gLExec Integration with the ATLAS PanDA Workload Management System

E Karavakis[1], F Barreiro[2], S Campana[1], K De[2], A Di Girolamo[1], M Litmaath[1], T Maeno[3], R Medrano[1], P Nilsson[3] and T Wenaus[3] for the ATLAS Collaboration

[1]CERN, European Organization for Nuclear Research, Switzerland

[2]University of Texas at Arlington, USA

[3]Brookhaven National Laboratory, USA

Email: Edward.Karavakis@cern.ch

Abstract. ATLAS user jobs are executed on Worker Nodes (WNs) by pilots sent to sites by pilot factories. This paradigm serves to allow a high job reliability and although it has clear advantages, such as making the working environment homogeneous, the approach presents security and traceability challenges. To address these challenges, gLExec can be used to let the payloads for each user be executed under a different UNIX user id that uniquely identifies the ATLAS user. This paper describes the recent improvements and evolution of the security model within the ATLAS PanDA system, including improvements in the PanDA pilot, in the PanDA server and their integration with MyProxy, a credential caching system that entitles a person or a service to act in the name of the issuer of the credential. Finally, it presents results from ATLAS user jobs running with gLExec and describes the deployment campaign within ATLAS.

ATL-SOFT-PROC-2015-011
10 May 2015

1. Introduction

The ATLAS experiment [1] at the Large Hadron Collider [2] has collected data during Run 1 and is ready to collect data in Run 2. The ATLAS data are distributed, processed and analysed at more than 130 grid and cloud sites across the world within the Worldwide LHC Computing Grid (WLCG) [3]. At any given time, there are typically more than 150,000 concurrent jobs running and about a million jobs are submitted on a daily basis on behalf of thousands of physicists within the ATLAS collaboration. The Production and Distributed Analysis (PanDA) [4] workload management system has proved to be a key component of ATLAS and plays a crucial role in the success of the large-scale distributed computing as it is the sole system for distributed processing of Grid jobs across the collaboration since October 2007.

The ATLAS user jobs are executed on Worker Nodes (WNs) by pilots sent to the sites by pilot factories. This pilot architecture has greatly improved job reliability and although it has clear advantages, such as making the working environment homogeneous by hiding any potential heterogeneities, the approach presents security and traceability issues [5] distinct from standard batch jobs for which the submitter is also the payload owner. Jobs initially inherit the identity of the pilot

submitter, typically a robot certificate with very limited rights. By default the payload jobs then execute directly under that same identity on a Worker Node. This exposes the pilot environment to the payload, requiring any pilot 'secrets' such as the proxy to be hidden; it constrains the rights and identity of the user job to be identical to the pilot; and it requires sites to take extra measures to achieve user traceability and user job isolation [6].

To address these security risks, the gLExec [7] tool and framework can be used to let the payloads for each user be executed under a different UNIX user identity that uniquely identifies the ATLAS user. gLExec is an open source application with a super-user privileged executable with the capability of modifying the UID and GID to provide a mapping between the Grid user and the local UNIX user accounts. This mapping is performed based on the LCAS and LCMAPS [8][9] security components.

2. Architecture

This section describes in detail the evolution of the security architecture within the ATLAS PanDA system (figure 1), including improvements in the PanDA Pilot, in the PanDA Server and their integration with MyProxy [10], a credential caching system that entitles a person or a service to act in the name of the issuer of the credential.

