

Table of Contents

pre-GDB on Volunteer Computing, CERN, November 11, 2014.....	1
Agenda.....	2
BOINC - N Høimyr.....	3
/XWHEP - O. Lodygensky.....	5
LHC@home - M. Giovannozzi.....	6
BOINC and CERNVM - B. Segal.....	7
Experiment Experiences.....	8
ATLAS - D. Cameron.....	8
CMS - L. Field.....	9
LHCb - F. Stagni.....	10
Discussion.....	10
Conclusions.....	13

pre-GDB on Volunteer Computing, CERN, November 11, 2014

Agenda

<https://indico.cern.ch/event/272793/>

BOINC - N Høimyr

Reason to use volunteer computing in HEP

- Volunteers: uncoordinated, unpledged/opportunistic, allow to engage with new people/communities
- Opportunistic use of institution desktop resources: coordinated, unpledged/opportunistic
- Small to medium cluster: coordinated, pledged, easier to deploy than grid MW

BOINC: main MW for volunteer computing.

- To be a volunteer with BOINC: 3 steps on Linux
 - ◆ Get BOINC client (RPM with BOINC 7.2.x on EPEL)
 - ◆ Get VirtualBox (installed by BOINC client on Windows)
 - ◆ Attach a project with boinc-client utility (user/password or key for silent connections on servers)
- Designed not to disturb the user: many (optional) features to ensure that job is suspended/killed when the machine is used by its main user
- Provide a GUI to manage the client
- Core client is a daemon

Key concept to motivate volunteers: credits

- Based on host benchmark, allow to calculate the contribution of a user to a project
- Credits collected from multiple projects
- Statistics produced exposing the credits of each user
- The 2 main projects (SETI@home and Einstein@home) typically operate with a simultaneous capacity of 600 TFlops/second

Use of virtualization (VirtualBox) pioneered at CERN in 2010-11 with Test4Theory

- Now mainstream BOINC code
- Several projects outside CERN using it now
- Classic BOINC still usable but more difficult: unpredictable client environment (Windows, Mac, Linux, Android..), labour intensive application building/testing.
 - ◆ At CERN, only SixTrack uses classic BOINC
- When using virtualization, BOINC acts as a VM dispatcher to volunteer clouds.
- @CERN, images based on CERNVM + CVMFS
 - ◆ Interaction with experiment frameworks (PanDA, DIRAC...)

Volunteers vs. institutional desktops/clusters

- With volunteers, publicity and marketing are essential when they are less relevant for institutional resources
- More controlled environments in institutional resources
- Volunteers/desktops: CPU-intensive apps mainly, as I/O is an issue outside our labs.

BOINC @CERN: now a generic service supported by IT/PES (0.5 FTE).

- BOINC servers (LHC@home) for CERN-related projects
 - ◆ Relying on DB on Demand for BOINC MySQL DB
- Application support: BOINC server configuration, upgrades operations and monitoring
- Using central NFS filer service
- Drupal portal <http://lhcatome.cern.ch>

Planned improvements

- Scalability: split BOINC server into several servers to scale out on upload/download, Ceph for data buffers
- Tools to automate app deployment
- Security: https rather than http, investigate alternative login methods (e.g. OpenID)

/XWHEP - O. Lodygensky

XtremWeb [↗](#) started as a project to support multiple projects on the same computing infrastructure and offering some authentication/security features

- Mainly targeting desktop grids and other environments. Uses push, rather than the pull approach of BOINC. No need for outreach and forums!
- Since 2006, forked as XWHEP, the actively developed production version based on XtremWeb.
- Many new features added in XWHEP for authorization and data management.
- 3-tier application: user wanted to run apps, the XWHEP server and the volunteer resources.
 - ◆ P2P direct communication possible between the app user and the volunteer resources.

