Performance of the LHCb heterogeneous software trigger

Alessandro Scarabotto
On behalf of the LHCb collaboration

21st October 2024

CHEP 2024

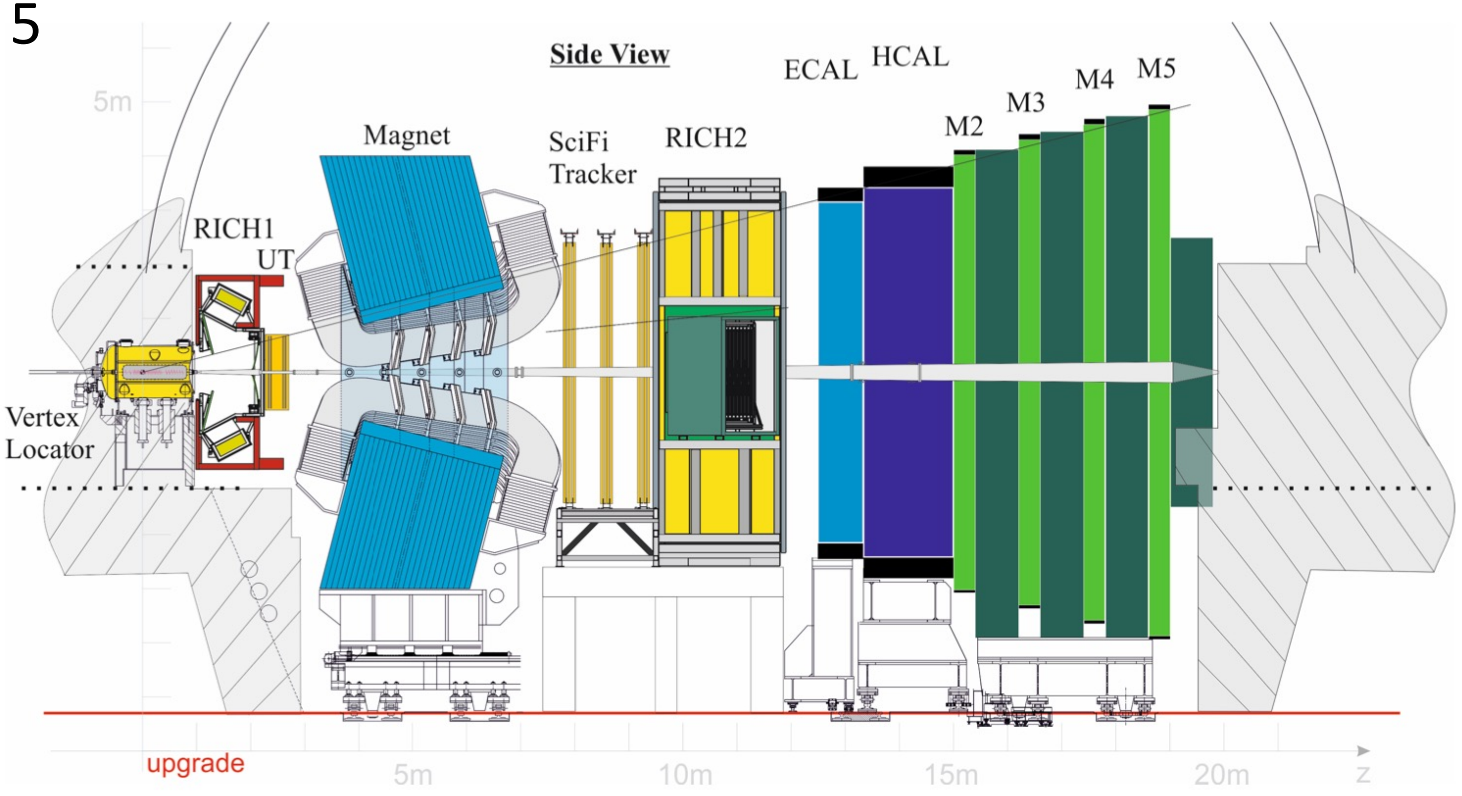
Krakow, Poland

The LHCb Run3 upgrade

- The LHCb detector was upgraded with the aim of collect data with 5 times higher luminosity:

$$\mathcal{L} = 2 \cdot 10^{33}\, cm^{-2} s^{-1}$$

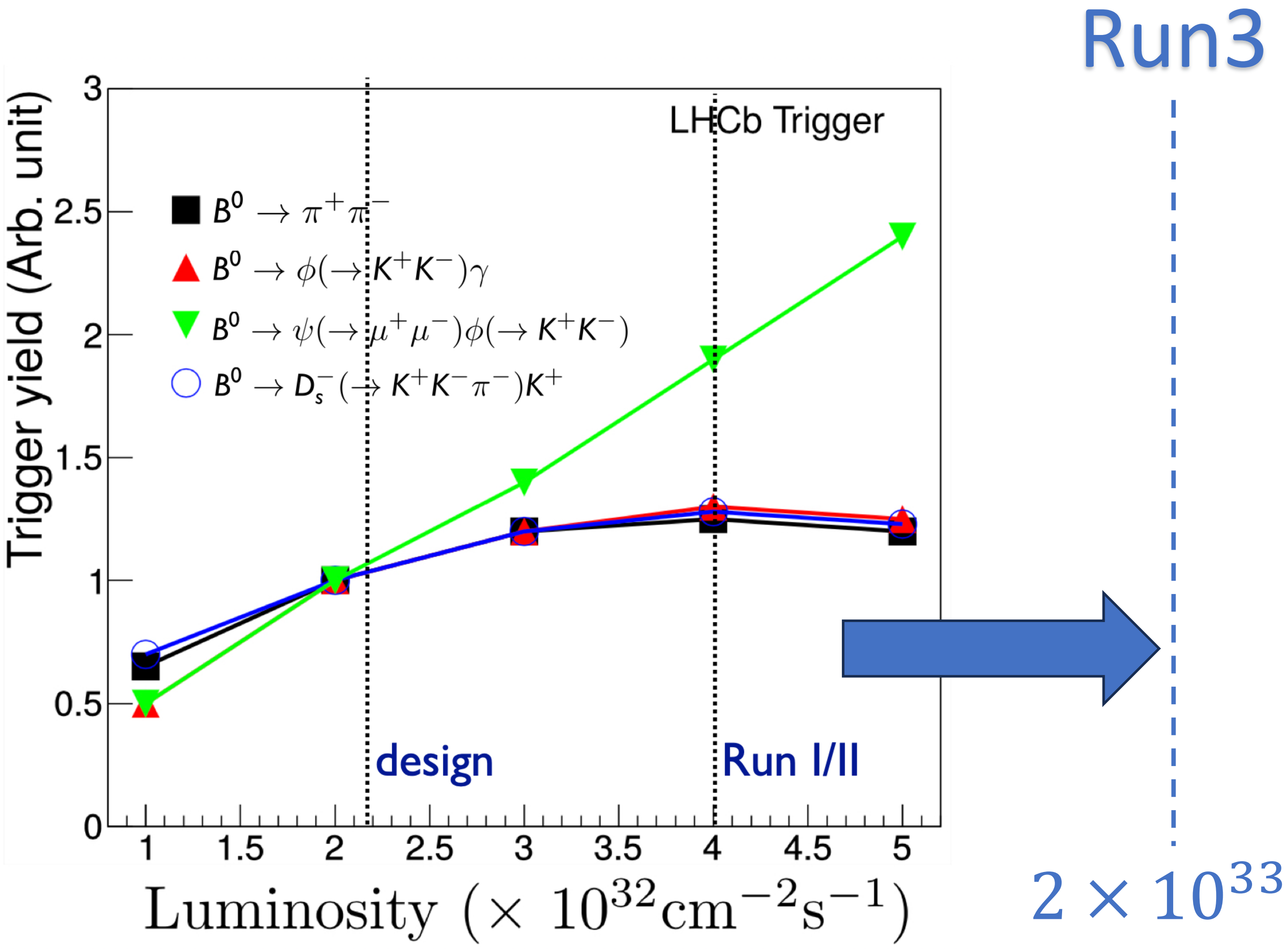
- Increasing pile-up $\langle \mu \rangle \sim 5$ (average pp collisions per bunch crossing)
- Full replacement of tracking detectors needed to deal with higher occupancy environment
- What about triggering?



