



CONTRACT NUMBER RI-222919

DEISA 2
DISTRIBUTED EUROPEAN INFRASTRUCTURE FOR
SUPERCOMPUTING APPLICATIONS

European Community Seventh Framework Programme
RESEARCH INFRASTRUCTURES
Integrated Infrastructure Initiative

Report on Integration of First Associate Partner

Deliverable ID: DEISA2-D3-2.1
Due date: April 30th, 2009
Author: Xavier Delaruelle, CEA

Project start date: May 1st, 2008
Duration: 3 years

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2011)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Content

Table of Content.....	1
List of Figures	1
List of Tables.....	1
1 Introduction	2
1.1 Executive Summary.....	2
1.2 References and Applicable Documents.....	2
1.3 Document Amendment Procedure.....	3
1.4 List of Acronyms and Abbreviations.....	3
2 Integration process of CEA	4
2.1 Integration of CEA OCRE computing environment	4
2.2 Helping CEA with its integration	4
3 Technical integration of CEA.....	5
3.1 Networking	5
3.1.1 Connection to the DEISA network.....	5
3.1.2 Access for DEISA services on public Internet	6
3.1.3 Network monitoring	6
3.2 Data services.....	6
3.2.1 MC-GPFS	6
3.2.2 GridFTP and RFT	7
3.3 Middleware.....	7
3.3.1 UNICORE	7
3.3.2 MC-LL.....	8
3.3.3 Globus toolkit services	8
3.4 Authentication, Authorization and Accounting.....	8
3.4.1 DEISA user registration.....	8
3.4.2 Accounting repository	9
3.5 User.....	9
3.5.1 DEISA Modules Environment.....	9
3.5.2 INCA monitoring.....	9
4 Conclusion	10

List of Figures

Figure 1 – CEA network layout	6
-------------------------------------	---

List of Tables

Table 1 – Properties of the DEISA users at CEA	8
Table 2 – Properties of the DEISA LDAP host at CEA.....	8
Table 3 – Overview of DEISA services integration at CEA.....	10

1 Introduction

1.1 Executive Summary

DEISA2 targets the integration of Associate Partners into the DEISA environment. It is currently planned to offer this service to CEA, the French Atomic Commission, to the Swiss partner ETH-CSCS, to the Swedish partner KTH/PDC. In addition the Computing Centre of the Russian Academy of Science (JSCC) will also be given the status of an Associate Partner. WP3 has a dedicated subtask (Task 6, Integration of Associate Partners) to contribute to this integration process.

This deliverable reports the work that has been achieved to integrate CEA into the DEISA infrastructure since September 2008. It first details the general strategy that has been defined for this integration process and then addresses the technical aspects of the integration: Network, Data, Middleware, AAA and User related services.

The integration process defines an initial evaluation phase of the different DEISA components to check that they do no conflict with the CEA local technical and security policy. CEA is assisted during all the integration process by IDRIS.

The Network integration is already in a well advanced stage. CEA is connected to the DEISA network and can reach and be reached by all the partners. Focus is now put on allowing access for DEISA services to public Internet. This integration step is not trivial since this kind of access basically conflicts with CEA security policies.

Concerning the data management services, CEA has planned to mount a remote GPFS (hosted by another partner) on its login nodes and to deploy a public GridFTP and RFT services.

CEA has also been actively involved in the middleware evaluations, especially for UNICORE 6 with a direct implication in task T4.3a. Its two PRACE prototype systems are already part of the UNICORE 6 test bed. A high priority will also be put on the deployment of GSI-SSH.

All the technical details regarding the AAA layer have been discussed and agreed with WP3. The LDAP integration has started and is in a well advanced stage. The accounting service is planned for July 2009.

Finally, for the User environment layer, the DEISA Modules environment and the INCA monitoring framework will be deployed after the data and middleware layers (June 2009 for INCA and November 2009 for Modules).

