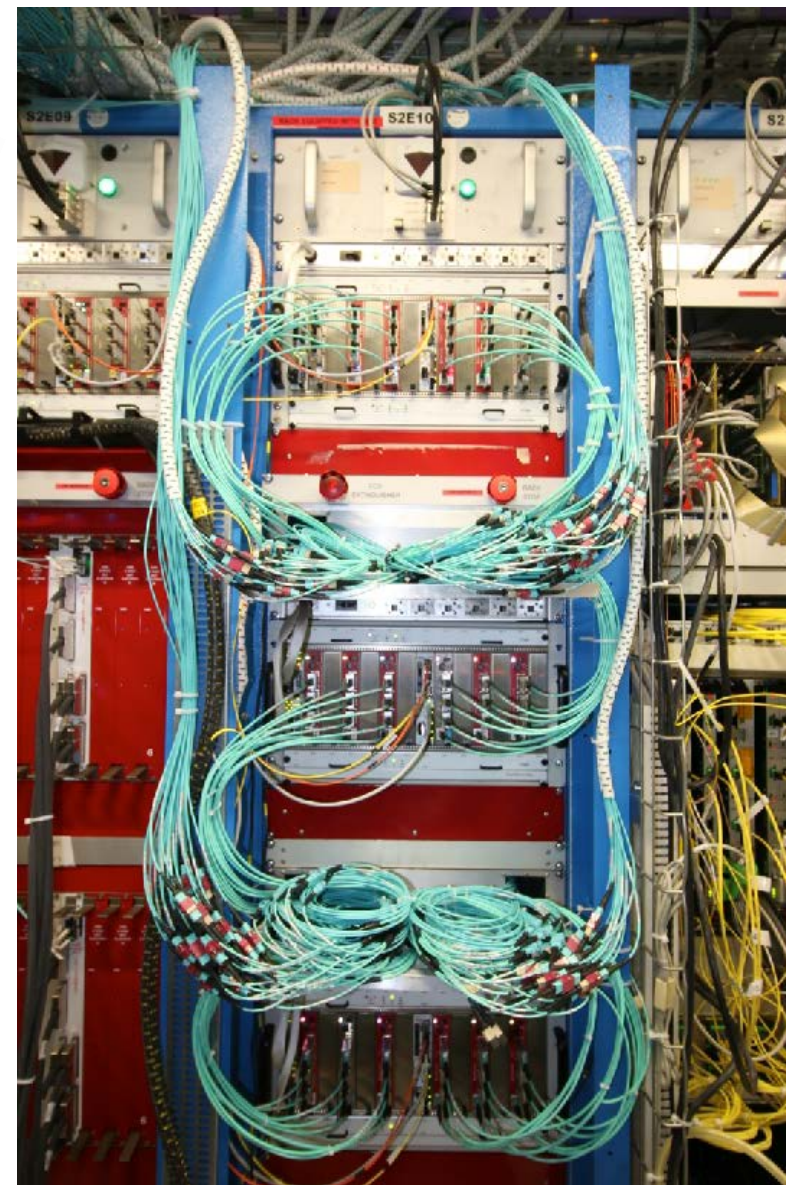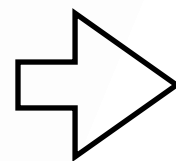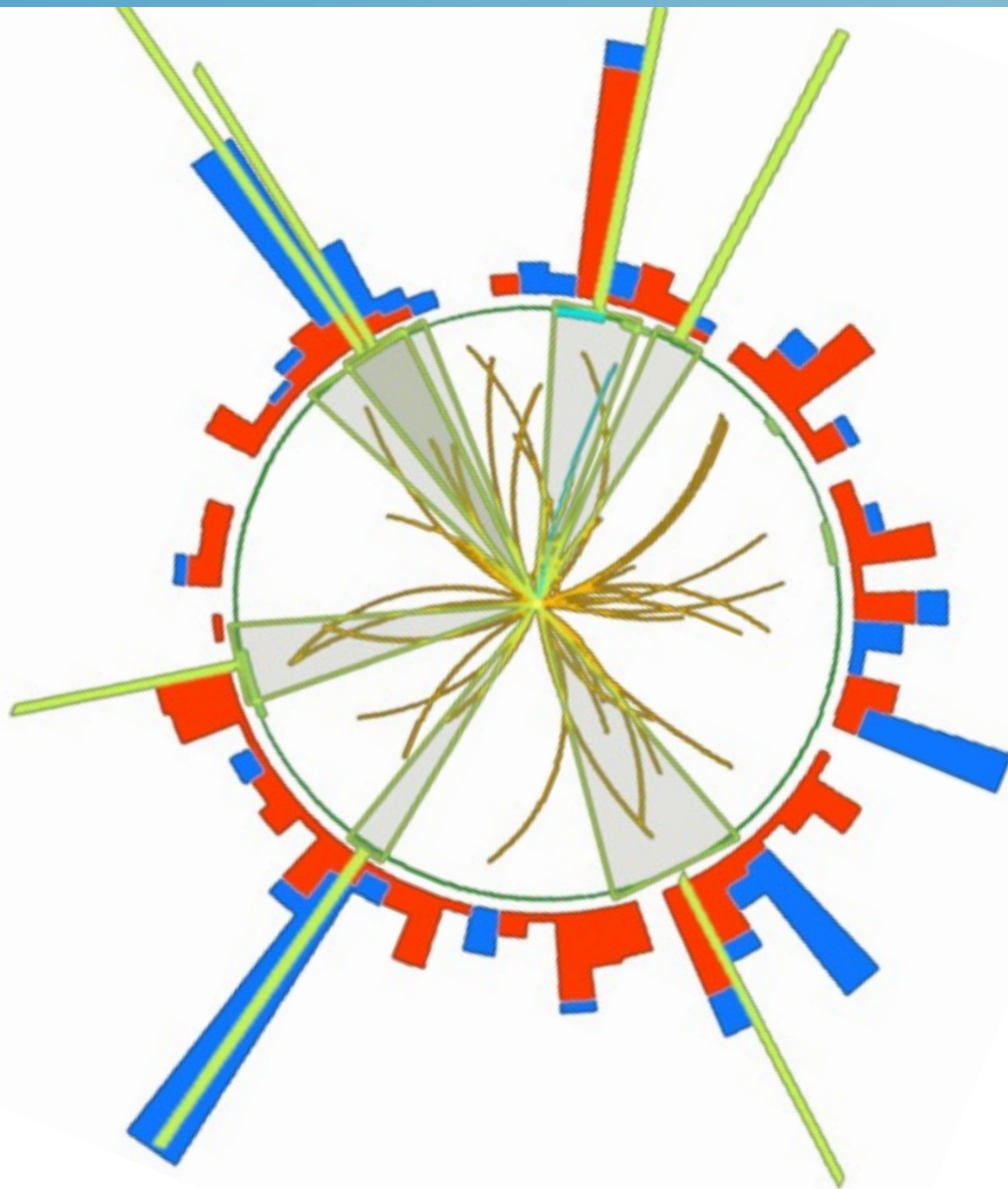# CMS Phase-1 Trigger

Phase-1 Layer-1 Calo Trigger (U. Wisconsin)

Dave Newbold

On behalf of the CMS L1 Trigger Group

Overview

- ▸ **A successful Run 1 for CMS**
  - ▸ Not least due to many years of R&D, construction, commissioning of L1 trigger system
  - ▸ One of the tougher aspects of the original design
  - ▸ 'Cutting edge' technology in mid-2000s
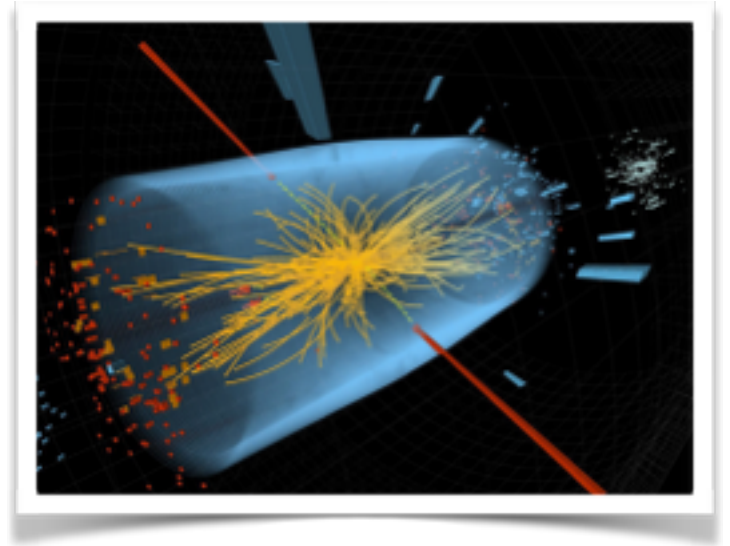
- ▸ **Decision taken to ~completely replace the system in 2013-15**
  - ▸ In parallel with major changes to timing, DAQ front end

