High level triggers of LHC Run-3 and their performance on hybrid computing architectures

XXV DAE-BRNS HEP Symposium 2022

IISER Mohali December 12th - 16th

Brij Kishor Jashal,

Tata Institute of Fundamental Research, Mumbai.

Instituto de Física Corpuscular (IFIC), Valencia.

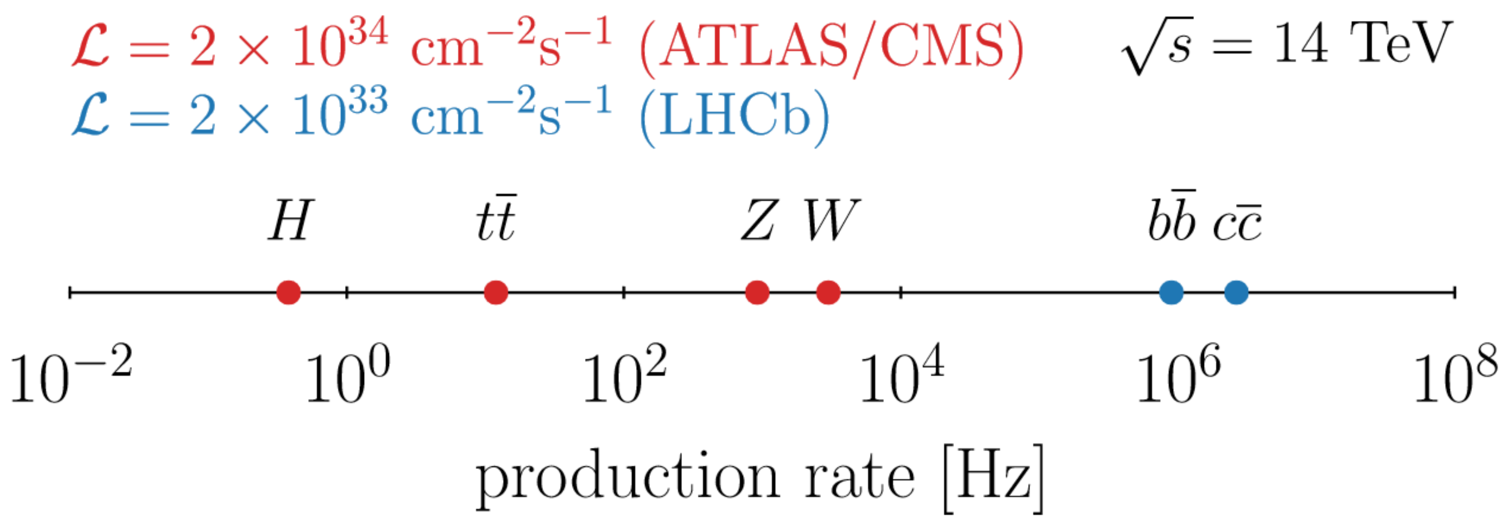
Outline

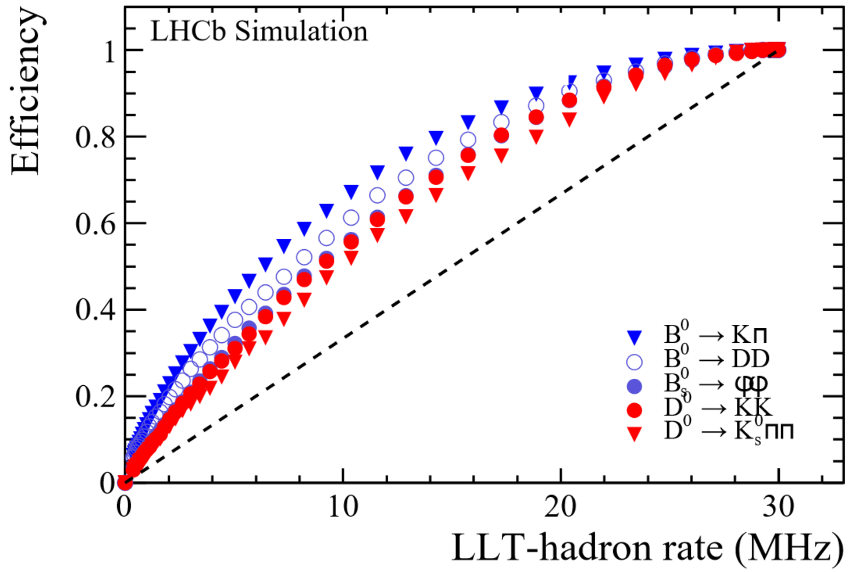
- Introduction
- Trigger systems for Run-3
- Commissioning and performance of LHCb and CMS HLTs
- Conclusion and prospects

Trigger strategy

[REF-LHCb-PROC-2022-010]

Trigger strategy for each experiment depends on the underlying physics program.





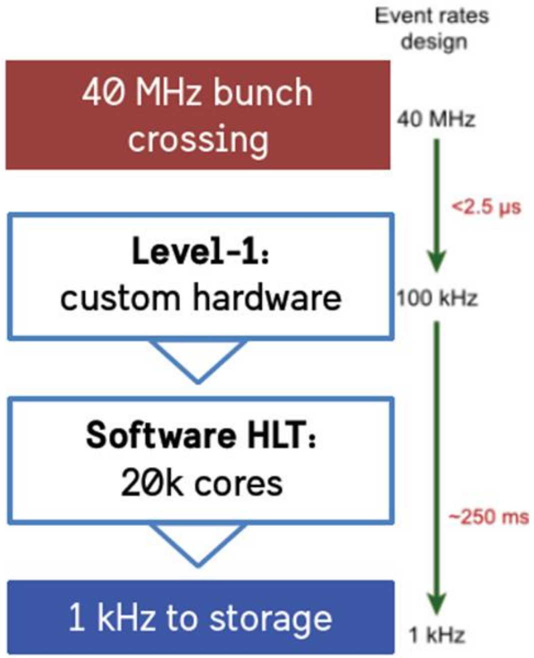
[CERN-LHCC-2014-016]

Triggering ➔ Real Time Analysis

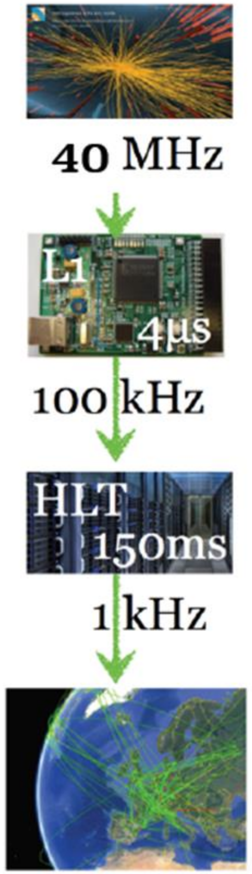
- Can't efficiently trigger on heavy flavor decays using hardware signatures
- Solution: process every event (30 MHz, 5 TB/s) in software

Run-3 Trigger systems

ATLAS

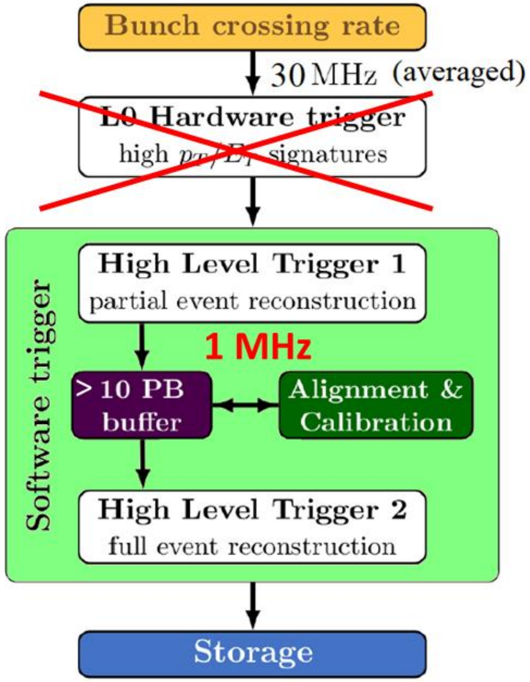


CMS



LHCb

(Run3: 2022)



