Read-out developments for the

LHCP 2024, Boston

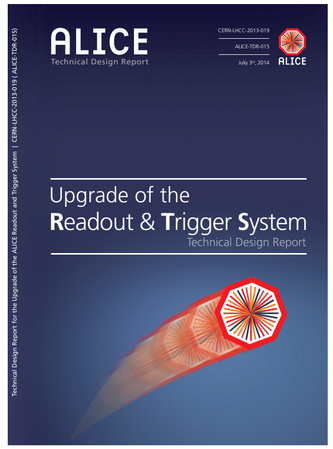
Jeroen Hegeman

Photo: © CERN

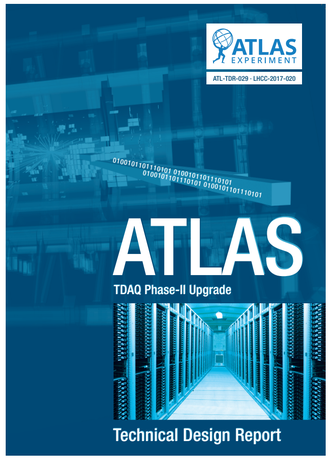
Setting the context

- **This will focus on the HL-LHC upgrades of the large LHC experiments, with an attempt to `place everything in context'**
  - This naturally leads to a look at the ECFA DRD7 efforts

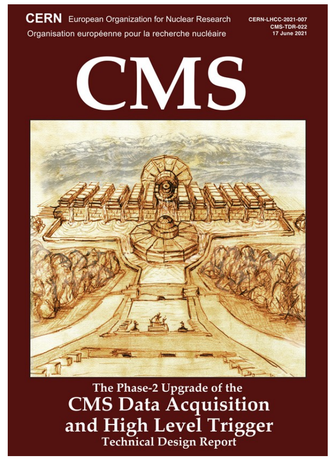
- I am grateful to the experts who shared their thoughts with me in preparation of this talk
  - The facts come from them, and from various publications
  - Interpretations and opinions were added by me (Misinterpretations and misguided opinions will be mine too)



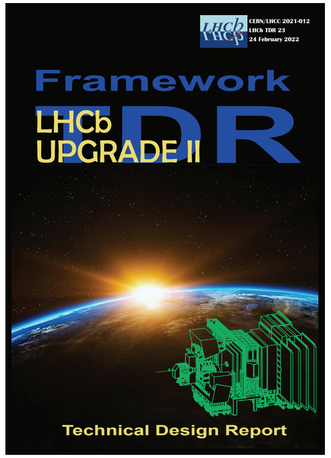
CERN-LHCC-2013-019



CERN-LHCC-2022-004



CERN-LHCC-2021-007

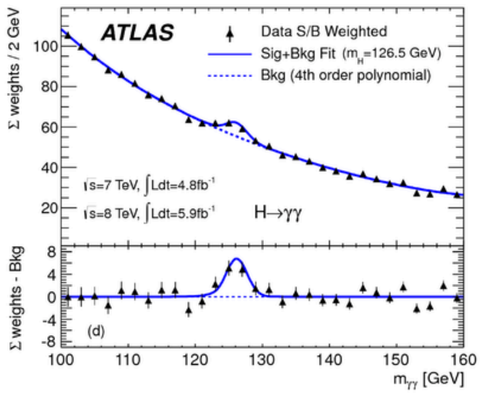


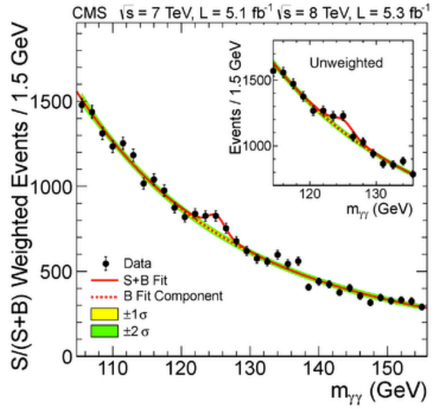
CERN-LHCC-2021-012

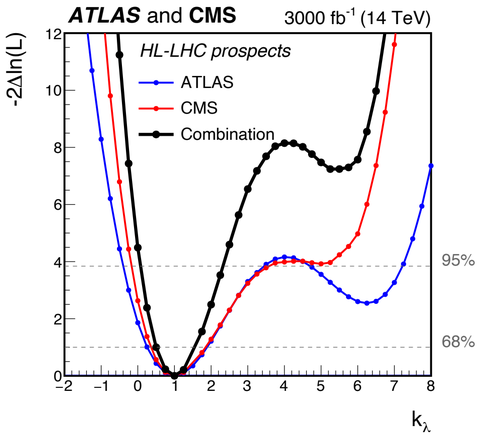
These developments are all driven by one thing

Continuing our search for smaller needles in larger haystacks

- More (rare events) and (more rare) events
- More statistics → more luminosity
  And really: higher instantaneous luminosity
- More granularity → more channels
- **Bottom line: more data, faster**
  → Larger events at a higher rate
  → More complex online reduction required

