



# GISELA

## REVISION OF SUPPORTED VRCs AND EXECUTION PLAN FOR THE 2<sup>ND</sup> YEAR

### EU DELIVERABLE: D3.3

---

|                          |   |
|--------------------------|---|
| Document Full name       | <b>GISELA-D3.3-v1.13</b>  |
| Date                     | <b>07/10/2011</b>   |
| Activity                 | <b>WP3 / User Communities Support</b>   |
| Lead Partner             | <b>INFN</b>   |
| Document status          | <b>APPROVED</b>   |
| Classification Attribute | <b>PU (Public)</b>  |
| Document link            | <a href="http://documents.gisela-grid.eu">http://documents.gisela-grid.eu</a> |

---

Abstract: This document reports the revision of supported VRCs and the execution plan for the second year of the project updated after the first project review.



#### 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 7<sup>th</sup> Framework Programme. **GISELA** began on 1<sup>st</sup> September 2010 and will run for 2 years.

For more information on GISELA, its partners and contributors please see [www.gisela-grid.eu](http://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](http://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   |
|--------------------|------------------|---------------------------------------|------------|---|
| <b>From</b>        | WP3              | INFN / WP3 - User Communities Support |            |   |
| <b>Reviewed by</b> | Technical Board  |                                       |            |   |
| <b>Approved by</b> | Management Board |                                       | 07/10/2011 | B. Marechal<br>Ph. Gavillet<br>S. Jalife Villalón<br>L. A. Trejo Rodriguez<br>R. Barbera<br>R. Ramos Pollán |

### Document Log

| Issue | Date       | Comment  | Author                            |
|-------|------------|--|-----------------------------------|
| 0-1   | 15/09/2011 | First draft  | B. Marechal                       |
| 0-2   | 23/09/2011 | Second draft   | D. Scardaci                       |
| 0-3   | 25/09/2011 | Draft of Section 3 included  | R. Mayo                           |
| 0-4   | 02/10/2011 | Re-shuffling of the document   | Ph. Gavillet & B. Marechal        |
| 0-5   | 03/10/2011 | Draft of Section 5.2 Training Plan & Training Events   | G. Díaz-Herrero                   |
| 0-6   | 04/10/2011 | Comments on strategy concerning Cloud Computing (section 4); Modified Section 5.2 considering CLARA input and view on Training. Section 3.2. | G. Díaz-Herrero                   |
| 0-7   | 04/10/2011 | Sections 4, 5.1 and 6 included   | D. Scardaci                       |
| 0-8   | 05/10/2011 | Executive Summary, Section 3.1 and Conclusion included   | D. Scardaci                       |
| 0-9   | 05/10/2011 | Minor comments and suggestions included  | R. Mayo                           |
| 0-10  | 06/10/2011 | Some revisions done throughout the document  | R. Barbera                        |
| 0-11  | 06/10/2011 | Minor revisions  | D. Scardaci                       |
| 0-12  | 06/10/2011 | Some revisions done throughout the document  | R. Diacovo, L. Nunez, D. Scardaci |



**REVISION OF SUPPORTED VRCS AND  
EXECUTION PLAN FOR THE 2ND YEAR**

*Document Full Name*  
GISELA-D3.3-v1.13

*Date:* 07/10/2011

---

|      |            |  |                                |
|------|------------|--|--------------------------------|
| 0-13 | 07/10/2011 | Some more comments on Grid versus Clouds | F. Brasileiro, G. Díaz-Herrero |
| 1-13 | 07/10/2011 | Final Review and Approval                | Ph. Gavillet & B. Marechal     |

**Document Change Record**

| <b>Issue</b> | <b>Item</b> | <b>Reason for Change</b> |
|--------------|-------------|--------------------------|
|              |             |                          |
|              |             |                          |

## TABLE OF CONTENTS

|  |           |
|--|-----------|
| <b>1. INTRODUCTION .....</b>   | <b>6</b>  |
| 1.1. PURPOSE OF THE DOCUMENT .....   | 6         |
| 1.2. DOCUMENT ORGANISATION .....   | 6         |
| 1.3. APPLICATION AREA .....  | 6         |
| 1.4. DOCUMENT AMENDMENT PROCEDURE .....  | 6         |
| 1.5. GLOSSARY .....  | 7         |
| <b>2. EXECUTIVE SUMMARY .....</b>  | <b>9</b>  |
| <b>3. STATUS OF VRCS AFTER 1<sup>ST</sup> YEAR .....</b>                                     | <b>10</b> |
| 3.1. ACHIEVEMENTS .....  | 10        |
| 3.1.1. <i>Virtual Research communities supported</i> .....                                   | 10        |
| 3.1.2. <i>Overall list of applications and VRCS</i> .....                                    | 11        |
| 3.1.3. <i>Supported groups in Latin America</i> .....  | 12        |
| 3.1.4. <i>Collaboration with EGI-InSPIRE</i> .....   | 12        |
| 3.1.5. <i>Collaboration with CHAIN and Regional Initiatives</i> .....                        | 13        |
| 3.1.6. <i>Training</i> .....   | 13        |
| 3.2. OBSERVATIONS AND LESSONS LEARNED .....  | 13        |
| <b>4. STRATEGY TO INTEGRATE NEW VRCS AND TO FULLY EXPLOIT THE E-<br/>INFRASTRUCTURE.....</b> | <b>16</b> |
| <b>5. PLAN FOR THE SECOND YEAR.....</b>  | <b>19</b> |
| 5.1. VRC-DRIVEN SCIENCE GATEWAY DEPLOYMENT AND EVOLUTION .....                               | 19        |
| 5.2. TRAINING PROGRAM & TRAINING EVENTS .....  | 19        |
| 5.2.1. <i>Tentative training agenda</i> .....  | 21        |
| 5.2.2. <i>Training and documentation supporting tools</i> .....                              | 23        |
| 5.3. FOCUSED DISSEMINATION.....  | 23        |
| 5.4. LIAISON WITH EGI-RELATED AND OTHER EXTERNAL ENTITIES .....                              | 23        |
| <b>6. OVERALL IMPACT &amp; PROGRESS EVALUATION .....</b>                                     | <b>25</b> |
| <b>7. CONCLUSIONS.....</b>   | <b>26</b> |

