Available on CMS information server **CMS CR -2022/180**

The Compact Muon Solenoid Experiment

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

13 October 2022 (v3, 18 October 2022)

Trigger-less readout and unbiased data quality monitoring of the CMS Drift Tubes muon detector

Matteo Migliorini, Jacopo Pazzini, Andrea Triossi, Marco Zanetti and Alberto Zucchetta on behalf of the CMS Collaboration

Abstract

The CMS experiment 40MHz data scouting project is aimed at intercepting the data produced at the level of the detectors' front-end without the filters induced by hardware-based triggers. A first implementation is realized by the trigger-less reading and processing of a fraction of the Drift Tube (DT) muon detector, equipped with a preliminary version of the so-called Phase-2 Upgrade on-detector electronics boards. The data are transferred via high-speed optical links to back-end boards independently from the central experiment data acquisition (DAQ), permitting real-time detector status monitoring via receiving all the signals produced at the front-end level, and providing an unbiased estimate of the CMS DT hit-rate under various data-taking conditions.

Presented at *TWEPP2022 Topical Workshop of Electronics for Particle Physics*

Prepared for submission to JINST

Topical Workshop on Electronics for Particle Physics
19-23 September, 2022
Bergen, Norway

Trigger-less readout and unbiased data quality monitoring of the CMS Drift Tubes muon detector

---

M. Migliorini,[a,b] J. Pazzini,[a,b,1] A. Triossi,[a,b] M. Zanetti,[a,b] and A. Zucchetta[b] on behalf of the CMS collaboration

[a] *Department of Physics and Astronomy "Galileo Galilei", Padova University, Via Marzolo 8, 35131 Padova, Italy*

[b] *National Institute for Nuclear Physics, Padova Division, Via Marzolo 8, 35131 Padova, Italy*

E-mail: `jacopo.pazzini@unipd.it`

Abstract: The CMS experiment 40MHz data scouting project is aimed at intercepting the data produced at the level of the detectors' front-end without the filters induced by hardware-based triggers. A first implementation is realized by the trigger-less reading and processing of a fraction of the Drift Tube (DT) muon detector, equipped with a preliminary version of the so-called Phase-2 Upgrade on-detector electronics boards. The data are transferred via high-speed optical links to back-end boards independently from the central experiment data acquisition (DAQ), permitting real-time detector status monitoring via receiving all the signals produced at the front-end level, and providing an unbiased estimate of the CMS DT hit-rate under various data-taking conditions.

Keywords: Data acquisition concepts, Trigger concepts and systems (hardware and software), Data processing methods

---

[1] Corresponding author.

