Available on CMS information server **CMS CR -2024/242**

The Compact Muon Solenoid Experiment

Conference Report

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

30 September 2024 (v4, 08 October 2024)

The CMS Level-1 Trigger Data Scouting system for the HL-LHC upgrade

Emilio Meschi for the CMS Collaboration

Abstract

To fully exploit the extended capability of its upgraded L1 trigger at the High-Luminosity LHC, CMS is pioneering a novel L1 Data Scouting (L1DS) system, capable of acquiring and processing the quasi-offline-quality trigger primitives produced by the upgraded L1 at the accelerator bunch-crossing rate of 40 MHz. The goal of the system is to give full access to potential physics signatures otherwise constrained by the L1 latency and accept rate limitations, or whose selection strategy diverges from that of the standard CMS physics program. To validate the concept and provide a development platform for the firmware and software required for the final system, a demonstrator system was assembled to operate with the current L1 trigger in the LHC Run 3. We discuss the L1DS demonstrator system architecture, performance, and some preliminary results, as well as the design of the final L1DS system and a summary of ongoing studies of its potential and competitiveness in selected physics channels.

Presented at *ICHEP2024 42nd International Conference on High Energy Physics*

PROCEEDINGS OF SCIENCE

The CMS Level-1 Trigger Data Scouting system for the HL-LHC upgrade

Emilio Meschi[a,*] **on behalf of the CMS collaboration**

[a] *CERN, 1 Esplanade des Particules, Meyrin, Switzerland*

E-mail: emilio.meschi@cern.ch

To fully exploit the extended capability of its upgraded L1 trigger at the High-Luminosity LHC, CMS is pioneering a novel L1 Data Scouting (L1DS) system, capable of acquiring and processing the quasi-offline-quality trigger primitives produced by the upgraded L1 at the accelerator bunch-crossing rate of 40 MHz. The goal of the system is to give full access to potential physics signatures otherwise constrained by the L1 latency and accept rate limitations, or whose selection strategy diverges from that of the standard CMS physics program. To validate the concept and provide a development platform for the firmware and software required for the final system, a demonstrator system was assembled to operate with the current L1 trigger in the LHC Run 3. We discuss the L1DS demonstrator system architecture, performance, and some preliminary results, as well as the design of the final L1DS system and a summary of ongoing studies of its potential and competitiveness in selected physics channels.

42nd International Conference on High Energy Physics (ICHEP2024)
18-24 July 2024
Prague, Czech Republic

*Speaker

1. Introduction

The High-Luminosity upgrade of the LHC (HL-LHC) will produce pp collisions at a peak instantaneous luminosity up to 7.5×10^{34} cm^{-2}s^{-1}, with up to 200 pileup interactions per crossing, entailing a major upgrade of the CMS detector [1]. The data acquisition system of the upgraded CMS will continue to feature two trigger levels, with a first hardware-based Level-1 trigger (L1) system and a High Level Trigger (HLT) using software algorithms running over the full detector data on a heterogeneous compute farm. The fully redesigned L1[2] will exploit finer granularity trigger primitives from the upgraded front-end of the Barrel Calorimeter, the Hadron Forward calorimeter (HF) and the new High Granularity endcap calorimeter (HGCAL), as well as refined and more efficient muon identification, thanks to new front-end electronics and the addition of new RPC and GEM detectors. Most importantly, tracking will be introduced at L1, thanks to the addition of p_{T} layers in the outer silicon tracker, capable of identifying with high efficiency track segments with $p_{\mathrm{T}} > 2$ GeV up to $|\eta| < 2.4$, which will be fed to and reconstructed by a L1 track processor, fast enough for reconstructed tracks to be used in a L1 with a global latency of 12.5 μs.

The overarching goal of the CMS upgrades is to optimally exploit the HL-LHC data and extend the experiment's sensitivity to new phenomena. For the L1 trigger this means maintaining or reducing the Run-3 thresholds for all basic trigger objects in the presence of high pileup, while providing maximum flexibility to efficiently select specific signatures. This is to be achieved within an increased L1 accept rate budget of up to 750 kHz.

