

D-Grid: A German Core Grid Infrastructure

Matthias Kasemann / DESY

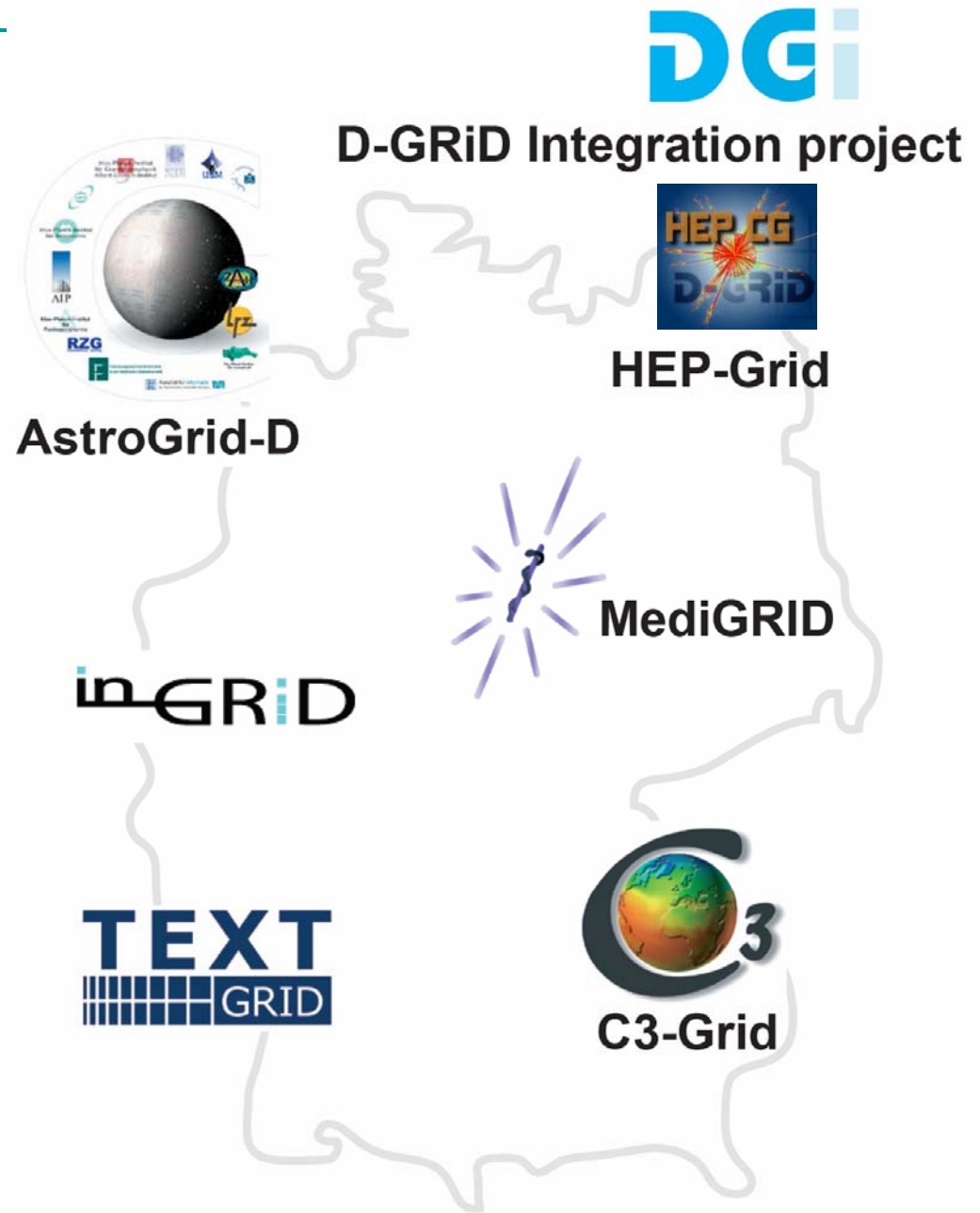
SPONSORED BY THE



Federal Ministry
of Education
and Research

D-Grid Projects: 6 Science Communities and Core Integration

- Funding: ~20M €
- Schedule:
9/05-9/08(07+)
- > 100 Partners
- Partner projects from
e-Learning



