Frascati Physics Series Vol. XXXVIII (2005), pp. 73–87 II WORKSHOP ITALIANO SULLA FISICA DI ATLAS E CMS – Napoli, October 13 - 15, 2004

THE HIGH LEVEL TRIGGERS IN ATLAS AND CMS

Alessandro Di Mattia Università "La Sapienza" and INFN Marco Zanetti Università and INFN Padova

Abstract

The High Level Trigger (HLT) systems of ATLAS and CMS provide a software based event selection after the initial Level-1 hardware trigger. It is implemented as software tasks running on large processor farms, and the foreseen rejection factor is about 10^3 for both experiments. Besides this commonalities, ATLAS and CMS have different approaches for its design which originate from an opposite philosophy of the DAQ architecture. An overview of the two architectures is presented together with examples of online reconstruction algorithms. The different trigger strategies foreseen for the LHC runs and their impact on the physics program are also discussed. Finally some remarks concerning the HLT systems commissioning are given.

1 Introduction

The ATLAS $^{(1)}$ and CMS $^{(2)}$ experiments are omni-purposes detectors that will operate at the Large Hadrons Collider (LHC), a p - p collider with a center of mass energy of 14 TeV and a design luminosity ranging from 2×10^{33} cm⁻²s⁻¹ (low luminosity phase) to 10^{34} cm⁻²s⁻¹ (high luminosity phase). Although having substantial different designs, ATLAS and CMS have to face the same extreme challenges provided by the collider, in particular the reduction of the data rate approximately by a factor of 10^6 , from 40 MHz, corresponding to the interaction frequency, down to a value of the order of 100 Hz which is manageable by the persistent storage and the off-line processing. Both ATLAS and CMS split the decision whether to retain or not an event, into different stages. In both cases, the first step (Level 1), given the high interaction rate and the limited buffering capabilities, exploits algorithms implemented on custom hardware boards which process only a subset of the event data (namely from muon detectors and calorimeters) with coarse granularity and lower resolution. The output event rate of the Level 1 is designed to be of the order of 100 kHz. Subsequently, the High Level Triggers (HLT) provide a software based event selection to further reduce the Level-1 rate to ≈ 100 Hz. At the high luminosity phase, this value corresponds to a cross section of ≈ 10 nb which matches the prediction for $\sigma(pp \to W^+ \to \mu^+ \nu)$; it is therefore evident how critical is the role of the HLT which has to perform *online* the first actual physics analysis.

2 DAQ and HLT Architectures

Being the identification and selection of the interesting events a real time procedure, the reliability of the entire DAQ and HLT systems is a critical issue for both the experiments in order to reach their full physics potential. The main guidelines that define the design and the development of the HLT of ATLAS and CMS are based on the following considerations:

- the performances and the working conditions of the collider and of the detectors are not precisely predictable and vary with time; in this sense the flexibility of the system is a crucial issue;
- the HLT should fulfill the physics programs as well as be inclusive enough

in order not to reject unexpected phenomena;

- the system should be robust, i.e. it should not rely too much on the changes of the calibration and detector alignment constants;
- to maximize the efficiency of the filtering process, an uninteresting event should be rejected from the data flow as soon as possible;
- it must be possible to validate the trigger and to compute the overall selection efficiency using only the data, relying as less as possible on simulation;
- the HLT should benefit of all the major features of the offline software, in particular it should be as close as possible to the offline reconstruction code;

Moreover, being the experiments life extended over 20 years, the long term maintainability is a critical issue too. Previous experiences have shown that custom electronics is more difficult to maintain and upgrade than comparable commercial products. Therefore the use of commercial computing and network equipment is required, because it helps to maintain the system over the full life time of the experiment.

2.1 ATLAS Architecture

After the initial Level-1 selection, the data coming in parallel from the detector readout are handled by the HLT/DAQ system. Taking into account the estimates of the bandwidth that will be used by the ATLAS subdetectors, the total readout bandwidth is of about 200 GBytes/s. This constitutes a formidable challenge for the switching network that accounts for the data transmission through the entire system: the use of a bare network schema would require to implement thousands of connections with a data throughput that reaches the performance limit of the current commercial equipment. Moreover, since the estimate of the total bandwidth depends on the working conditions of the detectors (luminosity delivered by LHC, fluctuations on the Level-1 rate, detector occupancy), the network must be dimensioned to face the fluctuations of the data rate.

