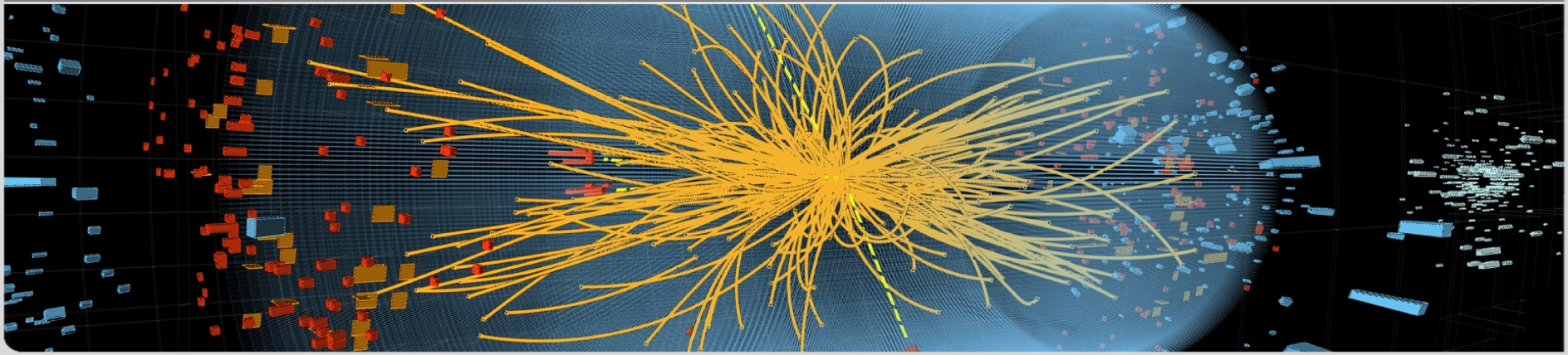
Karlsruher Institut für Technologie

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

FPGA-based real-time track reconstruction for the CMS phase-2 tracker upgrade

Luis Ardila

INSTITUTE FOR DATA PROCESSING AND ELECTRONICS (IPE)



Outlook

Introduction: High Luminosity Large Hadron Collider (HL-LHC)

- CMS Tracker Upgrade
- Track Finder Architecture

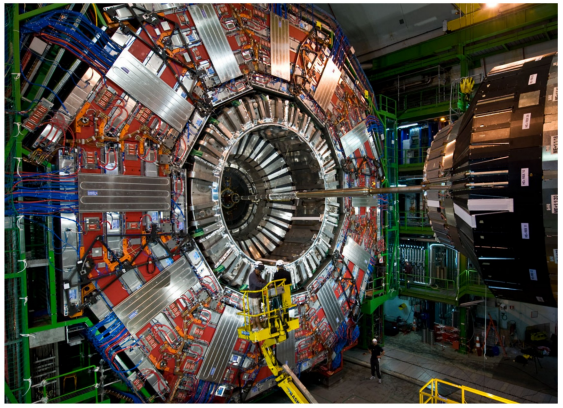
Track Finding Algorithms

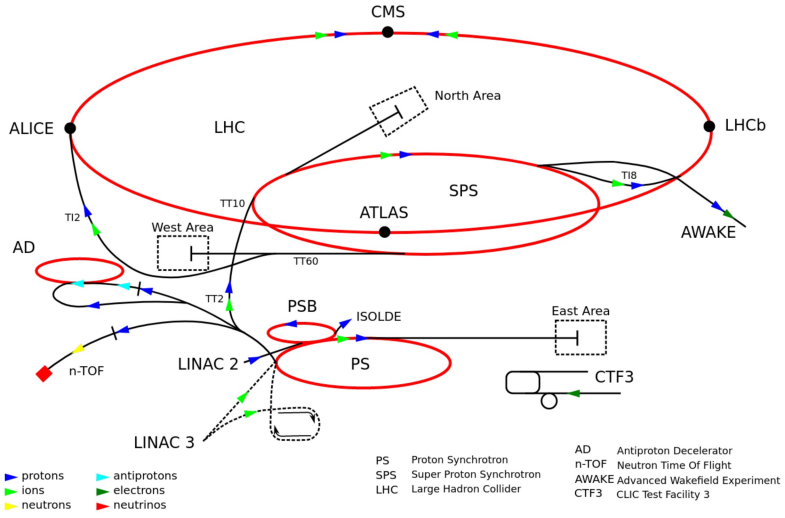
- Tracklet
- Time-Multiplexed Track Trigger (TMTT)
- Hybrid

Hardware R&D

- Hardware Prototyping Platforms
- Integrated Board Management Module

Summary





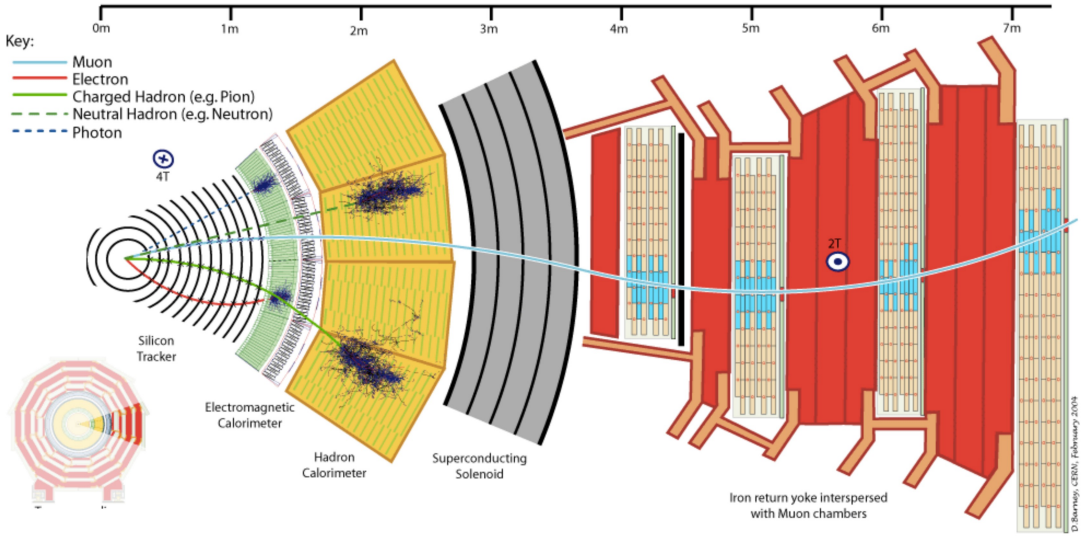
High Luminosity LHC – CMS

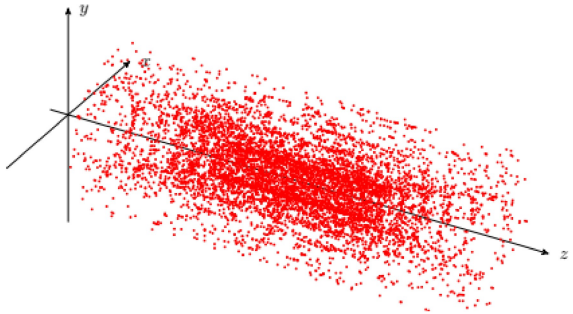
By **2026** the LHC will be upgraded in luminosity 5-7 x 10^{34} / cm^2 / s

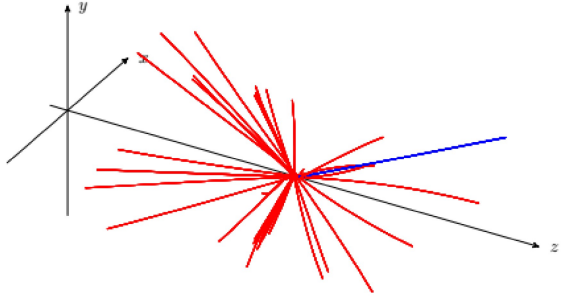
Silicon strip **tracker** will be **replaced**

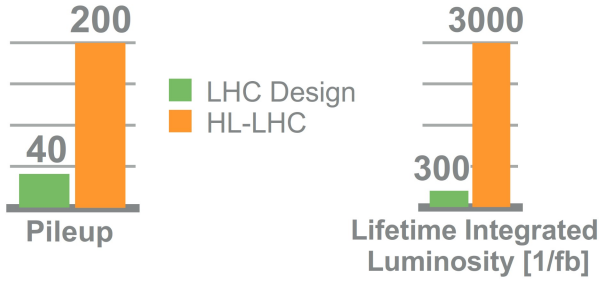
Challenging high occupancy conditions. ~10,000 charged particles per bunch crossing

Necessary to **include tracking** information at **first level** of **triggering**





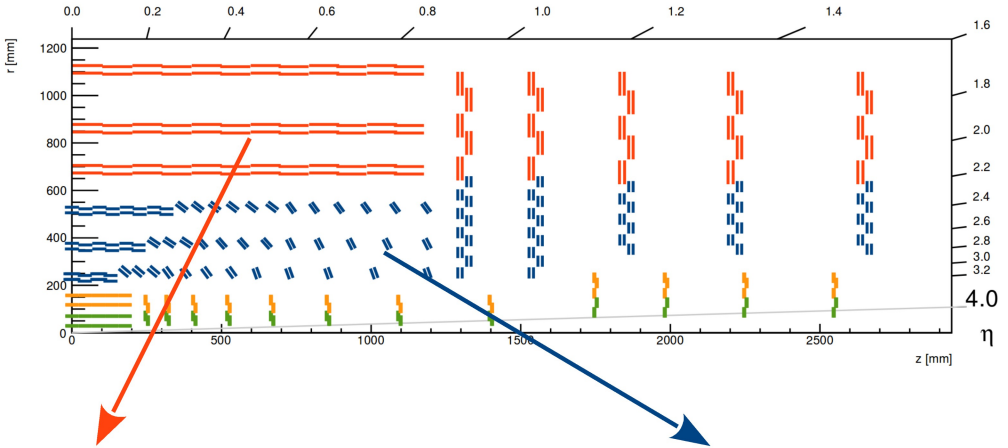


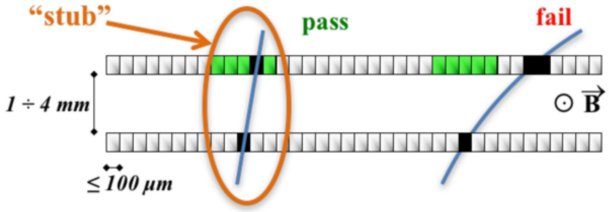


CMS Tracker Upgrade

p_T discrimination provided by use of special modules

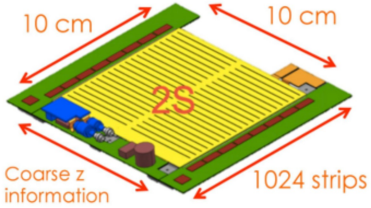
- Pairs of closely spaced silicon sensors, separated 1.6 - 4 mm
- Signals from each sensor are correlated
- Only hit pairs compatible with **p_T > 2 – 3 GeV/c ("Stubs") are forwarded off-detector**
- Factor ~10 data reduction ~15,000 stubs per bunch crossing



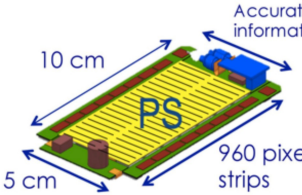


"2S" 2 Strip Modules r > 60 cm

Strip Sensor × 2:
5 cm × 90 μm
+
Strip Sensor × 2:
5 cm × 90 μm

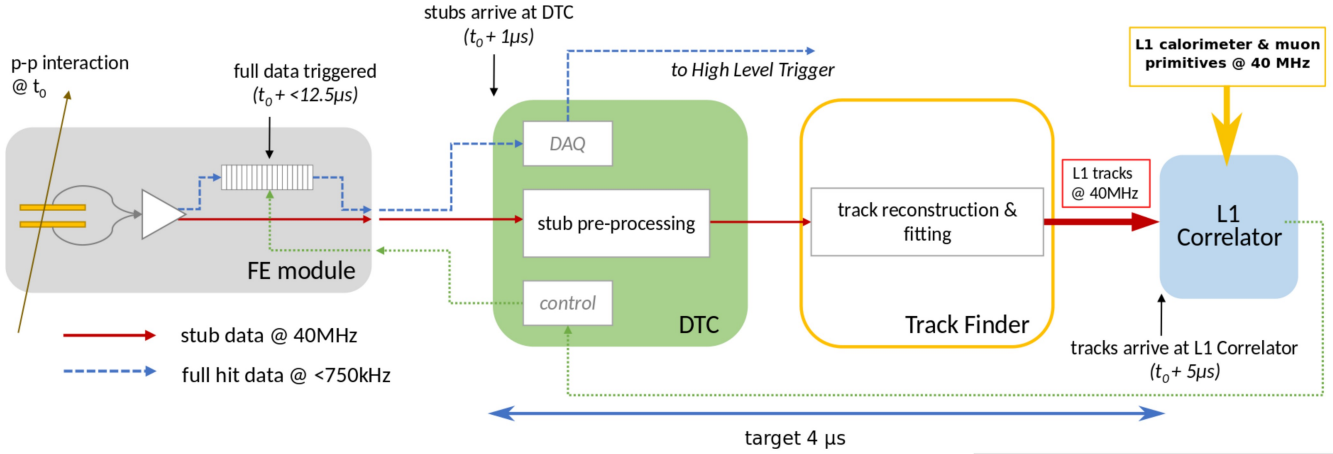


PS" Pixel + Strip Modules 20 < r < 60 cm



Strip Sensor × 2:
2. 5 cm × 100 μm
+
Pixel Sensor × 32:
1.5 mm × 100 μm

Tracker → Trigger Data Flow



Average 15,000 stubs every 25ns (200PU) → **Stub bandwidth O(20) Tb/s**

L1 hardware trigger reduces event rate from **40 MHz to < 750 kHz** using calorimeter, muon and tracker primitives

