ICHEP 2022 BOLOGNA

ICHEP 2022
XLI
International Conference on High Energy Physics
Bologna (Italy)

6
13 07 2022

LHCb HLT1: Tracking and vertexing at 30MHz with GPUs

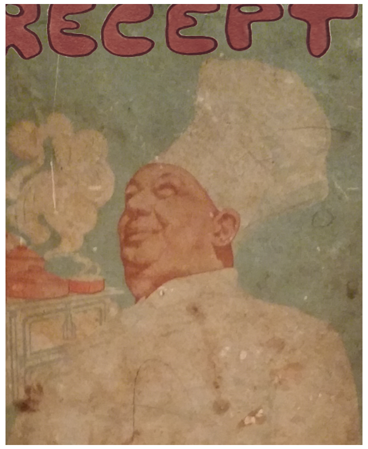
Christina Agapopoulou on behalf of the LHCb collaboration

ICHEP 2022 - 09/07/2022

cnrs
LPNHE PARIS
SORBONNE UNIVERSITÉ

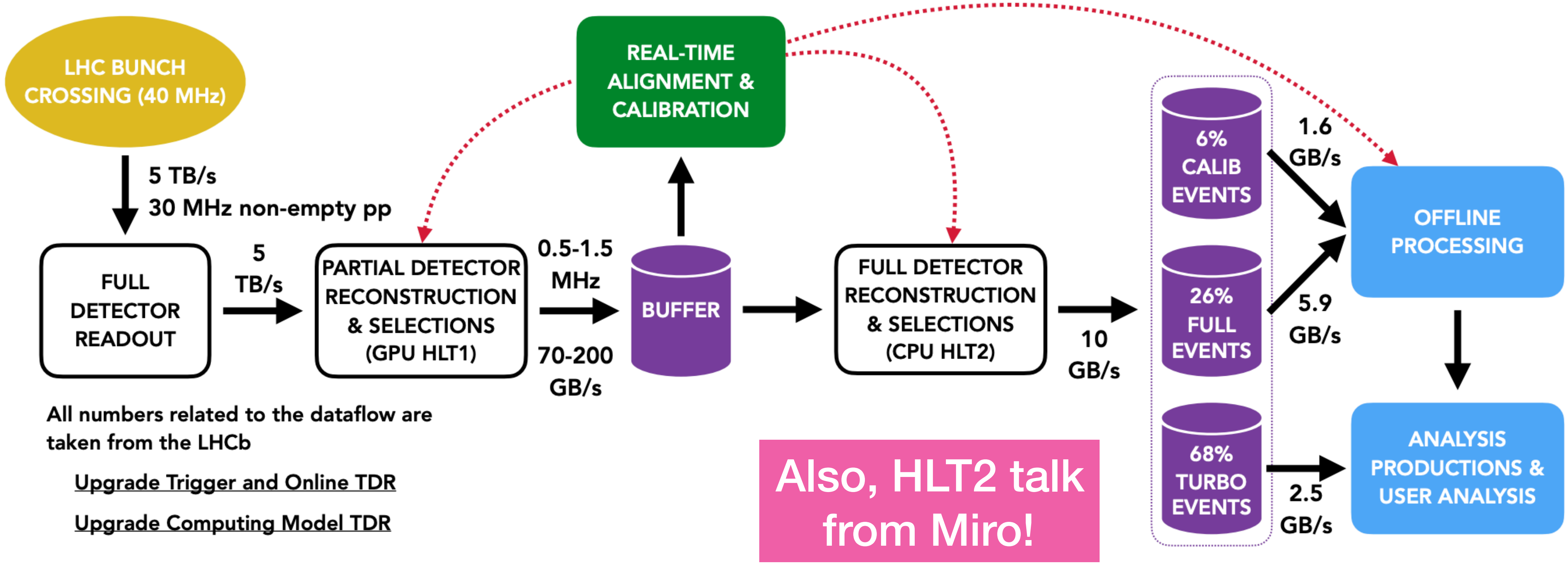
LHCb

erc
European Research Council
Established by the European Commission



The LHCb data-flow

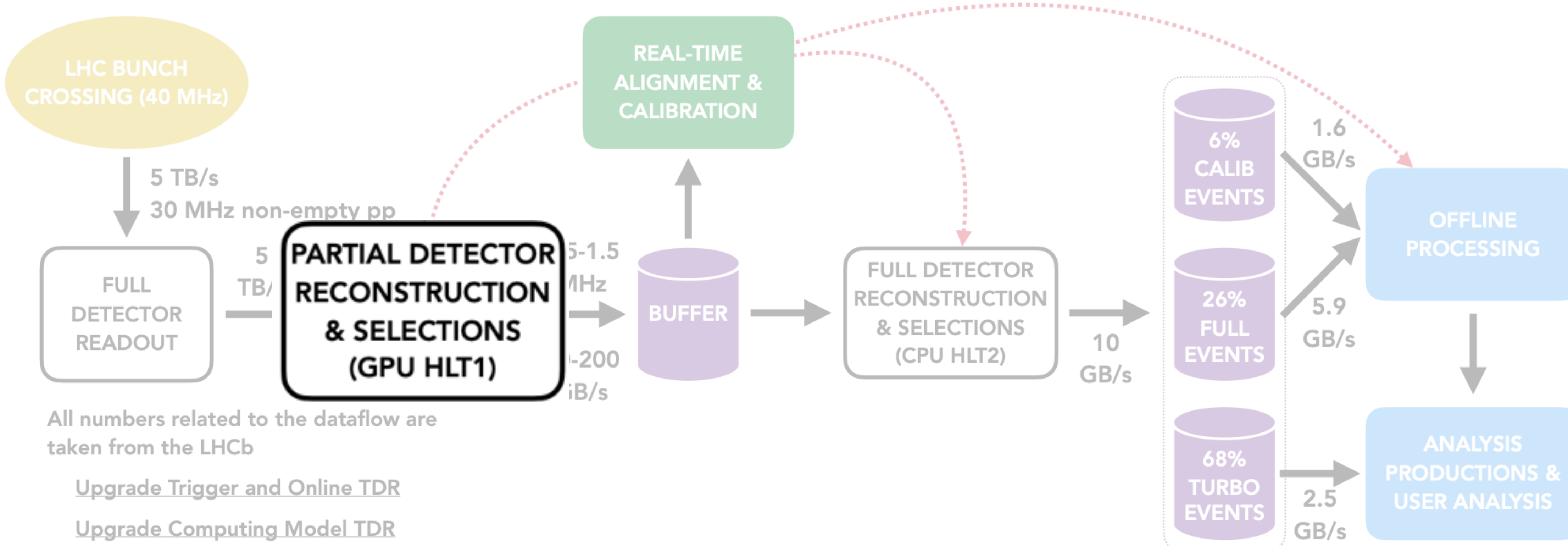
Checkout overview from Daniel!



Also, HLT2 talk from Miro!

- Detector data received by O(500) FPGAs and built into events in the event building (EB) farm servers
- 2-stage software trigger, HLT1 & HLT2
- Real-time alignment & calibration
- After HLT2, 10 GB/s of data for offline processing

The LHCb first level trigger

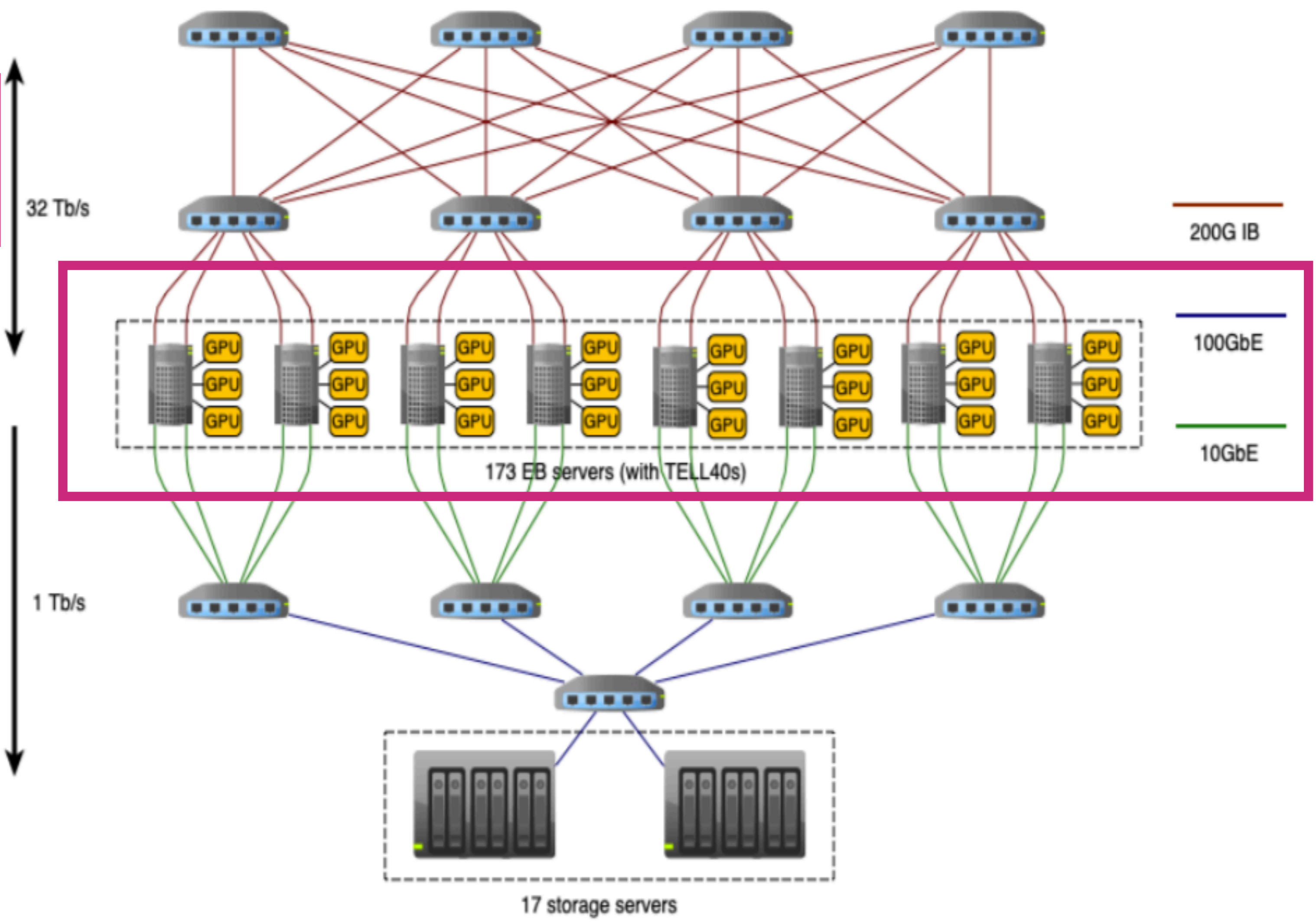


- **The goal of HLT1:**
  - Be able to intake the entirety of the LHCb raw data (5 TB/s) at 30 MHz
  - Perform partial event reconstruction & coarse selection of broad LHCb physics cases
  - Reduce the input rate by a factor of 30 (~ 1 MHz)
  - Store selected events in intermediate buffer for real-time alignment and calibration

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

Are GPUs a good fit?

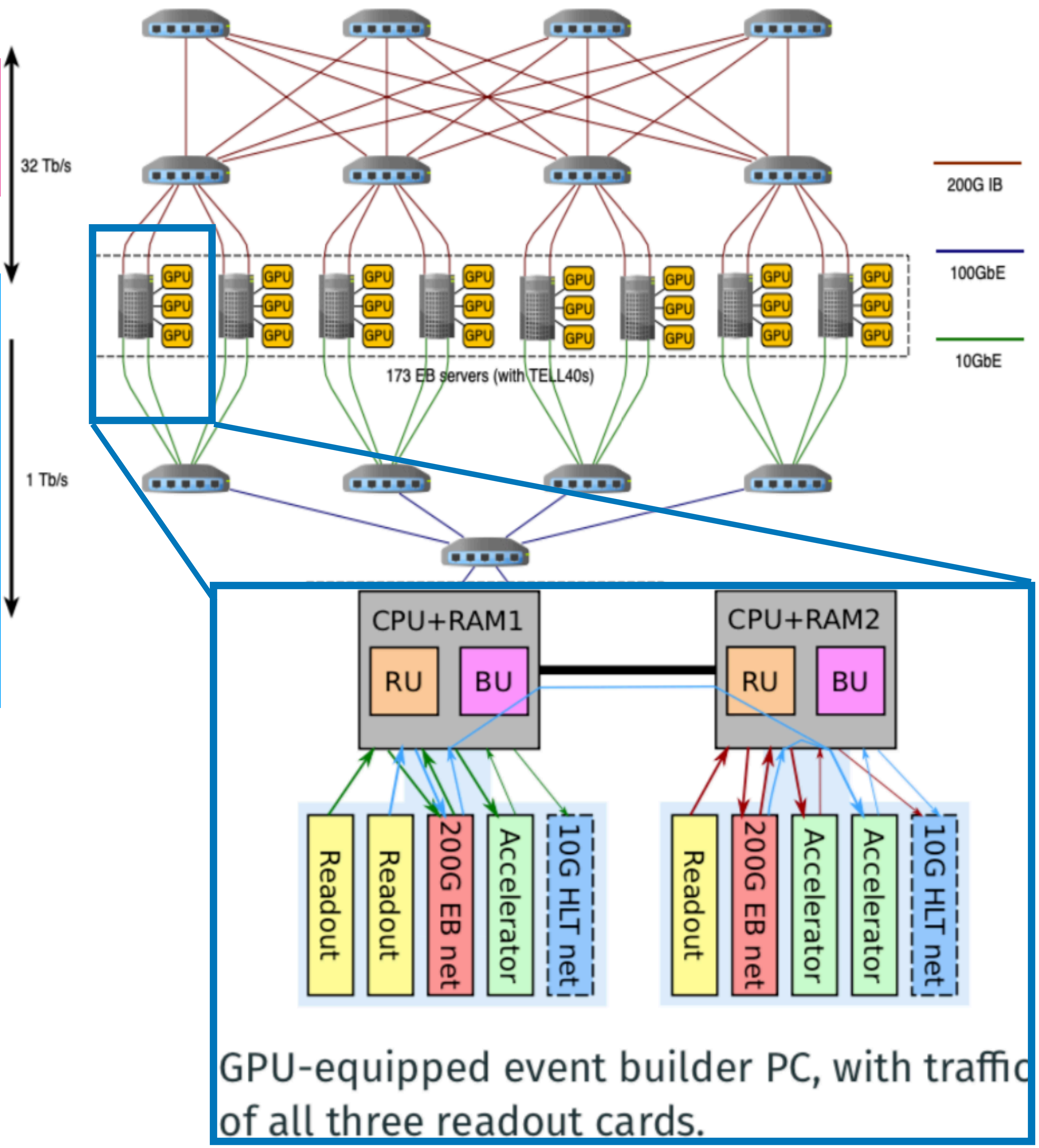
- Event builder farm equipped with 173 servers



Are GPUs a good fit?

- Event builder farm equipped with 173 servers

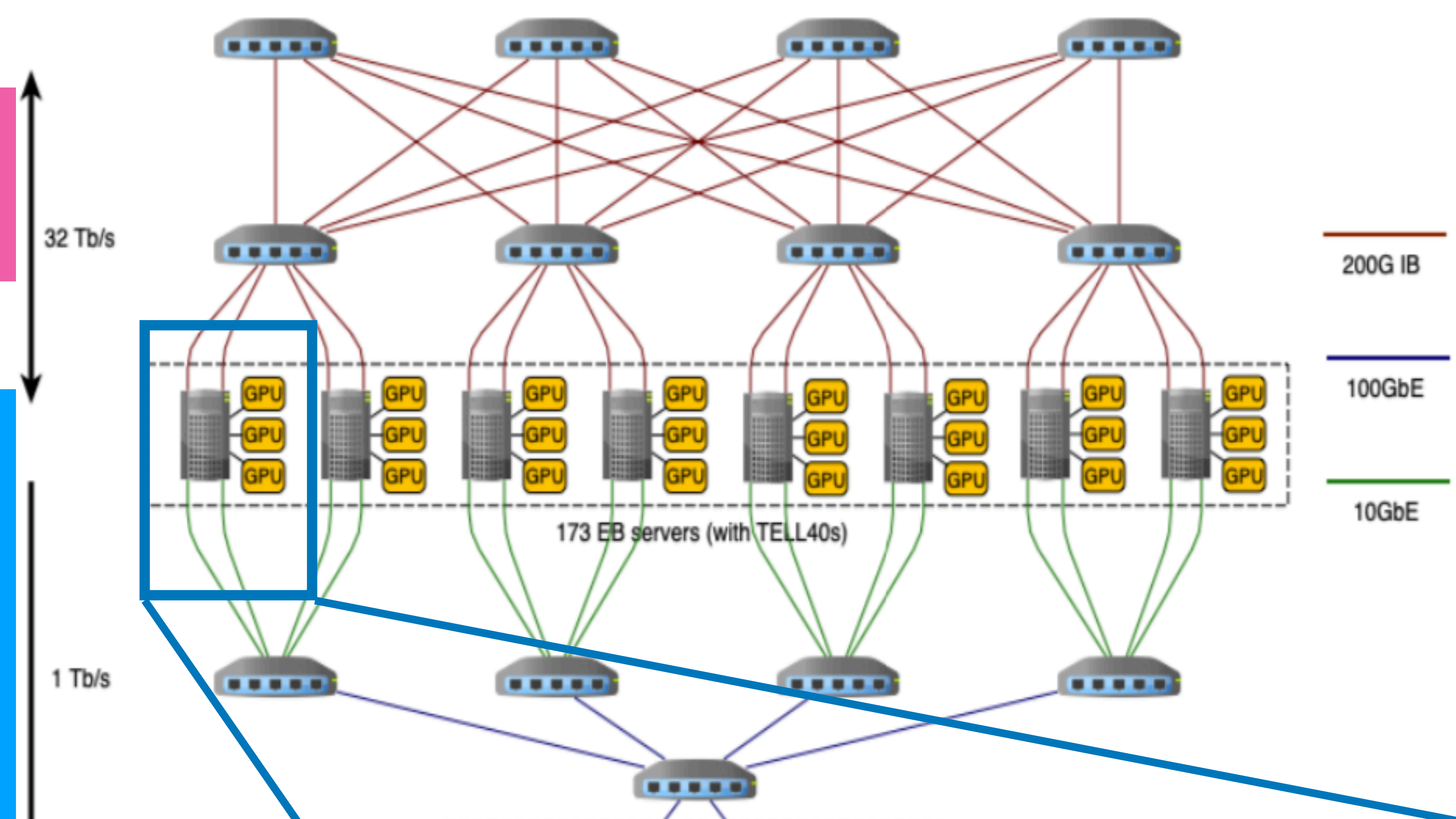
- Each server has 3 free PCIe slots
  - Can be used to host GPUs
  - Sufficient cooling & power
  - Advantageous to have GPUs as self-contained processors
  - Sending data to GPU is like sending data to network card

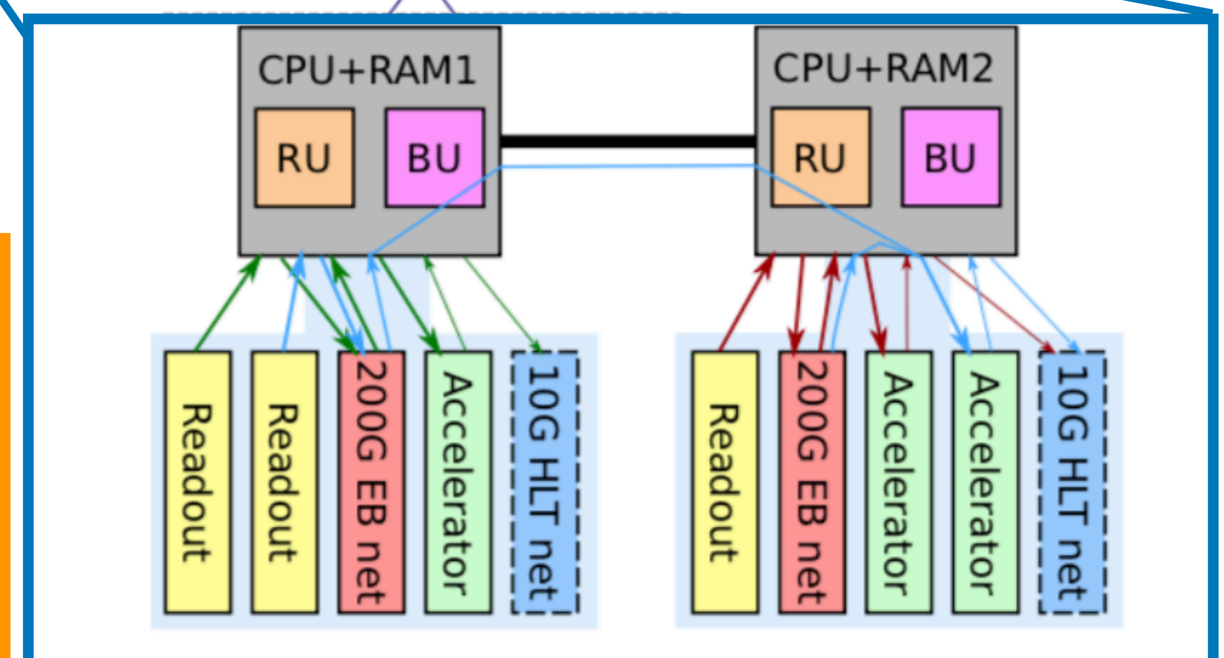


Are GPUs a good fit?

