CERN

NexTGen

Next Generation Triggers

Task 3.4: Optimal Calibration for CMS High-Level Trigger

M. Musich, J. Prendi, T. Tomei, M. Zarucki

CERN | NexTGen Next Generation Triggers

The Real-time Reconstruction Revolution (R3)

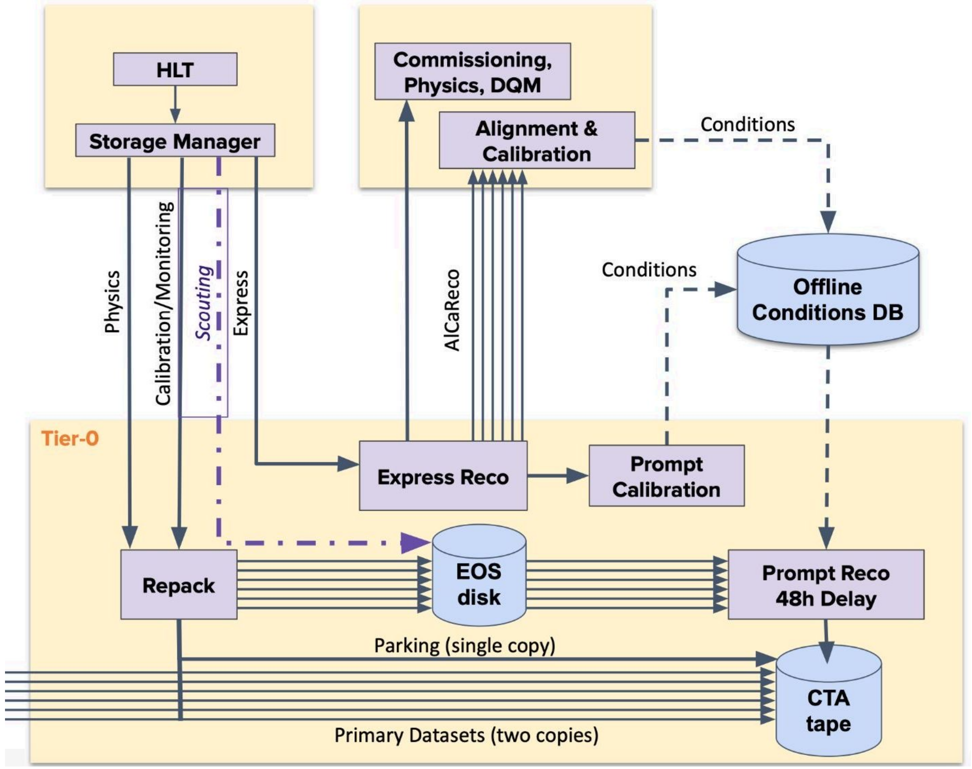
- **Overcome the two main limitations of the High-Level Trigger (HLT)**
  - Online reconstruction quality is limited by the HLT farm processing, as complex algorithms can be run only on a fraction of the events.
  - The HLT output rate is limited by the storage capacity and processing power of the offline computing infrastructure.
- **What if we could ...**
  - Have offline-like quality calibrations and reconstruction at the HLT?
  - Store all events in nano-AOD format?

The Real-time Reconstruction Revolution (R3)

- Overcome the two main limitations of the High-Level Trigger (HLT)
  - Online reconstruction quality is limited by the HLT farm processing, as complex algorithms can be run only on a fraction of the events.
  - The HLT output rate is limited by the storage capacity and processing power of the offline computing infrastructure.
- **What if we could ...**
  - **Have offline-like quality calibrations and reconstruction at the HLT?**
  - Store all events in nano-AOD format?

