Triggering strategies and heterogeneous computing

Marianna Fontana (INFN Bologna)
HSF workshop, 27 November 2025

INFN BOLOGNA

LHCb RTA

LHCb

Introduction

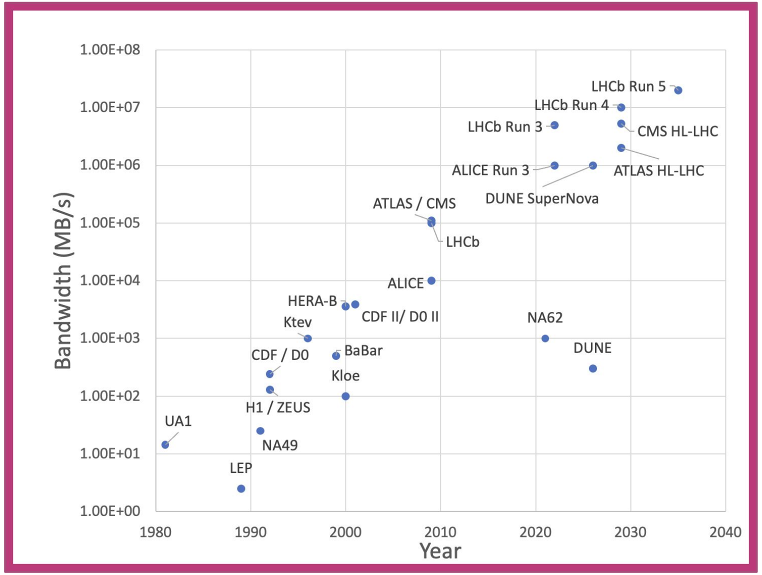
- The reconstruction of raw detector data and its processing in real time represents a **major challenge in HEP**
- **Demands for higher throughputs in upcoming years**

Two demands for Trigger:

- Decrease throughput to backend DAQ
- Keep trigger efficiency high

Two trends:

- Triggerless/continuous readout
- Higher-level reconstruction in hardware trigger



Courtesy of A. Cerri

Disclaimer

This presentation is in no way an exhaustive view of all LHC experiments trigger systems, rather a selection of some topic biased from my view and background

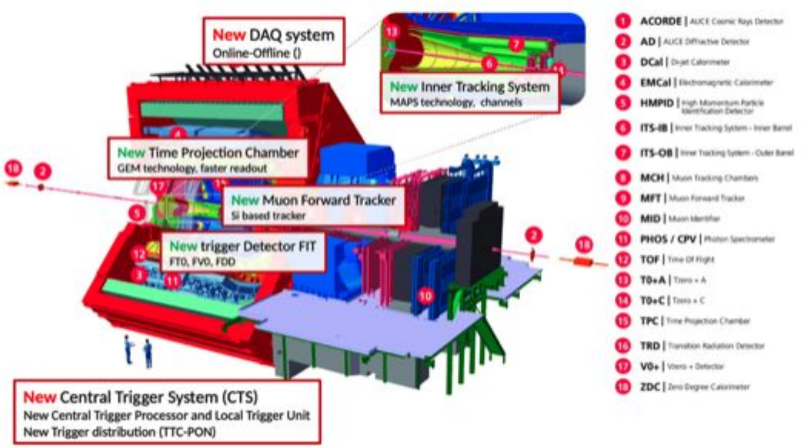
ALICE

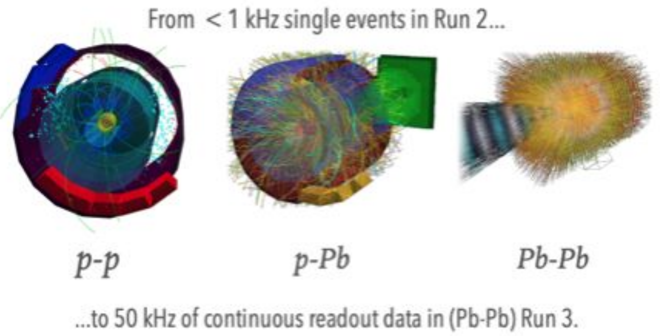
(material taken from CHEP24 talks)

ALICE in Run 3

Challenges for Run 3:

- Completely new detector readout and substantial detector upgrades: new ITS, MFT, FIT. New GEM for TPC readout
- ~100 x more data than Run 2
- Many important physics signals have very small signal-to-noise ratio
- Triggering (selection) techniques very inefficient if not impossible
- Needs large statistics
    - **Read the data resulting from all interactions**

