Long-term Archiving of Climate Model Data at WDC Climate and DKRZ

Michael Lautenschlager
WDC Climate / Max-Planck-Institute for Meteorology, Hamburg

Wolfgang Stahl
German Climate Computing Centre (DKRZ) Hamburg

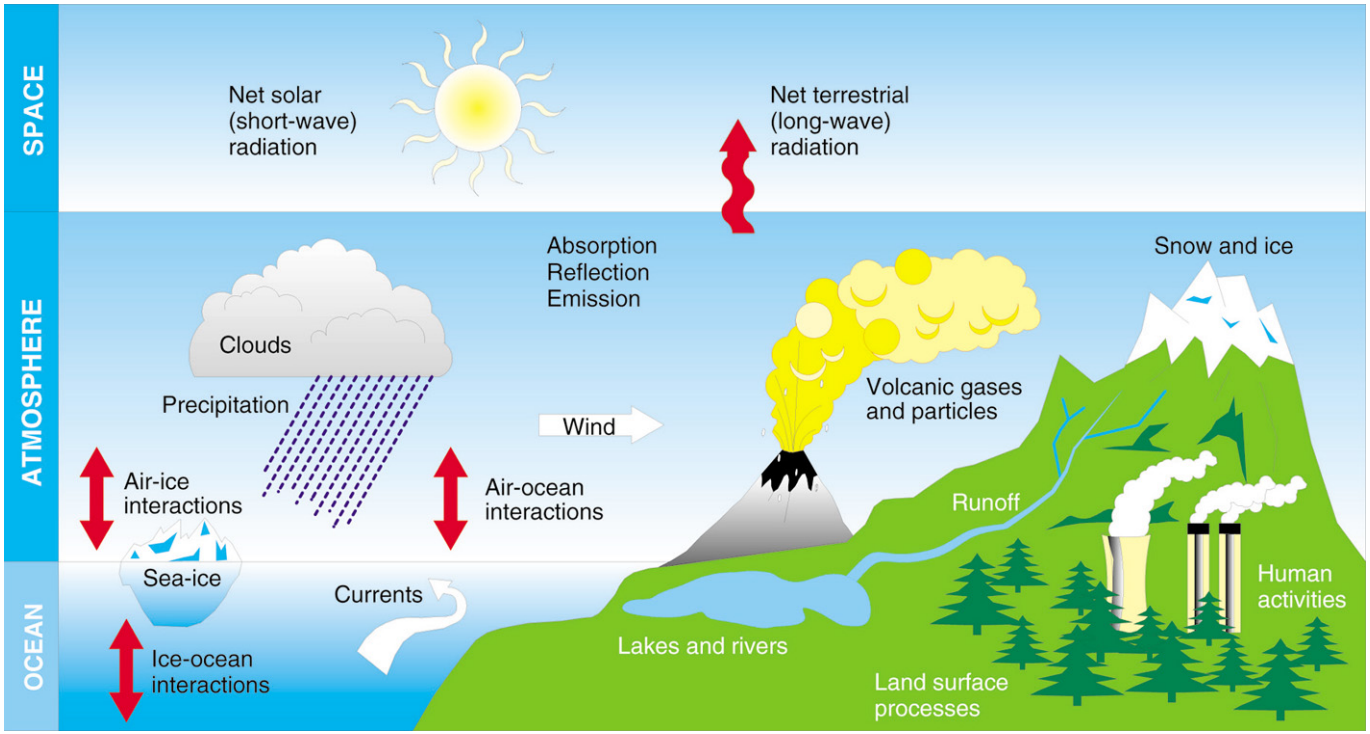
MPG eScience Seminar: Aspects of Long-Term Archiving
June 19th – 20th, 2008 in Göttingen, Germany

Next generation of compute server and of climate models leads to a problem at DKRZ:

Data production increase with implications for long-term archiving

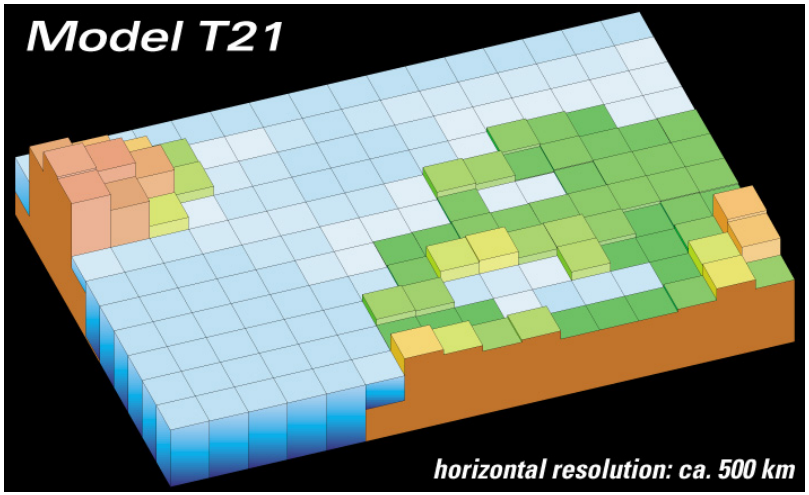
What is the problem?

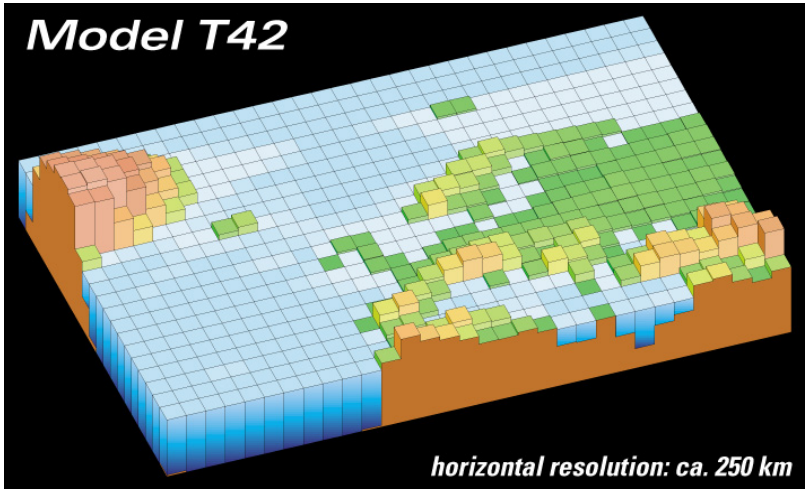
What could be the solution?

