ATL-SOFT-PROC-2012-067
29 September 2012

ATLAS DISTRIBUTED COMPUTING AUTOMATION

J. Schovancová[1,*], F. H. Barreiro Megino[2], C. Borrego[3], S. Campana[2], A. Di Girolamo[2], J. Elmsheuser[4], J. Hejbal[1,5], T. Kouba[1], F. Legger[4], E. Magradze[6], R. Medrano Llamas[2], G. Negri[2], L. Rinaldi[7], G. Sciacca[8], C. Serfon[4], D. C. Van Der Ster[2]
on behalf of the ATLAS Collaboration

[1] *Institute of Physics, Academy of Sciences of the Czech Republic, Na Slovance 2, CZ-18221 Prague 8, Czech Republic*
(*) Email: schovan@fzu.cz
[2] *CERN, CH - 1211 Geneva 23, Switzerland*
[3] *Universidad Autonoma de Madrid, Madrid*
[4] *Ludwig-Maximilians-Universitat Muenchen, Fakultaet fuer Physik, Am Coulombwall 1, DE - 85748, Garching, Germany*
[5] *Czech Technical University in Prague, Fac. of Nuclear Sciences and Physical Engineering, Brehova 7, CZ-11519 Prague 1, Czech Republic*
[6] *II. Physikalisches Institut, Georg-August Universitaet Goettingen, Friedrich-Hund-Platz 1, D-37077, Goettingen, Germany*
[7] *Instituto Nazionale Fisica Nucleare, Viale Berti Pichat 6/2, Bologna I-40127, Italy*
[8] *Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern*

Abstract The ATLAS Experiment benefits from computing resources distributed worldwide at more than 100 WLCG sites. The ATLAS Grid sites provide over 100k CPU job slots, over 100 PB of storage space on disk or tape. Monitoring of status of such a complex infrastructure is essential. The ATLAS Grid infrastructure is monitored 24/7 by two teams of shifters distributed world-wide, by the ATLAS Distributed Computing experts, and by site administrators. In this paper we summarize automation efforts performed within the ATLAS Distributed Computing team in order to reduce manpower costs and improve the reliability of the system. Different aspects of the automation process are described: from the ATLAS Grid site topology provided by the ATLAS Grid Information System, via automatic site testing by the HammerCloud, to automatic exclusion from production or analysis activities.

1 Introduction

The Large Hadron Collider (LHC) at CERN has been delivering stable beams colliding at the centre-of-mass-energy 7TeV since March 2010 and at the center-of-mass-energy 8TeV since April 2012. ATLAS Experiment [1], one of the general purpose detectors of the LHC, has accumulated over 4 PB of RAW data over past 2 years. ATLAS benefits of the World-wide LHC Computing Grid (WLCG Grid) to process data and simulations.

The ATLAS Distributed Computing [2] (ADC) infrastructure is a complex and heterogeneous 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 their storage either on disk or tape systems, with 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 good quality of service to the ATLAS Collaboration, the operations team of the ATLAS computing resources has to be able to easily identify issues with the infrastructure, and to address these issues. Challenging task to address requests for the monitoring of the ADC infrastructure is addressed by the ADC Monitoring team [3]. Even more challenging task to monitor the ADC infrastructure is covered by the ADC Operations teams: various ADC Shift teams [4], ADC Experts, and site administrators.

ATLAS sites may or may not be part of 3 ATLAS Activities: Data transfers, Data processing, Distributed analysis.

In Section 2 we describe motivation for automation efforts within ADC team. In Section 3 we briefly describe benefits of the ATLAS Grid Information System. In Section 4 we detail on what functional tests are available for automation.

2 Repetitive tasks & need for automation

There are 3 Shift teams monitoring status of the ADC infrastructure, each team focusing on different aspects of ADC Activities. The Shift teams are backed up by 2 ADC Experts. Site issues are addressed by site administrators.

Whenever a shifter on duty identifies an issue with the ADC infrastructure, he/she creates a bug report to expert, or activity requester, or to the site. There were over 6700 GGUS tickets created to the ATLAS grid sites since 1st January 2010, which leads to an average rate of 7 tickets per day. This amount of bug reports represents huge manual effort carried out by the ADC Operations team, ranging from the issue investigation by the ADC Shifter, creation of the bug report, addressing the issue by the site administrator or activity requester, resulting in functional testing of the reported service, and putting the service back into production for ATLAS Activities. The amount of manual work is the main motivation for automation of well known issues.

3 ATLAS Grid Information System

The ATLAS Grid Information System [5] (AGIS) collects site information from the GOCDB [6] and the OIM [7], and exposes it in a way convenient to the experiment. The AGIS provides topology information about the ATLAS grid sites, about services at sites, about downtimes of those services.

This unique information collection available in AGIS enables the ATLAS experiment to map between physical resources (CEs, SEs, LFCs, etc.) and ATLAS activity endpoints (PanDA [8] queues workload management endpoint, DDM [9] spacetoken endpoints), and additional logical layer in AGIS provides availability information of an ATLAS Activity at a particular site based on availability of subsequent physical resources at that site.

Having written what useful set of information AGIS provides, ATLAS benefits from several collectors, which collect downtime information for ATLAS activity endpoints, and exclude those activity endpoints for downtime period from corresponding ATLAS Activities.

First example of such a collector is the DDM collector, which excludes DDM spacetoken(s) from Data transfer activity, with granularity of sub-activities such as write/read/deletion, when a downtime of underlying SE starts, and re-enables those DDM spacetokens for the sub-activities once the SE downtime is over.

