The ATLAS Computing Model: Status, Plans and Future Possibilities

Shawn McKee *

University of Michigan, Ann Arbor, Michigan, USA

Abstract

The ATLAS Collaboration[1] has been preparing for Large Hadron Collider (LHC) running for more than 20 years. By summer of 2007 we expect the first colliding beams of protons and the start of a data deluge involving many Petabytes of data per year. The collaboration is eagerly anticipating the many possible scientific discoveries which may lie hidden in this massive amount of information. I will discuss how the ATLAS collaboration is planning to manage this large amount of data and provide physicists the infrastructure required to:
– Calibrate and align detector subsystems to produce well understood data
– Realistically simulate the ATLAS detector and underlying physics of interest
– Provide access to ATLAS data globally
– Define, manage, search and analyze data-sets of interest
There are also numerous research activities in networking and grids which could eventually have a significant impact on the ability of ATLAS (and LHC) physicists to quickly access distributed data sets and effectively manage and utilize their computational, storage and network resources. I will cover some of the research in this area and indicate how it might benefit ATLAS in augmenting and extending its infrastructure.

Key words: High-energy Physics, LHC, Computing, Applied Research, Network Research

1. Physics at the Large Hadron Collider

The ATLAS Collaboration is one of five particle detector experiments (ATLAS, ALICE, CMS, TOTEM and LHCb) being constructed at the Large Hadron Collider (LHC)[2]. The ATLAS experiment[3], scheduled for completion in 2007, is designed to measure phenomena involving massive particles, previously unmeasured, that might explain new theories beyond the Standard Model.

One of the most important goals of ATLAS is to investigate the final missing piece of the Standard Model: the Higgs mechanism[4]. The Higgs boson is predicted to exist in an energy range accessible to ATLAS and, through its interactions, would give mass to the other bosons and quarks.

2. ATLAS Computing Model

To explore the exciting physics possible at the LHC the ATLAS collaboration has been working to provide a computing model[5] that fully supports physicists in accessing and analyzing data. This data could be either simulated or real data from the ATLAS detector and amounts to many petabytes (10^{15} bytes) per year. The goal of the computing model is a production and analysis system which provides seamless access to all ATLAS data and resources.

Even though the computing model has seen significant effort and planning there are still a number of areas which must be addressed to deliver an effi-

* Corresponding author.
Email address: smckee@umich.edu (Shawn McKee).

cient, effective infrastructure for ATLAS physicists to utilize. One of the primary areas requiring additional focus is in physics data access and data distribution.

ATLAS has a hierarchical model for data production and distribution. ATLAS has a set of tiered computing centers, starting from the "Tier-0" at CERN and moving out to national scale "Tier-1's", regional "Tier-2's", institutional "Tier-3's" and ending at "Tier-4's" (desktop workstations). Data will be processed in stages from RAW to ESD (Event Summary Data) to AOD (Analysis Object Data) to TAG. To understand this process we need to examine the required flow of data from the detectors outword.

2.1. *ATLAS Data Flow*

The ATLAS detector produces an enormous amount of data while its in operation, equivalent to over 1 petabyte per second if all event data could be recorded. This is far beyond the capabilities of current or even near-term-future technology.

To create a manageable data flow, the ATLAS collaboration has implemented a fast layered trigger system designed to select out interesting events from the huge "background" of standard (well understood) events. The trigger system uses simple real-time information to identify the most interesting events out of the 40 million beam crossings that occur every second in the center of the detector.

There are three trigger levels, the first based in electronics on the detector and the other two primarily run on a large computer cluster near the detector. After the first-level trigger, about 100,000 events per second have been selected. After the third-level trigger, a few hundred events remain to be stored for further analysis. Since each event is about 2 megabytes (MB) in size this works out to a data rate of 200-400 MB/sec; equivalent to a few petabytes each year.

3. Offline Softare

Offline event reconstruction will be performed on final selected events, turning the pattern of electronic signals from the detector subsystems into physics objects, such as jets, photons, and leptons.

Grid computing will be extensively used for event reconstruction, allowing the parallel use of university and laboratory computer networks throughout the world for the CPU-intensive task of reducing large quantities of raw data into a form suitable for physics analysis. Grid middleware will be required to provide the ability to manage and access all the collaboration's distributed resources. The software for these tasks has been under development for many years, and will continue to be refined once the experiment is running.

