ATLAS job submission system for Salomon HPC based on ARC-CE

Michal Svatoš(✉)[1][0000−0002−7199−3383], Jiří Chudoba[2][0000−0002−6425−2579], and Petr Vokáč[3][0000−0002−3429−4778]

[1] Institute of Physics of the CAS, Na Slovance 1999/2, Prague, 18221, Czech Republic Michal.Svatos@cern.ch

[2] Institute of Physics of the CAS, Na Slovance 1999/2, Prague, 18221, Czech Republic Jiri.Chudoba@cern.ch

[3] Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Brehova 7, Prague, 115 19, Czech Republic petr.vokac@cern.ch

Abstract. The ATLAS experiment at CERN is using HPCs opportunistically to extend its computing capacity for years. To the Salomon HPC, ATLAS jobs come via ARC-CE machines located in the computing center of the Institute of Physics of the Czech Academy of Sciences. The ARC-CE serves as an interface between job management systems of the ATLAS and the HPC. Commands of the PBSpro batch system are submitted via ssh. Scripts and input files are shared between the ARC-CE and shared file system located at the HPC via sshfs. There are several aspects of interaction between ARC-CE machines and Salomon's batch system which are important for performance of the whole system. First, the allowed number of requests to PBSpro is limited and the ARC-CE needed to be adapted to this fact. Second, the sshfs connection speed seems to be a limiting factor for job turnaround. Some possibilities of sshfs parameter tuning were investigated. Moreover, monitoring allows quick detection of issues and therefore helps the performance of the system. The ARC-CE based job submission system has adapted to conditions of the Salomon HPC and utilizes successfully its resources.

Keywords: HPC · computing.

ATL-SOFT-PROC-2020-001
27 January 2020

1 Introduction

In recent years, physics experiments require more and more computing power to perform research. There are various resources which can be utilized if suitable job submission system is created and maintained.

1.1 The ATLAS Experiment

The ATLAS experiment [6] is located at the Large Hadron Collider (LHC) at CERN near Geneva. The detector that the experiment uses is cylindrical, 44m long, 25m in diameter, weighing 7000 tonnes (see Figure 1). The physics programme studies elementary particle properties by observing products of the high

energy particle beams collisions.

The collaboration comprises about 3000 scientific authors from 183 institutions, representing 38 countries. One of its greatest achievements was the discovery of Higgs boson. The Nobel price in 2013 was awarded to theorists who predicted it.

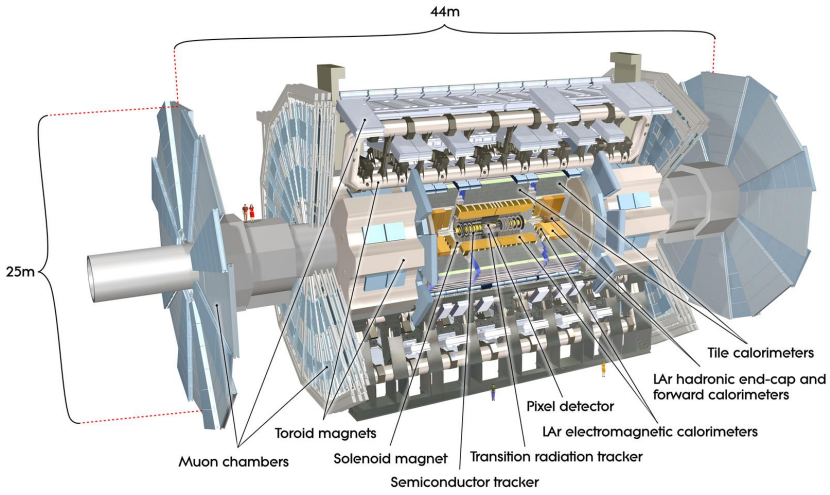

Fig. 1. The ATLAS Experiment.

1.2 ATLAS Distributed Computing (ADC)

The ATLAS Distributed Computing (ADC) group is responsible for the development and operations of infrastructure necessary to store and process data used in physics analysis.

CPU resources: The ADC uses heterogeneous computing resources - Worldwide LHC Computing Grid (WLCG) sites, cloud resources, HPCs, and volunteer computing (BOINC) resources. Some of these resources are used as part of a pledge (mainly WLCG sites) and they are guaranteed to be available to the ATLAS experiment. Others are used opportunistically, i.e. ATLAS jobs are allowed to run there without any dedicated allocation. Clouds, HPCs, and volunteer computing (BOINC) resources are amongst those. All these allow ATLAS to run stably with more than 300 thousand utilized cores (see Figure 2). The Salomon HPC cluster of the Czech national HPC center IT4Innovations is one such opportunistically-used group of HPCs.