Authorization based on user profiles: administrator, group administrator, standard user, worker user

- Standard user can submit apps, manage its private apps and retrieve results
 - ◆ A private app, until certified by a group admin, can be seen/used only by the user who submitted it
- Worker user: used only to deploy the worker, no other right
- Access rights, similar to Unix files, associated with all objects

Virtualization supported by XWHEP: users can submit apps with a VirtualBox image

- Based on XWHEP feature allowing volunteers to say the apps they support and ask payload requiring this app
 - ◆ Not restricted to virtualization support, can also be arbitrary apps
- First implementation is VirtualBox but will be used to implement other kinds of virtualization (e.g. libvirt)
- Implements VM contextualization based on HEPiX WG proposal
- Demonstrated the ability to deploy CERNVM + ROOT
- P2P feature used to give SSH access to VMs
- No access allowed to the local network (protection) but unrestricted access to internet
- Direct communications possible between different volunteer resources if allowed by volunteers
- Several different kinds of apps using the platform

XWHEP has both a portal and REST API

XWHEP using several medium-size infrastructure in best-effort mode

- GRID5000: French computer science infrastructure
- Qarnot: startup company providing heating systems for appartements based on computers

Question: Is XWHEP used at LAL, CNRS and IN2P3? LAL, yes, not IN2P3, but at EPFL in Switzerland and also in a lab in Tunis.

LHC@home - M. Giovannozzi

SixTrack⁷ simulations: studies about field quality in superconducting magnets

- Critical to understand the stability of the particles of the beam and predict beam lifetime and emittance growth
- Easy to split a simulation into multiple independent jobs
- Used originally during the design phase of the LHC magnets
- Now used during operations to compare with original simulations and refine understanding of beam with the measured magnet parameters
- Now used for HL-LHC studies

Key issue: numeric compatibility of results obtained on heterogeneous architectures

- Now solved in SixTrack

Attempt to keep track of volunteers and acknowledge them in all articles

- Also on LHC@home website

Future work

- Reduce the length of jobs: split jobs into smaller ones, better for efficiency
- Reduce size of output files
- Possibility of performing collimation studies

BOINC and CERNVM - B. Segal

Original challenge: run real LHC physics on BOINC

- Huge codebase, strict OS requirements

Early tests with virtualization (VMware) done in 2006 as a proof of concept but very far from a production solution

2008: CERNVM developed and became an enabling technology

- Solved the image size problem of previous prototype

2009: CoPilot added, solving the VM credential problem and allowing integration with experiment frameworks.

Philosophy: volunteer cloud

- BOINC is only used to start VMs
- Everything else done by experiment frameworks

Launched to the public in 2011 as LHC@home 2.0 Test4Theory (press reports etc) Many signed up, only a fraction continued running successfully.

No marketing/promotion effort since then but able to run a steady 2500 VMs per day

- Not at the same scale as SixTrack

Now called [vLHC@home [↗](#)] (Virtual LHC@home.)

- Aim to attract other use cases from other experiments

Experiment Experiences

ATLAS - D. Cameron

Motivations

- (Almost) free resources
- Public outreach

Design considerations

- Low priority CPU-intensive jobs, typically non-urgent MC jobs
- Virtualization environment for ATLAS SW, using CVMFS
 - ◆ Consequence: need a different image per version of ATLAS SW. Happens once every few months.
- No (grid) credential on volunteer host: ARC CE used for staging data and interacting with data catalogs
 - ◆ Grid credentials not leaving the ARC CE
 - ◆ All jobs using the same credential: ATLAS production
- Apart from these specificities, looks as a regular PanDA queue
 - ◆ ARC CE insert the job in BOINC queue
 - ◆ VM image to use is pre-loaded to the BOINC server: all jobs using the same image
 - ◆ ARC CE and BOINC server on separate machines with a NFS shared file system

Work started early this year in IHEP (Beijing)

- July: ARC CE + BOINC server moved to CERN
- Recently: combined ARC-CE and BOINC server replaced by a virtual LHC@home BOINC server, separate from ARC-CE.

Simulation job profile

- Full Athena jobs, 50 evts/job
- Image is 1.1 GB (500 MB compressed): CERNVM + pre-cached SW
- input tarball is 1-100 MB
- Output ~100 MB
- Memory requirement : 2 GB/job
- Some simple validation done at the end of the job
- Physics validation done later

Issues found

- Volunteer: entry barrier quite high, many volunteers never run a job
 - ◆ Requirement: 64-bit OS, 4 GB, decent bandwidth
- Unstability/failing jobs: people not credited and thus moving away
 - ◆ Virtualization/VMwrapper causing a lot of problems
 - ◆ Firewall at volunteer PC preventing access to condition data....

