



LHCb HLT1

Tracking and vertexing at 30 MHz with GPUs

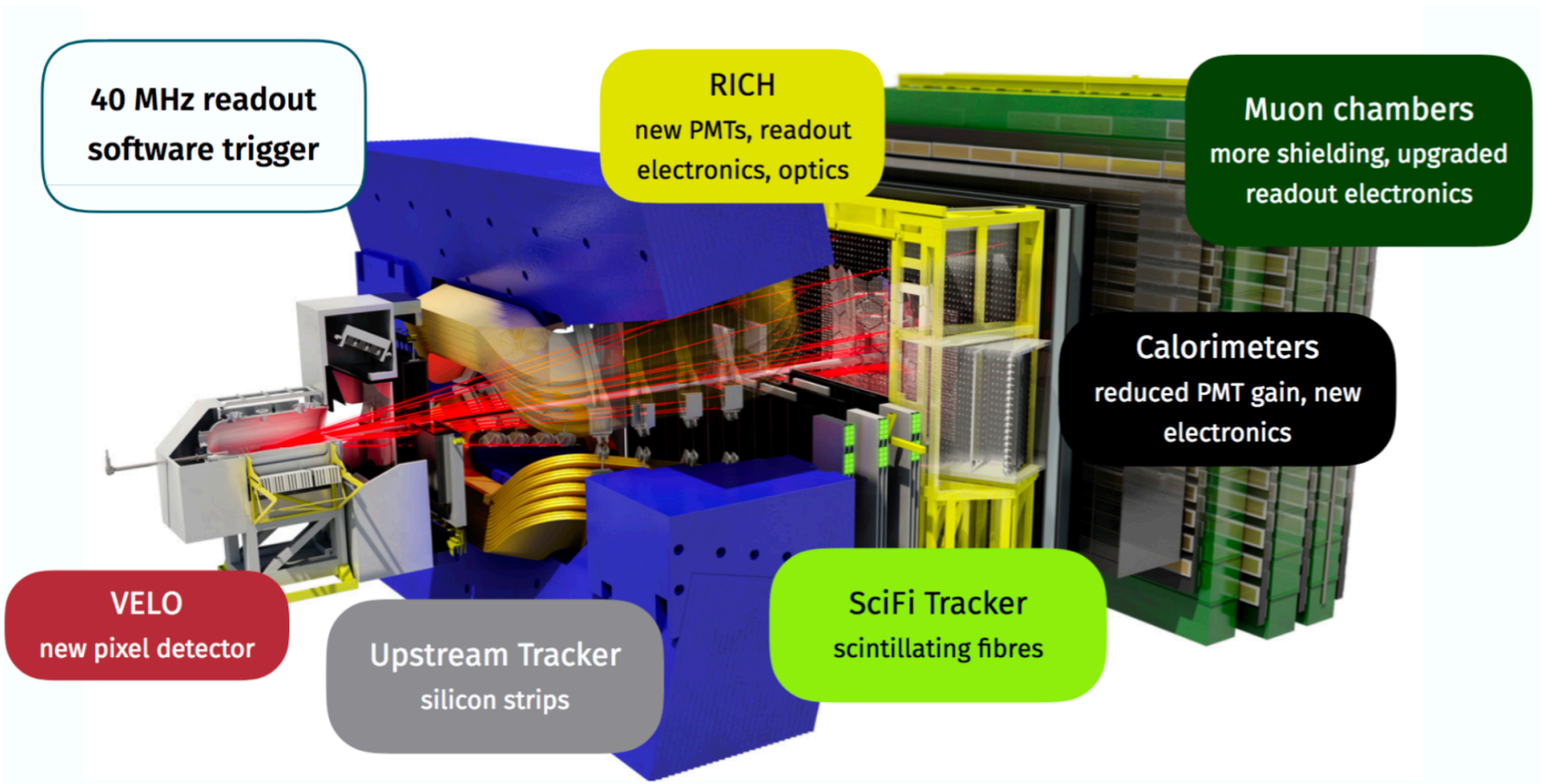
Marianna Fontana on behalf of the LHCb collaboration

CEPC 2022 - 27/10/2022

The LHCb experiment

CERN-LHCC-2012-007

- Forward spectrometer ($2 < \eta < 5$)
- Optimised for b physics -> nowadays a multipurpose experiment
- Run 1+2: 9 fb^{-1} of pp collisions

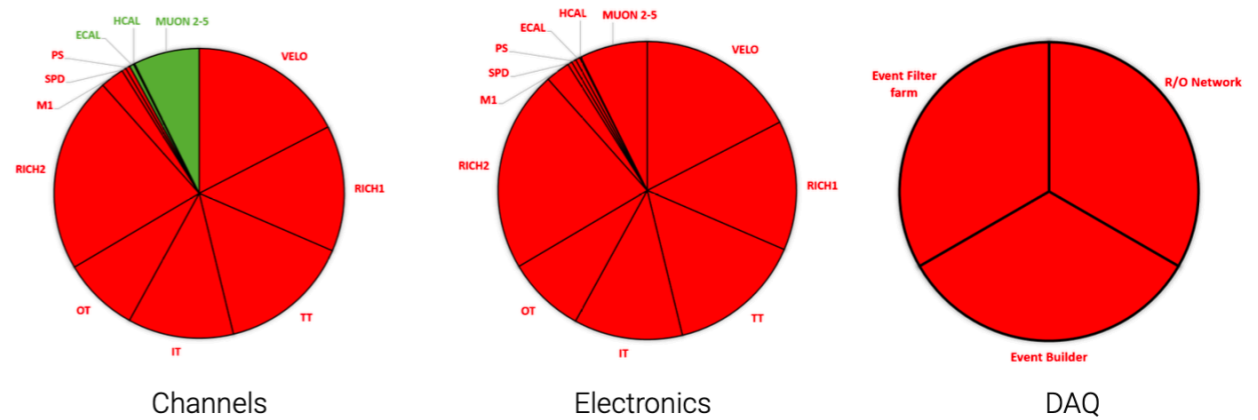


- Major upgrade of all subdetectors:
 $\mathcal{L} = 2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$
 pile-up ≈ 5
 Run 3+4 aiming to collect 50 fb^{-1}

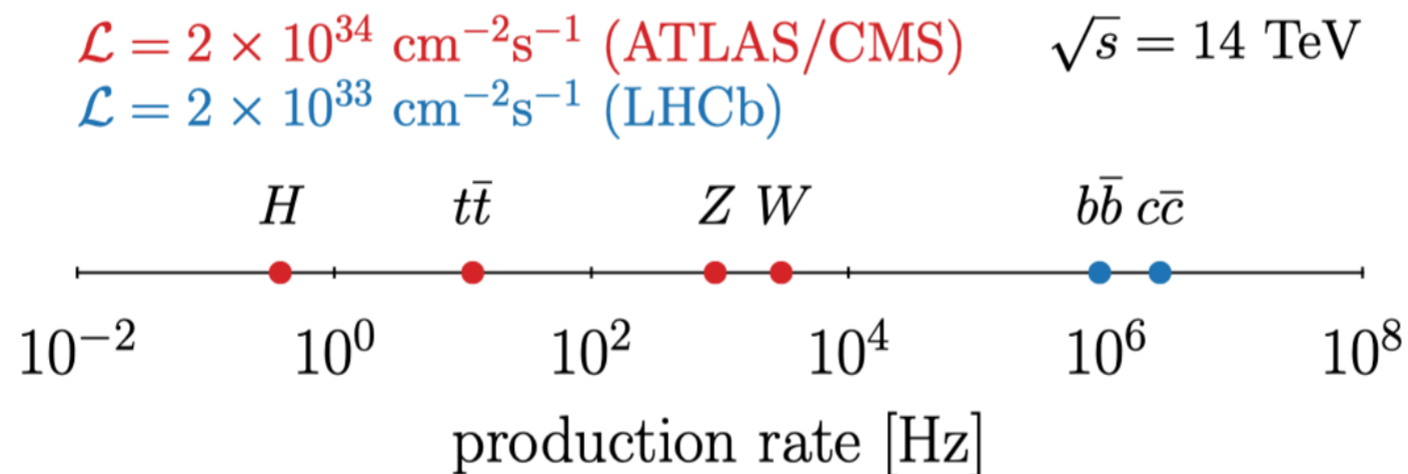
- 100% of the readout electronics replaced
- New data acquisition system and data center

- New pixel-detector **VELO**
- New **RICH** mechanics, optics, photodetector
- New silicon strip upstream tracker **UT**
- New **SciFi** tracker
- New electronics for **MUON** and **CALO**
- New luminometer **PLUME**

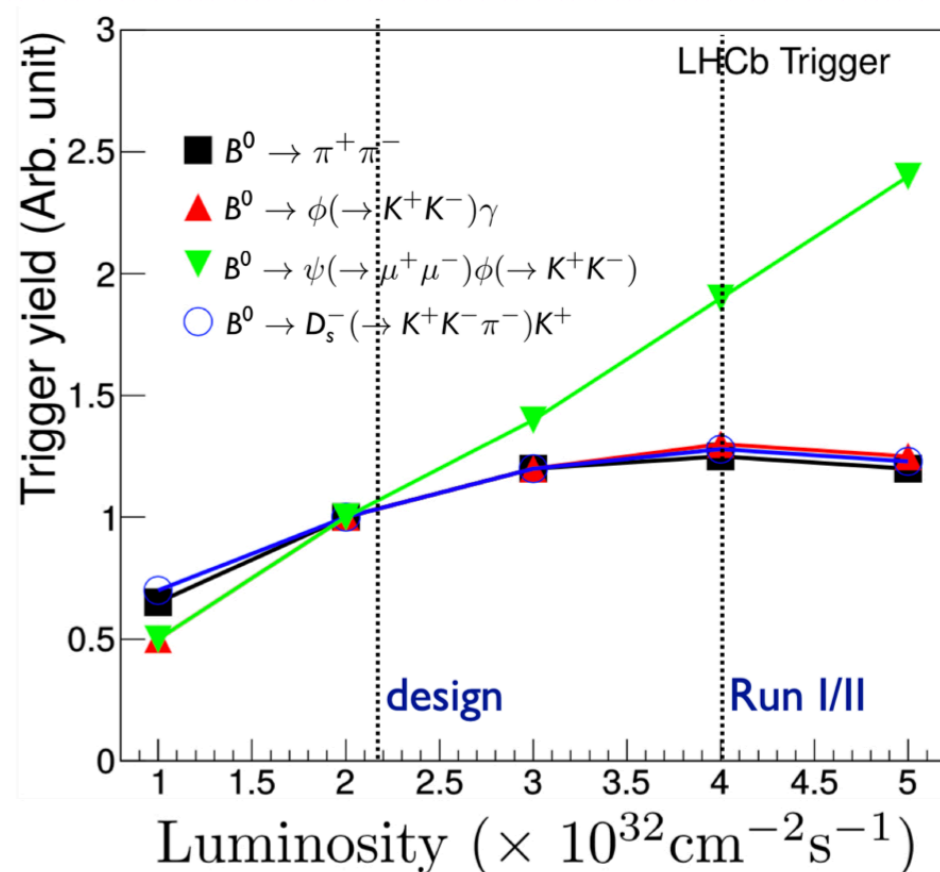
Upgraded; Kept



The need for an upgraded trigger

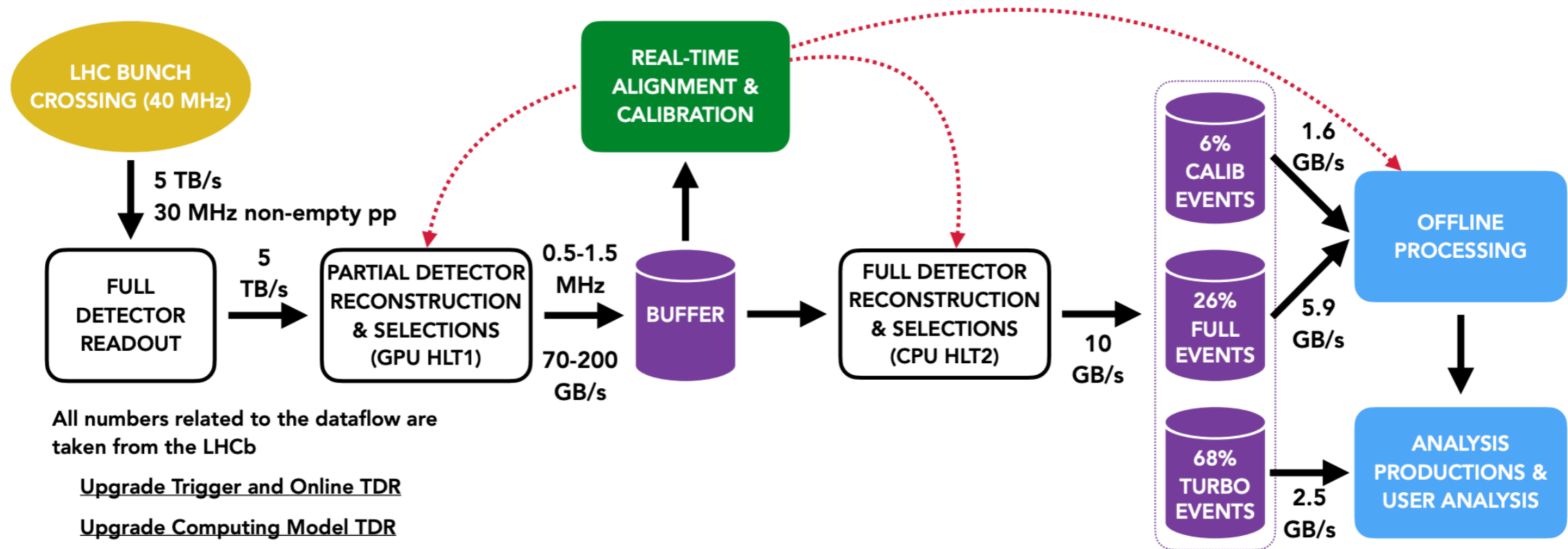


- Heavy flavour hadrons decay to final-state particles with momenta similar to those of particles from the underlying event
- LHCb has to distinguish between signal and signal-like background
- Cannot effectively trigger on heavy flavour using hardware signatures
- Trigger for many hadronic channels saturated already at Run 1-2 luminosity



