



CONTRACT NUMBER 508830

DEISA
**DISTRIBUTED EUROPEAN INFRASTRUCTURE FOR
SUPERCOMPUTING APPLICATIONS**

European Community Sixth Framework Programme
RESEARCH INFRASTRUCTURES
Integrated Infrastructure Initiative

**Operation of Resource Management in the DEISA
Infrastructure**

Deliverable ID: DEISA-SA3-2B
Due date: October, 31st, 2005
Actual delivery date: November 25, 2005
Lead contractor for this deliverable: CINECA, Italy

Project start date: May 1st, 2004
Duration: 4 years

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	X
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)	

Document Keywords and Abstract

Keywords:	DEISA, UNICORE, RMIS, HPC, Grid
Abstract:	<p>During the first eighteen month the Service Activity 3 (SA3) 'Resource Management' achieved the following results:</p> <ol style="list-style-type: none">1. The DEISA UNICORE infrastructure is up and running at the core sites and most other sites.2. The DEISA RMIS infrastructure is up and running for the core sites; some of the other sites still need to complete the configuration process. <p>This meets the objectives of Task 2 in SA3 as described in the Technical Annex [4].</p> <p>This document explains how the DEISA consortium reached the above mentioned project objectives. Considering that in the meanwhile three new sites joined the project, the actual status represents a big step ahead for the DEISA infrastructure.</p>

--

Table of Contents

Table of Contents.....	3
List of Figures.....	4
List of Tables.....	4
1. Introduction.....	5
1.1 Executive Summary.....	5
1.2 Document content	5
1.3 Document structure	5
1.4 References and Applicable Documents	6
1.5 Document Amendment Procedure	6
1.6 List of Acronyms and Abbreviations	6
2. The DEISA UNICORE infrastructure: status report.....	8
2.1 Introduction.....	8
2.2 UNICORE components	8
2.3 Installed UNICORE components at the DEISA sites.....	9
2.4 DEISA UNICORE infrastructure configuration.....	9
2.5 DEISA UNICORE infrastructure installation status.....	10
2.5.1 FZJ	12
2.5.2 RZG.....	13
2.5.3 IDRIS.....	14
2.5.4 CINECA.....	15
2.5.5 SARA.....	16
2.5.6 CSC.....	17
2.5.7 BSC.....	18
2.5.8 LRZ.....	19
2.5.9 ECMWF.....	20
2.5.10 EPCC.....	21
2.5.11 HLRS.....	22
3. The DEISA RMIS infrastructure: status report	23
3.1 Introduction.....	23
3.2 RMIS components.....	23
3.3 Installed RMIS components at the DEISA sites	23
3.4 DEISA RMIS infrastructure configuration	24
3.5 DEISA RMIS infrastructure installation status	24

List of Figures

Figure 1 - Actual DEISA UNICORE configuration.....	9
Figure 2 - UNICORE configuration at FZJ	12
Figure 3 - UNICORE configuration at RZG	13
Figure 4 - UNICORE configuration at IDRIS.....	14
Figure 5 - UNICORE configuration at CINECA.....	15
Figure 6 - UNICORE configuration at SARA.....	16
Figure 7 - UNICORE configuration at CSC.....	17
Figure 8 - UNICORE configuration at BSC	18
Figure 9 - UNICORE configuration at LRZ.....	19
Figure 10 - UNICORE configuration at ECMWF	20
Figure 11 - UNICORE configuration at EPCC.....	21
Figure 12 - RMIS configuration	24

List of Tables

Table 1 - List of acronyms.....	7
Table 2 - UNICORE configuration details at FZJ	12
Table 3 - UNICORE configuration details at RZG.....	13
Table 4 - UNICORE configuration details at IDRIS.....	14
Table 5 - UNICORE configuration details at CINECA.....	15
Table 6 - UNICORE configuration details at SARA.....	16
Table 7 - UNICORE configuration details at CSC.....	17
Table 8 - UNICORE configuration details at BSC	18
Table 9 - UNICORE configuration details at LRZ.....	19
Table 10 - UNICORE configuration details at EPCC	21

1. Introduction

1.1 Executive Summary

During the first eighteen month the Service Activity 3 (SA3) 'Resource Management' achieved the following results:

1. The DEISA UNICORE infrastructure is up and running at the core sites and most other sites.
2. The DEISA RMIS infrastructure is up and running for the core sites; some of the other sites still need to complete the configuration process.

This meets the objectives of Task 2 in SA3 as described in the Technical Annex [4].

This document explains how the DEISA consortium reached the above mentioned project objectives. Considering that in the meanwhile three new sites joined the project, the actual status represents a big step ahead for the DEISA infrastructure.

Intended audience are DEISA partners, EU officers and future partners that will find clear guidelines to adopt for joining the DEISA infrastructure.

1.2 Document content

SA3 is composed of three main layers:

- *Basic services* are those located closest to the operating system of the computing platforms and enable the operation of a single or a multiple cluster through local or extended batch schedulers and other cluster-like features; This is the basic system software needed for the initial operation of the core distributed platform.
- *Intermediate services* are the first-level Grid services that allow access to an enlarged GRID empowered infrastructure, dealing with resource and network monitoring and information systems;
- *Advanced services* use the previous layers to implement the global management of the distributed resources of the infrastructure.

Up to now SA3 has produced three deliverables: D-SA3-1A and D-SA3-1B are concerned with the operation of the DEISA homogeneous infrastructure, while D-SA3-2A covers the choice of the middleware useful for the infrastructure. In brief, D-SA3-1A and D-SA3-1B were addressing the *basic services layer* as defined in [4] and D-SA3-2A is concerned with the *intermediate services layer*.

This deliverable has been planned as the logical continuation of D-SA3-2A, so it will address only the DEISA UNICORE infrastructure status and the DEISA RMIS status (described in [3]).

1.3 Document structure

This section provides an overview of the structure and the intended audience of the document, plus references and acronyms. The second section describes the DEISA UNICORE infrastructure status. The third section illustrates the DEISA RMIS infrastructure status.

1.4 References and Applicable Documents

- [1] <http://unicore.sourceforge.net/>
- [2] <http://winnetou.matrix.sara.nl/deisa/certs/index.html>
- [3] DEISA deliverable D-SA3-2A <http://work.deisa.org>
- [4] DEISA technical annex Section 5.B.4 <http://work.deisa.org>
- [5] Ganglia: <http://ganglia.sourceforge.net>

1.5 Document Amendment Procedure

The initial document amendment procedure is via communication between members of DEISA SA3 team. The document is then submitted for review to the DEISA Executive and an Executive appointed DEISA reviewer. The document is then amended according to comments received from the Executive and the DEISA appointed reviewer. It is subsequently re-submitted to the DEISA Executive for submission to the EU.