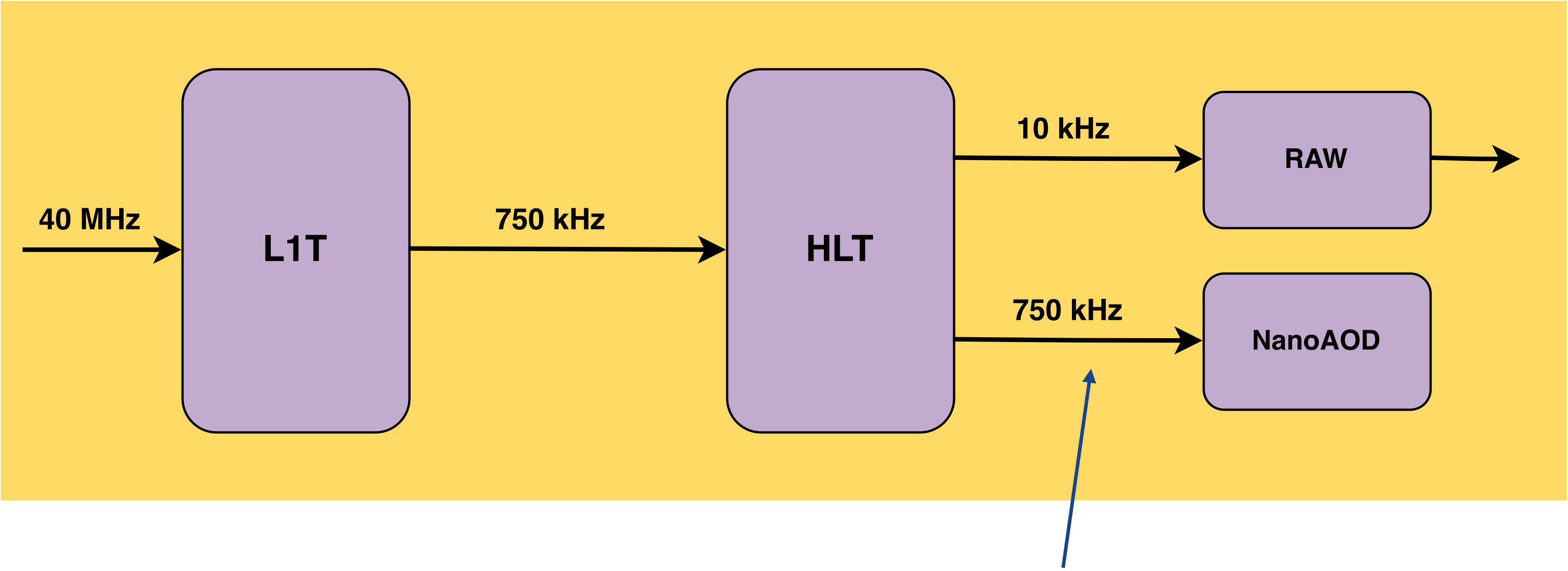
Alignment & Calibrations for Prompt Reconstruction

- **Current standard: prompt reconstruction ready ~48 hours after the data collection.**
  - To first approximation, all calibration/alignment constants (hereafter: conditions) are derived from the so-called **Prompt Calibration Loop (PCL).**
  - Some other approaches exist:
    - O2O-like, automation framework (see backup).
  - Constants derived from a given run are used in the reconstruction of that same run (with some exceptions).
  - The derivation of the constants come from the analysis of the Express Streams, a set of O(100) Hz raw data streams that feed the calibration workflows.
- **Prompt reconstruction is suitable for physics results.**
  - The problem would be mostly solved if we brought prompt reconstruction to the HLT.



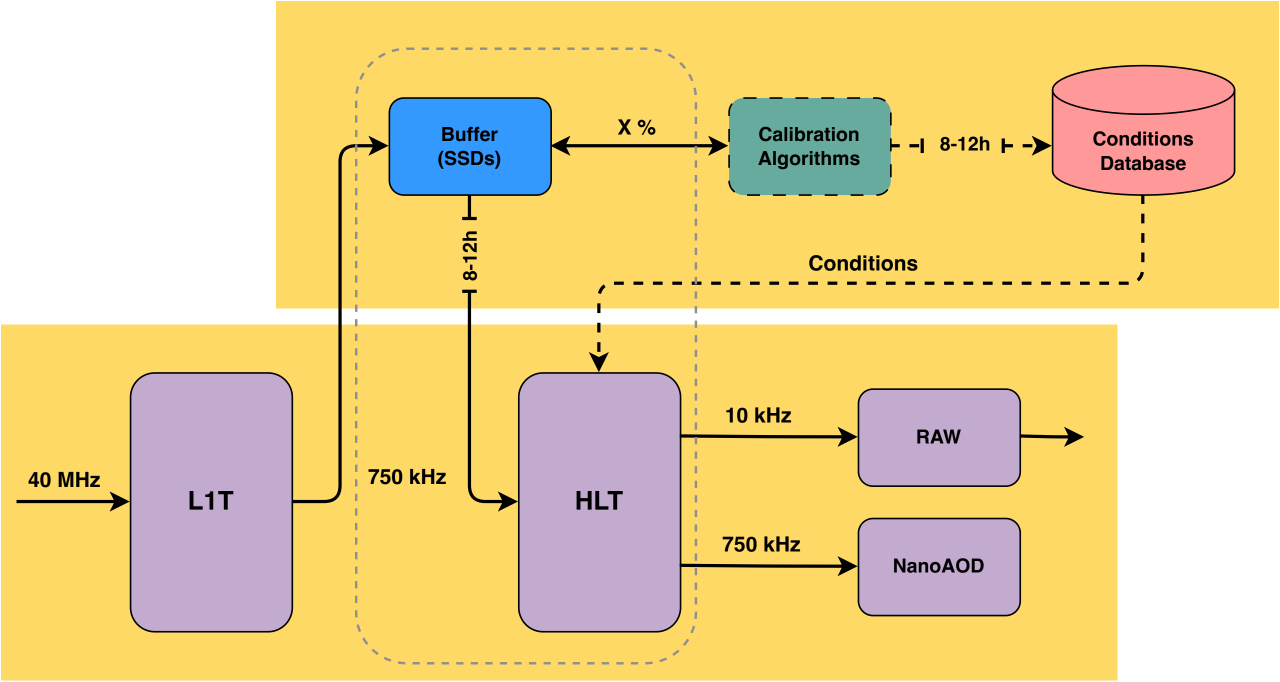
The Prompt Calibration Loop as it is today

