

Distributed Computing for the Pierre Auger Observatory

J. Chudoba¹ for the Pierre Auger Collaboration²

¹Institute of Physics of the CAS, Na Slovance 1999/2, Prague, 18221, Czech Republic

²Full author list: http://www.auger.org/archive/authors_2015_02.html

E-mail: chudoba@fzu.cz

Abstract. Pierre Auger Observatory operates the largest system of detectors for ultra-high energy cosmic ray measurements. Comparison of theoretical models of interactions with recorded data requires thousands of computing cores for Monte Carlo simulations. Since 2007 distributed resources connected via EGI grid are successfully used. The first and the second versions of production system based on bash scripts and MySQL database were able to submit jobs to all reliable sites supporting Virtual Organization *auger*. For many years VO *auger* belongs to top ten of EGI users based on the total used computing time. Migration of the production system to DIRAC interware started in 2014. Pilot jobs improve efficiency of computing jobs and eliminate problems with small and less reliable sites used for the bulk production. The new system has also possibility to use available resources in clouds. Dirac File Catalog replaced LFC for new files, which are organized in datasets defined via metadata. CVMFS is used for software distribution since 2014. In the presentation we give a comparison of the old and the new production system and report the experience on migrating to the new system.

1. Pierre Auger Observatory

Pierre Auger Observatory is the world's largest facility for measurements of ultra-high energy cosmic rays. The main objective of the Observatory is to determine properties of cosmic rays with energies above 10^{17} eV. Because cosmic rays with such energies are rare, an array of 1660 water-Cherenkov detectors (surface detector – SD) was built covering area of 3000 km². Good energy calibration is provided by 24 air fluorescence detectors (FD). Some new detectors were added after construction of these base detectors between 2002 and 2008 for an extension of energy range or for purpose of testing new detection techniques. Major upgrade of SD is planned to be realized between 2016 and 2019. Detailed overview of the Pierre Auger Observatory with its major scientific results can be found in [1].

2. Computing requirements and their realization

Large scale facility consisting of hundreds of surface detectors, tens of telescopes and several monitoring devices need various computing services for successful operation. Here we discuss only raw data storage and simulated data production and storage; onsite data communication and general data acquisition systems are described in [1].

2.1. Computing requirements

Typical size of file with raw data from one event has size of hundreds of kB. Total size of raw data files from SD and FD is under 10 TBs per year. One copy is stored in IN2P3 Computing Centre in Lyon and another copy is sent to FERMILAB. Derived data with reconstructed parameters are much smaller and



can be downloaded via a standard FTP server. Total size of real data does not present a challenge for current computing resources.

2.2. Simulated data production and storage

Reconstructed raw data are compared to simulated data. The most common framework for cosmic ray shower propagation through the atmosphere is CORSIKA [2]. Processing of one shower with energy of primary particle exceeding 10^{20} eV requires tens of CPU hours on current processors even when some simplifications (like thinning) are used. Several different models of interactions of particles at energies well above the reach of current accelerators can be used with different parameters. The number of combinations of input parameters is increased by several types of primary particles, different energy range and incoming direction angles and also by different versions of computing programs used for physical models. The output from CORSIKA simulations is processed by Offline framework [3], which simulates response of detectors. CORSIKA shower is placed on a random place within the detector array and can be reused several times to get different responses.

Number of simulated showers is limited by computing resources.

2.3. Local and grid resources

Local computing farms were used for simulations. It was difficult to verify results from various resources because different computing environments and job parameters were used. The access to results is limited only to users with local account. These resources are still used for special productions required only by a single user or a small group.

In 2006 we have created VO *auger* in order to access opportunistic resources available on sites supporting LCG and later gLite middleware. The first CPU challenge [4] used just 4 sites. A low number of successfully finished jobs (80% and 51% in phase 1 and phase 2) demonstrated that the grid resources should not be used without a system extending the basic middleware commands for a job and data management.

The first bulk production in 2007 used 7 sites from 7 countries and produced 8400 showers in 9500 jobs consuming 7000 CPUdays. The output data used 4 TB of disk space. We have created a set of production scripts for the bulk job submission, the retrieval of log files and the resubmission of failed jobs [5]. Later we have added a Django dashboard for the production monitoring [6,7].

