Design of the CMS calorimeter trigger upgrade from Phase I to Phase II of the LHC

Alexandre Zabi, on behalf of the CMS Collaboration
Laboratoire Leprince Ringet, Ecole polytechnique, Route de Saclay, 91120 Palaiseau, France
E-mail: Alexandre.Zabi@cern.ch

Abstract. The CMS experiment implements a sophisticated two-level triggering system composed of the Level-1, instrumented by custom-design hardware boards, and the software High Level Trigger. In 2017, the LHC delivered proton-proton collisions at a centre-of-mass energy of 13 TeV with a peak instantaneous luminosity larger than $2 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$, more than twice the peak luminosity reached during Run I and far larger than the design value. The CMS Level-1 calorimeter trigger was upgraded during the long shutdown 1 between 2013 and 2015, 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, 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 dynamic clustering techniques, pile-up subtraction and isolation requirements for electrons and tau leptons. In addition, the new global trigger is capable of computing complex variables such as those involving the invariant mass of trigger objects. The trigger selections used for a wide variety of physics signals during Run II are presented, ranging from simple single-object selections to more sophisticated algorithms combining different objects and applying analysis-level reconstruction and selection. The design and operation of the Phase I calorimeter trigger will be reviewed. The technological choices made influenced the path towards the Phase II upgrade system necessary for the LHC run at a center-of-mass energy of 14 TeV with luminosity of $5 - 7 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$, corresponding to 140-200 pile-up events. The addition of the tracker information at Level-1 and the enhanced calorimeter granularity will be used to maintain the trigger object thresholds at a similar level as the present system.

1. Introduction
With the intense LHC running conditions in terms of instantaneous luminosity and pile-up events, the calorimeter structures and their associated readout electronics have grown in complexity. As a result, trigger architectures also became increasingly complex in order to provide sophisticated selection algorithms to ensure the highest possible acceptance for physics already at the hardware level of the data acquisition system. The key technological features of the CMS upgraded calorimeter trigger systems are the following:
  • The extensive use of large computing power (FPGA) to achieve an optimised reconstruction, identification, isolation and energy calibration of trigger candidates using high granularity information from the detector.
  • The use of high-speed optical links to achieve a global view of the detector for the purpose of the precise evaluation of global quantities such as missing transverse energy or the
implementation of pile-up mitigation techniques. Moreover, the selection of specific topologies such as Vector Boson Fusion production mode (VBF) is also possible.

- The implementation of a flexible and modular architecture, which can be expandable to adapt to other possible LHC running conditions and physics needs. Extra resources allow more sophisticated quantities to be computed to give a richer physics menu and increase selectivity.

The Phase I \(^1\) upgraded CMS Level-1 calorimeter trigger architecture will be described. The challenges related to its implementation and operation are presented. The calorimeter trigger algorithms and their performance measured on Run II collision (2015-2018) data are detailed. The Phase II \(^2\) trigger system will have to face even more intense running conditions. The use of highly granular information from calorimeters and trackers will have implication on the overall architecture design. The scalability of the Phase I calorimeter trigger system will be discussed. Examples of sophisticated calorimeter trigger algorithms will be presented along with the first implementation of higher-level trigger object reconstruction using a Particle Flow algorithm.

2. Upgrade to the Level-1 trigger system for the Run II of the LHC

The CMS trigger system is organised in two consecutive steps: the hardware-based Level-1 (L1) trigger utilizes coarse energy deposits in the calorimeters and signals in the muon systems to reduce the rate from 40 MHz to about 100 kHz; this is followed by the software-based High Level Trigger (HLT), implementing selection algorithms based on finer granularity and higher resolution information from all sub-detectors \(^1\). The upgrade to the Level-1 calorimeter trigger system is motivated by the need to preserve the trigger acceptance for physics until the end of Run III. Until then, no fundamental modification of the current primitive generation scheme is required. One of the main constraints comes from the total Level-1 bandwidth of 100 kHz that will remain unchanged. In order to avoid a significant increase in trigger energy thresholds, which would be detrimental for physics, an upgrade of the L1 trigger system was required \(^2\).

