

Design of the CMS upgraded trigger from Phase I to Phase II of the LHC

Manfred Jeitler*

Institute of High Energy Physics of the Austrian Academy of Sciences

E-mail: manfred.jeitler@cern.ch

**Bernhard Arnold, Herbert Bergauer, János Erő, Sebastian Templ,
Claudia-Elisabeth Wulz, Mateusz Zarucki**

Institute of High Energy Physics of the Austrian Academy of Sciences

Dinyar Rabady, Hannes Sakulin

CERN, on behalf of the CMS Collaboration

The CMS experiment implements a sophisticated two-level triggering system composed of the Level-1, instrumented by custom-design hardware boards, and the High Level Trigger, a streamlined version of the offline reconstruction software running on a computer farm (more than 30k CPU cores). In 2017, the LHC delivered proton-proton collisions at a centre-of-mass energy of 13 TeV with a peak instantaneous luminosity larger than $2 \cdot 10^{34} \text{cm}^{-2} \text{s}^{-1}$, more than twice the peak luminosity reached during Run1 and far larger than the design value. The CMS Level-1 trigger was upgraded during the end-of-the year technical stop between 2015 and 2016, to improve its performance at high luminosity and large number of simultaneous inelastic collisions per crossing (pile-up). All the electronic boards have been replaced and the upgraded electronics tested and commissioned with data. Smarter, more sophisticated, and innovative algorithms are now the core of the first decision layer of CMS: the upgraded trigger system implements pattern recognition and MVA (Boosted Decision Tree) regression techniques in the trigger boards for p_T assignment, pile-up subtraction, and isolation requirements for electrons and tau leptons. In addition, the new global trigger is capable of evaluating complex selection algorithms such as those involving the invariant mass of trigger objects.

The High Level Trigger features a trade-off between the offline complexity of the algorithms and the available computing power, and between the selection efficiency the maximum sustainable output rate. The trigger selections used in Run-2 will be presented, ranging from simpler single-object selections to more sophisticated algorithms combining different objects and applying analysis-level reconstruction and selection.

This presentation will cover the design and performance of the Phase I trigger and how it influences the path towards the Phase II upgrade necessary for the LHC run at a center-of mass energy of 14 TeV with luminosity of $(5-7) \cdot 10^{34} \text{cm}^{-2} \text{s}^{-1}$, corresponding to 140–200 pile-up events. The addition of the tracker information at Level-1 and the enhanced computing resources at HLT will maintain the trigger efficiency at a similar level as the present one.

European Physical Society Conference on High Energy Physics - EPS-HEP2019 -

10-17 July, 2019

Ghent, Belgium

1. The upgrade of the CMS Level-1 Trigger

The first-level trigger of the CMS experiment has been implemented in hardware for latency reasons: a decision must be taken within a fixed time span to guarantee that the event data to be retrieved are still available in the on-detector hardware buffers. The somewhat conflicting requirements have been the following:

1. latency: the decision must be taken within a short, fixed time span
2. limited input data: the trigger decision must be based on a limited amount of data small enough to be read out from the detector at 40 MHz; this means that only part of the subdetectors could be used for the first-level trigger decision, and their data had to be used at reduced precision
3. high trigger efficiency: the amount of potentially interesting physics events lost by the trigger must be kept at a minimum
4. manageable rates: the total data rate must not exceed the limit given by the capabilities of the readout system (100 kHz in the current CMS setup)

The limitations listed under points 1 and 2 make it difficult to satisfy the requirements 3 and 4. The situation will become even more difficult under the conditions of the High-Luminosity LHC because the higher data rates required to investigate very rare physics channels and to carry out high-precision measurements entail a higher number of simultaneous proton-proton collisions (pileup) of up to 200 proton-proton collisions per bunch interaction. To take up this challenge, improvements on all four points listed above are envisaged: the latency (the length of the on-detector memories) will be increased from 4 to 12.5 μs , the precision of input data will be improved, more subdetectors will enter the Level-1 trigger (in particular, the silicon strip tracker of CMS), and the readout capacity and thus the maximum admissible Level-1 trigger rate will be increased from 100 to 750 kHz [1]. At the same time, it will be important to significantly increase the trigger efficiency for certain types of events, which may so far have been lost at the trigger level.

This paper is concentrating on the plans for the algorithmic trigger logic, which combines physics objects derived from various types of digitized detector data and is implemented in the so-called “Global Trigger”.

