Real-time analysis with LHCb

The 2019 International Workshop on the High Energy Circular Electron Positron Collider
IHEP, Beijing, China
2019/11/19

SAUR Miroslav
On behalf of the LHCb collaboration

(University of Chinese Academy of Sciences)

LHCb detector 2010-2018

→ Single-arm forward spectrometer focused on heavy flavor (*b*, *c*) physics

→ Run I (7/8 TeV, 3 fb^{-1}), Run II (13 TeV, ~6 fb^{-1}) + special runs (pPb, PbPb, SMOG)

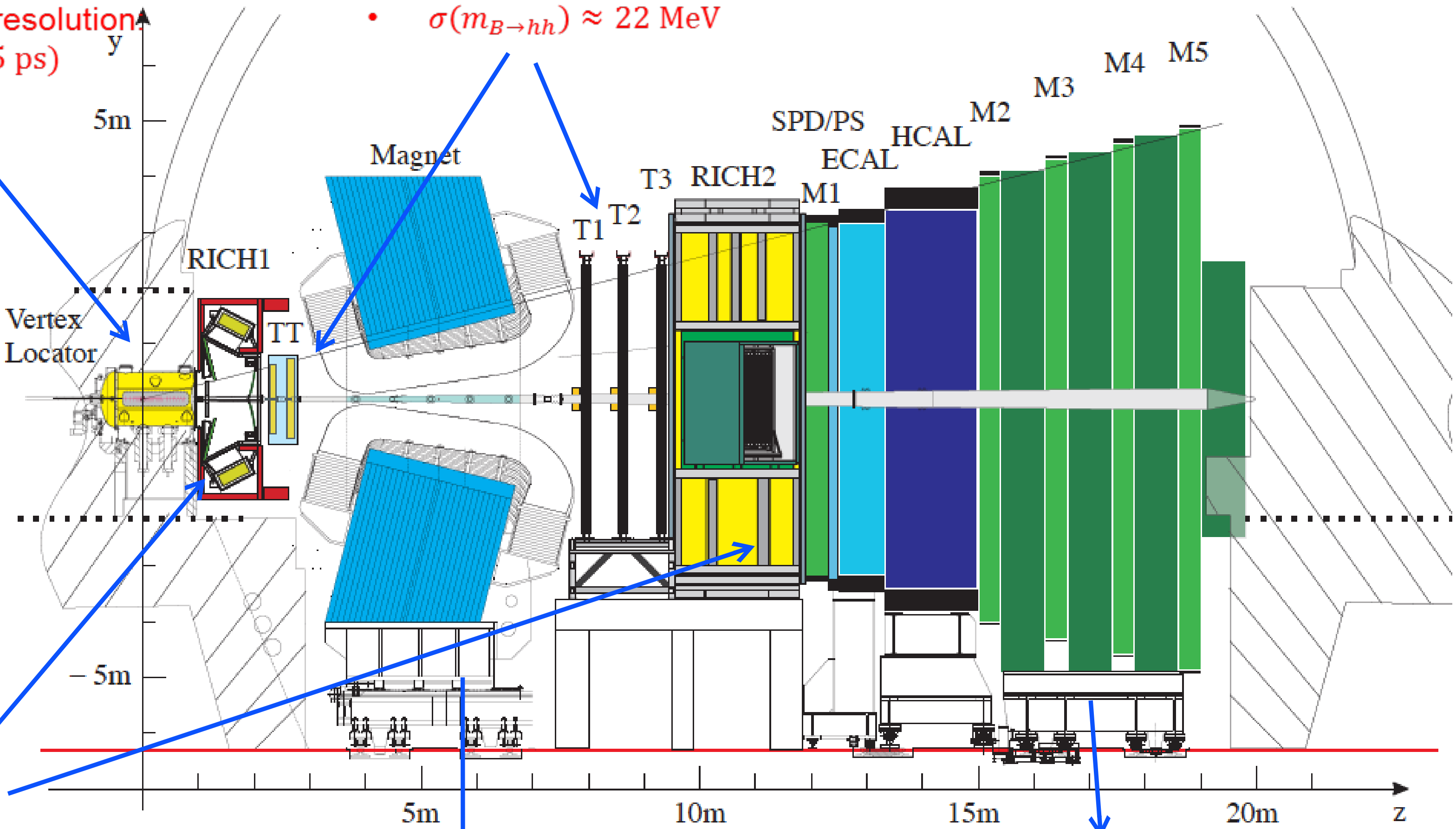
Vertex Locator(vertex reconstruction)

- Impact parameter resolution: 20 μm
- Decay time resolution: 45 fs (τ_B~1.5 ps)

Tracking system(particle reconstruction)

- ϵ(Tracking) ~96%
- $\delta p/p$ ~0.5%-1%(5-200 GeV)
- $\sigma(m_{B\to hh}) \approx 22$ MeV

JINST 3 (2008) S08005
IJMPA 30 (2015) 1530022



RICH: particle ID

- $\varepsilon(K \to K)$ ~95%
- Mis-ID: $\varepsilon(\pi \to K)$ ~5%

Magnet

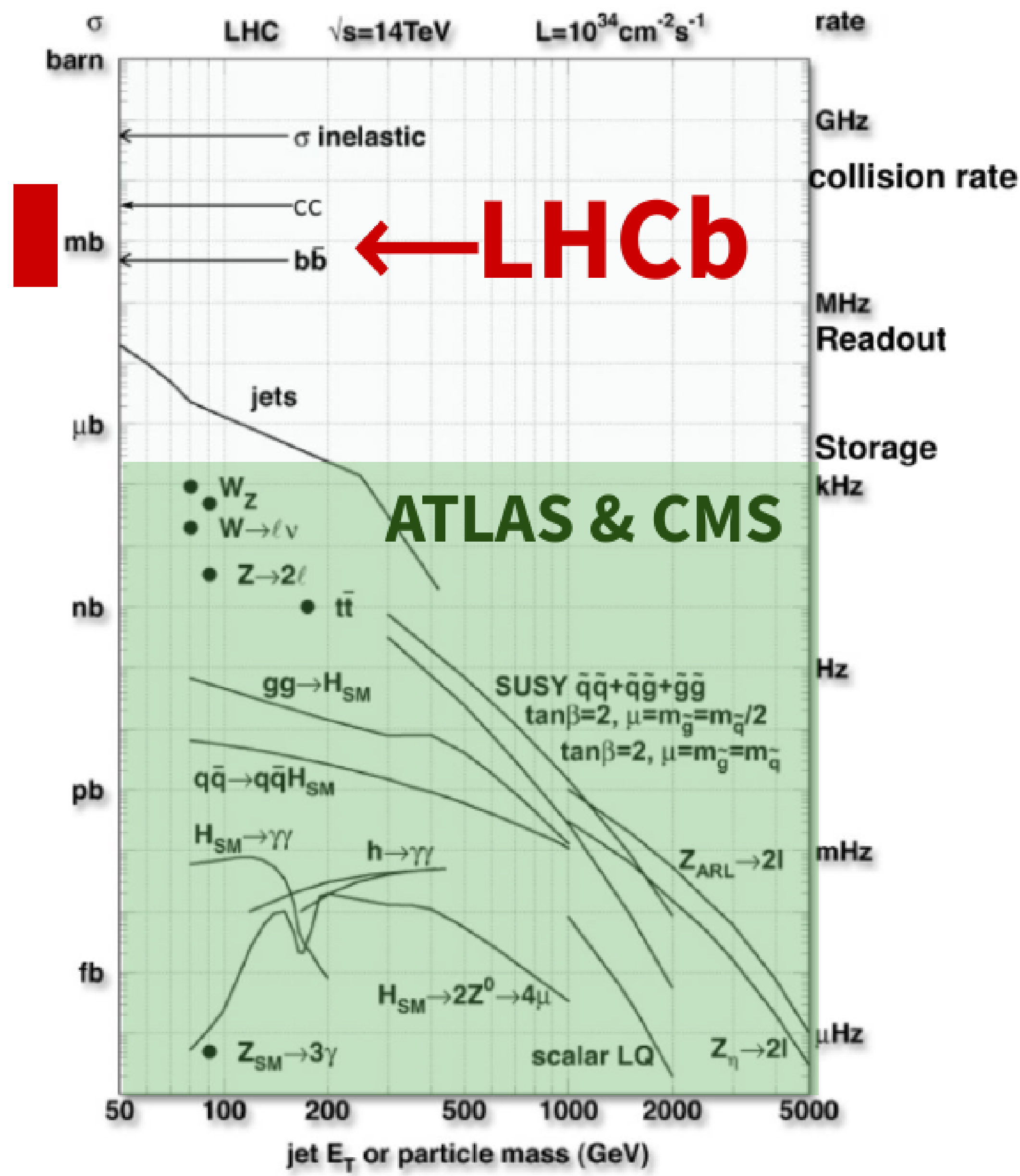
- Bending power: 4 Tm

Muon system

- μ ID: $\varepsilon(\mu \to \mu)$ ~97%
- Mis-ID: $\varepsilon(\pi \to \mu)$ ~1-3%

Requirements for trigger

- Triggering is a crucial part of data taking
- Decision of what physics can be recorded
- Resources demanding operation
- Hard constrains: Bandwith [GB/S] ≈ Accept Rate [kHz] × Event size [kB]
- Limiting factors: both hardware and software
- Raw data bandwidth scales up quadratically with luminosity
- During the Run II already significant rates: 45 kHz for bb, 1 MHz for cc

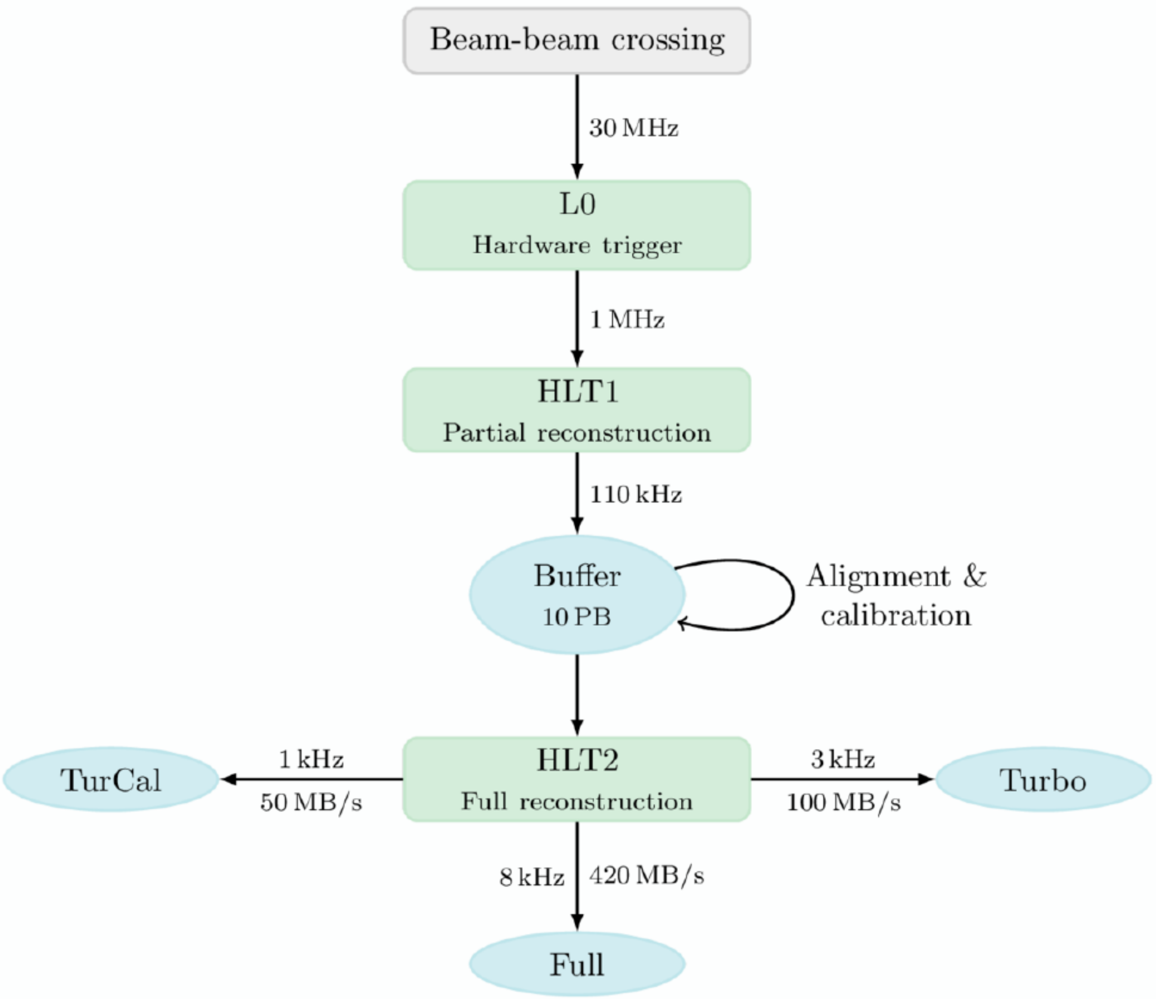


Trigger during Run II

- Run II (2015-2018) trigger system consisted of 3 stages

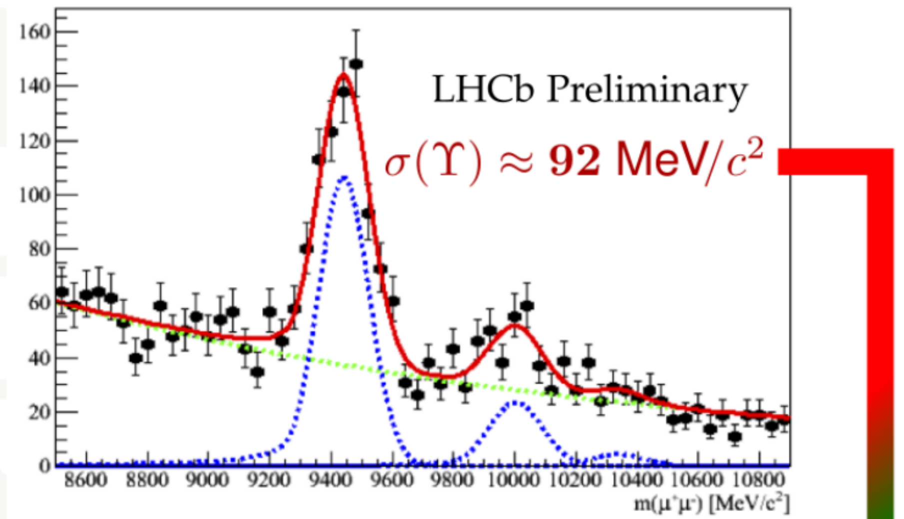
1) L0 Hardware trigger
   - fast detectors
   - CALO and MUON information
2) High Level Trigger 1
   - 10 PB disk buffer
   - Around two weeks of data taking
3) High Level Trigger 2

