



CMS Level-1 trigger Scouting system upgrades for LHC Run-3 and Phase-2

R. Ardino^{1,2} on behalf of the CMS collaboration

¹Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, I-35131 Padova, Italy

²CERN, Esplanade des Particules 1, Meyrin, 1211, Switzerland

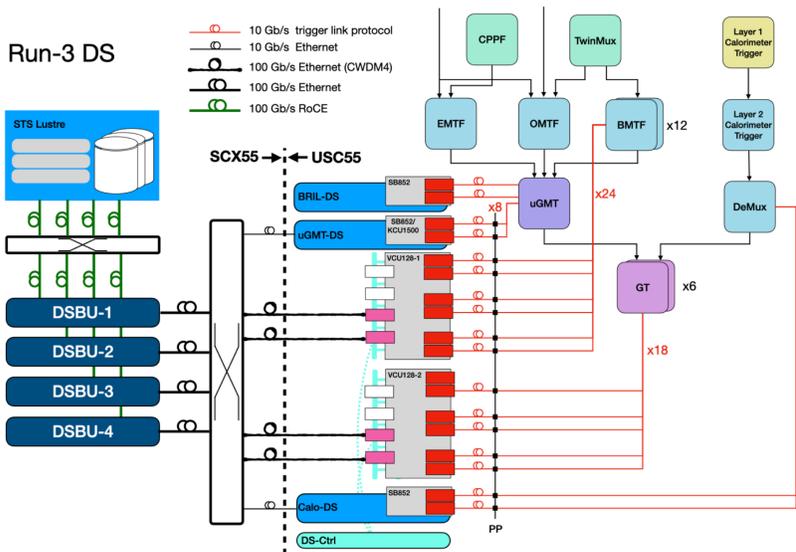


1. Overview

The Large Hadron Collider (LHC) will be upgraded in the second half of this decade, to provide an instantaneous luminosity of up to $7.5 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ to its two largest experiments, around seven times higher than the original design value. The upgraded collider, the High-Luminosity LHC (HL-LHC), will deliver an average number of proton-proton collisions per bunch crossing (pileup) up to 200. To cope with these extreme conditions the Compact Muon Solenoid (CMS) experiment will undergo a significant detector and hardware upgrade. Given a total event size of 7.5 MB, a full readout of the entire detector at the 40 MHz bunch crossing rate would not be feasible, due to limitations in readout, storage, and analysis capabilities. Therefore, CMS will continue to use two trigger levels: a Level-1 (L1) trigger based on field-programmable gate arrays (FPGAs), selecting events at 750 kHz; and a high-level trigger running on a farm of compute nodes performing the second level of selection in software. While the two-stage trigger of CMS is designed to provide excellent physics performance, the study of some physics processes can potentially benefit from an analysis of the full available dataset at 40 MHz, exploiting the near-offline performance of the upgraded L1 trigger reconstruction. Thus, a L1 trigger Data Scouting system (L1DS) is being developed [1]. The proposed architecture and hardware receive data from L1 trigger spare output links and perform a quasi-online analysis on them in a heterogeneous computing farm.

Physics signatures that would benefit from such a system are: displaced muon signatures, e.g. from Dark Matter (DM) signals; Heavy Stable Charged Particles (HSCPs), with event signatures covering multiple bunch crossings; $W \rightarrow \pi\pi\pi$ rare decay. Additionally, the Scouting system can significantly improve the diagnostic capability of the L1 trigger and online instantaneous luminosity measurements, as a result of the virtually unlimited dataset provided. Machine Learning (ML) inference can also be executed in the system hardware, improving the potential of the data for online analysis. Deep Neural Networks can be deployed directly on the FPGAs for muon primitives recalibration. Muon pair classifiers have also been designed to detect/reject "fake" muon pairs from the Global Muon Trigger, and are deployed on FPGA hardware as well. Future plans for the system include a Barrel Muon track finder.