To cope with this problem, the ATLAS architecture $^{3)}$ is designed to minimize the data movement towards the HLT processors. At the early stage



Figure 1: The ATLAS Trigger/DAQ system.

of the HLT processing, the event data fragments are held separately in memory buffers (the ROBIN) and wait for the decision whether to be assembled into a complete event for the final selection or to be discarded. This constitutes an intermediate step of the HLT trigger, which is called Level-2. The trigger algorithms make use of the full granularity data but access only the detector regions that has been flagged by the Level-1 as those containing the physics candidates (the Region of Interest, RoI). The RoI based access allows to use only 2% of the event data to take the Level-2 decision, thus limiting the required bandwidth of the dataflow. A further reduction of the data traffic is obtained by multiplexing the detector readout into the ROBINs in such a way as to gather together data from detector elements which are projective aligned towards the interaction vertex. This increases the probability to access a single ROBIN per RoI. The design foresees that the Level-2 will reduce the Level-1 rate by a factor of about 30 with a latency of about ~ 10 ms. To achieve this rejection power in such a small time the Level-2 algorithms will perform an approximate reconstruction of the physics candidate avoiding the use of the fully detailed calibration constants of the detectors.

After the Level-2 decision, the data are delivered to the Event Filter processor farm to take the final decision on the event. At this stage the algorithms employed are derived from the offline reconstruction software and access to the full event data. The increase of the reconstruction accuracy provides a rejection of the Level-2 accept rate of about a factor 10. But the use of more sophisticated procedures, which make use of complete detector calibrations, requires more time to execute the algorithms. Therefore the latency time of the Event Filter is estimated to be about ~ 1 s.

Both the Event Filter and the Level-2 algorithms run into a common software framework which reuses part of the offline software components. In particular all the interfaces towards the data, and the code providing the detector description and calibration is implemented by the same offline tools. This eases the development and the study of the selection algorithms optimizing the manpower and, at the same time, increases the long term maintainability of the code. But the use of pure offline components into the online environment clashes sometimes with the latency requirement thus requiring their replacement with highly optimized code, especially in the Level-2 environment.

The general schema of the ATLAS Trigger/DAQ is shown in figure 1. The detector data flows through ~ 1600 Read Out Drivers (RODs) into the Read Out System (ROS). This is realized by standard PCs hosting several ROBIN boards. The data concentration into the ROS's is about 10, therefore ~ 150 machines will be employed to buffer the data within the Level-2 decision. On the contrary the Level-1 trigger data are processed by the RoI Builder (ROIB), and then sent to a Level-2 Supervisor (L2SV), which manages the Level-2 operations. The Level-2 algorithms run into the Level-2 Processing Units (L2PU) of the trigger farm which is made of ~ 500 biprocessor machines. In case of a "Level-2 accept", the result is stored into the pROS as a part of the event data and then is sent through the Event Builder and the SFI to the "EF farm" together with the ROS's data under the supervision of the DataFlow Manager (DFM).

2.2 CMS Architecture

The main feature and peculiarity of the CMS architecture $^{4)}$ is that the Data Acquisition system performs the HLT event selection in a single farm of com-

mercial processors, the Filter Farm (FF). This design principle has been established in order to take advantage of the extraordinary and constant rate of evolution in computing technology either in processing power and in network speed. Avoiding a physical intermediate level in the selection chain allows the HLT to be entirely software implemented, and to access full resolution and granularity data as well as calibration and alignment monitored constants. More in details, the data from each detector front-end belonging to a "L1-accepted" event, are collected by a set of Read-out Units and then delivered to the Builder Units (BUs) through a large switching network (Read-out Builder Network). The network bandwidth required is of the order of 1 Tbyte/s. The BU receiving the data fragments is responsible for the actual building of the event and serves it to a Filter Unit (FU) via another switching system, the Filter Farm Network. The FUs are the components of the FF where the HLT code is executed and the selections are applied¹. The selected events are then forwarded to the computing services for storage or for further analysis. A picture of the CMS DAQ architecture is shown in Fig. 2. A key feature of the FF is that the raw data is delivered to the FU only if requested by the specific HLT algorithm; this allows to reduce either the data traffic and the HLT processing time. The computing power needed by the FF is estimated to be as high as $10^6 SI95$ corresponding to $\mathcal{O}(10^3)$ GHz processors. Given the average event size of the order of 1 MB and the manageable output rate of $\mathcal{O}(10^2)$ Hz, 1Tbyte of data will be written on the mass storage every day.