- Event builder farm equipped with 173 servers

- Each server has 3 free PCIe slots
  - Can be used to host GPUs
  - Sufficient cooling & power
  - Advantageous to have GPUs as self-contained processors
  - Sending data to GPU is like sending data to network card





GPU-equipped event builder PC, with traffic of all three readout cards.

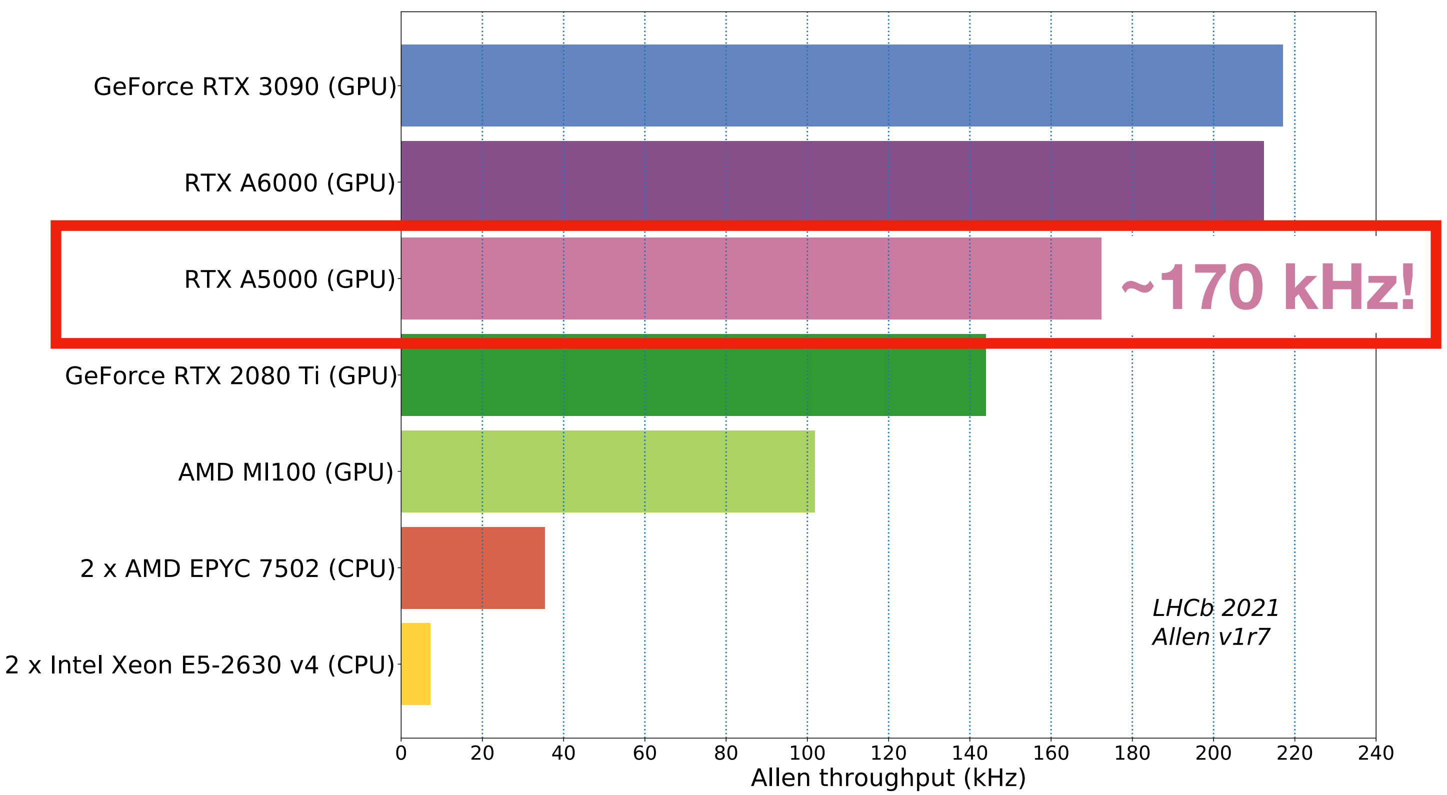
- GPUs map well into LHCb DAQ architecture
- HLT1 tasks inherently parallelizable
- Smaller network between EB & CPU HLT
- Cheaper & more scalable than CPU alternative

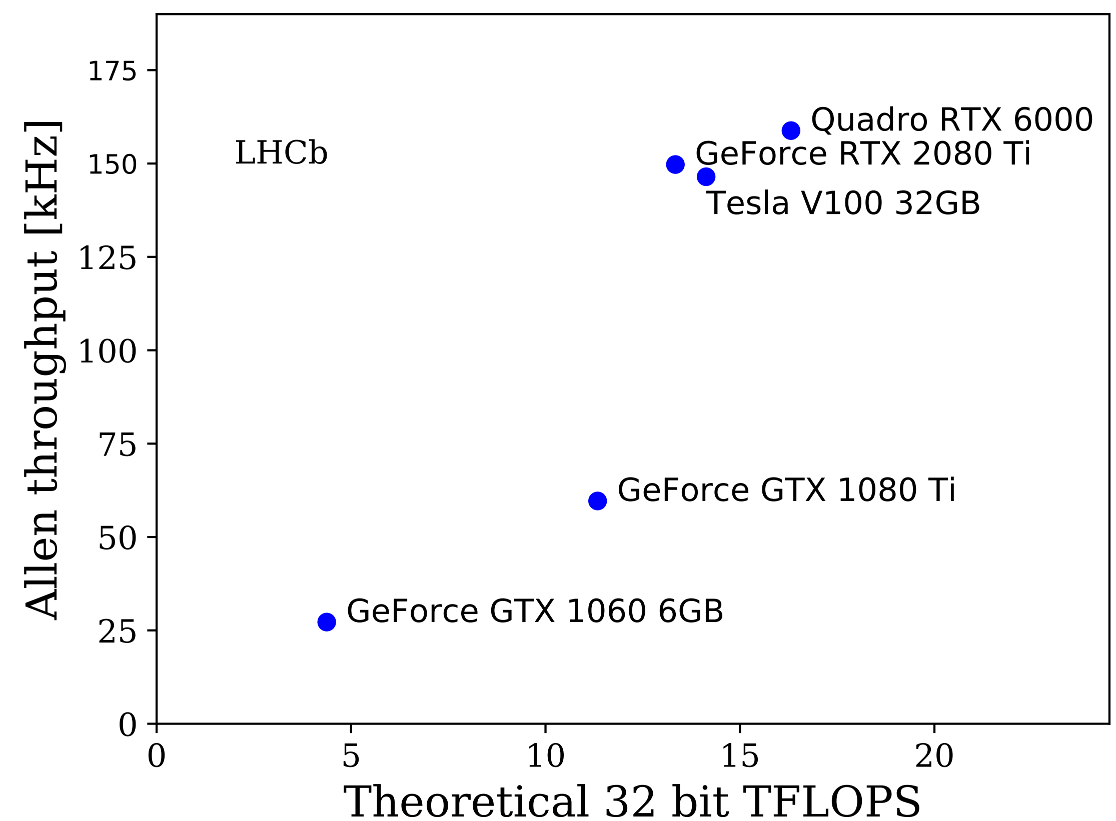
➡**Was chosen as the baseline for the upgrade!**

➡**Is implemented with O(200) Nvidia RTX A5000 GPUs**

Throughput

- 30 MHz goal can be achieved with O(200) GPUs (maximum the Event Builder server can host is 500)
- Throughput scales well with theoretical TFLOPS of GPU card
- Additional functionalities are being explored





LHCb-FIGURE-2020-014

HLT1 sequence

Public software project: [gitlab repo]()

Reconstruction

Tracking at the core of the HLT1 reconstruction

Relies on 3 sub-detector systems:

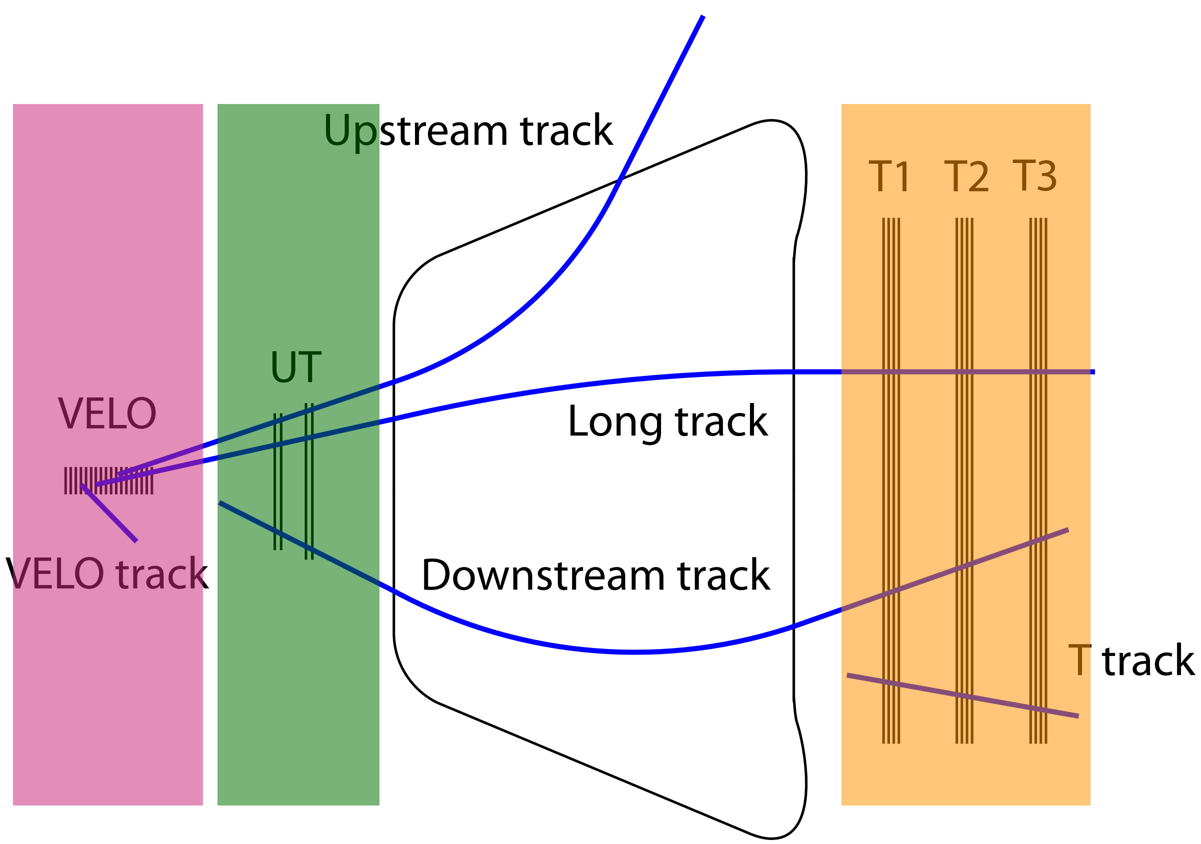
- VELO: clustering, tracking, vertex reconstruction
- UT: track reconstruction, momentum estimate, fake rejection
- SciFi: track reconstruction, momentum measurement

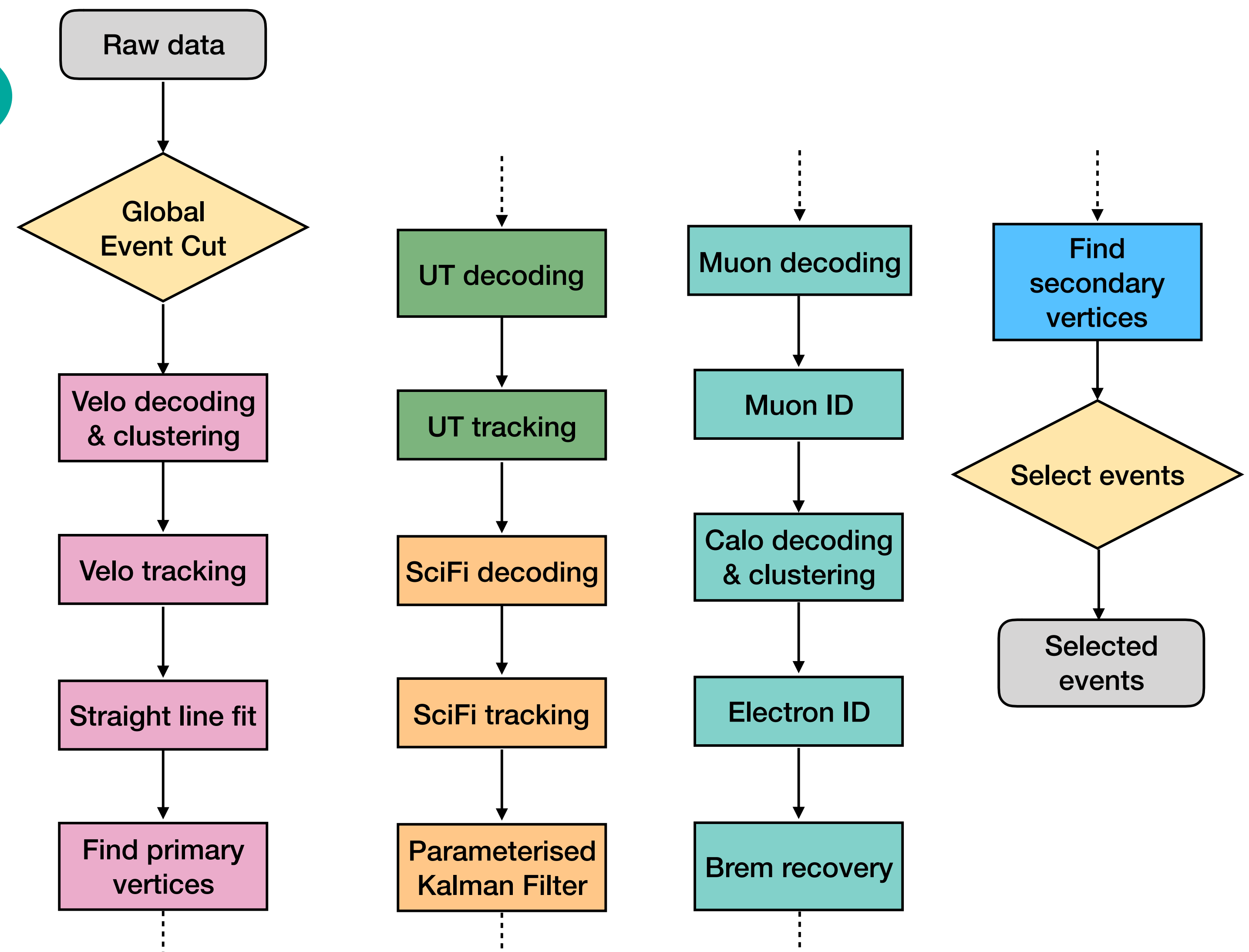
+

PID from muon stations + **Calo**

NEW!

