The CMS Framework for Alignment and Calibration

Gero Flucke[*][†]
University of Hamburg, Institute for Experimental Physics, Hamburg, Germany
E-mail: gero.flucke@desy.de

David Futyan
Imperial College, University of London, London, United Kingdom
E-mail: david.futyan@cern.ch

Rainer Mankel[‡]
Deutsches Elektronen-Synchrotron, Hamburg, Germany
E-mail: rainer.mankel@desy.de

The ultimate performance of the CMS detector relies crucially on precise and prompt alignment and calibration of its components. A sizable number of workflows need to be coordinated and performed with minimal delay through the use of a computing infrastructure which is able to provide the constants for a timely reconstruction of the data for subsequent physics analysis. The framework supporting these processes is presented together with alignment and calibration results in simulated LHC start-up scenarios and from cosmic muons recorded with the CMS solenoidal magnetic field being switched on and off.

---

[*]Speaker.

[†]On behalf of the CMS Collaboration.

[‡]Also at CERN, Geneva, Switzerland.

1. Introduction

The Compact Muon Solenoid (CMS) experiment [1] is located at the access point P5 of the Large Hadron Collider (LHC) [2] at CERN, close to the village of Cessy. From the beam line outwards, the detector consists of a silicon based inner tracking system, an electromagnetic and a hadronic calorimeter (ECAL and HCAL, respectively), a superconducting solenoid, providing a magnetic field of 3.8 T, and a system of three kinds of muon tracking chambers (DT, CSC and RPC) interleaved with an iron return yoke. The overall radius of the CMS detector amounts to $r = 7.5$ m, measured from the LHC beam line which is parallel to the solenoidal field and coincides with the z-axis. The azimuthal angle ϕ is measured in the plane perpendicular to the z-axis.

The analysis of the complex LHC collision events requires a reconstruction with ambitious goals in terms of resolution which can only be achieved by excellent alignment and calibration. To maintain fast availability of samples for physics analysis, the objective is a prompt calibration with a fast turnaround. Thus a robust and reliable concept is required to ensure processing of a large number of calibration events with short I/O latencies.

2. Offline Alignment and Calibration Workflow

The CMS alignment and calibration workflow has to ensure that the prompt offline event reconstruction can apply alignment and calibration constants that are already updated for possible rapidly changing data taking conditions. This sets the time scale for the determination of these constants to be about 24 hours. In addition, the framework has to support longer latency workflows for constants that change less rapidly and require a larger data set than the one that can be accumulated within this relatively short time. These constants can then be applied in a later re-reconstruction.

The key components of the alignment and calibration workflows are the online data processing at the CMS detector site (P5), including the software based High Level Trigger (HLT), and the offline processing at the Tier 0 (T0) and the CERN Analysis Facility (CAF), both located at the CERN Meyrin site. Special alignment and calibration datasets are called AlCaReco and AlCaRaw. The interplay between the key components and the data streams is sketched in Fig. 1 and detailed in the following.

P5 The HLT computers select events and feed them to the Storage Manager. Here the events are buffered and transferred to the T0 in a special streamer format. According to the HLT decision, different data streams are processed: one for the bulk data for physics analysis and a subset thereof as an express stream for the prompt alignment and calibration as well as for physics data quality monitoring. In addition, some calibration tasks require such a high data rate that special calibration streams (AlCaRaw) are created where the content of selected events is already reduced before sending the data.