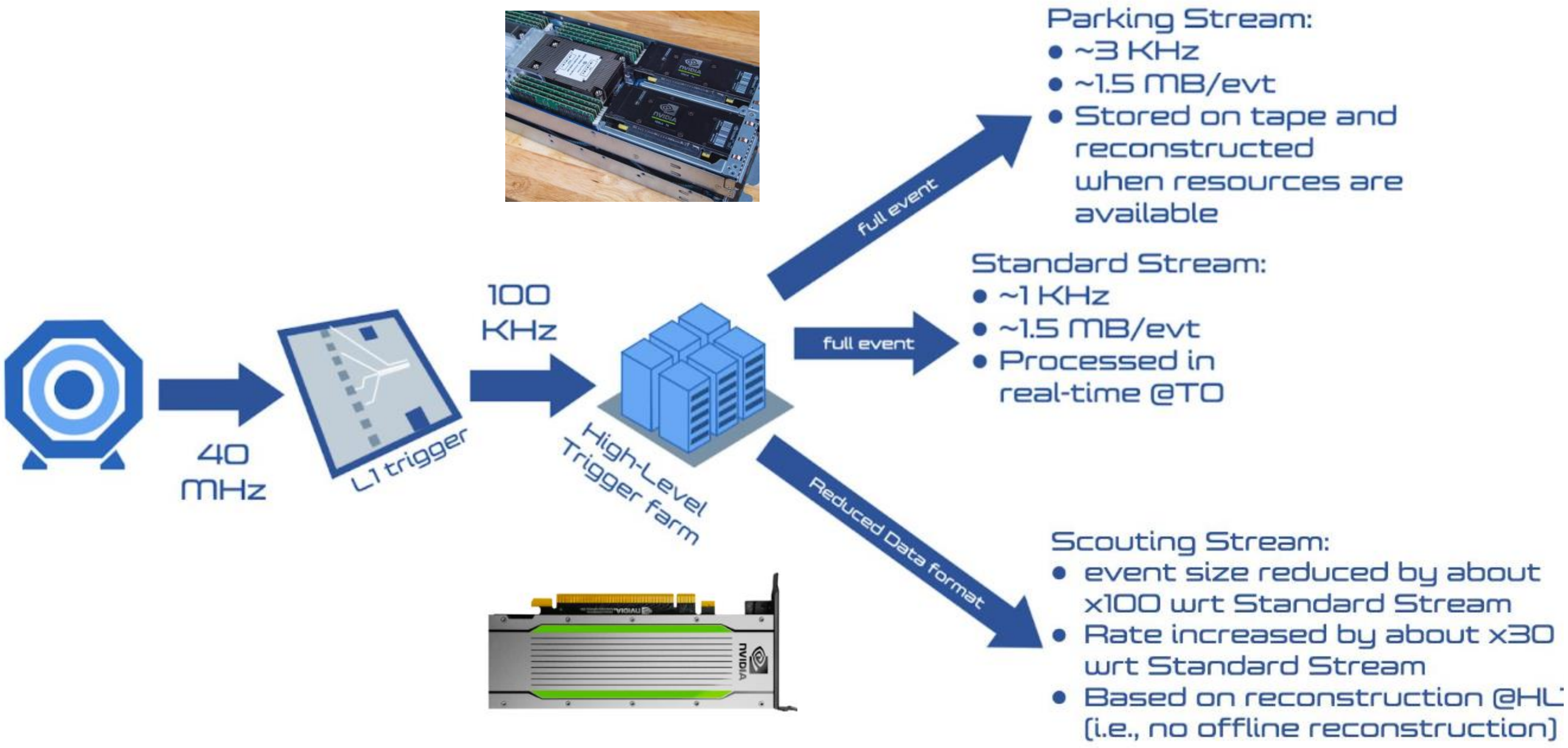
Bandwidth [GB/s] ~ Trigger output rate [kHz] x Average event size [MB]

~ 1 GB/s

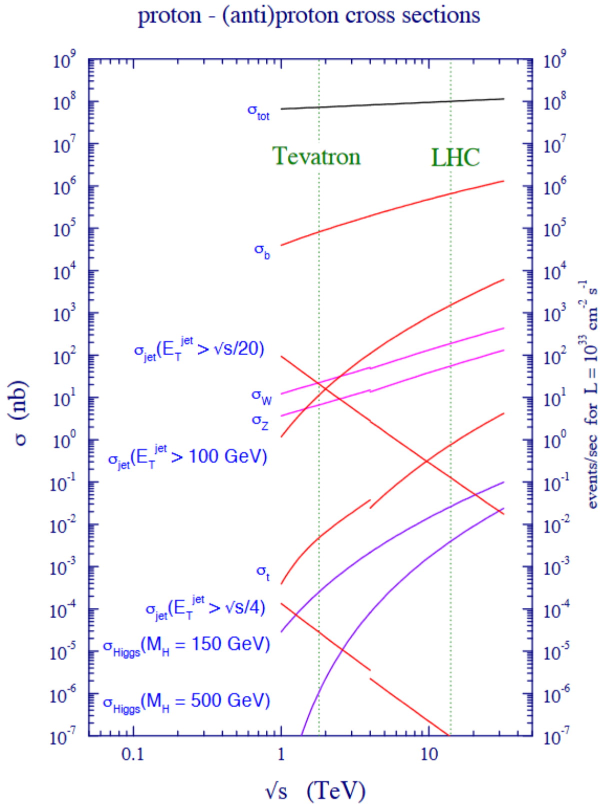
1 kHz (ATLAS & CMS)
12.5 kHz (LHCb)

Raw event data size
~1 MB (ATLAS and CMS)
~0.1 MB (LHCb)

CMS: Run 3 data taking strategy



Ref [LHCC Nov22]



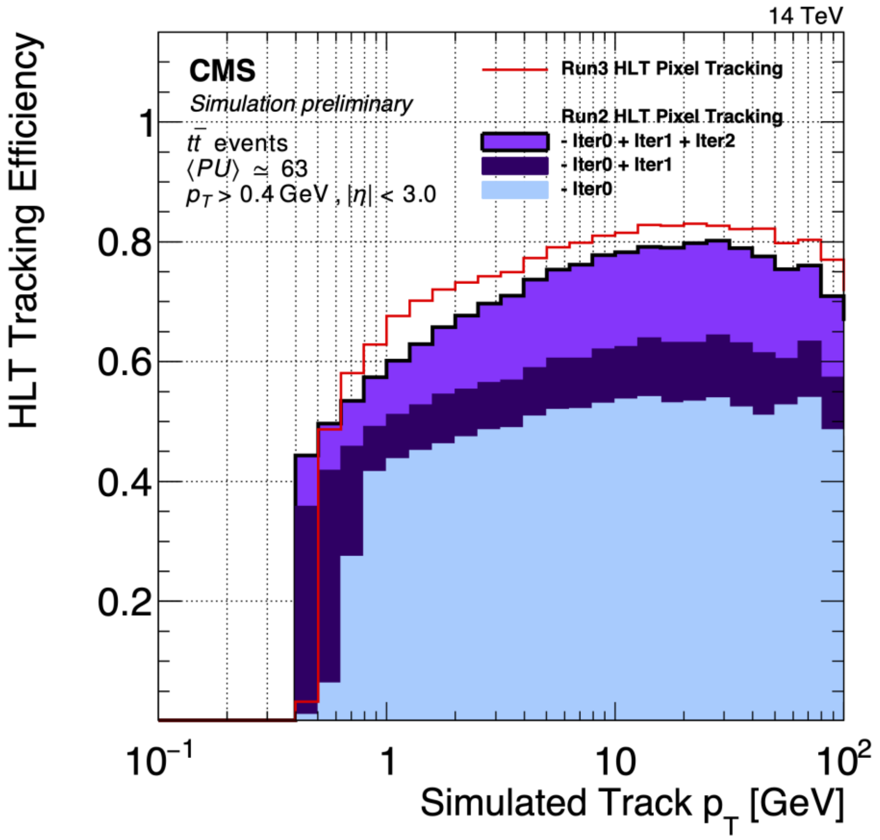
GPU offload 40% of event processing

- Yields 70% increase in throughput

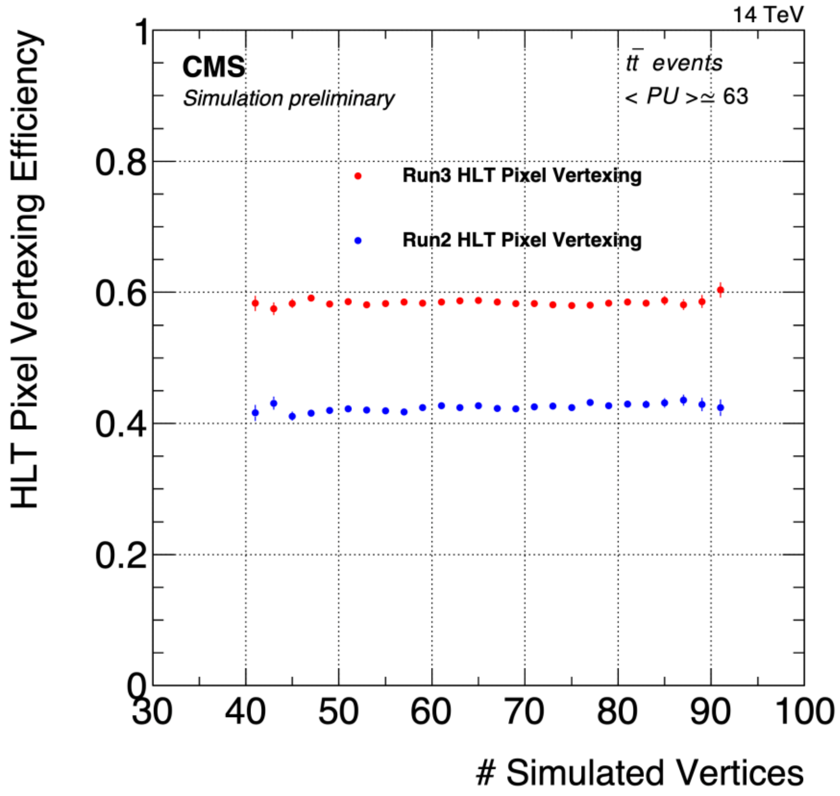
- The HLT farm consists of 200 machines with 2 sockets, equipped with AMD EPYC 7763 ”Milan” 64-core processors for a total of 128 physical cores and 256 hardware threads, as well as 2 low profile Nvidia T4 GPUs

CMS: Run-3 HLT performance

REF [CERN-CMS-DP-2022-014]



The pixel tracking efficiency as a function of the simulated track pt for the Run-2 HLT pixel tracks (stacked histogram) and the Run-3 HLT pixel tracks (red).