Trigger lines → Event selection

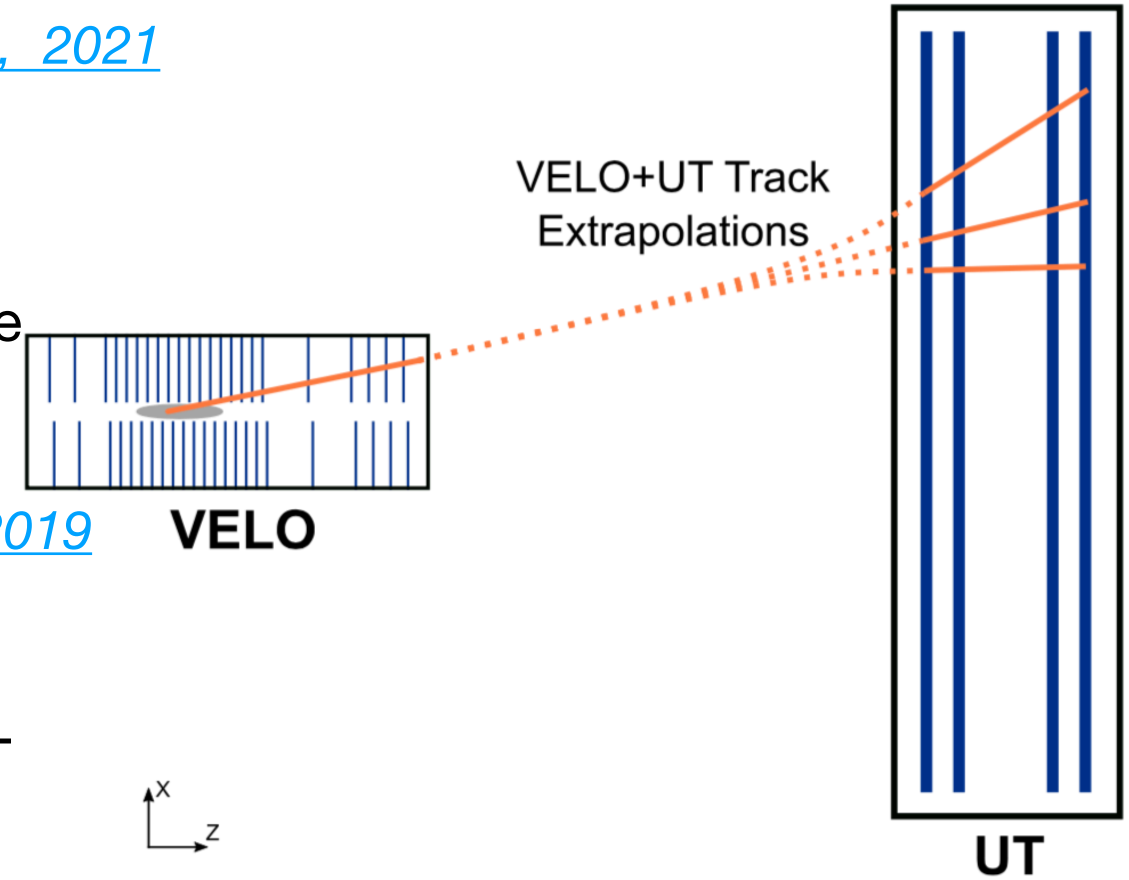




Track reconstruction with GPUs

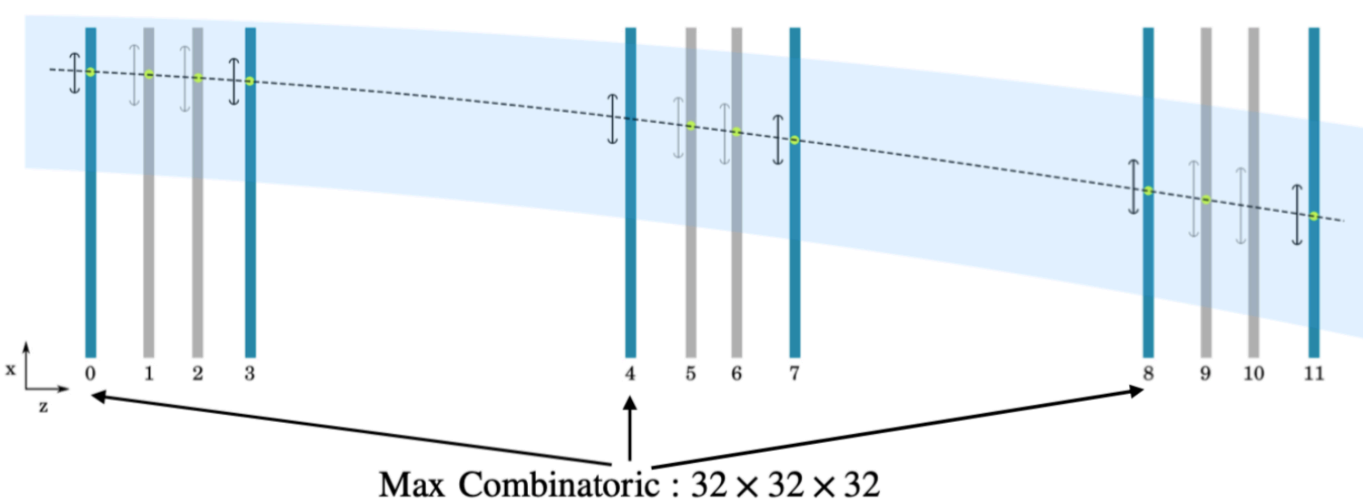
Velo tracking: [Journal of Computational Science, vol. 54, 2021]

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



UT tracking: [IEEE Access, vol. 7, pp. 91612-91626, 2019]

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

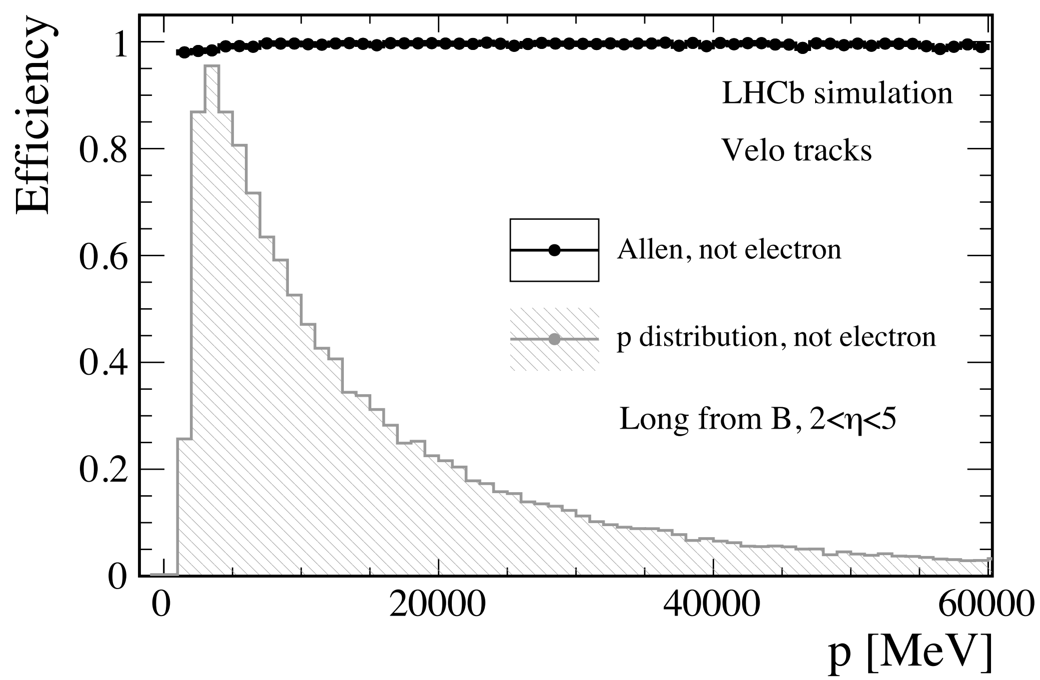


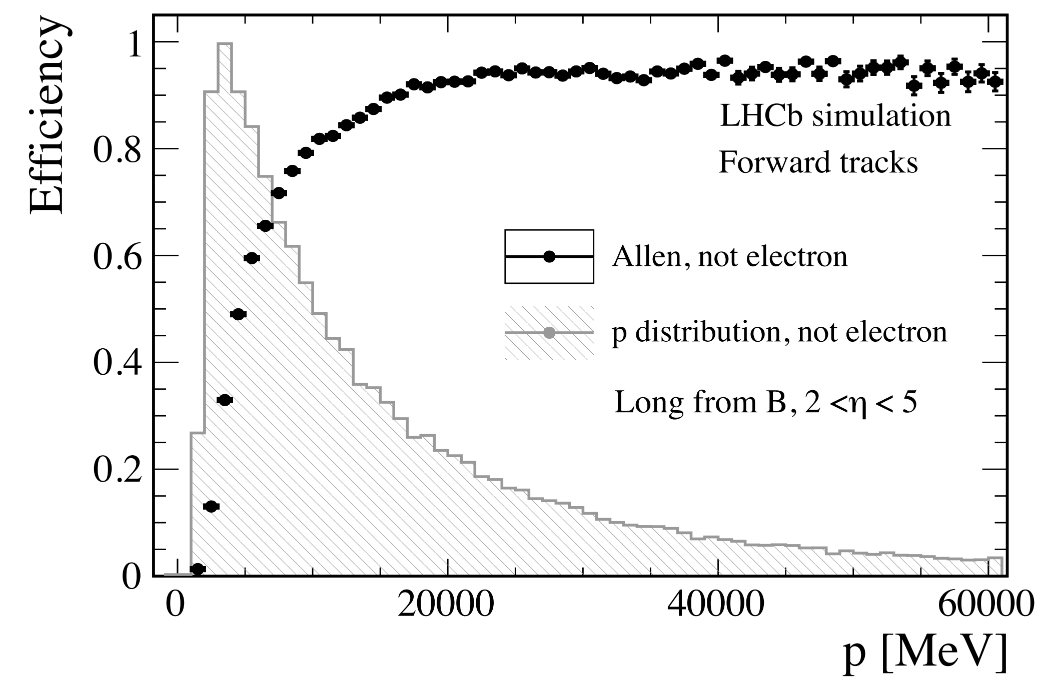
SciFi tracking: [Comput Softw Big Sci 4, 7 (2020)]

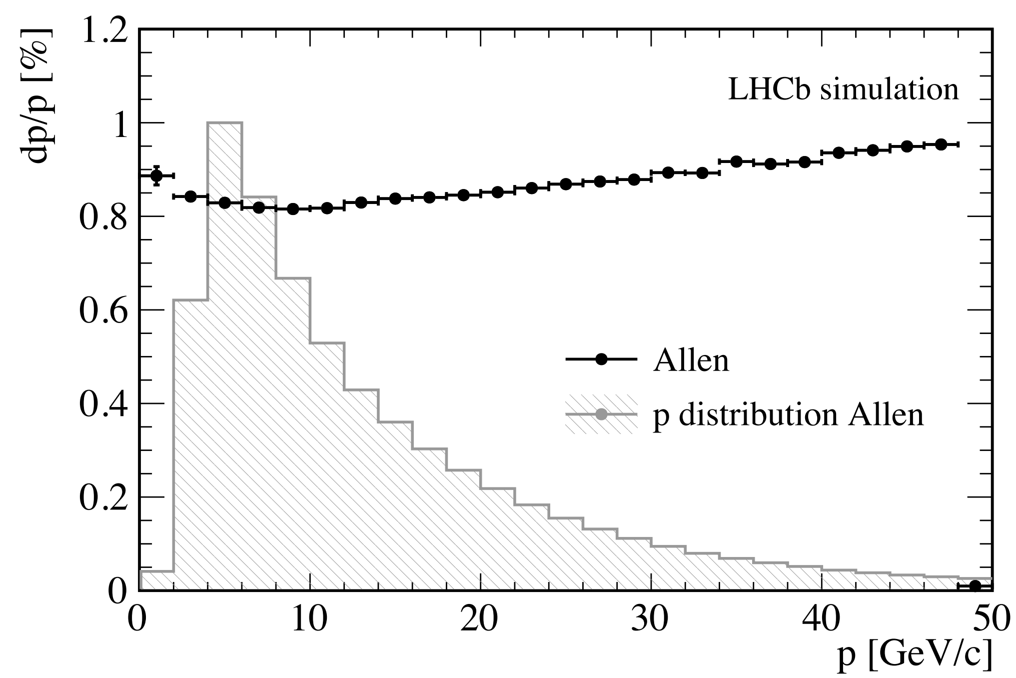
- 3 stations with 4 layers of Scintillating Fibres
- Velo-UT tracks extrapolated using parametrization
- Parallelized *Forward algorithm* to reconstruct **long tracks:**
  - Search windows from Velo-UT momentum estimate
  - Form triplets and extend to remaining layers

