PROCEEDINGS OF SCIENCE

# Experience in Grid Site Testing for HEP with HammerCloud

**Ramon Medrano Llamas**[1]
*European Organization for Nuclear Research*
*CERN CH-1211, Geneve 23, Switzerland*
*E-mail:* ramon.medrano.llamas@cern.ch

**Daniel Colin van der Ster**
*European Organization for Nuclear Research*
*CERN CH-1211, Geneve 23, Switzerland*
*E-mail:* daniel.colin.vandester@cern.ch

**Johannes Elmsheuser**
*Ludwig-Maximilians-Universitat Munich*
*Geschwister-Scholl-Platz 1, 80539, Munich, Germany*
*E-mail:* johannes.elmsheuser@physik.uni-muenchen.de

**Federica Legger**
*European Organization for Nuclear Research*
*Geschwister-Scholl-Platz 1, 80539, Munich, Germany*
*E-mail:* federica.legger@physik.uni-muenchen.de

**Andrea Sciaba**
*European Organization for Nuclear Research*
*CERN CH-1211, Geneve 23, Switzerland*
*E-mail:* andrea.sciaba@cern.ch

**Mario Ubeda Garcia**
*European Organization for Nuclear Research*
*CERN CH-1211, Geneve 23, Switzerland*
*E-mail:* mario.ubeda.garcia@cern.ch

HammerCloud is a grid site testing service for the ATLAS, CMS and LHC*b* experiments centered at CERN in Geneva. This tool, which is provided as an online service for operation managers, site administrators and, in general, grid experts, allows them to perform on-demand tests of their computing facilities in order to validate and measure their performance. In addition, HammerCloud runs automated tests to check the availability and reliability of the sites under different circumstances. The tests consist of real analysis code provided by the physics community to ensure real-world use cases for the grid sites. Indeed, HammerCloud has been employed in HEP for more than 2 years and has helped increase the performance and reliability

[1] Speaker

 

seen by the grid users. In this work we will present the lessons learnt while deploying, optimizing and evolving the system for the three VOs and the development plans for the near and mid-term future.

*EGI Community Forum 2012 / EMI Second Technical Conference*,
*Munich, Germany*
*26-30 March, 2012*

 

## 1. Introduction

Grid computing is a term usually referred to a federation of computer resources, that are not centralized, to deliver nontrivial qualities of service [1]. To deliver this nontrivial quality of service, testing has to be implemented in all the phases of the software development and operation of grid sites.

### 1.1 Grid functional testing

Testing is crucial in grid operations, since it ensures that all the components that form grid sites are performing as expected and with the functionalities that users require. Testing is being performed in all the steps and phases of development and operation of grid software, for instance:

- During the software development cycle, unit testing techniques are applied.
- When releasing software, integration testing, such as nightly build testind are being run.
- While operating sites, using metodologies like DevOps to backfill the developers with feature requests, performance improvements and bug reports.

There are two more strategies to test a site, that are user-centric: functional testing and functional monitoring. Theses techniques allow to test the test from a user wise perspective, that means that the testing tool should mimetize a user in order to have the same interface to access the resources and the same requirements, such a software version running in the site or access to specific datasets.

Functional monitoring is based on the metrics and data collectec by the functional tests, and allows to measure the site from a user perspective, changing administrative metrics, like “throughput of the storage element”, that can be less meaningful for a final user, by high level metrics, like “time to access hits dataset” or “events per second”, that measures the site as a black box.

This monitoring technique measures the grid sites while they are in production and allow site administrators to improve their sites “on live”. Also, this functional monitoring metrics allow to automate some of the procedures that the site operators (*shifters*) performn in case of problems or of some predetermined events.

## 2. HammerCloud

HammerCloud [2] is a functional testing engine for grid sites. Since it is performing functional testing, it behaves like a user, submitting to the grid sites jobs that perform real analysis or production tasks in the case of the HEP virtual organizations (VOs).

The principal users of HammerCloud are grid experts or site administrators, which can test their sites in two ways: scheduling stress tests or having functional tests running on them.

A stress test is a one-time test designed to benchmark the site, this is useful while performing upgrades or optimization to the site, in order to measure its impact. On the other hand, functional tests implement the availability monitoring described before, and they are tests

that are running continuously on the sites. The purpose of these tests is more focused on monitoring the availability and functionality of the sites than the pure benchmarking of the stress tests.

HammerCloud was originally developed inside the ATLAS Collaboration at the LHC to test the grid sites that were working with the collaboration. Shortly, other experiments of the LHC showed interest in having such a testing infrastructure: CMS to replace a similar project called JobRobot, and LHC*b* to integrate testing on the DIRAC framework.

Having three different codebases, infrastructures and operations was not an acceptable option, so in 2010, HammerCloud was rewritten to be a VO-independent tool. Actually, and although now is being used for HEP-centric testing on grid environments, the only requirement of the system under test is to be a job-oriented system; like batch systems or the grid itself, but testing can be performed on non-HEP VOs or non-grid systems, like a local batch system or resources residing the in the cloud. To have a new VO, a simple plug-in for HammerCloud is needed.

