# Institut des Grilles - Report to the scientific council

Dominique Boutigny, Vincent Breton, Michel Daydé, Frédéric Desprez, Xavier Jeannin, Fairouz Malek, Frédéric Suter, Guy Wormser

# Foreword

The present document provides a brief overview of the activities of CNRS Institut des Grilles in 2008 and 2009.

Its content is not aimed at being exhaustive: additional information will be provided through the oral presentations given on June 1st 2010.

Table des matières

Institut des Grilles - Report to the scientific council 1

Foreword 1

Introduction 2

Presentation of CNRS Institut des Grilles 2

Presentation of France Grilles, the French National Grid Initiative 3

Research grids 4

Introduction 4

Research topics 4

Structuration of research on grids in France 5

The Grid'5000 research community 5

Research activities in HEMERA 7

Some trends in grid / cloud computing 7

Production grids 8

Introduction 8

The gLite operated production grid 9

The Decrypthon production grid 9

Interface between research and production grids 10

Selected projects 10

Simulating Data-Intensive Grid Applications 10

SimGLite, when SimGrid meets gLite 11

Energetic Efficiency in Grids: from research to production 11

Virtual screening of seeds 12

XWHEP, A secure Computing Grid interconnected to EGEE 12

Modeling, Simulation and High Performance Computing for solar energy 12

High Performance Computing on GPUs for integrative biology 12

Perspectives 13

Networking activities 13

Activities within EGEE-III SA2 activity 13

Network support within EGI and the French NGI 14

Science on the grid 16

LCG France 16

Introduction 16

The state of the art: WLCG 16

The French contribution to this effort: 18

Conclusion 21

Bibliography 22

Other user communities 22

International collaborations 23

Collaboration with ESFRIs 23

EGI 23

ELIXIR 24

LifeWatch, 24

European projects 24

DEGISCO 25

EDGI 26

EGI-Inspire 27

EUMedGrid-Support 28

GISELA 29

SHIWA 30

StratusLab 31

Bilateral collaborations 32

Bilateral collaborations with Africa and Middle-East 32

Bilateral collaborations with Asia 32

# Introduction

## Presentation of CNRS Institut des Grilles

The CNRS Institut des Grilles was created in 2007. The present report reflects on activities that have been taking place up to May 2010.

The last two years were marked by major changes in the national and international landscapes. The most significant changes are the end of EGEE-III and kick-off of EGI at the European level and the creation of the Scientific Interest Group (Groupement d’Intérêt Scientifique, GIS) France Grilles to operate the French National Grid Initiative.

The CNRS Institut des Grilles (IDG) was created on September 2007 to federate the contributions to Grids deployment and Grids research within CNRS, to reinforce interaction between computing science research and production infrastructures and to represent the CNRS in European projects.

The management structure put in place at the institute creation reflected its missions: two deputy directors who were in charge respectively of production grids and computing science research supported the director (see figure 1). Following Guy Wormser resignation in December 2009, Vincent Breton was appointed as director. The management structure was modified as follows:

* Guy Wormser was appointed as Honorary Director
* A Charge de Mission position was created to specifically address the relationship between production grids and research grids. This position was taken by Frederic Suter

The present management structure is shown on figure 1.

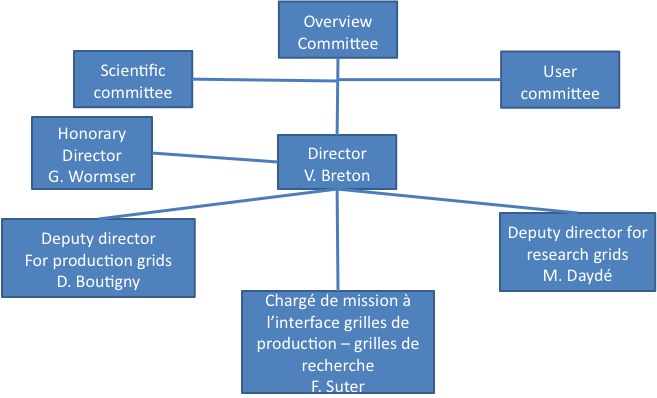


Figure 1: management structure of the Institut des Grilles

## Presentation of France Grilles, the French National Grid Initiative

The French National Grid Initiative formal structure is a "Groupement d'Intérêt Scientifique" signed between the 8 major research entities in France: Ministry of Research and Education, CEA, CNRS, CPU, INRA, INRIA, INSERM and RENATER. France Grilles represents more than 500 researchers and engineers coming from 50 different units and has the operational responsibility for the French Grid.

CNRS Institut des Grilles is the acting legal entity on behalf of all the partners. In the future, CNRS will sign bilateral conventions with some other partners, associations or private companies.

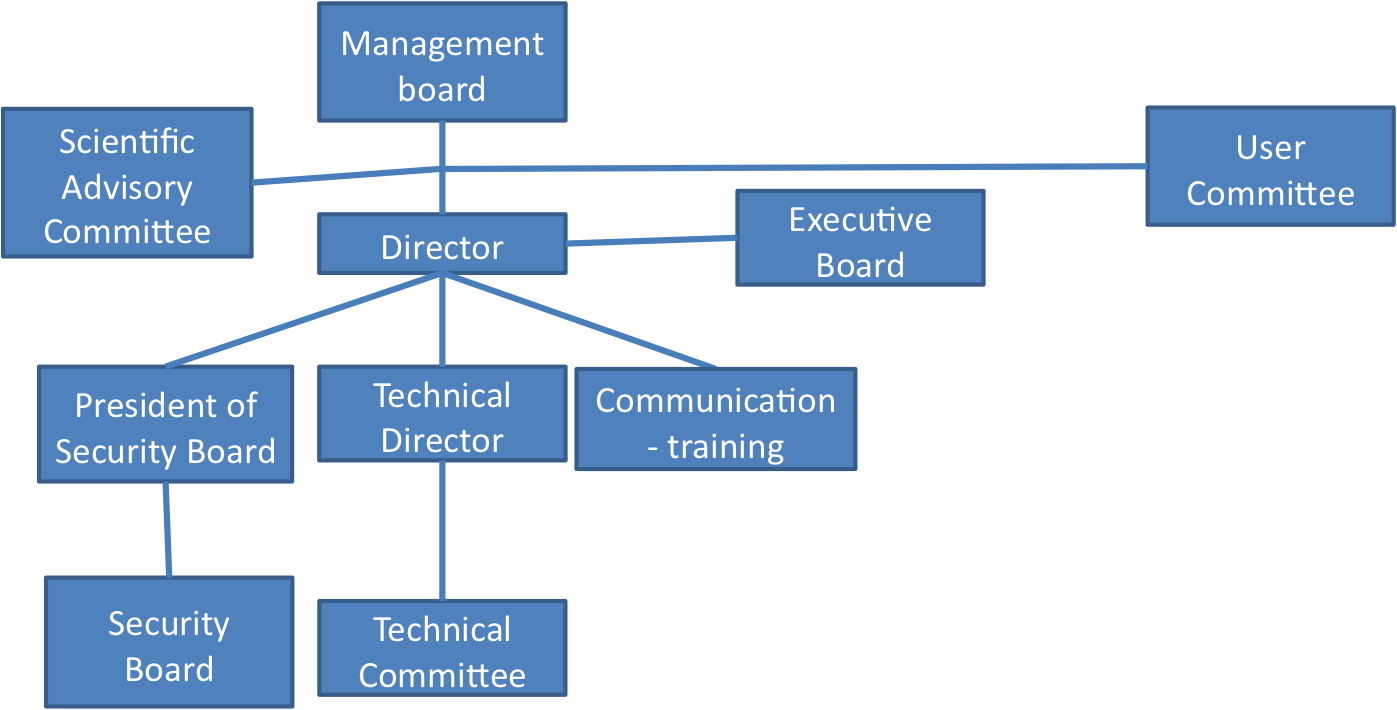
Figure 2 shows the management structure of France Grilles.

Figure 2: management structure of France Grilles

It is not yet finalized as the application sector requires to be organized. It is foreseen to adopt a structure inspired from EGEE NA4 activity with the equivalent of scientific clusters. The clusters would all be represented in an application management board.

# Research grids

## Introduction

For more than 10 years, grids and more generally speaking « Large scale distributed infrastructures » attract considerable attention. This is still the case with recent evolutions such as « Cloud Computing ».

Our goal is to make an overview of grid research in France in terms of topics, management structures and trends.

This overview does not pretend to be exhaustive as it is based on an analysis of both the work within the Grid'5000 community and of the web sites and activity reports of the laboratories that do not always reflect the fast evolution in our area.

### Research topics

Research around large scale distributed computed infrastructures is quite old and has started in the nineties for Grid Computing. Since 2008, Cloud Computing has emerged as a hot topic.

French researches can be classified into topics that not only consider grid and cloud but also more general large scale distributed platforms (large clusters, sensor networks, etc):

* Environment for large scale distributed platforms: middleware, efficient use, distributed systems and objects, distributed software infrastructures, large scale data management, security, virtualization, reconfiguration, autonomic management, etc
* Services: service oriented computing, management, dynamic adaptation, workflows
* Distributed applications: new concepts, languages and tools, context-sensitive applications, mathematical modeling, abstraction
* Objects or software components architectures for distributed intelligent systems
* Algorithms and scheduling: distributed algorithms, P2P, large scale computations, combinatorial and NP-hard problems, …
* Design of distributed applications, parallel programming, refinement, development and proof of parallel programs, application of formal languages, models

The community involved in these very general research topics represents about 530 researchers (around 450 academics, 50 INRIA researchers and 30 CNRS researchers).

