CASEST
CENTRE FOR ADVANCED STUDIES IN ELECTRONICS SCIENCE & TECHNOLOGY

UNIVERSITY OF HYDERABAD
सा विद्या या विमुक्तये

CMS
Compact Muon Solenoid

System Design and Prototyping for the CMS Level-1 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

IEEE NPSS NUCLEAR & PLASMA SCIENCES SOCIETY
IEEE
23rd Real Time Conference

High-Luminosity LHC (HL-LHC)

- Luminosity: indicate the performance of an accelerator
  - Proportional to: number of collisions that occur in a given amount of time
  - higher the luminosity: the more data the experiments can gather
- Aim: to deliver a much larger dataset for physics to the LHC experiments
- Pile-up: Number of simultaneous proton-proton interactions (~200)
  - With high pile-up, need more advanced selection algorithms at L1 trigger
- This increased datasets will help in the high precision measurements of:
  - Standard model (SM)
  - new territories beyond the SM (BSM)

	Instantaneous Luminosity	Pile-up (average)	Integrated luminosity
Run-2	2.1×10^{34} $cm^{-2}s^{-1}$	55	160 fb^{-1} (4 years)
HL-LHC (baseline)	5×10^{34} $cm^{-2}s^{-1}$	140	3000 fb^{-1} (10 years)
HL-LHC (ultimate)	7.5×10^{34} $cm^{-2}s^{-1}$	200	4000 fb^{-1} (10 years)

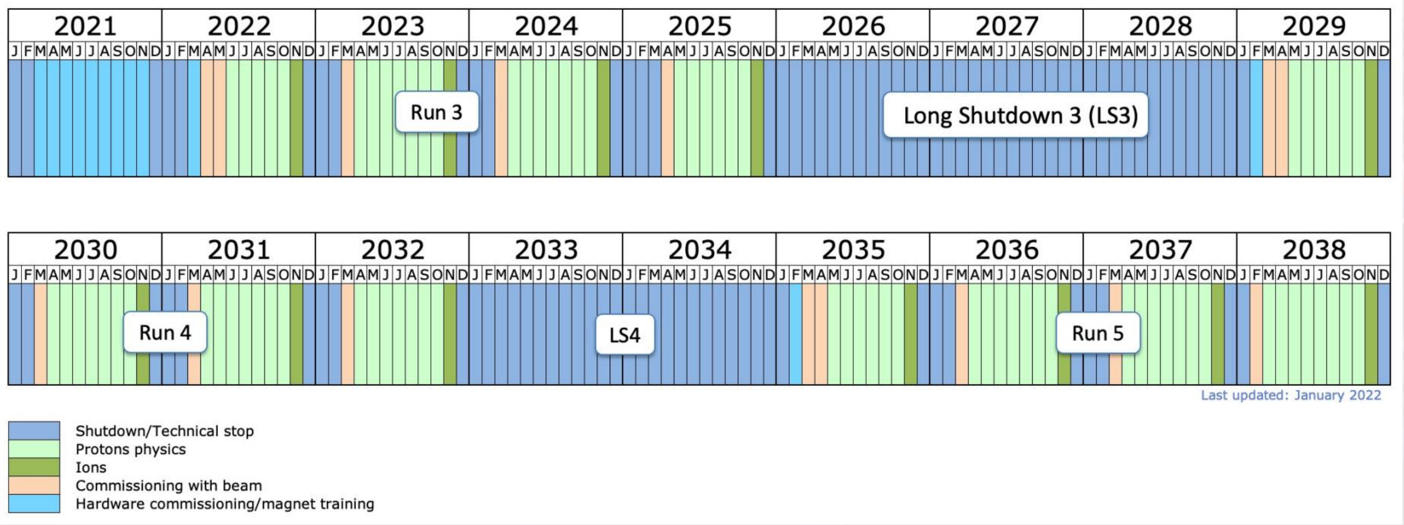


Fig: HL-LHC timeline

CMS HL-LHC upgrade

- The CMS detector planned upgrade for the HL-LHC era:
  - New pixel and strip tracking detector
  - New high-granularity calorimeter (HGCAL) of the endcap
  - New frontend/backend electronics for the:
    - Barrel calorimeter
      - Electromagnetic calorimeter (ECAL)
      - Hadronic calorimeter (HCAL)
    - Muon system
      - Drift tube (DT)
      - Cathode strip chambers (CSC)
  - 40 MHz Scouting system
    - can be used to scrutinize the collision events and identify potential signatures unreachable through standard trigger selection processes
  - L1 trigger:
    - Inclusion of the tracker information
    - Extensive usage of:
      - large FPGA (Virtex UltraScale+/Kintex UltraScale)
      - high-speed optical links (28 Gbps)

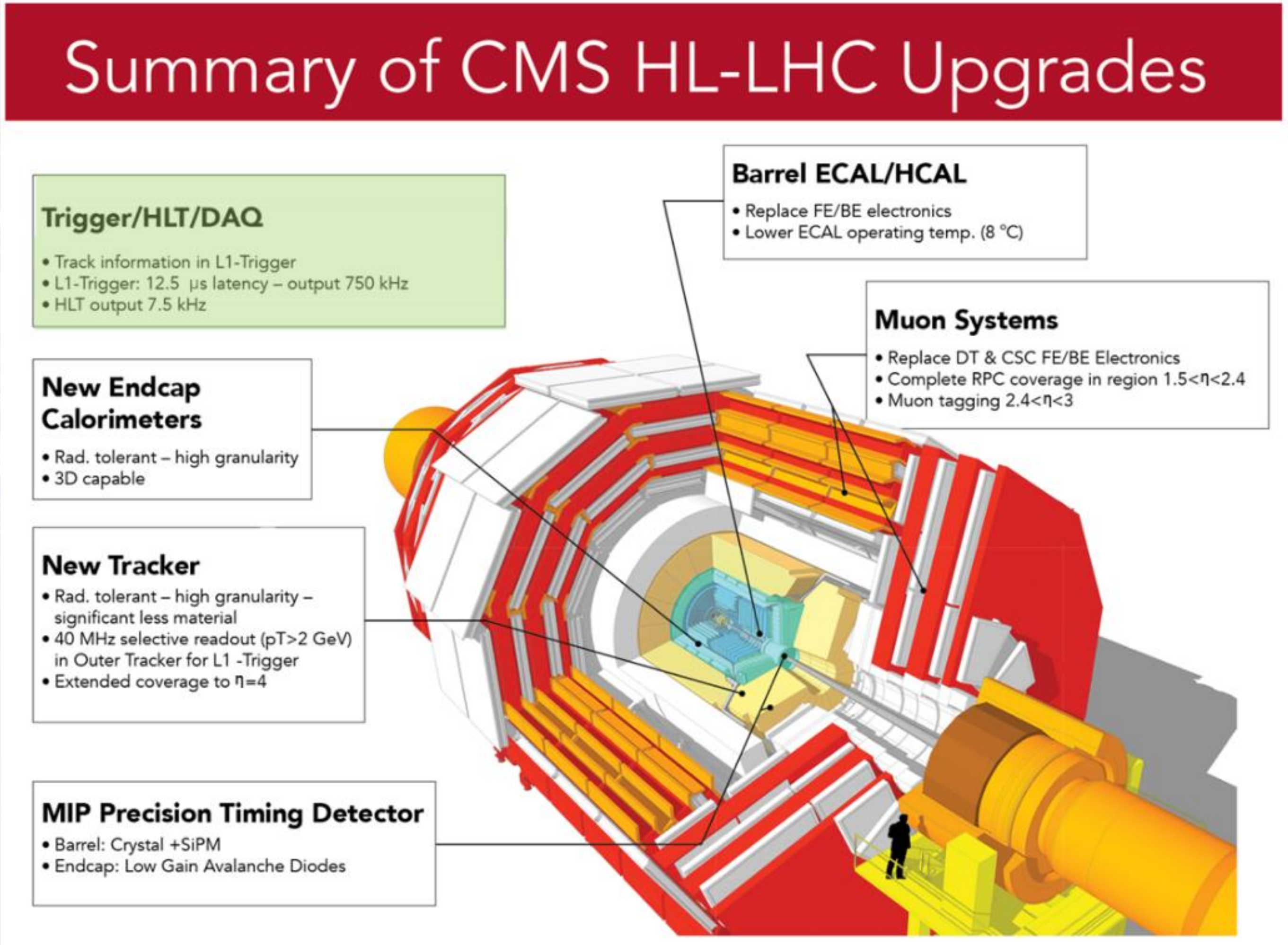


Fig: CMS detector HL-LHC upgrade

L1 trigger principle

- At design parameters the LHC produces:
  - ~10^9 events/second in CMS detectors.
  - each event is ~ 1 MB.

- 10^9 events/s x 1 Mbyte/events = 10^{15} bytes/s = 1 PB/s (1 Petabyte/second)

- Problem:
  - It is impossible to store and process this large amount of data
- Solution:
  - a drastic rate reduction has to be achieved
    - Level-1: 40 MHz to 750 kHz
    - High level trigger (HLT): 750 kHz to 7.5 kHz

- A trigger is designed to reject the uninteresting events and keep the interesting ones for physics.

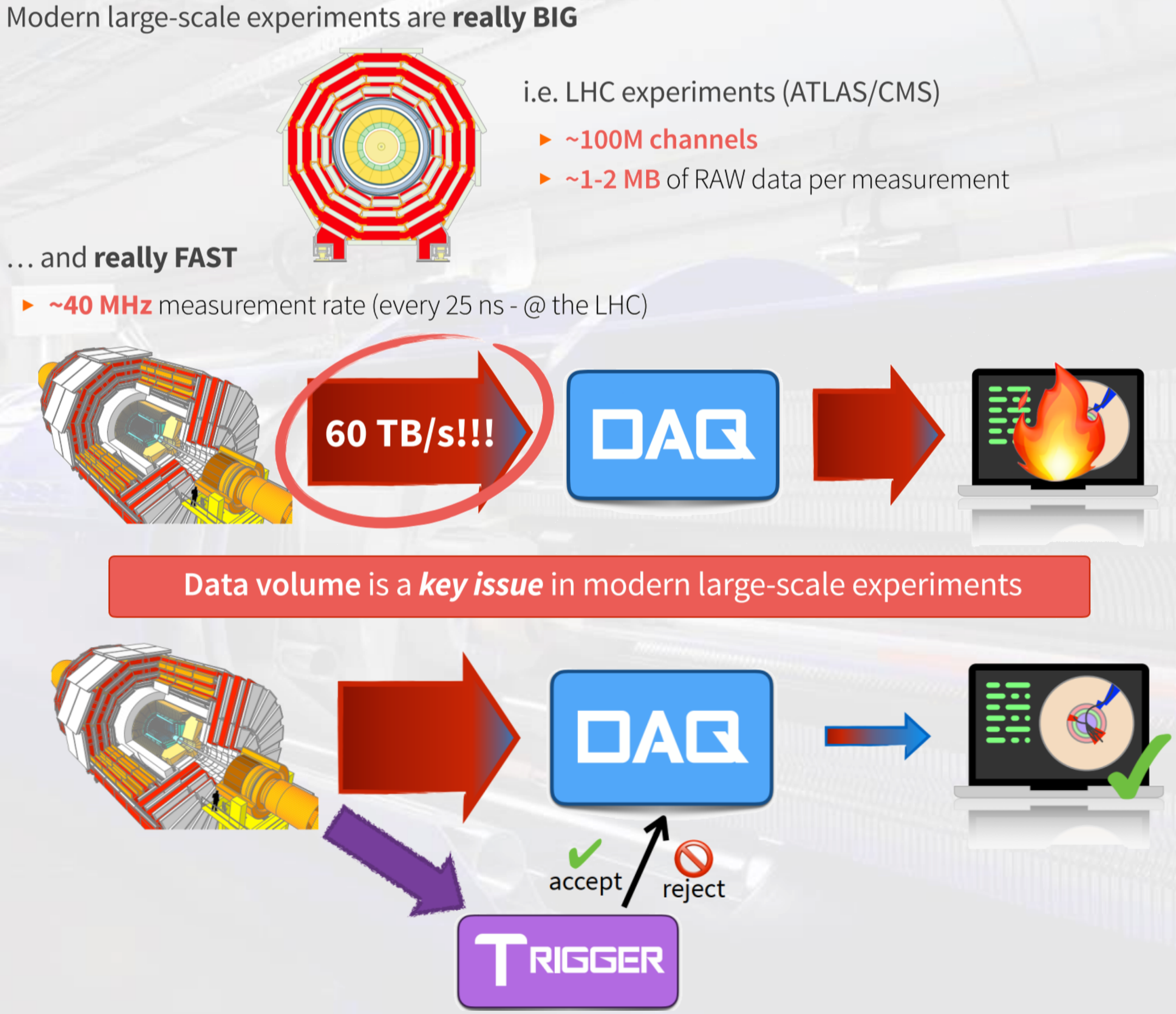


Fig: Trigger system

L1 trigger architecture

- The HL-LHC L1 trigger receives input from the backend electronics of:
  - Calorimeters
  - Muon spectrometers
  - Track finder

