Revised CMS Global Calorimeter Trigger Functionality & Algorithms

G. Iles[a], J. Brooke[c], C. Foudas[b], R. Frazier[c], G. Hall[b], M. Hansen[a], M. Hansen[c], G. Heath[c], J. Jones[b], J. Nash[a], A. Rose[b], M. Stettler[a], A. Tapper[b]

[a] CERN, Switzerland
[b] Imperial College London, UK
[c] Bristol University, UK

gregory.iles@cern.ch

Abstract

The Global Calorimeter Trigger (GCT) is a device which uses data from the CMS calorimeters to search for jets, produce isolated and non-isolated electron lists and compute all the transverse and missing transverse energy sums used for the Level-1 trigger decision (L1A). GCT performs these functions by receiving and processing the data from the Regional Calorimeter Trigger (RCT) and transmitting a summary to the Global Trigger (GT) which computes the L1A decision. The GCT must also transmit a copy of the RCT and GCT data to the CMS DAQ. The vast amount of data received by the GCT (230 Gb/s) as well as the necessity for data sharing required by the jet finder impose severe constrains on the GCT design. This paper presents an overview of the revised design, in particular, the algorithms, data flow and associated latency within the revised GCT.

I. Introduction

The GCT is the last stage of the calorimeter trigger chain. Its primary purpose is to reduce the number of calorimeter trigger objects that need to be processed by the GT to produce the L1A decision [1]. The pipeline memories that store the event information prior to a L1A decision have only a limited depth and thus the trigger latency is limited to 3.2μs.

A detailed description of the GCT is beyond the scope of this paper and is covered in detail in the CMS Trigger TDR [1] and several subsequent CMS internal notes [2,3].

The trigger objects computed by the GCT from data supplied by the RCT are listed below and described in subsequent paragraphs.

- 4 isolated and 4 non-isolated electrons of highest rank
- 4 central, 4 forward , 4 tau clustered jets of highest rank
- total transverse energy
- total missing transverse energy
- total jet transverse energy
- 12 jet counters

The 'rank' of an electron or jet is at present its transverse energy; however, in principle it could also be derived from jet location.

The electron sort operation must determine the 4 highest rank objects from 72 candidates, supplied by the RCT, for both isolated and non-isolated electrons from a significant data volume (29Gb/s per electron type).

To sort the jet clusters according to rank the GCT must first perform jet cluster finding and convert the clustered jet energies to rank. The clustered jets are created from regional transverse energies supplied by the RCT. The latter are the sum of contributions from both the hadronic and electromagnetic calorimeters. This is a substantial extension of the GCT capability beyond that specified in the Trigger TDR [1].

The jet cluster finding and subsequent sort is more challenging because of the larger data volume (172.8 Gb/s) and the need to share or duplicate data between processing regions to perform cluster finding. The latter can require data flows of a similar magnitude to the incoming data volume depending on the clustering method used. The clusters, defined as the sum of 3x3 cells, are located using a new method, which is described in Section III. It requires substantially less data sharing than the sliding window method.

Jets are subdivided into central, forward and tau jets based on the RCT tau veto bits and the jet pseudorapidity, η.

The GCT must also calculate some additional quantities. The total transverse energy is the sum of all regional transverse energies. The total missing transverse energy is calculated by splitting the regional transverse energy values into their x and y components and summing the components. The resulting vector, after a rotation of 180°, provides the magnitude and angle of the missing energy.

The total jet transverse energy is the sum of all found clustered jets. There are 12 jet counters that are calculated from rank and position criteria.

In addition to these tasks the GCT: (a) acts as a readout device for both itself and the RCT by storing information until receipt of a L1A and then sending the information to the DAQ via a SLINK64 interface (b) extracts trigger information for the muon system from the calorimeter data stream (c) monitors the LHC luminosity.

The revised design is discussed, although the hardware details are kept to a minimum as they are presented in a separate talk at this conference [4]. The data processing firmware, in particular the algorithms and associated latency within the revised GCT, are presented.

