ATLAS Distributed Computing Monitoring tools after full 2 years of LHC data taking

Jaroslava Schovancová for the ATLAS Collaboration

1 Institute of Physics, Academy of Sciences of the Czech Republic, Na Slovance 2, CZ-18221 Prague 8, Czech Republic
E-mail: schovan@fzu.cz

Abstract. This paper details a variety of Monitoring tools used within ATLAS Distributed Computing during the first 2 years of LHC data taking. We discuss tools used to monitor data processing from the very first steps performed at the CERN Analysis Facility after data is read out of the ATLAS detector, through data transfers to the ATLAS computing centres distributed worldwide. We present an overview of monitoring tools used daily to track ATLAS Distributed Computing activities ranging from network performance and data transfer throughput, through data processing and readiness of the computing services at the ATLAS computing centres, to the reliability and usability of the ATLAS computing centres. The described tools provide monitoring for issues of varying levels of criticality: from identifying issues with the instant online monitoring to long-term accounting information.

1. Introduction

The Large Hadron Collider (LHC) at CERN (European Organization for Nuclear Research) has been delivering stable beams colliding at the centre-of-mass energy of 7 TeV since the end of March 2010, and at the centre-of-mass energy of 8 TeV since mid April 2012. The ATLAS Experiment [1] is one of the general purpose detectors at the LHC. ATLAS has been efficiently taking data; over 4 PB of RAW data has been accumulated over the past 2 years. ATLAS heavily uses the Worldwide LHC Computing Grid (WLCG) to process the data and simulations. The ATLAS Distributed Computing (ADC) team [2] is responsible for optimizing usage of the ATLAS grid resources while addressing the evolving requirements coming from the Physicists of the ATLAS Collaboration.

The ATLAS Distributed Computing infrastructure is a very complex system: The ATLAS grid resources (CPU resources, storage systems, network links) are spread over more than 120 computing centres distributed worldwide. ATLAS grid computing centres host over 95 PB of storage either on disk or tape systems, employing different flavours of storage systems, and heterogeneous CPU resources available to accommodate over 100k job slots. ATLAS grid sites are organized within three different flavours of grid: EGI, OSG, and NorduGrid. To provide a high quality of service to the ATLAS Collaboration, the operations of the ATLAS computing resources must be able to easily identify issues with the infrastructure, and to address these issues. Such a challenging task is addressed by the ATLAS Distributed Computing Monitoring team.

In this paper we describe how the ADC activities are monitored. In Section 2, we describe the state-of-the-art of the ATLAS Distributed Computing Monitoring as of May 2012, after the full
2 years of LHC data taking. In Section 3, we summarize the standardization efforts conducted in the past 2 years. In Section 4, we briefly outline future challenges for the monitoring.

2. ATLAS Distributed Computing Monitoring at a glance
ATLAS Distributed Computing Monitoring tools cover every aspect of the day-to-day work of ADC Operations. As the requirements on ADC evolve, the monitoring tools respond with similar evolution in order to address all operational needs. In this section we describe the state-of-the-art of monitoring of key areas of ATLAS Distributed Computing: CERN Analysis Facility, Distributed Data Management, data processing, databases, and status of sites and services hosted at sites.

2.1. CERN Analysis Facility
The raw data acquired with the ATLAS detector are recorded at a nominal rate of 200 Hz. The average uncompressed event size is 1.6 MB; after compression it is 0.8 MB. Data is stored on tape directly at the Tier-0. In 2011, ATLAS decided to increase the rate of event recording up to 400 Hz. The raw data recorded with the ATLAS detector is processed at the Tier-0 and then the raw data and the derived data are exported from the CERN Analysis Facility to Tier-1 centres and several calibration Tier-2 centres. The data processing activity is monitored with the 
conTZole dashboard [3]. In Fig 1 the view of the overall status of the data processing at Tier-0, status of the underlying batch system, and status of data registration to the ATLAS Distributed Data Management system is shown. In addition, several views detailing the data processing progress are provided. The 
conTZole dashboard is one of the many monitoring tools based on the Experiment Dashboard framework [4].