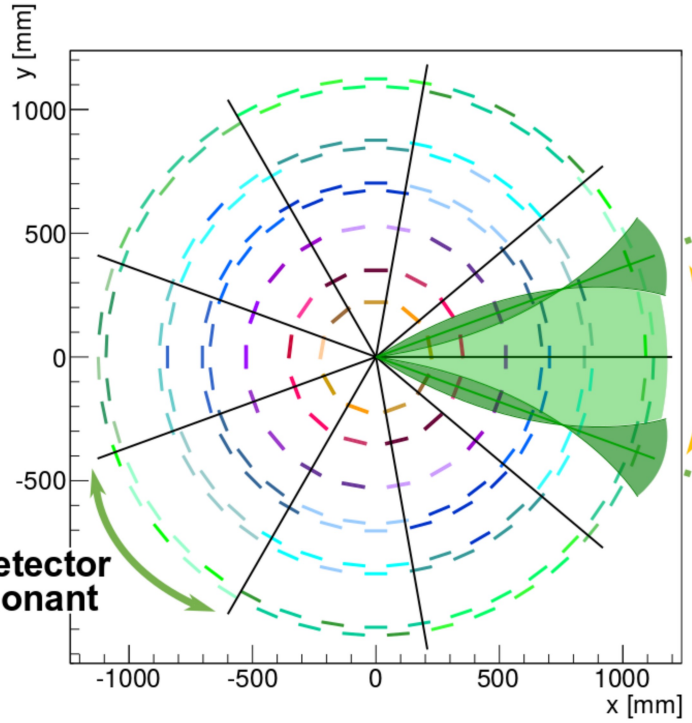
- **Tracker primitives are all tracks (pT > 2-3 GeV/c) from Outer Tracker**
- L1-Accept triggers all front-end buffers to read out to DAQ → HLT farm

FE L1 latency buffers limited to 12.5 μs

Transmission of stubs to back-end electronics	1 μs
Correlation of trigger primitives (inc. tracks)	3.5 μs
Broadcast of L1-Accept to front-end buffers	1 μs
Safety Margin	3 μs

→ Track finding from stubs must be performed in 4 μs

Track Finder Architecture



Detector nonant 1:
z+, z- (24 DTCs)

DTC

Processor nonant

DTC

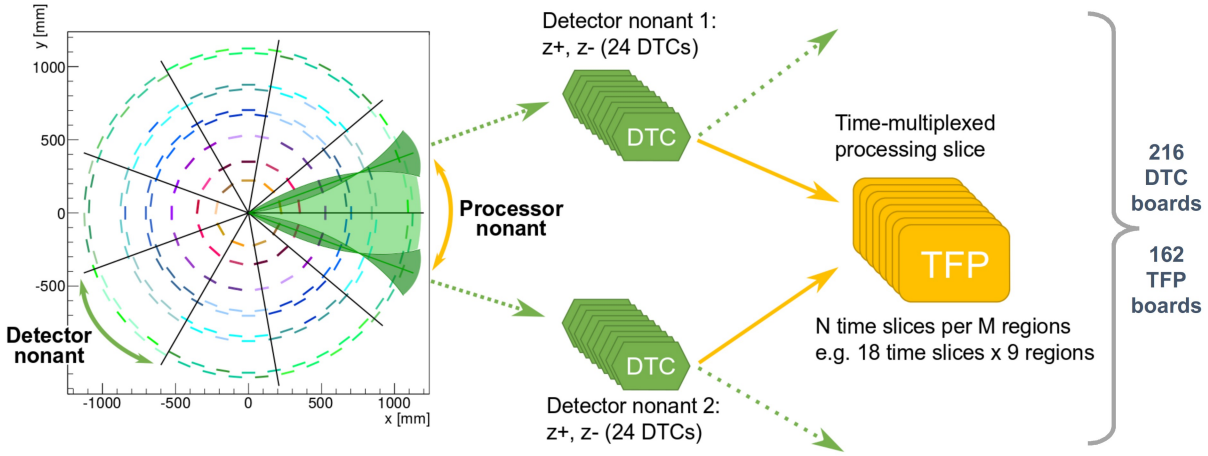
Detector nonant 2:
z+, z- (24 DTCs)

Time-multiplexed processing slice

TFP

N time slices per M regions
e.g. 18 time slices x 9 regions

Track Finder Architecture



Two stages of data processing

- **DAQ, Trigger and Control (DTC) layer**
- **Track Finding Processor (TFP) layer**
- **All-FPGA** processing system
- **ATCA** form factor; CMS standard dual-star backplane

Outer Tracker cabled into **nonants**

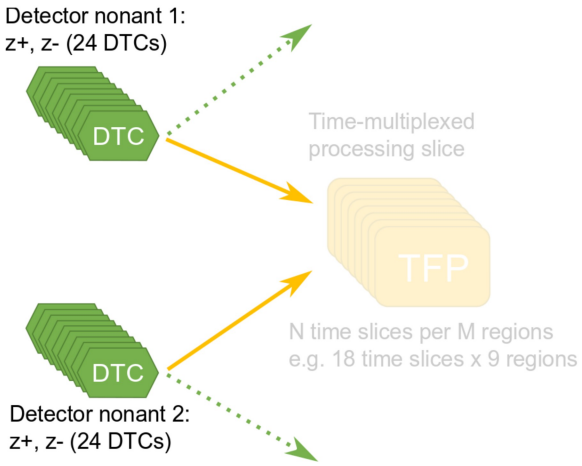
Use of time-multiplexing to increase parallelization

- Time-multiplexing directs data from multiple sources to a single processing node
- **1 event per processing node**

Processors are independent entities → simplifies commissioning and operation

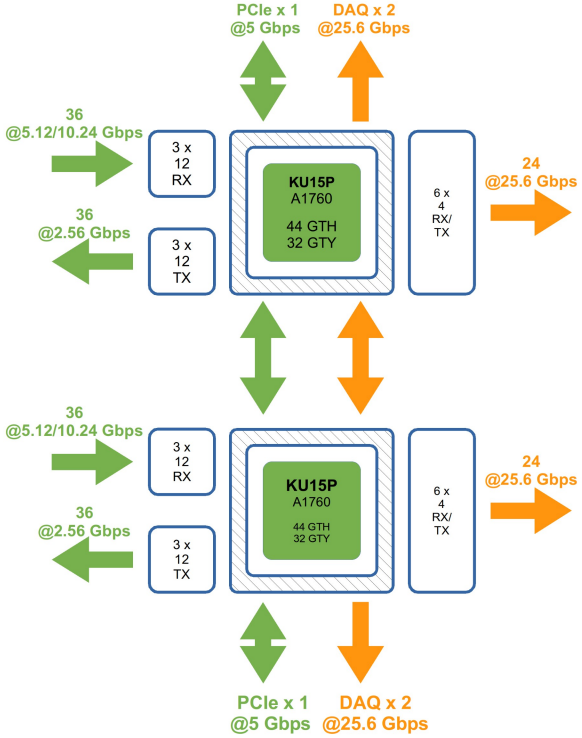
Spare nodes available for redundancy

Track Finder Architecture – DTC



Two stages of data processing

- **DAQ, Trigger and Control (DTC) layer**
- **Track Finding Processor (TFP) layer**
- **All-FPGA** processing system
- **ATCA** form factor; CMS standard dual-star backplane



→ **216 DTC boards, 18 crates, 1 rack/nonant**

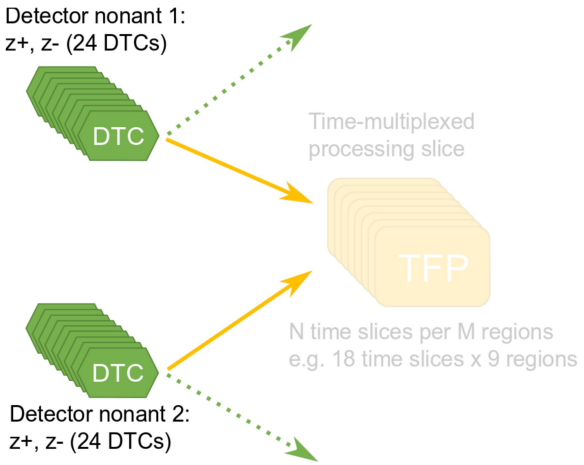
DTC card must handle

- **<=72 modules (5G/10G lpGBT opto-links)**
- Control/Readout for each module
- Direct L1 stream to central DAQ (16G/25G)
- Direct stub stream to TFPs (25G)

Stub pre-processing includes:

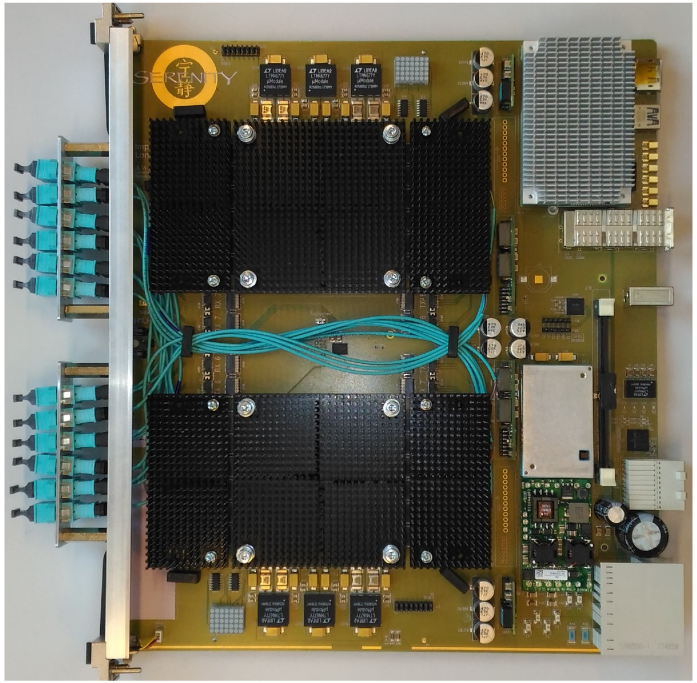
- **Local→ Global** look up, position calibration
- Sort and pre-duplication
- Time-multiplexing

Track Finder Architecture – DTC



Two stages of data processing

- **DAQ, Trigger and Control (DTC) layer**
- **Track Finding Processor (TFP) layer**
- **All-FPGA** processing system
- **ATCA** form factor; CMS standard dual-star backplane



DTC card must handle

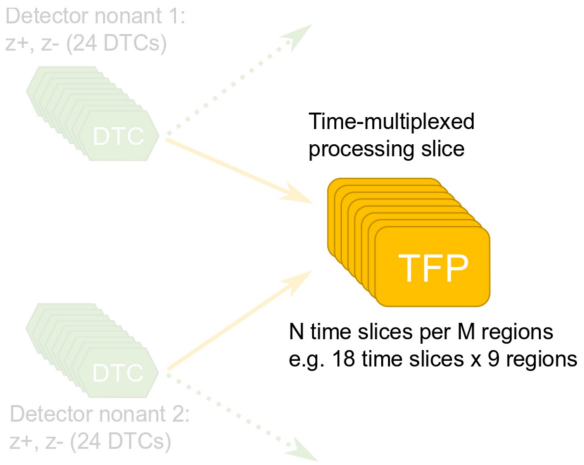
- **<=72 modules (5G/10G lpGBT opto-links)**
- Control/Readout for each module
- Direct L1 stream to central DAQ (16G/25G)
- Direct stub stream to TFPs (25G)

Stub pre-processing includes:

- **Local→ Global** look up, position calibration
- Sort and pre-duplication
- Time-multiplexing

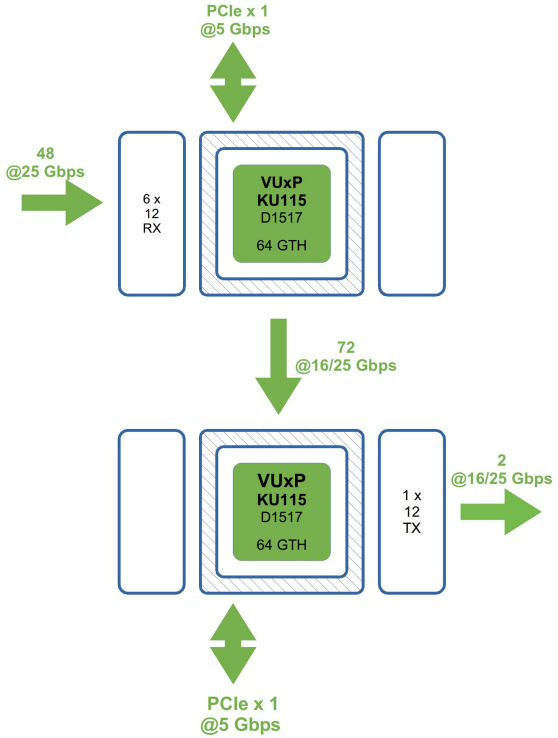
→ **216 DTC boards, 18 crates, 1 rack/nonant**

Track Finder Architecture – TFP



Two stages of data processing

- **DAQ, Trigger and Control (DTC) layer**
- **Track Finding Processor (TFP) layer**
- **All-FPGA** processing system
- **ATCA** form factor; CMS standard dual-star backplane



TFP card must handle

- **Up to 48 DTCs (25G optical links)**
- Track Finding from stubs
- Track Fitting
- Transmission to L1 Correlator Trigger

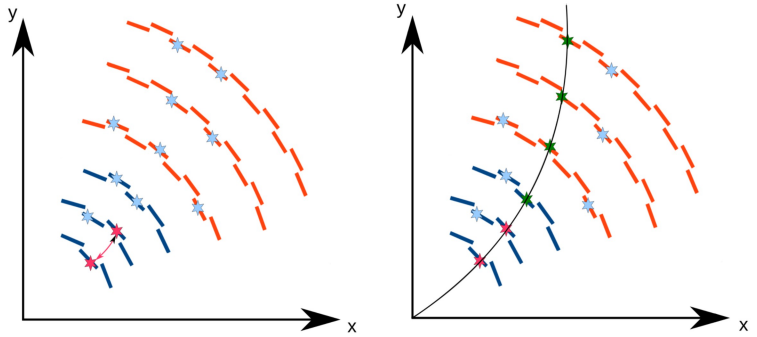
High bandwidth processing card

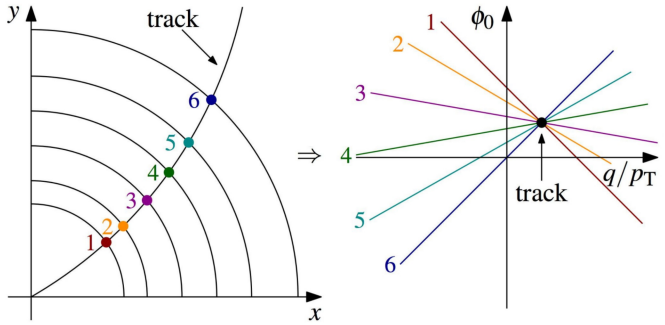
- **~1 Tb/s** processing bandwidth
- Rate to L1 Correlator much lower < 30 Gb/s

