Deutscher Wetterdienst
Wetter und Klima aus einer Hand

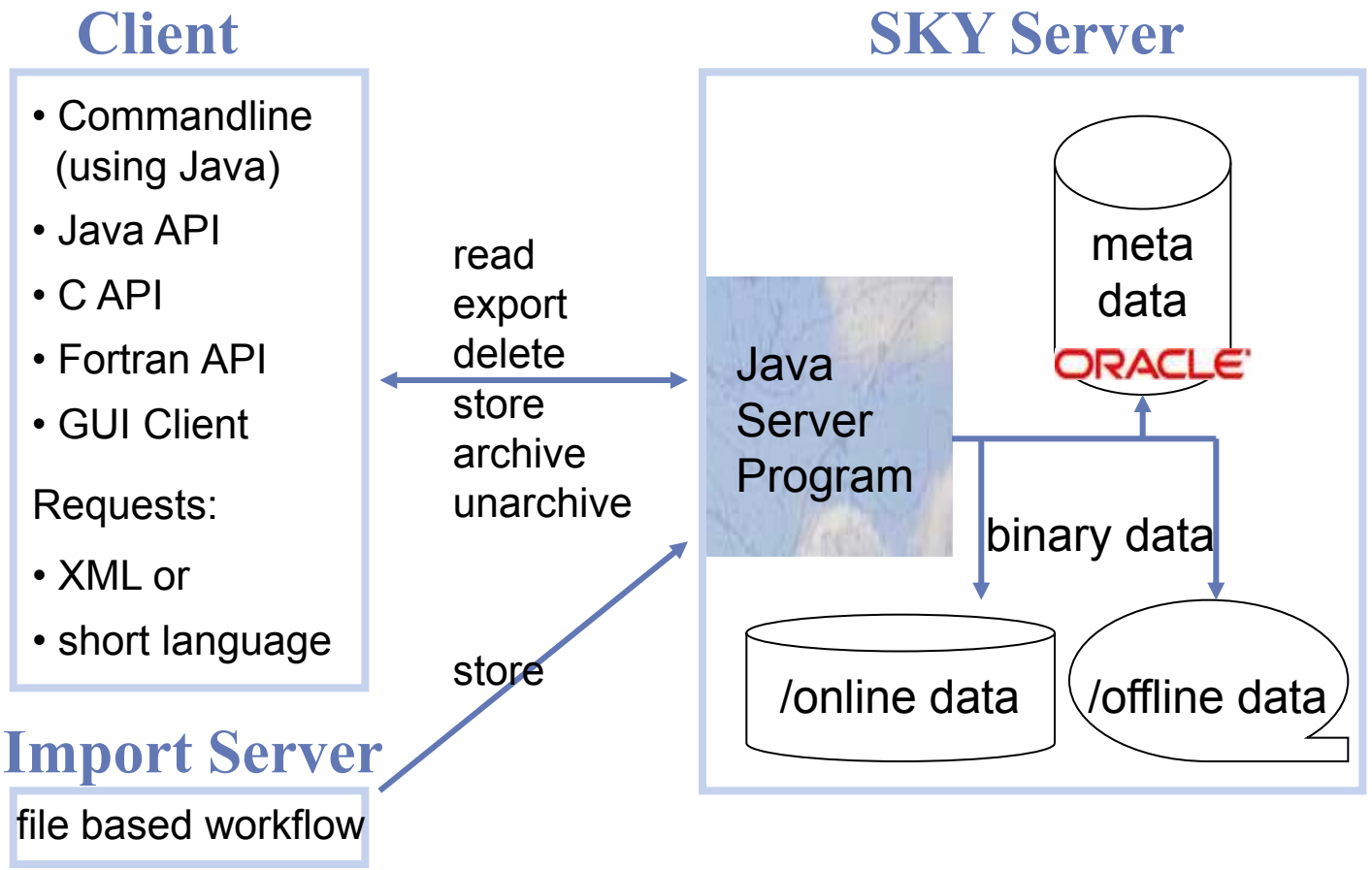
GRIB data handling with SKY at Deutscher Wetterdienst

H. Lemmin, Data Management Unit, Dept. Systems and Operations,
Business Area Technical Infrastructure and Operations, 2.11.2011

