Available on CMS information server

CMS CR -2022/197

The Compact Muon Solenoid Experiment

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

17 October 2022 (v3, 21 October 2022)

Architecture and Prototype of the CMS Global Level-1 Trigger for Phase-2

G. Bortolato, C. Deldicque, D. Gigi, B. Huber, E. Leutgeb, A. Lobanov, D. Rabady, A. Racz and H. Sakulin on behalf of the CMS Collaboration

Abstract

We present the architecture and current state of prototype firmware of the CMS Level-1 Global Trigger, the final stage of the new Level-1 trigger for Phase-2 of the operation of the LHC. Based on high-precision inputs from the muon, calorimeter, track and particle flow triggers, the Global Trigger evaluates O(1000) cut-based and neural-net-based algorithms in a system of up to thirteen Xilinx Ultrascale Plus based ATCA processing boards interconnected by 25 Gb/s optical links. In order to optimize the usage of resource, the main algorithms, including the DSP-based calculation of invariant masses, are implemented at 480 MHz.

Presented at *TWEPP2022 Topical Workshop of Electronics for Particle Physics*

Architecture and Prototype of the CMS Global Level-1 Trigger for Phase-2

G. Bortolato,[a,b] C. Deldicque,[a] D. Gigi,[a] B. Huber,[a,c] E. Leutgeb,[a,c] A. Lobanov,[d] D. Rabady,[a] A. Racz[a] and H. Sakulin[a,1] on behalf of the CMS Collaboration

[a] *CERN,*
Esplanade des Particules 1, P.O. Box, 1211 Geneva 23, Switzerland
[b] *Università degli Studi di Padova*
Via VIII Febbraio 2, 35122 Padova PD, Italy
[c] *Technical University of Vienna,*
Karlsplatz 13, 1040 Wien, Austria
[a] *University of Hamburg,*
Mittelweg 177, 20148 Hamburg, Germany
E-mail: Hannes.Sakulin@cern.ch

ABSTRACT: We present the architecture and current state of prototype firmware of the CMS Level-1 Global Trigger, the final stage of the new Level-1 trigger for Phase-2 of the operation of the LHC. Based on high-precision inputs from the muon, calorimeter, track and particle flow triggers, the Global Trigger evaluates O(1000) cut-based and neural-net-based algorithms in a system of up to thirteen Xilinx Ultrascale Plus based ATCA processing boards interconnected by 25 Gb/s optical links. In order to optimize the usage of resources, the main algorithms, including the DSP-based calculation of invariant masses, are implemented at 480 MHz.

KEYWORDS: Trigger Algorithms; Trigger concepts and systems (hardware and software)

[1] Corresponding author.

Contents

1. Introduction 1
2. Target Hardware 1
3. Architecture of the Level-1 Global Trigger 2
4. GT Algorithm Board Firmware Prototype 2
5. GT Final-OR Board Firmware Prototype 4
6. Summary 5

1. Introduction

For Phase-2 of the operation of the LHC, starting in 2029, CMS will undergo major upgrades to its detectors and readout electronics. A completely new first-level trigger system [1] will ensure that the excellent physics performance of CMS is maintained or improved under the challenging pile-up conditions in Phase-2. The new trigger system will exploit high granularity information from the calorimeters, muon systems and a track finder reconstructing tracks from the silicon strip tracker at the bunch crossing (BX) rate. A latency budget of 12.5 μs will be available to the level-1 trigger, compared to 3.8 μs in Phase-1, facilitating the use of sophisticated algorithms that previously were only employed in software at the higher trigger levels, such as vertex finding, particle flow reconstruction and the extensive use of neural networks. As illustrated in Figure 1, the final stage in the level-1 trigger, the Global Trigger (GT), will receive collections of high-precision trigger objects from the Global Muon Trigger, Global Calorimeter Trigger, Global Track Trigger and Particle Flow Trigger. The GT will determine the level-1 trigger accept decision by evaluating a menu of cut-based and neural-net-based trigger algorithms that can be tuned to target specific physics signatures. Cut-based algorithms may include conditions on event topology, such as the angle between particles or the invariant mass of a hypothetical mother particle of two trigger objects. To enable the search for certain types of long-lived particles, the GT will be able to correlate trigger objects within a window of ±3 BX. The GT will send the final trigger decision to the Phase-2 Trigger and Timing Control and Distribution System (TCDS2) [2].