1.6 List of Acronyms and Abbreviations

Term/Acronym	Definition
AJO	Abstract Job Object
BSCW	Basic Support for Cooperative Work
CA	Certificate Authority
CRL	Certification Revocation List
DMZ	Demilitarised Zone
GUI	Graphical User Interface
IDB	Incarnation Data Base
LL	IBM LoadLeveler
LL-MC	IBM LoadLeveler Multi Cluster
LSF	Platform Load Scheduler Facility
NJS	Network Job Supervisor
PBS	Altair Portable Batch System
RMIS	Resource Management Information System
SLURM	Simple Linux Utility for Resource Management
TSI	Target System Interface
Uspace	UNICORE space
UADB	UNICORE User Data Base

Vsite	(UNICORE) Virtual site
XML	eXtensible Markup Language

Table 1 - List of acronyms

2. The DEISA UNICORE infrastructure: status report

2.1 Introduction

Among the DEISA partners, FZJ coordinates the management of the DEISA UNICORE infrastructure and gives UNICORE administrators installation and configuration support.

UNICORE [1] provides a seamless interface for preparing and submitting jobs to a wide variety of heterogeneous computing resources.

In a first step in the DEISA project a homogeneous supercomputer cluster was built. This homogenous cluster consisted of IBM AIX PowerPC systems located at FZJ, RZG, CINECA and IDRIS. The unified access to this cluster with together almost 24 Tflops was realized via UNICORE and demonstrated successfully at the end of 2004 at the IST conference.

In a second step (project month 13) the DEISA supercomputer cluster became heterogeneous with the introduction of the SGI ALTIX Linux cluster located at SARA, the IBM PowerPC Linux system at BSC, and the Linux cluster from LRZ which will change to an SGI ALTIX Linux cluster in 2006. The unified access to this enhanced heterogeneous cluster is also established via UNICORE.

2.2 UNICORE components

The UNICORE software provides a client and a server component. The UNICORE server consists of the Gateway, Network Job Supervisor (NJS) including an Incarnation Database (IDB) and a UNICORE User Database (UUDB), and the Target System Interface (TSI).

UNICORE has a three-tier architecture:

- The UNICORE client GUI is used for the preparation, submission, monitoring, and administration of jobs.
- The Gateway is a site's point of contact for all UNICORE connections. Authentication is done by checking if the user's public certificate is signed by a trusted CA. All public user certificates are stored in the UUDB where they are mapped to the user's login name on the target system. Site specific information on computing resources, including the availability of applications, is provided by the NJS. This component dispatches the jobs to a dedicated target machine or cluster, and handles dependencies and data transfers for complex workflows. It transfers the results of executed jobs from the target machine. The abstract definition of the Job is translated to a concrete job in the NJS with the help of the IDB. The IDB contains all target system specific issues. Each NJS has its own IDB to access one specific target system.
- The TSI, which is running on the target machine, is the interface to the batch scheduler on the target machine (up to now in DEISA to the *LL-MC* at CINECA, FZJ, IDRIS and RZG, and *LL* at CSC and ECWMF, *LSF* at SARA, *SLURM* and *LL* at BSC, as well as *PBS* at LRZ).

2.3 Installed UNICORE components at the DEISA sites

Actually the following versions of the UNICORE servers (Gateway, NJS, TSI) are installed at almost each site.

- Gateway 4.1.0 build 1
- NJS 4.3.0 build 1
- UADB 1.0.0
- TSI 4.1.1 build 1
- UNICORE Client 5.3 build 1

All UNICORE servers and client versions are available at [1].

2.4 DEISA UNICORE infrastructure configuration

Figure 1 shows the actual DEISA UNICORE configuration.

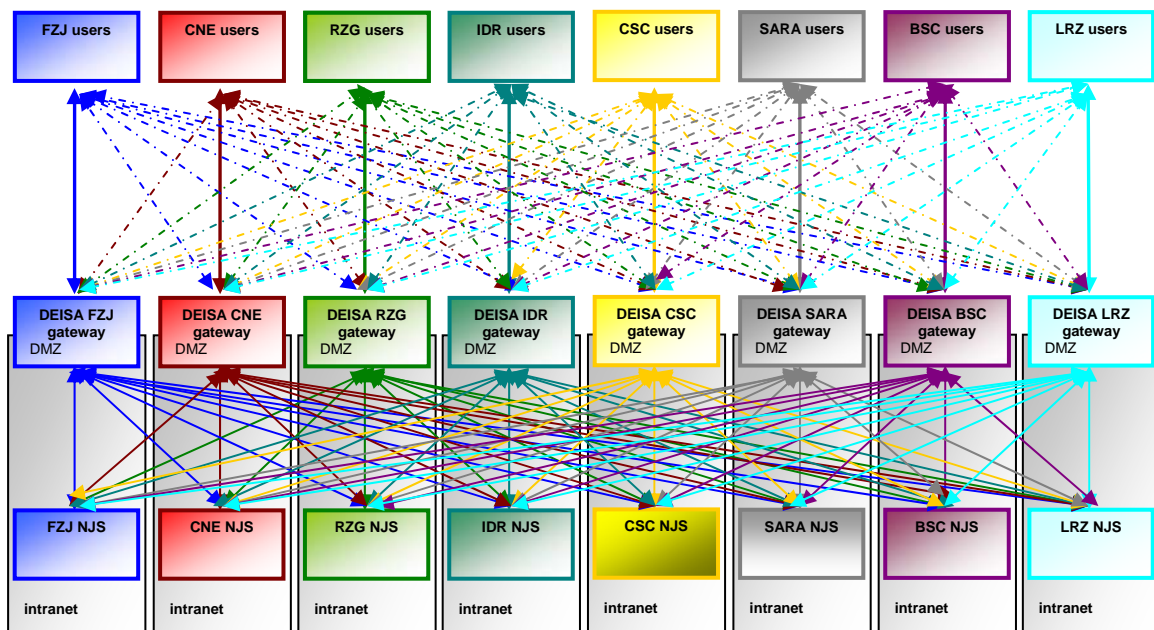


Figure 1 - Actual DEISA UNICORE configuration

In DEISA each site runs a UNICORE gateway that is configured in such a way that it can access all dedicated DEISA high performance computing resources. So there is no central access point to UNICORE but several distributed ones and thus no single point of failure. If necessary, more UNICORE gateways could be added. All NJSs from each site are allowed to register with each gateway, too.

On the one hand this has to be arranged of course consistent with local security policies regarding firewalls etc. because each site has to permit every DEISA user and all NJSs to access and register with each local gateway and each gateway has to trust all certificates they are presenting.

But on the other hand this UNICORE configuration promises the user a high flexibility. There is no single point of failure. If one gateway is down or heavily loaded other gateways can be contacted to submit UNICORE jobs to any target system within DEISA.

2.5 DEISA UNICORE infrastructure installation status

The following subchapters give a detailed plan of the installation at each DEISA site. A schematic diagram gives a general overview of the UNICORE configuration. For a successful communication between all DEISA UNICORE sites it is necessary to share some special information between the administrators. Some of the information has to be put in the configuration files of the UNICORE servers at all sites, e.g., to guarantee successful NJS registration at all Gateways and to guarantee that all DEISA users are able to connect to all Vsites. Other information is needed, e.g., for local firewall settings.

The shared information is:

- For the Gateway:
 - Machine name: the machine name is used by both the client to connect to the UNICORE site, and the NJS to register itself to the Gateway.
 - Machine IP: The IP is needed by some sites to setup firewall rules to allow connections only to the well-know gateways at the different sites.
 - Port: the port where the Gateway is listening to.