History of the D-Grid Initiative



- The 90s: Several Government funded projects for the area of “distributed computing”, like e.g. Unicores,...
 - 2000++: Many German research centres are part of European grid projects
 - 01/03: German scientists started D-Grid Initiative
 - 07/03: Initial result: strategic D-Grid paper
 - 03/04: BMBF announced e-Science Initiative for Germany
 - 08/04: BMBF Call for Proposals for e-Learning, knowledge management, and D-Grid
 - 01/05: D-Grid pre-project with final review
 - 09/05: Official start of the 6+1 D-Grid projects
 - 09/06: Additional funding (5M€) for Equipment becomes available
 - 03/07: Mid-term Review
 - 08/07: Scheduled end of DGI
 - 08/08: Scheduled end of 6 Community Projects
-
- D-Grid 2 Call in April 06
 - Deadline for short proposals June 06
 - 3 main areas to enhance current D-Grid:
 - Generic ('horizontal') Service Grids Provide services for heterogeneous user groups
 - Community ('vertical') Service Grids New communities using the existing D-Grid infrastructure and services, supported by service providers
 - D-Grid extensions, based on gap analysis of current D-Grid 1 environment
 - Funding foreseen: ~20M €
 - Start foreseen: beginning 2007

Objectives of e-Science Initiative



- Building a Grid Infrastructure in Germany
 - Combine the existing German grid activities for infrastructure, middleware, and applications
 - Integration of the middleware components developed in the Community Grids
- Development of e-science services for the research community
 - Science Service Grid
- Important:
 - Continuing production grid infrastructure after the end of the funding period
 - Integration of new grid communities (2. generation)
 - Business models for grid services

