CMS Calorimeter Trigger Upgrades

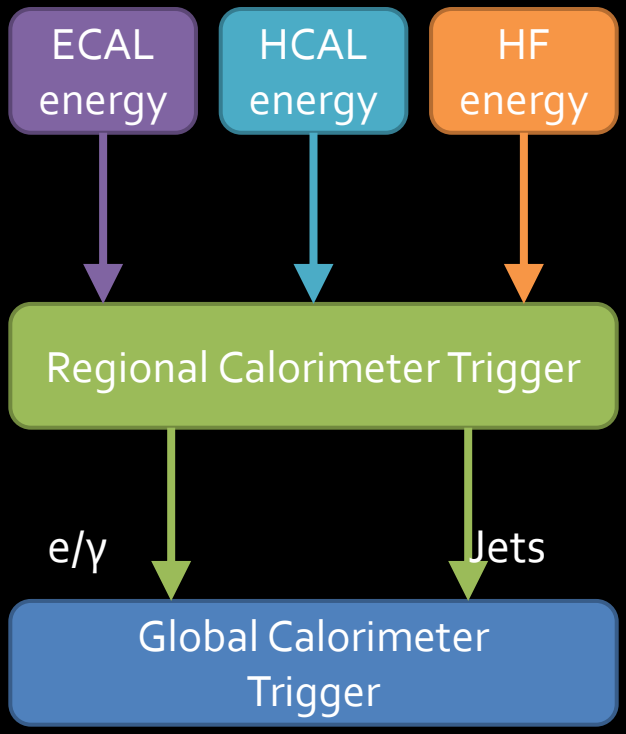
Andrew Rose – Imperial College, London
on behalf of the CMS Level-1 trigger collaboration

Overview

- So many triggers, so little time
- Hardware for upgrades
- Time-multiplexed triggering
- Status of the TMT concept
- Firmware
- Conclusions

So many triggers, so little time

The CMS Calorimeter Trigger

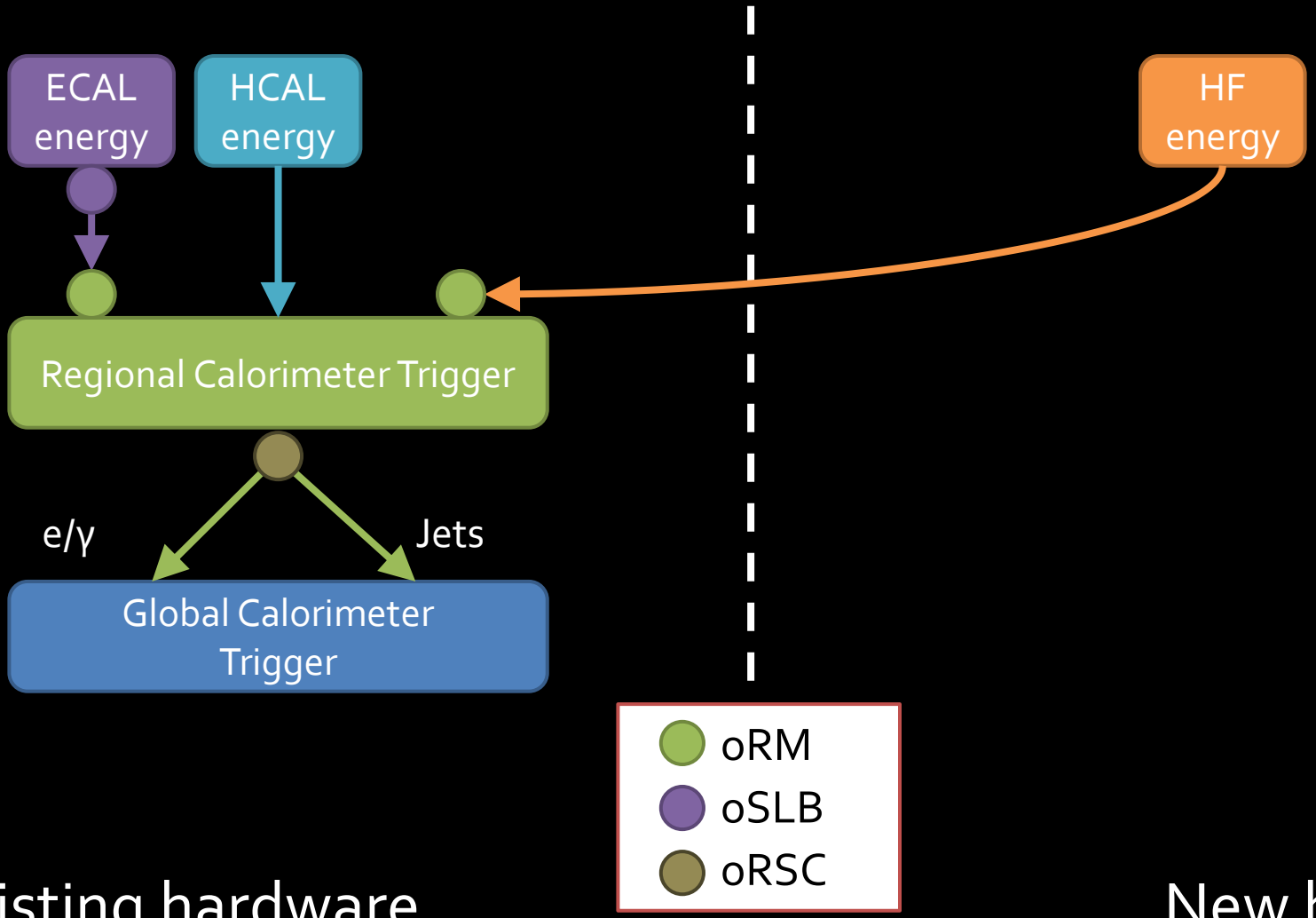


Mitigate risk by building upgrade in parallel:

- Current system, albeit modified, remains operational in parallel
- TDR upgrade
  - All inputs in optical format
  - Full tower-level information available in processor cards
- Interim trigger for 2015
  - Improve processing of existing RCT
  - Backup solution for 2015 running
  - For use if the current trigger cannot cope and full upgrade not yet available

