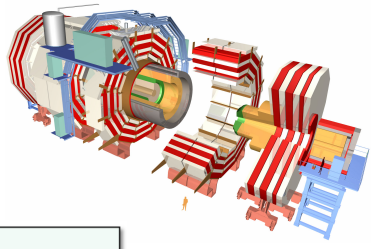
Commissioning the CMS Alignment and Calibration Framework

- CMS Collaboration -

17th International Conference on Computing in High Energy and Nuclear Physics, Prague, March 2009

Presenter: David Futyan, Imperial College London (david.futyan@cern.ch)

The CMS experiment has developed a powerful framework to ensure the precise and prompt alignment and calibration of its components, which is a major prerequisite to achieve the optimal performance for physics analysis.



General Strategy

Calibration/alignment workflows categorized by their required latency

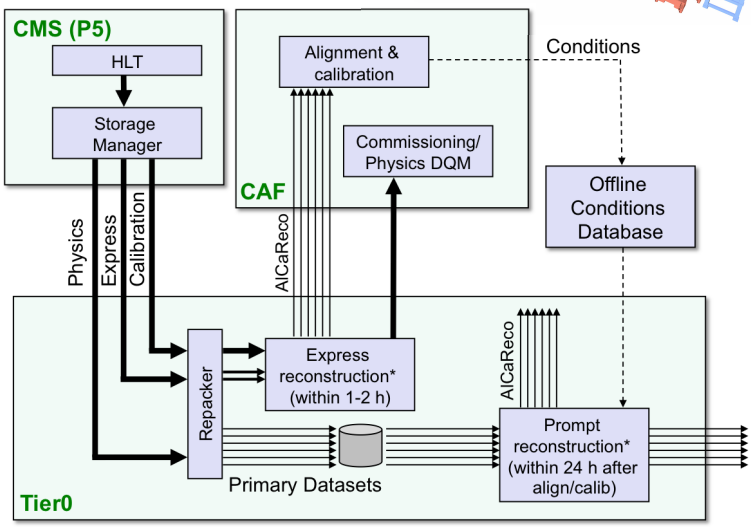
- ❑**Detector-near calibrations** (pedestal/pulser, laser runs etc): usually determined online
- ❑**Prompt calibration:** constants determined offline, synchronously with data-taking, and used in first full reconstruction of data sample
- ❑**Long-range calibrations:** constants determined using a data sample accumulated over a longer range of time and applied to the data in re-reconstruction campaigns

Prompt Calibration Loop

Goal: **provide good alignment/calibration constants in time for the first (prompt) reconstruction of the data sample**.

While the main data stream events are **buffered at the Tier0**, express and calibration streams are passed through an **express reconstruction** within 1-2 hours. Reconstructed data and **AlCaReco skims** (see below) are transferred to the **CERN Analysis Facility (CAF)**, where they are used as input to prompt alignment/calibration workflows to produce new constants. These constants are validated and uploaded to the **conditions database**.

When prompt alignment and calibration are completed (target latency: ~24h), the full data sample is passed through **prompt reconstruction** with the updated constants, and generally made available for analysis



Data Streams from CMS (P5) to Tier0

Data is transferred in three principal streams:

- ❑**Physics stream**: The main data stream intended for physics analysis
- ❑**Express stream:** Specially selected subset of events (~10%, initially more, of the total data bandwidth), chosen according to trigger bits, for prompt physics analysis, calibration and alignment, detector and trigger commissioning, fast physics validation
- ❑**Calibration Streams:** Dedicated streams for calibration and alignment (~10% of the total data bandwidth). Includes:
  - ❑**Hardware calibration streams:** Events taken during the LHC orbit gap, important e.g. for pedestal and noise calibration, when no signal is present
  - ❑**"AlCaRaw" streams**

AlCaRaw Streams

Three calibration procedures require a very high event rate (~1kHz each) which would saturate the affordable data bandwidth (~300MB/s) if the full event content were transferred:

- ❑HCAL ϕ-symmetry calibration
- ❑ECAL ϕ-symmetry calibration
- ❑ECAL calibration with π^0s

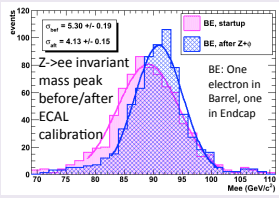
Instead of transferring full RAW event content, store only minimal information needed for the corresponding calibration algorithms. Special high rate triggers are processed on the HLT farm at P5, producing dedicated output modules with event content reduced by factor ~100 w.r.t. the full RAW event.

Commissioning with Monte Carlo

Computing Software and Analysis Challenge (CSA08), May 2008:

Test full scope of **offline** data handling and analysis activities needed for early LHC data-taking operations - First **full-scale challenge** with **large statistics** under conditions similar to LHC startup:

- ❑Initial **mis-alignments** and **mis-calibrations** as expected before collisions
- ❑**Full end-to-end offline workflows** for alignment/calibration (Tier0, CAF) exercised in "real-time": schedule matched to pace of reconstruction
- ❑Resulting constants used to **re-reconstruct** data for physics analysis
- ❑Full complexity of ~20 concurrent alignment & calibration workflows, with **Interdependencies** taken into account



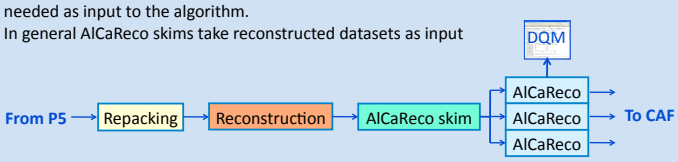
AlCaReco Skims

Calibration/alignment algorithms need to run over very large datasets, in some cases with many iterations, in a short time.

AlCaReco datasets are **skims of reconstructed data** (both event selection and event content selection), containing the **minimal information required as input to a given calibration/ alignment algorithm.**

All AlCaReco skims perform an initial filtering of events based on a set of trigger bits. In some cases additional processing is performed to produce specialized data collections needed as input to the algorithm.

In general AlCaReco skims take reconstructed datasets as input



A subset of AlCaReco skims is run on the express stream as needed for the prompt calibration loop

Commissioning with Cosmic Data

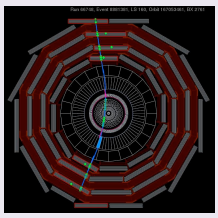
Full alignment and calibration workflows have operated extensively during data taking periods in 2008:

CRAFT: Major cosmic data taking period Oct/Nov 2008

- ❑4 weeks of continuous cosmic data taking, with the **complete detector** operating with **B field on** (3.8 Tesla)
- ❑375 million events collected

CRUZET: Cosmic data taking with B field off, July-Sept 2008

Beam commissioning 10th September



Steady ramping up of alignment/calibration workflows during CRUZET.

Alignment/calibration procedures operated for all subdetectors during CRAFT, including:

- ❑**AlCaRaw** streaming from P5
- ❑Centrally produced **AlCaReco** datasets at Tier0
- ❑Transfer to **CAF** for constants determination

Newly determined constants used for two subsequent **re-processings** of the full dataset, following formal validation and sign-off procedures.

Significant improvements obtained e.g. for tracker alignment w.r.t. survey measurements