II. System Overview

The GCT is a challenge to build because (a) the data volumes and processing require many large FPGAs and thus multiple PCBs (b) the PCBs must be connected together to allow data sharing and to eventually concentrate the data into a single location for the sort algorithms (c) the latency allowed is 24 bunch crossings for jets and 15 bunch crossings for electrons. Ideally one would use high speed serial links to transport the large data volumes between FPGAs; however, these typically require several clock cycles to serialise/deserialise the data and thus they have to be used sparingly to keep the latency low. The final architecture (fig. 1) uses high speed optical links (1.6 Gb/sec actual and 1.28 Gb/sec effective) to concentrate the data in the main processing FPGAs, followed by standard FPGA I/O to connect to downstream FPGAs. There are two main trigger data paths; electron and jet.

A. Dataflow

A block diagram of the GCT system is shown in figure 1. The input to the GCT is 18 RCT crates. The 72 Source cards retransmit the data on optical high speed serial links (shown by dashed arrows). For each RCT crate, the electron data is transmitted on 3 fibres, the jet data on 10 fibres and the muon data on 1 or 2 fibres (not shown).

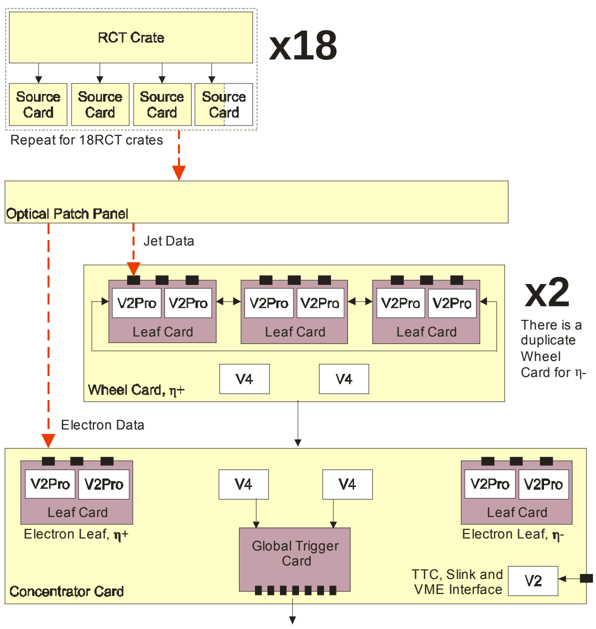

Figure 1: Block diagram of GCT system. Dashed arrows indicate high speed optical links. The 23 large FPGA devices in this design are indicated as squares with the either a symbol 'V2P' (Virtex-II Pro), 'V4' (Virtex-4) or 'V2' (Virtex-II) inside.

The jet data are sent to Leaf cards (configured for Jet finding) mounted on the Wheel cards. The Leaf cards are connected in a circle to search for clustered jets in one half of the CMS calorimeter (either in the positive or the negative η). The Wheel card collects the results from 3 Leaf cards, sorts the clustered jets and forwards the data to the Concentrator.

The electron data are sent to Leaf cards (configured for electron sorting) mounted on the Concentrator.

The Concentrator performs a final sort of electrons and clustered jets and send the results to the Global Trigger.

In addition to this the jet data is also used to calculate the transverse energy sums and jet counts; however, these tasks do not impose the significant constraints on the design, other than some extra FPGA I/O.

B. Source Card

The 6 differential ECL cables per RCT crate are fed into Source cards (fig. 2), which will be installed in the same racks as the RCT. The Source cards have been designed by Imperial College. They receive up to two RCT cables and transmit the data over 4 fibre links using 4 HFBR-5720 single fibre transmitters operating at 850nm. This has several advantages. It allows the Source cards to be electrically isolated from the main GCT system. It allows the different data within the RCT cables to be rearranged (e.g. electron data can be split from muon data so they can be sent to different destinations). It allows a large amount of information to be concentrated so that it can be delivered to the processing FPGAs on Leaf cards. It allows data to be duplicated.

