.

The Swiss ATLAS Grid Mid 2008 Progress Report for the SwiNG EB

2008-08-06

S. Gadomski[a], S. Haug[b], C. Topfel[b]

[a]Department of Nuclear and Particle Physics
24, Quai Ernest-Ansermet, CH-1211 Geneva
Switzerland

[b]Laboratory of High Energy Physics
Sidlerstrasse 5, CH-3012 Bern
Switzerland

ABSTRACT

Since 2005 the Swiss ATLAS Grid (SAG) is in production. Summer 2008 the infrastructure consists of about 1100 cores and 180 TB of disk space wich are clustered at four sites and interconnected by a dark fibre network operated at 10 Giga bit per second. The SAG working group is based on a charter which was accepted by the SwiNG Executive Board in 2007. In summary it says that the SAG working group supports, monitors and manages the usage of the SAG. In the first half of 2008 the group has monitored, supported and managed about 28 thousand wall time days on the infrastructure. The activity has been presented in several meetings and workshops. This is the report on the progress.

1 Introduction

The Swiss ATLAS Grid (SAG) is a computing infrastructure whose purpose is to serve Swiss ATLAS physicists. ATLAS is one of four experiments at the Large Hadron Collider in Geneva (CERN) [1]. The data from its detector is expected to answer fundamental questions about the universe. The prototype of the Swiss ATLAS Grid has been described in the ATLAS note "The Swiss ATLAS Computing Prototype" [2]. Since then both infrastructure and usage have evolved.

This first progress report from the Swiss ATLAS Grid (SAG) as a working group of the Swiss National Grid Association (SwiNG) may have a broader scope than other progress reports. The reason for this is the wish to also capture the work on SAG which was done before SwiNG was established. In this way we intend to document the history and continuity of SAG also within SwiNG.

2 The Swiss ATLAS Grid Working Group 2008

The SAG working group is based on a charter which was accepted by the SwiNG Executive Board in 2007 [3]. In summary it says that SAG working group supports, monitors and manages usage of the SAG. Since the foundation it consists of Sigve Haug, who leads the group in 2008, Szymon Gadomski and Cyril Topfel. Further are Peter Kunszt and Sergio Maffioletti associated members from the Swiss National Super Computing Center (CSCS). The group will reconsider its existence and personelle in the end 2008 progress report (see Figure 1).

Figure 1: The Swiss ATLAS Grid Working Group. From the left S. Haug (Bern and 2008 lead), C. Topfel (Bern), S. Gadomski (Geneva), P. Kunszt and S. Maffioletti (CSCS).

The group communicates via wiki pages which are hosted by CSCS [3], a mailing list which is provided by SwiNG (atlas-wg@swing-grid.ch) and monthly meetings which are normally virtual EVO meetings [4].

The group has five main outbound communication lines. It communicates with the Swiss T2 system administrators at CSCS via lcg@cscs.ch, with the German T1 via atlas-germany-computing@desy.de, with the SwiNG Executive Board via its mailing list, with the Swiss Institute for Particle Physics (CHIPP) and the T0 at CERN via personal contacts. For 2008 these lines seem to be sufficent.

Table 1: The clusters in the Swiss ATLAS Grid mid 2008. The ^{s} indicates that the resource is not ATLAS only but shared. The numbers have a 5% fluctuation following the actual status of the resources.

Cluster	WN Cores	Storage/TB	OS	Middleware
Bern ATLAS T3	30	30	SLC4	ARC
Bern UBELIX T3	512^s	0	Gentoo	ARC
Geneva ATLAS T3	152	75	SLC4	ARC
Manno Phoenix T2	400^s	80	SLC4	ARC/LCG
Sum	1094	185		

In the first half of 2008 the human activity of the group has been the following. There were seven virtual meetings whose minutes are collected on the group's wiki pages. There was one meeting in person at CSSC in coincidence with the T2 inauguration of the Phoenix cluster. S. Haug initiated and lead the organisation of the "ARC meets SwiNG 2008" event in Bern [5]. Allthough not directly relevant for the group's charter, this event and its consequences for the grid infrastructure in Switzerland will be of importance for the evolution of the SAG. In the period several presentations were given by the group members. In January C. Topfel presented the SAG at the CHIPP Winter School in Näfels. In March S. Gadomski gave a talk in Geneva at the annual meeting of the Swiss Physics Society. S. Haug presented the SAG at the monthly user support meeting ("Benutzerkonferenz") of the "Informatikdienste" at the Bern University and in the SwiNG Scientific Council meeting, 2008-05-23. S. Haug has also followed and represented the interests of the SAG in CHIPP and T1 meetings.

