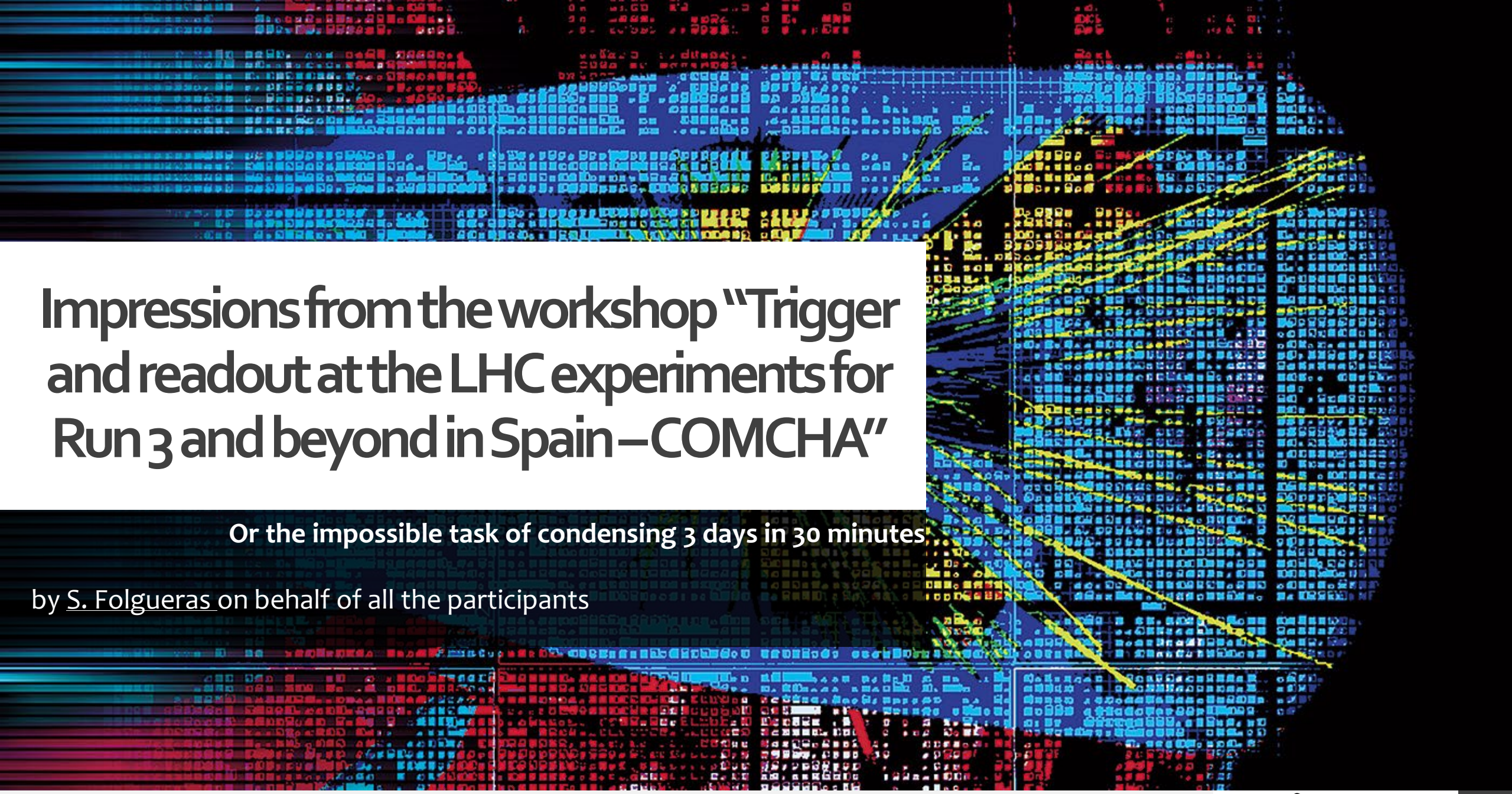


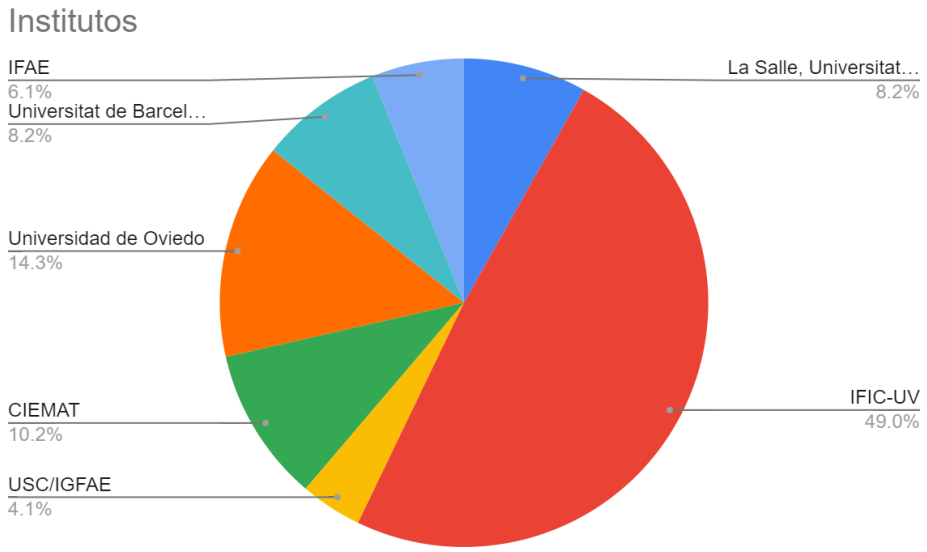
Impressions from the workshop "Trigger and readout at the LHC experiments for Run 3 and beyond in Spain –COMCHA"

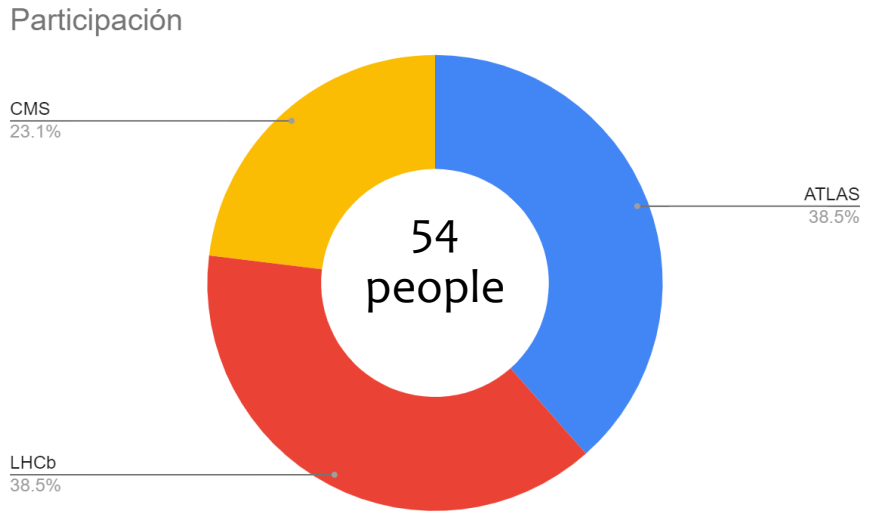
Or the impossible task of condensing 3 days in 30 minutes

by S. Folgueras on behalf of all the participants

Introduction

The goal of this workshop was to give an overview of the plans of the different groups working in the trigger and DAQ systems of the LHC experiments at CERN for Run 3 and beyond, as well as to establish new synergies…





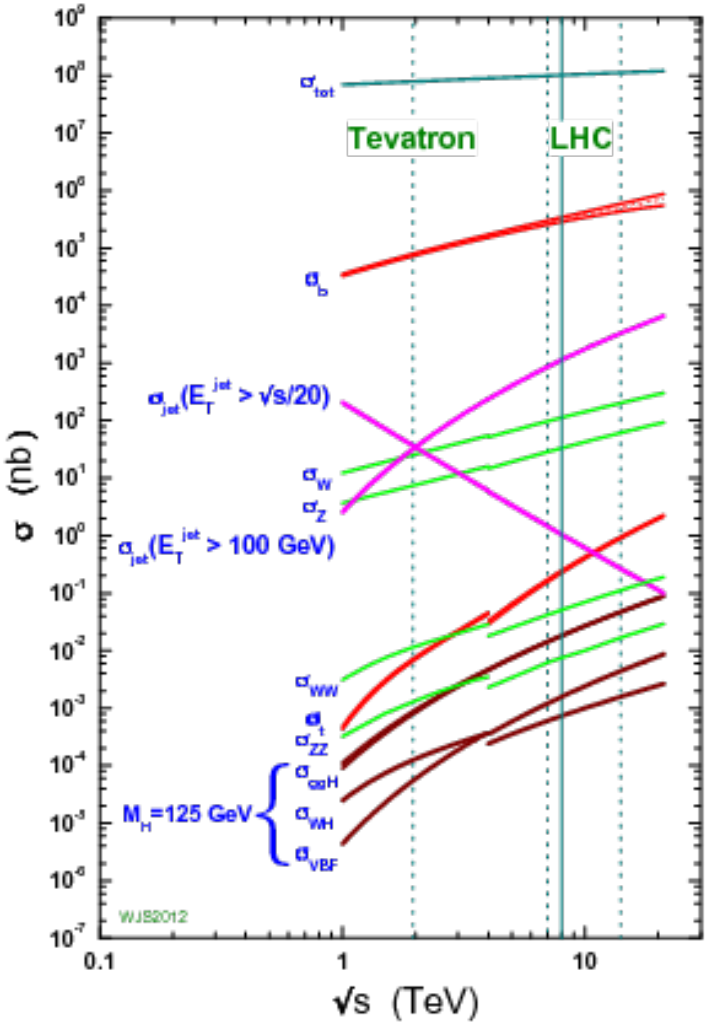
Great success and attendance!

What is a trigger system and why it is needed?

- Store all the possibly relevant data provided by the sensors is UNREALISTIC and often becomes also UNDESIRABLE.
- You **CANNOT / DON'T WANT TO** store every event

$$BW\ (kB/s) = rate\ (Hz) \times size\ (kB)$$

- Risk loosing interesting physics processes with low cross-sections.

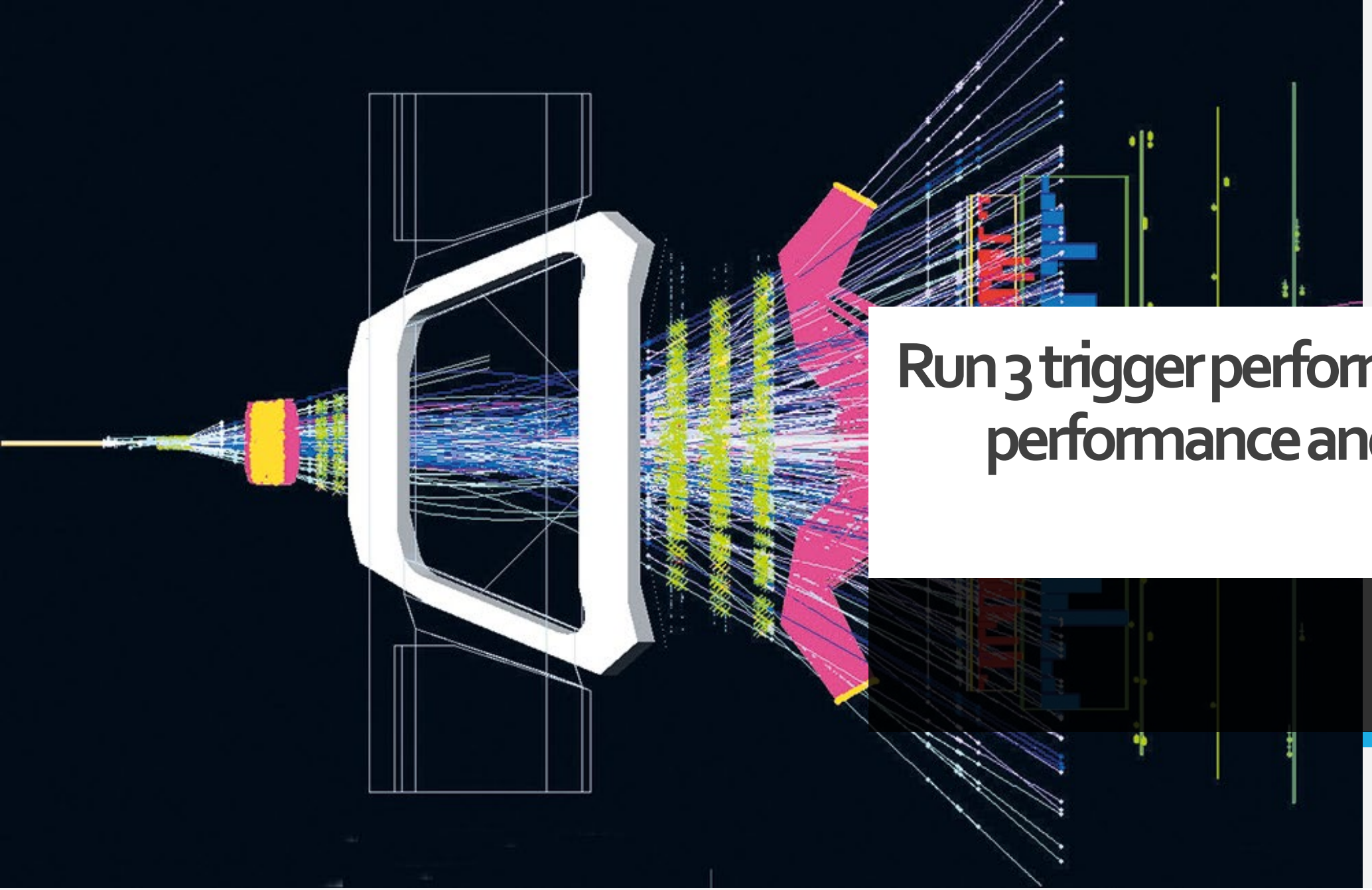


Where do we contribute to?

