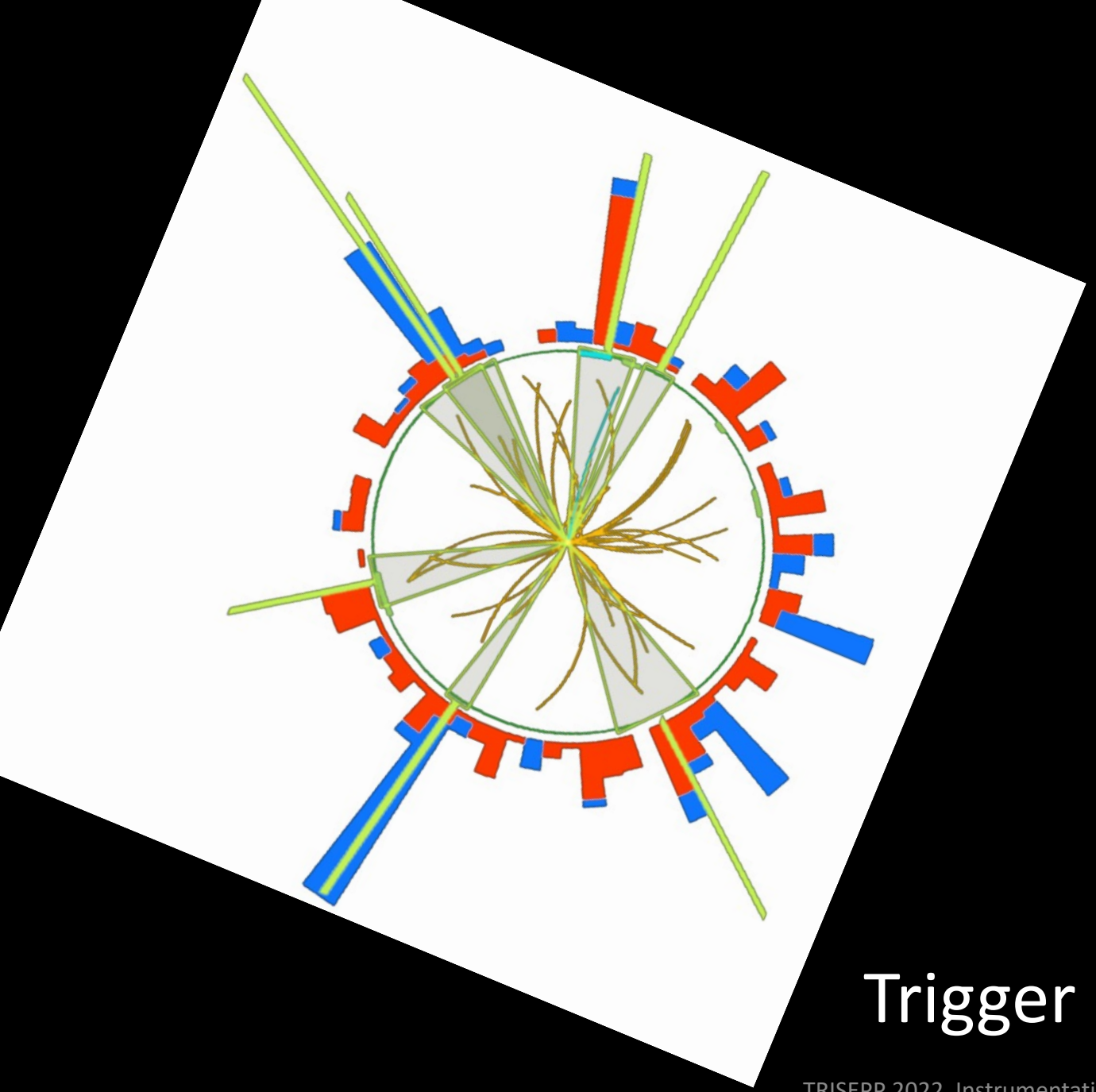
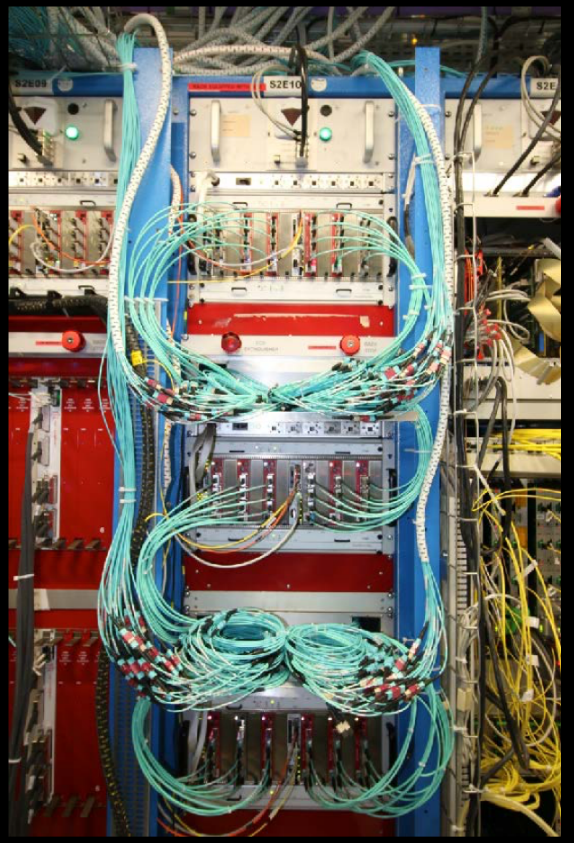




Trigger and DAQ

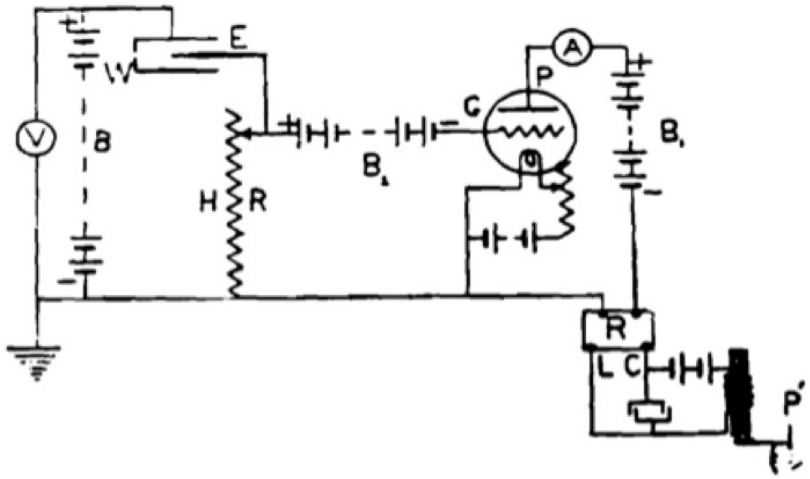
Overview

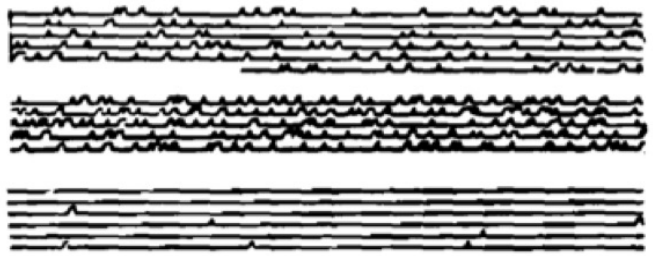
- The basic concepts
- Today's challenges
- Case studies from LHC
- Tomorrow's challenges
- Future evolution and work

- Front end:
  - The electronics within the experiment that digitises signals
- Data acquisition:
  - Captures, records, organises, stores the raw data from an experiment
- Trigger:
  - Makes decisions on what to store from the experiment, and when
- Offline:
  - The computing and software systems that handle the data after storage
- These boundaries are becoming increasingly blurred – more later

Centenary

ON THE AUTOMATIC REGISTRATION OF α-PARTICLES, β-PARTICLES AND γ-RAY AND X-RAY PULSES





- Physics TDAQ is 100 years old!
  - Phys. Rev. 13, 272 , 1st April 1919

"… visual or audible methods of counting are quite trying on the nerves … A self-recording device would therefore be an obvious improvement."

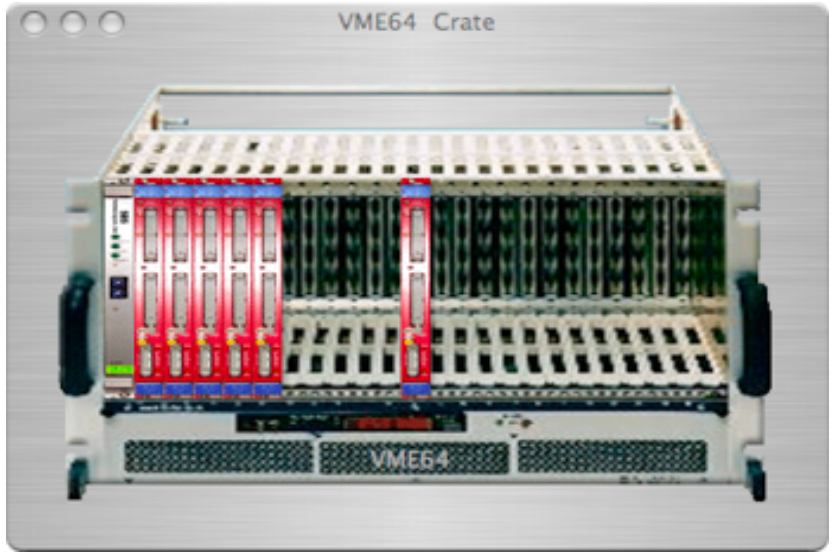
- A lot has happened since then
  - But trigger and DAQ can still be quite trying on the nerves

Why is TDAQ Interesting?

- Why do smart people spend time on such a 'technical' thing?
- Close to the physics
  - TDAQ underlies, and affects strongly, the physics output of experiments
  - Trigger is the 'first stage of analysis', and irrevocable
- Strong need for efficiency
  - 'Data-handling' is now the dominant cost of many projects
  - Strongly professional approach required for successful outcome
- Challenge
  - Very serious technological challenges for the next generation of projects
  - We are *required* to push the technology envelope (especially for trigger)
- Technological relevance
  - One of the most direct points of contact with industry and other fields
  - Excellent training ground for new researchers
- 45 minutes far too short to cover the topic
  - For a pedagogical introduction: https://indico.cern.ch/category/345
  - Cannot cover test beam DAQ, machine learning, new architectures, etc

Wait, but…?

- Can't you just *buy* this stuff?
  - For 'small' systems, off-the-shelf systems are certainly an option
  - Important to recognise total-cost-of-ownership (over many years)
- What does 'small' mean?
  - Limited throughput: <1GB/s
  - Non-distributed system
    - Rely on traditional 'local bus' for event building
  - No complex trigger requirements
  - No strong reliability requirements

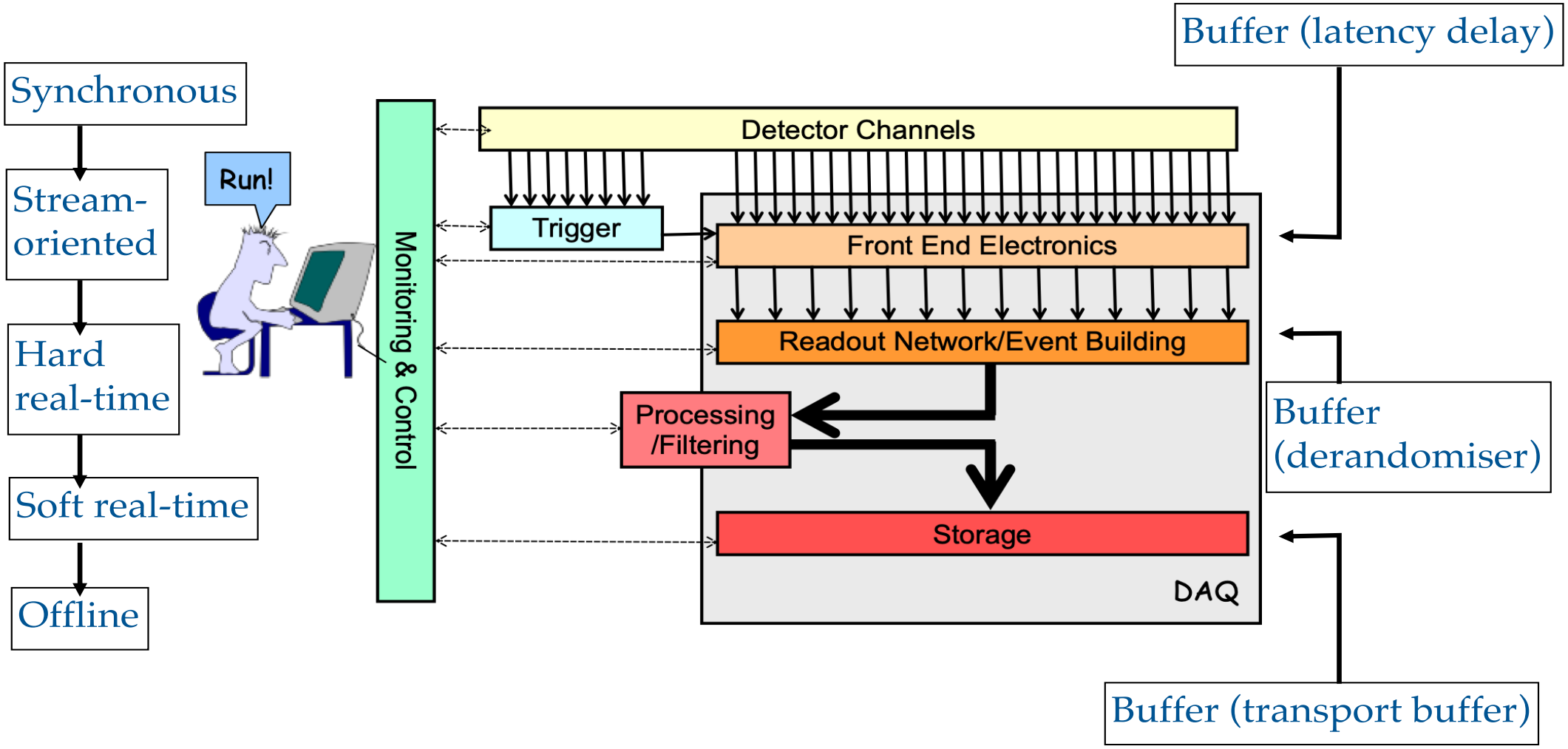


- In general, however:
  - Data rates are very large (at least at some points in the system)
  - Distributed hardware and computing is mandatory
  - Flexible and complex trigger functionality is needed
- Strong inclination to void 'vendor lock in' for long term projects

DAQ: Requirements

- Interface to, and control, front end ('readout')
  - Often including timing and synchronisation
- Organise data into coherent structure ('event building')
  - For high-level triggering and storage, need all related data in one place
  - Usually implies a channel-wise to time-wise ordering
  - Requires many-to-many network fabric; often the expensive part
- Record the data, correctly, at required rate ('logging')
  - With minimal dead time (time after trigger in which no new triggers)
  - Ensure data integrity and robust / reliable operation
- Configuration and control of the system ('run control')
  - Sometimes a rather complex distributed computing system in play
- Provide monitoring of system and data quality ('DQM')
- Usually: provide the basic layer of computing / IPC services
- Sometimes: operate and monitor the detector ('slow controls')