2. Target Hardware

The GT is planned to be implemented using multiple Serenity [3] generic ATCA processing boards developed by the CMS collaboration, each equipped with a single Xilinx Virtex Ultrascale+ VU13P FPGA. The board will provide up to 120 inputs and 80 outputs at 25 Gb/s via Samtec Firefly optics. For prototyping, also a version of the Serenity board with a VU9P FPGA is used. A firmware framework delivered with the board handles the 25 Gb/s links using a common protocol and provides interfaces to TCDS2 and for control and monitoring. The GT prototype firmware is developed as a payload firmware within this framework. Two prototype

Serenity boards with a VU9P and one with a VU13P, with some optical connectivity between the boards are available in a shared test setup, the CMS Integration Facility.

3. Architecture of the Level-1 Global Trigger

Given the availability of silicon tracker information and particle flow reconstruction, the GT will receive a significant number of trigger object collections – 26 are currently defined – from the upstream systems. These include muon, electron, γ, jet, τ and energy sum collections, most of them in standalone, track-matched and particle-flow varieties, as well as several meson collections (B_S, ρ, φ). The increase of object collections with respect to Phase-1[1] is expected to lead to a trigger menu of up to O(1000) algorithms. Novel algorithms such as neural networks will need additional FPGA resources. The GT architecture therefore needs to comprise multiple FPGAs for algorithms, ideally in an easily scalable way. While most upstream systems are implemented in a time-multiplexed [4] way with consecutive BXs processed by distinct boards, the GT needs to be able to access objects in preceding and subsequent BXs for long-lived particle triggers. It is thus natural to implement the GT in a conventional way.

Figure 2 shows the architecture of the GT. Up to 12 algorithm boards each receive a copy of all upstream system inputs, so that any trigger algorithm may be placed into any of the FPGAs. The algorithm outputs are collected by the Final-OR board, which provides the interface to TCDS2. The total latency budget planned for the GT is 1 µs. Figure 2 shows how the latency will be distributed over the two layers of the GT.

4. GT Algorithm Board Firmware Prototype

Prototype firmware for the algorithm board has been developed targeting both the VU9P and VU13P FPGAs. The firmware handles 66 out of 78 currently defined time-multiplexed 25 Gb/s input links. On these links, object collections, typically containing 12 objects, are received one after the other over the time-multiplexing period of 6 or 18 bunch crossings. These objects are 64,

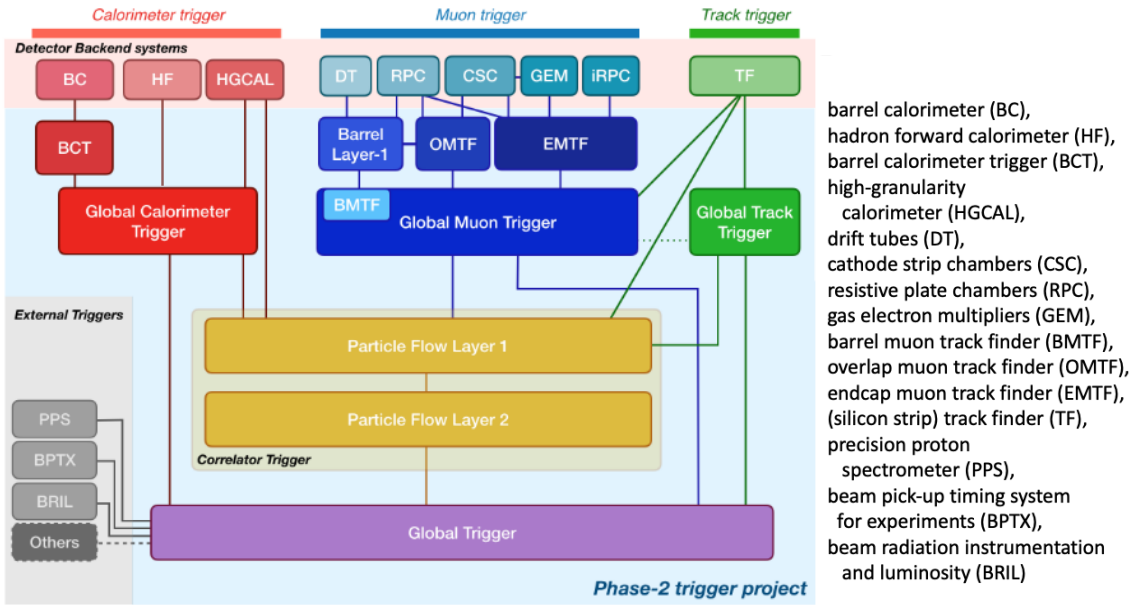

Figure 1. Overview of the CMS Level-1 Trigger for the Phase-2 upgrade.

[1] The Phase-1 GT [5] receives 6 object collections and currently applies a menu of ~400 algorithms.

