Swiss ATLAS Grid computing in preparation for the LHC collision data

S.Haug, C.Topfel, E.Cogneras
LHEP, University of Bern

P.Kunszt, S.Maffioletti, R.Murri
CSCS, Swiss National Supercomputing Centre

S.Gadomski
DPNC, University of Geneva
szymon.gadomski@unige.ch

Abstract. Computing for ATLAS in Switzerland has two Tier-3 sites with several years of experience, owned by Universities of Berne and Geneva. They have been used for ATLAS Monte Carlo production, centrally controlled via the NorduGrid, since 2005. The Tier-3 sites are under continuous development. In case of Geneva the proximity of CERN leads to additional use cases, related to commissioning of the experiment, which require processing of the latest ATLAS data using the latest software under development, normally not distributed to grid sites. The Swiss Tier-2 at the CSCS centre has a recent and powerful cluster, serving three LHC experiments, including ATLAS. The system features two implementations of the grid middleware, NorduGrid ARC and the LCG gLite, which operate simultaneously on the same resources. In this article we present our implementation choices and our experience. We will discuss the requirements of our users and how we meet them. We will present the status of our work and our plans for the ATLAS data taking period in 2009-2010.

1. Introduction
Computing for the ATLAS experiment [1] in Switzerland is a part of a globally distributed system, using multiple grids to implement a hierarchy of computer centers. The hierarchy, which is described in [2], has CERN as a Tier-0 center, followed by large Tier-1 centers, smaller regional Tier-2 centers and finally Tier-3 centers, which belong to universities and institutes of the ATLAS collaboration.

Contrary to what is done in other countries, the Swiss ATLAS computing has started from Tier-3 sites. The work began in 2005, when first prototypes, integrated with the ATLAS grid production, were put in place by the University groups of Bern and Geneva [3]. The first grid installation at the Swiss National Supercomputing Centre (CSCS) in Manno, near Lugano, was in operation in 2006. The CSCS plays the role of the Swiss Tier-2 center. As there is no Tier-1 in Switzerland, we are associated with the FZK in Karlsruhe.
All three Swiss sites have contributed steadily to the ATLAS grid, undergoing upgrades every year. As the activities of the ATLAS groups progressively shift from detector construction to commissioning and to preparation of data analysis, demand of local usage increases at the Tier-3 sites.

While physical distance to computing facilities matters less nowadays, the physical distance to people is still an important factor in a collaboration like ATLAS. This leads to additional use cases for the Geneva Tier-3. All the members of the Geneva group work every day at CERN. They have a leading role in the High Level Trigger system. It is useful that many other ATLAS physicists, not members of the Geneva group but working with us on the Trigger, use the same computing installation. Over sixty ATLAS physicists are now using the Geneva T3 cluster for their daily work. The cluster has become an important installation for software development and testing of the ATLAS Trigger.

2. Hardware and operating systems

Each of the three sites has been significantly upgraded two or three times since the first prototypes were put in place. The present status of the hardware is summarized in Table 1.

| Cluster     | CPU cores | Disk [TB] | OS  | Shared FS | Batch system |
|-------------|-----------|-----------|-----|-----------|--------------|
| Bern T3 A   | 30        | 30        | SLC4| NFS       | Torque       |
| Bern T3 B   | 1000      | 0         | Gentoo | Lustre   | Sun Grid Engine |
| CSCS T2     | 960       | 520       | SL4 | NFS       | Torque+Maui  |
| Geneva T3   | 188       | 150       | SLC4| NFS       | Torque+Maui  |

\(^a\) shared with other research domains
\(^b\) shared with other LHC experiments (CMS and LHCb)

The Tier-3 in Bern is composed of two clusters. The smaller cluster is owned by the ATLAS group, the larger one, named Ubelix, is owned by the University and is shared with other research activities. The operating system on the Ubelix, which can not be dictated by particle physics activities, is Gentoo Linux. Running ATLAS software on this system is possible, but can be costly in terms of manpower. It has often required significant efforts from the involved physicists. The Tier-2 at the CSCS is shared by the three LHC experiments, in which Swiss Particle Physics groups participate: ATLAS, CMS and LHCb.

3. Grid middleware and other software

The middleware used at the Swiss ATLAS grid sites, including Storage Element technology, is listed in Table 2.

| Cluster     | Middleware        | Storage Element |
|-------------|-------------------|-----------------|
| Bern T3 A   | NorduGrid ARC     | ARC SE          |
| Bern T3 B   | NorduGrid ARC     | -               |
| CSCS T2     | gLite + NorduGrid ARC | dCache     |
| Geneva T3   | NorduGrid ARC     | ARC SE + DPM   |

Choices of the middleware may seem exotic and require an explanation. The Tier-3 sites of Bern and Geneva have started in 2005 with rather small installations, where it would not be possible to dedicate multiple machines to running services. We also knew that we could not impose the choice of
Scientific Linux as the operating system of Ubelix (Bern T3 B), which was the only larger cluster involved at the time. This has led to the choice of Nordu Grid ARC middleware [4], which could meet all the mentioned constraints more easily than the LCG middleware at the time. One machine was sufficient to run the grid front-end services, there was no middleware on the worker nodes and various distributions of Linux were supported.

The Tier-3 installations have in the meantime grown substantially. From hardware point of view it would now be possible to install the EGEE gLite middleware [5] at the T3 sites. However, there remain important manpower considerations. The T3 sites are managed by physicists, who are not supposed to work full time on the grid projects. This is an important difference between a university group and a computer center, which typically employs two or three people responsible for system administration and for the grid. We believe that NorduGrid middleware is more manageable by a university group.

