The next step in real time data processing for large scale physics experiments

Sudarshan Paramesvaran for the CMS Collaboration

Abstract

Run 2 of the LHC represents one of the most challenging scientific environments for real time data analysis and processing. The steady increase in instantaneous luminosity will result in the CMS detector producing around 150 TB/s of data, only a small fraction of which is useful for interesting Physics studies. During 2015 the CMS collaboration will be completing a total upgrade of its Level 1 Trigger to deal with these conditions. In this talk a description of the major components of this complex system will be described. This will include a discussion of custom-designed electronic processing boards, built to the uTCA specification with AMC cards based on Xilinx 7 FPGAs and a network of high-speed optical links. In addition, novel algorithms will be described which deliver excellent performance in FPGAs and are combined with highly stable software frameworks to ensure a minimal risk of downtime. This upgrade is planned to take data from 2016. However a system of parallel running has been developed that will allow CMS to install, commission and operate it alongside the current Trigger to assess and validate performance with LHC collision data. This systems combination of state-of-the-art electronics, firmware and software could have many interesting applications for particle physics, astronomy and other areas.

Presented at ACAT2016 17th International workshop on Advanced Computing and Analysis Techniques in physics research
The next step in real time data processing for large scale Physics experiments

Sudarshan Paramesvaran, on behalf of the CMS Collaboration
University of Bristol, Bristol BS8 1TH, UK
E-mail: sudarshan.paramesvaran@bristol.ac.uk

Abstract. Run 2 of the LHC represents one of the most challenging scientific environments for real time data analysis and processing. The steady increase in instantaneous luminosity will result in the CMS detector producing around 150 TB/s of data, only a small fraction of which is useful for interesting Physics studies.

During 2015 the CMS collaboration will be completing a total upgrade of its Level 1 Trigger to deal with these conditions. In this talk a description of the major components of this complex system will be described. This will include a discussion of custom-designed electronic processing boards, built to the uTCA specification with AMC cards based on Xilinx 7 FPGAs and a network of high-speed optical links. In addition, novel algorithms will be described which deliver excellent performance in FPGAs and are combined with highly stable software frameworks to ensure a minimal risk of downtime.

This upgrade is planned to take data from 2016. However a system of parallel running has been developed that will allow CMS to install, commission and operate it alongside the current Trigger to assess and validate performance with LHC collision data.

This system’s combination of state-of-the-art electronics, firmware and software could have many interesting applications for particle physics, astronomy and other areas.

1. Introduction
The Large Hadron Collider at CERN has begun its second run after a highly successful first run which included the discovery of the Higgs Boson. In Run 2 the center-of-mass energy of the colliding protons increases from 8 TeV to 13 TeV, providing a new record-breaking energy scale in which to search for New Physics. Together with this increase in energy, the instantaneous luminosity and simultaneous interactions per crossing will also increase beyond original design specifications. Figure 1 shows an event with 78 collisions in one crossing.

These factors present significant challenges for the Level 1 Trigger systems of the LHC experiments, as they try to disentangle the “events of interest” from the plethora of collisions.

In this paper the focus is on the CMS Experiment [1], and how specifically the Level 1 Calorimeter Trigger was upgraded [2] during the 2 year Long Shutdown 1 to prepare for these new conditions.
2. Triggering at the LHC
The peak crossing rate of protons at the LHC is 40MHz. The ability to read out this information from the detectors is not possible due to bandwidth limitations. In fact, the production cross-section of interesting Physics signatures is so small that most of the collisions produce non-interesting events. This is clearly seen in Figure 2; Higgs processes occur at a much lower rate than inelastic scattering for example.

To deal with this, experiments use “Triggers” to filter out the events of interest and discard the rest. This is a crucial system for all of the LHC experiments as discoveries could be missed.
if you are not triggering on the “correct” Physics signatures. Therefore, a great deal of thought is spent on designing these systems, much of which centres on a few essential functions:

- **Real Time Processing:** The trigger system has to decide in a very short space of time, in CMS this is 3.2\(\mu\)s, whether to retain or discard an event. It has to be able to take a “quick look” and then make a decision.

