Semantic oriented data access and storage at MPIM/DKRZ

Michael Lautenschlager and Hannes Thiemann
Max-Planck Institute for Meteorology
Tel.: +4940 41173 297, Email: lautenschlager@dkrz.de
Tel.: +4940 41173 100, Email: thiemann@dkrz.de

Abstract

Problems in data access may be caused by large amounts of data as for climate research or Earth observation and by interdisciplinary data usage. Data organisation in files on the operating system level is not sufficient to fit the requirements of access performance and easy to use data extraction. Semantic oriented data handling is suggested as an alternative. Data access and storage are formulated within the semantics of the data model neglecting the details of physical storage.

Introduction

Two basic types of data are distinguished in climate research, (1)climate model and satellite data and (2) observational data. Climate model and satellite data are characterised by huge amounts of structured data which are produced by machines. Data normally reside on sequential storage devices (e.g. tapes) in a few locations. Data handling problems are related to the huge archive sizes. Satellite data archives normally supersede the climate data archive sizes by one order of magnitude.

Observational data are characterised by smaller amounts of heterogeneous data which are produced by ground-based measurements including human interventions. Machine readable data reside on random access devices (e.g. disks) in many different locations and structures. Data handling problems are related to the heterogeneous structures.

MPIM/DKRZ, as one of the European global climate modelling centres, focuses on climate model data management. The expected archive development is determined by the expected increase in compute power and related data production rates during the coming years.

Table 1. Expected archive increase at MPIM/DKRZ

End of year	Unix Files [Tbyte]	Database size [Tbyte]	Total archive size [Tbyte]
2000	120	10	130
2001	240	40	280
2002	480	120	600
2003	720	200	920
2004	1080	320	1400
2005	1440	440	1880

Data are organised and accessed in different units, data access units, data storage units and logical data units. The data access units are the smallest units for data access. For climate model data this is a horizontal global field of one variable for global climate modelling and as one number for instrumental observations. The data storage units are the file sizes which are used for data storage on the operating system level. This reads as one month accumulated model output for global climate modelling. The logical data unit summarises the entire model output which corresponds to one climate model integration. The separate data units are determined by the model resolution. Examples are given for the standard climate model ECHAM4 in T42-resolution (standard version) and in T106-resolution (high resolution version). The T42 version corresponds with a horizontal resolution of 2.81 degrees or 8192 points in the global grid, the T106 version with 1.13 degrees horizontal resolution or 51200 grid points. The separate data units then are in GRIB storage format for the 19 level version of the climate model:

Table 2. Climate data units

	ECHAM4 (T42)	ECHAM4 (T106)
Data access unit (1 global horizontal field)	17.1 kB	100.1 kB
Data storage unit (1 month model output)	120 MB	750 MB
Logical data unit (100 year model integration)	140 GB	880 GB

The results are stored with 2 Bytes accuracy in the data record of the GRIB format. If the results are given in IEEE floating point 4 Byte accuracy or in 8 Byte accuracy the values in the table have to be multiplied by 2 or 4, respectively.

Climate model output is normally stored in a raw data format (data storage unit) as it is directly produced by the models. The model raw data time series contains the overall amount of model variables at each time interval. The user access is perpendicular. The normal request is for time series of separate model variables in order to identify climate development in terms of temperature, precipitation or wind.

Extraction of individual time series from the model raw data infer the de-migration of 1200 files from the mass storage archive. The number 1200 corresponds with a one hundred year model integration. The total number of files increases if the data storage unit is less than one month or the time series increases. From each raw data file the small requested data fraction (less than 1%) has to be extracted. This procedure is as time consuming as the model calculation itself. Extracting 1% of data from 1200 raw data files bears 3000 hours processing time and 400 hours access time on a well equipped client work station. The data access and processing hours translate to 142 days if there are no problems with the environment. In practise a factor of two has to be added (problems with network and computing environment) yielding in more than 9 months user client process time for 100 years T42 model output. Even on the present-day DKRZ data server only a factor of approximately 7 can be saved (300 hours data processing and 200 hour data access).

But even for the relatively small mass storage archive at MPIM/DKRZ mid of 2000 with 100 Tbyte data organised in 3 million files data extraction problems are obvious.

