Distributed Data Analysis in the ATLAS Experiment: Challenges and Solutions

Johannes Elmsheuser[1], **Daniel van der Ster**[2] **on behalf of the ATLAS collaboration**

[1] Ludwig-Maximilians-Universität München, Germany
[2] CERN, Switzerland

E-mail: `johannes.elmsheuser@physik.uni-muenchen.de`

Abstract. The ATLAS experiment at the LHC at CERN is recording and simulating several 10's of PetaBytes of data per year. To analyse these data the ATLAS experiment has developed and operates a mature and stable distributed analysis (DA) service on the Worldwide LHC Computing Grid. The service is actively used: more than 1400 users have submitted jobs in the year 2011 and a total of more 1 million jobs run every week. Users are provided with a suite of tools to submit Athena, ROOT or generic jobs to the Grid, and the PanDA workload management system is responsible for their execution. The reliability of the DA service is high but steadily improving; Grid sites are continually validated against a set of standard tests, and a dedicated team of expert shifters provides user support and communicates user problems to the sites. This paper will review the state of the DA tools and services, summarize the past year of distributed analysis activity, and present the directions for future improvements to the system.

1. Introduction

The LHC at CERN is colliding protons or heavy ions at unprecedented center-of-mass energies and these collisions are recorded by several experiments including the ATLAS experiment. The large number of collision events and resulting data volumes require an analysis on distributed computing resources.

2. ATLAS distributed analysis overview

The challenges in a LHC data analysis are high: there are large data volumes from the detectors and simulation which require a large number of CPUs for processing. There is a complex experiment software structure and a high connectivity to the data that needs to be provided. LHC experiments produce and store several PetaBytes/year, the ATLAS experiment for example has recorded more than 1.65 billion events from proton-proton collision at a center-of-mass of $\sqrt{s}$ =7 TeV in the year 2011. The collision event complexity due to the large number of detector channels and number of analysis users demand at least 100k modern CPUs based on the computing models of the experiments. Therefore the distributed analysis tools used to analyze the data should be easy to configure, reliable, fast to work with, and analysis jobs should have a high success rate possibly at the first attempt.

ATLAS jobs on distributed Grid resources are typically classified into 'production' and 'analysis'. Production jobs consist of Monte Carlo (MC) generation, simulation, and reconstruction as well as real data reprocessing and production of physics group ntuples. These jobs process data which can be of interest to either the entire collaboration or large physics groups; in both cases it is submitted and monitored by an automated central production system. On the other hand, analysis jobs are submitted by individual physicists to analyse the centrally-produced data for her/himself or on behalf of a small group of physicists.

To illustrate, the distributed analysis model of the ATLAS experiment is explained in more detail in the following. Being based on the ATLAS computing model [1], ATLAS distributed analysis has the following key characteristics:

- The detector and simulation data are distributed worldwide by the data management system DQ2 [2] among the different computing centers which are organized in a hierarchical tier structure.
- The different releases of the experiment software Athena [1] are installed at the different sites worldwide by a central installation system.
- The user analysis code and tasks are sent to the sites where the data to be analyzed is located. The model is called: "Jobs to Data".
- The distributed analysis tools for handling all the user job management are Ganga [3] and the Panda clients tools [4].
- The output and results of the user jobs are written to the scratch disk of the site where the job was executed and can be transferred asynchronously as needed to a remote site.
- The results can be retrieved back to the user's local desktop computer using the DQ2 command line tools.

ATLAS foresees different analysis work-flows for the data formats produced by the central production system:

- Athena user code sequentially processes large MC or data stream samples on the Grid in parallelized jobs. These jobs produce ROOT tuple outputs which are further processed locally or on the Grid.
- TAG files which contain a condensed event summary can be used to process only events of interest and offer the possibility to seek through large data samples.
- Small MC samples can be produced using the production system transformations for special or official usage.
- Generic ROOT applications eventually with DQ2 access can be executed.

The analysis jobs can be executed on various resources. The primary and recommended way to access the Grid resources is through the Panda workload management system. This pilot job based system with a central queue operates on the LCG/EGEE/EGI [5], OSG [4] and NDGF/Nordu-Grid [6]. Analysis is also performed on non Grid resources via batch systems like the LSF system at CERN, SGE at the DESY NAF or PBS at many other sites. Figure 1 summarizes the various workflow paths.

A dedicated team of expert shifters provides user support and communicates user problems to the sites. ATLAS users report their problems with job or tool failures to a mailing-list which is monitored by a shift team which is active for 16 hours each day during the European and American day time.

