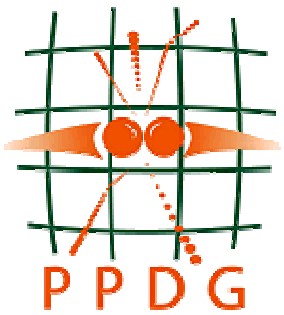
Particle Physics Data Grid: From Fabric to Physics

Quarterly Status Report of the Steering Committee,

January - March 2005

6 May 2005



1 Project Overview

The major focus of effort this quarter is on the preparation for deployment of applications, resources and services on Open Science Grid. PPDG effort is contributing leading roles in coordinating the deployment of the initial version of the OSG infrastructure as well as key services for storage management, security, and discovery. The DOE lab computing facilities are involved with integrating and deploying the OSG environment at their sites. ATLAS and CMS have been validating the OSG release candidate of grid services with their applications on the OSG Integration Test Bed (OSG-ITB).

1.1 Papers and Documents

Reflecting the focus on Open Science Grid, PPDG effort has contributed to many OSG documents[1] this quarter..

Reports, Documents and Papers	Date/Version
OSG bylaws Frank Wuerthwein	25 Mar 2005 OSG 70-v8
OSG Service Agreement Security Technical Group	25 Mar 2005 OSG 87-v4
Open Science Grid Operations Model Support Centers TG	25 Mar 2005 OSG 47-v8
OSG Registration Template Support Centers TG	18 Mar 2005 OSG 97-v3
OSG User Acceptable Use Policy Security Technical Group	14 Mar 2005 OSG 86-v6
Storage Element Authorization Architecture Abhishek Rana et. al.	08 Mar 2005 OSG 96-v1
Security in a "pull" architecture Frank Wuerthwein	07 Mar 2005 OSG 95-v1
Thoughts on Requirements for ad-hoc VOs Frank Wuerthwein	07 Mar 2005 OSG 94-v1
Arguments in favor of a "pull model" Frank Wuerthwein	07 Mar 2005 OSG 93-v1
Installation of SAMGrid on Grid3/OSG site Joel Snow	05 Mar 2005 OSG 92-v1
Open Science Grid Security Infrastructure Blueprint	05 Mar 2005 OSG 91-v1
OSG Operations Plan for 2005 Support Centers TG	25 Feb 2005 OSG 72-v3
Report of the OSG Integration and Operations-Provisioning Workshop OSG Partners	25 Feb 2005 OSG 89-v1
DRM-ITB Plan and Status Jorge Rodriguez	16 Feb 2005 OSG 82-v1
MCPS Task Description Gregory Graham	16 Feb 2005 OSG 79-v1
Finer Grain Authorization with VOMRS Ian Fisk	16 Feb 2005 OSG 78-v1
ATLAS applications on OSG ITB Yuri Smirnov	16 Feb 2005 OSG 77-v1

[1] http://osg-docdb.opensciencegrid.org//Static/Lists//FullList.html

OSG Incident Response Communications 2005.02.15 Doug Pearson	16 Feb 2005 OSG 76-v1
Open Science Grid Deployment Plan - Spring 2005 Deployment	14 Feb 2005 OSG 28-v5
Building an Open Science Grid Ruth Pordes et. al.	09 Feb 2005 OSG 71-v4
ATLAS Computing on Grid3 and the OSG Rob Gardner	08 Feb 2005 OSG 75-v1
US CMS and the OSG Ruth Pordes	08 Feb 2005 OSG 73-v1
A Blueprint for the Open Science Grid Blueprint	31 Jan 2005 OSG 18-v4
OSG Service Agreement Policy Policy Technical Group	12 Jan 2005 OSG 48-v1
OSG Integration Activity Report Rob Gardner	10 Jan 2005 OSG 61-v1

2 The Common Project

The common project effort has become deeply involved with readying the first deployment of Open Science Grid, including Dane Skow as co-chair of the OSG Provisioning activity[2]. Provisioning in OSG is the process of taking a tested release candidate software stack and turning it into the deployed grid-wide production services, including the documentation, software cache and support infrastructure that sites use to install and run the OSG production infrastructure.

The SRM effort has been working very hard on preparing the LBNL DRM for deployment at OSG sites. The dCache/SRM effort at FNAL is preparing a production storage service open to all OSG participants with an SRM interface to mass storage system at FNAL. The FNAL SRM service is expected to be deployed in June and the LBNL DRM available for deployment later this summer.

The Privilege Project products, PRIMA and GUMS, have undergone significant testing and deployment preparations and they have been functioning on some of the OSG Integration TestBed sites (OSG-ITB). These rely on a VOMS service, for which the software is distributed in VDT but is developed by INFN (Italy). VOMS is a key part of the infrastructure making security interoperable with Europe and LCG but there have been a number of compatibility issues arising from varying versions of deployed VOMS servers in the US and Europe.

The Web Service Discovery Service for OSG, based on Clarens and MonALISA, is now included in VDT and has been tested in at least one deployment on the OSG-ITB. It provides a registry for Web Service (WS) based services. At present there are no required WS services being deployed for OSG but it is expected there will be in the near future. SRM and VOMS-Admin for example, have a WS interface. The Caltech group has also provided a tool that can be used to register non-WS services so that it can be used as a directory of services for manual query in addition to automated WS discovery. Use of this discovery mechanism will be compared with the GridCat (provides Grid3 backward compatibility) and MDS (LCG compatibility) experiences.

Development work on a resource accounting system for OSG is being developed primarily at SLAC, and is called Gaspacio. It will not be ready for inclusion in the June deployment of OSG but will likely be tested on the OSG-ITB later this summer.

[2] http://osg.ivdgl.org/twiki/bin/view/Provisioning/

Related meetings:

OSG Integration Meeting, Feb. 15-17, Chicago, http://www.ivdgl.org/osg-int/workshop.html

3 CS-11 Data Analysis Working Group

Participants at one or more of the bi-weekly phone meetings included the following (with most attending almost all of the meetings): Doug Olson, Joseph Perl, Michael Thomas, Frank van Lingen, Conrad Steenberg, David Adams, Jerome Lauret, David Alexander, Tony Johnson, Bala Ananthan, Tahir Azim, Dane Skow.

The areas with notable progress include:

- Adopted Wiki as Better Way to Document our Progress.
  At request of the PPDG executive team, reviewed CS-11 objectives and milestones that were included in the SciDAC Year 4-5 Proposal. Determined that while CS-11 is functioning as intended, facilitating discussion and interchange of solutions for interactive analysis, the specific milestones were not an effective way to document our progress. Moved to use of a Wiki (hosted by our GAE collaborators). The Wiki has made it easy for our members to document their own accomplishments and so provides a comprehensive look at all of our work:
  http://ultralight.caltech.edu/gaeweb/wiki/index.php/ppdg-cs11
- Clarens Discovery Service.
  DIAL and JAS/Tech-X experimented with the Clarens Discovery Service and provided feedback and additional requirements through CS-11 discussions. Requested features were added and reported back through additional CS-11 discussions. Process then iterated further.
- Software Discovery Service.
  This is another service from the Clarens team has been of interest to many CS-11 members. The interface was presented to CS-11 and then enhanced to include additional features requested through the CS-11 process.
- Request Definition Language
  Tech-X Corporation and STAR collaborators reported on their work towards a common Request Definition Language which can serve as input to schedulers. Again CS-11 provided a route for testing and feedback by developers of other analysis systems.
- Additional significant presentations and discussions involved:
  Security Use Cases for Analysis Peer-to-Peer Services.
- ATLAS Dataset Management.
  Dataset Analysis Grid service (DAGS), a collaboration between Tech-X and JAS.