- **High Rejection Factor:** With current technology we can typically store \(\approx 100\)s of Hz of data - need to be able to discard \(\approx 10e5\) of events

- **High efficiency for interesting events:** The trigger needs to have algorithms capable of identifying interesting Physics signatures with high efficiency to ensure discoveries remain possible.

- **Flexibility:** The Physics goals and priorities can evolve with time and with new information, thus it is essential that changes to the Triggers can be made relatively easily.

An important observation in the context of the above points is that the LHC is a discovery machine - and therefore it is not always clear what the new Physics signatures will look like. It is therefore important to have the ability to have quite generic Triggers to be able to capture a wide range of possible events.

### 2.1. How Triggers work

The basic concept of a Trigger system works by transmitting information from detectors to electronics boards, where a first decision is made. This reduces the rate of accepted events to around 100KHz in the case of CMS. The information from this first level trigger is then sent for further processing to a computing farm, or High-Level Trigger - where the rate of accepted events is typically on the order of 100s of Hz. In CMS - the trigger paths are split by object. For example the information from the Calorimeters is processed a specific Calorimeter Trigger and likewise for the Muon detectors. This information is combined in a “Global Trigger” which makes a final decision on whether to keep or discard the event.

### 2.2. Run 1 CMS Calorimeter Trigger

Figure 3 shows the original CMS Calorimeter Trigger [3] which was used throughout Run 1 of the LHC. The technology available at the time drove the design of the systems - this is why ASICs and early FPGA’s were used for the data processing, along with parallel copper links to move data from place to place.

This trigger was extremely successful, able to withstand a 100KHz rate at Level 1 for the duration of the run. The Physics accomplishments of this run, Standard Model precision measurements, searches for Supersymmetric signatures, and of course the discovery of the Higgs boson [4] , all owed to the excellent performance of this Trigger.

However, as discussed earlier in this paper, Run 2 presents significant new challenges. Without upgrading the Trigger, the Physics events The maximum rate of acceptance at the first level of 100 KHz is a fixed maximum during Run 2 so within this boundary, the improvements have to made.

### 3. The road to upgrading...

The upgrade to the trigger necessitates first looking at the available technology and what could be coming in the future months and years, and evaluating how these technological improvements could yield an improved Trigger.

An important design consideration is the form factor of the electronics boards and crates. At the time of designing the upgrade, the \(\mu\)TCA standard [5] was very popular - especially in the telecommunication industry. It has several advantages: open modular standard, small form
The data used as input to the L1 trigger system as well as the input data to the global muon trigger, global calorimeter trigger and the global trigger are transmitted to the DAQ for storage along with the event readout data. In addition, all trigger objects found, whether they were responsible for the L1 trigger or not, are also sent. The decision whether to trigger on a specific crossing or to reject that crossing is transmitted via the Trigger Timing and Control system to all of the detector subsystem front end and readout systems.

### 1.4.2 Calorimeter Trigger

The calorimeter trigger begins with trigger tower energy sums formed by the ECAL, HCAL and HF upper level readout Trigger Primitive Generator (TPG) circuits from the individual calorimeter cell energies. For the ECAL, these energies are accompanied by a bit indicating the transverse extent of the electromagnetic energy deposit. For the HCAL, the energies are accompanied by a bit indicating the presence of minimum ionizing energy. The TPG information is transmitted over high speed copper links to the Regional Calorimeter Trigger (RCT), which finds candidate electrons, photons, taus, and jets. The RCT separately finds both isolated and non-isolated electron/photon candidates. The RCT transmits the candidates along with sums of transverse energy to the Global Calorimeter Trigger (GCT). The GCT sorts the candidate electrons, photons, taus, and jets and forwards the top 4 of each type to the global trigger. The GCT also calculates the total transverse energy and total missing energy vector. It transmits this information to the DAQ for storage.

![Figure 1.2: Overview of Level 1 Trigger](image)

**Figure 3.** A schematic of the dataflow for the original CMS Trigger, used during Run 1 of the LHC.

Due to the fact that a wide variety of AMC’s are supported, the Trigger could exploit this by designing and building a few cards that could cover the needs of the entire system. The decision to use μTCA was made earlier on - enabling the design of the cards to be started on early by the different groups involved in the upgrade.

