PROCEEDINGS OF SCIENCE

Overview of the HL-LHC upgrade for the CMS Level-1 Trigger

PoS(ICHEP2022)639

Benjamin Radburn-Smith[a,1,*]

[a] *Imperial College London, London, UK*

E-mail: benjamin@cern.ch

High-precision measurements of the standard model as well as searches for new physics beyond the standard model require the statistical power of the High-Luminosity LHC. The CMS experiment will perform an upgrade of the Level-1 hardware trigger in order to take full advantage of the High-Luminosity LHC. This Phase-2 Level-1 Trigger will utilise, for the first time, tracking and high-granularity calorimeter information and employ sophisticated algorithms such as particle flow reconstruction and Machine Learning techniques. An overview of the design of the Phase-2 Level-1 Trigger, along with two recent Machine Learning algorithm highlights, and benchmark examples of the extension to physics acceptance is presented.

41st International Conference on High Energy physics - ICHEP2022
6-13 July, 2022
Bologna, Italy

[1] on behalf of the CMS Collaboration
[*] Speaker

1. Introduction

We require the high statistics provided by the High-Luminosity LHC (HL-LHC) in order to search for new physics at the electroweak-scale by investigating any small deviations to the standard model. We can also probe challenging phase space by increasing the acceptance with new detectors and new capabilities. In order to achieve the high statistics projected to be delivered by the HL-LHC of around 4 ab^{-1} we need the instantaneous luminosity to increase to around 7.5×10^{34} $cm^{-2}s^{-1}$. However, as a consequence of this we expect the number of parasitic collisions, called Pileup (PU), to increase from around 40 in Run 3 to around 200 in Run 4. If we attempted to use the current Phase-1 Level-1 Trigger (L1T) system and algorithms within HL-LHC environment of 200 PU, we would need to output around 4 MHz in order to achieve the same Physics acceptance that we currently have. The planned total allowed Phase-2 L1T (P2 L1T) bandwidth will be 750 kHz.

The CMS experiment will undergo many upgrades, including those to the L1T, High Level Trigger, Data Acquisition, Tracker, Minimum Ionising Particle Timing Detector, Barrel Calorimeters, Endcap Calorimeter, Muon Systems, and Beam Radiation and Luminosity Instrumentation. For what concerns the L1T, there will be new inputs from the silicon strip tracker, the High Granularity Calorimeter, and Gas Electron Multiplier detectors in the muon system. There will also be upgrades from the Electromagnetic and Hadron Calorimeter back-end boards, and Drift Tube, Cathode Strip Chamber and Resistive Plate Chamber back-end boards that comprise the rest of the Muon detector.

The architecture of the L1T will be upgraded in order to allow data from all the new inputs. Figure 1 shows the planned architecture for the P2 L1T. We will have a dedicated Global Calorimeter Trigger, Global Muon Trigger, and Global Track Trigger which implements logic upon inputs from Trigger Primitives or Local Reconstruction. Then there is the Correlator Trigger (CT) which takes inputs from across the detector and the aforementioned Calorimeter, Muon and Track Triggers to perform more sophisticated algorithms. Finally the Global Trigger takes inputs from the Calorimeter, Muon, Track and Correlator Triggers, as well as some external sources, to make the final logic decisions of the L1T at a rate of 750 kHz. The allowed latency of the P2 L1T was increased from 3.8 μs in Phase-1 to 12.5 μs in Phase-2. This allows the L1T to read the inputs from across all the new detectors and implement more complex algorithms.

All the boards in the P2 L1T design plan to use the same Xilinx VU13P FPGA, with more than 100 high-speed optical links running at speeds of up to 28 GB/s. However, different board families will perform different functions and will have different firmware. The four main board families are the X20, APx-F, BMT-L1 and Serenity. The latency is planned to be less than 9.5 μs, which gives a 20% buffer.

