LIP

LABORATÓRIO DE INSTRUMENTAÇÃO E
FISICA EXPERIMENTAL DE PARTÍCULAS

SUPPORT FOR THE ATLAS AND CMS PORTUGUESE TIER-2 IN THE CONTEXT OF THE WLCG MOU

RECI/FIS-NUC/0115/2012

FINAL REPORT

January 2016

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

COMPETE
PROGRAMA OPERACIONAL FACTORES DE COMPETITIVIDADE

QREN
QUADRO
DE REFERÊNCIA
ESTRATÉGICO
NACIONAL

UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

TABLE OF CONTENTS

1 Objectives

The grand challenge of this project was to ensure that adequate computing services integrated in the WLCG infrastructure would continue available to the Portuguese physicists enabling them to participate in the ATLAS and CMS data analysis at the CERN/LHC collider.

The main expected results as defined in the proposal were:

- Allow the continuation of the Portuguese participation in the LHC physics research program.
- Partial renewal and maintenance of the WLCG Tier-2 components hosted at LIP.
- Improvements of the LHC analysis facilities at LIP.
- Fulfillment of the Portuguese obligations in the context of the WLCG MoU.
- Continuation of a unique state-of-the-art e-science infrastructure in the country.
- Support other research activities in diversified scientific domains that profit from the Tier-2.
- Keeping crucial know-how in a domain that is strategic and where Portugal has always played a leading role.
- Enable the participation of Portuguese researchers in large scientific collaborations that require access to distributed resources.
- A reference installation whose technological challenges, ideas and developments can be relevant and applicable in other scenarios.

2 Task reports

The project included four tasks whose activities are described in this section. Some tasks were initially foreseen to cover only some months of the project, however it was quickly understood that several of the actions required considerable effort and time to implement in a production system. Furthermore, the Tier-2/3 must be part of a continuous improvement process that considers new ideas and options regarding the technology and follows the evolution of the LHC computing model. Therefore, the task activities were extended to cover the project duration without additional funded person months.

2.1 Efficiency and technologies

This task was mainly oriented to pursue the efficiency aspects of the Tier-2 aiming at decreasing the total cost of ownership and improving the service quality. The task aimed to address the following topics from the perspective of efficiency: energy, IT services, and alternative technologies.

After careful evaluation of possible actions and benefits, it was concluded that one of the most effective improvements addressing both the costs and the service quality was the consolidation of the three Tier-2 sites at the NCG datacenter in Lisbon. A major factor that contributed to this decision was the increase of electricity costs in Portugal due to a VAT change from 6% to 23% together with the increase of commercialization costs. The NCG datacenter in Lisbon (also known as sala-grid) was built by FCCN (Portuguese NREN), LIP and LNEC (national Civil Engineering laboratory) in 2009/2010 and already hosted a significant fraction of the Tier-2 capacity. The NCG datacenter had the best Coefficient of Performance (COP) among the three datacenters supporting the Tier-2. Additionally, since it is a large facility (with 1 MW of power) it has the lowest energy price. LIP worked with FCCN to improve the datacenter and the Tier-2 housing with additional space, better separation between cold and hot aisles and better resiliency. Additional racks and power lines for the Tier-2 and Tier-3 were also deployed within the project. Currently (January 2016) a study to further enlarge the datacenter is finished and the new design will further improve energy efficiency. The Tier-2 operations effort also became more sustainable, with FCCN managing the datacenter while LIP focuses on managing the Tier-2 computing equipment and related services. Furthermore, the consolidation into a single facility enables economy of scale and avoids the duplication of data and services thus contributing to an easier operation and optimization of the available resources. The consolidation of the Tier-2 into a single location facilitated the planning, deployment and operation of the new storage system purchased within the project. In terms of services reliability and availability the NCG datacenter also had clearly the best record. Finally due to the economic crises and difficulties to renew contracts some computing team PhDs left the country, making the management of the Tier-2 across three datacenters more complex. The consolidation was performed in several steps during the three years of the project. In order to reduce costs during this period the datacenters were carefully monitored and parameters such as temperature and humidity set-points were continuously tuned. The LIP-Lisbon centre was still kept both as a backup solution and also as the housing place for the Tier-2 tape library supporting the data backups.

In order to decrease the energy costs of the farm, the Sparks farm power management system developed by the project team was significantly rewritten and enhanced, The system dynamically manages the power-up and power-down of Tier-2 worker nodes (computing nodes) according to the load and in accordance with several power management policies. The system aims at minimizing the power consumption. The system is modular, it supports equipments from multiple vendors and several power management interfaces. It is being used with the Son of Grid Engine (SGE) batch scheduler used at the Tier-2/3, but it can be extended to work with other scheduling systems. The results obtained from monitoring have shown a decrease of 15% in the power consumption. Figure 1 shows the activity of Sparks for a given real load across a time interval. It can be seen that the number of powered-down computing hosts evolves with the actual number of computing tasks.

The Sparks system was continuously enhanced along the project. It can manage the load across different time intervals with different prices. It can take into account machines with different power efficiencies, and address specific usage policies. Currently there are already plans to further extend the system to include the management of virtualized farm computing nodes to be delivered via an OpenStack IaaS cloud (see next sections).

An evaluation of the power consumption of all different element types in the datacenters was performed, and each new system type acquired is now evaluated in terms of power consumption. Together with benchmarking information that was already being collected for each system type, it can be used to characterize the power efficiency of each system. This information is used to decide about retiring systems due to low efficiency and also to tune the server power management system (Sparks).

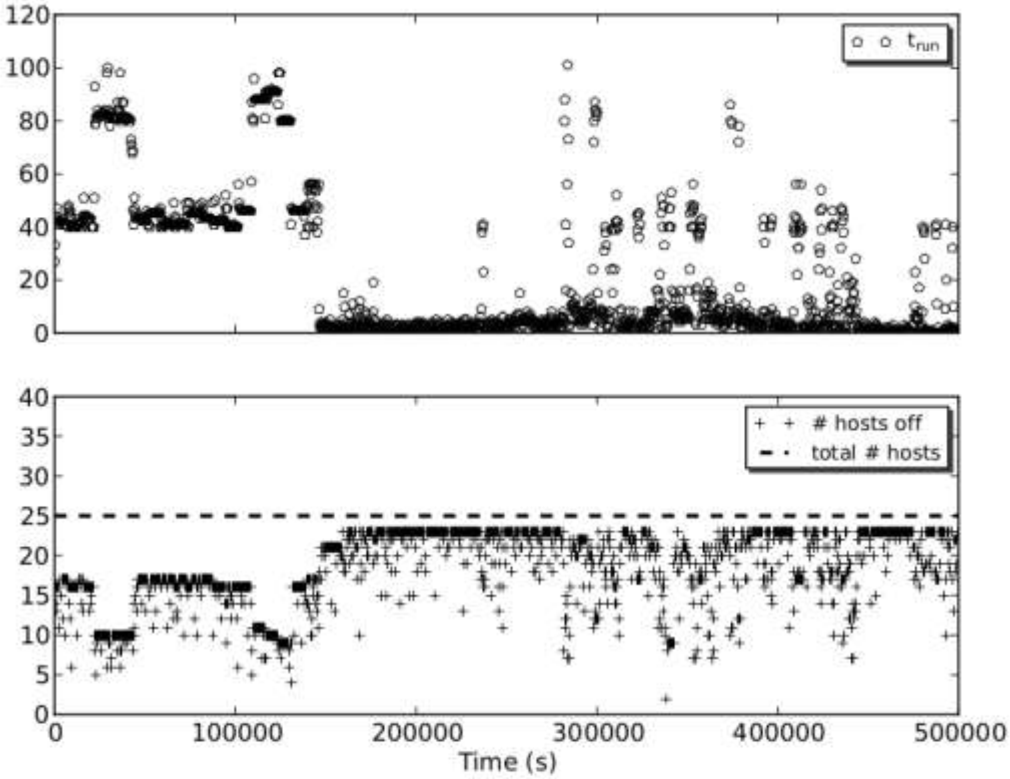

Figure 1 Sparks at work. Number of total executing tasks (t_{run} , ∇ in top figure) and power-off nodes (+ in bottom figure) **in LIP's LRMS system, as a function of time. The dashed line represents the total number of hosts in the system.**

In order to improve the data backup operations and reduce its associated costs, since 2014 all backups have been centralized at LIP-Lisbon. This step also enabled the execution and safe keeping of offsite backups for the data stored at NCG and LIP-Coimbra Tier2/3 sites. The capacity of the Tier-2/3 tape library hosted at LIP-Lisbon was very limited and within task 2 "Infrastructure improvement" its capacity was significantly increased during 2015.

In the Tier-2/3, considerable time is spent in detecting and solving problems. Therefore focus was put in the problem solving cycle: problem detection, incident management, impact assessment, and problem resolution. The Tier-2/ 3 information and operation procedures were streamlined. New computing wikis have been introduced for users and IT services with a much more coherent and exhaustive coverage of all aspects of the Tier-2 oriented towards problem solving. They are currently the focal point to centralize information regarding services and also the configuration. The automated configuration of systems and services was also improved by increasingly use of automated scripts and procedures. During 2015 the use of Ansible (http://www.ansible.com/) for automated configuration purposes was evaluated and adopted. The trouble ticketing system based on RT was improved and is now used for all Tier-2/3 incidents. In terms of the best practices the work was initially focused on following ITIL guidelines however ITIL is quite detailed and the effort required to implement it was quite high. Therefore a more lightweight approach is now being followed by pursuing some of the relevant ITIL recommendations and pursuing the FITSM (http://fitsm.itemo.org/) standard for service management where relevant. FITSM is also being used by the European Grid Infrastructure which is the organization coordinating the global operations for grid sites in Europe and in other countries (excluding US). It is expected that the understanding and compliance with FITSM will facilitate the provisioning and operation of the distributed computing services integrated in EGI (which includes the Tier-2/3).

A software tool to monitor and analyze the file access operations in distributed file systems was developed by the project team. The tool was conceived to enable a better understanding of how file-systems are used. Most data access monitoring tools work at the level of the block devices and do not allow the identification of individual file access operations, nor their related operating system processes and users. The Portuguese Tier-2/3 file-systems are based on Lustre deployments at each site. At each site the Lustre file system is distributed across multiple servers. Obtaining a global view of the file access activity is difficult. The tool makes use of kernel probes via SystemTap to monitor system calls and selected kernel functions in order to obtain detailed file system access information.