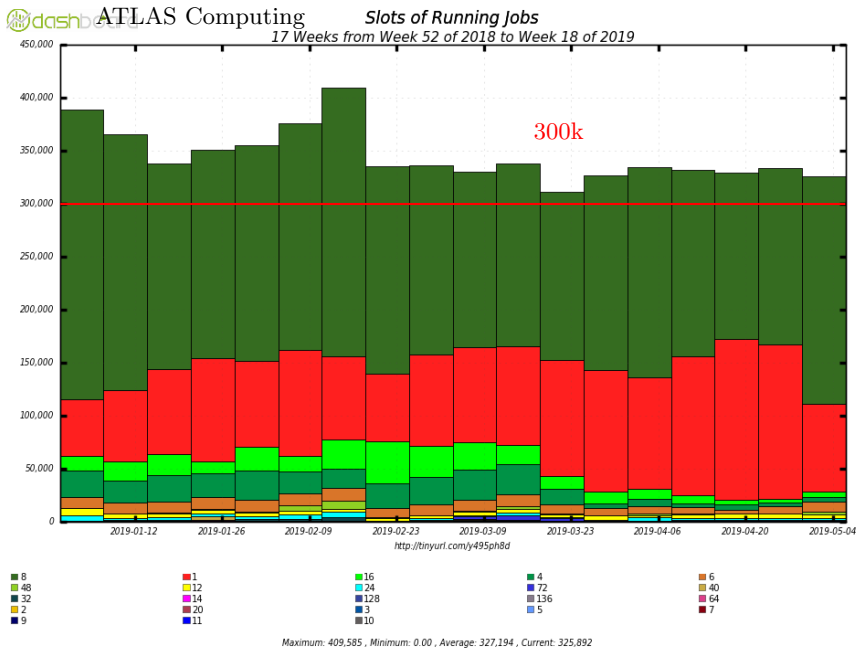

Fig. 2. Weekly average number of cores used by the ADC since the beginning of 2019.

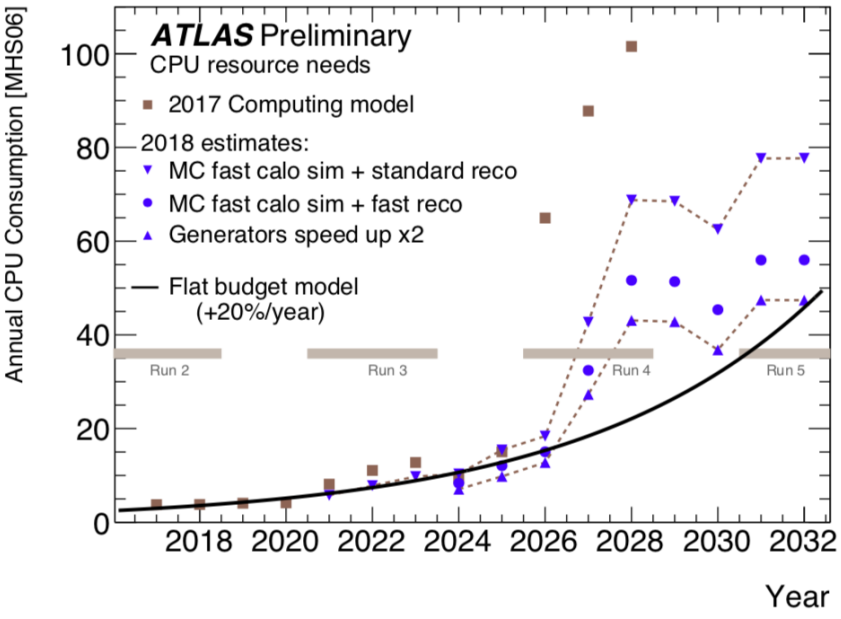

Fig. 3. Estimated CPU resource needs of the ADC [3] in HEPSPEC06 [7].

The reason for such a varied computing environment is the ever-increasing need for computing resources. In Figure 3, the prediction of computing needs is displayed. The, so called, Run 4 will bring major improvements in the LHC accelerator and therefore significantly more data. For example, using the computing model from 2017, the ATLAS experiment would need about one order of magnitude more computing resources than its budget allows to be able to use all the data. Further improvements of the ATLAS computing model somewhat decrease the required amount of resources.