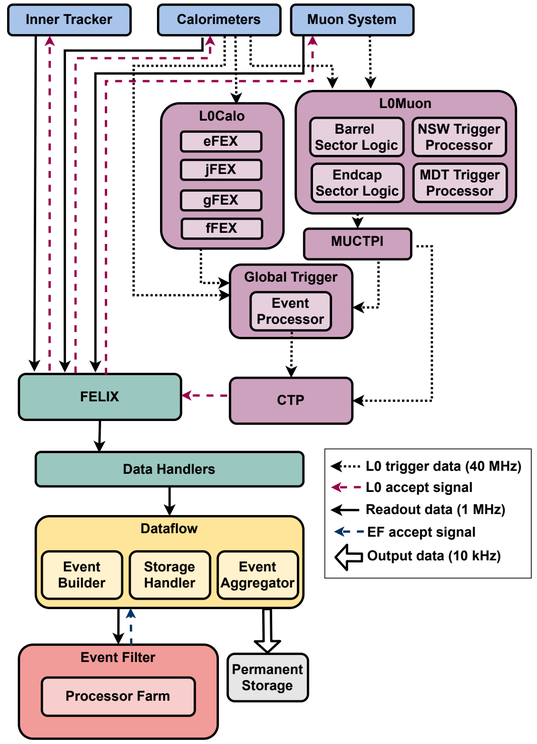




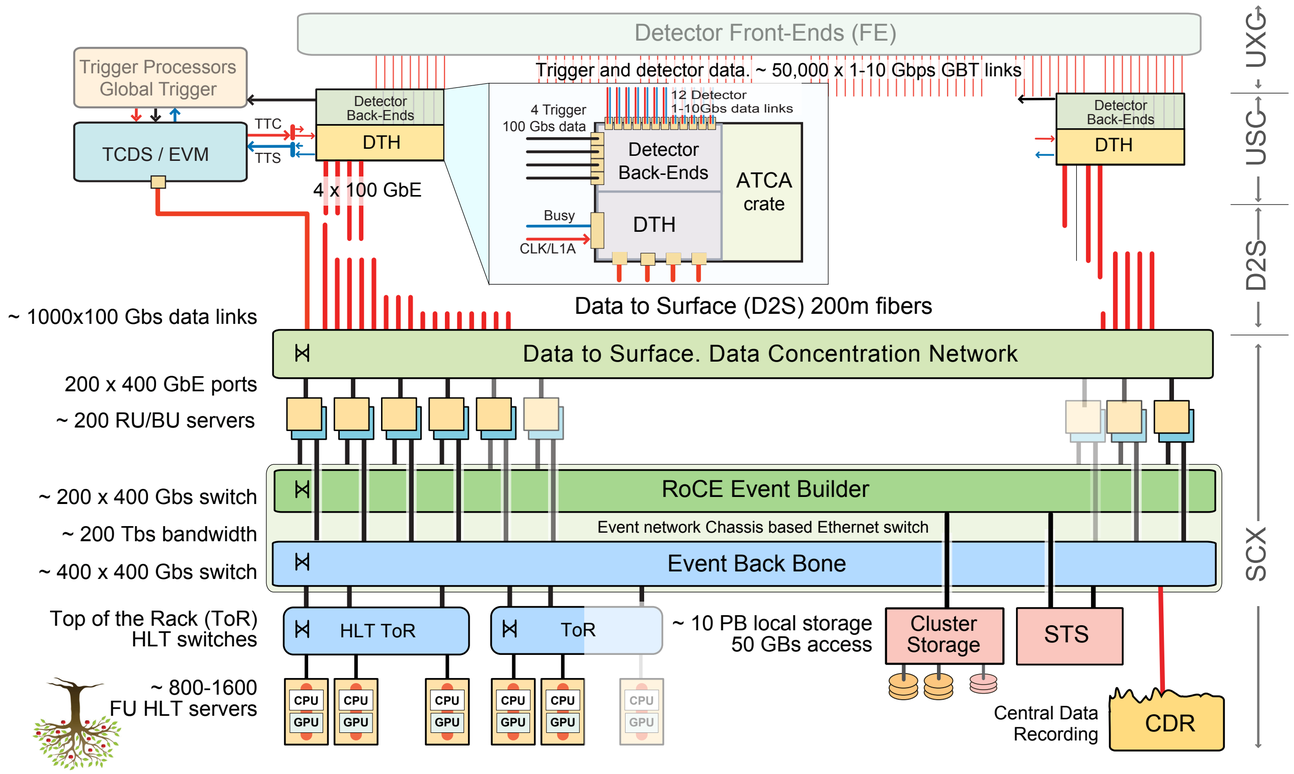


P.S. The detector read-out is by no means the biggest challenge here.

Traditional HEP experiment trigger-DAQ architecture

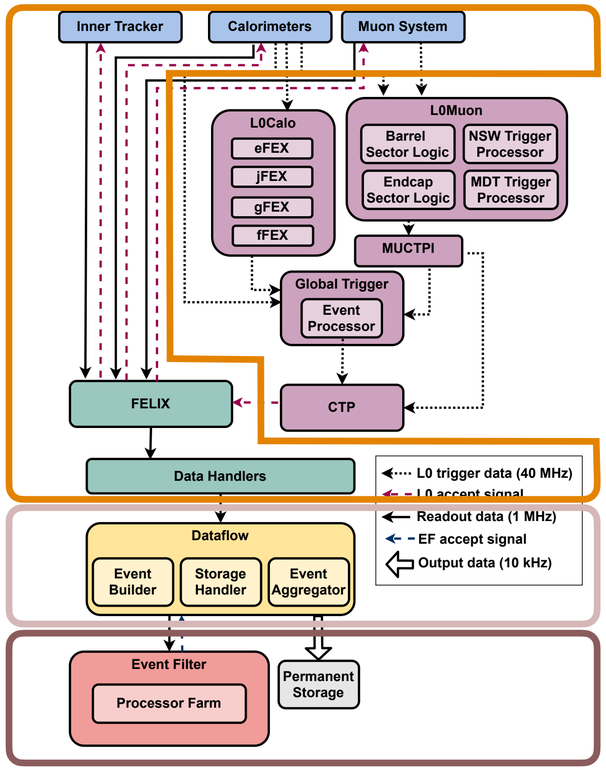


Phase-2 ATLAS

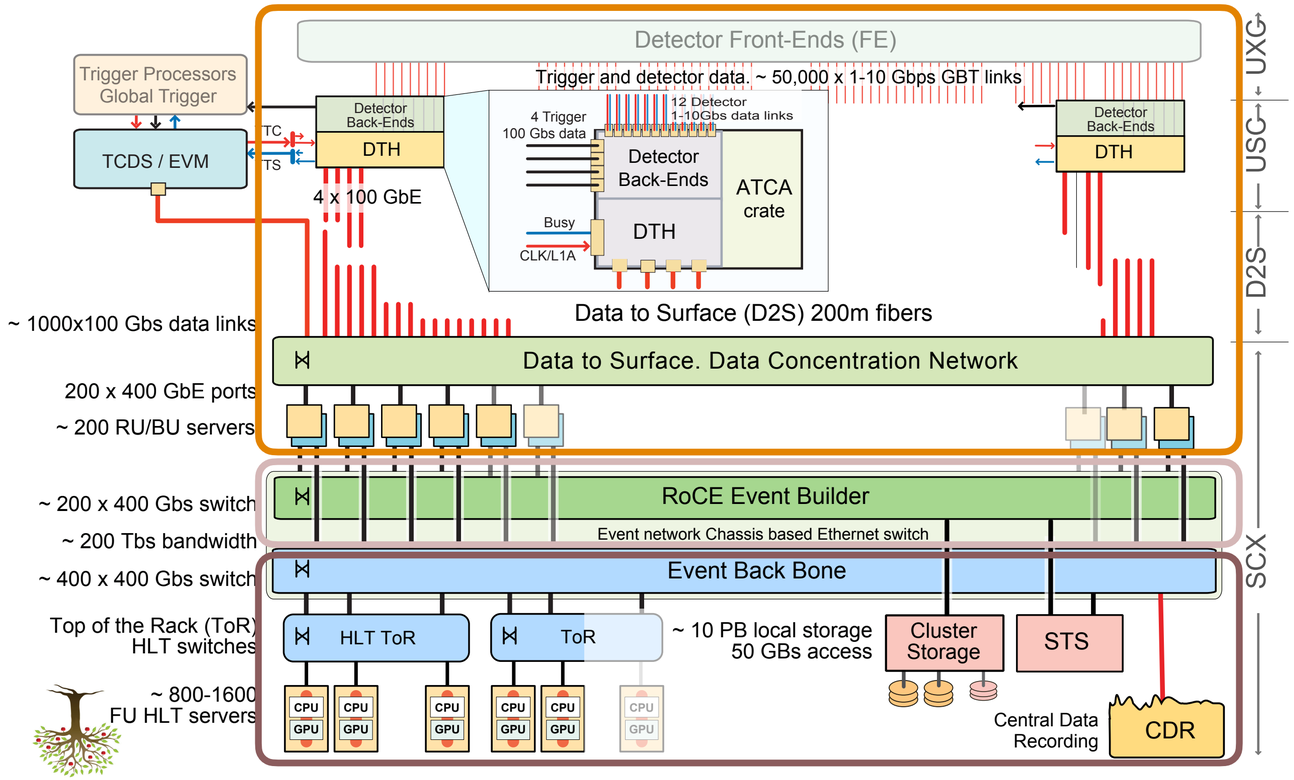


Phase-2 CMS

Traditional HEP experiment trigger-DAQ architecture



Phase-2 ATLAS



Phase-2 CMS

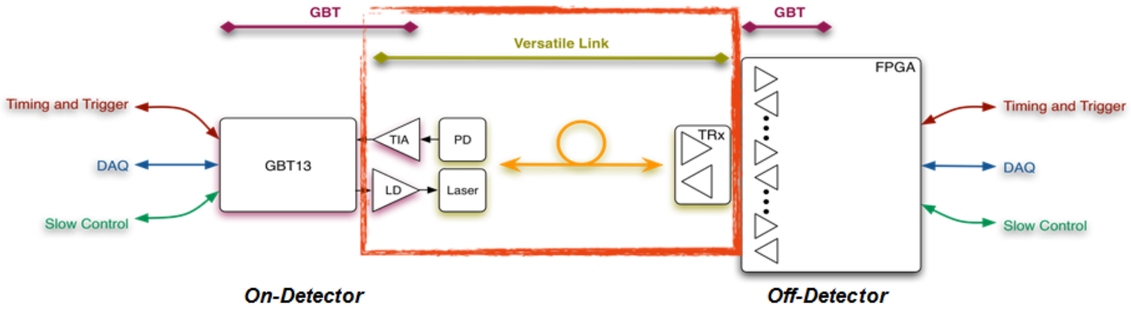
The concepts involved have not changed fundamentally for a while now:

data read-out, event building, event filtering and storage

Zoom on the front-end

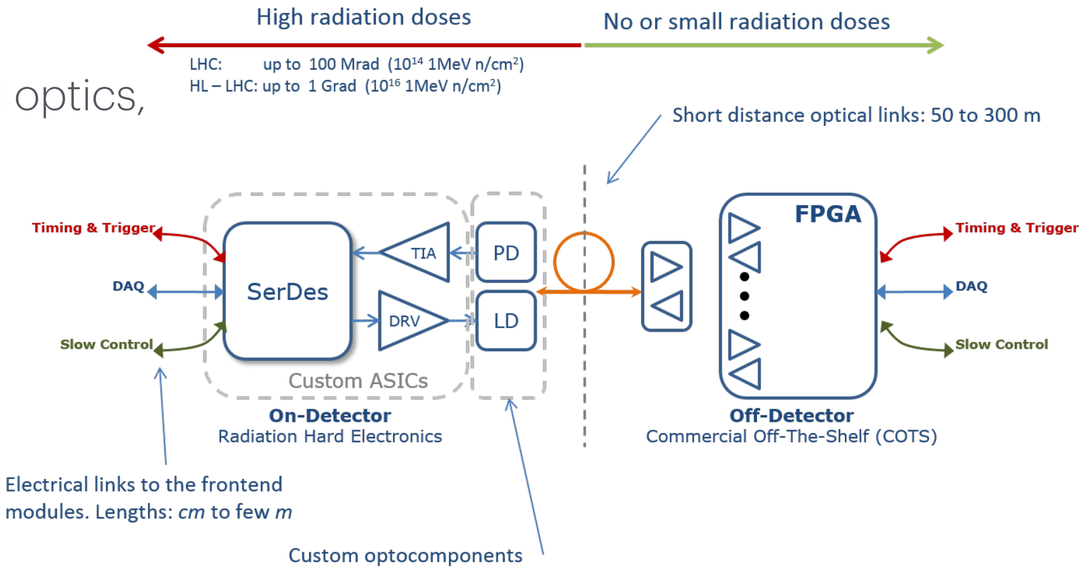
Getting the data off the detector

- The Phase-1 **Versatile Link**: first general front-end optical link solution (4.8 Gb/s, 10 kGy)
  - Combines fast control and slow control (downlink), as well as DAQ (uplink)
  - Pair of front-end ASICs (GBT and GBT-SCA) and front-end optics, **plus selection guides and application notes for implementation with COTS back-end optics**



Getting the data off the detector

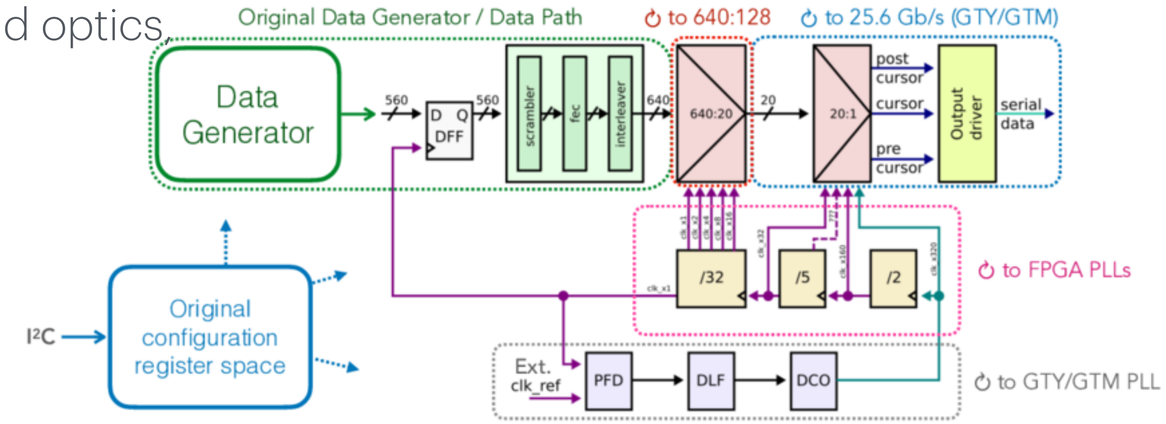
- The Phase-1 **Versatile Link**: first general front-end optical link solution (4.8 Gb/s, 10 kGy)
  - Combines fast control and slow control (downlink), as well as DAQ (uplink)
  - Pair of front-end ASICs (GBT and GBT-SCA) and front-end optics, **plus selection guides and application notes for implementation with COTS back-end optics**



- Phase-2 (i.e., HL-LHC) brings the **Versatile Link+**
  - Supercharged performance: 2.56 Gb/s downlink, 10.24 Gb/s uplink, 1 MGy TID
  - Front-end ASIC (lpGBT) and optics (VTRX+), **plus validated commercial back-end optics**
  - **Unique challenge: Pairing custom and commercial components over the full front-end lifetime**

Getting the data off the detector

- The Phase-1 **Versatile Link**: first general front-end optical link solution (4.8 Gb/s, 10 kGy)
  - Combines fast control and slow control (downlink), as well as DAQ (uplink)
  - Pair of front-end ASICs (GBT and GBT-SCA) and front-end optics, **plus selection guides and application notes for implementation with COTS back-end optics**



- Phase-2 (i.e., HL-LHC) brings the **Versatile Link+**
  - Supercharged performance: 2.56 Gb/s downlink, 10.24 Gb/s uplink, 1 MGy TID
  - Front-end ASIC (lpGBT) and optics (VTRX+), **plus validated commercial back-end optics**
  - **Unique challenge: Pairing custom and commercial components over the full front-end lifetime**

- Next-generation R&D is already ongoing
  Demonstrator ASIC (DART28) driving Silicon Photonics for a 100Gb/s radiation tolerant front-end link
  Target: compatible with COTS back-end solutions

Getting the data off the detector

The enormous growth in performance of these optical serial links was made possible by the adoption of industry tools and techniques, and their adaptation to our specific use case.

- HEP communications ASICs and on-detector optics are nowadays developed by dedicated collaborations of specialists in institutes and universities, developing common components and giving support to users. Small-scale link developments are no longer viable.

- Next-generation R&D is already ongoing
  Demonstrator ASIC (DART28) driving Silicon Photonics for a 100Gb/s radiation tolerant front-end link
  Target: compatible with COTS back-end solutions

Zoom on the back-end

The back-end environment throughout the years

- Originally, LHC off-detector electronics were predominantly VME (6U and 9U), and some CompactPCI
  - Boards were relatively sparsely populated
  - Layout dominated by peripherals
- The Phase-1 upgrade saw a trend towards MicroTCA
  - Significant increase in logic density and component miniaturisation made smaller boards possible
  - Access to peripherals became a challenge
- The Phase-2 upgrade introduced ATCA as crate system
  - Larger boards, and more power per slot
  - However: modern FPGAs require multiple supply voltages,
    high-speed lanes introduce routing constraints
  - **Board design has become a multi-faceted engineering challenge**
    → Powering, cooling, airflow, noise, vibration, plus the actual electronics part…
  - ATCA has actually already gone out of fashion in the TelCo industry
    (and higher link speeds have rendered the backplane obsolete)