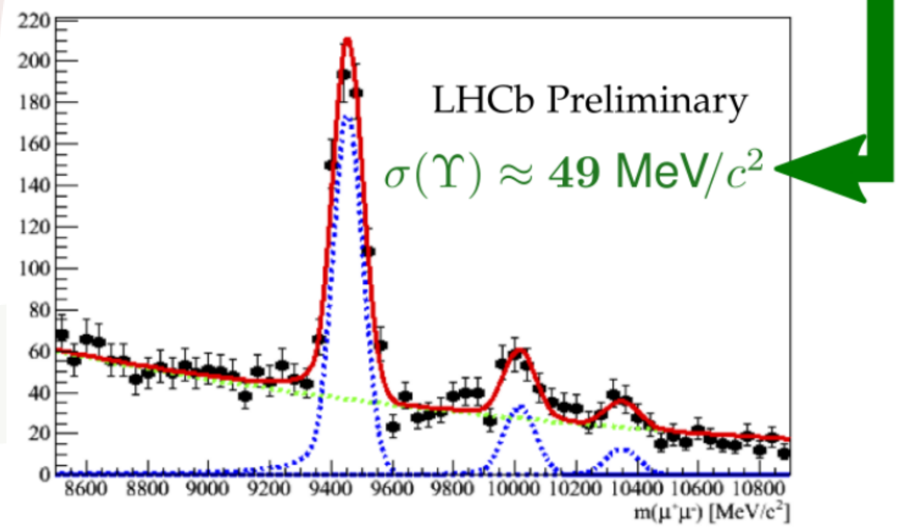
- Around 500 specific HLT lines (particular decays)
- Introduction of TESLA framework → Turbo stream

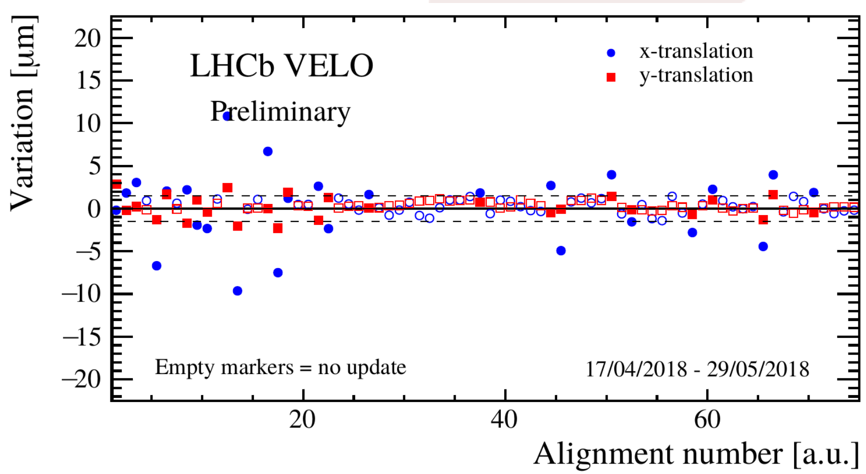


Trigger – alignment & calib.

- HLT1 samples are used for alignment and calibration
- Alignment procedure of the full tracker system run automatically at the beginning of each fill
- Based on Kalman filter
- Update if the variations are significant
- RICH calibration and alignment
- Time calibration of OT
- Calibration of ECAL

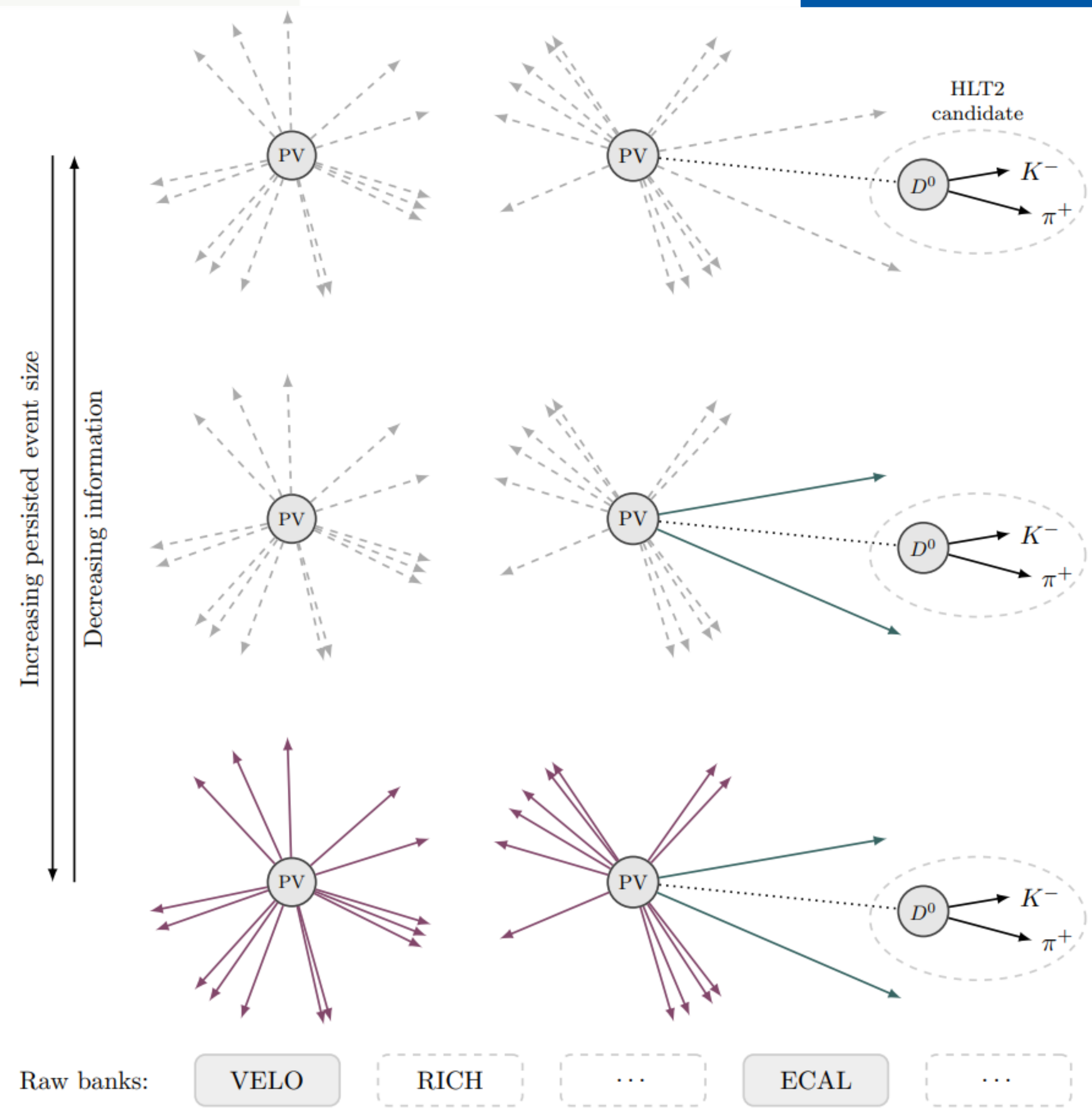






Turbo stream

- Given the bandwidth hard limits, do we need to save all information about all events?
- Select what we want to save
- Turbo (2015)
  - Keep only objects used for trigger
- Turbo SP (2017)
  - Objects used for trigger + special selection
- Turbo++ (2016)
  - All reconstructed events
  - Raw event is dropped



Turbo stream

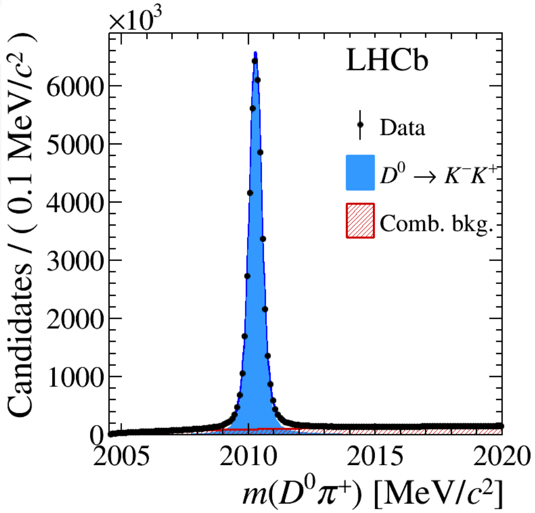
- Extensively used during the Run II
  - Around 30 % of the trigger rate is Turbo – almost all Charm physics
  - But only about 10 % of the bandwidth!
  - Approximately 2/3 lines keep raw detector information (Turbo SP)
- Significant reduction of data size → more events at same bandwidth

Persistence method	Average event size [kB]
Turbo	7
Turbo SP	16
Turbo++	48
Raw event	69

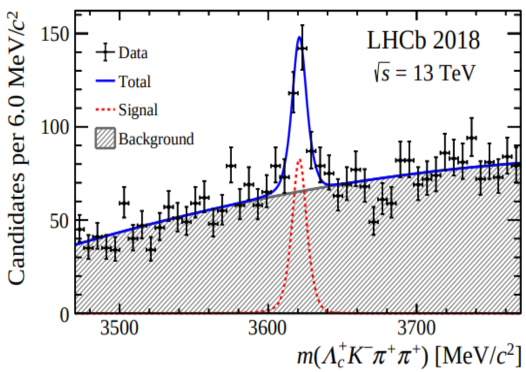
- Turbo stream relies on full detector alignment and calibration within the trigger phase

Accomplishments of Turbo

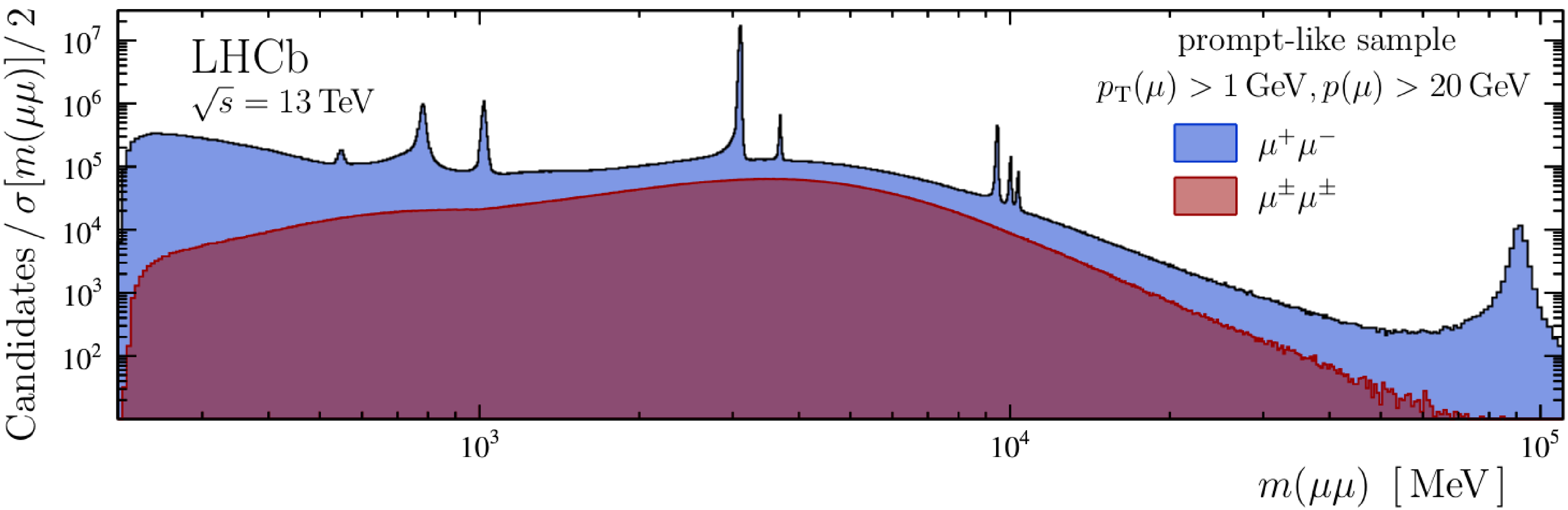
- Turbo lines proved to be necessary for keeping a significant charm program during Run II
- Suitable for a broad range of physics – from high to low rate
- One of the key ideas of Upgrade