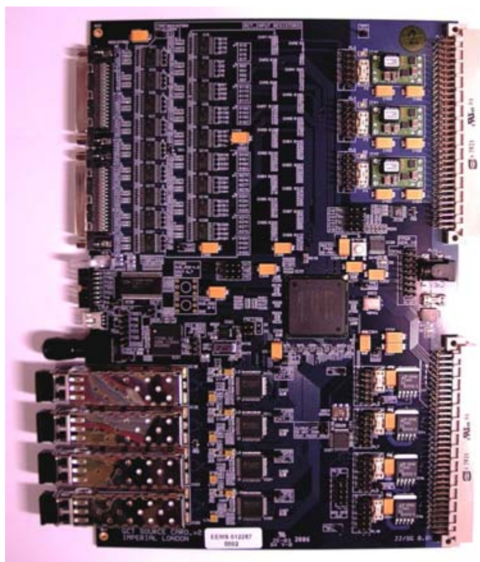

Figure 2: The Source card. The RCT inputs are visible in the top right of the picture. The 4 SFP cages are in the bottom right. All the signals are routed through a Xilinx Spartan 3 (XCS1000-5FG676C).

C. Leaf Card

A Leaf card (fig. 3) is the main processing block in the GCT design. It is a dual CMC card designed at CERN and produced by LANL. The most difficult task in the GCT is the jet clustering. This is made simpler by concentrating the data in as few FPGAs as possible. Consequently, the Leaf card has 2 Virtex II Pro FPGAs (XC2VP701513C), each with 16 multi-gigabit transceivers that are used to bring the raw data in. Three Agilent AFBR 742B 12 channel receivers provide the opto-electronic interface. The large standard I/O are used to transmit the data to the Wheel card. Each Leaf card in jet finding mode transmits to the Wheel card the 18 clustered jets, total transverse energy and missing transverse energy.

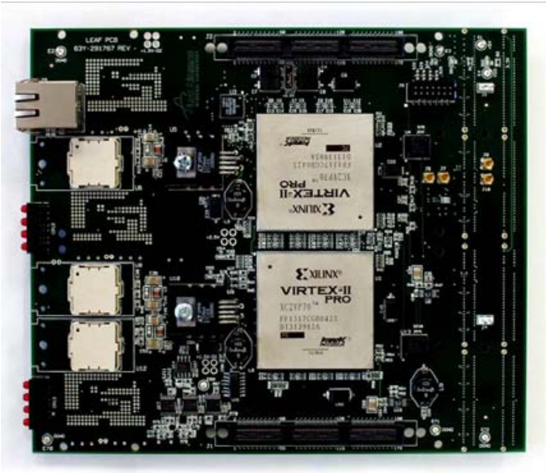

Figure 3: The Leaf card. The LVDS Samtec connectors that provide Leaf-to-Leaf card connection are visible at the top and bottom.

D. *Wheel Card*

There are 2 Wheel cards, one for each half of the detector. They are 9U, 400mm, VME cards (albeit without a VME interface) that act as carriers for 3 Leaf cards and further concentrate the data. They sum the energy values and jet counts (number of clustered jets found in η-φ regions). They also sort the 54 clustered jets by rank into the 3 types (forward, central, tau). The Wheel cards then forward the information to the Concentrator card via high speed Samtec LVDS cables.

E. *Concentrator Card*

The 9U, 400mm, VME Concentrator card performs similar actions to that of the Wheel Card after which it transmits the resulting trigger objects to the Global Trigger and stores the information in a pipeline until receipt of a L1A trigger. The Concentrator also carries 2 Leaf cards, which process the electron data. These Leaf cards record the incoming RCT data in a pipeline memory until receipt of a L1A and perform a fast sort on the incoming data.

The interface to the GT is another dual CMC card that contains National DS92LV16 SERDES chips, of which 14 are used to transmit data to the GT and 2 are used for loopback testing. Infiniband High Speed Serial Data 2 connectors and cables are used to receive/transmit the data in pairs at an effective rate of 2x1.28 Gb/s per pair.