Research topics targeting grids more specifically are the following:

* Environment for the Grid: middleware, languages, libraries, high performance execution support, networks, communications, energy, etc
* Management of resources: virtualization, autonomic administration, workflows, etc
* Data Management in large scale distributed systems (clusters, grid and P2P), distributed storage of data
* Scaling
* Algorithms and scheduling
* High performance networks
* Models and utilization of large scale data and computing infrastructures
* Gridification of applications: simulations and numerical algorithms in various areas (environment, electromagnetism, chemical engineering, biology and health, linear algebra, combinatorial problems, …)
* Pervasive grids

**As a conclusion, 12% of researchers in computer science are studying large scale distributed infrastructures. Among them, around 310 researchers (7% of the research community) are working on grids and clouds.**

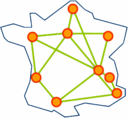
## Structuration of research on grids in France

### The Grid'5000 research community

Most of the research on grids in France (in fact 2/3 of the researchers) is structured around the GRID'5000 national platform and the ALADDIN-GRID'5000 action driven by INRIA that ensures the development and support of the GRID’5000 infrastructure.

The HEMERA initiative also supported by INRIA is in charge of the scientific animation around Grid'5000. It is organized around scientific challenges and working groups.

The Grid'5000 experimental platform involves 9 French sites (Bordeaux, Grenoble, Lille, Lyon, Nancy, Orsay, Sophia-Antipolis, Toulouse et Rennes) – connected using the French national network for research and education RENATER – and one Brazilian site (Porto Alegre). It provides over 5,000 cores.

The goal of Grid'5000 is to provide a large scale, nation-wide platform for distributed computing at large scale. Mode details can be found on http://grid5000.fr.

HEMERA involves 21 teams (15 are INRIA, the 6 left are in Strasbourg and Toulouse). It represents an amount of 174 researchers (119 academics, 38 INRIA and 17 CNRS) corresponding to about 2/3 of the researchers on grids in France.

Research teams involved in HEMERA are today (per site):

* Bordeaux : CEPAGE and RUNTIME (LaBRI).
* Grenoble :MESCAL (LIG).
* Lille : DOLPHIN (LIFL).
* Lyon : GRAAL and RESO (LIP).
* Nantes : ASCOLA (LINA).
* Paris : GRAND-LARGE (LRI) and REGAL (LIP6).
* Rennes : ASAP, KERDATA, MYRIADS, SAGE (IRISA).
* Nice-Sophia : OASIS (I3S).
* Strasbourg : ALGORILLE (LORIA), ICPS (LSIIT).
* Toulouse : ACADIE , APO , ASTRE (IRIT), MINC and MRS (LAAS)

However, quite a few research teams on grid / cloud computing are not involved in HEMERA today. Here is a non-exhaustive list:

* Team of David Hill (LIMOS, Univ. Clermont).
* AND and CARTOON (LIFC, Univ. Franche-Comté)
* MAP (LIFL, Lille).
* ATLAS-GDD (LINA, Nantes).
* PRV (LIFO, Orléans).
* Groupe Grilles (LAL, Orsay).
* PEQUAN (LIP6, Paris).
* T2I (LMA, Univ. Pau),
* Groupe SysCom (CRESTIC, Univ. Reims).
* MODALIS (I3S, Sophia).
* Team of Frédéric Suter (CC-IN2P3, Lyon)

### Research activities in HEMERA

HEMERA is organized around scientific challenges. Seven were initially identified:

* Network:

1. Traﬃc Awareness

* System:
  + 1. Energy proﬁling of large scale applications
    2. Robustness of large systems in presence of high churn
* Programming paradigm:

4. Large scale computing for combinatorial optimization problems

5. Scalable Distributed Processing Using the MapReduce Paradigm

* Application domain specific:

6. Multi-parametric intensive stochastic simulations for hydrogeology

7. Thinking GRID for Electromagnetic Simulation of oversized structures

Eight working groups have been set up:

1. Transparent, safe and efficient large scale computing
2. Energy Efficient Large Scale Experimental Distributed Systems
3. Bring the Power of Grids to Internet-Users thanks to Virtualization Technologies
4. Efficient exploitation of highly heterogeneous and hierarchical large-scale systems
5. Efficient management of very large volumes of information for data-intensive applications
6. Completing challenging experiments on Grid’5000
7. Modeling Large Scale Systems and Validating their Simulators
8. Network metrology and traffic characterization

## Some trends in grid / cloud computing

The working groups in HEMERA (see end of previous section) already give some of the major trends in terms of grid / cloud research topics for the next years:

* Models for large scale distributed infrastructures, experimental methodology, reproducibility of experiments
* Security on large scale platforms (cluster, grids, clouds, etc)
* Taking into account dynamicity at all levels: virtualization, autonomic management, monitoring, process automation, …
* Mastering scalability
* Energy-aware management of resources: algorithms, scheduling, monitoring, …

The underlying cloud architecture includes a pool of virtualized resources (compute, storage and networking, etc) that can be aggregated and used as a platform to process user requests with some Service-Level Agreement. Cloud architectures include provisions to both guarantee service delivery for clients and optimize efficiency of resources of providers. Efficient provisioning may imply scaling resources up/down, monitoring for managing the workload, studying fault-tolerance issues and energy optimizations, etc

Some emerging topics specific to clouds are the following:

* Cloud architectures and provisioning
* Grid / Cloud interoperability, private / public cloud interoperability
* Cloud and HPC
* Programming models, applications and middleware suitable for dynamic cloud environments.
* Novel human interfaces and browsers for accessing clouds: interaction of mobile computing, Commerce and Clouds. ...

Finally, a crucial issue lies with the management and exploitation of large amount of data (possibly heterogeneous) arising from numerical simulations at the Terascale (and soon ExaFlops) or from various applications in biology, high energy, semantic web, … These data are often processed on large scale distributed infrastructures (e.g. data arising from the LHC). Storing and indexing them as well as performing some post-processing for extracting information or for visualization purposes still require efforts and interactions between the various communities (grid / cloud, visualization, large scale databases, etc).

# Production grids

## Introduction

There are two production grids in France:

* the production grid which comes out of LCG and EGEE European projects, using gLite as its middleware
* Decrypthon , a production grid dedicated to medical applications which was developed through a collaboration between AFM (French Myopathy Agency), CNRS and IBM and which uses DIET as its middleware.

The majority of the nodes of these two production grids are located in academic research laboratories and computing centers.

Several regional grids provide a framework for the management of the resources at a regional level.

|  |  |  |  |
| --- | --- | --- | --- |
| Regional grid | Location | Number of sites | Starting date |
| AuverGrid | Auvergne (Clermont-Ferrand) | 6 | 2004 |
| CIMENT | Grenoble | 5 | 2004 |
| GRIF | Ile-de-France | 6 | 2005 |
| TIDRA | Rhone-Alpes (Annecy – Grenoble – Lyon) | 5 | 2008 |

Table of regional grids in France

Regional grids are currently under creation in Bordeaux, Marseille, Montpellier and Strasbourg.

## The gLite operated production grid

The following table provides an estimate of the resources provided by the nodes of the gLite operated production grid. About 2/3 of these resources are used by LCG France.

|  |  |  |
| --- | --- | --- |
| Regional grid or node | Number of cores | Storage in TOctets |
| Bordeaux | 300 | 2 |
| Clermont-Ferrand (AuverGrid) | 1000 | 300 |
| Ile-de- France (GRIF) | 5500 | 1500 |
| Lille | 226 | 2 |
| Annecy - Lyon - Grenoble (CIMENT - TIDRA) | 7000 | 50 |
| Marseille | 500 | 100 |
| Strasbourg | 1000 | 300 |
| Total | 15526 | 2254 |

Table of resources provided on the gLite operated production grid

## The Decrypthon production grid

The Decrypthon project ([http://www.decrypthon.fr](http://www.decrypthon.fr/)), started in 2004, comes out of a collaboration between CNRS, the French Muscular DIstrophy Association (AFM) and IBM. Its aim is to provide research groups in bioinformatics with access to computing and storage resources on a grid platform that includes 6 sites connected through RENATER: Bordeaux I University, Jussieu (Paris VII), Lille I University, Lyon ENS, Orsay (Paris XI) and Rouen (CRIHAN).

Decrypthon provides the necessary resources for exploiting the grid. It gives research grants to the groups selected though a call for projects and provides support for the technical aspects: modeling, application porting, data management, etc. The GRAAL team at ENS Lyon (INRIA) acts as grid expert support team since the project beginning. In 2006, the DIET middleware (<http://graal.ens-lyon.fr/DIET/>) developed in the GRAAL team was selected as the Decrypthon Grid operating system. This middleware allows transparent job submission on the six University computing centers through the RENATER network. Some particularly computing intensive applications are also deployed on the World Community Grid (WCG, [http://www.worldcommunitygrid.org](http://www.worldcommunitygrid.org/)) volunteer grid.

The Decrypthon University Grid is organized around key elements: the DIET middleware, a web portal to access the grid resources and local schedulers in each computing centers. The web portal located on a dedicated server in Orsay has a specific web application for each project that allows submission of scientific tasks on all Decrypthon resources. The portal uses the DIET functionalities for job submission on dedicated servers for each application. The DIET platform deployed for Decrypthon is made of a Master Agent hosted on a machine in Orsay and a Server Daemon (SeD) that is launched on each frontal access point of the computing centers. SeDs are connected to the Master Agent: they collect information on resource availabilities and performance and submit jobs to the local scheduler (Loadleveler, PBS, OAR, etc). The local batch scheduler is in charge of locally supervising the load balancing and proper job execution. Each site defines its own policy for accessing resources as Decrypthon dedicated cores are integrated into the local pool. This way, all the local resources are accessed and shared by the local and the Decrypthon users. Finally, the DIET middleware handles data migration for the input data needed for computation and the storage of output data on dedicated servers.

Today, the Decrypthon program opens up to CNRS IdG within the framework of a partnership that is currently being finalized.

# Interface between research and production grids

In September 2009, the Institut des Grilles and the INRIA action ALADDIN-GRID'5000 have launched a common call for proposal. This call, supported by Cécile Germain-Renaud and Frédéric Desprez, was equally funded by both partners. The total funding was of 20,000 euros. On March 2010, Frédéric Suter was entrusted with the management of this interface project.

