Trigger R&D for CMS at SLHC

G. Iles [a], C. Foudas [a], M. Hansen [b] , J. Jones [c]

[a] Imperial College, London, UK
[b] CERN, 1211 Geneva 23, Switzerland
[c] Princeton University, Princeton, NJ, USA

g.iles@imperial.ac.uk

Abstract

CERN has made public a comprehensive plan for upgrading the LHC proton-proton accelerator to provide increased luminosity commonly referred to as Super LHC (SLHC) [1]. The plan envisages two phases of upgrades during which the LHC luminosity increases gradually to reach between 6-7×10^{34} $cm^{-2}sec^{-1}$. Over the past year, CMS has responded with a series of workshops and studies which have defined the roadmap for upgrading the experiment to cope with the SLHC environment. Increased luminosity will result in increased backgrounds and challenges for CMS and a major part of the CMS upgrade plan is a new Level-1 Trigger (L1T) system which will be able to cope with the high background environment at the SLHC.

Two major CMS milestones will define the evolution of the CMS trigger upgrades: The change of the Hadronic Calorimeter electronics during phase-I and the introduction of the track trigger during phase-II.

This paper outlines alternative designs for a new trigger system and the consequences for cost, latency, complexity and flexibility. In particular, it looks at how the trigger geometry of CMS could be mapped onto the latest generation of hardware while remaining backwards compatible with current infrastructure.

A separate paper presented at this conference [2] looks at what could be possible if large parts of the trigger system were changed, or additional hardware added to create a time multiplexed trigger system.

I. INTRODUCTION

Plans are already well advanced for upgrades to the LHC machine that will provide increased luminosity. The current CMS experiment will fail to reap the full benefit of these upgrades for a number of reasons. One of these is that the current trigger system will be overwhelmed. It will not be possible to set sensible energy thresholds without the trigger rate exceeding the maximum Level-1 Accept (L1A) rate of 100kHz. Hence the Global Trigger would be forced to restrict the trigger rate by simply pre-scaling the trigger and thus effectively negating any benefit from increased luminosity. It is for this reason that work has started on trying to integrate a tracking trigger in a future trigger system.

This would help identify the most interesting events and bring the trigger rate back below 100kHz. A new trigger system could potentially have several others benefits such as improved flexibility because it would be based solely on FPGAs. The improvements in technology could also make the system easier to design, build and maintain, which could have a substantial impact not just on the cost of the hardware, but also on the manpower cost to test and operate it.

The phase I upgrade of the Hadronic Calorimeter (HCAL) electronics will precede that of the tracker and will provide lateral information of the energy depositions within the HCAL. An upgraded trigger system implemented at the same time would provide improvements to cluster-based triggers, such as the tau trigger, whilst at the same time preparing the trigger for track trigger information. This will enable CMS to make more stringent isolation cuts and provide triggers of higher purity early in the upgrade program. Consequently, the time seems ripe to begin consideration of a new trigger system.

II. CURRENT TRIGGER

The trigger in CMS is split into two stages; the L1T (Level-1 Trigger) operates on coarsely segmented data that is transmitted and analysed for every proton-proton bunch crossing; the HLT (High Level Trigger) operates on the high resolution data that is stored on-detector in pipeline memories and is only read out after receipt of a L1A. The L1T uses a mixture of ASICs and FPGAs to processes data from each bunch crossing (i.e. 40MHz), while the latter uses PCs to process events at up to 100kHz.

The L1T design is split into two paths. The calorimeter trigger path is decribed here, but there exists a similar path for the muon trigger.

The Trigger Primitive Generators (TPGs) provide coarsely segmented data from the detector front ends at “tower” resolution, which for the Electromagnetic & Hadronic Calorimeters (ECAL & HCAL) consist of energy depositions with some additional detail (e.g. energy spread). The RCT (Regional Calorimeter Trigger) uses a clustering algorithm to search for electron candidates. It also reduces the resolution further by building “regions”. These are then used by a clustering algorithm in the Global Calorimeter Trigger (GCT) to find jets. The GCT then sorts the electrons and jets into rank (i.e. in order of importance) and transmits the data to the Global Trigger (GT) which searches for physics signatures.

