

**The Particle Physics Data Grid Project -
From Fabric to Physics**
Empowering International Virtual Organizations for Frontier Science

Table of Contents

Table of Contents.....	i
List of Participants.....	ii
Budget Summary Table.....	iv
Abstract.....	1
1 Concept of the Proposal – From Fabric to Physics.....	2
2 Background and Accomplishments.....	3
2.1 High-Energy and Nuclear Physics.....	3
2.2 Achievements.....	4
3 From Fabric to Physics - Goals and Requirements.....	5
4 Program of Work: Strategy.....	6
4.1 End-to-End Analysis.....	8
4.2 Software and Middleware Methodology.....	8
5 Program of Work: Details.....	8
5.1 PPDG Common-Project Program of Work.....	8
5.2 Facilities.....	11
5.3 PPDG Experiments.....	13
5.4 Computer Science teams work plan.....	21
5.5 PPDG Milestones.....	24
6 Partnerships and Collaborations.....	27
6.1 SciDAC.....	27
6.2 Physics Collaborations.....	28
7 Budget Summary and Justification.....	29
8 Literature Cited.....	31
9 APPENDIX.....	34

The Particle Physics Data Grid Project - From Fabric to Physics

Empowering International Virtual Organizations for Frontier Science

Scientific Discovery through Advanced Computation Solicitation DE-FG01-04ER04-02 (Renewal Application from SciDAC Announcements LAB 01-06 and LAB 01-11 and Grant Notices 01-06 and 01-11)

DOE Laboratory Contact:

Richard P. Mount, SLAC MS 97, 2575 Sand Hill Road, Menlo Park, CA 94025, Tel: 650 926 2467, Fax: 650 926 3329

University Contact:

Miron Livny, Computer Sciences Dept., Room 5372, University of Wisconsin-Madison, 1210 West Drayton St., Madison, WI 53706, Tel 608-262-4794,

List of Participants

Computer Science Teams

Computer Science Department, University of Wisconsin

Miron Livny (PI), Peter Couvares, Alan De Smet, Alain Roy, Derek Wright

Mathematics and Computer Science Division, Argonne National Laboratory

Ian Foster, Greg Nawrocki

Information Sciences Institute, University of Southern California

Carl Kesselman, Ann Chervenak

Scientific Data Management Group, Lawrence Berkeley National Laboratory

Arie Shoshani, Junmin Gu, Alex Sim

San Diego Supercomputer Center

Reagan Moore, Wayne Schroeder

Physics Experiment Teams

US-ATLAS

John Huth¹², David Adams², Kaushik De¹⁸, Rob Gardner¹⁷, Stewart Loken⁵, David Malon¹, Ed May¹, Larry Price¹, Jim Shank¹⁶, Alex Undrus², Alexandre Vaniachine¹, Torre Wenaus²

BaBar

Roger Barlow¹³, Makoto Asai⁷, Tofigh⁷ Azemooon⁷, Adil Hasan⁷, Wilko Kroeger⁷, Joseph Perl⁷

US-CMS

Harvey Newman³ (PI), Lothar Bauerdick⁴, Anzar Afaq⁴, Paul Avery¹⁰, James Branson⁹, Julian Bunn³, Ian Fisk⁴, Gregory Graham⁴, Saima Iqbal³, Frank van Lingen³, Dan Nae³, Iosif Legrand³, James Letts³, Vladimir Litvin³, James Patton³, Sylvain Ravot³, Conrad Steenberg³, Michael Thomas³, Yang Xia³

CDF

Rick St. Denis¹⁴, Frank Wuerthwein⁹

D0

Wyatt Meritt⁴, Gabriele Garzoglio⁴, Igor Terekov⁴, Sinisa Veseli⁴

PHENIX

Barbara Jacak¹⁵, Dave Morrison², Scott O'Hare¹⁵, Andrey Shevel¹⁵

STAR

Jerome Lauret², Valeri Fine², Richard Casella², Efratios Efstathiadis², Eric Hjort⁵, Doug Olson⁵, Dave Stampf²

Jefferson Lab

(see facilities)

ALICE-USA

Lawrence Pinsky¹⁸

Facilities

BNL RHIC/ATLAS

Bruce Gibbard, Rich Baker, Gabriele Carcassi, Razvan Popescu, Jason Smith, Thomasz Wlodek, Dantong Yu

Fermilab

Don Petravick

Jefferson Lab

Andy Kowalski, Ying Chen, Bryan Hess, Michael Haddock-Schatz, Sandy Philpott, Chip Watson

NERSC/PDSF

Steve Chan, Shane Canon, Bill Kramer, Iwona Sakrejda

SLAC

Richard Mount (PI), Chuck Boeheim, Robert Cowles, Neal Adams, Booker Bense, Gary Buhmaster, Andy Hanushevsky, Randy Melen, Lois White, Wei Yang

Management – Executive Team

Ruth Pordes⁴ (PI) (Chair Steering Committee), Miron Livny¹¹ (Deputy, Computer Science), Doug Olson⁵ (Deputy, Experiments and Facilities)

Liasons

Paul Avery¹⁰ (GriPhyN and iVDGL Projects), Mike Wilde¹ (Globus/GriPhyN Project), Rob Gardner¹⁷ (iVDGL), Ian Bird^{CERN} (LCG), Torre Wenaus² (LCG), Chip Watson⁸ (Lattice-QCG SciDAC), Bill Kramer⁵ (DOE Science Grid)

Institutions

¹Argonne National Laboratory, ²Brookhaven National Laboratory, ³California Institute of Technology, ⁴Fermi National Laboratory, ⁵Lawrence Berkeley National Laboratory, ⁶San Diego Supercomputer Center, ⁷Stanford Linear Accelerator Center, ⁸Thomas Jefferson National Accelerator Facility, ⁹University of California at San Diego, ¹⁰University of Florida, ¹¹University of Wisconsin, ¹²Harvard University, ¹³University of Manchester, ¹⁴University of Glasgow, ¹⁵SUNY Stony Brook, ¹⁶Boston University, ¹⁷University of Chicago, ¹⁸UT Arlington, ¹⁹University of Houston, ²⁰University of Southern California

The Particle Physics Data Grid Project - From Fabric to Physics

Empowering International Virtual Organizations for Frontier Science

Budget Summary Table

The budget requests for the collaborators are as follows:

The University of Florida is not requesting funding but will collaborate to ensure coordination with the NSF-funded GriPhyN and iVDGL projects led by Florida. The coherence of the “Trillium” – PPDG, GriPhyN and iVDGL is a fundamental component of the PPDG strategy.

\$K	July 1, 2004 June 30, 2005	July 1, 2005 June 30, 2006	2 Year Total
Argonne National Laboratory	487	487	974
Berkeley Laboratory	420	420	840
Brookhaven National Laboratory	366	366	732
California Institute of Technology	194	194	388
Fermi National Accelerator Laboratory	833	833	1666
San Diego Supercomputer Center	72	72	144
Stanford Linear Accelerator Center	368	368	735
Thomas Jefferson National Accelerator Facility	176	176	352
University of California, San Diego	82	82	164
University of Florida	0	0	0
University of Wisconsin	540	540	1080
Total DoE Laboratory Request	2648	2648	5296
Total University	888	888	1776
Total Request	3538	3538	7074

The Particle Physics Data Grid Project - From Fabric to Physics

Empowering International Virtual Organizations for Frontier Science

Abstract

The Particle Physics Data Grid Collaboratory Pilot has benefited the scientific discovery potential of the collaborating experiments by delivering and deploying vertically integrated Grid systems for their applications. The Grid technologies of the collaborating Computer Science groups have been hardened and extended in response to the needs of the experiments' systems. The proposed two year extension will be used to build a Grid infrastructure across the major HENP DOE Laboratories and universities that are part of PPDG, which will additionally provide a contribution to the Open Science Grid. We will extend, develop and deploy well defined common interfaces and adaptors, based on common Grid technologies, to the services provided by the center facilities. The program of work will focus on further extending needed capabilities for data and job management, and on providing missing services in particular in the areas of accounting, authorization and end-to-end grid system monitoring and troubleshooting. We will also extend the performance and capabilities of our vertically integrated Grids to support the more dynamic and complex experiment analysis applications, with the experiment groups collaborating on providing necessary higher level services. Our goals and plans continue to be guided by practical immediate and medium-term needs of the now-running physics experiments as well as developments for the LHC experiments. This phase of PPDG will also complete the transition from the GT2 based Grid infrastructure to the next generation web service based technologies. PPDG will continue its active involvement in collaboration with other SciDAC programs - in particular the DOE SG, data middleware and security infrastructure projects, as well as with the LHC Computing Grid and ARDA project, and other data-grid projects in Europe and beyond.

The Particle Physics Data Grid Project - From Fabric to Physics

Empowering International Virtual Organizations for Frontier Science

1 Concept of the Proposal – From Fabric to Physics

The Particle Physics Data Grid (PPDG)[PPDG] is the focus of DOE support for the creation and exploitation of large heterogeneous computing environments in high-energy and nuclear physics (HENP). Since June 2001, the PPDG collaboration has successfully deployed elements of Grid technology in all participating experiments, culminating in strong PPDG involvement in the recent successful deployment and operation of a national Grid demonstrator [GRID2003]. Over the same period the close collaboration between PPDG computer scientists and the mission-driven PPDG experimental physics community motivated and guided research in the areas of distributed systems and networking as well as the development of new technologies and software tools.

PPDG intends to maintain its pragmatic focus and believes that it is now time to engage the principal DOE-HENP computing centers, linked by the ESNet network, in the deployment of a sharable infrastructure for U.S. HENP. Petabytes of mass storage, hundreds of terabytes of disk cache and many thousands of processors, interconnected by high speed links and supported by the best of current PPDG distributed technology will be a powerful resource for all ranks of the national physics community. More importantly, it will provide a powerful stimulus for the collaborative computer-science/experimental-physics developments required to extend usability of Grid technologies into the apparently chaotic world of physics analysis where we expect distributed computing to bring its greatest benefits.

PPDG strongly supports and plans to contribute at all levels to the emerging plans for a consortium of the laboratory-based computer centers, university facilities, application communities and Grid projects serving high energy and nuclear physics to join together in an Open Science Grid (OSG)[OSG03] to deploy a U.S. national persistent Grid infrastructure for science. This strategy and collaboration affords a unique opportunity for PPDG to continue its work in a larger community to provide and support vital Grid-enabled services to our experiments. Moreover, OSG will provide a natural channel for interfacing PPDG resources and the PPDG community with other continental and national Grid efforts like the EGEE.

We have set ourselves an achievable but daunting task. Moving beyond production-driven simulation – currently the principal application running on HENP Grids – means addressing the additional dimensions of data placement and storage management, heterogeneity of executable tasks, and an ill-controlled community of hundreds of mission focused, scientifically aggressive users. Moving into major DOE computing facilities requires a rigorous approach to assuring that DOE mission priorities and policies are met, and that facilities do operate in a sustainable rather than heroic way. Finally, robust and functional distributed physics analysis capabilities, on a scale targeted at the LHC, must be delivered by the PPDG experiments to their physicists.

This proposal is an evolutionary step that builds on accomplishments achieved and lessons learned by the PPDG partnership of experiments and computer science teams over the first three years of the SciDAC program. It is motivated and guided by the same principle as the original proposal – deployment and adoption of end-to-end, Grid-enabled capabilities by the PPDG experiments – and follows the same organizational structure. We plan to continue to exploit the unique set of inter-disciplinary partners and leverage the strong and productive links they have established and cultivated over the past five years. As we move to the next phase in our plan, to translate the potential of distributed computing into effective capabilities, we will focus our efforts on the following challenges:

- Deploy a shared distributed infrastructure capable of supporting a diverse workload of analysis applications.

- Deploy Grid-enabled storage and data management services.
- Migrate PPDG Grid-enabled tools to the next generation of Grid technologies.

The first challenge requires a change in the way our collaboratory operates. Deployment of cross-experiment capabilities mandates centrally managed collaboratory-wide efforts. We refer to such an effort as “common projects”. We plan to allocate 25% of our FTEs to common projects. The PPDG executive team will manage each of these projects.

The second challenge requires a refocusing of our current efforts. The members of PPDG have demonstrated a rich set of distributed storage and data management capabilities. Extensions to these capabilities must be developed and integrated to meet the additional needs of analysis applications.

The third “challenge” stems from our success in integrating common grid technologies into the software stacks of the experiments. As these common technologies evolve over time, PPDG faces the difficult task of upgrading an installed and operational software infrastructure. As the Grid community moves to a Web Services software architecture, most noticeably the transition of the Globus toolkit to a WSRF based suite of web services [WSRF], PPDG will have to carefully and methodically plan and execute the transition to new technologies in order to reap the benefits of emerging frameworks.

2 Background and Accomplishments

2.1 High-Energy and Nuclear Physics

HENP experiments, now taking data or under construction, are poised to break new ground in the understanding of the unification of forces, the origin and stability of matter, and structures and symmetries that govern the nature of matter and space-time in our universe. Among the principal goals at the high-energy frontier are to find the mechanism responsible for mass in the universe, the Higgs particles associated with mass generation, and the fundamental mechanism that led to the predominance of matter over antimatter in the observable cosmos. At the high-density frontier, studies of the properties of dense, strongly-interacting partonic matter are discovering exciting novel features, and will continue to be pushed further with upgraded and new facilities. The structure of the nucleon, and the mechanism responsible for confining the primordial quarks and gluons into the observed hadrons, is being explored, and the nature of the confining flux tube examined in upgraded hadronic physics facilities.

