January 17, 2018

CMS Level 1 Trigger Rate Upgrade for Heavy Ion Collisions Status Report

Austin Baty[1], Ivan Amos Cali[1], Gian Michele Innocenti[1], Yen-Jie Lee[1], Camelia Mironov[1], Gunther Roland[1], Zhaozhong Shi[1], Jing Wang[1], Ta-wei Wang[1], Boleslaw Wyslouch[1]

[1]Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139
[2]CERN
[3]LLR

Contents

1 Introduction

Relativistic heavy ion collisions allow experimental studies of the quark-gluon plasma (QGP), the equilibrated high-temperature state of deconfined quarks and gluons. At the Large Hadron Collider (LHC), the abundant production of hard probes such as vector bosons, heavy quarks, jets, and quarkonia produced in the highest energy heavy ion collisions has opened a new era in the characterization of QGP, providing new information on initial state properties, parton energy loss and color screening. The capabilities of the Compact Muon Solenoid (CMS) detector in the detection of high momentum photons, electrons, muons, charged particles, and jets have proven to be a unique match for the opportunities at the LHC. A key aspect of this success has been the convergence of detector needs for precision measurements in p-p discovery physics in a high luminosity, high pileup environment and for QGP studies in high multiplicity Pb-Pb collisions. This convergence has allowed CMS to adapt a large number of new analysis techniques to studies of heavy ion collisions, such as particle-flow jet reconstruction, studies of jet substructures, lifetime fits for secondary J/ψ studies, b- and c-tagging of jets, missing p_{T} measurements of W production and energy flow in dijet events.

The full power of these and other techniques will be exploited in future heavy ion studies beginning in 2018 and Run III, when an increase in the collision energy to $\sqrt{s_{NN}} = 5.02$ TeV and an eventual increase in Pb-Pb collision rate to as high as $\sim 30-50$ kHz will enhance the production rate of hard probes by more than an order of magnitude. In this high energy and high luminosity era, CMS will undertake precision studies of complex observables such as γ-, Z^0- and $\mathrm{W}^{\pm}$-jet correlations, heavy flavor jet quenching, multi-jet correlations and the azimuthal anisotropy of high p_{T} jets and quarkonia. Recently, extensive studies of fully reconstructed heavy flavor mesons such as D^0 and B^+ in heavy ion collisions also show that CMS is ideally suited for the studies of heavy flavor production and heavy quark energy loss in the quark-gluon plasma.

The CMS trigger and data acquisition system is the key to achieving the ultimate high p_{T} physics reach in heavy ion collisions. Uniquely, the CMS trigger system only has two main components, the hardware-based Level-1 (L1) trigger and the High-Level Trigger (HLT), which is implemented as "offline" algorithms running on a large computer farm with access to the full event information. The triggering on high p_{T} probes for Pb-Pb collision rates up to several kHz (i.e. the design value for Pb-Pb) is possible using event rejection solely or mostly at the HLT level. In the 2011 Pb-Pb run, this allowed a reduction of the event rate to storage by more than an order of magnitude compared to the collision rate, without prescaling the most interesting high p_{T} observables. In the 2015 Pb-Pb run, the heavy flavor meson trigger was deployed for the first time. Those triggers increased the high p_{T} D^0 meson statistics for example in p-p (Pb-Pb) collisions by a factor of 600 (20), opening a new era for the precision heavy flavor physics in heavy ion collisions.