- Calorimeter trigger: (creating clusters from the energy deposited by the particle in the calorimeter)
  - Regional calorimeter trigger (RCT)
    - Barrel ECAL and HCAL
  - Global calorimeter trigger
    - RCT, forward hadronic (HF), and HGCAL

- Correlator trigger (CT) receives input from all the trigger sub-system:
  - Aim: identifying and reconstructing all the particles with a particle flow algorithm

- Global trigger:
  - Aim: Issues the final L1 trigger decision

- Input rate: 40 MHz
- Increased output rate: 100 kHz => 750 kHz
- Increased latency: 3.8 μS => 12.5 μS

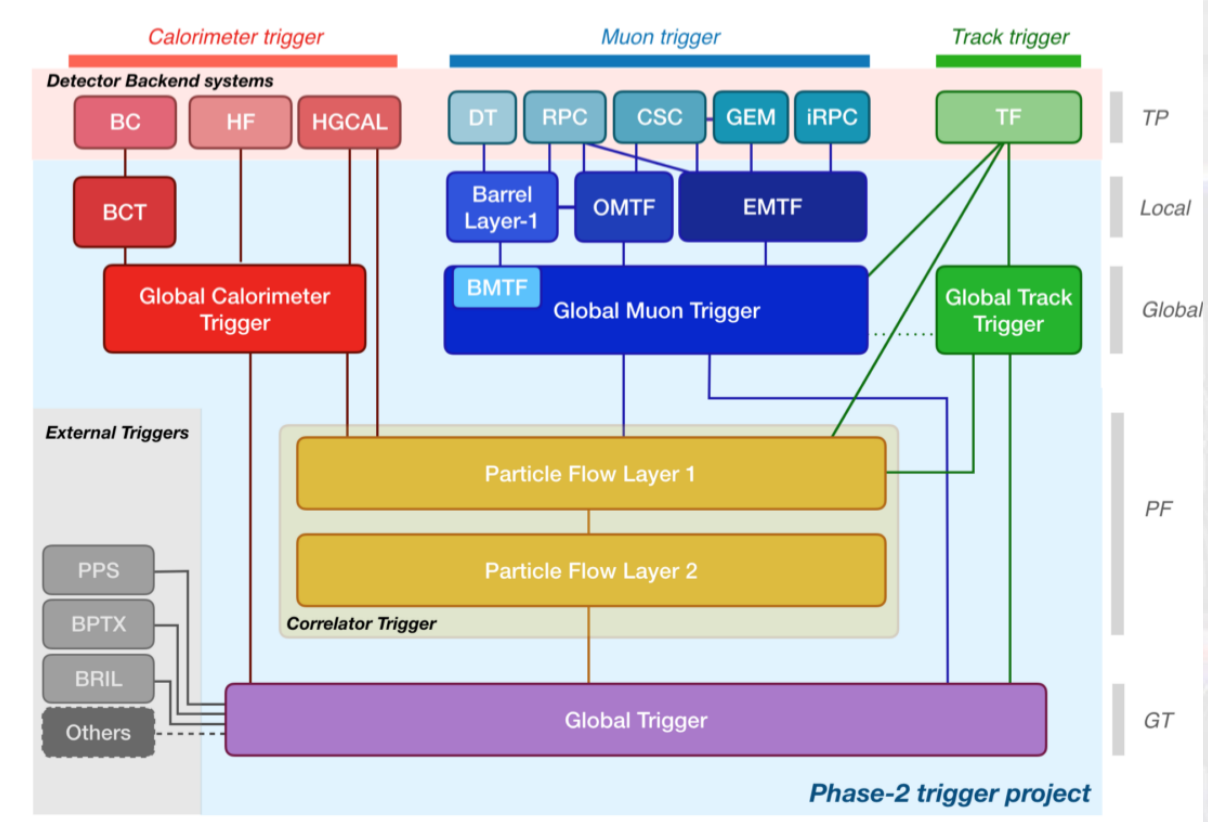


Fig: HL-LHC L1 trigger architecture

L1 trigger architecture

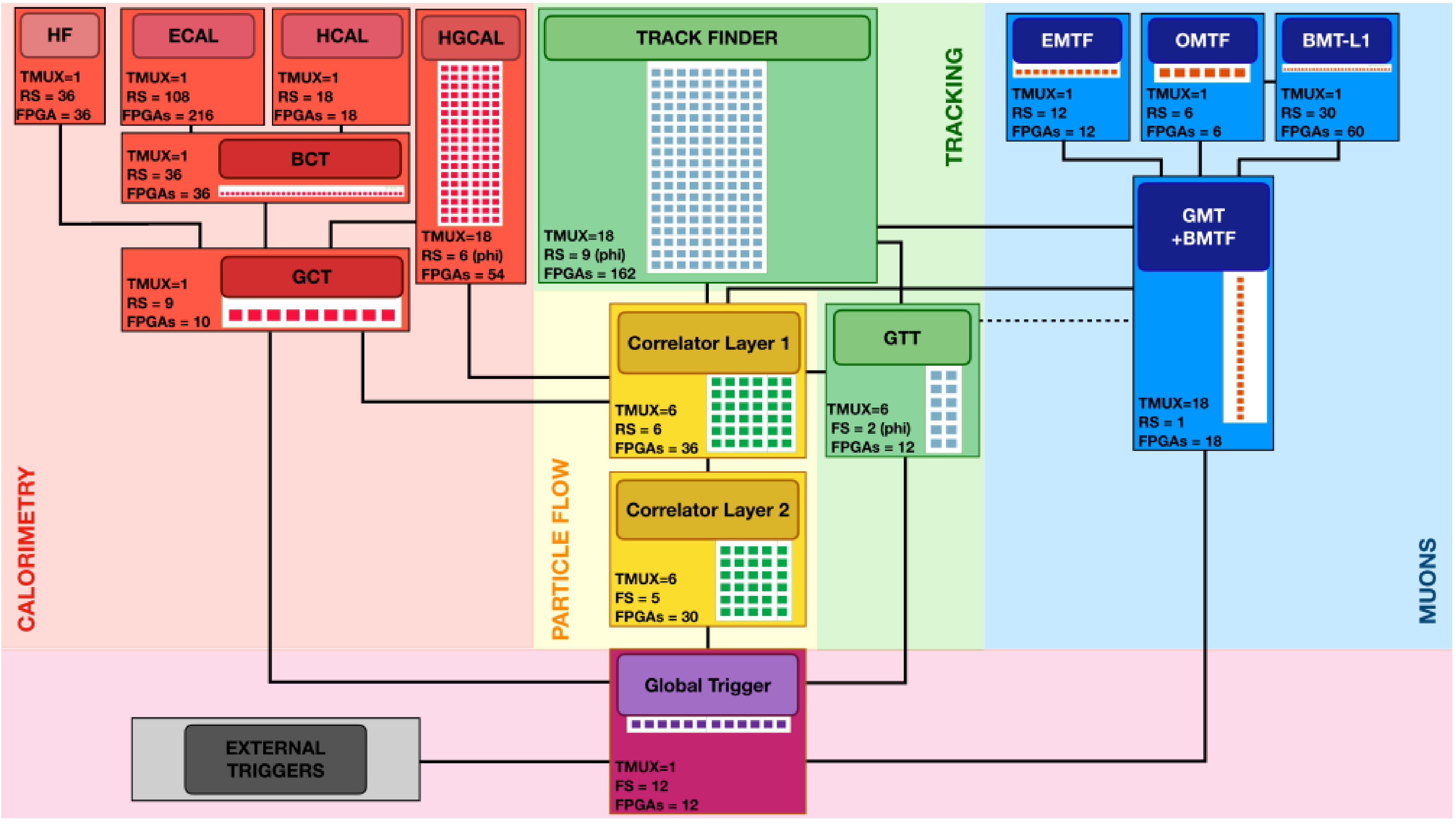


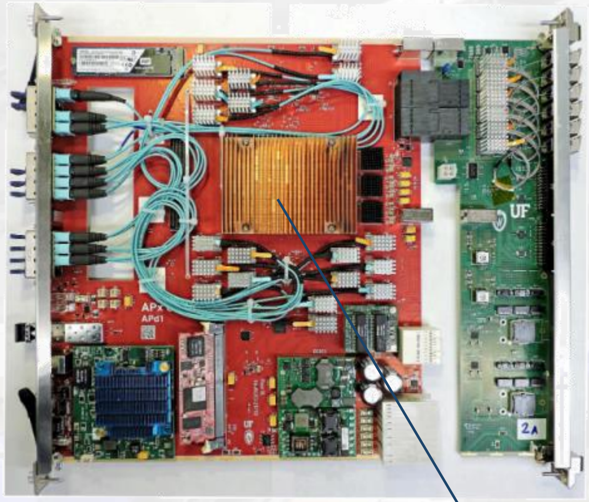
Fig: CMS Phase-2 L1 trigger design. Mentioning the time-multiplexing (TMUX) period, regional (RS) and functional segmentation (FS), and the number of FPGAs for each architecture component.

Technology R&D examples

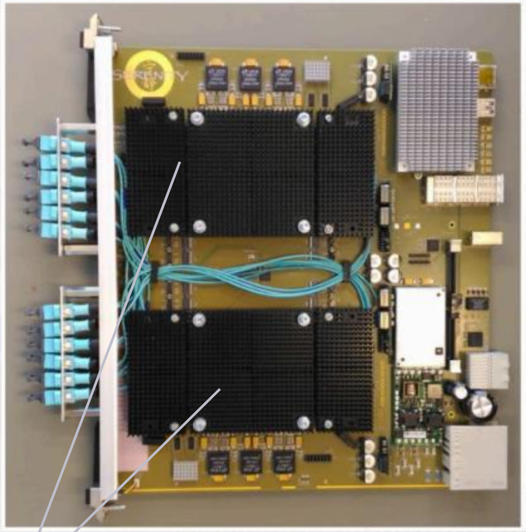
- ATCA based electronics
  - Generic high I/O (> 100) processing boards
  - One or two Virtex UltraScale+/Kintex UltraScale FPGA from Xilinx

- Wide range of testing and prototypes
  - Extensive link tests @ 28 Gb/s
  - endurance test (< 10^{-12} BER) of the FPGA quads.
  - Thermal performance test and simulation
    - Heat sink test (in order to keep operating temperature bellow 100^0C)
  - Algorithm firmware
  - Infrastructure firmware

BOARD FAMILIES

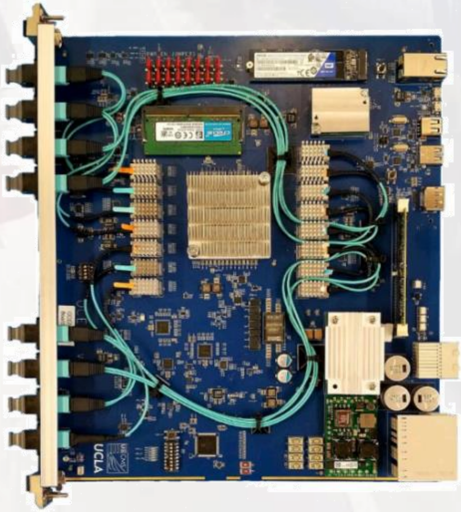


APx



Heat sink

Serenity



Ocean

BMT-L1

APx 25G quad eye scans
- 25.78125 Gbps
- Using pseudorandom binary sequence (PRBS31)
- Clock and data recovery (CDR) ON

