Open Science Grid
December 2008
Annual Report to DOE

The Open Science Grid Consortium

Open Science Grid

Miron Livny	University of Wisconsin	Principal Investigator
Ruth Pordes	Fermilab	Co-PI, Executive Director
Kent Blackburn	Caltech	Co-PI
Paul Avery	University of Florida	Co-PI

Table of Contents

1. Organizations

1.1. Partner Organizations

The members of the Council and List of Project Organizations

1. Boston University
2. Brookhaven National Laboratory
3. California Institute of Technology
4. Clemson University
5. Columbia University
6. Cornell University
7. Distributed Organization for Scientific and Academic Research (DOSAR)
8. Fermi National Accelerator Laboratory
9. Indiana University
10. Information Sciences Institute/University of South California
11. Lawrence Berkeley National Laboratory
12. Purdue University
13. Renaissance Computing Institute
14. Stanford Linear Accelerator Center (SLAC)
15. University of California San Diego
16. University of Chicago
17. University of Florida
18. University of Illinois Urbana Champaign/NCSA
19. University of Nebraska – Lincoln
20. University of Wisconsin, Madison

1.2. Participants: Other Collaborators

The OSG relies on external project collaborations to develop the software to be included in the VDT and deployed on OSG. Collaborations are in progress with: Community Driven Improvement of Globus Software (CDIGS), SciDAC-2 Center for Enabling Distributed Petascale Science (CEDPS), Condor, dCache collaboration, Data Intensive Science University Network (DISUN), Energy Sciences Network (ESNet), Internet2, LIGO Physics at the Information Frontier, Fermilab Gratia Accounting, SDM project at LBNL, U.S. LHC software and computing.

2. Activities and Findings:

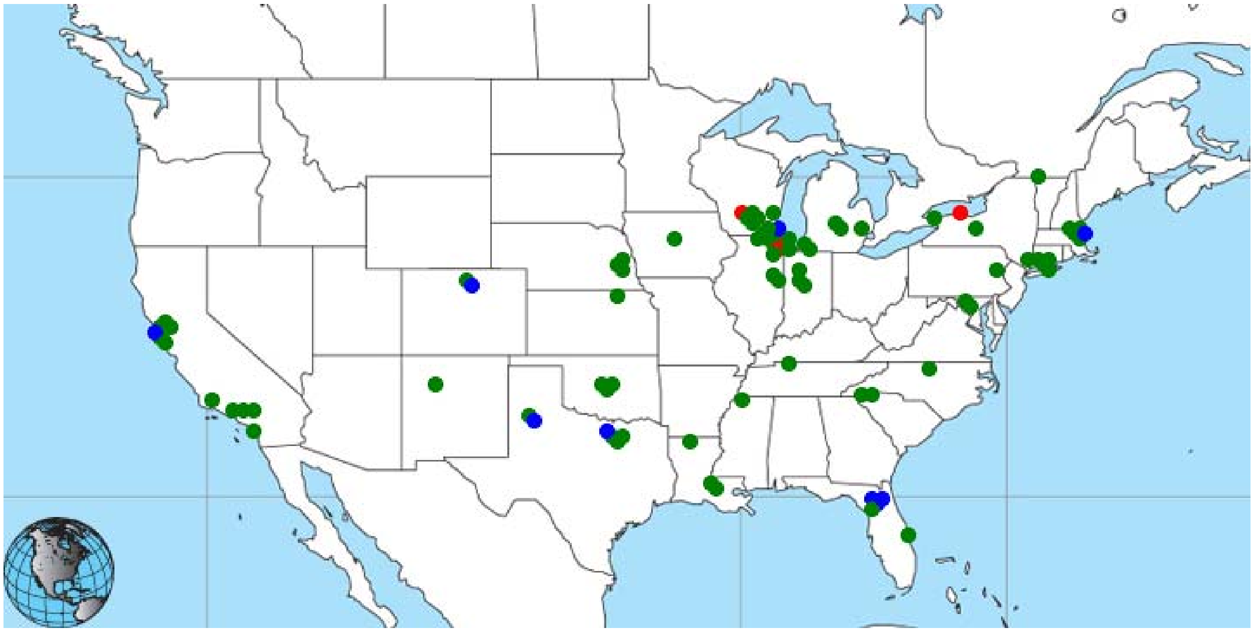

Figure 1: Map showing U.S. sites making up Open Science Grid

2.1. Research and Education Activities

OSG provides an infrastructure that supports a broad scope of scientific research activities, including the major physics collaborations, nanoscience, biological sciences, applied mathematics, engineering, computer science and, through the engagement program, other non-physics research disciplines. The distributed facility is quite heavily used, as described below and in the attached document showing usage charts.

OSG continued to provide a laboratory for research activities that deploy and extend advanced distributed computing technologies in the following areas:

- Integration of the new LIGO Data Grid security infrastructure, based on Kerberos identity and Shibboleth/Grouper authorization, with the existing PKI authorization infrastructure, across the LIGO Data Grid (LDG) and OSG.
- Support of inter-grid gateways which transport information, accounting, service availability information between OSG and European Grids supporting the LHC Experiments (EGEE/WLCG).
- Research on the operation of a scalable heterogeneous cyber-infrastructure in order to improve its effectiveness and throughput. As part of this research we have developed a comprehensive "availability" probe and reporting infrastructure to allow site and grid administrators to quantitatively measure and assess the robustness and availability of the resources and services.
- Scalability and robustness enhancements to Condor technologies. For example, extensions to Condor to support Pilot job submissions have been developed, significantly increasing the job throughput possible on each Gid site.

- Deployment and scaling in the production use of “pilot-job” workload management system – ATLAS PanDA and CMS glideinWMS. These developments were crucial to the experiments meeting their analysis job throughput targets.
- Scalability and robustness enhancements to Globus grid technologies. For example, comprehensive testing of the Globus Web-Service Gram which has resulted in significant coding changes to meet the scaling needs of OSG applications
- Development of an at-scale test stand that provides hardening and regression testing for the many SRM V2.2 compliant releases of the dCache storage software required by the WLCG MOU.
- Integration of BOINC-based applications (LIGO’s Einstein@home) submitted through grid interfaces.
- Further development of a hierarchy of matchmaking services, ReSS or REsource Selection Services, that collect information from more than 60 OSG sites and provide a VO based matchmaking service that can be tailored to particular application needs.
- Investigations and testing of policy and scheduling algorithms to support “opportunistic” use and backfill of resources that are not otherwise being used by their owners, using information services such as GLUE, matchmaking and workflow engines including Pegasus and Swift.
- Comprehensive job accounting across 76 OSG sites, publishing summaries for each VO and Site, and providing a per-job information finding utility for security forensic investigations.

The key components of OSG’s education program are:

- Organization and participation in more than 6 grid schools and workshops, including invited workshops at the PASI meeting in Costa Rice and the first US eHealthGrid conference, and co-sponsorship of the International Grid Summer School in Hungary.
- Active participation in more than 5 “Campus Infrastructure Days (CI Days) events. CI Days is an outreach activity in collaboration with Educause, Internet2, TeraGrid and the MSI institutions. Each event brings together local faculty, educators and IT personnel to learn about their combined needs and to facilitate local planning and activities to meet the cyber-infrastructure needs of the communities.
- Invited participation in the TeraGrid Supercomputing 08 education workshop, participation in the Grace Hopper Conference GHC08 October 1-4, Colorado and Applications of HPC, Grids, and Parallel Computing to Science Education Aug 15, 2008, U of Oklahoma
- Support for student computer science research projects from the University of Chicago, performing FMRI analysis and molecular docking, as well as evaluating the performance and usability of the OSG infrastructure.

2.2. Findings

- Scientists and researchers can successfully use a heterogeneous computing infrastructure with job throughputs of more than 20,000 CPU days per day (an increase of an average of 5,000 CPU days per day over the last six months), dynamically shared by up to ten different research groups, and with job-related data placement needs of the order of Terabytes.

- Initial use of opportunistic storage in conjunction with opportunistic processing provides value and can significantly increase the effectiveness of job throughput and performance.
- Federating the local identity/authorization attributes with the OSG authorization infrastructure is possible. We know there are multiple local identity/authorization implementations and it is useful to have an exemplar of how to integrate with at least one.
- The effort and testing required for inter-grid bridges involves significant costs, both in the initial stages and in continuous testing and upgrading. Ensuring correct, robust end-to-end reporting of information across such bridges remains fragile and human effort intensive.
- Availability and reliability testing, accounting information and their interpretation are proving their worth in maintaining the attention of the site administrators and VO managers. This information is not yet complete. Validation of the information is also incomplete, needs additional attention, and can be effort intensive.
- The scalability and robustness of the infrastructure has reached the performance needed for initial LHC data taking, but not yet reached the scales needed by the LHC when it reaches stable operations. The goals for the commissioning phase in FY09 have been met and are only now being sustained over sufficiently long periods.
- The job "pull" architecture does indeed give better performance and management than the "push" architecture.
- Automated site selection capabilities are proving their worth when used. However they are inadequately deployed. They are also embryonic in the capabilities needed – especially when faced with the plethora of errors and faults that are encountered on a loosely coupled set of independent computing and storage resources used by a heterogeneous mix of applications with greatly varying I/O, CPU and data requirements.
- Analysis of accounting and monitoring information is a key need which requires dedicated and experienced effort.
- Transitioning students from the classroom to be users is possible but continues as a challenge, partially limited by the effort OSG can dedicate to this activity.
- Many communities are facing the same challenges as OSG in educating new entrants to get over the threshold of understanding and benefiting from distributed computing.