The LHCb data flow

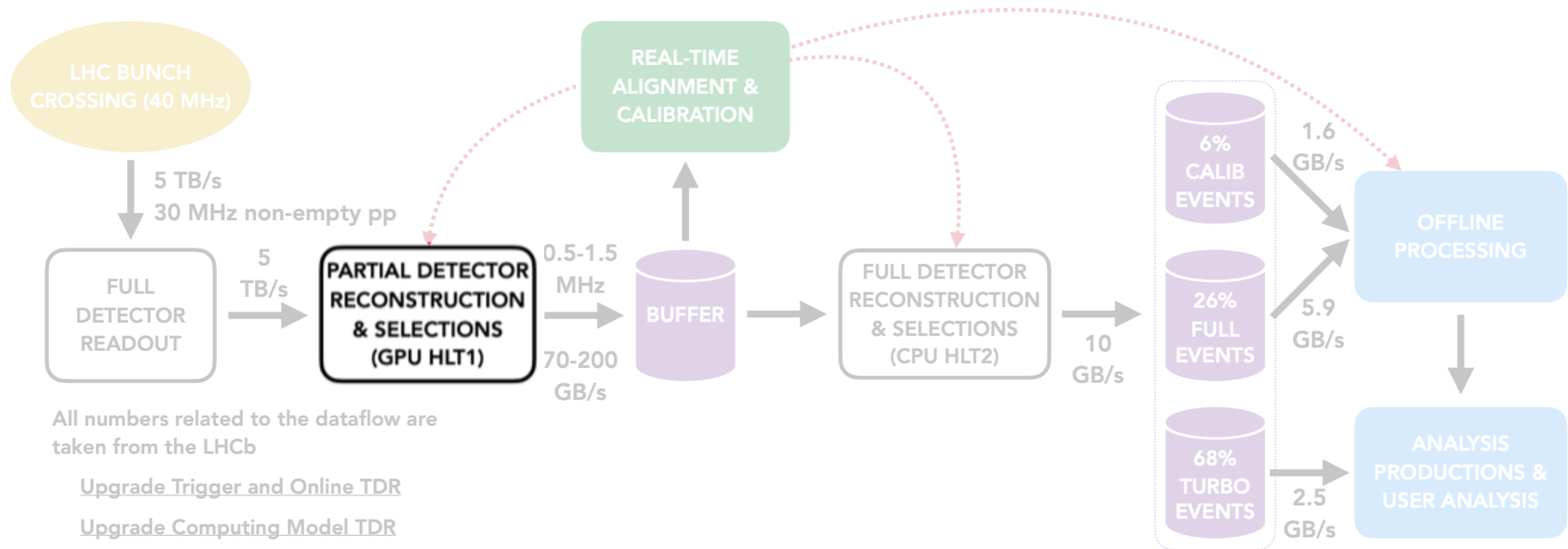
LHCb-FIGURE-2020-016



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

The LHCb data flow

LHCb-FIGURE-2020-016

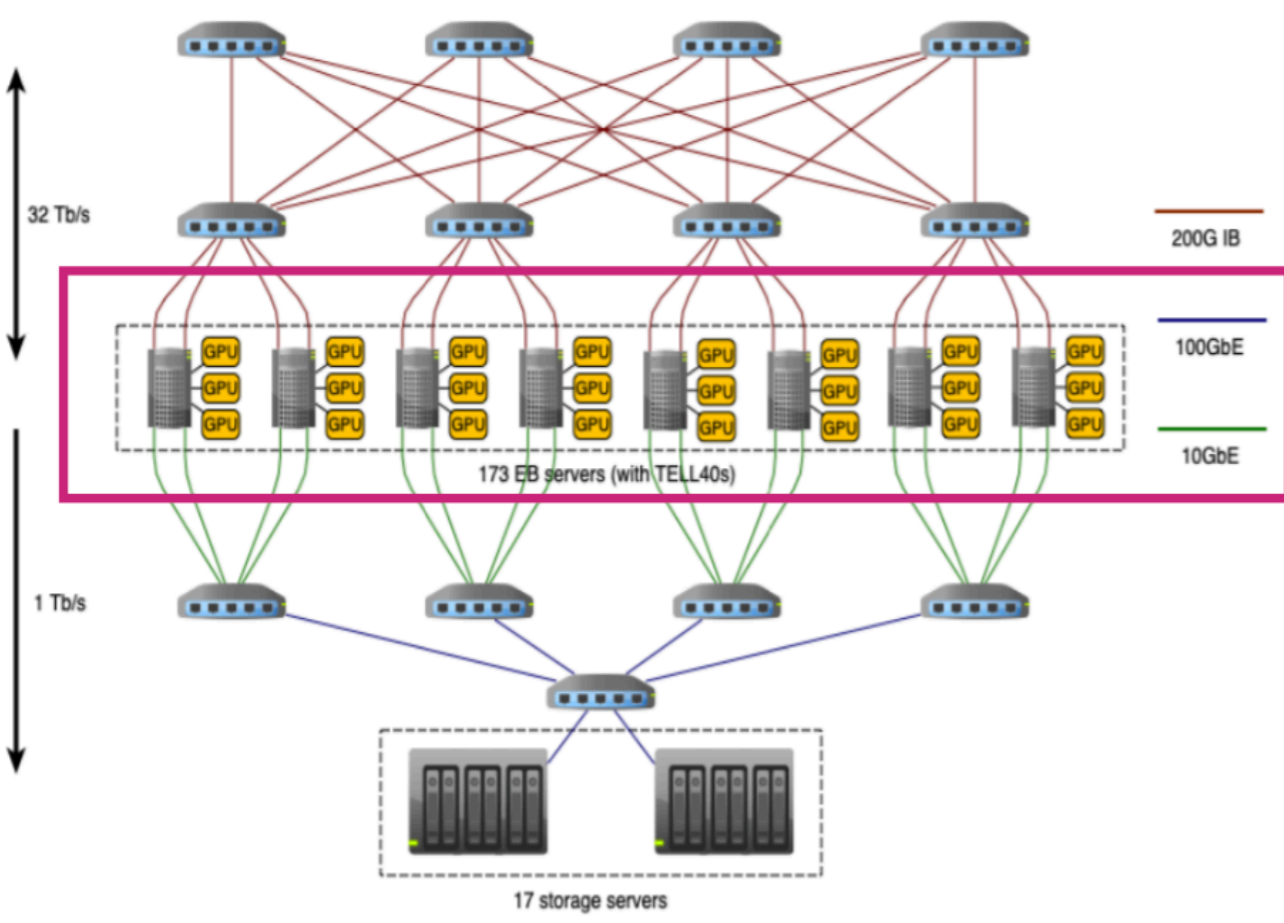


The goal of HLT1:

- Be able to intake the entirety of the LHCb raw data (5 TB/s) at 30 MHz
- Perform partial event reconstruction & coarse selection of broad LHCb physics cases
- Reduce the input rate by a factor of 30 (~ 1 MHz)
- Store selected events in intermediate buffer for real-time alignment and calibration

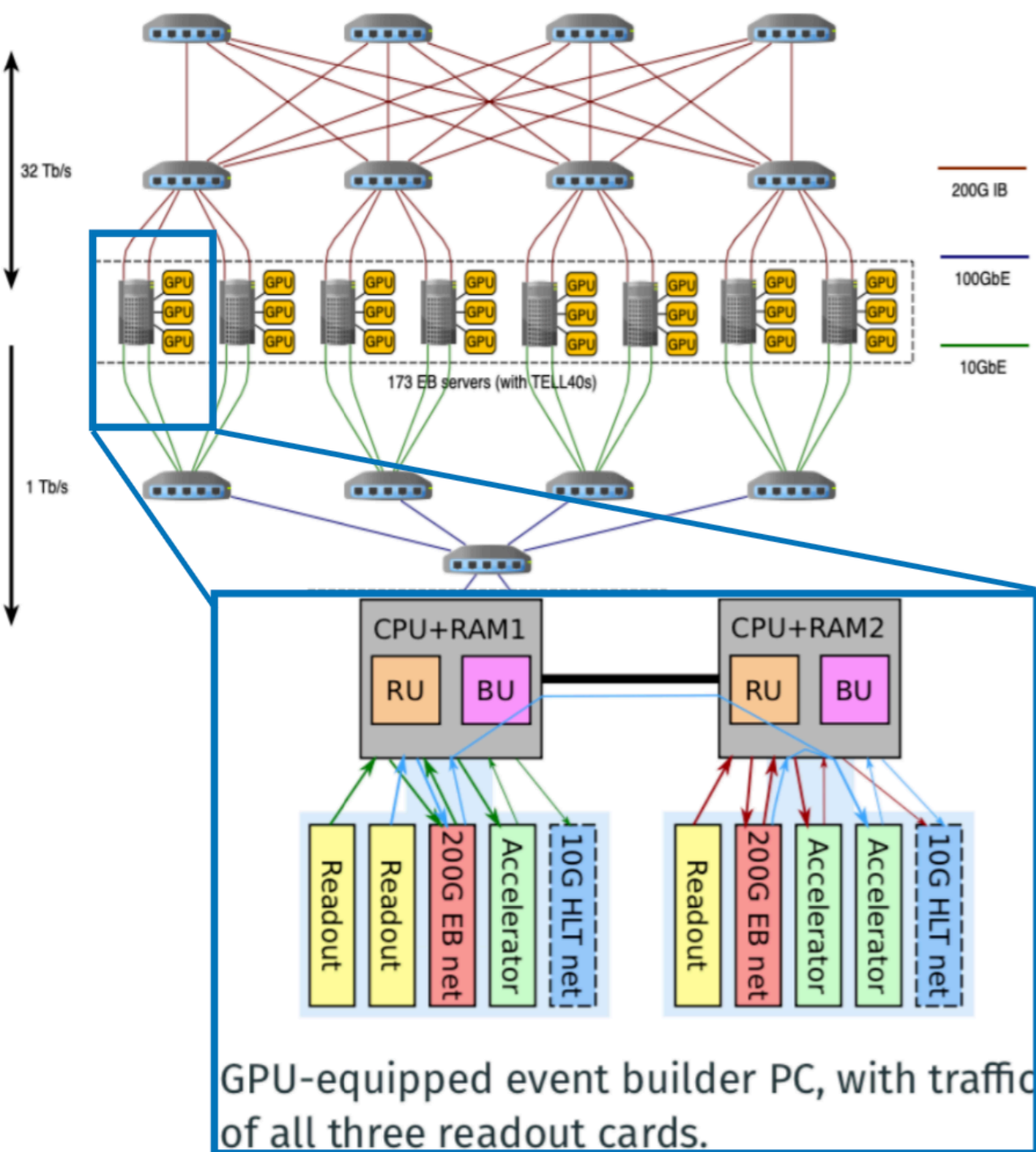
First complete high-throughput GPU trigger for a HEP experiment!

The convenience of GPUs



- Event builder farm equipped with 173 servers

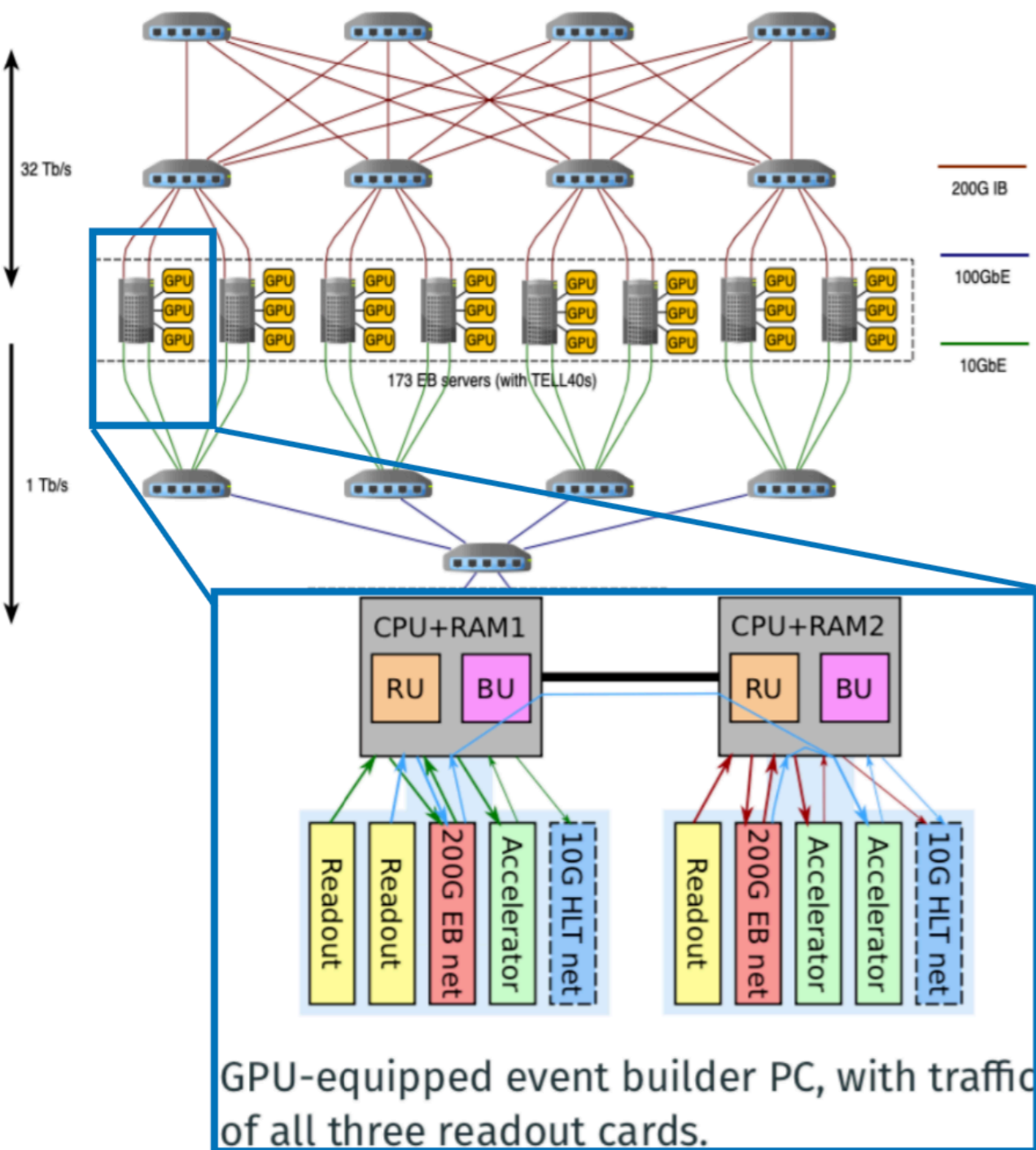
The convenience of GPUs



- Event builder farm equipped with 173 servers

- Each server has 3 free PCIe slots
 - can host GPUs
 - sufficient cooling and power
 - advantageous to have GPUs as self-contained processors
 - sending data to GPUs is like sending data to network card

The convenience of GPUs



- Event builder farm equipped with 173 servers

- Each server has 3 free PCIe slots
 - can host GPUs
 - sufficient cooling and power
 - advantageous to have GPUs as self-contained processors
 - sending data to GPUs is like sending data to network card

- GPUs map well into LHCb DAQ architecture
- HLT1 tasks inherently parallelizable
- Smaller network between EB & CPU HLT
- Cheaper & more scalable than CPU alternative
- Implemented with O(200) Nvidia RTX A5000 GPUs

Allen: a GPU HLT1 trigger platform

- Public software project: [gitlab repo](#)
- Supports three modes:
 - Standalone
 - Compiling within the LHCb framework for data acquisition
 - Compiling within the LHCb framework for simulation and offline studies
- Runs on CPU, Nvidia GPU (CUDA, CUDACLANG), AMD GPUs (HIP)
- GPU code written in CUDA
- Cross-architecture compatibility (HIP, CPU) via macros

Allen

pipeline **passed**

Welcome to Allen, a project providing a full HLT1 realization on GPU.

Documentation can be found [here](#).

Mattermost discussion channels

- [Allen developers](#) - Channel for any Allen algorithm development discussion.
- [Allen core](#) - Discussion of Allen core features.
- [AllenPR throughput](#) - Throughput reports from nightlies and MRs.

Performance monitoring

- [Allen throughput evolution over time in grafana](#)
- [Allen dashboard with physics performance over time](#)

[Home](#) » Welcome to Allen's documentation!

[Edit on GitLab](#)

