AUTOMATED WORKFLOWS FOR CRITICAL TIME-DEPENDENT CALIBRATIONS AT THE CMS EXPERIMENT

Gianluca Cerminara (CERN) on behalf of the CMS Collaboration

Prompt Calibration Concept

Low latency calibration workflows run immediately after the data-taking:
- conditions which need continuous updates
- conditions which need monitoring

Update strategy based on delay between "express reconstruction" of a sub-set of the data and "prompt reconstruction" of the bulk of the data for Physics Analysis

Run unattended on the Tier0 farm:
- process a subset of the data of each run
- compute the calibration conditions
- upload to the condition DB (CondDB)
- conditions are consumed by prompt reconstruction of the same run

Goal: best calibrations already available when data are promptly reconstructed

- data analysis ready a few hours after they have been acquired
- reduce the need for re-reconstruction passes

Express & Automated Calibration running on Tier0 Computer Farm

example of workflow running on a single run-X

RECO + skimming (parallel processing)

Prompt Calibration Algorithms

AlCa Harvesting
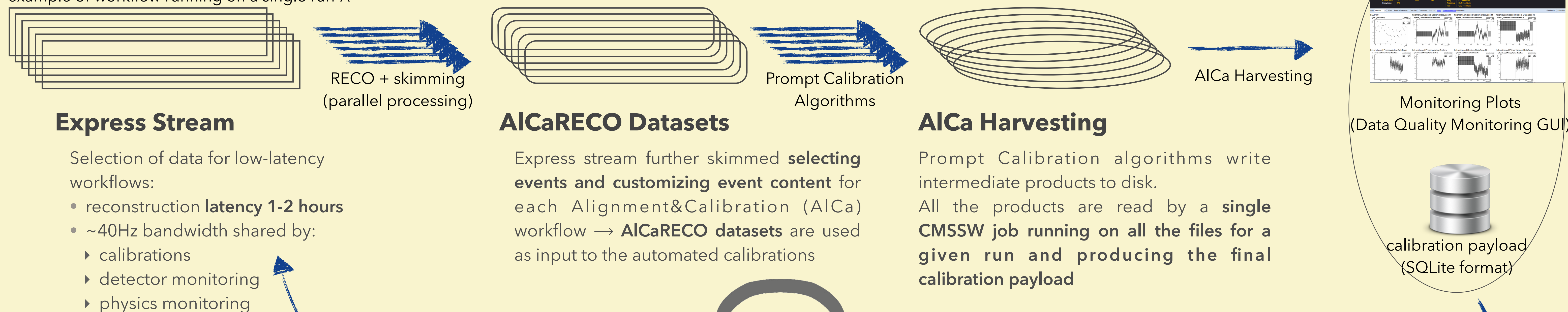
Express Stream

Selection of data for low-latency workflows:
- reconstruction **latency 1-2 hours**
- ~40Hz bandwidth shared by:
  - calibrations
  - detector monitoring
  - physics monitoring

AlCaRECO Datasets

Express stream further skimmed **selecting events and customizing event content** for each Alignment&Calibration (AlCa) workflow → **AlCaRECO datasets** are used as input to the automated calibrations

AlCa Harvesting

Prompt Calibration algorithms write intermediate products to disk.
All the products are read by a **single CMSSW job running on all the files for a given run and producing the final calibration payload**

Monitoring Plots (Data Quality Monitoring GUI)

calibration payload (SQLite format)

Calibrations consumed by Prompt Reconstruction

The reconstruction of the bulk of the data for physics analysis starts **48h after the acquisition run** profiting of updated calibrations.

The offline applications retrieve the condition data for the Interval of Validity (IoV) corresponding to the run-X (already processed by the Prompt Calibration) as C++ objects.

The FroNTier middleware takes care of the distribution and caching layer.

The performance of the automated algorithms and of the reconstruction is monitored via DQM applications producing histograms displayed on a web-based GUI.