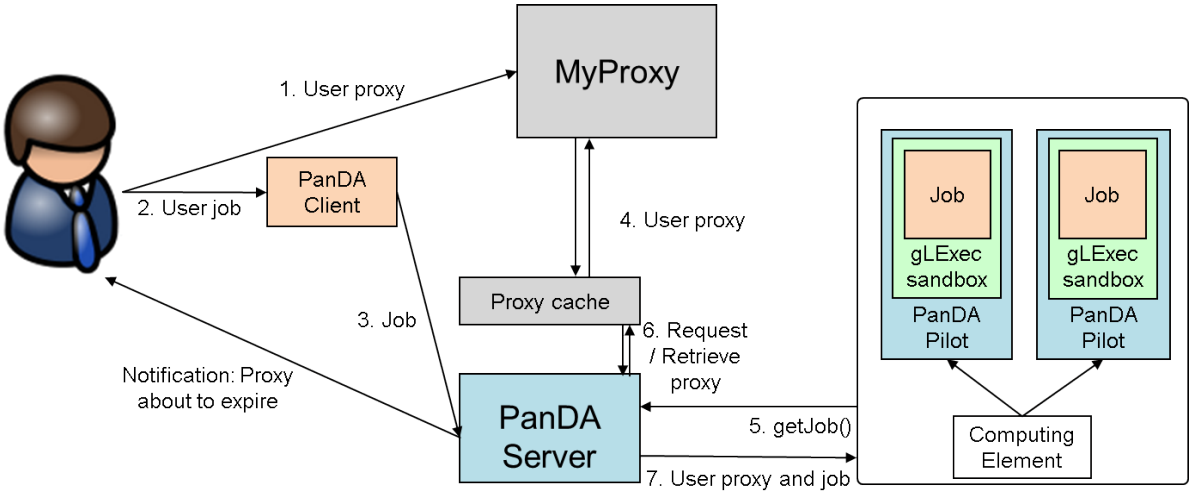

Figure 1. Architecture of the gLExec integration within the ATLAS PanDA Workload Management System.

2.1. Proxy Delegation

A delegated X.509 [11] proxy credential entitles a person or a service to act in the name of the issuer of the credential. The ATLAS users will be asked to upload a long-lived proxy on a MyProxy service every time they get a new user certificate. The PanDA host certificate is authorised to retrieve delegated proxies from that MyProxy service. The PanDA host also needs to provide a secret key when retrieving user proxies from MyProxy. This limits potential security issues from misuse of a compromised PanDA host certificate.

2.2. Proxy Cache

It provides an API to the PanDA Server to retrieve proxies and it is a simple mechanism that reduces the load on MyProxy. On the PanDA Server, a cronjob running every ten minutes gets the list of active ATLAS analysis users and then uses the API of the proxy caching mechanism to retrieve user proxies.

For every single active user, it checks the proxy cache for the existence of a proxy. If the proxy is there and valid for three or more days, that proxy will be used when required. In all other cases it tries

to retrieve a plain proxy from MyProxy and add ATLAS VOMS [12] attributes with a validity of four days.

The proxies are stored in the file system of the PanDA Server and their names are hashed with the SHA1 [13] algorithm.

2.3. PanDA Pilot

The architecture of the PanDA Pilot can be seen in figure 2. A functional call to get a job from the PanDA Server returns the proxy along with the job if the gLExec flag is set for that site in the ATLAS Grid Information System (AGIS) [14]. Hence, the user proxy is shipped with the Grid job/payload to the WN.

If the gLExec flag is set, the jobs will run under a temporary unique gLExec sandbox. If the flag is not set, the pilot will work in normal mode without using gLExec. The pilot does a quick "ping" test of the gLExec infrastructure, with a retry mechanism. If the result is not OK, it falls back to the normal mode. This feature allows the gLExec infrastructure of a site to be tested in production without having jobs fail. If the user's proxy was not provided, the pilot can still run in gLExec mode by using the proxy of the pilot itself. This ensures that not all users have to upload their proxy before the site can be used with gLExec. Finally, the pilot switches back to the original environment at the completion of the job and deletes the retrieved credentials from the local disk of the WN.

