GISELA

NGI / LGI INFRASTRUCTURE SERVICE ASSESSMENT OF THE 2ND YEAR

EU DELIVERABLE: D4.3

Document Full name	**GISELA-D4.3-v1.4**
Date	**31/08/2012**
Activity	**WP4 / NGI-LGI Infrastructure Services**
Lead Partner	**UFRJ**
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 by the Work Package 4 "*NGI-LGI Infrastructure Services*" during the second project-year.



Delivery Slip

	Name	Partner/Activity	Date	Signature
From	WP4	UFRJ / WP4 - NGI-LGI Infrastructure Services		R. Diacovo
Reviewed by	Technical Board			
Approved by	Management Board		31/08/2012	B. Marechal Ph. Gavillet S. Jalife Villalón Luis A. Nuñez R. Barbera R. Diacovo

Document Log

Issue	Date	Comment	Author
0-1	28/06/2011	First draft	B. Marechal
0-2	03/08/2012	Sections 2, 3, 4 and 5 added	R. Diacovo
0-3	06/08/2012	First review and minor corrections	B. Marechal & R. Diacovo
0-4	27/08/2012	Updates	R. Diacovo
1-4	31/08/2012	Final review and approval	B. Marechal

Document Change Record

Issue	Item	Reason for Change

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1. INTRODUCTION

1.1. PURPOSE OF THE DOCUMENT

This deliverable presents the evolution of the GISELA infrastructure during the second project-year. It covers the pledged Success Thresholds and Milestones for the referred period, while also providing extra useful information, such as the infrastructure usage.

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

An executive summary is provided in Section 2. The Work Package achievements are detailed in Section 3, where a thorough analysis of WP4 success thresholds is made (Section 3.1) and other activities are explained (Section 3.2). The table of WP4 human resources follows in Section 0. Finally, some conclusions are shown in Section 5.

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)

1.5. GLOSSARY

DoW	Description of Work
EGI	European Grid Initiative
GOC	Grid Operation Centre (Regional Operations, GISELA context)
GSC	Grid Support Centre (Central Operations, GISELA context)
IGALC	IGALC - Iniciativa de Grid de América Latina – Caribe - Regional Operating Centre (www.igalc.org)
RC	Resource Centre
ROC_LA	Centro de Operaciones Regional Latino Americano
UNAM	Universidad Nacional Autónoma de México
VO	Virtual Organisation
VRC	Virtual Research Community
WP	Work package
WP3	User Communities Support
WP4	NGI / LGI Infrastructure Services

2. EXECUTIVE SUMMARY

During the second project-year, WP4 focus was to keep the e-Infrastructure in top shape, allowing the best possible end-user experience in order to support the push to the disseminate grid e-Infrastructures in the Latin American region. This objective translates into the metrics set in the DoW and the support for other Work Packages, in particular WP3.

The DoW metrics are as follows:

- Number of implemented Grid Support Centres (GSC);
- Average ticket resolution time;
- Average Resource Centre availability;
- Number of cores integrated;
- Number of implemented Grid Operation Centres (GOC).

The first three metrics have been achieved as planned. Although WP4 managed to integrate every CPU core that has been made available in due time, unfortunately some of the GISELA partners did not fulfil their pledge to the Project, either due to unresponsiveness or to technical difficulties at some institutions. Thus, not all of the pledged cores have been integrated. In some cases, although the pledge was fulfilled, the resources have been concentrated in fewer institutions than those foreseen in the original plan. Consequently, the two GOCs already in place and the GSC were able to handle the workload with no issues, and the project opted for not deploying the third GOC, as it would be redundant. However, it is worth mentioning that non-GISELA institutions have provided about 3,000 cores.

This deliverable dwells on other important results such as the usage of the e-Infrastructure by the supported VRCs. With more than 60% of occupancy (non-idle time), it shows a good amount of user interest, while still holding available resources for new users or applications with burst-like demands. A list of the deployed software is also provided, since it is key for the proper functioning of the GISELA Science Gateway. The e-Infrastructure handover status is provided.

