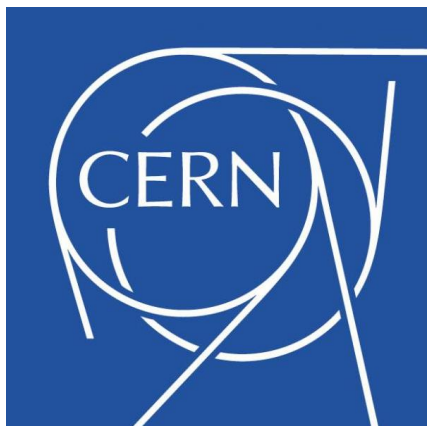
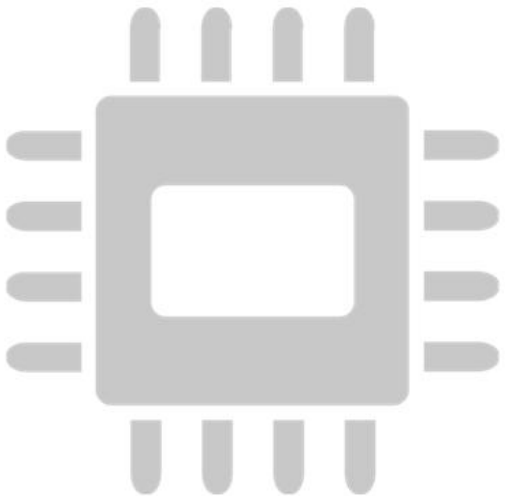
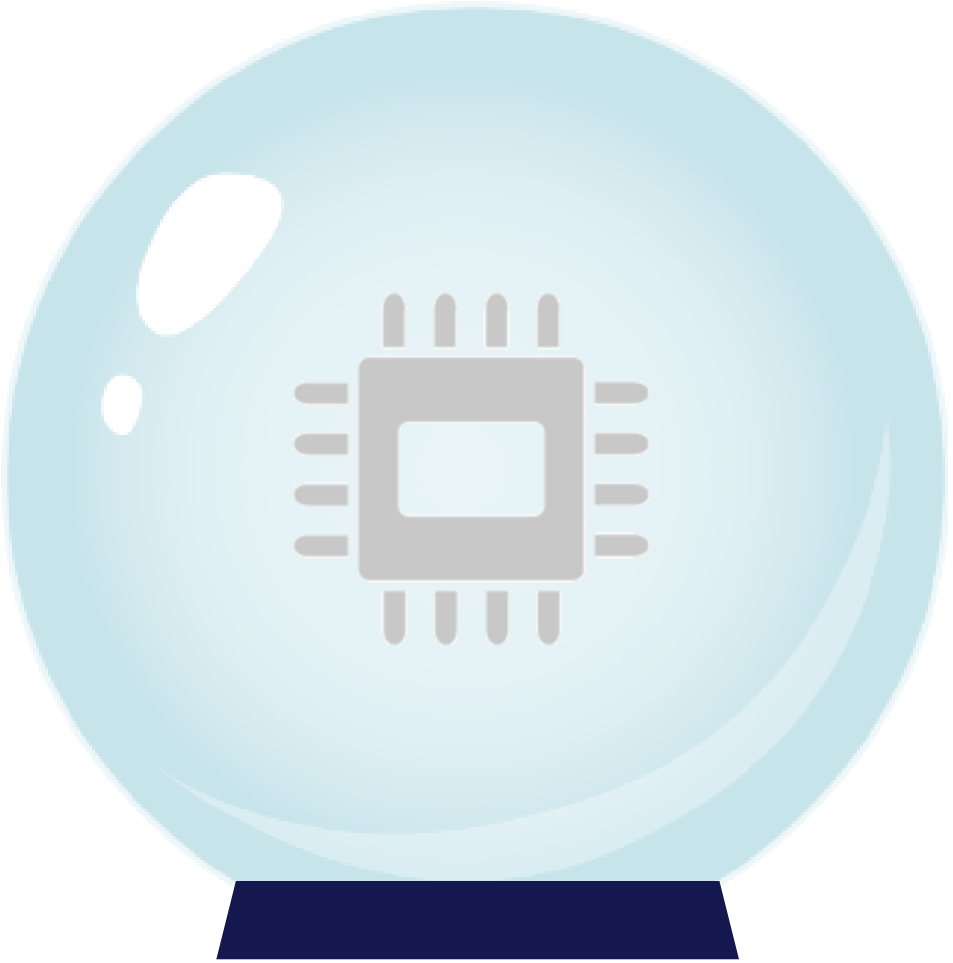


