J Grid Computing (2011) 9:259–277
DOI 10.1007/s10723-011-9183-2

DEISA—Distributed European Infrastructure for Supercomputing Applications

Wolfgang Gentzsch · Denis Girou · Alison Kennedy · Hermann Lederer · Johannes Reetz · Morris Riedel · Andreas Schott · Andrea Vanni · Mariano Vazquez · Jules Wolfrat

Received: 27 July 2010 / Accepted: 10 February 2011 / Published online: 1 March 2011

Abstract The paper presents an overview of the current research and achievements of the DEISA project, with a focus on the general concept of the infrastructure, the operational model, application projects and science communities, the DEISA Extreme Computing Initiative, user and application support, operations and technology, services, collaborations and interoperability, and the use of standards and policies. The paper concludes with a discussion about the long-term sustainability of the DEISA infrastructure.

Keywords DEISA · e-Science · e-Infrastructures · Grid computing · Supercomputing · High performance computing · Interoperability · Security

1 Introduction

About 10 years ago, Grid infrastructures for pooling and sharing of compute resources and for fostering collaboration among scientists started to emerge world-wide. Today, scientific Grids have matured, their scope extended, and "e-Infrastructures" or "Cyberinfrastructures" built up as a subspecies of "research infrastructures". Powerful Grids of continental scope evolved, as TeraGrid and Open Science Grid (OSG) in the US, Enabling Grids for E-SciencE (EGEE) [1] and the Distributed European Infrastructure for Supercomputing Applications (DEISA) in the EU, NAREGI in Japan, and Australian Partnership for Advanced Computing (APAC) in Australia, but also as large distributed desktop Grids [2].

Those Grids were defining, selecting, testing, optimising and further developing their portfolio of middleware according to their specific needs and strategic goals. In that way, Grids have

W. Gentzsch (✉) · H. Lederer · J. Reetz · A. Schott
RZG-Max Planck Society, Garching, Germany
e-mail: gentzsch@rzg.mpg.de

D. Girou
IDRIS-CNRS Orsay, Paris, France

A. Kennedy
EPCC, University of Edinburgh, Edinburgh, UK

M. Riedel
Forschungszentrum Jülich, Jülich, Germany

A. Vanni
CINECA Bologna, Bologna, Italy

M. Vazquez
BSC Barcelona, Barcelona, Spain

J. Wolfrat
SARA, Amsterdam, The Netherlands

matured and reached production state quality, supporting scientific communities which are becoming increasingly distributed geographically, and their members now have access to computational and data resources as well as excellent user support. To enhance and facilitate co-operation among science communities in different parts of the world, interoperability of infrastructures is currently being improved, and synergies between different infrastructures and science community projects are being exploited.

e-Infrastructures with European scope are well settled network organisations and projects such as DANTE/GEANT/TERENA, the High Performance Computing (HPC) oriented infrastructures DEISA and the Partnership for Advanced Computing in Europe (PRACE, [27]), and the High Throughput Computing (HTC) oriented infrastructures EGEE and EGI. These organisations and projects jointly formed the so-called European E-Infrastructure Forum (EEF) in 2009. According to its mission statement [3], this is a forum for the discussion of principles and practices to create synergies between e-Infrastructures. The goal of the EEF is to achieve seamless interoperation among leading e-Infrastructures serving the needs of the user communities that require services which can only be provided by collaborating e-Infrastructures.

This paper aims at presenting an actual overview and current research enabled by the DEISA project, with a strong focus on major contributions to the field of e-Infrastructures. The paper, however, does not (again) present a complete overview of the DEISA project, instead we selected major actual areas of interest for e-Infrastructures. The remainder of this contribution is as follows. After the introduction is Section 1, we present a brief overview of DEISA in Section 2. In Section 3, we highlight the general concept and the operational model. We offer a discussion on current projects and science communities in Section 4, and DEISA services and infrastructure in Section 5. Collaborations and interoperability are discussed in Section 6. The paper concludes with a discussion about the long-term sustainability of infrastructures like the one built in the DEISA project.

2 A Brief Overview of DEISA

The DEISA project consortium has deployed and operated the Distributed European Infrastructure for Supercomputing Application (DEISA, [4]), co-funded through the EU DEISA project, from 2004 to 2008. Since May 2008, the consortium continues to support and further develop the distributed high performance computing infrastructure and its services through the DEISA2 project with funds for another 3 years until 2011. Activities and services relevant for applications enabling, operation, and technologies are continued and further enhanced, as these are indispensable for the effective support of computational sciences in the HPC area. The resulting infrastructure is unmatched world-wide in its heterogeneity and complexity, enabling the operation of a powerful Supercomputing Grid built on top of national services, facilitating Europe's ability to undertake world-leading computational science research.

DEISA has already proved its relevance for advancing computational sciences in leading scientific and industrial disciplines within Europe and has paved the way towards the deployment of a cooperative European HPC ecosystem. The existing infrastructure is based on the tight coupling of leading national supercomputing centres, using dedicated network interconnections of GEANT2 (2008) and the National Research and Education Networks (NRENs). DEISA members and associate partners are: eleven DEISA members from seven countries, BSC (Barcelona, Spain), CINECA (Bologna, Italy), CSC (Espoo, Finland), ECMWF (Reading, UK), EPCC (Edinburgh, UK), FZJ (Juelich, Germany), HLRS (Stuttgart, Germany,) IDRIS (Orsay, France), LRZ (Garching, Germany), RZG (Garching, Germany) and SARA (Amsterdam, The Netherlands). Additionally, further centres joined DEISA as associate partners: CEA-CCRT (France), CSCS (Manno, Switzerland), KTH (Stockholm, Sweden) and JSCC (Moscow, Russia).

Launched in 2005, the DEISA Extreme Computing Initiative (DECI, [5]) regularly selects leading grand challenge HPC projects, based on a peer review system and approved by the DEISA Executive Committee (Execom), to enhance DEISA's impact on the advancement of computational sciences. By selecting the most appropriate supercomputer architectures for each project, DEISA is opening up the currently most powerful HPC architectures available in Europe for the most challenging projects. In addition, this service provisioning model has been complementary extended from short-term project support to supporting longer-term Virtual Science Communities. Collaborative activities are also being carried out with new European and other international initiatives. Of strategic importance is the cooperation with the Partnership for Advanced Computing in Europe (PRACE, [27]) which is installing a limited number of leadership-class Tier-0 supercomputers in Europe.

DEISA Vision The DEISA infrastructure aims at delivering a turnkey operational solution for a future persistent European HPC ecosystem which integrates national Tier-1 supercomputer centres (today in the TeraFlop/s range) and the new Tier-0 centres (today in the PetaFlop/s range).

DEISA Mission In the current second project phase of DEISA, the following two-fold strategy is applied: First, consolidation of the existing infrastructure developed in the initial DEISA project by guaranteeing the continuity of those activities and services that currently contribute to the effective support of world-leading computational science in Europe. Second, the evolution of this infrastructure towards a robust and persistent European HPC ecosystem, by enhancing the existing services, by deploying new services including support for European Virtual Science Communities, and by cooperating and collaborating with new European initiatives, especially PRACE that will enable shared European PetaFlop/s supercomputer systems.

