

System Design and Prototyping of the CMS Level-1 Calorimeter Trigger at the High-Luminosity LHC

Piyush Kumar & Bhawna Gomber (on behalf of the CMS collaboration)

CASEST, School of Physics, University of Hyderabad, Hyderabad, Telangana, India

24th IEEE Real Time Conference - ICISE, Quy Nhon, Vietnam

Date: 25 April 2024

Talk in 23rd Real Time Conference:

https://ieeexplore.ieee.org/document/10049675

Trigger Systems

System Design and Prototyping for the CMS Level-1 Trigger at the High-Luminosity LHC

Oral Presentation

to 20m

Speaker

& KUMAR, Piyush (University of Hyderabad, India)

圖

DAQ System & Trigger - I

High Luminosity LHC and Phase-2 Upgrade

❑ **High Luminosity LHC**

- \Box Higher the luminosity: the more data the experiments can gather.
- \Box Helps in the searches of new physics.
- **Phase-1:** Current CMS architecture and data taking.
- ❑ **Phase-2: architecture during HL-LHC, proposed to be start during year 2029 and onwards.**

❑ L1 trigger upgrade:

- ❑ More granular information
	- ❑ 25 times increment for calorimeter trigger.
- ❑ Advanced and more complex algorithms.
- ❑ Usage of:
	- ❑ Large FPGAs.
	- \Box High-speed optical links (25 Gbps).
- ❑ Latency: 12.5 µs
- ❑ Replacement of electronics infrastructure
	- \Box from μ TCA to ATCA [7] standard
		- ❑ Rack
		- ❑ Crate
		- ❑ Board 3

Summary of CMS HL-LHC Upgrades

[Ref: 1]

Trigger system

- □ At design parameters the LHC produces:
	- ❑ ~billions of events per second in CMS detectors .
	- ❑ Corresponds to petabyte of data per second .
- ❑ Problem :
	- □ It is <mark>impossible to store</mark> and process this large amount of data.
	- □ Most of the events are not interesting.
- ❑ Solution :
	- ❑ a drastic rate reduction must be achieved .
		- ❑ TRIGGER SYSTEM .
- ❑ Two level of triggering strategy
	- ❑ Level-1 (Hardware): 40 MHz to 750 kHz
	- □ High level trigger (Software): 750 kHz to 7.5 kHz.
- □ Phase-2 Level-1 Trigger
	- ❑ Processing CMS three sub -detector system.
		- ❑ Calorimeter
		- Tracker (First time inclusion in the system)
		- ❑ Muon system
- \Box Final decision by global trigger (GT).

4

CMS L1 trigger - a hardware perspective

CMS detector and Front-end electronics

Backend and trigger system at CMS USC: electronic rack hosting multiple crate

Trigger algorithm on FPGA APx FPGA board

Crate hosting 12 APx FPGA boards

APx Philosophy

- ❑ Aim is to provide generic trigger processing board to fulfill the requirement of Phase -2 in terms of
	- □ Computational power.
	- ❑ High bandwidth.
- ❑ Designed at University of Wisconsin.
- \Box Features single FPGA as the main processing element.
- ❑ APd1 First demonstrator board. ❑ Based on XCVU9P FPGA (3 SLR)
	- ❑ SLR: Silicon logic region [Ref: 8]
	- ❑ Support 100 of 25 Gbps optical links for algorithm.
- \Box SLR division and links capacity of the APd1 board have larger design implication.

APd1 board [Ref: 2]

Phase-2 Level-1 Calorimeter Trigger Architecture

❑ Two level architecture.

- ❑ Regional calorimeter trigger (RCT).
- ❑ Global calorimeter trigger (GCT).
- \Box Proposed to implement using 46 APx boards.
- \Box The levels are further subdivided into three layers.
	- ❑ Layer-1
		- ❑ RCT: 36 APx boards.
	- ❑ Layer-2
		- ❑ GCT barrel: 3 APx boards.
		- ❑ GCT endcap: 6 APx boards.

❑ Layer-3

- ❑ GCT sum: 1 APx boards
- ❑ Latency requirement:
	- \Box Output to correlator trigger: 3 μ s
	- \Box Output to global trigger: 4 μ s
- ❑ Device Utilization: < 50%
- \Box Main feature: Identical division of the calorimeter detector.
	- \Box Can be prototyped using only 4 boards. The contract of t

Calorimeter Trigger Objects

- \Box The L1 trigger algorithms are developed to identify physics objects, i.e., electrons, photons, hadrons (jets) and missing transverse energy (MET).
- Electromagnetic and hadrons particles interact differently in the calorimeter.
	- ❑ Different clustering scheme
		- ❑ Different cluster sizes

Major Design and Prototyping Constraints

Latency (highest priority)

- ❑ Algorithm
- Data transmissions.

