Synergy between the CIMENT tier-2 HPC centre and the HEP community at LPSC in Grenoble (France)

To cite this article: C Biscarat and B Bzeznik 2014 *J. Phys.: Conf. Ser.* **513** 032008

View the article online for updates and enhancements.

Synergy between the CIMENT tier-2 HPC centre and the HEP community at LPSC in Grenoble (France)

C Biscarat[1] and B Bzeznik[2]

[1] Laboratoire de Physique Subatomique et de Cosmologie (IN2P3/CNRS, UJF, Grenoble INP), 53 rue des Martyrs, Grenoble, France
[2] Université Joseph Fourier, CIMENT, Laboratoire d'Informatique de Grenoble, France

E-mail: `Catherine.Biscarat@lpsc.in2p3.fr,bruno.bzeznik@ujf-grenoble.fr`

Abstract.   Two of the most pressing questions in current research in Particle Physics are the characterisation of the newly discovered Higgs-like boson at the LHC and the search for New Phenomena beyond the Standard Model of Particle Physics. Physicists at LPSC in Grenoble are leading the search for one type of New Phenomena in ATLAS. Given the rich multitude of physics studies proceeding in parallel in ATLAS, one limiting factor in the timely analysis of data is the availability of computing resources. Another LPSC team suffers from the same limitation. This team is leading the ultimate precision measurement of the W boson mass with DØ data, which yields an indirect constraint on the Higgs boson mass which can be compared with the direct measurements of the mass of the newly discovered boson at LHC. In this paper, we describe the synergy between CIMENT, a regional multidisciplinary HPC centre, and the HEP community in Grenoble in the context of the analysis of data recorded by the ATLAS experiment at the LHC collider and the DØ experiment at the Tevatron collider. CIMENT is a federation of twelve HPC clusters, of about 90 TFlop/s, one of the most powerful HPC tier-2 centres in France. The sharing of resources between different scientific fields, like the ones discussed in this article, constitutes a great asset because the spikes in need of computing resources are uncorrelated in time between different fields.

1. The scientific landscape in Grenoble

Grenoble, located in south-east of France, is a medium size city of about 160,000 inhabitants also known as the Capital of the French Alps. Grenoble is a major scientific centre, especially in the fields of physics, computer science, and applied mathematics. With a well developed collaboration between industry, the university and research laboratories, Grenoble has been rated in the top five of the most inventive cities in the world [1].

The University of Grenoble was established in 1338 and today it accounts for more than 60,000 students. It is particularly active in scientific domains. Together with other major players of public research in Grenoble - several "Grandes écoles" (higher education schools with competitive entrance exams) and the French National Centre for Scientific Research (CNRS) - they represent 10,000 jobs in more than 200 research laboratories. Many fundamental and applied scientific research laboratories are operated jointly by the Scientific branch of the Grenoble University (University Joseph Fourier, UJF), the Grenoble Institute of Technology (Grenoble INP) (which trains engineers in key technology disciplines) and by CNRS. Many laboratories in other domains

are operated or co-operated by CNRS or by the French National Institute for Research in Computer Science and Control (INRIA). Grenoble also hosts several high-profile scientific centres like the European Synchrotron Radiation Facility (ESRF), the Laue-Langevin Institute (ILL), which provides the highest intensity neutron source in the world, and a research facility of the Nuclear Energy and Alternative Energies Commission (CEA). The MINATEC centre of innovation, created jointly by CEA and Grenoble INP, focuses on micro- and nano-technologies thus reinforcing Grenoble's position as a European Scientific centre.

2. The CIMENT HPC centre and the CIGRI grid

The CIMENT [2] High Performance Computing (HPC) centre of Grenoble University was created in 1998. It is a joint centre and its main partners are UJF, Grenoble INP, and about 30 research laboratories from CEA, CNRS and INRIA. It federates 12 HPC clusters in the Grenoble area and it accounts for 5,700 cores of 89.3 TFlop/s, 19 TB of memory, 760 TB of disk space and about 200 active users. CIMENT is a tier-2 centre in the HPC pyramid, gathering HPC resources of medium scale (intermediate between laboratory resources [tier-3] and national/european [tier-1/tier-0] platforms), cf. figure 1. The main objective of CIMENT is to enable the development of HPC projects in the Grenoble academic research communities by providing researchers and engineers with an easy access to local HPC resources to develop, test and run their codes; to strengthen trans-disciplinary collaborations around HPC; and to train end-users to the methods and tools of HPC for the subsequent use of national/european HPC resources whenever this is required for larger projects. The federated clusters belong to different poles: Biology and Health, Chemistry, Environment and Climate, Numerical Physics, Earth and Planetary Sciences, and Distributed Computing. In 2013, a brand-new HPC cluster shared by the CIMENT partners, accounting for 72 TFlop/s and offering GPU and Xeon Phi technologies, has reinforced the existing computing power.