The main goal of this call for proposal was to foster research activities at the interface between research on Grids and production grids. The funding granted to the selected project (up to 5,000 euros) was intended as a support to organize meetings between partners. Such meetings are a real opportunity to build concrete synergies and to prepare answers to more important call for proposal. Typically a project funded through this particular call should lead to ANR or FUI projects, or smaller interdisciplinary projects from CNRS or INRIA.

No specific restrictions were made with regards to the topics covered by the proposition. The only requirement was to clearly express some research issues and either the need for a production infrastructure (large computing power or huge amount of storage) or some expected benefits for the production grid community.

## Selected projects

This call received 9 proposals involving more than 25 distinct paterns. Among them, 7 were selected and granted different amounts of money. In what follows we shortly describe the selected projects.

### Simulating Data-Intensive Grid Applications

**Project Leader**: M. Quinson, assistant professor at Nancy University

**Partners**: LORIA, CERN, CC IN2P3

**Grant**: 5,000 euros

**Project Description**:

Data-intensive applications such as those related to the LHC experiments rely on distributed management systems composed of several data repositories across multiple sites. The optimization of such application is difficult as such data management systems are used in a production mode. Any experiment on the infrastructure itself would result in unacceptable service degradations.

The evaluation of new ideas and algorithms through simulation thus constitutes an appealing approach. The performance of several general-purpose grid simulators were evaluated at CERN. The purpose of this project is to build upon the good performance of the SimGrid toolkit, and add the missing features, such as the modeling of storage elements, for the evaluation of data-intensive applications.

### SimGLite, when SimGrid meets gLite

**Project Leader**: F. Suter, CNRS junior researcher at IN2P3 Computing Center

**Partners**: CC IN2P3, LORIA, I3S, CREATIS

**Grant**: 5,000 euros

**Project Description**:

Understanding the behavior of large scale distributed systems under load, such as production grids is a major challenge for different reasons: scientific (optimizing performance), operational (service dimensioning), and applicative (development cycle reduction, performance prediction, problem identification). But such systems do not allow for experimentations or reproducible measures.

This project aims at gathering the expertise of the partners on research grids (stochastic modeling and simulation) and production grids (administration and application deployment) to define tools and methods to reproduce the behavior of a production grid under controlled conditions. This project will model load and errors of the EGEE/EGI production grid, inject these models into the SimGrid toolkit, emulate some gLite components on the ALADDIN-GRID'5000 platform. The resulting framework will be evaluated with real applications.

### Energetic Efficiency in Grids: from research to production

**Project Leader**: L. Lefevre, INRIA junior researcher at LIP-ENS Lyon

**Partners**: LIP, IRIT, LIG

**Grant**: 4,000 euros

**Project Description**:

The impact of large scale distributed systems such as production grids on the world energy consumption becomes more and more important. To face environmental threats and energy costs of production grids, this project aims at studying software solutions that could reduce the energy consumption in grids.

Models and algorithms have already been proposed by the partners in the more controlled environment of research grids such as ALLADIN-GRID'5000. In this project, these results will be brought face to face with the operational reality of production grids (CIMENT, GridMIP, PireGrid). It will show the technical feasibility and the economical interest of saving energy without loosing on performance.

### Virtual screening of seeds

**Project Leader**: G. Da Costa, Assistant professor at University of Toulouse

**Partners**: IRIT, CERFACS, LIMAGRAIN, RAGT, THEODORE, UPETEC

**Grant**: 2,000 euros

**Project Description**:

Companies selling corn seeds have a strong interest in determining with lineage is the most suited to the requirements of their customers. But growing all the lineage is expensive, time consuming, and hazardous. Consequently, resorting to virtual screening is very appealing. This consists in selecting lineages according to the correlations between different genetic markers thanks to algorithmic techniques such as data mining, classifications, or data assimilation.

This project requires access to a production infrastructure because of the huge amount of needed computing and storage resources. The required grant will partly fund a preliminary phase of 6 months conducted by a consortium of research laboratories and agronomic companies to prepare an ANR or FUI project.

### XWHEP, A secure Computing Grid interconnected to EGEE

**Project Leader**: O. Logydensky

**Partners**: LAL, LIP

**Grant**: 2,000 euros

**Project Description**:

A good alternative to production grids composed of several clusters such as EGEE/EGI are desktop grids. This kind of infrastructure is based on volunteer computed connected on the Internet. This raises several security issues that will be tackled by this project. Moreover, the interconnection between a Desktop Grid, managed by the XtremWeb and BitDew software packages, and the EGEE production grid will be considered. The expertise gained during the EDGES FP7 European project will be reused, while a particular emphasis will be put on security aspects.

### Modeling, Simulation and High Performance Computing for solar energy

**Project Leader**: Marc Daumas, Professor at University of Perpignan

**Partners**: ELIAUS, PROMES, LIPN, LIG

**Grant**: 1,000 euros

**Project Description**:

This project considers the simulation of turbulent flow submitted to large thermal gradients. Such numerical simulations can take advantage of an execution on hardware accelerators such as GPU processors. The design of new algorithms, such as 3D cellular automata, will first be done on a research grid before a passage to production.

### High Performance Computing on GPUs for integrative biology

**Project Leader**: D. Hill, Professor at University of Clermont Ferrand

**Partners**: LIMOS, LGME, LPC

**Grant**: 1,000 euros

**Project Description**:

The involved partners study the design of a multi-scale multi-agent system for the simulation of transport in tumorous tissues. This problem can be tackled with Monte-Carlo simulations run on the EGEE/EGI production grid. An interactive part will make use of GPU processors.

A preliminary study will be conducted on a local cluster of GPUs wich will later be included in a grid, which could be a research Grid (ALADDIN-GRID'5000) or a production Grid (AuverGrid).

## Perspectives

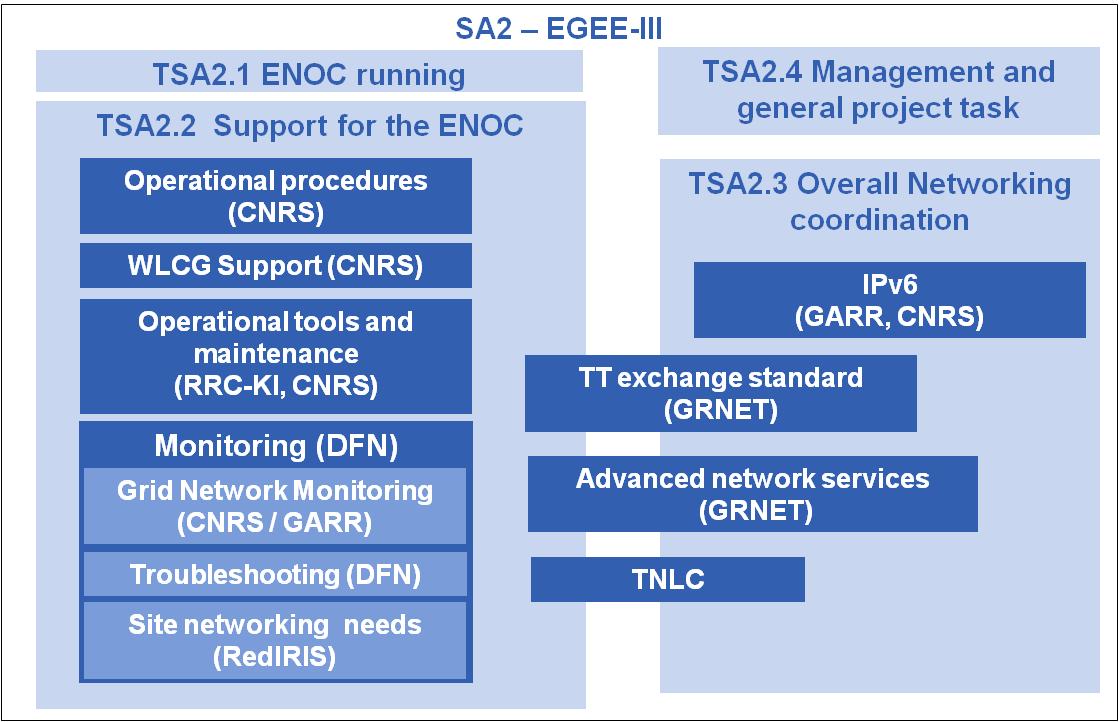
At Fall 2010, a second call for proposal will be edited. We expect to have the same global funding, and aim at selecting the same number of ambitious projects. Nevertheless we would like to change the way the grant is used by the selected projects. Indeed, most of the laboratories have recurrent budgets for missions. Having a few thousand euros more has not a great impact. We strongly believe that funding, at least partially, a master student internship in each project will help their kickoff. Moreover this could be a way to find good student to keep in PhD in the bigger projects that should naturally follow.

# Networking activities

## Activities within EGEE-III SA2 activity

Part of networking activities at CNRS IdG took place within the framework of EGEE-III SA2 networking support activity which was led by CNRS-UREC and involved 7 partners coming from 6 countries plus the international organism DANTE that operates the European network GEANT. SA2 provided both an operational interface ensuring the daily relations with the network infrastructures (EGEE NOC, SLA, IPv6, etc) and a relational interface ensuring the “higher level” of interactions with the network providers and users: LCG, NRENs (DANTE, DFN, GARR, GRNET, RedIRIS, RRC-KI).

|  |  |  |
| --- | --- | --- |
| **Country** | **Total PM planned at M24** | **Total FTE** |
| **France** | **96** | **4.0** |
| **Germany** | **12** | **0.5** |
| **Greece** | **18** | **0.8** |
| **Italy** | **12** | **0.5** |
| **Russia** | **6** | **0.3** |
| **Spain** | **6** | **0.3** |
| **DANTE (GEANT2)** | **3** | **0.1** |
| **Total PM planned at M24** | **153** |  |
| **Total FTE** |  | **6.4** |



All grid projects rely on a low-level layer, the network, but, unfortunately, in the context of multi-domain, the “network operations” (advances network services, network trouble impact assessment, network monitoring) are poorly deployed. SA2 with the support of the NRENs designed and maintained an EGEE NOC in order to provide network support within EGEE. For instance, the ENOC team designed the pioneering federal operational model for the LHC Optical Private Network.

