A service oriented architecture to integrate external earthsystem data centres into the grid

Kerstin Ronneberger & Stephan Kindermann (Deutsches Klimarechenzentrum, Hamburg)

**Motivation**

**Grid as a common platform**

- Connect external data sources with complex metadata to the grid
- Share data, that is produced in the grid, with traditional community

**Climate research is data & collaboration intensive**

- Models use and produce diverse data
- Data is needed and analysed by scientists of diverse disciplines

**Climate data is organized heterogeneously**

- Data vary tremendously in description detail and accessibility
- Data are stored in heterogeneous, organisational separated data centres

**Architecture**

Interconnecting legacy systems and grid resources

- Modular SOA to flexibly integrate external data sources with complex metadata into the grid
- Collaboration with German C3Grid[1] to integrate climate community

**Standards to hide implementation details**

- Authentication & and data discovery
- Data description
- Metadata harvesting
- Data & processing request
- Data access
- Metadata provenance

**Reference Installation**

**Grid Perspective**

Connect to off-grid communities:
- expose data on the grid to external community
- offer interfaces to use external data in the grid

**Provider Perspective**

To expose data resource to grid:
- transfer metadata to ISO format
- connect local data access to WS interface
- deliver data to a gridftp storage

**User Perspective**

Use and process data of various sources:
- consistently discover data from various sources
- select a subset and directly download it, or upload it to EGEE
- trigger offered processing with automatic republishing of results in ISO
- or use update interface to automatically republish own processing results in ISO

**Current Status**


**Added value**

Contact: ronneberger@dkrz.de, kindermann@dkrz.de

References

[1] Collaborative Climate Community Data and Processing Grid: http://www.c3grid.de