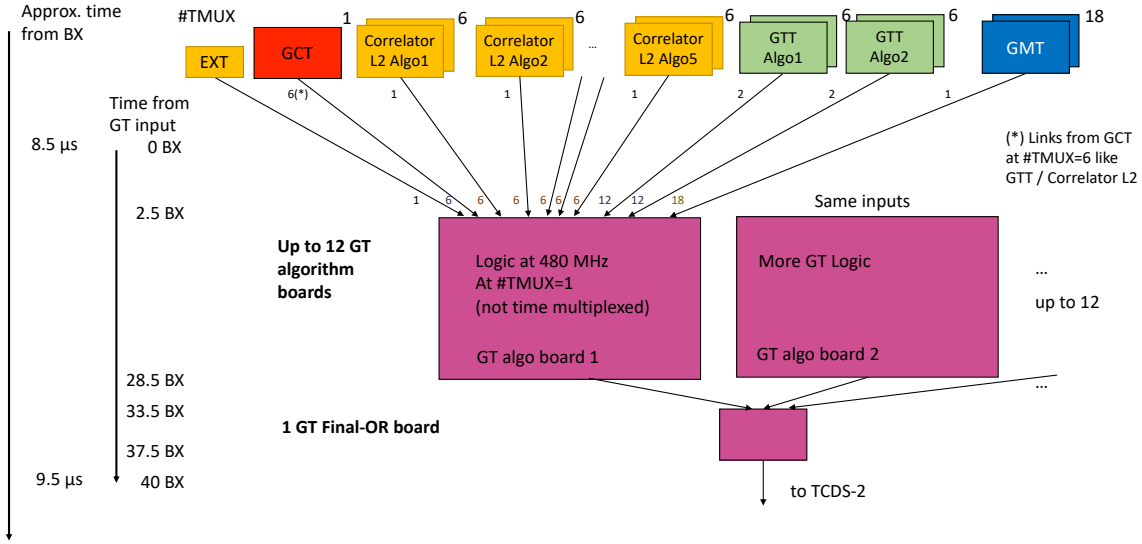

Figure 2. Architecture of the CMS Level-1 Global Trigger for Phase-2. #TMUX denotes the time multiplexing factor, i. e. the number of processing boards over which the event stream is distributed.

96 or 128 bits wide. An effort was made to unify their data formats as much as possible so that algorithms in the GT can be developed independently of the object type they are applied to with as few as possible scale conversions in the GT. At the interface to the payload firmware, data from an input link are received as 64-bit words at 360 MHz. In the *demultiplexing* step, data from each link are sequentially written into dual-port memories – one per collection with a width corresponding to the object width. These memories are then read out in parallel at 480 MHz so that the 12 objects of all collections are streamed to the algorithm units within one BX. The read-out then continues with the memories of the next time slice (next link) so that the algorithms receive a continuous stream of objects at 480 MHz. Where necessary, scale conversions are applied immediately after reading the memories so that all streams distributed to the algorithm units follow a uniform internal format.

The target FPGA contains four individual dies, so-called Super Logic Regions (SLR) that are connected via an interposer. Signals crossing from one SLR to another incur a delay that needs to be considered in the design. At higher clock frequencies, timing closure is facilitated by using transmit and receive registers in the so-called laguna sites at the boundary of the SLR (in a similar way to using input and output flip-flops in the I/O blocks of an FPGA). In order to successfully place and route the prototype firmware, all inputs from an upstream system are assigned to the same SLR and the demultiplexer logic is constrained to that SLR as illustrated in Figure 3 (left). Demultiplexed collections are distributed to all other SLRs passing through the registers in the laguna sites and several additional flip-flop stages to cross an SLR. Having all collections available in all SLRs allows algorithms to be placed freely within the FPGA.

Implementing algorithms at 480 MHz results in significantly reduced resource usage compared to the 40 MHz implementation used in Phase-1. At 480 MHz, a single comparator is used to compare all objects of a collection to a threshold, while 12 comparators are needed in a 40 MHz implementation. When evaluating correlational di-object conditions, each object in the first object collection needs to be compared to each object in the second collection calculating for example the distance $\Delta R^2 = \Delta\phi^2 + \Delta\eta^2$ between the objects using two DSPs for the calculation. While at 40 MHz, all these calculations are done in parallel using 12x12x2=288 DSPs, the 480 MHz implementation deserializes the first collection and then compares all objects of the first

collection with one object of the second collection at each 480 MHz clock, needing only 12x2=24 DSPs, albeit at the cost of additional latency.