- ▸ **In this talk**
  - ▸ Focus on technical developments
    - ▸ For algorithms and physics performance, see past and upcoming CMS conference talks
    - ▸ e.g. *The CMS Level-1 Trigger for the LHC RUN-II*, C. Foudas, EPS-HEP 2015
  - ▸ Why is the trigger challenging to build and upgrade?
  - ▸ What were the key technologies?
  - ▸ What worked, and what didn't?
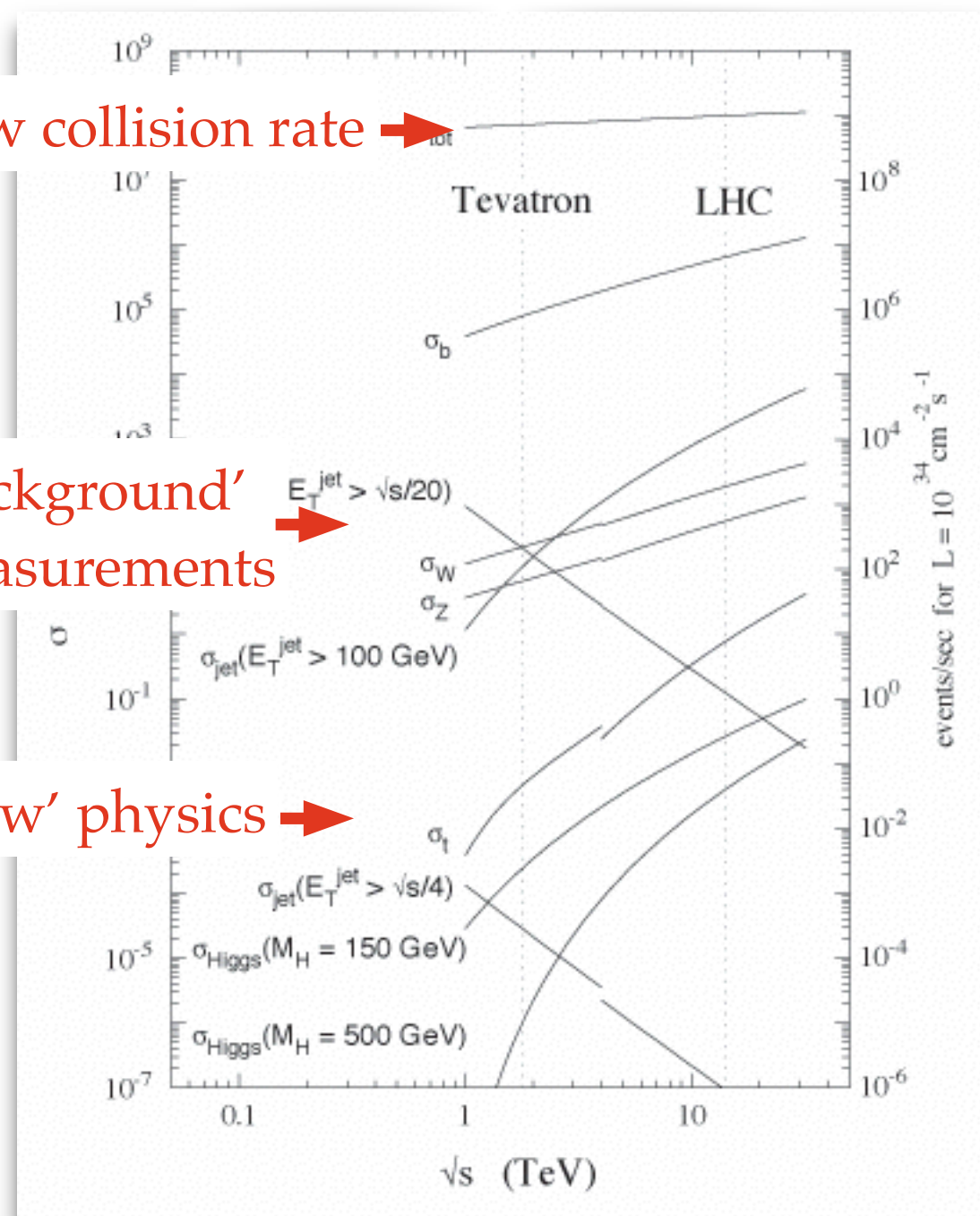 - ▸ What next?

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Raw collision rate ➡

'Background' measurements ➡

'New' physics ➡

(Figure: cross-section vs √s (TeV) for Tevatron and LHC, showing σ_b, $E_T^{jet} > \sqrt{s}/20$, σ_W, σ_Z, $\sigma_{jet}(E_T^{jet} > 100\ GeV)$, σ_t, $\sigma_{jet}(E_T^{jet} > \sqrt{s}/4)$, $\sigma_{Higgs}(M_H = 150\ GeV)$, $\sigma_{Higgs}(M_H = 500\ GeV)$; right axis events/sec for $L = 10^{34}\ cm^{-2}\ s^{-1}$)

‣ **Raw CMS data rate:**
  ‣ 40MHz @ >1MB per event

‣ **L1 trigger must:**
  ‣ Select collisions of interest at rate of O(100kHz)
  ‣ Make decisions within limited latency: O(3μs)
  ‣ Work on a limited subset of detector data

‣ **After this…**
  ‣ Software-based High Level Trigger further reduces rate to O(100Hz)

‣ **Theme of this talk:**
  ‣ Clever algorithms from HLT → L1

Requirements for L1 Trigger

‣ **Operational:**

    ‣ Guarantee a hard limit on the data rate from the detector

    ‣ Provide negligible dead time

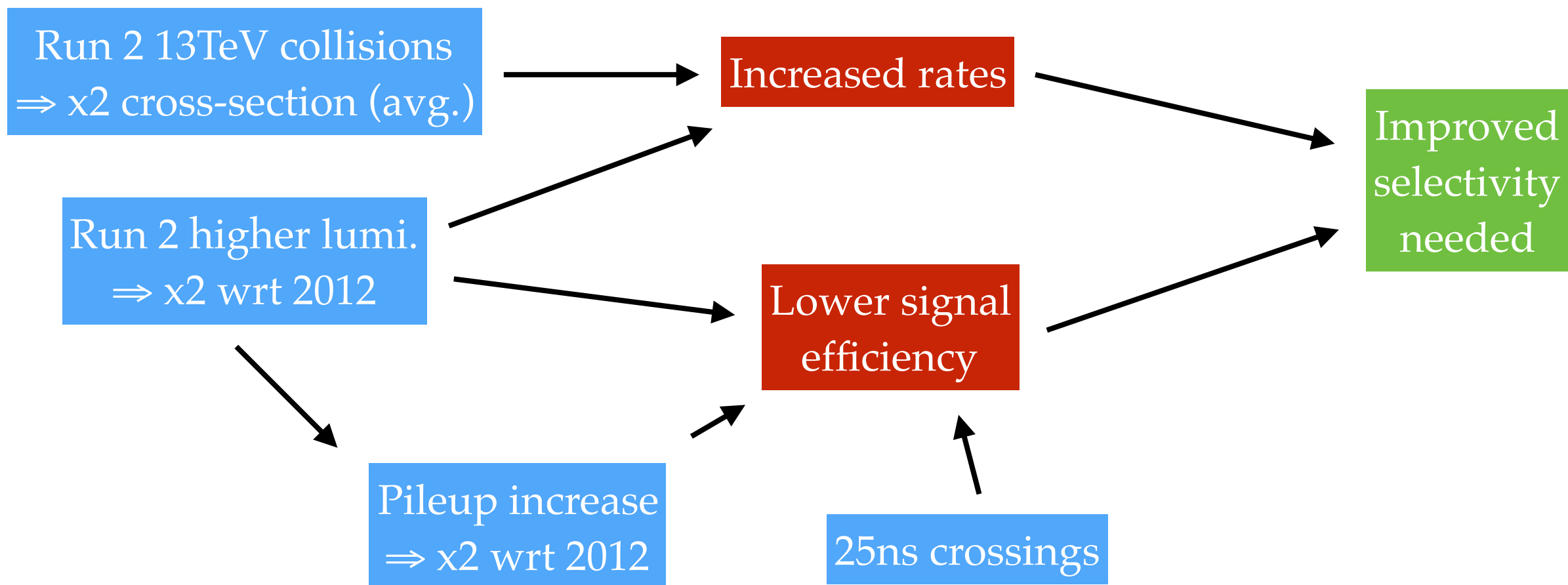   ‣ Provide robust 'handles' for controlling rate in presence of background

‣ **Physics:**

    ‣ Trigger efficiency must be unbiased, measurable, reproducible

    ‣ In practice: provide handles to measure efficiency / purity from data

‣ **Technical:**

    ‣ Extreme reliability – without L1 trigger, there is no data taken

        ‣ This includes rapid ('instant') detection of faults – can lead to biased trigger and useless data

    ‣ Extreme flexibility – changing machine conditions and physics priorities require new selection algorithms, sometimes new data flows

    ‣ Extreme performance – meeting operational requirements means processing $O(10\text{Tb/s})$ of data in real time

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Run 2 13TeV collisions ⇒ x2 cross-section (avg.)