During the next three years, the data volumes of the running experiments at the US HENP laboratories will increase and the LHC experiments will ramp up towards first operations. Physics analysis activities during LHC operation will be characterized by thousands of requests per day for subsets of data and processed results from the Gigabyte to the Terabyte range, drawn from multi-Petabyte distributed data stores. These requests will be posed by individuals as well as by small and large workgroups located across the U.S. and around the world, with varying task requirements for computing, data handling and network resources and varying levels of priority.

While the ensemble of computational, data-handling and network resources foreseen is large by present-day standards, it is going to be limited compared to the potential demands of the physics user community. Each physics collaboration, large as it is, has a well-defined management structure with lines of authority and responsibility. Scheduling of resources and the relative priority among competing tasks becomes a matter of policy rather than moment-to-moment technical capability alone. The performance (efficiency of resource use; turnaround time) in completing the assigned range of tasks, and especially the weekly, monthly and annual partitioning of resource usage among tasks at different levels of priority must be tracked, and matched to the policy. There will also be site-dependent policies on the use of resources at each facility, negotiated in advance between each site-facility and the Collaboration and Laboratory managements. These local and regional policies need to be taken into account in any of the instantaneous decisions taken as to where a task will run, when a set of files will be moved, or where a data set will be stored.

2.1.1 Production and Analysis

HENP computing tasks divide into “production” and “analysis” which have comparable processing needs but require quite different levels of sophistication from Grid systems.

Production tasks are highly organized activities where a small number of executables are run repeatedly by a very small number of well-organized teams to perform data processing or simulation on behalf of the whole physics-experiment collaboration. The simplest production task is simulation, which is invariably CPU-dominated. Production tasks normally run on resources that are, at least temporarily, dedicated to the task. In a Grid environment, it is easy for the production teams and the Grid-operation team to cooperate at a personal level to ensure success. Planning, managing and accounting of resource use can, at the limit, involve only telephone, pencil and paper. Troubleshooting is limited to a well defined software stack and turnaround is measured in weeks and months. The early success of GRID2003 is closely related to the fact that simulation production has been its dominant application.

Analysis is optimized when physicists at all ranks and locations in a collaboration have the opportunity to try out their ideas – even the apparently crazy ones – subject to ensuring that a collaboration’s physics policies are reflected in the overall consumption of available resources. Each of these physicists may use a different home-grown set of executables and may manipulate a different data set. An analysis task may be CPU-intensive, I/O-intensive, memory intensive or all of the above. It may require huge datasets of whose location the physicist has little or no knowledge or just a small locally stored dataset. In one evening a single graduate student can submit work that would keep the entire Grid occupied for a decade. As the Grid begins to facilitate and empower distributed analysis, it will become more and more important to have automated approaches to monitoring and controlling resource consumption so that the priorities of the physics collaborations and of the resource providers are reflected in the allocation of resources.

2.2 Achievements

The progress of PPDG is documented in quarterly reports and over 40 other documents and papers [PPDGDOC].

During the past three years the PPDG has continued to make significant progress in the deployment and adoption of Grid-enabled end-to-end capabilities in the production data processing infrastructure of the participating experiments. We have seen an across-the-board increase in the integration of common Grid technologies into the software stacks of the experiments. The success of PPDG in actually enhancing the daily data-processing activities of the experiments stems from our focus on end-to-end functionality and the proactive management and execution style of the collaboratory. During the current (third SciDAC) year of the PPDG Collaboratory Pilot we have shifted our focus and commitment to cross-experiments and cross-projects efforts and started the transition to the new generation of a service based Grid middleware.

The last two years have seen a significant change towards the acceptance of Grids as a viable technology to address the experiments’ distributed computing needs. Success of early focused integration efforts resulted in the hardening of common Grid middleware and led the way to broader joint physics-computer sciences efforts. Today, all PPDG experiments use Grid technologies in their simulation/production infrastructure. The developments and deployments of distributed computing and Grid technologies involve many more people than those on the “PPDG team” *per se*.

Specific achievements by PPDG over the past two years are:

- Robust, sustained, hands-off, production data transfer of terabytes of data using GridFTP and SRM implementations. This activity crosses all experiments, with each using a combination of common components and specific variants of the middleware.
- Grid-based job scheduling and execution based on Condor-G, DAGMan and GRAM, in particular for simulation jobs for U.S. ATLAS, U.S. CMS, D0, and CDF, and production data processing for Run II experiments by SAM. Up to a hundreds thousand jobs – covering test and production – have been executed.

- Deployment and use of end-to-end Grid systems for experiment applications. In particular the recent Grid2003 shared by U.S. ATLAS and U.S. CMS (among others) and SAMGrid for Fermilab RunII. This includes Grid system information gathering using MDS and monitoring using the agent-based MonaLisa, in particular for Grid2003 and STAR.
- Procedures, integration and support for DOE Science Grid certificate-based authentication at all the facilities. Hundreds of user certificates and more than a thousand host certificates have been issued.
- Demonstrations of interoperability and compatibility across the US and European Grid infrastructures of the experiments, in particular BaBar, D0 and CDF, US ATLAS and US CMS.

Close cooperation with the NSF-funded Grid projects targeted at HENP applications has become a way of life in PPDG. Coordination with GriPhyN and iVDGL is maintained at the management level, and in the field, to ensure that the complementarities of the projects that was “designed-in” at the proposal stage is strengthened as the projects progress and evolve. PPDG experiments have also adopted the Virtual Data Toolkit (VDT) [VDT00] packaging and distribution of the common Grid middleware. Build and testing tools and procedures developed by the National Middleware Initiative (NMI) are used to package the VDT that includes most of the components distributed by this initiative. The most recent and relevant example is GRID2003[GRID2003], where a working Grid of over 2000 processors spread over 27 sites is now operating and delivering simulations urgently needed by the LHC experiments. The success of GRID2003, and the cost of that success in terms of initial and ongoing effort, is a central force behind the program of work present by this proposal.

PPDG has taken a major role in the international coordination of Grid projects relevant to HENP. The excellent coordination of the US Grid projects has been of great value in establishing productive relations, particularly with European partners involved in the CERN LHC program. Moreover, our contributions to the functionality and overall quality of Grid technologies have a profound impact on the national and international Grid communities. It will be imperative that this coordination continues. We will continue to maintain our role in leading the HENP community in broadening the deployment and adoption of common Grid technologies.

3 From Fabric to Physics - Goals and Requirements

The PPDG Collaboration maintains its vision that the multiple experiment-computer science partnerships should evolve into a common approach across many experiments. This partnership brings together demanding applications, advanced middleware and powerful computing, storage and networking resources. The cross fertilization between the different computer science teams and the experiments coupled with the constraints of existing facilities and externally driven timelines hold the key to the effectiveness of the PPDG approach. Progress in many areas of PPDG has reached a point where a push for a broader user base and more commonality is timely. What better way to initiate this than to offer a common US Grid fabric shared by all members of the PPDG collaborative and capable of supporting analysis tasks? Opening some of the dedicated computing and storage fabric to managed shared access and opening Grid-enabled services to a diverse analysis workload requires a coordinated effort that brings together a powerful collection of resources and flexible and robust services.

The computing centers at BNL, JLab, Fermilab, SLAC, and NERSC/LBNL linked by an increasingly proactive ESnet, form an outstanding computing resource for data-intensive science. The dedicated facilities at the high-energy and nuclear physics labs bring multi-tera-op computation linked to multi-petabyte data storage. These facilities are undergoing aggressive expansion to meet programmatic needs. The dominant users of these facilities are PPDG members all of whom are committed to Grid computing as a long-term maximizer of scientific productivity.

The use of a common Grid fabric as a resource shared by the entire PPDG community brings benefits and requirements. It will contribute to the Open Science Grid strategy, which relies on the partnership and contributions of existing computing resources, and an alignment of the current Grid efforts. The principal benefits are:

- Acceleration of progress towards a hardened software stack and a Grid environment that is productive for individual physicists;
- Early exposure of design and implementation shortcomings of common components that will require changes in allocation of resources and time lines;
- Heightened ability, based on demonstrated deployment and aggregated computing and storage resources, to partner with the international Grid community in setting future directions;
- Improved coherence of the US HENP computing environment;
- Greatly improved ability to respond to urgent programmatic needs for massive data-analysis resources.

Exploitation of an open Grid fabric will also bring requirements for new services and capabilities. We must also pay attention to the increased need for global services for monitoring, tracking, strategic planning, policy-matching etc. These requirements will drive PPDG activities in the following areas (a subset of the program of work outlined later in the proposal):

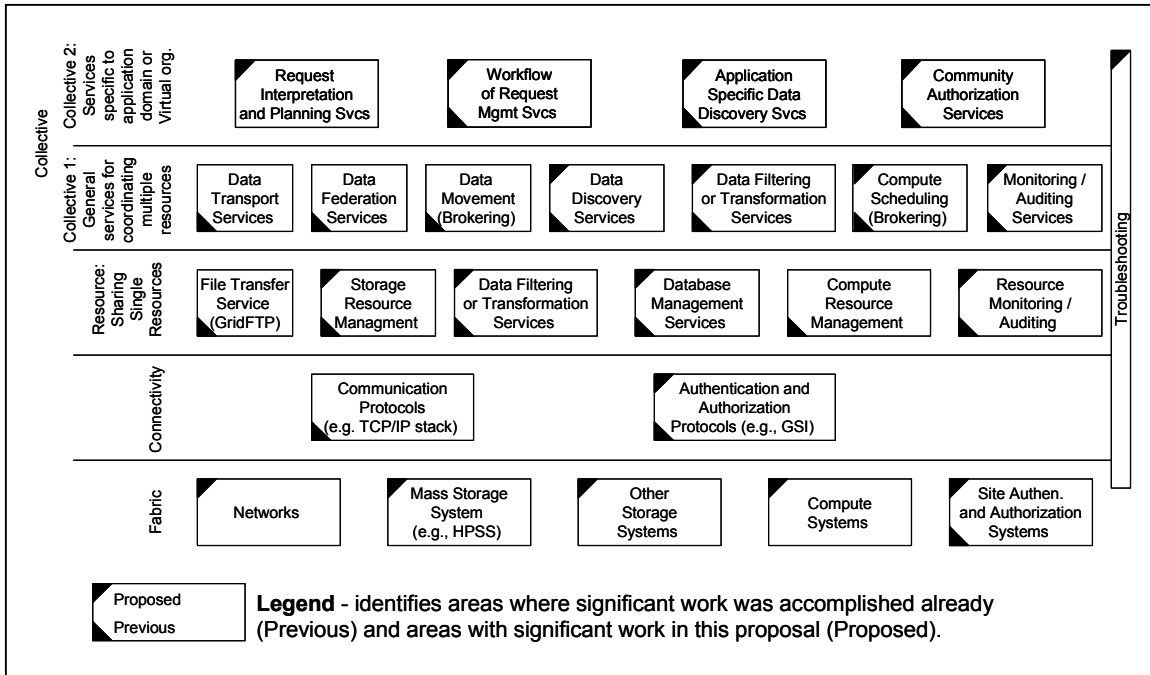
- **Accounting** - Detailed and reliable recording of resource consumption will be a prerequisite as the Grid moves into production. Almost all the fabric elements have been funded and built to meet the documented needs of particular experimental programs and there is no existing process that will fund general facilities for HENP. Committing the experiment-specific resources to the Grid is justified by the benefits this will bring, but it will remain essential that the usage of resources by experiments can be tracked and reported. Data management resources cost 50% or more of the total expenditure on HENP computing, so the accounting of storage usage must accompany the accounting of computing usage.
- **Authorization** - A dynamic approach to authorization, capable of implementing the mission-driven policies of the resource owners, the virtual organizations and research subgroups within organizations and groups will be essential. At its simplest, this approach will be capable of giving preferred services to the science for which a fabric element was funded and constructed.
- **System Instrumentation and Monitoring** – System-wide monitoring and information gathering will become increasingly important as a tool for managing and ensuring a robust distributed system. Instrumenting components in the system with a consistent “event identifier” and persistent “state recorder” model is essential to enable tracking, diagnosis and troubleshooting [TIE03]. The project will continue to use, and influence the hardening and extensions of, MDS[MDS], MonaLisa [MONA], and NetLogger [NETL]. as well as continue the use of other systems. These are part of the global information, monitoring and logging services which must eventually be extended to include more complex capabilities of strategic planning and policy driven matching.
- **Troubleshooting** - Detecting and resolving deficiencies in the deployed software stack are also vital to the effective operation of a distributed system. We refer to this detection and resolution as “troubleshooting” the system. In this very broad area of work it is expected that the focus of the PPDG effort in the next two years will be driven by deployment of basic capabilities and emerging operational experience, moderated by regular reviews of the activities for consistency with long-term architectural integrity.
- **Data Management** - Effective access to large amounts of data continues to be a PPDG focus. Components in support of PPDG have two major aspects: Replica Management Services (RMS) to provide robust and managed multi-file data movement; and Storage Resource Management (SRM) services needed for managing disk space and mass-storage systems space dynamically.

