Real-time analysis in Run 3 with the LHCb experiment

Peilian Li on behalf of the LHCb collaboration

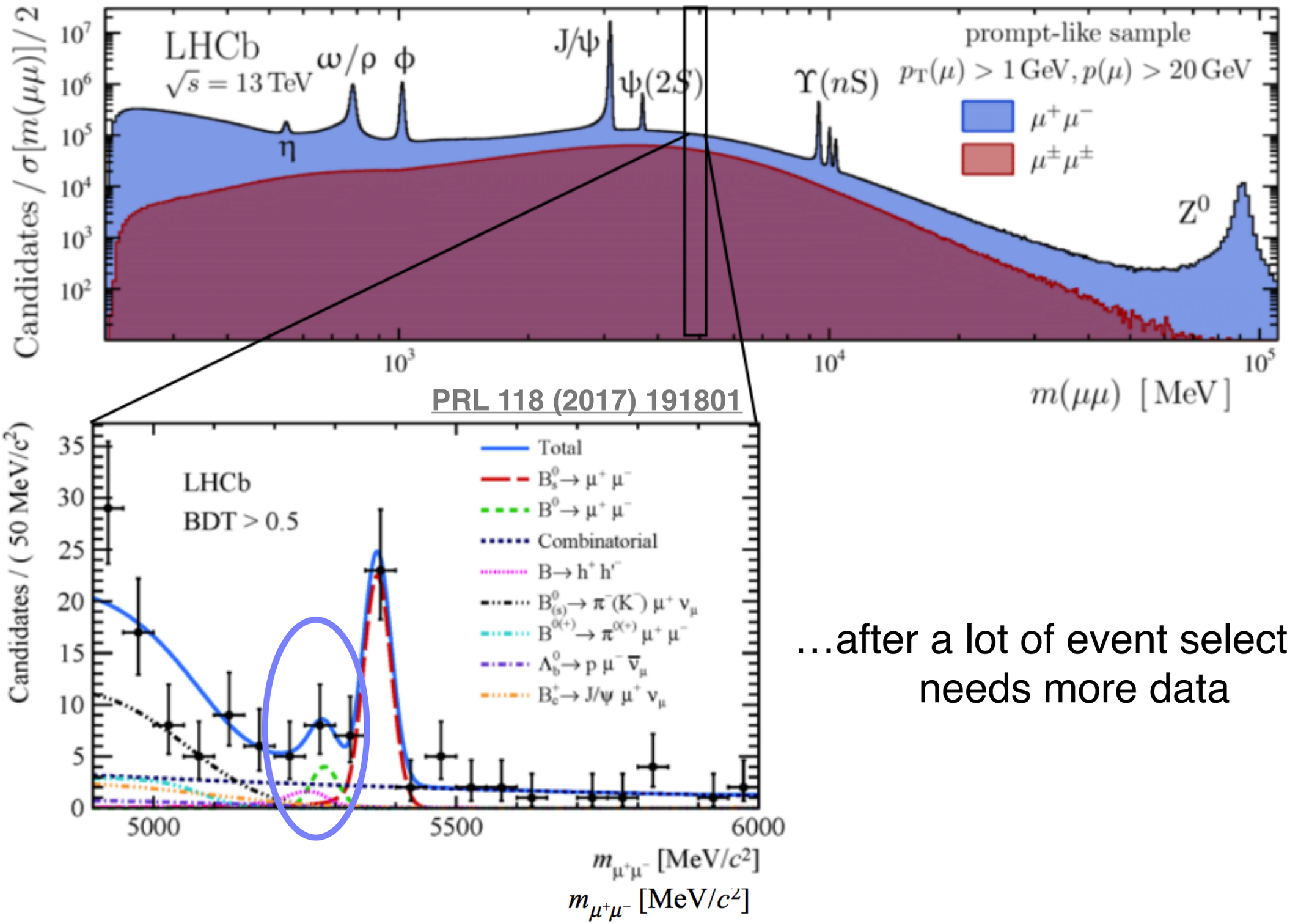
Heidelberg University

July 28th, 2021

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Opportunities in Run 3

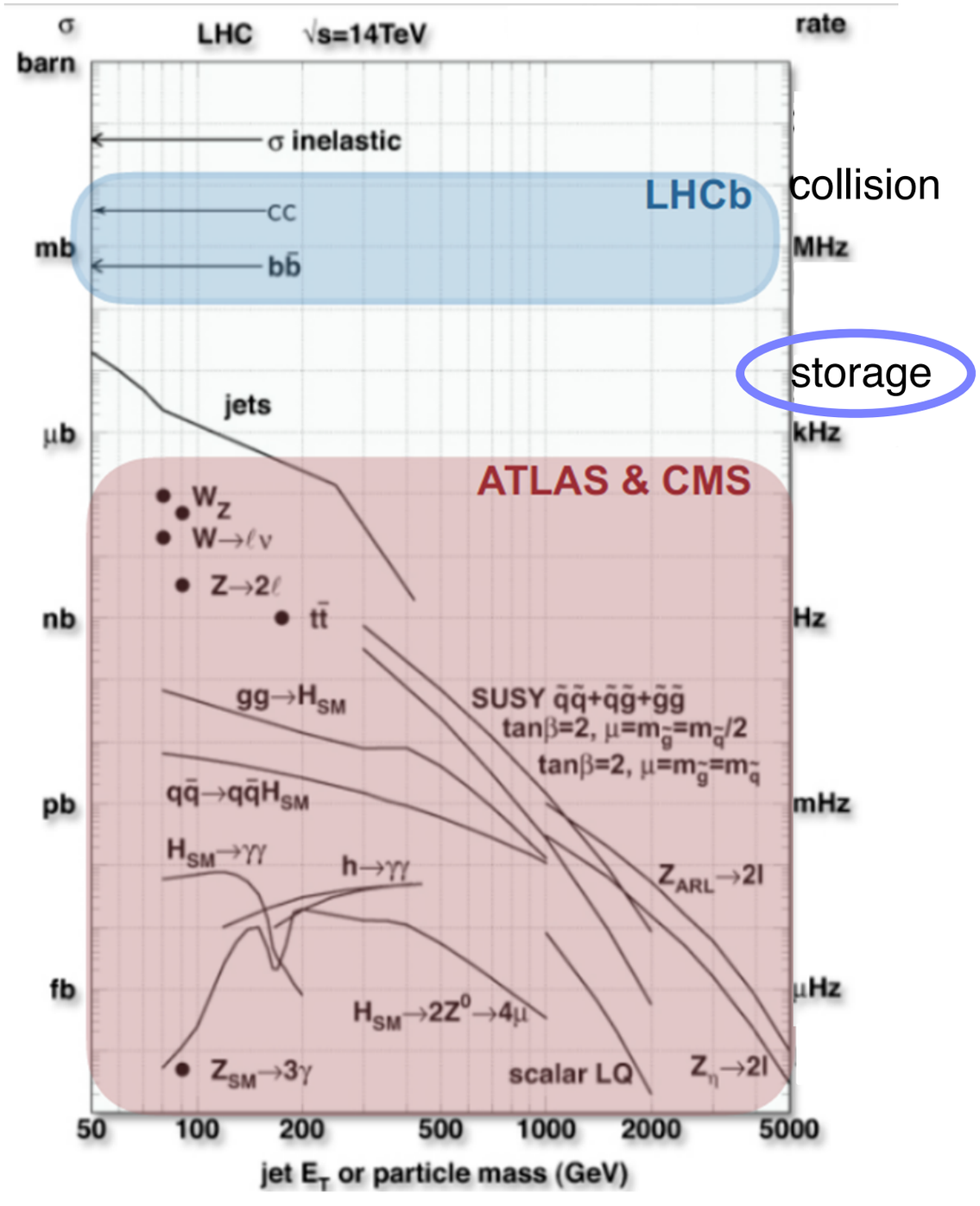
PRL 120 (2018) 061801



…after a lot of event selection, needs more data

LHCb Upgrade

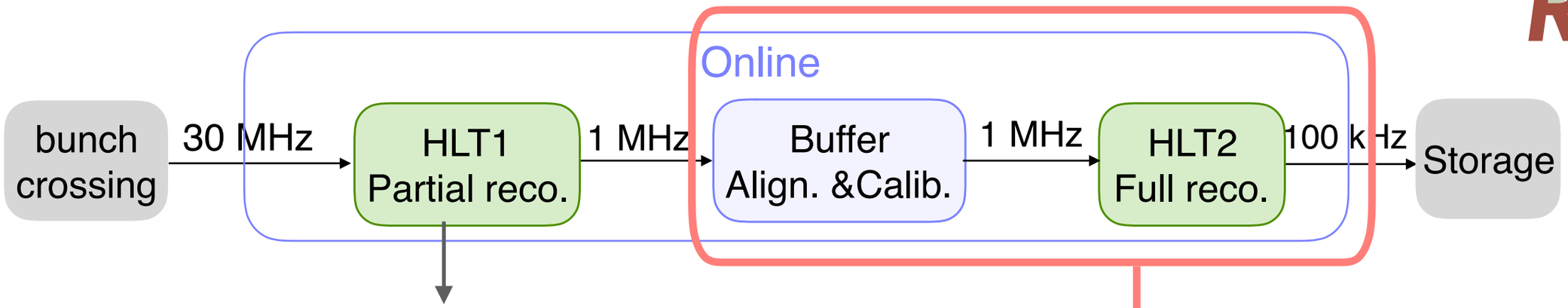
- Luminosity of $2 \times 10^{33}\ cm^{-2}s^{-1}$, $\sqrt{s} = 14$ TeV, visible collisions per bunch $\mu \sim 5$
  - More PVs, more tracks, almost all events will have b or c hadron



- ATLAS/CMS mainly look at the very rare event → ~10 -100 Hz of event rate