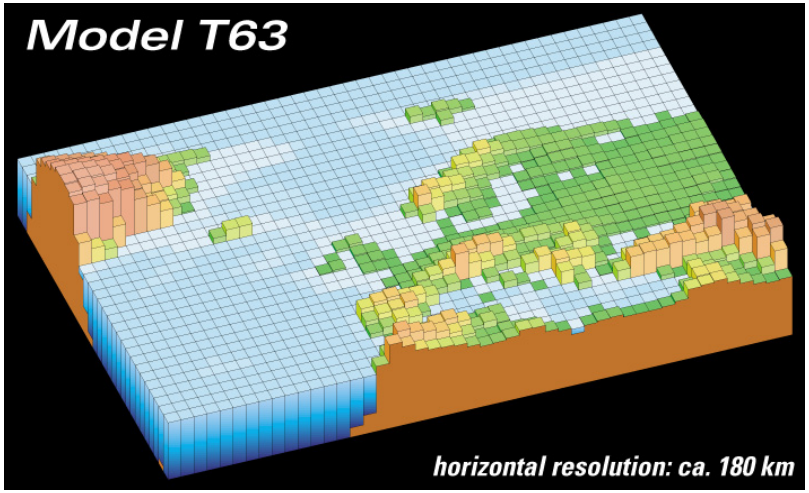


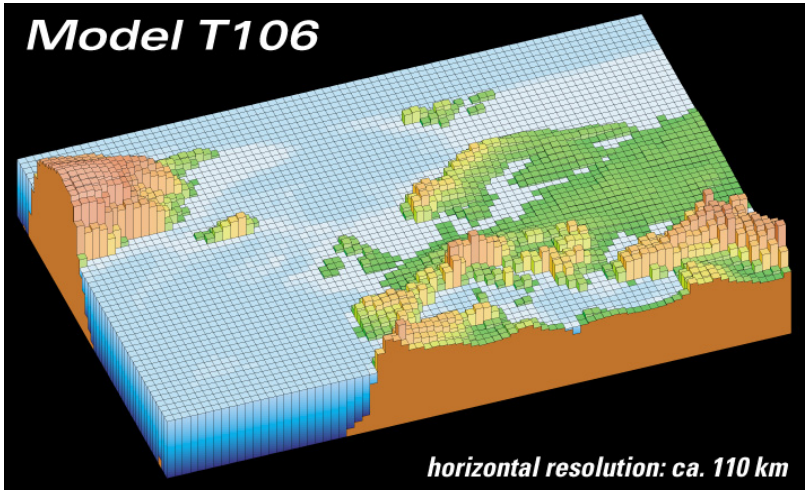
Increase in installed compute power motivates

- finer spatial and temporal model resolution and
- integration of additional physical and chemical processes into climate models.





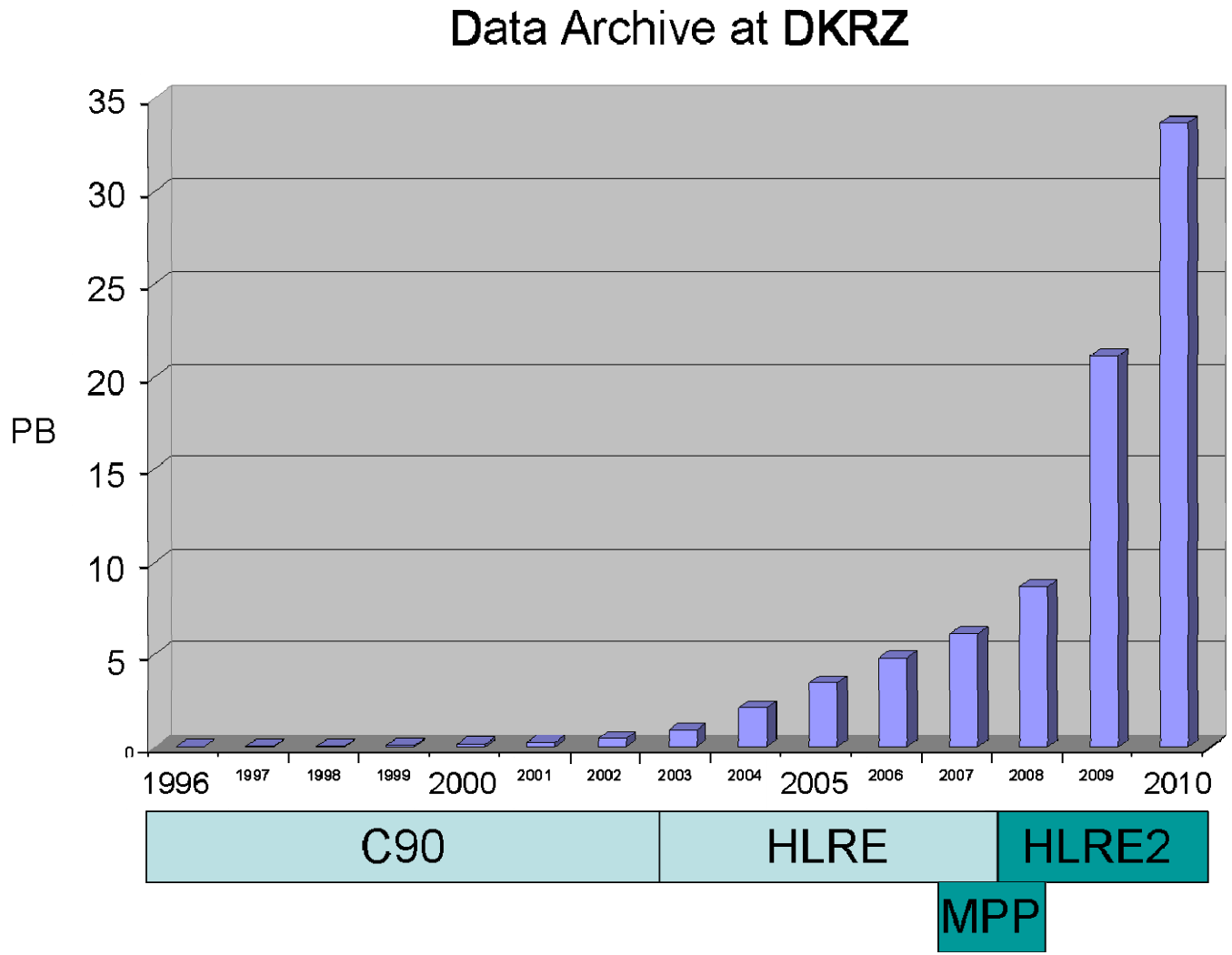




- Finer spatial (+ temporal) model resolution or
- Execution of climate model ensemble runs

- Experience at DKRZ: Linear increase in data production with installed compute power
- Previous data storage strategy: all data migrated to the long-term mass storage archive
  (that means archive increase follows directly the compute power increase)
- Resulting problem: Since the total amount of money for investment and for operations is fixed the cost relation between compute service and data service shifts towards data service while reducing the compute service fraction

- Example: archive increase and explosion of media costs with next generation of compute services assuming 10 * HLRE

