Calibration

Yves Schutz

Tuesday, June 27, 2006

Preamble

Elaboration and implementation of the calibration strategy has started since about a year within the Offline project together with the detector experts.
In a first step the calibration algorithms have been included in the data reconstruction algorithms of each detector, the data structure of the relevant calibration parameters has been defined and implemented as a Data Base (DB) called the Offline Calibration Data Base (OCDB) and the mechanism to write into and query the OCDB has been implemented. It must be emphasized first that the OCDB does not hold the data itself but instead the unique file identification of the data files, stored permanently in the MSS (MSS), together with validity time stamps; secondly the file format is exclusively the ROOT format. This first step is completed.
The second step was launched at the same time by the Offline project. It consists in the definition of the strategies to generate the relevant calibration parameters either online or offline, the strategy to collect these parameters and populate the OCDB and the implementation of the various algorithms within the appropriate environment. To the purpose of collecting data from the online environment into the OCDB, the Offline project has created the SHUTTLE program which is made available to the detector groups as a prototype. Detector experts are still in the process of defining their strategies to generate calibration parameters. Concurrently, several discussions have been held about the implementation of the calibration between the DAQ, DCS, ECS and the detector experts.
As coordinator for the calibration issues within the Commissioning Task Force, I am helping in defining the strategies together with the experts of the detector projects, and the DAQ, DCS, ECS, HLT and Offline experts. My strategy to approach this issue was to select a few detectors and try to elaborate a limited number of use cases which would cover all the requirements.
In the present document, I first describe the requirements of the selected detector (SPD, because its strategy is simple and is already implemented, PHOS because it is a particular case that will be tested during beam tests in July, TPC because they have started the commissioning and should test their calibration strategy and TRD because they have reported a rather well understood strategy) together with a few open questions.
I then discuss the use cases and the tasks required by the various projects to implement these use cases.
The document has been discussed by the detector experts, and the DAQ, DCS, ECS, HLT, Offline experts.

It is tentatively planned to organize a one day workshop as soon as possible (to be discussed at the ALICE week in Bologna) to discuss the present document and elaborate an agenda. First feedback should be discussed during the Offline week in October.

Detector requirements

SPD

A meeting took place on March 24th and I had private discussions with Ivan Amos Cali. Two modes are implemented to create and collect the calibration parameters required for the offline analysis (thresholds, noise levels and list of dead and noisy pixels). In both modes data taking is done in special runs and is triggered by a trigger internal to the detector. Raw data are not kept in permanent data storage. This could be a problem for offline checks. How to send the raw data to MSS is an open question!

- In the first mode raw data are sent to a DCS worker node, which runs analysis tools. The analysis tools produce several ROOT files with the required data. The ROOT files are shipped by the analysis tools to a file exchange server provided by the DCS group. The SHUTTLE collects these ROOT files and stores them in the CDB. This mode can also be used on a partition of the detector without interrupting the data taking of the rest of the detector.

- In the second mode the raw data follow the DAQ flow without event building (but do not go to MSS) and a series of analysis scripts (derived by the DCS analysis tool) are running in LDCs. The resulting ROOT files are saved temporarily in a second file exchange server. The SHUTTLE collects them. The results are also recorded in the DCS Configuration DB through the DAQ-DCS file exchange server.

Both modes are now being implemented and will be exercised before the June offline week.
Some of the calibration results are important for knowing the status of the detector. Although they will not be used offline for the reconstruction they will be stored in the OCDB.

PHOS

Two meetings took place on April 21st and May 24th.

- The share of bandwidth allocated to PHOS (40M Bytes/s as stated in the TRG, DAQ, HLT TDR) is not sufficient to accommodate the whole PHOS raw data, even for one module only: 2 channels per crystal, 128 samples per channel, 10 bits per sample, i.e. 600kBytes/event with zero suppression or 1.2GBytes/s dor a trigger rated of 2K Hz! Only a sampled set of raw data will be kept in the MSS. The consequence of this new requirement would be that in Modes A or B (data unchanged by the HLT) the PHOS would run at a maximum rate of a few 10s of Hz.