Alongside deciding which standard to use for infrastructure, an analysis of the current processors had to be undertaken. This clearly showed that the XiLinx Series 7 FPGA’s were the cutting edge for processing, and a major advance in what was used in the legacy trigger for Run 1. These new FPGAs had up to 2 Million logic cells, with a total serial bandwidth of up to 2.8TB/s. For the calorimeter trigger, two cards were designed using these new FPGA’s, and are shown in Figures 4 and 5. Both make use of the high-end Virtex 7 FPGA and have high-speed optical links.

The available space in these new processors led to a change in thinking about how to process the data from the detector. Previous triggers adopted a conventional “regional” architecture - meaning data from the detectors is split and sent to different processors depending on which physical location the data is coming from. This necessarily leads to the problems associated with overlap regions and boundary conditions between physical locations. With these new large FPGAs- it seemed possible to put all algorithms into one board. In order to for this to be useful for processing, the architecture had to be rethought.

3.1. **Time-Multiplexed Triggering (TMT)**

The advantages of using a single processor to process the entire CMS Calorimeter are evident:
There would be no overlap/boundary conditions to consider
Each board with algorithms would be identical
Each board would have access to full event information
Could rotate through a series of boards sequentially to increase efficiency.

It’s clear that to achieve this goal, the latest FPGA’s combined with high-speed optical links would be needed. The data coming from the detector would need to be organised in such a way that full calorimeter information for one bunch crossing would be sent to a the processor with algorithms in turn. The following bunch crossing would then be sent to the next processor, and this would continue in a round-robin fashion.

Figure 6 shows the design of the CMS TMT. Layer 1 of the TMT receives the calorimeter information, applies preliminary input calibration, before arranging the data and sending the data to Layer 2 where the algorithms find the physics objects. One of the most interesting parts of this upgrade was devising a method to send the data between Layer 1 and Layer 2. The data from Layer 1, is transmitted on 12-fibre ribbons, with each fibre providing information from the same physical location in the calorimeter, but on a different bunch crossing. For example, if the fibres on the ribbon are numbered 1-12. The first fibre will bring information on bunch crossing 1, and the second fibre information on the second bunch crossing etc. However each of these fibres has to go a different processing node, fibre 1 will have to go processor 1, fibre 2 to processor 2 and so on so that each processing node has complete information on a single bunch crossing. Doing this multiplexing using single fibres would have taken an entire rack of space with more than 1100 fibres. A novel solution was proposed which only took up the space of 3 pizza boxes. Using Molex Flexplane [7] technology, all the routing is done on mesh-fibre plane inside the patch panel and the outside only has inputs and outputs for the 12-fibre ribbons. This is show in Figure 7.

3.2. Parallel Running
In order to minimise the effect of upgrading the trigger on Physics studies, the decision was taken to upgrade the Trigger while maintaining the Run 1 Trigger. This translated into being able to send the calorimeter information to both the legacy and upgraded trigger simultaneously. This proved extremely useful as direct comparisons could be made between the Triggers, and enabled us evaluate the performance with collision data. For this reason the Calorimeter Trigger
**Figure 6.** A schematic of how a time-multiplexed trigger operates, showing the Layer 1 and Layer 2 processors.

**Figure 7.** The Molex Flexplane [7] custom-designed multiplexer optical patch panel
Hardware was installed and commissioned during 2014 and 2015. During the 2015 LHC run over 3 billion collision events were collected with the upgraded Trigger. These events were analysed prior to the start-up and provided useful input and confidence to start the 2016 run with the upgraded Trigger.

4. Conclusion
The LHC performance expected for Run 2 required an upgraded CMS Trigger in order to maintain the high performance achieved during Run 1. Using the latest technology, a novel time-multiplexed architecture was researched and developed for the Calorimeter Trigger. This was installed and commissioned during 2015, including new µTCA infrastructure, state-of-the-art electronics boards and optical fibres. A system of parallel running was developed which enabled this Trigger to be operated in parallel with the Run 1 Trigger in 2015. This gave us the confidence required to start the 2016 run with this new Trigger.

The new technology advances drove us to think of new and improved ways of real time data processing. This also showed the importance of researching emerging technology well in advance and to continually rethink the ways we can use the technology. The next major step in real time data processing will be trying to include the data from the Trackers in the Trigger systems, and many people have been working on this problem for some time already.

References