

# Proposal for Alignment and Calibration procedure

## Summary of the DQ procedures [↗](#)

Data are permanently monitored at the pit by monitoring processes running in the Monitoring Farm. This includes subsystem-specific monitoring as well as full reconstruction of a small sample. Produced histograms are analysed automatically as well as scrutinised by one of the pit shifters as explained in this page [↗](#).

Every day a dedicated DQ meeting decides on the quality of the data taken since the previous meeting and gives eventually the green light for offline processing using the same CondDB tag as was used in the Monitoring farm. The green light is given by setting the DQ flag of the fills (or runs) to "Good" in the Bookkeeping. DIRAC processing is on hold until this flag is set.

This proposal is an attempt at describing a possible procedure in case the red light is set.

## Getting new alignment/calibration constants

Hereafter "calibration" is used generically for calibration and/or alignment.

When the result of the Online Monitoring indicates that there is a need for re-calibrating one or more sub-detectors, its(their) representative(s) at the DQ meeting trigger the process of re-calibration. This activity has to be co-ordinated by a "Calibration coordinator" as it may involve complex and/or sequential processes.

The Computing Project is setting up a dedicated cluster (LHCb-CAF, for Calibration and Alignment Facility) at the computing center for providing the necessary resources. The CAF consists of a set of dedicated batch nodes, accessible through the lhcbcaf queue by a restricted number of users. This queue allows to submit jobs that can then create an interactive session on the WN. AFS scratch disk space is also foreseen for copying temporarily the necessary files (space handled by the calibration team). To start with we have requested 2 WNs with 8 cores each, but this can be easily extended if needed.

The result of the re-calibration is a set of calibration constants that are provided as SQLite small CondDB file(s).

After re-calibration it is necessary to assess the quality of this calibration before committing and tagging the constants in the CondDB. It is thus necessary to reconstruct and monitor results obtained with this calibration on a representative sample of data covering the whole data taking period that has not been processed yet offline. This cannot happen on the Online Monitoring farm and must thus be done and followed up offline.

## Certifying a new calibration

We propose that a small rate of raw data be streamed on a dedicated data stream (Calibration stream/files) during data taking. It may be either a random sampling of events or the result of a specific selection in order to enrich the sample in representative events. This stream would then be reconstructed using the candidate calibration, produce monitoring histograms whose analysis would allow to go to the next step.

Note that this applies as well to problems encountered with the application (Brunel) rather than the calibration (e.g. too frequent crashes, of small bug fixes). New software implying better (incompatible) performance of the reconstruction should be treated differently as it would possibly imply a re-processing of previous data.

The Calibration files are transferred from the pit onto Castor, registered in the LFC and the Bookkeeping as such (specific namespace and BK configuration). Taking as an example a rate of 5 Hz with 40 kb events, this stream represents 8 GB per day (assuming a 50% duty cycle of LHC).

Open questions:

- How is the selection made? According to online streaming possibilities, this should be based on the decision of a dedicated HLT alley (which can be very simple and just sample events either using another alley or all events).
- When is the streaming done? A dedicated data writer can be instantiated that takes events from the main buffer managers, or a dedicated process can read main stream files and produce the selected stream. In any case the files have to be registered in the Run Database in order to make it to the offline bookkeeping.
- What should be the size of these files? With the small amount of data, one could imagine one file per run. However this would not easily allow a fast processing when verifying a calibration. Hence probably the preference would be for smaller files (compatible with 1 to 2 hours processing time) in order to process them in parallel.
- Which Castor pool? The regular lhcbraw Castor pool is not adequate for small files, hence the proposal would be for using a non-migrated pool with permanent disk storage. Files could possibly be archived when necessary by creating tarballs e.g. per run or fill.
- Should the calibration stream files be systematically reconstructed? This would allow a comparison when being re-reconstructed after calibration fixes. Should DSTs be saved and kept?
- Who manages the verification jobs? Appropriate functionality has to be made available in the Computing infrastructure (DIRAC) in order to launch these jobs in a controlled manner. An alternative could be to use the CAF for running these jobs, but here again tools have to be developed for launching the processing in a semi-automated way.
- For the above reconstruction, should we ask for a dedicated batch CAF queue accessible through a CE?

The results of the calibration sample processing is then analysed by the offline DQ shifter who reports on the result to the Calibration coordinator. If OK, the new calibration is entered into the CondDB, tagged (who does this? Calibration coordinator?) and the tag is transmitted to the Computing shifter who creates a new reconstruction production using this tag. The former production is stopped. This tag is also set for usage on the monitoring farm and (if necessary) on the HLT farm.

Fills/runs are then marked as "Good" in the BK and the processing can start with the newly created production. Note that the calibration validation jobs have to be launched on all runs that have been taken until the new tag is made available on the Online Monitoring farm, not only the runs that triggered the re-calibration. Other changes in the calibration may imply a further loop for more recent runs, as this process will certainly take several days.

## Offline DQ

In addition to the above, the DQ offline shifter is in charge of checking the histograms produced by the offline reconstruction production. (S)he may spot problems with calibration or applications that have not been seen in the Online Monitoring. In this case a red light may be set to the current production. It should then be discussed whether a "bad quality" flag has to be set on the faulty already-processed data samples (RDSTs). In this case the raw files concerned have to be set as "to be processed" for the fixed production.

The above proposal implies a very good collaboration of a set of people and projects:

- Online shifters
- Offline shifters (DQ and Computing)
- Physics software project
- Core software
- Core Computing (DIRAC)

-- PhilippeCharpentier - 27 Aug 2008

This topic: LHCb > Calibration

Topic revision: r3 - 2008-10-02 - unknown



Copyright &© 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback