HPC Services Report
January 26, 2009: *bcloud, dshealy, jprobinson*

Agenda Overview

1. Briefing
2. Compute Expansion
   a. UAB HPC Update
   b. OSG Exploration
   c. Research data storage
   d. MatLab
3. Special Initiatives
   a. caBIG
      - LabKey/CPAS
      - HSIS & CTSA Engagement
   b. RUST Co-location
4. Research Networking
5. Upcoming Events
6. Other Items
7. Supporting Materials

Agenda Details

1.0 Briefing

2.0 Compute Expansion

Campus Compute Resources

Updates:

2.a UAB HPC Update

January 2009 announces Cheaha availability as a general resource for campus computing. (See attached announcement) This update marks the completion of "Phase 1" development of Cheaha as a shared resource for research computing at UAB.

This update sets the stage to further develop research services through the development of on demand reservation, special purpose appliances, and expansion into cloud services.

This update includes:

- Meta-scheduling with to harness the power of multiple clusters and the ability to build scientific workflows leveraging the Distributed Resource Management Application API (DRMAA) with Java, C, Python, and Perl bindings.

- 10GigE connectivity to a inter-cluster networking back-plane which supports efficient job staging across clusters.
- 192 3.0 GHz Intel and 120 1.6 GHz AMD x86-64 cores in a dedicated 2.7TFlop local compute pool. This pool is available via SGE to facilitate transparent migration and backward compatibility with existing user experiences and applications.
- Transition to a continuous resource improvement model using a development model that is open to community participation.
- Transition to the UABgrid wiki-based documentation service (http://docs.uabgrid.uab.edu)

Next Steps:

- Implement improved data services for research activities using grid technologies
- Develop resource reservation and research support appliance services
- Improve and transition existing documentation to the UABgrid Docs wiki.

Pending:

- ASA grid connectivity update – SC08 conversation with Charles Wright on next steps and need to define "production" options
- Compute expansion for Fluent jobs for Engineering with Alan Shih and Puri (regional centers like LSU and TACC)
- Compute expansion for life science applications work with Puri to leverage TACC Life Sciences outreach.

2.b SURAgrid and OSG Integration

- Compute expansion via OSG and SURAgrid – SC08 and SURAgrid working group updates on Meta-scheduling and OSG

2.c Research Data Storage

- Currently have 6TB data storage (dell 1950's) that has been moved recently from LHL to BEC and plan to configure with parallel file structures to use for HPC scratch space.
- Add more research home directors with option for researchers to use Rust data storage for expansion and regular backups. *This still need to be fleshed out for access and backup.*

2.d MathLab: running on HPC clusters at UAB - running in batch mode on CIS cluster. Puri Bangalore is scheduling visit to campus of MatLab to assist planning expansion to more clusters and users

3.0 Special Initiatives

3.a *caBIG*

3.a.1 LabKey/CPAS

LabKey (http://www.labkey.com/) is the data management system used by Dr. Jim Mobley, Director of Research in the Division of Urology at UAB. LabKey is a web-based, open source application that supports collaborative biomedical research and is used for the collection, analysis storage and distribution of mass spectrometry data derived from tissue samples.

Update:

- Successfully connected Computational Proteomics Analysis System (CPAS) data service to production caGrid infrastructure. (See attached excerpt from UAB Urology site.)

This data service connects the production LabKey (v8.3) instance running in Mobley's lab to caBIG. The lab has direct control over which data is shared via the standard authorization functionality provided by the LabKey management interface. Currently, only animal data is shared to demonstrate functionality while data sharing policies are further developed. UAB IT produced the grid interface using the LabKey provided caBIG API after assisting LabKey with the debug of the previous release.

Link to mention of caBIG access on UAB Urology site:

http://www.urology.uab.edu/urologyresearch/Mobley Lab2_files/page0001.htm

- LabKey effort used to understand operation and generation of caGrid data services in preparation for building additional data services.

Data service tooling is understood and we can use UML+Data Model from an originating application to build data services for caGrid. There is potential to leverage this in a variety of ways including construction of data services in the style of database "reports" (ie. views) in style of Crystal Reports or Jasper Reports.