2. Particle Flow

One of the key features of the P2 L1T is the CT which can implement sophisticated algorithms allowing for higher level objects. The CT utilises particle-flow (PF) which has proven to be useful in offline reconstruction during Run 2. A two layer system is used to achieve this. The first layer produces PF candidates constructed from the matching of calorimeter clusters and tracks. It also runs the Pileup Per Particle Identification (PUPPI) algorithm which mitigates the degradation of

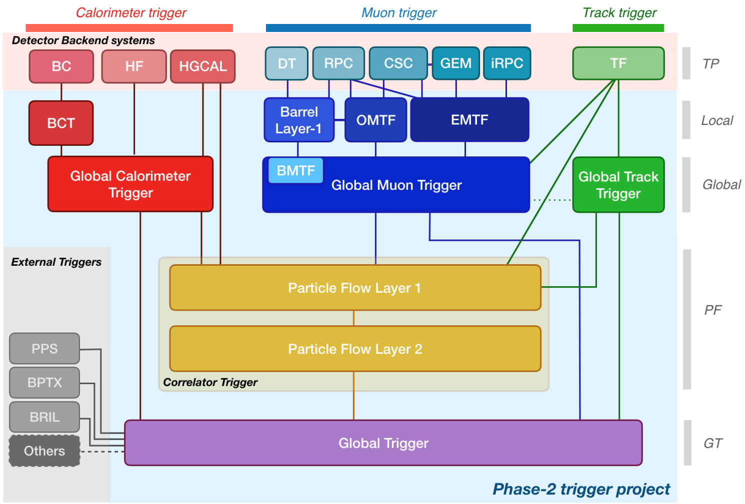

Figure 1: The planned architecture for the P2 L1T which receives inputs from the calorimeters, muon spectrometers and tracker. There are three global reconstruction level triggers: the Global Calorimeter Trigger, the Global Muon Trigger, and the Global Track Trigger. The Correlator Trigger is comprised of two layers dedicated to particle-flow reconstruction. The Global Trigger takes all the objects formed in the reconstruction, as well as some external sources, to make the final logic decisions [1].

energy resolution due to PU. The second layer builds and sorts the final trigger objects. It also applies additional identification and isolation requirements to those objects. We require both PF and PUPPI in order to sustain Jets and Missing Energy thresholds similar to those used in Run 2.

All of the sophisticated PF algorithms still need to fit within the resources of the FPGAs. For the Layer 1 we have the fully working PF and PUPPI algorithms working on Xilinx VU9P-2 boards, and plan to deploy and test on the larger VU13P-2 boards. Despite the complexity of these algorithms we find there are still enough resources to expand with other algorithms in the future. For the Layer 2 we have two well performing algorithms for the Jet finding: the SeedCone, and the Histogram algorithms. We are currently working to expand the scope of identification and isolation for the Neural Network Tau algorithms, NNTau, and the egamma identification. Figure 2 shows representations of the resource usage for the Layer 1 Barrel, Layer 2 Jet SeedCone algorithm, and Layer 2 NNTau algorithm, highlighting the space to potentially include other algorithms in the future.

3. Highlights

Two recent highlights of the P2 L1T algorithm developments come in the form of an End to End Neural Network (NN) for vertex finding [2], and a NN to identify jets originating from b-quarks [3].

The vertex finding algorithm is referred to as End to End as the track to vertex association is optimised during the training of the network. In order to fit on the resources of the FPGA the NN was both quantised, in which the weights, biases or activations were constrained to fewer values, and pruned by removing redundant connections. In comparison to the baseline histogramming algorithm we find the NN vertex finder has lower tails in the primary vertex (PV) z_0 position residual, given by z_0^{PV}(True)$-z_0^{PV}$(Reconstructed), for $t\bar{t}$ Monte Carlo (MC) simulation, where z_0

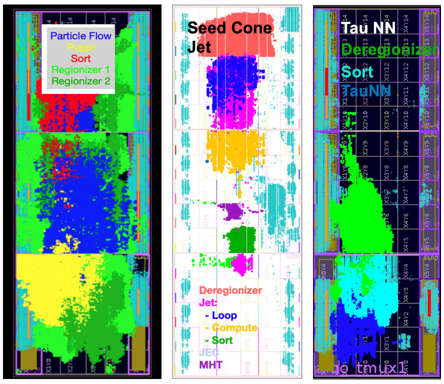

Figure 2: Example floor-plans of resource usage created with Vivado High Level Synthesis software on a Xilinx VU9P FPGA. Representations of resource usage in the Layer 1 Barrel (left), the Layer 2 Jet SeedCone algorithm (middle), and Layer 2 NNTau algorithm (right) are shown.

