



GISELA

2ND INTERMEDIATE REPORT ON THE JOINT RESEARCH ACTIVITY

EU DELIVERABLE: D6.2

Document Full name	GISELA-D6.2-v1.2
Date	31/08/2012
Activity	WP6 / Infrastructure and Applications-oriented Services for User Communities
Lead Partner	UFCG
Document status	APPROVED
Classification Attribute	PU (Public)
Document link	http://documents.gisela-grid.eu

Abstract: This document contains an assessment of the work carried out in the context of the Joint Research Activity, Work Package 6 “*Infrastructure & Application-oriented Services for User Communities*” of the GISELA project during the second year of its execution. It describes the main achievements attained and compares these results with the outcome planned for the period in the Description of Work.



Copyright notice

Copyright © Members of the **GISELA** Consortium, 2010

GISELA (“Grid Initiatives for e-Science virtual communities in Europe and Latin America”) is a project co-funded by the European Commission as an Integrated Infrastructure Initiative within the 7th Framework Programme. **GISELA** began on 1st September 2010 and will run for 2 years.

For more information on GISELA, its partners and contributors please see www.gisela-grid.eu.

You are permitted to copy and distribute, for non-profit purposes, verbatim copies of this document containing this copyright notice. This includes the right to copy this document in whole or in part, but without modification, into other documents if you attach the following reference to the copied elements: “Copyright © Members of the **GISELA** Consortium, 2010. See www.gisela-grid.eu for details”.

Using this document, in a way and/or for purposes not foreseen in the paragraph above, requires the prior written permission of the copyright holders.

The information contained in this document represents the views of the copyright holders as of the date such views were published.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS “AS IS” AND ANY EXPRESSED OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE MEMBERS OF THE **GISELA** COLLABORATION, INCLUDING THE COPYRIGHT HOLDERS, OR THE EUROPEAN COMMISSION BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THE INFORMATION CONTAINED IN THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Delivery Slip

	Name	Partner/Activity	Date	Signature
From	WP6	CLARA / WP6 - Infrastructure and Applications-oriented Services for User Communities		
Reviewed by	Technical Board			
Approved by	Management Board		31/08/2012	B. Marechal Ph. Gavillet S. Jalife Villalón Luis A. Nuñez R. Barbera Ramon Diacovo

Document Log

Issue	Date	Comment	Author
0-1	14/06/2012	First draft (template)	B. Marechal
0-2	05/08/2012	Sections 2, 3, 4 and 5 inserted	F. Brasileiro, V. Hammar and H. Castro
1-2	20/08/2012	Final review and approval	B. Marechal

Document Change Record

Issue	Item	Reason for Change

TABLE OF CONTENTS

1. INTRODUCTION	5
1.1. PURPOSE OF THE DOCUMENT	5
1.2. DOCUMENT ORGANISATION	5
1.3. APPLICATION AREA	5
1.4. DOCUMENT AMENDMENT PROCEDURE	5
1.5. GLOSSARY	6
2. EXECUTIVE SUMMARY	7
3. ACHIEVEMENTS	8
3.1. SUPPORT AND CUSTOMISATION OF ALREADY AVAILABLE SERVICES	8
3.1.1. <i>DIRAC</i>	8
3.1.2. <i>OurGrid</i>	8
3.2. DEVELOPMENT AND SUPPORT OF NEW SERVICES	9
3.2.1. <i>Execution of data-intensive applications based on the MapReduce paradigm</i>	9
3.2.2. <i>Seamless execution of applications in hybrid e-Infrastructures</i>	10
3.3. DISSEMINATION ACTIVITIES	12
4. HUMAN EFFORT	14
5. CONCLUSIONS	15

TABLE OF FIGURES

Figure 1: OurGrid JSAGA Adaptor for job submission	9
--	---

TABLE OF TABLES

Table 1: WP6 Human Resources	14
------------------------------------	----

1. INTRODUCTION

1.1. PURPOSE OF THE DOCUMENT

This document describes the main achievements of the research and development activities carried out in the context of the Joint Research Activity, Work Package 6 (WP6) of the GISELA project during its second year of execution - from 01/09/2011 (M13) to 31/08/2012 (M24). It also discusses the main difficulties that have been faced and how they have been addressed.