1.2 References and Applicable Documents

- [1] Description of Work (Annex I of the Grant Agreement)
- [2] DEISA web pages: <http://www.deisa.eu>
- [3] Deliverable DEISA2-D1.1: Initial Report on Management
- [4] Lustre web pages: <http://www.lustre.org>
- [5] SLURM web pages: <http://www.llnl.gov/linux/slurm/>

1.3 Document Amendment Procedure

This document is prepared according to the guidelines defined by the management of DEISA2. These rules can be found in section 2.7 of the deliverable DEISA2-D1.1 [3].

1.4 List of Acronyms and Abbreviations

AAA	Authentication, Authorisation, Accounting
BSC	Barcelona Supercomputing Center
CEA	Commissariat à l'Énergie Atomique
CCC	CEA Computing Complex
DEISA	Distributed European Infrastructure for Supercomputing Applications
DoW	Description of Work
FTP	File Transfer Protocol
FZJ	Forschungszentrum Jülich
GÉANT2	European backbone Network connecting national research networks
GPFS	General Parallel File System
GPCPU	General Purpose Graphics Processing Unit
GSI-SSH	Grid Security Infrastructure – Secure Shell
HPC	High Performance Computing
IDRIS	Institut du Développement et des Ressources en Informatique Scientifique
INCA	User Level Grid Monitoring
LDAP	Lightweight Directory Access Protocol
LRZ	Leibniz Rechenzentrum
MC-GPFS	Multi Cluster General Parallel File System
MC-LL	Open Computing Research Environment Multi Cluster Load Leveler
OCRE	Open Computing Research Environment
OGF	Open Grid Forum
OGF-UR	Open Grid Forum - Usage Record
PRACE	Partnership for Advanced Computing in Europe
RENATER	Réseau National de télécommunications pour la Technologie l'Enseignement et la Recherche
RFT	Reliable File Transfer
RZG	Rechenzentrum Garching
SLURM	Simple Linux Utility for Resource Management
UNICORE	Uniform Interface to Computing Resources
XNJS	eXtended Network Job Supervisor
WAN	Wide Area Network

2 Integration process of CEA

Integration of CEA started in September 2008 with a first step of evaluation of the whole DEISA Services. The integration of CEA is supported by IDRIS, which acts as the privileged contact providing help.

2.1 Integration of CEA OCRE computing environment

Before integrating a production-level computing center, CEA wishes to technically evaluate DEISA components to verify if they fit into its technical architecture and security model.

CEA integration will first consist in the evaluation of the DEISA Services. For that purpose CEA will integrate its Open Computing Research Environment (OCRE), dedicated to computing research and development. OCRE is the CEA testbed for assessing new HPC technologies for its next-generation supercomputers.

During this evaluation period, there will be no computing cycle available for DEISA users. PRACE prototypes of CEA, the Intel Nehalem/Xeon InfiniBand cluster and the NVIDIA Tesla GPGPU cluster, will be integrated in the CEA testbed for DEISA, but only to assess behaviours between Tier-0 and Tier-1 machines.

2.2 Helping CEA with its integration

During the first months of its integration, CEA was introduced to the DEISA world by RZG. Authorizations to the DEISA BSCW and subscription to DEISA technical mailing lists were made. Thanks to this, CEA people were able to access the DEISA deliverables and were able to understand the project internals.

For a smoother integration, DEISA Operation team nominated a DEISA site to be the privileged contact for technical questions and other kind of help. IDRIS was chosen to provide this help.

The IDRIS DEISA staff introduces the DEISA technical architecture to the CEA staff. When specific questions were raised by CEA, IDRIS provided their knowledge of the DEISA teams to forward these questions to the relevant DEISA partners. IDRIS provided CEA staff with DEISA staff accounts to give them access to DEISA internal workspaces (wiki and BSCW), which contains helpful documentation about technical procedures and software installations.

This help was very fruitful, since it allows CEA to quickly achieve a general understanding of the DEISA platform. And thanks to this understanding, technical integration is going smoothly.