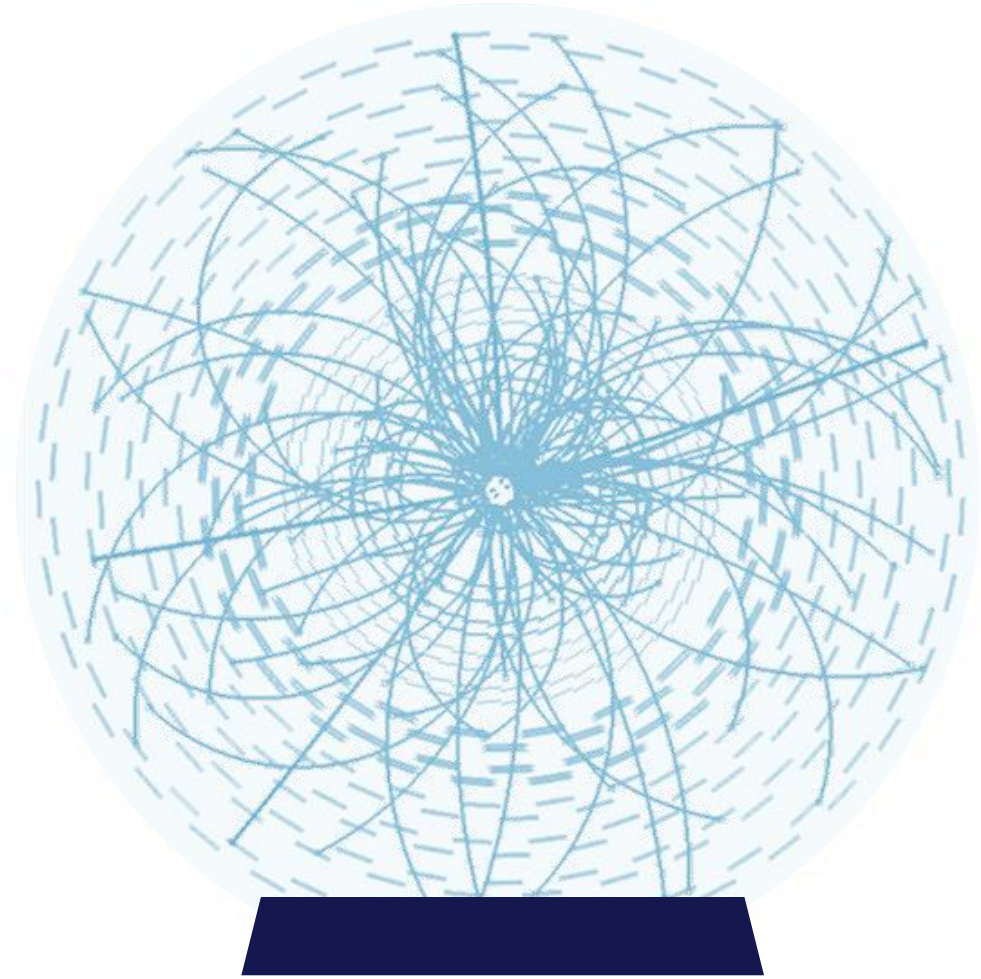
FDF



THE FUTURE OF FPGAS



THE FUTURE OF FPGAS

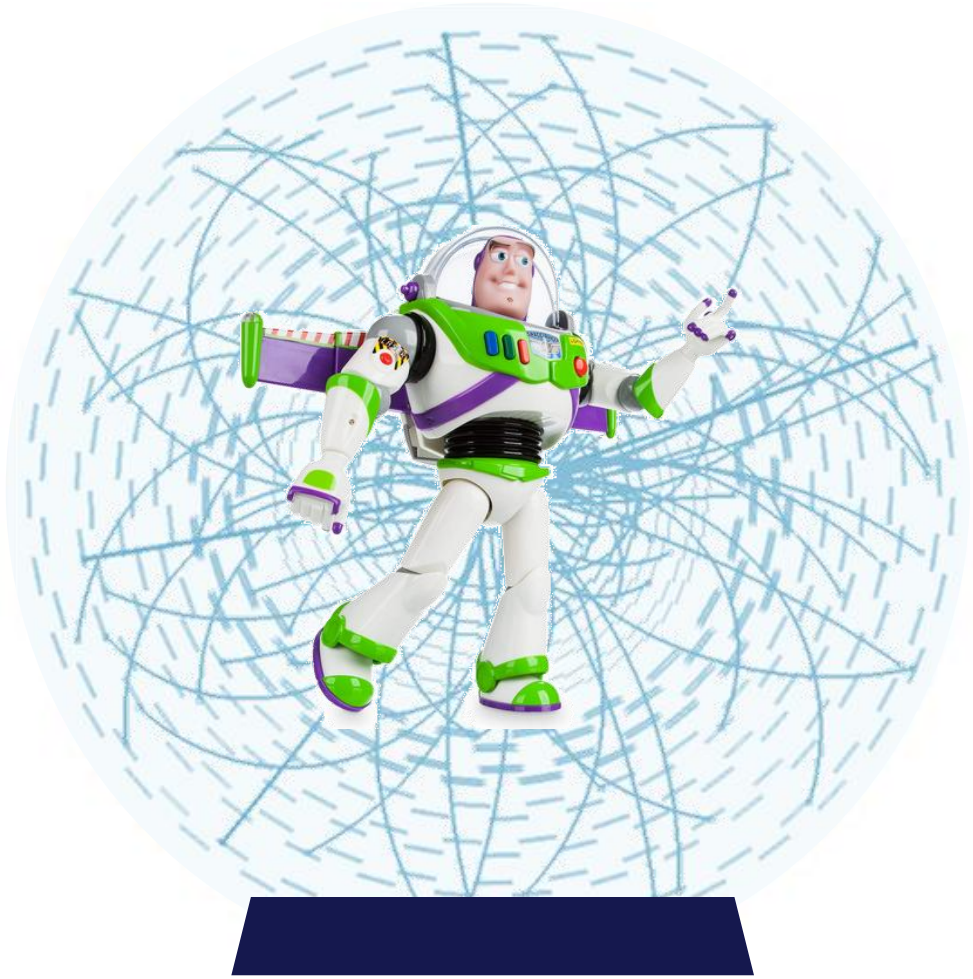


IN HEP* DETECTORS, FOR FCC*

*HEP: High Energy Physics

*FCC: Future Circular Collider

THE FUTURE OF FPGAS



IN HEP* DETECTORS, FOR FCC*

AND BEYOND

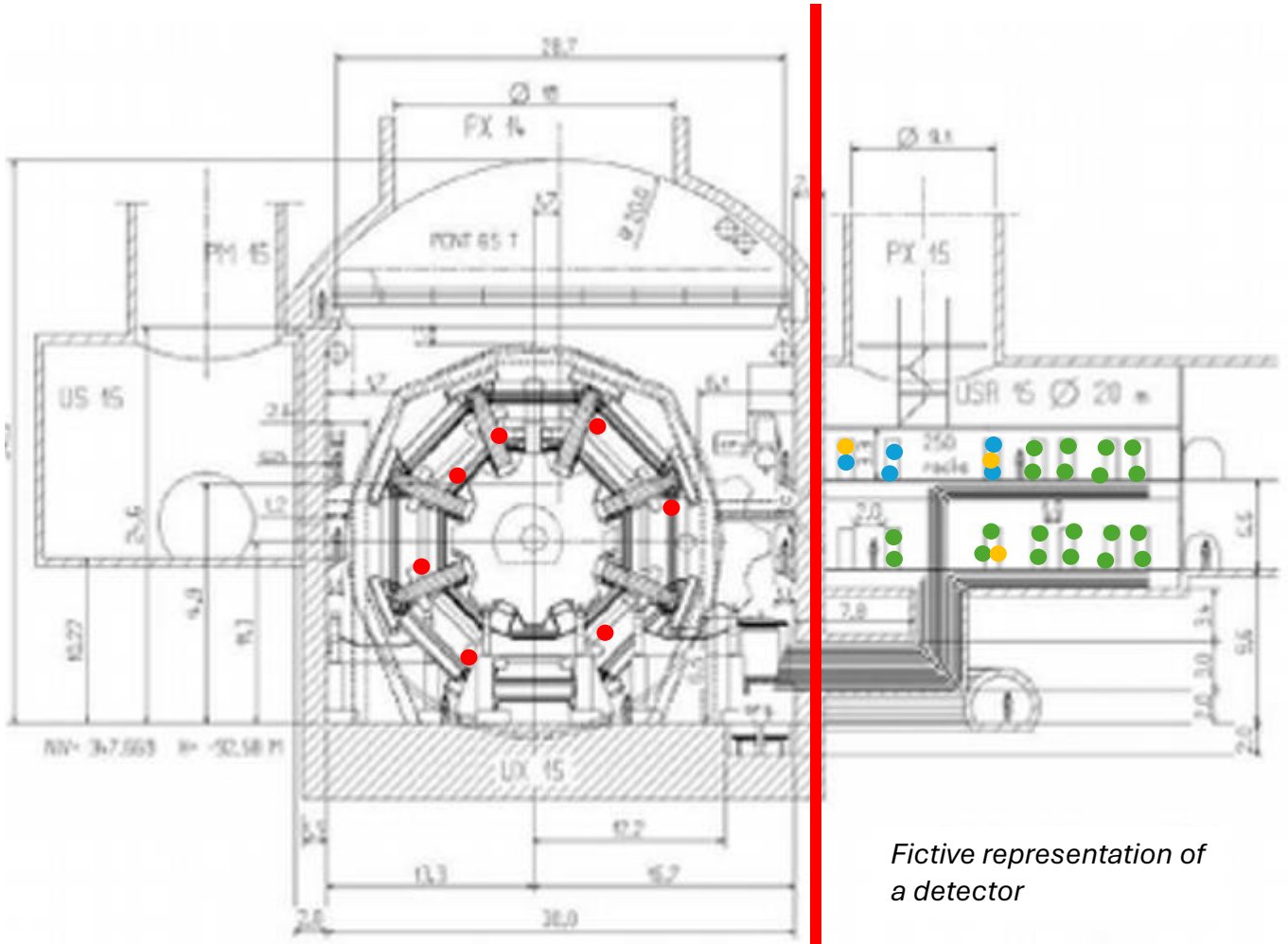
*HEP: High Energy Physics

*FCC: Future Circular Collider

FPGAs in HEP Detectors

How do they look like today?