For a comprehensive view of the Project and of the GISELA Consortium, the Description of Work (DoW)¹ and the Consortium Agreement (CoA)² should be consulted.

1.2. DOCUMENT ORGANISATION

The work developed by the WP6 team is mainly related to the support of the application-oriented and infrastructure-oriented services developed in the context of the EELA-2³ project, and the research leading to the development of new services that can enhance the experience of both application users and system administrators that use the GISELA infrastructure and other similar systems. We present an Executive Summary in Section 2. In Section 3, we describe the activities performed in the second year of the project execution. Then, in Section 4, the human effort invested to conduct these activities is shown. Finally, Section 5 presents some conclusions.

1.3. APPLICATION AREA

The target audience for this document is:

- The members of the Project;
- The European Commission Services;
- The Project Reviewers;
- The External Advisory Committee (EAC);
- The general public.

1.4. DOCUMENT AMENDMENT PROCEDURE

Amendments to this document can be requested by any Project Member to the Project Coordinator, via the Project Office (hlp-gisela@hlpdeveloppement.fr).

¹ <http://documents.gisela-grid.eu/record/32?ln=en>

² Consortium Agreement (CoA) available upon request to the GISELA Project Office (hlp-gisela@hlpdeveloppement.fr)

³ <http://www.eu-eela.eu/>.

1.5. GLOSSARY

API	Application Programming Interface
BeeFS	Beehive File System
BOINC	Berkeley Open Infrastructure for Network Computing
CNRS	Centre National de la Recherche Scientifique
CPPM	Centre de Physique des Particules de Marseille
CoA	Consortium Agreement
CPU	Central Processing Unit
DCI	Distributed Computing Infrastructure
DIRAC	Distributed Infrastructure with Remote Agent Control
DoW	Description of Work
EAC	External Advisory Committee
GFS	Google File System
GISELA	Grid Initiatives for e-Science in Europe and Latin America
GSG	GISELA Science Gateway
HDFS	Hadoop Distributed File System
IN2P3	Institut National de Physique Nucléaire et de Physique des Particules
INFN	Istituto Nazionale di Fisica Nucleare
JSAGA	Java Simple API for Grid Applications
LAN	Local Area Network
MPI	Message Passing Interface
NFS	Network File System
POSIX	Portable Operating System Interface for Unix
SAGA	Simple API for Grid Applications
UFCEG	Universidade Federal de Campina Grande
UNIANDES	Universidad de los Andes
VM	Virtual Machine
WP2	Work Package 2: Dissemination and Outreach
WP3	Work Package 3: User Communities Support
WP4	Work Package 4: NGI / LGI Infrastructure Services
WP6	Work Package 6: Infrastructure and Applications-oriented Services for User Communities
WQR	Work Queue with Replication

2. EXECUTIVE SUMMARY

The goal of this activity is to develop and support grid services that facilitate the porting and execution of applications in the e-infrastructure. For that, on one side actions have been taken to facilitate the access to the e-infrastructure through the development of Web portals that served as scientific gateways. In addition to the DIRAC portal that was deployed in the first year of the project, in the second year of the project WP6 team has joined efforts with WP3 team to deploy the GISELA Science Gateway⁴, which is currently the main entry point for GISELA users. In particular, work has been focusing on developing a JSAGA Adaptor for the OurGrid middleware. This adaptor allows the seamless execution of application in both gLite and OurGrid resources. On the other side, new infrastructure-oriented services were developed, tested, and deployed in the GISELA infrastructure. Finally, actions to help in the dissemination of the services have also been conducted. The main achievements during the second year of execution of the project are briefly summarised below.

New infrastructure-oriented services were developed, tested, and deployed. They provide support for:

- Efficient execution of data-intensive applications based on the Map-Reduce paradigm over resources exploited opportunistically;
- Seamless execution of applications in hybrid e-Infrastructures.