PRL 122 (2019) 211803



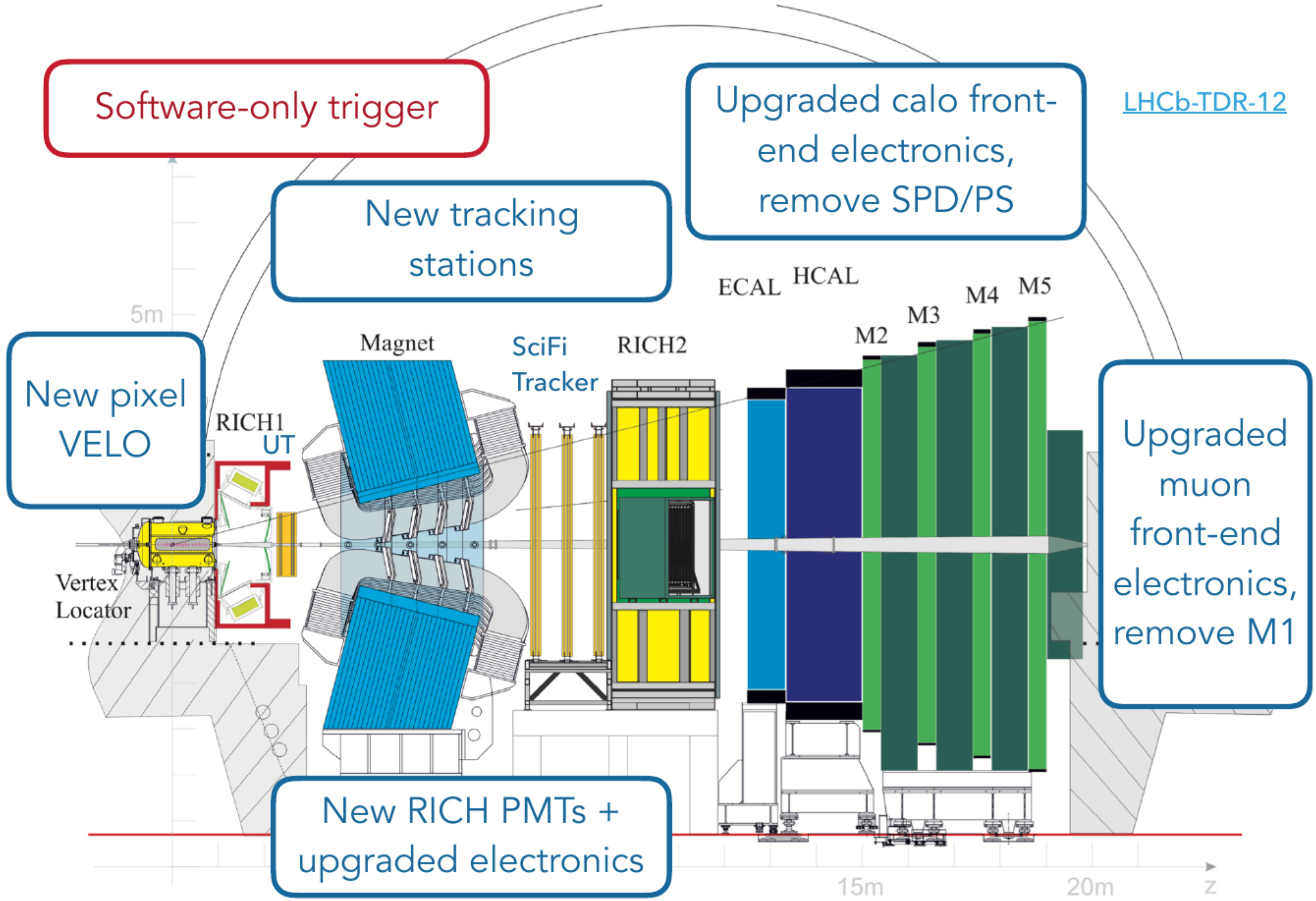
arXiv:1909.12273



PRL 120 (2018) 061801

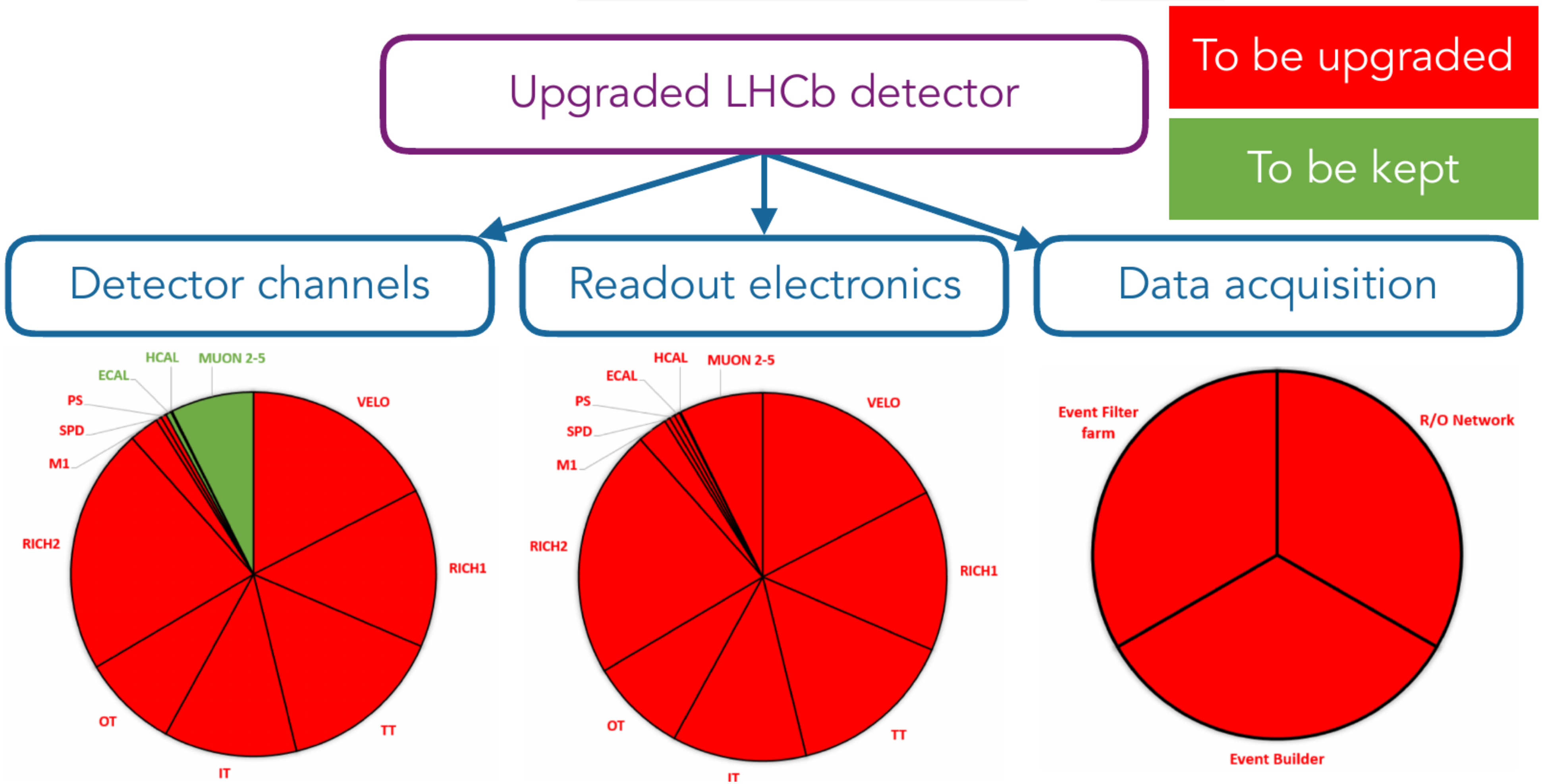
LHCb Upgrade I (Run III)

- Luminosity will increase 5x times and collision energy to 14 TeV
- Aim is to maintain the same performance as during Run II



LHCb Upgrade I (Run III)

- Luminosity will increase 5x times and collision energy to 14 TeV
- Aim is to maintain the same performance as during Run II



<u>A NEW DETECTOR AT LHC</u>

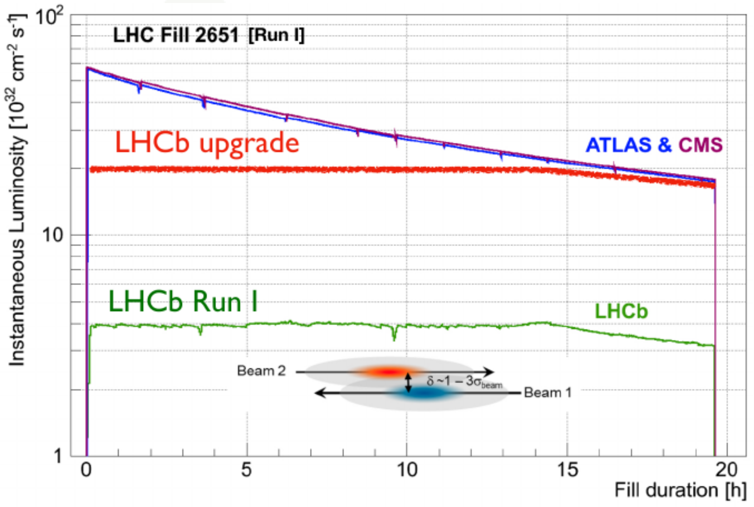
LHCb Upgrade I - Physics

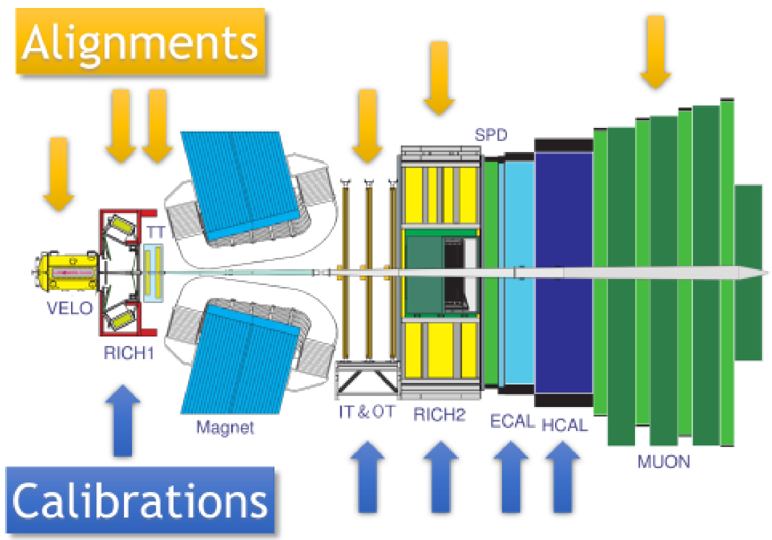
Type	Observable	Current precision	LHCb 2018	**Upgrade** ($50\,\text{fb}^{-1}$)	Theory uncertainty
B_s^0 mixing	$2\beta_s\ (B_s^0 \to J/\psi\ \phi)$	0.10 [9]	0.025	0.008	~ 0.003
	$2\beta_s\ (B_s^0 \to J/\psi\ f_0(980))$	0.17 [10]	0.045	0.014	~ 0.01
	$A_{\text{fs}}(B_s^0)$	6.4×10^{-3} [18]	0.6×10^{-3}	0.2×10^{-3}	0.03×10^{-3}
Gluonic penguin	$2\beta_s^{\text{eff}}(B_s^0 \to \phi\phi)$	–	0.17	0.03	0.02
	$2\beta_s^{\text{eff}}(B_s^0 \to K^{*0}\bar{K}^{*0})$	–	0.13	0.02	< 0.02
	$2\beta^{\text{eff}}(B^0 \to \phi K_S^0)$	0.17 [18]	0.30	0.05	0.02
Right-handed currents	$2\beta_s^{\text{eff}}(B_s^0 \to \phi\gamma)$	–	0.09	0.02	< 0.01
	$\tau^{\text{eff}}(B_s^0 \to \phi\gamma)/\tau_{B_s^0}$	–	5 %	1 %	0.2 %
Electroweak penguin	$S_3(B^0 \to K^{*0}\mu^+\mu^-; 1 < q^2 < 6\,\text{GeV}^2/c^4)$	0.08 [14]	0.025	0.008	0.02
	$s_0\,A_{\text{FB}}(B^0 \to K^{*0}\mu^+\mu^-)$	25 % [14]	6 %	2 %	7 %
	$A_{\text{I}}(K\mu^+\mu^-; 1 < q^2 < 6\,\text{GeV}^2/c^4)$	0.25 [15]	0.08	0.025	~ 0.02
	$\mathcal{B}(B^+ \to \pi^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$	25 % [16]	8 %	2.5 %	~ 10 %
Higgs penguin	$\mathcal{B}(B_s^0 \to \mu^+\mu^-)$	1.5×10^{-9} [2]	0.5×10^{-9}	0.15×10^{-9}	0.3×10^{-9}
	$\mathcal{B}(B^0 \to \mu^+\mu^-)/\mathcal{B}(B_s^0 \to \mu^+\mu^-)$	–	~ 100 %	~ 35 %	~ 5 %
Unitarity triangle angles	$\gamma\ (B \to D^{(*)}K^{(*)})$	~ 10–12° [19, 20]	4°	0.9°	negligible
	$\gamma\ (B_s^0 \to D_s K)$	–	11°	2.0°	negligible
	$\beta\ (B^0 \to J/\psi\ K_S^0)$	0.8° [18]	0.6°	0.2°	negligible
Charm CP violation	A_Γ	2.3×10^{-3} [18]	0.40×10^{-3}	0.07×10^{-3}	–
	ΔA_{CP}	2.1×10^{-3} [5]	0.65×10^{-3}	0.12×10^{-3}	–

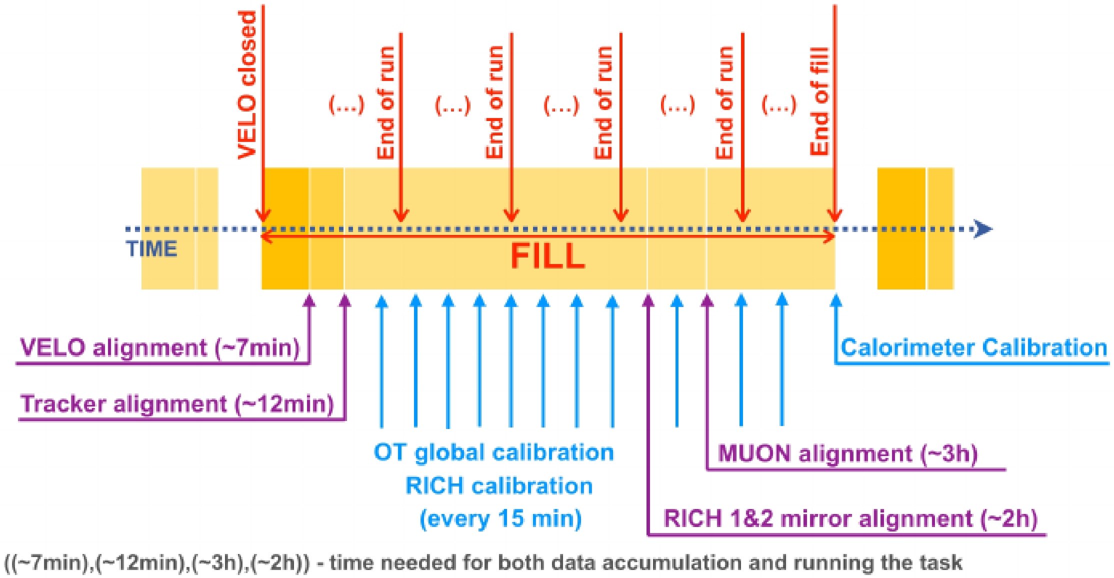
CERN/LHCC 2012-007

LHCb Upgrade I (Run III)

- LHCb has a very broad physics program
- High quality data requires a perfectly calibrated and aligned detector
- Have to process 5x bigger events at 30 times the rate, L0 removed
- From Run 3 all alignments and calibrations will be fully automatic and incorporated to the software trigger
- Around 70 % of data will go to Turbo