→ **162 TF boards, 18 crates (one per time-node)**

Track Finding Algorithms

Two main algorithms for reconstructing tracks, plus a number of hybrids, variation and options





TRACKLET + CHI2 FIT APPROACH

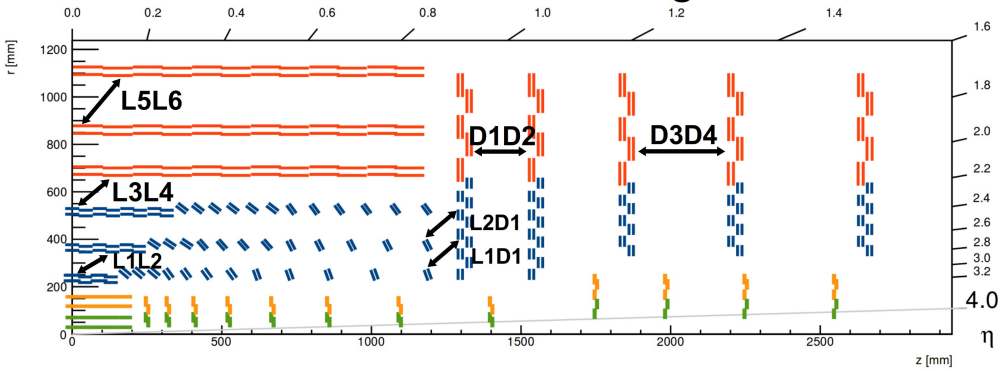
- Combinatorial approach using **pairs of stubs as seeds**
- **Extrapolation** to other layers → hit matching
- **Linearized χ2** fit on candidates
- Uses **full resolution stubs** at earliest stage of processing
- N time-slices x M regions → 6 x 24, 9 x 18

HOUGH TRANSFORM + KALMAN FILTER APPROACH

- Uses a **Hough Transform** to detect coarse candidates
- Candidates are filtered and fitted in a single subsequent step using a **Kalman Filter**
- Combinatorial problem pushed to latter stages of processing
- N time-slices x M regions → 18 x 9

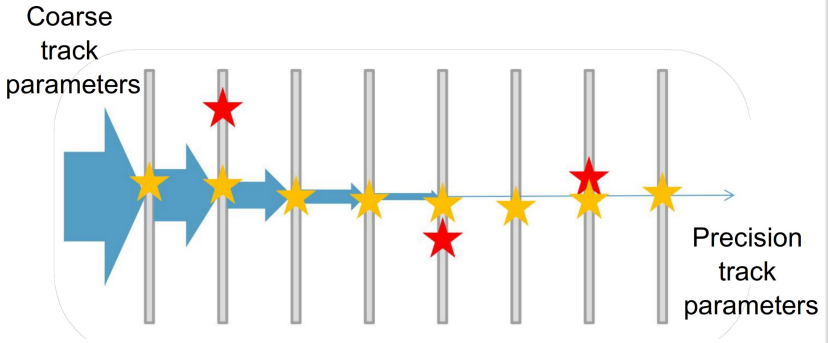
Hybrid Algorithm

Tracklet Seeding



+

Kalman Filter



1. Tracklet Seeds formed in multiple layer combinations

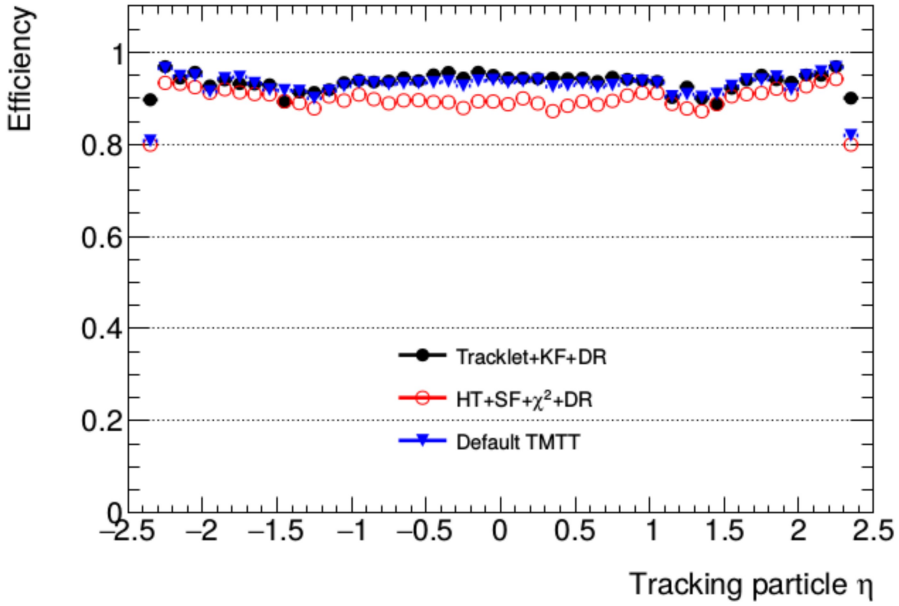
2. Projection and match in other layers

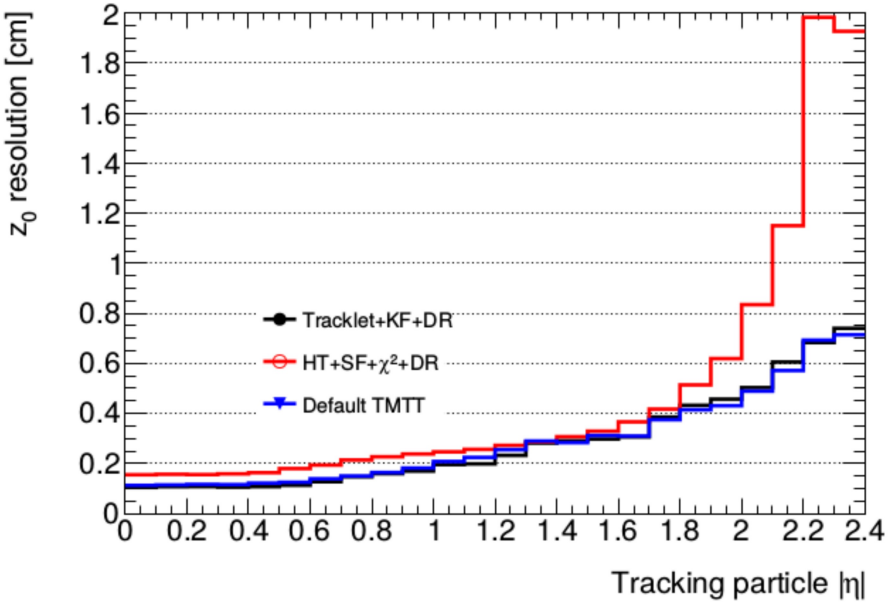
3. Candidates sharing stubs are merged

4. Kalman Filter fits track and selects best stubs candidates

- 9 Φ sectors
- No η division
- Time-Multiplex Period of 18

Hybrid Comparison



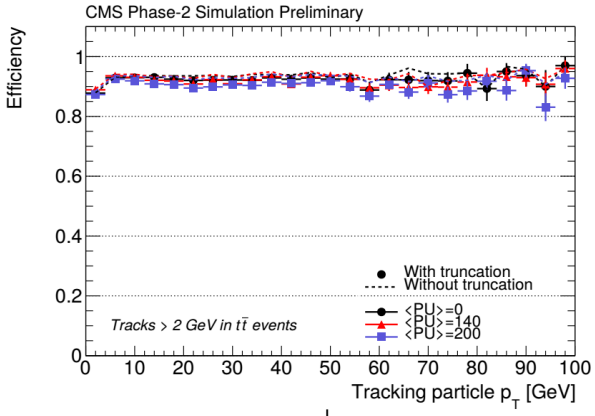


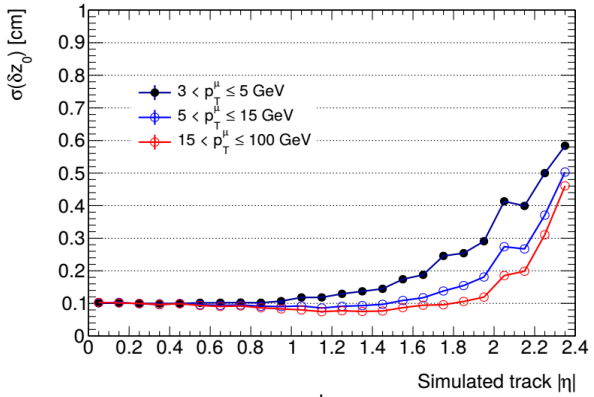
Efforts have started to **merge** the two approaches

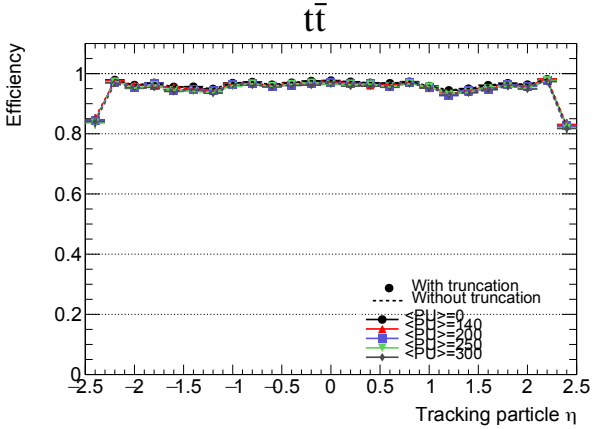
- Working on defining a **reference algorithm**

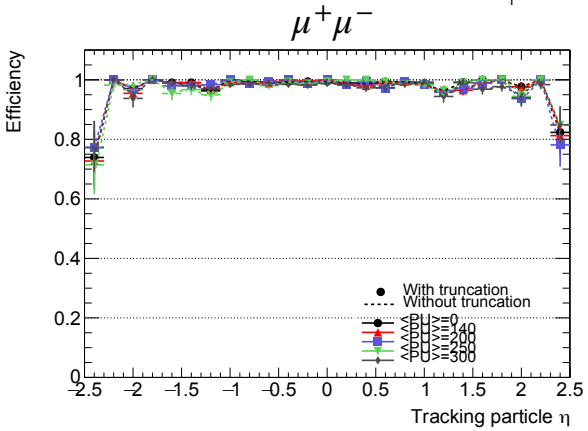
Hybrid Performance

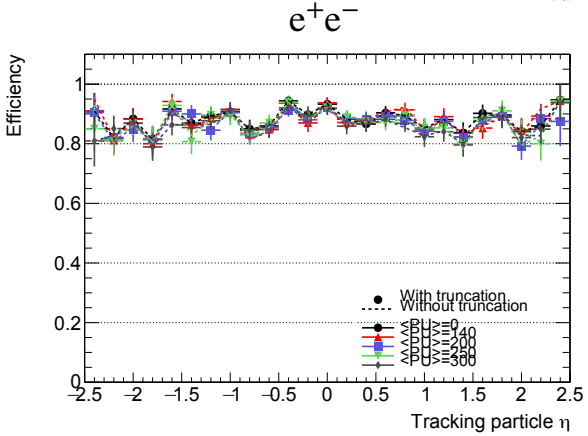
- Average track finding **efficiency** for $t\bar{t}$ tracks **> 95% (> 3 GeV)**
- **z_0** resolution **~1 mm** (barrel)
- **p_T** resolution **~1%** (barrel)
- Per event average **~60 tracks (3 GeV) ~200 (2GeV)** ($t\bar{t}$ at 200 PU)





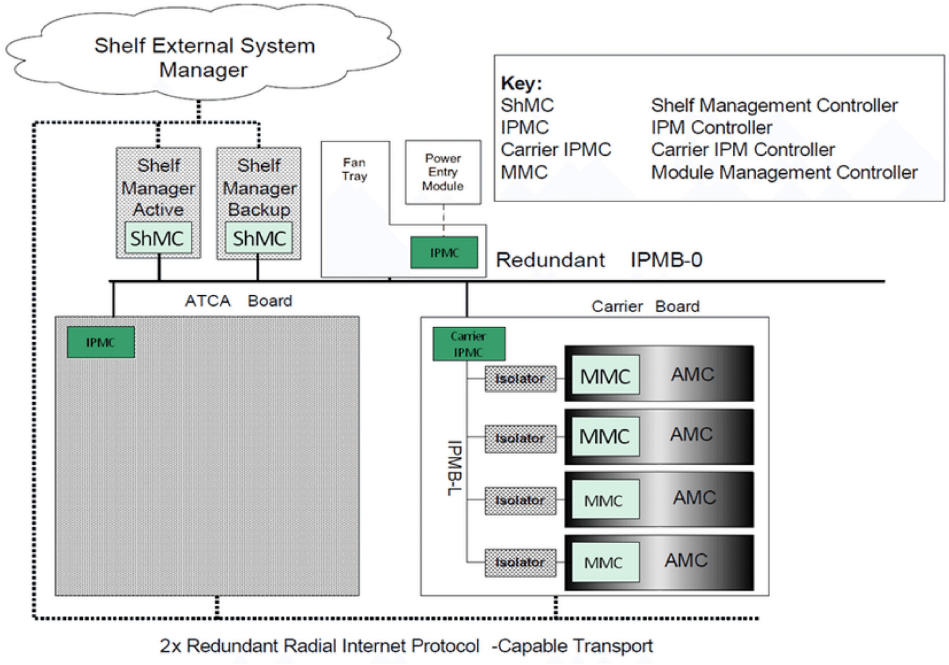






Hardware R&D: ATCA

Advanced Telecommunications Computing Architecture (ATCA)