Detailed minutes of all of our discussions are included in the CS-11 email archives.

http://www.ppdg.net/archives/ppdg-idat/

4 Collaborations

4.1 Open Science Grid

Dane Skow represented Open Science Grid for a presentation[3] at GlobusWorld 2005 presenting the Grid3 successes and our plans moving forward on OSG. The talk was very well received and sparked several

[3] http://computing.fnal.gov/docdb/osg_documents/0000/000071/004/DDS050209-OSG-GlobusWorld.ppt

discussions with persons from Canada, UK, Japan on similar activities in their nations. Considerable PPDG effort towards OSG is described elsewhere throughout this report.

Many members of PPDG are in leadership roles in the OSG Technical Groups and Activities[4] as well as reprensted on the OSG Council and Interim Executive Board[5].

4.1.1 Security

Dane Skow and Bob Cowles continue to serve as co-chairs of the TG-Security in OSG. A test of the OSG Incident Response procedure was successfully performed at the ITB workshop. To our knowledge this represents the first test of a Grid incident response procedure anywhere in the world. This first step identified areas for future work and sparked plans for future tests and collaboration with other Grid projects on tests of more complex scenarios.

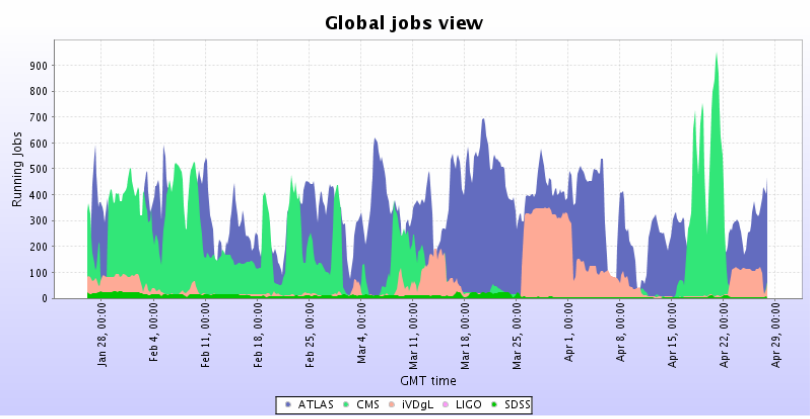
Basic User and Service agreements were developed for OSG in consultation with our LCG colleagues. We expect that the basic requirements (and most of the text of the agreements) will be common as the LCG have agreed to take these documents as the basis of their usage requirements agreements. We have continued to participate in the Joint Security Policy Group working on trying to harmonize such policies and requirements between OSG and LCG. Bob Cowles and Dane Skow participated in the LCG MiddleWare Security Group (MWSG) meeting at CERN in February 23-25 and started detailed discussions about agreements and developments needed for the migration to webservices based grids. It is clear that there are many new aspects of this technology and developments both in the US (Globus) and Europe (gLite, OMII) that need to be integrated. One particular aspect (a web services based credential delegation method) was identified as a first start and is presently underway. There is work needed on the gene ral authorization framework and a start is underway between Globus and PDC in Sweden. An evaluation is needed soon to determine whether this new development meets the needs identified in experience with previous toolkits.

A letter was drafted and sent by the OSG Executive Board to the 3 regional Profile Management Associations (EUGrid-PMA, APGrid PMA, TAGPMA) asking them to serve as Certification Authority accreditation bodies for OSG. Defacto, the EUGrid-PMA has been serving this function for LCG. We sought to extend this role to OSG and make more explicit the services we were requesting. This letter has been well received by the associations and sparked agreements at GGF13 for the 3 bodies to form a global federation to coordinate their activities and accreditation profiles.

4.2 Trillium and Grid3

Grid3 has continued its role as a production facility for ATLAS and CMS as illustrated in the diagram below.

[4] https://plone3.opensciencegrid.org/opensciencegrid/
[5] http://osg-docdb.opensciencegrid.org/cgi-bin/RetrieveFile?docid=70&extension=PDF



Running Jobs				
Farm	**Last value**	**Min**	**Avg**	**Max**
ATLAS	465	0	329.6	784
CMS	62	0	157.1	1.031 K
iVDgL	60	0	62.62	425
LIGO	0	0	0	1
SDSS	0	0	8.509	41
Total	**587**	**0**	**557.8**	**2.253 K**

4.3 EGEE, LCG

Interoperability discussions and work with LCG and EGEE are increasing with the formation of an OSG Activity focusing on interoperability with other major grids (EGEE and TeraGrid as first two partners). PPDG participates in this activity as well as direct efforts through the various experiments. This activity meets now weekly and has begun plans for interoperability testing during the next quarter.

Ruth Pordes and Razvan Popescu serve on the joint Baseline requirements team for the LCG helping to define the basic required services needed for robust initial operation in 2006 of the LHC Computing Grid.

Miron Livny continues his involvement with the gLite development helping to ensure interoperability between the EGEE and PPDG/OSG efforts. Kate Keahy from the Globus team has been working on developing a dynamic account allocation system that can be used in both EGEE and OSG software stacks.

Bob Cowles and Dane Skow represented Open Science Grid at the MiddleWare Security Group meeting (MWSG) at CERN in February for detailed discussions about security interoperability of new planned features of the software kits being developed in EGEE, INFN, Globus, and PPDG. They also continue to represent OSG on the Joint Security Policy Group and have reached agreement on common usage and service policies for EGEE, LCG, and OSG. Planning has begun on a joint incident response drill to build upon the OSG successful test this quarter.

4.4 DOEGrids PKI

There are two significant aspects of PPDG activities related to the DOEGrids PKI this quarter, some relating to policies and their implementations, and another about command line script interfaces providing a simplified usability in a grid context.

There was a lack of clarity in the definitions of uniqueness of grid certificate identifiers about whether unique meant that only a single instance of a certificate should exist or if the binding between the distinguished subject name (DN) in a certificate to an end entity (EE) is unique. This has been clarified in discussions and updates to the official policy document (CP/CPS[6]) is being updated to clarify that it is the

[6] http://www.doegrids.org/Docs/CP-CPS.pdf

binding between the DN and the end entity (person or computer) is unique but that multiple instances of certificates with the same DN are allowed as long as the unique binding to the EE is maintained. There are also updates to the policy document to introduce the concept of ownership of the DN for service certificates (those certificates that identify services running on a computer rather than a person). Previously it has been assumed that different unrelated persons or groups requesting a service certificate with the same DN would not be allowed but this was not actually documented in the policy.

A set of command line scripts[7] providing a simple interface to DOEGrids from a unix grid environment were developed in previous quarters (and mentioned in previous reports) but these are now included in the distribution of the VDT[8] and being used in the OSG ITB.

5 Single Team Reports

5.1 Experiments

5.1.1 ATLAS

The ADA/DIAL team with help from the ATLAS facilities group at BNL was very busy the first quarter of this making DIAL release 1.0 and preparing for release 1.1. The latter includes the long-promised transformation catalogs, provides up-to-date snapshots of the Rome AOD samples as they are produced and is supported by stable analysis services running at BNL and CERN. This is the first release intended to provide support for serious ATLAS analysis. It provides support for catalog browsing both on the web and at the root command line.

5.1.2 BaBar

Wilko Kroeger - BaBar root-files are routinely transferred between SLAC and IN2P3. The transfer is very stable and only a few problems were encountered. We spend a little bit of time to improve our tools in particular to ensure that the SRB contains all BaBar root-files.

Most of my effort was spend on testing the new SRB release, version 3.3. We are planning to put it into production by the end of April to be used for the distribution of the new and reprocessed BaBar data. We tested the new version to ensure that all of our tools are still working properly. The new SRB version offers new features that makes it easier to do mass loading of the BaBar root-files and setting the proper permissions. These features were tested and incorporated into our tools.

From January to March Matteo Melani had been working on an accounting prototype for the Open Science Grid. This effort is part of the PPDG common project contribution to Open Science Grid.

