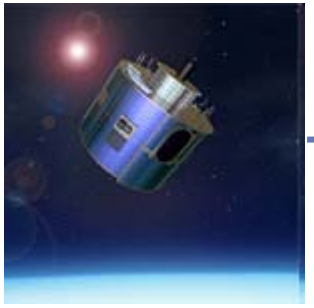
SKY - the new data management system at Deutscher Wetterdienst

H. Lemmin, Data Management Unit, Dept. Systems and Operations,
Business Area Technical Infrastructure and Operations, 3.11.2009

SKY - the new data management system at DWD

- **Observational data flow using SKY**
  - **Store**
  - **Read**
- User friendly query language
- Flexible storage architecture
- Query processing
  - small query
  - large query
- About SKY
- Future

Observational data flow using SKY – Store



FM…
BUFR

HDF
NetCDF
Bzip2 ...

GloBUS decoder/ encoder

metadata generators

Standard BUFR + metadata

… + metadata

SKY Server „CIRRUS“

Java Server Program

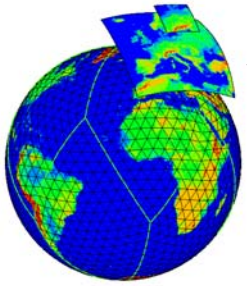
meta data
4 billion records
ORACLE

binary data

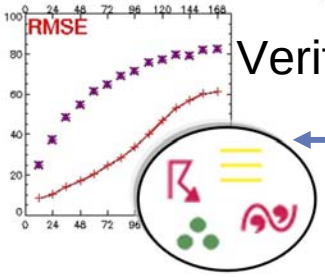
/online data
1.5 TB

/offline data
65.5 TB

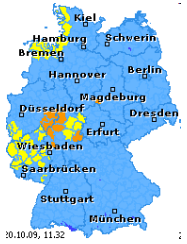
Observational data flow using SKY – Read



Numerical Weather Prediction



Verification



And many others…

Clients

- Commandline (using Java)
- Java API
- C API
- Fortran API
- GUI Client

Requests:

- XML or
- short language

70.000 requests per day

SKY Server „CIRRUS“

Data categories (59)

- BUFR
  - Dispersal and Transport
  - Radar Data
  - Radiances Satellite
  - Single Level Upper Air Data
  - Single Level Upper Air Satelli
  - Surface Data Land
    - Aviation Report
    - GNSS Data
    - National Data
    - Synoptical Data

SKY - the new data management system at DWD

- Observational data flow using SKY
  - Store
  - Read
- **User friendly query language**
- Flexible storage architecture
- Query processing
  - small query
  - large query
- About SKY
- Future

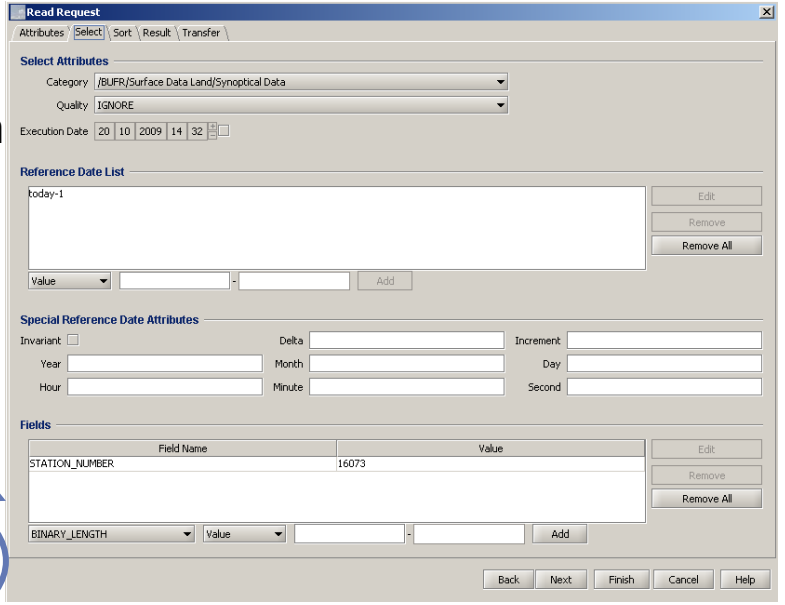
User friendly query language

➔ as XML: self descriptive

```
<requestCollection>
   <read database="cirrus">
      <select category="/BUFR/Surface Data Land/Syn
         <referenceDate>
            <value>today-1</value>
         </referenceDate>
         <field name="STATION_NUMBER">
            <value>10637</value>
         </field>
      </select>
   </read>
</requestCollection>
```