Another big change in the production framework was in the usage of a MySQL database for job states and PHP based web for production status [8]. This system is able to handle several thousands of simultaneously running jobs. For several years the VO *auger* became the biggest user of EGI computing resources just after the 4 LHC collaborations. More than 1 million of jobs per month were processed in peak periods and equivalent of 4000 constantly used cores for periods of months was reached. Jobs statistics for period from 2010 to 2014 is illustrated in Fig. 1 and 2. Complete statistics is available at the EGI accounting portal [9].

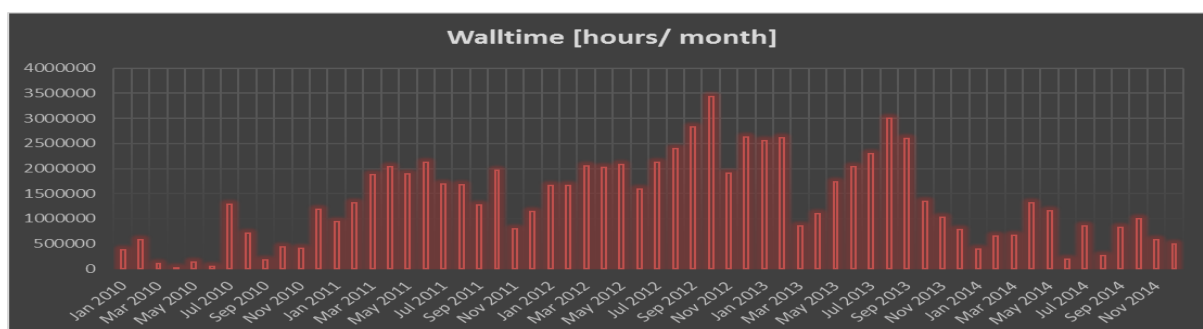


Figure 1: Total processor time used by VO *auger* from 2010 to 2014.

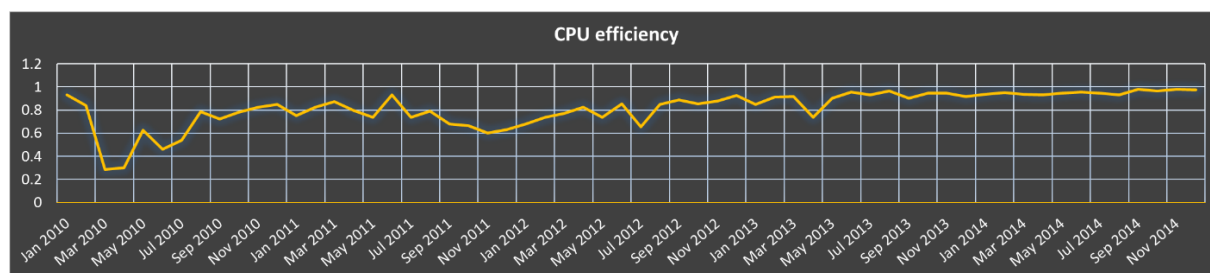


Figure 2: CPU efficiency (ratio cputime/walltime) for auger jobs for period 2010 – 2014.

3. Towards common solutions

The Auger production system based on the set of scripts and the MySQL database proved to be able to produce required simulations and utilize thousands of simultaneous jobs. Limitations in the manpower for further development and operation lead us to look for a system shared with other big international collaborations. In 2014 we chose the DIRAC interware [10] as the basis for the next generation of the production system. DIRAC has relatively long history of usage by the LHCb [11] experiment and was rewritten to remove LHCb specific parts. Several other big international collaborations like Belle II [12] and CTA [13] have started to use it. In the future their computing requirements will be much larger than the Auger requirements. The DIRAC consortium was created by CERN, University of Barcelona and CNRS with firm promises to further contribute to the development and maintenance of the DIRAC software.

