

CALICE software status

Sebastian Schmidt

19-JAN-06

Software releases (Roman)

- calice_lcioconv v02-00-pre3
 - Conversion of native Raw Data to LCIO
- calice_reco v02-00-pre3
 - Interface classes to access converted data
 - no real HCAL reconstruction software so far
- calice_online v01-03
 - Software tools needed to access the native raw data in the conversion job
- calice_cddata v01-02
 - Small routines to populate the CALICE database with conditions data entries
- calice software archives:
<http://www.listserv.cclrc.ac.uk/archives/calice-sw.html>

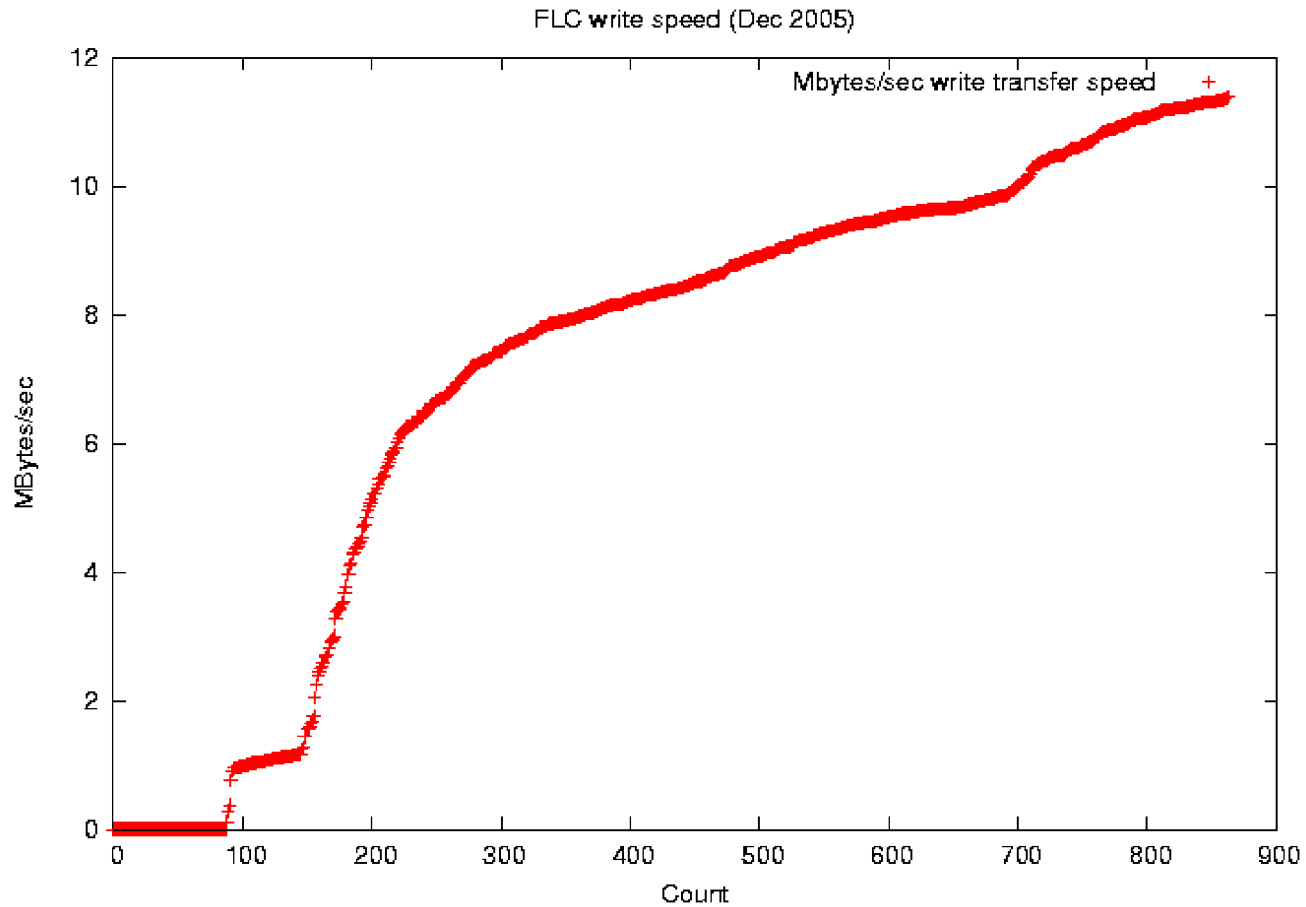
Experiences from Xmas Run

- Approx. 250GB of data taken
- Storing of binary files into dcache
 - Done automatically by perl script
 - Initial problems seem to be solved, needs still to be watched closely
- Conversion to Icio-files
 - Done automatically by perl script on different machine (ilc-log01)
 - Once new files are detected converter job is launched
 - One major bug found plus few smaller ones, but no new conversion necessary
 - Expect 'real' release soon
- Storage of binary and Icio-files in dcache pool
 - Analysts experienced some delays during usage of the Icio-files
 - Investigations performed