Around the ticketing problem, SA2 investigated an original statistic approach to assess the impact of a network problem reported by network trouble tickets. SA2 also contributed to trouble ticket normalization by proposing a RFC draft[[1]](#footnote-0) and providing software that converts tickets to the normalized format.

SA2 fostered the gLite migration toward IPv6 and this work was congratulated by the European reviewers[[2]](#footnote-1).

SA2 was not in charge of the network monitoring design and suffered from the absence of a solution. In order to solve this issue, SA2 has contributed to the deployment of a network monitoring tool - PerfSONAR MDM – within the Spanish EGEE sites and designed and prototyped a lightweight solution based on grid jobs. The network troubleshooting was also addressed by the development of the tools called PerfSONAR-Lite TroubleShooting Services.

The advanced network services that could be used by EGEE are right now at prototype stage, nevertheless, a deep study was realized on the advanced network services available in Europe and in USA (AMPS, AutoBHAN, GLIF, Phosphorus, IDC and Sherpa) in order to ease the usage of this kind of services by the European grid.

The management of SA2 maintained the collaboration with the NRENs by leading the Technical Network Liaison Committee meeting and by organizing with TERENA a joint NRENs & Grids Workshop[[3]](#footnote-2). Finally, SA2 prepared the migration of the network activity towards EGI, and especially the ENOC, by publishing publicly all the code produced by SA2, helping EGI to install those tools and contributing to the setting-up of the network task of EGI.

## Network support within EGI and the French NGI

Garr will be in charge of **network coordination at EGI level** and the network activity has been sketched out:

* Mapping of planned network hardware interventions and faults to the corresponding Grid sites (alarm system)
* Network trouble support / ticketing and troubleshooting
* Network multi-domain monitoring
  + Focus on main network links and big EGI sites
  + Investigate monitoring tools and solutions
* Deployment and management of network support services: a network information collaborative portal, network troubleshooting, monitoring and alarm system tools
* Network performance support (PERT Performance Enhancement and Response Team)
* Network collaboration between NRENs, DANTE, EGI.eu and NGIs.

**The French NGI network activity project** will foster the development of network operations and the availability of new network services within the French NGI and the EGI. The network operations aim to be able to know how the network behavior impacted, impact or will impact the grid operations in term of performance, failure problem, resource availability, by using reporting, alarm and proactive system. The work will focus in developing network tools and integrating them into an operational dashboard of the French NGI. Advanced network services are currently tested on different infrastructures (Sherpa tool within National LambdaRail), those services together with the virtualization trend will allow the virtualization of the computing infrastructure used by science: massive data center, computing center, virtual observatory and laboratory. The network support task, with the help of network providers, shall address the availability of these resources at network level and also help the experiments to take benefit of these new services of hybrid network when available.

**The first objective** will be to develop a network monitoring solution for the end users:

1. French NGI sites connectivity assessment
2. Network troubleshooting tool deployment
3. Building a topology used by the French NGI and the tools to maintain it
4. Network monitoring for the French NGI sites

A platform including a portal that will host these network tools will be set-up including: troubleshooting tool, monitoring tool, statistics. The access to this tools and data will depend on the identity and the role of the user; for instance, an Atlas scientist could access to data and tools related to the EGI sites that host the Atlas virtual organization. The tools developed within this project must be able to be reused in other contexts, for instance in the context of CEA, of a CNRS institute or within another NGI.

More specifically, the network mission could be defined within the French NGI at two levels:

1. **Network tools development**
   1. Topology design and setting-up
   2. Monitoring tools for the French NGI
      1. Network connectivity assessment
      2. Network performance
   3. Integration into the operations portal of the developed network tools
   4. Other subjects could be investigated in the future (impact assessment of network troubles)
2. **Network operation** 
   1. Network support for the French NGI (Performance, network troubleshooting, …)
   2. Maintaining of the network tools for NGI operations
   3. Maintaining of a testbed site for the development and testing network of grid network technologies

This task should be handled through a tight collaboration between the French NGI partners: laboratory (CC-IN2P3, etc), EGI (GARR), RENATER, CNRS/UREC but also Grid Observatory. The collaboration of regional networks could also be necessary; this is why coordination is required within this activity. UREC plans to devote 1.5FTE for this task, the contribution of the other partners is not defined at the present time.

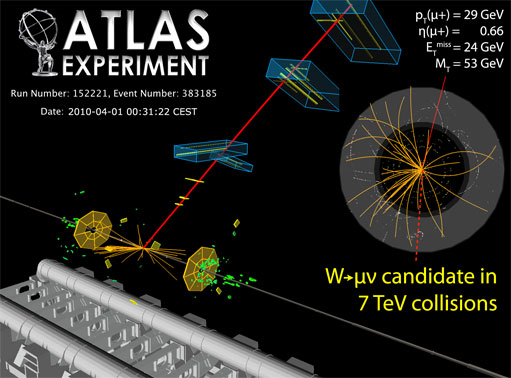
# Science on the grid

## LCG France

### Introduction

The LHC is the Large Hadron Collider, located at CERN (European organization for particle physics) near Geneva. It is the most powerful accelerator in the world in terms of energy, 14 TeV foreseen in the centre of mass of the collision of both hadrons (proton-proton).

The particles are accelerated at a speed close to that of the light and they circulate in a ring of 27 km of circumference buried 100 m underground. The LHC and the associated experimental equipments were conceived and built by international collaborations. It has delivered its first collisions at low energy before 2009 Christmas and it is now operating successfully at a higher energy (7 TeV at c.m.), see **Fig. 3.**



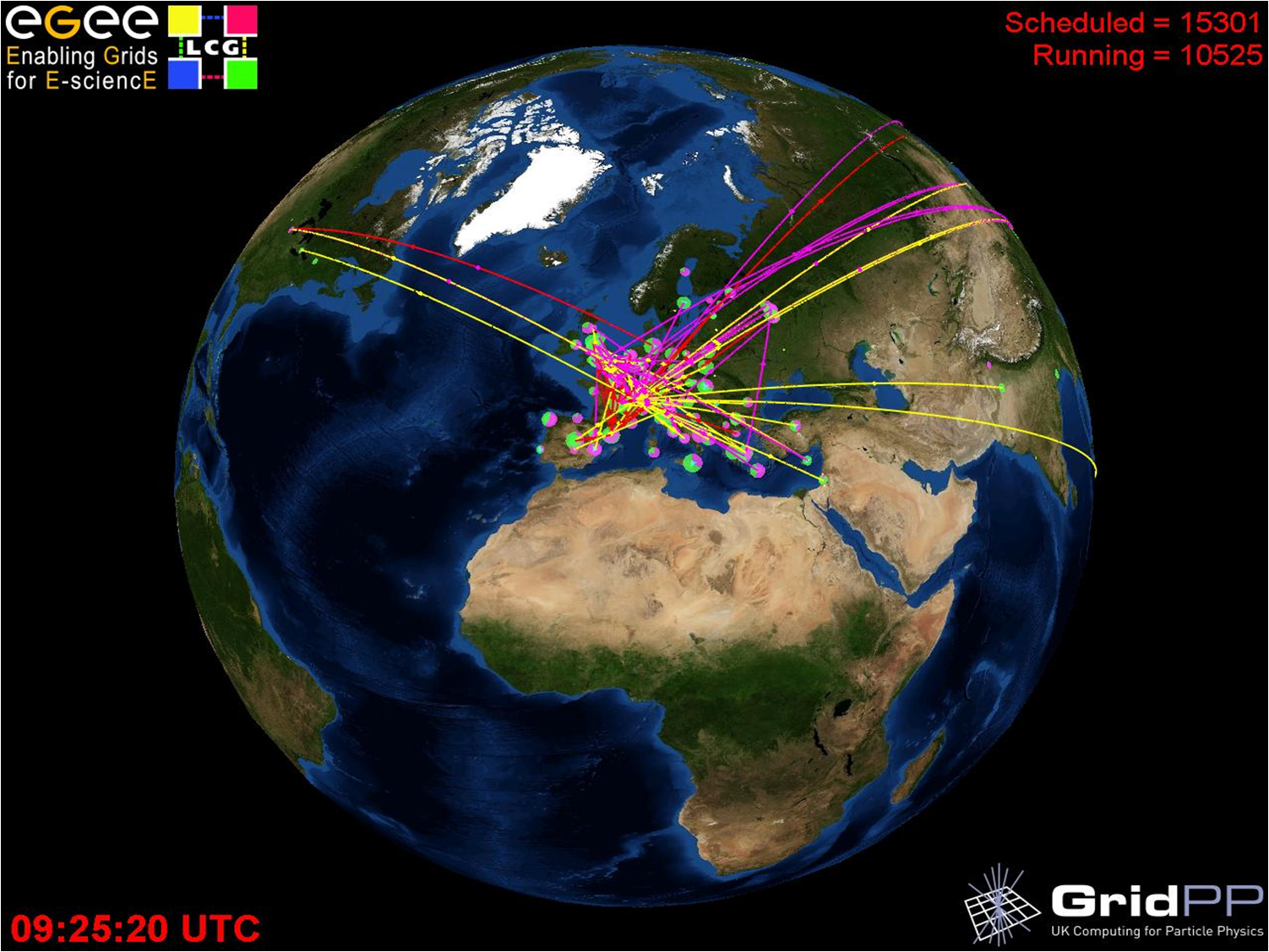
**Fig.3**: An event at 7 TeV collisions within ATLAS experiment, April 2010. This is a W candidate decaying to a muon and a neutrino

It has required titanic works of infrastructure. It is also a human and intellectual adventure, both unique in the history of science. The analysis of the enormous quantity of data recorded by the four experiments exploit a new worldwide infrastructure, the Grid, comparable in CPU power to the infrastructure of companies such as Google targeted towards an efficient transfer and storage of the data.