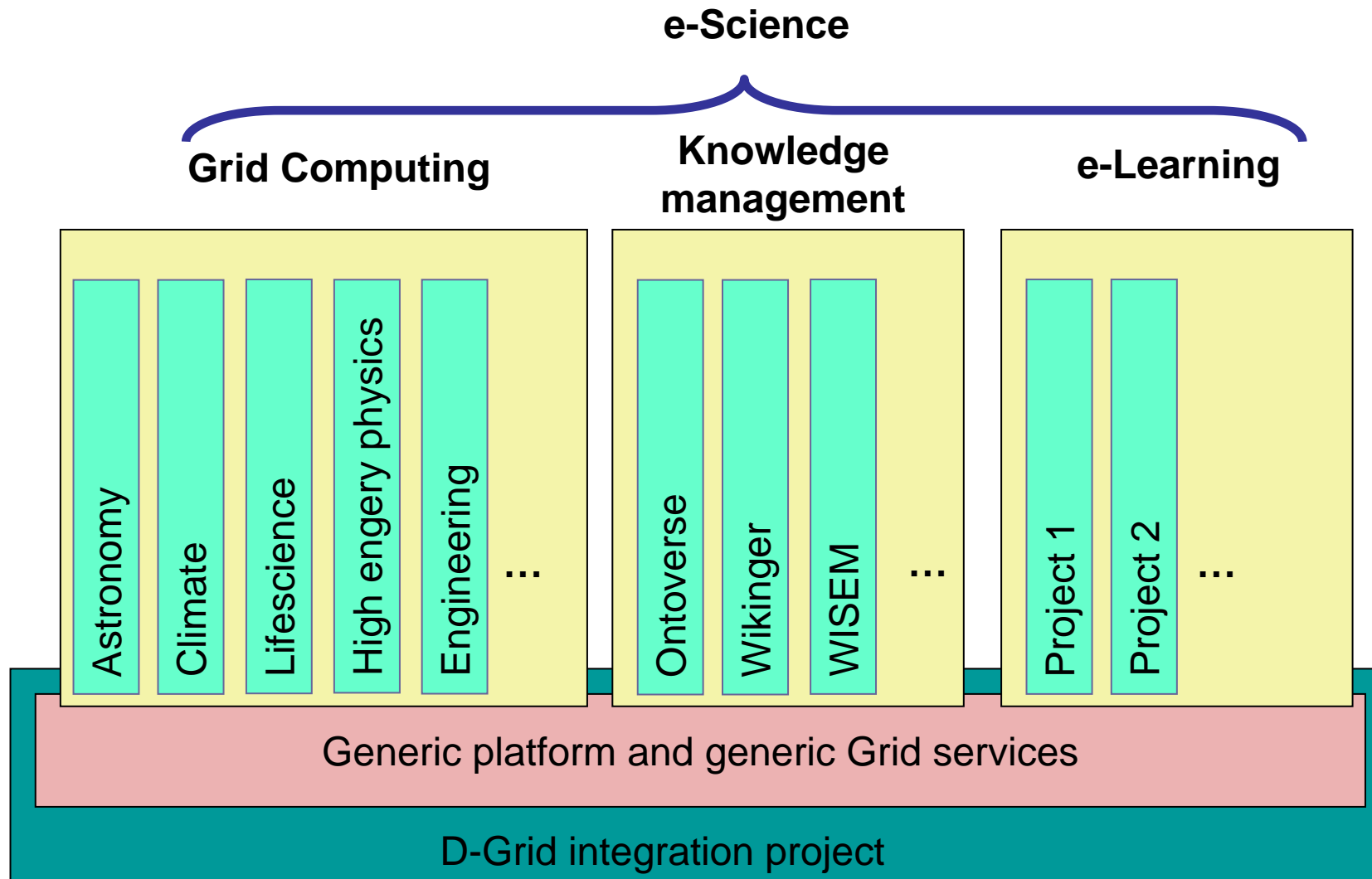
D-Grid – Goals



The D-Grid initiative should:

- provide very soon a nationwide reliable, sustainable and robust Grid infrastructure – build around the “Kern-D-Grid” (Core-D-Grid)
- develop this “Kern-D-Grid” to an infrastructure for all e-Science applications
- demonstrate the usability of Grids for (initially) six 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. GLOBUS, EGEE/LCG, UNICORE)
- Longterm goal: "Services for Scientists"

D-Grid e-Science architecture



HEP-Grid - Goals



- Focus on tools to improve data analysis for HEP and Astroparticle Physics:
 - Data management
 - Advanced scalable data management
 - Job-and data co-scheduling
 - Extendable Metadata catalogues for Astroparticle and Lattice QCD Physics
 - Jobmonitoring and automated user job support
 - Information services
 - Improved Job failure treatment
 - Incremental results of distributed analysis
 - End-user data analysis tools
 - Physics and user oriented job scheduling, workflows
 - Automatic job scheduling
- Integration with the existing Tier-structure of LCG
- Strengthen the participation in international Grid projects
 - Use of the gLite middleware from the EGEE project

HEP Grid results after PM12



Metadata catalogue for Lattice QCD in production, including API for concurrent access

- > 70k documents („configurations“)
- Huge user international community in Lattice QCD

Datamanagement with dCache, see www.dcache.org

- Load optimized dynamic performance for tape I/O in HSM mode
- Packaging and documentation, example installation provided
- Provision of Test- and evaluation infrastructure
- Optimized job- and data co-scheduling using a WMS plugin for HSM information

Job-Monitoring Tools for LCG applications

- Instrumented user applications, using R-GMA and visualisation GUI
- Technically to be integration into LCG software stack, operationally into Grid User Support (GGUS)

Interactive analysis with Grid resources (in the ATLAS and ALICE context)

- Dynamic splitting and distribution of analysis jobs (ATLAS)
- PROOF – gLite extensions to use Grid catalogue