## TABLE OF TABLES

|   |    |
|---|----|
| Table 1: GISELA Computing Elements supporting HEP VOs .....                           | 11 |
| Table 2: VRC Applications .....   | 11 |
| Table 3: List of training events held in the first year .....                         | 13 |
| Table 4: Description of training actions to be performed during the second year ..... | 20 |
| Table 5: Training events tentative schedule and Characteristics.....                  | 22 |
| Table 6: WP3 (upgraded) quality metrics .....   | 25 |

## TABLE OF FIGURES

|  |    |
|--|----|
| Figure 1: Number of application groups from Latin America..... | 12 |
|--|----|

## **1. INTRODUCTION**

### **1.1. PURPOSE OF THE DOCUMENT**

This document aims at describing the WP3 plans for the second year of the Project, properly revised after the first periodic review.

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

### **1.2. DOCUMENT ORGANISATION**

Section 2 gives an Executive Summary of the document. Section 3 reminds the achievements of the work carried out by WP3 during the first year. Section 4 describes the new strategy to increase the e-Infrastructure utilisation. Section 5 shows the plan that will be followed to accomplish the new strategy line. Section 6 lists the new metrics introduced to better measure the WP3 progress and the new plan status. Finally, conclusions are presented in Section 7.

### **1.3. APPLICATION AREA**

The target audience for this document is:

- The members of the Project;
- Any VRC already, or eager to become, user of the GISELA e-Infrastructure;
- The European Commission (Project Officer, Reviewers, etc.);
- 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](mailto:hlp-gisela@hlpdeveloppement.fr)).

---

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

<sup>2</sup> Consortium Agreement (CoA) available upon request to the GISELA Project Office ([hlp-gisela@hlpdeveloppement.fr](mailto:hlp-gisela@hlpdeveloppement.fr))

## 1.5. GLOSSARY

|         |   |
|---------|---|
| AA      | Authentication & Authorisation                                |
| CE      | Computing Element (Core in case of multicore node)            |
| CLARA   | Cooperación Latino Americana de Redes Avanzadas               |
| CHAIN   | Co-ordination and Harmonisation of Advanced e-Infrastructures |
| DCI     | Distributed Computer Infrastructure                           |
| DoW     | Description of Work   |
| EGI     | European Grid Initiative                                      |
| EPIKH   | Exchange Programme to advance e-Infrastructure Know-How       |
| EU      | European Union  |
| gLite   | Lightweight middleware for Grid Computing                     |
| IaaS    | e-Infrastructure as a Service                                 |
| LA      | Latin America   |
| LGI     | Latin American Grid Initiative                                |
| MPI     | Message Passing Interface                                     |
| NGI     | National Grid Initiative                                      |
| NREN    | National Research & Education Network                         |
| OurGrid | <a href="http://www.ourgrid.org/">http://www.ourgrid.org/</a> |
| PDCA    | Plan, Do, Check and Act cycle                                 |
| PC      | Project Coordinator   |
| PM      | Person Month  |
| RC      | Resource Centre also sometimes called Site                    |
| SaaS    | Software as a Service   |
| SAGA    | Simple API for Grid Applications                              |
| SG      | Science Gateway   |
| TP      | Training Program  |



**REVISION OF SUPPORTED VRCS AND  
EXECUTION PLAN FOR THE 2ND YEAR**

*Document Full Name*  
GISELA-D3.3-v1.13

*Date:* 07/10/2011

---

|      |                              |
|------|------------------------------|
| VO   | Virtual Organisation         |
| VRC  | Virtual Research Community   |
| VRE  | Virtual Research Environment |
| WP3M | Work Package 3 Management    |



## 2. EXECUTIVE SUMMARY

The scope of this Deliverable is to update the support of VRCs provided by WP3 and to describe the work plan of the 2<sup>nd</sup> year of the project. It shortly analyses the WP3 achievements over the 1<sup>st</sup> year, in order to well identify what has already be granted as of the VRC support and, moreover, what is not fully satisfactory and has to be recovered over the 2<sup>nd</sup> year and the means for it.

Although the WP3 achievements fulfilled the first year quality metrics defined in the Description of Work, the reviewers of the first Periodic Review did not consider the results obtained fully satisfactory, mainly due to the relative lack of Research communities with sufficient regional impact in Latin America to guarantee the sustainability of the e-infrastructure. A stronger impact on the Latin American scientific world has to be considered essential to reach the long-term sustainability target. For this reason the Project is proposing a new strategy to increase its impact on the Latin America Scientific communities.

Nowadays, the problem of increasing the number of e-Infrastructure users is common, not only in Latin America, but also in Europe and other regions of the world. The reasons considered as the biggest hindrances towards a larger uptake of the e-Infrastructures are the complexity of the middleware, especially its security infrastructure based on digital certificates, the lack of intuitive high-level user interfaces and the very little adoption of standards. Authentication & Authorisation (AA) models based on digital certificates and quite complex interfaces to interact with the e-Infrastructures are, often, insurmountable barriers for scientists that choose to continue to work with their old tools without e-Infrastructure capabilities.

As an attempt to overcome the AA and interface obstacles, GISELA is developing a Science Gateway (SG)<sup>3</sup> providing simple authentication mechanisms and intuitive user interfaces to execute applications on the e-Infrastructure.

The GISELA Science Gateway will be VRC-driven and will not be only an information container. It will allow users to submit applications on the e-Infrastructure via a simple Web interface. The SG will contain one section for each VRC supported, with specific information on the applications, such as application characteristics, case studies and success stories. These web-site sections will become specialised Science Gateways for the various supported VRCs, playing the role of entry points to access and use the e-Infrastructure.