2.2. Distributed Data Management
Since the very first collisions at 900 GeV in 2010 and at 7 TeV in 2011 the ATLAS data reconstruction at Tier-0 has produced over 10 PB of raw and derived data. This data is registered to the ATLAS Distributed Data Management system (DDM) [5] and distributed to the ATLAS sites worldwide. The data distribution status and progress is monitored with the DDM dashboard. In the past 2 years the DDM dashboard UI has been significantly improved to address several requests, e.g. ability to visually distinguish between source and destination failures (source/destination matrix replaced the destination-oriented list), introduction of flexible filtering based on string patterns, and flexible grouping of DDM endpoints. The backend of the DDM dashboard had to cope with larger data volumes, and generation of the visualisation data is faster and scales better. Currently, the DDM dashboard is able to process and visualise the amount of transfer events (start/end) with a rate of 100 Hz. The average callback load is around 30 Hz. In figure 2 the source/destination matrix is shown. This matrix shows the efficiency of the transfers between each pair of groups of the storage endpoints, the throughput of such transfers and the number of successes and errors. It also offers a view to show data on the registration process: number of datasets, number of files, and number of errors. Sources and destinations are grouped. The first level of grouping is per cloud (group of geographically close Tier-2 and Tier-3 sites which are connected to the same Tier-1), followed by grouping endpoints on the level of sites. The most detailed group lists all DDM endpoints (DDM spacetokens) provided by a site. Thanks to the elegant way of grouping storage endpoints, it is very straightforward to spot a failing endpoint and to see whether that endpoint is failing to transfer data as a source, or as a destination, or as both. The DDM dashboard matrix can be adjusted in terms of time interval of the presented data, filtering sites as sources or destinations, based on filtering by cloud, tier, site name, or spacetoken, as well as filtering by one or more transfer activities.

Each cell of the transfer matrix leads to a special view with details about the transfers. For erroneous transfers, this detailed view provides groups of errors determined by the error
Figure 1. Overview of the data processing at the CERN Analysis Facility shown by the conTZole dashboard. The top-left plot shows uncompressed RAW data write rates: physics streams (blue), express stream (green), calibration streams (grey), and commissioning data (black). The top-middle and top-right plots show the evolution of the number of running and failed reconstruction jobs.

pattern, number of occurrences, and a detailed list of attempts of failed transfers with a list of transfer metadata such as file transfer placement time, file name, exit code and status of the transfer, GUID, counter of attempts, transfer requestor identifier, source and destination SURL, transfer ID, FTS channel, error message, and activity name. Such a set of detailed information is very useful to provide in a bug report, especially to clarify if the issue is at the source, at the destination or just for a pair of sites (usually a network issue).

The DDM dashboard provides not only textual information, but also a set of plots with history trends. There are sets of Efficiency, Throughput, Transfer Successes, and Transfer Failures history plots available for Sources, Destinations, and Activities. The overall throughput of the DDM system since early 2010 is shown in figure 3.

The DDM dashboard is based on the Experiment Dashboard framework. The DDM dashboard uses the xbrowse framework [6] to visualize data. Plots are made with the Highcharts library. A WLCG Transfers Dashboard used across the LHC experiments is built using the same architecture as the DDM dashboard.

If data transfer or Grid jobs indicate a missing, inaccessible or corrupted file at the source, data consistency checks are performed. The consistency monitoring consists of a set of tables listing DDM endpoints with suspicious data. For each problematic file, a list of the SURL, GUID and dataset name is provided and the status of the recovery procedure is displayed.
Figure 2. ATLAS Distributed Data Management transfer matrix. The current transfer efficiency and throughput between each group of storage endpoints is shown.

In order to increase the number of dataset replicas according to its access frequency, data popularity is monitored. Weekly reports with plots of the number of file accesses vs. the number of file replicas are then provided, as well as time series plots for dataset popularity per dataset project name or per user.