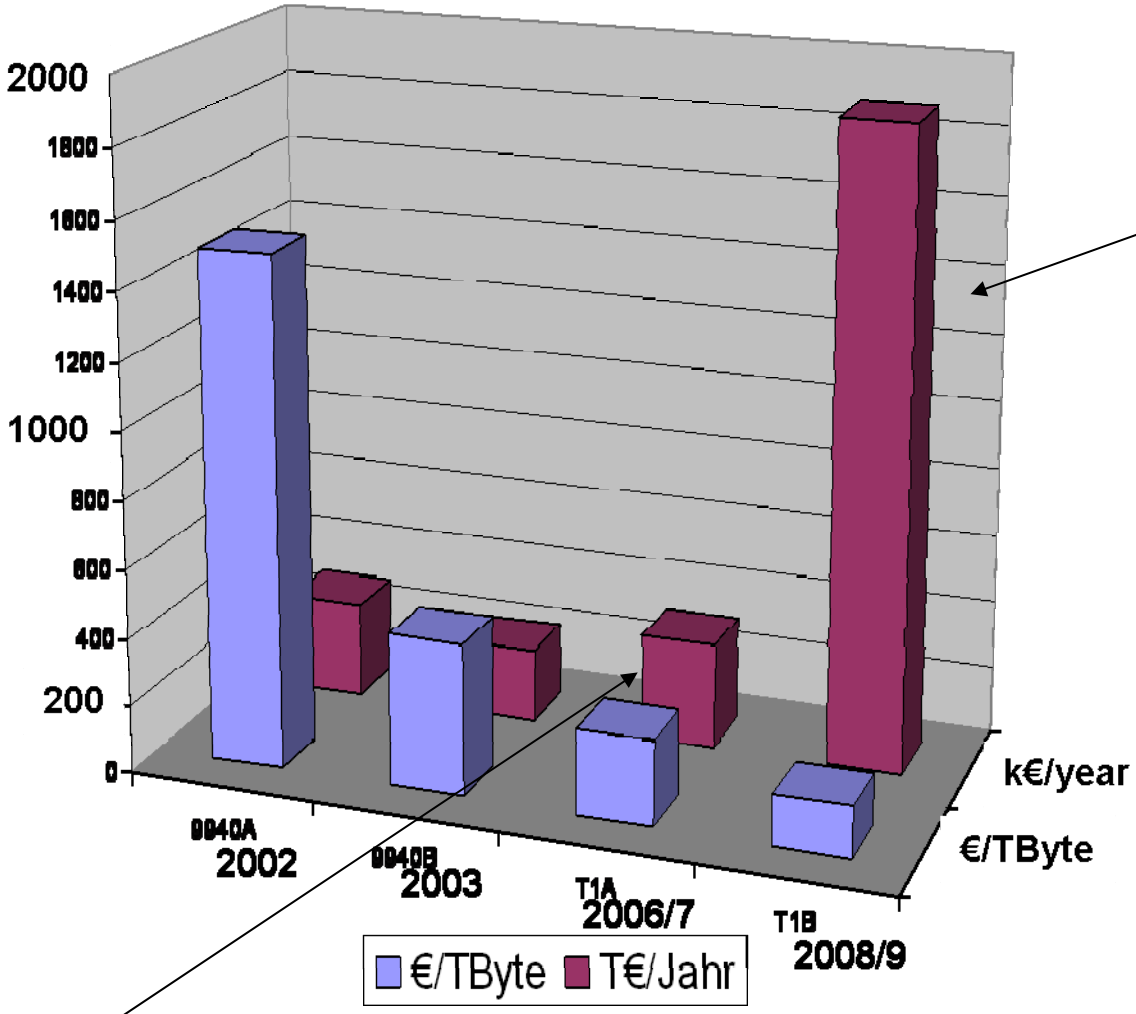


Extrapolated HLRE2 linear archive increase (10 times HLRE)

Compute server architectures:
C90: Cray C90 / HLRE: NEC SX-6 / MPP: SUN-Cluster / HLRE2: new system

(HLRE: Höchstleistungsrechnersystem für die Erdsystemforschung)

Development of media costs vs. total costs per year (Feb. 2007)



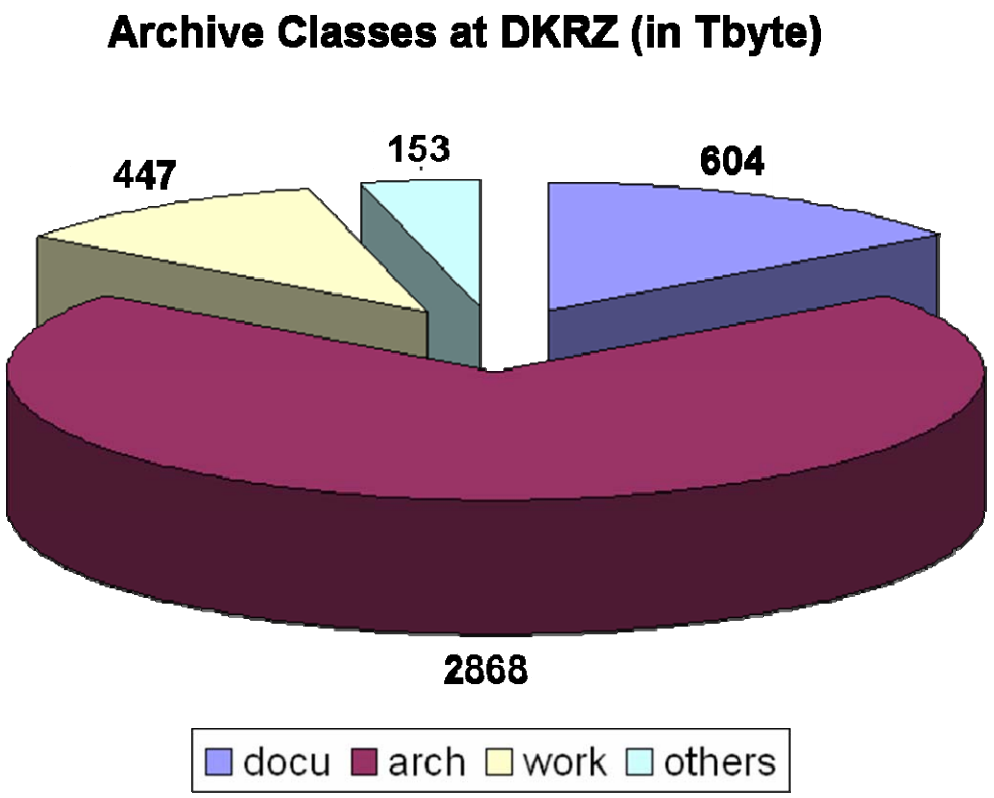
Not accepted projection

Discussion about total media costs starts

What could be the solution for mass storage cost balancing?