LHCb-TDR-12

The LHCb Run3 trigger

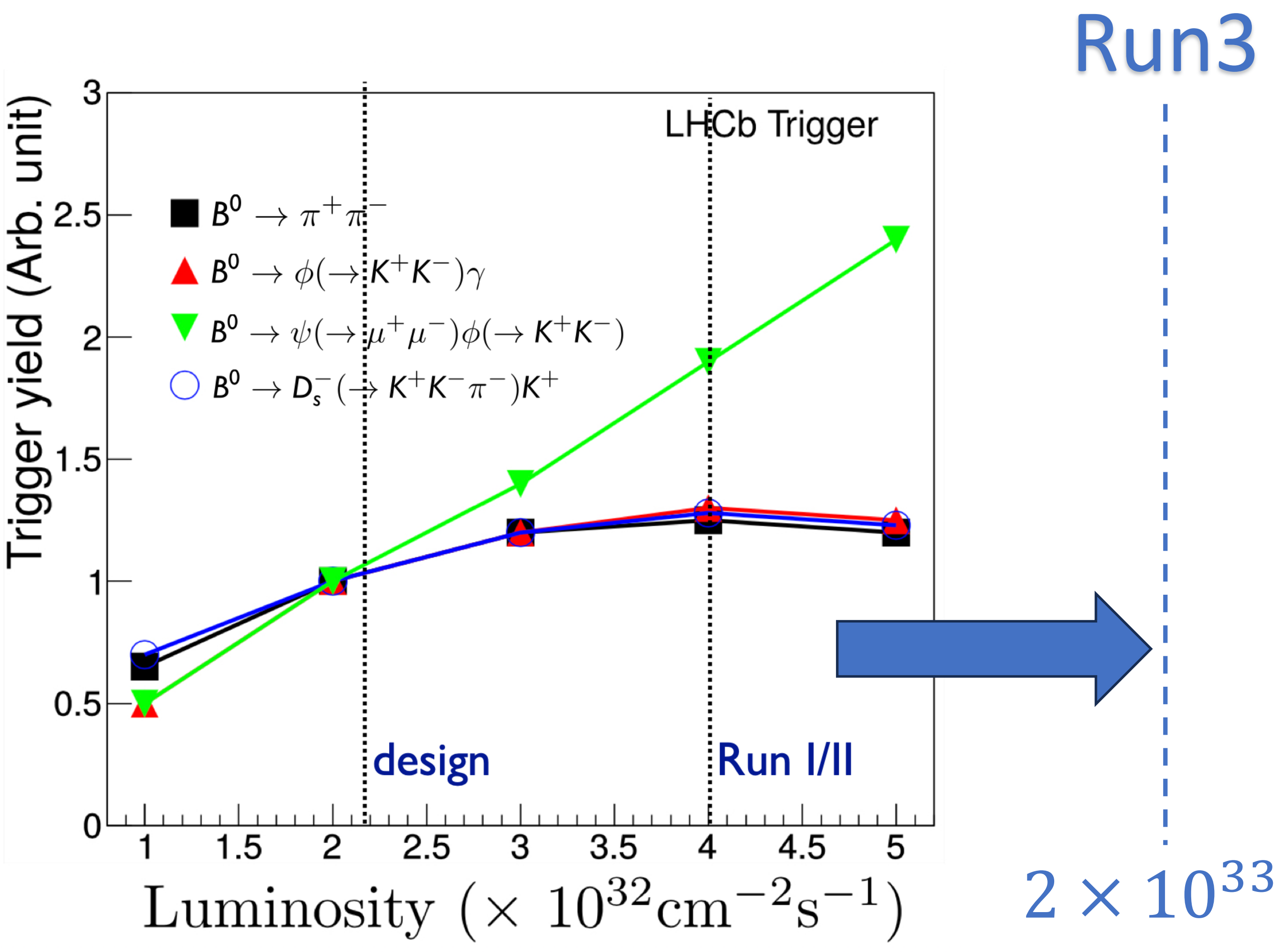
- Limitation of Run2 trigger is the first-level hardware stage (L0)
- Saturation of trigger yields by increasing luminosity
- Caused by tight momentum/energy requirements at L0 selections



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

The LHCb Run3 trigger

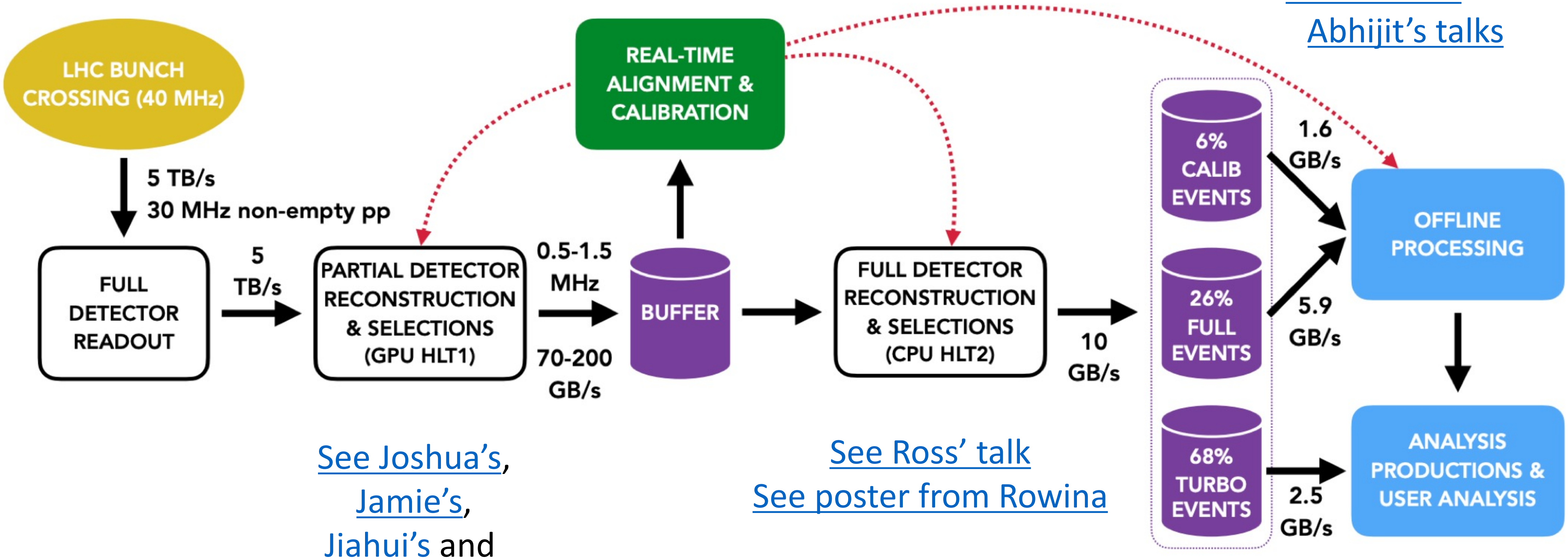
- Limitation of Run2 trigger is the first-level hardware stage (L0)
- Saturation of trigger yields by increasing luminosity
- Caused by tight momentum/energy requirements at L0 selections
- The Run3 LHCb:
  - Removal of L0! But …
  - Reconstruction at 30 MHz LHC pp collision rate for the High Level Trigger (HLT)



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

The Run3 LHCb dataflow

See Nicole's and
Abhijit's talks



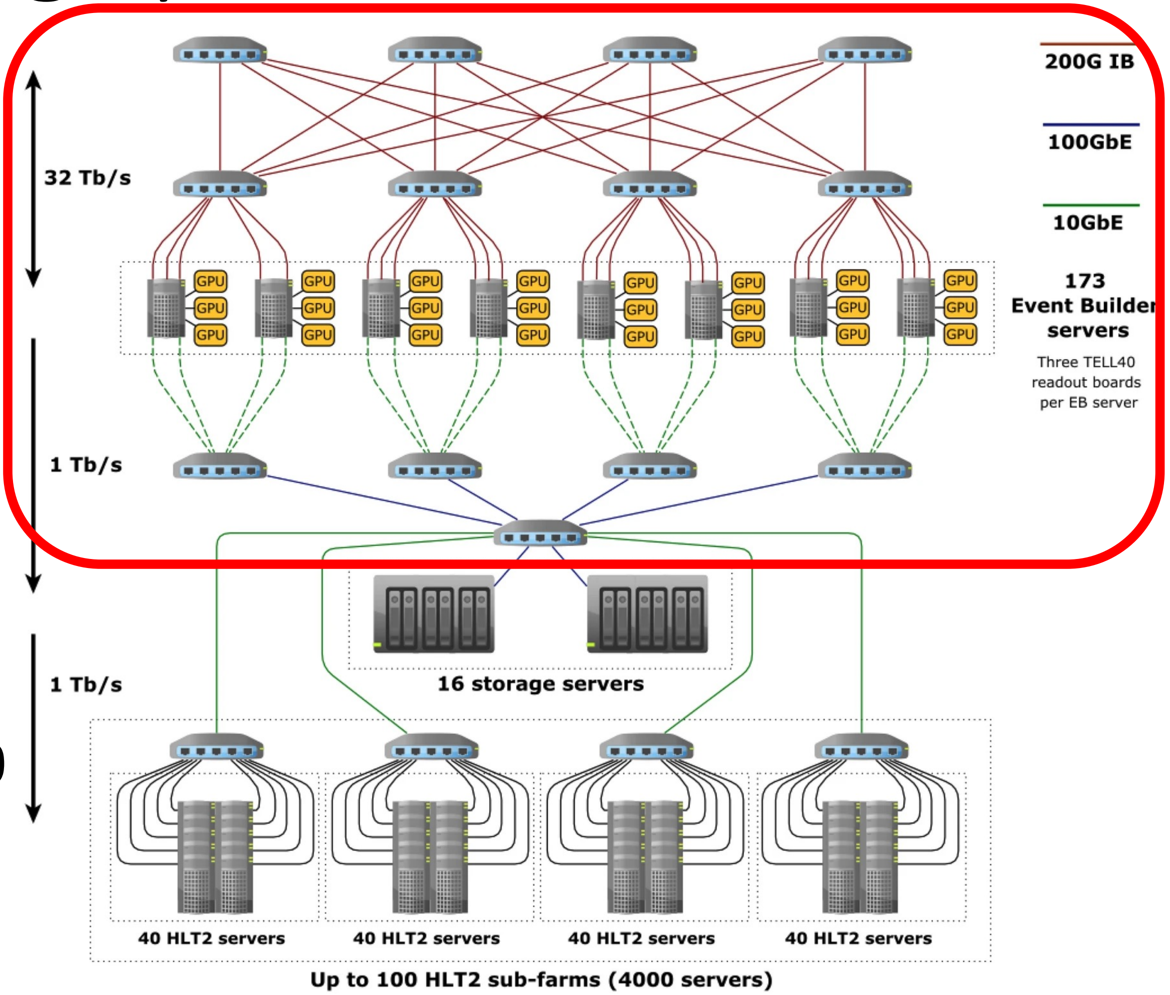
See Joshua's,
Jamie's,
Jiahui's and
Valerii's talks