2.2.1. Findings enabled by the Distributed Infrastructure: Science Deliverables

Physical Sciences:

CMS: US-CMS relies on Open Science Grid for critical computing infrastructure, operations, and security services. These contributions have allowed US-CMS to focus experiment resources on being prepared for analysis and data processing, by saving effort in areas provided by OSG. OSG provides a common set of computing infrastructure on top of which CMS, with development effort from the US, has been able to build a reliable processing and analysis framework that runs on the Tier-1 facility at Fermilab, the project supported Tier-2 university computing centers, and opportunistic Tier-3 centers at universities. There are currently 13 Tier-3 centers registered with the CMS datagrid in the US which provide additional simulation and analysis resources to the US community. In addition to common interfaces, OSG has provided the packaging, configuration, and support of the storage services. Since the beginning of OSG the operations of storage at the Tier-2 centers have improved steadily in reliability and performance. OSG is playing a crucial role here for CMS in that it operates a clearinghouse and point of contact between the sites that deploy and operate this technology and the developers. In addition, OSG fills in gaps left open by the developers in areas of integration, testing, and tools to ease operations. The stability of the computing infrastructure has not only benefitted CMS. CMS' use of resources has been very much cyclical so far, thus allowing for significant use of the resources by other scientific communities. OSG is an important partner in Education and Outreach, and in maximizing the impact of the investment in computing resources for CMS and other scientific communities.

In addition to computing infrastructure OSG plays an important role in US-CMS operations and security. OSG has been crucial to ensure US interests are addressed in the WLCG. The US is a large fraction of the collaboration both in terms of participants and capacity, but a small fraction of the sites that make-up WLCG. OSG is able to provide a common infrastructure for operations including support tickets, accounting, availability monitoring, interoperability and documentation. As CMS has entered the operations phase, the need for sustainable security models and regular accounting of available and used resources has become more important. The common accounting and security infrastructure and the personnel provided by OSG is a significant service to the experiment.

ATLAS: US ATLAS continues to depend crucially on the OSG infrastructure. All our facilities deploy the OSG software stack as the base upon which we install the ATLAS software system. The OSG has been helpful in improving usability of the grid as seen by US ATLAS production and analysis, and mitigating problems with grid middleware. Examples include

- GRAM dependency in CondorG submission of pilots, limiting the scalability of PanDA pilot submission on the grid. The OSG WMS program has developed a 'pilot factory' to work around this by doing site-local pilot submission without every pilot seeing the gatekeeper and GRAM.
- glexec for analysis user tracing and identity management, now deployed for production by FNAL/CMS and planned for EGEE deployment soon. US ATLAS will benefit from its addition to the OSG software stack, and has benefitted from OSG WMS support in integrating glexec with PanDA.
- OSG-standard site configuration, providing a 'known' environment on OSG WNs. This has lessened the application-level work of establishing homogeneity.
- Tools for resource discovery. We use OSG tools to gather the information on resource availability, health, and access rights that is required to fully utilize the resources available.

In the storage area, US ATLAS Tier-2 sites running dCache are benefiting from the support provided by the OSG Storage Group for code integration and operational help. At BNL (ATLAS and STAR), and at some Tier-2 sites we have Xrootd deployed. The integration and support effort that has started as part of the Storage Group's program is much appreciated. We have also benefited from having OSG help to integrate the SRM 2.2 storage interface component into the middleware stack.

We greatly benefit from OSG's Gratia accounting services, as well as the information services and probes that provide OSG usage and site information to the application layer and to the

WLCG for review of compliance with MOU agreements. We rely on the VDT and OSG packaging, installation, and configuration processes that lead to a well-documented and easily deployable OSG software stack, and OSG's validation processes that accompany incorporation of new services into the VDT. This year's additions requested by ATLAS included the LCG-Utils package and client libraries and API bindings for the LFC file catalog service which US ATLAS has just adopted. US ATLAS and ATLAS operations increasingly make use of the OSG trouble ticketing system (which distributes tickets originating from OSG and EGEE to the US ATLAS RT tracking system) and the OSG OIM system which communicates downtimes of US ATLAS resources to WLCG and International ATLAS. We also benefit from and rely on the infrastructure maintenance aspects of the OSG such as the GOC that keep the virtual US ATLAS computing facility and the OSG facility as a whole operational. We have PanDA running on ALL OSG sites (not just ATLAS) and are poised to take advantage of opportunistic resources on the OSG to do further Monte Carlo simulation and distributed analysis for ATLAS.

The US-developed PanDA distributed production and analysis system based on just-in-time (pilot based) workflow management is now in use ATLAS-wide, and is (since 2006) a part of the OSG's workload management effort as well. Both ATLAS and OSG have benefited from this activity. The OSG WMS effort has been the principal driver for improving the security of the PanDA system, in particular its pilot job system, bringing it into compliance with security policies within the OSG and WLCG. OSG WMS effort has also deepened the integration of PanDA with the Condor job management system, which lies at the foundation of PanDA's pilot submission infrastructure. For the OSG, PanDA has been deployed as a tool and service available for general OSG use. This year a team of biologists used PanDA and OSG facilities for the protein folding simulation studies (using the CHARMM simulation code) underpinning a research paper accepted for publication. We are now working to increase the scope of PanDA's offerings to the OSG community. Enhancements will include mechanisms by which VOs and the OSG itself can monitor and control resource usage; record, track and address failure modes at the application and facility levels; and gather and publish VO/site specific attributes that guide workload brokerage for VO applications. Reciprocally the OSG WMS effort will continue to be the principal source for PanDA security enhancements, further integration with middleware and particularly Condor, and scalability/stress testing of current components and new middleware integration.

LIGO: LIGO has advanced its capabilities to deliver large workflows for the binary inspiral searches. These workflows, consisting of tens of thousands of individual jobs, are necessary to analyze LIGO dataset in the search for gravitational waves from the inspiral of compact binary stellar objects such as neutron stars and black holes. Working closely with the developers of the Pegasus Workflow Planner, LIGO scientists, post-docs and technical staff have also been able to develop common abstract workflow generation tools, which can be used to submit to either the LIGO Data Grid or the Open Science Grid. This effort continues to identified further technological advances needed in the areas of data storage and data management which will allow even greater opportunities for scientific deliverables associated with the data inspiral analysis in the coming year.

In addition, LIGO identified the Einstein@Home application late in 2008 as an excellent candidate application to port onto the Open Science Grid based on the previous experiences with the migration of the Einstein@Home BOINC application onto the German D-Grid. This was particularly appealing as the Einstein@Home Project has an established policy regarding acceptance of scientific results from computational resources not under the management of the LIGO Scientific

Collaboration. A science value bookkeeping system based on a computational unit called a "credit" (one CPU hour ~ 10 credits) is provided by the Einstein@Home Project. This credit system allows LIGO and others to account for the meaningful scientific results collected from computational resources being employed. Beginning in November 2008, use of the OSG for Einstein@Home jobs was shifted to a dedicated user account under the credit accounting system, allowing for direct measurement of OSG's contributions to the search for this important gravitational wave source. Within a week of running under the dedicated user account, the OSG was in the top 10 daily contributors to science credits to the Einstein@Home analysis worldwide. Contributions from the OSG had been limited to OSG sites supporting WS-Gram. In early December 2008 changes were made to the code base to allow running on Globus Toolkit 2 Gram sites on the OSG. This new code base is currently being tested and is expected to be ready for production quality running in January 2009, increasing the contribution being made by OSG to this search for gravitational waves from periodic sources associated with spinning neutron stars.

D0 at Tevatron: The D0 experiment continues to heavily use OSG infrastructure and resources in order to achieve the computing demands of the experiment. The D0 experiment has used OSG resources for over two years and plans on continuing this very successful relationship into the future.

All of the Monte Carlo simulation that is required for the experiment is generated at remote sites and OSG continues to be a major contributor in producing Monte Carlo events. Over the past year, approximately 4.5 million events/week were being generated by OSG sites. In an effort to increase the number of events being produced, the production inefficiency of sites was studied. It was determined that sites that did not have local storage elements were less efficient. A request to OSG was made to have sites that D0 uses to implement local storage elements. D0 also worked with Fermilab computing to improve the infrastructure on the experiment's side. After the improvements made by Fermilab computing and the implementation of additional storage elements by OSG sites, the efficiency of Monte Carlo production greatly increased. Recently, D0 had the highest efficiency of Monte Carlo production ever.

Over the past several months, the average number of Monte Carlo events produced by OSG has risen by approximately 2 million events/week. In September, over 10 million Monte Carlo events/week were produced by OSG. D0 plans to continue to work with OSG and Fermilab computing to continue to improve the efficiency of Monte Carlo production on OSG sites.

The primary processing of data continues to be run using OSG infrastructure. One of the very important goals of the experiment is to have the primary processing of data keep up with the rate of data collection. It is critical that the processing of data keep up in order for the experiment to quickly find any problems in the data and to keep the experiment from having a backlog of data. D0 is able to keep up with the primary processing of data by reconstructing nearly 6 million events/day. Over the past year D0 has reconstructed over 1.8 billion events using the OSG infrastructure.