HammerCloud is different approach to existing monitoring systems deployed in the grid, such as Nagios, SAM or Dashboard, since it performs active testing by submitting actual jobs to the sites using the whole tool chain that a user would do. Also, HammerCloud is able to interface those systems (particularly Nagios) to provide information about the testing status. The overhead to be running jobs in production sites depends on the size of the site, in terms of job slots. Currently, a set of 4 jobs is running in each site tested will run in parallel. For a site with 10,000 job slots this mean a 0.04% of capacity used on functional testing.

### 2.1 HammerCloud as a common solution

The rewrite performed on HammerCloud in 2010 suppose a new architectonic approach to design a tool that can be generic enough to support any kind of VO to get tested.

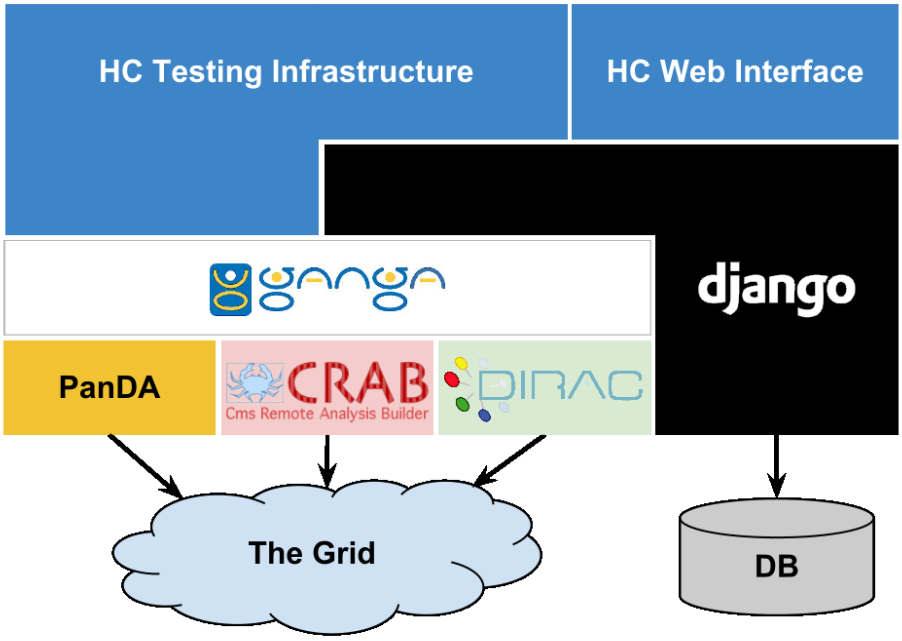

**Figure 1:** HammerCloud architecture.

As can be observed in Fig. 1, jobs are submitted to the grid using the native systems that the users of the Vos would use, like PanDA [3] for ATLAS, CRAB [4] for CMS or DIRAC [5] for LHC*b*. All these submission systems are adapted and put into a common interface using the Ganga framework [6].

HammerCloud offers a web interface for user to schedule tests and monitor test status. The test infrastructure, which is a set of scripts based on Ganga that, besides running the tests themselves, provides metric analysis and alarms systems. Both components are based on the Django framework for storage abstraction, which also provides independency of the storage technolgy used, currently MySQL.

The test infrastructure is running in a cluster of machines, with a load balancing scheme that allows running between 10 and 20 tests per machine. Mostly, each test is an isolated process that runs the Python script which simulates the behavior of a user submitting jobs to each site is testing.

### 2.2 HammerCloud results on 2012

During the first quarter of 2012, HammerCloud has been running three instances of the service for ATLAS, CMS and LHC*b*.

In the case of ATLAS, more than 5,000,000 grid jobs have been completed. This VO is using HammerCloud as a central service for operations, in which automatic site exlcusion is performed in base of the results of some functional tests called Analysis Functional Tests (AFT) and Production Functional Tests (PFT), this feature allows the automation of some of the procedures performed by the shifters that monitor the grid status. When a site is missing some software, or the storage element is not accesible, user jobs will start failing in this site. This problem is reported to the shifter, that sets the site in a status (broker off) in which no new jobs will arrive while the site is being put in order. After the intervention, the site is set back in an active state.

HammerCloud is performing this procedures automatically based on the results of those AFT and PFT tests, for analysis and production queues. Also, since the jobs submitted by HammerCloud have the maximum priority in PanDA, they will be the first ones to be executed, being also the first ones to fail in case of problems. In this way, the time to detect the problem is minimum and the site gets excluded from the queues earlier. Therefore, the user jobs that are waiting in the queues will be brokered to another site that is functional and the effect of the problem is reduced. Almost 3,000 exclusions have been performed this year.

Also, HammerCloud is being used in ATLAS to validate the software releases in the grid sites by deploying a *nightlies* test. The software developers will issue a daily version of the software to the grid sites and HammerCloud will send jobs that use this special release, tesing the new iterations of the software in real environments. This allows to reduce the risk of a software release on grid sites, since all the problems are already tested by the time of the milestone.