Storage: The ADC manages more than 400 PB of data (on both disk and tape). They are spread over more than 700 storage endpoints (on more than 150 sites located around the world, see Figure 4).

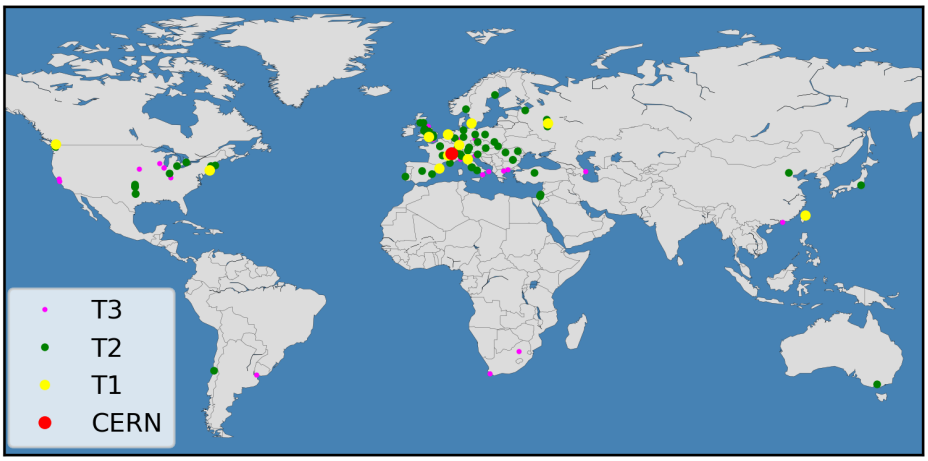

Fig. 4. Geographic locations of ATLAS computing centers: Tier 0 is CERN (red), Tier 1 sites are yellow, Tier 2 sites are green and Tier 3 sites are purple.

2 ARC-CE Submission System Overview

The ARC-CE [5] is a job submission framework which has proven to work under various conditions in many places. It has been modified to allow submission of jobs to Salomon (Figure 5).

The submission of a job starts at the ARC Control Tower (aCT) in CERN. It obtains the job description from the ATLAS workflow management system and submits it to one of the ARC-CE machines located at the computing center of the Institute of Physics of the Czech Academy of Sciences in Prague (praguelcg2) [4]. The ARC-CE translates the job description into a PBS script and stores it with all other dependencies in the session directory which is shared with Lustre

scratch space on the Salomon in Ostrava via sshfs. The shared space is also used for input files which are either linked to the session directory from a cache directory (also on the scratch of the Salomon) or copied there from local DPM grid storage. Then the ARC-CE submits the job to the PBSpro batch system [2] via ssh connection to Salomon's login node. When a job has the opportunity to run in the batch system, it uses software which is also stored on the scratch. When the job finishes, the output and log files are located in the session directory. They are accessed by the ARC-CE via sshfs. The job output and log files are stored in CERN S3 Object Store and a copy of the log files is stored in the local DPM grid storage. More details of the setting can be found in [8].

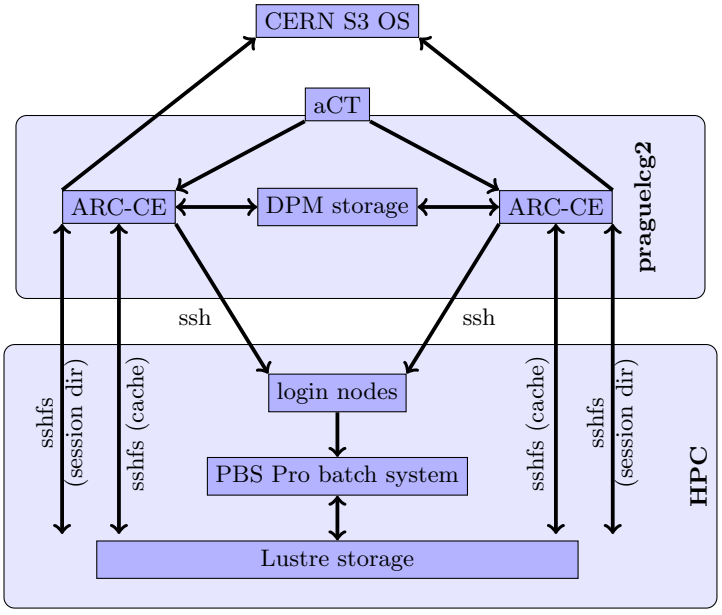

Fig. 5. Schematic of job submission system.

