GES 2007 German e-Science
Available online at http://www.ges2007.de

The Collaborative Climate Community Data and Processing Grid (C3Grid) – Architecture and Implementation Issues

W. Hiller and B. Fritzsch

Alfred Wegener Institute Foundation for Polar and Marine Research, Bremerhaven
Am Handelshafen 12, 27570 Bremerhaven, Germany
Email: {Wolfgang.Hiller, Bernadette.Fritzsch}@awi.de
phone: (+49 471) 4831 1506 *fax:* (+49 471) 4831 1590

Abstract

The project „Collaborative Climate Community Data and Processing Grid – C3Grid“ is one of the community projects in D-Grid (sponsored by German government BMBF) and aims at linking distributed data archives in several German institutions. The fundamental architecture is discussed, which includes C3Grid specific components beside the standard middleware. The implemented infrastructure for scientists in climate research provides tools for effective data discovery, data transfer and processing.

1 Introduction

Climate research and Earth System sciences strongly depend on different data sources from simulation runs of coupled earth system models and a multitude of observational data. These data sets are distributed over many archives and sites worldwide and have different states with respect to general accessibility and data quality.

On the other hand the data volume is continuously increasing. Therefore the daily work of a researcher in this field consists of tedious procedures for data access, data quality control and extraction, before he is able to even start the process of scientific evaluation of theses data sets. It is obvious, that only a common platform with uniform access mechanisms and standardized data descriptions can substantially improve the scientists' ability to use this data in a successful way to explore and explain the Earth system.

The **C3Grid Mission Statement** essentially reflects this situation by

- enhancing ***uniform and easy access to distributed and heterogeneous data sources,***
- adopting a ***workflow oriented approach,***
- provisioning ***standardised data descriptions based on ISO standards*** and
- delivering ***combined storage and processing facilities for scientists.***

The C3 consortium consists of major German earth science institutions, including data providers as well as workflow developer and scientific user. Within the data providers are the world data centres WDC Climate, WDC

RSAT and WDC Mare as well as Germany's National Meteorological Service (DWD). Furthermore, partners with grid experience from Zuse Institute Berlin (ZIB) and University Dortmund develop community specific components in C3Grid.

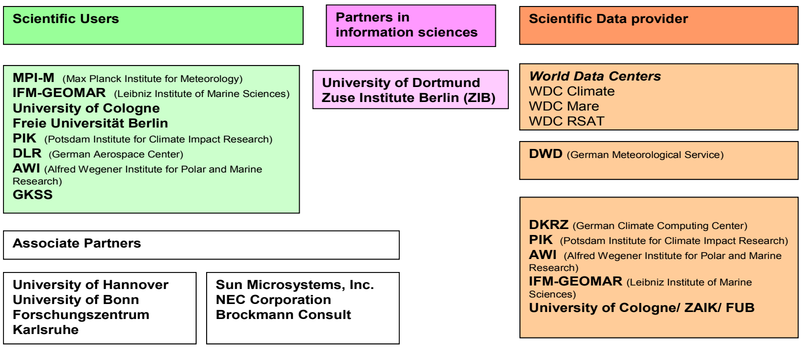

Figure 1: C3Grid Consortium

2 Metadata

A precondition for uniform access is the definition of a common metadata system. Until the start of the C3 project, the data providers had different metadata schemes for their data. Based on the international schema required for describing geographic information and services (ISO 19115 / 19139) an own C3 metadata profile was established. The process of ISO 191xx standardisation was finished in late 2006, so it could be adopted as the general C3Grid ISO format just in time.

It provides a general description framework for geographic data products in Earth System research. Some C3Grid adaptations and usage agreements were necessary to implement the standard for the C3Grid Community, but these are complementary and do not limit the international compatibility. These form the specific C3Grid Profile of ISO 19139/ ISO 19115 Geographic Information Metadata and consist of C3Grid Data Access Interface References, Data Aggregation and Derivation Info and Data Content Description based on the Climate Forecast Convention. Further details see in [1]

Data providers have to map their individual metadata schemes to that profile. Initial automated transformations of proprietary metadata formats into the common C3Grid metadata format were implemented at Hamburg (for data in the CERA database system as well as the DKRZ file archive), the University Bremen (for the Pangaea database) and for some test cases at DWD. Some data providers mapped their metadata manually to the global C3Grid metadata profile, an interactive web-form based interface to generate ISO conforming C3Grid metadata descriptions was developed at Potsdam Institute for Climate Impact Research.

