

The Particle Physics Data Grid Collaboratory Pilot

Year 1 Project Plan

V.02 7/4/01

Table of Contents

1 Introduction.....	1
1.1 PPDG in Year 1	1
1.2 End-to-End Applications	2
1.3 Coordination with other (Data) Grid Projects	2
1.4 Project Definitions	2
1.5 Architecture	3
1.6 Interfaces.....	3
2 Project Activities.....	3
2.1 Grid Data Management Pilot (GDMP) - CMS, Globus	3
2.1.1 GDMP Components:.....	4
2.1.2 End-To-End Architecture.....	4
2.2 Global Job Management - D0, Condor	4
2.2.1 Components:	4
2.3 ATLAS- Distributed Data Management	4
2.4 STAR Site-to-Site Data Distribution and Storage Management.....	5
2.4.1 Components	5
2.4.2 Architecture	6
2.5 JLAB-SRB Data Grid Web Services	6
2.6 CMS-University of Wisconsin Prototype Distributed Production System for Simulated Events.....	6
3 Project Efforts.....	6
4 End-to-End Applications and Integrated Systems.....	7
4.1 ATLAS Tier 1,2.....	7
4.2 BaBar – SLAC & IN2P3.....	7
4.3 CMS Tier 1,2	8
4.4 D0 - Fermilab & NIKHEF	8
4.5 JLAB – JLAB & SSU	8
4.6 STAR – BNL & LBL.....	8
5 Collaboration Activities.....	8
5.1 Project Documents	8
Editor	8
5.2 Meetings and Reporting.....	8
5.3 Participation in Standards	9

5.4 Coordination with Related Projects	9
5.4.1 Resource Management for Data intensive Grid Applications (SRM).....	9
5.4.2 Grid Physics Network (GriPhyN) http://www.griphyn.org	9
5.4.3 EU DataGrid (EDG) http://www.eu-datagrid.org/	10
5.4.4 MONARC http://monarc.web.cern.ch/MONARC/	10
5.4.5 Distributed Terascale Facility	10
5.5 Investigation of Commercial Technologies	10
5.6 Education and Outreach	10
6 Effort and Schedules.....	10
7 Project Activities from Proposal.....	11

1 Introduction

This document is a working document to present our plans for Year 1 of the three year Particle Physics Data Grid (PPDG) Collaboratory Pilot Project¹. PPDG's goals are to contribute to meeting the distributed computing needs of High Energy and Nuclear Physics Experiments by extending, integrating and deploying into production common grid software in experiment applications. The PPDG Collaboratory meets its goals through collaboration between Computer Science groups leading the development of Grid middleware in the US and six experiments in HENP whose data processing strategies stress the importance of distributed computing.

The PPDG project is organized as:

- “*Project Activities*” which are well defined projects between one or more experiments and one or more computer science groups;
- “*Project Efforts*” that typically involve one PPDG collaborator or institution;
- Deployment and testing of end-to-end “*Applications*” and integrated systems;
- “*Collaboration wide activities*”, most notably developing and promulgating Grid software standards and collaborating with other Grid projects within the field. These typically involving work with related projects e.g. Grid Forum, GriPhyN.

PPDG funded effort is spread across ten institutions and the resources available to the project are clearly a small fraction of the total effort needed to meet the experiment and computer science deliverables and goals. The project relies on leveraged activities across the larger groups - experiment data handling, core CS development projects and the other data grid projects. PPDG pays attention to the coordination and collaboration activities necessary to achieve coherent system developments in such a complex environment.

The PPDG executive team provides cross-project coordination. Other PPDG effort is funded through each Experiment or Computer Science groups. PPDG faces the challenge of bringing different experiments to consensus on the development and use of the same software modules. Our goal in the earlier stages of the project is to provide the groundwork so as to increase the development and use of common software in the second and third years.

Most PPDG information is maintained and available from the web pages <http://www.ppdg.net>. A project cvs repository is available for general use.

1.1 PPDG in Year 1

In Year 1 of PPDG ATLAS, BaBar, CMS, D0, JLAB and STAR will each participate in one or more Project Activities which will develop software and deploy Applications in the experiments data systems. PPDG effort includes bringing these Applications into production use – that is to allocate effort (in collaboration with “off-project effort”) to the integration, robustness, monitoring and deployment tasks that are an essential part of moving software to production use. Twelve months is not necessarily a natural milestone for most such projects, and within this Year 1 Project Plan specific timelines will be given as applicable.