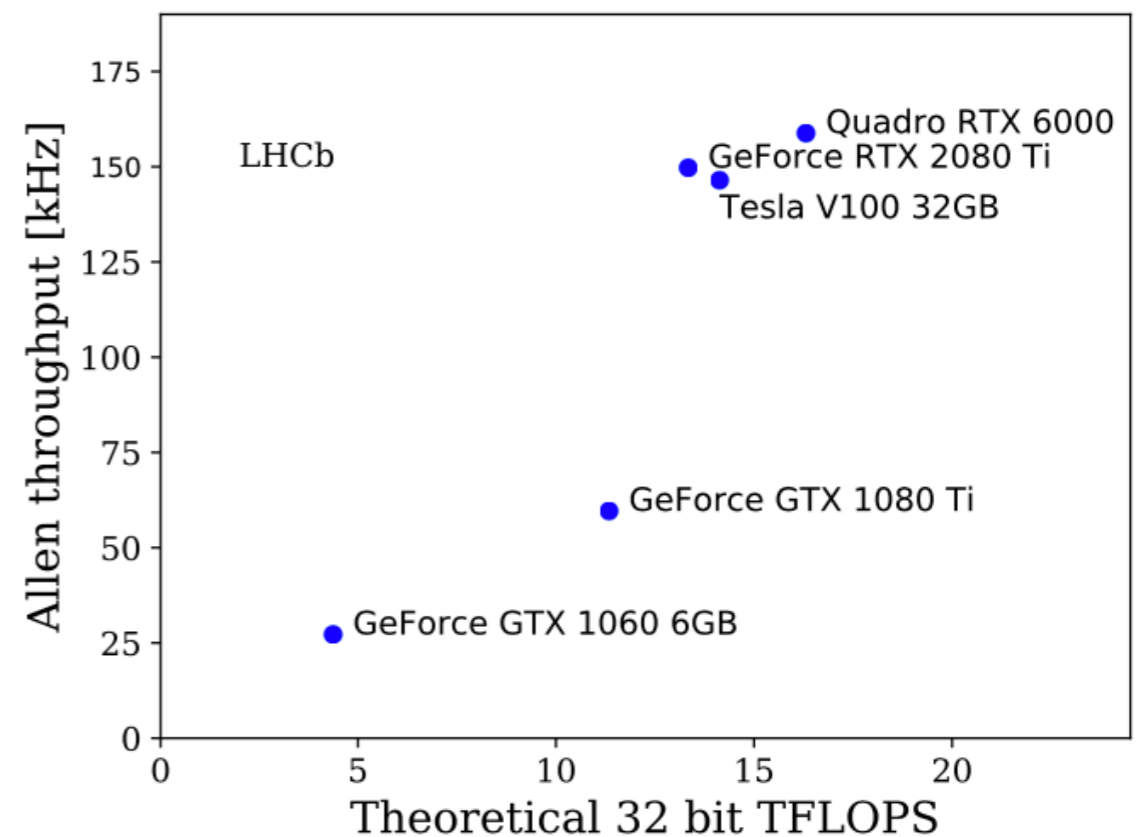
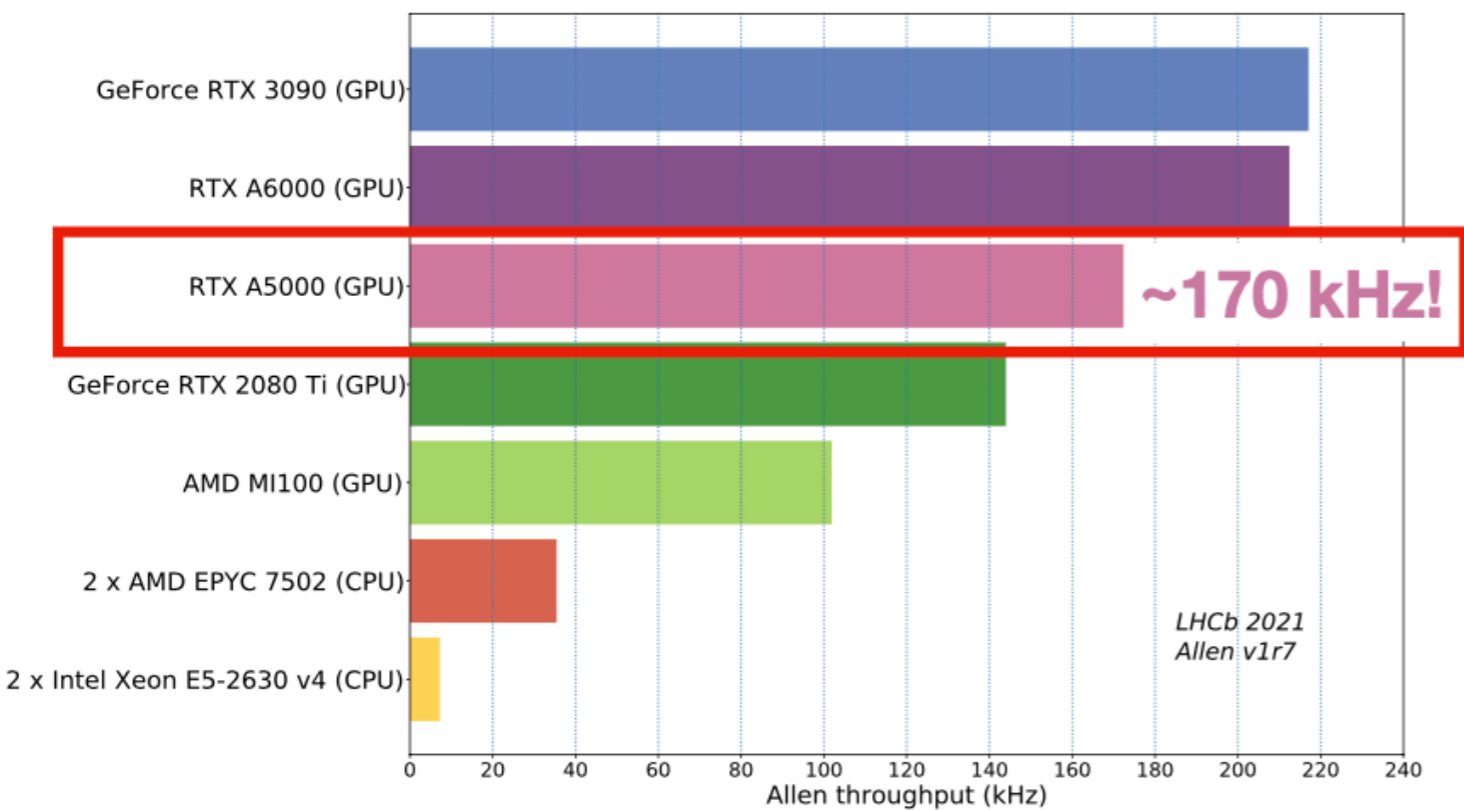
Welcome to Allen's documentation!

Allen is the LHCb high-level trigger 1 (HLT1) application on graphics processing units (GPUs). It is responsible for filtering an input rate of 30 million collisions per second down to an output rate of around 1-2 MHz. It does this by performing fast track reconstruction and selecting pp collision events based on one- and two-track objects entirely on GPUs.

This site documents various aspects of Allen.

Throughput

- 30 MHz benchmark can be achieved with O(200) GPUs (max number EB can host is 500)
- Throughput scales well with theoretical TFLOPs of GPU card
- Additional functionalities are being explored



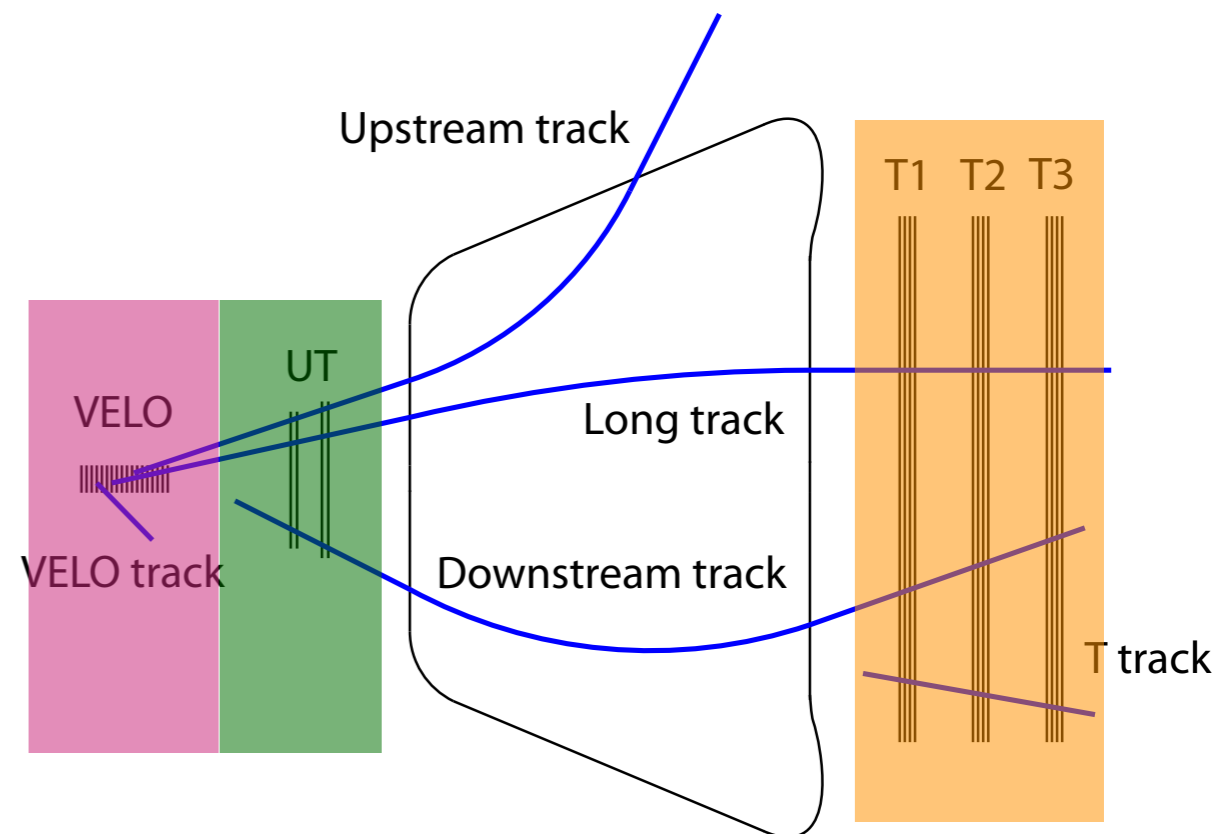
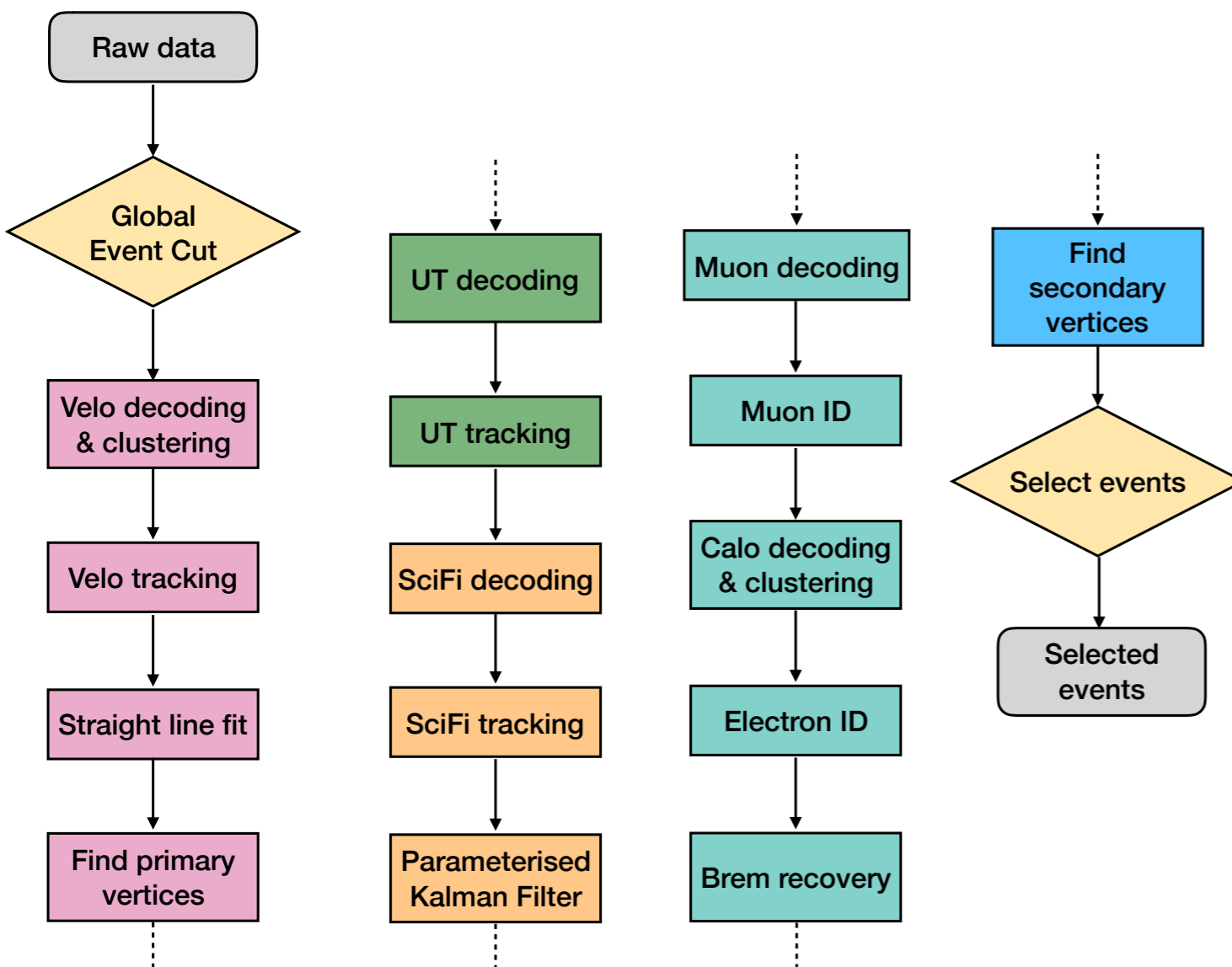
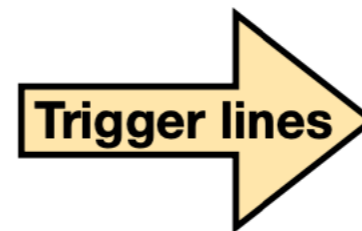
LHCb-FIGURE-2020-014

HLT1 sequence

Tracking relies on

- **VELO**: clustering, tracking, vertex reconstruction
- **UT**: tracking, momentum estimate, fake rejection
- **SciFi**: tracking, momentum measurement

+
PID from **MUON** and **CALO** systems

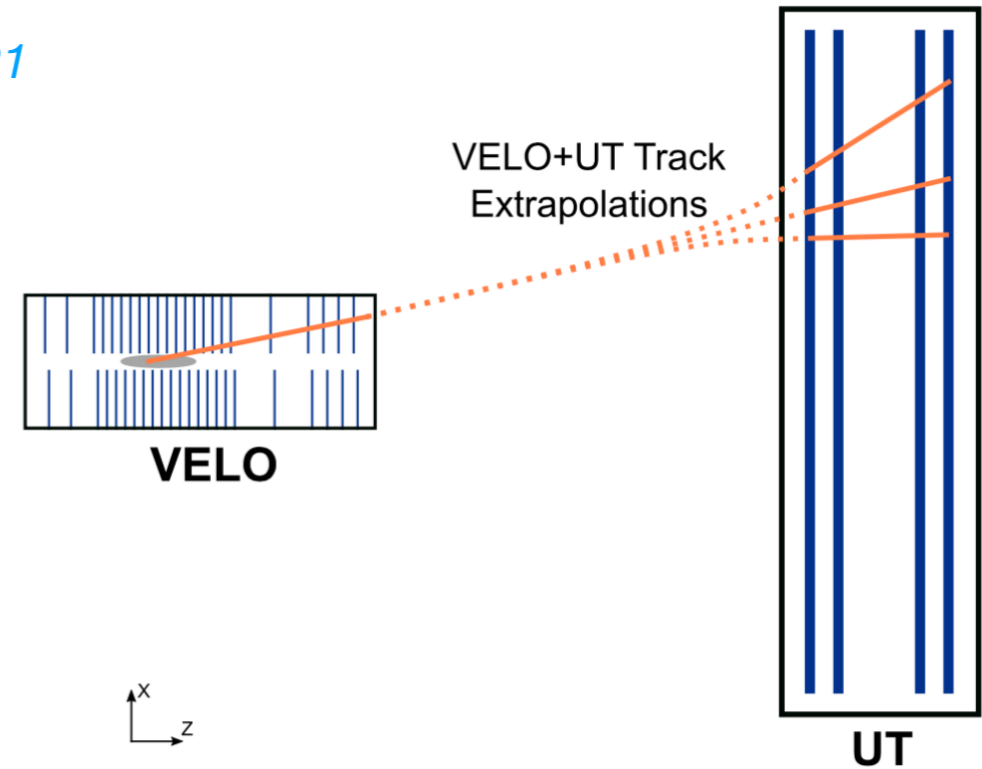


Track reconstruction

Journal of Computational Science, vol. 54, 2021

Velo tracking

- 26 silicon pixel modules with $\sigma_{x,y} \sim 5 \mu m$
- Local paralleled clustering algorithm (Search by Triplet)
- Tracks fitted with simple Kalman filter assuming straight line model

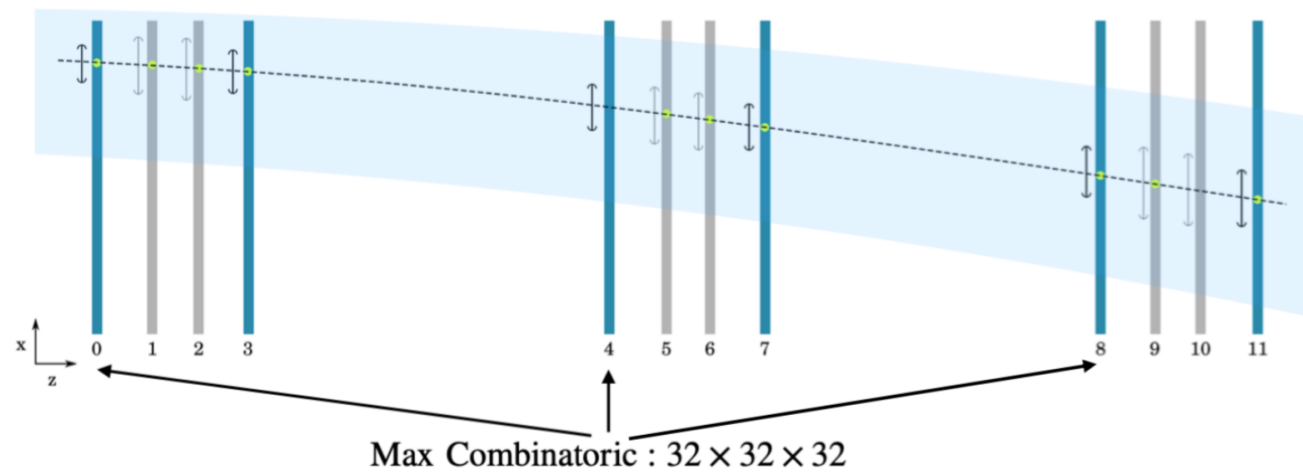


IEEE Access, vol. 7, pp. 91612-91626, 2019

UT tracking

- 4 layers of silicon strips
- Velo tracks extrapolated to UT taking into account B field
- Parallelized trackless finding inside search window requiring at least 3 hits