In addition, a number of scientific papers were published, some of them in high-impact journals and conferences, such as:

- Paulo Ditarso Maciel Jr., Francisco Vilar Brasileiro, Ricardo Araújo Santos, David Candeia, Raquel Vigolvino Lopes, Marcus Carvalho, Renato Miceli, Nazareno Andrade, Miranda Mowbray: Business-driven short-term management of a hybrid IT infrastructure. *J. Parallel Distrib. Comput.* 72(2): 106-119 (2012).
- Juan D. Osorio, Harold Castro, Francisco Vilar Brasileiro: Perspectives of UnaCloud: An Opportunistic Cloud Computing Solution for Facilitating Research. *CCGRID 2012*: 717-718.
- A. Tsaregorodtsev, R. Graciani Diaz, A. Casajus Ramo, F. Stagni, V. Hamar, M. Sapunov. Status of the DIRAC Project, *Computing in High Energy and Nuclear Physics (CHEP), USA 2012*.
- A. Tsaregorodtsev, V. Hamar. MPI support in the DIRAC Pilot Job Workload Management System, *Computing in High Energy and Nuclear Physics (CHEP), USA 2012*.
- Paulo Ditarso Maciel Jr., Francisco Vilar Brasileiro, Raquel Vigolvino Lopes, Marcus Carvalho, Miranda Mowbray: Evaluating the impact of planning long-term contracts on the management of a hybrid IT infrastructure. *Integrated Network Management 2011*: 89-96.

Besides these tangible results, important outcomes were also achieved in strengthening the Latin American network of research in the area of Distributed Computing Infrastructures (DCIs). The development of the GISELA project enabled new collaborations to flourish, resulting in actions to allow mobility of students and researchers among Latin American institutions. This is an indication that the research activity in the area of DCIs will continue to happen in a cooperative way beyond GISELA's lifetime.

⁴ <http://gisela-gw.ct.infn.it/>

3. ACHIEVEMENTS

3.1. SUPPORT AND CUSTOMISATION OF ALREADY AVAILABLE SERVICES

A number of improvements have been performed in the services that belong to the GISELA's portfolio. They are detailed in the next sub-sections.

3.1.1. DIRAC

A lot of improvements have been added to DIRAC's core software, which are now available to the GISELA users. Among these new features one can mention:

- The DIRAC File Catalog, the one used by default;
- Numerous improvements in DIRAC Web Portal such as user proxy upload, support of parametric jobs, pilot submission to BOINC Desktop resources, and improved MPI support.

During the reporting period, the MPI support in the DIRAC Pilot Job Workload Management System was further improved. This new version includes using network file systems in order to simulate a shared file system required for some applications. The CCTOOLS package⁵ was used to emulate a user space distributed file system. The main functioning of the system, however, did not change: MPI jobs should be able to run even at sites where MPI support is not officially provided. A poster discussing this development was presented at the Computing in High Energy and Nuclear Physics Conference (CHEP 2012) at New York City, NY, USA.

Also, two new DIRAC agents were developed, an agent to submit pilot jobs to sites using Grid Engine as a batch system manager, as well as an agent to monitor the availability of GISELA sites.

In order to optimize maintenance of the GISELA DIRAC services, the central workload and data management systems are now provided by the France-Grilles DIRAC installation at CC/IN2P3, Lyon. The GISELA DIRAC Web Portal as well as several services ensuring the system high availability are still running at GISELA sites.

WP6 team contributed to the porting of two applications:

- Bowtie - an ultrafast, memory-efficient short read aligner;
- The MyLims⁶ Portal.

This experience of porting user applications in GISELA has been described in a paper presented at the GISELA-CHAIN Conference (<http://indico.ceta-ciemat.es/conferenceDisplay.py?confId=26>) in Mexico City, Mexico.

3.1.2. OurGrid

As mentioned before, the GISELA Science Gateway (GSG) is a scientific gateway that enables users to maximize the e-Infrastructure capabilities using a standard Web browser. Applications are automatically submitted by the Science Gateway to the GISELA e-Infrastructure via a library based on the SAGA standard. Currently GISELA e-infrastructure has adopted JSAGA as a SAGA implementation allowing it to integrate resources using different grid middleware, for which a JSAGA adaptor is available.

⁵ <http://nd.edu/~ccl/software/>

⁶ LIMS stands for Laboratory Information Management System, and the portal address is <http://co.mylims.org>

In order to allow jobs to be submitted in the GSG to the OurGrid part of the infrastructure, an implementation of JSAGA adaptors for the OurGrid middleware was required. During the second year of the project we have implemented a JSAGA job adaptor that allows the use of the SAGA standard to transparently submit jobs to a peer-to-peer grid infrastructure powered by the OurGrid middleware. This job adaptor enables grid developers and grid users to build interoperable end-user applications, access CPU and storage resources over OurGrid middleware with a single and simple API.

The JSAGA engine uses the appropriate adaptor for connecting to the grid structure, which is capable of interacting with the grid. Initially the job adaptor translates the job into a language understood by the targeted middleware and it submits and monitors the job until it is done or cancelled. Finally the adaptor returns the grid response to the client application (see Figure 1).

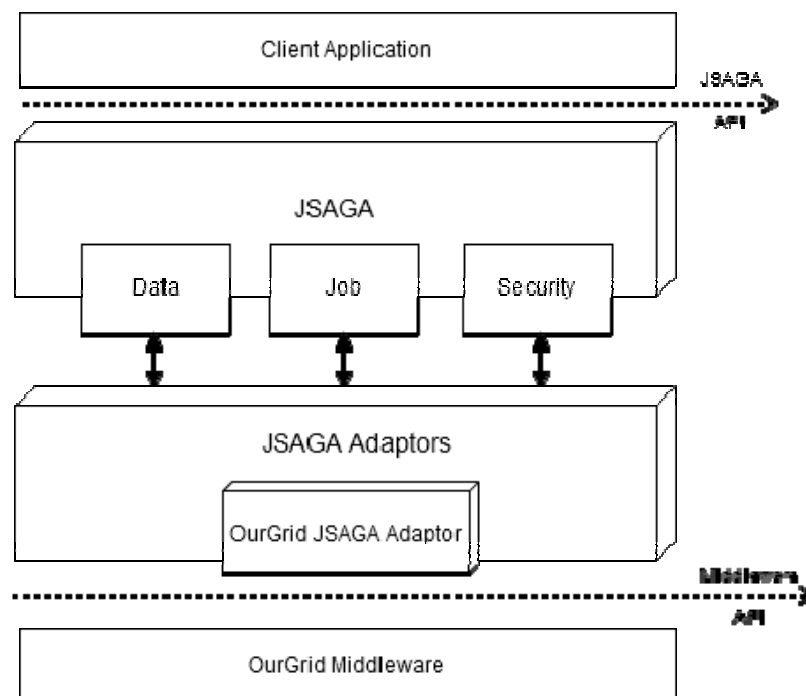


Figure 1: OurGrid JSAGA Adaptor for job submission

A working prototype has been deployed allowing seamless execution of applications both in gLite, as well as in OurGrid resources.

3.2. DEVELOPMENT AND SUPPORT OF NEW SERVICES

3.2.1. Execution of data-intensive applications based on the MapReduce paradigm

Recently several organisations are using the MapReduce programming model to face the increasing demand for the processing of data-intensive applications. MapReduce is an emerging programming model that enables easy development of scalable parallel applications to process vast amounts of data on large clusters of commodity machines. It was originally proposed by Google to handle large-scale

web search applications and has been used by other companies and universities in a wide variety of applications, like data mining, machine learning, and natural language processing.

Currently, the two best-known frameworks for the development and execution of MapReduce applications are the proprietary one developed by Google, and Hadoop, an open source implementation supported by the Apache Software Foundation. Google's implementation uses the Google File System (GFS) as its underlying storage infrastructure, while Hadoop is supported by the Hadoop Distributed File System (HDFS). Both implementations are meant to execute on clusters of commodity processing nodes, with each node incorporating a fair amount of local secondary storage space where the corresponding companion distributed file systems store the files to be processed.

However, with the increasing popularity of MapReduce, there are many efforts trying to support it on other architectures and exploring various ways to make it more suitable for a wider user community and range of applications. In this work, we address the problem of how to easily and efficiently execute MapReduce applications in standard desktops available in conventional local area networks (LANs). Such approach is feasible, given that a notably fraction of desktops shows low CPU load and disk usage. In this direction, in 2010, Erik Frey has described the implementation of a very simple framework for the implementation of MapReduce applications over POSIX-compliant systems, called BashReduce. BashReduce was designed for simplicity of use and imposes very little requirements from the underlying infrastructure. In particular, it does not assume that a distributed file system is in place. Obviously, this characteristic substantially limits the performance that can be attained when running BashReduce applications.