2.1. Conceptual choices for the upgraded trigger

The performance required to select collision events efficiently in a much harsher environment than the LHC Run I led to the conceptual choices discussed below. In order to sustain single physics object triggers, higher level of background rejection needed to be achieved through the implementation of sophisticated reconstruction and identification algorithms using the full tower granularity. Large FPGAs such as Xilinx Virtex-7 have been chosen to provide enough computational power. The evaluation of global quantities such as missing transverse energy or pile-up energy can be performed precisely by removing all boundaries. A total of 1152 10 Gbps high-speed optical links were installed to collect rapidly all calorimeter primitives hence providing a full field view of the detector. An increased selectivity of the system is required to cope with the constant raise of instantaneous luminosity and pile-up. The micro Global Trigger (\(\mu\)GT) was designed to be expandable to many more possible conditions and more sophisticated quantities, to give as rich a menu as the HLT. A flexible and modular architecture based on the \(\mu\)TCA telecom standard was selected to instrument the upgraded system. All hardware was replaced including the timing control system and all software and databases. One of the main key technological changes is the implementation of the novel TMT architecture \(^3\) visible on Figure 1. The upgraded trigger is organised in two consecutive processing layers. The layer-1 is composed of 18 Calorimeter Trigger Processor boards (CTP7) that are used to perform the pre-processing and data formatting. Tower level operations are executed such as the sum

\(^1\) The Phase I of the LHC refers to the period of running, which extends from 2009 to 2023

\(^2\) Phase II is also called the High-Luminosity LHC period starting in 2023 where we expect 5 to \(7 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}\) of instantaneous luminosity along with 200 pile-up events
of ECAL and HCAL energies and energy calibration as well as the computation of the H/E ratio. The 9 Layer-2 master processor cards (MP7) host all calorimeter algorithms that find particle candidates and compute global energy sums. Each MP7 has access to a whole event at trigger tower (TT) granularity. The benefit of time multiplexing is the removal of boundaries. The latency associated with the TMT depth allows to accommodate sufficient time for complex trigger algorithms to run. These algorithms are fully pipelined and start processing as soon as the minimum amount of data is received. The trigger candidates are sent to a demultiplexer board, also an MP7, that formats the data for the upgraded \( \mu \)GT. The \( \mu \)GT is responsible for allowing the event to be further scrutinised by the HLT.

![Figure 1. The upgraded Time-Multiplexed-Trigger architecture. The outputs of the Layer-1 CTP7 corresponding to one event are transmitted to single processing nodes in the second layer, Layer-2.](image)

2.2. The improved Level-1 Calorimeter trigger algorithms

The algorithms of the upgraded Level-1 trigger system have been designed to exploit the full trigger tower granularity and the global calorimeter view provided by the TMT architecture. Novel reconstruction techniques inspired by the offline algorithms have been implemented at the firmware level. Well-reconstructed single trigger objects can then be used to compute complex quantities and other correlated variables at the \( \mu \)GT level. Another challenge addressed by the new Level-1 system is the online determination of the pile-up energy without the information from the tracking. What follows present the algorithms and their performance. A more complete description is available in [3]

2.2.1. The electron, photon and tau lepton finders

In order to obtain a better containment of the electron energy, the reconstruction of the associated cluster should accommodate an extension in the \( \phi \) direction to recover energy loss through Bremsstrahlung emission. Inspired by the offline reconstruction, the dynamic clustering was designed to adapt the size of the cluster to precisely match the electron footprint in the calorimeter to optimize the trigger response. As compared to the Run I algorithm [1], the energy resolution is improved by almost 30% in the transition region between the barrel and the endcap where the tracker material is maximum. The enhanced granularity (Trigger tower instead of 4\( \times \)4 TT regions) allows to improve the trigger object position resolution. In the case of hadronically decaying \( \tau \) leptons, several energy clusters associated with each decay product may be produced. The dynamic clustering is used