4 Program of Work: Strategy

PPDG must achieve a significant shift of emphasis towards common efforts to ensure that its physics experiments can run production and analysis on the Open Science Grid. At the same time, the ongoing

more experiment-specific activities must necessarily receive reduced PPDG support. This will be achieved by moving some of these activities out of PPDG to become the responsibilities of the experiments, and by realigning other activities with the new PPDG emphasis.

Each PPDG group agrees to provide resources that will support the exploitation of the common fabric under PPDG executive management – the PPDG Common Project. The resources to support ongoing collaborative and joint activities will be the responsibility of the groups concerned.



The diagram above illustrates a view of a data grid infrastructure[ARCH]. One can see a concentration of previous effort in the Connectivity, Resource and Collective 1 layers. The areas emphasized in this proposal for the next two years illustrate the new focus extending into the fabric interfaces as well as to the higher level application services. While there will be necessarily less effort available for the Resource and first Collective layers, an important point of view enabling diagnosis and troubleshooting will be maintained throughout the layers.

PPDG deliverables and milestones will continue to be driven by the computing needs¹ and time lines of the experiments. Deliverables and milestones will clearly evolve as PPDG deployment brings the costs and capabilities of the Grid into clearer view.

PPDG-funded effort has always been complemented by an often larger effort from each physics experiment and has been leveraging software tools developed and maintained by projects supported by other funding sources. As PPDG focuses on enabling experiments to exploit shared Grid resources with common Grid technologies there will be major interactions with the teams managing the fabric of the facilities serving HENP. It is a PPDG management goal that these interactions are sufficiently productive that all parties involved view them as beneficial.

¹ Computing “needs” may be defined as the most productive physics computing environment that can be created with resources that are available and are reasonable in comparison with the total project cost.

4.1 End-to-End Analysis

The success of Grid-based analysis must be assured. In addition to providing the fabric plus the accounting, resource-management and troubleshooting capabilities, attention must be paid to ensuring that the end-to-end requirements of physics analysis are met. PPDG effort, augmented by a much larger effort from the experiments, must ensure that evolving analysis requirements are understood and that missing functionality, or opportunities for revolutionary improvement, receive appropriate attention. Besides the activities given above, analysis requirements can be expected to drive additional common work in the area of job definition languages, workflow management and scheduling, finer granularity privilege and access control authorization, performance prediction and system optimization.

4.2 Software and Middleware Methodology

The end-to-end and system-wide focus of PPDG leads to the adoption of a software-testing and acceptance process integrated with the experiments' practices. The required development and testing of components are performed in a Development or Testbed Grid environment, integration of new services with the existing environment and with the facilities are performed on an Integration or Certification Grid environment, and lastly, the new functionalities are moved into the Production Grid. This requires strategies for updating the production environment, of quality control and certification, and for partitioning the environment in a flexible way.

The common middleware technologies are subject to a process of validation and regression as part of the Virtual Data Toolkit (VDT) release procedures. VDT releases have well-identified version numbers, release notes and support mechanisms. Given the wide acceptance by the HENP community of the VDT, the PPDG collaboratory will align, to the extent possible, middleware deployments and timelines with the testing and release procedures of the VDT. PPDG experiments contribute to additional testing of VDT by integrating pre-release and development versions into their test beds for end-to-end validation and system tests.

PPDG will continue a staged transition, started in the third year of the project, to the next generation of service based Grid technologies. This transition will require changes at the client software side as well as extensions in functionality and hardening of the new middleware in a similar fashion to that required for the currently deployed software. Services adopted or developed by PPDG are likely to follow common Web-Service based interfaces so that they meet the increasing system requirements for scalability, robustness and complexity management. Attention will be paid to service-based components offered through the VDT as they emerge and evolve.

PPDG will keep aware of, make use of, and contribute to, the standards and practices of the wider community through attendance at the Global Grid Forum and participation in the Particle and Nuclear Physics Application Research Group of the GGF, in which members of PPDG play a leadership role.

5 Program of Work: Details

The program of work is described in four groups – Common Project, Facilities, Experiments and Computer Science. The work is organized and executed in teams with participants from more than one group as described in the PPDG management plan.

5.1 PPDG Common-Project Program of Work

Common deliverables by their very nature require a cooperative effort, managed by the PPDG Common Project, but drawing heavily on hardware resources of the PPDG experiments and the computer centers. Consistency with the milestones of the individual experiments and the computer science groups will be maintained as all parties develop a clearer vision of the costs and benefits of the planned common

capabilities. We characterize many of these deliverables as “PPDG-OSG”, indicating their relevance to the requirements for the Open Science Grid.

The deployment of Grid computing on existing well-managed computing facilities will, initially, increase the workload of the non-PPDG groups that develop and operate these facilities. This will be offset by coherent engineering support provided by PPDG. Achieving this tradeoff will be one of the responsibilities of the PPDG Common Project team.

5.1.1 PPDG-OSG overall deliverables and capabilities

Ongoing: At six-month intervals the availability of service based middleware technologies for PPDG common services will be evaluated. Planned transitions to these will be supported by the Common Project and will require additional designated effort from the computer science and experiment groups.

January 2005: Sustained and ongoing multi-experiment exploitation of over 2000 processors and over 100 terabytes of disk space for production simulation and a small total number of more data-intensive production tasks. Limited “early adopter” usage for unplanned analysis tasks submitted by individual physicists.

June 2006: Sustained multi-experiment exploitation of a significant fraction of the DOE-HENP resources in OSG for production processing and planned analysis tasks serving analysis groups. Analysis usage by individual physicists subject to appropriate guidelines aimed at ensuring system stability.

5.1.2 Accounting Services

The effort to develop and deploy accounting services will initially focus on addressing the need for the contributing facility centers to record the shared usage of their resources, especially storage, by application and experiment groups. Initially we will concentrate on the business and information models, adapting the existing accounting systems in use at the computing centers and recording and providing the information in a uniform fashion.

- PPDG-OSG version-1 accounting specification, **October 2004**, including: Definition of the economic, resource cost and value model; definition of requirements for facility and experiment accounting; agreement on schema for accounting reporting; architecture for storage and access to accounting information; investigative research into design principles and technologies.
- PPDG-OSG version-0.1 accounting implementation, **December 2004**. The subset of the version 1 accounting specification required for initial operation: Initial limited instrumentation and automated information gathering; limited system for storage of and access to the information; definition of common command-line interface to accounting information; initial definition of granularity of information; first analysis of information gathered.
- PPDG-OSG version-2 accounting specification, **October 2005**: Review of utility of accounting information; refinement of the cost and value model; adoption of language, accounting information and metrics; upgrade of the service interface; improved functionality of acquisition of the information; ability to interoperate with the European HEP Grids; enhanced analysis and reporting capabilities.
- PPDG-OSG version-1.1 accounting implementation, **December 2006**. The implementation will aim at consistency with the version-2 specification, but with scope closer to the version-1 specification.
- PPDG-OSG version-2 accounting implementation, **June 2006**.

5.1.3 Authorization

To the extent possible, PPDG will extend and harden existing authorization technologies being developed in the computer science and Grid project communities. The requirements of the HENP community, especially in support of analysis, are likely to stimulate new technologies in this area. PPDG expects to continue to evaluate technologies from and collaborate with several outside groups (e.g. NCSA, Virginia Tech, EGEE and Internet2).

- PPDG-OSG version-1 authorization specification, **July 2004**. The goal will be to allow authorization and prioritization for processing and storage facilities at the Virtual Organization (VO) level, and outline the approach to address the issue of renewable certificates, more granular authorization and more flexible prioritization.
- PPDG-OSG version-0.1 authorization implementation, **January 2005**. Implementing simple authorization and prioritization for processing and storage facilities at the VO level. The delivery date should be adhered to even if the automation is incomplete.
- PPDG-OSG version-2 authorization specification, **September 2005**. Authorization and prioritization at the subgroup level within VOs plus interface specifications allowing policy information and policy decisions to be provided by external agents and processes.
- PPDG-OSG version-1.1 authorization implementation, **May 2006**.

5.1.4 System Instrumentation and Monitoring

Extensions to existing system monitoring capabilities will be required to address the scale and performance of the common Grid infrastructure.

- PPDG-OSG monitoring version 1, **Dec 2004**. Monitoring of the system infrastructure across all the Laboratory facilities, and presentation of the information in a uniform and accessible form.
- PPDG-OSG instrumentation version 1, **Dec 2004**. Instrumentation of selected data transport and replication applications. Possibilities are, pyGlobus[PYGLO] interface to GridFTP (already instrumented with Netlogger), SRM and SRB.
- PPDG-OSG monitoring and instrumentation roadmap, **May 2005**, based on experience deploying and using the existing tools.
- PPDG-OSG monitoring version 2, **Dec 2005**. Migrate system monitoring with MonaLisa to enable clients to obtain feedback (through Clarend or other web service infrastructures) about the state of the system.
- PPDG-OSG instrumentation version 2, **Dec 2005**. Instrumentation of selected end-to-end data analysis applications consistent with the data transport instrumentation infrastructure.

5.1.5 Troubleshooting

Our troubleshooting activities will involve a pragmatic mix of manual and automated procedures and services. As new tools and approaches are developed by the distributed computing community, PPDG will focus on developing and deploying tools to collect analyze and display information produced by existing middleware and application components. Experience will drive information and monitoring requirements as well as error reporting and response services.

- PPDG-OSG version-1 specification of baseline troubleshooting capabilities required by users and administrators, and assessment of reporting and logging capabilities of existing software components, **September 2004**.
- Implementation of baseline troubleshooting capabilities complete, **July, 2005**.
- PPDG-OSG specification of enhanced troubleshooting capabilities, **November 2005**.
- Evaluation of “state of the art” of distributed troubleshooting technologies and of highest priority enhanced troubleshooting services by **June 2006**.

5.1.6 Data Management

Replica Management Services (RMS) require a secure file transfer service (GridFTP), robust multi-file data movement services from disks and mass storage systems (DataMover and Reliable File Transfer services), Replica Location Service (RLS) in order to keep information about replicated files, and Replica Registration Service (RRS) in order to register or delete multiple files into an RLS and/or expanded

metadata file catalogs during or after their movement. Development and implementation work on GridFTP and RLS is funded partially through the SciDAC High Performance Data Grid Toolkit middleware project.

- PPDG-OSG File Data Transfer. Deployment and integration of new versions high performance data transfer GridFTP (delivered through the VDT). **December 2004. Increased capabilities December 2005.**
- PPDG-OSG Replica Registration Service version 1 specification, **December 2004.** Replica Registration Service implementation of a production version1, **June 2005.**

Storage Resource Management (SRM) services are needed for managing disk space and mass-storage systems space dynamically. They are necessary to support reporting and allocating space for job planning, job execution, accounting and job monitoring. SRMs are used to support robust large-scale data movement from mass-storage systems, and are essential to support analysis jobs by providing dynamic storage management. The SRM concept development and implementation work is supported partially by the SciDAC supported Storage Resource Management middleware project, and the results deployed in the PPDG-OSG framework.

- Implementation of Storage Resource Management, SRM version 2 that supports storage reservation and directory management, **November 2004.**
- Specification of Storage Resource Management, SRM version 3 that supports authorization, monitoring, and access control, **June 2005.**
- Implementation of SRM version 3 over various disk and mass-storage systems, **May 2006.**

5.1.7 Remote Workload Management and scheduling

In support of the additional requirements for Grid-based analysis applications, common services in support of job management, execution and workflow will be extended as part of the common project to:

- Develop a portable and generic Process Wrapper (ProRap) for remote execution of HENP applications. The ProRap will be based on the Condor “starter” and will provide monitoring, control, and remote I/O services via authenticated and secure communication links. A prototype of the ProRap is currently being tested. (delivered via VDT) (**October 2004 and April 2005**)
- Enhance the scalability of Condor-G (support 10^5 jobs) and to provide High-Availability services. Scalability will be achieved via dynamic clustering of jobs, restructuring of the ClassAd representation of job that takes advantage of new features of the ClassAd language and multi threading of services that only require read access to the job information. (delivered via VDT) (**December 2004**)

5.2 Facilities

PPDG end-to-end systems and applications and the services supporting them will need to be integrated with and deployed in the Grid environment that the physics experiments are using to provide the computing for their science mission.

PPDG will continue to help integrate end-to-end services into the facilities environments at labs and universities that are providing their resources to the Grid, and also with the set of already existing Grid services, with the goal of creating end-to-end capabilities that are coherently and persistently deployed.

We will develop an understanding of the sufficient set of interfaces and principles of operation between the PPDG computing-center facilities and the common Grid environment. Together with the computing-center facility groups we will identify additional integration and system services relevant to their operational needs, including clearly distinguishing between required, optional and experimental features of the service interface. Facility-oriented areas that will receive attention include:

- Robustness – ensuring that the integrated systems work toward the goal of managed, persistent and end-to-end deployment and support. These properties include: operational support

procedures; problem reporting, dissemination, response and tracking; alarm and fault reporting and response services.

- Evolution – of the facilities infrastructure towards a Grid-enabled environment via moving the facilities interfaces into the Grid system stack.
- Change Control – the methods, culture and habits related to operating and using a large data center are constrained by the need to maintain production and support non-Grid use of the facilities;
- Simplicity – as measured by the difficulty of providing independently implemented versions. This can be driven by the need to retro-fit legacy systems, or to preserve a path for innovation of the underlying facilities.