3 Architecture and Middleware Components

The C3Grid architecture (see Figure 2) reflects the requirements of the potential users. The basic middleware is Globus Toolkit 4.x complemented by some C3Grid specific components. Central components will be described in the following sections.

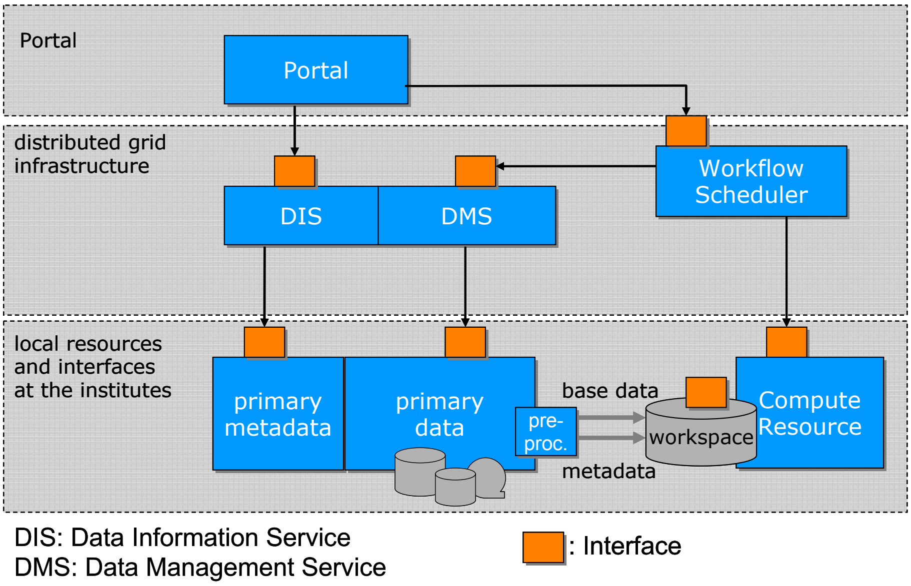

Figure 2: C3Grid architecture

The overall strategic decision for the C3Grid architecture was to adhere to standards and GGF or D-Grid/DGI supported software components whenever possible. Only in areas where no substantial software offerings or packages were on the market, C3specific software was decided to implement. Still there was general agreement to use definition standards whenever possible. The chosen software stack is displayed in Figure 3.

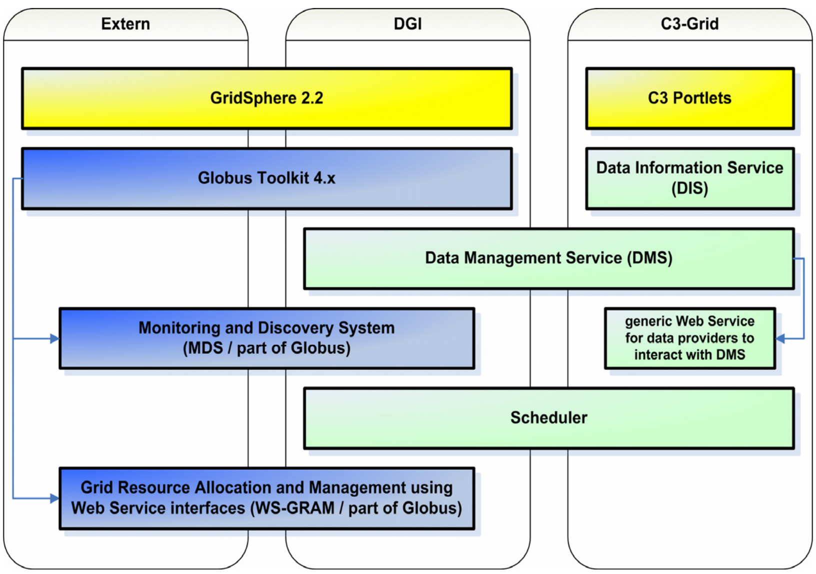

Figure 3: C3Grid software stack

3.1 Portal

The portal is the central focus point for user interaction with the C3Grid resources. It allows the user to initiate data queries for all accessible distributed data sets in C3Grid and to formulate data selection and pre- or post-processing steps in a suitable workflow oriented way. The portal serves as a single sign-on point, where all necessary user credentials are generated and addresses all necessary needs for authentication, authorisation (AA) and attributes for selective data access, by a Shibboleth based AA security infrastructure, with its implementation on the way in 2007.

