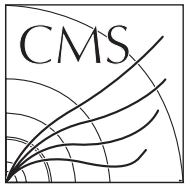
Available on CMS information server

CMS CR -2024/282



The Compact Muon Solenoid Experiment

Conference Report

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

26 October 2024 (v3, 05 November 2024)

Level-1 Tracking at CMS for the HL-LHC

Mark Steven Larson on behalf of the CMS Collaboration

Abstract

The success of the CMS physics program at the HL-LHC requires maintaining sufficiently low trigger thresholds to select physics processes. With an average expected 200 pileup interactions, it is critical to achieve this while maintaining manageable trigger rates by including tracking information in the Level-1 (L1) trigger. A 40 MHz silicon-based track trigger on the scale of the CMS central tracking system has never been built and can enable entirely new physics studies. The main challenges of reconstructing tracks in the L1 trigger are the large data throughput at 40 MHz and the need for a trigger decision within 12.5 μs. To address these challenges, the CMS Outer Tracker for the HL-LHC uses modules with closely spaced silicon sensors to read out only hits compatible with charged particles above 2 GeV. This contribution will discuss the L1 tracking algorithm and its implementation, present simulation studies of the estimated performance, and show results from hardware studies at the CERN Tracker Integration Facility.

Presented at *LHCP2024 12th Large Hadron Collider Physics Conference*

PROCEEDINGS OF SCIENCE

Level-1 Tracking at CMS for the HL-LHC

Mark Larson[a,*] **on behalf of the CMS Collaboration**

[a] *Northeastern University,*
360 Huntington Ave, Boston, United States of America

E-mail: mark.steven.larson@cern.ch

The success of the CMS physics program at the HL-LHC requires maintaining sufficiently low trigger thresholds to select physics processes. With an average expected 200 pileup interactions, it is critical to achieve this while maintaining manageable trigger rates by including tracking information in the Level-1 (L1) trigger. A 40 MHz silicon-based track trigger on the scale of the CMS central tracking system has never been built and can enable entirely new physics studies. The main challenges of reconstructing tracks in the L1 trigger are the large data throughput at 40 MHz and the need for a trigger decision within 12.5 μs. To address these challenges, the CMS Outer Tracker for the HL-LHC uses modules with closely spaced silicon sensors to read out only hits compatible with charged particles above 2 GeV. This contribution will discuss the L1 tracking algorithm and its implementation, present simulation studies of the estimated performance, and show results from hardware studies at the CERN Tracker Integration Facility.

**** Large Hadron Collider Physics Conference*
**** 3-7 June 2024 ****
**** Northeastern University, Boston, United States of America ****

[*] Speaker

1. Introduction

The LHC will be upgraded to the High-Luminosity LHC (HL-LHC), with operation expected to begin in 2030. The HL-LHC will provide unprecedented amounts of data, providing up to an average of 200 proton-proton interactions, or pileup (PU), at a rate of 40 MHz, corresponding to an instantaneous luminosity of up to 7.5×10^{34} cm^{-2} s^{-1} [1]. This will greatly improve the potential of searches for new physics, as well as precision measurements and searches for rare processes of the Standard Model. However, this increase in data rate makes for challenging detector conditions. The Phase-2 upgrades [2] to the CMS detector [3, 4] aim to adapt and improve the CMS detection capabilities for the HL-LHC environment.

To take full advantage of the increase in instantaneous luminosity while maintaining manageable data readout rates, the Level-1 (L1) hardware trigger system will be upgraded. The upgraded L1 trigger will increase its maximum output rate to 750 kHz, along with an increased maximum latency of 12.5 μs compared to 4 μs [5]. One important component of this upgrade, which will both improve the quality of data collected and provide improved event rejection, is the inclusion of tracking information into the L1 trigger system. This information will allow to improve transverse momentum (p_T) resolution by reducing overestimation of muon momenta, electron identification, and the reconstruction of hadronic tau decays. Additionally, the effects of increased pileup will be significantly mitigated by allowing for the identification of the primary vertex at L1, and, with calorimeter L1 input, a simplified Particle Flow event reconstruction can be run at L1 [6].