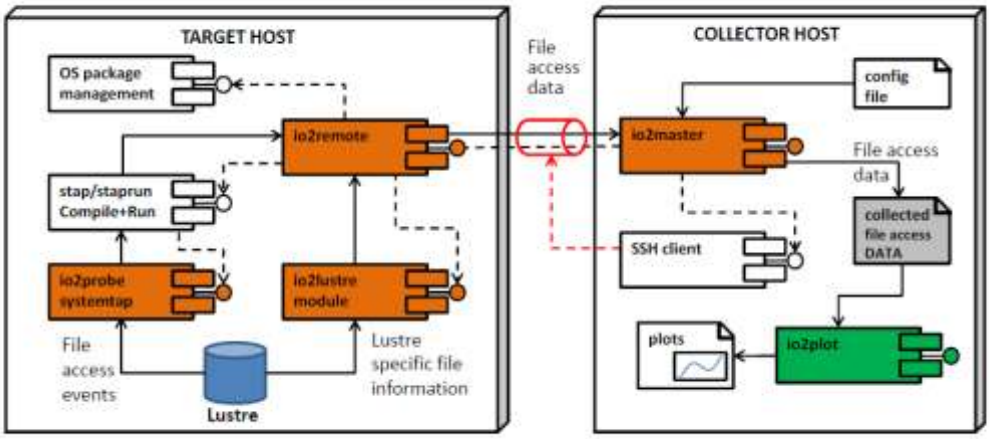

Figure 2 File system operations monitoring tool architecture

The tool is generic and can be used to monitor any Linux file system. A specific module was developed for Lustre enabling to capture and map information related to volume and server identifiers for each file access. The tool was conceived to remotely monitor hundreds of file-system clients. It was used to capture live information from the three Lustre installations at LIP-Lisbon, LIP-Coimbra and NCG. Several types of plots can be produced for analysis. Figure 3 illustrates typical charts produced by the tool. The top figure shows the number of file open operations split by read and write mode across Lustre storage servers at different sites. The bottom figure shows the number of file open operations per team over a given time period. The use of SystemTap kernel probes has an impact on performance therefore the tool cannot be used for continuous monitoring but instead it can be used during a specific time frame (few days) to capture information for analysis.

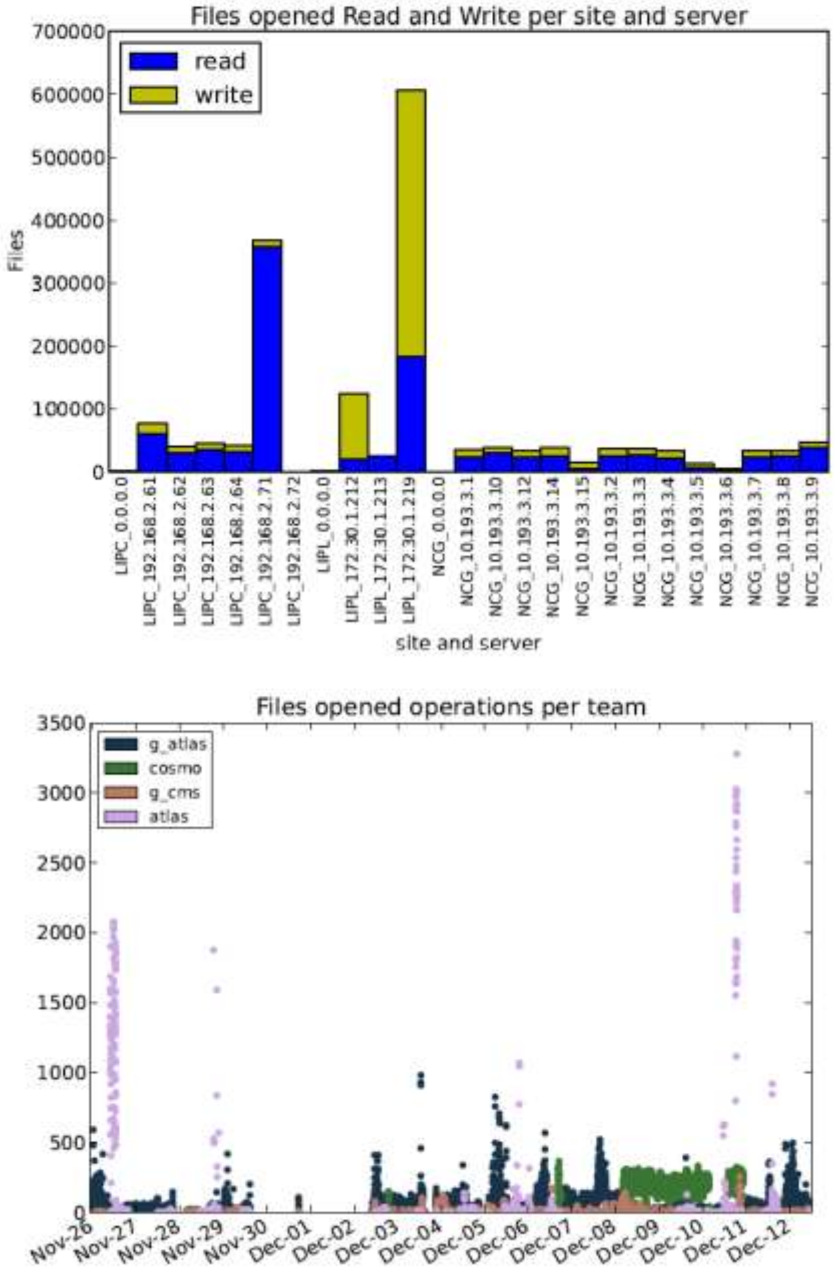

Figure 3 Example of file open operations per server and per team

The evolution of storage software was followed revealing trends associated with the raise of cloud computing and object storage systems. These trends were carefully followed especially taking into account the evolution of the LHC computing model where cloud computing is gaining considerable interest. Among the file systems with high potential Ceph and GlusterFS were identified as very promising since they have support for object storage, block storage and POSIX file storage. The Ceph system seems to be the most scalable also very well adapted to the needs of a future storage service. However the file-system support for Ceph was and still is announced as not ready for production, therefore its evaluation was postponed. On the other hand the Lustre file system also continued to exhibit considerable potential, new features and strong interest namely by the HPC community.

Although object storage can be a future alternative to the conventional storage in the Tier-2, its deployment and integration in the current grid would be complex. Furthermore the computing model of the LHC experiments is evolving therefore it was too early to perform a major change in the Tier-2 storage architecture. The Portuguese Tier-2 needs to be based on solutions that are reliable and can be operated without significant effort. Furthermore the LIP computing facilities and operations team has to support a wide range of users and communities including astro-particles, medical physics, space applications, engineering, and also other users many of them external to LIP. Therefore each solution/system needs to address a broad range of requirements. The POSIX file-systems are so far the solution that better addresses this wide range of user requirements. Breaking with this solution would also introduce problems to the local physics groups and their applications. Therefore it was decided to proceed in parallel with two approaches:

a) Continue using Lustre as the production storage system for the Tier-2/3 moving forward to a more recent version of Lustre.
b) Further evaluate Ceph and other storage alternatives.

The evaluation of Ceph took place in the context of a master thesis in the department of informatics of Universidade Nova de Lisboa (FCT/UNL). The thesis entitled “Enabling and Sharing Storage Space Under a Federated Cloud Environment” was aligned with the future evolution of the Tier-2 towards a cloud computing model. The work started in 2014, and the thesis was successfully defended in December 2015. The thesis was jointly coordinated by FCT/UNL and LIP. The work evidenced Ceph as the best candidate solution. Benchmarks have been performed on Ceph lower level capabilities namely the block storage. The benchmarks aimed to understand and evaluate Ceph’s basic capabilities and behavior. The work developed within the thesis helped to identify how to deploy Ceph within the constraints of the hardware available to the Tier-2, and how to plan its deployment in production. The results were positive although the thesis couldn’t go as far as desired due to lack of time. Some clear conclusions are that: Ceph scales well, is robust and has very good fault-tolerance built-in, however although Ceph can use hardware very similar to the Tier-2, the requirements in terms of memory and networking are higher, also its fault tolerance depends on data replication across servers potentially requiring a higher volume of raw disk space. A potential solution to minimize the replication can be the use of “Ceph erase polls” a mechanism similar to the traditional RAID but that can work across Ceph disk servers. Nonetheless, this mechanism incurs a strong performance penalty and needs to be further evaluated. Overall Ceph can be deployed and managed effectively with the effort available at the LIP computing team.

The evolution of storage hardware was also followed. The technology evolution has enabled a storage density increase. When the Tier-2 was initially deployed in 2009/2010 the maximum disk capacity available from major vendors was 1TB, along the evaluation performed within this task it became clear that the purchase of disks with 4TB could be economically feasible. The overall power consumption benefit would be at least a three times decrease in power consumption. The analysis of usage also showed that the Tier-2 could cope with such a reduction of spindles pointing to the purchase of disks with higher capacity. Together with the other energy efficiency improvements a very significant reduction of power consumption and associate costs could be achieved. The focus was put on this approach. Other much more complex options such as using hierarchical storage systems (HSM) or automated disk spin-down were set to low priority. The HSM option was still considered however the following technical disadvantages and potential issues were identified:

- High rotation of data in the Tier-2. Data is being constantly produced, transferred and deleted in short cycles and thus requires constant file availability thus defeating the benefit of offline storage.
- The HSM would require a considerable number of tape drives and would still require a large pool of disks making this option expensive to acquire.
- It would imply file access delays with impact on batch jobs efficiency (higher elapsed time, less CPU time delivered and more energy consumption in the farm for the same work).
- The identification of files to be moved to tape also raises issues as the last access time of each file is required to understand which files have not been accessed within a given timeframe. At LIP this functionality is disabled for performance reasons as it implies a larger increase of metadata write operations. Metadata access is a known bottleneck in Lustre.

The HSM option was not fully discarded. The choice to continue with Lustre and moving towards version 2.5 was also performed considering this option. Lustre 2.5 supports HSM although only offers disk pool data management, the actual data storage in tape and its management requires external software such as HPSS or TSM (which are commercial products).

Exploiting disk spin-down also lost much of the initial motivation due to the possibility of simply reducing the number o spindles by using higher capacity disks, and because the increase of capacity per disk also increases the likelihood of data access per volume thus decreasing the opportunities for disk spin-down. Still hardware solutions that allowed disk spin-down of configured volumes were also investigated.

2.2 Infrastructure Improvement

The task main focus was the replacement of the Tier-2 storage system with the goal of maintaining its reliability and availability, while reducing its operational costs.

Taking into account foreseen market developments it was decided to postpone the purchase of the storage renewal until late 2014. This decision aimed to take advantage of the appearance of new products in the market that would enable a better price/capacity ratio thus maximizing the return of the investment. This strategy also provided more time to follow the technology and market evolution and to prepare the requirements and specifications for the purchase. Meanwhile the behavior of the old storage system was carefully monitored and thanks to the capacity reduction performed in 2013 enough spare parts were available to perform repairs.

The new storage architecture was established according to the input of the efficiency and technologies task. The strategy was set to pursue a solution based on disk storage but with higher density and therefore less power consumption. The storage architecture aimed at the following goals:

- Maximize the capacity within the budget
- Maintain (or improve) reliability and availability
- Keep a performance level compatible with the load
- Reduce energy costs
- Provide a scalable solution
- Minimize system management costs

A survey of the hardware solutions of the major vendors was performed to identify their offerings and the communalities between those offerings, so that the purchase requirements could match several vendors and increase the chances for competitive bids. Similarly several typical technologies offered by the vendors were considered:

a) Attachment via servers with SATA/SAS HBA and directly attached disks.
b) Attachment via servers with SAS RAID controller and SAS disk expanders.
c) Attachment via servers with fibre channel and fibre channel disk arrays.