- LHCb is interested in b and c hadrons → about 10^3 times higher event rate (MHz), 100 times higher than storage

storage is tight bottleneck

- Pioneering role of LHCb in real time analysis & novel storage concepts



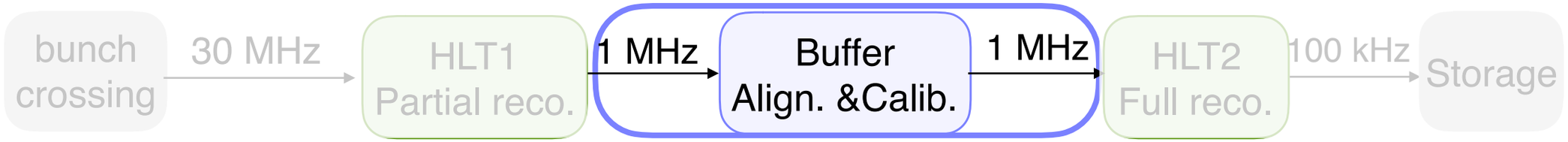
See dedicated talk by Marianna Fontana

Run1&2 trigger:
background rejection

Upgrade trigger:
background rejection & signals classification

Real-time Alignment & Calibration

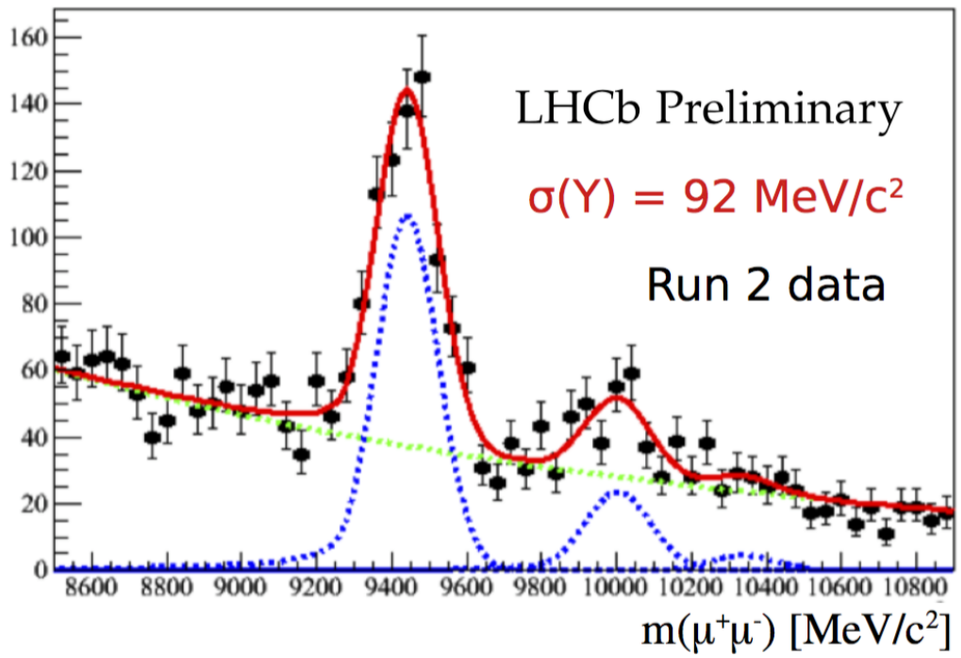
Giulio Dujany et al 2015 J. Phys.: Conf. Ser. 664 082010



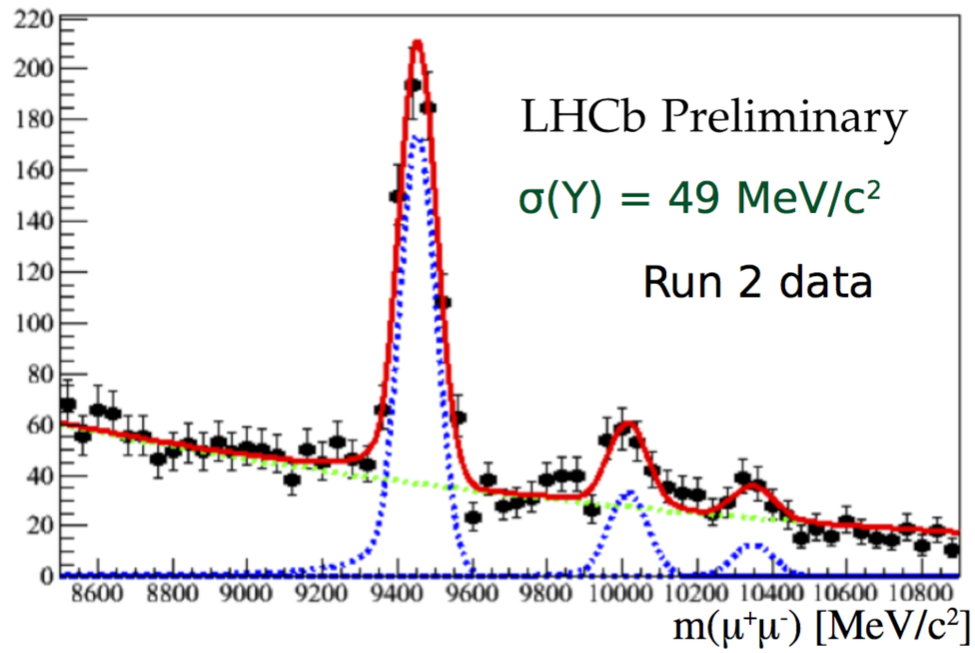
- Efficient and pure selections require offline-quality reconstruction at the HLT2 level ← Aligned and calibrated detector
  - Better mass resolution
  - Better particle identification
  - Less background