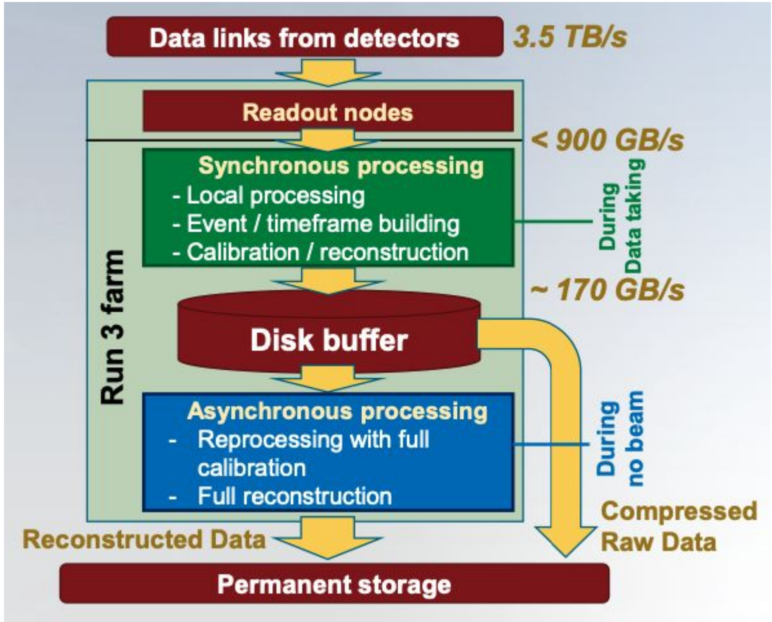




Run 3: Online and offline processing

D. Rorh @CHEP '24

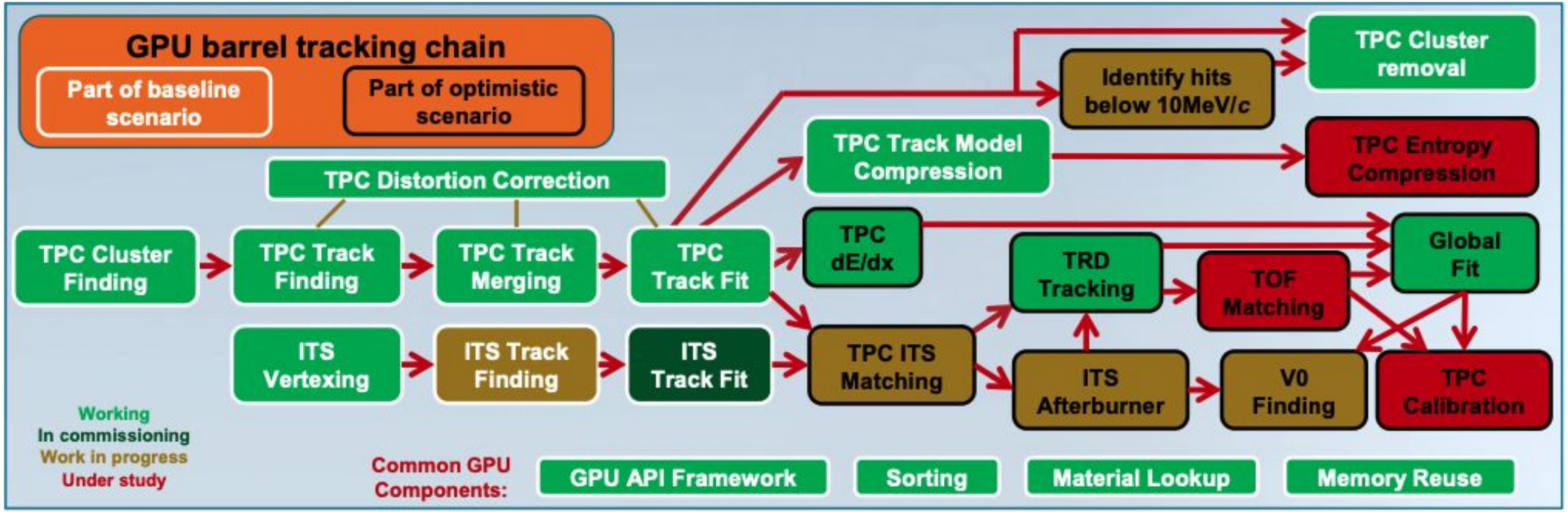
- Use **GPUs to speed up online (and offline)** processing
- Reconstruction is two-stepped
  - Synchronous phase (beam circulating):
    - online processing on GPUs
    - calibration and data compression stored to disk buffer
  - Asynchronous phase (no beam):
    - full processing of data staged in the disk buffer on online farm.
    - optionally use GPUs when possible
- TPC track reconstruction is the most time consuming during synchronous reconstruction and is therefore performed on GPUs (the most cost effective solution)
- Try to offload more algorithms to GPU for better GPU usage in offline



D. Rorh @CHEP '24

Experience running with GPUs

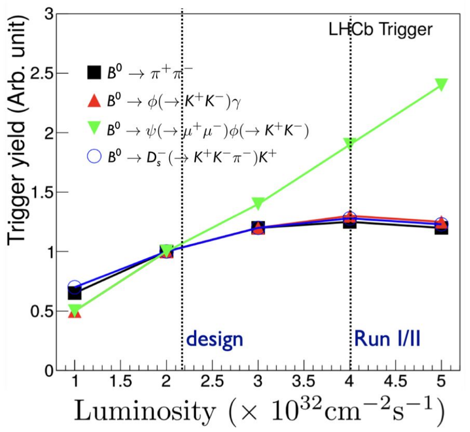
- Running the GPU-enabled online workflow successfully for pp and Pb-Pb from 2022 to 2024
  - During 2023 Pb-Pb had 17% free GPU resources at highest interaction rate
- Vendor-independent GPU usage via generic common C++ Code
  - Can run on CUDA, OpenCL, HIP, and CPU (with pure C++, OpenMP, or OpenCL)
- Planning to run full barrel tracking on GPU in optimistic scenario, to raise fraction on GPU from 60% to 80%, aiming for 5x speedup