Utilization

❑ Device utilization is preferred to be low $< 50\%$.

❑ 10 years of operations.

Input/Output constraints

- ❑ 100 links available in APd1
- Decides the algorithm processing region.
- ❑ The transceivers (MGTs) are distributed throughout the FPGA.
	- \Box Algorithm require early partitioning and floorplanning.

❑ Two important cases in Prototyping

❑ Case-1:

 \Box Algorithm can cover the detector in identical region with one region require > 100 links

❑ Case-2:

- \Box Algorithm can process a detector region using < 100 links.
- ❑ However, the algorithm coverage is not identical.
- ❑ Need to develop multiple algorithm.

Making prototype difficult.

❑ Need to find a balance between detector coverage, required links, and identical division.

Make prototyping easier.

RCT region coverage

Regional Calorimeter Trigger (RCT) Prototyping

RCT8x4 **HCAL** RCT9x4 \Box RCT v2.0 functionality is sub-2 links ECAL 36 links ECAL 36 links divided into three individual $|12|$ $|12\rangle$ 12 **IP1: clustering** IPs IP1 for IP1 for **IP2 : sorting** \Box IP1 • Crystals input • Crystals input processing processing **IP3 : ECAL + HCAL** \Box IP2 • Making clusters • Making clusters and towers and towers **APx-FS** ❑ IP3 ❑ Implementation in 2 SLRs. IP2 for P2 for ❑ SLR2 Stitching in eta · Stitching in eta Sorting clusters • Sorting clusters **BUFFER** ❑ SLR1 SLR₂ $\sqrt{2}$ SLR1 **RCT** IP₃ Buffering to hold **SLR** η • ECAL + HCAL merging **HCAL input till ECAL** ❑ The HCAL towers information crossing • Final stitching and sorting processing in IP1 via SLL and IP2 at IP3. $\overline{\smash{\succ}}$ Output to GCT **Algorithm post Utilization** ❑ Bitstream is successfully **Prototyping and** tested **FMAX** SLIR ❑ Using LHC events LUT 16% (standalone mode) FF 13% \Box 3 unique physics events. F_{MAX} 361 MHz \Box Latency and F_{MAX} : 1.2 µs, 361

GCT Barrel

❑ Process 16 RCT boards.

❑ Functionality division (SLR2 and SLR0):

❑ IP1

- \Box IP2
- ❑ A time-slice module developed in VHDL for sending the data to correlator trigger.

❑ SLR1:

❑ CaloObject algorithm

❑ Jets

- ❑ Taus
- ❑ Partial MET

❑ Input: 64 links

- ❑ Output:
	- ❑ 48 links to correlator trigger
	- ❑ 6 links to GCT sum board
- ❑ Bitfile is generated and tested with LHC events of test vector.

GCT Endcap

- ❑ HGCAL backend sends detector information in 18 time-slices (TM18).
	- ❑ Algorithm at calorimeter trigger runs without any time-multiplexing (TM1).
	- \Box Demultiplexing (TM18 => TM1) is required before GCT endcap algorithm.
- ❑ DEMUX algorithm is developed in VHDL
	- \Box serving two time slices per SLR.
	- \Box Reduces the utilization by half.
- ❑ Prototyping one SLR can easily scale to one board
	- ❑ One board can easily scale to 6 boards.
	- \Box Make prototyping faster.
- □ Floor planned in two separate regions to reduce the net delays
	- ❑ DEMUX: 2 clock regions
	- ❑ Jet algorithm: 6 clock regions
- Bitfile is tested for multiple LHC events.