3 Number of PBS Requests

The ARC-CE interacts intensely with the PBSpro batch system. Apart from job submission and deletion, it performs rounds of checking via `qstat`, `pbsnodes`, and `qmgr` to get overview of the batch system status. These queries can be quite heavy and Salomon limits the number of PBS requests to protect the system and exceeding this limit puts a user into a blacklist (which can lead to failed submissions).

When a default setting is used and PBS commands executed by the ARC-CE are logged, the ARC-CE usually performs 8 PBS commands per minute (Figure

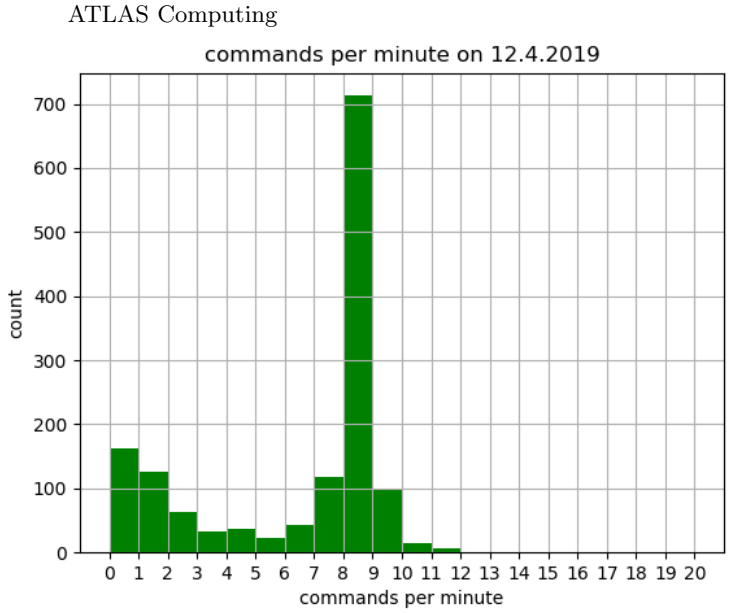

Fig. 6. Histogram of number of PBS requests before modification of the submission script.

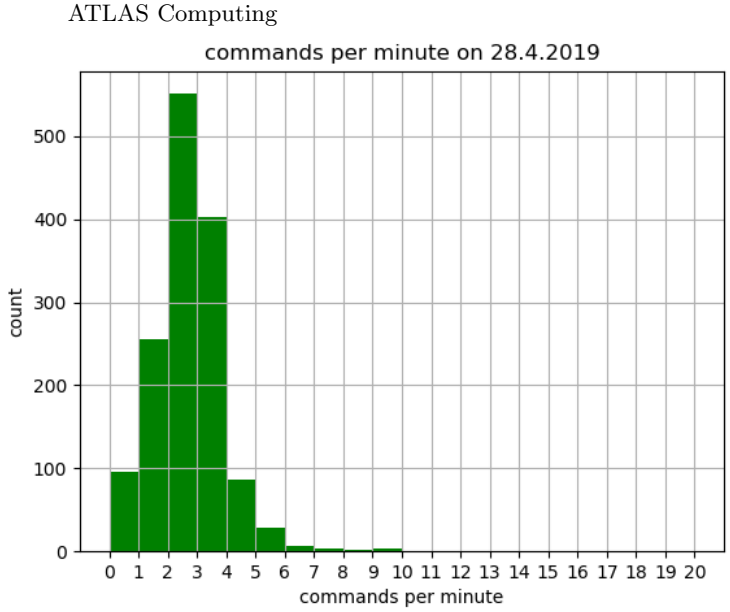

Fig. 7. Histogram of number of PBS requests after modification of the submission script.

6). This is above the limit and eventually leads to failed submission. The ARC-CE commands were modified to slow down the round of checking. Specifically, a sleep of 30 seconds was put before any execution of `qstat`, `pbsnodes`, or `qmgr`. After that, the number of PBS commands is usually below 6 per minute (Figure 7) which is within the limits allowed by Salomon.