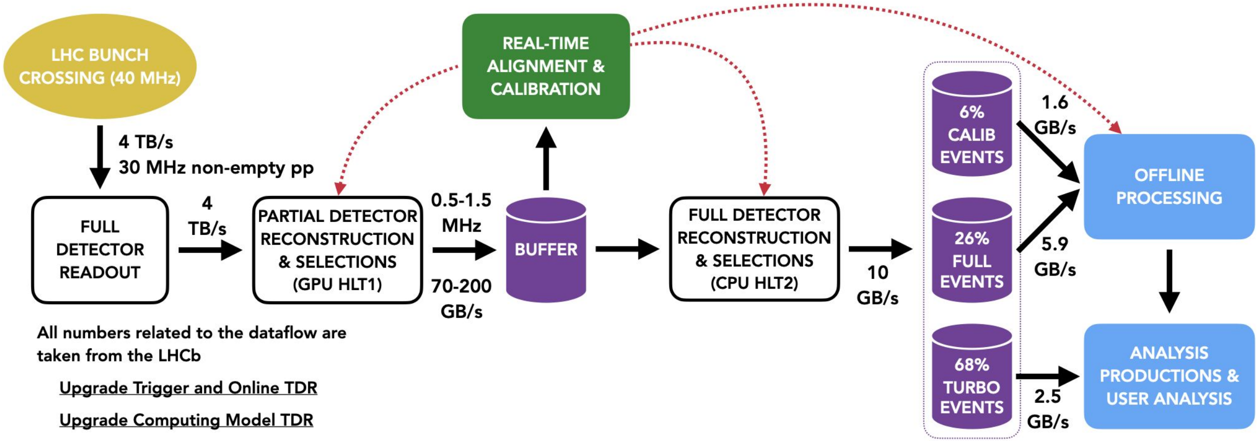
LHCb

The Run 3 data flow

J. Phys.: Conf. Ser. **878** 012012



LHCb-FIGURE-2020-016



- In Run 1-2 couldn't efficiently trigger on heavy flavour using hardware signatures
- Trigger for many hadronic channels saturated
- Solution: **fully software trigger**

- Detector data @30 MHz received by O(500) FPGAs
- 2-stage software trigger, HLT1 (GPU) & HLT2 (CPU)
- Real-time alignment & calibration
- After HLT2, 10 GB/s of data for offline processing

HLT1 trigger

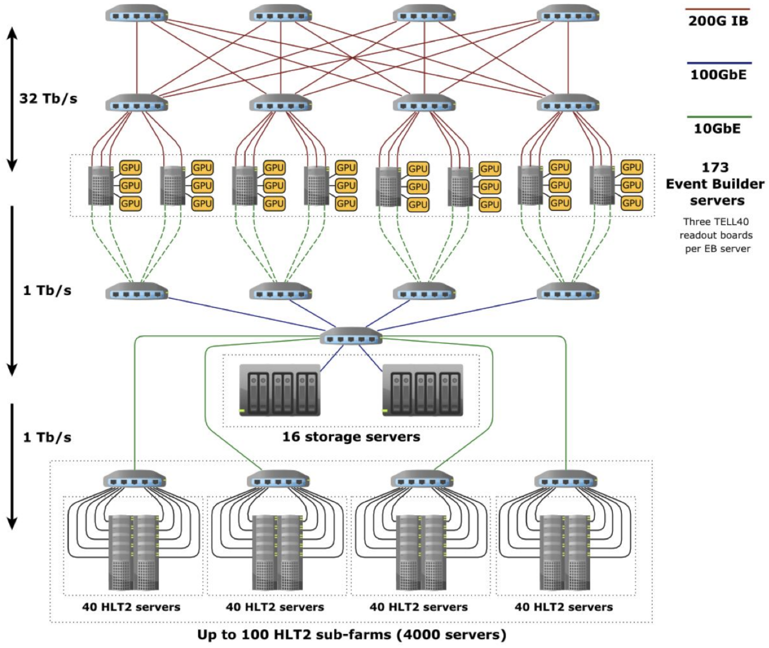
- Take as input LHCb raw data (**4 TB/s**) at 30 MHz
- Perform partial event reconstruction & coarse selection to cover the full breadth of LHCb physics
- Reduce the input rate by a factor of 30 (~1 MHz)
- **~ 500 GPUs NVIDIA RTX A5000 GPUs** installed
  - The baseline TDR design could be achieved with 300 GPUs
  - Extra GPU power used to extend the improvements beyond-TDR

The GPU choice matches the DAQ architecture of LHCb

- GPUs can be hosted by the Event Builder Nodes via PCIe slots
- reduced costs due to shared powering and cooling and smaller network

HLT1 tasks are suited for parallelisation:

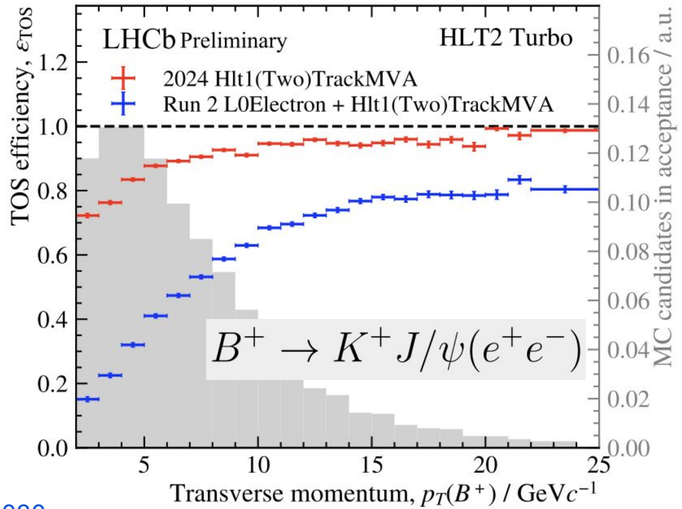
- Events can be treated independently
- Objects of reconstruction (tracks, vertices, ...) are independent



