THE LHC COMPUTING GRID

Andrea Sciabà
CERN, Switzerland

Abstract

The LHC experiments will start taking data in about two years and will produce about 15 PB of data each year. The goal of the LHC Computing Grid (LCG) is to provide the infrastructure and the environment for the preparation and the operation of the experiment computing systems. In this contribution, the global architecture, the main services and functionalities, and the challenges to be addressed by the LCG are described.

1 Introduction

The Large Hadron Collider will become operational at CERN in 2007, and the four major experiments (ALICE, ATLAS, CMS, LHCb) will collect an unprecedented amount of data, estimated to be of the order of 15 PB/year. These data will have to be processed, calibrated and analyzed several times,
requiring a huge amount of computing and storage resources. In order to meet the computing requirements of the LHC experiments, the LHC Computing Grid project (LCG) was created.

The LHC computing infrastructure follows the Grid paradigm of a federation of heterogeneous computing and storage resources. This model maps very well with the structure of the LHC collaborations (4000+ physicists from 100+ of research institutes worldwide) and was motivated by the impossibility to physically concentrate in a single location all the necessary resources and by the desire to integrate the local resources of the participating institutes \(^1\). This implies to have an adequate network between the different sites, a uniform interface to heterogeneous resources and an efficient and fair access to them. The project is developing in two phases: in the first one, a prototype is built and experience is accumulated on how to operate a world-wide Grid, while in the second one, beginning in 2006, the full-scale production service will become operational. As we write, LCG includes approximately 170 sites, 18000 CPUs and 4 PB of storage capacity.

In the following, a description of the current status of the LCG project will be given.

2 The LCG architecture

All the experiments have chosen to implement for their computing system a multi-tiered, hierarchical model, where different sites have different roles as a function of the services offered and the support level. A schematic data flow among tiers is depicted in fig. 1 (left).

A single Tier-0, located at CERN, reconstructs the raw experimental data coming from the data acquisition systems and distributes the raw and the reconstructed data according to the experiment policies to each Tier-1 centre.

Each Tier-1 centre supplies a complete range of services (disk and tape storage, computing farms, databases) and support to end users and to a given set of Tier-2 centres. They perform data analysis and reconstruction, and provide archiving for raw, reconstructed and Monte Carlo data.

The primary roles of Tier-2 centres are event simulation and end-user analysis. No archival service is required, but partial copies of specific data samples can be hosted if required.
3 Overview of the LCG middleware

The middleware adopted by the LCG project is developed by external sources, research projects and institutions. It is currently based on the Virtual Data Toolkit \(^2\) and the DataGrid middleware \(^3\). Other important components include the GLUE information schema \(^4\), which allows interoperability with other Grids, the dCache storage system \(^5\), the LCG File catalogue \(^6\) and the Disk Pool Manager \(^6\). The middleware developed by the EGEE project \(^7\), gLite, is gradually being introduced: the File Transfer Service (FTS) is an example of a gLite service already used in LCG.

Resources in LCG are of two kinds: computing elements (CE), which provide a Grid interface to a whole farm and its batch system, and storage elements (SE), which provide a uniform interface to a disk server or a tape-based mass storage system. An information system maintains an updated view of the status of the Grid resources through a hierarchical structure of information providers and publishers. Information about all files stored in the LCG SEs is stored in file catalogues, global or local, which associate logical names to physical files. Data are accessed via transfer protocols, like GridFTP \(^8\),
using reliable file transfer tools (e.g. FTS), or directly via secure POSIX-like protocols⁶). The distribution of jobs to CEs is done in a load-balanced way by a central service, the resource broker.

A number of accounting and monitoring systems displays, records and analyzes the information available from the information system, the fabric monitoring and the result of periodic test jobs to give an accurate view of the status of LCG.

4 LCG deployment and operations

In order to integrate the middleware components in a coherent distribution, LCG has set up a thorough testing and certification procedure. If the tests are passed, a candidate release is defined and made available to the experiments for further testing. Then, a distribution for deployment is prepared, adding installation tools, configuration files, release notes and documentation, and is installed world-wide⁹).

The operation of the LCG infrastructure is done by a distributed support structure¹), managed together with the EGEE project. It includes continuous site monitoring, problem solving through a multi-level support team, coordination of middleware installation and upgrade, user support and documentation. Regional Operations Centres (ROC) coordinate and support sites in their region, Core Infrastructure Centres (CIC) maintain central services and perform, in rotation, the daily monitoring, while the Operations Management Centre, which is located at CERN, oversees the LCG deployment and provides high level expertise. The User Support Centre (GGUS) coordinates the user support by providing a central portal.

5 LCG and the Experiments

The LHC experiments are performing an increasingly substantial part of their computing activities on LCG resources, from prototyping of their computing models to running standard applications like Monte Carlo production or analysis of simulated data. This strongly helps in identifying the limitations of the system and steers the middleware development based on requirements derived also from actual experience.
In 2004 and 2005, all experiments have performed a series of Data Challenges, simulating the data processing flow during the LHC data taking at a sizeable fraction of the full scale, and Monte Carlo productions \(^{10)\ 11)\ 12)\ 13)\); in fig.1 (right) the number of jobs per day run by ATLAS during data challenges is shown.

These data challenges served multiple purposes: they were milestones towards the final commissioning of the experiment computing systems, but they were also effective in exposing some weaknesses of the system. They showed that site misconfiguration was responsible for a large fraction of the failures (failure rates ranged from 5% to 40% depending on the application), which lead to a big effort being invested in tools to identify and minimize the effects of site problems. Other issues were related to insufficient scalability and reliability of some services (RLS file catalogue, data management tools).

An important observation made from such large-scale exercises was that the system still requires a significant human effort to be operated.

In 2005, experiments have been focusing more on using LCG to run user analysis jobs with user-friendly interfaces like CRAB \(^{14)\} or Ganga \(^{15)\}.

The transformation of the LCG in a real production service is happening gradually through Service Challenges, which must attain certain goals in terms of inter-site data transfer and job submission rates and stability of services, both from the point of view of LCG and the experiments. The schedule foresees to have the initial LHC service in stable operation by September 2006, ready for the beginning of the data taking in April 2007.

References


14. F. Fanzago et al, CRAB: a tool to enable CMS Distributed Analysis, in: these proceedings