CMS Conference Report

14 July 2004

The CMS High-Level Trigger Selection

Domenico Giordano

*Dipartimento di Fisica dell'Università degli Studi di Bari e INFN Sezione di Bari,
Via Amendola 173, 70126, Bari, Italy*

on behalf of CMS Collaboration

Abstract

The CMS High-Level Trigger (HLT) is designed to reduce the Level-1 accept trigger rate of 100 kHz to a final output rate of 100 Hz, at which selected events are written to permanent storage. In order to satisfy the very high performance requirements of CMS, the HLT will consist of a farm of commercial processors performing the online event filtering. The flexibility provided by a fully programmable environment implies that algorithmic changes can be easily introduced to improve the selection of various physics channels, as well as to deal with unforeseen experimental conditions. The filter farm software includes all major features of the offline reconstruction code. In this report, the selection algorithms and physic performance of the CMS HLT are presented.

*Presented at 9th Topical Seminar on Innovative Particle and Radiation Detectors
Siena - Italy, May 23-26, 2004*

1 Introduction

The Compact Muon Solenoid (CMS) [1] is a general purpose detector that will operate at LHC [2], a proton-proton collider with a centre-of-mass energy of 14 TeV and a bunch crossing rate of 40 MHz. At the design luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, each crossing results in an average of about 20 inelastic pp events producing approximately 1 MB of zero-suppressed data. These figures are many orders of magnitude larger than the archival storage capability of $O(10^2)$ Hz at data rates of $O(10^2)$ MB/s. For this reason, CMS needs of an excellent trigger system in order to reduce the event rate to the storage rate, while maintaining a high selection efficiency for the physics phenomena that the experiment intends to study.

2 CMS Trigger System

The CMS trigger consists of two physical levels: a Level-1 trigger [3] and a High-Level Trigger (HLT) [4].

The Level-1 trigger receives data at the full LHC bunch crossing rate of 40 MHz and takes the trigger decision for each bunch crossing after a latency time of $3.2 \mu\text{s}$. During the latency time, the full detector data are stored in front-end pipeline memories. The output rate is limited by the capabilities of the CMS data acquisition system to 100 kHz. The Level-1 runs on custom synchronous processors and has access to a coarse granularity information from calorimeters and muon detectors.

The HLT is the second step of the trigger chain. It is designed to reduce the Level-1 output rate of 100 kHz to a final output rate of $O(100)$ Hz. The HLT code runs on a farm of commercial processors and performs the reconstruction and selection of physics objects using the full event data with fine granularity and matching information from different sub-detectors. These objects are electrons and photons, muons, missing energy, hadronic jets, tagged jets, the selection/rejection of which allows the output rate of $O(100)$ Hz to be achieved. Data from the front-end electronic modules are assembled by an event builder switching network that dispatches complete events to the processing nodes of HLT farm by means of asynchronous protocols. The switching network has a high bandwidth of 1 Tbit/s.

The choice of executing the HLT on a single processor farm provides the maximum flexibility and modularity to the trigger system, because it has neither architectural nor design limitations other than the total bandwidth and CPU that the experiment can acquire. In a fully programmable environment, the physics algorithms have the maximum freedom in what data to access and in the complexity of the reconstruction tools. Algorithmic changes can be easily introduced to improve the selection of various physics channels as well as to deal with unforeseen experimental conditions. Moreover, the system can benefit from the evolution of network, processor and memory technologies. The use of commercial components, wherever requirements are satisfied, results in the reduction of production and maintenance costs.

The system is sufficiently modular to deal with the evolution of both the accelerator and detector performance ensuring, at any time, the minimal requirements adequate to fulfil the CMS physics program.

3 Physics requirements for HLT

The HLT code runs on a single processor for a given event, and has to select/reject the event after a total processing time less than 1s. The real-time nature of the selection imposes significant constraints on the resources that the algorithms can use. The reliability of these algorithms is a key issue, since the events rejected by the HLT are lost forever.

The main physics guidelines that direct the HLT code development are the following:

- the HLT has to fulfil the CMS physics program with high efficiency;
- the event selection must be inclusive in order to select unexpected new phenomena;
- selection conditions must be independent of the detailed knowledge of calibration constants and other run conditions;
- trigger and reconstruction efficiencies must be measurable from data alone;
- events selected should be tagged in order to indicate the reasons for their selection;
- the code should be as close as possible to the offline reconstruction code.