4 Tuning of the sshfs

When observing how fast the HPC resources can be filled and refilled by the ARC-CE, it appeared there is a bottleneck in the sshfs connection between the ARC-CE machine and scratch space of Salomon. When output files are being produced and then shared back with the ARC-CE machine, the new jobs cannot start because they are waiting for input files to become available. The scale of the problem can be guessed from Figure 8. The traffic reaches a plateau around 60 Mbps. But in a simple throughput test using scp, the plateau appears at the level of 500 Mbps. The underlying hardware of the ARC-CE (it is a virtual machine) has a 10 Gbps network card and the WAN connection between IT4I and FZU is also 10 Gbps.

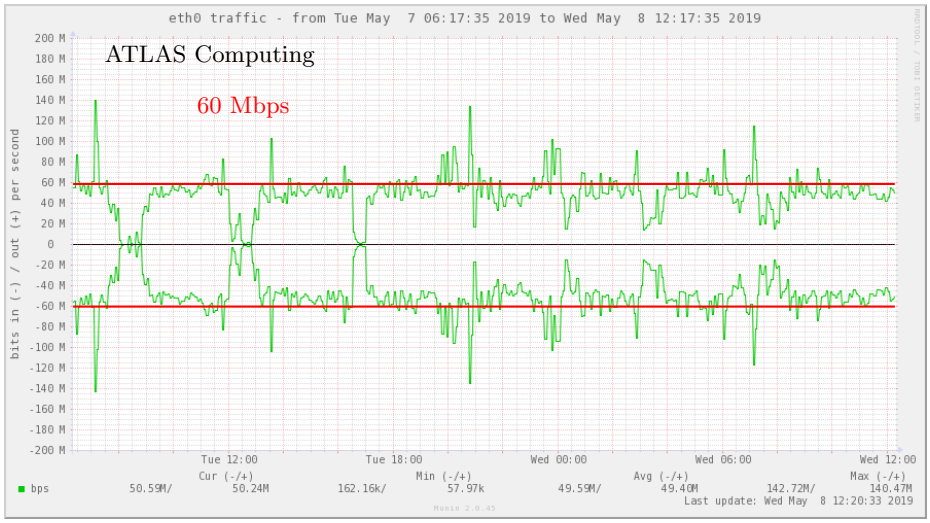

Fig. 8. eth0 traffic of the ARC-CE machine.

Several sshfs parameters were tweaked, one by one, and their effect observed. First, a compression was turned off. In theory this should decrease the CPU load and make transfers faster. But no significant difference was observed. Second, usage of a faster encryption algorithm (aes128-ctr) was investigated with the same motivation (lower CPU usage). Again, no significant difference was

observed. Another possibility is usage of various caching settings. This is under investigation.

The current assumption is that the low speed is caused by a high number of files. Each running job can reach more than thousand files, i.e the sshfs has to handle up to $O(100)$k files.

5 Monitoring

To guarantee submission system stability, many parameters of involved hosts and services are monitored. In case of problems, these data provide necessary insights that help with investigation and understanding of observed issues.

5.1 Jobs in the Batch System

Job states in the Salomon's batch system are monitored via the munin framework [1] installed at the praguelcg2. The information are obtained via the `qstat` command running once every 30 minutes. This granularity is selected to avoid overloading the PBS servers (see section 3). The default sampling frequency of the munin framework is 5 minutes.