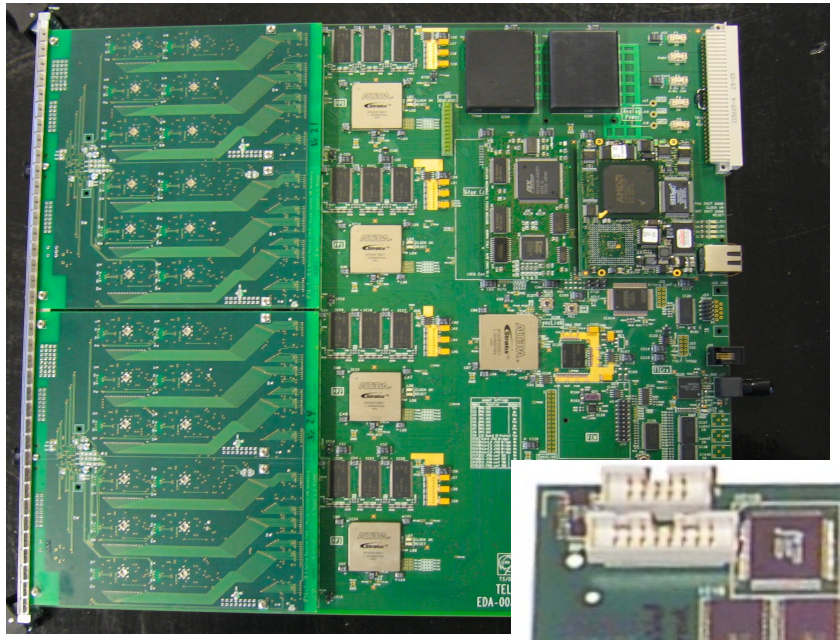
DAQ: Architecture



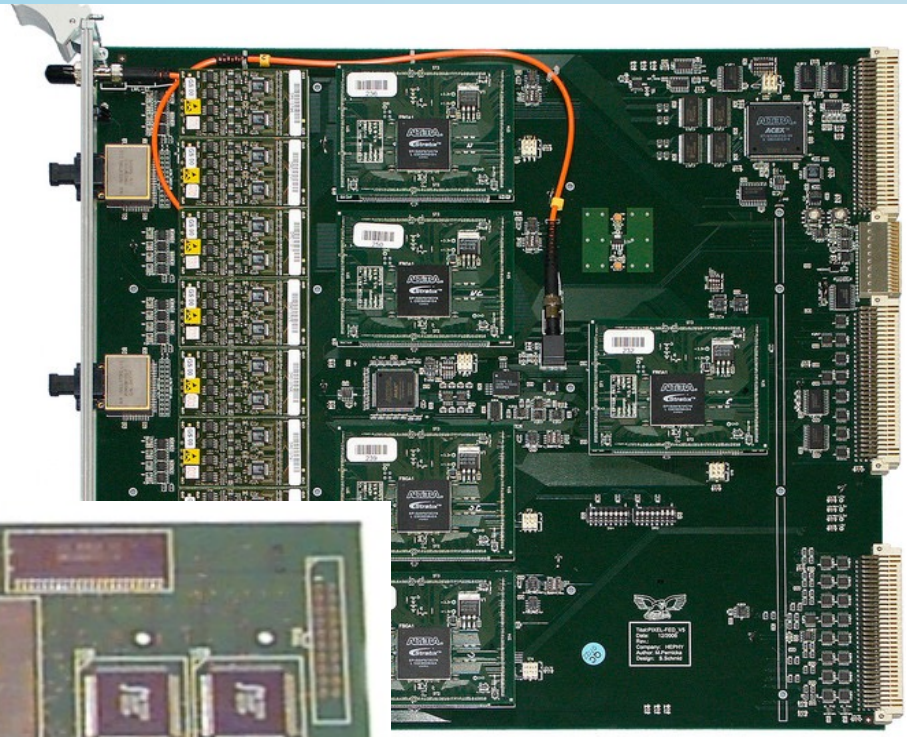
DAQ: Implementation

- Increasing reliance on 'commodity' tech
  - Commodity really means data-centre quality, at least for large systems
- Computation
  - Almost uniform use of x86 / Linux
    - Increasing use of GPU / FPGA / instruction set extensions
  - I/O subsystem is now often the bottleneck
  - RAM bandwidth also an issue in some systems
- Storage
  - Until recently, spinning disks mainstream
  - 'Continuously updated rolling buffer' is not a typical business computing use case
  - SSDs now entering the field; but wearout?
- Networking
  - Until ~2010: HPC-class interconnect (Myrinet, infiniband) in key places
  - All Data Centre Ethernet (10/40/100GBE) now

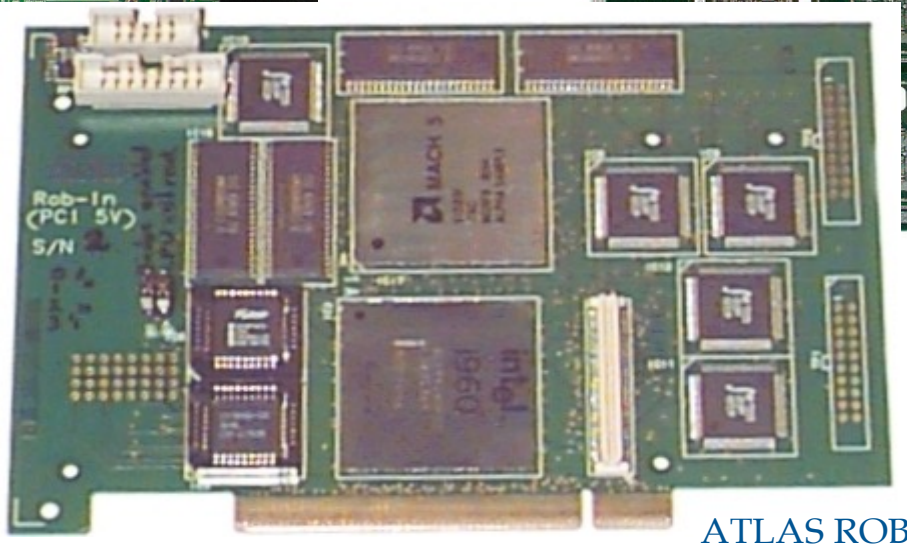
DAQ: Implementation



LHCb Tell1



CMS Pixel FED



ATLAS ROBIN

- Often a thin layer of custom hardware interfacing to 'computing world'

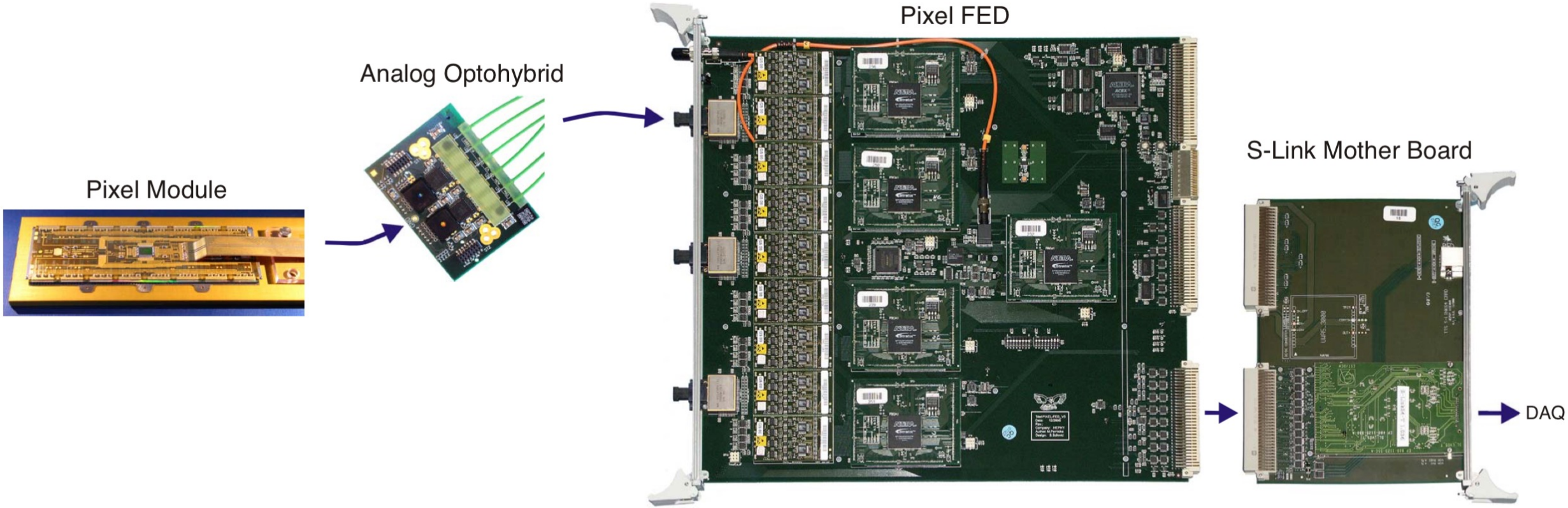
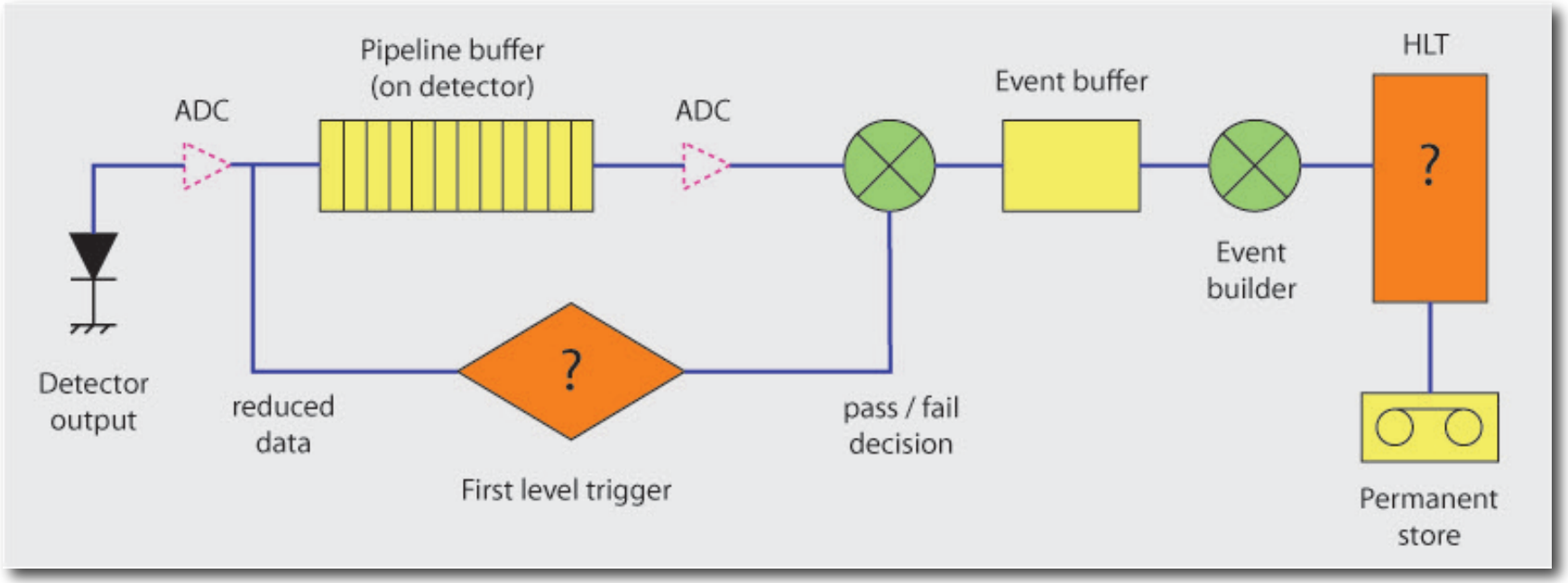


Figure 1: The Pixel Detector Readout Chain

Trigger: Architecture

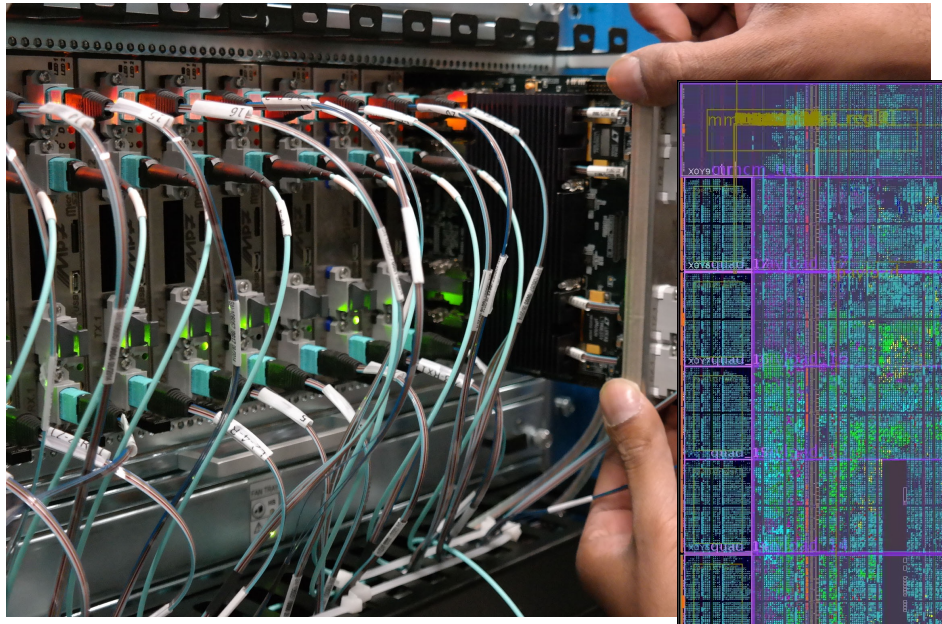


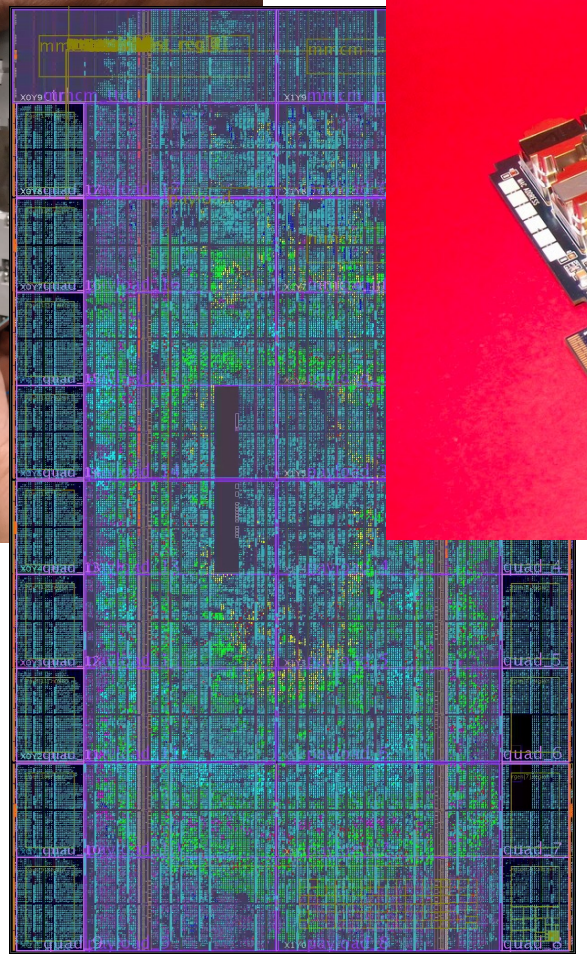
- 30+ years ago
  - Problem usually readout time; slow detectors need to be cleared if event is rejected
  - Dead time a major issue
- Currently
  - Key problem is data volume – systems are usually pipelined, deadtimeless
- In the future
  - Key problem will be readout bandwidth from the detector, limited by power and space

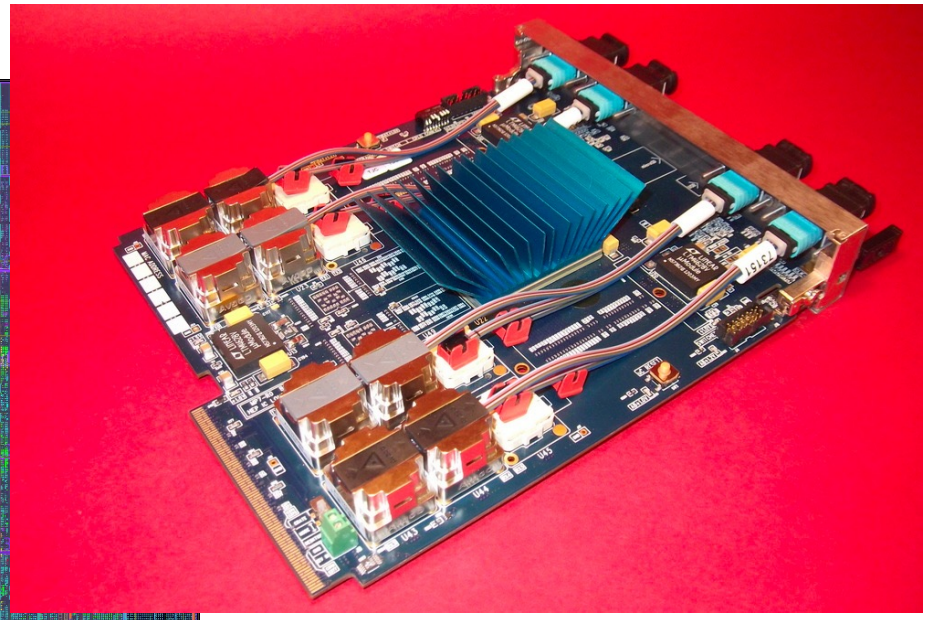
Trigger: Requirements

- Carry out real-time data reduction
  - Channel data aggregation / truncation / compression
  - Sometimes zero suppression; but often the zeroes are important!
- Feature extraction
  - Look for 'physics signatures' in the reduced data – always imperfectly
- Select full data to record based on reduced data
  - This may mean which events to record (for collider detectors)
  - In other cases, may also select Regions of Interest / Space-Time Ranges
- Make selection within strict time / throughput limits
  - Most triggers operate with fixed latency, which dictates buffer lengths
  - Limits can be stringent [O(μs)], requiring custom hardware systems
- Make selection with high efficiency, well-quantified bias
  - Trigger acceptance often studied using less-restrictive triggers with rescales
- Operation must be continuous, and robust
  - The trigger is a single point of failure for the whole experiment