3. ACHIEVEMENTS

3.1. SUCCESS THRESHOLDS (FROM THE DOW)

The progress of WP4 is to be measured by five success thresholds. Table 1 enumerates them, as well as the status at the end of the project. More details are given through Sections 3.1.1 to 3.1.5.

Table 1: Success Thresholds for WP4

Quality Metric	Foreseen at year 2	Status
Number of implemented GSC	1	1
Number of implemented GOCs	≥3	2
Number of cores integrated	>2600	2563
Average ticket resolution time	≤2 weeks	2 weeks
Average Resource Centre availability	≥80%	88%

3.1.1. Grid Support Centre (GSC)

The GSC, as discussed in the WP4 Deliverable D4.2[3], is responsible for all central operation activities. Over the second project-year, its main focus was the optimisation of the resource allocation – getting Resource Centres to support as many VRCs via the VOs as possible.

A study has been conducted in order to find a VO allocation schema that maximises the utilisation of the resources, spreading the support between as many VOs as possible. Based on that, it has been proposed to some Resource Centres to extend their support to different VOs. While positive feedback was received from part of them, which have deployed other previously unsupported VOs, most Resource Centres alleged that their current VO schema already kept their computing power completely loaded, and that they would not be able to support more VOs.

At the time, this was a counter-intuitive outcome, as the study indicated room for different VO support. Due to this, allied with the recommendation received during the first Project Review to present accounting data in a way that made it possible to infer the e-Infrastructure occupancy, WP4 re-evaluated the accounting methods for the second year. As a result, data presented for the current reporting period is calculated differently from the previous one (see Section 3.2.1).

The remainder of the tasks attributed to the GSC began the current reporting period in either completed or steady-flow state. One is encouraged to refer to WP4 previous Deliverable for their last updates.

3.1.2. Grid Operation Centre (GOC)

The GOCs are regional instances of the GSC, with a reduced number of attributions and a regional mandate only. The statuses of the Mexican GOC and the Colombian GOC have not changed since the previous reporting period. For more details, one should refer to the WP4 Deliverable D4.2[3].

[3] http://documents.gisela-grid.eu/record/249?ln=en

The foreseen third GOC has not been deployed, however not causing any impact to the e-Infrastructure. Indeed, the GOCs most critical task is to relief the GSC from the load of supporting new Resource Centres during the integration process and the first operation months, when is it most probable that untrained administrators would need external support. However, as detailed in Section 3.1.3, the number of integrated Resource Centres has been lower than what was foreseen in the DoW. This fact, allied to the sparse integration times, has made the deployment of a third GOC to be of little use to the quality of the e-Infrastructure. Thus, WP4 opted for not carrying on with this plan.

3.1.3. Number of cores integrated

Table 2 is displaying the current contribution of the GISELA partners to the common e-Infrastructure in terms of Resources Centres (RC), CPU cores and Terabytes (TB) of data storage. Concerning CPU cores and storage, the situation at the end of the Project is quite satisfactory. However one can observe that the current number of integrated RCs is lower than the committed one. This is mainly due to the fact that several beneficiaries were expecting to integrate resources of former EELA-2 partners that are not GISELA partners or third parties. Furthermore it was preferred, for cost-effective support, to concentrate the computing resources in a few Resource Centres only.

Table 2: Pledged Resources Integration

	DoW			Situation in August / 2012		
Partner	CPU Cores	Storage (TB)	RCs	CPU Cores	Storage (TB)	RCs
CEDIA	100	1	5	150	1	1
CIDETYS	100	1	2	0	0	0
CIEMAT	100	20	2	320	53	2
CNRS/IPGP	40	0	1	388	0	1
CUBAENERGIA	50	0	1	0	0	0
INFN	100	30	1	100	14	1
INNOVA-RED	130	0	5	17	0.6	1
RAAP	110	8	7	0	0	0
UdelaR	100	10	1	40	0	1
UFCG	450	4	7	280	1	2
UFRJ[4]	762	21	2	762	21	2
ULA	120	2	3	54	1	1
UNAM	198	2	8	170	7	1
UNIANDES	200	2	6	176	7	1
UPORTO	100	0	3	106	0.3	2
Total	**2660**	**101**	**54**	**2563**	**105.9**	**16**

All of the resources made available by the partners up to June, 2012 have been integrated. Unfortunately, a non-negligible amount of the pledged computing power has not been made available to the project, due to unresponsiveness from partners or technical issues at some institutions.