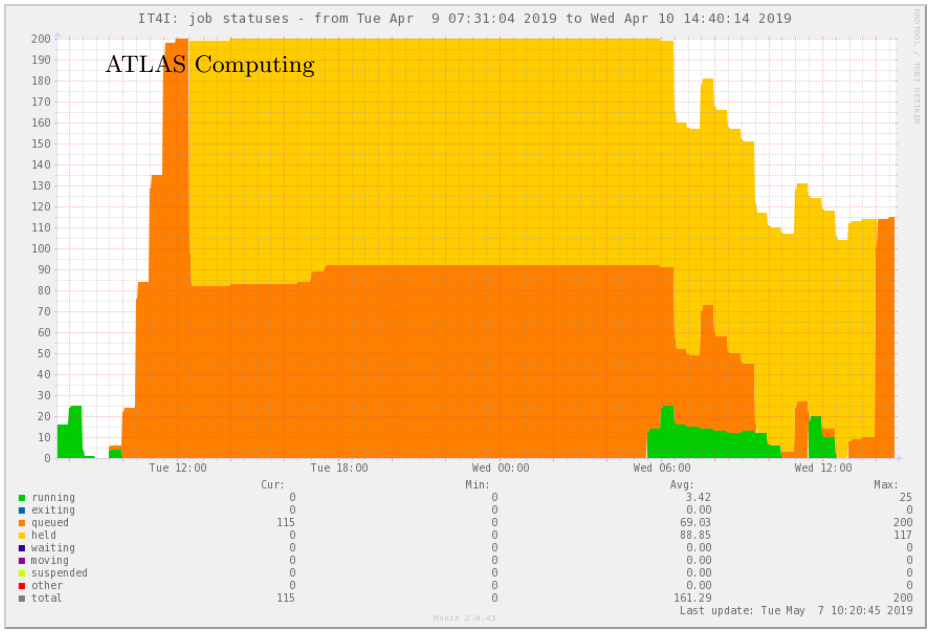

Fig. 9. Distribution of job states in the PBSpro [2].

Figure 9 illustrates the monitoring of job states. In this case, the plot shows

a batch system scheduler problem which caused many jobs to be stuck in held state. This problem would otherwise not be visible and held state jobs would use part of the job quota until they would eventually be killed, i.e. they would block submission of new jobs.

Figure 10 illustrates the monitoring of running jobs per user. Two user accounts are used to submit jobs (each from a separate ARC-CE machine) and each has a limit of 100 jobs. If running jobs from one user disappear, it would indicate that one of the ARC-CE machines could have problems. The structure with many peaks is caused by the sshfs limitations (see Section 4), i.e. while the output of finished jobs are transferred to the ARC-CE machine, a new job cannot push its input file to start running.

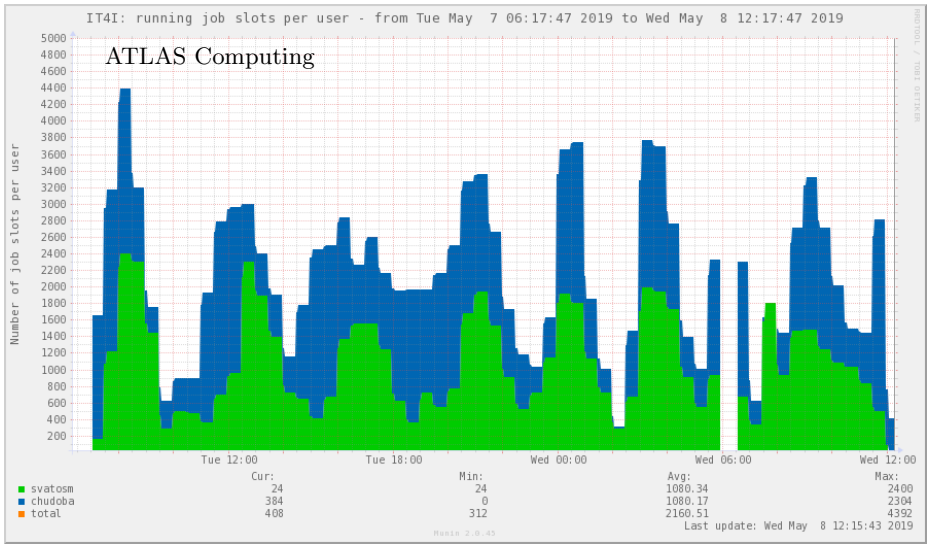

Fig. 10. Distribution of running jobs per user in the PBSpro.

5.2 Jobs in the ARC-CE

Job states in the ARC-CE are also checked by the munin framework. Figure 11 shows an example of one-day activity. It shows the following summary states:

- "running jobs", the number of jobs sent to Salomon;
- "accepted jobs", the jobs which arrive to the ARC-CE machine (being prepared or running);
- "all jobs", which includes finished and deleted jobs.

The figure illustrates the 100-job limit as well as how long it takes to refill the quota after some jobs have finished.