Introduction to Optimal Calibrations for HLT



How do we get these data to have the highest possible quality?

R3 Optimal Calibrations



Conceptual Design of R3 Optimal Calibrations

- **Design accelerated calibration workflows to achieve at HLT the same accuracy as the offline reconstruction:**
  - Optimize the calibration process for the CMS detectors.
  - Introduce **data buffering** online, similar to the LHCb HLT1+HLT2 approach.
- **Synergy with Run-3 operations:**
  - Deploy a **prototype** applied to the HLT Scouting data during the last year of **Run-3**.
  - Rethink the hardware and software infrastructure for the calibration workflow.

Key Questions to Answer

- **Which conditions make sense to include in the NGT workflow (especially for the Run 3 prototype)?**
  - Needs a full survey of subdetector conditions in CMS (see preliminary results next slide)
    - Partial results already highlighted a set of candidate conditions.
  - Critical for reconstruction quality ←→ correlate with a condition's update frequency.
  - NGT aims at HL-LHC: think about current subdetector plans and about future subdetectors (HGCAL, MTD).
- **What is the effect of improved conditions in online reconstruction?**
  - Needs a test harness to run HLT with different conditions and compare.
  - Explore the "needed dataset size X needed computing power X quality of conditions space".
  - Preliminary code in place, work-in-progress – not discussed further here.
- **What is the framework for deriving and deploying the conditions online?**
  - Connected directly to the new HLT architecture as distributed application (Task 3.2).
  - Joint discussion with DAQ group.
- **Most critical: <u>deployment of a prototype by late 2025/early 2026</u> (last year of Run 3).**
- **<u>Keep focus on this reduced scope.</u>**