- **Run 3 trigger performance, early performance and challenges**
  - Upgrades, commissioning and operations (CMS, ATLAS, LHCb)
- **High-speed electronics devices for trigger and DAQ**
  - Board design, assembly, validation, production, and commissioning: off-detector electronics for TileCal (ATLAS), on-detector board for DT (CMS)
  - Board validation, production, and commissioning: back-end electronics for the barrel muon trigger (CMS)
  - Board commissioning: overlap muon track-finder (CMS)
- **Trigger Algorithms, ML / AI applications running on both non-CPU HW and CPUs**
  - Long-lived particles (LHCb, CMS)
  - Muon tracking (LHCb, CMS)
  - Calorimeter reconstruction (LHCb, ATLAS)
  - Topological trigger (ATLAS, LHCb)

Will show a glimpse of what was discussed and showed in Valencia

https://indico.ific.uv.es/event/6923/timetable/?view=standard



Run 3 trigger performance, early performance and challenges

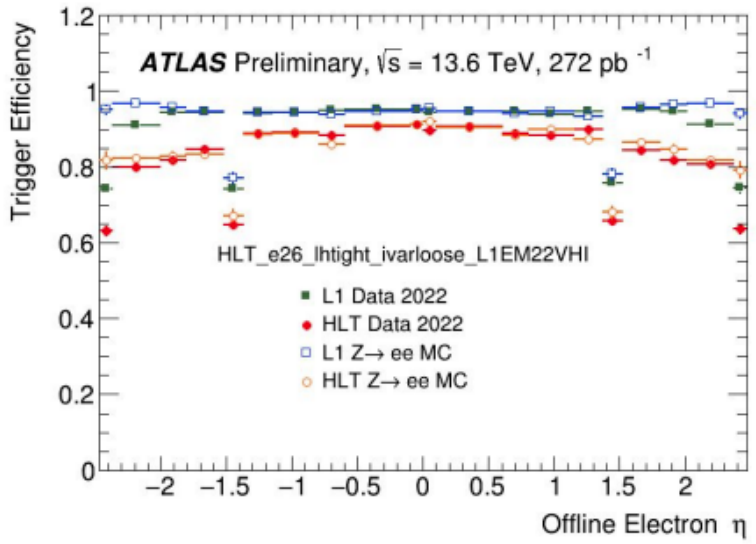
"First" look at Run-3 data: ATLAS

V. Senthilkumar

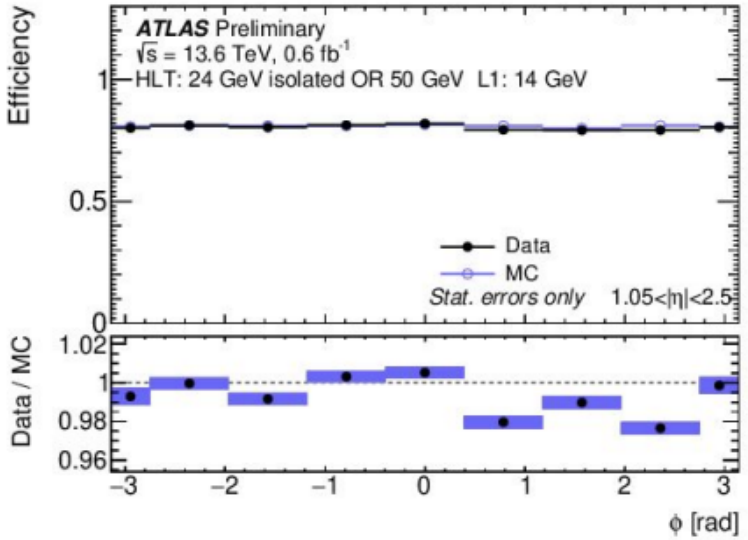
Good efficiency observed at L1 & HLT for the lowest unprescaled single electron triggers

Isolation of electrons from jet clusters at L1

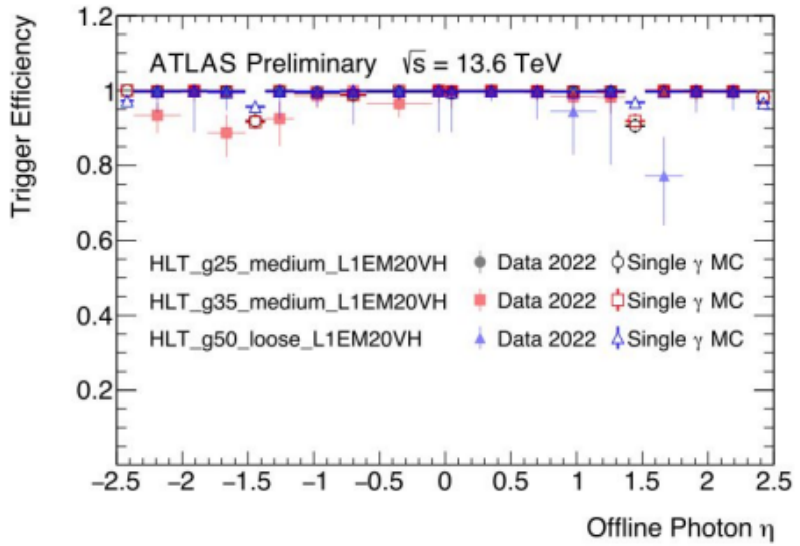
Data efficiency not corrected for backgrounds



Good efficiency observed at HLT for lowest unprescaled single muon triggers in the barrel region



Good efficiencies for photon chains using the lowest threshold L1 EM triggers without isolation

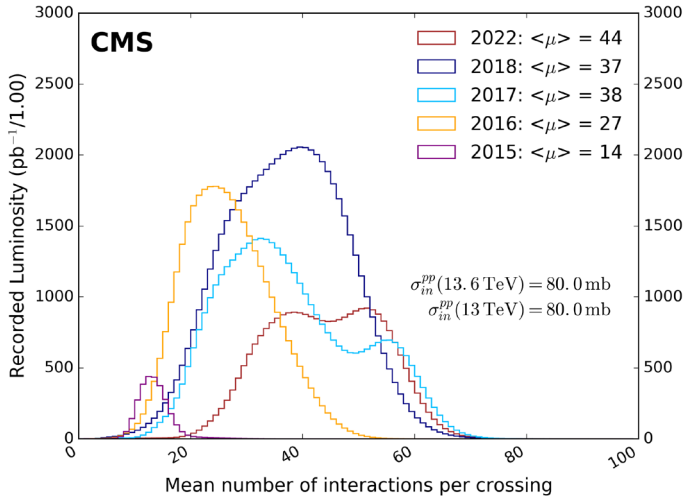


EgammaTriggerPublicResults, MuonTriggerPublicResults

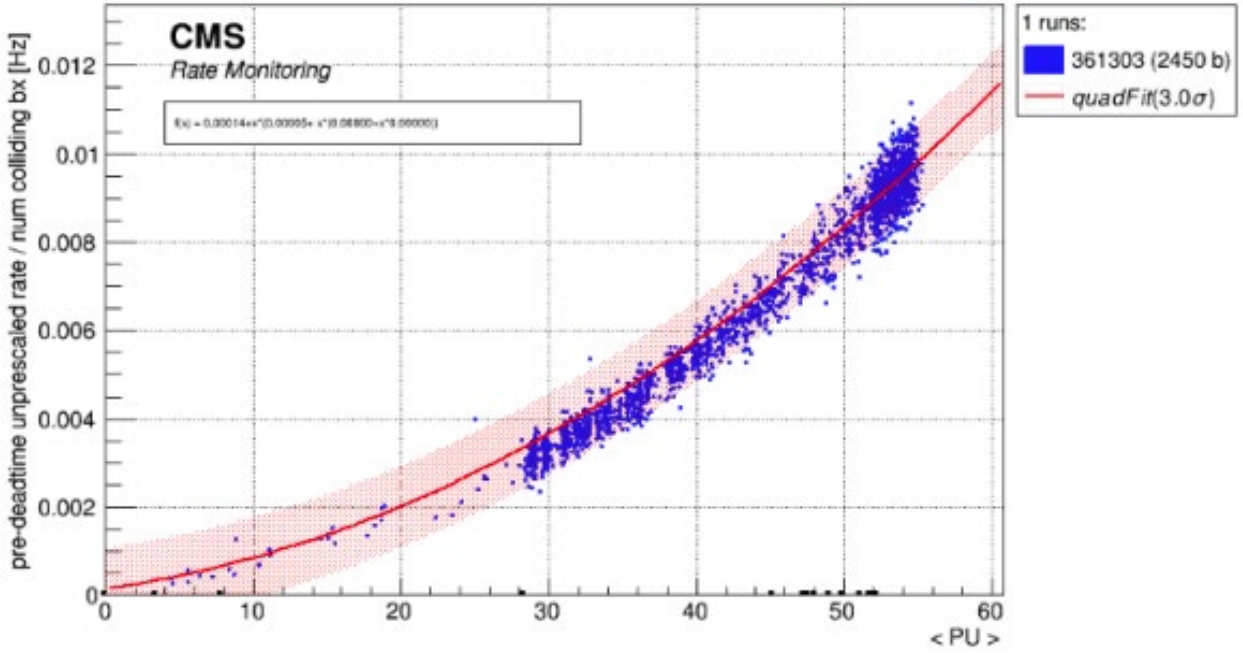
"First" look at Run-3 data: CMS

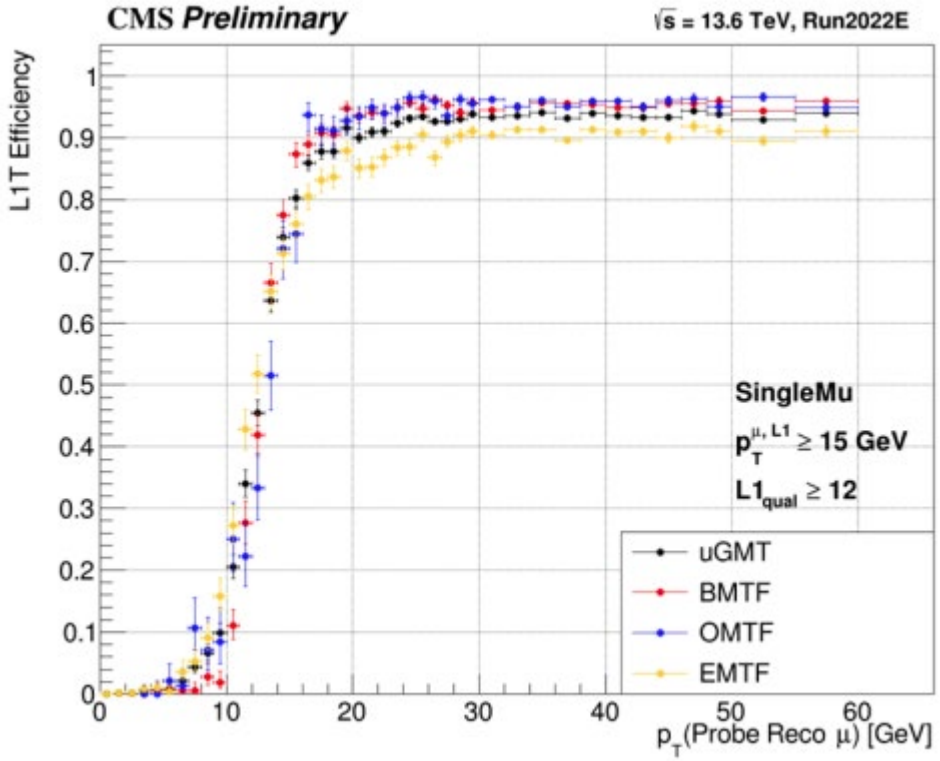
- Excellent performance of muon triggers both for prompt and displaced muon signatures.
- Exciting physics program ahead of us.

A. Soto



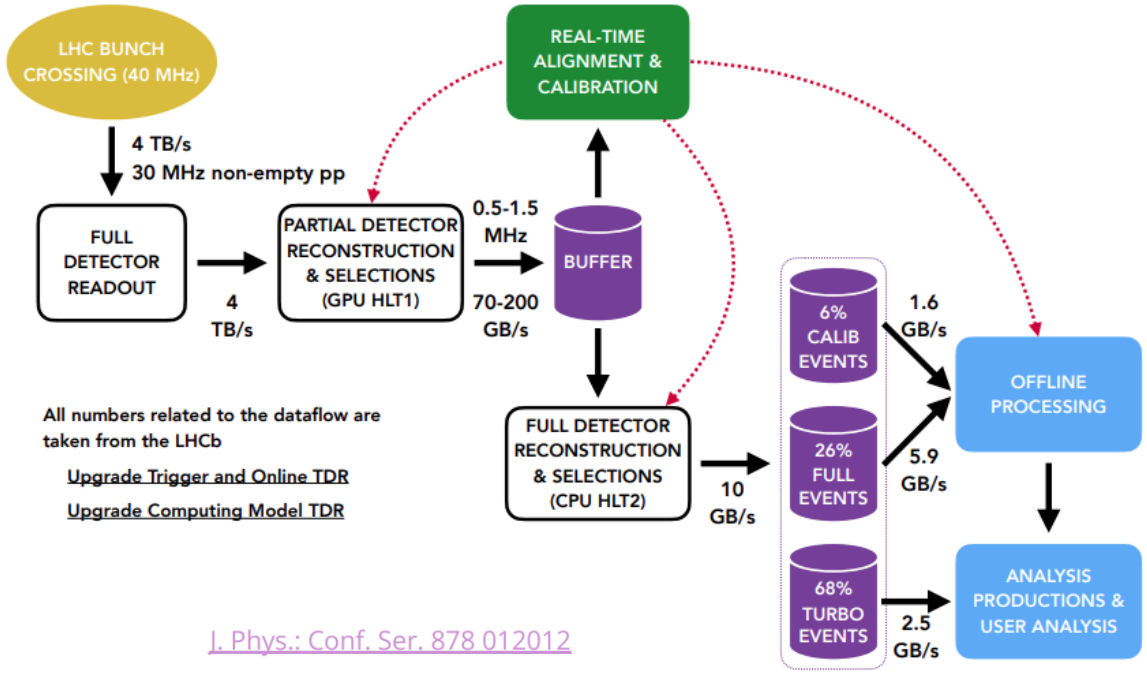
HLT_DoubleL3Mu16_10NoVtx_DxyMin0p01cm_v1





The first upgrade

- LHCb switched to a full readout model at 30 MHz to overcome limitations of the two-level system.
- Relying on GPUs to perform tracking in different subdetectors (VELO, UT, SciFi)