Trigger: Implementation

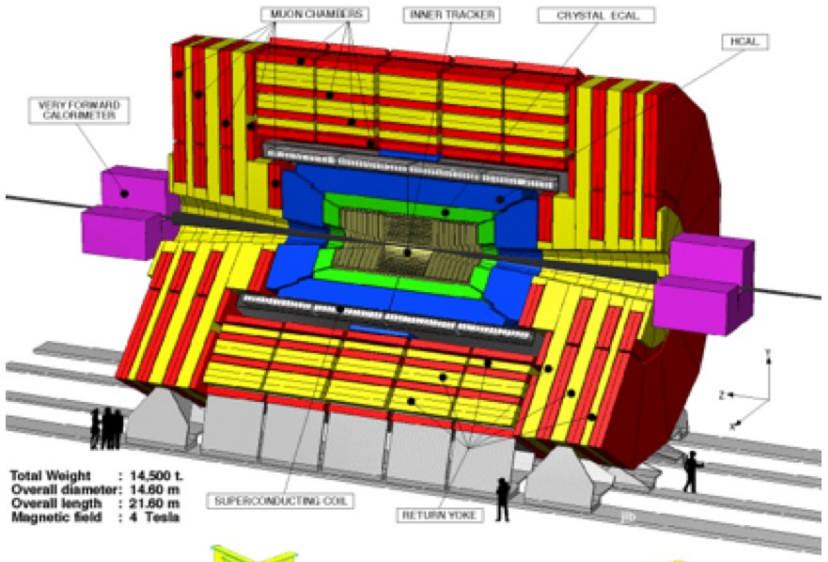


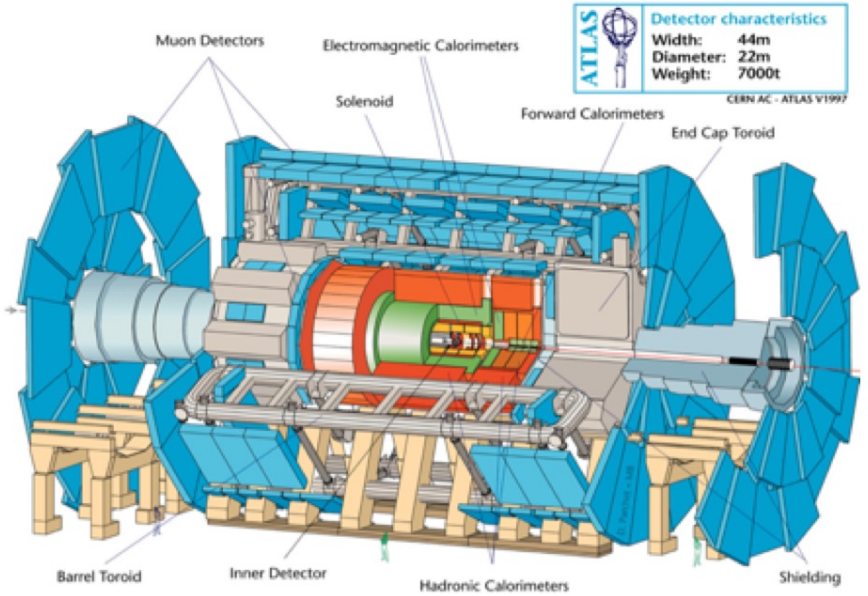


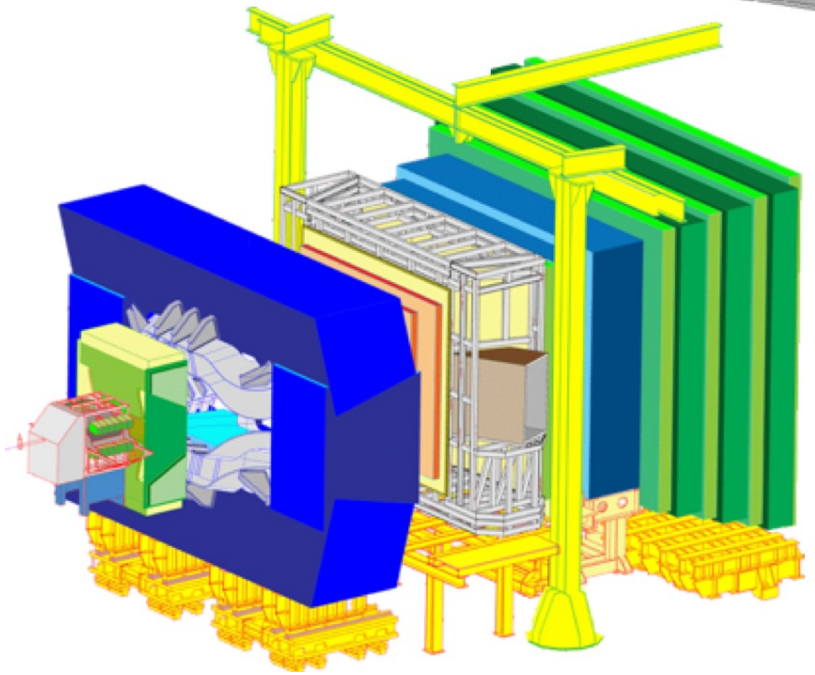


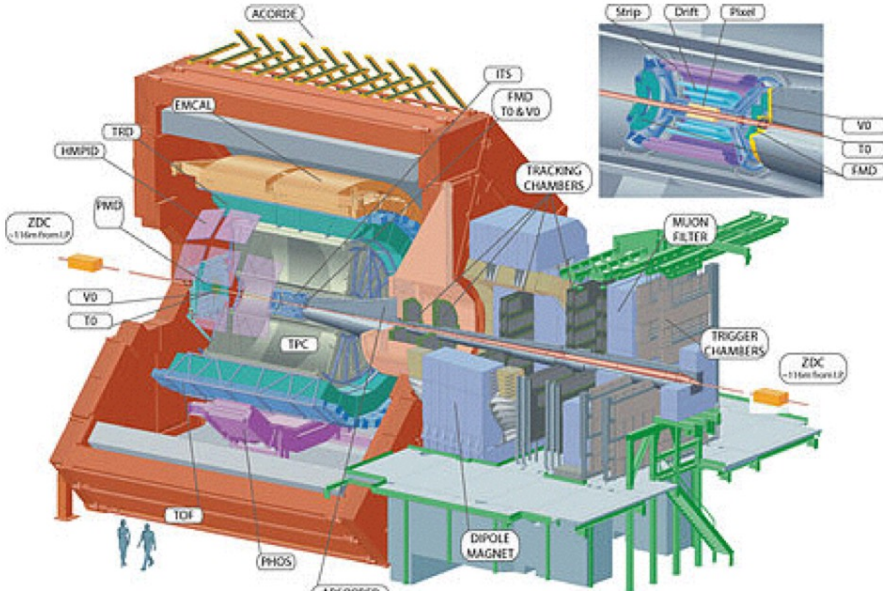
- Large FPGAs, short-haul medium-speed (10Gb/s) NRZ optolinks ubiquitous
  - No more ASICs; development cost and risk too high, not flexible, low gate count
  - Modular electronics the norm, but no real 'common standard' at present

Case Study: LHC

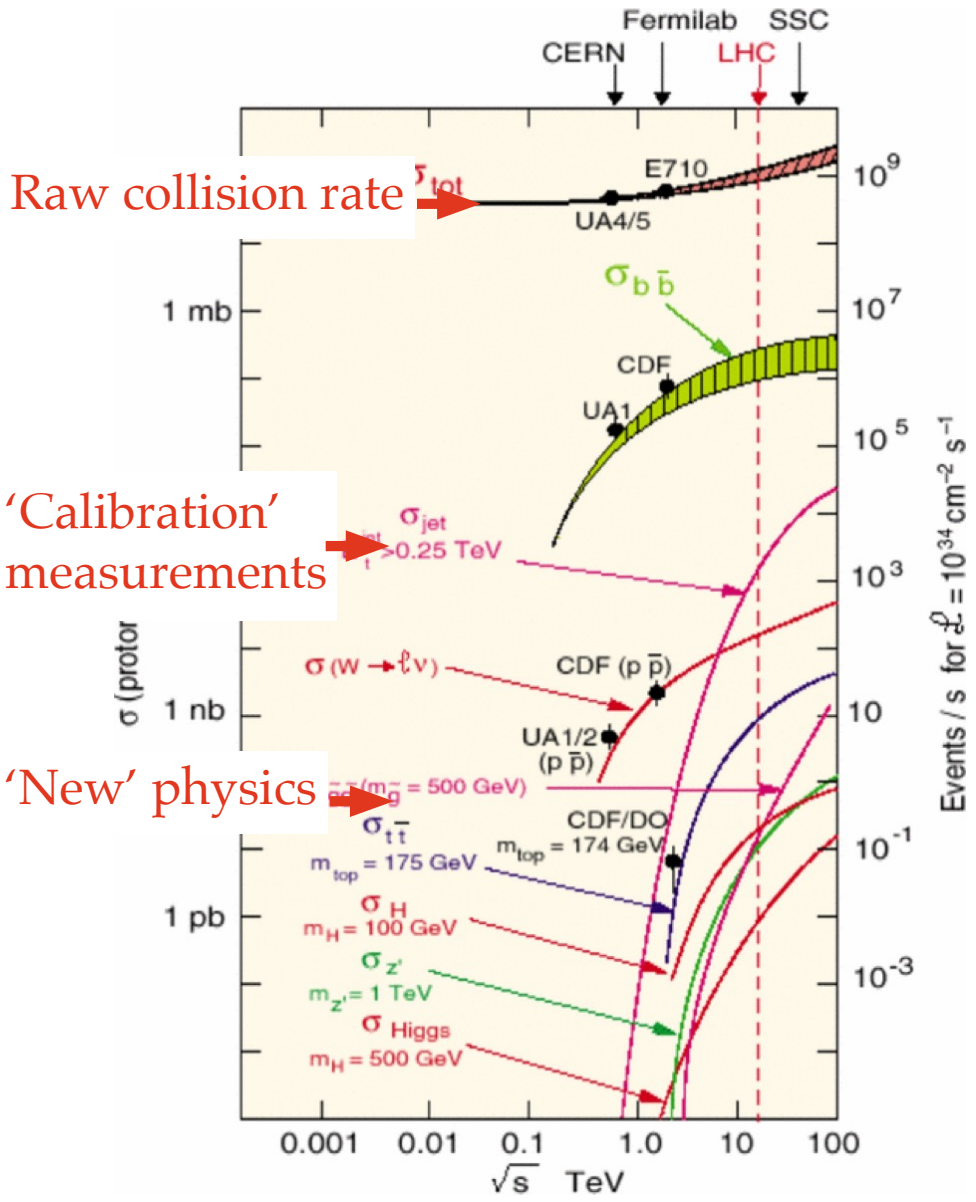






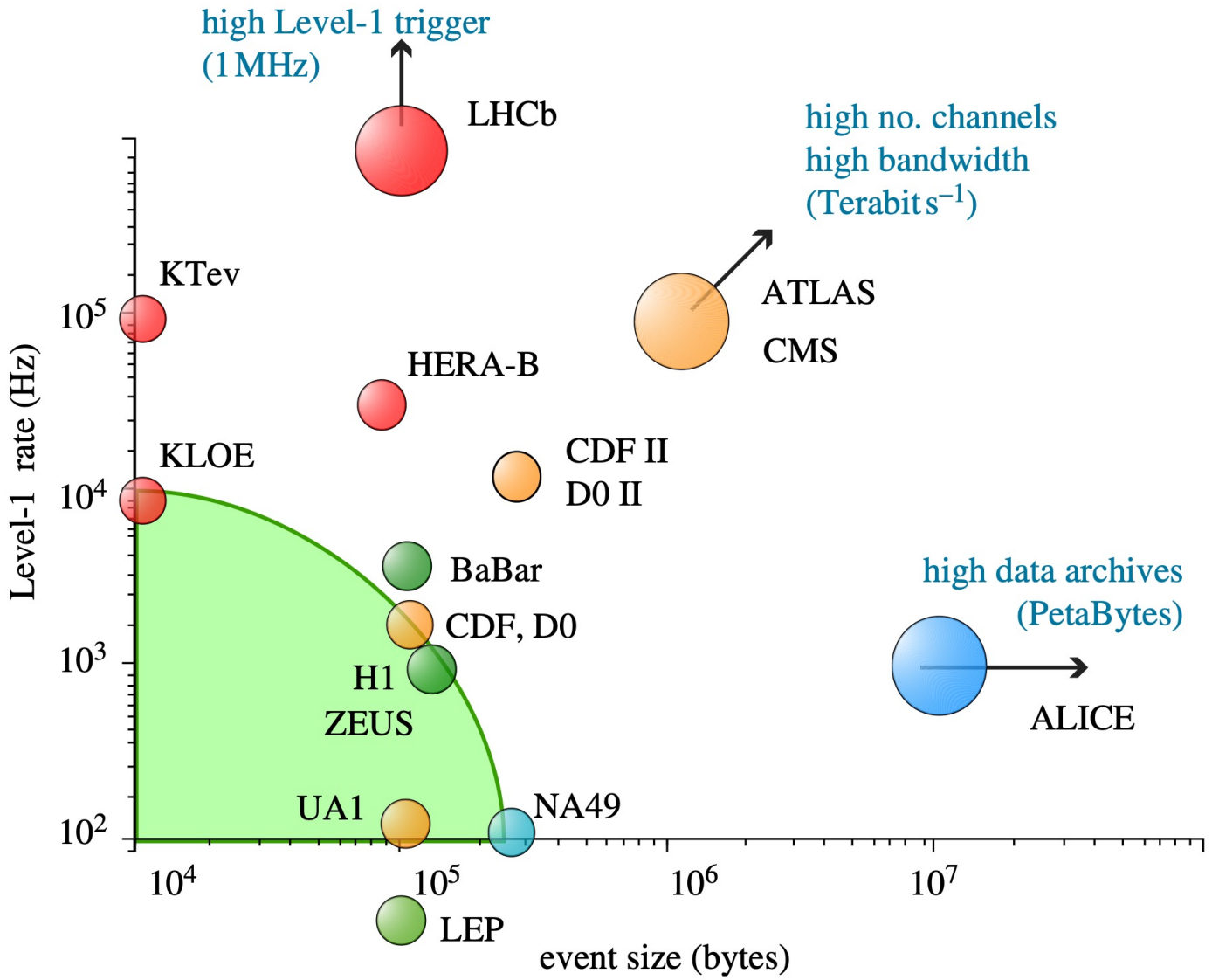


GPD Trigger Problem



- Typical detector parameters
  - 100M channels, 40MHz sampling, 12b resolution, 100 days per year = 50ZB/yr
  - ZS might buy factor 100…
- A GPD trigger must
  - Reduce recorded event rate by factor of ~200,000
  - Operate with ~1μs latency
- Restricted data available
  - Only calorimeter and muon tracking systems
- Physics reach ultimately dictated by trigger efficiency
  - And the uncertainties in it