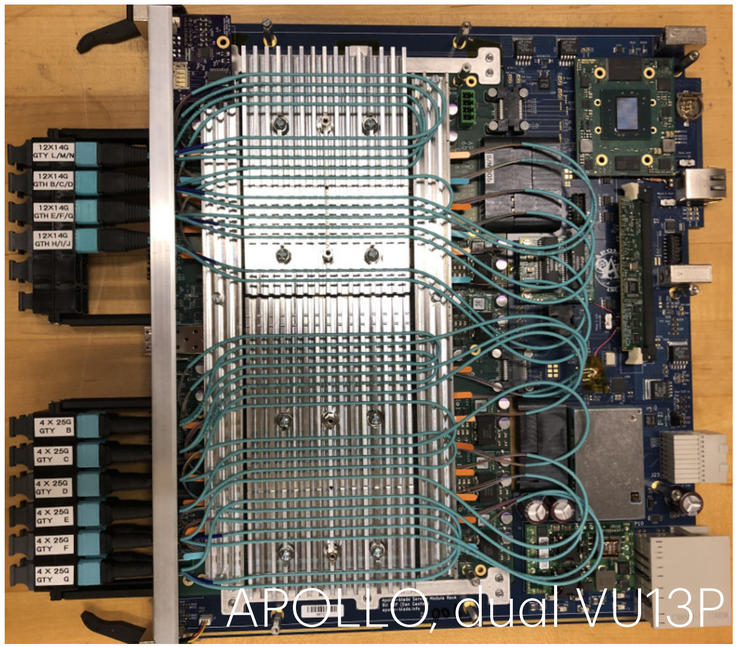
The back-end environment throughout the years

- Originally, LHC off-detector electronics were predominantly VME (6U and 9U), and some CompactPCI
  - Boards were relatively sparsely populated
  - Layout dominated by peripherals
- The Phase-1 upgrade saw a trend towards MicroTCA
  - Significant increase in logic density and component miniaturisation made smaller boards possible
  - Access to peripherals became a challenge
- The Phase-2 upgrade introduced ATCA as crate system
  - Larger boards, and more power per slot
  - However: modern FPGAs require multiple supply voltages,
    high-speed lanes introduce routing constraints
  - **Board design has become a multi-faceted engineering challenge**
    → Powering, cooling, airflow, noise, vibration, plus the actual electronics part…
  - ATCA has actually already gone out of fashion in the TelCo industry
    (and higher link speeds have rendered the backplane obsolete)

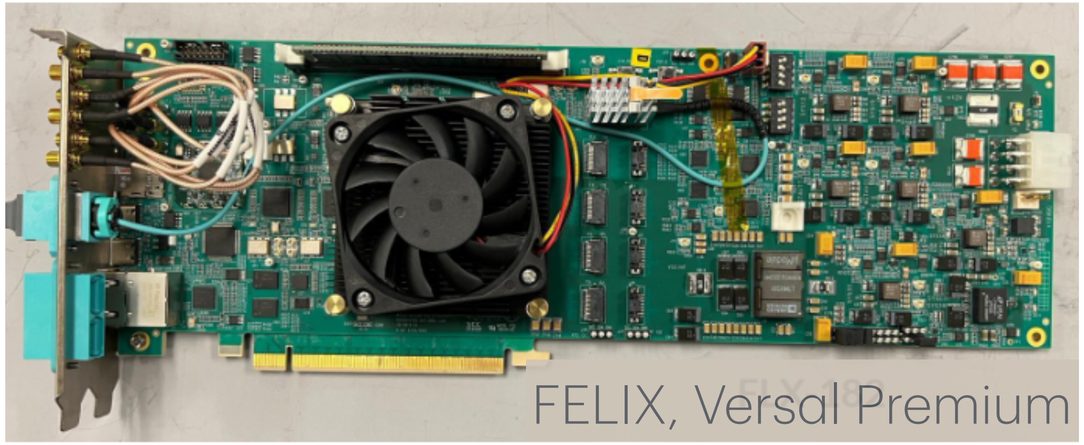
- **VME was likely the last standard that was adopted `CERN-wide'**
- **HEP experiments don't profit from most of the advanced features of these standards**

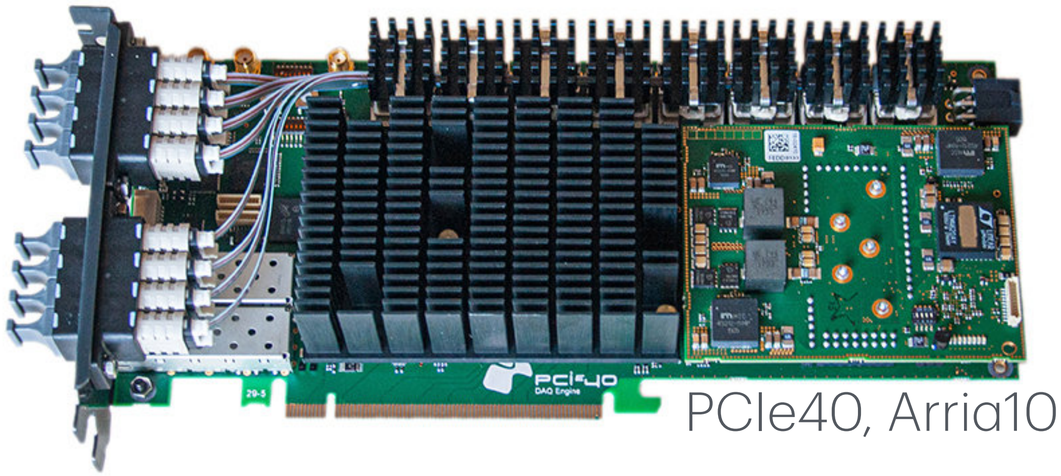
The back-end — Some out-of-the-crate thinking

- Only CMS adopted ATCA for all Phase-2 back-end, read-out, and L1 trigger boards
- ATLAS uses ATCA for L1 trigger
- ALICE, ATLAS, and LHCb all developed back-ends that live in the DAQ read-out/builder unit
  - **Closest integration with COTS hardware so far**
  - May be the best fit for a crate-less future approach



