instance, ☐ Virtual Network, ☐ Image

The VNC connection window shows a MATLAB R2012a virtual machine. The interface includes a file explorer, a command window, and a workspace. The command window displays the following output:

```
%>>
Command History
--
09/06/2011 06:08:14
%>> 09/06/2011 06:08:14
%>> 09/06/2011 06:08:15
%>> 10/03/2011 12:02:11
-File = 'http://www...'
-fullList = urload('...')
-File = 'http://www...'
-fullList = urload('...')
-setpref('Internet',...)
-sendmail('jpr@ub...'
Figure:
Figure: plottools;
--
07/15/2012 03:03:14
08/14/2012 05:50:11
08/15/2012 10:51:14
08/15/2012 10:52:14
08/15/2012 10:53:14
```

# What's Motivating Us?

- Research Reproducibility
- Pattern Recognition
  - Applications
  - Computers
  - Networks
- Researcher Autonomy

# Goals into 2013

- Build out storage services
  - Eventually bake in new features like redundancy, NSF4, AFP, and on and on
- Open virtualization fabric for general use
- Document operations to grow trust
- Build Community
  - hpc-announce and hpc-users lists
- Number one challenge is scaling effort!



# Special Discussion

## Research Storage Services and Cost

# Research Storage Cost Models

- Campus Cloud Storage (raw storage):
  - 1 TB – free for researchers as announced on Sept 15, 2011 (no backup)
  - >1 TB- Cost: \$150/TB/Yr
  - Condo model: \$12,000 for 36TB Raw for 3Yrs
- Hosted Internet2 Net+ services with box.net
  - Data synchronization solution for all platforms
  - Targeting rate of \$2.00/GB/Yr

# Box.net

- Hosted Internet2 Net+ box.net service
- Targeted Rate \$2.00/GB/Year
- InCommon(BlazerID) based user accounts
- Security Compliance
- Account Administration by UAB Research Computing
- Features: Online workspace, file sharing, mobile access, file commenting, group discussions, online documentation, updates, version history, user management, security controls, easy-to-use, box sync
- <http://www.internet2.edu/netplus/box/>