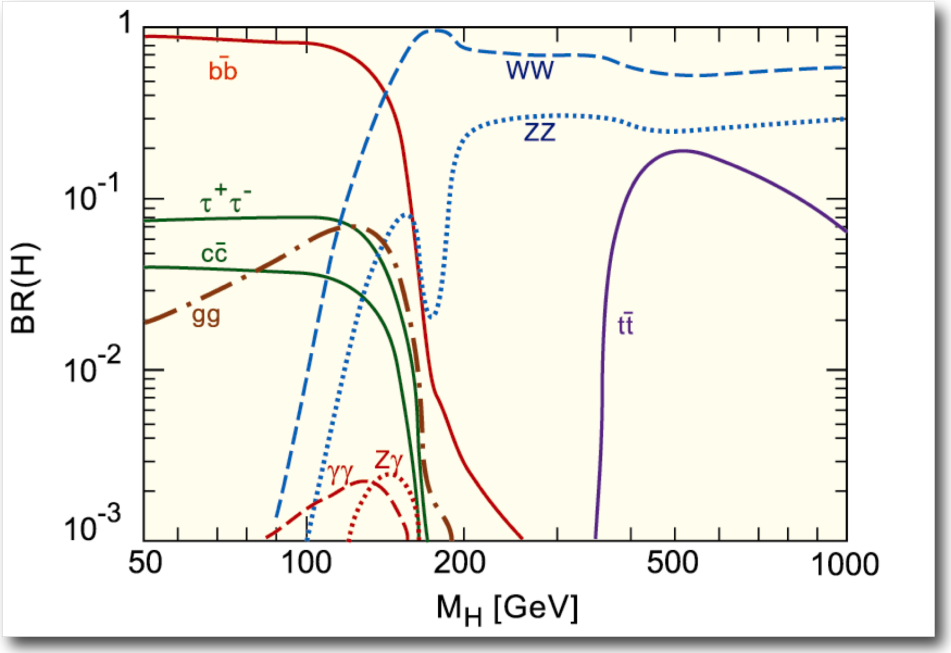
LHC Data Rates



- Recorded raw data from GPDs ~1PB/yr (factor of a few for copies, reco data)

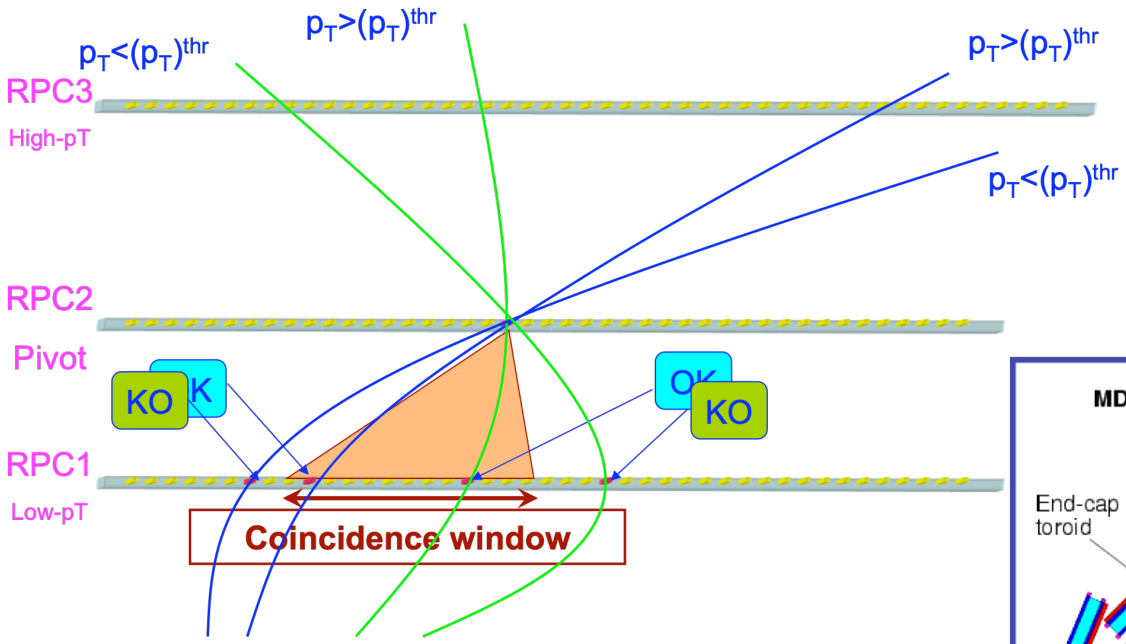
GPD Trigger Strategy

- Working > electroweak scale
  - Need good sensitivity for decays of objects of mass O(100GeV)
- Basic 'trigger objects'
  - Leptons – with / without isolation
    - Isolation suffers badly as pileup increases
  - Photons – hard to distinguish from e
  - Jets of sufficiently high E_t; hadronic tau
  - Global variables: E_t, missing E_t, H_t, etc

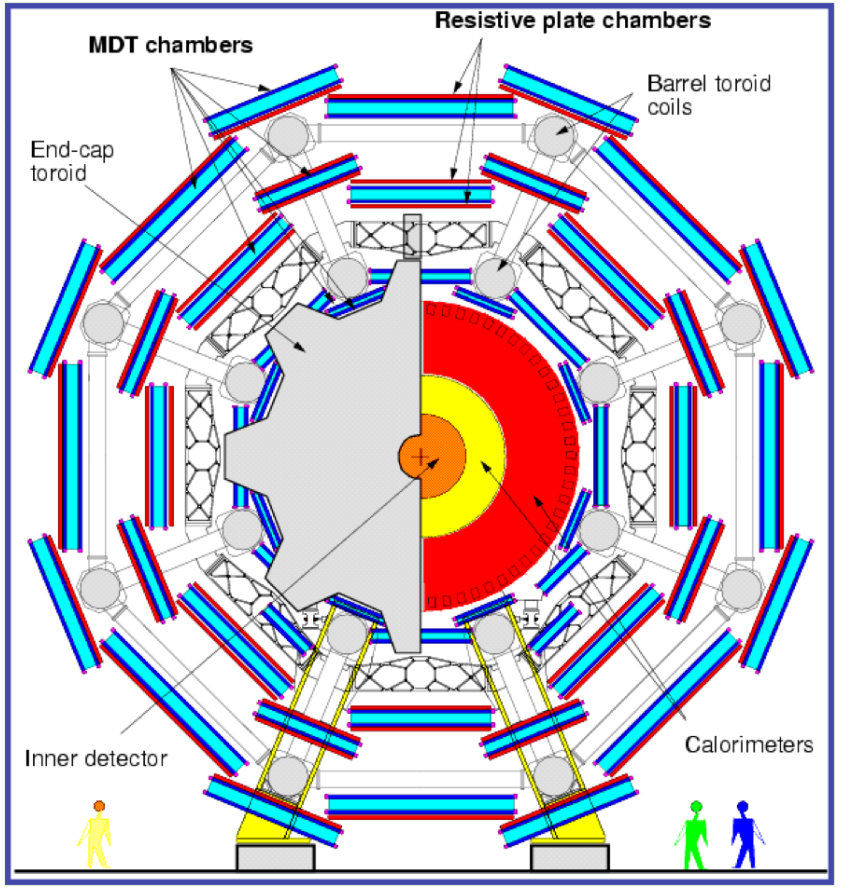


- Backgrounds
  - Real leptons from heavy flavour production, typically soft
  - Mis-measured jets; photons in jets; fake electrons from conversions
  - Punch-through in muon system
- Making the decision
  - Require combinations of trigger objects above a set of E_t thresholds
  - More clever combinations are possible with sufficient trigger resolution
    - Event topology cuts, invariant mass, etc

L1 Trigger Algorithms: Muons

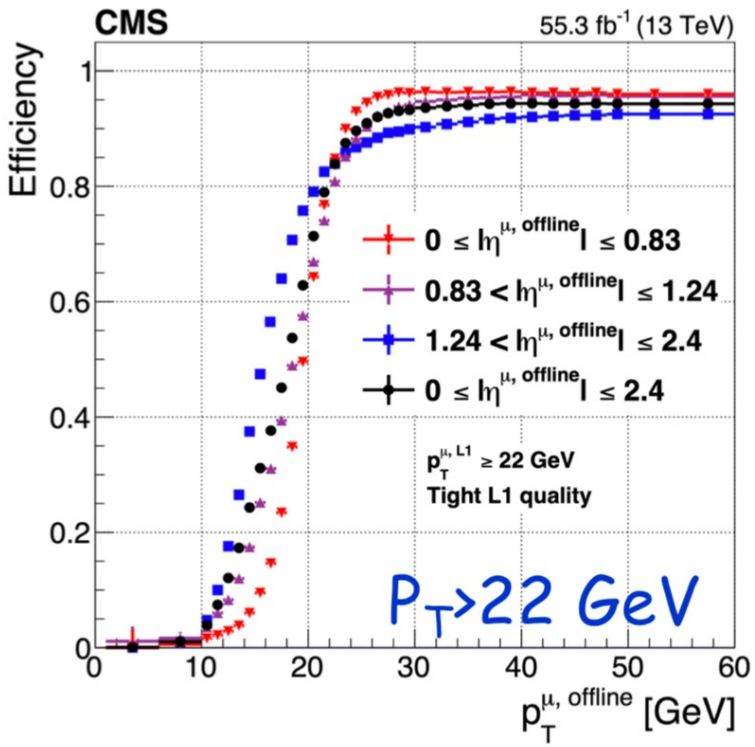


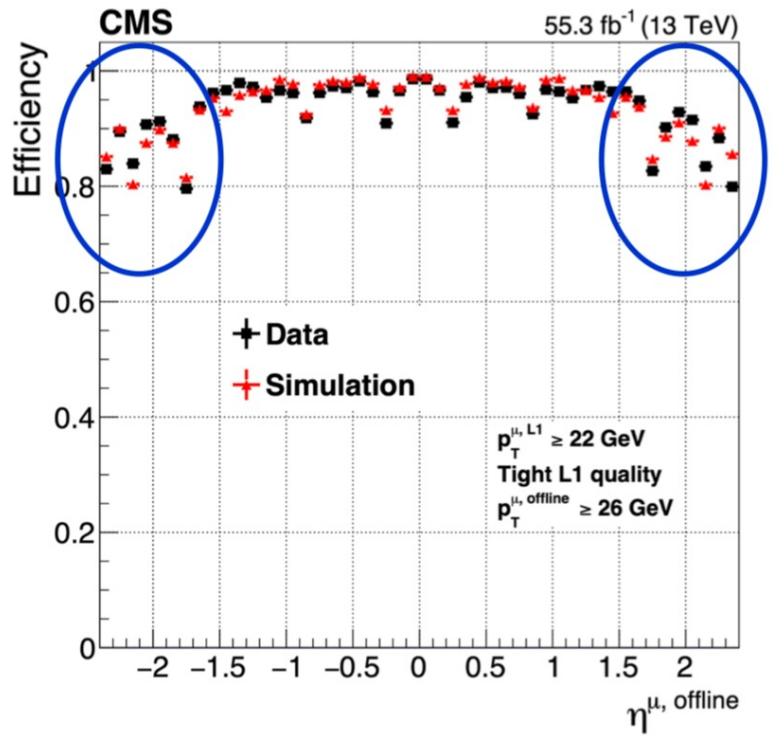
- Not as simple as it looks!
  - Hit correlation in 4D is necessary
  - Muon detector spacing is large compared to time-of-flight
  - Combine detectors with good space resolution (tracking) with good time resolution (bunch-crossing assignment)

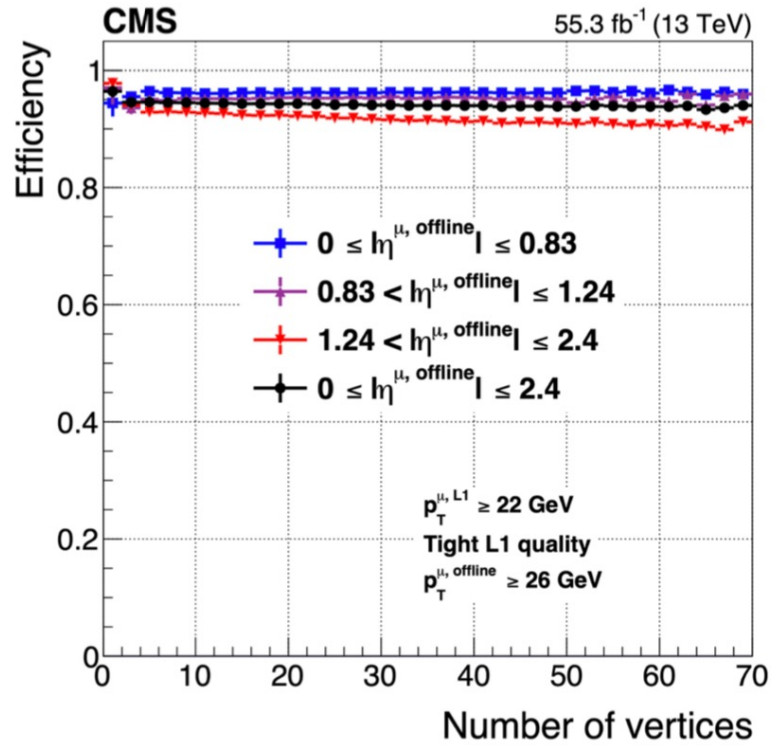


Level-1 Muon Trigger Performance

JINST 15, P10017 (2020)

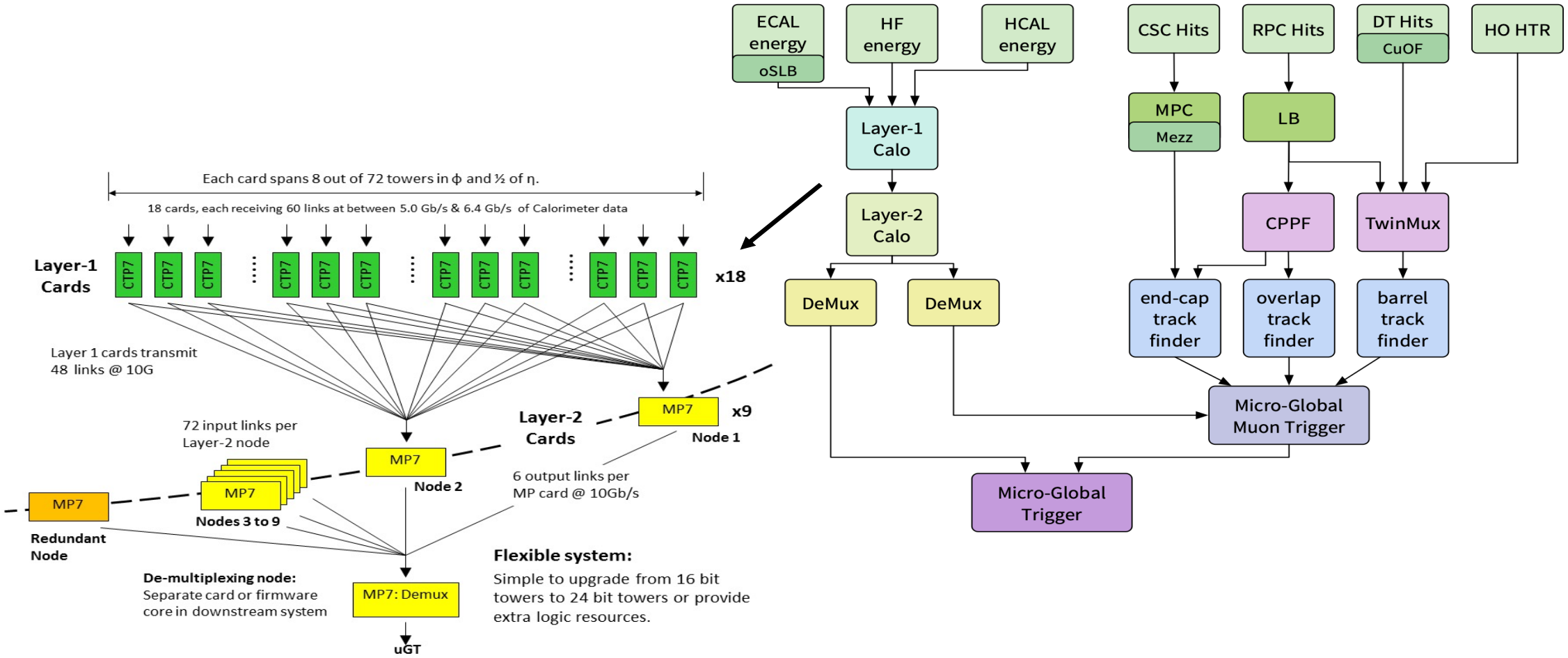






- Efficiency falls slightly in forward region
  - Less redundancy → should improve for Run 3 with new GEM detectors
- Efficiency flat with pileup