III. UPGRADE PATH

A new trigger system would replace the RCT and GCT. It would be highly desirable if this could be achieved with little impact on the rest of the CMS detector. The minimal changes would probably require upgrading the TPG and GT interfaces to use multimode optical links running at speeds comparable to the latest iteration of FPGAs (i.e up to 6.5Gb/s, perhaps up to 11 Gb/s).

This was foreseen over a year ago and thus when a replacement had to be designed for the GCT-GT links it was based on a Xilinx Virtex 5 with multimode optics [3]. The Optical Global Trigger Interface (OGTI) design (fig. 1) is essentially the first step in an upgrade of the trigger. A beneficial aspect of the card is that there is spare link bandwidth and thus it would be possible to drive two GTs. An upgraded GT could therefore be developed in parallel with the existing GT without having an impact on normal CMS running.

Figure 1: OGTI Card. Xilinx XC5VLX110T FPGA and 4x POP4 optics providing 16 channels at 3.2Gb/s in a dual CMC form factor.

It might be useful to use the same concept for the TPGs, which would need their links upgraded (i.e. they would have dual outputs). This is relatively easy because the links between the TPGs and the RCT reside on a daughter card known as the SLB. Hence the second step in an upgrade program would probably be to switch these links to use optical multimode links and an FPGA.

A new RCT and GCT could then be developed in parallel with the output going to a new GT, which could then be fed into the existing GT as a technical trigger without comprimising normal CMS operation.

Upgrading the links in CMS is relatively straight forward, but not the data on them. The latter would require changing the TPGs and while this is planned for the HCAL there is currently no plan for ECAL. A second option might be to build adapter cards, however this would impose a latency penalty that may or may not be acceptable. The following is therefore a consideration for a new trigger system design in which the data flowing from the TPGs remains unchanged, albeit concentrated onto faster optical links where possible.

IV. TRIGGER GEOMETRY

The CMS coordinate system (fig. 2) has its origin centred at the nominal collision point. The azimuthal angle φ (0 to 2π radians) is measured in the plane perpendicular to the beam. The polar angle θ ($-\pi/2$ to $\pi/2$) is measured from the plane perpendicular to beam, although it is more normally expressed in terms of pseudorapidity, η, because at a hadron collider particle production is roughly constant as a function pseudorapidity.

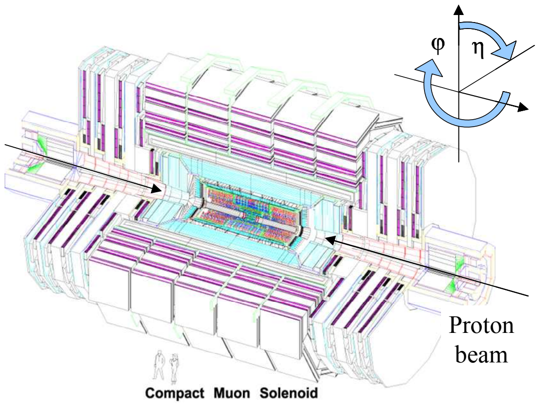

Figure 2: The φ and η coordinate system used in the CMS detector.

The TPGs, provide coarsely segmented data at “tower” resolution, which has an η, φ coverage of 0.087 x 0.087 rad up to $\eta = 1.74$. Beyond that the towers are larger [4]

The trigger geometry (fig. 3) is split into 18 regions in φ and ±11 regions in η, however regions ±8 and above (i.e. psueudorapidity > 3.0 and < 5.0) are only covered by the Forward HCAL.