In the first year of the GISELA project we have worked on the implementation of a distributed file system that harnesses the free disk space of desktops connected by a LAN. Similarly to GFS and HDFS, the Beehive File System (or BeeFS, for short) has been designed following a hybrid architecture that employs a centralised server to store metadata and manage file placement and replication and multiple data storage servers that collaboratively store data. This architecture reduces the bottleneck on the central server, and allows the incremental growth of the storage capacity with the addition of new desktops to the system. On the other hand, differently from GFS and HDFS, the primary purpose of BeeFS is to provide a POSIX-compliant distributed file system that is not only more efficient than the prevalent approach based on dedicated servers (eg. NFS), but also cheaper and naturally scalable.

Building on this work we have made a simple modification to BashReduce's original scheduling algorithm that allows it to fully explore the distributed nature of BeeFS to substantially boost its performance. In our experiments we have measured increases in the performance of the applications ranging from 17 to 42 times, when comparing executions that store files in a standard NFS file system (the current approach normally followed by BashReduce users) with those that use BeeFS instead. Given that BeeFS is useful on its own as an alternative to NFS, we envisage that, in time, executing BashReduce applications with the support of BeeFS file systems may turn out to be a common configuration for POSIX-compliant desktops connected by a LAN.

3.2.2. Seamless execution of applications in hybrid e-Infrastructures

The WP6 has conducted two different deployments that aim at supporting the execution of grid applications on hybrid infrastructures. In the first case, the hybrid infrastructure is composed of a service grid and a desktop grid, while in the second one it is a peer-to-peer grid that is integrated to a Cloud computing infrastructure. These developments are described in the following.

3.2.2.1. Elastic integration of a dedicated grid infrastructure with an opportunistic cloud infrastructure

This is a work developed within the Systems and Computer Engineering department of Universidad de los Andes that aims at increasing the infrastructure resources of the university grid (known as site Uniandes) with the opportunistic resources of the desktop computer labs provided by the project UnaCloud. The elastic characteristic allows site Uniandes to change the processing capabilities dynamically and according to the current workload, utilizing only extra resources when needed.

The objective of this integration is to increase the collaboration between research groups at a local, national and international arena. Since site Uniandes collaborates within international grids, partners' projects also leverage from this elastic integration. This integration does not incur in infrastructural extra costs, motivating projects that do not count with budgets to provide dedicated resources for grid initiatives, to support them with desktops. Another objective is to utilize the resources available in the desktops labs in an efficient way.

As mentioned, the opportunistic usage of the desktop computers is only present when site Uniandes is fully loaded and does not have resources left to process the current demand. In this case the elastic integration proposed implements an elastic system that calculates, based on the current demand and certain policies, the number of opportunistic instances needed to lighten the load. In detail, the elastic system monitors the scheduler of site Uniandes looking for queue jobs. When the number trespass a given threshold, the system communicates with UnaCloud via a web service to turn on opportunistic virtual machines (VMs). On the other hand, the system monitors the scheduler looking for the current state of the opportunistic VMs and if the state has been idle for more than five minutes, the resources shut down. Also, due to the volatility of the opportunistic VMs, the elastic system checks for failures and if detected, redirects jobs that were being executed on failed VMs.

Validation took place to analyze the reactiveness and performance of the system, the viability for using the opportunistic VMs (type of resources and volatility) and to evaluate the strategies and policies used. The results showed that the elastic integration provided site Uniandes a good reactiveness to unexpected spikes in the workload, being able to satisfy the demands of the associated virtual organization with a minimum delay.

3.2.2.2. Cloudbursting of a peer-to-peer grid

With the emergence of the CLOUD computing paradigm come new opportunities for the capacity planning of distributed computing infrastructures target for the execution of e-science workloads. One possibility is to run these applications at resources acquired on-demand from cloud providers. However, although very low, there is a cost associated with the usage of cloud resources. Besides that, the amount of resources that can be simultaneously acquired is, in practice, limited. Another possibility is the not new idea of composing hybrid infrastructures in which the huge amount of computational resources shared by the grid infrastructures is used whenever possible and extra capacity is acquired from cloud computing providers. We here investigate how to schedule e-Science activities in a hybrid infrastructure comprising a peer-to-peer grid, so that deadlines are met and costs are reduced.