Existing hardware

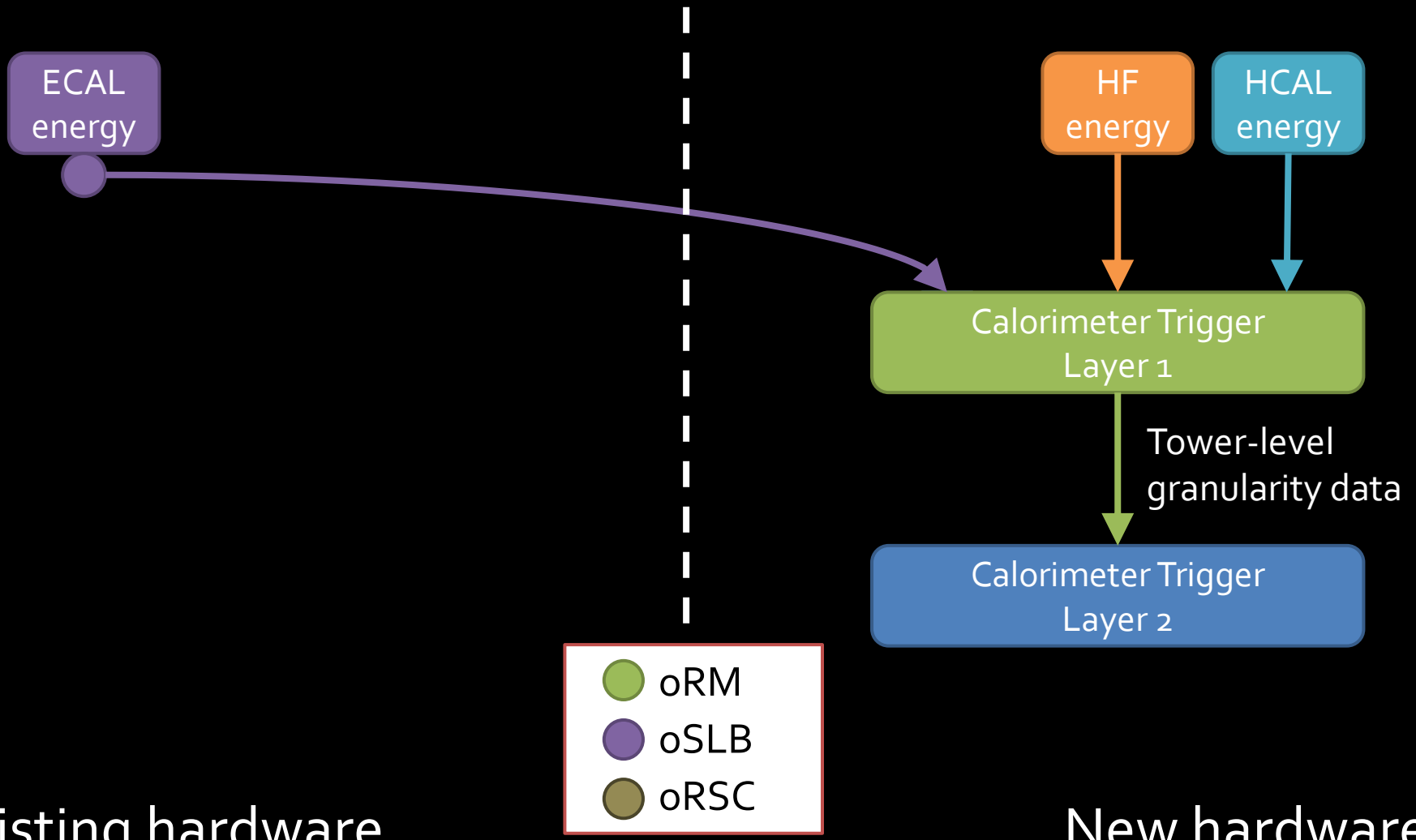
Recommissioned existing trigger



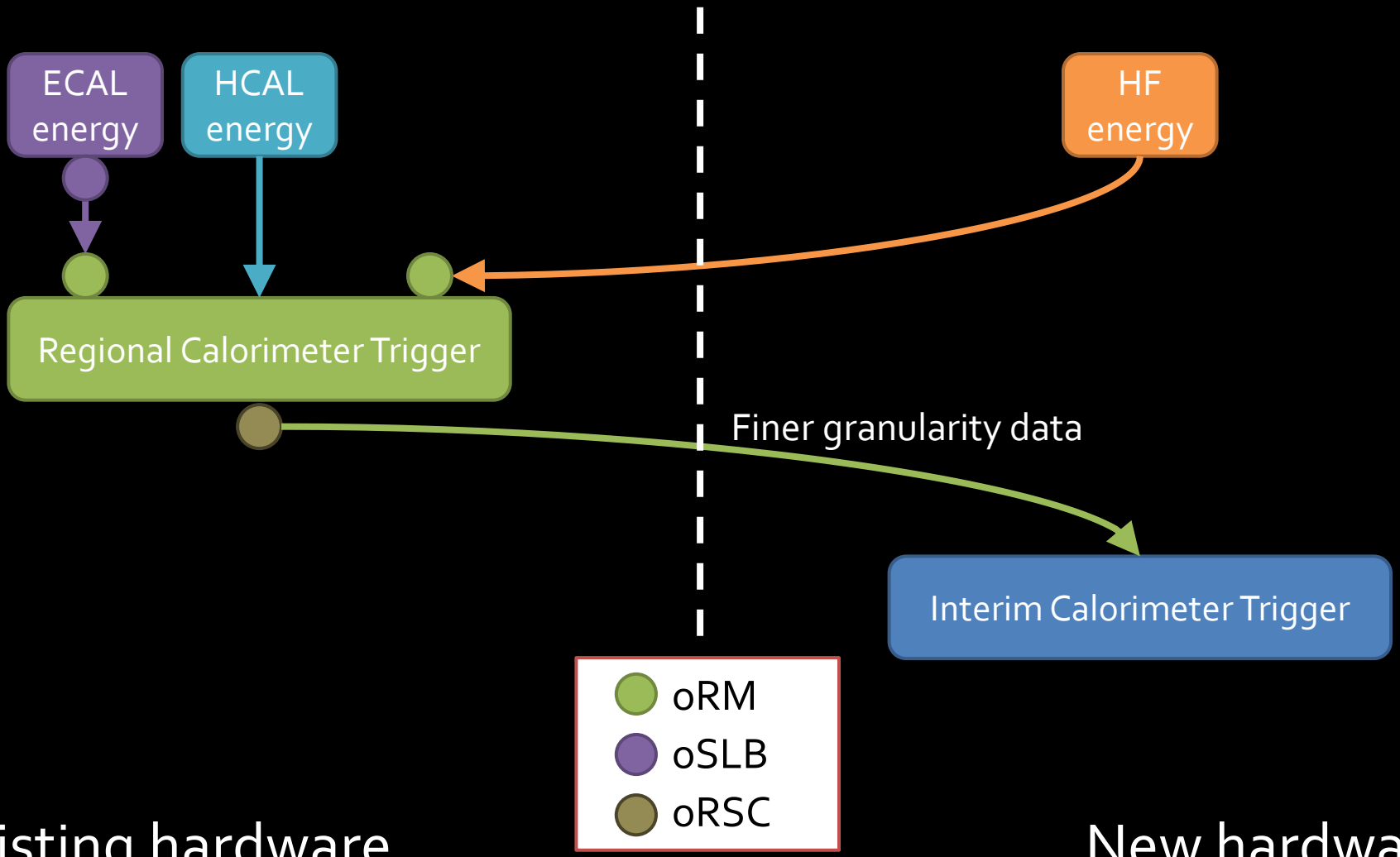
Existing hardware

New hardware

TDR trigger upgrade



Interim (back-up) trigger for 2015



Schedule

The schedule for the calorimeter trigger upgrades is defined in the TDR:

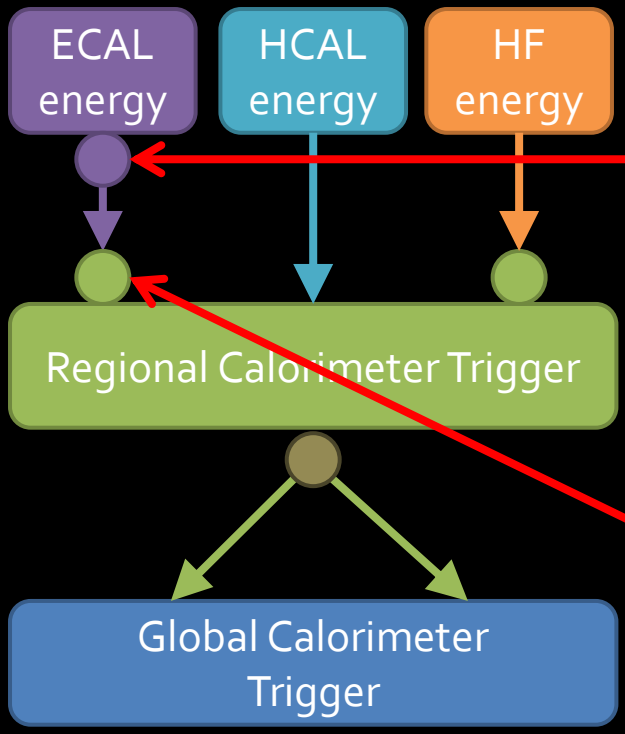
- Slice-test of the TDR trigger design – July 2014
- All hardware installed at P5 – November 2014
- Commissioning in parallel with the existing trigger – 2015
- Fully commissioned and operational – January 2016

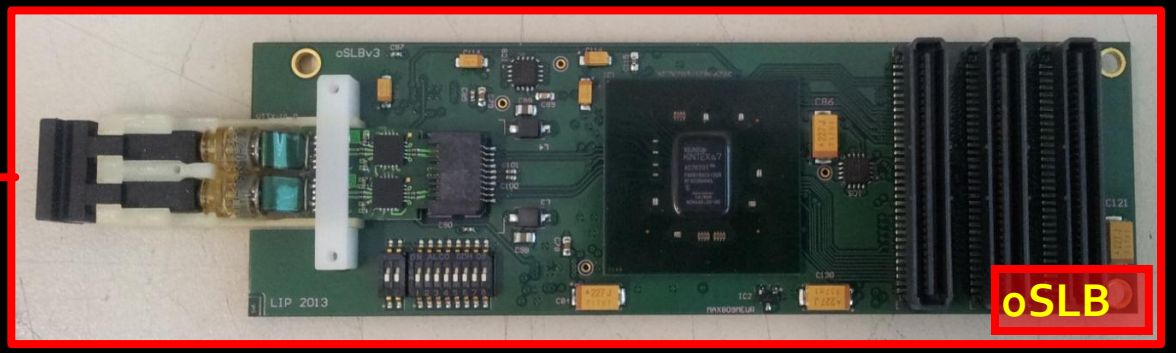
The interim trigger for 2015 has a deadline for the completion of commissioning of 31st March 2015

Hardware for upgrades

LABORATÓRIO DE INSTRUMENTAÇÃO E FÍSICA EXPERIMENTAL DE PARTICULAS

oSLB and oRM

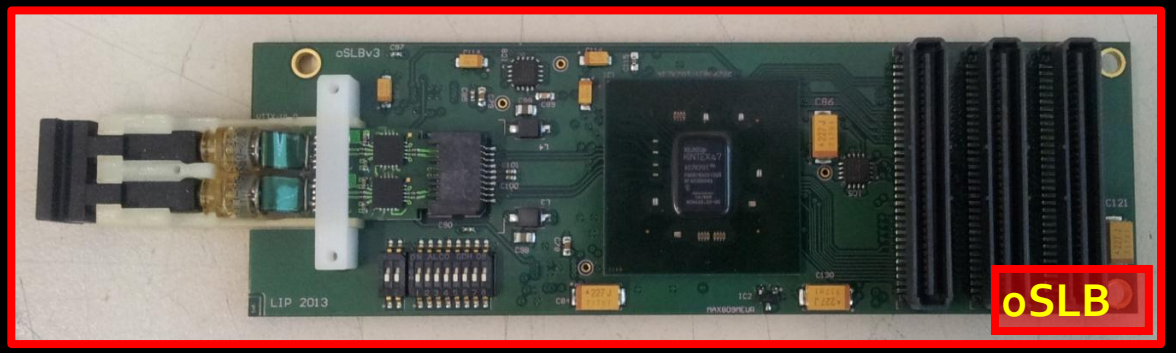


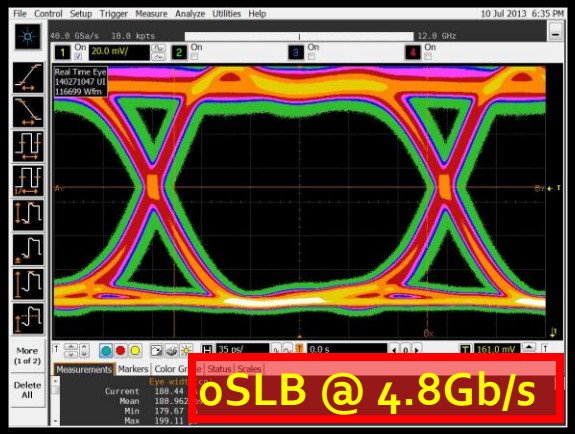


LABORATÓRIO DE INSTRUMENTAÇÃO E FÍSICA EXPERIMENTAL DE PARTICULAS

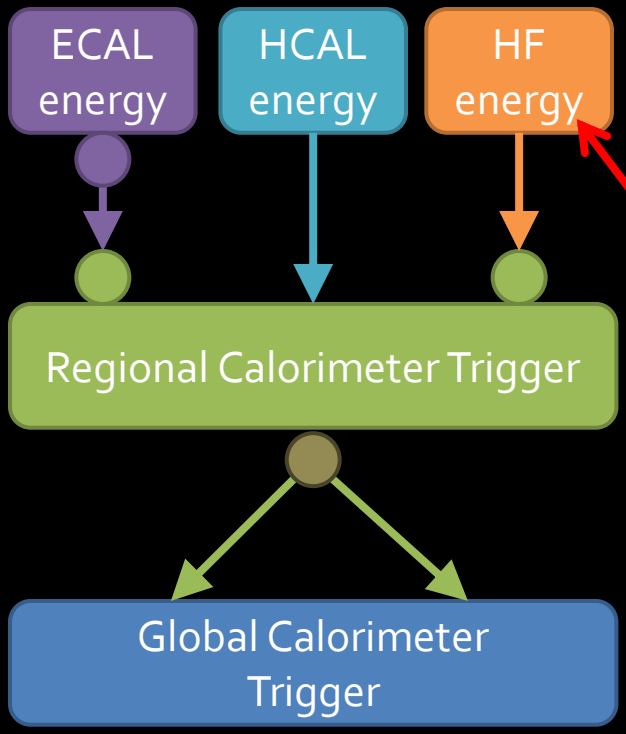
oSLB and oRM

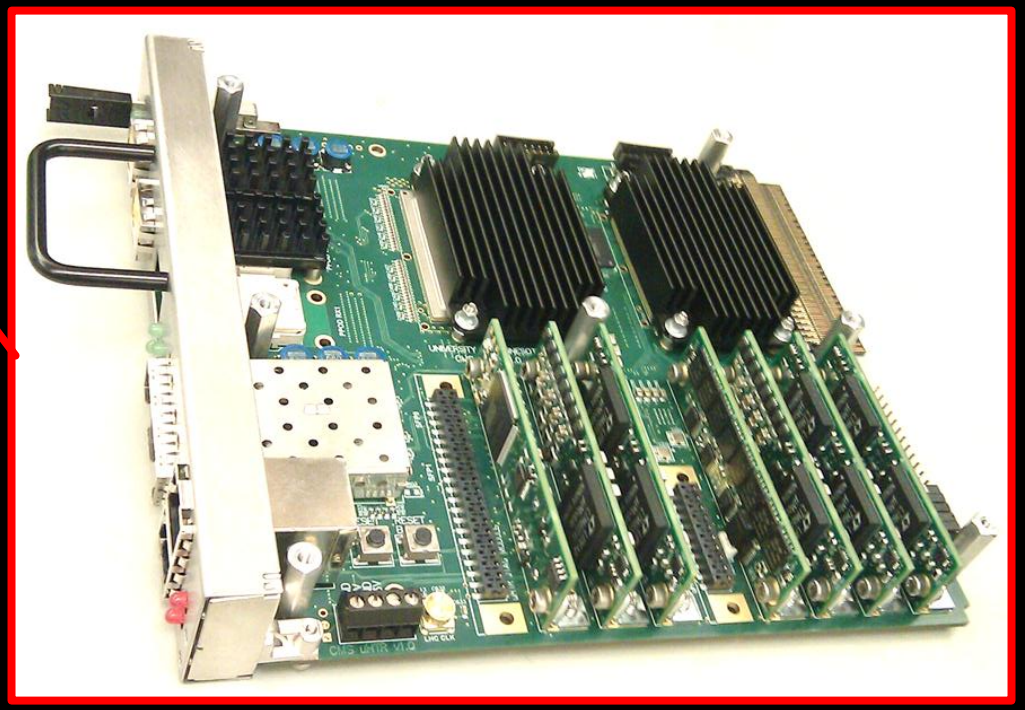
- Kintex-7 FPGAs
- Mount on existing VME cards, replacing copper links
- Have now been manufactured
- Installation pending an installation readiness review