It has now become clear that the LHC will be able to significantly exceed the Pb-Pb design luminosity in future runs, possibly reaching up to 50 kHz already in Run III. This will place an even greater emphasis on the CMS trigger and DAQ system. The CMS during Pb-Pb collisions is being operated close to its hardware limit in terms of readout rate (see Sec. 2 for details). As a consequence, CMS will not be able to benefit from the increased luminosity without a dedicated upgrade. In 2017, a new 4-layer pixel detector was installed in CMS that will greatly improve the impact parameter resolution of the charged tracks. This is a unique opportunity to combine the capability of CMS for the studies of high p_{T} probes with the heavy flavor program down to $p_{\mathrm{T}} \sim 0$, complementary to the physics goal of the major ALICE upgrade foreseen for Run III. Moreover, a stage-2 upgrade of the Level-1 trigger system will be commissioned for Pb-Pb data-taking. This requires improvement in the front-end detector readout bandwidth in Pb-Pb collisions, as well as significant development on the trigger strategy for data-taking in 2018 and beyond, such that CMS could provide unbiased single track, heavy flavor meson, jet triggers that are critical for the CMS physics program, and at the same time record a large statistics Minimum-Bias triggered sample that could be used for studies of low p_{T} heavy flavor mesons. This necessitates the development of underlying event background subtraction for the stage-2 L1 trigger system and improvements in the detector readout to increase the maximum L1 trigger accept rate during the Pb-Pb data-taking period.

In this report, we will present the CMS L1 rate upgrade design and demonstrate that the planned upgrade delivers the needed capability to record heavy flavor mesons and jets over a very wide kinematic range at the full delivered rate in high luminosity Pb-Pb runs in 2018 and beyond. This new system will need to be commissioned in 2017-2018 before the next heavy-ion running period at the end of 2018, to make successful Pb-Pb data-taking possible.

2 Hardware Upgrade

2.1 Introduction

The CMS experiment is mainly designed to record high luminosity, high pile-up (PU) proton-proton collisions events. However, with several detector adjustments/configuration changes over the last several years, the CMS experiment has been able to record successfully also the high multiplicity heavy ion collisions delivered by the LHC. The main changes required to allow CMS to operate in the HI environment can be summarized as follows:

- A dedicated silicon pixel Front End Driver (FED) firmware for a high multiplicity environment. The overall FED data flow has been re-designed, changing the readout scheme of the input links and redistributing internal buffers.

- The silicon tracker zero-suppression was bypassed and all channel information was forwarded to the CMS DAQ system. Zero-suppression was then performed in the HLT. A new tracker FED firmware was also designed to reduce the event payload by $\sim 30\%$ without information loss. The reduction was obtained by stripping unused information out from the standard data format.

- DAQ reconfiguration / rebalance to deal with big data volume and high throughput. In standard p-p collision operation, an average event size corresponds to $\sim 1 - 1.5$ MB. During Pb-Pb collision operation, the average event size at HLT input is around 17 MB (14 MB of the event are due to the reading out of all channels of the silicon tracker).

- The Level-1 trigger firmware was redesigned to perform a dedicated underlying event subtraction algorithm and to include specific HI algorithms.

- Significant adjustment in the overall CMS dataflow.

With the specific configuration described above, and with a series of other small adjustments not mentioned here, CMS was able to run smoothly during the 2015 Pb-Pb data-taking with an average L1 rate of ~ 10 kHz. A few hours of the Pb-Pb data-taking period were also devoted to exploring the limit of CMS in terms of the maximum L1 rate allowed. With the 2015 CMS configuration for Pb-Pb, the absolute L1 rate limit was observed at 12 kHz. In the next sections, this limit, as well as solutions to increase it, will be discussed.

During the 2015 run, the collision rate was ~ 20 kHz. For the 2018 Pb-Pb run we expect an increase in the collision rate to ~ 30 kHz. Without applying any changes to the L1 trigger mix used in 2015, this extra factor would mean operating CMS above its limit, with a significant impact on the physics program. In addition, it would be beneficial to also further increase the number of minimum bias events collected. The goal of the poposed work is to increase the CMS readout rate to 30 kHz as well as an higher number of events to disk.

It is important to bare in mind that also if the latter two projects results vital for the Minimum Bias event collection and the HF program, these two projects have also a fundamental importance in the ability of CMS to exploit the potential of the luminosity delivered by LHC in case of PbPb collisions.