1-2h

t=0

48h
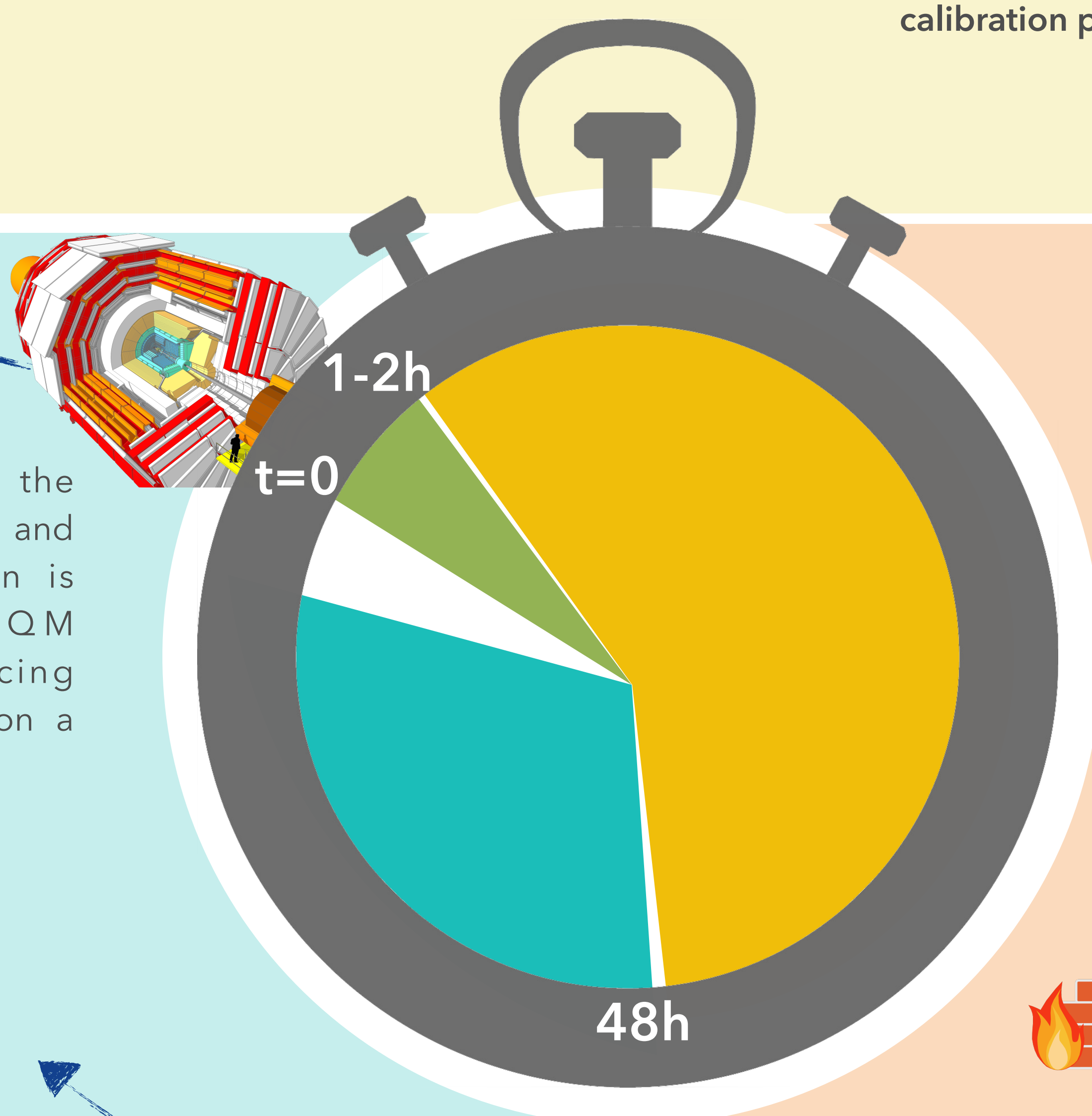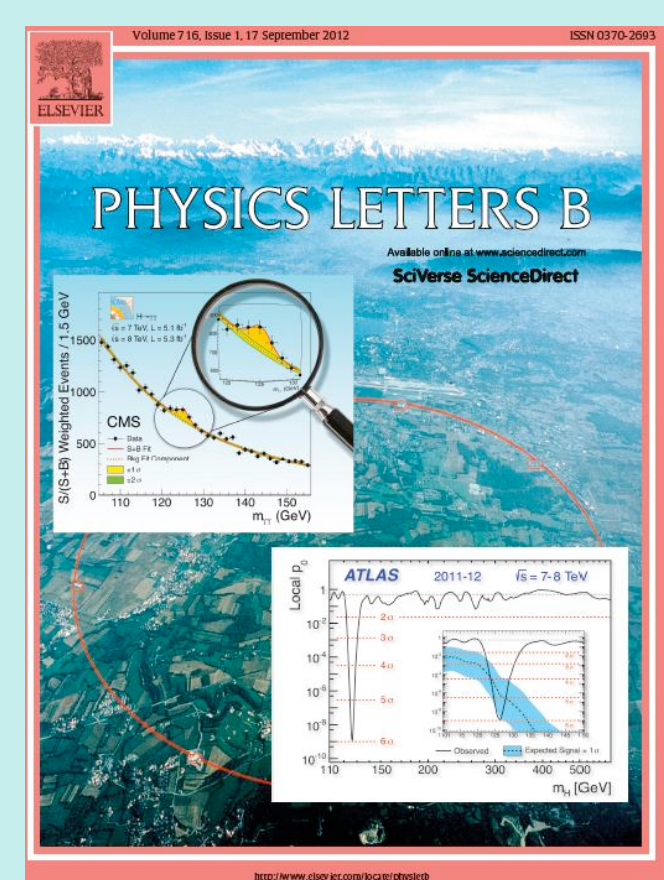
Upload to Condition DB