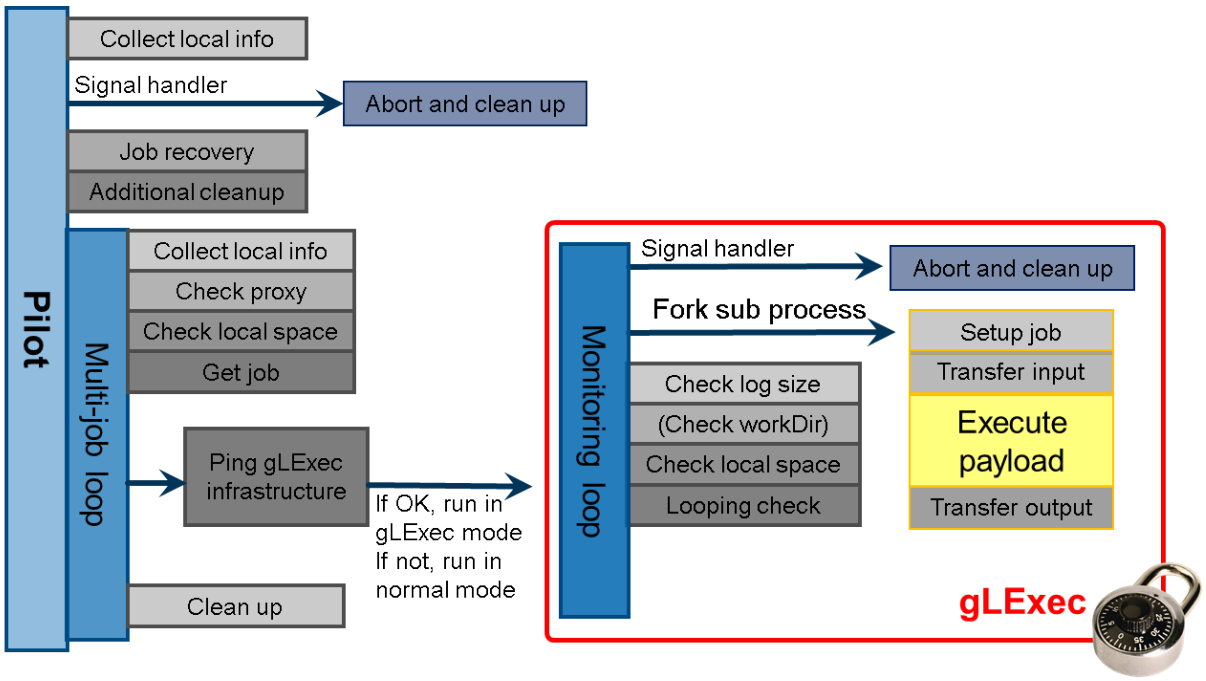

Figure 2. The integration of gLExec in the PanDA Pilot.

3. Deployment

In order to create a new gLExec queue or to change an existing one to gLExec, an authorised ATLAS expert needs to change the gLExec flag in AGIS from 'False' to 'True' or 'test' (figure 3). It was clear from the start that a quick and easy way of creating and modifying queues for gLExec was needed in order to boost the deployment of gLExec within ATLAS.

Setting the flag to 'test' will force the PanDA Pilot to hard-fail in case of gLExec related errors. This mode is used mostly for debugging and testing proposes.

By setting the flag to 'True', the PanDA Pilot will first test the gLExec infrastructure and in case of an error, it will switch back to the normal running mode without utilising gLExec. This ensures that normal ATLAS activities will not be affected due to any temporary glitches of the gLExec infrastructure at a given site.

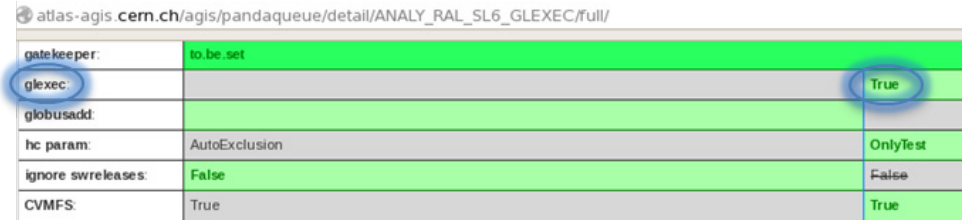

Figure 3. Setting up a gLExec queue within AGIS.