280 x 322 mm board size

All CMS Phase-2 back-end electronics will be ATCA-based

- Dual-star backplane

Standardized the use of the backplane for clocks, and timing and throttling signals

- LHC bunch-crossing clock(40.08MHz)
- Precision crossing clock(320.64MHz)
- TTC2 trigger and fast-control stream (from DTH to back-ends)
- TTS2 throttling stream (from back-ends to DTH)

HW - R&D

Bristol University, Imperial College, Ioannina, INFN, KIT, RAL, SACLAY, TIFR

ATCA infrastructure

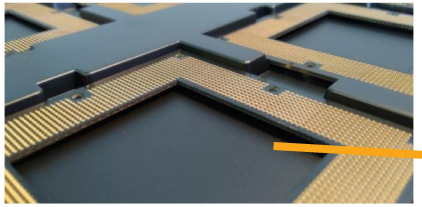
- Systematic **thermal studies** about air cross-section and impact on opto-lifetime
- Backplane signal integrity → important for DAQ/timing

Use of **interposer technology**

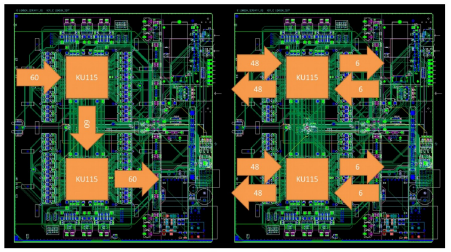
- Flexibility (e.g. FPGA)
- Mitigate losses/costs due to yield issues
- **Modularity**; separate complex and simpler part of the board design

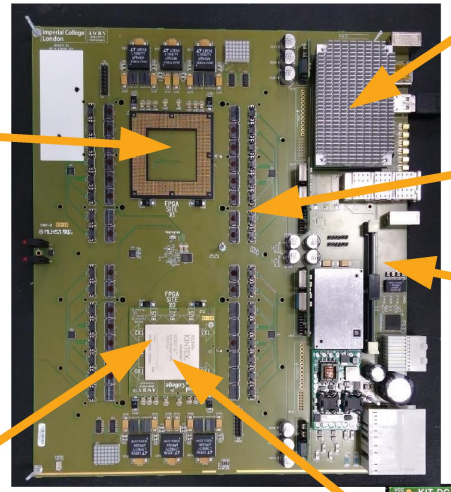
On-board computing and control variety

- Standard on-board PC (COM Express mini)
- ZynqUS+ SoC
- Intelligent Platform Management Controller (IPMC)



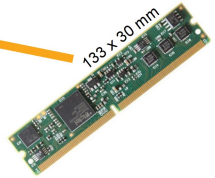
Samtec Z-RAY interposer





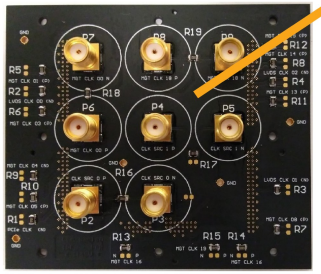
COM Express

Samtec Firefly x12 RX/TX pairs



CERN-IPMC

Clock test daughtercard



FPGAs
KU115
KU15P
VU9P
daughtercards

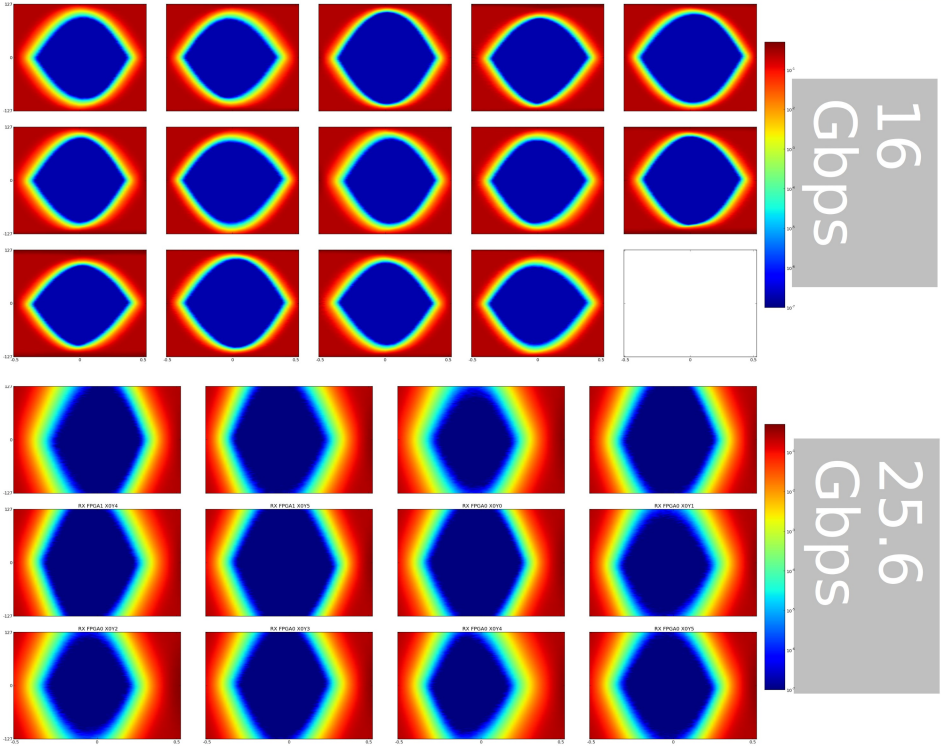
HW - R&D

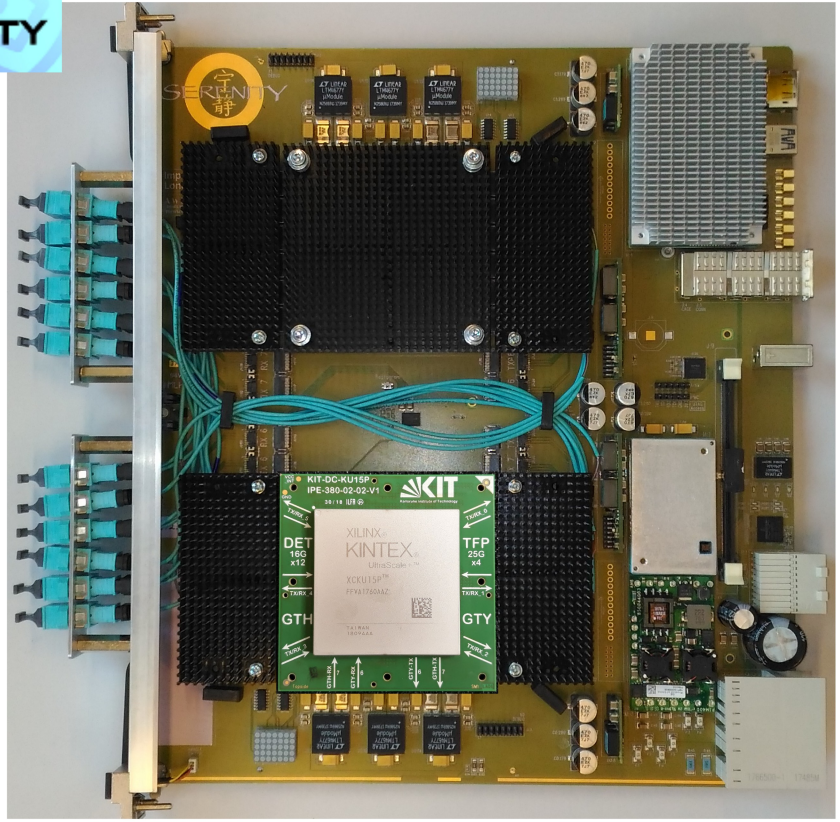
Bristol University, Imperial College, Ioannina, INFN, KIT, RAL, SACLAY, TIFR

SERENITY

DUAL KU15P

Karlsruher Institut für Technologie





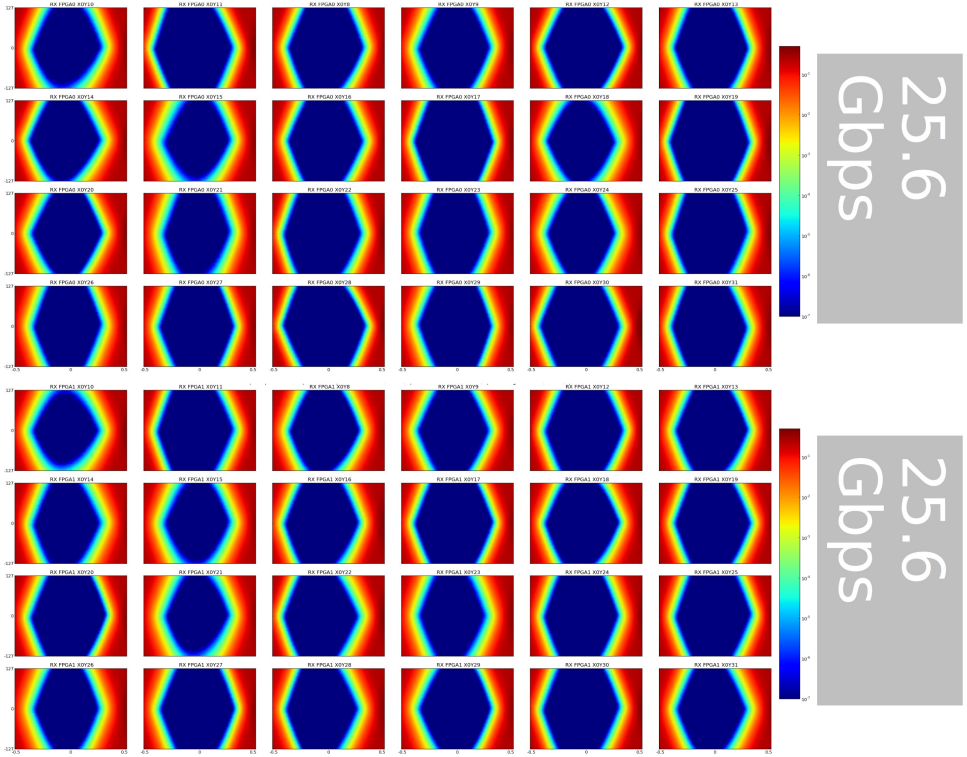
HW - R&D

Bristol University, Imperial College, Ioannina, INFN, KIT, RAL, SACLAY, TIFR

SERENITY

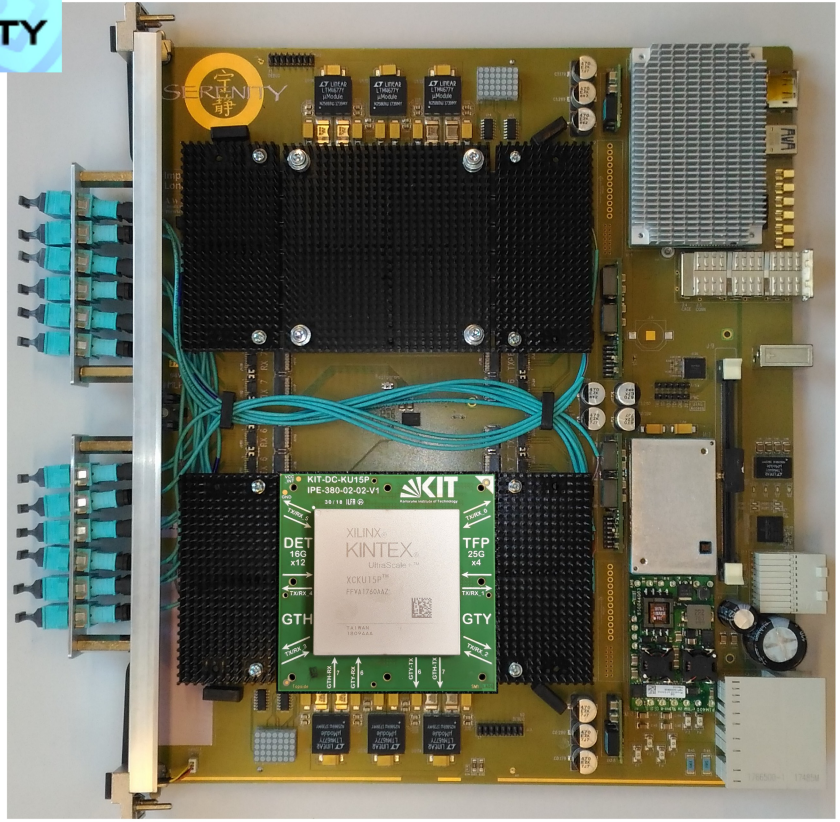
2x DUAL KU15P

Karlsruher Institut für Technologie



25.6 Gbps

25.6 Gbps



HW - R&D

APOLLO uses coplanar PCBs with Back-Plane Connectors in between

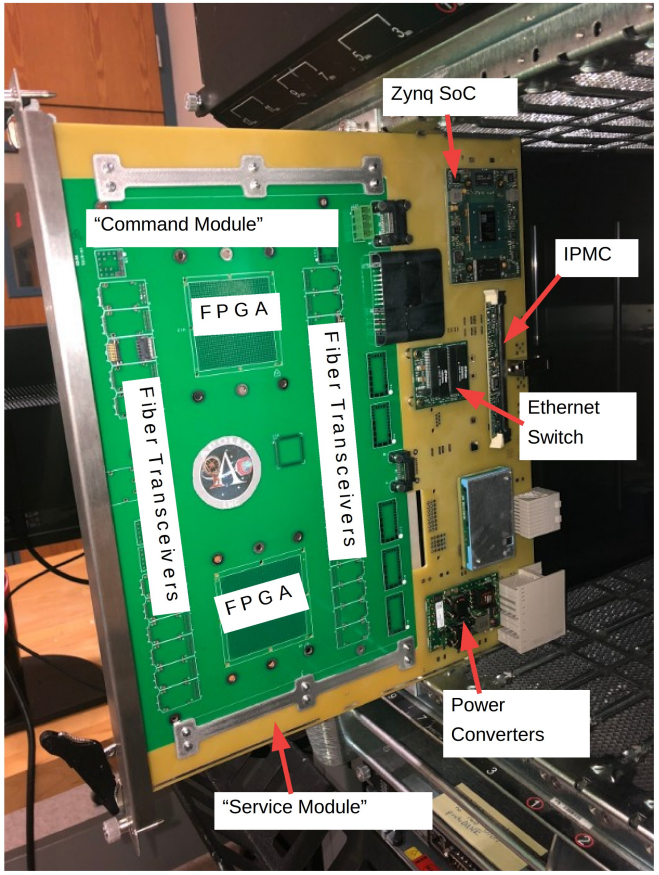
- Flexibility (e.g. FPGA+Optics)
- Modularity; separate complex and simpler part of the board design

