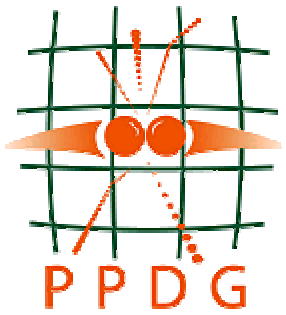
Particle Physics Data Grid: From Fabric to Physics

Quarterly Status Report of the Steering Committee,

October - December 2004

31 January 2005



1 Project Overview

1.1 Highlights

During the last quarter of 2004 the Steering Committee welcomed the new PPDG Technical Coordinator, Dane Skow, who is responsible for the management of the Common Project and is an additional member of the executive team. Dane started to ramp up his effort towards the end of the quarter. Work proceeded for the common project and culminated in a successful Technical Meeting at Brookhaven in early December.

PPDG ramped up its contributions to the Open Science Grid Technical Groups and Activities towards the evolution of Grid3 to deployment of the Open Science Grid scheduled for Spring 2005. Additional services for Storage Access and Management, Site Identify Mapping, Role based Authorization, and Discovery will be added to the core infrastructure as part of the OSG deployment.

Many demonstrations at the SC2004 conference showed the progress being made in use of Grid infrastructure and middleware by our collaborators. As part of setting the Bandwidth Challenge record of 101 Gbps[1], CMS demonstrated a globally distributed data analysis using the Grid Analysis Environment.

Many of the items in this Quarterly Report, as well as the Sc2004 demonstrations, many Chep2004 papers with PPDG participation, and the ongoing deployments of the experiments, give evidence that PPDG collaborating experiments are increasingly valuing the Grid tools and are providing additional (non-externally funded) effort towards meeting their needs for distributed data processing and analysis on the Grid.

Dane Skow represented PPDG at the National Collaboratories Opportunities for Distributed Science Workshop[2] and submitted a White Paper [PPDG-48] on behalf PPDG.

1.2 Papers and Documents

Reports, Documents and Papers		Date/Version
PPDG-48	PPDG White paper Input to the National Collaboratories Distributed Science meeting	doc, pdf
PPDG-47	Selected Accomplishments	doc, pdf (22Nov04)
PPDG-46	Development and use of MonALISA high level monitoring services for the STAR Unified Meta- Scheduler (SUMS)	pdf

2 The Common Project

A face to face meeting of the Common Project participants was held at Brookhaven National Lab on December 8-9, 2004. The meeting gathered technical leads from all the PPDG Common Project activities for purposes of sharing experiences and technical issues and reviewing status and dependencies for the Spring 2005 OSG deployment. Discussions about the migration to Web Services based services identified a few pressing issues and some decisions were taken. First, a decision was taken not to base Web Service development on the GT3 environment given its ties to the deprecated OGSI tooling. Recommendations at this point are to prefer WS-I profile tools, but not require a common development environment. Reviewing status (including the GT4 release) is expected for the Summer quarter of FY05 after the OSG deployment.

Second, discussions among the developers, and with our EGEE colleagues, revealed consensus that GSI delegation over HTTPS (aka HTTPG) is deprecated and that separating the delegation function from authentication is the correct design decision moving forward. There are efforts underway to address this

[1] http://ultralight.caltech.edu/sc2004/BandwidthRecord/
[2] http://dsd.lbl.gov/Collaboratories/NCWorkshop/agenda.htm

within Globus, EGEE, Clarens and others. We agreed to work with our EGEE colleagues to identify compatible solutions and anticipate work in this area for the next quarter

Clarens discovery service has deployed a test release at Caltech and is getting feedback from the CS-11 group. Small modifications based on this feedback are underway and conversations begun with the VDT team for packaging for OSG, and with the OSG Integration activity to develop an integration test plan (per the OSG process). A meeting was held between the Clarens and Globus MDS developers reviewing interoperability with the Clarens Discovery Service. In general, the latter is much more focused on a specific type of monitoring data access and dissemination, namely for web service locations and attributes. The former is a much more general monitoring framework. The consensus of the participants was that it should be possible to provide a two-way bridge between the two systems to share the discovery information using existing publish and polling facilities in MDS. An implementation was postponed until a formal release of the Globus toolkit version 4 is made. This project is on track for the Spring release.

The SRM-DRM package, including the LBL DRM implementation of an SRM file service based on a single filesystem disk store, is also well along track for OSG deployment. The service is currently integrated with the VDT 1.3 release and has an integration test plan being worked in the OSG Integration activity. Development of a configuration tester and compliance tester are both underway and on track for a Spring release.

The Privilege Project is well along development needed to support Role Based Access Control (RBAC) for Compute Elements. Unit testing has completed and development of an Integration test plan has begun. The client modules were rewritten to remove dependency on the Globus Toolkit v3.x web services environment per the decision taken at the face to face meeting not to rely on this deprecated environment. Deployments of GridFTP using the gatekeeper authorization should also be able to use these RBAC features. Development of integration with the dCache Storage Element has begun. There are general issues being raised about integration of authorization and provisioning functions and they are under discussion now. This is likely a more general issue than storage, but has immediate needs there. See these references for additional details about this authorization and identity management work[3].

It was agreed at the December OSG Technical meeting to base the near-term accounting function on the support of VO-level accounting from the information reported to the MonALISA monitoring system – basically a continuation of the Grid3 accounting system – and focus development activity on understanding requirements and design for the next stage. An extension to the MonALISA schema to support a common accounting set of information for storage systems has begun with target completion prior to the Spring deployment. It was agreed that the primary function of the accounting system is to support the economic model of the Grid. With that in mind, encouragement was given for OSG to develop an economic model for its operations.

3 CS-11 Data Analysis Working Group

Several papers on grid-based data analysis written for the CHEP 2004 conference proceedings included "*ATLAS Distributed Analysis*", D. Adams, "*Grid Enabled Analysis for CMS: prototype, status and result*s", F. Van Lingen, "*Interactive Data Analysis on the Grid using Globus 3 and JAS3*", T. Johnson.

The latter paper describes collaborative work with an SBIR funded project with Tech-X Corp with one person at SLAC. Also Tech-X has begun collaborating with STAR on a user level job description language, a continuation of previous JDL work in PPDG.[4]

The Clarens Rendezvous service[5] is being discussed and tested within the group, and is to be contributed at the Discovery Service for OSG.

[3] http://grid.racf.bnl.gov/GUMS/,
http://computing.fnal.gov/docs/products/voprivilege/documents/motivation-2005-01-14.pdf,
http://computing.fnal.gov/docs/products/voprivilege/documents/transition-to-privilege.html
[4] http://www.star.bnl.gov/STAR/comp/Grid/scheduler/rdl/index.html
[5] http://hepgrid1.caltech.edu/GAE/services/rendezvous/

4 Collaborations

4.1 Leveraging Experiment Effort

The Experiments collaborating on PPDG all have effort engaged in deploying data distribution and distributed data processing and analysis components as part of their data handling strategies and roadmaps. As planned from the start of the project, the PPDG funded effort is integrated with the main experiment software and development teams in order to facilitate both flow and leverage of common grid technologies within the experiments.

Examples of this include:

1) ATLAS Distributed Computing Environment. The PPDG effort at ANL (Jerry Gieraltowski) is working as a core part of the US ATLAS DCE team which includes iVDGL, GriPhyN as well as US ATLAS funded members. DCE is used by the common ATLAS production scripts where jobs are submitted across three Grid infrastructures - NorduGrid, LCG and Grid3 - and data is collected from all execution sites and archived through Brookhaven to the CERN tape store.

2) CMS. The PPDG effort at Fermilab (Anzar Afaq) works as a core part of the US CMS production team who develop and run production across the CMS Tier-1 and Tier-2 sites as well as on Grid3. Several of the products developed by this team are used by the whole CMS collaboration as part of the simulation production services (e.g. mcrunjob).

3) STAR. The PPDG funded effort working on STAR at BNL (Levente Hajdu) is gradually converting the STAR data processing systems to make use of the Globus and Condor-G technologies for use by the rest of the STAR production support team.

4) BaBar. The US based BaBar PPDG effort (Wilko Kroeger) is working with the global BaBar collaboration and the analysis support teams on the distribution of full datasets from SLAC to 3 sites in Europe and integration of the US developed software with the European software stack.