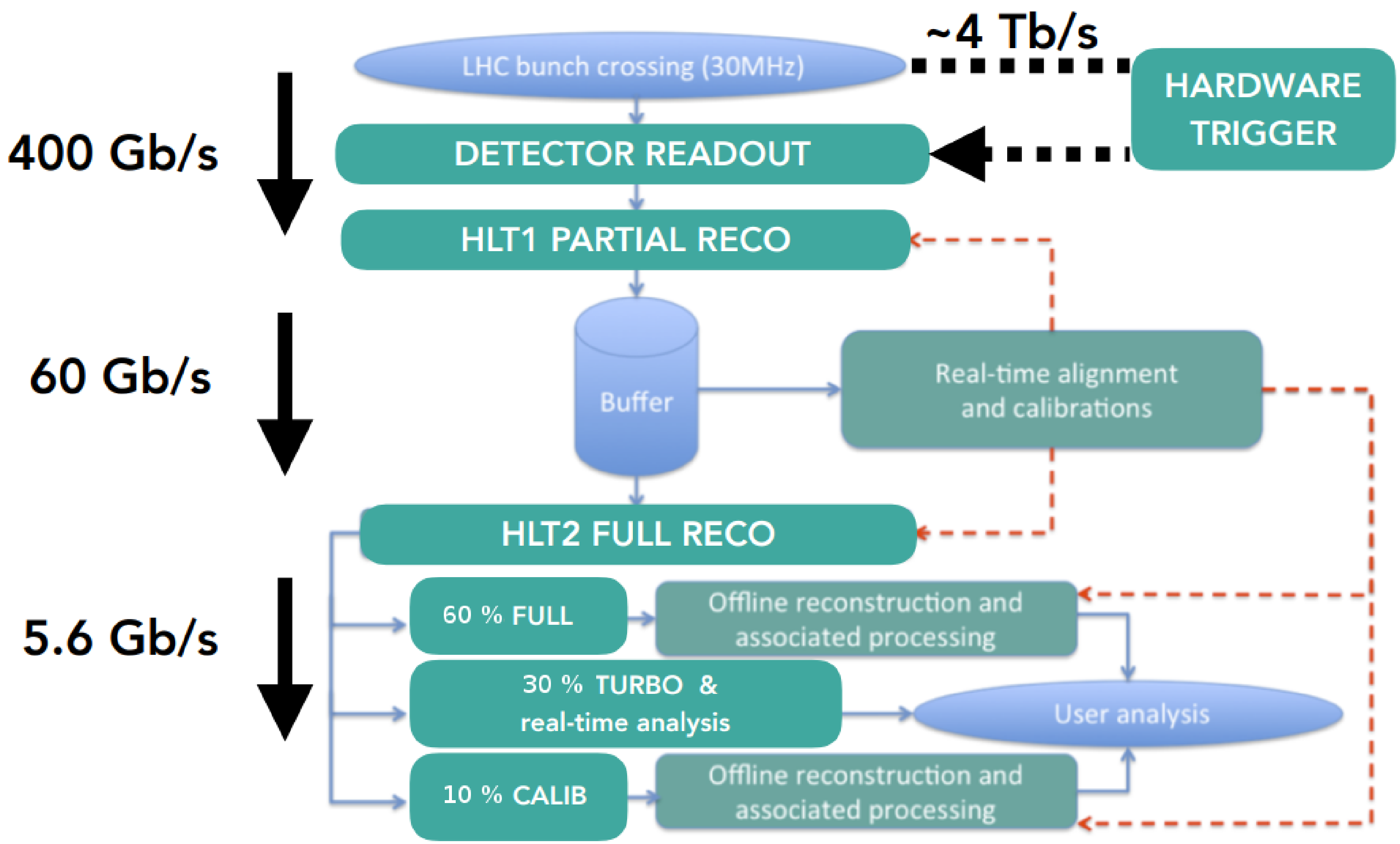






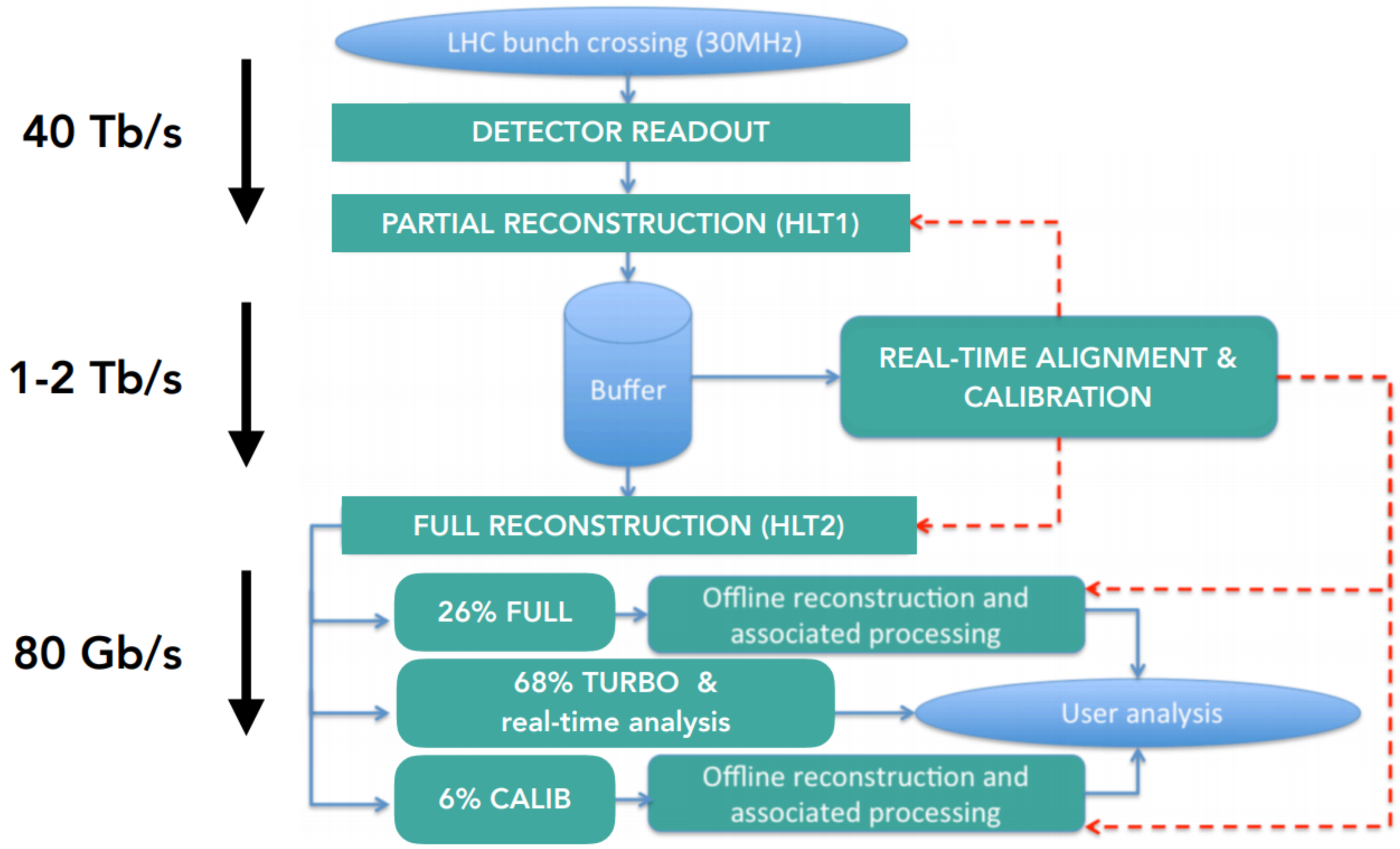
Idea of Real-Time Analysis

- Real-Time Analysis – efficient decision about data in the full online mode
- Keeping only a signal and suppress any unnecessary information about event



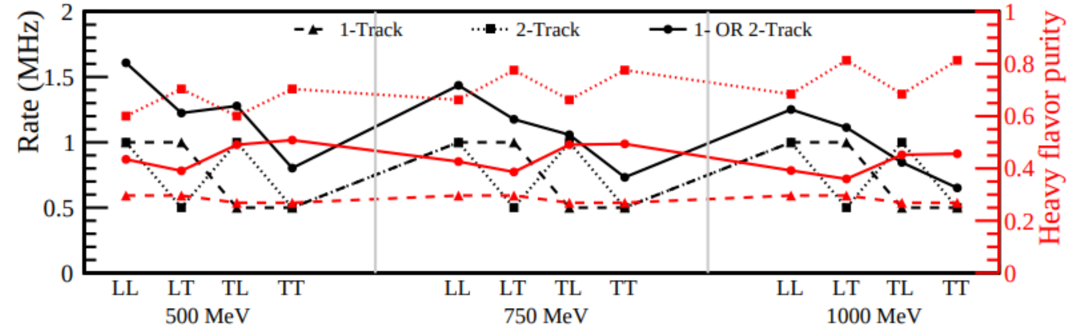
Idea of Real-Time Analysis

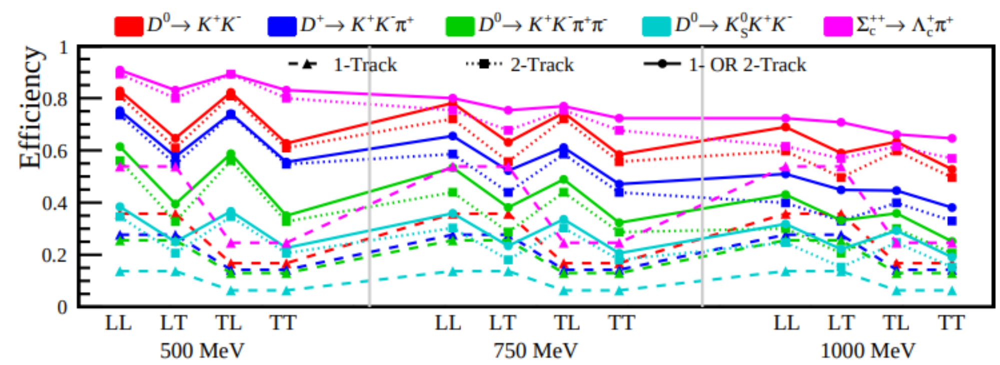
- Real-Time Analysis – efficient decision about data in the full online mode
- Keeping only a signal and suppress any unnecessary information about event
- Triggerless readout, full software trigger on 30 MHz (readout 40 MHz, around 40 Tb/s)

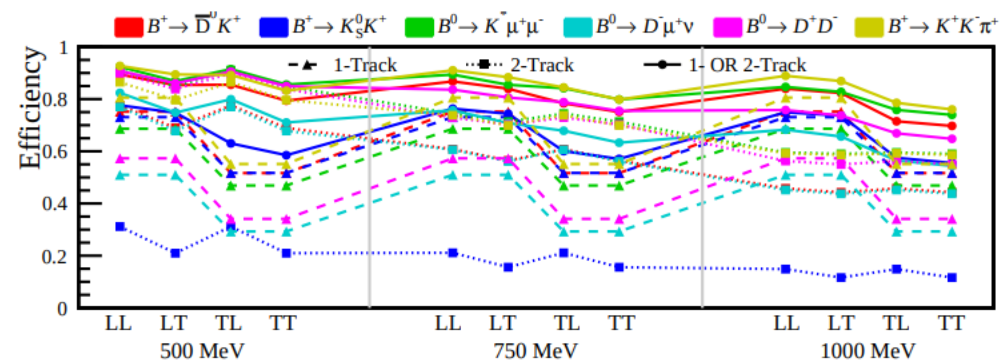


Run III - HLT1

- Full charged particle track reconstruction
- Some inclusive selection
  - 1-Track trigger based on individual displaced tracks
  - 2-Track trigger based on secondary vertices
- Different kinematic thresholds for each configuration
- Reduction of event rate approximately by factor 30
- Simplified Kalman filtering in VELO stage



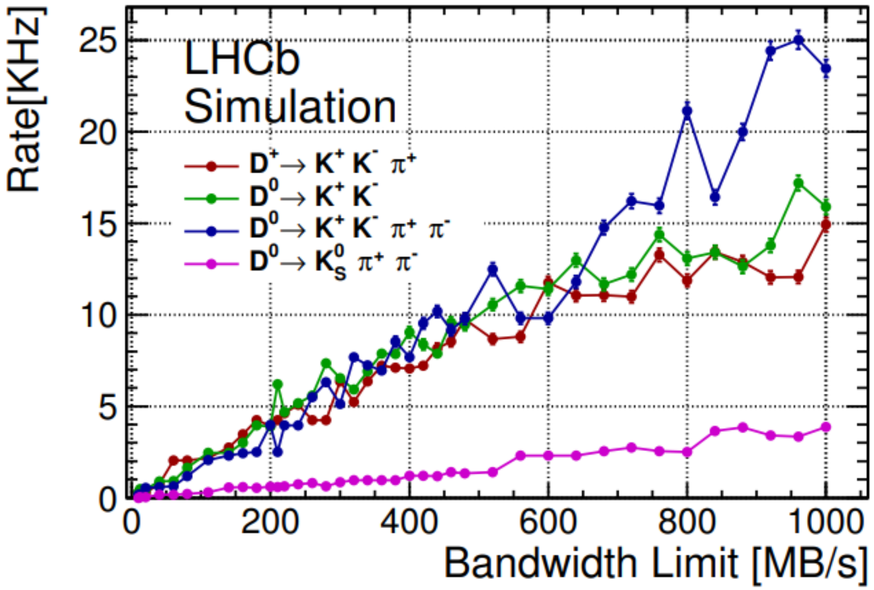


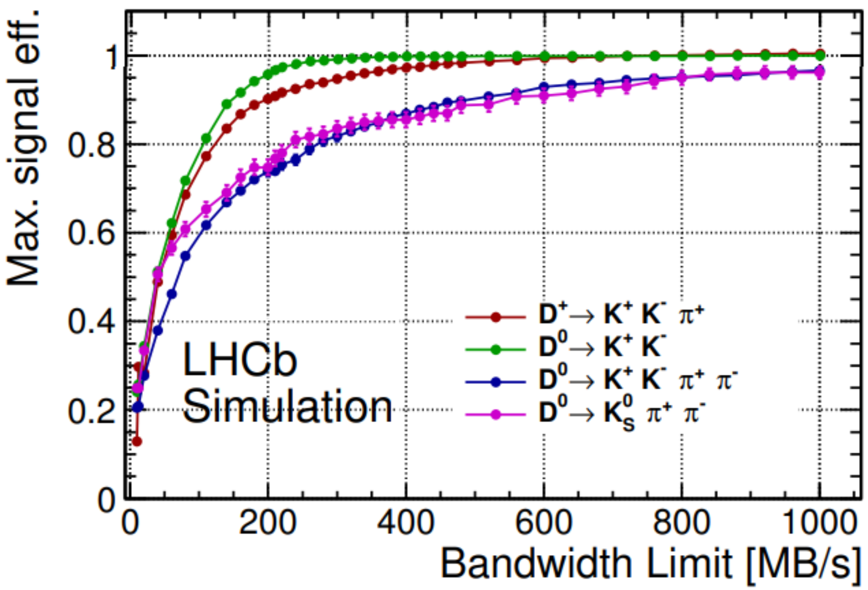


LHCb-PUB-2017-006

Run III - HLT2

- Fully aligned and calibrated detector on this stage
- HLT2 should achieve offline-quality track reconstruction
- Main limit is a bandwidth → more than 500 specific lines expected
- Extensive usage of MVA-based lines is expected
- Ongoing studies on general tracks from B and D decays (arXiv: 1903.01360)
- With full reconstruction on HLT2 stage, offline CPU can be used for another purposes



