

# ATLAS@Home: Harnessing Volunteer Computing for HEP

C Adam-Bourdarios<sup>1</sup>, D Cameron<sup>2</sup>, A Filipčič<sup>3</sup>, E Lancon<sup>4</sup> and W Wu<sup>5</sup> for the ATLAS Collaboration

<sup>1</sup> Laboratoire de l'Accélérateur Linéaire, BP 34, F-91898 Orsay Cedex, France

<sup>2</sup> Department of Physics, University of Oslo, P.b. 1048 Blindern, N-0316 Oslo, Norway

<sup>3</sup> Jozef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

<sup>4</sup> CEA Irfu/SPP, 91191 Gif/Yvette Cedex, France

<sup>5</sup> Computer Center, Institute of High Energy Physics, CAS, 19B Yuquan Road, Beijing, China, 100049

**Abstract.** A recent common theme among HEP computing is exploitation of opportunistic resources in order to provide the maximum statistics possible for Monte Carlo simulation. Volunteer computing has been used over the last few years in many other scientific fields and by CERN itself to run simulations of the LHC beams. The ATLAS@Home project was started to allow volunteers to run simulations of collisions in the ATLAS detector. So far many thousands of members of the public have signed up to contribute their spare CPU cycles for ATLAS, and there is potential for volunteer computing to provide a significant fraction of ATLAS computing resources. Here we describe the design of the project, the lessons learned so far and the future plans.

## 1. Introduction

Volunteer computing is the concept of using spare computing cycles on computers when they are not in use to perform a computational task for someone else. People typically “volunteer” their computers for public scientific projects in order to contribute to the greater good (such as searching for extra-terrestrial life or large prime numbers). The first big volunteer computing project was SETI@Home [1], where a program installed on volunteers’ computers searched for evidence of extra-terrestrial life in radio signals from telescopes. Since then the volunteer community has grown to over 50 recognised scientific projects and hundreds of thousands of volunteers around the globe.

In parallel to the growth of volunteer computing, the computing resources required by high energy physics experiments for data processing is growing much faster than can be provided through traditional funding by national government research agencies. Increasingly, computing power is provided by opportunistic resources for which the experiments do not pay but are unreliable in nature. Volunteer computing is a prime example of opportunistic resources where volunteers offer their computing cycles for no charge, but there is no guarantee that a certain level of resources will be available at any given time.

The marriage of volunteer computing and high energy physics started in 2004 with the LHC@Home project [2] which was set up as part of CERN’s 50th birthday celebrations. Volunteers ran a program simulating the LHC beam to test the stability of the circulating



particles under a variety of conditions. The project proved much more successful than expected and is currently still active despite being planned to last only a few weeks. In 2011 a new version of LHC@Home dubbed Test4Theory [3] was launched to simulate collisions of particles in the LHC. This was one of the first projects to make use of virtualisation, enabling the software to run on many platforms without recompilation.

In early 2014 a volunteer computing project was started within the ATLAS experiment [4] for two purposes: to provide extra opportunistic resources and to involve the general public directly in ATLAS data processing. The latter reason is a very important part of volunteer computing, allowing people have a direct connection with scientific research. The rest of this paper describes the ATLAS@Home project in detail. In Section 2 the architecture and implementation of the project is described, then in Section 3 the experience and results so far are shown, and finally in Section 4 some conclusions are drawn.

## 2. Architecture and Design

The majority of volunteer computing projects use software called BOINC (Berkeley Open Infrastructure for Network Computing) [5]. The software comprises a server, which hosts tasks or *work units* to be processed and a client which volunteers install and configure to pull and run work units from specified projects. Once a work unit is processed the client sends the result back to the server which validates the result, and if the result is good awards credit to the volunteer. Credit is simply a measure of how much computation has been done and has no monetary value, however it provides motivation for many volunteers.