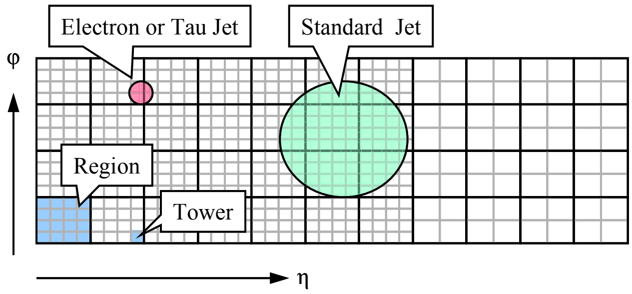

Figure 3: A portion of the RCT input geometry. Only 4 of the 18 regions in φ are shown and only ½ of η. The approximate size of an electron, tau and normal jet are shown to give the reader an indication of size.

Each region is sub-divided into 4x4 towers except for the HF that is divided into 2x2 towers. In the case of ECAL, these towers are further subdivided into 5x5 crystals. Electrons have a width of less than 2 towers in both dimensions. Tau jets are similar, although they can extend to 3 towers in the φ dimension. Standard jets span up to 9-12 towers in both dimensions. Both systems transmit 8bits of energy and one extra bit. ECAL transmits the Fine Grain Veto bit, which is asserted when 90% of the energy within a tower is not contained within two crystals in η (i.e. it is designed to identify a single electron/photon, while allowing for the fact that an electron might emit bremsstrahlung radiation in the magnetic field). HCAL transmits the Minimum Ionising Particle (MIP) bit, which indicates that the energy deposited was compatible with a muon passing through it.

The tower information arrives at the RCT in the form of cables with 4 channels (ABCD). Channels AB and CD both span a single tower in η, but 4 towers in φ and when combined they span 2 towers in η. The links currently run at 1.2Gb/s with each bunch crossing comprising 2x9bits of tower data, 5bits of hamming code and a single bit for BC0 identification.

The 4 links would combine nicely to create a single 4.8Gb/s link with room for additional information if the Hamming code and BC0 were discarded in favour of a once per orbit CRC check and a special 8B/10B k-code to indicate BC0. This would provide 8 towers per bunch crossing. However, there are some special circumstances in which channels ABCD do not originate from the same location and thus forming a single 4.8Gb/s link would not be possible. Instead there would have to be 2x 2.4Gb/s links which would require additional FPGA I/O.

V. TECHNOLOGY CHOICE

The two major advances over the last 5 years that are particularly useful for a trigger system are the continuing advances in both FPGA technology with embedded SerDes blocks operating a multi Gb/s rates, and the move to the optical interconnects necessary to transmit these signals over distances of more than a few feet.

Despite the latest FPGAs now having an I/O bandwidth of several hundred Gb/s they are still approximately an order of magnitude below what would be needed to absorb all the TPG data of several Tb/s in a single FPGA.

The challenge is therefore to concentrate the data into multiple FPGAs with sufficient boundary condition data for the cluster algorithms to operate efficiently and within a timescale of < 1µs.

If we assume that in an upgrade there should be some spare capacity for additional tower information (e.g. improved energy resolution) and thus allocate 12bits rather than 9bits per tower and we also assume a 4.8Gb/s, 8B/10B link synchronised to the LHC clock then we can transmit 8 towers (i.e. half a region) per bunch crossing (25ns). It is of course possible to slightly improve the efficiency of the link by going to 64B/66B encoding. We may also prefer to run with a slightly faster asynchronous clock, at perhaps 5.0Gb/s, however these are just details. The basic architecture should not be determined by these details and the data packing on the fibres should not be optimised so that it becomes imposible to easily understand the system. Consequently, we require approximately 4 links per region to accept HCAL and ECAL data. It is assumed that any tracking trigger, possibly even muon trigger would require substantially less bandwidth because it is only transmitting location information, however for modularity reasons they may require multiple input links and perhaps a lower speed interface to the FPGA (i.e. < 1Gb/s).

VI. INITIAL CONCEPT

The original concept behind a new trigger system was to place all the ECAL, HCAL, muon and tracking trigger information into a single FPGA at tower resolution so that coincidences between different subsystems could be used to improve physics object recognition. The baseline design consisted of finding trigger objects centred within a single region that was bounded by a region on all sides and all corners so that an array of 3x3 regions was constructed (fig. 4). The boundary information would be provided by duplicating data where necessary. This led to the development of the Matrix card [2] that incorporated a 72x72 cross-point switch for data duplication.