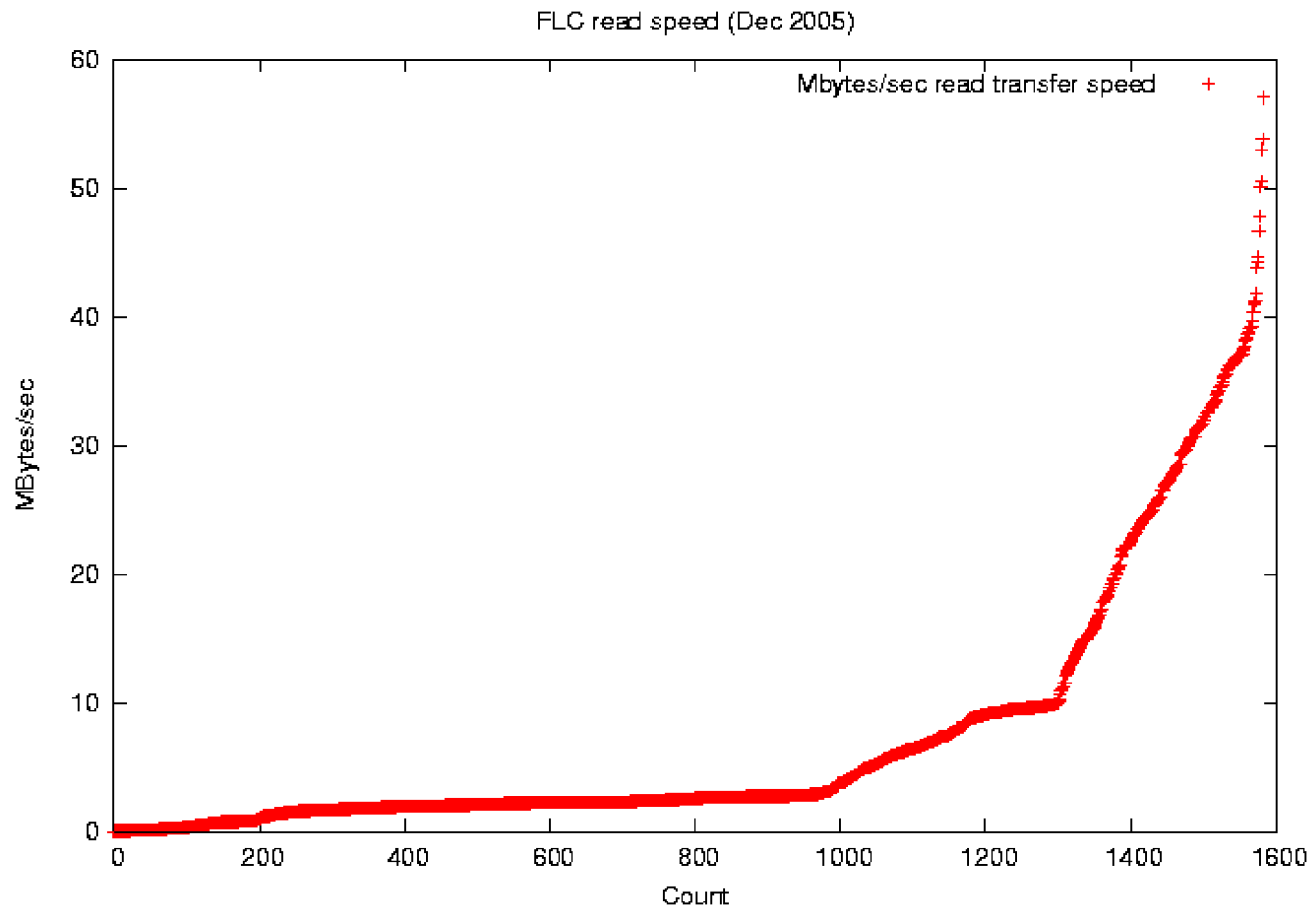
dcache performance issues

- dcache
 - Everything is stored on tape
 - To access a file the corresponding tape needs to be staged and copied by the robot
 - To speed up the access rate files which are used often are kept on disk (cache)
- Till December: FLC made parasitic use of other groups cache, files without access stayed in the cache for about four days
- Since early January: FLC got some (borrowed) dcache space (0.5TB), files up to 0.5TB stay on disk till other FLC files need the space, more or less smooth running
- IT recommends FLC to buy its own dcache server (4-8TB, 5K€), under consideration