---

3 The ratio of HCAL and ECAL energy for a given tower is used as a veto to discriminate hadron from electron energy deposits.
here to reconstruct individual clusters, which can be subsequently merged. An extra collection of isolated candidates can be produced for both e/γ and τ leptons. The isolation energy is evaluated as the energy deposited into ECAL and HCAL in a 6×9 TTs window around the cluster seed after subtracting the candidate’s E_T. The pile-up energy dependence of the isolation energy is taken care by introducing a threshold cut that depends on the transverse energy, the position in η and the multiplicity of trigger towers with non-zero E_T in the event. The cluster shapes produced by the e/γ and τ finder algorithms are categorized to provide further background discrimination. The collision data collected during 2017 have been used to study the performance of single object triggers. Figure 2 (left) shows the trigger efficiency curve for various isolated tau lepton thresholds as a function of the offline reconstructed tau p_T.

2.2.2. The jet finder and energy sums The reconstruction of jets is based on a 9×9 TTs sliding window centered around a local maxima. In order to avoid double counting of overlapping jets without efficiency loss, the trigger towers are required to satisfy a set of inequalities [3]. The size of the window is chosen to correspond to the cone radius of 0.4 used for the offline anti-k_T reconstruction algorithm. Other global quantities are computed with full calorimeter granularity such as the total E_T, the missing transverse energy (ME_T) and H_T, the jet-based equivalent of the total E_T as well as MHT, the jet-based equivalent of the ME_T. In order to preserve the energy resolution performance, a local pile-up correction technique called “chunky donut” is used to estimate the pile-up energy around the jet candidate on an event-by-event basis. An additional pile-up subtraction technique for missing transverse energy has been implemented to keep the thresholds low enough for physics. This technique consists of E_T thresholds applied to trigger towers entering the ME_T calculation, which depend on η and the tower size. Figure 2 (right) shows the performance of such technique on 2017 data.

Figure 2. (Left) Trigger efficiency curve for isolated τ lepton thresholds of 30, 34 and 38 GeV as a function of the offline reconstructed τ p_T. A data sample with Z → ττ → µτ selected events is used where the µ is the tag and the hadronic τ is the probe. (Right) Efficiency curves for the upgraded Level 1 missing transverse energy trigger using events with a single muon required offline as a function of the Particle Flow missing transverse energy reconstructed offline, which is the magnitude of the negative vector sum of the transverse momenta of all Particle Flow candidates reconstructed in an event excluding muons. Efficiency curves with and without pile-up subtraction are compared for trigger thresholds that give the same rate: L1 ME_T HF PUS on > 80/100/120/150 GeV produce the same rate as L1 ME_T HF PUS off > 103/125/144/170 GeV. ME_T HF refers here to ME_T computation including the HF Trigger Towers.

Figure 2. (Left) Trigger efficiency curve for isolated τ lepton thresholds of 30, 34 and 38 GeV as a function of the offline reconstructed τ p_T. A data sample with Z → ττ → µτ selected events is used where the µ is the tag and the hadronic τ is the probe. (Right) Efficiency curves for the upgraded Level 1 missing transverse energy trigger using events with a single muon required offline as a function of the Particle Flow missing transverse energy reconstructed offline, which is the magnitude of the negative vector sum of the transverse momenta of all Particle Flow candidates reconstructed in an event excluding muons. Efficiency curves with and without pile-up subtraction are compared for trigger thresholds that give the same rate: L1 ME_T HF PUS on > 80/100/120/150 GeV produce the same rate as L1 ME_T HF PUS off > 103/125/144/170 GeV. ME_T HF refers here to ME_T computation including the HF Trigger Towers.