Because of OSG resources, D0 is able to meet its computing requirements in both Monte Carlo production and in data processing. This has directly led to D0 publishing a large number of papers in 2008. Over 40 science papers have been published in 2008 from the D0 experiment.

CDF at Tevatron: The CDF experiment continues to use OSG infrastructure and resources in order to provide the collaboration with enough Monte Carlo data to keep a high level of physics results. CDF, in collaboration with OSG, aims to improve the infrastructural tools in the next years to increase the Grid resources usage.

During last six months CDF adopted the pilot-based Workload Management System (glideinWMS) as submission method to OSG sites. The usage of the new infrastructure system provides powerful scalability and makes Grid computing working like in a local batch environment with ability to handle more than 10000 running jobs at a time. This new submission system will become the default one also for jobs running at the Fermilab Tier 1 in a near future.

Currently CDF is collaborating with OSG to make opportunistic usage of disk space. Monte Carlo data produced using resources outside Fermilab are moved to the tape storage in the Feynman center using old, not scalable protocols. The new method consists in writing the output of the job in the closest disk space using SRM tools available in OSG and then organize the movement to Fermilab depending on the network and final storage availability. This will allow to increase the CDF Monte Carlo production efficiency and the cpu resources usage. The implementation of the new tools is in progress and preliminary tests are very encouraging.

When the new Monte Carlo data movement method from worker nodes to Fermilab will be in usage, CDF will be able to use it also to move data from Fermilab to worker nodes and use OSG resources for detector data analysis if necessary.

Thanks to OSG resources and infrastructure CDF has been able to publish about 50 physics papers during this year with more than 10 world's best results.

Nuclear physics: The STAR experiment has continued to use the OSG data movement capabilities between LBNL, BNL and new sites on the OSG at Wayne State University and/or NPI/ASCR in Prague (two fully functional Tier 2 centers) and in tests at the University of Illinois at Chicago. At Prague / Bulovka, data transfers are handled using a BeStMan SRM client interoperating with a *Disk Pool Manager* (DPM) SRM door. Data rates between BNL and Prague, reaching 300 Mb/sec at the moment, are sufficient to sustain the local needs. Local data access in Prague rests on the use of the *STAR Unified Meta-Scheduler* (SUMS), which provides users a common interface for job submission. SUMS offers a transparent submission approach to both Grid and non-Grid resources and remains at the heart of STAR’s strategy to migrate an entire class of jobs to Grid resources. STAR plans to utilize the Prague resources for opportunistic Monte-Carlo event processing in 2009 as well as study and test a multi-site data transfer paradigm coordinating movement of datasets to and from multiple locations (sources) in an optimal manner that is, using a planner taking into account network and site performance. It is noteworthy to mention that analysis of data sets now entirely relies on access to Scalla/Xrootd data aggregation at BNL (since 2006) and DPM/rfio access at Prague (2007/2008). Users make extensive use of SUMS abstraction to seamlessly launch jobs on the respective farms; the same job description works on both farms.

STAR has continued to leverage and help consolidate the BeStMan SRM implementation and have engaged in active discussions, steering and integration of the messaging format from the *Center for Enabling Distributed Petascale Science* (CEDPS) Troubleshooting team, in particular targeting the use for BeStMan client/server troubleshooting for faster error and performance anomaly detection and recovery. Stability and scalability testing have been regularly conducted using BeStMan between BNL and LBNL with increasing beneficial results (downtimes are minimal and to the level of one issue per month with a few hours recovery). BNL’s gatekeepers are also being upgraded and the services will be re-tuned once again upon completion (end of

2008) with the prospect of massive data transfer outside/in from the *Korea Institute of Science and Technology Information* (KISTI), a facility and team now member of the STAR collaboration. KISTI holds potentials for serving as a Tier-1-like center for STAR. Setup and data transfer testing are in progress with a first trial in real data transfer mode planned for early 2009 (during RHIC Run-9). Access to job submission at KISTI will likely make use of standard Grid technologies.

STAR grid data processing and job handling operations have continued their progression toward a full Grid-based operation relying on the OSG software stack and the OSG Operation Center issue tracker. The STAR operation team and support has been efficiently addressing issues and stability overall of the grid infrastructure seemed have increased. To date, STAR has however mainly achieved simulated data production on Grid resources. Since reaching a milestone in 2007, it has become routine to utilize non-STAR dedicated resources from the OSG for the Monte-Carlo event generation pass and to run the full response simulator chain (requiring the whole STAR framework installed) on STAR's dedicated resources. STAR's primary dedicated resources in the US are at BNL's RACF and NERSC's PDSF. These and other resources are specifically allocated to and secured for STAR and are aggregated into a virtual facility via the use of the OSG interfaces and the continuous use of SUMS for job scheduling. In the case of jobs submitted to PDSF from BNL, job output is returned to BNL using BeStMan SRM for data transfers.

The relative proportions of processing contributions are at the moment at the level of 10-15% for Monte-Carlo event generation (and mainly pre-allocated use of resources from Fermi-Grid) for non-dedicated resources and 85-90% on STAR dedicated sites for full signal reconstruction. This proportion is explainable by the fact that the complete STAR software stack and environment, which is difficult to impossible to recreate on arbitrary grid resources, is necessary for the full event reconstruction and hence, access to generic and opportunistic resources are simply unpractical and not matching the realities and needs of running experiments in Physics production mode. Those like STAR have as requirements a reproducibility of all results ensured by thoroughly validating the software stack for each variant of OS and platforms via regression testing. The resources needed for running all simulated data productions (both contributions of STAR and non-STAR dedicated resources) is equivalent to 10% of a one pass real-data production and has, since 2008, been entirely handled on a Grid based operation.

STAR has also worked closely, to the extent possible, with the CEDPS Virtualization activity, seeking the benefits of truly opportunistic use of resources by creating a complete pre-packaged environment (with a validated software stack) in which jobs will run. Such approach would allow STAR to run any one of its job workflow (event generation, simulated data reconstruction, embedding, real event reconstruction and even user analysis) while respecting STAR's policies of reproducibility implemented as complete software stack validation. The multitude of combination and the fast dynamic of changes (OS upgrade and patches) make the reach of the diverse resources available on the OSG, workforce constraining and economically un-viable. While this work has potential for dramatic positive impacts for STAR (and other VOs), our progresses were hampered due to the fact that while STAR has the knowledge and will toward the innovative, it also has limited workforce availability (none of which is OSG funded and all of which has as primary mandate to support the immediate need of RHIC operations) and hence, has not been able to sustain significant effort for this activity.

Finally, since all of STAR's physics results are built on a foundation of simulated data, all STAR's physics publications acknowledge the resources provided by the OSG.

Astrophysics: SDSS and the Dark Energy Survey (DES) together used approximately 40,000 hours of OSG resources in 2008, with DES simulation activities ramping up in the latter part of the year. The most recent DES simulation produced 3.34 terabytes of data. Along with 2,600 mock telescope images and 240 images of known stars, it produced 500 calibration images that help scientists identify unwanted distortions due to the atmosphere and the telescope optics. DES processed 15,000 simulations in their largest run, where a single job simulates one 3-degree-square pointing of the DES camera.

Multi-Disciplinary Sciences: The Engagement team has worked directly with researchers in the areas of: biochemistry (Xu), molecular replacement (PRAGMA), molecular simulation (Schultz), genetics (Wilhelmsen), information retrieval (Blake), economics, mathematical finance (Buttimer), computer science (Feng), industrial engineering (Kurz), and weather modeling (Etherton).

The computational biology team led by Jinbo Xu of the Toyota Technological Institute at Chicago uses the OSG for production simulations on an ongoing basis. Their protein prediction software, RAPTOR, is likely to be one of the top three such programs worldwide.

A chemist from the NYSGrid VO using several thousand CPU hours a day sustained as part of the modeling of virual coefficients of water. During the past six months a collaborative task force between the Structural Biology Grid (computation group at Harvard) and OSG has resulted in porting of their applications to run across multiple sites on the OSG. They are planning to publish science based on production runs over the past few months.

Computer Science Research: A collaboration between OSG extensions program, the Condor project, US ATLAS and US CMS is using the OSG to test new workload and job management scenarios which provide "just-in-time" scheduling across the OSG sites using "glide-in" methods to schedule a pilot job locally at a site which then requests user jobs for execution as and when resources are available. This includes use of the "GLexec" component, which the pilot jobs use to provide the site with the identity of the end user of a scheduled executable.

2.2.2. Findings of the Distributed Infrastructure: The OSG Facility

OSG Facility: The facility provides operational, security, software, integration, and engagement capabilities and support. The facility continues to evolve with improvements to operational processes, more robust software platforms and tools, new resource providers, and better management of storage resources. In the last year, the focus has been to provide a stable production grade environment; to a great extent this was driven by planning for the LHC turn-on and enabling CMS and ATLAS to be ready for data taking. This included: (1) improvements to RSV for reporting site availability; (2) improvements to Gratia for accounting data to WLCG; (3) improvements to GIP to enable better matchmaking; and, (4) new tools needed by ATLAS and CMS, specifically lcg-utils (for data management) and LFC (for data replication tracking). In addition, the OSG management team increased focus and velocity on identifying and resolving production related issues; we improved project communication via periodic meetings and assignment of dedicated staff to interface to key stakeholders. The OSG platform was successfully exercised via the various data challenges of ATLAS and CMS in preparation for the LHC startup.