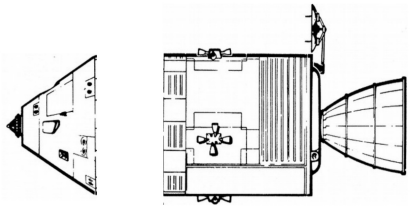
On-board computing and control variety

- Zynq SoC
- CERN-IPMC or UW-IPMC

PCB Characteristics:

- 16 layers / Megtron-6 / 1.8 mm
- Apollo analogy: Split into “Command” and “Service” modules





Boston University, Cornell University, Rutgers University, Ohio State University, University of Notre Dame, Northwestern University, University of Colorado

HW - R&D

APOLLO uses coplanar PCBs with Back-Plane Connectors in between

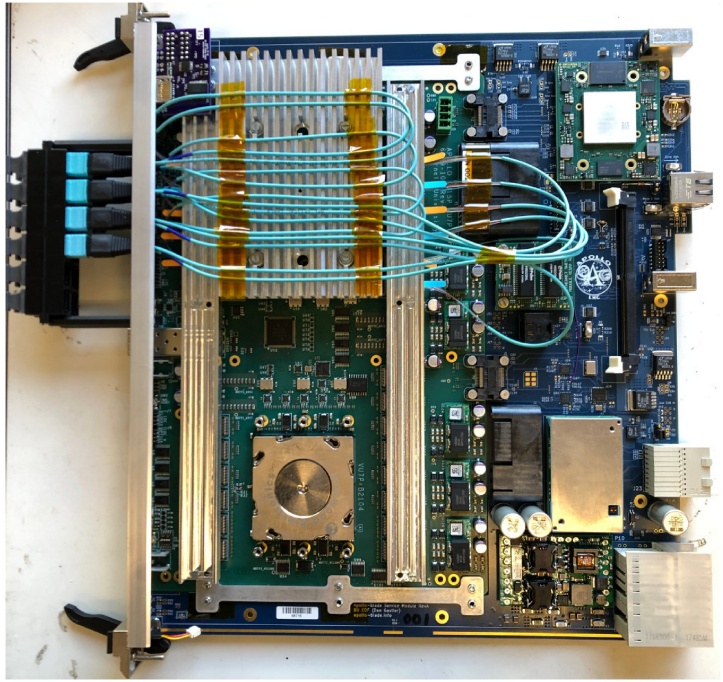
- Flexibility (e.g. FPGA+Optics)
- Modularity; separate complex and simpler part of the board design

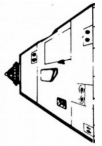
On-board computing and control variety

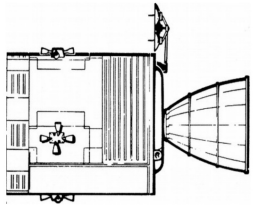
- Zynq SoC
- CERN-IPMC or UW-IPMC

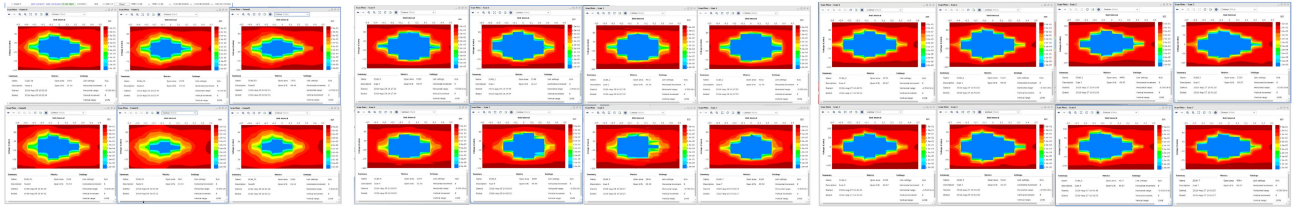
PCB Characteristics:

- 16 layers / Megtron-6 / 1.8 mm
- Apollo analogy: Split into “Command” and “Service” modules





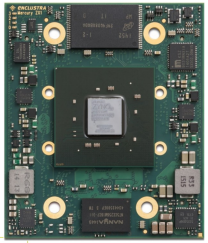


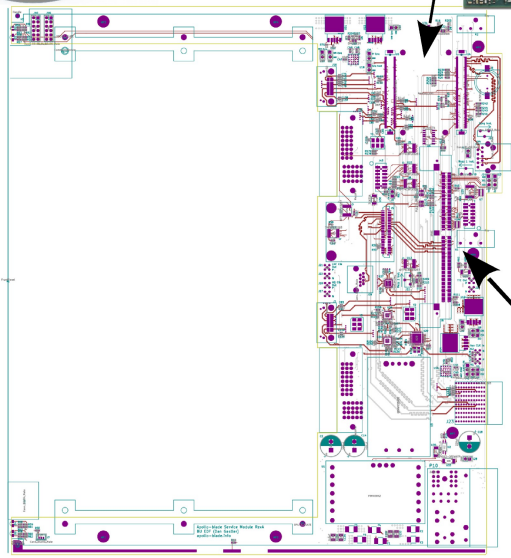


Boston University, Cornell University, Rutgers University, Ohio State University, University of Notre Dame, Northwestern University, University of Colorado

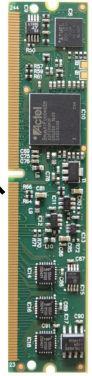
Slow Control at Tracker R&D Boards

Enclustra ZX1
- AXI-C2C
- TCDS
- Eth



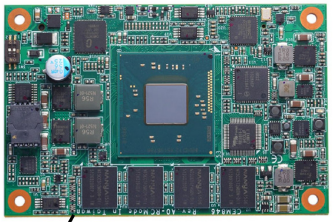


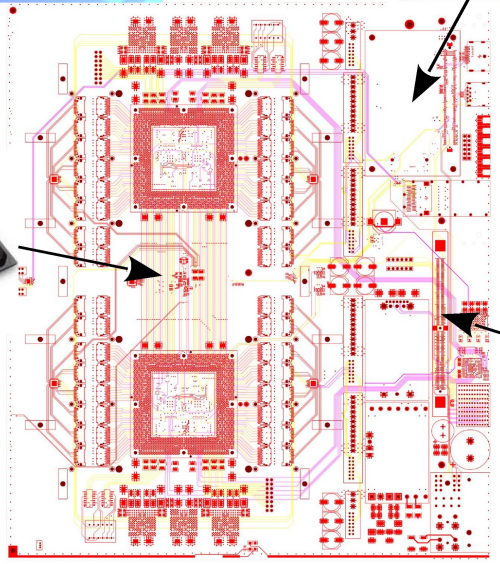
CERN-IPMC
- IPMI
- Eth
- PMBus



SERENITY

Com-Express
- PCIe
- Eth



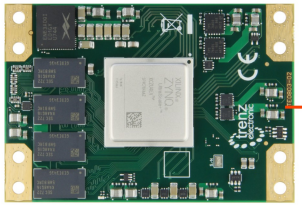


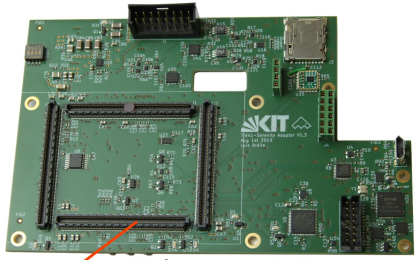
Artix 7
- TCDS
- Eth
- I2C
- PMBus

CERN-IPMC
- IPMI
- Eth
- PMBus

Proof Of Concept: Adapter On Serenity

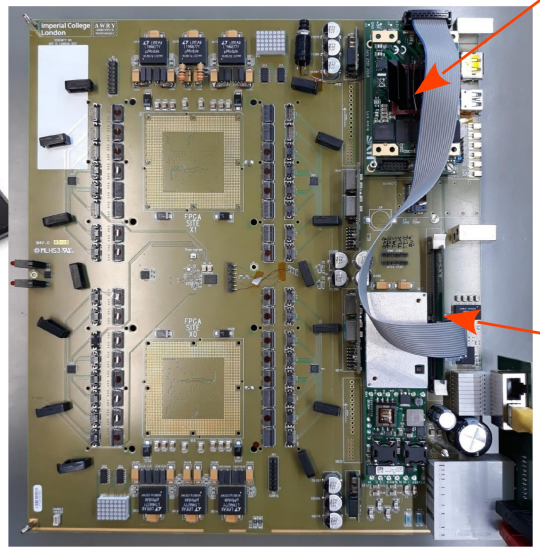
Trenz
ZynqUS+
4EG





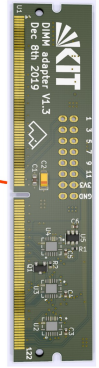
Adapter v1.3

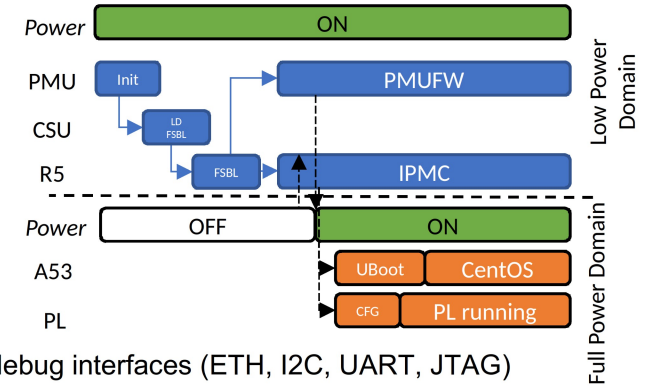
Artix 7
- TCDS
- Eth
- I2C
- PMBus



Serenity v1.0 by Imperial College

IPMC signals





☑ Test and debug interfaces (ETH, I2C, UART, JTAG)

☑ **Pigeon Point IPMC** standalone software compiled for the ARM-R5 processors

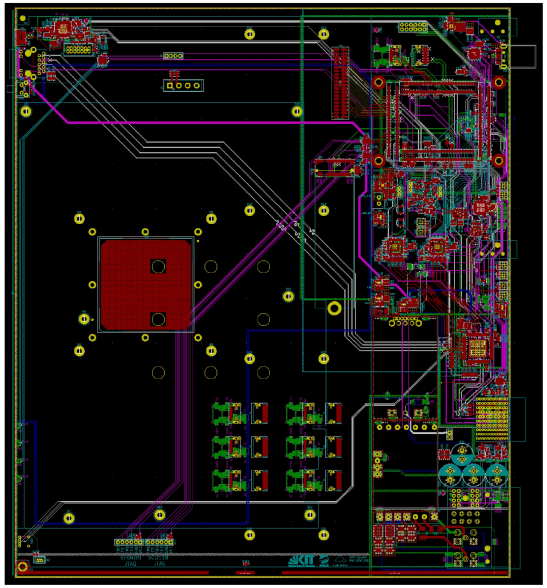
- ☑ IPMC boot, ☑ Com with Shelf Manager, ☑ Board activation/deactivation, ☑ Power-up/power-down sequence, ☑ Read of IPMC sensors, ☑ Cold reset, ☑ Initiating boot of Linux, ☑ Coexistence of Linux & IPMC, ☐ JTAG on Linux (XVC, not integrated)

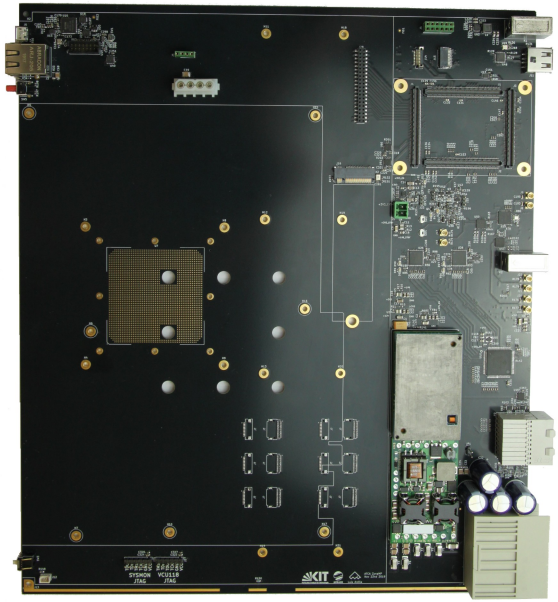
☑ **Petalinux & CentOS** running on the ARM-A53 Processors

