

ATLAS Data Challenges: large-scale productions on the Grid

Luc Goossens^{*}, Gilbert Poulard, Dario Barberis

CERN

Geneva 23, Switzerland

E-mail: luc.goossens|gilbert.poulard|dario.barberis@cern.ch

The ATLAS Collaboration is preparing for data taking and analysis at the CERN LHC scheduled to start operating in 2007. As part of this preparation, a series of computing Data Challenges (DC) was started in 2002, with the goal of validating the Computing Model, the complete software suite, and the data model, and to ensure the correctness of the technical choices to be made for the final offline computing environment. DC1 was conducted during 2002-03; the main goals achieved were to set up the production infrastructure in a real worldwide collaborative effort, and to gain experience in exercising an ATLAS wide production model. DC2, run from May until December 2004, was a similar exercise but deployed a new Production System able to submit jobs on three different Grids: LCG, OSG/GRID3 and Arc/NorduGrid. Right after DC2, a new Monte Carlo production was necessary to produce the needed event samples for an ATLAS Physics workshop (called Rome workshop) taking place in June 2005. In this paper we present our experience with these large-scale productions conducted on the three aforementioned Grids.

European Physical Society

HEP2005 International Europhysics Conference on High Energy Physics

EPS (July 21st-27th 2005) in Lisboa, Portugal

^{*} Speaker

1. Introduction

In 2007 the Large Hadron Collider (LHC) [1] is due to come into service at the European Particle Physics Laboratory (CERN) in Geneva. In the LHC two proton beams, each of energy 7 TeV, are steered to collide head-on at the centre of large complex detectors. These collisions are expected to reveal fundamental new processes in particle physics.

The ATLAS [2] (A Toroidal LHC ApparatuS) detector is one of four detectors that are being constructed to exploit the LHC. The expected volume of data recorded for off-line reconstruction and analysis will be of the order of 1 PB (10^{15} bytes) per year, to be analysed by institutes around the world. The needed computing power to process this data is at a scale that is one to two orders of magnitude larger than previous experiments in particle physics.

To prepare for this task, the ATLAS collaboration decided to carry out Data Challenges (DCs) of increasing size and complexity. The purpose of these DCs is to validate the ATLAS Computing Model, software suite, and data model, and to ensure the correctness of the technical computing choices made. The DCs are performed in the context of the LHC Computing Grid project, LCG [3], to which ATLAS is committed, but ATLAS also uses both the middleware and the resources of two other Grid projects: OSG/GRID3 [4] and ARC/NorduGrid [5].

The ATLAS Production System is described in Section 2. Section 3 gives some details on the achieved productions. Section 4 reports on our experience and an overall conclusion and outlook is given in section 5.

2. ATLAS Production System

In order to handle the task of ATLAS DC2 an automated production system [6] was designed. All jobs are defined and stored in a central database (prodDB). Supervisor agents (Windmill) pick them up, and send their definition as XML message to an executor via a Jabber server. Executors are specialised agents, able to convert the XML job description into a Grid specific language (e.g. JDL, job description language, for LCG). Five executors have been developed, two for LCG (Lxor and Lxor-CG), one for NorduGrid (Dulcinea), one for GRID3 (Capone) and one for legacy systems, allowing the Data Challenge to be run on different Grids.

For data management, a central component, Don Quijote (DQ) offers a uniform layer over the different replica catalogues of the 3 Grid flavors.

3. Details on Achieved Productions

DC2 started in July 2004 and lasted until December, immediately followed by the "Rome" production that lasted until May 2005. Figure 2 shows the number of jobs run per day. The maximum number of parallel running jobs reached on all Grids together was 5800 jobs.

For DC2 in total 10 million events were processed ~340 thousand jobs, consuming ~200 kSI2K years of CPU and producing ~60 TB of data. Eighty-seven sites spread over 20 countries participated. For the Rome production 8.5 million events were simulated and over 6 million reconstructed, in ~570 thousand jobs. Here 114 sites in 22 countries participated.

	DC2			Rome			
Grid	LCG	NorduGrid	OSG/Grid3	LCG	LCG-CG	NorduGrid	OSG/Grid3
Sites	40	21	26	67		13	34
Jobs	116215	104959	117312	191207	179197	59383	131238
%	34	31	35	34	32	11	23