Applications integrated in the SG will be properly chosen in cooperation with CLARA and NRENs, with focus on emerging or large communities and high impact applications to enlarge the infrastructure utilisation with the aim of the sustainability in mind.

To exploit this new strategic line, the training activity has to be properly adapted to increase the Science Gateway “power” and to spread the adoption of the new paradigm. The location of the training events will be suggested by CLARA and NRENs.

Finally, new metrics, as “Number of Active Users in the VRC-driven Science Gateway”, “Number of Applications integrated in VRC-driven Science Gateways”, etc., have been introduced to better measure the WP3 progress and the execution of the new plan.

---

<sup>3</sup> “A Science Gateway is a community-developed set of tools, applications, and data that is integrated via a portal or a suite of applications, usually in a graphical user interface, that is further customised to meet the needs of a specific community.”

### 3. STATUS OF VRCS AFTER 1<sup>ST</sup> YEAR

In this section we give a short summary of the GISELA achievements over the 1<sup>st</sup> year, in order to well identify what has already be granted as of the VRC support and, moreover, what is not fully satisfactory and has to be recovered over the 2<sup>nd</sup> year and the means for it.

#### 3.1. ACHIEVEMENTS

##### 3.1.1. Virtual Research communities supported

The GISELA infrastructure currently supports 3 VRCS. The support activity for each VRC is summarised as follows:

- **Life Sciences:**
  - **WeNMR** (collaboration started at the EGI Technical Forum 2010):
    - [www.wenmr.eu](http://www.wenmr.eu)
    - Access is provided to the WeNMR application portal;
    - Several LA sites are opened to “enmr.eu” VO;
  - **HealthGrid** (agreement reached during the EGI User Forum 2011):
    - [www.healthgrid.org](http://www.healthgrid.org)
    - Support of the “biomed” VO in LA;
    - Exchange of software and experiences;
    - Users could be anywhere in LA or EU;
- **Earth Sciences:**
  - Special interest detected in WRF model;
  - Leveraging WRF4G from EELA-2 to new groups in Brazil and Cuba:
    - <http://www.meteo.unican.es/software/wrf4g>
    - Simulations in Amazon region;
    - Applications to detect pollution affecting air quality in atmosphere;
  - World-wide interest (EGI Technical Forum 2011);
- **High Energy Physics (HEP):**
  - 4 experiments @LHC accelerator supported: ALICE, ATLAS, CMS and LHCb;
  - Pierre Auger Observatory  
(see [http://applications.gisela-grid.eu/application\\_details.php?l=20&ID=35](http://applications.gisela-grid.eu/application_details.php?l=20&ID=35));
  - Table 1 lists the GISELA Computing Elements supporting HEP VOs;

**Table 1: GISELA Computing Elements supporting HEP VOs.**

| HEP Virtual Organisation | GISELA CEs   |
|--------------------------|--|
| <b>ALICE</b>             | ce01-tic.ciemat.es<br>cream-ce.ct.infn.it<br>grid012.ct.infn.it<br>tochtli.nucleares.unam.mx<br>tochtli64.nucleares.unam.mx                            |
| <b>ATLAS</b>             | ce.labmc.inf.utfsm.cl  |
| <b>Pierre Auger</b>      | cale.uniandes.edu.co<br>ce01.eela.if.ufrj.br<br>ce02.eela.if.ufrj.br<br>grid012.ct.infn.it<br>tochtli.nucleares.unam.mx<br>tochtli64.nucleares.unam.mx |
| <b>CMS</b>               | cale.uniandes.edu.co<br>grid012.ct.infn.it   |
| <b>LHCb</b>              | ce01-tic.ciemat.es<br>ce01.eela.if.ufrj.br<br>grid012.ct.infn.it   |

### 3.1.2. Overall list of applications and VRCS

Table 2 shows the applications associated to each supported VRC.

**Table 2: VRC Applications**

| VRC                   | Legacy applications  | Virtual Organisations             |
|-----------------------|--|-----------------------------------|
| <b>Life Sciences</b>  | Haddock, Cyana, Xplor-NIH, CS-ROSETTA, MD, Amber, MDD, Bowtie, etc.                              | biomed, enmr.eu                   |
| <b>Earth Sciences</b> | CAM and WRF  | prod.vo.eu-eela.eu                |
| <b>HEP</b>            | Applications managed by each VO  | ALICE, ATLAS, AUGER, CMS and LHCb |
| <b>Catch-all</b>      | CLARA Communities, Mexican Industry Community, Applications inherited from EELA and EELA-2, etc. | prod.vo.eu-eela.eu                |

Moreover, 70 Grid applications inherited from EELA & EELA-2 are available on the GISELA application database (see [http://applications.gisela-grid.eu/app\\_list.php?l=20](http://applications.gisela-grid.eu/app_list.php?l=20)), of which 34 are in a production stage. However, this does not mean that all ported applications are currently running on the infrastructure on a production basis. That list should be interpreted as “available applications”, which can be requested by the GISELA users whenever needed.

### 3.1.3. Supported groups in Latin America

Figure 1 gives the evolution of supported groups in Latin America. The figure also shows the two levels of support: the Central and the Local ones. While the former regards the support offered by the WP3 task leaders, the latter refers to the support of local experts inside the institution that develops the application. Such a distinction of support levels is foreseen in the Latin American Grid Initiative (LGI) sustainability model<sup>4</sup>: Local support being at RC or NGI/GOC levels; Central support being at the LGI/GSC level.