Prototyping Calorimeter Trigger System

- □ Prototyping of the complete CALO system using 4 APd1 board.
- ❑ The boards are aligned with MC input pattern filled at
	- ❑ ECAL Crystals and HCAL towers: input buffers of RCT board
	- ❑ HGCAL tower and clusters: input buffers of GCT endcap board.
- ❑ Output is captured at GCT SUM and matches with \Box MC == HLS == RTL == Bitfile

Summary

- ❑ An architecture of Phase-2 calorimeter trigger is proposed which uses 46 APx FPGA boards.
- \Box The architecture comprises two level and further sub-divided in three layers.
- ❑ Algorithms are developed for all board flavors which satisfies the main constraint of latency and utilization.
- ❑ Algorithms are tested and validated on APd1 board. \Box MC == HLS == RTL == Bitfile
- \Box Following are the multi-board tests are performed and validated successfully.
	- \Box 2 board test: RCT v2.0 => GCT barrel
	- \Box 3 board test: GCT barrel, GCT endcap => GCT sum
	- ❑ 4 board test: RCT, GCT barrel, GCT endcap, GCT sum.
- ❑ With availability of algorithms prototyped at each layer and input test vectors, entire Phase-2 calorimeter trigger is prototyped using 4 APd1 boards and 16 optical links.

References

- ❑ [1] The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments. J. Phys.: Conf. Ser., 706(2):022002, 2016. doi: 10.1088/1742-6596/706/2/022002. URL https://cds.cern.ch/record/2263093.
- ❑ [2] The Phase-2 Upgrade of the CMS Level-1 Trigger. (2020). [Techreport]. CERN. https://cds.cern.ch/record/2714892.
- ❑ [3] Dris, Stefanos & Foudas, C & Troska, J.. (2010). Performance of the CMS Tracker Optical Links and Future Upgrade Using Bandwidth Efficient Digital Modulation.
- $□$ [4] Julia Handl. Jet Energy Calibration of the CMS Detector with Z(→e+e) + Jet Events at \sqrt{s} = 13 TeV, 2016.
- ❑ [5] https://home.cern/resources/image/experiments/cms-images-gallery
- ❑ [6] https://en.wikipedia.org/wiki/Large_Hadron_Collider
- ❑ [7] https://www.picmg.org/openstandards/advancedtca/
- ❑ [8] https://docs.xilinx.com/v/u/en-US/wp380_Stacked_Silicon_Interconnect_Technology

THANK YOU…

BACKUP…

GCT sum Prototyping

- \Box Final layer of the calorimeter trigger
- \Box Combine all the information of the calorimeter trigger.
- \Box Re-assemble the data according to the global trigger requirements and send it in six time slices (TM6).
- \Box 32 input links
	- ❑ 8 x Three GCT barrel board.
	- ❑ 8 from GCT endcap board.
- ❑ Calculate the final MET
	- ❑ Using CORDIC algorithm.
		- \Box cosφ
		- ❑ sinφ
- ❑ 6 output links.

GCT Sum including CORDIC 19

RCT and GCT barrel region coverage

Design flow

CMS Phase-2 L1 Calorimeter trigger

❑ Processing four calorimeter subsystem

> ❑ Electromagnetic calorimeter (ECAL).

- ❑ Hadronic calorimeter (HCAL).
- \Box Forward hadronic (HF).
- ❑ High-granularity calorimeter (HGCAL).
- ❑ Implemented in two level
	- ❑ Regional calorimeter trigger (RCT)
	- ❑ Global calorimeter trigger (GCT).
- \Box Further sub-divided in three layers.

Regional Calorimeter Trigger (RCT)

- ❑ Algorithm summary:
	- ❑ IP1:
		- ❑ Prepare clusters and towers
	- ❑ IP2:
		- ❑ Cluster merging
		- ❑ Sorting of clusters
	- ❑ IP3
		- ❑ Final stitching
		- ❑ Final sorting
		- ❑ ECAL and HCAL merging.

□ IP1 performs the stitching of the RCT region

❑ Latency of the algorithm in all the direction ❑ 3 clock cycles

□ Create full tower information. ❑ Cluster + tower

❑ Full tower latency: 2

IP2: PF Clustering

❑ Input: Full tower of 10 RCT region.

❑ Procedure:

❑ Creating overlapping region \Box 17x4 -> 21x8

- ❑ Finding peak tower in 21x8 region.
- ❑ Calculating 3x3 PF cluster energy.
- ❑ Removing the peak tower for next iteration
- ❑ Repeating the steps eight times.
- ❑ Output: 48 PF cluster in 6 RCT boards.
- ❑ Latency of making eight PF cluster: 80 clock cycle.

Algo

component

Stitching in eta

Stitching in phi

48 PF clusters

Total latency

LUT

 FF

 F_{MAX}

Full towers

Jet and tau algorithm

□ Input: Full tower of 16 RCT region (34 x 32).