The DDM deletion activity is mainly triggered by transient data cleaning or to reduce the number of dataset replicas. There are on average 3 M files to be deleted every day. In figure 4 the deletion activity during a busy week is shown: number of files deleted – peak around 6 M deleted files per day on 2nd October 2011, and amount of data deleted – peak around 750 TB deleted on 30th September 2011.

In order to prevent ATLAS from oversubscribing data to storage endpoints ATLAS monitors how much space the site pledged to the experiment and how much space is remaining at a site (figure 5). Storage accounting reports are available to the ATLAS Collaboration.

There are three possible ways to replicate data in the ATLAS DDM systems. Firstly, the data is replicated centrally according to the pledged resources availability. Such a replication follows WLCG Memorandum of Understanding (MoU) pledges of the sites. The second possibility is a replication triggered by an active subscription of data by physics groups or by physicists to a particular grid site. The third possibility is PanDA Dynamic Data Placement (PD2P) [10], based on data popularity and availability in several places. PD2P is triggered by user jobs running on the grid. The aim of all data replication policies is to make data available to the ATLAS
physicists in a timely manner. The data transfers are monitored with the DDM dashboard. In addition, accounting reports and weekly activity reports are provided.

Functional testing of the DDM system at a marginal level is an important activity. DDM functional tests [7] are submitted regularly with a 1 week period probing availability of a storage endpoint to write and read data with files of different sizes ranging from several MB (small files), through several hundred MB (medium files), to several GB (large files). In addition to
storage sanity checks, such a functional test provides information about network throughput between each pair of disk storage endpoints used by ATLAS. In figure 6 an example of network throughput evolution between the ATLAS CERN Analysis Facility and an ATLAS Tier-1 site is presented. More precise network link properties measurement can be accomplished with the perfSONAR-PS [8] infrastructure.

2.3. Data processing
The ATLAS Experiment uses PanDA [9] as the workload management backend. ATLAS is able to run over 100k production and analysis jobs simultaneously on the grid, averaging 200k-300k jobs per day. There are three points of view for instant monitoring of data processing: site-centric, production task-centric, and user task-centric.

The site-centric view is provided by the Panda Monitor [10]. Panda Monitor provides detailed information about the workload management backend. Primarily, it provides information about jobs, and about site performance with respect to running jobs from the Production and Analysis activities.

The task-centric monitoring for production and analysis activities [11] comes with many useful summaries to monitor the progress of a task. Summaries are provided as a comprehensive table or as a plot. The comprehensive table contains basic task description information, task requestor, task progress information (duration, progress status, processing time). The summary plots describe key metrics of the task progress: evolution of job status with failures distinction, evolution of number of job attempts to better track job tails, distribution of processing time, summary of failures. Task-centric monitoring uses the Experiment Dashboard and hBrowse [12] frameworks.

The instant job monitoring tools provide all necessary information to follow up possible job failures and to file a bug report.
Figure 6. Time evolution of the network throughput between 2 DDM endpoints. Measured with large files (size of several GB).

The workload accounting is visualized in the Historical Views [13] dashboard. The Historical Views dashboard provides views which focus on accounting information about completed, submitted, pending and running jobs, CPU consumption and efficiency, processed data, success/failures, activities, and resource utilisation. The Historical Views dashboard provides a set of powerful filtering options: filter by sites (selected by Tier level, cloud, country, or simply a list of sites), by physics groups requesting a workload to be processed, by input data types (e.g. AOD, ESD, NTUP, RAW and their derivatives) and data projects (e.g. data11_7TeV, mc11_7TeV, etc.), by cloud, by activity (e.g. Data processing, Production triggered by a physics group, Analysis triggered by a physics group, Monte Carlo simulations, User analysis, Validation and Testing). The filters allow the user to define the time interval for the plots, time binning, and grouping by all the filters listed above. The plots in each of the views contain the time evolution plots as instant and cumulative evolution in time, and a pie chart with drill down to categories given by the filter. The Historical Views dashboard provides useful source of information for data mining the performance of the ATLAS grid resources, as well as a platform to interconnect information about the computing resources from different sources, e.g. the information about jobs is taken from PanDA, while information about pledged CPU resources is taken from WLCG REBUS, and the site topology information is taken from the ATLAS Grid Information System (AGIS) [14]. An example of such an interconnection of the CPU resources information is shown in figure 7: time evolution plot of the walltime (in HEPSPEC06 hours) spent on all ATLAS grid sites since late 2010 is shown, the line of amount of walltime pledged by the sites is depicted. In addition, this plot shows relative share of different activities: User analysis, Monte Carlo simulations, Group production and analysis, Data processing, Testing and Validation. Each of the bins show the monthly-averaged sum. One can easily see the steep increase in the User analysis activity share around the summer conferences of 2011.
2.4. Databases