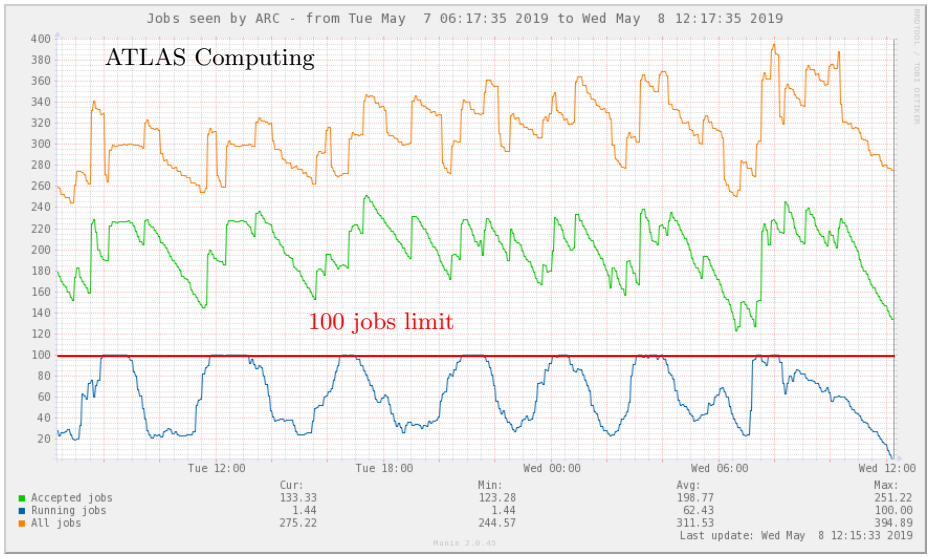

Fig. 11. Distribution of selected job groups in the ARC-CE.

5.3 Used Space

There are also other important factors to monitor, other than job states. One of them is the available storage. More precisely, if all data fit into quotas on size and number of files per user. The number are obtained from output of `it4i-disk-usage` command which is executed once a day. Monitored storage directories are `/scratch/work`, where the software is located and `/scratch/temp` with session+cache directories in it. Figure 12 shows the number of entries, which is the most significant for the software part (about 14M entries). Figure 13 shows the most significant amount of disk is occupied by session+cache directories (most of it is taken by the cache).

5.4 Other Monitoring

Finally, the monitoring pages of IT4I also need to be mentioned.

First, IT4I provides a publicly available monitoring dashboard of Salomon (https://extranet.it4i.cz/dash/salomon). It provides an overview of storage and cluster occupancy. So, when no jobs are running, one can cross-check if the machine is full or if the problem is somewhere else.

Second, there is the Message of the Day (MotD). It provides information about various seminars organized by the IT4I and more importantly about downtimes and outages. The submission system state is adapted to the MoTD information. The crucial attribute of the MotD is that the content can be collected programmatically which allows it to appear in the praguelcg2 monitoring. The MotD was available at https://extranet.it4i.cz/motd/all. But this page was decommissioned and https://scs.it4i.cz/api/v1/motd/all replaced it.

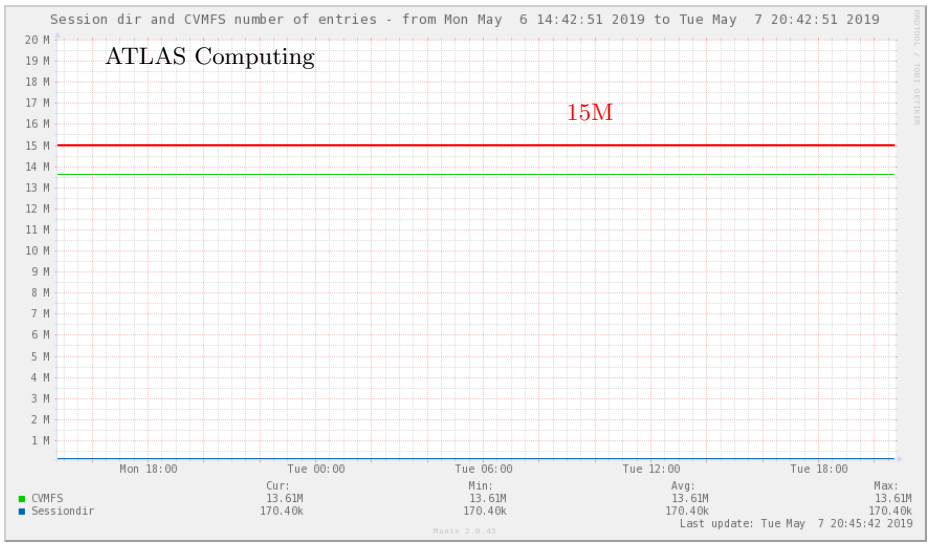

Fig. 12. Distribution of number of entries in the Lustre.