3 General Concept and Operational Model

The DEISA project [4] has established a Virtual Distributed HPC Centre by expanding the local HPC centre teams to European-wide teams, organized in three operational activities:

- Operations Team (aka Operations)—operating the infrastructure
- Development and Technology Team (aka Technology)—advancing the infrastructure
- Applications and User Support Team (aka Applications)—helping user to make best use of the infrastructure.

Together these three complementary teams advance the existing infrastructure to provide an operational model, which can be integrated as a proven turnkey solution in the future European HPC ecosystem.

The *operations team* coordinates the daily operation of the production services. This includes the implementation of changes in services and systems, the resolution of problems and the handling of security incidents.

The 'development and *technology team*' makes services available for provision to the infrastructure and the end users. The activity is responsible for the evaluation of new products (any forms of software) and testing them in a pre-production phase. The team thus prepares recommendations about taking new products into production. After acceptance by management and the operations team the 'development and technology' team will prepare the installation, configuration and user documentation for the product.

The '*application and user support team*' provides support for users and their applications. To better serve the users' needs this activity is supported by the operations team by adapting configurations of production services and by the technology team in evaluating and testing

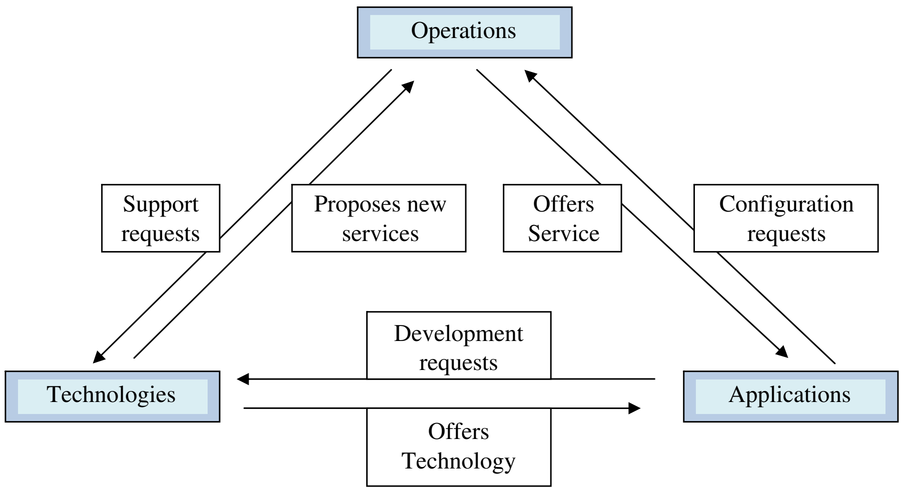


Fig. 1 Interactions between operational DEISA teams

new products useful for and requested by the users.

The three service activities "Operations", "Technologies", and "Applications" are tightly interconnected and interact as shown in Fig. 1, which highlights the most common interactions between the teams.

4 Projects and Science Communities

4.1 DEISA Extreme Computing Initiative

The DEISA Extreme Computing Initiative (DECI, [5]) is a scheme through which collaborative projects, headed by European scientists, can apply for single-project access to world-leading computational resources in the European HPC infrastructure for a period of up to 12 months per project.

DECI [5] was introduced in 2005 to enable European computational scientists to obtain access to the most powerful national computing resources in Europe regardless of their country of origin or work, and to enhance DEISA's impact on European science and technology at the highest level. Through an annual call, a number of capability computing projects are selected by peer-review on the basis of innovation and scientific excellence. The consortium has designed and deployed, and operates a complex, distributed, heterogeneous supercomputing environment with an aggregate peak performance in excess of two PetaFlop/s.

Successful projects are given access to the exceptional resources in the DEISA infrastructure (on an HPC architecture selected for its suitability to run the project's codes efficiently) and are offered applications support to enable them to use it productively. The number of proposals received has grown, from 51 at DECI-1 to 122 at DECI-6, with particularly rapid growth over the past 3 years. The number of CPU cycles requested per project has grown steadily since DECI's inception. DECI will continue to operate until 2011 through the funding of the current DEISA project funding.

DECI is of key importance in continuing to build a European HPC user community, supported in their use of top-level HPC facilities by applications experts from leading European HPC centres. More than 330 different researchers from 25 European countries took already advantage of this scheme to date as lead investigator or as named collaborator, along with named collaborators from four other continents. The scheme aims to enhance Europe's international standing in science, by providing access to world-class resources which are complementary to resources available locally. DECI is a very successful instrument for the European Union research policy of enabling smaller countries to take part in research at the cutting-edge and enables all European researchers to participate in internationally competitive computational science.

DECI also aims to facilitate a better understanding of the likely requirements of future users of the Tier-0 systems (European leadership-class supercomputers) by collecting real use-case information about what European computational scientists want and about the differences between usage of national and European resources and facilities.

In 2010, the DECI-6 call attracted 122 proposals for challenging European computational science projects, requesting more than half a billion processor(-core)-hours (in P4 processor units) and asking for significant application support. The call was very heavily oversubscribed both in requests for CPU and for applications enabling effort. 91 million processor(-core)-hours (in P4 processor units) were available for distribution on 16 European HPC systems, and these were allocated to 55 projects, which have been given production access to the infrastructure between 1 October 2010 and 30. April 2011.

Figure 2 shows how both the median and the average amount of CPU requested by projects has grown year on year. The growing divergence between the average and the median CPU requested reflects the increasing number of larger, collaborative projects applying to DEISA for computational and applications enabling resources. For example, particle and plasma physics projects (where there is a long tradition of cooperation at the European and international level) generally request many more resources on average per project than other applications areas. The materials science community, however, is much more disparate and submitted with a much larger number of smaller projects.

The following section gives an overview of the projects running on the DEISA infrastructure during 2010–2011 to show the range of science supported.

The 55 DECI projects selected for DECI-6 were self-categorised (by the PIs) into six broad scientific disciplines—Astronomical Sciences, Biological Sciences, Earth Sciences, Engineering, Materials Science and Plasma/Particle Physics.