The institutions have been requested several times to make the pledged resources available, and the partners responsible for the unresponsive institutions have been notified as well. Despite the multiple attempts, GISELA does not technically own the computing resources pledged to it, meaning that WP4 has no mechanisms to ensure 100% integration.

Although the number of integrated Resource Centres is lower than the originally foreseen, this fact actually brings more advantages than disadvantages, provided that the total amount is the same. From

[4] After the integration, about 61% of the UFRJ Resource Centre has been temporarily shut down due to air-conditioning issues. At the time of this writing, UFRJ had 293 online cores.

the technical point of view, managing a given number of resources scattered over many institutions is much more difficult and expensive than managing the same resources when they are concentrated in fewer institutions. In addition to that, with fewer Resource Centres, the overhead of auxiliary services decreases, leaving more hardware dedicated to actual computation. Of course, concentrating resources is not the spirit of the grid, but one must realize that expenditure minimization is always a priority in any institution.

Despite all of the above, GISELA managed to capture the interest of other institutions, either due to good relationships established in previous projects or via new partnerships. These institutions have greatly helped to mitigate the integration deficit, and have actually boosted the available resources to more than the original pledge. Details are shown in Table 3.

Table 3: "Extra" Resources

Resource Centre	CPU cores	Storage (TB)
CERN-PROD	2000	0
IN2P3-CC	0	11
csTCDie	800	13
UNICAN	200	0.2
IEETA	10	0.5
EELA-UTFSM	44	1
INPE-CPTEC	6	0
INSU01-PARIS	0	1
Total	3060	26.7

3.1.4. Resource Centre availability

Similarly to the previous reporting period, most of the Resource Centres have complied with the availability requirements. The actual numbers can be seen in Table 4 and
Table 5, which show that the 80% target has been met. Some comments on outlier results follow:

- CEFET-RJ has underperformed for three consecutive months and has been suspended. This shows WP4 commitment with the quality of the offered e-Infrastructure, even if it means temporarily having fewer resources available;
- CETA-GRID has been integrated by mid October 2011. Thus, there is no accounting data for this RC prior to that;
- IN2P3-CPPM, although active for a longer time, only began supporting the prod.vo.eu-eela.eu VO (one of the requirements to be part of the GISELA e-Infrastructure) in 2012, which explains the lack of data for the previous year;

- During the second quarter of 2012, UFRJ-IF experienced some problems with its monitoring system, negatively affecting the availabilities of EELA-UNLP, UFCG-LSD and ULA-MERIDA. The presented numbers do not take this fact into account.

Table 4: gLite Resource Centre Availability

Resource Centre	Sep 2011	Oct 2011	Nov 2011	Dec 2011	Jan 2012	Feb 2012	Mar 2012	Apr 2012	May 2012	Jun 2012	Jul 2012	Avg
CEFET-RJ	-	85%	78%	90%	72%	42%	59%	-	-	-	-	**71%**
CETA-GRID	-	-	88%	90%	95%	15%	76%	84%	97%	88%	92%	**83%**
CIEMAT-TIC	99%	97%	90%	84%	85%	90%	77%	99%	91%	87%	93%	**91%**
EELA-UNLP	80%	95%	95%	82%	89%	83%	80%	54%	32%	88%	83%	**78%**
ICN-UNAM	64%	57%	90%	65%	84%	39%	58%	71%	69%	44%	64%	**64%**
IN2P3-CPPM	-	-	-	-	100%	100%	99%	98%	91%	94%	99%	**99%**
INFN-CATANIA	95%	93%	92%	99%	91%	94%	87%	96%	97%	73%	88%	**91%**
UFCG-LSD	82%	62%	81%	92%	100%	84%	96%	46%	98%	79%	60%	**80%**
UFRJ-IF	77%	88%	90%	95%	77%	85%	93%	75%	96%	76%	67%	**84%**
ULA-MERIDA	90%	81%	86%	56%	99%	75%	89%	52%	74%	79%	81%	**78%**
UMinho-CP	92%	94%	96%	100%	58%	93%	79%	88%	89%	70%	82%	**86%**
UNIANDES	91%	62%	61%	99%	98%	87%	75%	92%	86%	45%	59%	**78%**
UPorto	100%	95%	99%	99%	100%	99%	98%	97%	100%	100%	100%	**99%**
OVERALL AVERAGE												**83%**