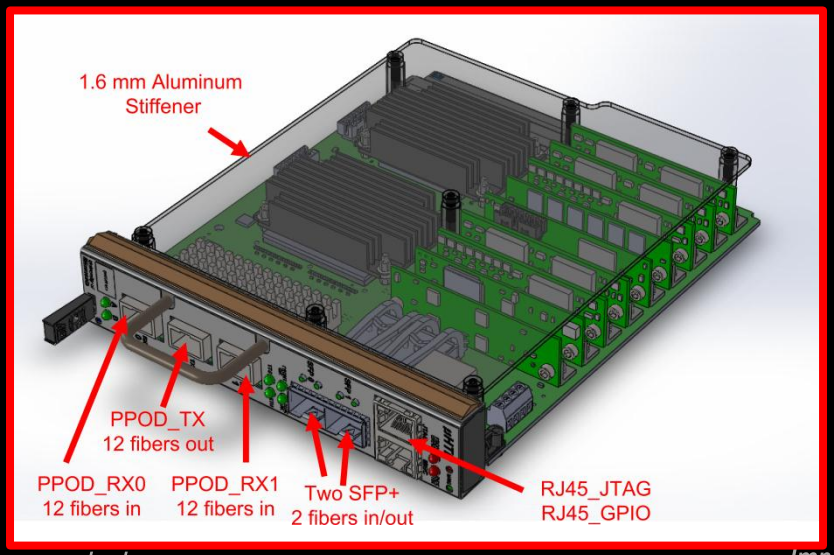
μHTR

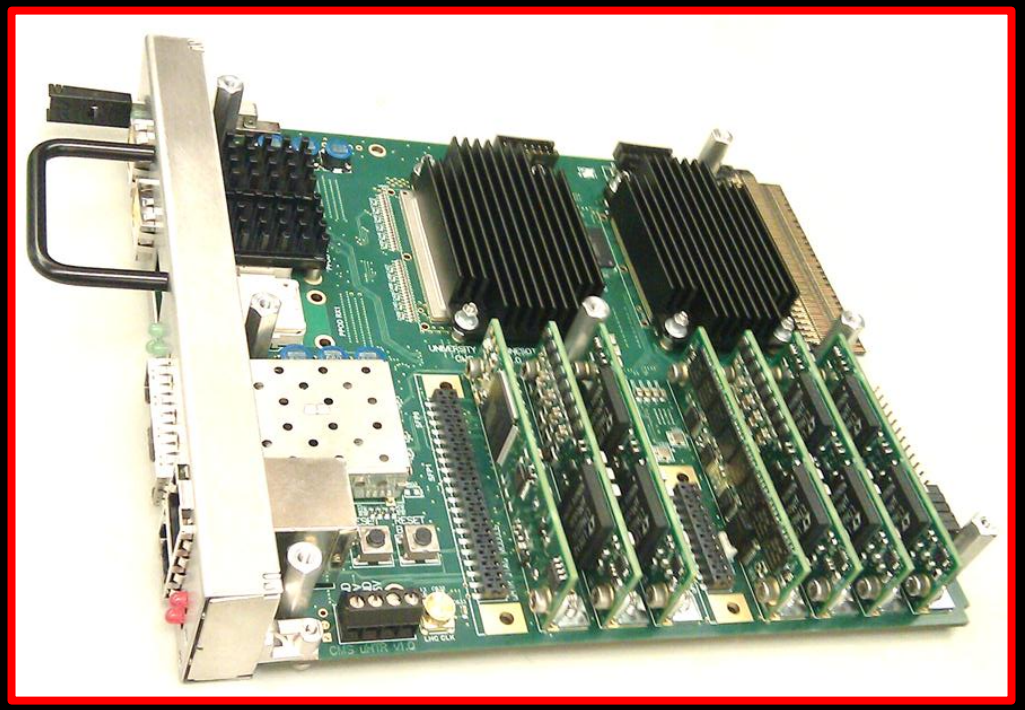

ECAL energy
HCAL energy
HF energy
Regional Calorimeter Trigger
Global Calorimeter Trigger



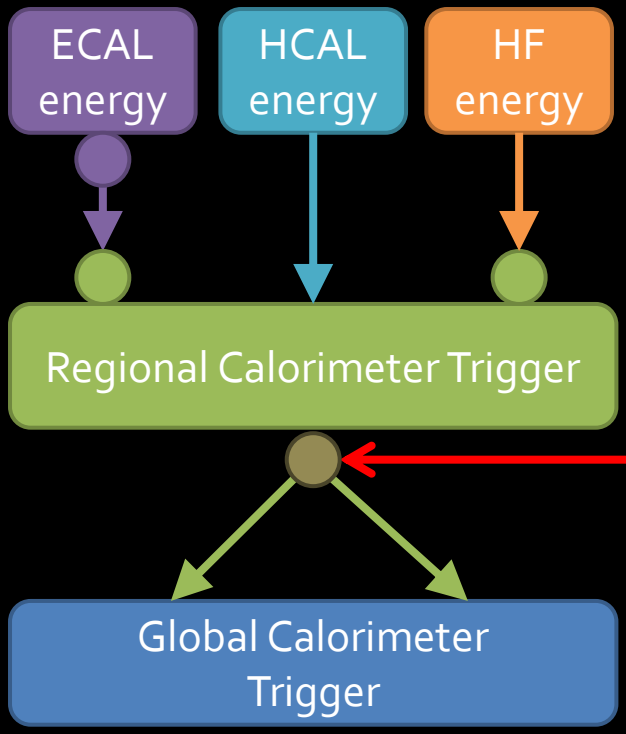
µHTR

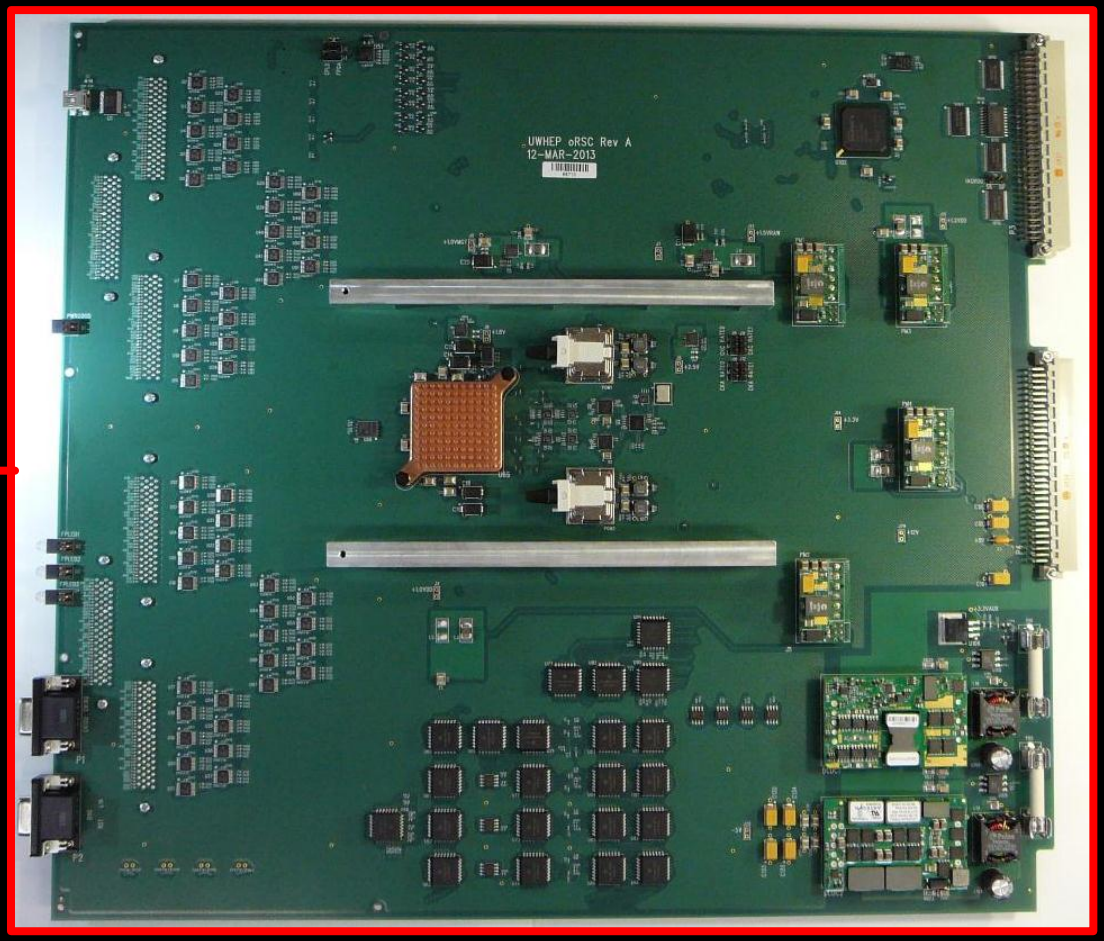
- 2× Virtex-6 FPGAs
- Cost-optimised
- One-for-one replacement of existing HTR cards
- Modular design with services on mezzanine cards
- Mixed PPod and SFP+ optics