→ Use limited storage more efficiently

Before alignment



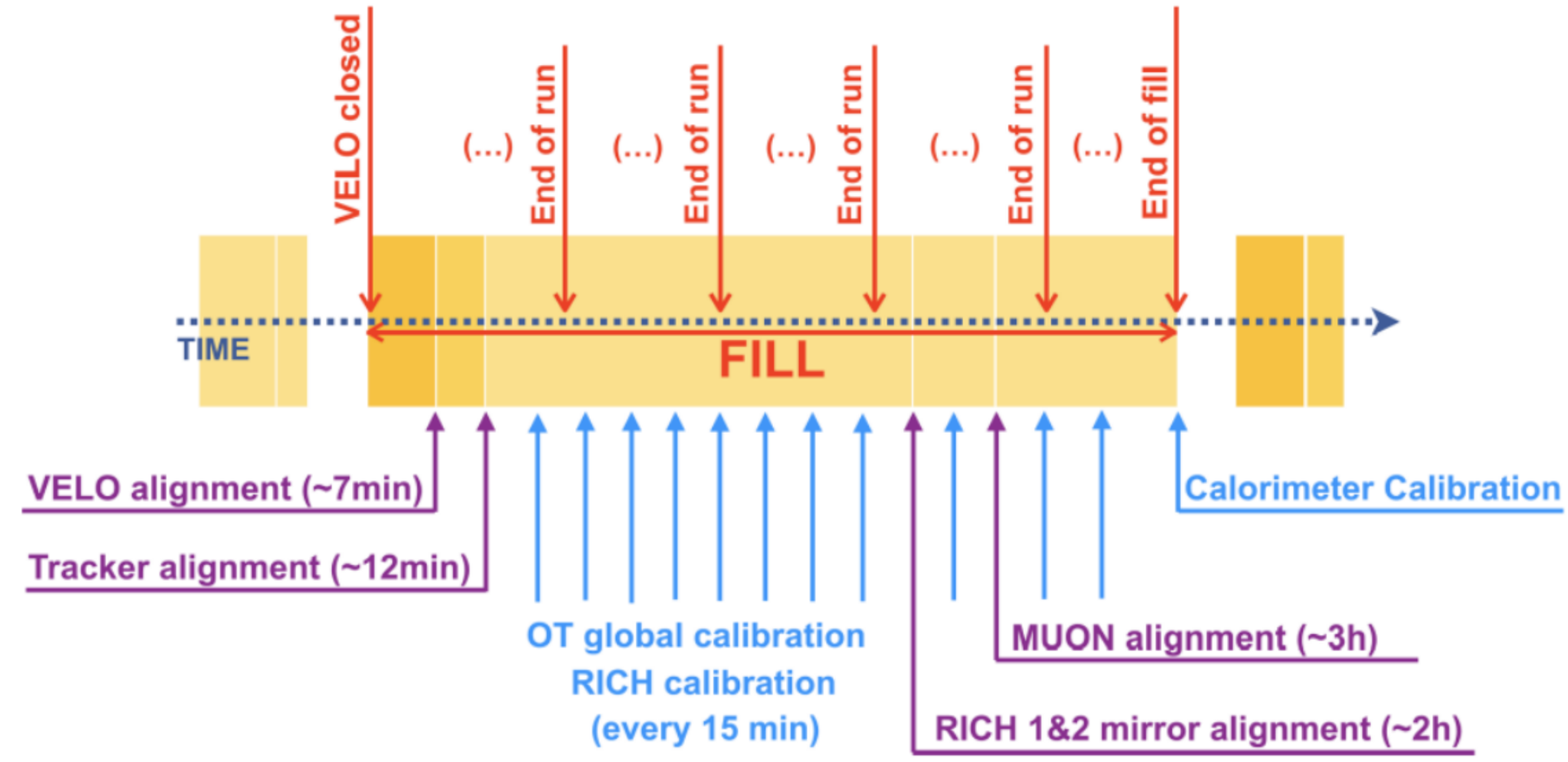
After alignment



Real-time Alignment & Calibration

See poster by Florian Reiss

- Same disk buffer as Run 2 but 10x higher data rate
  - Alignment not only trackers but also RICH, MUON, CALO
  - Should be very fast!



Second High Level Trigger (HLT2)

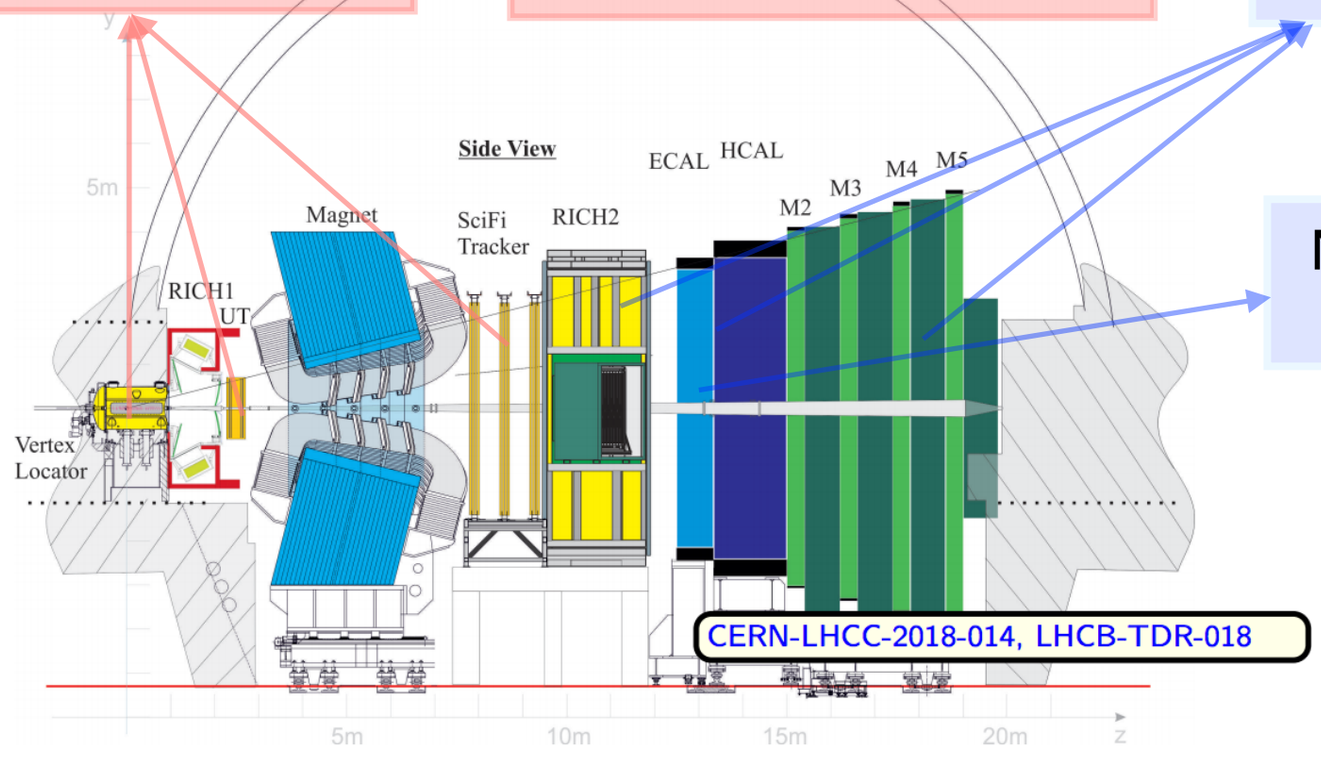
- HLT2 reconstruction is critical to both physics output and physics quality
  → Full, offline-fidelity event reconstruction on at least 1 MHz