3 Reconstruction Algorithms

The HLT selections are based on the precise and efficient reconstruction of the physical objects, i.e. e, γ, μ, τ , *jets* and *b-jets*. The higher the stage of the selection chain, the larger the amount of data and the greater the time available by the reconstruction algorithms. In the following, as examples of the way physical object are measured and identified by ATLAS and CMS HLT systems, two reconstruction algorithms will be illustrated.

¹Actually the event building and filtering is performed in eight independent slices, each one of which could in principle perform the whole precess by itself. This address the issue of the scaling of the system



Figure 2: Schema of CMS Data Acquisition system.

3.1 Online muons reconstruction in ATLAS

The HLT muon selection in ATLAS comprises the Level-2 trigger and the Event Filter. The Level-2 identifies muon objects and estimates their physics properties with a set of optimized algorithms. At this stage the event selection is mostly based on the physics of the single muon, therefore the algorithms are tuned to provide the best physics performance while reconstructing the muon features around the selection threshold values. On the contrary, the Event Filter employs a complete muon reconstruction program, based on offline packages, that provides very good performance over the full spectrum of muon events. This allows to select the events using also tight invariant mass criteria. The task of the Level-2 muon trigger can be decomposed into a number of broad steps: validation of the Level-1 muon RoI, combination of the muon track with the Inner Detector tracks, check for isolation in the calorimeter and recovery of the very low- $p_{\rm T}$ muons not triggered by the Level-1 (i.e. search for secondary RoIs). The aim of the first step is to reject the fake Level-1 triggers and to operate a first reduction of the Level-1 rate by means of a more precise measurement of the muon transverse momentum² $(p_{\rm T})$. The algorithm doesn't make use of time consuming fit methods: muon hits are recognized by means of geometrical criteria and the track is reconstructed with a set of linear segments fitted on each muon station. Nevertheless the resolution of the transverse momentum reconstruction is 6% at the low- $p_{\rm T}$ threshold (6 GeV) and 4% at the

²The better quality of the momentum measurement, with respect to that provided by Level-1, allows for a sharper $p_{\rm T}$ threshold.

high- $p_{\rm T}$ threshold (20 GeV). The good quality of the momentum measurement allows to reduce the Level-1 input rate by a factor of about 2 at 6 GeV and by a factor of about 10 at 20 GeV. The rejection of the Level-1 fake triggers is about 10^3 and is provided by requiring at least two segments per track. After the RoI is confirmed, the selected sample is refined depending on the given trigger threshold. For high- $p_{\rm T}$ muon triggers the calorimeter energy deposition around the track direction is analyzed to confirm the track is isolated. On the contrary, the low- $p_{\rm T}$ muon triggers undergo to a sharp refinement, which involve the use of the Inner Detector data. Extrapolating backward the muon flight direction, a small slice of Inner Detector is identified where to search for tracks compatible with the muon one. The matching candidate allows to refine the estimate of the muon $p_{\rm T}$ thus yielding a reduction factor of about 2.5 of the Level-1 low- $p_{\rm T}$ rate. A further reduction is obtained requiring the muon event is compatible with a $J/\psi \to \mu^+\mu^-$ decay. In this case a wide region of the Inner Detector is reconstructed to search for the second decay muon track. To be identified as a muon, the Inner Detector track is demanded to be consistent with hits in the innermost MDT station. A loose cut on the invariant mass of the dimuon system is also applied. The Inner Detector data are also employed to confirm the high- $p_{\rm T}$ triggers, but the criteria used are less stringent because the rate is not demanding.