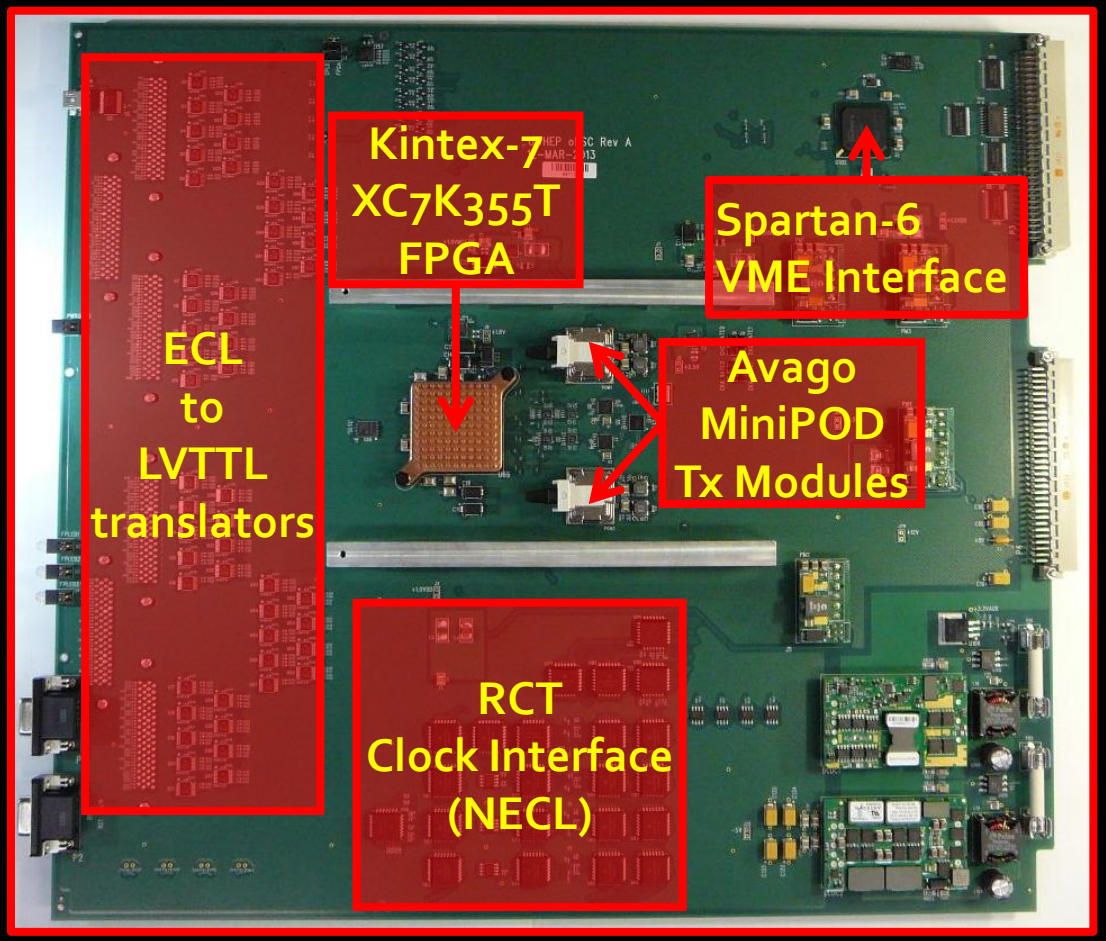
oRSC


ECAL energy
HCAL energy
HF energy
Regional Calorimeter Trigger
Global Calorimeter Trigger

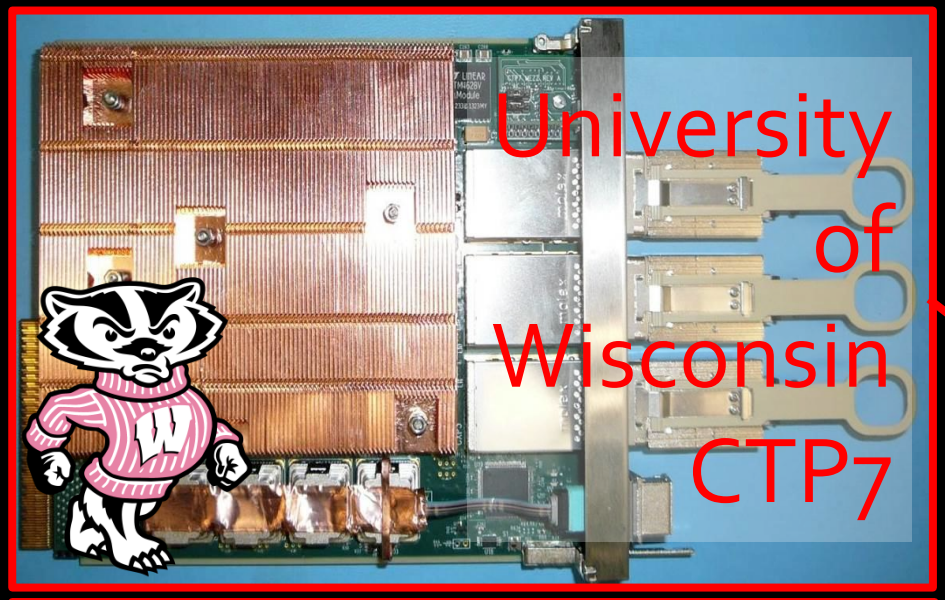

UWHEP oRSC Rev A
12-MAR-2013

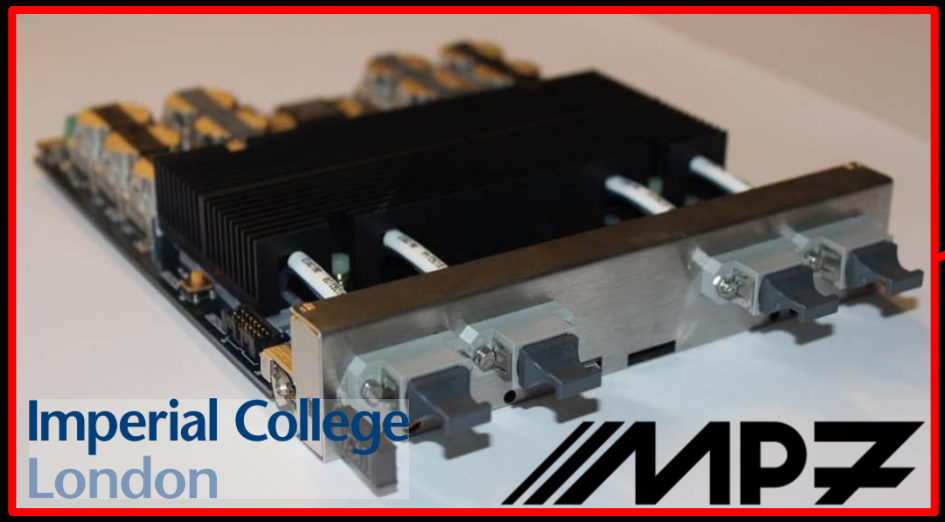
oRSC

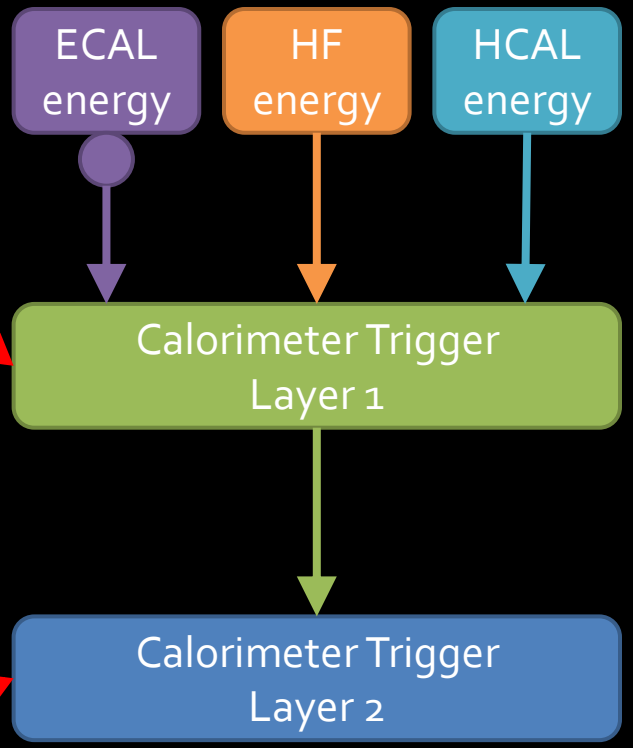
- Kintex-7 FPGA
- VME interface card
- Fits in current RCT crates
- 18 cards in total
- Provides RCT output in optical format to:
  - Existing GCT Leaf cards
  - Interim trigger processor
- Waiting on integration tests before final manufacturing ordered



