Status of the LCG Project

CERN-C-RRB-2005-002
Les Robertson
4 April 2005

Introduction

This status report covers the period from October 2004 to March 2005. Further details on progress, planning and resources, including detailed quarterly progress reports, can be found in the documents linked to the LCG Planning Page on the web.

Chart 1: Number of LCG Level 1&2 Milestones (cumulative)

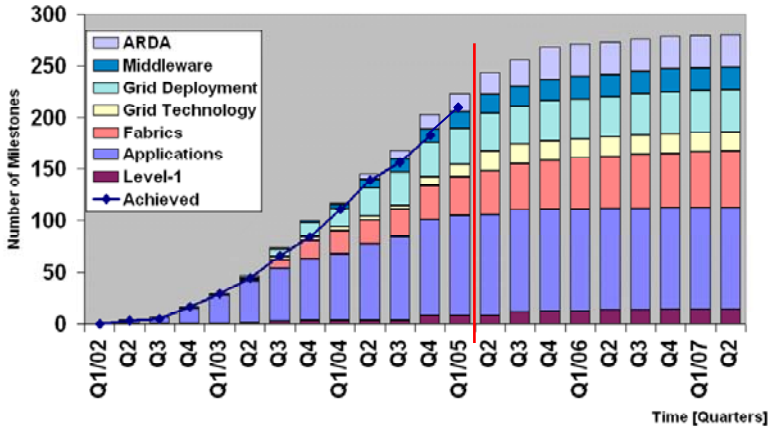

Chart 1 shows the milestone performance at the end of the last quarter. This year the project is in a state of evolution from Phase 1 to Phase 2. While Phase 1 was concerned with evaluation and development, Phase 2 will focus more on supporting the common tools developed in Phase 1 and installing and commissioning the computing services required at LHC start-up. The planning for Phase 2 will be completed with the publication of the Technical Design Report, due at the end of June. This process will result in the definition of new milestones in all areas of the project.

The comprehensive review of the project by the LHCC took place in November. The referee's report is available as CERN/LHCC 2004-032. Concerns were expressed with delays in the delivery of the new version of the CASTOR mass storage system at CERN, delays in the availability of the first release of the gLite grid middleware, and the difficulties resulting from the multiple grid environments through which computing resources are being provided for LHC experiments. These issues have been followed up in subsequent meetings with the LHCC referees and are referred to in later sections of the present status report.

The number of LHC computing centres accessible through grid technology has already reached the scale required for the Tier-1 and Tier-2 facility at LHC start-up. An agreement has been reached between the experiments and the major centres on the two year schedule to build up the production LHC service. Defined as a series of *Service Challenges*, this builds up the service from the core of Tier-1 and major Tier-2 centres.

A working group, including representatives of the experiments and regional computing centres, is presently tackling the definition of the baseline services to be provided to support the computing models during the initial period of LHC operation.

Applications

Pere Mato (CERN) took over responsibility as Applications Area manager at the beginning of March from Torre Wenaus (BNL), who returned full time to ATLAS at the end of his three year mandate. Torre leaves the Applications Area in good shape:

- the POOL object persistency system has been successfully used in large scale production in ATLAS, CMS and LHCb data challenges in 2004, with more than 400 TeraBytes of POOL data produced; the objective of a quickly-developed persistency system leveraging the storage management technologies of the ROOT project (for bulk data) and of Relational Database Management Systems has been fulfilled;

- Geant4 has been firmly established as baseline simulation in successful ATLAS, CMS, LHCb productions; electromagnetic and hadronic physics validation has been completed; the system has been shown to be very stable in use - ~1 Geant4-related crash per O(1M) events generated;
- the components developed in the SEAL (*Shared Environment for Applications at Lhc*) project have underpinned POOL's success, in particular the dictionary system that is now entering a second generation with a new implementation called *Reflex;*
- the Savannah project portal and external software service are accepted standards inside and outside the project (e.g. it is also used by the EGEE project).

The high level programme of work for the Applications Area for the next four years is now being prepared, and the major elements of this have been discussed at an internal Applications Area Review at the end of March. The intention is to reorganise this part of the project to prepare for the commissioning phase of the experiments' software systems, and to lay the groundwork for the support of end user data analysis. Among the options being considered is the close integration of the ROOT package with the rest of the common tools developed in the framework of the LCG Applications Area.

CERN Fabric

The LCG Level 1 milestone to demonstrate sustained data recording at 450 MBytes/sec. for seven consecutive days was successfully completed on 8 March. This was performed in close collaboration with the ALICE experiment. The test had been delayed in order to be able to use the new version of the CASTOR mass storage management system, which had suffered long development delays. When finally available at the beginning of the year, the new CASTOR version was brought into operation very rapidly, with a good collaboration between the IT developers and the online group in ALICE. Figure 1 shows the tape recording rate during the data challenge. Note in particular the area outlined by the green loop - showing the resilience of the system to hardware failures. The CASTOR team has made up about two thirds of the 9-month delay measured at the end of last year, and expects to be back on schedule by the middle of the year.