One of the most important new features of the upgraded L1 is the use of a Particle Flow (PF) algorithm in the Correlator (CT). The first layer of the CT links information from the calorimeter and muon standalone processors with charged tracks, identified by the Track Finder (TF), to improve momentum and energy resolution and classify energy depositions. The PF candidates reconstructed in the first layer of the CT are later used to reconstruct electron, photon, jet candidates, and compute energy sums in the second layer [3]. Information on the primary vertex reconstructed by the Global Track Trigger (GTT) is used to carry out pileup subtraction using the PileUp Per Particle Identification (PUPPI) algorithm [2]. The use of PF candidates provides significant improvements to the p_{T} resolution of all types of physics objects, in particular jets and energy sums. Combined with PUPPI pileup subtraction, they allow Run 3 thresholds on energy sums and missing energy to be maintained or even reduced despite the HL-LHC maximum pileup level being more than three times larger, thanks to a much sharper turn-on.

2. Level-1 Data Scouting concept and motivation

After the upgrade, the L1 will be capable of reconstructing exclusive signatures, in many cases with close-to-offline resolution, thanks to tracker tracks, the particle flow reconstruction, and control over pileup. Displaced objects will be selected with improved efficiency in the muon track finders and in the tracker. Correlations among different bunch crossings (BX) will be possible within ±2 BXs, providing improved sensitivity to heavy stable charged and long lived particles. The readout bandwidth of some detectors, the HLT, and the offline processing capacity will still dictate the maximum L1 accept rate, limiting the access to signatures that are too common, such as low E_{T} jets. Signatures whose combinatorics or complexity exceeds latency constraints, or computing capacity of the L1 processors (the former dictated by the length of readout pipelines, the latter by the limited, albeit very large, amount of logics available in FPGAs) will remain elusive. Exotic configurations with orthogonal requirements to “mainstream” physics may also be difficult to trigger on.

A Level-1 Trigger Data Scouting (L1DS) system, collecting and analysing L1 physics objects produced for every bunch crossing, has the potential to extend the access of CMS to physics otherwise constrained by the L1 latency and maximum accept rate. It potentially enables the exploration of additional exotic signatures, and is a powerful tool to study correlations over several BXs, for diagnostics and physics searches. The L1DS is designed to acquire and analyse the trigger objects produced by the L1 processors at the accelerator bunch-crossing rate of 40 MHz. As trigger primitives are designed to provide best and well-understood efficiency for physics objects and to control the accept rate, they are not necessarily immediately usable for physics. For example, a sharp turn-on at threshold does not always translate into best accuracy of the object parameters; features could be introduced by limitations of the processing hardware and/or the transport protocols and their bandwidth. In general, calibrations may not be optimal, and detailed simulation studies are needed to explore the capabilities of a scouting system.

3. Level-1 Data Scouting architecture and physics case

The L1DS baseline architecture is illustrated in Fig.1. Data from the trigger links will be collected using ATCA data acquisition boards (DAQ800) that have been designed for the CMS DAQ upgrade, and are equipped with two Xilinx VU35P FPGAs [4]. These boards are capable