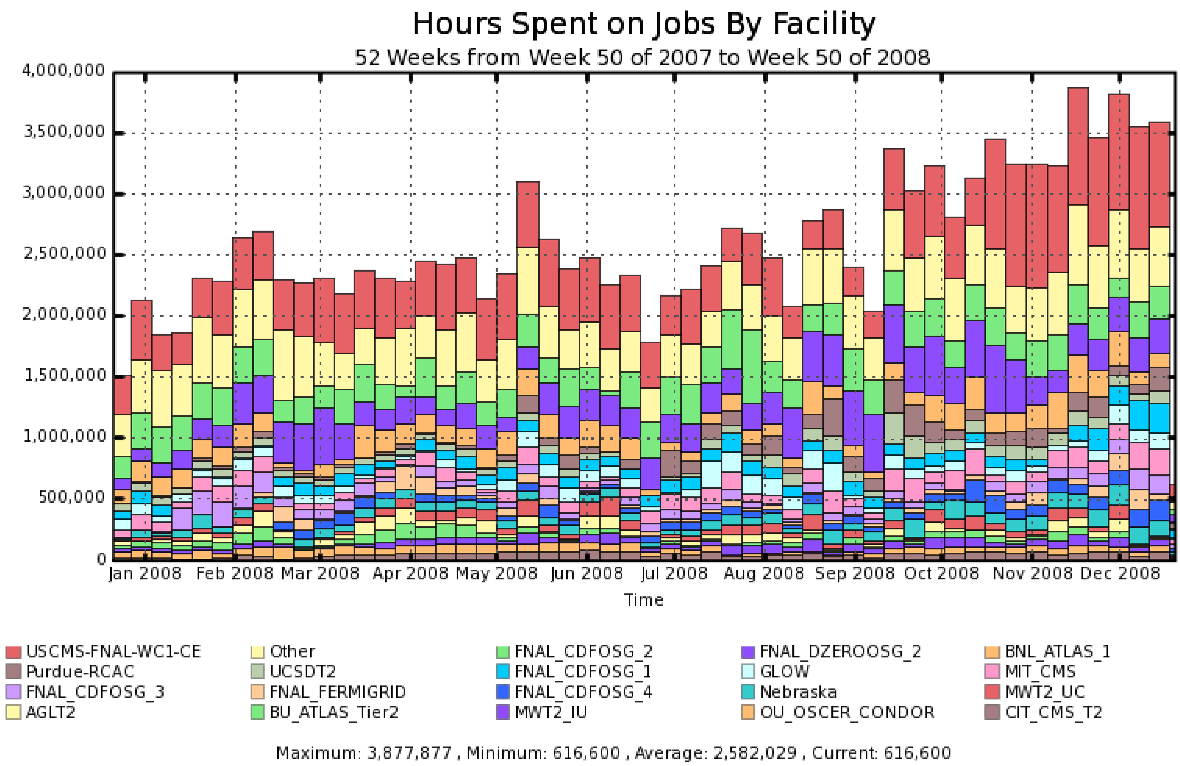

Figure 2: OSG facility usage vs. time broken down by site

The usage of the facility continues to grow; the usage varies depending on the needs of the stakeholders and during stable normal operations is providing over 500,000 CPU wallclock hours a day with peaks occasionally exceeding 600,000 CPU wallclock hours a day with approximately 100,000 to 150,000 opportunistic wallclock hours available on a daily basis for resource sharing. The success rate is variable and hard to measure and depends as much on the user end-to-end success and failure measurements as on those at the level of the infrastructure; currently the infrastructure reports success rates to be over 90%. Figure 2 (usage by week) and Figure 3 (cumulative) document the steady growth in OSG usage and the contributions of many individual sites to the overall OSG facility during 2008.

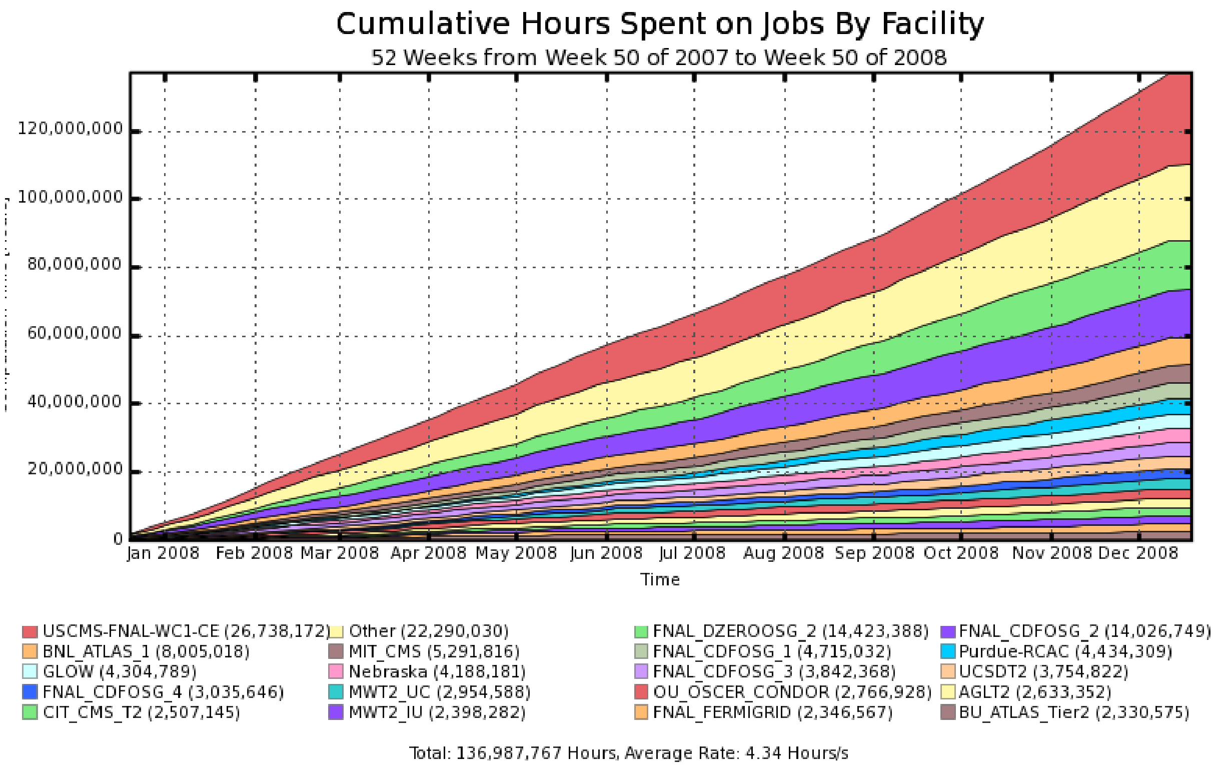

Figure 3: Cumulative CPU usage of the OSG facility by site

Middleware: Early in 2008, OSG's software efforts focused on two areas: supporting the existing OSG 0.8 release, and developing OSG 1.0. Like all software distributions, significant effort must be given to ongoing support.

OSG 0.8 was updated ten times in 2008. Four of those updates were for security updates, while others fixed various bugs that affected OSG users. While the details of each of the updates aren't remarkable for this report, they were all important updates for the continued smooth operation of OSG sites and VOs.

In early 2008, OSG 1.0 was developed. True to OSG's principle of incremental shifts, it was an evolutionary (not revolutionary) improvement from OSG 0.8. Several improvements were made, including:

- A significant update to dCache (version 1.8) with better performance and support for opportunistic storage. CMS was the primary beneficiary. In addition, D0 and CDF both started benefiting from opportunistic storage at OSG sites that deploy dCache.
- BestMan was updated, and the first LHC Tier-3 sites are now benefiting from this storage software. BestMan is viewed as an alternative to dCache, primarily for smaller sites.
- The addition of lcg-utils, which are client utilities for storage management. They are used by CMS and ATLAS.
- Support for new platforms, particularly Red Hat Enterprise Linux 5 (used by several VOs, particularly LIGO) and Debian 4 (used by LIGO).

- Many other smaller, but significant improvements that improved security and reliability for OSG sites and users.

In OSG 1.0, there has been an emphasis on being able to provide incremental updates. In past releases, it has been difficult to deliver significant updates to the OSG software stack without requiring users to fully reinstall and reconfigure the software from scratch. Since VDT 1.10.1 (the basis for OSG 1.0) was released in May 2008, there have been fourteen updates to the software stack. Some of them have been significant updates, such as the addition of the LFC software (requested by ATLAS). Under development right now is a major (but incremental) update to provide significant improvements to the BestMan/XRootd storage management software, also requested by ATLAS.

We have found some difficulties with incremental updates, and work is underway right now (December 2008) to make the process significantly easier for users, and we hope this will be made available to users in the first quarter of 2009.