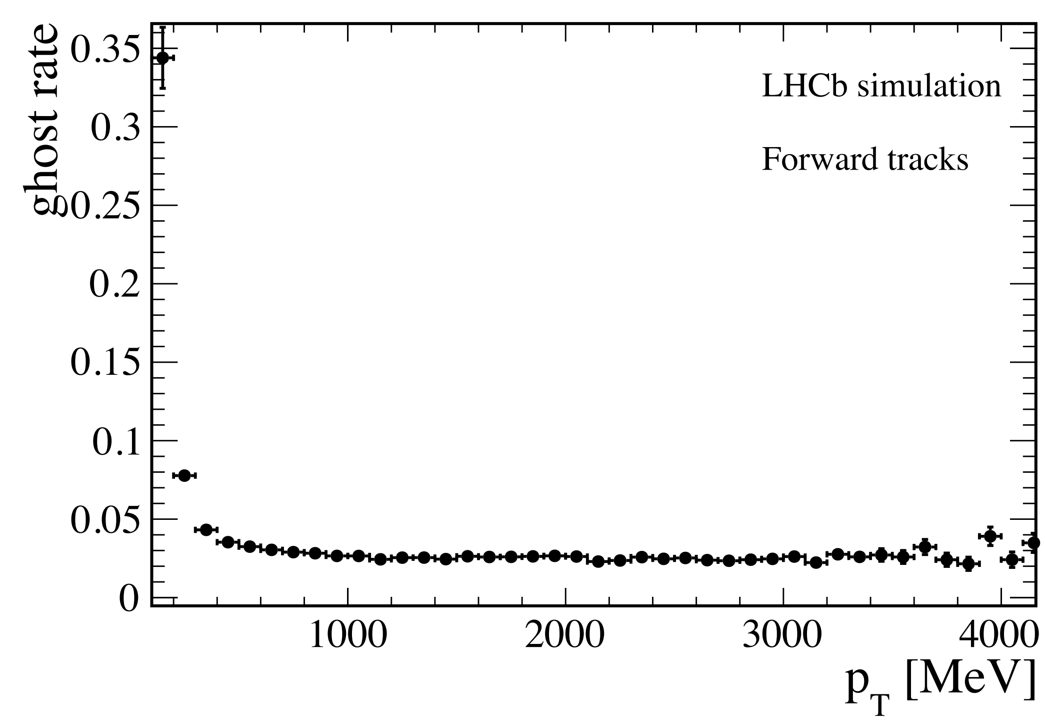
HLT1 tracking performance

- **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





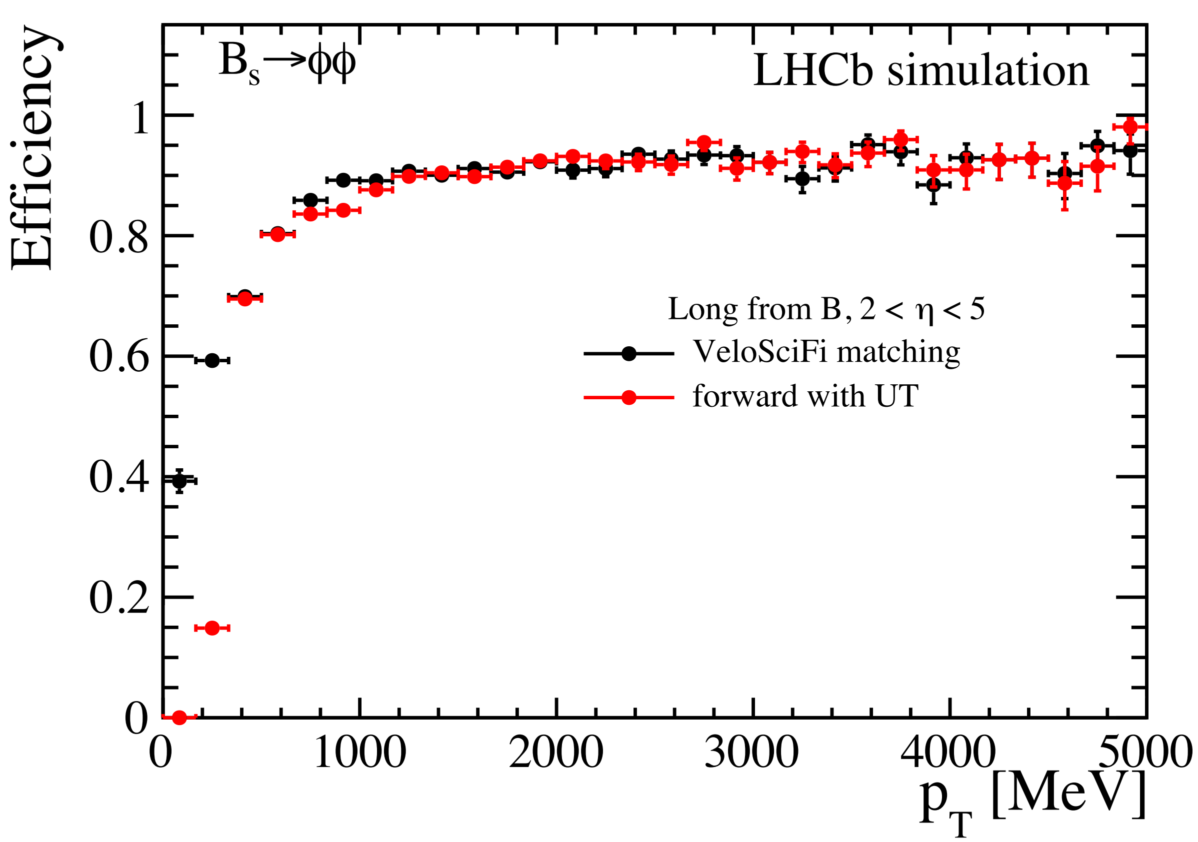


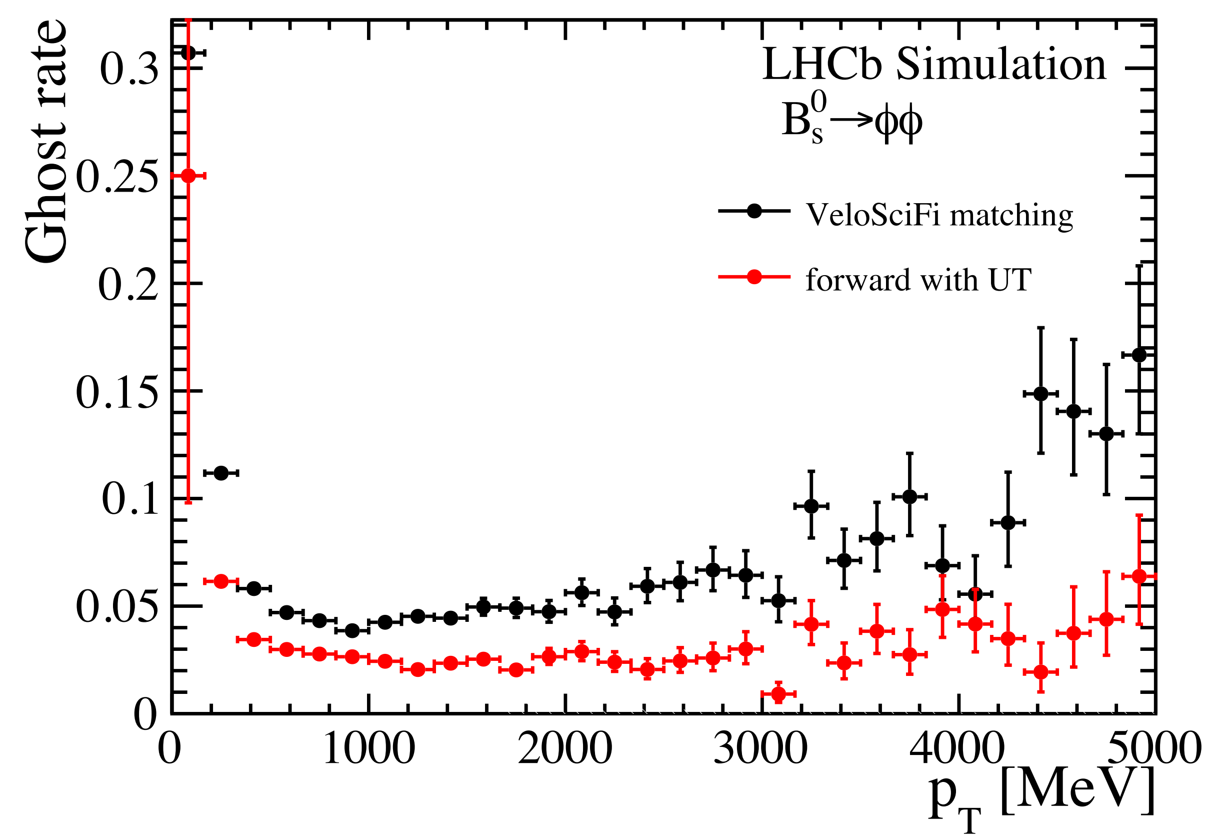


LHCb-FIGURE-2020-014

Tracking without the UT

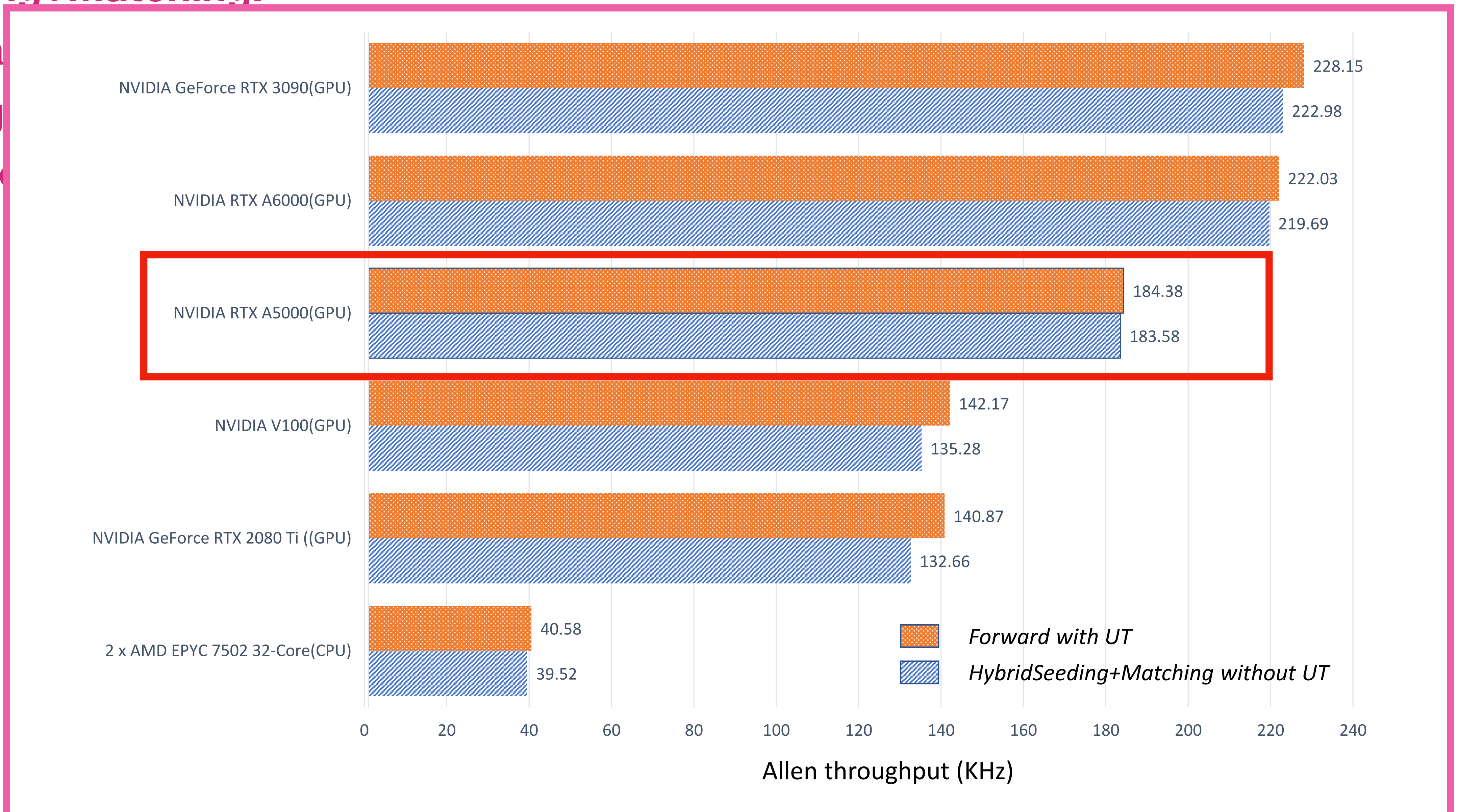
- In 2022, the UT detector will unfortunately not be available for data-taking
- Tracking performance and throughput maintained, at the cost of larger fake rate
- Opportunity to commission 2 options, which **both maintain the current throughput:**
  - Forward without UT -> check out Daniel's talk
  - **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)





Tracking without the UT

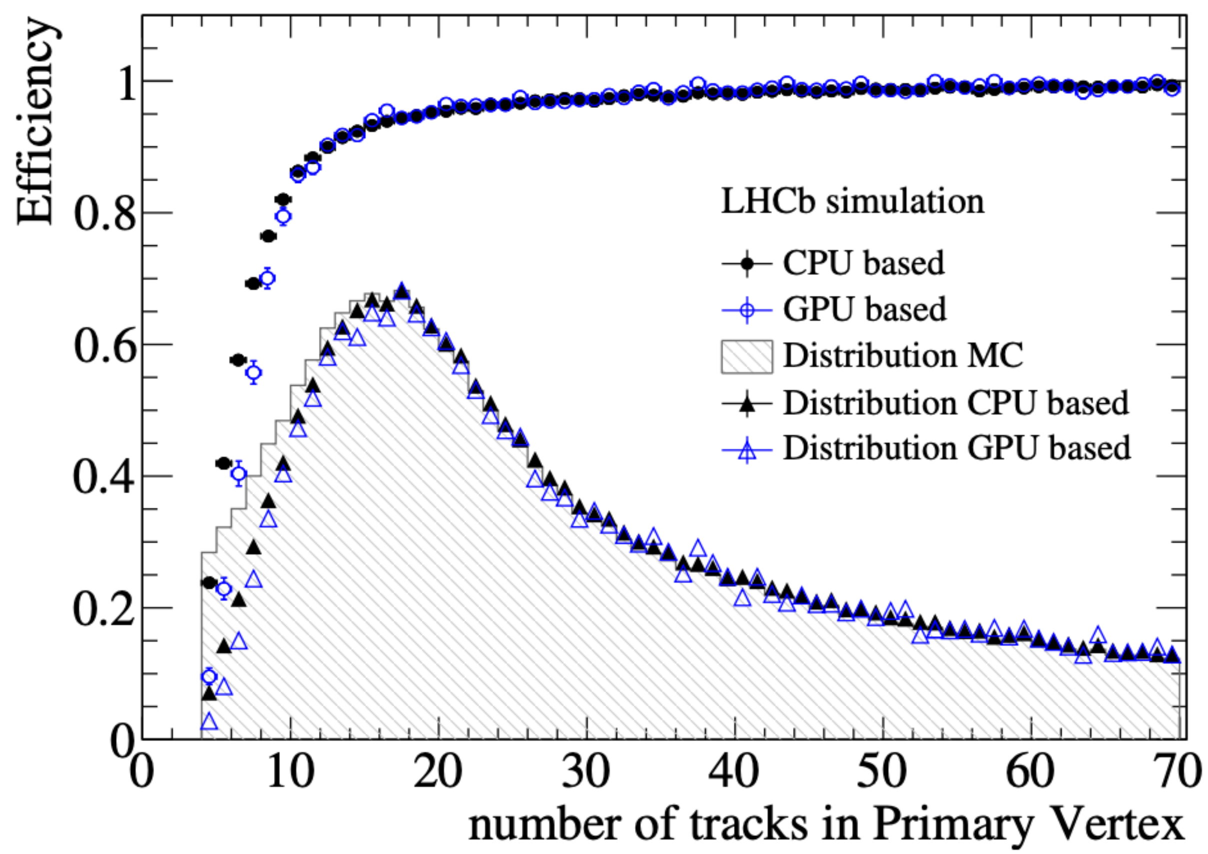
- In 2022, the UT detector will unfortunately not be available for data-taking
- Tracking performance and throughput maintained, at the cost of larger fake rate
- Opportunity to commission 2 options, which **both maintain the current throughput:**
  - Forward without UT -> check out Daniel's talk
  - **Seeding+Matching:**
    - Sta
    - Hig
    - Op



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 N. tracks > 10

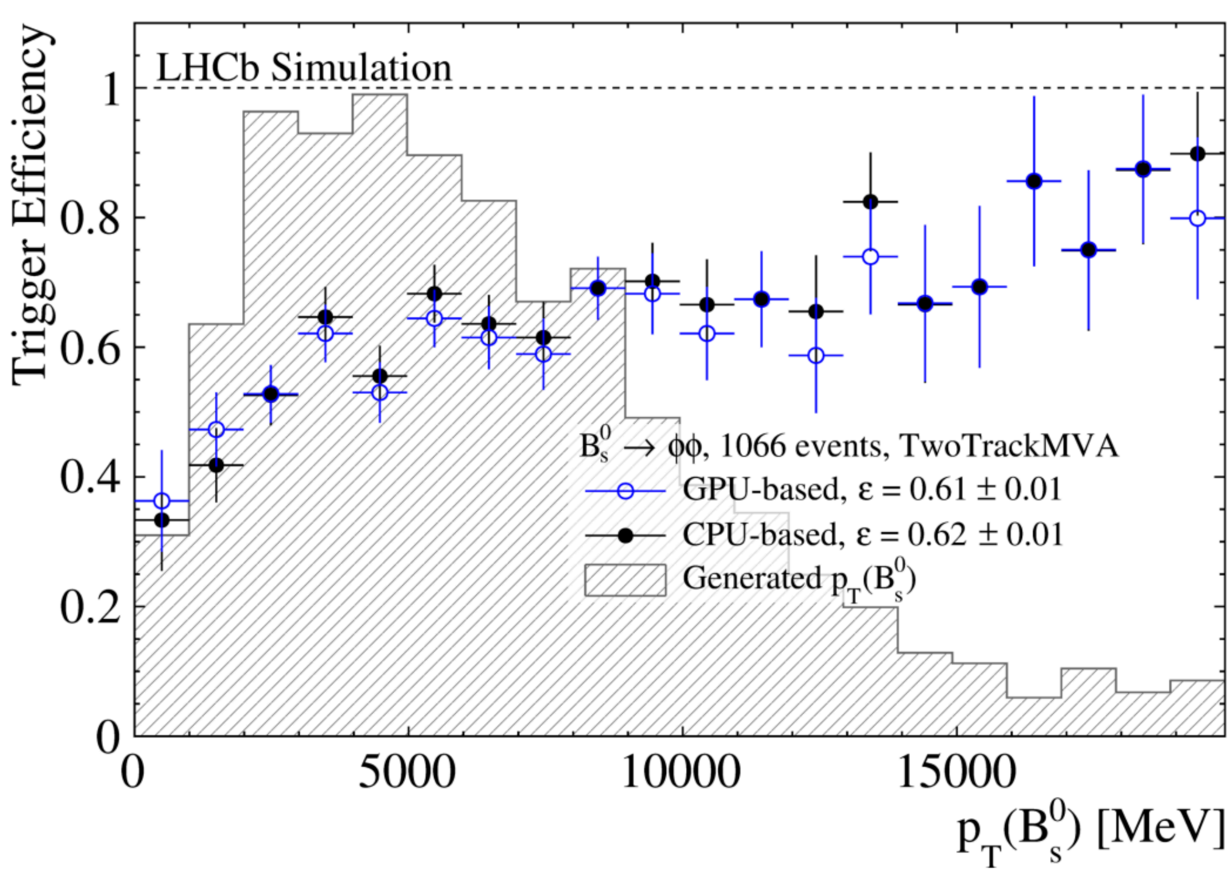
[Comput. Softw. Big Sci. 6 (2022) no.1, 1]



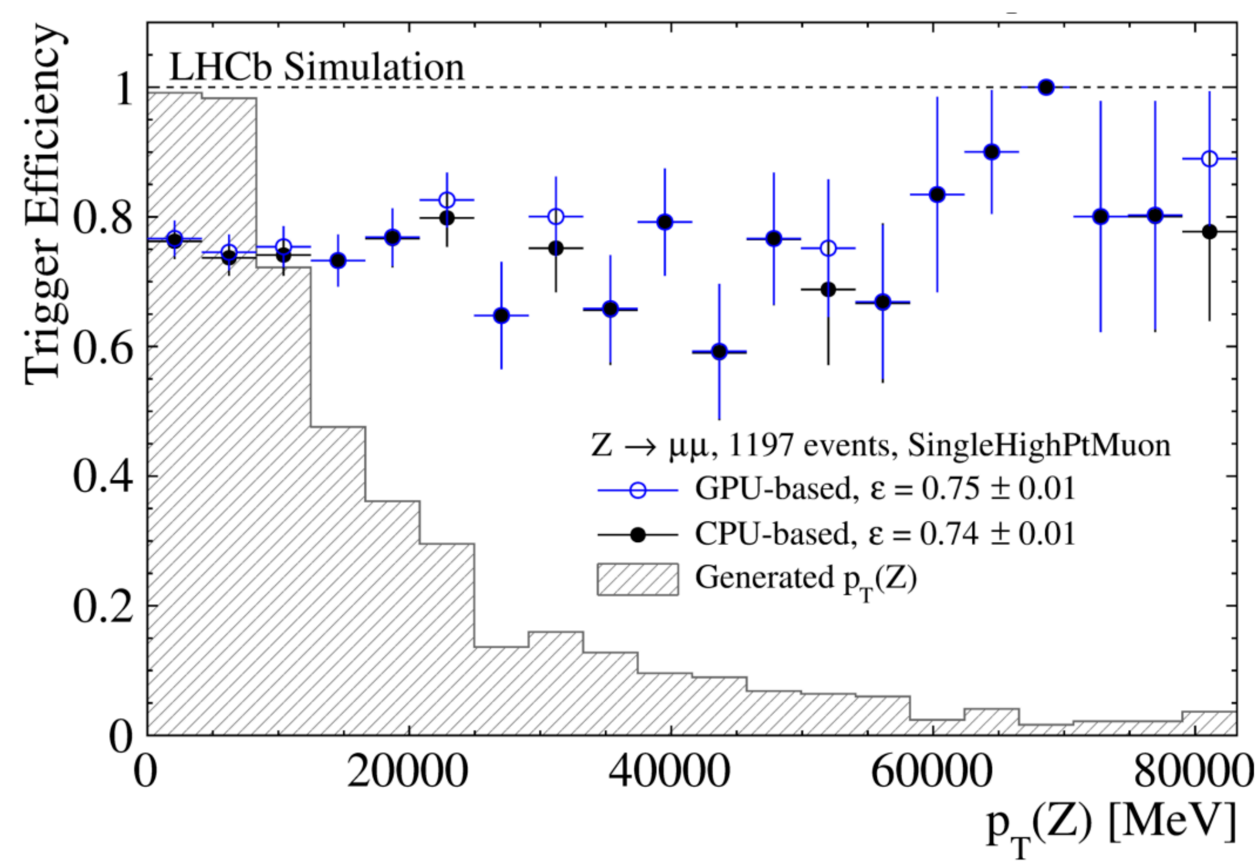
HLT1 selection performance

- Inclusive rate for the main HLT1 lines ~ 1 MHz
- O(30) lines implemented so far:
  - Cover majority of LHCb physics program (B, D decays, semileptonic, EW physics)
  - Special lines for monitoring, alignment and calibration
  - Additional trigger lines under development
- Compatible performance between CPU and GPU

