Open Science Grid News

MARCH 2007

In This Issue

Meetings and Events

EGEE User Forum
Manchester, United Kingdom
May 7-11, 2007

TeraGrid '07
Madison, Wisconsin
June 4-7, 2007

HPDC 2007: IEEE International Symposium on High Performance Distributed Computing
Monterey Bay, California, USA
June 27-29, 2007

View Full Calendar

Spotlight on an OSG Contributor

Mats Rynge, of the Renaissance Computing Institute (RENCI) participates in OSG's Engagement activity, to bring in new scientific user communities to the yberinfrastructure. Mats introduces new users to the environment, and helps them get their applications up and running.

“Working on the Engagement activity has been a bigger challenge than I anticipated.” says Mats. “Many of our users are new to the grid, and we often need to spend as much time

From the Executive Director

Image courtesy Ben Tolo of the San Diego Supercomputer Center

A big thank you to Frank, Penelope and all the organizers of the all hands meeting at the San Diego Supercomputer Center – it was great! Certainly not a time for relaxation; discussions and thinking went on through the lunches on the lawn and evening meals, and we came away with more to do and a much clearer direction than when we arrived.

Among the issues discussed were the continuing support for D-Zero reprocessing, and the new LHC Tier-3s that are registering and will be part of the ongoing Operations meeting agendae. We're pleased that we appear to have enough interest for a Biologist Interest Group.

In this issue we include summaries from a few of the activities at the meeting, and we'll include some more next month.

~ Ruth Pordes

OSG 0.6.0 Released

OSG 0.6.0 was released on March 7, 2007 with several new updates that significantly improve that quality of information and the ease of administration. It now includes the Gratia software, which collects detailed accounting information so that both local system administrators and OSG management can understand what is happening. dCache has been added to the VDT to help with data management. A new tool (vdt-control) has been added to simplify the life of system administrators, and makes it very easy to enable or disable components of the VDT. In addition, several new Linux distributions are supported. Of course, the usual raft of bug fixes and tweaks have been applied. Upgrade today and we promise a ten-fold reduction in tooth decay and twenty-fold improvement in your Sudoku abilities! View release notes.

All Hands and Associated Meeting Summaries

Meeting Summary

Enabling User Communities
The user group coordinators found an enthusuastic response among VOs validating the OSG 0.6.0 release. The session elicited input on software stack requirements and expectations, validation, and support for grid operations and processes.

We will continue to monitor VO experience with subreleases over the next few months. We started collecting requirements for the next release, 0.8.0. We expect to solidify resource advertising, and have started soliciting VO input on the current attributes and schema, for the attribute subcommittee.

OSG Security -- Reference Document Available
The security activity, led by Don Petravick, has been developing a role-based security program for OSG. The program is split into "Core" (paid by or contributed to OSG) and "non-Core"(site/VO owned and maintained). Roles are identified within each, and security responsibilites vary by role. The adopted security philosophy ensures that computer security, like safety, is not an arbitrary set of prescriptive rules imposed from the outside, but rather a part and parcel of all OSG activities.

training them as we do getting their applications to run."

Mats and his colleagues have the opportunity to try out new integration techniques and technologies to make life easier for grid novices as well as power users. One challenge is getting OSG to handle parallel code that requires Message Passing Interface. MPI is working at a couple of sites, but usability issues are still being resolved.

Mats' experience participating in the Engagement activity has been positive. "There are difficulties, sure. But we are having a lot of fun, and it is rewarding to help new users achieve their goals." He finds most new users excited to be introduced to this new technology, to have more resources available to them and to be part of a growing international scientific community

As the OSG moves forward, efforts by Mats and others in the Engagement activity are essential for recruiting new researchers who will become involved and increase the productivity of the OSG.

Application: The CMS "Top 100"

Lowering of the 16 meter high CMS endcap disc into the experimental cavern. *Image courtesy CERN.*

It takes audacity to throw away all but a hundred of every 40 million data points. How can you be sure you're saving the precious few that may lead to discovery? For the CMS high-energy physics experiment, the response involves gambling – fortunately, these scientists have a good hand.

Read more

Supported By

Office of Science/ U.S. DOE

Read More

Effectiveness of the OSG Facility
Easy and dependable access by all VOs to all sites is the cornerstone of the OSG vision. Translating this vision into an effective High Throughput facility requires that each and every OSG site actively monitor the quality of the services it offers and address any deficiencies and problems. It is important that each site publish accurate information about its capabilities and services. It is the responsibility of the OSG facility to provide sites with a validation framework and a suite of tools to test and advertize the functionality of CEs, SEs and the WNs.

U.S. ATLAS Tier-2 Workshop
The U.S. ATLAS Tier-2 Workshop agenda centered on moving the Computing Facilities from development and deployment into stable and continuous operation. The goal is to enable ATLAS Physics Analysis in time for U.S. physicists to participate in the worldwide effort in a competitive manner.

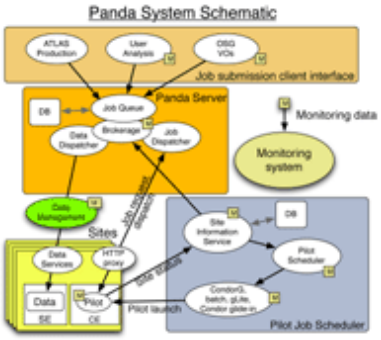
Read More

OSG Council Meeting
Almost all the OSG Council members attended the Council meeting, held on March 8 in conjunction with the All Hands meeting. LIGO, ATLAS, CMS, CDF, D0, STAR, Nanohub, GADU , WLCG and TeraGrid described their current use of and experiences with OSG. These projects reported a high value from the OSG. They all expressed, however, a need for easier access to resources on a short term basis, and that they expect their future requirements to be significant.

Read More

Technical: Managing Workloads with PanDA

Data processing for the LHC ATLAS experiment, which will start taking data late this year, places challenging requirements on throughput, scalability, robustness, efficient utilization of computing resources and operations manpower, and efficient data flow. In response, the US contingent of ATLAS has developed the PanDA distributed production and analysis system.



(Click for larger image)

Development began in August 2005 and PanDA began production four months later. In the last year Panda production on the OSG accounted for 28% of ATLAS production, well in excess of the US quota of 19%. Panda's analysis interface, pathena, is popular throughout ATLAS as a simple means of running analysis on the grid. Panda currently processes up to ~10k ATLAS jobs/day, expected to rise to ~100k/day as ATLAS ramps up.

A few key design elements are worth highlighting here:

- the job submission client is lightweight and easily adapted to new users.
- the job queue provides uniform job management across the grids, facilities and VOs that Panda serves.
- data is automatically pre-placed at processing sites.
- 'just in time' workload delivery via pilot jobs allows Panda jobs to avoid the latency and failure modes of grid submission.
- pilot submission can use a range of scheduling systems (most often CondorG)
- a comprehensive monitoring system flags problems and allows drill-down investigation.

Panda is being generalized for use by any VO, and the first OSG customer, CHARMM, is now in production. Through collaborations with Condor and the CMS glide-in WMS effort, the already successful integration of Condor in Panda is being deepened. PanDA now operates on over 200 computing elements across the OSG and LCG, and Panda-based ATLAS production and analysis is being deployed across these domains.

For more information or to try Panda for your VO, contact Torre Wenaus.

~ Torre Wenaus