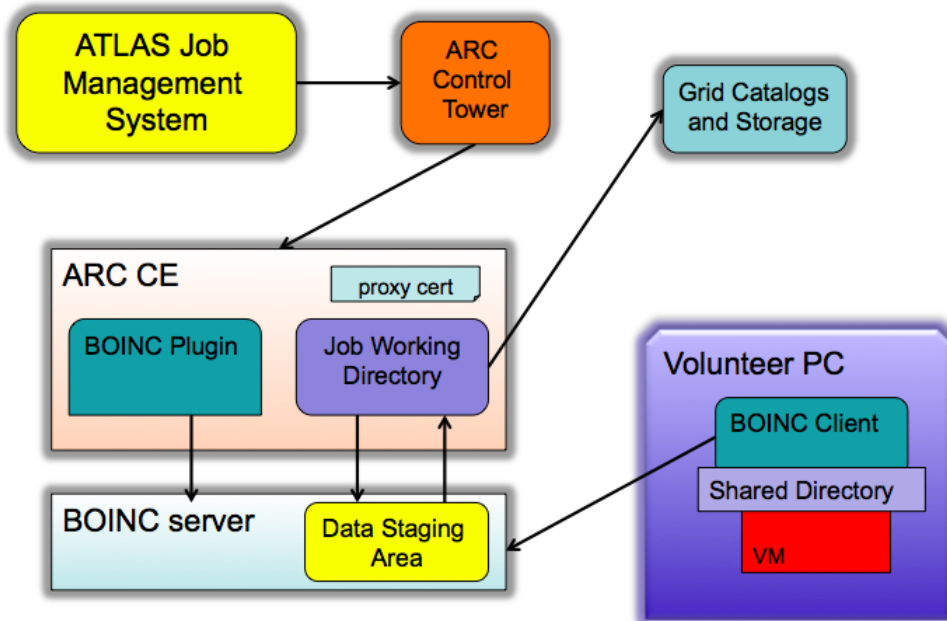
In the early years of volunteer computing the work units were executable programs compiled for the types of operating systems run by volunteers. A technique pioneered by Test4Theory is the use of virtualisation to allow software to be compiled for one operating system but run on many. For each work unit the BOINC client boots a *Virtual Machine* (VM) which runs the required operating system and the work unit is processed inside the VM. In addition to easing the life of application developers this provides an extra layer of security by containing applications in an environment isolated from the volunteers' main operating system.

When deciding the type of work units to run in ATLAS@Home there were several considerations to take into account. Firstly the unreliable nature of the resources means that only low-urgency tasks can be run. In contrast to regular Grid sites running ATLAS tasks the volunteers may be located far from the data on a poor Internet connection which means that tasks processing large amounts of data are also not suitable. The type of tasks which fit best these requirements are low priority *Monte Carlo Simulation* tasks. These tasks involve generating random collisions inside a simulated ATLAS detector, and modelling their trajectory and interactions with the various detector components. This is a vital part of the overall process of analysing data from ATLAS as it provides both detailed information on the detector itself and a means of comparing the observed data against theoretical models.

On a technical level there were further issues to consider in the design of ATLAS@Home. The software used for Monte Carlo simulation is extremely complex and only runs on specific operating systems seldom used by the general public and so virtualisation is the only way to attract significant numbers of volunteers. The CERNVM project [6] provides virtual images tailored for the LHC experiments' software and these images can run seamlessly inside the virtualisation layer provided by BOINC. ATLAS software is provided to the VM through the CERNVM File System (CVMFS) [7], a remote read-only filesystem using aggressive local caching which is mounted inside the image. To avoid downloading the software every time the VM is started, the CVMFS cache inside the image is pre-filled with the required software, by running an example job, saving a snapshot of the image, and using that snapshot as the final image to distribute to volunteers.

One critical requirement was that no ATLAS credentials should be distributed to volunteers.