- For the NJS:
 - Machine name: Used for firewall settings at some sites.
 - Machine IP: Used for firewall settings (i.e. allow connections only from/to well-known NJSs to/from well-known Gateways).
 - Port: Has to be known by all Gateways in order to forward requests.
 - Common name of the NJS certificate: used at each gateway to authenticate and identify which NJSs are authorised to register.
 - NJS certificate (has to be added to the UADB for UNICORE to work properly).

- For the TSI:
 - No information is needed outside each site.

At each site only the Gateway port has to be opened for the whole outside world in the central firewall of the centre. All connections from user clients and NJSs located not in the centre connect to this port. While the principle of UNICORE is not to touch the autonomy of the UNICORE sites all DEISA partners are free to choose the Gateway and NJS port. This is the reason why not all DEISA sites have chosen the same ports for UNICORE in their configuration.

The NJS port listens for connections from all DEISA Gateways and expects from the local one it has to be opened just for connections from and to those Gateways. Because the DEISA sites are limited, the number of connections from other Gateways to the local NJS is limited, too.

A subset of the shared information is also needed for client configuration.

By now FZJ, CINECA, RZG, IDRIS, SARA, CSC, BSC, ECMWF, EPCC, and LRZ have installed the UNICORE server components Gateway, NJS, and TSI to access the local supercomputer resources of each site in the heterogeneous DEISA cluster.

The following subchapters explain the current DEISA installation and configuration status at each site. The DEISA partners have set up their DEISA UNICORE environment in a quite different way concerning local policies. All Gateways are protected from Internet by a firewall. But at some sites the Gateway additionally is located in a Demilitarised Zone (DMZ) to get extra protection.

Usually the Gateway and NJS are on different machines and in some cases the NJS is protected from an internal firewall, too. At most sites the NJS and TSI are separated for security reasons.

Some sites only allow DEISA Gateways to access the local NJS but others have opened the concerning ports for the whole world.

The different setups of DEISA UNICORE at the partner sites do not have any bearing on the DEISA UNICORE functionality which is the same at each site.

2.5.1 FZJ

FZJ has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on different Linux machines. The TSI runs on the target system which is an IBM AIX super-cluster.

The following schematic diagram shows the FZJ configuration.

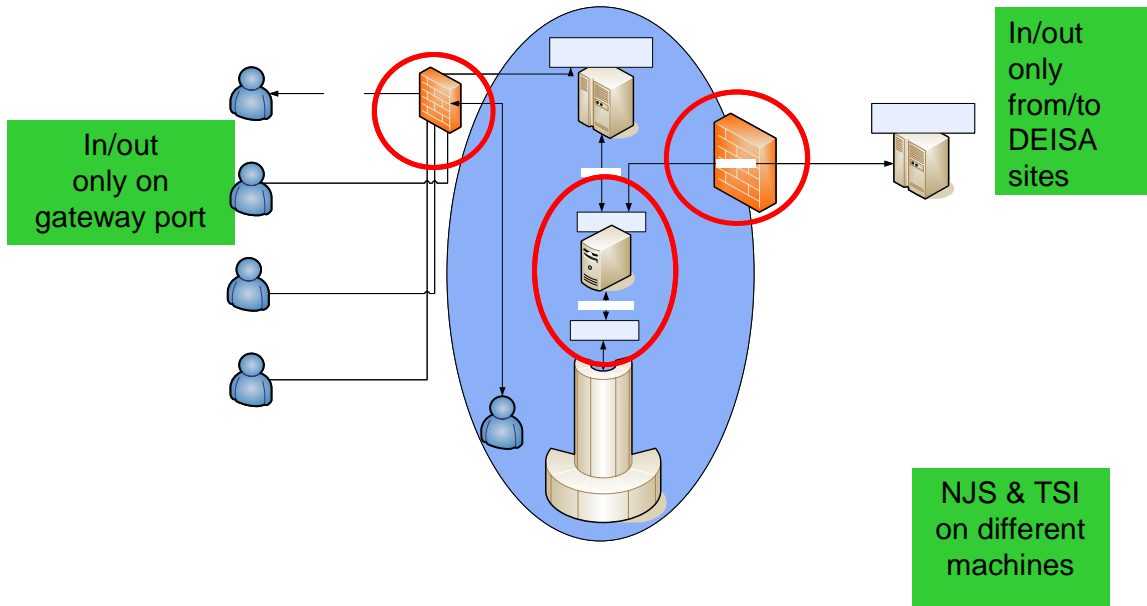


Figure 2 - UNICORE configuration at FZJ

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	zam177.zam.kfa-juelich.de
Gateway IP address	134.94.168.56
NJS machine name	zam178.zam.kfa-juelich.de
NJS IP address	134.94.168.57
CN of NJS certificate	DEISA Demo NJS
Client (NJS)-Gateway port	4000
Gateway-NJS port	9785

Table 2 - UNICORE configuration details at FZJ

2.5.2 RZG

RZG has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on different Linux machines. The Gateway is located in a DMZ. The TSI runs on the target system which is an IBM AIX super-cluster. The following schematic diagram shows the RZG configuration.

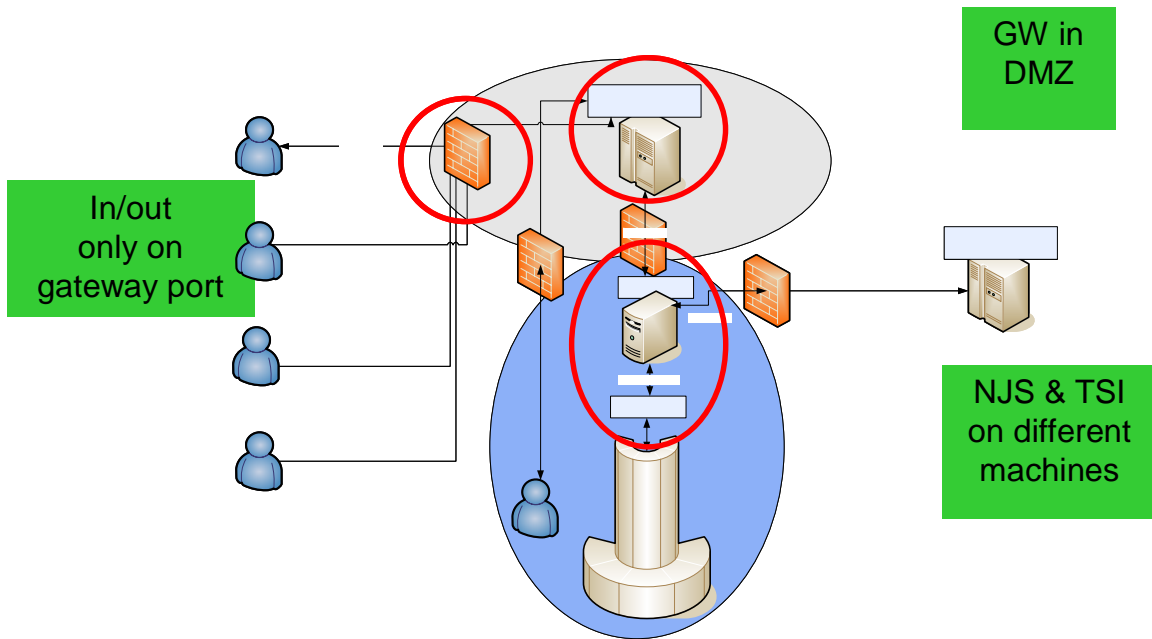


Figure 3 - UNICORE configuration at RZG

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	unigate.rzg.mpg.de
Gateway IP address	130.183.3.42
NJS machine name	unicorn.rzg.mpg.de
NJS IP address	130.183.8.180
CN of NJS certificate	unicorn.rzg.mpg.de
Client (NJS)-Gateway port	4433
Gateway-NJS port	8181

Table 3 - UNICORE configuration details at RZG

2.5.3 IDRIS

IDRIS has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on different Linux machines. The Gateway is located in a demilitarized zone (DMZ). The TSI runs on the target system which is an IBM AIX super-cluster.

