

UTFSM/CCTVal Data Center (10 Years of Experience)

Yu. P. Ivanov^{1,2,a}, L. Salinas^{1,b}

¹Universidad Técnica Federico Santa María,
Avenida España 1680, Casilla 110-V, Valparaíso, Chile

²Joint Institute for Nuclear Research,
Joliot-Curie 6, Dubna, Moscow region, 141980, Russia

E-mail: ^a yuri.ivanov@usm.cl, ^b lsalinas@inf.utfsm.cl

During the last 10 years the data center of Universidad Técnica Federico Santa María (UTFSM, Valparaíso, Chile) has been providing its computational resources for users from UTFSM and other Chilean universities. Started as a Chilean part of the international project EELA on creation of a computational infrastructure distributed between Europe and Latin America, the cluster was significantly extended after creation of the Scientific and Technological Center of Valparaíso CCTVal in 2009 and became the UTFSM/CCTVal Data Center. Local users have direct access to the cluster's computational resources via the batch system. The cluster facilities are also available in the frame of Grid Computing. Grid users have access to the part of the cluster resources in the frame of the EGI (European Grid Infrastructure) and WLCG (Worldwide LHC Computing Grid) infrastructures. These resources are provided to users from certain "Virtual Organizations" (VO). The main supported VO is "ATLAS", which joins researchers of the ATLAS experiment at the Large Hadron Collider (LHC, CERN). The data center was the first computer center in Latin America working for the ATLAS Collaboration. Another supported VO is "EELA Production" which includes users from Latin America, Italy, Portugal and Spain. In addition to ordinary processors, the cluster also allows powerful Graphics Processing Units to be used for calculations. The data center participates in a big Chilean project NLHPC (National Laboratory for High Performance Computing). Computational facilities provided by this project are integrated into the data center infrastructure. The range of problems being solved at the cluster is huge, from fundamental and applied problems in different branches of Physics, Chemistry, Astronomy, Computer Science to the educational and training purposes. The paper presents the history and the current status of the cluster including configuration and some statistics of usage.

Keywords: distributed computing, high performance computing, grid computing

© 2016 Yuri P. Ivanov, Luis Salinas

1. Introduction

In this paper we present the history and the current status of the UTFSM/CCTVAL data center, beginning with just a few words about Universidad Técnica Federico Santa María [UTFSM] and Centro Científico Tecnológico de Valparaíso [CCTVal].

UTFSM is a private, not-for-profit university and belongs to the 25 traditional universities in the Chilean University Council of Rectors (CRUCH). The beginnings of the University go back to the altruist dream of Federico Santa María, who set the foundation of the Institution in his will and testament in 1920. He stated his wish to contribute to the progress and increase the cultural horizon of Chile to his executors. And in the mid-1930s his dream of having a world class engineering University came true. UTFSM is one of the most academic selective universities in the country. The University specializes in the following areas: Electronics, Mining, Mechanical Engineering, Metallurgy, Electrical Engineering, Industrial Engineering, Informatics, Business, Basic Sciences (Physics, Chemistry, and Mathematics), Aeronautics, Architecture, Construction, and Environmental Sciences. It has four campuses in Chile and one in Ecuador. The University has PhD and Master's Degree Programs in around 20 areas including Physics, Chemistry, Informatics, Electronics Engineering, and Biotechnology. UTFSM also hosts a number of research centers, including CCTVal.

The idea to merge experience in the development of the Particle Physics, Computing, and Electronics research areas led to the foundation of CCTVal in 2009. The Center was created and acknowledged by the National Commission of Scientific and Technological Investigation (CONICYT). CCTVal has several research groups: Theoretical Elementary Particle Physics, Experimental High Energy Physics, Informatics and Computing, Power Electronics and Systems and Signals. One of the main objectives of CCTVal is to comply with the commitment of global collaboration, strengthening the links between Chile and world renowned laboratories such as CERN, Jefferson Lab, and Fermilab.