AstroGrid-D - Goals



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

AstroGrid-D - Scope

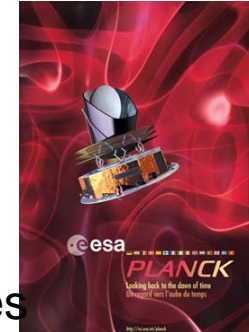


Workpackages:

- 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 interaktive job steering
- User interface and Grid application toolkit

International Projects:

- Planck satellite
- Lofar – radio astronomy
- VO - Virtual Observatories
- LIGO – Gravitational waves
- Numerical relativity
-



Technology:

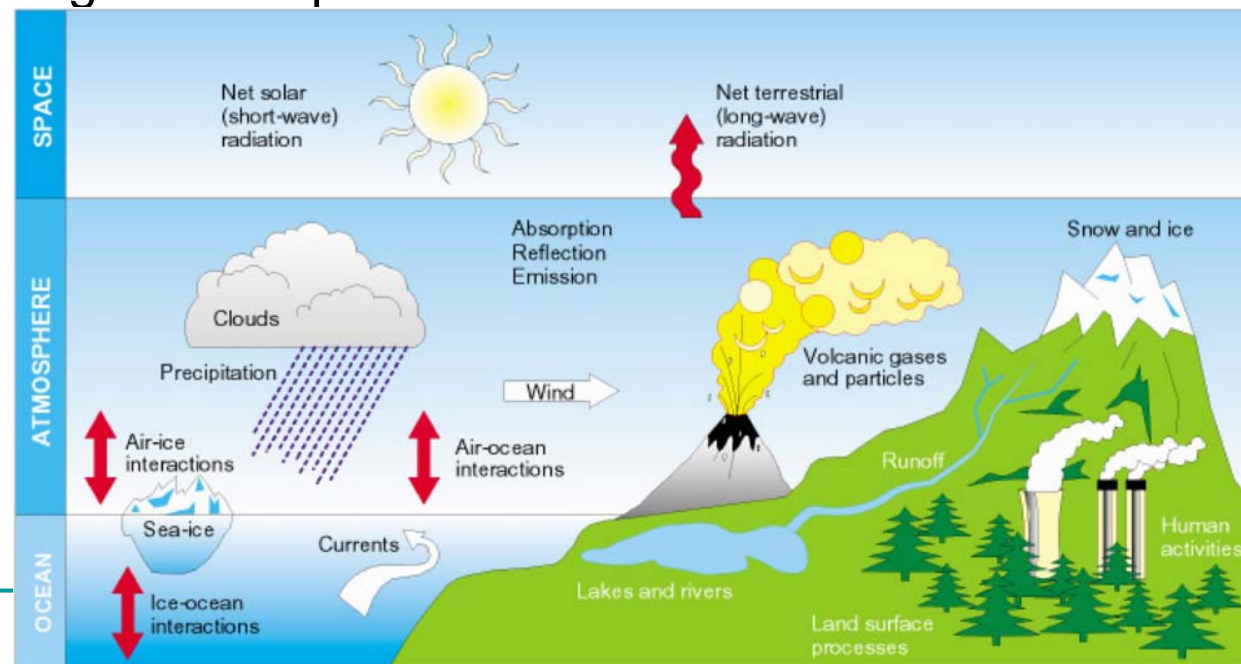
- Gridshpere
- Workflow Engine from the Planck satellite mission (Process Coordinator ProC)
- Cactus
- ...