### The state of the art: WLCG

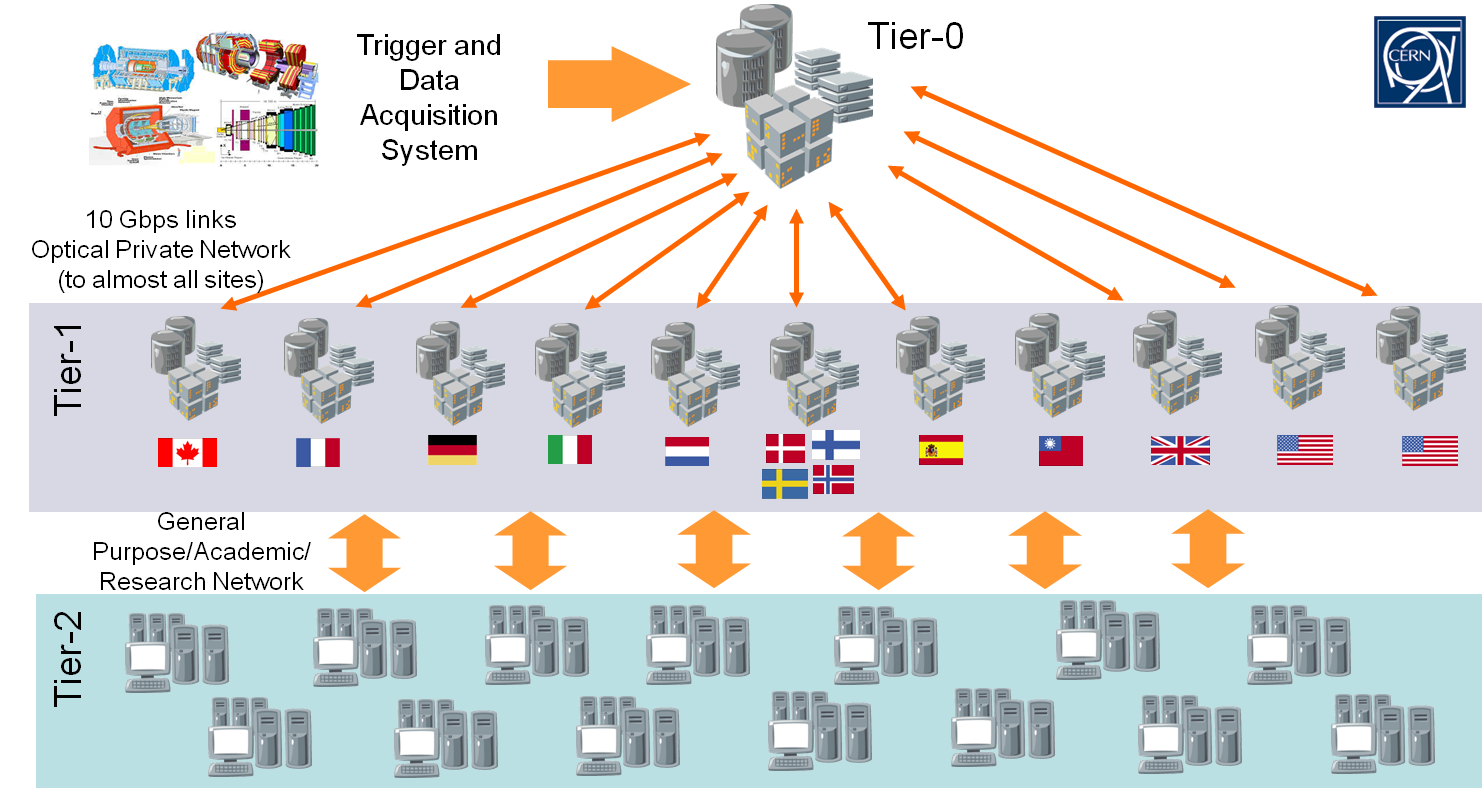
At the nominal conditions of functioning of the LHC, i.e. 14 TeV proton-proton collisions with luminosity of around 1034 particles per second per cm2, the experiments will record every year some 1010 collisions, a number comparable to the amount of stars in the Milky Way. Data handling represents a real IT challenge, in terms of data flow (of the order of Gigabytes per second) as well as of volume (several tens of Petabytes every year). At any time, thousands of researchers from any part of the world will look for computing resources and the associated storage capacity to be able to analyze the data. To face this challenge, the solution of the distributed calculation, or Grid calculation, was imperative. It was implemented by the collaboration WLCG (World wide LHC Computing Grid).

The WLCG technical Design Report [1] was issued in June 2005, describing the grid infrastructure, the common tools to be developed and the computing models for each of the four LHC experiments. The computing centres providing resources for the Worldwide LHC Computing Grid are also connected to other grids, in particular Enabling Grids for E-sciencE (EGEE) in Europe and Open Science Grid (OSG) in the United States, but also several national and regional grid structures such as GridPP in the UK, INFN Grid in Italy and NorduGrid in the Nordic countries. A Memorandum of Understanding [2] was established in October 2005 between the WLCG collaboration and the participating nations and funding agencies. This MoU guarantees the resources, the quality of services and looks 5-year forward for the computing resource planning. The quality of services includes a guarantee of operations 24 hours a day and 7 days a week with intervention to services essential to the running of a centre within a time laps of 4 hours. For any site in the first level of the hierarchy (also known as tier-1s), the target reliability is 98%. The Worldwide LHC Computing Grid combines the computing resources of more than 100,000 processors from 150 institutions in 33 countries, producing a massive distributed supercomputer that will provide more than 7000 physicists around the world with near real-time access to LHC data and the computing capability to process them. **Fig. 4** shows a snapshot of the WLCG real time monitoring where the worldwide computing centres are processing data and exchanging information.



**Fig.4**. The real time GridPP monitoring showing the WLCG computing facilities processing jobs and transferring data.

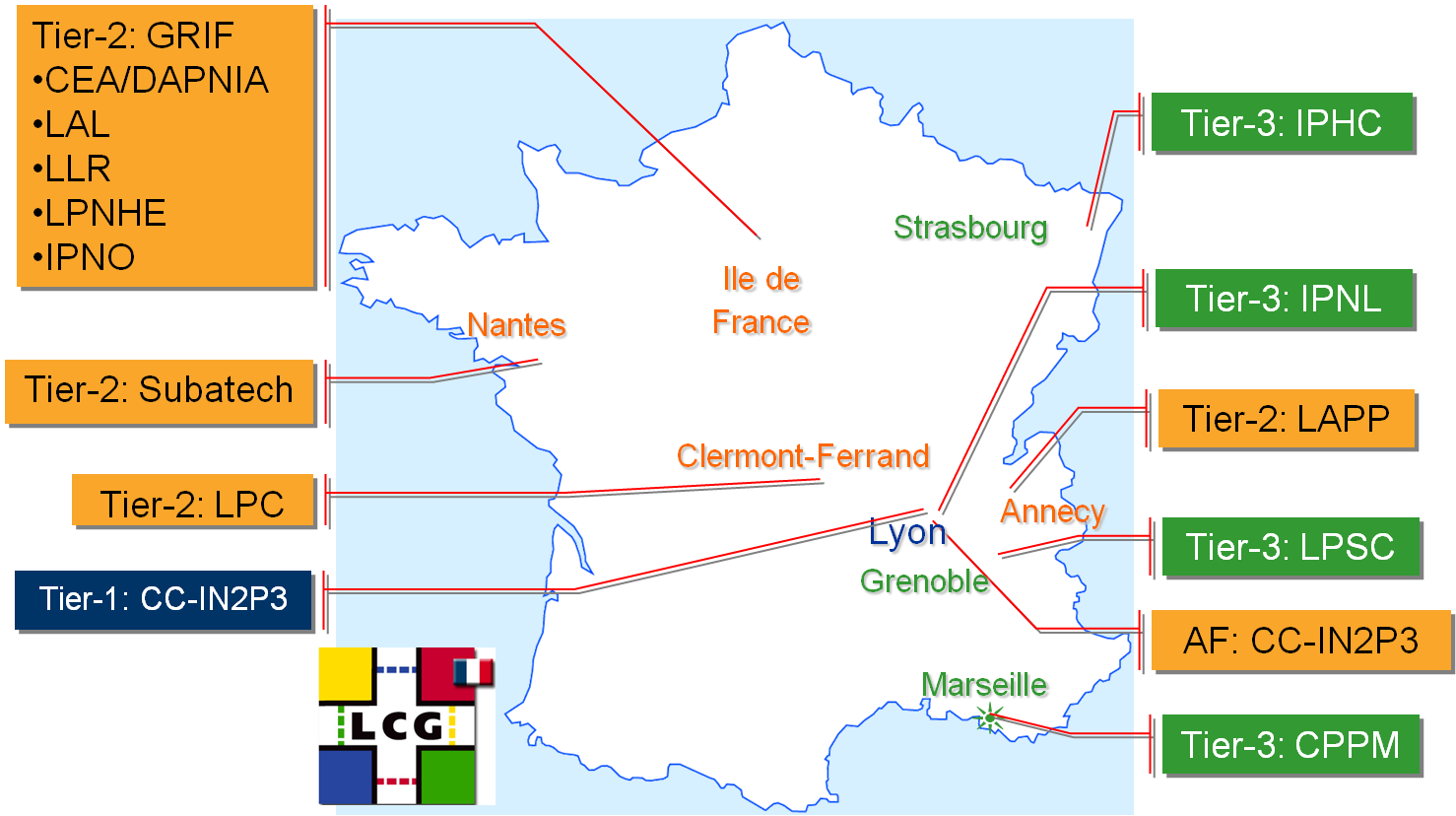
The LHC will produce around 15 petabytes (15 million gigabytes) of data every year for ten to fifteen years. This is enough to fill 3,000,000 DVDs every year. Viewing 3,000,000 DVDs would take around 500 years. If LHC data were to be burned to CD, a tower of CDs around 20 kilometres high would be created within a year. The WLCG infrastructure is based on three ‘’tiers” (Tier0, Tier1, Tier2) and 33 countries are formally involved. The CERN facility (Tier0) is linked to other major national computing centres (Tier1s) using 10 Gigabit per second optical wide area links, as well as to the general global education and research network infrastructure. The four LHC experiments computing models rely on the distributed computing facilities sketched on **Fig.5** which share the responsibilities of the needed computing services: raw data archiving, data distribution, simulated data production, data analysis, etc. More details can be found in the computing models described in the technical design reports for each of the experiments [3].



