



CONTRACT NUMBER 031513

eDEISA
**EXTENDED DISTRIBUTED EUROPEAN
INFRASTRUCTURE FOR
SUPERCOMPUTING APPLICATIONS**

European Community Sixth Framework Programme
RESEARCH INFRASTRUCTURES
Integrated Infrastructure Initiative

Formation of Operation Team
Definition and Setup of Activity

Deliverable ID: eDEISA-eSA2-1

Due date: November, 30th, 2006
Actual delivery date: December 21, 2006
Lead contractor for this deliverable: RZG, Germany

Project start date: June 1st, 2006
Duration: 2 years

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
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	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Content

Table of Content.....	1
1. Introduction.....	2
1.1 Executive Summary.....	2
1.2 References and Applicable Documents	2
1.3 Document Amendment Procedure	2
1.4 List of Acronyms and Abbreviations	2
2. The Operation Team.....	3
2.1 DEISA before the Operation Team.....	3
2.2 Requirements of the Operation Team	3
2.3 Formation of the Operation Team	3
3. Definition of Responsibilities.....	4
3.1 Information Exchange and Planning.....	4
3.2 Service Activity Coordination.....	4
3.3 Cross-Site Coordination	4
3.4 Resource Coordination.....	5
3.5 Emergency Response Team	5
4. Definition and Setup of Communication	5
4.1 Video Conferences.....	5
4.2 Document Exchange	5
4.3 Contact Partners.....	5
4.4 Trouble Ticket System.....	6
5. Quality Assurance.....	6
6. Summary of the Work	6
6.1 Security Discussion	6
6.2 Maintenance Rules.....	6
6.3 Common Production Environment.....	7
6.4 GPFS Upgrade.....	7
6.5 DEISA Users	7
7. Overview	7

1. Introduction

1.1 Executive Summary

The eSA2 activity, named Operation Team, is responsible for a smooth operation of the DEISA [1] research infrastructure integrating 11 leading HPC centres in Europe. In this first deliverable the formation of the Operation Team is described as well as the definition and setup of its responsibilities and work flows, by which the quality of the services provided through DEISA is assured. Since all service activities of DEISA and all HPC centres participating in DEISA are covered this document is of global interest to everybody involved with DEISA.

1.2 References and Applicable Documents

- [1] DEISA project homepage, <http://www.deisa.org>
- [2] DEISA Deliverable D-SA2-5A, "Integration of Altix and PPC Systems and Improvements in MC-GPFS for Production",
<http://work.deisa.org/bscw/bscw.cgi/d46523/D-SA2-5A.doc>
- [3] Operation Team Status Report, "Report on Current Status and Agreements of Operation Group", Version 2006-12-08,
<http://work.deisa.org/bscw/bscw.cgi/d30517/Status-Report.doc>

1.3 Document Amendment Procedure

The initial document amendment procedure is via communication between members of eDEISA eSA2 team. The document is then submitted for review to the DEC and an internal DEISA reviewer appointed by the DEC. The document is then amended according to comments received from the DEC and the internal DEISA reviewer. It is subsequently re-submitted to the DEC for submission to the EU.

1.4 List of Acronyms and Abbreviations

AIX	Advanced Interactive eXecutive (IBM's derivative of UNIX OS)
Altix	Multi-Processor compute node from SGI
ATASKF	Application Task Force in DEISA
DCPE	DEISA Common Production Environment
DEC	DEISA Executive Committee
DECI	DEISA Extreme Computing Initiative
DEISA	Distributed European Infrastructure for Supercomputing Applications
GPFS	General Parallel File System
HPC	High Performance Computing
IBM	International Business Machines (Computer Manufacturer)
Linux	Free UNIX-like Operating System
MC-GPFS	Multi-Cluster GPFS

PPC	PowerPC (a CPU chip type)
SA	Service Activity
SGI	Silicon Graphics Incorporated (Computer Manufacturer)
WAN	Wide Area Network

2. The Operation Team

2.1 DEISA before the Operation Team

Before the eSA2 activity had been defined, DEISA already had the need to operate the infrastructure. The responsible persons for different topics were the naturally involved ones from the affected SAs and from the local responsables for operations. Since communication between the different SAs was always required some implicit form of an Operation Team was active from the very beginning of DEISA.

After the initial setup of DEISA more and more scientists used the infrastructure, for which it was intended for. Soon it was obvious that for the real day to day operation of the DEISA infrastructure the more or less informal structure of the implicit Operation Team was not sufficient any more.

2.2 Requirements of the Operation Team

Thus with eDEISA the eSA2 activity formalized the need of an operating infrastructure group involving highly experienced system administrators of each site, who are able to coordinate the operational environments existing at each site as well as the interactions influencing the operational behaviour of the infrastructure between the different SAs in DEISA.