1.1 CMS Outer Tracker Upgrade

The inclusion of tracking information into the L1 trigger is made possible by the upgrade to the CMS Outer Tracker (OT) [7]. The tracker geometry in the $r - z$ plane is presented in Fig. 1. This upgrade will increase radiation tolerance, improve granularity, and allow for the readout of hits consistent with $p_T > 2$ GeV at the bunch crossing frequency of 40 MHz through the use of silicon-based p_T modules. These double-sided p_T modules will be implemented in 6 barrel layers and 5 disk endcaps, and will come in two types: pixel-strip (PS) and strip-strip (2S). The PS modules will be placed in layers 1 – 3 and in the inner half of the disks, and 2S modules will be placed in layers 4 – 6 and in the outer half of the disks. These modules estimate the track curvature, and hence p_T, by correlating hit positions in the top and bottom halves of the modules. The hits in these modules consistent with $p_T > 2$ GeV are read out as 'stubs' at 40 MHz.

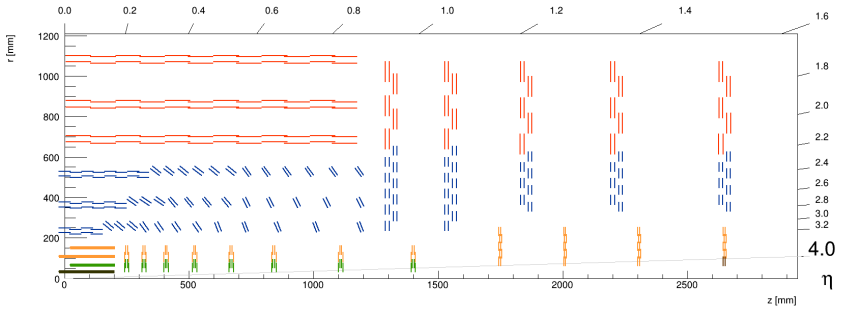


Figure 1: Depiction of the upgraded CMS tracker for HL-LHC in the $r - z$ plane, with PS modules in blue, and 2S modules in red. The Inner Tracker is depicted in black, yellow, and green.

2. Level-1 Track Finding Algorithm

The goal of the L1 Track Finding (TF) algorithm is to reconstruct all tracks with $p_T > 2$ GeV and $|\eta| < 2.4$ and send track parameters like p_T, impact parameters, and others to the L1 trigger. This algorithm will be implemented on Track Finder Processor (TFP) boards using field-programmable gate arrays (FPGAs). Each TFP board takes input from Data, Trigger and Control (DTC) boards. The detector is split into 9 ϕ nonants, for which there are 24 DTC and 18 TFP boards each.

The algorithm is split into a pattern recognition and a track fitting stage. First, pairs of stubs in adjacent layers and disks are formed into seeds (tracklets). Next, the trajectory of each tracklet is projected inward and outward to other layers and disks, and stubs consistent with these projections are matched to the tracklet. These first three steps of the tracklet algorithm are presented in Fig. 2. A Kalman Filter is then used to identify the best stub candidates and track parameters, and lastly a Boosted Decision Tree is used to evaluate track quality. The implementation of this algorithm in hardware and firmware, and its simulated performance, are presented in the following sections.

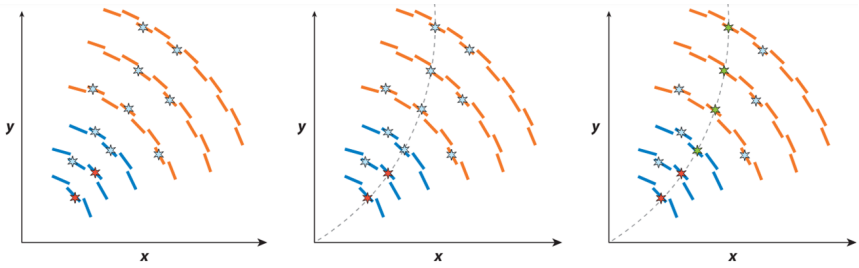


Figure 2: Illustration of the tracklet algorithm showing first the formation of seeds (in red) from stubs (marked as stars), their projection to other layers/disks, and lastly the matching of consistent stubs (in green) to tracklet projections.

3. Hardware Implementation

The TF algorithm is implemented on the 'Apollo' platform, a custom dual-FPGA board [8]. The algorithm on this board must run within 4 μs. This board is based on the Advanded Telecommunications Computing Architecture (ACTA) platform, and consists of two large Virtex Ultrascale 13 Plus (VU13P) FPGAs, 25 GB/s 'firefly' optical links, and a Zynq micro controller.