TDR trigger upgrade

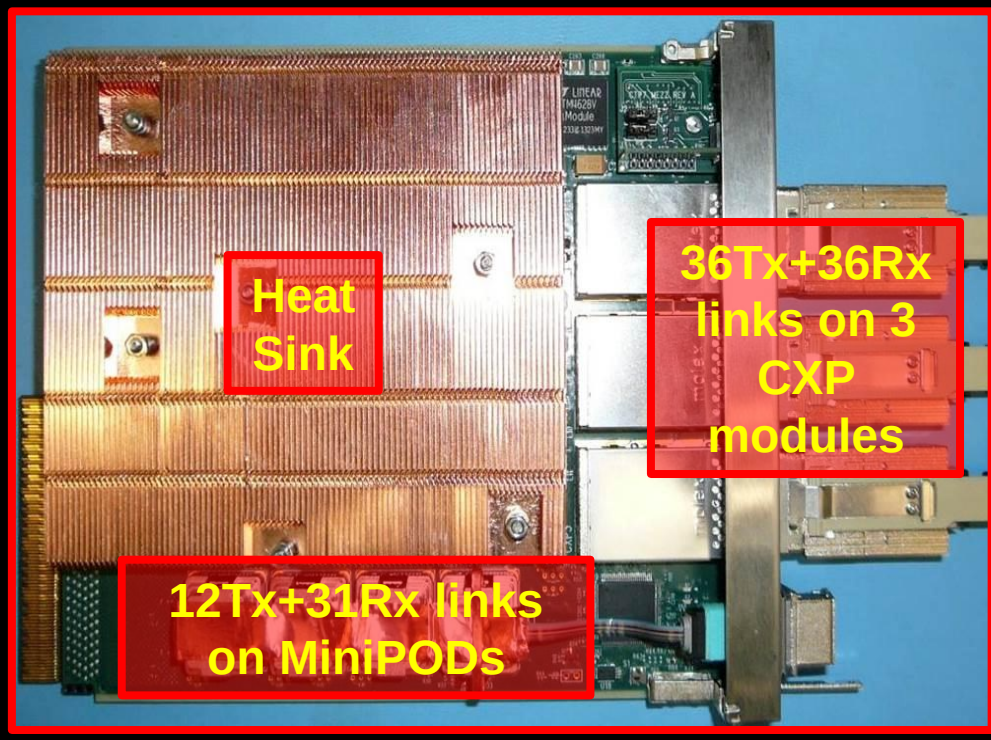






CTP7

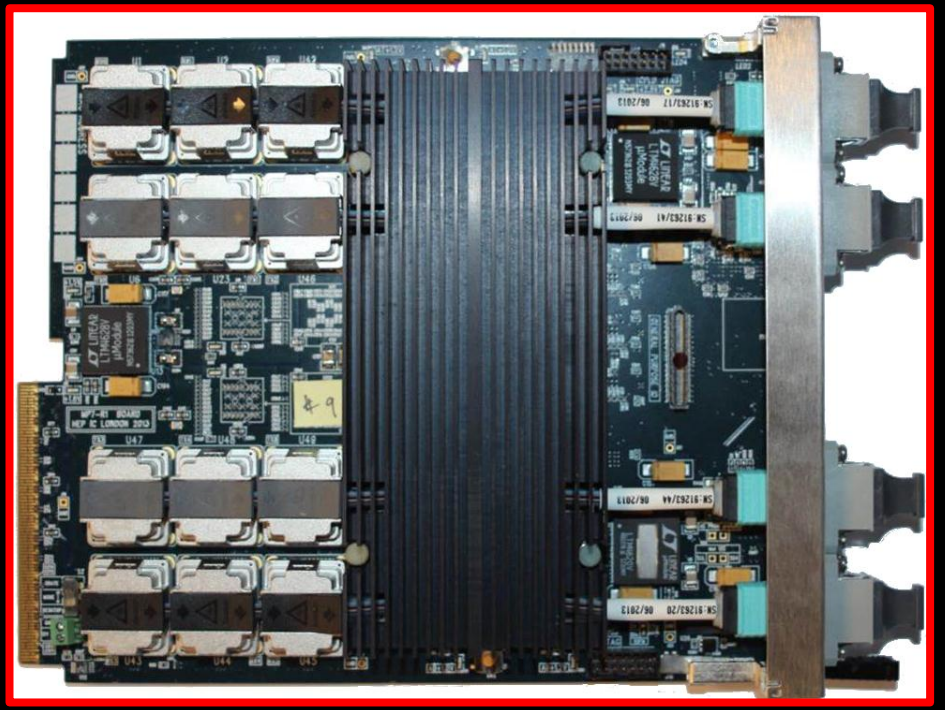
- Based on Virtex-7 XC7VX690T FPGA
- Zynq processor running Xilinx PetaLinux for service tasks
- Baseline Layer 1 hardware
- Backplane communication on Ports 12-15 & 17-20 for use with semi-custom VT894 crate
- Auxiliary power connector on front-panel
- Custom-designed heatsink
- Two prototypes in hand

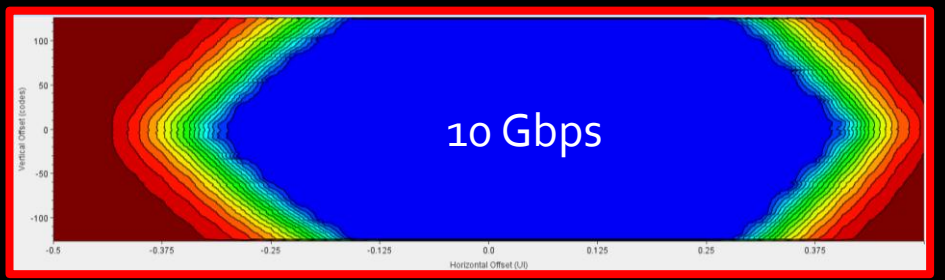


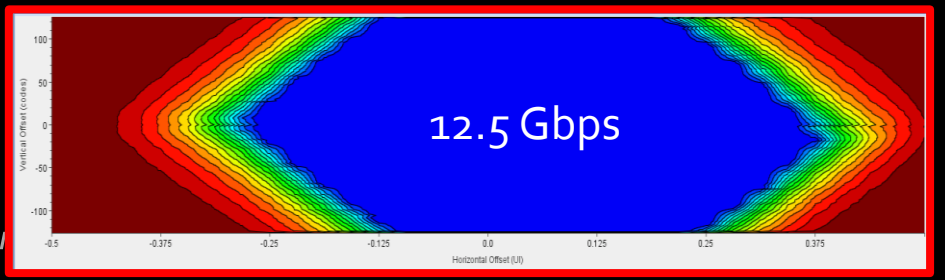
Imperial College London

MP7

- 1.8Tb/s optical signal processor
- 72Tx+72Rx links at 12.5Gbps
- Xilinx Virtex-7 FPGA:
  - XC7VX485T or XC7VX690T
- On-board firmware repository
- 2×144Mbit 550MHz QDR RAM
- Been in hand for nearly two years
  - Continuous testing over that period
  - Very well understood
- Extensive software and firmware
- Currently in manufacture, formally procured through CERN purchasing







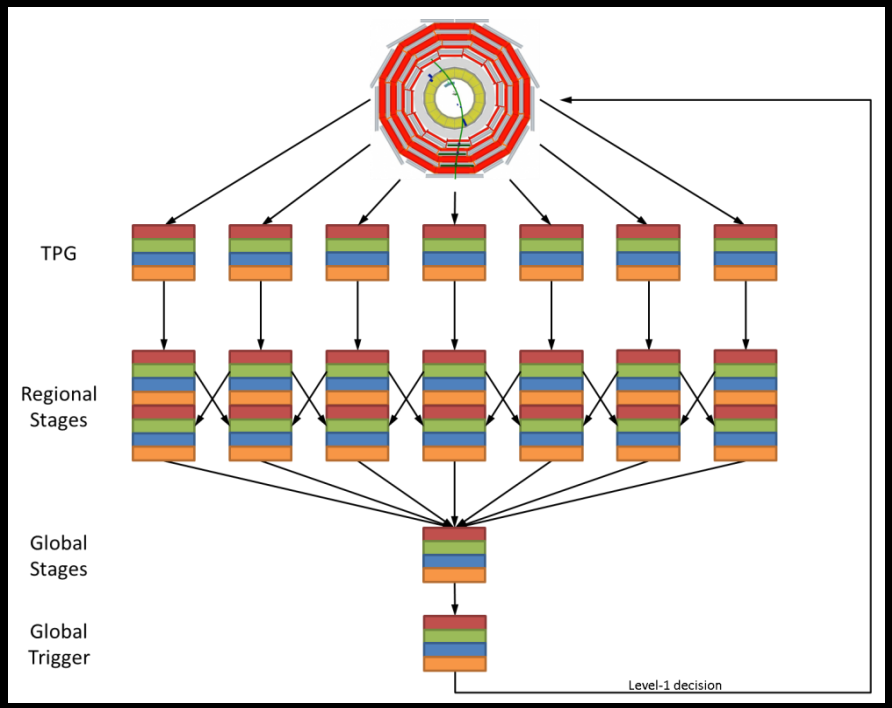
An observation…

With the recent up-turn in the global economy, many boards which have gone out for manufacturing recently are suffering long lead-times on components!

Manufacturing, testing and installation schedules need to take this into account!

Time-multiplexed triggering

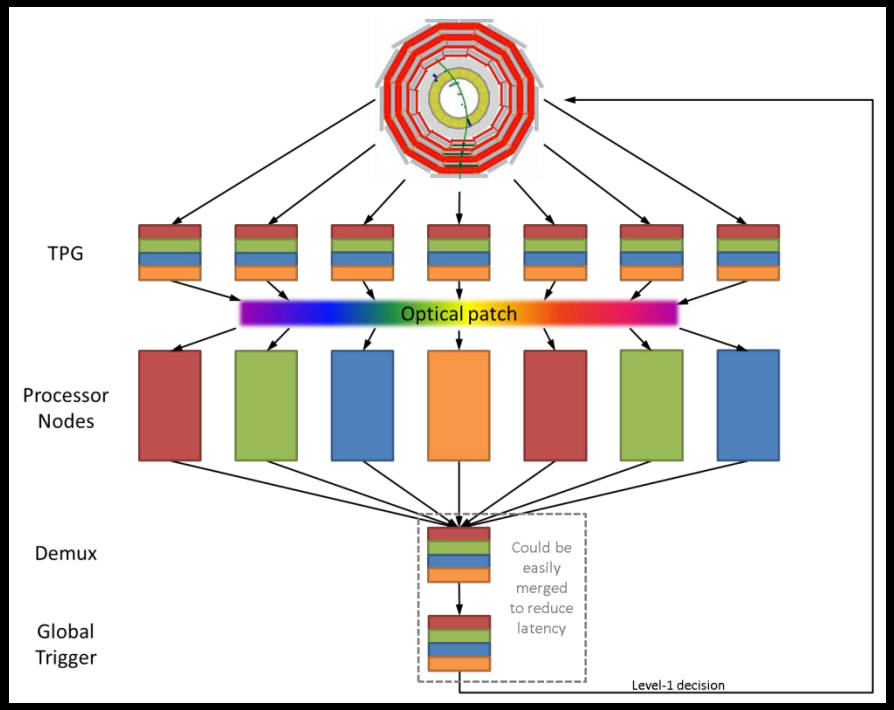
Conventional trigger



Conventional Trigger

- Data is processed in regions
- Boundaries between regions must be handled by sharing or duplicating inputs
- Volume of data reduced at each stage by selecting and discarding candidates
- When volume of data has been sufficiently reduced it can be passed to the global trigger

Time-multiplexed trigger



Time-Multiplexed Trigger

- Data from an event is buffered and retransmitted to the first processing node over *N* bunch crossings
- Data from the next event is buffered and retransmitted to the second processing node, again, over *N* bunch crossings
- Process is repeated in round-robin fashion across ≥*N* processing nodes
- Because both algorithm latency and data volume are constant, dataflow is fully deterministic and no complex scheduling mechanism is required

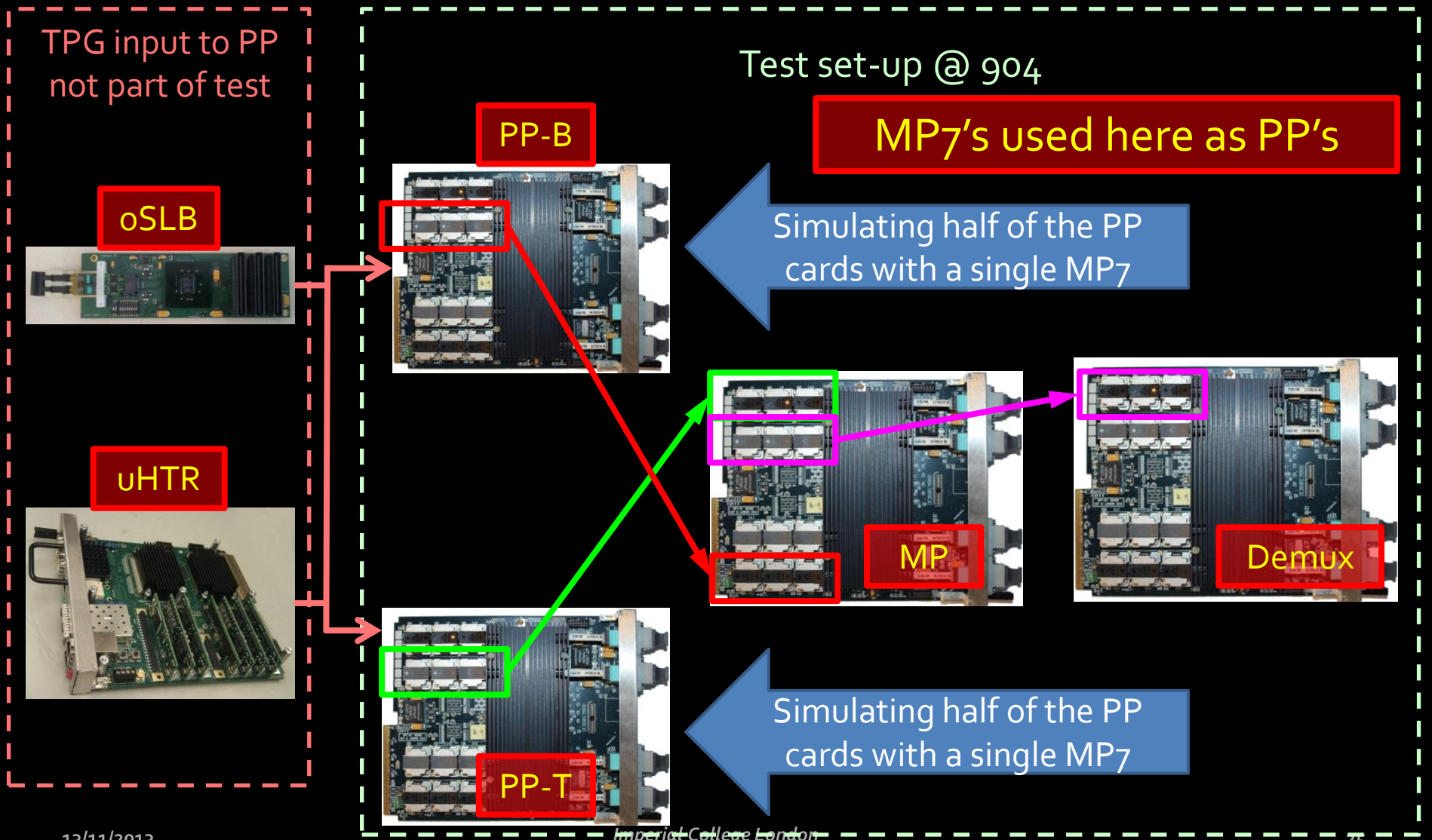
Why TMT?