Being seeded by the Level-2 result, the muon algorithm in the Event Filter starts to reconstruct the spectrometer data. This standalone reconstruction is implemented with offline algorithms that make use of combinatorial technique to identify the muon hits and involve the complete magnetic field map in the fit. To improve the performance on high- $p_{\rm T}$ momentum reconstruction, the multiple scattering effect along the track path is recovered taking into account the material distribution crossed by the muon. As in Level-2, the muon track is propagated backward to the Inner Detector to search for the muon hits. Once identified, these are entered in the global fit to refine the track reconstruction. After the combined reconstruction the complete definition of the muon track through the ATLAS detector is available to be used for the final trigger menu selection. Altogether the Event Filter reconstructs the muon momentum with a resolution close to that provided by the offline program, i.e. ~ 2.5% for muons up to $p_{\rm T} \simeq 200 GeV$. The sophisticated techniques employed by the Event Filter require a large amount of CPU time that, at present, clashes with the

latency requirement. To limit the CPU usage, the reconstruction is performed only in a wide region around the muon RoIs. Optimization studies are ongoing to see if it is possible to execute the standalone reconstruction over the full spectrometer.

3.2 Online electrons and gammas reconstruction in CMS

The CMS HLT selection of electrons and gammas proceeds in three steps. At the beginning, namely $L2^3$ the electron/photon candidates are reconstructed exploiting only the calorimetric information with the full granularity. The reconstruction is performed in the regions indicated by the Level 1 candidates. In order to recover the energy radiated by electrons and converted by photons in the tracker material, "super clustering" algorithms are used.

The second step, L2.5, demands hits in the pixels vertex detector consistent with a L2 candidate. The expected hit position on the pixels layers is estimated by propagating inward the energy weighted average impact point of the candidate to the nominal vertex position. If at least two hits are found, the candidate is classified as an electron, otherwise as a γ . The rate of the photons candidate is further reduced applying higher thresholds energy cuts than in the electron stream. The γ selection can also use isolation requirements, lateral shower shape for π^0 s rejection and reconstruction of converted photons. In the final step, L3, the algorithm to select the electron candidates has enough time to use the tracker hit in order to perform a full track finding and reconstruction. Cuts are then applied on E/p and on difference in η between the extrapolated track and the supercluster position. Isolation is required for both electrons and photons.

4 Triggers

A major issue concerning the HLT selection is *what* to save permanently on the mass storage. More precisely, the question about which trigger streams has to be settled and how much bandwidth needs to be allocated for each of them has to be addressed. The answer is of course a compromise between the request of

 $^{^{3}}$ the name L2 does not represent a physical layer in the HLT chain. Given the entirely software implementation of the HLT, an arbitrary high number of intermediate stages could be implemented

	ATLAS		CMS	
Streams	Thresholds	Rate	Thresholds	Rate
	(GeV)	(Hz)	(GeV)	(Hz)
Single μ	20	40	19	25
Double μ	10	10	7	4
Single e	25	40	29	33
Double e	15	< 1	17	1
Single Photon	60	25	80	4
Double Photon	20	2	40,25	5
Single Jet, 3 Jets, 4 Jets	400, 165, 110	30	657, 247, 113	9
$Jet + missing E_T$	70, 70	20	180, 123	5
τ jet + missing E_T	35, 45	5		
au jet			86	3
Double $ au$ jet			59	1
$e + \tau$			19, 45	2
b-jet			237	5
<i>b</i> - <i>physics</i>	topological	10		
Prescaled, calibration		20		10
Totals		200		105

Table 1: ATLAS and CMS High Level Trigger tables.

maximal efficiency for the physic program and the total bandwidth and CPU power available. It also depends strongly on the phase of the experiment and the actual conditions of the LHC and the detectors.

ATLAS and CMS performed the exercise⁴ of listing a feasible set of selections for standard LHC running condition at $\mathcal{L} = 2 \times 10^{33} \text{ cm}^{-2} \text{s}^{-1}$. The established outcome is summarize in Table 1. Within the streams listed in Table 1, tree kind of triggers can be identified, each one responsible of a different part of the physics program.

⁴Neither the thresholds nor the associated rates must be taken as the final ones; it is probable that a better knowledge of the detectors and of the phenomenology will lead to a better tuning of the selection cuts. Moreover the associated rated should be compared only by order of magnitude since they rely on different physics assumptions



Figure 3: Output of dimuon stream of CMS HLT for $h \to W^+W^- \to \mu^+\nu\mu^-\bar{\nu}$ and main backgrounds.