2. The Global Trigger

The task of the Global Trigger has been to receive and process trigger “objects” (such as particle candidates for electrons, muons, tau leptons and jets) with values for their coordinates in azimuth (φ) and pseudorapidity (η) as well as for their transverse momentum and certain additional parameters such as charge and reconstruction quality of muons. There are also other trigger objects not linked to individual particle candidates, such as sums of the transverse momenta of particles in the detector calculated over different ranges in pseudorapidity and by including different types of subdetector information. These input data have been provided by the muon system and the

*Speaker.

calorimeters while data from the silicon tracker have not been available for the Level-1 trigger. The Global Trigger sets thresholds on the received parameters and combines trigger objects to trigger “algorithms” according to requirements defined in a flexible “trigger menu”. The decisions of the various algorithms are combined to a final verdict, the so-called “Level-1 Accept” signal, to either trigger the readout of an event or else discard the event and allow the corresponding data to be overwritten. Fig. 1 shows the setup of the CMS Level-1 Trigger after the Phase-2 upgrade and the position of the Global Trigger in the overall system.

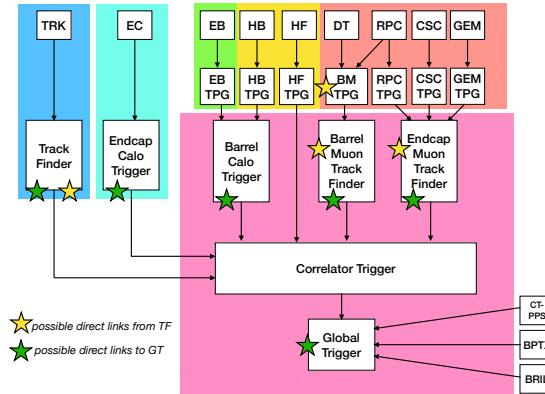


Figure 1: Setup of the CMS Level-1 Trigger after the Phase-2 upgrade

3. Triggering on non-prompt decays

So far, the time resolution for all input data to the Global Trigger has been given by the bunch collision frequency in the LHC of 40 MHz. In other words, timing was only given in units of about 25 ns (referred to as “bunch crossings” in LHC experiments). Most trigger algorithms have used only data from the current “bunch crossing” although in principle there is already a way to use data from preceding or subsequent time units.

The situation is meant to change drastically with the upgrade of the CMS experiment to High-Luminosity LHC conditions. Not only will trigger data be available at higher resolution, including information from the silicon strip tracker (more precisely, the “outer” tracker up to a pseudorapidity of 2.4), but they will also contain timing information at higher resolution than the LHC bunch collision rate of 40 MHz. So far, the investigated physics scenarios have largely assumed that new particles created in a proton-proton interaction transform instantaneously into the final decay products (stable or long-lived “messenger” particles that can be detected by the CMS detector such as leptons, photons and hadrons). As the recent searches have not resulted in the discovery of the hypothetical particles predicted by Supersymmetry and other theories, it is mandatory to look for new scenarios. In particular, there are predictions for relatively long-lived particles (with lifetimes of many nanoseconds or more), which would be missed by exclusively looking for prompt decays. For such particles, it may be important to combine information from different times. A typical scenario could be that a particle travels from the interaction point to some part in the detector, such

as a calorimeter or the experiment's solenoidal magnet, where it comes to rest and decays after a certain time. So, a delayed signal in the muon system or the hadron calorimeter could be correlated to an earlier signal in the silicon tracker or the electromagnetic calorimeter. It could also be that the original particle is not seen in the detector (a neutral particle would not leave a signal in the tracker). The track of its decay product would in general not point back to the interaction vertex. Therefore it will also be imperative to reconstruct such decays without imposing on a track in the outer detector (the muon system) the condition that it must point to the interaction vertex.

The multitude of possible decay scenarios will make it hard to find optimal trigger conditions to cover all cases. Here, novel techniques such as machine learning will become very important. Decays predicted by certain theory scenarios will be calculated by Monte Carlo programs and used to “teach” the trigger to detect such decays in an approach of “supervised” machine learning.

Studies for such novel scenarios must in general not be limited to only one step of the Level-1 trigger. They have to track a mother particle and its daughter particles all the way from the interaction vertex to the point where they are stopped in the detector or leave it. Such studies will have to pay particular attention to the performance of the “Correlator”, whose task will be to combine the information on one particle or object obtained from the tracker and the other subdetectors (the calorimeters and the muon system). In this way, the particle flow approach, which uses all the information provided by various subsystems of the detector for reconstructing a given decay product and has proved to be extremely useful and powerful in data analysis will also enter the trigger world.