In Year 1 of PPDG the Computer Science groups, Condor (University of Wisconsin), Globus (Argonne), Storage Resource Broker (SDSC), and Storage Resource Management (LBL), will each have participated in one or more Project Activities. They will have extended existing software, developed new grid middleware services, and supported the integration and deployment of the experiment applications.

In Year 1 of PPDG the focus is on Applications and *Common Software Components* in the following areas.

- Grid Data Replication - Grid Data Transfer and Replication Services
- Grid Job Submission and Resource Management - *Grid Job & Resource Management*
- Disk and Tape Resource Management - *Storage Resource Management*
- Grid Authentication and Authorization – *Grid Security Infrastructure*

The emphasis will be to develop and deploy these basic services in the experiments integrated systems. It is a goal of PPDG to extend and or foster increasing the sophistication of these components to provide more predictive, adaptive and next-generation services later in the lifetime of the project.

Each of the six experiments will deploy one or more such Applications as part of PPDG using a mix of common software and their experiment specific code. At minimum the following Common Software Components will be used in one or more experiment Applications.

- GridFTP - Grid Data Transfer
- Globus replica catalog - Grid Replication Services
- MCAT meta-data catalog - Grid Data Transfer and Replication Services
- SRB replication services - Grid Data Transfer and Replication Services
- Globus Replica Management Services - Grid Data Transfer and Replication Services
- GDMP extended Replica Management Services - Grid Data Transfer and Replication Services
- Condor ClassAds Library - Grid Job & Resource Management
- Condor-G , Condor - Grid Job & Resource Management
- DRM, HRM, SRM - Storage Resource Management
- GSI authentication - Grid Security Infrastructure

Clearly the above list covers only a small fraction of the total software deployed in any one of the experiments end-to-end applications. Other major components such as catalogs, data processing frameworks, experiments data persistency libraries etc, will appear in the descriptions below and will require effort. There is no attempt to have the descriptions describe the complete applications that will be deployed and it is clear that non-PPDG effort will be necessary to achieve our project goals.

In Year 1 especially, many experiment applications will not use all, or even some, of the common software components. Experiments have existing data handling systems which already fully or partially provide these services, and there are continuous related developments not associated with PPDG. One of the goals of PPDG as Year 1 progresses, and into the second and third years of the project, is to understand the experiment systems, foster software developments and coordinate with all the collaborating groups, such that the list of shared components grows.

1.2 End-to-End Applications

PPDG end-to-end applications provide the *Test Beds* for the PPDG project activity and effort deliverables. As said earlier, they involve many more people and projects than are collaborators on PPDG. They are typically systems that span institutions and continents. In PPDG these integrated systems are an integral part of the project. The work done to bring our delivered software to use in experiment production systems is an essential ingredient to make it reusable and maintainable.

1.3 Coordination with other (Data) Grid Projects

The experiments on PPDG are collaborators on other significant Grid projects both in the US and Europe. PPDG will actively work with these projects and with coordinating groups in the field to ensure consistency of deliverables and promote common solutions and architecture to enable the best possible systems to the experiments themselves.

It is part of PPDGs mission to work with other Grid projects in the High Energy and Nuclear Physics Application domain on the development and research needed to bring Grid services to be the accepted norm in their distributed applications.