- Raw data will be pedestal suppressed (mean minus five sigma from mean). Pedestal's mean and sigma are calculated in dedicated run and are not kept for offline use.
- Data will be compressed in the HLT by performing a pulse shape analysis that leaves only the pulse amplitude (energy) and the pulse offset (time of flight) sent back to the DAQ. The HLT will not run AliRoot, however the modules running in HLT can also run within the AliRoot framework. There are still three candidates for the fitting algorithm. One fitting routine requiring the knowledge of the pulse shape has been implemented in Fortran and satisfactorily tested on simulated data. A peak finder method has been implemented in C++ and tested on simulated data. Both are being implemented in AliRoot. The final algorithm will be selected after it has been evaluated with test beam data.
- If the fitting algorithm is used offline a set of calibration parameters are required: APD voltage, shape of the pulse (rise time and order of the Gamma function) set by the FEE shaper, pedestals mean and RMS, APD gain factor, mean amplitude of the LED. It is not clear where to get the APD voltage from. The other parameters are calculated online in the HLT or quasi-online in the CAF and collected by the SHUTTLE.
- The dead channels maps of APD and FEE are constructed from pedestal run data and from LED monitoring. It will be updated every run. The maps will be filled online by the LDCs or the HLT nodes, stored temporarily on a DAQ or HLT file exchange server and collected by the SHUTTLE.
- During the first run, raw data will be kept and the level of online processing will be gradually increased.
- The online monitoring will be done in HLT as well, similarly to what is done for the TPC. The display being based on ROOT and Qt.
- The HLT will communicate to three different DB, the DCS Configuration and Condition DBs and the Offline Calibration DB. The possibility for the HLT to write/read into the DCS configuration DB needs too be determined; it would be done only via detector software. The HLT must pass data to the OCDB via the HLT-Offline exchange server.
- The online calibration will be done in the HLT as well (cell-to-cell and absolute calibration) and the resulting parameters stored in the DCS Archive DB where the SHUTTLE will collect the parameters to store them in the OCDB. This possibility does not exist. The HLT-Offline file exchange server must be used instead.
- Enough disk space is provided for transient storage of the histogrammed parameters by the HLT distributed file storage. Users will have access to these histograms for monitoring purposes.

Several issues remain to be clarified or done.

- Select and implement the fitting algorithm in the AliRoot framework. What will be the data sent back to DAQ: 1 or 2 amplitudes/times? Amplitudes will be given in ADC units of already energies?
- The creation of the calibration parameters is understood and some code exists but nothing is implemented yet in the final framework.

- It is not clear if DCS and SHUTTLE are aware of the parameters that will be stored in the CDB.
- The CPU load imposed by the various processes (fitting, calibration, monitoring) needs to be evaluated.
- How will HLT collect the initial calibration values stored in the OCDB?
- The list of parameters, such as APD voltages, monitored by DCS, needed by offline calibration and to be collected by the SHUTTLE must be defined.

TRD

The TRD experts have provided a detailed list of the calibration parameters together with information on their source, the way to create the parameters etc.