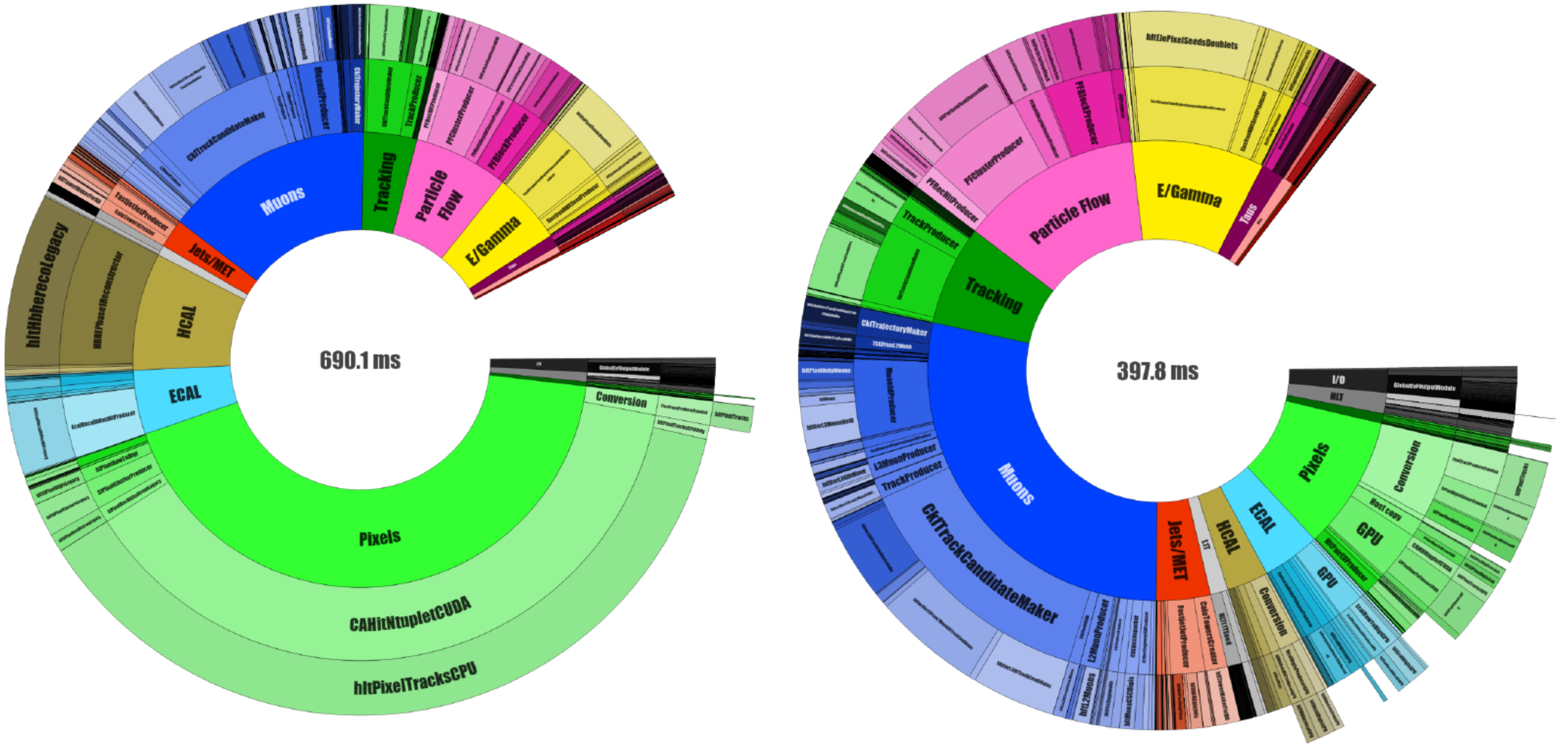
The pixel vertex efficiency as a function of the number of simulated vertices for the Run-2 HLT (blue) and the Run-3 HLT (red) vertex reconstruction algorithms.

CMS: Run-3 HLT performance

Average time per event on CPU vs GPU [REF- CMS_ACAT_2022]

Parts of the reconstruction are offloaded to GPUs: Pixel Tracking, ECAL, HCAL

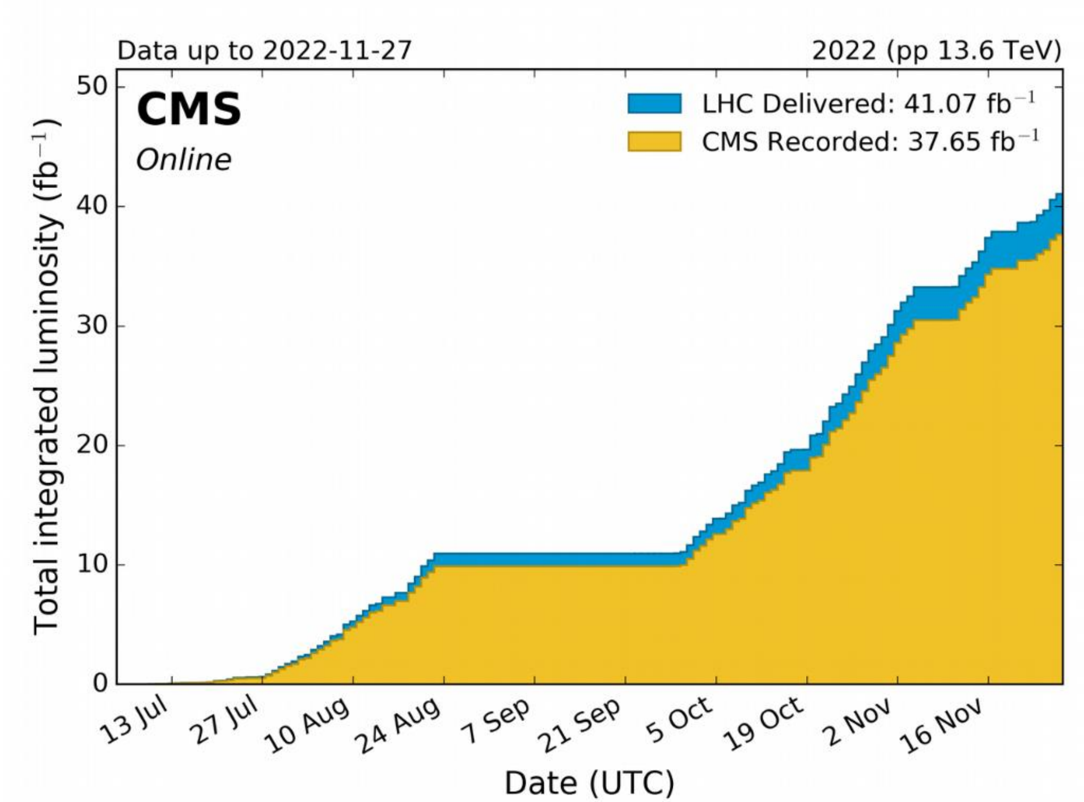
- Currently, the GPU code is running at 90kHz (pixel reconstruction is run on 88% of events, ECAL on 70% and HCAL on 65%
- In addition to releasing the CPU from time consuming tasks, it offers the ability to use global pixel tracks at the HLT for the first time

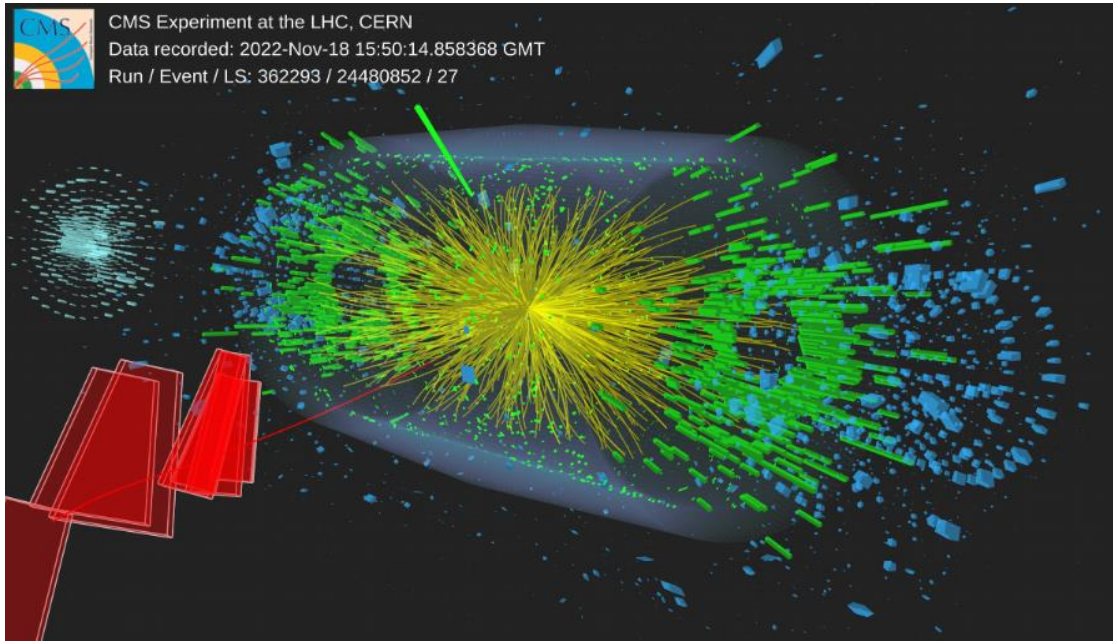


- The execution time per event of the HLT step was reduced on average by a factor 1.7
- Graph Neural Network for jet tagging