4 HLT Reconstruction and Selection

Minimizing the CPU time required to process each event is useful to discard background events as soon as possible. Therefore, the HLT reconstruction and selection is arranged in a chain of virtual trigger levels, which consist of algorithms of increasing complexity and CPU time consumption: (i) the Level 2, which uses calorimeter and muon detector information; (ii) the Level 2.5, which additionally uses the tracker pixel information; (iii) the Level 3, which accesses the full event information, including all tracking detectors. At the end of each level a set of selection criteria reject a significant fraction of the events selected by the previous step.

Other selection strategies used in implementing the software are the *reconstruction on demand* and the *regional reconstruction*. Physics objects in an event are reconstructed only if they are requested, and the track reconstruction is performed only in a region of interest of the detector. The region of interest is pointed out by objects reconstructed in the previous trigger level. Additionally, the concept of *conditional reconstruction* allows more CPU time to be saved during track reconstruction. Because the ultimate resolution is not needed at HLT, the reconstruction is performed with a reduced number of hits, and is stopped as soon as the desired resolution is achieved.

4.1 Electrons and photons

The HLT selection of electrons and photons proceeds in three steps.

At Level 2 electron/photon candidates are reconstructed using only the calorimeter information with the full detector resolution. The reconstruction is performed in a region specified by the Level-1 candidates. The key issue in the calorimetric reconstruction is to include the energy radiated by electrons and converted by photons in the material between the interaction point and the ECAL. In fact, when an electron radiates in the tracker material, the energy detected by the ECAL is spread in the ϕ direction because of the bending of the electron in the 4T magnetic field. Therefore, the initial electron energy can be collected by making a cluster of clusters (“super-cluster”) along a ϕ road. The super-cluster is required to fall within the fiducial region of the ECAL, and to have an E_T above a given threshold.

The next step, the Level 2.5, demands hits in the pixel detectors consistent with a Level-2 candidate. The expected hit position on the pixel layers is estimated by propagating the electron inwards to the nominal vertex using the magnetic field map. The matching of an ECAL super-cluster with at least two of three pixel hits divides the electromagnetic trigger into two streams: electron candidates (single and double), and photon candidates in case of unmatched clusters. This procedure is highly efficient and provides a large rejection factor, since most of electron bremsstrahlung and photon conversions take place after pixel detectors. The rate of photon candidates is further reduced by applying higher threshold cuts than in the electron stream. The selection of photons can use isolation cuts, rejection of π^0 s based on lateral shower shape and reconstruction of converted photons.

In the final step, the Level 3, the selection of electrons uses full track reconstruction, seeded from the Level-2.5 pixel hits. To maintain high efficiency, the track finding is made with very loose cut parameters. Cuts are then made on E/p and on the distance between the super-cluster position and the extrapolated track position in the ECAL. In the endcap, a cut on the energy found behind the super-cluster in the HCAL, expressed as a fraction of the super-cluster energy (H/E), is found to give useful additional rejection.

4.2 Muons

The muon selection proceeds in two steps. At Level 2, muon candidates are reconstructed in the muon chambers, using the full detector resolution. The algorithm performs a regional reconstruction starting from the Level-1 candidates. An iterative Kalman Filter method is used to reconstruct track segments and extrapolate to the closest detector layer, where local reconstruction is performed only in the neighbours of extrapolated state. The trajectory building works from inside out, then the track fitting is performed from the outermost muon station inwards, including the interaction region. The p_T resolution obtained for muons from W decays is 10% in the barrel region ($|\eta| < 0.8$), 16% in the endcaps ($1.2 < |\eta| < 2.1$) and 15% in the overlap region. The key issue is to take into account the propagation of tracks through the iron yoke, in a not uniform magnetic field and with multiple scattering effects.