A very simple solution would be to follow a greedy approach in which as many cloud resources as possible are acquired at the beginning of the application execution, aiming at completing the application as soon as possible. This solution obviously solves the problem of meeting deadlines, however, does it improve the gains associated with cloud resources acquisition? Other solution would be to follow an online scheduling approach in which scheduler makes decision along the application execution. This online solution seems to be more sophisticated than the greedy approach, but is it more efficient?

We studied this problem by using a business-driven approach. A set of heuristics was proposed to predict grid behavior based on its recent behavior and then purchase cloud provider resources. Besides, a simple greedy heuristic was also studied. Our methodology includes simulation experiments in which different application profiles are considered, leading to different utility function shapes and maximum utility values. The availability of the grid resources was modeled using OurGrid traces and tasks were assigned to resources using an algorithm very similar to the WorkQueue with Replication (WQR), with at most three replicas being created (OurGrid's default).

We concluded that cloudbursting applications with very low Utility/Cost relations is hard and requires intelligent approaches. It can also be observed that greedy heuristic results are very inefficient for these applications with low Utility/Cost values. This occurs because this heuristic spends a lot buying instances, but the gain obtained by reducing the finish time is not enough to compensate the investment. In all scenarios studied, the online scheduling approaches presented better results for those applications. Moreover, the greater is the relation Utility/Cost, the greater are the efficiencies achieved by each heuristic, including the greedy one, which sometimes outperforms the others. This occurs because the cost of acquiring resources is compensated by the utility gains resulted from the cloudburst. The wrong decisions made by the schedulers have less impact in the total profit as incomes increase. So, heuristics efficiencies approximate to the results of the optimum solution. In summary, the greedy heuristic, which is very simple and prematurely acquires resources, is, on average, the best approach to cloudburst such kinds of urgent applications if the Utility/Cost relation is high. It was not effective, however, for applications with linear shaped utilities that present low Utility/Cost relation. Compared to an optimal strategy, the other proposed heuristics achieved good results in most of the scenarios. Efficiencies in the range from 50% to 90% were observed.

3.3. DISSEMINATION ACTIVITIES

Dissemination of the services supported by WP6 has been achieved through the communication channels put in place by the WP2 team, i.e. training events, grid schools, newsletter, Web site, etc. In addition, a considerable amount of scientific papers, listed below, have been produced:

- Juan D. Osorio, Harold Castro, Francisco Vilar Brasileiro: Perspectives of UnaCloud: An Opportunistic Cloud Computing Solution for Facilitating Research. CCGRID 2012: 717-718.
- Paulo Ditarso Maciel Jr., Francisco Vilar Brasileiro, Ricardo Araújo Santos, David Candeia, Raquel Vigolvino Lopes, Marcus Carvalho, Renato Miceli, Nazareno Andrade, Miranda Mowbray: Business-driven short-term management of a hybrid IT infrastructure. J. Parallel Distrib. Comput. 72(2): 106-119 (2012).
- Paulo Ditarso Maciel Jr., Francisco Vilar Brasileiro, Raquel Vigolvino Lopes, Marcus Carvalho, Miranda Mowbray: Evaluating the impact of planning long-term contracts on the management of a hybrid IT infrastructure. Integrated Network Management 2011: 89-96.
- A. Tsaregorodtsev, R. Graciani Diaz, A. Casajus Ramo, F. Stagni, V. Hamar, M. Sapunov. Status of the DIRAC Project, Computing in High Energy and Nuclear Physics (CHEP), USA 2012.
- A. Tsaregorodtsev, V. Hamar. MPI support in the DIRAC Pilot Job Workload Management System, Computing in High Energy and Nuclear Physics (CHEP), USA 2012.
- Aquino de Carvalho, Marcus Williams ; Brasileiro, Francisco. Modelagem da Carga de Trabalho de Grades Computacionais Baseada no Comportamento dos Usuários, Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, 2012, Ouro Preto. Anais do XXX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos. Porto Alegre (Brazil) : Sociedade Brasileira de Computação, 2012. v. 1. p. 800-813.