In option a) each server would have a set of internal disks accessible through an HBA. This approach maximizes the overall disk throughput and is often the best approach to implement storage systems such as Ceph. However to provide resiliency with good performance it requires data duplication across disks in multiple servers. This approach also has the disadvantage of forcing physical disk migration between boxes in case of major server failure. Each box is a server and to provide a good granularity for the allocation of storage each server cannot have a high number of disks, this increases the number of servers and their related costs.

The option b) is similar to the Tier-2 setup initially acquired in 2009/2010. This option provides a reasonable tradeoff between fault-tolerance, capacity, performance and price. It allows fault-tolerance at the level of the disk volumes with a maximization of the storage capacity via RAID 5, 6 or others. The RAID controllers constitute a limitation in performance, however if this performance penalty can be acceptable they provide the best solution to maximize capacity with redundancy and minimize CPU impact. Alternatively, software RAID can be used. In case of server failure, the server can be replaced and the new server can be easily attached to the same storage enclosures without moving disks. The use of external enclosures with a reasonable number of disks (eg. 12 – 24) can be valuable to manage the data capacity and how it is assigned to servers.

The option c) is similar to option b) it has all its advantages and can provide more flexibility and dynamic management of the data capacity via the fibre channel fabric. However it implies hardware with higher costs and the introduction of fibre channel switches. The additional layer of fibre channel networking also increases the system management complexity.

Table 1 shows a comparison of the several options, with a classification ranging from 1 (low) to 3 (high) according to the set of established goals.

	Maximizing the capacity within the budget	Maintain reliability and availability	Keep performance compatible with the load	Reduce energy costs	Provide scalability	Minimize system management costs	Total
attachment via servers with SATA/SAS HBA and directly attached disks	3 only with software RAID, ZFS or Ceph erasure	3	3	2 only with software RAID, ZFS or Ceph	3	2 likely more servers to manage	16
attachment via servers with SAS RAID controller and SAS disk expanders	3	3	3	3	3	3	18
attachment via servers with fibre channel and FC disk arrays	2 expensive solution	3	3	2 additional hardware implies more energy	2 depends on fibre channel hardware	2 flexible but requires management of FC fabric	14

Table 1 Comparison of different storage options for the Tier-2

Within the goals previously defined the option that offered the best match was the option b) “attachment via servers with SAS RAID controller and SAS disk expanders”. The requirements for a public tender were established accordingly. The capability of controlling the volumes spin-down was also defined in the requirements as an optional functionality that if present would contribute to a higher bid punctuation.

Changes in the law that regulates public tenders in Portugal caused a delay in the preparation of the tender. The LIP administrative procedures had to be reviewed and adjusted. An electronic platform through which the process needed to be conducted had to be chosen and the personnel had to be trained. This was the first public tender at LIP to go through this process and the computing team participated actively in the implementation of the process flows. Still it was possible to accomplish the full process before the end of 2014. All bids received fulfilled the requirements and had similar technical characteristics.

The tests performed over the first proposal have shown full compliance with the intended performance with a raw capacity of 700TB (the best among all bids). The tests have covered several scenarios stressing the disks, RAID volumes, system performance and network performance. The solution was fully deployed in December 2014 and comprises:

- 5x servers
  - 2x Intel Xeon E5-2630 v3 (8 core processors)
  - 32GB of RAM
  - 2x 10Gbase-SR interfaces (Intel X520)
  - 2x 1000Base-T interfaces
  - 1x management card
  - 1x internal RAID controller
  - 2x 300GB SAS internal system disks
  - 1x external RAID controller (2GB de cache)
  - Redundant power supplies
- 14x disk expanders
  - 12x SATA 4TB disks
  - Redundant controllers and power supplies
- 2x network switches (mainly for management)
  - 24x port

The new hardware was configured to provide about 600TB of usable storage. This capacity is higher than the initially planned. It enabled a full renewal of the Tier-2 including the capacity that had been temporarily reduced in 2013 which was consequently restored in 2015, and will be reflected in the next pledge. This hardware also has the capability of allowing volumes to spin-down when not in use. Although tested this feature is not in use because of the continuous data access to the volumes.

Based on the input from efficiency and technologies task, the Lustre installation was upgraded to version 2.5. Some limitations were found on the enforcement of ACLs when manipulating extended attributes. The team had to develop a patch to fully implement ACL validation in the Lustre kernel components for the manipulation of file extended attributes. This functionality was need by the StoRM grid middleware used to implement the interface between the grid and the Lustre file-system. Additionally several deployment improvements were implemented including:

- Fault-tolerant storage for Lustre metadata based on DRDB volume replication.
- Redistribution of file-system metadata across different metadata servers.
- Reorganization of all storage spaces.
- Merging of the storage systems migrated in the context of the datacenters consolidation.

The Lustre storage at NCG now contains a Tier-2 component for ATLAS and CMS fully hosted on the new storage system, and a Tier-3 component also supporting ATLAS and CMS which is still partially hosted on old storage hardware.

After this purchase an analysis of the most relevant needs was performed. The remaining budget from the storage purchase together with savings obtained from the datacenters consolidation, energy efficiency and from the energy at the LIP Coimbra Tier-2/Tier-3 component (supported by other funds) was shifted to buy complementary equipment. The first budget reorganization included the following items:

- Storage complement: to increase the storage space and achieve a better distribution of the disk storage expanders among the servers. Some additional servers to act as storage network gateways.
  - 1x additional disk expander with 12x 4TB SATA disks
  - 5x additional storage server
  - Purchase of five disks to complete the "last" array purchased in the storage tender

- Computing complement: minor reinforcement of the computing capacity (the first since 2010), especially aimed at improving the Tier-3 data analysis infrastructure for the Portuguese Physicists. Includes two computing servers with larger memory, additional memory for some of the old servers, and small capacity GPGPUs to increase the computing capacity of the servers and for the users to evaluate the benefits of the technology.
  - 2x computing servers with:
    - 2x Intel Xeon E5-2680 v3 (24x cores per server)
    - 192GB of RAM (4GB per thread)
    - 3x 1.2TB SAS disks
  - 48GB of RAM to duplicate the memory capacity of two servers from 24GB to 48GB
  - 5x Nvidia QUADRO K2200 GPUs to install on rack mount 1U compute servers

- Several components: to replace broken parts to keep the old servers still operational.
  - 1x RAID controller
  - 14x batteries for RAID controller caches
  - 14x network controllers to replace old NetXen NICs that frequently fail
  - 24x disks for repairs and enlargement of computing nodes capacity

- Improvement of the network equipment: necessary to properly interconnect the new systems, provide redundancy and spares for failures. Since the Tier-2 uses Force10 C300 modular switches the same type of equipment was purchased.
  - Force10 C300 components

- Increase of the tape library capacity: the Tier-2 tape library had only 240TB of capacity which was insufficient for the Tier-2 and Tier-3 needs. The tape library was improved with two new higher capacity tape drives and some higher capacity cartridges. New fibre channel interfaces were also acquired to connect the two drives to the server.
  - 1x fibre channel HBA
  - 2x LTO-6 fibre channel drives for IBM TS3310

  - 60x LTO-6 data cartridges (150TB) and labels to start the cartridge replacement process

A final budget change request was performed in the last quarter of 2015, to execute the remaining funding obtained from savings achieved in the previous purchases, and also to reuse funding still available in the services budget line. The final budget reorganization approved by FCT in December 2015 included the following items:

- Tape cartridges: acquisition of the remaining LTO-6 tape cartridges to fill in all slots of the Tier-2 tape library.
  - 220x LTO-6 tape cartridges (550TB) and labels for IBM TS3310 tape library

- Additional GPGPU to reinforce computing capacity: A small number of GPGPUs was purchased in the previous budget reorganization aiming at enhancing the Tier-3 computing services. This GPGPU of higher capacity will enable to evaluate a more powerful solution.
  - 1x GPUs Quadro K5200

- SSD disks for caching and acceleration of Tier-2 services. Being used to evaluate the impact of SSD disks on data storage on several scenarios, e.g. with Lustre metadata, for Ceph metadata or Ceph caching pool, for data analysis as local storage space.
  - 13x Intel S3700 SDDs 200GB

- Human resources: to pay the last 3 months of the contracted project member Mario David thus effectively covering his contribution until the end of the project (October, November and December 2015).
  - Keep the Tier-2 operations coordination until the end of the project
  - Finalize documentation
  - Execute the last budget reorganization

2.3 Operation

This task was mainly focused at the continuous operation of the Tier-2 and fulfillment of the WLCG service level agreements.

The consolidation of the three datacenters was a complex operation that was performed in small steps over the three years. The Tier-2 had to continue operational and available to the users during the migration. The LIP Tier-3 is an integral part of the Tier-2, such integration improves efficiency but made the migration more difficult to plan and perform. Furthermore the exit of personnel that maintained the LIP-Coimbra Tier-2/Tier-3 resources reduced the human resources, and left the site without local staff. The first step was an urgent reorganization of the LIP-Coimbra Tier-2/Tier-3 resources so that it could be more efficiently managed remotely by the team based in Lisbon (at a 200Km distance).

The Tier-2 capacity pledge agreed with CERN was temporarily renegotiated in 2013 and reduced by about 15%. This action was needed to decrease the energy costs and to free hardware both for spare parts and to ease the migration of the datacenters. A data consolidation was then performed concentrating the data into a smaller number of disk arrays. Still the team made a huge effort to comply much as possible with the initial agreed pledges. This was achieved by reconfiguring policies and optimizing the resources.

The Tier-2 consolidation started in 2013 with the Tier-2 hosted at the LIP-Lisbon datacenter. Some of the capacity was left at the LIP-Lisbon site as a backup precaution, however the services where shifted to NCG. The consolidation continued during 2014 with the migration of the LIP-Coimbra Tier-2. In 2015 the focus was on Tier-3 computing resources and remaining Tier-2 hardware. The last part of the planned migration was achieved in December 2015. The Portuguese Tier-2 is in fact two Tier-2s that support the ATLAS and CMS experiments managed in an integrated way. The CMS Tier-2 and Tier-3 was fully migrated to NCG in 2014. The ATLAS Tier-2 migration was completed in 2014, while the Tier-3 migration was performed in the last months of 2015. A fraction of the ATLAS Tier-3 equipment and functionalities housed at the LIP Lisbon datacenter, have been kept and will be moved to NCG during 2016.

Over the three years of the project (2013 to 2015) the Tier-2 has delivered about 110% of the capacity pledged to the CERN Worldwide LHC Computing Grid (WLCG), thus fulfilling and even exceeding the agreed capacity. The capacity was delivered both to the ATLAS and CMS experiments with a relation of about 52% to ATLAS and 48% to CMS. Overall 17,267,856 hours were delivered and more than 5,040,000 jobs were executed by both experiments. Figure 4 shows the evolution of the number of processing hours delivered to ATLAS and CMS from 2013 to 2015.