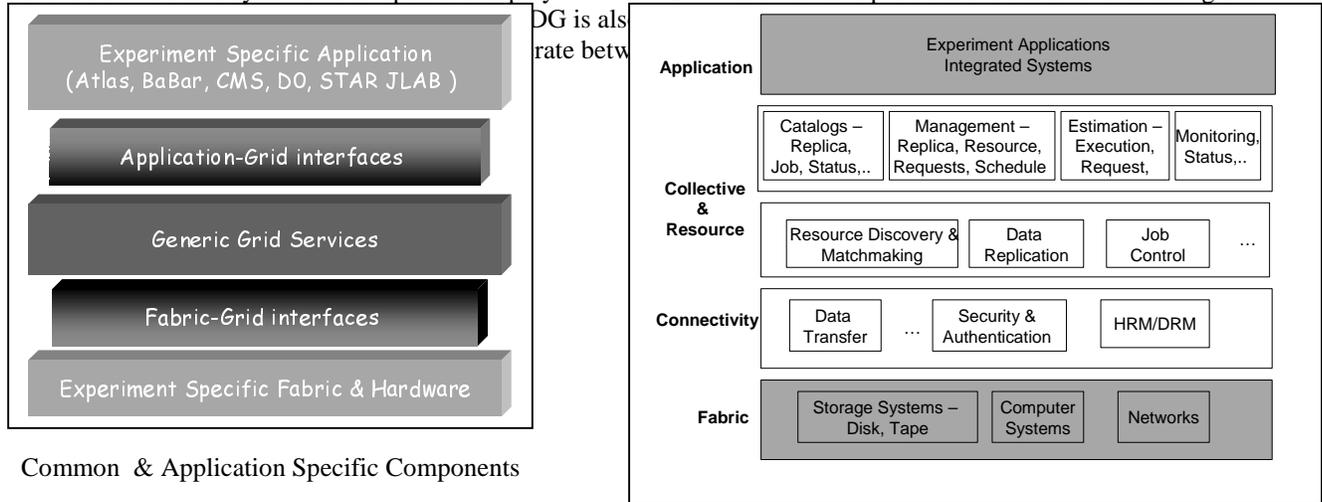
1.4 Project Definitions

Project Activities, Efforts and End-to-end applications will be assigned IDs that identify the CS area they are associated with, the name of the activity and/or the main experiment working on the project and/or a .

E.g. GDMP is P5-GDMP-CMS. Deliverables, schedules and effort are defined in a variation of the Atlas Xproject definitions²

1.5 Architecture

PPDG uses architecture definition and discussion as a means to understand experiment, application and systems requirements and to promote the modularity, interfacing, reusability, integrability, extensibility and maintainability of the software and systems developed and deployed. The DGRA³ architecture provides the basis for discussing the



Common & Application Specific Components

1.6 Interfaces

The following interfaces are defined for the First year architecture with references to how they are or will be defined:

- Storage System – including Disk Management – HRM –
- File Replication – GDMP –
- Data Transfer – GridFTP -

2 Project Activities

We will only briefly describe each Project Activity here. Each one will have a link from the PPDG web pages with a more complete description of the activity.

2.1 Grid Data Management Pilot (GDMP) - CMS, Globus

GDMP is project activity coordinated within PPDG by CMS and Globus. The Project itself is also an EU Data Grid project and the first milestones are defined to be in common between the 2 projects and to meet the DataGrid Month 9 Deliverables schedule, from “The DataGrid Month 9 Deliverable V1” Document.

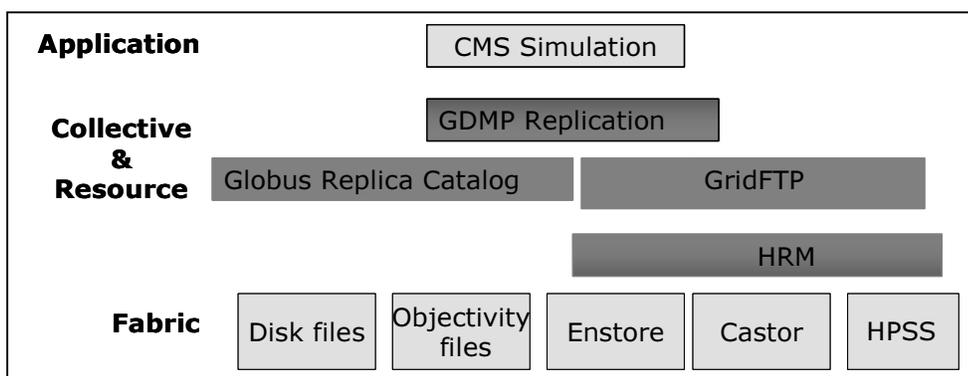
- Will provide: *Grid Data Replication Services*
- Will use: *Grid Security Infrastructure, Grid Data Transfer*
- Will test interfaces to: *Storage Resource Management*

2.1.1 GDMP Components:

- File Copier: gsi-wuftp. Gsi-ncftp
- Replica Catalog: Globus LDAP Catalog,
- User directed replica creation and deletion.
- Integration with mass storage systems : HRM to HPSS and Enstore

For PPDG the Milestones are taken from the CMS CPT/CAS WBS:

2.1.2 End-To-End Architecture



2.2 Global Job Management - D0, Condor

The D0 –University of Wisconsin Global Job Management will be integrated with the D0 SAM system such that there are well defined interfaces between the SAM services for data access and file transfer, the SAM job control, and the global job management developed under PPDG.

