

DEISA : status, strategies, perspectives



Towards cooperative extreme computing in Europe

Victor Alessandrini
IDRIS - CNRS
www.deisa.org



DEISA Symposium
Bologna, May 4-5, 2006.

Victor Alessandrini, IDRIS - CNRS



DEISA objectives



- *To enable Europe's terascale science by the integration of Europe's most powerful supercomputing systems.*
- *Enabling scientific discovery across a broad spectrum of science and technology is the only criterion for success*
- **DEISA is an European Supercomputing Service built on top of existing national services. This service is based on the deployment and operation of a persistent, production quality, distributed supercomputing environment with continental scope.**
- **The integration of national facilities and services, together with innovative operational models, is expected to add substantial value to existing infrastructures.**
- **Main focus is High Performance Computing (HPC).**

Participating Sites



BSC	<i>Barcelona Supercomputing Centre</i>	Spain
CINECA	<i>Consorzio Interuniversitario per il Calcolo Automatico</i>	Italy
CSC	<i>Finnish Information Technology Centre for Science</i>	Finland
EPCC/HPCx	<i>University of Edinburgh and CCLRC</i>	UK
ECMWF	<i>European Centre for Medium-Range Weather Forecast</i>	UK (int)
FZJ	<i>Research Centre Juelich</i>	Germany
HLRS	<i>High Performance Computing Centre Stuttgart</i>	Germany
IDRIS	<i>Institut du Développement et des Ressources en Informatique Scientifique - CNRS</i>	France
LRZ	<i>Leibniz Rechenzentrum Munich</i>	Germany
RZG	<i>Rechenzentrum Garching of the Max Planck Society</i>	Germany
SARA	<i>Dutch National High Performance Computing and Networking centre</i>	The Netherlands

The DEISA supercomputing environment

(21.900 processors and 145 Tf in 2006, more than 190 Tf in 2007)

- IBM AIX Super-cluster
 - FZJ-Julich, 1312 processors, **8,9 teraflops peak**
 - RZG – Garching, 748 processors, **3,8 teraflops peak**
 - IDRIS, 1024 processors, **6.7 teraflops peak**
 - CINECA, 512 processors, **2,6 teraflops peak**
 - CSC, 512 processors, **2,6 teraflops peak**
 - ECMWF, 2 systems of 2276 processors each, **33 teraflops peak**
 - HPCx, 1600 processors, **12 teraflops peak**
- BSC, IBM PowerPC Linux system (MareNostrum) 4864 processeurs, **40 teraflops peak**
- SARA, SGI ALTIX Linux system, 416 processors, **2,2 teraflops peak**
- LRZ, Linux cluster (2.7 teraflops) moving to SGI ALTIX system (5120 processors and **33 teraflops peak in 2006, 70 teraflops peak in 2007**)
- HLRS, NEC SX8 vector system, 646 processors, **12,7 teraflops peak.**
- **Systems interconnected with dedicated 1Gb/s network – currently upgrading to 10 Gb/s – provided by GEANT and NRENs**