2.2 Silicon Strip Tracker

In standard p-p collision operation, the tracker detector sends the full detector information to 430 Front-End Driver (FEDs). They apply common mode noise subtraction and strip zero-suppression. However, the common noise subtraction algorithm implemented in the Tracker FEDs firmware doesn't allow compensation of the baseline distortion observed in the presence of Highly Ionizing Particles (HIP), with the consequent potential loss of clusters associated with the affected readout chip. Due to the stringent requirements on tracking for high multiplicity HI events, the cluster-finding inefficiency propagates directly to the track-finding inefficiency, depending on the centrality of the collision.

The solution adopted up to now was to bypass the tracker zero-suppression implemented in the tracker FEDs and send the full detector information to the DAQ/HLT. HI-specific common mode noise subtraction and zero-suppression algorithms were implemented as HLT processes. However, this solution results in a heavy load on the links between the FEDs and the DAQ and on the DAQ itself. The links are based on S-Link technology with a maximum speed of 200 MB/s. Some FEDs (connected to sensors in which higher multiplicity is expected) have duplicate links, allowing a maximum of $\sim$ 400 MB/s per FED. Considering that on the average fragment size is 32 kB/event (VR10 version), the readout limit is at around 12 kHz. It has already been demonstrated that the FED itself is designed to sustain up to 520 MB/s of throughput. The limitation is then only coming from the links to the DAQ.

In agreement with the Tracker group, the only solution available to significantly increase the overall L1 rate is to modify the FEDs firmware reducing the data output to DAQ. The possible solution envisaged were two:

1. Design a new zero suppression algorithm
2. Create an hybrid format in which only the data of APVs with distorted baseline are sent to the DAQ in Virgin Raw mode

In sections 2.2.1 and 2.2.2 a detailed description of the two solutions is presented including the preliminary studies performed. After evaluation, it was decided that solution (2) represents the most interesting in term of performances and feasibility.

For completeness, we should also report that other possible strategies were studied/considered to increase the L1 rate without modifying the tracker FED FW. However, they were rejected as being either too complicated to be implemented or as underperforming. The two other solutions envisaged were either to increase the number of links between the tracker FEDs and the DAQ or to reduce the data payload. The first solution appears to be too complicated, adding an extra $\sim$ 400 links. Apart from the technical constraints, there would also not be enough time to install and commission the links during the available shutdown time. The second solution is to reduce the data payload. A FED firmware version implemented already in 2015 allows the reduction from 10 to 8 in the number of bits read out by each strip ADC. However, this strategy would allow only a moderate increase of $\leq 20\%$ of the overall L1 rate at the cost of a reduction of the detector resolution/sensitivity.

2.2.1 Derivative follower algorithm

In 2013 a specific HI zero-suppression algorithm designed for high multiplicity environments that would satisfy the FPGA constraints was designed offline; it is named a "Derivative Follower". It scans the APVs and it finds gradients. A positive slope can represent the beginning of a cluster while a negative slope can represent the end of a cluster. Checks on the cluster shape are included in the algorithm.

Implementing the algorithm in firmware would give the maximum data reduction to HLT having an average event size of roughly MB/ev. However this approach require a significant rewriting of the tracker FED firmware.

Performance studies on the algorithm using 2015 PbPb data were carried out and a summary of the results can be found in the presentation [1]. As shown in fig. 1 the Derivative Follower algorithm is inefficient in finding high-pt tracks compared to standard HI algoritm used in HLT during the 2015 Pb-Pb collision run. The algorithm optimization can still be performed. However, considering the complexity of implementing the algorithm in firmware and the fact that the whole HI physics program would be based on this algorithm, it was decided to reduce the priority to this option but rather invest on the hybrid data format solution described in section 2.2.2