- Will provide: Grid Job Specification and Management
- Will use: Grid Job and Resource Management

2.2.1 Components:

Initial deliverables for this activity are probably:

- Job Definition Language
- Distributed Management Features – extensions to existing SAM and Condor languages
- Interface of new s/w to SAM
- Demonstration application

2.3 ATLAS- Distributed Data Management

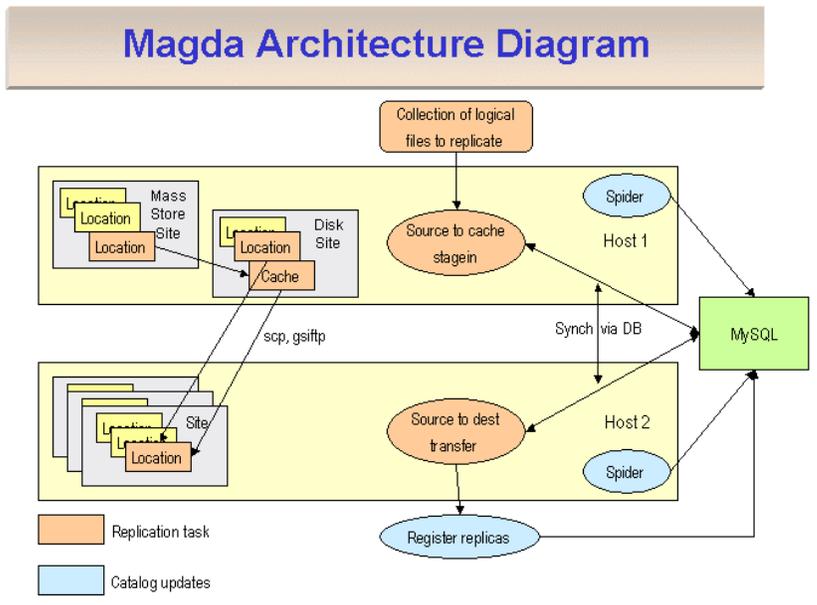
The principal ATLAS PPDG deliverable for year 1 is a production distributed data system deployed to users. MAGDA provides user interaction via web interface and command line and includes as principal components:

- **A File catalog** covering arbitrary range of file types
- **Data repositories** organized into *sites* and *locations* Computers with repository access: a *host* can access a set of *sites* Logical files can optionally be organized into *collections*
- **Replication, file access** operations organized into *tasks* To serve environments from production (DCs) to personal (laptops)

MAGDA will use

- *Grid Data Transfer and Replication Services*
- *; Grid Security Infrastructure*

-



2.4 STAR Site-to-Site Data Distribution and Storage Management

The STAR experiment sites at Brookhaven and Berkeley labs, in conjunction with the Storage Resource Management project at LBNL are developing and deploying a data distribution and management service for STAR computing operations at the RCF at BNL and at PDSF/NERSC at LBNL. The roles of the BNL and LBNL computing centers for STAR are analogous to the Tier0-Tier1 centers in the LHC computing model. The goal of this PPDG project activity is to automate and integrate the data transfer between BNL and LBNL (10's of TB per year) with the STAR data management system where the Computer Science focus is on the storage management issues of such an activity.