□ Subdivided in two half (positive and negative eta).

❑ Uses supertower methodology to reduce the geometry in threefold.

> ❑ Supertower: group of 3x3 tower.

 \Box 34x32 tower region =>:

 \Box + η: 12 x 8 supertowers: 6 jets

 \Box - η: 12 x 8 supertowers: 6 jets

❑ One jet region: 9x9 in tower ❑ 3x3 in supertower.

GCT Endcap

- ❑ Prepare jet and taus cluster.
- ❑ Internal input links: 108
- ❑ Geometry: 20 x 72 endcap towers
- ❑ Create supertower for jet calculation
	- \Box 20x72 => 6x24
- ❑ Jet calculation step
	- ❑ Finding peak in the region of 6x24
	- ❑ Calculating jet cluster of size 3x3 around the peak
	- ❑ Removing the peak for next iteration
- ❑ Total 6 jets
- ❑ Output: 4 links

ĥ

- The SSI technology integrate multiple Super Logic Region (SLR) components placed on a passive Silicon Interposer (fig 3).
- Each SLR contains the active circuitry common to most Xilinx FPGA (Field programmable gate array) devices. This circuitry includes large numbers of:
	- 6-input LUTs (Look-up tables)
	- Registers
	- I/O components
	- Gigabit Transceivers(GT)
	- Block memory
	- DSP blocks
	- Other blocks
- The device we are using for our synthesis and implementation is based on Xilinx SSI technology and support three SLRs.
	- Xilinx Virtex UltraScale + xcvu9p flgc2104-
		-

Xilinx Stacked Silicon Interconnect (SSI) Technology

Fig 3: Xilinx FPGA Enabled by SSI Technology*

1-e FPGA **: UG872 Large FPGA Methodology Guide*

Key design challenges…

- Throughput is increasing almost 9 times (0.28 to 2.41 Tbps) per board compared to the Phase-1 system.
	- Tackled by employing high-end FPGA devices such as XCVU9P: can handle 3.93 Tbps of bandwidth.
- However, the device capability in terms of logic performance, serial bandwidth, and on-chip memory are not scaling in the same fashion.

ATCA

- \Box Board total Area ~140 in² (~12.68" x 11 .02")
- ❑ Size advantage over microTCA
	- ❑ 50% more area, 100% more front panel
- ❑ Power and associated Cooling advantage over microTCA ❑ 400% more available power.
- ❑ Crate: the 12U (U is rack unit equivalent to 1.75 inches) tall chassis delivers power, backplane connectivity, cooling, and the slots to maintain up to 16 boards.
- ❑ Rack: A rack (generally 46U high) delivers a rigid framework abiding up to three shelves.

APd1

- ❑ Controls in the APx are provided by the IPMC and the ELM. The IPMC board is responsible for crate power on/off control
- ❑ IPMC also communicate with the crate IPM controller, known as the Shelf Manager.
- \Box Once booted, the ELM Linux system provides configuration and operational support for the platform. This includes initialization of FPGAs
- ❑ (bitfile loading and register/memory initialization) and configuration of support devices such as Firefly optical modules

Phase-1 Architecture and its Drawback

- ❑ Current architecture support 6.4 Gbps of link bandwidth which is insufficient to handle the 25 times increase in input granularity.
- ❑ The current architecture lack the support for the new radiation hardened HGCAL trigger.
- \Box The FPGA in CTP7 board viz. Xilinx Virtex-7 is unable to satisfy the demand of the Phase-2 calorimeter trigger
	- \Box which requires additional MGTs with high bandwidth
	- \Box more logic capacity in terms of
		- \Box LUTs
		- ❑ FF and
		- ❑ DSPs
- \Box On the algorithm side, the current system works over the tower's granularity
	- ❑ requires redesigning to work on more precise input, such as ECAL crystals.
	- \Box The current doesn't support the particle-flow (PF) clustering vital for the PF algorithm at the correlator trigger.

Phase-1 Architecture and its Drawback

- ❑ Input primitives: ECAL, HCAL, and HF tower
- Input bandwidth of 6.4 Gbps.
- ❑ Organized in two layer
	- ❑ Calo Layer-1 using 18 Calorimeter Trigger Processor board (CTP7).
	- ❑ Calo Layer-2 using 10 Master Processor (MP7) board.
- ❑ Layer-1 adopted the regional architecture approach
	- \Box Each board process one phi segment of the detector.
	- \Box Employs identical algorithm.
- \Box Layer-2 adopted the TM approach and works in TMUX9 scheme.
	- ❑ One MP7 board for demultiplexing before sending the final output to GT.