- Survey data for super module and chamber position data will be stored in the offline Calibration DB as AliAlignObject.
- Several parameters with a year or run update frequency will be calculated from data collected in physics runs and require large statistics 10^8 (10^6) and 10^5 (10^3) pp (HI) events respectively. The procedures to calculate the parameters will be run on the HLT farm online to provide a first-pass estimate and iterated later on if necessary. The procedure itself provides input to the local CDB, which then can be fed into the global one. (This should be done with the SHUTTLE, right?). The analysis procedure are implemented and run also in the AliRoot framework.
- Several parameters related to the FEE are stored in the DCS Configuration DB. The scenario currently envisioned is the following:
    1. Configuration consists of four parts and is characterized by four tags:
        - basis configuration (clock frequency, number of time bins, filter settings...) common to all chambers
        - lists of dead MCMs, pads...
        - command coder version
        - gain tables
    2. Configuration is stored in the DCS configuration DB and in the OCDB at the same time
    3. the configuration is loaded into the electronics from the DCS configuration DB
    4. during run the DCS boards report their tags to the DCS archive DB
    5. during run the DCS boards send messages to DCS archive DB whenever they found a misconfigured MCM and configure it again (bit flip)
    6. at the end of run the SHUTTLE fetches the 4 tags and the bit flip messages and stores it in the OCDB
    7. during offline analysis one reads the electronics configuration tags from the offline DB and loads the respective electronics configuration datasets
- External parameters (about drift velocity, voltages and currents, atmospheric pressure, luminosity and magnetic field) will be monitored every minute by DCS and stored in the DCS Archive DB. The SHUTTLE will collect these parameters and make them available to offline through the Calibration DB.

- Most but not all of these parameters are used by simulation and reconstruction and accessed through the CDB framework.

Apart of the questions already raised in the above items, there are the following additional questions:

- What is the local HLT CDB? Is the SHUTTLE aware that it has to collect data from the HLT?
- How do you collect parameters available in the DCS Config DB to make them available offline through the CDB. This is not possible with the shuttle.
- Not all calibration parameters are used at simulation and reconstruction time? Will this be implemented later? When?

MUON

A document has been prepared and updated in March 2006 by the MUON experts. It can be found at http://aliceinfo.cern.ch/static/Offline/dimuon/CalibrationAlignmentforMUON/. I had phone discussion with the person in charge of calibration.

The parameters required by offline have been identified together with the related, although incomplete information.

- Pedestal parameters will be calculated online in the LDCs with pedestal run data during dedicated standalone runs. Results are so far stored in ASCII files. The parameters will be pre-pended to raw data event header. Other options are considered such as storing the parameters in online files to be collected by the SHUTTLE. The calculation procedure is ready, but outputs need still to be converted in the appropriate FEE format.
- The electronic calibration will be calculated from data acquired during dedicated standalone runs. The gain correction factors for tracking are calculated from pulser data quasi-offline in the CAF or online in the HLT, depending on the CPU resources required. They will be stored in the OCDB using the framework implemented in AliRoot or fetched by the SHUTTLE on HLT. They will be updated once per day. The procedure is not yet implemented.
- There are FEE parameters that might or might not be needed offline. It is planned to store them in the Construction Data Base. The format is XML. These data are needed for the calculation of the gain correction factors. Can the Construction DB be accessed by online? These parameters are not needed offline.
- Dead channels map are calculated online from pedestal run data for MUON tracking and from calibration run data and physics run monitoring for MUON trigger. The procedure is understood but not implemented. Results will be stored in online files to be collected by the SHUTTLE and inserted in the offline CDB.
- The MUON trigger requires masks applied to dead channels. Status is the same as the in the previous item. Although it duplicates the information, both are required for convenience.
- Look-up-tables (only for MUON trigger) with momentum thresholds are calculated from Monte-Carlo data of physics run data. They are stored as ROOT

files or flat files to load into the trigger electronics of MUON tracking. They have to be archived somewhere, best place would be in the event header.

- Work on the various procedures and strategies are progressing: the status will be reviewed during the October Offline week.

TPC

Several discussions took place by e-mail, in TPC meetings and in face to face meetings.