See Ross' talk
See poster from Rowina

LHCB-FIGURE-2020-016

Heterogeneous computing system

- Raw detector info is sent to the data processing center
- FPGA cards receive data at average 5 TB/s
- 163 Event Builder (EB) servers produce the packets of events
- Each EB server has 3 PCIe slots in which a GPU is installed and where **HLT1** is run (zero overhead costs)
- 3 GPUs per EB server → ~ 500 Nvidia A5000



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

First-level trigger on GPUs: HLT1

- Event reconstruction respecting the tight throughput constraints (30 MHz)

	LHCb	ATLAS	CMS	ALICE
$\mathcal{L}\ [cm^{-2}s^{-1}]$	2×10^{33}	2×10^{34}	2×10^{34}	6×10^{27}
pile-up	5	60	60	1
reconstruction rate	30 MHz	100 kHz	100 kHz	50 kHz
reconstructed tracks/s	1800 M	90 M	90 M	10 M

- Exploiting the parallelization power of GPUs for reconstruction algorithms

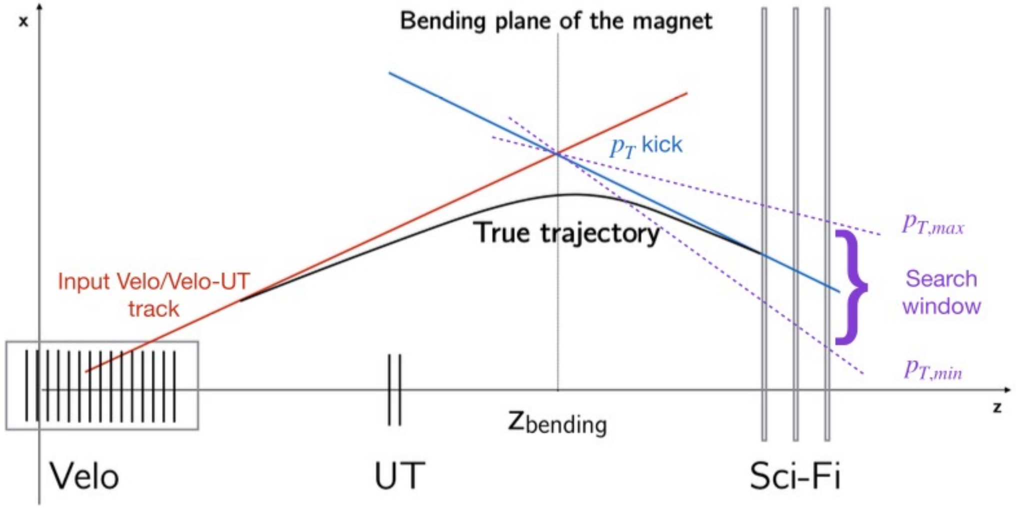
First-level trigger on GPUs: HLT1

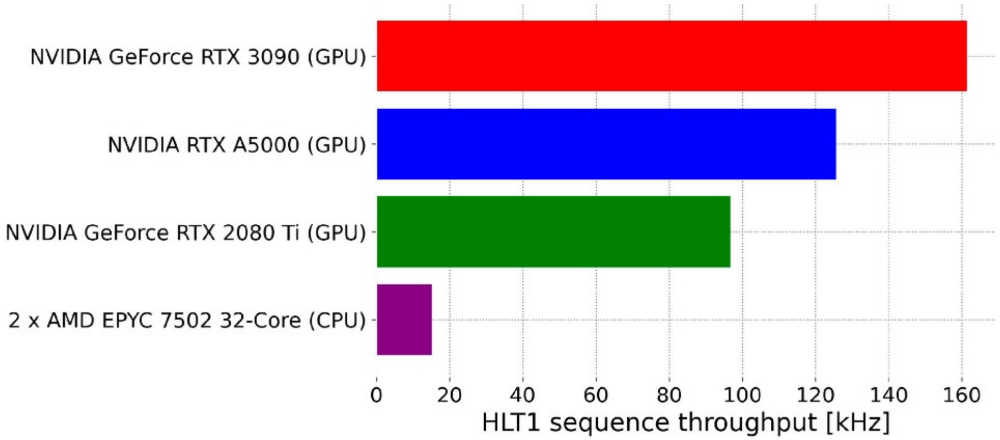
- Event reconstruction respecting the tight throughput constraints (30 MHz)

	LHCb	ATLAS	CMS	ALICE
$\mathcal{L}\ [cm^{-2}s^{-1}]$	2×10^{33}	2×10^{34}	2×10^{34}	6×10^{27}
pile-up	5	60	60	1
reconstruction rate	30 MHz	100 kHz	100 kHz	50 kHz
reconstructed tracks/s	1800 M	90 M	90 M	10 M

- Exploiting the parallelization power of GPUs for reconstruction algorithms
- Example: long tracks reconstruction algorithms

A. Scarabotto at al. IEEE Access (2024) Vol12, 10.1109



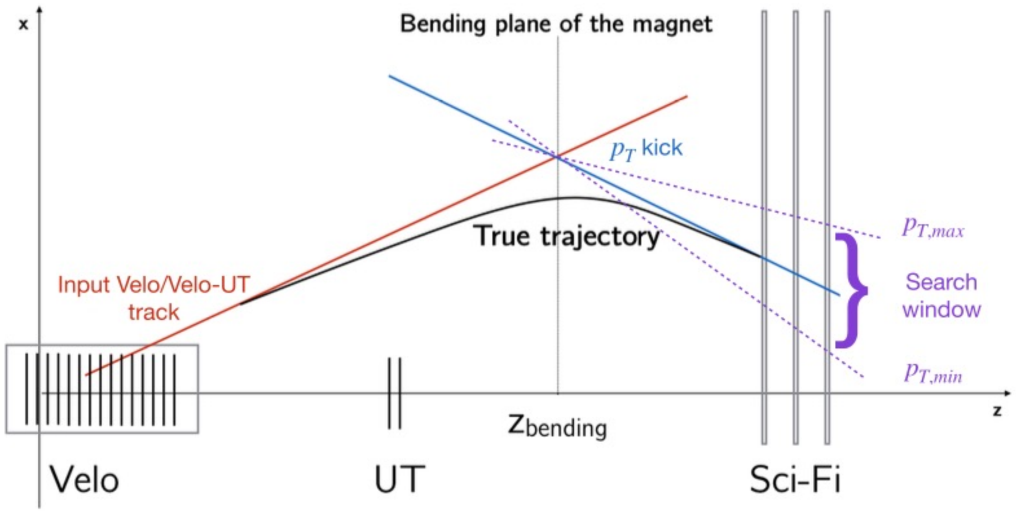


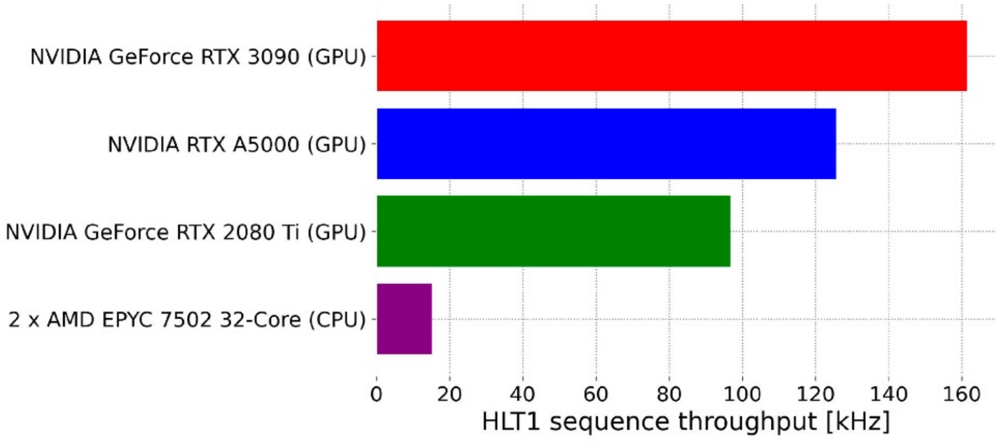
First-level trigger on GPUs: HLT1

- Event reconstruction respecting the tight throughput constraints (30 MHz)

	LHCb	ATLAS	CMS	ALICE
$\mathcal{L}\ [cm^{-2}s^{-1}]$	2×10^{33}	2×10^{34}	2×10^{34}	6×10^{27}
pile-up	5	60	60	1
reconstruction rate	30 MHz	100 kHz	100 kHz	50 kHz
reconstructed tracks/s	1800 M	90 M	90 M	10 M

- Exploiting the parallelization power of GPUs for reconstruction algorithms
- Example: long tracks reconstruction algorithms [A. Scarabotto at al. IEEE Access (2024) Vol12, 10.1109](#)



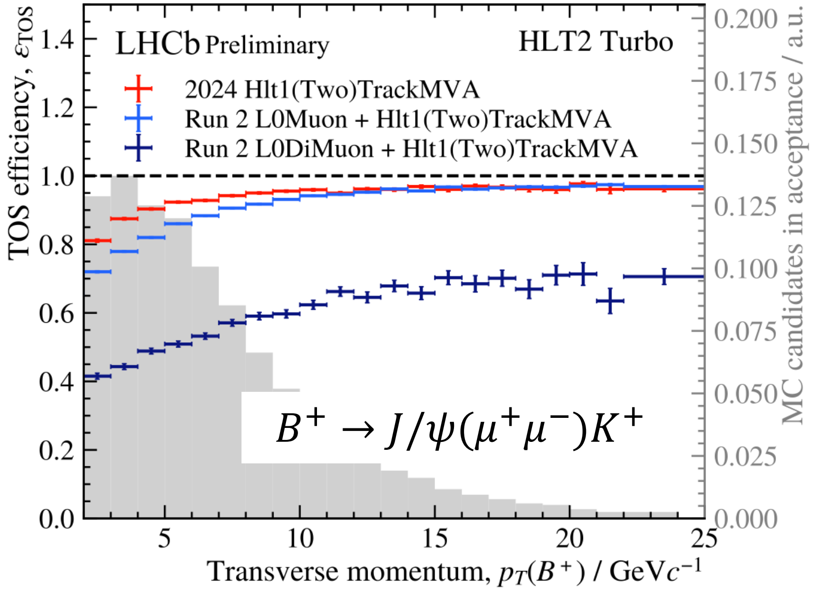


The current full HLT1 sequence is running at 70 kHz per GPU (x 500 GPUs = 35 MHz !)