4.1 Inclusive Triggers

By means of these, most of the physics program of the experiment will be covered as well as eventual unexpected phenomena at the TeV scale. In normal LHC run condition, most of the bandwidth will be dedicated to such streams. One of the golden processes for the Higgs boson discovery at the LHC is the decay $h \to Z^0 Z^0 \to e^+ e^- e^+ e^-$. The ATLAS HLT will selected these events with high efficiency in the single and double isolated electron streams (97% for event with 2 *e* with $p_T > 20 GeV/c$, $|\eta| < 2.5$).

In the case the Higgs particle has a mass within the range $[150-170] \ GeV/c^2$, because of the favorable branching ratio, the decay into W boson pairs dominates and the processes $h \to W^+W^- \to \mu^+\nu\mu^-\bar{\nu}$ becomes one of the most appealing channel for a fast discovery. In CMS the efficiency for selecting these kind of events with μ s within detector acceptance ($p_{\rm T} > 3 GeV/c$, $|\eta_{\mu}| < 2.5$) ranges from 93% for $M_h = 150$, to 95% for $M_h = 170$. In Fig. 3 the contribution of various processes to the "dimuon" stream is shown. The invariant mass of the dimuon system is reconstructed by the L3 muon algorithm.

4.2 Exclusive Triggers

Some interesting physics processes have kinematic features that do not allow the standard inclusive triggers to select them efficiently. In order not to loose the chance to investigate such phenomena, dedicated triggers are then needed to save those kind of events for off-line analysis. The most important example is the physics related to the quark beauty whose rare decays could reveal new scenarios beyond the Standard Model. The most suitable final states, i.e. for a hadronic collider like the LHC the muonic ones, although having peculiar topological and kinematic properties, are mainly populated by low $p_{\rm T}$ muons, usually well below the single and double inclusive thresholds. In order to retain these kind of events, the HLT selection exploits part of the inner detectors information from the beginning, otherwise utilized only at the last stage.

At low luminosity $(\mathcal{L} < 2 \cdot 10^{33} \text{ cm}^{-2} \text{s}^{-1})$ the ATLAS B-physics High Level Trigger is initiated by a low- $p_{\rm T}$ muon trigger, confirmed by Level-2, in combination with an electromagnetic trigger or a jet trigger. Then, a fast track algorithm performs an unguided search for tracks in a wide region of the inner detector. These tracks serve as seeds for the semi-exclusive reconstruction of the interesting decay channels, e.g. $B_d \to \pi^+\pi^-$, $D_s^- \to \Phi(K^+K^-)\pi^-$, $J/\Psi \to \mu^+ \mu^- (e^+ e^-)$. It proceeds combining couple of opposite charged tracks to identify a specific parent particle on the basis of the invariant mass. The combinatorial background is reduced by means of cuts on the scalar sum of the transverse momenta and on the difference of the z-intercept of the two tracks. The minimum $p_{\rm T}$ required to a track for entering in this procedure depends on the decay channel: it is 4 GeV for events with pions, but it drops down to 2 GeV for events with electrons. These latters are the most challenging for the track reconstruction and imply the use of the TRT data at Level-2 to identify the electrons down to very low- $p_{\rm T}$. The boundary between the Level-2 and the Event Filter is set by the techniques used to improve the track reconstruction quality. The recovery of the electron bremsstrahlung and the primary vertex reconstruction are CPU intensive calculations which are executed at the Event Filter stage. By means of these the Event Filter can apply a tighter invariant mass cut on the sample selected by the level-2.

The CMS dedicated HLT selections for *b*-physics are described in Ref ⁴). An important example of those is the exclusive trigger for the decay $B_s \rightarrow \mu_+\mu_-$. It happens in two stages, the first one based only on the vertex detector measurements, the second exploiting the whole tracking system. As first step a fast track and vertex reconstruction is performed looping on pixel hit pairs; starting from the obtained track-seeds, the algorithm execute the track reconstruction in the tracker. If and only if two opposite charged tracks are founded, the dimuon invariant mass is required to met the B_s mass within a range of $\pm 150 \ MeV/c^2$. In order to suppress combinatorial background, χ^2 and decay length transverse distance criteria are applied. The selection efficiency for the signal is 33.5% with an average execution time of 240 ms on 1 GHz CPU.