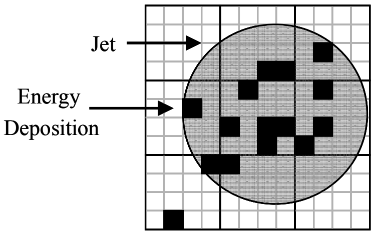

Figure 4: The 3x3 regions required to encompass a jet centred on a tower somewhere in the central region. If a single tower is considered the centre of the jet then the algorithm could sum energy depositions from up to 9 towers in each dimension.

This architecture has several disadvantages. The design is very inefficient because only 1/9 of the data is processed in any given processing card. Furthermore, duplicating and distributing such a large quantity of data is not trivial. For example, if we use our earlier assumption of 4 links per region to bring ECAL and HCAL data into the FPGA we would require 36 (9x4) links running at 4.8Gb/s. The largest Xilinx Virtex 6 FPGAs do have this many links, however there is little spare capacity for extra trigger input.

Furthermore, it is currently envisaged that the data duplication would take place with a combination of large, high speed serial, protocol agnostic, cross-point switches and optical / µTCA backplane interconnects. It is not clear whether the links would be able to pass through many of these components, as they might have to, without regeneration to avoid the jitter becoming too large. The inefficient nature of the design would require a large number of cards (> 252). Lastly, the large number of cards would require the sorting stage to consist of two stages (i.e. passing through 2 cards) because of the large fan-in. This would impose additional latency.

VII. SPLIT FINE/COARSE PROCESSING

An alternative approach was therefore considered. It is the requirement to fully contain a jet that requires such a large overlap between processing regions. It was therefore decided to split the fine and coarse processing into two parts. The fine processing would have the bandwidth to provide an overlap of just one tower in the first dimension and have an entire region of overlap in the second dimension. The fine processing would concentrate on electron and tau detection whereas the coarse processing would be used for jet detection.

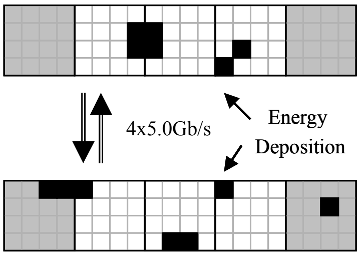

Figure 5: Two processing cards exchanging data to perform fine processing (i.e. creating electron/tau clusters). The two shaded regions on either side provide data to build clusters centered within the 3 middle regions.

The basic concept (fig. 5) is to receive 5 regions of data in η, although potentially it could be φ, and locate electrons and taus centred on the 3 central regions (or 3+1 regions when one region is at a η limit). Hence 4 cards could span from η = -3.0 to +3.0 (i.e. where there is both ECAL & HCAL coverage). The 4 cards would cover η regions -7 to -4, -3 to -1, +1 to +3, and +4 to +7. If we assume that we need 4 links at 4.8Gb/s to receive 12bits of data for both HCAL and ECAL information then we would expect to require 20 input links excluding any tracking information. However, the barrel/endcap boundary is arranged in such as way that it is probably not possible to merge the 4x1.2Gb/s links into a single 4.8Gb/s link (i.e. the data sources are in different locations) and it would be necessary to use 2x2.4Gb/s links. Hence we expect that the cards covering η = -3 to -6 and η = +3 to +6 would require 22 links, however this would need verification from ECAL and HCAL cabling experts.

In the second dimension, which would nominally be φ, 4 bidirectional links would provide either the overlap information or possibly pre-clustered objects. The latter potentially offers far more useful information to be transferred, possibly even allowing full size jets to be built, however this requires study because it would require a more complex algorithm. A very similar concept is used in the current GCT to sucessfully cluster jets. The 4 bidirectional links would be transmitted over either a custom µTCA backplane or QSFP optical cables.