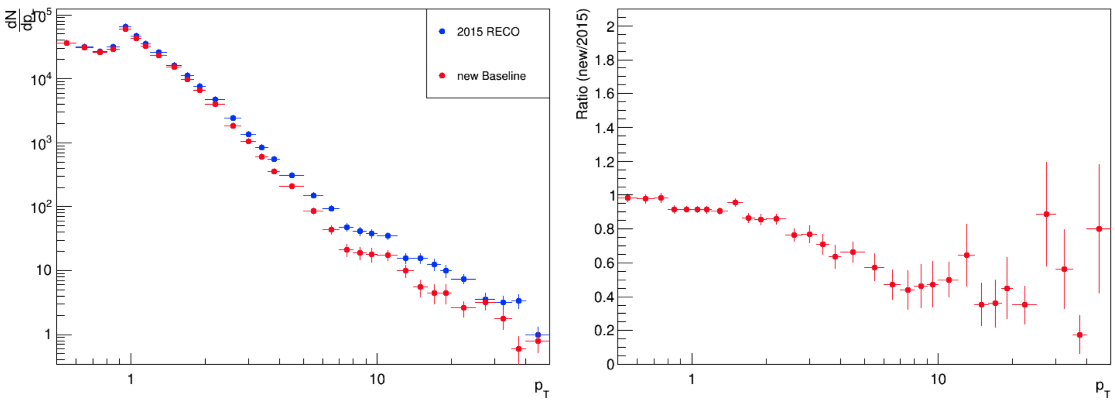

Figure 1:

2.2.2 Hybrid data format

The Hybrid data format strategy is based on the principle that the standard FED zero suppression fails only for the fraction of APVs in which the baseline is either heavily shifted or distorted. When none of the two conditions are met, the Common Mode Noise calculated as a constant value to be subtracted to all the ADC of a single APV, is good enough for cluster findings. In fig. 2 it is shown the fraction of modules for which the full baseline needs to be reconstructed as function of centrality. For very central Pb-Pb collisions $\sim$ 10 k APVs require a specific handling while for Minimum Bias events only $\sim$ 100 APVs require a specific handling.

In this optic, it is possible to receive a significant data reduction to DAQ sending an hybrid data format containing the Virgin Raw data only for the APVs that require a special care. The APVs without pathology in the baseline are zero suppressed in the tracker FED using the standard zero suppression algorithm. In HLT, the tracker data will be unpacked and the standard HI Zero Suppression algorithm (Baseline Follower) will be used to zero suppress the problematic APVs.

The code needed to emulate the full chain is still being written so it is not yet available a precise measure of the actual data size in hybrid data format. However, a rough estimation can be performed considering that each APV sent to the DAQ in VR format adds a weight of $\sim$ 1280 bits. The average HI zero suppressed event event size for the tracker is of $\sim$ 500 kB. This means that DAQ system should be able to handle an avent size that can vary from $\sim$ 500 kB in case of peripheral Pb-Pb collisions to an event size that can be up to $\sim$ 12 MB in case of very central Pb-Pb collisions. Using a conservative approach, we consider that the average tracker event size to DAQ will be of roughly $\sim$ 1 MB. The goal of 30 kHz should then be achieved.

The hybrid data format strategy does not allow the same level of data format reduction as for the Derivative Follower algorithm strategy. However, it has the advantage that allows data reduction reusing the already well