Given the testing and developing nature of part of the applications run on CIMENT, the clusters are not always fully occupied, resulting in a fair number of empty cycles. Therefore, in 2000, the CIMENT clusters usage has been optimised by integrating all the CIMENT platforms in a local computing grid. The middleware used, CIGRI (cf. figure 2), and the task scheduler, OAR [3], are local products developed by CIMENT, INRIA and the Grenoble Informatics Laboratory (LIG). The use of OAR in an opportunistic mode (tasks are aborted if the "owner" of the cluster needs it) with the automatic re-submission of aborted tasks on another cluster is ideal for running short opportunistic tasks with low parallelism. The CIMENT grid is also equipped with 520 TB of distributed storage based on the IRODS middleware, providing a transparent access to data spread over different physical locations and heterogeneous storage technologies.

3. High Energy Physics in Grenoble

In France, fundamental research in Particle Physics is led by a branch of CEA near Paris and by the National Institute for Nuclear and Particle Physics (IN2P3), part of CNRS. One of the 19 research laboratories of IN2P3, the Laboratory for Subatomic Physics and Cosmology (LPSC), is located in Grenoble. The LPSC is also affiliated with the UJF and the Grenoble INP. Regarding High Energy Physics, the activities of the LPSC are focused on the greatest unsolved mysteries of the universe, like the unification of forces and the origin of the mass of particles.

LPSC physicists are active in the DØ experiment [4] where they lead a high-profile analysis aimed at the indirect determination of the Higgs boson mass via a precise measurement of the W boson mass. This analysis is complementary to the direct observation of a Higgs-like boson at the LHC and it allows for a stringent test of the internal consistency of the Standard Model of Particle Physics (SM) [5]. The measurement of the W boson mass is based on data accumulated by the DØ detector at the Tevatron proton-antiproton collider at Fermilab (US). Despite the challenging environment of the hadronic collisions at the Tevatron, the precise
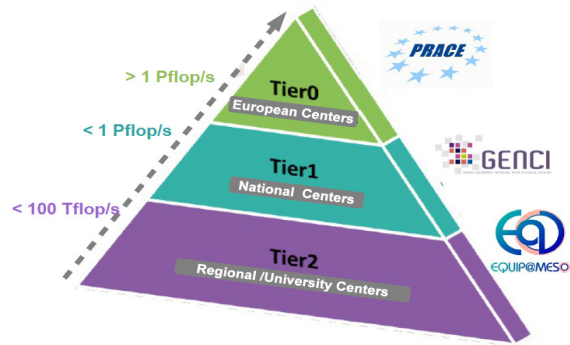
Figure 1. The HPC pyramid in Europe with PRACE, GENCI and EQUIP@MESO providing the governance by layer.
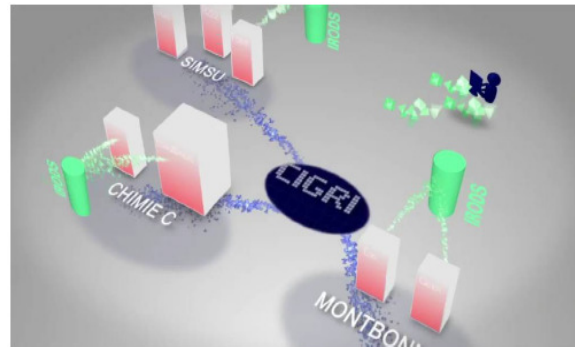

Figure 2. Schematic view of the CIMENT grid, CIGRI, which integrates the HPC clusters of the CIMENT federation.

measurements of the W boson mass by CDF and DØ (the two detectors at the Tevatron) dominate the world average measurement and they will be highly competitive for at least the next decade. The centralised computing tasks (like data reconstruction and event simulation) at the DØ experiment are executed using distributed computing as well as the grid. End-user analyses are executed locally at Fermilab on a dedicated analysis farm. The computing resources available on the local analysis farm have always been a bottleneck in the timely publication of physics results but LPSC physicists have explored new possibilities and moved part of their jobs away from Fermilab to CIMENT.