Comput. Softw. Big Sci 4, 7 (2020)

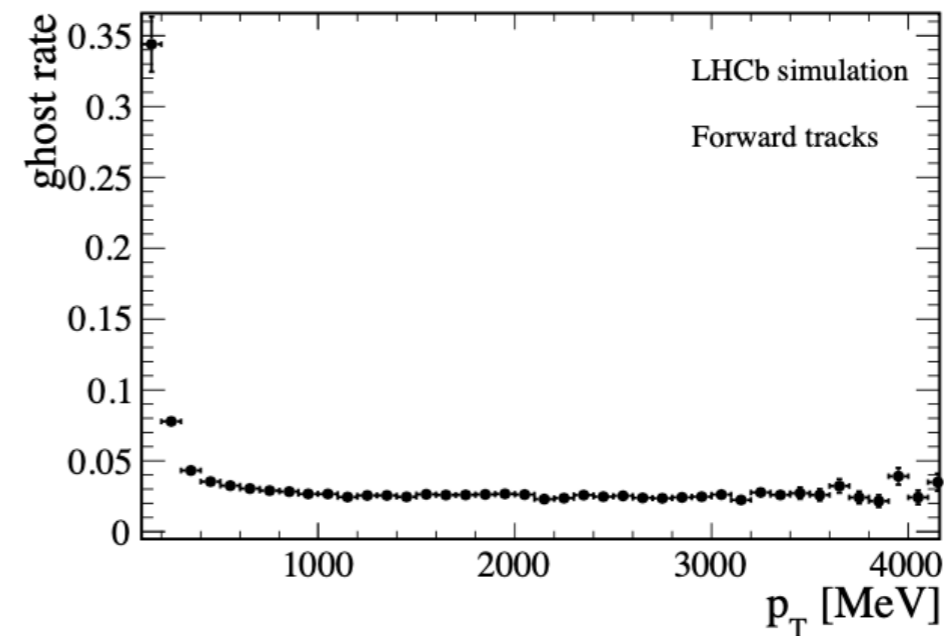
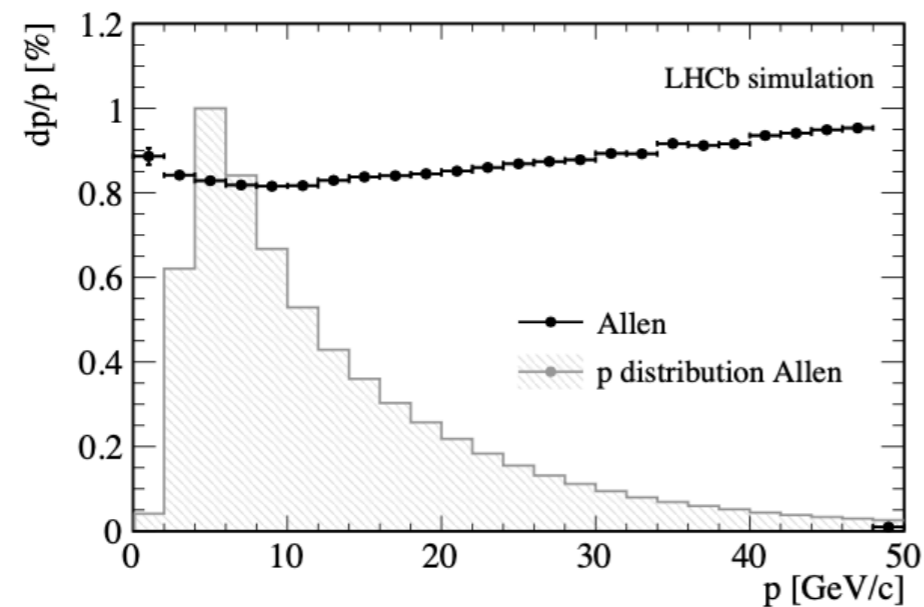
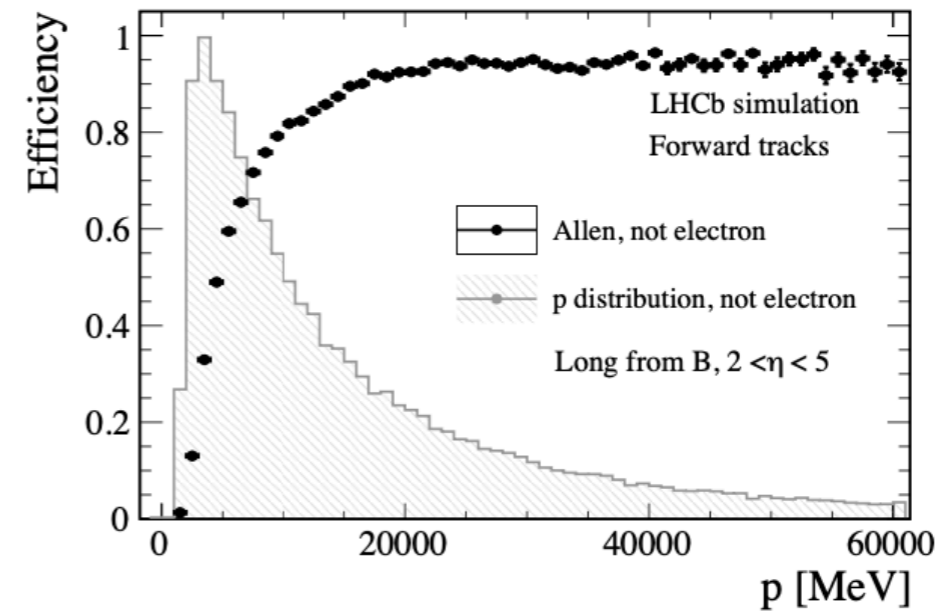
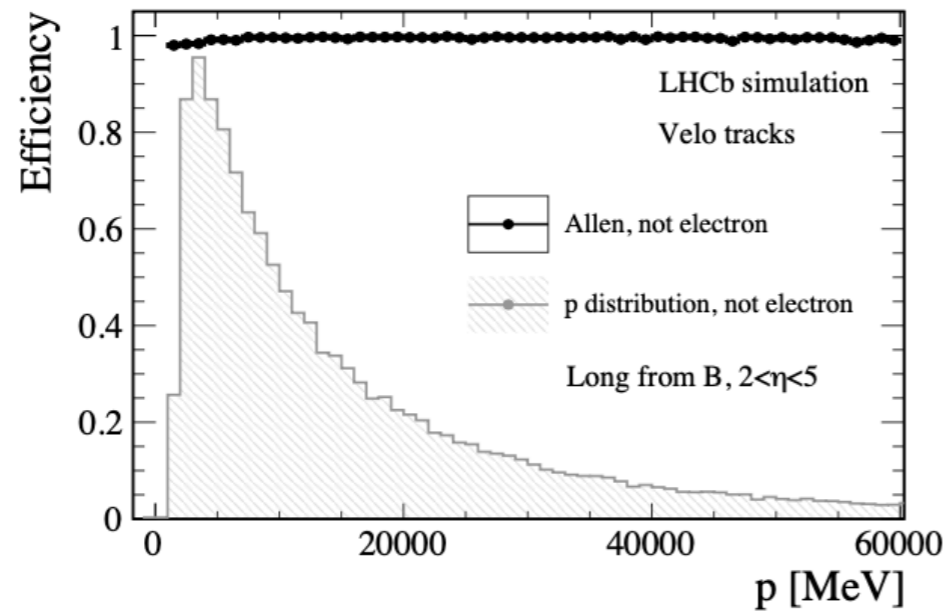


SciFi tracking

- 3 stations with 4 layers of Scintillating Fibres
- Velo-UT tracks extrapolated using parametrisation
- Parallelized Forward algorithm to reconstruct long tracks
 - Search windows from Velo-UT momentum estimate
 - From triplets and extend to remaining layers

Tracking performance

- **Run 2 performance maintained at x5 instantaneous luminosity**
- Excellent track reconstruction efficiency (> 99% for VELO, 95% for high-p forward tracks)
- Good momentum resolution and fake rejection



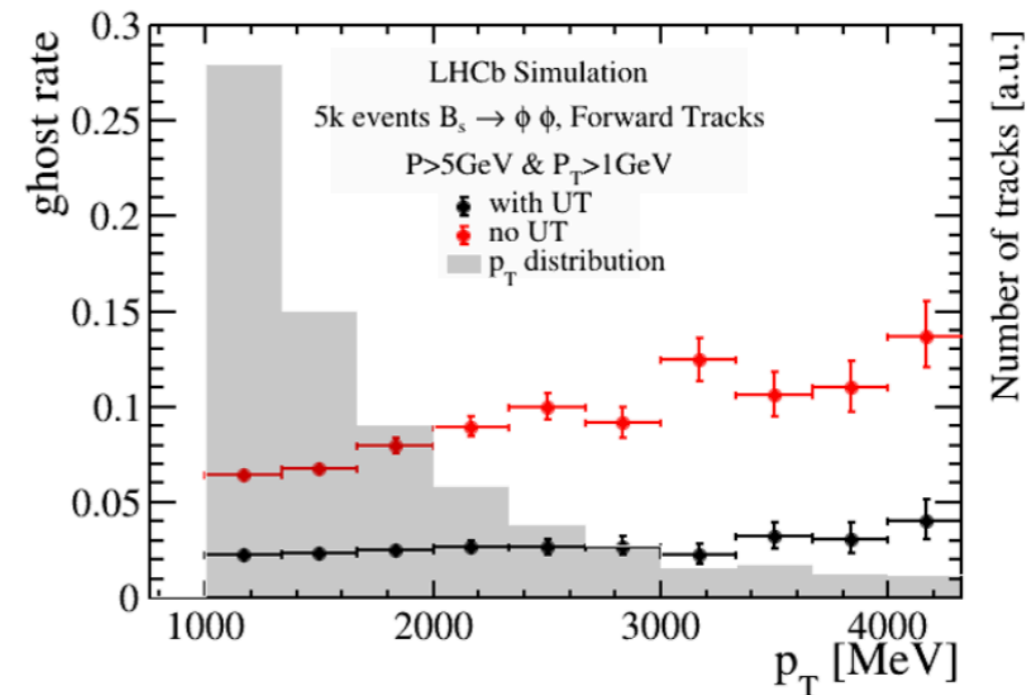
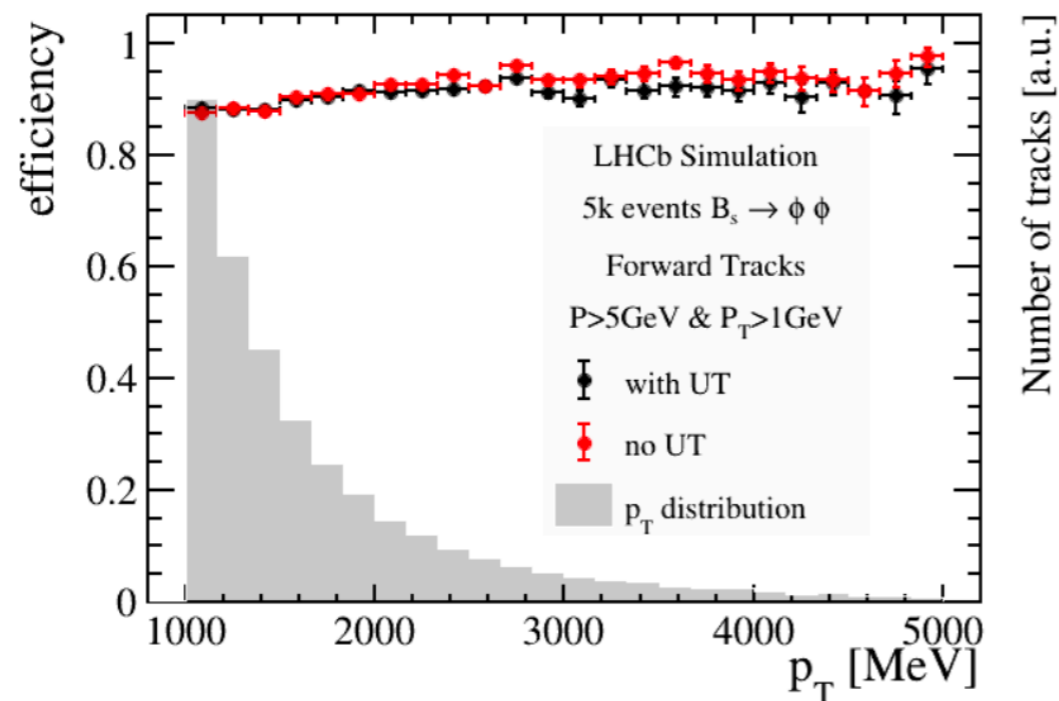
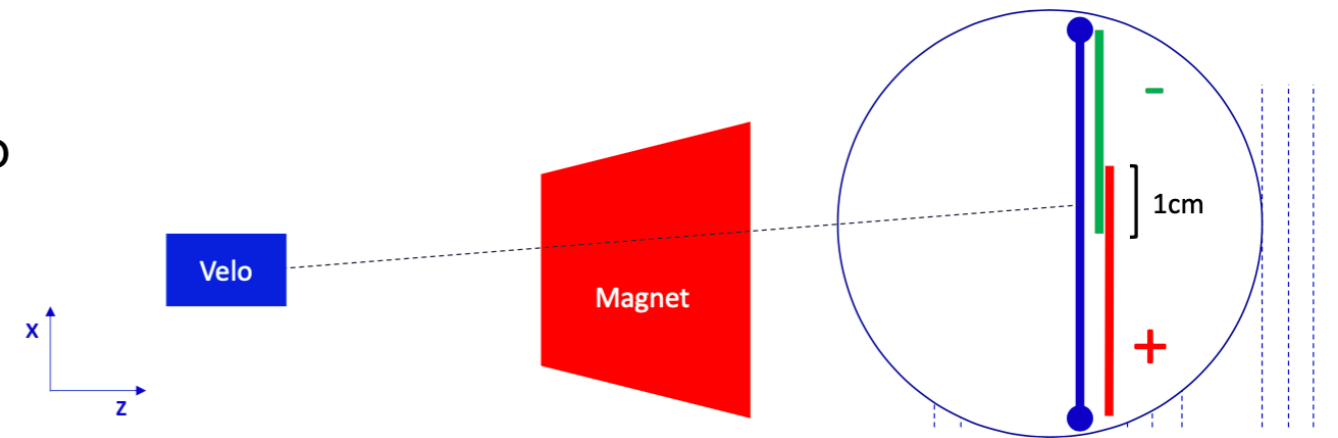
LHCb-FIGURE-2020-014

Tracking without the UT

- In 2022, the UT detector is unfortunately not be available for data-taking
- Tracking performance and throughput maintained, at the cost of larger fake rate
- Commissioning two options, which **both maintain the current throughput**

1. Forward without UT

- Extrapolate VELO track as a straight line, make two windows — assuming positive/negative charge
- Assume $p > 5 \text{ GeV}$, $p_T > 1 \text{ GeV}$ (low-p tracks get bent out of the SciFi acceptance anyway)