The Level-3 muon reconstruction consists in extending the muon trajectories to include the full tracker system. A region of interest is defined around the Level-2 track after vertex constraint. The trajectory building starts from pixel hits found in the region of interest, and performs the track reconstruction using the Kalman filter technique. Trajectories are grown in parallel, and are rejected as soon as their χ^2 exceeds a fixed threshold, or if no hit is

found in more than four consecutive layers. Finally, the selected trajectories are fitted including also the hits of the Level-2 candidate reconstructed in the muon detectors. The improvement in muon resolution respect to Level 2 is substantial. For muons coming from W decays, the resolution is 1% in the barrel, 1.7% in the endcaps, and 1.4% in the overlap region.

After each step, isolation criteria are applied to the muon candidates in order to suppress muons from b, c, K and π decays, which are produced in jets. Two isolation techniques have been studied: (i) calorimeter isolation, applied at Level 2; (ii) pixel and tracker isolation applied at Level 3. The calorimeter isolation is based on the energy sum in a cone around the muon in the (η, ϕ) plane, the track isolation algorithms are based on the measurement of the sum of the p_T of the tracks around the muon. The isolation effect on the muon rate is higher for muons below 30 GeV/c, where the rate is dominated by b, c, K and π decays.

4.3 Jets and Missing Energy

The HLT jet reconstruction is done using a simple and fast iterative cone algorithm in the (η, ϕ) plane. The algorithm looks for a protojet in a cone around the direction of a seed calorimeter tower. The direction and the E_T of the protojet are updated recursively, by computing the E_T -weighted angles of the towers in a cone around the protojet direction, and the sum of towers energies in the cone. The procedure is repeated between iterations until the energy of the protojet changes by less than 1% and the direction of the protojet changes in the (η, ϕ) plane by less than 0.01, or until 100 iterations are reached. Jet finding does not exploit the regional reconstruction strategy. It is currently carried out using all calorimeter towers instead. The offline algorithm will be based on more sophisticated algorithms that use more information and improved energy resolution.

The bandwidth allocated for the HLT selection of events with only one, three and four jets is 9 Hz at low luminosity ($2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$). Due to the high jet rates at LHC, the thresholds for the event selection based only on jets are also very high, but they can be reduced at an acceptable level with additional trigger conditions, as missing transverse energy cut or some lepton selection.

The identification of neutrinos and other non-interacting particles is performed using the calorimeter information to look for missing transverse energy (E_T^{miss}). The current algorithm calculates E_T^{miss} as a vector sum of the tower energies over a threshold of $E_T(\text{tower}) > 500 \text{ MeV}$. Studies at generator level show that most of the E_T^{miss} rate below 60 GeV is due to the finite coverage of the calorimeter ($|\eta| < 5.0$), while above 60 GeV the rate is dominated by neutrinos produced in the decays of heavy flavours. The missing E_T selection is done in association with other requirements, like an energetic jets or a lepton.

4.4 τ tagging

The High-Level Trigger algorithms for τ identification are designed to be used in the selection of isolated τ leptons expected as a signature of MSSM Higgs boson decays as $A/H \rightarrow \tau\tau$ and $H^\pm \rightarrow \tau\nu$. The final-state signatures involve events with a lepton and a τ jet, two τ jets or only one τ jet. The τ lepton decays hadronically 65% of the time, producing a τ jet, which is a narrow jet-like cluster in the calorimeter containing a relatively small number of charged and neutral hadrons. In 77% of hadronic τ decays, the τ jet consists of only one charged hadron and a number of π^0 s (one prong decays).

The identification of τ jets employs information from both the calorimeter and tracking detectors. At Level 2, τ -jet identification involves the calorimetric jet reconstruction and isolation in a narrow region centred on the Level-1 candidate τ jet. At Level 3, the τ tagging is performed with tracker detectors, using the regional and conditional reconstruction. The direction of a candidate τ jet is defined by the axis of the Level-2 calorimeter jet. The track-finding algorithm reconstructs all tracks candidates inside a matching cone around the jet direction, and looks for a leading signal track. Any other track in a narrow signal cone around the leading track is assumed to come from the τ decay. Then the isolation condition is applied, by requiring no tracks outside the signal cone above a threshold in p_T . The event is rejected as soon as possible, if no leading track is found, or if the isolating criterion is not fulfilled. The algorithm gives an efficiency of about 50% with a background rejection of 10^3 .