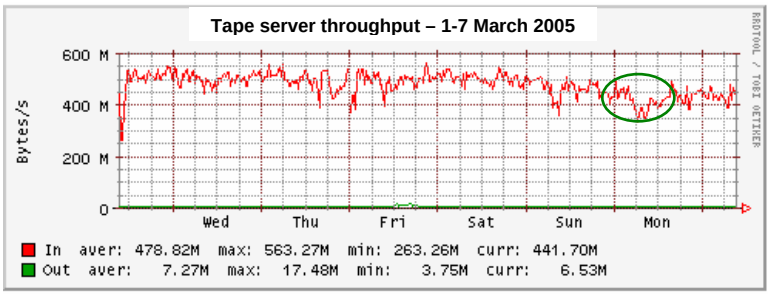

Figure 1 – Tape server throughput during data recording challenge

Networking

The acquisition of the internal network equipment for the CERN facility is now in the final stages of the tender process and it is expected that installation will take place on schedule, starting in the second half of this year. Progress is also being made in the provision of the wide area networking infrastructure for the Tier-0 and Tier-1 centres. A special working group of the LCG Grid Deployment Board was set up to bring together the regional centres and the national and regional research network providers in order to develop an architecture and implementation plan to satisfy the LHC requirements and timetable. Fortuitously this work was concurrent with the completion of the planning of the next phase of the European Research Network Backbone, GÉANT, which takes LHC needs into account. Work is well advanced and it is expected that agreement will be reached in time to permit the first elements to be in production by the end of this year.

Grid Deployment

The resources available and planned for LHC computing are embedded in three independent operational grids - NorduGrid with about 30 sites mainly in the Nordic countries, the Open Science Grid (OSG) with about 25 sites in the USA, and the EGEE grid with more than 120 sites in Europe and the rest of the world. The EGEE grid integrates many large national grid organisations and provides the majority of the resources available to the experiments. The LCG project has a formal and close relationship with the EGEE grid, which grew out of the earlier LCG grid operation of 2003-04, and the operation of EGEE is managed

jointly by the two projects, with decisions affecting the HEP sites being subject to the LCG Grid Deployment Board. LCG does not have formal relations with OSG and NorduGrid, but the sites in these grids that provide resources for LHC are represented in the project through the Grid Deployment Board.

At an informal level there has been important progress on inter-operability of the OSG and EGEE grids. Resources in each grid are now visible to the other, and cross submission of work has been demonstrated. Work on inter-operability is continuing, and it is hoped that the gLite middleware (discussed below) will bring new opportunities for inter-operation. Some discussions have taken place with the NorduGrid software developers, but so far there have not been any concrete steps towards inter-operability at this level.

In terms of scale of grid operation, the target for the end of 2004 was to have 20 sites with 2,000 processors operating as a coherent grid. This has been achieved independently by the three grid organisations, with the EGEE grid itself now connecting over 120 sites, presenting over 10,000 processors. In terms of scale this is already close to that needed for the full LHC grid. This rapid expansion has been possible through development work last year to improve the scalability of the middleware, but is also a result of the EGEE project's investment in a solid distributed operations structure.

An LCG Grid Operations workshop took place at CERN in November, leading to an agreement on the basic processes and policies for a range of operations management and user support issues. The workshop was attended by EGEE, OSG and NorduGrid operations staff. The EGEE grid operation is now well under way, with day-to-day operations management responsibility cycled on a weekly basis around the operations centres at CNAF, IN2P3-Lyon, RAL, MSU (Russian Data Intensive Grid), and CERN. ASCC-Taipei should join in this rota in the near future. The role of the user support centre at FZK is developing within this structure. A centre at the University of Indiana provides similar operations services for the OSG grid.

The major issue for the EGEE grid remains reliability, with a mixture of software problems, and operational failures at individual sites. The new operations management activity will help in monitoring and tracking down problems, and new more realistic tests are being added to the standard monitoring framework. Close cooperation of the experiments is also essential, as problems are often only visible to the application itself. This will be a major focus of work for at least the next six months.

The accounting system (responsibility of RAL) is now in operation, collecting data automatically from EGEE sites, and accepting manual input from sites in OSG and NorduGrid. At present, however, only the major sites are reporting resource usage centrally.