DEISA activities and technology leadership



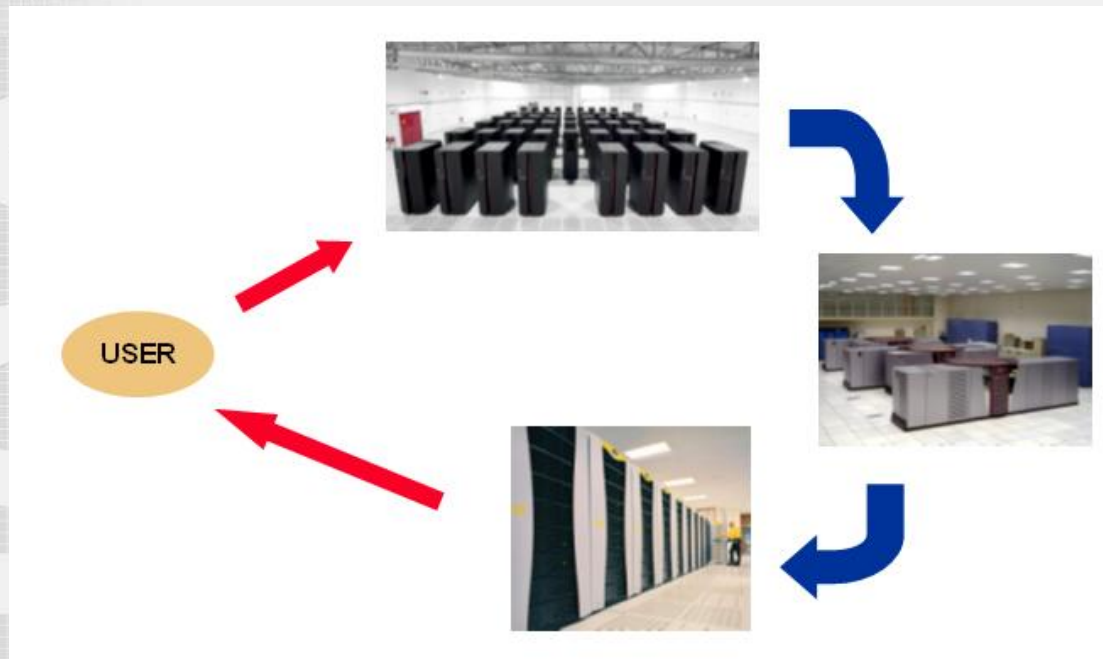
- Service provision model is a trans-national extension of national supercomputing centres operational model
- Service Activities based on trans-national working teams with distributed leadership:
 - Network Deployment and Operation R. Niederberger, FZJ (Germany)
 - Infrastructure Operation A. Schott, RZG (Germany)
 - Global File Systems. S. Heinzl, RZG (Germany)
 - Middleware P. Malfetti, CINECA (Italy)
 - User Support D. Girou, IDRIS (France)
 - Applications Enabling H. Lederer, RZG (Germany)
 - Security J. Wolfrat, SARA (The Netherlands)
 - JRA on Middleware R&D T. Sloan, EPCC (UK)
- Scientific Joint Research Activities
 - Materials Sciences, Plasma Physics, Life Sciences, CFD, Cosmology, Coupled Applications
- The Extreme Computing Initiative

How is DEISA enhancing HPC services in Europe?



- **Running larger parallel applications** in individual sites, by a cooperative reorganization of the global computational workload on the whole infrastructure, or by the operation of the **job migration service** inside the AIX super-cluster.
- Enabling **workflow applications** with UNICORE (complex applications that are pipelined over several computing platforms)
- Enabling coupled multiphysics Grid applications (when it makes sense)
- Providing a **global data management** service whose primordial objectives are:
 - Integrating distributed data with distributed computing platforms
 - **Enabling efficient, high performance access to remote datasets** (with Global File Systems and stripped GridFTP). We believe that this service is critical for the operation of (possible) future European petascale systems
 - Integrating hierarchical storage management and databases in the supercomputing Grid.
- **Deploying portals** as a way to hide complex environments to new users communities, and to interoperate with another existing grid infrastructures.

Basic services: workflow simulations using UNICORE



UNICORE supports complex simulations that are pipelined over several heterogeneous platforms (workflows).

UNICORE handles workflows as a unique job and transparently moves the output – input data along the pipeline.

UNICORE clients that monitor the application can run in laptops.

UNICORE has a user friendly graphical interface. DEISA has developed a command line interface for UNICORE.

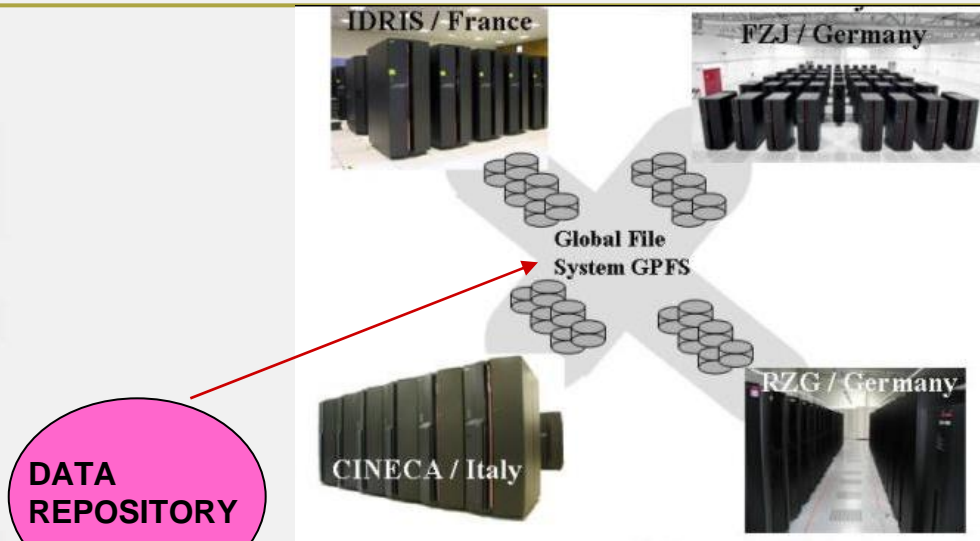
**UNICORE infrastructure including all sites has full production status.
It has proven to be very stable during the last few months.**