Run 2 higher lumi. ⇒ x2 wrt 2012

Pileup increase ⇒ x2 wrt 2012

25ns crossings

Increased rates

Lower signal efficiency

Improved selectivity needed

‣ From technical point of view

   ‣ Vastly improved processing capacity per $ in modern devices

      ‣ Allows also for future flexibility beyond LS2

   ‣ Substitution of copper cables with robust optical transmission

   ‣ Replacement of ageing electronics & removal from expt. cavern

‣ Also an opportunity to bring in the new generation of experts

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Upgrade Strategy

▸ **Challenges for Run 2:**

 ▸ Pileup reduces the effectiveness of simple threshold-based algorithms

 ▸ Muon pt mis-measurement causes rate blowup

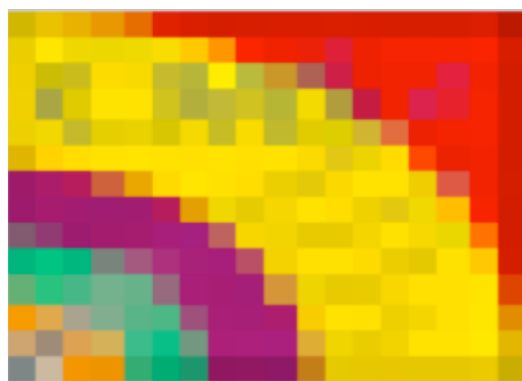 ▸ Position / energy resln. at global trigger limits final decision performance

▸ **Upgrade strategy:**

 ▸ Increase resolution of detector information entering trigger

  ▸ Substantially increased data flow within the system

 ▸ Use higher granularity to select on local cluster shape for e/g, tau

  ▸ Increased algorithm complexity and gate count

 ▸ Perform on-the-fly pileup subtraction for calo objects

  ▸ Increased algorithm complexity and gate count; data locality issues

 ▸ Combine muon system information at the earliest possible stage

  ▸ Complete re-working of data flow in muon trigger system

 ▸ Improve muon track-finding algorithms, including in 'overlap' region

  ▸ Large LUTs required, increased algorithm complexity

 ▸ Increase number and complexity of GT selection
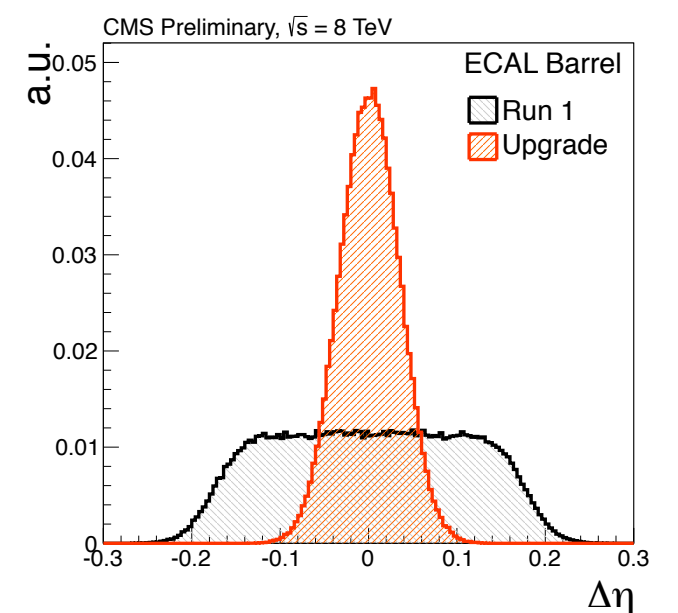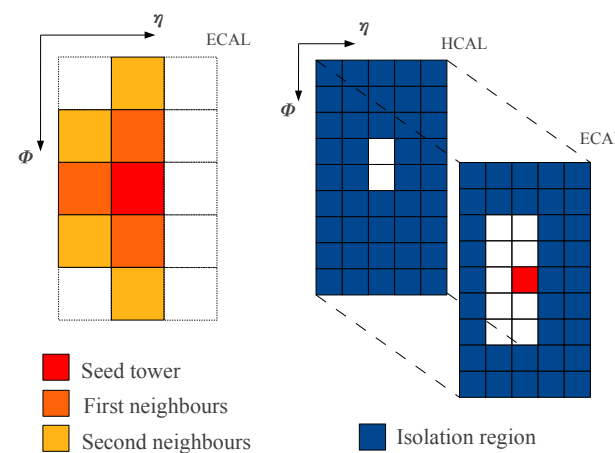
  ▸ Increased data flow to GT

Calorimetry

Muons

Global trigger

Roadmap
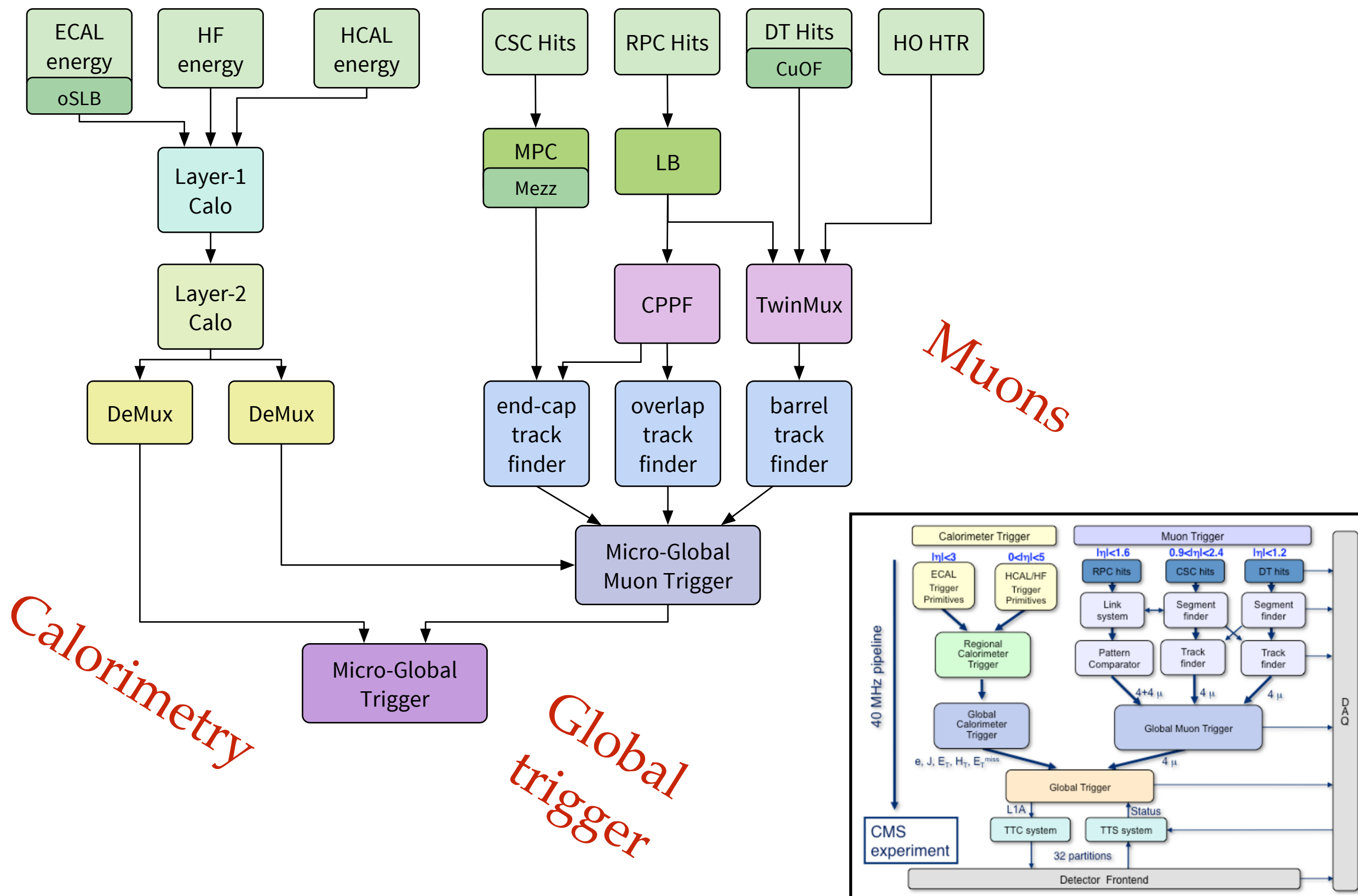
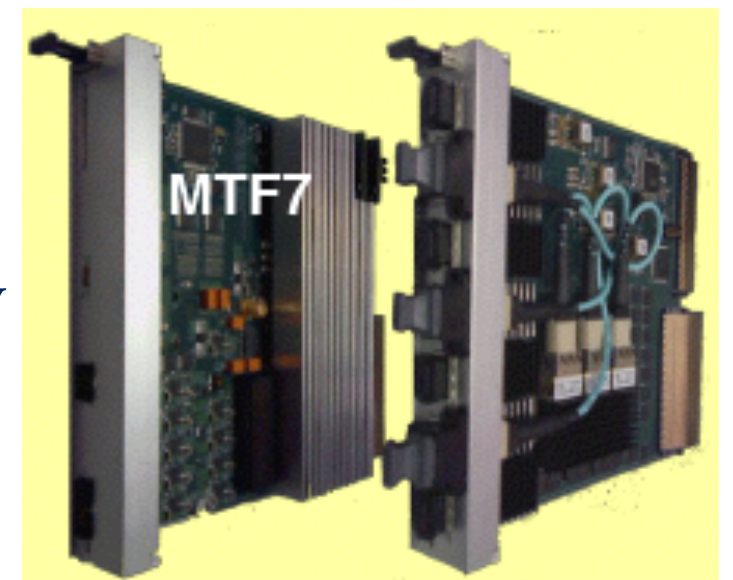|	Run 1	Run 2	Run 3	Phase-2
ECAL / HCAL granularity	Regions / Regions	Towers / Towers	Towers / Towers	Crystals / Towers
Detector information	Calo + muon	Enhanced calo / unganged muons	+ additional muon coverage	+ inner tracking
L1 Trigger rate	100kHz	100kHz	?	1MHz
GT algorithms	Cut and count + topological	+ Invariant mass	?	Particle flow, track isolation