Contents

1 Trigger and detector data scouting at CMS

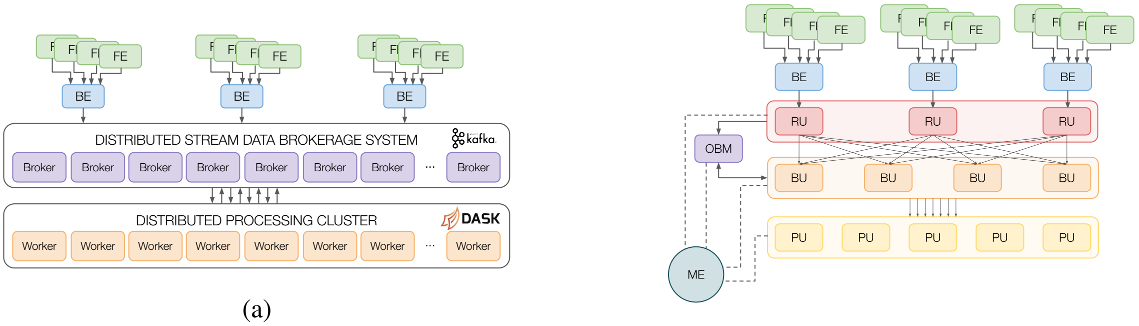
The High-Luminosity phase of the Large Hadron Collider (HL-LHC), currently planned to be operational starting in 2029, will see a sizable increase of the instantaneous luminosity delivered to the experiments, $7.5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{sec}^{-1}$, resulting in an expected pileup of up to 200 inelastic collisions per event. Overall, the HL-LHC is expected to deliver an integrated luminosity of $3\,\mathrm{ab}^{-1}$ over about twelve years of operation, totaling about ten times the integrated luminosity expected to be collected after the first twelve years of the LHC operations [1]. The Compact Muon Solenoid experiment (CMS) [2] will undergo a major campaign of upgrades (referred to as Phase-2 Upgrade) to cope with the increased luminosity and pileup conditions and to improve the detector performances. The CMS Phase-2 Upgrade will include: a full replacement of the tracking system, along with a track-finder for the Level-1 trigger [3]; the installation of a new timing detector [4] and new high-granularity calorimeter in the endcap region [5]; the replacement of large fractions of the front-end (FE) detector electronics to improve radiation hardness and readout bandwidth; and a new Level-1 trigger based on state-of-the-art FPGAs (field programmable gate array electronics) capable of accepting events at a rate of 750 kHz [6]. While CMS will continue to employ a two stages trigger system, the upgrades in FE electronics and trigger boards will make possible to reconstruct higher-quality physics objects already at the Level-1 hardware trigger level, resulting in resolutions close to the one achieved offline. The CMS 40MHz Level-1 trigger scouting project [7] proposes to use spare optical outputs of Level-1 trigger boards, processing data quasi-online on a set of computing resources without the latency limitations of the Level-1. Several physics studies are expected to benefit from the trigger scouting system data. Notable examples are processes with low trigger efficiencies due to the Level-1 trigger thresholds and rare processes, both of them could benefit from the analysis of the full available statistics. Additional examples include all the topologies where a non-standard reconstruction is required, such as the searches for long-lived particles with lifetimes spanning multiple bunch crossings, as well as appearing- or disappearing-tracks. The upgrades in the detector FE electronics will also allow CMS to extend the scouting project approach, intercepting data as close as possible to the detector FEs, depending on the expected throughput. The 40MHz *data scouting* is an additional CMS project investigating the possibility of collecting and processing the data stream at the level of the detectors' FE before any trigger stage, performing the online reconstruction at full resolution. An additional advantage of the data scouting project is the continuous monitoring and diagnostic of the experiment conditions, based on the totality of the signals produced by the detector without any suppression induced by the trigger chain.

2 Data scouting implementation

The first implementation of the 40MHz data scouting has been deployed in CMS in 2022. The Drift Tube (DT) detector sub-system [8] was chosen for the first deployment, as four chambers (MB1 to MB4 of the DT sector 12 of wheel +2) have been instrumented with Phase-2 On-Board DT readout boards (OBDT) [9], performing the Time-to-Digital Conversions of the DT hits time in FPGA with nanosecond resolution. In total, the DT Phase-2 Upgrade demonstrator is equipped with 13 OBDT boards, and 3120 individual channels. The streams of hits produced by the OBDTs are shipped via high-speed optical links using the GBT [10] protocol to back-end (BE) devices, where they are collected and processed in parallel to the standard CMS data acquisition (DAQ) and trigger chains, to enable an event-by-event comparison of the legacy and the upgrade demonstrator systems. Two Xilinx KCU1500 development boards, equipped with a Kintex UltraScale XCKU115-2FLVB2104E FPGA, are used as BE devices. Each BE board is equipped with two QSFP transceivers and is capable of receiving up to 8 OBDT input links. The BE boards are installed in a dedicated Dell PowerEdge R730 server using the PCIe Gen3 x8 interfaces on a bifurcated x16 edge connector. The FPGA firmware implements the links deserialization with the GBTx-FPGA protocol [11] and the gearboxing of all links into a single data stream. The data transfers to memory are performed through a Direct Memory Access (DMA) engine via the Advanced eXtensible Interface (AXI) stream protocol over the PCIe Gen 3 bus to avoid burdening the server CPU. The data stream is further continuously transferred to a computing farm where it is processed on-the-fly. Two alternative architectures for the transfer and online processing of data have been designed and implemented.