The tasks to be covered include the identification of requirements in the setup of the single sites. It also means the coordination of actions which need to be performed at each site. This includes planning for configuration changes concerning DEISA and discussion of influences of different options and reporting them to the DEC.

Beside these actions, which require regular communication, the Operation Team should also act like an emergency team, in cases where a quick reaction on problems arising in the DEISA infrastructure are needed. Thus the Operation Team and its members have to have the technical knowledge and administrative privileges at each site to accomplish the task.

2.3 Formation of the Operation Team

After having defined these main objectives for the Operation Team, each site named one person and a substitute responsible for the above mentioned tasks of the Operation Team. The following persons have been named:

Site	Operation Team Member
BSC	Sergi Girona
CINECA	Stefano Martinelli
CSC	Pekka Lehtovouri
ECMWF	Ricardo Correa
EPCC	Jeremy Nowell
FZJ	Klaus Wolkersdorfer
HLRS	Thomas Boenisch
IDRIS	Philippe Collinet

LRZ	Horst-Dieter Steinhöfer
RZG	Andreas Schott (team-leader)
SARA	Vincent van Elzen

3. Definition of Responsibilities

3.1 Information Exchange and Planning

One of the major tasks for the Operation Team is the exchange of information. This covers not only the information required for the coordination of the service activities or actions affecting more than one site. It is also the distribution of knowledge one partner gained to all the others. Since the members of the Operation Team are all senior staff members at their site, they often participate in conferences or other meetings (e.g. SPXXL), where information of general interest is provided.

Such information not only helps understanding local operation problems but is useful for all partners also, since local problems often occur similarly at one ore more of the other sites. An example may illustrate that: FZJ was upgrading the batch software (LoadLeveler) used in DEISA, when some problems occurred, which were later solved by IBM. Therefore, all other IBM sites in DEISA avoided the upgrade before the fix was available by IBM.

Furthermore new technical information gathered by one partner can be communicated to all the others. This is especially useful for planning possible actions influencing the DEISA infrastructure. Thus the Operation Team can provide useful advice for the service activities and the executive committee of DEISA concerning technical decisions for the future.

3.2 Service Activity Coordination

All service activities have well defined communication and operating channels concerning their own activity. But often decisions and actions of one service activity have an influence on another service activity. Thus coordination effort between different service activities is required on a case by case basis. The Operation Team is the instance to establish this.

Another example may describe this. The setup of the Global File System GPFS by SA2 required specific network settings. Most of these parameters have been exchanged directly between SA1, providing the network, and SA2, responsible for the file system. But there was the need of changing configurations in the firewall which are seen as security critical by some sites. So a joint video conference of all relevant SA members as well as the Operation Team was held to discuss the problem and finally a solution was developed and implemented which respected all local site security policies.

3.3 Cross-Site Coordination

The above example also shows already that not only the coordination between service activities but also the coordination of actions between the sites is required. In all cases where the actions of one site have an influence on the others often not only an exchange of information is necessary but also some cooperative actions have to be achieved.

Another example shall explain this. Since all the DEISA sites operate in principle autonomous, the actions at one site may have an influence at another site. When a site is going into maintenance, its part of the Global File System will not be available at any other site. So no jobs of users of that site may run on another site as their

DEISA home and data directories are not available. This requires a dismount of the file system and a closing of the batch system for the affected users.

3.4 Resource Coordination

One of the tasks relevant for the DECI projects assigned by the ATASKF after scientific approval and acceptance by the DEC is the provision of resources to such jobs. The Operation Team members take local responsibility to ensure availability of adequate disk space and proper batch queue configurations.

3.5 Emergency Response Team

Last but not least in a complex configuration like DEISA it may happen that the usually smooth operation is disrupted. The analysis of such situations may be done in a first step by the single sites. If the analysis proves that another site has to perform some changes, it is outside of the scope of the local site responsible. In such cases the Operation Team takes responsibility and triggers the required actions.

4. Definition and Setup of Communication

4.1 Video Conferences

The general exchange of information between all partners has to occur quite regularly. To minimize travelling but having almost face-to-face communication, it was decided to hold video conferences, which usually occur bi-weekly, but at least once in a month. From all these video conference minutes are taken and distributed inside DEISA using the BSCW server, which is the document sharing facility in DEISA.

4.2 Document Exchange

Inside the BSCW server a sub-tree was setup for the Operation Team. In this folder all documents related to the work of the Operation Team are available. Beside the already mentioned minutes of the video conferences a continuously updated Operation Team Status Report document [3] contains all relevant agreements and other information relevant to the service activities across the sites.

4.3 Contact Partners

Since the Operation Team has a more global view on the whole DEISA infrastructure it is clear that it is quite useful to have short communication channels. For this reason one of the first actions of the Operation Team was the gathering of all names, email-addresses and phone numbers of all persons working for and in DEISA at all sites. This list then was distributed among all Operation Team members via the BSCW server.