CMS: Run 3 performance

[REF: LHCC_CMS_Nov_2022]





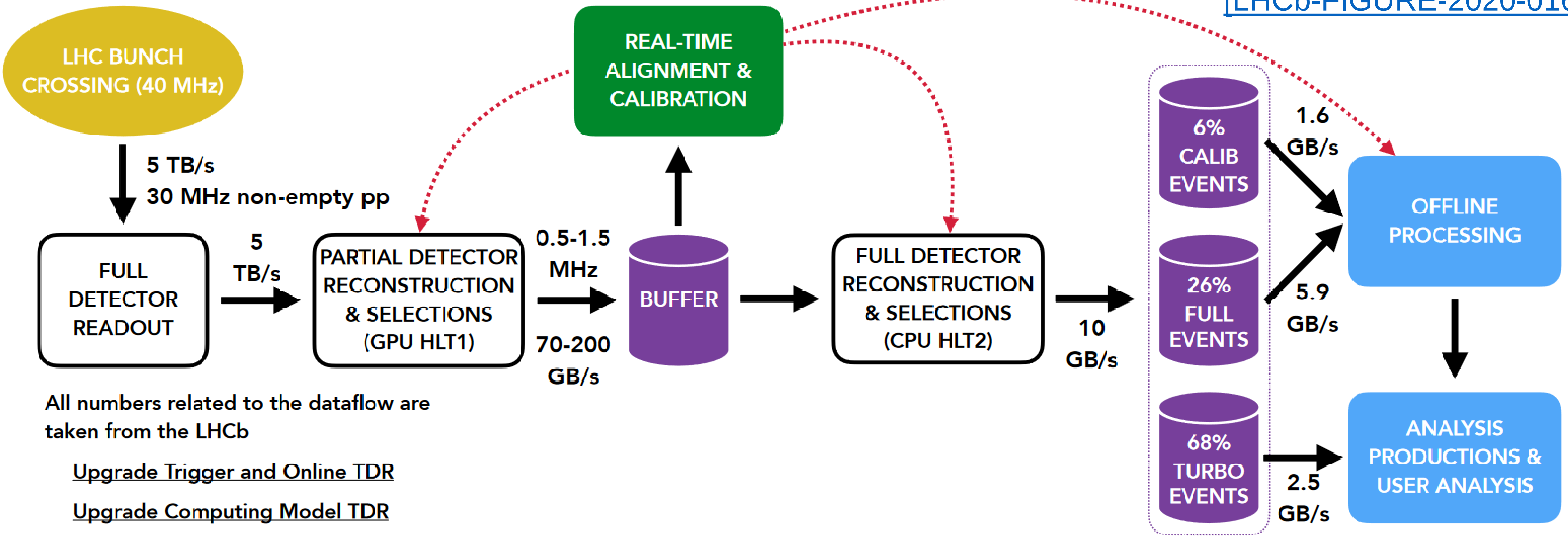
The trigger menu comparison with and without GPU reconstruction

- From the ~ 700 trigger paths in the menu, ~ 400 show no difference at all
- A dedicated trigger path was added to monitor differences in yield between CPU and GPU after the full ParticleFlow reconstruction, and store these events: *HLT PFJet40 GPUvsCPU v1*

First measurement of the top quark pair production cross section in proton-proton collisions at 13.6 TeV

LHCb: Run-3 Real Time Analysis Data flow

[LHCb-FIGURE-2020-016]