C3-Grid - Goals



Collaborative Climate Community Grid

- 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





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
- Preprocessing of data
- Grid data management
- Grid scheduling
- C3 Grid infrastructure
- Sustainable integration
- Project management

Innovative Grid developments for engineering science applications

- 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

MediGrid - Goals



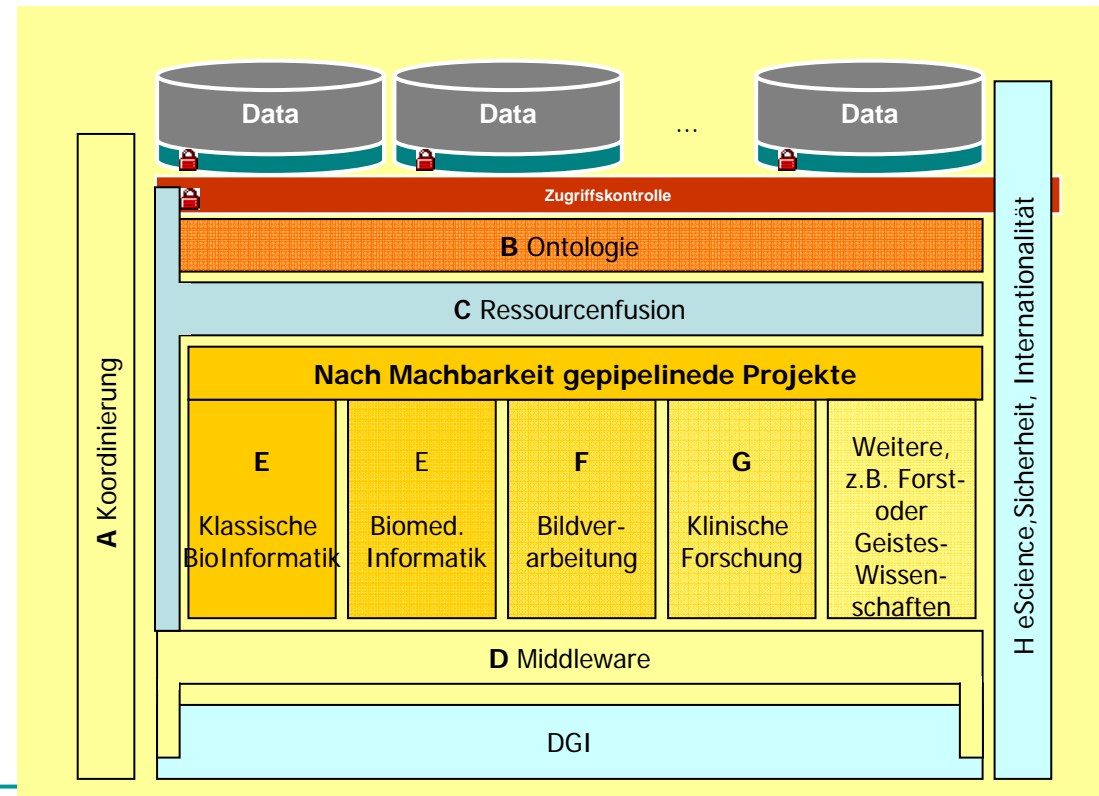
- Demonstrate the usability of Grid services for medicine and life science by connecting large and high dimensional data sets
- Setup of a Grid environment for medical research
- Transfer of project results into the generic D-Grid platform
- MediGrid is open for new medical applications

MediGrid - Scope



Workpackages:

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



Partner Projects:



Knowledge Management

- **WISENT:** Knowledge network on energy meteorology
- **WIKINGER:** Scientific libraries and cooperation
- **ONTOVERSE:** Cooperative knowledge management in life sciences
- **Im Wissensnetz:** Social Semantic Desktop
- **eSciDoc:** Platform for e-publications
- **TextGrid:** Workbench for Publishing, Text Processing, Text Retrieval, Linking, and Workflow

TextGrid - Goals



Modular Platform for distributed cooperative scientific text processing for the humanities

Workbench for Publishing, Text Processing, Text Retrieval, Linking, and Workflow

Focus of TextGrid is on:

- Quantity: Full text instead excerpts, text and pictures
- Speed of reprography: stability of text, volatility of medium
- Precision: Maximum requirements on correctness
- Availability: international open standards

DGI - D-Grid Integration Project



Integration of

- new developments from the D-Grid communities
- developments of other projects (UNICORE, EGEE,...) 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 to reach the longterm vision of