Charged track reconstruction in full momentum range: VELO, UT, SciFi

Full track fit: best resolution, background rejection

Particle identification: RICH, MUON, ECAL

Neutral reconstruction: ECAL



Current HLT2 Throughput

LHCb-FIGURE-2021-003

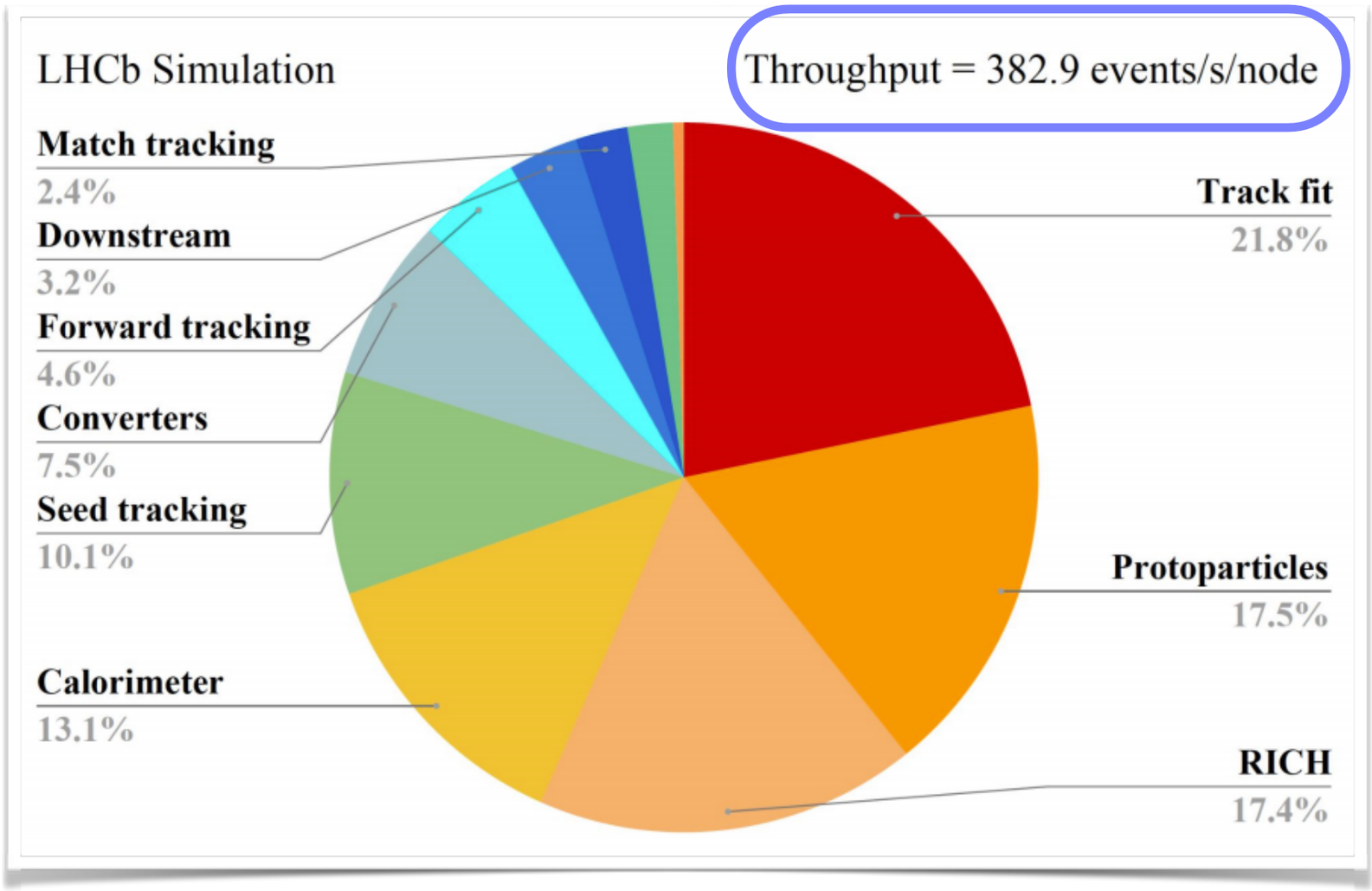
- Significant achievement in the past years (~ 2.5x higher throughput)

  → Vectorize the algorithms with structure-of-array & remove redundancy in the track reconstruction JINST 15 (2020) 06, P06018

  → Apply track fit only to these tracks needed by physics analyses

  → Selective track-calo matching

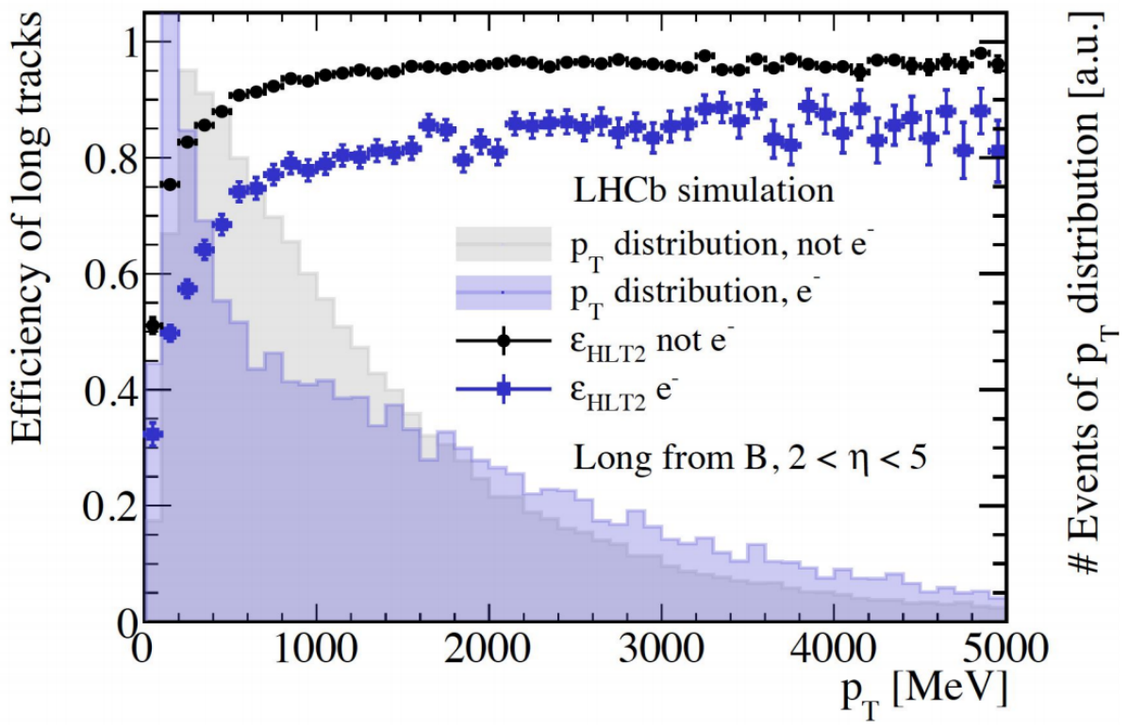
  → **Caveat**: Selection and saving of trigger candidates not accounted yet