Since the first days of its creation, the UTFSM/CCTVal data center has been actively used by researchers and students from UTFSM and other Chilean universities. The cluster facilities are also available in the frame of the Grid Computing. The Data Center is a Tier-2 level site in the frame of the European Grid Initiative [EGI] and Worldwide LHC Computing Grid [WLCG] infrastructures.

The short history of the cluster development is presented in the second Section. The third Section contains information on the current cluster layout for the Grid and High Performance Computing (HPC). Conclusions are presented in the last Section.

2. Short history of the Data Center

The Data Center started as a small Chilean part of the international project "E-infrastructure shared between Europe and Latin America" [EELA]. UTFSM joined this project due to the initiative of the Informatics and Physics Departments that were badly in need of modern computational facilities. At the end of 2006 UTFSM received a dozen of servers in a quite reasonable configuration for that time (dual CPU 1.6 GHz, 4 GB RAM and 140 GB HDD per server). One machine had even three SAS hard drives with a total capacity of around 1 TB raw disk space. This equipment allowed launching the first computational cluster in UTFSM.

The first operating system (OS) used at the cluster in 2006 was Scientific Linux 3 (SL3). In a year it was replaced by Scientific Linux 4 (SL4). This system was used approximately up to 2010 despite the fact that Scientific Linux 5 (SL5) was already available since 2007. Such delays with OS upgrades are related with the requirement to keep the OS compatible with the systems used in big scientific research centers like CERN and Fermilab.

During the first few years the cluster worked mostly for local users. All Grid-related activities during that period were on a level of simple tests. One should mention here that the Grid in Latin America is rather specific: there are just a few computational centers working in this area, most

of which are dedicated only to one experiment or project. So, instead of coordination of the Grid activities on a national level via the National Grid Infrastructures (NGI), it was necessary to create the "Latin American Regional Operation Centre" (ROC-LA). It was done only at the end of 2008 in the frame of the second stage of the EELA project (EELA-2). And in March 2009 the UTFSM cluster successfully passed all Grid tests and was certified as a Tier-3 Grid Resource Center with the site name "EELA-UTFSM". One of the first supported VOs was "EELA Production" which included users from Latin America, Italy, Portugal, Spain and some other countries originally through the project EELA and since 2010 through the project "Grid Initiatives for e-Science virtual communities in Europe and Latin America" [GISELA].

In 2009 the cluster still had around 40 CPU computational cores, while the number of even local users had already grown up to more than a hundred. The situation became worse when Grid job processing started. It was clear that the existing computational facilities were not enough. Fortunately, that year the CCTVal center was founded. Funds provided by the center allowed the computational facilities and storage capacities to be significantly increased during the next years. In 2010 the cluster got additional computational servers with 256 CPU 2.8 GHz cores and a storage server with the 20 TB disk space, and more servers (around 200 additional CPU 3.1 GHz cores) with storage servers of total capacity around 180 TB in 2012. Computers with only 40 CPU cores can fit in a small rack, but several hundreds require not only a few racks, but also a proper environment, i.e. air conditioning systems, powerful electric power supply, etc. All this was done with CCTVal funding, and, as a result, the UTFSM computer cluster became the UTFSM/CCTVal Data Center.

Noticeable increase in computational power brought another problem to the fore — bandwidth of the external network connection. In most cases, processing of Grid jobs also requires data transfer from the external Grid storage. When in 2009 the cluster started processing of grid jobs, it had only 10 Mbps external bandwidth shared with the Informatics Department. And that situation was typical of universities and research centers in most Latin American countries. Fortunately, after 2010 the situation with network connectivity started to improve. New optical lines (including submarine cables) to Europe and North America appeared. Finally, in 2011 UTFSM got a direct network link from the National University Network provider REUNA [REUNA]. The UTFSM/CCTVal cluster obtained the total external bandwidth of 54 Mbps (40 Mbps dedicated and 14 Mbps shared). Such bandwidth extension allowed the cluster to start processing of the full range of Grid jobs. With a low bandwidth, the cluster was suitable for processing jobs without big data transfers like ATLAS Monte-Carlo simulations. With the ability to process hundreds of jobs, the cluster started to meet requirements of local users, and it also became possible to include the center into ATLAS data processing. In 2012 the cluster passed the ATLAS certification and was included in ATLAS Production and Analysis as a Tier-3 Grid Ready site. The UTFSM/CCTVal data center was the first computer center in Latin America working for the ATLAS collaboration.