JINST 14 P04006

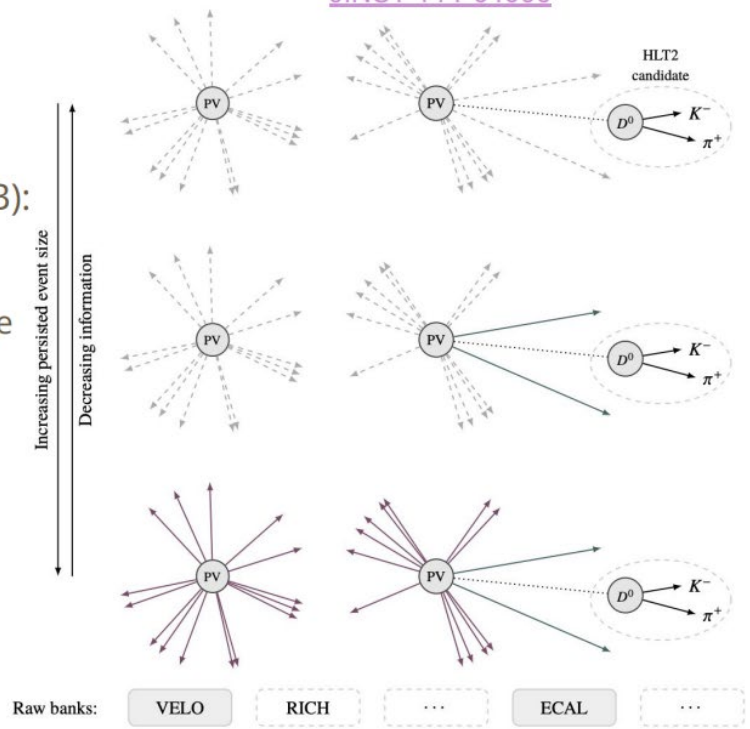
The turbo model

BW(kB/s) = rate(Hz) x event size(kB):
4 TB/s input → 10 GB/s offline limit

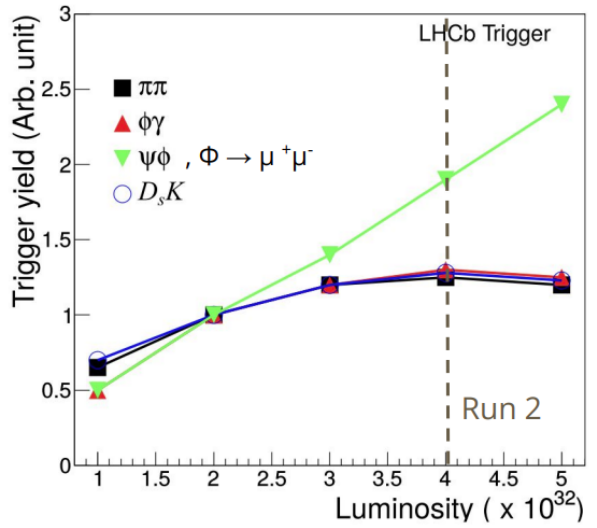
- huge signal rate → reduce evt size

Flexible persistence model:

- **Turbo** (35 kB): signal only
- **Full** (70 kB): all reco'ed objects
- **Selective**: signal + selection of reconstructed objects and raw banks



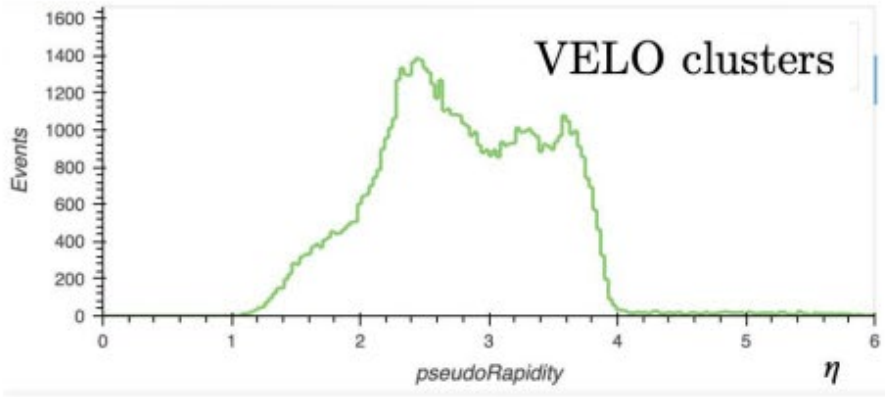
J. Phys.: Conf. Ser. 878 012012

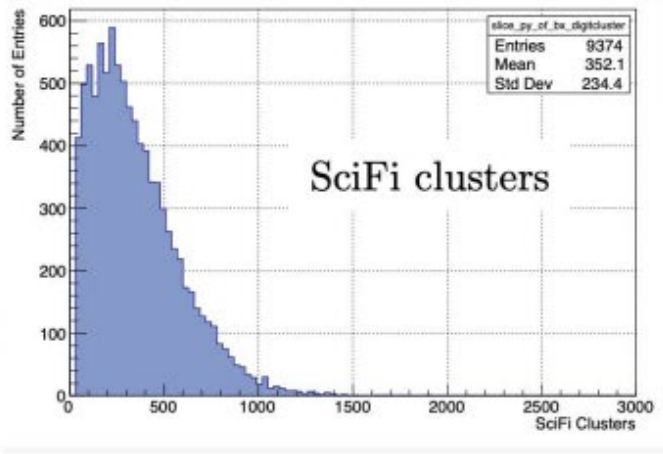


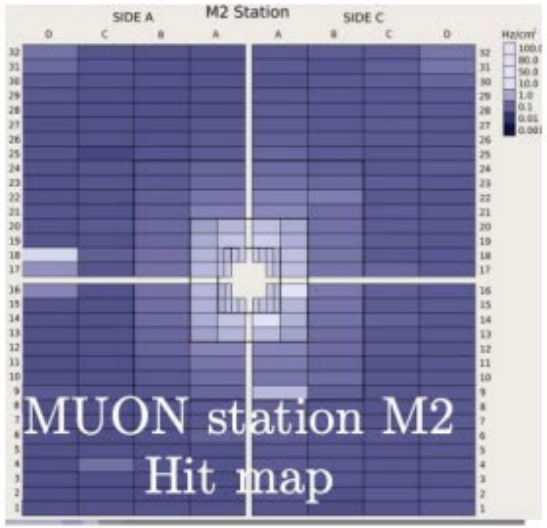
C. Marin

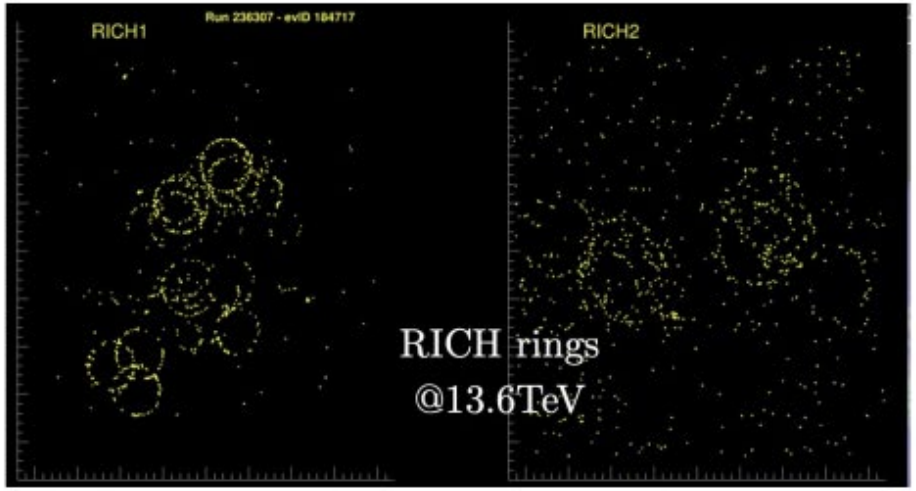
Now it's time for commisioning

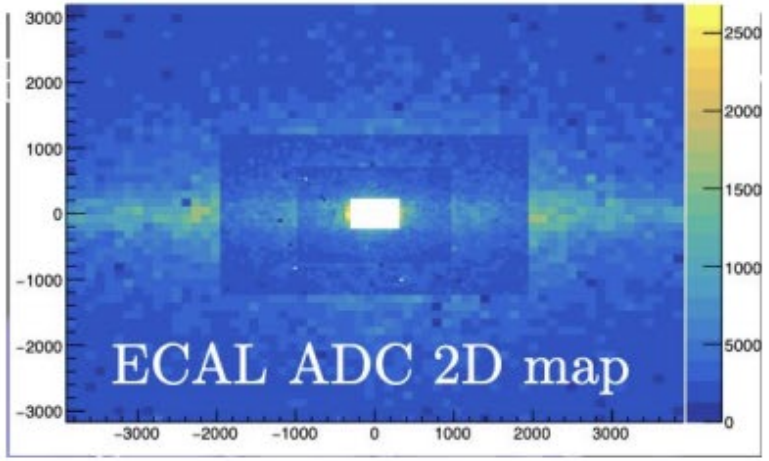
First Run 3 data











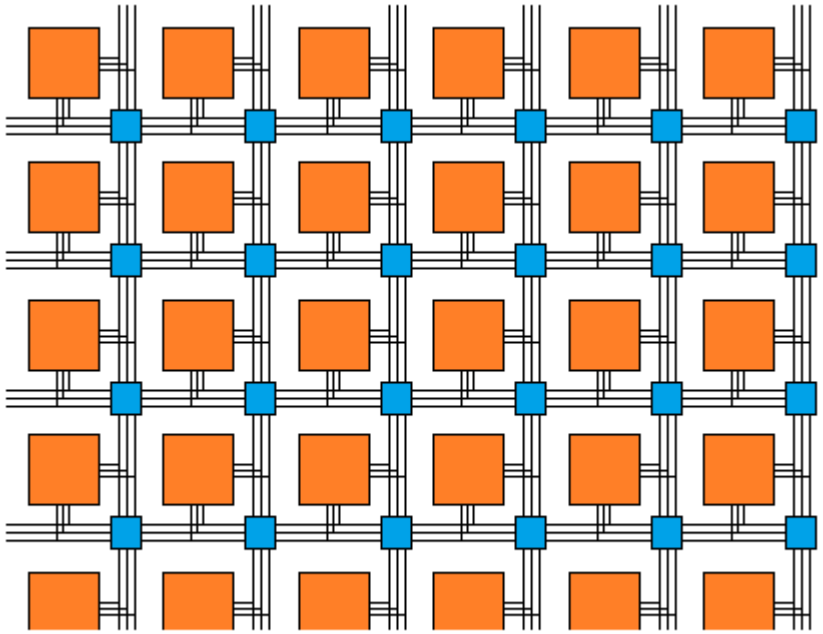
C. Marin

25

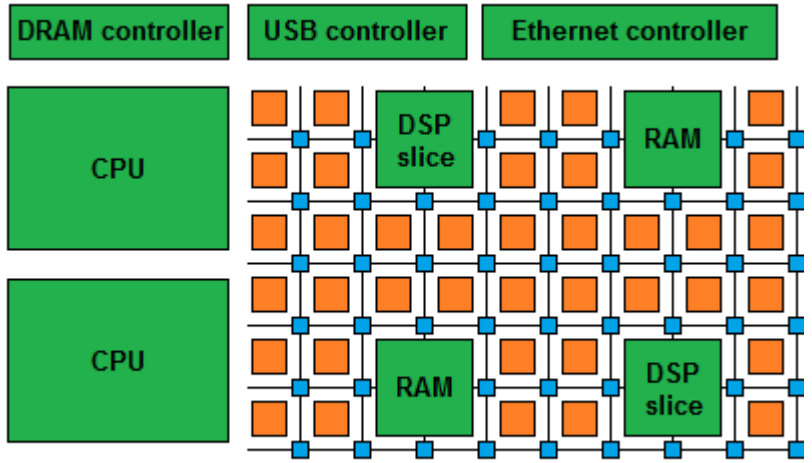
High-speed electronics for trigger and DAQ

What is an FPGA?

- ***FPGA stands for Field Programmable Gate Array**, essentially it's a piece of hardware that can be programmed as many times as the user wants and it can convert or implement any arbitrary equation into the form of the boolean equation, consequently, implement this as combinational and sequential logic. Simply put, an FPGA can be used to implement any logic function.*
- FPGAs are extensively used at the LHC for both on- and off-detector electronics where high-speed processing of signals is required.



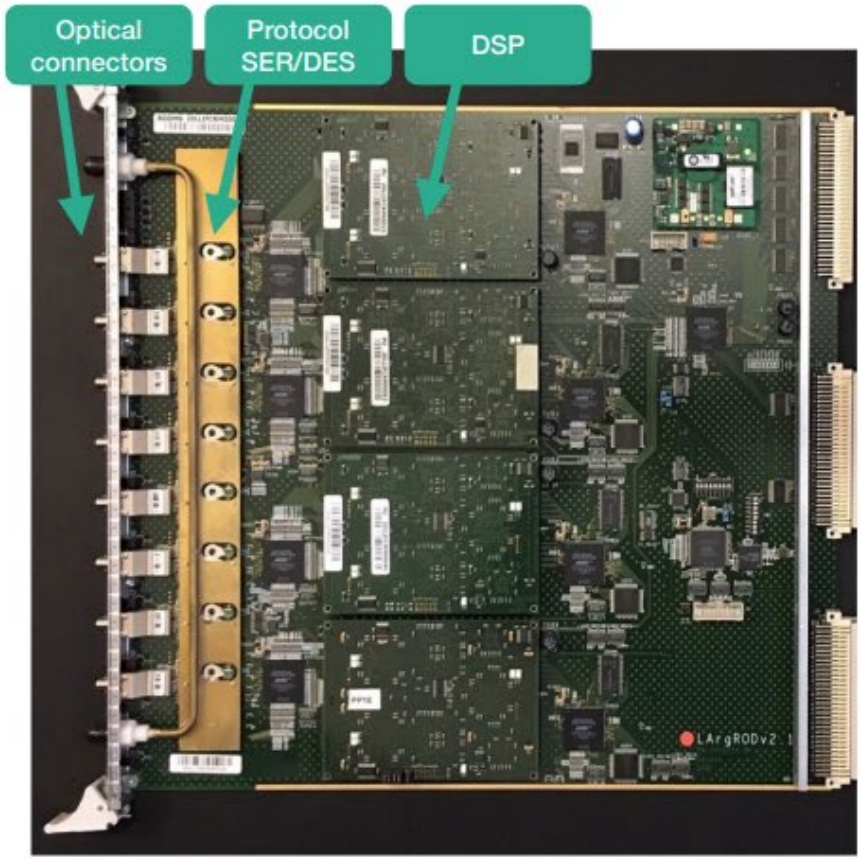
"Clean slate" FPGA: programmable gates and routers



Modern FPGA: lots of hard, not-field-programmable gates

ATLAS Tile Calorimeter – upgrades

ROD (2000): VME9U ~ 34 x 36 cm – DSP processing
Throughput (Max 1 Gb/s x line): RX: 5.12 Gb/s ; TX: 4 Gb/s
8 optical receivers, 4 G-Link ser/des, 22 FPGAs / 8 DSPs