Thermal studies were conducted at the CERN Tracker Integration Facility (TIF) to compare the temperatures of each FPGAs and in each physical slot within the larger ACTA shelf (shown on the left in Fig. 3) at high computing load by increasing the power consumption of each FPGA until reaching 85 °C, while using a constant fan level. The presented results indicate how best to balance the computing load across each FPGA and ACTA slot to avoid overheating. It was found that at higher power consumptions FPGA 1 tends to run hotter, whereas at relatively lower power consumptions FPGA 2 tends to run hotter. As illustrated on the right side of Fig. 3, the distribution of temperatures across ACTA slots was confirmed to be inversely proportional to the expected airflow (not shown) within each slot.

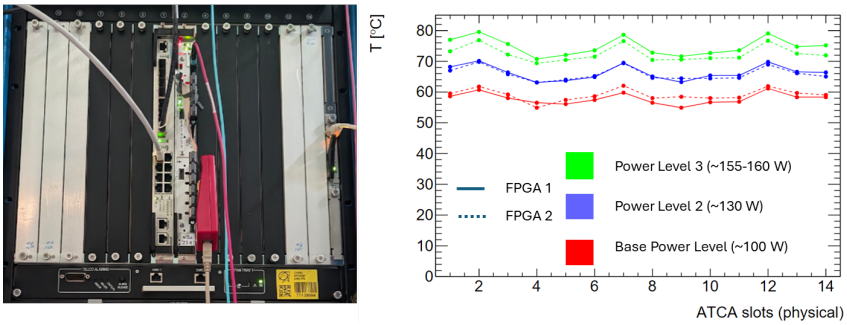


Figure 3: Thermal test setup at TIF (left), results of thermal studies showing temperature for each FPGA at different FPGA power consumption levels and across all ATCA slots (right).

4. Firmware Implementation

To program the L1 tracking algorithm onto the Apollo board, Vivado High-Level Synthesis (HLS) is used for the pattern recognition and projection of tracklets, and VHDL is used for the Kalman Filter track fit. HLS allows for the algorithm to be written in C++ and programmed directly onto the FPGA. Within the algorithm, processing modules perform calculations (e.g., projecting tracklets, matching stubs to tracklets), and memory blocks are used to transmit relevant data between processing modules.

The latest version of the TF algorithm has been implemented in firmware for the barrel region on only one FPGA. The final algorithm must be split across the two FPGAs of the Apollo board to take full advantage of the full Apollo board. Major progress has been made toward this dual-FPGA project and toward including the full detector. The updated dual-FPGA project pipeline is presented in Figure 4, with an emphasis on the new and modified processing modules and memory blocks that were required to be implemented to allow splitting the algorithm across the two FPGAs.

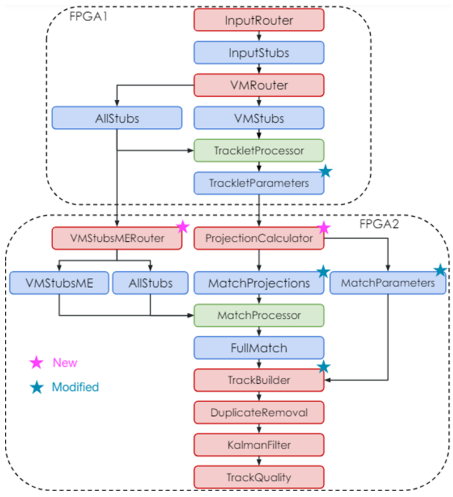
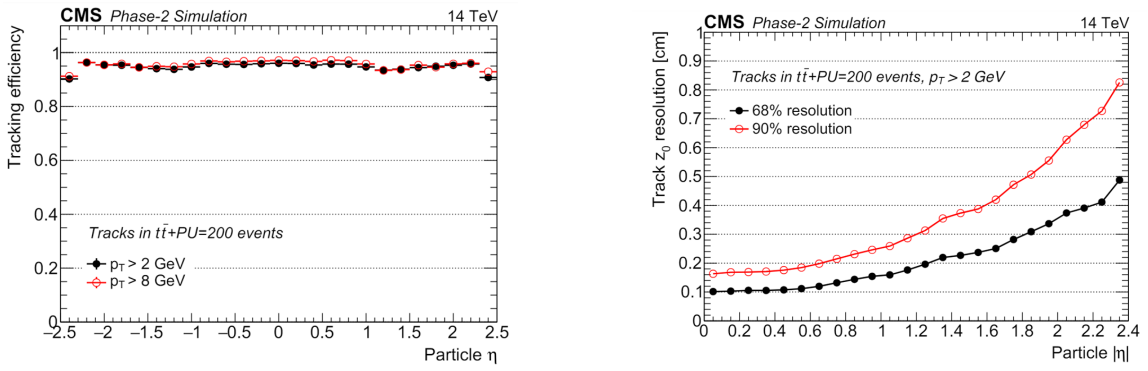


