Technical Details

Metadata:

Format: Metadata are written in the ***XML implementation 19139 of the ISO Standard 19115***, which has been adapted to the needs of the climate community. These metadata describe discovery aspects of the data (such as content details, spatial and temporal coverage etc.) and some use aspects (such as data format, location and size).
For the reference implementation, so far 8 German earth science data centres are providing their metadata in ISO format

Harvesting: The ***OAI-PMH protocol*** forms the interface to harvest the ISO metadata. It is a common protocol for metadata harvesting in earth science portals and thus allows for easy integration into existing web portals.
All ISO metadata produced in the reference implementation are made available via OAI-PMH and are harvested to and can be browsed in the C3Grid web portal (http://www.c3grid.de/portal/grid).

Storage: No specific metadata storage technology is required. So far, we have gained experience with several implementation variants:
In the C3Grid reference installation the ***Lucene catalogue*** is used and accessed through the web portal; for metadata storage in EGEE the ***AMGA catalogue*** was used so far, interfacing a MySQL database. Currently, as the relational database shows up its limits in terms of efficient XML-storage, we are exploring the performance of an ***eXist*** database, interfaced to the grid via ***OGSADAI.***

Provenance: A ***java API*** to manage, store and update ISO metadata during processing is developed, interfacing either AMGA (assuming a relational database) or OGSA-DAI (using ***Xquery***, assuming a XML database). In the reference implementation the java API is used to update, manage and republish metadata belonging to data, stored in EGEE.

Data:

Request: In the C3Grid project a ***webservice interface*** was developed to hide the proprietary access and storage details of different data resources. It offers a flexible way to submit data and processing requests to the respective data and/or resource provider; not only are the necessary parameters to describe the requested data submitted, but also a description of the required pre-processing (such as selecting of data subsets, calculating of mean values and/or transformations of the geographic representation of the data). In the reference installation the webservice client has been implemented at each data provider as well as on an EGEE user interface.

Access: The requested or processed data is made available via ***gridftp***. This offers a commonly used interface to the grid-world. In the reference installation, this interface is also used to directly upload and register requested data to EGEE via lfc-tools. Each data provider has a gridftp server installed to provide the requested data. The C3Grid web portal "speaks" gridftp to provide data to the user.

Security: A security concept, based on ***shibboleth (gridshib)*** and ***short-lived-grid-credentials*** is currently developed in the C3Grid and will be expanded to be interoperable with EGEE. In the reference installation we are not yet in a production mode and thus still work with collective accounts.