The accounting system prototype, named Gaspacio, serves two purposes: one is to study and understand the main stakeholder requirements for an accounting system in the grid environment; two is to give Matteo the chance to familiarize with some of the technologies that are part of the grid environment (Java, XML, Web Service, Glue Schema…). Another important purpose of the Gaspacio system is to complete the Open Science Grid monitoring subsystem by adding storage accounting. Details on the prototype can be found at: http://www.slac.stanford.edu/~matteom/osg/accounting/accounting.html.

5.1.3 CMS

Michael Thomas:

Summary

=======

[7] http://www.ppdg.net/RA/cert-scripts/
[8] http://www.cs.wisc.edu/vdt/releases/1.3.5/contents.html

This quarter was focused on further development of the JClarens Grid Service Framework and its integration with the OSG and Ultralight activities.

JClarens

========

Significant progress has been made on the JClarens Grid Service Framework, making it more robust and easier to use for both site administrators and grid service developers. Internal API changes have greatly simplified the process of writing new grid services for JClarens. Other important changes include:

* A software discovery service prototype was written to help application authors locate grid sites that contain specific versions of analysis software packages. Applications or schedulers can use this information to direct data analysis jobs to those sites so that the overhead of installing new software is not incurred. The software discovery service makes use of the same MonALISA/JClarens integration in the same way as the discovery service.

* The concept of Background Processors was added to JClarens. Background Processors are plugins that add new functionality to the JClarens server without exposing a new Web Service interface. A simple example of a Background Processor is a hit counter that keeps track of the number of web service requests and rate of web service requests on the server. This hit counter monitors these rates and publishes them to a local log file or to a remote monitoring tool such as MonALISA.

* As a host for grid services, JClarens must provide a rich and easy to use API to assist software developers during the development of new Grid services. The barrier to entry for writing new services must be sufficiently low so that developers see immediate benefit from using JClarens. Numerous improvements to reduce this barrier to entry have been made. This includes the addition of numerous APIs to assist service developers, the creation of a generic xmlrpc binding layer to remove the need to write service-specific bindings to the xmlrpc protocol, and the creation of a sample 'Hello World' service that can be used as a template for writing new services.

* JClarens was added to the Virtual Data Toolkit, a collection of grid middleware that can be easily installed and configured. The VDT is the primary software delivery vehicle for grid middleware in Grid3 and the OSG. Inclusion in the VDT should help increase the exposure of JClarens to the Grid Service developer community.

As a result of these improvements, JClarens was recognized as an important piece of software for several different projects. LambdaStation is using JClarens to host grid services to schedule and control optical light path connections. The Monte Carlo Processing Service at FNAL is using JClarens to host Web Services for end-user job submission requests. The SPHINX Grid Scheduling Middleware team is developing their novel grid scheduling software as a JClarens Grid service. The Open Science Grid is using the Discovery Service as the standard Web Service registry. JClarens is also an important component of the Ultralight project, acting as a Grid Service framework for many of the services developed in this project.

OSG

===

Two major projects also started using the JClarens Framework during this past quarter. The Open Science Grid (OSG) is adopting the JClarens Discovery Service as the standard for Web Service registries. The Discovery Service was presented at the OSG Integration Workshop in February 2005. As a result, a number of OSG grid sites are now running JClarens and the Discovery Service to provide a Web Service registry to local applications. The OSG does not currently utilize a wide variety of Grid Services, but this is expected to change, and the Discovery Service was adopted early on in anticipation of this change. This early success of the Discovery Service in the OSG has prompted discussions of creating new distributed global registries similar to the Discovery Service to store non-Web Service information.

Ultralight

==========

The Ultralight project is a collaboration of experimental physicists and network engineers whose purpose is to provide the network advances required to enable petabyte-scale analysis of globally distributed data. One of the goals of Ultralight is to develop and deploy prototype global services which broaden existing Grid computing systems by promoting the network as an actively managed component. The JClarens Framework is being used in the Ultralight project to host such global services. A workshop was held at Caltech in January 2005 to discuss and plan applications that can run on the Ultralight testbed. As part of the Ultralight project, a testing framework was set up to help identify misconfigured or unavailable services in the Ultralight testbed. Nightly tests are now run against many of the Clarens servers in the Ultralight testbed, the results of which are sent to the server administrators. An important side effect of this testing framework is that an increased set of compatibility tests are now run between various web service clients and the python and java-based Clarens servers. These compatibility tests have already helped to locate and identify potential problems in the implementations of Web Services hosted in Clarens servers.

E&O

===

The significant Education and Outreach activity this quarter was the ongoing collaboration with the National University of Science and Technology in Pakistan. I have been serving as Ashiq Anjum's technical supervisor for his PhD thesis project. Ashiq's thesis project involes the development of a dataset location service that will locate remotely available datasets that are ranked by various access criteria, such as estimated transfer time or grid access policies. We anticipate that this will become an important component of the Ultralight project and CMS in general. Tahir Azim is another student from NUST. He has been working closely with us on developments in the core JClarens Framework. Tahir arrived at Caltech in February 2005 for a 6 month visit so that he can work more closely with the rest of the GAE team. Tahir's contributions to JClarens include:

* authentication and authorization modules for JClarens

* a windows installer that enables easy installation of JClarens on Windows machines.

* support for accessing JClarens services through the Simple Object Access Protocol (SOAP)

* a 'shell' service to enable secure, authenticated execution of shell commands on Linux based JClarens servers.

* support for authentication using proxy credentials and proxy delegation.

* a Java webstart graphical client to access Clarens services.

* various utility classes and features that hugely simplify the process of writing new JClarens services for new developers.

Tahir has also been an important conduit between us and other developers at NUST working on GAE components. Their work includes:

* UDDI-based discovery: A discovery service implementation based on UDDI was done by Faisal Khan from NUST to conform to the generally accepted Web Services standards.

* Steering service: Adeel Zafar has been leading the effort to develop a steering service for greater control over jobs submitted to the Grid, by allowing users to move, restart, and reschedule submitted jobs.

* Job monitoring service: Waqas ur Rehman wrote a job monitoring service which uses information from Condor to report monitoring statistics related to a running job.

* Estimator service: Atif Mehmood developed a prototype estimator service to provide advance estimates of expected job completion and data transfer times.

* Database access service: S. Yousaf Shah and Tahir have developed a prototype service to provide transparent access to heterogeneous relational databases.

* Ahsan Ikram and Tahir developed a number of applications and prototype services to enable handheld devices to access Grid services and data using Clarens as a web service wrapper. These included several added features in the PocketPC JAS software enabling it to access and plot hbook data, and submit jobs on Condor pools for data analysis, using Clarens web services.

Meetings and Conferences

==========

I have been attending all of the PPDG phone conferences, and recently started attending the PPDG Common Project phone conferences as well. In addition, I have started participating in the OSG Integration and MIS-CI activities.

Suresh Singh: Here are two main accomplishments in this quarter (Jan-Mar, 2005):

1. Caltech-PG Readiness for OSG Deployment:

With the target of Spring OSG Deployment in mid-May, necessary preparations are underway to make the Caltech Production Cluster (Caltech-PG) a full-fledged OSG site. This 33 node dual Intel Xeon based cluster has been upgraded to RHEL 3 from RH 7.3 with Rocks 3.3. Recently, a latest dCache storage system with SRM interface has been deployed on this cluster. PostgreSQL has been used for the dCache and SRM transfer related databases. We used an Intel 3.2 GHz node for dCache admin as well as SRM/GridFTP door node. A third party SRM transfer was verified by transferring some files from FNAL. We also deployed Phedex (Physics Experimental Data Export) data transfer system on this cluster by registering it to "CMS Production Topology". The operation was verified by transferring 235 Digi and Hits files from FNAL to Caltech which can be viewed in main Phedex monitoring page at http://cms-project-phedex.web.cern.ch/cms-project-phedex/. All of these tranferred files are registered to l ocal MySQL catalogue. We are making a list of datasets of interest for the user analysis. Though, the latest OSG release have been installed for the test purpose, we are waiting for the final stable production quality OSG release from OSG Integration Group.