The number of CPU cores provided by the cluster for the Grid computing at that time allowed it to be considered close to the Tier-2 level, but the technical characteristics of the site were not the only issue. The most important difference between Tier-3 (i.e. entry Grid level) and Tier-2 is that resources provided by a Tier-2 site are "pledged". These "pledges" are stipulated in a special agreement between the resource provider and the WLCG. In 2013, the coordinated efforts of the Latin American research centers led to the creation of the "Tier-2 Latin-American Federation". The Memorandum of Understanding between Centro Latinoamericano de Física [CLAF] and the WLCG was signed in September 2013. CLAF is an international organization aimed to promote and coordinate efforts in development of physics in Latin America (Argentina, Brazil, Colombia, Chile and others). This organization works as an "umbrella" to all ROC-LA sites. The financing is coming from the hosting institutions, not from central sources. As a result of all those efforts, in November 2013 the

UTFSM/CCTVal cluster was officially presented as an ATLAS Tier-2 site at the ATLAS International Computer Board (ICB).

In addition to computations with ordinary CPUs, the cluster hardware allows data processing with the powerful GPU cards. The first servers with nVidia Tesla C1060 cards became available for cluster users in 2010. Next year a couple of servers with nVidia Tesla M2050 were added. And in 2014 the cluster received five more servers with powerful nVidia Kepler K20m GPU cards in the frame of the Chilean project "Tsunami".

In 2014 the cluster was seriously extended in the frame of the big Chilean project called "National Laboratory for High Performance Computing" [NLHPC]. The main computational facilities of this project are located in the Center for Mathematical Modeling [CMM] of Universidad de Chile, Santiago. But other universities participating in this project also obtained some computational equipment. As a result, the UTFSM/CCTVal data center increased its computing power by around 240 CPU cores of 2.9 GHz and disk storage by around 60 TB.

Cluster computational facilities grew with time. Figure 1 shows the amount of the CPU time used by local users with direct access for High Performance Computing (HPC) and by remote users via EGI/WLCG Grid infrastructures since 2007.

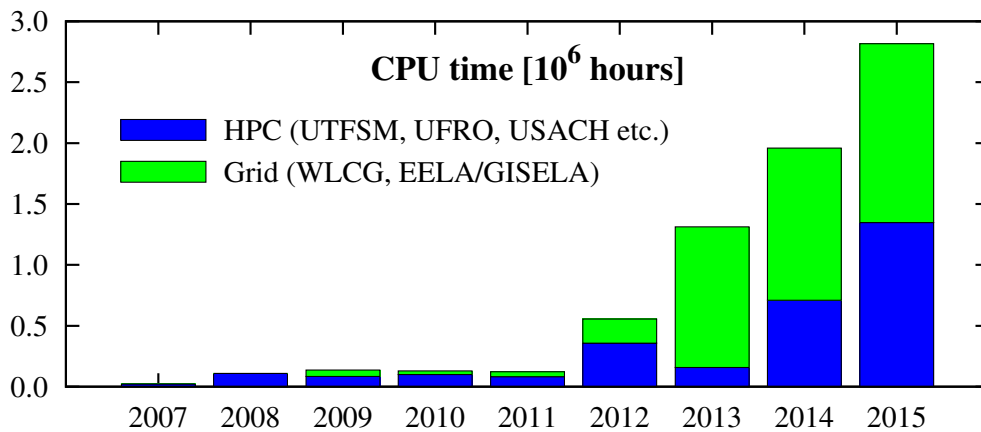


Fig. 1. CPU time used yearly for HPC and Grid computing

In the last two years the whole cluster infrastructure has been seriously updated: noticeable extension of all computational facilities required serious upgrading of the whole air conditioning system and reconstruction of the electric power supply. The internal network infrastructure has also been changed: now all storage servers and most of the computational servers have 10 Gigabit Ethernet connection.