The ATLAS Distributed Computing products store their status information and metadata in a central database, which is served by an Oracle 11g cluster hosted at CERN. The cluster is monitored by the DB Administration experts. ATLAS Distributed Computing shifters and experts can monitor the status of the databases hosted by the Oracle 11g cluster on a simple dashboard. This DB dashboard shows the status of each of the databases, distinguishing service degradation with a simple colour scheme. The DB dashboard also provides a very compact set of parameters to describe the cluster health. More detailed information can be retrieved from the Service Level Status (SLS) [15] monitoring, which is referred from the DB dashboard.

The ATLAS grid jobs need to access several databases to retrieve metadata essential for data processing. At the beginning of 2010, there was a replica of such a database at the CERN Analysis Facility and each of 10 Tier-1s. Only direct access to this database from a hosting site was possible. Subsequently ATLAS started using Frontier [16] in order to optimise data access and cache the content at the CERN Analysis Facility and every Tier-1. In addition, a squid cache is deployed at every site, so that the metadata is cached, and access to metadata at Frontiers at Tier-1 sites is further optimised.

Frontier and Squid monitoring consists of a set of common monitoring tools shared between the ATLAS and CMS experiments. Frontier monitoring consists of an SLS probe which is used mainly by the ADC shifters and experts, and the more detailed awstats information targeting the Frontier experts. Squid monitoring consists of a standardised set of MRTG plots provided for the squids at every site. An example of such monitoring is shown in figure 8.

![WallClock HEPSPEC06 Hours](image)

**Figure 7.** ATLAS grid resources utilization since late 2010. This plot shows walltime (in HEPSPEC06 hours) spent on all ATLAS sites. Data bins are aggregated by activity. Blue bars denote the Monte Carlo production, red bars User analysis, violet bars Group production, grey bars Group analysis, yellow bars Data processing, and green bars Testing activity.
2.5. Status of Sites or Services

In 2010 ATLAS started to use the Site Status Board (SSB) [17] to aggregate monitoring information from many different sources, and visualize the status of ATLAS grid sites on a single page. The ATLAS SSB [18], [19] proved to be a very useful tool providing not only the current status of the ATLAS Distributed Computing infrastructure, but also historical trends of the aggregated monitoring data. The core feature of the ATLAS SSB, status information aggregated on a single page, is used by different groups of consumers ranging from the ADC Shifters, through experts, to ATLAS Site administrators. In figure 9 an example of the Cloud view is shown. A simple colour scheme used in the ATLAS SSB views enables the user to rapidly spot problems and focus on their escalation or solution. The ATLAS SSB provides links to the original monitoring tool, where the shifter, or expert, or site administrator can get more detailed information about the metric status.

The ATLAS SSB views provide aggregated information for a variety of different metrics, such as site topology information, downtime, status in DDM, status in workload management, status in functional tests (HammerCloud [20], SAM/Nagios [22], DDM FT), exclusion status for various activities, number of critical software releases installed at a site and progress of SW installation.

The ATLAS SSB history page provides a simple history plot, a ranking plot, a history table with plot and list of all possible states, and a plot with the number of sites entering each of the states. Each plot can be configured with a powerful set of filters: time interval, group of sites (one or multiple sites, filtering sites by their Tier level or cloud), and choice of time-binning. An example plot showing ATLAS Site Usability in Production and Analysis for the CERN Analysis Facility and all the Tier-1, and Tier-2 sites since August 2011 is shown in figure 10. The ATLAS Site Status Board is one of the applications based on the Experiment Dashboard framework.

ATLAS uses the Site Usability Monitor (SUM) [21] to visualize the latest and historic results of the SAM/Nagios tests since February 2012. The ATLAS SUM provides availability and reliability plots for the critical CE and SE metrics, as a quality or as a ranking plot. These plots are provided for availability and reliability of a site, of a service registered in GOCDB [23] or OIM DB [24], and for test results. In figure 11 an example of a site reliability plot for the CERN Analysis Facility and ATLAS Tier-1 sites is shown. The visualisation part benefits from the Experiment Dashboard framework, while the data is provided by the SAM framework API.
ATLAS Distributed Computing workload management uses pilot jobs submitted to ATLAS grid sites. The pilot jobs then execute useful job payloads. Pilot job submission is handled by a pilot factory. The pilot factory can be operated centrally or by an expert local to the cloud. Each site defines several queues. The pilot factory then submits jobs to these queues. The health of each of the queues, submission, runtime, and exit status is monitored for each of the pilot factories.

In 2010 ATLAS increased the number of ATLAS grid sites available for physics analysis with the addition of the Tier-3 sites. A Tier-3 site is a site which provides non-pledged computing resources (CPU, storage, network connectivity) to the local community, but a Grid one can be used centrally by ATLAS in an opportunistic way. Such a Tier-3 site may be completely off-grid, or can be part of the ATLAS grid site family. The non-grid Tier-3 sites are monitored with the T3MON package [25], which provides a set of very useful monitoring tools for site administrators. The T3MON package monitors local computing resources at a off-grid Tier-3 site. The T3MON provides also a global aggregation of the monitoring probes.

3. Standardization

The ATLAS Distributed Computing Monitoring tools evolved over the past 5 years. Historically, there were many tools which originated at different times with different technologies available at those times and used a variety of visualisation elements. Over the past two years the monitoring tools converged in terms of data communication and data visualisation. Data is provided by a backend in the lightweight data-interchange format of JSON (JavaScript Object Notation). JSON data is then rendered by a visualisation framework based on the JavaScript library jQuery and some of its plug-ins. Examples of the visualisation frameworks used are the xbrowse framework (DDM dashboard, WLCG Transfers dashboard) and the hBrowse framework (Production and User task monitoring). Examples of plotting libraries used are Highcharts (Production and User task monitoring), graphtool (Historical Views dashboard), flot (Panda monitor), Google charts (PD2P monitoring). The ATLAS Distributed Computing Monitoring
tools converged also in terms of visual identity. All the monitoring tools follow common colour schemes based on the sRGB colour space.

4. Future challenges
Identification of common solutions in the monitoring field can help us to save manpower needed to develop and maintain the monitoring tools. Another concern is and will be the scalability of a solution. In order to decrease the load on production databases we are going in two distinct ways: distribution of load on multiple clusters, and moving from relational DBs to NoSQL DB solutions. One of the major concerns for the future is not only to be able to monitor every

Figure 10. ATLAS Site Status Board – sum of the number of the ATLAS Tier-1, and Tier-2 sites (including the CERN Analysis Facility) available for Production activity (left) and for Analysis activity (right).
possible activity of the ATLAS Distributed Computing, but also to be able to preserve the monitoring data or their summaries on a long-term scale.

5. Conclusions
The ATLAS Experiment has already been collecting data for more than 2 years. The ATLAS data is processed and analysed at more than 120 grid sites distributed worldwide. Such a complex system is continuously monitored. The ATLAS Distributed Computing successfully fulfils its mission to deliver data to the ATLAS physicists. The available monitoring tools help to identify possible issues in a timely manner, as well as mine long-term data series, or monitor and effectively utilize available computational resources. The ATLAS Distributed Computing Monitoring tools provide a comprehensive way to monitor the infrastructure, identify and address issues, and to improve automation of repetitive tasks. The ADC Monitoring tools have to maintain a high level of quality, and be flexible enough to support the high-energy physics community by adapting to the ever-evolving environment. The future challenges to face come with advances in data storage technologies, and the need to preserve data. There are many challenging years ahead of us.