- Data classes
  - Test data from model code development,
    life cycle: weeks to months
  - Project data from scientific model evaluation and research projects (DKRZ resources at project level),
    life cycle: 3 – 5 years
  - Final results as data products for international projects (IPCC) and scientific publications,
    life cycle: 10 years and longer

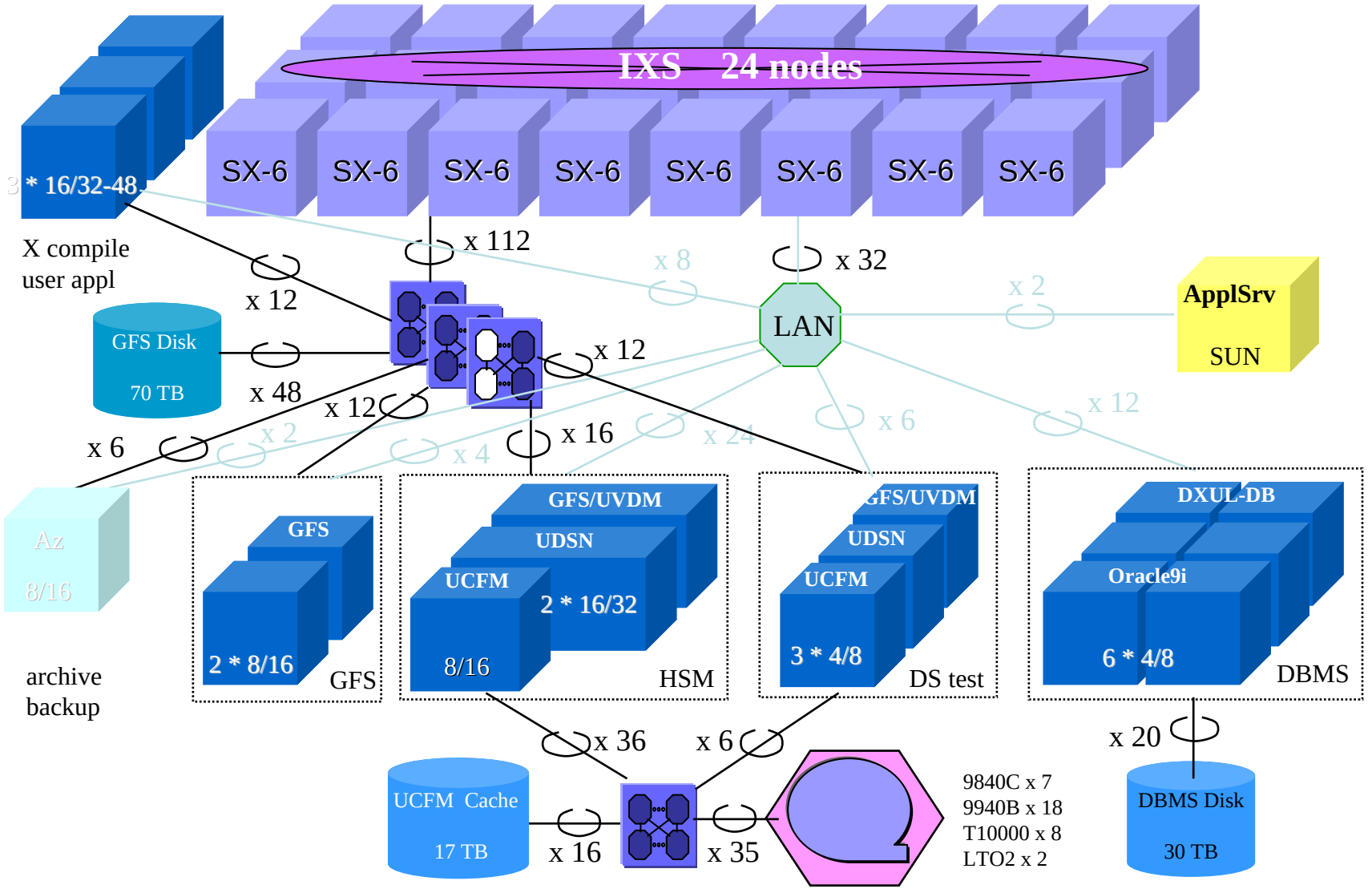
- Data hierarchy levels
  - Temp(orary)
    scratch discs at compute server
  - Work
    fixed disc space at project level for evaluation
  - Arch(ive)
    tape storage space (single copy) with expiration date for project data beyond available disc space
  - Docu(mentation)
    documented, long-term tape archive (security copy) for data products, focus on interdisciplinary data utilisation,
    data are fixed and no longer matter of change



Tape space distributon to archive classes at DKRZ begin of 2007:

- part of the “work” space on tape because GFS too small
- “docu” domain consists of WDCC
- no expiration dates in “arch” domain, parts of “arch” domain belongs to “docu” but not yet documented

HLRE System architecture at DKRZ



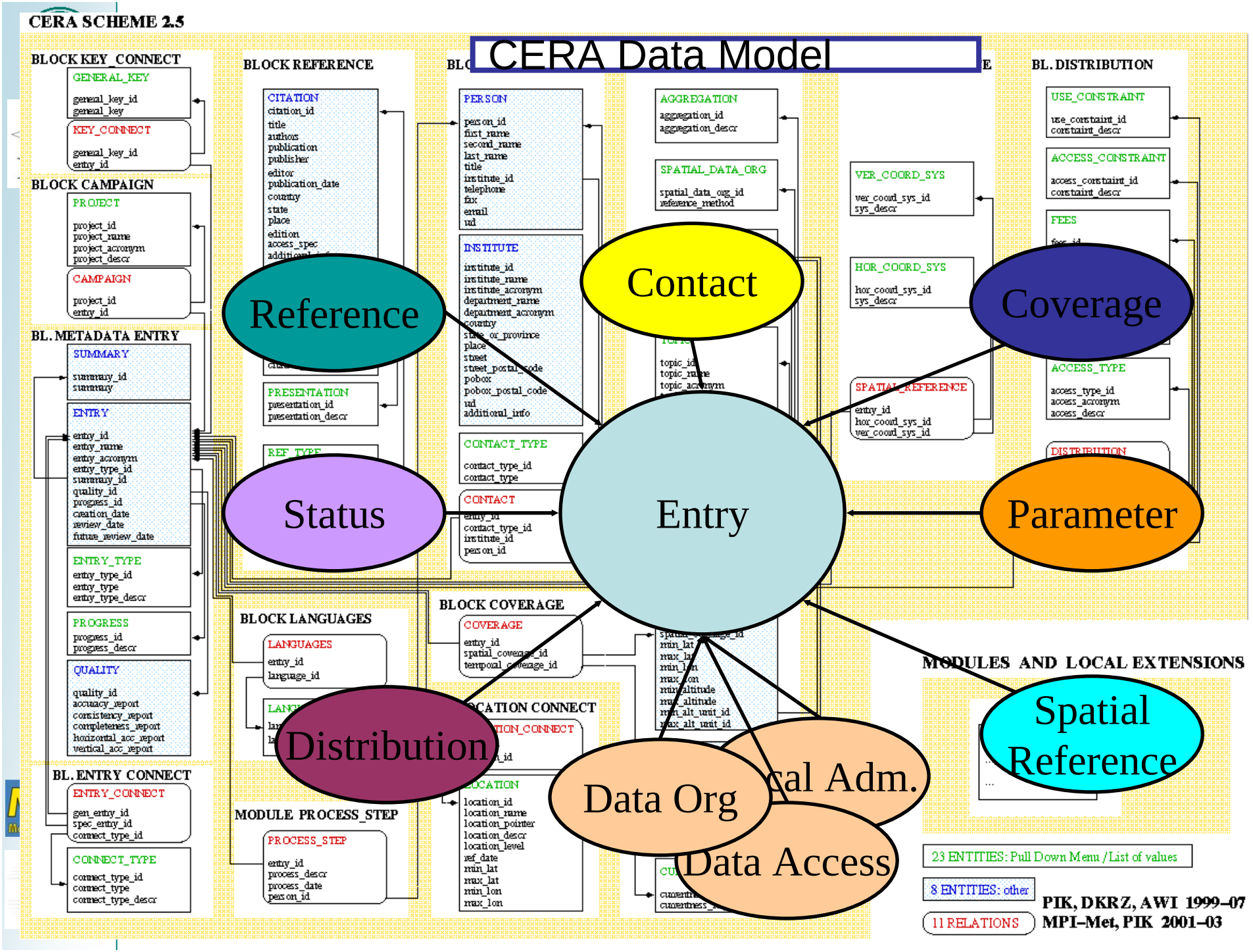
- Project based data storage strategy and resource assigment at DKRZ:
  - Separation of project data and long-term archive
  - Expiration date for project data
  - Aware, scientific decision to move data into the long-term archive
  - Data documentation requirements for long-term archive
  - Long-term data archive (“docu” hierarchy level) accomplishes the rules for good scientific practice