T0 The data is first repacked from the streamer format into the ROOT based CMS Event Data Model (EDM), and primary datasets are built based on HLT decisions. The physics data is cached on disk for about 24 hours until the prompt alignment and calibration constants are available for reconstruction. Express and calibration streams are reconstructed with lower
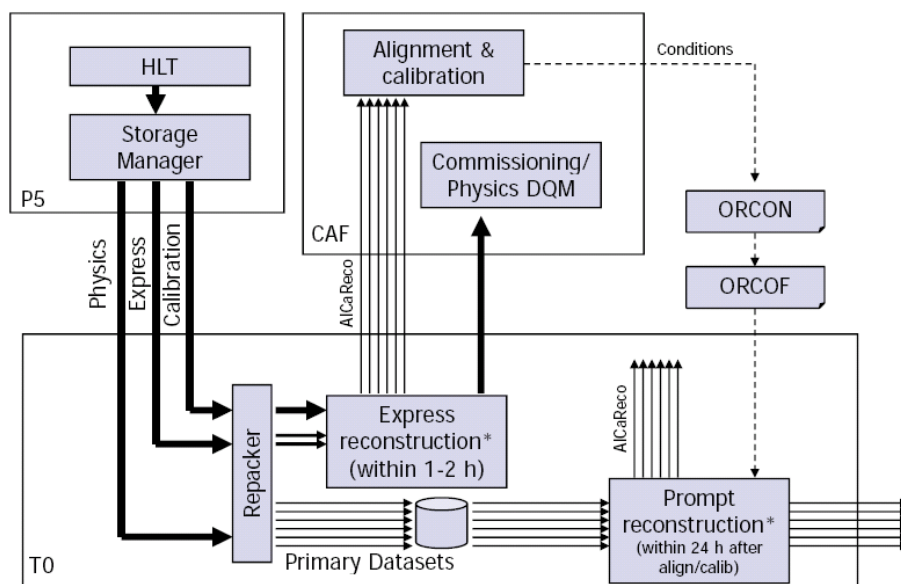
Figure 1: *Sketch of the CMS alignment and calibration workflows.*

latency. AlCaReco datasets are created both from the calibration streams and normal data. The number of events and the event content for these datasets are reduced to exactly match the requirements of the different alignment and calibration tasks. In that way I/O latencies can be avoided.

CAF The CAF is the prime platform for offline alignment and calibration algorithms. The Al-CaReco datasets are stored on a dedicated CAF disk pool. This ensures fast access by the algorithms that run on CAF farm queues to compute constants and validate them.

Databases Alignment and calibration constants derived at the CAF are transferred to the online environment at P5. There they are uploaded to an online database (ORCON) from where they are automatically streamed to the offline counterpart (ORCOF). The prompt reconstruction at T0 reads the constants, similarly to analysis jobs, via intermediate caching layers, called Frontier [3]. These http based proxy servers avoid overloading the underlying database.

3. Simulation Exercise

The CMS alignment and calibration framework has been extensively tested during the Computing, Software and Analysis Challenge that took place in May 2008 (CSA08). To make this simulation exercise as realistic as possible, two scenarios from the LHC commissioning schedule have been chosen, corresponding to 43 and 156 colliding bunches in both beams (see Tab. 1). For both scenarios, one week of data taking has been simulated, assuming misalignment and miscalibration as expected at the start-up of the LHC. Alignment and calibration has been performed at the CAF. To achieve a quasi real-time exercise, constants have had to be computed and validated within one week after arrival of the data and have then been used for a re-reconstruction of the data for physics analysis.

3

Name	Bunch schema	Luminosity	Duration [effective]	Integrated Luminosity	HLT Output	Events
S43	43x43	$2 \cdot 10^{30}$ cm^{-2}s^{-1}	6 days	1 pb^{-1}	300 Hz	150 M
S156	156x156	$2 \cdot 10^{31}$ cm^{-2}s^{-1}	6 days	10 pb^{-1}	300 Hz	150 M

Table 1: *LHC commissioning scenarios used in CSA08.*

Task	Workflow	Task	Workflow
Tracker Alignment	Millepede algorithm	HCAL Calibration	ϕ-symmetry
	HIP algorithm		use of isolated tracks
	Kalman algorithm		di-jet balancing
Muon System Alignment	HIP algorithm	Muon DT Calibration	time pedestal
	Standalone algorithm		drift velocity
ECAL Calibration	ϕ-symmetry	Tracker Calibration	Pixel Lorentz angle
	use of $\pi^0 \to \gamma\gamma$		Strip Lorentz angle
	use of $Z \to e^+e^-$		Strip charge response
Muon RPC Monitoring			

Table 2: *Alignment and calibration tasks and specific workflows tested during CSA08 (S43 and/or S156).*

Various alignment and calibration workflows have been run, as listed in Tab. 2. For the first time, interdependencies have been taken into account, e. g. the muon system alignment with the HIP algorithm could not start before tracker alignment (since the algorithm uses extrapolated tracks from the tracker) and DT calibration (influencing reconstructed hit positions) have been finished.