3. ATLAS distributed analysis usage

In the year 2011 more than 1500 users have submitted distributed analysis jobs to the Panda system. Figure 2 (left) shows the total number of completed analysis and productions jobs from April 2011-April 2012. More than 133 million analysis jobs have been executed, which averages to approximately 366000 jobs per day. Figure 2 (right) shows the average number of

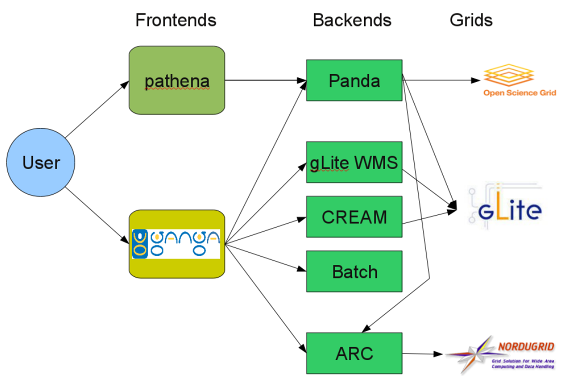

Figure 1. Overview of the different Grids and Job submission mechanisms offered to the users in the ATLAS experiment.

simultaneous running jobs from April 2011-April 2012. This corresponds to the available number of total job slots or CPU cores in the Grid available for the ATLAS experiment. Between 80000-120000 analysis and production jobs have been running together simultaneously, with 20000-40000 analysis jobs only. The share of analysis jobs on the total available resources has been limited to give production jobs higher priority. At Tier1 sites it was initially set to 20% and lowered at the beginning of 2012 to 5%. At Tier2 sites there is a share of 50% for analysis jobs. Analysis jobs are usually rather short compared to MC production jobs. They only last

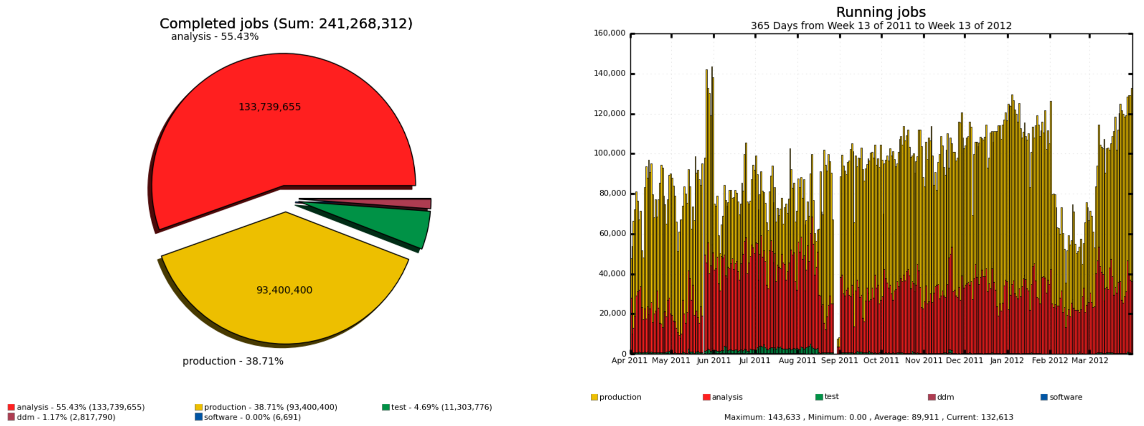

Figure 2. Left: Number of completed analysis and productions jobs from April 2011-April 2012. Right: Number concurrently running jobs for analysis and production in the same time period.

between 10 to 60 minutes on average. This can be followed from Figure 3 which shows the wall-clock consumption time for analysis and production jobs in the time range from April 2011 until April 2012. The total analysis share is only approximately 22% of the total consumed job wall-clock time, while there is a share of approximately 55% of analysis in the total number of jobs. Analysis jobs have many different types of input formats. Athena type of jobs use

Journal of Physics: Conference Series **396** (2012) 032035 doi:10.1088/1742-6596/396/3/032035