3 Technical integration of CEA

Since joining DEISA as associate partner in September 2008, actions have been taken at CEA to integrate DEISA. Well progressed topics are the connection to the GÉANT2 network and the integration to the UNICORE 6 testbed.

3.1 Networking

Connection to the DEISA network was the first integration step achieved by CEA. The access for DEISA services on the public Internet is currently ongoing. When this is completed, the integration of the network monitoring services will start.

3.1.1 Connection to the DEISA network

The CEA 10 Gigabit Ethernet network was connected to the DEISA network in October 2008, using RENATER infrastructure. IDRIS and CEA are sharing a 10 Gigabit Ethernet link between the RENATER node in Orsay and GÉANT2. No 10 Gigabit fiber is available on the RENATER infrastructure to build a dedicated link for both partners.

The CEA router has been connected to the DEISA central switch-router in Frankfurt since November 2008. The IP network address between those two routers is 10.31.1.128/30.

CEA has been interconnected with each DEISA site since the beginning of 2009. The CEA public IP range dedicated to DEISA is 132.167.248.0/22. This IP range is reachable from other partner sites and the partner IP ranges are reachable from the CEA machines. The DEISA network connectivity monitoring tool, installed at FZJ, monitors CEA connectivity to the host with address 132.167.248.1. The network infrastructure between CEA and Frankfurt is shown in Figure 1.

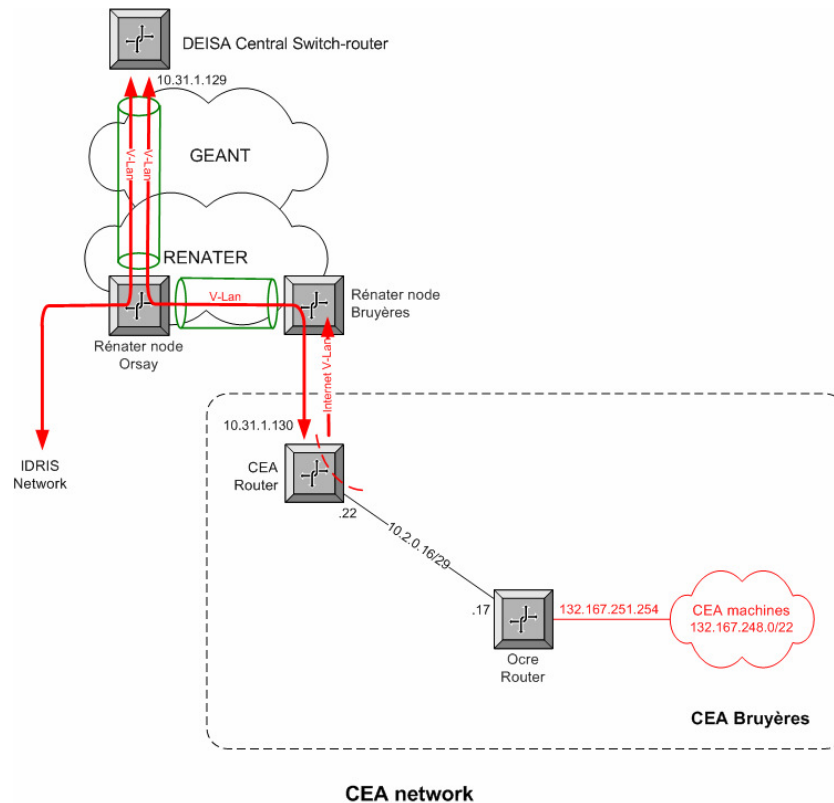


Figure 1 – CEA network layout

3.1.2 Access for DEISA services on public Internet

Giving access to DEISA services on the public Internet is currently ongoing at CEA. This kind of public access does not fit the CEA security model. So having these facilities available implies that the DEISA architecture must be explained to the CEA security officers and that a risk analysis must be built for each DEISA service which needs this access.

CEA has now precisely identified the network flow of each DEISA service. The risk analysis will now be conducted. Internet access for accepted services will be available in May 2009.