The following schematic diagram shows the IDRIS configuration.

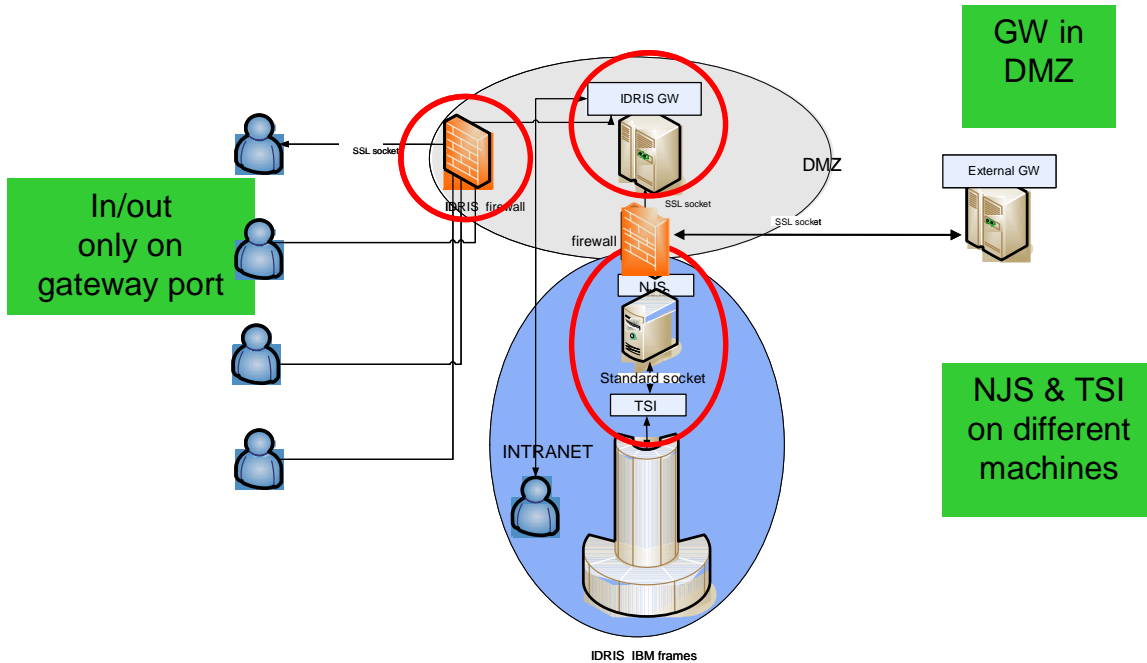


Figure 4 - UNICORE configuration at IDRIS

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	sirius.idris.fr
Gateway IP address	130.84.37.20
NJS machine name	canopus.idris.fr
NJS IP address	192.54.160.20
CN of NJS certificate	njszahir
Client (NJS)-Gateway port	4433
Gateway-NJS port	8193

Table 4 - UNICORE configuration details at IDRIS

2.5.4 CINECA

CINECA has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on the same Linux machine. The TSI runs on the target system which is an IBM AIX super-cluster.

The following schematic diagram shows the CINECA configuration.

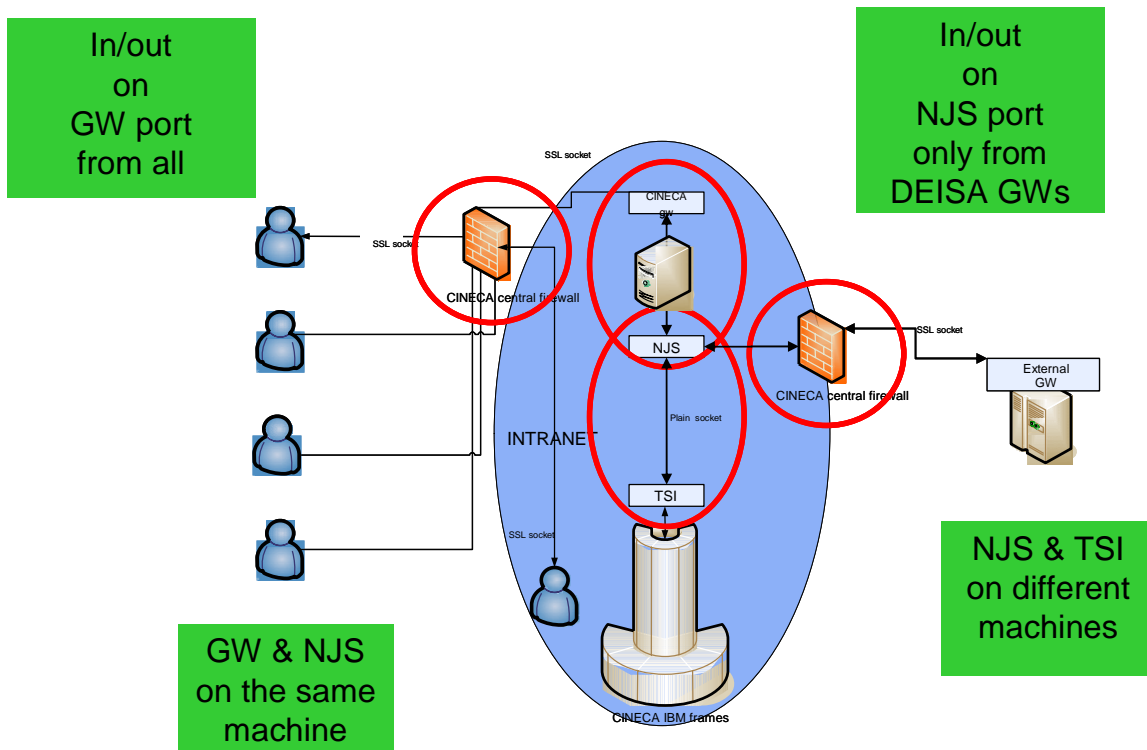


Figure 5 - UNICORE configuration at CINECA

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	reunion.cineca.it
Gateway IP address	130.186.1.43
NJS machine name	reunion.cineca.it
NJS IP address	130.186.1.43
CN of NJS certificate	login011.sp4.cineca.it
Client (NJS)-Gateway port	4433
Gateway-NJS port	8181

Table 5 - UNICORE configuration details at CINECA

2.5.5 SARA

SARA has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on different Linux machines and are located both in a DMZ. The TSI runs on the target system which is a SGI Altix Linux cluster. The following schematic diagram shows the SARA configuration.