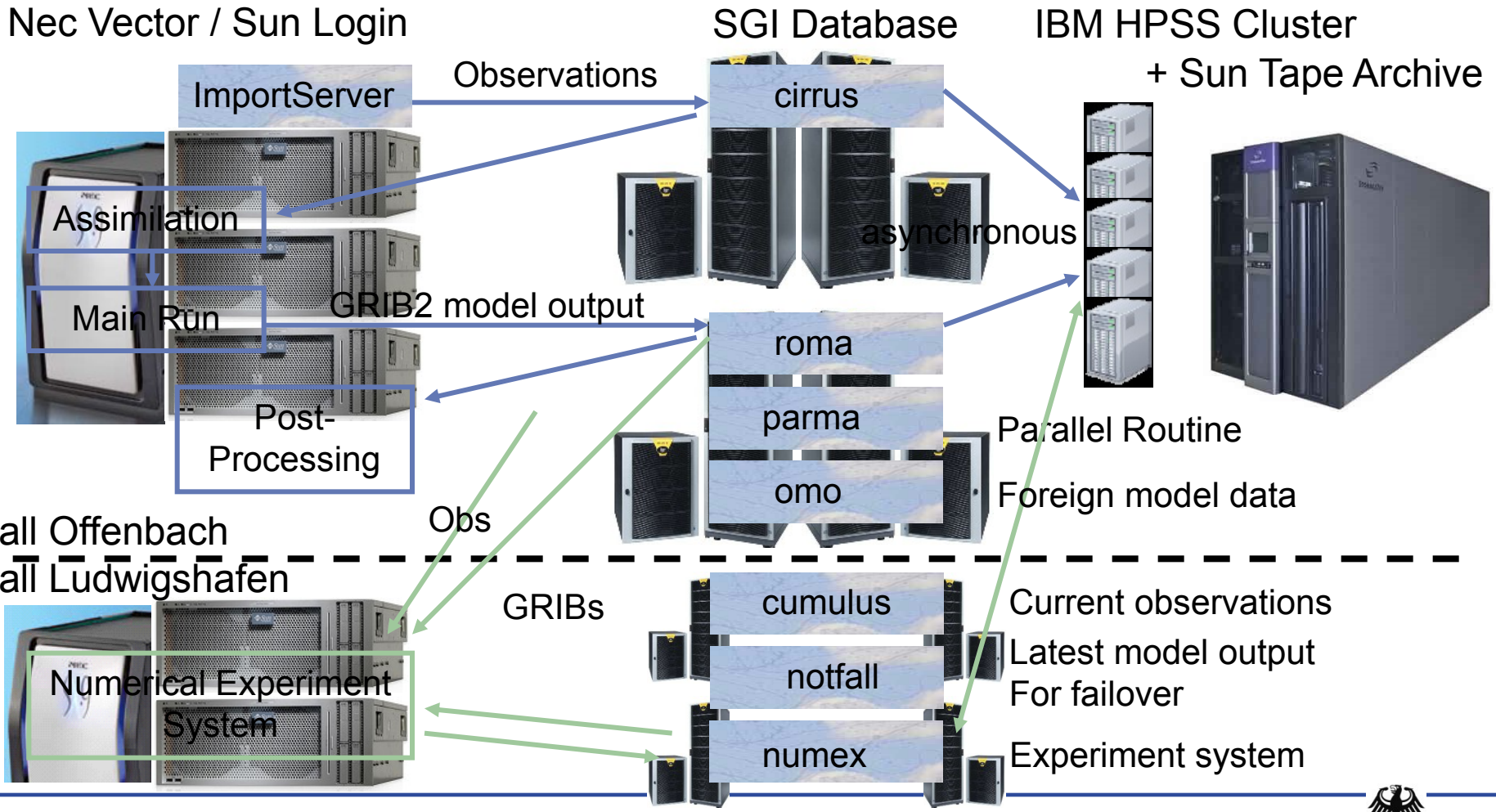
GRIB data handling with SKY at Deutscher Wetterdienst

- **SKY system components**
- **NWP data flow using SKY (Summer 2012)**
- Data migration CSOBANK -> SKY
- Migration GRIB1 to GRIB2
- Hierarchical storage management (HSS)
- Update of backup system

