DEISA and the Supercomputing View
Status of the European HPC Infrastructure

Vision of the European DEISA/eDEISA projects in FP6

To enhance Europe’s capability computing and science by the integration of Europe’s most powerful supercomputing systems Tier-1 in a European HPC e-infrastructure

DEISA built a European Supercomputing Service 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

Vision of the European Strategy Forum on Research Infrastructures (ESFRI)

Delivering a turnkey operational solution for a future persistent European HPC ecosystem integrating national Tier-1 centres and the new European Tier-0 centres
Vision of a European HPC Ecosystem in FP7

DEISA2 - The Infrastructure for the European HPC Ecosystem

Deep operational and technological integration of European HPC (T-0 and T-1) centres and systems
providing efficient seamless access to shared HPC resources and large data repositories
designing and approving an operational model for a large European Virtual HPC Centre.

Providing scientists access to a large distributed HPC environment via integrated services.
DEISA is paving the way to the efficient operation of the T-0 and T-1 ecosystem

PRACE
Building a world-class pan-European High Performance HPC
Ecosystem
which is operated under the umbrella of an European Legal Entity
adopting operational and technological concepts and services
designed and approved by DEISA2.
Towards a European HPC Infrastructure – DEISA2

Objectives
Enhancing the existing distributed European HPC environment (DEISA) to a turnkey operational infrastructure
Advancing the computational sciences in Europe by supporting user communities and extreme computing projects
Enhancing the service provision by offering a complete variety of options of interaction with computational resources
Integration of T1 and T0 centres
The Petascale Systems need a transparent access from and into the national data repositories
Bridging worldwide HPC projects
DEISA Partners

DEISA: May 1st, 2004 – April 30th, 2008

Three new partners joined June 2005 (eDEISA)

DEISA2: May 1st, 2008 – April 30th, 2011
## Partners/ Associate Partners

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSC</td>
<td>Barcelona Supercomputing Centre</td>
</tr>
<tr>
<td>CINECA</td>
<td>Consortio Interuniversitario per il Calcolo Automatico</td>
</tr>
<tr>
<td>CSC</td>
<td>Finnish Information Technology Centre for Science</td>
</tr>
<tr>
<td>EPCC</td>
<td>University of Edinburgh and CCLRC</td>
</tr>
<tr>
<td>ECMWF</td>
<td>European Centre for Medium-Range Weather Forecast</td>
</tr>
<tr>
<td>FZJ</td>
<td>Research Centre Juelich</td>
</tr>
<tr>
<td>HLRS</td>
<td>High Performance Computing Centre Stuttgart</td>
</tr>
<tr>
<td>IDRIS</td>
<td>Institut du Développement et des Ressources en Informatique Scientifique - CNRS</td>
</tr>
<tr>
<td>LRZ</td>
<td>Leibniz Rechenzentrum Munich</td>
</tr>
<tr>
<td>RZG</td>
<td>Rechenzentrum Garching of the Max Planck Society</td>
</tr>
<tr>
<td>SARA</td>
<td>Dutch National High Performance Computing</td>
</tr>
<tr>
<td>KTH</td>
<td>Kungliga Tekniska Högskolan</td>
</tr>
<tr>
<td>CSCS</td>
<td>Swiss National Supercomputing Centre</td>
</tr>
<tr>
<td>JSCC</td>
<td>Joint Supercomputer Center of the Russian Academy of Sciences</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSC</td>
<td>Spain</td>
</tr>
<tr>
<td>CINECA</td>
<td>Italy</td>
</tr>
<tr>
<td>CSC</td>
<td>Finland</td>
</tr>
<tr>
<td>EPCC</td>
<td>UK</td>
</tr>
<tr>
<td>ECMWF</td>
<td>UK (int)</td>
</tr>
<tr>
<td>FZJ</td>
<td>Germany</td>
</tr>
<tr>
<td>HLRS</td>
<td>Germany</td>
</tr>
<tr>
<td>IDRIS</td>
<td>France</td>
</tr>
<tr>
<td>LRZ</td>
<td>Germany</td>
</tr>
<tr>
<td>RZG</td>
<td>Germany</td>
</tr>
<tr>
<td>SARA</td>
<td>Netherlands</td>
</tr>
<tr>
<td>KTH</td>
<td>Sweden</td>
</tr>
<tr>
<td>CSCS</td>
<td>Switzerland</td>
</tr>
<tr>
<td>JSCC</td>
<td>Russia</td>
</tr>
</tbody>
</table>
The basic DEISA infrastructures and services