In regular Grid jobs a proxy certificate is passed through to the worker node on which the job runs, so that the job can communicate with the central workload management system and stage data to and from Grid storage. Having this certificate accessible to volunteers would provide access to and even the ability to delete ATLAS data on the Grid which is obviously unacceptable. The solution was to use the model deployed in NorduGrid [8] and other environments such as High Performance Computing centres which have restricted access to the outside world from the job worker nodes. In this model the ARC Computing Element (ARC CE) [9] takes care of data staging before and after the job runs and the ARC Control Tower [10] provides the link with the ATLAS workload management system, PanDA [11].



**Figure 1.** Architecture of ATLAS@Home.

The overall architecture is shown in Figure 1. Jobs which are assigned to ATLAS@Home by PanDA are picked up by the ARC Control Tower, and sent to an ARC CE connected to the BOINC server. ARC CE copies the required input files from Grid storage to a staging area inside the BOINC server. ARC CE supports many batch systems and a new plugin for a BOINC “batch system” was written to allow injection of jobs as work units in the BOINC server. At this point the job is ready to be picked up by a BOINC client running on a volunteer’s PC. This client only has access to the BOINC server data staging area and no access to Grid storage or Grid credentials and so there is no chance of accidental or deliberate tampering with ATLAS data.

Once the BOINC client has picked up a work unit and downloaded the necessary input files it starts a VM. The VM image itself is only downloaded once and reused for each work unit but each work unit requires of order 100MB of input data which is downloaded each time. When the image starts it automatically sets up the ATLAS software environment and starts the ATLAS job. The input data resides in a directory shared between the host and the VM. The result of the job is also written to this shared directory and once finished the BOINC client uploads it to the BOINC server staging area. ARC CE then uploads the data from there to Grid storage.

As the jobs are running on untrusted resources, it is very important to ensure that bad results are not uploaded to Grid storage and eventually used by physicists. There are therefore several validation steps in place. Firstly each result returned by a volunteer is checked to make sure the

expected files and filenames are present. If not the job is failed and the volunteer receives no credit. At the end of each job a small metadata file is created with information on each output file such as size and checksum and after ARC CE uploads the file to Grid storage the checksums are compared to those in the metadata file. If they differ the job is rejected and retried.

A clever volunteer could of course tamper with the metadata file to upload junk data, however at the next stage of the ATLAS task the results of the ATLAS@Home simulations are merged together into larger files. Typically 10 results are merged into 1 file and this merging is done on regular Grid resources. As part of the merging each file is checked that it contains sensible data and if not the merge fails and the results are not used. A really clever volunteer with intimate knowledge of the ATLAS software could still in theory modify the event data to produce for example fake tracks in the detector but to mitigate against this, large-scale validation of the physics results is done comparing against similar tasks run on regular Grid sites. In many BOINC projects each work unit is sent to a few different hosts and the result is only considered valid if the results match byte for byte. However due to various random factors inside ATLAS software running the same job twice can never produce exactly the same result.