FPGAs in HEP Detectors



Fictive representation of a detector

FPGAs for...

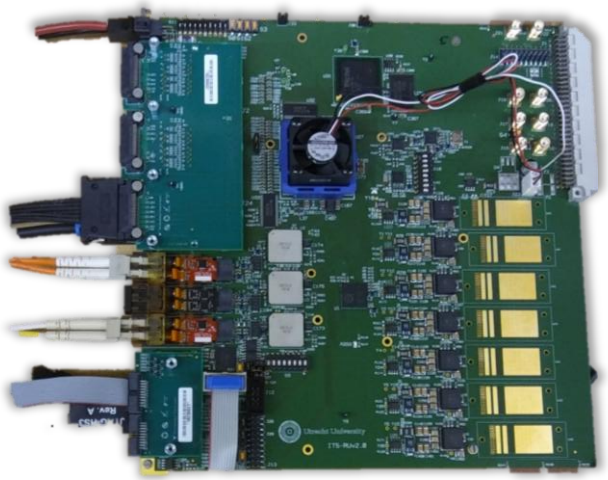
- Back-End
  - Trigger
  - DAQ
  - Control & Timing
- Front-End
  - Data Readout

Front end **Back end**

*Majority of Back End FPGAs are AMD-based, but there are also some Altera based boards

FRONT-END FPGAs

- First line of digital processing—right after the analog/digital conversion of signals from sensors.
  - initial data conditioning & formatting like zero suppression, pedestal subtraction, hit finding, clustering..
- Harsh environments (High Radiation dose & fluence, high magnetic field, limited access)
  - Not the ideal place to install commercial components
    - Avoid them as much as we can when Rad-Hard ASICs available
    - FPGAs are placed meters away from the interaction point to lower radiations
  - Preferably Flash or Antifuse based FPGAs (Microchip ProASIC, Polarfire, RTG4 or RTSX).
    - Rad Tol or Rad Hard
    - Expensive and logic limited.
  - Also SRAM-Based like AMD 7 or Ultrascale series,
    - Radiation hardening techniques required like Tripple Modular Redundancy (TMR) or configuration scrubbing



Example: ALICE ITS2 Readout Unit

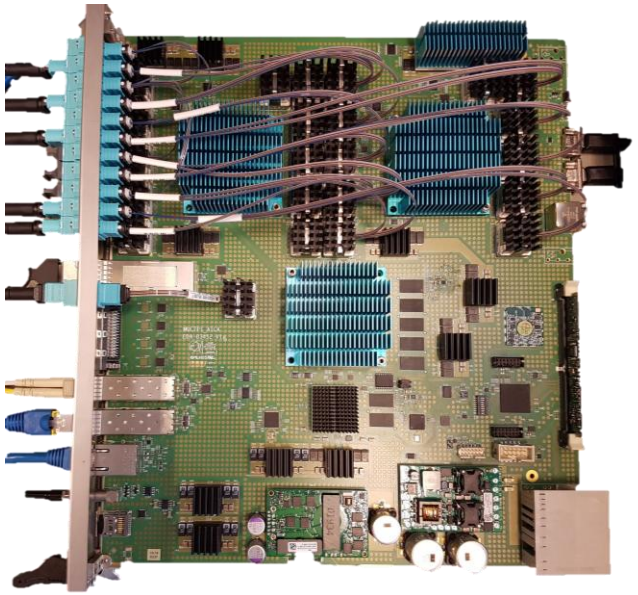
- 5m from beam
- 1 Ultrascale FPGA
- 1 Pro-Asic FPGA for scrubbing

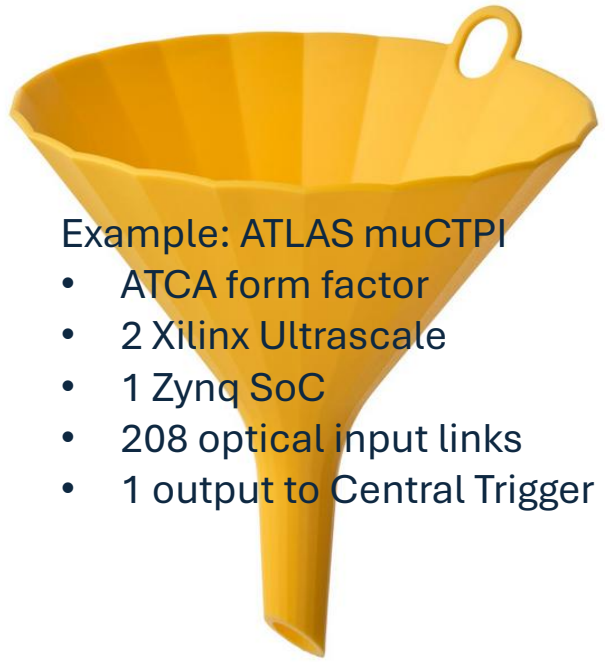
- Magnetic Field: 0.5T
- TID: 10krad
- 1MeV neq fluence: 1.6.1011pcm-2
- SEU cross section 103 cm-2s-1

TRIGGER FPGAs

- Identify interesting events out of billions occurring per second
  - select which events to keep (event rate 40MHz)
  - do that very quickly and with limited latency (us)
  - propagate down the decision to front-end to drastically reduce the data rate
- Complex Tasks
  - Synchronization (timestamping)
  - Trigger primitive identification (clusters, energy sum..)
  - Pattern recognition (tracks, showers..)
  - Kinematic reconstruction (momentum, energy, angles..)
  - Pipelining (parallelizing tasks)
  - Trigger decision (threshold, ..)
  - Latency control (all this with fixed latency)
- Very demanding and challenging requirements
  - Many (100+) 10Gbps transceivers
  - High-density logic (LUTs), many DSP slices, and large RAM
  - Fast processing power in logic

  => Ideally: Ultrascale+ series from AMD
- Large Boards Form factor
  - ATCA
  - Few boards required (<10)