- The Time-multiplexed architecture allows all data to arrive in geometric order:
  - Towers at given ϕ always occupy the same bits on the same optical links
  - Towers arrive in order of increasing $|\eta|$
- This converts a 2D geometric problem to a 1D problem
- This allows all algorithms to be fully pipelined†:
  - The processing is localised
  - Fan-outs reduced
  - Routing delays minimised
  - Register duplication eliminated
  - Also only need to develop one FPGA design

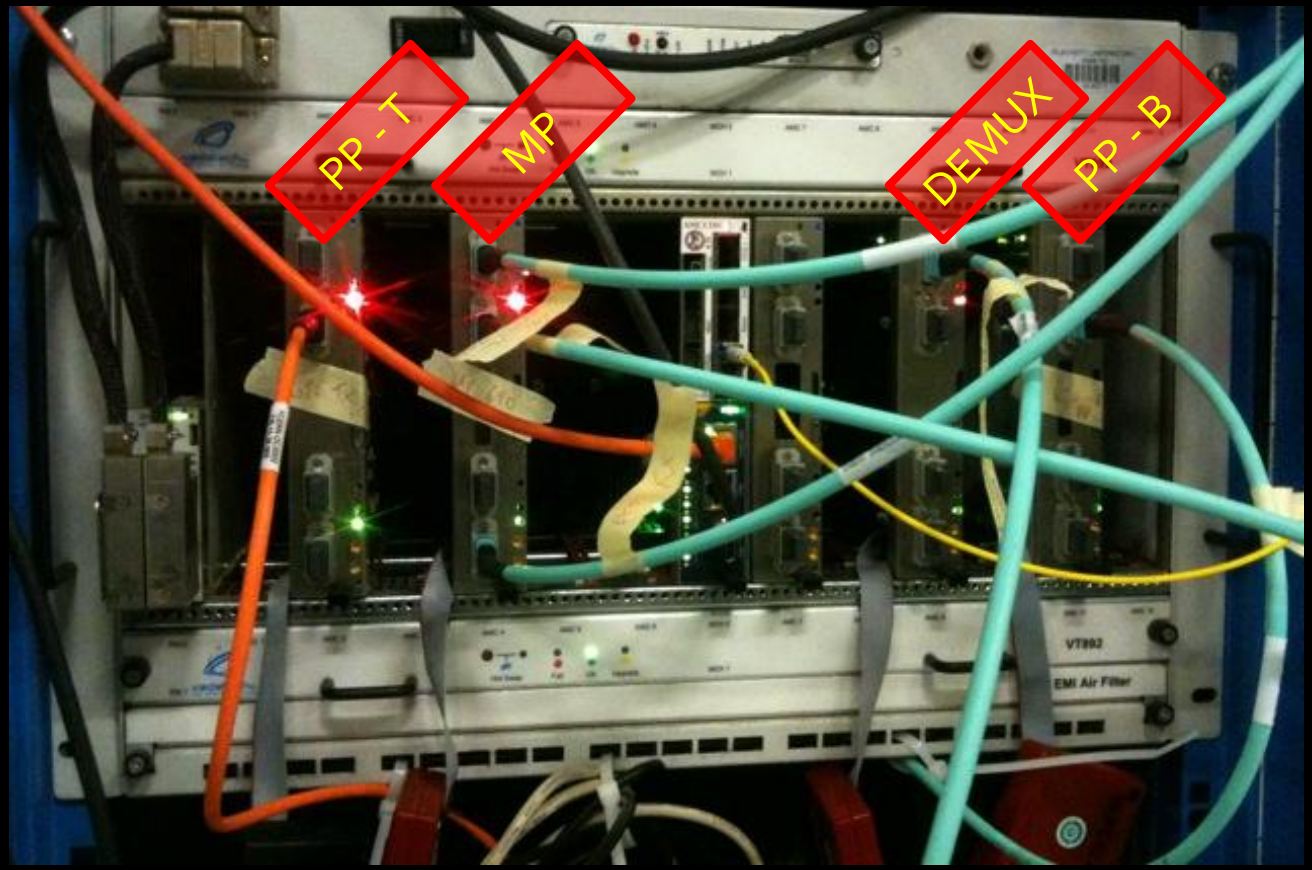
†That is, pipelined at the full data rate, not at the bunch-crossing rate

Status of the TMT concept

Status of TMT: Integration test @ B904



Status of TMT: Integration test @ B904



720 Gbps of data going through the MP board
The MP is processing data from the entire calorimeter

Status of TMT: Algorithms

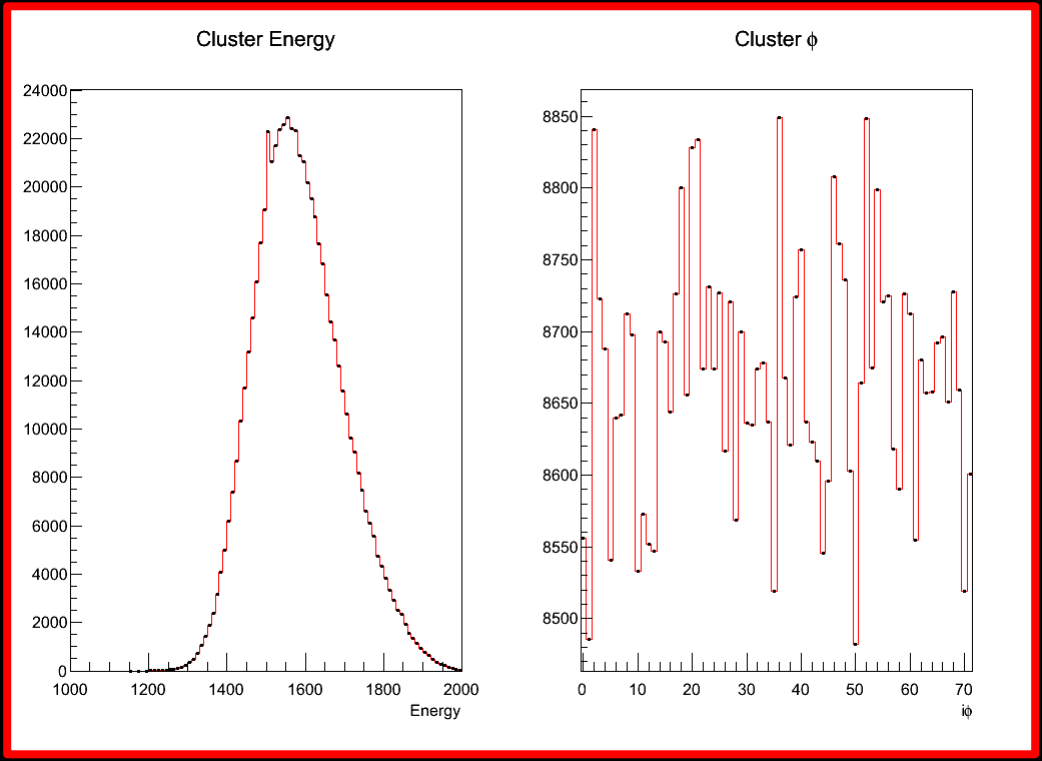
- All algorithms operate on tower-level information at 240MHz
- Compact objects – e/γ/τ
  - Use the 2×2 Wisconsin clustering algorithm
  - Baseline algorithm currently in CMSSW
- Extended objects – Jets
  - Use an 8×8 tower circular jet
  - Equivalent to cone jet (r=0.35)
  - Other algorithms available
- Global objects – Ring sums
  - Use full granularity equivalent of what is done in current trigger

Status of TMT: Results

- Random data passed through an emulator was used in the testing of the algorithms

Data injected into PP → Time-multiplexed → Optical Fibre → **Clustering algorithm (2x2)** → Sort → Capture

- Compared emulated results (solid line) with those from the MP7 (markers): C++ emulator and hardware match precisely
- System latency lower than TDR prediction
- No link-errors, no CRC-errors and maintained alignment after a week of operation