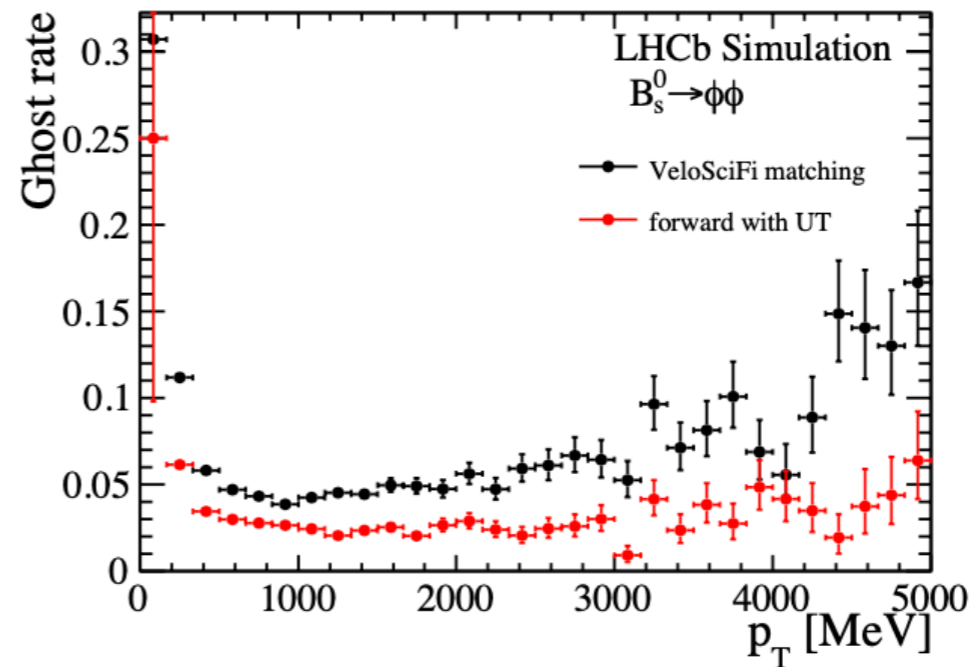
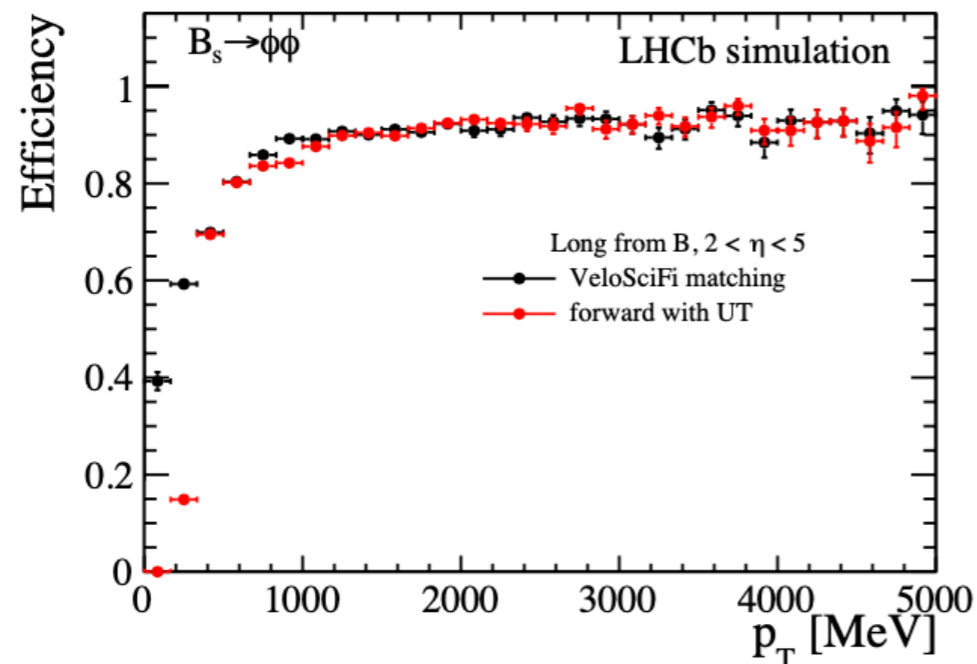
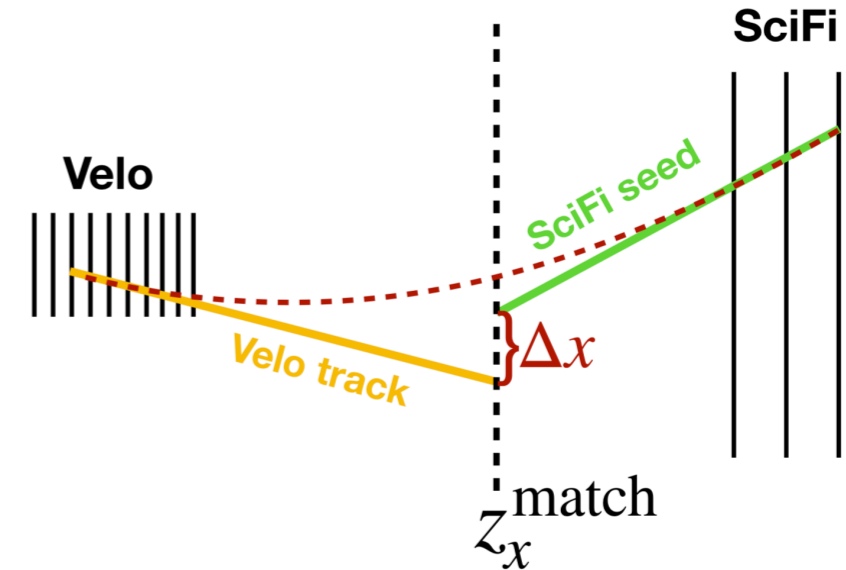
LHCB-FIGURE-2022-007

Tracking without the UT

- In 2022, the UT detector is unfortunately not be available for data-taking
- Tracking performance and throughput maintained, at the cost of larger fake rate
- Commissioning two options, which **both maintain the current throughput**

2. Seeding+matching

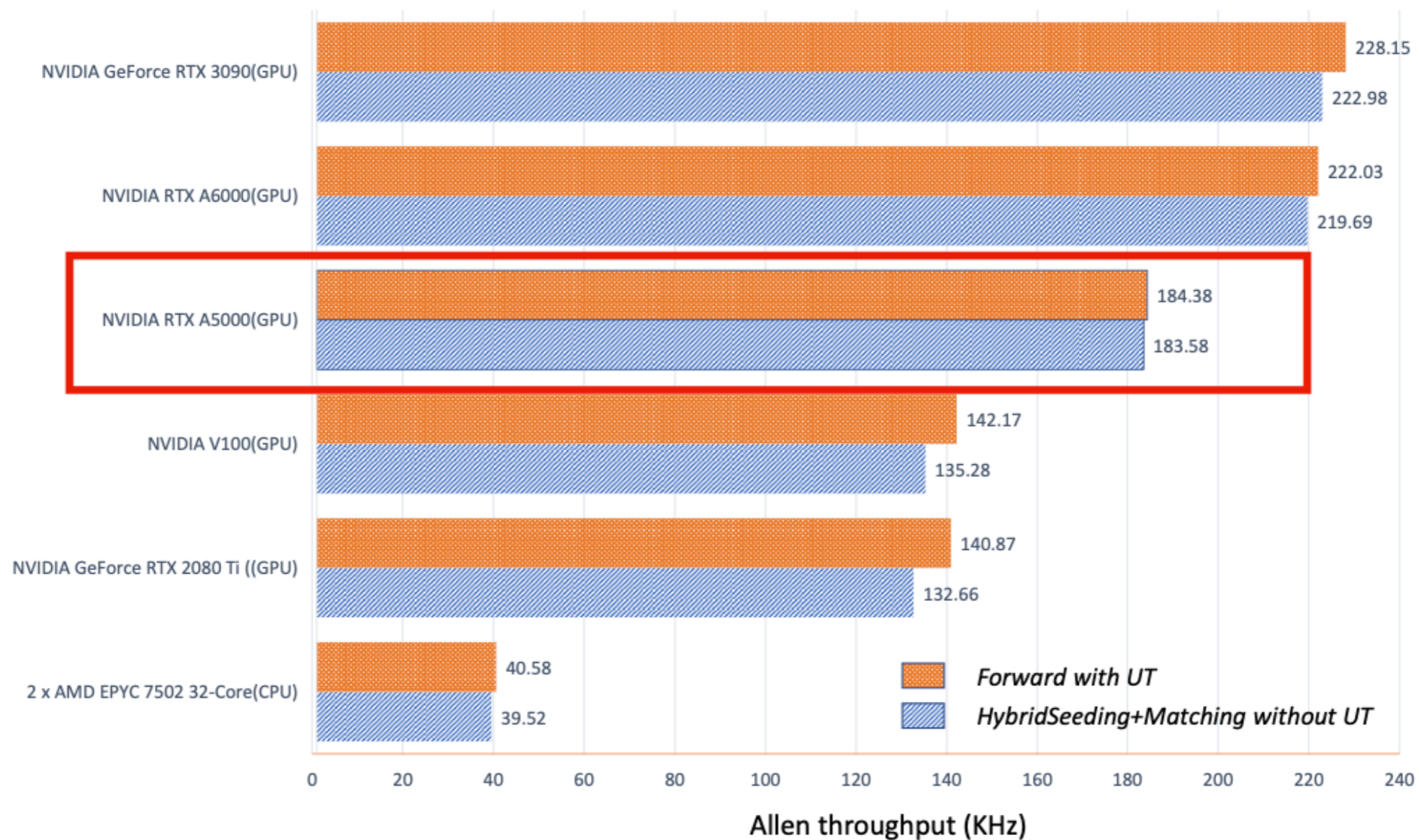
- Standalone SciFi reconstruction & matching to VELO seeds
- Highly efficient for low momenta
- Opens the door to additional physics cases in HLT1 (downstream and SciFi tracks)



LHCb-FIGURE-2022-010

Tracking without the UT

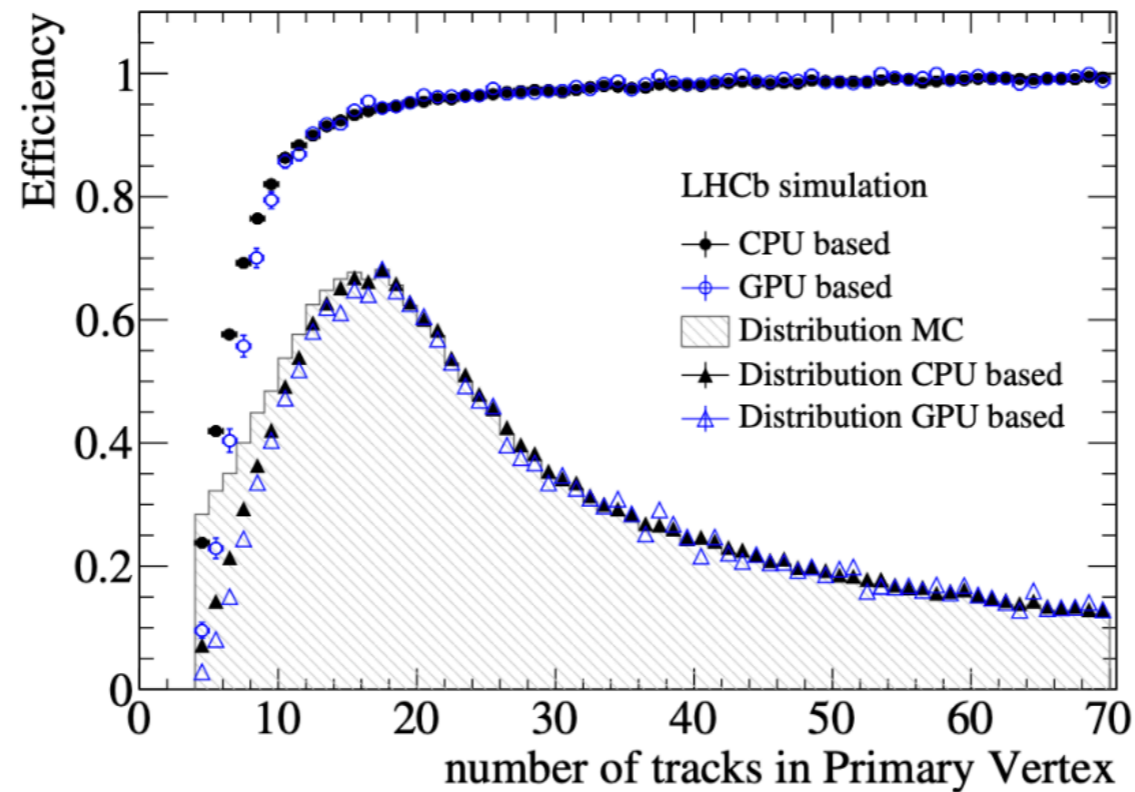
- In 2022, the UT detector will unfortunately not be available for data-taking
- Tracking performance and throughput maintained, at the cost of larger fake rate
- Opportunity to commission 2 options, which **both maintain the current throughput**



LHCB-FIGURE-2022-010

Vertex reconstruction

- Primary vertices found from **clusters** in the closest approach of tracks to the beamline
- 1-1 mapping between tracks and vertices requires **serialization**
 - Instead, every track assigned to every vertex based on **weight**
- **Efficiency > 90%** for vertices with number tracks > 10

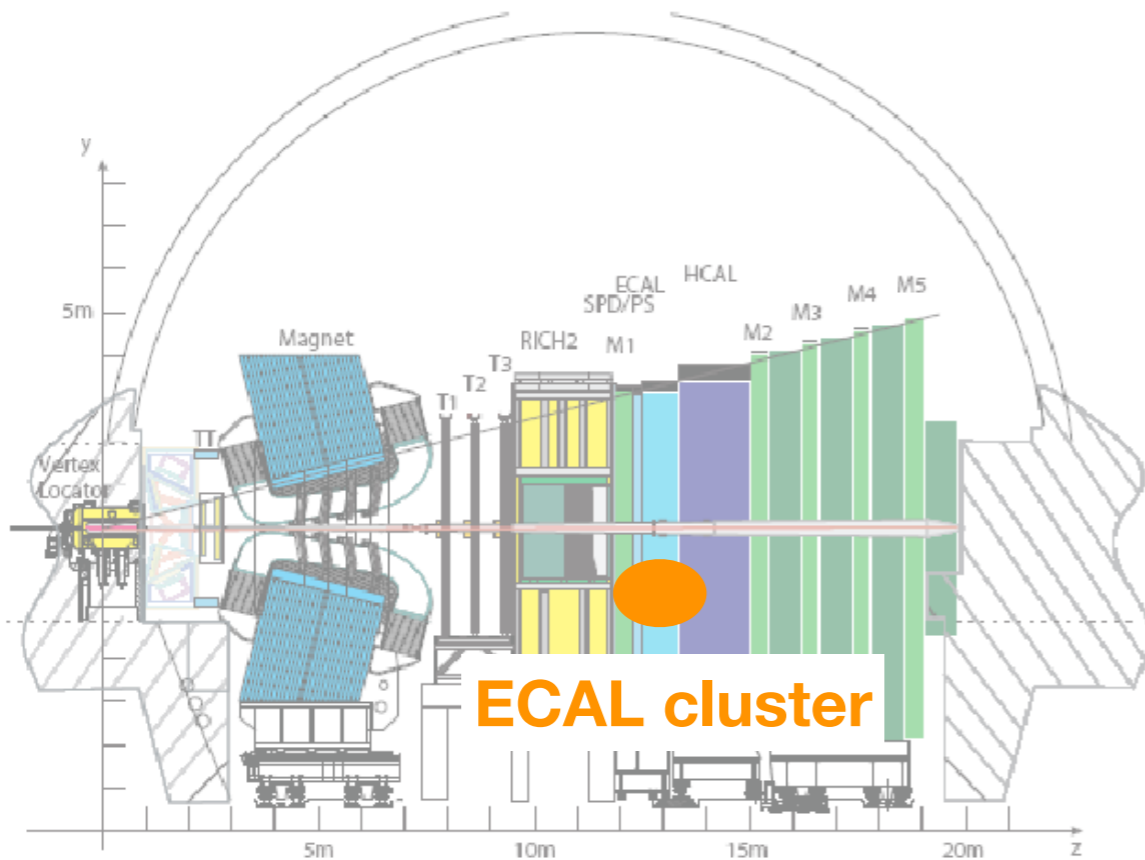
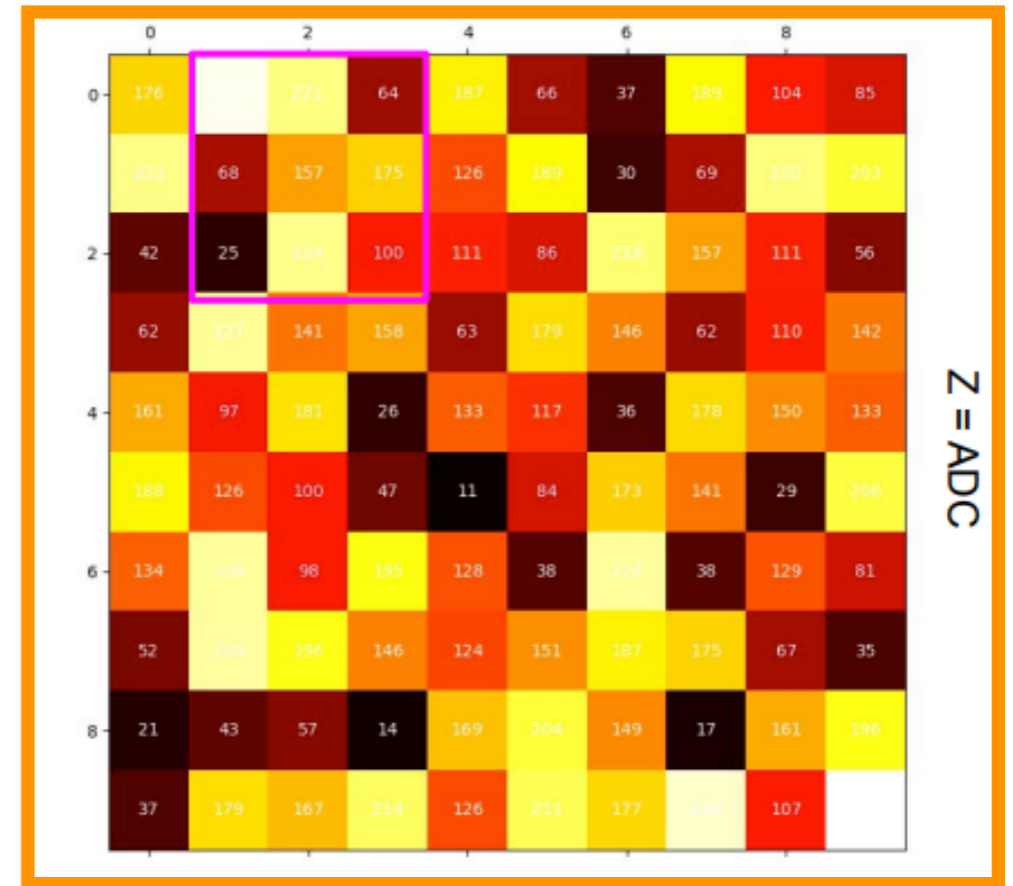