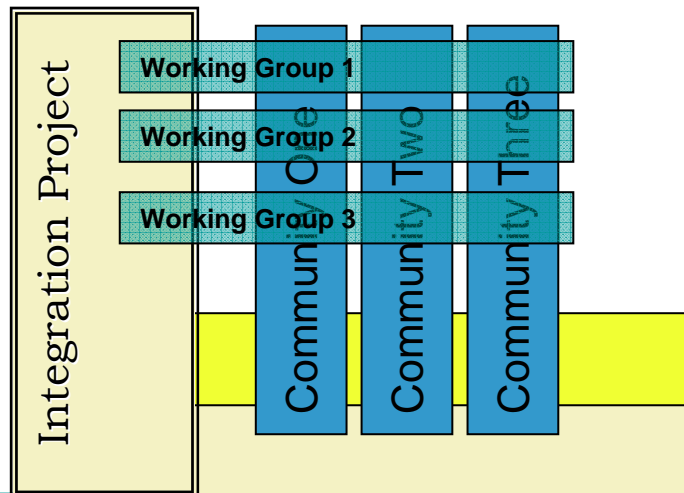
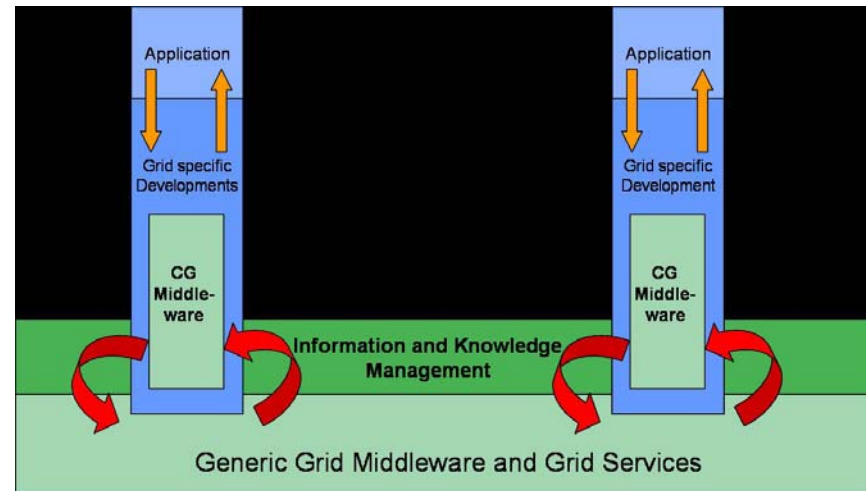
„Grid, knowledge and e-Learning services for Scientists“

DGI – Scope



Idea:

- development of new and enhanced services in the Community Grid (CG) projects
- Generalisation of services and integration into the common D-Grid platform will be done by the DGI together with the CGs



Working Groups:

- DGI will coordinate different groups together with the CGs
 - Grid scheduling
 - Data & Meta data management
 - VO management
 - AAI
 - Security

DGI: Middleware & tools



Middleware:

- Globus 4.x
- gLite (LCG)
- UNICORE
- GAT and GridSphere

Data Management:

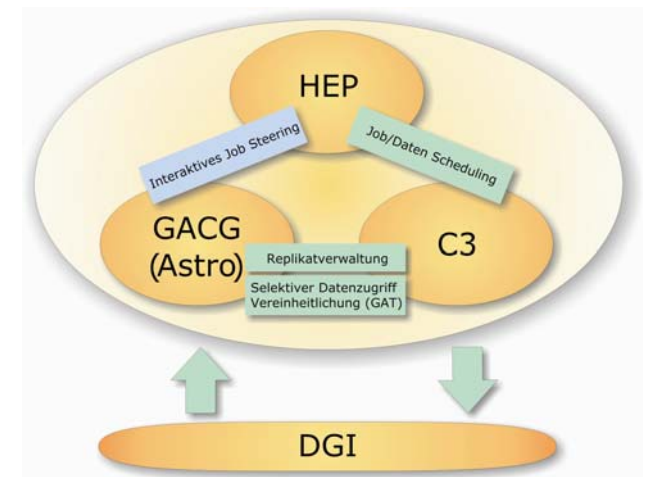
- SRM/dCache
- OGSA-DAI
- Meta data schemas