5) D0. D0 supports a fully distributed data handling, distribution and job scheduling system. The contributions are merged into the core SAMGrid system with the PPDG effort integrated into the larger development and support project.

6) JLAB. The JLAB sponsored effort works together with the off-site groups to ensure that the grid based data access and distribution is effective and usable at the remote universities.

4.2 Open Science Grid

The PPDG extension proposal includes contributions to the Open Science Grid Consortium through the Common Project, through facility contributions to the resources presented to the common grid infrastructure, and through engagement of all teams, including the Executive Team, in the OSG roadmap. Evolution of Grid3 into the first deployment of the Open Science Grid is planned for Spring 2005. PPDG is contributing, chiefly through the Common Project, to the extension and deployment of new services on the common infrastructure.

The Charter[6] of the Open Science Grid has been agreed to by the Governance Technical Group.

The Blueprint[7] has been published as an OSG reference document detailing the principals that will guide the development and building of the Open Science Grid infrastructure.

A Security Incident Handling and Response Plan[8] has been agreed to in collaboration with the European EGEE/LCG projects.

[6] http://computing.fnal.gov/docdb/osg_documents/0000/000025/004/Charter.pdf
[7] http://computing.fnal.gov/docdb/osg_documents/0000/000018/004/OSG-Blueprint.pdf
[8] http://computing.fnal.gov/docdb/osg_documents/0000/000019/002/OSG_incident_handling_v1.0.pdf

Documents covering the Operations Model, the Deployment and Operations Plans, and the By-Laws of the Consortium are in progress.

Among other contributions, iVDGL is contributing leadership and effort to the Integration Activity. PPDG is contributing to this activity and to the "Release Process" which has been defined to move new services from development, through integration, testing and validation, into production.

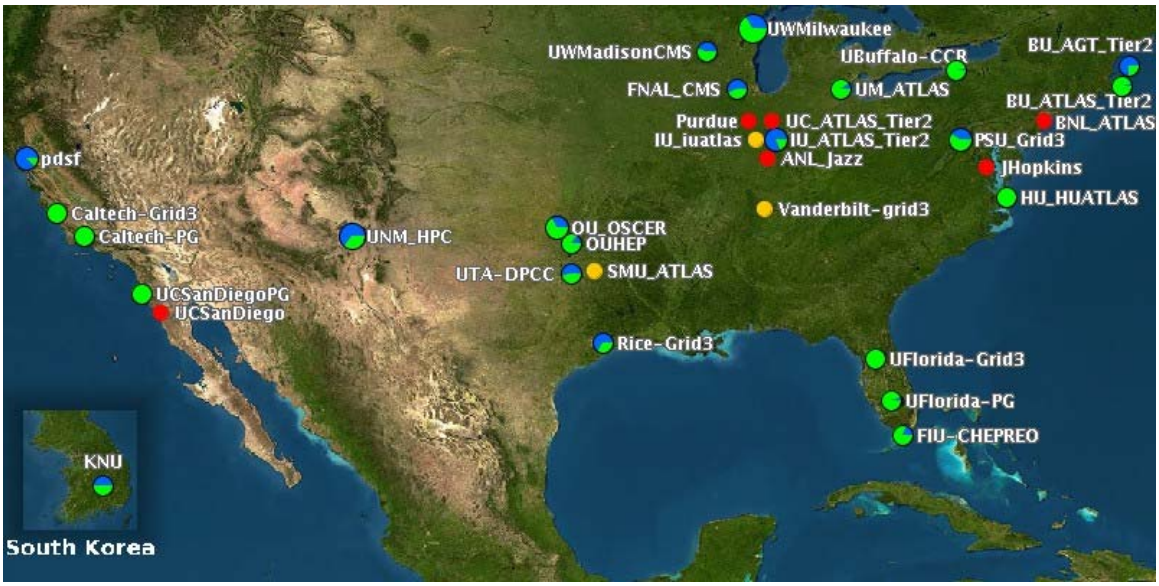
A presentation[9] was given at the TeraGrid GIG review and a list of collaborative milestones proposed.

The PPDG proposal roadmap of further engaging the DOE Laboratory Facilities as part of a common grid infrastructure is making progress, with co-chairmanship of several groups from the Laboratories and commitment of the Lab Facility infrastructures to present resources and provide interfaces to the common grid infrastructure.

4.3 Trillium and Grid3

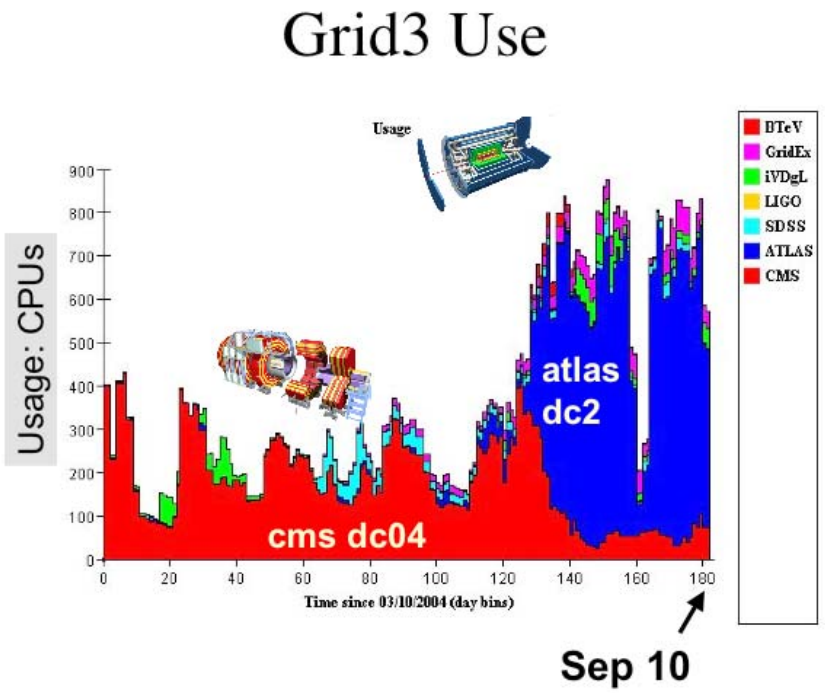
The stakeholders continue to benefit from the Grid3 infrastructure. The number of sites grew from 26 to 36 during 2004, and the number of available CPUs from 2600 to 3600. The Grid3 infrastructure remained stable and in this mode has required operational effort at the level of about 2 FTEs additional effort for Grid wide operations, and 2 FTE effort in support of the software stack.

Status of Sites on Grid3 January 24th 2005



[9] http://computing.fnal.gov/docdb/osg_documents/0000/000027/002/GIG-OSG.pdf

Illustration of shared use of resources in the transition from the CMS to ATLAS data challenges.



4.4 EGEE, LCG

PPDG continues its focus on collaboration with the LCG and EGEE both through the PPDG project and through the Open Science Grid Consortium.

Participation in the Joint Security Group bore fruit through the development of a common Incident Response and Handling Plan. Joint work on User Registration, User VO Registration Management, Security Policies and Certificate Management continued.

During the OSG Blueprint meeting in October a joint session was held with the ARDA workshop, ARDA[10] being the LCG/EGEE project to accept the EGEE/gLite[11] middleware stack and demonstrate end to end experiments distributed data analysis systems.

Miron Livny continued contributions to the gLite project, and developments of the Condor Project are contributing to the gLite service developments.

4.5 DOEGrids PKI

Participation in the DOEGrids PKI continued this quarter. An issue arose in the context of the transition to web services and the use case of agent services. The concept of service certificates for software agents,

[10] http://lcg.web.cern.ch/LCG/peb/arda/
[11] http://glite.web.cern.ch/glite/

acting as both client and server, is not well described in the certificate policy resulting in some ambiguity in procedures for how to validate requests for service certificates. A discussion of this was started in the PMA, and other forums, and an update and clarification of the policies and procedures document will occur in early 2005.

Additionally, a set of command-line scripts[12] for interacting with the CA for getting new certificates was prepared and packaged, based upon previous work done in the DOE Science Grid and FusionGrid projects. These are contributed by PPDG to the DOEGrids PKI and will be distributed as part of VDT, beginning version 1.3.1.