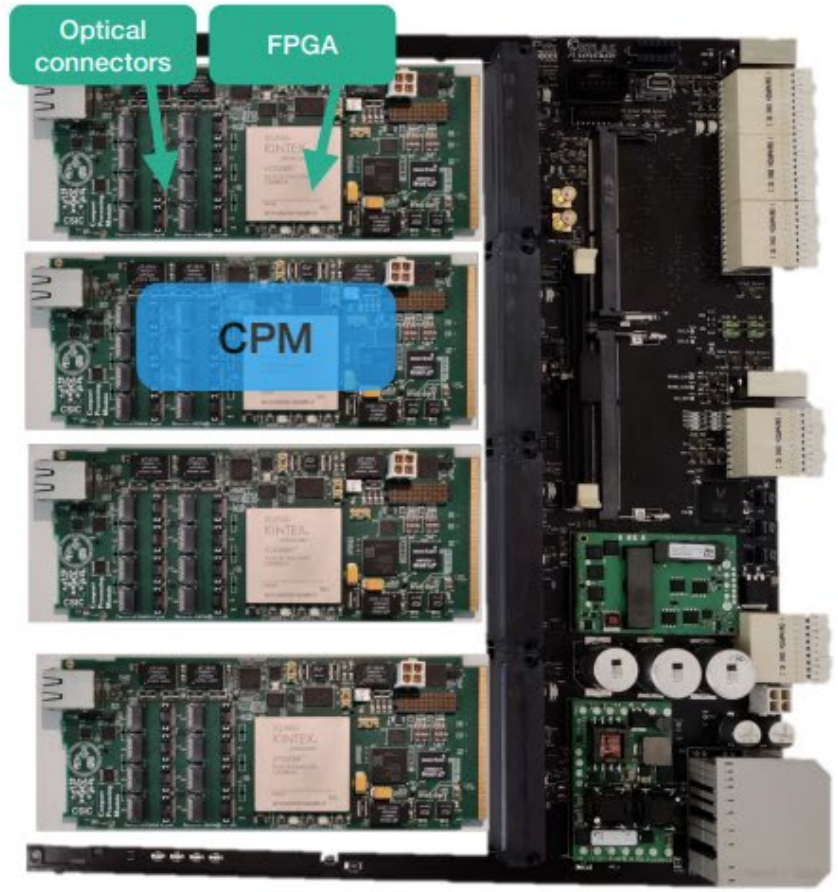


Huge increase in aggregated data rates not only due to individual line speeds

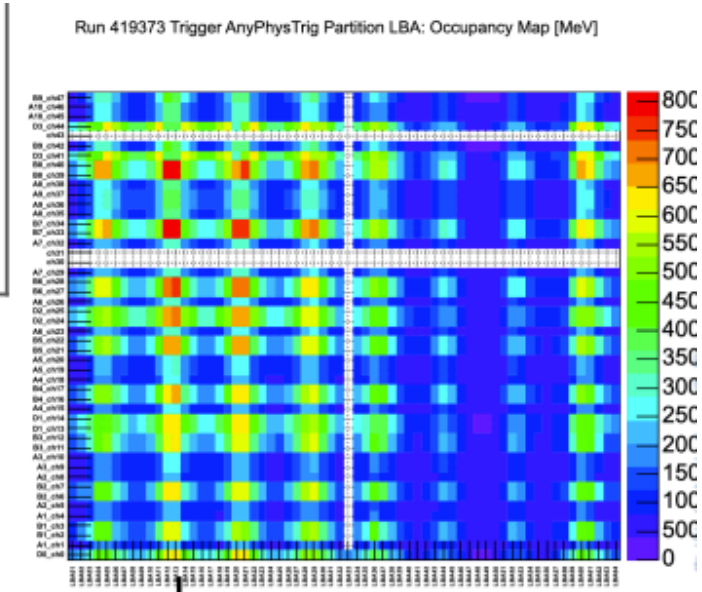
	Current	Phase II
Total BW	165 Gbps	40 Tbps
# fibers	256	4096
BW/module	800 Mbps	160 Gbps
# cards	32 (VME)	32 (ATCA)
# crates	4 (VME)	4 (ATCA)
BW/card	6.4 Gbps	1.28 Tbps

8

TillePPr (2020): ATCA8U ~ 35 x 28 cm - FPGA processing
Throughput (16 Gb/s line): RX: 2 Tb/s Gb/s ; TX: 1 Tb/s
24 Firefly (x4 : 96 links) , 10 FPGA

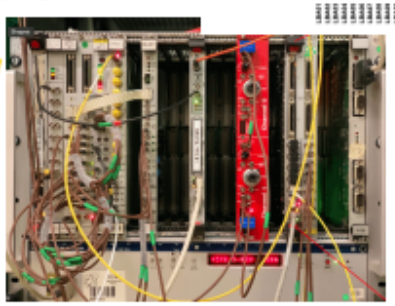


PPr	PreProcessor
FELIX	Front End Link eXchange
ROD	Read Out Driver
TTC	Timing, Trigger and Control
ALTI	ATLAS Local Trigger Interface
DBv6	DaughterBoard version 6
YETS	Year-End Technical Stop



Demonstrator

- New hybrid electronics for Phase-II Upgrade
- Module inserted in LBA14 in July 2019
- Tested in ATLAS since M1 (May 2020)
- Backward compatibility with the current ATLAS TDAQ system
- Very useful to exercise new electronics and identify potential problems before Phase-II
- Recent upgrade to DBv6 during YETS



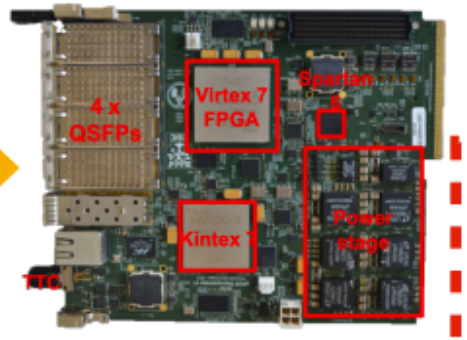
TTC / ALTI system

Demonstrator during splashes in 2022

L1A/TTC - 80 Mbps

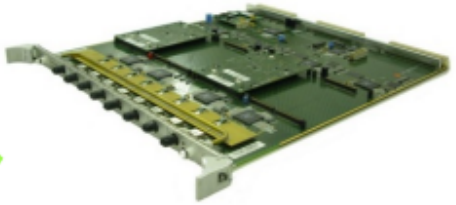
Demonstrator module (4 x mini drawers)

4.8/9.6 Gbps



PPr Demonstrator

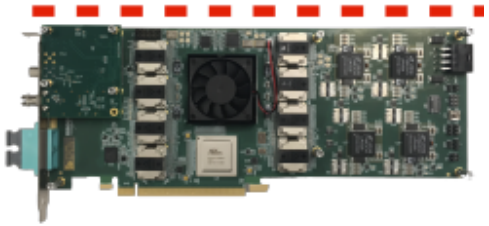
800 Mbps



Legacy ROD

4.8/9.6 Gbps

Next Step



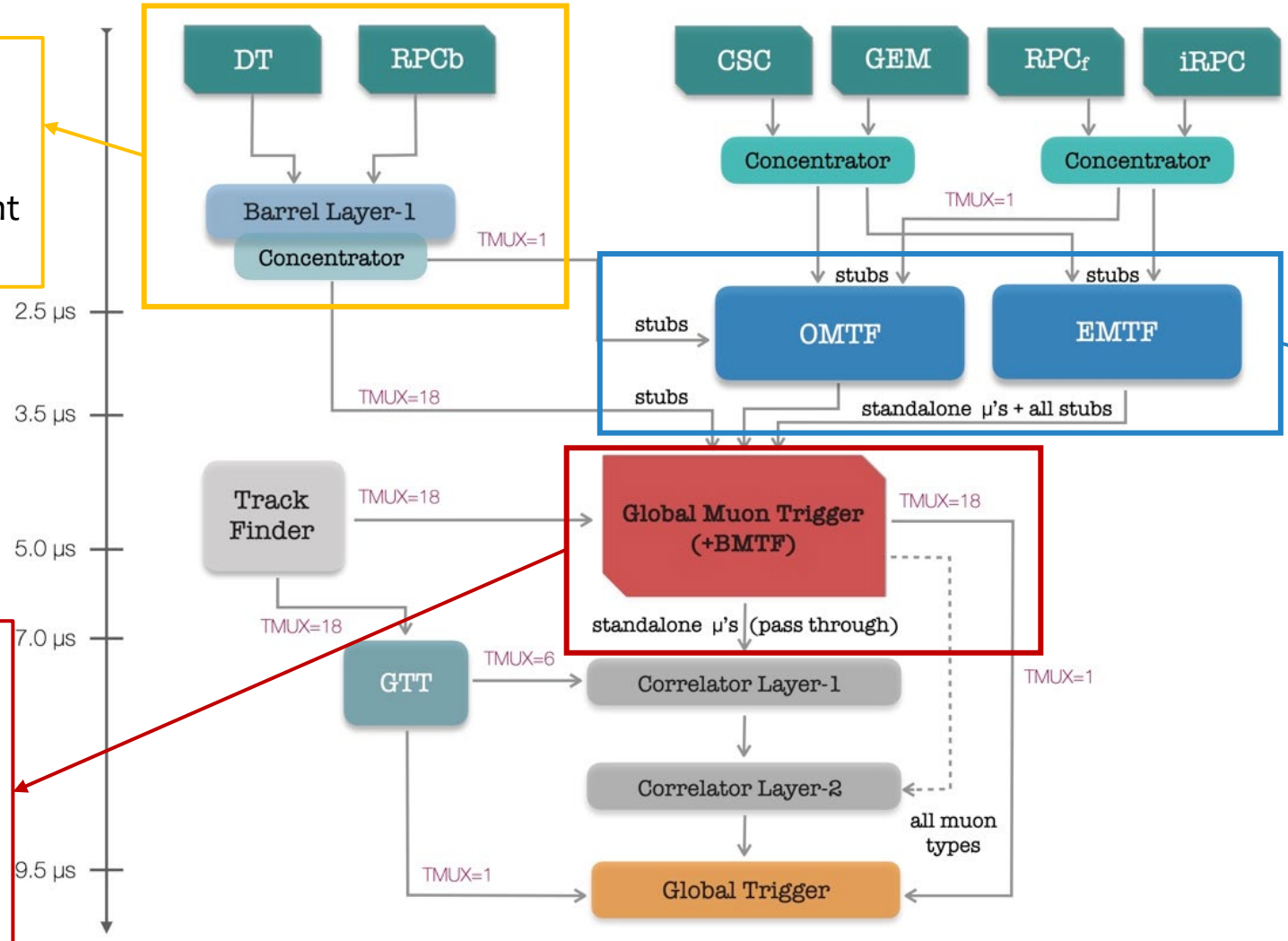
FELIX System

CMS Phase-2 Muon trigger architecture

Barrel Layer-1: local information, reconstructs segment direction in the different muon stations.

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas · UAM Universidad Autónoma de Madrid · Universidad de Oviedo

Hardware, software and firmware development.



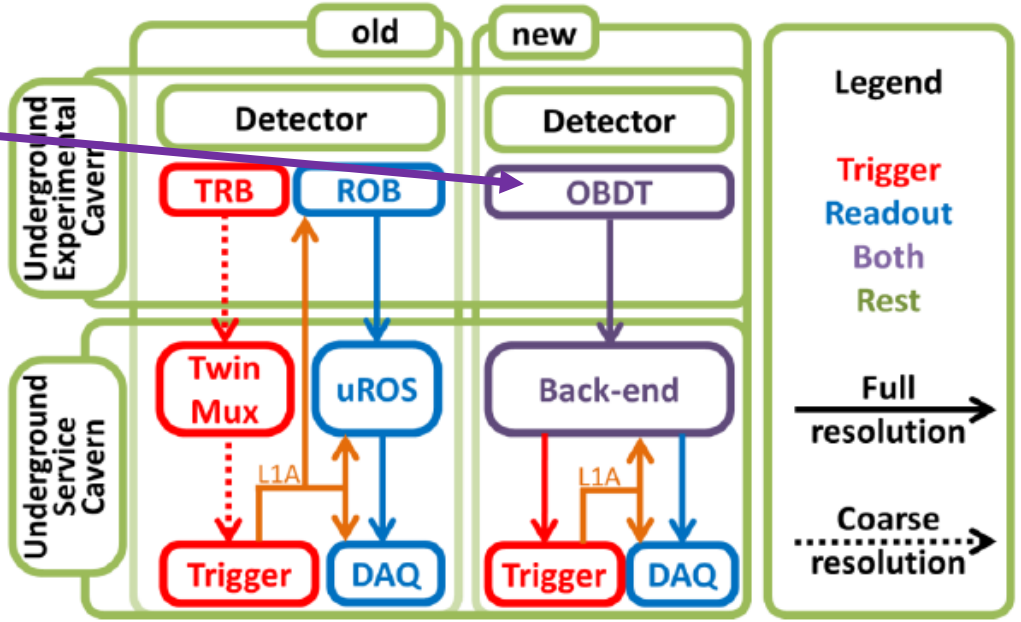
Global muon trigger: Combines muon info (local and regional) with tracker information to reconstruct muons with ultimate momentum precision.

Muon Track Finders: Combine local info from different stations and subdetectors to estimate muon trajectory and momentum.

Software and firmware development.

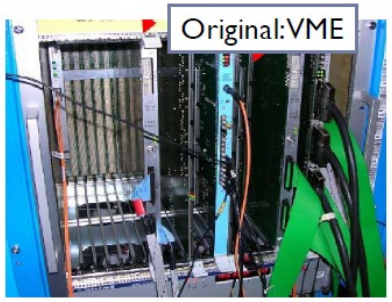
Universidad de Oviedo

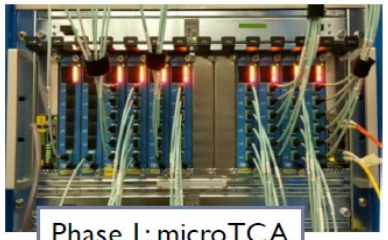
CMS DT Phase-2 Upgrade for HL-LHC



- Legacy system limitations:
  - Electronics not able to deal with increased processing demand
  - Links not able to deliver the increased data rates
  - Trigger algorithm works with coarse TDC resolution

DT Backend





- Fast-forward 20 years, we can take advantage of technological progress
- Not just dealing with increased rates… Also improve performance
- Availability of radiation-tolerant FPGAs and high-speed serial links over optical fibre means we can stream everything out of the experimental cavern
- Availability of high-performance FPGAs means we can perform a much finer trigger algorithm using full TDC resolution

CMS Muon DT trigger HW evolution

Original
50 000 on-detector ASICs
300 Actel ProAsicPlus FPGAs
(~ 2M equivalent Logic Cells)
Luminosity 10^{34} cm^{-2} s^{-1}

Phase 1
50 000 on-detector ASICs
60 Xilinx Virtex7 FPGAs
(~ 20M Logic Cells)
Luminosity $2 \cdot 10^{34}$ cm^{-2} s^{-1}

Phase 2
42 Xilinx UltraScale Plus FPGAs (~ 70M Logic Cells)
Luminosity 10^{35} cm^{-2} s^{-1}

Trigger algorithm data flow

Grouping:
Each new hit is paired with other hits in its vicinity. Combinatorial explosion under high noise. Has to keep up with hit input rate.

Laterality provider:
Hypotheses on wire side laterality. Finer prediction saves wasting expensive fitter time.

Fitting. Linear regression. Computationally expensive.