It is an open question as to which services require treatment of this sort. Batch-submission services, permanent-storage services, accounting and authorization services are seen to be candidate services.

Several parallel Grid environments must be sustained to successfully integrate new systems and services into production. Commissioning and integration work has to proceed in a realistically complex and to-scale environment without disrupting or negatively impacting ongoing production work. This requires providing tools to allow a flexible and reconfigurable Grid environment – that is to configure how the Grid maps to the facility services, and to configure and setup the end-to-end services that allows the facility to manage and operate the Grid.

5.2.1 BNL

The RHIC/ATLAS Computing Facilities at BNL will, over the course of the next two years, establish an interface between its storage resources and the Grid-based on SRM. This interface will be common to both the RHIC experiments and US ATLAS and will make use of solutions common to other participants in PPDG. The interface will be evolved to a services-based architecture and connected to appropriate registration and brokering services as rapidly as this can be reliably and robustly accomplished. The local fabric storage technologies behind these services will include the HPSS Hierarchical Storage Management System, NFS served disk and one or two yet to be determined alternated disk technologies. To the extent allocated by the need to maintain a production quality environment for RHIC Runs and ATLAS Data Challenges, efforts will be made to integrate this storage service with evolving Grid authorization and accounting technology.

5.2.2 Fermilab

The Fermilab Computing Division facilities groups will continue the evolution of enabling access from the Grid to the resources provided by to the experiments. Developments will continue to provide access to and use of the storage and processing facilities from the common Grid infrastructure, while allowing local management, control and policy implementations.

In particular, participation in the SciDAC SRM project will provide enhanced capabilities for privilege and access control to disk based storage (DCache) and the petabytes of tape storage through Enstore.

The Fermilab Facilities groups will be working to define and implement the operational interfaces between the local resources and the Open Science Grid, and will work with the PPDG experiments locally to integrate, test and deploy these capabilities.

5.2.3 Jefferson Lab

Jefferson Lab will continue to concentrate on integrating its local storage and batch management systems, Jasmine and Auger respectively, with grid services to support the lab's experimental and theory programs. Development on existing grid services will concentrate on improved robustness as well as functionality. As a contributor on the Storage Resource Manager (SRM) project, Jefferson Lab will develop and deploy an implementation of the version-2 interface definition for use with Jasmine. Jefferson Lab will continue to work with the SRM project on future specifications.

In addition to data grid activities, work on a web services based job submission system and its integration with Auger will continue. This prototype system will manage the submission of jobs to the Jefferson Lab batch farm. It will also provide a means of monitoring and tracking submitted jobs. An API will be provided to the Jefferson Lab experiments so that they can develop custom job submission interfaces and monitoring systems.

5.2.4 NERSC/PDSF

The NERSC/PDSF facility is used by the STAR and ATLAS experiments in PPDG even in the absence of any Grid-related activities. Establishing production-level availability for certain Grid services has been a NERSC goal and exists today for basic Globus-gatekeeper and GridFtp services at PDSF as well as a GSI-enabled ftp service for the HPSS system. X.509 Grid identifiers are integrated in with the user-management database system (NIM). Collaboration in the PPDG-OSG common project brings additional features and services.

NERSC already acquires CPU usage (PDSF) and storage usage (HPSS) information by user and group. We will work with the PPDG-OSG project to work out the means by which this information, in appropriate form, is available for Grid-wide aggregation to enable VO-level summaries of resource usage.

NERSC will work, in the context of the DOE Science Grid project, on the site authorization aspects at PDSF to enable the application of authorization policies appropriate to the VO-wide authorization requirements. The DOE Science Grid project will also collaborate on the system instrumentation and monitoring and troubleshooting aspects of the PPDG-OSG project as it relates to activities at PDSF.

NERSC and DOE Science Grid will work with the PPDG-OSG project as it relates to the “shared” aspect of the resources with regard to policies and user management issues. We understand that the justification for changes to policies and their implementation (business practices) is largely driven by the user community rather than the facilities operators.

5.2.5 SLAC

SLAC facilities are dominated by storage and computing installed for the BaBar experiment. Members of the SLAC Systems Group and Security Group have already been active in PPDG, and in trial deployments of Condor/Globus-based Grid middleware. Experience with preparing reports for the agencies funding BaBar’s US-Europe distributed data analysis environment has already raised many of the CPU and storage accounting issues that will be faced by PPDG-OSG. These groups are already participating in PPDG-OSG planning and have identified the team that will undertake the initial PPDG-OSG deployment at SLAC.

The goals of the initial deployment will be to support compute-intensive production jobs. Simultaneously, the SLAC groups intend to collaborate on the design of the PPDG-OSG approaches to accounting and resource management for storage, expecting that this will have benefits for the operation of the SLAC facility in addition to the benefits for the wider Grid community. SLAC has not, up to now, deployed the SRM storage resource management middleware that has been successful in PPDG. As the most mature PPDG experiment, BaBar had little to gain from an early replacement of its SLAC-developed storage management technology. Serious consideration will be given to the use of SRM.

5.2.6 Exploitation of Ultranet

The PPDG facilities will exploit capabilities emerging from research and developments on Ultranet which plan to introduce awareness and exploitation of advanced networking into data management services into the common data management and transport layers. Production use of “hands-off” sustained data flow with efficient exploitation of dedicated or shared network links is expected by **November, 2005**.

5.3 PPDG Experiments

It is understood that the PPDG experiment teams all work within the context of experiment-specific milestones. While the diversity of Physics experiments, each with its own time scale, would *a priori* suggest that a common goal would be difficult to identify, a closer look clearly shows commonality: all experiments identify production (Monte Carlo or data mining) as a priority with more loosely defined

requirements for user analysis on the Grid as later milestone. Hardening of the existing data management, job scheduling (and convergence toward a common service for job submission) and production readiness starting as early as 2005, or 2007 for the LHC experiments, are strongly correlated with the respective experiments deadlines. PPDG feels that it is part of its mission to support the experiments in their production and user-analysis goals as well as to encourage, extend and foster joint projects across the experiments. Besides the deliverables of the Common Project, and other common work fostered by the computing facility centers, additional PPDG experiment joint projects include:

- SAM-GRID[SAM] was developed for D0 and is now also used by CDF, and being adopted by the MINOS experiment
- JLab, STAR and PHENIX job-scheduling and job-submission service.
- U.S. ATLAS and U.S. CMS collaborative work on prototyping distributed analysis environments in the framework of the LCG-ARDA roadmap for analysis[ARDA].
- JLab, Fermilab, LBNL SRM interface prototyping and implementation and interoperability demonstrations

5.3.1 Common Analysis Services

PPDG participants are uniquely positioned to help ensure both interoperability and sharing of solutions with European Grid projects, and with LCG in particular. A portion of the PPDG-funded ATLAS and CMS effort will be directed expressly toward developing, integrating, and demonstrating deployment of such interoperable and shared solutions. Consonant with this approach, PPDG collaborators will work together specifically in the definition, development, and deployment of common components and services upon which end-to-end analysis can be based.

The groundwork for such collaboration has already been laid. In the PPDG CS-11 a Grid Services Requirements document[CS11-1] has been written that broadly outlines the types of services needed for interactive analysis of complex datasets. Through a series of CS11 related workshops, a list of 18 required common interfaces has been developed[CS11-2, CS11-3], and work has begun to precisely specify those interfaces and to prepare reference implementations beginning with a Dataset Catalog Service Interface[CS11-4, CS11-5], developed to CS11 specifications by SLAC and Tech-X corporation and now under consideration by PPDG and ARDA participants. CS11 will continue to serve as a forum where developers come together to specify these interfaces and share their implementations.

The LHC experiments (ATLAS, CMS, ALICE) have, in the context of HEPCAL II, agreed to a common set of use cases as starting points for specifying common analysis service requirements. The experiments have further assented to a shared Architectural Roadmap toward Distributed Analysis (ARDA) sufficient to initiate an LHC-wide common project in this area, with a kickoff workshop in January 2004. They are incorporating a common persistence technology, LCG POOL [POOL], into their experiment-specific frameworks, and collaborating in POOL's specification and development including support for collections and extended metadata capabilities.

It is beyond the reach of PPDG to propose that all analysis environments be identical, common approaches to higher-level analysis services have been prototyped in earlier PPDG work in CS-11 (JAS[JAS], CAIGEE[CAIGEE] and DIAL[DIAL]), and expanded scope for joint work promises to be fruitful.

Collaboration with the LCG improves PPDG's ability to deliver shared solutions. What is proposed by ARDA is a collaborative program of work, with common specifications and shared approaches to development and deployment of many of the components and services upon which end-to-end analysis will be based. . The approach taken by ARDA has a lot in common with the driving principles of the PPDG collaboratory. Participation in ARDA provides both a forum for shared service definition and a vehicle for ensuring commonality in the PPDG distributed analysis initiative.

5.3.1.1 Analysis Service Definitions – CS-11

CS-11 will continue with the staged definition and implementation of the agreed-upon service interfaces, defining milestones as effort is available. CS-11 members will work with ARDA on defining common interfaces:

- **June 04:** Implementation of the CS11 Dataset Query Service by at least Caltech and SLAC
- **June 04:** Complete specification of three more of the interfaces from <http://www.ppdg.net/mtgs/20mar03-cs11/APIsBriefDefs.doc> (building on existing grid interfaces where available, creating entirely new interfaces only as needed).
- **Dec 2004:** Multiple implementations of the three interfaces selected.
- **May 2005:** Complete specification of remaining ~15 interfaces.
- **Dec 2005:** First reference implementation of all interfaces.

5.3.2 High Energy Physics Experiments

5.3.2.1 BaBar

BaBar's work within PPDG has focused on a collaboration with SDSC to deploy SRB in support of BaBar data movement.

Over the next two years we intend to move BaBar's usage of the SRB to distribute data from a testing and limited use system to a more production quality system capable of handling high user load with maximal reliability. This will require the development of applications to monitor the SRB servers, maintain catalog consistency, automatically register new users from the BaBar VO and manage disk and staging requests. Many of these applications will be built on lower-level applications provided by SRB and other middle-ware applications. In the course of building the full production system for distributing BaBar data we anticipate the possible absence of some lower-level applications. In such cases we will provide the missing applications ourselves for integration into the SRB standard release. In addition, some of our higher-level applications may be of sufficiently general interest to warrant integration back into the standard release.

The BaBar SRB metadata catalog will be related to the BaBar metadata catalog (containing experiment specific information) allowing users to query a richer set of attributes in order to extract the dataset of interest. BaBar will make use of the new SRB peer-to-peer federated environment that will permit users easier access to the data, since one MCAT would have knowledge of the data stored in other MCATs that form part of the same federation.

The task of developing the major applications to achieve and maintain a production system for data distribution is expected to take 1 year from deployment since the need for many of the applications will only become apparent with increased and sustained usage of the system.

Inter-operation of the SRB MCAT with the RLS will permit the possibility of access to the SRB from a wider range of Grid applications such as those provided by LCG or CS11 participants. This would be especially useful for analysis applications that need to submit to a site containing the data of interest. The very first step would be a set of applications that publish in the RLS the SRB MCAT knowledge on the data location. This first stage is expected to be completed in 6 months. The next step would be to allow the RLS to make SRB calls to move data from one site to another. This second step is expected to take more than a year; however, the first step is thought to be of more interest in BaBar. For both these steps BaBar plans to adopt and integrate the applications being developed in the UK.

The experience of using the SRB in production in BaBar will also allow an understanding of the Grid-services needed for data access and distribution. We plan to work on replacing, or wrapping, the services with OGSA/WSRF compliant services. This should allow BaBar resources to be accessible by OGSA/WSRF compliant applications from other experiments. Work on this is expected to start during the first year and a majority of the services identified and in place within two years (depending on the emerging OGSA/WSRF standards).

BaBar also intends to exploit PPDG-OSG resources for production simulation and will encourage a limited number of BaBar physicists to help in the testing of PPDG-OSG analysis facilities.

5.3.2.2 D0 and CDF

D0 and CDF will team to use and consolidate a common set of tools serving both experiments. The experiments have in place a common program of work to continue development of the SAMGrid data-

handling system. D0 and CDF have already converged on a common schema for file metadata and process tracking, which is in production use for both experiments. Both experiments use the data-caching, storage-management and job-scheduling services of SAM for data processing and analysis. The SAM-JIM job-scheduling extensions will be used collaboration wide in 2004 for Monte-Carlo production as well as for the later data re-processing. SAMGrid will be further extended to provide resource management within regional Grids that are part of more local D0 Remote Analysis Centers. Both D0 and CDF will work on developing and solidifying interfaces to the workflow management tool, additionally part of a GridPP[GRIDPP] deliverable, components of which are in common with CMS.

Additional areas of development work in the project include interfacing to SRMs and further refining the system architecture towards a service-based infrastructure. The PPDG work for this proposal will include accounting and authorization. There will continue to be a strong focus on integration with production facilities. As we evolve SAMGrid service-based components to interoperate with other common infrastructures, applications will be deployed on the Open Science Grid in the US and the common Grid infrastructures used by collaborators in Europe and elsewhere. We will continue to work with our Computer Science partners toward hardening and enhancing job submission, authorization and accounting capabilities. We also plan to work on end-to-end application through fabric monitoring, and software in support of troubleshooting. There is a strong ongoing partnership of the D0 and CDF PPDG team with the GridPP project in the UK. From this project both experiments expect to benefit from further development of a common workload-management tool, RUNJOB, also in collaboration with CMS.