The implementation of the C3Grid portal is based on the GridSphere open source portal framework [2]. The Gridsphere Software is supported by D-Grid DGI and is completely Java Specification Request 168 compliant. This ensures an extensible architecture because of the modular and flexible portlet structure and the integration of the Java Commodity Grid Kit for direct use of Globus Toolkit features (e.g. GridFTP).

Two interfaces connect the portal to the rest of the C3Grid architecture, one for the Data Information Service (DIS) and one interface for the Workflow Scheduling Service (WSS). As it was possible to integrate the class libraries of the DIS directly into the portal framework, a Java API based web service interface of DIS was implemented. The portal is also connected as a client through a conventional axis web service to the WSS.

Via these interfaces the portal's main functionalities are realised. The user can freely define data queries by making specific selections for regions of interest, time constraints or variable names (based on the Climate Forecast Convention thesaurus).

The interface of the scheduler allows the user to directly specify data downloads and corresponding preprocessing functionality, as well as to initiate the staging of data for later analysis in the workspace system described below. Aside from this rather elementary submission of scheduled jobs, the user can also submit complete predefined workflows e.g. Stormtrack and Humidity Flux analysis (which are available at the moment as 'proof of concept workflows').

The task/ workflow control is implemented via Job Submission Description Language (JSDL). In order to describe workflows consisting of several sequential but mutual dependent chains of elementary tasks, a Workflow Specification Language (WSL) was defined. Another portal functionality is the monitoring of tasks/ workflows and cancellation of tasks via handle.

3.2 Workflow Scheduling Service (WSS)

Climate research requires the execution of complex modular applications. The scheduling system must be able to handle sets of interdependent tasks which comprise multiple steps in the analysis of a scientific problem. Moreover, high volume data requirements emerge from long-term experimental recordings. Therefore, the scheduler has to work in cooperation with the data management system (DMS) which has the full information about the location of data and their replicas. Sophisticated strategies for the orchestration of data availability and compute resource usage have to be applied in order to allow seamless co-allocation of user workflows. The overall architecture and the specific interfaces of the WSS are displayed in Figure 4. As can be seen, the DMS and compute resources are both coupled to the WSS via web services.

The description of atomic tasks uses the Job Submission Description Language (JSDL) and contains staging, execution and publishing directives. The description of workflows uses the C3Grid-specific Workflow Specification Language (WSL) and defines descriptions of dependencies between atomic tasks.

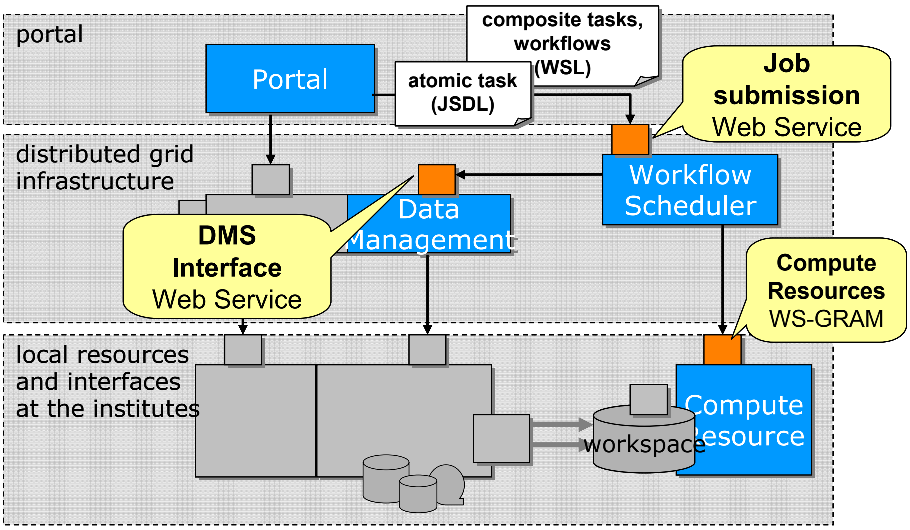

Figure 4: Definition of Interfaces between Scheduling and portal, Data Management and compute resources.