TwoTrackMVA line for $B_s^0 \to \phi\phi$

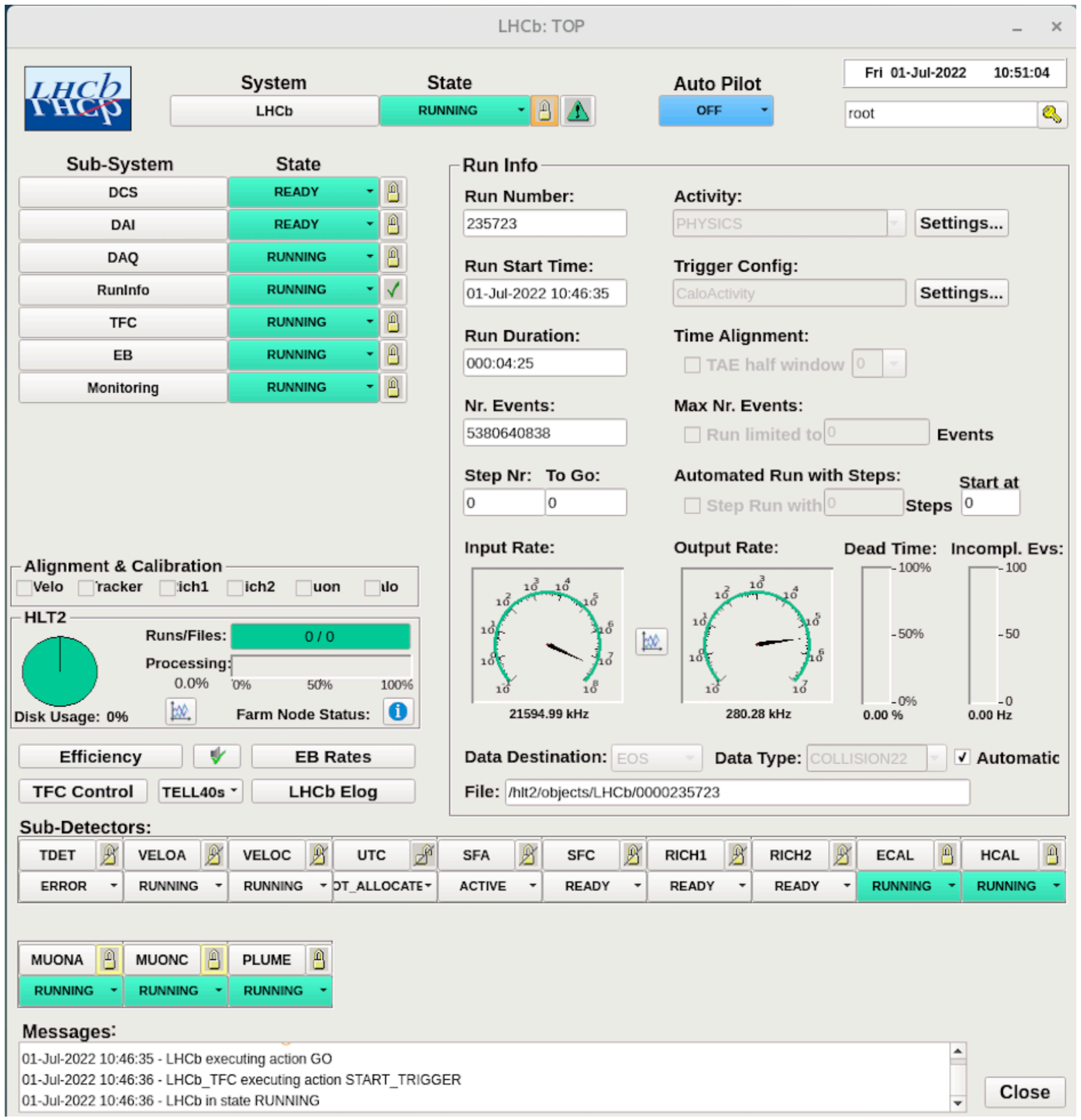


SingleHighPtMuon line for $Z \to \mu^+\mu^-$



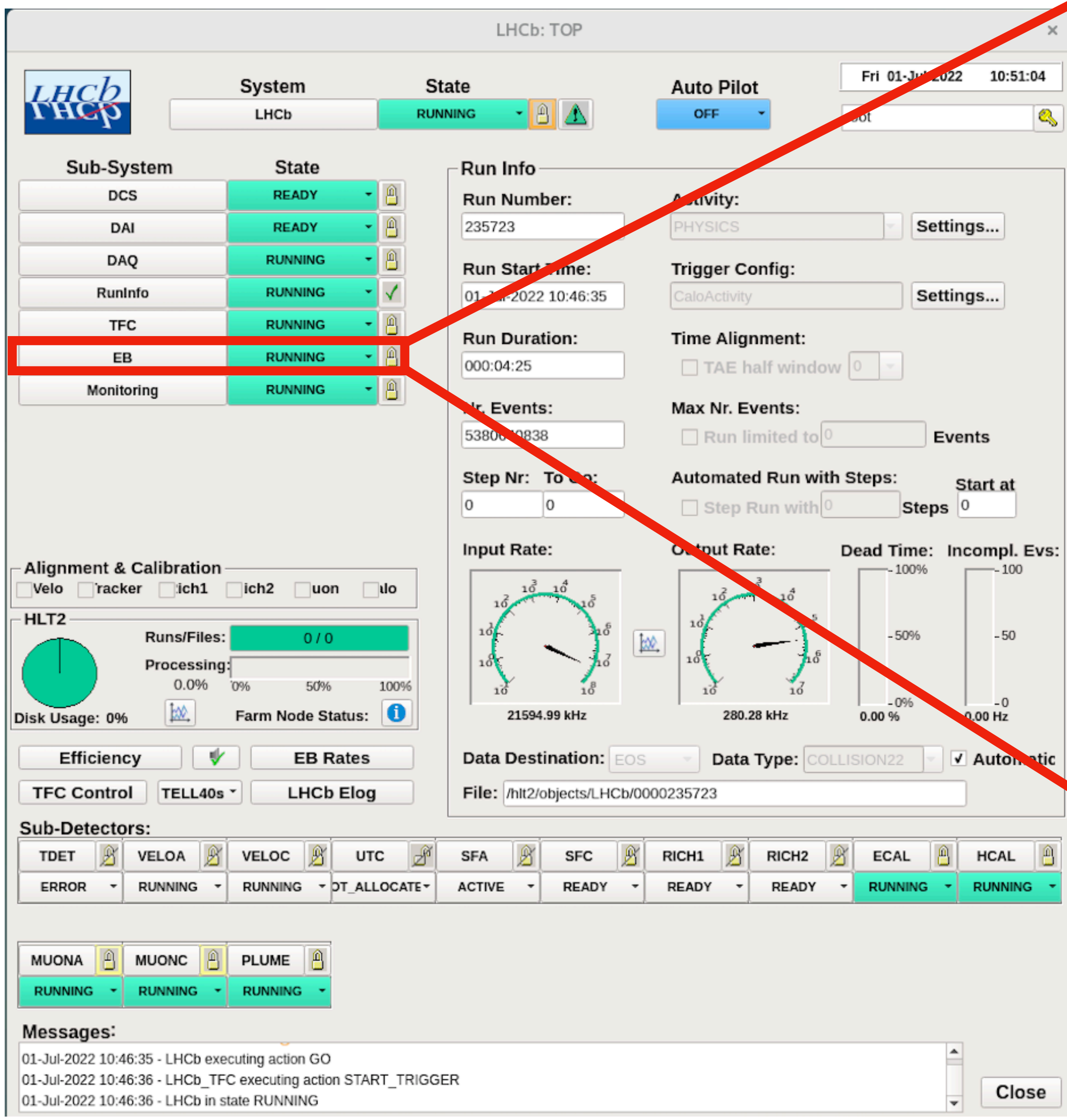
Comput. Softw. Big Sci. 6 (2022) no.1, 1

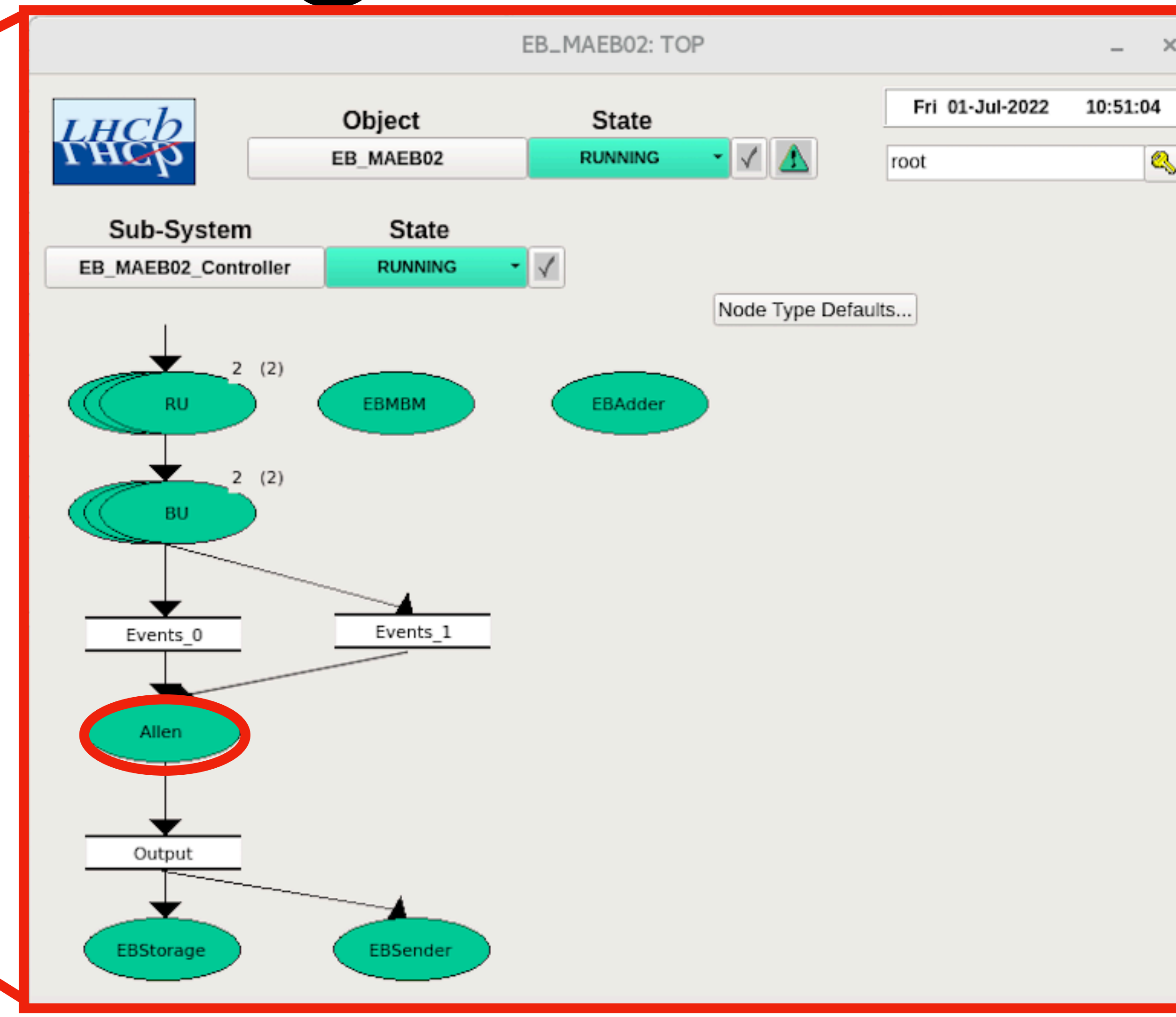
HLT1 commissioning



- 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

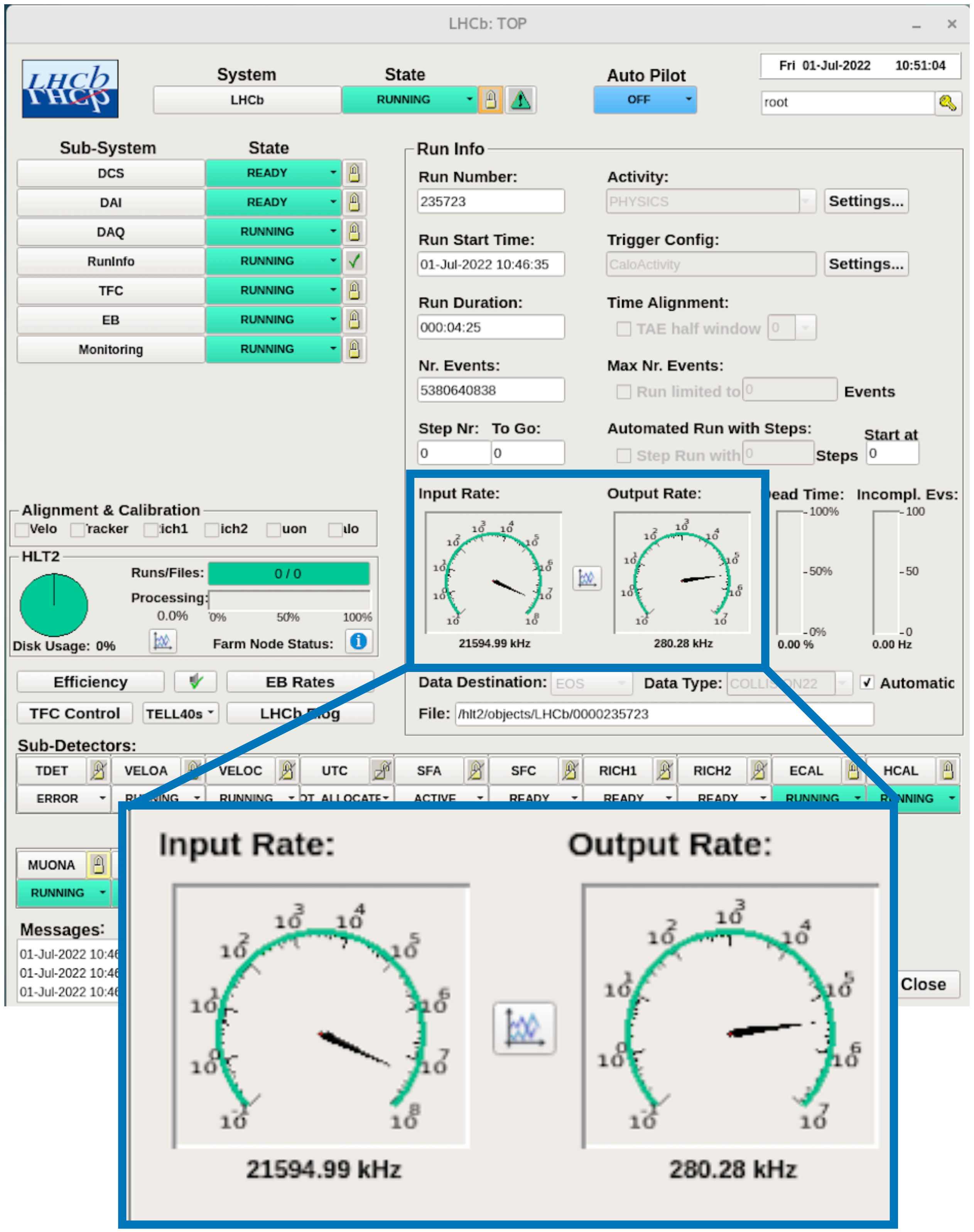
HLT1 commissioning

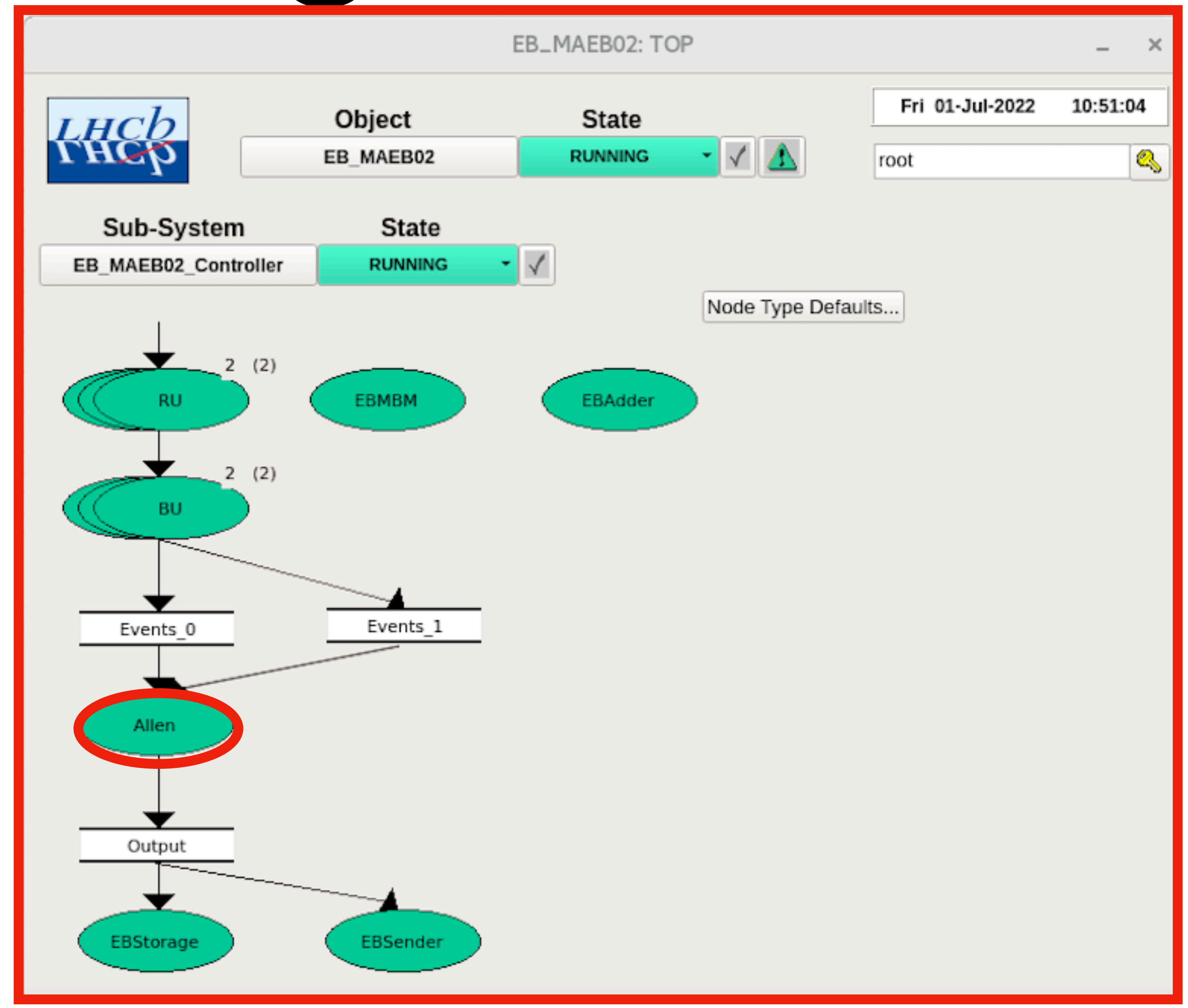




- ~200 GPUs are installed in the EB
- HLT1 is already included in the global partition

HLT1 commissioning





- ~200 GPUs are installed in the EB
- HLT1 is already included in the global partition
- Triggering on calorimeter clusters @ 20 MHz!
- Next steps:
  - Test full trigger sequence when trackers are ready
  - Monitoring in progress

Conclusions

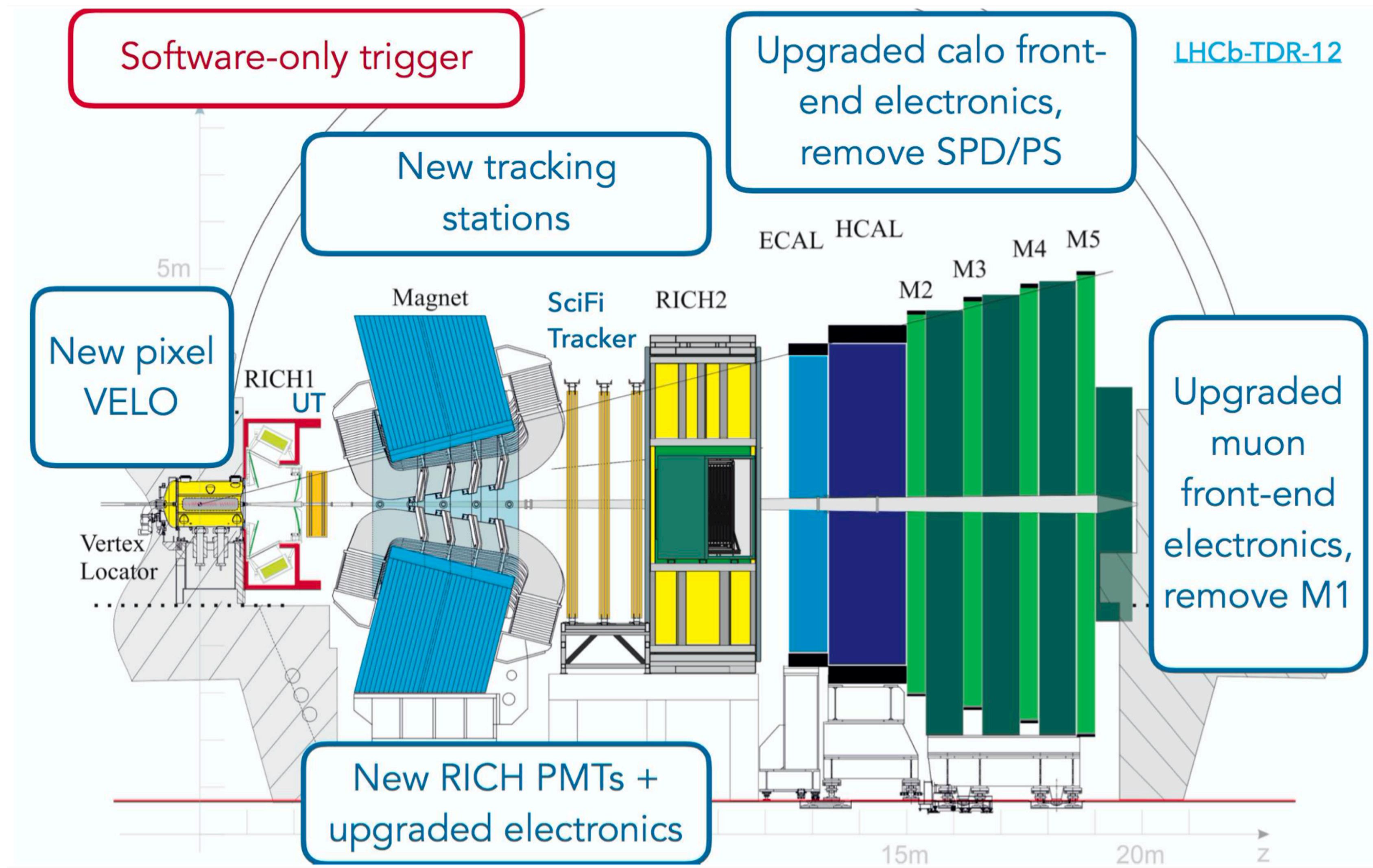
- LHCb is currently undergoing its first major upgrade in order to increase its instantaneous luminosity by x5
- Major changes on the trigger strategy:
  - Remove L0 hardware trigger, read-out full detector at 30 MHz
  - New, software-only first level trigger based on GPUs
- Partial event reconstruction and trigger selection lines implemented, excellent physics performance
- Throughput ~170 kHz → system can be realised with around 200 GPUs
- GPUs are installed in the EB server and commissioning ongoing with first collisions!