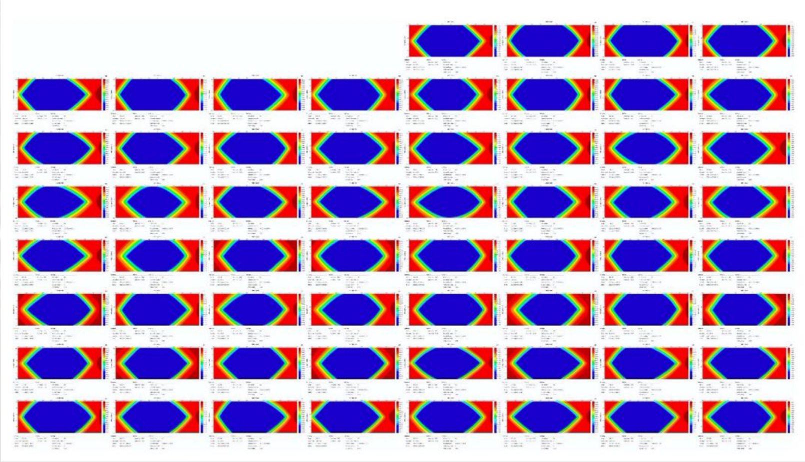


Fig: APxF 25G eye scans, quads 121-135

Trigger Algorithms Development

- The trigger algorithms are implemented by using Xilinx Vivado-HLS (high level synthesis) tool
  - Rapid prototyping
  - Codes are written in C++
  - HLS synthesizes the code to generate the RTL and
  - Provide an **early estimate** of **latency** and **resource utilization**
  - Increased **ease of collaboration** and code sharing for algorithm design

- Downstream:
  - Integration of the algo with the firmware shell (orange box) that provides
    - MGT link instantiation
    - Timing and Control Distribution System (TCDS) connectivity
    - DAQ support
    - and an AXI interface to the controlling system
  - Uses HDL wrapper for integration (magenta box)

== Performance Estimates

+ Timing (ns):
  * Summary:

Clock	Target	Estimated	Uncertainty
ap_clk	4.17	2.917	1.25

+ Latency (clock cycles):
  * Summary:

Latency min	Latency max	Interval min	Interval max	Pipeline Type
32	32	6	6	function

== Utilization Estimates

* Summary:

Name	BRAM_18K	DSP48E	FF	LUT	URAM
DSP	-	-	-	-	-
Expression	-	-	0	4	-
FIFO	-	-	-	-	-
Instance	-	-	49827	78752	-
Memory	-	-	-	-	-
Multiplexer	-	-	-	56	-
Register	0	-	3360	32	-
Total	0	0	53187	78844	0
Available SLR	1440	2280	788160	394080	320
Utilization SLR (%)	0	0	6	20	0
Available	4320	6840	2364480	1182240	960
Utilization (%)	0	0	2	6	0

Fig: Vivado-HLS performance estimates of trigger algorithm

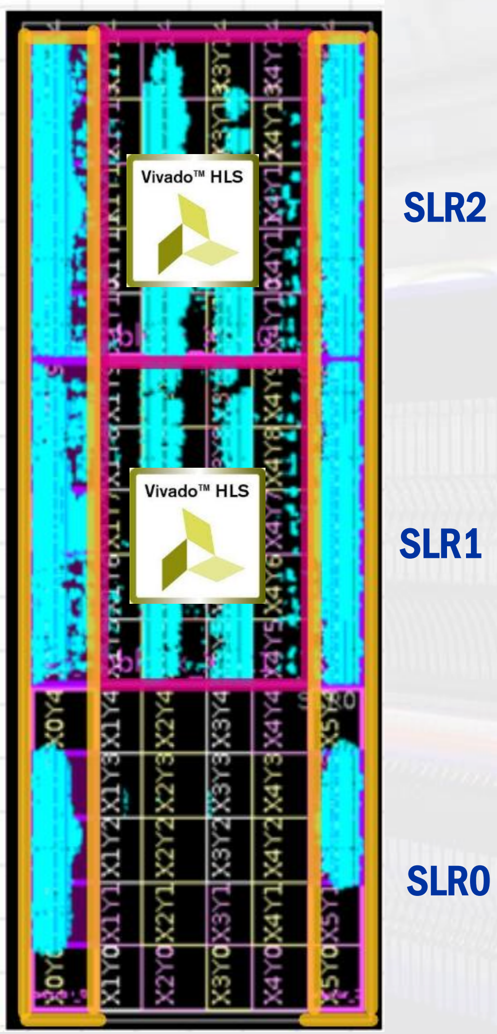


Fig: Trigger algo device implementation

Aim is to write HLS algorithms in a framework agnostic way

Barrel calorimeter segmentation

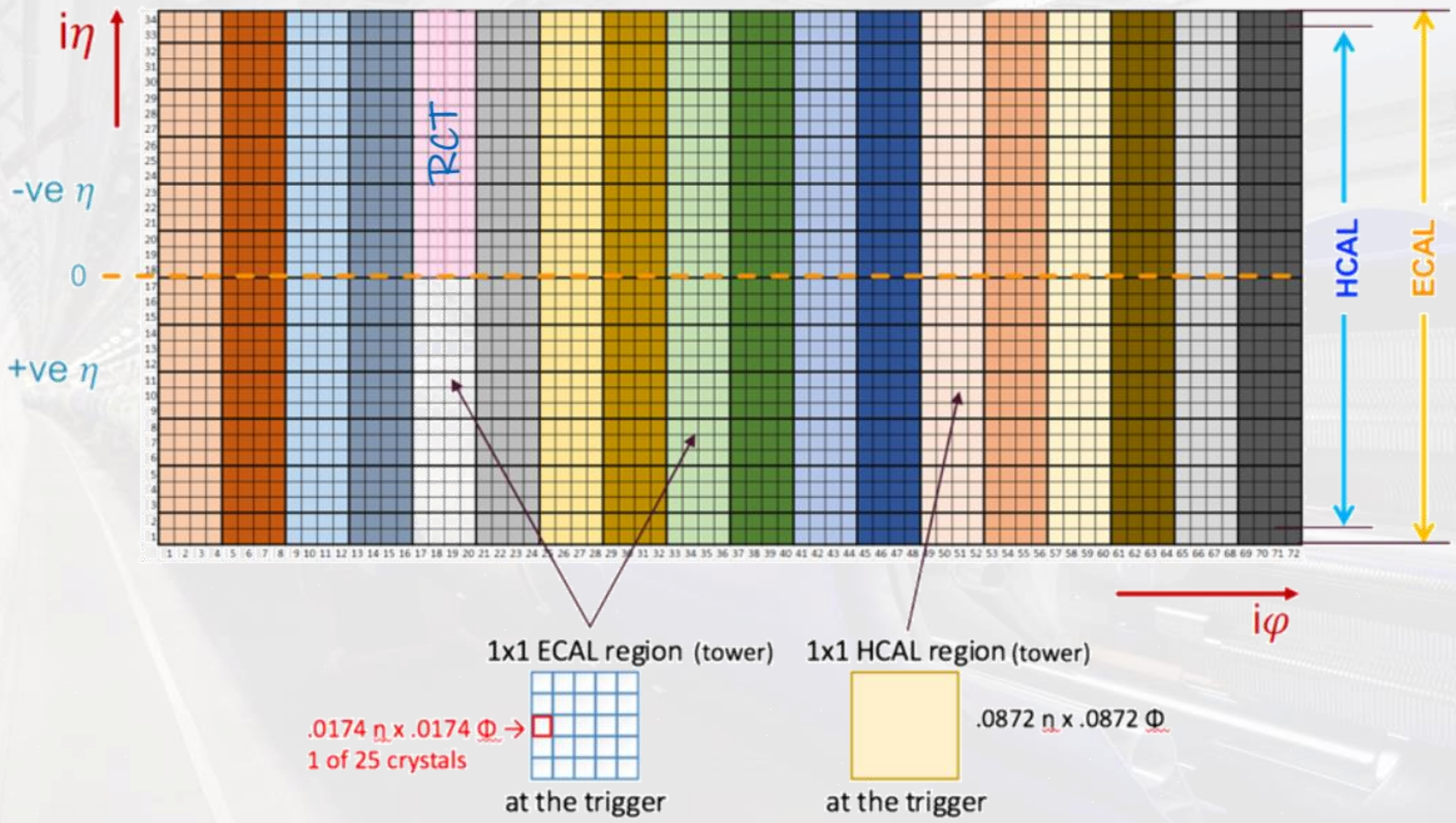


Fig: Barrel calorimeter segmentation

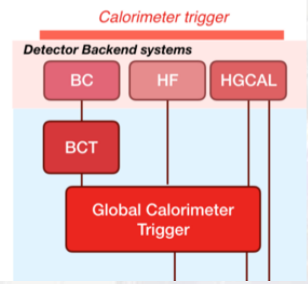
L1 Trigger Algorithms

Calorimeter Trigger

RCT geometry for the FPGA processing: 17η×4ϕ of the barrel (total 36 APx cards)

Regional Calorimeter Trigger (RCT) creates electrons/photons energy clusters and towers and sends them to Global Calorimeter Trigger (GCT)

- The Xilinx UltraScale+ XCVU9P FPGA supports 3 super logic regions (SLR).
- For efficient implementation, the algorithm is partitioned SLR wise in 2 SLR (SLR2 and SLR1)
- RCT algorithm is divided in three part
  - RCT8x4:
    - Implemented in SLR1
    - Processes the 8η x 4ϕ RCT regions
    - only ECAL.
  - RCT9x4
    - implemented in SLR2
    - processes the 9η x 4ϕ RCT regions
      - ECAL
      - 16η x 4ϕ HCAL data.
  - RCTSUM
    - implemented in SLR2
    - combines both the algorithm and sends the output to the GCT.



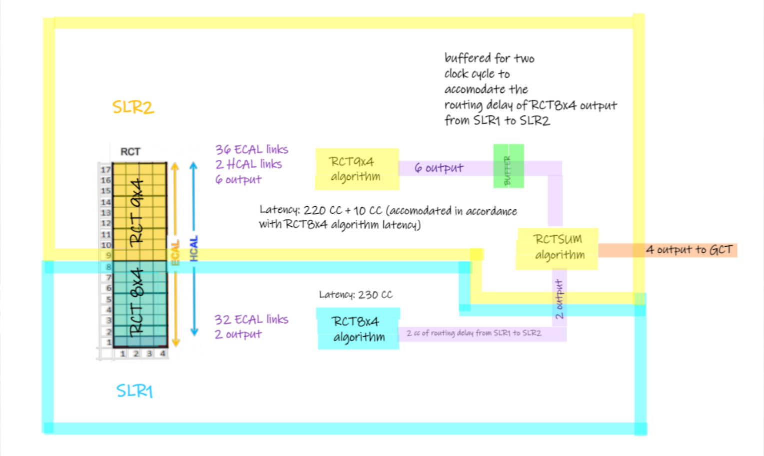


Fig: RCT algorithm organisation and dataflow

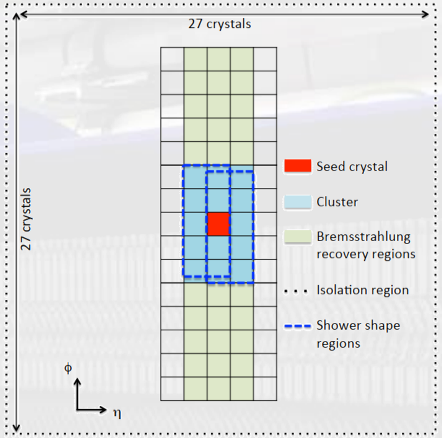


Fig: e/gamma cluster making in RCT algorithm

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	4.17	3.491	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
230	230	6	6	function