The deployment campaign started mid-September 2014 with CERN being the first site where the initial tests were performed. Slowly more sites were added, first covering the Tier 1 sites and then by adding Tier 2 and Tier 3 sites. As of early April 2015 gLExec was enabled for special queues in 61 analysis sites out of 92 sites with an online analysis queue (according to AGIS as of 1st of April 2015). The remaining 31 sites did not have a gLExec infrastructure yet. ATLAS user jobs running in gLExec mode have been thoroughly tested in 10 Tier 0+1 sites, 49 Tier 2 sites and 2 Tier 3 sites and most of them worked "out of the box". 18 GGUS tickets were opened against sites because of issues with their gLExec infrastructure or configuration. The most common gLExec deployment issues encountered were the following:

- Misconfiguration of the ARGUS instance [15] at a site. ARGUS is a centralised authorisation service for distributed services based on policies.
- Problems related to the gLExec-specific infrastructure at a given site:
    - Wrong glexec.conf configuration file on the WNs, either not having the account running the pilot whitelisted and thus, not allowing the pilot to call gLExec, or not allowing the pilot to copy the user proxy from the working directory to the gLExec user sandbox.
    - Missing gLExec utility scripts (mkgltempdir) or wrapper/unwrapper scripts on the WNs.
    - Outdated and buggy version of mkgltempdir installed on the WNs.
- Wrong Access Control Lists (ACLs) on the ATLASSCRATCHDISK "spacetoken"; only the pilot was allowed to store the output of a job into the ATLASSCRATCHDISK spacetoken although, in principle, any ATLAS user should be able to write into it.

The Service Availability Monitoring (SAM3) Experiment Dashboard application [16] was extensively used in order to check if the gLExec infrastructure was working on the ATLAS sites where it was installed.

Sites are constantly tested with HammerCloud [17] functional and stress tests and are monitored within the Experiment Dashboard [18] ATLAS Job Accounting system [19]. The status of the deployment campaign can be seen in figure 4. At any given time throughout the deployment campaign, there were on average 172 gLExec-related test jobs running with a peak of 972 concurrent running jobs on the 11th of February (figure 5). The steady load in figure 5 comes from the HammerCloud functional tests whereas the peaks come from the HammerCloud stress tests.