5 Single Team Reports

5.1 Experiments

5.1.1 ATLAS

Jerry Gieraltowski: The production system in Grid3 follows the ATLAS approach that involves a supervisor interacting with the production database and one or more executors that manage all Grid interactions. The Grid3 executor system, Capone, communicates with the supervisor and handles all the interactions with Grid3 resources and services. Capone utilizes the GCE-Client middleware package to construct an abstract directed-acyclic graph (DAG) which is used to define the workflow of the job. Previous versions of the GCE-Client package incorporated specific handling for each type of transformation/executable requested to be executed by the user. These transformations were specific to the tasks of event generation, simulation, digitization, pileup, and reconstruction. My work in the past quarter has been to construct a general transformation which can be used by the user to submit a job request to execute any type of executable. The user can specify via (name, value) pairs the name and location (both local or remote) of the actual executable script the user wants to execute, sets of input parameters expected by the user executable, sets of desired input and output data files (both local and remote), and sets of outputs to be automatically returned to the user if so desired. This general transformation has been shown to be backward compatible with all of the existing ATLAS transformations used in DC2 and provides an easy way to extend the functionality of Capone to the general user community. Further extensions of this work to incorporate an interface to DIAL and ADA are in the planning stage. Additional work with GriPhyN project personnel is being discussed to extend the functionality to include new storage resource management (ex., SRM) services.

5.1.2 BaBar

Wilko Kroeger: An SRB zone that provides access to the BaBar conditions snapshots (set of coherent files) has been setup. The loading of BaBar conditions snapshots into this SRB zone has been automated and a tool that makes it easy to download a snapshot has been written.

As usual, some effort was spent to support the data distribution of BaBar root-files from SLAC to IN2P3 using the SRB. We are transferring constantly root files from SLAC to IN2P3. The transfer works very smoothly, but it isn't completely automated yet. We also observed some infrequent problems with the SRB servers that require a restart. We did some debugging of the problem but couldn't resolve it yet. I also started testing how to upgrade our production SRB to a newer version. Some changes for our tools are needed and we also would like to incorporate some of the new features of the latest SRB version.

5.1.3 CMS

Abhishek Rana: Provided dCache deployment effort for the PEAC demo by Caltech/UCSD/FNAL/MIT at Supercomputing'04. Also started to work on understanding authorization mechanisms in a distributed storage management solution, and related development work in collaboration with dCache collaboration. Helped organize the OSG Technical meeting held locally.

[12] http://www.ppdg.net/RA/cert-scripts/

Suresh Singh

The tasks accomplished in this quarter can be grouped in two categories.

1. Maintaining Caltech Grid Clusters for productions (US CMS/iVDGL work)

The DOEGrids host certificates of two main Caltech clusters (Caltech-PG and Caltech-Grid3) participating Grid3 productions had been renewed just before their expiration. A 1.8 Terabyte NAS storage unit with XFS filesystem has been added to Caltech-PG. All the nodes in the Catech-PG have been upgraded to Linux 2.4.28 kernel. The two Dell switches making Caltech-PG have been upgraded to latest switch operation code supplied by Dell.

The other important aspects were to deal with security incidences of Caltech-DGT and MonALISA server. The headnode of Caltech-DGT (citcms) was found to be running Internet Relay Chat (IRC) BOT after the intruder had accessed the system with a compromised user's ssh key. The compromise of the ssh key had happened in another system prior to accessing the Caltech-DGT. However, the compromise was not found to be in the ROOT level. After removing IRC Bot related files and compromised user's key, the problem had been solved. Also the firewall of the system has been made much tighter by using stateful firewall.

Similarly, the MonALISA server at Caltech had been observed to be sending spam emails to other mail servers at Caltech. This server is under intense scrutiny by security experts at Caltech including Iosif Legrand. Any mail traffic coming to SMTP port (25) has been blocked at ITS router. After having done that, no further incidence of spamming has been observed on this system. A backup MonALISA server is in place should this server to be taken off-line.

Internetworking between Caltech Tier2 and Caltech TeraGrid has been done by connecting respective two Force10 switches with a single mode fiber to their 10 Gigabit interfaces. This will enable us to use TeraGrid resources in future.

2. Assisting UERJ.

I had spent significant time to help HEP group of University of Rio in Brazil remotely to build their Linux cluster just before SC2004. I had provided necessary advices to design, construct and operate their cluster using ROCKS cluster management software. In order to make their cluster available for SC2004 demonstrations, I had installed Grid3 software and configured cluster nodes for Condor. This was in turn used for distributed grid analysis demonstration.

Michael Thomas: Oct.-Dec. '04 quarter was focused on E&O and the Clarens Discovery service. A lot of time was spent working with the NUST, Pakistan team in their Clarens contributions. UERJ (Rio de Janeiro) also occupied quite a bit of time with the setup of their new computing cluster. The Clarens Discovery service was improved and made public.

E&O

As part of his ongoing PhD program, I mentored Ashiq Anjum from NUST in his design and prototype implementation of a Dataset Location Service for CMS. This service takes a logical filename as input and returns optimal locations for obtaining the data. Optimal data locations are determined by a matchmaking algorithm that makes use of monitoring data from MonALISA. While the initial matchmaking algorithm was trivial, the service was written so that it could support more complex algorithms as they are needed. I worked closely with Ashiq during the design and implementation of this service, which resulted in a presentation at the Nov. 26 APROM meeting at CERN.

Another NUST student, Faisal Khan, came up with a new implementation of the Clarens Discovery service that uses a UDDI registry to store service publications. This provided an important link from Clarens to UDDI, allowing Clarens to make use of the standard service registry in use by the Web Service community. This UDDI implementation of the Discovery service was added to the jClarens codebase.

A significant E&O effort involved assisting the UERJ (Rio de Janeiro) HEP department in becoming a real Tier2 site for CMS. I made two trips to UERJ during this quarter and worked closely with their software and hardware teams in order to design and implement a cluster architecture for their site. Both MonALISA and Clarens were installed on their new cluster, which enabled them to participate in SC2004 as part of the

job monitoring demonstration (more on SC2004 later.) The two site visits were followed up with numerous email conversations in an effort to continue to make progress on their involvement in CMS production.

Discovery Service

The Clarens Discovery Service received a lot of attention this quarter. The public Discovery Service installation (http://discover.gridservice.info/) was officially presented to the PPDG CS11 working group at the Oct 14 meeting. One request that came of this meeting was the desire to publish a service with additional key-value data. This feature was added to the discovery service WSDL (http://discover.gridservice.info/wsdl/rendezvous.wsdl) in late December, but has not yet been implemented.

Balamurali Ananthan discovered some nasty bugs in the SOAP binding of the Discovery service. The server-side stubs were not 100% compatible with the client side stubs generated from the WSDL. This was fixed at the end of the quarter and is pending further testing.

The discovery service was also presented at the PPDG Common Project meeting at Brookhaven on Dec. 8-10. Another feature request came out of this meeting: the ability to restrict who is allowed to publish services to a VO. This is a very useful security feature and will be added to an as-yet undetermined release of the Discovery Service.

The Discovery Service has been committed to the OSG-0 Spring rollout. I'm working with Rob Gardner to be a guinea pig for the "Service Readiness Plan".

jClarens

As mentioned earlier, a UDDI implementation of the Discovery Service was added to jClarens. During this work significant improvements were made to the way that add-on services get built and packaged with jClarens. It is now possible to build new jClarens services with installation RPMs by following the examples of the UDDI and Jini implementations of the Discovery service. It is no longer necessary to install the entire jClarens source code package in order to build and test add-on services.

In collaboration with Tahir Azim from NUST, a RMI binding was added to jClarens. This allows optimized access to jClarens services from localhost connections. We also added support for accessing services without any authorization. This gives the site administrator the ability to let anyone access a particular set of services without the need to first login.

Other work

As part of the SC2004 demonstration, the Clarens/BOSS adapter was upgraded to work with the recently released BOSS v3.4.0. I also generated Openpkg RPMs for BOSS in order to simplify installation of this adapter.

Work was started on a SRM adapter for Clarens. Abhishek Singh Rana helped to get dCache with a SRM door setup on the GAE testbed cluster. Work will continue on this adapter next quarter.