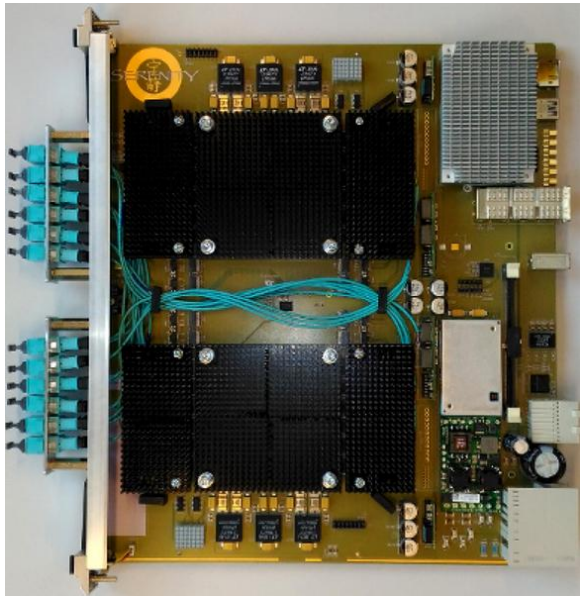


Example: ATLAS muCTPI
- ATCA form factor
- 2 Xilinx Ultrascale
- 1 Zynq SoC
- 208 optical input links
- 1 output to Central Trigger

DAQ FPGAs

- Readout of all detector data from front-end electronics to event-building farms
  - high-throughput, reliable, and lossless data transmission
- Typical tasks: reorganizing data
  - Decoding & correcting Front-End packets
  - Unpacking, deserializing
  - Clustering (grouping adjacent detector hits that likely originated from the same physical particle or interaction)
  - Compression, Buffering & Congestion management
  - Formatting to standard protocols (PCIe, GbE..)
- Requirements
  - High aggregation ability: LUT+, little processing
  - Mix-type of transceivers (48+ 10Gbps & a few 100+Gbps transceivers)
  - Hard IPs for standard protocols (PCIe, 100GbE, ..)
  - Typical FPGA throughput ~100-400Gbps
  - Large buffering ability, fast memory access (RAMs, HBMs..)

  =>Ideally: Versal Premium series
- Boards Form factor
  - More flexible: ATCA or PCIe
  - 100s of boards required



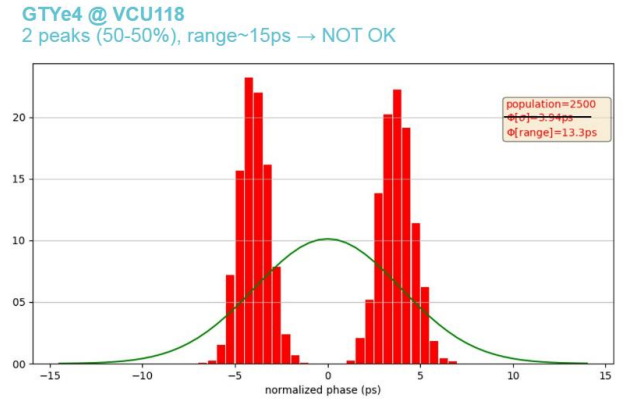
CMS SERENITY

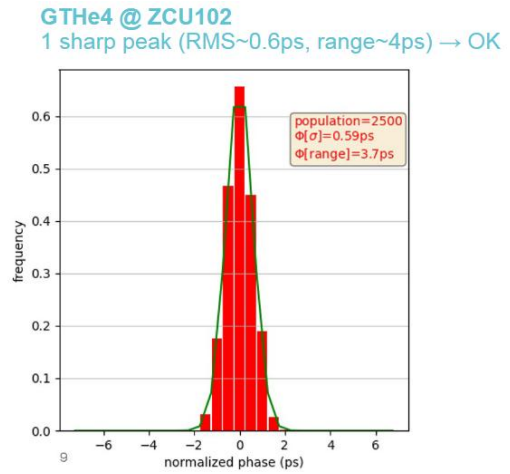
LHCb & ALICE PCIe40

ATLAS FELIX FLX155

Control & Timing FPGAs

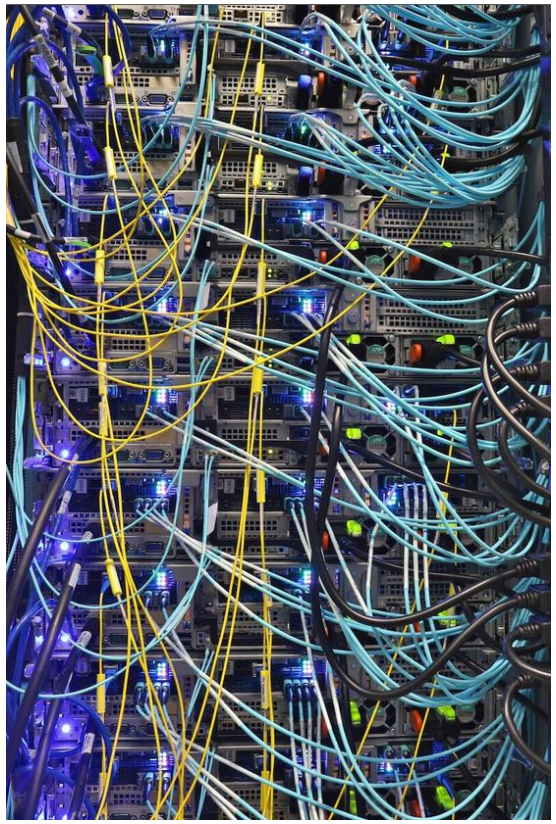
- Control
  - Now mostly provided by SoC
    - In dedicated FPGAs whenever there is space or several FPGAs to control (Zynq)
    - Embedded within DAQ FPGAs for PCIe boards
- Timing
  - In charge of Bunch Clock (ie collision clock) distribution to FE
    - Often merged with Trigger or DAQ Hardware
  - Requires dedicated optical links
    - Clock Embedded in serial stream (trigger and/or slow control)
    - Bunch Clock is recovered thanks to CDR circuits embedded in transceivers
    - Bunch Clock is cleaned and re-embedded in serial streams to Front-End
  - New Timing detectors very demanding
    - O(ps) precision/stability and ~1ps rms jitter in all conditions
    - Not all transceiver/FPGA types are able to garanty this, especially between resets
      - AMD FPGAs seem to provide more features & garanties so far (U & U+)
      - Each transceiver family performs differently
    - Sensitive to temperature, firmware layout, …





