

Particle Physics Data Grid: From Fabric to Physics

Selected Science Benefits

28 February 2004



1 Introduction

This is a compilation of “science benefits” resulting from PPDG, and sent to DOE.

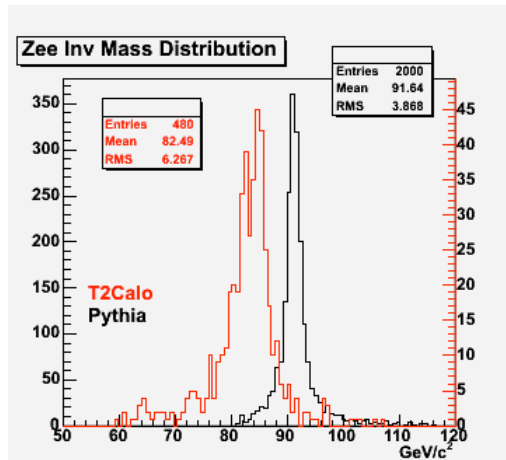
2 ATLAS Data Challenge 2: Grid-based production and analysis

Collaboratory: Particle Physics Data Grid (www.ppdg.net)

Contact: ppdg-exec@ppdg.net

ATLAS experiment: <http://atlas.ch/>

The ATLAS experiment is being prepared to take data at the Large Hadron Collider (LHC) at CERN, near Geneva, Switzerland beginning in 2007. A central goal of the LHC program is the discovery and characterization of the Higgs boson, an as yet unseen elementary particle predicted to exist and necessary to the theory describing why particles have mass. The ATLAS collaboration consists of 2000+ scientists around the world with hundreds of scientists participating in the U.S. (USATLAS).



The ATLAS Data Challenge DC2 is a major milestone in the preparation of the ATLAS world wide software environment that includes global distribution of data and access to computing resources in a distributed system. The goals of DC2 include the operation of the full software environment for ATLAS with multiple, interoperating grids. The results from DC2 will provide input for the ATLAS computing model as well as providing input to the overall resource estimates of needs. Physicist Ed May at Argonne National Laboratory says “DC2 produced simulation data and reconstruction has been and continues to be used for the study of the Level 2 trigger efficiency for high PT electrons, using data samples from $Z \rightarrow ee$.” The Level 2 trigger is part of the real-time filter that selects one interaction out of 10 million for recording by the experiment. Understanding the effectiveness of this filter is essential understanding physics meaning of the recorded data. The present grid-based simulation and production is providing the capability to refine this understanding. The figure at the left shows the simulated energy deposited in the level 2 trigger calorimeter compared with the mass of the Z-meson that decayed into two electrons.

The persistent shared grid infrastructure in the U.S., Grid3, is providing about 30% of the global resources used for DC2. The grid-enabling of the ATLAS applications and deployment of the Grid3 environment is a result of USATLAS participation in the Trillium consortium, a cooperative effort of the NSF funded iVDGL and GriPhyN projects and the DOE funded Particle Physics Data Grid project.

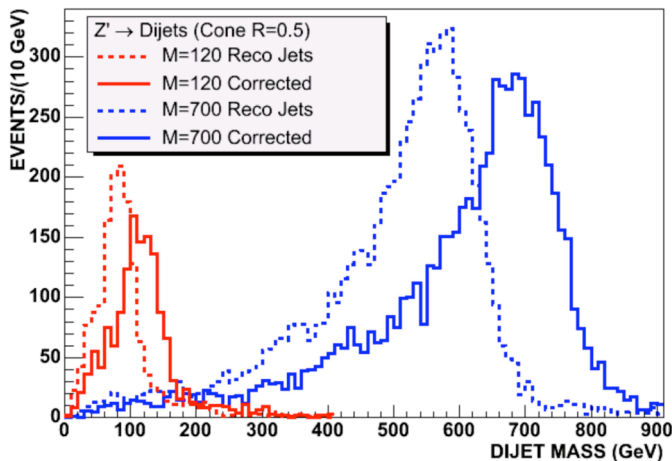
3 CMS Simulation Production on Grid3

Fermilab Computing Division. Collaboratory: Particle Physics Data Grid (www.ppdg.net)

US CMS project: www.uscms.org

The US CMS software and computing team is bringing the CMS data from CERN to the US to allow the 400 US Physicists to analyse the data at their home institutions. CMS is a companion experiment to ATLAS with similar needs in terms of data distribution and computing requirements. Both experiments make use of the Grid3 common grid infrastructure for event simulation. In a "first," the two experiments shared university and laboratory resources, even changing the priority of their batch queues to accommodate the other experiments jobs. The CMS Data Challenge DC04 was one of the milestones of the experiment scoped to ensure the experiment is ready with its global data distribution and analysis systems for the start of data taking at the Large Hadron Collider at CERN in 2007.