Filter: reduce the combinatorial explosion. Keep only the highest-quality segment among the ones that share hits.

In each superlayer (x2)

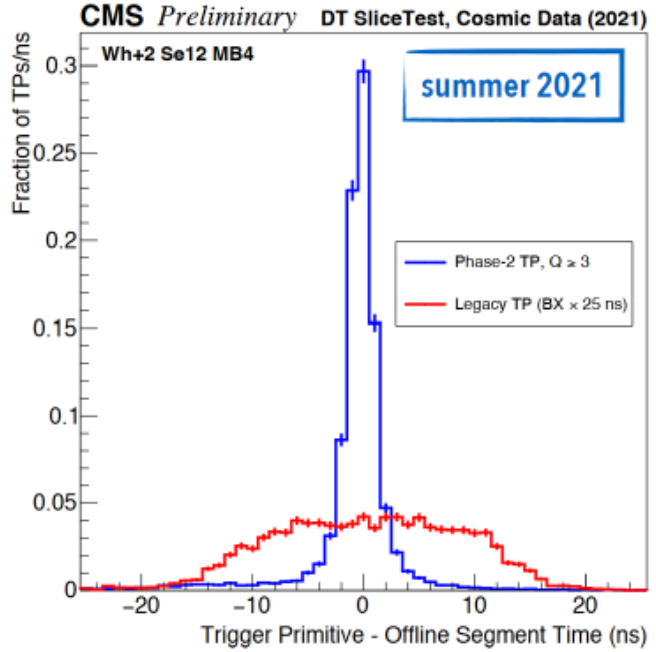
Matcher: do all possible (viable) pairings between segments from each superlayer. Combinatorial explosion, again.

Fitter. Linear regression, again. But with more hits.

Filter. Reduce the combinatorial explosion, again.

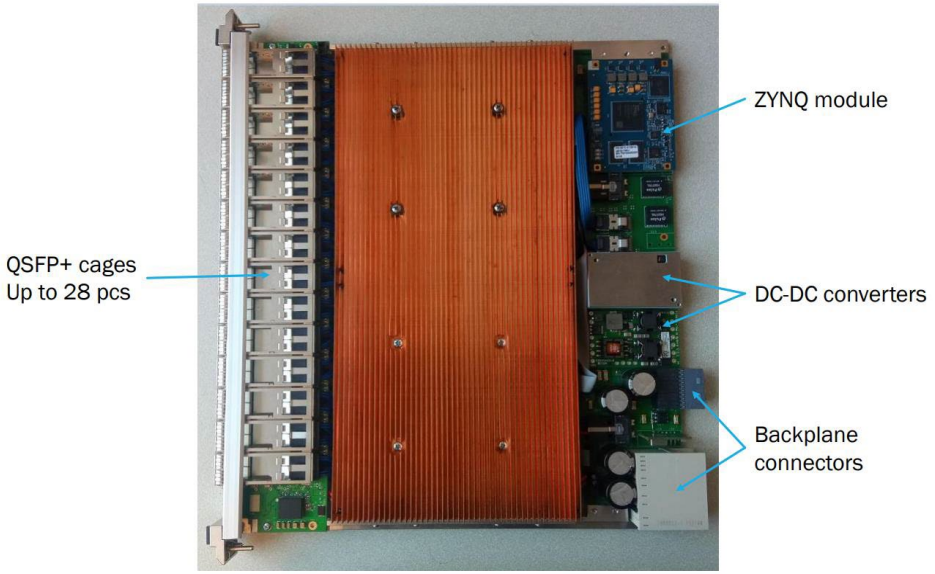
- Latency ~12 BX, occupancy 45 kLUT*
  * ~200 € (2.6%) worth of VU13P FPGA
- $3 \cdot 10^8$ Fits/s, 10^{10} Filter Comparisons/s, $7 \cdot 10^9$ matcher pairings/s

12 A. Navarro - Electronics for CMS DT



Muon track finder in CMS

- We have implemented an optimized version of the OMTF algorithm in HLS, with lower latency and area.
- Modularized small IPs provide a wider vision at RTL level of what's going on and allows to debug faster
- Firmware testbench has been done, algorithm works as expected, timing closes.

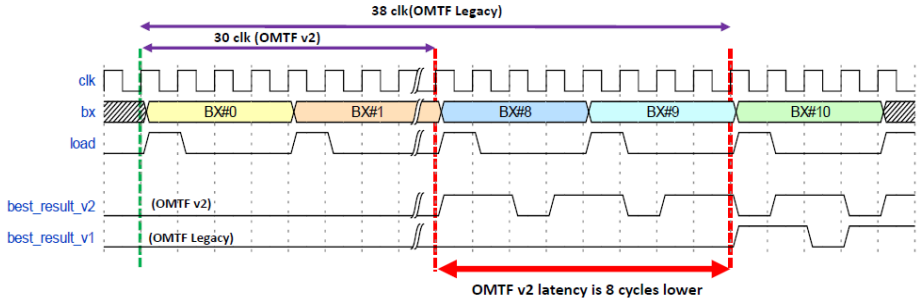


Latency results

P. Leguina

Universidad de Oviedo

Module	InputProcessor	GoldenProcessor	BestPatternSelector
Latency (cycles)	4	12	12



- Total latency of the module is 28 + 2 (FFs) cycles.
- Algorithm maximum frequency is ≈ 200 MHz (with this latency)
- The new HLS version shows an 8-cycle improvement when compared to the legacy (VHDL) version.

Latency has improved compared to OMTF legacy, reduced by 8 cycles due to the algorithm being fully pipelined and all data structures partitioned.

Implementation

Universidad de Oviedo

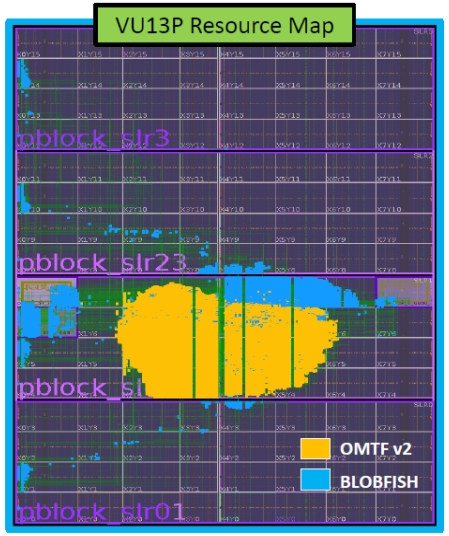
- Successful implementation and validation of OMTF algorithm in HLS.
- Acceptable resource utilization (1 SLR, < 25%).
- Latency and timing constraints are met.
- OMTF has been implemented along with BLOBFISH[1] framework.

(1) Basic Level Object Firmware Integration Shell (BLOBFISH) is a package made by the UCLA team which provides firmware infrastructure adapted to the board by using custom IP cores and tcl scripting. The core of the framework is the slr module that is an IP core that talks to the outside world through a 64-bit AXI interface.

OMTF	LUT	FF	DSP	BRAM
Legacy*	9,13%	4,08%	0,00%	8,70%
v2	7,66%	4,01%	0,00%	9,10%

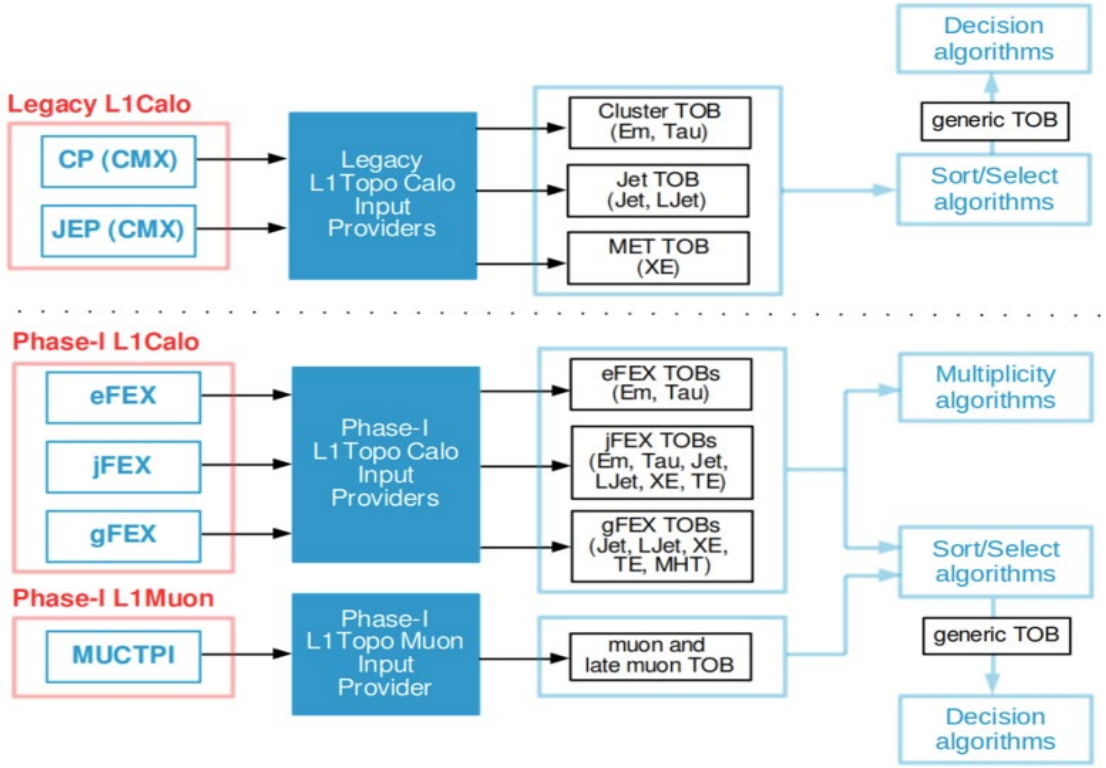
Utilization of OMTF v2 < 25% of a VU13P

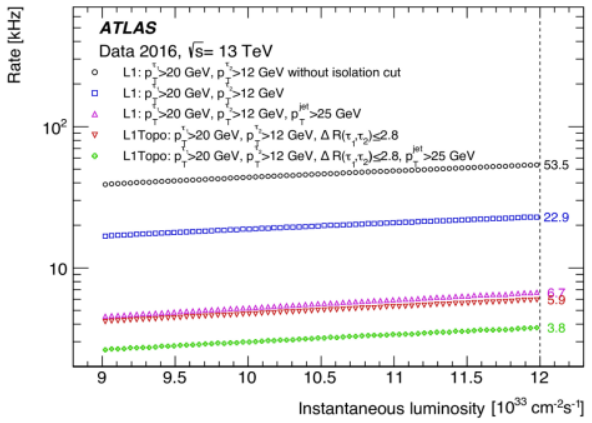
(*) Utilization % scaled from *xcvc1502 to xcvu13p*

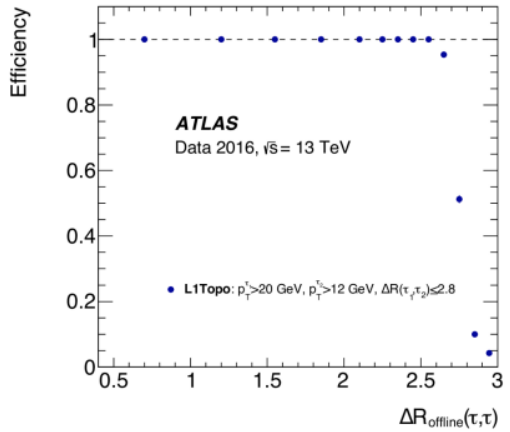


The ATLAS L1 Topological trigger

- The L1Topo system takes inputs from the Level-1 Calorimeter (L1Calo) and Level-1 Muon (L1Muon) systems and outputs decisions to the Central Trigger Processor (CTP).
- Level-1 Topological system has undergone several changes in Run-3 (Phase-1) to improve the trigger system.
- The upgrades bring:
  - Improved input granularities from L1Calo/Muon systems
  - More precise LUTs
  - New topological algorithms
  - Multiplicity calculation



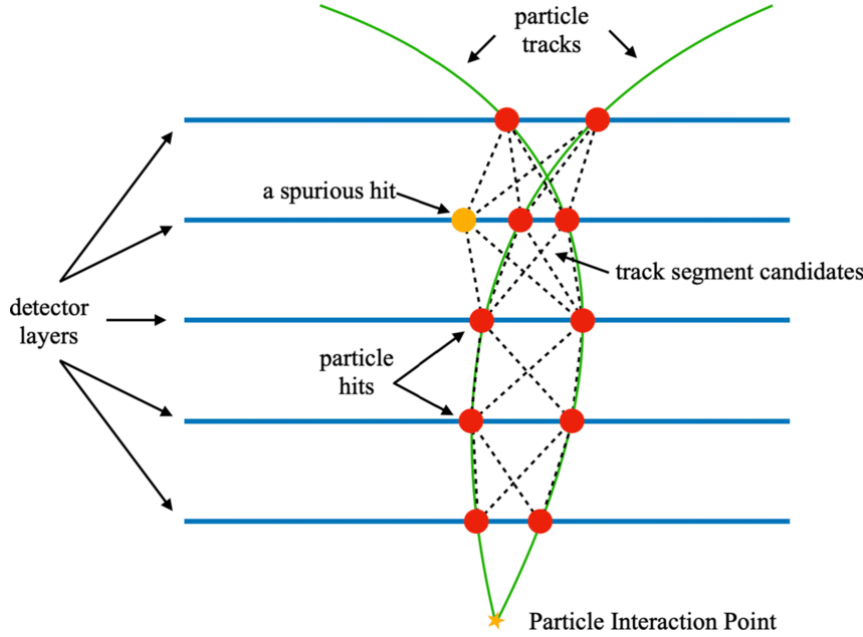




Efficiency relative to the equivalent non-topological trigger.

Algorithms: an overview

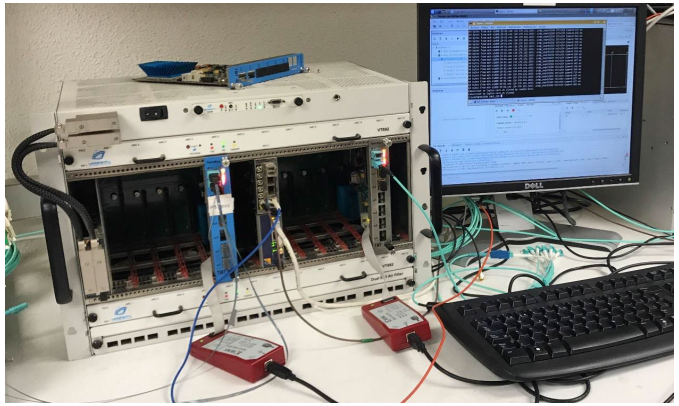
- Great expertise in particle tracking in different architectures: CPUs, GPUs, FPGAs
  - Exploring event reconstruction with GPU-only
  - Mixed approaches with CPU+GPUs
  - The use of FPGAs is completely standard in the hardware trigger
- Trend goes towards heterogenous infrastructures!
- **LHCb has completely re-designed the trigger application on GPUs.** [https://allen-doc.docs.cern.ch/index.html] Filtering an input rate of 30MHz down to 1-2MHz. It runs fast track reconstruction and selecting pp collision events based on one- and two-track objects entirely on GPUs.
- **CMS** uses GPUs for dedicated areas of the reconstruction (tracking) and CPUs for the rest of the High-Level-Trigger. FPGAs are used for hardware trigger.