- Storage in classical file systems is not sufficient to organise the climate model data. Descriptive information about the file content (data catalogue) is widely missing.
- The access performance is limited due to storage on sequential devices and the 99% overhead in data download.

In order to improve the data access the concept of the semantic oriented data access has been developed. The access interface is the data model (data model, data catalogue). Information requests are formulated within the semantics of the data model. The data system then collects and processes the required data and disseminate the extracted information to the user. The data extraction is supported by an application adapted climate model data hierarchy. Both, the data model and the data hierarchy are part of a database system and can be uniquely accessed by the same application interface.

The access example is substantially improved by the semantic oriented data access. The 1% requested data may correspond with the surface climatology, namely the four time series for temperature, precipitation and the two wind components. These data are retrieved by one tape mount. The same interface allows for catalogue queries and for data extraction. SQL selection of parts of the time series is supported, no data processing and no raw data movement is performed. The data access is improved by at least a factor of 1000 compared to the raw data access example. The data movement is reduced by a factor of 100 because only the requested 1% data are transferred.

Moreover, in designing such a system one should be aware that a file system invariably has higher performance for a single file access than any more sophisticated system. The bottom line is simple: if you do not need the service, then there is no need to pay for it in terms of lower single performance.

Semantic oriented data handling

The semantics of the data are stored in the metadata. The corresponding data model is the interface to the archive. Information requests are formulated within the semantics of the data model. The data system then collects and processes the required data and disseminates the extracted information to the user. Data, especially global climate model data, are stored automatically in the archive within the framework of the data model semantics and data hierarchy. Essential is that this automatic storage has to be performed before the climate model results are migrated to slow sequential storage devices (tapes).

The semantic oriented data handling at MPIM/DKRZ is realised in the CERA system (Climate and Environmental data Retrieval and Archiving system). The CERA system is implemented using the Oracle 8i database management system and consists of three major parts:

- data model,
- data hierarchy and
- standard data processing.

The CERA system improves data access and consistency, supports data security and provides content and quality information.

The CERA data model has been developed in co-operation between PIK, DKRZ and AWI in order to implement in some sense a standard for climate and environmental data models and to incorporate climate model and observational data. The CERA data model incorporates the basics of the IEEE reference model. Meta data are defined as the information required to make scientific data useful:

1. query, browse, retrieval;
2. ingest, quality assurance, reprocessing;
3. application to application transfer;
4. storage, and archive.

Relevant international description standards like CSDGM (Content Standards for Digital Geospatial Metadata developed by Federal Geographic Data Committee) and DIF (Directory Interchange Format developed by NASA) are included to allow for an information exchange with other data archives. A documentation of the data model can be obtained from under Report 15: "The CERA-2 Data Model", the current development is documented at .

The CERA-2 data model defines an open, extendable structure which allows not only for the storage of climate model data but also of observational data. Any hierarchical and network data organisation can be implemented and relevant description standards with respect to database interoperability are matched. The CERA-2 core, that is the mandatory information, is defined as small as possible to implement a basic data description.