Another focus of LPSC physicists is the search for physics beyond the SM. In the ATLAS experiment, LPSC physicists lead the search for extra dimensions in di-photon events. This search will potentially give insight on the nature of the fundamental theory that underlies the well known SM of Particle Physics as we know it today. The computing model of the LHC experiments is well established. It is organised in the framework of WLCG with more than 150 participating computing centres integrated in a grid. The LPSC is part of this global effort since 2008. LPSC hosts a WLCG tier-2/3 centre supporting the ALICE and ATLAS experiments with, as of 2013, about 800 cores and 700 TB of storage space (DPM and xROOTd technologies are deployed). The site policy aims at keeping the site busy (cf. figure 3) and consequently there is a large latency to get jobs for end-user analysis running. In particular, operational experience shows that the grid does not provide enough computing resources to absorb peaks in end-user analysis ahead of the most salient physics conferences. This is why we have taken the opportunity offered by CIGRI to run analysis jobs on CIMENT.
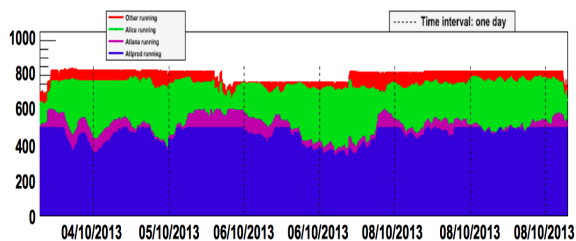

Figure 3. Number of running jobs on the LPSC site during the week prior to CHEP2013. The site is fully occupied with ALICE (green) and ATLAS tasks (production in blue and analysis in magenta). A few percents of slots are dedicated to non-LHC VOs (red), mainly used by the VO "biomed".

4. Collaboration between HPC and HEP

The CIMENT grid hosts several scientific research areas different from Particle Physics. The spikes in the need for computing resources in Particle Physics tend to occur before the major

conferences of this field. They are not correlated in time with the spikes of computing resource usage on CIMENT. Therefore the use of the CIMENT grid for end-user analyses of HEP physicists from LPSC contributes an important synergy between CIMENT and LPSC. The sharing of expertise between the HEP and the HPC communities, regarding the usage of GPUs and supercomputers or the management of large datasets, is also an asset for providing diverse appropriate capabilities to the different scientific communities.

4.1. First HEP use case on CIGRI

The precise measurement of the W boson mass needs an enormous computing power in order to simulate precisely the response of the DØ detector. The PMCS application, a DØ product, is a parametrised simulation used by the W-mass team at DØ. PMCS is highly customised in order to achieve the thousands level of precision needed for this measurement. Since PMCS can be executed as large series of short tasks, it is an ideal application for the opportunistic execution on the CIMENT grid. We have ported PMCS to CIMENT, along with the standard HEP tools (ROOT, LHAPDF ...) that PMCS depends on. The input files needed by PMCS (sets of four-vectors generated by a HEP event generator) are small and can easily be stored in IRODS. The CIGRI middleware needs a JDL file as well as a parameter file describing the input parameters of each job described in a bag of tasks. Extensive tests have proven that the PMCS simulation executed on CIMENT leads to results that are statistically consistent with calculations executed at Fermilab (cf. figure 4). Running hundreds of PMCS jobs lasting about one hour each, we have observed a 100% success rate on CIMENT. Given these results, we have developed the machinery to feed CIMENT with thousands of such jobs.

The scientific result obtained by the DØ Collaboration in 2012 [6] was the most precise single W boson mass measurement at the time. The community has celebrated this important result, which together with the direct observation of a Higgs-like boson at CERN, allows for a stringent test of the internal consistency of the SM. For this occasion, an artistic view of the result was presented on the front cover of the prestigious PRL journal, cf. figure 5. The published result includes the calculations performed on CIMENT.
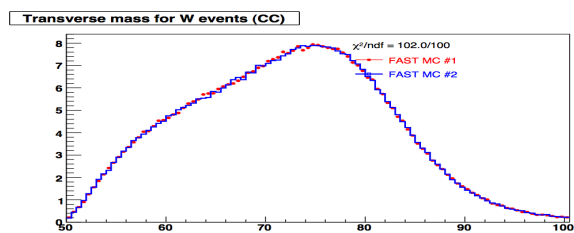

Figure 4. Comparison between the W boson transverse mass computed on CIMENT (blue) and on the Fermilab analysis farm (red) during the PMCS porting phase. Statistical fluctuations come from the use of different random seeds.