Stay tuned for more updates!

Thank you for your attention!

Backup

The LHCb U1 upgrade



Will start in 2022!

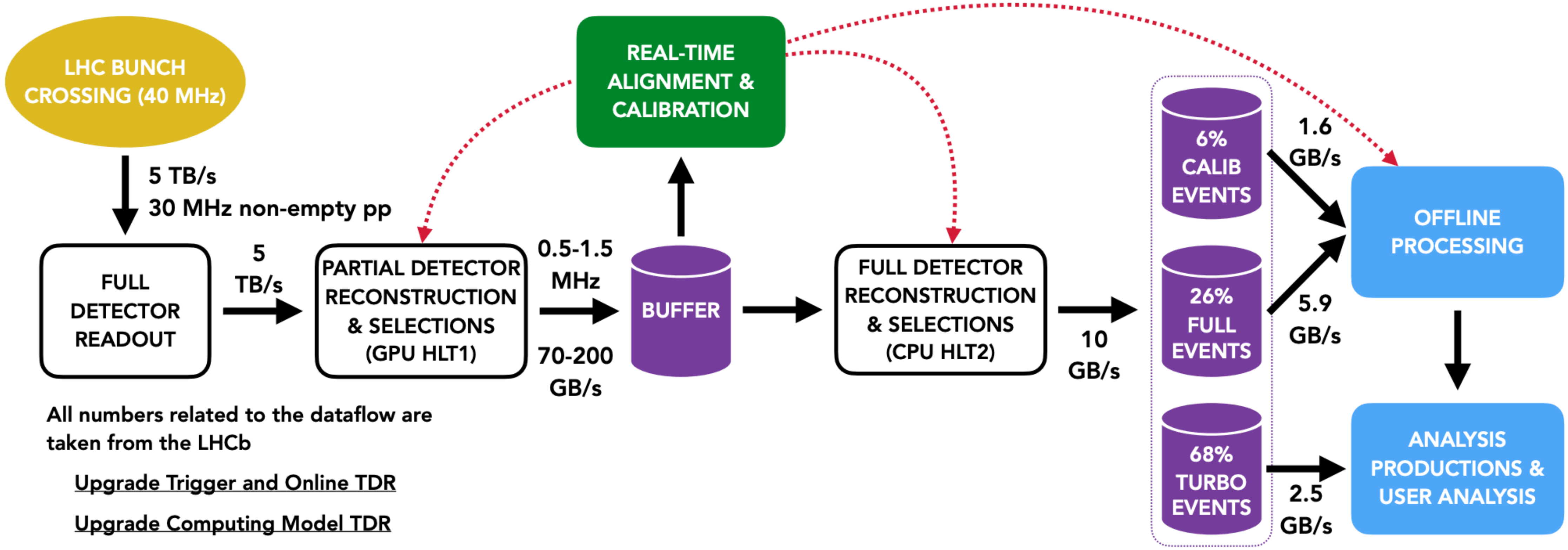
The LHCb detector at CERN:

- Single-arm forward spectrometer for high-precision flavour physics
- High precision tracking and vertexing
- Complemented with excellent PID

The U1 upgrade

- Instantaneous luminosity will increase by x5
- Major upgrade in all sub-detectors to handle increased rates
- Software-only trigger!

LHCb data-flow in Run 3



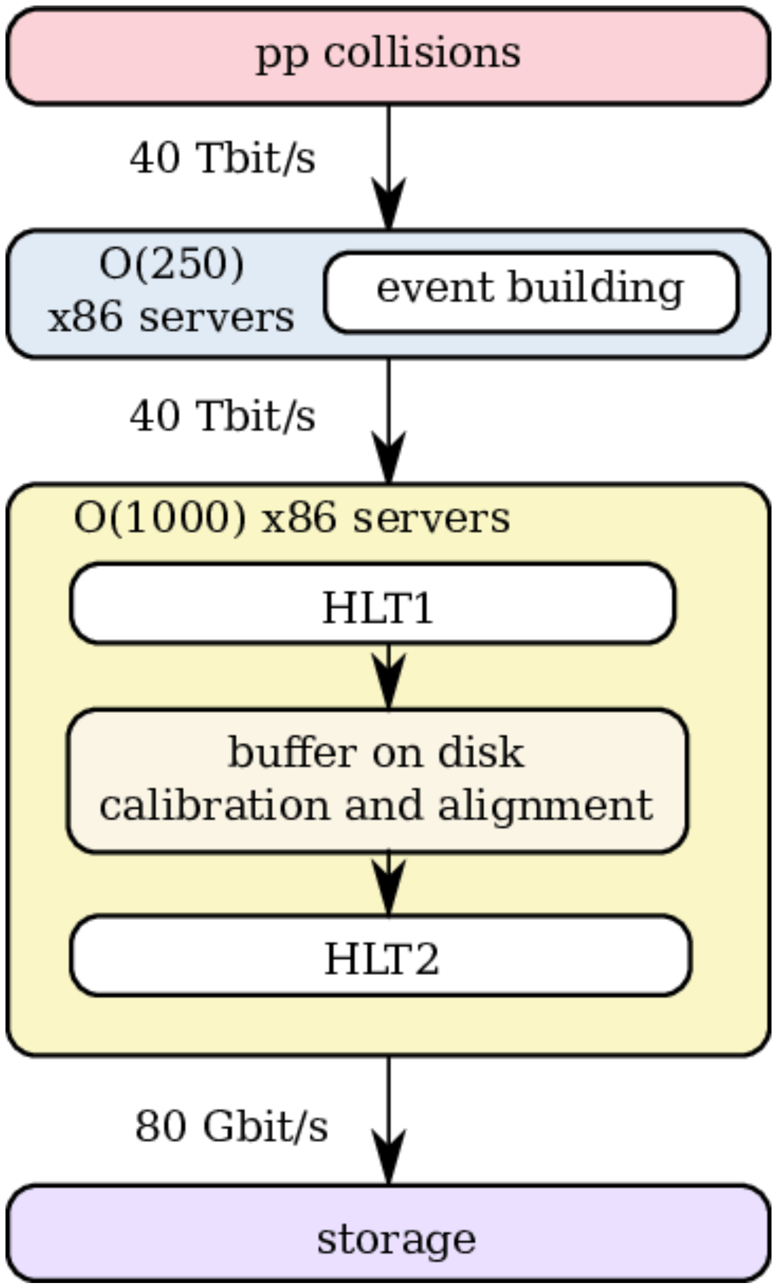
All numbers related to the dataflow are taken from the LHCb

[Upgrade Trigger and Online TDR](#)

[Upgrade Computing Model TDR](#)

- Detector data received by O(500) FPGAs and built into events in the event building (EB) farm servers
- 2-stage software trigger:
  - HLT1: partial event reconstruction and coarse selection, reduces rate to ~ 1 MHz
  - HLT2: full event reconstruction and O(1000) selection lines
  - Buffering between HLT1 & HLT2 → real-time alignment & calibration
- After HLT2, 10 GB/s of data for offline processing

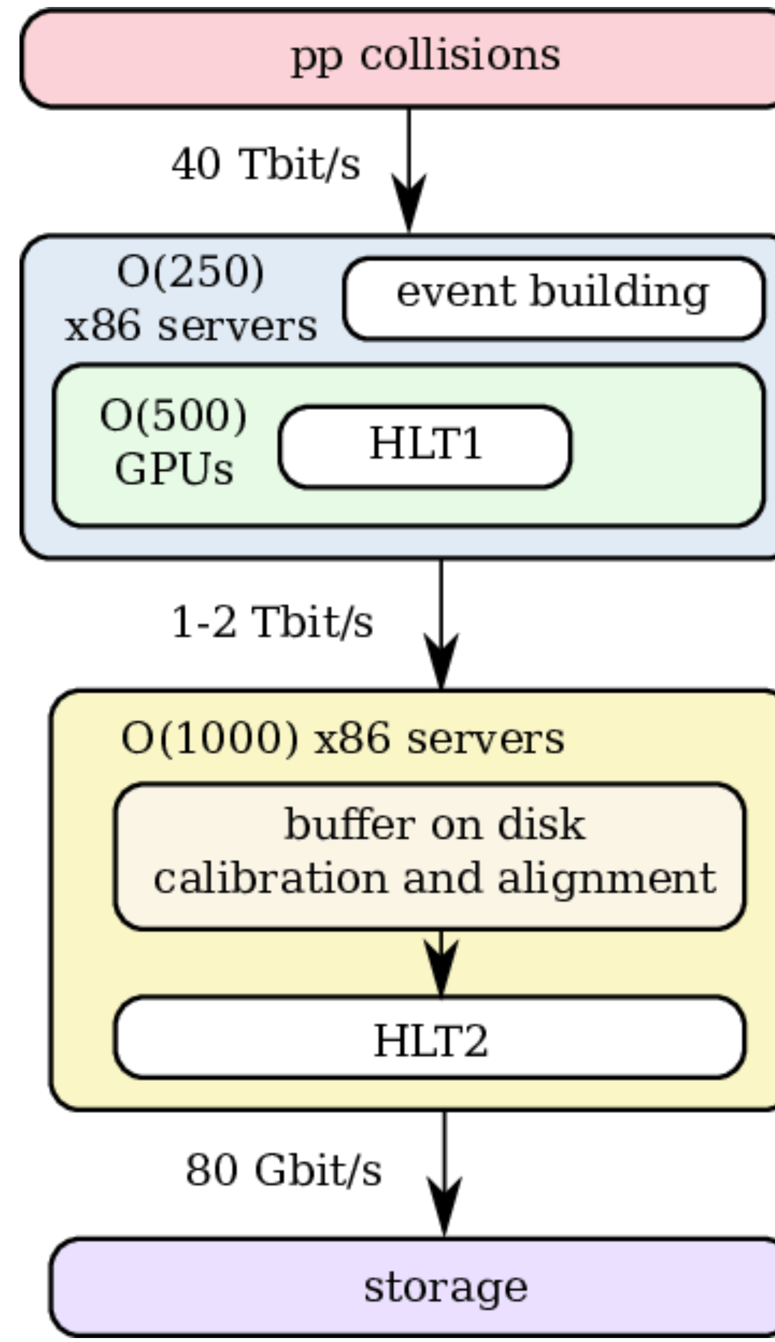
Architecture upgrade options



Detector data received by O(500) FPGAs and built into events in the EB servers

Two options:

1. Send full 40 Tb/s to a CPU processing server → extra network needed
2. Fill extra EB slots with GPUs → reduce rate locally to 1 Tb/s before full processing



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

[Documentation]

Edit on GitLab

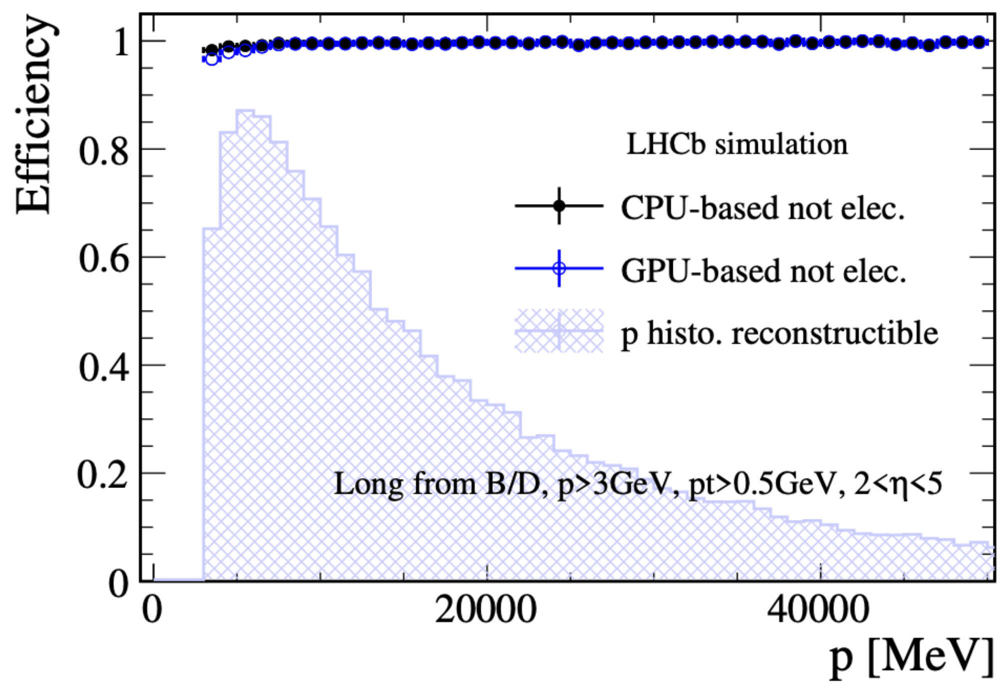
Welcome to Allen's documentation!

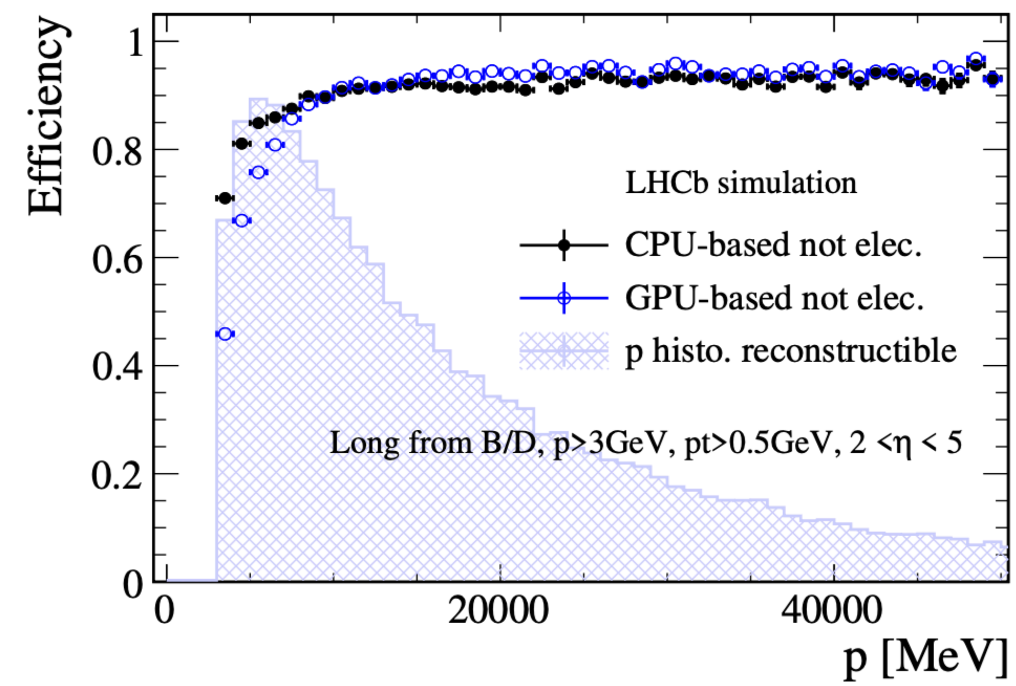
Allen is the LHCb high-level trigger 1 (HLT1) application on graphics processing units (GPUs). It is responsible for filtering an input rate of 30 million collisions per second down to an output rate of around 1-2 MHz. It does this by performing fast track reconstruction and selecting pp collision events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

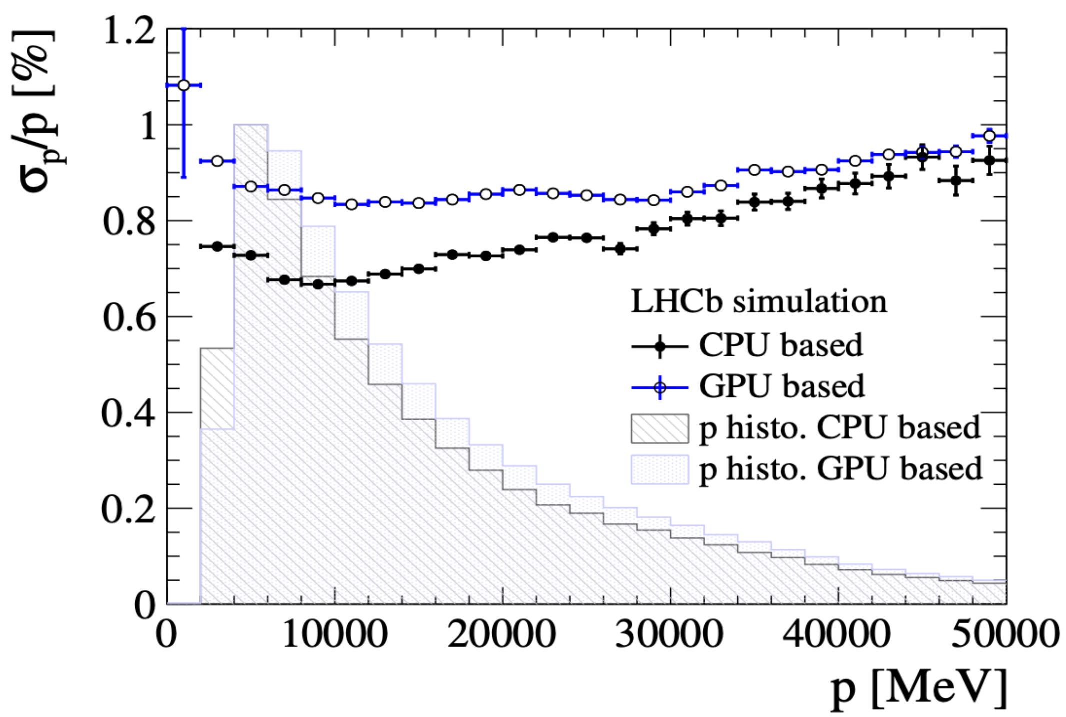
This site documents various aspects of Allen.