- ☑ SSH to Linux on ZynqUS+

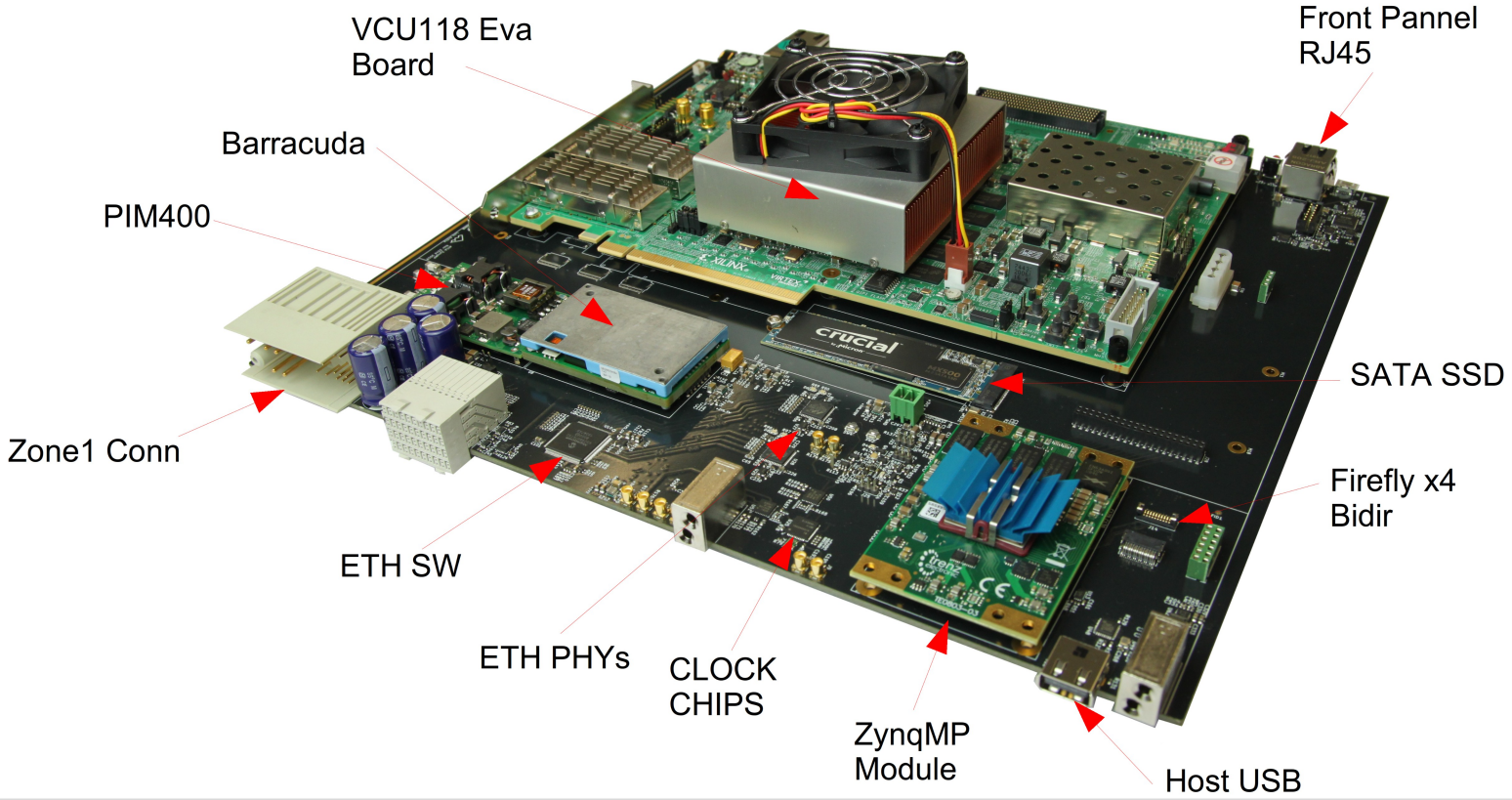
ZynqMP-IPMC ATCA Test Board

- **Power tree** tailored to examine different **configurations** (e.i. always active, only low-power domain, partial reconfiguration)
- independent Ethernet interfaces, one for the management control and one for the Linux OS
- Front panel access to UARTs and JTAG interfaces via FTDI
- Mechaniical mounts for Xilinx VCU118 Evaluation board
- **PCB Assembled this week**, currently under bring up tests

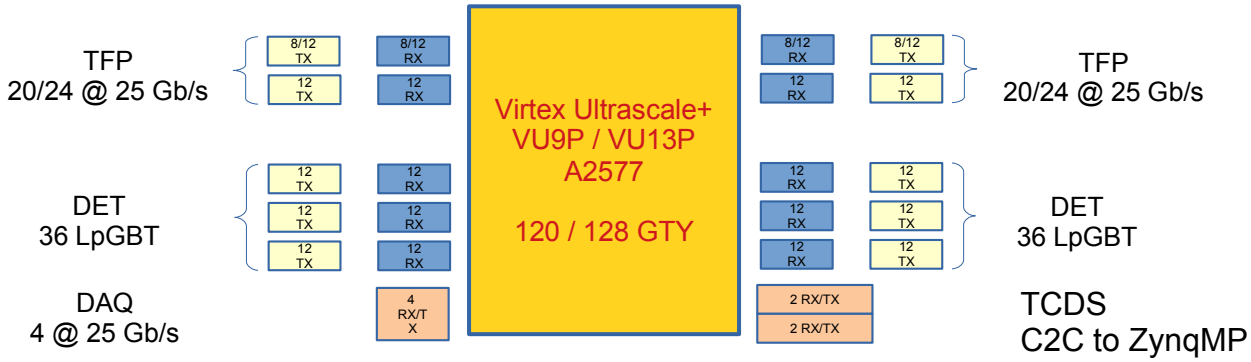




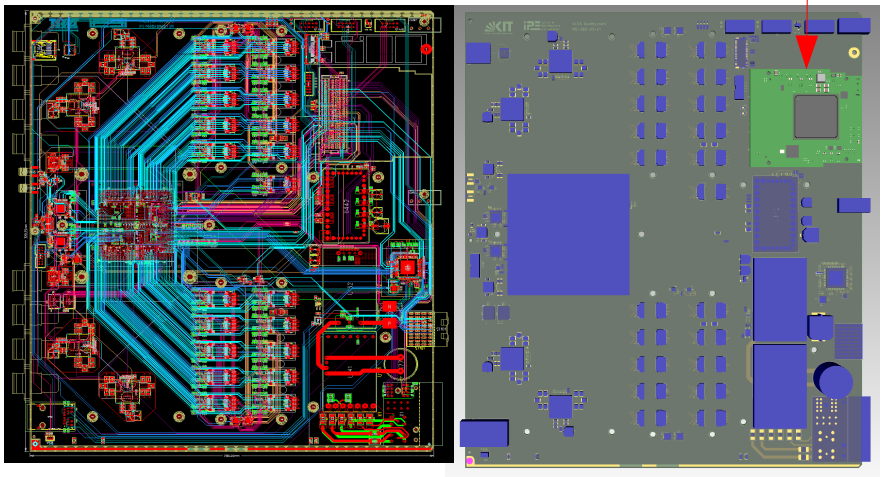
ZynqMP-IPMC ATCA Test Board



KIT – ATCA Board & Management Module

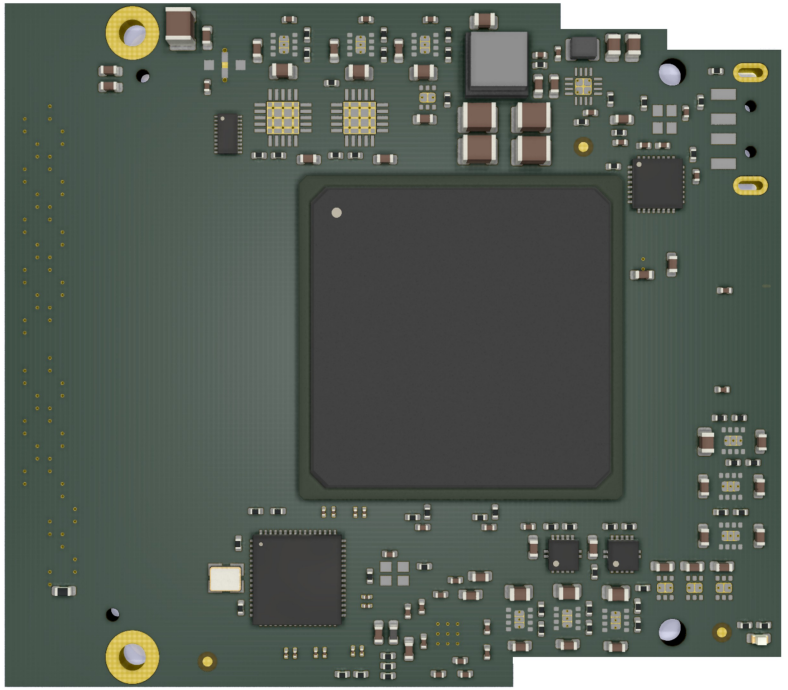


- One FPGA design **VU9P / VU13P**
- **No inter-FPGA** communication
- **Monolithic heat sink** for all optics
- Clean Optical Cable management (**one type of module**)
- TCDS backplane signals **directly** routed to main FPGA
- All Firefly connectors populated with 12x RX or TX signals
- 16 Gbps Firefly Y cable available Today
- 25 Gbps Firefly Y cable expected to be available in Q2 2020



Integrated Management Module

- **Interface defined** with **FMC+** Form factor
- ZU4EG-B900 device with **16 MGTs @ 16 Gbps**
- 2GB DDR4
- CMS Clock distribution with Si5397 chip
- Host USB PHY
- RGMII ETH Phy + SGMII
- SATA x1 lane
- **PL and PS independent power supply**
- IPMC functionality
- Schematic & Placement at 90%



Summary

- CMS **needs** tracks at L1 trigger to cope with HL-LHC pileup conditions
  - p_T modules provide **first layer** of efficient **data reduction**
- **Highly flexible** track-finder/pattern recognition algorithms were demonstrated in hardware
- **Highly scalable**, time/physical segmentation could be as large/small as required based on data rates
- **Proven** with **currently available hardware**, that a level-1 track-trigger based on **FPGAs** is **feasible**
- Lots of **flexibility** with an **all-FPGA solution**

Common infrastructure R&D

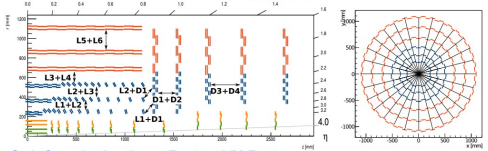
- Current prototypes showed very good **optical performance**
  - Evaluation of other optical drivers is planned in the future
- Realizing all management functionality in a single MPSoC is an **exciting** solution for next generation boards in the DAQ chain
- Proof of principle **successfully built** and **tested** for the unified controller architecture

Track Finding Algorithms

Karlsruher Institut für Technologie

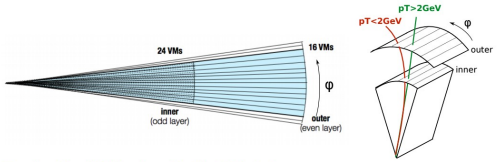
Data Flow

Tracklet



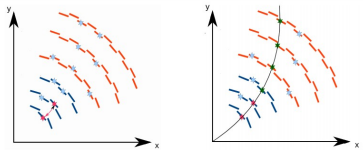
Stub Organization: Layer Router - VM Router

Sorts input stubs into each layer/disk and splits processing module into virtual modules (VM)



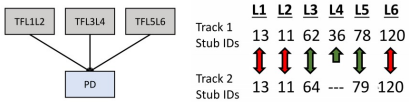
Seeding: Tracklet Engine - Tracklet Calculator

Forms a tracklet from pairs of stubs in adjacent layers according to the VMs, then calculates exact tracklet parameters for each pair



Projections and fitting: Match Engine/Calculator & Linearized χ2 fit

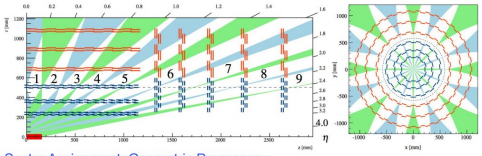
Looks for stubs within a window around the projected track and fits it



Duplicate Removal

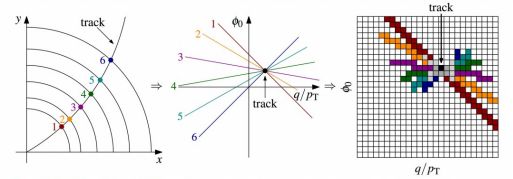
Compares tracks pairwise and eliminates those with less than 3 unique stubs

TMTT



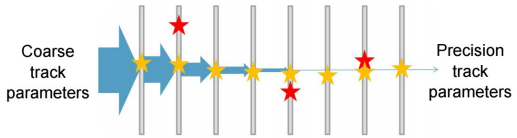
Sector Assignment: Geometric Processor

Processes stub data, and sub-divides the nonant into 36 finer sub-sectors



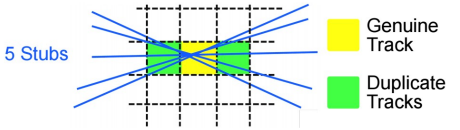
Track Finding: Hough Transform

Track finder that identifies groups of stubs consistent with a track in the r-φ plane



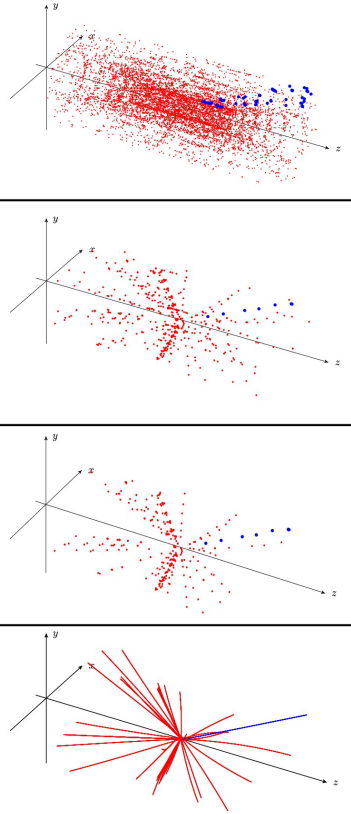
Track Fitting: Kalman Filter

A candidate cleaning and precision fitting algorithm

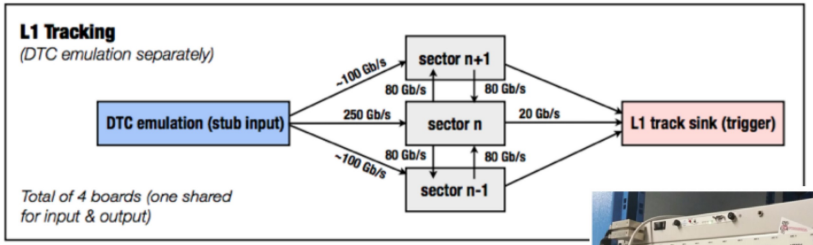


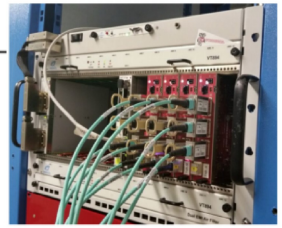
Duplicate Removal

Uses precise fit information to remove duplicate tracks generated by the HT



Hardware Demonstrators - 2016

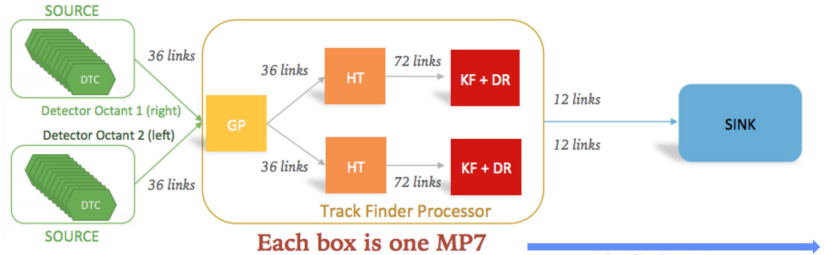


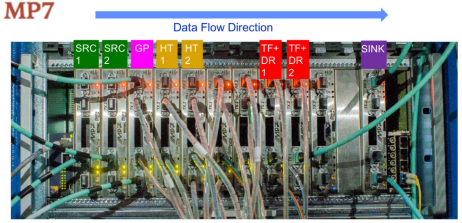


Half a barrel demonstrator in hardware, verified using emulation software

Hardware demonstrator has been built to validate the algorithm and measure latency