In the following, two example workflows and their results will be described in more detail.

3.1 Tracker Alignment

The CMS Tracker is an all-silicon detector with 1440 pixel sensors at radii up to $r < 11$ cm and 15148 strip modules at $20 < r < 115$ cm [1]. A track from the interaction point crosses about 13 layers, varying with its polar angle. With a resolution ranging from 9 to 60 μm in the sensitive direction, precise alignment of the tracker is mandatory to fully exploit its potential for physics analysis. Taking into account four degrees of freedom per pixel sensor and three to four per strip module[1] results in more than 44000 parameters to be determined. In both CSA08 scenarios, misalignment has been introduced as it is expected to remain after application of survey and laser alignment system measurements as well as a first track based alignment using cosmic muon tracks [4].

Three different algorithms have been run in CSA08. All minimise the χ^2 of the track fits to determine the alignment parameters. They share a common infrastructure for parameter handling and the calculation of the needed derivatives. Results obtained using the global fit algorithm Millepede II [5] have performed best in terms of χ^2-minimisation and comparison with ideal geometry. Fig. 2 shows the resulting normalised χ^2 and momentum resolution distributions. More than $4 \cdot 10^6$

[1]Two silicon strip modules mounted back-to-back with a stereo angle of 100 mrad between the strip orientations are treated as one entity which is referred to as "2D module".
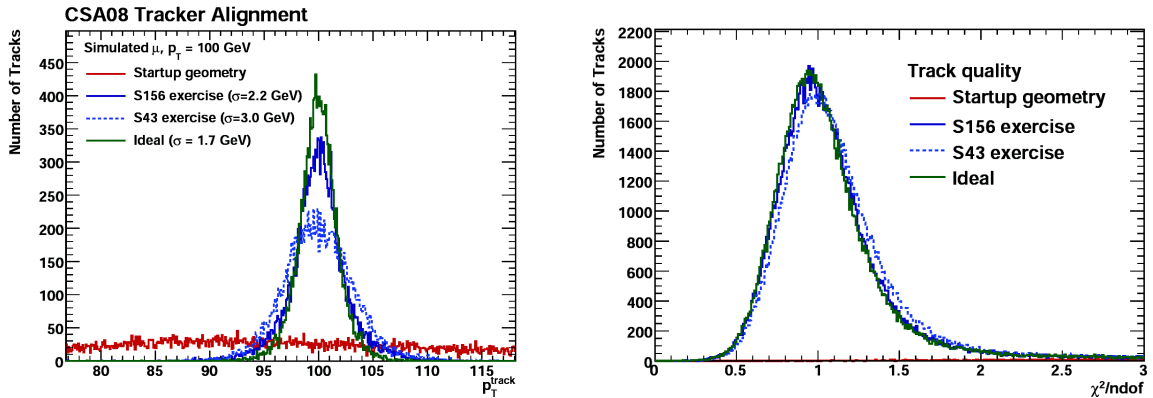
Figure 2: *Transverse momentum (p_t, left) and χ^2 per degrees of freedom (right) distributions of simulated muon tracks with $p_t = 100$ GeV for geometries obtained by Millepede II in the CSA08 exercises compared with the ideal and start-up case.*

tracks have been processed in 50 parallel jobs of about 30 minutes CPU time on the CAF, followed by one job of about 5 hours CPU. The main AlCaReco datasets used have been those selecting minimum bias tracks, isolated muons, muons from Z^0 and J/Ψ decays and cosmic ray muons. The fact that the cosmic sample did not arrive in time for the corresponding alignment workflow is a major reason for its worse momentum resolution.

Overall, the CSA08 Tracker alignment exercise has been a very successful test of the production of five AlCaReco datasets for collision data taking, three for special data taking like cosmic runs, achieved a high accuracy and provided constants well in time for the depending workflows like Muon HIP alignment and tracker Lorentz angle calibration. However, it has to be noted that this simulation study did not include all possible misalignment effects expected in data, e.g. relative mispositioning of the components of the 2D strip modules. An extension of the alignment procedures to cover this degree of freedom is currently under study in CMS.