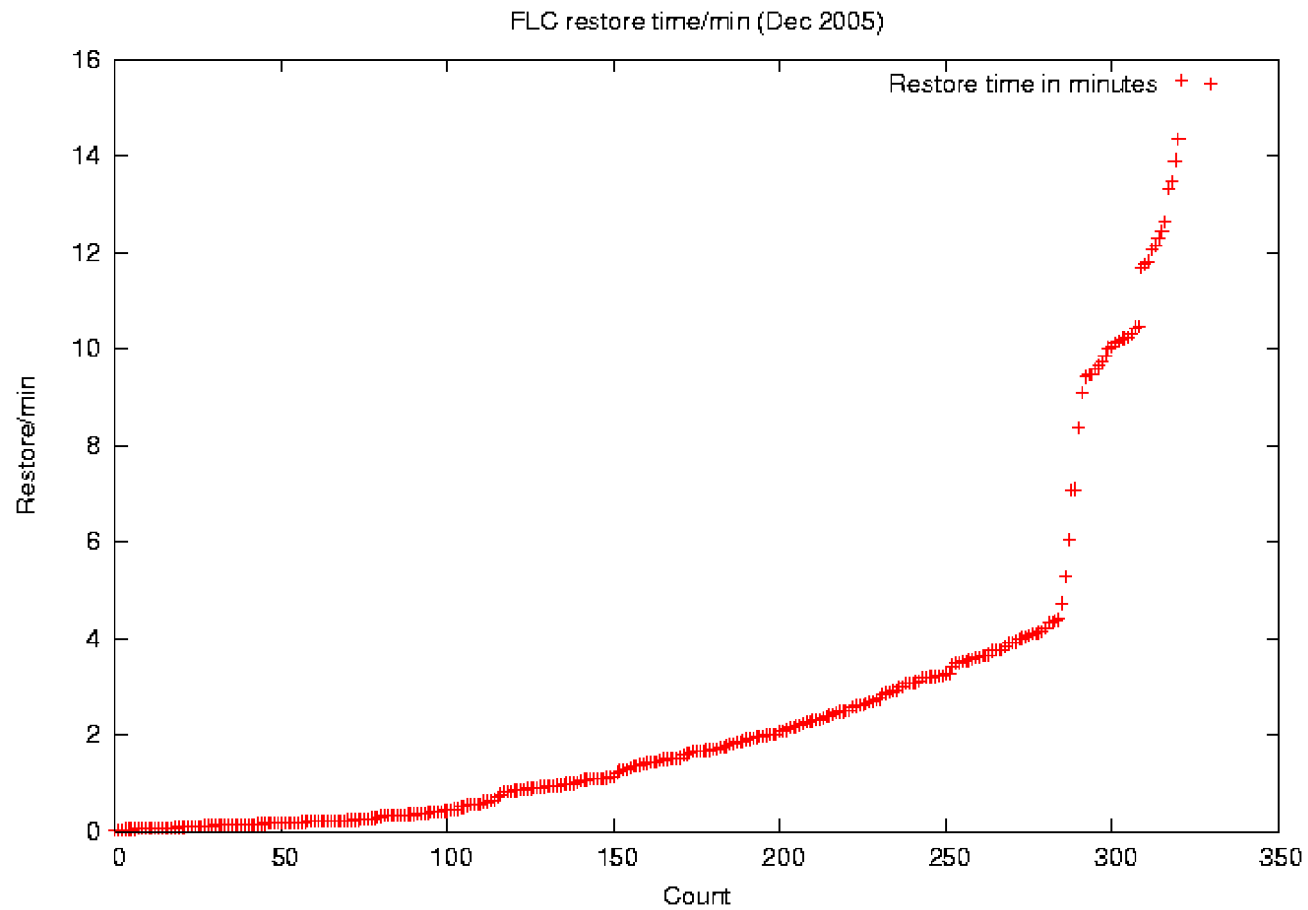
dcache write speed



dcache read speed



dcache restore time

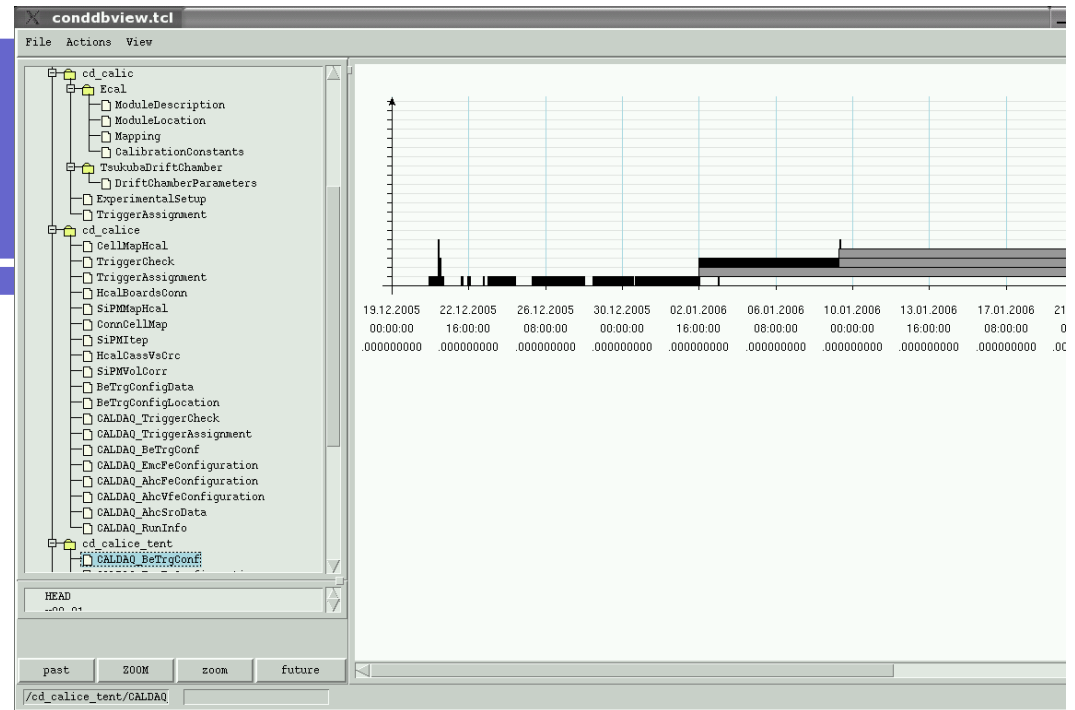


Database issues

- Added second database server
- flccaldb01.desy.de
 - Works as master
 - For writing to the database (reading also possible but not recommended)
 - Nightly database dump and backup
- flccaldb02.desy.de
 - new state-of-the-art server
 - Works as slave
 - Only for reading from the database
 - Faster server, thus programs which access database run faster when using this machine

Database GUI

- Fully functional
- For now available from [/flc/flcl01/data/sschmidt/lccd/v00-09 ...](#)



CERN testbeam data flow considerations

- Quick backup of raw data from DAQ machine necessary
- What network connection is available for DAQ machine?
- Access to CERN network (CASTOR)?
- Local conversion or conversion at DESY?
- Local conversion for 'online' analysis at CERN?

Reconstruction issues

- Tried to get an overview over existing code (not considering physics for now ...)
- Quite large amounts of data have to be treated
- Memory consumption
 - 8000 cells x 16 bit ADC converter (i.e. approx. 33000 bins) x 16 bit (minimum) = 500MB
 - Recommend at least 1 GB extra for flcl01 (approx 150€/GB)
- Time consumption
 - 8000 cells x 2s gain fit = 4.5h
- Calibration conditions data
 - One collection of conditions data has to be written as a whole at once
 - Calibration jobs should be usable module-wise (less memory and time consuming, maybe not all modules in use at a certain time)