-
- SILVA, Jonhunny W. ; PEREIRA, Thiago Emmanuel ; Brasileiro, Francisco. Computação intensiva em dados com MapReduce em ambientes oportunistas, Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, 2011, Campo Grande. Anais do XXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos. Porto Alegre (Brazil) : Sociedade Brasileira de Computação, 2011. p. 1-14.
 - Tsaregorodtsev, Andrei; Hamar, Vanessa. DIRAC experience with porting user applications in GISELA, Proceedings of the Joint GISELA-CHAIN Conference, R. Barbera et al. (Eds.), COMETA 2012.
 - Patricia Alanis Maldonado, Francisco Vilar Brasileiro, Abmar Grangeiro de Barros. Design and Implementation of a JSAGA Adaptor for the OurGrid Middleware, Proceedings of the Joint GISELA-CHAIN Conference, R. Barbera et al. (Eds.), COMETA 2012.
 - Jonhunny Wesley Silva, Francisco Brasileiro. When BashReduce Met BeeFS, Proceedings of the Joint GISELA-CHAIN Conference, R. Barbera et al. (Eds.), COMETA 2012.
 - Thiago Emmanuel Pereira, Jonhunny Wesley Silva, Alexandro Soares, Francisco Brasileiro. BeeFS: A Cheaper and Naturally Scalable Distributed File System, Proceedings of the Joint GISELA-CHAIN Conference, R. Barbera et al. (Eds.), COMETA 2012.
 - Araujo, Ricardo; Candeia, David; Lopes, Raquel; Brasileiro, Francisco. Cloudbursting a Peer-to-Peer Grid, Proceedings of the Joint GISELA-CHAIN Conference, R. Barbera et al. (Eds.), COMETA 2012.
 - Nicolas Ortiz Gonzalez, Nathalia Garcés Ferrera, German Sotelo Arevalo, David Méndez Lopez, Fabio Hernán Castillo-Coy, Harold Castro. Multiple Services hosted on the opportunistic Infrastructure UnaCloud, Proceedings of the Joint GISELA-CHAIN Conference, R. Barbera et al. (Eds.), COMETA 2012.

4. HUMAN EFFORT

The human resources allocated to WP6 are listed in Table 1.

Table 1: WP6 Human Resources

Name	Role	Partner
Francisco Vilar Brasileiro	WP6 Manager TWP6.1 and TWP6.3 Task Leader	UFCG
Vanessa Hamar	WP6 Deputy Manager TWP6.2 Task Leader	CNRS / CPPM
Lívia Campos	TWP6.2 staff	UFCG
Rodrigo Duarte	TWP6.2 staff	UFCG
Raquel Lopes	TWP6.3 staff	UFCG
Carla Souza	TWP6.3 staff	UFCG
Harold Enrique Castro Barrera	TWP6.3 staff	UNIANDES
Mario Villamizar	TWP6.3 staff	UNIANDES
Germán Sotelo	TWP6.3 staff	UNIANDES
Arthur Oviedo	TWP6.3 staff	UNIANDES
Diego Scardaci	TWP6.2 staff	INFN

5. CONCLUSIONS

Overall, the outcomes from the second year of execution of the activity have been very good. We were able to upgrade in many ways the services already available in the GISELA portfolio and enhance the portfolio with the addition of two new services. Most importantly, the services supported are already fulfilling their aim of facilitating the implementation and deployment of applications.

We were able to cooperate with both WP3 and WP4 teams to understand the users' requirements and cater to their needs. The best example of this integrated effort is the successful deployment of the GISELA Science Gateway, which is thoroughly described in WP3's final deliverable.

According to the DoW, the accomplishments of WP6 are to be measured using the following Quality Metrics: *number of scientific papers published, and percentage of applications using at least one WP6 service*. We expected to have at least 4 scientific publications by M24. This expectation was largely exceeded. The percentage of applications using services from the portfolio was also very large, mainly because most of the effort involved the development of infrastructure-oriented services, which are used by most, if not all applications ported.

It is worth mentioning that the activities developed during the execution of the project fostered cooperation among different Latin American groups, strengthening the ties between these groups, and leading to actions, such as the exchange of students and the submission of joint proposals to regional calls. This is a very positive indication that the cooperative work that has been started in GISELA and in its predecessor EELA and EELA-2 projects will probably continue, in many cases with funds provided by Latin American scientific agencies.