3. Cluster Layout

The present-day cluster configuration includes around 800 CPU cores with a total amount of disk space on storage servers around 300 TB. Figure 2 presents the current cluster layout. Local users have direct access to the cluster computational resources via User Interface (UI) servers. For authentication and authorization of users LDAP and Kerberos services are used. Using development tools and packages installed on the UI servers, users can submit jobs to the Portable Batch System (PBS) server for the execution via the batch system. The PBS server (PBS version "Torque" with scheduler "Maui" is used) distributes jobs among the Worker Nodes (WN). All computational nodes have processors with frequencies from 2.8 GHz to 3.1 GHz and at least 3 GB RAM per each core.

In addition to the ordinary computational facilities provided by CPUs, the cluster allows special jobs processing with powerful GPU cards on the GPU Worker Nodes (nodes `gp01..gp09` in Figure 2). Disk storage for local users is provided by the File Servers (FS) with the total capacity of around 100 TB via the NFS and GlusterFS [GlusterFS] distributed file systems. Local and grid users also have direct access to the programs and libraries provided by CERN via the CernVM File System [CernVM-FS].

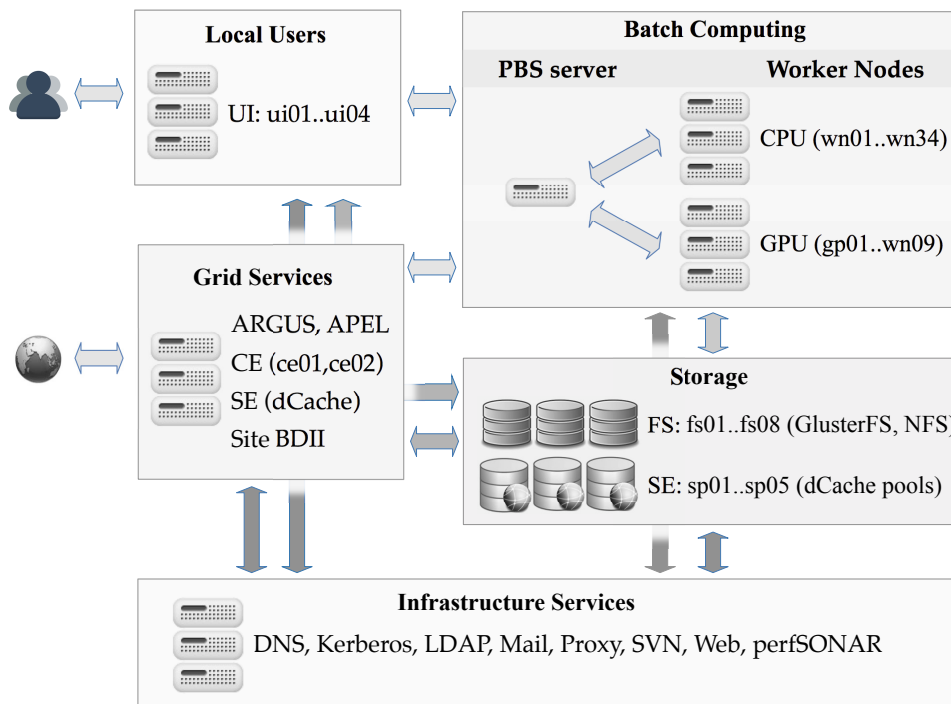


Fig. 2. UTFSM/CCTVal cluster layout. Local users have direct access via the UI servers to the batch system. Access for remote users is provided via the Grid services. See text for more detailed explanations.

Remote users from the supported VOs have access to the cluster resources via the EGI/WLCG Grid infrastructure. Integration into the cluster computational structure is provided by the Grid services: the Computing Element (CE) servers deliver Grid jobs to the local batch system and the Storage Element (SE) server (dCache system) gives access via the Grid access protocols to the SE dCache disk pools with a total size of around 200 TB. The "Site BDII" server informs the Grid infrastructure on the current status of the computational and storage cluster resources. For authorization of Grid users the ARGUS server [ARGUS] is used. The APEL client [APEL] collects accounting data on the cluster work and transfers this information to the EGI/WLCG servers.