Table 5: OurGrid Resource Centre Availability

Resource Centre	Sep 2011	Oct 2011	Nov 2011	Dec 2011	Jan 2012	Feb 2012	Mar 2012	Apr 2012	May 2012	Jun 2012	Jul 2012	Avg
LCC2	86%	95%	96%	100%	100%	96%	68%	100%	100%	100%	100%	**95%**
AESA	100%	95%	77%	80%	100%	84%	81%	100%	100%	100%	100%	**96%**
LCC1	96%	95%	96%	100%	100%	83%	74%	100%	100%	100%	100%	**95%**
LSD	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	**100%**
OVERALL AVERAGE												**96%**

3.1.5. Average ticket resolution time

This section presents the resolution time for the tickets created during the reporting period. A total of 142 tickets have been issued for the relevant categories, namely:

- Tickets for IGALC Resource Centres, the ones directly operated by GISELA;
- Tickets for IGALC management from EGI;
- Tickets related to the prod.vo.eu-eela.eu VO:
  - Requests from the VO operations team to Resource Centres;
  - Requests from users to the VO operation teams;
  - Requests from Resource Centre administrators to the VO operations team.

In Figure 1, each vertical bar denotes a ticket, with length proportional to the time between the ticket issue and its closure, in days.

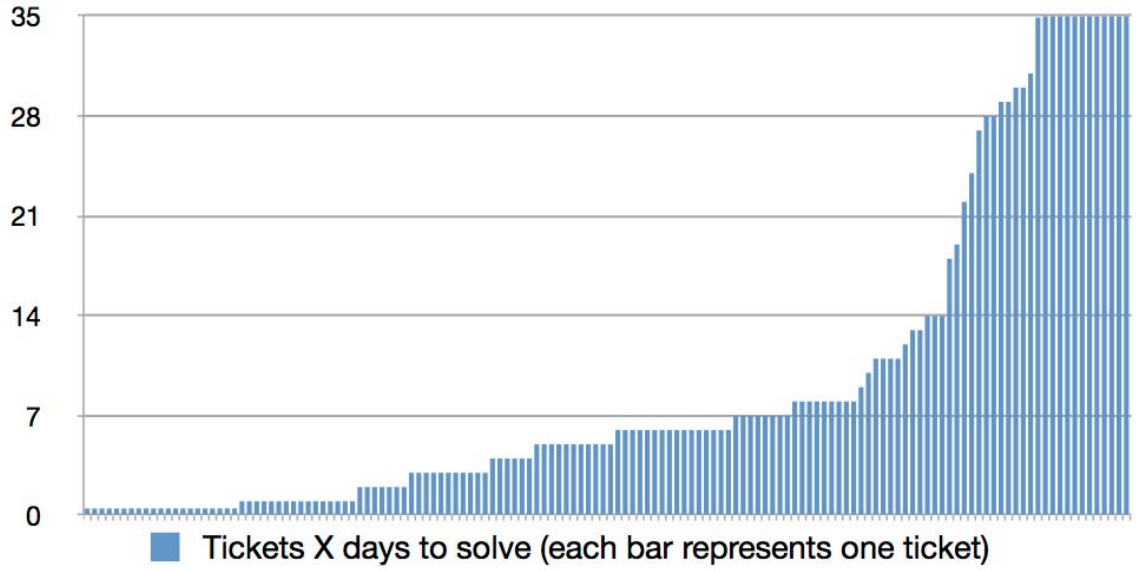

Figure 1: Ticket solving time (days)

The *Average ticket resolution time* over the second project-year, computed by averaging the resolution time of the 142 tickets, equals 14 days – a satisfactory result according to the DoW.