Current status

- Running 2000-3000 jobs continuously
- 75% successful jobs
- 50% efficiency

- 28th simulation site in ATLAS

Very active message board: eating a lot of time to answer...

- In particular volunteers asking why they fail job or are not credited for their contribution

Lessons learned

- Requires a lot of effort, in particular the interaction with volunteers
 - ◆ A "community follower" is needed
 - ◆ Some competent volunteers can help others
 - ◆ It's why it is not completely free resources...
- Number of running jobs has reached a plateau: due to current setup, technical reasons understood and solutions being discussed with CERN IT
 - ◆ Mostly an issue of scale out (adding new boxes) + a few adjustments in the workflow (increased caching)
- Main problems caused by vboxwrapper but BOINC developers very enthusiastic and helpful
- Not ready yet for pushing ATLAS@home in production but this is the goal
 - ◆ Could be an alternative to a grid CE for small sites
- Outreach: ATLAS has prepared a dedicated outreach site, and also plan "badges" for volunteers who contribute a lot of CPU. No active campaign has started yet, first the infrastructure needs to be scaled to handle a high number of jobs.

Discussion about pushing up the outreach effort

- Infrastructure not yet ready but clearly needed to expand the number of volunteers
 - ◆ Involve outreach team at CERN? Make much more visible the need for contributors
- Main effort is to build the vLHC@home platform as a a platform usable by all experiments: not yet there
 - ◆ This coordinated outreach effort should be done at the platform level rather by each experiment

CMS - L. Field

Test4Theory model

- BOINC server submits a job wrapper that is in charge of starting and stopping the VM
 - ◆ Stopping after 24h to get the contribution credited as it happens only at the end of jobs
- VM contains an agent that contacts CoPilot to get the actual job and the data (data bridge)
- Challenge: CoPilot support
 - ◆ Relies on no longer maintained dependencies

CMS took the same model but with a different implementation

- HTCondor used to push jobs to BOINC: backend needs to be written
- Authentication using standard web server authentication
 - ◆ BOINC puts user credentials in a MySQL DB and push them to VM through /dev/fd0
 - ◆ mod_auth_ssl (Apache) used to authenticate against BOINC credentials
- Data: http-based data bridge, allowing to use http dynamic federations
 - ◆ DataBridge is no more than http Dynamic Federation + mod_auth_sql to authenticate against BOINC
 - ◇ No development needed
 - ◆ Monitoring LHCb tests in Canada, promising results so far
 - ◆ Great scalability potential due to the redirect capabilities and the use of an S3 backend

- ◆ Asynchronous stage-out after job completion through FTS built upon standard behaviour/feature of CMS jobs
- BOINC is just used to instantiate VM: could technically leave without it but important to keep it for the community of volunteers
 - ◆ BOINC is the main way to attract resources

Basic workflow, based on DataBridge, can be generalized

- PUTjob description in data bridge
- PUTdata in data bridge
- Create a job agent with input and output parameters
- Async. stage-out of output bucket
- DataBridge could be used also for storage-less IaaS resources

Status: still an early prototype, work started recently

- Too early for reporting results

LHCb - F. Stagni

In fact was the first LHC experiment to start with BOINC but have not gone as far as ATLAS and CMS

- Several generation of early prototypes done and abandoned in 2013 and 2014

Started with the very simple of VMs contacting DIRAC but not possible to use it in production in the volunteer computing environment

- Requires credentials on the volunteer PC

Several issues found like job length and resource volatility: solutions applicable to other environments too

- Job masonry to use the most of an available slot: running short MC to fill the slot if needed
- Smaller output: partly results from shorter job but also splitting simulation and RECO in different jobs
- Some issues not solved yet: job monitoring, accounting
- CVMFS: no pre-caching, depend on volunteer location, suffering lack of Shoal support until now
 - ◆ May consider doing the same as ATLAS: images specialized for BOINC, with pre-cached SW
- grid credentials: no solution adopted yet, very interested by DataBridge

Discussion

- E. Lançon: importance to have a BOINC server in charge of managing volunteers to avoid exposing the pilot framework to (malicious) volunteers
 - ◆ LHCb also convinced of this!

