Particle Physics Data Grid Collaborative Science

Dane Skow, Fermi National Accelerator Laboratory, dane@fnal.gov

This White Paper attempts to collect the results and vision of PPDG collaborators looking toward the future for the next 10 years.

Science Area Summary

The Particle Physics Data Grid (PPDG) brings together High Energy and Nuclear Physics experiments from currently operating, world-class facilities and the next generation of experiments facilities. Science performed by these experiments is done in a competitive global environment where speed to understanding is important for establishing priority and leadership. Detectors are large and complex objects and experiments often run for 10 years or more. The data accumulated is of order of Petabytes per year for modern large collaborations.

Collaborations are intensely international and have been for decades. Collaboration size has been growing steadily for the past several decades and now routinely has ~1000 scientists per experiment. The Large Hadron Collider (LHC) collaborations are expected to include 2500 scientists each. These collaborators will mostly be located at universities and not co-located with the apparatus.

There has been a move in the usage of the analysis systems from an electronic presence focus to a computing/communications fabric idea where resources can be integrated wherever they are located. The presumption is still strong that students in the field will spend a significant amount of time with the apparatus and be co-located. However that is not longer the assumption for the full tenure of students and current vision is of active participation without requiring significant periods of co-location. There have been experimental efforts at creating remote control rooms and this continues to be a possibility of interest, both to capture the expertise of remote apparatus experts (often the builders are remote) and to "follow the sun" and provide convenient shift opportunities in various time zones.

The field has largely completed a migration from FORTRAN as the lingua franca to a polyglot of primarily C++ but large pools of Java, Python. The efforts and costs in this transition were large with yet to be fully accrued benefits, but there is a fresh understanding of the requirements, capabilities and costs of large scale computing paradigm change. Platforms for computing have almost entirely moved from commercial SMP systems to workstations. Operating systems of choice are Linux, Windows, Mac OS and (for server applications) Solaris. Almost all other uses are legacy systems.

Experimental analyses are frequently limited by available effort and/or ability to digest large datasets. This has the physics result of delaying or preventing high statistics, precision measurements. At the colliders today, this would include some Bottom and Charm physics; at the LHC, it would be Top and Bottom physics. Improvements in analysis efficiency, ability to do more speculative or alternate analyses, and reduced turnaround times in investigating alternatives would enable faster and/or more physics results. Such measurements have in the past spotted effects of physics beyond the direct production threshold, thus increasing the discovery reach of the accelerators. As the cost and timescale of collider construction grows, these measurements become increasingly valuable.

Theoretical computation is focused on computationally intensive methods of Lattice calculations (similar to many other finite element techniques) and Monte Carlo simulation of perturbative calculations.

Comparisons with data are limited by computational capabilities in mathematic solving tools and by understanding of how to map experimental data reduction/extraction techniques to calculable quantities.

Accelerator and detector builders need more simulation capability to rapidly return results of simulations to tune device operation and to guide design decisions for next generation apparatus. Current scale of simulations is of order 100s of CPU days per trial. Would like to reach regular production of >1 trial/shift.

Current network connectivity expected to large-scale research institution is .5-1 Gbps with first forays into dedicated optical network connections between major research institutions.

Challenges, Priorities and Requirements

PPDG was started 4 years ago and has been funded at the level of ~$3M/yr as an incremental pilot to engage computer scientists with high energy physics collaborations to integrate grid technologies and concepts into experiment collaboration specific computation software and infrastructure. The base operations investment in each of the 6 experiments is of order $4M/yr. In a collaboration with NSF projects Griphyn and iVDGL, PPDG collaborated with the experimentsto join resources into a shared ensemble [Grid3] resulting in 30-100% computational throughput improvements over dedicated resources alone for simulation production jobs.. This is considered to have reached the breakeven point and current efforts are focused on improving the operational characteristics of the shared resources. It is expected that follow-on deployments in new application areas would likely repeat much to the first year's community building experience, but should expect to reach breakeven within 18months from project initiation given similar level of effort. To this end the Grid3 community is evolving the existing infrastructure, adding capabilities and building a sustained, production system. The successful operation of Grid3 for more than a year has shown that the benefits obtainable from a coherent common infrastructure can be achieved with a supportable operational cost. The contributors to Grid3 plan to present their existing and emerging resources to the Open Science Grid.

The majority of resources available today are devoted to production uses for extant experiments with immediate data demands and competitive environments. These capabilities must continue to function while the next generation of facilities is constructed. The challenge will be to utilize common tools and techniques to build to a common infrastructure that will enable resource sharing and facilitate new initiatives (e.g. accelerator simulation, data searches, etc.) both within the field and without

Conceptually, the experimenters are looking for middleware to be as useful and ubiquitous as ROOT is in the current analyses. Determining what roles the Grid middleware should play and at what level of the software stack is an ongoing area of investigation. PPDG science is strongly committed to the open source software model, both due to the needs to understand and modify common code in detail and the need for widespread and "free" dissemination of software.