Basic services



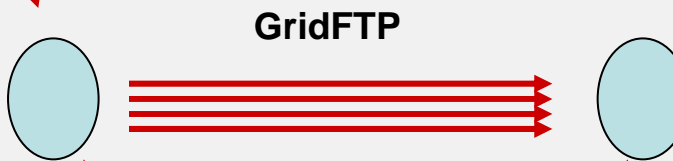
- **Job migration inside the AIX super-cluster.** Based on LoadLeveler Multi-Cluster, it allows system administrators to reroute jobs to other sites, in a way transparent for the end users. Used to move away simple jobs of « implicit users » to make place for a bigger application in a site. **Full production status.**
- **Co-allocation.** We are starting to prepare a first generation co-allocation service on the full heterogeneous infrastructure, using LSF Multi-cluster. Important for coupled Grid applications and for data movement. **Service in development phase, prototype expected in 6-9 months**
- **Remote I/O using Global File Systems and fast data transfers.** See next transparency
- **Integrating hierarchical data management and databases in the supercomputing Grid.** **In progress**
- **Deploying and monitoring a Common Production Environment.** **Operational over the whole infrastructure.**
- **DEISA training.** Six DEISA training session scheduled during the next 18 months.

Accessing remote data: high performance remote I/O and file transfer

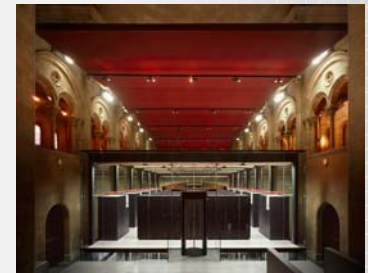


Remote I/O with global file systems **implicitly** moves data across platforms (in production today)

DEISA will also deploy **explicit** high performance data movers, using a parallel implementation of GridFTP



Co-scheduled, parallel data mover tasks



DEISA Global File System integration in 2006

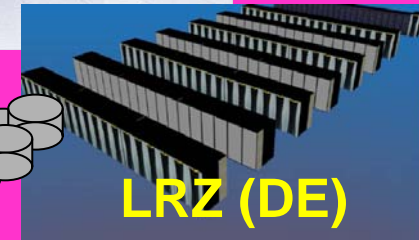
(based on IBM's GPFS)



AIX IBM domain



Linux SGI



High Performance Common Global File System
various architectures / operating systems
High bandwidth (up to 10 Gbit/s)



LINUX Power-PC

Global File System Interoperability demo during Supercomputing Conference 2005 in Seattle