For CMS, HammerCloud is also a critical operations service, which has run more than 1,400,000 grid jobs this year. Software validation is also being done, having in the horizon the validation and stress test of CRAB 3, the new generation of the CMS analysis suite.

Finally, the LHC*b* instance of HammerCloud is being integrated quickly into operations and the development of DIRAC tests is ramping up.

## 3. Automatic site exclusion and reliability

Automatic site exclusion has proven to be a very useful feature. It reduces the time to detect problems and the time needed to put back a site in normal operations after the interventions. Also, it has been very effective to improve the reliability of the grid.

The exclusion is based on deterministic policies based on the job failures of the the AFT and PFT tests.

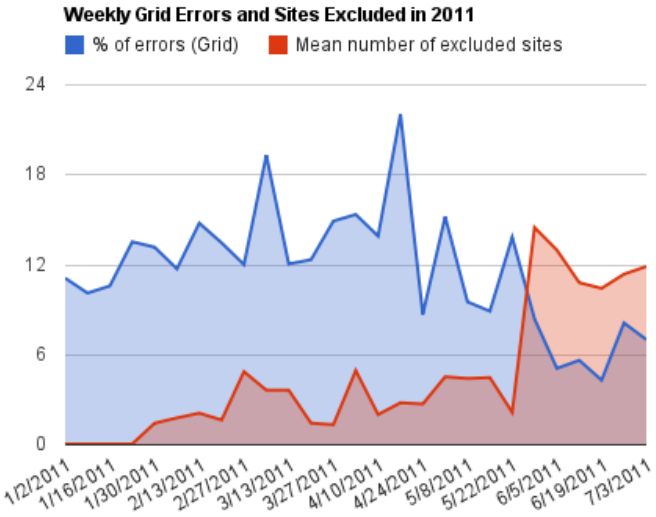


**Figure 2:** Site exclusions and fail rate.

Fig. 2 shows a relation between the site exclusions and the fail rate of the grid, for ATLAS sites, in 2011. The servive started to run in production at the end of January 2011 and a major policy review was performed on May 2011.

As the plot shows, the new policy sets reduced the failure rate on ATLAS sites to almost a half, since the new policies are more strict, so more sites are excluded after not fulfilling the functional requirements.

## 4. Correlation of results

At this moment HammerCloud has a big advantadge since it is a common solution between VOs. The benefit is mainly showed on development and maintenance efforts, and operations from the users. However, taking this avantanges is possible by correlate the data from different Vos to have inter-VO learning.

HammerCloud executes millions jobs in hundreds of grid sites in the World. Daily, it submits more than 60,000 jobs to sites that are different sized, have different backend

technologies and different communication resources. In the case of HEP, there are many sites that are running jobs of different VOs, for instance, ATLAS and CMS.

By applying machine learning algorithms to these streams of data, HammerCloud would be able to predic failures on one VO from the data obtained in other VO, for instance gatekeeper failures, storage elements downtimes or network issues. Of course, there will still be some errors that are specific to each VO.

To do this comparison, all the details of the behavior of the VO have to be taken into account, like the size, the job submission ratio, the *shape* of the tests (for instance, ATLAS deploys tests by kind, such a software version to test, and CMS by region) and the underlying software stack.

## 5. Scalability of the service

The main operations problem faced this year was related to the scalability of the service. In 2012 HammerCloud will submit about 30,000,000 grid jobs that will generate in the order of 3,000,000,000 metrics to store and analyze.

For this huge amounts of data, analysis and correlation is harded (and slower) and data curation is out of manual scale. Also, relational databases get slower to perform queries.

The first solutions came from the new MySQL service from the CERN IT department and a deep optimization of the schema and storage backends. In the future, is possible to apply NoSQL techniques and distributed analysis technologies like the MapReduce implementation of Hadoop.

## 6. Conclusions

HammerCloud has become the tool of choice for functional testing in grid environments. Both for testing and availability monitoring from a user perspective. Also, HammerCloud has proven to be effective to other validations, such a software releases.

Finally, automating operations through HammerCloud allows to reduce the time to detect failures and improves the reliability of grid sites.

## References

[1] I. Foster, *What is the grid? A three point checklist*, *GRID today* **01** (6) 2002.

[2] D.C. van der Ster, J. Elmsheuser, M. Ubeda, M. Paladin, *HammerCloud: A stress testing system for distributed Analysis*, CERN IT 2011.

[3] T. Maeno *et. al.*, *Overview of ATLAS PanDA workload management*, *ATLAS communications on software* 2011.

[4] D. Spiga *et. al.*, *Automation of the user analysis workflow in CMS*, *CMS* 2009.

[5] A. Tsaregorodtsev *et. al.*, *DIRAC, the LHC*b *data production and distributed analysis sytem*, *LHC*b 2006.

[6] J.T. Moscicki *et. al.*, *Ganga: A tool for computational-task management and easy access to grid resources*, *Computer Physics Communications* **180** (11) 2009.