oSLB
oRM
oRSC

HCAL Optical Splitters

→ Copper
→ Optical

HCAL Energy HTR

ECAL Energy TCC

HF Energy μHTR

HCAL Energy mHTR

Regional Calo Trigger

EM Candidates

Region Energies

Entire Summary

Global Calo Trigger

Layer 1 Calo Trigger — CTP

Layer 2 Calo Trigger — MP7

Region Granularity

Trigger Tower Granularity

Seed tower
First neighbours
Second neighbours

Isolation region

ECAL

HCAL

ECAL

CMS Preliminary, √s = 8 TeV

ECAL Barrel
Run 1
Upgrade

Science & Technology Facilities Council
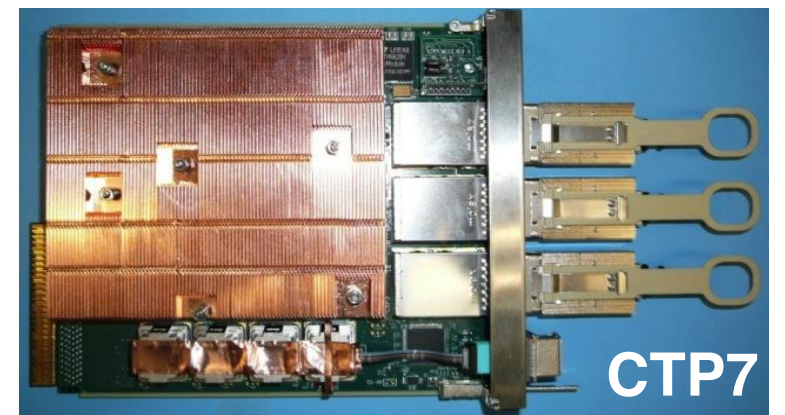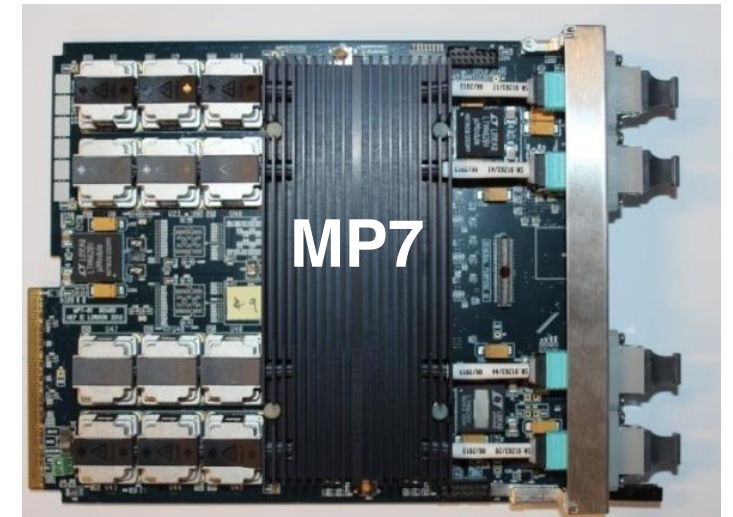Rutherford Appleton Laboratory
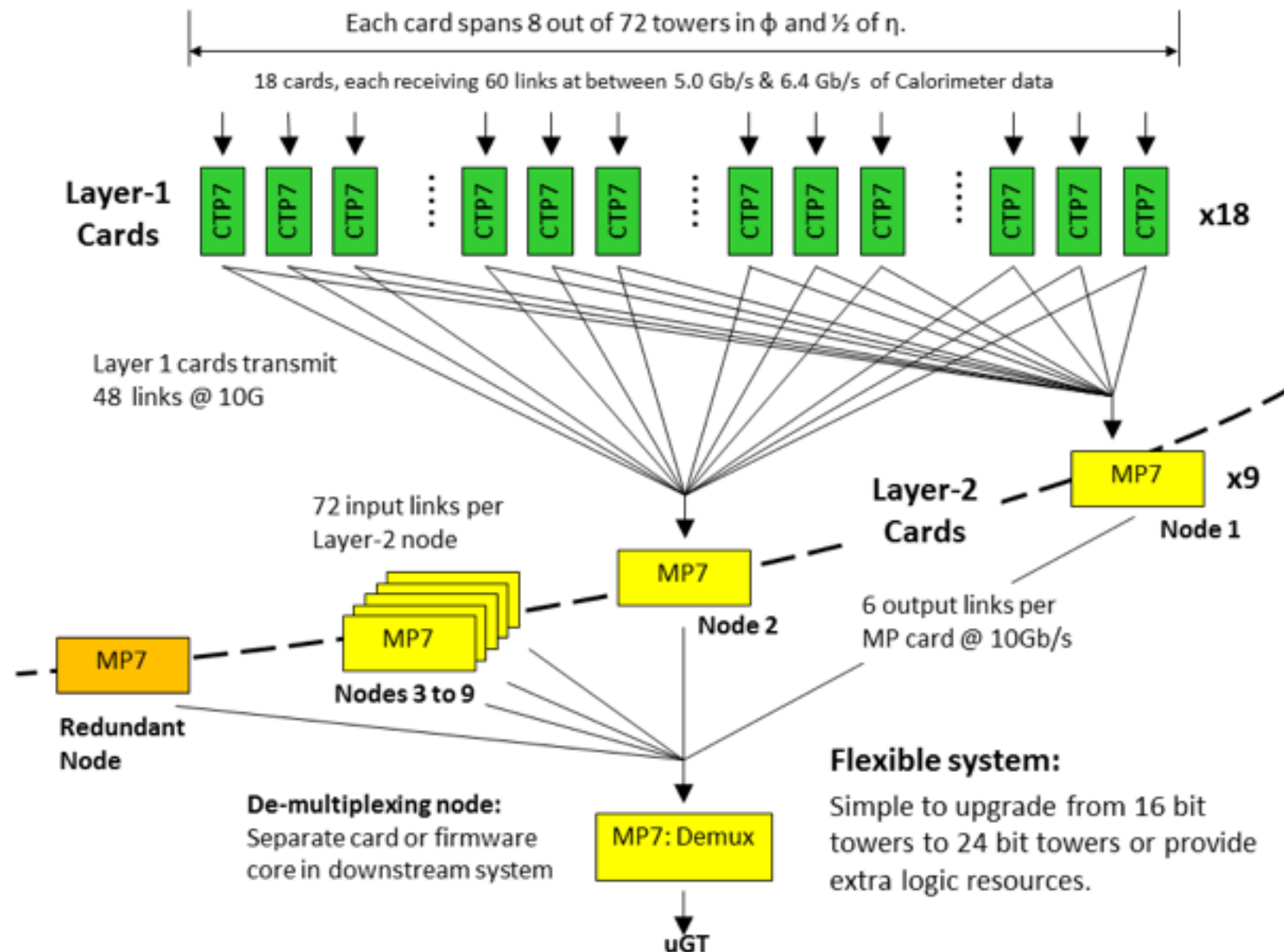
University of BRISTOL

Hardware Processor Platforms

MP7

- ## MP7 (calo Layer-2, BMTF, GMT, GT)
  - 144Tx/Rx 10Gb/s optical links
  - V7 690 FPGA

- ## CTP7 (calo Layer-1)
  - 67Tx, 48Rx 10Gb/s optical links, backplane IO
  - V7 690 FGPA

CTP7

- ## MTF7 (Endcap, overlap track finders)
  - Large input IO (84 Rx 10Gb/s links)
  - Large 1GB LUT in external RAMs

- ## All boards in microTCA format
  - Common interface to DAQ, timing, etc
  - Modular design with optical IO for max. flexibility
 - microTCA telecoms format chosen to give access to commercial infrastructure components

MTF7

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Each card spans 8 out of 72 towers in φ and ½ of η.