A prototype menu consisting of 78 single-object and di-object conditions per SLR, including 24 di-object conditions in each SLR with ΔR or invariant mass correlational conditions, was successfully placed in the target FPGA as illustrated in Figure 3 (right). With a latency of 410 ns the prototype meets design requirements. In this prototype, LUT usage is at 30%, DSP usage at 23% and Block RAM usage at 70%, the latter in large part due to buffers allocated by the firmware framework. The demultiplexing logic accounts for 4% of LUT usage. The prototype was tested against a C++ simulation in the CMS Integration facility by injecting simulated events into playback memories and capturing the algorithm results in spy memories. 100% agreement with the C++ simulation was found. At this density of algorithms, and assuming that the prototype menu is representative of a future physics menu, 3.5 FPGAs would be needed to implement 1000 cut-based algorithms. The scalable architecture of up to 12 boards allows for sufficient contingency and the implementation of advanced resource-intensive algorithms such as neural-net-based triggers.

5. GT Final-OR Board Firmware Prototype

The output of the GT is a vector of multiple bits per BX, each denoting whether a certain physics trigger sub-type fired. Physics trigger sub-types are distributed to the front-ends by TCDS2 and may be used to adapt the readout behaviour. The exact list of sub-types remains to be defined. It may include sub-types for minimum bias, high priority or long-lived particle triggers. In the GT, each algorithm bit is assigned a vector of trigger sub-types that are combined with a logical OR across all algorithms. Before the assignment of the trigger sub-types, each algorithm bit is optionally pre-scaled so that its rate can be adapted to the instantaneous luminosity

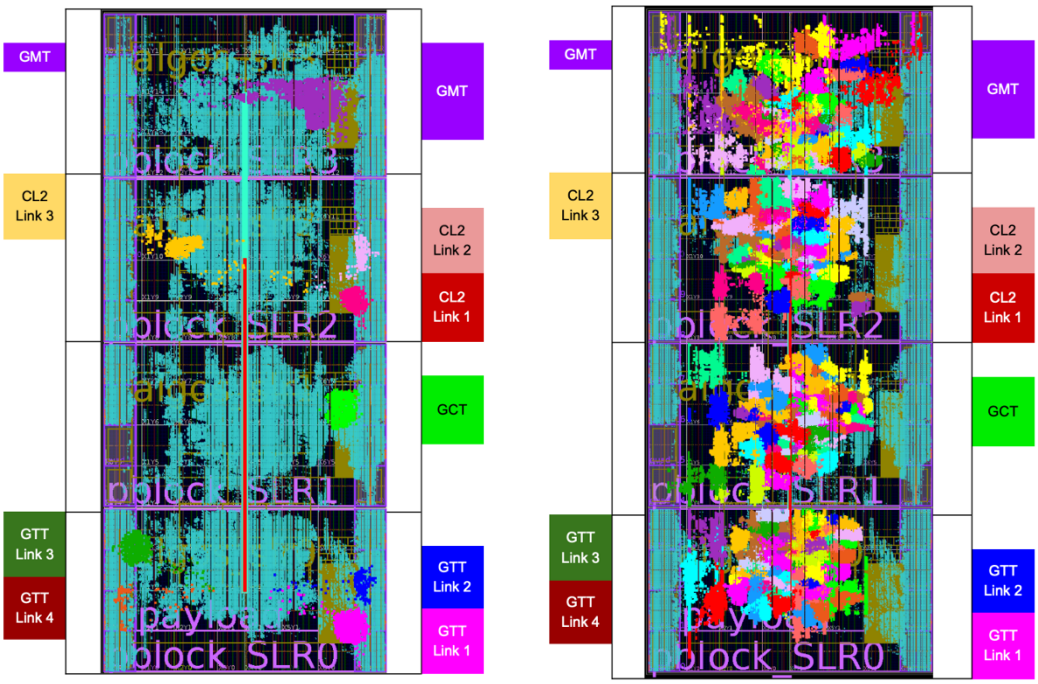

Figure 3. Floorplan of the prototype Global Trigger algorithm firmware on a Xilinx Virtex Ultrascale+ VU13P FPGA with 4 Super Logic Regions (SLR). Left: Demultiplexing logic is constrained to the same SLR as the inputs (in the same color). Right: 78 single or di-object algorithms are placed per SLR.

of the LHC. To be able to make fine-grained rate adjustments, a fractional pre-scale logic, that can for example apply a pre-scale factor of 1.5 in order to accept 2 out of 3 triggers, is used. The rate for each algorithm needs to be monitored before and after pre-scaling, after pre-scaling with a preview factor and after considering dead-time. Most of the described logic may be implemented either in the algorithm FPGAs or in the Final-OR FPGA.