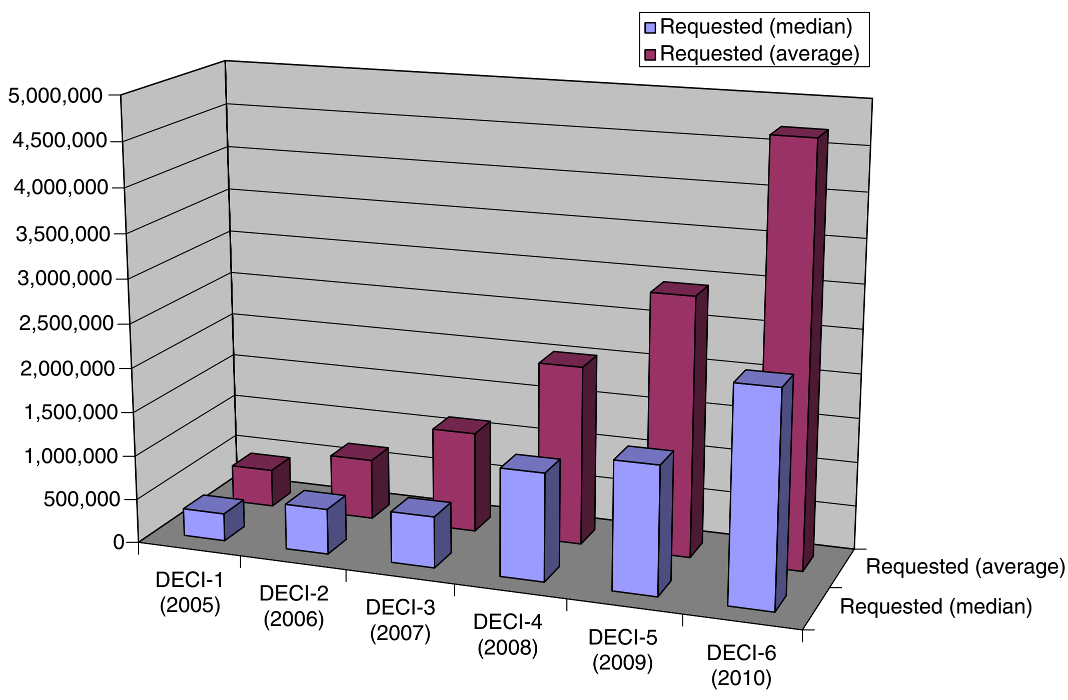


Fig. 2 Average and median amount requested during the lifetime of DECI

Overall, only 45% of the proposals were able to be accepted and the amount of resources awarded to many projects was scaled back to enable as many projects as possible to be supported. Figure 3 shows the resources awarded to each broad scientific discipline.

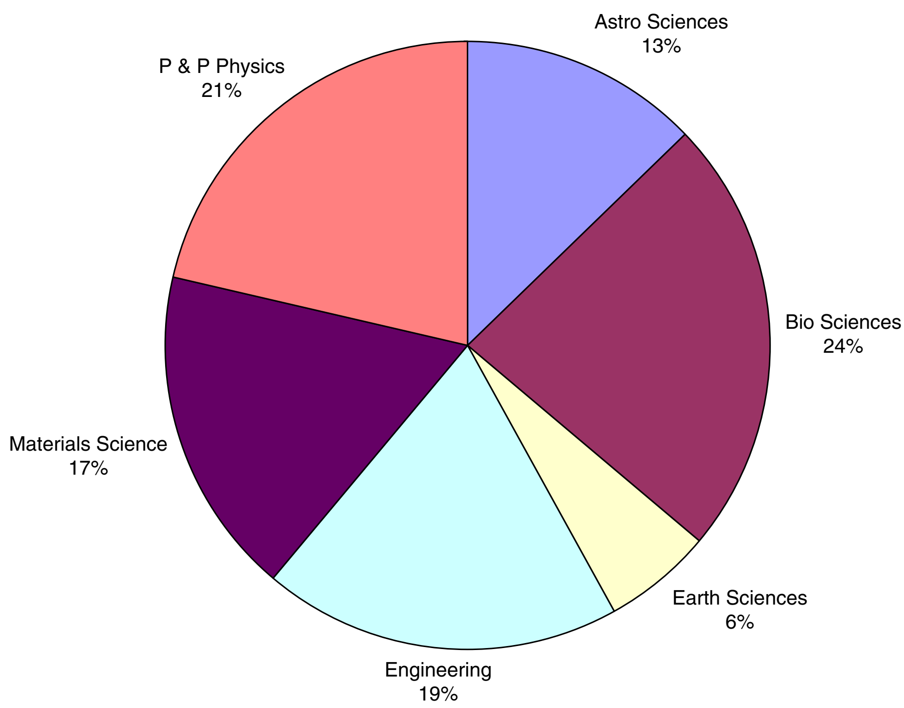
There were 151 investigators named in the 55 accepted projects. Of these investigators, 106 (70%) came from countries with a DEISA partner site, 30 investigators (20%) came from other European countries and 15 (10%) from large sites outside of Europe (China, Israel, Japan, Saudi Arabia, USA). This suggests that successful project proposals were generally based on collaborative European science. All of the 58 PIs (three projects had two PIs) were Europeans, offering further proof that the research being undertaken focuses on European research priorities, with 49 (84%) of the PIs coming from countries with DEISA sites and 9 (16%) coming from other European countries.

Since many of the countries with large computational research communities are DEISA partners, it is to be expected that a large majority of the PIs and Co-PIs will continue to come from countries with DEISA sites. However, calls are publicised extensively throughout Europe.

In an effort to find out more about the degree of co-operation and collaboration within DECI projects, we analysed the information on investigators in further detail. The number of investigators in DECI-6 projects varied from one to seven. During the 6 years in which DECI has been operating, the average number of investigators has risen steadily from 1 to 3.

Finally, we looked at the number of countries collaborating in successful proposals. This showed that 38 (69%) of the projects involved scientists from more than one country, with 14 (25%) involving scientists from three or more countries. This indicates that DECI is supporting pan-European scientific collaboration, and that the DEISA infrastructure is attractive to European researchers. Of the 364 scientific investigators who have used the DEISA infrastructure via DECI since 2005, 281 (77%) have been involved in one project to date with a further 83 (23%) involved in two or more projects. This shows that DEISA is being successful in reaching out to new groups and collaborations. It also

Fig. 3 Proportion of CPU resources awarded to each broad scientific discipline in DECI-6



suggests that DEISA resources are recognised as being complementary to national facilities, with applications only being made to DEISA when the circumstances warrant it.

Overall, the statistics which we have collected suggest that DECI is being successful in attracting high quality collaborative proposals involving scientists from more than one European country, and in involving partners from outside Europe where they add value to the collaboration (Fig. 4).

The 55 DECI-6 projects specified 51 (different) named codes which they wanted to use plus two unnamed "own" codes. Of the 50 codes which specified their communications method, 39 (78%) of the codes were parallelised with MPI, with a further 8 (16%) having hybrid OpenMP/MPI implementations and three (6%) using MPI and pthreads. Of the 51 named codes, 45 were used by just one DECI project, with a further three being used by two projects, two (Chroma and VASP) by three projects and one (NAMD) by five projects.

In particular, the large number of different codes used by DECI projects indicates the wide variety of science being undertaken via DECI and reinforces the need for close partnership with users to help them to achieve good performance via applications enabling and code tuning. There is obviously a strong demand for the sort of complementary computing support offered by DECI which crosses scientific domains.

4.2 DEISA Benchmark Suite

Applications benchmarking is an important tool in scientific and high performance computing as it enables the study of system performance under real-life workloads. The DEISA Benchmark Suite [6], designed to reflect the most commonly used scientific applications covering several scientific areas, is used to obtain a performance profile of the DEISA infrastructure.