The Swiss National Supercomputing Centre operates an EGEE Tier-2 site, attached to the Karlsruhe Tier-1 cloud. In addition the CSCS T2 has a NorduGrid front-end, which submits jobs to the same worker nodes. This redundancy has already proven useful in the past.

In order to receive data distributed centrally by ATLAS computing operations, a site must provide an SRM interface. One of the possible implementations is dCache [6], which was installed at the CSCS T2. Installation and maintenance of this system has required a significant effort. It would be difficult to manage such an installation in a university environment.

Unfortunately an ARC Storage Element, which does not implement an SRM interface, can not be integrated into the Distributed Data Management system of ATLAS. At the T3 in Geneva we are currently testing the Disk Pool Manager software [7]. As the disk servers of the Geneva T3 run Solaris, the DPM software was ported to this operating system by its main developer.

4. Applications and results

In terms of resource utilization, the dominant application running at the Swiss ATLAS Grid sites is the ATLAS Monte Carlo production. The production is operated via the Nordu Grid at the T3 sites. In case of the T2 in Manno the production comes mostly by the EGEE channel. The total computing time, calculated in days per calendar year, is shown in Figure 1 for years 2006-2008. One should note that the Tier-3 sites, which are not formally required to contribute their CPU time to central ATLAS operations.

![Figure 1. Computing time measured in days (wall time) used by ATLAS jobs running at the four cluster of the Swiss ATLAS grid in years 2006-2008.](image)

The Tier-3 clusters in Bern and in Geneva are also used by members of the local ATLAS groups for various studies done in preparation to data analysis. Such studies lead to developments of techniques and tools, which will later be used for analysis of real data. Local usage of the cluster is increasing in both Bern and Geneva University groups.
In our experience running ATLAS software on Gentoo Linux (Ubelix cluster, Bern T3 B) has proven manpower-intensive on a few occasions. The solution recently adopted is to use the ‘chroot’ command to provide an environment similar to the Scientific Linux.

The Geneva T3 has additional use cases, because the group plays a leading role in the development of the ATLAS Trigger system. The group members usually work at CERN, where they work closely with other Trigger developers. In the interest of the Trigger project, all Trigger developers can have accounts at the Geneva T3, whether or not they are associated with the University. The Tier-3 is used for general development and testing work on the ATLAS Trigger. This work requires using nightly builds of the Trigger software, which are normally not distributed to grid sites. In Geneva the software is available on the AFS file system, supported by CERN.

5. Plans for first collision data
The analysis preparations will continue at the T3 sites. They will be replaced by intensive analysis of the LHC data as soon as the data will be available. We expect nevertheless that all four clusters will continue to be used for ATLAS Monte-Carlo production.

The Geneva Tier-3 site will continue as a Trigger development and testing facility. We are planning to participate in commissioning and in subsequent data quality studies of the Trigger. Unlike detectors, the Trigger will need to continuously evolve during the data taking. We will need to continue testing new configurations of the Trigger, preparing new releases of the Trigger software and verifying them before they are used online by ATLAS. Access to small subsets of recent ATLAS data will be necessary. The DPM-based Storage Element, mentioned in section 3, will enable data transfers to Geneva controlled by the Distributed Data Management system of ATLAS [8].

At present a typical analysis in our University groups starts by copying datasets to a Tier-3 site. We expect that as data samples will become larger, notably when processing real data, the physicists in Bern and Geneva will be sending their jobs to the grid. The results will be copied to Tier-3 sites for further analysis. This analysis model, which is different from the current practice, is still largely untested for us. It is important that we start testing before the real data analysis will start.

6. Summary
The Swiss ATLAS computing is, as one might expect, less centralized than ATLAS computing in other countries. Relatively strong Tier-3 centers in Bern and Geneva have been contributing their resources to ATLAS grid computing since 2005. The Tier-2 center at the CSCS near Lugano is active since 2006 and has recently reached the scale of 960 CPU cores. In case of Geneva the proximity of CERN leads to additional use cases for the Tier-3, related to the commissioning of the High Level Trigger system. The Bern Tier-3 uses two clusters, one of which is shared with other disciplines of science.

The two ATLAS groups and one computer center are making an effort to have their systems well integrated with the ATLAS grid systems, so that they can be efficiently used for analysis of early LHC data. We are actively involved in the development and promotion of grids in Switzerland via the SwiNG association [9], where we can contribute our experience.

References
[1] The ATLAS experiment, http://atlas.ch.
[2] ATLAS Collaboration, Computing Technical Design Report, CERN-LHCC-2005-022, ATLAS-TDR-017.
[3] S. Gadomski, et al., “The Swiss ATLAS Computing Prototype”, ATL-SOFT-PUB-2005-03, CERN-ATL-COM-SOFT-2005-07.
[4] NorduGrid, http://www.nordugrid.org/
[5] gLite middleware, http://glite.web.cern.ch/glite/.
[6] dCache software, http://www.dcache.org/.
[7] Disk Pool Manager software, https://twiki.cern.ch/twiki/bin/view/LCG/DpmGeneralDescription.
[8] S.Campana, “Experience Commissioning the ATLAS Distributed Data Management system on top of the WLCG Service”, these proceedings.
[9] Swiss National Grid Association (SwiNG), http://www.swing-grid.ch/.