The Particle Physics Data Grid Collaboratory Pilot
End-to-end Grid Systems for High Energy and Nuclear Physics Experiments

PPDG Steering Committee
www.ppdg.net, February 4th 2004

Summary
The PPDG mission continues to be: to enhance and enable High Energy and Nuclear Physics (HENP) Experiment scientific productivity by applying recent advances in the Computer Science of distributed, data-intensive computing (Grids) to current and developing needs; to advance the Computer Science of distributed, data-intensive computing through exploitation of current and developing HENP needs for end-to-end production Grid services; and to use the National and International scale of HENP collaborations to drive Grid middleware development and production service deployment to a National and International scale.

1. Introduction

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.

2. Achievements

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*.

3. Benefits to the Experiments

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 hundred 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 ESnet/DOEGrids 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.

4. Benefit to Computer Science Groups
PPDG integrates computer science groups into multi-disciplinary teams and also facilitates collaboration across CS groups towards common technical deliverables. Through PPDG new features have been added to Condor, Globus, SRB and SRM to successfully meet the achievements mentioned above. Additional benefits to the common software are increases in:

- Robustness - through deployment of Grid middleware in "production" environments, which expose limitations (functional and/or scale) and bugs in software tools.
- Research directions - requirements (physics use cases) provided by the experiments and limitations exposed by the deployment teams triggered new IT research directions and activities.
- Usability - where the combined integration teams and hands on approach have provided immediate feedback to the developments teams.

Many of these improvements have been incorporated in the Virtual Data Toolkit (VDT) that serves the wider scientific community.

5. Partnerships
PPDG continues to work closely with the DOE Science Grid Collaboratory Pilot (http://www.doesciencegrid.org/) and ESNET in the support and operation of the Certificate Authority and Security infrastructure. Close collaboration with the Storage Resource Management project and SDM ISIC continues (http://sdm.lbl.gov/srm/). The SRM specification has been adopted by computing centers that support high energy and nuclear physics users as well as by CERN as part of the LHC Computing Grid Project. PPDG contributes to interoperability tests across multiple implementations of SRM in the U.S. and Europe.
Coordination with GriPhyN and iVDGL is maintained at all levels. PPDG experiments have adopted the Virtual Data Toolkit (VDT) packaging and distribution of the common Grid middleware. (http://www.griphyn.org/vdt/), based on the underlying National Middleware Initiative (NMI) packaging and distribution.

6. Grid2003 – a shared science grid
Grid2003 (http://www.ivdgl.org/grid2003). is a collaboration between PPDG, iVDGL, GriPhyN and the U.S. ATLAS and U.S. CMS computing projects that deployed a multiple organization shared grid infrastructure and environment (Grid3) for physics and computer science applications. Grid2003 consists of over 2000 processors spread over 27 sites running applications from 6 different experiments and computer science groups. The Grid3 environment has been adopted for the U.S. ATLAS and U.S. CMS production simulation Data Challenges in 2004. U.S. CMS is currently achieving a factor of two increase in throughput by taking advantage of the Grid3 resources.

7. International Collaboration
PPDG continues to take a major role in the international coordination of Grid projects relevant to HENP. The project has strong ties with European partners involved in the CERN LHC program.
Members of PPDG are contributing to the work of the Global Grid Forum by participating in Working Groups and in leadership of the Particle and Nuclear Physics Application Research Group.

The progress of PPDG is documented in quarterly reports and over 40 other documents and papers
http://www.ppdg.net/docs/documents_and_information.html

For further information contact the PPDG Executive Team at email: ppdg-exec@ppdg.net