Via S. Haug the group has also engaged in another SwiNG working group, the Swiss Multi-Science Computing Grid group. In this group's SWITCH/AAA project "Swiss Multi Science Computing Grid" he has taken the lead of the work package 7a which will bring the ATLAS applications, among others, onto this upcoming infrastructure. The SAG has several years of grid experience and thus may contribute and benefit from a new grid infrastructure in Switzerland.

3 Machines and Network

The Swiss ATLAS Grid consists of four clusters which by the mid of 2008 have about 1100 computing cores and 180 TB disk space. The cluster hardware is summarized in Table 1. The machines range from several years old 32 bit one core PC boxes up to new 64 bit 4 core blade servers. Both Intel and AMD processors are represented. The storage systems are all disk based. All clusters use Gigabit ethernet for interconnections, and at least 1 GB RAM is available per core (an complete upgrade to 2 GB RAM per core is ongoing in order to meet the ATLAS RAM requirement).

The sites are connected by the SWITCHlambda dark fibre network, i.e. the bandwith can be adjusted by illuminating the optical fibres with multiple frequencies. The network is currently operated with one 10 Gb/s channel, but can deliver severals [3]. The network maps are shown in Figure 2. As the foreseen ATLAS TDAQ output is about 2.4 Gb/s, the Swiss capacity of the network meets the estimated ATLAS requirements for connectivity by far [2]. The international BelWu 1 Gb/s connectivity to the Tier 1 in Karlsruhe may have to be increased since not only ATLAS, but also LHCb and CMS are served by this connection. However, the direct and redundant SWITCHlambda connection to CERN provides a natural fallback possibility.

A speciality is the direct and dedicated network link between the CERN T0 and the Geneva T3. It enables data transfers directly from the CERN IT to the Tier-3 in Geneva. As only as small fraction of the data can be processed in Geneva, this option is not of interest for final physics analysis of the data, which will need to start with large datasets on Tier-1 sites. Instead the direct line will enable the users of the Geneva Tier-3 to participate more effectively in the commissioning of the ATLAS experiment. During regular data taking the direct line can be used for data quality monitoring, which can be done by regular processing of the order of 1% of the data.

A proposition to use the Geneva Tier-3 for data quality work, focusing in particular on the Trigger system of ATLAS, in which the University of Geneva has a significant involvement, is currently under discussion. If this idea is implemented, the Tier-3 in Geneva will become an extension of the CERN Analysis Facility. Such a development would be beneficial not only for the Geneva group, but also for other Swiss ATLAS physicists, working in Berne. It would also be beneficial to the entire ATLAS collaboration, adding computing resources where they will be much in demand.

4 Middleware and Software

As shown in Table 1 all SAG clusters which are dedicated to LHC applications, have deployed the operating system (OS) Scientific Linux CERN 4 (SLC4). The general purpose cluster UBELIX is running Gentoo. It is of course not possible for SAG to enforce a particular OS on clusters not dedicated to LHC. The experience has shown that this causes signifcant additional work for every new release of the SAG application software. These are developed and validated on SLC4, hence, not guaranteed to work on other operating systems. The workaround is to create runtime environments with the necessary libraries. This overhead in the workload should be weighted with the gain in resources. At the moment the resource situation is very good, and the need for a hundred cores from the non-SLC cluster UBELIX is almost questionable. However, this clusters and others from the emerging SwiNG infrastructure will in the coming