SKY system components



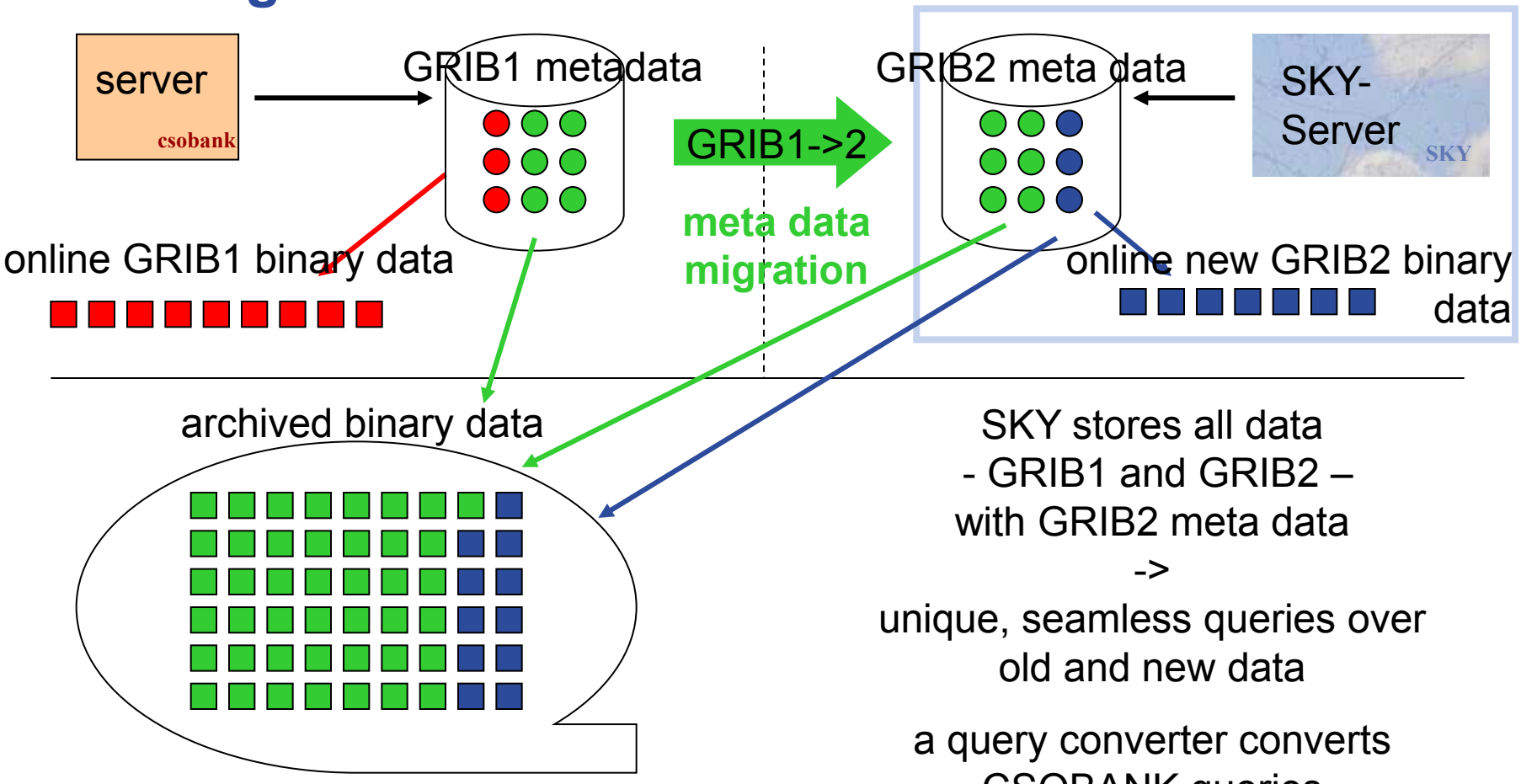
NWP data flow using SKY (Summer 2012)



GRIB data handling with SKY at Deutscher Wetterdienst

- SKY system components
- NWP data flow using SKY (Summer 2012)
- **Data migration CSOBANK -> SKY**
- **Migration GRIB1 to GRIB2**
- Hierarchical storage management (HSS)
- Update of backup system