III. Processing Methods & Algorithms

A. *Clustering*

Traditionally, the clustering method used in the CMS trigger design was based on a 3x3 sliding window. The jet clusters, defined as the sum of 3x3 cells, were found by a sliding window that traversed the η-φ space of 22x18 regional transverse energy values. In this method a cluster was created if the central value of the 3x3 array was larger than the 8 neighbouring cells.

This method was simple and symmetric; however, the processing could not easily be split into regions without sharing or duplicating the boundaries of the regions.

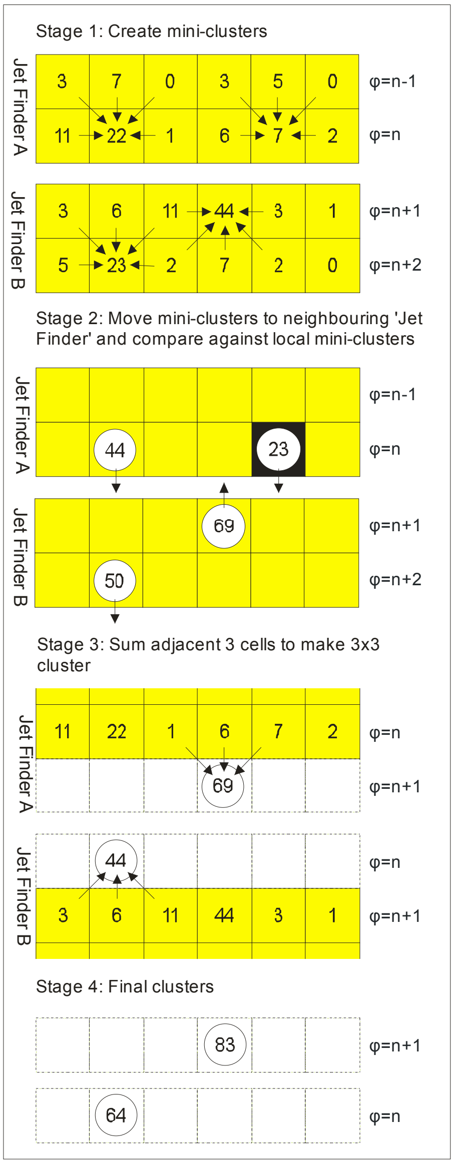

Figure 4: The Jet cluster method. There are 6 cells in η shown, but there are 11 cells per 1/2 detector.

The current method, called the 'Jet Finder', receives data from a single RCT crate that spans 2 cells in φ (40 degrees) and 11 cells in η (half the detector). It also receives duplicated data from the next 2 adjacent cells in η (i.e. at η = 0) that are located in the other half of the detector (i.e. these 4 cells span the same φ region, but have η with the opposite sign). Duplicating the cells at η = 0 allows the processing regions to be independent and thus in physically different locations. The current method only uses the immediately adjacent cells in η and thus the Jet Finder operates on a 2x12 region in φ and η. A 1x12 region of constant φ is known as a φ strip.

The 1st stage of the current method (fig. 4) is to create 2x3 mini-clusters, when the central cell is larger than neighbouring cells. When a central cell is larger than or equal to neighbouring cells the situation is more complicated (e.g. a jet divided equally between two cells). In this scenario a clustered jet should still be created. The equality statements between cells are set-up so that in particular directions the central cell must be greater than its neighbours, whereas in other directions it must only be greater than or equal to. This ensures that clustered jets are never formed adjacent, including diagonally adjacent, to each other.

The 2nd stage of the process is for the Jet Finder to transfer the 3 largest mini-clusters in each of the two φ strips in opposite φ directions. In practice this can be an inter-FPGA, inter-Leaf or Leaf-to-Leaf connection because the 9 Jet Finders for a given η sign are distributed across 3 Leaf cards. On the Leaf card the 3 Jet Finders are split 2/1 across the 2 FPGAs.

