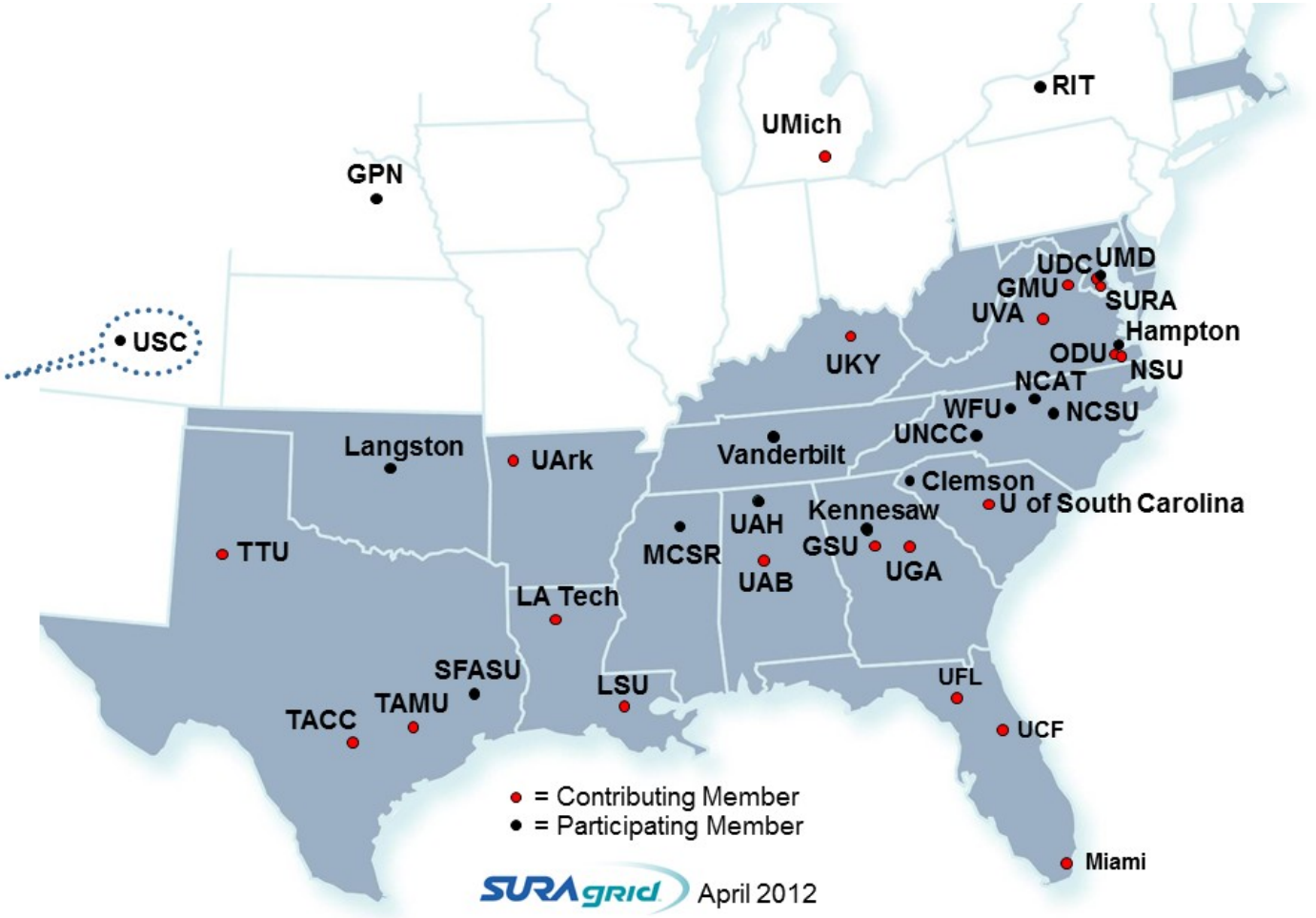
OSG Council Teleconference: November 13, 2012