There are 18 regions in φ and thus a full system would require 72 cards distributed across 8 µTCA crates, with a pair of crates for each η segment.

The simplest way of handling the jets is to coarse grain the data into 2x2 tower squares and transmit them to a jet processing stage. The 2x2 tower resolution is more than sufficient for jet processing and would combine very nicely with the jet information from the HF which is already at a 2x2 tower resolution. The jet cards would cluster jets centred on an area that spanned ½ of η and 2 regions in φ, but they would have access to 1 extra region in both η and φ so that jet clusters could be built with a size up to 10x10 towers. The electrons and jets would then be sorted in terms of rank (i.e. importance) before being forwarded to the GT. It would require 4 cards to sort the electrons and 2 cards to sort the jets. The GT would receive up to 16 electrons and 16 taus (4 per η segment), 8 central jets from the HCAL Barrel & Endcap, and 8 forward jets from the Forward HCAL).

The design currently uses 22x 5 input links and 8 sharing links running at 5.0Gb/s. There would also need to be a link for slow control over Ethernet and another for DAQ. Hence 32 links are used. It is assumed that the bandwidth for a tracking trigger would be substantially less as it is simply indicating the presence of a high transverse momentum track. A single input link would be sufficient to provide 1bit of information per tower.

A minimum of 36 links are therefore necessary if we wish to reserve up to 4 links for a tracking and possibly even muon information.

The Xilinx XC5VTX150T has 40x 5.0Gb/s links and the latest announcements from Xilinx for the Virtex 6 range include up to 36x 6.5Gb/s links (XC6VLX550T) for the LXT series and 48x 6.5Gb/s links, plus 24x 11Gb/s links for the HXT series (XC6VHX565T) .

VIII. PROCESSING CARDS

The Mini-T5 (fig. 6) is an attempt to build a processing card with the capabilities necessary to realise the system described above. The same card would be used for the fine (electron/tau) processing, coarse (jet) processing and subsequent sorts.

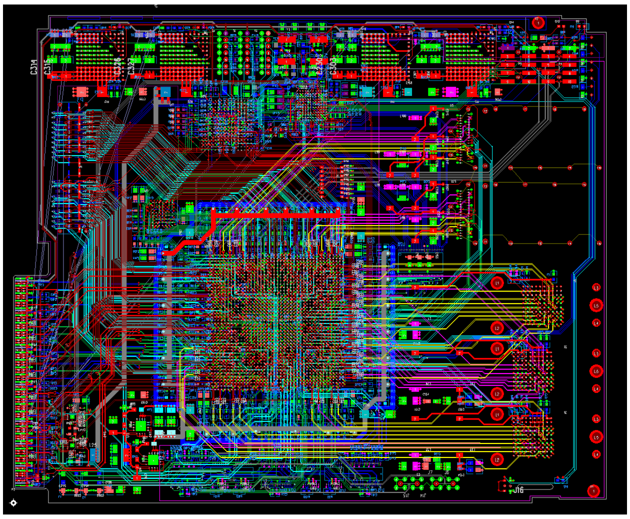

Figure 6: The Mini-T5 technology demonstrator card. SNAP12 optics would be mounted bottom right. QSFP optics are mounted in the middle of the right hand side. Power supplies are at the top. The Samtec differential headers and the AMC card edge connector are on the left hand side.

It is based on a Xilinx Virtex-5 XC5VTX150T-2FFG1759C in a double width AMC form factor. The FPGA offers 40 links running at up to 5Gb/s. It is pin compatible with the XC5VTX240T if extra logic or links are required. It also uses the same GTX transceivers used in the Virtex-6 and thus it should be possible to upgrade the board with minimal changes to the firmware when the large Virtex-6 FPGAs become available.