All numbers related to the dataflow are taken from the LHCb

Upgrade Trigger and Online TDR

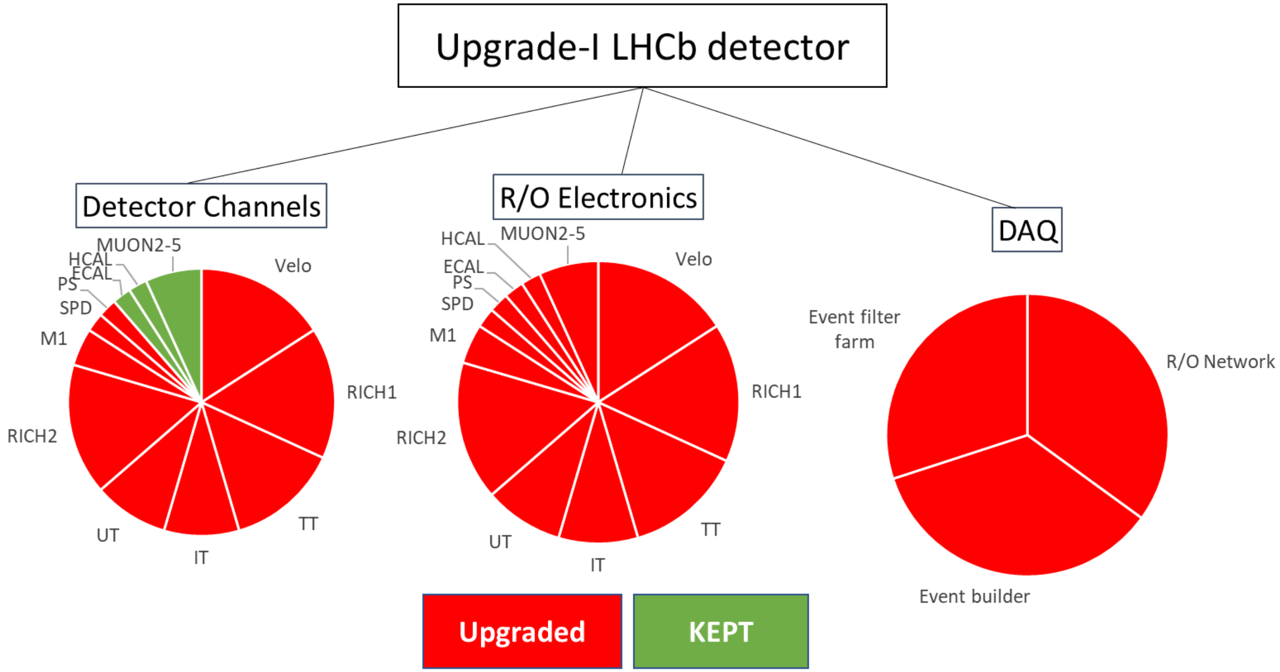
Upgrade Computing Model TDR

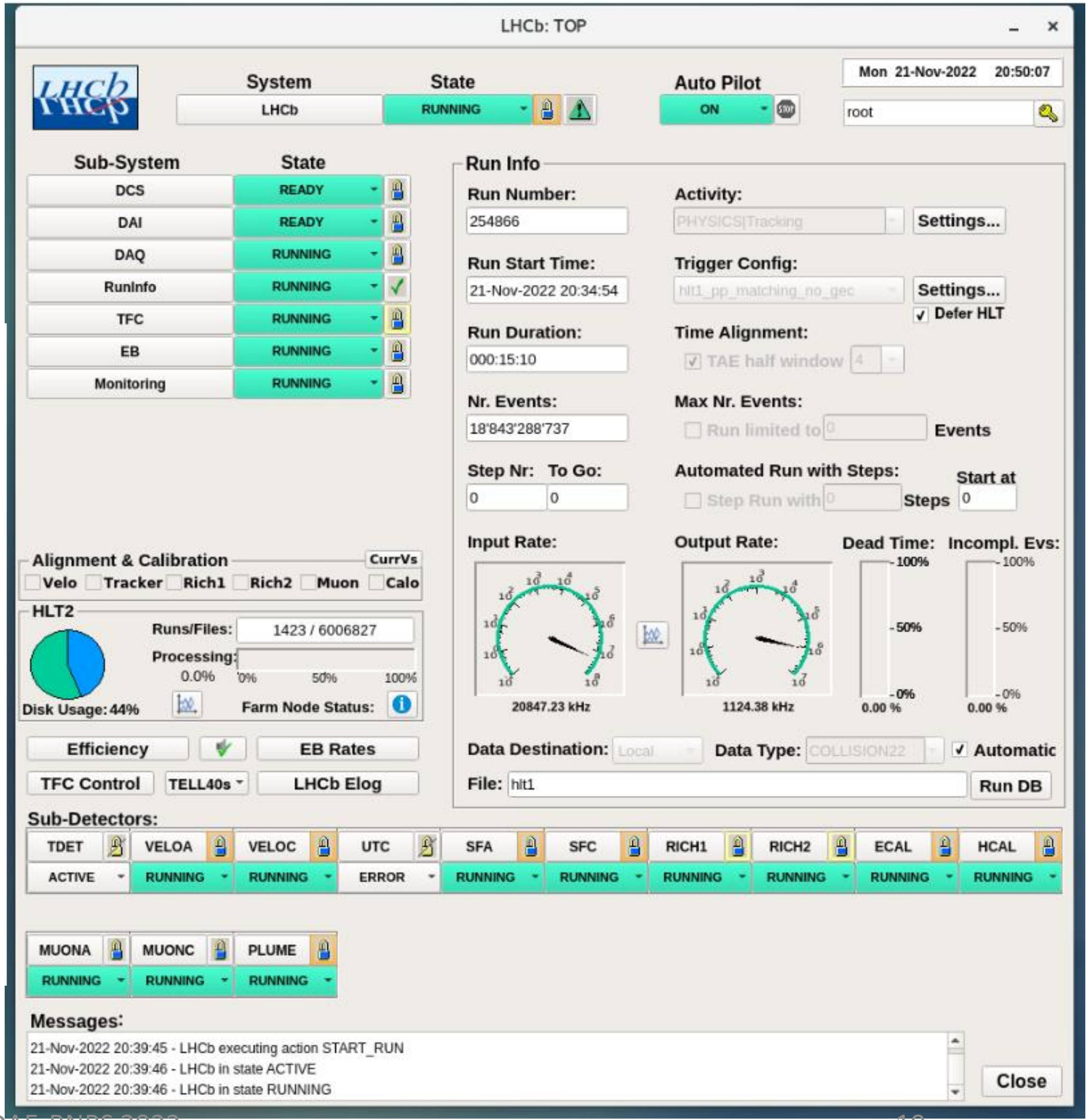
LHCb: Commissioning Upgrade-I

Upgraded LHCb detector for Run-3

- **5x** higher instantaneous luminosity $2x10^{33} cm^{-2} s^{-1}$
- **10x** per unit time signal yield
- **6x** more pileup

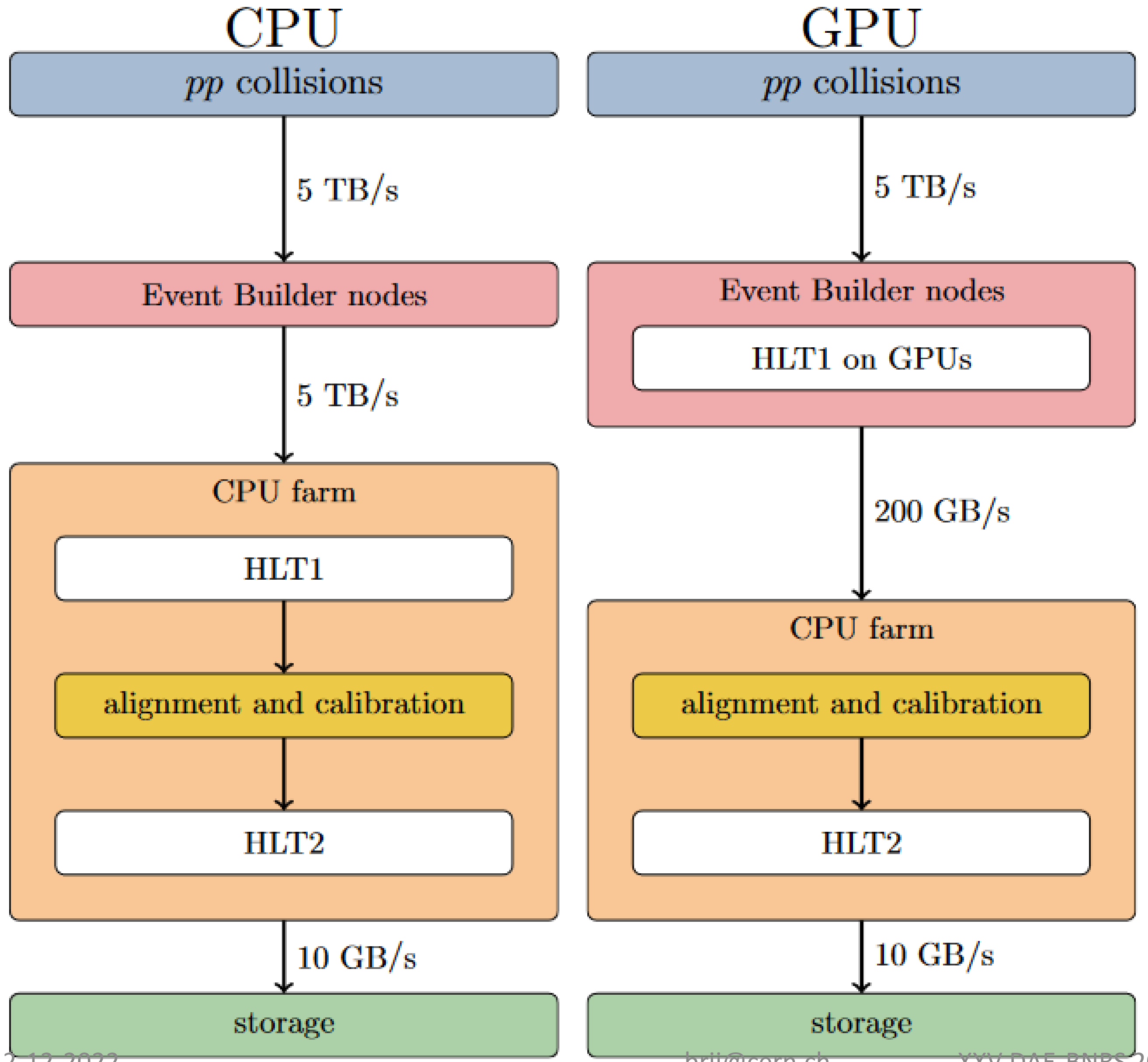
[LHCB-TDR-018]





LHCb: Technology decision -CPU vs GPU

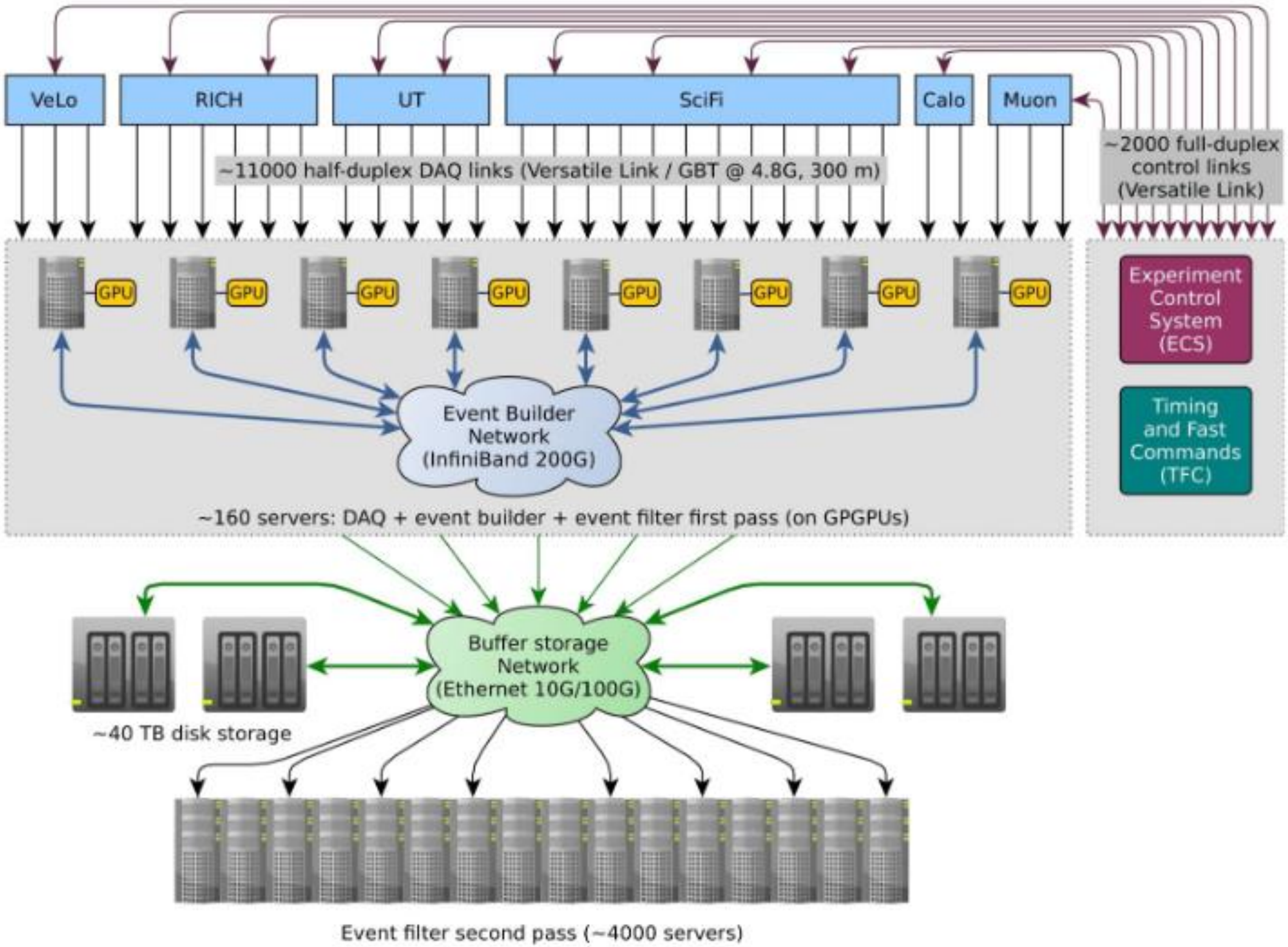
[Comput. Softw. Big Sci. 4, no.1, 7(2020)]
[CSBS6 (2022) 1, 1]