Data Flow



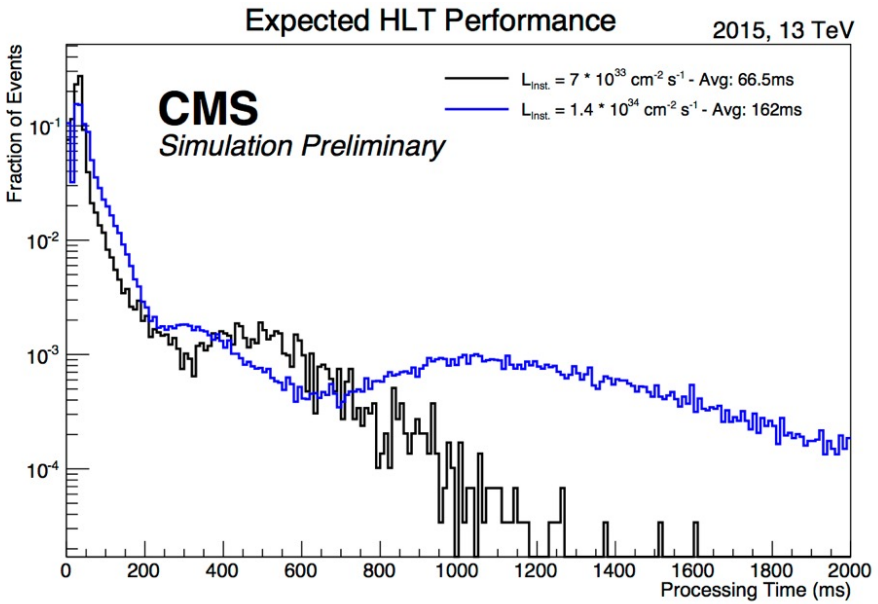
- Shown is CMS calo 'run 2' trigger
  - The scale: around €5M kit, ~100 staff-years of work (€10M) over three years
  - Much less effort than original trigger, due to some sharing of common hardware

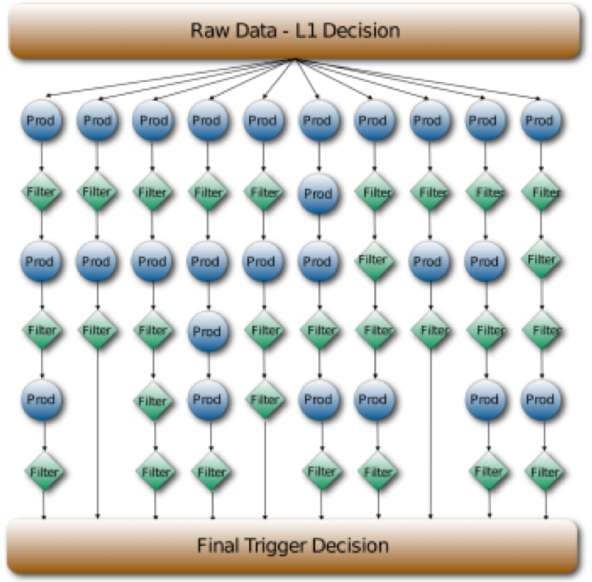
Final Trigger Selection

- Final selection based on 'cut and count' today
  - Some obvious ways to improve on this…?
  - Significant bandwidth given to monitoring / 'backup' triggers

```
L1_SingleMu3 (4000) : Indiv.: 3.2 +/- 2.5
 L1_SingleMu5 (2000) : Indiv.: 3.2 +/- 2.5
 L1_SingleMu10 (1) : Indiv.: 496.7 +/- 17.1
 L1_DoubleMu3 (1) : Indiv.: 316.1 +/- 20.3
 L1_TripleMu3 (1) : Indiv.: 7.0 +/- 2.5
 L1_Mu3_Jet15 (20) : Indiv.: 200.0 +/- 17.1
 L1_Mu5_Jet20 (1) : Indiv.: 1282.5 +/- 36.0
 L1_Mu3_IsoEG5 (1) : Indiv.: 922.0 +/- 35.6
 L1_Mu5_IsoEG10 (1) : Indiv.: 57.4 +/- 7.0
 L1_Mu3_EG12 (1) : Indiv.: 82.9 +/- 9.2
 L1_SingleIsoEG8 (1000) : Indiv.: 19.2 +/- 6.5
 L1_SingleIsoEG10 (100) : Indiv.: 82.8 +/- 13.5
 L1_SingleIsoEG12 (1) : Indiv.: 4003.4 +/- 93.0
 L1_SingleIsoEG15 (1) : Indiv.: 1757.9 +/- 61.3
 L1_SingleIsoEG20 (1) : Indiv.: 574.8 +/- 34.8
 L1_SingleIsoEG25 (1) : Indiv.: 232.1 +/- 22.0
 L1_SingleEG5 (10000) : Indiv.: 13.3 +/- 5.5
 L1_SingleEG8 (1000) : Indiv.: 21.9 +/- 7.0
 L1_SingleEG10 (100) : Indiv.: 99.8 +/- 14.8
 L1_SingleEG12 (100) : Indiv.: 53.4 +/- 10.7
 L1_SingleEG15 (1) : Indiv.: 2471.9 +/- 72.3
 L1_SingleEG20 (1) : Indiv.: 925.5 +/- 43.7
 L1_SingleEG25 (1) : Indiv.: 456.7 +/- 30.7
 L1_SingleJet15 (100000) : Indiv.: 10.3 +/- 4.9
 L1_SingleJet30 (10000) : Indiv.: 18.7 +/- 6.5
 L1_SingleJet70 (100) : Indiv.: 34.2 +/- 8.5
 L1_SingleJet100 (1) : Indiv.: 588.3 +/- 34.7
 L1_SingleJet150 (1) : Indiv.: 66.4 +/- 11.0
 L1_SingleJet200 (1) : Indiv.: 19.5 +/- 6.0
 L1_SingleTauJet40 (1000) : Indiv.: 0.0 +/- 0.0
 L1_SingleTauJet80 (1) : Indiv.: 723.1 +/- 38.4
 L1_SingleTauJet100 (1) : Indiv.: 214.5 +/- 20.8
 L1_HTT100 (10000) : Indiv.: 16.3 +/- 6.0
 L1_HTT200 (1000) : Indiv.: 22.3 +/- 7.0
 L1_HTT250 (100) : Indiv.: 60.6 +/- 11.3
 L1_HTT300 (1) : Indiv.: 1739.1 +/- 59.8
 L1_HTT400 (1) : Indiv.: 158.5 +/- 17.4
 ETM45 (1) : Indiv.: 527.6 +/- 33.8
 ETM45_Jet30 (1) : Indiv.: 511.6 +/- 33.3
 ETM50 (1) : Indiv.: 190.0 +/- 20.0
 L1_DoubleIsoEG8 (1) : Indiv.: 740.4 +/- 39.2
 L1_DoubleEG10 (1) : Indiv.: 0.0 +/- 0.0
 L1_DoubleJet70 (1) : Indiv.: 733.9 +/- 38.8
 L1_DoubleJet100 (1) : Indiv.: 150.3 +/- 17.4
 L1_DoubleTauJet40 (1) : Indiv.: 2970.4 +/- 78.9
 L1_IsoEG10_Jet15 (20) : Indiv.: 345.4 +/- 27.4
 L1_IsoEG10_Jet30 (1) : Indiv.: 3990.7 +/- 92.2
 L1_IsoEG10_Jet70 (1) : Indiv.: 472.8 +/- 31.0
 L1_IsoEG10_TauJet20 (1) : Indiv.: 3697.9 +/- 88.7
 L1_IsoEG10_TauJet30 (1) : Indiv.: 2389.5 +/- 70.9
 L1_TauJet30_ETM30 (1) : Indiv.: 3570.6 +/- 88.3
 L1_TauJet30_ETM40 (1) : Indiv.: 587.7 +/- 35.4
 L1_HTT100_ETM30 (1) : Indiv.: 0.0 +/- 0.0
 L1_TripleJet50 (1) : Indiv.: 349.7 +/- 26.1
 QuadJet40 (1) : Indiv.: 192.9 +/- 19.3
 QuadJet50 (1) : Indiv.: 43.7 +/- 8.9
 L1_ExclusiveDoubleIsoEG6 (1) : Indiv.: 467.1 +/- 32.3
 L1_ExclusiveDoubleJet60 (1) : Indiv.: 158.5 +/- 18.6
 L1_ExclusiveJet25_Gap_Jet25 (1) : Indiv.: 776.4 +/-
42.7 seqPure:
 L1_IsoEG10_Jet20_ForJet10 (1) : Indiv.: 2130.9 +/- 67.6
 L1_MinBias_HTT10 (1) : Indiv.: 0.4 +/- 0.1
 L1_ZeroBias (1) : Indiv.: 0.6 +/- 0.1
```

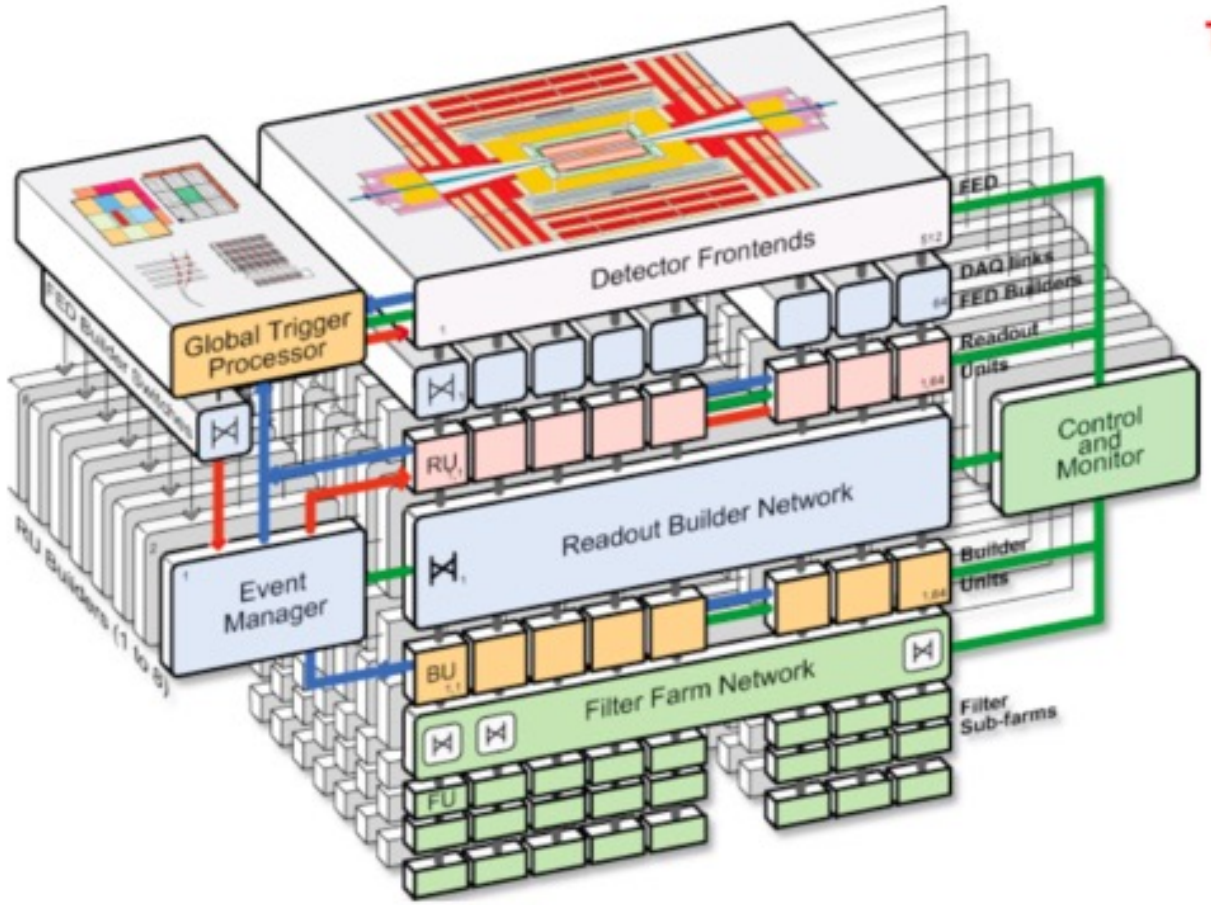
GPD High-Level Trigger

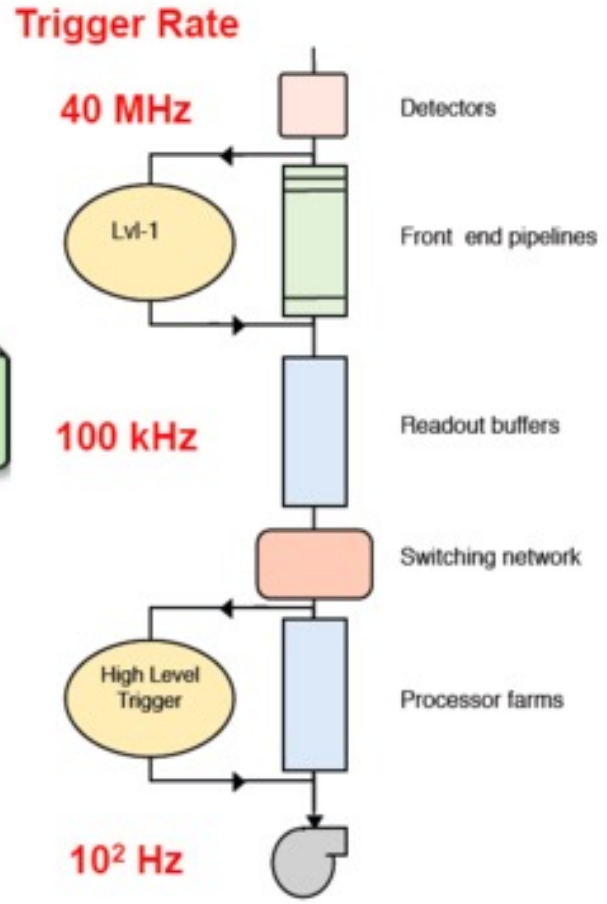




- Multiple higher trigger levels are implemented in software
  - More 'funnel' than 'filter' – implemented on CPUs (~2MW power by 2025)
  - ATLAS has an additional L2 stage using full granularity in RoI
- Note that few events are 'fully reconstructed'
  - Throw everything away as fast as possible; 'no trigger' is not an option
  - Unpacking and basic manipulation of raw data is a large overhead
- Future HLT upgrades will involve co-processing of various types