A prototype for the Final-OR board firmware has been developed for a Serenity with a VU13P FPGA, performing all the above processing for 1152 algorithms (the capacity of two 25 Gb/s links with the standard protocol) in the Final-OR FPGA. In order not to limit the number of algorithms that can be placed into a single GT algorithm board, each GT algorithm board outputs the full set of 1152 algorithm bits via 2 fibres, carrying half of the algorithm bits to SLR 2 and the other half to SLR 3 of the Final-OR FPGA (Figure 4, left). A second set of 2 fibres is used to avoid latency due to SLR crossings in the algorithm FPGA. SLRs 2 and 3 in the Final-OR FPGA receive up to 24 links with algorithm bits each. These are deserialized to 40 MHz and then combined across links with a logical OR. The algorithm bits are then pre-scaled and monitored before assigning the trigger sub-types as described above. Only the trigger sub-types fired by the first half of the algorithms need to pass from SLR 3 to SLR 2 where they are combined with the trigger sub-types fired by the second half of the algorithms and sent to TCDS2. Figure 4 (right) shows the floor plan of the Final-OR board. The LUT usage in the two top SLRs is just below 50%, while the flip-flop usage is at 40%. The firmware has successfully been tested in the integration facility with inputs from a GT algorithm FPGA via optical fibre.

6. Summary

The architecture of the CMS Global Level-1 Trigger for Phase-2 consists of a layer of up to 12 GT Algorithm boards and a GT Final-OR board. Both layers target a Serenity generic processing board with a single VU13P FPGA. Firmware prototypes for both boards, covering basic functionality, have been developed and successfully tested in prototypes of the Serenity board.

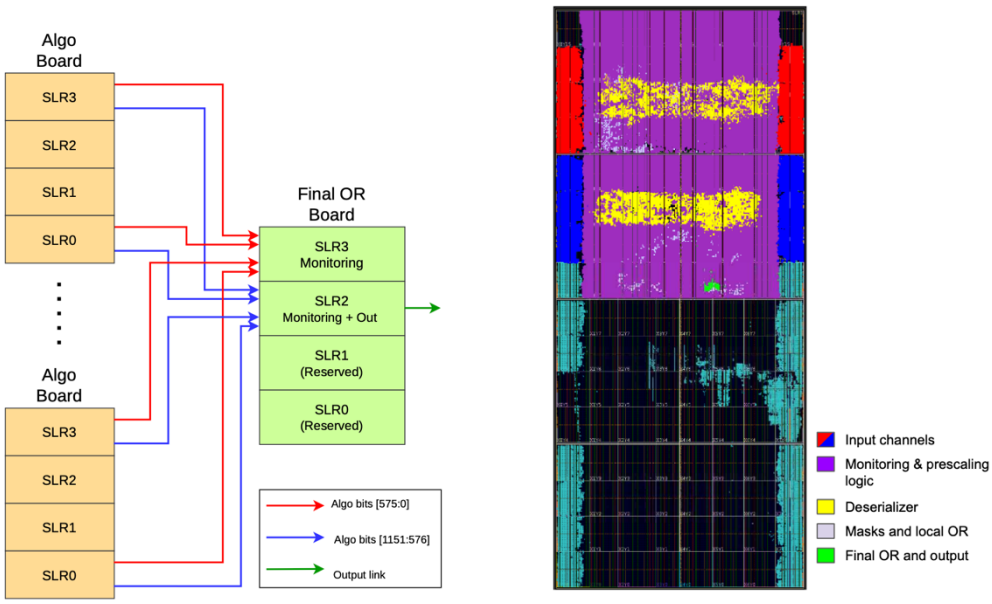

Figure 4. Left: Connectivity between Algo boards and Final-OR board. Right: Floorplan of Final-OR board on a Xilinx Virtex Ultrascale+ VU13P FPGA with 4 Super Logic Regions.

References

[1] CMS Collaboration, *The Phase-2 upgrade of the CMS level-1 trigger*, 2020, CERN-LHCC-2020-004, CMS-TDR-021

[2] CMS Collaboration, *The Phase-2 upgrade of the CMS Data Acquisition and High Level Trigger,* 2021, CERN-LHCC-2021-007, CMS-TDR-022

[3] A. Rose et al., *Serenity: An ATCA prototyping platform for CMS Phase-2, PoS* TWEPP2018 (2019) 115

[4] G. Iles et al., *A demonstration of a time multiplexed trigger for the CMS experiment,* 2012 *JINST* **7** C01060

[5] J. Wittmann et al., *Design and performance of the phase I upgrade of the CMS Global Trigger,* 2017 *JINST* **12** C01046