- Pedestal parameters (mean and sigma) are computed on line either in the DAQ or HLT farm from dedicated runs data. Given that the pedestals will be loaded to the FEE through the DDL, it seems more logical to perform the pedestal computation in the DAQ LDCs. The operation is repeated every run. The results are temporarily stored in a TPC local DB and are fed to the FEE through the DDL. There is no need to make these data available to offline. Eventually the pedestal files could be archived in the mass storage and the pointer to the file location stored in the DAQ logbook where the SHUTTLE can collect the information and store it in the OCDB. The same considerations apply to the relative time alignment of the electronics.
- Parameters of the FEE required for reconstruction, such as the tail cancellation, zero suppression, sampling frequency and acquisition window status bits, will be written by the RCU into the event trailer.
- The list of active channels is calculated with physics data offline.
- The pad to pad gain calibration is calculated in special runs by injection Kr86 in the TPC. The calculation is performed offline and is repeated rarely.
- The drift velocity is monitored during data taking both with laser data (electrons drifting from the central membrane) and using two GOOFY (the TPC drift velocity monitoring system) systems. The calculation can be performed online either in the DAQ or the HLT farm. Results will be collected by the SHUTTLE and stored into the OCDB. The drift velocity map is calculated once and updated very rarely.
- The distortions map due to $E \times B$ is calculated offline with physics data. This is a lengthy process and will not be available for first pass reconstruction.
- Various parameters, such as the temperature values, the Oxygen content, high voltage values, gating voltages and magnetic field, are monitored by DCS and stored in the DCS archive DB. The SHUTTLE collects this information and stores it into the OCDB.

I do not have a good appreciation on the status of the implementation of this strategy.

ACORDE

- During data taking, the counts of each element (60) of ACORDE are continuously incremented.
- At the end of the run, the average count rate is calculated making use of the begin of run and end of run time stamps available from the DAQ LogBook. The result is stored in the DAQ FES.

- The SHUTTLE collects the data from the DAQ FES after rootification by the SHUTTLE pre-processor.
- The DCS collects the data from the DAQ FES, compares the values to reference values stored in the DCS Configuration DB and calculates new settings of the phototubes high voltage. The new values are set by DCS.

General comments

- The OCDB does not hold the data itself but instead the unique file identification of the data files, stored permanently in the MSS (MSS), together with validity time stamps.
- The calibration file format is exclusively the ROOT format.
- The source of calibration parameters can be data from specialized runs or from physics runs.
- The calibration parameters are calculated with specialized algorithms not necessarily integrated in the AliRoot framework.
- They can run either online in the DCS, DAQ or HLT farm or quasi-online on the CAF.
- Results are permanently stored in the CERN MSS and catalogued in the offline calibration DB where they are accessed from by the reconstruction algorithms.
- Parameters calculated online are temporarily stored on disks local to DCS, DAQ or HLT where they will be collected by the SHUTTLE.
- There will be one SHUTTLE pre-processor per detector implementing all necessary procedures for data manipulation from all input sources.
- Parameters calculated in the CAF can be directly stored and catalogued using the AliRoot CDB framework.
- Only parameters monitored by DCS will be stored in the DCS Archive DB.
- Three exchange file system writeable by the owner system (DCS, DAQ, HLT) and readable by everybody else (+offline)

Use cases

I tried to find a set of use cases that should cover the various requirements by the detectors. I suggest to:

- discuss these use cases with the detector experts, the DCS, ECS, DAQ, HLT and Offline experts (tentatively end of June)
- define what has to be done by detectors and systems (should be attached to this document)
- prototype and test (wrap up session during October offline week)

Use case 1:

Parameters are computed from the DAQ LDC/GDC data during physics run or dedicated run. The parameters are needed offline by the reconstruction algorithms.

- In the case of a dedicated run, the calculations requiring sub-events only (data on a single DDL) are run on the relevant LDC, data requiring the whole event are run on any GDC. In the case of calibration triggers during a physics run, the calculations will be performed by a monitoring program running on one DAQ node and accessing remotely the LDC or GDC.

- During the run data are processed to produce the required parameters. Results are stored temporarily on the DAQ file exchange server. The detector provides the processing module; the DAQ provides the file system and the module storing the data. The format is preferably a ROOT file but can be any kind of format.
- The DAQ writes in the DAQ Logbook the address of the produced files together with the start and end of run timestamps.
- At the end of the run, after the processing of the data, the ECS sends a signal to the SHUTTLE. The SHUTTLE queries the DAQ Logbook for file name and timestamps and fetches the appropriate parameter files. It stores the files in the MSS and adds an entry (timestamps and unique identifier of the files) into the OCDB. Offline provides the SHUTTLE framework. The detector provides an optional pre-processor algorithm where new parameters can be calculated for example, and the parameter file can be ROOTified if necessary.