Acknowledgments
Jaroslava Schovancová gratefully appreciates support from the Academy of Sciences of the Czech Republic. Support from the grant LA08032 of the MEYS (MŠMT), Czech Republic, and from ACEOLE: A Marie Curie ITN Project is greatly acknowledged. The work was supported by the grant SVV-2012-265309 of the Charles University in Prague.
References
[1] ATLAS Collaboration, 2008 JINST 3 S08003
[2] Jézéquel S et al. for the ATLAS Collaboration, “ATLAS Distributed Computing Operations: Experience and improvements after 2 full years of data-taking”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[3] Ueda I et al. for the ATLAS Collaboration 2011 J. Phys.: Conf. Series 331 072034
[4] J. Andreeva et al. 2010 J. Phys.: Conf. Series 219 062003
[5] M. Branco et al. 2008 J. Phys.: Conf. Series 119 062017
[6] D. Tuckett et al., “Designing and developing portable large-scale JavaScript web applications within the Experiment Dashboard framework”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[7] S. Campana et al. for the ATLAS Collaboration, “Evolving ATLAS computing for today’s networks”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[8] P. Laurens et al., “Monitoring the US ATLAS Network Infrastructure with perfSONAR-PS”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[9] T. Maeno for the ATLAS Collaboration 2008 J. Phys.: Conf. Series 119 062036
[10] T. Maeno et al. for the ATLAS Collaboration, “Evolution of ATLAS PanDA System”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[11] L. Sargsyan et al. for the ATLAS Collaboration, “ATLAS job monitoring in the Dashboard Framework”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[12] L. Kokoszkiewicz et al., “hBrowse - Generic framework for hierarchical data visualization”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[13] E. Karavakis et al., “User-centric monitoring of the analysis and production activities within the ATLAS and CMS Virtual Organisations using the Experiment Dashboard system”, to appear in Proceedings of the EGI Community Forum 2012 / EMI Second Technical Conference, 2012, PoS(EGICF12-EMITC2) 110
[14] A. Anisenkov et al. for the ATLAS Collaboration, “AGIS: The ATLAS Grid Information System”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[15] S. Lopienski 2008 J. Phys.: Conf. Series 119 052025
[16] A. Dewhurst et al. for the ATLAS Collaboration, “Evolution of grid-wide access to database resident information in ATLAS using Frontier”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[17] P. Saiz et al., “Collaborative development. Case study of the development of flexible monitoring applications”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[18] C. Borrego et al. for the ATLAS Collaboration, “Aggregated monitoring and automatic site exclusion of the ATLAS computing activities: the ATLAS Site Status Board”, appeared in Proceedings of the 5th Iberian Grid Infrastructure Conference (IBERGRID), 8-10 June 2011, Santander, Spain, ISBN 978-84-9745-884-9
[19] E. Magradze et al. for the ATLAS Collaboration, “Automating ATLAS Computing Operations using the Site Status Board”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[20] D. C. van der Ster et al. 2011 J. Phys.: Conf. Series 331 072036
[21] A. Di Girolamo et al., “New solutions for large scale functional tests in the WLCG
infrastructure with SAM/Nagios: the experiments experience”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference

[22] P. M. Rodrigues De Sousa Andrade et al., “Service Availability Monitoring framework based on commodity software”, to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference

[23] J. Gordon et al. 2010 J. Phys.: Conf. Series 219 062021

[24] R. Pordes et al. 2008 J. Phys.: Conf. Series 125 012070

[25] J. Andreeva et al., “Tier-3 Monitoring Software Suite (T3MON) proposal”, ATL-SOFT-PUB-2011-001, 2011, https://cdsweb.cern.ch/record/1336119