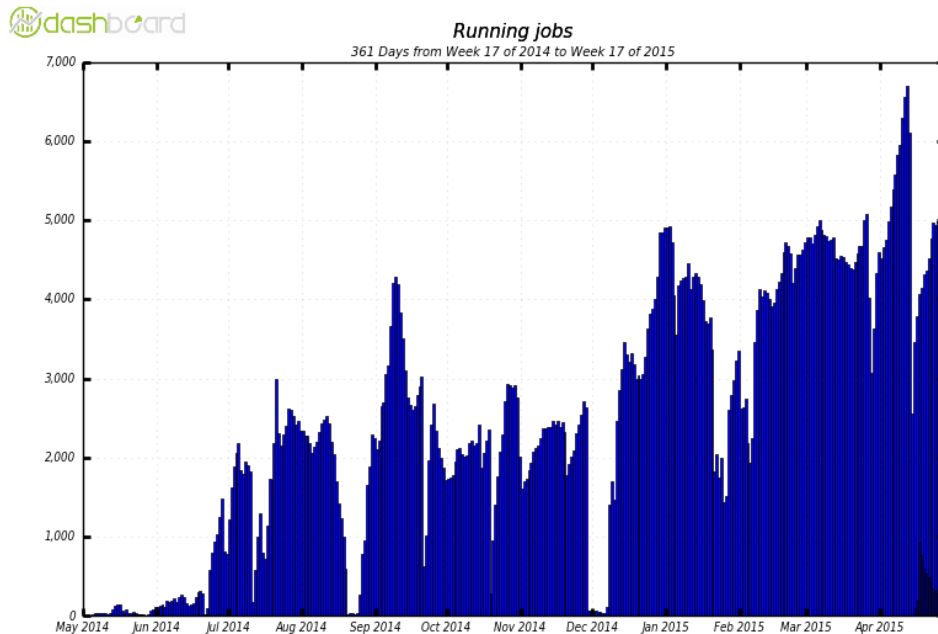
### 3. Experience and Results

Work on ATLAS@Home started early in 2014 with a test server in Beijing. Even though this server was not advertised the BOINC community found out about it and several volunteers registered and started running work units. The first interactions with those volunteers showed a surprising amount of enthusiasm in the community for the ATLAS@Home project and for CERN experiments in general. After quickly reaching the infrastructure limits of the test server, a new server was set up at CERN with the help of the CERN IT department in summer 2014 and the existing volunteers were migrated there. A proper website and an alias, atlasathome.cern.ch, were created and soon after ATLAS@Home was added to the list of official BOINC projects. This greatly increased the visibility of the project within the existing BOINC community.

As the number of volunteers increased some scalability issues appeared, mainly related to the input and output data handling. For example for each job the input data file was copied to the BOINC server staging area but as the same file is used by many jobs this was changed to create hard links instead of copying, which greatly reduced the pressure on the disks. Around October 2014 ATLAS@Home was moved to the BOINC platform used by the other CERN volunteer computing projects, thus removing the need to maintain a separate BOINC server and database. In addition a large storage area was provided which allowed extra headroom to expand the volunteer base.

Figure 2 shows the number of running jobs in ATLAS@Home over the period of around one year, as seen by the ATLAS job accounting dashboard. A big jump can be seen in July 2014 corresponding to the point where ATLAS@Home was made an official BOINC project and from that point until the end of the year an average of between 2000 and 3000 running jobs were continuously running. At the start of 2015 this jumped up to 5000 jobs and until the time of writing the number of jobs stayed around this level. The sharp increase in early April 2015 is probably due to the publicity surrounding the restart of the LHC for Run-2. The dips in the figure were all due to technical issues and not due to lack of volunteers.

During the year shown in the figure the volunteers simulated over 20 million events in total corresponding to almost 200 CPU years of processing. Throughout the first few months of 2015 ATLAS@Home consistently ranked in the top 3 Grid sites running ATLAS jobs, however it should be noted that an ATLAS@Home CPU core is not the same as a regular Grid site CPU core. A “running” job may actually be paused while the volunteer is using the PC or has switched it off and many volunteers configure BOINC to use less than the full CPU capability for power saving reasons. Traditional Grid resources are quantified by fixed benchmark numbers but due to the wide variety of volunteer resources and the above reasons it is impossible to



**Figure 2.** Running jobs on ATLAS@Home from May 2014 to April 2015.

obtain such numbers for ATLAS@Home. Even so the capacity of ATLAS@Home is at the time of writing roughly equivalent to a large Tier-2 Grid site.

The experience dealing with volunteers has been both rewarding and challenging. Many volunteers expect a certain level of support in return for donating their CPUs and take the allocation of credit very seriously. One thing learned early on is that volunteers should always be given credit for their work units no matter what. This led to several changes in the workflow, for example still giving credit even if a work unit failed through a problem in ATLAS software, and never cancelling ongoing work units even if the jobs were cancelled by PanDA. The BOINC server comes with a web message board which acts as the main communication channel between volunteers and administrators. The ATLAS@Home message board is very active with several messages per day, but the volunteers tend to help each other a lot so not every message requires administrators to answer.