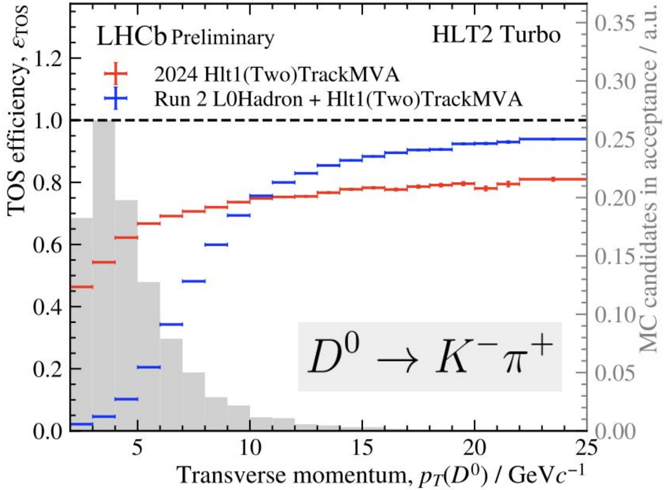
Comput.Softw.Big Sci. 6 (2022) 1, 1

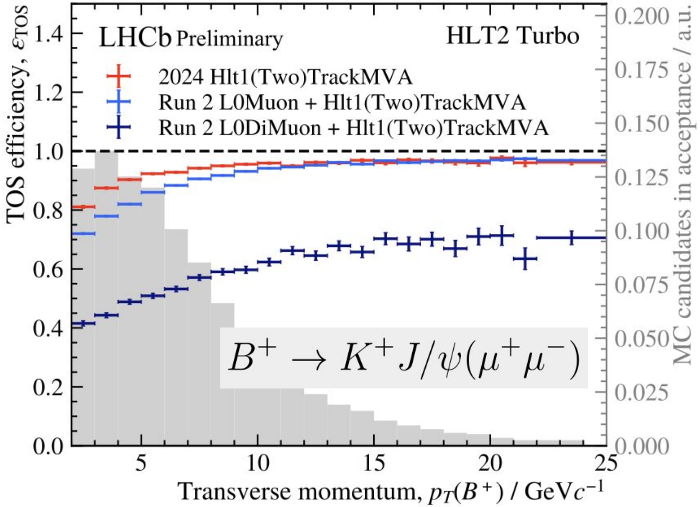
HLT1 performance

- **The real-time analysis philosophy proved to be valid**
- Significant improvements in trigger efficiencies
- Huge gain a low-pT
  - Beneficial for the charm and strange physics programme
- Large impact for electron channels
- Muon channels gained from the removal of the global event cuts



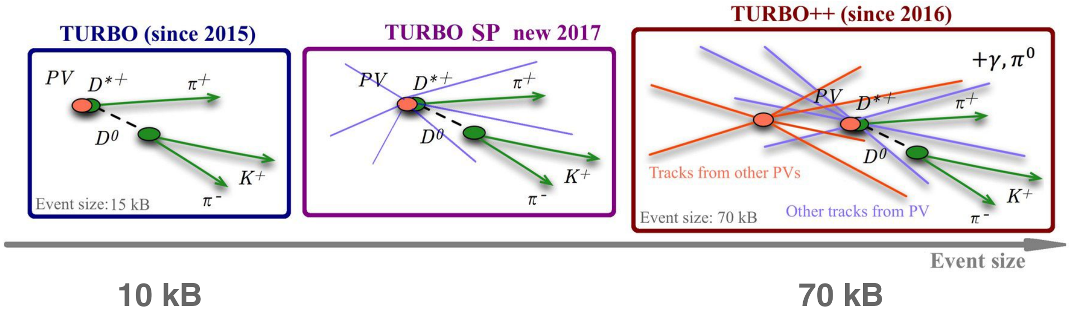
LHCb-FIGURE-2024-030

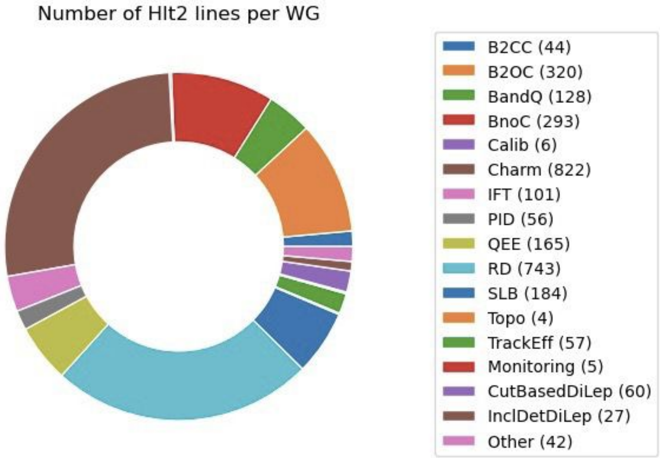




HLT2 trigger

- HLT2 runs a full reconstruction and all the necessary selections (inclusive but mostly exclusive) for the wide LHCb physics programme (~3000 lines)
- Given the hard limit on bandwidth (10 GB/s to tape and 3.5 GB/s on disk) and expected signal rate, event size is the only free parameter
- Need to "persist" all the reconstructed objects for offline analysis
- The **successful strategy of the Turbo paradigm used at full speed also in Run 3**

