The German Grid Initiative D-Grid and the D-Grid Integration Project

Klaus-Peter Mickel
Institute for Scientific Computing
Forschungszentrum Karlsruhe
Co-Coordinator of the D-Grid consortium
National grid initiatives now include…

Slide from Mike Mineter from Monday's EGEE workshop - Thanks
Outline

• e-Science in Germany
• D-Grid Initiative
• Community projects
• D-Grid Integration project
• Summary and Outlook
Grid projects in the past

- Many German research institutions participated in various Grid projects
- Germany was strongly involved in the development of UNICORE middleware
Grid projects in the past

- Many German research institutions participated in various Grid projects
- Germany was strongly involved in the development of UNICORE middleware
- National Grid initiatives were started by many (large and small) countries
- But Germany was still a "white" area on the Grid map until 2001
• Some activities on national level:
  – 2001: At Forschungszentrum Karlsruhe GridKa was launched as a Tier1 centre for LHC computing [www.gridka.de](http://www.gridka.de)
  – 2002: The German Ministry for Education and Research (BMBF) disseminated "IT-2006", a funding programme for enhancing the Information Technology in Germany [www.it2006.de](http://www.it2006.de)
  – 2003: The idea of a national Grid initiative emerged from Forschungszentrum Karlsruhe – including the name “D-Grid” – and many institutions joined this initiative [www.d-grid.de](http://www.d-grid.de)
From D-Grid to e-Science

• ~300 scientists from ~100 institutions joined D-Grid
• All aspects of e-Science were discussed in working groups
  → A „white paper“ was presented to the ministry (BMBF):
  „e-Science in Germany“
• This paper became the base for the BMBF funding programme for e-Science
• Aim of e-Science:
  "To build an infrastructure in order to enhance net-based collaboration in key areas of science." (John Taylor, 2001)
  „e-Science“ covers three areas:
  ➢ Grid computing
  ➢ Knowledge management
  ➢ e-Learning
Knowledge Management & e-Learning

• Knowledge Management
  – IT based handling and management of information and knowledge
  – Open access to scientific information
  – Project "eSciDoc" (FIZ Karlsruhe & Max-Planck-Society)
    • 6 million Euro, 09/2004 - 08/2009
  – More projects in the pipeline (Ontoverse, WISENT, …)

• e-Learning
  – e-Learning = Using new media in university teaching
  – Seeking methods for enhancing innovation potential given by e-Learning and e-Teaching techniques within universities
    • 40 million Euro, 01/2005 - 12/2007

→ a common platform is needed → D-Grid
e-Science architecture

Generic platform and generic Grid services

D-Grid integration project

Grid Computing

Knowledge management

e-Learning

e-Science

Astronomy
Climate
Lifescience
High Energy Physics
Engineering
Social Science
Ontoverse
Wikinger
WISEM
Project 1
Project 2

...
The D-Grid initiative should:

• Provide very soon a nationwide reliable, sustainable and robust Grid infrastructure – built around the "Core-D-Grid" (in German: “Kern-D-Grid“)

• Develop this Core-D-Grid to an infrastructure for all e-Science applications

• Demonstrate the usability of Grids for many different communities

• Contribute in the development of Grid middleware and tools by integrating the needs of the scientific communities

• (re-)use existing technical and middleware solutions from the international context (e.g. GGF, EGEE, UNICORE)

• . . . .

• Long-term goal: "Services for Scientists"
D-Grid Initiative Roadmap

**D-Grid is becoming reality**
- Build up a national Grid infrastructure
- Define operating and support
- Find models for long term sustainability

**Becoming more concrete:**
- Describe concrete projects
- Find suitable application communities
- Seek further cooperating institutions
- Find funding agencies

**Making concepts:**
- Who are the partners?
- What are the aims?

**CY03** | **CY04** | **CY05** | **CY06**
Today: 7 projects

- 6 Community projects:
  - Astrophysics, Climate research, Medicine, Engineering, High Energy Physics (HEP), Social Science
- D-Grid integration project (DGI) to build a commonly usable platform for e-Science in Germany
AstroGrid-D

Aims:

• Creation of a nation wide collaborative environment for astronomical research institutions using Grid technology
• Setup and operating of a Grid infrastructure for the Astro-Community to optimise the usage of the existing resources
• Establishing the access to distributed astronomical data archives
• Integration of astronomical instruments and experiments in the research infrastructure
• Close collaboration with the international Virtual observatory
**AstroGrid-D**

**Workpackages:**
- Management
- Resource integration
- Meta data management
- Distributed data management
- Distributed data base access and distributed data flows
- Resource management for Grid-jobs
- Job-monitoring and interactive job steering
- User interface and Grid application toolkit

**International Projects:**
- Planck satellite
- Lofar – radio astronomy
- VO - Virtual Observatories
- LIGO – Gravitational waves
- Numerical relativity
- ...

**Technology:**
- Gridshpere
- Workflow Engine from the Planck satellite mission (Process Coordinator ProC)
- Cactus
- ...

---

Klaus-Peter Mickel  
D-Grid  
ISGC '06, 03.05.06
C3-Grid

C3-Grid = Collaborative Climate Community Grid

Aims:
• Access to distributed climate data archives
• Establishing scientific workflows in earth science community
• Setup of a productive Grid infrastructure for climate research in Germany
• Integration of existing heterogeneous systems (computing and data)
• Enable global cooperation
C3-Grid

Scientific Workflows:
- Modelling of earth systems (Climate, Oceans, Biosphere)
- Acquisition of measuring data (classical measurement, satellites)
- Comparisons of models and measurements
  - model A and model B
- Access to Data essential!!
  - Meta data
  - distributed archives
- Long-term availability of data must be guaranteed!!