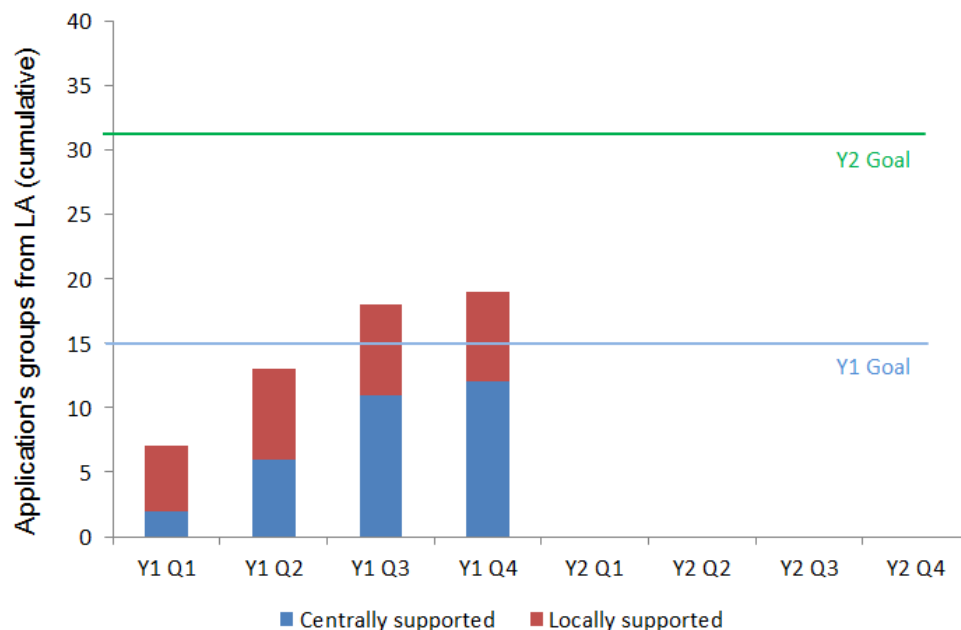


Figure 1: Number of application groups from Latin America.

### 3.1.4. Collaboration with EGI-InSPIRE

During the first year of the project a Memorandum of Understanding<sup>5</sup> between GISELA and EGI-InSPIRE has been signed. Among others goals, it foresees that:

- VRC members in Europe could exploit GISELA resources;
- VRC members in LA could exploit EGI resources;
- EGI and GISELA will coordinate VRC support in every EGI forum;
- GISELA applications in production stage are published in <http://appdb.egi.eu/>.

<sup>4</sup> <http://documents.eu-eela.eu/record/1333/files/>

<sup>5</sup> <http://documents.gisela-grid.eu/record/171/files/GISELA-MOU05-EGI-InSPIRE.pdf>

### 3.1.5. Collaboration with CHAIN and Regional Initiatives

Two Workshops on “Resource Infrastructure Providers meet VRCs” were jointly organised with the CHAIN project with the goal of:

- Precisely identifying the requests of VRCs;
- Adapting the e-Infrastructure support accordingly;
- Harmonizing VRC support across Regional initiatives under the auspices of EGI;

The main outcome of this activity has been the inventory of technical specification per VRC/Application and the agreement to establish direct one-to-one communication.

### 3.1.6. Training

The training program for the first year was chosen amongst GISELA Partners to include a Site Administrator and an Application porting sessions to expand and consolidate the expertise inherited from EELA-2. Two workshops were held to disseminate the interest of using e-Infrastructures towards new communities.

The two training events were jointly organized with EPIKH and CHAIN projects. The two workshops were locally organised by UNAM at Hermosillo and Guaymas (Sonora, Mexico). Table 3 lists the training events held so far.

**Table 3: List of training events held in the first year.**

| <b>Event</b>   | <b>Date</b>                                     | <b>Venue</b>            |
|--|---|-------------------------|
| 1 <sup>st</sup> Joint GISELA/EPIKH School for Application Porting and Grid Site Administrators       | November 2010                                   | Mexico City, Mexico     |
| 2 <sup>nd</sup> Joint CHAIN/GISELA/EPIKH School for Application Porting and Grid Site Administrators | November 2010                                   | Valparaiso, Chile       |
| Workshop: <i>Uso de Infraestructuras Grid Computacional em Aplicaciones de Ciencia y Tecnologia</i>  | 8 <sup>th</sup> - 10 <sup>th</sup> , June 2011  | Hermosillo, Mexico      |
| Workshop: <i>Uso de Infraestructuras Grid Computacional em Aplicaciones de Ciencia y Tecnologia</i>  | 13 <sup>th</sup> - 17 <sup>th</sup> , June 2011 | Guaymas, Sonora, Mexico |

## 3.2. OBSERVATIONS AND LESSONS LEARNED

Beyond the results achieved, a few observations can be drawn, while thinking of the second year of the project, in particular:

- Metrics thresholds were met but they are not enough to reach the objectives of sustainability and users support by the end of the project;
- More relevant measures need to be used in order to monitor all planed actions, therefore supporting progress towards the overall project goals;

- The impact of e-Infrastructures is not yet at the expected level:
  - Not all RCs can support those VO's that generate highest demands of resources:
    - Some cannot fulfil specific requirements (disk size, network bandwidth, 64-bit architectures, etc);
    - Some just need to upgrade their configurations, or install additional software components;
  - Smaller research communities do not make an intensive nor constant use of resources:
    - Some may actually prefer to deal with local resources, or even envisage hiring infrastructure as a service to commercial providers, just to avoid dealing with underlying complexities of Distributed Computing Infrastructure (DCI) in general, and gLite middleware in particular.
- Identified barriers and problems reclaimed by researchers, are hindering a larger usage of grid resource, in particular:
  - The complex and extremely long Authentication and Authorization process, based on X.509 certificates;
  - The difficulty to deal with gLite Middleware issues;
  - The fact that some middleware components require special network connection that can be blocked by security firewalls at many research institutions (e.g., Universities);
  - The limitation of the gLite Middleware not providing tools to measure actual usage of e-Infrastructure in terms of active users or running applications.
- The lack of support and commitment by policy makers or funding agencies at the national level:
  - Only two NRENs in LA (Brazil and Venezuela) are financially supported by public funds;
  - Difficulty to make them fully understand GISELA Grid as a tool for e-Science and a cost-effective instrument against digital divide;
  - Not enough political interest in those LA VRCs that generate the highest demand of e-Infrastructure resources, such as the VRCs related to LHC experiments or WeNMR.
- The Research Institutions in LA currently pay NRENs only for the videoconferencing service. CLARA needs to show NRENs the usefulness and potentialities of the e-Infrastructure by offering high added-value services to researchers, apart from just offering plain raw access to Grid Resource Centres.
- The sustainability relies also on the CLARA TT acquiring knowledge and expertise to become capable of providing user support services beyond the project term:
  - CLARA TT should identify human resources at NRENs available by the end of project to punctually provide on-demand training/support services;
  - CLARA TT has to gain control on the management of some core services and assets, for instance on the new VRC-driven Web site targeted to users/researchers not belonging to the GISELA Consortium;
- CLARA has to identify the most relevant research communities, especially in countries where more resources are/can be committed to e-Infrastructures and in smaller countries of Central