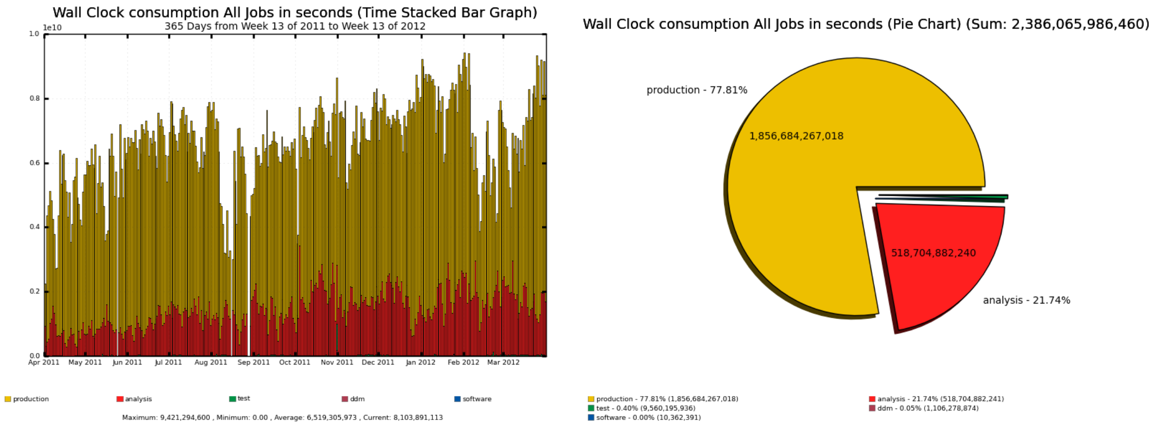

Figure 3. Left: Wall-clock consumption time for analysis and production jobs in the time range from April 2011 until April 2012. Right: Total Wall-clock consumption time in the same time period.

AOD, ESD and RAW data formats as input which are together about 1/3 of the analysis jobs as can been seen in Figure 4. Other input types are private ROOT n-tuples which are marked "unknown" in Figure 4 and many different type of physics group ROOT n-tuples marked with "NTUP_". Jobs running at the Grid sites trigger the so called "PD2P" algorithm in the Panda system: the requested input data is automatically replicated to another site which currently has available resources. If the jobs wait in the Grid queues too long they can re-brokered to the new site [7].

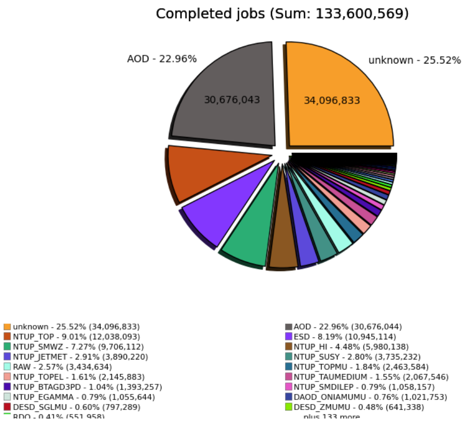

Figure 4. Distribution of data formats used in analysis jobs in the time range from April 2011 until April 2012.

4. ATLAS distributed analysis efficiency

Figure 5 (left) shows the distribution of successful, cancelled and failed distributed analysis jobs from April 2011-April 2012. About 75% of all jobs succeed, while about 12% of jobs have failed. The distribution of error categories is shown in Figure 5 (right). A larger fraction of job failures is due to temporary site and infrastructure problems: input files cannot be copied or accessed

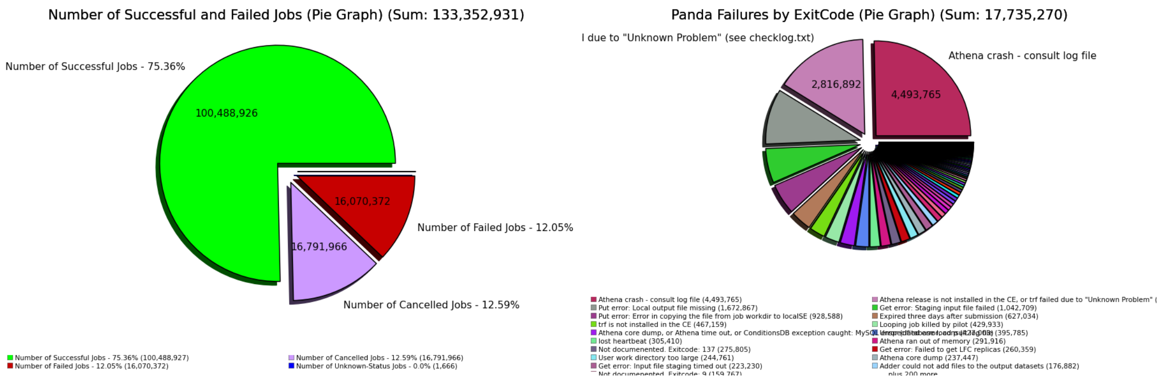

Figure 5. Left: Total number of successful, failed and cancelled analysis jobs from April 2011-April 2012. Right: Error category break-down for the failed jobs in the same time range.