3.2 ECAL Calibration using $\pi^0 \to \gamma\gamma$

The ECAL consists of 61200 Lead Tungstate ($PbWO_4$) crystals in the central barrel part and 7324 in each endcap [1]. The startup intercalibration precision of the ECAL barrel is apt to provide an acceptable performance of the detector for startup physics and can hardly be improved after 1-10 pb^{-1} of integrated luminosity. On the other hand, a fast improvement on the initial miscalibration of the ECAL endcaps is desirable and the CSA08 has been an excellent opportunity to test the data handling procedures.

The π^0 calibration concept consists in obtaining a channel-by-channel (inter-)calibration of the ECAL response by reconstructing the invariant mass of $\pi^0 \to \gamma\gamma$ decays. The target precision of 0.5% requires the accumulation of about 2000 photons per crystal from π^0 decays where $E_t(\pi^0) \sim$ 5 GeV. The high rate of 1 kHz for useful π^0 candidates requires the use of an AlCaRaw stream from P5 to T0 to avoid saturating the bandwidth. A dedicated HLT filter stores only the energies of crystals in a small area surrounding the π^0 candidates. Such candidates are selected with cuts on the energies, transverse shower profiles and separation between the two photon candidates.
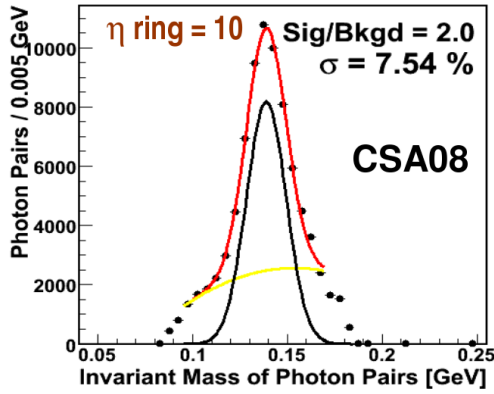
5

Figure 3: *CSA08 ECAL π^0 calibration: Invariant mass of photon pairs selected at a fixed η value. The black and yellow lines represent the signal and background part, respectively, of the combined fit (red line).*
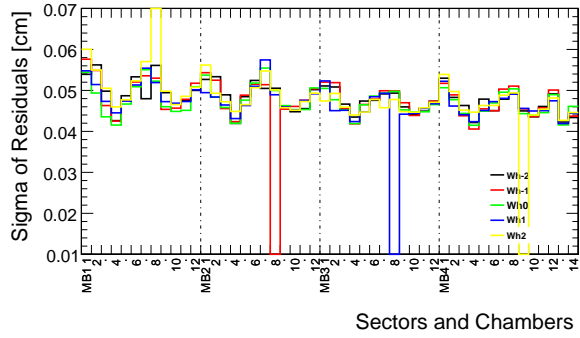

Figure 4: *CRAFT DT calibration: Widths of residual distributions for chambers obtained after t_{trig}-calibration. The dash-dotted lines separate chambers at different radii. For each radius, the chambers are sorted in ϕ, starting at $\phi = 0$. Colours represent different z positions.*

In CSA08 the offline part of this workflow has been exercised. The HLT filter has been run for simulated "minimum bias" data and for jet events with $E_t > 20$ GeV, requiring in both cases an emulated level 1 jet trigger accept. The AlCaReco dataset has been produced from output of this HLT filter and further analysis, including a refined selection and the derivation of the calibration constants (details see [6]) has taken place at the CAF.

Fig. 3 shows the resulting π^0 peak for one fixed pseudorapidity (η) value which is fitted to a combination of a Gaussian and a second-order polynomial for the signal and the background, respectively. The background distribution is cut-off close to the signal due to the preselection which was judged to be somewhat tight in CSA08, and will be extended for real LHC data taking. No noticeable systematic effect is observed that would compromise the ability of the method to reach the envisaged 0.5% calibration precision. In addition, the experience from CSA08 has been used to tune the selection procedures to require no more than 10-15 CPUs allocated for the π^0 calibration at the CAF.