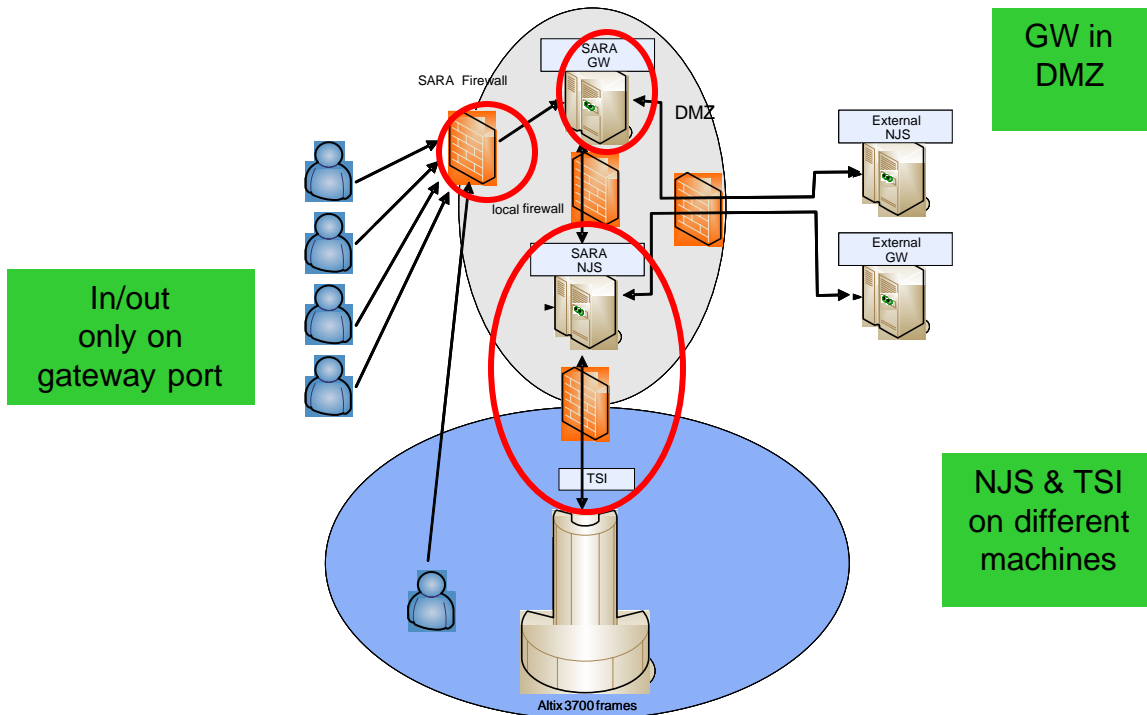


Figure 6 - UNICORE configuration at SARA

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	uni-gw1.sara.nl
Gateway IP address	145.100.29.233
NJS machine name	uni-njs1.sara.nl
NJS IP address	145.100.29.234
CN of NJS certificate	uni-njs1.sara.nl
Client (NJS)-Gateway port	4004
Gateway-NJS port	4444

Table 6 - UNICORE configuration details at SARA

2.5.6 CSC

CSC has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on different Linux machines. On the NJS machine a local firewall is running. The TSI runs on the target system which is an IBM AIX super-cluster. The following schematic diagram shows the CSC configuration.

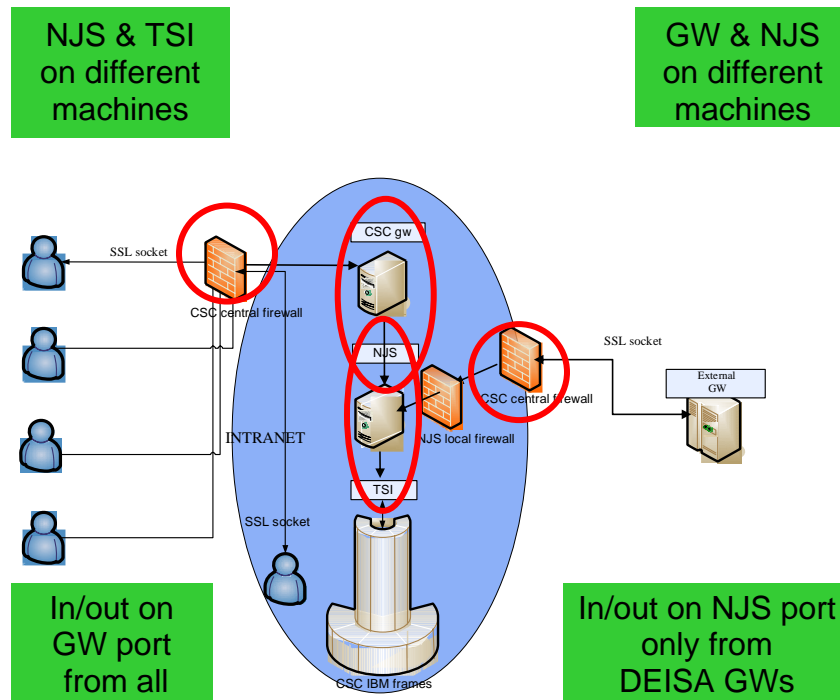


Figure 7 - UNICORE configuration at CSC

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	uni.csc.fi
Gateway IP address	193.166.7.126
NJS machine name	hiekkka.csc.fi
NJS IP address	193.166.7.127
CN of NJS certificate	hiekkka.csc.fi
Client (NJS)-Gateway port	4000
Gateway-NJS port	9785

Table 7 - UNICORE configuration details at CSC

2.5.7 BSC

BSC has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on different Linux machines and are located both in a separate DMZ. The TSI runs on the Target system which is an IBM PowerPC Linux system. The following schematic diagram shows the BSC configuration.