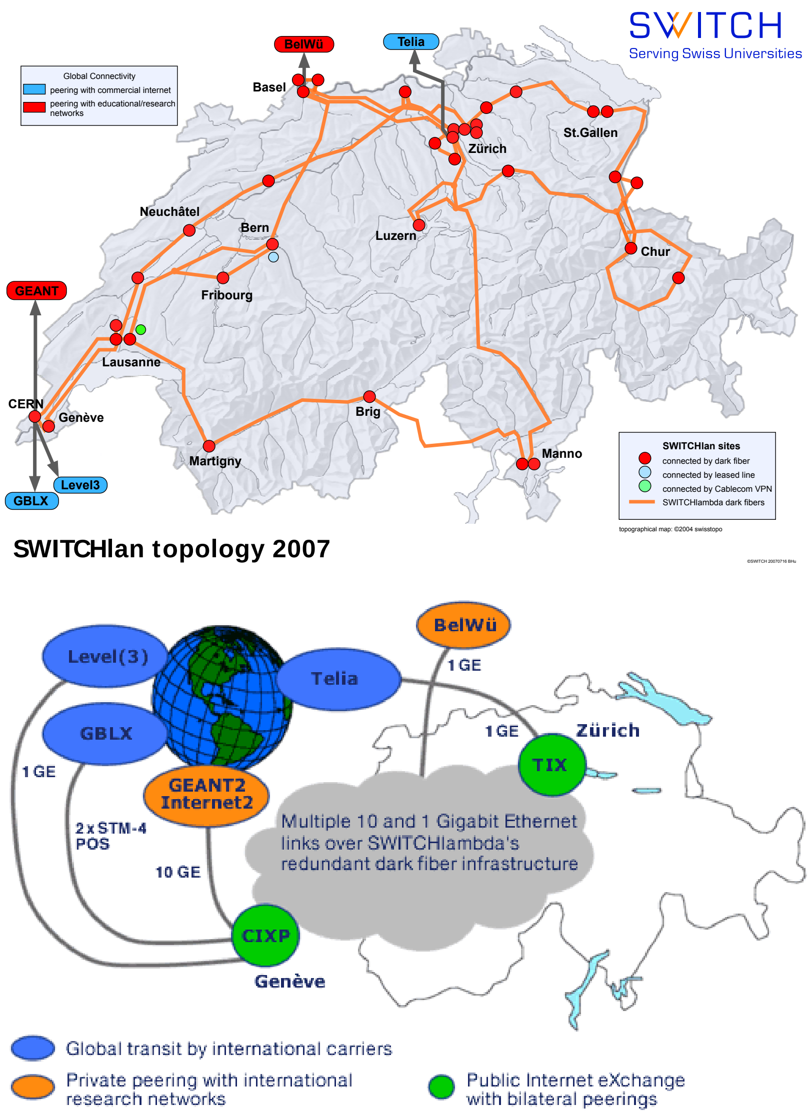

Figure 2: The upper figure shows the SWITCHlan Network. The network has a redundant topology of dark fibres currently operated with 10 Gb/s channels. The SAG sites are connected with 10 Gigabit Ethernet to the backbone. The lower figure shows the international connections of the SWITCHlan network. The current 1 Gigabit Ethernet connections can be leased at a higher speed if necessary [6].

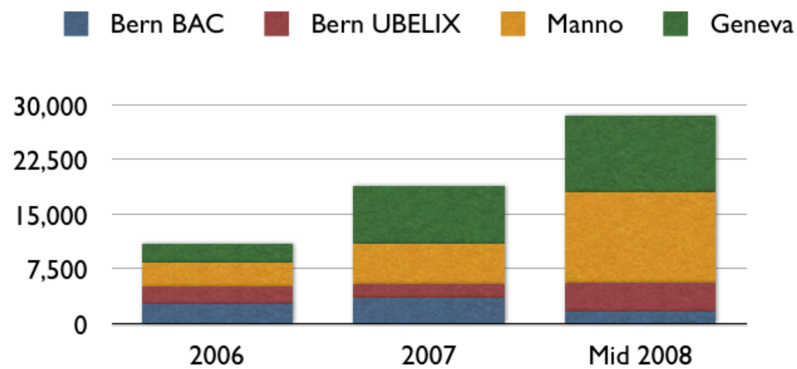

Figure 3: Wall time days on the Swiss ATLAS Grid. The numbers are taken from the accounting files of the batch systems.

years represent a considerable resource gain for the Swiss ATLAS Grid.

The software stack required for the ATLAS applications is relativly large and complex. Meanwhile one kit, which is the name of a precompiled release necessary to provide a runtime environment, occupies more than 7 GB disk space in almost 300 thousand files. Further roughly ten kits need to be available at the same time in order to serve a sufficent broad spectrum of ATLAS applications, not all using the newest software.

The software connecting the sites into a grid, the middleware, used on the SAG is mainly NorduGrid's Advanced Resource Connector (ARC). The T2 cluster has also the LCG middleware deployed. For smaller and not LHC dedicated sites, ARC is almost an inevitable choice. LCG is very manpower intensive, especially if not being deployed on Scientific Linux 4 which is the only supported platform. This has been the reason for the SAG to use ARC and not LCG on smaller and non-SLC sites. ARC does provide all needed features and also benefits from a small developer community with large flexebility and fast response.