Use case 2:

Parameters are computed in the HLT farm from physics data or dedicated data e.g. pedestals. The parameters are needed offline by the reconstruction algorithms. This case is similar to the previous one. Only differences are reported:

- Calculations are running on any HLT node.
- Results are stored temporarily on the HLT file exchange server. The detector provides the processing module which will run in the HLT environment. It should be strictly identical to the same module running in the DAQ environment. The HLT provides the file system and the module storing the data.
- An HLT Logbook, or more precisely a file exchange server, must be created.

Use case 3:

Parameters are computed in the DCS farm from dedicated data. The parameters are needed offline by the reconstruction algorithms. This case is similar to the previous one. Only differences are reported:

- Calculations are running on a dedicated DCS node.
- Results are stored temporarily on the DCS file exchange server. The detector provides the processing module which will run in the DCS environment. It should be strictly identical to the same module running in the DAQ environment. The DCS provides the file system and the module storing the data.
- It is assumed that the DCS has write access to the DAQ Logbook. If not, it could be made available in the DCS archive DB.

Use Case 4:

Parameters are monitored by DCS and are required offline by the offline reconstruction algorithms.

- These parameters are stored in the DCS Archive DB and updated according to the specified frequency.

- At the end of the run, the SHUTTLE queries the DAQ Logbook for the start and end of run timestamps and then queries the Archive for the relevant parameters. The SHUTTLE pre-processor ROOTifies the parameters, the ROOT files are added to the file exchange server.
- It is the responsibility of the detectors to provide DCS and Offline the list of parameters.

Use Case 5:

The FEE parameters are required by the offline reconstruction algorithms.

- The setting parameters of the FEE are formatted in a file and fed to the electronics either trough the DCS or through the DAQ DDL. If DCS is in charge, the file is stored in the DCS Configuration DB and can be made available in the DCS file exchange server, the notification being done through the DCS Archive DB. If DAQ is in charge, the file can be stored directly in the MSS and its unique identification registered in the DAQ Logbook.
- At the end of the run, the ECS notifies the SHUTTLE which queries the DCS Archive DB and/or the DAQ Logbook for timestamps and file identification, collects the relevant data, optionally pre-processes the data (ROOTification,…), stores them in the MSS and updates the OCDB.

To do

In order to implement these use cases the following items have to be provided:

- **Detector:**
    - i. Processing algorithms running both in a DAQ LDC/GDC node and in a HLT node.
    - ii. Pre-processing algorithms of the SHUTTLE.
    - iii. Make DCS and Offline aware of the parameters to be collected from the DCS Archive DB.
    - iv. Evaluate the CPU cycles required during a physics run for the parameters generation.
- **DAQ and ECS:**
    - i. A file exchange server with write access by the DAQ and read access by DCS, HLT and Offline. The DAQ file exchange server is implemented. DCS and HLT could reuse the same design: sql structure and code (it's just a table definition and a simple query to populate it). The file access mechanism can use either SCP or SFTP.
    - ii. A file system similar to the ones of the DCS and HLT.
- **DCS:**
    - i. A file exchange server with write access by the DAQ and read access by DCS, HLT and Offline.
    - ii. A file system similar to the ones of the DAQ and HLT.
- **HLT:**
    - i. A file exchange server with write access by the DAQ and read access by DCS, HLT and Offline.
    - ii. A file system similar to the ones of the DAQ and DCS.

- **Offline:**
    i. Provide the SHUTTLE, the OCDB and the communication framework of AliRoot and the OCDB and the supporting MSS. All these tasks are done.
    ii. Make the SHUTTLE aware of the various file exchange servers.
    iii. Provide an environment which will allow the detector experts to simulate the entire strategy for the creation and archiving of the calibration parameters.
- **All:** implement and exercise the strategy in real while commissioning the detectors.