- Data documentation requirements are accomplished by using the WDCC infrastruture
  - CERA-2 metadata model developed in 1999
    - Catalogue interface: cera.wdc-climate.de
    - Input interface: input.wdc-climate.de
  - CERA-2 metadata content is complete with respect to browse, to discover and to use climate data which are stored in the database system or outside in flat files
  - The WDCC matches international description standards like ISO 19115, Dublin Core or GCMD and is integrated in international data federations
  - Data storage structure assembles storage of climate time series per variable in BLOB data tables. This allows for web-based data catalogue search and data access in small data granules.

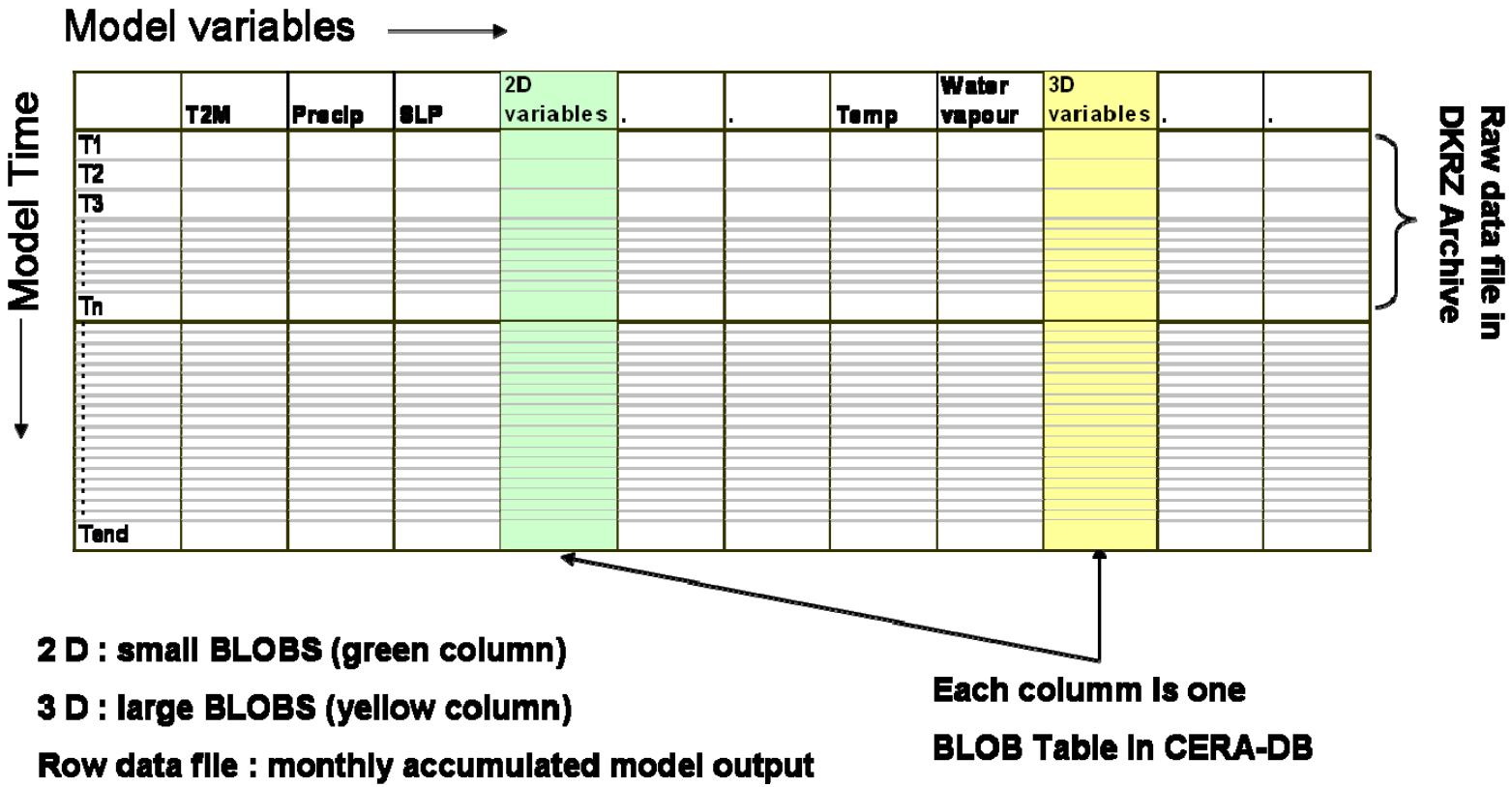
WDCC / CERA: General Statistics at 01-06-2008 00:00:10

- Database Size (TByte): 357
- Number of blobs: 6242417840 (6.2 billion)
- Number of experiments: 1081
- Number of datasets: 136606
- Total size divided by number of BLOBs gives the average size of data access granules: 60 kB/BLOB