The Alignment and Calibration constants are stored in the Condition Database (ORACLE).
Two instances of the Condition DB exist and are synched via the Active Data Guard (ADG) technology:

- **ORCON** serves condition data needed by HLT, DQM and offline reconstruction - only writeable from CMS dedicated network
- **ORCOFF** serves offline applications - read-only

ORCON

ADG   .cms technical network
       .cern.ch general network

ORCOFF

{JSON} +
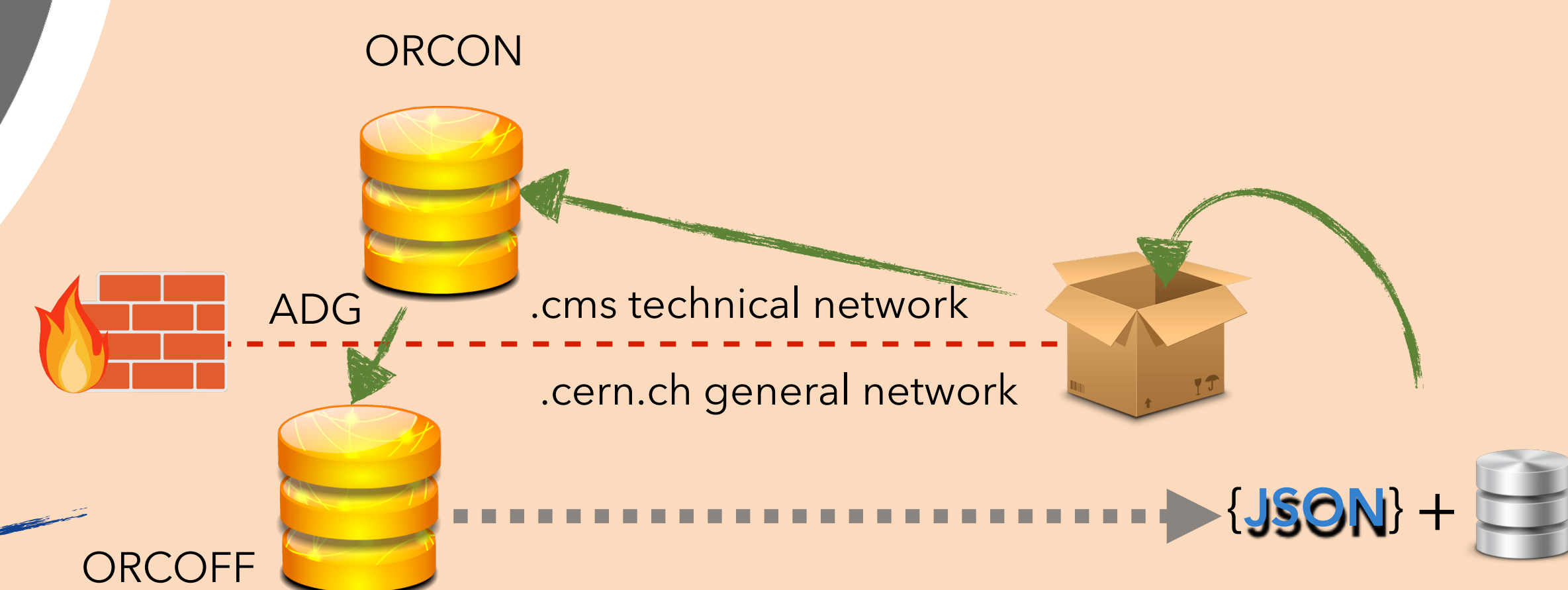
Offline DropBox

Uploads from Tier0 are performed using the Offline Dropbox Service: calibration payloads are exported from SQLite to ORACLE by this service sitting **across the firewall**.