The LIP Tier-2 also participated actively in pilot activities to exploit multicore job execution within WLCG. The NCG site was configured to support the execution of multicore jobs. Considerable capacity has been delivered through this type of jobs. However this is a recent feature within WLCG and is not yet properly reported in the WLCG production accounting. This type of jobs allocates entire computing nodes and allows a better exploitation of the computing nodes hardware capacity. One of the observed advantages is a reduction of the memory footprint in comparison to single core jobs. This is import for the Tier-2 which is still based on computing nodes with only 24GB of RAM. The computing nodes were purchase in 2009/2010 and their replacement will need to be addressed in a future project.

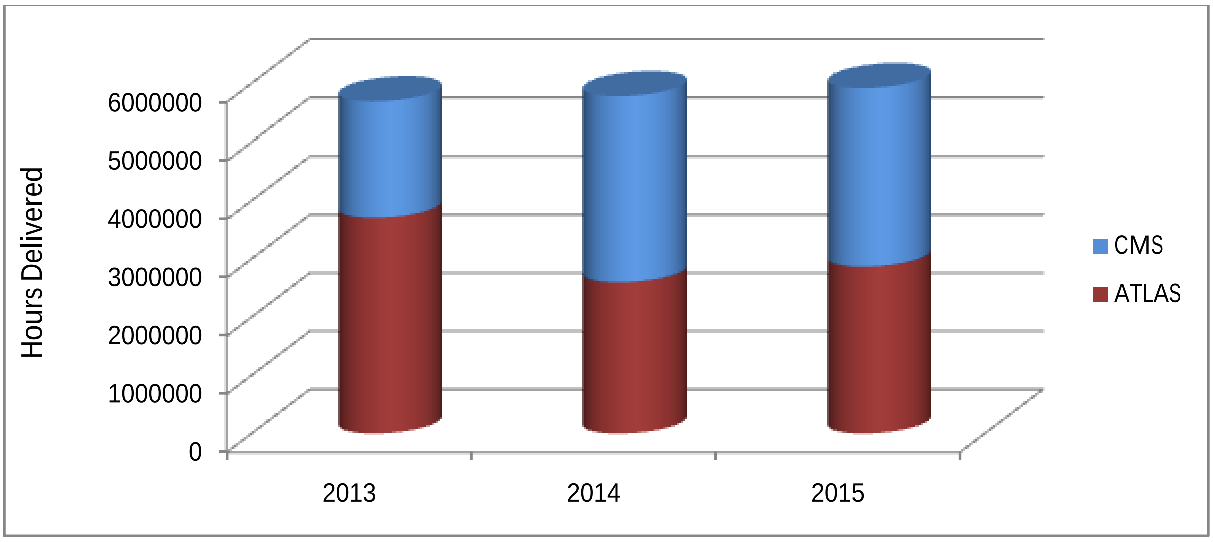

Figure 4 Elapsed processing time in the Tier-2 from 2013 to 2015

Over the three years the average reliability of the Tier-2 was about 95%. The intended target for WLCG Tier-2s is above 90%. Regarding trouble tickets the Tier-2 operations team received 105 tickets via the Global Grid User Support (GGUS) helpdesk. The average response time of the tickets opened against the LIP Tier-2 was generally better than the average response time of all other WLCG Tier-2s, evidencing the dedication and fast response of the LIP grid support team. The Figure 5 shows the evolution per quarter of the average response time measured in working days.

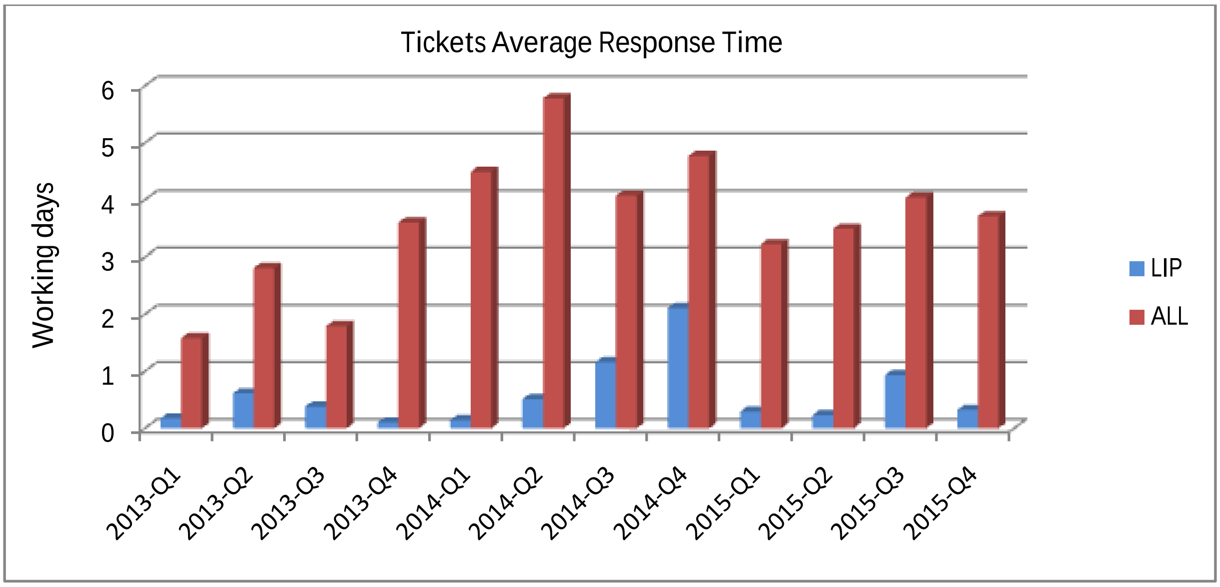

Figure 5 Average response time for tickets opened against LIP and all other WLCG Tier-2s

The Figure 6 shows the average number of trouble tickets opened against the LIP Tier-2 in comparison with the average of all other ATLAS and CMS Tier-2 sites. It is important to understand that most of the WLCG communication is performed via tickets meaning that not all tickets correspond to incidents. In the plot the fourth quarter of 2014 shows an unusually high number of tickets corresponding to a major upgrade and reorganization of the LIP Tier-2 that took place in the Christmas of 2014. Most of the major interventions in the LIP Tier-2 are performed between Christmas and the New Year, in order to minimize the impact on users.

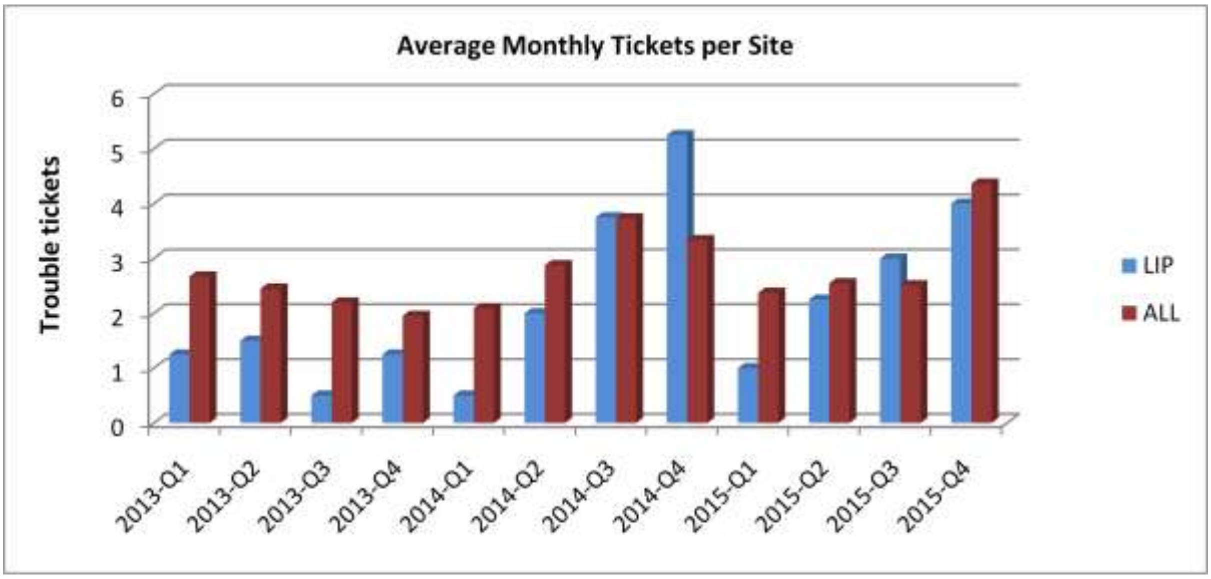

Figure 6 Number of opened tickets against LIP and all other WLCG Tier-2s

The Portuguese Tier-2 is part of a larger national infrastructure composed of computing centres in universities and research organizations, created as a direct result of the Tier-2 R&D activities. This infrastructure shares computing capacity with users from multiple scientific domains both via grid computing interfaces and via local access to the computing resources. Figure 7 shows the percentage of jobs executed in the Tier-2 submitted by several grid virtual organizations (user communities).

Support for additional virtual organizations is performed on request from Portuguese researchers. The supported virtual organizations share the Tier-2 capacity. In total more than 16,500,000 grid jobs were executed during the 2013-2015 period. Of those, about 21% were from non-LHC users, these include among others the Auger experiment, biomed (life sciences) and enmr (structural biology).

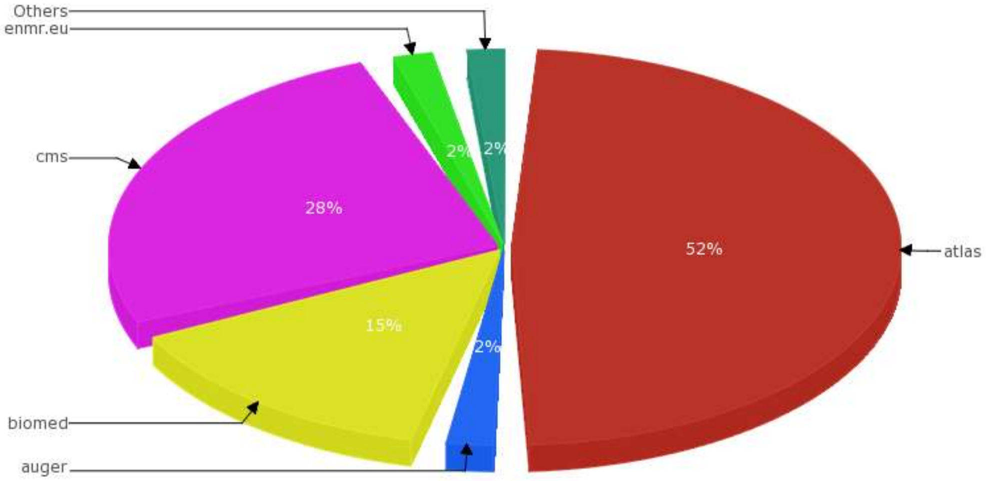

Figure 7 Number of jobs executed in the Tier-2 infrastructure from 2013 to 2015

Figure 8 shows the evolution of the usage of processing time of the Tier-2 by several grid virtual organizations. The Auger experiment was a considerable consumer of the Portuguese Tier-2 resources. The figure also shows Tier-2 usage from thematic virtual organizations of the Iberian Grid Infrastructure (IBERGRID) which joins Portuguese and Spanish computing centres, of which the Portuguese Tier-2 is also member.