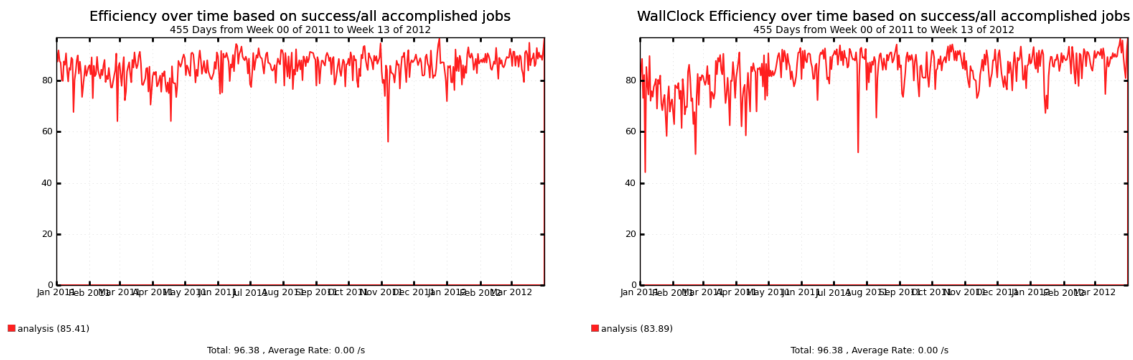

Figure 6. Left: Relative analysis job efficiency from January 2011-April 2012. Right: Relative wall-clock time job efficiency in the same time range.

at the local storage element or output files cannot be stored. All errors have been categorized into retriable and non-retriable errors. In this first version of the categorization 20% of failures are retriable, i.e. jobs are automatically rescheduled and 66% succeed after the second attempt.

To improve the job efficiency and protect users against temporary malfunctioning sites, the HammerCloud system is used [8, 9]. This system sends single analysis test jobs to all sites, so that there is always one job executed at all sites. If more than 2 jobs in a row fail at a site, this site is temporarily blacklisted, so that no further analysis jobs are sent or executed there. If the test jobs succeed again the queue is whitelisted for analysis. Figure 6 shows the overall analysis job efficiency from January 2011-April 2012 (left) and the wall-clock time job efficiency in the same time period (right). The automatic black- and whitelisting algorithm was introduced in June 2011. The improved overall job efficiencies due to this system can clearly be seen in both plots.

5. Future plans and conclusions

As shown in the previous section the overall job efficiency has improved with the introduction of the HammerCloud system. Continuous monitoring of sites and the central infrastructure with new test cases will help to increase the efficiency even more. Furthermore increasing the number of retriable errors will also lead to higher efficiencies. For this the error categorization and identification in the various tools needs to be improved.

Slow job submission speed is sometimes reported by users. This can be improved by moving parts of the job splitting and input data detection algorithms from the client tools to the Panda server.

More and more analysis is done using ROOT with ROOT n-tuple input. Improved job error detection and input file handling for this work-flow is planned and in the works by dedicated client tools or job wrapper on the Grid.

In conclusion, the distributed analysis system of the ATLAS experiment is processing 300000-400000 jobs per day and is working well. Steady improvements in the tools have been and will be made to improve the user experience and job efficiency in this the very complex world wide system.

References

[1] G. Duckeck *et al.* [ATLAS Collaboration], "ATLAS computing: Technical design report," ATLAS-TDR-017 ; CERN-LHCC-2005-022.
[2] M. Branco *et al.* [ATLAS Collaboration], J. Phys. Conf. Ser. **119** (2008) 062017.
[3] J.T. Moscicki *et al.*, "Ganga: a tool for computational-task management and easy access to Grid resources", Computer Physics Communications, Volume 180, Issue 11, (2009), doi:10.1016/j.cpc.2009.06.016
[4] T. Maeno [ATLAS Collaboration], J. Phys. Conf. Ser. **119** (2008) 062036.
[5] P. Andreetto *et al.*, J. Phys. Conf. Ser. **119** (2008) 062007.
[6] M. Ellert *et al.*, Future Generation Computer Systems **23** (2007) 219.
[7] T. Maeno, "PD2P : PanDA Dynamic Data Placement for ATLAS", these proceedings.
[8] D. van der Ster *et al.*,"Experience in Grid Site Testing for ATLAS, CMS and LHCb with HammerCloud", these proceedings.
[9] F. Legger *et al.*,"Improving ATLAS grid site reliability with functional tests using HammerCloud, these proceedings.