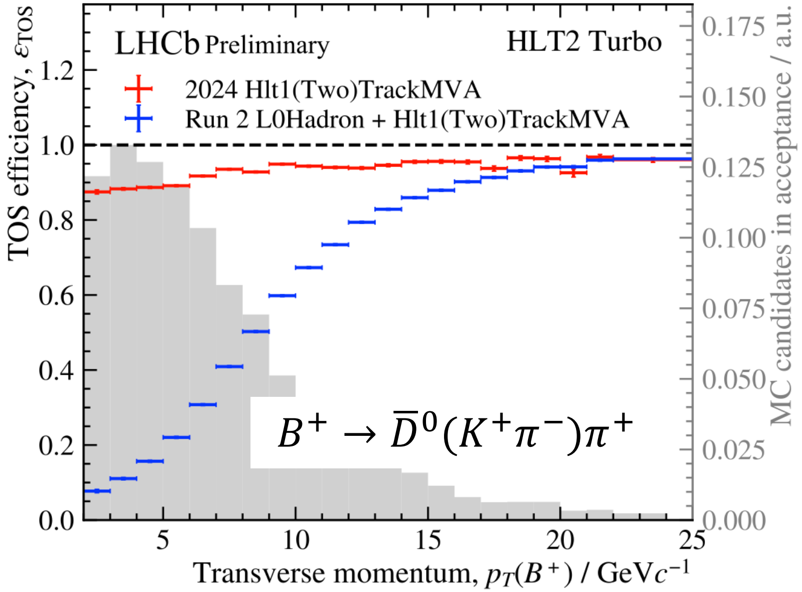
HLT1 performance on 2024 data

- Exploiting reconstructed objects to select decays of interest
- Output rate must be around 1MHz
- Comparison with Run2 trigger efficiencies, limited by L0 selections
- Clear gain at low momentum for hadronic and electronic B-mesons modes

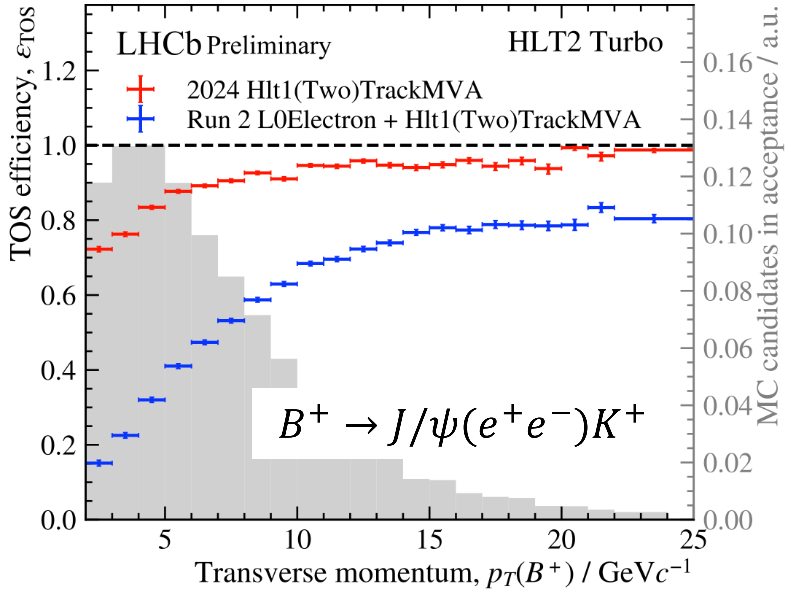
Muonic



Hadronic



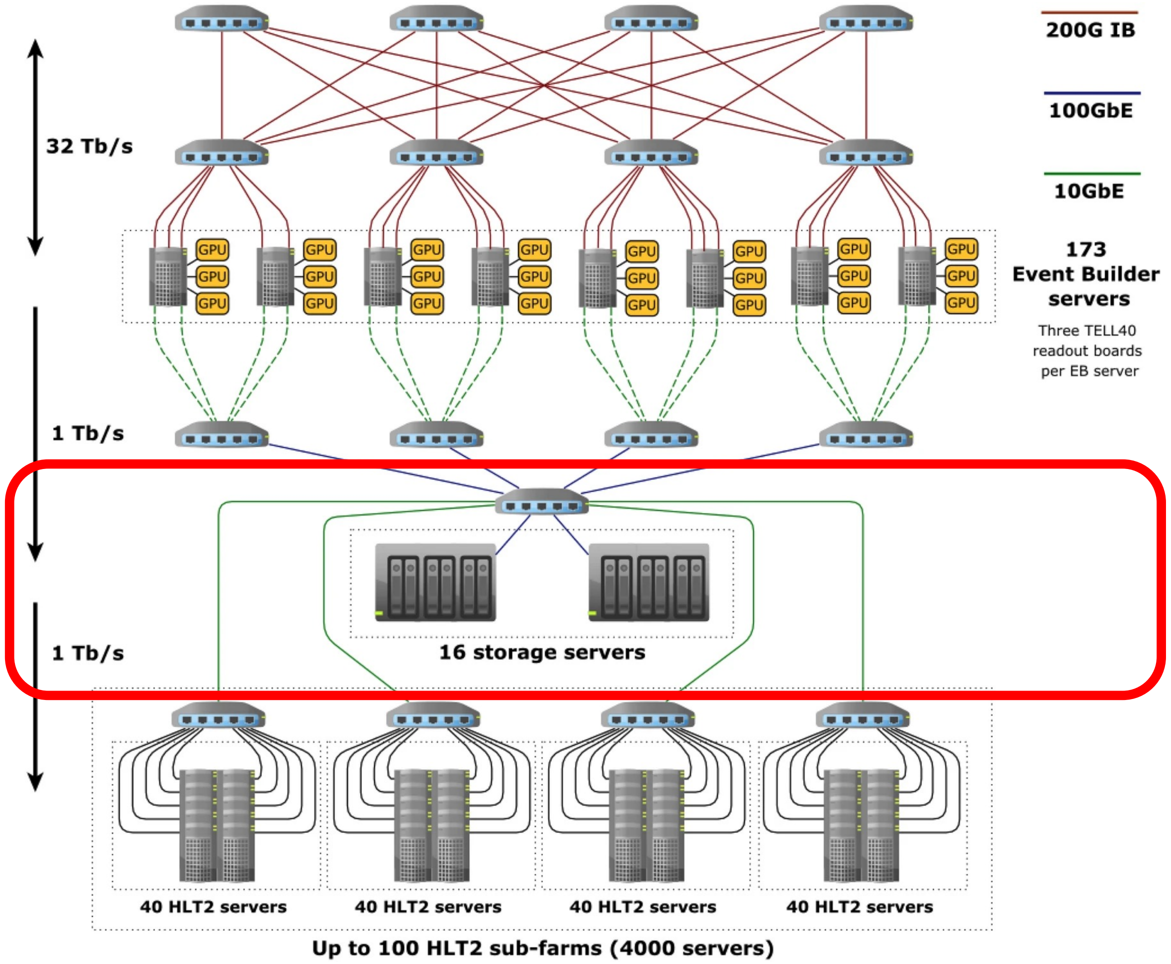
Electronic



LHCB-FIGURE-2024-030

Alignment & Calibration

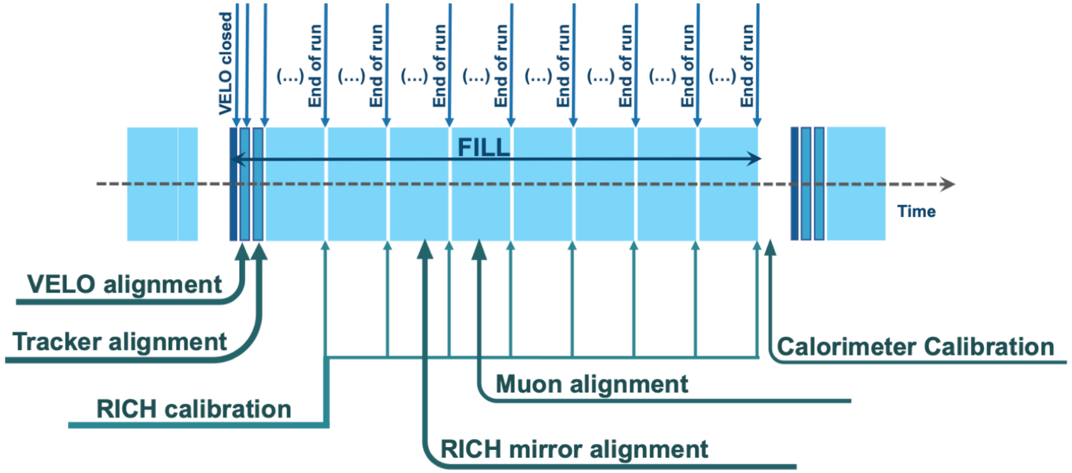
- Output of HLT1 is sent to a 40 Petabytes storage
- While data is stored, perform full detector alignment and calibration