Summary

Name	BRAM_18K	DSP48E	FF	LUT	URAM
DSP	-	-	-	-	-
Expression	-	-	0	24202	-
FIFO	-	-	-	-	-
Instance	8	0	303544	464948	-
Memory	-	-	-	-	-
Multiplexer	-	-	-	16292	-
Register	30	-	23821	1813	-
Total	38	0	327365	507255	0
Available	4320	6840	2364480	1182240	960
Available SLR	1440	2280	788160	394080	320
Utilization (%)	~0	0	13	42	0
Utilization SLR (%)	2	0	41	128	0

Fig: RCT algorithm HLS results

The implementation is scalable for the region of 17η×6ϕ (can use 3 SLRs). RCT APx board will reduce from 36 to 24

RCT to GCT slice test

- The GCT algorithm (merging the energies between the RCT cards in phi direction) is synthesized in Vivado-HLS
- The RCT (SLR2 and SLR1) and GCT (SLR0) is implemented together in XCVU9P FPGA.
- Tested on a single card:
  - Replicate the 4 RCT output links x5 (20 input) ~ GCT processing 5 RCT cards
- Implementation details
  - XCVU9P-FLGC2104-1-E FPGA
  - Clock: 240 MHz
  - Link bandwidth: 16 Gbps

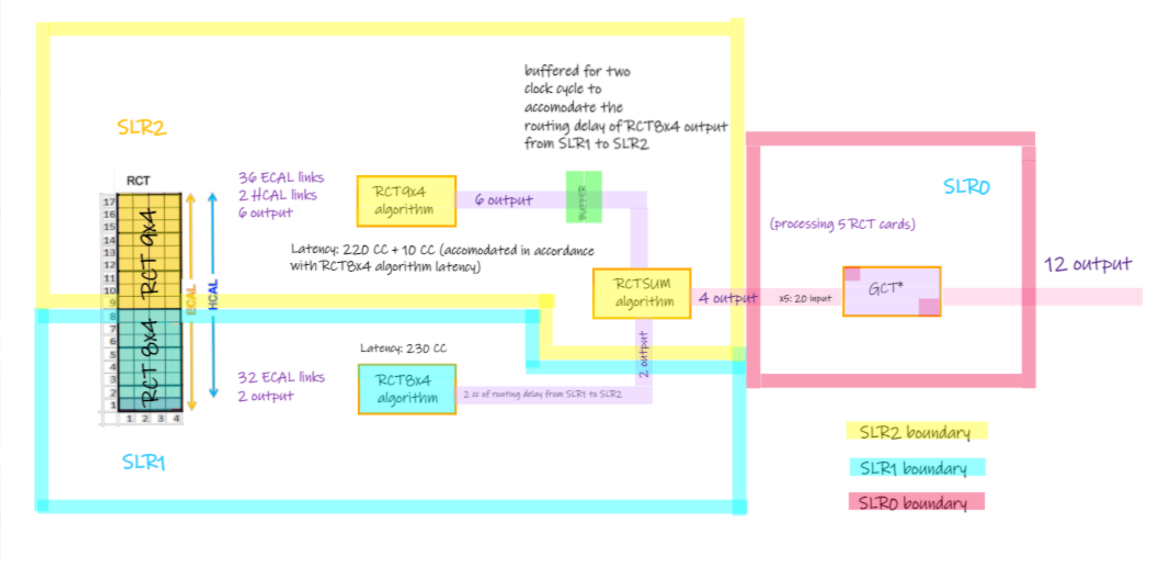


Fig: RCTTDR and GCT algorithm implementation in three SLR

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	4.17	2.909	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
120	120	6	6	function

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT	URAM
DSP	-	-	-	-	-
Expression	-	-	0	1444	-
FIFO	-	-	-	-	-
Instance	-	-	27703	146555	-
Memory	-	-	-	-	-
Multiplexer	-	-	-	56	-
Register	0	-	82036	36864	-
Total	0	0	109739	184919	0
Available	4320	6840	2364480	1182240	960
Available SLR	1440	2280	788160	394080	320
Utilization (%)	0	0	4	15	0
Utilization SLR (%)	0	0	13	46	0

Fig: GCT algorithm HLS results

RCT to GCT slice test

- The bitstream is generated and the project passes the timing constraints.
- Following are the algorithms device placement:
  - RCT8x4: SLR1
  - RCT9x4: SLR2
  - RCTSUM: SLR2
  - GCT: SLR0
- Post implementation device utilization is within the boundary.
- Bitstream is successfully tested on the APd1 (APx demonstrator board) board
  - Test vector generated via Monte Carlo physics simulations for different physics models.

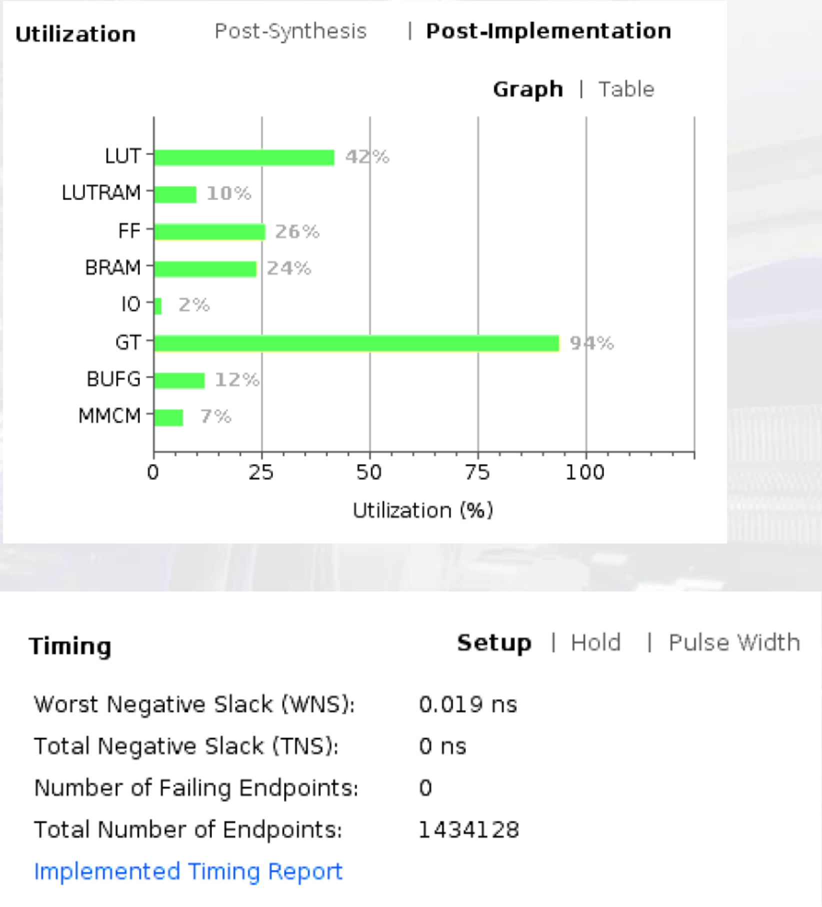


Fig: Utilization and timing summary (setup)

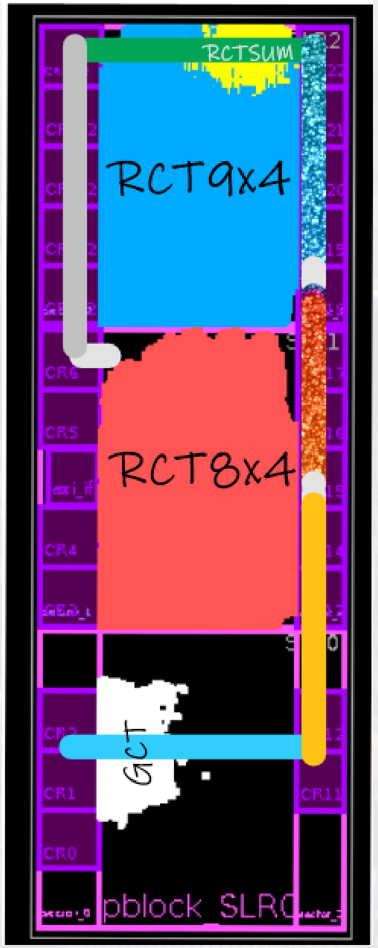


Fig: GCT device implementation

$F_{max} = 1/(4.167-0.019) \sim 241$ MHz

Muon trigger

- The function of the muon trigger:
  - Identification of the muon tracks
  - Measure momenta
- inputs in the form of muon stubs (32-64 bits each)
- Inputs (stubs) are relaying through various electronics regions:
  - Barrel:
    - Drift tube (DT)
    - Resistive plate chambers (RPC)
  - Endcap:
    - very forward extension iRPC
    - cathode strip chambers (CSC)
    - gaseous electron multiplier (GEM)
- Full implementation of the barrel algorithm
  - Tested on small KU040 FPGA
  - Algorithm clock: 160 MHz
  - BMT latency: 2.25 μS