- 4 CTP7 boards with Virtex-7 FPGA – 3 CTP7 cover 3 Φ sectors – 1 CTP7 emulate DTC
- 1 AMC13 card for clock and synchronization
- 240 MHz internal fabric speed
- Measured latency of **3.33 µs** in agreement with latency model – without duplicate removal step





Both Demonstrators were tested with samples from PU 0→ 200

Demonstrator in hardware and emulation

- One per time multiplexing and detector nonant
- Each box is one MP7 board with Virtex-7 FPGA
- Can compare hardware output directly with software
- 240 MHz internal fabric speed
- Latency verified to be **3.5 µs**

Integrated Board Management

Low Power Domain – always active:

- **Intelligent Platform Management Interface (IPMI)** application running in one of the R5 cores.
- It uses the On-chip-Memory and the I2C peripherals
- IPMC memory region protected through system configuration
- SPI and PMBus

High Power Domain – active upon request of full power from the crate:

- Runs Yocto/CentOS based Linux
- FPGA configuration and monitoring
- Slow-control to FPGAs
- Test patterns to firmware

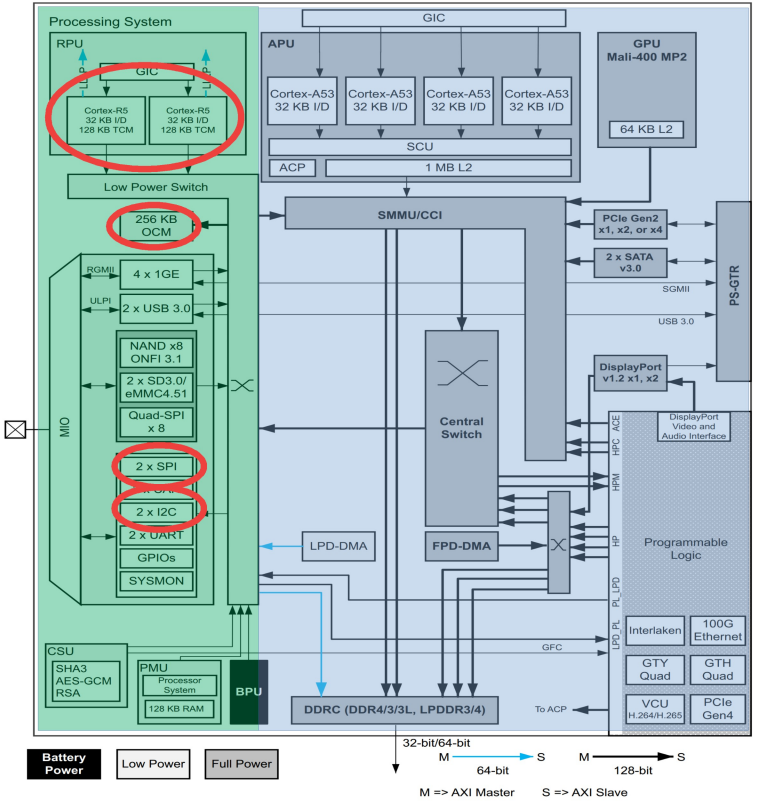
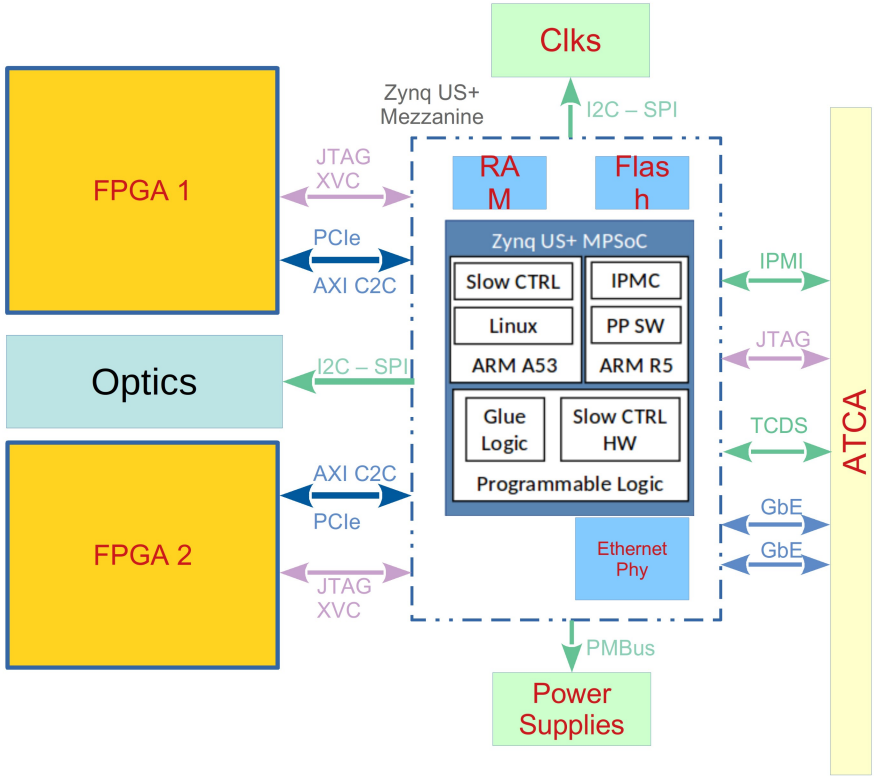


Figure 2-1: **Zynq UltraScale+ MPSoC Device Hardware Architecture**

Integrated Board Management

Integrate IPMC, GPP based slow control functionality and FPGA in a **single heterogeneous MPSoC** (Zynq Ultrascale+)

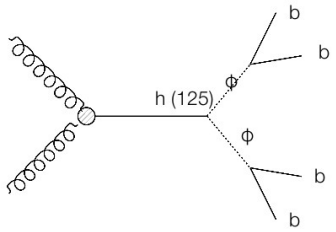
- **Intelligent Platform Management Interface (IPMI)** in ARM-R5 processor running freeRTOS
- **Timing and Control Distribution System (TCDS)** in PL-FPGA
- Xilinx Virtual Cable (**XVC**) JTAG
- **AXI Chip2Chip** slow control capable

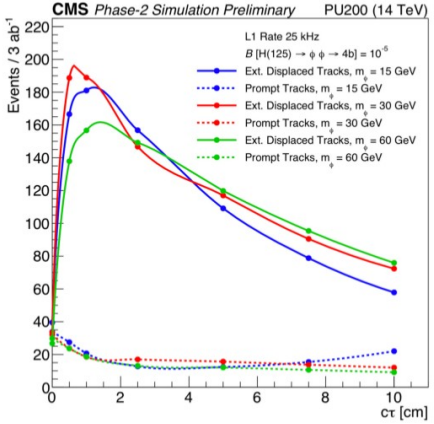


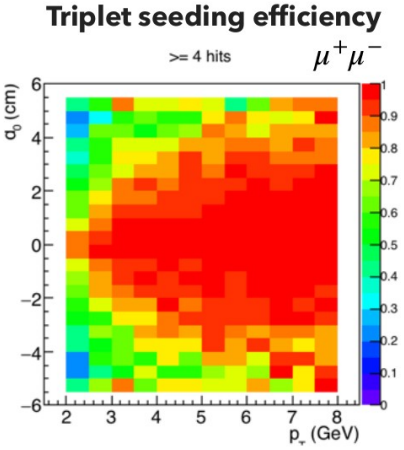
Displaced Track Finding

http://cds.cern.ch/record/2647987

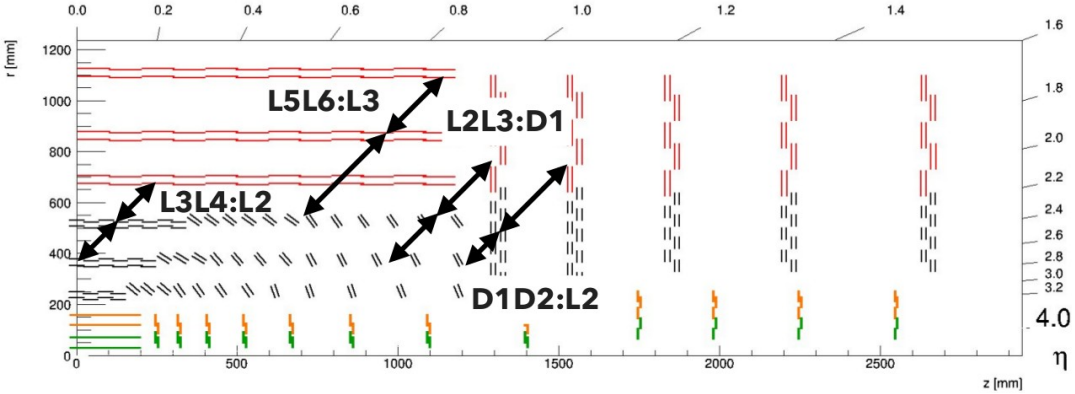
- **Motivation**
  - Lots of **interesting physics** with displaced tracks e.g rare Higgs decay to a long lived (dark matter) ϕ (~no background)
  - Alternative to **expensive dedicated experiments**





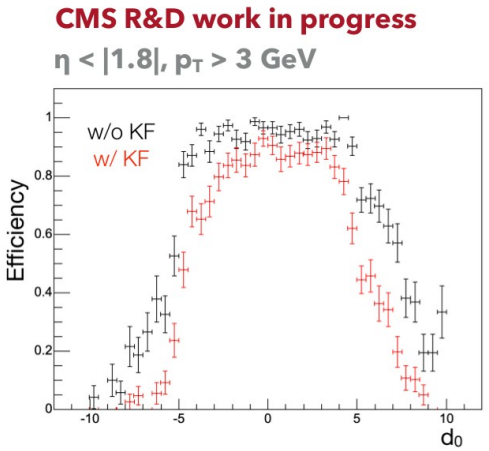


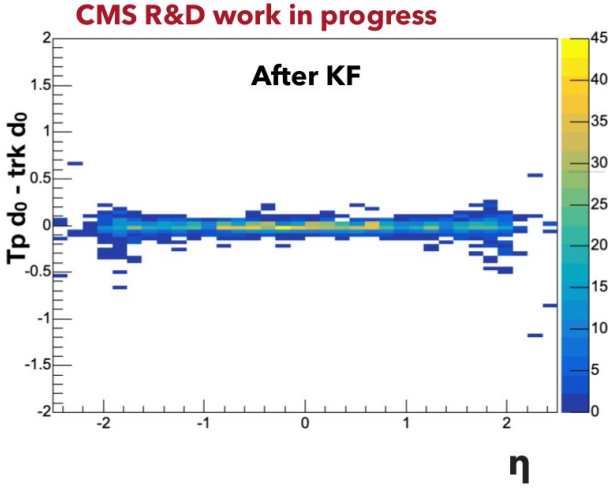
- **Challenges**
  - No beam point constraint -> higher (but manageable) fake rates
  - **Increased processing requirements** - truncation vs FPGA resources
- **Adaptations**
  - Seed with stub triplets
  - Fit with 5 param fit (d_0)

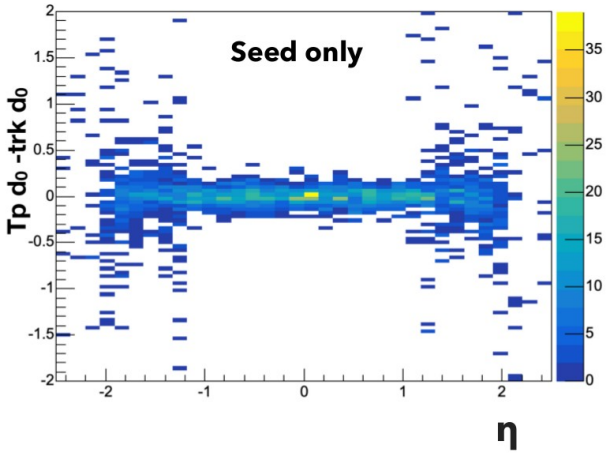


Displaced Track Finding

- **Efficiency**
  - Up to 5 - 10 cm
  - Limitation is bend cut on FE in inner layers
- **Rate**
  - 1.2 x increase in rate w.r.t 5 param prompt
  - 1.4 x increase w.r.t 4 param prompt