Data migration CSOBANK -> SKY



SKY stores all data
- GRIB1 and GRIB2 –
with GRIB2 meta data
->
unique, seamless queries over old and new data

a query converter converts CSOBANK queries

Migration GRIB1 to GRIB2

- Deutscher Wetterdienst migrates its forecasting and postprocessing routines to ECMWF grib_api to ease migration
- As SKY is implemented in JAVA, grib_api is not used directly
- SKY uses grib_api configuration files
  - for migration
  - for parameter selection: element short names can be used
- additionally complex GRIB1 to GRIB2 conversion rules are implemented

GRIB data handling with SKY at Deutscher Wetterdienst

- SKY system components
- NWP data flow using SKY (Summer 2012)
- Data migration CSOBANK -> SKY
- Migration GRIB1 to GRIB2
- **Hierarchical storage management (HSS)**
- Update of backup system

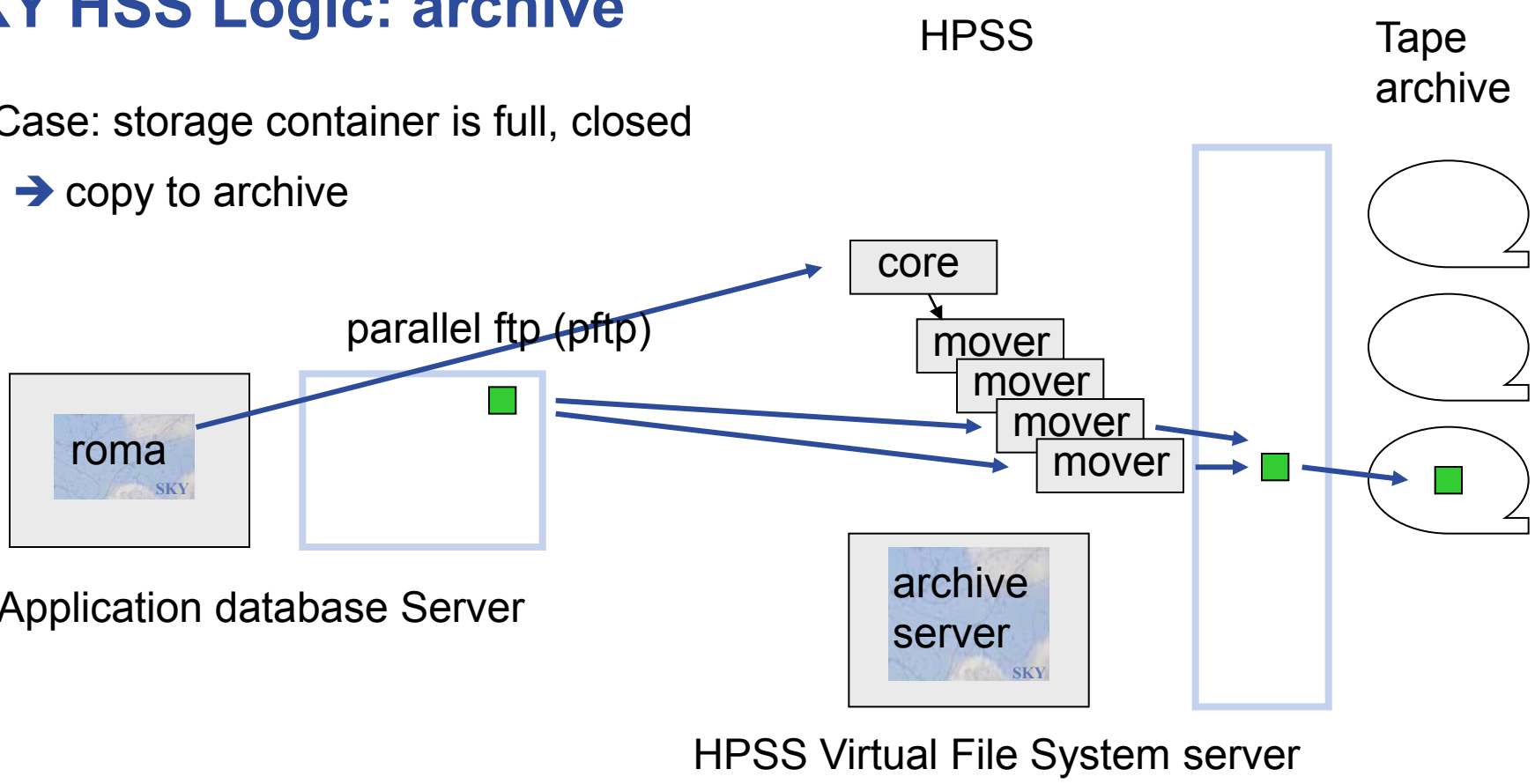
Hierarchical storage management (HSS)