2. Participation on ITB:

A six node cluster, originally named as Caltech-Gri3 has been used for OSG ITB (Integration Testbed) related activities. This was also being upgraded to RHEL 3 with Rocks 3.3 and was used during the last ITB workshop held in Chicago. Presently, we are actively invloved in testing and validating every new ITB release. This validation includes core VDT components such as Globus toolkit, condor, Monalisa monitoring etc and optional service components such as Core-MIS, Prima authorization module, jClarens Discovery, DRM/SRM etc. We observed that some service component in VDT are not fully supported in Rocks built cluster and have been reported to VDT team which they have assured to support in future ITB releases.

Iosif Legrand:

<u>Developments for the MonALISA framework</u>

The storage part in MonaALISA service and repository was improved significantly. A memory cache module is used based on available resourced detected on the running system. It adjusts itself to available memory on the system. It is interfaced with MySQL and Postgres DB systems which are deployed with the MonALISA distribution as embedded storage systems. The performance of the data storage part is improved, and the new storage part allows recording more than 5 000 monitoring records per second and service for large databases.

A new GUI was developed and it provides many new functions: - On demand loadable high resolution maps

- A new panel for a WAN topology

- A new panel for optical switches to show the state of each link and the signal the active links

- Possibility to plot aggregated values from multiple sites in real-time or user selected history (average, integral, sum ?)

Two distributed services were implemented using the agent mechanism in MonALISA:

Distributed Intrusion Detection System. Each site is using a local system to detect attacks and once a system is identified performing strange requests, its IP is distributed to all the other sites in less than 1s and it can be blocked at the system or router level.

On Demand Optical Path Connections. We developed dedicated modules to monitor and control optical Switches (Calient and Glimmerglass) based on the TL1 or GMPLS/TL1. The connectivity map for each device is known and controlled by an agent. These agents can interact with end user application and can create an end to end optical connection, on demand is less than 1s independent on the number of switches involved or their location.

Support for OSG-ITB and GRID3

We help in deploying MonALISA services to all the OSG-ITB sites and provided support to the site administrators

Improvements were done for the VO modules :

- load dynamically grid-vo_maps files when they are changed and to reconfigure the VO maps
- added the functionality to get job information from Sun Grid Engine

Now, the VoJobs module allows to collect information from (Condor, LSF, PBS, FBS, SGE).

In OSG-ITB distribution we included the WAN topology agents. These dedicated agents are used to construct in real-time the WAN topology graph. From each site a tracepath like measurement (using a multithreaded requests with different TTL) are periodically performed to all the other sites in a group or virtual organization. The topology graph is constructed from all these measurements, identifying all the routers, the network?s IDs and corresponding AS. The system evaluates the mean latency on each on each segment and can easily detect asymmetric routing. Agents are dynamically configured from servelts that use the discovery mechanism in MonALISA and they also detect the IP of the requesting agent.

We setup and configured a dedicated OSG-ITB repository http://gocmon.uits.iupui.edu:8888/

A web page to administrate the sites in OSG-ITB and to configure dynamically the ABping (network performance monitoring) measurements is also part of the OSG-ITB repository.

We had the opportunity to deploy a MonALISA service at LBL (on the PSDF STAR cluster) and is running on a dual cpu AMD system in native 64 bits. This was very useful to verify on real systems that the MonALISA distribution works correctly on both 32 and 64 bits architectures.

Abhishek Singh Rana –

Common Project areas: Data Management and Security

Storage Security Design

We worked on understanding the existing security frameworks for the Compute Element (CE) in grid environments. Based on our experience working with distributed storage (SRM-dCache), drafted a specification of initial stage Storage Element (SE) authorization architecture within the context of USCMS and Open Science Grid. This is work in progress, and is illustrated in OSG-doc-96.

Storage Security Development

Motivated by a need for fine-grain security (VO Role based Authorization) in SE's and for providing a choice of pluggable security mechanisms to Sites and VO's, we introduced gPLAZMA (grid-aware PLuggable AuthoriZation MAnagement) architecture. This has been implemented as gPLAZMA module, targeted towards SRM-dCache. The module also introduces Storage callouts: SRM callout and GridFTP callout.

An internal release v0.1 of gPLAZMA module within scope of SRM-dCache was made. This release is functional with SRM server and GridFTP door in dCache, integrated as SRM callout and GridFTP callout to the module. There is a provision for using different authorization mechanisms simultaneously, fine-tuned with switches and priorities of mechanisms. Of the four mechanisms supported: one is an integration with services from the Privilege Project (GUMS and PRIMA), others are built-in as a lightweight suite of services (gPLAZMAlite). A quasi-firewall option is included for Grid Incident Response.

The implementaion provides full backward-compatibililty. Site administrators can continue to use legacy security (dcache.kpwd) by not enabling the module. When enabled, module still supports legacy mechanism.

The module is currently being deployed with SRM-dCache on a small-scale and tested at UC-San Diego and Fermilab.

A brief illustration of this work is in OSG-doc-108. An executive summary is available at http://home.fnal.gov/~rana/gPLAZMA_v0.1_executive_summary.pdf.

This work was undertaken as a collaborative effort between PPDG Common, Fermilab Privilege project, Fermilab and UC-San Diego USCMS storage groups, and OSG Storage technical group.

USCMS Tier-2 Storage (SRM-dCache) Deployment

UC-San Diego Tier-2 Center continues to remain active in deploying the most recent internal releases of SRM-dCache, and upgrades the deployment weekly. Active participation in LCG Service Challenge (SC2) is in progress. We also provide ad hoc consulting to other Tier-2 centers on storage deployment and configuration related matters.

5.1.4 D0

Description: (Gabriele Garzoglio)

In the past three months I have concentrated on the SAM-Grid system.

SAM-Grid has been deployed to about a dozen sites in US, EU, and Canada. The infrastructure is going to be used in the next 6-9 months for the DZero data reprocessing effort. This consists of 250 TB of physics data that need to be re-processed to take advantage of a better understanding of the detector, in particular reconstruction algorithms and calibration constants. To complete the task in time, the effort requires 1600 GHz-CPU year of computing power.

Special care has been given to the scalability of the infrastructure at each site: the grid can support running on the order of 1000 concurrent jobs at every site, if the resources are available.

We are also concentrating in the integration of the SAM-Grid infrastructure with OSG and LCG.

We are supporting CDF work to use the SAM-Grid infrastructure for montecarlo production.

5.1.5 STAR

5.1.5.1 General

Members of the STAR Software and Computing team and the PPDG Grid supported personnel met at BNL for a STAR Grid workshop at the end of March. We would like to mention that our grid efforts is currently supported at a proportion PPDG:STAR equivalent to 1:3 so, the manpower is currently mostly driven from internal resources.

The workshop was mainly aimed to discuss the roadmap toward running on OSG and/or other opportunistic resources and help bringing all the pieces together. We discussed of the work still to be done in order to fully exploit a model which would be based on delegating the job output to an SRM so we could envision real-data production on the grid rather than the traditional simulation runs on grid. The basic idea is to "hand" off the output to an SRM service (a DRM cache) and let that service transfer the file back the submitter site (in our case, would be BNL) where a SRM/RRS would also be running. The file would then

be pushed into a MSS and registered using RRS, completing the chain of production. During the workshop, we exercised this scheme but did not entirely achieve the expected goal.

Another topic was the current state of our production level data transfer; while most parts from Tier0 to Tier1 transfer are automated, there are a few operations which require manual interventions: for example, reproduced data should be re-transferred when a job crash recovery occurs (same output in principle expect that it ends up more complete). Also, the re-indexing of the tag database is not automated for now (this is used by the GridCollector) nor is the handling of updated records from the Catalog (related to reproduced files, the size may change and be updated). In general, the benefits provided by RRS are being fully exploited with fine tuning of the special cases.