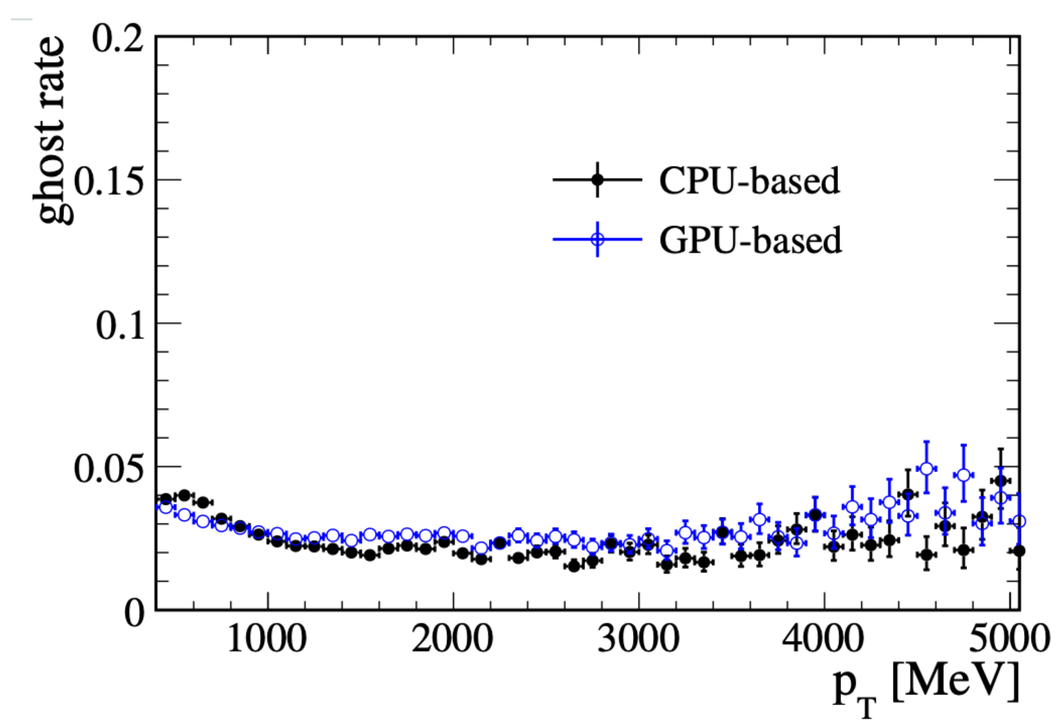
HLT1 CPU/GPU tracking performance

Compatible performance between CPU and GPU!





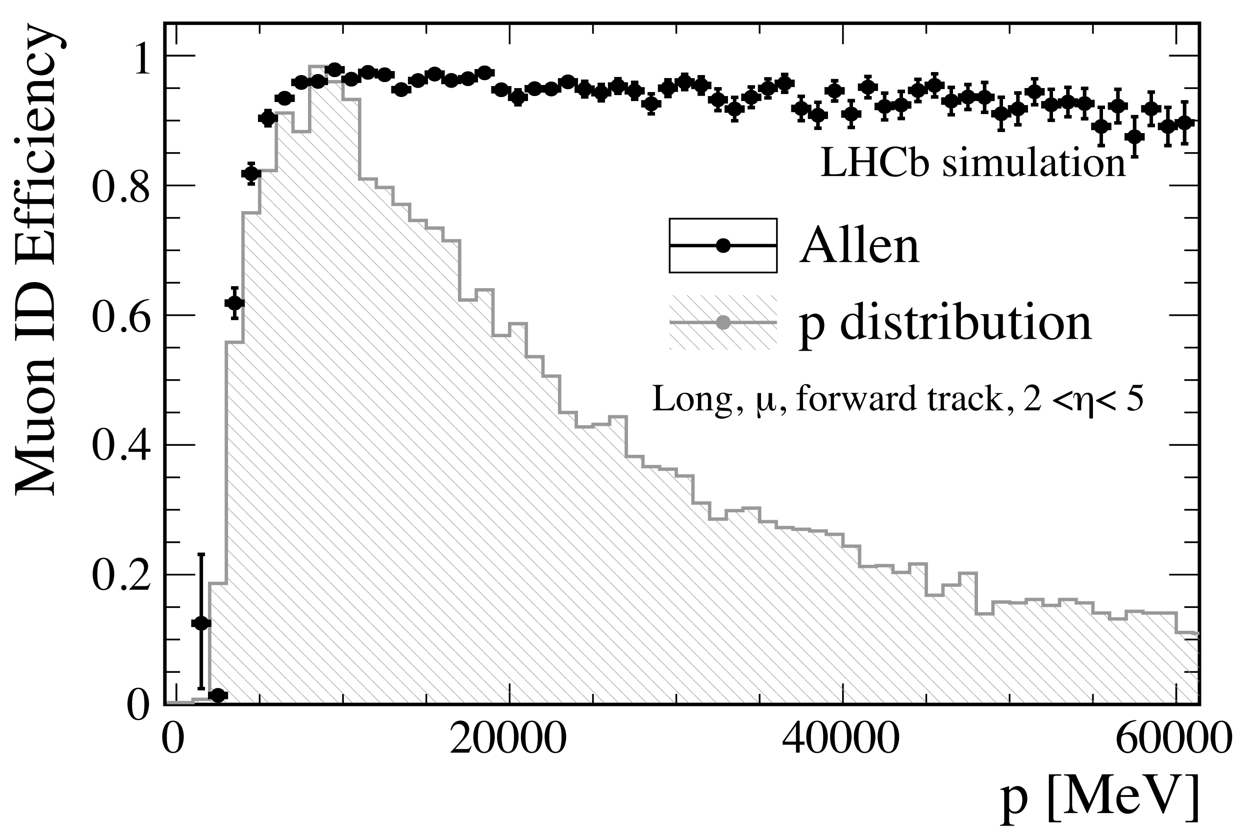


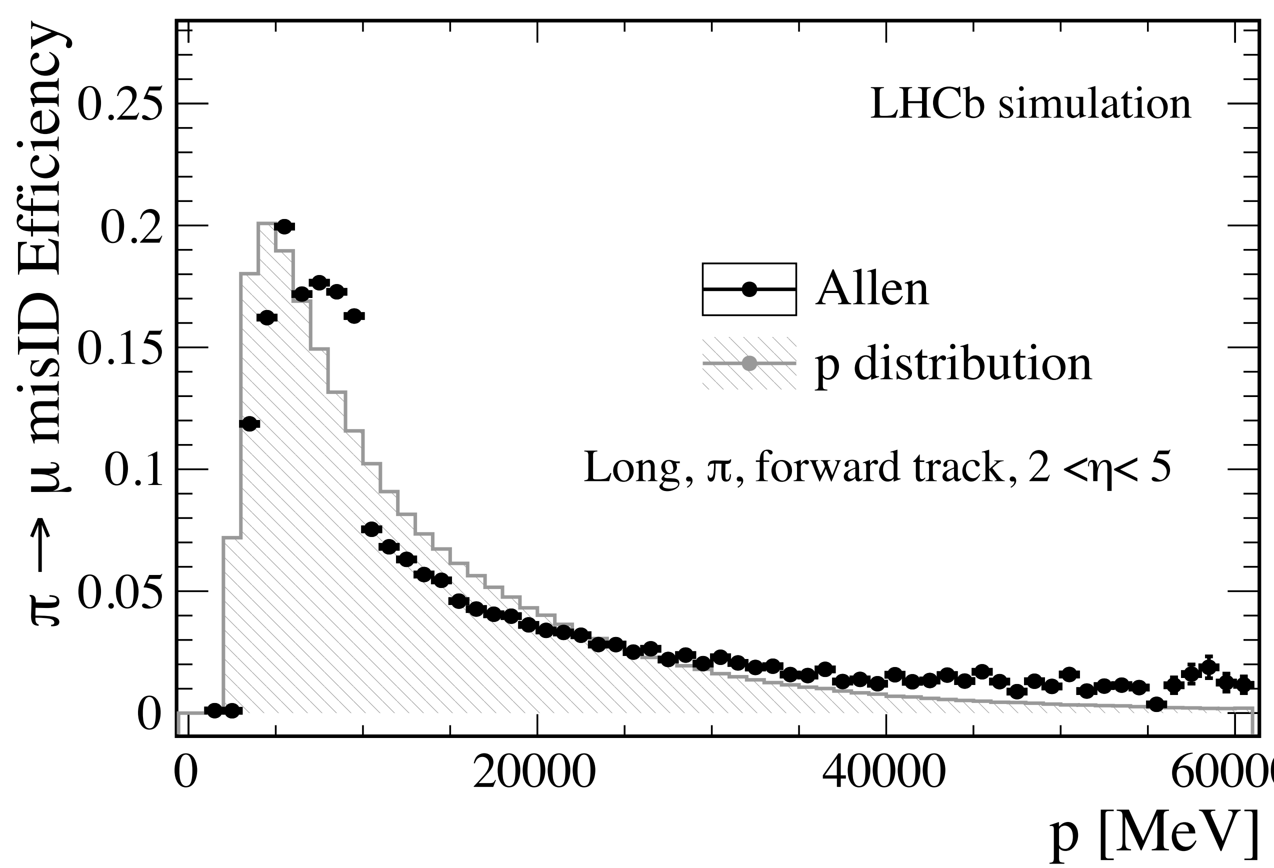


Comput. Softw. Big Sci. 6 (2022) no.1, 1

HLT1 muonID performance

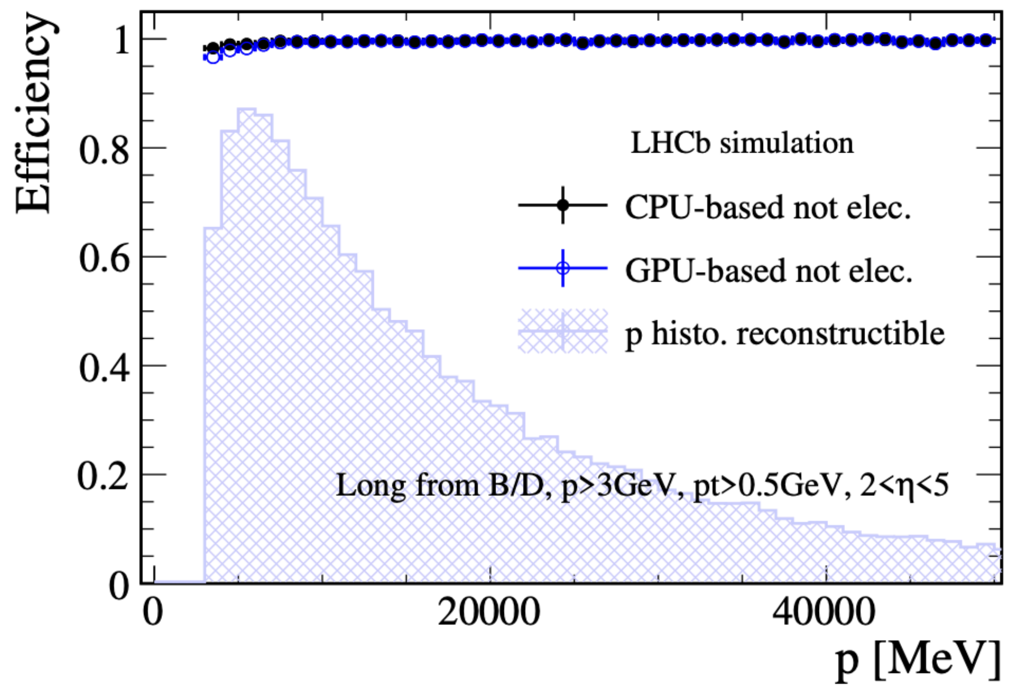
- Excellent muon identification and misID background rejection