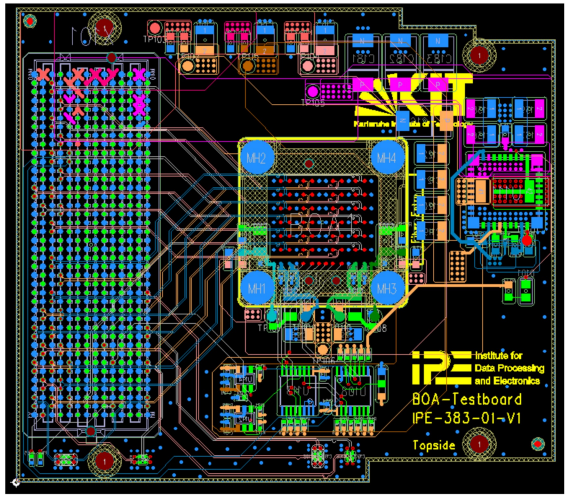


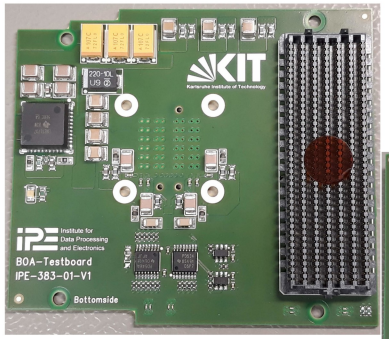


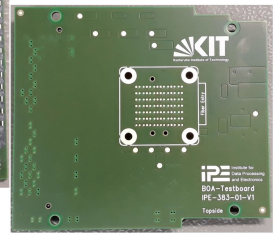


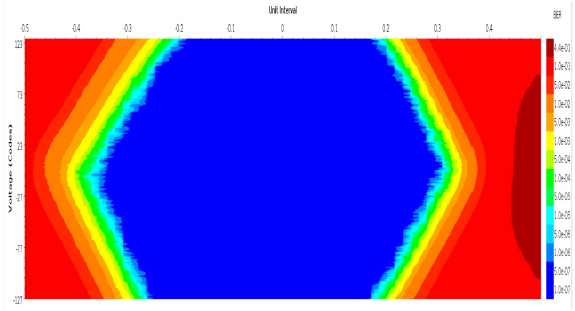
High-Speed Optical Evaluation

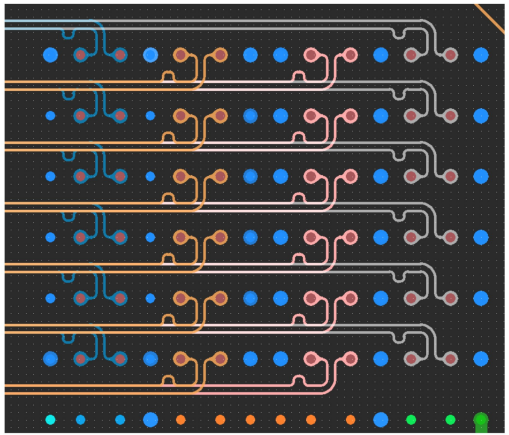
- FMC+ sized board for evaluation of the Finisar BOA 25 Gb/s transceiver
- 12 TX and 12 RX integrated in the same package
- 4 Electrical loop-back channels capacitively coupled with different features
- Skew < 20 µm
- MT ferrule optical interface
- Performance of capacitively coupled lanes looks good











Thermal Simulation And Tests

Simulation setup

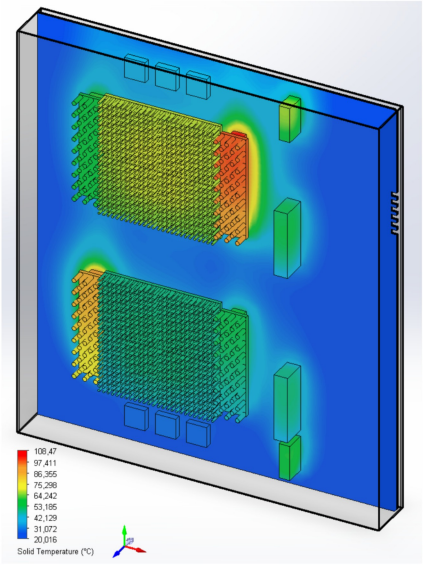
- PCB imported from PADS
- Placed in a 33 mm deep tunnel
- 4 m/s airflow from bottom (20 °C) to top

Placed components

- KU15P (50 W) doubled θ_{JB} to take interposer into account
- Firefly banks 25 G (30W) and 16 G (12 W)
- **Total power 205.4 W**

Test setup

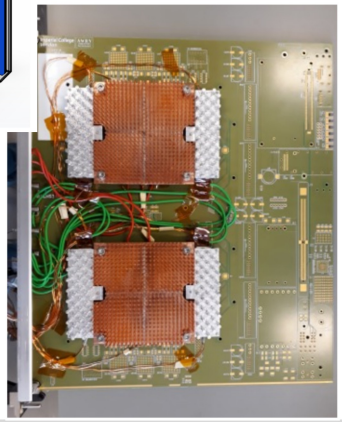
- Two heat-pads **45 mm x 45 mm** and **12 mm x 70 mm**
- Just one mockup board is present, it will be put in between two additional soon
- **~11 W for 6x block** of 16 Gbps optics
- **~10 W for 6x block** of 25 Gbps optics

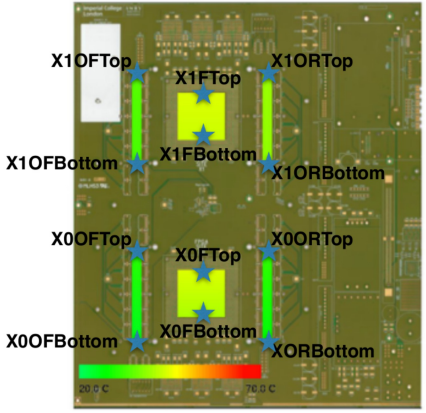


Test1 (°C)
4xFan-block speed=50%
Exhaust temp ~17°C (~amb)
Power on FPGA heaters = 86 W
Power on Optics heaters = 41 W

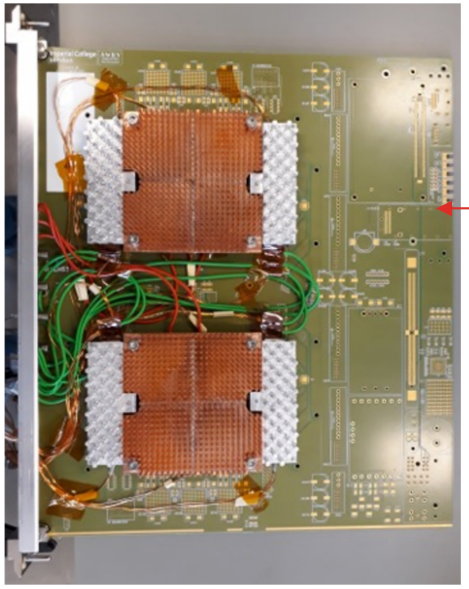
X1FTop = 60.7
X1FBottom = 59.1
X1ORTop = **50.8**
X1ORBottom = **49.7**
X1OFTop = 43.1
X1OFBottom = 41.7

X0FTop = 53.7
X0FBottom = 50.1
X0OFTop = 35.8
X0OFBottom = 28.2
X0ORTop = 37.2
X0ORBottom = 31.1





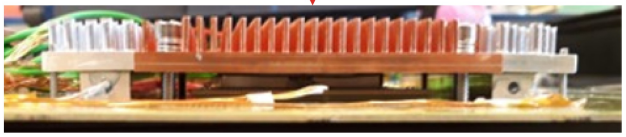
THERMAL & MECHANICAL TESTS

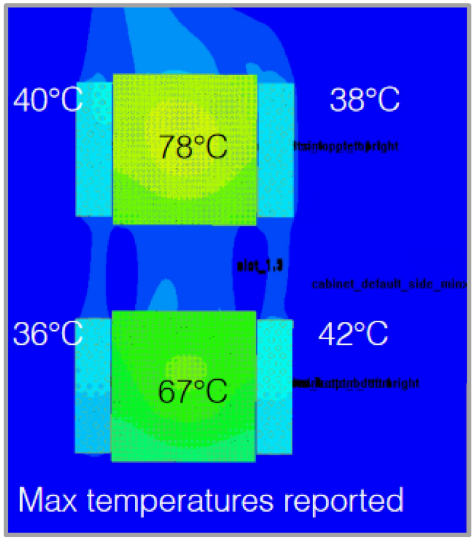


Thermal simulations

Physical thermal studies at CERN

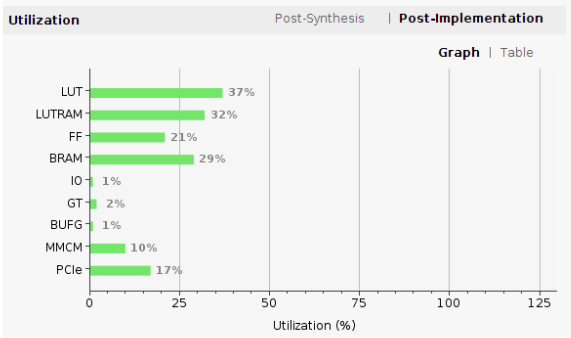
Mechanical component design studies into stress on FPGA solder balls and stress on PCBs at IC



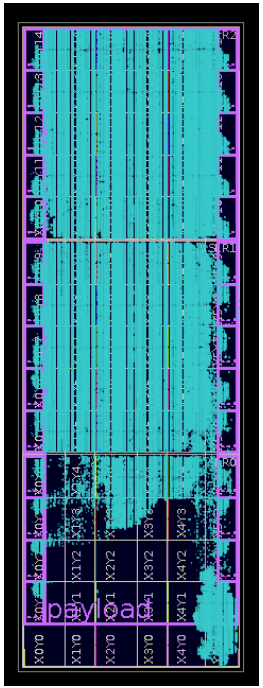


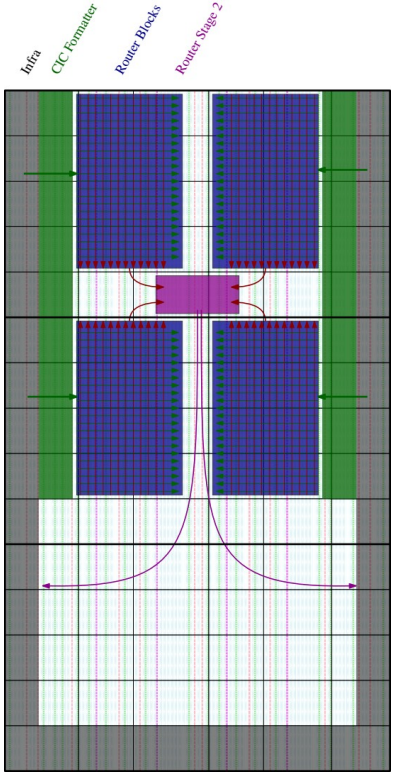
DTC Firmware in VU9P

- Target single FPGA (VU9P) to avoid side wise communication
- Use of EMP framework with 64 bit frames at 320 MHz
- Input packet structure of 58 data frames + 6 frames gap
- 5 gbps modules send up to 16 stubs per CIC in 8 BX packets
- 10 gbps modules send up to 35 stubs per CIC in 8 BX packets
- One 64 bit word contains a Stub from CIC0 and a Stub from CIC1

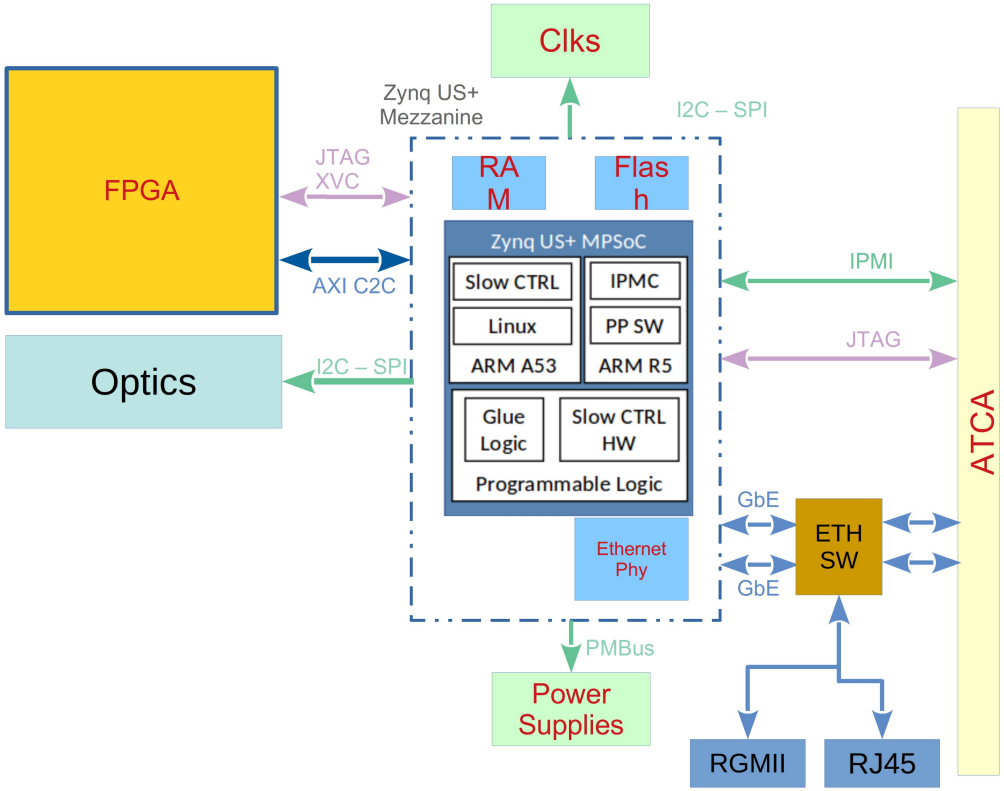


Taken from Thomas Schuh





Integrated Management Module Architecture



- **IPMI** (standalone/RTOS) on ARM-R5 processor
- **Slow Control** (Linux) on ARM-A53 Cores
- Xilinx Virtual Cable (**XVC**) JTAG
- **2 Links to main FPGAvia PL-MGTs** (AXI C2C)
- **I2C-SPI** to configure Optics/Clocks
- **PMBus** to configure Power Supplies
- **Eth** and **I2C** backplane connection

ATCA Layout

- One FPGA design **VU9P / VU13P**
  - **No inter-FPGA** communication
- **TCDS** backplane signals directly routed to main FPGA
- Integrated **IPMC** slow control solution
- **Monolithic** heat sink for all optics
- Clean Optical Cable management
- **Only one type of Firefly cable x10**