The CMS High Level Trigger

Daniele Trocino
Northeastern University, 360 Huntington Ave, 111 Dana Research Center, Boston, MA 02115, USA
E-mail: Daniele.Trocino@cern.ch

Abstract. The CMS experiment has been designed with a two-level trigger system: the Level-1 Trigger, implemented in custom-designed electronics, and the High-Level Trigger (HLT), a streamlined version of the CMS offline reconstruction software running on a computer farm. A software trigger system requires a tradeoff between the complexity of the algorithms running with the available computing power, the sustainable output rate, and the selection efficiency. We present the performance of the main triggers used during the 2012 data taking, ranging from simple single-object selections to more complex algorithms combining different objects, and applying analysis-level reconstruction and selection. We discuss the optimisation of the trigger and the specific techniques to cope with the increasing LHC pile-up, reducing its impact on the physics performance.

1. The CMS Trigger System

The Compact Muon Solenoid (CMS) experiment [1] has been designed with a two-level trigger system. The Level 1 Trigger (L1) is based on FPGA and custom ASIC technology, and reduces the LHC collision rate, about 16 MHz in 2012, to the maximum rate allowed by the CMS readout system, 100 kHz. The system has a fixed latency of less than 4 µs (3.2 µs in 2012) and uses information from the calorimeters and muon spectrometer to accept or reject an event.

The High Level Trigger (HLT) is implemented in software running on a farm of commercial computers, including over 13,000 CPU cores, and reduces the L1 output rate to a sustainable level for storage and physics analysis, about 1 kHz in 2012. The HLT software consists of a streamlined version of the offline reconstruction algorithms, optimised to comply with the strict time requirements of the online selection. The average processing time in 2012 was about 200 ms per event, two orders of magnitude less than the offline reconstruction. The HLT splits the selected events into various non-exclusive streams, with different purposes and event content. Most of the bandwidth is devoted to data for physics analysis. Streams of data with limited event content are also saved for data quality monitoring (DQM), calibration, and trigger studies.

2. The HLT Software

The HLT must ensure a large acceptance for physics signals, while keeping the output rate and CPU time under control. This is achieved by exploiting the same sophisticated software used for offline analysis, which ensures high reconstruction efficiency and allows for stringent identification criteria and quality selections to reduce the trigger rate.

The modular structure of the trigger paths helps to keep the processing time within the tight requirements of the online selection. A trigger path is a sequence of reconstruction and filtering...
blocks of increasing complexity: faster algorithms are run first and their products are filtered. If a filter fails, the rest of the path is skipped. Another important feature used to reduce the CPU time is the regionality of the HLT: the detector read-out and reconstruction are restricted to narrow regions around the L1 or higher-level candidates.

3. Instantaneous Luminosity and Pile-Up
During the 2012 LHC run, the instantaneous luminosity and average pile-up reached peaks of about $7.5 \cdot 10^{33} \text{cm}^{-2}\text{s}^{-1}$ and 35 interactions per bunch crossing. As the average processing time increases linearly with the pile-up, the HLT farm was extended and upgraded. The current farm includes more than 13,000 CPU cores with hyper-threading and can sustain luminosities up to $8 \cdot 10^{33} \text{cm}^{-2}\text{s}^{-1}$.

Pile-up also affects the performance of many triggers. The additional tracks and calorimeter deposits increase the rate of jet triggers and reduce the efficiency of lepton isolation. The rate of multi-object triggers also increases due to combinatorial effects. In order to mitigate these effects, minimum $p_T$ thresholds and vertex constraints are applied to tracks and other object constituents. The average energy density is measured in each event and subtracted from jet areas and lepton isolation cones [2].

Due to the intense irradiation, the lead-tungstate crystals of the CMS electromagnetic calorimeter (ECAL) lose transparency, compromising the efficiency of electron and photon triggers. A sophisticated laser system [3] is used to monitor the transparency of each crystal and compute energy corrections. In 2012, these corrections were updated weekly and applied to the ECAL endcaps, where the effect of the transparency loss was more relevant.

4. Particle Flow in HLT
“Particle flow” (PF) [4] is a reconstruction technique widely used in CMS offline analyses. It uses the full detector information to describe the global collision event by identifying particles individually and clustering them into more complex objects. In 2011 PF was introduced in the HLT for $\tau$ lepton reconstruction, and in 2012 it was extended to jets and missing transverse energy (MET). The use of PF improved the energy resolution of trigger objects, increased their efficiency with respect to offline selection, and provided more refined methods for pile-up mitigation.

5. Performance of HLT Objects
This section shows the reconstruction and identification performance of the main physics objects used in HLT paths.