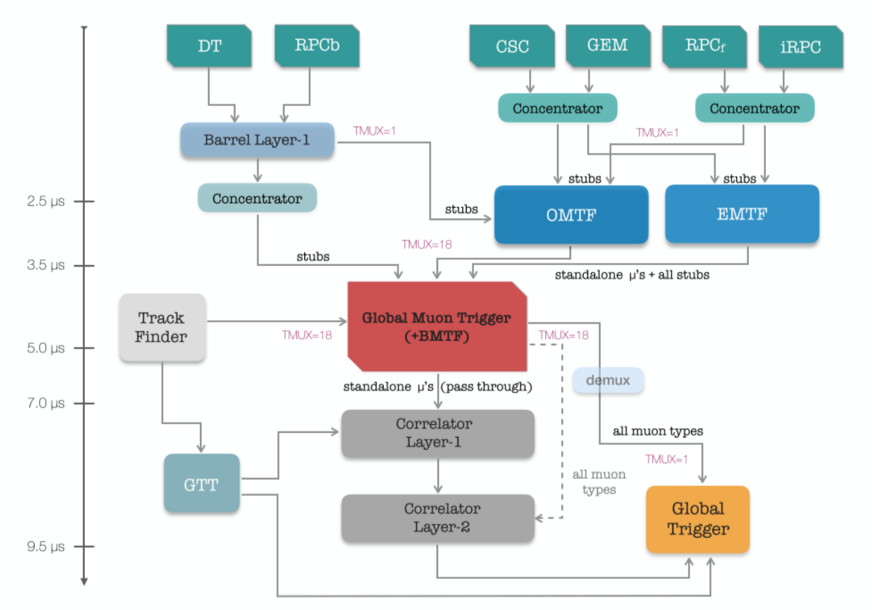


Fig: Muon trigger architecture

DSP	FF	LUTs	BRAM
10%	17%	37%	46%

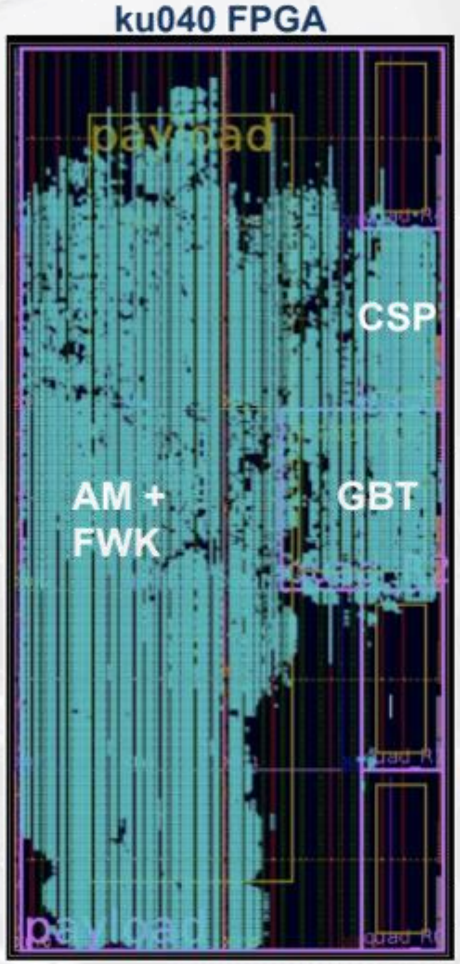


Fig: barrel algorithm implementation

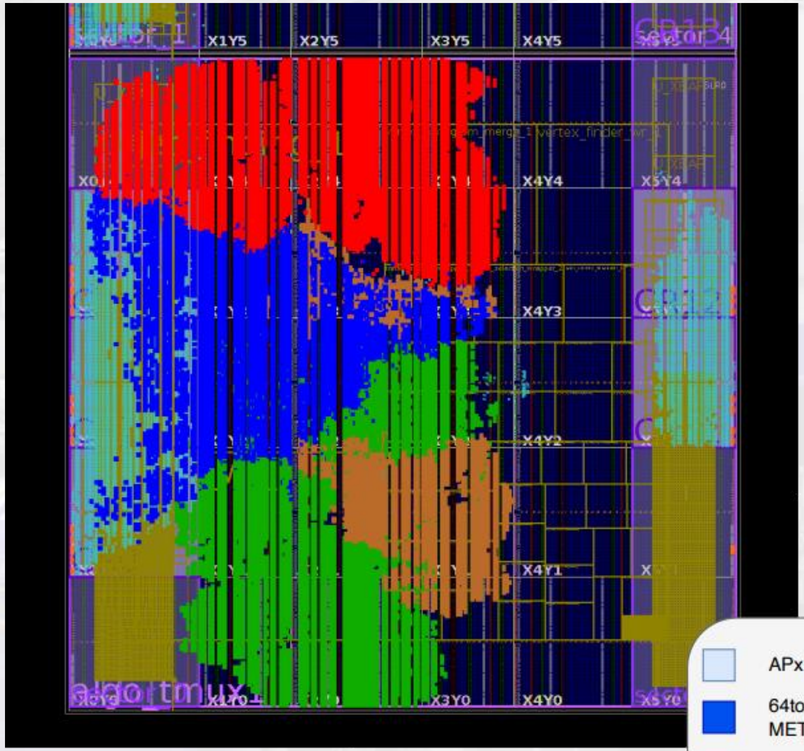
Stubs: position, bend angle, and timing information of the muons

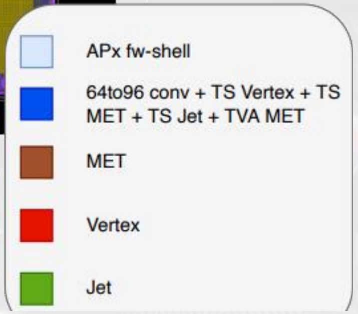
Track trigger

Global track trigger algorithm:

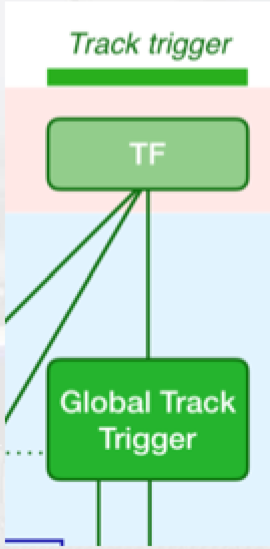
- Aim:
  - Reconstruction of the primary vertices
  - Identify track-only objects
- Uses 6 APx and 6 serenity board
- Primary Vertex (PV) Finding:
  - Origin of tracks constrained to ~1mm
  - Remove pileup to maintain manageable rates
- Track-Vertex Association:
  - Select tracks consistent with the PV
- Track-based Jet Finding:
- Track-based missing transverse energy (MET)
- Track-based Missing H_T^*

Whole algorithm fits in one SLR





Implemented in XCVU9P



Single board, GTT framework + multiple algorithms

DSP	FF	LUTs	BRAM
1%	11%	17%	21%

	Vertexing	MET/Jets
Latency	540 ns	1530 ns

H_T : *scalar sum p_T of jets*

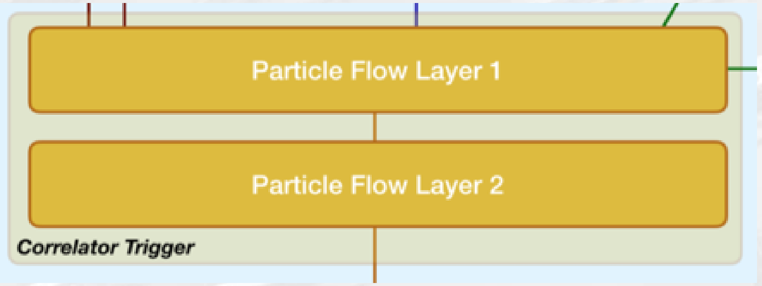
Correlator trigger

Correlator trigger layer-1

- Aim: Collect information from calorimeters/muon systems/tracker, combine them
  - reconstruct the particles and identify them.

- Employs algorithms for:
  - Particle Flow (PF) and Particle per pile-up identification (PUPPI) (barrel + endcap)
  - Jets/Missing transverse energy (MET)/H_T
  - Taus, Isolation, NN MET, electron/photon (egamma)

- Two layers:
  - Correlator Layer-1: Performs full PF+PUPPI create particle-flow candidates
  - Correlator Layer-2: use PF candidates to reconstruct physics objects



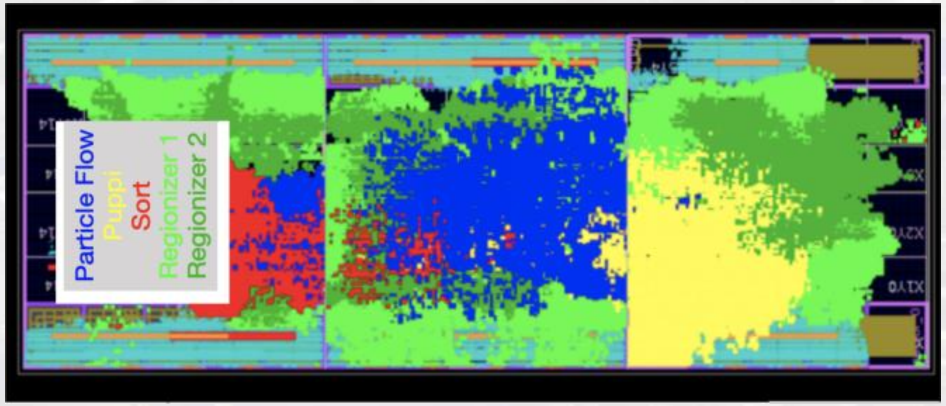


Fig: Layer-1 barrel

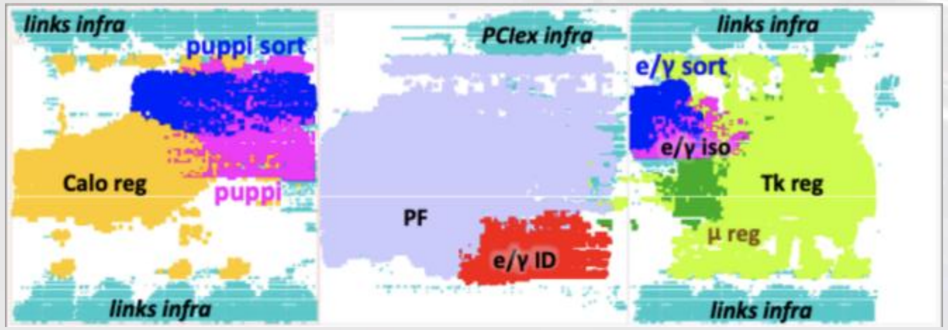


Fig: Layer-1 endcap

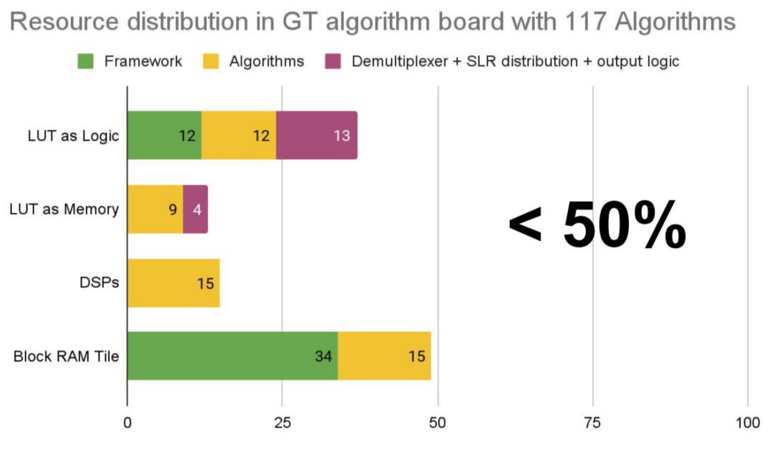
- Full working PF+PUPPI
- Barrel/endcap implemented using VU9P-2

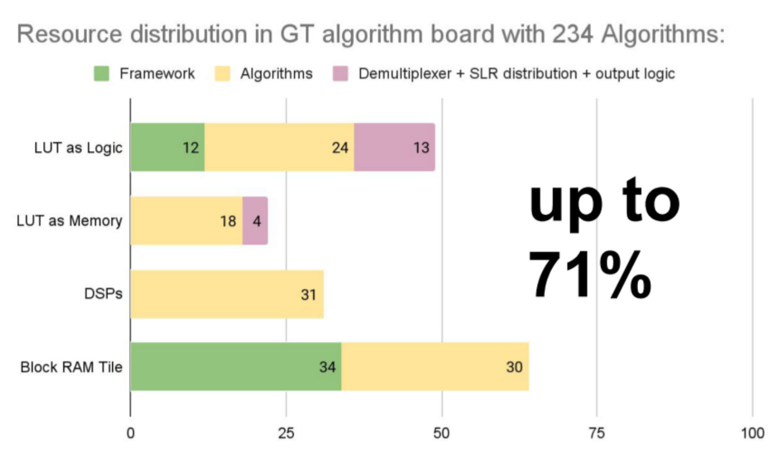
VU9P	DSP	FF	LUTs	BRAM
Barrel	33%	36%	46%	38%
Endcap	24%	24%	30%	32%

	Barrel	Endcap
Latency	1120 ns	1030 ns

Global trigger

- Final stage of the Level-1 trigger
- Aim: responsible for implementing the trigger menu
- Based on serenity board
  - XCVU9P FPGA
- Flexible design:
  - can be adapted for future algorithms
- 480 MHz algorithm clock
- Total latency of the GT Algorithm
  - ~250 ns (10 Bunch-crossing)
  - Budget: 40 BX (1000 ns)





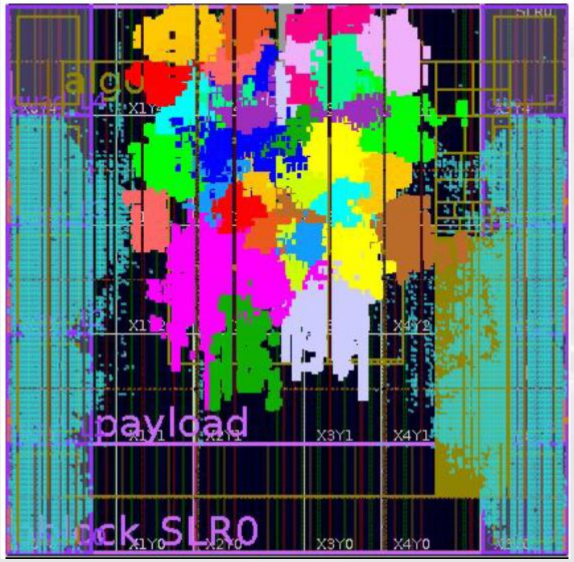


Fig: 39 algorithm placed in 1 SLR (total 117 algorithms for 3 SLR)