Near Term Expectations (2004-2007)

Currently leading experiments are in full operation and striving to take full advantage of their leadership position to scoop the LHC on any and all observable signals that will stretch the capabilities of the detectors and analysis systems to their limits (and, in calculated gambles, beyond).

LHC construction projects will complete and data analysis systems will need to be ready for operation. Current generation experiments will begin to loose manpower to new experiments, but will sit on the crown jewels of the field for experimental data. Improvements in efficiency and accessibility will be essential to enable full extraction of information from the gathered data.

The priority with in PPDG is to make available large catalogs of data and metadata to a global ensemble of computing resources. The immediate priority is the computational clusters >100 CPUs, but eventually this will extend down to the researchers desktop.

We expect middleware to complete the migration to Services Oriented Architecture and look forward to development of common core standards on which several interoperable software kits are built.

We expect network developments to continue to rapidly improve bandwidth availability in point-to-point systems with rather slower improvements in wide area general network connectivity. It is likely that a mesh of dedicated links between major centers and local connectivity to a regional center is a viable model for our science and the path that current plans are following. Efforts on optical networking will need to progress and by the end of this term we will need to be reaching expected network connectivity levels of order 100Gbps for major institutions.

We expect that Linux will continue to be a large presence operating system. We continue to have concerns about the high level of effort required to properly maintain and configure operating systems. We expect that improvements will come slowly from the base OS providers.

We need to increase efforts on modularity of experiment specific code allowing for ready reuse of common packages and services. This will be a natural consequence of a move to a shared infrastructure. It is not clear what price in performance experiments are willing to pay over dedicated solutions, but costs for customized solutions are being more clearly accounted and watched.

We will continue our commitment to a sustained common shared infrastructure with the building and support for the Open Science Grid [OSG]. As contributors to the global infrastructure needed by our experiments, we will continue to work on interoperability with the LCG and EGEE in Europe, and other national and regional grid infrastructures in the US and elsewhere

It is too hard to operate these grids securely and at high performance. The systems need to be simplified and made more robust. This will be tough to do until stability comes. We see no light at the end of that tunnel yet.

Planning Horizon (2008-2010)

Networking connectivity expectations will continue to increase with regional loops running at 1000Gbps to deal with hub and spoke architecture and support of regional analysis efforts.

The LHC has turned on and data analysis is underway with keen competition toward first result analysis of result from opening the energy frontier. Previous generation facilities continue to take data and exploit their advantages of known detectors and accelerators to pursue specialty physics topics. Most of the national only facilities have stopped operation as other than training and test beam facilities. The remaining ones are regional and/or international.

The Next Linear Collider (NLC) decision has been made and construction activities have begun. Global network of accelerator designers and operators are developing a community to share resources, tools and design ideas.

A Vision for the Future of Particle and Nuclear Physics (2010+)

In the future of 2010, most of the world-class facilities will be global enterprises (the LHC, the Linear Collider, …) and the computing environment will be globally interconnected. It is likely that computing infrastructures will be federations of national or regional grids shared with other research communities. No single grid implementation is likely to be universal; standard interfaces will support federation.

The bulk of analysis time is in understanding systematic errors associated with a particular method of extracting an experimental result. This requires great detailed understanding of precisely how a result was obtained (the provenance of analysis) and what conditions were used to estimate the systematic uncertainties (which set is appropriate for the given analysis). Improvements in matching these two and automatically cataloging relevant parameters in the provenance would go a long way toward improving analysis efficiency.

Phenomenology has been an important part of our science with an industry to do composite analysis of data from different sources. As the number of apparatus decreases, it is likely that this sort of derivative analysis will increase and there will be increasing pressures to standardize data outputs and make them more widely available. There are probably lessons from the biologists and astronomers relevant here.

Most analysis goals are factorable and the old debate of whether to move the data to the executable (simpler but bandwidth intensive) or the executable to the data (more complex scheduling and job decomposition demands) continues. Further effort in this area may be productive.

Many experiments are currently limited by I/O rates. Typically output rates to storage systems and input rates into analysis software. Research on how to optimize these would enable a broader range of science to be done – as well as likely increase the breadth of participation of distributed collaborators.

References

[PPDG] Particle Physics Data Grid Renewal Proposal http://www.ppdg.net/

[OSG] Open Science Grid http://www.opensciencegrid.org/

[ESNET] DOE Science Networking: Roadmap to 2008 http://www.es.net/hypertext/welcome/pr/Roadmap/

[Grid3] http://www.ivdgl.org/grid2003