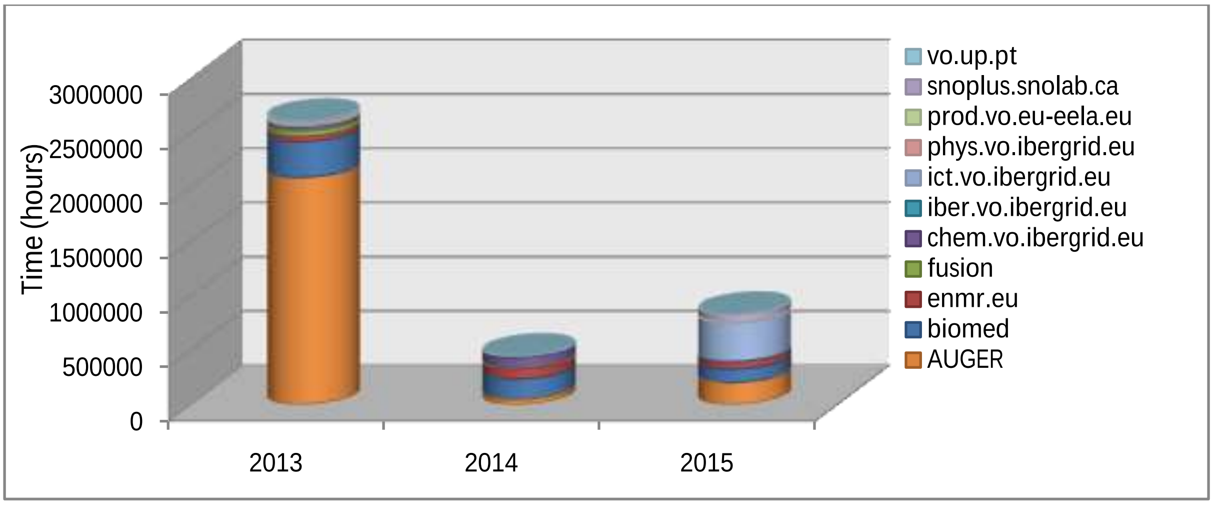

Figure 8 Elapsed computing hours by the several non-LHC communities

Besides the High Throughput Computing (HTC) services a High Performance Computing (HPC) service with low latency Infiniband and Ethernet interconnects is also available within the Tier-2 to support parallel applications. From 2013 to 2015 the HPC service delivered 16% of the total Tier-2 processing time. The HPC service makes use of a specially configured Lustre file system tuned for high performance parallel I/O. This service is delivered as a pilot aimed only at Portuguese researchers mostly from other scientific domains and its provisioning constitutes an additional contribution of the Tier-2 to the community. The service was established based on the observation that many of the national users from other scientific domains require access to parallel processing. The service is delivered using the same farm and storage systems that supports the Tier-2.

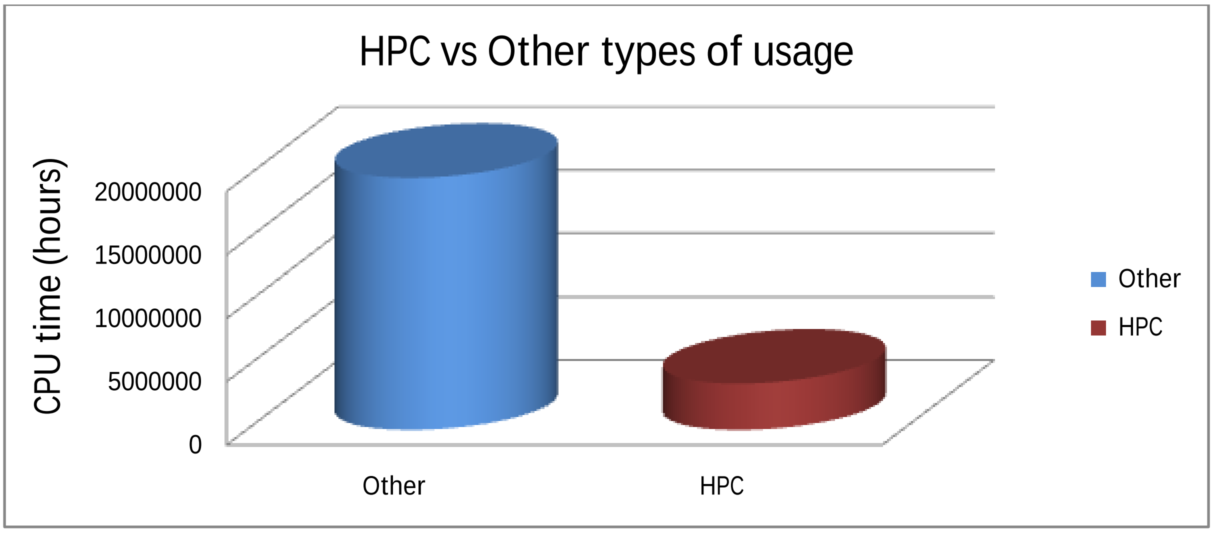

Figure 9 HPC usage in the Tier-2 from 2013 to 2015

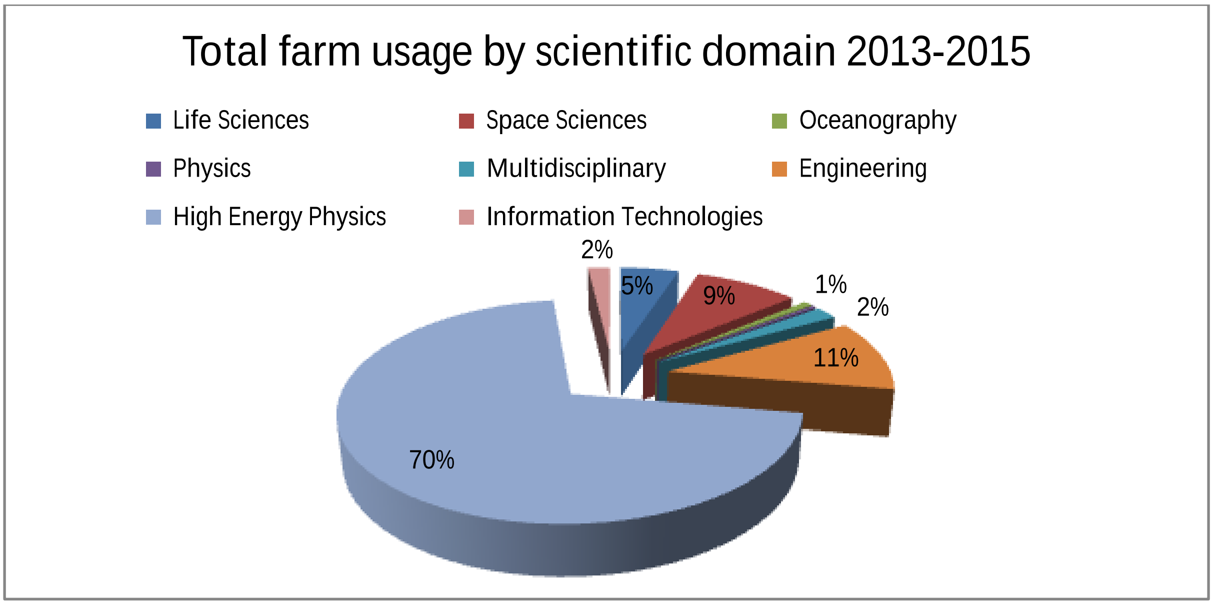

Figure 10 Farm usage by scientific domain including grid and local job submissions from 2013 to 2015

Figure 10 shows the total farm processing time in percentage over the three project years. The data is split per scientific domain and includes all job submissions (grid and local job submissions of both parallel and sequential types). About 30% of the processing time was used by non-HEP users.

The computing team maintains the following set of internal production services at the Tier-2/3 sites:

- Lustre: is the main Tier-2 storage system. The Lustre file-systems are directly mounted on the computing farm thus delivering very good performance and reducing bottlenecks.
- SGE: Son of Grid Engine is the batch scheduler that supports the computing farm. It was chosen due to its rich and flexible scheduling policies. The team contributes to the support of the grid middleware components for the SGE batch scheduler in the European Grid.

- NFS: is used mainly to provide shared storage for common software, also provides storage for user home directories supporting the job submission interactive machines that local users can use to access the farm directly. HPC users have their home on Lustre.
- CVMFS clients and caches: read-only wide area network file system developed by CERN used for distributing the software of the LHC experiments across WLCG sites. The team also maintains CVMFS instances for other communities.

The computing team maintains the following grid oriented services at the Tier-2/3 sites:

- CREAM-CE: service that provides an interface between the grid and the local farm scheduling system. LIP participated in the development of the CREAM-CE for the SoGE scheduler and is also providing international support for the SoGE CREAM-CE.
- StoRM: service that provides an SRM interface between the grid and the local file systems. At the Portuguese Tier-2 the StoRM service is used to allow grid access to the Lustre file systems.
- SiteBDII: component of the grid information systems that allows service discovery and resource information. The SiteBDII caches information from the information systems attached to each service making this information reliably available to the grid.
- ARGUS: authorization policy service meant to render consistent authorization decisions for the grid services. It is used to authorize and ban users at each site.
- GridFTP, WebDAV and xrootd: data access protocols implemented by specific dedicated services usually associated with StoRM (GridFTP and WebDAV) or ran as independent services (xrootd).
- PHEDEX: data management service specific to the CMS experiment which must be available at each site.

In addition the team also maintains the necessary national services to integrate the Portuguese grid sites into a coherent infrastructure, and which support the Tier-2 and all other sites and communities in the country.
Following the evolution of the LHC computing model and of these services is of utmost importance and was an integral part of the activities. Clearly there will be a need to change and/or adapt the current setup in the future. WLCG is in continuous evolution and the Tier-2/3 needs to follow it.

Monitoring improvements were performed at the level of service testing (Nagios), network (Cacti), service monitoring (Ganglia), and new monitoring tools were evaluated namely (Grafite and Grafana) and are currently under test. The Tier-2 team also supports the regional backup Nagios instances which supports the continuous testing of the grid sites in Portugal and Spain and feeds this information into the global EGI monitoring. Also in a related area the team has developed and maintains nsupdater a service used in the Iberian region to implement the TOPBDII grid information system as a fully distributed service with fault tolerance and load balancing across sites in the region. The TOPBDII is the key element that enables the browsing and discovery of grid services including all Tier-2 services. This service has significantly improved the reliability and availability of the grid sites in the region including the WLCG sites. Figure 11 shows the nsupdater architecture. Several nsupdater agents monitor distributed TOPBDII instances and dynamically update a local DNS server with round robin resource record entries for the best TOPBDIIs. The DNS servers are automatically selected by the DNS resolvers based on the round trip time, thus getting the view of closest nsupdater agent.