4. Real Data Experience: Cosmic Mun Runs

In 2008 the CMS detector has regularly been run around the clock in so called Global Runs to collect muon data from cosmic ray events. Since July most relevant components, including the Pixel Tracker, have participated. Runs with the magnetic field being switched off are called "Cosmic Run at Zero Tesla" (CRUZET), while from mid October the "Cosmic Run at (almost) Four Tesla" (CRAFT) took place for four weeks.

The alignment and calibration framework has been steadily commissioned during these data taking periods. Since July the first AlCaReco datasets have been produced centrally by the data operations team. At the end of CRAFT, nine datasets have been part of the prompt production workflow. Also the four calibration streams from P5 to T0 (AlCaRaw) have been successfully put into operation. Full workflows are already established as explained in more detail for two example alignment and calibration tasks presented in the following.
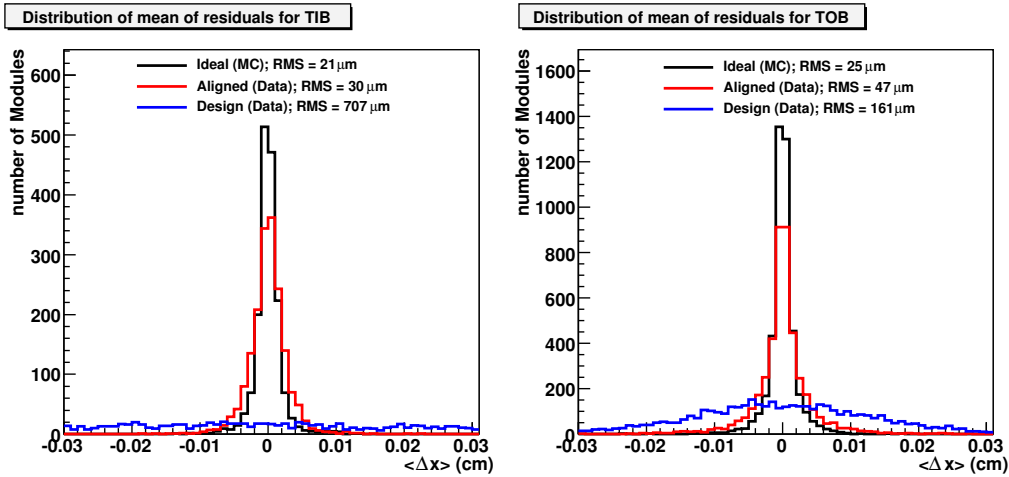
Figure 5: *Mean values of the distributions of the residuals Δx for the modules in the inner (left) and outer (right) barrel of the silicon strip detector with at least 100 hits, comparing the situations before (blue) and after (red) alignment with a simulation under ideal conditions (black).*

4.1 Muon Drift Tube Time Pedestal Calibration

The CMS Barrel Muon system primarily consists of drift tube (DT) detectors [1]. The DT chambers are interleaved with the magnet return yoke and located at four different radii. Most chambers comprise three superlayers, two measuring in ϕ-, one in z-direction.

As for a drift detector, the spatial measurement precision depends strongly on the exact time at which a signal is obtained, making the precise determination of the time pedestal (t_{trig}) essential. This is done by investigating, at a superlayer granularity, the time distribution of the digital data. Since the digital data is produced during reconstruction, but not stored in its output, the corresponding AlCaReco dataset cannot simply be a skim, but runs in parallel with the reconstruction. Currently, to avoid any delay for the reconstruction, the dataset is produced separately re-running the reconstruction code.

The rising edge of the time distribution is fitted with the integral of a Gaussian with mean μ and width σ. The pedestal is determined as $t_{trig} = \mu - k \cdot \sigma$ where the factor k is tuned by investigating the resulting residuals between hit and track segment predictions, see [7] for details of this procedure. The widths of the resulting residual distributions obtained from CRAFT data, shown in Fig. 4, are about 500 μm, except for some chambers with known hardware problems.