Comput. Softw. Big Sci. 6 (2022) no.1, 1

Calorimeter and muon PID

• CALO reconstruction

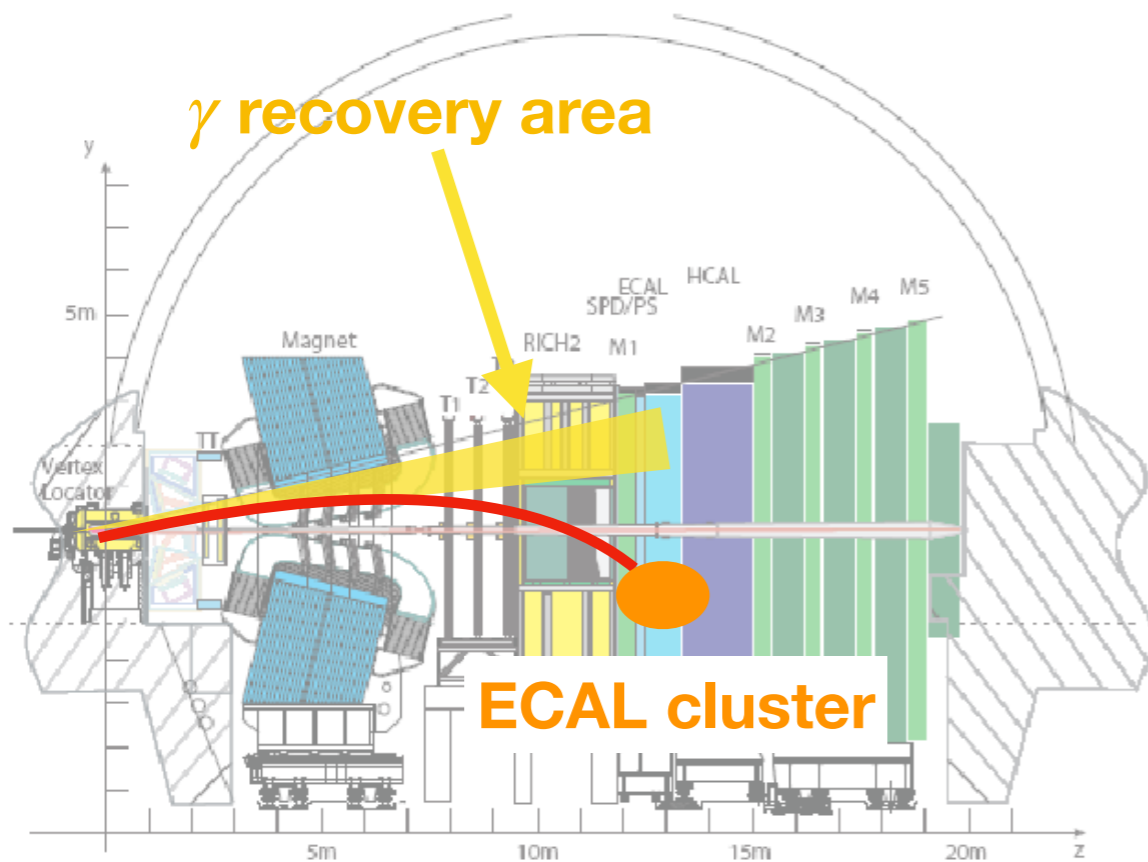
- Loop over calorimeter cells and look for energetic clusters
- Originally not foreseen within the baseline TDR, but outcome of ambition and good design (and lots of optimisation)
- The very first algorithm that was tested with real Run 3 data!



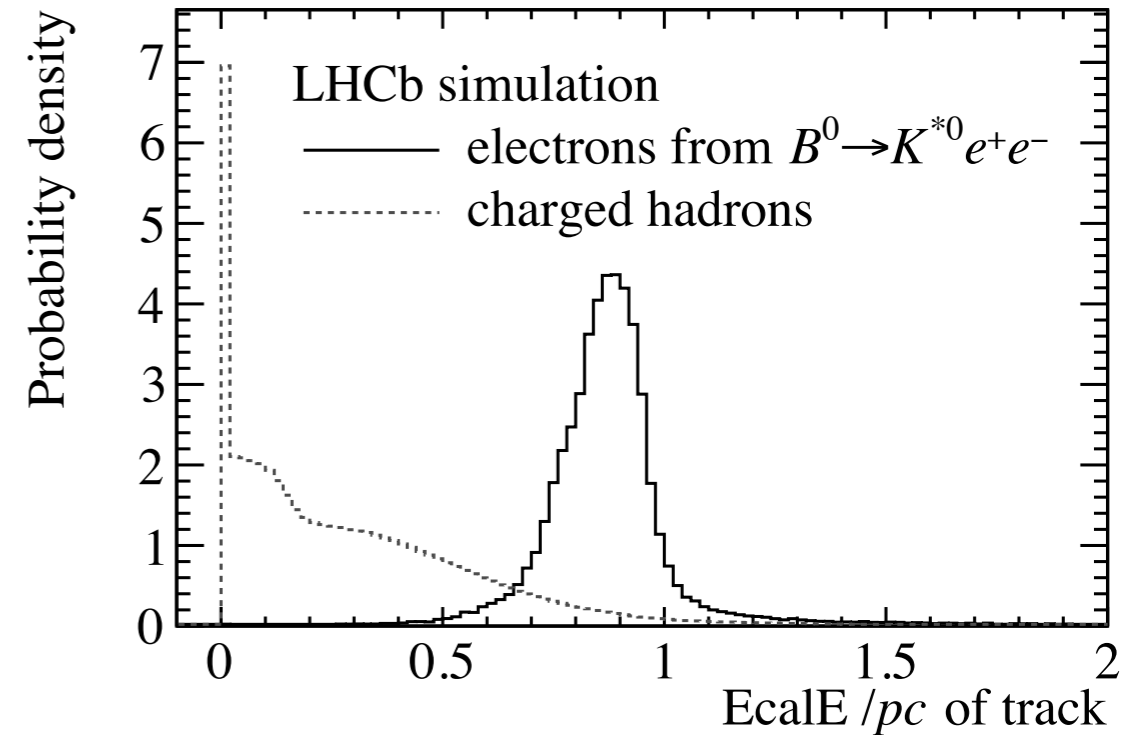
Calorimeter reconstruction

• CALO reconstruction

- Loop over calorimeter cells and look for energetic clusters
- Originally not foreseen within the baseline TDR, but outcome of ambition and good design (and lots of optimisation)
- The very first algorithm that was tested with real Run 3 data!



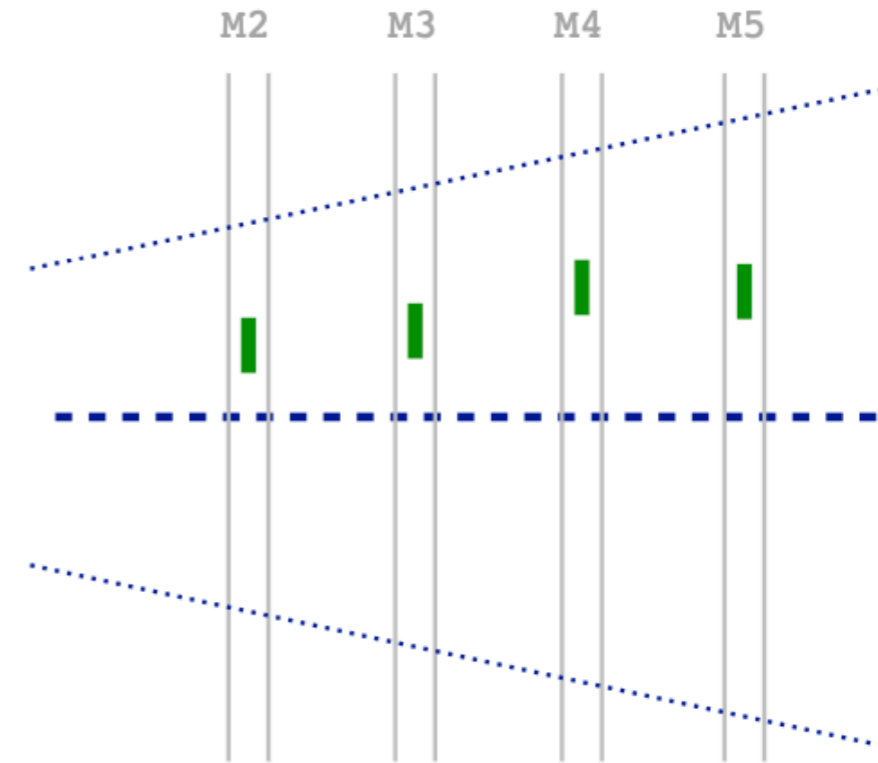
LHCb-FIGURE-2021-003



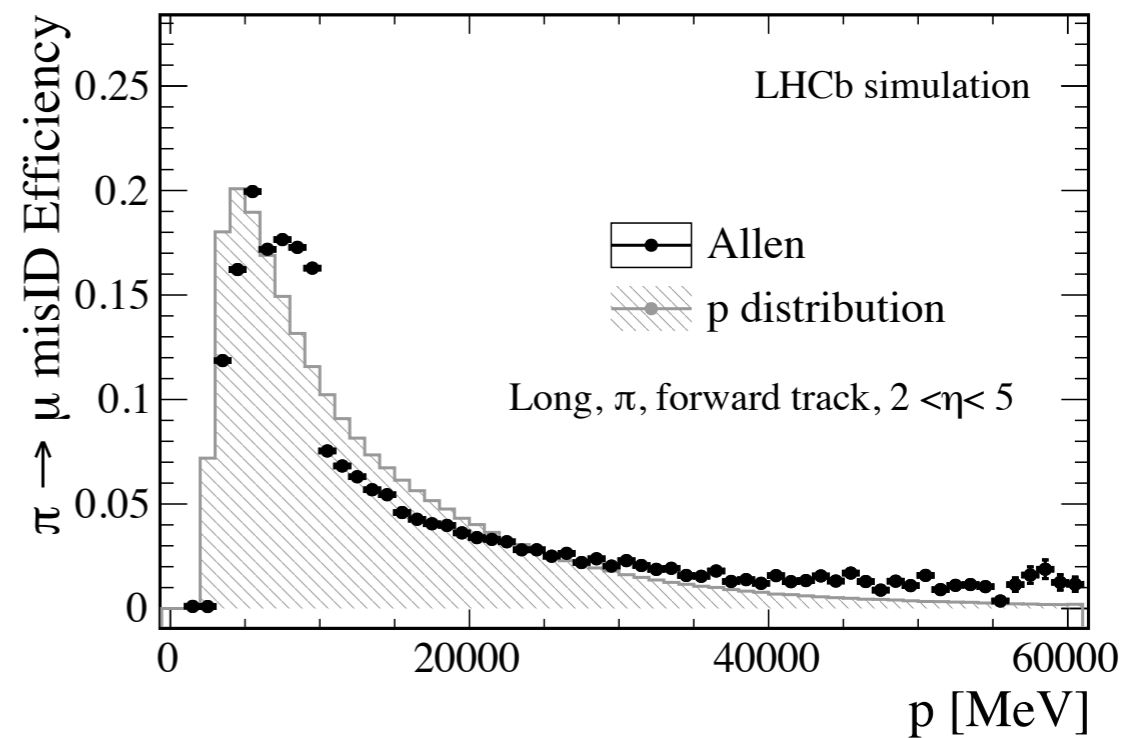
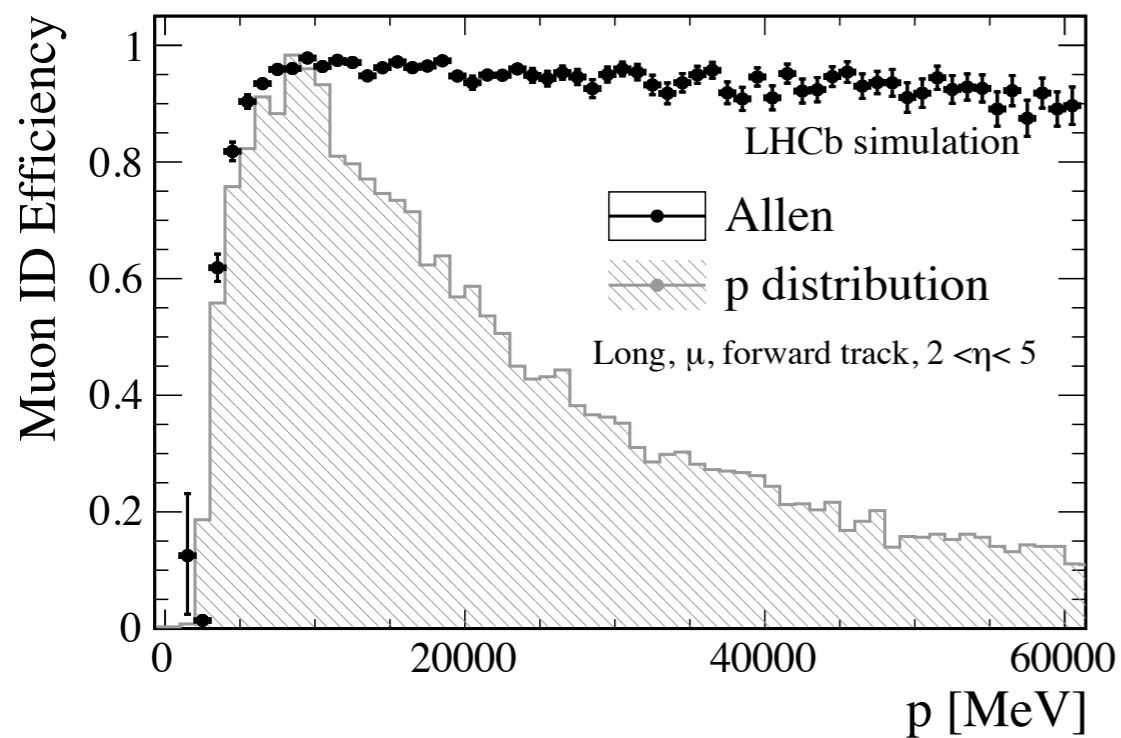
- Calo digits attached to long tracks for electrons
- Momentum is corrected if clusters are found in the Bremsstrahlung recovery area

Muon PID

- **Muon particle identification**
 - Extrapolate tracks from SciFi to Muon stations
 - Match hits to tracks in a field of interest
 - Excellent muon identification and misID background rejection