2. Level-1 Data Scouting demonstrator for LHC Run-3



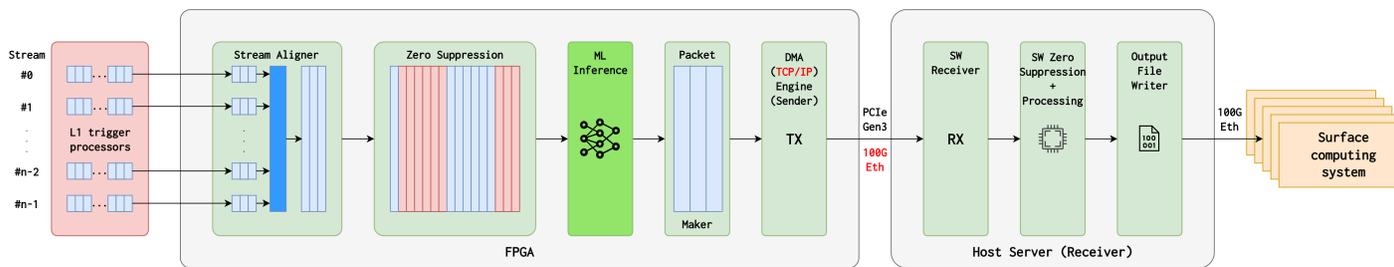
- L1DS system demonstrated for the first time at the end of LHC Run-2 in 2018
- For LHC Run-3, demonstrator expanded to readout multiple sources of the CMS L1 trigger
- **Global Muon Trigger (μ GMT)**
 - $8 \times 10 \text{ Gb/s}$ links from 1 trigger processor
 - Up to 8 muons
- **Layer-2 Calorimeter Trigger (DeMux)**
 - $7 \times 10 \text{ Gb/s}$ links from 1 trigger processor
 - Up to 12 jets, e/γ 's, taus
 - Missing Transverse Energy, Sums
- **Barrel Muon Track Finder (BMTF)**
 - $24 \times 10 \text{ Gb/s}$ links from 12 trigger processors
 - Stub primitives in Muon Barrel system
- **Global Trigger output (μ GT)**
 - $18 \times 10 \text{ Gb/s}$ links from 6 trigger processors
 - Global Trigger algorithm decision bits

3. FPGA Hardware for Run-3 demonstrator



- **Xilinx KCU1500**
 - Ultrascale XCKU115 FPGA
 - GTH transceivers
 - 2 on-board QSFPs
 - PCIe Gen3 \times 8 \times 2
 - Already used for Run-2 L1DS
- **Micron SB-852**
 - Ultrascale+ XCVU9P FPGA
 - GTY transceivers
 - 2 on-board QSFPs
 - PCIe Gen3 \times 16
 - 64 GB DDR4 RAM
- **Xilinx VCU128**
 - Ultrascale+ XCVU37P FPGA
 - GTY transceivers
 - (4 + 6 w/mezzanine) QSFPs
 - PCIe Gen4 \times 8 (or Gen3 \times 16)
 - 8 GB of HBM

4. L1 Scouting Hardware and Software design

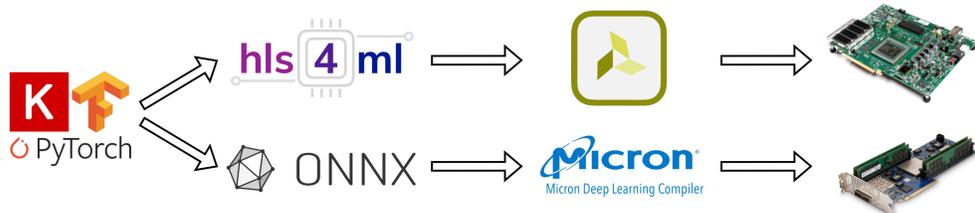


- **FPGA (Firmware)**
 - L1 trigger processors optical link receiver
 - Alignment of $n \times 32$ -bit word streams at 250 MHz
 - Zero Suppression of bunch crossing data with no interesting primitives (e.g. empty muons, low energy jets, ...)
 - Machine Learning inference applications
 - Packet maker for transfers to host server
- **Host Server (Software)**
 - DMA or TCP/IP transfers receiver
 - SW zero suppression and processing
 - Write output file, compression and transfer to computing farm

5. Machine Learning for L1 Scouting

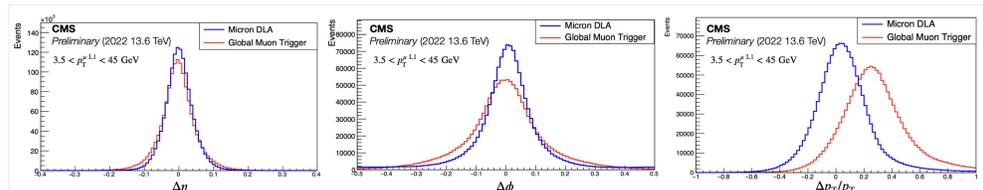
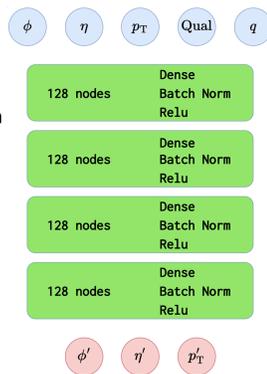
ML integration on FPGA Hardware