- Will provide: *Storage Resource Management*
- Will use: *Grid Data Transfer & Replication Services*

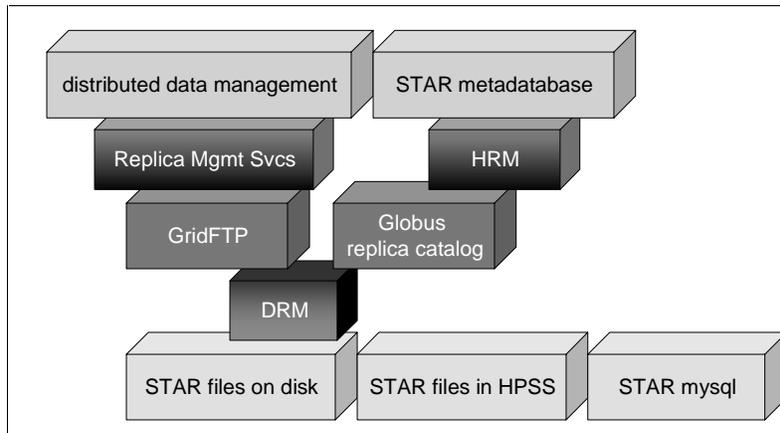
2.4.1 Components

In the model of PPDG project activities, there is a focus of pressing the development and integration of one domain of middleware components while making use of other stable "near production grade" middleware components from the Computer Science projects participating in PPDG, as they become available. In addition, there are application specific software components (STAR experiment) including in putting together the end-to-end application for this project. The list of components below are those expected to be available during the first year (June 2001 – May 2002). It is expected that standardized components from other activities may replace some of the application specific STAR components in the future.

- SRM – storage resource management components
 - HRM
 - DRM
- "near production grade middleware" components
 - GridFTP from GLOBUS
 - GLOBUS replica catalog
- application specific components from STAR
 - metadata catalog, including file replica information

- distributed data management, including replica management functions

2.4.2 Architecture



2.5 JLAB-SRB Data Grid Web Services

Develop XML based web services in support of a Data Grid Portal, including XML schema to describe data files and replicas, and transactions (possibly using SOAP) to manipulate a data grid via web services (HTML form or applet or application as web services client). Should show high correspondence with HRM IDL so that the back end could be implemented over HRM (actual file transfers using separate file transfer utilities such as gridFTP). These services will be extended for resource discovery and management and the development of XML schemas corresponding to ClassAds OR schema for wrapping ClassAds descriptors so that they can be carried by web services / XML / SOAP

- Will provide: *Grid Data Specification and Replication Services*
- Will use: *Grid Security Infrastructure, Grid Data Transfer*
- Will test interfaces to: *Storage Resource Management*

2.6 CMS-University of Wisconsin Prototype Distributed Production System for Simulated Events

3 Project Efforts

PPDG includes evaluation and development efforts that are localized to a single collaborator or site. Each experiment has a different set of immediate needs and schedule. The experiment and computer science groups provide leveraged manpower to work with their PPDG colleagues on experiment specific grid applications. PPDG sponsors these activities, participates in discussions on their design and implementation, and encourages reports of the work. PPDG will monitor the technical success and directions with a goal of proposing and fostering software that can be reused for other experiments, or that can be extended or modified for more general use.

Year 1 Project Efforts that are currently underway (this list will be updated during the year):

Collaborator	Title	Schedule	Description
Atlas	Tool for distributed data services - MAGDA	8/1/01	Transition to a Project Activity with Globus - integration of Globus replica catalog , GDMP and Globus replica management services.
BaBar	Intra-site data replication	10/1/01	Prototype a file replication system tailored to the BaBar objectivity dataset files. Initially this application will be deployed local to SLAC for performance and robustness tests. This work will follow on from tests with the SRB MCAT catalog and file replication service, and file request redirection techniques using HTTP redirection and the Globus LDAP replica catalog. Following this the effort will transition to a Project Activity to develop Inter-site replication between SLAC and IN2P3.
CMS/Caltech	Remote data analysis prototype		Remote data analysis using JAS and GSI authentication
D0	SAM Information Services and Test Harness	12/1/01	Provide common status and information services throughout the SAM system. Present and analyze information to understand performance and availability aspects of the system. Complete Test Harness application to allow simulation of system, stimulation of error conditions, and configuration of system parameters to the boundary conditions.
CMS/Caltech	Partitionable Execution Service for Distributed Processors - TQS		High Availability , Fault Tolerant Job Queuing service
CMS/Caltech	Optimized database query tags for Tier 2 data set		In collaboration with GriPhyN - Simulate the grid environment with Belief Desire Intention (BDI)-based software agents. Test different algorithms for optimising query planning, execution and long-term load balancing of the data grid
LBNL	Disk Resource Manager – DRM – IDL and prototype implementation	12/1/01	In collaboration with the Resource Management for Data intensive Grid Applications SciDAC project.