**Fig. 5**: The computing model

### The French contribution to this effort:

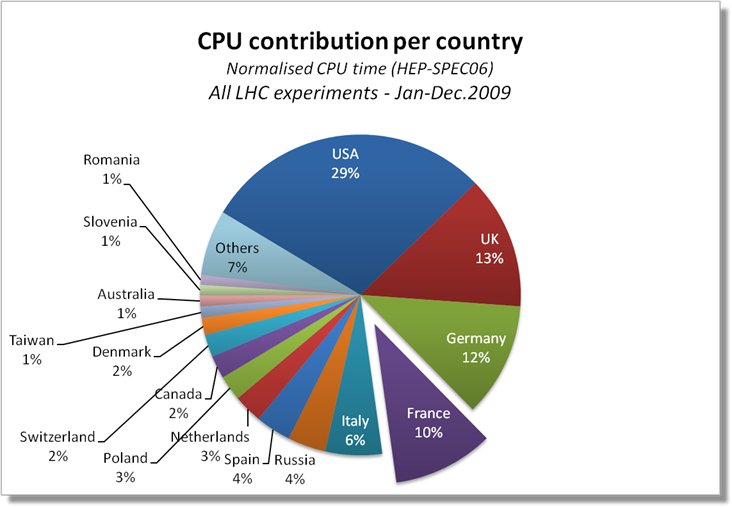
**Fig. 5** shows the French infrastructure which comprises a Tier1 located in Lyon at the IN2P3 computing centre (CC-IN2P3), and the Tier 2/Tier 3 located in Annecy, Clermont-Ferrand, Grenoble, Ile-de-France, Lyon, Marseille, Nantes, and Strasbourg. This infrastructure and the resources pledged to WLCG are managed through LCG-France project [4].



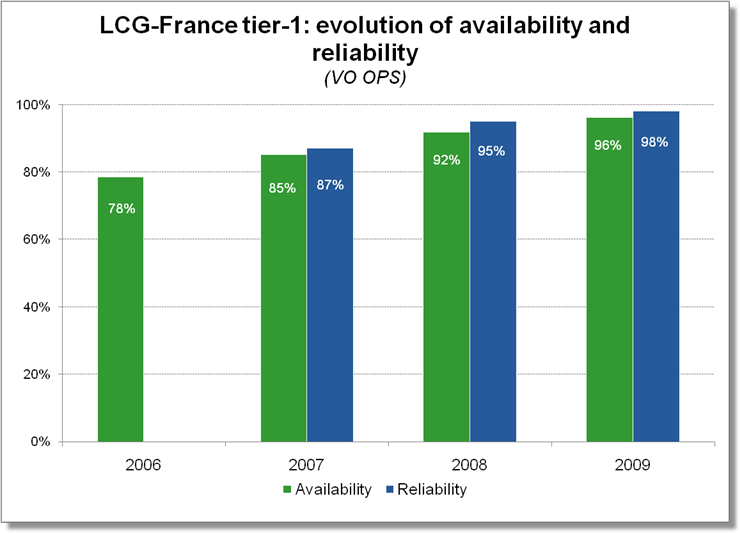
**Fig. 6** The French infrastructure operating centres in 2010

The target contribution of the French sites to the worldwide global effort is close to 10%. Fig. 7 shows the actual CPU contribution per country in 2009.

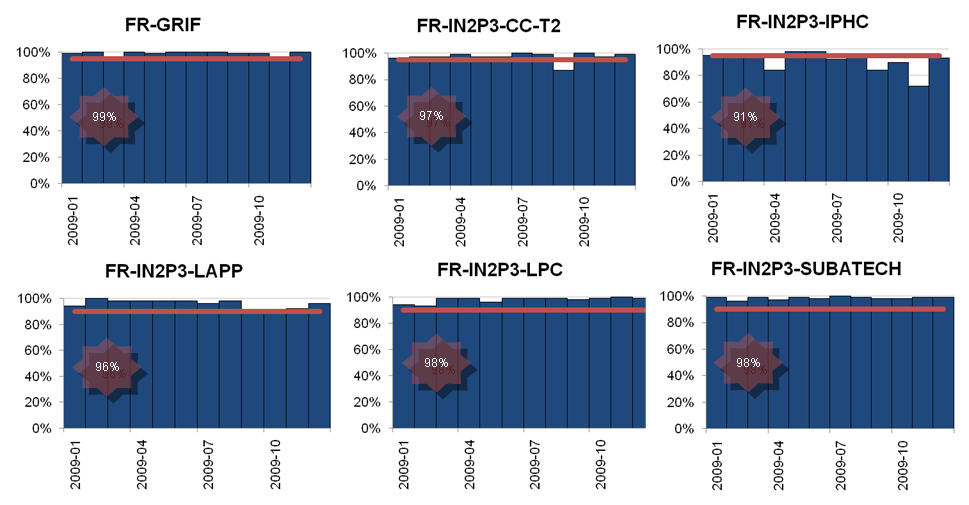
Also, to meet WLCG requirements, reliability and availability tests are performed regularly. **Fig. 8** shows the evolution of those metrics for the French Tier1 since 2006. Actually, to meet the 96% availability means for the site to stay unavailable less than 14 days over the year, including the unavoidable scheduled quarterly shutdown periods for maintenance purposes. The target reliability for the Tier2s sites is 95%. **Fig. 9** shows the monthly reliability scores measured at the French sites in 2009. It is excellent for the majority of sites.



**Fig. 7**: LCG-France CPU contribution to the worldwide symphony in 2009.

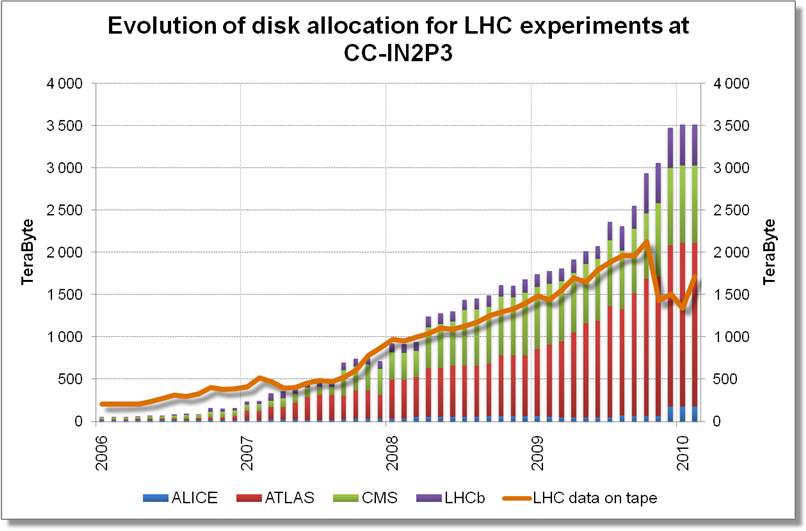


**Fig. 8**: Reliability and availability of the French Tier1 site since 2006.



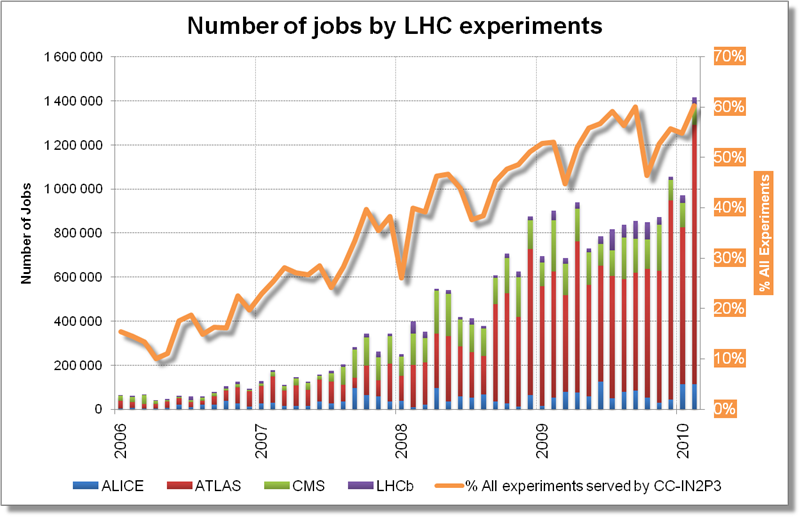
**Fig. 9**: Reliability of the French Tier2 sites since 2009.

As more data is gathered from collisions inside the LHC, there will be increasing demand upon the storage and processing capacity of the Worldwide LHC Computing Grid. These challenges will include the increase of available resources in response to planned upgrades to the LHC accelerator, as well as to the increasing data requirements of the four LHC experiments. This is increase is already observed, as sketched on **Fig. 10** for the French Tier1 site. As of today, it has reached 3.5 PBytes and is planned to double fall 2010.