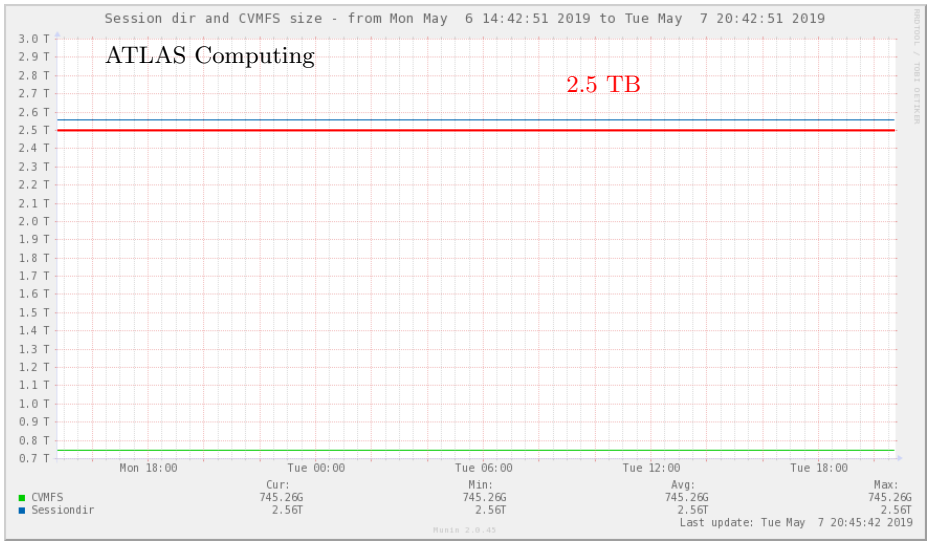

Fig. 13. Distribution of used space in the Lustre.

6 Summary and Conclusions

The submission system for ATLAS jobs into Salomon was configured and it can automatically process jobs from the ATLAS workflow management system. The necessary grid middleware runs in the computing center of the Institute of Physics of the Czech Academy of Sciences in Prague. It is successfully submitting jobs since the end of 2017 and therefore contributing to progress of the physics programme of the LHC experiments.

Several aspects of interaction between the computing center and Salomon have been investigated. The number of PBS requests performed by the ARC-CE has been decreased to protect Salomon and to avoid job submission failures. Tuning of sshfs parameters is ongoing because the changes in compression and encryption had no significant effect. Various aspects of the system are monitored and frequently checked.

Acknowledgement

Computing resources are co-financed by projects Research infrastructure CERN (CERN-CZ) and OP RDE CERN Computing (CZ.02.1.01/0.0/0.0/16013/0001404) from EU funds and MŠMT.
We gratefully acknowledge National Supercomputing Center IT4Innovations (IT4I) for computing time provided by the Salomon supercomputer.

Copyright

References

1. http://munin-monitoring.org/
2. Altair pbs professional 19.2 administrator's guide. https://www.pbsworks.com/pdfs/PBSAdminGuide19.2.1.pdf
3. Computing and software - public results. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ComputingandSoftwarePublicResults
4. Chudoba, J., Adam, M., Adamová, D., Kouba, T., Mikula, A., Říkal, V., Švec, J., Uhlířová, J., Vokáč, P., Svatoš, M.: A multipurpose computing center with distributed resources. J. Phys. Conf. Ser. **898**(8), 082034 (2017). https://doi.org/10.1088/1742-6596/898/8/082034
5. Ellert, M., Grønager, M., Konstantinov, A., Kónya, B., Lindemann, J., Livenson, I., Nielsen, J.L., Niinimäki, M., Smirnova, O., Wäänänen, A.: Advanced Resource Connector middleware for lightweight computational Grids. Future Gener. Comput. Syst. **23**, 219–240 (2007). https://doi.org/10.1016/j.future.2006.05.008
6. ATLAS Collaboration: The ATLAS Experiment at the CERN Large Hadron Collider. JINST **3**, S08003 (2008). https://doi.org/10.1088/1748-0221/3/08/S08003

7. Michelotto, M., et al.: A comparison of HEP code with SPEC benchmarks on multi-core worker nodes. J. Phys. Conf. Ser. **219**, 052009 (2010). https://doi.org/10.1088/1742-6596/219/5/052009
8. Svatos, M., Chudoba, J., Vokac, P.: ATLAS utilisation of the Czech national HPC center. EPJ Web Conf. **214**, 03005 (2019). https://doi.org/10.1051/epjconf/201921403005