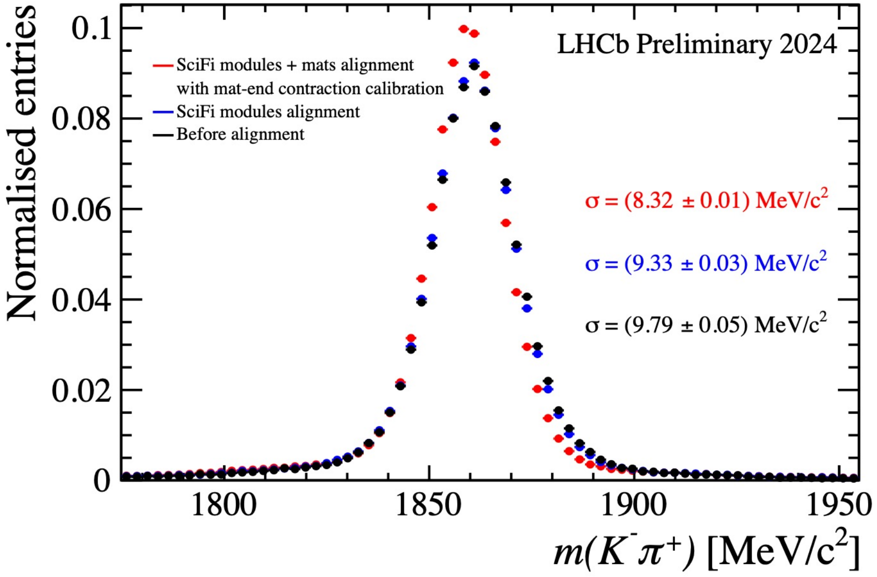


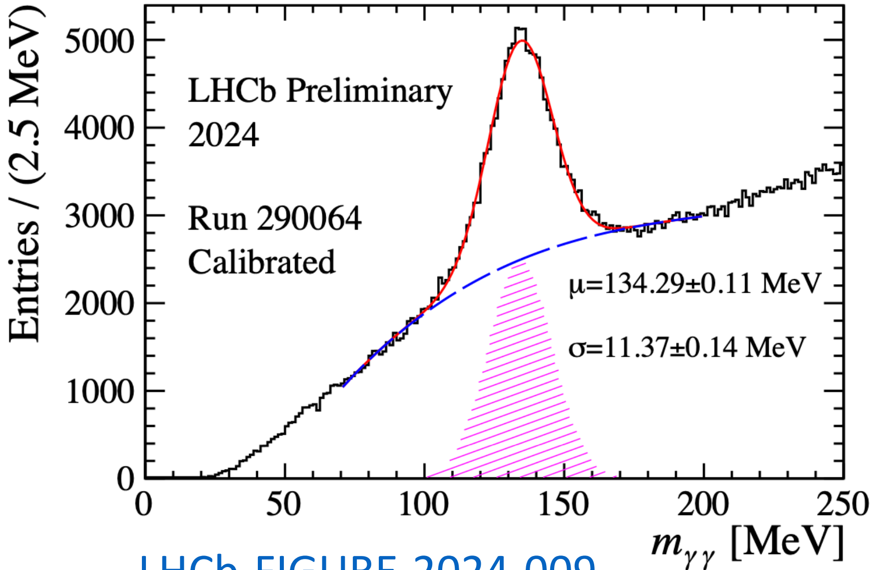
[Comput.Softw.Big Sci. 6 (2022) 1, 1]

Alignment & Calibration

- Alignment of the tracking detectors, muon chambers and RICH mirrors
- Calibration of RICH detectors and calorimeters
- Performed at each LHC fill or more frequently
- Critical to ensure offline-like quality of HLT2 reconstruction



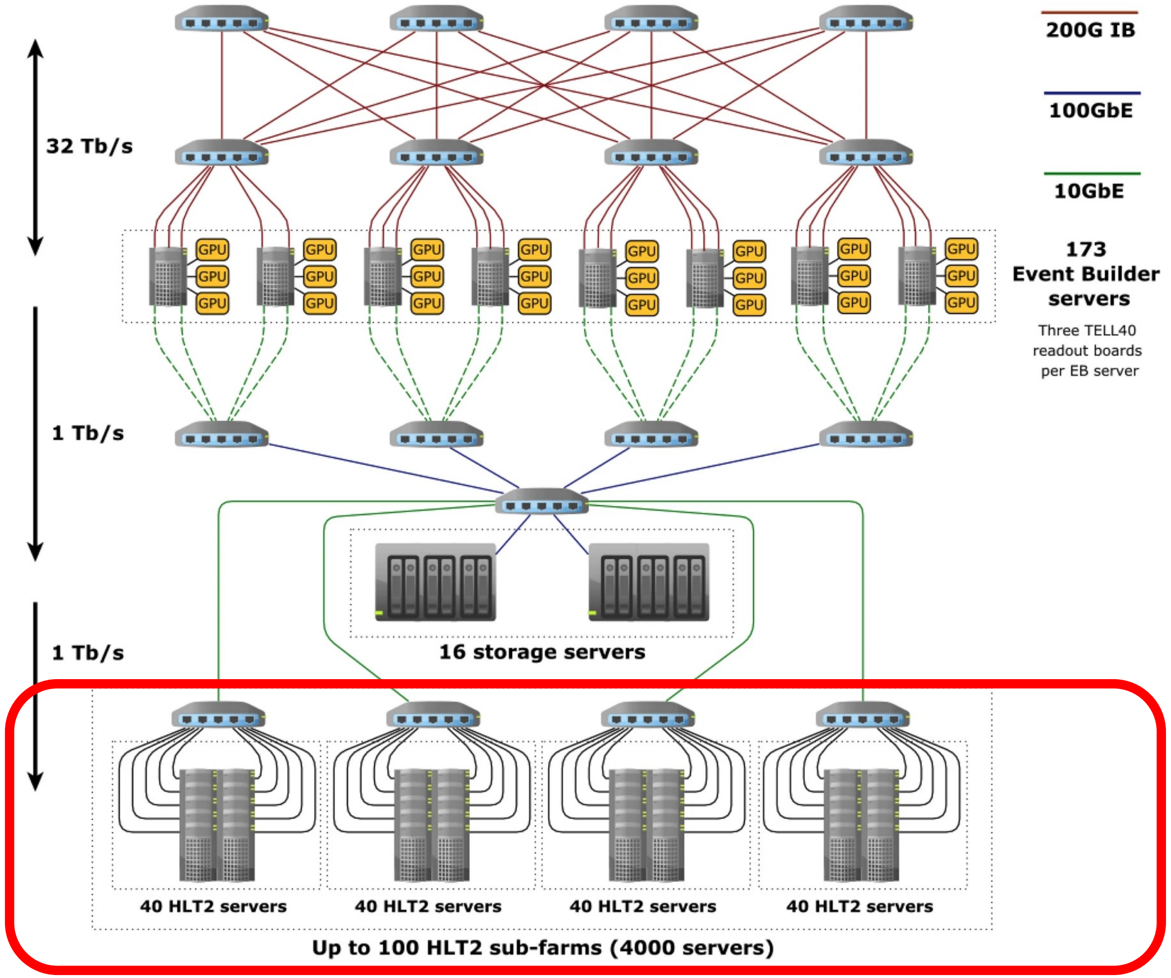




LHCb-FIGURE-2024-009

The second high-level trigger: HLT2

- HLT2 can be run asynchronously to HLT1 once the full alignment&calibration is performed
- HLT2 needs to process data at rate greater than half of the HLT1 output (1 MHz): minimum 500 kHz
- Dedicated trigger selections O(3000) to cover broad LHCb physics program
- Limited bandwidth of 10 GB/s of data saved in memory