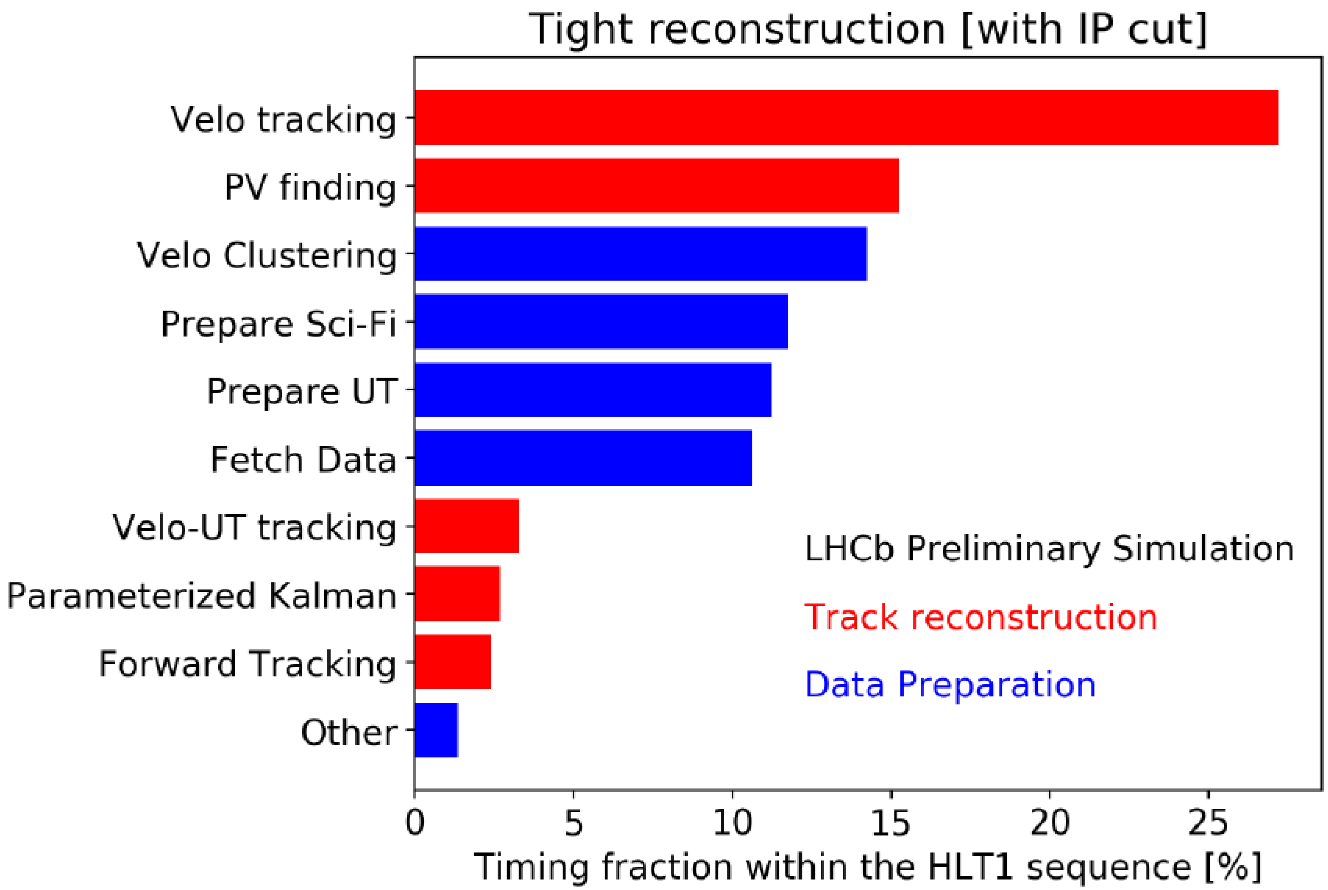


LHCb-PUB-2017-006

A way to 30 MHz ...
(2017 - 2018)

Data Preparation

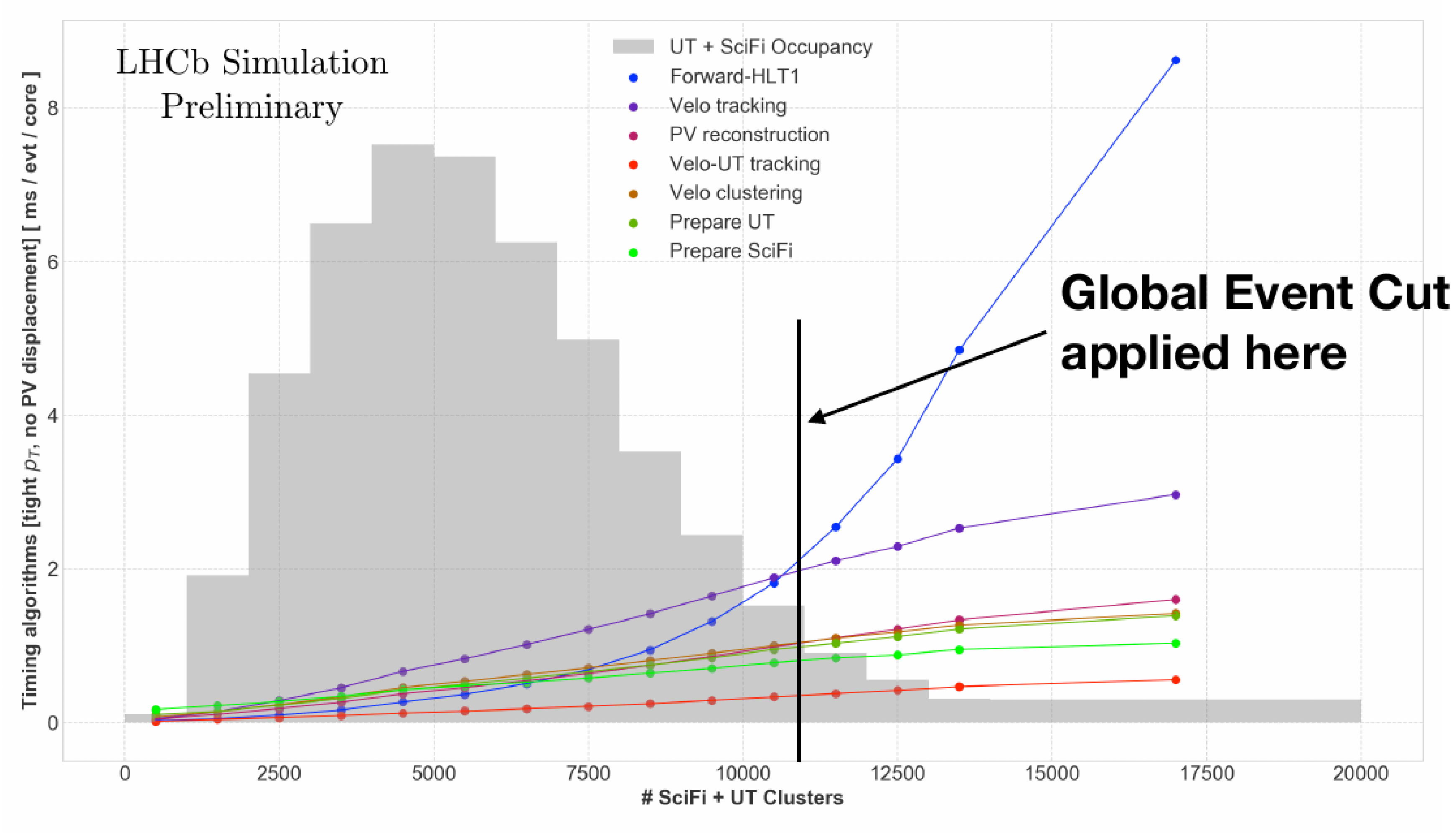
- Data must be prepared for algorithms
  - Very CPU intensive – up to 50 % of data processing time
- Need of fast and optimized code
  - Trade-off between efficiency and speed? How much efficiency we can lose?



LHCb-TDR-017

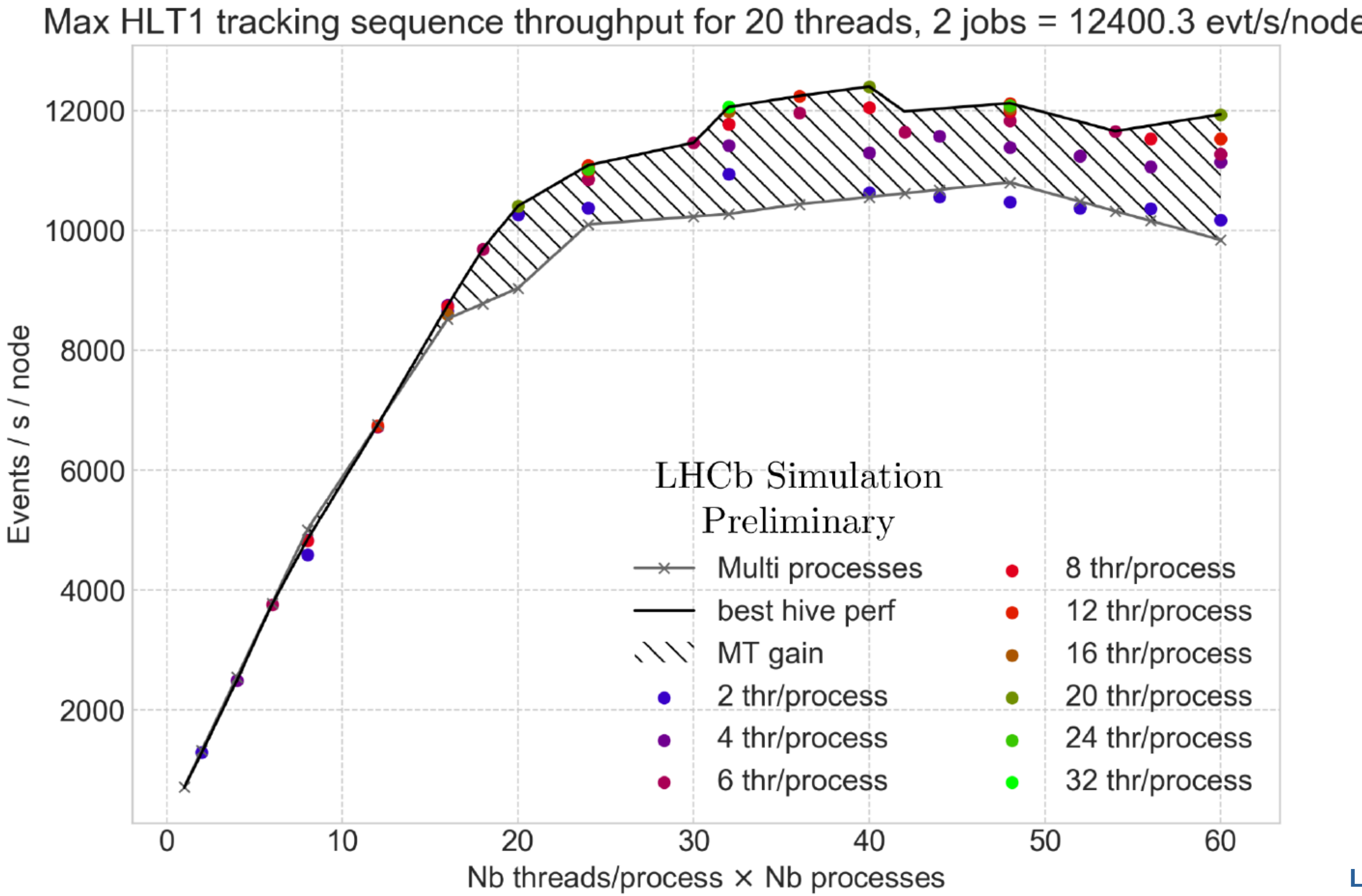
Data Occupancy effect

- Events with higher occupancy take longer to process – possibly critical effect
- However, such an events typically are not representative of signal topologies
- Most of algorithms scale linearly, issue with forward tracking



Multi-thread HLT1 sequence

- Run II HLT1 framework is a single-thread algorithm
- Porting to multi-thread architecture: 20 % gain just from more threads

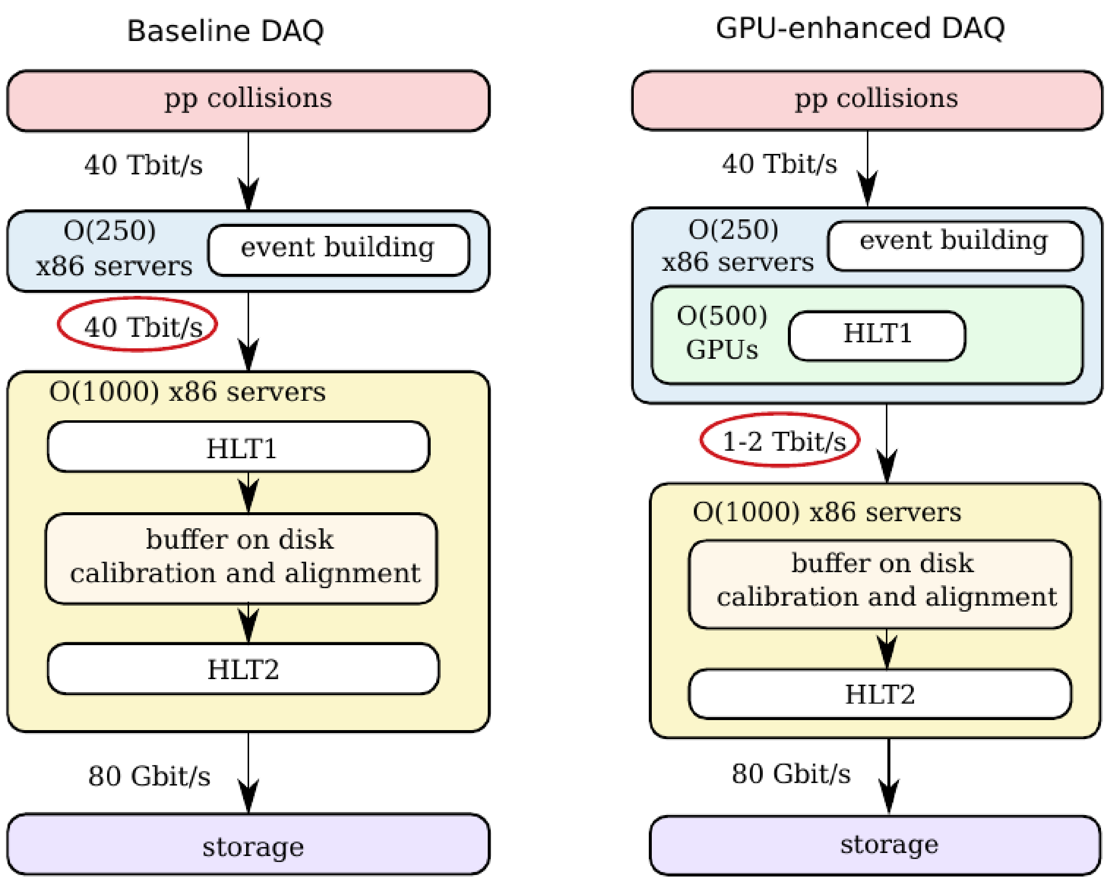


LHCb-TDR-017

A possible non-x86 approach to HLT1?