The experiment milestones which are driving the PPDG and other Grid work for CDF and D0 are:

- Use JIM job scheduling for D0 Monte-Carlo production: March 2004. Test JIM job scheduling at CDF: **June 2004**. Use JIM job scheduling on remote resources for next round of D0 reprocessing: **June 2004**. Begin using JIM job scheduling as part of the D0 RAC infrastructure: **December 2004**.
- Prototype services which use more general Grid resources, in particular run tests for CDF and D0 executables on Open Science Grid resources: **December 2004**.
- Use remote resources for 50% of CDF user analysis cycles: **June 2005**. Further develop and solidify interfaces to the workflow-management tool for D0 and CDF (a GridPP deliverable): **June 2005**.
- Begin to add accounting and authorization capabilities to SAMGrid: **June 2005**.
- Further develop the SAMGrid operational support model, including the specification of performance metrics: **June 2005**.
- Further develop monitoring and troubleshooting capabilities: **December 2005**.
- Further develop the SAMGrid operational support model, and raise targets on performance metrics: **March 2006**.
- Achieve broader analysis and production access to general HENP Grid resources for CDF and D0 by evolving SAMGrid to greater interoperability with these resources: **June 2006**.
- Further develop accounting and authorization capabilities: **June 2006**.

5.3.2.3 ATLAS

The DC2 data flow planned for the first half of 2004 follows ATLAS mandates for computing model testing. The Tier-1 centers and the distributed Grid2003 environment which include the prototype Tier-2 centers will be used to generate up to 10^7 events from April to June. Validated data will then be transferred back to the CERN Tier-0 center where they will be reconstructed. DC2 is also envisioned to test distributed analysis on the Grid via access to event and non-event data flow anywhere in the world both in organized and chaotic ways.

The second ATLAS production milestone will be to support its DC3. Envisioned to be a similar exercise than DC2, it will however differ by its scale and scope: the target is to deliver a full set of ATLAS software running in the Grid environment – an increase in complexity and data throughput. During DC3, interoperability with the LCG will be demonstrated. Both DC2 and DC3 have a production and an analysis phase. The production phase will be carried out largely under central control, using the common US Grid fabric used by US ATLAS and US CMS. This work will also be coordinated with CMS and distributed analysis efforts.

By 2006, the ATLAS collaboration will undergo its final approach towards LCG startup. The current schedule calls for a Physics readiness report by mid 2006 and make any final corrections for the beginning of the LCG run in 2007. By then, the infrastructure for distributed production and analysis should be fully in place and operational.

To reach those goals, ATLAS proposes the following PPDG:

- Implement authentication and authorization mechanisms for ATLAS services.
- Deploy storage-management services (SRM) for Tier-1 and other sites in the US. This is described in the facilities section for BNL.
- Prototype activities for Grid metadata services and access to and management of distributed datasets.
- Draft a specification for a hosting framework for ATLAS services required for distributed processing.
- Define or adopt interfaces and specifications for service behavior and workflow systems. High level candidates include AJDL (Analysis Job Description Language) that defines the scope of and interface for the high-level analysis and catalog services directly available to users. An ATLAS implementation of AJDL suitable for its immediate needs, e.g. to process AOD data with Athena. This will be coordinated with GriPhyN/iVDGL/US ATLAS work to produce a process execution environment employing Chimera+CondorG. Much of the integration effort will come from outside PPDG.
- Develop an interactive analysis service meeting the above interface. The analysis service itself will rely on common Grid transformation services which, for batch mode processing, may be shared with production services. We will use DIAL to implement an interactive analysis service meeting the above interface.
- Develop a ROOT/CINT client that would enable access to this and other AJDL services from ROOT.
- In 2005, the effort may need to concentrate on the performance of the system(s), in collaboration with the PPDG Common Project. These activities include:
 - Defining and implementing performance metrics, including rates for job submission, data production and transfer, response time for interactive analysis, query time for cataloged data, etc.
 - Identifying and addressing scaling problems associated with CPU counts or data volume.
 - Identifying and addressing latencies in the interactive analysis services. There will likely be different issues for different dataset sizes

Successful deployment of ATLAS analysis applications is predicated upon integration with the ATLAS collaboration's data and metadata management. Such deployment and integration is therefore an essential focus of U.S. ATLAS PPDG activities. ATLAS and CMS both rely on LCG POOL software for their persistence infrastructure. Successful POOL integration requires resolution of a number of technical issues such as replica catalog compatibility. One must come to terms with the fundamental role that event collections, often hosted in relational databases, play in physics analysis. The U.S. ATLAS PPDG team will deliver:

- **December 2004:** integration and deployment of a prototype system sufficient to assess POOL-based ATLAS applications on OSG-managed resources;
- **June 2005:** access for Grid-based applications to ATLAS metadata services based upon joint ATLAS/CMS development done in the context of POOL and ARDA, in close collaboration with Globus/GriPhyN providers of metadata services and functionality;
- **June 2006:** OSG deployment integrated with more general ATLAS analysis environment developed in collaboration with ARDA, and with U.S. CMS collaborators in particular.

5.3.2.4 CMS

The CMS experiment plans several demonstrations and data challenges on the Grid, to get the distributed computing environment and software ready for the start of the LHC in 2007. The US CMS PPDG Team will continue to be directly involved in these and to develop and maintain the US Grid Testbed infrastructure and software required and used by CMS. Areas of work will include MOP, MCRUNJOB, and the development of the Open Science Grid, and helping with use and support of Grids in the experiments.

CMS will work towards interoperation between the US Grid (with Fermilab providing major Tier-1 resources) and the LCG, and will federate its US Grid resources with other LHC computing resources under US CMS governance. US CMS will work with the regional centers and Grid projects to define a roadmap towards the US federated Grid, to build the Open Science Grid. To achieve this objective, US CMS will work with PPDG to define and develop a sustainable operation, integration and development model for the US Grid, work on a plan for selecting and integrating middleware components and Grid services and on hardening of the Grid environment and enhancements of its functionalities.

The CMS program of work schedules a major CMS data challenge in 2004, DC04, which will demonstrate the capability and scalability of a sustained dataflow to Tier-1 centers, with Fermilab in the US. DC04 will also have a data analysis component with calibration and analysis demonstrations at the US Tier-1 and Tier-2 centers. Data samples will be made available to remote Grid users by mid 2004. Data challenges follow in 2005 and 2006, related to the regional center and physics technical-design reports. The LHC Grid data centers will be brought to operation at incremental fractions from 20%, 50% to full capability in respectively 2007, 2008 and 2009.

Together with the PPDG analysis efforts, with US ATLAS and the LCG, US CMS will work towards a distributed analysis environment on the Grid. The interfaces between Grid services and CMS metadata services will be developed, and a CMS data model for distributed analysis will be defined and implemented. US CMS plans to work with US ATLAS on POOL metadata and experiment metadata services and event collections and integration of these services into the analysis environment and the experiment frameworks.

Development of the Grid Analysis Environment (GAE) at Caltech will continue in collaboration with other US CMS institutions. Particular PPDG deliverables will include:

- *Extensions to the Clarens Web Services Framework.* Clarens provides a specification for and implementation of Grid-enabled web services using standard protocols and tools. The Clarens framework is gaining acceptance and use by a variety of applications for LHC experiments. The development of Clarens in an open and collaborative fashion in response to the needs of the user community will continue, with a focus on interoperability, standards compliance and robustness. The current Clarens implementation is as a classical passive web service, so it responds to requests in a client-server system. Work on the Caltech GAE has shown that for support of interactive-analysis use, the server needs to become a more active participant in its environment. This requires it to seek out service peers through lookup services and in turn to advertise those services, as well as to provide mechanisms for asynchronous server-client communications through persistent two-way connections to service users that support this capability. This active mode also implies more active internal state management of sessions, connections, as well as policy enforcement on resource usage. Specific services will be developed based on this framework as part of the GAE, as well as the client applications that make use of these services. Specific emphasis will be given to interacting with the ROOT analysis framework, and

providing an extensive browser-based interface to services. Releases of server and client packages will be made as new functionality is implemented for testing and deployment as part of common software distributions such as the Virtual Data Toolkit (VDT), and supported in collaboration with the Clarens user community. Milestones:

- Enable clients to access the Pool metadata catalog via the Clarens web service environment. Enable IGUANA (CMS visualization client) to browse meta-data in POOL catalog via web services. **(May 2004)**.
- Deployment of second version of lookup service for dynamic discovery. Multiple instances of service will be deployed. Users and services can discover services, minimizing single point of failure. **(September 2004)**.
- *Global Interactive Grid System-User Dialogs*: Dialogs will be developed to allow users to view the state of the Grid and its evolution, giving feedback on the progress of work scheduled, and to allow users giving input to decisions that affect task scheduling and workflows. This can be extended to allow the user to also give guidance for alternative strategies for completing their work more efficiently. Using the interactive input and constructing user-system dialogs treating a broad set of work-scheduling and problem situations, will lead to progressively more effective decision support and guidance to users performing Grid-enabled Analysis.
- *Distributed System Optimization*: Once initial strategies are developed and effective, self-learning adaptive optimization algorithms, such as self-organizing neural networks [SONN1] will be used to optimize system performance, according to metrics that take into account the policies for priority and resource usage of the Virtual Organization.

In order to make the proposed Grid Analysis Environment more effective and responsive, we will also integrate networks as an actively managed resource into the overall Grid system.² We will integrate these developments into the mainstream data-management services of these experiments, and take the further development steps necessary to do this successfully, using advanced network infrastructures such as the DOE Science Ultranet as the key prototyping network.

5.3.3 Nuclear Physics

The Nuclear Physics experiments, with up to Petabyte-scale data samples every year (and growing) as well as with large number of potential early users of the Grid, offer a superb and immediately available test-bed for Grid infrastructure and developments. In fact, early running experiments brings to the project a unique and rich (user) environment in which base components of the Grid can be stress tested and therefore, allow for catching possible problems at an early stage (including redesign and/or design requirements change in directions). This is as a major strength in the path toward robustness and scalability as components can be immediately assessed and rectified by our community prior to operation by experiments which, by international agreements, rely upon Grid performance for their success.

5.3.3.1 Jefferson Lab

At Jefferson Lab, the highest Grid priority for the laboratory for both the experimental program and the theory program is a robust data Grid. As Jefferson Lab has been a key contributor of ideas and prototyping for the Storage Resource Manager (SRM) project, expanding this effort into related data Grid components (Replica Catalog, Meta-data Catalog) is a high priority for the coming years. This data Grid, consisting of

² This will be addressed by introducing awareness and exploitation of advanced networking, embodied in recent developments in high performance TCP protocol stacks geared for fair-sharing up to the 1-10 Gbps range ([NetL1]; see netlab.caltech.edu/FAST/), global end-to-end monitoring systems ([MONA], see monalisa.cacr.caltech.edu), and a new network paradigm involving the dynamic construction of optically switched paths over national and transoceanic distances (see for example <http://www.startup.net/translight/>, and ultralight.caltech.edu), with realtime routing and switching decisions to handle massive dataflows between multiple sites efficiently while directing small and medium-sized traffic flows to shared, statically provisioned production networks.

interoperable implementations of community-standardized interfaces, will support the movement of data-analysis products and simulation products (detector and theory) alike. The PPDG effort at Jefferson Lab is a collaborative effort between the lab's Computer Center, which supports the experimental program, and the High Performance Computing groups, whose main customer is currently Lattice QCD. PPDG resources augment the programmatic efforts being expended towards their distributed computing goals, and provide for valuable interactions with others within NP and HEP doing similar work.

To reach a production quality and interoperable Grid environment, JLab will implement and deploy SRM version 2.x as a web service. This service will interact with user files, disk cache and Jasmine [JASMINE]. As part of the joint STAR-JLab common project, a prototype for describing user requests and WSDL defined interface for the submission, tracking and modifications of jobs will be delivered as a basic batch service supporting job translation to local batch systems. A consolidation of this work would include the delivery of the STAR-JLab common project schema specification to the PPDG collaboration.

Part of the 2004/2005 effort will concentrate on the hardening of the production environment deliverables from the previous years. Toward the end of 2004, a prototype implementation of SRM, based on specification 2.x, will be delivered. Subsequent versions of the SRM interface specification will be modified to support more advanced access control. Prototype implementations of future specifications will be deployed in a multi-site context. Additionally, JLab intends to demonstrate a multi-site operation for experimental physics batch simulations. This demonstration will include 3 to 5 sites utilizing the STAR/JLab common job description, batch web services, and the SRM. Later, more advanced batch web services, taking advantage of multi components, will be put into operation on resources from JLab, FNAL and BNL. This will include additional functionality for job tracking, job control, reporting, file migration using the SRM, and the registration of job results into the Replica Catalog. Finally, in 2006, JLab will start to integrate into its production environment VO tools and advanced Grid components for accounting, debugging, and scheduling.

5.3.3.2 STAR

The *STAR* experiment faces several immediate challenges driving its Grid efforts. The current RHIC run (Run4) will produce a data set larger by factors of 5 to 10 in the number of events, leaving the experiment with longer production cycles and the need to empower its users with tools to seamlessly access computing resources across its facilities.