Meetings and Conferences

I attended almost all PPDG-CS11 phone conferences, except where it conflicted with another conference. I prepared and ran a Distributed Processing / Job Monitoring demonstration in collaboration with USP (Sao Paulo, Brazil) and UERJ (Rio de Janeiro) to show job submission and job monitoring across multiple sites. I attended the PPDG Common Project meeting at Brookhaven on Dec. 8-10. A presentation on Clarens and the Discovery service was given.

Anzar Afaq

CMS Production started to gear up for DST Production in early January 2005. DST Production is next stage in chain of Production cycle. This important step required key changes in MOP (USCMS Grid submission tool) and McRunjob. Beside basic support for the GRID in these tools, McRunjob itself required a major development to support Dataset Manipulation and Publication Tools, deprecating DC04 era temporary developed tools. This was solely designed and implemented by me and called Publish Service. Later Publish Service was integrated to McRunjob Job Creation infrastructure. DST Production was tested to work fine even with MOP, instruction and documentations were provided to users.

A sideline task was to Publish Datasets for analysis users, I dev eloped the tools, which are later used and enhanced to Publish these datasets by USCMS. These datasets are being used by CMS Users within FNAL and other USCMS.

SC2004 Demo Preparation: I coordinated the demo deployment activities, working with developers and Facility. A CLARENS based MOPService is deployed and integrated on demo machine, and tested to work with other components. We could submit jobs remotely and browse results (root files). This task involved, development of MOPService, testing and deployment of DPE based CLARENS releases and working with rest of the team of deployment issues. The SC2004 demo was finally a success.

USCMS is moving to a new tool called MCPS for Private user analysis and production. There are plans to use this tool for Official CMS Production as well. I am actively working to keep MOP and MOP-interfaces up to date to work in changing environments. MOPService developed in CLARENS is being viewed as new interface to GRID for MCPS, while maintaining current non-webservice interface.

I worked on a prototype tool called MOP-db to track MOP GRID jobs during execution and manage them as per user needs. MCPS will have an enhanced flavor of MOP-db where Jobs submitted by MOP will be tracked for debugging and debugging purposes.

I am also looking at Providing a common interface to MOP and LCG jobs within MCPS, to delay environment specific details as much and generalize job creation to fit any environment.

5.1.4 D0

Andrew Baranovski: During Oct-Dec 2004 period my role was to provide the infrastructure and tools needed to accomplish distributed merging of the reconstructed data in D0 experiment. The essential part of the task was to implement ordered streaming of the input data in a formalized way using SAM Metadata catalog and SAM data handling according to the experiment specific requirements. Control the actual merging flow via MCRunjob and ensure consistency and uniqueness of the produced data across all execution sites.

5.1.5 STAR

5.1.5.1 Infrastructure

Our RCF Grid infrastructure was under upgrade to Scientific Linux. None of our BNL/RCF grid gatekeepers have been fully functional for most of this quarter and our Grid R&D had to be moved to NERSC/PDSF where the maintenance and upgrade seem to proceed in a more suitable way (in phase) for our schedule. Delays have however impacted some area of activities and namely

- Grid job production testing has been brought to a stall.
- GridCollector progress was planned (in the previous quarter) to be tested on a larger scale on a dedicated resource. This has not happen. Ongoing discussion and action plan may include shipping the node to a more time flexible site.
- File transfer and network rate testing has not been completed: one gatekeeper was purchased installed with 2 NIC card in an attempt to test a local network issue which affected our Grid job submission (ports would close after a while and the call-back would fail) and reported in a previous quarterly report. This node was also meant to re-test a direct path to PDSF as per the file transfer rate. It is now operational and hopefully will remain as such so this part of our program may proceed.

On another front, STAR loaned a node to to serve as a generic Grid gatekeeper. stargridgk01.rhic.bnl.gov (aka rgridgk01) is accessible by all RHIC experiment members who would want to be involved in Grid activities but would not have dedicated resources for it. We hoped to make individual efforts and Grid endeavor at BNL easier.

5.1.5.2 Data Management

Work continued on the HRM production usage and maintenance: Upgrades made to the STAR-specific part of the HRM data mover that began in the previous quarter have been completed and are working well in production mode now. Transfers run continuously with very few interruptions. Typical performance is about 500 GB (2000 Files) per day. The prototype RRS we are using continues to perform without any problems on the NERSC/PDSF side. The service is being deployed at the RCF for completeness and fully automated file transfer and registration both ways.

5.1.5.3 OSG, SRM, VDT

Eric Hjort is following closely the tg-storage OSG activity that is, participated to the DRM Integration/Readiness meetings and email list (organized by Jorge Rodriguez). This OSG activity is concentrating on testing and deployment of DRM (through VDT) on Grid3 sites with a later OSG-0 objective. Our OSG-0 plans currently relies on the presence and smooth integration of SRM tools into VDT.

Part of his contribution has been to provide feedback and testing of the SRM component in VDT as it is provided. Documentation are currently being written and will be available from our Grid Web site[13] soon. However, feedback was immediately sent to both the SDM (Alex Sim) team and at the TG Storage. Some ongoing and not yet documented work is to test the java components of the VDT-DRM release (i.e. the Web Services). More progress on this front is expected along with matrix of applicability of the SRM use.

5.1.5.4 Grid job production and SUMS

As mentioned previously, Grid job production testing has been brought to almost a stall on our resources due to longer than expected upgrade cycles. However, several action items were needed as per the development of the SUMS (STAR Unified Meta-Scheduler) project and we met the schedule on this front. Amongst the most recent development, a Sun Grid Engine (SGE) dispatcher was developed and put into operation in support for a NERSC/PDSF move to SGE. This implementation was rendered easy as we had already transition to a new queue object (see the July-Sep 04 quarterly report) and the notion of virtual clusters allowing for fine grain sub-division of even local resources. In essence, a node (or set of nodes) is determined by the SUMS dispatcher based on requested data set and location of the Physical files on distributed disk; a node being part of a cluster and a cluster being a property of a dispatcher, the proper dispatcher is painlessly determined. Even if the physical files spans over two overlapping clusters (and dispatchers), the last reshape also allowing for priority ordering the virtual queues and dispatchers in case of multiple choices.

On the Grid job submission, we fixed the mapping between the U-JDL file copy element and its translation to the proper `globus-url-copy` command(s). The previous implementation introduced a bug due to syntax miss-understanding and we got around it by disabling this feature and embedding the proper commands in our job submission script. We also took the opportunity to implement a wild-card support for our U-JDL output-copy (local or remote copy, Grid copy). Later implementation will be using SRM and RS for file copy.

SUMS release 1.7.5 with the above features was released soon to be followed by a 1.7.7 version. In the latest, we tuned the scheduler parameters for better performance of SUMS at PDSF based on feedback from Iwona Sakrejda and Eric Hjort. We also recently found a problem with Java threads and its interaction with LSF dispatching mechanism and had to re-code SUMS in that regard (in release 1.7.7): LSF dispatching shell command `bsub` also creates threads and initiate a server/client mechanism. It is therefore unsafe to attempt, from SUMS, to kill and retry the dispatching process if no response is received within a time interval. This may in fact lead to double job submission (rare but observed especially on high system load where the response of the system is slower).

5.1.5.4 From U-JDL to RDL.

Collaboration with Tech-X on the SBIR Phase-I proposal have continued. We meet on a weekly basis and lots of progress has been made as per coding the initial RDL 0.5 proposed schema. The development

[13] http://www.star.bnl.gov/STAR/comp/Grid/

project remains as usual available from the same SUMS project CVS repository (public access) and we carefully keep the namespace separated to avoid clashes. The current implementation allows as planned for support of both the (old) U-JDL and (new) proposed RDL. We expect an initial command line driven of version to be completed within weeks as well as a prototype Web Service based version.

An initial document was presented to the `ppdg-idat` (CS-11) and we are waiting for feedback from this working group. All documentation and progress are available from our RDL Web Link[14].

5.1.5.5 Other activities

A full implementation of a Grid-oriented logger utility is ready for daily testing. We are planning to finalize this work (see CHEP04 480 s23) within weeks now and integrate it to our grid based production testing.

