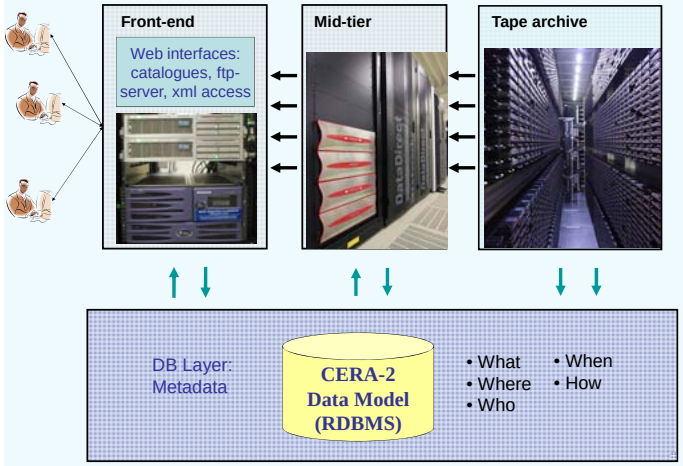
Technical Aspects of Data Life Cycle Management at WDCC

IN53A-1606

WDCC repository architecture

The WDCC provides fine-granular field-based access to climate data. The volume and age of the data managed (current as well as expected) implies that tape storage is the most reliable and economically efficient solution. An appropriate interface to the tape archive is thus needed.

An hierarchical storage management system has been developed, wherein time step data of a few dozens of kilobytes are packed to datasets, and datasets are packed into container files. Such container files are stored on HPSS tapes, whereas all access information is kept in the CERA-2 RDBMS (currently Oracle RAC).



The CERA-2 Data Model

...can represent Metadata structures of ISO, Dublin Core (DC) and many other standards. It has been used at WDCC for more than ten years. Only minor changes and few supplements have had to be made.

Repository hardware/software

HPSS: High Performance Storage System

- 6 Sun StorageTek, SL8500 tape libraries, 8 robots per library
- 73 tape drives
- 60.000 media (T10000 A/B, LTO4, 9940B and 9840C)
- 500 Terabytes disk cache (DDN)
- bidirectional bandwidth: 3 Gigabytes/s (sustained), 5 Gigabytes/s (peak)

Mid-tier organization

The mid-tier consists of a disk-only distributed file-system that is used as a cache for data-set files on tape, and an internally developed software system named Lobster, which presents a conventional JDBC interface to the front-end system. The front-end can thereby access each data-field present in the archive by submitting an SQL-query of the form:

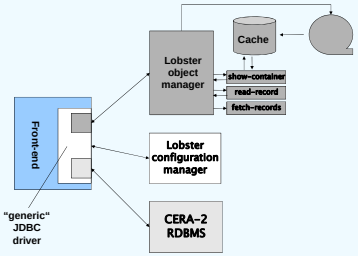
```
SELECT  blob_id, blob_data  FROM
   FROM CERA_BLOB.dataSetName
WHERE blob_id >= x and blob_id <= y  ORDER BY blob_id;
```

through JDBC and obtain the contents of the column named blob_data by calling the getBlob method of the ResultSet object for each row of a result set. If the requested data-set is not present on disk it is automatically retrieved from tape by the Lobster software.

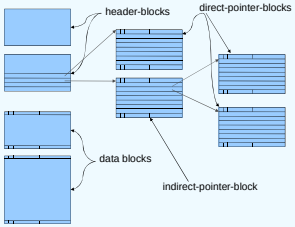
The lobster configuration manager is the first point of contact for the JDBC driver, and indicates to each client which additional drivers to load and with which hosts to communicate.

The communication with the lobster configuration manager and loading of additional drivers is performed by the "generic" JDBC driver (see figure).

Thereby, front-end software need not be concerned with the physical organization of data, location of services, *etc*. Indeed, different users can be directed to interact with completely different RDBMS and lobster object managers.



Storage format



Time-series of single variables are stored in so-called *container files* (the unit of transfer to and from tape). Large variabilities in both record counts and record sizes led to a block-based storage format with multiple block-sizes (2^k- byte blocks, for k = 8,...,62), and an index structure for records similar to that of i-nodes and indirect blocks in Unix file-systems.

Each direct-pointer block keeps information about 186 different records, including references (location & size) to 9 data-blocks for each record. Each indirect-pointer block contains references to 2338 direct- or indirect-pointer blocks.

Finally, all information in direct- and indirect-pointer blocks is duplicated and all blocks are doubly labelled, to improve recovery prospects in case of accidents such as torn tapes, *etc*.

The DKRZ GmbH is financed by the German Government and big research organisations. It offers High Performance Computing facilities to the Earth System Research Community.

Climate analytics on distributed exascale data archives

Climate science demands on data management are growing rapidly as climate models grow in the precision with which they depict spatial structures and in the completeness with which they describe a vast range of physical processes.

This project will explore the challenges of developing a software management infrastructure which will scale to the multi-exabyte archives of climate data which are likely to be crucial to major policy decisions in by the end of the decade. Support for automated processing of the archived data and metadata will be essential.

Partners

British Atmospheric Data Centre · Princeton University · DKRZ · University of Toronto · Institut Pierre Simon Laplace · UCLA · Centro Euro-Mediterraneo per i Cambiamenti Climatici

Climate Data Operators (CDO)

CDO is a collection of command line Operators to manipulate and analyse Climate and NWP model Data.

ExArch Web Processing Service

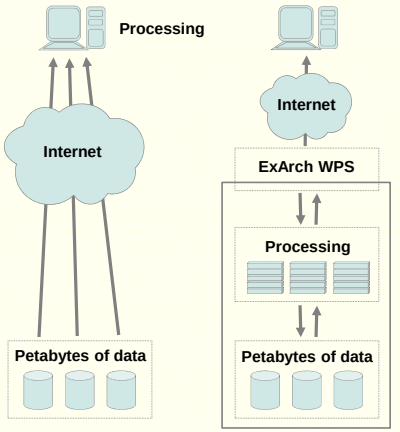
Past: ExArch:

The ExArch approach

The ExArch Web Processing Service brings CDO to an exascale archive.

Features & Benefits

- Flexible transparent and uniform taskmanagement due to a standardised Web Processing Service.
- Performance enhancement in consequence of distributed computing using dedicated hardware and fast local data access.



N. Peter Drakenberg, Gregory Foell, Frank Toussaint, Michael Lautenschlager, Stephan Kindermann, Hannes Thiemann

World Data Centre for Climate at Deutsches Klimarechenzentrum, Bundesstr. 45a, 20146 Hamburg
drakenberg@dkrz.de foell@dkrz.de

Abstract Reference Number 1185994