There are two types of optics. SNAP12s are uni-directional devices providing either 12 inputs or outputs at up to 6.5Gb/s. An interesting alternative is the PPOD from

Avagotech, which is very similar, but rated up to 10Gb/s, however questions remain over availability to relatively low volume science experiments. QSFPs offer 4 bidirectional links at up to 10Gb/s, but often in only a cable format (i.e. no MTP connector). This doesn't allow the fan in/out of fibres often required by a physics experiment. The Mini-T5 has 2xSNAP12-Rx, 1xSNAP12-Tx and 2xQSFPs.

Additional high speed link I/O is provided on the backplane on ports 0-7 (i.e. common options and fat pipes on the µTCA specification). Ports 1 and 3 have the option of being switched to LVDS ports on the FPGA to allow for reception/transmission of fast control such as Timing, Trigger & Control (TTC) and Trigger Throttle System (TTS).

The card also has Samtec QTH/QSH series headers on either side of the card, which are each connected to up to 40 LVDS pairs that can operate up to 1.25Gb/s. Samtec offers flex cables for these connectors and thus it is possible to hook adjacent cards together with very low latency and with a bandwidth similar to that of the QSFP optical inter card connection. Alternatively, it is possible to install daughter cards for additional tracking trigger I/O.

The card also has an external AT32UC3A microprocessor for offloading appropriate tasks and for AMC card functionality. The design is finished and is passing through pre-manufacture checks before being submitted for manufacture.

IX. LATENCY

The latency associated with serial links is unpleasant (typically ~100ns for both transmission and reception), however it offers an excellent way of bringing large amounts of data into an FPGA and offers electrical isolation between sub-systems. The CMS TDR allocates < 1µs for both RCT and GCT including input and output links. Hence if we wish to retain a reasonable amount of time for processing within FPGAs we must have a maximum of 2 serial link transmissions within a combined RCT and GCT.

In the Mini-T5 example the first serial link period is used to provide the overlap area for the electrons and pass the coarse 2x2 tower information to the jet processing cards. The second serial link period is used for transmitting the data to sorting cards.

X. SERVICES

The MCH in a µTCA crate (fig. 7) provides GbE and clock distribution to each slot, however CMS would probably require additional functionality. For example the LHC clock needs to be extracted from the biphase mark encoded TTC signal, which is distributed at 1310nm on single mode fibre. The fast control information (i.e. Channels A/B) encoded on the TTC signal needs to be distributed in a constant latency, upgradeable manner (i.e. LVDS at 400 or 800Mb/s). Some systems (e.g. trigger) have a very high data bandwidth, but generate a relatively small amount of data. For these systems it would be useful to have a data concentrator or DAQ channel per card.

Figure 7: The Vadatech VT891 crate with 12 full size AMC slots and redundant MCH/PM slots may be a good choice for a standard CMS µTCA crate.

Trigger systems also need a lot of inter card data sharing. This can be accomplished by modifying an existing µTCA backplane. This is standard practice in the µTCA community and relatively inexpensive.

XI. CONCLUSIONS

A compact trigger architecture has been presented that remains backwards compatible with the current CMS experiment. It could be easily extended to incorporate a tracking trigger. A single card design is used for the entire system, albeit loaded with 4 different firmware versions, of which 2 are very simliar.

XII. ACKNOWLEDGEMENTS

We would like to thank Sarah Greenwood (Imperial College) for layout of the Mini-T5 card and STFC for financial support.

XIII. REFERENCES

[1] F. Zimmermann et al., "CERN Upgrade Plans for the LHC and its Injectors", CERN-sLHC-PROJECT-Report-0016, 2009

[2] J. Jones et al., "The GCT Matrix Card and its Applications", These Proceedings, Paris, France, 2009

[3] G. Iles et al., "Performance and lessons of the CMS Global Calorimeter Trigger", TWEPP-08, Naxos, Greece, 2008

[4] The CMS Collaboration, S Chatrchyan et al., "The CMS experiment at the CERN LHC", JINST 3 S08004, 2008

[5] The Trigger and Data Acquisition Project, Vol. I, The Level-1 Trigger, CERN/LHCC 2000-038, CMS TDR 6.1, 15 December 2000.