The Jet Finder therefore receives 3 mini-clusters from each φ direction from neighbouring Jet Finders. These are compared against the existing mini-clusters on the receive φ strip. Mini-clusters adjacent or diagonally adjacent to a larger mini-cluster are removed. The equalities are again setup in such a way so that should two mini-clusters have the same value only one will be removed.

In the 3rd and 4th stages the received mini-clusters that survive have their 3 adjacent cells in the receive φ strip added to themselves to make a 3x3 cell. A consequence of this method is that the Jet Finder that initially operates on φ strips n and n-1 produces full clusters centred on φ strips n+1 and n-2.

The current jet clustering method takes advantage of the fact that in a given strip of constant φ, viewed by one RCT crate, in either positive or negative η it is sufficient to consider a maximum of 3 clusters rather than 6. It is this possibility that enables the current method to reduce the data sharing requirement from 66% to 25% of the aggregate input data in the revised GCT design. This allows the Leaf-to-Leaf transfers to operate at 40MHz DDR (80MHz) rather than 160MHz DDR (320MHz). The latter is possible given the LVDS Leaf-to-Leaf connections; however, the timing constraints and thus the design is more difficult.

B. Sort

Sorts must be performed on the Leaf, Wheel and Concentrator. The Wheel card has the most challenging sort. It has up to 54 clustered jets, which must be reduced to 4 with the highest rank. In the Concentrator the sorts must produce the 4 highest ranked objects in order of rank.

The generic sort operates on 16 bit words of which the rank (on which the words are sorted) occupies 6 bits. A diagram of the sort is shown in figure 5. The 1st stage of the sort is the creation of a comparison matrix in which every input is compared against every other input. The matrix is constructed so that two different inputs are only compared once and the result inverted for the other comparison (i.e. $M_{ij} = M_{ji}^{-1}$). This has the benefit of not simply reducing the amount of logic, but more importantly that where two input values have the same value then one is given preference (e.g. if two input values were the same and were both in 2nd place then one value would be placed 2nd and the other 3rd). In the resulting matrix an input value which had the highest rank has a row in the matrix consisting entirely of logic '0's. The next highest ranked object has a row consisting entirely of '0's except for one logic '1' and so on.

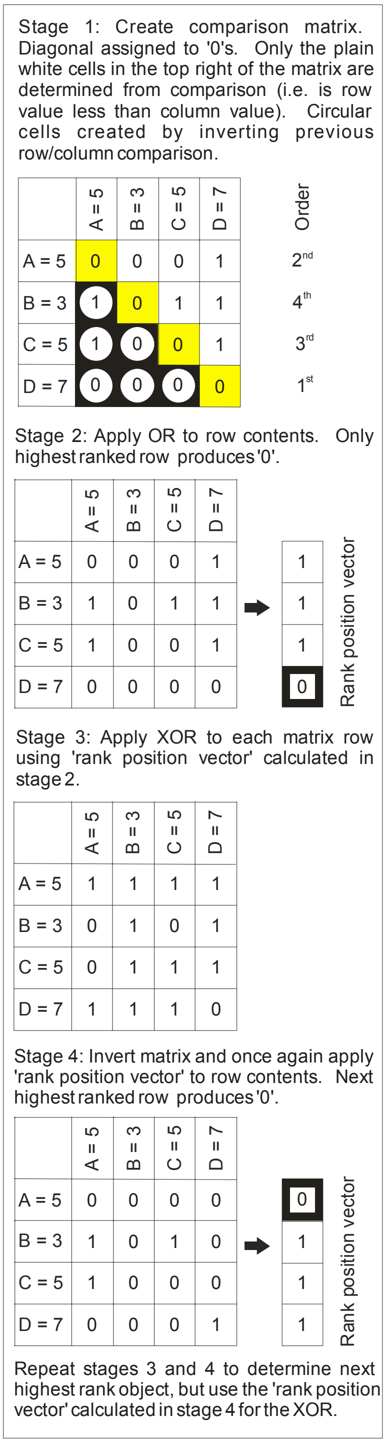

Figure 5: An example of the sort algorithm applied to 4 inputs; however, it is optimised for handling many more inputs.