America where policy makers can be more easily approached. Five main domains have been determined so far, although CLARA will keep on prospecting new research groups, VRCS and e-Science trending topics all during the second year of the project:

- Industry;
- Bioinformatics;
- Medical image processing;
- Climate change and weather forecast simulations;
- Epidemiology and Natural disaster management in Central America.

#### 4. STRATEGY TO INTEGRATE NEW VRCS AND TO FULLY EXPLOIT THE E-INFRASTRUCTURE

Although WP3 achievements fulfilled the first year quality metrics defined in the Description of Work, we cannot consider the results obtained fully satisfactory, mainly due to the relative lack of Research communities with sufficient regional impact in Latin America to guarantee the sustainability of the e-infrastructure. A stronger impact on the Latin American scientific world has to be considered essential to reach the long-term sustainability target.

Due to the multi-country membership nature of CLARA and, the multi-diversity of technical skills of its members, the situation become more complex when we ask for a sustainability of the e-infrastructure. We should aim at GISELA services for a wider number of users with high impact applications in order to increase NRENs interest in the e-Infrastructure and, then, to spontaneously obtain NRENs commitment to keep it up and running.

Nowadays, the problem to find a way to increase the number of e-Infrastructure users is quite common. In Europe, except for HEP, EGI is facing a similar problem, obviously specific to the different local situation, and the communities are looking for new solutions to increase the e-Infrastructure's spreading. The technical difficulties to use e-Infrastructures are generally considered the biggest hindrance towards a larger uptake. Authentication models based on digital certificates and quite complex interfaces to interact with the e-Infrastructures are, often, insurmountable barriers for scientists that choose to continue to work with their old tools without really knowing and appreciating e-Infrastructure capabilities

To overcome the above constraining and/or limiting factors, we have decided to develop a new strategy during the second year of the project, changing our approach to attract and support user communities. In this sense we will focus our efforts in two directions:

- Provide users with a friendly and intuitive environment to easily use the e-infrastructure
- Build and certify technical capacities to run centralised services.

The first group of actions will be addressed by the implementation of a Science Gateway environment. The second one will be accomplished by a Training Program allowing several technical teams in Latin America to run the centralised services.

The GISELA Science Gateway will be developed starting from the existing application Web site using new technologies and adapting the contents to become VRC-driven. This new applications centre will not be any more an information container only but it will allow users to exploit the e-Infrastructure too. Considering as main requirements simple authentication mechanisms and intuitive user interfaces to execute applications on the e-Infrastructure, we are working to provide the GISELA e-Infrastructure with a Science Gateway, a Web site allowing users to fully exploit the e-Infrastructure capabilities through a normal Web browser. Users will obtain access to the Science Gateway using a pair username/password (provided by an Identity Federation) and submit applications simply filling a Web form. The e-Infrastructure authentication mechanism based on digital certificates will be hidden to the end users through a shibboleth based authentication mechanism<sup>6</sup> and the use of robot certificates<sup>7</sup> still keeping the needed security level. Applications will be automatically submitted by the Science

---

<sup>6</sup> Shibboleth web site, see <http://shibboleth.internet2.edu/>

<sup>7</sup>A grid portal with robot certificates for bioinformatics phylogenetic analyses, see <http://dl.acm.org/citation.cfm?id=195241>



---

Gateway on the e-Infrastructure by using a library based on the Simple API for Grid Applications (SAGA) OGF standard<sup>8</sup>.

The main functionalities that will be implemented are the following:

- Authentication based on username/password pair provided by an Identity Federation, hiding digital certificates;
- Access to the information on WP3, VRCS and applications supported by GISELA;
- Availability of a new concept of application registry allowing users to directly submit selected high impact applications on the GISELA e-infrastructure in an easy way;
- Creation of one section for each VRC supported, containing specific information on the VRC applications such as application characteristics, case studies and success stories. These Web-site areas will become specialised Science Gateways for the VRCS assuming the role of access point to submit some VRC applications on the e-Infrastructure;
- Collaborative tools;
- Training Materials;
- List of publications inherent infrastructure usage and exploitation.

These functionalities will be updated, reviewed and integrated according to the users feedback collected through collaborative tools as mailing lists or forum and during the training events.

Initially, the number of applications executable through the Science Gateway will be limited. We plan to start with 3-5 applications selected according to their potential impact on the LA scientific communities. This number will be increased, during and after the project lifetime, continuing the selection of high impact applications and providing applications developers with support to integrate the application on the Science Gateway.

According to what stated above, the training activity will be properly adapted to increase the Science Gateway “power” and to spread the new paradigm. The target of the schools devoted to applications will differ from the old training events, gridifying applications using directly middleware interfaces. The objective we want to reach at the end of a school is twofold:

- Integrate new applications on the Science Gateway (an integrated application is an application in which a user can submit jobs to the infrastructure directly from the Science Gateway);
- Demonstrate the Science Gateway capabilities to Latin American scientists.