18 cards, each receiving 60 links at between 5.0 Gb/s & 6.4 Gb/s of Calorimeter data

Layer-1 Cards — CTP7 × 18

Layer 1 cards transmit 48 links @ 10G

72 input links per Layer-2 node

Layer-2 Cards — MP7 × 9 — Node 1

6 output links per MP card @ 10Gb/s

MP7 — Node 2

MP7 — Nodes 3 to 9

MP7 — Redundant Node

De-multiplexing node:
Separate card or firmware core in downstream system

MP7: Demux

uGT

Flexible system:
Simple to upgrade from 16 bit towers to 24 bit towers or provide extra logic resources.
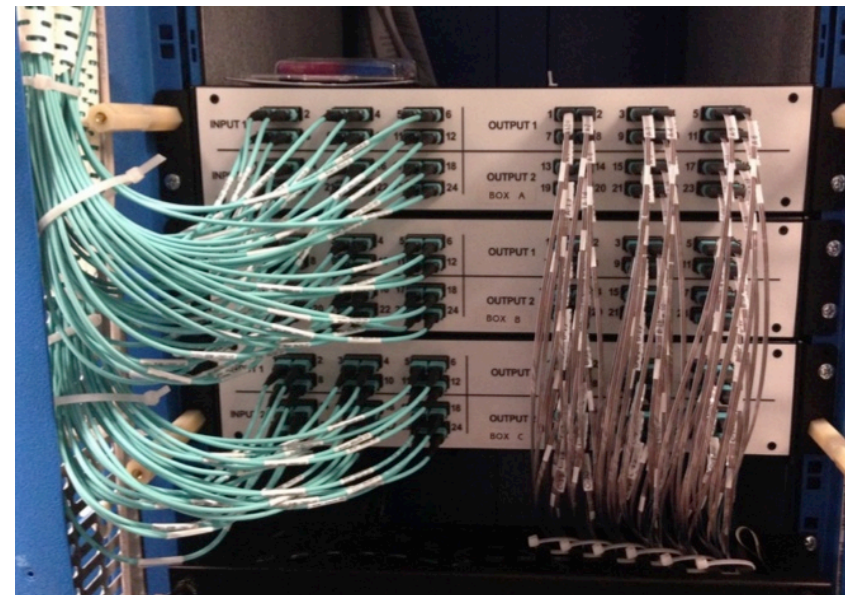
‣ One processing FPGA sees the entire detector for one event
  ‣ Advantages: 'seamless' coverage of detector; optimum use of logic elements; redundant nodes for testing and fail-over
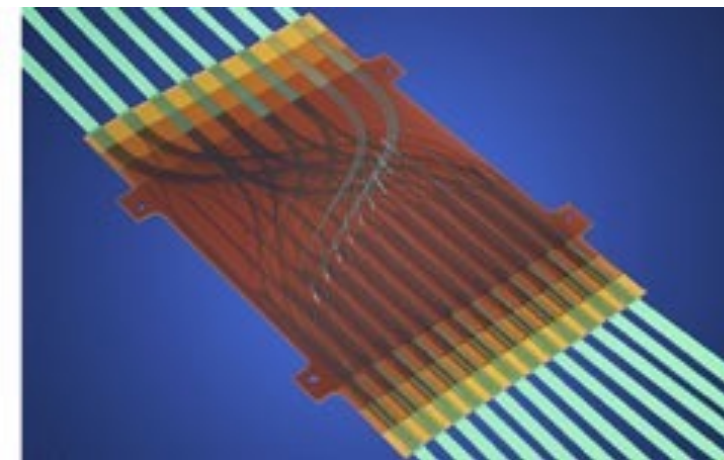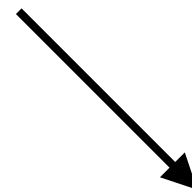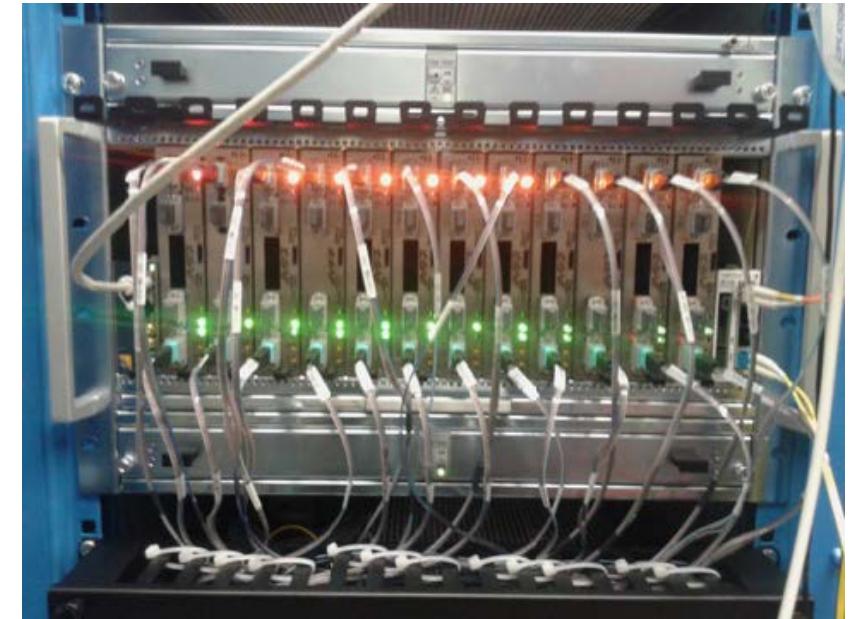 ‣ Disadvantages: large many-to-many optical IO system; large IO per node; demultiplexing stage required
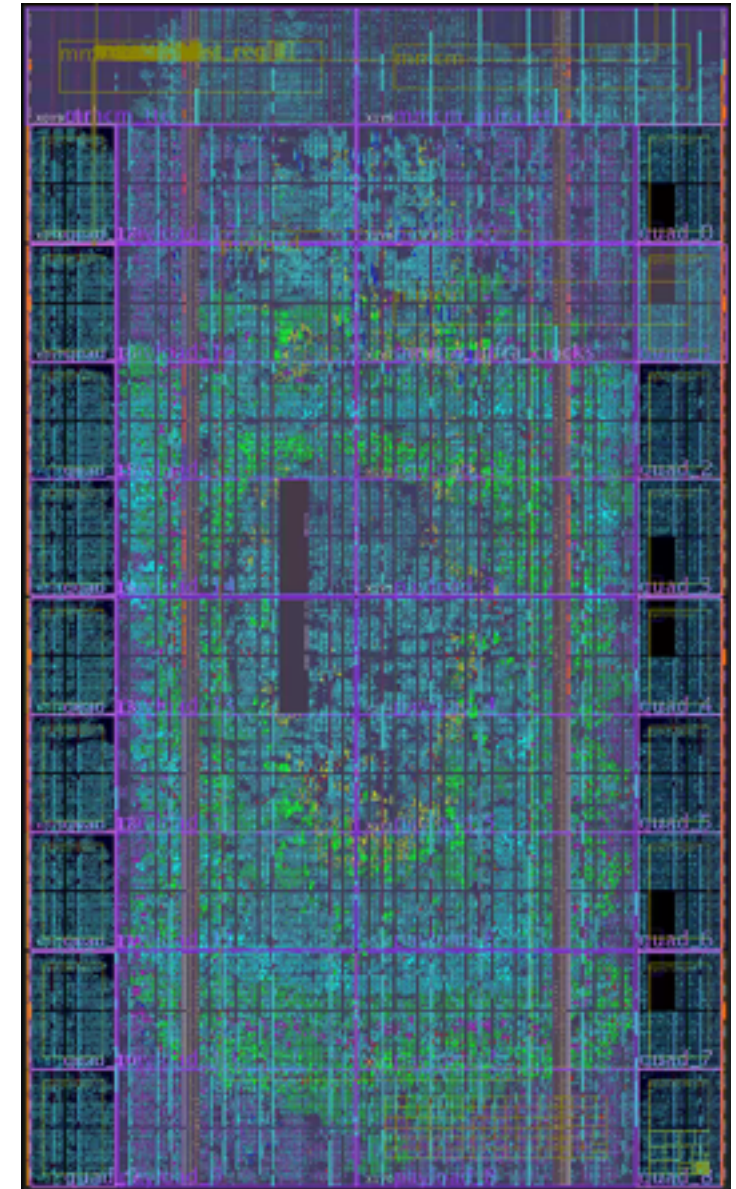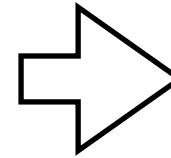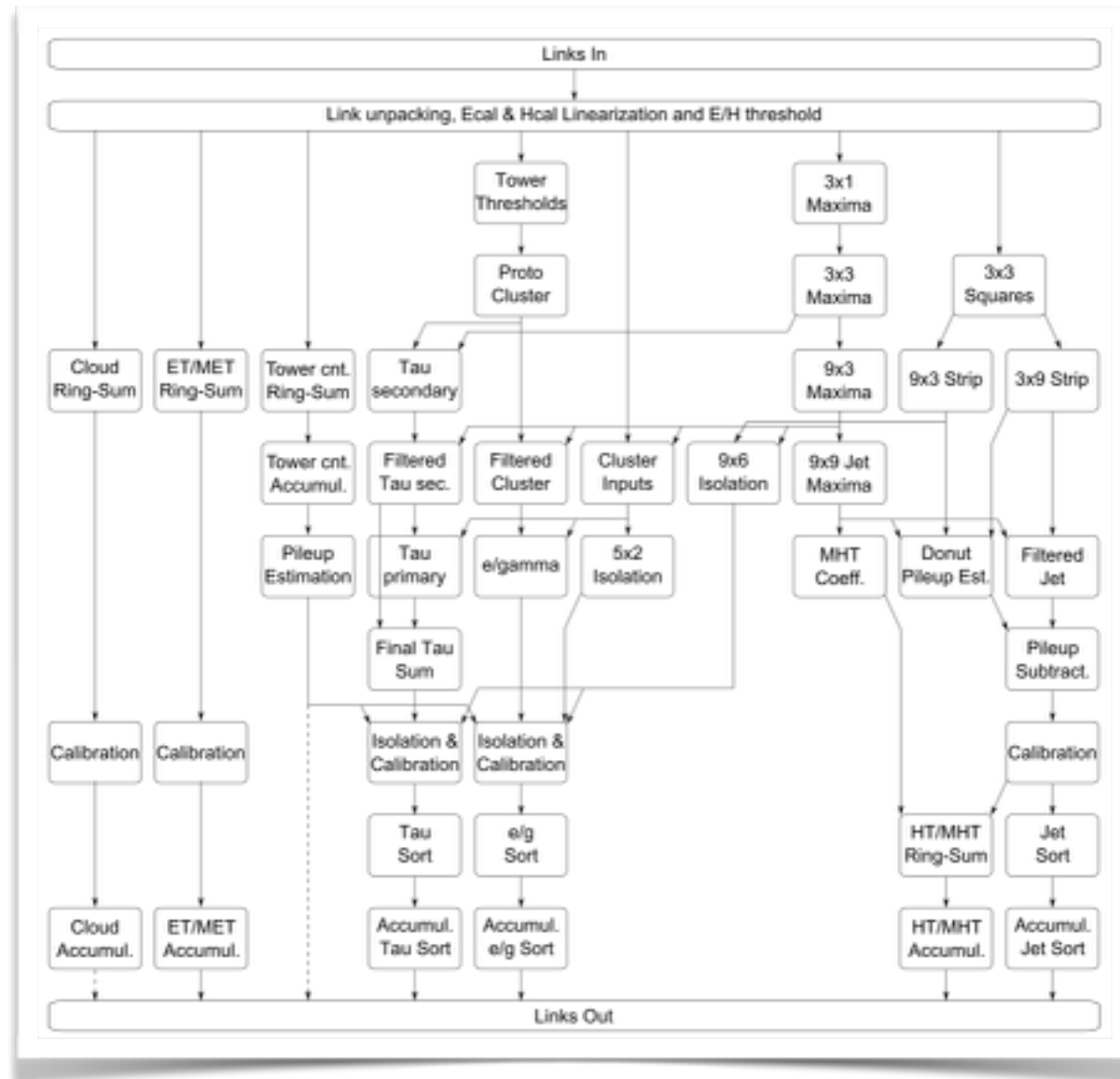
Optical IO
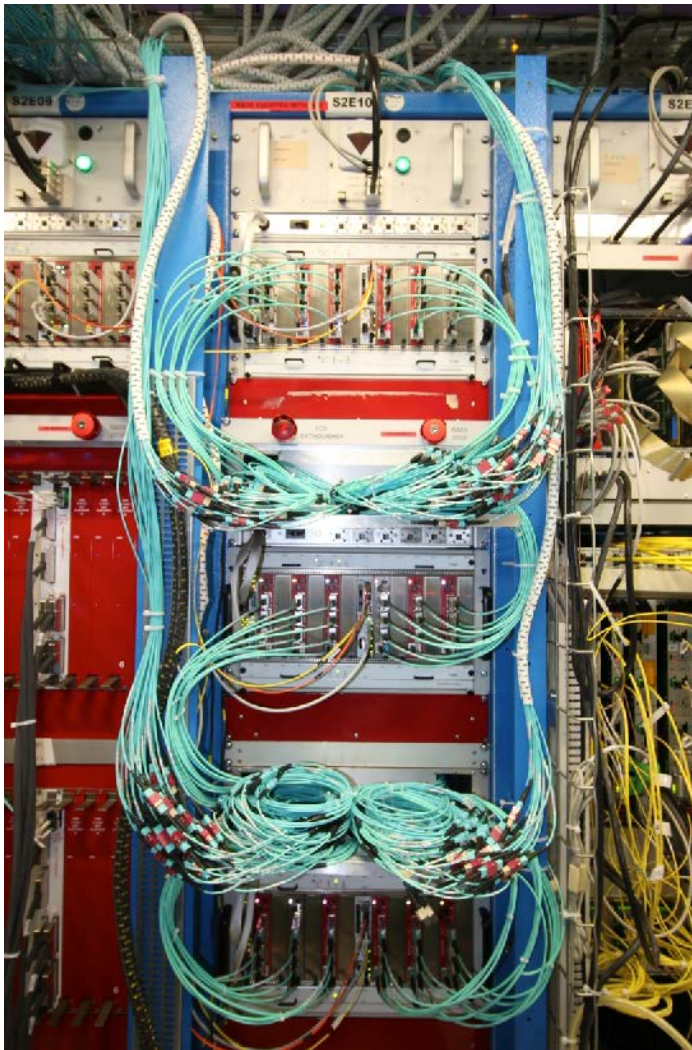
Calo Layer-1