2.1 Distributed processing and messaging systems

A first architecture (see Figure 1a) is designed based on the usage of horizontally scalable distributed computing frameworks for both the brokerage and processing of the FE signals. Dask [12] is used as the distributed processing engine to schedule the data reconstruction workload over a set of worker nodes. Apache Kafka [13], a distributed event streaming platform based on a pub-sub messaging model, is used as the data brokerage system to efficiently serve data from the BE units to the Dask worker nodes. In each BE unit, an Apache Kafka producer process continuously polls the server memory and publishes the new data to a remote topic, distributed in several partitions over a cluster of remote broker nodes. The cluster of Dask worker nodes acts as a set of consumers, continuously reading from the Apache Kafka partitions to collect raw data and create distributed data structures. The producer/consumer architecture entirely decouples the writing and reading processes, thus permitting the Dask worker nodes to read and process data independently from the BE readout of the FE data, while using the Apache Kafka brokers as a buffering stage. This processing model is



(a)

(b)

Figure 1: Schematic representations of (a) the distributed processing data scouting architecture and (b) the multistage data scouting architecture. FE: front-end devices; BE: back-end devices; RU: readout units; BU: builder units; PU: processing units; OBM: orbit manager; ME: master entity process.

designed to accommodate the asynchronous nature of the 40MHz data scouting readout and to avoid inducing back-pressure to the earlier stages of the readout chain. Simple aggregated quantities, useful for the monitoring of the detector, are computed directly by exploiting the data-analytic functionalities of the distributed computing framework. An event-building stage is performed in parallel by subdividing the data batches by the LHC Orbit ID (where a LHC orbit corresponds to 3564 divisions of 25 ns each). Both the distributed data brokerage and the processing systems have been tested to efficiently scale their performance with the number of computing nodes, as discussed in detail in [14].

2.2 Multistage architecture

A multi-stage data aggregation architecture (see Figure 1b) has been developed as an alternative to the distributed cluster-based system. This design is focused on merging data fragments collected by several BE devices to create self-consistent data structures (events), before serving those to the processing units. As the 40MHz data scouting system is continuously and asynchronously collecting data from several FE devices, no event-like structure can be identified before the processing of the collected signals. The LHC orbit identifier is thus used as the main event “key”. Several Readout Units (RUs), deployed directly in the servers hosting the BE boards, act as in-memory key-value stores by temporarily caching the collected DT hits into arrays indexed by LHC Orbit ID. A single Orbit Manager process (OBM) performs the bookkeeping of the LHC Orbits IDs available on the entire set of allocated RUs. This is achieved by the RUs by sending a new message to the OBM every time a new Index is created into the RUs cache. Another set of processes is in charge of merging all hits pertaining to a given LHC orbit, stored in the cache of multiple RUs, thus creating a single collection consistent with a given LHC Orbit ID. Each such process, named Builder Unit (BU), is a multi-threaded process interfacing directly with both the Orbit ID list managed by the OBM, and with the RUs to fetch the cache records. Upon completion of the event-building phase, the BUs push the aggregated sets of hits to a number of Processing Units (PUs). The individual PUs are finally the processes performing the analysis of the consistent set of hits, reconstructing all the features corresponding to the passage of the muons through the detector. To control all the other processes (RU, BU, PU, OBM) a Master Entity (ME) process is instantiated. The ME exposes a programming interface acting as the entry point to interact with all the system's components, including the start/stop of data taking, and the collection of the metrics from all processes. In this architecture, all communications both inter- and intra-processes are handled via the ZeroMQ [15] messaging library.

3 Unbiased data quality Monitoring of the CMS DT

The scouting readout is completely opportunistic, as the data can be collected independently from the central DAQ and trigger chain of the experiment. During a scouting data acquisition run, the entire stream of DT hits produced by the FE devices is acquired and processed on-the-fly without any online filter. So far, the system has been already used to collect and process data under several LHC filling scheme conditions. Figure 2 shows the data collected during five minutes of the first LHC collisions of 2022 (LHC fill number 7920). The data highlights both the presence of the two bunches colliding at the LHC Interaction Point 5, as well as the effects of beam halo induced by the passage of non-colliding bunches. The nature of the data collected with this trigger-less DAQ offers the opportunity to develop data quality monitoring studies based on 100% of the hits produced by the detectors' FE, and completely unbiased thanks to the absence of dedicated trigger selection logic. For instance, the 40MHz data scouting readout will provide CMS with an alternative unbiased monitor of the instantaneous luminosity delivered by the accelerator based on the raw counts of all the detectors' signals. In addition, it can enable the in-situ measurement of the detector occupancy both using the data collected in collisions and outside LHC bunches collisions, as shown in Figure 3. The off-collision data can also be used to perform an unbiased estimation of the detectors' background noise rate, and its evolution with the integrated luminosity. Finally, the trigger-less data scouting opens up the possibility to investigate transient detector effects such as the aftermath of the accelerator dumps, as shown in Figure 4.