4 End-to-End Applications and Integrated Systems

In Year 1 of PPDG the following integrated systems will be deployed. The experiment team leads will coordinate the tests and monitoring activities and will keep the PPDG collaboration informed as to the progress, problems, and results. What is learned from these integration and deployment efforts will be used to improve the quality of the software and drive future extensions and developments.

4.1 ATLAS Tier 1,2

The US ATLAS grid testbed involves PPDG sites (ANL, BNL, LBNL) and several university sites (Boston U, LBNL, Indiana U, U of Michigan, U of Oklahoma, and U of Texas-Arlington). The experiment will perform tests using Globus and other tools to integrate a system that provides file replication, transport, and a security infrastructure. Data samples of order (1 TB) from testbeam data recorded at CERN will be used. Further information on the US ATLAS grid testbed can be found at <http://www.usatlas.bnl.gov/computing/grid/>.

4.2 BaBar – SLAC & IN2P3

BaBar is already transferring ~100Gbytes data /day from SLAC to IN2P3 using scripts that require manual intervention and monitoring. Before the end of 2001 BaBar will deploy production data replication services to fully automate the transfer of data to IN2P3, integrating a security infrastructure.

4.3 CMS Tier 1,2

Simulation production will be run on the PPDG sites (Fermilab, Caltech, UCSD and University of Wisconsin) as part of the Spring and Fall 2001 production activities.

4.4 D0 - Fermilab & NIKHEF

Support for Monte Carlo production at NIKHEF. Bbftp has been integrated with SAM and deployed on the D0 Central Analysis System. Sustained file transfer rates of above 20 Mbit/sec on SURFNET are measured by our collaborators from NIKHEF to Fermilab.

4.5 JLAB – JLAB & FSU

A multi-site test bed data grid will be deployed based upon JLab developed disk and tape management software and portal / web services. Portal software will conform to the XML schemas developed within the JLAB-SRB Data Grid Web Services activity described in section 2.5 and will attempt to track schemas coming from the Global Grid Forum. Both web (XML+XSL => HTML) and custom client (Java) user interfaces will be prototyped as part of this testbed. Portal software for batch interactions will also be prototyped as part of this testbed.

4.6 STAR – BNL & LBL

The Tier 1, 2 (BNL, LBL) centers in STAR will provide the testbed for a production system for data replication for the experiments. Of the order of 10Tbytes /year will be transferred.

5 Collaboration Activities

5.1 Project Documents

In Year 1 the following documents will be delivered to the collaboration. Each document has an editor who is responsible for gathering the information, publishing the revisions, coordinating the reviews and delivery of the final text:

Document	Schedule	Editor
Year 1 Project Plan (this document)	Review: 8/15/01; Final: 10/15/01	Ruth Pordes
Replication Requirements		
Replication Use Cases		Mike Wilde
Storage Management Experiment Use Cases		

5.2 Meetings and Reporting

This Year 1 Project Plan document supplements other PPDG planning and reporting activities:

- PPDG submits quarterly reports to the DOE on its progress.
- Bi-weekly collaboration phone meetings at which status of each activity is reported and discussed.
- Face-to-face collaboration meetings are opportunities for everyone on the project to discuss and input at a technical level.
- More focused Activity Meetings are an opportunity for detailed design and implementation plans to be made.
- Presentations and papers from PPDG collaborators.

5.3 Participation in Standards

- Grid Forum – Individuals on PPDG chair several working groups⁴ in the Grid Forum and others are active in the working groups themselves. PPDG is a supporter of GGF being the repository of Grid Standards. (<http://www.gridforum.org>). In the first year PPDG activities are most related to Applications, Grid Computing Environments, Remote Data Access, Scheduling and Security Working Groups.
- Networking - Individuals on PPDG lead and others are active on the HENP Network Users Group and Trans Atlantic Networking Committee. These committees will define the requirements and help acquire the network bandwidth needed by the HENP experiment community in the next few years. (<http://www.transpac.org>)
- IETF - The GridFTP protocol is being submitted as an Internet Working Draft to the IETF (<http://www-fp.mcs.anl.gov/dsl/GridFTP-Protocol-RFC-Draft.pdf>) (does it have an RFC number yet?)