Operations: OSG Operations provides a central point for operational support. The Grid Operation Center (GOC) performs real time monitoring of OSG resources, supports users, developers and system administrators, maintains critical information services, provides incident response, and acts as a communication hub. Goals of the OSG Operations group are the preservation and strengthening the autonomy of OSG resources, building operational relationships with peering grids, providing reliable grid infrastructure services, ensuring timely action and tracking of operational issues, and quick response to security incidents. In year 3, the GOC continued to provide the OSG with a reliable facility infrastructure while at the same time improving services to offer more robust tools to the users of the OSG.

The GOC continues to provide numerous services to the OSG. The OSG Information Management (OIM) database provides the definitive source for listing OSG entities at the person, resource, support agency, or virtual organization level in year 2; this system has been used to provide operations automation, simplifying and reducing some time consuming administrative tasks. Effort in the operations automation areas will allow us to better handle the need of the stakeholders during the LHC data-taking. The Resource and Service Validation (RSV) monitoring tool that was released and put into production during year two has undergone refining based on the needs of the ATLAS and CMS VOs, and the WLCG. This includes expanding the probe set, friendlier configuration and update tools, and a centralized results display. At the request of CMS, the GOC has moved to implement redundant BDII (Berkeley Database Information Index) servers in Bloomington and Indianapolis. This provides redundancy for the critical information service vital to many collaborators and allows load balancing during heavy usage and failover. Several other hardware and service upgrades have taken place during year 2: the TWiki environment used for collaborative documentation was updated with new functionality and with security fixes, the BDII was updated to improve performance, power and networking was updated in the racks holding OSG services, migration to a common RHEL 5.1 environment for all services is underway, and CA Certificates were moved to the GOC from the VDT.

The year 3 focuses for OSG Operations are LHC support, service reliability, and operations automation. We are actively preparing for the stress of the LHC turn up on services by testing and putting proper failover and load-balancing mechanisms in place and implementing administrative ticketing automation. Service reliability for GOC services has always be very high, during year 3 we will begin gathering metrics that can show the reliability of these services and implement more redundancy mechanisms to assure even higher availability of compute services. Operations automation is very important if the GOC will be scalable into the future, more research into the best ways to allow process automation and problem alerts will allow us to move into the future assured operations manpower can keep up with the growth of OSG.

Integration and Site Coordination: The mission of OSG integration is to improve the quality of grid software releases deployed on the OSG and enable greater success by the sites in achieving effective production.

In the last year, the Integration effort delivered high quality software packages to our stakeholders resulting in smooth implementation of the OSG 0.8.0 and OSG 1.0 releases; several process innovations were key to these results. During the release transition of OSG 0.6 to 0.8, and from OSG 0.8 to OSG 1.0, several iterations of the Validation Test Bed (VTB) were made using a 3-site test bed which permitted quick testing of pre-release VDT updates, functional tests, and install and configuration scripts. The ITB was deployed on 12 sites providing compute elements and four sites providing storage elements (dCache and BestMan packages implementing SRM v1.1 and v2.2 protocols); 36 validation processes were defined across these compute and storage resources in readiness for the production release. Pre-deployment validation of applications from 12 VOs were coordinated with the OSG Users group. Other accomplishments include both dCache and SRM-Bestman storage element testing on the ITB; delivery of a new site configuration tool; delivery of a facility-wide logfile collection infrastructure for troubleshooting; a program of client/server tests for WS-Gram with functional and scalability measurements for typical site configurations.

The OSG Release Documentation was significantly upgraded; this collection of wiki-based documents capture processes that enable install, configure, validation methods used throughout the integration and deployment processes. These documents were updated and received review input from all corners of the OSG community (33 members participated for the OSG 1.0 release) resulting in a higher quality output. A new initiative is being launched for 2009 to align site administrator's documentation with other groups in OSG to promote re-use and consistency.

The community of resource providers comprising the OSG Facility is diverse in terms of the scale of computing resources in operation, research mission, organizational affiliation, and technical expertise, leading to a wide range of operational performance. The Sites Coordination activity held two face-to-face workshops, one of these hands-on covering several technical areas for both new and advanced OSG administrators.

Virtual Organizations Group: A key objective in OSG is to enable Science communities to *produce* Science using the OSG Facility. The Virtual Organizations Group (VO Group) provides direct support to the various OSG science communities in this mission. The VO Group collaborates with the staff of each VO to address requirements, feedback, issues and roadmaps for production-scale operations of our at-large (i.e. all VOs except CMS, ATLAS, and LIGO) Science community.

Broad goals of this team are to: (a) Improve efficiency and utilization of OSG Facility; (b) Provide an avenue for operational, organizational, and scientific discussions with each at-large stakeholder; (c) Facilitate broad stakeholder participation in the OSG software engineering lifecycle; (d) Enable tactical methods for sustenance of communities that have a newly formed VO.; and (e) Provide a platform for OSG Storage group to work directly with all stakeholders, and thus to strengthen Data-Grid capabilities of OSG.

Some of the major work items in the last year were:

- The Science communities were canvassed for inputs and their requirements were compiled for OSG Release1.0, with 14 VOs outlining needs and expectations.
- The *Science Validation* of sites on Integration Testbed (ITB) was completed for OSG Release1.0. In partnership with OSG Integration, a rigorous OSG process has been designed and is regularly executed for each software release to assure high quality. Each participating Science stakeholder exercises a stake in vetting, suggesting changes, and signaling an official approval toward readiness of a major OSG release. In recent ITB validation, 12 VOs participated, 7 VOs ran real Science applications, 6 VOs participated in storage validation, of which, 4 VOs conducted introductory validation of opportunistic storage. In terms of process execution, this was a coalition of 36+ experts, 20+ from VO communities. After meticulous validation and feedback, official 'green flags' toward OSG 1.0 were given by ATLAS, CDF, CIGI, CMS, DES, DOSAR, Dzero, Engage VO, Fermilab VO, LIGO, nanoHUB, SBGrid, SDSS.
- *Joint Taskforces* were structured and executed for ALICE, D0, and SBGrid. These were planned to enable wide-range technical and procedural matters by rallying concerted effort from VOs, external partners, as well as from OSG groups. The ALICE-OSG Taskforce evaluated LHC AliEn grid paradigm for possible future deployment on OSG. The D0-OSG Taskforce led to a significant improvement in D0's procedures, D0's grid infrastructure, and in overall D0 Monte-Carlo event production on OSG. The SBGrid-OSG Taskforce worked closely together to enable SBGrid resource infrastructure and to evolve design and implementation of the SBGrid Molecular Replacement science application, to strengthen each of these to a production-ready level. Recently started, the CDF-OSG Taskforce is targeting similar goals as D0's, and nanoHUB-OSG Taskforce is investigating possible ways to make workflow improvements to increase nanoHUB job volume on OSG.
- Production-scale *Opportunistic Storage* provisioning and usage was initiated on OSG. With OSG Storage group, a technological model was designed and enabled on select storage sites of CMS and ATLAS, followed by its sustained active usage by D0.
- The Annual OSG Users meeting was organized at BNL in June 2008, with emphasis on VO security and policy. In addition, *Weekly VO Forum* teleconferences were organized on a regular basis. The scope was highly-focused technical discussions with each VO, selectively inviting 3-4 VOs and relevant OSG groups every week. These have resulted in strong functional associations of OSG groups with each of CDF, D0, DES, Engage, Fermilab VO, FermiGrid, ILC, nanoHUB, NYSGrid, OSG Education, SBGrid, and STAR. *Bi-monthly VO Forum* teleconferences were started with scope of a 'stakeholder virtual round table' and wide attendance, providing opportunity for technical discussions and concerns to be brought forth by each VO. The first such virtual gathering was organized in October 2008, with active representation from 16 VOs and 10 OSG groups.

The VO Group continues to provide bidirectional channels between Science communities and all facets of the OSG, to assure that the *needs and expectations of Science* communities are absorbed and translated into work activities in OSG to strengthen grid-operations of all stakeholders.

Engagement: A major priority of Open Science Grid is helping new user communities benefit from the infrastructure we are putting in place by working closely with these communities over periods of several months. The Engagement activity brings the power of the OSG infrastructure to scientists and educators beyond high energy physics and uses the experiences gained from working with new communities to drive requirements for the natural evolution of OSG. To meet these goals, engagement helps in: providing an understanding of how to use the distributed infrastructure; adapting applications to run effectively on OSG sites; engaging the deployment of community owned distributed infrastructures; working with the OSG Facility to ensure the needs of the new community are met; providing common tools and services in support of the engagement communities; and working directly with and in support of the new end users with the goal to have them transition to be full contributing members of the OSG.

During this program year, the Engagement team has successfully brought three lead researchers and their teams into full production use of the Open Science Grid (Xu, Blake, Wilhelmsen). These users represent, respectively, the science domains of Biochemistry, Information and Library Science, and Genetics. A Research Highlight has been published for one of these groups, and there is a commitment from the other two for Fall of 2008; and OSG was cited by the winner of the SC07 Storage Challenge as a result of Engagement efforts. There are roughly ten emerging users we are engaged with at various stages that we are working to convert into production users.

Additional activities of the Engagement team during the reporting period include: hosting the 2008 OSG annual meeting; convening a computational Biology workshop at the 2008 OSG annual meeting; representing OSG at CI-Days events; participating in the Campus Champions program in partnership with TeraGrid and CI-Days; contributing the match making technology used to assist new users as an open source project on SourceForge; assisting PRAGMA with OSG usage; demonstrated OSG/EGEE interoperability via the WISDOM application.