As stated earlier, one of the two main reasons for starting this project is to engage the public in the science that ATLAS is exploring. A public outreach page [atlasphysathome.web.cern.ch](http://atlasphysathome.web.cern.ch) exists to inform the volunteers of the science behind the simulations they are running and also to give some background on the project and the people behind it. At the time of writing there has been no large campaign to attract volunteers and most have come through word of mouth and message forums in the BOINC community, or are from the ATLAS community. The outreach pages will act as the entry point for interested volunteers once such a campaign is launched but before that there are a few improvements to make to the system. The most important is to provide some visual feedback to the volunteers, showing them what the events they are simulating look like in the ATLAS detector. In addition to encourage people to stay in the project, BOINC features such as challenges (goals for volunteers or teams of volunteers to reach in a limited time) and badges (rewards for reaching milestones such as number of work units run) will be set up.

The potential for ATLAS@Home lies not only in the general public, but in spare computing resources at universities, labs or other institutes already involved in ATLAS. Many people in

these places leave desktop computers running all night and ATLAS@Home could run on those machines outside working hours. These institutes often have small computing clusters for local use and as an alternative to setting up the infrastructure necessary to connect them to the Grid, installing BOINC and connecting to ATLAS@Home could be a much easier way to run ATLAS processing. To facilitate this, part of future work will involve creating standardised installation mechanisms to manage ATLAS@Home on clusters of computers.

#### 4. Conclusion

In this paper the ATLAS@Home project has been presented. In little over a year it has grown to provide a significant fraction of computing resources for ATLAS simulation and has attracted thousands of volunteers. Even though the volunteers come for free, the manpower required to set up and maintain ATLAS@Home is not negligible, however it is a fraction of that needed for a regular Grid site. There is huge potential for ATLAS@Home to make up a significant part of ATLAS computing resources but further work must be done to increase the appeal to the general public.

#### Acknowledgements

We wish to thank our colleagues in the CERN IT department for providing the hardware and support for integrating ATLAS@Home into the LHC@Home infrastructure. We would also like to thank the developers of BOINC for their helpful responses to our questions and rapid implementation of feature requests.

#### References

- [1] Korpela E, Werthimer D, Anderson D, Cobb J and Leboisky M 2001 *Computing in Science & Engineering* vol **3** pp 78–83
- [2] Herr W, McIntosh E, Schmidt F and Kalchev D 2006 *Proc. 10th European Particle Accelerator Conference*
- [3] Lombrana González D, Karneyeu A, Buncic P, Segal B, Grey F, Blomer J, Harutyunyan A, Skands P and Marquina M 2012 *Proc. Int. Symp. on grids and clouds (PoS(ISGC 2012)036)* pp 36–49
- [4] ATLAS Collaboration 2008 *J. Inst.* **3** S08003
- [5] Anderson D 2004 *Proc. 5th IEEE/ACM Int. Workshop on Grid Computing, GRID 04* pp 4–10
- [6] Buncic P, Sanchez C A, Blomer J, Franco L, Harutyunian A, Mato P and Yao Y 2010 *J. Phys.: Conf. Series* vol **219** p 042003
- [7] Aguado Sanchez C, Bloomer J, Buncic P, Franco L, Klemer S and Mato P 2008 *Proc. XII Advanced Computing and Analysis Techniques in Physics Research (PoS(ACAT08)012)* p 52
- [8] The NorduGrid Collaboration web site URL <http://www.nordugrid.org>
- [9] Ellert M *et al.* 2007 *Future Gener. Comput. Syst.* vol **23** pp 219–240
- [10] Filipčič A on behalf of the ATLAS Collaboration 2011 *J. Phys.: Conf. Series* vol **331** p 072013
- [11] Maeno T on behalf of the ATLAS Collaboration 2008 *J. Phys.: Conf. Series* vol **119** p 062036