Figure 2. (Left) Trigger efficiency curve for isolated τ lepton thresholds of 30, 34 and 38 GeV as a function of the offline reconstructed τ p_T. A data sample with Z → ττ → µτ selected events is used where the µ is the tag and the hadronic τ is the probe. (Right) Efficiency curves for the upgraded Level 1 missing transverse energy trigger using events with a single muon required offline as a function of the Particle Flow missing transverse energy reconstructed offline, which is the magnitude of the negative vector sum of the transverse momenta of all Particle Flow candidates reconstructed in an event excluding muons. Efficiency curves with and without pile-up subtraction are compared for trigger thresholds that give the same rate: L1 ME_T HF PUS on > 80/100/120/150 GeV produce the same rate as L1 ME_T HF PUS off > 103/125/144/170 GeV. ME_T HF refers here to ME_T computation including the HF Trigger Towers.
2.3. Operating the Phase I calorimeter trigger during Run II
Since the start of operation, the instantaneous luminosity has reached an unprecedented level that required regular adjustments to be made to the physics selection strategy. Thresholds as well as other parameters such as the lepton isolation and calibration were derived many times over. The LHC reached an instantaneous luminosity of $2.1 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$ (in 2017) along with 55 average pile-up events. An increase of the out-of-time pile-up was experienced due to the special bunch structure used during the fall of 2017. CMS was able to record collisions a total of 45 fb$^{-1}$ at high pace with very little time to perform optimizations. The single electron trigger threshold was kept below $\sim 34$ GeV and the double electron trigger required $\sim 25$ GeV and $\sim 14$ GeV thresholds on both legs. A double isolated $\tau$ lepton trigger was deployed with a threshold that remained under 32 GeV. The extension of the $\mu$GT hardware from 3 to 6 boards in the winter 2017 offered a lot more flexibility in the implementation of cross-triggers and topological triggers based on invariant masses for example. The first dedicated VBF and W-boson trigger algorithms were deployed in 2017. The 2017 trigger menu has implemented twice as much correlation triggers than the previous year in order to maintain the thresholds for physics. The flexibility of the system allowed to mitigate other effects induced by the ageing of the subdetectors such as noise or loss of response. These running experiences are used to guide the design of the Phase II system presented in the next section.

3. The Phase II trigger structural organisation and key features
The CMS Phase II trigger system is designed to benefit from the new features provided by the upgraded subdetectors to sustain high efficiency of physics event selection in very high luminosity regime. A possible architecture and data flow of the Phase II trigger system is presented in Figure 3. As for the Phase I system, information from calorimeters and muon detectors are used but the 12.5 $\mu$s latency budget allows the inclusion of tracking information. The total output bandwidth considered is 750 kHz [4]. Given the complexity and large data volume produced by the detector, a significant fraction of the computing takes place in the detector Back-End electronics. The key features of the proposed system is the introduction of a Correlator layer which implements sophisticated algorithms producing high-level trigger objects to achieve an enhanced selectivity, comparable to the HLT. The design exploits the availability of new technologies to ensure the needed computing power along with high-speed data transfer (up to 28 Gb/s) to provide a global detector view. The system will handle a huge amount of data (50 Tb/s compared to 1.8 Tb/s today). Inspired by the technological choices made in the Phase I, the processing units are not designed to perform specific tasks but are rather built as generic stream-processing engines allowing to consider original trigger architectural choices adapted to the requirement inherent of a flexible Correlator design. The algorithm implementation in firmware greatly benefits from the introduction of the High-Level-Synthesis (HLS) software that could be used to design the first machine learning trained variables or even iterative processes in the core of the trigger system. What follows describe in more details the inputs to the Phase II system and their utilisation by the sophisticated trigger reconstruction algorithms under development. Preliminary performance results for both standalone and higher-level objects are described.