Accelerators - GPU

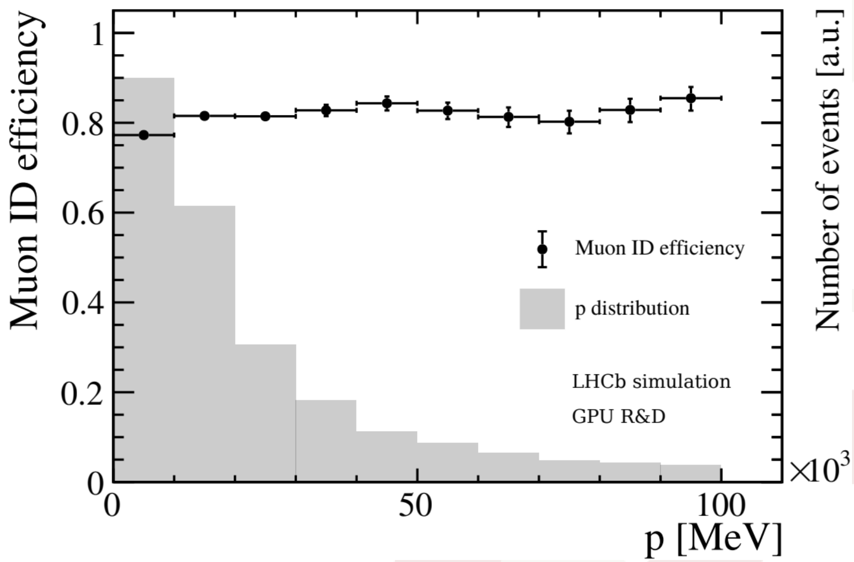
- HLT1 is by definition a parallel system with huge computation load
- Each raw event is relatively small (~ 100 kB)
- Highly parallel computation – a perfect match with modern GPU
- Usage of GPUs in HLT1 → The Allen project (gitlab)

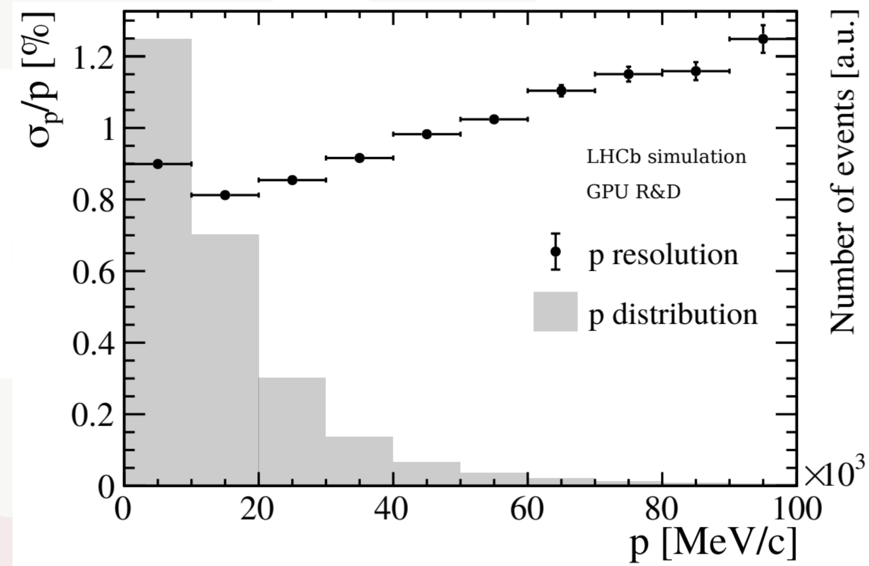


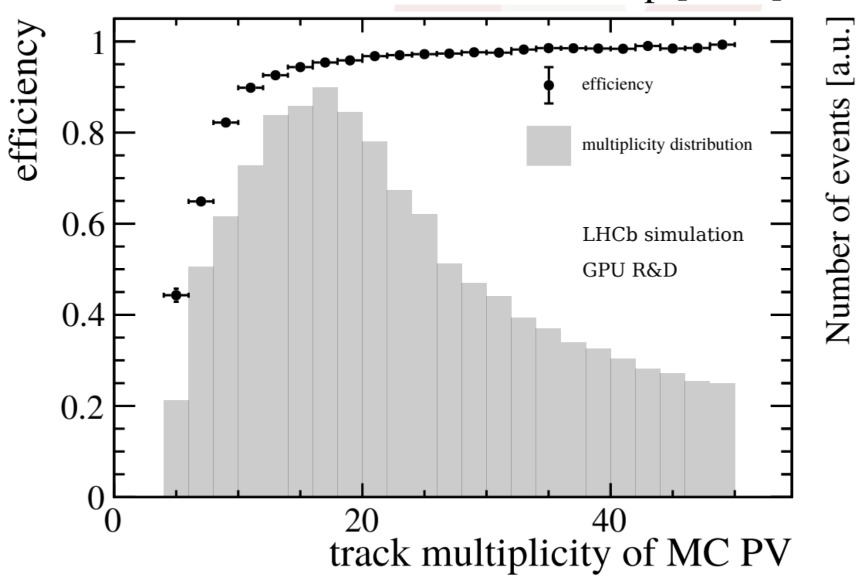
Accelerators - GPU

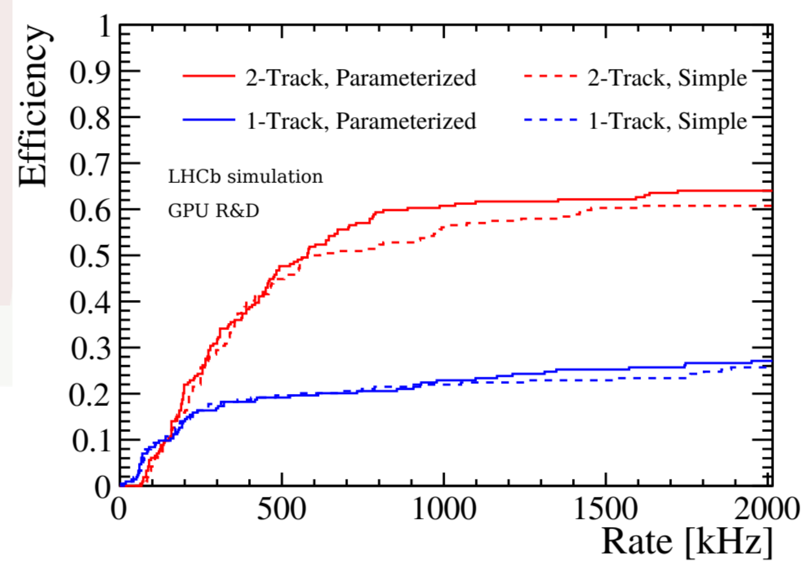
- Usage of GPUs in HLT1 → The Allen project (gitlab)
- Possible solution for Run III

LHCb-FIGURE-2019-009





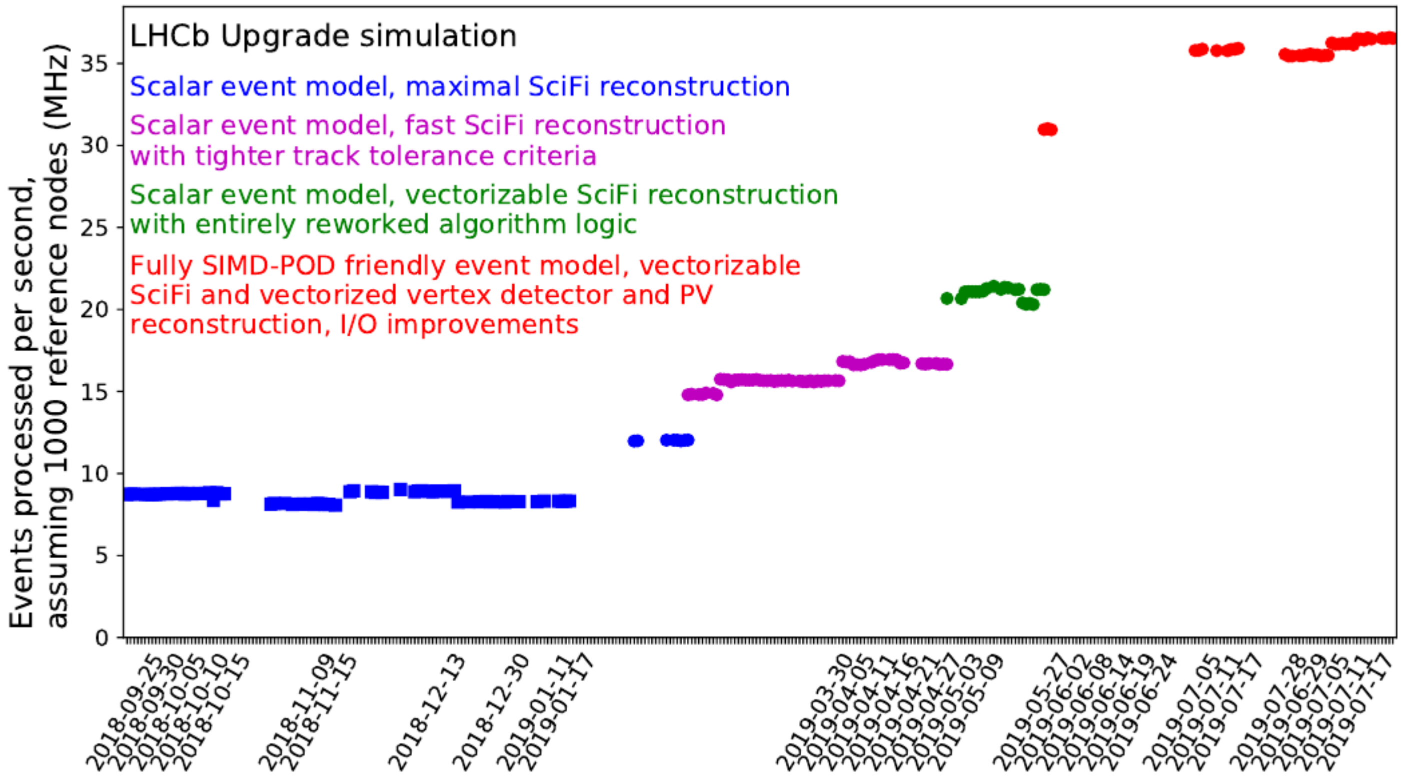




30 MHz era
(2019 onwards)

HLT1 on 30 MHz

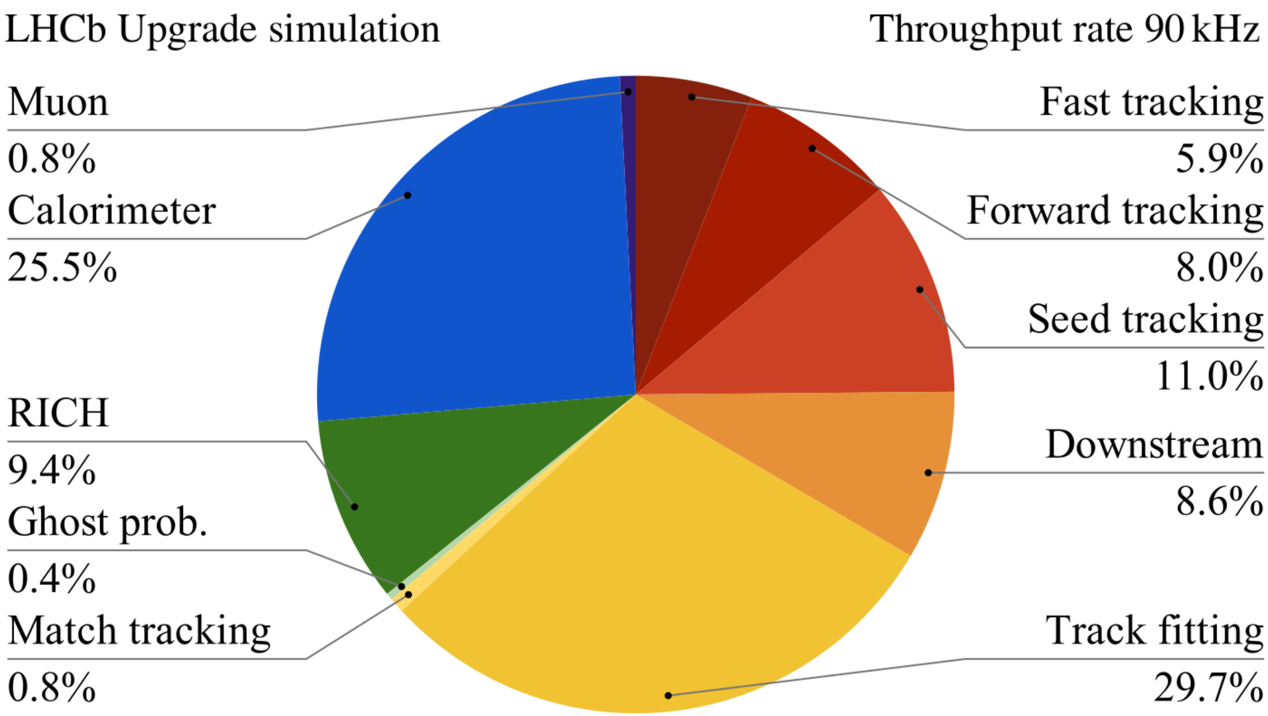
- HLT1 throughput evolution between the autumn of 2018 and summer of 2019
- Reference system: Intel Xeon E5-2630
- Nominal upgrade data taking conditions, automatic nightly test of throughput
- We are in the process of testing the new AMD EPYC architecture and see a major price/performance improvement, precise numbers to be confirmed in the next months



LHCb-FIGURE-2019-002

HLT2 throughput

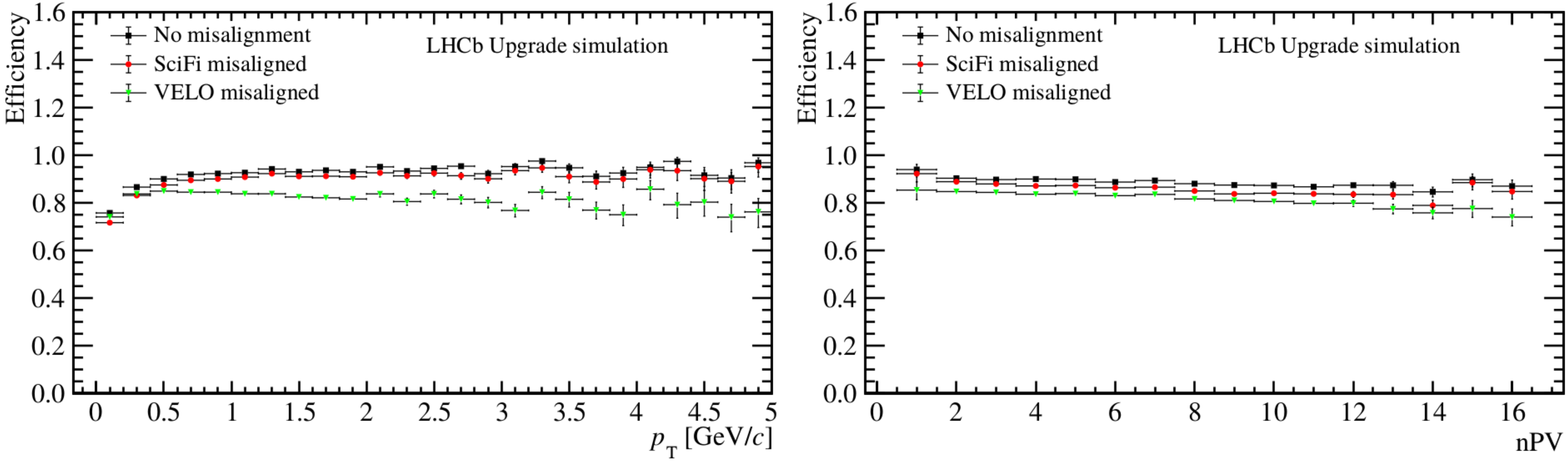
- Status of the HLT 2 reconstruction throughput as on 2019/09/01
- Reference system: Intel Xeon E5-2630
- Nominal upgrade data taking conditions, automatic nightly test of throughput
- Based on simulated minimum bias sample passing HLT1 selection