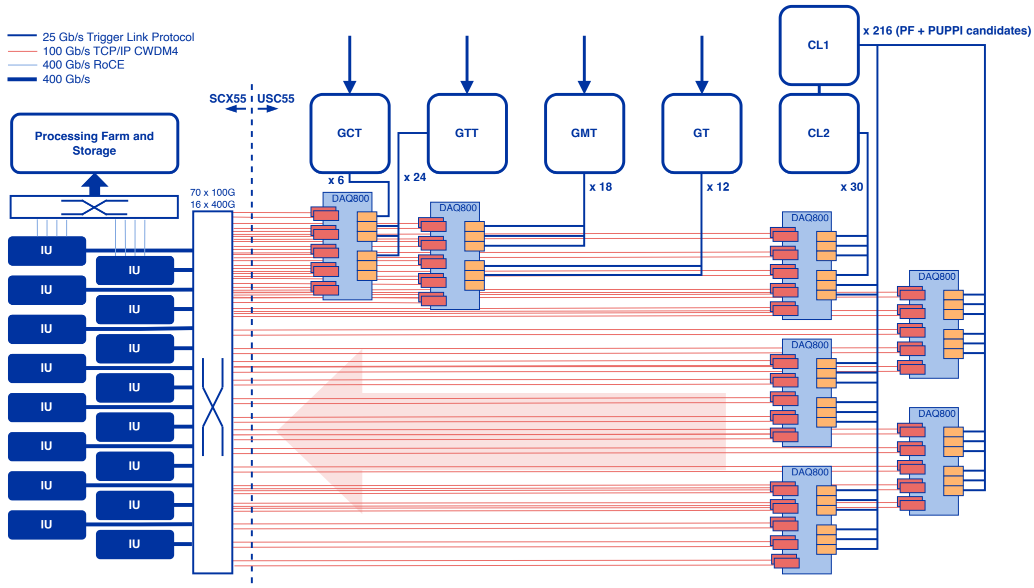

Figure 1: Baseline architecture of the L1DS.

of collecting data from up to 48 optical links at a speed of 25 Gb/s, the link speed of the trigger boards outputs. A modified firmware is being developed to interpret and aggregate the trigger protocol, apply minimal pre-processing (including zero-suppression, required to reduce the data throughput to match the board output maximum bandwidth of 800 Gb/s), and buffer data into the on-chip High-Bandwidth Memory (HBM), before transferring them to a commercial switched network in the surface counting room, using a standard TCP/IP protocol over long-range optical links at 100 Gb/s. The data streams, aggregated into 400 Gb/s links, will be received by a set of Ingestion Unit (IU) servers, that will assemble them into records containing data from all BXs in an LHC orbit (about 89 μs), and make them available to a distributed processing farm that will combine all the information from the different streams and carry out the online analysis and save the results to permanent storage. The system is sized to collect all candidate objects feeding the final GT algorithms, i.e. calorimeter and muon objects from the Global Muon and Calorimeter triggers (GMT and GCT), the primary vertex and multi-track objects from the GTT, as well as electrons, jets, etc. from the layer-2 of the CT. Additional links are devoted to the collection of pileup-subtracted (PUPPI) candidates or (optionally) PF candidates prior to processing in the CT layer-2. The latter require zero-suppression at the source, which limits the number of usable input links in the DAQ800 but makes it possible to use software algorithms for pileup subtraction and reconstruction of jets, for example, without latency or hardware resource limitations.