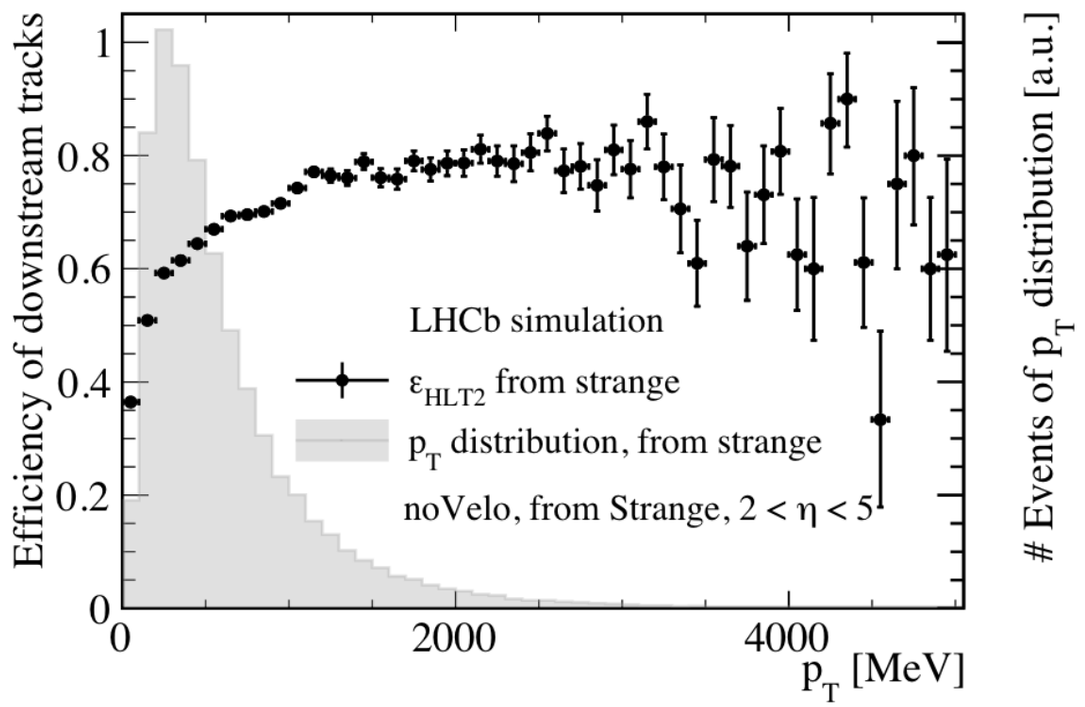


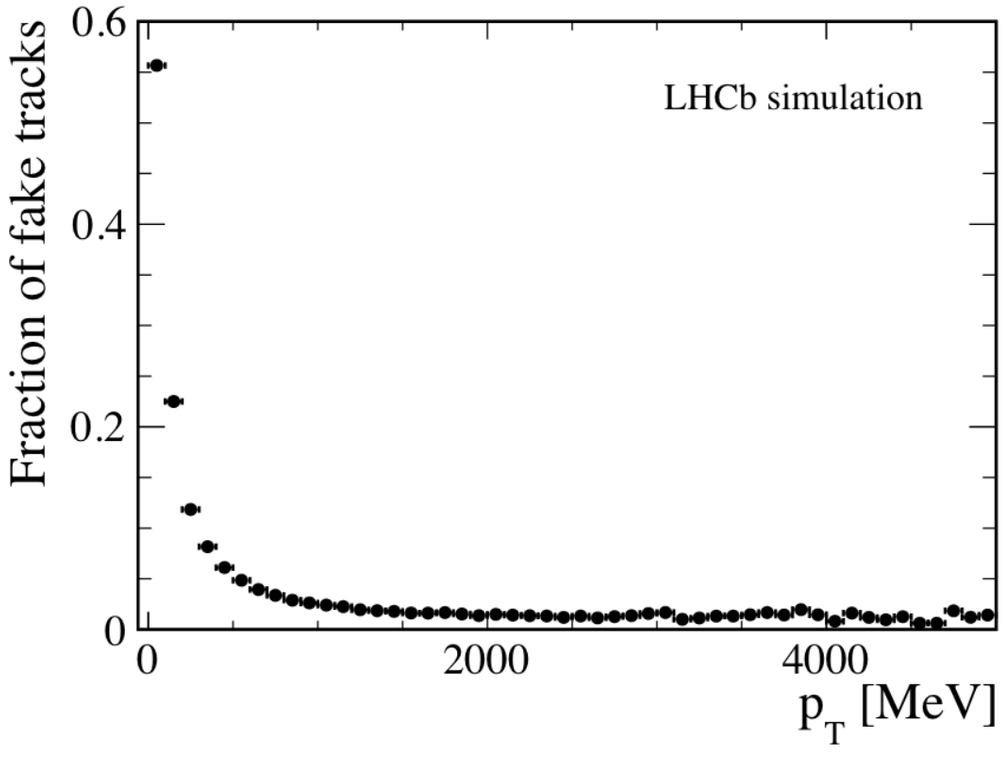
Allowed maximum HLT1 output rate: ~2 MHz