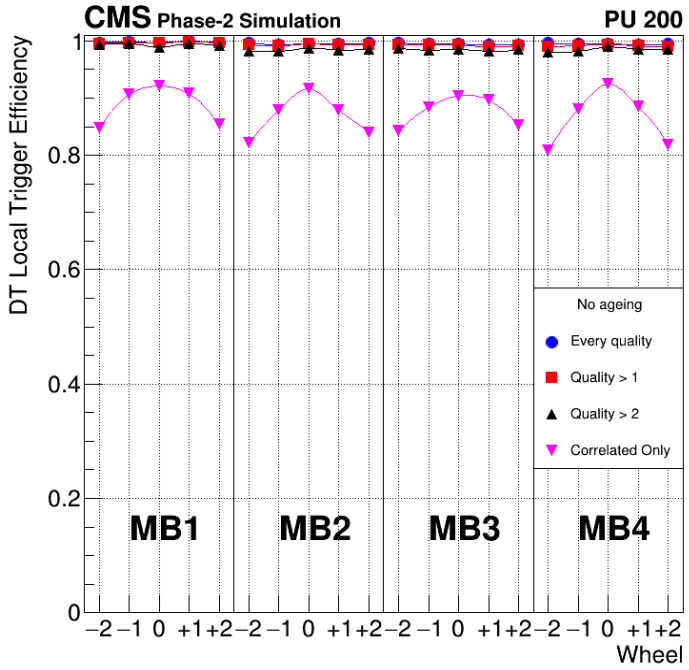


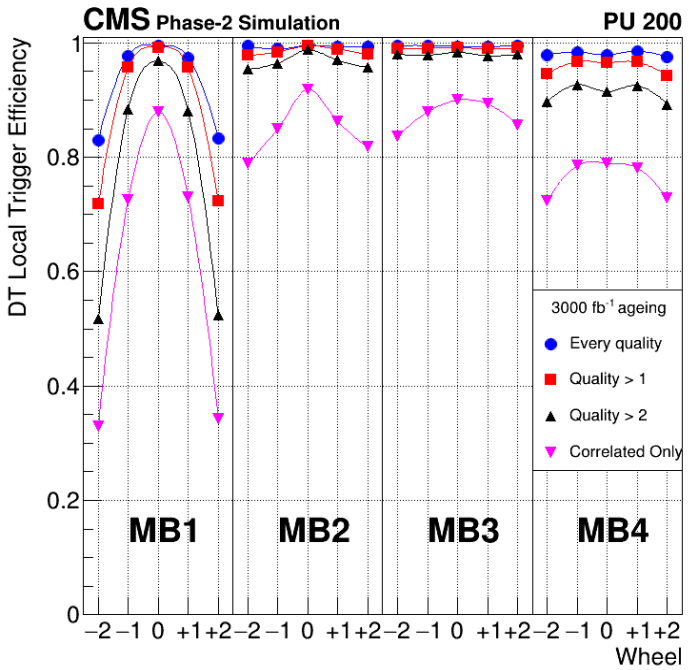
- **ATLAS** uses FPGAs for the hardware trigger and CPUs for the High-Level-Trigger.
- Our teams have successfully developed (sophisticated) algorithms for all the architectures.

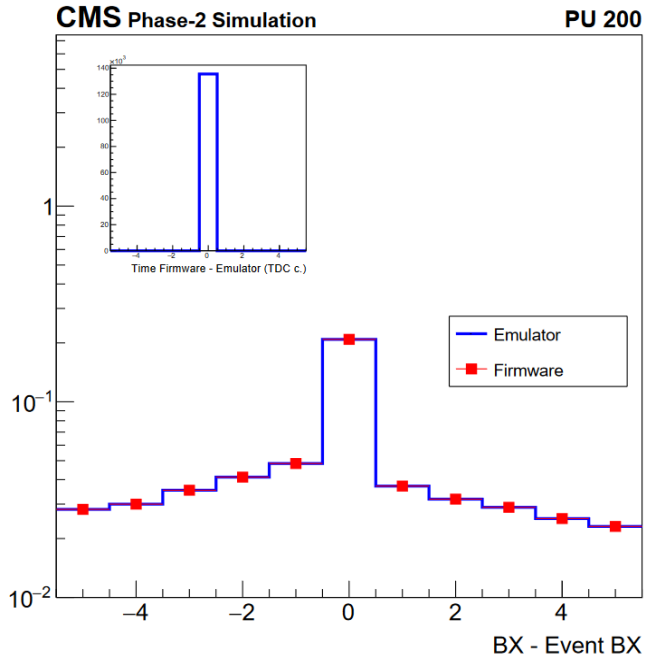
L1T Muon tracking algorithm in the CMS DT for the HL-LHC

- The Analytical Method is an algorithm implementing analytical solutions for reconstructing the DT trigger primitives for HL LHC.
- Performance studies show very good results, both in efficiency and in resolution.
- The AM firmware is on a good shape, showing a very good agreement with the AM emulator and providing good results at the DT Slice Test (already with Run 3 collision data).



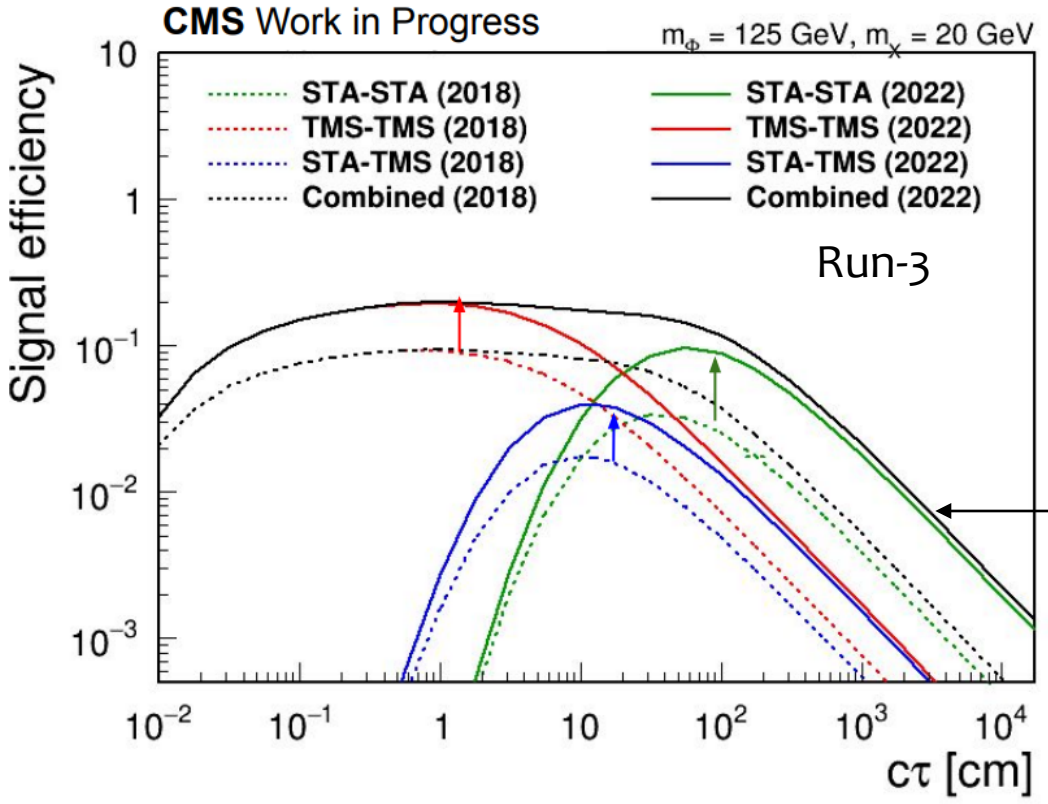


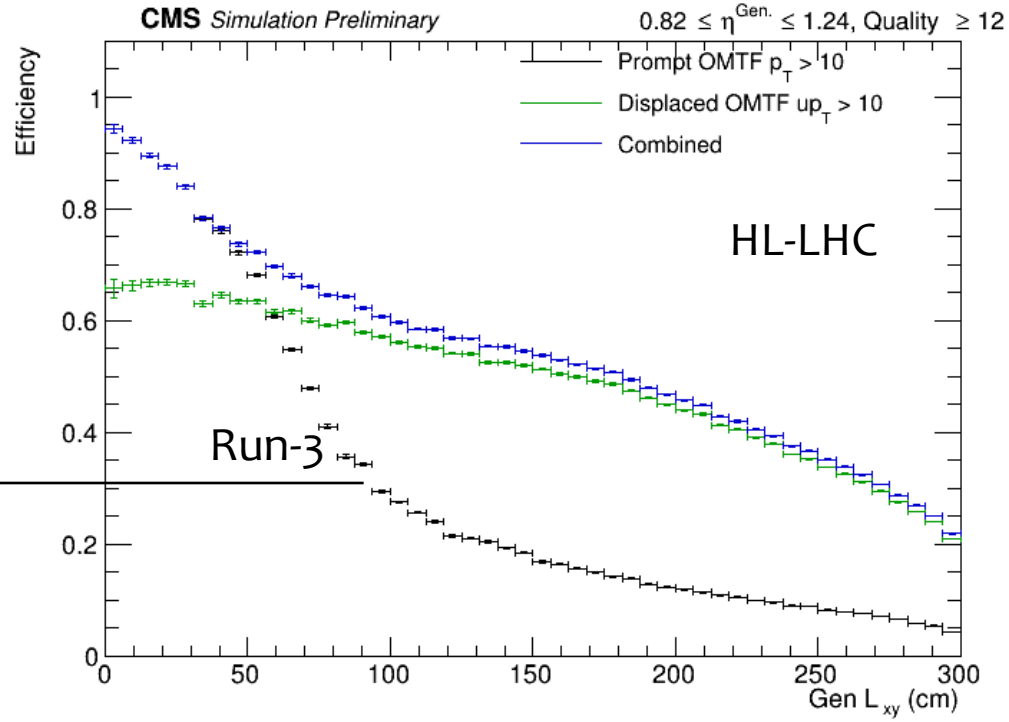




Algorithms: long-lived particles (CMS)

- Work to improve Run-3 sensitivity, but still limited by L1 efficiency.
- HL-LHC should be able to overcome this limitation with improved displaced (muon) reconstruction in the track finder.

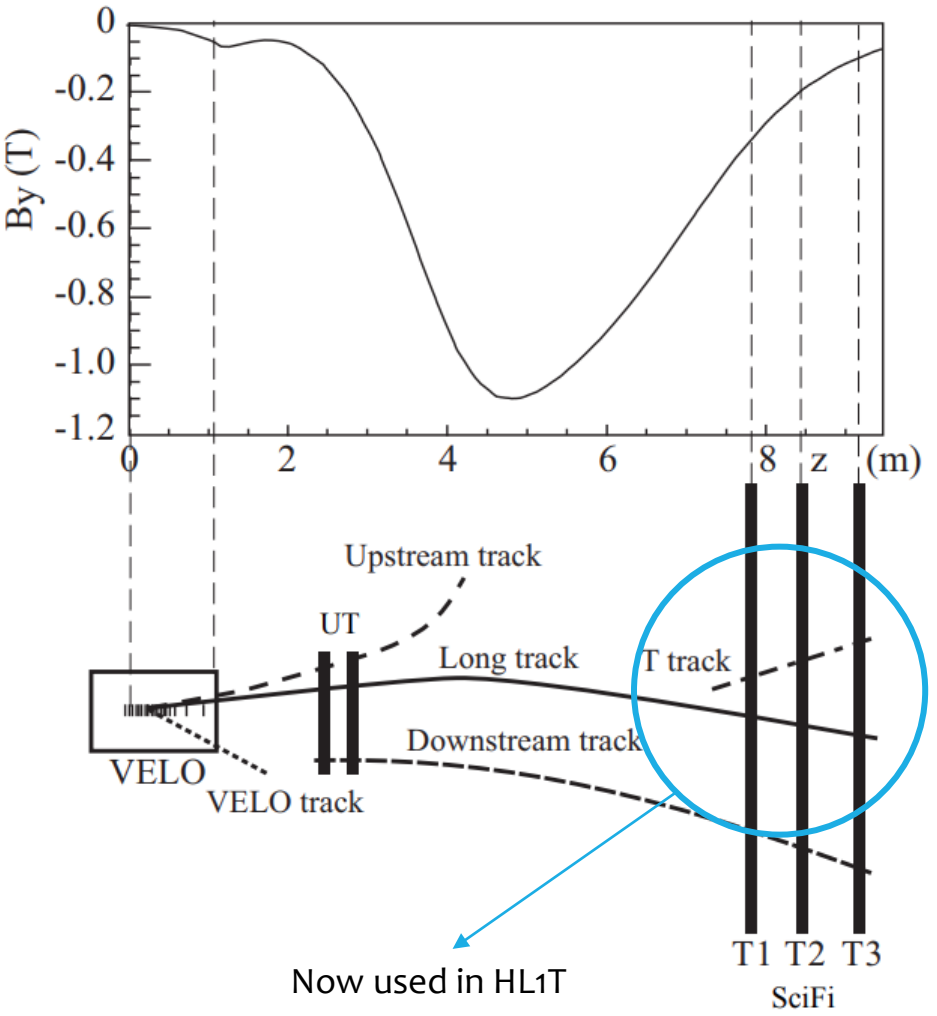




Algorithms: long-lived particles (LHCb)

- In LHCb tracks are reconstructed from segments in the different 3 tracking subdetector:
  - Long: signal in VELO and SciFi (minimum) + UT (full)
  - Downstream: signal in UT and SciFi
  - T: hits only in SciFi
- It is possible to reconstruct particles with a decay flight distance > 2.6 m using only T tracks
  - Up to now, there have been no analyses with particles reconstructed only from T tracks
- By using tracks made exclusively from hits downstream of magnet, particles decaying up to 7.6 m from interaction point can be reconstructed, corresponding to lifetimes ~few ns
- This enables new types of physics analyses, including electromagnetic dipole moment measurements, and searches for BSM LLPs

