Status of the Real Time Analysis project at LHCb

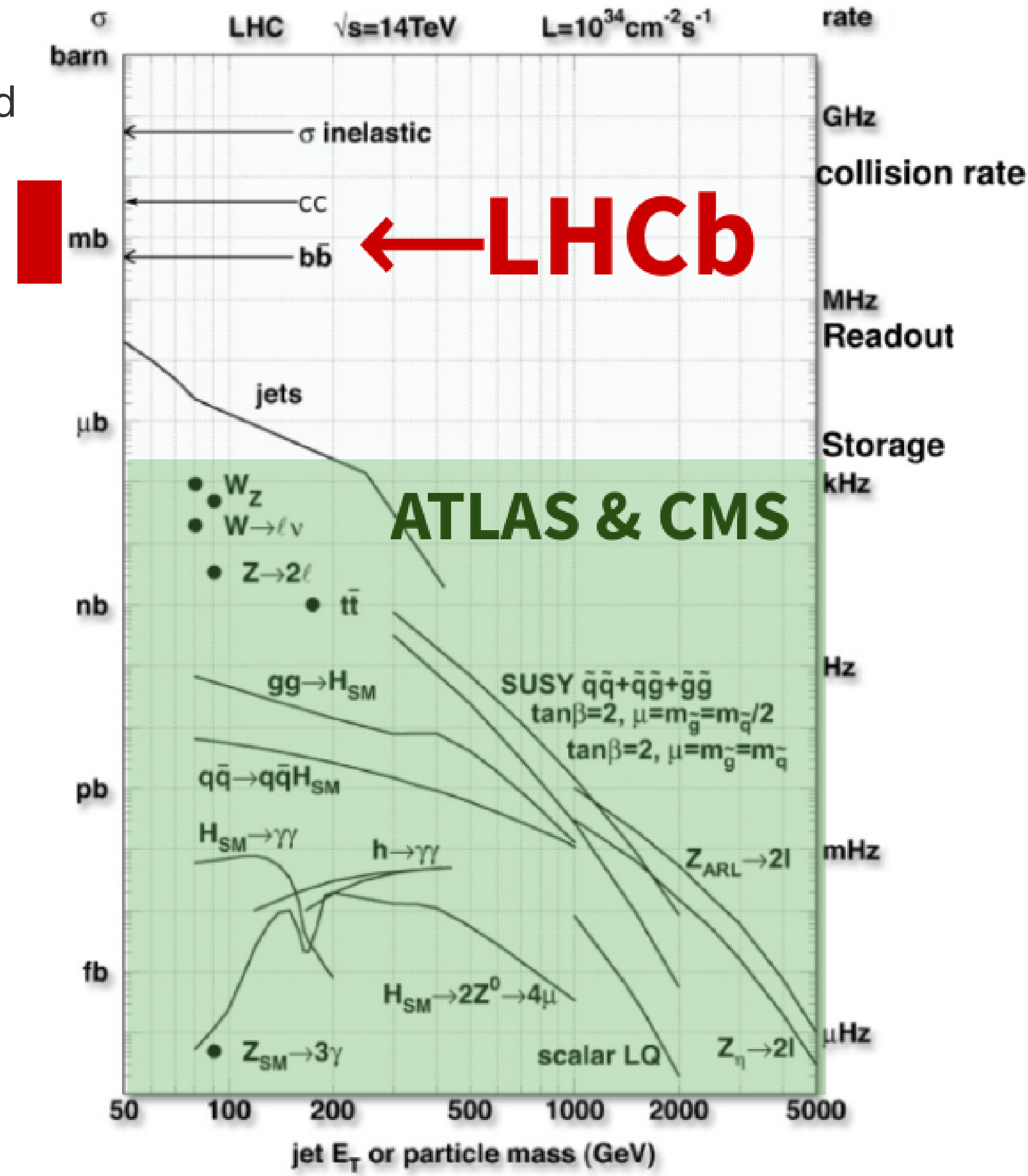
The 6th China LHC Physics Workshop
Online, 2020/11/08