Back-End FPGAs in their environment

DAQ rack with PCIe boards in servers



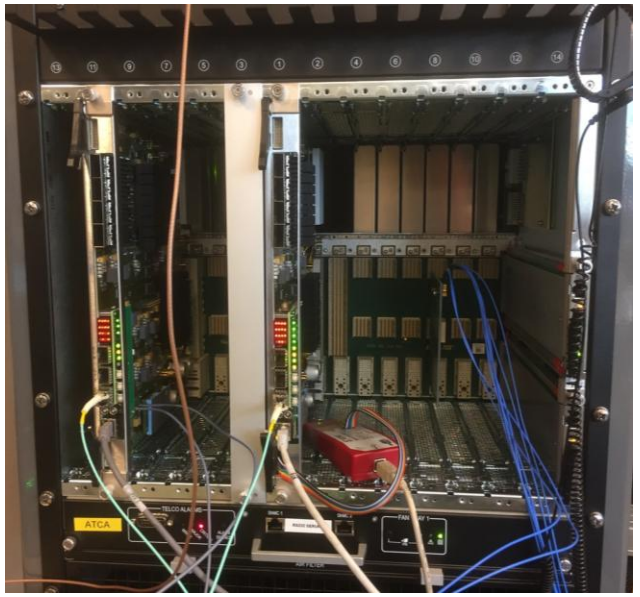
ARRIA10

2-3 boards /server

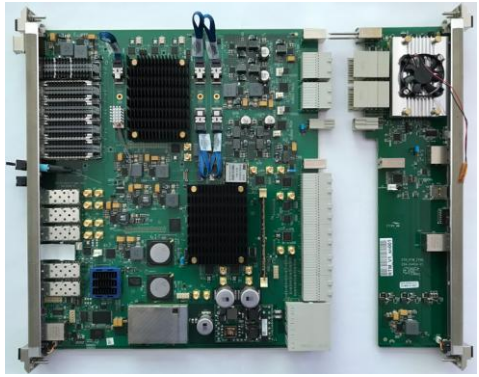
Credit S. Chapeland

ALICE TPC rack with 12 server & 30 CRU (PCIe40)

Partial view of a DAQ rack with ATCA boards



One ATCA crate with DTH boards in CMS lab



Max 14 blades / crate often less

Future of FPGAs in HEP

(A lot of) questions and (some) elements of answers

FUTURE OF FPGAs IN HEP

FPGA Market trends

Physics strategy

Detector trends

Future
Radiation Levels

Timing Precision

R&D trends
for HEP

Triggering strategy

HEP projects organization

Bandwidth
requirements

Questions

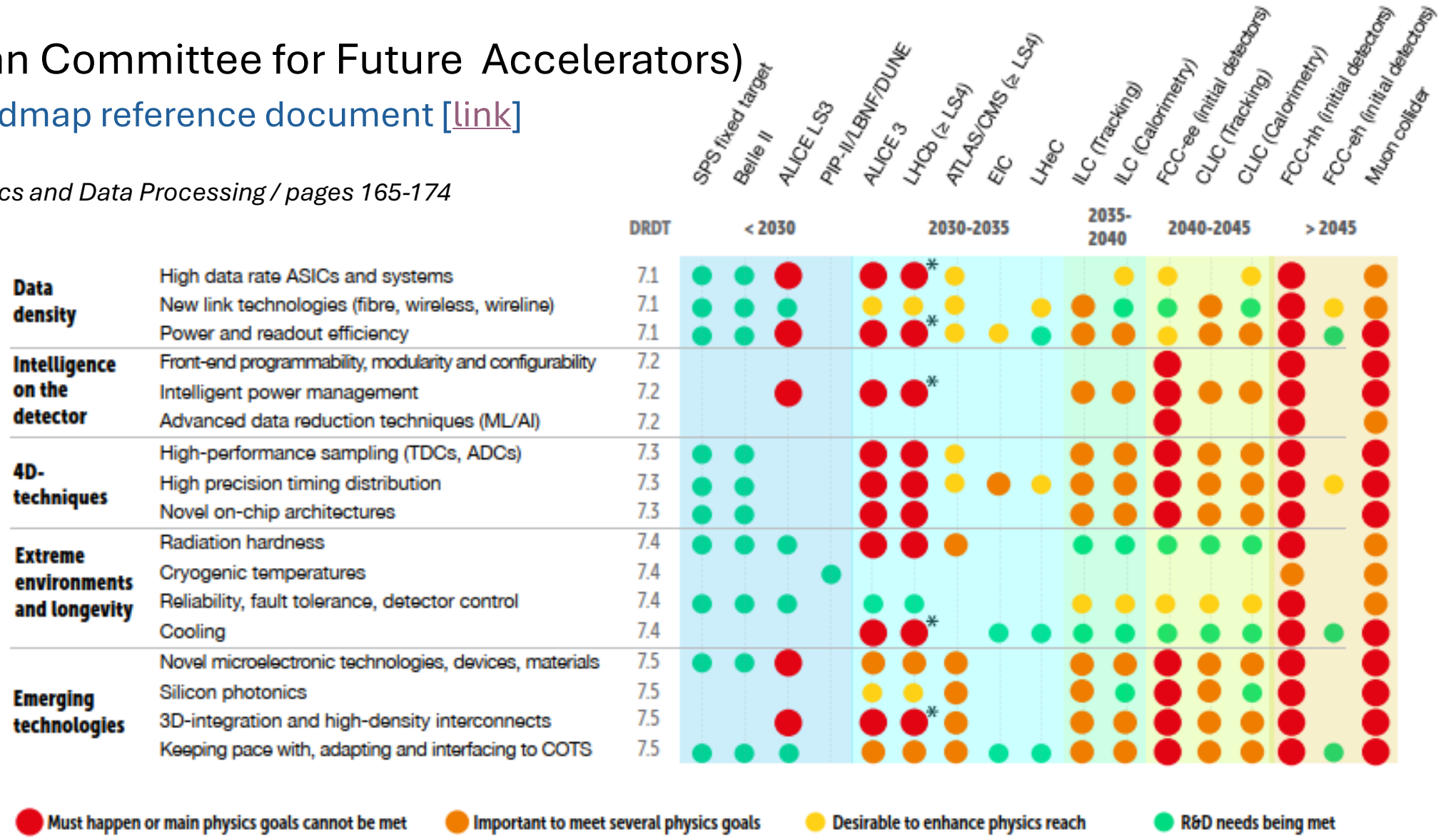
1. Will we still have trigger FPGAs in future detectors?
2. Will we replace L1Trigger FPGAs with GPUs?
3. Could we get rid of FPGAs in future DAQs?
4. Will we be able to use FPGAs in the Front-End?
5. Does the transceivers trend match the needs of future detectors?
6. Will future FPGAs be compatible with high precision timing?
7. Is the Chiplet-trend interesting for FPGAs in future detectors?
8. On which types of boards will our FPGAs sit?

Sophie.baron@cern.ch

Detectors R&D Trends for Future Accelerators

- ECFA (European Committee for Future Accelerators)
  - Detector Roadmap reference document [link]

Chapter 7 / Electronics and Data Processing / pages 165-174



Focussing on HL-LHC & FCC

2030 → 2047 ? → 2070?

	HiLumi HL-LHC PROJECT Discovery machine	FUTURE CIRCULAR COLLIDER ee Precision machine	FUTURE CIRCULAR COLLIDER hh Energy Frontier Discovery machine
Energy	14TeV	<0.3TeV	100TeV
luminosity	$7x10^{34}$	10^{34}-10^{36}	4-$30x10^{34}$
Bunch Crossing Freq.	40MHz	<50MHz	40MHz
Collision type	proton-proton Dirty, complex	Electron-positron Clean, simple	Proton-proton Dirty, complex
Pileup	~200	~ <5	~1000
Highest Total Ionizing Dose	~1–5 MGy ~100-500 MRad	<10 kGy <1 MRad	~10–20 Mgy ~1-2 GRad
Highest Fluence (neq/cm^2)	~2×10^{16}	~10^{10} -10^{12}	~2×10^{17}

Q1

Will we still have trigger FPGAs in future detectors?