The SURAgrid VO

Steve Johnson
Texas A&M University


RIT
UMich
GPN
UDC
UMD
GMU
SURA
UVA
Hampton
USC
ODU
UKY
NSU
NCAT
WFU
NCSU
UNCC
Langston
UArk
Vanderbilt
Clemson
U of South Carolina
Kennesaw
UAH
GSU
MCSR
UGA
TTU
UAB
LA Tech
SFASU
UFL
LSU
TAMU
TACC
UCF
= Contributing Member
= Participating Member
Miami
April 2012

History

- Began as NMI Testbed ~ 2002
- BridgeCA provided internal PKI
- Install fests of Globus to get sites running
- A brief foray into GRAM4
- Deployed a wide range of apps, some using Globus, some not
- Governance committee of 9 members
- Joined OSG about a year ago as SURAgrid VO

Past Applications

- SURA Coastal Ocean Observing and Predicting Program – ADCIRC, Wave Watch 3
- Cactus
- Turbulence Studies
- GROMACS
- Options Pricing
- Virtual Parasite
- Virtual Screening for Computation Chemistry
- CAM3 Climate Modeling
- CH3D Storm Surge Monitoring
- SURAgrid Teaching Environment
- Bio-electric Simulator for Whole Body Tissue
- Simulation-Optimization for Threat Management in Urban Water Systems
- Dynamic BLAST
- Multiple Genome Alignment

Current and Future Applications

- Statistics w/ R (using DAGMan for WMS)
- High Energy Physics (non-CMS/ATLAS phenomenlogical app)
- Schubert Calculus (custom program)
- Octave for general purpose use

Other Activities

- Gradual adoption of rpm-based OSG distro, mostly as sites upgrade to RHEL6.
- Getting ready for new PKI – 3 RA's, 3+ GA's, working on sponsors.
- Cloud Options Working Group
  - Futuregrid
  - NIST SAJACC
  - Cloud Plugfests
- XSEDE Partner
  - Increase and sustainability of XSEDE access to HBCU's/MSI's.
  - Regional workshops for scientific visualization

R on SURAgrid

- Six resources as of October 2012
- SGVO R administrator (SJ) installs and maintains R on the sites.
  - R, LAPACK, BLAS
  - R packages ncdf & SuppDists
- Keep it simple for all involved!
  - Use $OSG_APP/suragrid/etc/profile inside shell script to define R env vars for each site.
  - Sites may have custom high-performance build of R, so we need to take advantage of it.

Courtesy: Google Maps

- As of October, 2012, the following sites have a functioning R along with *ncdf* and *SuppDists* packages installed in $R_LIBS_SITE:
  - TAMU_BRAZOS
    Custom build with Intel compiler and MKL, RHEL5, very fast
  - TAMU_Calclab
    GotoBLAS2/gcc, openSUSE 12.1
  - TTU_ANTAEUS
    GotoBLAS2/gcc, RHEL5
  - Grid_UNESP
    GotoBLAS2/gcc, RHEL6
  - UTA_SWT2
    GotoBLAS2/gcc, RHEL5
  - FNAL_FERMIGRID
    GotoBLAS2/gcc, RHEL5