With the opening up of the European HPC ecosystem to scientists from all countries resulting in more choice, the assessment of the real-life performance of HPC systems through application benchmarking is becoming more important

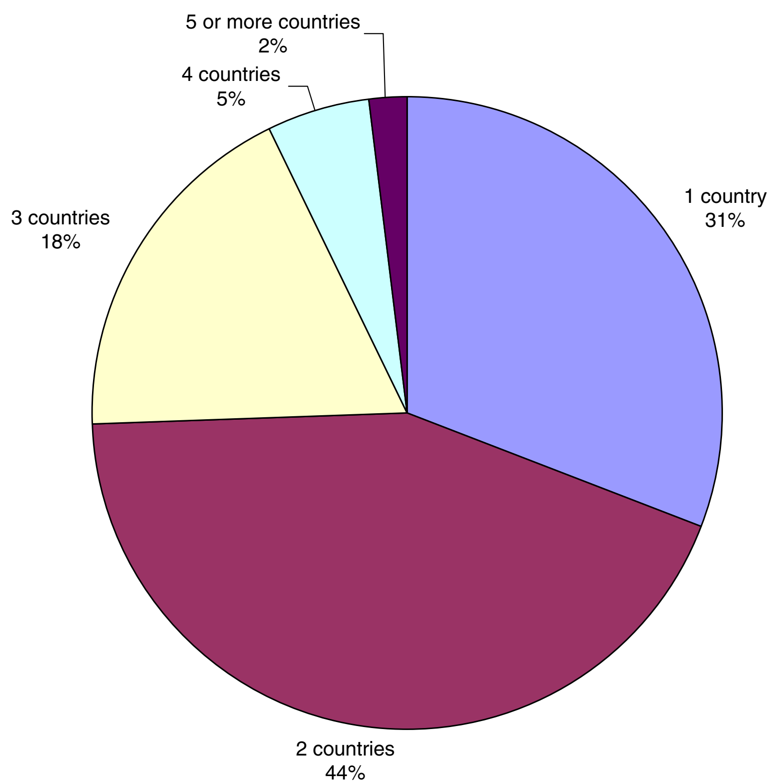


Fig. 4 Number of countries involved in each DECI-6 project

in helping to match scientific codes to the most appropriate applications platforms. A better understanding of how applications behave on different platforms will enable applications experts and scientific researchers to get the best performance from current systems, while an understanding of how the current generation of HPC systems is actually used and what performance can be achieved with real application codes informs the design of next generation many-core and exascale systems.

DEISA ports the applications in the Benchmark Suite to all the systems in the infrastructure. The applications are then run from the point of view of an average user of an HPC system (who would get basic help form a user guide) in order to reflect the performance such a user can expect to see. A recent survey carried out by DEISA among DECI users revealed that 73% of the users considered help with ‘code scaling and performance enhancements’ the most important aspect of the DECI initiative. Most users will not (or cannot, due to lack of knowledge and experience) perform any tuning or optimisation, but will largely rely on the default behaviour of a system.

One of DEISA’s interests in applications benchmarking and tuning lies in finding a good match between science and architecture, and thus providing the scientific community with the best and most effective service. However, results are also used to improve the focus of education and training needs so we can equip users with the knowledge to efficiently exploit the potential of the hardware.

4.3 Virtual Science Communities

In the second phase of the DEISA project the consortium has extended its service provisioning model from individual project support (i.e. DECI), to persistent service provision to some specific European user communities. A call for expressions of interest, published at the end of 2008, resulted in expressions of interest of several important communities from four different science areas: Fusion Energy, Cosmology/Space Science, Climate Research, and Life Sciences.

These Scientific Communities have asked for support in several ways: allocation of CPU cycles, access to state-of-the-art supercomputing architectures, help on technological requests, support on software porting and application enabling, and data repository management. Meanwhile, support for communities in all four science areas has been continued for the second year. The major difference to support individual projects via DECI is better sustainability, a longer-term commitment at higher flexibility: established science communities can better plan and determine on their own priorities how to make best use of the resources. As these communities traditionally consist of numerous pan-European collaborations and institutions we call them ‘virtual (science) communities’.

5 Services and Infrastructure

5.1 User and Application Support

An important resource of DEISA is the Applications Task Force, which helps scientists and virtual communities with providing user support and enabling of complex applications for effective use of DEISA resources. By “application enabling” we mean the following tasks:

- Provision of expertise and technical assistance for developing and porting new applications to new architectures, and for accessing new supercomputing architectures.
- Enabling activity to run large-scale computations on new architectures, helping scientists to make the necessary changes to the application to adapt it to each new system.
- Porting applications to relevant systems of existing and new architectures in the DEISA environment.
- Reengineering of applications to match the operating system (AIX, Linux, UNICOS/lc, SUPERUX), the CPU type (scalar, vector, single core, multi-core), the memory hierarchy (data caches, shared or distributed memory), and the interconnecting network.
- Adapting applications to multi-core systems: support for hybrid shared and distributed-memory approaches, such as OpenMP+MPI.
- Support for efficient and portable parallel I/O: analysis of the application’s I/O requirements

including data types, data distribution, and data access patterns; selecting, implementing and testing one of the standardized file formats such as HDF5 or NetCDF, depending on the requirements and community-specific preferences and conventions (optional: enabling conversion between old and new data formats).

- Optimization and reengineering parallel I/O in data intensive applications for target architectures, in case conversion to a portable data format is not a solution; optional implementation of highly optimised parallel I/O schemes for best adaptation to DEISA's global parallel file system.

The application enabling effort is a key concept in DEISA and it is estimated and quoted (usually in terms of person months) for every project or proposal. Due to the heterogeneity of the European trans-national supercomputing resources, any enabling effort must be coordinated among all the partners in an efficient way, according to the platform(s) where a given application is due to run. Once the resources are allotted, the Home Site (HS) monitors how these resources are used at the (maybe more than one) Execution Site(s) (ES). Both the HS and ES are in charge of adapting the application to the ES. Through frequent contacts with the project's PI, the enabling process develops until the application is running in production.

Since the early days of the second phase of DEISA, every selected DECI project represented an enabling effort to port, analyze, and optimize research codes in order to be adapted to the DEISA infrastructure, in areas as diverse as Materials Sciences, Biotechnology, Earth Sciences, or Nuclear Fusion.

Some cases are specially selected to go one step beyond: performance enhancement. As the applications are enabled, their parallel behaviour is assessed by probing them using performance analysis tools. Thanks to a long experience, the DEISA technical staff can detect opportunities for efficiency improvements in many cases. When the effort is reasonable according to the manpower available, these improvements are carried out in close collaboration with the code developers. This is the mission of the application and user support team. It is worth mentioning that every improvement is transferred to the next release of the enhanced code.

5.1.1 Helpdesk

A centralised service called the *DEISA Trouble Ticket System* has been set up to manage support requests, based on the *Request Tracker* (RT) software [7], which allows to record support requests and the responses to them by means of *tickets*, i.e. entries in a database with identification numbers to which the requestor can refer. This allows each user to access the full trace and history of the problems that they submitted and to staff members to have a complete overview of all the requests concerning the entire infrastructure. The preferred method of creating a ticket is via the web interface, although a generic email address is also available. Experience within DEISA has shown that managing support requests this way greatly facilitates the detection of common problems and allows the solution process to be monitored efficiently.

5.1.2 Documentation

DEISA documentation is freely available via the main website with the most important document for new users being the *DEISA Primer* [8]. Apart from this introductory document, there are also separate ones for more specific topics (see the related descriptions in Section 5.2 such as the DCPE, UNICORE, DESHL, interactive access, etc. Within DEISA, documentation is kept in a central repository with a unique version for each document. Project staff maintains the documents via the website using the *Plone* Content Management tool [9], thus allowing the possibility of instant publication. PDF versions of the *DEISA Primer* and the dedicated technical manuals, with a formatting especially adapted to ease the consultation of paper documents, are also available.

5.2 Operations and Technology

The infrastructure integrates services between some of the largest supercomputer systems in

Europe. These services include a dedicated high speed network between the systems, high performance data sharing and data transfer facilities, common tools for accessing the systems, interactively and by compute job submission, shared user accounts and a common programming environment.

The operations team is responsible for keeping these services available. In the first place it is the responsibility of the local teams to monitor and maintain the operational status of the services. However since the dependencies of these integrated services between sites a tight collaboration between those sites is required. Therefore, responsibilities, procedures, and internal services are defined.