CERA Data Model



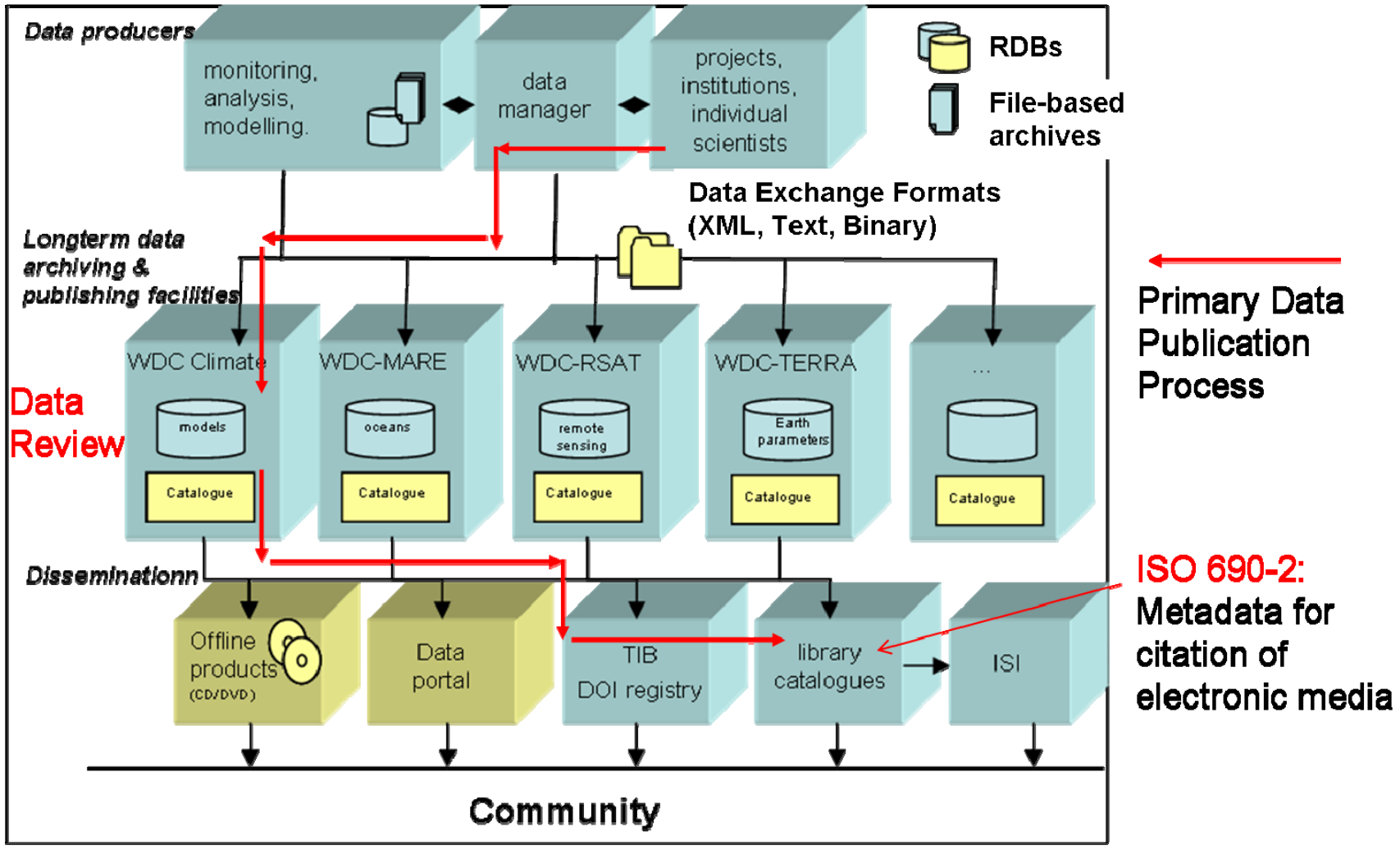
Data matrix of climate model experiment



Coloured columns correspond to BLOB data tables in WDCC.

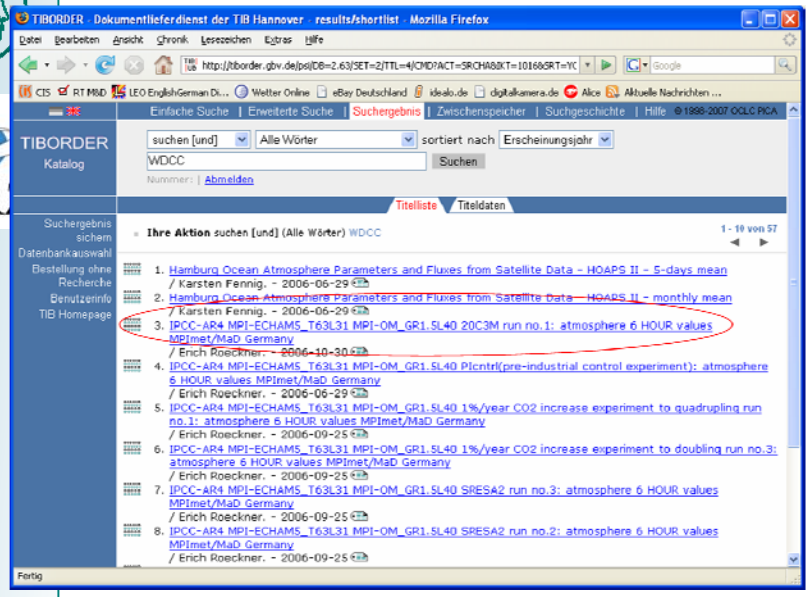
Collections of matrix rows represents storage in model raw data files (complete model output storage time step by storage time step).