STAR proposes to converge and finalize the U-JDL specification [PPDG-39] and will prototype, as part of the STAR-JLab joint project, a batch scheduling service in accordance with emerging common interfaces. The high level U-JDL is viewed as a key component to unify the way we describe and execute a job on the Grid.

Grid-based Monte-Carlo production is seen as achievable in early to mid-2004 and will start to migrate from the centralized Tier-0 center to a Tier-1 distributed processing model. Real data processing will be tested as a way to better understand and extract the full set of requirements for stable production. STAR experiment data mining on the Grid raises several issues not currently addressed in its infrastructure (database unified service and discovery). Internal (non PPDG) manpower will be devoted to the consolidation of the production environment and the introduction of new components. Batch oriented user analysis will start by mid-2004. At this stage STAR will start to integrate and make use of services addressing accounting, authorization and troubleshooting emerging from the PPDG-OSG common project deliverables as applicable. Interactive analysis will also undergo further investigation by expanding on the collaborative work with the SDM center: a first version of a STAR analysis framework on Grid, called the GridCollector, was released, tested and put into operation. Further development and integration of this analysis framework within its core toolset is an essential component for STAR analysis completeness.

Capabilities for scheduling algorithms with the inclusion of monitoring information will be employed to increase the efficiency of job execution. We will continue the support, enhancement and consolidation of the data migration tool (HRM/SRM/DataMover). STAR will collaborate with facility personnel in support of SRM deployment. Given the STAR experience with the initial Replica Registration Service (RRS) implementation within HRM and its dependence on RRS for its production on the Grid, STAR will continue this activity in collaboration with the SRM project [SRM]. Finally, STAR will also start to integrate more facilities within a common Grid infrastructure as the base for user analysis and simulation.

We foresee the success of running on non-STAR PPDG-OSG resources as depending on the availability of several components and will draft a requirement and specification document in late 2004.

Beginning in 2005, STAR will undergo detector and machine developments and will be in need of significant CPU cycles for simulation. By then, data production will be moved to the Grid and a first prototype of an interactive user analysis framework using Grid components will be delivered and operational. The Batch submission system will introduce job tracking and control and will be deployed in other participating facilities and demonstrated as interoperable with other Grids. It is intended to use the PPDG-OSG resources as available for a larger scale production at this stage.

In 2006, STAR will aim to achieve full access to Grid resources at universities and collaborating institutions (in the US or across the world). HRM/SRM/RRS based data movement on the Grid will be the technology of choice. Hardening and enhancing the capabilities of the resource brokering will also be implemented. Inclusion of local policies for accounting and authorization will be gradually phased into use at the federated sites.

5.3.3.3 PHENIX

PHENIX, as all RHIC experiments, will concentrate on Run4 data taking, producing a data volume 5-10 times larger than in previous years. Run4 data production will be done on RCF and the PHENIX computing facility in Japan; successful reconstruction, analysis and publication of Run4 results are the primary goal for the PHENIX computing group. We will, however, continue joint development with STAR on the Batch Submission System, adding a PHENIX user interface. We will deploy Grid production of simulated data using these tools, with a demonstration in summer 2004, and utilization by PHENIX users by the end of the calendar year.

PHENIX has begun collaborating with facility personnel in support of SRM deployment. We will continue this work, with the goal of developing a solution applicable to all experiments using the BNL facility.

In 2005, we aim for production use of Grid tools and infrastructure.

5.3.3.4 ALICE

An intention by ALICE-USA to collaborate with PPDG is included in the Appendix. It describes how the ALICE-USA grid-related activities align with the PPDG program and will begin as soon as the level of available effort increases sufficiently (pending approval from the funding agency).

5.4 Computer Science teams work plan

Four computer-science teams – ANL, LBNL, SDSC, and UW – participate in the PPDG collaboratory. In addition to their contributions to the common projects, the four teams will continue to facilitate the integration of common Grid technologies into the software stacks of the PPDG experiments. Most of these technologies are developed and maintained by members of the PPDG team under other funding sources. Changes to the software stack of an experiment mandates close coordination with the experiment and may entail enhancement to or adaptation of existing software tools. Given the goals of our collaboratory, we plan to allocate only limited resources to the development of new software. We expect that most of the computer sciences effort will continue to be devoted to the hardening, deployment and support of existing software. Given the wide acceptance by the HENP community of the Virtual Data Toolkit (VDT), we will align to the extent possible our activities and timelines with the testing and release procedures of VDT. Including PPDG software components in the VDT is likely to broaden the impact of PPDG, enhance the overall quality and functionality of the toolkit and reduce the “ownership cost” of our software.

In addition to the requirements for end-to-end capabilities generated by the experiments, the over all program of work of the computer-science teams will be guided by the ongoing transition of Grid technologies to a service centric model. As a new generation of Grid technologies that follow this model becomes available, an effort will be made to transition the PPDG software stacks to the new technology. Formulating all the new middleware developed by PPDG in terms of services and re-factoring of existing middleware components into services will simplify such a transition. The LCG has recently launched an effort to prototype a common set of services for distributed analysis. Members of the PPDG computer science teams will participate and contribute to this effort.

The following four subsections outline the program of work for each of the computer science teams. The milestones and deliverables do not cover activities performed as part of the common projects.

5.4.1 Globus Toolkit (ANL team)

Services provided by the Globus Toolkit[GT] can be found today in all the software stacks of PPDG. Enhanced services from the Toolkit are planned for and required by PPDG experiments and facilities. These enhancements will be realized by the integration of more recent revisions of the Globus Toolkit into the VDT and into the software stacks of the experiments. This includes GT3.2, an OGSA based release of the toolkit that offers significant improvements for stability, usability and performance; and GT4.0, the evolution of the toolkit that will support WSRF based web services components. Experience shows us that hardening of new capabilities and components requires planning, effort, careful tracking and support as well as a commitment to support versions of the Globus Toolkit still in use until transitions are made by all PPDG member experiments. The ANL team will continue to offer the necessary guidance and support to insure that the PPDG experiments fully benefit from the common Grid technologies offered by the Globus Toolkit. The team will focus their efforts on the following three service categories:

Accounting and Security – Improvements to accounting and security include the addition of the Community Authorization Service[GT-AS]. The Community Authorization Service or CAS is an authorization system that allows for fine-grain (e.g. per file) access control of Grid Resources. The CAS Grid Service allows for creating and administering rights, client tools and libraries for using CAS, and a CAS-enabled GridFTP service. CAS uses the OASIS Security Assertion Markup Language (SAML) standard. Improvements to the Globus Toolkit Security component (GSI)[GT-GSI] include third party libraries, Certificate Revocation List (CRL) support in Java, support for pluggable authorization in GridFTP and improved documentation.

High Performance Data Transfer – GridFTP[GT-FT] improvements include support for platform independent, machine parseable, file information (MLST and MLSD commands), a new checksum command (CKSM), and support for file globbing and directory moves. Globus-url-copy will no longer be restricted to a single file and will have command line options for all the equivalent client library functionality. The new version of the Reliable File Transfer Service (RFT)[GT-RFT] will take advantage of the enhanced functionality of GridFTP and will expose them as services. The schemas of the service will be reworked to increase the scalability of the service. A new service instance level option will be added that turns on data channel reuse if your request consists of multiple transfers between same hosts, enhancing performance. In past RLS [GT-RLS] versions, RLIs summarized mapping information from one or more Local Replica Catalog (LRC) nodes. The new version of RLS will support an arbitrary hierarchy of RLI nodes, where each RLI summarizes state information for one or more lower-level LRC and/or RLI nodes.

Resource Management - Improvements to existing GRAM[GT-RM] functionality include local scheduler scripts that reduce the chance of errors due to NFS caching delays and increase scalability by reducing memory consumption and internationalization of MJS fault reporting

5.4.2 Storage Resource Management (LBNL team)

Integration of DRM v2.0 with NeST and inclusion in VDT: The current version v.1.x of SRMs has the basic functionality of allocating space by automatic quota assignment, supporting default lifetime policies. The new version v.2.x of SRM [GSS'03] has explicit space allocation negotiation capability, explicit lifetime negotiation, and full support of directory capabilities. This specification which was led by the LBNL team, is a joint effort of people from JLab, Fermilab, and the European Data Grid (EDG). The LBNL team plans to implement SRM v.2 and use NeST (developed by the UW) to perform the space allocation using the “lot” assignment capability. This joint effort will culminate in the inclusion of the SRM-NeST product in VDT. In addition, the LBNL team plans to develop a new HRM version for HPSS based on the SRM v.2 specification. This implementation will be compatible with other SRM implementations in JLab, Fermilab, CERN, and Rutherford Lab.

Interaction of SRMs with other Grid middleware services: SRMs are essential for Request Planners and Request Executors. They are needed to provide information on storage resource availability, for supporting space reservation with lifetime policies, and for interacting dynamically with Request Executors to pin and release files. SRMs can also provide to monitoring tools the state of storage usage levels, queues of jobs waiting for file movement, and information on files they currently store. Another essential capability that

SRMs can provide is the enforcement of access-control policies that are provided by components that support virtual organizations, such as Community Authorization Service (CAS). Features to support such interactions with other middleware components are planned to be added in SRM version v.3.x. As these middleware products mature, we will work on the active interaction of SRMs with Grid components that perform storage planning, file movement, storage monitoring, and enforcing storage policies for executing Grid jobs.

Replica Registration Service: The need for a Replica Registration Service (RRS) was observed as a result of large-scale file replication performed routinely by PPDG experiments. Since in practice each of the PPDG experiments maintains a “File Catalog” specific to its needs, it became apparent that a uniform RRS that can be applied to various catalogs is a valuable concept. RRS implementations may need to be specialized for the catalogs they target. For example, some catalogs may only be able to register a file-at-a-time, while others may be able to register in bulk. The RRS can also be used to un-register files. It is common that shared disk caches are used for replicated files. However, when space is needed files or entire directories are removed either by an administrator or by an automatic process that manages the cache. Here again, there is a need for a service that will un-register the files to be removed. The RRS can be designed to perform this service as well. The LBNL team has developed a simple prototype of the RRS for the STAR experiment. This work is still in progress, but has shown its value in streamlining the file replication task. We plan to define a formal interface specification for RRS, and involve multiple PPDG experiments in using the same interface for their specialized catalogs. This will facilitate the development of a common Replica Management Service that consists of interacting with Replica Catalogs including RLS (developed by the Globus Alliance), file movement services including the SRM DataMover and RRS as sub-services.

5.4.3 Storage Resource Broker (SDSC team)

The Storage Resource Broker (SRB) Data Grid provides data management mechanisms for the organization of distributed files into a collection. Data Grids implement distinguished name spaces for storage resources, users, files, and metadata about the files or collection context. Each distinguished name space supports mappings from a logical name to the physical name that represents the entity. SRB has been used by two of the PPDG experiments to support their data movement and file cataloging needs.

As new versions of SRB are released SDSC team will continue to facilitate and support the integration of the new versions with the software stack of the BaBar experiment. The most recent version - Version 3.0 - provided the fundamental mechanisms needed for peer-to-peer federation. To improve support for the next phase of data challenges, the data grid federation mechanisms of SRB will be extended to enable:

- specification of operations that can be performed on digital entities when dealing with both raw experimental data and derived data products. This includes support for a web services-based data flow processing system, which is able to retrieve result sets from collections, process each member of the result set, and store the derived products back in a SRB collection. The system will be upgraded to support interaction with Open Grid Services Architecture services.
- integration with the next generation GridFTP transport mechanism

The schedule for these development efforts will track technology developed within the Global Grid Forum, and development of the Open Grid Services Architecture.

5.4.4 Condor Project (UW team)

Directed Acyclic Graph Manager (DAGMan) – DAGMan is used by several PPDG experiments to manage the execution of inter related jobs. The current implementation of the manager addresses the most basic capabilities required to support real life scenarios. Over the last two years the UW team has accumulated a long list of request for enhancement in functionality, scalability and performance. The UW team will continue to rank these requests and implement them in priority order. Functionality features currently at the top of the list are – support for dynamic DAGs, conditional arcs and logical grouping of nodes. As storage

management and data placement services become available, DAGMan will be extended to support data placement jobs and to interact with the Stork scheduler developed by the Condor project.

Computing on Demand (CoD) – The interactive nature of analysis activities requires schedulers capable of supporting the “time sharing” of resources. At SC03 we demonstrated a PROOF-based analysis environment that used the CoD capabilities of Condor to multiplex analysis and simulation tasks on the same set of Condor-managed resources. The UW team plans to continue the effort to integrate CoD with different distributed-analysis prototypes and to develop interfaces to schedulers that support advance reservations of storage and compute resources. The interaction between the timesharing requirements of CoD and the remote-execution services supported by Condor-G will be studied and mechanisms to support interactive workloads will be developed. The UW team will perform the work on CoD in close coordination with the LCG-ARDA project and PROOF team at MIT and CERN.

Classified Advertisement Language (ClassAd) – The ClassAd language has gained popularity among the HENP community as means to describe jobs and resources and as a means to implement resource-allocation tools that are based on policy driven “matchmaking”. The UW team will be involved in the effort to develop a Job Description Language for analysis jobs and will implement extensions and enhancement to the language as needed. An important element of the runtime support library of the ClassAd language is collections of persistent ClassAd objects. The UW team will work on improving the functionality and performance of the support library. As part of this effort a prototype of a relational backend for such collections will be developed and evaluated. The recent evaluation of an AML backend points to a relational approach as promising solution.