Discussion

Helge: is calling for "institutional" volunteers energy efficient? Not convinced. May be better to set up a small dedicated cluster...

Image pre-caching: probably no alternative as VM must be restarted very regularly to get contribution credited

- Without caching, will have to pay the price of first download of application from CVMFS at every instantiation
- Optimizing the cost of image management: may be worth to evaluate the real impact of using a SW version different from the one pre-cached
 - ◆ Experience at grid sites shows that there is only a very minimal amount of different files between releases

Currently 3 different BOINC projects: SixTrack, vLHC and ATLAS

- Still under discussion whether merging them is a goal or not: trying to experiment with app-level credits in BOINC. Merging forums is not an option.)

Michel: avoid the word "desktop grid", tends to designate both desktops and clusters but they are 2 very different resources

- Institutional desktop remains a desktop, i.e. an untrusted environment
- L. Field: must concentrate on volunteer computing (rather than opportunistic usage) as the real challenge is attract and manage volunteers

Volunteers: how to get more of them? Do we really want more?

- Must understand the "price" of volunteers in term of support, manpower... Must find the right balance. There is an initial investment, but after that you get resources with little cost compared to renting capacity e.g. at Amazon.
- Institutional desktops: low hanging fruits, BOINC client installed by default, not caring about credits
- Costs of running desktops for volunteer computing in HEP institutions compared to dedicated clusters
 - ◆ Several studies/presentations on this have been made, e.g. at BOINC workshops on this but difficult to get a precise idea.

Opportunistic resources: volunteer computing is not necessarily the appropriate response as they are dependable resources seen as volatile resources

- VAC or clouds may be better alternatives
- L. Field: the real showstopper for a technology to be used as a provisioning method for WLCG T2s is the possibility to do accounting
 - ◆ Still not understood how to do proper accounting in cloud and volunteer computing world
- A. McNab: running multiple jobs in a VM is not that different from running multiple jobs in a pilot job...
- Agreement that we must concentrate on the volunteer computing as the first goal it to get additional ("free") resources, as an alternative to commercial cloud providers.
 - ◆ Probably a similar effort is necessary to to adapt our application environment to support the use of commercial clouds but the volunteer computing resources are free.

F. Stagni: can we say the level of resource we must be able to provision from the volunteers to be worth the effort?

- L. Field: ATLAS said 500 usable cores was a minimum
 - ◆ O. Lodygensky: remember that 10% of volunteers actually providing resources is already a very good result

O. Lodygensky: may be interesting to look at 3GBridge developed by EDGeS/EDGI projects and allowing to bridge between many different computing infrastructures, including BOINC and grid

Outreach: clearly need a meeting similar to this one between people involved in this effort around volunteer computing. F. Grey: We are planning to have a meeting on Citizen participative in science, e.g. volunteer computing, crowdsourcing etc.

Conclusions

A lot of progress has been made over the last 6 months, and notably ATLAS has shown that volunteer computing has real potential for experiment software. BOINC as a vehicle to distribute CernVM images getting applications via CVMS is a common denominator for all projects. Volunteer Cloud can be seen as an extension of the cloud model of the experiments, and we should work together to share experience and use as many common components as possible.

Use [project-lcg-gdb-clouds-wg@cernNOSPAMPLEASE.ch](mailto:project-lcg-gdb-clouds-wg@cern.ch) mailing list for further discussions on the technology and cloud related matters (CernVM, CVMS etc) + [project-lhcathome@cernNOSPAMPLEASE.ch](mailto:project-lhcathome@cern.ch) for things specific to the LHC@home platform.

Probably another similar meeting next spring.

This topic: LCG > GDBMeetingNotes20141111
Topic revision: r4 - 2014-11-13 - NilsHoeimyr



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback