



Contribution ID: 162

Type: Poster

## **A service oriented infrastructure to integrate earthsystem databases into the grid**

**Describe the scientific/technical community and the scientific/technical activity using (planning to use) the EGEE infrastructure. A high-level description is needed (neither a detailed specialist report nor a list of references).**

Climate research is data- and collaboration-intensive. Climate and earthsystem models are calibrated and driven by data of different scientific and technical sources. Model results describe several spheres and are needed and analysed by various scientists of diverse disciplines. Moreover, multi model comparisons gain in importance to evaluate uncertainty of models and results. Yet, most of the data is stored in large archives and central databases and analysis is done locally and individually.

**Report on the experience (or the proposed activity). It would be very important to mention key services which are essential for the success of your activity on the EGEE infrastructure.**

A prototype, developed in collaboration with the German, community driven Grid initiative C3-Grid (<http://www.c3-grid.de/>), was setup to demonstrate the feasibility of the developed system. Data of different German earthsystem science data centres can be discovered, browsed and uploaded to the EGEE infrastructure via a central Web portal. Via the same Web portal an example workflow can be triggered to run on EGEE; the results are automatically described in ISO 19115 and republished to a central metadata catalogue.

The administration, update and republishing of processed data is based on the data management services of EGEE, such the lfc-tools, the lfn catalogue and the storage elements. To find and retrieve the data of different data centres, currently tools of the C3Grid are used. The system can be easily expanded by further international data providers. Respective collaborations with the British NERC datagrid and the US-American earth system grid are ongoing.

**Describe the added value of the Grid for the scientific/technical activity you (plan to) do on the Grid. This should include the scale of the activity and of the potential user community and the relevance for other scientific or business applications**

The Grid offers a common platform to share data, tools and resources, which could be useful to the entire climate community. Even though most large climate and earthsystem models are designed for specialized computer architectures, the pre- and post-processing of input and output data could be done on the grid, once data and tools were accessible. But still some effort is required to seamlessly and efficiently integrate data - described by complex metadata and e.g. stored in databases - into the grid.

We develop a service-oriented architecture to integrate external data sources with complex metadata into a grid infrastructure. The system is built modular and based on common standards such as webservice technology and gridftp for data access, ISO 19115 for data description and OAI protocol for metadata harvesting. This

makes the system easily adaptable and expandable and thus potentially beneficial also to other communities that wish to integrate databases and their describing metadata

**Primary author:** Dr RONNEBERGER, Kerstin (DKRZ)

**Co-author:** Dr KINDERMANN, Stephan (DKRZ)

**Presenters:** Dr RONNEBERGER, Kerstin (DKRZ); Dr KINDERMANN, Stephan (DKRZ)

**Track Classification:** Demo and Poster session