APOLLO, dual VU13P


FELIX, Versal Premium



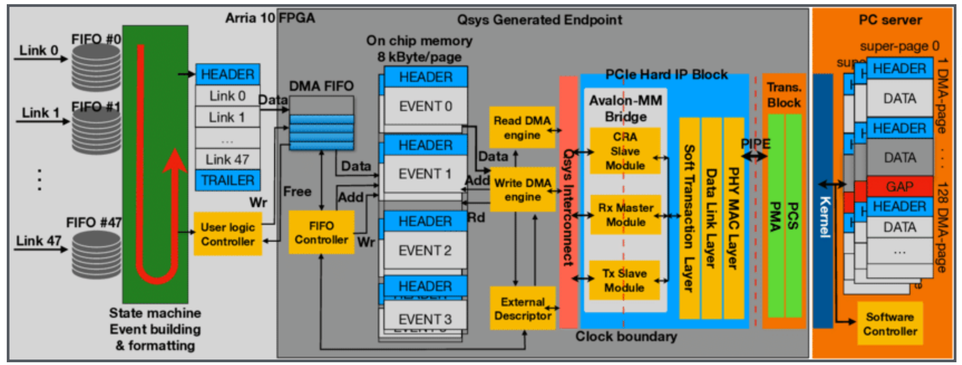
PCIe40, Arria10

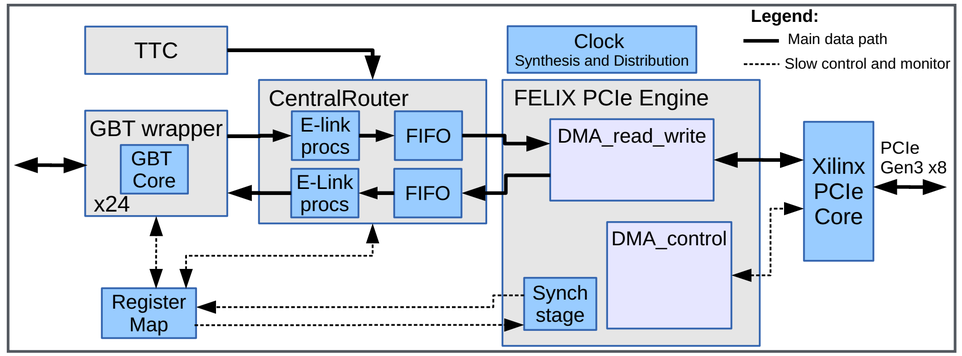
The collaborative back-end

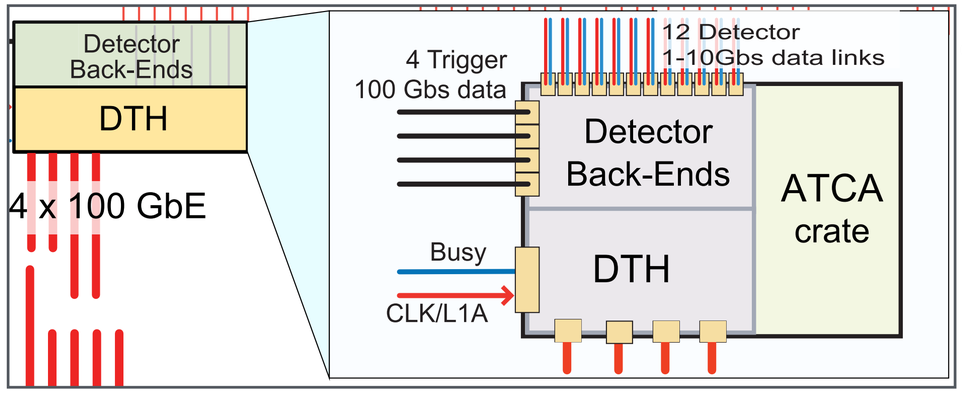
- **State-of-the-art hardware development for trigger-DAQ has outgrown the traditional HEP development model where single institutes deliver an ASIC/board**
  - Modern FPGAs require more and more care in powering, decoupling, signal routing
  - High-speed serial links require more and more specialist design, routing, and even testing
  - High power and signal density requires dedicated PCB design experise
  - Many components (incl. the PCBs) have become `almost too expensive to prototype'
    → Split design into functional parts for prototyping/proof-of-principle
    (Makes for easier sharing too.)

- **Board development (e.g., Apollo, FELIX, PCIe40, Serenity) nowadays often done by consortia of universities and institutes**
  - Expert groups can focus on aspects of functional design, interfaces, mechanics, cooling, etc.
  - Requires a change in focus when presenting our contributions to our funding agencies

Back-end firmware design choices





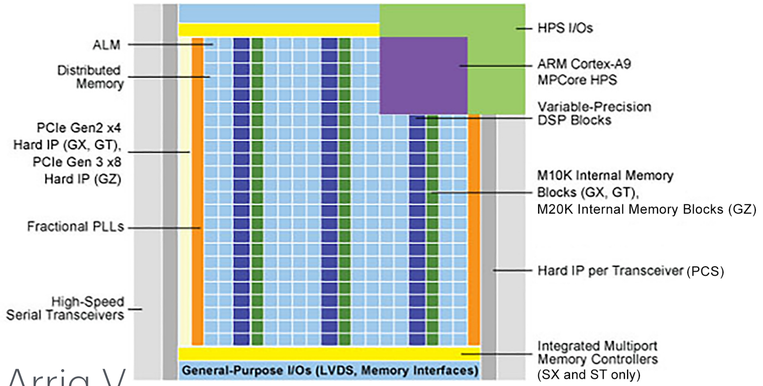


- **ALICE and LHCb**
  - Standardised on a single hardware device
  - Customised this with detector-specific firmware embedded in a common framework
- **ATLAS**
  - Chose a single hardware device
  - Customised this with flexible configuration
- **CMS**
  - Chose a hardware form factor, and allowed detector-specific back-end hardware implementations
  - Delivers common firmware for all central interfaces

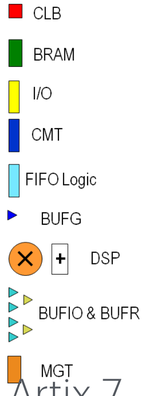
Back-ends become (pre)processors?

- FPGAs used to scale predictably:
  - More of the same blocks
  - Faster I/O

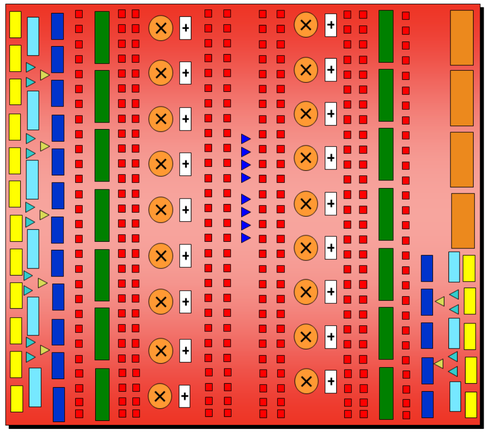
  → Facilitated design evolution



Arria V



Artix 7

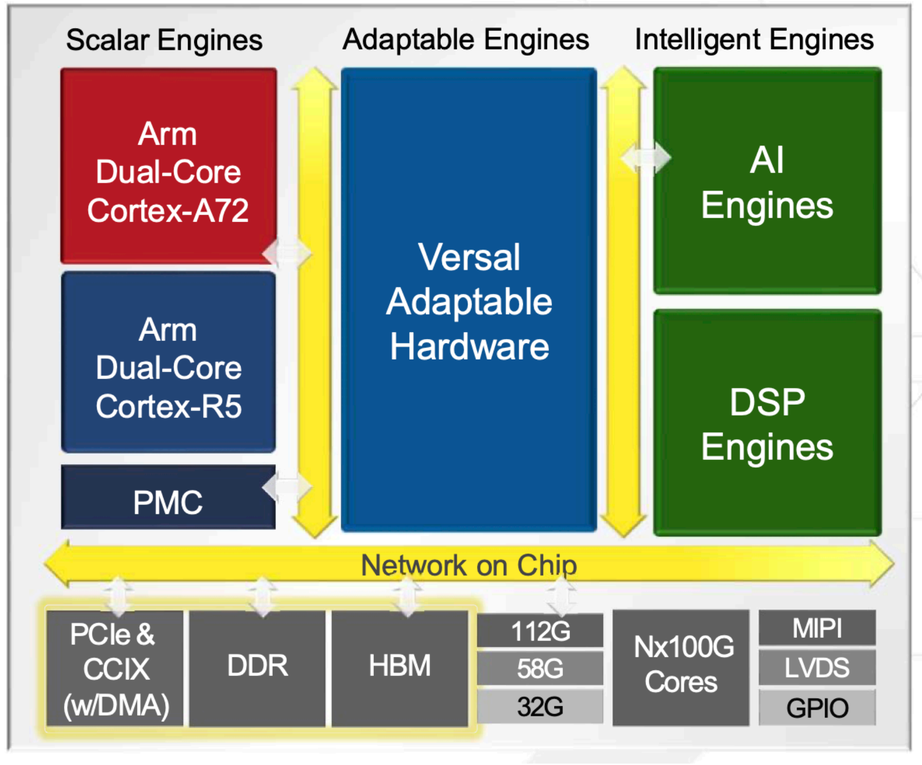


Back-ends become (pre)processors?