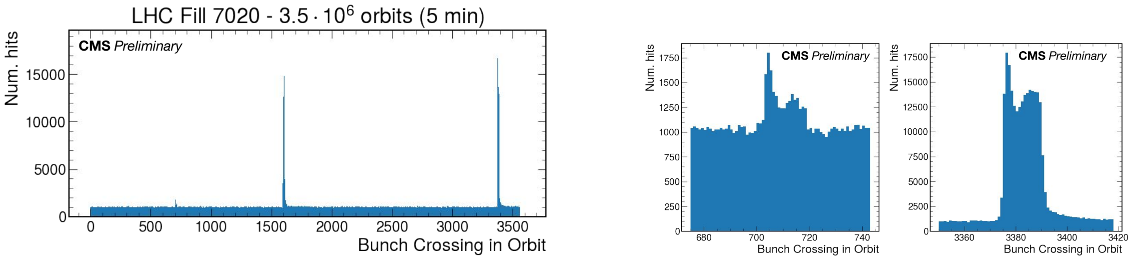


Figure 2: Left: projection of the hits collected during five minutes of the first LHC Fill (7920) in 2022, with three circulating bunches and two colliding at the LHC Interaction Point 5. The counts of all collected hits are projected in bins of bunch crossing within the LHC orbits, highlighting the filling scheme. Right: the hits clustering around colliding and a circulating bunches are shown. The patterns formed by the hits collection times are compatible with the drift time of muons crossing the DT detectors.

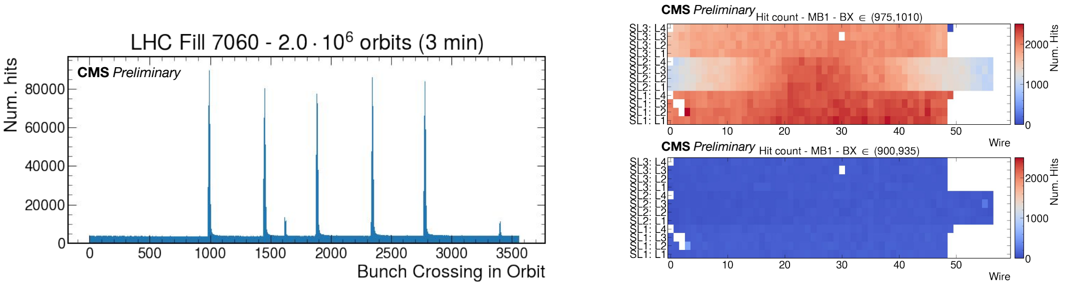


Figure 3: Left: projection of the hits collected during three minutes of the LHC Fill 7960. At the LHC Interaction Point 5, five trains of 12 contiguous bunches (spaced by 25 ns), and two single bunches were provided for collisions. Right: the occupancy of one DT chamber comparing a window of bunch crossings (BX) compatible with the collision of a train of LHC bunches (from BX 975 to 1010) and in a window with no collisions (from BX 900 to 935).

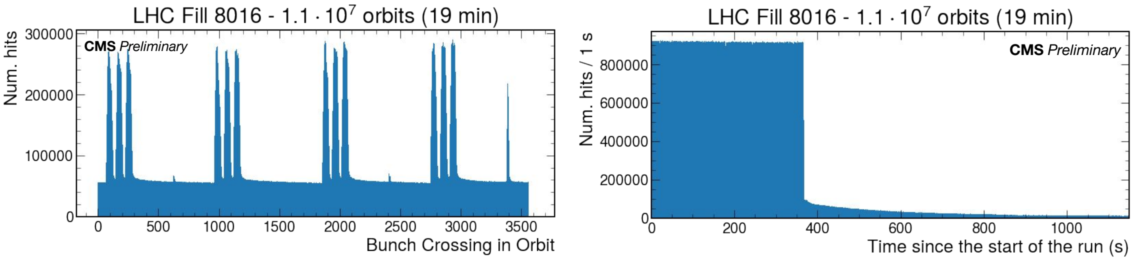


Figure 4: Left: projection of the hits collected during 19 minutes of the LHC Fill 8016. At the LHC Interaction Point 5, 13 trains of 48 contiguous bunches (spaced by 25 ns) and two single bunches were provided for collisions. Right: the number of hits collected per unit of time as a function of the time since the start of the scouting data acquisition run. The LHC Fill 8016 was dumped after about six minutes, and the dump aftermath is visible as a decaying trend of the number of hits collected with time, as the result of the experiment's material activation.