The basic management structure is given in Fig. 5. Each site has a site representative, the so-called site deputy, who is responsible for the status of the DEISA services at the site. Any issues with the site can be discussed between the operations leader and the site deputy as a first escalation step. Six subtasks, five of which together cover the whole range of external services to be supported, have been defined:

(1) Network-Related Services
(2) Data-Related Services
(3) Compute-Related Services
(4) Authentication, Authorization and Accounting (AAA) Related Services
(5) User-Related Services
(6) Integration of new partners

Each of these subtasks (1–5) or service domains are coordinated by a subtask leader, who is responsible for the coordination of actions between sites for the services in a domain.

The integration subtask (6) is organised differently. This subtask is responsible for the integration of a new site (e.g. Site N in Fig. 5) into the infrastructure. For each new site a supporting site (e.g. Site 2 in Fig. 5) will be appointed from existing sites and this site will be responsible for the integration.

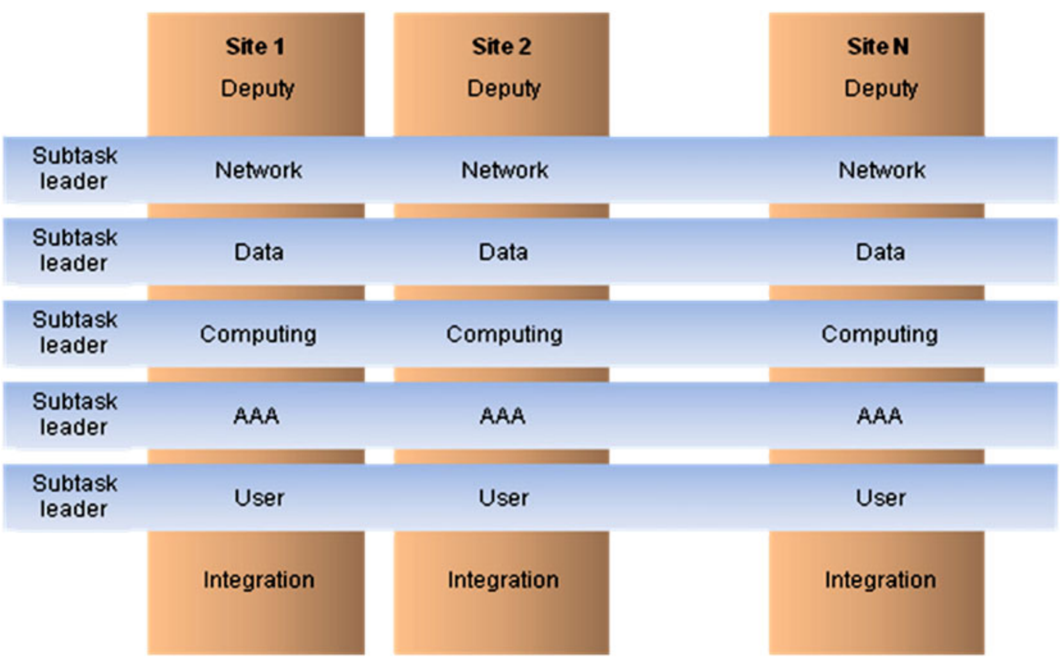


Fig. 5 Internal management structure of operations

Common to all subtasks is that for a high quality support well-defined procedures and internal services are needed. These services include monitoring facilities, a trouble ticket system and collaborative workspaces. Procedures include those defined for change management, the 'Operator on Duty task' and security methods.

Several collaborative facilities are in use to support the communication and information exchange between sites:

- BSCW [10] as the tool for document sharing between partners and the scheduling of meetings.
- The TWiki® software [11] for a wiki service with easy update facilities. Each staff member can update information, e.g. the configuration information for a site, scheduled maintenances, and changes in software/hardware.
- Subversion, a versioning system [12] for storing configuration files and tools, and as a traditional versioning system for software development. Additionally, the tool Viewvc [13] is installed which provides a web front-end to browse through the files stored by Subversion.

For the reporting and tracking of problems a centralized trouble ticket system is implemented [14]. This system is available both for staff and users. The difference is that staff members have more permissions than end users. This facility is based on the Request Tracker (RT) open source product [7].

Each week a site is responsible for overlooking the operational status of the infrastructure, known as the Operator on Duty (OoD) service. The main responsibilities are: signalling problems, the creation of trouble tickets, and the escalation of open problems without activity for defined periods.

Change management procedures are defined for changes in the production services, both hardware and software, and for new services, evaluated and prepared for production by the technology team. The basic steps of a change are:

- the preparation of the change
- the announcement of the proposed change with a planning of dates
- the acceptance of the proposed change
- the implementation of the change itself
- the closing of the change

The preparation phase includes: providing the documentation, also possible updates to user documentation, and for changes with a high impact the running of tests and a fallback plan in case the change is not successful. It is in the first place the responsibility of the subtask to decide on the impact of the change. Changes only at a single site, such as hardware or software changes, are

DEISA Common Production Environment

	BSC	CINECA SP6	CSC	ECMWF	EPCC HECTOR	FZJ JUGENE	FZJ JUROPA	HLRS LAKI	HLRS NEC-SX-8	HLRS NEC-SX-9	IDRIS BABEL	IDRIS VARGAS	LRZ	RZG GENIUS	RZG VIP	SARA
Applications	3/3	4/4	3/3	4/4	4/4	2/2	1/1	n/a	2/2	2/2	n/a	4/4	2/2	1/1	4/4	3/3
Compilers	4/4	4/4	4/4	4/4	4/4	4/4	4/4	7/7	3/3	3/3	4/4	4/4	4/4	4/4	4/4	4/4
Libraries	17/17	19/19	8/8	19/19	11/11	9/9	7/7	6/6	5/5	5/5	7/7	19/19	12/12	8/8	19/19	17/17
Shells	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
Tools	11/11	13/13	11/11	12/12	13/13	9/9	9/9	7/7	6/6	6/6	8/8	13/13	8/8	7/7	13/13	11/12

DEISA Globus

	BSC	CINECA SP6	CSC	ECMWF	EPCC HECTOR	FZJ JUGENE	FZJ JUROPA	HLRS	IDRIS VARGAS	LRZ	RZG VIP	SARA
Globus	1/1	1/1	1/1	1/1	1/1	1/1	1/1	n/a	1/1	1/1	1/1	1/1
Globus GSISSH	14/14	14/14	14/14	8/9	14/14	14/14	14/14	n/a	14/14	14/14	14/14	14/14
Globus GridFTP	13/13	13/13	13/13	8/10	13/13	n/a	13/13	10/10	13/13	13/13	13/13	13/13
Globus RFT	5/5	n/a	5/5	4/4	5/5	n/a	n/a	n/a	5/5	5/5	5/5	5/5
Globus WS_GRAM	11/12	n/a	11/12	9/10	11/12	n/a	n/a	n/a	11/12	11/12	11/12	11/12

Fig. 6 INCA summary screen for users

the responsibility of the site. In contrast, if the local changes have impact on the infrastructure services, then also the DEISA procedure must be followed. In the acceptance phase specifically the security risks of the change will be reviewed too.

The INCA tool [15] is used for providing monitoring information for most services. Part of the information is also available for the end users, so they also can see the status of software and services.