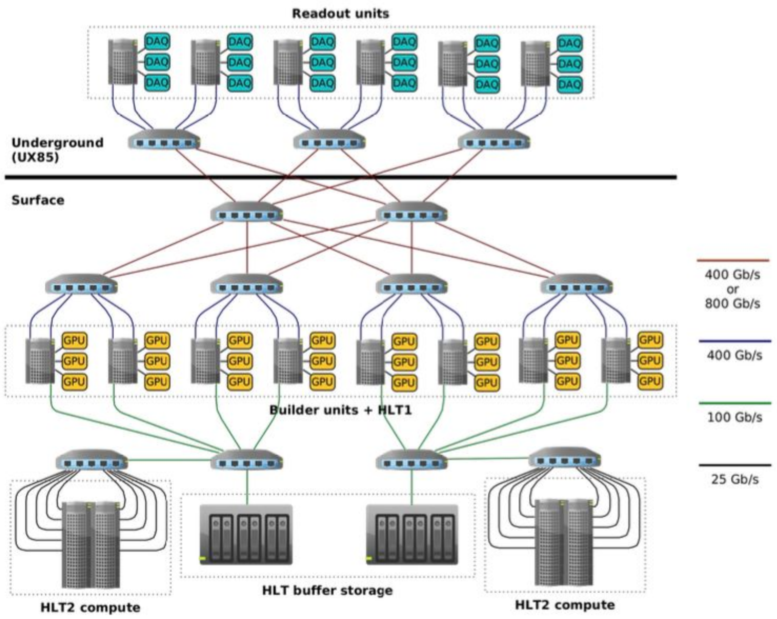




LHCb-FIGURE-2024-034

The trigger evolution: Run 5

- Triggerless design philosophy will remain correct and scalable
- Exciting challenges in trigger and DAQ
  - 200 TB/s of data, to be processed in real time and reduced by ~4 orders of magnitude before sending to permanent storage
  - data processing will be based around pile-up suppression
- Partial and full detector reconstruction (and selections?) both on GPUs
- Complementary R&D activities focusing on two main areas
  - Building subdetector primitives, for example tracks or calorimeter clusters, on FPGAs [LHCb-PUB-2024-001]
  - Exploiting other architectures such as the IPUs



ATLAS

(material prepared by C. Antel)

ATLAS Run 3 Evolved SW Triggering Strategies

Fully CPU

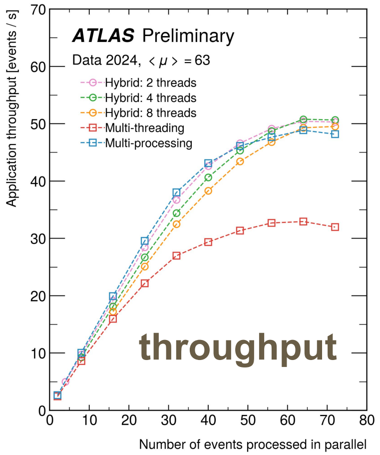
HLT operates 60 000 real CPU cores, having replaced all Run 2 cores @ 22.8 HS06/core performance with AMD EPYC 7302 CPUs @ 36.2 HS06/core performance by 2023:

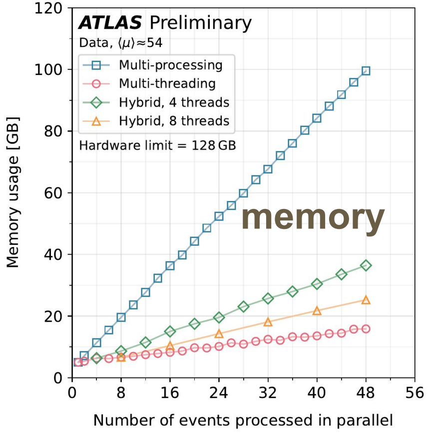
year	2018	2022	2023
HS06	1.2M	1.7M[1]	2.0M

CHEP 2023

Adopted & optimised multithreading + multiprocessing:

- Trade off between memory (multithread better) and throughput (multi-process better).
  - 2022: HLT - Fully multiprocess, Grid sim/reco - fully multithreaded.
  - 2024 HLT: hybrid (multi-process + 4 threads/process)

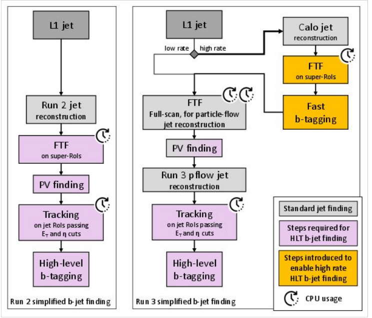




TriggerCoreSWPublicResults twiki

Implemented particle flow reconstruction:

- Better agreement with offline, lower pile-up rate.
- Required full scan tracking @ 8 kHz for jets and MET triggers.
- Resulting CPU limits required redesign of hadronic trigger scheme:
  - First stage selection for early, fast event rejection:
    - Jets/MET: Calo-based jet/MET reco.
    - B-jets: RoI fast tracking + dedicated ML.