5.4 Coordination with Related Projects

All the PPDG collaborating computer science and experiments groups and many of the individuals on our project are participating in other Grid development and deployment projects. We must ensure, for example, that the PPDG deliverables for ATLAS and CMS Tier 1-2 sites integrate well with the Cern Tier 0 center and to be well integrated with the Virtual Data Toolkit from the GriPhyN project. The PPDG deliverables for the BaBar data replication service must meet the collaboration needs of distributed data archiving between California, France & England. The PPDG deliverables for the D0 global job and resource management must integrate well with the existing SAM data access system. PPDG coordinates and collaborates with the following data grid projects:

5.4.1 Resource Management for Data intensive Grid Applications (SRM)

The SRM project – a collaboration between LBL and Fermilab – will perform research and development on the tape and disk management and interfaces needed by Grid applications. SRM plans to deliver the software developed to PPDG collaborators for integration and deployment into PPDG end-to-end applications.

5.4.2 Grid Physics Network (GriPhyN) <http://www.griphvn.org>

Taking advantage of the overlap in collaborators, PPDG is following the GriPhyN research and developments with interest, and plans to continue close coordination with the requirements and test bed activities of the experiments the two projects have in common. The GriPhyN Virtual Data Toolkit (VDT) V1.0 will be of interest to PPDG as it provides grid enabling services that include security, information, metadata, CPU scheduling, and data transport, and will be based on much of the same Grid middleware as PPDG experiments are using.

We anticipate that the GriPhyN research will pave the way for the more sophisticated Grid services that PPDG experiments need to develop and deploy in the second and third years of the project. From the GriPhyN all hands meeting the following working groups overlap areas of current work in PPDG, and we are collaborating with GriPhyN in these areas: Collection and Storage Management (hierarchical storage); Application interfaces to Data Grid; Replication requirements. The following working groups are addressing areas that PPDG is not yet working on: Performance, instrumentation, monitoring, simulation; Planning (fault tolerance, grid-wide event handling); Architecture including virtual data.

The following current areas of research that were reported at that meeting are of interest to PPDG:

- Dynamic Replication Strategies for a High Performance Data Grid – *Grid Data Transfer and Replication*
- Agent Grid Interfaces – *Grid Job and Resource Management*
- Query Optimisation – *Grid Job and Resource Management* – also a PPDG effort
- Grid Database techniques and issues
- Issues in Fault Tolerance and the Grid
- NeST: Network Storage Flexible Commodity Storage Appliances – *Storage Resource Management*
- Prophecy: Performance Analysis & Modeling for Distributed Applications - *Grid Job and Resource Management*
- High-performance bulk data transfers with TCP – *Grid Data Transfer and Replication*

- Location-Transparent Adaptive File Caching – *Storage Resource Management*

5.4.3 EU DataGrid (EDG) <http://www.eu-datagrid.org/>

PPDG Year 1 activities are related to the EDG Working Groups as follows:

- Grid Job and Resource Management – WP1 – Grid Work Scheduling
- Grid Data Transfer and Replication – WP2 – Grid Data Management
- Storage Resource Management – WP5 – Mass Storage Management
- End-to-end Applications – WP6 – Testbed and Demonstrators & WP8 – HEP Applications
- Future work in Grid Information Services in PPDG - WP3 – Grid Monitoring Services

PPDG is also participating in the US/European Datagrid Coordination group.

5.4.4 MONARC <http://monarc.web.cern.ch/MONARC/>

5.4.5 Distributed Terascale Facility

5.5 Investigation of Commercial Technologies

Members of PPDG have many contacts to different commercial distributed software and integrated system vendors. During the first year of the project we anticipate that people will report on such contacts as they pertain to PPDG, and that we will naturally discover whether there is a commercial company that can meet some of the needs of the experiments and/or is a potential partner in our developments.