Lessons learned from test

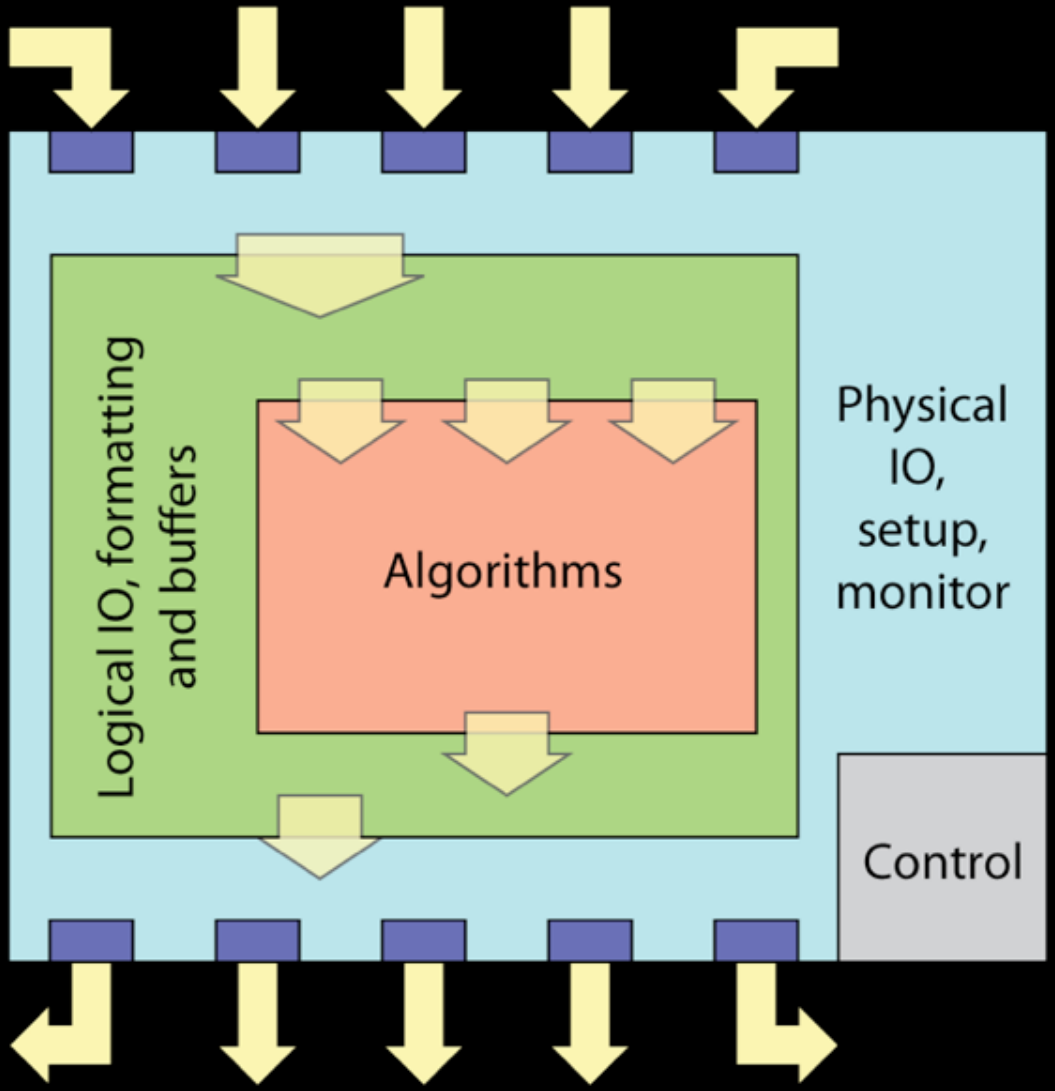
- Floor-planning
  - Huge impact on algorithm design
  - Structure the algorithm to map optimally onto the FPGA
  - Reduces risk that after many hours of routing 6 million nets just 1 or 2 fail to meet timing - exceedingly annoying
  - Significant timing improvement
  - Only viable if signals remain relatively local
- Full pipelining of algorithms is essential - even relatively innocuous looking fan-outs in chips this large have the potential to kill off the entire design
- uHAL and IPbus are working well, and have made debugging and multi-user crate access both possible and easy
- AMC13 has worked extremely well for us during the test

Firmware

Firmware

- MP7 is a completely generic optical-stream processor
- Being used in at least 5 roles in the CMS trigger upgrades
  - Interim trigger processing hardware
  - TDR trigger processing hardware
  - Drift-Tube Track-Finder
  - Global Muon Trigger
  - Global Trigger
- Having common hardware platform obviously helps with maintainability
- How to collaboratively develop firmware is not a trivial problem!

“Base Firmware” Concept



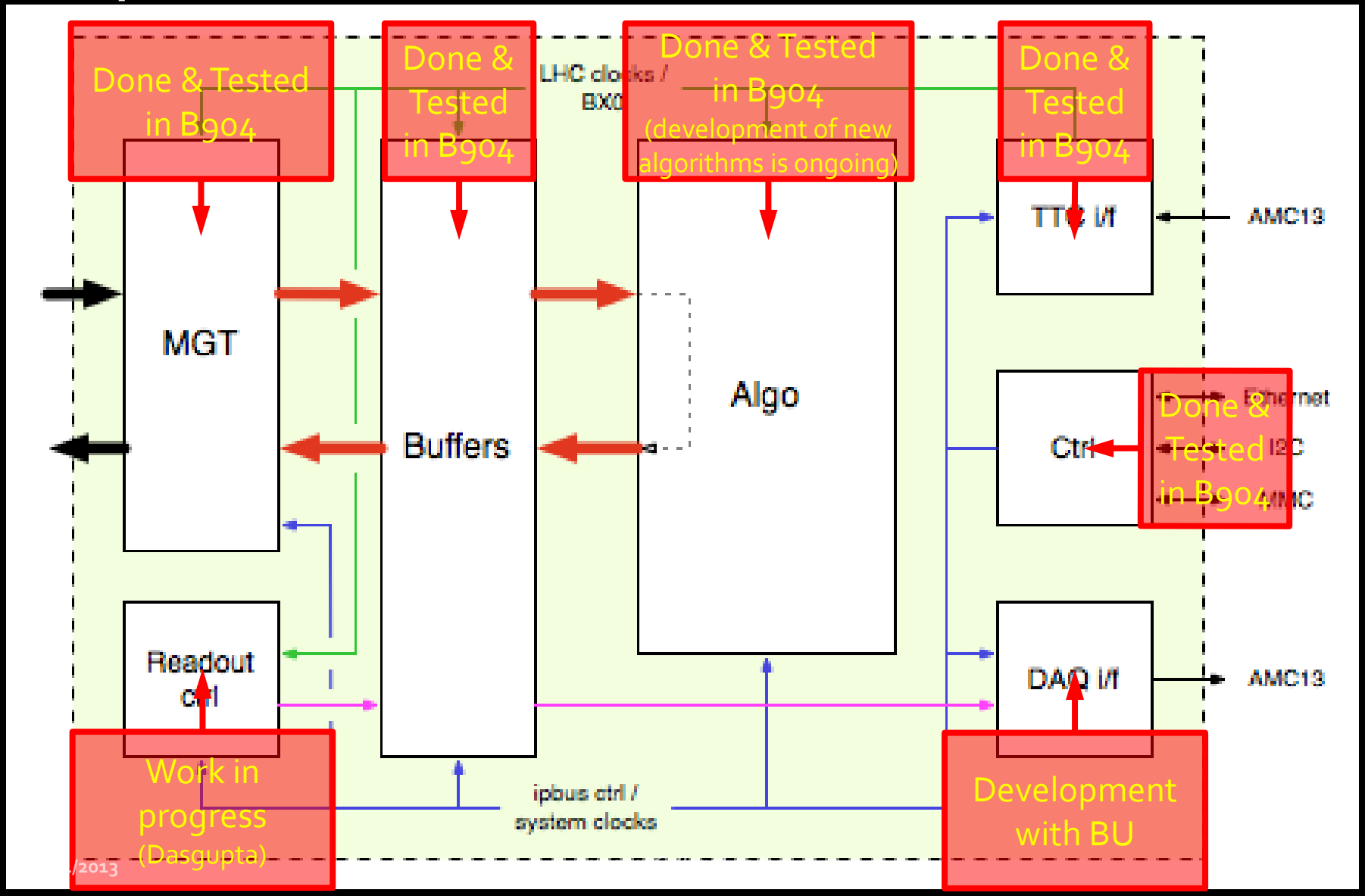
Trigger emulator
Open Development

System setup and test
Common across trigger

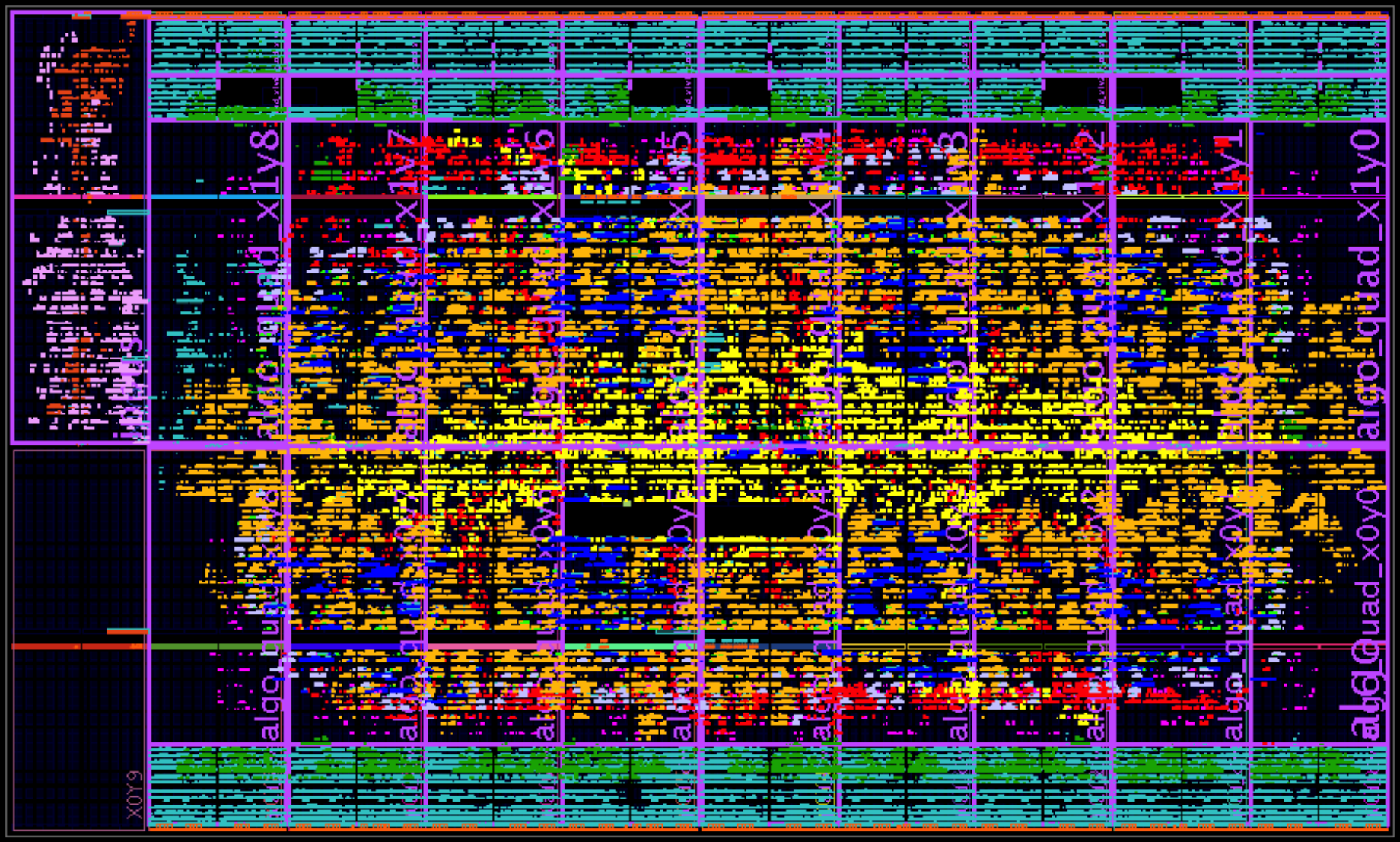
Low-level control
Hardware-specific development