GPD DAQ



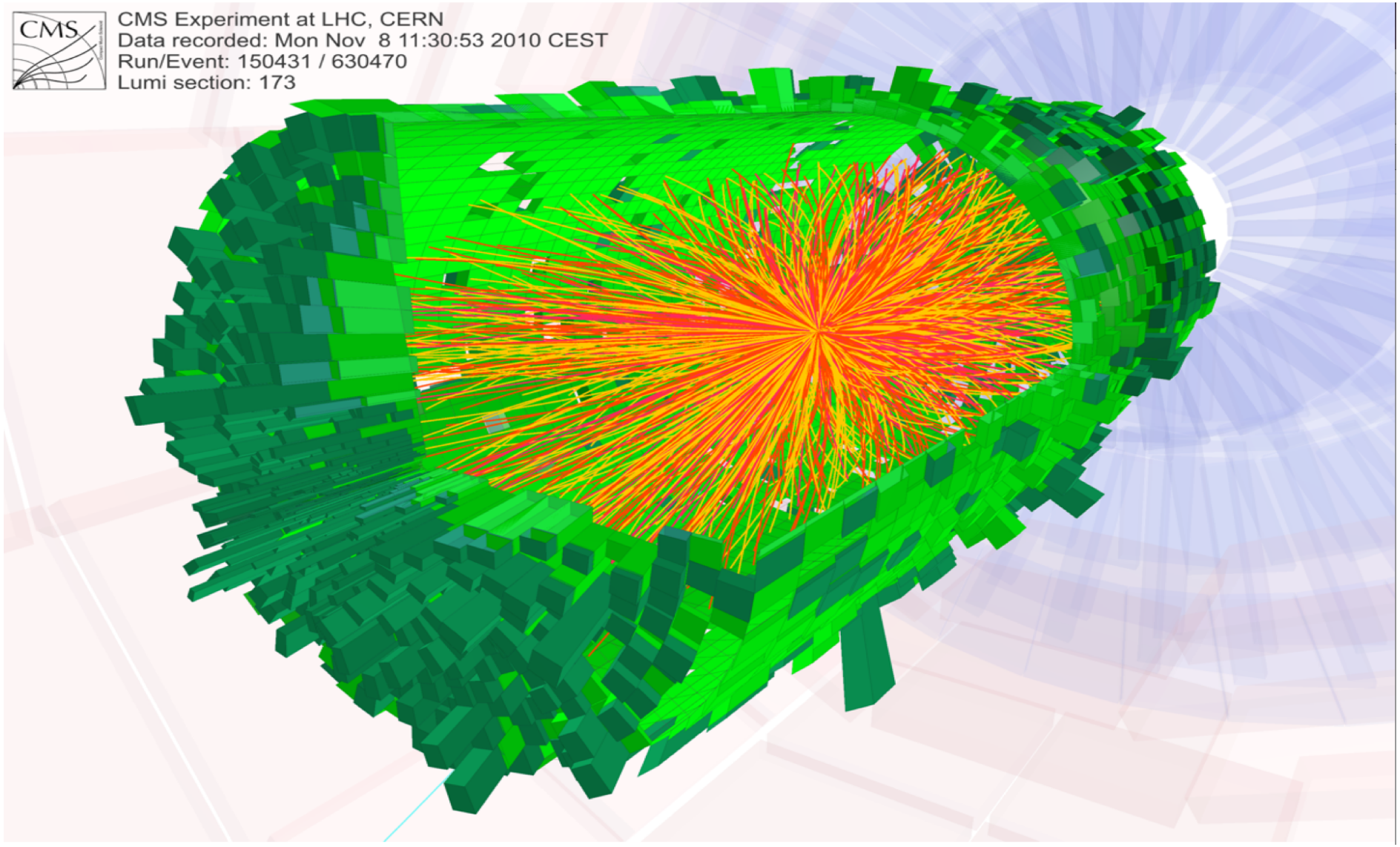


- Note the presence of multiple independent *partitions*

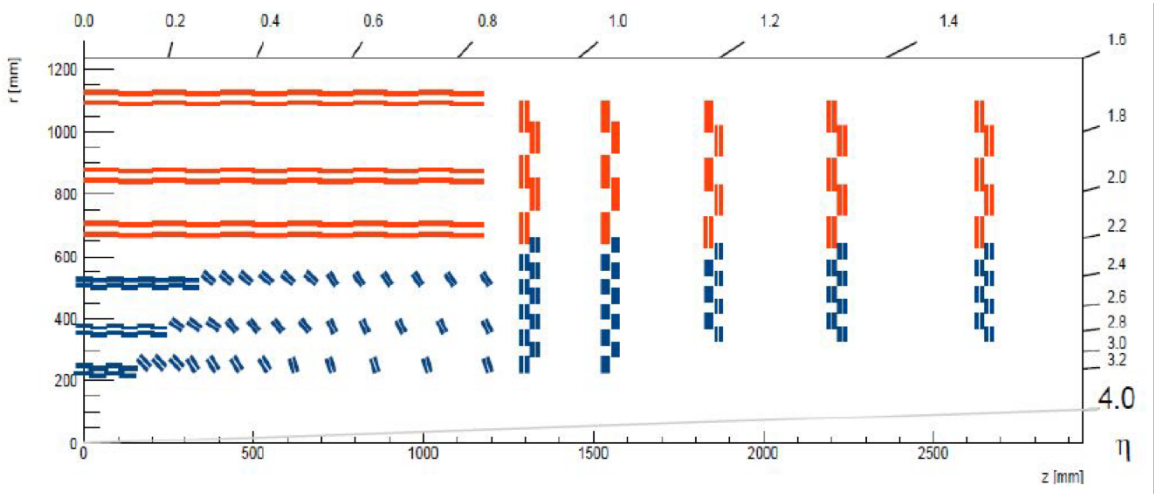
Future Challenges

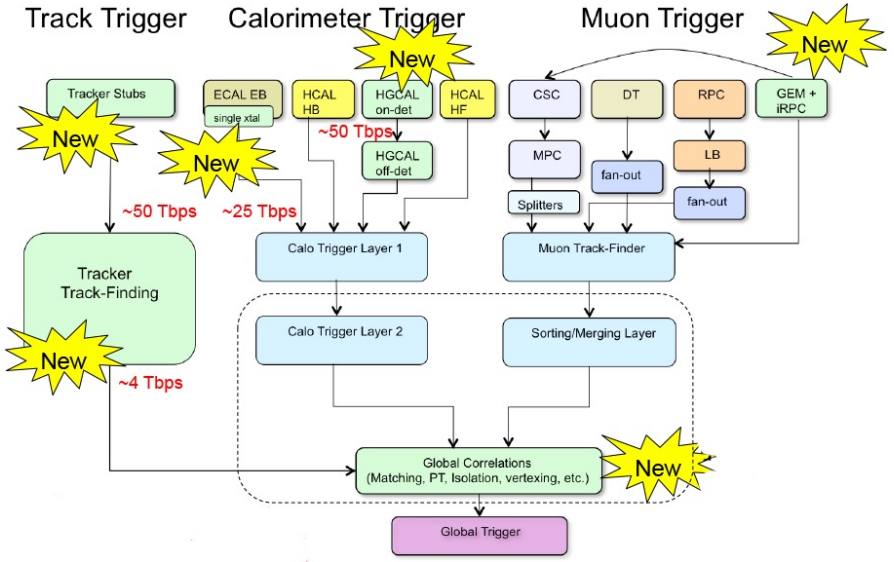
- Future projects bring new problems
- HL-LHC (10x lumi increase)
  - Far more detector occupancy, existing trigger algorithms will not scale well
  - Use of tracking information in trigger now seems mandatory
- Ultra-large long baseline neutrino experiments
  - Not just for oscillation measurements, also astrophysics observatories
  - Channel counts approaching LHC Run-1, signal-to-noise low
- Ultra-high-beam-rate 'rare process' experiments
  - e.g. mu2e, COMET, NA62, SHiP
  - Data rates are large, budgets not LHC-like
- Future colliders
  - e+e- / FCChh: can we go triggerless?
  - Detectors will have very high channel count, severe power limitations
    - Readout bandwidth will be the limitation here

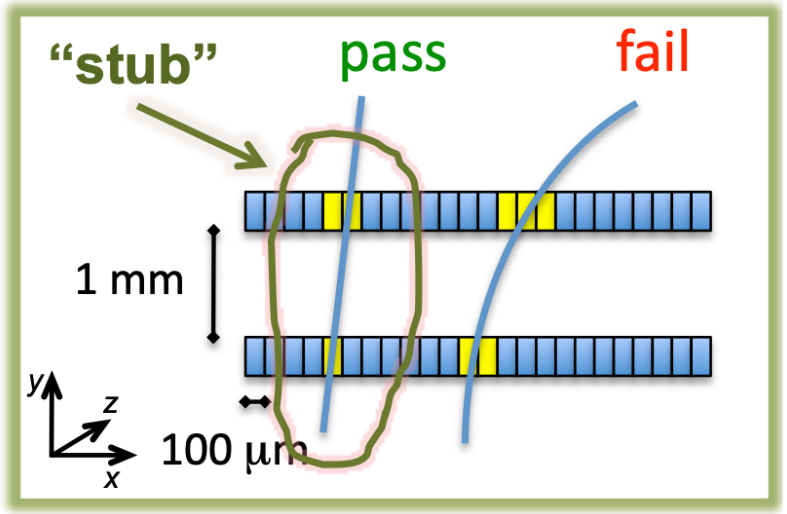
This is what 'Challenging' Looks Like



LHC Phase-2: Track Triggering



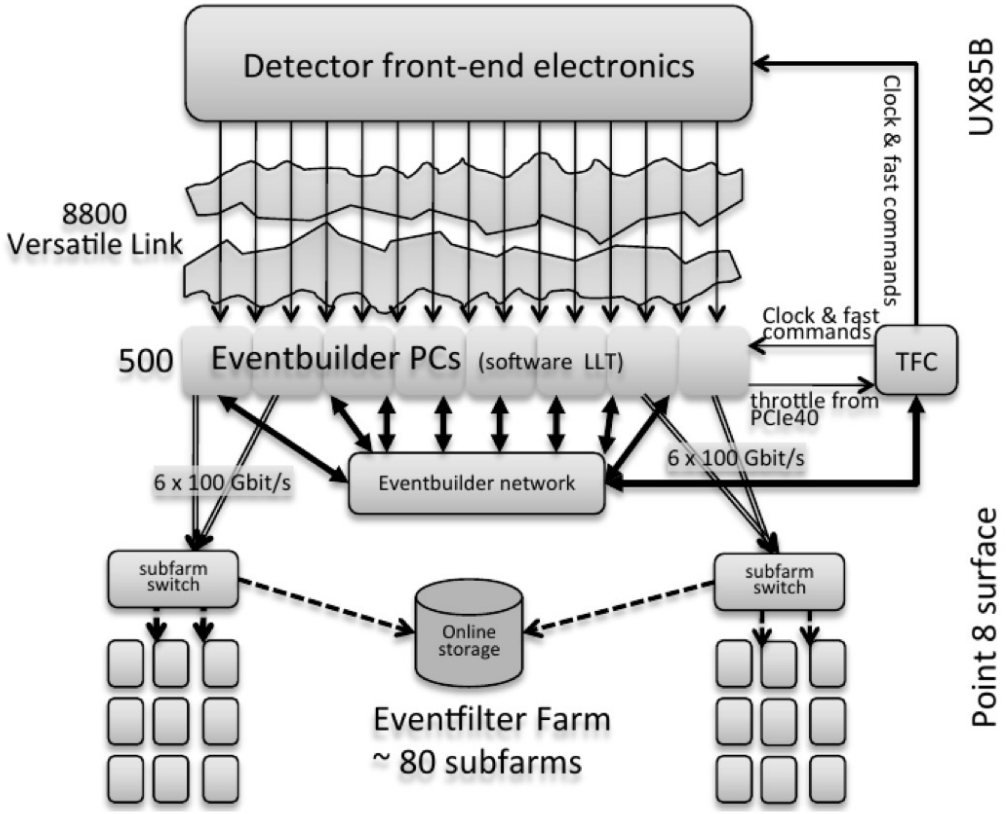




Only high-pt stubs
are sent off-detector

- CMS: Emphasis is on-detector data reduction
  - First stage of L1 tracking in front-end ASICs; second stage in FPGA custom hardware

LHCb: What, no Trigger?



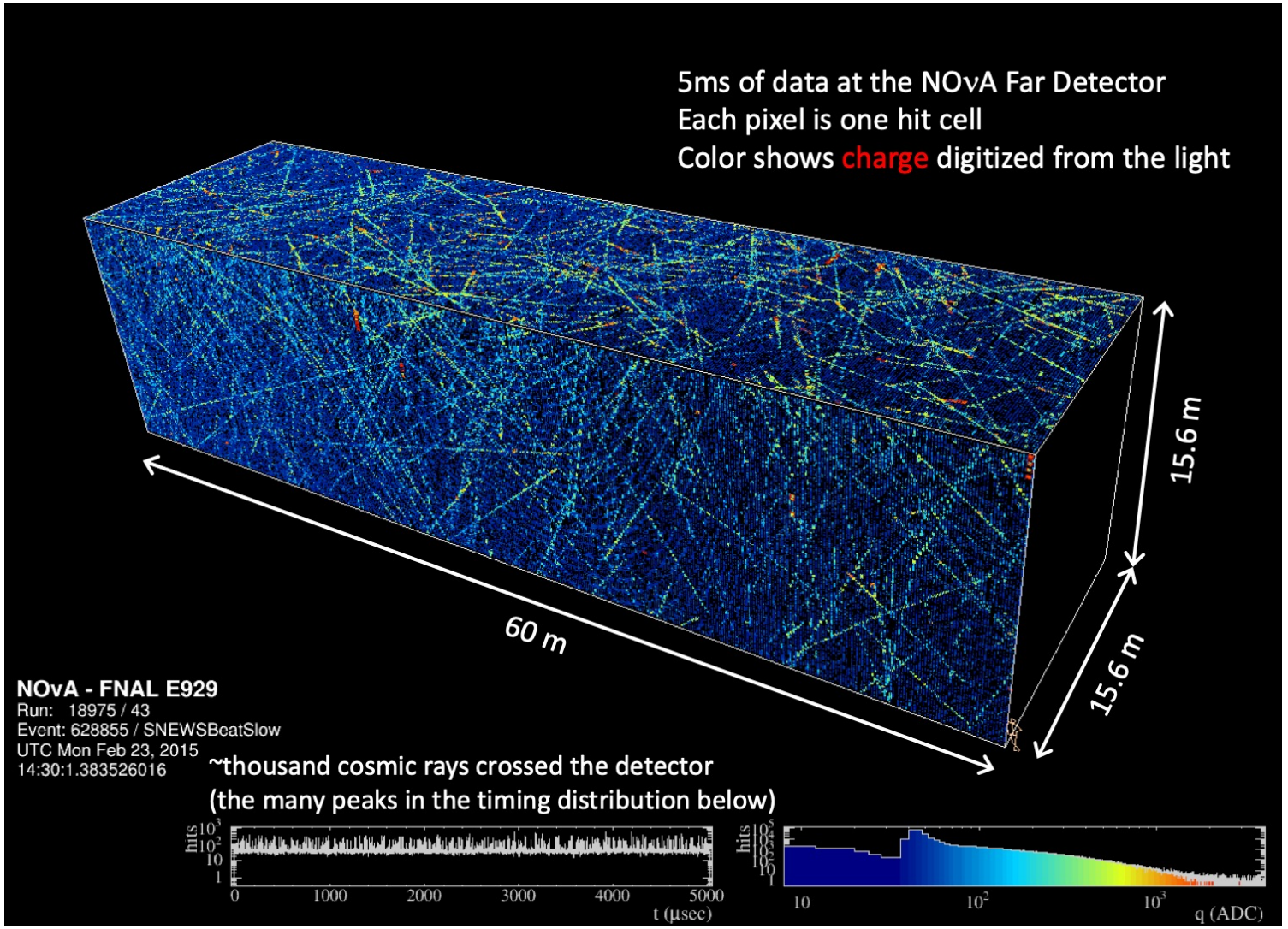
- LHCb taking 'no L1 trigger' approach from LHC Run-3
  - Full software trigger, all data read from detector at 40MHz rate
    - Some 'hardware assist' for data unpacking is also proposed – watch the power! (>2MW)
  - 30Tb/s COTS (ish) network (10x CMS / ATLAS) – feasible now

Space COTS Data Acquisition

Successful COTS Solution for Space Vehicle Onboard Recording & Audio Handling

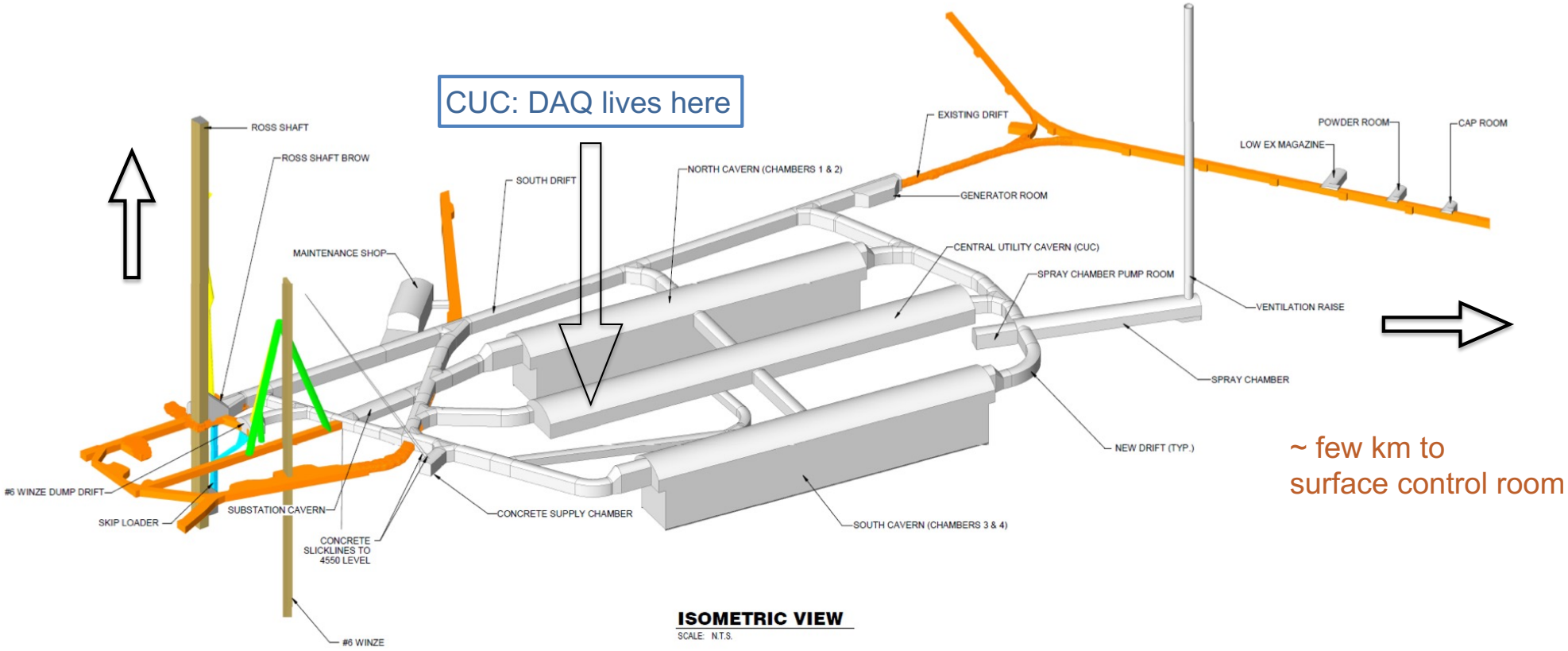
A space vehicle designer was looking for a data acquisition solution that could collect all of the data from the onboard computers. Given that the vehicle was designed for applications such as resupplying the international space station, it was necessary that the equipment important for the mission met stringent reliability requirements, including predictable operation in a low earth orbit (LEO) radiation environment. The solution was to use Space qualified COTS, which were believed to be about three times less expensive than the traditional approach.