The refined interaction of the WSS with the DMS is based on an abstract negotiation protocol [3]. The DMS provides an abstraction layer to underlying services and communicates this information to the scheduler. By means of the negotiation protocol between WSS and DMS agreement is reached on the provision of services such as data staging und preprocessing or data transport. Also the general supervision and control - as for example status requests or cancellation of requested services - is carried out in this way. Figure 5 gives an example of such an interaction between WSS and DMS.

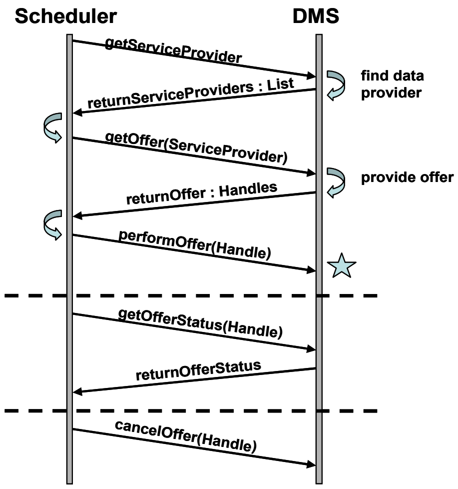

Figure 5: Negotiation between Scheduler and Data Management System

3.3 Data Management System

The task of the C3Grid data management is to provide means for making data available as input for grid applications. Actually, this comprises a number of different assignments: staging of data from primary data archives to local grid workspaces, transferring data between workspaces, delivering output data to the caller of the grid application and publishing results in the central data information service.

The C3Grid data management differs in some respect from common mainstream distributed data management systems. It actively supports its main client, the workflow scheduler, by performing its own planning of future transfers. For this task, it takes all information available about its current environment into account and corresponds with the scheduler to find agreements on transfer proposals.

The decision for exclusive access and management of grid workspaces as well as the design of a uniform interface to data archives are the key building blocks for making this possible. The latter interface keeps the type of storage system hidden from the grid environment but, nonetheless, provides a defined set of preprocessing operations which can be mapped, e.g., on queries of a relational database systems.

The web service based interfaces of the DMS to the primary data sources via the C3Grid specific Archive Interface and the WSS is displayed in Figure 6. The Workspace access is done by GridFTP and scheduled transfers. A system of naming conventions for consistent logical path names in the distributed C3Grid Workspace was established in such a way, that the logical path name of a dataset is valid throughout the whole C3Grid Domain (i.e. Logical file abstraction, location independent) and mapped to its physical location only when necessary, e.g. when a processing task or transfer request needs to know the physical location for direct access to this data item.

Not yet implemented is the DMS functionality for the planning of future transfers and the registration of generated ISO metadata for intermediate data products, which are generated in the workflow processing chain e.g. data items the user wants to keep for future references.