3.1. DIRAC file catalog

The VO *auger* uses an LFC file catalog. This product was recently abandoned by the LHC experiments. There is no new development and its further maintenance is limited. We have tested the DIRAC file catalog (DFC) [14] as a possible replacement of the LFC. About 30 million entries belonging to the VO *auger* were copied to a separate instance of the DFC catalog. A performance test on bulk queries of replicas, queries of replicas in directories and directory listings were done. The results presented at ACAT 2014 [15] showed better performance of the DFC in these tests.

We decided not to migrate all entries to the new catalog. Only new files will be registered in the DFC and the requested sets of files will be migrated. In this way we plan to get rid of the old data, which are difficult to identify in the LFC.

3.2. Datasets in DIRAC File Catalog

Libraries of generated showers may contain thousands of output files. Users do not want to work with individual files for processing, they want to use lists of files. Such lists of files belonging to each shower are generated for each numerated production - run. These lists are static and are created when the given production is finished.

The DIRAC File catalog offers also the functionality of a metadata catalog. Attributes of the type string, integer, float and date can be assigned to each file. These attributes can be used in dynamic definition of datasets. Users will be able to process whole datasets or replicate them to the selected locations. We developed a first implementation, which is available in the AugerTestCatalog. An example of the behaviour of frozen and released datasets when new files are added to the catalog is demonstrated in Fig. 3. Behaviour of frozen datasets in a case of lost files belonging to the datasets are under discussion and will be fixed according to the users' feedback.

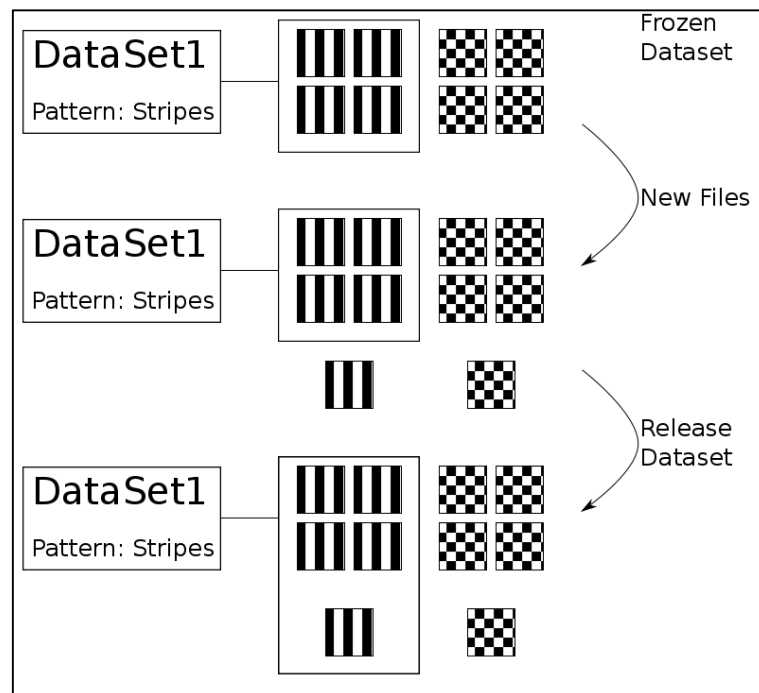


Figure 3: Behaviour of frozen and released datasets: a file fulfilling definition criteria is added to a frozen dataset only after a “*release dataset*” command.

3.3. Bulk production

Current production system for bulk simulations is operated by a production manager. The production manager has access to the MySQL database and has an overview of running production jobs. The complete view of all running and waiting jobs in the VO is available only via information from BDII, which is often not reliable. Statistics of usage reported on EGI accounting portal are gained with a delay of several days or even weeks. DIRAC offers web interface for monitoring of all jobs for a given VO in all states and also accounting information. It will be used for monitoring of individual users’ productions. Users can be immediately notified if they use too much resources or if their jobs behave incorrectly. This is a much easier approach in comparison with a setup of fair-share ratios for production role on each site.

3.4. VO management

All VO users must be registered in the VOMS server. Registration for the VO *auger* is done via Perun [16] developed by CESNET. Perun was used for administration of users and also of resources in the Czech NGI. Current version extends VOMRS features for user administration. Users are authenticated by the X509 certificates. During registration they must choose a unique username which is used for creation of an account on a common User Interface. This feature is especially handy for the DIRAC system, where each user must be registered also under a unique username. Perun periodically propagates users to VOMS servers and deletes users with expired membership. Automatic propagation to DIRAC servers will be developed in the future.