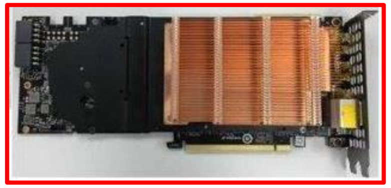
Comprehensive cost benefit analysis detailed in **Ref [CSBS6 (2022) 1, 1]**

The final verdict:
GPUs in the software trigger will allow LHCb to expand its physics reach in Run 3 and beyond

LHCb: DAQ Architecutre



1-3 GPUs

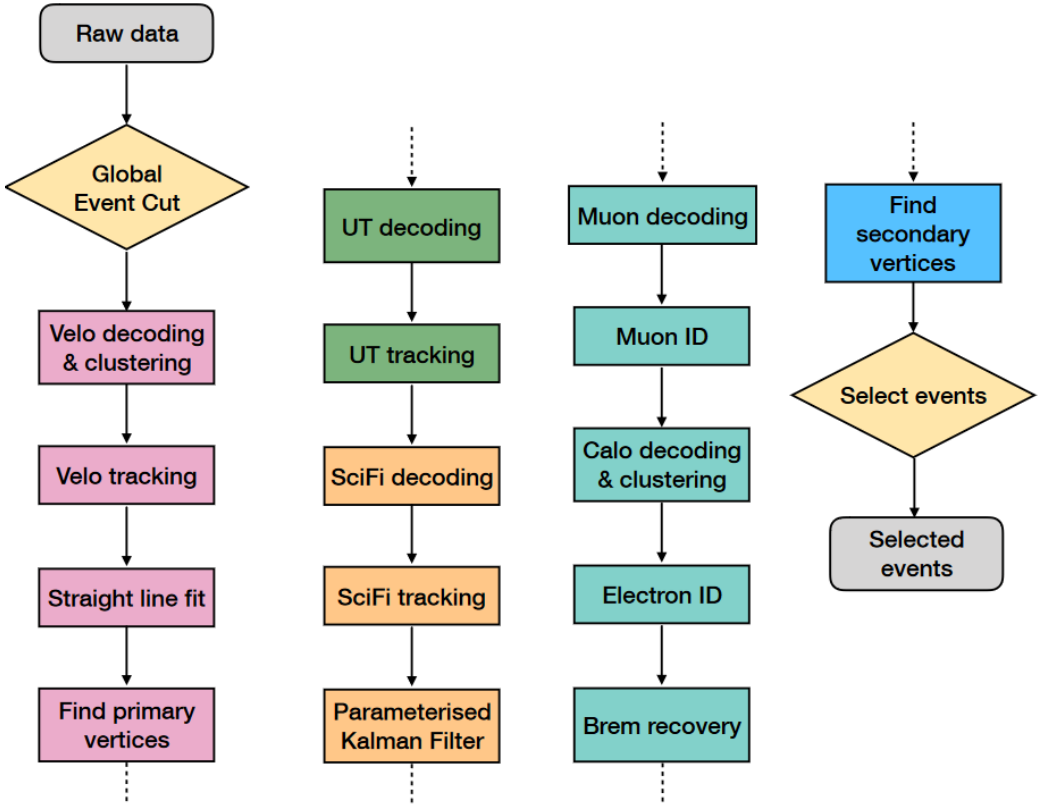


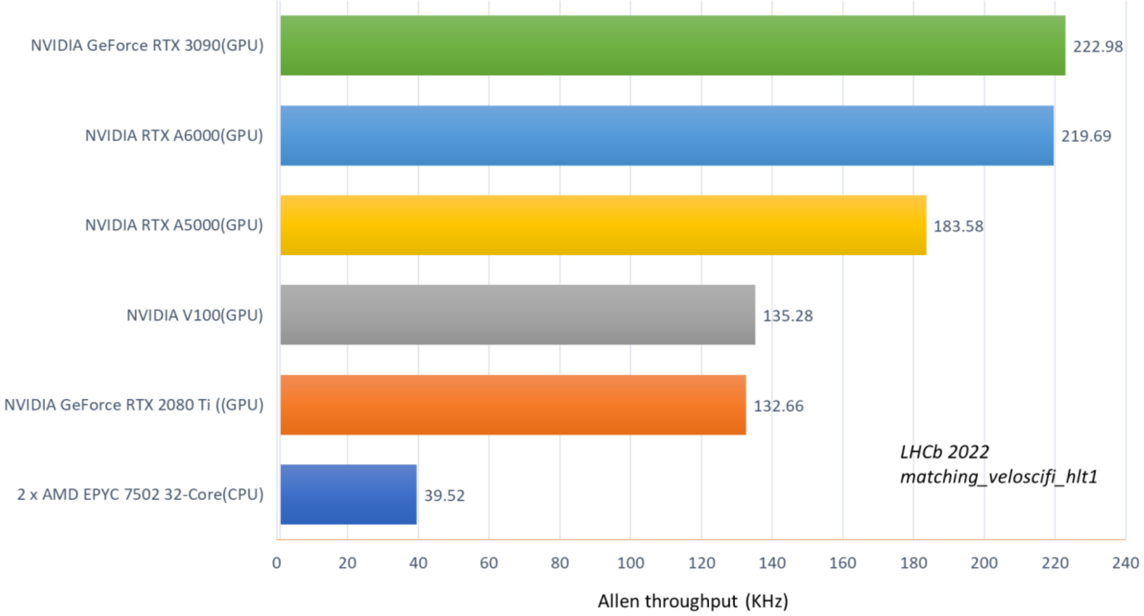
2 network connections

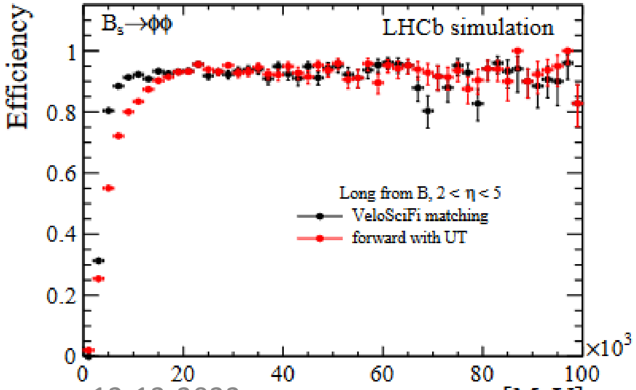
3 PCIe40 (FPGAs)

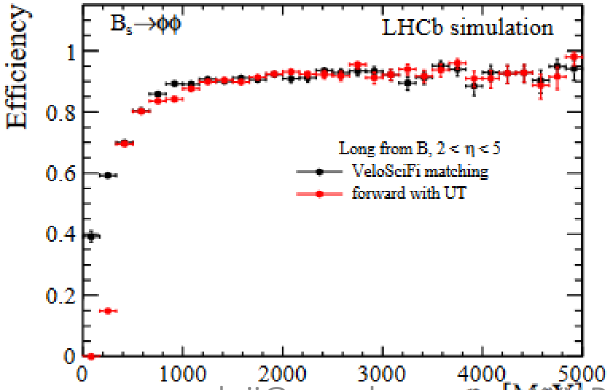
LHCb: HLT1

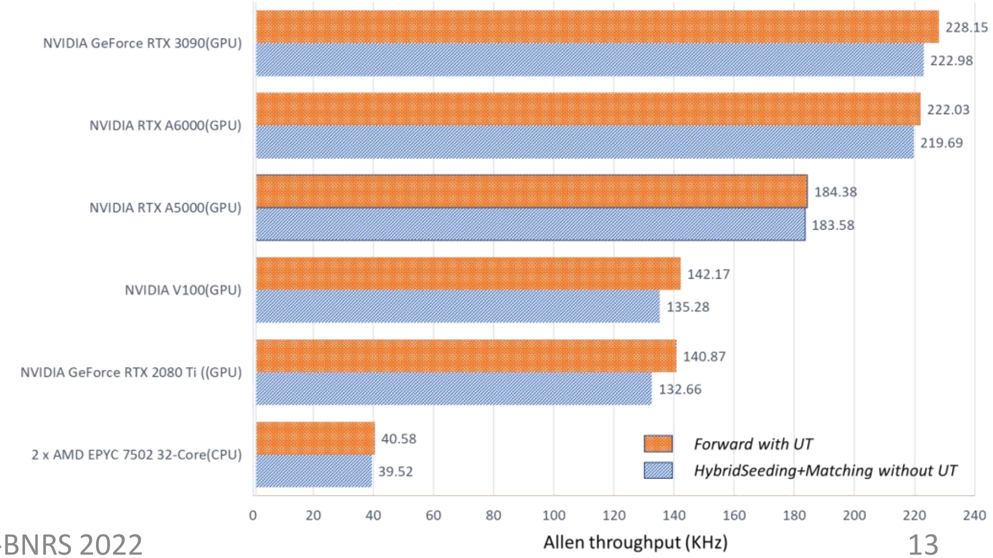
[CERN-LHCb-PROC-2022-013]