is the position along the beam axis. The algorithm returns a likelihood of a track belonging to the PV with a flexible threshold, which can be used by downstream algorithms. The performance of the NN algorithm in terms of making the correct association between the track and PV is increased compared to the baseline. The PV z_0 position residual and track association performance is shown in Figure 3.

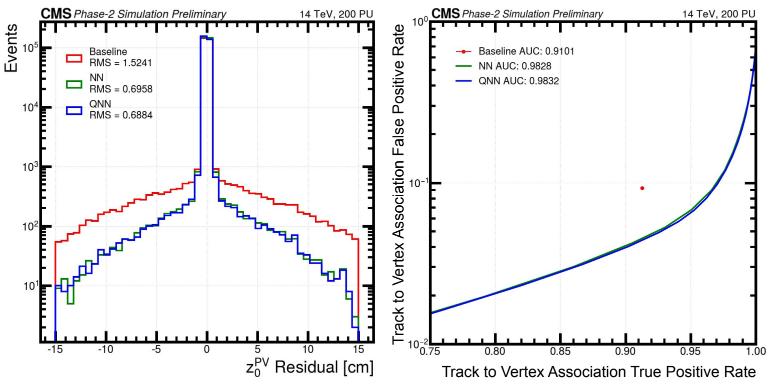

Figure 3: The PV z_0 position residual (left) and track association performance (right) of the End to End NN vertex finding algorithm as compared to the baseline histogramming algorithm measured in $t\bar{t}$ MC [2].

The NN used to identify jets originating from b-quarks is implemented in the Correlator Layer 2. The algorithm runs on the PUPPI particles in each jet and discriminates between jets originating from b-quarks and those from light quarks or gluons. Using the NN we gain better performance compared to conventional triggers for di-Higgs masses less than 500 GeV. Figure 4 shows the L1 trigger efficiency as a function of di-Higgs mass for HH→bbbb events. For di-Higgs masses less than 500 GeV the NN has a 68.9% efficiency compared to 61.4% efficiency using a QuadJet+HT trigger.

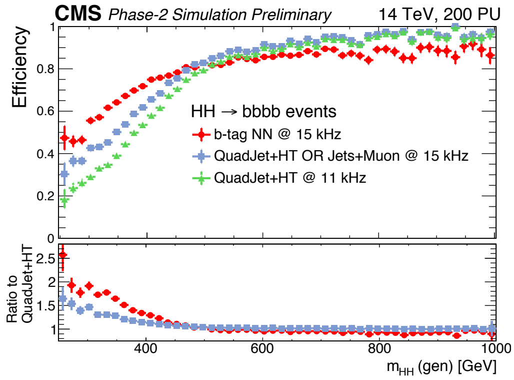

Figure 4: The L1 trigger efficiency as a function of di-Higgs mass for HH→bbbb events [3].

4. Physics Reach Examples

With the upgraded L1T system we now have access to tracker tracks, PF objects, and PV information at the L1T, as well as an extension of muon and electron acceptance. This allows us to trigger on rare B-meson decays, displaced jets, soft and correlated muons, Higgs candidates produced through Vector Boson Fusion (VBF), displaced muons and leptons at high pseudorapidity.

An example of triggering on rare B-meson decays is with $B_S^0 \to \Phi(K^+K^-)\Phi(K^+K^-)$ decays [1]. This decay is a rare flavour changing neutral current process which is forbidden at the tree level in the standard model. We can potentially trigger on the fully hadronic final state with L1 tracker tracks. The aim would be to reconstruct the Φ candidates through pairs of oppositely charged tracks originating from the same vertex. The B_S^0 candidates could then be reconstructed from pairs of Φ candidates which originate from the same vertex. Figure 5 shows the possible resolution of B_S^0 candidates within the L1T system with respect to offline selection and background processes at the L1T.