Optical Multiplexer

Calo Layer-2

Molex Flexplane interconnect

‣ 864 x 864 10Gb/s optical patch panel reduced from 56U to 6U

‣ Optical links running custom packet protocol, async. to LHC clock
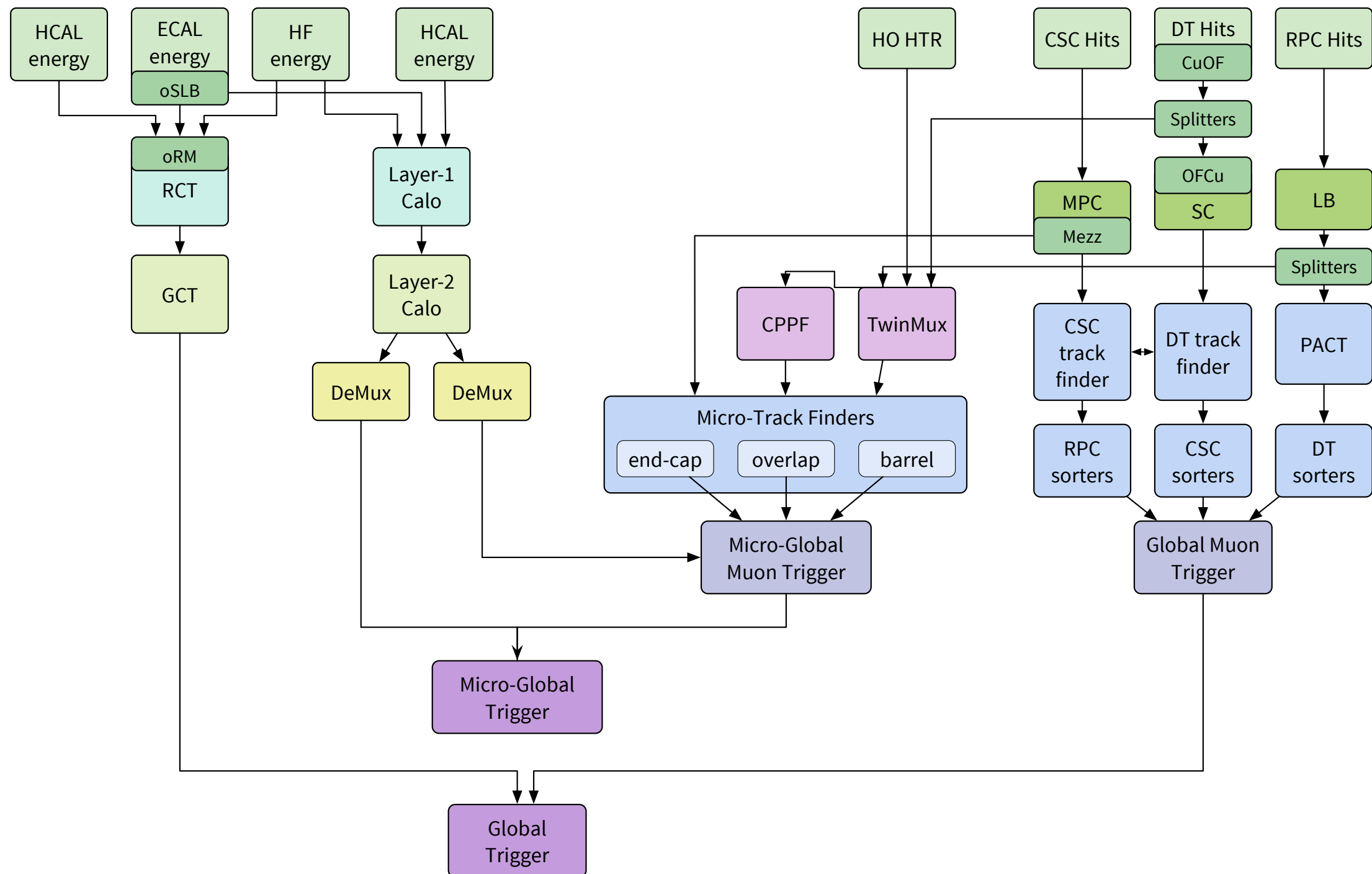
Algorithms in Practice