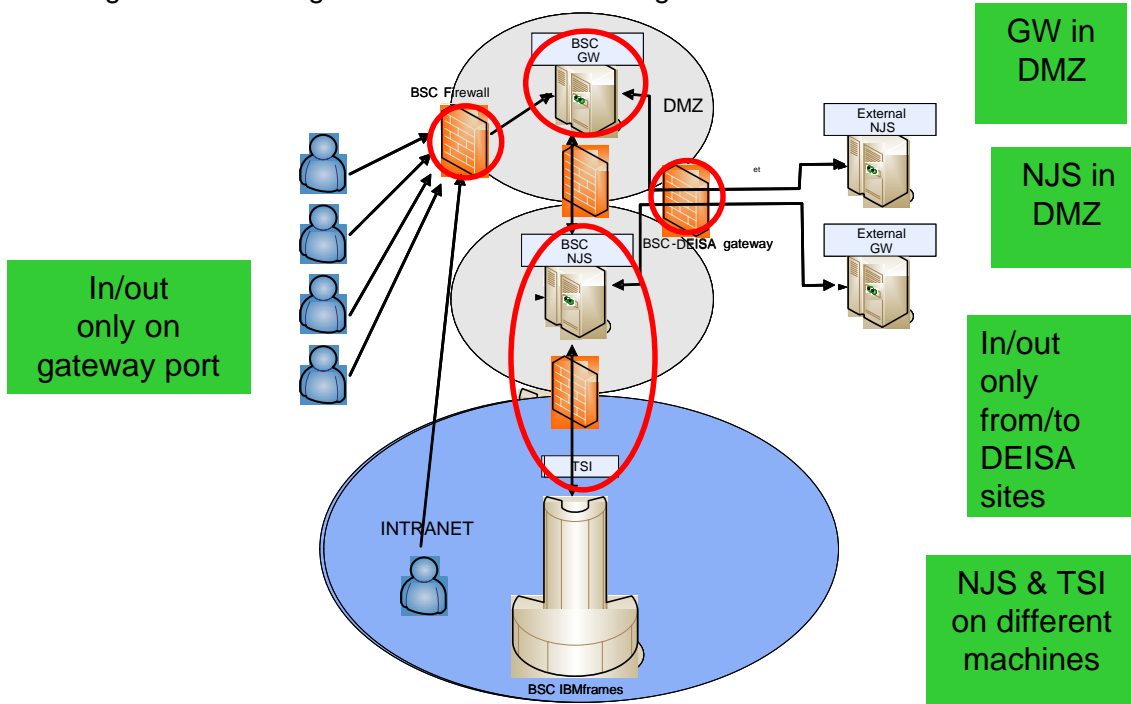


Figure 8 - UNICORE configuration at BSC

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	opsuni01.bsc.es
Gateway IP address	84.88.52.131
NJS machine name	opsuni02.bsc.es
NJS IP address	84.88.52.132
CN of NJS certificate	unicore02.bsc.es
Client (NJS)-Gateway port	4433
Gateway-NJS port	4422

Table 8 - UNICORE configuration details at BSC

2.5.8 LRZ

LRZ has installed the DEISA UNICORE environment and everything runs properly. Both Gateway and NJS run on different Linux machines. The TSI runs on the Target system which is currently a Linux cluster, but will be replaced by an SGI Altix in 2006. The following schematic diagram shows the LRZ configuration.

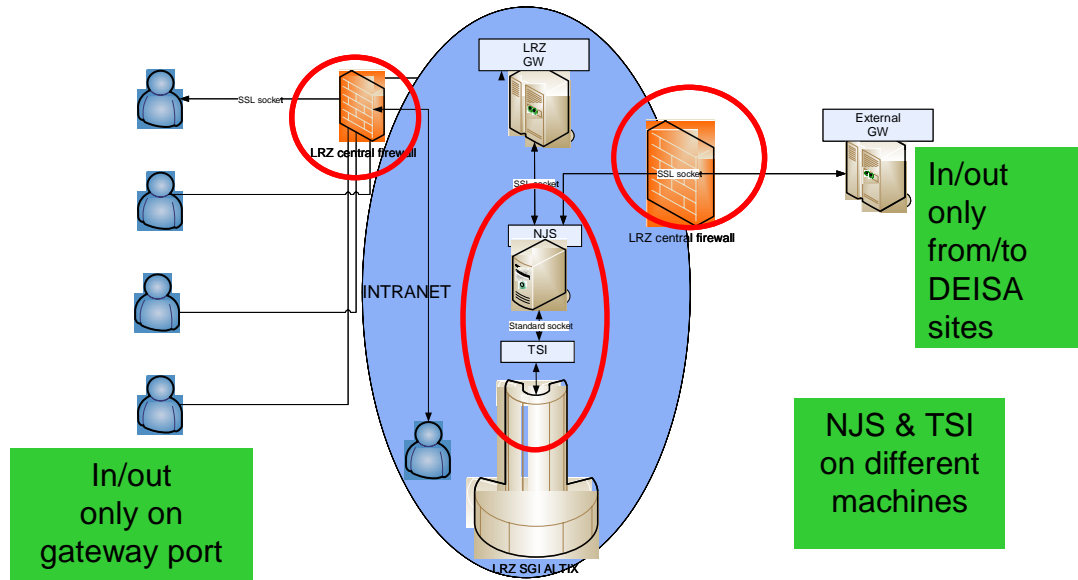


Figure 9 - UNICORE configuration at LRZ

All necessary information for the other UNICORE sites is collected in the following table.

Gateway machine name	unicore.lrz-muenchen.de
Gateway IP address	129.187.254.70
NJS machine name	lxsrv0.lrz-muenchen.de
NJS IP address	129.187.20.237
CN of NJS certificate	lxsrv0.lrz-muenchen.de
Client (NJS)-Gateway port	4433
Gateway-NJS port	8186

Table 9 - UNICORE configuration details at LRZ

2.5.9 ECMWF

ECMWF has installed UNICORE but needs to complete a final internal security risk assessment before opening the service to the rest of the infrastructure. ECMWF is the only current DEISA partner to have additional security requirements where clients are required to be strongly authenticated using locally validated Smart Card passcodes or very short lived X.509 certificates. To comply with the design of the current DEISA UNICORE infrastructure ECMWF has found an interim solution by deploying two Network Job Supervisors (NJS). One of the NJS will be fully connected to the other DEISA gateways providing information about ECMWF resources but will not accept jobs to be submitted, the reason being that authentication would be done at the remote gateway contrary to its security policy. A second NJS will be connected externally to ECMWFs gateway, which will accept jobs to be submitted, but forces authentication to be done locally as per its security policy.

The following schematic diagram shows the future ECMWF DEISA UNICORE configuration.