3.1. Higher granularity calorimeter trigger primitive information
The Phase II upgraded calorimeter trigger will benefit from an enhanced granularity provided by the upgraded Barrel and Endcap calorimeters. The ECAL barrel (EB) Very Front-End (VFE) and Front-End (FE) electronics will be fully replaced with boards equipped with high-speed optical links (5 Gb/s, compare to 800 Mb/s in the Phase I system) in order to transmit the crystal energy information for triggering purposes. This modification will allow an increase of granularity by a factor 25. The HCAL barrel (HB) also plans to replace its on-board electronics...
readout to retrieve the depth information providing then 7 times more granular information
with respect to the previous system. The Endcap calorimeters will be replaced by the High-
Granularity Calorimeter (HGCAL) instrumented with over 6 million readout channels. It is
a 3D sampling calorimeter with alternating layers of silicon sensors/scintillator and copper/lead
absorbers. The goal is to achieve unprecedented spatial resolution and shower separation to
optimise the matching with tracks. About half of the 52 layers (28 electromagnetic and 24
hadronic), representing a total of 900,000 channels, can be exploited for triggering purposes
hence increasing the granularity by a factor of more than 500 compared to the Phase I system.
The HGCAL trigger primitives consist of the summed energy of 4 adjacent channels. One
of the key features of the Phase II trigger upgrade system is the addition of the tracking
information. Standalone calorimeter-based objects will suffer from high trigger rates and
therefore, reconstruction algorithms will greatly benefit from matching with tracks. Tracking
information can also be used to compute lepton isolation, improve jet and energy sum trigger
reconstruction, provide vertexing, estimate the level of pile-up (not only calorimeter-based
estimators as in Phase I) etc. CMS plans to replace entirely both the Pixel and the Tracker
detectors. The outer tracker on detector electronics will generate "stubs" that the track trigger
system will associate into track candidates to be transmitted to the Level-1 trigger system. The
overall latency is estimated to 5 µs (including 1 µs for data transmission). A total of 200
tracks with p_T > 3 GeV can be reconstructed per event at 200 pile-up in the Phase II LHC
environment.

3.2. Advanced calorimeter trigger algorithms and their performance
As for the offline object reconstruction, the Phase II trigger system can provide a large variety
of trigger objects such as standalone, which are based on single detector information (including
track-based objects), standalone objects matched to tracks and Particle Flow objects. The trigger
decision can rely on the complementarity of these objects to achieve the best possible
efficiency while maintaining the trigger rate under control.

Figure 3. High-level view of the Phase II Level 1 trigger. The main data flow is
shown with solid lines. Additional data paths are under study, including direct connections
from systems up-stream of the Correlator Trigger to the Global Trigger, and paths that
allow Tracker data to be passed to the Muon Triggers. Shown: Outer Tracking Detector
(TRK), the Endcap Calorimeter (EC), the ECAL Barrel (EB), the HCAL Barrel (HB),
the HCAL Forward Detector (HF), the Muon Drift Tube Detectors (DT), the Resistive
Plate Chambers (RPC), the Cathode Strip Chambers (CSC), the Gas Electron Multiplier
Chambers (GEM).

4 A similar upgrade has been successfully performed on the forward HCAL (HE) in 2018. The entire readout
boxes were replaced and equipped with SiPM replacing the existing Hybrid Photo-Diodes. The use of the depth
information in the trigger algorithms are being studied with Run II data
5 An HGCAL channel represents a readout cell of roughly 1 x 1 cm^2
6 Stubs are computed from closely-spaced silicon-sensor modules building each tracker layers. The bending of
tracks between each side of the modules is used to discard low p_T tracks.
3.2.1. **Triggering on electrons, tau leptons and jets**

Many standalone object reconstruction techniques are currently being investigated to optimise both the response and the position resolution for the purpose of realising the highest possible track matching efficiency. The algorithms propose both identification criteria and isolation variables based on calorimeter as well as tracking information to reduce the background level. Given the intense running conditions foreseen, the algorithms are designed to be pile-up resilient. Taking the electron finder case for example, in order to evaluate the gain in using crystal information in the ECAL Barrel compared to trigger towers, a simple $5 \times 3$ crystal matrix is used to define the maximum size of the electron footprint in the ECAL. An improved position resolution is achieved by calculating with a weighted-energy sum around the seed crystal (The seeding threshold used is $E_T > 1$ GeV). Extra shower shape features are used as identification criteria and a matching of the clusters with tracks is performed using an extrapolation to the ECAL surface. Figure 4 shows the performance of the algorithm in terms of rate compared to the Phase I algorithms. These results demonstrate the large rate reduction achieved with track matching while keeping high efficiency (> 90%). The reconstruction of hadronically decaying tau leptons cannot rely on calorimeter only information as in Phase I. To demonstrate this, the algorithm considered here, uses Phase I type clusters (see section 2.2) that are confirmed by tracks to build tau lepton candidates. A significant rate reduction is achieved by introducing of a track-based isolation around the tau candidate and primary vertex requirement. The Endcap Calorimeter trigger algorithms are designed to exploit the 3D information provided by the HGCAL. The associated Back-End electronics are organised in two consecutive layers: Stage 1, reconstructing 2D clusters on each layer from trigger-cells produced by the Front-End electronics boards [5] and Stage 2, combining 2D clusters into 3D clusters. The trigger architecture is Time-Multiplexed to provide extra latency to implement sophisticated algorithms. Along with 3D clusters, the longitudinal shower profile is exploited to distinguish electromagnetic from hadronic objects. The pile-up profile per layer differs from that of the signal and is also used to reduce the rate. Figure 4 shows the performance of the jet algorithm in the HGCAL with 200 pile-up events. Standalone HGCAL jets are formed from multiple 3D clusters using an anti-$k_T$ algorithm (cone of 0.2). The lateral granularity can be used to exploit the dense core of jets to mitigate the effect of pile-up. Pile-up corrections and calibration are details in [5].