The dropbox **enforces reproducibility** policy avoiding to overwrite conditions already consumed by HLT or Prompt Reconstruction.

The **upload is steered by a metadata file** (JSON dictionary format) specifying:
- destination account
- target tag name
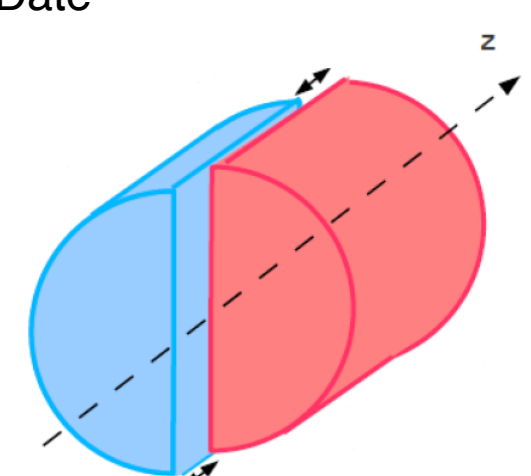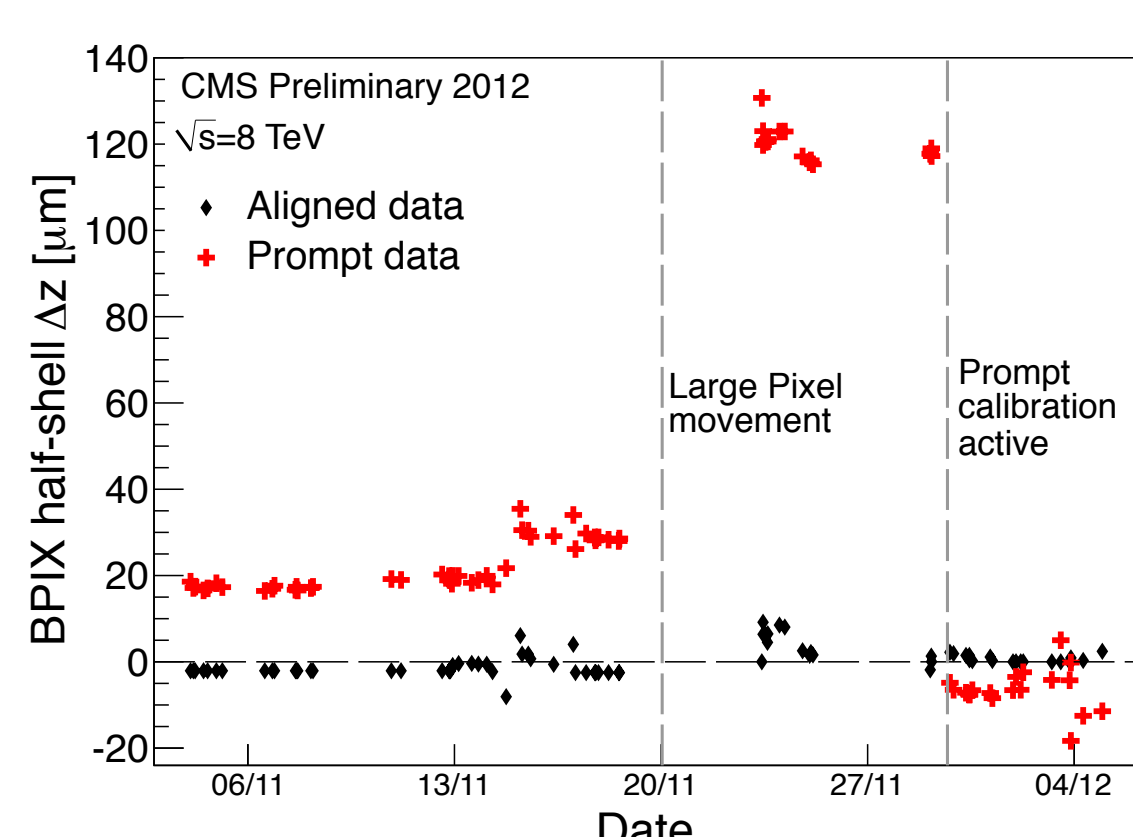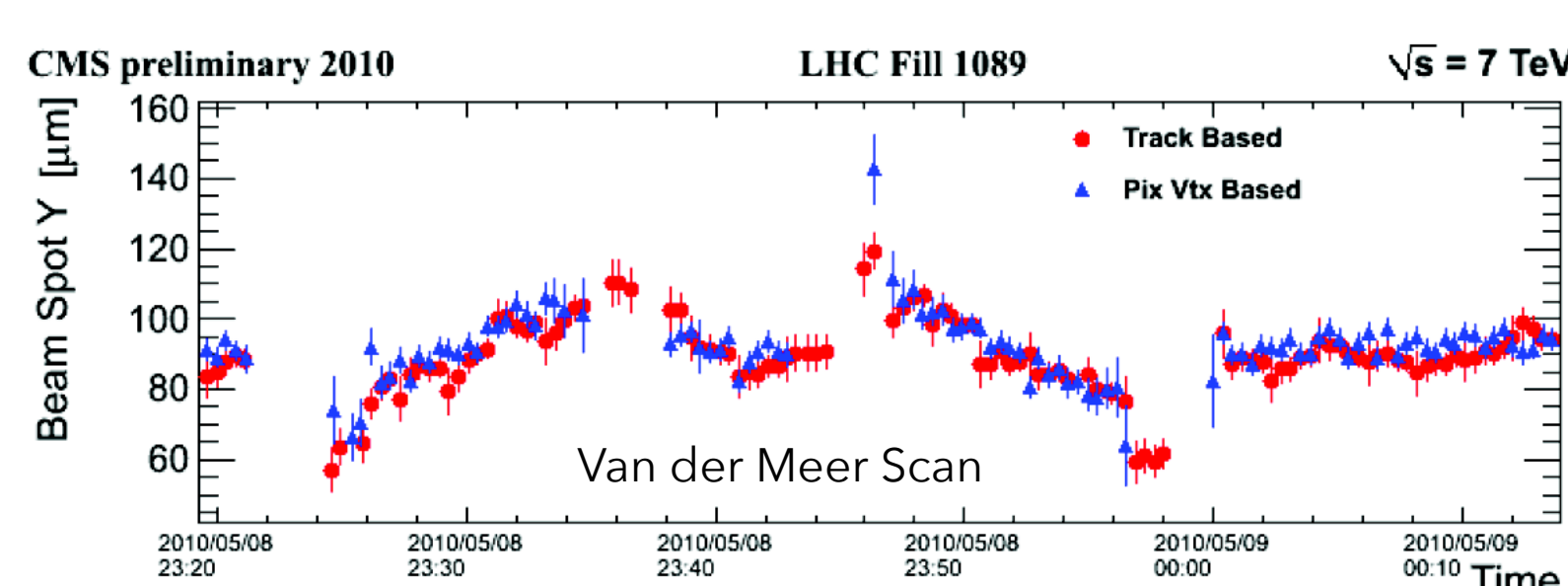- Interval of Validity (IoV) of the new payload