For triggering on VBF Higgs events either decaying to invisible particles or pairs of b-quarks, we require the L1 PF reconstruction [1]. NN algorithms were separately trained on the two different Higgs decays as signal events and using minimum bias MC as the background. The models were pruned to remove connections with low weights, in order to fit onto the resources of an FPGA. It is observed in simulations that the NN outperforms the cut-based trigger algorithms. Figure 6 shows the possible improvement in signal efficiency when using the NN algorthims at the L1T compared to baseline cut-based algorithms. There are larger gains in the VBF Higgs→bb signal where missing transverse energy triggers could not be used.

5. Summary

The High Luminosity LHC (HL-LHC) is required to gain large statistics for physics analyses, but it comes at the price of a high number of parasitic collisions. The CMS experiment will undergo many developments in order to capitalise on the HL-LHC, including utilising entirely new detectors, as part of a Phase-2 upgrade. The Level-1 Trigger (L1T) will be upgraded in order to not only cope with the harsh environments provided by the HL-LHC, but also to increase the physics acceptance. Novel triggering solutions are being developed now for Phase-2. This includes

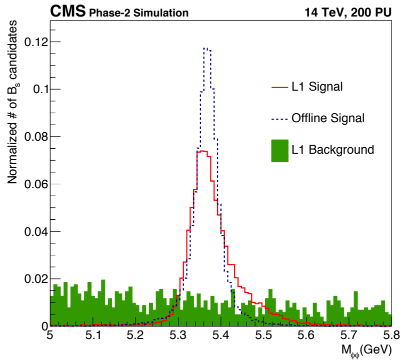

Figure 5: The possible resolution of B_S^0 candidates at the L1T with respect to offline selection and background processes [1].

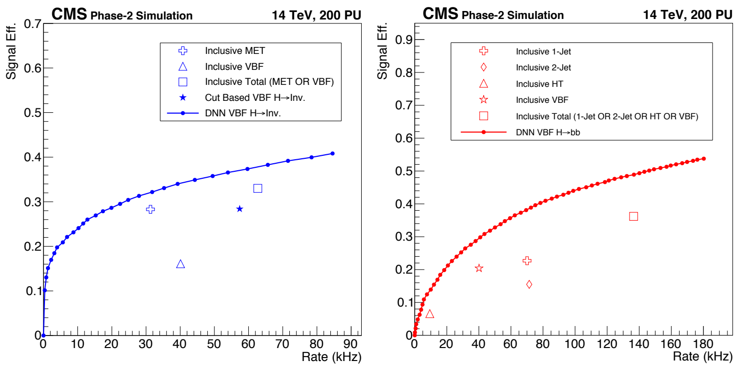

Figure 6: The possible improvements in signal efficiency as a function of rate when using dedicated NN algorithms at the L1T to trigger on VBF Higgs→invisible (left) and VBF Higgs→bb (right) decays [1].

developing Machine Learning (ML) techniques which have to fit onto the resources of the FPGAs. Modern ML is possible due to the more powerful FPGAs and the tools needed to synthesise ML in those FPGAs such as High Level Synthesis. It allows us to move from using simple algorithms at the L1T to more complex offline-like algorithms. Examples of recent ML algorithm development for the Phase-2 L1T have been shown. The physics reach will be extended with the upgraded L1T system. A number of benchmark channels have been studied in order to understand the potential improvement compared to Run 2 triggers.

References

[1] *The Phase-2 Upgrade of the CMS Level-1 Trigger*. CERN-LHCC-2020-004, CMS-TDR-021. CMS, 2020. URL: https://cds.cern.ch/record/2714892.

[2] *Performance of the End-to-End Neural Network Approach to Phase-2 Level-1 Trigger Primary Vertex Reconstruction and Track to Vertex Association*. 2021. URL: https://cds.cern.ch/record/2792619.

[3] *Neural network-based algorithm for the identification of bottom quarks in the CMS Phase-2 Level-1 trigger*. 2022. URL: https://cds.cern.ch/record/2814728.