generates, edits

➔ GUI Client: Query Builder



➔ as short language: short, writeable

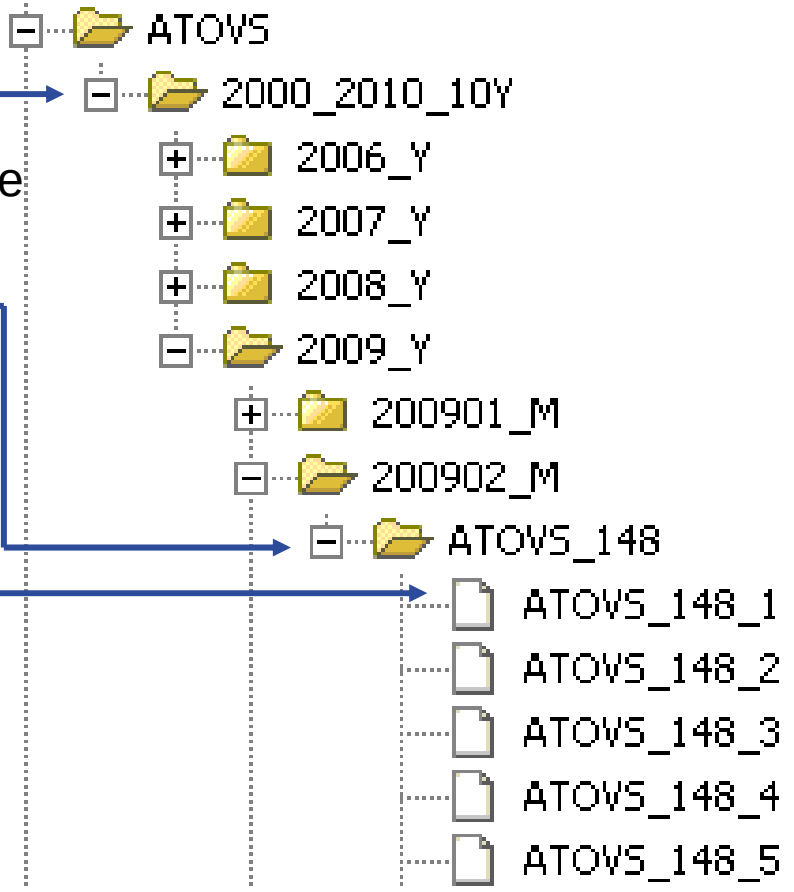
```
read db=cirrus cat=sdlsynop d=today-1 stnr=10637
```

SKY - the new data management system at DWD

- Observational data flow using SKY
  - Store
  - Read
- User friendly query language
- **Flexible storage architecture**
- Query processing
  - small query
  - large query
- About SKY
- Future

Flexible storage architecture

- Configurable for each data category
- Period nodes
  - group the data by time and determine the size of the metadata containers
- Metadata container
  - stores the metadata information in one database table
  - feasible table size of ~5 million records
- Storage container
  - stores the binary data in one file
  - feasible file size of 4-8 GB
  - online or offline

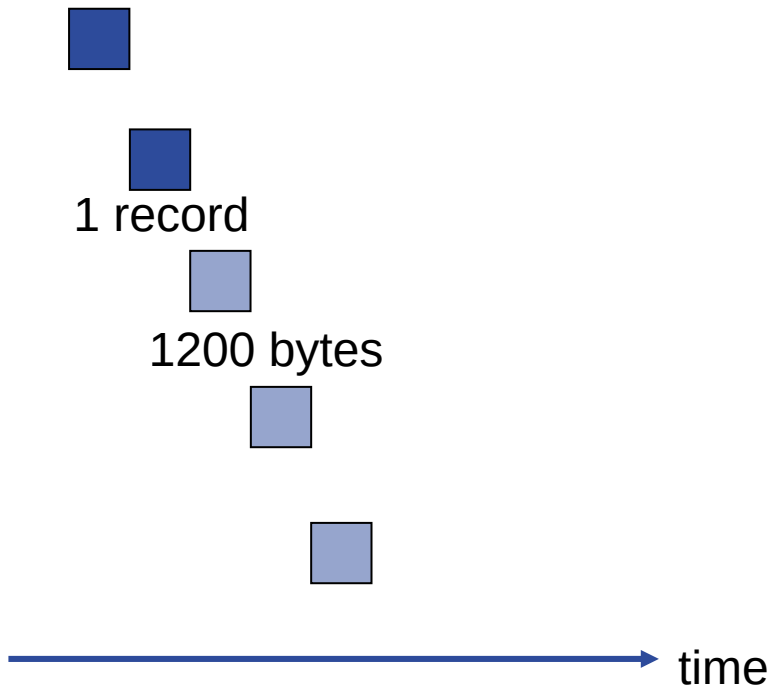


SKY - the new data management system at DWD

- Observational data flow using SKY
  - Store
  - Read
- User friendly query language
- Flexible storage architecture
- **Query processing**
  - **small query**
  - **large query**
- About SKY
- Future