5.1. Muons
Muon reconstruction starts with track fitting in the outer muon spectrometer. The outer tracks are then used to seed track reconstruction in the inner tracker. Matching inner-outer track pairs are finally combined and fitted together. Muon isolation can be required by combining tracker and calorimeter information, and subtracting the average energy from pile-up, as explained in Section 3.

Figure 1 shows the efficiency of an isolated muon trigger as a function of $p_T$, before and after the introduction of pile-up energy subtraction. The efficiency reaches a plateau of about 90% and the pile-up corrections make the turn-on sharper. Figure 2 compares the cross sections of a non-isolated muon trigger in 2011 and 2012. In 2011, a pile-up-dependent fake rate caused the cross section to increase with luminosity. In 2012 the fake component was suppressed with a tighter muon identification, based on track quality and impact parameter cuts, restoring a constant cross section. This also improved the trigger purity (Figure 3), reaching 85% for isolated triggers, stable with pile-up.
Figure 1. Efficiency of an isolated single-muon trigger with $p_T$ threshold of 24 GeV/c and $|\eta| < 2.1$, before (black squares) and after (red triangles) applying pile-up corrections, as explained in Section 3. Note the sharper turn-on curve with pile-up subtraction.

Figure 2. Cross sections of a non-isolated single muon trigger with $p_T$ threshold of 40 GeV/c in 2011 (blue open circles) and 2012 (red full circles). A tighter muon identification in 2012 restores a flat cross section.

Figure 3. Purity of an isolated single-muon trigger with $p_T$ threshold of 24 GeV/c, defined as the fraction of triggered events that satisfy typical offline muon selections used in analysis. The overall purity is about 85% and stable with pile-up.

5.2. Electrons and Photons
Electrons and photons are reconstructed from clusters of ECAL deposits with isolation requirements in the tracker and hadronic calorimeter (HCAL). For electrons, ECAL clusters are also required to match a track in the pixel detector, with cuts on the cluster energy to track momentum ratio. Further identification criteria can be applied, based on isolation, cluster shape, etc. Figure 4 shows the efficiency of a single-electron trigger versus $E_T$ in the ECAL barrel, and in the endcaps before and after transparency corrections. It can be seen that in the endcaps laser corrections are essential to ensure full efficiency and a sharp turn-on, and to keep low thresholds. Figure 5 shows the efficiency of two single-photon triggers with different identification criteria versus $E_T$, and their combination (logical OR). The total efficiency is close
to 100% and pile-up independent.

Figure 4. Efficiency of a single-electron trigger with $E_T$ threshold of 33 GeV/c for barrel ($|\eta| < 1.5$, black triangles), and for endcaps ($|\eta| > 1.5$) before (red open circles) and after (blue full circles) applying transparency corrections. Applying such corrections brings the turn-on curve closer to the barrel one.

Figure 5. Efficiency of photon reconstruction in HLT with different identification criteria, based on calorimeter isolation (red squares), cluster shape (blue triangles), and the OR of the two paths (black circles). The combination of the two identification criteria recovers all inefficiencies and brings the plateau close to 100%.

5.3. Hadronic Taus

Hadronic $\tau$'s are reconstructed using the PF algorithm. Figure 6 shows the efficiency of a single-$\tau$ object as a function of $p_T$. The turn-on is sharp and the plateau above 90%. Typical triggers, e.g. used in the search for a Higgs boson decaying to a $\tau$ pair, require a hadronic $\tau$ in association with an isolated electron or muon.

5.4. Jets and MET

Jets and MET can be reconstructed from ECAL and HCAL deposits (calorimetric jets/MET) or with PF reconstruction (PFJets, PFMET). The latter are characterised by better resolution and sharper turn-on curves, especially at low energies, where the tracker resolution exceeds that of the calorimeters. Figure 7 shows the efficiency of single-jet triggers with different thresholds, comparing PF and calorimetric jets.

6. Data Parking and Scouting

During 2012, CMS collected on average 400 Hz of data that were promptly reconstructed for physics analysis. In addition to these “core” data, about 600 Hz were recorded and temporarily saved in raw format at CERN’s Tier-0 and Tier-1 grid computing centres, in order to be processed and analysed during the 2013-14 LHC shutdown, when more computing power would be available [5]. These “parked” data are intended to extend the CMS physics programme, providing more statistics for precision measurements or new physics searches. The trigger paths used to collect these data are either looser versions of core triggers, or new paths extending to kinematic regions not covered by core triggers. Examples include triggers for Vector Boson Fusion, multi-jet final states, and low-mass muon pairs for B-physics.
Figure 6. Efficiency of a single-\(\tau\) trigger with \(p_T\) threshold of 20 GeV/c in data (black circles) and simulation (blue diamonds).