School location and applications involved will be suggested by CLARA and NRENs. Emerging or large communities and high impact applications will be chosen to enlarge the infrastructure utilisation with the aim of the sustainability in mind. In this way, the Science Gateway will become a way to attract new VRCS towards the GISELA infrastructure.

Also the system administrator training events will be organised according to the CLARA needs. In particular we will organise two kinds of events:

- Events to teach Central Service management, organised in the competence centres chosen by CLARA to manage e-infrastructure Central Services after the end of project;
- Events to deploy/manage, in situ, Resource Centres that will be integrated on the e-infrastructure managed by CLARA after the end of project.

---

<sup>8</sup> Simple Api for Grid Applications (SAGA) OGF specification: [www.ogf.org/documents/GFD.90.pdf](http://www.ogf.org/documents/GFD.90.pdf)

The validity of the new approach to increase the infrastructure utilisation just described will be reviewed every three months, validated through Science Gateway usage statistics and feedbacks collected from users. The strategy could be opportunely modified and updated according to the outcomes, positive or negative, reached implementing in this way the Plan-Do-Check-Act cycle.

Finally we should comment that GISELA, according to its intention expressed in its DoW, will investigate, during its second year, the potentialities of Cloud computing in the Latin American context. In practice, we intend to explore the possibility of “*seamless execution of CPU-intensive applications in hybrid e-Infrastructures augmented with the capability of interfacing with Cloud computing providers*” (from Annex-3 of the DoW). This should contribute to form a more complete long-term strategy as of the future CLARA offer of Advanced Computing Services.

---

## **5. PLAN FOR THE SECOND YEAR**

In this section we describe the plan for the second year to implement the new strategy previously described.

The plan will be realised in two phases:

- October - November 2011: deployment of the first release of the VRC-Driven Science Gateway integrating several high impact applications;
- December 2011 - end of the project: Aggressive training and user support to fully exploit the Science Gateway model by integrating new applications and, possibly, involving new VRCs.

### **5.1. VRC-DRIVEN SCIENCE GATEWAY DEPLOYMENT AND EVOLUTION**

The Science Gateway deployment and its evolution will proceed as follows:

First phase (October-November 2011):

- Identification of the first set of applications to be integrated in the Science Gateway, including applications promoted by CLARA and the Latin American NRENs and other high impact applications of the Application database;
- Deployment of the first release of the VRC-Driven Science Gateway for a single selected VRC.

Second phase (December 2011- end of project):

- Carry on the CLARA and NRENs identification of new applications to be integrated (they could already be in the Application database) and of new VRCs to be supported. An application could be evaluated for its integration on user request too;
- Accordingly to the result obtained in the previous steps, organise training events or/and provide remote support to integrate new applications on the Science Gateway or to create new specialised Science Gateways;
- Drive the creation of Identity Federations in Latin America.

The steps of the second phase will be executed in an iterative way and could overlap. Moreover, CLARA could continue this interaction on the long term, making the Science Gateway a tool that could evolve to satisfy new users, new VRCs or new requirements (for example a new application to be supported in a VRC).

### **5.2. TRAINING PROGRAM & TRAINING EVENTS**

The Training Program (TP) for the second year is of outmost importance for reaching the long-term sustainability objective.

The TP was designed in close cooperation with CLARA representatives, who much helped in defining specific purposes and selecting dates, venues and targeted audiences of the proposed Training Events.

Accordingly, the TP for the second year aims to support CLARA in the achievement of three main objectives:

- Increase the usage, and thus the utility and added value, of the e-Infrastructures in LA;
- Increase number of available resources offered to VRCs and user communities in LA;

- Prepare the CLARA TT to provide sustainable user support services beyond the project end. For this purpose, special Site Administrator training will be organised for those RCs capable to provide GOC Services.

The CLARA TT will help WP3 to maximise the impact of all Training Events, by means of:

- Outreaching those VRCs and user communities in LA more prone to generate long-term demand of computing and storage services;
- Identifying most relevant applications to be converted into Scientific Gateways, like those that may attract more user communities, and/or those to which funding agencies are more sensitive.

The TP will be quarterly updated, in the context of PDCA cycles, in order to cope with VRC real needs and business opportunities detected by CLARA in terms of long-term sustainability.

Table 4 describes the training actions scheduled for the second year, which will be tailored to three different target audiences.

**Table 4: Description of training actions to be performed during the second year**

| Target audience                        | Training actions   | Training artefacts   | Expected outcome  |
|--|--|--|---|
| <b>CLARA TT selected members</b>       | Teach how to take control of new VRC-driven application site<br>Teach on EGI-procedures, tools and means concerning user support services  | Virtual meetings referring to existing training material                                   | Guarantee continuation of user support services after project ends  |
| <b>Site administrators</b>             | Teach how to install and operate RCs<br>Teach how to enable support of high impact VOs<br>Teach standard procedures involved in managing a GOC, maintaining central services operational, and provisioning user support services | Virtual meetings<br>Schools for system administrators<br>Schools for system administrators | Increase number of available resources<br>Increase overall usage of e-Infrastructures<br>Guarantee continuation of user support services after project ends                                   |
| <b>Latin American User communities</b> | Introduce Scientific Gateways as new model to port applications to grid<br>Teach how to integrate grid applications as Scientific Gateways<br>Teach how to use applications integrated as Scientific Gateways                    | Schools for application porting  | Increase overall usage of e-Infrastructures<br>Satisfy user needs in a better ways<br>Increase number of supported application groups<br>Increase number of users exploiting e-Infrastructure |

---

Schools for application porting will follow a new approach, not anymore based on gLite usage. The schools will aim at satisfying user needs in a better ways and therefore at increasing the overall usage of e-Infrastructures as well as the number of users exploiting them.

These schools will last for two weeks and will be organised along the following pattern:

- During the first week, Java developers will be invited to integrate their application into a Scientific Gateway and make their application exploiting the GISELA resources;
- During the second week, users will be taught on how to register, access and use their applications by interacting with their respective Scientific Gateway.