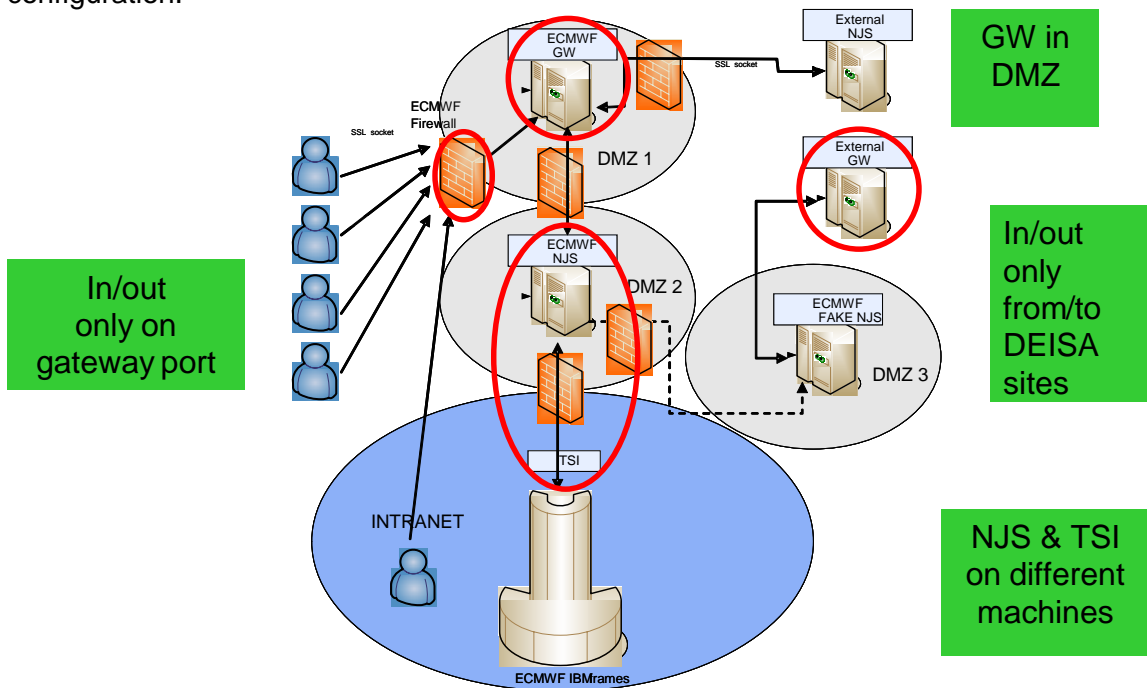


Figure 10 - UNICORE configuration at ECMWF

2.5.10 EPCC

EPCC have installed the DEISA UNICORE environment on the HPCx system. The gateway and NJS run on the same restricted access machine. The TSI is installed on the target system which is an IBM AIX Power4 cluster. This cluster will be replaced during November 2005 by a Power5 cluster, onto which the TSI will be migrated.

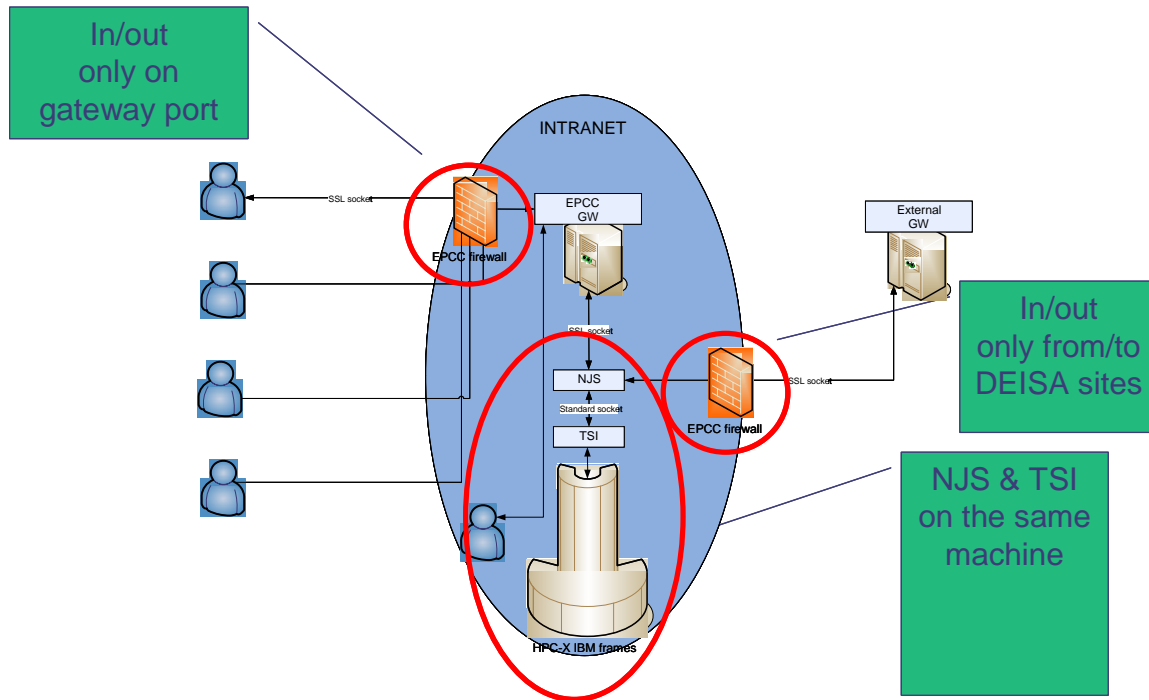


Figure 11 - UNICORE configuration at EPCC

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	admin.hpcx.ac.uk
Gateway IP address	193.62.122.2
NJS machine name	admin.hpcx.ac.uk
NJS IP address	193.62.122.2
CN of NJS certificate	ladmin.hpcx.ac.uk
Client (NJS)-Gateway port	4433
Gateway-NJS port	8181

Table 10 - UNICORE configuration details at EPCC

2.5.11 HLRS

HLRS has installed the DEISA UNICORE environment and everything runs properly. The Gateway runs on its own Linux machine, which is located in the HLRS DMZ. The NJS also runs on its own Linux system which is located in the HWW network where also the TSI and the Target Systems are located. The TSI for the NEC SX-8 runs on the frontend system, a NEC TX-7. The following schematic diagram shows the HLRS configuration.

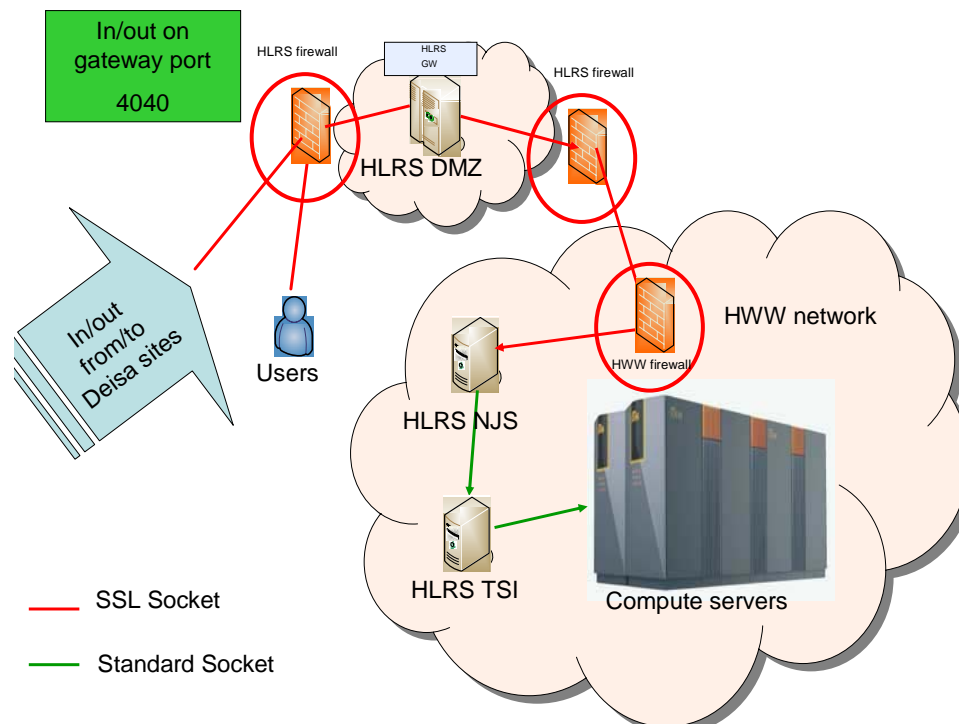


Figure 12 – UNICORE configuration at HLRS

All needed information for the other UNICORE sites are collected in the following table.