The CERA-2.4 SCHEME

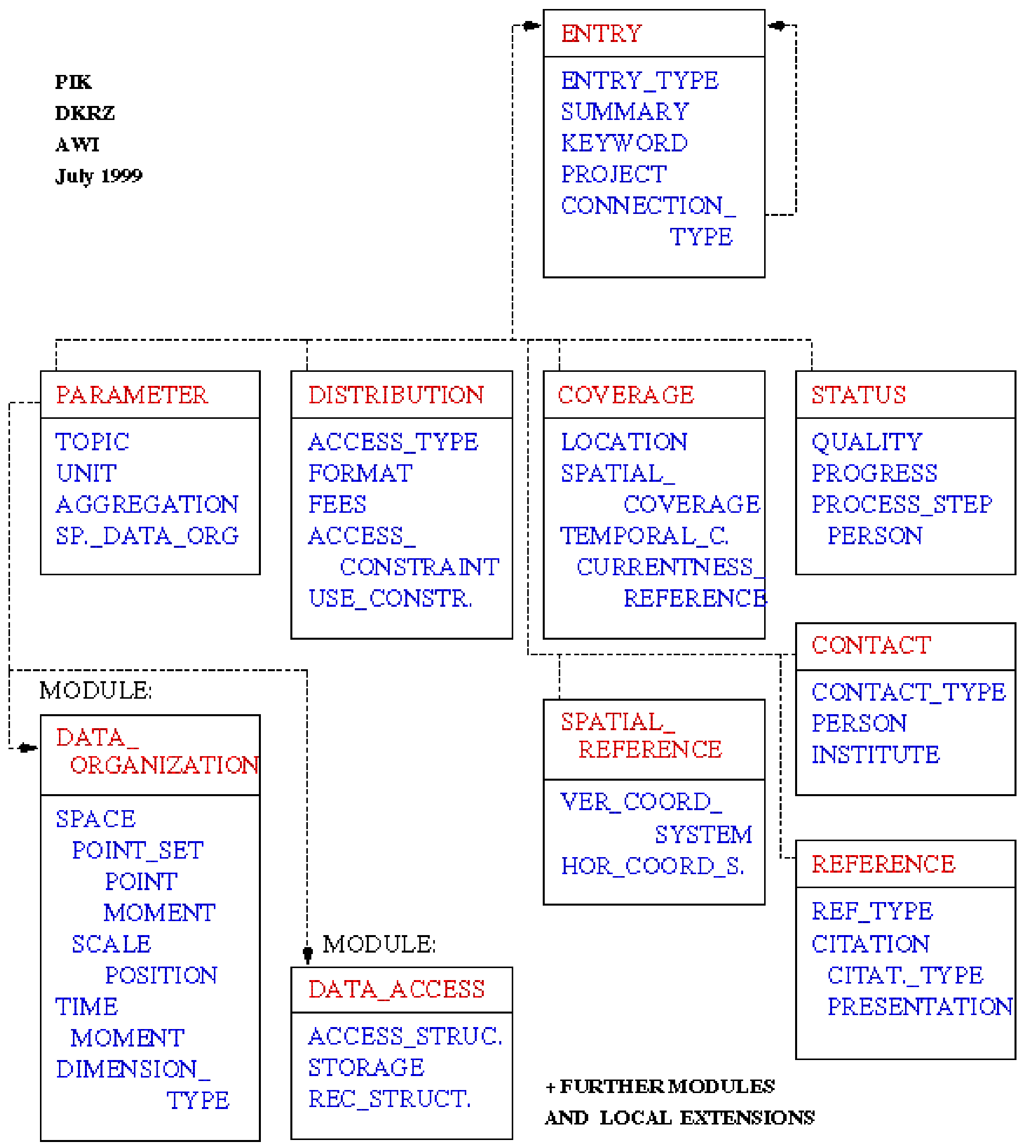

Figure 1: Block diagram of the CERA data model.

The CERA data hierarchy has been defined based on user group discussions and multi-year experience with climate model data access. The data hierarchy is implemented in the data model and consists of mainly three layers, the model raw data, the processed data and the meta data.

The model raw data contain climate model data as they are calculated by the numerical models. The raw data files are time series of 4D data blocks. All model variables at all points and in all layers are stored at a specific point in time. They are stored as 4D BLOB's (Binary Large Object) in database tables. The table represents the time series. These data are stored preferably on sequential access devices (e.g. tapes). They are used for backup and for special evaluations which are not fit by the processed data layer. The effort to extract data from this raw data structure has been discussed in the previous section under access problems.

The processed data layer contains climate data which are split up into time series of separate variables. The data are stored in database tables as 2D BLOB's. For the atmospheric part of the climate model this means global horizontal fields of for example temperature, precipitation or wind speed. The transition from the raw data structure to the processed data requires a number of operations on the raw data: extraction of specific variable, calculation of diagnostic relations (e.g. wind components from divergence and vorticity), vertical interpolation from model into pressure levels, and spectral transformation into grid points. The entire amount of data is not reduced by these operations but the access to the data is usually much faster. In order to get a time series of one model variable only one data and tape access is necessary and the entire time series or parts of the data can be transferred to the user client machine. In comparison, the same operation on the raw data structure requires data extraction month by month and piecewise extraction and creation of the requested time series. These processed data are preferably stored on random access devices (e.g. magnetic disks) in order to reduce the tape mounts and to increase the access performance.

The meta data contain the data catalogue and information data quality, data access and data processing. This inventory information is small compared to the climate data and is stored exclusively on magnetic disks. The access should be as fast as possible at any time in order to get the data request answered and to specify the data retrieval.