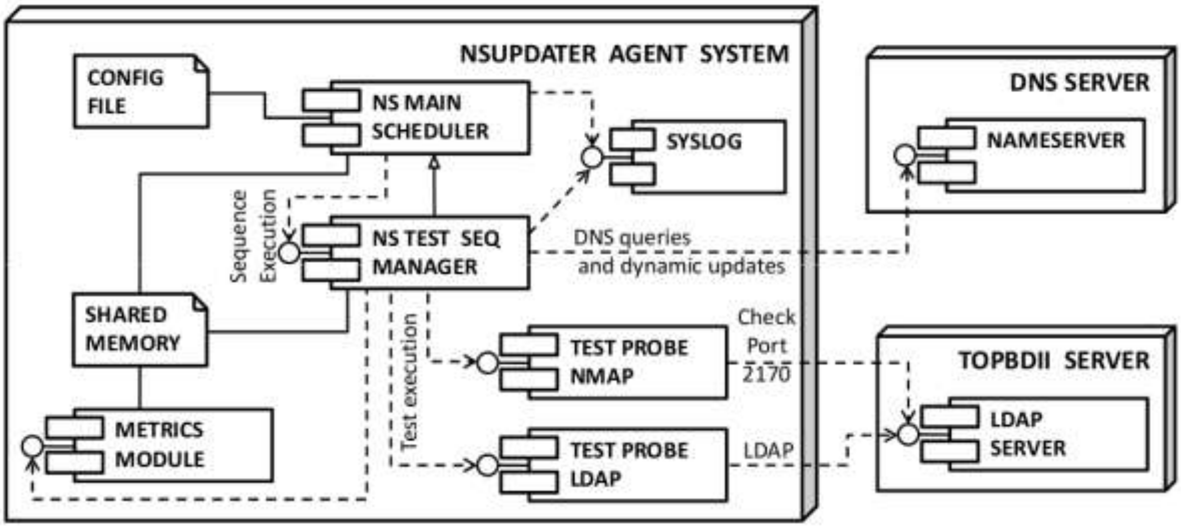

Figure 11 The nsupdater architecture

The Tier-2 network includes a wide area network composed of 10GbE links that interconnect the three Tier-2/3 datacentres enabling also the LHC researchers at LIP-Lisbon and LIP-Coimbra to access both the Tier-2 resources and the WLCG infrastructure. This layer two network infrastructure is operated in partnership with the Portuguese NREN using the NREN private fibre infrastructure. It is connected to the wider Portuguese academic network and to the Geant network at the NCG centre.

In 2014, within the project, a study started to evaluate the possibility of connecting the Tier-2 to the LHCONE private network in collaboration with the NREN and Geant. However the need to clearly separate LHC traffic from the traffic produced by other users raised difficulties. More recently a plan and topology was agreed for LHCONE, but it was postponed until the required hardware would become available and the network traffic separation implemented. This separation is now being implemented by segregating the LIP commodity traffic from the grid traffic. Again this required careful planning and collaboration from the NREN. Several improvements were introduced in the switching infrastructure which was reorganized to support the datacenters consolidation. This included a new smaller switch/router for LIP-Coimbra to enable the migration to NCG of the core switch housed at the centre and additional ports for the NCG core switches. Together with these changes the Tier-2 connectivity to the internet is now being upgraded from 3Gbps to 10Gbps to be finalized in the first quarter of 2016. Backup fibre connectivity was also implemented at the Tier-2/3 locations during 2014/2015.

In 2013 a new firewall/router for the LIP-Tier-2 centres was developed exploiting iptables, ipset, quagga and other related functionalities and aimed at replacing commercial hardware which had expensive maintenance costs. The implementation was successful but was built using old servers that didn't have adequate packet forwarding capacity. The upgrade was thus postponed until better hardware would be purchased, in addition the evolution of NFTABLES as a potentially better alternative to IPTABLES was followed. With the consolidation of the datacenters the network load of the Coimbra site decreased substantially, therefore the router/firewall of the Tier-2 at LIP-Coimbra was replaced in 2014 using this same approach. The LIP-Lisbon site router/firewall which still host part of the ATLAS Tier-3 is also being replaced by making use of more recent and performing Linux server purchased within the project late in 2015. The connectivity of LIP Lisbon is more complex and requires additional capacity as the site holds the backup tape library for the daily off-site backups. The firewalls at the NCG centre are also based on the same approach.

The operation task also included the maintenance and improvement activities need to keep the datacenters operational during the project. These included the maintenance of network equipment, HVAC systems, and replacement of UPS batteries.

The Tier-2 relies on a set of middleware and software services previously described in this section that needs to be delivered in reliable and flexible way. Many of the Tier-2 supporting services are delivered via virtualization. Initially The Tier-2 relied on Xen hypervisors whose data storage was backed by proprietary iSCSI storage arrays. The iSCSI Storage Area Network together with a GFS file-system enabled easier migration of the services virtual machine images across a cluster of physical hosts. This setup was present at each of the three datacenters. During the project this setup received several improvements. The hypervisors were migrated to KVM facilitating the deployment under CentOS and the overall system management. Several performance and reliability issues affected the iSCSI based storage forcing the evaluation of alternative solutions. In order to decrease costs and improve sustainability it was decided to focus on open source storage solutions. A new approach towards the delivery of storage for the virtualized services was designed based on DRDB block devices replicated across the network. With this approach images are stored locally in the physical hosts. Simultaneously the data written to the local storage is replicated to another host enabling faster recovery in case of host failure. This solution has improved significantly the performance, scalability and reliability of the virtualized services. After careful testing was deployed at all sites and later was also applied to the Lustre metadata servers.

2.4 Physics Analysis Improvement

This task aimed at improving and supporting the LIP physics analysis environment for the LHC experiments that sits on Tier-2/Tier-3 resources.

Taking advantage of the hardware freed with the storage upgrade the Tier-3 storage was improved. The best storage arrays and servers were reused for the Tier-3 storage. A new Tier-3 Lustre storage was deployed and data migrated. To address the data access limitations of some applications running in the Tier-3 improvements were performed such as increasing the number of spindles, tuning the disk servers and using faster RAID levels.

Optimizations to run the ROOT analysis framework were considered. The interactive Linux cluster used to run most of the interactive analysis including ROOT was enlarged. This was possible thanks to the hardware that resulted from the Tier-2/3 consolidation and thus results in a more scalable environment for data analysis. The installation of Parallel ROOT (PROOF) was also considered. The recommended way to use PROOF is now via PROOF on demand (PoD), this was the selected approach to support PROOF at the LIP Tier-3 via an SGE parallel environment. Since the farm is integrated with the Tier-3 storage no further improvements were required besides the ones already mentioned for the Tier-3 storage.

The virtualization and cloudification of resources via Openstack was evaluated in order to support and improve the data analysis. Two approaches were identified.

The first approach consists in the migration of the interactive Linux cluster to the cloud resources which would facilitate the instantiation of additional nodes or different machine configurations when needed. In fact several interactive clusters are being operated to address specific sets of users, and some of them are already being delivered on virtualized resources. The major issue observed is the I/O performance especially in terms of network operations. The use of several virtual machines per physical host implies the use of software network switches which together with the two network stacks of the host and virtual machine introduce performance constraints. In this aspect several tests were performed with Openstack to understand its network behavior. The use of VLANs together with Single Root I/O Virtualization (SRIOV) was tested with Openstack (Neutron networking) using hardware provided by FCT-FCCN. This enabled the provisioning of virtual PCI devices to the virtual machines providing almost line rate speed with all the benefits of the Openstack network virtualization. This approach is depicted in Figure 12 and will be probably pursued do deliver the interactive Linux clusters for analysis, for this purpose adequate hardware will have to be purchased. The same approach may be applied to the virtualization of the Tier-2/3 compute nodes. The direct use of VLANs also seems the best approach to interconnect the Openstack virtualized machines with the remaining Tier-2 services including the storage network.

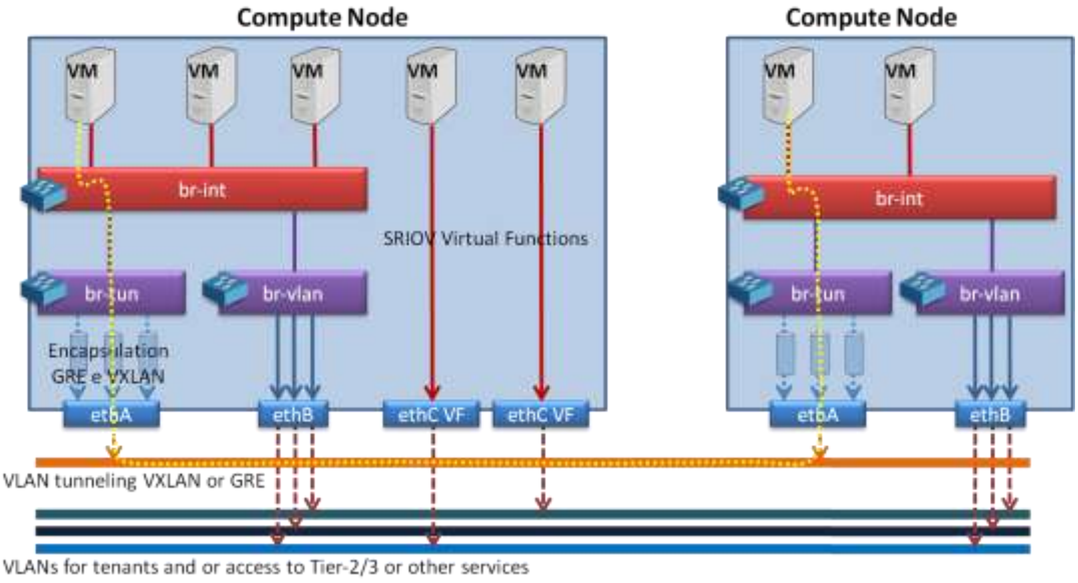

Figure 12 Overview of tested scenario for Openstack networking

The second approach to virtualization is the possibility of allowing the users (or research groups) to instantiate their own virtual machines in the cloud. This approach also looks viable due to the Openstack multi-tenancy capabilities. It may allow research groups to build specific services for their own use without the intervention of the LIP computing team. However few users and groups will have the human resources (time and knowledge) to manage such virtualized environments. Therefore the first approach seems the most suitable to the Tier-3 virtualization and this will be pursued.