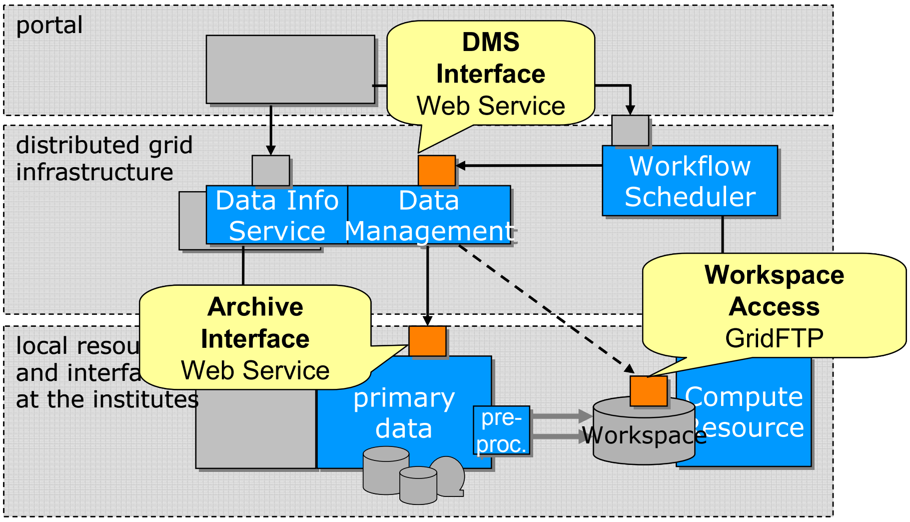

Figure 6: DMS interfaces

3.4 Data Information Service

The C3Grid data information service DIS uses the information from metadata to discover the data provided in the grid. The DIS is based on the Open Archives Initiative Protocols (OAI-PMH) and Apache Lucene, which provides a fast full text search engine what is displayed in Figure 7. The algorithms for the required range search capabilities were optimized to enhance the performance of range queries. Being fully Java-based the DIS is directly integrated in the user portal to avoid performance losses.

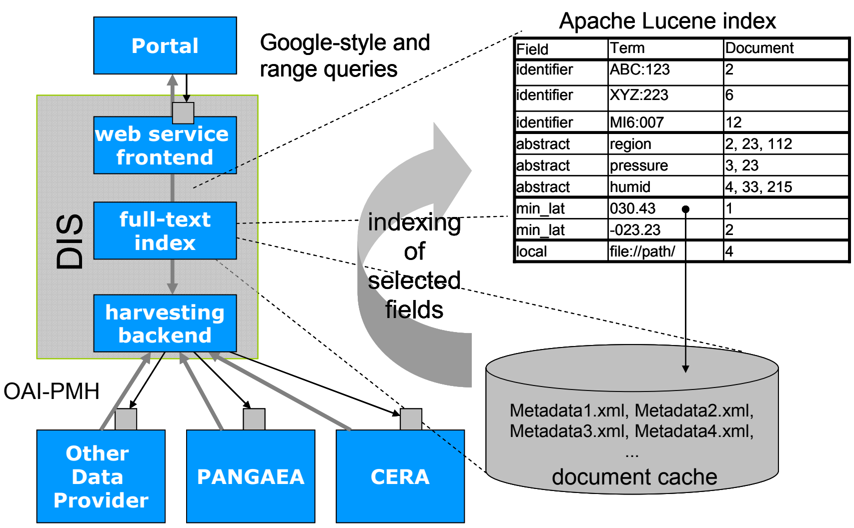

Figure 7: Architecture of Data Information Service (DIS)

Via a web service the portal user can issue Google-like queries for specific data sets, or query the index with physical variable names or select other constraints to the query he is interested in [4].

3.5 Data access

The data access strategy is based on standard web service technology and a community interface which specifies the users' time, space and variable constraints for the requested data as well as additional processing functionalities supplied by the provider. The provider specific implementation of the data access as well as of the data preparation is hidden behind a common WSDL defined web service interface.

The first implementation of interface between data management and data provider was made in Hamburg at the WDC Climate and the DKRZ. Thus the access to data held in the CERA database system and to data held as flat files in the DKRZ file archive was provided. The data request triggers appropriate pre-processing and filtering functionality.

3.6 VO Management and C3Grid security

The security infrastructure is an essential point for the sustainability of C3Grid. The implementation phase in 2007 is a stepwise approach towards a seamless overall security structure of C3Grid, which leverages the institutional Identity Management systems (IdM) of the C3 member institutions in a Shibboleth based AA infrastructure.

Initially, a simple Virtual Organisation (VO) structure is realised by setting the LDAP attribute eduPersonEntitlement to a common value ("c3-member") for each C3Grid member in the IdM systems of the corresponding home institutions. Based on a Shibboleth-enabled version of GridSphere the authentication and authorisation scheme is integrated at the Portal. The user has to pass an initial AA cycle before it is given access to the portal and mapped to a C3Grid account.

In this step of development, authorisation is done at the portal. The resource providers trust the portal. In late 2007 it is planned to replace the mapping to a single Grid account by individual accounts. At the same time, the authorisation process will be moved to the resource providers by adopting a GridShib use case, which was developed for TeraGrid Science Gateways [5]. The chosen approach will utilise the variant "Attribute Push with Local SSO" as in [6].