- FPGAs used to scale predictably:
  - More of the same blocks
  - Faster I/O

  → Facilitated design evolution

- FPGA have now become `Adaptive SoCs'
  - Have to learn how to best use these for HEP purposes
  - Data reduction? Data reformatting?
    → Will have effects on both trigger and DAQ

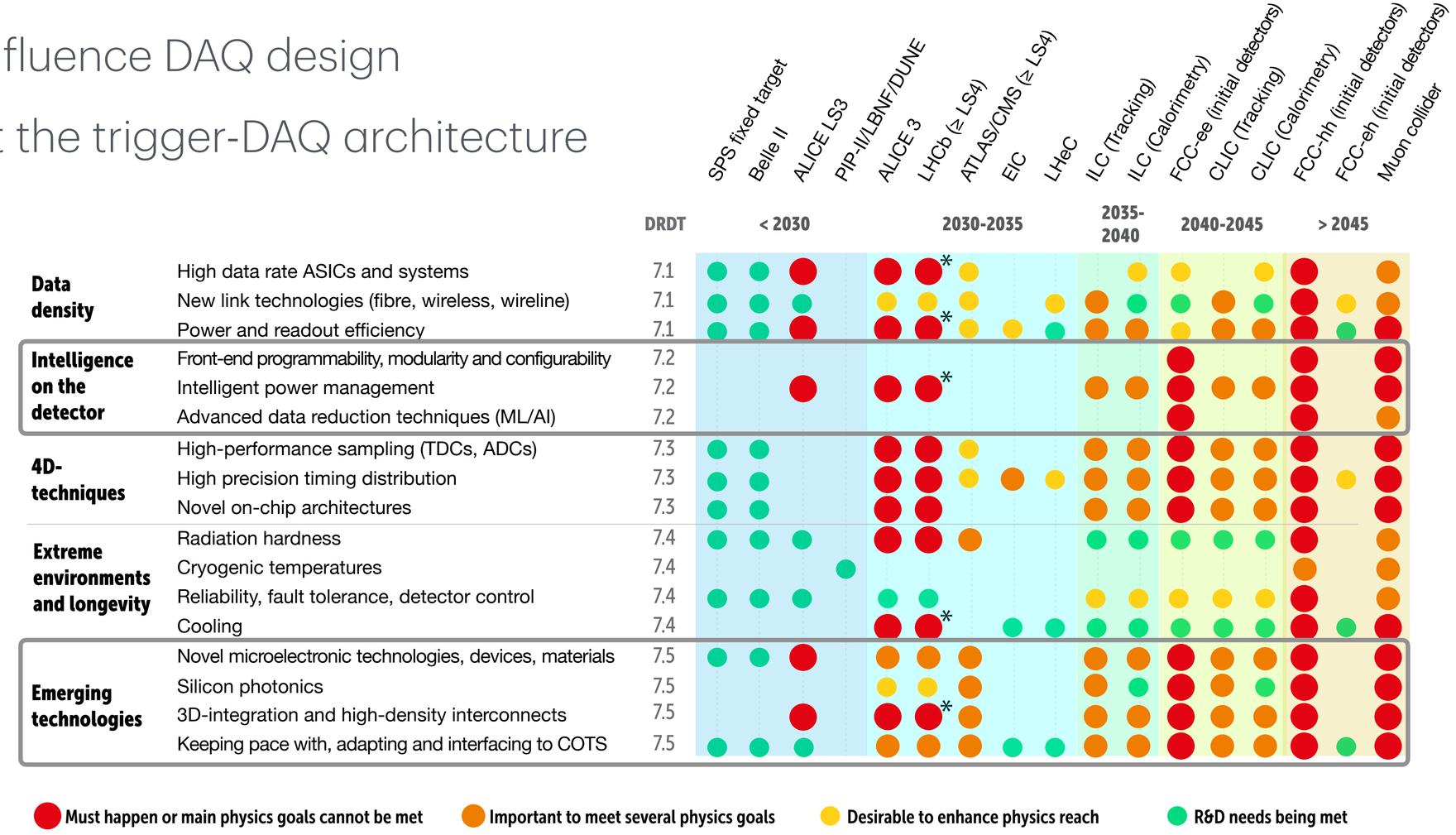


Focus on the future

The global trigger-DAQ strategy may be about to change

ECFA DRD7: electronics and on-detector processing

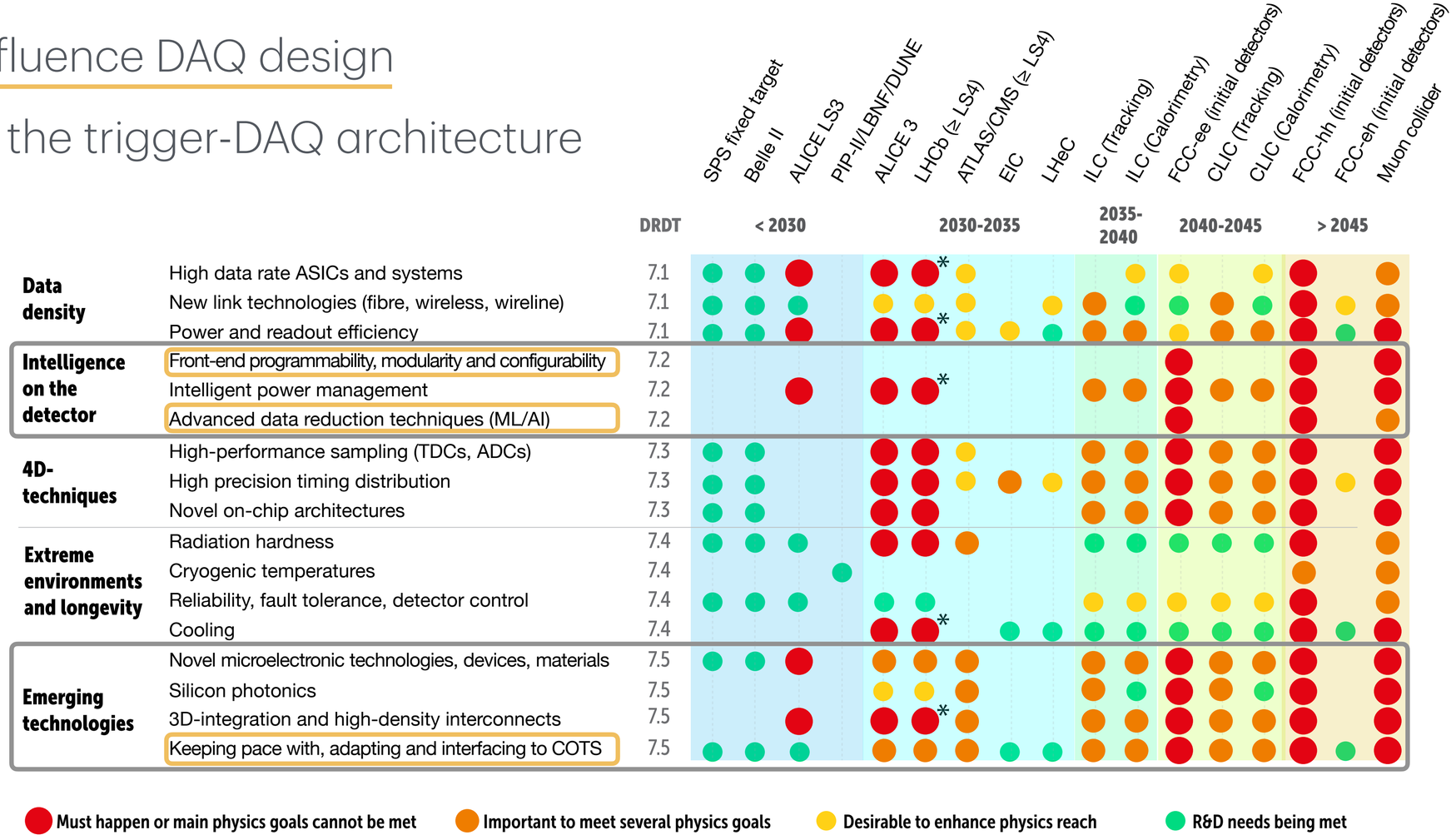
- Some topics directly influence DAQ design
- Others indirectly affect the trigger-DAQ architecture