The initial path of exploiting the TIMBUS project developments for digital process preservation conducted to an evaluation of the LIP environment covering both the organization as a whole and also the data analysis processes. An identification of stakeholders their relations and the risks associated to the processes was performed. However the actual software solutions developed by TIMBUS were oriented to more static processes that take place in the corporate world, and did not seem adequate to the physics analysis environment. However several of the recommendations regarding data preservation were taken into account and incorporated. Some examples are given below.

Users are now warned not to delete old files without consulting with the LIP computing services. They are warned about the implications in terms of data preservation and options are discussed. Profiting from the hardware made available with the storage renewal, a system aimed at long-term storage was implemented allowing users to copy data to storage spaces that are managed by the LIP computing and that are powered up on demand. In addition a new storage service is being planned to take advantage of the old LTO-4 tape drives and tape cartridges (240TB) recently made available (December 2015) with the full upgrade of the tape library to LTO-6. This service will allow data to be copied to these tapes for long-term storage.

The preservation and replay of analysis steps was also considered. The use of digital logbooks is now recommended to the users so that they can more quickly and effectively store and keep trace of their actions and research. Based on the Openstack experimentation, the use of cloud technology was considered to provide access to older computing environments and enable their instantiation by the users (either directly as previously mentioned or automatically through the virtualized compute nodes in the Tier-2 farm). This may constitute a good approach to digital processes preservation and their replay, it will enable a much easier execution of past analysis steps developed in older operating system versions and environments. Images of past versions of compute nodes are already being kept.

The use of Linux containers is being pursued both as a way to provide more efficient virtual compute nodes and as a way to encapsulate applications and their environments without the overheads of full virtualization. The use of a driver for the execution of Docker containers has been tried in the Openstack testing cloud with very good success. The team is collaborating with the INDIGO-DATACLOUD project where some of these technologies will be applied and further developed. In this context a tool was created to facilitate the execution of Docker containers by end users. The tool allows the download and basic execution of Docker containers without using Docker and without some of its data storage overheads (related to layered file systems). The tool makes use of PRoot (http://proot.me/) to create user level chroot like environments where containers can be executed without Docker. It has been successfully tested in a wide range of operating systems and can be deployed by the end user without the system administrator intervention. The execution of Docker in traditional batch systems is also being evaluated both using this tool and also through full Docker integration in the batch systems. This aims at proving a more flexible and efficient compute environment for the users while addressing some of the preservation concerns.

3 Other activities

3.1 International collaboration

IBERGRID is a joint collaboration of Spain and Portugal in the area of distributed computing. Since the early days of the European grid, both countries have shared responsibility for the infrastructure, combining expertise to provide the best support to their users. In 2007 IBERGRID came along at the level of official political agreement between the governments of both countries to formalize the collaboration. The combined effort has been very successful in getting priority research areas onboard. These include life sciences, environment, astrophysics, materials science, meteorology, satellite remote detection, seismology and biomedical research. The current infrastructure is distributed across more than 20 sites, with a total capacity of ~ 24K CPUs and supporting grid and cloud computing. LIP has a major role in the IBERGRID operations. During the 2013-2015 period the computing resources in Portugal and Spain (including the WLCG resources in both countries) continued to be operated seamlessly under the IBERGRID consolidated infrastructure. The infrastructure services are fully distributed across both countries and the operations effort was shared.

Is through IBERGRID that the Portuguese and Spanish grid sites including the Tier-2 and other WLCG sites are integrated in the European Grid Infrastructure (EGI). EGI provides the operations and organizational integration that joins the European WLCG sites. Portugal is an official member of EGI through the Portuguese Science Foundation (FCT). During the timeframe of this project, LIP represented the Portuguese Science Foundation in the EGI council. LIP also acted as technical liaison and coordinator for the integration of Portuguese sites in EGI.

The team submitted international bids for the provisioning of EGI global services such as the middleware coordination and the 1st and 2nd level user support. The team won all the bids. For the 2015 bid the team opted for focusing the effort in the middleware coordination dropping the 1st and 2nd level support which was requiring too much effort. Again the team won the bid and will continue delivering middleware services to the EGI infrastructure. These responsibilities result directly from the expertise of the LIP computing team in the technical aspects of distributed computing related to the Tier-2 operations. They are also a positive direct contribution to WLCG and to the international research community using EGI.

The participation in the middleware validation and quality assurance activities has been extremely valuable for the Tier-2. They allowed the team to have early access to new middleware versions, learn about their capabilities, flaws, plan their deployment in advance and provide direct feedback to the developers. The collaboration with developers and other sites in these activities enabled the team to establish close collaborations and links. These advantages are translated to the Tier-2 in better quality of service and faster response times.

Figure 13 shows the IBERGRID grid and cloud resource centers in Portugal and Spain. Many of these sites are also contributing to WLCG. Their operation and management is jointly coordinated by LIP and IFCA. The Portuguese Tier-2 sites are fully integrated in the structure IBERGRID => EGI => WCLG.

The Portuguese Tier-2 centres appear as “LIP LX”, “LIP C” and “NCG”. The centres marked with a different color (LIP LX and IFCA) represent the coordination centres.

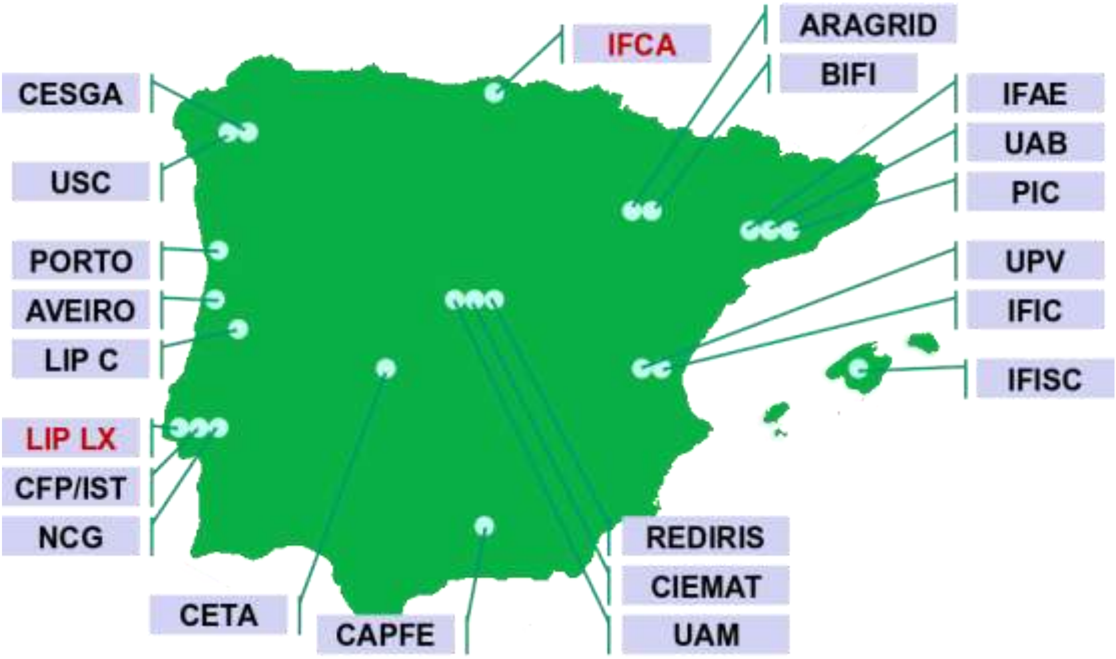

Figure 13 IBERGRID map

The team also participated in several activities related to authentication, authorization and security for the grid. These included: the participation in the EGI CSIRT during 2013 and 2014 coordinating intrusions and vulnerabilities in the grid sites, the operation of the LIP grid Certification Authority that issues digital certificates to Portuguese grid users and services, and the participation in IGTF/EUgridPMA activities defining international authentication policies for grid computing. The fulfillment of these activities is extremely important to keep the Tier-2 infrastructure operational.

3.2 Events

The LIP team jointly with IFCA organized the EGI technical forum and 7th Iberian grid infrastructure conference that took place in Madrid in September of 2013. The event joined 500 distributed computing experts and researchers from Europe and elsewhere.

Together with the University of Aveiro organized the 2014 edition of the IBERGRID conference. The conference joined 70 distributed computing researchers and users from Portugal and Spain. Also collocated with the conference, a workshop and tutorial on Openstack was organized by the team in partnership with FCCN (Portuguese NREN).

Together with the University of Minho organized the 2014 edition of the CERN School of Computing which took place in Braga and counted with 61 international students.

Jointly with EGI the team organized the EGI Conference 2015 in Lisbon. The conference joined 300 distributed computing experts and researchers and users from Europe and elsewhere.

The 2016th edition of the WLCG workshop will take place in Lisbon in February 2016 and is being organized by the team.

4 Future

In September 2013, a proposal to evolve the Tier-2 into a larger national infrastructure with a broader range of users and services was submitted in the context of the Portuguese Science Foundation Roadmap of Research Infrastructures. The proposal for a National Distributed Computing Infrastructure (INCD) was coordinated by LIP and included FCCN (Portuguese NREN) and LNEC (Portuguese Civil Engineering Laboratory) as partners. The proposal was submitted with the support of the Universities of Porto, Minho and Aveiro which have been collaborating with the LIP computing team in the context of national grid activities.

The INCD infrastructure will encompass the current Tier-2 infrastructure as well as additional services to be developed such as cloud computing and data analysis oriented services. INCD aims to support the Portuguese research and academic community and also its participation in large flagship projects such as the LHC and the ESFRI infrastructures. In this context the team is already engaging with Portuguese communities involved in the ESFRIs. The delivery of the Tier-2 services to external (non-LHC) users is already starting to take place under the INCD branding. This is especially true for the cloud related pilot activities which also exploit and share Tier-2 resources.

The INCD proposal was approved as a digital infrastructure and received the maximum evaluation. During 2014 and 2015 additional evaluation steps have been performed by FCT to evaluate the maturity of the infrastructures and prioritize the investment. Overall, the INCD propositions have been well received and a restructured budget for funding was finally submitted in December 2015.

Unfortunately the Portuguese Science Foundation Roadmap of Research Infrastructures cannot fund the infrastructures operation. Due to funding restrictions the roadmap is focused (at least during the next years) in the research and development towards the creation of the infrastructures. This is a major issue for the Tier-2 which is an operational infrastructure that requires continuous support. The R&D components of the Tier-2 will be as much as possible supported through INCD, at least in what concerns generic functionalities common to other user communities. The development of LHC specific functionalities and the actual operation of the Tier-2 will have to be funded by other sources. Furthermore it is yet not known how INCD will be funded and when will the funding become available. It is therefore fundamental to cover the gaps between this project that finished in December 2015, and a future INCD project. Furthermore the operations of the Tier-2 need to be funded in complement to INCD.

5 Conclusions

The project was fundamental to keep the Tier-2/3 operational, fulfilling the obligations assumed by LIP and by the Portuguese authorities in the LHC computing (Memorandum of Understanding for Collaboration in the Deployment and Exploitation of the Worldwide LHC Computing Grid), and to support the research activities of Portuguese researchers in the ATLAS and CMS experiments. The computing support is an integral component of the LHC research program. Without the Portuguese Tier-2/3 the last 20 years of investment in the LHC would have been lost. The LHC confirmed the existence of new particle compatible with the Higgs Boson and Portugal through LIP has participated in this huge result.