Survey to the Physics Groups

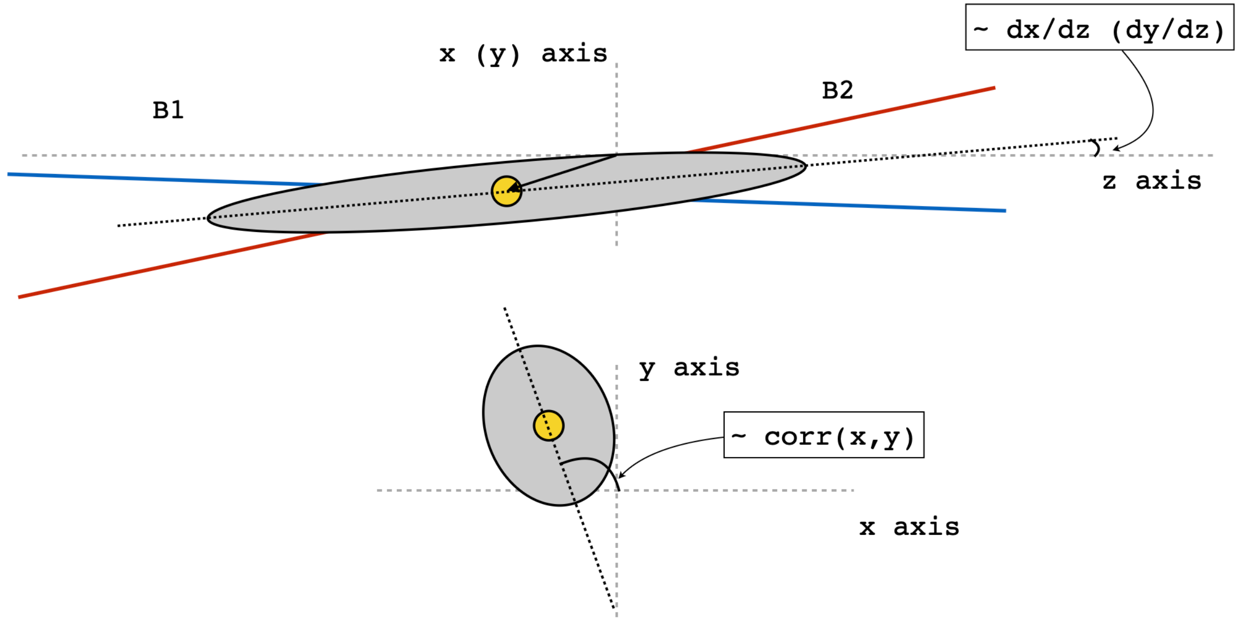
HLT-GT

	Record	Label	Tag	Type of workflow (O2O, PCL, popcon...?)	Critical?	Link to Configuration	How often is it updated (roughly	Consumed at L1 repack (fi	Consumed at HL
2	AlCaRecoTriggerBitsRcd	-	AlCaRecoHLTpaths8e29_1e31_v15_hlt	Manual	yes (for AlCaReco workflows	N/A	1-3 / year		
3	AlCaRecoTriggerBitsRcd	JetMETDQMTrigger	AlcaRecoTriggerBits_JetMET_DQM_v0_hlt	technical tags to encode in DB list of triggers used	no	link			
4	AlCaRecoTriggerBitsRcd	MuonDQMTrigger	AlCaRecoTriggerBits_MuonDQM_v1_hlt	to filter events for JME-Muon DQM	no				
5	BTagTrackProbability2DRcd	-	TrackProbabilityCalibration_PDF2D_hlt						Yes
6	BTagTrackProbability3DRcd	-	TrackProbabilityCalibration_PDF3D_hlt						Yes
7	BeamSpotObjectsRcd	-	BeamSpotObjects_PCL_byRun_v0_hlt	PCL, but only consumed in case of failure of the "BeamSpotOnline" workflow	medium	ALCARECOPromptCalibProdBeamSpotHP_cff.py	Run based condition populated by PCL		Yes
8	BeamSpotOnlineHLTObjectsRcd	-	BeamSpotOnlineHLT	DQM Clients, O2O like	yes	beamhlt_dqm_sourceclient-live_cfg.py	Every 5 LS during SB		Yes
9	BeamSpotOnlineLegacyObjectsRcd	-	BeamSpotOnlineLegacy	DQM Clients, O2O like	yes	beam_dqm_sourceclient-live_cfg.py	Every 5 LS during SB		Yes
10	CSCAlignmentErrorExtendedRcd	-	CSCAlignmentErrorExtended_6x6_hlt	Manual (Muon Alignment are responsible)				Yes	Yes
11	CSCAlignmentRcd	-	CSCAlignment_2009_v1_hlt	Manual (Muon Alignment are responsible)				Yes	Yes
12	CSCBadChambersRcd	-	CSCBadChambers_2009_v1_hlt	Manual	no (irrelevant - MC only)				Yes
13	CSCBadStripsRcd	-	CSCBadStrips_hlt	Manual	no (irrelevant)				
14	CSCBadWiresRcd	-	CSCBadWires_hlt	Manual	no (irrelevant)				
15	CSCChamberIndexRcd	-	CSCChamberIndex_hlt	Manual	no (irrelevant)				
16	CSCChamberMapRcd	-	CSCChamberMap_hlt	Manual	no (irrelevant - MC only)			Yes	Yes
17	CSCChamberTimeCorrectionsRcd	-	CSCChamberTimeCorrections_hlt	Manual	yes		once per year		Yes
18	CSCCrateMapRcd	-	CSCCrateMap_hlt	Manual	yes		~never	Yes	Yes
19	CSCDBChipSpeedCorrectionRcd	-	CSCDBChipSpeedCorrection_hlt	Manual	yes		rarely		Yes
20	CSCDBCrosstalkRcd	-	CSCDBCrosstalk_v2_hlt	Manual	yes		once per year		Yes
21	CSCDBGainsRcd	-	CSCDBGains_hlt	Manual	yes		once per year		Yes
22	CSCDBGasGainCorrectionRcd	-	CSCGasGainCorrections_v2_hlt	Manual	no (irrelevant - run1 only)				
23	CSCDBL1TPParametersRcd	-	CSCDBL1TPParameters_hlt	Manual	not sure - but probably not				
24	CSCDBNoiseMatrixRcd	-	CSCDBNoiseMatrix_hlt	Manual	yes		once per year		Yes
25	CSCDBPedestalsRcd	-	CSCDBPedestals_hlt	Manual	no (irrelevant - MC only)				Yes
26	CSCDDUMapRcd	-	CSCDDUMap_hlt	Manual	no (irrelevant)				
27	CSCL1TPParametersRcd	-	CSCL1TPParameters_2010_hlt	Manual	not sure - but probably not				
28	CSCRecoDigiParametersRcd	-	CSCRECODIGI_Geometry_v2_hlt	Manual	yes		~never	Yes	Yes
29	CSCRecoGeometryRcd	-	CSCRECO_Geometry_v2_hlt	Manual	yes		~never	Yes	Yes
30	CTPPSOpticsRcd	-	PPSOpticalFunctions_hlt_v10	Manual	Y (for Reco)		1-2 / year		Yes
31	CTPPSPixelAnalysisMaskRcd	-	CTPPSPixelAnalysisMask_Run3_v1_hlt	Manual	Y (for Reco)		in case of problems with HW		Yes
32	CTPPSPixelDAQMappingRcd	-	CTPPSPixelDAQMapping_Run3_v1_hlt	Manual	Y (for Reco)		in case of problems with HW		Yes
33	CTPPSPixelGainCalibrationsRcd	-	CTPPSPixelGainCalibrations_Run3_v1_hlt	Manual	Y (for Reco)		1-3 / year		Yes
34	DTAlignmentErrorExtendedRcd	-	DTAlignmentErrorExtended_6x6_hlt					Yes	Yes
35	DTAlignmentRcd	-	DTAlignment_2009_v1_hlt					Yes	Yes
36	DTCCBConfigRcd	-	DTCCBConfig_V06_hlt	Not used (AFAIK)	no			Yes	
37	DTDeadFlagRcd	-	DT_dead_cosmic2009_V01_hlt	Manual	yes		Not in the foreseeable future		