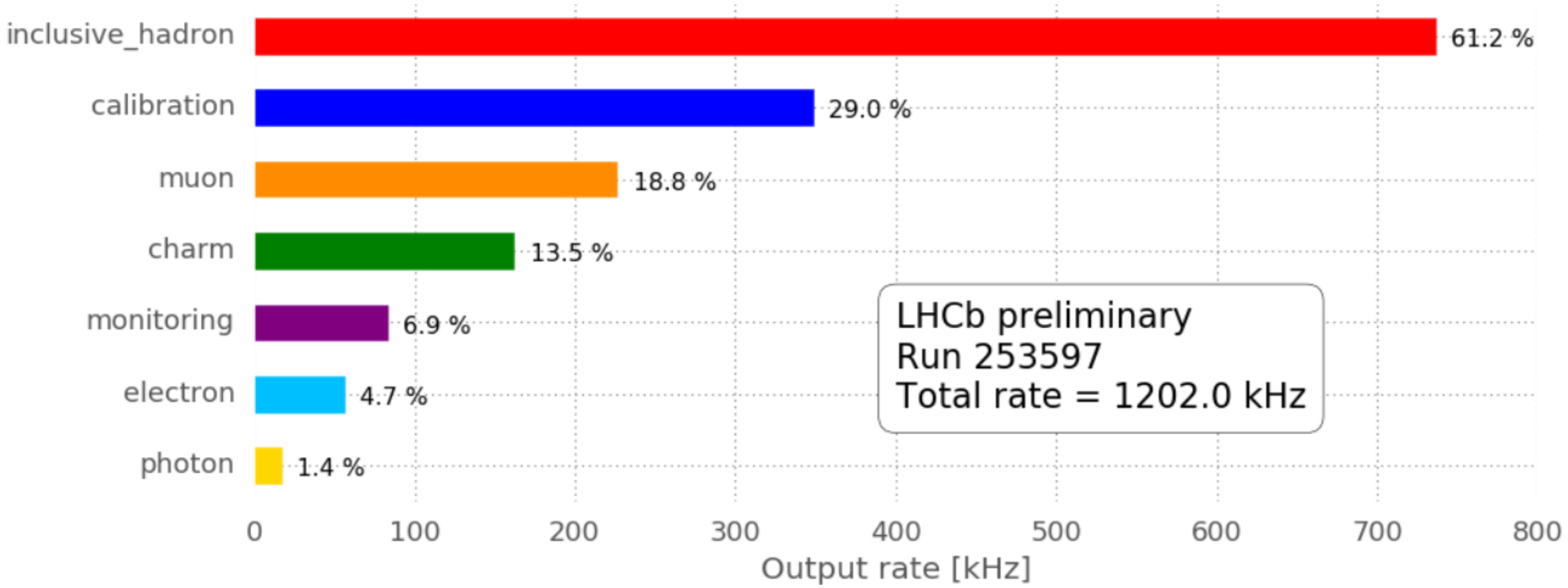


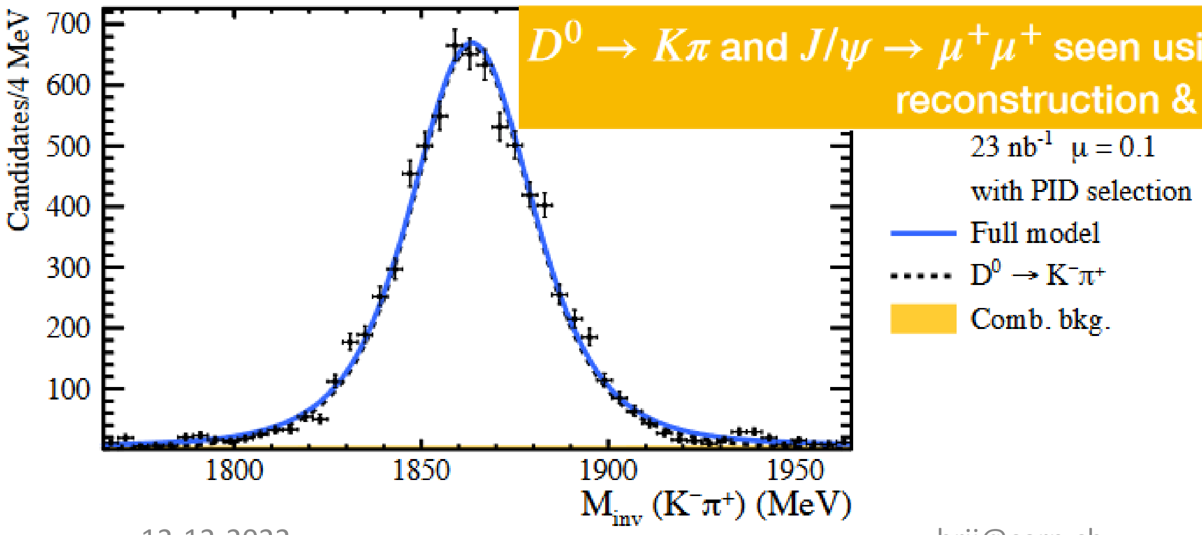


LHCb: Commissioning Upgrade-I Run-3

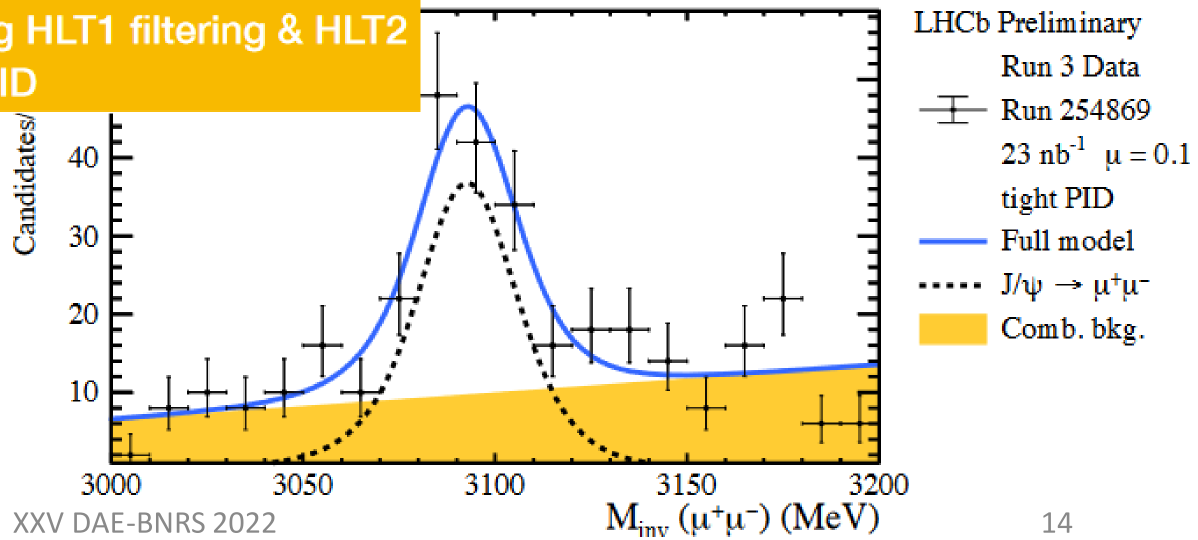
HLT1 able to handle max LHC input rate of this year and select at design output rate

[LHCC_open_LHCb_Nov_2022]





$D^0 \to K\pi$ and $J/\psi \to \mu^+\mu^+$ seen using HLT1 filtering & HLT2 reconstruction & PID