The SAG do not consider the kits and the middleware as critical data since they can be reinstalled within a day or two. Consequently the sites decide themself if they prefer to back up this software.

5 Usage up to Mid 2008

Figure 3 shows the usage on the clusters in wall time days. Compared to 2006 the usage was approximately doubled to about 19000 wall time days in year 2007. This corresponds to about 0.5% of all the accounted ATLAS computing [7]. However, neither on the Swiss ATLAS Grid nor on the full ATLAS grid is all usage, e.g. interactive usage, accounted. Considering the mid 2008 numbers it seems that the usage will more than double in 2008. It is difficult to predict the usage for 2009 since measured data will be then be analysed. However, a non-linear increase is expected.

The accounting of the disk space usage is not very well implemented. Nevertheless tries the SAG to keep track of the free space on our resources. In July 2008 the situation looks satisfactory with more than 70 TB free disk space. This should be sufficient for the 2008 needs.

6 Evaluation of the 2008 Objectives

The 2008 goals, which were staked out in the charter of the SAG group, are repeated and commented below.

- *"Monitor, account and manage the ATLAS data transfer in Switzerland, i.e. the T1-T2and the T2-T3s transfers. Ensure dataset completeness, sufcient rates, LFC-dCacheconsistency. It includes work with ATLAS DQ2 system, FTS, LFC, SRM, GridFTP and dCache. Communication with T1, T2 and among T3s is required."* Some of the mentioned tasks are being taken care of by the T1 operators. However, the SAG working group needs to improve on these objectives. Appropriate actions have to be taken for the second half of 2008.

- *"Monitor and account the ATLAS jobs on the Swiss ATLAS Grid.It includes work with the batch systems, ARC front-end, LCG/gLite CE and deployment of ATLAS releases. Communication as for goal 1."* As was seen from Figure 3 this goal has already been met. Work can be continued without change in procedure.

- *Support local users, i.e. Geneva and Bern physicists,in issues related to 1 and 2.* This goal has been met by personal contact between the users in Geneva and S. Gadomski and between the users in Bern and S. Haug. Further wiki pages have been created for the SAG at CSCS and for the Geneva T3 at CERN [3][8]. The SAG also has some old wikipages at CERN [9]. These pages should updated, and merged or linked into one single location.

- *Enable 5% of the Tier 3 CPU resources to at least one "SwiNG application". Access will be given via the Advanced Resource Connector of the NorduGrid collaboration.* This goal will be addressed after the first successful tests of the first applications on the Swiss Multi Science Computing Grid infrastructure. These tests are expected in October.

- *Apart from uncoordinated group communitcation, the group meets virtually monthly, preferably the same day as the ATLAS FZK T1 meeting.* This goal has been met.

- *The group will reconsider its existencein the latest 2008 report to the SwiNG executive board.* This goal will be addressed in December.

7 Summary

The Swiss ATLAS working group manages a Grid system in which up to 1100 processors are potentially available. During the first six months of 2008 the Swiss ATLAS Grid has provided 28 thousand wall time days to the computing of the ATLAS experiment. Compared to 2007 our data processing capacity has doubled. The system is under continous development end testing, in order to meet the challenfges of real experimental data, expected this Autumn.

The activity has been reported in four talks at the CHIPP Winter School, the annual meeting of the Swiss Physics Society, at the SwiNG Scientic Council meeting, at University of Bern. Further the group has engaged in other SwiNG projects.

The main challenge for the next six months will be to provide reliable transfers of large quantities of data.

References

[1] ATLAS Collaboration, *Computing Technical Design Report*, CERN-LHCC-2005-022, ATLAS-TDR-017.

[2] S. Gadomski *et al.*, *The Swiss ATLAS Computing Prototype*, ATL-SOFT-PUB-2005-03, CERN-ATL-COM-SOFT-2005-07.

[3] https://twiki.cscs.ch/twiki/bin/view/LCGTier2/SwissATLASGridWorkingGroup.

[4] Enabling Virtual Organizations, http://evo.caltech.edu/evoGate/about.jsp.

[5] ARC meets SwiNG, www.lhep.unibe.ch/gridbern08.

[6] http://www.switch.ch.

[7] The ATLAS Dashboard, http://dashboard.cern.ch/atlas/.

[8] https://twiki.cern.ch/twiki/bin/view/Atlas/GenevaATLASClusterDescription

[9] https://twiki.cern.ch/twiki/bin/view/Atlas/SwissAtlasComputing