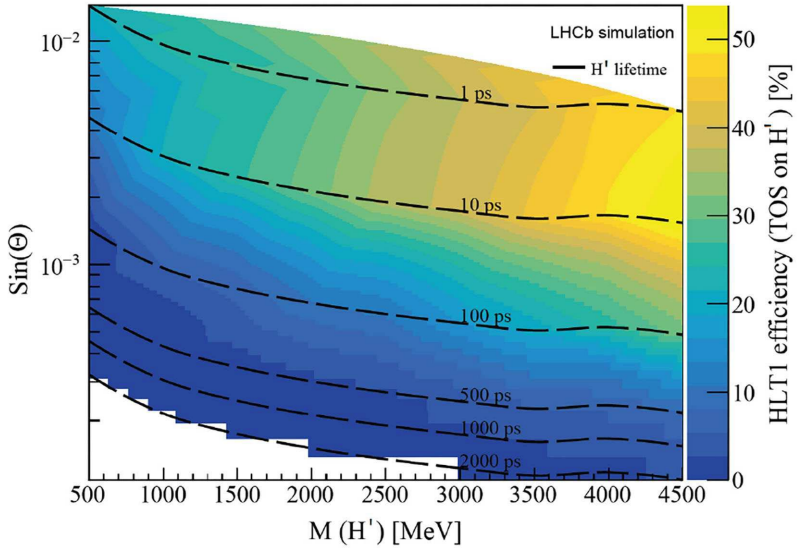
Source Int.J.Mod.Phys.A 30 (2015) 07

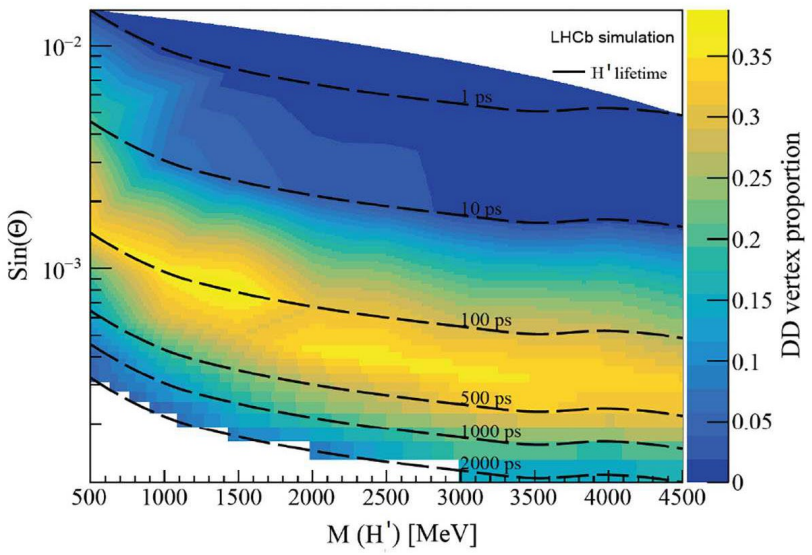


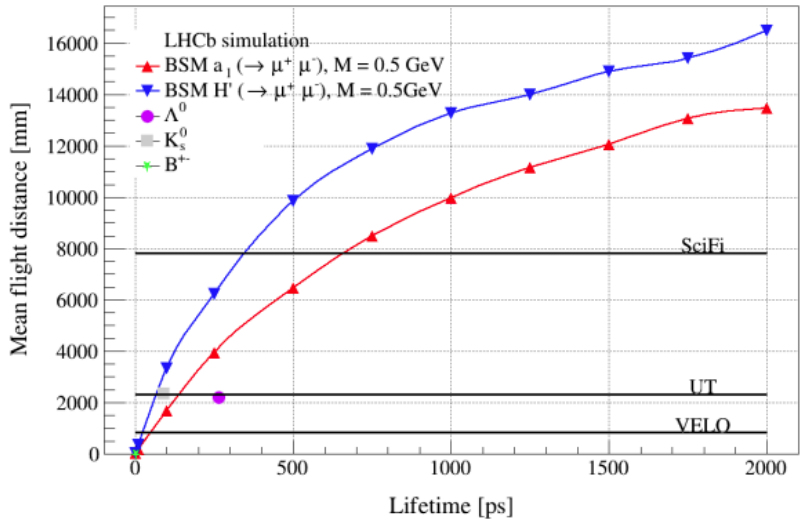
fdata.2022.1008737

Algorithms: long-lived particles @HLT1 (LHCb)

- Current HLT1 uses only long tracks:
  - Decent efficiency (30-50 %) for low lifetime; Poor efficiency (< 10 %) for $\tau > 100$ ps
  - Loss in sensitivity for small H' mass
- Use downstream and T tracks instead @HLT1
  - Extrapolate SciFi seeds to UT
    - Take the output of SciFi seeding
    - Filter out the used seeds
    - Extrapolate to UT stations





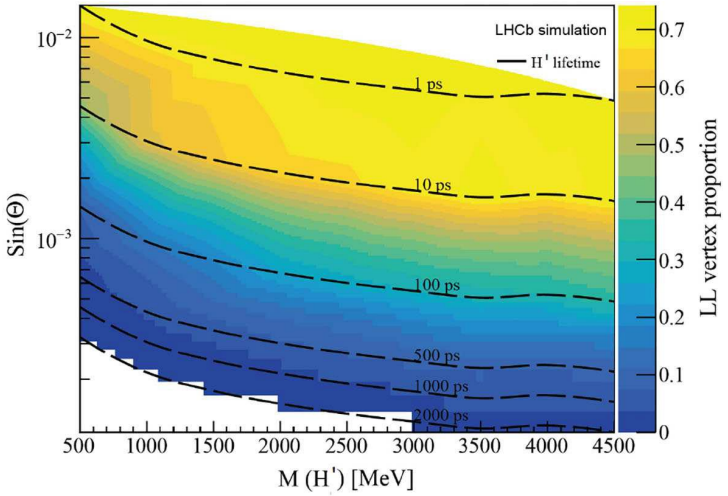


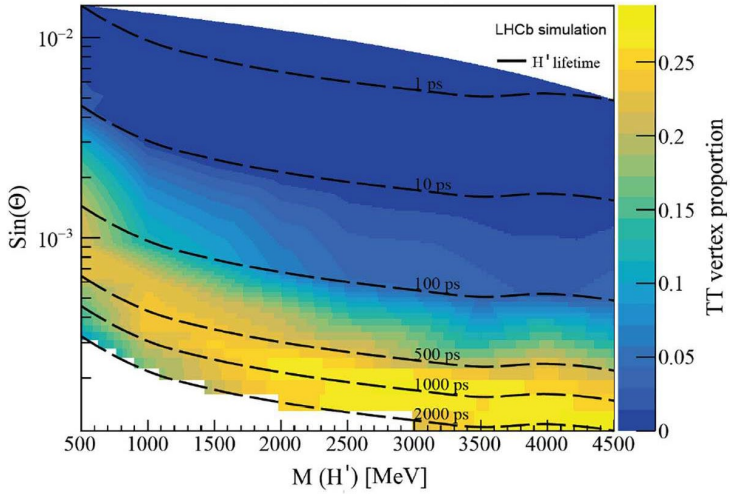
Downstream at HLT1 level : Challenges

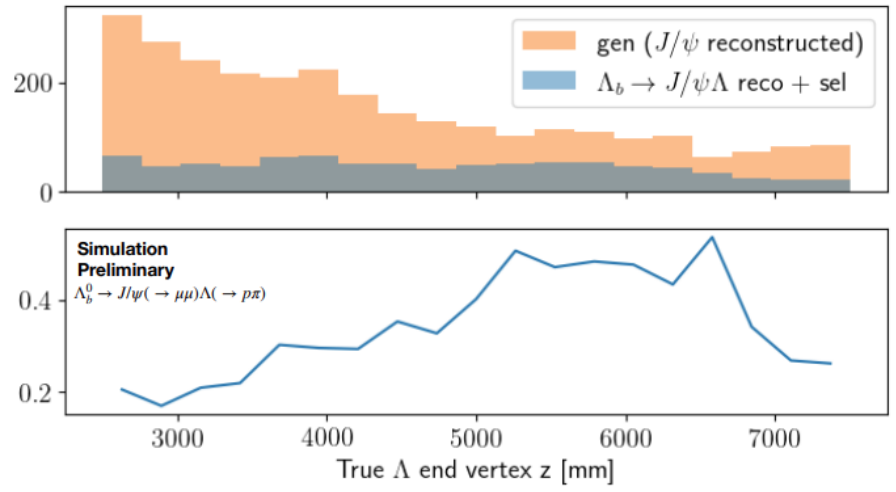
- Main challenges are:
  - Purity of Inputs to the algorithm (Input SciFi seeds)
    - ~28% ghosts in unmatched SciFi seeds
  - Algorithm throughput:
    - Must fit in the HLT1 budget
  - Reconstruction efficiency and ghost rate:
    - The number of layers and type of trackers in the UT defines its capabilities.

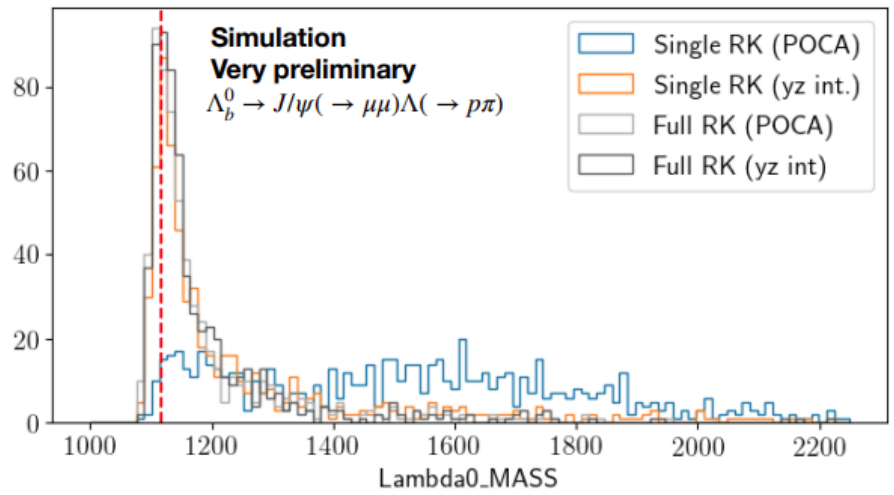
Algorithms: long-lived particles @HLT2 (LHCb)

- In HLT2 LHCb's Particle Vertex Fitter algorithm not suitable for T-tracks due to inhomogeneous field.
  - Runge Kutta (RK) extrapolation to fit correct vertex position
- Algorithm has been adapted to perform one RK extrapolation in first iteration, linear in subsequent
  - Mass resolution is still poor, there is still a significant z-bias
- Room for improvement to better resolve mass and z-position



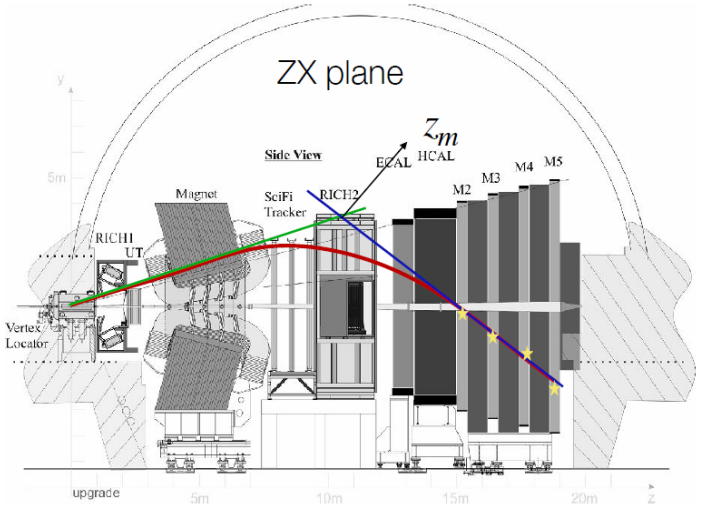


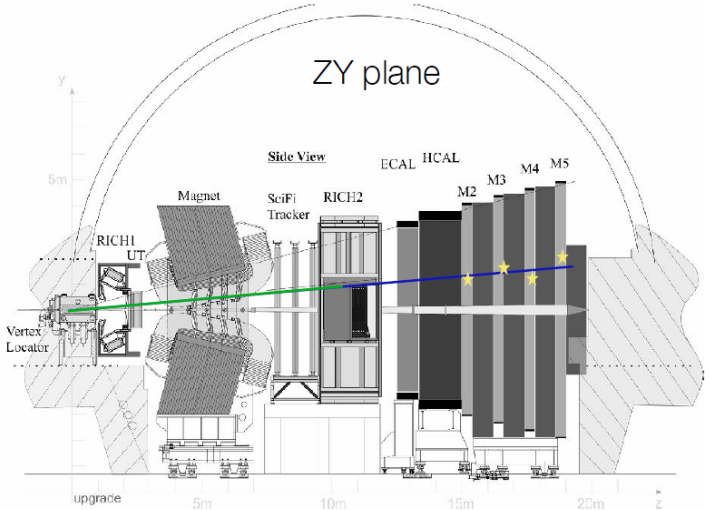




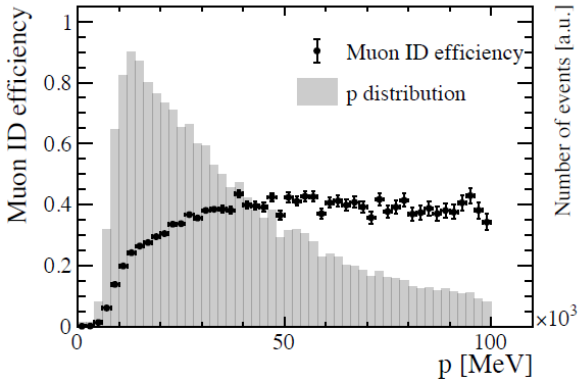
Algorithms: muon reconstruction @HLT1 (LHCb)

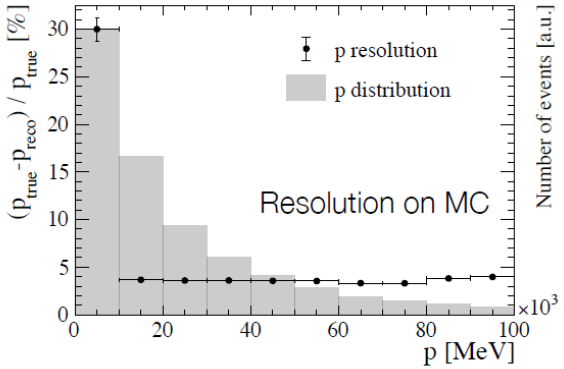
- VELO-MUON tracking
  - Standalone muon tracking
    - No seeding from other tracking detectors)
  - VELO-Muon matching
    - VELO tracks + Standalone MUON tracks
    - Momentum and charge: B field parametrisation
  - Secondary vertex fitting
    - Use the VELO segment to obtain very precise secondary vertices
  - 4% momentum resolution
  - Default muon identification
- Usability:
  - Tag-probe track efficiency studies
  - High multiplicity events
  - Strange physics: lower momentum threshold needed
  - Commissioning: validate MUON time alignment