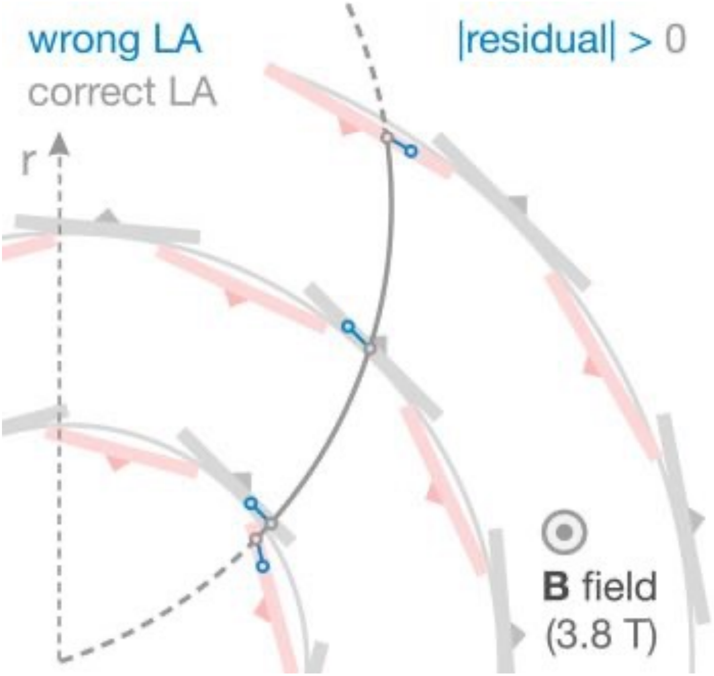
***Nota bene*: there are 300+ conditions in the set currently used by HLT.**

Candidate Conditions for NGT (1)