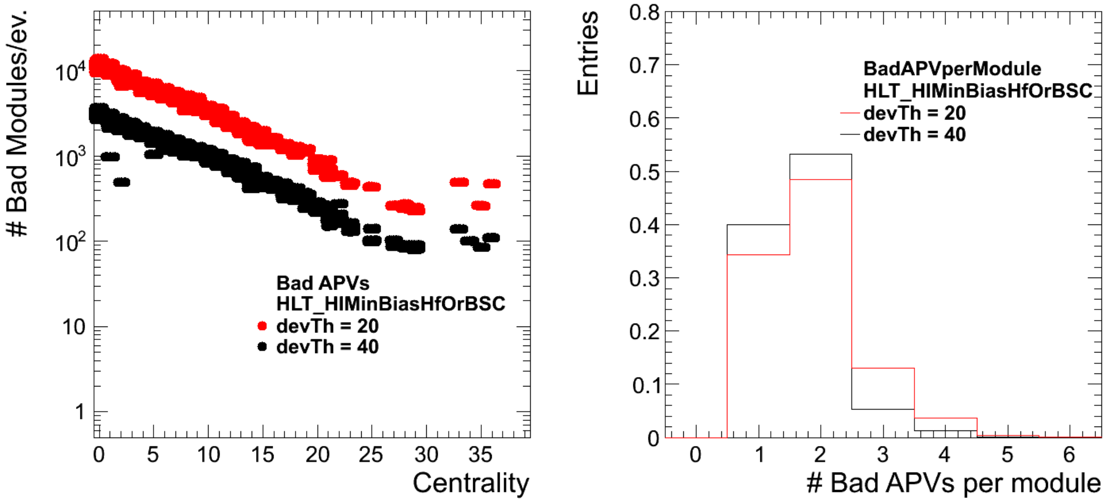

Figure 2:

known HI Derivative Follower algorithm. The performances are then preserved over the several HI runs.

At the time of writing is also being discussed which strategy to be used to mark the APVs as problematic. Two approaches has been evaluated. The first uses the CMN value shift. It is assumed than in case of patological baselines, the CMN shift is higher than for for not patological ones. A method using a threshold can then be envisaged. A second approach consists in counting the number of strips with ADC above a certain threshold. It gives an indication of the APV signal occupancy.It is indeed demonstrated that higher is the multiplicity and higher it is the probability of having deformed baselines.

Once the emulation code will be completed, the final strategy will be defined with also the corresponding fine tuning. An over efficient algorithm to define problematic APVs would not be an issue as long as a data reduction of at least a factor 3 respect to the VR data size is obtained.

2.3 Silicon Pixel Detector

A new silicon pixel detector with 4 layers was installed and commissioned for the 2017 p-p run. The detector front end electronics and sensors had some improvements but for the purpose of this document can be considered equivalente to the legacy detector ones. The whole off-detector electronics was re-designed and the FEDs are based on the latest Xilinx FPGA technology (Virtex-7). The new firmware designs already take into account high multiplicity events. However, the performance of the FEDs needed to be evaluated in the HI multiplicity environment.

An electronics board to emulate the detector response was designed to commission the off-detector electronics. The test board is able to receive input files with hit positions and produces at its output link a response equivalent to one of the actual FE electronics. Random and constant patterns, were injected in one FED and it was measured the throughput with a multiplicity up to 14 hits/ROC/channel. It corresponds to a multiplicity expected in pp collision with PU = 500. Figure 3 shows the results of the test. A rate of 44 kHz at L1 can be archived in this condition. Very central PbPb events have a multiplicity equivalent of pp collisions with

PU=300. More details on the test performed can be found in the presentation [2].

It is important to bare in mind that in case of Pb-Pb collisions, not only the average multiplicity should be considered but the peak multiplicity associated to the single event. The buffer present in the FPGA should be adjusted accordingly. During the Xe-Xe run collected the 12th of October, the buffer size was adjusted passing from 520 to 982 hits/event/channel. Also if central Xe-Xe events have a multiplicity that is 30 % lower respect to Pb-Pb central events, the Xe-Xe run was a good test to verify the functionality of the new pixel FEDs on high multiplicity environment.

In order to complete the studies and fully validate the FED operations also for Pb-Pb collisions, it is planned to inject Pb-Pb simulated data into the system. The test could underline bottlenecks not seen up to now.