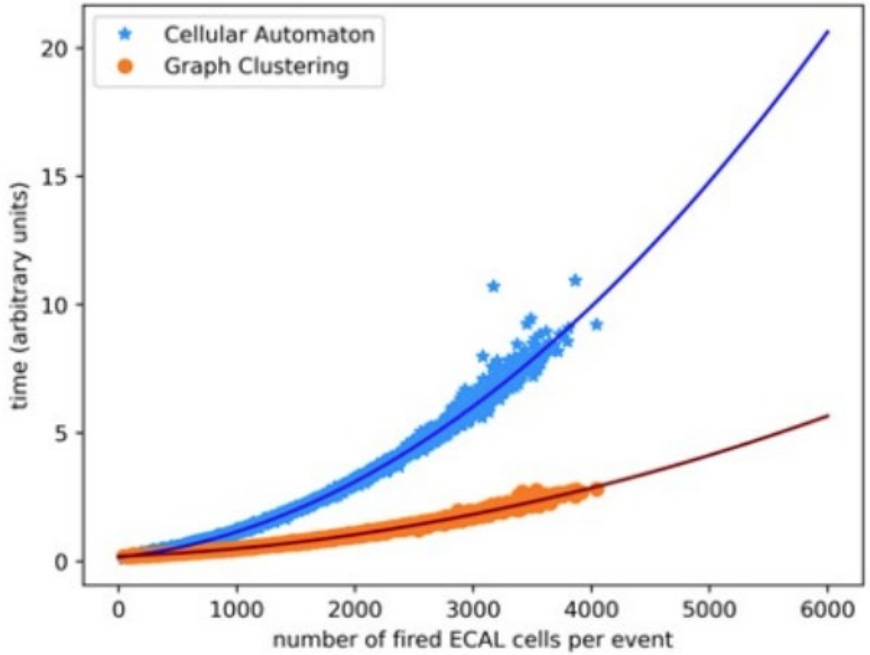
$$\frac{q}{p} = \frac{1}{|\int d\vec{l} \times \vec{B}|} \left(\frac{t_{x,\mathrm{MUON}}}{\sqrt{1 + t_{x,\mathrm{MUON}}^2 + t_{y,\mathrm{MUON}}^2}} - \frac{t_{x,\mathrm{VELO}}}{\sqrt{1 + t_{x,\mathrm{VELO}}^2 + t_{y,\mathrm{VELO}}^2}} \right)$$



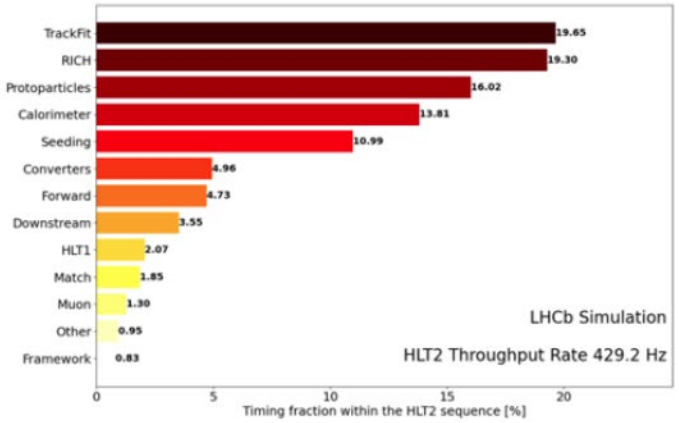


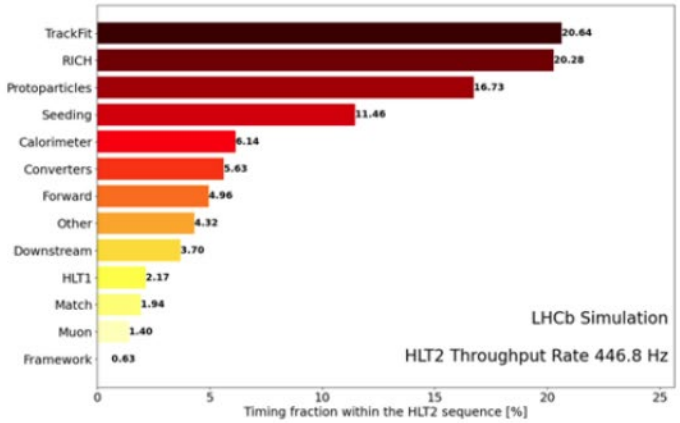
Algorithms: clustering algorithms @HLT2

- Algorithms also needed to adapt to new conditions, simply increasing the complexity of the model (ML, DL) may not fulfil all specifications (time, area…)
- Using graphs as data structure to store digis, with an insertion under certain rules, digits from the same cluster are already grouped together.
- Overlap cases are contained into independent connected components of the graph



- `Moore_hlt2_reco_baseline` →change of -0.71%, calorimeter change from 20.74% to 18.98%.
- `Moore_hlt2_fastest_reco` → change of **4.09%**, calorimeter change from **13.81%** to **6.14%**.
- `Moore_hlt2_pp_thor` → change of **3.21%**, calorimeter change from **10.59%** to **4.71%**.

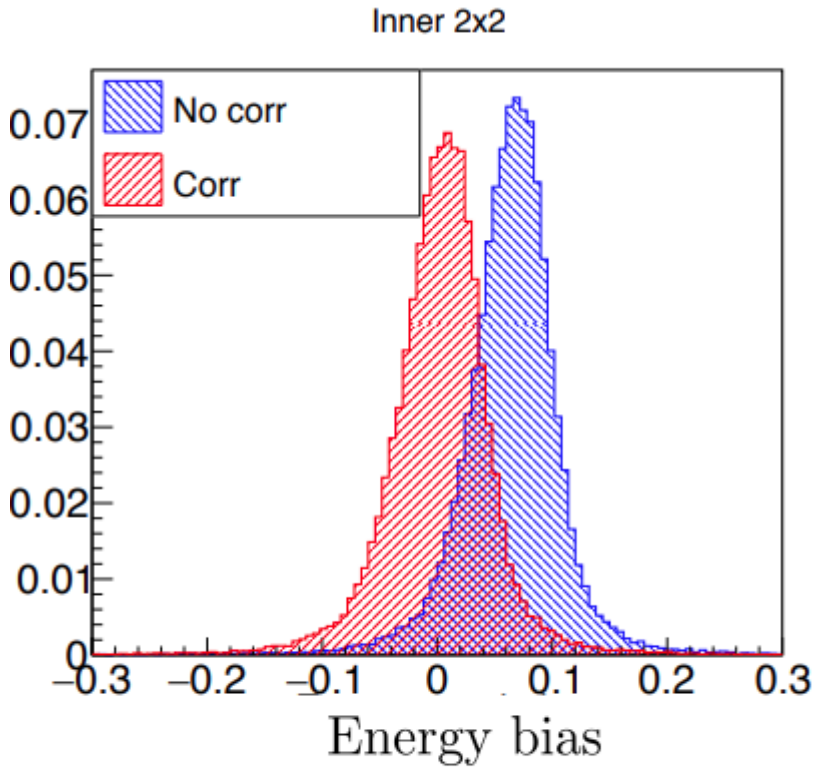




P. García

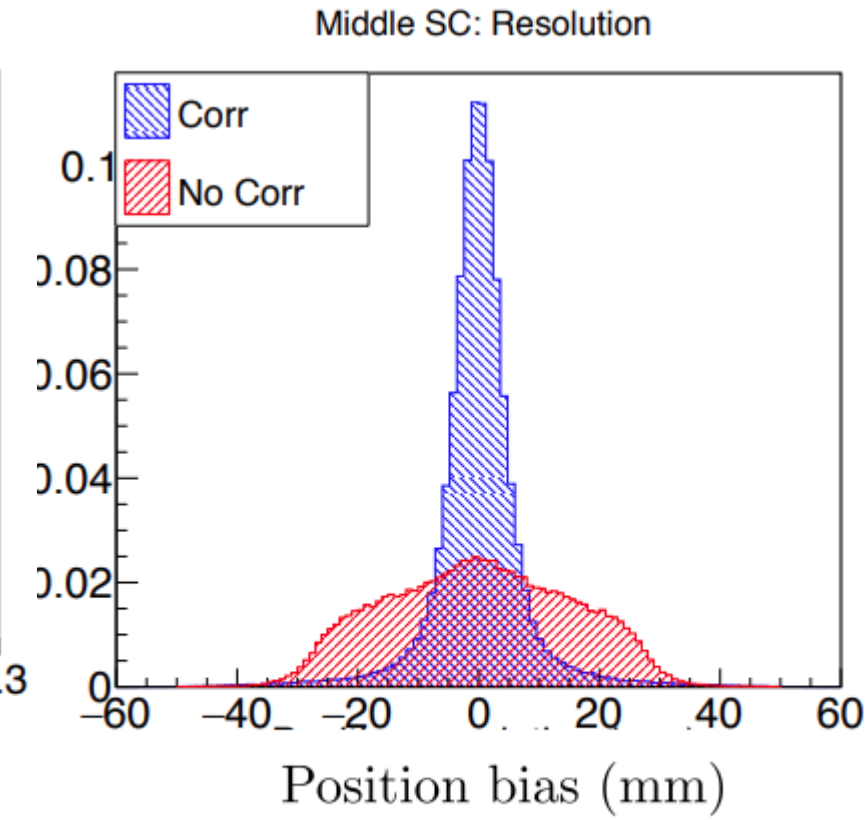
Algorithms: calorimeter calibrations for energy and position (LHCb)

- With the increased occupancy of Run 3, the effect of pile-up and external particles hitting the detector cell will be also increase. This will lead to a loss in resolution for both position and energy measurements.
- Using smaller cluster shapes (still larger than the Molière's radius), such as 2x2 or SwissCross, can reduce those effects resulting in better measurement resolution.



No corr: bias = -6.1%

Corr: bias = 0.02%

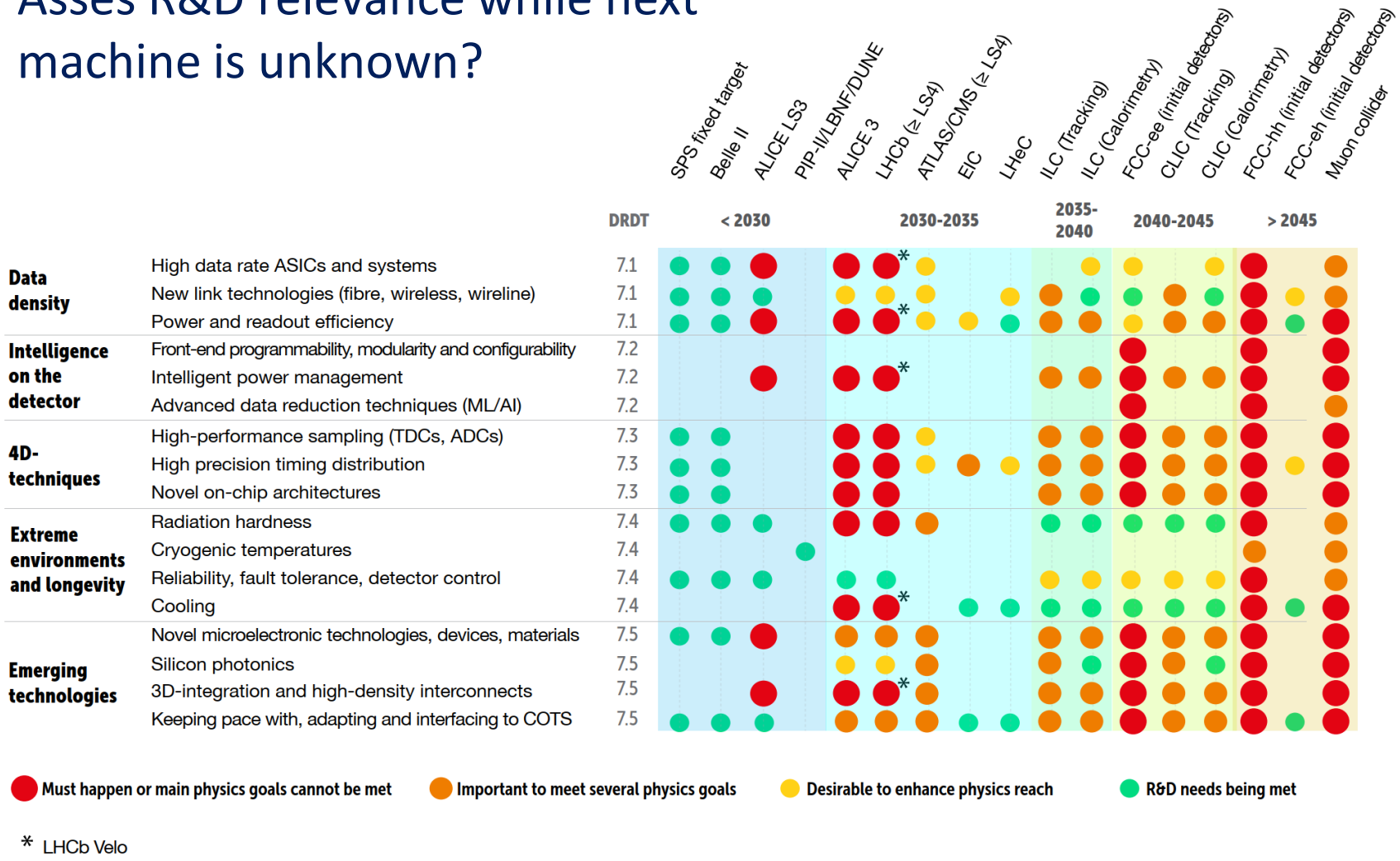


resolution = 3.598 ± 0.019 mm

A word about DRD7

- An R&D Collaboration on Electronics and On-detector Processing
- Spanish-wide DRD meeting in Barcelona ➔ A role for CPAN
- Is ultra-high radiation hardness relevant? Or on the other hand, low material detectors?
- Conservative/simple frontends vs. adding sophisticated technological complexity on-detector?
- Is there some interest from the Spanish community in some of these topics? YES

Asses R&D relevance while next machine is unknown?



Outcome

- Great interest of the Spanish community on these topics, long-standing expertise in high-speed electronics and accelerator processors. Well stablished, internationally recognized community.
- Start a collaboration across different institutions. There is a significant interest from our community to work on high-speed electronics, heterogenous computing architectures and trigger.
- Elaborate a document summarizing the interest of the different Spanish groups and express interest of our participation in the DDR7 collaboration: https://www.overleaf.com/read/kpczrjtwyzbn

If you wish to contribute to it, just let me know.

3rd COMCHA School in Oviedo

https://indico.uniovi.es/e/comchaschool

- Registration open as of today
- Encourage your students to register soon!
  - Only 35 places available