In Fig. 6 an example is shown of the Inca summary screen which is displayed for users. The first number in each cell gives the number of successful tests and the second number is the total number of tests run at the site as given on top of the column. The user can zoom in on the details by clicking on one of the cells, so for instance for the example it will be shown to which particular site the WS_GRAM service is failing (the summaries only show from which site the test is run). The full matrix of results for a particular test suite can be displayed too. For easy identification of problems the cells are coloured depending on the results, from dark green for all tests passing to red if all fails.

In Fig. 7 an example is shown of weekly success rates of the different INCA tests, both for the different systems and the middleware suites. The results for the last seven days, given by the top bar for each site, are also compared with those for the preceding week, the lower bar. In general the results are in a satisfying high nineties percent range. These figures are only available for staff members and can be used to get a quick overview of the general status of the infrastructure. Detailed history graphs for longer periods can be produced too.

Network monitoring information is managed separately. A full matrix of network availability and performance between all sites is maintained. The DEISA monitoring information is in addition to local monitoring facilities, which in general will be more detailed on a system level. The DEISA specific monitoring is needed to detect problems between sites.

Fig. 7 Example of INCA history

Security The relative open access to the infrastructure and the tight integration between sites require good coordination between sites about the security of the infrastructure and the set up of adequate procedures. Trust among the partners can only be maintained if well established procedures and policies exist.

At the basis of the collaboration on security is a security policy document, accepted by all partners. The main subjects of this document are:

- The handling of security incidents
- The risk review of changes in the infrastructure
- The auditing of the security set-up
- The roles and responsibilities of persons and teams

The document only describes the basic requirements and policies. The details of the implementation of the different tasks to fulfil the policies are described in separate documents. Although there is trust between partner sites it is important that there is a tight collaboration between the security experts of the sites. Objectives of this collaboration are to enhance the understanding of each others policies and to further develop the common policies. Two teams are defined:

- The Operational Security Coordination Team (OSCT) with as main objective to review the long term risks of services and their configurations.
- The Computer Security Incidence Response Team (CSIRT) for the coordination of all incoming incident reports.

Furthermore, the need for and content of other policy documents is reviewed by a working group, which presents proposals to all site representatives. Examples are the Acceptable Use Policy (AUP) for users and the policies for the registration of users. Site security officers and site representatives meet twice a year to discuss the status of the security activity.

Technology In order to keep updated, to enhance and empower the DEISA infrastructure, it is mandatory to have an action concerning the technologies assessment, improvements or deployment of additional services and functionalities. This is the responsibility of the technology team.

The technology activity is driven by both user requirements and technological innovation. Software requests originating from requirement by users and virtual communities are being evaluated, assessed, installed, tuned for the DEISA infrastructure, and finally deployed on it. Technology innovation-driven refers to new software releases of production services and new software technologies that are being investigated. Objectives are to fulfil users' needs, to update the infrastructure, to deploy new enabling technologies, and to simplify infrastructure usage.

In order to establish this approach the following methodology based on three main pillars has been employed: *Objectives, Structural Organization,* and *Means*.

Objectives There are five objectives which have to be addressed:

1. Monitor those technologies already in use in the project,
2. To identify and select those technologies of potential interest to the project,
3. Evaluate technologies for pre-production deployment purposes,
4. Plan and design specific sub-infrastructures to deliver a new service by using the selected technologies,
5. Deploy and test the specific sub-infrastructures on the DEISA production service.

Structural Organization As shown in Fig. 8, the HPC ecosystem has been split into five technological areas: Network, Data, Compute, AAA, and User related technologies, which directly match the five subtasks and services of Operations (defined above), to underline the coherency and strong inter-operation between Operations and Technology. Thus, according to a well-defined process among Technology and Operations, new technologies which have been evaluated and published by the technology team, are then handed over to the operations team for implementing it as a new production service. In addition, the

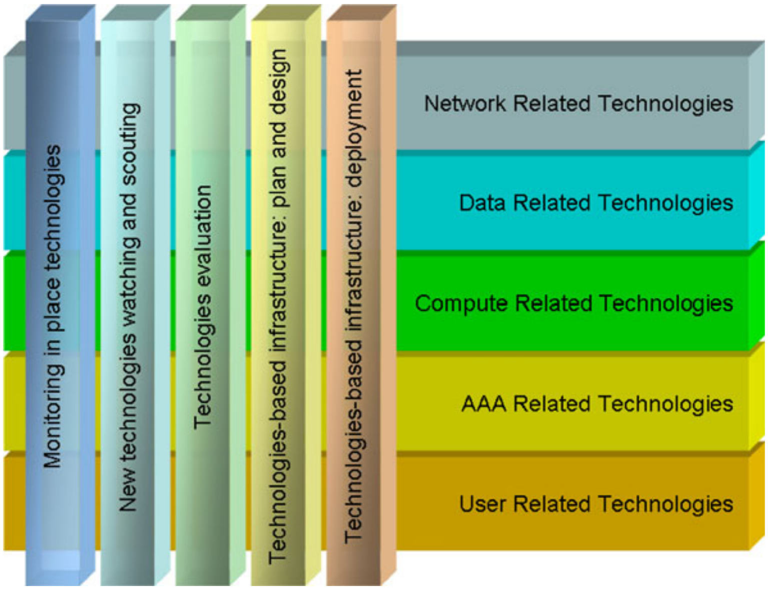


Fig. 8 Objectives in context of the major technological areas of DEISA

separation into five distinct technological areas also allows assigning tasks to those DEISA partners who have an in-depth knowledge about a specific area. As Fig. 8 reveals, the objectives stated in the previous paragraph can be applied to all shown technological areas.

Means The second phase of DEISA uses a project-based methodology to explore the spectrum of potential technologies that may meet the DEISA objectives for improving the European HPC Infrastructure production service. The operations team adopted a procedure-based approach that is necessary for a day-to-day service-oriented activity. Consequently, tasks are organised by means of "Internal Specific Targeted Projects" (ISTPs). ISTPs can be broadly classified as:

1. Simple technology evaluation for addressing objective 1 or 2 above;
2. Technology evaluation with subsequent planning and design of a sub-infrastructure to deliver a new (or upgraded) service for addressing above objectives 1 or 2, 3 and 4;
3. Technology evaluation with planning and design of a sub-infrastructure to deliver a new (or upgraded) service and sub-infrastructure deployment.

A rigorous approach based on the ISTP mean is necessary to ensure that most (if not all) technology, infrastructure, and service-related aspects are handled and that clearly identifiable results are obtained and made available to all DEISA activities. For these reasons each ISTP contains at least one mandatory deliverable. These internal projects, their work breakdown structure, Gantt charts, and documentation are not available externally to the project but their major findings are being made publicly available via deliverables. Each ISTP has a project manager. In the past, to achieve more successful outcomes and to speed up projects DEISA partners have worked closely with companies or technology providers to adapt their software to their EU projects requirements. For instance, DEISA and Platform Computing customized and extended the LSF batch system in order to explore possibilities for HPC resource co-reservation and co-allocation. A similar approach has been adopted by the DEISA ISTPs. As an example, the DEISA2 project supported IBM for the evaluation of GPFS WAN-caching functionality.