A more detailed analysis reveals that 25 (~18%) of the tickets generated during the reporting period took more than 14 days to be solved. The primary reason for the delay has been traced to Resource Centre administrator slow response in 18 of them. The remainder 7 were related to problems that were either very difficult or rare, which understandably take longer to solve than the bulk of the ticket flow. Resource Centres with underwhelming response time have been instructed on how to increase their efficiency, and most have improved to reaction times within expectations.

If the 25-outlier tickets are not considered, the *Average Ticket resolution time* drops to approximately 4 days.

3.2. OTHER ACHIEVEMENTS

3.2.1. Infrastructure Usage

During the First Project Review, and over the course of the VO allocation task, it became clear that reporting Normalised CPU Hours and not taking into account the fluctuations on the e-Infrastructure size due to downtimes has given the false impression that the resources were underused. In addition to that, the VO-allocation task described in Section 3.1.1 also hinted that the methods could be improved. For this reporting period, the following has been adopted:

- Data are presented in terms of wall-clock time, resulting in fewer distortions due to time spent by the jobs on I/O and file transfers;
- The dynamic size of the e-Infrastructure is taken into account, meaning that, when a given RC has part of their resources offline, only the online part counts toward the available resources;
- The Resource Centre Availability, as reported in Section 0, is also considered when calculating the amount of available resources.

Figure 2 depicts the contribution of the GISELA resources to the supported VRC VOs. The calculation method defined in the GISELA Deliverable D4.1[5] still applies, and the figure reflects the contribution to all supported VOs from consortium resources added to the contribution to the prod.vo.eu-eela.eu VO from any RC that supports it.

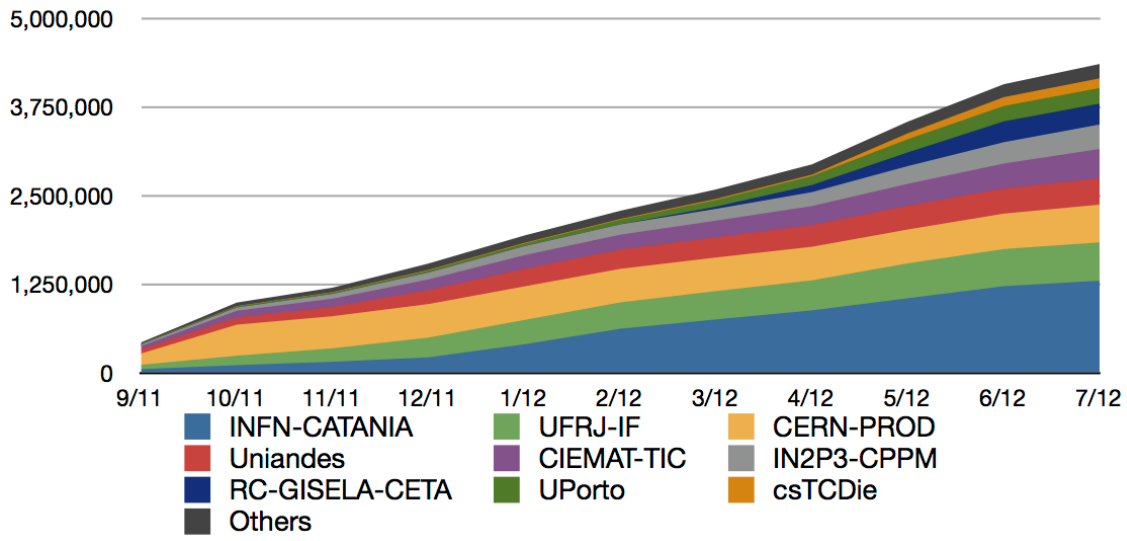

Figure 2: Elapsed wall time hours

In order to provide a better insight concerning the activity level in the e-Infrastructure, Figure 3 shows the theoretical available time compared to the actual used time. One should note that, in this figure, RCs that are not part of the GISELA consortium are not accounted for. The reason is the impossibility to establish the fraction of non-consortium resources available to GISELA in these RCs, since there is no pledge. The average occupancy over the current reporting period is close to 57%, with a noticeable

