D-Grid is Germany's long-term Grid initiative and is building a network of distributed integrated, and virtualised high-performance resources and services to enable the processing of large amounts of scientific data and information as **Professor Uwe Schwiegelshohn** explains

The communities of Germany's D-Grid

The German Grid Initiative (D-Grid) aims on achieving sustainability by incorporating a broad spectrum of different communities. The origins of these communities are in both academia and industry. It is the goal of D-Grid to explore synergies between the approaches of these communities so that communities can benefit from each other without giving up their own identity. Further, various service and integration projects assure that the same basic components and services are available to all communities. Finally, the German Federal Ministry of Education and Research (BMBF) did not only fund all the community projects but also provided investments to generate an initial Grid infrastructure located at many centres throughout the republic. In this article, we present some of the 22 D-Grid projects that are presently active.

DGI-2 at the core

The D-Grid Integration Project 2 (DGI-2) forms the core of the D-Grid base. It has the goal to put the German Grid infrastructure on a sustainable basis for long-term use.

The contents and the structure of DGI-2 are mainly community driven. This has led to five important focal points of DGI-2: Support and operation of the core Grid infrastructure, security, data management and sustainable business models. Although DGI-2, like DGI-1, is not a development project but a consolidation project, some specific development measures are indispensable and also form a work package of DGI-2.

From 2005 to 2007 the precursor DGI-1 has created a sizable core Grid infrastructure in Germany that is used by about 20 community projects today. This already existing D-Grid infrastructure is a considerable achievement even on an international scale. DGI-2 will build on these excellent foundations, and establish a complete, sustainable Grid infrastructure in Germany which can be used in the long run as a production system by science and industry, and which will contribute to strengthening science and economy in Germany.

The German Grid Initiative (D-Grid) aims on achieving sustainability by incorporating a broad spectrum of different communities. The origins of these communities are in both academia and industry

Dissemination

The project Sustainable Grid Infrastractures (SuGI) complements the work of DGI by disseminating the knowledge of Grid technology and thus enhancing its use. Therefore, SuGI addresses all academic computing centres as well as enterprises, which still have not adopted Grid technology. It makes the research experiences gained in the D-Grid projects available to the public. To this end, SuGI offers training courses, supports courses given by other projects, creates video and audio recordings and provides them online to all D-Grid communities.

In addition to the dissemination of the necessary Grid know-how, SuGI also works on the applicability of the used Grid middleware by simplifying the installation and maintenance tasks. Not only standard middleware installations are affected, but also the configuration effort will be minimised to the absolutely necessary topics supporting the standardisation of middleware installation and enhancing the general Grid productivity.

Finally, the main aspects for the sustainability of the German Grid infrastructure will be investigated. Many organisational and legal questions will be addressed prior to the execution of workflows across institutional or state boundaries. Legal aspects covering software license conditions will be investigated enabling service providers to offer a wide spectrum of Grid services in Germany. Likewise, economic aspects such as profitability and standardisation are being taken into account. New insights will lead to case studies showing potential new service providers opportunities on how to offer their services in an easy and economic way.

Climate research into the Grid

The Collaborative Climate Community Data and Processing Grid project (C3Grid) represents one of the academic research communities. In climate research, there is a growing focus on the evaluation of multi-model ensemble simulations, which at the same time comprise of data with higher spatial and temporal resolution, meaning strongly growing amounts of data. Before a regular user can process these data, he must identify the location of their storage systems as well as the formal and logical data formats. In most cases, processing of the data cannot be done at the storage system itself, so that transfer of the data to a suitable processing location is necessary. These steps are part of the regular workflow of data analysis which is time consuming and hampers the scientific process. The C3Grid project has been formed to address this problem.

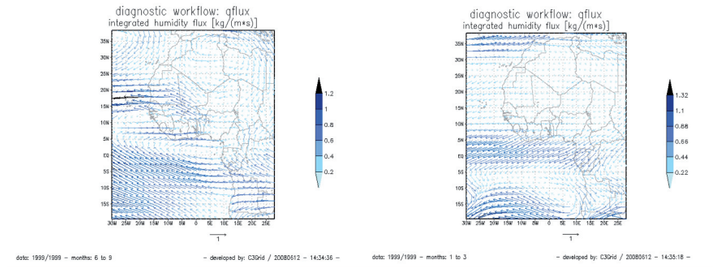

Figure 1: Graphical result of a workflow executed at C3Grid

The software services of C3Grid permit an identification of datasets and uniform data access, irrespective of the individual location. A Data Information Service collects the metadata of all data resources and provides miscellaneous search features for users. Furthermore, C3Grid offers tools to extract and process data. Up to now, some prototypical workflows for analysis of meteorological data are already implemented. In addition to these standard workflows a wide range of individual workflows can be included consisting of data extraction, data transfer and execution of some kind of calculation jobs. The user can submit these workflows at the portal. The underlying middleware of C3Grid, including Workflow Scheduling and Data Management Services, takes care of the distribution of tasks within the Grid infrastructure, thereby optimising the allocation of computing and data extraction resources as well as the automatic handling of intermediary data transfers. The implementation of different C3Grid services is based on the Globus Toolkit 4.0.x Framework. This ensures compliance to widely used standards in service definition and communication and provides a basis for interoperability to other Grids, e.g. the NERC DataGrid.

As an example, the analysis of the integrated transport of humidity between selected levels in the atmosphere is presented here. Based on a simulation the integrated transport of humidity in Africa was analysed in C3Grid. In the figure above, the flow directions indicate the moisture period at west Africa in summer due to monsoon (left) as well as the dry period upcountry in winter (right).

Grid for the arts and humanities

TextGrid is the first project in the arts and humanities in Germany creating an academic reasearch community Grid for collaborative editing, annotating, analysing, and publishing of specialist text resources. Providing a computational infrastructure, a collective network, and a comprehensive generic toolset for specialists in the arts and humanities, it is based on e-Science methods and forms a cornerstone in the emerging e-Humanities. The project establishes an interdisciplinary platform, a virtual workbench for research and a trustworthy repository for research data. The resulting architecture is shown in Figure 2.

Open interfaces open the door for other projects to plug into TextGrid. Thus, any arts-and-humanities specialist can adopt TextGrid for his work. In its core functionality, however, TextGrid is, at this stage, focused on text as a data type since there is considerable demand in the community for processing text data.

In spite of modern information technology and a clear thrust towards collaboration, researchers in the arts and humanities cannot currently make full use of the potentials of this development. For example, text scholars researching into the relations between language and discourse and into the complex processes in the genesis of literature, still mostly work in local systems and project-oriented applications. Current research initiatives also lack integration with already existing text corpora, and they remain unconnected to resources such as dictionaries, lexica, secondary literature and text processing tools. This integration and interconnection, though, bears a wealth of opportunities. With its architecture and integrated tools and services that satisfy these requirements, TextGrid is able to provide such forms of integration.

To ensure technological compatibility and data interoperability and to quickly respond to new requirements arising in

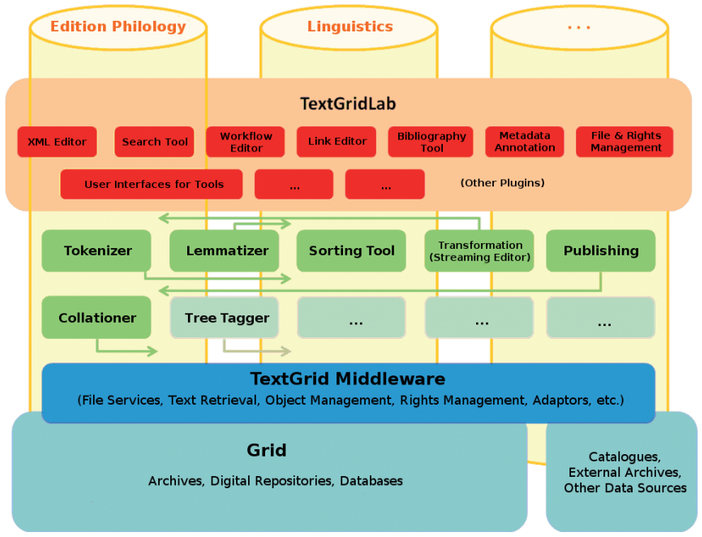

Figure 2: TextGrid architecture

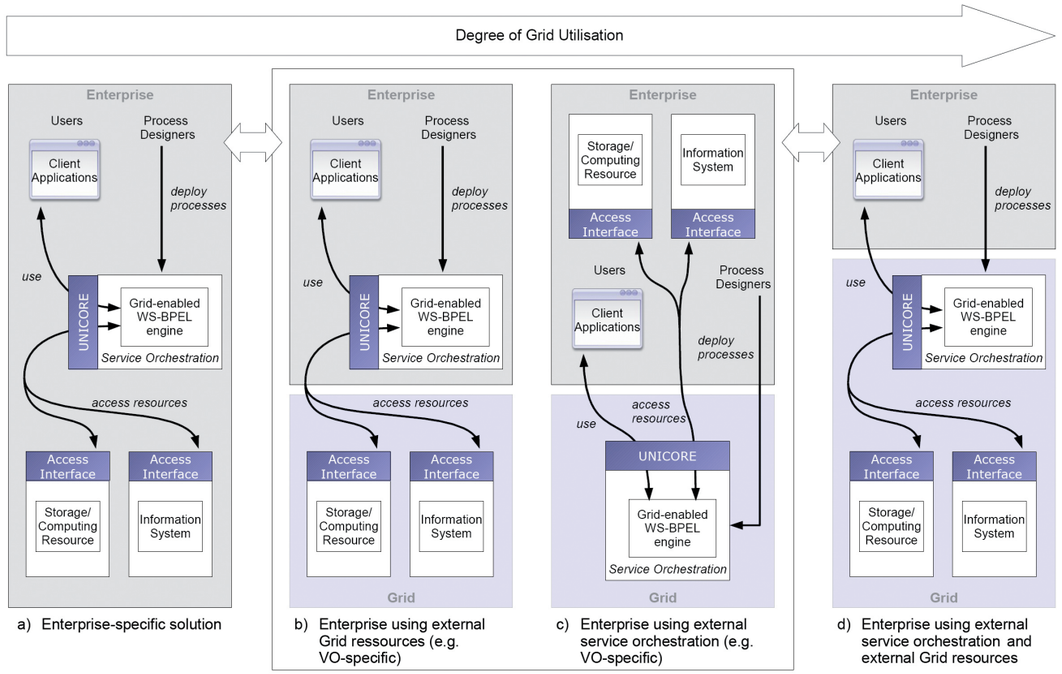

Figure 3: Dynamic utilisation of Grid resources using the BIS-Grid middleware

the arts and humanities, TextGrid participates in several international projects concerning the e-Humanities (DARIAH, interedition, Bamboo) and Grid technologies as well as standardisation issues (Open Grid Forum, IEEE – Repositories Track).

Grid means good business

The project Business Information Systems in Grids (BIS-Grid) focuses on employing Grid technologies for information system integration. It has the goal to enable small and medium enterprises (SMEs) to integrate heterogeneous business information systems and to use external Grid resources with the least possible additional effort. To this end, BIS-Grid especially considers industrial security requirements on Grid utilisation and integration as a Service distribution model. It particularly develops Grid workflow middleware that is capable to orchestrate WSRF Grid Services. This middleware is based upon service extensions to the UNICORE 6 Grid middleware using an arbitrary WS-BPEL workflow engine without modification. This integrates the WS-BPEL language, the industry de-facto standard for service orchestration, into Grid technology. As the exchangeability of the WS-BPEL engine in the BIS-Grid middleware is guaranteed this approach avoids incompatibility with commercial WS-BPEL engines.

The BIS-Grid middleware enables enterprises to dynamically switch between enterprise-specific Grid-based EAI and Grid utilisation. While in the first case resources and applications are both located within an enterprise (inhouse providing), it is also possible to outsource application providing and/or resources (Grid application providing), and even the integration of self generated services into the Grid while guaranteeing appropriate quality of service requirements. This allows the use of external Grid Services in enterprise-specific workflows, and helps to create specific Virtual Organisations (VOs) that may traverse enterprise boundaries within a defined range.

Geospatial data and the Grid

Vast amounts of geospatial data have been gathered in the last decades. With the advent of commodity applications such as Google Earth, usage of these geospatial data has become ubiquitous. However, most raw data must be processed before it can be put into a meaningful context. In the Spatial Data Infrastructure Grid project (GDI-Grid), computationally demanding tasks such as disaster and noise dispersion simulation will be deployed on existing Grid resources to dramatically decrease computation times. GDI-Grid plans to combine two complementary technologies: the well-established Spatial Data Infrastructures (SDI), which mainly provide access to spatial data resources, and the Grid, which performs processing and storage of enormous amounts of data. From a technical point of view, it is necessary to develop interfaces between the current base technologies of SDI and the Grid middleware to enable seamless processing of spatial data in Grids. In order to achieve this goal, data and models must be coupled with Grid services and security mechanisms to create and enable entire SDI-specific workflows in Grid environments.