Flexible System

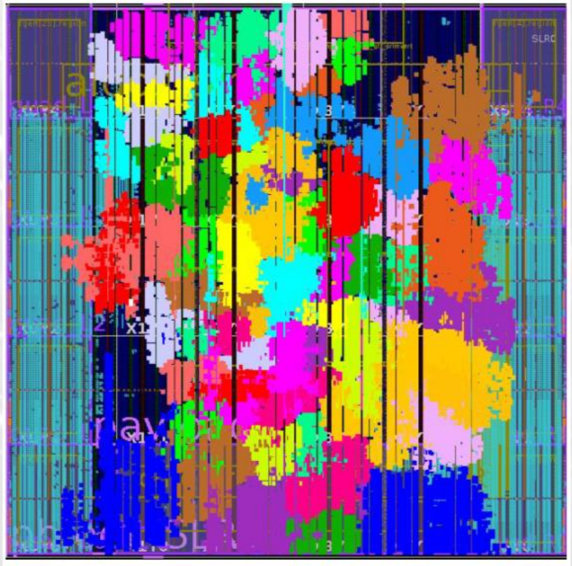


Fig: 78 algorithm placed in 1 SLR (total 238 algorithms for 3 SLR)

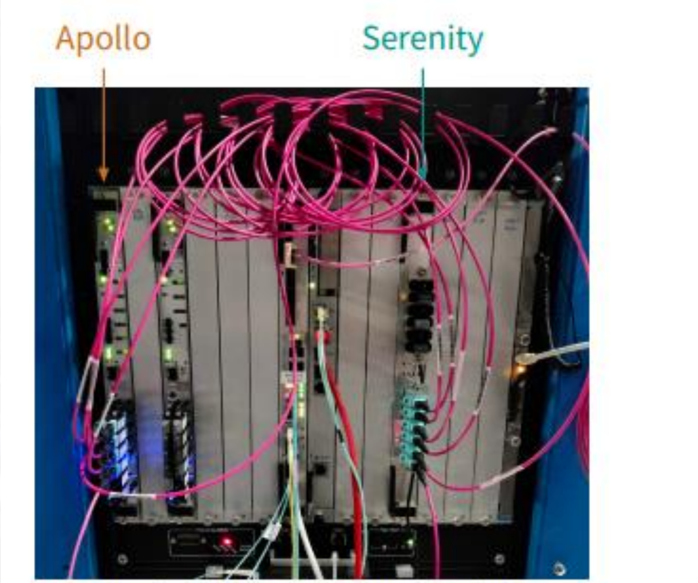
Slice test

Track finder (backend) => Global track trigger (GTT)

- VU7P Apollo => KU15P Serenity test
  - Apollo algo firmware: Final subcomponent of track finder
  - Serenity algo firmware: Vertexing algorithm
  - Tracks sent over 18 links
- inputs is injected into the buffers on Apollo
  - Generated via CMS software (CMSSW)
- Outputs is captured on the Serenity buffer
  - Compared with expectations: 100% agreement

Correlator layer 1 (Serenity) → Layer 2 (Serenity)

- Layer-1 algo input: HGCAL => jets
- Layer-2 algo output: electron/photon (egamma)
- 100% agreement with emulator



TIF crate (Apollo connected to Serenity)

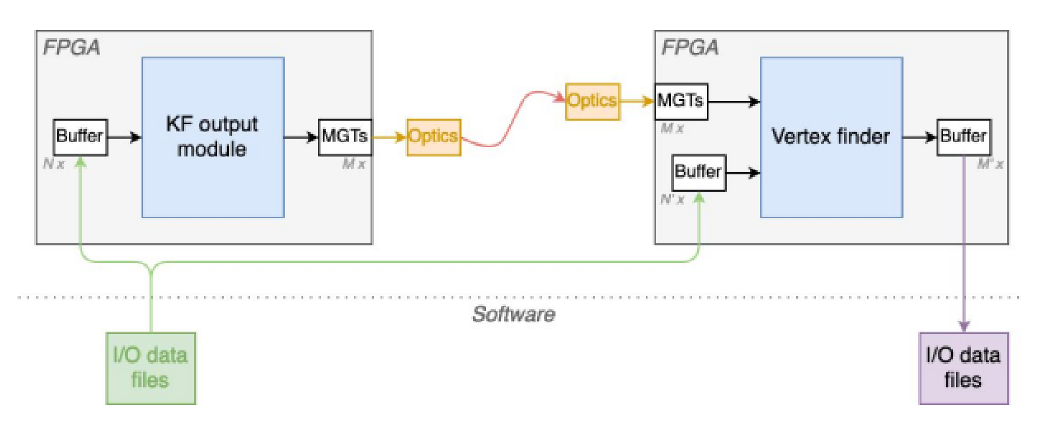


Fig: Track finder and GTT board placement in the TIF crate

Summary

- Key technological choices to leverage the HL-LHC high data-taking environment:
  - High-speed optical links (from ~10 Gbps to ~28 Gbps)
  - Large FPGAs (from Virtex-7 to Xilinx Virtex UltraScale+/ Kintex UltraScale)
  - Modular and scalable algorithm firmware
- Several FPGA boards are being developed and various tests were performed, such as:
  - The links eye scan (@25 Gbps) and
  - endurance test (< 10^{-12} BER) of the FPGA quads.
  - FPGA thermal test to explore various heat sinks options.
- Following trigger algorithms are being prepared and tested successfully on their corresponding prototyped board:
  - RCT and GCT
  - Barrel muon trigger and global muon trigger (GMT)
  - Global track trigger (GTT)
  - Correlator Layer-1 and Layer-2
  - Global track trigger
- The latency and resource utilization is well within the desired limit.
- All the testing/development is going in time with the HL-LHC schedule.

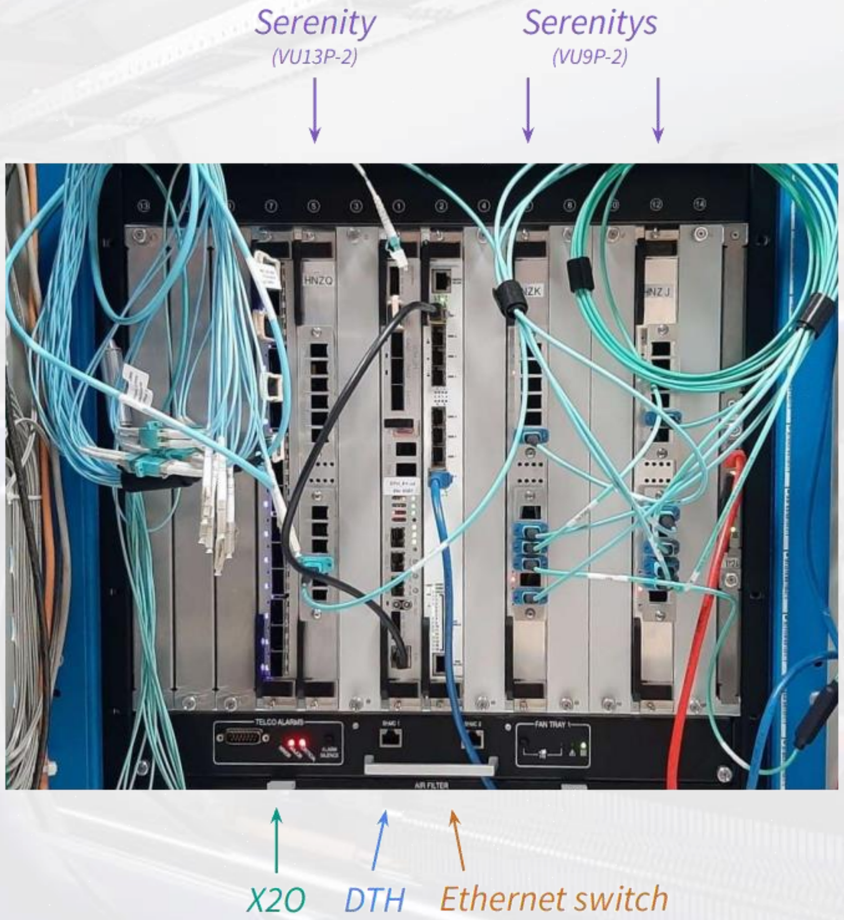


Fig: L1 trigger crate installed at CERN that houses three Serenity, X2O, and DTH (DAQ and TCDS hub) board (for multi-board testing)

The Phase-2 Upgrade of the CMS Level-1 Trigger. Technical Report CERN-LHCC2020-004. CMS-TDR-021, CERN, Geneva, Apr 147 2020. URL http://cds.cern.ch/record/2714892

धन्यवाद

Thank You

Acknowledgement

- Piyush Kumar and Bhawna Gomber acknowledges the support from IOE, University of Hyderabad through Grant Number UOH-IOE-RC2-21-006

CASEST

CENTRE FOR ADVANCED STUDIES IN
ELECTRONICS SCIENCE & TECHNOLOGY

UNIVERSITY OF HYDERABAD

सा विद्या या विमुक्तये

CMS

Compact Muon Solenoid

BACKUP...

- 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 flgc2104-1-e FPGA

Xilinx Stacked Silicon Interconnect (SSI) Technology

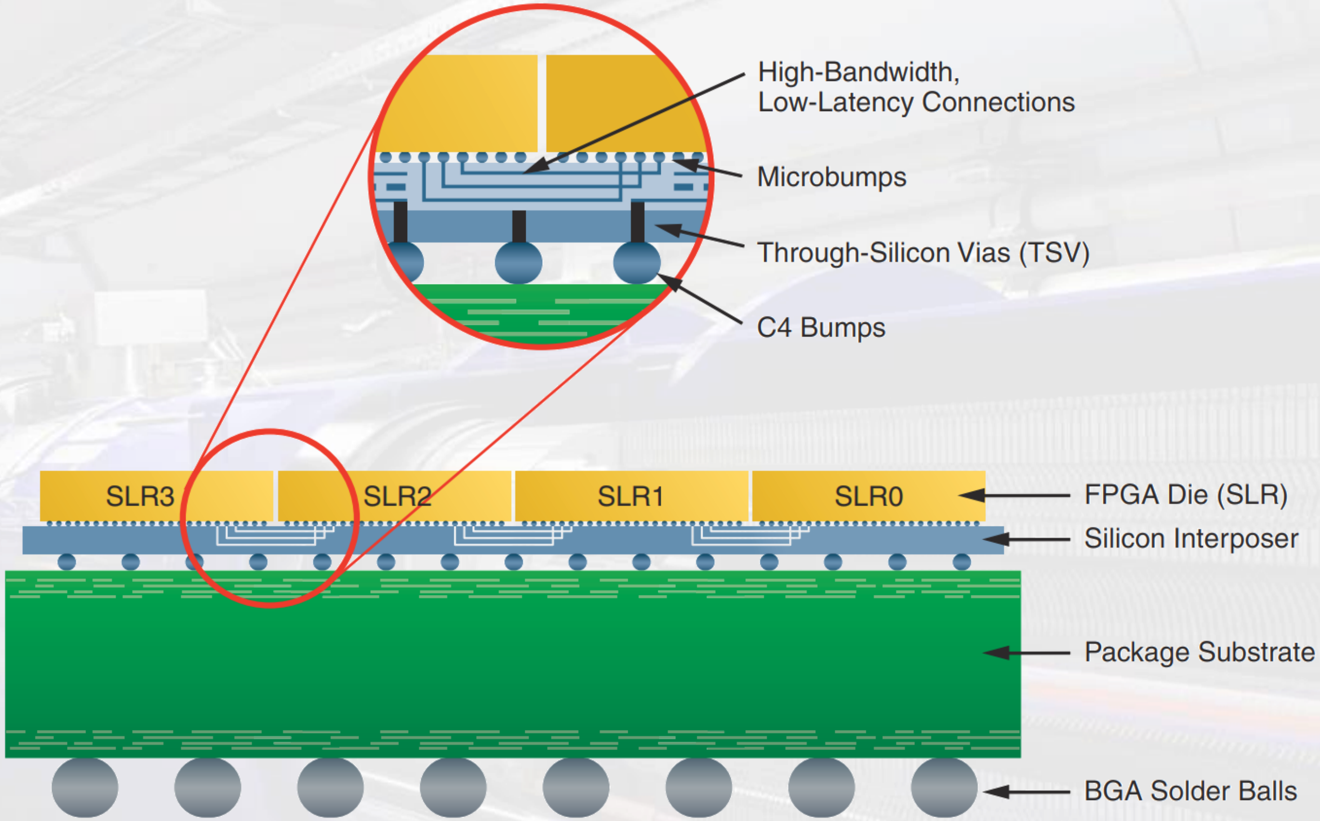


Fig 3: Xilinx FPGA Enabled by SSI Technology*

*: *UG872 Large FPGA Methodology Guide*

Barrel Calorimeter Segmentation (New)