LHCb-FIGURE-2019-004

Data challenges

- Most of the data will be using Turbo lines, mistake in any part of data processing can be fatal
- The Upgrade framework must be thoroughly tested as whole
- Testing is an iterative procedure closely following progress of the RTA project and Upgrade itself
- Tracking efficiencies using the first estimates of misalignment values for VELO, SciFi and combination of both, without running alignment

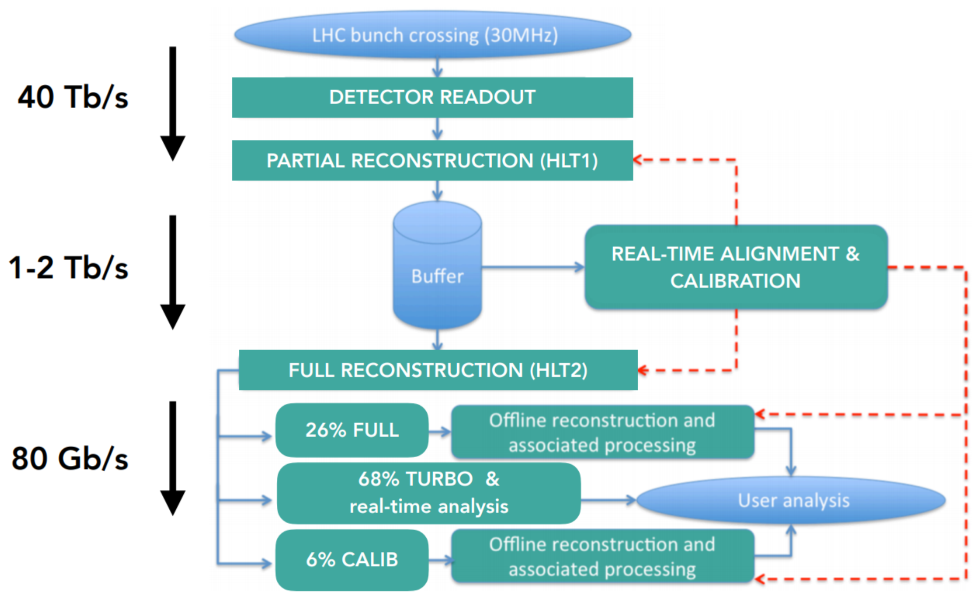


LHCb-FIGURE-2019-005

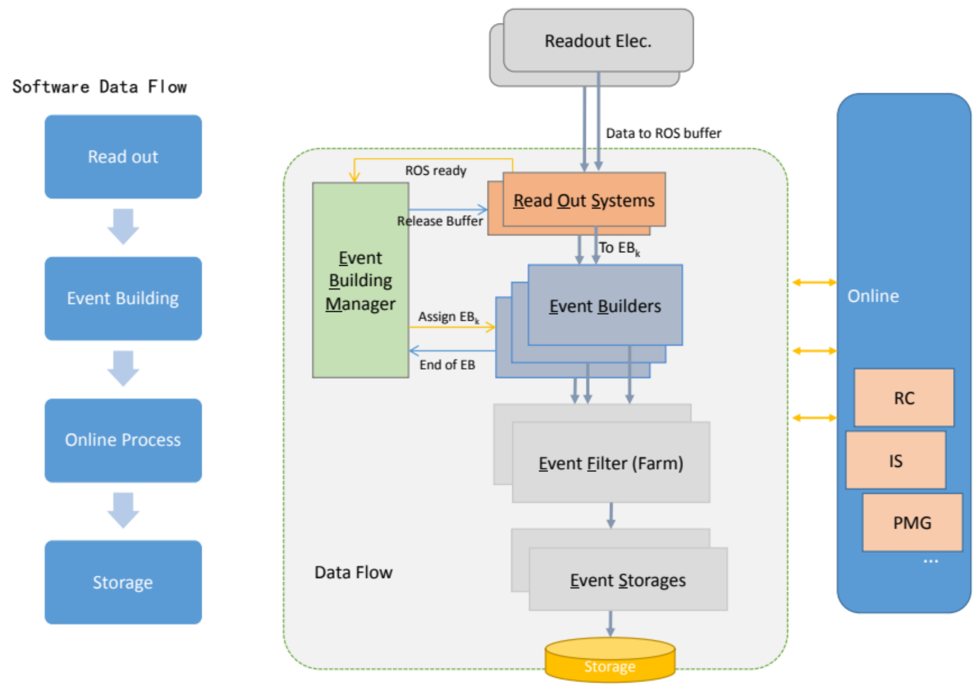
RTA and CepC

- Possibility of using Turbo-like and RTA approach in CepC experiments?
- CepC: cleaner environment and smaller event size
- Expected total raw date rate ~ 2 TB/s on 100 kHz L1 trigger (CepC TDR)
- RTA and Turbo-like system could save computing resources and improve versatility of detectors

LHCb (RUN III)



CepC (TDR)



IHEP-CEPC-DR-2018-02

Conclusion

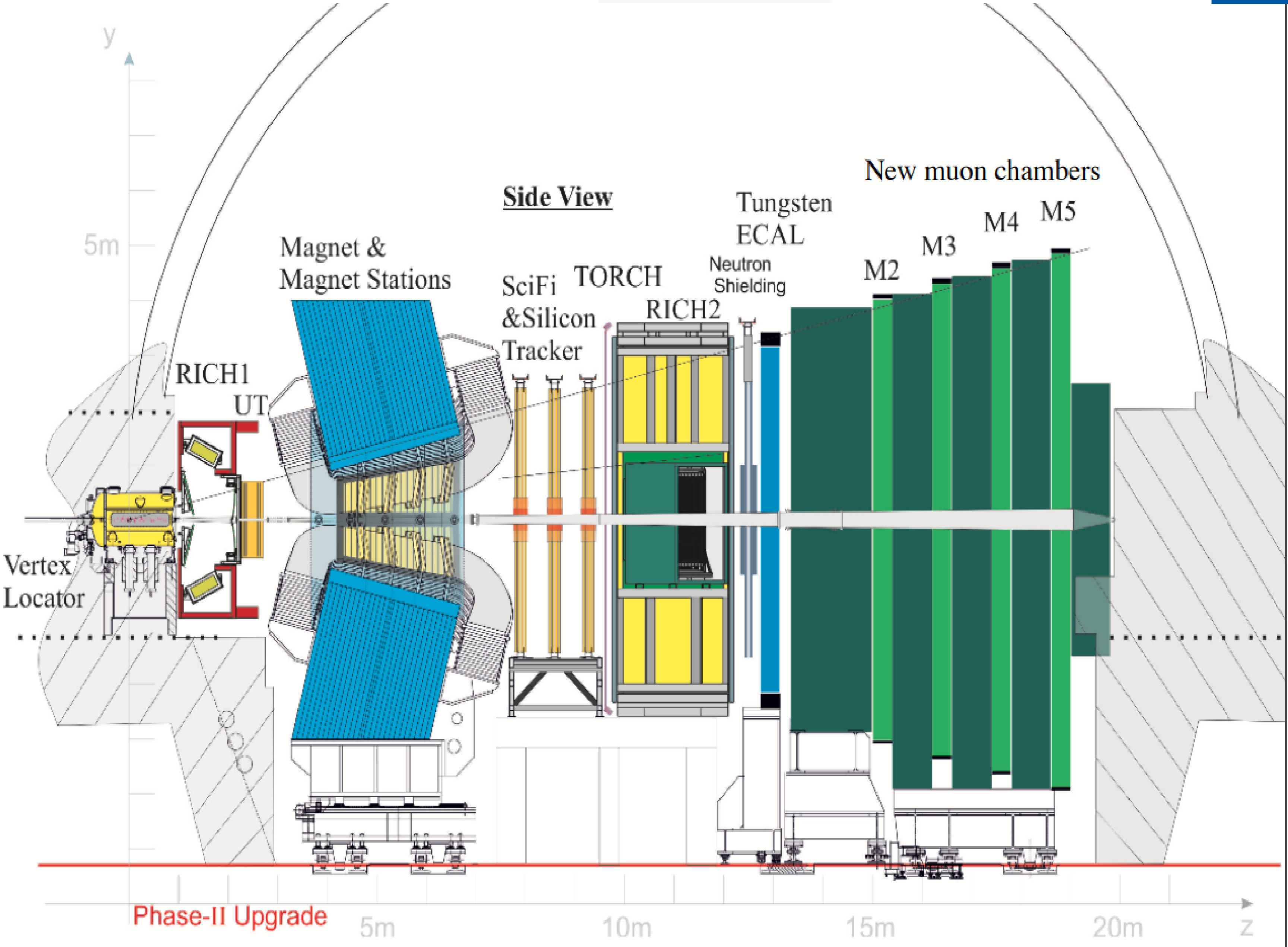
- LHCb detector finished two successful periods of data taking
- During the current long shutdown LHCb detector is going through a large-scale upgrade of both hardware and software part of detector
- 5x higher luminosity during Run III
- Run I utilized a ‘HEP-standard’ trigger strategy
- Run II shown a need for faster trigger system and usefulness of a online data reconstruction
- RTA is a novel approach for hadron collider experiments enabling significant increase of the recorded data
- CepC aims to be state-of-the-art Higgs factory
- RTA approach can help utilize a full potential of this new accelerator and detectors

Thank you for your attention

BACKUP Slides

Planned LHCb upgrades

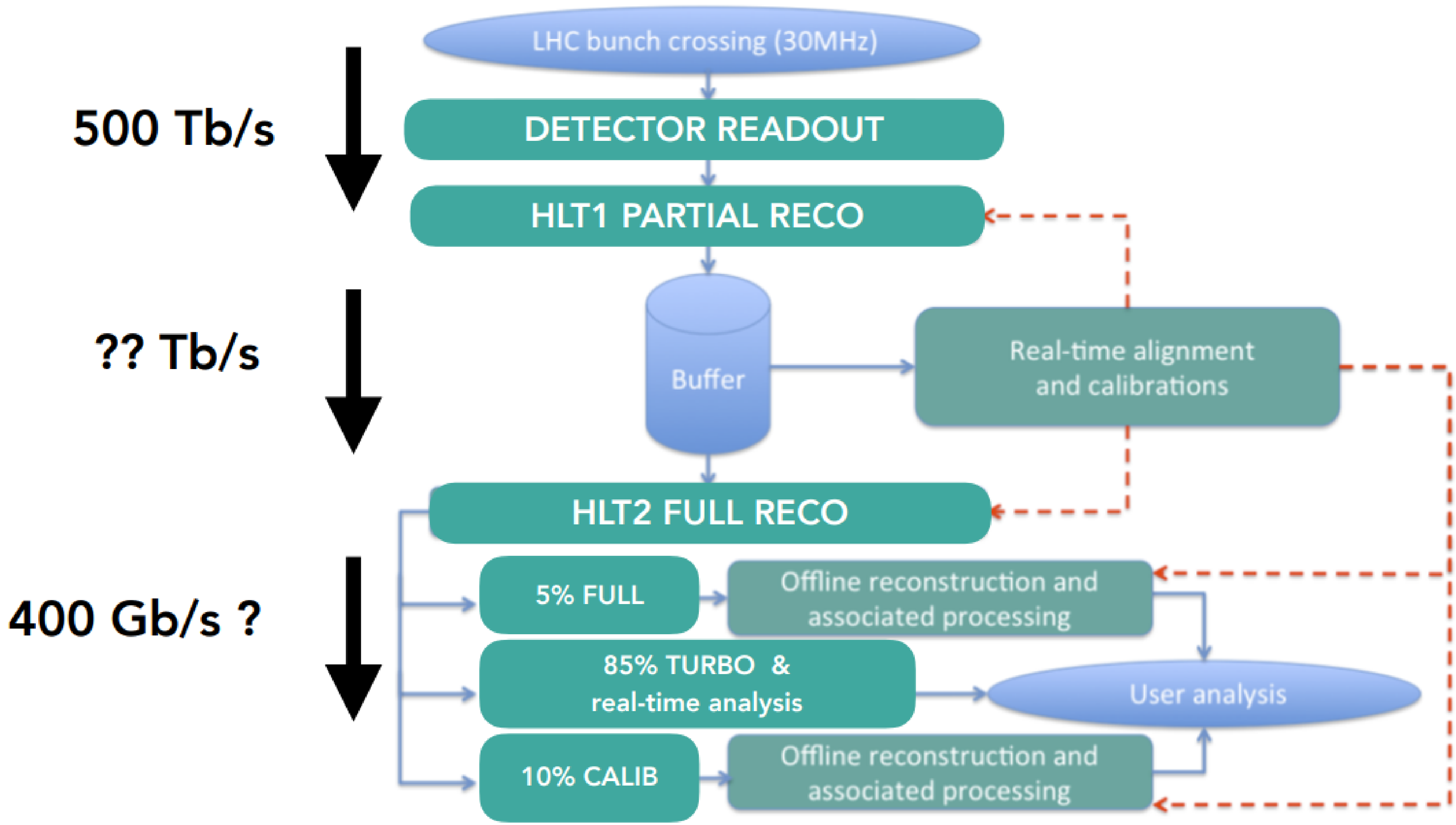
LHCb upgrade Phase II (Run V)



CERN-LHCC-2017-003

Real-Time Analysis at Run V

- Real-Time Analysis – efficient decision about data in the full online mode
- Run V – HL-LHC



LHCb upgrade Phase II (Run V)