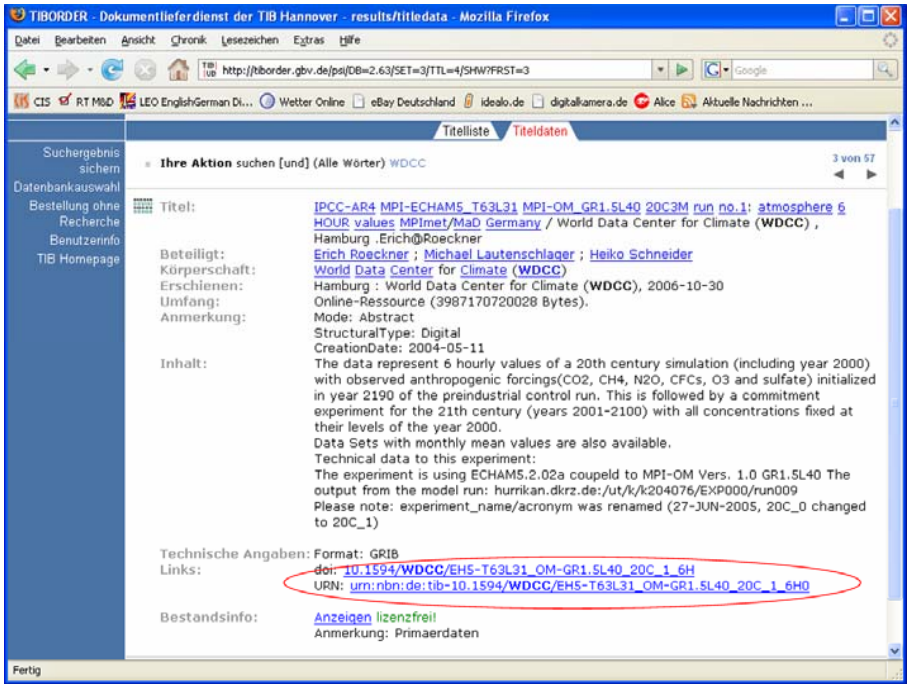
- Additionally WDCC offers the primary data publication service
  - Following the STD-DOI concept (Scientific and Technical Data – Digital Object Identifier, URL: www.std-doi.de)
  - Important aspects of the publication process are
    - The identification of independent data entities which are suitable for publication at the level of scientific literature,
    - The execution of an elaborated review process for metadata and climate data,
    - The assigment of additional metadata for electronic publication (ISO 690-2) and of persistent identifiers (DOI / URN) and
    - The integration of publication metadata and persistent identifiers into the TIB library catalogue (Technical Information Library, Hannover) so that primary data entities are searchable and citable together with scientific literature.
    - Quality characteristic is presently “approved by author”, future development should be “peer reviewed”.