Individuals and groups within the collaboration will write their own code to perform further analysis of these objects, searching in the pattern of detected particles for particular physical models or hypothetical particles. These studies are already being developed and tested on detailed simulations of particles and their interactions with the detector. Such simulations give physicists a good sense of which new particles can be detected and how long it will take to confirm them with sufficient statistical certainty.

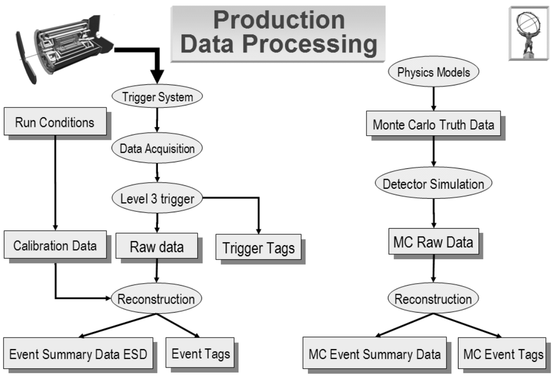

Fig. 1. The ATLAS real and simulated data production flow

The data flow for both real and simulated data is shown in figure 1 and demonstrates the high-degree of similarity between the real data path and the simulated one.

3.1. *Distributed Data Management*

One of the most critical tasks facing the collaboration is the management and distribution of data. With approximately 1800 PhD physicists distributed world-wide this is a non-trivial task. The ability of ATLAS physicists to expeditiously analyze ATLAS data will be dependent upon the capabilities of the underlying infrastructure (both hardware and software) to move and process data.

There has been a significant amount of effort put into distributed data management which includes data movement, providence, registration and cataloging. ATLAS uses DQ2[6] (Don Quijote 2) for its distributed data management system and the gLite FTS[7] service for data movement. The emphasis to date has been upon creating a working functional system, but not neccessarily one which emphasizes performance or works with the network components to insure overall quality of service.

4. Related Research Efforts

The concept of grid-computing would never have developed without robust, ubiquitous networks to enable consistent access to distributed resources. Networks play the role of the virtual data bus for the virtual computer the grid can assemble.

Because of the central importance of the network to grid computing, and therefore ATLAS, we need to make sure we can rely upon and effectively use networks to enable our computing model. In addition to those working to deliver the core software for ATLAS there are a number of physicists working along with network engineers and computer scientists to develop high performance capabilities which can significantly improve the overall effectiveness of the ATLAS computing model.

An important area of focus is on enabling a "managed network" to put the network on an equivalent footing with computing and storage resources. This is important because modern high-energy physics experiments produce a huge amount of data the must be made available to physicists all over the world. Analyzing this data requires significant computing power and storage for both the data and its transformations. All three primary aspects of grid infrastructure (storage, compute and network) must work together to deliver an efficient, high-performing system for collaborations like ATLAS.

In the next few sections I will highlight some of the research into aspects of a managed network for high-energy physics.

4.1. *UltaLight*

The UltraLight[8] collaboration was formed in late 2003 with a goal of enabling the network as a managed resource. Their four year program was funded in September 2004 by the National Science Foundation's Mathematics and Physical Science (MPS) division an is focused upon application driven network research and development. The intent is the by enabling the network as a managed resource they really will enable physics analysis and discoveries which could not otherwise be achieved.

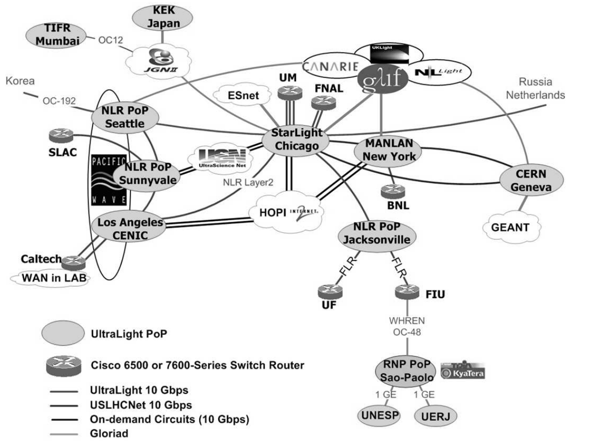

Fig. 2. The UltraLight network testbed as of August 2006

The primary areas of focus for UltraLight are in optical networks and control planes, agent mediated network path discovery and setup, end-to-end data transport and network and end-host monitoring.