A further important task will be the implementation of such new triggering methods in the electronics firmware. While the processing in the so-called “High-Level trigger” (the computer farm that receives the detector data in full resolution after a Level-1 Accept has been issued, and whose task is the further reduction of the data volume to be sent to permanent storage) uses largely the same software routines that are also employed by physicists in the final data analysis, the processing algorithms to be implemented in the Field-Programmable Gate Arrays (FPGAs) of the fast electronics that make up the Level-1 trigger system have to be reformulated in a way that is closer to the pipelined architecture of these systems and takes into account the stringent latency requirements of this trigger stage. So far, hardware description languages such as VHDL or Verilog have been used by engineers to program the FPGAs in a manual way. For the future and already starting now, attempts are underway to use the approach of so-called “High-Level Synthesis” and base much of the algorithmic description on a high-level computing language such as C++. Such an approach might not only ease the task of programming the devices but also mitigate the risk of mistakes or of mismatches between the firmware and the computer-based programs.

4. Machine Learning for the Level-1 Trigger

So far, the Level-1 Trigger has used relatively simple criteria for accepting or rejecting an event, such as the transverse momentum and the coordinates (in azimuth and pseudorapidity) of object candidates (such as electrons/photons, muons, jets etc.) and correlations between such variables. In order to keep up with increasing collision rates while maintaining an adequate efficiency for physics objects, more sophisticated combinations of input variables will be required.

Due to the inclusion of additional detectors into the Level-1 Trigger path, such as the silicon tracker, more types of input variables will be available. In particular, the silicon strip tracking detector of the CMS experiment will deliver information for every bunch crossing and early enough to be used by the Level-1 trigger, allowing for a “particle flow” approach to be used at Level-1. There will still be a possibility to trigger on stand-alone calorimeter or muon objects without tracker information and this option will be pursued for objects with very high reconstructed transverse momentum. Such objects are relatively rare and therefore do not constitute a rate problem. On the other hand, there may also be interest to trigger on tracks only reconstructed in the silicon tracker and not leaving traces in other subdetectors, which may be important in certain physics scenarios involving “disappearing tracks” caused by “long-lived particles”.

With trigger algorithms getting more and more complex, physicists started implementing them by using machine learning techniques where the computer “learns” by receiving a large number of sample events with their input parameters and the desired decision (the “verdict”). This “knowledge” is used to calculate the coefficients for a number of matrices forming an “artificial neural network” that eventually allow to calculate the verdict from the input parameters. Such a network can involve arbitrarily complex correlations between these parameters that can be evaluated only by using a significant number of nodes distributed over several “hidden layers” of the network. Due to the increasing complexity, machine learning is playing an increasingly important role in this field.

5. Implementation of Neural Networks in hardware

However, implementing such a system in an online trigger environment such as the CMS Level-1 Trigger is not straightforward due to the “latency” requirements of the trigger system, which has to deliver its verdict during the short time during which the input data are available in the on-detector memories. One approach is to simply store the verdict for each combination of input parameters in a lookup-table (such a solution using Boosted Decision Trees has been successfully used in the Endcap Muon Track Finder of the CMS Level-1 Trigger [3]). While this approach is straightforward, it becomes impractical when too many variables are involved.

Another way to approach this problem is by simplifying the neural network and removing nodes that have no or little effect on the result. There are tools to automatically translate programs written in a high-level language (such as C++) to hardware description languages (such as VHDL or Verilog) for Field Programmable Gate Arrays (FPGAs) used in the trigger electronics. This approach called “High-Level Synthesis” is now also being used in CMS [4]. While the network is initially “trained” on a computer (usually by using Monte-Carlo data) the framework then allows the experts to simplify it to an extent that can be implemented in firmware.

Certain types of events with salient features (such as very high transverse momentum) will continue to be accessible by using simple conventional triggers. However, in order to keep thresholds on energy and transverse momentum at a reasonable level that does not create the danger of losing important events cause by “New Physics”, sophisticated trigger approaches involving artificial neural networks and machine learning will become increasingly important not only at the level of the computer-farm based event selection by the High-Level Trigger but also at the initial Level-1 Trigger implemented in custom processors and Field Programmable Gate Arrays.

References

- [1] CMS Collaboration. The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report. Tech. rep. CERN-LHCC-2017-013. CMS-TDR-017. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2283192>.
- [2] CMS Collaboration. “Particle-flow reconstruction and global event description with the CMS detector. Particle-flow reconstruction and global event description with the CMS detector”. In: JINST 12.CMS-PRF-14-001. CMS-PRF-14-001-004. 10 (2017), P10003. 82 p. DOI: 10.1088/1748-0221/12/10/P10003. URL: <https://cds.cern.ch/record/2270046>.
- [3] D. Acosta et al., Boosted Decision Trees in the Level-1 Muon Endcap Trigger at CMS; Compact Muon Solenoid Experiment Conference Report CMS CR -2017/357
- [4] J. Duarte et al., Fast inference of deep neural networks in FPGAs for particle physics; arXiv:1804.06913v3 (to be published in JINST)