Accounting – Effective and reliable accounting holds the key to the acceptance of a shared facility that consists of distributively owned resources and services. Accounting information will be needed to keep “bean counters” happy as well as to empower policy driven schedulers. Supporting such an accounting system in a widely distributed environment will require an integrated collection of secure and reliable information recording, storage and transfer services. The UW team will develop principles and technologies to address these requirements, taking advantage of existing results in these areas. This activity will be closely aligned with the functionality and timeline needs of the common project.

The Condor project hosts the integration, packaging and support infrastructure of the VDT. The UW team will contribute resources at the level of .5 FTE to the VDT effort and will work closely with the VDT team to support the integration of PPDG-developed middleware and to coordinate the content and timelines of new releases. The VDT activity will also facilitate a technical communication channel the LCG/EGEE communities. We expect that the PPDG experiments will also play an important role in the testing and hardening of the components of the VDT. This may lead to a common testing framework, testing tools and procedures, and testing facility across all major Grid projects.

5.5 PPDG Milestones

We show a summary of the milestones for the project:

Milestone Summary for Year 1		
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	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	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	PPDG-STAR	Batch oriented user analysis will start.

Milestone Summary for Year 1		
Sep-04	PPDG-CMS	Deployment of second version of lookup service for dynamic discovery. Multiple instances of service will be deployed.
Sep-04	PPDG-OSG-Troubleshooting	Specify Baseline Capabilities
Oct-04	PPDG-OSG-Accounting	Specification and Requirements for Accounting Version 1
Oct-04	PPDG-OSG-WMGT	Process Wrapper (ProRap) for remote execution of HENP applications V1
Nov-04	PPDG-SRM	SRM Implementation V2
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 105 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
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 Interfaces

Milestone Summary for Year 2		
Jun-05	PPDG-ATLAS	Access for Grid-based applications to ATLAS metadata services
Jun-05	PPDG-STAR	STAR, data production will be moved to the Grid and a first prototype of an interactive user analysis framework using Grid components
Jun-05	PPDG-SAMGrid	Use remote resources for 50% of CDF user analysis cycles
Jun-05	PPDG-SAMGrid	Begin to add accounting and authorization capabilities to SAMGrid
Jun-05	PPDG-OSG-RRS	Replica Registration Service Implementation V1
Jun-05	PPDG-BaBAR	RLS to make SRB calls to move data from one site to another
Jun-05	PPDG-SAMGrid	SAM operational Model
Jun-05	PPDG-SRM	Specification V3
Jul-05	PPDG-OSG-Troubleshooting	Implementation of Baseline Capabilities
Sep-05	PPDG-OSG-Authorization	Specification and Requirements V2
Sep-05	PPDG-OSG-Monitoring	Updated Specification and Requirements V2
Oct-05	PPDG-OSG-Accounting	Specification and Requirements for Accounting V2
Nov-05	PPDG-OSG-Troubleshooting	Specification of enhanced capabilities
Dec-05	PPDG-OSG-Monitoring	Instrumentation of selected end-to-end data analysis applications consistent with the data transport instrumentation infrastructure
Dec-05	PPDG-DMGT	Use of Enhanced GridFTP
Dec-05	PPDG-OSG-Accounting	Implementation of Accounting V1.1
Dec-05	PPDG-OSG-Monitoring	Enable MonaLisa clients to obtain feedback (through Clarens or other web service infrastructures)
Dec-05	PPDG-SAMGrid	Monitoring and Troubleshooting V1
Dec-05	PPDG-CS-11	First reference implementation of all interfaces
Mar-06	PPDG-SAMGrid	SAMGrid operational support model enhancements
May-06	PPDG-OSG-Authorization	Implementation V1.1
May-06	PPDG-SRM	Multiple implementations of SRM V3

Milestone Summary for Year 2		
Jun-06	PPDG-OSG	Sustained exploitation of a significant fraction of the DOE-HENP resources
Jun-06	PPDG-OSG-Accounting	Implementation of Accounting V2
Jun-06	PPDG-SAMGrid	Analysis and production access to general HENP Grid resources for CDF and D0
Jun-06	PPDG-SAMGrid	Further develop accounting and authorization capabilities
Jun-06	PPDG-ATLAS	OSG deployment integrated with more general ATLAS analysis environment developed in collaboration with ARDA
Jun-06	PPDG-JLAB	Integrate VO tools into the production environment
Jun-06	PPDG-STAR	STAR will aim to achieve full access to Grid resources at universities and collaborating institutions
Jun-06	PPDG-OSG-Troubleshooting	Upgraded implementation
Jul-06	PPDG-OSG-Authorization	Implementation V2

5.5.1 PPDG Management Plan

The PPDG Collaboration believes that the management structure described in the original PPDG SciDAC proposal has worked well. The proposed management plan builds on this experience.

U.S Physics Experiment computing management and Computer Science project management joined the PPDG proposal with a common vision to work together to develop and apply common Grid technology to the benefit of each experiment, and with a commitment to adapt, extend and reuse information technology and software used in one experiment for the benefit of other experiments. The PPDG management plan is designed to provide sufficient coordination of the specific project deliverables and milestones, as well as show our commitment to a process of working together towards the common goal.

The obvious challenge of PPDG, the optimal coordination of all Grid-related efforts in spite of the time pressures and physics-focused mission of the experiments, is also an outstanding opportunity. The collective resources that the experiments and CS projects can bring to bear on Grid enabled developments of common interest provide a many-fold leverage of the proposed PPDG funding. Working together closely on these developments will provide both communities with the best environment to meet their scientific goals.

The PPDG management strategy has top-down and bottoms-up components. The bottoms-up strategy involves exploiting the mission-oriented drive of the physics experiments, part of which is already directed towards strengthening collaborations with PPDG-CS and with U.S. and European Grid projects outside of PPDG. The involvement of the CS projects is driven by their desire to test and evaluate their technology in a real-life setting and to be exposed to real-life distributed applications.

The top-down strategy must ensure the steering of the powerful bottoms-up forces to ensure that, while the individual goals of each of the experiments and the CS projects are met, most developments are of value to other PPDG collaborators, and with little or no additional generalization, to a wide user community. In this renewal proposal, the collaboration is strengthening the top-down managed component of PPDG, by adding one FTE of funding for technical management of the Common Project, including contributions to the Open Science Grid, and by agreeing that 25% of the remaining PPDG resources will be devoted to and managed by the PPDG Common Project.

PPDG will continue its two-component management structure with:

1. An **executive team** composed of Ruth Pordes (PPDG Steering Committee Chair), Doug Olson (Deputy for Experiments and Facilities) and Miron Livny (Deputy for Computer Science). An additional member of the Executive Team (to be identified in the next three months) will have responsibility for the day to day management of the Common Project. Members of this team have extensive experience in managing software development and deployment projects and will each make PPDG a principal activity. They will be tasked to track the goals and work of the physics-CS activities teams and provide guidance to ensure overall coherence of the PPDG Collaboratory Pilot. Miron Livny will steer the project deliverables towards maximal commonality (and wide usability) in the components of each experiment's vertically

integrated Grid software. The team will work together to advise the Steering Committee to best meet the short and long term goals of the Collaboration.

2. A **Steering Committee** (SC) comprising one representative from each physics experiment, one representative of each Computer Science group, and a representative for each of the DOE lab computing facilities participating in the PPDG-OSG. The project PIs and the executive team will be ex officio members of the Steering Committee. In the interests of keeping the SC small and effective, the project PIs and members of executive team may also act as experiment or CS group representatives if they wish. A current list of members is maintained at www.ppdg.net. A number of liaison representatives are included on the PPDG steering committee mail list to ensure good ongoing communication.

The steering committee is structured to provide decisive management and to balance the Computer Science, software technology, physics-experiment needs and budgetary constraints. In light of dynamic nature of the technologies we use, requirements we face and time lines we have to follow, the steering committee has been playing a crucial role in changing our work plan and the allocation of funds as we go.

6 Partnerships and Collaborations

6.1 SciDAC

6.1.1 DOE Science Grid Collaboratory

Like PPDG, the DOE Science Grid project (DOESG) is aiming at a persistent sustainable Grid infrastructure for science at a set of non-HENP DOE labs (NERSC, PNNL, ORNL, ANL, LBNL) and with a very useful overlap with PPDG at the NERSC/PDSF facility. PPDG will interact with the DOESG on the whole range of PPDG-OSG common project issues as they relate to the Grid services at PDSF and, consequently, to standardizing interfaces to Grid infrastructure across a broad range of DOE facilities. The usage accounting information now being collected across NERSC will be a very useful example as PPDG works to achieve useful accounting information for a shared Grid.

In addition to the facilities overlap with PPDG, the DOESG has been and continues to be very helpful to PPDG in the area of identity management. DOESG was instrumental in establishing the DOEGrids CA (www.doeGrids.org) that is widely used by PPDG. The future DOESG work planned at improving the usability of PKI in the Grid context will be very valuable to PPDG.

6.1.2 Storage Resource Management SciDAC Project

The Storage Resource Management (SRM) middleware project [SRM] at LBNL initiated the development of dynamic storage management concepts and support for multi-file requests. This work benefited from the cooperation and close collaboration with PPDG members, especially from Fermilab (who are participant of the SRM project), Jefferson Lab, and BNL. This collaboration resulted in benefiting the PPDG as well, by providing a common interface definition for implementations of compatible storage resource managers of different mass storage systems (HPSS at BNL and LBNL/NERSC, Enstore and dCache at Fermilab, JASMine at Jefferson Lab, and Castor at CERN). These SRMs were applied in several experiments including LQCD, CMS, and STAR. Because of the collaboration with the European Data Grid there is now strong interest by the LCG to use SRM interfaces to various storage systems that support SRMs.

Another example of a mutual benefit between the SRM and PPDG projects is the development of the DataMover on top of the SRM-HPSS system developed by the LBNL team. This software now uses two SRM-HPSS systems, one at BNL and one at LBNL/PDSF to routinely moving entire directories of file robustly (with recovery from failures) as part of the production system.

The SRM work also benefited from the NetLogger technology. It was found to be ideal for tracking the behavior of file movement. By keeping track of events over files (such as staging from tape to disk, transfer over the network, and archiving to tape) it is possible to see patterns of behavior and identify bottlenecks. We have this technology to monitor the scaling of the DataMover to thousands of files, and to verify the correct recovery behavior of DataMover when failures occur.

6.1.3 Scientific Data Management (SDM) Center

Part of the work at the SDM center [SDMC] includes a special indexing method (called FastBit) based on bitmap technology. It is designed to provide a fast index to high-dimensional data, which is typical of summary data in HENP applications. In particular, searching over millions or even billions of events by their properties (such as total energy, momentum, and number of particle of certain types) can be very slow. The LBNL team involved in the SDM center has collaborated with the PPDG members of the STAR project (at BNL and Kent State) to apply FastBit to this problem. The result is a component that was incorporated into the STAR analysis framework, called the GridCollector. We expect this work to continue, and bring to fruition the ability by physicists to select dynamically the subset of events they are interested in based on event properties, rather than the current practice of an inflexible pre-selection of event subsets into rigid smaller data sets for analysis.

6.1.4 Other SciDAC Projects

The Jefferson Lab team serves as the interface between PPDG and the Lattice QCD SciDAC project National Computational Infrastructure for Lattice Gauge Theory. The LQCD project envisages a distributed meta-facility of multiple dedicated large parallel machines, with Grid software providing the user transparency. This LQCD distributed facility will exploit developments coming from PPDG such as the SRM web service specification and other appropriate components as they appear, including both batch and data Grid components. JLab also provides the co-chair of the middleware working group of the International Lattice Data Grid, a planned Grid-of-Grids of which the LQCD SciDAC project is a key participant and contributor.

The High Performance Data Grid Toolkit project develops and implements GridFTP, RLS and MDS. Developments and implementation of GridFTP are contributed partially through this project as well as PPDG by the Globus team at ANL, PPDG will continue the collaboration with the Globus team at ISI, who develop and support MDS and RLS.

PPDG maintains contact with numerous other SciDAC projects as seen at http://www.ppdg.net/scidac_contacts.htm.

6.2 Physics Collaborations

6.2.1 Open Science Grid

As explained throughout the text PPDG plans to contribute to and collaborate with the Open Science Grid roadmap for deploying a persistent Grid infrastructure for the U.S., in particular initially utilizing existing and planned facilities at the DOE Laboratory Computing Centers. This strategy additionally provides a platform for the necessary planning, negotiation and international agreements necessary to bring Grids to fruition as a global infrastructure meeting the needs of the globally distributed scientific collaborations of today and tomorrow.

6.2.2 Trillium: PPDG, GriPhyN and iVDGL

PPDG will continue coordination and collaboration with the other Trillium Grid projects as described in section 2.2. As GriPhyN computer science deliverables such as Chimera reach a level of maturity for potential use in production Grids we expect further joint projects for common work on hardening and deployment.

6.2.3 LHC Computing Grid Project, EGEE

The CS-11 group in PPDG looking at issues related to interactive data analysis on the Grid has had continuing collaboration with the relevant activities centered at CERN for the LHC experiments. We were active in contributing to the HEPICAL[HEPCAL] documents defining a common application layer to the Grid for the LHC experiments as well as interaction with the ARDA process[ARDA] (architectural roadmap for distributed analysis).