4.2. *TeraPaths*

TeraPaths[9] is a DOE MICS/SciDAC-funded project to address the data transfer needs of the high energy and nuclear physics scientific community and effectively support data flows of various levels of priority through modern high-speed networks. TeraPaths is rapidly evolving from a last-mile, LAN QoS provider to a distributed end-to-end network path QoS negotiator through multiple administrative domains. It bridges the gap between data transfer intensive applications and high performance heterogeneous network. Developed as a web service-based software system, TeraPaths automates the establishment of network paths with QoS guarantees between end sites by configuring their corresponding LANs and requesting MPLS paths through WANs on behalf of end users. The primary mechanism for the creation of such paths is the negotiation and placement of advance reservations across all involved domains. Thus, TeraPaths provides end-to-end quality of service users with a one stop network solution and a widely adopted web-services interface.

One example from TeraPaths is shown in figure 3 demonstrating prioritized access to dCache systems at Brookhaven from Michigan in the presence of competing traffic. Terapaths has worked closely

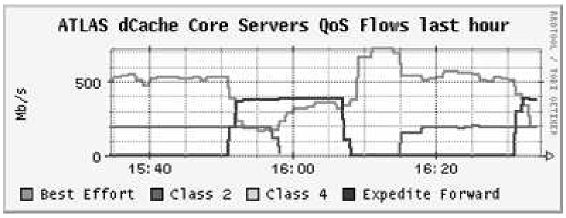

Fig. 3. A plot showing Terapath's managed bandwidth by flows for different classes of service between BNL and Michigan.

with the OSCARS project (see below) and UltraLight in its research efforts.

4.3. *OSCARS*

The ESnet On-demand Secure Circuits and Advance Reservation System [OSCARS[10]] is developing and deploying a prototype service that enables on-demand provisioning of guaranteed bandwidth secure circuits within ESnet. OSCARS leverages existing (and developing) products, services, and code (both from the industry and academia) to accomplish its goals.

The OSCARS objective is to develop and deploy an intra-domain service that can be used by ESnet attached sites, but that will eventually be able to be used by a bandwidth broker to set up inter-domain QoS paths.

Interoperability with emerging standards, in particular the OASISs Web Services Resource Framework (WS-RF) and the Global Grid Forums Open Grid Services Architecture (OGSA), are a focal point in the implementation of this project. The research aspects of this project are the investigation of how all of the various elements of the OSCARS service can properly interact with deployed network tools, and how the overall service can co-exist with the production network.

5. Outlook and Future Possibilites

The research examples shown above represent critical areas of functionality for the creation, deployment and control of a managed network. Such a network could significantly improve the power of the ATLAS computing model by providing the capability to prioritize data transport of varying importance and guarantee needed bandwidth for data transport matching the capabilities of the corresponding storage and compute resources.

The ATLAS experiment is approaching its first real data in about one year and there remain a number of significant challenges to delivering a robust, high-performing infrastructure to expdite the process of physics discovery. It is my expectation that a number of research areas in networking, grids and storage can be exploited to produce such an infrastructure and I look forward to participating in its development, deployment and use.

Acknowledgements

I would like to acknowledge the support of both the Department of Energy and the National Science Foundation who have funded my efforts in ATLAS, UltraLight and Terapaths.

References

[1] The ATLAS webpage is located at http://atlas.web.cern.ch/Atlas/index.html.

[2] LHC, "large hadron collider." Wikipedia, the free encyclopedia. 27 Sep. 2006. Reference.com http://www.reference.com/browse/wiki/Large_Hadron_Collider.

[3] An overview of ATLAS for general audiences can be found at Wikipedia, the free encyclopedia. Retrieved September 27, 2006, from Reference.com website: http://www.reference.com/browse/wiki/ATLAS_experiment.

[4] P. Higgs, Phys. Rev. 145 (1966) 1156.

[5] The ATLAS Computing Technical Design report is located at http://doc.cern.ch/archive/electronic/cern/preprints/lhcc/public/lhcc-2005-022.pdf.

[6] The ATLAS Distributed Data Management (DDM) page has information on DQ2 at http://atlas.web.cern.ch/Atlas/GROUPS/DATABASE/project/ddm/.

[7] The LCG gLite FTS Twiki has information at https://uimon.cern.ch/twiki/bin/view/LCG/GliteFTSInformation.

[8] The official UltraLight web page is at http://www.ultralight.org.

[9] The official TeraPaths web page is at http://www.atlasgrid.bnl.gov/terapaths/.

[10] The OSCARS web page is located at http://www.es.net/oscars/index.html.