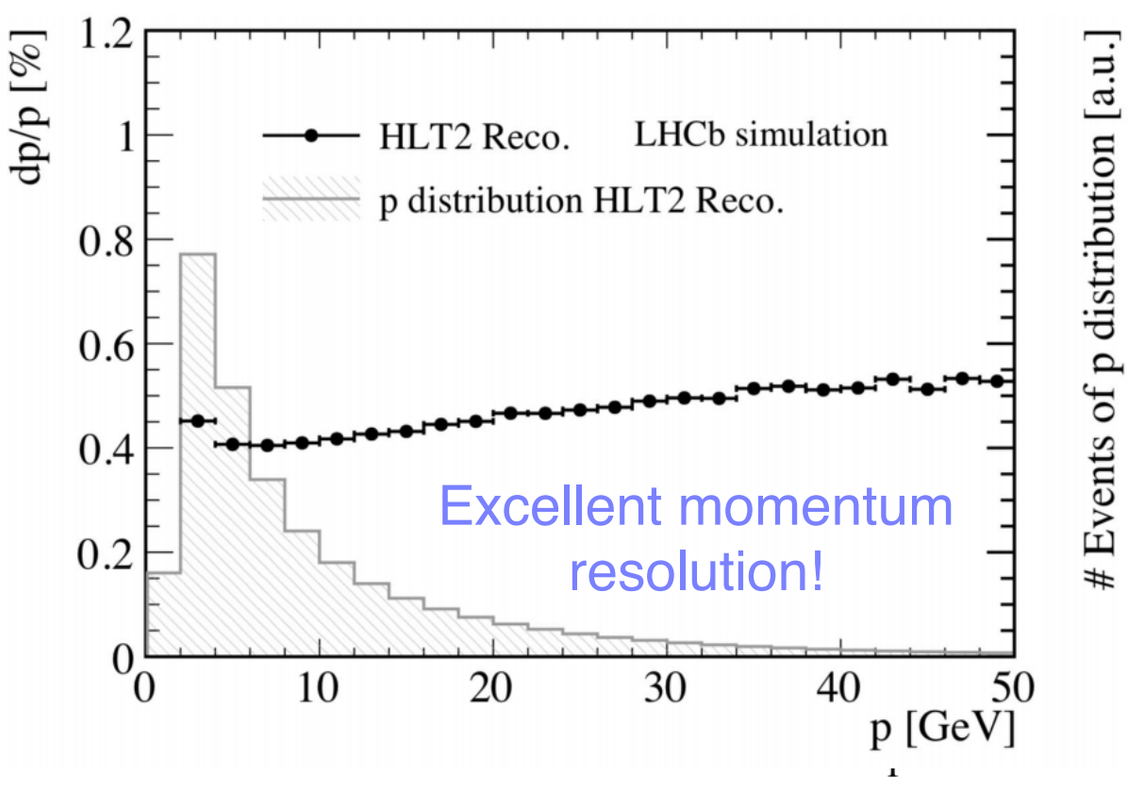
Physics Performances - Tracking

LHCb-FIGURE-2021-003





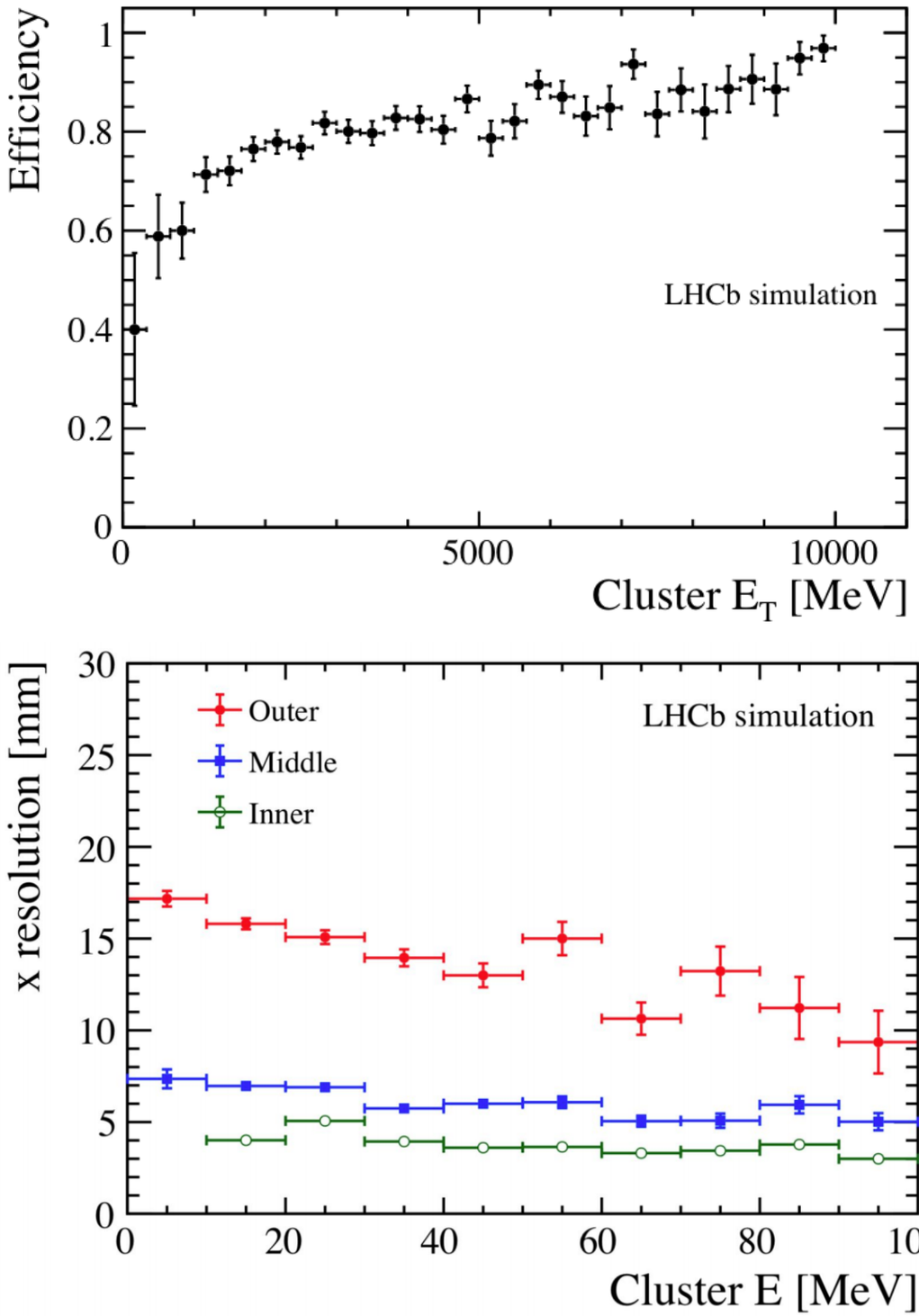




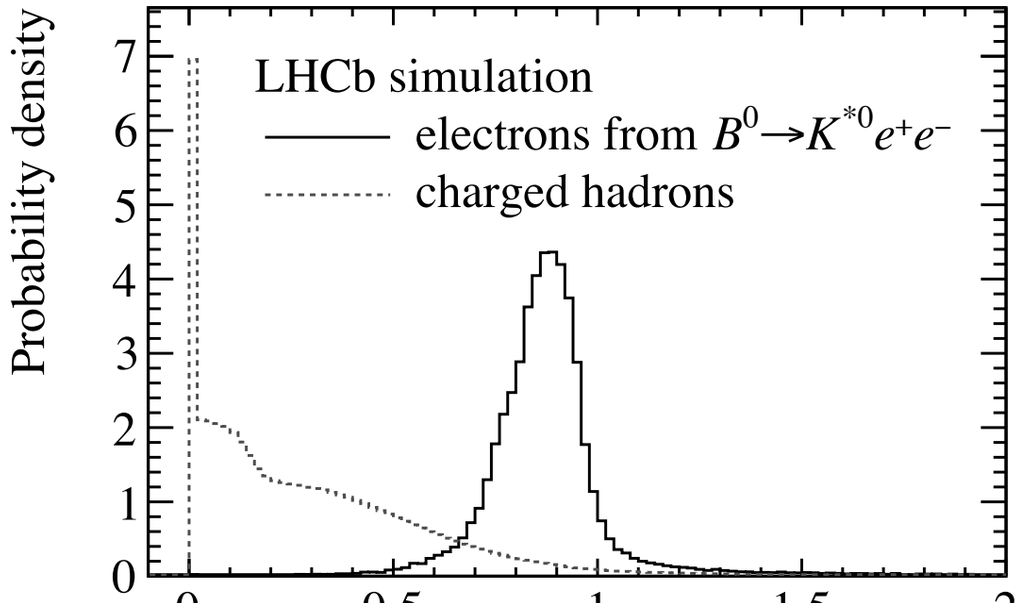
Physics Performances - ECAL

LHCb-FIGURE-2021-003

MC Samples: B → K*γ



MC Samples: B → K*e+e-

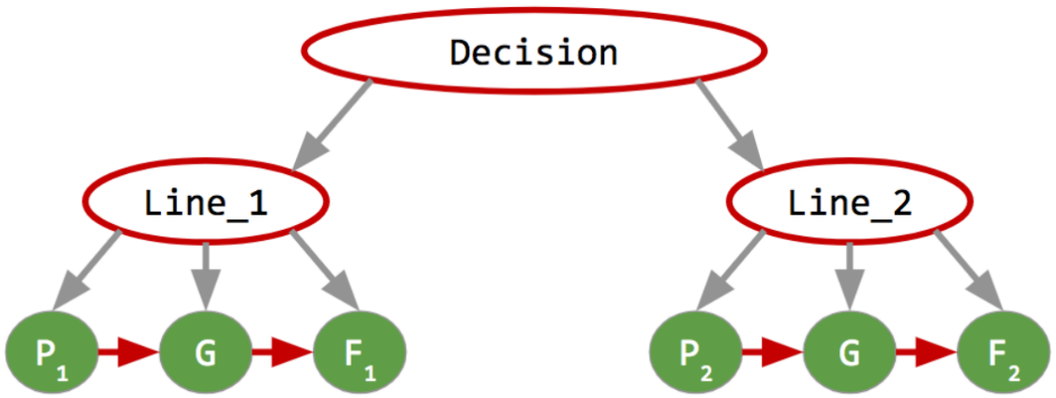


- Efficiency of reconstructing ECAL clusters from energy deposited by photons
- Cluster resolutions in three ECAL regions (different size of cells)
- Electron identification