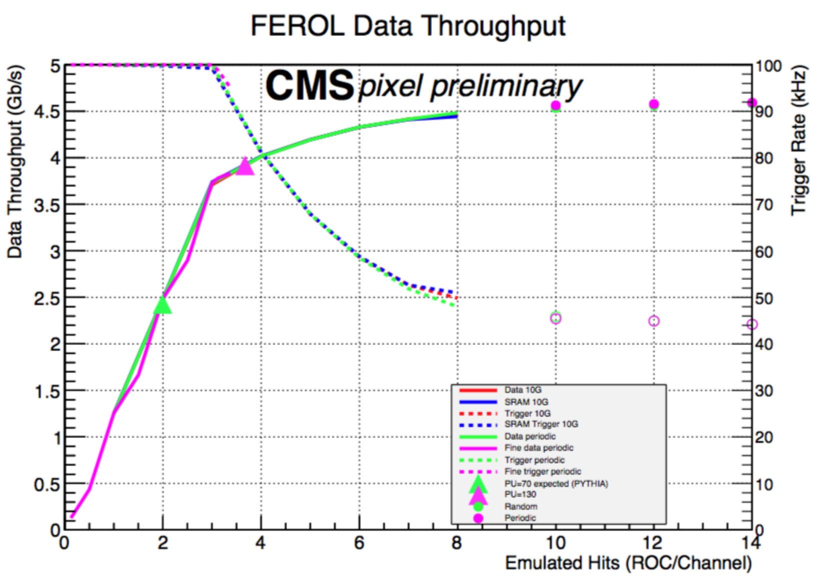

Figure 3:

2.4 L1 trigger

For the 2015 Pb-Pb run an upgraded L1 trigger system was designed. The Stage-1 L1 trigger was operated with specific HI algorithms. However, in the middle of 2016 the full trigger upgrade was installed and commissioned. The upgraded trigger system has a new layer-1 and the single MP7 board layer-2 was replaced with 9 MP7 boards operating with a time-multiplexed architecture. Also, the firmware was completely redesigned by the L1 team, including a series of new algorithms specific to p-p collision operation.

For the 2018 HI run, the specific HI algorithms must be ported to the new system, including a specific background subtraction, centrality triggers, single-track triggers and Q2 triggers. It was agreed with the L1 team that the Stage-2 engineers would take care of the firmware implementation and commissioning. However, we anticipate the CMS HI group will conduct the performance studies of the algorithms and a specific description of the implementation. Furthermore, several tests on the bench or at P5 would be performed by the CMS HI group. Also, the compilation of the L1 and HLT menus would remain responsibility of the CMS HI group.

3 Management Plan, Personnel, Schedule

The overall organization of the upgrade effort of the tracker and ECAL for the 2018 HI run will be conducted within the existing management structure of CMS. HI personnel supported by this request will be part of the two corresponding detector groups and follow the established management procedures within CMS. The project will include the following deliverables:

- Detailed performance studies of the alternative algorithm for the tracker zero suppression. It includes the commissioning of the offline software dedicated to the local reconstruction code;
- Participate in the re-design of the tracker FED firmware including;
- Participation in the phases of validation, test and commissioning of the HI tracker FED firmware in the CMS experiment;
- Performance studies on ECAL data reduction, adjusting ECAL operational conditions. This activity implies also a significant re-design of the offline ECAL local reconstruction

The schedule of the whole project is driven by the need to complete the project by the beginning of Fall 2018. Then a few months are left for online operation and final commissioning phases before the 2018 HI run. We have identified the following milestones in the proposed schedule:

- Silicon Tracker zero suppression algorithm performance studies to be finished by June 2017;
- Silicon Tracker FW implementation to be completed by April 2018;
- Commissioning of the Silicon Tracker FED and corresponding software by September 2018;
- ECAL preliminary performance studies to be concluded by June 2017;
- ECAL calibration and corresponding offline software by August 2018;
- Commissioning of the system in CERN/P5, October 2018;

References

[1] URL https://twiki.cern.ch/twiki/pub/CMS/MITHIGAnalysis2017/Baty_DerivativeFollowerPerformance.pdf

[2] Kilpatrick M Pixel ferol throughput URL https://indico.cern.ch/event/670423/contributions/2744625/attachments/1536139/2406497/FEDTester_10_04_17.pdf