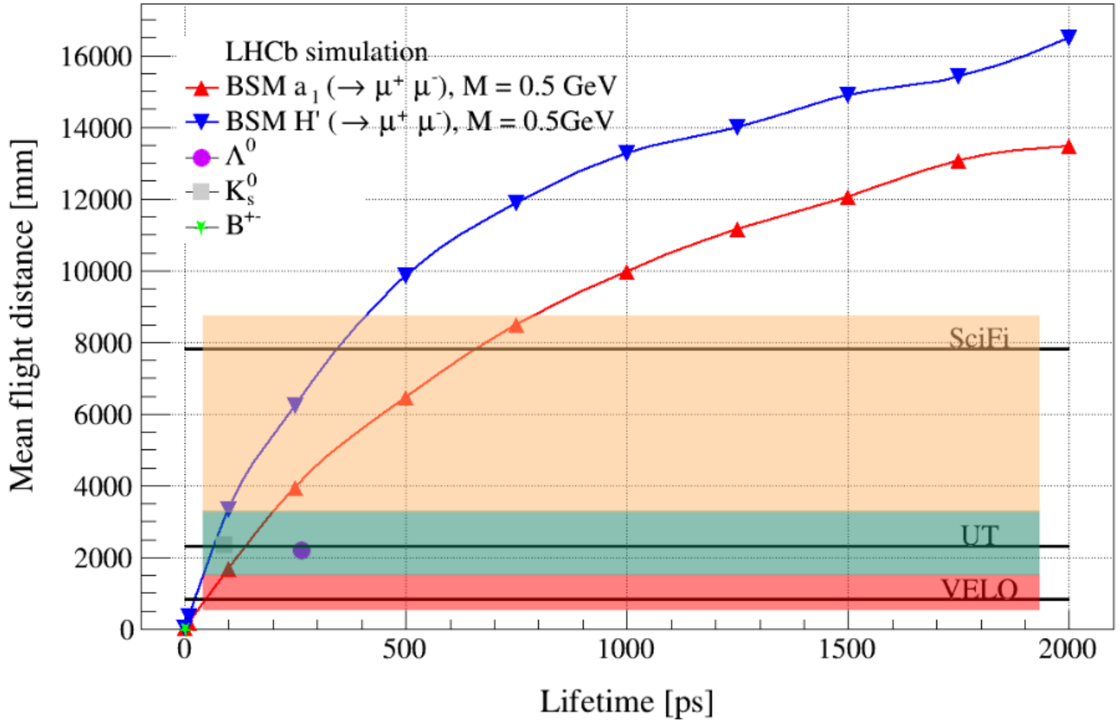


Conclusion and prospects:- Triggers role in extending physics programmes

[Front. Big Data, 07 November 2022]

LLPs present in SM and many BSM extensions
Expected track types depend on LLP flight distance

- Hybrid computing technologies and architectures are helping in extending the physics programs of LHC experiments while keeping the costs in check.
- More powerful then ever Tracking and reconstruction algorithms at High level triggers.
- Moving from Triggering to Real Time analysis - Exciting times ahead. !!!

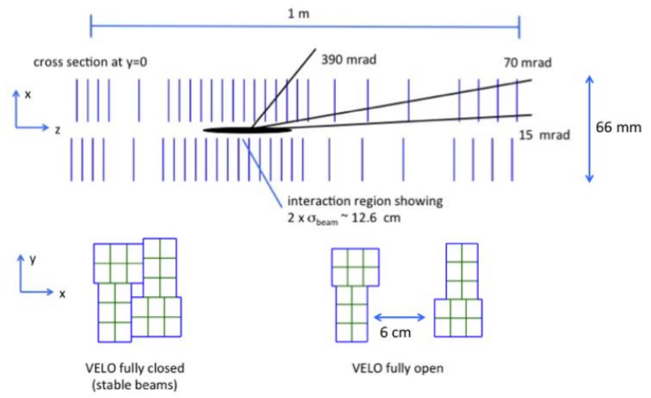


Thank you

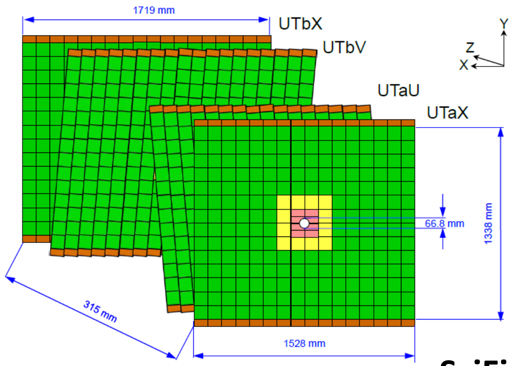
LHCb: Tracking system

[CERN-LHCb-PROC-2022-013]

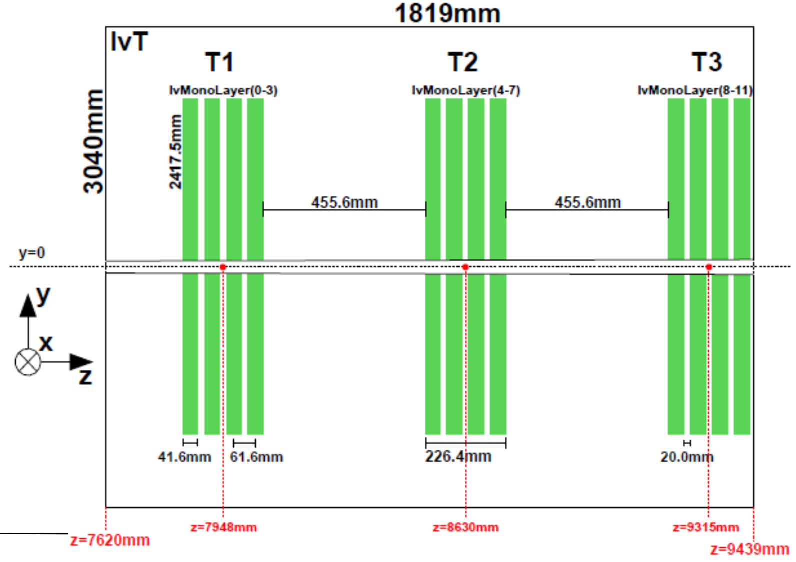
VELO: Silicon pixel detector, 52 planes

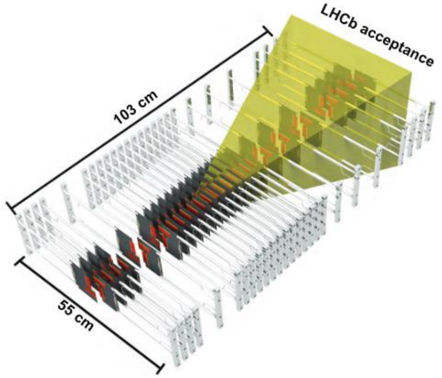


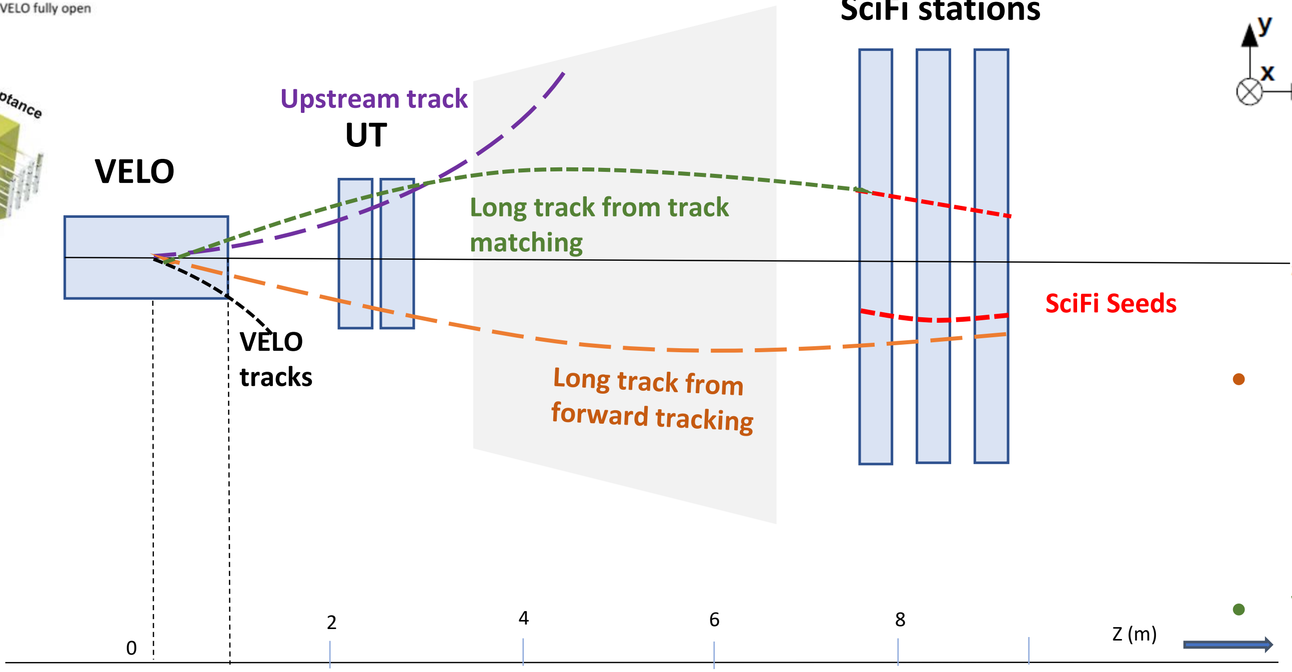
UT: Silicon strip detector, 4 planes



SciFi: Scintillating Fibre detector
(12 planes of 2x2.5 m long))







- Forward tracking: Velo tracks are extended to UT and then to SciFi to create long tracks (Baseline HLT1)

- VeloSciFi Track matching: SciFi seeds are matched with Velo tracks to create long tracks. (UT hits can also be added)