Members of PPDG present information to and participate in meetings of the LCG and EGEE. PPDG collaborators are making contributions to the ARDA middleware activities, and there are joint activities between the VDT and the LCG/EGEE. Additional collaboration on joint projects between LCG and Trillium, LCG and U.S. ATLAS and U.S. CMS strengthen the ties.

7 Budget Summary and Justification

The budget requests for the collaborators are as follows:

The University of Florida is not requesting funding but will collaborate to ensure coordination with the NSF-funded GriPhyN and iVDGL projects led by Florida. The coherence of the “Trillium” – PPDG, GriPhyN and iVDGL is a fundamental component of the PPDG strategy.

\$K	July 1, 2004 June 30, 2005	July 1, 2005 June 30, 2006	2 Year Total
Argonne National Laboratory	487	487	974
Berkeley Laboratory	420	420	840
Brookhaven National Laboratory	366	366	732
California Institute of Technology	194	194	388
Fermi National Accelerator Laboratory	833	833	1666
San Diego Supercomputer Center	72	72	144
Stanford Linear Accelerator Center	368	368	735
Thomas Jefferson National Accelerator Facility	176	176	352
University of California, San Diego	82	82	164
University of Florida	0	0	0
University of Wisconsin	540	540	1080
Total DoE Laboratory Request	2648	2648	5296
Total University	888	888	1776
Total Request	3538	3538	7074

The funding request from each collaborating institution is given in the table.

The functional breakdown of this funding is:

- \$461k/year funding as direct support for PPDG management and the PPDG common project. These funds provide approximately 50% support for Ruth Pordes and Doug Olson and will also support a full-time Common Project Manager.
- \$1,168k/year of funding for the PPDG computer-science teams, representing about 6 FTEs of PPDG-funded effort
- \$1,908/year of funding for the experiments and the PPDG-OSG facilities teams, representing about 9 FTEs of PPDG-Funded effort
- CDF (Fermilab), PHENIX (BNL) and Globus (ISI) are not requesting funding in this proposal renewal. They will collaborate on the program of work because of their

commitment to leverage, contribute to and benefit from common solutions. The ALICE collaboration is also interested in collaborating on the PPDG program of work, subject to a decision on funding for US participation in the ALICE experiment itself.

The PPDG common project will draw resources from the PPDG-funded computer-science, experiment and facilities teams at a committed level of 25%.

The involvement of the over 90 scientists and engineers whose names appear on this proposal illustrates the PPDG strategy of using the PPDG-funded effort to leverage a much wider involvement from all the relevant communities.

8 Literature Cited

[AAA2003] Results of Particle Physics DataGrid Site Requirements on Authentication, Authorization, and Auditing Project. <http://www.slac.stanford.edu/econf/C0303241/proc/pres/424.PPT>

[AAA2003-1] Using CAS to Manage Role-Based VO Sub-Groups

[AAA2003-2] Site Grid Authorization Service (SAZ) at Fermilab
<http://www.slac.stanford.edu/econf/C0303241/proc/papers/TUBT007.PDF>

[AAA2003-3] Kerberized Certificate Authority (KCA) use at Fermilab
<http://www.slac.stanford.edu/econf/C0303241/proc/pres/422.PPT>

[ARCH] GriPhyN 2001-12, <http://www.griphyn.org/documents>

[ARDA] <http://project-arda.web.cern.ch/project-arda/>

[ARDA] <http://www.uscms.org/s&c/lcg/ARDA/>, 2003

[BaBar2003] Distributing BaBar Data using the Storage Resource Broker (SRB)
<http://www.slac.stanford.edu/econf/C0303241/proc/pres/392.PDF>

[CAIGEE] <http://pcbunn.cithec.caltech.edu/GAE/CAIGEE/default.htm>

[CIAR] C. Steenberg, E. Aslakson, J. Bunn, H. Newman, M. Thomas, F. van Lingen, "The Clarens Web Services Architecture", Proceedings of CHEP, La Jolla, California, March 2003

[CS11-1] Grid Services Requirements for Interactive Analysis,
http://www.ppdg.net/pa/ppdg-pa/idat/papers/analysis_use-cases-grid-reqs.pdf

[CS11-2] APIs for Interactive Data Analysis Using the Grid
http://www.ppdg.net/mtgs/20mar03-cs11/APIsDiagram_rev3.pdf

[CS11-3] APIs for Interactive Data Analysis Brief Descriptions
<http://www.ppdg.net/mtgs/20mar03-cs11/APIsBriefDefs.doc>

[CS11-4] Dataset Catalog Service Interface
<http://www.ppdg.net/archives/talks/2003/ppt00077.ppt>

[CS11-5] Dataset Catalog Service Reference Implementation
<http://grid.txcorp.com/des>

[DIAL] <http://www.usatlas.bnl.gov/~dladams/dial>

[GAE] web page, <http://ultralight.caltech.edu/gaeweb/>, 2003

[Grid2003] <http://www.ivdgl.org/grid2003>

[GRIDPP] <http://www.gridpp.ac.uk/>

[GSS'03] The Storage Resource Manager Interface Specification, Edited by *Junmin Gu, Alex Sim, Arie Shoshani*, available at <http://sdm.lbl.gov/srm/documents/joint.docs/SRM.spec.v2.1.final.doc>.

[GT] <http://www.globus.org/toolkit>

[GT-AS] <http://www-unix.globus.org/toolkit/3.2alpha/release-notes.html#CAS>

[GT-FT] <http://www-unix.globus.org/toolkit/3.2alpha/release-notes.html#GridFTP>

[GT-GSI] <http://www-unix.globus.org/toolkit/3.2alpha/release-notes.html#Security>

[GT-RM] <http://www-unix.globus.org/toolkit/3.2alpha/release-notes.html#GRAM>

[GT-RFT] <http://www-unix.globus.org/toolkit/3.2alpha/release-notes.html#RFT>

[GT-RLS] <http://www-unix.globus.org/toolkit/3.2alpha/release-notes.html#RLS>

[HEPCAL] <http://lcg.web.cern.ch/LCG/SC2/GAG/HEPCAL-II.doc>

[HOS] W. Hoschek, A Unified Peer-to-Peer Database Framework for XQueries over Dynamic Distributed Content and its Application for Scalable Service Discovery (Dissertation), Technical University of Vienna, 2003

[JAS] <http://jas.freehep.org>

[JASMINE] <http://cc.jlab.org/scicomp/JASMine/>

[LCG] <http://lcg.web.cern.ch/LCG/default.htm>

[MDS] <http://www.globus.org/mds>

[MONA] H.B. Newman, I.C. Legrand, P. Galvez, R. Voicu, C. Cirstioiu, "MonaALISA: A Distributed Monitoring Service Architecture", Proceedings of CHEP, La Jolla, California, March 2003.

[NETL] <http://www-didc.lbl.gov/netlogger>

[NEW] H. B. Newman, I. C. Legrand, "A Self-Organizing Neural Network for Job Scheduling in Distributed Systems", Presented at: ACAT 2000, FERMILAB, USA, October 16-20, 2000

[ORCA] "Object Oriented Reconstruction for CMS Analysis" CMS internal note submitted to the LHCC referees in January 1999

[OSG03] http://openscienceGrid.org/Documents/OSG-whitepaper-v2_3.pdf

[PNPA] <https://forge.Gridforum.org/projects/pnpa-rg/>

[POOL] <http://lcgapp.cern.ch/project/persist/>

[PPDG] <http://www.ppdg.net/>

[PPDG2003] PPDG presentation at SC2003. <http://www.ppdg.net/docs/presentations/ppdg-olson-sc03.pdf>

[PPDG-39] Document PPDG-39 at http://www.ppdg.net/docs/documents_and_information.htm.

[PPDGDOC] http://www.ppdg.net/docs/documents_and_information.htm

[PYGLO] <http://www-itg.lbl.gov/gtg/projects/pyGlobus>

[ROOT] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework, Proceedings AIHENP'96 Workshop", Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86

[SAM] <http://www-d0.fnal.gov/computing/grid/>

[SAM2003] Grid Job and Information Management for the FNAL Run II Experiments. <http://www.slac.stanford.edu/econf/C0303241/proc/papers/TUAT001.PDF>

[SAM2003-1] Grid Data Access for the Run II experiments at Fermilab <http://www.slac.stanford.edu/econf/C0303241/proc/papers/TUAT002.PDF>

[SAM2003-2] The SAMGrid Testbed: First D0 and CDF Experiences with Analysis on the Grid <http://www.slac.stanford.edu/econf/C0303241/proc/pres/342.PPT>

[SDMC] <http://sdmcenter.lbl.gov>

[SONN1] Self-Organizing Neural Network for Job Scheduling in Distributed Systems, Harvey B. Newman, Iosif C. Legrand, ACAT 2000, Chicago, http://clegrand.home.cern.ch/clegrand/MONARC/ACAT/AI406_paper.pdf, CMS NOTE 2001/009, January 8, 2001 http://clegrand.home.cern.ch/clegrand/SONN/note01_009.pdf

[SRB2002] Real-life Experiences with Data Grids: Case Studies in using the SRB, Arcot Rajasekar, Michael Wan, Reagan Moore, Arun Jagatheesan, and George Kremenek, The 6th International Conference on High Performance Computing (HPCAsia-2002) Bangalore, India, December 16-19, 2002. <http://www.npaci.edu/DICE/Pubs/hpcasia2002.doc>

[SRB2003-1] Storage Resource Broker - Managing Distributed Data in a Grid, Arcot Rajasekar, Michael Wan, Reagan Moore, Wayne Schroeder, George Kremenek, Arun Jagatheesan, Charles Cowart, Bing Zhu, Sheau-Yen Chen, Roman Olschanowsky, accepted Computer Society of India Journal, special issue on SAN, 2003. <http://www.npaci.edu/DICE/Pubs/CSI.doc>

[SRB2003-2] Operations for Access, Management, and Transport at Remote Sites, Reagan Moore, submitted to the Global Grid Forum, December, 2003. <http://www.npaci.edu/DICE/Pubs/Transport-operations.doc>

[SRB2004] Data Grid Federation, Rajasekar, A., M. Wan, R. Moore, W. Schroeder, submitted to PDPTA 2004 - Special Session on New Trends in Distributed Data Access, January, 2004. <http://www.npaci.edu/DICE/Pubs/Zone-SRB.doc>

[SRM] <http://sdm.lbl.gov/srm> and <http://sdm.lbl.gov/srm-wg>

[SSG'02] Storage Resource Managers: Middleware Components for Grid Storage, *Arie Shoshani, Alex Sim, Junmin Gu*, Nineteenth IEEE Symposium on Mass Storage Systems, 2002 (MSS '02).

[SSG'03] Storage Resource Managers: Essential Components for the Grid, *Arie Shoshani, Alexander Sim, and Junmin Gu*, in Grid Resource Management: State of the Art and Future Trends, Edited by Jarek Nabrzyski, Jennifer M. Schopf, Jan weglarz, Kluwer Academic Publishers, 2003.

[TIE03] B. Tierney and J. Hollingsworth, "Instrumentation and Monitoring", Chapter 20, The Grid Blueprint for a New Computing Infrastructure, 2nd Edition, Elsevier 2003

[VDT00] <http://www.griphyn.org/vdt>

[WSRF] <http://www.globus.org/wsrff/>

9 APPENDIX

The ALICE-USA collaboration is seeking funding from DOE NP for participation in the ALICE experiment at the CERN LHC. The program of work envisaged by ALICE-USA includes grid-related development and production activities and it is the intention to collaborate closely with PPDG once the approval from DOE NP is granted.

ALICE is the primary heavy ion experiment being mounted at the LHC at CERN. ALICE-USA [URL] is proposing to extend the ALICE experiment capabilities in high pt physics by adding an electromagnetic calorimeter (EMCal) to the experimental apparatus. Approval and initial funding for this effort is anticipated in FY2004. As for ATLAS and CMS, the analysis and simulation tasks will necessarily be distributed and will make use of a data Grid infrastructure. ALICE-USA has facilities at LBNL (NERSC), OSC and UH that will anchor the U.S. portion of ALICE computing resources. Along with approval for the overall ALICE-USA effort, it is anticipated that 1 FTE starting in FY2004 will work closely with PPDG on the data Grid aspects of ALICE-USA computing.

The AliEn software suite has been developed by the ALICE software team at CERN as a prototype for such a Grid infrastructure. This software is has been chosen by LCG to form the basis of a prototype infrastructure suite for the Architectural Roadmap for Distributed Analysis (ARDA) plan within the CERN LCG program. The ALICE-USA Collaboration has a direct interest in making AliEn and its progeny compatible with and able to take advantage of the VDT as well as interface and integrate with the US facilities. As the future development of AliEn moves it to a more general Grid infrastructure there will likely be some security and operations issues important to Grid operations at these facilities. Two specific examples of work are interface to Condor-G and end user (Grid ID) tracking of identity from initial job submission to execution process on a compute node. The general working plan is to start with the ARDA prototype, 1) identify and then develop interfaces to VDT components, and 2) analyze issues for site facilities and security integration with subsequent feedback into the AliEn development to address them. The situation with respect to the resource centers in the US will be of primary interest to the ALICE-USA Collaboration. In addition, the ALICE-USA Collaboration will endeavor to explore with our PPDG colleagues other areas of common interest that may present themselves with sufficient urgency to ALICE and LCG interests to warrant our more focused attention.