3.2.2. **High-level calorimeter trigger objects using Particle Flow algorithms**

The availability of tracking information and the enhanced calorimeter granularity at the Level-1 trigger allow us to contemplate the exciting possibility to propose algorithms matching the performance of those implemented at the HLT. The benefits of introducing a full event description such as Particle Flow reconstruction algorithm has been demonstrated successfully with the Phase I HLT system. The Phase II trigger Correlator processor (see Figure 3) will be used to combine all detector information to produce a list of candidates from which, higher-level trigger objects are constructed such as identified prompt leptons, photons and jets, as well as global quantities such as missing transverse energy or hadronic transverse energy sum ($H_T$). In addition to the Particle Flow algorithm, a simplified version of the offline algorithm, called PUPPI for pile-up per particle identification (Ref in [4]) can be implemented to mitigate the degradation of the energy resolution caused by pile-up. Figure 4 shows the trigger efficiency of the $H_T$ as a function of the generated quantity using the combination of the Particle Flow and PUPPI algorithms. The study was conducted using minimum-bias collisions and $t\bar{t}$ simulated events decaying semileptonically, corresponding to an average PU of 140. A successful implementation of these algorithms in firmware (Xilinx Ultrascale+VU9P FPGA) was performed using HLS.

---

7 As in the Phase I algorithm approach, the extension in $\phi$ is motivated by the necessity to recover energy loss through bremsstrahlung.

8 The total estimated resources using the Vivado HLS (v2016.4) is about 30%.
4. Conclusion
The Phase I calorimeter trigger is showing excellent performance. The thresholds for single objects are maintained low enough for the CMS physics program to be carried out. The flexibility of the system is demonstrated by its capabilities to adapt to higher instantaneous luminosity and harsher conditions by introducing more topological variables. The impact of the technological choices and the running experience acquired with this system most certainly guide the design of the Phase II system. The introduction of finer calorimeter granularity and tracking information allows to consider the implementation of higher-level object reconstruction using Particle Flow algorithm leading to enhanced selectivity.

References
[1] S. Chatrchyan et al. (CMS Collaboration), The trigger systems CERN/LHCC 2000-38 CMS-TDR-006-1
[2] S. Chatrchyan et al. (CMS Collaboration), The level-1 trigger upgrade CERN-LHCC-2013-011 CMS-TDR-12
[3] A. Zabi et al. (CMS Collaboration) The CMS Level-1 Calorimeter Trigger for LHC Run II JINST 12 C01065
[4] S. Chatrchyan et al. (CMS Collaboration) The Phase-2 Upgrade of the CMS Level-1 Trigger Interim Technical Design Report CERN-LHCC-2017-013 CMS-TDR-017
[5] S. Chatrchyan et al. (CMS Collaboration), The Phase-2 Upgrade of the CMS endcap calorimeter CERN-LHCC-2017-023 CMS-TDR-019 and T. Strebler et al. (CMS Collaboration) these proceedings