New hadronic triggering scheme enabled particle flow based triggers for
HH -> bbbb/bbττ
(a key physics driver for HL-LHC)

JINST 18 (2023) 11, P11006

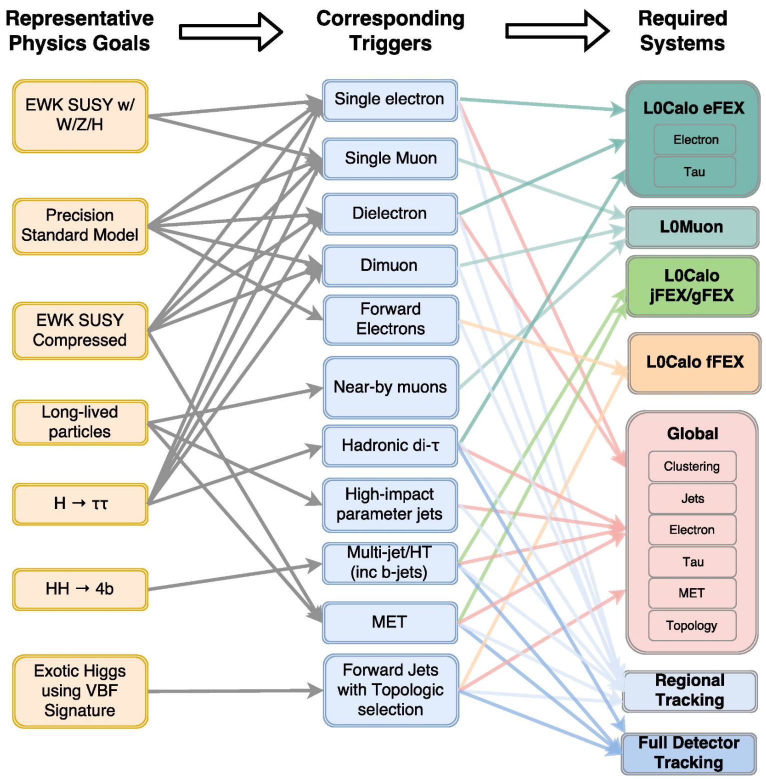
ATLAS Phase II TDAQ Overview

Key Physics drivers: SM precision measurements, rare SM processes (di-Higgs coupling, H -> μμ), beyond SM, forward physics tagging.

Detector upgrades: New inner tracker out to |eta| 4 (currenty: |eta| 2.5), new forward High Granularity Timing Detector, upgraded muon detectors to reduce fake rate.

TDAQ specs: 4.6 TB/s DAQ (5 MB/evt) for 1 MHz full detector readout (Run 3: 100 kHz) to Event Filter (EF) farm, 10 kHz final recording rate (Run 3: 3 kHz).

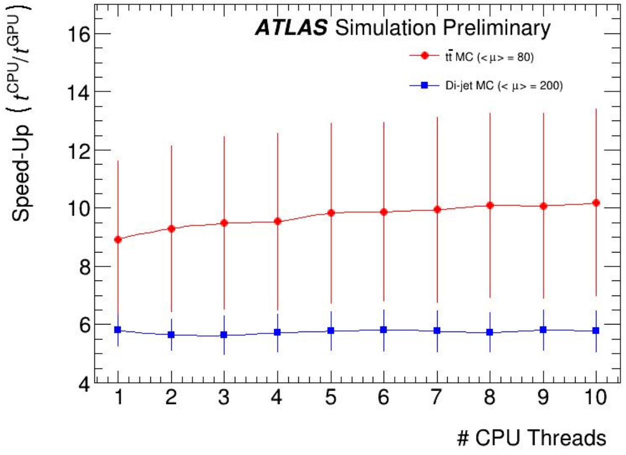
Software challenges: Maintain good event filter performance and low pT trigger thresholds in pile-up 200 environment (Run 3: pile-up 60).



ATLAS TDR

ATLAS Phase II Technology Choice

- Currently undergoing CPU/GPU/FPGA technology choice process for Event Filter Farm.
- Primarily driven by tracking needs: 1 MHz regional tracking, including ~150 kHz full scan tracking.
- GPUs/FPGA could be key to:
  - Affording target physics thresholds
  - Dealing with higher than expected hadronic rates (large uncertainty)
  - Simplifying triggering scheme (need for extra filtering steps)
- Technology choice expected mid 2025.
- Examples of technology choice criteria: Minimal tracking performance, cost, power consumption, maintainability, flexibility, trendy.
- R&D highlights (that are public):
  - GPU: towards GNN Track reco, calo clustering
  - FPGA: towards GNN track reco, towards full pipeline, high throughput flavour tagging (deep sets NN)