**Fig. 10**: The evolution of disk allocation and storage for the LHC experiments since 2006 at the French tier-1.

Also interesting to observe is the data processing activity in France, sketched on **Fig. 11** which shows the number of jobs processed since 2006. The amount is reaching 1.5 Million jobs per month in 2010 for the four LHC experiments, the ATLAS experiment being the big consumer.



**Fig. 11**: Number of jobs processed since 2006 at the LCG-France tier-1.

Last but not least, the French infrastructure is very well performing and very much appreciated worldwide thanks to the support of the funding agencies (CNRS/IN2P3 and CEA/Irfu) and to the appropriate budget given to the sites. It has to be noticed that, not less than 5 Million Euros are necessary every year for equipment and operations of the LCG-France infrastructure, not including salaries.

### Conclusion

With the start-up of the LHC, we are about to enter the operational phase of the WLCG infrastructure. After intensive successful testing in the last few years, we will see what it delivers in real circumstances. Priority is now to make the grid more useable for the individual scientist and in particular for data analysis.

The WLCG applications support and service meet with satisfaction the requirements and the baseline services are in production. There is, as expected, a continuously increasing capacity and workload and the general site reliability is improving. It has already reached a quality level of 98% for most of the Tier1s, the target value fixed by the WLCG MoU.

Besides this, adoption of new technologies is important. Continuing to optimize the applications to fully exploit multi-core processors, where cores increase beyond two or four or eight processors to 16- or 32-core processors is essential. In the medium term, some technical work will also focus on improving the performance and scalability of access to data, especially for analysis use cases. And finally, working with limitations in terms of the cooling and power requirements of large data centres, is an ongoing issue shared by large data centres all over the world.

### Bibliography

[1] CERN-LHCC-2005-024, see also: http://lcg.web.cern.ch/LCG/tdr/.

[2] CERN-C-RRB-2005-01.

[3] ALICE-TDR-012: CERN-LHCC-2005-018;

ATLAS-TDR-017: CERN-LHCC-2005-022;

CMS-TDR-007: CERN-LHCC-2005-023;

LHCb-TDR-11: CERN-LHCC-2005-019.

[4] LCG-France project: http://lcg.in2p3.fr/

## Other user communities

Production grids have become in a few years a mandatory tool for the national research communities in several disciplines, especially subatomic physics, life sciences and planet sciences. Production grids appear very clearly as providing computing resources complementary to those offered by supercomputers. A large effort is still needed in terms of information and training as all scientific disciplines have not been exposed at the same level to this new technology.

Five thematic groups met during year 2008 to study the use of production grids in medical and life sciences, chemistry, human and social sciences, engineering and computer sciences, planet and universe sciences. For this purpose, each working group surveyed the community and more than 1000 answers were collected and analyzed. The groups expressed very similar requirements with different highlights:

* The life sciences group stressed the need to hire engineers who would act as mediators between end users and middleware developers and site administrators.
* The planet-universe working group stressed the importance for the production grid to get integrated besides supercomputers into the present ecosystem of computing resources, insisting on the necessity to have the right interfaces to the existing working environments and data centres used by the community
* The particle physics working group insisted on the necessary scaling up of the production grid resources with LHC kick-off
* The Computer and Engineering sciences working group highlighted the need to build bridges with the research community on grids and the necessity to deploy licensed software on production grids
* The chemistry working group stressed the need to create a hard core of users of the production grid to increase its adoption
* The human and social sciences working group highlighted the potential structuring role of the production grid for its community and its interest to strengthen the activity of Digital Resource Centres (Centres de Ressources Numériques)

The out come of the working groups is documented in the Production Grid White Paper published by CNRS IdG in 2009 and available online at <http://www.idgrilles.fr/IMG/pdf/livre_blanc_draft1_8FINAL.pdf> .

# International collaborations

## Collaboration with ESFRIs

CNRS INstitut des Grilles is directly involved in the European Grid Initiative. It also actively seeks to establish contacts with ESFRI design studies in order to foster the adoption of the production grids by multiple scientific disciplines. Direct contacts have been established with two ESFRI projects, ELIXIR (distributed infrastructure for biological information) and LifeWatch (biodiversity).

### EGI

CNRS IdG is involved in the European Grid Initiative on behalf of the French National Grid Initiative. The France Grille representative seating at EGI council is Michel Spiro (CNRS-IN2P3) and his alternate is Laurent Crouzet (CEA Saclay). CNRS IdG director acts as an expert beside them during EGI councils. He also represents France in several EGI council working groups dealing with voting rights and the transformation of EGI.eu into a European Research Infrastructure Consortium (ERIC).

In 2009, CNRS IdG has been actively involved in the creation of EGI.eu as a foundation established under Dutch law to create and maintain a pan-European Grid Infrastructure in collaboration with National Grid Initiatives (NGIs) in order to guarantee the long-term availability of a generic e-infrastructure for all European research communities and their international collaborators.

CNRS IdG also coordinates the involvment of the French National Grid Initiative in the EGI-Inspire European Project (see below).

### ELIXIR

The mission of ELIXIR (European Life Sciences Infrastructure For Biological Information) is to construct and operate a sustainable infrastructure for biological information in Europe to support life science research and its translation to medicine and the environment, the bio-industries and society. Through its involvement in the EMBRACE Network of Excellence (<http://www.embracegrid.info>) led by the European Bioinformatics Institute, CNRS IdG has established links with the European bioinformatics community that supports the ELIXIR project. At a national level, the bioinformatics community is structured through the RENABI network that federates 7 bioinformatics platform in France. Discussions are under way with both RENABI and EBI to understand how the French National Grid Initiative can contribute resources and services to the ELIXIR distributed infrastructure.

### LifeWatch,

LifeWatch, the e-Science and Technology Infrastructure for Biodiversity Data and Ecosystem Research, aims at constructing and bringing into operation the facilities, hardware, software and governance structures for all aspects of biodiversity research. It will consist of:  facilities for data generation and processing; a network of observatories; facilities for data integration and interoperability; virtual laboratories offering a range of analytical and modelling tools; and a Service Centre providing special services for scientific and policy users, including training and research opportunities for young scientists.

Institut des Grilles has established contacts with the LifeWatch consortium through the HealthGrid association and the EGEE project. Vincent Breton is a member of the Scientific Advisory Committee that evaluated the LifeWatch master plan in April 2010. Contacts have been established with major French actors in the field of biodiversity (CNRS Institute of Environment and Ecology, Paris Museum of Natural History) to help foster the emergence of a national LifeWatch initiative.

## European projects

Since 2001, grid and now cloud technologies have been strongly supported by the European Commission. One of the mandates of the institute is to provide an administrative support to the CNRS laboratories involved in European consortia related to grids. This role was extended to other French public and private partners through the signature of bilateral agreements and the creation of a Joint Research Unit in 2009.

European projects play an important role in the life of the institute:

* they are at the front edge for the development and deployment of these technologies
* they allow advances and development of new services that can be later deployed at a large scale
* they provide financial resources particularly useful for hiring engineers

Seven projects involving the institute are starting in 2010. A short description of them is provided in the following sections.

### DEGISCO

|  |  |
| --- | --- |
| Project name | Desktop Grids for International Scientific Collaboration |
| Project short abstract (10 lines max.) | DEGISCO |
| Project coordinator | Robert Lovas <[rlovas@sztaki.hu](mailto:rlovas@sztaki.hu)> |
| Project budget | Total Cost : 675,261 Euros  EC contribution : 799,915 Euros  (non European partners have no eligible costs) |
| Project duration | 24 months |
| Project provisional starting date | June 1st, 2010 |
| Scientific coordinator at CNRS  (name – lab - email) | Oleg Lodygensky < [lodygens@lal.in2p3.fr](mailto:lodygens@lal.in2p3.fr)>  Laboratoire de l’Accelerateur Lineaire |
| Laboratories involved | LAL; HealthGrid |
| Budget for the CNRS contribution | Total Cost : 101,400 Euros  EC contribution : 90,415 Euros |
| Main responsibilities in the project (20 lines max.) | DEGISCO aims to extend EDGeS SG/DG infrastructure by involving new non european partners. The project will encourage to deploy and interconnect new desktop grid infrastructures. It will also encourage and help the usage of the applications.  CNRS is involved in infrastructure as well as user and application supports and in dissemination activities. |
| Additional information (help needed, requests on NGI services) |  |

### EDGI

|  |  |
| --- | --- |
| Project name | European Desktop Grid Initiative |
| Project short abstract (10 lines max.) | EDGI |
| Project coordinator | Kacsuk Peter <[kacsuk@sztaki.hu](mailto:kacsuk@sztaki.hu)> |
| Project budget | Total cost : 2,615,770 Euros  EC requested : 2,300,170 Euros |
| Project duration | 24 months |
| Project provisional starting date | June 1st, 2010 |
| Scientific coordinator at CNRS  (name – lab - email) | Oleg Lodygensky < [lodygens@lal.in2p3.fr](mailto:lodygens@lal.in2p3.fr)>  Laboratoire de l’Accelerateur Lineaire |
| Laboratories involved | LAL |
| Budget for the CNRS contribution | Total cost : 242,000 Euros  EC requested : 211,300 Euros |
| Main responsibilities in the project (20 lines max.) | EDGI aims to extend EDGeS SG/DG infrastructure by extending bridges technologies to ARC and UNICORE. It will also propose a new DG QoS by interconnecting DG to clouds. LAL is involved in joint research activity on bridging as well as QoS technologies and in support activities to deploy and maintain SG-DG-cloud infrastructures as well as to help user to port their applications and use the plaftorm. |
| Additional information (help needed, requests on NGI services) |  |