Workpackages:
- User interfaces
- Grid Information services
- Access to distributed local databases/meta data bases
- Pre-processing of data
- Grid data management
- Grid scheduling
- C3 Grid infrastructure
- Sustainable integration
- Project management
Aims:

- Development of applications and components for the analysis of data of international high energy physics and astroparticle physics experiments in a national Grid environment
- Integration with the existing Tier-structure of LCG
- Development of components for data analysis using Grid resources
- Strengthen the participation in international Grid projects
  - Use of the gLite middleware from the EGEE project
- Three main focuses:
  - Data management
  - Tools for job monitoring and automatic support
  - Distributed data analysis by end users („physicists“)
• Innovative Grid developments for engineering science applications

• Aims:
  – Common use and community specific development of Grid environments for engineering applications
  – Efficient use of common resources for:
    • Modelling
    • Simulation
    • Optimisation
  – Basic and application research
  – Strong cooperation with industry
Applications:
Five different engineering domains
• Foundry industry
• Metal forming
• Flow physics
• Transportation processes in ground water
• Magneto-hydrodynamics coupling

Methods and models:
Engineering specific problems should be solved on the Grid
• Knowledge based support for decision processes
• Support for engineer specific workflows
• Distributed simulation based optimisation of products and processes

Grid specific developments:
• Security and Trust models in economic scenarios
• Cooperation and Business models
• Licensing
Aims:
- Demonstrate the usability of Grid services for medicine and life science by connecting large and high dimensional data sets
- Setup of a Grid environment for medical research
- Transfer of project results into the generic D-Grid platform
- MediGrid is open for new medical applications
MediGrid

Workpackages:
- Project management
- Ontology tools
- Resource fusion
- Middleware
- Enhanced Trust and Security
- Medical e-Science framework
  - Legal aspects
  - Biometric aspects of high dimensional data
- Clinical research
- Biomedical informatics
- Image processing
TextGrid

Aims:

- Establish a community Grid for text based science (social science) to support
  - tools for edition work
  - modules for text data processing
  - the definition of standardized interfaces to publication software
  - modules to control and monitor access to text data and text data tools
- TextGrid will be a part of Semantic Grid
Integration of:

- developments of other projects (UNICORE, EGEE,...)
- new developments from the D-Grid communities
  - to establish a generic platform for e-Science which
    - contains the software stack needed by the communities
    - provides a reliable and **sustainable** Grid infrastructure
      - "Core-D-Grid"
    - offers Grid services for new application domains
  - in order to reach the long-term vision of
    "Grid, knowledge and e-Learning services for scientists"
- DGI consists of four Working Groups
DGI – WG1: Middleware & Tools

Middleware:
- Globus 4.x
- LCG
- UNICORE
- GAT and GridSphere

Data Management:
- SRM/dCache
- OGSA-DAI
- Meta data schemas

VO Management:
- VOMS and Shibboleth

Tasks:
- Support communities
  - with training
  - with installation packages
  - with support for new operating systems
- Coordinate together with communities the development of new tools
  - Prevent „reinventing the wheel“

Needed by the communities

Klaus-Peter Mickel

D-Grid
Core-D-Grid will …
- Integrate all participating resources and providers
- Operate the Core-D-Grid
  - including User Support
- Provide monitoring and accounting tools for all middleware platforms
- Prepare the accounting and – later – the billing
- Contribute to the development of a general resource description language
- Offer Grid services for new communities
DGI – WG3: Network & Security

Aims:
• Extend the existing German research network ("DFN") by Grid specific elements
• Evaluate alternative network transport protocols
• Build and operate an infrastructure for authorisation and authentication (AAI)
• Evaluate and/or develop Grid suitable firewall concepts
• Provide a Grid specific CERT team
  – CERT = Computer Emergency Response Team
DGI – WG4: Management & Sustainability

Tasks:
- Coordinate the integration procedures within the DGI
- Coordinate the cooperation between the DGI and the community projects
- Establish connections between D-Grid and knowledge management/e-Learning
- Support new applications domains to the Grid
- Evaluate methods for facilitating sustainability

Sustainability:
- Take the different roles into account
- Evaluate the legal framework concerning sustainability
D-Grid – collaborations

- D-Grid collaborates with many other projects
  - EGEE
  - DEISA
  - Globus
  - other national initiatives
  - …
- D-Grid will contribute to standards
  - GGF, …
- D-Grid will disseminate the results to new user communities
- D-Grid wants to start an effort together with industry
  - In April 2006, a BMBF call came out to encourage „e-Science service providers“ (commercial and academic) to propose longterm solutions for „e-Science services“
Conclusion and Outlook

- D-Grid started successfully in September 2005
- The aim is to have a nationwide Grid infrastructure for scientists in autumn 2007
- Grid technology will be established in many scientific communities
- All e-Science domains are working in a close collaboration
- DGI will provide a general infrastructure for e-Science in Germany
- All three e-Science domains will organize the
  "German e-Science Conference"
  (GES2007, May 02 - 04, 2007, Baden-Baden)

➤ All the work presented is funded by the BMBF