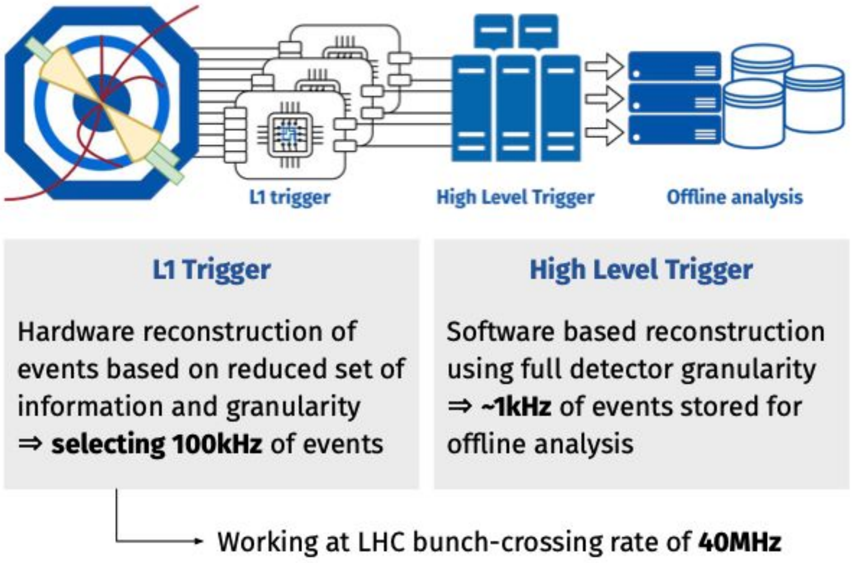


EF Calo clustering CPU/GPU speed ups

CMS

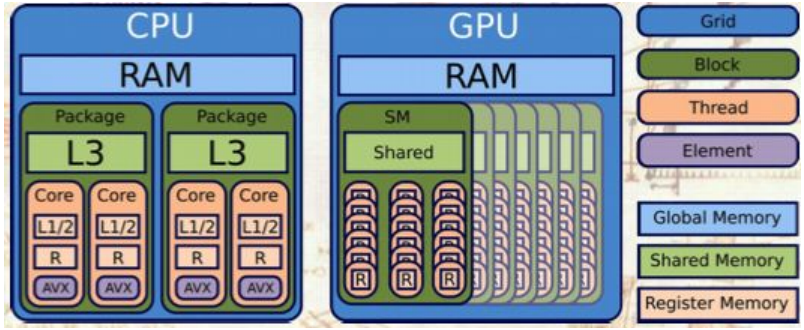
(material taken from CHEP24 talks)

The Run 3 path

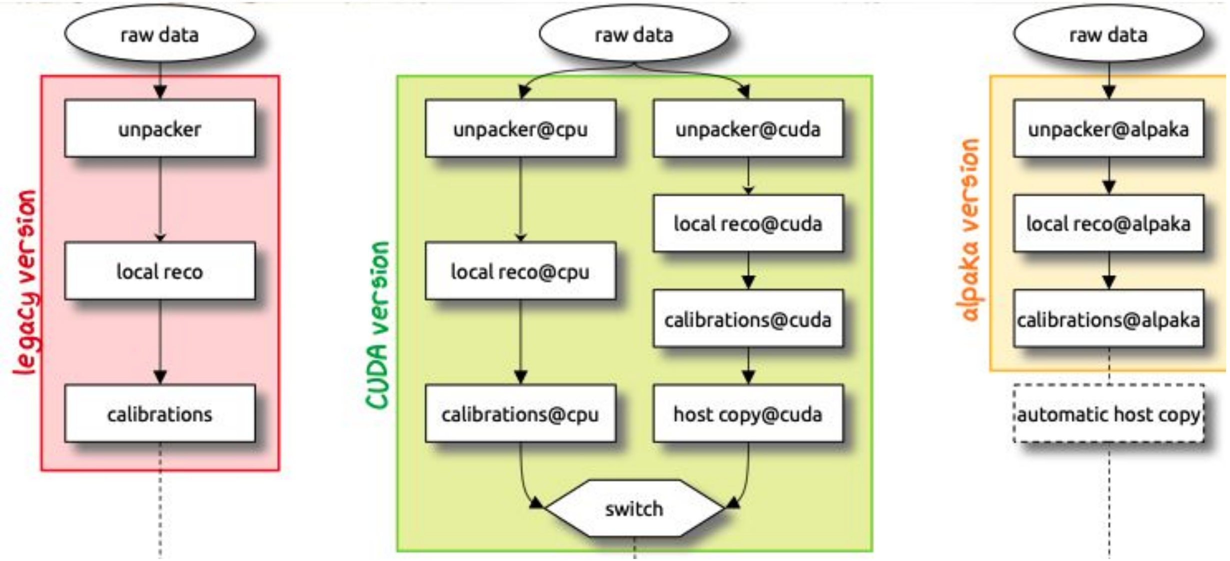


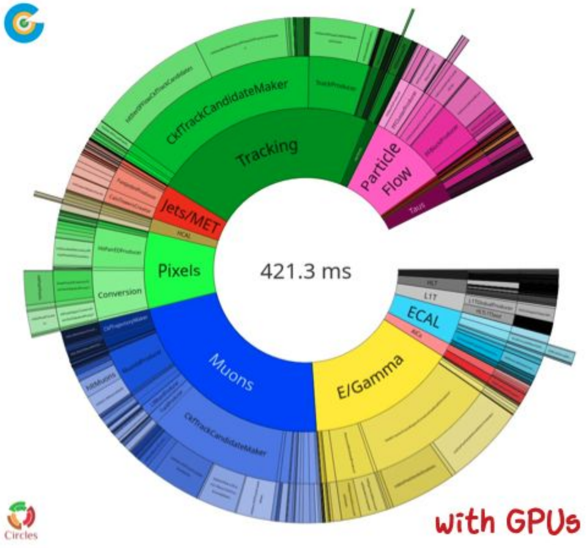
Software based trigger with full event information available running on CPU + GPU based farm

- Run 3 HLT farm composed of **200 nodes:** each node equipped with two AMD Milan 64-core CPUs and two NVIDIA Tesla T4 GPUs
  - +20% extention in 2024 with 18 nodes:
  - 2 × AMD EPYC “Bergamo” 9754 processors
  - 3 × NVIDIA L4 GPUs
- Increasing usage of GPUs at Run 3
  - Offloading **30% of the HLT reconstruction to GPU**
- GPU reconstruction implemented and fully commissioned
  - The **execution time per event** was **reduced by ~40%**
- HLT throughput requirement ~500Hz:
  - Throughput increases by a factor of ~1.80
  - Power Consumption (per throughput) reduced by ~30%



Improved heterogeneous framework





Use of Alpaka to provide performance portability across accelerators through the abstraction of the underlying levels of parallelism