As expected, the observed widths are considerably larger than the intrinsic spatial resolution of the chamber, which is due to the arbitrary arrival time of the cosmic muons relative to the bunch clock steering the readout. Horizontal chambers have broader widths due to lower statistics. Nevertheless, the t_{trig}-calibration workflow, including validation, is fully commissioned.

4.2 Tracker Alignment

The Tracker alignment algorithms have been exercised already with the very first data taken by the tracker in early July 2008. A coherent data set of almost $3.5 \cdot 10^5$ events in the dedicated AlCaReco dataset has then been collected in the CRUZET run at the end of July. The absence of the

magnetic field inhibits a correct treatment of multiple scattering effects since track momenta cannot be measured. A momentum assumption of $p = 5$ GeV/c is used to estimate the uncertainties of the track hit residuals. Despite this obstacle, the first tracker alignment results have been achieved within a week. Since the underground topology constrains the cosmic muons to arrive dominantly with roughly vertical incidence, the hit statistics in the endcaps is very limited and the results mainly cover the barrel part. Here the mean values of the residual distributions per module become much smaller after alignment with the HIP algorithm [8], as can be seen in Fig. 5 for the strip barrel. The width of this distribution has already the same order of magnitude that is obtained in simulation with an ideal geometry, i.e. 21 μm (25 μm) in the inner (outer) part of the strip barrel.

Meanwhile the much larger CRAFT dataset ($> 4.5 \cdot 10^6$ AlCaReco events for tracker alignment) is being analysed, increasing the precision and extending the region aligned on module level to the complete barrel, including the pixel detector, and the majority of the modules in the endcaps.

5. Conclusions

Physics analysis of the data collected with the CMS detector will require excellent alignment and calibration. To allow a fast turnaround, the offline framework for alignment and calibration has to provide condition constants already for the prompt reconstruction pass, i.e. within 24 hours. The backbone of this framework are the AlCaReco and AlCaRaw data skims, containing only selected events and event content, dedicated for certain alignment and calibration tasks. This ensures, together with a disk storage at the CAF, that I/O latencies are avoided.

The framework has been extensively tested on simulated data during the CSA08 campaign in May 2008, taking care of interdependencies between the 17 tested workflows and providing condition constants within a week. Since June 2008, production of AlCaReco datasets has been steadily commissioned in CMS Global Runs collecting cosmic ray data. At the end of CRAFT data taking, nine AlCaReco datasets have been produced, including those based on the calibration streams (AlCaRaw). Using these skimmed data sets, complete workflows like DT time pedestal calibration and tracker alignment have been accomplished and provide condition constants valuable for the understanding of the detector and the analysis of the cosmic data.

Together both, CSA08 and experience with cosmic data taking, show that the CMS framework for alignment and calibration is well set to face the challenges expected once the LHC starts to provide collisions in 2009.

References

[1] R. Adolphi *et al.* [CMS Collaboration], *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[2] L. Evans and P. Bryant (Eds.), *LHC Machine*, JINST **3** (2008) S08001.

[3] S. Kosyakov *et al.*, *FroNtier: High Performance Database Access using Standard Web Components in a Scalable Multi-Tier Architecture*, in proceedings of *Computing in High Energy Physics*, Interlaken, Switzerland (2004).

[4] T. Lampén, N. de Filippis, F.-P. Schilling, A. Schmidt, M. Weber, *Comprehensive Set of Misalignment Scenarios for the CMS Tracker* [CMS NOTE-2008/029].

[5] V. Blobel, *Software Alignment for Tracking Detectors*, Nucl. Instrum. Meth. A **566** (2006) 5.

[6] M. Gataullin *et al.*, *Intercalibration of the CMS Barrel Electromagnetic Calorimeter Using Neutral Pion Decays* [CMS DN-2007/013].

[7] N. Amapane *et al.*, *Offline Calibration Procedure of the Drift Tube Detectors* [CMS NOTE-2007/034].

[8] V. Karimaki, T. Lampén, F.-P. Schilling, *The HIP Algorithm for Track Based Alignment and its Application to the CMS Pixel Detector* [CMS NOTE-2006/018].