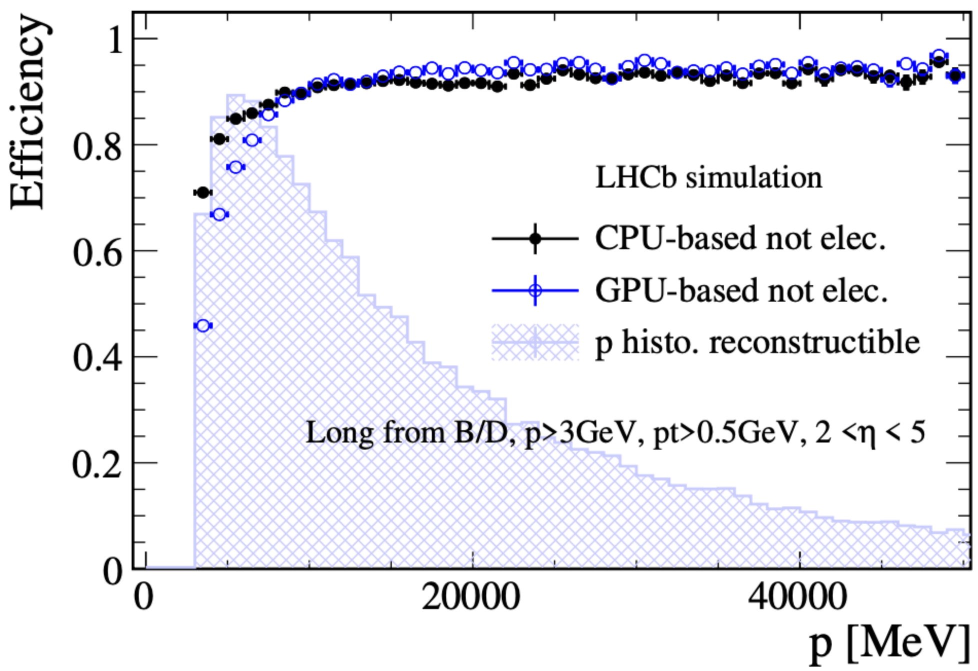


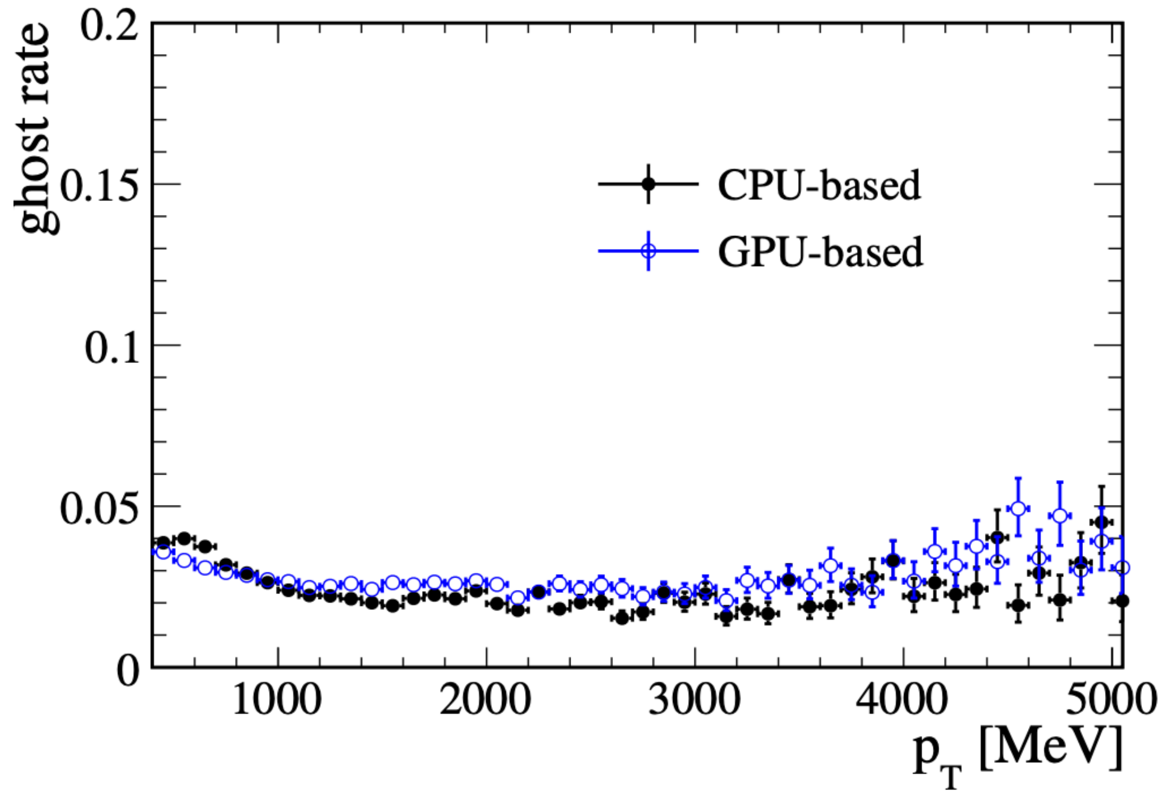


HLT1 tracking performance

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



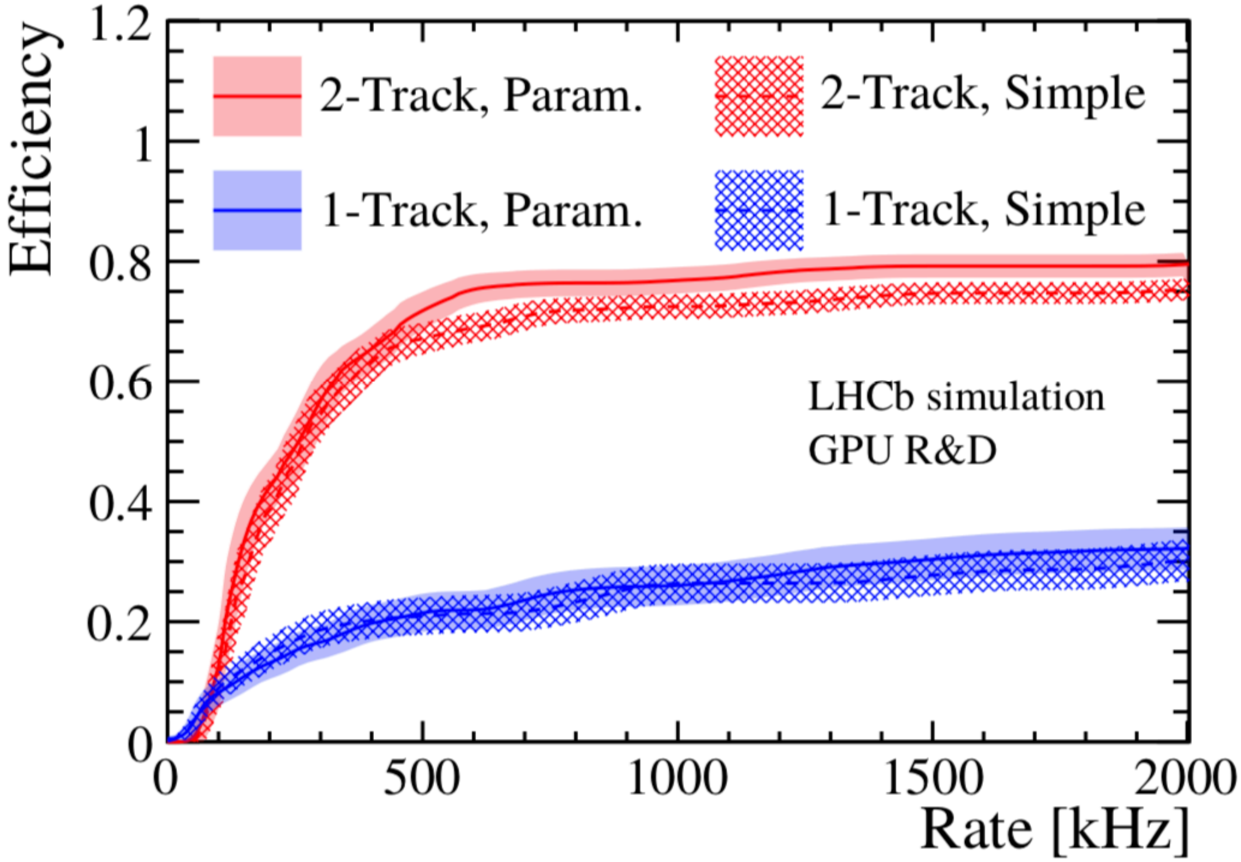




Kalman filter

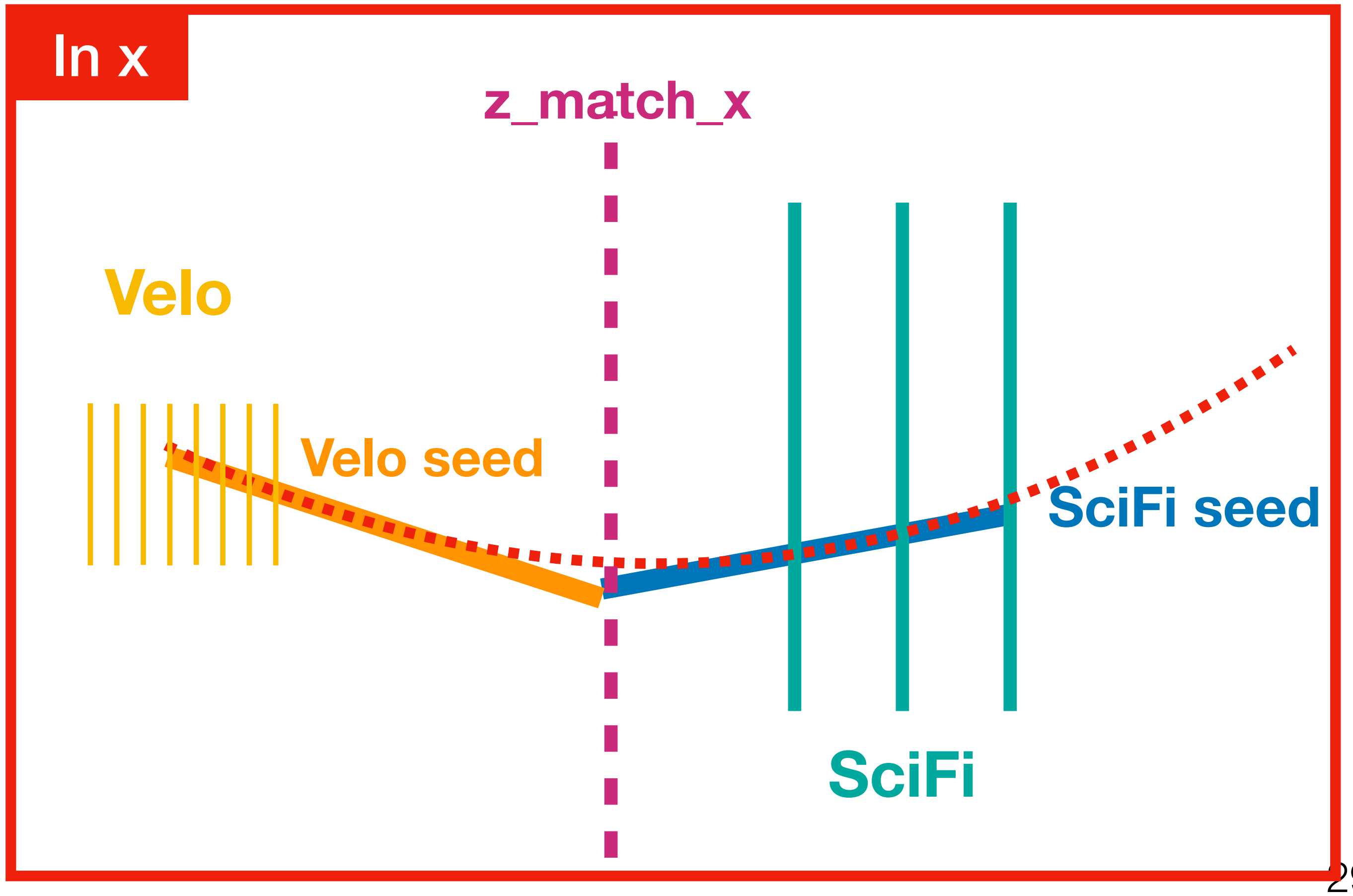
- Improve Impact Parameter (IP) resolution and reduce ghosts
- Nominal LHCb Kalman filter uses Runge - Kutta extrapolator + detailed detector description

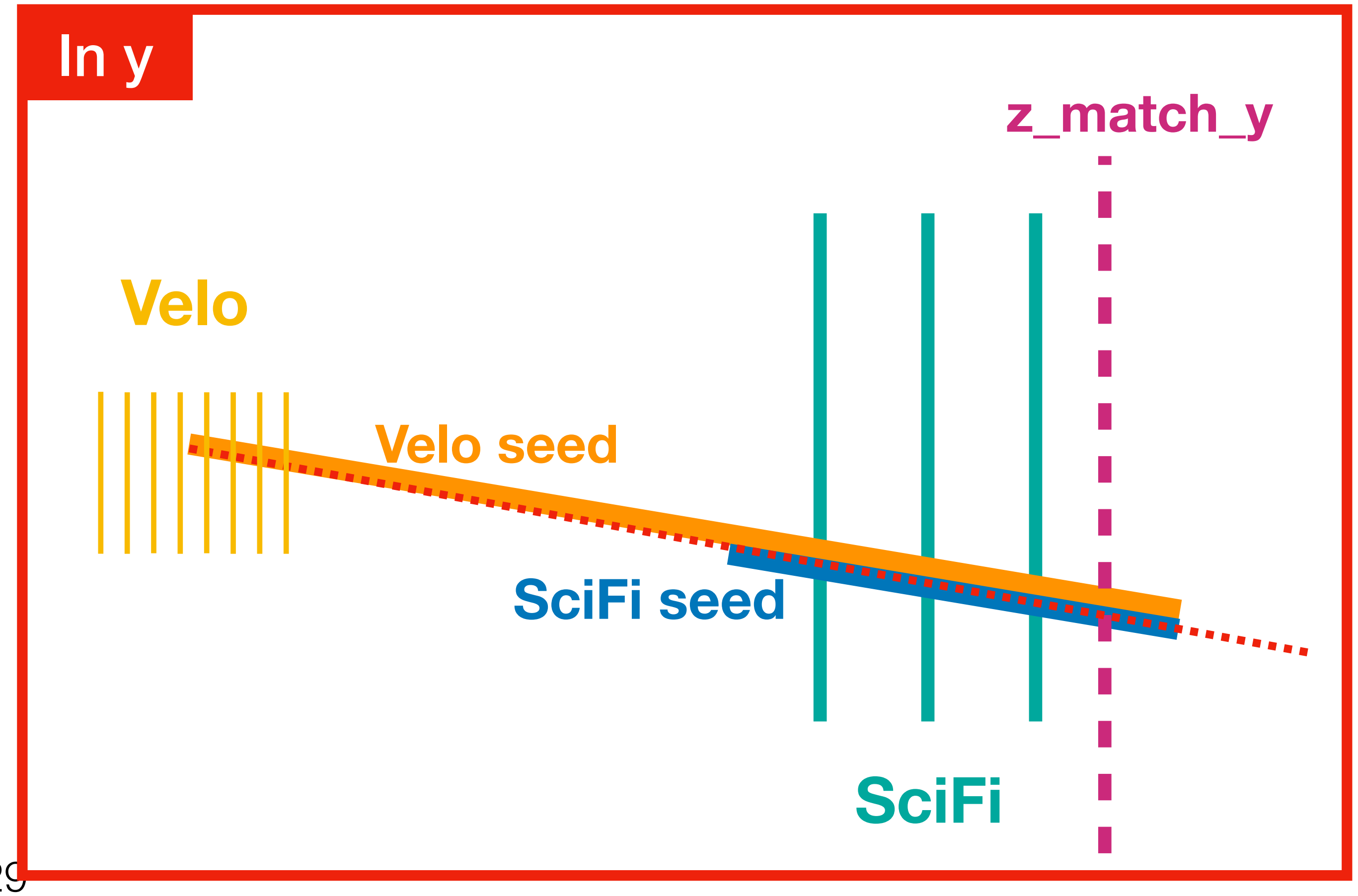
- In HLT1, for performance reason two alternatives based on parametrizations:
    - Full detector Parametrized Kalman Filter
    - Velo-Only Kalman Filter (fits only Velo segment, momentum estimate from full track)
    - IP resolution mostly impacted by Velo measurement -> Velo-Only option chosen, which significantly improves throughput



The track matching algorithm

- Two main inputs: **SciFi** and **VELO** seeds
- Algorithm approach:
  - “Kink” approximation: Velo/SciFi seeds extrapolated to matching position as straight lines
  - Magnetic field and bending in y parametrised with truth simulation to calculate z_match(x,y)



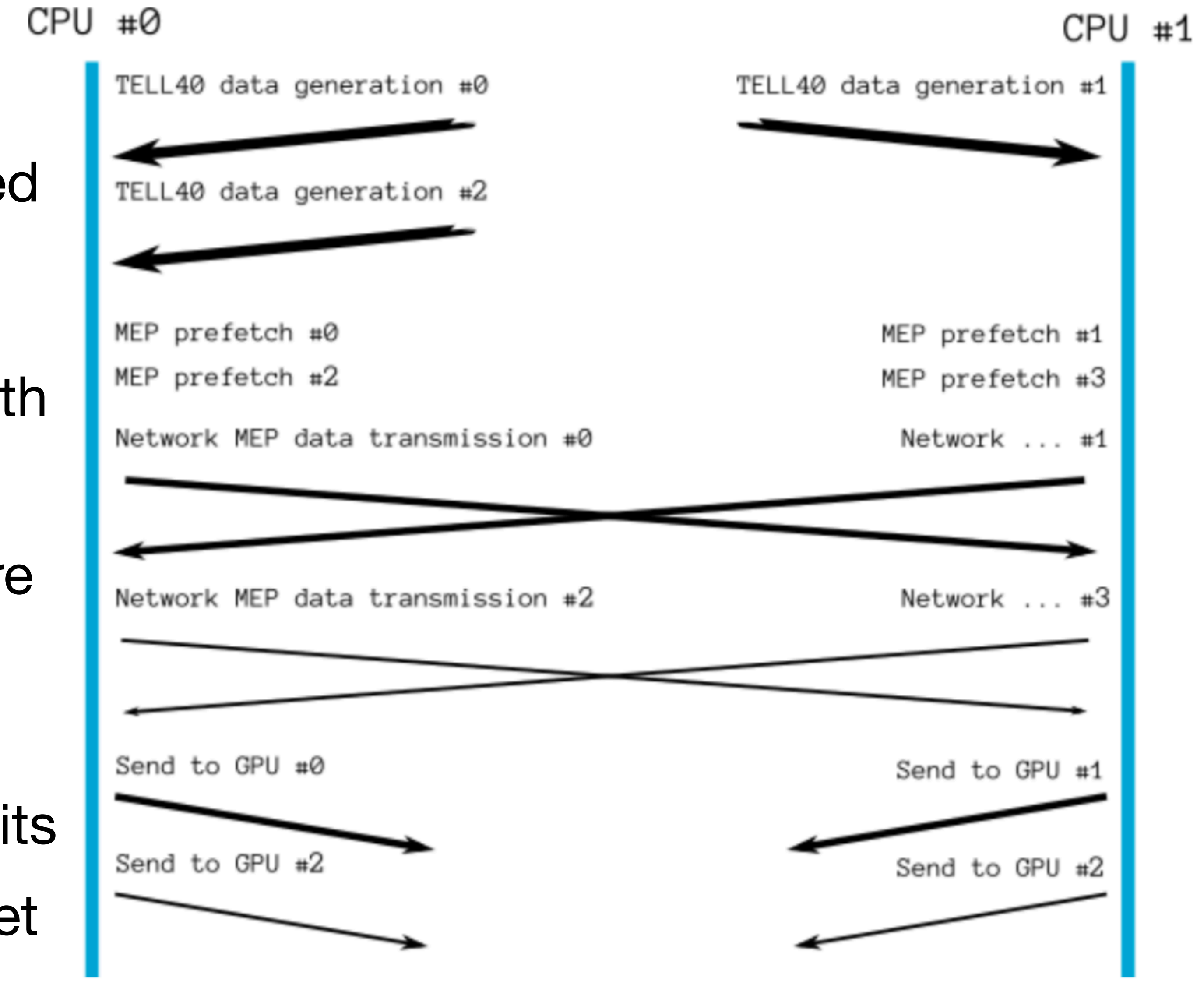


Towards the integration of Allen in the online system

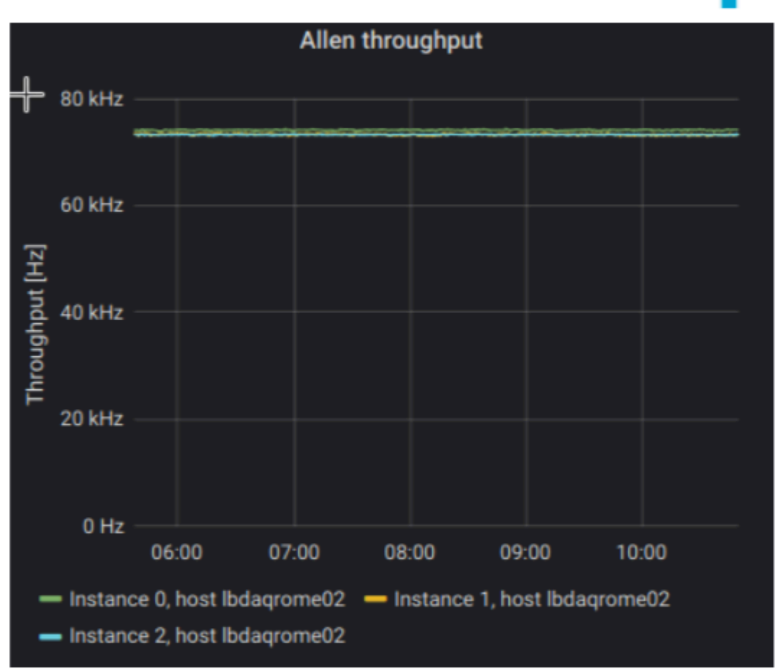
Challenge of fully commissioning Allen: we need the real detectors and EB server first!

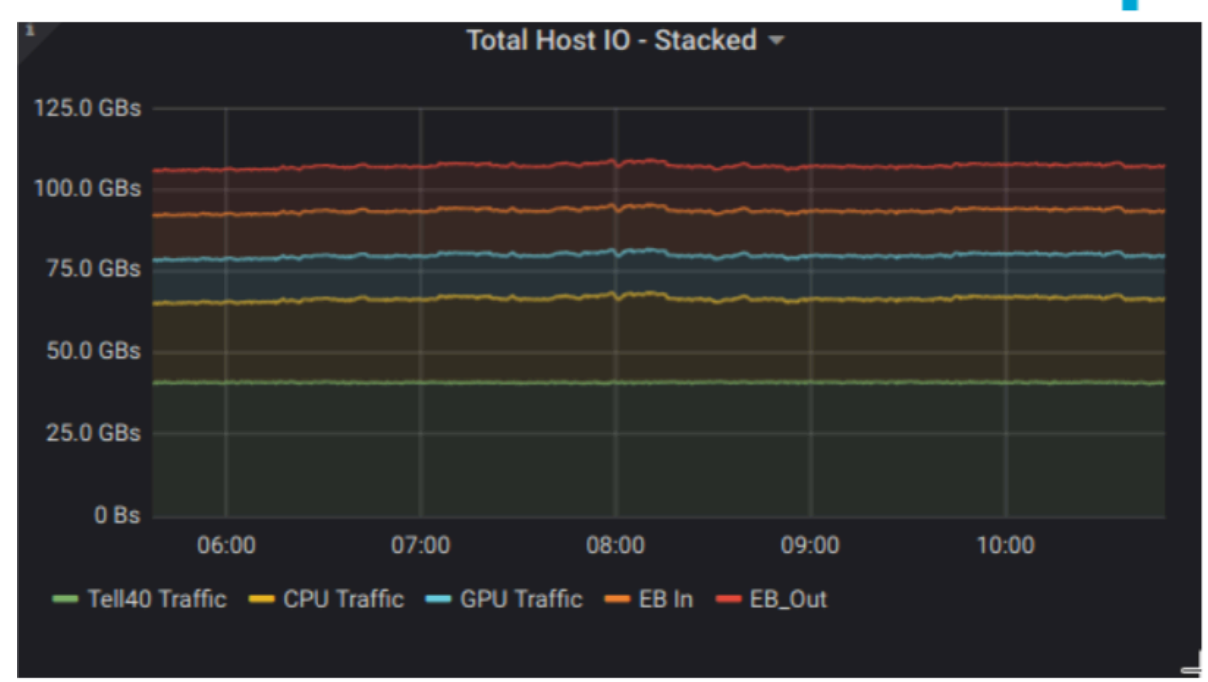
First integration tests in smaller-size servers with pre-loaded simulation data

- 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!





LHCb-TDR-021