Several physics channels where the L1 scouting approach can potentially be beneficial are currently under study. **Soft hadronic final states** are characteristically difficult to trigger on, owing to the large QCD rates. For example, classic dijet resonance searches will benefit from the absence of rate limitations and from PF-jet p_{T} resolution to cover the low-mass region inaccessible to standard L1 dijet triggers, complementing studies using triggered data, boosted jets, and jet substructure. Multiple soft jet final states in general, where a cut-and-count approach would be sufficient and less sensitive to L1 jet features, as well as high multiplicity unclustered hadronic final states (either model-driven, such as hidden sector with compressed spectra, or model-independent) are being actively investigated. **Exclusive rare Standard Model bosons decays** are characterized by tiny predicted branching fractions, with reach critically dependent on trigger efficiency. For example, existing analyses searching for $\mathrm{H} \to \mathrm{J}/\psi\gamma$, $\mathrm{H} \to \phi\gamma$, and $\mathrm{H} \to \rho\gamma$ use single photon triggers, with thresholds entailing significant efficiency loss. Similarly, rare radiative W decay searches, such as $\mathrm{W} \to \pi\gamma$, $D_s\gamma$, currently (Run 3) use Ws from $t\bar{t}$ with significant efficiency loss. All-hadronic SM boson decays, for example $\mathrm{H} \to \phi\phi$, $\mathrm{W} \to \pi\pi\pi$ are currently inaccessible and, even with tracks at L1, trigger strategies for them are potentially challenging computationally- and latency-wise, due to the large combinatorics. **Lepton flavor anomalies and τ physics** can benefit from scouting because of notorious difficulties in controlling trigger rate. In particular, the $\mathrm{B}_s \to \tau\tau$ decay requires high efficiency τ selection at low-p_{T}. The partial reconstruction of $\tau \to 3\pi + X$ decays, an arduous combinatorial task in hardware, is under development for the L1DS. For the decay $\tau \to 3\mu$, acceptance for low-p_{T} muons (not necessarily fully reconstructed in the muon chambers) is key: techniques to recover these low-p_{T} muons are currently being studied. **Other new physics searches**, such as dark photon decays to low-p_{T} electron-positron pairs, may profit from the L1DS access to a lower-mass range, thanks to less-aggressive electron identification. Long-lived and very long-lived particle searches can potentially be extended on the one hand by bridging the gap between small displacement (tracks) and large displacement (standalone objects), for example by relaxing muon-track matching; on the other by exploiting multi-BX correlations e.g. among segments in the muon chambers. Anomaly Detection techniques, using all available L1 information at the BX rate, are also being investigated as a means to isolate exotic signatures that may escape classic algorithms. One important goal of these physics studies is to demonstrate that the analyses can be performed in real time with a reasonable amount of computing resources. The entire acquisition and analysis flow is emulated, starting from the L1 processors, using simulated data. The L1 objects are then injected into the acquisition chain, assembled into orbit blocks and distributed to a set of processing units, to evaluate the computing needs for each channel.

Some of the signals above depend on extensions to the baseline L1DS design, such as acquiring all L1 tracks, muon or calorimeter primitives before their processing in the GTT, GMT and GCT. These extensions require additional throughput capacity and computation power, both at the level of the acquisition boards and the online processing farm. Thanks to the modular nature of the L1DS design, it will be possible to extend the system even at a later stage in the project, fully profiting of new FPGA families with increased I/O capability, as well as improvements in the distributed real-time processing from progress in CPU and accelerators. R&D is continuing in all these areas.

4. The Run-3 L1DS demonstrator

To demonstrate the concept of the L1DS acquisition and processing chain, from the L1 processors to the DAQ, online processing, and storage, a demonstrator system was assembled to collect primitives from the current CMS trigger [5]. This system was completed and commissioned in 2023 and is currently operating in CMS Run-3. It is schematically shown in Fig.2. It collects global

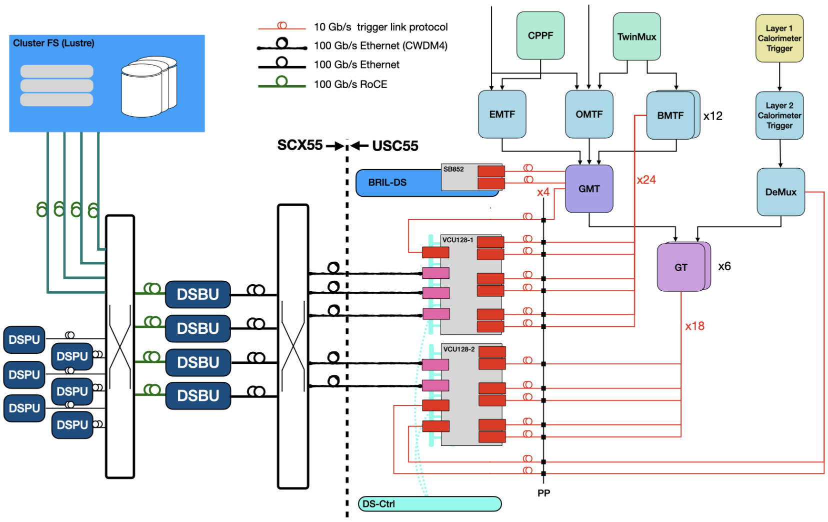

Figure 2: Schematic view of the Run-3 L1DS demonstrator.