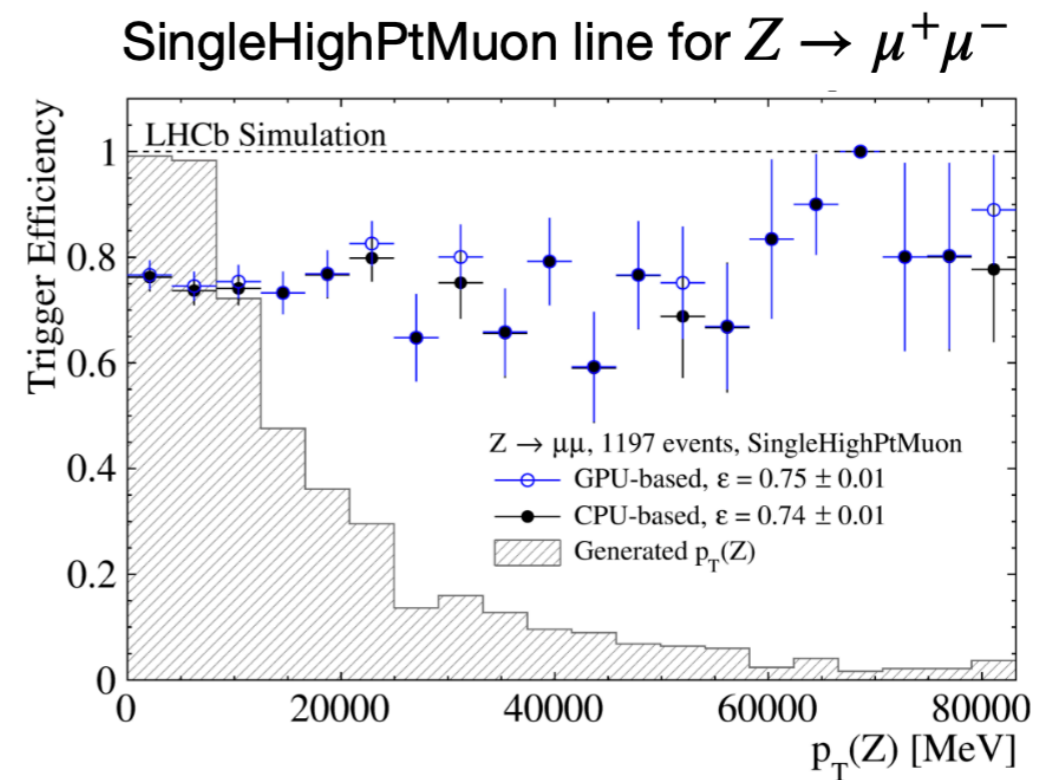
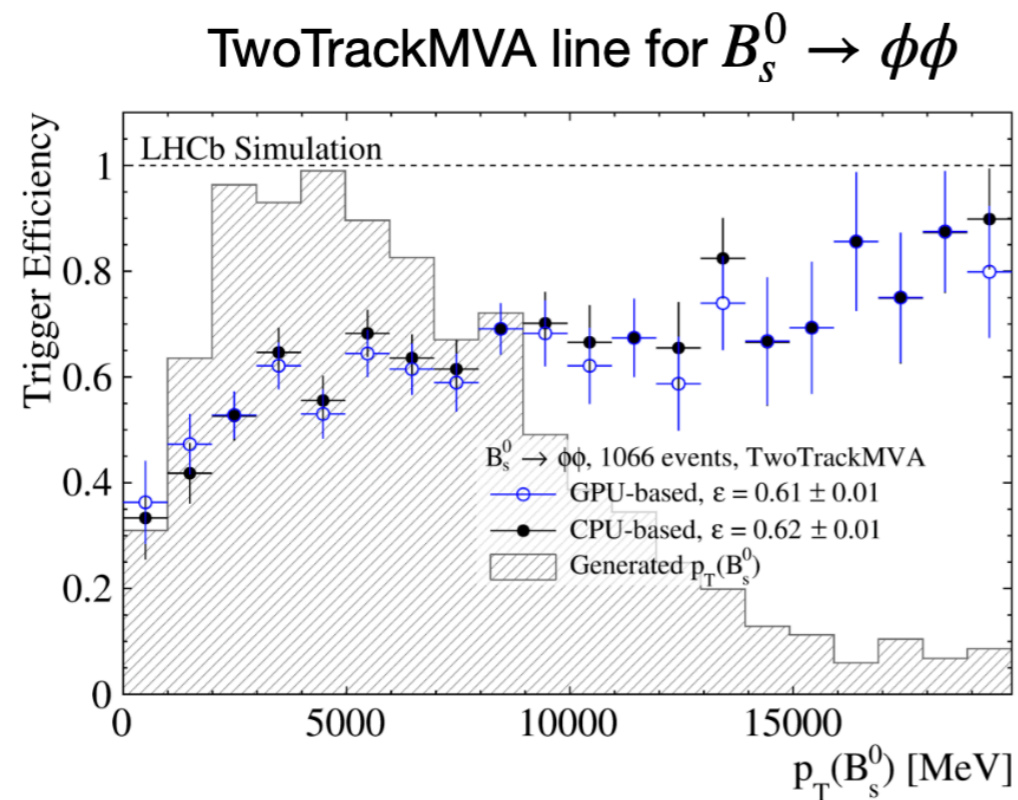


LHCB-FIGURE-2020-014



HLT1 selection performance

- Inclusive rate for the main HLT1 lines ~ 1 MHz
- O(30) lines implemented so far:
 - Cover majority of LHCb physics programme (B, D decays, semileptonic, EW physics)
 - Special lines for monitoring, alignment and calibration



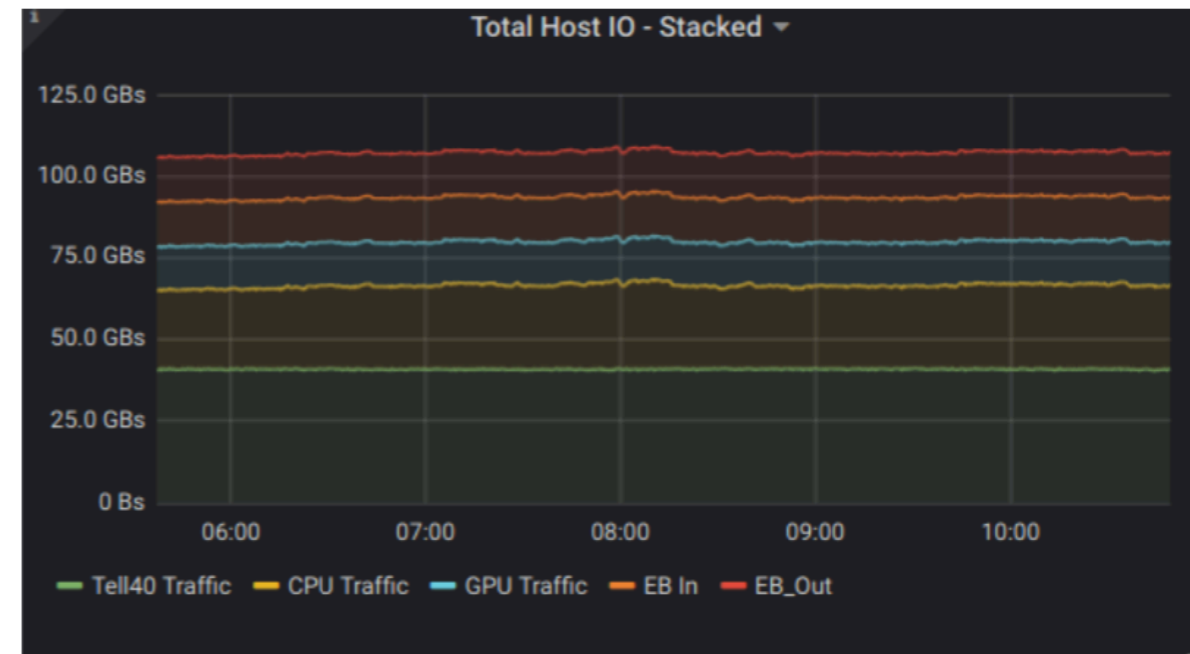
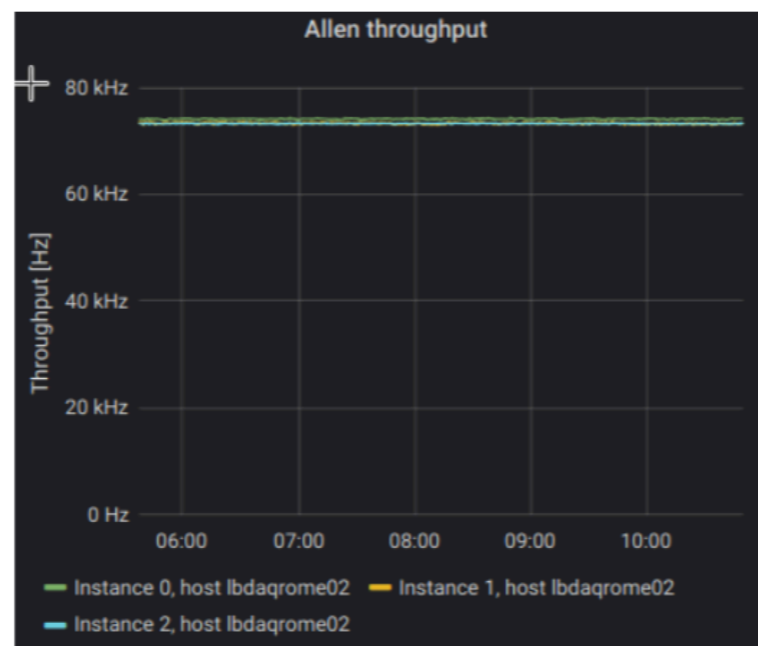
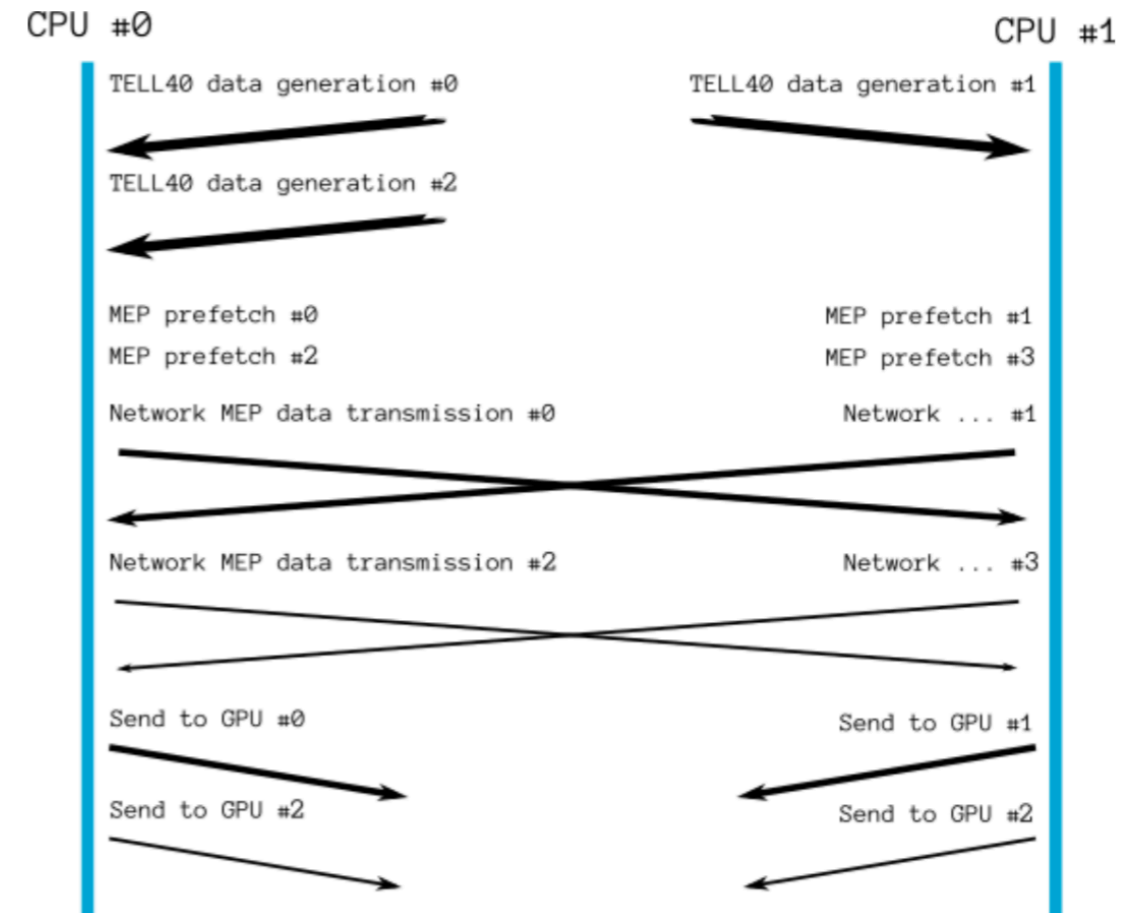
Comput. Softw. Big Sci. 6 (2022) no.1, 1

Allen commissioning without real data

Challenge of fully commissioning Allen: the real detectors and EB server are needed!

First integration tests in smaller-size servers with pre-loaded simulation data done in the past

- Emulate network traffic and memory pressure with mock-up data from FPGAs
- Stable throughput at 70 kHz
- I/O memory bandwidth stable and within limits
- Cooling and memory usage requirements met
- Proof of principle!



Tests in 2019
A lot of performance
improvements since!

Allen commissioning with data

- LHCb has been exercising its DAQ in parallel to the LHC commissioning
- Sub-set of detectors (Calorimeters, Muon stations, PLUME) already in the global partition of the Experiment Control System (ECS)
- System running 24/7 in parallel to sub-detector commissioning activities

The screenshot displays the LHCb TOP control interface. At the top, the system is labeled 'LHCb' and is in a 'RUNNING' state. The 'Auto Pilot' is set to 'OFF'. The date and time are 'Fri 01-Jul-2022 10:51:04'. Below this, a table lists sub-systems and their states: DCS (READY), DAI (READY), DAQ (RUNNING), RunInfo (RUNNING), TFC (RUNNING), EB (RUNNING), and Monitoring (RUNNING). The 'Run Info' section shows 'Run Number: 235723', 'Run Start Time: 01-Jul-2022 10:46:35', 'Run Duration: 000:04:25', 'Nr. Events: 5380640838', and 'Step Nr: To Go: 0 0'. It also displays 'Input Rate: 21594.99 kHz' and 'Output Rate: 280.28 kHz'. The 'Data Destination' is 'EOS' and 'Data Type' is 'COLLISION22'. The 'Sub-Detectors' section shows the status of various detectors: TDET (ERROR), VELOA (RUNNING), VELOC (RUNNING), UTC (OT_ALLOCATE), SFA (ACTIVE), SFC (READY), RICH1 (READY), RICH2 (READY), ECAL (RUNNING), and HCAL (RUNNING). The 'Messages' section at the bottom shows log entries: '01-Jul-2022 10:46:35 - LHCb executing action GO', '01-Jul-2022 10:46:36 - LHCb_TFC executing action START_TRIGGER', and '01-Jul-2022 10:46:36 - LHCb in state RUNNING'.

Allen commissioning with data

The screenshot shows the LHCb TOP control interface. The main window is titled "LHCb: TOP" and displays the system status as "RUNNING". The "Auto Pilot" is set to "OFF". The "Run Info" section shows the Run Number as 235723, Run Start Time as 01-Jul-2022 10:46:35, and Run Duration as 000:04:25. The "Nr. Events" is 5380640838. The "Input Rate" is 21594.99 kHz and the "Output Rate" is 280.28 kHz. The "Data Destination" is set to "EOS" and the "Data Type" is "COLLISION22". The "File" path is "/hlt2/objects/LHCb/0000235723".