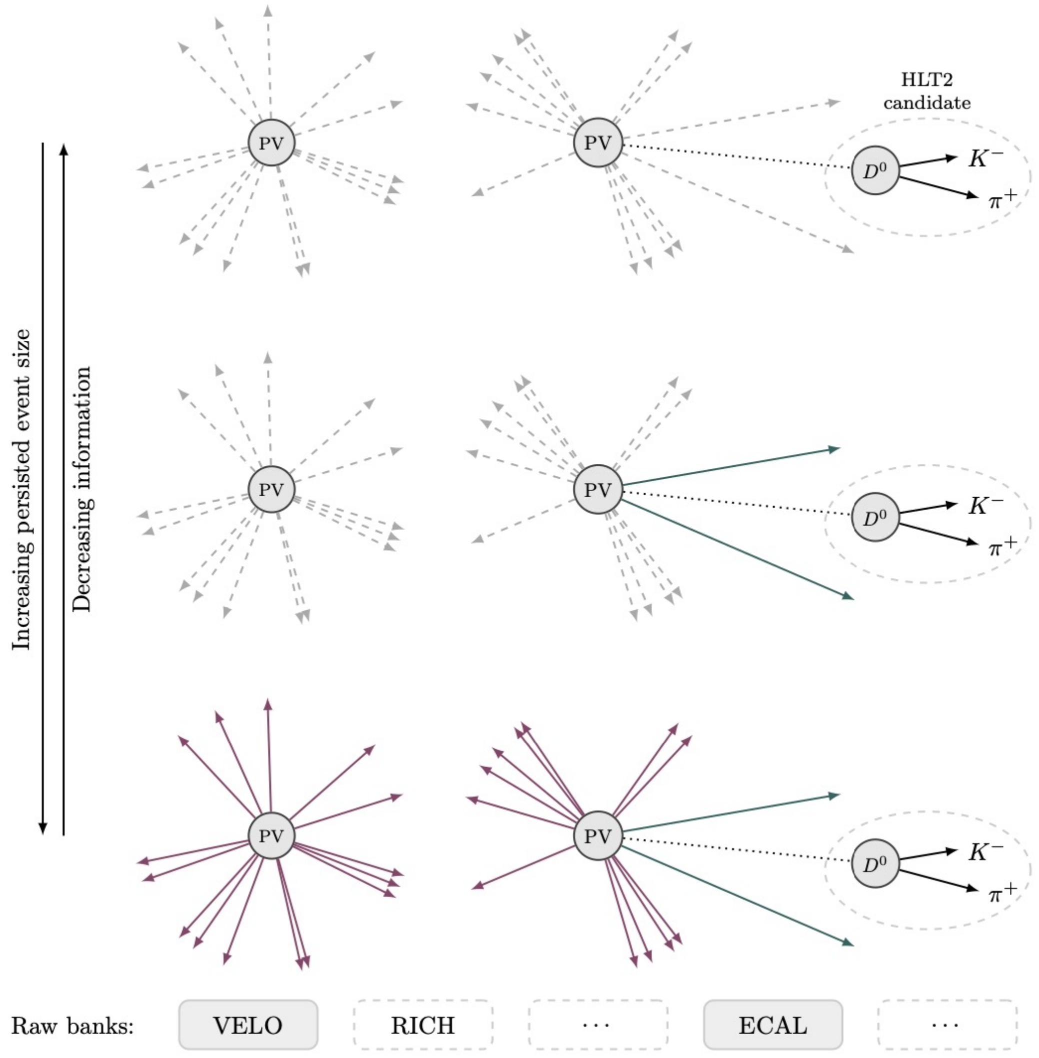


[Comput.Softw.Big Sci. 6 (2022) 1, 1]

HLT2 throughput and bandwidth

JINST 14 (2019) P04006

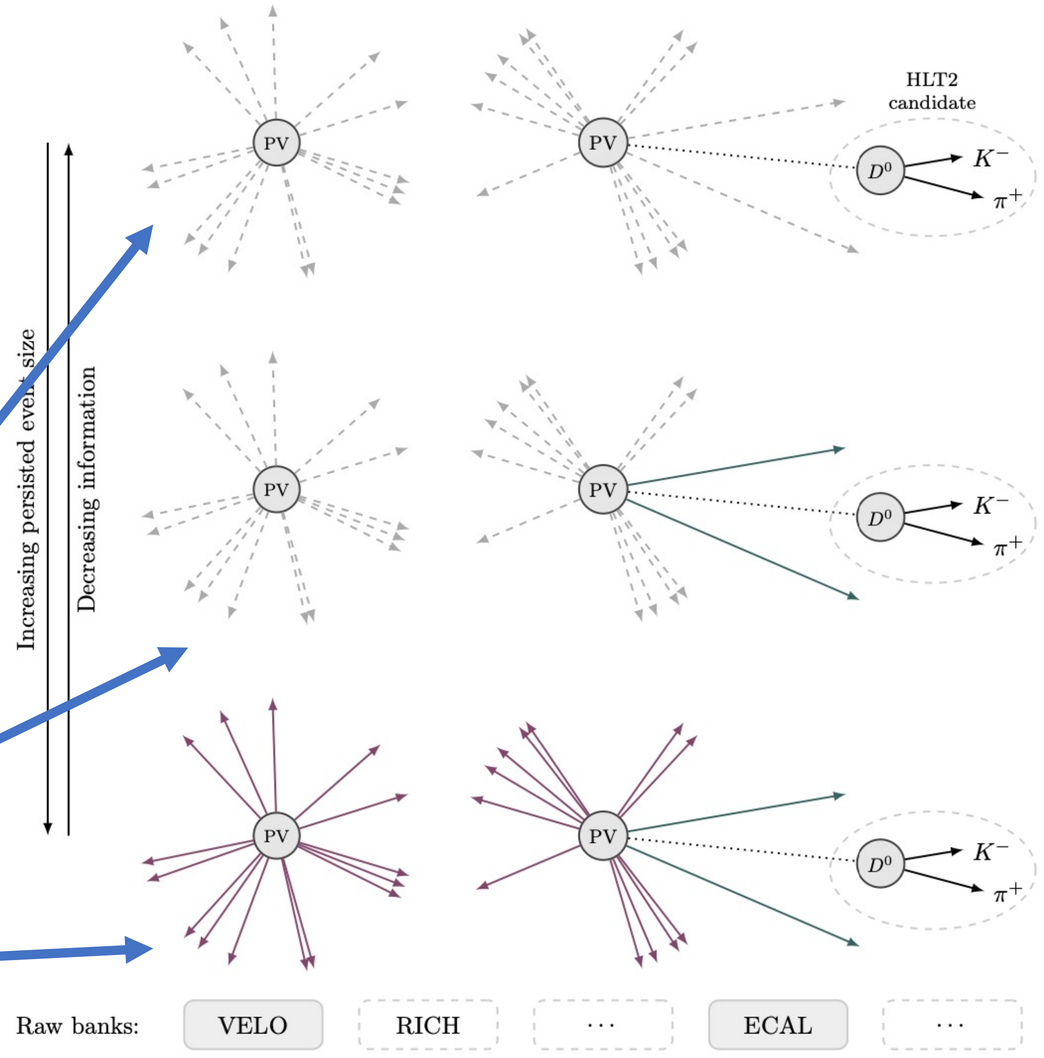
- Throughput (minimum 500 kHz):
  - Structure-of-Arrays collections to exploit vectorisation and multi-threading
  - Throughput-Oriented selections (Thor functors): built at compile time into cache memory and agnostic on I/O type



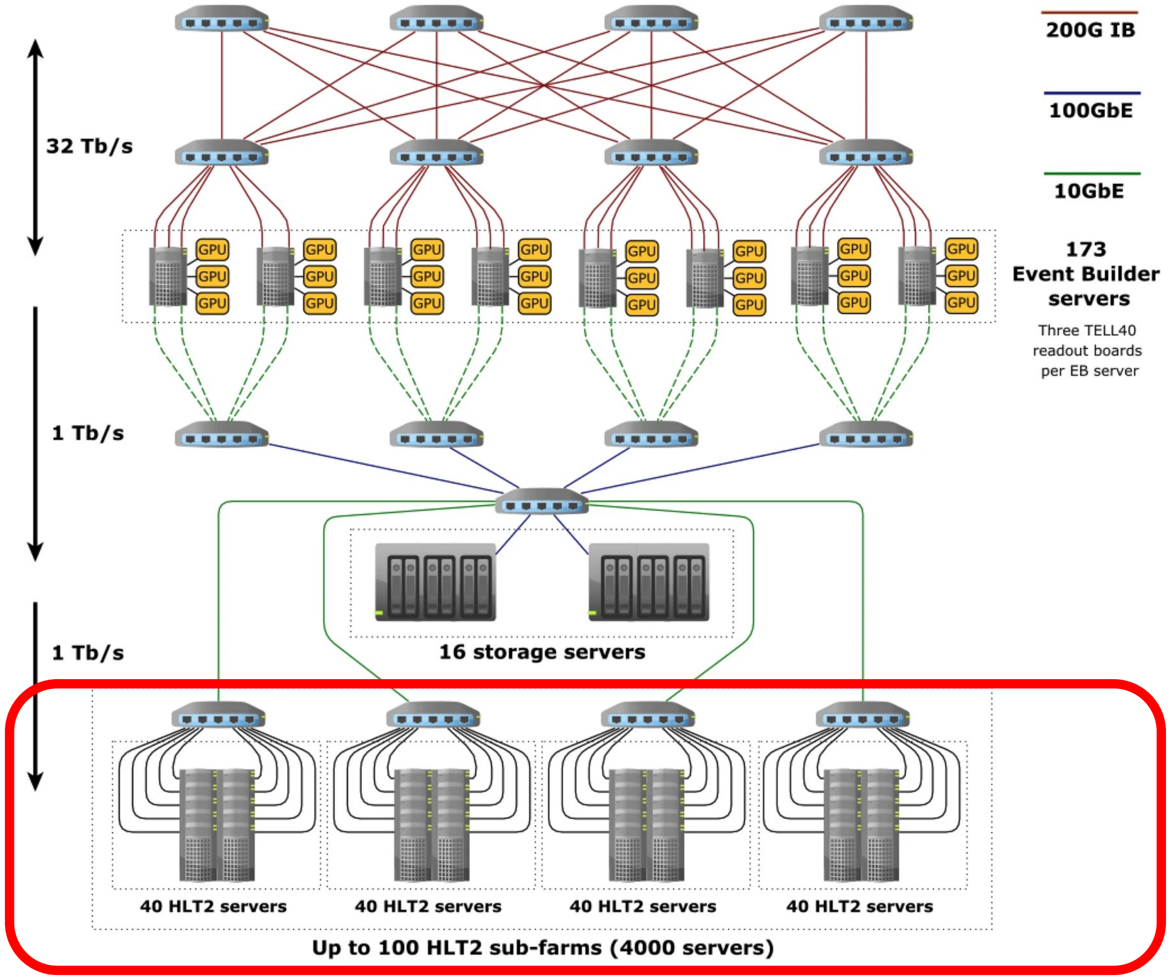
HLT2 throughput and bandwidth

JINST 14 (2019) P04006

- Throughput (minimum 500 kHz):
  - Structure-of-Arrays collections to exploit vectorisation and multi-threading
  - Throughput-Oriented selections (Thor functors): built at compile time into cache memory and agnostic on I/O type
- Bandwidth (maximum 10 GB/s):
  - Turbo: saving only info related to the signal candidate reducing the event size by a factor 10 (about 70% of all LHCb selections)
  - Selective persistency: save additional objects relative to the signal
  - Full persistency