Having all email and phone numbers available the Operation Team is able to contact every person concerned with DEISA in a fast and direct way. This is helpful for all situations when a quick action is required.

Operation Team members often attend face-to-face meetings of other DEISA or eDEISA service activities. Along with those meetings the Operation Team also has informal face-to-face meetings. Since the Operation Team is mostly a coordinating group, extra face-to-face meetings besides the regular video conferences do not seem to be necessary.

4.4 *Trouble Ticket System*

In DEISA already a Trouble Ticket System (TTS) was setup for the single service activities to document problems potentially crossing the sites. This TTS can easily be used for problems involving more than one service activity. The Operation Team regularly checks the open calls in the TTS to eventually coordinate service activities or trigger actions at the sites, which are beyond the scope of a single service activity.

5. *Quality Assurance*

Currently there is no written document yet available, describing the procedures, methods, and actions to assure the quality of service provided by this service activity. The eSA3 service activity is working on a document which describes the quality assurance requirements and the process for eSA3 and SA3.

As soon as this document is finalized it will be adopted to eSA2. But the main criteria of quality assurance are already in place. The quality assurance quadruple "Plan, Do, Check and Act" is in place.

"Plan" means, what has to be done and how. This is in principle documented here in section 3. "Do" covers the work being done and the documentation of that work. This is covered mainly by the usage of the means described in the previous section 4. "Check and Act" are the corrective actions implied by analyzing the "Plan" and the "Do" actions and improving them. This is an implicit action performed in the regular video conferences. Improvements may also be triggered by requests from other service activities or the DEC.

6. *Summary of the Work*

6.1 *Security Discussion*

As already mentioned above as an example, there was a long lasting discussion about firewall restrictions to be applied between the sites. Since this topic is mainly implemented by SA1 (Network) it has also an influence on all other service activities. For example: SA2 (Global File System) requires a special setup for the internal communication of the MC-GPFS. SA3 (Middleware) is also heavily relying on open communication channels between the sites in order to submit or migrate jobs to different sites. Last but not least SA5 (Security) is naturally concerned with all issues concerning security.

To solve a current problem concerning such firewall rules a joint video conference with all relevant persons from the single service activities together with the Operation Team was held on July, 13th. The discussion and agreements are documented in the already mentioned Status Report, available from the BSCW server.

6.2 *Maintenance Rules*

Maintenance actions at one site affect the operation of the whole DEISA infrastructure. In order to inform all affected partner sites, a mailing list solely dedicated to announcements of planned maintenance actions has been setup. It is also documented in the Status Report.

During maintenance parts of the MC-GPFS may not be available, so batch jobs requiring such parts of the MC-GPFS may not run. A policy for handling such cases has been discussed and is implemented for automation of proper actions as far as possible. The policy is documented in the Status Report, too.

6.3 Common Production Environment

The DEISA wide Common Production Environment (DCPE) developed by SA4 (Applications and User Support) available to all users includes many software packages and setup. Installation or upgrade of versions almost always requires coordination between all sites. Thus this part of SA4 is a regular topic at the video conferences of the Operation Team. Mostly the regular checking of the DCPE with the INCA system requires local setup actions in the operating system, which have to be coordinated by the Operation Team.

6.4 GPFS Upgrade

The greatest challenge for the Operation Team was the coordination of a software upgrade concerning the MC-GPFS affecting almost all DEISA sites. Before deciding that the software has to be upgraded the Operation Team discussed in detail the influences and effects, whether the upgrade should be performed or not. Especially an upgrade affected the integration of non-AIX sites into the homogenous MC-GPFS. This could not be decided inside SA2 solely since the decision had a strong influence on DEISA as a whole. Finally it was decided that the upgrade was necessary. The upgrade itself is described in detail in the latest deliverable of SA2 (D-SA2-5A) [2].

In order to provide continued and reliable DEISA services, the upgrade had to be performed in a time period as short as possible. To ensure the availability of all experts the date was chosen after the holiday season. Actually almost all sites could perform the upgrade even within one day, the 26th of September 2006.

6.5 DEISA Users

One of the main aims of DEISA is to provide the whole infrastructure transparently to all users. Thus to achieve the possibility to migrate jobs submitted at one site to another one, some rules on the requirements of the available resources had to be setup. These definitions influenced the configuration of batch systems on all AIX sites. The modified setup allows so called DEISA Standard Users (DSUs) to just submit a job not knowing which site will actually compute the results.

Another discussion concerns jobs submitted by portals. The eSA3 portal maps many individual users to just one actual userid inside DEISA. The discussion of eSA3 was escalated to the Operation Team to find agreements on how this behaviour can be aligned with differing site policies.

7 - Overview

The following viewgraph visualizes the integration of the Operation Team (eSA2) with the single DEISA service activities.