Virtex-7 690T, ~70% occupancy

- ‣ Billion-transistor firmware designs now the norm
  - ‣ Code management of 50k line VHDL is a non-trivial exercise
  - ‣ Proactive floor planning / partitioning / clocking strategy mandatory
    - ‣ With care, >90% local resource occupancy is possible
  - ‣ Many bugs / 'features' in vendor tools found and worked around

Technical Context

- Scope of CMS Phase-1 upgrade larger than just L1 trigger
  - DAQ front end upgrades & use of AMC13 common module
  - TCDS upgrade, replacing TTC system
  - Detector readout -> trigger links upgraded to multiple optical links
  - Some early detector front-end changes

- A substantial re-commissioning project for all of CMS TDAQ
  - Interactions between system elements are non-trivial
  - In particular, interface between GT and 'trigger control' completely new

- Commissioning strategy
  - 'Do no harm' – always have a fallback in place to guarantee functional L1
  - Parallel running – commission trigger with data during physics running
    - Implies operation of new and old trigger systems in parallel
  - Use 2015 run as the testbed for 2016

- Advanced enough to profit from Stage-1 calo upgrade for 2015

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Parallel Running

▸ Passive or active splitting of detector signals used throughout the system

Commissioning Steps

- Commissioning steps over the last 24 months:
  - Step 1: Stand-alone module tests
  - Step 2: Interconnection tests
  - Step 3a: System 'dataflow commissioning' and timing in local mode
  - Step 3b (parallel): Final algorithm development and tuning
  - Step 4: System commissioning with data
  - Step 5: Final switch-over to new system

  We are here

- Substantial online software effort required
  - Online framework for system of this size is large and complex
  - Software must also support:
    - Commissioning operations as above, with scriptable interfaces
    - 'Expert mode' operations and special test modes

- This required a completely new framework: SWATCH
  - Constructed building from a low base of online software effort
  - Key is to maximise common interfaces & codebase across the L1 trigger subsystems
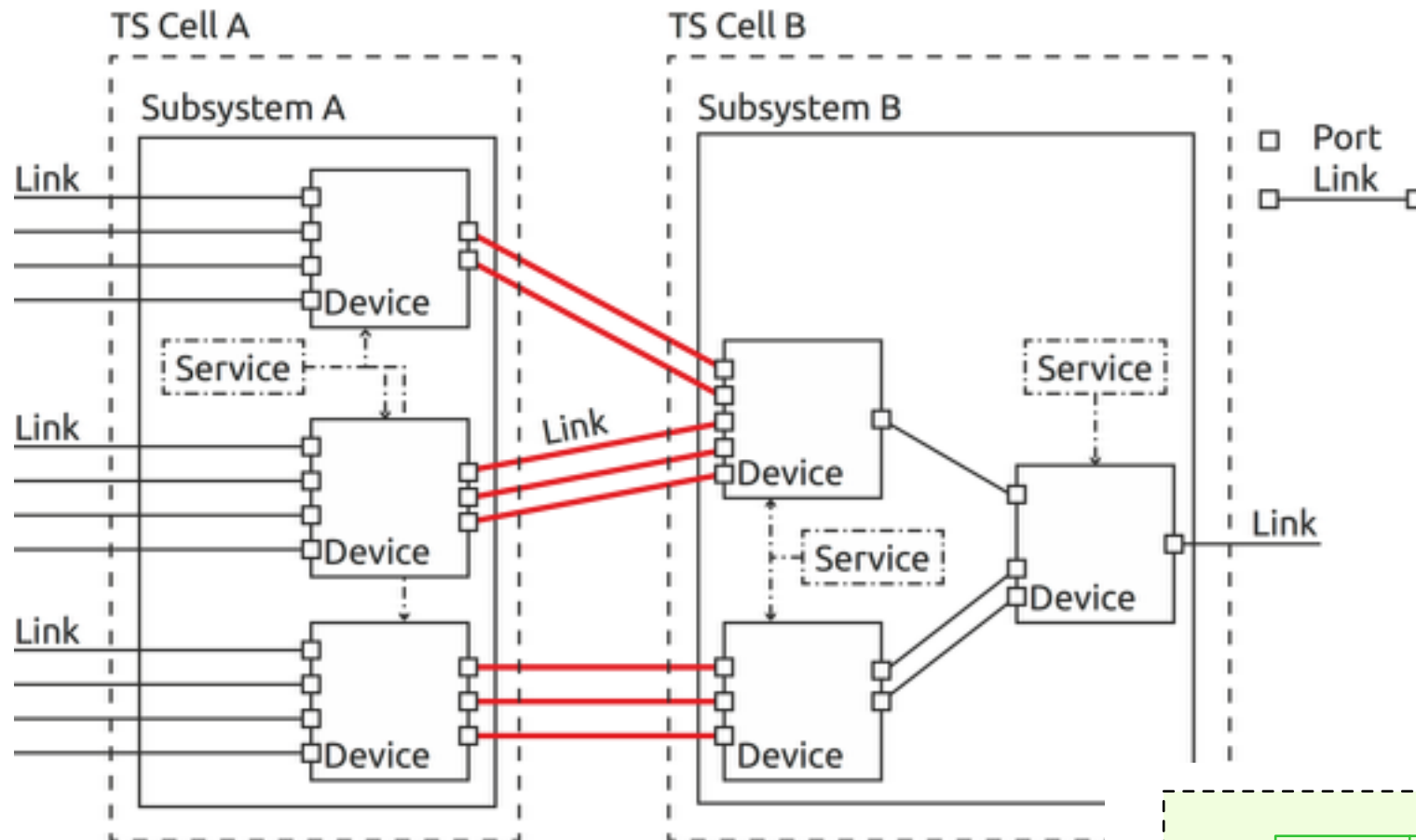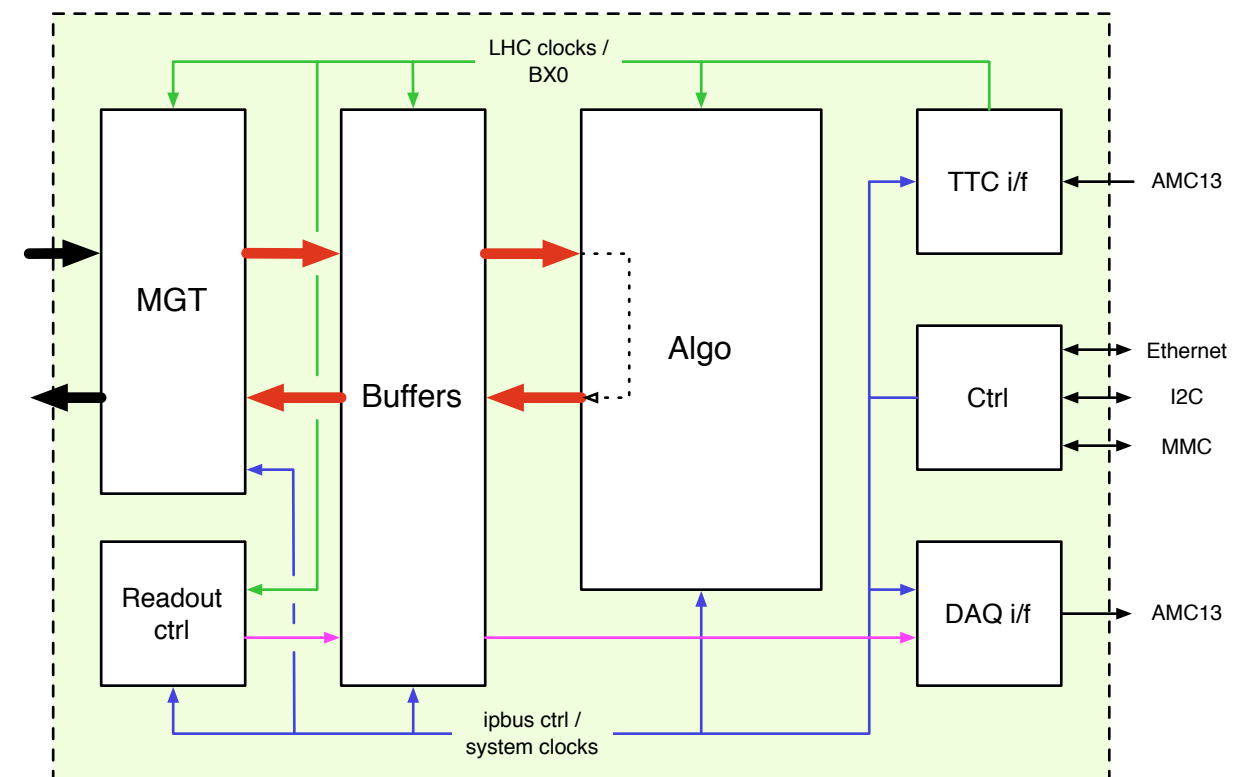
- 7 brand new subsystems
- O(100) boards
- O(3000) optical links
- 3 uTCA processors