Rob Gardner was very kind to grant us some of his time to guide us through the procedures and documents available describing and necessary for an ITB deployment. We would like to thank him for his time, help and support and appreciated once again working with him. As an immediate outcome, we took care of creating/declaring a STAR VO for OSG and believe we have all of the necessary procedure in place. While we were missing at the time a security representative for the STAR VO registration, Gabriele Carcassi has kindly accepted to be our contact person.

Additionally, Doug and Iwona also discussed with Gabriele about GUMS and how to use/deploy. We intend to use GUMS at PDSF. This is described in the NERSC/PDSF section. STAR already has a VOMS server running at BNL, thanks again to Gabriele's good work.

We also took the opportunity to revive our MonaLisa support now in the hands of Levente Hajdu. Efstratios Efstathiadis from the IT department has now moved to new responsibilities and will not be able to support this activity. We would like to thank him once more as well as the IT director Thomas Schlagel for his help and contributions to our common efforts.

5.1.5.2 OSG, SRM

Eric Hjort has continued to support the STAR production data transfer between BNL and LBNL. About 17 TB and 85k files were transferred in this using HRMs + Replica Registration Service or RRS. As mentioned above, work is in progress to further automate the transfers of files which may be recovered from a crashed job. While only of small fraction of those occur, we intend to build a system which would not require manual intervention.

Eric has also pursued his participation in the in the OSG-DRM activity helping to figure out how to install and configure DRMs. The documentation was greatly enhanced based on his knowledgeable and expert feedback.

5.1.5.3 SUMS and STAR jobs on the Grid

We reported in the last quarterly some difficulties in getting our infrastructure back in shape after an OS upgrade required by Cyber-Security constraints. Related to relatively trivial issues such as miss-configuration, changes in the network interface or routing, port range mismatch, missing shell declaration in /etc files, wrong hostname setting and other issues of this order, the previous quarter network transfer testing was the first effort to correct those issues we found to be present on all of our RCF maintained gatekeepers. New waves of effort were then launched to also address and restore the Grid based job submission. STAR/PPDG and STAR S&C personnel concentrated on understanding and resolving the issues ourselves. The people involved in this activity included Iwona Sakrejda, Levente Hajdu,, Lidia Didenko, Doug Olson with help from myself & the RCF personnel adjusting the setting as instructed or found as improperly set. The main issue was an interesting side effect of having Condor set to a certain port range as the local RMS while a different port range was expected for communication with Condor-G (i.e. another globus port range issue).

The outcome was very positive. We finally restored our Grid based simulation production. Tested and exercised by the STAR production coordinator, Lidia Didenko, she reported 100% success over 500 jobs and 50k events by waves of 50 jobs at a time for a period of three weeks. In this case, success means successful completion of jobs at PDSF submitted from BNL as well as successful transfer of the output back to BNL. It is noteworthy to mention that the data transfer mechanism in its native form in SUMS (using the U-JDL syntax) was buggy and is now fixed. The test mentioned in this section incorporated the changes made to this effect.

On another front, Levente has implemented a new dispatcher, allowing for executing jobs on the local node (equivalent of a "fork" or local thread). This special dispatcher was mostly developed to test SUMS and help with the development cycles of module developers outside of STAR who do not have access to the grid or even a computing farm.

Finally, and since SUMS is in operation for many years now, we bumped into a difficulty with its statistics database which has grown out of reasonable proportion: more than 3 million rows corresponding to more than 3 million jobs submitted by users via SUMS. A redesign was necessary to handle the demand and preserve the full historical data. The usage statistics was therefore reshaped and split by years rather than a big monolithic table. This required a reshape of the API so the access to a multi-table layer would be transparent to the JSPs, Michael DePhillips (STAR database leader) helped and advised in this task. A few graphs were added that show the amount of time it takes to dispatch a job via a given dispatcher.

5.1.5.4 Other activities and plans

Iwona Sakrejda has been working on delivering a Sun Grid Engine (SGE) job-manager to VDT. Based upon the initial version developed by the London e-Science center, Iwona tested the version and made sure the job manager was working well. As mentioned in the past, we have moved to a SGE based RMS at our NERSC/PDSF site and support for this job-manager is now essential. Iwona and Alain Roy (VDT) were in contact so the job-manager would be included into the base VDT release. We volunteered to be the official testers for this job-manager if need be. It is our understanding from Alain that the LSF job-manager is not thoroughly tested at releases either; we are in a position to do that as well shall we be provided with basic instructions of test expectations.

Infrastructure wise, and considering the past difficulties in relying on external to STAR resources to maintain our gatekeepers, more internal manpower will be drawn to the Grid effort to cover for hardware and software stack installation. Wayne Betts, our computer support specialist will be accommodating his schedule and provide such support. Wayne will slowly get accustom with Grid software and start with deploying an OSG-ITB gatekeeper at BNL.

Mid-march, the S&C management team met with the Alice software team who came for a visit to discuss computing model and tools used in running experiment. We presented our light-weight Grid approach, catalog design (there were interested in our federated catalog approach) as well as our implementation of an interactive framework on Grid (the GridCollector). Alice and CERN in general are extremely interested in this later technology.

Contact with Jean-Yves Nief from Babar/In2p3 was made in Nantes (France) end of April. We discussed briefly the possibility of working together on the SLAC Xrootd/SRM project we have initiated with the Xrootd team (Andy Hanushevsky and Peter Elmer). While interested, it appeared that In2p3 will not be in an immediate position to help. Our work on this front will resume in summer.

5.1.5.5 Grid computing outreach

In this quarter Levente was asked to present two talks. The first was given at the STAR Collaboration meeting and was giving an overview of the current development and status with details on the improvements and the ongoing research. It was interesting to observe that users seem to be more interested than usual, the interest shifting toward trying to better understand how a "user" would run his jobs on the Grid. We would need to concentrate again on this aspect soon as the demand and need is growing.

The later talk was presented at the BNL Technology Meeting and was entitled "*STAR Job Scheduling*".

5.1.6 PHENIX

5.1.7 ALICE

5.2 Facilities

5.2.1 Jlab

Michael Haddox-Schatz - In the January-March 2005 timeframe I continued to investigate a new version of the SRM (v.3) based on ideas presented at the SRM meeting at LBL in December 2004. The main design goals are that the SRM be simple, practical, and modular. Modular means that an SRM implementor just has to implement the components they want to provide. For example if a site wanted to provide read access without write access, they would provide an SRM Get module but not provide an SRM Put module. Simple and practical means that the SRM protocol only defines what is necessary, and that it is straightforward for sites to provide an SRM interface to there resources. To help ensure these "simple and practical" qualities, I am creating a prototype SRM implementation in conjuction with writing a proposed SRM v3 specification. During the first quarter of 2005 work progressed on both of these fronts.

5.2.2 BNL RCF/ACF

Gabriele Carcassi has been acting as the Privilege Project representative in the PPDG Common Project group. He worked to make sure the deliverable from the Privilege Project would be ready for the OSG-0 milestone, by:

- Releasing the 1.0.0 and 1.0.1 versions of GUMS
- Making GUMS available through the VDT
- Installing and testing the full chain of the system (VOMS Admin/VOMS/voms-proxy-init/PRIMA/GUMS)
- Making available a test GUMS server for other site to experiment with
- Participating at the OSG workshop in February
- Writing documentation on GUMS and OSG sites
- Following compatibility problems between OSG and EGEE
- Providing support

5.2.3 FNAL

FermiGrid Project Status

Keith Chadwick

The Fermilab Facilities are working on an IntraGrid – FermiGrid – to facilitate common interfaces between the clusters used by Fermilab users and provide common interfaces to the Open Science Grid. FermiGrid has contributed to the OSG Integration activity at the same time as commissioning and deploying services for the local community.