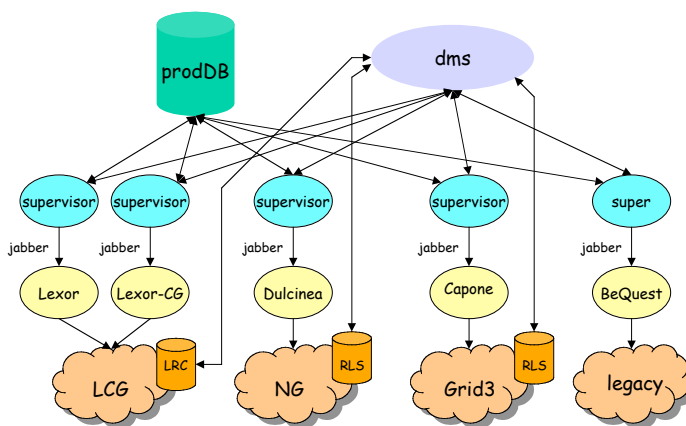


Fig. 1. The ATLAS production system architecture

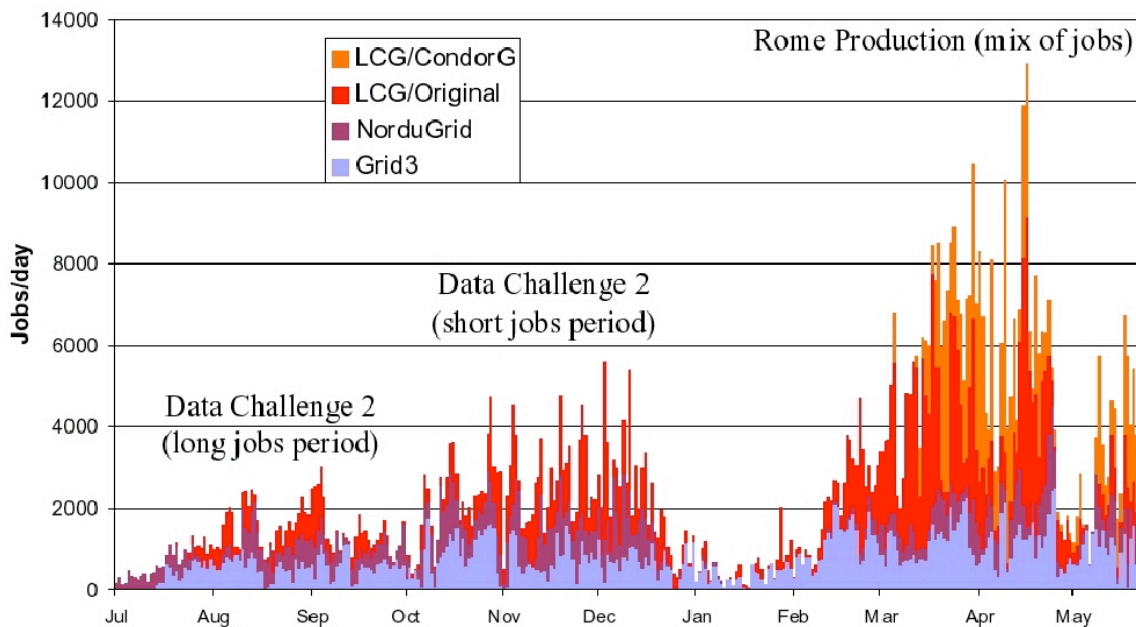


Fig. 2 Number of successful jobs per day

POS (HEP2005) 394

4. Experience

By design the ATLAS production system had a very high dependency on the Grid middleware. During DC2 many problems were identified and reported to the Grid developers, forming the basis for improvements. For some problems a workaround needed to be included in the production system. One example of the latter was the slowness of the job submission in the LCG. Even when using several submission points, it was not possible to saturate the LCG. This led to the development of a second executor called Lxor-CG (for CondorG) that submits jobs directly to the LCG computing elements (bypassing the resource broker). This promptly doubled the number of jobs we were able to run on the LCG resources (see figure 2).

Also on the ATLAS side, DC2 started before development was finished and many problems were found and corrected while running the DC. In general, our software was not prepared for the unreliability of the middleware: error handling, logging, etc. were too minimal. In addition, due to manpower shortage, the data management component DQ2 was never finished, lacking important functionality to manage the huge amount of files involved.

5. Conclusions and Outlook

The three Grids in use in ATLAS have proven to be usable for large-scale productions. It was possible, however not easy, and requiring a non-scalable amount of manpower. All involved components both on the Grid and the ATLAS side are clearly in need of substantial improvements. Problems were coped with, but often with work-arounds or ad hoc fixes

An improved version of the production system is in pre-deployment phase. It features the same architecture, but has a new implementation of the crucial supervisor component (Eowyn). The data management project has been split off and now has more and separate manpower. It also features a new design that relies less on the middleware.

In autumn/winter 2005/2006 ATLAS will participate with this new production system in the so-called Service Challenge 3. The new production system will also form the basis of the ATLAS Tier-0 Management System that will orchestrate the operation of the ATLAS Tier-0.

Acknowledgements

We thank all members of the ATLAS collaboration who participated in this effort, the supporting people of the three Grids, and the people involved in the maintenance of each site.

References

- [1] CERN-LHC: <http://www.cern.ch/lhc>
- [2] ATLAS: <http://atlas.web.cern.ch/Atlas/index.html>
- [3] LCG: <http://lcg.web.cern.ch/LCG>
- [4] The Grid 2003 Project: <http://www.ivdgl.org/grid2003/index.php>
- [5] NorduGrid: <http://www.nordugrid.org>
- [6] L. Goossens, K De, *ATLAS Production System in ATLAS Data Challenge 2*, in proceedings of CHEP 2004