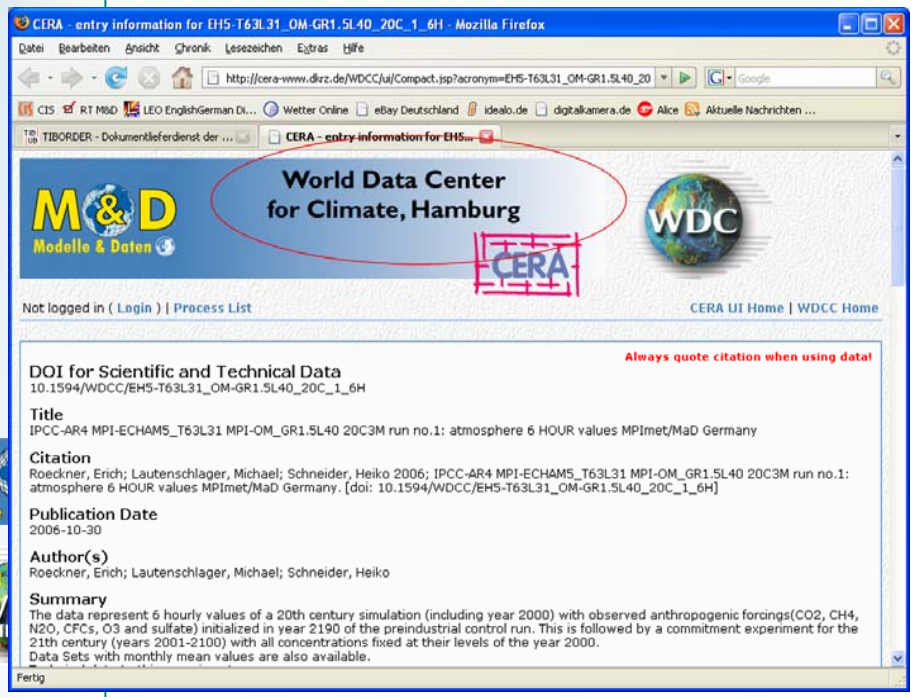


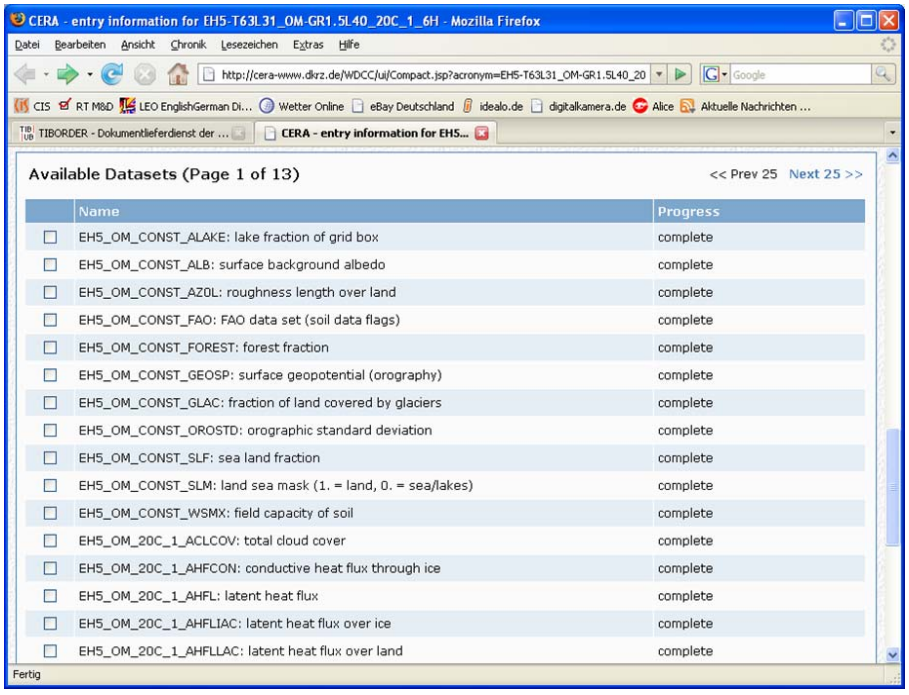
STD-DOI data publication workflow

URL: http://tiborder.gbv.de/psi/DB=2.63/LNG=DU/ , keyword: WDCC

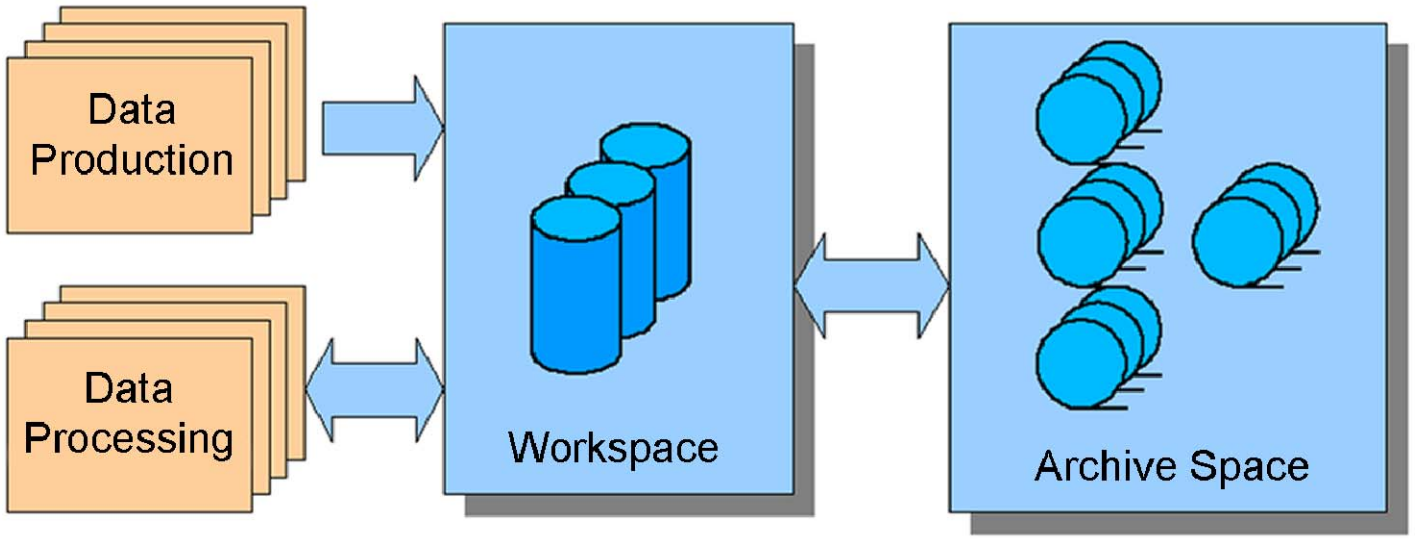








Data Flow at DKRZ



Different architectures for production and processing

Different tape media for performance and data security

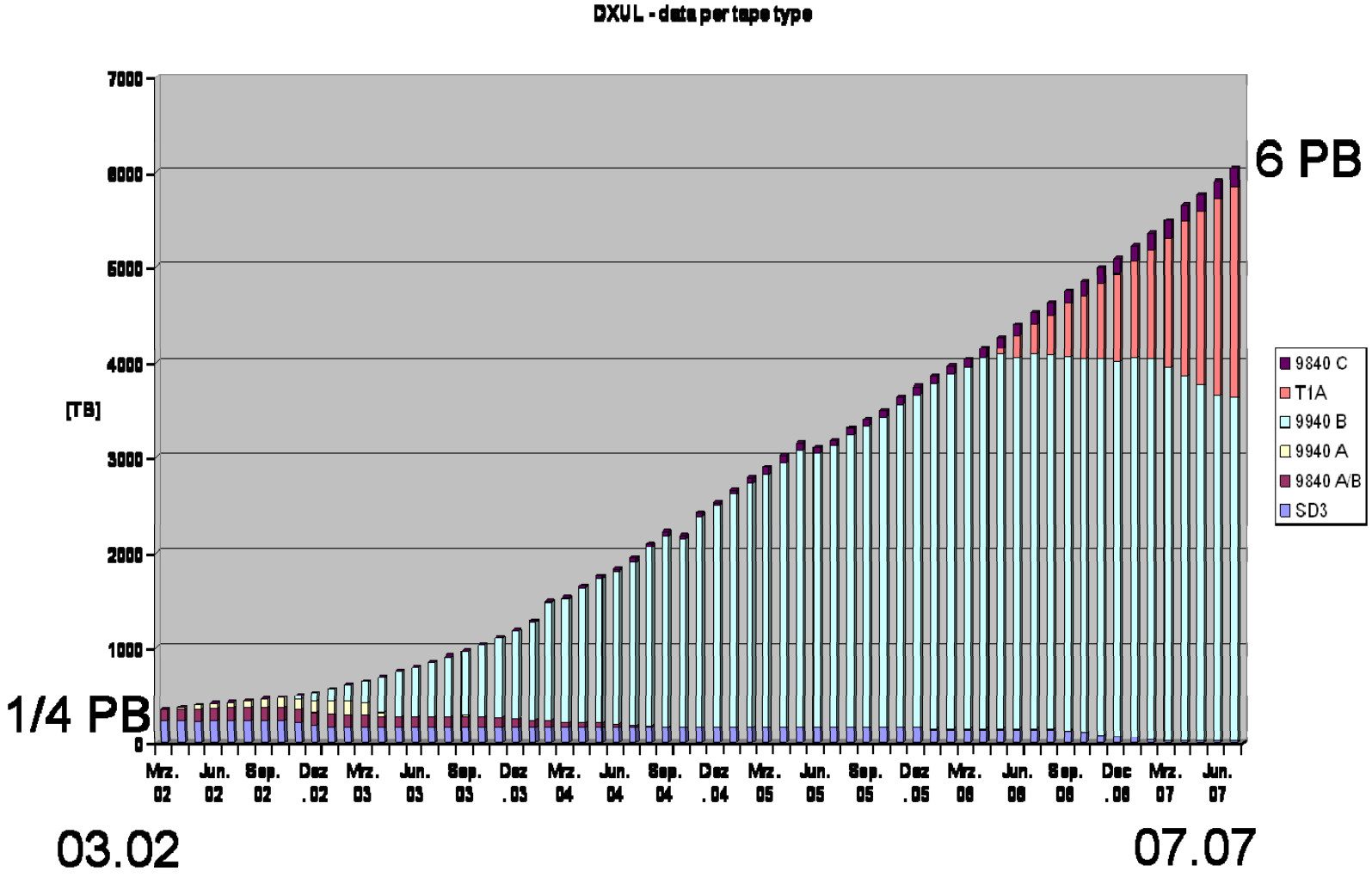
Data infrastructure integrates data stewardship in the long-term archive

- Bit-stream preservation
- Quality assurance
- Usability enabling

Long-term archive data stewardship

- Bit-stream preservation
  - Secondary tape copies on different tapes and technology at separate location
  - Copy to new tapes after maximum number of tape accesses are reached (Refreshment)
- Quality assurance
  - Semantic examinations: behavior of a numerical model compared to observations and to other models, part of the scientific evaluation process
  - Syntactic examinations: formal aspects of data archiving and ensurance that data archiving is free of errors as far as possible
    - Consitency between metadata and climate data
    - Completeness of climate data
    - Standard range of values
    - Spatial and temporal data arrangement

Media Distribution: Capacity



History: DKRZ archive increase and transition in tape technology

Long-term archive data stewardship (continued)

- Usability enabling
  - Complete and searchable documenation of climate data entities (database tables and flat files) in the catalogue system of the WDCC
  - WDCC offers web-based data access to small data granules (individual entries in BLOB DB tables)
  - Archive technology transfer must be downward compatible to keep old data technically readable
  - Data processing tools and data format access libraries must be migrated to new architectures

Summary

- DKRZ long-term data archive will still grow but slower than linear with the installed compute power
  - Target increase rates are
    10 PB/year for the tape archive,
    1 PB/year for the WDCC.

- Improvement of reliability of long-term archive because of more emphasis on data stewardship than on technical data service operations

- At the end the new data archive concept will result in a completely documented and searchable long-term data archive.
  (Data without documentation are only numbers)