The following GridShib components will be used: Within the portal the GridShib SAML Tools will be used to embed a user’s authentication and authorization assertion into a X.509 proxy certificate. At the resource provider GridShib for GTK provides a SAML Policy Information Point (PIP) to support the parsing of the SAML assertion.

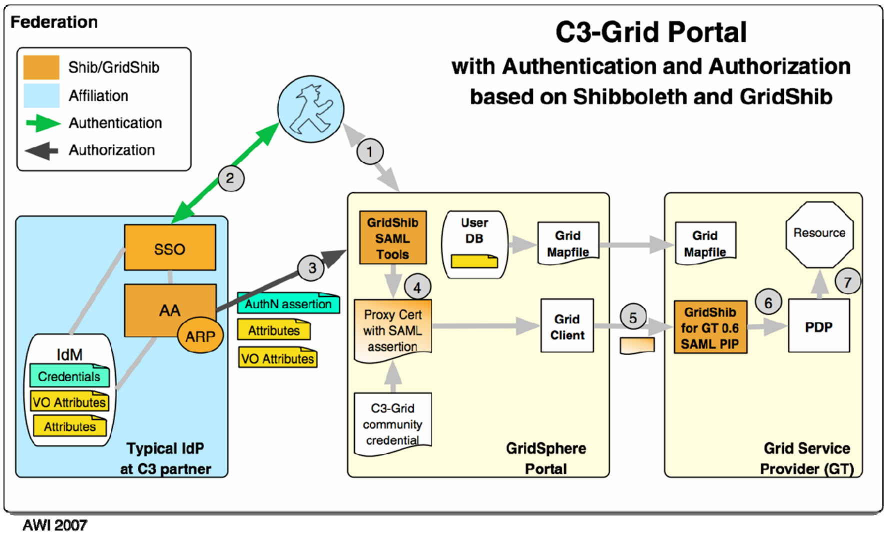

Figure 8: C3Grid Portal AA infrastructure based on Shibboleth and GridShib

An overview of the authentication and authorization process in C3Grid is shown in fig.7. The user enters the portal (step 1) and authenticates with the corresponding home institution (step 2), then the user's authentication and attribute assertion is transmitted to the portal with an attribute push operation (step 3), after verification of the assertion at the portal, the GridShib SAML Tools are used to embed the assertion into a proxy certificate, which will be signed by a community credential (step 4). In step 5 the Grid client presents the proxy certificate along with the job request to the Grid Service Provider, where the GridShib for GTK SAML PIP parses the assertions (step 6). Based on the PIP output the Policy Decision Point (PDP) makes an access control decision in step 7. If the decision is positive, access to the resource is granted and the job request will be processed (step 8).

In 2008, it is planned to integrate a central VO management system. By then, results of the IVOM GAP project should be available. IVOM was started within D-Grid to work on the integration and interoperability of VO management technologies in D-Grid.

4 Status and Perspectives

The C3Grid infrastructure will be built up in various phases. First, high attention was paid to the development of the infrastructure for data discovery and retrieval. Typical workflows for data analysis were defined. As an example, the analysis of stormtracks and the humidity flux analysis were implemented.

Actually, further data providers will be integrated in the C3Grid to broaden the data basis for the scientists. Similarly, further compute resource providers will join the grid in the next months. New workflows which will be implemented in the next phase will give the users more flexibility in their work.

Acknowledgements

C3Grid is a collaborative work of all participating institutions. A list of all staff members can be found under www.c3grid.de.

References

1. S. Kindermann, M. Stockhause, and K. Ronneberger: “Intelligent Data Networking for the Earth System Science Community” GES 2007
2. www.gridsphere.org
3. C. Grimme, T. Langhammer, A. Papaspyrou and F. Schintke: “Negotiation-based Choreography of Data-intensive Applications in the C3Grid Project“, GES 2007.
4. U. Schindler, B. Bräuer and M. Diepenbroek: “Data Information Service based on Open Archives Initiative Protocols and Apache Lucene”, GES 2007.
5. https://spaces.internet2.edu/display/GS/TeraGrid
6. https://spaces.internet2.edu/display/GS/NanoHUBTestbed