[5] http://documents.gisela-grid.eu/record/84?ln=en

increase towards the end of the period. This indicates that, while the current resources are enough to satisfy the demand, it is very likely that more computing power will be necessary on the mid-term, assuming the demand continues to behave in the same way.

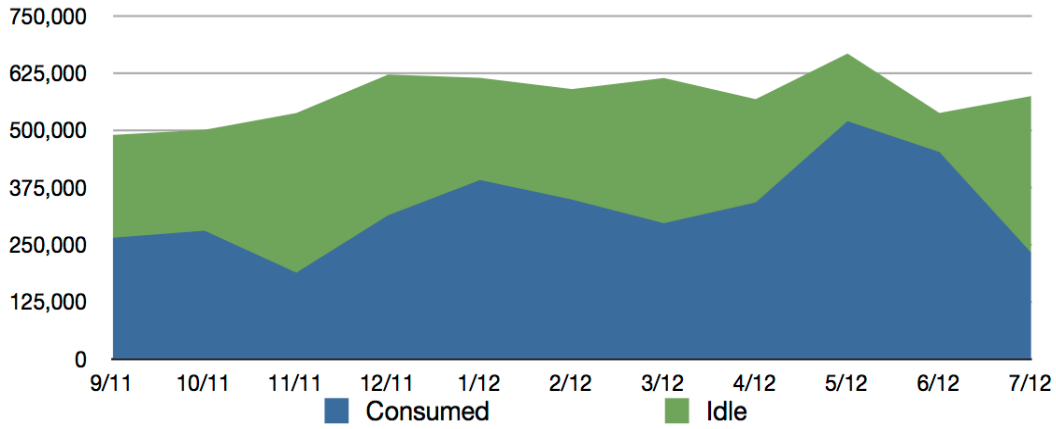

Figure 3: Infrastructure Occupancy (wall time hours)

3.2.2. Number of TeraBytes of storage Integrated

Although the metrics established in the DoW measures the resources only as far as the computing power goes, the ability to store results and applications on the grid is very important and appreciated by the VRCs. Table 2 (located in Section 3.1.3) depicts that, the 101 TB pledged by the participating institutions have been integrated with a surplus, totalling 106TB. If the 26.7 TB made available by the non-GISELA Resource Centres are also considered (shown in Table 3, Section 3.1.3), the storage amount reaches more than 132TB, exceeding the pledge by almost one third.

3.2.3. Base support for Applications

With the advent of the GISELA Science Gateway, the responsibility of installing some of the base packages and libraries has been moved from the users to the e-Infrastructure operations team. Requests are received from WP3, and WP4 then proceeds with the deployment in as many sites as possible, considering hardware and available VO-area size. In some cases, it is necessary for the RC administrator to deploy software in every node of his site, which lengthens the process due to the amount of work involved. Table 6 enumerates the software installed in the prod.vo.eu-eela.eu VO during the current reporting period. For further details concerning the applications, one should refer to the GISELA WP3 Deliverable D3.3[6].

[6] http://documents.gisela-grid.eu/record/267?ln=en

Table 6: Base support deployment

Package/Library	Description
Astra 2.0	Sound Reconstruction Application
Gate 6.0.0	Medical Imaging tool
Gromacs 4.5.3	Molecular Dynamics package
Java 1.6	Programming platform
Octave 3.2.4	Numerical Calculation package
Povray 3.6	3-D Graphics tool
R 2.13.1	Statistical computing package
R 2.15.1	Statistical computing package

3.2.4. Handover status

Following the long-term sustainability plans, GISELA is handing over the operation of both the Resource Centres under IGALC and of the prod.vo.eu-eela.eu VO to ROC_LA and UNAM, respectively. Although this handover will happen after the end of the project, this achievement is well inline with the proposed plan for sustainability and deserves to be mentioned here.