The FermiGrid project has Four Equal Elements

FermiGrid Common Grid Services: Supporting common Grid services to aid in the development and deployment of Grid computing infrastructure by the supported experiments at FNAL

FermiGrid Stakeholder Bilateral Interoperability:Facilitating the shared use of central and experiment controlled computing facilities by supported experiments at FNAL– CDF, D0, CMS, General Purpose Farms.

FermiGrid Development of OSG Interfaces for Fermilab: Enabling the opportunistic use of FNAL computing resources through Open Science Grid (OSG) interfaces.

FermiGrid Exposure of the Permanent Storage System: Enable the opportunistic use of FNAL storage resources (STKEN) through Open Science Grid (OSG) interfaces

The Status of

FermiGrid – Common Grid Services is that the hardware to host the FermiGrid Common Grid Services has been delivered and installed. FermiGrid1 has been configured to point to fngp-osg as its condor master. VOMS has been installed on FermiGrid2.

GUMS has been installed on FermiGrid3. Ganglia monitoring has been established on FermiGrid 1,2,3. FermiGrid 1,2,3 are being backed up using the TiBS system. Fermilab and dzero VOs have been established on FermiGrid2.

We are researching Issues with IP aliases and Grid Service Certificates in order to allow "service named" IP aliases such as voms.fnal.gov and gums.fnal.gov. VOMRS should be installed on FermiGrid2 this week (pending RPM availability from Tanya Levshina),

GUMS will be configured with two “default” user mappings:

• 1 to 1 – between all “active” Kerberos accounts and assigned Fermilab wide UIDs.

• Many to 1 – (i.e. the entire VO mapped onto a single UID). This will principally be used in the GP Farms and D0 SAMGrid to maintain the current functionality.

Additional customized mappings will be made available as needed by VOs to perform grid related administration tasks and manage privileges.

FermiGrid – Stakeholder Interoperability: Biweekly FermiGrid stakeholder

meetings are taking place to coordinate this work. FermiGrid is working with the genoimics group at the University of Madison Wisconsin to test the provision of storage and data access for the researchers.

5.2.4 NERSC/PDSF

PDSF has continued to provide resources to Grid3. Work has started on setting up a gatekeeper for OSG. One special issue for NERSC is relating the authorized user lists published with the VOMS servers of the various VO’s with the user database at NERSC, called NIM. This issue was analyzed in discussions with Gabriele Carcassi at a STAR grid workshop at BNL in March and a mechanism using GUMS to provide an interface that will allow VO groups and role information to be combined with the authoritative user information in NIM was identified. Effort to work out the implementation details is continuing.

5.2.5 SLAC

In the past 3 months SLAC staff has worked to set up two Open Science Grid environments: one for production and on for testing and development. The plan for the coming weeks is to have MC simulation production for CMS and Babar.

5.2.6 Collaboration with IEPM, Network Performance Monitoring

5.2.6.1 Bandwidth/Throughput Monitoring

The DataGrid Wide Area Network Monitoring Infrastructure (DWMI) toolkit now has IEPM-BW monitoring successfully installed, making measurements, collecting, analyzing and reporting results at:

- BNL
- Caltech
- CERN
- FNAL
- SLAC

After some difficulties in getting ports opened at BNL, all sites are up and running successfully. We are working with BNL to get a faster host at BNL, currently it is limited to 100Mbps. All the other hosts are running at 1Gbpss.
The measurements include traceroutes at 10 minute intervals, pings and capacity/throughput. To reduce load on the network and remote sites being monitored we are mainly using the lightweight ABwE/abing monitoring tool developed by the INCITE project. It provides rough estimates of capacity, cross traffic and available bandwidth while using only 20 packets. For monitoring between the monitoring sites we are also using the more intensive iperf tool. We are looking at an alternative to iperf to get achievable throughput measurements. This uses thrulay a new tool from Stanislav Shalunov of Internet2. This appears to be easier to control than iperf. It has been installed at SLAC and initial experiences are positive. We need to compare and contrast the results from iperf and thrulay before we deploy it elsewhere, We presented a status report on DWMI at the Internet2 Joint Techs meeting in Salt Lake City in February.
We are working on providing maps of the IEPM-BW deployment. We are also working with Iosif Legrand of CERN/Caltech to decide how to provide MonALISA access to the data via a graphical web interface. With the new version of IEPM-BW (version 3), the web services access to the data no longer works. We will be reviving this to enable application access to the data.
We need to work with BNL to understand how we can monitor and compare MPLS/QoS circuits with shared best-effort circuits.

5.2.6.2 Anomalous Event Detection

We are working with FNAL to evaluate various methods for detecting anomalous events in the monitoring data. These include the Plateau Algorithm, Kolmogorv-Smirnoff, Holt-Winters, Principal Component Analysis, and a technique from Mark Burgess. Avoiding false positives and missing events is made difficult by the diurnal behavior of the data and the noise in the ABwE/abing data. We have presented a status report at the ESCC in Salt lake City and are working on a publication.

5.2.6.3 PingER and Developing Regions Monitoring

We put together and submitted the 2005 ICFA/SCIC Network Monitoring Working Group Report on Internet Performance.
Two members of the IEPM team visited NIIT in Pakistan for two weeks each, to further the MAGGIE-NS collaboration that is building and extending tools to provide sustainable monitoring of the Internet.
We have set up PingER monitoring sites in Pakistan and India which will assist in providing information about performance within, between and from Developing Regions, assisting the exiting measurements from Brazil and Russia. We put together a presentation for the World Bank on Quantifying the Digital Divide

5.2.6.4 Testbeds

To facilitate usage of the UltraScienceNet (USN) and UltraLight 10GE circuits being connected to Sunnyvale, we have acquired (from CENIC) colocation space and power on a temporary basis ($1K/month) at the Sunnyvale Level/CENIC Point of Presence. The plan is that when ESnet provides 10Gbits/s access from SLAC to Sunnyvale (planned for July 2005), we will review whether we need to retain / can afford this colocation space.

At Sunnyvale we have installed an UltraLightCisco 6509 router switch, which is connected to the 10GE UltraLight circuit and has access to StarLight and Caltech. In addition, from BaBar, we have loaned four SunFire V20Z's each with two 1.8GHz AMD Opteron 64bit cpus. Two of these have been configured with Red Hat EL 4 Linux and SLAC owned Neterion/S2io 10GE Network Interface Cards (NICs). These have been installed at Sunnyvale and connected to the Cisco 6509. To enable remote management at Sunnyvale we have also installed remote power cycling control, console access via a terminal server, and 10Mbps management access to the hosts. We have exchanged accounts with other UltraLight sites at CalTetch, CENIC-LA and StarLight and are currently working on optimizing the configurations to achieve optimum throughput between the V20zs. Currently the throughput achieved at Sunnyvale is 5.7Gbits/s for the Neterion NICs. At SC2004 we achieved 7.4Gbit/s with 2.4GHz V20zs and S2io and Chelsio TOE NICs. We need to understand the discrepancy.
The remaining two V20z's are still at SLAC. They are being configured with Chelsio TCP Offload Engine (TOE) 10GE NICs loaned from Chelsio. We are also working with Sun to install Solaris 10 on one of the hosts so we can evaluate its performance for 10GE.
Next week (4-8 April 2005) USN plans to make the physical installation of the UltraScienceNet at Sunnyvale. Initially the USN circuits will support IP traffic so we will plug it into the Cisco 6509.
We published a paper on Characterization and Evaluation of TCP and UDP-based Transport on Real Networks at the Protocols for Fast Long-Distance Networks in Lyon in February 2005.

5.3 Computer Science & Middleware

5.3.1 Condor

5.3.2 Globus – ANL

5.3.2.1 MDS work

The MDS4 scalability analysis work continued, looking at both multiple client work and stability. The most noted result in the past quarter was that an MDS4 index server stayed up over 2 weeks, serving over 98 Million requests, and only stopped running when the server was rebooted by mistake.