Uniform algorithms and data structures

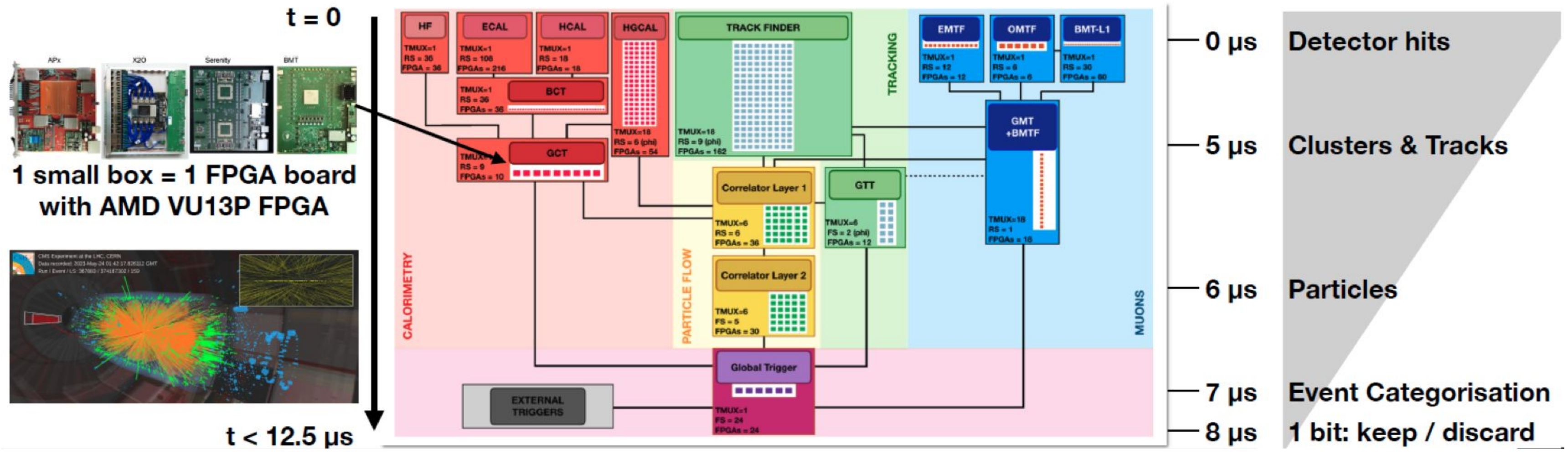
- framework can automatically schedule tasks on the CPU or on the GPUs
- framework can automatically schedule copies (to and) from the GPUs

Thanks to the use of GPUs

- 50% better event processing throughput
- 35% less processing time per event
- 15% - 20% better performance at initial cost
- 15% - 25% better performance per kW

The Phase-2 Trigger Upgrade

- Benefit from the upgrade of the CMS detector: high granularity information and tracking information
- The system allows a throughput of > +64 Tb/s using top-of-the-line FPGAs and ultra-fast optical links (25 Gbps)
  - Adapt and evolve as needs of experiment change
  - Increased bandwidth to 750 kHz at increased latency of < 12.5 μs
- Incorporate sophisticated algorithms and advanced techniques to extend CMS physics acceptance
- Design philosophy: Custom ATCA-boards



Testing the ideas in Run 3

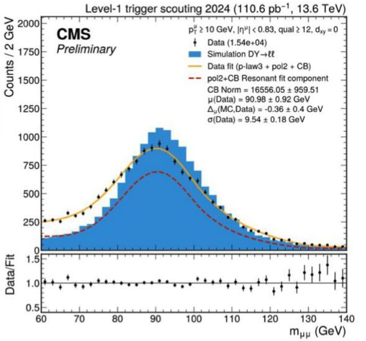
Hardware demonstration ongoing and some tests in Run-3 data taking

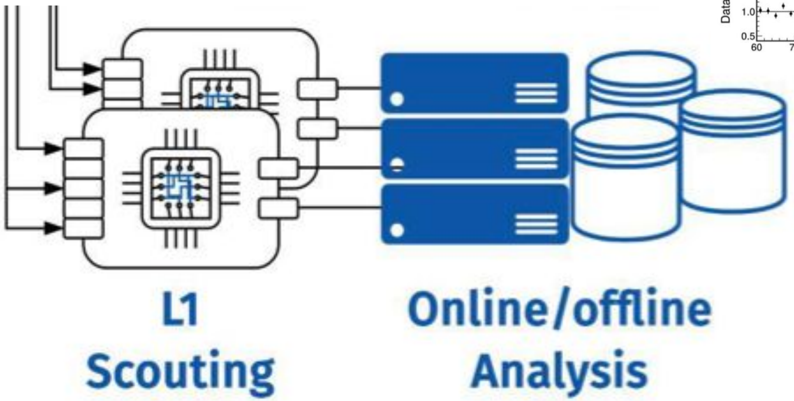
- new algorithms, optimisation techniques, hardware inspired from the phase-2 upgrade project
- LLPs triggers: displaced muons, muon showers, delayed jets…
- 40 MHz scouting (real-time data analysis)
- Inclusion of the first anomaly detection trigger on live data: AXOL1TL and CICADA

Triggerless analysis

- Storing and analysing events at L1 or HLT (x100 smaller event size)
- Crucial for very low-mass bump-hunt searches, compressed spectra or b-physics

L1 scouting: standard L1 rejects 99.75% events. L1 scouting will allow us to have a look at those events





- The development of performant software will be vital for the future of HEP to address the demand for higher throughput in the coming years
- This requires a change of culture in the community, to consider the software projects as milestones when building new experiments

Thank you