3.1.3 Network monitoring

Integration of monitoring services will start when access for DEISA services on public Internet is available. This service will be available in June 2009.

CEA intends to join the network monitoring testbed based on perfSonar, which is currently evaluated by WP4. Once the current DEISA monitoring client is available, this task will be started.

3.2 Data services

Integration of data services at CEA will start with the deployment of GridFTP and RFT. Then the deployment of GPFS clients will follow.

3.2.1 MC-GPFS

CEA does not have a GPFS server infrastructure available. Instead, parallel file systems are provided to users by Lustre [4] at CEA.

Thus, CEA will rely on another DEISA site to host the MC-GPFS DEISA mount points relative to its users. It is planned to have the MC-GPFS mount points available on the login node of each CEA prototype:

```
inti.ocre.cea.fr  
uchu.ocre.cea.fr
```

CEA is currently working to obtain a MC-GPFS licence. The MC-GPFS client deployment will be ready for October 2009.

3.2.2 GridFTP and RFT

The GridFTP service, for data transfers between DEISA sites using the DEISA network, will be available at CEA in June 2009. CEA is willing to follow the “public GridFTP” proposal currently drafted in DEISA to deploy this service, in order to have its GridFTP setup for the DEISA network not conflicting with the upcoming “public GridFTP” recommendation.

As suggested by DEISA, at least two GridFTP backend nodes will be deployed to ensure good performance. CEA plans to serve the GPFS DEISA mount points but only those used by CEA users to avoid exposing the GPFS servers of another DEISA site to any potential vulnerability of our GridFTP deployment.

CEA will also configure its GridFTP service to serve its Lustre file systems. This will enable that CEA users can stage their data to a GPFS file system hosted elsewhere for CEA if there is no GPFS client available on the CEA systems.

The installation of the RFT service, working on top of GridFTP to ensure users of the reliability of the data transfers, will follow the GridFTP deployment at CEA and should be ready for July 2009.

3.3 Middleware

Up to now middleware integration focused on UNICORE 6 at CEA. Now that the UNICORE integration is well advanced, CEA will continue by deploying GSI-SSH.

3.3.1 UNICORE

DEISA currently proposes to its users the UNICORE middleware version 5. An evaluation task of the version 6 of UNICORE is also in progress. Since there is a major gap between the two versions and since UNICORE 5 will soon be considered as deprecated, CEA has directly invested in UNICORE 6.

Since January 2009, CEA is taking part in the T4.3a activity and has joined the UNICORE 6 evaluation testbed. The UNICORE Gateway is currently restricted to the use of the DEISA task members but once the technology is approved by the CEA security officer, the Gateway will go public.

Two resources are available through this UNICORE testbed at CEA. These two correspond to the PRACE prototype machines hosted in the OCRE environment at CEA:

```
https://gruget.ocre.cea.fr/CEA_INTI/  
https://gruget.ocre.cea.fr/CEA_UCHU/
```

CEA works currently to make UNICORE 6 fit its environment well. Some features have been asked to the UNICORE developers, like the support of web proxy on UNICORE clients. Moreover CEA will do some development work, like interfacing the UNICORE XNJS component with the LDAP server for user authorization.

3.3.2 MC-LL

CEA does not plan to integrate the Multi Cluster LoadLeveler platform since the LoadLeveler batch scheduler is not in use at CEA.

3.3.3 Globus toolkit services

Concerning the integration of the services from the Globus toolkit, CEA will give the deployment of GSI-SSH high priority. This software facilitates user access by using user's grid certificate as a credential instead of a password. Thus CEA will make it available in June 2009. Then registered users will be able to connect with their grid certificate to the login nodes of each CEA prototype from the Internet by using the following addresses:

```
inti.ocre.cea.fr
uchu.ocre.cea.fr
```

The deployment of WS-GRAM will be given low priority, since this software is optional in DEISA for the moment and CEA prefers to focus on UNICORE 6.