Dedicated high speed network infrastructure

Common AAA infrastructure

Global data management infrastructure

Integrating distributed data with distributed computing platforms, including hierarchical storage management and databases. Major highlights are:

- High performance remote I/O and data sharing with global file systems, using full network bandwidth
- High performance transfers of large data sets, using full network bandwidth

DCPE (DEISA Common Production Environment)

- The job management service
- The science gateways (portals) to supercomputing resources

Common Operation Environment

- Common monitoring and Information systems
- Common system operation
- Common help desk

Global Application Support
DEISA network infrastructure
H. Lederer and J. Wolfrat, DEISA

DEISA 1GE site
ECMWF via GEANT2 GRE tunnel
DEISA Data Services - Global File System
(based on IBM’s GPFS)

AIX IBM domain
- ECMWF (UK)
- RZG (DE)
- IDRIS (FR)
- CSC (FI)
- CINECA (IT)
- FZJ (DE)

Linux SGI
- SARA (NL)
- LRZ (DE)

Linux SGI
- IDRIS (FR)

HPC Common Global File System
various architectures / operating systems
High bandwidth (up to 10 Gbit/s)
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.
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 national HPC evaluation committees.

Supported by the Applications Task Force (ATASKF)

Hyperscaling of huge parallel applications, data oriented applications

Workflows and coupled applications

Production of an European Benchmark Suite for HPC systems
DEISA Extreme Computing Initiative

**DECI call 2005**
- 51 proposals, 12 European countries involved
- 30 mio cpu-h requested
- 29 proposals accepted, 12 mio cpu-h
  (standardized to P4+ at FZJ)

**DECI call 2006**
- 41 proposals, 12 European countries involved
- 28 mio cpu-h requested
- 23 proposals accepted, 12 mio cpu-h

**DECI call 2007**
- 63 proposals, 14 European countries involved
- (US, Canada, Brazil, Israel)
- 70 mio cpu-h requested
- 45 proposals accepted, 30 mio cpu-h
<table>
<thead>
<tr>
<th>Austria</th>
<th>Finland</th>
<th>France</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungary</td>
<td>Italy</td>
<td>Netherlands</td>
<td>Poland</td>
</tr>
<tr>
<td>Portugal</td>
<td>Romania</td>
<td>Russia</td>
<td>Spain</td>
</tr>
<tr>
<td>Sweden</td>
<td>Switzerland</td>
<td>UK</td>
<td></td>
</tr>
</tbody>
</table>

With collaborators from four other continents

| Asia | Australia | North America | South America |
The purpose of the DEISA2

Consolidation of the existing DEISA infrastructure

Guaranteeing the continuity of those activities and services that currently contribute to the effective support of world-leading computational science in Europe

*DEISA2 providing a lean and reliable turnkey operational solution for a persistent European HPC ecosystem*

Evolvement this European infrastructure towards a robust and persistent European HPC ecosystem

Enhancing the existing services, by deploying new services including support for European Virtual Communities, and by cooperating and collaborating with new European initiatives, especially PRACE that will enable shared European PetaFlop/s supercomputer systems

*DEISA2 as the vector for the integration of Tier-0 and Tier-1 systems in Europe*
## Integrated Tier-1 Systems