Second example of a collector taking action when a service is on downtime, is the Switcher. The Switcher manipulates Panda queues when a downtime of a CE or a SE affects Data processing or Distributed analysis activity at an ATLAS grid site.

Both collectors, DDM collector and Switcher, take automatic action ca 30 times per week. The main benefit of both collectors is saving ADC Operations manpower when a site declares unscheduled downtime, secondary benefit is for scheduled downtime.

Third example of a collector is the DDM space collector, which based on DDM spacetoken occupancy excludes a DDM spacetoken for write when a very small fraction of its size (several TBs) is left. When a fraction of free space at that DDM spacetoken is cleaned, at least up to limit which enables uninterrupted ATLAS Activities at that site, DDM spacetoken is enabled for writing again.

4 Functional tests for Services and Activities

ATLAS experiment runs a continuous flow of functional tests at each site. The tests are marginal with respect to normal ATLAS Activity at a site; Fraction of functional tests with respect to the overall activity is of the order of percent.

ATLAS experiment uses the HammerCloud [10], [11] framework to test how a site performs in the Data processing and Distributed analysis Activities. The HammerCloud test jobs simulate behaviour of an usual ATLAS data processing or analysis job. The HammerCloud uses the same environment as usual ATLAS jobs, access input data and installed SW in the same way, and stages out the output data in the same way. The HammerCloud then provides a very useful probe in the site health for real Activities. When several HammerCloud tests fail, site is excluded from an Activity for period of time, and recovered for that Activity once a set of jobs in a row succeeds. The HammerCloud takes ca 240 exclusion/recovery actions per week. The HammerCloud framework is used as the recovery framework for the Switcher exclusions.

ATLAS experiment probes NxN endpoint-to-endpoint transfers functionality with the Sonar [12] test. Purpose of this testing is to find optimal path for the transfers. In the past, ATLAS used strictly hierarchical topology of DDM endpoints. ATLAS sites are grouped in 10 clouds, each cloud is a set of geographically-close grid sites. The most powerful site in each cloud is a Tier-1 site. There are usually several Tier-2 sites in each cloud. Cloud may host also Tier-3 grid sites (sites with no pledge to WLCG).

In the past transfer between 2 Tier-2 sites, which belong to different clouds, was possible only through 2 Tier-1 sites, transfer path then was T2(Cloud A) → Tier-1 (Cloud A) → Tier-1 (Cloud B) → Tier-2 (Cloud B). This transfer path may not be very optimal, due to 3 additional sites being filled with data on the way.

Currently, ATLAS relaxes a bit the strictly hierarchical tier mode, and direct transfers between Tier-2 sites from different clouds with a very good network connectivity are enabled.

About 20 ATLAS grid sites are taking part in the LHCONE [13] network project. Such sites are running perfSonar [14] tests. As of September 2012 there is no automatic action taken based on perfSonar test results.

ATLAS uses the WLCG SAM framework [15] to test resources [16] registered in GOCDB or OIM. Currently, the SAM test results are used as additional sanity check when manual recovery of a service is necessary.

Conclusion

The ATLAS experiment has been successfully collecting data for more than 2 years. ATLAS data is processed and analysed at more than 120 grid sites distributed worldwide, taking into account Tier-1s, Tier-2s, and Tier-3 sites. The ATLAS Distributed Computing successfully fulfils its mission to deliver data to the ATLAS physicists. Current monitoring tools enable the ADC Operations team address issues in a timely manner. Level of automation of the ADC Operations helps to save manpower, and to focus on more urgent issues first.

Acknowledgements

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 the grant SVV-2012-265309 of the Charles University in Prague is greatly acknowledged. This work was partly supported by the JINR and FZÚ AS CR Common Project The GRID infrastructure for the physics experiments.

References

[1] The ATLAS Collaboration, "*The ATLAS Experiment at the CERN Large Hadron Collider*," *JINST* **3** (2008) 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] Schovancová J. for the ATLAS Collaboration, "*ATLAS Distributed Computing Monitoring tools after full 2 years of LHC data taking*", to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference

[4] Schovancová J. et al. for the ATLAS Collaboration, "*ATLAS Distributed Computing Shift Operation in the first 2 full years of LHC data taking*", to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference

[5] Anisenkov A. 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

[6] J. Gordon *et al.*, "*GOCDB, A Topology Repository For A Worldwide Grid Infrastructure*", *J. Phys. Conf. Series* **219** (2010) 062021

[7] R. Pordes *et al.*, "*New science on the Open Science Grid*", *J. Phys. Conf. Ser.* **125** (2008) 012070

[8] Maeno T. for the ATLAS Collaboration, "*PanDA: Distributed production and distributed analysis system for ATLAS*", *J. Phys. Conf. Ser.* **119** (2008) 062036

[9] Branco M. et al. for the ATLAS Collaboration, "*Managing ATLAS data on a petabyte-scale with DQ2*", *J. Phys. Conf. Ser.* **119** (2008) 062017

[10] Van Der Ster D. C. et al. for the ATLAS Collaboration, "*Experience in Grid Site Testing for ATLAS, CMS and LHCb with HammerCloud*", to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference

[11] Legger F. for the ATLAS Collaboration, "*Improving ATLAS grid site reliability with functional tests using HammerCloud*", to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference

[12] Campana S. 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
[13] Fisk I, "*New computing models and LHCONE*", to appear in Proceedings of the Computing in High Energy and Nuclear Physics 2012 International Conference
[14] Laurens P. 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
[15] Rodrigues De Sousa Andrade P. M. 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
[16] Di Girolamo A. 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