In the 2nd stage the highest ranked input is found by applying an OR to all values in a matrix row. Only the row with highest rank will produce logic '0'. The result from all rows is stored in a rank position vector.

In the 3rd stage the next highest ranked input is found by applying an XOR between the rank position vector, just calculated, and the rows of the comparison matrix.

In the 4th stage the entire matrix is inverted. Stages 3 and 4 modify the comparison matrix so that the same method can be used to find the next highest ranking object (i.e. An OR is applied to the values in each matrix row. Only the row with highest rank, or now in our case next highest rank, will produce logic '0').

The 3rd and 4th stages are repeated to find next highest rank input, albeit with the latest rank position vector (i.e. that calculated in the 4th stage).

The output stage of the sort (i.e. selection of the highest ranked inputs) is performed by a wide OR rather than multiplexers because this enables the synthesis tool to use the 5 input LUTs rather than the 2 input multiplexers and thus make the combinatorial chain shorter. The maximum so far sorted is 36 to 4 in a Virtex 2 Pro (XC2VP70-7FF1517C) in a single LHC bunch crossing (25ns).

IV. LATENCY

The latency is 15 bunch crossings, excluding cable delays between the RCT/GCT and GCT/GT, and excluding any contribution from the serialising/deserialising the data over GCT/GT links. The jet data arrive 9 bunch crossing cycles earlier and thus there is more time to process this data

	Jet latency in bunch crossings	Electron latency in bunch crossings
Source Card, including RCT cable	2.5	2.5
Source to Leaf Fibre	(4)*	(4)*
Leaf Deserialise	3	4
Leaf Processing A	3	1
Exchange Leaf Data	3	N/A
Leaf Processing B	2	N/A
Move to Wheel	2	N/A
Wheel Processing	3	N/A
Move to Concentrator	2	2
Concentrator Processing	1	1
Move to Serialiser	0.5	0.5
Total latency	23	11
Allocated latency	24	15

Table 1: The contributions to the latency for both the electron and jet data paths. Note that the "Source to Leaf Fibre" delay is not included because the equivalent delay is excluded in the Trigger TDR [1].

Table 1 shows the various contributions to the latency for the two different trigger paths. Achieving a nominal 15 bunch crossing latency for the Jet processing is difficult despite the data arriving significantly earlier. It should be noted that this table is based on synthesis of key components of the system. The full firmware is not yet complete and the system has yet to be commissioned and thus the latency may change slightly.

The DS92LV16 latency is not included in the table yet it will obviously contribute to the total latency of the final trigger system. The latency from registering the data with the transmit clock to presenting the data with the receive clock is 3 bunch crossings with no allowance for cable delays.

V. STATUS

Two Source cards have been manufactured and extensively tested. They have passed a successful integration stage with the RCT. A further 8 cards have been produced and are under test.

The Leaf card has already passed preliminary tests and testing of the Multi-Gigabit serial links will commence shortly, followed by Source-Leaf integration.

The Concentrator is in manufacture and is expected to return in mid-October. The Wheel card manufacture will commence after the design aspects that are common between the two cards have been verified.

The core parts of the firmware have already been simulated and synthesised to verify that the latency target can be met. The electron trigger path is scheduled for installation in the first months of 2007, with the jet data path following soon after.

VI. ACKNOWLEDGEMENTS

We would like to thank the following; Pascal Vulliez (CERN), William Billereau (CERN), Sarah Greenwood (Imperial) and Tom Cote (Los Alamos) for their layout of the designs; PPARC for financial support; Wesley Smith and Sridhara Dasu for their help in understanding the GCT requirements.

VII. REFERENCES

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

[2] CMS-IN 04-009, Updated Interface Specification for the CMS Level-1 Regional Calorimeter Trigger to Global Calorimeter Trigger.

[3] CMS-IN 02-069, CMS Level-1 Global Calorimeter Trigger to Global Trigger and Global Muon Trigger Interfaces.

[4] M. Stettler et al., "The CMS Global Calorimeter Trigger Hardware Design", These proceedings.