In April 2008, the Engagement team was awarded a grant from the NSF CI-TEAM proposal to continue these efforts for an additional three years. Current initiatives include scaling up our efforts to more effectively manage the Engagement process, working more closely with the OSG EOT and Users group efforts, communicating our activities to the broader OSG community, actively recruiting a more diverse set of OSG users, and helping new users implement systems that leverage OSG.

Campus Grids: The Campus Grids team's goal is to include most US universities in the national infrastructures. By helping universities understand the value of campus grids and resource sharing through the OSG national framework, this initiative aims at democratizing cyberinfrastructures by providing all resources to users and doing so in a collaborative manner.

In the last year this team has enabled several campuses to start the process of joining the OSG. Following the NYSGrid CI Days in December 2007, Rochester Institute of Technology (RIT) has deployed a condor based campus grid of several thousand nodes. The SBgrid project (http://www.sbgrid.org) from Harvard has also become very active in OSG and deployed two new resources; a team of OSG experts from Clemson, RENCI and Wisconsin engaged with Harvard to help finish the configuration of these two sites. Clemson and RENCI staff went to Duke University and helped bring up a new ATLAS Tier-3 site. The New Jersey Higher Education Network (http://www.njedge.net/) contacted us after they concluded that OSG had the framework to help them build their regional cyberinfrastructure; a plan has been setup to enable their users on OSG and add a new resource before the end of summer 2008.

Security: Security team continued its multi-faceted approach to successfully meeting the primary goal of maintaining operational security, developing security policies, acquiring/developing necessary security tools and software, and disseminating security knowledge and awareness.

As part of operational security, we evaluated our findings from the last year's Security Test & Evaluation (ST&E) controls, and modified the controls to make them better suitable for our community. Incident response and containment took a large chunk of our operational activities. We had 7 security incidents, 4 of which are classified as high, 3 as medium, and 1 as low severity. We responded to the incidents on average within 1.4 days. Our initial response consists of classifying attack severity, identification of affected OSG members, and a mitigation and containment plan that is sent to the identified members. We keep an incident open until we receive confirmation from our members on enactment of the mitigation plan. On average, it took us 10 days to close an incident completely. This average includes actions we suggested to our members such as applying patches, disabling credentials, etc. As a result of a high-level incident, we requested the International Grid Trust Federation (IGTF) to form an incident response team and we participate in this team on a voluntary basis.

In total, we have had 25 security tickets, which we closed in an average of 5.2 days. In order to help our members and ourselves, we developed an incident response procedure to follow. We are adding to our training materials. The cookbooks describe how to determine whether a compromise occurred, how to report and investigate an incident.

We identified the key infrastructural needs to maintain operational security and generated a roadmap document, based on our users' requests. We acquired two security tools. The first tool is a unified interface for managing certificates. The second is a banning tool for denying access at computing resources. Both tools have been presented at site admins meeting and got very positive feedback. For security monitoring purposes, security team developed security-oriented RSV probes. These probes help site admins as well as security team by monitoring the security infrastructure. In order to involve our users more with their daily usage of OSG, we requested OSG accounting team to send individualized usage records directly to VO managers and users. Inspired from credit-card statements, we hope that OSG users monitoring the jobs submitted under their names would help us detect suspicious behaviors. We received very good feedback from our users.

We are continuing working on getting better communication channels with OSG members. This year, we started regularly attending VO and site meetings. We started working with Education VO to disseminate our training materials to OSG members.

The progress in developing security policies was not as fast as we had hoped. This was due to the different operational views held between OSG and our partner grids. We arranged frequent meetings with OSG executive board to understand and resolve these differences. We started involving our site and VOs early in the policy development process such that they can give us feedback over contentious issues.

Education & Training: Training and outreach to campus organizations, and the development of the next generation of scientist is a core part of the OSG program. The OSG Education and Training program brings domain scientists and computer scientists together to provide a rich training ground for the engagement of students, faculty and researchers in learning the OSG infrastructure, applying it to their discipline, and contributing to its development.

In the last year, the OSG has sponsored and conducted numerous training workshops for students and faculty, and participated in additional invited workshops. This has enabled grid computing training and education for about 300 participants, of which about 30% were from under-represented groups and about 20% were women.

Four major OSG sponsored "Grid School" training events (3 full days each) were conducted in the past 12 months: (1) Florida International University Grid School, January 2008; (2) Tuskegee University Grid School, Feb 2008; (3) Georgetown Grid School, April 2008, and (4) Midwest Grid School, September 2008. To further OSG outreach goals, the FIU and Tuskegee Grid Schools brought leading grid experts and scientific colleagues to two important minority serving institutions. Following each of these workshops, feedback is gathered, tabulated and studied, and follow-on research and engagement assistance opportunities are provided with access to the hands-on curricula and to OSG EOT staff.

In addition to these major training events, OSG staff conducted numerous smaller training and outreach workshops and was a co-organizer of the International Summer School on Grid Computing (ISSGC08) with EGEE which was held in July 2008 in Hungary. The OSG Education team coordinated the selection of students who participated at ISSGC 08, and arranged sponsorship for US-based students to attend this workshop. In addition, OSG staff provides direct contributions to the International Grid School by attending, presenting, lab exercise development, and student engagement. And OSG staff conducted a grid training school at The University of Wittswatersand, Johannesburg, South Africa during July 2008. Another noteworthy event was co-organizing the Second Brazilian LHC Computing workshop in December 2008 held in Sao Paolo, Brazil.

The content of the training material has been enhanced over the past year to include a full module for OSG Grid site administrators, and the end-user training content was updated to include training on SRM/dCache and BestMan storage applications. Work continues to enhance the online delivery of user training and we are seeing an increased number of students and faculty that are signing up for the self-paced online course (about 50 active participants). In addition to individuals, we have made the course accessible as support material for graduate courses on grid computing at universities around the country, such as Rochester Institute of Technology, University of Missouri at St. Louis and Clemson University.

OSG collaborates with Educause, Internet2 and TeraGrid to sponsor day-long workshops local to university campus-wide CyberInfrastructure (CI) Days. These workshops bring expertise to the campuses to foster research and teaching faculty development, IT facility planning, and CIO awareness and dialog. Ongoing dialog and coordination between the EOT programs of TeraGrid, OSG and the Supercomputing conference education programs take place frequently during the year.

Metrics and Measurements: The metrics and measurements activity aims to give the OSG and external agencies a view of the OSG's progress. The formal reporting activity in this area was a definition of the Year 2 metrics (OSG DocDB #740-v3, at http://osg-docdb.opensciencegrid.org/0008/000811/001/Year%202%20Final%20Metrics-2.pdf), a mid-year report focusing on the WLCG VOs (OSG DocDB #769), and a final report covering more OSG-specific metrics (to be posted). These documents follow the progress on 16 high-level metrics throughout the year. Of these high level metrics, OSG was able to make measurements on all the defined goals and was able to achieve the goal value on 9 of them. Most of the goals that were not achieved were mostly due to the measurements being very close (but not quite achieving) or goals which were not well defined. We plan to use the experience in defining and measuring metrics during Year 2 to set out improved goals during Year 3.

In addition to formal reports, there are continuous and monthly activities supported by the metrics and measurements activity. The area maintains a repository of historical data from the OSG information services; this allows the OSG to track the number of active running jobs throughout the year, as well as track the number of cores deployed and the types of CPU on the distributed facility. Starting in the spring of 2008, the area has been generating monthly "metrics thumbnail" that show high-level LHC VO activity on the OSG. These graphs, along with others, are featured on the opensciencegrid.org homepage. Starting in late summer, this area has also generated monthly reports to the OSG JOT that cover availability and CPU accounting statistics for WLCG sites on the OSG.

During the year, several development tasks were done in this area. Upon request from the OSG executive board, software for generating RSV-based reports was written. The top-level package used to generate graphs is also maintained; it indeed has found a niche in the HEP area, and is reused by CMS, ATLAS, LHCb, and the ARDA Dashboard project. Several webpages have been created to serve specific OSG areas; one is designed to be a well-refined interface which is ready for being viewed by the JOT. Others serve specific stakeholder VO needs, such as being able to view the current month's WLCG reports before they are released. Finally, plots are designed upon request in order to investigate OSG performance characteristics.

2.2.3. Findings of the Distributed Infrastructure: Extending Science Applications

In addition to operating a facility, the OSG includes a program of work that extends the support of Science Applications both in terms of the complexity as well as the scale of the applications that can be effectively run on the infrastructure. We solicit input from the scientific user community both as it concerns operational experience with the deployed infrastructure, as well as extensions to the functionality of that infrastructure. We identify limitations, and address those with our stakeholders in the science community. In the last year of work, the high level focus has been threefold: (1) improve the usability and scalability, as well as our understanding thereof; (2) establish and operate a workload management system for OSG operated VOs; and (3) establish the capability to use storage in an opportunistic fashion at sites on OSG.

In 2008, we made an organizational adjustment by moving the "Virtual Organizations Group" from the "Extending Science Applications" area to the "Facility" area in order to reflect better its day-to-day operational focus, while still continuing to coordinate the longer term planning aspects of science application support within the "Extensions" program of work. The new management plan can be found at http://osg-docdb.opensciencegrid.org/0003/000314/004/OSGmanagementPlan-v5-3.pdf.