  CTP7, MP7, MTF7 & AMC13
- 2/3 "satellite" systems

SW

FW

‣ Online SW a huge task

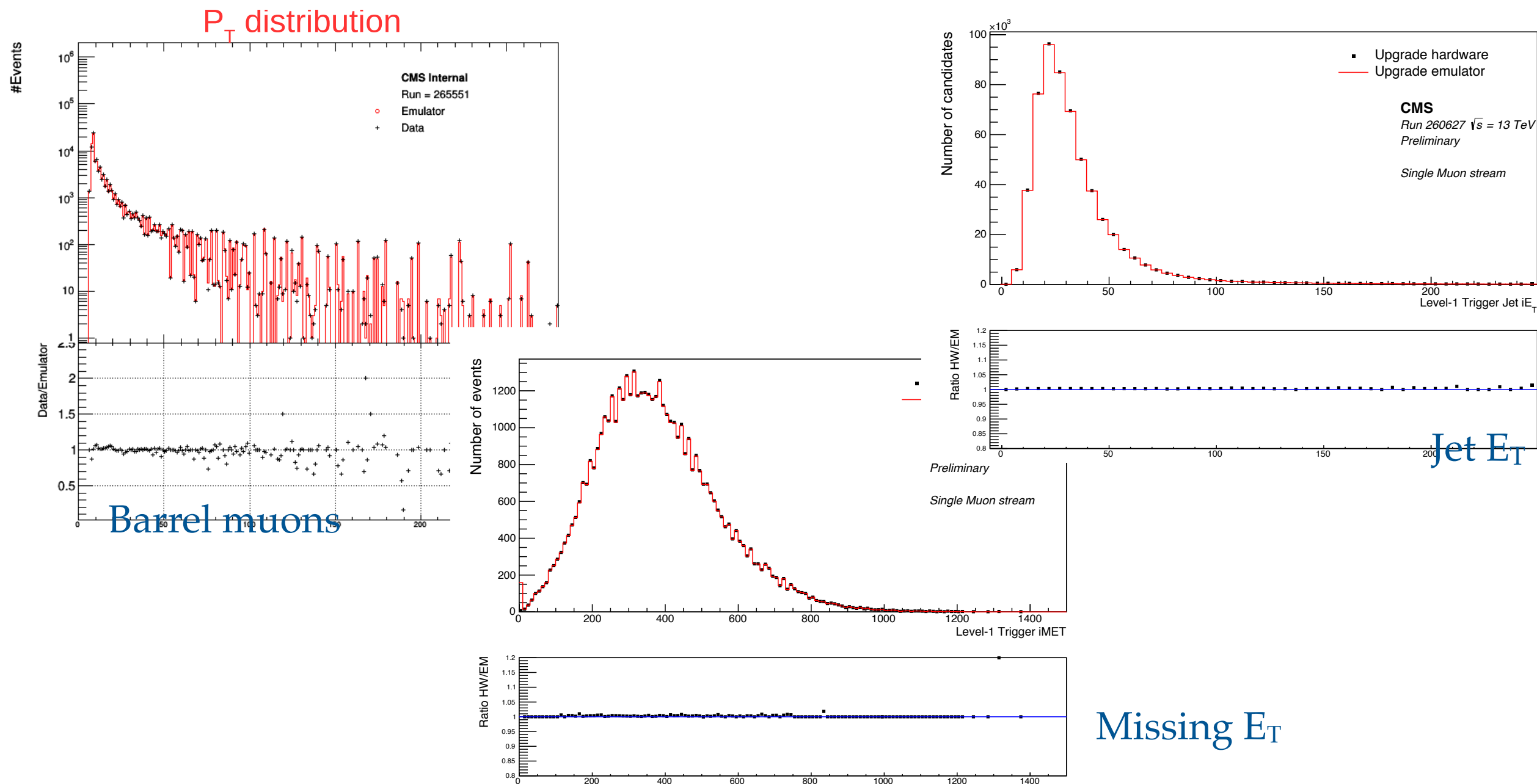 ‣ Complex distributed control and monitoring mandatory for L1
 ‣ Without a new common approach, we would have failed

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Control, Debug, Monitoring

‣ Variety of control approaches used across the system

‣ CTP7: Embedded processing via Xilinx ZYNQ platform

  ‣ Full linux OS system on combined hard CPU / FPGA device on board

  ‣ Control via ethernet; many embedded functions possible

‣ MP7: IPBus lightweight ethernet control protocol

  ‣ Reliable UDP-based ethernet control with software API and on-chip bus

  ‣ The ~minimal way to solve the problem; now in use in all LHC experiments

‣ MTF7: PCIe communication from external host PC

  ‣ Uses embedded PCIe blocks in FPGA for low-overhead solution

  ‣ High throughout allows loading of large LUTs rapidly at system start

‣ Pros and cons to each of these approaches

  ‣ See xTCA workshop later today for more discussion

  ‣ Common higher level software model hides the differences

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

node

agreement

P_T distribution

Jet E_T

Barrel muons

Missing E_T

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Successes

- Modular electronics based on large FPGAs ✅
  - Have already seen the benefits in flexibility
  - The future is 'lego'

- Common firmware and software ✅
  - Some firmware blocks (links, interfaces, DAQ) in wide use across CMS
  - Key is standardisation of on-chip bus interface
  - Modular common online software now mandatory for project of this scale
  - Final integration of calo trigger took ~6 months, muon trigger ~few days

- Mass deployment of high speed MM parallel optics ✅
  - Performance outstanding, cost is not huge compared to processing elements
    - Though, latency consumption is non-trivial; compensated by faster processing on FPGAs

- Parallel commissioning ✅
  - Required much upfront work during LS1, but otherwise impossible to commission trigger on schedule
  - The 'split links' remain for testing of new ideas in coming years

- Time-multiplexed architecture ✅
  - This approach is likely to be used for future trigger upgrades

Challenges
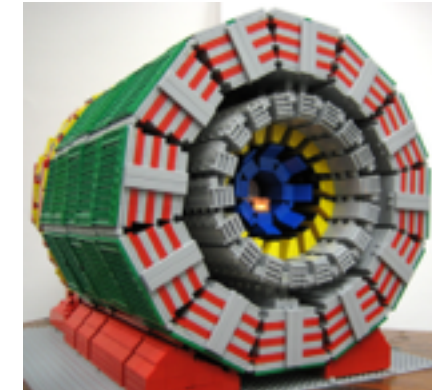
▸ **A very large technical step during LS1**

  ▸ 'Seamless transition' between R&D and deployment of 7-series modules

    ▸ Still learning much about the technology during early commissioning

  ▸ Board manufacturability required careful attention throughout the project

  ▸ Procurement also painful at times for few / advanced / expensive boards

▸ **'Re-learning CMS'**

  ▸ A lot of deep voodoo was uncovered (and expunged) during the upgrade

  ▸ Parallel running forced a more programmatic approach to timing in

▸ **Schedule was tight**

  ▸ We took some risk in deployment during LS1 – but always a way back

▸ **Effort for online software insufficient**

  ▸ Only heroic (and not sustainable) efforts have brought us to where we are

  ▸ Appears to be a chronic problem; the solutions are political, not technical

  ▸ The job does not end when the hardware is finished (it never ends)

▸ **microTCA not a panacea**

  ▸ We entered the microTCA world with high hopes, and learnt some lessons

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL

Lessons Learnt

- **Common components make sense**
  - The 'new world' applies to hardware, firmware, software
  - Much upfront effort in 'soft' work: specification, standards, interfaces, testbenches, etc
  - Cannot bring about this approach by legislation, only by consensus

- **microTCA advantages**
  - A key enabling technology behind our successful modular appro
  - Commodity ethernet control links were a success
  - Adoption across CMS allowed exchange of experience

- **microTCA disadvantages**
  - Form factor not optimal for future more power-hungry FPGAs
  - Physical, electrical, and logical interface specification has issues
    - Including an unreasonably complex and fragmented specification
  - No well-defined approach to backplane extensibility
  - Vendor support mostly good, but serious issues with cross-vendor interoperability
  - Reliability and COTS quality claims not (yet) substantiated

- **Bottom line: microTCA not likely to be the baseline for Phase-2**

Conclusions

- **Phase-1 trigger upgrade for CMS successfully deployed** ✅
  - A marathon effort over a number of years by many people
  - Substantial benefits for CMS Run 2physics programme and operations

- **Successful new developments**
  - Modular processing platform approach based on FPGAs / parallel optics
  - Mass deployment of microTCA electronics
  - Splitting of detector data and parallel commissioning
  - Time-multiplexed architecture

- **Many lessons learnt**
  - Common components pay off, but do not come 'for free'
  - microTCA served us well for this project, but search for 'the new VME' continues
  - Software continues to be an existential threat to projects of this scale

- **The future**
  - Trigger design allows for flexibility, expansion – will make much use of this
  - Absorbing lessons as we embark upon Phase-2 design choices for CMS TDAQ

Science & Technology Facilities Council
Rutherford Appleton Laboratory

University of BRISTOL