The data center also has infrastructure level services like the Domain Name Server (DNS) and the Mail, Proxy, and Web servers. For monitoring the quality of the external network connectivity two "perfSONAR" nodes [perfSONAR] (for monitoring bandwidth and latency) are used. All computational and storage servers use "bare" hardware, while for some services (like APEL, ARGUS, CE *etc.*) different virtualization techniques (Kernel Virtual Machine [KVM] and Container Virtualization provided by OpenVZ [OpenVZ]) are used.

Currently most of the cluster servers have Scientific Linux 6 (SL6) installed in the 64-bit mode. Users have access to compilers (C, C++, FORTRAN) and other programming languages (Perl, Python).

The cluster has a lot of specialized program packages and libraries installed (GEANT, OpenFOAM, Pythia, ROOT *etc.*). For the parallel programming the Message Passing Interface (MPI) with the MPICH or OpenMPI packages can be used. For the GPU programming the nVidia CUDA compiler is available.

4. Conclusions

The UTFSM/CCTVal Data Center is used by researchers from UTFSM, Universidad de Chile, Universidad La Frontera (UFRO), Universidad Santiago de Chile (USACH) and other Chilean universities. Being also a Grid Resource Center of the Tier-2 level, the cluster is under permanent monitoring by different control systems, starting from local cluster monitoring and up to external monitoring by the Grid infrastructures, including systems provided by each VO. This control imposes serious restrictions on the possible time of cluster unavailability for users. The Tier-2 level requires high values for such cluster parameters as "Accessibility" and "Reliability": their values should be not lower than 95%. High reliability of the data center provides effective cluster usage in a wide range of different research projects:

- Computations in high energy and particle physics, including ATLAS analysis
- Biomedical image processing (like Digital Pathology on breast cancer *etc.*)
- Satellite image processing for environment protection
- Modeling of mechanical constructions (turbulent flow around bridge piers *etc.*)
- Project "Tsunami"(modeling of tsunami hydrodynamics and implementation of an operational database integrated with the National Tsunami System)

Creation and maintenance of this center require serious efforts from the cluster support team. Note that here the term "maintenance" includes not only regular replacement/repair of hard drives, UPS batteries, hardware and software updates, but also a proper user support via web and online tutorials *etc.* During all the years passed the cluster has been working as a reliable computational tool. The cluster support team works hard to meet all needs of the data center users. And all its members hope that it will be possible to continue this work in the future to satisfy in a reliable way all new arising requirements of the ongoing and upcoming projects.

References

- APEL accounting tool [Electronic resource]: <https://wiki.egi.eu/wiki/APEL>
Argus authorization service [Electronic resource]: https://www.gridpp.ac.uk/wiki/Argus_Server
CernVM File System [Electronic resource]: <https://cernvm.cern.ch/portal/filesystem>
CCTVal Research Center <http://cctval.cl>
Centro Latino-Americano de Física [Electronic resource]: <http://www.claffisica.org.br>
Center for Mathematical Modeling [Electronic resource]: <http://www.cmm.uchile.cl>
Project EELA [Electronic resource]: <http://www.eu-eela.eu>
European Grid Infrastructure [Electronic resource]: <https://www.egi.eu>
The GISELA Project [Electronic resource]: <http://www.gisela-grid.eu>
GlusterFS storage file system [Electronic resource]: <https://www.gluster.org>
Kernel-based Virtual Machine [Electronic resource]: http://www.linux-kvm.org/page/Main_Page
National Laboratory for High Performance Computing [Electronic resource]: <http://www.nlhpc.cl>
OpenVZ container-based virtualization [Electronic resource]: <https://openvz.org>
"perfSONAR" test and measurement infrastructure [Electronic resource]: <http://www.perfsonar.net>
Red Universitaria Nacional (REUNA) [Electronic resource]: <http://www.reuna.cl>
UTFSM University [Electronic resource]: <http://www.usm.cl>
Worldwide LHC Computing Grid [Electronic resource]: <http://wlcg.web.cern.ch>