Scalability, Reliability, and Usability: As the scale of the hardware that is accessible via the OSG increases, we need to continuously guarantee that the performance of the middleware is adequate to meet the demands. There were four major goals in this area for the last year and they were met via a close collaboration between developers, user communities, and OSG.

- At the job submission client level, the goal is 10,000 jobs running simultaneously and 100,000 jobs run per day from a single client installation, and achieving in excess of 95% success rate while doing so. The job submission client goals were met in collaboration with

CMS, CDF, Condor, and DISUN, using glideinWMS. This was done via a mix of controlled environment and large scale challenge operations across the entirety of the WLCG. For the controlled environment tests, we developed an "overlay grid" for large scale testing on top of the production infrastructure. This test infrastructure provides in excess of 10,000 batch slots across a handful of OSG sites. The large scale challenge operations was done in the context of CCRC08, the main LHC computing challenge in May 2008. Here we submitted a typical CMS application to 40 sites distributed worldwide, at a scale of up to 4000 simultaneously running jobs. Two main lessons were learned. First, Condor scalability limitations across large latency networks were discovered. This led to substantial redesign and reimplementation of core condor components, and subsequent successful scalability testing with a CDF client installation in Italy submitting to the CMS server test infrastructure on OSG. Large scale testing under production conditions is expected to start in January 2009. Second, we identified storage failure as the dominant cause of job failure. During CCRC08, we found that glideinWMS successfully isolates the end-user from grid failures, except for storage failures. Success rates in excess of 95% drop down to the 70-80% level due to storage failures on the WLCG production infrastructure. We plan to focus on this issue in 2009 once we have demonstrated that the Condor scalability limitations have been resolved even during realistic operations on the production WLCG grid.

- At the storage scheduling level, the present goal was to have 1Hz file handling rates. For Gbyte file sizes this would translate into close to 10Gbps data transfer capacity. An SRM scalability of 5Hz was achieved in collaboration with the dCache developers, and demonstrated at the CMS Tier-1 center at FNAL. While we clearly exceeded our goals, we also realized that significant further increases are needed in order to cope with the increasing scale of operations by the large LHC VOs Atlas and CMS. This is thus an area that will continue to be high priority in 2009.
- At the functionality level, the 2008 goal was to roll out the capability of opportunistic space use. The roll-out of opportunistic storage was exercised on the OSG ITB preceding OSG v1.0. It has since been deployed at several dCache sites on OSG, and successfully used by D0 for the production operations on OSG. CDF is presently in the testing stage for adapting opportunistic storage into their production operations on OSG.
- OSG has successfully transition to a "bridge model" with regard to WLCG for its information, accounting, and availability assessment systems. This implies that there are aggregation points at the OSG GOC via which all of these systems propagate information about the entirety of OSG to WLCG. For the information system this implies a single point of failure, the bdii at the OSG GOC. If this service fails then all resources on OSG disappear from view. Atlas and CMS have chosen different ways of dealing with this. While Atlas maintains its own copy of the information inside Panda, CMS depends on the WLCG information system. To understand the impact of the CMS choice, OSG has done scalability testing of the bdii. We find that the service is reliable up to a query rate of 10 Hz. The OSG GOC is planning to deploy monitoring of the query rate of the production BDII in early 2009 in response to this finding. The goal is for the GOC to monitor this rate in order to understand the operational risk implied by this single point of failure.

In addition, we have worked on a number of lower priority objectives:

- On WS-GRAM scalability and reliability in collaboration with LIGO, DISUN, CDIGS/Globus, and OSG.
- On testing of a condor client interface to the CREAM compute element (CE) in support of Atlas and CMS. CREAM is a webservices based CE developed by INFN in EGEE. WLCG sites in Europe and Asia are considering replacing their globus GRAM with CREAM on some yet to be determined timescale. Both Atlas and CMS require condor client interfaces to talk to CREAM. We have successfully completed the first of three phases of testing, and are now waiting for CREAM developers to finalize their proxy delegation interface reimplementation before we resume testing in early 2009.
- In the area of usability, an “operations toolkit” for dCache was started. The intent was to provide a “clearing house” of operations tools that have been developed at experienced dCache installations, and derive from that experience a set of tools for all dCache installations supported by OSG. We expect this to significantly decrease the cost of operations, and lower the threshold of entry. Site administrators from both the US and Europe have uploaded tools, and a first release was derived from that. This release has been downloaded by a number of sites, and is in regular use across the US, as well as some European sites.
- Work has started on putting together a set of procedures that would allow us to automate scalability and robustness tests of a Compute Element. The intent is to be able to quickly “certify” the performance characteristics of new middleware, a new site, or deployment on new hardware. Once we have such procedures, we can then offer this as a service to our resource providers so that they can assess the performance of their deployed or soon to be deployed infrastructure.

Workload Management System: The primary goal of the OSG Workload Management System (WMS) effort is to build, integrate, test and support operation of a flexible set of software tools and services for efficient and secure distribution of workload among OSG sites. There are currently two suites of software utilized for that purpose within OSG: Panda and glideinWMS, both drawing heavily on Condor software.

The Panda system continued as a supported WMS service for the Open Science Grid, with reports of good results from the user community. We worked to increase the resilience and scalability of Panda. To that end, we created a working prototype of the Panda Pilot Factory, which enhances scalability of the pilot job framework utilized in Panda. We added resource usage tracking and enhanced logging capabilities of the Panda monitoring system, which allow the operators to identify and eliminate performance bottlenecks.

To foster wider adoption of Panda in the OSG user community, we started development of a generic data movement mechanism in Panda so as to make easier its utilization by individual users, building on ATLAS work that added built-in data management capability to Panda.

Development work continued on the glideinWMS system in approaching the project goal of pilot-based large-scale workload management. Enhancements to monitoring, stability, reliability, and scalability were performed. The maintenance of the glexec software product, a collaborative effort with EGEE, was added as a project responsibility. The CMS collaboration ramped up its use of glideinWMS on OSG sites. The MINOS collaboration (part of Fermilab VO) has scaled up their operations by an order of magnitude and the CDF VO is preparing to integrate the glideinWMS product with their large-scale computational work as well.

In the area of Security enhancements, the Panda Pilot job software has been updated with the capability to change the effective user ID from the identity of the Pilot Job submitter to one of the end user, based on glexec privileged executable. Successful testing has been performed on properly configured OSG sites. In addition, similar functionality was built into the glideinWMS system and successfully tested. We have maintained a close working relationship with the WLCG collaboration in order to ensure that a common set of security approaches is agreed upon, which is necessary to maintain the high level of interoperability of the OSG workload management software we have enjoyed so far.

This program of work is important for the science community and OSG for several reasons. First, having a reliable WMS is a crucial requirement for a science project involving large scale distributed computing which processes vast amounts of data. A few of the OSG key stakeholders, in particular LHC experiments ATLAS and CMS, fall squarely in that category, and the Workload Management Systems developed and maintained by OSG serve as a key enabling factor for these communities. Second, drawing new entrants to OSG will provide benefit of access to opportunistic resources to organizations that otherwise wouldn't be able to achieve their research goals. One example of such a research groups is the CHARMM team of biologists, who, thanks to simulations run on the Panda system, have successfully produced results used in a publication. As more improvements are made to the system, Panda will be in a position to serve a wider spectrum of science disciplines. And finally, recent security enhancements will ensure compliance with OSG security policies and guarantee alignment of OSG WMS software with policies of member sites of ATLAS and CMS collaborations, thus allowing for a wider deployment and use of their software.

Storage Extensions: The Storage Extensions area contributes to the enhancement of software used in Storage Elements on the Open Science Grid, and software used to discover, reserve, and access those Storage Elements. This includes additional features needed by users and sites, as well as improvements to the robustness and ease-of-use of middleware components. During most of this year we focused on robustness and access. In the last quarter we began work on improving discovery and reservation of storage in order to facilitate opportunistic use of available storage.

Robustness is an important aspect of the grid Facility, both in terms of computational throughput and attraction of new communities to the OSG. Robustness includes the features of software quality, ease-of-use (both in installation and operation), interoperability, and monitoring. We improved robustness by: (1) negotiating interoperability conventions for Glue Schema 1.3 Storage entries; (2) creating Gratia probes for dCache Storage Elements; (3) collecting tools from OSG sites and providing the OSG Storage Operations Toolkit. Easing VOs' access to the grid increases the grid's value. Access is facilitated not only by the mechanics of authorization, but also by better dissemination of the knowledge needed to make use of OSG capabilities. Our work with Role Based Access Control and Space Reservation supported enabling technologies in this area, and direct work with sites and VOs is leading to increased use of Storage Elements in OSG, in ways that lead to greater scientific production. In the area of access, we: (1) supported testing of Space Reservation with the Fermigrid, BNL, UChicago, UCSD and LBNL sites; (2) supported testing of opportunistic storage with the ATLAS, CDF, CIGI, SBGrid, and SDSS VOs; (3) worked with ATLAS on Bestman/xrootd and Bestman RBAC/GUMS support; and (4) updated Generic Information Provider scripts for Glue Schema 1.3.