- hls4ml [2] for generation of neural networks in Hardware Description Language for Xilinx VCU128
- Micron Deep Learning Accelerator (MDLA) proprietary engine for deployment on Micron SB-852



Muon primitives online recalibration

- Trigger primitives are calibrated for a given efficiency at a threshold
- Useful for triggering, not for physics analysis!
- Train fully connected Neural Network (NN) for L1 objects recalibration
- Dataset for training: Run-3 ZeroBias
- Offline reco objects as target for recalibration of trigger level objects
- NN input: μ GMT muon primitives ϕ , η , p_T , quality Q , charge q
- NN output: correction terms for recalibration ϕ' , η' , p_T'
- $\Delta\phi$, $\Delta\eta$, Δp_T are the differences between
 - prediction (or μ GMT extrapolated) values...
 - ...and offline muon tracks for matched muons ($\Delta R < 0.1$ at 2nd muon station)
- Results for VCU128 using hls4ml yet to be approved



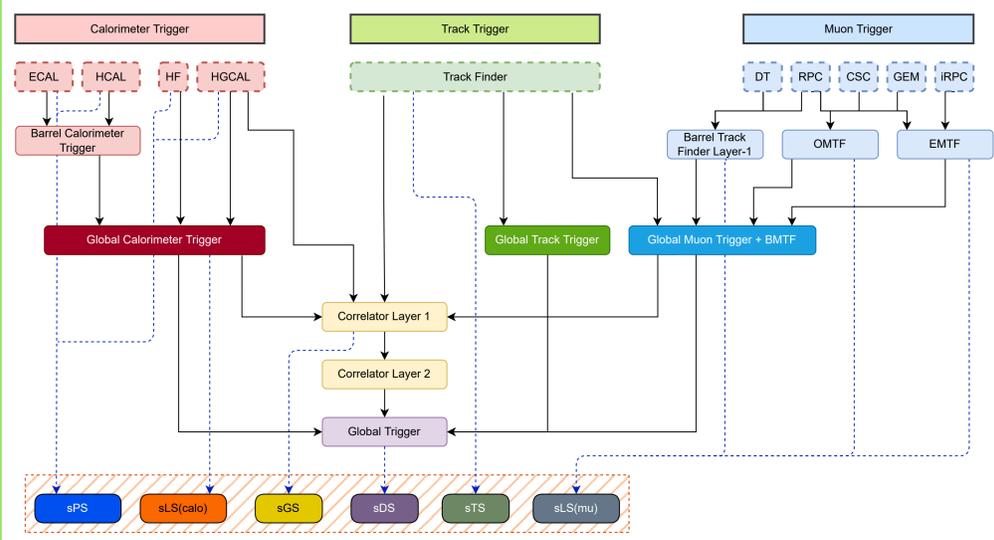
6. L1 Scouting for CMS Phase-2

New L1 trigger for CMS at HL-LHC [3]

- More sophisticated algorithms
- Approaching offline resolutions
- Vertex finding, Particle Flow
- Kalman Filter track reconstruction

L1DS will have a stageable architecture

- **GT inputs & outputs (sDS)**
- **Calo & Muon local reco (sLS)**
- **Tracker tracks (sTS)**
- **Calo primitives (sPS)**



References

[1] D. Rabady et al. "A 40 MHz Level-1 trigger scouting system for the CMS Phase-2 upgrade". In: *NIM-A* (2022). doi: <https://doi.org/10.1016/j.nima.2022.167805>.

[2] J. Duarte et al. "Fast inference of deep neural networks in FPGAs for particle physics". In: *JINST* 13.07 (2018). doi: [10.1088/1748-0221/13/07/P07027](https://doi.org/10.1088/1748-0221/13/07/P07027).

[3] CMS collaboration. "The Phase-2 Upgrade of the CMS Level-1 Trigger". In: (2020). Final version. URL: <https://cds.cern.ch/record/2714892>.