uTCA infrastructure
CMS common standard

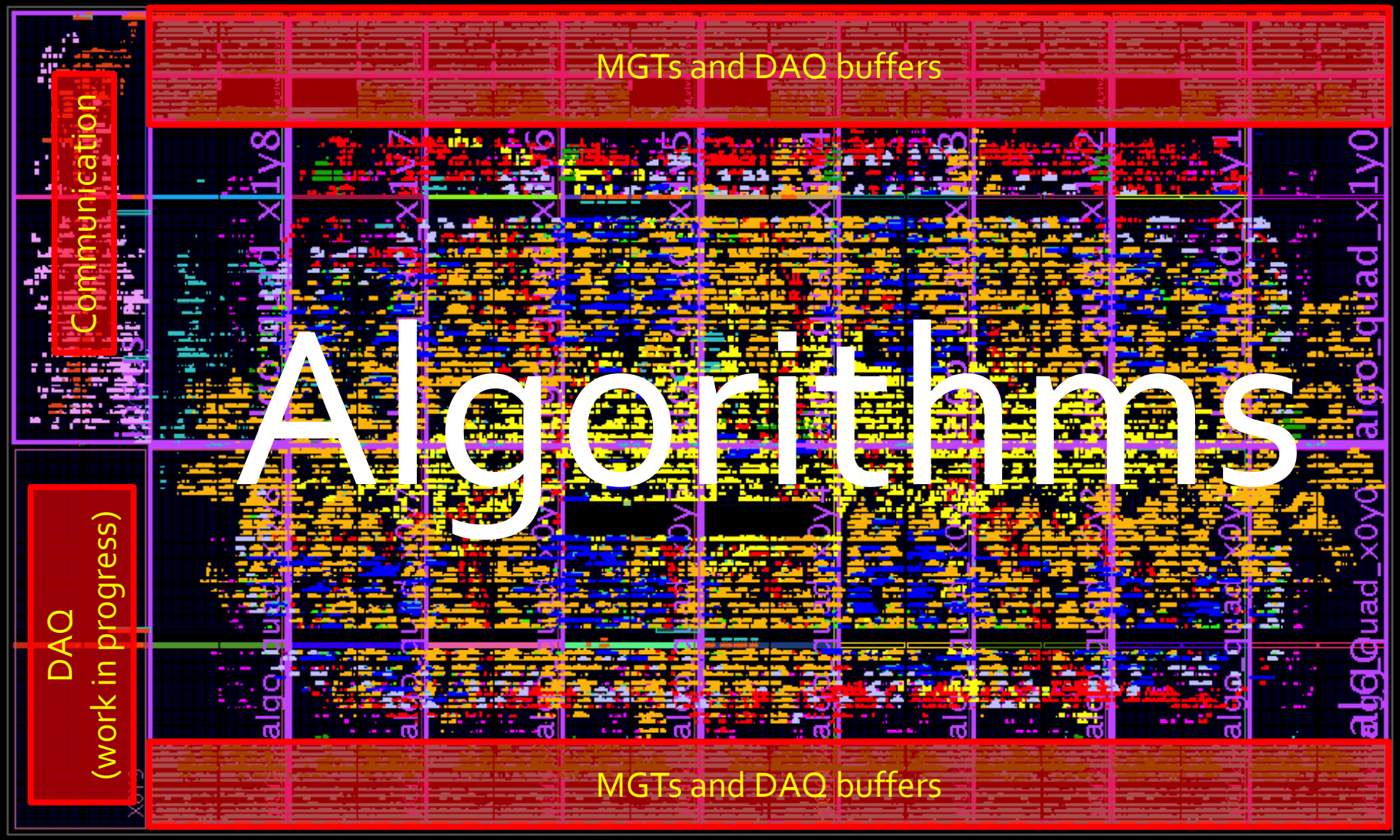
Implementation Status



Floorplan of FPGA: Integration Test



Floorplan of FPGA



Resource usage of FPGA

Resource	Fraction of Entire FPGA		Fraction of Algorithm Area used
	Used for infrastructure	Used for algorithm + infrastructure	
Registers	9%	21%	20%
LUTs	19%	36%	35%
DSPs	0%	8%	13%
BRAM	12%	12%	0%

All Algorithms and Infrastructure:

7 hour build-time

Much better than the >24 hours reported elsewhere

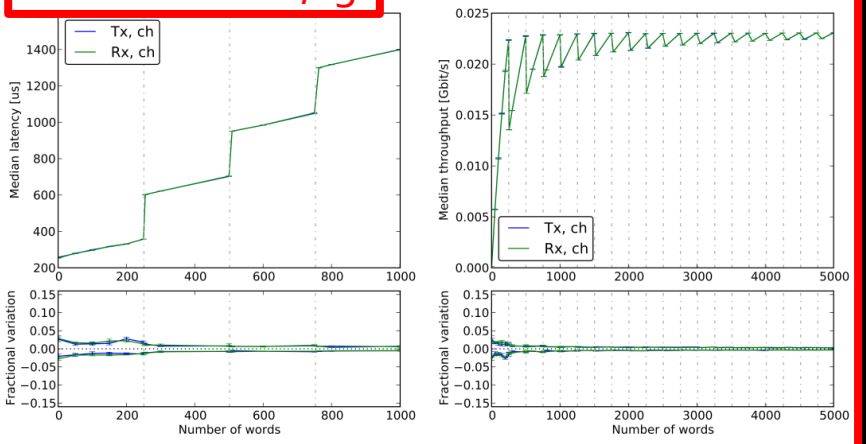
Software

IPbus/uHAL

- IPbus and uHAL, which are supported as part of the UK calorimeter trigger upgrade program, have been widely adopted outside its original scope
- A very large effort has gone into optimization & reliability over the last year:
  - Read/write bandwidths both now exceed 0.55Gbps
  - Retry mechanism works perfectly on networks with 10% packet loss. Mechanism should work identically for higher loss rates.

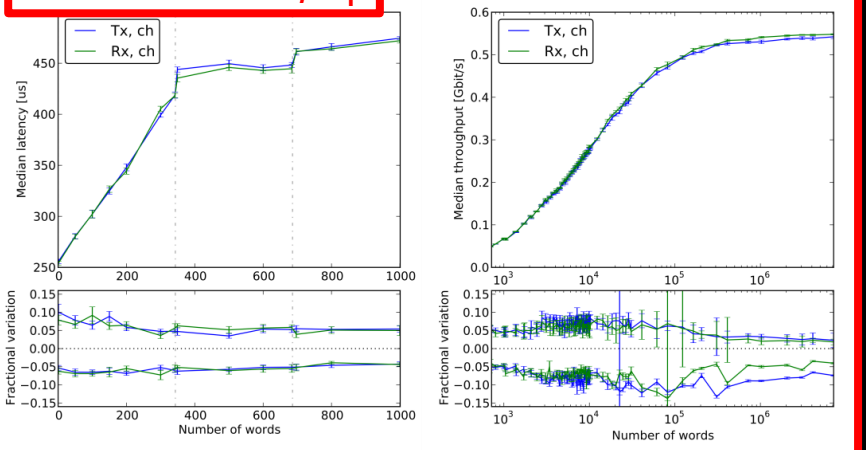
- University of Wisconsin have a preliminary TCP implementation for embedded processors (e.g. ZYNQ)
- Bandwidth tests show rates of 3-4MB/s (0.024-0.032Gbps)

IPbus 2.0 – 01/13



~2.5 FTEs

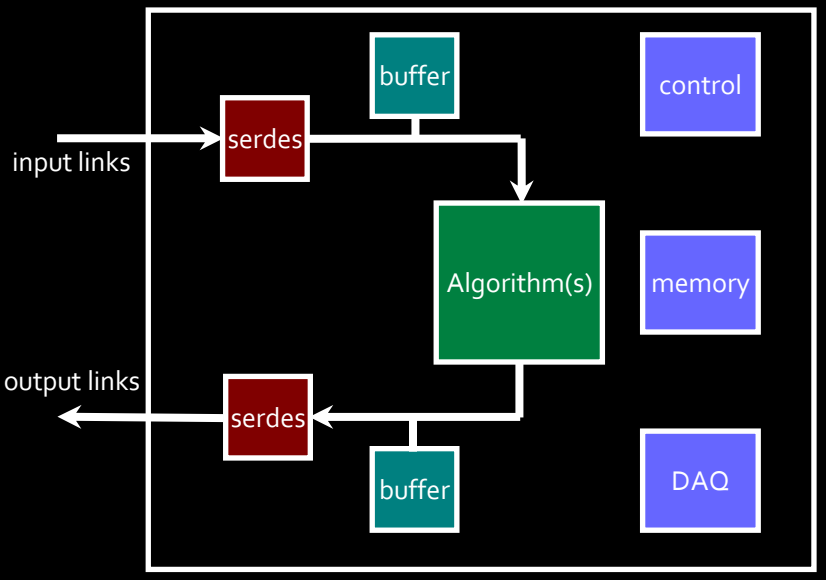
IPbus 2.2 – 01/14



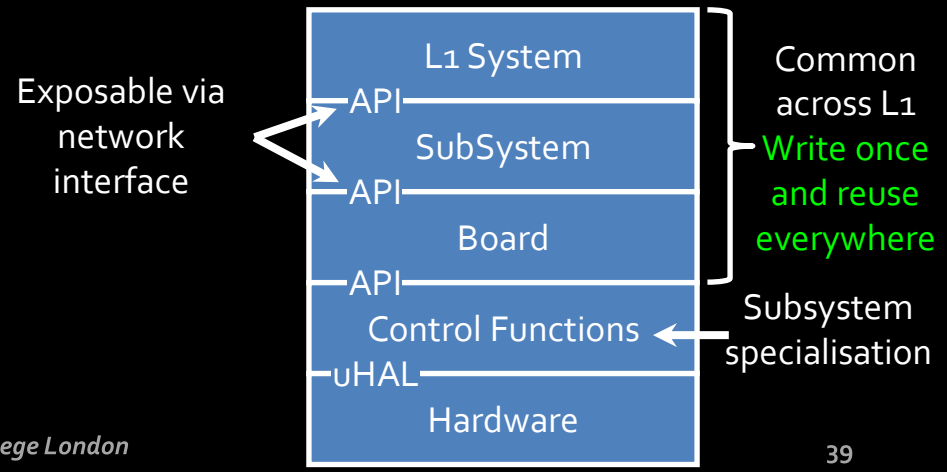
Plots courtesy of T. Williams

SoftWare for Automating conTrol of Common Hardware

- IPbus and uHAL have rationalized hardware access
- Can we rationalize board control S/W?
  - MP7, CTP7, MFT7 are all very similar
  - One software class to control them all?
- Can we use the same software model for production and testing?
  - Standardize the concept of 'test' (what goes in, what goes out)
  - Introduce flexible system-wise tests: from {system A, buffer X} to {system B, buffer Y}
- Requires rationalizing DAQ, data-packing and data-unpacking!



Above: Common Board Model, Below: System Model



Conclusions

Conclusions

- All hardware for the CMS Calorimeter Trigger upgrade now in hand, albeit having undergone varying degrees of testing:
  - TDR upgrade: μHTR, oSLB, oRM, CTP7, MP7
  - Interim trigger for 2015: oRSC, MP7
- The baseline architecture is a Time-Multiplexed Trigger
- The Time-Multiplexed Trigger Integration test at B904 last September was a great success
- Many lessons were learned which reaffirmed the motivations for choosing the TMT architecture over a conventional architecture
- Have developed a model for collaborative development of firmware
- uHAL and IPbus v2 work very well and are being widely adopted outside their original scope