Internet2 Joint Activities: Internet2 partnered with OSG to develop and test a suite of tools and services that would make it easier for OSG sites to support their widely distributed user community. A second goal is to leverage the work within OSG to create scalable solutions that will benefit the entire Internet2 membership.

Identifying and resolving performance problems continues to be a major challenge for OSG site administrators. A complication in resolving these problems is that lower than expected performance can be caused by problems in the network infrastructure, the host configuration, or the application behavior. Advanced tools that can quickly isolate which problem(s) exist will go a long way toward improving the grid user experience and making grids more useful to more scientific communities.

In the past year Internet2 has provided OSG with two advanced client/server based tools that can perform the function described in the previous paragraph. These tools are targeted at both the site administrator and the end user, allowing anyone to begin debugging a potential performance problem. The client executable is being incorporated into the VDT package ensuring that the required software will be available anywhere inside the OSG. In addition two additional measurement tools will also be incorporated into the VDT package. These tools, geared more toward the site administrator, will allow admins to begin determining if peer network domains need to be contacted to resolve a specific network problem.

While deploying client tools is an essential task, it is not sufficient if the OSG community is to take complete advantage of them. Therefore Internet2 is working with OSG site administrators to deploy perfSONAR boxes that provide a suite of advanced services in addition to hosting the servers needed by these clients. These perfSONAR boxes automatically register their existence in a global database, making it easy to find new servers as they become available. In addition to hosting servers for on-demand testing, these perfSONAR services can be used to schedule periodic test between 2 points, allowing for continuous monitoring of that infrastructure.

The US-ATLAS community is assisting Internet2 in testing the steps needed to deploy these perfSONAR boxes. Once this test phase has been completed, Internet2 will work with other OSG members to refine the existing tasks, and begin a larger deployment.

2.3. Outreach Activities

2.3.1. U.S. Outreach

- "Megajobs: How to Run One Million Jobs" Birds of a Feather session Feather at Supercomputing '08, Austin, November 2008, http://scyourway.nacse.org/conference/view/bof118.
- Joint EGEE and OSG Workshop at the High Performance and Distributed Computing (HPDC 2008): "VO Management in Production Grids" http://indico.cern.ch/conferenceDisplay.py?confId=27789
- The fourth annual Cybersecurity Summit for NSF Large Research Facilities will be held May 7–8, 2008 at the Sheraton Crystal City in Arlington, Virginia. http://roadrunner.lternet.edu/drupal/?q=node/15 http://www.educause.edu/About+EDUCAUSE/MemberCommittees/CommitteeMembers/959?CODE=CYB08

- Contributions to the DOE Grass Roots Cyber Security R&D Town Hall Meetings and mathematics and security workshop.
- Contributions to the Building Effective Virtual Organizations (BEVO) workshop in January 2008, http://www.ci.uchicago.edu/events/VirtOrg2008/index.php?pg=main.
- Presentation at the Joint Genomics: GTL Contractor-Grantee Workshop VI & Metabolic Engineering Working Group Interagency Conference on Metabolic Engineering 2008 February 10-13, 2008, Bethesda, MD
- Contributions to the 2008 NRC report, "Integrated Computational Materials Engineering: A Transformational Discipline for Improved Competitiveness and National Security," http://books.nap.edu/catalog.php?record_id=12199.
- Participation in and presentations for the Campus Infrastructure Days, a collaboration with Internet2, Educause, TeraGrid and other projects. http://cidays.org/Presentations
  - Clemson University CI Days, May 19-21, 2008
  - New Mexico CI Days; March 10-11, 2008
  - Elizabeth City CI Days, January 4, 2008
  - NYSGrid CI Days Workshop - December 2007

2.3.2. International Outreach

- Co-sponsorship of the International Summer School on Grid Computing in Hungary (http://www.issgc.org/). OSG sponsored 6 students to attend the 2 week workshop, provided a key-note speaker and 3 teachers for lectures and hands-on exercises.
- Co-editorship of the highly successful International Science Grid This Week newsletter, www.isgtw.org which is becoming a recognized communication vehicle for publishing of science and technologies of grids. OSG is very appreciative that DOE and NSF have been able to supply funds matching the European effort starting in January 2009. This will enable additional effort to be applied in the US to this effort.
- Training at Cyberinfrastructure for International Collaborative Biodiversity and Ecological Informatics Pan-American Advanced Studies Institute, http://ciara.fiu.edu/eco/
- Presentations at the online International Winter School on Grid Computing http://www.iceage-eu.org/iwsgc08/index.cfm

3. Publications and Products

3.1. Journal publications

3.2. Book(s) and/or other one time publication

Chapter titled "Challenges Facing Production Grids" in High Performance Computing and Grids in Action", ed. L. Grandinetti., IOS Press, Amsterdam, March 2008. Also available as FERMILAB-PUB-07-323-CD.

"New Science on the Open Science Grid", Ruth Pordes et al. Published in J.Phys.Conf.Ser.125:012070,2008.

"The CMS experiment at the CERN LHC", by CMS Collaboration (R. Adolphi et al.). 2008. 361pp. Published in JINST 3:S08004,2008.

"The Open Science Grid status and architecture", Ruth Pordes et al. Published in J.Phys.Conf.Ser.119:052028,2008.

3.3. Other specific products

3.3.1. Teaching aids

OSG has developed web based training materials for Grid Schools. Gregor von Laszewski, Associate Professor in Computer Science at RIT, uses the material in his classes on an ongoing basis, The University of Missouri, St Louis, has also used the class material.

OSG has published an online course where students work through the class material and hands-on exercises. The OSG instructor is available for help through chat, email and phone.

3.3.2. Technical Know-How

OSG is developing an experienced and expert workforce in the operational, management and technical aspects of high throughput production quality distributed infrastructures. This experience includes the use, diagnosis, security and support of distributed computing technologies including Condor, Globus, X509 based security infrastructure, data movement and storage, and other technologies included in the Virtual Data Toolkit.

3.4. Internet dissemination

OSG co-sponsors the weekly newsletter International Science Grid This Week: http://www.isgtw.org/. The other major partner in this newsletter is the Enabling Grids for EScience (EGEE) project in Europe. Additional contributions, as well as a member of the editorial board, come from the TeraGrid. The newsletter has been very well received, having just published issue 104 with subscribers totaling 4,116. It covers the global spectrum of science, and projects that support science, using distributed computing.

OSG research highlights each describe a science result from the project: http://www.opensciencegrid.org/About/What_We%27re_Doing/Research_Highlights

The results currently published are:

- Protein Structure: Taking It to the Bank (12/2008)
- Opportunistic Storage Increases Grid Job Success Rate (11/2008)
- Simulating Starry Images-Preparing for the Dark Energy Survey (7/2008)
- Friendly but fierce competition for the Higgs (6/2008)
- Simulating Starry Images-Preparing for the Dark Energy Survey (7/2008)
- Matchmaking for Science: RENCI connects researchers with computing resources on the Open Science Grid (3/2008)
- Stormy weather: grid computing powers fine-scale climate modeling (3/2008)
- Open Science Grid crunches through CMS simulations (12/2007)

OSG has a comprehensive web site and information repository: http://www.opensciencegrid.org.

4. Contributions

4.1. Contribution to Resources for Science and Technology

The OSG infrastructure currently provides access to the following resources. It must be remembered that OSG does not own any resources. They are all contributed by the members of the OSG Consortium, and are used both locally and by the owning Virtual Organization. Only a percentage that varies between 10 and 30% are in general available for use by the OSG.

Number of processing resources on the production infrastructure	79
Number of Grid interfaced data storage resources on the production	25
Number of Campus Infrastructures interfaced to the OSG	4
Number of National Grids interoperating with the OSG	2
Number of processing resources on the Integration infrastructure	20
Number of Grid interfaced data storage resources on the integration infrastructure	3
Number of Cores accessible to the OSG infrastructure	~44,000
Size of Tape storage accessible to the OSG infrastructure	~10 Petabytes @ LHC Tier1s
Size of Disk storage accessible to the OSG infrastructure	9 Petabytes
CPU Wall Clock usage of the OSG infrastructure	Average of 20,000 CPU days/ day during November 2008

4.1.1. The OSG Virtual Data Toolkit

The OSG Virtual Data Toolkit (VDT) provides the underlying packaging and distribution of the OSG software stack. VDT continues to be the packaging and distribution vehicle for Condor, Globus, myproxy, and common components of the OSG and EGEE software. VDT packaged components are also used by EGEE, the LIGO Data Grid, the Australian Partnership for Advanced Computing, GridUNESP the Sao Paulo state grid, and the UK national grid, and the underlying middleware versions are shared between OSG and TeraGrid.

Much of the work for the OSG Virtual Data Toolkit (VDT) has been focused on the needs of the OSG stakeholders and the OSG software release (described previously in Section 2.2.2).

However, the VDT continues to be used by external collaborators. EGEE/LCG uses portions of VDT (particularly Condor, Globus, UberFTP, and MyProxy). The VDT team maintains close contact with EGEE/LCG due to the OSG Software Coordinator's (Alain Roy's) weekly attendance at the EGEE Engineering Management Team's phone call. TeraGrid and OSG continue to maintain a base level of interoperability by sharing a code base for Globus, which is a release of Globus, patched for OSG and TeraGrid's needs. The Earth Science Grid (ESG) has investigated adoption of the VDT, and several discussion have been had with them about it.