5.3 Production Services

A classification of the production services in three service levels is used: Core Services, Additional Services, and Optional Services, having the following characteristics:

- The *Core Services* are implemented and supported at each site, where technically possible. They are highly robust and will persist in the long term. DEISA provides a high support level with a high priority to solve the reported problems.
- The *Additional Services* are also supported at every DEISA site if technically possible, but are services for which not the same level of reliability and long term availability can be guaranteed as for the Core Services. They are supported on a best effort basis.
- The *Optional Services* might be supported on each DEISA platform, but are not guaranteed to be supported in the near future, and have a low priority level of support.

The list of Core Services (further details in [16]) is

- The network services, including a dedicated network between sites, provided by the GÉANT infrastructure;
- The data services provided by Multi-Cluster GPFS and GridFTP; The GPFS service provides users with a shared file system on most systems. GridFTP provides a high bandwidth data transfer service between sites and external locations.
- The services provided by the batch scheduler available on each supercomputer;
- The Authentication, Authorization and Accounting (AAA) services provided by GSI-SSH for interactive access, the DEISA User Administration System (DUAS), the DEISA Accounting Report Tool (DART) and the shared *Public Key Infrastructure* (PKI) based on IGTF [17] membership;
- The User services provided by the DEISA Common Production Environment (DCPE), the monitoring framework for applications availability (INCA) and the DEISA Trouble Ticket System (DTTS) to report and track problems.

The Additional Services are UNICORE and DESHL (a command-line interface and API to UNICORE), uniform job and workflow managers which allow submitting jobs on heterogeneous platforms in a common and seamless way. Application workflows form one large class of applications which are well suited for Grids, see e.g. [18].

The Optional Services are the SRB and iRODS data management tools, and the RFT (Reliable File Transfer) tool. For the compute services, WS-GRAM from the Globus Toolkit, which allows submitting jobs on heterogeneous platforms in a uniform way, is available on some sites. For the AAA services, the SSH and UNICORE-SSH, alternative tools to provide an interactive access to the platforms, are available in some places.

5.4 DEISA Common Production Environment

The unified *DEISA Common Production Environment* (DCPE) is another key feature in DEISA. It defines a coherent set of software accessible on the various sites of the infrastructure and offers a common interface to the users, independent of the target platform employed. The main components of the DCPE are:

- A coherent set of software packages for each architecture, called *the Software Stacks*;
- A uniform interface to access the software on all the platforms, provided by the *Modules* tool [19] ;
- A framework to monitor the software, provided by the *INCA* tool [15].

The *Modules* tool interface allows users to execute commands and to define scripts in a portable way, without having to adapt them from one site to the other according to the local procedures chosen to install the software. In the context of DEISA, we have added some extra features to prevent some common usage errors, helping users to keep a coherent software environment. This framework has been fully appreciated by the users and has motivated several sites to offer locally a DEISA-like *Modules Environment* on their platforms for all their local users.

To continuously monitor the status of the software environment on all our platforms, we are using the *INCA* tool (*Test Harness and Reporting Framework*) [15], initially developed by the *San Diego Supercomputer Center* for the TeraGrid project [20]. *INCA* is specifically designed to periodically run a collection of validation scripts to frequently check the version of the software installed and the availability and the correct operation of them.

6 Collaborations and Interoperability

The DEISA Consortium is collaborating on an international level with leading organisations and projects in the HPC and Grid area, and is highly engaged in efforts towards interoperability of leading compute infrastructures and standardisation of their way of usage.

6.1 Collaborations

6.1.1 PRACE

In Europe, DEISA is collaborating with PRACE to establish an HPC ecosystem of Tier-0 and national Tier-1 supercomputers. PRACE, the Partnership for Advanced Computing in Europe, is the EU member countries' approach towards a persistent European HPC ecosystem, and has started with the installation of leadership supercomputers, so-called Tier-0 systems, with a Petaflop/s system at FZJ in Jülich, Germany. National supercomputers, so-called Tier-1 systems, on which the DEISA infrastructure is built on, will become integrated under one umbrella. The well established and production oriented DEISA services will play a key role for the operation of the distributed infrastructure.

6.1.2 European E-Infrastructure Forum

DEISA is an active member of the European E-Infrastructure Forum [3] which develops a common view and language about future unified European e-Infrastructures, together with its partner projects DANTE, EGEE, PRACE, TERENA and the more recently formed European Grid Initiative (EGI) [R].

6.1.3 TeraGrid

DEISA is traditionally closely collaborating with the NSF project TeraGrid, the US counterpart of DEISA, [21]. TeraGrid started in 2001 with a different business model concerning hardware funding, and continues until 2011 with a planned follow-on called Extreme Digital (XD). A special collaboration between DEISA and TeraGrid was on high-performance wide-area global file systems, [21].

6.1.4 International Exascale Software Project

DEISA is participating in IESP, the International Exascale Software Project, a world-wide effort to address software bottlenecks which prevent the effective usage of next generation supercomputers, [22].

In the meantime the European Exascale Software Initiative (EESI) started to build a European vision and roadmap to address the challenge of Petascale and Exascale computing and simulation, and to coordinate the European contributions for future Exaflop/s computers.

6.2 Interoperability and Standardisation

6.2.1 Open Grid Forum

DEISA's major standardization needs are related to High Performance Computing, job submission, job and workflow management, data management, data access and archiving, networking and security (including authentication, authorization, and accounting).

DEISA participates in Open Grid Forum (OGF, [23]) standardization activities, such as Job Submission and Description Language (JSDL) and the OGSA-Basic Execution Services (BES) for job submission, Usage Resources (UR), Resource Usage Service (RUS) for accounting purposes and Database Access and Integration (DAIS) for data-oriented Web Services. In addition, DEISA actively contributes to GLUE2 significantly increasing its abilities to describe state-of-the-art HPC systems. In terms of network, DEISA is active in the Firewall Research Group (FI-RG), Firewall Virtualization for Grid

Applications (FVGA), as well as the Grid High Performance Networking group (GHPN). Further efforts are taking place in the Network Mark-Up Language (NML), Network Measurements (NM), and Network Service Interface (NSI) groups.

Under the Grid Interoperation Now (GIN) umbrella, DEISA contributes to the continuous interoperation efforts in using the High Performance Computing Basic Profile (HPC-BP) where multiple different international middleware systems and infrastructures exchange jobs. Insights of this complement the lessons learned from DEISA in terms of using standards in a large-scale HPC infrastructure that in turn are given as an input to the Production Grid Infrastructure (PGI) working group in order to actively contribute to the evolution of open standards in different technical areas. For more details see [24].

6.2.2 Infrastructure Policy Group

DEISA is a member of the Infrastructure Policy Group IPG, comprising leading infrastructure projects worldwide (TeraGrid, OSG, EGEE, DEISA, NAREGI). The aim is to collaborate towards the goal of seamless interoperation of leading Grid Infrastructures world-wide. Regular meetings have taken place co-scheduled with OGF. Topics dealt with so far are:

- Authentication, authorization, accounting, and auditing
- Resource allocation policies and procedures
- Handling of security incidents
- Portals and science gateways
- Data services

The policy documents are available at the Infrastructure Policy Group's website at [25].

7 Sustainability of the DEISA e-Infrastructure

After more than 6 years of successful operation, the DEISA Consortium has developed, enhanced and optimized the distributed European Tier-1 compute infrastructure and its services which are indispensable for the effective support of computational sciences in the HPC area: applications enabling and user support, operation, and technology testing and deployment. The service provisioning model has been extended from one that supports single projects to one supporting European Science Communities.