VO Management:

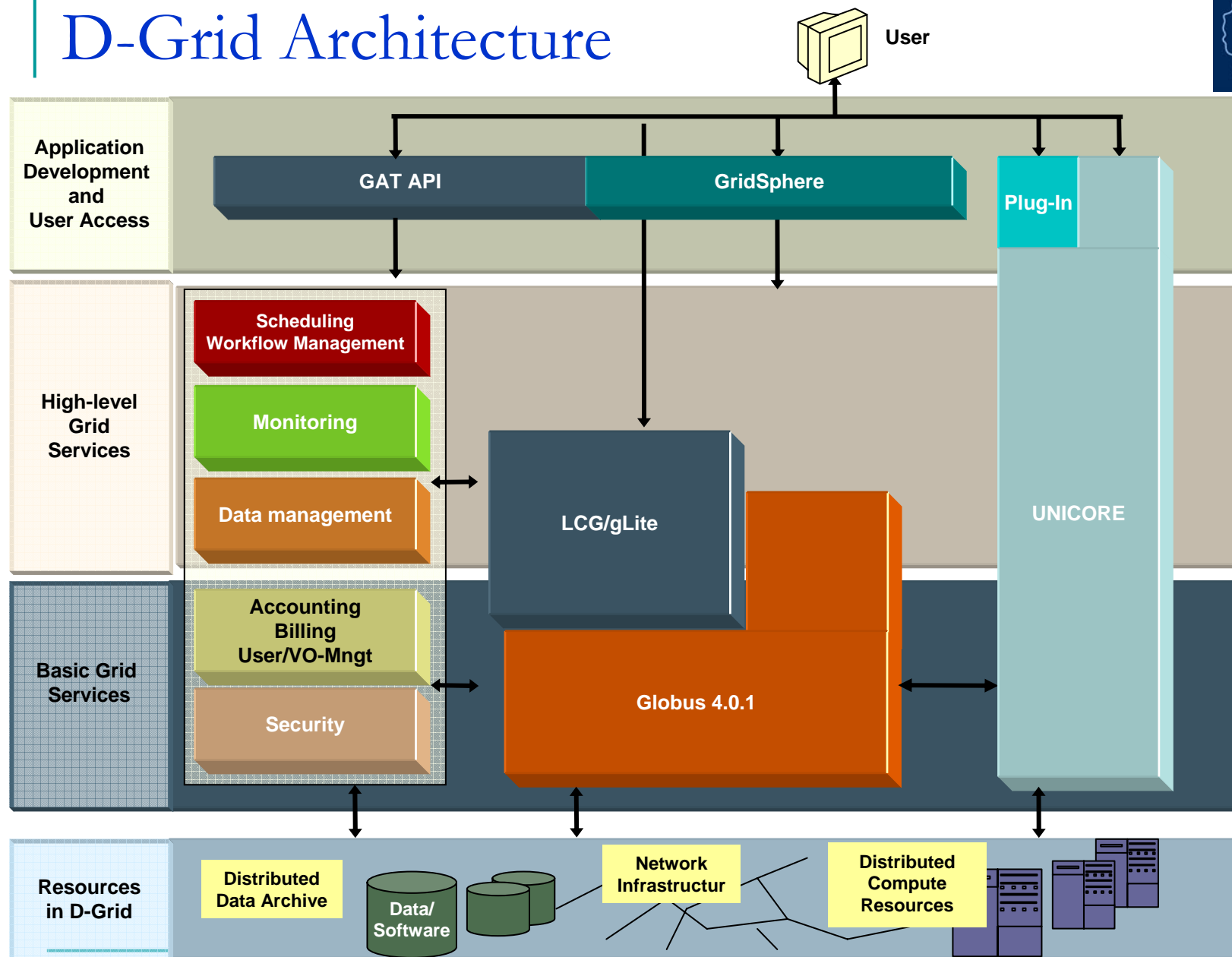
- VOMS and Shibboleth

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“



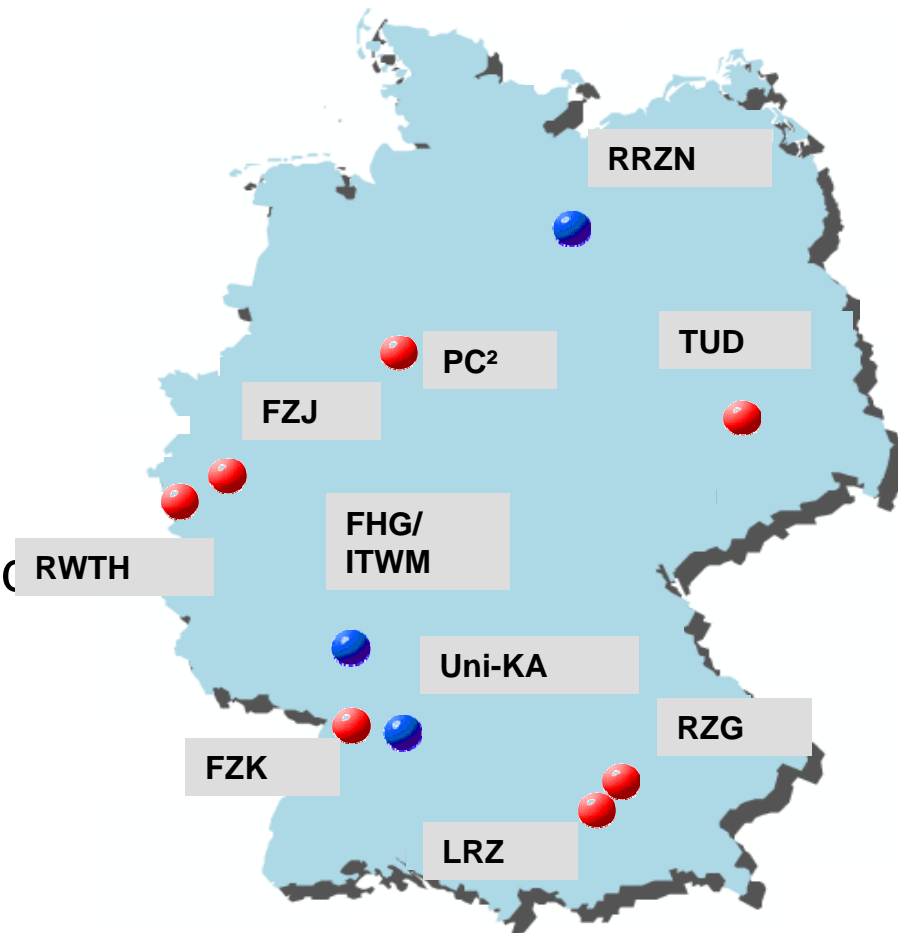
D-Grid Architecture



DGI - Infrastructure



- 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 – billing
- Contribute to the development of a general resource description language
- Offer services for new communities



Uni-KA	PC-Pool	assoz.
FHG/ITWM		assoz.

Goals:

- Extend the existing German network 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 – 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

Conclusion and Outlook



- D-Grid started finally in September 2005
- Grid technology will be established with (initially) six scientific communities
 - Some investment funding became available
 - New projects will join starting 2007
- All e-Science domains will work in a close collaboration
- DGI will provide and operate a general infrastructure for e-Science
- All three e-Science domains will organize the



- First results of the D-Grid available to communities!!
 - All the work presented is funded by the BMBF



Bundesministerium
für Bildung
und Forschung