3.5. CVMFS

Important improvement in the software distribution was achieved by using CVMFS [17]. While the CORSIKA program written in the FORTAN language is easily installed from a downloaded tarball, The offline software requires many packages and is not relocatable. We installed it on selected sites often with a help from local administrators. In cooperation with EGI CVMFS task force we were able to use RAL resources as stratum-0 under a generic domain *auger.egi.eu*. Now all EGI sites supporting VO

auger configured this CVMFS domain and we have exactly the same software environment on all these sites.

4. Conclusions

Computing requirements on simulations for the Pierre Auger Observatory are mostly covered by grid resources. We plan to migrate the bulk production to the DIRAC system during 2015 and start to use the DIRAC File Catalog for new entries. The old LFC catalog will be maintained for sufficiently long time to ensure the access to the old files. The DIRAC interware will enable unified access to new resources like Cloud and HPC.

Acknowledgment

The DIRAC dataset implementation was done by M. Adam after many detailed discussion with DIRAC technical coordinator A. Tsaregorodtsev. The VO *auger* bulk production is operated by G. Rubio. Transition to the new production system with DIRAC and CVMFS is supported also by G. Isar and A. Sevcenco. DIRAC central services are operated by NGI France, DIRAC catalog and VO management servers are provided by CESNET. This work was partially supported by the Ministry of Education of the Czech Republic, projects LM2010005 and LG13007 (J. Chudoba) and by a grant of the Minister of Education and Scientific Research of Romania, Programme for research - Space Technology and Advanced Research - STAR, project number. 83/2013 (G. Isar and A. Sevcenco).

References

- [1] Aab A *et al* 2015 The Pierre Auger Cosmic Ray Observatory, *Nucl. Instrum. Meth. A* **798** pp 172–213
- [2] Heck D *et al* 1998 CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers, *Report FZKA* **6019**
- [3] Argiro S *et al* 2007 The Offline Software Framework of the Pierre Auger Observatory, *Nucl. Instrum. Meth. A* **580** 1485-1496
- [4] Chudoba J and Schovancova J 2007 Simulations and Offline Data Processing for the Auger Experiment, 2nd EGEE User Forum, Manchester
- [5] Schovancova J *et al* 2008 VO auger Large Scale Monte Carlo Simulations using the EGEE Grid Environment, 3rd EGEE User Forum, Clermont-Ferrand
- [6] Schovancova J and Chudoba J 2009 Django dashboard and Grid production monitoring for the VO AUGER, 4th EGEE User Forum, Catania
- [7] Chudoba J *et al* 2010 Simulation and reconstruction of cosmic ray showers for the Pierre Auger Observatory on the EGEE grid *J. Phys.: Conf. Series* **219** 072033
- [8] Bahilo J L 2012 Mass production of extensive air showers for the Pierre Auger Collaboration using Grid technology, *J. Phys.: Conf. Series* **368** 012015
- [9] EGI accounting portal, <http://accounting.egi.eu>
- [10] DIRAC interware, <http://diracgrid.org>
- [11] LHCb experiment, <http://lhcb.web.cern.ch/lhcb/>
- [12] Belle II experiment, <http://belle2.kek.jp/>
- [13] CTA experiment, <https://www.cta-observatory.org>
- [14] Tsaregorodtsev A and Poss S 2012 DIRAC File Replica and Metadata Catalog *J. Phys.: Conf. Series* **396** 032108
- [15] Chudoba J and Tsaregorodtsev A 2014 Comparison of LFC and DFC grid file catalogues International Grid Symposium, Taipei (2014)
- [16] Prochazka M and Licehammer A and Matyska L 2014 Perun – Modern Approach for User and Service Management, IIMC International Information Management Corporation Ltd ISBN: 978-1-905824-44-1

- [17] Blomer J *et al* 2011 Distributing LHC application software and conditions databases using the CernVM file system *J. Phys.: Conf. Series* **331** 042003