Architecture choices

- FPGA boards parallelly process the detector geometry (a segment of the detector) to every bunch-crossing.
- Each board employ similar algorithms.
- ❑ May require sharing of data between the FPGAs to process the overlapping region.

Regional approach Time-multiplexing (TM)

- \Box FPGA boards run identical/different algorithms on different time slice.
- \Box The TMUXN (where N is a natural number) denotes the round-robin scheduling interval of the trigger processing board.
- Requirement for data duplication can be eradicated.
- \Box The data arrive and are processed over a more extended time than the regional approach.

Design flow – from emulator to hardware

- The trigger algorithms are implemented by using a high -level synthesis tool
	- ❑ Rapid prototyping
	- \Box Codes are written in C++
	- ❑ HLS Synthesis: generate the HDL
- ❑ Provide of latency and very early estimate of resource utilization
- \Box Can implement (place and route) the design without any pin constraint.

Integration of the algo with the APx firmware shell (FS) that provides

- Serial link instantiation (MGT hard block).
- ❑ LHC clock connectivity
	- ❑ Trigger timing and control distribution (TCDS)
- \Box AXI interface to the controlling system
- \Box Support to test and ILA debug the design
	- \Box Playback (input) and capture (output) long buffer
	- ❑ Able to emulate 113 bunch -crossing of LHC data.

ALGO

Performance Estimates

\equiv Timing (ns)

□ Summary

Clock Target Estimated Uncertainty ap_{-} clk 4.17 2.917 1.25

\Box Latency (clock cycles)

□ Summary

RCT v2.0 IP1

- ❑ Primary cluster function
	- ❑ Region: 19x24
	- ❑ Creating 3x5, 2x5, and 5x5 cluster
	- \Box This computation is done by calculating the energy of the five strips in the eta direction
	- \Box As the peak position can be arbitrary in the crystal space
	- \Box The overall latency is 6 clock cycles.

RCT v2.0 IP2

- ❑ A 16-input bitonic sorting unit is used.
- ❑ 10 CAE level.
- ❑ Considering the unit is pipelined such that the individual stage takes one clock cycle, it can render the sorted results in 10 clock cycles and can initiate a new sort each cycle

IP2

❑ Two output links

❑ Each output link carries

- ❑ 18 towers and
- ❑ 3 cluster information
- ❑ The latency measured for IP2 is 62 clock cycles.

 \Box F_{MAX}: 368 MHz

❑ LUTs: 3%

❑ FFs: 2%

IP3

- \Box Receives the information from the IP2 of both the subregions and HCAL towers.
- ❑ Perform the stitching of the clusters at the central eta boundary
- \Box Sort the 12 clusters from both the sub -regions using the *bitonic sort* algorithm.
- ❑ The final merging of the ECAL and HCAL tower is
- \Box Packs 4 output links to the GCT. With 17 towers and 2 clusters of information per link.

RCT v2.0 IP3

- ❑ The merging of ECAL and HCAL also includes the calculation of H/E. ❑ HCAL and ECAL energy ratio.
- The H/E calculation is stored in 4-bit.
- \Box The LBS bit indicates the possibility that either the ECAL or HCAL energies are zero or
	- ❑ HCAL energy is greater than or equal to the corresponding ECAL tower energy.

place and

LUT

- \Box A logarithmic scale is employed
	- ❑ cover a wide range of differences between the HCAL and ECAL energy profiles
- \Box Packs 4 output links to the GCT. With 17 towers and 2 clusters of information per link.

Algo component Latency (clock

Global Calorimeter Trigger (GCT) Barrel

Global Calorimeter Trigger (GCT) Endcap

Continue…

VHDL based design

9 words / BX

Fig 2: RTL simulation

and 360 MHz clock with \sim 2 Clock cycle

Total Negative Slack (TNS): 0 ns Number of Failing Endpoints: 0 Total Number of Endpoints: 21549 Implemented Timing Report

Fig 3: synthesis and implementation results

Continue…

Continue…

- DEMUX code is developed in VHDL
	- 4 input links
	- 24 output links
	- Clock: 360 MHz
- A 2x1 MUX is used to select the input (BX0 and BX9) for the DEMUX IP using a "sel" signal
	- Sel 0: BX0
	- Sel 1: BX9