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MPG-RZG</td>
<td>IBM BlueGene P</td>
<td>27.0</td>
<td>8,192</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>IBM Power6</td>
<td>120.0</td>
<td>6,400</td>
<td>20.0</td>
</tr>
<tr>
<td>BAdW-LRZ</td>
<td>SGI-Altix</td>
<td>62.3</td>
<td>9,728</td>
<td>39.0</td>
</tr>
<tr>
<td></td>
<td>Itanium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSC</td>
<td>IBM PowerPC</td>
<td>94.2</td>
<td>10240</td>
<td>20.0</td>
</tr>
<tr>
<td>CINECA</td>
<td>IBM Blade Center</td>
<td>26.6</td>
<td>5120</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>IBM Power5</td>
<td>3.9</td>
<td>512</td>
<td>1.0</td>
</tr>
<tr>
<td>CSC</td>
<td>Cray XT4 dc</td>
<td>10.1</td>
<td>2,024</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Cray XT4 qc</td>
<td>70.0</td>
<td>6,736</td>
<td>7.8</td>
</tr>
<tr>
<td>FZJ</td>
<td>IBM BlueGene P</td>
<td>223.0</td>
<td>65,536</td>
<td>32.0</td>
</tr>
<tr>
<td></td>
<td>IBM p690 cluster JUMP</td>
<td>8.9</td>
<td>1,312</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>JUMP successor</td>
<td>&gt; 50</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>IDRIS-CNRS</td>
<td>IBM p690+ Cluster</td>
<td>6.7</td>
<td>1800</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>IBM p690+ successor</td>
<td>&gt; 40</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>SARA</td>
<td>IBM Power5+</td>
<td>14.6</td>
<td>1920</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>IBM Power6</td>
<td>60.2</td>
<td>3328</td>
<td>15.6</td>
</tr>
<tr>
<td>UEdin-EPCC</td>
<td>Cray XT4</td>
<td>65.0</td>
<td>11,520</td>
<td>35.0</td>
</tr>
<tr>
<td>UStutt-HLRS</td>
<td>NEC SX-8</td>
<td>13.0</td>
<td>576</td>
<td>9.0</td>
</tr>
</tbody>
</table>
Authentication

DEISA vision is one global federated namespace, enabling Single Sign-on facilities.
- Direct (interactive) access to all systems required for
  - Testing/debugging of code on specific architectures
  - Submission and checking of production jobs

DEISA AuthN is X.509 based PKI
- DEISA trusts IGTF accredited CAs
- Difficult to handle by users

Shibboleth technology interesting
- User doesn’t have to handle authN information – handled by middleware

No interoperability issues for AuthN
- Although DEISA may need more attribute information than IdPs deliver (including attribute certificates)
Authorization (1)

Vision is simple authorization scheme
Same policies everywhere

Access to systems based on subjectName (DN) from X.509 certificate
DN mapped to UNIX uids
No pool accounts – access to data (global file system) based on uid.

Access through UNICORE (v5) and GT4
Authorization information through UUDB (UNICORE) and grid-mapfile (Globus).

Authorization information based on LDAP based repository
No dynamic retrieval of authZ information – UUDB and gridmapfile are updated statically – different from VOMS model where authZ information is added dynamically (to attribute or proxy certificates).

Basically not very different from VOMS model – both systems give attribute information about users - VOMS repository could be filled from LDAP system

But middleware must be able to handle the information
Authorization (2)

AuthZ information providers

- LDAP repository managed by sites, not by VOs! Within HPC/DEISA accounts are managed by sites, so would also be done for VOMS servers.
- Partners/sites are the VOs
- Policies needed for managing authZ information, like for authN through PMAs

Interoperability -

- DEISA can operate a VOMS server for use by other infrastructures
- DEISA can import data to LDAP from external VOMS server (but additional attribute information needed)
Vision: uniform model for usage and budget information
Usage information published using OGF UR-WG format recommendation
Developed our own system for publishing usage records – existing systems did have dependencies on other tools
  Each site publishes data locally in DB (eXist)
  Access based on role – user, PI (principal investigator), site admin
  Based on X.509 SubjectName
  GUI for producing reports (DART)
  Conversion (comparison) between systems based on CPU performance
Accounting (2)

Wishlist

Publishing summary records
All data in central repository (automatically)
Budget check based on usage information
  Planning implementation – no job submission possible if budget is exceeded

Interoperability

Data can be exported in XML
  For importing usage information a separate DB could be set up
  But only subset of UR-WG attributes is supported
Thank you!