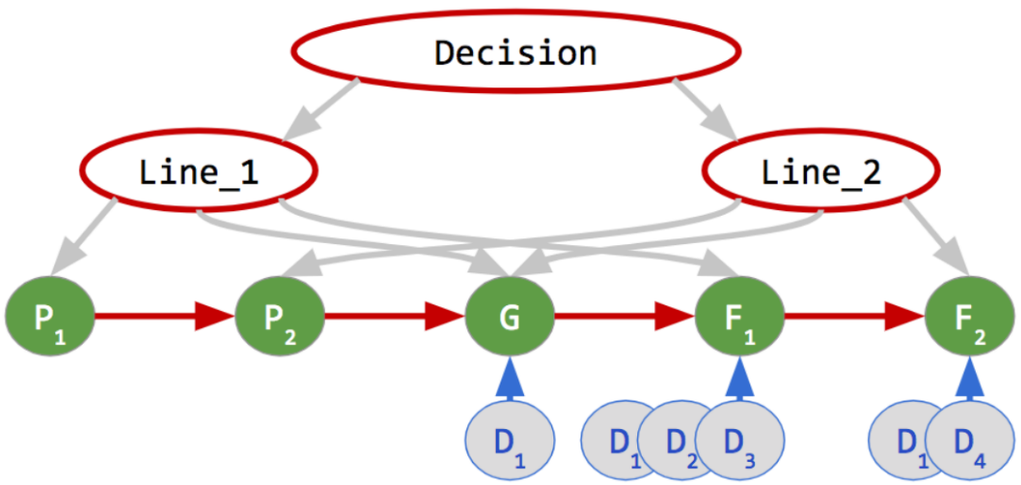
Managing O(1000) HLT2 Selections

LHCb-Proc-2020-003

Example dependency tree with two lines



Configuration during initialization



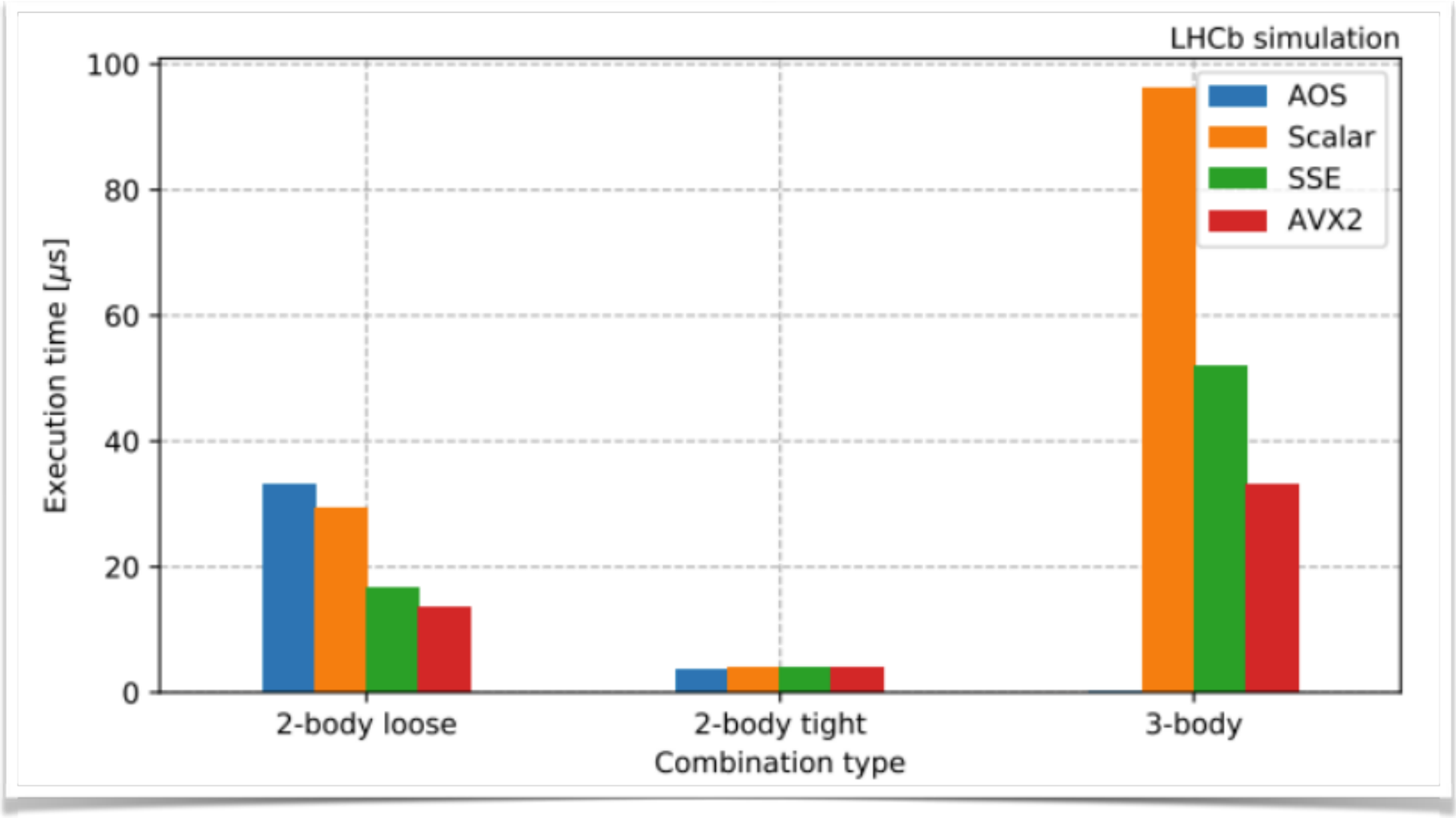
- New algorithms scheduler
  - multithreading friendly & **automatically** handle the data flow with specific logical types
  - **order the basic nodes** with specific control flow
  - **resolve data dependencies**

Static graph with **ordered nodes**

Managing O(1000) HLT2 Selections

LHCb-FIGURE-2020-018

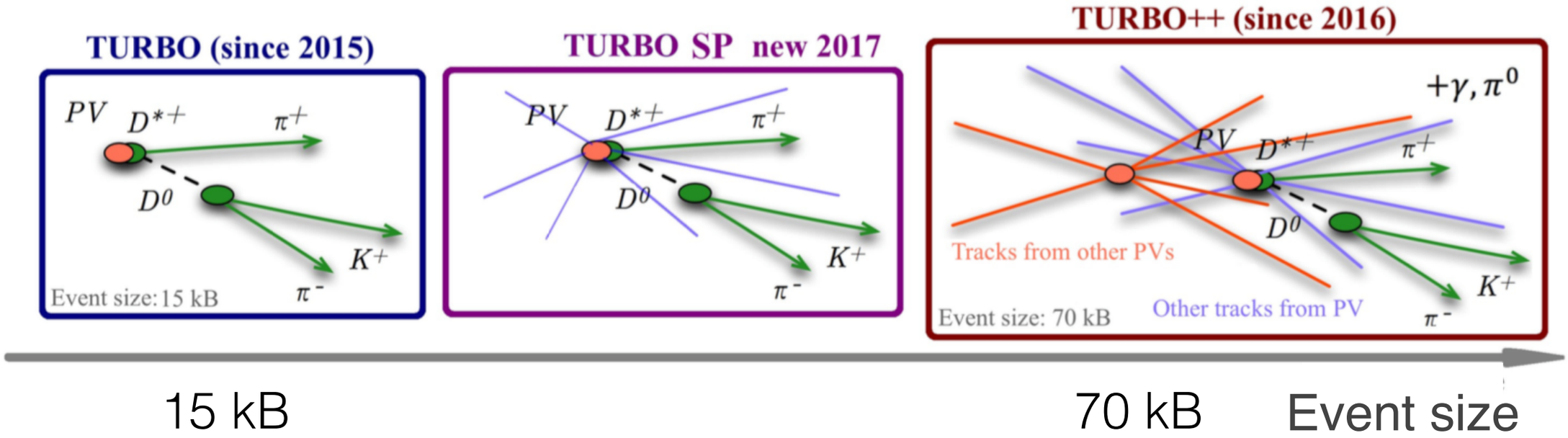
Vectorized selections ~ 65% gains up



* 2-body tight algorithm has tight selections on its input objects and subsequently cannot fully fill the vector instruction registers during execution

Persistency Model: what is saved to the disk

- Fixed output bandwidth of 10 GB/s
- Bandwidth [MB/s] ~ Trigger output rate [kHz] × average event size [kB]



- Turbo stream developed and commissioned in Run 2 as baseline for Run 3
  - Reduced event format: throws away the raw event info & reduces event size by saving only objects needed for physics analyses
    → Higher persistence efficiency for the same bandwidth

See offline data processing talk by Nicole Skidmore

Summary

- ✓ LHCb is almost ready to face the MHz signal era, changing the trigger paradigm and pioneering in the real time processing
- ✓ From background rejection → signal selection and characterization
- ✓ Full offline-quality event reconstruction in the trigger stage
- ✓ Turbo-mode with reduced event size will dominate the selective persistency in the upgrade