In addition the Tier-2 is sharing capacity with many other research projects and scientific domains, both from LIP (section 6.2) and from other national research organizations (section 6.3), such as the Portuguese SNO team, which recently received the Fundamental Physics Breakthrough Prize for their participation in the discovery of the neutrinos oscillation.

Besides the scientific outcome of these and other supported research activities, the project enabled a complete renewal of the Tier-2 data storage system which was getting obsolete and unreliable. In addition the project enabled a complete consolidation and reorganization of the Tier-2/3 services contributing to a better efficiency and sustainability. Through gains in efficiency the projects further enabled the renewal of the offline tape storage and several other important components.

The scientific activity included several new topics of relevance for the evolution of the Tier-2, including new storage technologies, cloud computing, GPU computing and others. Although not initially planned a master thesis on storage systems which included the evaluation of the Ceph storage system was successfully performed.

The project paved the way towards a future national computing infrastructure that will integrate the Tier-2 and many of its technologies and resources to deliver services open to the academic and scientific communities.

6 Appendix - Metrics

The project aims to support the participation of the Portuguese physicists in the LHC experiments ATLAS and CMS, by providing the computing capacity needed for their research activities and by fulfilling the LHC computing pledges agreed with CERN in the context of the Memorandum of Understanding for Collaboration in the Deployment and Exploitation of the Worldwide LHC Computing Grid. Therefore its main scientific impact is the successful Portuguese participation in the ATLAS and CMS experiments and the related results. In this section an overview of several metrics is provided.

6.1 Project metrics

Summary of project metrics related to the project computing team activity as defined in the proposal.

Metric	Value	Comment
Books:	3	conference proceedings
Articles international journals:	2	with published status
Communications international	5	communications in conferences
Reports	3	Tier-2 yearly reports
Seminars and conferences	5	events organized
Tier-2/3 centres	3	centres consolidated along the project
Pilot installations	2	grid and cloud installations
Computing applications	4	Sparks – farm management Nsupdater – TOPBDII balancer Nodocker – Docker without Docker io2 – file system operations monitoring
Thesis	1	Msc thesis
Portuguese physicists benefiting from Tier-2/3	165	In average along the three years
LHC experiments benefiting from Tier-2/3	2	ATLAS and CMS Other non-LHC experiments also benefitted via grid: AUGER, SNO, COMPASS via local access: AMS, HADES, etc

6.2 LIP users

All LIP research groups benefit from the Tier-2 and Tier-3 facilities either directly by submitting jobs and using shared storage space, either indirectly by using services and resources created for the Tier-2 but shared such as the network infrastructure.

Project	Researchers	Post-Docs	PhD students	Master students	Graduate students
ATLAS	16	3	10	5	2
CMS	7	4	7		
LHC phenomenology	13	2	1	2	
COMPASS	3	3	2	1	
HADES	2	2			
Computing research	10	1		2	
AMS	1	1		2	
SNO+	5				
Dark Matter Search	4	3	1	2	
High Energy Cosmic Rays	14	5	2	2	
RD51	5		1		
Neuland – R3B	3				
Neutron detectors	5	1	1		
NEXT	7		1		
Ion Transport Processes	5	1	1		
ICNAS	3		2		
PET mammography	5	1	6	1	
Human PET	5	1	6	1	
MC in Medical Physics	7	1		1	1
Orthogonal Ray Imaging	1		2	1	
Gamma cameras	6	2	2		
Space Physics	4	2	1	1	
AHEAD	6	1	1	2	
Total different researchers	94	27	36	19	3

Table 2 LIP researchers and groups benefitting from the Tier-2 and Tier-3 infrastructure

6.3 Portuguese users

The following Portuguese research and academic organizations used Tier-2 resources between 2013 and 2015. This list does not contain any LIP related activities or collaborations.

- Instituto de Medicina Molecular (IMM) – multiple projects related to genome and cancer
- Universidade do Algarve (UALG) – genome assembly
- Instituto Gulbenkian the Ciência (IGC) – biology, genome
- Fundação Champalimaud – neurology, genome
- Instituto de Biofísica e Engenharia Biomédica (IBEB) – electroencephalogram classification
- Instituto Superior de Agronomia (ISA) – modelling of climate impact on forestation
- Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência (INECTEC) – brain research
- Rede de Química e Tecnologia (REQUIMTE) – chemistry and biochemistry
- Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência (LIAAD) – artificial intelligence, machine learning
- Laboratório Nacional de Engenharia Cívil (LNEC) – civil engineering, earth sciences, coastal and river simulations
- Universidade de Coimbra (DEI) – desktop grids, EDGES
- Instituto Superior Técnico (LASEF) – fluids simulation and turbulent plane jets
- Instituto Superior Técnico (GDNL) – fluids simulation CDF framework (GERRIS, GTS)
- Centro de Ciências do Mar (CCMAR) – biophysics simulation, protein folding
- Centro de Investigação Marinha e Ambiental (CIMA) – ocean pollution simulation (SWAN)
- Faculdade de Ciências da Universidade de Lisboa (BIOFIG) – microbiology, biotechnology, Genomics (Velvet, Cortex, SAMTools,ABySS, SOAP)
- Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (FCT-UNL) – chemistry, nuclear magnetic resonance (NMR)
- Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (FCT-UNL) – transactional memory systems
- Centro de Investigação em Materiais Cerâmicos e Compósitos (CICECO) – computational chemistry (VASP, Gromacs, Gaussian)
- Instituto de Tecnologia Química e Biológica (ITQB) – transport in biological membranes protein-protein interaction, protein folding, protein Dynamics
- Universidade do Porto – Multiple scientific domains via UP grid VO
- Instituto de Plasmas e Fusão Nuclear (IPFN) – Nuclear fusion

6.4 Publications by the project team

Publications with direct contribution of the team members:

- Published - Validation of Grid Middleware for the European Grid Infrastructure
  Mário David, Gonçalo Borges, Jorge Gomes, João Pina, Isabel Campos Plasencia, Enol Fernández-del-Castillo, Iván Díaz, et al., Journal of Grid Computing, 2014, 10.1007/s10723-014-9301-z
- Published - Analyzing File Access Patterns in Distributed File-systems
  J.Gomes, J.Pina, G.Borges, J.Martins, N.Dias, H.Gomes, C.Manuel
  7th Iberian Grid Infrastructure Conference Proceedings pp:89-101, ISBN:978-84-9048-110-3
- Published - SPARKS, a dynamic power-aware approach for managing computing cluster resources
  J.Martins, G.Borges, N.Dias, H.Gomes, J.Gomes, J.Pina, C.Manuel
  7th Iberian Grid Infrastructure Conference Proceedings pp:3-15, ISBN:978-84-9048-110-3
- Published - Phenomenology tools on cloud infrastructures using OpenStack
  I.Campos, E.Fernández-del-Castillo, S.Heinemeyer, A.Lopez-Garcia, F.Pahlen, G.Borges
  The European Physical Journal C, vol 73 (4), pp:1-17, ISSN: 14346044
  10.1140/epjc/s10052-013-2375-0
- Published - IBERGRID 2012 6th IBERIAN GRID INFRASTRUCTURE CONFERENCE PROCEEDINGS
  Ignacio Blanquer, Isabel Campos, Gonçalo Borges, Jorge Gomes,
  ISBN: 978-989-98265-0-2
- Published - 8th Iberian Grid Infrastructure Conference proceedings;
  Ilidío Oliveira, Jorge Gomes, Isabel Campos, Ignacio Blanquer, IBERGRID 2014 ISBN: 978-84-9048-246-9, http://www.lip.pt/~jorge/PAPERS/IBERGRID2014-conference-proceedings.pdf
- Published - Exploring Containers for Scientific Computing;
  J.Gomes et al, IBERGRID 2014, ISBN: 978-84-9048-246-9, pp 27-38 , Sep 2014,
  http://www.lip.pt/~jorge/PAPERS/Ibergrid2014_Exploring_Containers.pdf
- Published - Py4Grid, a user centred tool for the long tail of science;
  G.Borges et al,
  IBERGRID 2014, ISBN: 978-84-9048-246-9, pp 65-76 , Sep 2014,
  http://www.lip.pt/~jorge/PAPERS/Ibergrid2014_Py4Grid.pdf
- Submitted - Running high resolution coastal models in forecast systems: moving from workstations and HPC cluster to cloud resources;
  J. Rogeiro, M. Rodrigues, A. Azevedo, A. Oliveira, João Paulo Martins, Mário David, João Pina, Jorge Gomes, Nuno Dias
  International Journal of Advances in Engineering Software (ADES)
- Published - Enabling and Sharing Storage Space Under a Federated Cloud Environment
  Pedro Miguel Simões Miranda, coordinator at LIP: Jorge Gomes
  Informatics Engineering Master Thesis
  Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (FCT/UNL)

6.5 Publications from other LIP beneficiaries

Publications metrics from LIP members of ATLAS, CMS and others made possible through this project are highlighted in Table 3. The table covers the period of 2013 to 2014. Data for 2015 was not yet available in January 2016 when this report was produced.

All LIP research groups benefit from the Tier-2 and Tier-3 facilities either directly by submitting jobs and using shared storage space, either indirectly by using services and resources created for the Tier-2 but shared such as the network infrastructure.

LIP group	Publications			Conferences			Thesis		
	Jrn-I	Jrn-II	other	int.o	int.p	nat	Bsc	Msc	PhD
ATLAS	170	14	24	10	2	7		2	1
CMS	172	17	30	28	1	2			
LHC Pheno	12	12	5						2
COMPASS	14	12	23	17		2			
HADES	18	2							
AMS	4		9	3	1			1	
SNO+	5	2	5	6	1				
Dark Matter Search	8	5	20		7			1	
HECR	20	5	13	13					1
RD51	7	6							
NeuLand	2	2		1					
GEMs	3	1							
NEXT	6	3		1					
Ion transf processes	5	5							
ICNAS	1	1	1		1	1			
PET- mammography	6	5	5	10	8				4
Human PET	3	3		2					1
MC in medical phys	5	5	8	8	4	5			
Orthogonal ray imag	2	2	7	2	3	1			
Gamma cameras					3			1	
RAD4LIFE	2	2		2		2			2
Space	2	2	1	6	3				
Dual	1	1							
AHEAD			1		3	1			

Table 3 Publications and thesis from LIP research groups

Legend:

- Jrn-I: Publications in international journals with scientific peer review co-authored by LIP members
- Jrn-II: Subset of publications Jrn-I in which LIP members had a major responsibility
- Other: Internal notes, conference proceedings, etc. with direct involvement of LIP members
- Int.o: Oral presentations by LIP members in international conferences
- Int.p: Poster presentations by LIP members in international conferences
- Nat.: Presentations by LIP members in national conferences