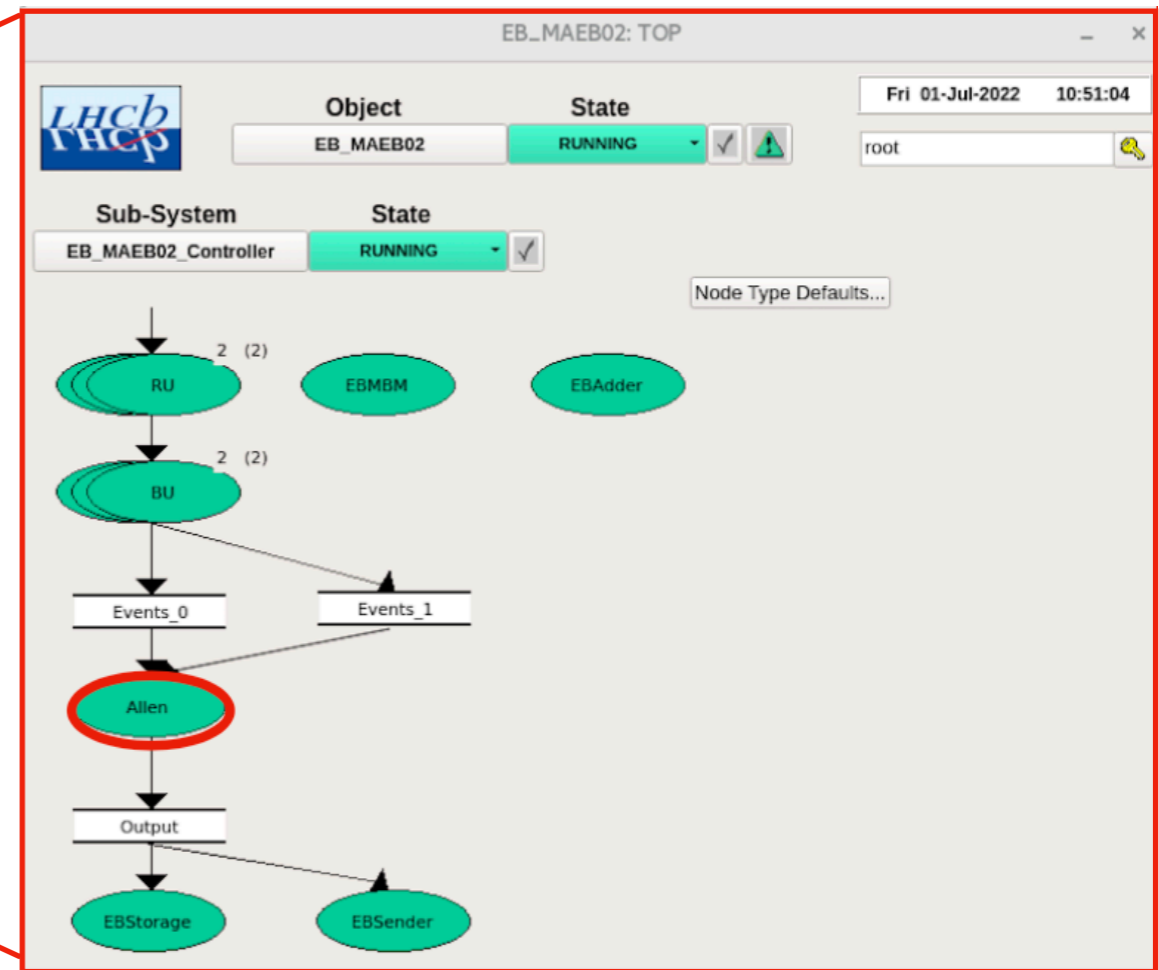
Sub-System	State
DCS	READY
DAI	READY
DAQ	RUNNING
RunInfo	RUNNING
TFC	RUNNING
EB	RUNNING
Monitoring	RUNNING

Sub-Detectors:

Sub-Detector	State
TDET	ERROR
VELOA	RUNNING
VELOC	RUNNING
UTC	OT_ALLOCATE
SFA	ACTIVE
SFC	READY
RICH1	READY
RICH2	READY
ECAL	RUNNING
HCAL	RUNNING
MUONA	RUNNING
MUONC	RUNNING
PLUME	RUNNING

Messages:

- 01-Jul-2022 10:46:35 - LHCb executing action GO
- 01-Jul-2022 10:46:36 - LHCb_TFC executing action START_TRIGGER
- 01-Jul-2022 10:46:36 - LHCb in state RUNNING



- ~200 GPUs are installed in the EB
- HLT1 is included in the global partition

Allen commissioning with data

The screenshot displays the LHCb TOP control interface. At the top, the system is identified as 'LHCb' and is in a 'RUNNING' state. The 'Auto Pilot' is set to 'OFF'. The date and time are 'Sat 01-Oct-2022 17:54:32', and the user is 'root'. A table of sub-systems shows their states: DCS (READY), DAI (READY), DAQ (RUNNING), RunInfo (RUNNING), TFC (RUNNING), EB (RUNNING), and Monitoring (RUNNING). The 'Run Info' section shows Run Number 247579, Run Start Time 01-Oct-2022 17:53:50, Run Duration 000:00:39, and Nr. Events 738'289'366. The 'Trigger Config' is set to 'hit1_pp_no_gec_no_ut_veloSI'. Two gauges show the input rate at 18160.72 kHz and the output rate at 838.34 kHz. The 'Sub-Detectors' section shows the status of various detectors: TDET (NOT_READY), VELOA (RUNNING), VELOC (RUNNING), UTC (NOT_READY), SFA (RUNNING), SFC (RUNNING), RICH1 (RUNNING), RICH2 (RUNNING), ECAL (RUNNING), HCAL (RUNNING), MUONA (RUNNING), MUONC (RUNNING), and PLUME (RUNNING). The 'Messages' section at the bottom shows the system executing 'GO' and 'START_TRIGGER' actions.

Sub-System	State
DCS	READY
DAI	READY
DAQ	RUNNING
RunInfo	RUNNING
TFC	RUNNING
EB	RUNNING
Monitoring	RUNNING

Sub-Detectors	State
TDET	NOT_READY
VELOA	RUNNING
VELOC	RUNNING
UTC	NOT_READY
SFA	RUNNING
SFC	RUNNING
RICH1	RUNNING
RICH2	RUNNING
ECAL	RUNNING
HCAL	RUNNING
MUONA	RUNNING
MUONC	RUNNING
PLUME	RUNNING

Full HLT1 tracking sequences

Input rate ~20 MHz
(max with current filling scheme)

All installed subdetectors included

Allen commissioning with data

The screenshot displays the LHCb TOP control interface. At the top, the system is labeled 'LHCb' and 'LHCb TOP'. The main status bar shows 'System: LHCb', 'State: RUNNING', and 'Auto Pilot: OFF'. The date and time are 'Sat 01-Oct-2022 17:54:32'. Below this, a table lists sub-systems and their states: DCS (READY), DAI, DAQ, Runinfo, TFC, EB, and Monitoring. A 'Run Info' section shows 'Run Number:' and 'Activity:'. A large 3D model of the LHCb detector is centered, showing various components like the VELO, TPC, and calorimeters. The interface also includes sections for 'Alignment & Calibration', 'HLT2' (with a pie chart and 'Disk Usage: 48%'), 'Efficiency', 'TFC Control', and 'Sub-Detectors' (listing TDET and VELOA). At the bottom, there are 'Messages' and a 'Close' button.

Next steps:

- Currently validating the tracking sequences
- Take some good quality data for physics until the end of this year run!

Conclusion

- LHCb is currently undergoing its first major upgrade in order to increase its instantaneous luminosity by x5
- Major changes in the trigger strategy:
 - Remove L0 hardware trigger, read-out full detector at 30 MHz
 - New first level trigger run on GPUs
- Partial event reconstruction and trigger selection lines implemented with excellent physics performance expected
- The system can be realised with ~200 GPUs (throughput ~170 kHz)
- GPUs are installed in the EB server and the commissioning is ongoing
- LHCb is almost ready to start collecting physics data with the brand new detector

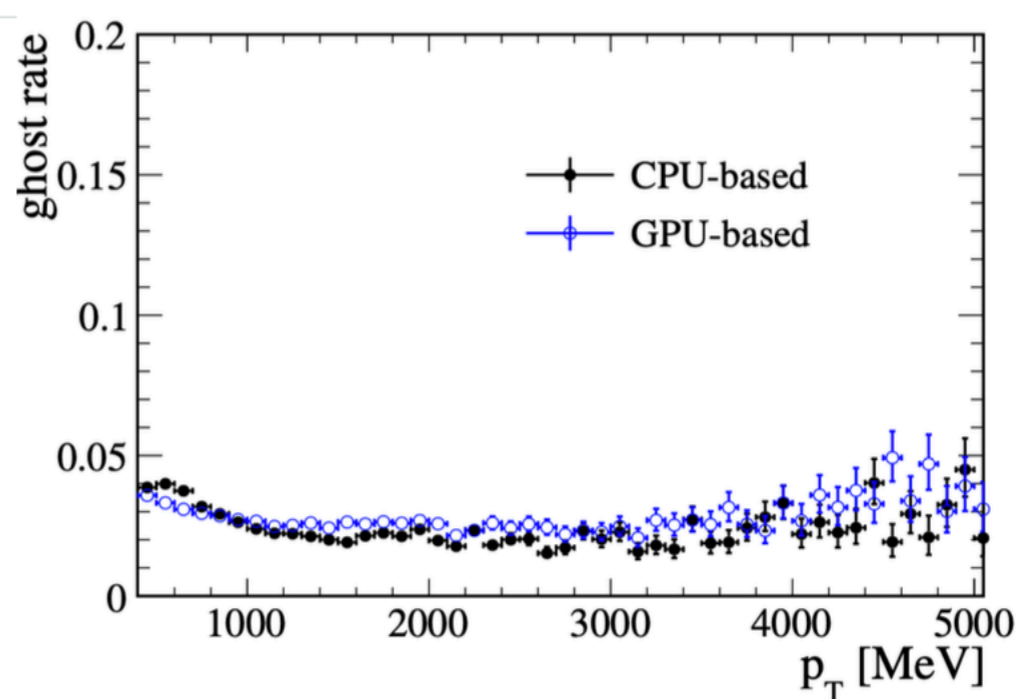
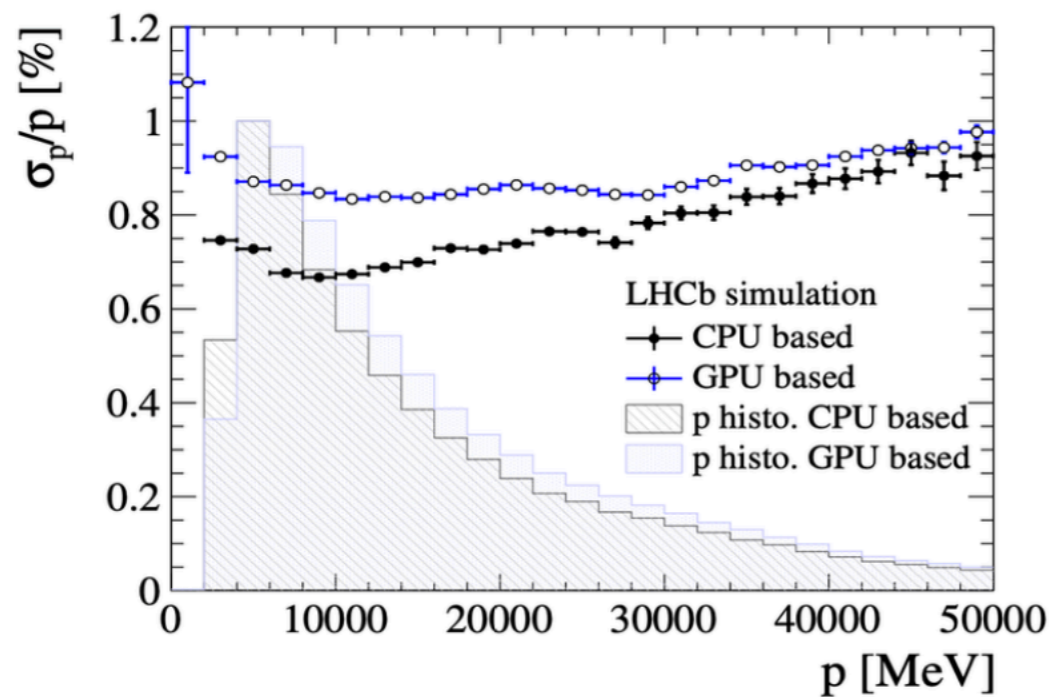
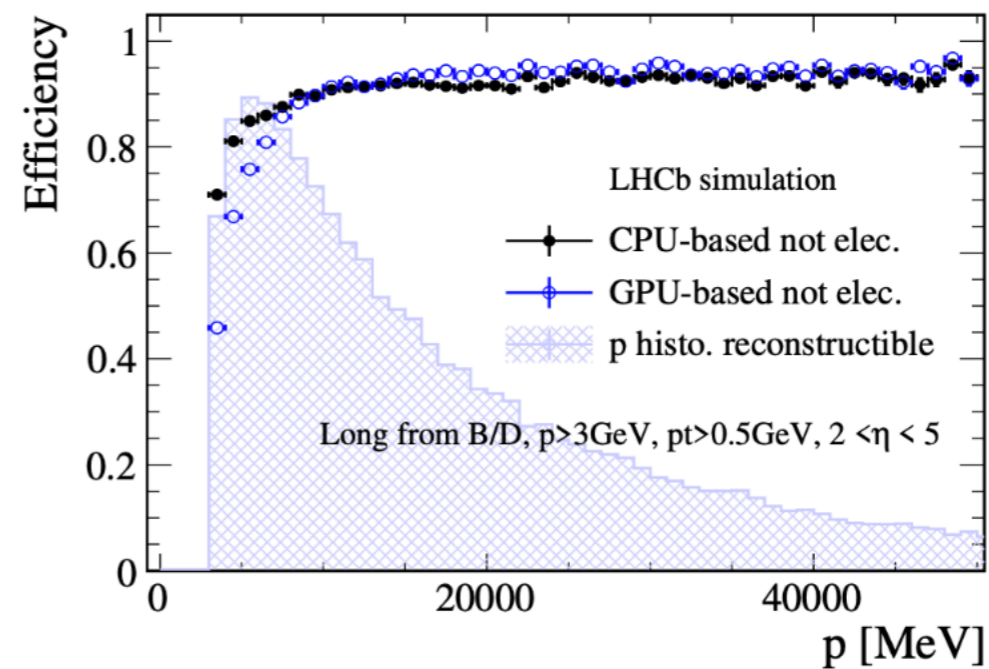
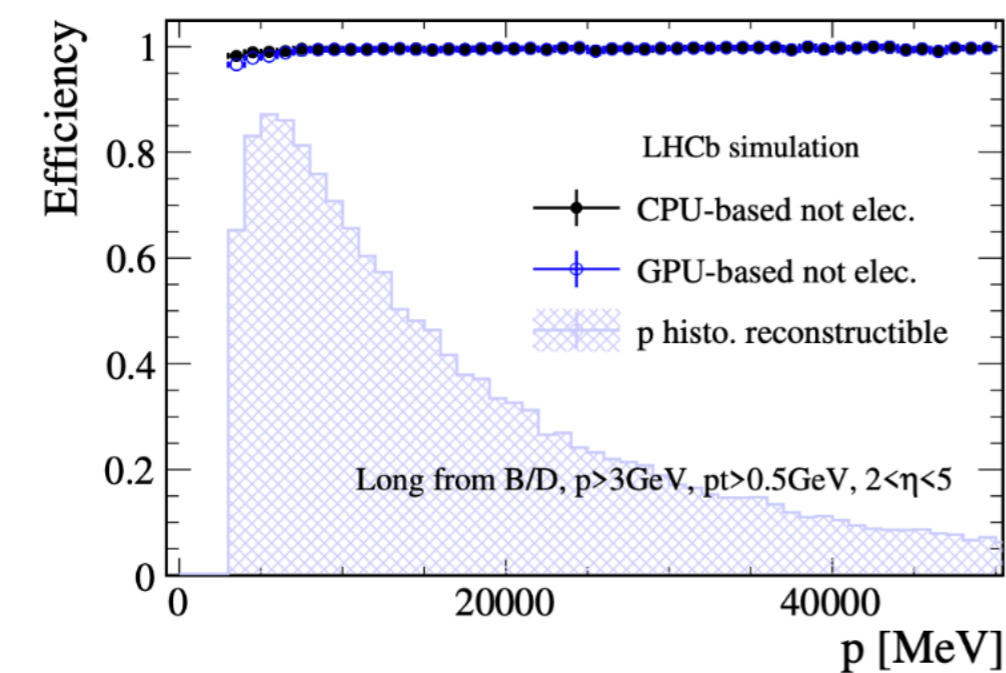
Stay tuned!



Backup

HLT1 CPU/GPU comparison

Compatible performance between CPU and GPU!



[Comput. Softw. Big Sci. 6 \(2022\) no.1, 1](#)

HLT1 CPU/GPU comparison

- Compatible performance between CPU and GPU [Comput. Softw. Big Sci. 6 \(2022\) no.1, 1](#)