- Balance between benefit (raw data reduction) & cost (technical constraints)
- HL-LHC : Definitely!
  - ATLAS, CMS : high collision rate (40MHz), high Pileup: 200 => Trigger!
  - LHCb will remain triggerless (software trigger only): Collision rate: ~30MHz, Pileup: ~5
- FCC-ee : not clear yet
  - Debate started in the TDAQ Community: a triggerless architecture is tempting but depends on parameters not yet known (Beam background, etc)
  - Pros-and cons for trigger/less architecture
    - 50MHz Bunch Crossing rate with very low Pileup
    - «physics events» can go down to 200kHz (when differentiated from background event)
    - Very dense detectors type: less space for fibers & services
  - Triggerless data rate not well known yet
    - Estimations vary from 1.3 TB/s to 14 TB/s (different background simulations)
    - Remains on the order of magnitude of current systems

Excellent talk at TDAQ@LHC 2025 [link]

Q1

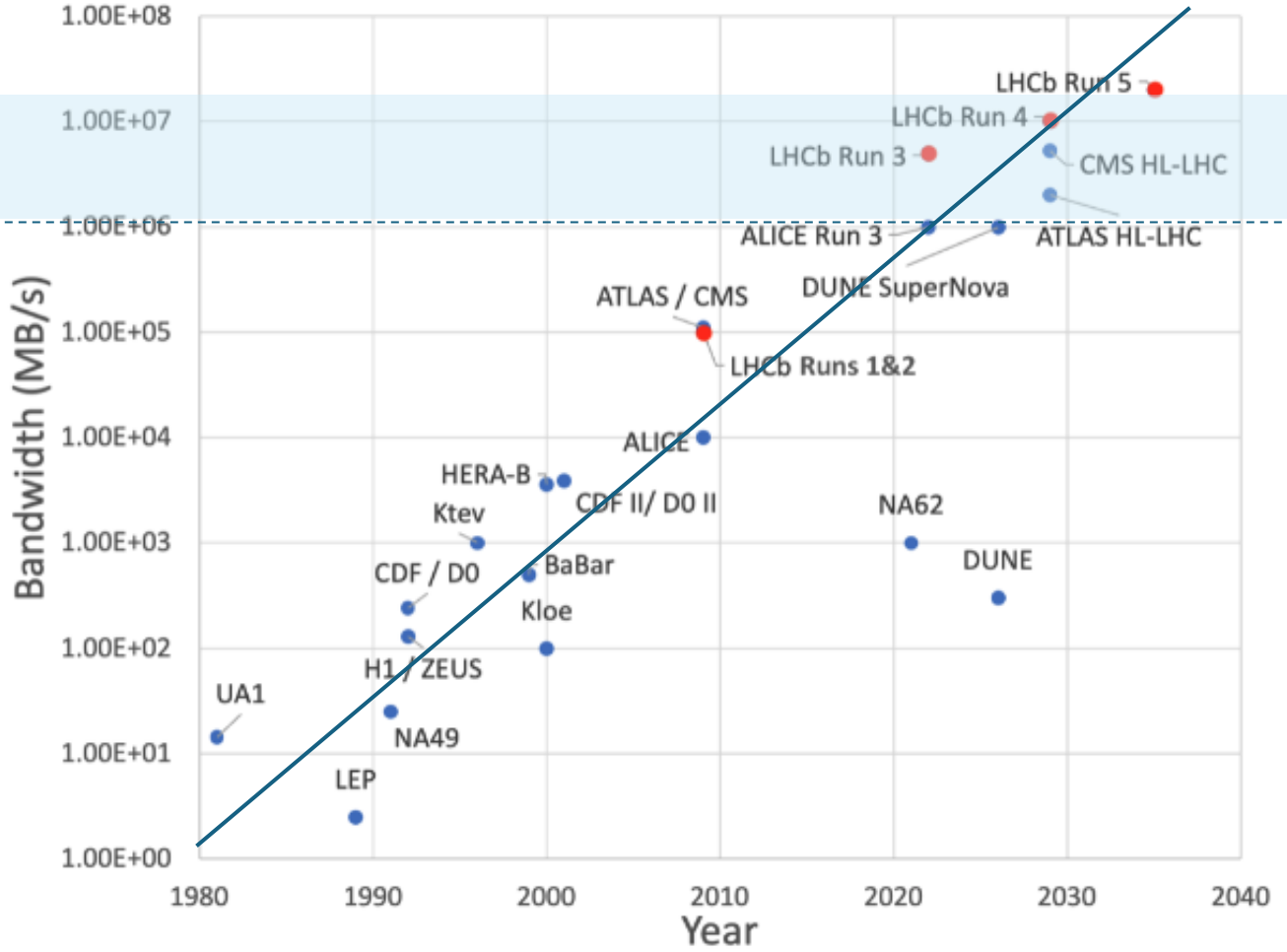
Experiment bandwidths today and tomorrow

1PB/s

Triggerless bandwidth for FCC-hh experiments

FCC-ee Experiments Triggerless Bandwidth estimation

1TB/s



[link]

Figure 1: Evolution of experiment bandwidths in High Energy Physics as a function of time [graph courtesy of A. Cerri, U. of Sussex]. The various LHCb experiment phases are highlighted in red.

Q2

Will we replace L1Trigger FPGAs with GPUs?

- Trigger FPGAs = **compute focused**, time critical
  - ATLAS and CMS implement neural network models on FPGAs
    - Tools like hls4ml or FINN allow model conversion between HLS and HDL.
  - Offline training of NN models in Python that can be compiled for both FPGA and GPU backends.

- BUT: Trigger FPGAs = compute focused, **time critical**
  - Next generation detectors will still have a high Bunch Crossing Rate (50MHz)
  - Latency: we still need ultra-low latency (~us)
  - Determinism: only with hard logic & DSPs

- If we decide to keep L1Triggering, we shall still need FPGAs
  - As long as manufacturers provide series like U and U+ families
    - a lot of serdes, a lot of DSPs

Q3

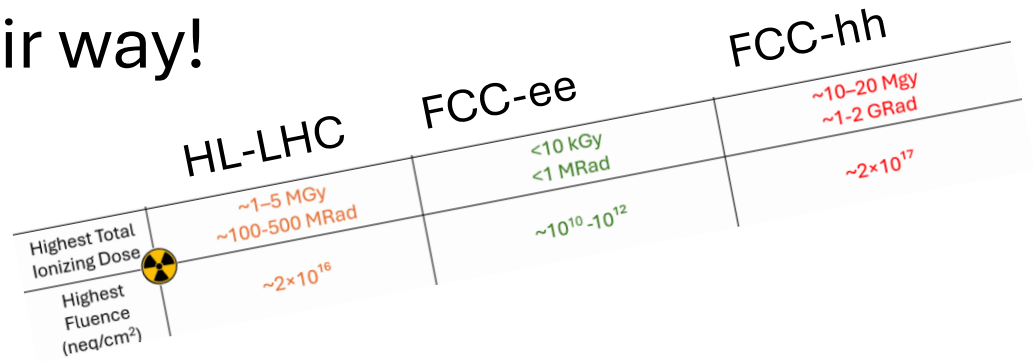
Could we get rid of FPGAs in future DAQs?