Query processing – small query

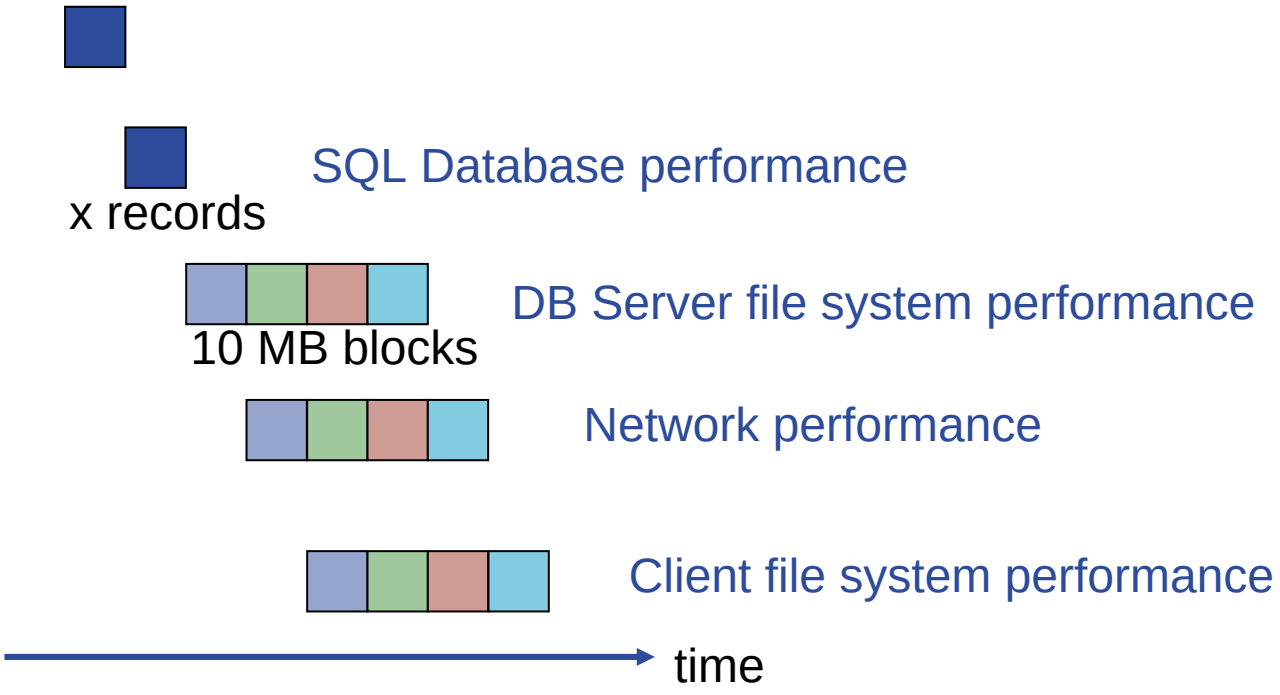
- Client: read and send query
- Server: read meta data stream
- Server: read binary data block
- Server: send binary data block
- Client: write binary data block



- de facto sequential processing – server side 40 ms – 30 KB/s

Query processing – large query

- Client: read and send query
- Server: read meta data stream
- Server: read binary data block
- Server: send binary data block
- Client: write binary data block



- asynchronous processing + streaming – server side top rates up to 350 MB/s
- performance is limited by performance of components and total load

SKY - the new data management system at DWD

- Observational data flow using SKY
  - Store
  - Read
- User friendly query language
- Flexible storage architecture
- Query processing
  - small query
  - large query
- **About SKY**
- **Future**

About SKY

- SKY server components are being developed by Deutscher Wetterdienst
- SKY GUI components have been developed by Ernst Basler + Partner GmbH
- Quality test suite contains more than 3600 automated test cases
- Performance test system for stress testing and performance analysis
- SKY is used by Deutscher Wetterdienst and KNMI
- SKY and GloBUS are marketed by EuMetSys:

EuMetSys c/o Ernst Basler + Partner GmbH
Tuchmacher Str. 47, 14482 Potsdam, Germany
Phone +49 331 74759 0
Fax +49 331 74759 90
www.ebp.de
info@ebp.de

Future

- SKY will store GRIB1 / GRIB2 and replace the CSOBANK system
- Queries will use GRIB API shortnames
- Categories will describe themselves according to ISO Metadata Standard
  - Integration of SKY Systems into the GISC Architecture

Questions

Harald Lemmin
Deutscher Wetterdienst
TI 12c
Frankfurter Strasse 135
D-63067 Offenbach
Phone +49 (0)69 8062 - 2556
Fax +49 (0)69 8062 - 3829
Harald.Lemmin@dwd.de