These metadata for the PCL workflows can be configured at run time in CMSSW via the Condition DB infrastructure.

Alignment & Calibration Workflows run in Prompt Calibration

Several calibration workflows are already running or being integrated in the Tier0 Automated Prompt Calibration system.

- **fit of the luminous region**: track beam-spot 3D position and width as a function of time → up to 1 fit every 1LS (=23s)
- identification of **SiStrip problematic channels** → up to 1 set per run
- determination of **SiStrip gains** → 1 set per run
- track based **alignment of SiPixel large structures** → up to 1 geometry per run

In addition: monitoring of **ECAL crystal radiation damage** running on dedicated resources in online → 1 measurement every ~30 minutes

CMS preliminary 2010   LHC Fill 1089   √s = 7 TeV

Beam Spot Y [μm]

Track Based
Pix Vtx Based

Van der Meer Scan

CMS Preliminary 2012 √s=8 TeV

BPX half-shell Δz [μm]

Aligned data
Prompt data

Large Pixel movement   Prompt calibration active

Monitoring Applications

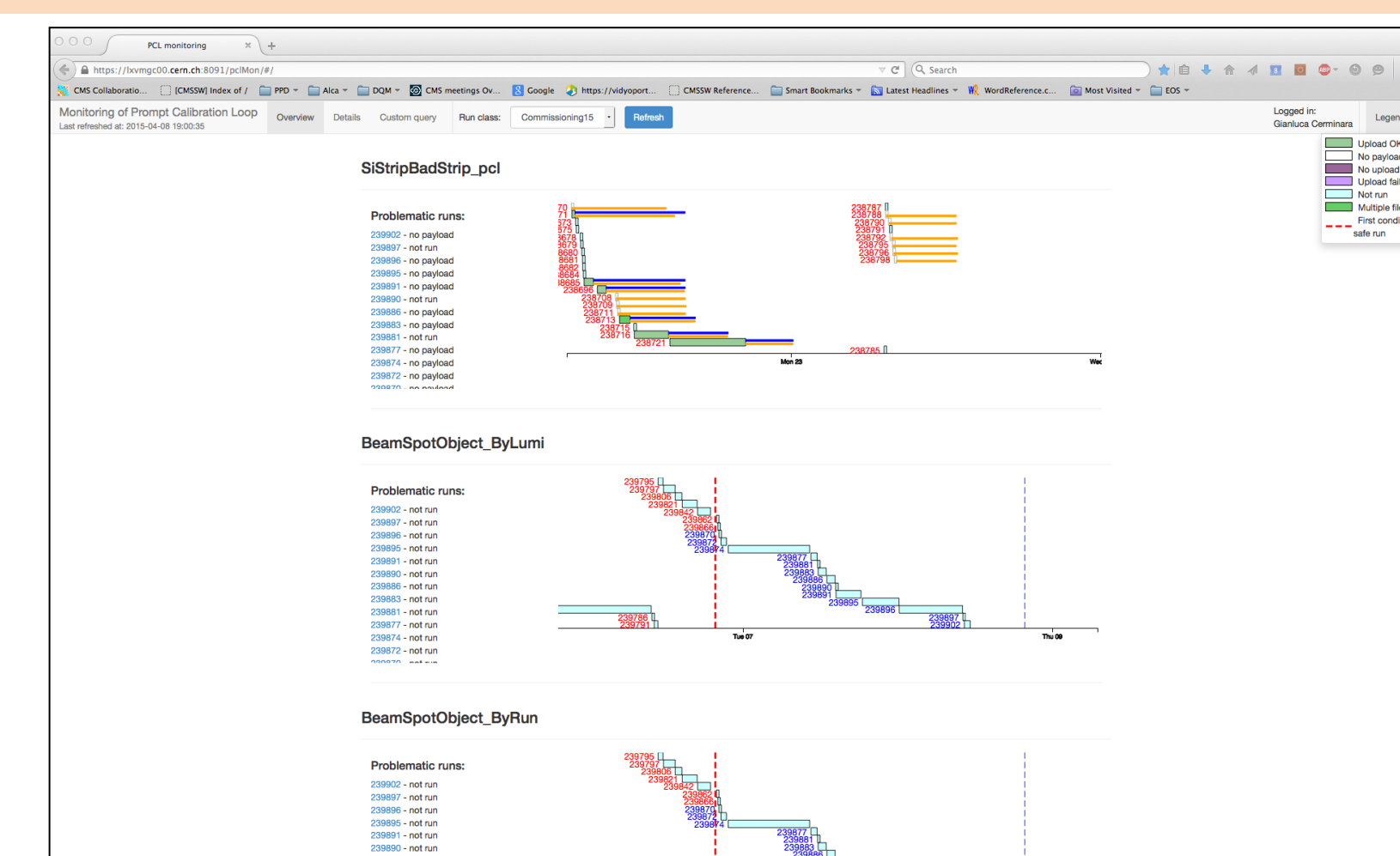Monitoring and alarm system crucial for reliable unattended operation.

pclMon

Web-based application aggregating information from several sources:
- CMS WBM and Run Registry
- Tier0 monitoring API (Tier0DAS)
- Offline DropBox logs
- ORACLE Condition DB

Crucial features:
- automatic e-mail alerts
- latency monitoring
- error acknowledgment by operator

Data Quality Monitoring

Monitoring of Physics Performance integrated in standard DQM running both during constant derivation and during reconstruction.

Outlook

Prompt Calibration Workflows successfully utilized already in RunI.
New workflows being integrated and/or commissioned for RunII.
Revamped monitoring applications for increased reliability.

Prompt Calibration: key asset to deliver analysis-ready data with short turnaround.

Gianluca Cerminara - CERN (gianluca.cerminara@cern.ch)
CHEP2015 - Okinawa - Japan