- «No FPGA in DAQ»
  - FE directly connected via switches to Network Interface Cards in Event Building Farms
  - …An old dream
- This pushes all the constraints down to Front End electronics:
  - Intelligence
  - Buffering and fast memory access
  - Reconfigurability
  - High speed and standard communication links directly from the Front-End
- Well…. This may end-up not being completely impossible!
  - R&D programmes exist on each of the points above (see spare slide)
  - AND Expected low levels of radiation could simplify our life
  - Still remains highly challenging

Q4

Will we be able to use FPGAs in the Front-End?

- FCC-ee detectors will less suffer from radiations
  - SRAM FPGAs may be usable in some locations
- FCC-hh: Rad-Hard SoC & eFPGA* are on their way!
  - Rad Tol RISC-V SoC
    - Ideal for slow control
  - Rad Hard eFPGA ASIC
    - Ideal for data aggregation & clustering
    - Configurable Logic Device
    - Rad Hard by design
      - Triplicated by default
      - Rad-Hard RAM blocks
  - Big challenge: Custom software tools

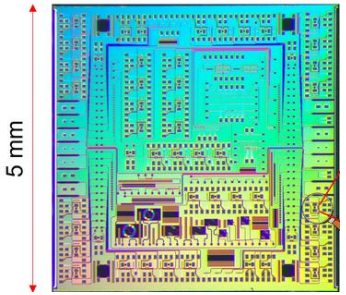


DRD7 WP7.2 [link]
CERN EP R&D WP5 & WP6 [link]

*eFPGA: embedded FPGA in an ASIC

Q5 Will the transceivers trend match the needs of future detectors?

- «slow» front-end links (Trigger and DAQ):
  - Currently «standard» Front-End links relatively low bandwidth
    - 10Gbps (aka lpGBT links)
  - Front End link will go up to 100Gbps (Silicon Photonics) but not faster
    - Due to Front-End limitations



DRD7 WP7.1 [link]
CERN EP R&D WP6 [link]

- “High bandwidth” back-end links (DAQ)
  - Standard protocols, ie Hard IPs needed
  - Typically now targeting 400Gb links (4x100Gbps)
  - Will probably increase with bandwidth need and availability

- FPGA Trend
  - See next slide

Q5 Will the transceivers trend match the needs of future detectors?

		type	Max Gb/s	Max #
AMD	Versal	GTY/GTM	32.7/58	44/52
AMD	Versal Premium	GTY/GTM	32.7/112	28/70
AMD	Virtex U+	GTY/GTM	32.7/58	128/438
AMD	Kintex U+	GTH/GTY	16/32.7	44/32
ALTERA	Agilex	GXE	58	32
ALTERA	Stratix-10	GX/GTX	29/58	24/120

→ (too) Good for DAQ

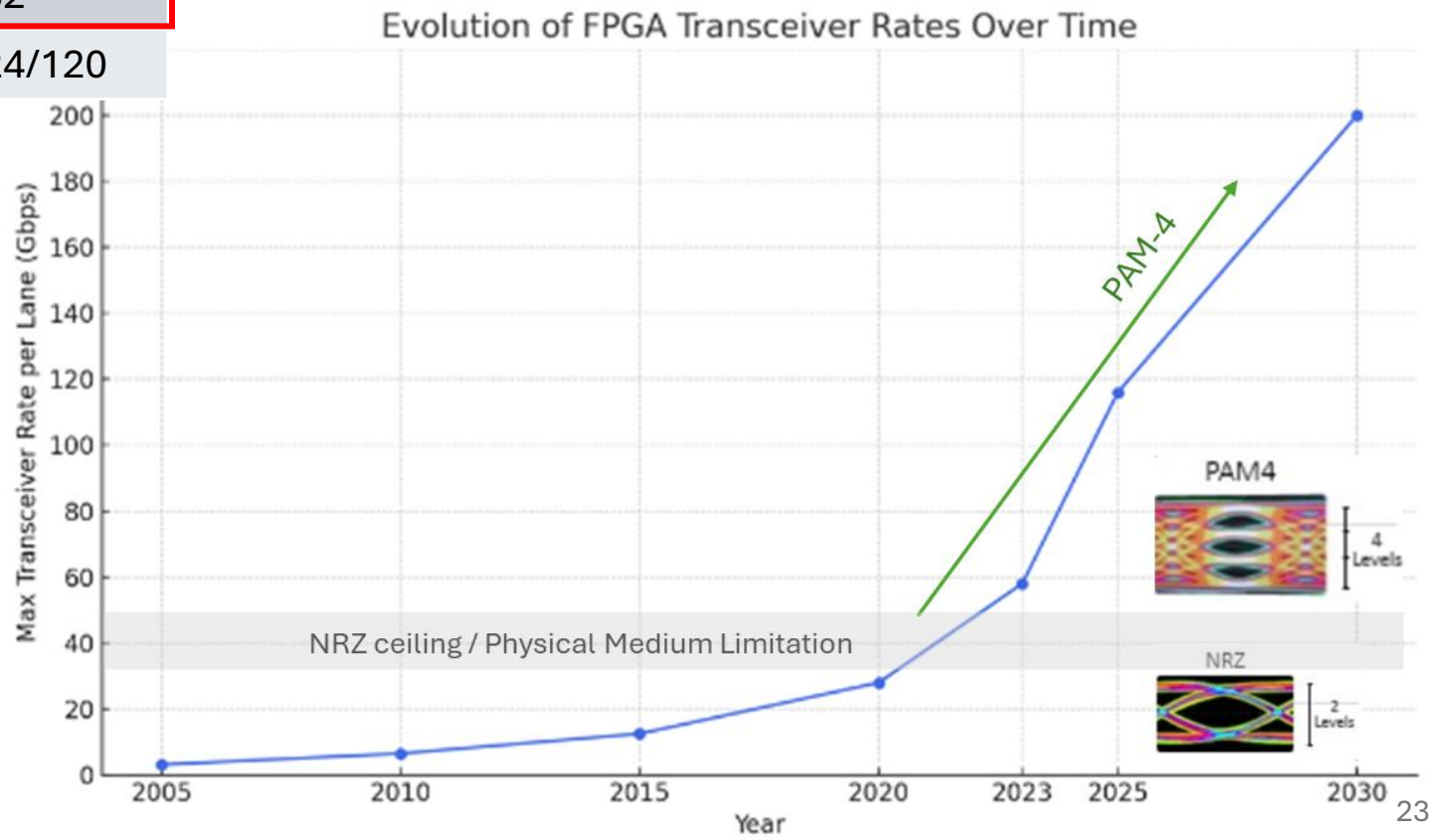
→ (too) Good for Trigger

$$$ - High-end families ! -$$$

=> HEP needs to join forces ahead to increase quantities and lower cost => «Standardize» FPGA type