Contacts were made with the Babar Xrootd team (Andrew Hanushevsky and Peter Elmer). We invited Andrew to give a presentation of the Xrootd system at the BNL technology meeting after talking informally at the CHEP04 conference about our goals. Our plan is to proceed with testing and use of the system in STAR as-is (including scalability testing) and later, provide a consolidation of the product by introducing an SRM component.

SRM enabling this product would immediately Grid-enable the Xrootd system and bring the better of the two worlds in what we believe to be an innovative and scalable solution to distributed data management. We made arrangement to have a dedicated STAR person to work on this project, the first phase of which will begin on the second week of February. We also contacted and informed the BNL/RCF personnel of our plan as concerned were raised as per potential security issues related to the introduction of an unknown component in out environment. The Babar team was particularly kind to offer direct help and assistance shall the security FAQ for Xrootd not answer the RCF's questions. Finally, we contacted the SDM team as per our plan to introduce SRM component in Xrootd. Further discussion and meeting with SLAC/SDM/STAR is envisioned to put this later activity well on its track.

5.1.6 PHENIX

5.1.7 ALICE

Larry Pinsky: Work continued on porting AliEn and the version to be used within gLite to the g5 platform. The issue of how to recompile executables in the absence of a Condor client is under consideration. Monte Carlo FLUKA-based codes have been recompiled for the g5 and preliminary benchmarking has been initiated. Substantial efforts have been spent on debating and planning the ALICE-USA role within the overall ALICE Computing Plan. Recent escalations in the estimates of the nominal requirements for an ALICE Tier-1 center have caused ALICE-USA to reconsider whether or not we will be able to provide such a capability in the US. Rather, our efforts have been focused on the potential roles that ALICE-USA-supplied resources might play within the overall ALICE Grid-based distributed computing environment, and the impact that such a decision might have on the middleware needed, as well as the impact on the mix of resources that might be supplied.

5.2 Facilities

5.2.1 JLab

Michael Haddox-Schatz: In the October-December 2004 timeframe I investigated a new version of the SRM (v.3). This effort consisted of two main focuses, design and security evaluation. The design goals were to come up with a simple core portion of an SRM, with well defined ways for adding on extensions. One of the design decisions was to make the SRM v3 web services based and to use X509 certificates for authentication and authorization. Globus Toolkit 3 and JClarens were both investigated with respect to their ability to provide certificate based security with web services.

[14] http://www.star.bnl.gov/STAR/comp/Grid/scheduler/rdl/

5.2.2 BNL RCF/ACF

Gabriele Carcassi has concluded the consolidation of the ATLAS VO with the USATLAS VO. ATLAS users have now a single place to register. This has been done in coordination with the CERN team and the ATLAS VO manager, Alessandro De Salvo. He has continued the development of GUMS, and, as part of the privilege project testbed, setup a SAML authorization service door, which was contacted to retrieve user mapping by a gatekeeper at FNAL. He presented the work at CHEP04 and HEPIX. He has started discussions for possible inclusion of GUMS in VDT. As part of GUMS development had been conducted on GT3, he gave feedback pointing out a number of details where the Globus Container could have improved. He has been participating in the PPDG Common Project, and he helped organizing the PPDG Common Project meeting at BNL, in which a number of issues related to the OSG Spring release were discussed. He collaborates with BNL ITD cybersecurity, to understand the impact of web services, and the possible use of SAZ in the facility.

5.2.3 FNAL

5.2.4 NERSC/PDSF

PDSF continued operating as a Grid3 production site running about 37K CPU hours of jobs for Grid3 during the quarter. Iwona Sakrejda of NERSC User Services attended the OSG Operations workshop at Indianapolis Dec. 1-3.

5.2.5 SLAC

Matteo Melani: 66% of my time was dedicated to working on the development of an accounting system for the OSG and to study the portability of the Babar Monte Carlo simulation production system to the OSG infrastructure. The activities that I performed were:

1) Reading and studies the current literature on accounting in Grid system, distributed system and networks;

2) Writing a draft of a requirements document.

3) Studying the current Babar MC production system, in particular how jobs are assigned and monitored and how data are moved across different administrative domains.

%25 of my time is sponsored by PPDG, the rest came from Babar.

5.2.6 Collaboration with IEPM, Network Performance Monitoring

5.2.6.1 Bandwidth/Throughput measurement (IEPM-BW)
We submitted and had accepted an extended abstract on Transport Fairness Characterization and Evaluation - comparison of variants of TCP and UDP based transport with respect to transient and steady state traffic on real networks for the Protocols for Fast Long-Distance networks in Lyon.

5.2.6.2 Lightweight Bandwidth Estimation
We cleaned up the ABwE web site that describes the lightweight available bandwidth estimation toolkit, and put into production an ABwE server at SLAC. The server is now installed at about 40 sites.

5.2.6.3 Bandwidth performance anomalous events
We completed implementation, tuning, and checking out of the modified NLANR "plateau" algorithm. We built a library of interesting events as well as a canonical data set of time series bandwidth measurements from SLAC to 40 sites for 100 days. This will be used for comparing the effectiveness of the Plateau versus other algorithms such as Kolmogorov-Smirnov, Holt-Winters and Principal Component Analysis. We also set up a collaboration between Loughborough University (K-S), FNAL (H-W), NIIT/Pakistan (PCA) and SLAC (Plateau and leadership) for implementing and comparing the effectiveness/applicability of the various algorithms against a common data set. We put together a draft paper describing this work.

5.2.6.4 Traceroute Analysis and Visualization

5.2.6.5 PingER

We worked with Florida International University and CalTech to install a PingER monitoring site at FIU. This will be particularly used for monitoring sites on the Americas Path Network (AMPATH).
Working with an Institute in Bangalore, India we set up a PingER monitoring site there. This will be valuable to see the performance of paths from and within the developing world.
Hosted visit and talk by geographer from UNE Armitage Australia, who has been using the PingER data to detect a wave of congestion that moves around the world with daytime. This was used very succesfully as a demo for SC2004.

5.2.6.6 SC2004 Bandwidth Challenge

We co-led (with CalTech) a collaboration with Caltech, FNAL, University of Manchester, England, several companies (e.g. Chelsio, S2io, Sun), ESnet, National Lambda Rail and others to participate in this year's SC2004 Bandwidth Challenge. We secured the loan of over $400K of routers/switches/servers/10GE NICs plus two 10Gbits/s links from Sunnyvale to Pittsburgh, and space at the QWest and Level(3) colocation facilities in Sunnyvale. We successfully defended our Bandwidth Challenge for sustained throughput achieving over 100Gbits. This was reported in the national news including Yahoo, CCNews, SlashDot as well as Internationally. For more details see the web site at http://www-iepm.slac.stanford.edu/monitoring/bulk/sc2004/hiperf.html.

5.2.6.7 Proposals and Representation etc.

SLAC is a co-PI on the US Department of State and the Pakistan Ministry of Science and Technology funded project for *Measurement and Analysis for the Global Grid and Internet End-to-end performance (MAGGIE)*. We have regular fortnightly phone meetings We will exchanging visits in the near future to forward the collaboration. SLAC is a partner in the *TeraPaths: A QoS Enabled Collaborative Data Sharing Infrastructure for Peta-scale Computing Research*. A two page summary of the new proposal can eb found at http://www.slac.stanford.edu/grp/scs/net/proposals/iepm-bw/dgnmi-2p.doc. SLAC is a partner in the UltraLight optical testbed proposal (led by Caltech) which was funded by the NSF. We attended the kick-off UltraLight meeting at CalTech in December 2004. The proposal with Texas A&M, NASA and others to NASA on developing and monitoring IP based protocols for NASA satellites etc. was rejected. Submitted paper with FNAL to CHEP04 on *Wide Area Networking System for HEP Experiments at FNAL*.

We served on the Program Committees of the Protocols for Fast Long Distance Networks (PFLDnet 2005) workshop to be held in Lyon in February 2005, and the Passive and Active Measurements workshop (PAM2005) in Boston in 2005. We reviewed about 20 papers and made recommendations.