- **Beamspot (luminous region, where the two beams interact)**
  - Used to bootstrap tracking.
  - Already updated close to in real time, with a latency of about 2 minutes.
- **Fit tracks and vertices (T&V) for ellipsoid position, displacement and rotation.**
- **Separated workflows for online and offline**
  - HLT T&V in the HLT farm
  - Pixel T&V in the DQM farm
  - Offline T&V with regular Express (Legacy)
  - Offline T&V with dedicated data stream (High Precision)
- **HLT uses one of the online algorithms (with arbitration).**



Candidate Conditions for NGT (2)

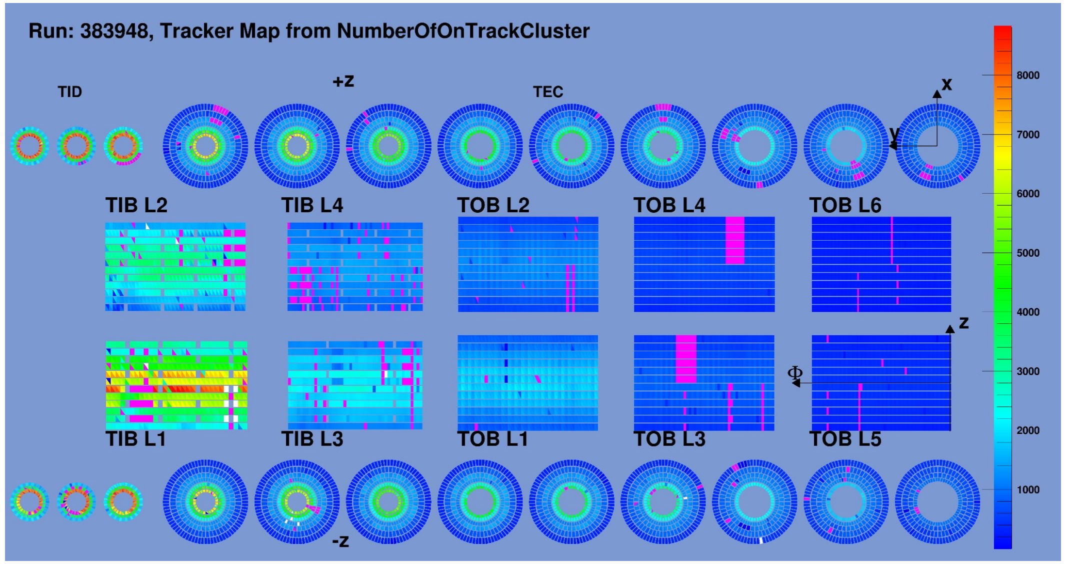


LA miscalibrations and alignment intertwined

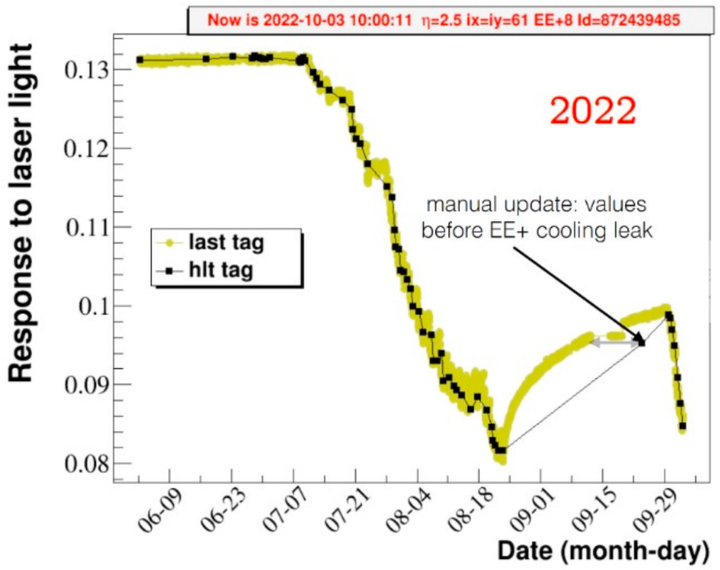
- **Silicon Pixel Alignment**
  - Tracker Alignment is coupled with the Pixel cluster position estimation (CPE) conditions, because it “overcorrects” for Lorentz Angle (LA) miscalibrations.
  - In offline reconstruction: optimize the pixel reconstructed hit position bias via the “high-granularity PCL”.
    - Automatic procedure to align the pixel detector at ladder / panel level (LA effect opposite on adjacent ladders).
- **The Offline PCL conditions produced for Prompt Reconstruction cannot be used directly for HLT,**
  - HLT CPE algorithm is different (“Pixel CPE Fast”)
  - The CPE-induced bias is CPE-algorithm dependent.
- **Plan from Tracker Alignment group:**
  - Parallel version of the offline PCL using HLT tracks.
  - Refitting with the HLT CPE algorithm.
  - Pick alignment to manually upload.
- **Could build upon this plan for NGT.**