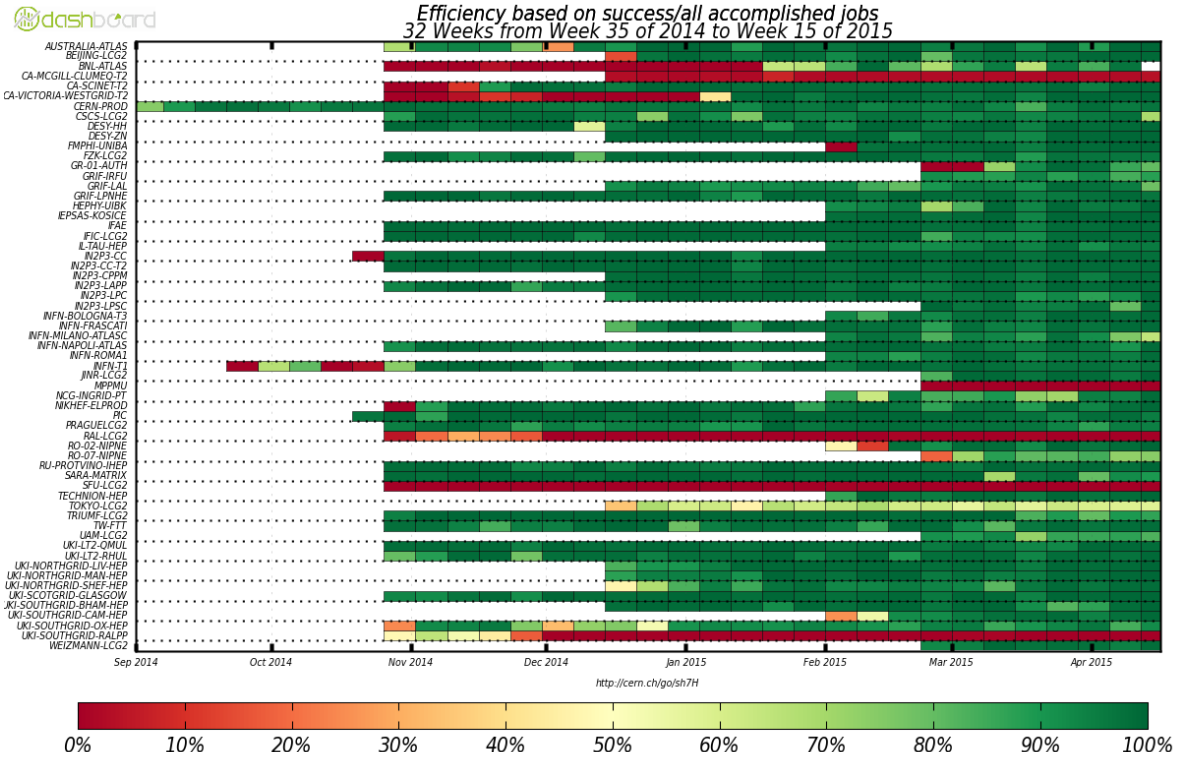

Figure 4. Deployment status of gLExec within ATLAS. Please note that the errors are not only due to the gLExec infrastructure at a site but also due to user or other site-related problems.

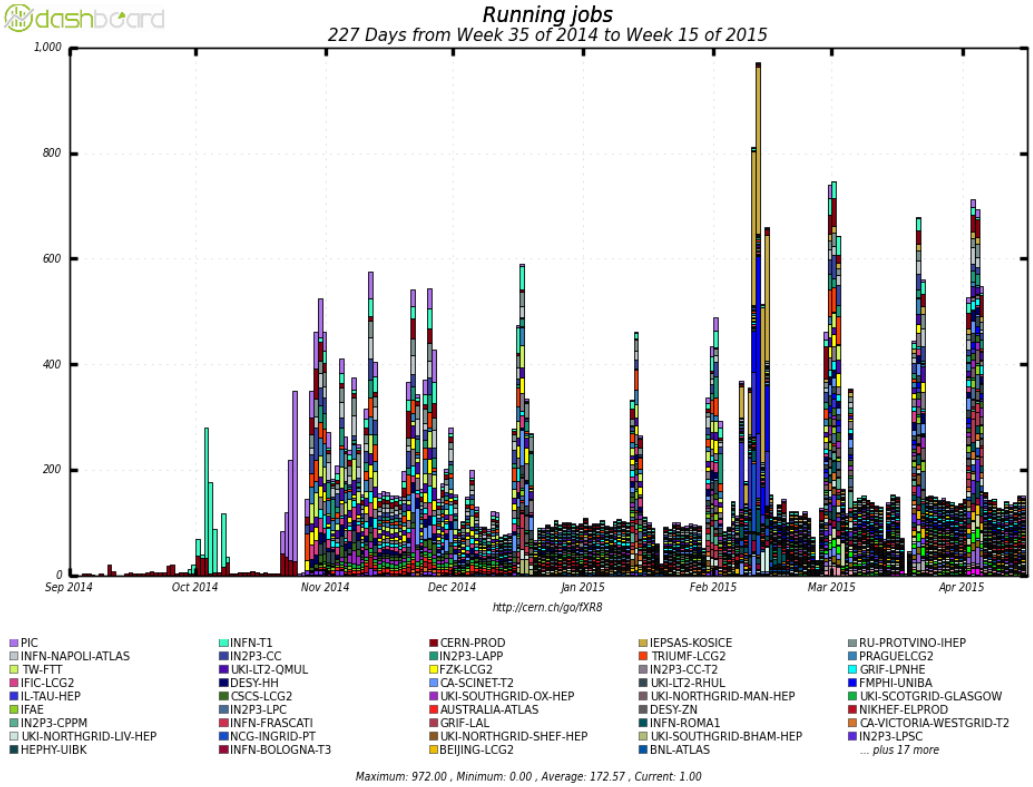

Figure 5. Daily average concurrent gLExec running jobs throughout the deployment period.

Up to now, there were more than 1.5 million successful jobs with a success rate of more than 88% (figure 6). Not all job failures were due to the gLExec infrastructure of the ATLAS sites as most of the failures were either due to user or other site-related problems or due to a temporary bug in our implementation. Also throughout our deployment campaign, the gLExec flag in AGIS was set to 'test' instead of 'True' for all the sites as we wanted the jobs to hard-fail for debugging purposes. Around 15% of these errors were due to problems related to the gLExec infrastructure, taking also into account the sites that still have minor issues with their gLExec setup.