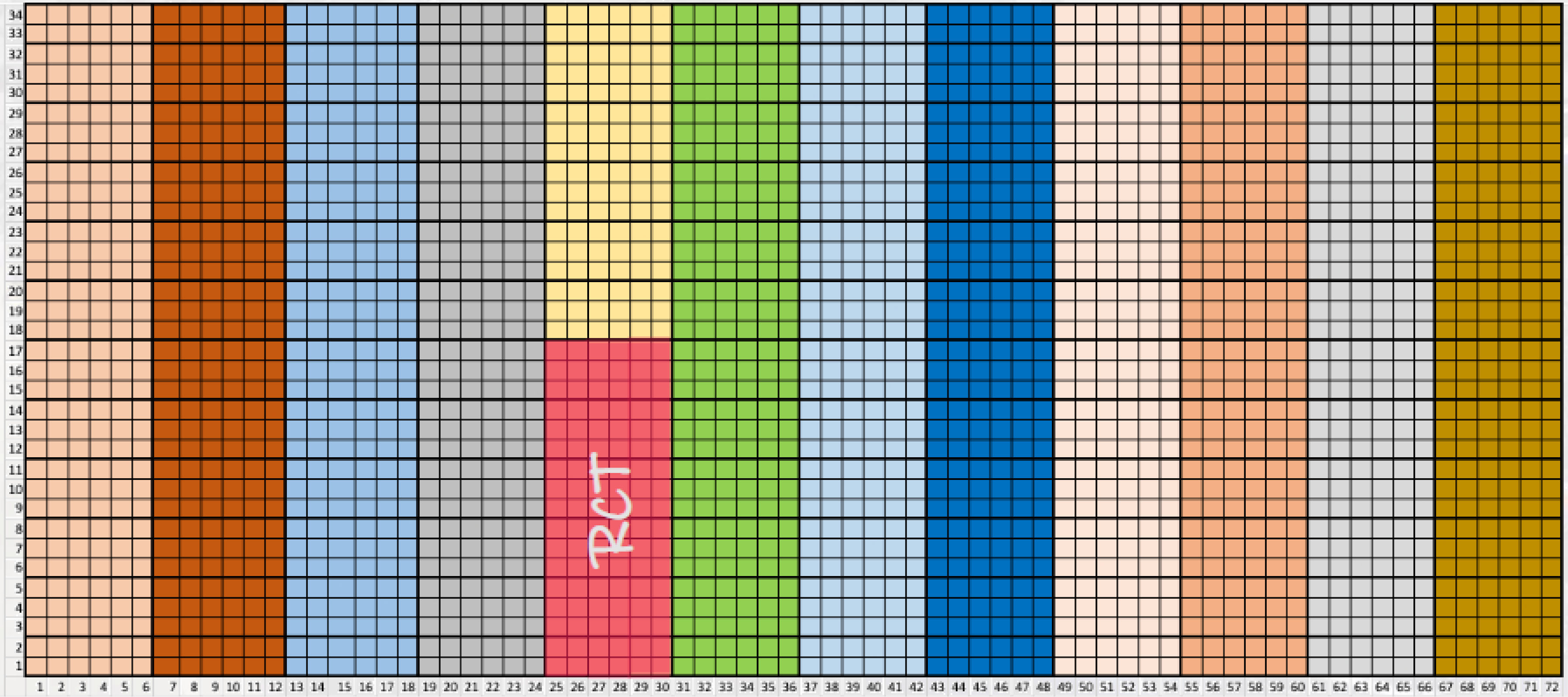


Fig 2: Barrel calorimeter segmentation (new)

Physics bandwidth vs Algo clock @ 16G

LHC BC Clock [MHz]	40.08
Word Bit Size	66
Line Rate [Gbps]	**16.00000**
Max Theoretical Words/Bx	6.04851

	TM1			TM6			TM18		
Bx Frame Length (TM interval)	**1**	**1**	**1**	**6**	**6**	**6**	**18**	**18**	**18**
Words/Frame	**4**	**5**	**6**	**24**	**30**	**36**	**72**	**90**	**108**
Equiv. Words/Bx	4.00	5.00	6.00	4.00	5.00	6.00	4.00	5.00	6.00
Equiv. Bits/Bx	256	320	384	256	320	384	256	320	384
Data Rate [Gbps]	10.58	13.23	15.87	10.58	13.23	15.87	10.58	13.23	15.87
Filler Rate [Gbps]	5.42	2.77	0.13	5.42	2.77	0.13	5.42	2.77	0.13
Average Filler Words/Bx	2.05	1.05	0.05	2.05	1.05	0.05	2.05	1.05	0.05
Average Filler Words/Orbit	7300.89	3736.89	172.89	7300.89	3736.89	172.89	7300.89	3736.89	172.89
Average Filler Words/Frame	2.05	1.05	0.05	12.29	6.29	0.29	36.87	18.87	0.87
Payload Bits/Frame	**256**	**320**	**384**	**1536**	**1920**	**2304**	**4608**	**5760**	**6912**
Algo Clock @ 64b i/f [MHz]	**160.32**	**200.4**	**240.48**	**160.32**	**200.4**	**240.48**	**160.32**	**200.4**	**240.48**

Physics bandwidth vs Algo clock @ 25G

LHC BC Clock [MHz]	40.08
Word Bit Size	66
Line Rate [Gbps]	**25.78125**
Max Theoretical Words/Bx	9.74613

	TM1			TM6			TM18		
Bx Frame Length (TM interval)	**1**	**1**	**1**	**6**	**6**	**6**	**18**	**18**	**18**
Words/Frame	**7**	**8**	**9**	**42**	**48**	**54**	**126**	**144**	**162**
Equiv. Words/Bx	7.00	8.00	9.00	7.00	8.00	9.00	7.00	8.00	9.00
Equiv. Bits/Bx	448	512	576	448	512	576	448	512	576
Data Rate [Gbps]	18.52	21.16	23.81	18.52	21.16	23.81	18.52	21.16	23.81
Filler Rate [Gbps]	7.26	4.62	1.97	7.26	4.62	1.97	7.26	4.62	1.97
Average Filler Words/Bx	2.75	1.75	0.75	2.75	1.75	0.75	2.75	1.75	0.75
Average Filler Words/Orbit	9787.22	6223.22	2659.22	9787.22	6223.22	2659.22	9787.22	6223.22	2659.22
Average Filler Words/Frame	2.75	1.75	0.75	16.48	10.48	4.48	49.43	31.43	13.43
Payload Bits/Frame	**448**	**512**	**576**	**2688**	**3072**	**3456**	**8064**	**9216**	**10368**
Algo Clock @ 64b i/f [MHz]	**280.56**	**320.64**	**360.72**	**280.56**	**320.64**	**360.72**	**280.56**	**320.64**	**360.72**