The performance metrics for DC04 were to provide a baseline to give the experiment input to the Physics and Computing Technical Design Reports in the next two years. These design reports will form the baseline to which the production data processing and analysis systems will be built and must perform.



Fermilab scientist Lothar Bauerdick says "Grid3 was a real success for us! For the US CMS DC04 data challenge we got a 50% increase in throughput using opportunistically available grid resources - and now we are able to continue to generate several million events a month when CMS gives us the requests. We will depend on a production grid infrastructure for our data processing and analysis in the US - we are working closely with the Condor and Globus groups to make

sure their middleware meets our needs and we are using the Open Science Grid Consortium as the means to work with many groups on a common grid infrastructure for us all"

4 Distributed Computing helps resolve Top quark mass in CDF

Collaboratory: Particle Physics Data Grid (www.ppdg.net)

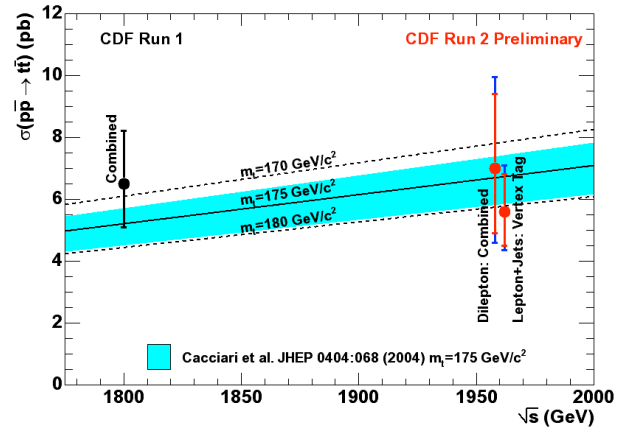
Contact: ppdg-exec@ppdg.net

CDF Experiment: www-cdf.fnal.gov

The CDF experiment at Fermilab is undergoing an aggressive Trigger and DAQ upgrade which will increase the maximum event rate out of the detector from 80Hz to 360Hz by 2006. The rationale for this increase is to be able to maximize physics output at any luminosity by dynamically filling the bandwidth with low threshold triggers. The most important physics measurement that benefits from the dynamically filled bandwidth is the measurement of Bs mixing. However, a broad range of other physics topics will also benefit. The increased data volume translates into a doubling of computing needs which had previously not been budgeted for. To satisfy those needs, CDF has embarked on an ambitious global computing program with the goal of having 50% of all of CDF's computing outside of FNAL by Summer 2005.

Professor Marjorie Shapiro from UC Berkeley says “Large samples of simulated data are essentially for understanding the backgrounds and systematic uncertainties for processes such as top quark production. The increased computing available through the grid allows detailed studies that will eventually reduce the uncertainties in the key physics measurements such as the top quark mass.” These studies rely entirely on the distributed computing resources outside of Fermilab and are an example of a new physics capability enabled with help from PPDG in the CDF transition to globally distributed computing.

Future directions are focused in the US on SAMGrid within the context of the Open Science Grid. However, as 50% of future CDF hardware funding presently appears to be located outside the US, interoperability between and federation of grids will be a primary concern for CDF in the future.



5 Grids: now an essential and standard part of getting physics from DØ

Collaboratory: Particle Physics Data Grid (www.ppdg.net) Contact: ppdg-exec@ppdg.net

DØ experiment: www-d0.fnal.gov

The Fermilab RunII program is the world-leading elementary particle physics program colliding protons and anti-protons, until the LHC program takes over at the end of this decade. The physics program for RunII is testing the limits of the “Standard Model” of particle physics (known to be incomplete), investigating the nature of the electroweak interaction, and CP symmetry violation. The RunII experiments (DØ, pronounced dee-zero, and CDF) each have about 700 physicists participating from around the world. One of the significant challenges of the RunII program compared to the previous run is a greatly increased luminosity and data rate from the experiments, leading to a much greater need for computing resources to process and analyze the data. The strategy adopted by DØ to meet this challenge is to integrate advances in grid computing technology with a distributed computing model so that data analysis and simulations activities can be shared at many computing facilities around the world.

Fermilab physicist Gustaaf Brooijmans says “Using grids is now an essential and standard part of the process to getting physics out of DØ data from the Tevatron. We could no longer get physics results out without using grid middleware to distribute data and schedule jobs. There is of course a lot left to be done to have grids provide the robustness and universality of our ESnet and other networks - but the potential is proven. The Fermilab Result of the Week showing B(c) mesons mass and lifetime measurement at DØ, is just one of many topical results. DØ is one of the founding member experiments in the PPDG SciDAC Collaboratory project.

http://www.fnal.gov/pub/today/archive_2004/today04-12-02.html”.