The global trigger-DAQ strategy may be about to change

ECFA DRD7: electronics and on-detector processing

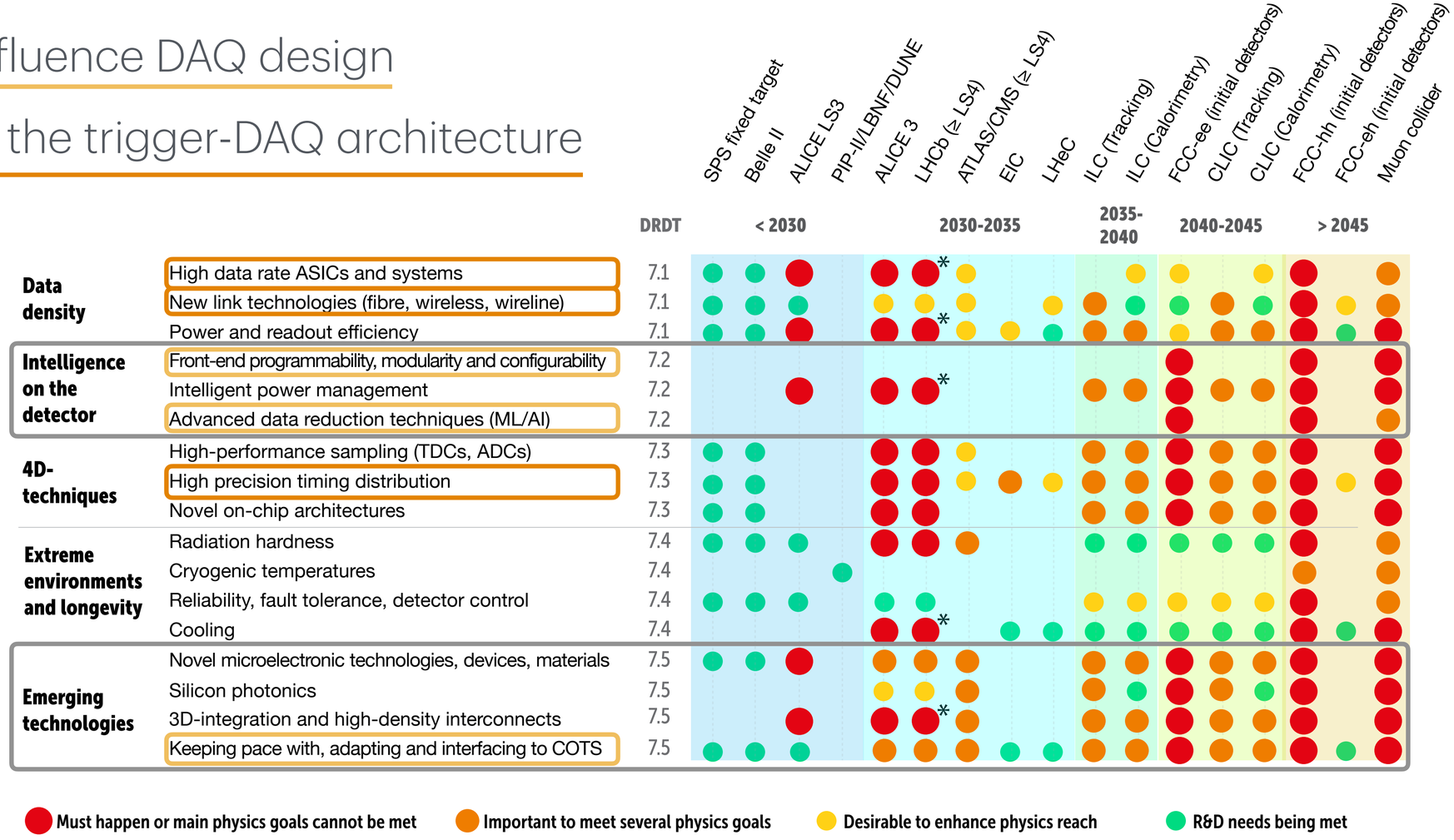
- Some topics directly influence DAQ design
- Others indirectly affect the trigger-DAQ architecture



The global trigger-DAQ strategy may be about to change

ECFA DRD7: electronics and on-detector processing

- Some topics directly influence DAQ design
- Others indirectly affect the trigger-DAQ architecture



ECFA DRD7: electronics and on-detector processing

- Programmable front-ends could trigger a paradigm change:
  process data at the front-end
  → reduce amount of data to be moved, and optimise for read-out

- Adopt and adapt (existing and emerging) industry standards and techniques

- Study effective use of new technologies in back-ends
  (SoCs, silicon photonics, GPUs, FPGAs as accelerators)
  → reduce amount of data to be moved

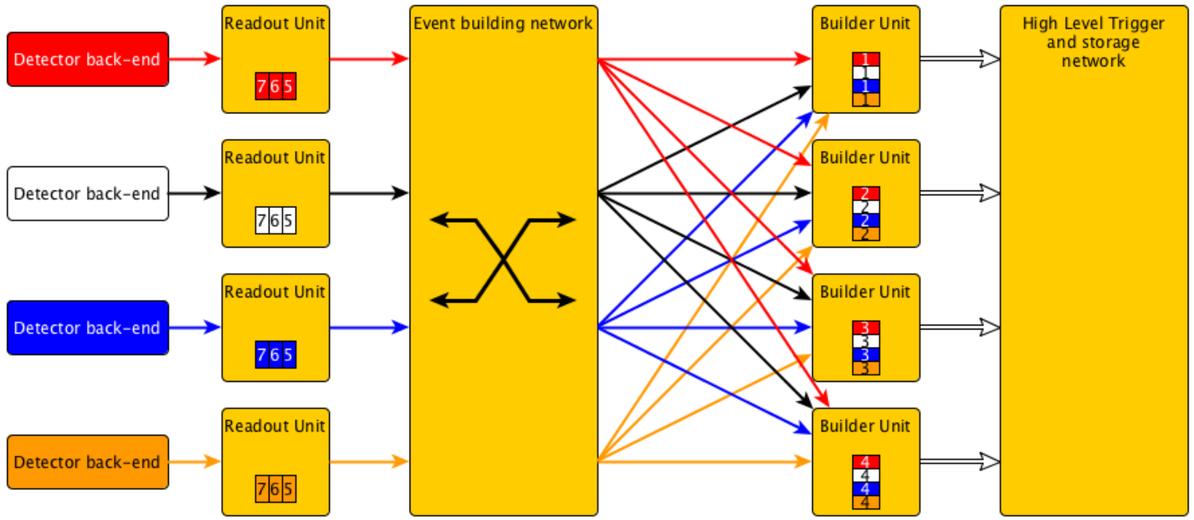
- Maybe no longer need for a (custom) back-end at all?
  Connect front-ends directly to COTS hardware, using standard protocols
  - **The next logical step after the PCIe40 and the FELIX?**
  - May require separating slow control from trigger-DAQ again
  - DUNE are going in this direction, important proving ground

		DRDT
Data density	High data rate ASICs and systems	7.1
	New link technologies (fibre, wireless, wireline)	7.1
	Power and readout efficiency	7.1
Intelligence on the detector	Front-end programmability, modularity and configurability	7.2
	Intelligent power management	7.2
	Advanced data reduction techniques (ML/AI)	7.2
4D-techniques	High-performance sampling (TDCs, ADCs)	7.3
	High precision timing distribution	7.3
	Novel on-chip architectures	7.3
Extreme environments and longevity	Radiation hardness	7.4
	Cryogenic temperatures	7.4
	Reliability, fault tolerance, detector control	7.4
	Cooling	7.4
Emerging technologies	Novel microelectronic technologies, devices, materials	7.5
	Silicon photonics	7.5
	3D-integration and high-density interconnects	7.5
	Keeping pace with, adapting and interfacing to COTS	7.5

Recent architecture example that optimise COTS hardware use

HEP data in the DAQ network — there and back again?