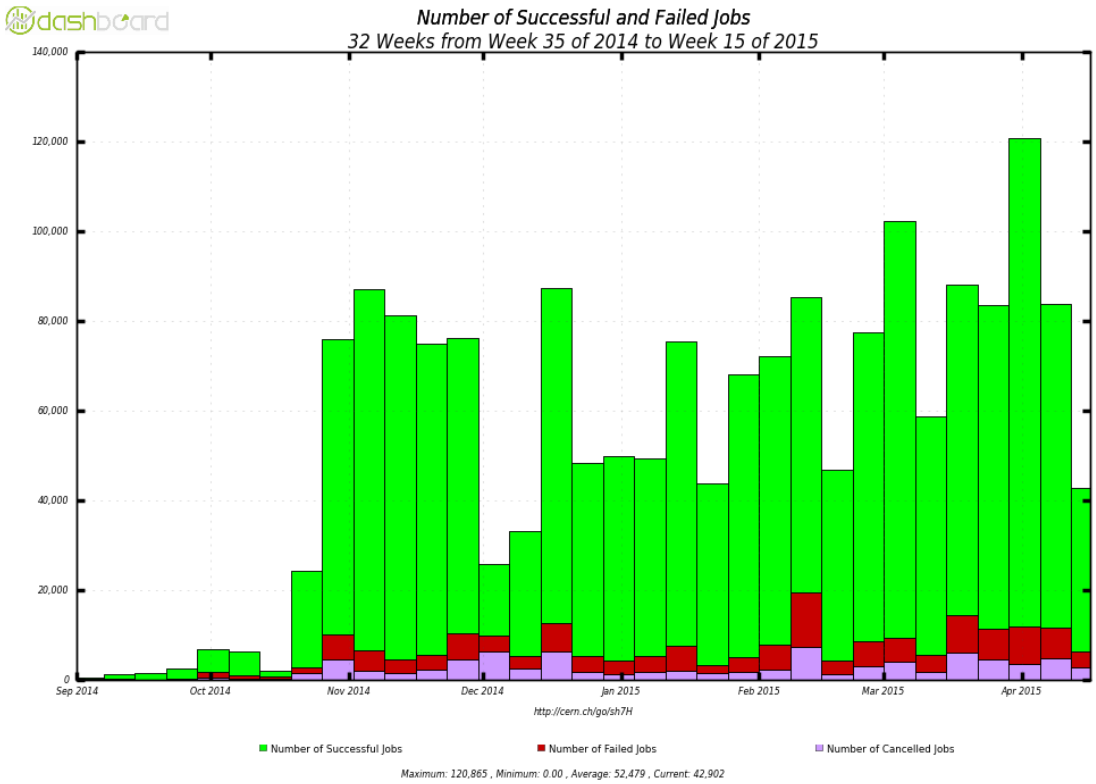

Figure 6. Number of successful, failed and cancelled ATLAS analysis jobs using gLExec during the deployment campaign.

4. Conclusions

The approach of pilot jobs executing Grid jobs directly under the same local user account on a WN has security consequences, in particular with respect to traceability. A pilot job running under the same user account as its Grid jobs cannot be protected from tampering or other illegitimate access. Similarly, it is not possible to protect and isolate different Grid jobs from each other in that case. Our development effort enables an identity switch once a pilot job starts the execution of a Grid job. The local identity switch would normally follow the Grid job submitter's identity after an authentication and authorisation step, while the implementation still allows user jobs to succeed when the gLExec infrastructure turns out to be defective at a given site.

The development work and the deployment campaign have been a big success as more than half of the ATLAS analysis sites were covered without affecting normal ATLAS activities thanks to the flexibility of the implemented system. This is an ongoing effort that will be handed over to the ATLAS Distributed Computing (ADC) Operations eventually. Some minor issues with specific sites are under investigation.