Project hierarchy and floor planning

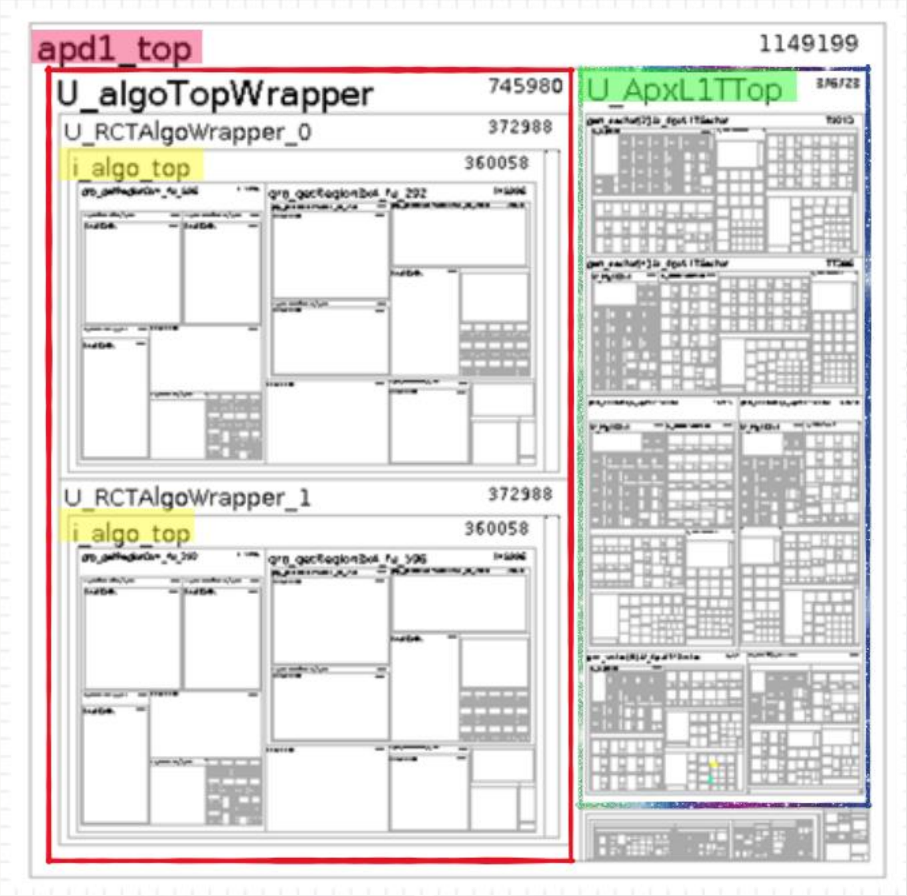


Fig 22: Project hierarchy in Vivado

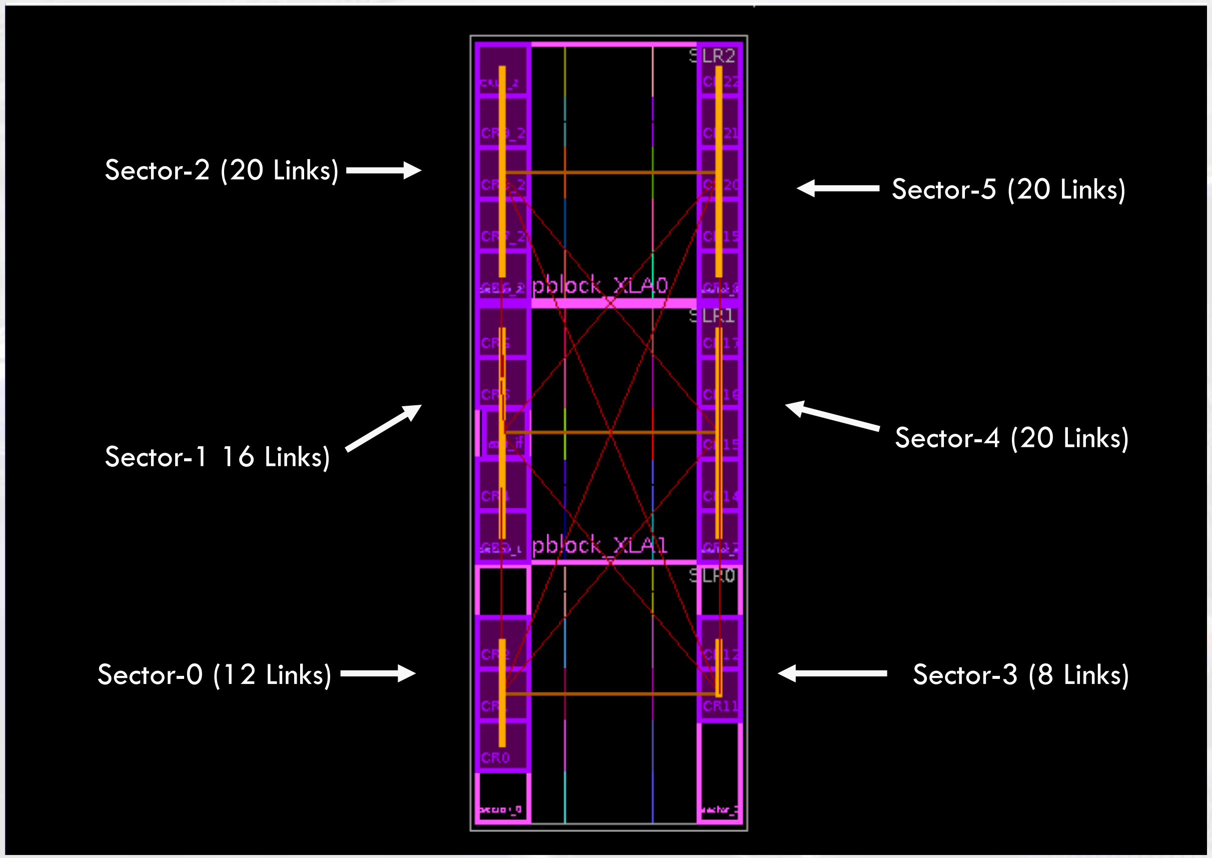
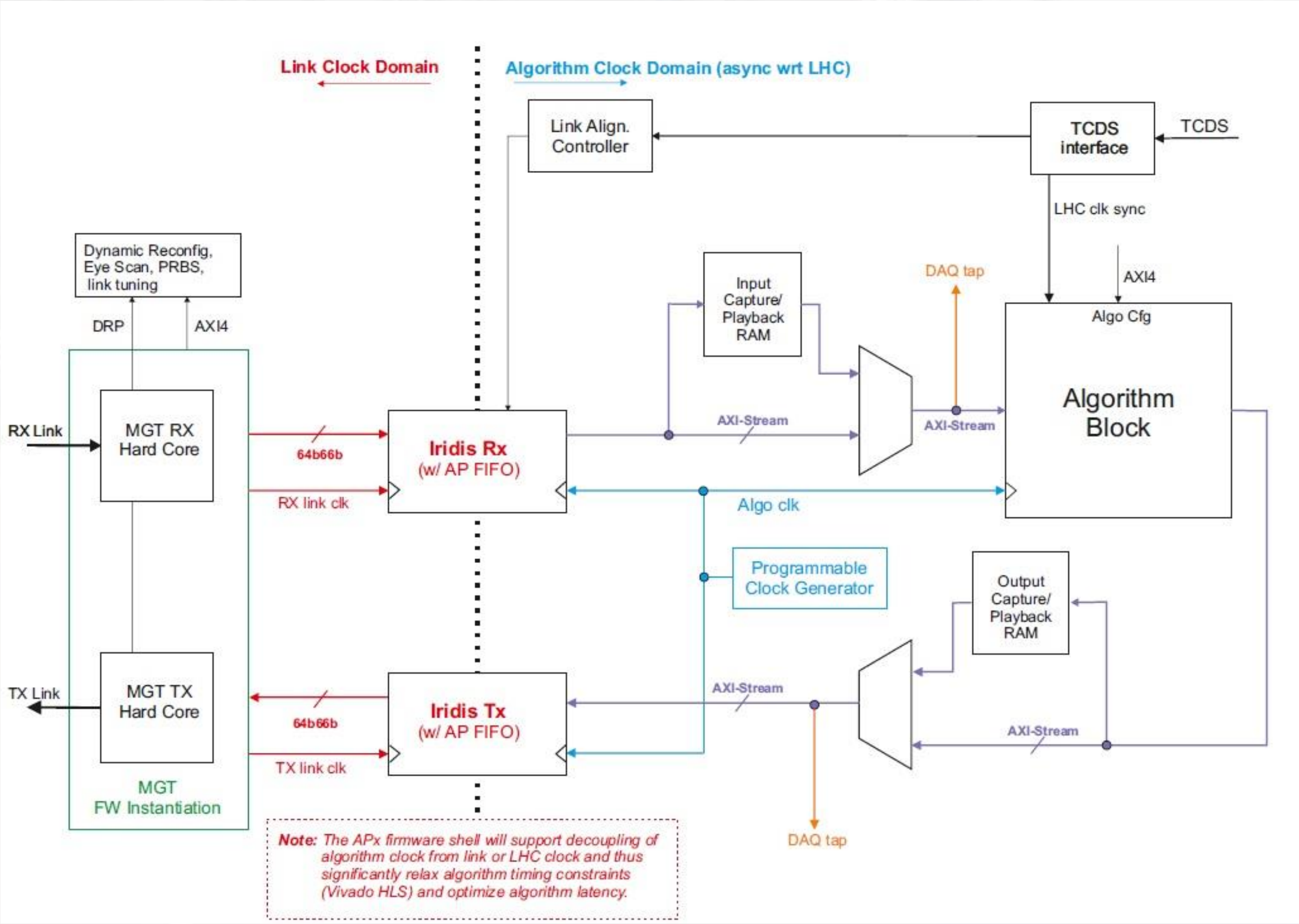


Fig 23: Project floor planning

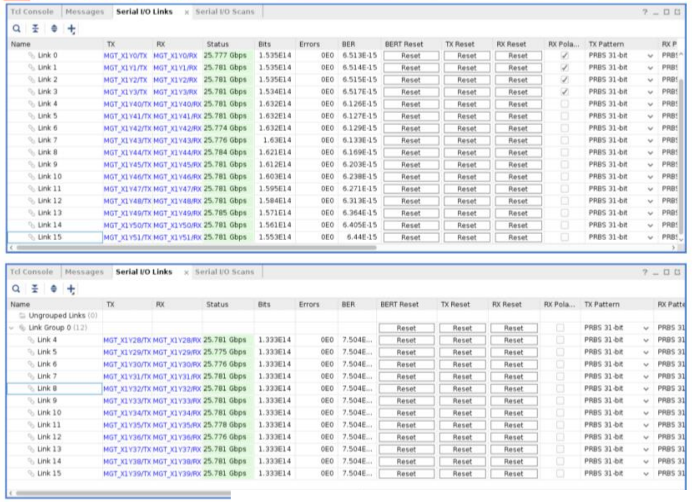
APx Firmware shell

Iridis – 64b66b-based optimized signaling method and firmware cores for CMS Trigger applications



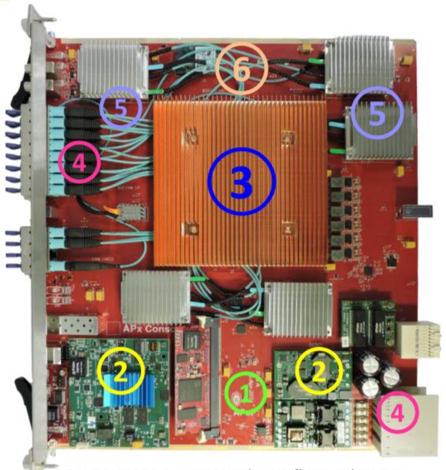
APx test

25G Endurance Tests



- Using both Firefly 25X12 Alpha module sets
- 515.625 MHz refclk frequency (zero rem.)
- All 124 paths tested to ≥ 1E14 bits of PRBS31 data with zero errors
- Some tweaking of fiber connections necessary for 25X12 modules

ATCA Cooling, APxF Example



Designing to maximize slot airflow utilization

1. Low restriction airflow path to FPGA heat sink
2. Low-profile, low load flyover zone
3. VU13P FPGA Heat Sink 12.5×12.5 cm, 16% fill fin pattern Measured 3.4 W/°C (0.29 °C/W) at full 450 Watt fan power (lidded A2577 package)
4. Significant airflow obstructions
5. Optical module heat sinks located for pressure balance
6. FPGA exhaust heat zone

THERMAL PERFORMANCE (APX)

- At 16 W/cm a 12.5cm heatsink provides 200W of cooling potential - assuming no significant ducting of air within the card
- APxF has 3.4 W/C heat sink performance so 200W load will increase temperature by 59 degrees (25C to 84C) with cooling at full power
- Observations:
  - 200W FPGA power limit feasible at full fan power
  - Care required when balancing design tradeoffs - e.g. heat sink dimensions vs MGT route length
  - Lower die temperature -> capacity to reduce fan speed

Sector-0 (12

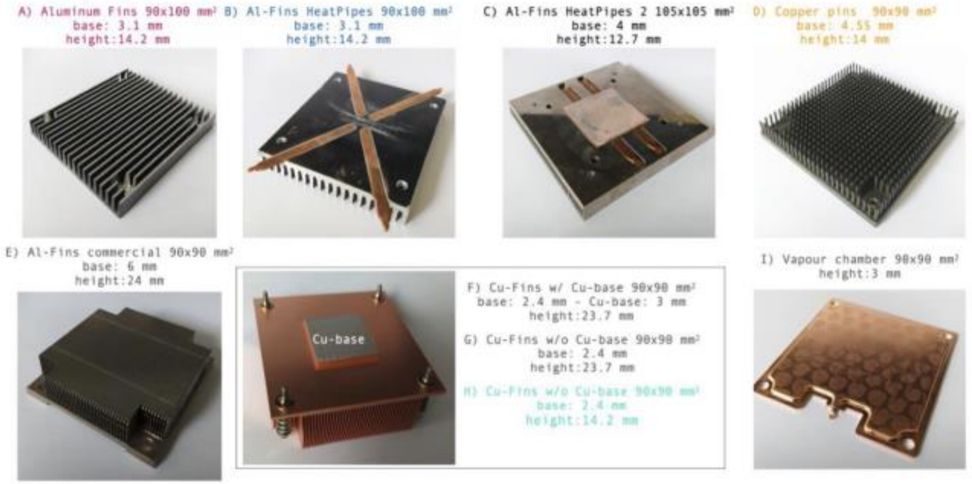
VU13P Lidless Package Option

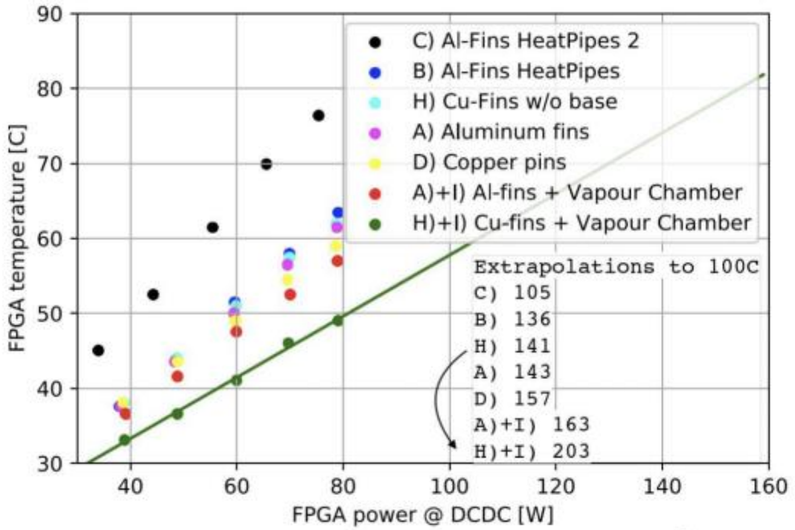
- Xilinx Data for A2577 Θ_{JC} (die to case):
  FLGA (lidded): 0.05 °C/W
  FSGA (lidless): 0.01 °C/W
- At 200W, up to ΔT ≈ 8 °C savings versus the lidded package
- Comments:
  - Lidless interface a more exacting design—APx has a lidless heat sink design on file
  - Would optimize other thermal design aspects first (board layout, heat sink geometries)
  - When a device is operating near the thermal limit, small °C improvements → a large % increase in thermal margin

Serenity tests

THERMAL PERFORMANCE (SERENITY)

- Explored using heat pipes and vapour chambers to allow "small" heatsinks
- Vapour chambers allow 200W dissipation with expected fan speed 10 of 15 with 90mm x 90 mm heatsink (i.e. compact)





- See large variation depending on heatsink
- Want to keep FPGA temperature at 100 degrees or lower