5.3 Computer Science & Middleware

5.3.1 Condor

Jeff Weber: The Condor Project had several major releases in Q4, 2004: v6.6.7 incorporated several significant bug fixes into the latest v6.6 stable series release. v6.7.2 added several new features to the developer series, including use of multiple collectors for high-availability pools; and fine-grained specification of encrypted file transfer, for improved encryption performance. v6.7.3 added the new "Condor to Condor" Condor-C capability which allows specification of a hierarchy of job schedulers, nested VO groups, and full ports of Condor to Redhat Fedora Core 1, 2 and 3. The nested VO capability allows VOs to define intra-VO group resources, with corresponding policies and accounting. Work continues on adding Globus-3.9.4 support for the next developer series release.
(http://www.cs.wisc.edu/condor/ , http://www.cs.wisc.edu/condor/manual/v6.7/5_3Condor_C.html)

The Virtual Data Toolkit group released two versions of the VDT stable series: v1.2.2, and v1.2.3 which also offers support for Redhat Fedora Core 3, and the DOE-EDG-Certificates certificates package. Work also progressed on adding VOMS and Clarens to the VDT development series.
(http://www.cs.wisc.edu/vdt/)

The Stork data placement service is evolving from a successful research topic, to a mature solution. v0.9.1 was released with support for Redhat 9 and Enterprise Linux 3. Work has begun on integrating Stork directly into future Condor releases. (http://www.cs.wisc.edu/condor/stork/)

CDF has been completing a migration of all analysis farms at FNAL to run under Condor. This has in turn driven the development of the nested VO capability in Condor v6.7.3, and also required significant technical support over the Quarter.

The Condor project participated in the OSG Operations Workshop at IU Dec. 1-3, PPDG Common Projects Dec. 9-10 meeting at BNL, and the OSG technical meeting Dec 15-17 at UCSD. Research continues on several topics of benefit to grid computing. The Condor job queue database mirror provides an independent RDBMS mirror of a Condor job queue. The NeST storage solution continues to move from an advanced research project to a mature software product Development of web services interfaces via SOAP for Condor is nearly complete. (http://www.cs.wisc.edu/condor/nest/, http://www.cs.wisc.edu/condor/birdbath/)

5.3.2 Globus – ANL

5.3.2.1 Coordination and Support

Primary coordination vehicle is via the Open Science Grid blueprint activity. Support through the discuss lists and Bugzilla are available to all experiments.

5.3.2.2 Globus Toolkit 2.x Updates, Bug Fixes and Open Issues

Bug 1725, severity trivial, was closed as WONTFIX. The user was trying to used named instances, something that is not supported under GT3, but its (via a GUID) in GT4.

Bug 2022 was submitted as severity BLOCKER, but after several exchanges we have not received any response from our last query on October 25th, 2004. Based on that fact, I closed this bug out.

Bug 2026 was entered as critical, but there has been no activity, from either side, since it was submitted. I recently sent a follow-up via Bugzilla to determine if this bug was still an issue. It was very similar in nature to 2022, and I suspect that the information provided there may have resolved this issue as well.

Additional information about Bugzilla bugs can be found at http://bugzilla.globus.org.

5.3.2.3 Globus Toolkit 3.2

GT 3.2 continues to be the current stable release for the Globus Toolkit. It contains both the OGSI based web service components, and the pre-WS components (the components similar to the 2.x releases). The WS components offer some substantial performance improvements over the 3.0.x versions. The pre-WS contain mostly bug fixes. However, GridFTP did include some additional, very useful functionality. It added support for structured directory listings in the protocol, which enabled a multi-file globus-url-copy as well as the ability to specify a directory and have it recursively moved in both globus-url-copy and the Reliable File Transfer service.

5.3.2.4 GridFTP / XIO

We continue to work on the new GridFTP server. A new development release (GT3.9.4) was released on December 17th. This release is extremely stable. Our testing to date has been extremely promising. Our bandwidth numbers have been excellent. We achieved 27 Gbs over a TeraGrid 30 Gbs link, using 32 nodes memory-to-memory. Disk results were 17.5 Gbs, and were limited by the disk sub system. Long term stability testing is also promising. The 3.9.4 release has a significant memory leak which has been resolved in the trunk of the CVS tree. Current tests indicate that for a continuous running transfer, we are leaking about 30 KB of memory per 24 hrs. Interoperability tests with previous versions of the server have been successful. Given that earlier servers were never subjected to this kind of testing, it is quite likely that our new implementation is at least as stable, and quite probably more stable than our previous implementation that was based on the wuftpd server. GT3.9.4 also has a known issue with logging to syslog that has been fixed, but not merged into the trunk yet.

XIO work is largely compete at this point. We are adding support for Extended Store and Extended Retrieve (ESTO/ERET) to the GridFTP driver. We have done very little explicit XIO testing. However, since XIO is used as the underlying IO system for GridFTP, the C Web Services Core, and the C GRAM client, it is heavily exercised.

5.3.2.5 Monitoring and MDS work

Initial performance testing was conducted on the Index Service. We ran queries against the DefaultIndexService (7.5 KB message) and achieved 16ms response time average per query. The ContainerRegistryService (32KB message) achieved 29 ms response time per query. The trigger service was completed for 3.9.4, but no performance data is available yet. The archiver has been dropped from GT4.0 due to insufficient resources.

5.3.3 SRM

Alex Sim, Junmin Gu, Viji Natarajan, Arie Shoshani

5.3.3.1 Enhancement activities

The following features were added to our implementation of the DRM and HRM v1.1.

- Support for GSI enabled HSI

The HRM was enhanced to support GSI enabled HSI connection to BNL HSI access is used by the HRM for getting information on the file placement on tapes. This information allows the HRM to access files in a tape-optimized order.

- Support for a default storage system by HRM

An HRM can manage both the space on its own disk, or the space on the MSS, specifically, HPSS. The site URL used for accessing HPSS requires some information needed to be provided to the HPSS system. That made the URL long and confusing to the users at BNL. At BNL's request, we added a parameter to the configuration file, so that this additional information needed by HPSS can be set up initially, thus avoiding the extra long URL.

- Enhancement to support easier debugging

The SRM's client programs were enhanced to provide better formatted log information that make debugging easier.

5.3.3.2 Support activities

Support activities included production level deployment at BNL and working with various people on the initial deployment, helping with problems of usage, and general support at multiple site. Specifically, the following activities took place.

- Support for BNL production level deployment and enhancements on HRM (with Junmin), compile and deliver new code, local installation and maintenance efforts.

- Working with Zhenping (Jane) Liu at BNL for BNL-HRM to understand their requirement. This led to several enhancements to the DRM and the HRM

- Providing technical support to individual users for BNL (David Adams, ATLAS) (Wensheng Deng, BNL, for accessing data from NorduGrid), ANL (Ed May, Atlas), Wayne State U (Elisabeth Atems, STAR)

5.3.3.3 Deployment activities

- Deployment of DRM in VDT

An improved version of DRM that includes all the latest enhancements - DRM v.1.2.2 – was deployed in VDT. This work was performed in cooperation with Alain Roy at U. of Wisconsin. New versions were compiled for the platforms: RH 7.3, RH 8, and RH WS 3 (enterprise Workstation version 3). We produced a "meta-package" for easy installation.

- OSG DRM deployment efforts

-- Alex Sim became DRM deployment Activity Group Co-Chair

-- Collaborated with Ed May for Grid3dev deployment – provided testbed on LBNL machine, testing various client tools, deployed DRM at ANL on a cluster.

-- Collaborated with Paul Sheldon (Vanderbilt) and Jorge Rodrigez (U of Florida) – give support, config file, instructions, …

-- Worked with OSG integration Group (Rob Gardner) on monitoring and accounting planning for development in DRM

-- Worked with OSG deployment group – OSG-0 spring release to include DRM.

5.3.3.4 Development activities

- DRM-NeST project

In this project, we use NeST to do the disk cache management. Typically, a client submits a request to DRM, identifying file source and file size, and DRM asks NeST to reserve a lot for the file, with the size specified, and transfer the file. One thing that DRM relies on NeST to do is when the file size submitted is smaller than the actual file size, the transfer should fail.

In order to have an environment that is easier to debug for the DRM-NeST, a DRM was deployed on node at U of Wisc. So far, we have tested 1000 files that transferred correctly. That required a fix for the NeST log file. Now, we are testing the transfer of a large number of files whose provided size is incorrect (i.e. the file is larger than specified and the transfer need to fail).