### EGI-Inspire

|  |  |
| --- | --- |
| Project name | EGI-Inspire |
| Project short abstract (10 lines max.) | Federate efforts of National Grid Initiatives in Europe and beyond |
| Project coordinator | Steven Newhouse |
| Project budget | 25.016.226 € |
| Project duration | 48 months |
| Project provisional starting date | May 1st 2010 |
| Scientific coordinator at CNRS  (name – lab - email) | Vincent Breton – CNRS Institut des Grilles – vincent.breton@idgrilles.fr |
| Laboratories involved | CCPM, CC-IN2P3, CREATIS, I3S, IdG, IPSL, LAL, LAPP, LPC, LPNHE, LRI, UREC |
| Budget for the CNRS contribution | 1.641.110 € |
| Main responsibilities in the project (20 lines max.) | Responsibilities in SA1, SA3 and JRA1   * leader of task TJRA1.5 : Hélène Cordier * leader of task TSA3.4 (Life sciences): Johan Montagnat |
| Additional information (help needed, requests on NGI services) |  |

### EUMedGrid-Support

|  |  |
| --- | --- |
| Project name | EUMedGrid-Support |
| Project short abstract (10 lines max.) | Raise awareness on grids in Mediterranean countries |
| Project coordinator | Federico Ruggieri |
| Project budget | 740.000€ |
| Project duration | 24 months |
| Project starting date | January 1st 2010 |
| Scientific coordinator at CNRS  (name – lab - email) | Vincent Breton – CNRS Institut des Grilles – vincent.breton@idgrilles.fr |
| Laboratories involved | IdG, LPSC |
| Budget for the CNRS contribution | 50076 € |
| Main responsibilities in the project (20 lines max.) | Responsibilities in WP2: |
| Additional information (help needed, requests on NGI services) |  |

### GISELA

|  |  |
| --- | --- |
| Project name | GISELA - Grid Initiatives for e-Science Virtual Communities in Europe and Latin America |
| Project short abstract (10 lines max.) | The GISELA objective is to guarantee the long-term sustainability of the European-Latin American e-infrastructure and ensure the continuity and enhancement of the Virtual Research Communities (VRC) using it. GISELA will implement the Latin American Grid (LGI) in association with CLARA and collaboration with EGI and provide VRCs with the e-infrastructure and Application-related Services. GISELA will concentrate on the sustainability of the LA part of the EU-LA e-infrastructure; it will support the VRCs in providing access to multi-disciplinary VOs, user training and Workshops. |
| Project coordinator | Bernard MARECHAL |
| Project budget | 850 000 € |
| Project duration | 24 months |
| Project provisional starting date | September 1, 2010 |
| Scientific coordinator at CNRS  (name – lab - email) | Vincent BRETON |
| Laboratories involved | CPPM (UMR 6550, Clause 10), IPGP (UMR 7154), IPG (Clause 10) |
| Budget for the CNRS contribution | 47 810 € |
| Main responsibilities in the project (20 lines max.) | IPGP contributes to NA3 (Virtual Community support) in Earth Sciences and seismology in particular. gCMST (Grid Centroid Seismic Moment Tensor), a tool for the earthquake mechanism determination, and SEMUM3D, a fast 3D simulation of wave propagation (Geneviève Moguilny, Jean-Pierre Vilotte).  CPPM contributes to the JRA1 : Infrastructures and Application oriented Services for Virtual Communities. CPPM will provide support and customization of DIRAC, a Web Portal Project as a Community Grid solution in a wide range of applications in various scientific domains (Andrei Tsaregorodtsev). |
| Additional information (help needed, requests on NGI services) | The budget of 48 k€ allocated to the DIRAC project corresponds to a 14 PM recruitment in France; the remaining 10 PM are expected to be provided by our LA collaborators. It is mandatory that this contribution is initiated with a presence of at least 3 months in Marseille, under the supervision of Andrei Tsaregorodtsev. |

### SHIWA

|  |  |
| --- | --- |
| Project name | SHIWA |
| Project short abstract (10 lines max.) | The SHIWA project aims to leverage existing solutions and enable cross-workflow and inter-workflow exploitation of DCIs by applying both coarse- and fine-grained strategies. The coarse-grained approach treats workflow engines as distributed black box systems, where complete sub-workflows are sent to pre-existing enactment engines. The fine-grained approach addresses language interoperability by defining an intermediate representation to be used for translation of workflows across systems (currently selected: ASKALON, Pegasus, P-Grade, MOTEUR, Triana). SHIWA will develop, deploy and operate the SHIWA Simulation Platform to offer users production-level services supporting workflow interoperability following both approaches. A Repository will facilitate publishing and sharing of workflows, and a Portal will enable their actual enactment. Three use cases based on medical imaging applications will serve to drive and evaluate this platform from a user's perspective |
| Project coordinator | Péter Kacsuk |
| Project budget | 1.8 M€ EC funding |
| Project duration | 2 years |
| Project provisional starting date | July 1st, 2010 |
| Scientific coordinator at CNRS  (name – lab - email) | Johan Montagnat, I3S, [johan@i3s.unice.fr](mailto:johan@i3s.unice.fr) |
| Laboratories involved | I3S (DR20), CREATIS (DR7) |
| Budget for the CNRS contribution | 294.5 k€ EC funding |
| Main responsibilities in the project (20 lines max.) | MOTEUR workflow manager interoperability with other workflow managers (P-GRADE, Triana, Pegasus, Askalon).  Porting of medical image analysis workflow-based applications to multiple platforms. |
| Additional information (help needed, requests on NGI services) | The project will search for applications from different domains with need for workflow interoperability solutions. Some of the application group will be partly funded as sub-contractors and some will be supported by the project development team. |

### StratusLab

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Project name | StratusLab: Enhancing Grid Infrastructures with Virtualization and Cloud Technologies | | Project short abstract (10 lines max.) | The project is aimed at service provisioning, networking and research of cloud and virtualization technologies to simplify and optimize the use and operation of existing distributed computing infrastructures (e.g. EGEE/EGI). The project will incorporate cloud and virtualization innovation into existing Grid infrastructures by integrating cloud technologies and services within Grid sites; and by enriching existing computing infrastructures with “Infrastructure as a Service” (IaaS) cloud-like delivery paradigms. | | Project coordinator | Charles LOOMIS | | Project budget | Total cost: 3,317,233 Euros  EC contribution: 2,299,575 Euros | | Project duration | 24 months | | Project provisional starting date | 1 June 2010 | | Scientific coordinator at CNRS  (name – lab - email) | Charles LOOMIS — LAL — [loomis@lal.in2p3.fr](mailto:loomis@lal.in2p3.fr)) | | Laboratories involved | LAL (Orsay), IBCP (Lyon) | | Budget for the CNRS contribution | Total cost: 623,039 Euros  EC contribution: 469,618 Euros | | Main responsibilities in the project (20 lines max.) | LAL coordinate the project and will lead the WP2 activity which will manage contacts between the project and its targeted communities (system administrators and scientists), contribute to the project dissemination, and participate in WP5 by providing one of the two production grid sites running the StratusLab distribution.  IBCP will provide input on features useful for the Bioinformatics community, and will also act as a StratusLab site and extensively test the software at the different stages of the project, before they are being made available to the public. IBCP will also contribute to dissemination activities. | | Additional information (help needed, requests on NGI services) | In order to reduce the budget to the requested level and keep the foreseen personnel, significant cuts needed to be made in the travel budgets. Additional travel funds from the NGI would be helpful in order to carry out the goals of this project and ensure that it is successful.  Sites within the French NGI are significantly involved in the project both in terms of deployment and utilization. While we expect these contacts to be primarily direct contacts between the project and the sites/users, some formal coordination at the French NGI level may be helpful in reaching the full French user community. In addition, if the StratusLab distribution is successful (and we expect it to be), then this may be a useful common base for site configuration in France. | |

## Bilateral collaborations

Institut des Grilles has been actively pursuing the development of collaborations outside Europe. A particular effort is made to support the adoption of grids in developing countries because they are identified as a powerful tool to address the digital divide. Indeed, all users of a virtual organization enjoy the same services.

### Bilateral collaborations with Africa and Middle-East

The Institute was involved in the organization of grid tutorials in Senegal and South Africa in 2009. Engineers helped install grid services on local clusters in both countries. In South Africa, the WISDOM production environment for in silico drug discovery was installed on a local cluster.

In 2010, within the framework of EUMedGrid-Support and the EPIKH Marie Curie network, participation to grid tutorials in Congo Democratic republic and Tunisia is foreseen while the first grid event (tutorial – workshop) in Lebanon is under preparation for May – June 2011;

### Bilateral collaborations with Asia

Collaboration with Asia takes place mainly within the framework of International Associated Laboratories (IAL). IALs are "laboratories without walls" that bring together laboratories from CNRS and from one other country. These laboratories contribute human and material resources to a common, jointly-defined project designed to "add value" to their individual pursuits. An LIA agreement is for 4 years, renewable twice.

CNRS-IN2P3 has set up four IALs with asiatic countries : France Japan Particle Physics Laboratory (FJPPL), France China Particle Physics Laboratory (FCPPL), France Korea Particle Physics Laboratory (FKPPL) and France Vietnam Particle Physics Laboratory. Three of the IALs are running projects in the field of e-science :

* FCPPL has one grid project on High Energy Physics with Beijing IHEP. Attempts to develop projects in the field of life sciences
* FKPPL has three projects between french and korean institutes related to grids. A deidcated FKPPL virtual organization was created in 2009 which was extremely successful for hosting korean users. Collabration between french and korean laboratories in the fiel dof grid-enabled drug discovery led to 5 aptents and significant scientific production
* FVPPL is not yet formally created, but already two grid schools were organized in 2007 and 2009 in Hanoi, leading to the installation of grid services in several institutes in Hanoi. Projects dealing with the moniroting of emerging diseases and the search for new drugs out of local biodiversity are under way.

1. <http://www.rfc-editor.org/internet-drafts/draft-dzis-nwg-nttdm-02.txt> [↑](#footnote-ref-0)
2. <http://egee-technical.web.cern.ch/egee-technical/EGEE-III/FirstReview-09/index.html> [↑](#footnote-ref-1)
3. <http://www.terena.org/news/fullstory.php?news_id=2484> [↑](#footnote-ref-2)