🤞lots of physics

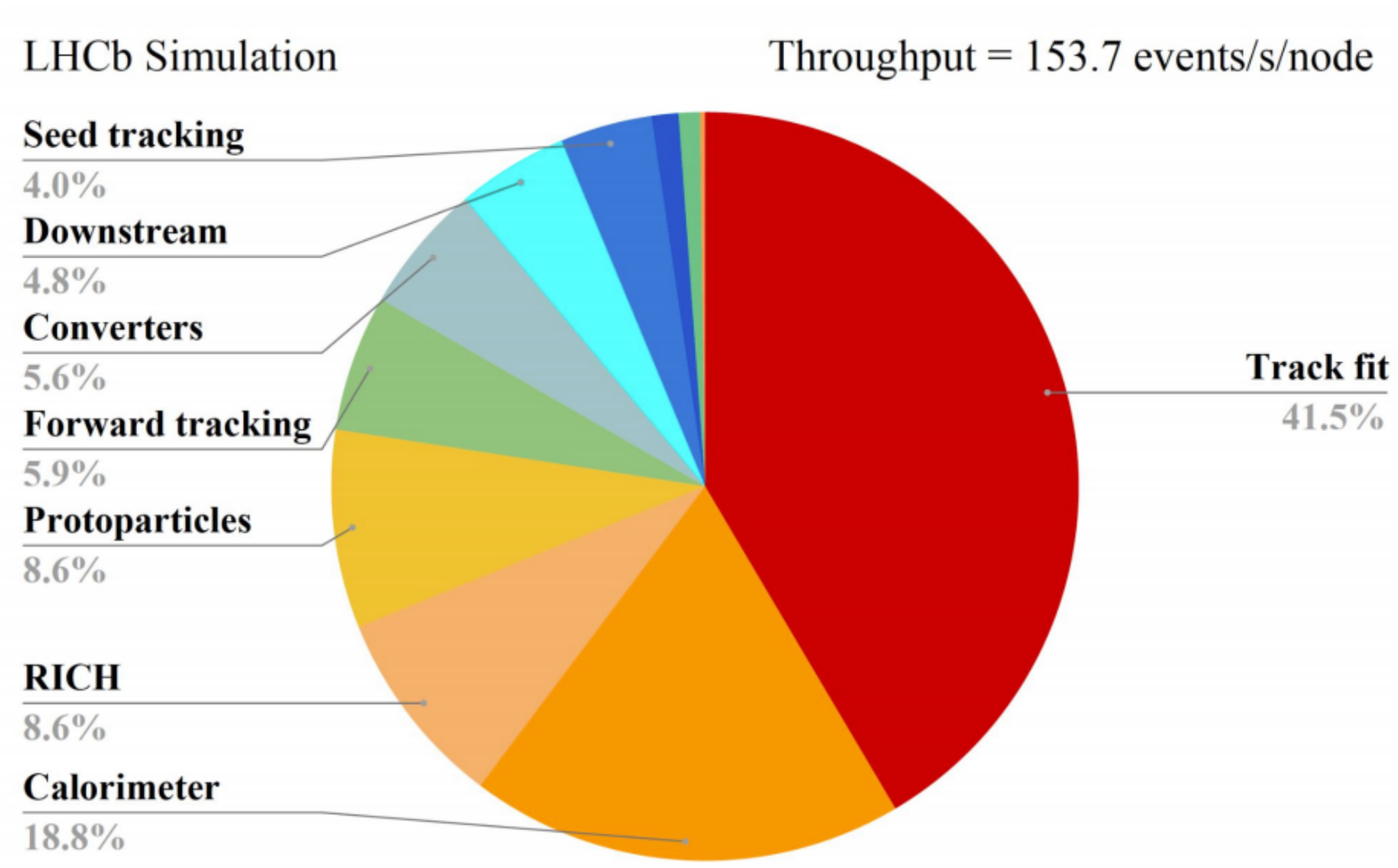
😬lots of data to process

Ambitious real time analysis scheme is on track thanks to huge effort and innovation with the design of software and algorithms!

Thanks for your attention!

Back Up: Current HLT2 Throughput

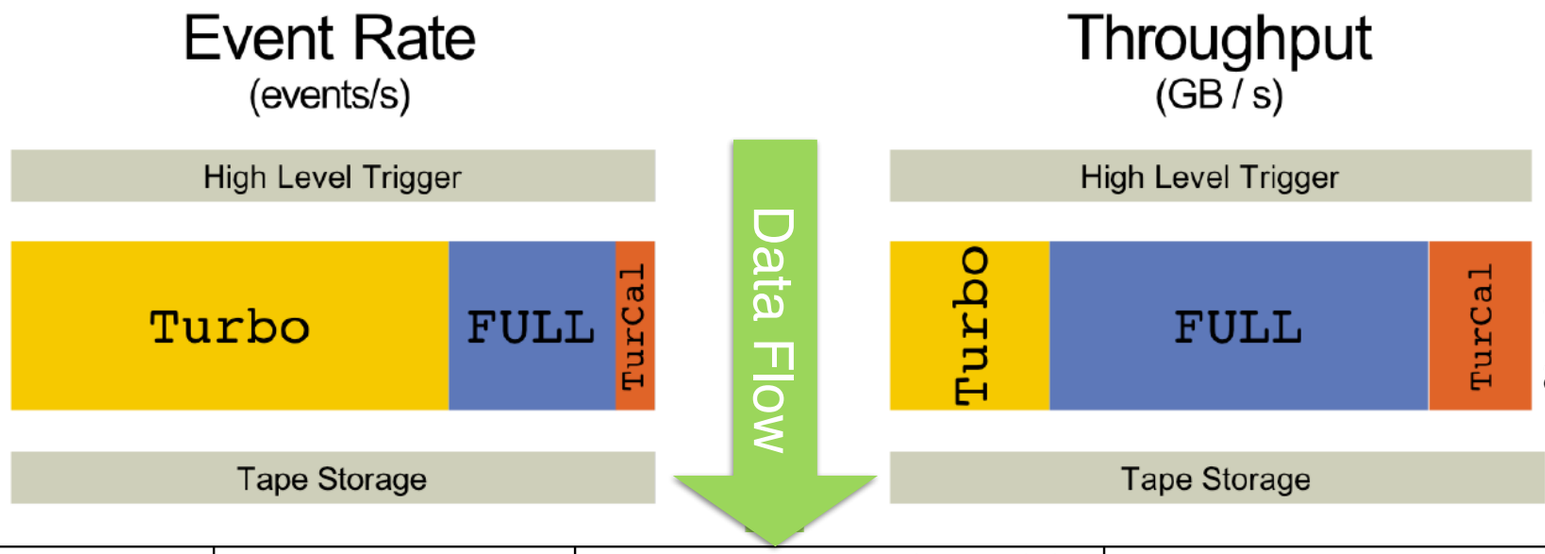
LHCb-FIGURE-2020-007



Allowed maximum HLT1 output rate: ~0.83 MHz

Back Up: Persistency Model

LHCB-TDR-018



= event rate × average event size

stream	rate fraction	throughput (GB/s)	bandwidth fraction
FULL	26%	5.9	59%
Turbo	68%	2.5	25%
TurCal	6%	1.6	16%
total	100%	10.0	100%

- Turbo stream: only HLT2 signal candidates (minimum output)
  - Optionally: (parts of) pp vertex (e.g. “cone” around candidate for spectroscopy)
  - Limitations: cannot refit tracks and PVs offiline, rerun flavor tagging etc.
  - Advantage: Event size O(10) smaller than RAW
- Full stream: all reconstructed objects in the event + selected RAW banks
- TurCal stream: HLT2 candidates and RAW banks
  - Used for offline calibration and performance measurement