- Developed utilities requested by PPDG common project

We developed configuration and testing tool . This includes a tool for generating the configuration file for DRM. Available as a command line and a GUI. This tools guides the user in setting about 20 possible parameters for DRM to use, such as the cache size available to DRM, it's location, lifetime defaults, quota per user default, etc.

5.3.3.5 Other activities

- Kurt Stockinger, Alex, and Arie developed a document for the functional design specification of Replica Management Service, having three services that it uses: the Replica Selection Service (RSS), the Replica Copy Service (RCS) and the Replica Registration Service (RRS). This was intended for the specification of the general framework to perform all the functions expected fo4r Replica Management

- The same people also developed a more detailed functional design specification of Replica Registration Service (RRS). This is expected to be implemented first by LBNL (for STAR), and by CERN (for LCG). The goal is to have an implementation independent specification that will allow many implementations to interoperate, including RLS implementations, file catalogs, and SRB's MCAT.

- Demo at SC 2004 with STAR experiment achieves new milestone

In this demo that was set up with Eric Hjort, the DataMover was successful in moving 17,870 files robustly in a single directory-to-directory request from BNL to LBNL and automatically registering all the file in the file catalog at LBNL..

- Participated in two OSG storage group face-to-face meetings: Arie attended the Blueprint meeting in San Diego, and Alex attended the OSG storage meeting in Fermilab

5.3.4 Caltech

Conrad Steenberg: During this period I was involved in the Open Science Grid collaboration in two ways:

1. As a co-chair of the Monitoring and Information Systems (MIS) working group charged with putting forward standard monitoring software infrastructure for OSG.

2. As a technology provider for the Discovery Service used to act as a central web service discovery mechanism based on Clarens and the MonALISA monitoring system. In this regard a deployment of the discovery service was done on a web server at "http://discover.gridservice.info/" http://discover.gridservice.info/. This service is provided as a service to OSG for development and production purposes.

This service has also been further developed using valuable feedback from the PPDG interactive analysis (CS-11) group.

3. I attended the OSG Blueprint meeting at Fermilab from October 19 to 21.

4. A collaboration with the LambdaStation project was started to develop a Clarens web service to enable for the interactive monitoring and reconfiguration optical network routing to optimize data flows.

5.3.4.1 Demonstrations

A major effort during this period involved preparing three demonstrations for the Supercomputing 2004 conference held in Pittsburgh, PA:

1. The CMS Monte Carlo Processing Service (MCPS) acts as a web services and normal web (browser) front-end to the CMS MCRunJob package that acts as a workflow manager for Monte Carlo simulation data production.

The developers of this service were assisted in constructing the web and web services interfaces, including the addition of better proxy certificate management capabilities in the core Clarens server.

The result of this development was demonstrated as part of the Fermilab/SLAC booth at the conference.

2. An earlier demonstration called Proof Enabled Analysis Cluster (PEAC) produced by the CDF experiment was updated in collaboration with the authors to use Clarens web services as a glue layer to facilitate interaction between users and the computational machinery. This machinery include a Global Manager (GM) that allocates jobs to different clusters, a Local Manager (LM) that takes these jobs and uses PROOF to process jobs, and a SAMGrid service for storage. More information is available at

http://supercomputing.fnal.gov/SC2004/WorkingGroupFiles/content/peac.pdf
http://supercomputing.fnal.gov/SC2004/WorkingGroupFiles/content/peac.pdf

3. A Grid-enabled Analysis Environment (GAE) demonstration was developed in collaboration with the University of Florida (UFL) and the Rio State University (UERJ). A prototype GAE was put in place and used to submit CMS analysis jobs to clusters at Caltech, UFL and UERJ using the BOSS submission system interfaced to Clarens web services. Job queuing and running information was provided to a Clarens server by BOSS for users to query and track the status of jobs.

4. During the Bandwidth Challenge event at the conference, a simple analysis and ROOT-based visualization of CMS Data Challenge '04 (DC04) data distributed across 4 sites in the U.S., Europe and Brazil was shown as an example of a bandwidth-hungry application.

5.3.4.2 Papers

Three different academic papers were co-authored and submitted for publication during this period.

1. HotGrid: Graduated Access to Grid-based Science Gateways Roy Williams, Conrad Steenberg, Julian Bunn, Proceedings of IEEE Supercomputing Conference, Pittsburgh, 2004.

This paper describes a Clarens-based web service and web interface to an astronomical image compositing application developed at the Caltech Center for Advanced Computing Research (CACR) TeraGrid installation This focuses on creating a class of weakly authenticated and authorized users that may use strictly limited amounts of CPU, with the goal of allowing a wider community to be introduced to Grid-based technologies, including certificates for authentication, job scheduling and tracking, as well as remote file handling.

2. The Clarens Web Service Framework for Distributed Scientific Analysis in Grid Projects, Frank van Lingen, Conrad Steenberg et al., submitted to the Workshop on Web and Grid Services for Scientific Data Analysis (WAGSSDA) to be held on Oslo, Norway.

3. Matchmaking, Datasets and Physics Analysis, Heinz Stockinger, Flavia Donno, Giulio Eulisse, Mirco Mazzucato, Conrad Steenberg, also submitted to WAGSSDA. This paper describes a Clarens-based DataLocation web service using the SOAP protocol to query a data replica catalog containing sets of CMS data for analysis. This service is the first to demonstrate interoperability of a Clarens service with the gSOAP toolkit used within the gLite middleware project. Apart from the SOAP interoperability work, my input also resulted in a tenfold increase of performance in this service.

5.3.4.3 Education and Outreach

A need was identified in the CDF experiment at Fermilab for a software tool that allows remote monitoring and interaction with jobs running on cluster worker nodes. This was envisioned as a "shell" that would handle bi-directional interaction with the job.

A student at Fermilab was assigned to implement this "shell" based on the message-based prototype Clarens implementation described in the CHEP paper "The Clarens Grid-Enabled Web Services Framework: Services And Implementation" by Steenberg et al. and the previous status report. In this regard I am acting as a co-adviser to this student.

5.3.5 SRB

Wayne Schroeder

The activities at SDSC in support of high-energy physics data grids were focused on improvements in SRB administration, support for GridFTP access to the SRB, and improvements in data grid federation.

SRB administration:

The SRB MCAT interface and Java Admin Tool were enhanced to improve and simplify SRB administration. Arcot Rajasekar and I implemented the concept of "domain admin" users who are allowed to create other (non-privileged) SRB user accounts in the domain, and "group owners" who can add and remove users to/from groups.

As part of this, I extended the Java Admin Tool in various ways. User update functions were split into two separate classes and windows, one for general operations and one for group-related operations. Code was developed to allow the update of users' email, address, and phone records and to add/remove users as owners of groups. A new subwindow was added for changing a user's type, allowing admins to change existing users to 'domainadmin' types. Additionally, 0 to n sets of resource access constraints can now be displayed.

GridFTP interfaces to the SRB:

Mike Wan and I hosted a visit by John Bresnahan to complete 1) a GridFTP driver (SRB accessing remote sites through GridFTP) and 2) a GridFTP SRB DSI (Data Storage Interface) module for accessing SRB collections using GridFTP. Creating a test/development environment required fixes and/or workarounds to a few GSI problems and a couple of fixes to very small SRB code problems. John Breshahan, with help from SDSC (including development of a new SRB routine by Mike Wan), developed and tested the GridFTP SRB DSI.Basic operations were tested and debugged. I later checked the build procedure, corrected problems discovered due to different user environments, and developed some preliminary internal documentation. The new GridFTP interfaces will be released with the next version of the SRB.

SRB data grid federation:

A simplified procedure for creating a remote zone was developed. I wrote a script, zoneingest.pl, that creates a remote zone (and related items) in the admin's MCAT using as input the output of a 'Stoken Zone' command run at the remote zone. This text could be sent via email from the other SRB administrator.

The zonesync.pl script was updated to ignore inactive zones (especially handy since zones can't be deleted) and also to automatically create and cd into a subdirectory for all the working files.

Reagan Moore: participated in two high-energy physics data grid meetings: EGEE conference, Nov 24-26, 2004; UK e-Science data grid symposium presentation, Dec 1, 2004 SDSC (Mike Wan) also participated

in a meeting at KEK and demonstrated data replication across seven federated data grids for the BELLE experiment. Data was replicated between Japan, Australia, the Far East, Europe, and the US.