References

[1] ATLAS Collaboration, "*The ATLAS Experiment at the CERN Large Hadron Collider*", 2008,

JINST **3** S08003 doi:10.1088/1748-0221/3/08/S08003
[2] Evans L and Bryant P, "*LHC Machine*", 2008, *JINST* **3** S08001 doi:10.1088/1748-0221/3/08/S08001
[3] Bird I, "*Computing for the Large Hadron Collider*", 2011, *Annual Review of Nuclear and Particle Science*, **61**:99–118 doi:10.1146/annurev-nucl-102010-130059
[4] De K et al., "*The Future of PanDA in ATLAS Distributed Computing*", 2015, Proceedings of the 21st International Conference on Computing in High Energy and Nuclear Physics CHEP2015, J. Phys.: Conf. Ser.
[5] Sfiligoi I et al., "*Addressing the Pilot security problem with gLExec*", 2008, J. Phys.: Conf. Ser. **119** 052029 doi:10.1088/1742-6596/119/5/052029
[6] Bagnasco S et al., "*WLCG Grid job and Worker Node security assessment*", 2012, https://twiki.cern.ch/twiki/pub/LCG/AAIOnTheWorkerNodes/WLCG_WN_Security-05.pdf retrieved 2015-03-31
[7] Groep D, Koeroo O and Venekamp G, "*gLExec: gluing grid computing to the Unix world*", 2008, J. Phys.: Conf. Ser. **119** 062032 doi:10.1088/1742-6596/119/6/062032
[8] Alfieri R et al., "*Managing Dynamic User Communities in a Grid of Autonomous Resources*", 2003, in Proceedings of the Computing in High Energy and Nuclear Physics conference, 24-28 La Jolla, California, USA (TUBT005, ePrint cs.DC/0306004)
[9] Röblitz T et al., "*Autonomic Management of Large Clusters and Their Integration into the Grid*", 2004, J. Grid Computing **2** 247260
[10] Basney J, Humphrey M and Welch V, "*The MyProxy online credential repository*", 2005, Software: Practice and Experience, **35**(9), 801-816
[11] Tuecke S, Welch V, Engert D, Pearlman L and Thompson M, "*Internet X.509 Public Key Infrastructure (PKI) Proxy Certificate Profile*", RFC 3820, https://tools.ietf.org/html/rfc3820
[12] Alfieri R et al., "*From gridmap-file to VOMS: managing authorization in a Grid environment*", 2005, Future Generation Comput. Syst. **21** (4) 549-558
[13] Eastlake D and Jones P, 2001, "*US Secure Hash Algorithm 1 (SHA1)*", RFC 3174, https://tools.ietf.org/html/rfc3174
[14] Anisenkov A, Di Girolamo A, Alandes M and Karavakis E, "*AGIS: Evolution of Distributed Computing information system for ATLAS*", 2015, Proceedings of the 21st International Conference on Computing in High Energy and Nuclear Physics CHEP2015, J. Phys.: Conf. Ser.
[15] ARGUS Authorization Service, https://twiki.cern.ch/twiki/bin/view/EMI/Argus retrieved 2015-04-01
[16] Saiz P et al., "*WLCG Monitoring, consolidation and further evolution*", 2015, Proceedings of the 21st International Conference on Computing in High Energy and Nuclear Physics CHEP2015, J. Phys.: Conf. Ser.
[17] Boehler M et al., "*Improved ATLAS HammerCloud Monitoring for local Site Administration*", 2015, Proceedings of the 21st International Conference on Computing in High Energy and Nuclear Physics CHEP2015, J. Phys.: Conf. Ser.
[18] Andreeva J et al., "*Experiment Dashboard - a generic, scalable solution for monitoring of the LHC computing activities, distributed sites and services*", 2012, J. Phys.: Conf. Ser. **396** 032093 doi:10.1088/1742-6596/396/3/032093
[19] Karavakis E et al., "*Common Accounting System for Monitoring the ATLAS Distributed Computing Resources*", 2014, J. Phys.: Conf. Ser. **513** 062024 doi:10.1088/1742-6596/513/6/062024