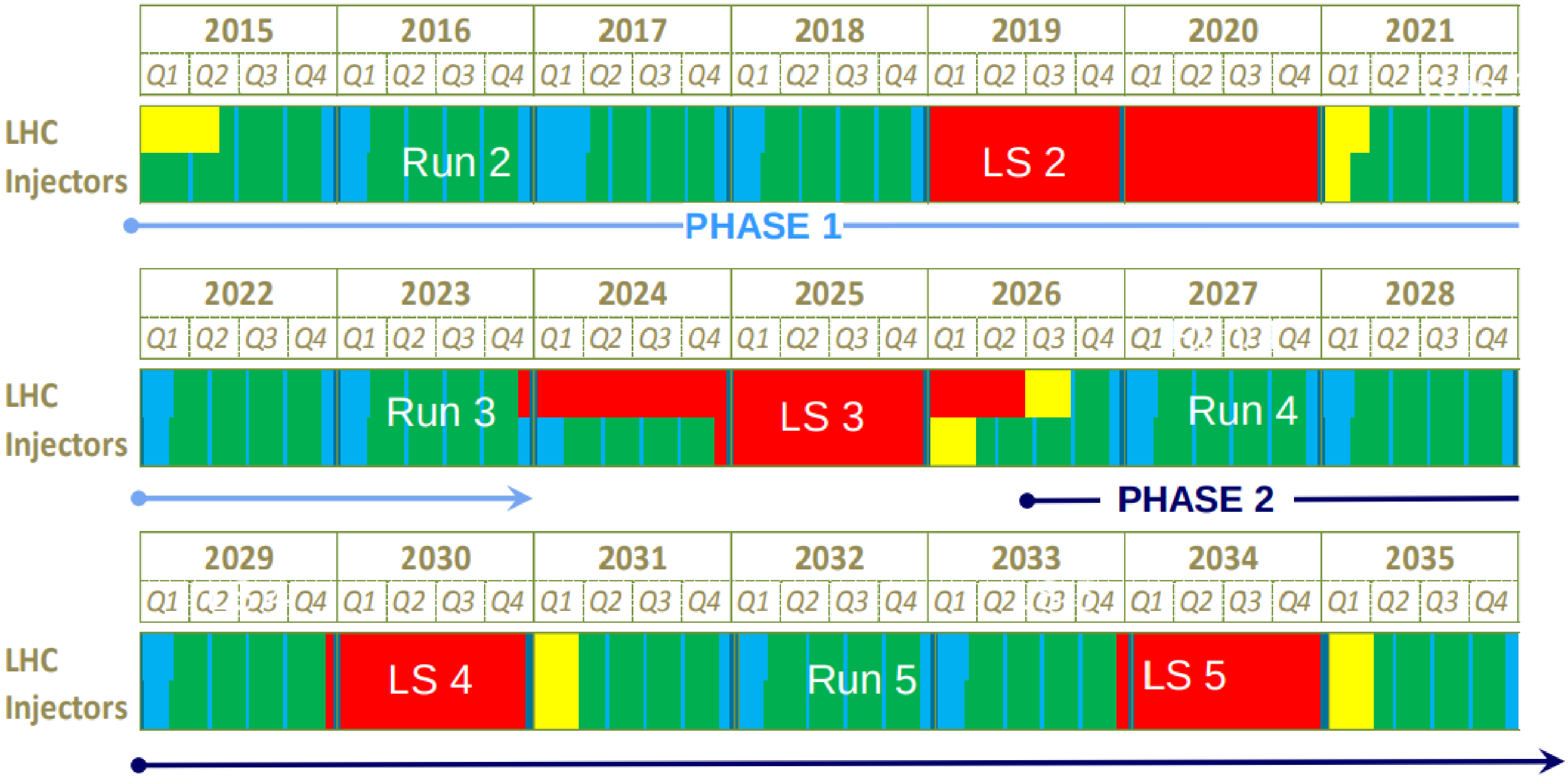
Topics and observables	Experimental reach	Remarks
EW Penguins		
Global tests in many $b \to s\mu^+\mu^-$ modes with full set of precision observables; lepton universality tests; $b \to dl^+l^-$ studies	*e.g.* 440k $B^0 \to K^*\mu^+\mu^-$ & 70k $\Lambda_b^0 \to \Lambda\mu^+\mu^-$; Phase-II $b \to d\mu^+\mu^- \approx$ Run-1 $b \to s\mu^+\mu^-$ sensitivity.	Phase-II ECAL required for lepton universality tests.
Photon polarisation		
$\mathcal{A}^\Delta$ in $B_s^0 \to \phi\gamma$; $B^0 \to K^*e^+e^-$; baryonic modes	Uncertainty on $\mathcal{A}^\Delta \approx 0.02$; $\sim 10k$ $\Lambda_b^0 \to \Lambda\gamma$, $\Xi_b \to \Xi\gamma$, $\Omega_b^- \to \Omega\gamma$	Strongly dependent on performance of ECAL.
$b \to cl^-\bar{\nu}_l$ lepton-universality tests		
Polarisation studies with $B \to D^{(*)}\tau^-\bar{\nu_\tau}$; τ^-/μ^- ratios with B_s^0, Λ_b^0 and B_c^+ modes	*e.g.* 8M $B \to D^*\tau^-\bar{\nu_\tau}$, $\tau^- \to \mu^-\bar{\nu_\mu}\nu_\tau$ & $\sim 100k$ $\tau^- \to \pi^-\pi^+\pi^-(\pi^0)\nu_\tau$	Additional sensitivity expected from low-p tracking.
$B_s^0, B^0 \to \mu^+\mu^-$		
$R \equiv \mathcal{B}(B^0 \to \mu^+\mu^-)/\mathcal{B}(B_s^0 \to \mu^+\mu^-)$; $\tau_{B_s^0 \to \mu^+\mu^-}$; *CP* asymmetry	Uncertainty on $R \approx 20\%$ Uncertainty on $\tau_{B_s^0 \to \mu^+\mu^-} \approx 0.03\,\text{ps}$	
LFV τ decays		
$\tau^- \to \mu^+\mu^-\mu^-$, $\tau^- \to h^+\mu^-\mu^-$, $\tau^- \to \phi\mu^-$	Sensitive to $\tau^- \to \mu^+\mu^-\mu^-$ at 10^{-9}	Phase-II ECAL valuable for background suppression.
CKM tests		
γ with $B^- \to DK^-$, $B_s^0 \to D_s^+K^-$ *etc.*	Uncertainty on $\gamma \approx 0.4°$	Additional sensitivity expected in *CP* observables from Phase-II ECAL and low-p tracking.
ϕ_s with $B_s^0 \to J/\psi K^+K^-$, $J/\psi\pi^+\pi^-$	Uncertainty on $\phi_s \approx 3\,\text{mrad}$	
$\phi_s^{s\bar{s}s}$ with $B_s^0 \to \phi\phi$	Uncertainty on $\phi_s^{s\bar{s}s} \approx 8\,\text{mrad}$	
$\Delta\Gamma_d/\Gamma_d$	Uncertainty on $\Delta\Gamma_d/\Gamma_d \sim 10^{-3}$	Approach SM value.
Semileptonic asymmetries $a_{\text{sl}}^{d,s}$	Uncertainties on $a_{\text{sl}}^{d,s} \sim 10^{-4}$	Approach SM value for a_{sl}^d.
$\lvert V_{ub}\rvert/\lvert V_{cb}\rvert$ with Λ_b^0, B_s^0 and B_c^+ modes	*e.g.* $120k$ $B_c^+ \to D^0\mu^-\bar{\nu_\mu}$	Significant gains achievable from thinning or removing RF-foil.
Charm		
CP-violation studies with $D^0 \to h^+h^-$, $D^0 \to K_S^0\pi^+\pi^-$ and $D^0 \to K^\mp\pi^\pm\pi^+\pi^-$	*e.g.* 4×10^9 $D^0 \to K^+K^-$; Uncertainty on $A_\Gamma \sim 10^{-5}$	Access *CP* violation at SM values.
Strange		
Rare decay searches	Sensitive to $K_S^0 \to \mu^+\mu^-$ at 10^{-12}	Additional sensitivity possible with downstream trigger enhancements.

LHC timeline

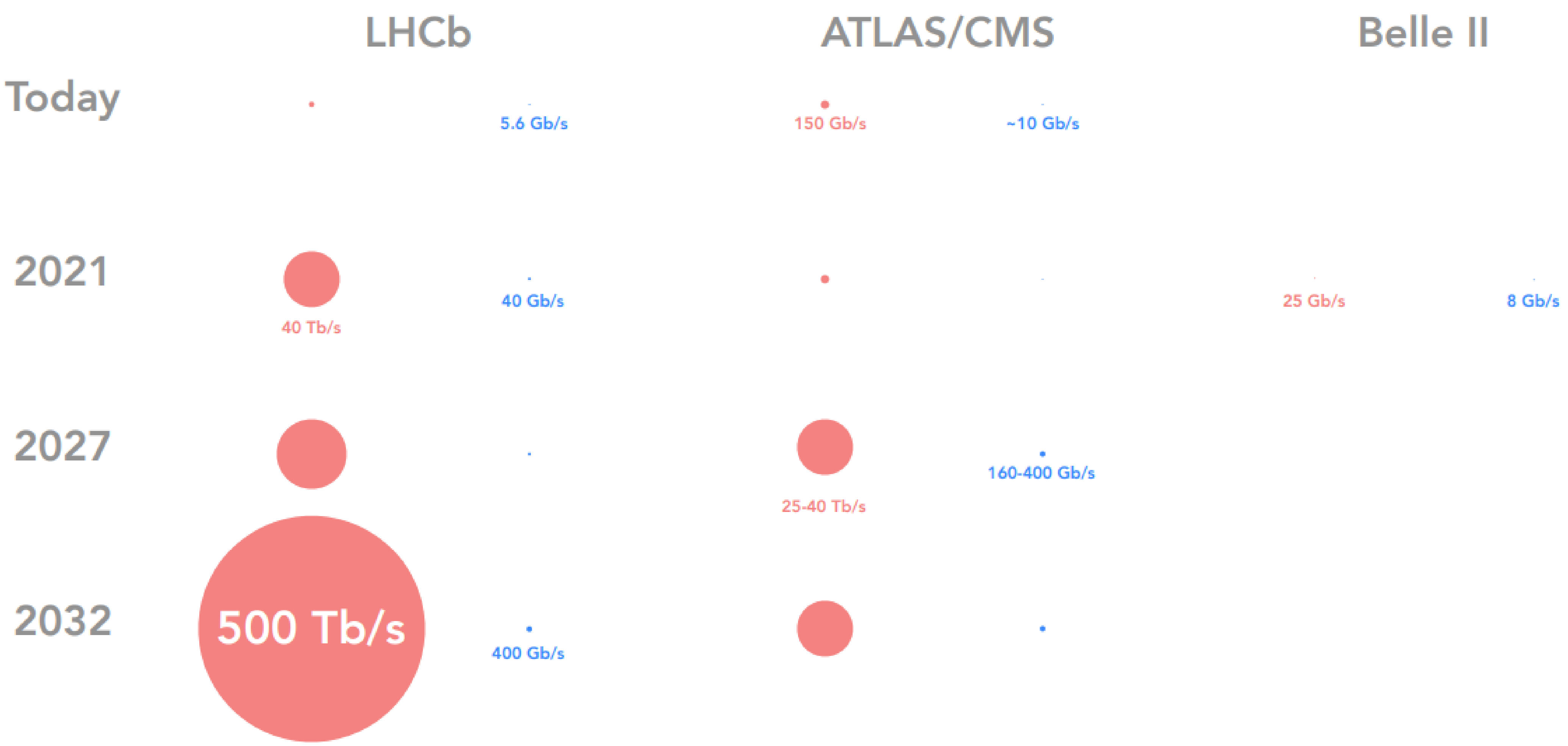
LHC roadmap: according to MTP 2016-2020 V1

LS2 starting in 2019 => 24 months + 3 months BC
LS3 LHC: starting in 2024 => 30 months + 3 months BC
Injectors: in 2025 => 13 months + 3 months BC

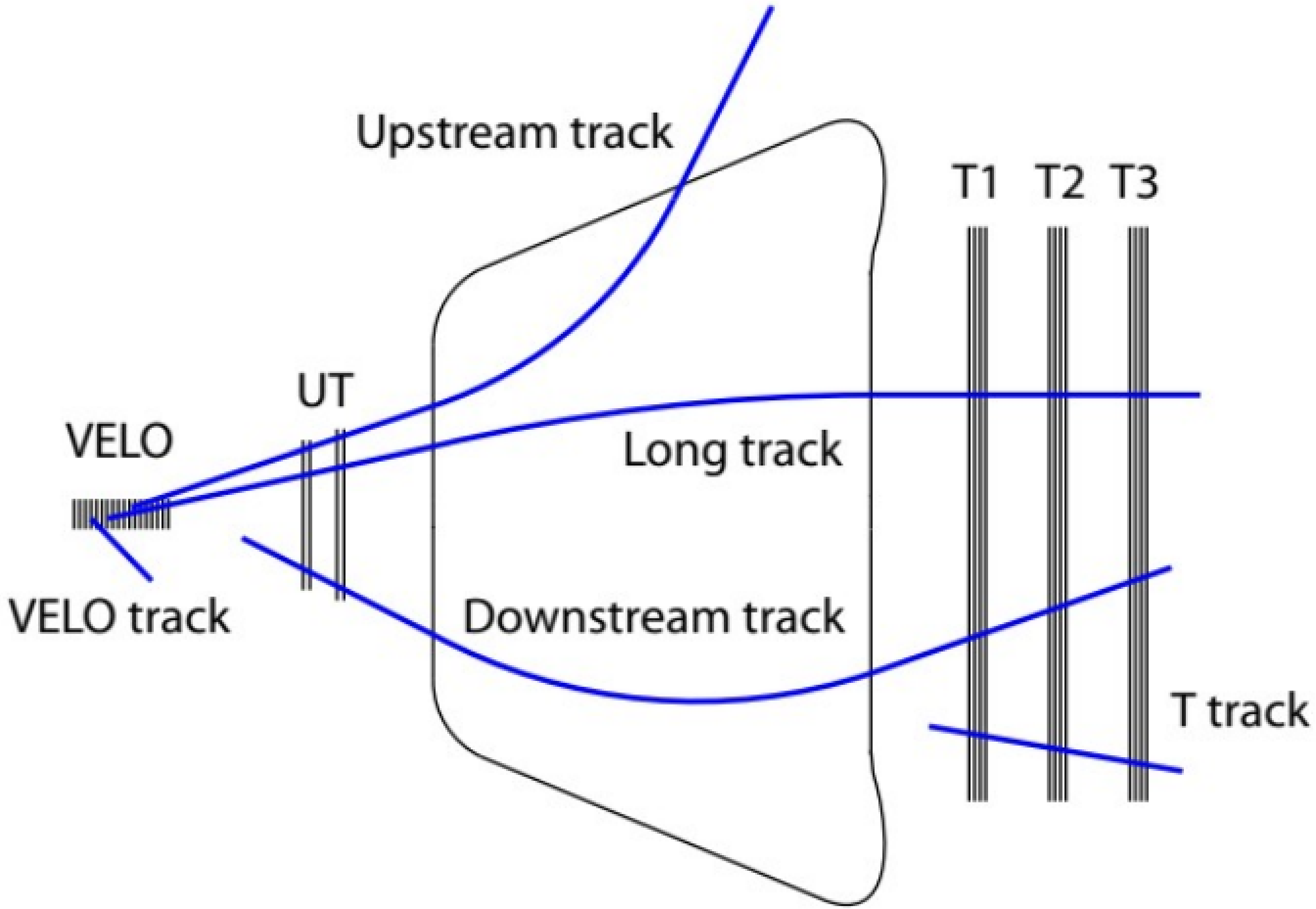
- Physics
- Shutdown
- Beam commissioning
- Technical stop



Amount of data in HEP



Type of track in the LHCb



Structure of RTA project

- Subdetector-like organization
- 36 institutes, around 50 FTE (>100 people involved)
- A long term project also in charge of maintenance after deployment