NOvA: Neutrinos – Somewhere?



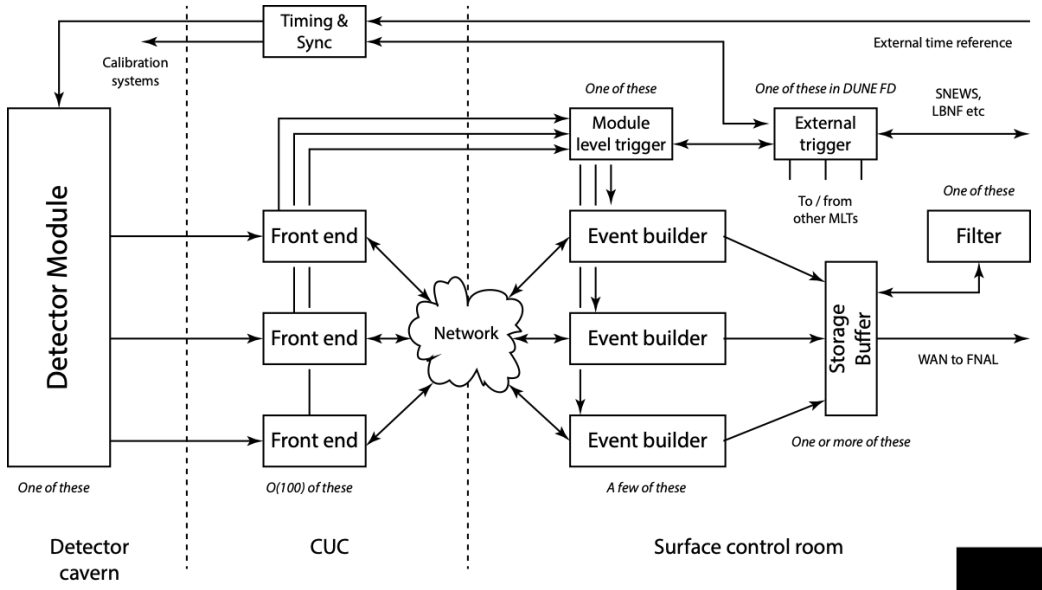
- Long-baseline neutrino expt, 15kT, 400k channels, *on surface*
  - Reconstruct background, the rest is signal!

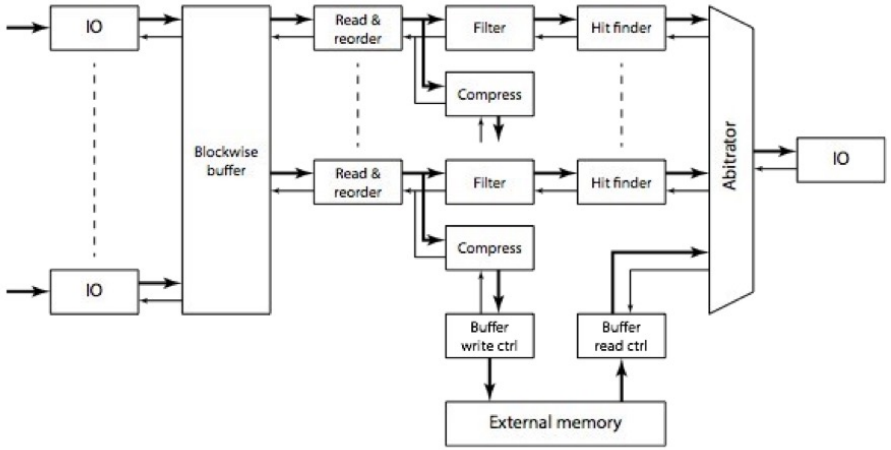
DUNE: Neutrinos, with Less Background

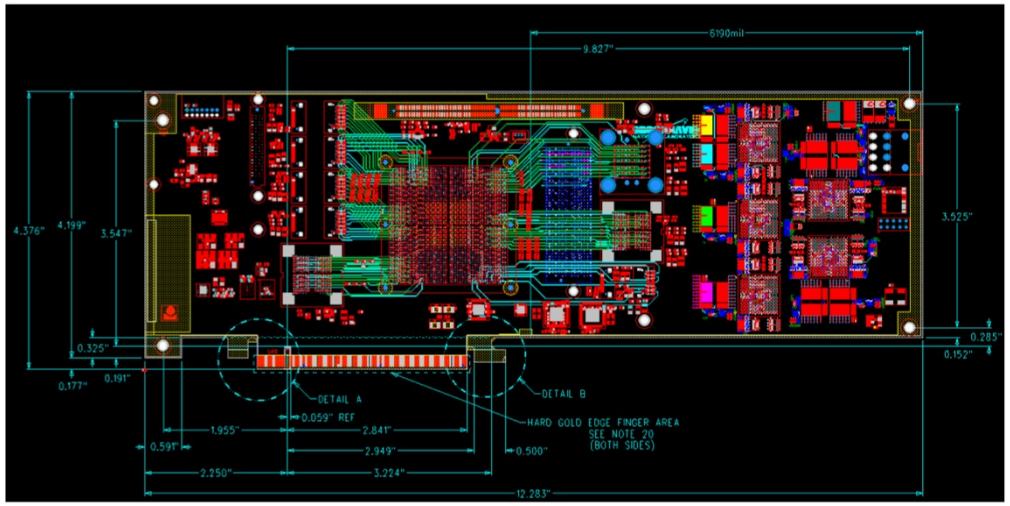


- DUNE: 40kT LAr TPC, 50Tb/s readout rate, 6GB event size, 30PB/yr
  - No on-detector data reduction (electronics is at LAr temperature, power is tight)
- DAQ split between 4850ft level and surface; minimise traffic to surface
  - Strong constraints on cooling (450kW), space (~50 racks) for all four modules
  - Strong schedule constraints on installation and commissioning

DUNE

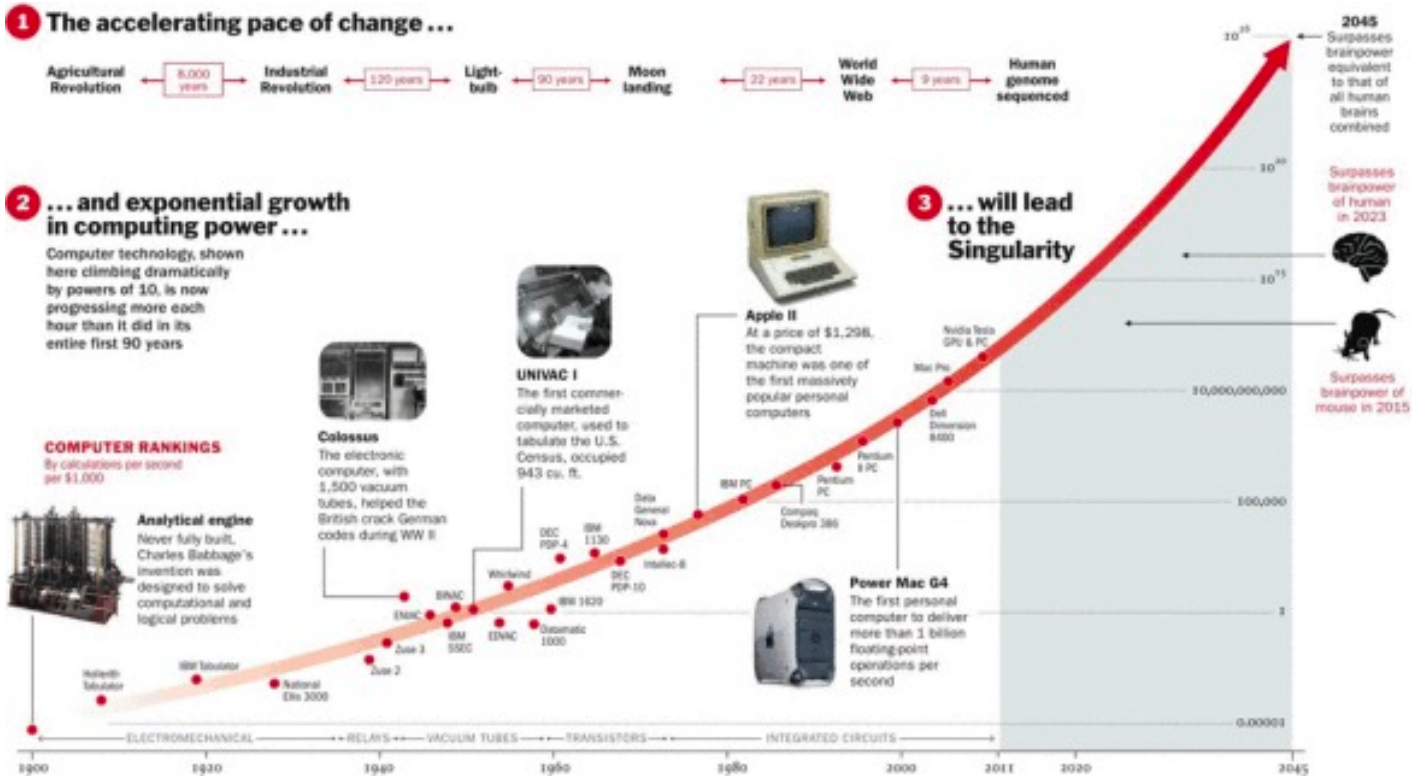






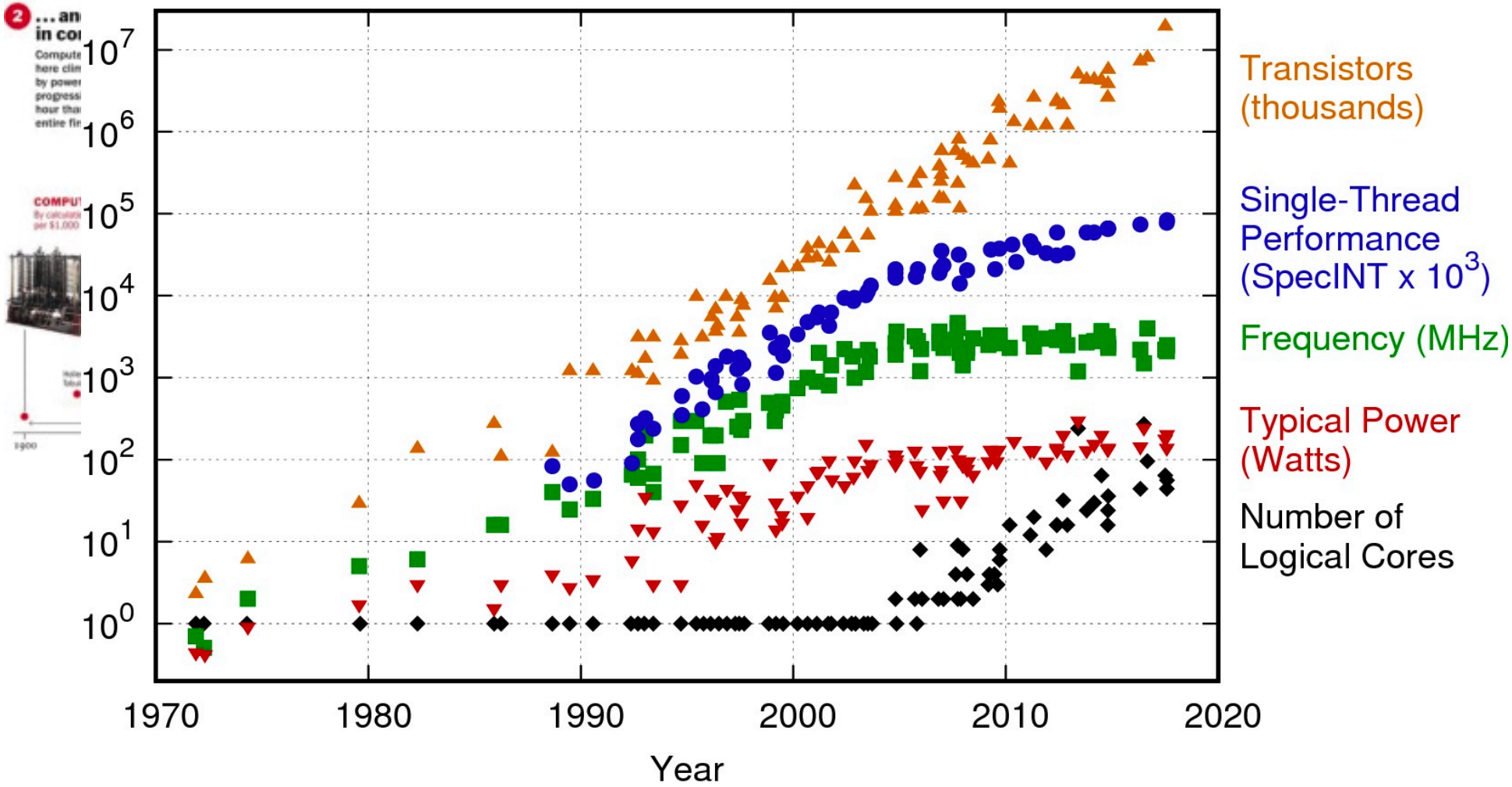
- DUNE DAQ *intentionally* similar to LHC Phase-2 systems
  - FPGA stream-processing front end
  - Full software 'data selection system'
    - There are no events per se in an LAr TPC
  - ArtDAQ back end software
- Lossless data compression is the key
  - Detector occupancy is low, but SNB trigger requires ~full information with no ZS
- Re-use / co-devt of ATLAS FELIX readout; but reliability demands are very high

Moore's Law…


1 The accelerating pace of change …
Agricultural Revolution
8,000 years
Industrial Revolution
120 years
Light-bulb
90 years
Moon landing
22 years
World Wide Web
9 years
Human genome sequenced
2 … and exponential growth in computing power …
Computer technology, shown here climbing dramatically by powers of 10, is now progressing more each hour than it did in its entire first 90 years
COMPUTER RANKINGS
By calculations per second per $1,000
Analytical engine
Never fully built, Charles Babbage's invention was designed to solve computational and logical problems
Colossus
The electronic computer, with 1,500 vacuum tubes, helped the British crack German codes during WW II
UNIVAC I
The first commercially marketed computer, used to tabulate the U.S. Census, occupied 943 cu. ft.
Apple II
At a price of $1,298, the compact machine was one of the first massively popular personal computers
Power Mac G4
The first personal computer to deliver more than 1 billion floating-point operations per second
3 … will lead to the Singularity
2045
Surpasses brainpower equivalent to that of all human brains combined
Surpasses brainpower of human in 2023
Surpasses brainpower of mouse in 2015
ELECTROMECHANICAL
RELAYS
VACUUM TUBES
TRANSISTORS
INTEGRATED CIRCUITS
1900
1920
1940
1960
1980
2000
2011
2020
2045