4.3 Prescaled, calibration and monitoring triggers

These triggers will play a crucial role for understanding, validating and debugging, either the detectors and the first LHC data during the first months of running.

The prescaled triggers are meant to extend the physics coverage of the online selection by enlarging the kinematic reach of the various measurements, e.g. towards smaller values of transverse momentum. A typical example is the measurement of the jet cross section over the full kinematic range, starting from the lowest achievable E_t value up to the region covered by the inclusive trigger. Prescaled triggers will be also crucial for determining the trigger efficiency from data, e.g. via bootstrap methods.

Calibration of the various subdetectors and monitoring of their performances are critical issues for every kind of physics measurement. As an example a strategy for a fast interacalibration of the different parts of the electromagnetic calorimeter must be developed. Ref ⁵) addresses the latter item. The method suggested there makes use of the ϕ -symmetry of deposited energy to intercalibrate the CMS electromagnetic calorimeter (ECAL) crystals within rings at constant η^5 . Single jet events triggered by L1 with a threshold of 120 Gev are used, the region within $\Delta R < 1.0$ of the trigger jet being excluded to avoid the most obvious trigger bias. Of these events, only the ECAL data are processed by the HLT, where the threshold is raised to 150 GeV. A dedicated high frequency (1 kHz) bandwidth is allocated for this calibration trigger. Eleven million jet trigger events, i.e. few hours of data taking during low luminosity phase, are sufficient to perform the intercalibration to a precision between 2% and 3%, depending on η .

⁵This method needs to be used in conjunction with another method to intercalibrate the ϕ rings - $Z^0 \rightarrow e^+e^-$ has been suggested

5 ATLAS and CMS commissioning

In order to be ready for the data taking and analysis in 2007 at the scheduled LHC startup, the certification and checking of the functionalities, the expected performances of the various sub-detectors and the detector as a whole must start well in advance. This should be done by means of real data to the maximum extent. In the period before the LHC installation ATLAS and CMS plans foresee to exploit cosmic muons. Later, as soon as the collider will provide the first single beam, the beam halo and the particles produced in the interactions of the beam with the gas in the vacuum pipe will be used too.

During this commissioning phase, the DAQ/HLT system carries out a twofold role: as part of the detector, it has to be commissioned as well as the other subdetectors, moreover it is a crucial tool for the commissioning of the latters. The first step is to verify the correctness of the data flow. In this context the detectors front ends synchronization, the event building from the data fragment and the actual event data flow through the HLT chain are the main issues. Because of the additional physical layer in the HLT system, ATLAS will be more focused on the precise understanding and debugging of the latter item, in particular of the RoI mechanism. On the contrary the most critical feature of the CMS HLT, the event building, will be stressed only when high rate of sizable events will be provided by the Level 1 trigger, i.e. during standard LHC runs. Because of the very low event rate in this phase no selection will be applied by the HLT. This allows part of the online reconstruction algorithms to be tested and debugged. The HLT commissioning will be completed during the LHC p-p runs, when the whole infrastructure will be in place on the basis of physics performances, i.e. efficiency for interesting processes and rejection for minimum bias and underlying events.

6 Conclusions

The ALTAS and CMS HLT systems, although based on different approaches, the former on two physical layers in order to reduce the data throughput, the latter implemented on a single processors farm for exploiting the maximum flexibility, are designed to reduce the LHC event rate to $\mathcal{O}(100)$ Hz with high efficiency for the whole physics program. In both experiments the HLT system development has being going on successfully and it is ready to be commissioned during the incoming period of cosmic muons tests and the first days of LHC activity.

References

- 1. the ATALS collaboration, Technical Proposal, CERN/LHCC 98-4 (1998).
- 2. the CMS collaboration, Technical Proposal, CERN/LHCC 94-38 (1994).
- the ATLAS collaboration, ATLAS High-Level Trigger Data Acquisition and Controls Technical Design Report, CERN/LHCC 2003-022, ATLAS TDR 016 (2003).
- the CMS collaboration, Data Acquisition and High-Level Trigger Technical Design Report, CERN/LHCC 2002-26 (2002).
- 5. D. Futyan and C. Seez, J. Phys. G: Nucl. Part. Phys. 29 (2003) 1299-1326