muon candidates, calorimetric candidates (Jets, e/γ, τ_h, energy sums), stubs from the barrel muon drift tubes, and global trigger bits. The system uses two VCU128 boards from Xilinx. These boards were chosen for their similarity to the DAQ800 board targeted by the final design. They assemble and concentrate data from up to 24 trigger links into TCP/IP streams, subsequently routed to surface into a standard 100 GE network switch. LHC orbit records are assembled in a set of Buffer Unit servers (DSBU). The online processing is performed in a small computing farm using the standard CMS reconstruction framework. The system can collect data for every bunch crossing with close to 100% efficiency, giving insights into the detailed response of the trigger chain to the accelerator

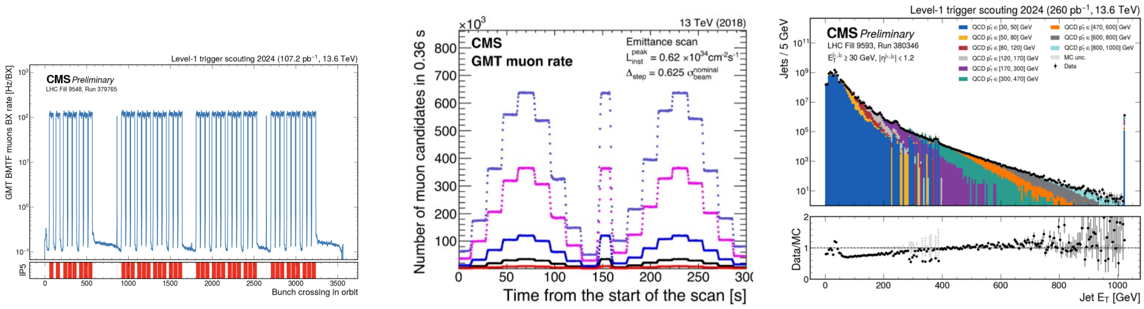

Figure 3: Left: muon rate in the Barrel Muon Track Finder (BMTF) for each BX in an LHC orbit (figure adapted from [7]). Center: muon rate in the Global Muon Trigger as measured by the scouting demonstrator, as a function of time in an LHC emittance scan, illustrating the L1 scouting capability as a luminometer (figure adapted from [2]). Right: E_T spectrum of L1 jets: the data (points) compare well to a QCD monte carlo, demonstrating the well-understood response of the system (figure adapted from [7]).

bunch structure (Fig.3 left). The rate of specific objects, in particular muons, can provide alternative measurements of bunch-by-bunch luminosity, thanks to the ability to collect all BXs (Fig.3 center). The response of the trigger can be accurately modeled by the emulator (see Fig.3 right). Despite the limited resolution, jet data from the L1DS demonstrator 2024 run is being used in searches of low-mass dijet resonances, while a study aims at reconstructing slow charged particles by linking muon chamber stubs from successive BXs [6].

5. Conclusions and outlook

The CMS L1 Trigger Data Scouting promises to complement the "standard" CMS HL-LHC physics program. The baseline design is well-established, and uses technology and approaches mediated in part from the CMS DAQ upgrade. A number of physics cases are being studied in detail and will be used to steer further development. A Run-3 L1DS demonstrator is successfully operating in CMS since 2023 and constitutes an important testing ground, with the potential to produce physics results.

References

[1] see e.g. G. Pásztor for the CMS Collaboration, PoS(LHCP2022), 045, (2023) and references therein.

[2] CMS Collaboration, CERN-LHCC-2020-004, CMS-TDR-021, (2020).

[3] CMS Collaboration, JINST 12, P10003, (2017).

[4] V. Amoiridis et al., IEEE Trans. Nucl. Sci. vol. 70, n. 6, 914-921, (2023).

[5] R. Ardino et al. Nucl. Inst. Meth. A, vol. 1067.169719 (2024).

[6] S. Giorgetti on behalf of the CMS Collaboration, CMS CR-2024/207, (2024).

[7] CMS Collaboration, CMS-DP-2024-056, (2024).