Eventually, once relevant Scientific Gateways get integrated in the GISELA e-Infrastructures, it could be possible to fully dedicate Application Schools to user communities willing to exploit already available Scientific Gateways.

The schools for System Administrators will last one (1) week and may be jointly organised with Schools for application porting, depending on actual needs of RCs in the country where the School for application porting takes place.

Tutors from GISELA member institutions and with experience in previous EELA-2 training events shall be requested to attend those organised by GISELA during the second year. The CLARA TT members will participate also in order to gain experience in organising this sort of training events.

In the same way as it was done during the first year, tutors attending training events will be in charge of either generating new training materials or improving existing ones. These materials will be saved and referred both in the Events and Documents Servers.

In addition, as a follow up of the new strategy, tutors, users and administrators will help providing contents, both in English and Spanish, to be added into the new VRC-driven application site such as success stories, training guidelines, referencing publications of researchers using results achieved using the infrastructure, etc..

### **5.2.1. Tentative training agenda**

Training Schools will start after December 2011, when the new tools will be made available (VRC-driven Web site and software to support the deployment of Science Gateways).

A tentative agenda of these training Schools (events #3 to #6) will be proposed to the EPIKH project (with which a MoU was signed<sup>9</sup>) to jointly co-organise them. EPIKH already confirmed that it might support more LA countries (apart from Mexico, Brazil or Chile), if duly justified. Table 5 lists the training events foreseen for the secondproject-year.

---

<sup>9</sup> See <http://documents.gisela-grid.eu/record/166>

**Table 5: Training events tentative schedule and Characteristics.**

| # | Month       | Date              | Venue                                 | Target audience   | Focus on  | Co-Organisers          |
|---|-------------|-------------------|---------------------------------------|---|---|------------------------|
| 1 | M14         | Oct'11            | Ecuador                               | CEDIA site administrators   | Installing RC   | CEDIA <sup>10</sup>    |
| 2 | M14-<br>M15 | Oct'11-<br>Nov'11 | <i>Virtual-meetings</i> <sup>11</sup> | NRENs and other GISELA members, site administrators   | Installing RCs<br>Upgrading RCs<br>Supporting VOs                   | NRENs                  |
| 3 | M17         | Jan'12            | Ecuador or Colombia                   | System administrators   | Supporting VOs<br>Managing GOCs <sup>12</sup>                       | EPIKH, CEDIA or RENATA |
| 4 | M17         | Jan'12            | Colombia <sup>13</sup>                | Application developers<br>Users communities on bioinformatics, biomedicine and medical imaging            | Supporting new applications;<br>Integrating new Scientific Gateways | EPIKH, RENATA          |
| 5 | M18         | Mar'12            | Argentina <sup>14</sup>               | System administrators from LA institutions selected by CLARA  | Installing RCs<br>Upgrading RCs<br>Managing GOCs                    | EPIKH, INNOVA-T        |
| 6 | M18         | Mar'12            | Central America <sup>15</sup>         | Policy makers<br>Users from bioinformatics, tropical diseases and natural disaster management communities | Supporting new applications;<br>Integrating new Scientific Gateways | EPIKH, CLARA           |

Although four training Schools will be jointly organised with EPIKH (two Schools for System Administrators and two Science Gateways Schools), more training Schools could be arranged during the last two quarters of the project. This will depend on actual needs identified by CLARA and measured effectiveness of the initially proposed training schools.

<sup>10</sup> CEDIA will fund travel of one tutor to support integration of recently acquired hardware.

<sup>11</sup> In cooperation with WP4, several virtual meeting will be held with system administrators in order to accelerate the integration of resources of those institutions who didn't manage yet to do so. See D4.2, <http://documents.gisela-grid.eu/record/249/files/GISELA-D4.2-v1.3.pdf>

<sup>12</sup> Ecuador and Colombia NRENs are candidates to offer GOC services at their respective countries.

<sup>13</sup> High interest detected in Colombia concerning "biomed", "bioinformatics" and "medical imaging" VRCs.

<sup>14</sup> CLARA considers important to have a School for Grid Administrators in Argentina, as there are several RCs who still have to integrate pledge resources.

<sup>15</sup> CLARA detected a big opportunity in Central America, as policy makers are much easier to be contacted and eventually convinced. This Applications School will be mainly targeted to user communities with interest in these topics: bioinformatics, tropical diseases and natural disaster management.

## 5.2.2. Training and documentation supporting tools

The following sections describe the tools that will support the training process.

### 5.2.2.1. GISELA Event Server

Non-virtual training actions such as Schools, tutorials and workshops will be published in the GISELA Events Server (<http://indico.gisela-grid.eu/>). Training schedule, training materials, enlisted tutors and attendants will therefore be publicly consulted.

### 5.2.2.2. GISELA Document Server

Training material produced by any of the training actions will be validated, categorised and uploaded to the GISELA Documents Server (<http://documents.gisela-grid.eu/>), following the same rules as stated at Section 6.3.2.4 of deliverable D3.1 (<http://documents.gisela-grid.eu/record/88>).

### 5.2.2.3. New VRC-driven Applications Web site

New contents will be provided substituting those currently available at <http://applications.gisela-grid.eu>. The CLARA TT will learn how to manage this new VRC-driven site, as it is to be maintained after the project end. Moreover, the CLARA TT will facilitate an alias for this site within redclara.net, to be disseminated to the e-Infrastructure users and administrators.

Applications leveraged from EELA-2, as well as those being used by high impact VOs, will be described and grouped by VRCs. This site will also publish Success Stories and refer to publications from user communities actively exploiting e-Infrastructures in LA.

Simple and easy to follow training materials will be made available through this new site, like (but not limited to): general guidelines introducing to Grid concepts and Science Gateways, instructions on how to interact with Science Gateways, how to register new user accounts, how to convert grid ported applications into Science Gateways. The site will include links to Wiki pages leveraged from other projects, as well as links to EGI sites providing more training resources and further VRC support.