Figure 5. Cover of the PRL journal displaying an artistic view of the precise measurement of the W boson mass obtained at the Tevatron and its prediction for the mass of the Higgs boson.

4.2. CIGRI as an analysis farm

The key ingredient in searches for signatures of New Phenomena beyond the SM at the LHC collider is the production rates of background events of the SM processes. In the search for extra dimensions based on ATLAS data [7, 8] lead by LPSC physicists in the ATLAS Collaboration (cf. figure 6), the theoretical calculations implemented in the DIPHOX event generator [9] are

used to predict the background rates. Unfortunately, DIPHOX requires very large amount of computing power to predict the event rates at large di-photon invariant mass which are relevant for this ATLAS analysis. We have ported DIPHOX to CIMENT and we have run thousands of independent jobs of up to 12 hours each. This analysis has been elaborated during a period of six months. During this time, we have used 196 cores/day on average for these calculations, with usage spikes which would have exceeded the total capacity of the LPSC if we had chosen to execute our jobs on this site, cf. figure 7. The results of the DIPHOX calculations (histograms) were stored in IRODS and analysed later at LPSC.

For this analysis, we also used CIMENT to calculate the final exclusion limits. These computations do not require an enormous amount of computing power, a hundred a short tasks at a time is enough though this exercise has to be repeated many times and a fast turn-over is needed for the end-user. To this end, CIMENT was extremely satisfactory and the result of the calculations always became available in a timely manner.
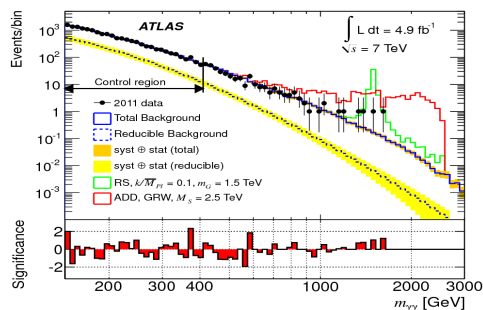

Figure 6. Di-photon events invariant mass spectrum for data (full circle) and simulated events (SM processes are shown in blue with two extra dimensions processes on top).
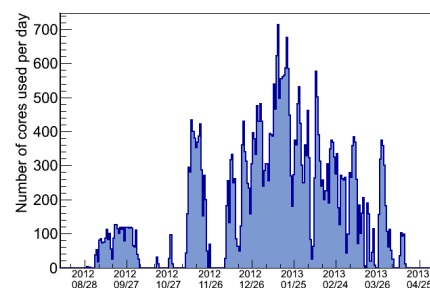

Figure 7. Number of sequential tasks per day (with one core each) run on CIMENT for the prediction of background rates in the ATLAS search for extra dimensions.

5. Conclusion

In this article we have described the synergy between the HEP and HPC communities in Grenoble. This synergy is twofold: the sharing of expertise regarding HPC technologies and the management of large datasets, and the optimisation of the usage of local resources shared between different scientific fields. The existence of the CIMENT computing grid CIGRI for an opportunistic usage of the CIMENT HPC clusters is an asset to exploit resources for the HEP end-users and it is a powerful tool for a timely elaboration of high-profile scientific results such as searches for New Phenomena in ATLAS and, in DØ, the characterisation of the newly discovered Higgs-like boson. Based on this success, more end-users at LPSC are eager to make use of CIGRI to address their computing needs.

References

[1] *Forbes magazine* July 2013 http://www.forbes.com/
[2] http://ciment.ujf-grenoble.fr/
[3] http://oar.imag.fr
[4] http://www-d0.fnal.gov/
[5] Kotwal V and Stark J 2008 *Annu. Rev. Nucl. Part. Sci.* **58** 147–75
[6] Abazov V M *et al* 2012 *Phys. Rev. Lett.* **108** 151804
[7] Aad G *et al* 2013 *New J. Phys.* **15** 043007
[8] Buat Q 2013 PhD dissertation https://cds.cern.ch/record/1620821
[9] Binoth T *et al* 2000 *Eur. Phys. J.* **C16** 311–30