The second high-level trigger: HLT2

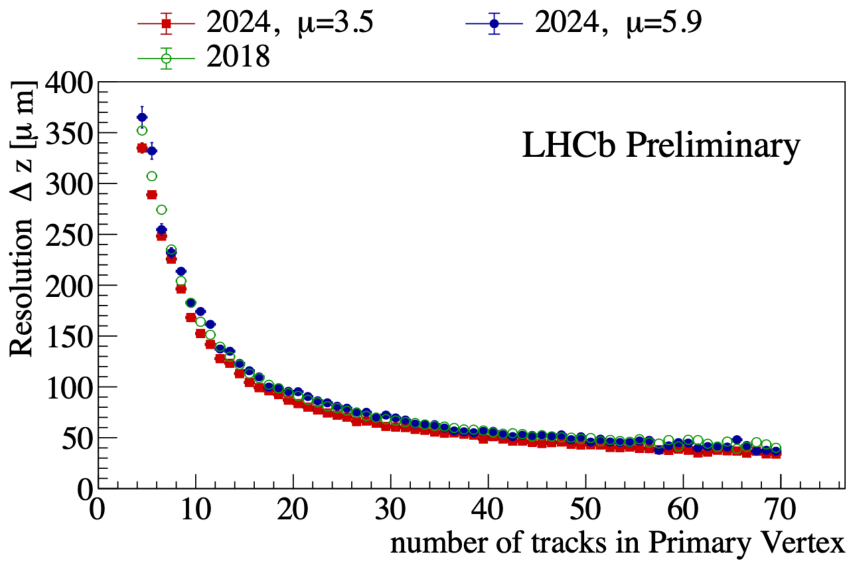


- With the current ~4500 CPUs we achieved a HLT2 throughput of 900 kHz (well above the 500kHz minimum) by keeping the bandwidth around 9-10 GB/s!

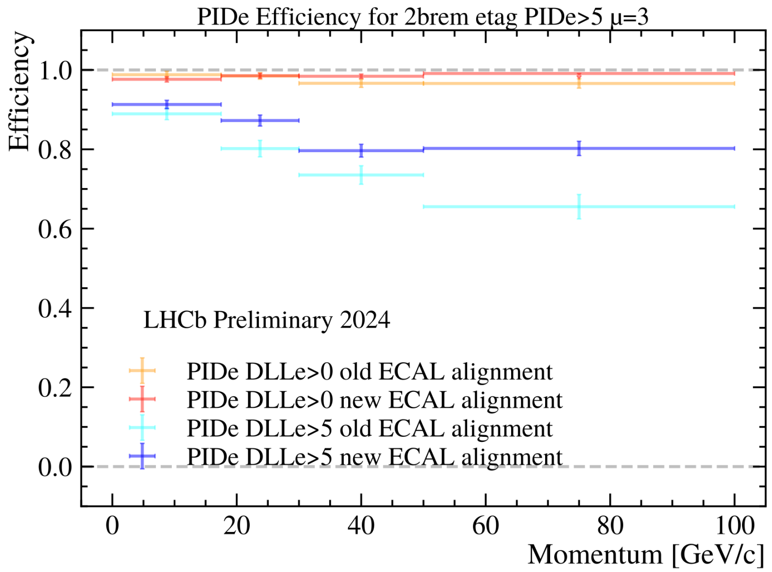
[Comput.Softw.Big Sci. 6 (2022) 1, 1]

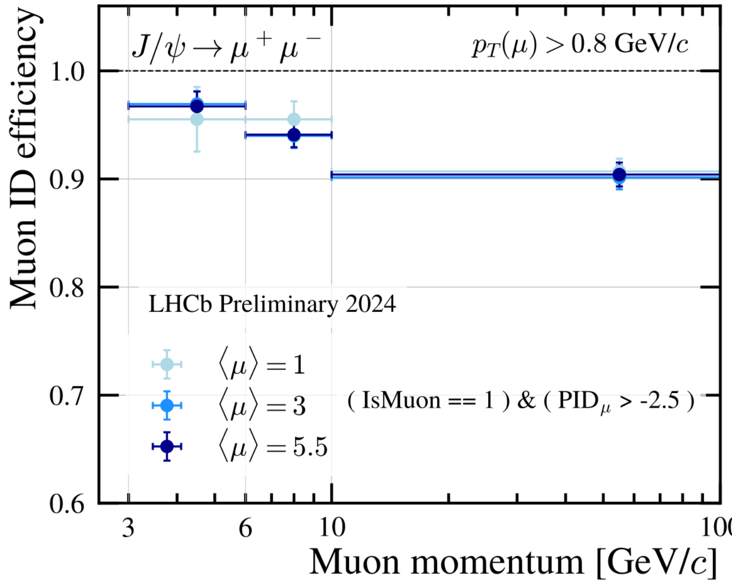
Performance of HLT2 on 2024 data

- Achieving excellent vertex resolutions, good track reconstruction and stable PID performance for muons, hadrons and electrons



[LHCB-FIGURE-2024-011]



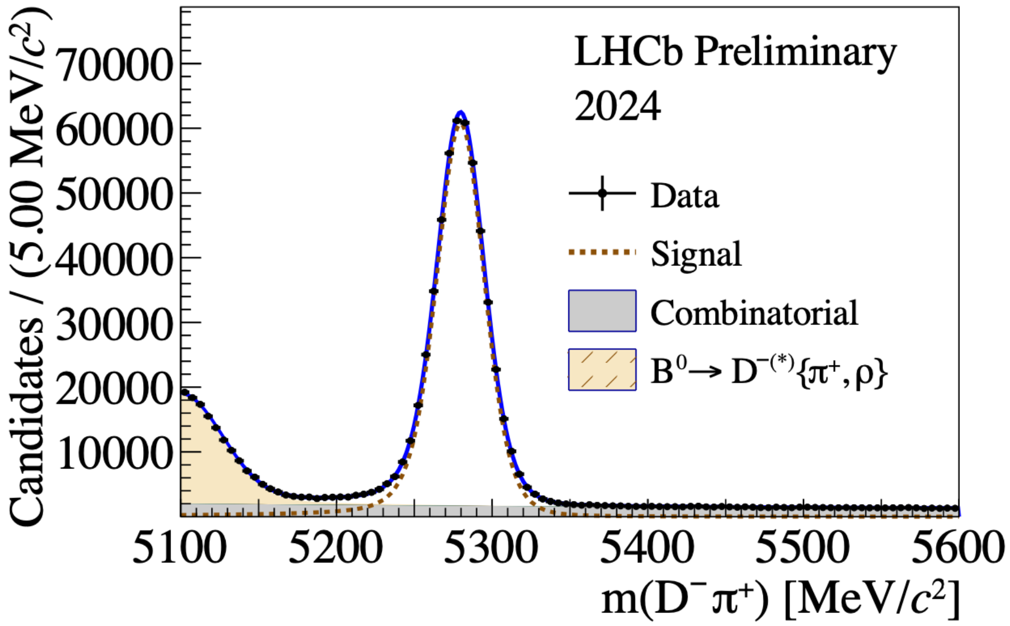


[LHCB-FIGURE-2024-010]

Performance of HLT2 on 2024 data

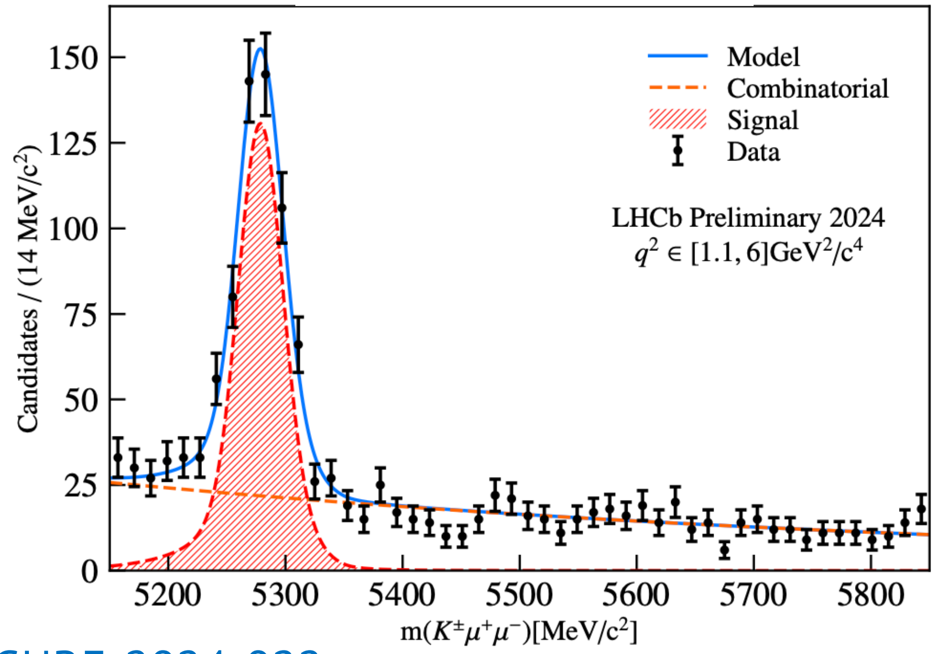
- Achieving excellent vertex resolutions, good track reconstruction and stable PID performance for muons, hadrons and electrons
- Leading to reconstruction and selection of decays of interest

