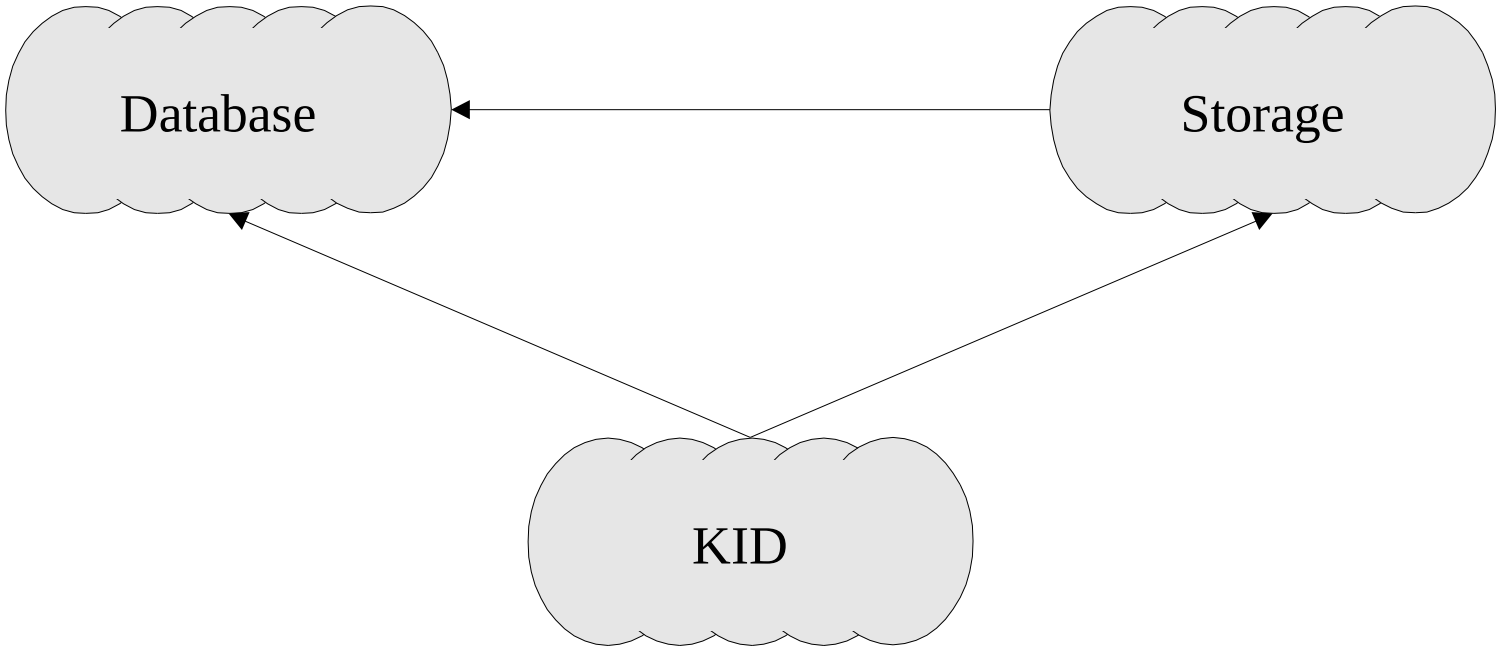
KLOE Data Handling



KLOE (online) database

- based on IBM DB2
- interfaced via SQLremote daemons/libraries
- 4 SQLremote daemons
  - SQLr_select
    - general SQL queries
  - SQLr_maps
    - cached maps access
  - SQLr_logger
    - data insertion
  - SQLr_update
    - superuser access

KLOE (online) database *(cont.)*

- fully functional
  - ✔ last SQLr_logger crash problems solved in December
- no major upgrades foreseen
- ✘ structural problem
  - ➡ since SQLr_logger keeps the connections open for long periods, upgrades are hard to deploy

SQLr_select

- libraries available for both C and FORTRAN
- tcl commands available (SQLremotetclsh and SQLremotewish)
- command line access to DB also possible via **dbonl**
- **Dbmaps** can be used for simple graphical browsing of the DB

SQLr_maps

- central daemon caches maps of recently used runs
  +
  the current maps
- a daemon reading from a file is available for stand−alone systems
- most of the DB traffic to the offline applications (in terms of data size) is handled by this daemon

DB proxies at section level

- needs a reliable computing node running a modified version of SQLr_select, SQLr_maps and SQLr_logger
- would reduce network traffic due to maps loading

Modified SQLremote daemons not yet implemented!

Anybody feels the need for such proxies?

Storage management

- archival of files produced by DAQ, datarec and MC is fully automated
  - l3spy still an exception
- archived files are automatically deleted when more disk space is needed
- recall on demand
- special disk areas reserved for the recalled data
- tape library accessed only if a copy of the file cannot be found on any disk
- data prefetching
  - currently, 1GB of data is moved to disk for each file not on disk

Storage management *(cont.)*

- interface to tape library via ADSM
- a local daemon, giving access to ADSM services, is running on each tape server
- disk space is also managed by daemons running on disk nodes
- KLOE database is used to coordinate all the processes
  - contains the list of all the instances of the interested files

User interface

- The recall daemon is the only user interface to the KLOE storage system
- given a list of files and a list of possible disk areas, it recalls one file at a time to a disk area of the list and returns the pointer to the file back to the user

Future developments

- The storage system has several problems both in terms of performances and reliability
- A radical rewriting of the archival and recall daemons is foreseen in the near future
- The user interface is expected **not** to change

KID

- flexible user interface to KLOE data
- uses URLs to select the data source
- allows access to data
  - on local/NFS disks
  - local circular buffers
  - remote disks/circular buffers
- data can also be accessed without knowing its current location
  - interface with the storage system

KID interface

- C libraries available
- an A_C input module makes it accessible to the offline and analysis software

KID remote access

- in order to access remote resources, a KID daemon must be run on every node that exports such resources
- all KLOE DAQ farm nodes and disk servers are configured to have one running