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
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		PPDG-OSG-Accounting	Implementation of Accounting V1.0
Dec-04		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		PPDG-DMGT	Use of Enhanced GridFTP

Target Date	Completed	Responsible Task	Milestone Deliverable
Dec-04		PPDG-OSG-Monitoring	Instrumentation of selected data transport and replication applications
Jan-05		PPDG-OSG-Authorization	Implementation V1.0
Jan-05		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.

6.2 Meetings

October 19-21 Open Science Grid Blueprint Meeting, Fermilab

December 9-10th PPDG Common Project Meeting, BNL

December 15-17, 2004 Open Science Grid Technical Meeting, UCSD

6.3 Papers

There were 52 presentations at the ***Computing in High Energy and Nuclear Physics 2004*** (CHEP'04) conference by teams participating in PPDG, and 49 papers.

"A Scalable Grid User Management System for Large Virtual Organizations", G. Carcassi, et al., BNL. (abstract, slides, paper)

Storage Resource Manager, T. Perelmutov, et al., FNAL. (abstract, slides, paper)

Testing the CDF Distributed Computing Framework, V. Bartsch, Oxford, et al. (abstract, poster, paper)

ATLAS Distributed Analysis, D. Adams, BNL. (abstract, slides, paper)

Networks and Grids for High Energy and Nuclear Physics, H. Newman, Caltech. (abstract, slides, paper)

Grid3: An Application Grid Laboratory for Science, R. Gardner, U. Chicago, et al. (abstract, slides, video)

Breaking the 1 GByte/sec Barrier? High speed WAN data transfers for science, J. Bunn, et al., Caltech. (abstract, slides, paper)

Grid Enabled Analysis for CMS: prototype, status and results, F. Van Lingen, Caltech, et al. (abstract, slides, paper)

The Clarens Grid-enabled Web Services Framework: Services and Implementation, C. Steenberg, et al., Caltech. (abstract, slides, paper)

Secure Grid Data Management Technologies in ATLAS, M. Branco, CERN, et al. (abstract, slides, paper)

Managed Data Storage and Data Access Services for Data Grids, J. Bakken, FNAL, et al. (abstract, slides, paper)

Virtual Organization Membership Service eXtension (VOX), Y. Wu, FNAL, et al. (abstract, slides, paper)

The Open Science Grid (OSG), R. Pordes, FNAL, et al. (abstract, slides, paper)

Lattice QCD Data and Metadata Archives at Fermilab and the International Lattice Data Grid, E. Neilsen, et al., FNAL. (abstract, slides, paper)

FroNtier: High Performance Database Access Using Standard Web Components in a Scalable Multi-tier Architecture, L. Lueking, FNAL. (abstract, slides, paper)

Middleware for the next generation Grid infrastructure, E. Laure, CERN, et al. (abstract, slides, paper)

Job Interactivity using Steering Service in Grid Enabled Analysis Environment, A. Anjum, NIIT, et al. (abstract, poster, paper)

Job Monitoring in Interactive Grid Analysis Environment, A. Anjum, NIIT, et al. (abstract, poster, paper)

Predicting Resource Requirements of a Job Submission, A. Anjum, NIIT, et al. (abstract, poster, paper)

JIM Deployment for the CDF Experiment, M. Burgon-Lyon, Glasgow. (abstract, poster, paper)

Use of Condor and GLOW for CMS Simulation Production, S. Dasu, U. Wisc, et al., submitted to CHEP'04. (abstract, slides, paper)

The STAR Unified Meta-Scheduler project, a front end around evolving technologies for user analysis and data production., J. Lauret, BNL, et al. (abstract, slides)

The Grid Collector: Using an Event Catalog to Speedup User Analysis in Distributed Environment, K. Wu, LBNL, et al. (abstract, slides, paper)

A Condor-based, Grid-aware batch software for a large scale Linux Farm, T. Wlodek, et al., BNL. (abstract, poster, paper)

Production mode Data-Replication framework in STAR using the HRM Grid, E. Hjort, LBNL, et al. (abstract, slides)

Tools for GRID deployment of CDF offline and SAM data handling systems for Summer 2004 computing, A. Kreymer, FNAL. (abstract, poster, paper)

Storage Resource Managers at Brookhaven, R. Popescu, et al., BNL. (abstract, slides, paper)

Wide Area Network Monitoring system for HEP experiments at Fermilab, M. Grigoriev, FNAL, et al. (abstract, slides, paper)

LambdaStation: A forwarding and admission control service to interface production network facilities with advanced research network paths, P. Demar, et al., FNAL. (abstract, slides, paper)

Globally Distributed User Analysis Computing at CDF, A. Sill, Texas Tech., et al. (abstract, slides, paper)

Experience producing simulated events for the DZero experiment on the SAM-Grid, G. Garzoglio, FNAL, et al. (abstract, slides, paper)

Cross Experiment Workflow Management: The Runjob Project, P. Love, Lancaster, et al. (abstract, poster, paper)

The Condor based CDF CAF, F. Wuerthwein, UCSD, et al. (abstract, poster, paper)

Development and use of MonALISA high level monitoring services for Meta-Schedulers, E. Efstathiadis, BNL, et al. (abstract, slides, paper)

Interactive Data Analysis on the Grid using Globus 3 and JAS3, T. Johnson, SLAC, et al. (abstract, slides, paper)

Building Global HEP Systems on Kerberos, M. Crawford, FNAL. (abstract, slides, paper)

SAMGrid Monitoring and Information Service and its Integration with MonALisa, A. Lyon, FNAL, et al. (abstract, poster, paper)

Application of the SAMGrid Test Harness for Performance Evaluation and Tuning of a Distributed Cluster Implementation of Data Handling Services, A. Lyon, FNAL, et al. (abstract, slides, paper)

SAMGrid Integration of SRMs, R. Kennedy, FNAL, et al. (abstract, slides, paper)

The SAMGrid Database Server Component: Its Upgraded Infrastructure and Future Development Path, S. Veseli, et al., FNAL. (abstract, slides, paper)

Mis-use Cases for the Grid, D. Skow, FNAL. (abstract, slides, paper)

The status of Fermilab Enstore Data Storage System, A. Moibenko, et al., FNAL. (abstract, slides, paper)

CHOS, a method for concurrently supporting multiple operating system, S. Canon, LBNL, submitted to CHEP'04. (abstract, slides, paper)

Generic logging layer for the distributed computing, V. Fine, BNL, et al. (abstract, poster, paper)

SAMGrid Experiences with the Condor Technology in Run II Computing, I. Terekhov, FNAL. (abstract, poster, paper)

Grid2003 Monitoring, Metrics, and Grid Cataloging System, B. Kim, UFL, et al. (abstract, slides, paper)

Distributed Computing Grid Experiences in CMS DC04, A. Fanfani, INFN, et al. (abstract, slides, paper)

ATLAS Production System in ATLAS Data Challenge 2, L. Goossens, CERN, et al. (abstract, slides, paper)

Resource Predictors in HEP Applications, S. Grinstein, Harvard, et al. (abstract, slides, paper)

PHENIX Job Monitoring/Submission In Transition To The Grid Infrastructure, A. Shevel, SUNY SB, et al. (abstract, slides, paper)

MonALISA: An Agent Based, Dynamic Service System to Monitor, Control and Optimize Grid based Applications, I. LeGrand, et al., Caltech. (abstract, slides, paper)

6.4 Related Publications

HotGrid: Graduated Access to Grid-based Science Gateways Roy Williams, Conrad Steenberg, Julian Bunn, Proceedings of IEEE Supercomputing Conference, Pittsburgh, 2004.

The Clarens Web Service Framework for Distributed Scientific Analysis in Grid Projects, Frank van Lingen, Conrad Steenberg et al., submitted to the Workshop on Web and Grid Services for Scientific Data Analysis (WAGSSDA) to be held on Oslo, Norway.

Matchmaking, Datasets and Physics Analysis, Heinz Stockinger, Flavia Donno, Giulio Eulisse, Mirco Mazzucato, Conrad Steenberg, also submitted to WAGSSDA.