4 Summary

The first implementation of the 40MHz data scouting system of the CMS experiment has been presented. The system is designed to read raw data from the front-end of the detector, before the filters of the Level-1 trigger, and with the full detector resolution. Two alternative processing architectures have been implemented to analyze the data stream on-the-fly. The 40MHz data scouting system has been deployed in CMS during 2022, reading the data produced by the Phase-2 Upgrade demonstrator of the Drift Tube detector. This system is independent of the central CMS DAQ and trigger systems, not interfering with the operations of the experiment. A number of physics channels currently limited by low trigger efficiency, or the impossibility of performing an online selection within the bounds of bandwidth and latency, are expected to benefit significantly from the sophisticated on-the-fly data analyses. The data scouting system is also going to be an invaluable tool for the live monitoring of the detectors, benefiting from unbiased and unlimited statistics. A preliminary set of data quality monitoring studies have been presented with the first data collected during the early Run-3 operations of the LHC in 2022.

References

[1] I. Zurbano Fernandez et al. High-Luminosity Large Hadron Collider (HL-LHC): Technical design report. 10/2020, 12 2020. doi: 10.23731/CYRM-2020-0010.

[2] CMS Collaboration. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.

[3] CMS Collaboration. The Phase-2 Upgrade of the CMS Tracker. 6 2017. doi: 10.17181/CERN.QZ28.FLHW.

[4] CMS Collaboration. A MIP Timing Detector for the CMS Phase-2 Upgrade. 2019. URL `https://cds.cern.ch/record/2667167`.

[5] CMS Collaboration. The Phase-2 Upgrade of the CMS Endcap Calorimeter. 2017. doi: 10.17181/CERN.IV8M.1JY2.

[6] CMS Collaboration. The Phase-2 Upgrade of the CMS Level-1 Trigger. 2020. URL `https://cds.cern.ch/record/2714892`.

[7] Gilbert Badaro et al. 40 MHz Level-1 Trigger Scouting for CMS. *EPJ Web Conf.*, 245:01032, 2020. doi: 10.1051/epjconf/202024501032.

[8] CMS Collaboration. Performance of the CMS Drift Tube Chambers with Cosmic Rays. *JINST*, 5: T03015, 2010. doi: 10.1088/1748-0221/5/03/T03015.

[9] Javier Sastre Álvaro, Andrea Triossi, Antonio Bergnoli, Alessandro Griggio, and David Redondo Ferrero. The OBDT board: A prototype for the Phase 2 Drift Tubes on detector electronics. *PoS*, TWEPP2019:115, 2020. doi: 10.22323/1.370.0115.

[10] P. Moreira et al. The GBT Project. In *Topical Workshop on Electronics for Particle Physics*. CERN, 2009. doi: 10.5170/CERN-2009-006.342.

[11] S. Baron, J. P. Cachemiche, F. Marin, P. Moreira, and C. Soos. Implementing the GBT data transmission protocol in FPGAs. In *Topical Workshop on Electronics for Particle Physics*. CERN, 2009. doi: 10.5170/CERN-2009-006.631.

[12] Dask Development Team. Dask: Library for dynamic task scheduling. available at `https://dask.org`.

[13] Apache kafka. available at `https://kafka.apache.org/`.

[14] Matteo Migliorini, Jacopo Pazzini, Andrea Triossi, Marco Zanetti, and Alberto Zucchetta. A horizontally scalable online processing system for trigger-less data acquisition. *Nucl. Instrum. Meth. A*, 1036:166869, 2022. doi: 10.1016/j.nima.2022.166869.

[15] Zeromq, an open-source universal messaging library. available at `https://zeromq.org/`.