4.5 b tagging

The goal of b tagging is the inclusive selection of b jets, obtained by benefitting from the large impact parameters of b-hadron decay tracks with respect to the production vertex. First the jet tracks are reconstructed, using a regional approach, within a cone around the Level-1 calorimeter candidate, and the jet direction is measured again as the

p_T -weighted sum of the track directions. The jet is tagged as a b jet if at least two tracks exceed a threshold on the impact parameter significance (defined as the ratio between the value of the impact parameter and its uncertainty). The threshold value is chosen according to the jet- E_T range, in order to ensure a 55% efficiency for b jets and a rejection factor of about 10 almost independently of the jet E_T .

5 HLT Performance

Table 1: HLT requirements at low luminosity ($2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$) for various channels: selection thresholds, corresponding storage rates, CPU time required for the selection, Level-1 trigger rates. The CPU figures refer to a 1 GHz Intel Pentium-III CPU. Computing time and Level-1 rate are cumulative for e - γ and for jets - jet $\times E_T^{\text{miss}}$ channels.

Trigger	HLT Threshold (GeV)	HLT Rate (Hz)	CPU time per Lvl-1 event (ms)	Lvl-1 Rate (Hz)
1e, 2e	29, 17	33, 1	160	4.3
1 γ , 2 γ	80, (40,25)	4, 5		
1 μ , 2 μ	19, 7	25, 4	710	3.6
1 τ , 2 τ	86, 59	3, 1	130	3.0
jet $\times E_T^{\text{miss}}$	180 $\times$ 123	5	50	3.4
1, 3, 4 jets	657, 247, 113	9		
e $\times$ jet	19 $\times$ 45	2	165	0.8
Inclusive b jets	237	5	300	0.5

At the LHC start-up luminosity of $2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$, CMS will have a DAQ system capable of reading a maximum of 50 kHz of events accepted by the Level-1 Trigger. Actually, only 16 kHz of the total 50 kHz bandwidth are allocated. The remaining part is a safety factor to take into account the rate underestimate due to all possible simulation uncertainties.

Table 1 shows the HLT requirements (threshold, allocated bandwidth, processing time) at low luminosity for the selection of various simulated streams, in order to have a cumulative storage rate of $O(10^2)$ Hz. The selection is highly efficient for the benchmark physics channels. Nevertheless, it remains inclusive by avoiding specific topological requirements.

Due to the real-time nature of the HLT selection, a key issue is the CPU power required for the execution of the algorithms. The fourth column in Table 1 reports the CPU time needed to process events on a Pentium-III 1 GHz CPU. The current requirements vary from 50 ms for jet reconstruction to 700 ms for muon reconstruction. Weighting the CPU needs of the algorithms by the frequency of their application (the Level-1 Trigger rate), a mean CPU time of 271 ms is found per event that passes the Level-1 Trigger. This mean time implies that the CMS HLT farm must consist of 15,000 CPUs, in order to run the HLT with 50 kHz input rate. Extrapolating these figures to the LHC start-up (2007), on the basis of Moore's Law [5], the CPU units are expected to be a factor height more powerful than at the time of these studies. Therefore, at the LHC start up, the HLT system will need of about 2,000 CPUs.

6 Summary

The HLT code will run on a farm of commercial processors performing the online event filtering. The selection algorithms and physic performance of CMS HLT have been described and show that the system can fulfil the CMS physics program with high efficiency, providing a selection of 1:1000. The fully programmable environment is suitable to design a flexible and modular system, which can deal with the evolution of both the accelerator and detector performance, as well as with unexpected new physics phenomena.

References

- [1] CMS Collaboration, "The Compact Muon Solenoid, Technical Proposal", **CERN/LHCC 94-38**.
- [2] The LHC Study Group, "The Large Hadron Collider, Conceptual Design", **CERN/AC/95-05**.

- [3] CMS Collaboration, " *The TriDAS Project Technical Design Report, Volume 1: The Trigger Systems*", **CERN/LHCC 2000-038**.
- [4] CMS Collaboration, "*Data Acquisition and High-Level Trigger Technical Design Report*", **CERN/LHCC 2002-26**.
- [5] G.E. Moore, "*Cramming more components onto integrated circuits*", **Electronics, 38 , 114 (1965)**.