In GDI-Grid, industry-leading partners team up with academic research in an effort to leverage the advantages of Grid technology and spatial data infrastructures. Combining cutting edge technologies from both disciplines leads to a significant improvement of disaster simulation and recovery. Special attention is given to the detailed simulation of flood catastrophes – an important scenario considering the very current discussion about global climate change.

Better use of renewable energy

In general the generation of usable energy increasingly depends on renewable energy sources whose availability varies significantly over time and space. Nevertheless, it is highly desirable to plan the deployment of new solar and wind power plants based on reliable data and to operate the existing facilities efficiently. Moreover, energy distributors need economically competitive energy generation capacities based on the renewable energy sources. These tasks require very detailed information about the available solar and wind supply and the resulting energy production. Very large data sets, near-realtime computations with physical models such as numerical weather prediction models, and simulations of solar and wind power plants' optimisations to the estimated supply are direct consequences of these needs. As a result, energy meteorology is not just interdisciplinary, but also a highly compute-intensive area of research.

The Knowledge Network Energy Meteorology project (WISENT) pursues as its main objective the construction of a high-performance IT infrastructure for the young energy meteorology community. Grid technology enables efficient numerical simulation and the processing of large amounts of data; thus, allows the improvement of prediction methods and other applications for energy meteorology such as determining optimal locations for different types of power plants. For example, archived solar irradiation data combined with geographical information and technical-economical information can be used to select sites for solar power plants. Such simulations and analytical computations are based on heterogeneous data sets from various sources such as satellites and meteorological ground stations. The large amount of data and the complex computations cause unacceptably large processing times on a single computer. In addition, newer generations of satellites will deliver higher resolution images, which will further increase the data volume. As a consequence, new solutions for looking up data stored in the Grid are required. Grid technology promises an efficient solution for the processing of such large amounts of data and for executing complex simulations. With this technology, the WISENT will be capable to contribute to the securing of tomorrow's energy generation based on renewable energy sources.

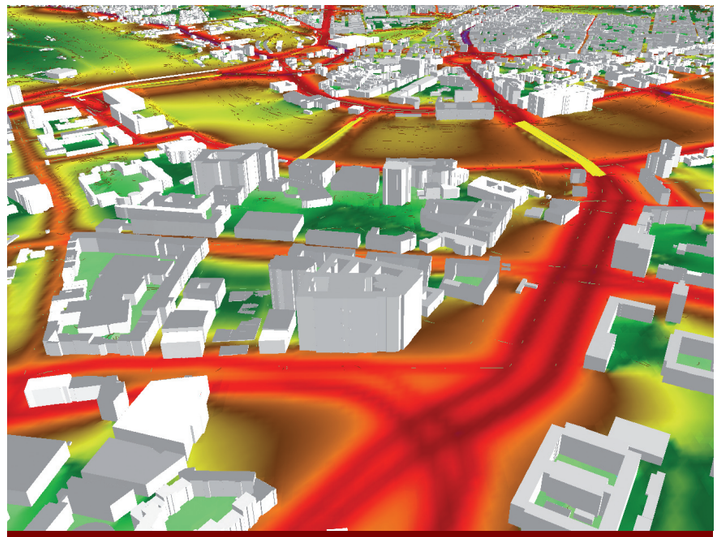

Figure 4: Noise dispersion simulation for urban areas

Grid helps collaboration

Industrial design and production processes are increasingly based on the division of labour. Companies concentrate on their core competences and buy parts, equipment and services from third parties, instead of covering the full value chain. To support this trend the PartnerGrid project develops a collaboration platform based on the D-Grid infrastructure and D-Grid services that facilitate the co-operation of companies with a special focus on small and medium enterprises (SME). The functionality and the benefits of this collaboration platform are demonstrated by using it on two typical scenarios of different industrial areas:

In the founding industry, rising quality requirements and demand of very short development processes generate high competitive pressures. To meet these challenges a close collaboration between the foundry and its customers as well as efficient optimisation processes of cast parts design based on cast-technical simulations are needed. PartnerGrid develops a Grid-based virtualisation strategy with special focus on licensing issues. The implementation of this strategy enables foundries and their customers to efficiently use all available resources resulting in a better design and an optimisation of their casting processes.

Virtualisation technologies are also used to carry out crash- and plastic deformation simulations in today's metal-processing industry. This requires a fast and secure data management infrastructure and collaboration support for tasks like project management, result presentation and expert discussions. The PartnerGrid consortium develops a Grid-enabled collaboration platform that integrates these simulation runs with the associated project management tasks.