Candidate Conditions for NGT (3)

- **Silicon Strip Bad Components**
  - Needs dedicated unbiased data stream to detect new dead or hot components.
  - Runs in the PCL, allows for dynamic monitoring.
  - Tracking can use this information to know if a missing hit is really missing or if it's sitting on a "inactive" detector.
- **Only static masks are used at HLT.**
  - Updated manually every once in a while, and there's no automatic book-keeping.
- **Nota bene: we will not have a Silicon Strip system in Phase-2, but still useful for the demonstrator + the big picture is still valid for NGT.**

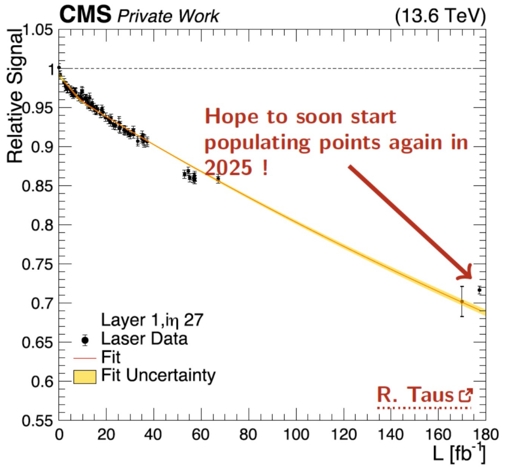


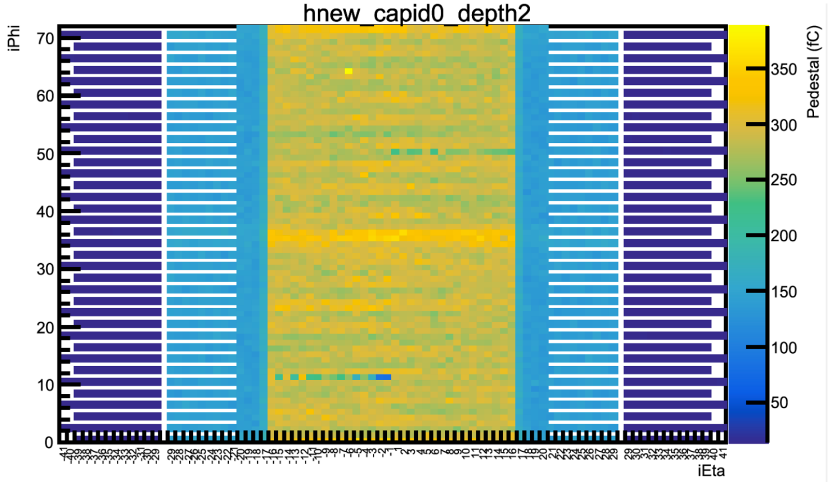
Candidate Conditions for NGT (4)



- **ECAL transparency**
  - Laser monitoring with dedicated framework in place.
  - Conditions updated with 40 min granularity (time for the laser to cycle through all crystals) for offline reconstruction + linear interpolation for LS granularity.
- **Available at HLT per fill only.**

Candidate Conditions for NGT (5)





- **HCAL Gains (HE and HF)**
  - Correction for radiation damage of active material.
  - Based on laser data in orbit gap + parametrization of exponential decay.
  - Frequency: weekly (every 2/fb) via automation framework (planned).
- **HCAL Pedestals**
  - Electronics noise measurements offset to avoid energy measurement bias
  - based on orbit gap data during collisions.
  - Frequency: weekly via automation framework.

Integrating the Calibration in the Online Environment

- **Buffer all the data of a run**
  - Approximate target for Run 4: 500 kHz × 12 hours × 6.1 MB × 1.5 safety ~ **200 PB buffer.**
    - More detailed calculations in progress, to be done for the report.
    - LHCb experience shows that the reading speed goes down when the buffer is close to saturation → safety.
  - Some comparable timescales:
    - PCL average turnaround time: 8 hours.
    - Tier-0 starts the Prompt Reconstruction in at most 48 hours.
    - Have also to consider the HL-LHC duty cycle.
  - Comparable buffer sizes:
    - LHCb buffer size: 30 PB, see 2305.10515
    - CMS DAQ-HLT TDR 1-day buffer: 3.3 PB, see https://cds.cern.ch/record/2759072/

Run the optimal calibration

  - In which/how many machines exactly?