Moore's Law… is Dead

42 Years of Microprocessor Trend Data



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2017 by K. Rupp

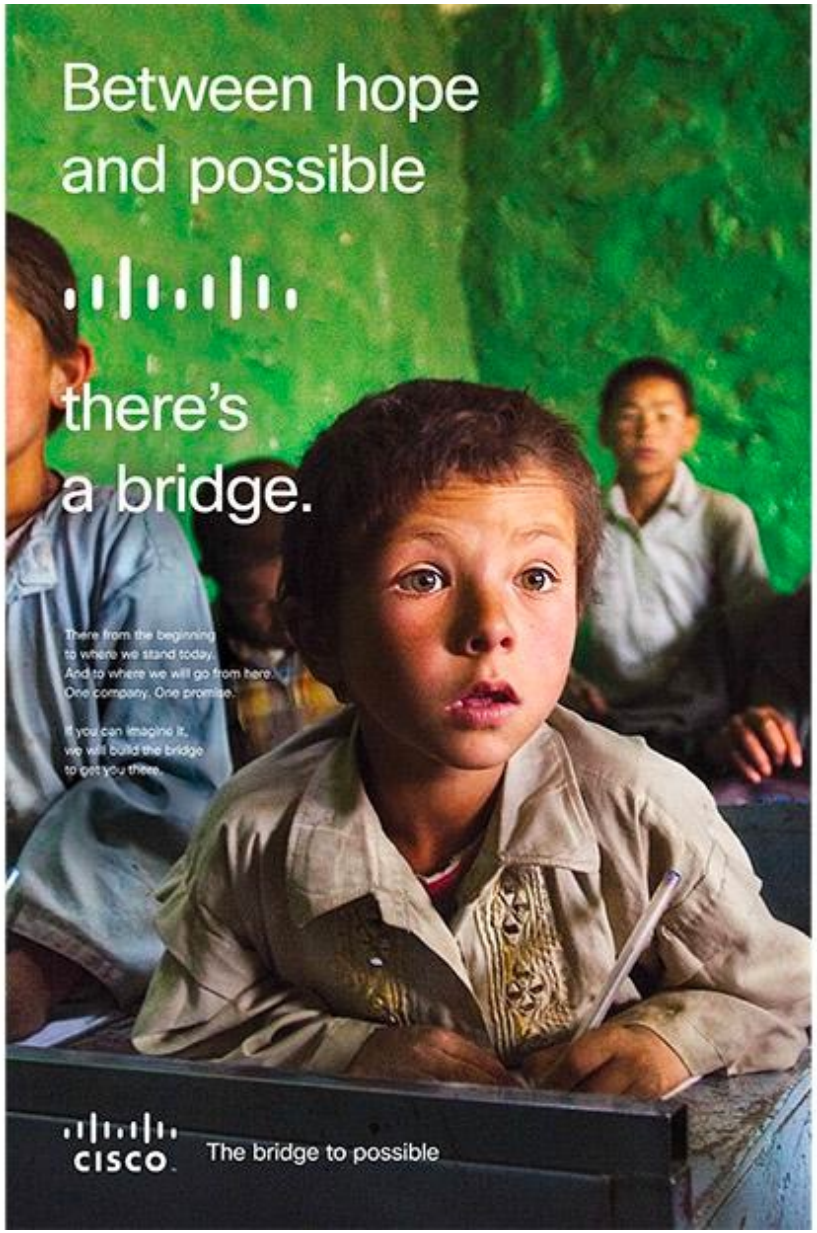
Don't Panic Just Yet

TABLE 1. Summary of techology options for extending digital electronics.

Improvement Class	Technology	Timescale	Complexity	Risk	Opportunity
Architecture and software advances	Advanced energy management	Near-Term	Medium	Low	Low
	Advanced circuit design	Near-Term	High	Low	Medium
	System-on-chip specialization	Near-Term	Low	Low	Medium
	Logic specialization/dark silicon	Mid-Term	High	High	High
	Near threshold voltage (NTV) operation	Near-Term	Medium	High	High
3D integration and packaging	Chip stacking in 3D using thru-silicon vias (TSVs)	Near-Term	Medium	Low	Medium
	Metal layers	Mid-Term	Medium	Medium	Medium
	Active layers (epitaxial or other)	Mid-Term	High	Medium	High
Resistance reduction	Superconductors	Far-Term	High	Medium	High
	Crystaline metals	Far-Term	Unknown	Low	Medium
Millivolt switches (a better transistor)	Tunnel field-effect transistors (TFETs)	Mid-Term	Medium	Medium	High
	Heterogeneous semiconductors/strained silicon	Mid-Term	Medium	Medium	Medium
	Carbon nanotubes and graphene	Far-Term	High	High	High
	Piezo-electric transistors (PFETs)	Far-Term	High	High	High
Beyond transistors (new logic paradigms)	Spintronics	Far-Term	Medium	High	High
	Topological insulators	Far-Term	Medium	High	High
	Nanophotonics	Near/Far-Term	Medium	Medium	High
	Biological and chemical computing	Far-Term	High	High	High

- But: consistent marketing-driven underestimation of time to market
  - Where's my memristor hard drive? (or my holographic memory?)

New Tech – but When?

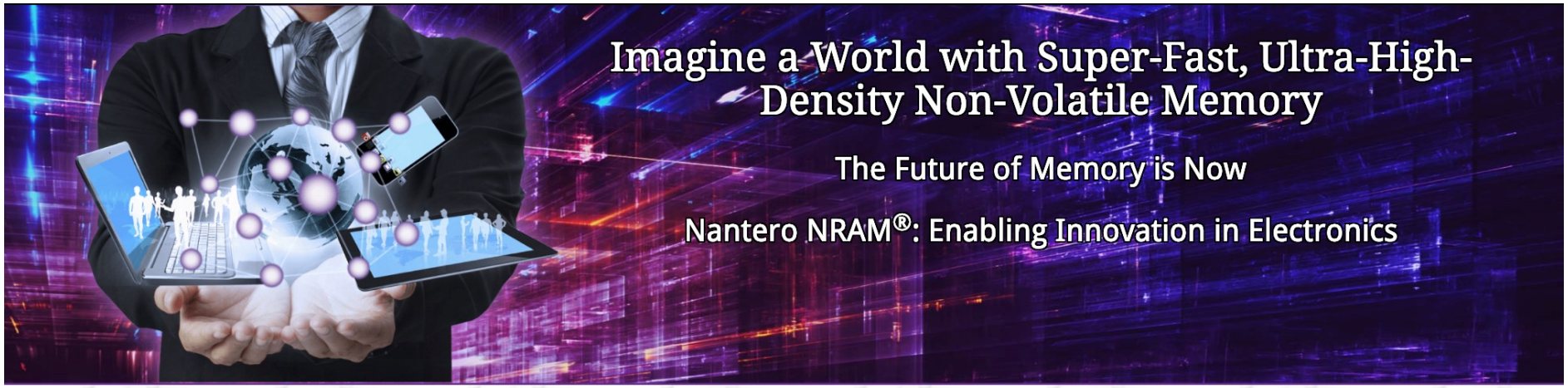


New Tech

- What could really make a difference?
- Co-processing without the pain
  - GPU programming is bad enough, FPGA interfacing is horrible, ISExt is voodoo
  - HLS is not the answer for low-level functionality; you have to understand the chip
- Open source firmware / ASIC development tools
  - Or even just good development tools of any sort
- Low power, low mass, high speed data links
  - Fixing the readout bandwidth limitations needs mW per Gb/s
  - Photonics, wireless, free-space optics – but needs to be rad-hard
  - Appears to be no commercial application – yet
- SSDs without wear-out
  - Much promised, not much seen working in practice
- Filling the 'RAM gap' – the big one
  - Between (fast, expensive, small, random) DRAM / (slow, cheap, large, block) mass storage
  - Would directly address most of our buffer needs very efficiently
  - Much promised, nothing delivered

An 'ASIC crisis' may be upon us

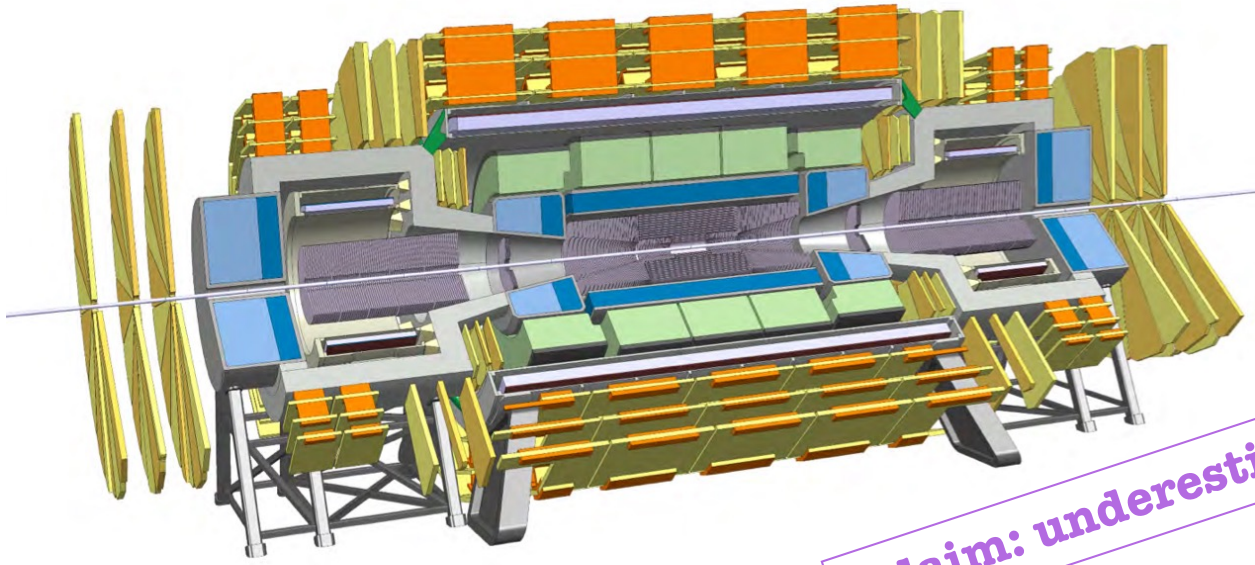
New Tech – but When?



- In the approximation 'now' = '2030'

Going to Extremes: FCChh

Reference detector for the CDR



- 4T 10m solenoid
- Forward solenoids
- Silicon tracker
- Barrel ECAL Lar
- Barrel HCAL Fe/Sci
- Endcap HCAL/ECAL LAr
- Forward HCAL/ECAL LAr

I claim: underestimated

$L = \sim 2.5 \times 10^{35}$; $E = 100\text{TeV}$

Front end detector data rates:

- **Tracker : ~800 TB/s**
- **LAr+Tile Calo : ~200 TB/s**
- **Si/W Calo : 5x LAr+Tile Calo ?**

Implications:

- **>1 M optical fibers @ 10Gb/s**
- **> 10 Pb/s event builder network**
- **Material budget to extract the data and bring in power and cooling**
- **Processing farm requirements (CPU, power and cooling)**

Options:

- Triggerless continuous readout
- 1-Level Triggered readout
- Multi-Level Trigger, Regional Readout schemes

Going to Extremes: FCChh

WARNING: Britain faces internet RATIONING as UK grid struggles to cope with web demand

BRITONS could see the unwelcome return of RATIONING as the UK power grid and communications network struggles to cope with the country's growing demand for internet access, an expert has warned.

By AARON BROWN
PUBLISHED: 07:42, Wed, May 6, 2015 | UPDATED: 16:51, Fri, May 8, 2015

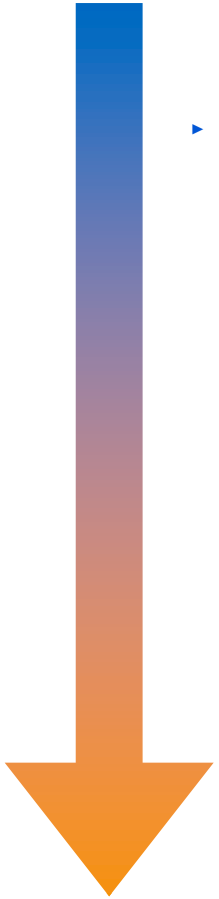
SHARE TWEET 425 22

The days of unlimited internet access may be numbered

- Data communications now use >2% of UK electricity supply
  - The servers use less power than the infrastructure…
- With today's tech, FCChh requires >2MW *on detector* to transmit data
  - Of course, today we are at ~20x the fundamental limits of power efficiency

Possible Approaches

- 'Conventional trigger'
  - Extreme processor performance
  - On-detector primitives logic
  - On-detector front end buffers
  - Emphasis on on-detector processing



- 'Sequential readout'
  - Stage out event to multi-level trigger
  - Successive levels of details with time
  - All data through event-builder network
  - Trigger implemented in software
  - Implement large 'bulk memory' in low radiation zone of detector
  - Emphasis on on-detector buffer

- 'Triggerless'
  - Massive bandwidth
  - Little on-detector logic
  - Small front end buffers
  - Emphasis on data transmission

That RAM gap again!

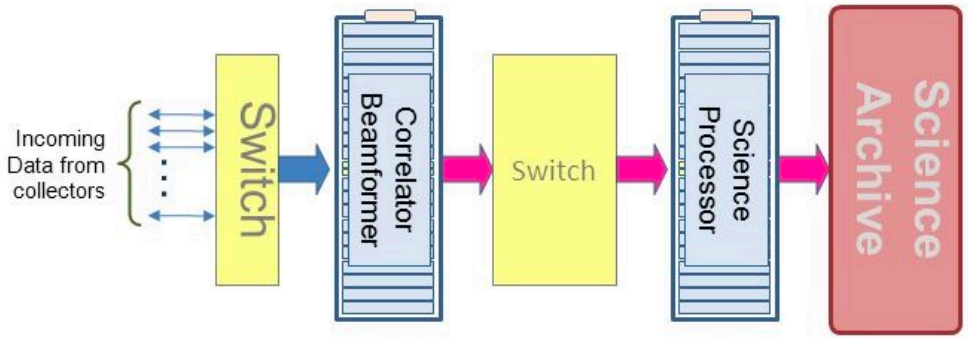
New R&D, New Approaches

- R&D we really need
  - Low-power links, signalling beyond NRZ
    - Me, I like the mm-wave power + signal approach aka 'wireless tracker' - photonics also interesting
  - Advanced statistical methods (MVA/ML) for global triggering
    - Surely this is low-hanging fruit for the educated youth of today?
  - Lossless and lossy compression algos for detector signals
  - Common firmware development and verification tools -> code re-use
  - Fault-tolerant, self-healing cluster-distributed online processing pools
    - If you can solve this, please come to work on DUNE
  - Follow closely the new storage devices (ReRAM, NRAM, 3d Xpoint, etc)
- New approaches we really need
  - Stop re-inventing the wheel (hardware, firmware, software)
    - Mature common projects (e.g. test beam DAQ) are out there – let's focus
  - Open source everything: tools, code, hardware
  - Engage a new generation of researchers in the interesting problems
    - This means development, not just tuning of the parameters
  - Find commonalities with other disciplines (mainly: astrophysics)

Other disciplines have very challenging DAQ requirements

The Data Challenge

	MeerKAT	SKA Phase 1	SKA Phase 2*
Into Correlator	2 Tbps	50 Tbps	up to 5 Pbps
Into Science Processor	0.4 Tbps	20 Tbps	up to 500 Tbps
Into Archive	35 Gbps	300+ Gbps	up to 2 Tbps
Compute load	200 TFlops	30+ PFlops	3+ EFlops



* SKA Phase 2 data rates are still fairly speculative

- Can also mention FAIR, XFEL, synchrotron sources, high intensity lasers, etc etc

TDAQ Conclusions

- Happy birthday to TDAQ
  - Still interesting after 100 years
- TDAQ developments underlie all major experiments
  - Success in this area is vital to enable the science
- The systems are more complex than they look on the surface
  - Ideally, you only notice them when they don't work
  - A lot of electronics, modelling, computing science, pragmatism is needed
- Next generation of projects will be harder than LHC
  - Not necessarily 'bigger', but demanding in other ways
  - New tech will help, but get ready to wait…
- Significant R&D needed towards future facilities
  - Ideally by a new generation of experts