5.6 Education and Outreach

In the PPDG Proposal it states “We will coordinate with existing ATLAS and CMS construction project outreach to include PPDG activities and work in their programs. Through the Quarknet program physicists mentor high school teachers. We will participate in this program to increase teacher and student awareness and understanding of distributed computing and Grid middleware.” In Year 1 CMS will collaborate with the Fermilab Trac Teachers Program (http://www-ed.fnal.gov/lasso/program_search/show_program_link.lasso?name=TRAC). We will allocate 2 teachers from a particular school – one Computer Science and one Physics to work together with the CMS PPDG team on deployment and testing of the distributed simulation production system. This will enable the teachers to learn about data grids in practical terms.

6 Effort and Schedules

During the first few months of the project we will be sorting out the issues of funding and who is working on what project. Several sites said in the proposal that they will need to hire in order to staff the PPDG projects. This section provides a placeholder for starting to list the schedule and effort that is being applied across the project in order at a later time to be able to understand the expended and required effort for our deliverables. Schedules will be determined by the collaborator team lead(s) (Experiment or CS Group Team Leads) associated the project activity or effort in consultation with the *Project Activity or Effort Leader* and the PPDG executive team. We will mark “*On-Project Effort*” – that is funded by PPDG, and “*Leveraged Effort*” – that is funded elsewhere. It is not expected that the named individuals are necessarily delivering all the effort listed – in development groups and collaborations there is generally sharing of the task assignments in order to maximize overall efficiency.

7 Project Activities from Proposal

Project Activity	Experiments	Yr1	Yr2	Yr3
CS-1 Job Description Language – definition of job processing requirements and policies, file placement & replication in distributed system.				
P1-1 Job Description Formal Language	D0, CMS	X		
P1-2 Deployment of Job and Production Computing Control	CMS	X		
P1-3 Deployment of Job and Production Computing Control	ATLAS, BaBar, STAR		X	
P1-4 Extensions to support object collections, event level access etc.	All			X
CS-2 Job Scheduling and Management - job processing, data placement, resources discover and optimization over the Grid				
P2-1 Pre-production work on distributed job management and job placement optimization techniques	BaBar, CMS, D0	X		
P2-2 Remote job submission and management of production computing activities	ATLAS, CMS, STAR, JLab		X	
P2-3 Production tests of network resource discovery and scheduling	BaBar		X	
P2-4 Distributed data management and enhanced resource discovery and optimization	ATLAS, BaBar			X
P2-5 Support for object collections and event level data access. Enhanced data re-clustering and re-streaming services	CMS, D0			X
CS-3 Monitoring and Status Reporting				
P3-1 Monitoring and status reporting for initial production deployment	ATLAS	X		
P3-2 Monitoring and status reporting – including resource availability, quotas, priorities, cost estimation etc	CMS, D0, JLab	X	X	
P3-3 Fully integrated monitoring and availability of information to job control and management.	All		X	X
CS-4 Storage resource management				
P4-1 HRM extensions and integration for local storage system.	ATLAS, JLab, STAR	X		
P4-2 HRM integration with HPSS, Enstore, Castor using GDMP	CMS	X		
P4-2 Storage resource discovery and scheduling	BaBar, CMS		X	
P4-3 Enhanced resource discovery and scheduling	All			X
CS-5 Reliable replica management services				
P5-1 Deploy Globus Replica Catalog services in production	BaBar,	X		

P5-2 Distributed file and replica catalogs between a few sites	ATLAS, CMS, STAR, JLab	X		
P5-3 Enhanced replication services including cache management	ATLAS, CMS		X	
CS-6 File transfer services				
P6-1 Reliable file transfer	ATLAS , BaBar, CMS, STAR, JLab	X		
P6-2 Enhanced data transfer and replication services	ATLAS, BaBar, CMS, STAR, JLab		X	
CS-7 Collect and document current experiment practices and potential generalizations	All	X	X	X

¹ The PPDG proposal is available from http://www.ppdg.net/docs/scidac01_ppdg_public.pdf

² Xproject , courtesy of Torre Wenaus <http://atlas.bnl.gov/atlas/Planning/XProject/README>

³ A Data Grid Reference Architecture, Ian Foster, Carl Kesselman

⁴ for a list of the Grid Forum Working groups see <http://www.gridforum.org/groups/WG.html>