- **How do we make available these calibrations to the online reconstruction?**
  - Regular database? HDF5 files? Binary blobs over HTTP?
- **Reconstruct all the physics objects (Tasks 3.1.1 and 3.1.2)**
  - No filtering – save all the events in reduced data formats (Task 3.3).
  - If it were to save 500 kHz of Scouting-like data (25 kB/event): additional **12.5 GB/s output.**
  - For comparison: the TDR throughput (concurrent recording + transfer) was 51 GB/s.

Feasibility Check of NGT for the High-Level Trigger

- **Assumptions:**
  - 200 PB buffer
    - Enterprise SSDs are rated for "3 drive writes per day" for five years.
    - Continuous operation → 24/3 = 8 hours. ✓
  - The HLT farm ITSELF will be running NGT as well as the regular HLT.
    - Order of 400 4U nodes
    - Proposal: have the storage IN the farm: 200 PB / 400 nodes = 500 TB / node.
    - 60 TB disks exist today, and 120 TB will be available for Run 4 → 5 disks / node. ✓
    - Alternative: have a central storage (like we have today) → okay, but larger dataflow.
- **We conclude that, in principle, there seem to be no technical showstoppers.**
  - We are in the process of discussing the architectural layout of the DAQ system.

Minimum Minimorum NGT Demonstrator

1. **The HLT runs as normal**
2. **Add to the menu a copy of the Scouting input, prescaled by (example) O(100): from ~30 kHz to ~300 Hz**
3. **Send this prescaled copy of the Scouting input to a dedicated area at P5**
    1. Lustre? Dedicated machine(s)?
       Back-of-the-envelope math: 300 Hz x 1.2 MB x 48 hours (for PCL) x 1.5 safety = 90 TB storage.
       So, one dedicated HLT farm node with 90–120 TB of storage should do it.
4. **Use the regular workflow to derive the candidate calibration**
    1. At P5 within 8 hours if we we manage to port the workflows.
    2. Using the PCL within 48 hours if we don't manage but we can afford the disk space.
5. **Re-run the Scouting paths on the buffered data**
6. **Compare the performance of the original vs re-run Scouting**

Conclusions and Outlook

- **Already done:**
  - Initial hirings (2 Fellows + 1 Doctoral Student)
    - Heavy training for the student on both HLT and alignment/calibration matters.
  - Survey the physics groups of CMS w.r.t candidate conditions for NGT demonstrator.
  - Initial set of candidate conditions for NGT identified.
    - Beamspot, SiPixel Alignment, SiStrip Bad Components, ECAL laser transparency, HCAL Pedestals.
  - Initial version of the test harness for evaluation of impact of different conditions.
- **End of 2024:**
  - Continuing polling CMS for information on Phase-2 detectors.
  - Report illustrating the current calibration workflows + initial evaluation of impact (contractual milestones).
  - Initial discussion with DAQ group w.r.t NGT requirements in the system design for Phase-2 **AND** prototype integration in the Run-3 system.
- **2025 plans**
  - Further exploration of impact of different conditions for the NGT prototype and Phase-2.
  - Design and construction of the NGT prototype.

CERN

NexTGen

Next Generation Triggers

Backup

Quick Jargon Explanation

- **O2O: stands for "Online to Offline", a framework which synchronises online configurations of the detectors (e.g. voltage settings, channel masks, …) into the offline conditions database.**
    - In CMS, online data (that is, data written by processes using the running detector) is stored in subdetector-specific schemas on the online database called OMDS (Online Master Data Storage). Offline calibration and alignment data are (...) persistently (stored as) C++ objects in a relational database such as Oracle (...). Because the format of the online and offline data is different it necessitates the O2O transform process for this kind of data."
- **Automation framework: a finite state machine implemented through Jenkins, Influxdb and Grafana for monitoring. Deployed with the Openshift instance provided by CERN-IT. Also provides a small python package to provide the interface between the CMS ecosystem, the user jobs and the framework.**
    - See S. Pigazzini's talk in ACAT 2022.