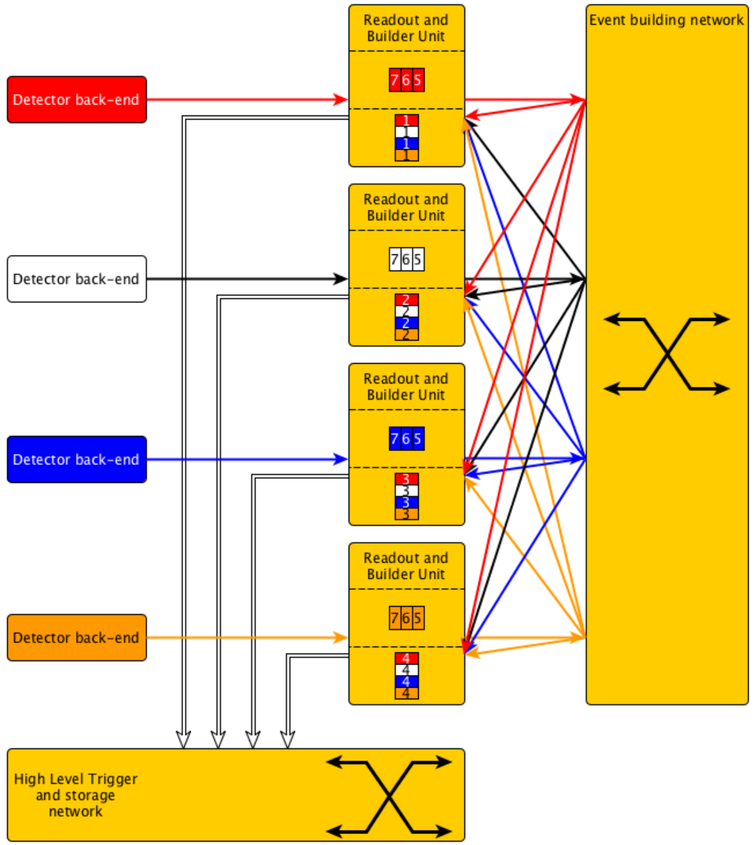
- Functionally, data move from front-end, to back-end, to read-out, to event building, to event filtering
- If read-out, protocol translation, etc. is heavy, the presence of the readout-only node leads to unidirectional network use
- If this load can be lightened (or if the processors improve faster than the network bandwidth) a folded approach uses the NICs more efficiently
  → `half the network ports'
- Note that in both approaches the network settings can be fine-tuned, probably with varying results



Recent architecture example that optimise COTS hardware use

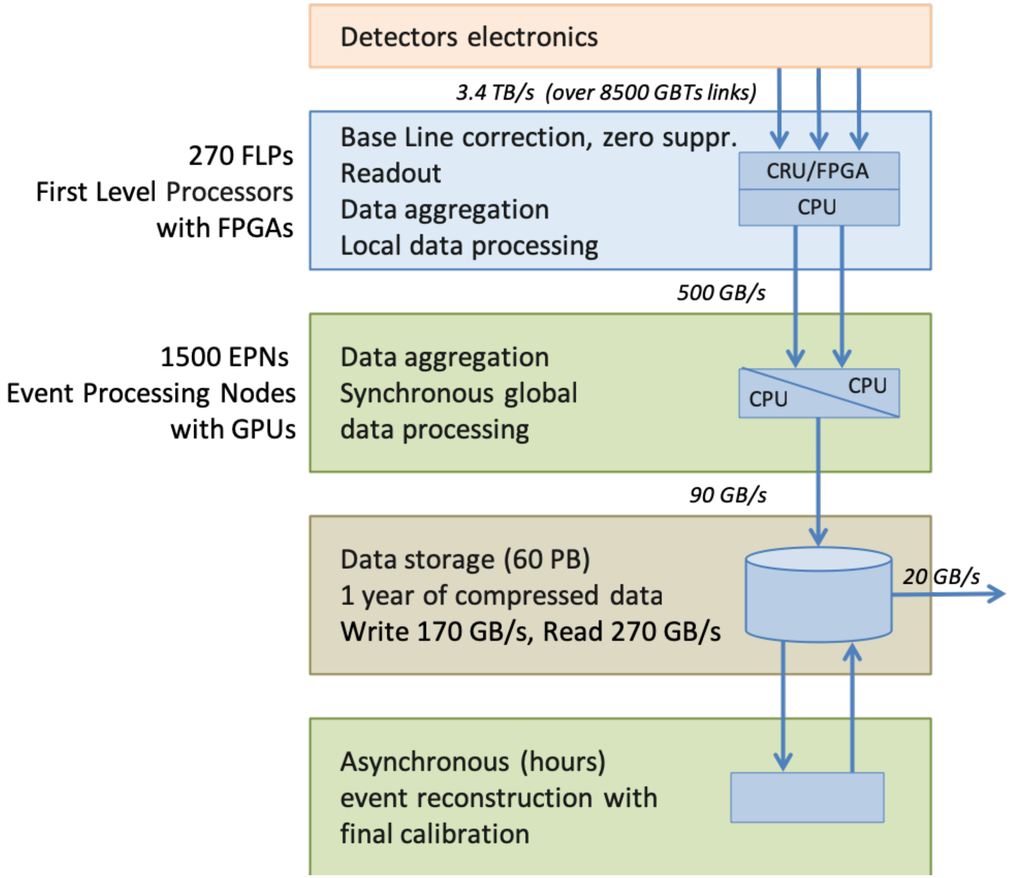
HEP data in the DAQ network — there and back again?

- Functionally, data move from front-end, to back-end, to read-out, to event building, to event filtering
- If read-out, protocol translation, etc. is heavy, the presence of the readout-only node leads to unidirectional network use
- If this load can be lightened (or if the processors improve faster than the network bandwidth) a folded approach uses the NICs more efficiently
  → `half the network ports'
- Note that in both approaches the network settings can be fine-tuned, probably with varying results



Recent example that maximises the use of COTS equipment

ALICE O^2 turns the (annual) LHC duty-cycle into an advantage



- The upgraded ALICE read-out uses only a single custom processing and read-out board: the Common Readout Unit (CRU/ PCIe40)
- PCIe board, housed in COTS host, equipped with versatile firmware
- COTS equipment handles all further stages of data acquisition
- Finely-tuned, custom software orchestrates data-taking workflow
- LHCb uses a similar approach, on an LHC fill-by-fill basis
  → Use the inter-fills as asynchronous processing time
  → May require redefining the meaning of 'raw data'

Food for thought

- **No read-out revolution is necessary. Evolution is doing just fine.**
  - COTS components are moving closer to on-detector (e.g., some muon systems already use FPGAs and SFPs)
  - Decreasing need for custom hardware
  - Increasing need for smart software and system design to optimise the use of COTS solutions

- **Some of the ECFA DRD7 projects have the potential to more radically change future detector read-out, as well as trigger-DAQ architectures**
  - No more back-end? Configurable front-ends? Intelligent front-ends?
  - Or, can we make the smart NIC the next GPU?

- **Continued evolution of our trigger-DAQ systems for the benefit of physics would really profit from a more `holistic' approach to detector design**
  (e.g., ensure that increasing granularity 'maps well' to trigger requirements and read-out aggregation)
  This could be a great by-product of the recent ECFA DRD efforts

Thank you for your interest

Questions and comments are welcome
Now, or over coffee later

Detector read-out *has* changed subtly

- Front-ends remain built around custom ASICs and links
  - Common component development optimises design and testing efforts, and maximises chances of success
  - Custom-to-COTS interoperability is a new challenge for every new technology generation
- Back-ends remain custom hardware, based on COTS components
  - Recent development very closely integrates the custom hardware with COTS hosts
- Event building is (and has been for a while) based on:
  - custom software, running on on COTS hardware, using standard interconnect technologies
  - Main challenge: keeping up with COTS developments, and effectively adopting/adapting those