All contents of this new VRC-driven applications site will be provided both in English and Spanish.

## 5.3. FOCUSED DISSEMINATION

As baseline support to the new strategy, it is anticipated that the Dissemination actions carried out by WP2 will more focus, over the second year of the project, on the announcement and the implementation of the Science Gateways-based new strategy. The actual plan is to take advantage of the commitment of the LA NRENs to the dissemination of the project initiatives. In practice, the WP2 Management and its NRENs collaborators will

- Orientate their dissemination activities towards publicizing the introduction of the Science Gateways as new paradigm;
- Suggest new actions to sustain the new SG approach.

## 5.4. LIAISON WITH EGI-RELATED AND OTHER EXTERNAL ENTITIES

As already underlined, the dependency of the sustainability of the GISELA e-Infrastructure to its cost-effective use by numerous users is, to a large extent shared by other Regional Initiatives, including EGI.

The established relation with EGI and Regional Initiatives (through the signed MoUs) has shown to be efficient in finding common solutions to common problems as illustrated by the two Workshops

---

“Resource Providers meet VRCs” co-organised with chain at the occasion of the EGI User Forum (Vilnius 2010) and the EGI Technical Forum (Lyon 2011).

Keeping the momentum of these recent successes, CHAIN and GISELA have decided to co-organise a new issue of their Workshop called “Sustainability & Interoperability” at the forthcoming “International Symposium on Grids and Clouds 2012,” to be held in Taipei on 26<sup>th</sup> Feb. to 2<sup>nd</sup> Mar. 2012. The Workshop program includes a session on “Status and plans of sustainability” in which the various approaches will be confronted and a session on “Status and plans of interoperability” which, amongst others subjects, foresees a debate on “Science Gateways and Authentication & Authorisation Infrastructures”. Having in mind that the Symposium will have a full session dedicated to “Business Models & Sustainability”, it is expected that ISGC’2012 will much help in consolidating new sustainability strategies being put at work.



## 6. OVERALL IMPACT & PROGRESS EVALUATION

To have a better measure of the outcomes of the new strategy described in the previous sections we decided to add five (5) new quality metrics to the WP3 work plan (see Table 6):

- Number of Active Users in the catch-all VO (prod.vo.eu-eela.eu): as active user we consider somebody exploiting regularly the infrastructure, i.e. not sporadically. In any case, this metric doesn't measure the whole set of the e-Infrastructure users due to other VOs exploiting the infrastructure;
- Number of Active Users in the VRC-driven Science Gateway: users submitting regularly applications from the VRC-driven Science Gateway;
- Number of Applications available on the e-Infrastructure, i.e. applications just ready to be executed on the e-Infrastructure;
- Number of Applications integrated in VRC-driven Science Gateways i.e. application that an user can submit to the infrastructure from the Science Gateway;
- Number of specialized Science Gateway for VRC i.e. Science Gateway sections playing the role of access point to submit some specific VRC applications on the e-Infrastructure.

Other measurements, like the “overall number of active users” (coming from all VOs using the infrastructure) and “number of applications actually used”, could be useful to better evaluate the WP3 progress but the current lack of accounting tools able to measure these numbers prevents us to add them in the quality metrics table.

**Table 6: WP3 (upgraded) quality metrics**

| Quality metric  | Current status Y1Q4 | Expected outcome Y1 | Expected outcome Y2 |
|---|---------------------|---------------------|---------------------|
| 1 Number of Active Users in the catch-all VO (prod.vo.eu-eela.eu)     | N.A. <sup>16</sup>  | N.A.                | 100                 |
| 2 Number of Active Users in the VRC-driven Science Gateway            | N.A.                | N.A.                | 100                 |
| 3 Number of Applications in production status on the e-Infrastructure | 34                  | N.A.                | 40                  |
| 4 Number of Applications integrated in VRC-driven Science Gateways    | N.A.                | N.A.                | 15                  |
| 5 Number of specialized Science Gateway for VRC                       | N.A.                | N.A.                | 2                   |
| 6 Number of supported VRCs  | 3                   | ≥ 2                 | ≥ 3                 |
| 7 Number of application's groups from Latin America                   | 19                  | ≥ 15                | ≥ 36                |
| 8 Number of training events   | 3                   | ≥ 1                 | ≥ 4                 |
| 9 Number of self-training material                                    | 40                  | ≥ 20                | ≥ 40                |
| 10 Number of papers published (WP3 related)                           | 5                   | ≥ 5                 | ≥ 20                |

<sup>16</sup> The plan is to start measuring this metric starting on M14.

## **7. CONCLUSIONS**

The first year WP3 objectives, as described in the DoW, have been reached and the WP metrics have been fulfilled. However, notwithstanding these metrics, GISELA suffers from problems, to a large extent, common to all the e-Infrastructures, such as the difficulty to attract new users and consequently to guarantee the sustainability of its e-Infrastructure.

The technical barriers to use e-Infrastructures are considered the biggest hindrance towards a larger diffusion of the e-Infrastructures. In an attempt to overcome this weakness, GISELA has chosen to orientate its VRC support strategy towards the adoption of the Science Gateway whose main goal is to offer a simple user friendly Web interface to exploit efficiently the e-Infrastructure services.

The Science Gateway model has been presented and the road map of its implementation has been sketched. The success of the new strategy implies the adaptation of the Training plan and also of the Dissemination initiatives towards all GISELA stakeholders, to publicize the reasons and the expected results of the new User support approach.

Moreover, a strong commitment of CLARA has been planned to drive this new kind of support and training activities. CLARA will identify emerging or large communities and high impact applications to be supported with the aim to integrate their applications on the Science Gateway or to create specialized VRC Science Gateways. Training venues and formats will be chosen accordingly.

Finally it is expected that the deployment of the Science Gateway model to get over the technical obstacles to use e-Infrastructures will much gain from the ongoing collaboration with EGI and the Regional Initiatives.