Figure 4: The pipeline for the dual-FPGA workflow is presented, with processing modules in red and green, and memory blocks in blue. The modules marked as "new" were developed within the past 6 months to allow for splitting the algorithm between both FPGAs, with the generation of tracklets to be done in FPGA 1, and the projection and matching of stubs to these tracklets, as well as track fit and quality evaluation, in FPGA 2.

5. Software Emulation

To evaluate performance and track quality for the TF algorithm, a bit-level accurate (> 99% agreement for $t\bar{t}$+ 200 PU events) software emulation integrated into the CMS software (CMSSW) framework is used. Track finding efficiency and longitudinal impact parameter, z_0, resolution results from this emulation are presented in Fig. 5. Recently, the emulation has been updated to include the modifications for the dual-FPGA project, which has minimal effect on performance.



(a) **(b)**

Figure 5: Figure 5 shows track finding performance for $t\bar{t}$ with 200 PU. In Fig. (a) the track finding efficiency across η is shown to be high (> 95%), and slightly improved for tracks with $p_T > 8$ GeV relative to $p_T > 2$ GeV. In Fig. (b) the z_0 resolution of tracks as a function of $|\eta|$ is presented. It is shown to worsen in the endcap regions which is a result of the orientation of the endcap tracker modules.

An under development future software emulation will allow for direct incorporation of code from the HLS processing modules into CMSSW, ensuring accuracy and synchronicitiy between software and firmware. Significant progress has recently been made in developing this new emulation. Wrapper code has been written for each HLS processing module from the single FPGA chain, and performance plots from the current CMSSW emulation are successfully reproduced.

6. Conclusions

Incorporating tracking information into the L1 trigger decision at CMS for the HL-LHC is crucial to reduce the data throughput and maximize the quality of the data collected. This is made possible through the use of double-sided p_T modules that provide p_T discrimination at the front-end readout stage. The track finding algorithm takes local track segments as input, performs pattern recognition, track fitting and an evaluation of track quality, before finally tracks are output to the L1 trigger. The algorithm is implemented on a dual-FPGA board, for which the firmware is written in HLS and VHDL. A simplified version of the algorithm has been demonstrated in hardware, and the final implementation is in progress. A software emulation exists and indicates high tracking efficiency and good track quality.

References

[1] HL-LHC Collaboration, *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report*, Tech. Rep. CERN-2020-010, CERN, 2020.

[2] CMS Collaboration, *Technical Proposal for the Phase-II Upgrade of the CMS Detector*, Tech. Rep. CERN-LHCC-2015-010, CERN, 2015.

[3] CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[4] CMS collaboration, CMS Collaboration, *Development of the CMS detector for the CERN LHC Run 3*, *JINST* **19** (2024) P05064.

[5] CMS Collaboration, *The Phase-2 Upgrade of the CMS Level-1 Trigger*, Tech. Rep. CERN-LHCC-2020-004, CERN, 2020.

[6] A. Ryd and L. Skinnari, *Tracking Triggers for the HL-LHC*, *Annual Review of Nuclear and Particle Science* **70** (2020) 171–195.

[7] CMS Collaboration, *The Phase-2 Upgrade of the CMS Tracker Technical Design Report*, Tech. Rep. CERN-LHCC-2017-009, CERN, 2017.

[8] A. Albert, Z. Demiragli, D. Gastler, K. Hahn, E. Hazen, P. Kotamnives et al., *The Apollo ATCA design for the CMS track finder and the pixel readout at the HL-LHC*, *Journal of Instrumentation* **17** (2022) C04033.