Gateway machine name	hww-gw1.hlrs.de
Gateway IP address	141.58.2.11
NJS machine name	njs1.hww.de
NJS IP address	193.196.155.196
CN of NJS certificate	njs1.hww.de
Client (NJS)-Gateway port	4040
Gateway-NJS port	confidential

Table 11 - UNICORE configuration details at HLRS

3. The DEISA RMIS infrastructure status

3.1 Introduction

The DEISA Resource management Information System (RMIS) is an information system which gives site administrators an up to date real-time view of the status of all DEISA resources.

Among the DEISA partners, RZG coordinates the management of the DEISA RMIS infrastructure and gives RMIS administrators installation and configuration support.

3.2 RMIS components

As described in an earlier deliverable [3], the current RMIS deployment consists of the following components:

- Ganglia: The gmond (see [3]) is installed on all compute nodes which are available to DEISA at each site. The gmond gathers data about the system status and availability and usage of those nodes. These data are collected at a central gmond or gmetad depending on the local installation.
- MDS2 ganglia backend: the collecting ganglia daemon is queried by a special Globus MDS2 backend and translated from XML to the LDIF format in order to be hooked into the general LDAP structure.
- MDS2 LoadLeveler backend: the LoadLeveler resource management status is queried by a special Globus MDS2 backend which transforms the text output into LDIF.

3.3 Installed RMIS components at the DEISA sites

As stated above, RMIS consists of several software components. The deployment of exactly those components can be seen as a (strong) suggestion for the each of the DEISA partners, since the provision of the required information can also be achieved by deploying different sets of tooling.

Hence, in the contrary to the UNICORE deployment, it is the interface which is standardized in the RMIS deployment. That means that each site has to provide some reliable LDAP service implementing the DEISA RMIS schema. Furthermore, each site has to guarantee the quality of the provided data.

This will make it possible for sites like HLRS, which have special security requirements, to deploy a non default installation of the RMIS components.

The LDAP schema is available from the DEISA BSCW work space at <http://work.deisa.org/>. This workspace includes also example deployment configuration for the default RMIS deployment.

3.4 DEISA RMIS infrastructure configuration

In an earlier deliverable [3] the use of the Ganglia gmetad has been envisaged. Further analysis and first experience showed that the history information gathered by the gmetad demon with the help of the rrdtool library [5], is not needed at remote sites in order to analyse problems. Most sites gather history information in other ways (e.g. by means of proprietary tools) and have no need of accessing gmetad's repository. Should the use of gmetad become necessary in the future, it is straight forward to reintegrate it.

Figure 13 sketches the current deployment at most sites. gmond daemons are running on all nodes which need to be monitored (usually the compute nodes). In that figure, one gmond is configured to receive information from all other daemons and assembles a status report which can then be queried from the MDS2 by the help of special scripts. In large clusters, such as BSC's MareNostrum, a hierarchy of gmonds can be implemented in order to enhance the scalability of the Ganglia deployment and avoid a strong load on the server hosting the gathering gmond in a single gmond environment.

A user, per definition a privileged administrator, of a remote site A is now able to connect to the MDS2 of site B and use the provided information, e.g. in order to analyse a problem with a compute job at that particular site.

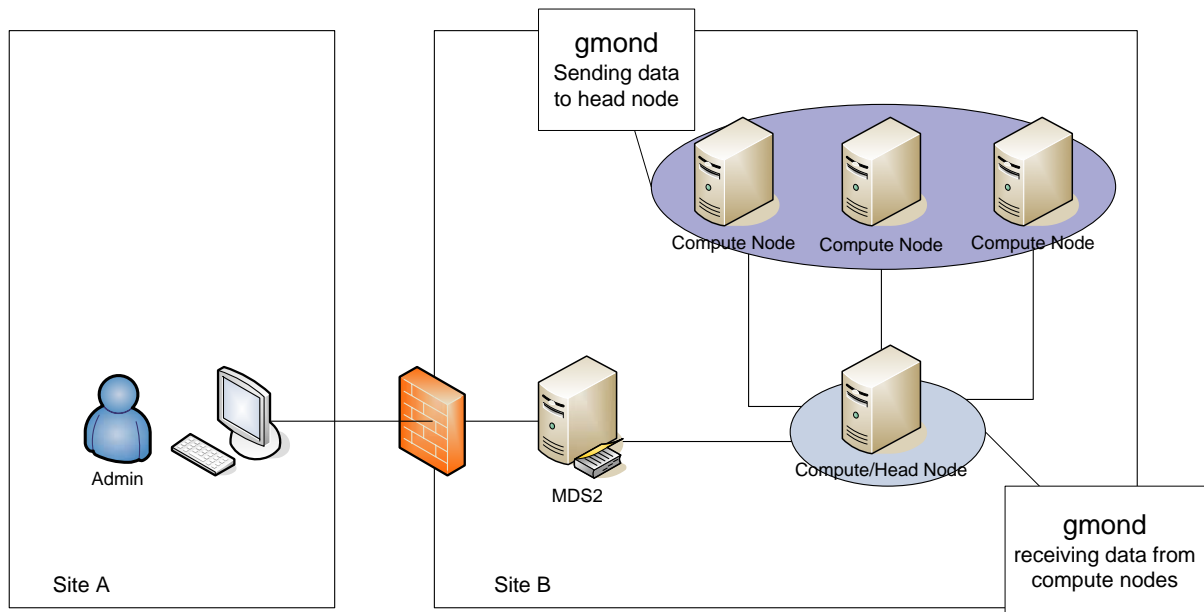


Figure 13 – Default RMIS configuration

3.5 DEISA RMIS infrastructure installation status

Currently RMIS is fully deployed at the four core sites, while others are in the installation and deployment process.

Site	MDS2 address	components
CINECA	Ldap://ldeisap.cineca.it:2135	All
FZJ	Ldap://ldap2.zam.kfa-juelich.de:2135	All
IDRIS	Ldap://orion.idris.fr:2135	All

RZG	Ldap://unicorn.rzg.mpg.de:2135	All
SARA	Ldap://mds.deisa.sara.nl:2135	MDS2, ganglia
CSC	Ldap://hiekkka.csc.fi:2135	MDS2, ganglia
EPCC		None
ECMWF		None
LRZ	Ldap://altix.lrz-muenchen.de:2135	MDS2
BSC		ganglia
HLRS		none

EPCC are in the process of replacing the current HPCx IBM Power 4 based system with an IBM Power 5 system. In the context of this migration the RMIS components will not be installed on the old system, however they are being considered with respect to installation onto the new system. In particular, the MDS has already been built and installed, so the LoadLeveler information provider should be able to be installed reasonably easily onto the new machine.

ECMWF will consider starting the installation of RMIS components after assessing locally the security and other implications that installing these components on its HPC resources may have.

LRZ will install all RMIS components as soon as possible.

The local RMIS installation will be completed in BSC as soon as possible. BSC is currently preparing the hardware and dedicated network for being connected to DEISA; because of this, the RMIS integration with the rest of the DEISA sites will wait till this infrastructure is ready

HLRS is running its HPC Systems in cooperation with industry in a special dedicated environment. For this, a specific security setup has been agreed and constituted. For the RMIS installation, we need first a security analyses, including an analyses of the impacts on the current environment. Basing on this, an agreement between the partners in the cooperation has to be concluded.