Discussions have begun with individual groups on transitioning from MDS2 to MDS4. While MDS4 does not, as yet, present new radical new functionality to the end user, it lays the foundation for new features, such as registration, which will allow entities to request notification of changes in information instead of having to poll for such changes. Interfaces to additional data sources, such as Hawkeye and RFT, are already available as part of the beta release. Discussions have taken place with the Clarens group about possible collaboration and to clarify interactions. Also, John McGee will be participating in the OSG calls to make sure these needs are being met.

5.3.2.2 Data Management

The data management components in Globus that are of interest to PPDG are GridFTP, the Reliable File Transfer (RFT) Service, and the Replica Location Service (RLS). Work over the last quarter for all components has focused on debugging, hardening, and documentation.

GT4 will include a completely re-designed and re-implemented, but 100% backward compatible GridFTP server. There are two major new features which include striping support (for bandwidth greater than a single node can achieve) and IPV6 support. The new server is slightly faster than the old server in single node performance and hit 27 Gbs on the TeraGrid 30 Gbs network (90% utilization) with 32 nodes memory to memory. That is nearly perfect linearity. We achieved 17.5 Gbs disk to disk limited by the storage system. The new server is already much more stable than the old server. It is supported in threaded builds which the old server was not. In one particular GriPhyN work flow, the old server had O(1000) failures, when the new server was installed, it had zero failures. We have also had a single host supporting 1800 simultaneous clients. The documentation provided for the new server is greatly expanded and much more detailed than it has been in the past.

The RFT service has also been significantly hardened. Originally designed for large (1000's of files) transfers, GRAMs use of RFT for staging pushed us into an entirely new operational regime. We are now handling large numbers of small requests coming over a short period of time. This exposed some race conditions that have been tracked down and resolved. We also recently completed a move of the Sloan Digital Sky Survey DR3 archive (900,000+ files, 3 TB) with a single command. RFT provides a fully Web Services Resource Framework (WSRF) compliant interface for data movement. It exposes a variety of resource properties which can either be polled or subscribed to. This includes overall service statistics, request statistics, and file by file state changes.

The Replica Location Service has also had significant hardening. It has been exposed to extremely high loads in both test and production environments. The LIGO experiment uses GridFTP and RLS to move a TB of data per day. They maintain 40 million files, 4 million LFNs across 10 sites. There is no WSRF compliant interface to RLS at this time. We are however prototyping WSRF compliant higher level reliable replication services that combine data movement and RLS updates.

5.3.2.3 Execution Management

The Web Services compliant GRAM (WS-GRAM) has made extraordinary strides in terms of performance and scalability through a focused effort from the entire Globus team. We are currently performing significantly better than GT3 GRAM. We have significantly improved over GT2 GRAM performance in some cases (such as head node load), and are about even in others (such as job throughput).

We have also been working with CERN (Kate Keahey) on a prototype of a new virtual machine workspace technology. Early experience has been positive and we will be proceeding with further development.

5.3.3 SRM

Participants: Alex Sim, Junmin Gu, Viji Natarajan, Arie Shoshani, Kurt Stockinger

DRM in VDT

We supported several people in deployment of DRM out of VDT as part of OSG activities: Eric Hjort (LBNL) for running the STAR analysis scenario, Ed May (ANL) as part of ATLAS, Jorge Rodrigez (U of Florida) as part OSG, Neha (Fermilab) as part OSG, and Elisabeth Atems (Wayne State) as part of the STAR experiment.

A new SRM testing tool

We developed a new SRM testing tool, called SRM-tester, packaged it, and deployed as part of DRM in VDT. Enhancements were made to the SRM-tester for testing SRM-dCache. Three SRM systems were tested using the new SRM-tester: DRM (VDT version), FNAL SRM-dCache and BNL SRM-dCache (with Viji and Junmin).

New option added is to provide source url, target url, and logging path. Previously, these were fixed.

An important addition to the testing tool is for checking the interoperation of DRM and SRM-dCache. This tested the peer-to-peer SRM interaction getting files from/to each other's space.

Enhancements and new features added to DRM/HRM

We added new features to prevent errors.

- This includes checking that a file is indeed in cache when a request to a cached file is made. This is done in case that files are removed by an external agent, not the DRM.
- We added an explicit remove function from user space. Now, "release" does not cause a file to be removed after it is released.
- A new configuration option for the drm*maxFileSize was added to allow for files that are larger than the 2 GB default.
- A new policy option was added. If drm*disableUserReleaseEnforcement=true, then files whose lifetime expired are not put back into the request queue.

- Startup time for large DRM cache was reduced by using of hash-map on the entries in the files log.
- Support for the HTTP transfer protocol was added in addition to the GSI-ftp and bbftp and ftp protocols.
- A new Startup script was developed. An appropriate entry in the VDT directory $DRM_HOME will start up the DRM service when the machine is booted.

Enhancements to the web service (WS) component

- accommodate all JDK versions with java.endorsed.dirs option
- added ability to use service certificate (previously it was only possible to use a host/personal certificate)
- The hostname and ip address is now collected automatically from the OS (previously it had to be provided by the user starting up the WS)

Developed a new DRM configuration generator

The number of items in the DRM configuration files has grown over time with multiple options. We developed two versions of a configuration generator: one using a GUI and one using only an interactive shell.

Developed new client DRM components

The command line client components are currently the most heavily used. The basic functions of srm-ls srm-copy were initially provided. However, now the usage is more sophisticated and it required the addition of two new clients: srm-release and srm-remove for both linux and solaris. We also added an exit code in case of client's crash at the request of users.

Meetings

Alex attended the OSG Integration meeting in February at the Univ. of Chicago. OSG deployment, installation and testing took place.

Alex and Arie attended GGF13 in Korea. In addition to running two sessions as part of the Grid Storage Management (GSM) working group, we gave talks at KISTI and a nearby university about SRMs and their uses.

Reports

A revised version of the Replica Registration Service (RRS) document was generated reflecting comments we received. It will shortly be distributed to the PPDG list.

5.3.4 Caltech

Work continued on the development, and deployment of the Clarens web services framework and services built using this framework.

Technical developments

Additions were made to the base Clarens server so that each instance advertises itself using the rapidly developing Discovery service, as well as the local multicast DNS protocol.

A major effort was undertaken to adapt the current web browser user interface to work with more types of browsers, including Apple Safari, Opera, and Konqueror. Work is continuing to add support for Internet explorer.

Part of this effort required the implementation of a new data serialization protocol that can be used by browsers that do not support XML, namely the Javascript Object Notation. Using this serialization format data objects can be very efficiently de/serialized using the browser's built-in Javascript interpreter.

In response to limitations in Grid middleware used in the CMS Monte-Carlo Processing Service (MCPS) at FNAL, the proxy credential service was expanded to allow remote users to make these credentials available in a more flexible manner. Requirements from the MCPS also prompted an expansion of the support for HTML form-based file uploads and remote service invocation. This enables users to build fully-featured Grid service interfaces using standard HTML forms in addition to the more complex Javascript methods hitherto available.

As part of the HotGrid project undertaken in collaboration with the Caltech Center for Advanced Computing Research (CACR), a simple to use web interface as well as a back-end HotGrid service was developed and a prototype deployed on the CACR TeraGrid node. This service allows user's in the Astronomy community to obtain limited Grid certificate credentials without administrator intervention using only a valid e-mail address. This effectively creates a class of weakly authenticated users that can be restricted to use only pre-determined amounts of CPU and storage resources. This so-called "graduated" security implementation allows large numbers of users to experiment and become familiar with the Grid and it's security environment, namely using certificates for authentication in a web browser and doing job submission etc.