- Standard methods of hierarchical storage management are not used so far at Deutscher Wetterdienst, because
  - the total amount of today db server online storage is 760 TB
  - no HPC vendor offered this functionality during invitation to tender so far
  - decision against combination of db and archive server

- SKY HSM functionality had to be developed to bridge between
  - Database servers with SKY server and online storages
  - Archive server front end to Tape robot archive system with a real HSM of 480 TB
    - HPSS with 4 mover and 3 virtual file system server

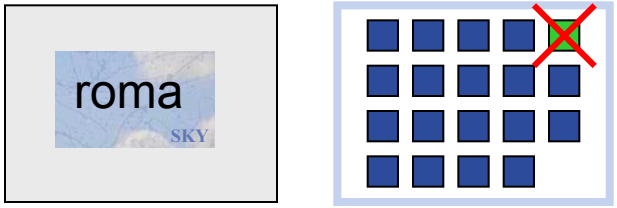
SKY HSS Logic: archive

- Case: storage container is full, closed
  - copy to archive

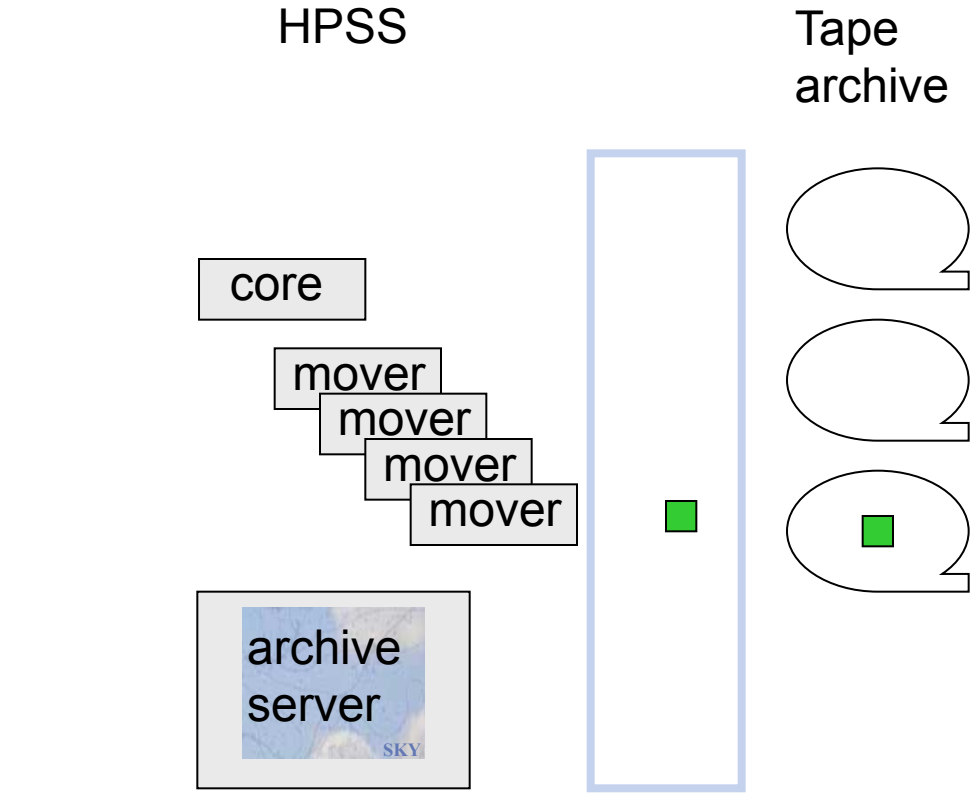


SKY HSS Logic: delete

- Case: file system is almost full
  - delete already archived and less used storage container (weight based approach)