Next Steps:

- Complete documentation for producing connector to contribute to caBIG/LabKey communities.
- Deploy CPAS data service to production hardware in RUST.
- Facilitate internal dialogs related to research data sharing and security policy.

Opportunities:

- Explore integration of UABgrid compute resources to speed data analysis using LabKey's Pipeline Data Processing interface to Globus.
- Explore potential to develop customized modules in collaboration with Mobley and LabKey that supports caBIG data queries directly from LabKey web interface to facilitate analysis of data collected by other CPAS data stores across caBIG, eg. Fred Hutchison Cancer Center.

3.a.2 HSIS and CTSA Engagement

In the context of caBIG support, UAB IT has been engaging with HSIS and CTSA representatives via regular meetings of the UAB Grid Applications Working Group.

Update:

- General agreement has been reached with HSIS on an architectural approach to data sharing that leverages the caGrid infrastructure to facilitate data sharing needs for on-campus collaboration and engagement in national efforts like caBIG and CTSA.

This solution builds upon the foundations of UABgrid and extends the data exchange architecture used by HSIS. De-Identified patient data can be made available from the Horizon system (the data aggregation point within the hospital information management system) via a caGrid data service that interfaces with the SUN JCAPS data exchange infrastructure of HSIS.

This approach enables combination with other grid data and analysis services facilitating information aggregation from autonomous sources (eg. UAB Urologoy's LabKey/CPAS, UAB HPC resources, and HSIS patient data) and supports automation of established research workflows within UAB as well as

the development of translational research services both on and off campus via caBIG and CTSA.

Next Steps:

- Further demonstrations of architecture and build buy-in with HSIS and UAB IT support communities.
- caBIG Task Force: Inaugural Meeting January 28, 2009. Provide overview of established caBIG services and current activities

Opportunities:

- Engagement with industrial and research partners to build support for advanced research infrastructure

3.b RUST Co-location

3.b.1 CIS

- Bob Cloud prepared quote for CIS to locate the Forensics cluster at Rust in lieu of 4th floor-UBOB. The UBOB floor loading capacity doesn't meet HPC specifications and Facilities has quoted several hundred thousand dollars and many months delay for an up fit of UBOB. *We met Friday, 1/23/09, and this looks like a done deal.They want the 3-year prepay option.*

3.b.2 BioInformatics Resource Center

- Bob Cloud working with Elliot Lefkowitz and Jim Moon to prepare submittal for five-year renewal. They requested UAB campus cyber-infrastructure background information similar to shandouts we used on CI-Day. He's also interested in reviewing IT-Infrastructure pricing for Storage: $/GByte, Colo Space: $/Rack, and Cheaha-HPC: $/cycle. *Quoted Hosting (colo) rate. Still need to validate the proposed $1,000/core/year rate which has not yet been quoted.*

4.0 Research Networking

Update:

- Announcement of Cheaha and Ferrum 10GigE connectivity as part of Cheaha update announcement.

Next Steps:

- Document 10GigE interface connectivity
- Develop reliable bandwidth metrics and performance monitoring
- Connect additional clusters
- Expand 10GigE connectivity on- and off-campus
- Develop 1GigE Desktop Computing pilot
- *Develop Firewall Security Policy for Research Network.*

5.0 Upcoming Events

- SURAgrid+OSG Integration Exploration Kick-Off Meeting
- Jan 28: Inagural meeting of caBIG Task Force in Cancer Center

6.0 Other Items

- Interview with Ying Zhang by JPR and DLS-- need for senior lead in research computing applications support
- Section on Statistical Genetics (SSG) – collaboration on improvements to their workflow scale and performance. Need gather production statistics of new workflow. Goal is publication in 2009.
- Needs to expand IMSL and software from departmental (or ASC systems) to campus wide HPC support

7.0 Supporting Materials

December 8, 2008 Meeting Agenda

UAB HPC Update, January 2009

UAB Urology caBIG connectivity

IT HPCS Current Activities http://docs.uabgrid.uab.edu/index.php/HPC_Services_Activities_FY2009

SURAgrid Strategic Plan