Service Challenges

A program of Service Challenges has been agreed to guide the setting up of the core LHC computing services over the next two years, with the aim of having the service fully commissioned at all Tier-1 and major Tier-2 sites by 1 April 2007. Four Service Challenges have been defined, progressively building up the functionality and capacity of the LCG service, with defined milestones for integrating specific sites and achieving performance targets. Service Challenge 1 (SC1) was scheduled to complete in December 2004, demonstrating sustained data movement at 500 MB/sec between CERN and three Tier-1 sites. This was achieved with FNAL in November and with FZK in January, and at a rather lower data rate with NIKHEF/SARA. Although the SC1 goals were not achieved, a great deal was learned at CERN and at other sites, and this groundwork enabled Service Challenge 2 (SC2) to get up to speed rapidly during March. The SC2 target was to demonstrate reliable file transfer between CERN and seven Tier-1s (BNL, CNAF, FNAL, FZK, IN2P3, NIKHEF/SARA and RAL), sustaining an aggregate throughput of 500 MB/sec at CERN. SC2 finished on 2 April with all sites active for 10 days, averaging 605 MBytes/sec, with a 12-hour peak of 820 MBytes/sec (see Figure 2). This is a major step forward in demonstrating the close technical collaboration between the operations teams at these sites.

While SC2 demonstrates data distribution from CERN to Tier-1 centres, a fundamental component of the LCG service, Service Challenge 3 (SC3) aims to go much further towards providing a service that can be used to check out the experiments' computing models. It will include all of the Tier-1 centres with the exception of the Nordic facility, access to mass storage (magnetic tape), and will integrate a number of Tier-2 sites. The target is to complete the first phase (setup and throughput demonstration) by end July, and open a stable test service for the experiments in September. A final Service Challenge (SC4) is being planned for 2006, culminating in the opening of the initial LHC Computing Service in September 2006.

This will be operated as a stable operational service, progressively integrating additional sites and resources to achieve full design capacity and performance at least three months before the first physics run.

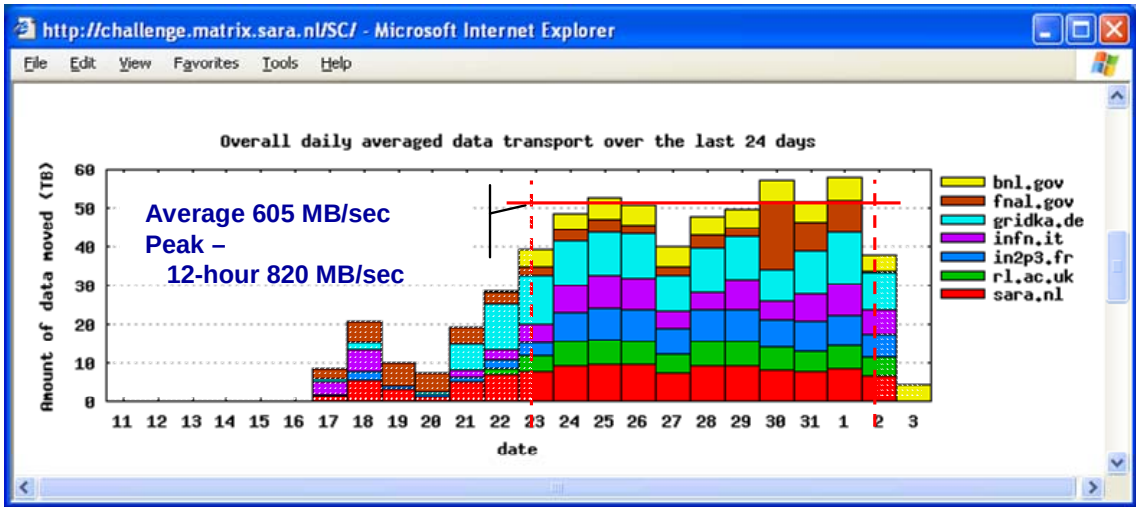

Figure 2 - Daily data transfer (TeraBytes) between CERN and Tier-1s

A summary of the current Service Challenge plans is available (CERN-LCG-PEB-2005-05)

Computing Models and Resource Requirements

The experiments' computing models and requirements were published in December and reviewed by a special meeting of the LHCC in January. These requirements, revised in the case of ALICE in the light of the reviewers' recommendations, now form the guidelines for the planning in regional centres. The detailed schedule for processing of the data from the heavy ions run needs further discussion, to see the extent to which peak processing demands can be reduced.

The EGEE gLite Grid Middleware Development

The EGEE project undertook a management review of the *gLite* middleware development activity at the end of 2004 and decided in January to proceed with a single software stack, effectively the stack on which the LCG-2 toolkit is based. This should have the advantage of enabling new *gLite* components to be introduced progressively into the LCG-2 environment. It should also help to retain a certain level of compatibility with OSG in the US. On the other hand, discontinuing support for the AliEn software base will cause real difficulties in the short term for the ALICE experiment, and so LCG will support effort to help ALICE in the transition of their software to the *gLite* base.

The components of the first release of *gLite* have been defined and a very tight schedule agreed to test and deliver them to grid deployment. The deadline for availability to the experiments (on the pre-production service) is the beginning of April, and the schedule foresees distribution to major sites at the end of May. At the time of writing this report it is not clear which components of gLite will meet the end May deadline to be installed in the production service for SC3.