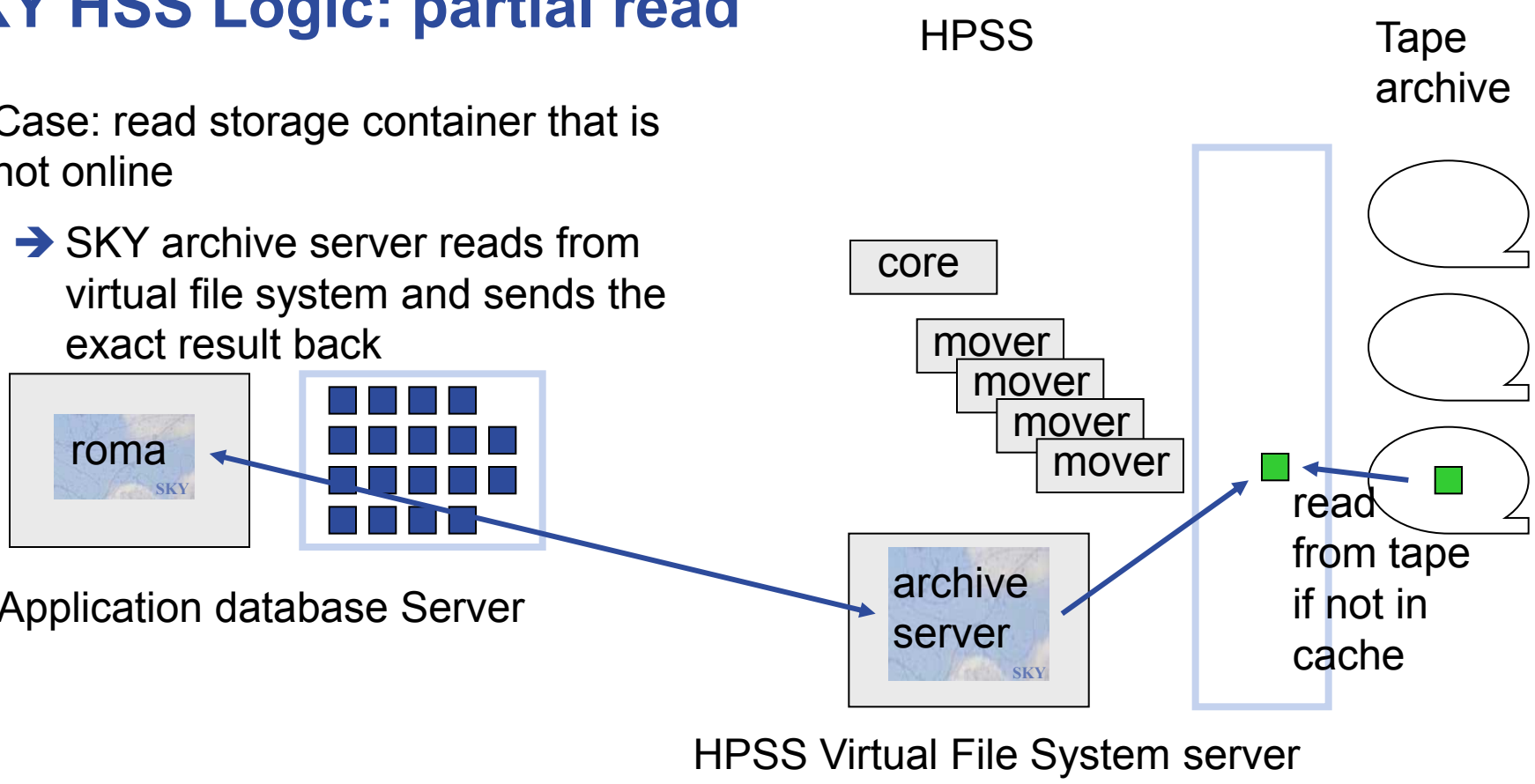
Application database Server



HPSS Virtual File System server

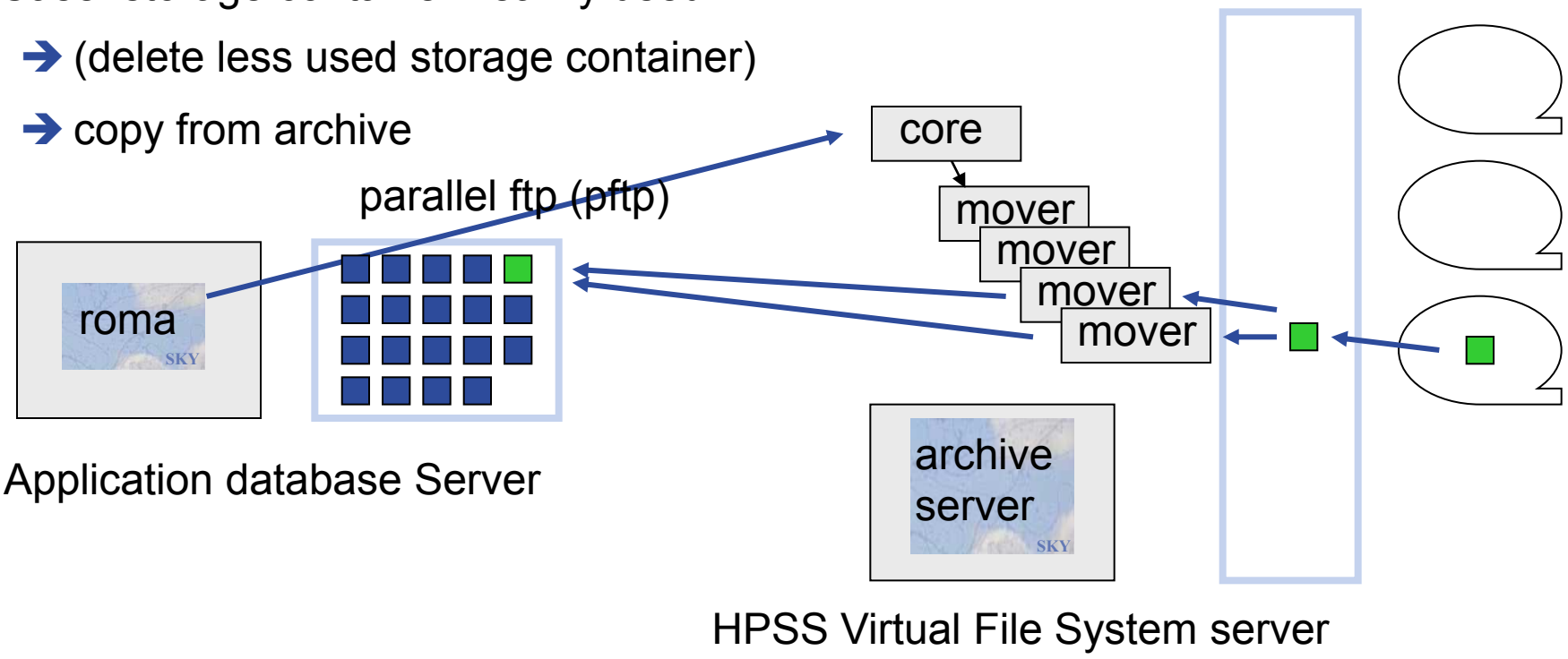
SKY HSS Logic: partial read

- Case: read storage container that is not online
  - SKY archive server reads from virtual file system and sends the exact result back



SKY HSS Logic: unarchive

- Case: storage container heavily used
  - (delete less used storage container)
  - copy from archive



GRIB data handling with SKY at Deutscher Wetterdienst

- SKY system components
- NWP data flow using SKY (Summer 2012)
- Data migration CSOBANK -> SKY
- Migration GRIB1 to GRIB2
- Hierarchical storage management (HSS)
- **Update of backup system**

Update of backup system

- backup system has to be updated constantly to minimize switch over time and keep up to the tight model schedule
- if a storage container (4 GB) is full or if the data set (ensemble member, forecast step) is closed:
  - copy meta data and binary data of those storage containers to the backup systems
- if the backup system is the primary system, the backup system feeds the main system

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