IGALC Resource Centres full migration to ROC_LA is scheduled to September 10th, 2012. ROC_LA has also been running e-Infrastructure operations for a long time, and thus already have all of the necessary ROC services deployed. The migration involves only minor reconfigurations by the RC administrators and central RC catalogue staff, meaning it should be transparent to the users.

The VO services migration, at the time of this writing, is ongoing. The VO services are being deployed at UNAM, and are expected to become operational by the end of September 2012.

For a list of ROC and VO services, one should refer to D4.2[7].

[7] http://documents.gisela-grid.eu/record/249?ln=en

4. HUMAN EFFORT

The current human resources allocated to WP4 are listed in Table 7.

Table 7: WP4 Human Resources

Name	Role	Partner
Jhoanna Serpa	CEDIA site administrator	CEDIA
Querube Urriola	CIDETYS site administrator	CIDETYS
Antonio Rubio	CIEMAT-TIC site administrator	CIEMAT
David Weissenbach	INSU-PARIS site administrator	CNRS
Andres Barbieri	CA Operator	INNOVA-T
Fernando Lopez	EELA-UNLP site administrator	INNOVA-T
Joaquín Bogado	EELA-UNLP site administrator	INNOVA-T
Lia Molinari	Partner Representative	INNOVA-T
Matias Banchoff	CA Operator	INNOVA-T
Yuri Ivanov	EELA-UTFSM site administrator	REUNA
Sergio Nesmachnow	Partner Representative	UdelaR
Sérgio Afonso	UPorto site administrator	U. Porto
Antonio Rodrigues	LSD-UFCG site administrator	UFCG
Rúbia Ramos	LSD-UFCG site administrator	UFCG
Bruno Azevedo	CEFET-RJ site administrator	UFRJ
Gabriel Almeida de Oliveira	Core services administrator	UFRJ
Gabriel dos Passos Gomes	UFRJ-IF site administrator	UFRJ
Mariana Sampaio	UFRJ GSC staff	UFRJ
Maurício Lamarão Mota	UFRJ GSC staff	UFRJ
Ramon Diacovo	WP4 Manager TWP4.1 and TWP4.2 Task Leader	UFRJ
Gilberto Diaz	ULA-MERIDA site administrator	ULA
Girley Margiotta	ULA-MERIDA site administrator	ULA
Luciano Diaz	ICN-UNAM site administrator	UNAM
Jesus Cruz Guzman	CLARA Transition Team	UNAM
Rene Luna-Garcia	IPN-GRID site administrator	UNAM
Andres Holguin	UNIANDES site administrator	UNIANDES
Artur Oviedo	Uniandes site administrator	UNIANDES
Daniel Burbano	Uniandes site administrator	UNIANDES
Natalia Garces	Uniandes site administrator	UNIANDES
Harold Enrique Castro Barrera	WP4 Deputy Manager TWP4.3 Task Leader	UNIANDES

5. CONCLUSIONS

The overall self-assessment is that WP4 has successfully played its part in GISELA, ensuring that a production quality e-Infrastructure has been made available to the users at all times. Here follow some highlights:

- The GSC and GOCs were able to handle every operational issue encountered efficiently, and sufficient knowledge has been spread to ensure that the GOCs in Colombia and Mexico are able to sustainably run the infrastructure after the end of the project;
- The prod.vo.eu-eela.eu VO grew in importance with the advent of the GISELA Science Gateway, hosting a fair amount of software for the VRCs;
- The hardware resources provided both by partners and third-parties have helped disseminating grid technologies and empowered researchers with computational power they would not have access otherwise.

WP4 is certain that GISELA has played a major role in ensuring the future of grid e-Infrastructures in Latin America and that CLARA has all the tools it needs to provide an ever-improving service in the long term.

Last but not least, the handover of operations of both the Resource Centres under IGALC and the prod.vo.eu-eela.eu VO to ROC_LA and UNAM, respectively, is almost completed.