3.4 Authentication, Authorization and Accounting

The DEISA user registration is mandatory to provide access to the DEISA users. So it will be one of the first DEISA services integrated at CEA. The accounting repository deployment will then follow.

3.4.1 DEISA user registration

The DEISA user registration service mainly consists in the availability of the DEISA LDAP infrastructure at the site. Integration of this LDAP service is well-advanced at CEA and will be ready before May 2009.

CEA has obtained a compatible UID/GID number range for its DEISA users. The whole DEISA UID/GID range does not conflict with ranges employed locally at CEA. A three-letter acronym has also been chosen and is "CCC", which stands for CEA Computing Complex.

Site	Acronym	First UID/GID number	Last UID/GID number
CEA	ccc	1200000	1299999

Table 1 – Properties of the DEISA users at CEA

The LDAP host at CEA is up and running and is currently being integrated among the DEISA hosts of the other partners.

Site	Host	Base
CEA	diron.ocre.cea.fr	ou=cea.fr,ou=ua,dc=deisa,dc=org

Table 2 – Properties of the DEISA LDAP host at CEA

3.4.2 Accounting repository

Once the DEISA user registration is available, CEA will start the integration of the accounting repository service. It will first be available to DEISA partners only for July 2009. This service will then be available to users once approved by the CEA security officer.

The accounting repository service relies on the accounting information extracted from the local batch scheduler and transformed into the Usage Record format, standardized by the OGF. Batch schedulers CEA uses, like SLURM [5], do not natively provide this accounting description format. Some developments were initiated few months ago at CEA to have a tool that takes SLURM accounting information and transforms it into the OGF-UR format. This tool called URGE will be ready for June 2009.

Contact has also been made with BSC, which also uses the SLURM batch scheduler, to learn how they provide the accounting data from SLURM in the repository.

3.5 User

The integration of the user aspects will be started once the basic infrastructure services are well-advanced.

3.5.1 DEISA Modules Environment

The integration of the DEISA Modules environment will be started once the data and middleware services will be ready. It will be available before November 2009.

CEA is willing to implement the complete DEISA Modules environment and merge it with its own. CEA will investigate technical solutions used by DEISA partners that have performed this environment merge, like LRZ.

3.5.2 INCA monitoring

The INCA monitoring, which supervises DEISA services, will help CEA to validate the integration of the different DEISA services. INCA monitoring will be available before June 2009.

The INCA monitoring could be performed following two different manners: one by delegating its execution on the site to LRZ, the other by installing the INCA reporter manager at the site. CEA will implement the latter solution.

4 Conclusion

Integration of DEISA services is going smoothly at CEA and will be finished on its evaluation testbed at the end of 2009.

From September 2008 to February 2009, most time was devoted to the understanding of the DEISA infrastructure. This task was unexpectedly tough since no global overview of the technical infrastructure was available.

At that time, CEA especially misunderstood the use and the purpose of the DEISA network and thought it was used by most of DEISA services, whereas it is dedicated to data services. Therefore CEA's first design of its DEISA testbed did not include Internet access for services which was not valid.

Having a privileged contact for its technical questions helped CEA and is still helping. Thanks to this a lot of technical clarifications were achieved in a short period of time.

Now DEISA infrastructure understanding is achieved and CEA is currently deploying the services. The following table summarizes the availability of each DEISA service.

Service	Availability
Networking	
Connection to the DEISA network	Available
Access for DEISA services on public Internet	May 2009
Network monitoring	June 2009
Data services	
MC-GPFS	October 2009
GridFTP	June 2009
RFT	July 2009
Middleware	
UNICORE	UNICORE 6 available
MC-LL	Not to be installed
GSI-SSH	June 2009
WS-GRAM	Not planned for the moment
Authentication, Authorization, Accounting	
DEISA user registration	May 2009
Accounting repository	July 2009
User	
DEISA Modules Environment	November 2009
INCA monitoring	June 2009

Table 3 – Overview of DEISA services integration at CEA