R Application Status

10/29/2012 17:17:09 CDT

refresh 10 minutes

Resource Group	Globus Resource	Free Slots	Max Jobs	Walltime (minutes)	R package tests lattice	SuppDists	ncdf
TTU-ANTAEUS	antaeus.hpcc.ttu.edu:2119/jobmanager-sge-suragrid	80	80	9999999	OK	OK	OK
	antaeus.hpcc.ttu.edu:2119/jobmanager-sge-serial	51	432	9999999	OK	OK	OK
TAMU_Calclab	calclab-ce.math.tamu.edu:2119/jobmanager-pbs-weekend	0	894	2520	OK	OK	OK
	calclab-ce.math.tamu.edu:2119/jobmanager-pbs-night	0	894	720	OK	OK	OK
GridUNESP_CENTRAL	ce.grid.unesp.br:2119/jobmanager-pbs-long_d	0	1	43200	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-medium	18	2032	1440	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-short_d	1	1	180	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-medium_s	0	2032	1440	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-long	18	2032	1440	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-short	18	2032	1440	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-long_s	0	2032	43200	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-medium_d	0	1	1440	OK	OK	OK
	ce.grid.unesp.br:2119/jobmanager-pbs-short_s	18	2032	180	OK	OK	OK
FNAL_FERMIGRID	fermigridosg1.fnal.gov:2119/jobmanager-condor-default	16	30000	1440	OK	OK	OK
FNAL_FERMIGRID_ITB	fgitbgkp2.fnal.gov:2119/jobmanager-pbs-batch	2	2	1440	OK	FAIL	ALERT
FNAL_GPGRID_1	fnpcosg1.fnal.gov:2119/jobmanager-condor-group_suragrid	1	1	1440			
SWT2_CPB	gk01.atlas-swt2.org:2119/jobmanager-pbs-input_q	2	2560	1440	OK	OK	OK
UTA_SWT2	gk04.swt2.uta.edu:2119/jobmanager-pbs-input_q	4	1140	1440	OK	OK	OK
TAMU_BRAZOS	lnnr.tamu.edu:2119/jobmanager-pbs-grid	256	256	480	OK	OK	OK
SPRACE	osg-ce.sprace.org.br:2119/jobmanager-condor-default	11	0	1440			
TTU-TESTWULF	testwulf.hpcc.ttu.edu:2119/jobmanager-sge-hep	312	320	9999999			
	testwulf.hpcc.ttu.edu:2119/jobmanager-sge-hase	32	32	9999999			
	testwulf.hpcc.ttu.edu:2119/jobmanager-sge-suragrid	80	80	9999999			
	testwulf.hpcc.ttu.edu:2119/jobmanager-sge-heptest	8	8	3600			
	testwulf.hpcc.ttu.edu:2119/jobmanager-sge-serial	51	432	9999999			
	testwulf.hpcc.ttu.edu:2119/jobmanager-sge-ttu-users	432	432	9999999			
	testwulf.hpcc.ttu.edu:2119/jobmanager-sge-cms	336	336	9999999			

[Summary | Detail | OSG BDII | Applications]

http://www.math.tamu.edu/osg/Rstatus.php

Looking ahead

- Become a *production* grid for R apps

  - Multiple maintainers of the software repository
  - Deployment by the book
  - Frequent health checks of participating resources
  - We may still need our own portal of accurate resource info
  - A *submit.suragrid.org* host with GSISSH access would be useful
  - Create knowledge base of best practices, sample scripts, new user guides, etc.
- Extend this to other common apps
  - Octave may be easy, BLAST may be difficult

Challenges

- Site deployment
  - Perceived steep learning curve – we may still need “install fests”
  - New PKI should help today; perhaps InCommon Silver down the road
  - Batch schedulers with solid Globus integration is very important
  - Glidein WMS is not universally welcome
  - Internet access from compute nodes is not guaranteed
- Attracting the users
  - Show some success stories
  - Develop knowledge base
  - Keep site admins involved and enthused
- Funding

URLs

- Website: http://www.suragrid.org
- Wiki: http://docs.uabgrid.uab.edu/suragrid
- BDII Status: http://www.math.tamu.edu/osg/sgstatus.php
  - Information in Globus Resource column may not be accurate.
- R Status mock-up: http://www.math.tamu.edu/osg/Rstatus.php

Q?

Appendix: Installation of R Packages

- R packages are stored in VO-specific location on each resource - $R_LIBS_SITE variable $OSG_APP/suragrid/R/2.15.1/libs/
- Owned by sgvoadmin on those resources with this user, otherwise owned by shared suragrid user
- Build and deploy from central site for each OS
  - RHEL5, RHEL6, SLES11
- Initial packages:
  - SuppDists (for running R benchmark)
  - ncdf (for TTU - depends on NetCDF/HDF5)
- R_LIBS_SITE is set in profile to pick these up at runtime

Appendix: User Installed R Packages

- User can install their own packages
- Create a shell script that sources $OSG_APP/suragrid/etc/profile
- The profile sets R_LIBS_USER=$HOME/.R/2.15.1/libs
- R CMD INSTALL pkgname.tar.gz
- This installs pkgname into $R_LIBS_USER
- Fairly easy with Condor-G