 - Thus: save calibration data into database in conditions data collections separately for each module (not so elegant, but that's life!)
 - Don't write conditions data extracted from a certain time period into the database till infinity!! (Make them valid only till the end of the time period)

Reconstruction issues

- **Success rates of calibration jobs**
 - Important to have very high success rates of fits during calibration
 - 99% success rate means: Somebody has to look after 80 fits every n hours
 - Clear criteria must exist e.g. if a fit converged correctly in a valid region of phase space and the calibration can be written to the database
 - It sounds easy to allow only “reasonable values” to be written, but it’s not easy to define the criteria
- **Visualization of large amounts of calibration data**
 - How does one display the time dependence of 8000 gains/pedestals?
 - Stamp plots (approx. 50 pages)
 - Visual comparison with example plots (Kolmogorov ?)
 - Define mathematical/logical criteria for status ok/not ok
 - Integrate status over samples which depend on the same hardware to get a hint for the problem source as fast as possible
- **Database accesses**
 - Some conditions data exist in 15s intervals
 - Accesses limit program performance in the moment
 - Got better with new database server, but still ...
 - Issue is under study

Conclusions

- New software releases
- Conversion jobs runs smoothly
- dcache, database performance improved
- First considerations for storage/data transfer at CERN
- Calibration, reconstruction: Consider the size of the project before programming!