[AMD link]

[ALTERA link]



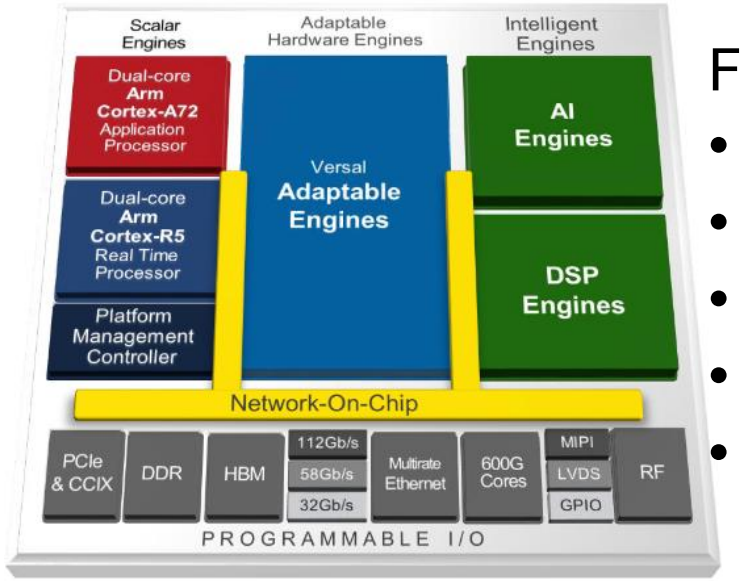
Q6 Will future FPGAs be compatible with high precision timing?

- Probably not ☹
  - Determinism is not a requirement from industry
  - Tx: Requires specific features in transceivers like «Phase Interpolators»
    - Not available in all transceivers/FPGAs (GTM, Altera FPGA)
    - High bandwidth transceivers are less and less configurable
  - Rx: each receiver reacts differently with resets
- Jitter in FPGAs tends to increase with technology evolution
  - The smaller the node, the higher the noise (especially pink noise, 1/f)
  - Clock quality is very sensitive to cross talk
    - A crowded FPGA degrades clock quality
- Potential solutions & avenues for R&D
  - Dedicate low end FPGAs to timing
  - Implement phase monitoring and noise reduction out of FPGAs

R&D in High Precision Timing [Link]
Paper from White Rabbit Team [link]

Q7 Is the Chiplet-trend interesting for FPGAs in future detectors?

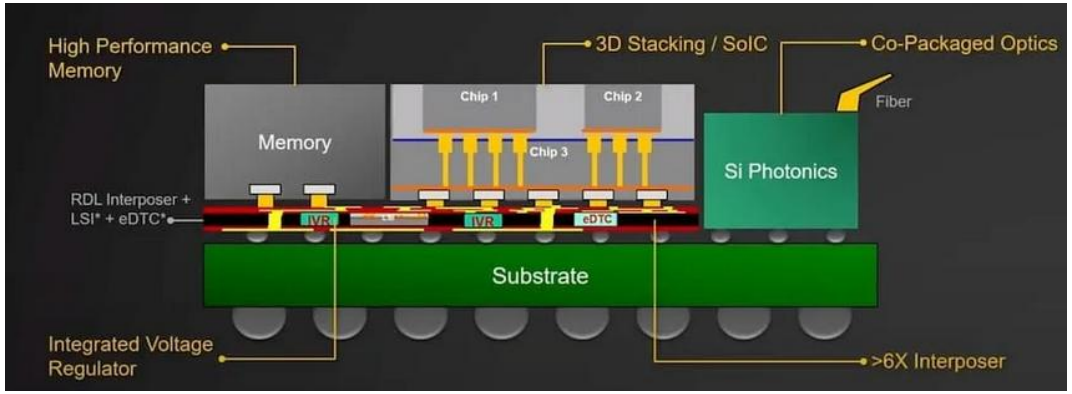
TODAY



FPGAs are now already built as chiplets

- ARM Processors: useful for management
- DSP Engines: a must for Trigger
- AI Engines: too slow & not deterministic enough for trigger
- HBM for fast & large memory access
- Hard IP for transceivers, PCIe or Ethernet

TOMORROW



Upcoming:

- Copackaged optics
  - Optical signal straight from FPGA
  - No copper limitation anymore
- Faster interconnect
  - Chip-to-chip for Triggering?
  - Faster ethernet links for DAQ

Q7

Chiplets: The special case of HBMs

- HBM: High Bandwidth Memories
  - Co-packaged in some High-End FPGAs
  - **VERY** interesting especially for DAQ purpose
- Manufactured by a handful of companies
  - SK hynix and Samsung are the dominant suppliers
- Super short turnaround time between Generations
  - 2-3 years
  - HBMx remains on the market max 7-10 years <<< FPGAs (~25 years)
  - Risk of early EOL if market drops
- Reduces the lifecycle of FPGAs with HBMs wrt «traditional» FPGAs
  - Contrast with HEP requirements - long development cycles
  - Risk mitigation for HEP: HBM lifecycle tracking & early procurement

	Versal™ Adaptive SoC	UltraScale+	UltraScale™
Block RAM	239 Mb	94 Mb	36 Kb
UltraRAM	717 Mb	360 Mb	–
High Bandwidth Memory (HBM)	32 GB	16 GB	–
External Max Data Rate	8533 Mb/s	2667 Mb/s	2400 Mb/s

Q8

On which types of boards will our FPGAs sit?

- So far, tried to base form factors to existing standards
  - Long time with VME crates
  - Then moved to uTCA (very small modular shelves) and ATCA (very large, for HL-LHC)
    - Now EOL !
  - In parallel, FHFL PCIe AIC turned to be an interesting & dense solution
    - Typical form factor of GPU & AI accelerator cards
    - High density & directly connected to the server CPU
    - But
      - Cooling quite challenging
      - Specific servers required
- May have to find a new/custom form factor if servers not available anymore
  - Currently looking at standalone boxes
  - Like 1U-rack mountable «pizza-box»
  - Cooling will be a challenge

Q9

Should I stop asking questions now?

THANKS

Feel free to ask yours or share your thoughts!

Special thanks to the numerous colleagues with whom I spoke to put this material together

Spare Slides

References: R&D on Front-End enabling no DAQ FPGAs

Intelligence on the Front-End	DRD & CERN R&D programs on -intelligence on the detector -Radiation Tolerant RISC-V System-On-Chip	DRD7 WP7.2 [link] CERN EP R&D WP5 [link]
Buffering and fast memory access	Low radiation levels in FCCee detectors could make this possible	
Reconfigurability	CERN R&D programme on Radiation Hard eFPGA	CERN EP R&D WP6 [link]
High speed and standard communication links directly from the Front-End	DRD & CERN R&D programmes on -100Gbps FE links (Silicon Photonics) -100GbE4FE	DRD7 WP7.1 [link] DRD7 WP7.5 [link] CERN EP R&D WP6 [link]