$B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$



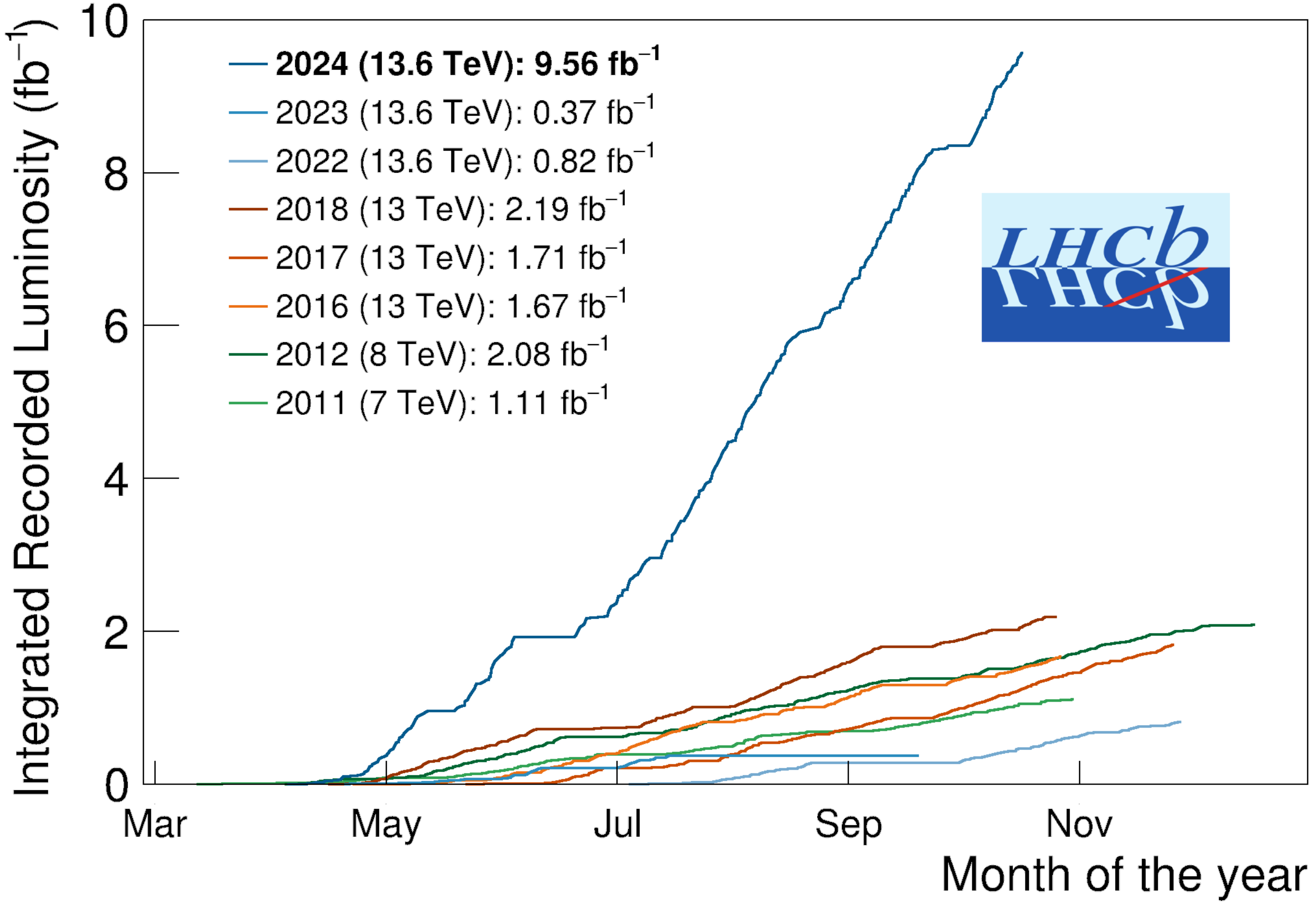
LHCB-FIGURE-2024-021

$B^+ \to K^+\mu^+\mu^-$



LHCB-FIGURE-2024-022

Very successful 2024 data taking year



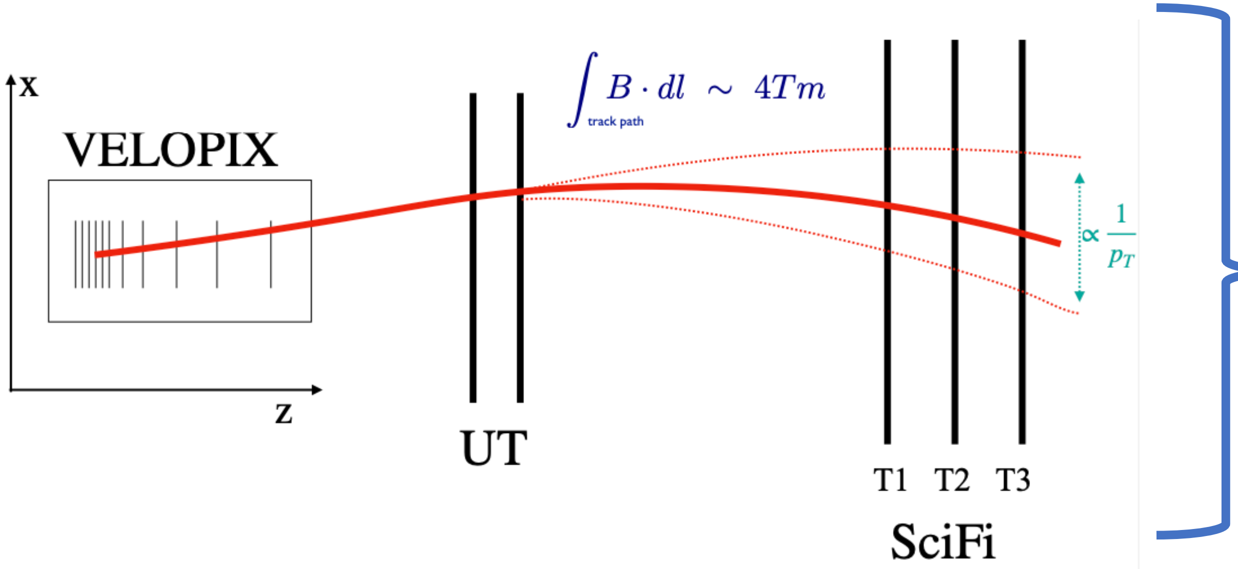
Conclusions

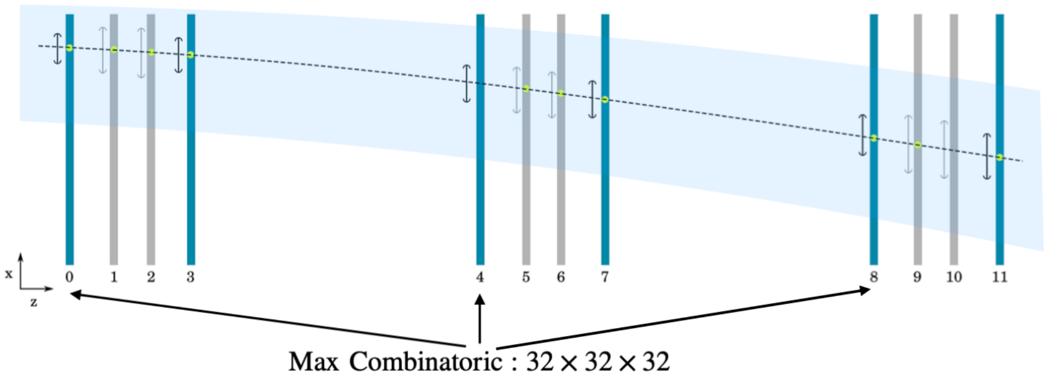
- LHCb taking data with a fully-software trigger: successfully in Run3!
- Heterogeneous system:
  - First trigger stage (HLT1) optimised on GPU dealing with 30 MHz LHC input rate
  - Performing alignment&calibration before running second stage (HLT2)
  - CPU-based HLT2 performs offline-like reconstruction including PID information
- HLT1, alignment and HLT2 achieve expected performance during 2024 data taking (more than 9 fb^{-1} collected)
- For the future: exploiting our knowledge on heterogeneous systems to port also HLT2 fully or partially on GPUs (probably Upgrade2 from 2032)

Backup

Reconstruction in GPU

- How to fully exploit parallelization power of GPUs?
- Parallelization levels when reconstructing tracks traversing the whole LHCb detector:
  1. Over events, independent p-p collisions
  2. Over input tracks, extrapolate straight tracks in VELO+UT into the magnetic field reaching the SciFi
  3. Over hits in SciFi, meaning possible extrapolations segments





[IEEE Access 12 (2024) 114198-114211]