One of the main goals of DEISA has been to become a world-class e-Infrastructure as part of a strategy to achieve European leadership in e-Science. To achieve this, one of the main requirements is to ensure long-term sustainability of its main results and achievements:

- An operational, distributed European HPC infrastructure ready for use in the future European HPC ecosystem,
- Well established European expert teams able to provide the necessary services,
- Support for grand challenge computational science projects
- European Science Communities benefiting from a European HPC infrastructure

To reduce or eliminate existing technological, cultural, economic, and political barriers and to foster wider acceptance of the European HPC infrastructure, DEISA is ensuring the sustainability of the different areas: technology, infrastructure, operations and support services, expertise, communities, and collaborations. For these areas, the DEISA Consortium has identified the following assumptions and a wide variety of actions and measures:

Technology and Infrastructure

- The DEISA infrastructure is built on existing, proven, sustainable technology components, including: the GÉANT2 and NRENs based high performance network; access to all major types of state-of-the-art supercomputers, based on national supercomputer services; a homogenized global software environment on top of the heterogeneous HPC architectures.
- DEISA2 has continued, consolidated and extended the pioneering work of the former DEISA project, to deliver and operate a European supercomputing infrastructure and related services, ready for use in a European HPC ecosystem.

Operations and Services

- DEISA operations benefitted from the many-years operations of the individual European supercomputers centres. The gained experience in the collaborations has improved the operational services of the national HPC centres.
- For the effective support of computational sciences in the HPC area, activities relevant for applications enabling, operation, and technologies have been developed and further enhanced

Expertise

- DEISA stimulated tight collaborations of the expert groups in the different HPC centres: existing and developing expertise within the individual centres has been united through DEISA and will be provided in the future to the wider European HPC communities.

Extreme Computing and Science Communities The DEISA Extreme Computing Initiative (DECI, [5]), launched in 2005, supports the most challenging supercomputing projects in Europe which require the special resources and skills of DEISA [26].

- In 2008, the DECI service provisioning model based on single project support has been extended to the support of European Science Communities.
- For supporting world-wide Science Communities across existing political boundaries, DEISA participates in the definition, evaluation and implementation of standards for interoperation.
- Short- to mid-term, DEISA has developed a collaboration strategy to encourage European users to use its HPC infrastructure, and to encourage non-European users through DECI to jointly apply, together with their European research colleagues, for HPC resources and related services.

Collaborations

- Collaborative activities are carried out with related European and other international initiatives.
- Most important is the collaboration with the PRACE project [27]. PRACE has already started to install a limited number of leadership-class Tier-0 supercomputers in Europe.
- Emphasis has been put on contacts to research infrastructure projects established by the European Strategy Forum on Research Infrastructures (ESFRIs), and the European Grid projects such as EGEE and EGI.
- DEISA has contributed to the creation of the European E-Infrastructure Forum [3] with the leading European research infrastructure projects as members, with the goal to facilitate seamless interoperation of the various existing compute and network infrastructures.
- The activity has reinforced the relations to other European HPC and leading international HPC centres and initiatives world-wide, such as TeraGrid and NAREGI. These successful collaborations with existing infrastructure projects around the world have driven the interoperability of the different infrastructure services, enabling users world wide to flexibly use the best suited resources for solving scientific grand challenge problems.

In summary, the DEISA Consortium has initiated and contributed to the awareness of the need for a persistent European HPC infrastructure as recommended by the ESFRI in its 2006 report [28]. As a consequence, the European Commission has launched a call for the continuation of the DEISA services within the upcoming HPC ecosystem to be implemented by PRACE.

Acknowledgements This work is supported by the DEISA project, funded through the European Commission contracts RI-508830, RI-031513 and RI-222919.

References

1. Floros, E., Germain-Renaud, C., Harris, F., Lorenzo, P. M. (eds.): Special issue on EGEE applications and

supporting Grid technologies. J. Grid Computing. **8**(2), (2010)
2. Kacsuk, P., Kovacs, J., Farkas, Z.: Attila csaba marosi and gabor gombas, et al., SZTAKI Desktop Grid (SZDG): a flexible and scalable desktop Grid system. In: Kondo, D. (ed.) Special Issue: Volunteer Computing and Desktop Grids. J. Grid Computing., vol. 7(4), pp. 439–461 (2009)
3. European E-Infrastructure Forum: http://www.einfrastructure-forum.eu (2009, EEF)
4. Gentzsch, W., Kennedy, A., Lederer, H., Pringle, G., Reetz, J., Riedel, M., Schuller, B., Streit, A., Wolfrath, J.: DEISA: e-Science in a collaborative, secure, interoperable and user-friendly environment. In: Proceedings of the e-Challenges Conference e-2010. See also the DEISA Website at www.deisa.eu
5. DEISA Extreme Computing Initiative, DECI: www.deisa.eu/deci. Accessed Jan 2011
6. DEISA Benchmark Suite: http://www.deisa.eu/science/benchmarking. Accessed Jan 2011
7. Vincent, J., Spier, R., Rolsky, D., Chamberlain, D., Foley, R., Essentials, R. T.: O'Reilly Media (2005)
8. DEISA Primer: http://www.deisa.eu/usersupport/primer. Accessed Jan 2011
9. Plone Content Management tool: http://plone.org. Accessed Jan 2011
10. BSCW—Basic Support for Cooperative Work: http://public.bscw.de/en/index.html. Accessed Jan 2011
11. TWiki®: The Open Source Enterprise Wiki and Web 2.0, Application Platform. http://twiki.org/. Accessed Jan 2011
12. Subversion: http://subversion.tigris.org. Accessed Jan 2011
13. Viewvc: http://www.viewvc.org/. Accessed Jan 2011
14. DEISA Trouble Ticket System (DTTS): https://tts.deisa.eu/. Accessed Jan 2011
15. INCA monitoring tool: Test Harness and Reporting Framework. http://inca.sdsc.edu/drupal/home. Accessed Jan 2011
16. DEISA user support: www.deisa.eu/usersupport. Accessed Jan 2011
17. IGTF: http://www.igtf.net/. Accessed Jan 2011
18. Bittencourt, L.F., Madeira, E.R.M.: Towards the scheduling of multiple workflows on computational Grids. J. Grid Computing. **8**(3), 419–441 (2010)
19. Modules: http://modules.sourceforge.net/. Accessed Jan 2011
20. The TeraGrid project: http://www.teragrid.org/. Accessed Jan 2011
21. Andrews, P., et al: Exploring the Hyper-Grid Idea with Grand Challenge Applications: The DEISA-TERAGRID Interoperability Demonstration, pp. 43–52. CLADE, IEEE (2006)
22. The Exascale Initiative: http://www.exascale.org. Accessed Jan 2011
23. OpenGridForum: http://www.ogf.org. Accessed Jan 2011
24. DEISA Interoperability: http://www.deisa.eu/services/interop. Accessed Jan 2011
25. The Infrastructure Policy Group: http://forge.gridforum.org/sf/wiki/do/viewPage/projects.ipg/wiki/HomePage. Accessed Jan 2011
26. DEISA—Advancing Science in Europe: http://www.deisa.eu/press/Media/DEISA-AdvancingScienceInEurope.pdf. Accessed Jan 2011
27. PRACE project: http://www.prace-project.eu. Accessed Jan 2011
28. ESFRI 2006: http://cordis.europa.eu/esfri/. Accessed Jan 2011