The CERA standard data processing is based on the GRIB format which is used at DKRZ for data storage of the atmospheric part of the climate model. The PINGO package (Procedural Interface for GRIB-formatted Objects) is a first version of standard data processing routines which can be used inside and outside of the CERA database system. The PINGO routines are based on the GRIB format and integrate a collection of a few hundred processing routines, which are predominantly merged from former packages of processing routines. The PINGO package allows a unified access to all routines. The package is written in ANSI-C and transportable to different platforms. The PINGO routines are an independent program package. They allow for data format conversion, information of file content, arithmetic operations and some statistical calculations. The PINGOs fit to the CERA data structure.

So far implemented is the semantic data access. Under development is the semantic oriented data storage. The implementation is necessary in order to keep up with the expected increasing archive rates.

The actual size of the CERA database is 2.5 Tbyte (11.05.00) which is already too much to store on magnetic disks. Therefore the direct connection of the ORACLE DBMS to a to tape of a mass storage system has to be implemented in order to migrate and de-migrate data. This mechanism is not part of the DBMS (database management system) or of the HSM system (hierarchical storage management) which administers the mass storage archive. A solution based on the ORACLE DBMS but independent from a special HSM system has been developed and implemented at DKRZ/MPIM.

User application interface

The user application interface is mainly restricted to the semantic oriented data access. Data storage is directly organised between the climate models and the CERA database system. The storage process is not visible for most of the users, only a few data model applicators <applications?> are involved.

Different user application interfaces are implemented for the retrieval, graphical user interfaces and a batch interface. Two graphical user interfaces allow for inspection of the complete data model. Question like what data are available and are these data useful for my application can be answered. The graphical user interfaces allow for three different catalogue access possibilities, namely browse, keyword or specific access.

The browse-window allows for a hierarchical catalogue access. Data are organised in climate modelling experiments and related data sets. The data sets contain the implemented data hierarchy which is an integral part of the CERA database system (BLOB data tables). The experiments are organised in different projects. They are additionally searchable by general keywords and spatial coverage. The browse-window allows for the specification of experiments which are related to a special project, a keyword and a spatial coverage. The resultant list of experiments can be inspected one by one in terms of the general description, the related institutes and the included data sets.

The key word access allows for data inspection by specification of search criteria within a three layer keyword hierarchy. The query result is a list of data sets which match the search criteria. For each data set the metadata as well as the related experimental information can be obtained. Useful data sets can be marked for a following data processing.

The most direct data search is the specific access. It allows for direct specification of data set names or of a part of the name. The query result lists all data set names which match the criteria. The data set handling possibilities themselves are the same as for the keyword access.

Finally in the process window the marked data sets or parts of the data sets can be transferred from the database server to the local client machine.

Figure 2: CERA-GUI browse window

Technically speaking, these graphical user interfaces are realized in two different ways. The first is based on Oracle Forms, an Oracle specific 4GL. Due to license restrictions this GUI is not available for all platforms and therefore only accessible within the intranet. The second is a very simple, CGI based Web interface. It allows the access of the metadata as well as download of public domain data in small amounts via the ftp-server. Flexibility and performance of this interface is low. The access to this interface is open to the internet.

The batch interface allows for the direct access of CERA data sets or parts of these data sets in UNIX scripts. The exact name or identifier of the data set has to be specified. This program uses oracle specific extensions and is therefore, like the Oracle Forms GUI, not available for all platforms and is accessible only to the intranet.

Access permissions are managed by the database server and not by the user interfaces (Applications.) Nevertheless access rights are shown to the users as well as sufficient information of whom to contact in order to get access. All metadata is open to the public.