As part of a quality control effort to ensure the availability of running servers and functionality of new services as they are developed, an effort was undertaken to do regular automated tests of a set of Clarens servers using two independently developed testing frameworks. The Python-based testing framework was significantly expanded to allow for greater code coverage, while the Java-based framework was updated to make it compatible with both Python and Java-based server versions.

A new release of the Clarens server and base services was also made during February, and currently more than twenty instances of the server are currently running worldwide.

Collaboration

I was involved in the Open Science Grid as a co-chair of the Monitoring and Information Systems group, where a monitoring environment and implementation for the consortium is being developed.

As part of the PPDG common project, a version of the Clarens server is being integrated into the Virtual Data Toolkit deployed as part of many Grid installations. Two updated versions of the Clarens server was also incorporated in the CMS Distributed Production Environment (DPE) for use by CMS Grid sites.

I attended the regular CMS week from March 14 to 18, as well as bi-weekly CMS Analysis Program (APROM) meetings held at CERN using VRVS. I also attended an MCPS integration meeting held at Fermilab during February.

The CMS Physics Shell (PHYSH) virtual filesystem interface and associated RefDB, PubDB and Phedex services developments was continued and expanded during the period, with a total of 1.5 FTEs dedicated to its development. A significant part of my own time was used to support this development, in the form of co-development, trouble-shooting and adding needed features to the

Clarens server.

Continued support was also given to the CAVES versioned analysis environment being developed at the University of Florida using a Clarens client/server architecture within the ROOT analysis environment.

A ROOT Clarens client developed for the SC2004 conference has been checked into that project's source repository, and is being maintained by the ROOT developers. It will thus be available to all users of ROOT by default sometime in future.

A Clarens-based job wrapper shell was co-developed with the CDF experiment for their User Analysis Facility (UAF). This shell allows users to obtain limited secured access to the runtime environment of jobs running on worker nodes. This allows users to obtain output files, do process management and associated tasks.

5.3.5 SRB

Reagan Moore: I participated in the 13th Global Grid Forum meeting in sessions related to remote procedures, file systems, and preservation. The Grid standards in these areas are being strongly influenced

by both academic prototypes and commercial systems. As yet, there is no Grid standardardization effort for data grids, the software infrastructure that self-consistently manages distributed shared data collections. On the other hand, additional projects are now using the Storage Resource Broker to support federation of data grids on an international scale. These include the KEK high energy physics experiment (federation of 7 data grids across Japan, South Korea, Australia, Poland, and the US) and the Worldwide Universities Network which will have links to the UK e-science data grid and high-energy physics collaborations.

Wayne Schroeder: SRB version 3.3 was released on February 18 and SRB 3.3.1 on April 6. These include the interfaces between GridFTP and SRB developed last quarter, support for GT3 GSI (in addition to the long-existing SRB support of previous versions of GSI), an interface from MDSPlus to SRB, and many other new features and bug fixes. Release notes are available at
http://www.sdsc.edu/srb/CurrentSRB/ReleaseNotes3.3.1.html and
http://www.sdsc.edu/srb/CurrentSRB/ReleaseNotes3.3.html

The pre-release SRB testing procedures have been greatly enhanced, allowing us to find and resolve a number of bugs before distributing the code. I developed a master script and various support scripts to do automated testing in various modes, performing different subsets of tests, on multiple platforms, and in various configurations. We added many new tests and also updated and revived some existing testsuites and integrated them into this system. We are in the process of formulating this to run automatically and continuously, via the Tinderbox system, on the SDSC QA lab hosts (using current snapshots from our CVS repository). Bugs and feature requests are now being tracked via bugzilla (http://srb.npaci.edu/bugzilla/).

The SRB installation script and administration tools have also been enhanced in various ways.

See the release notes for more information.

6 Appendix

6.1 Milestones

Target Date	Completed	Responsible Task	Milestone Deliverable
May-04	√	PPDG-CMS	Enable clients to access the Pool metadata catalog via the Clarens web service environment
Jun-04	√	PPDG-SAMGrid	Test JIM job scheduling at CDF
Jun-04	√	PPDG-SAMGrid	JIM job scheduling on remote resources for next round of D0 reprocessing
Jun-04	√	PPDG-SRM	SRM Specification V3
Jun-04	√	PPDG-BaBAR	SRB based production system for data distribution
Jun-04	√	PPDG-CS-11	Implementation of the CS11 Dataset Query Service by at least Caltech and SLAC
Jun-04	[1]	PPDG-CS-11	Complete specification of 3 more of the standard interfaces for Analysis Services
Jul-04	√	PPDG-OSG-Authorization	Specification and Requirements for Authorization Version 1
Jul-04	[2]	PPDG-OSG-Monitoring	Specify Requirements V1
Jul-04	√	PPDG-STAR	Grid-based Monte-Carlo production start to migrate from centralized Tier-0 center to a Tier-1 distributed model.
Jul-04	[1]	PPDG-STAR	Batch oriented user analysis will start.
Sep-04	partial	PPDG-CMS	Deployment of second version of lookup service for dynamic discovery. Multiple instances of service will be deployed.
Sep-04	[1]	PPDG-OSG-Troubleshooting	Specify Baseline Capabilities

Target Date	Completed	Responsible Task	Milestone Deliverable
Oct-04	√	PPDG-OSG-Accounting	Specification and Requirements for Accounting Version 1
Oct-04	[1]	Oct-04 PPDG-OSG-WMGT	Process Wrapper (ProRap) for remote execution of HENP applications V1
Nov-04	√	PPDG-SRM	SRM Implementation V2 Interfaces
Dec-04	[3]	PPDG-OSG-Accounting	Implementation of Accounting V1.0
Dec-04	[4]	PPDG-OSG-Monitoring	Monitoring of the system infrastructure across all the Laboratory facilities
Dec-04		PPDG-ATLAS	Deploy prototype sufficient to assess POOL-based ATLAS applications
Dec-04	√	PPDG-JLAB	A prototype implementation of SRM, based on specification 2.x,
Dec-04	√	PPDG-OSG-RRS	Replica Registration Service Specification V1
Dec-04		PPDG-OSG-WMGT	Enhance the scalability of Condor-G (support 10^5 jobs)
Dec-04	√	PPDG-SAMGrid	JIM job scheduling as part of the D0 RAC infrastructure
Dec-04	√	PPDG-SAMGrid	Prototype services which use more general Grid resources, in particular run tests for CDF and D0 on OSG resources
Dec-04		PPDG-CS-11	Multiple Implementations of the three interfaces selected
Dec-04	√Mar-05 Testing	PPDG-DMGT	Use of Enhanced GridFTP
Dec-04		PPDG-OSG-Monitoring	Instrumentation of selected data transport and replication applications
Jan-05	√ Mar-05 OSG deployment	PPDG-OSG-Authorization	Implementation V1.0
Jan-05	Partial: >2000 CPUs on OSG ITB	PPDG-OSG	Multi-experiment use of 2000 CPUs, 100TBytes
Jan-05		PPDG-BaBAR	Applications that publish in the RLS the SRB MCAT knowledge on the data location.
Apr-05		PPDG-OSG-WMGT	Process Wrapper (ProRap) for remote execution of HENP applications V2
May-05		PPDG-OSG-Monitoring	Monitoring and instrumentation roadmap
May-05		PPDG-CS-11	Complete specification of remaining ~15

[1] We identify this as a slipped milestone and will review this task during the next quarter.
[2] We identify this work as being done elsewhere in Open Science Grid.
[3] Storage accouning infrastructure added to MonaLisa. In test and considered for deployment on OSG ITB.
[4] Extensions to MonaLisa for monitoring of common grid infrastructure.

6.2 Meetings

PPDG participates in many OSG Technical Group and Activity meetings. The PPDG Executive Team is again sponsoring a meeting in the OSG Blueprint series with a US LHC Technology Roadmap meeting in Madison in May.

OSG Integration and Operations-Provisioning Meeting, February 15-17, 2005
http://www.ivdgl.org/osg-int/workshop.html.