The German Federal Ministry of Education and Research (BMBF) did not only fund all the community projects but also provided investments to generate an initial Grid infrastructure located at many centres throughout the republic

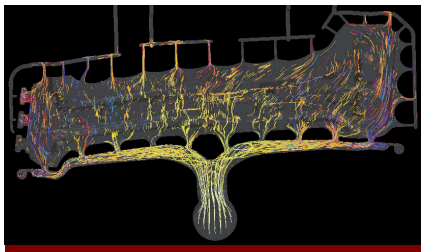

Figure 5: Simulation of a founding process (MAGMA Giessereitechnologie GmbH)

Industry into the Grid

The Biz2Grid project aims to migrate industrial business applications to existing Grid-middleware systems. To this end, it is necessary to design the organisational structure and an economically reasonable billing and pricing approach for Grid resources in combination with a prototype implementation in commercial example settings.

At the associate automotive partner BMW Group, two applications are selected for this purpose. Besides the pure technical implementation the project wants to determine how the existing IT structures in planning, procurement and maintenance can dynamically be affected by market mechanisms while maintaining a stable control and business environment. On the one hand Biz2Grid contributes to commercial use of Grid applications. On the other hand resulting services can be used for D-Grid as well as for developing commercial products with proprietary components.

AeroGrid project flying high

The AeroGrid project also addresses users from the German mechanical industries and corresponding research centres. It plans to provide an efficient Grid based working environment for the aerospace research community. The AeroGrid environment will be a permanent and effective Grid infrastructure for cooperation between industry, research centres, and universities in aerospace engineering and research. The environment allows virtual organisations to cooperate in research and development projects, to always use up-to-date program versions, data, and compute resources across all locations, and to document and trace the detailed history of a computational process that leads to a certain result.

The design of the AeroGrid environment addresses two important requirements: First, suitability for daily use, especially for the productive operation by service providers after the end of the project and second, applicability to similarly organised communities in science and industry. The integration of a Provenance service which records detailed information of all conducted execution steps, increases the dependability of the results and improves the user's confidence in their quality. The AeroGrid environment will consist of two different user interfaces, a Web portal based on GridSphere and VINE, and the data management client application DataFinder. Both user interfaces will be customised to support the various usage scenarios. The Grid middleware used for the AeroGrid infrastructure is UNICORE 6, but the user interfaces allow the usage of Grids based on the Globus Toolkit too.

Banking on Grid

Increasing competition in the German banking sector is similarly leading to a high pressure for restructuring and further automation in IT-related business processes in banks and financial services providers. Additionally, new legal regulations such as Basel II and changing customer needs towards highly customised on-demand financial products enhance this pressure. To address these challenges, the Financial Business Grid (FinGrid) project strives to identify suitable services and processes in the financial services sector and to develop Grid-based systems that enable financial service providers to reorganise their processes efficiently and to realise applications that have been impossible so far in terms of computational requirements.

To guarantee relevance for the target industry, the research is performed jointly with leading financial industry partners. Based on the technical foundations of D-Grid, new prototypes for banking service provisioning together with integrated pricing and accounting structures are being developed, tested, and implemented within the financial services sector. These prototypes are developed in cooperation with Deutsche Bank, Dresdner Bank, IBM, and other partners.

Sustainability of the research results is guaranteed due to the direct transfer of results into the financial services sector and the expected publications in national and international academic peer-reviewed media.

D-Grid contains more projects to cover an even wider range of possible Grid communities. At the moment, the German Federal Ministry of Education and Research is in the process to install a third group of projects to bring Grid technology to more users and therefore improve sustainability of Grid technology. ★

At a glance

Project Title

D-Grid

The D-Grid initiative currently runs 22 projects involving 117 partners hailing from industry, academia and research institutions. D-Grid is funded by the German Ministry of Education and Research (BMBF).

The overall objective of the D-Grid project is to establish a sustainable Grid infrastructure in Germany that provides access to high-level IT services for industry and academia.

The D-Grid funding period started in 2005 and will last until 2011, although individual projects will run for three years at most.

Contact details

Professor Uwe Schwiegelshohn
Technische Universität Dortmund
Robotics Research Institute
Dortmund
Germany
T: +49 231 755 2634
E: uwe.schwiegelshohn@udo.edu
W: www.d-grid.de

Professor Uwe Schwiegelshohn

Head of the Robotics Research Institute at the TU Dortmund

Professor Uwe Schwiegelshohn assumed his current position as managing director of the D-Grid Corporation in 2008. He participates in several D-grid projects including the integration project, in which he is responsible for the sustainability of the basic infrastructure.