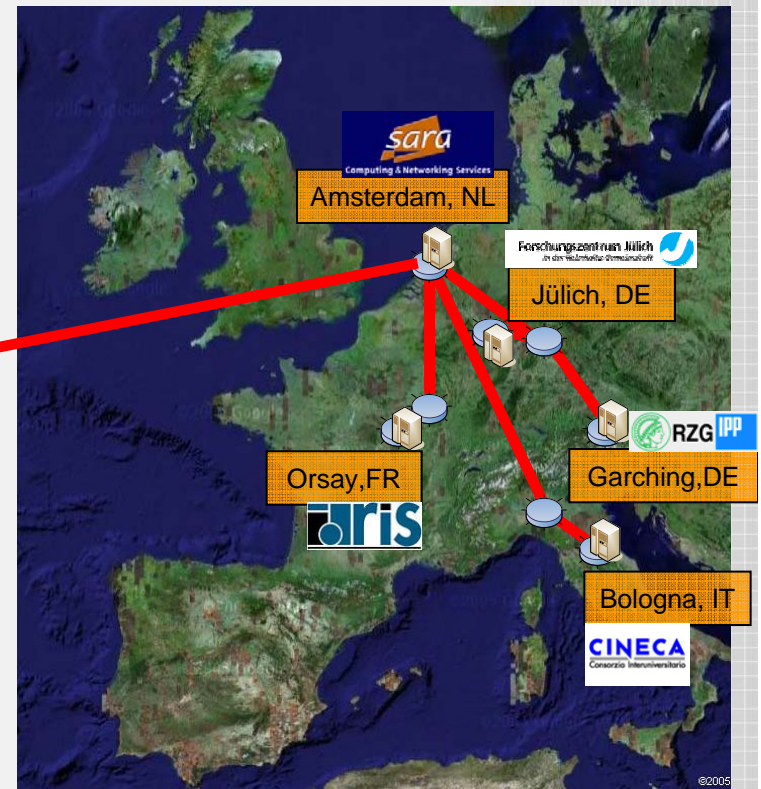


American and European supercomputing infrastructures linked: bridging communities with scalable, wide-area global file systems

TeraGrid Sites



DEISA Sites



Enabling science



- The basic service providing model for scientific users is the **Extreme Computing Initiative** (see www.deisa.org)
- Identification, deployment and operation of a number of « flagship » applications requiring the infrastructure services, in selected areas of science and technology.
- European Call for proposals in May-June every year. Applications are selected on the basis of scientific excellence, innovation potential and relevance criteria, with the collaboration of the HPC national evaluation committees.
- There are 29 projects in operation in 2005 – 2006 (see www.deisa.org)
- Supported by the **Applications Task Force** (ATASKF), whose objective is to enable and deploy the Extreme Computing applications. The activities of the ATASKF are focused on:
 - Hyperscaling of huge parallel applications, data oriented applications
 - Workflows and coupled applications
 - **Production of an European Benchmark Suite for HPC systems** (in collaboration with the HPC-EUR initiative, to be used in future procurements of European supercomputers).

Prospective issues



- **DEISA is deeply concerned with three strategic issues:**
 - *How should HPC be structured in Europe, to further enhance capabilities*
 - *How to move towards persistent, sustainable research infrastructures*
 - *How to make of DEISA non exclusive but efficient organization.*

How should a global HPC infrastructure in Europe look like?



- Emerging concensus about a global, strongly coupled infrastructure « à la DEISA » integrating the current national terascale systems (Tier 2) with new generation petascale systems (Tier 1) into a unique service providing environment, with a unique operational model.
- **Ideally, T1 should provide exceptional capability computational resources**, with a significant performance difference with respect to T2. We expect at any time a small number of T1 systems.
- Ideally, T1 systems should be shared and European. DEISA is working with the HPPC-EUR initiative, that is building a model and a proposal for that.
- A basic requirement is **cooperative operation of T1 and T2** including high performance access to existing national data repositories. This is why the data management services that DEISA is deploying are highly relevant.
- Ideally, exceptional resources will be concentrated in a very limited number of sites, but **services, know how, competence and leadership should remain distributed**.
- The boundaries of this « HPC island » in a global European infrastructure ecosystem follow from its unique operational and service provisioning models. **Interoperability with the rest of the ecosystem is a major priority that is already being addressed.**

Comments



- Funding and operational models for T1 have not yet been completely resolved. This is a very challenging issue for the future of HPC in Europe . **The eIRG contribution to this discussion could be important.**
- DEISA will try to evolve, whenever possible, to a consortium of national HPC funding organizations (rather than supercomputing centres) as a way to implement sustainability of the infrastructure.
- Partial steps have already been taken in this direction (example, the EPSRC partnership).
- This may also help to have a non exclusive organization, maintaining however a limited number of partners.
- The EU maintains a profound structuring action in all the eInfrastructures initiatives. The eIRG

Conclusions



- DEISA aims at deploying a **sustained and persistent**, basic European infrastructure for general purpose high performance computing.
- **We expect services and existing synergies to be persistent.** We do not claim persistency of the current organizational model. The DEISA Consortium is ready to adapt to the new FP7 strategies and establish a roadmap incorporating cooperation or merging with new HPC initiatives.
- Our next challenge is **establishing an efficient organization embracing all relevant HPC organizations in Europe.**
- One possibility is moving, whenever possible, from a consortium of HPC sites to a consortium of HPC national organizations.
- Interfaced with the other grid-enabled complementary infrastructures, DEISA expects to continue to contribute to a global European infrastructure for science and technology
- *Integrating leading supercomputing platforms with Grid technologies and reinforcing capability with shared petascale systems is needed to open the way to new research dimensions in Europe.*