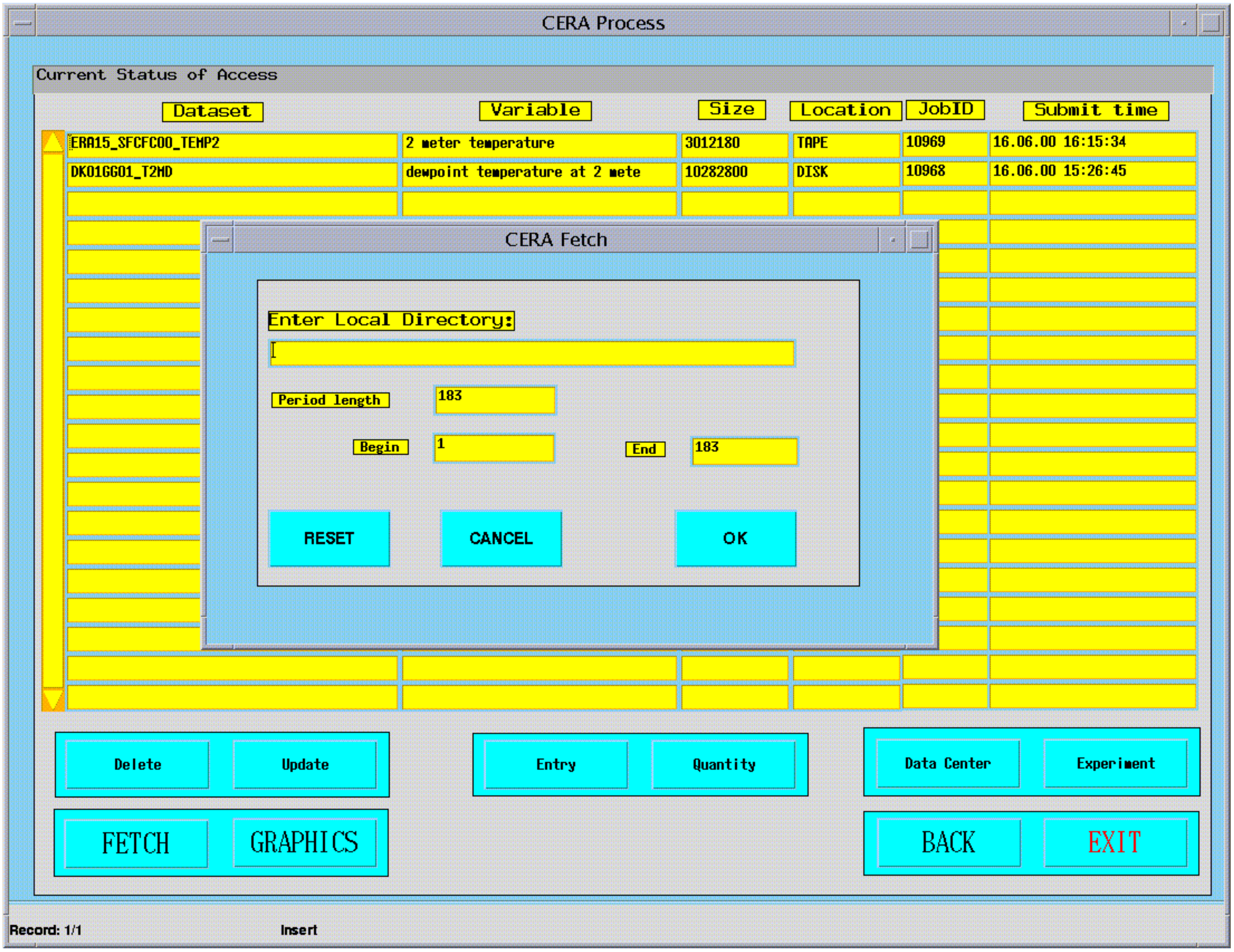

Figure 3: CERA-GUI process window

To overcome restrictions of all of these interfaces the development of new user interfaces is necessary. The successor generation of user interfaces should be hardware independent as the support of different platforms is no longer practical and due to license restrictions not affordable. Therefore Java resp. JavaApplets (Java version 1.2) have been chosen to fulfil these requirements. A prototype for the GUI is currently being tested. Within this development a flexible approach has been selected to allow for further changes within the underlying data model. The batch interface has been functionally tested but a prototype is not yet available. Future plans include the implementation of all existing services into the CORBA model, including a data-processing capability. Special attention has to be given on security considerations as well as on performance issues.

Conclusion and outlook

Currently available is a WWW-based interface for the dissemination of IPCC climate scenario data. This interface does not allow for a general database access. It is restricted to the IPCC DDC which is part of the CERA database system and the included climate model results from the international climate modelling community.

Emphasis will be on the implementation of the semantic oriented data storage of climate model data in order to represent the complete archive in the CERA DB and on the WWW-based graphical user interface in order to open the archive for general internet access. Based on CORBA mechanisms other data services can potentially integrated into the CERA database environment.