Figure 7. Efficiency of single-jet triggers with \(p_T\) thresholds of 320 (blue), 370 (red), and 400 GeV/c (black). The first and last paths use PF jets, while the middle path (red) uses calorimetric jets. The PF jets have better resolution, and thus sharper turn-on.

In addition to data parking, CMS extended its physics reach by also using “data scouting” [5]. This strategy allowed CMS to explore kinematic regions characterised by very high rates for Standard Model processes, e.g. low-\(p_T\) or low-mass regions, and thus normally excluded from data taking. A dedicated trigger path was designed to keep events with \(H_T\) (scalar sum of jet transverse energies) greater than 250 GeV. This trigger has a rate of the order of 1 kHz. In order to keep the bandwidth usage at a manageable level, the content of the events recorded by this trigger was reduced by roughly a factor of \(10^4\), including only HLT-reconstructed jets, and no raw data nor offline-reconstructed objects. With this information, simple analyses can be performed in a DQM-like framework, searching for possible deviations from the Standard Model. Should any indication of new physics arise, CMS would have the possibility to change the trigger menu and dedicate more bandwidth to record fully-reconstructed events and perform detailed studies. The data scouting was put in place for analyses like di-jet final state, razor, and \(\alpha_T\) searches.

7. Summary and Conclusions
The CMS HLT system was designed to select a great variety of physics signals, as well as data for detector monitoring and maintenance. Thanks to data parking and scouting, CMS was able to record events even beyond its standard rate constraints. With the challenge of increasing luminosity and pile-up in 2012, the HLT showed a remarkably stable behaviour. The current computer farm allows a maximum processing time of 200 ms per event at 100 kHz L1 input rate, and is able to sustain an instantaneous luminosity up to \(8 \cdot 10^{33} \text{cm}^{-2}\text{s}^{-1}\). The full software implementation and sophisticated algorithms give the CMS HLT a great flexibility and adaptability with respect to the evolving LHC conditions. All HLT objects showed high efficiencies, sharp turn-on curves, and stable rates and cross sections. Table 1 summarizes the thresholds and rates of the main triggers used in 2012. Figure 8 shows the rates of core, parked,
| Object                      | Threshold [GeV] | Rate at 6 Hz/nb [Hz] | Physics                      |
|-----------------------------|-----------------|----------------------|------------------------------|
| Single muon                | 40              | 24                   | Searches                     |
| Isolated single muon       | 24              | 48                   | Standard Model               |
| Double muon (17, 8)        |                | 14 [18]              | Standard Model, Higgs        |
| Single electron            | 80              | 10                   | Searches                     |
| Isolated single electron   | 27              | 65                   | Standard Model               |
| Double electron (17, 8)    |                | 9                    | Standard Model, Higgs        |
| Single photon              | 150             | 6                    | Searches                     |
| Double photon (36, 22)     |                | 7                    | Higgs                        |
| Muon + electron (17, 8)    |                | 13                   | Standard Model, Higgs        |
| Tau + isolated muon (20, 17)|           | 12                   | Higgs                        |
| Single PFJet               | 320             | 12                   | Standard Model               |
| Quad Jet (4 × 80)          |                | 9 [129]              | Standard Model, Searches     |
| Six Jet (6 × 45)           |                | 3.4                  | Searches                     |
| MET                        | 200             | 1.8                  | Searches                     |
| HT                         | 750             | 7                    | Searches                     |

Table 1. Summary of thresholds and rates of the HLT menu for luminosity up to 7 Hz/nb.

and full trigger menus as functions of the instantaneous luminosity. Thanks to the stability of all triggers with pile-up, the rate shows a linear dependence.

Figure 8. Total rate of the 2012 HLT menu as a function of the instantaneous luminosity. “Core” and “parked” trigger menus are also shown separately. The rate exhibits a linear dependence on the luminosity. The step around 6.5 Hz/nb is due to trigger prescales.
References

[1] CMS Collaboration, “The CMS experiment at the CERN LHC”, JINST 3 (2008) S08004.
[2] Cacciari M and Salam G P, “Pileup subtraction using jet areas”, Phys. Lett. B 659 (2008), 119–126.
[3] Anfreville M et al., “Laser monitoring system for the CMS lead tungstate crystal calorimeter”, Nucl. Instr. and Meth. A 594 (2008), 292–320.
[4] CMS Collaboration, “Commissioning of the Particle Flow Event Reconstruction with the First LHC Collisions Recorded in the CMS Detector”, CMS-PAS-PFT-10-001 (2010).
[5] CMS Collaboration, “Data Parking and Data Scouting at the CMS Experiment”, CMS-DP-2012-022 (2012).