SAUR Miroslav
on behalf of the RTA project at LHCb

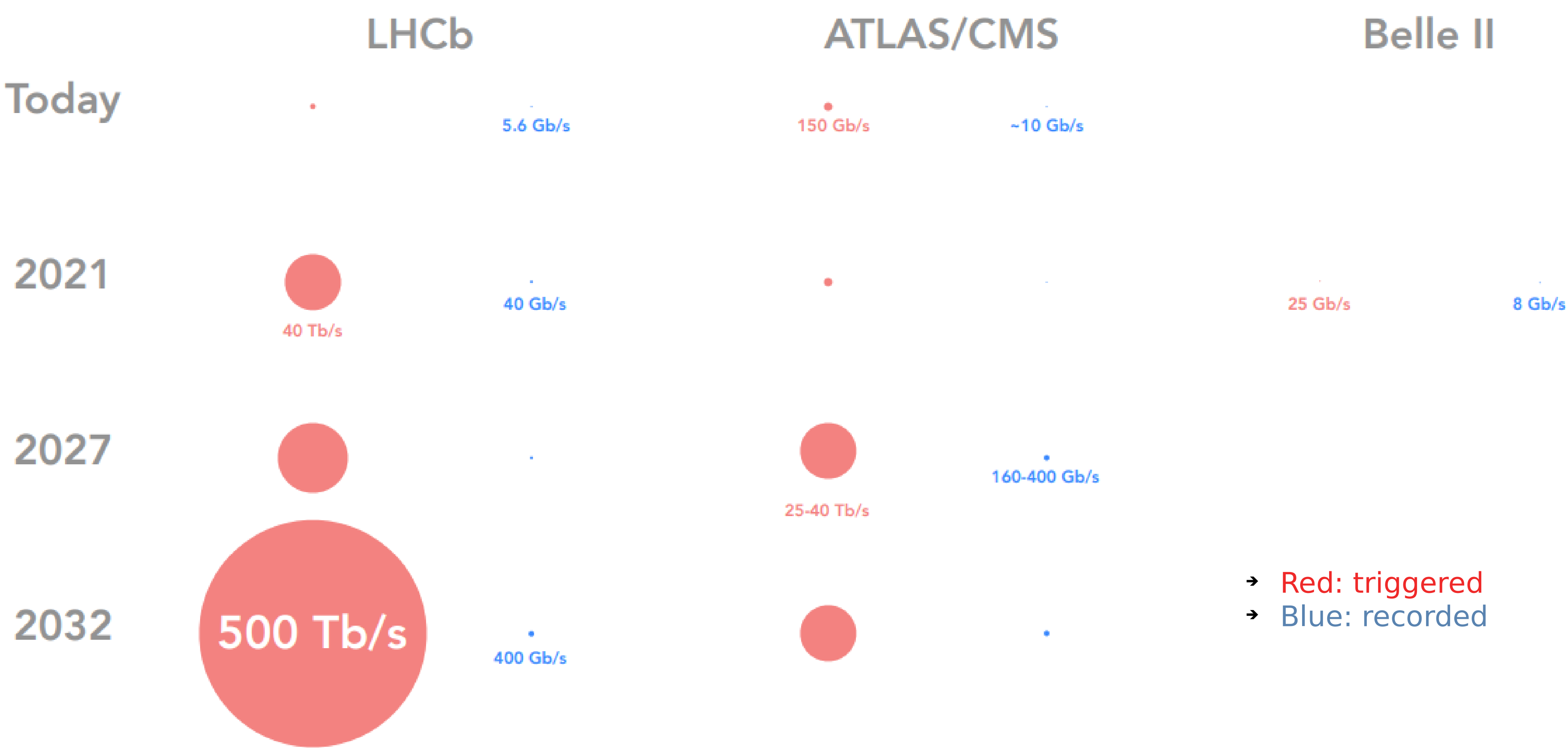
University of Chinese Academy of Sciences
and
TU Dortmund

Requirements for trigger

- Triggering is a crucial part of data taking
- Decision of what physics can be recorded
- Resources demanding operation
- Hard constraints: 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



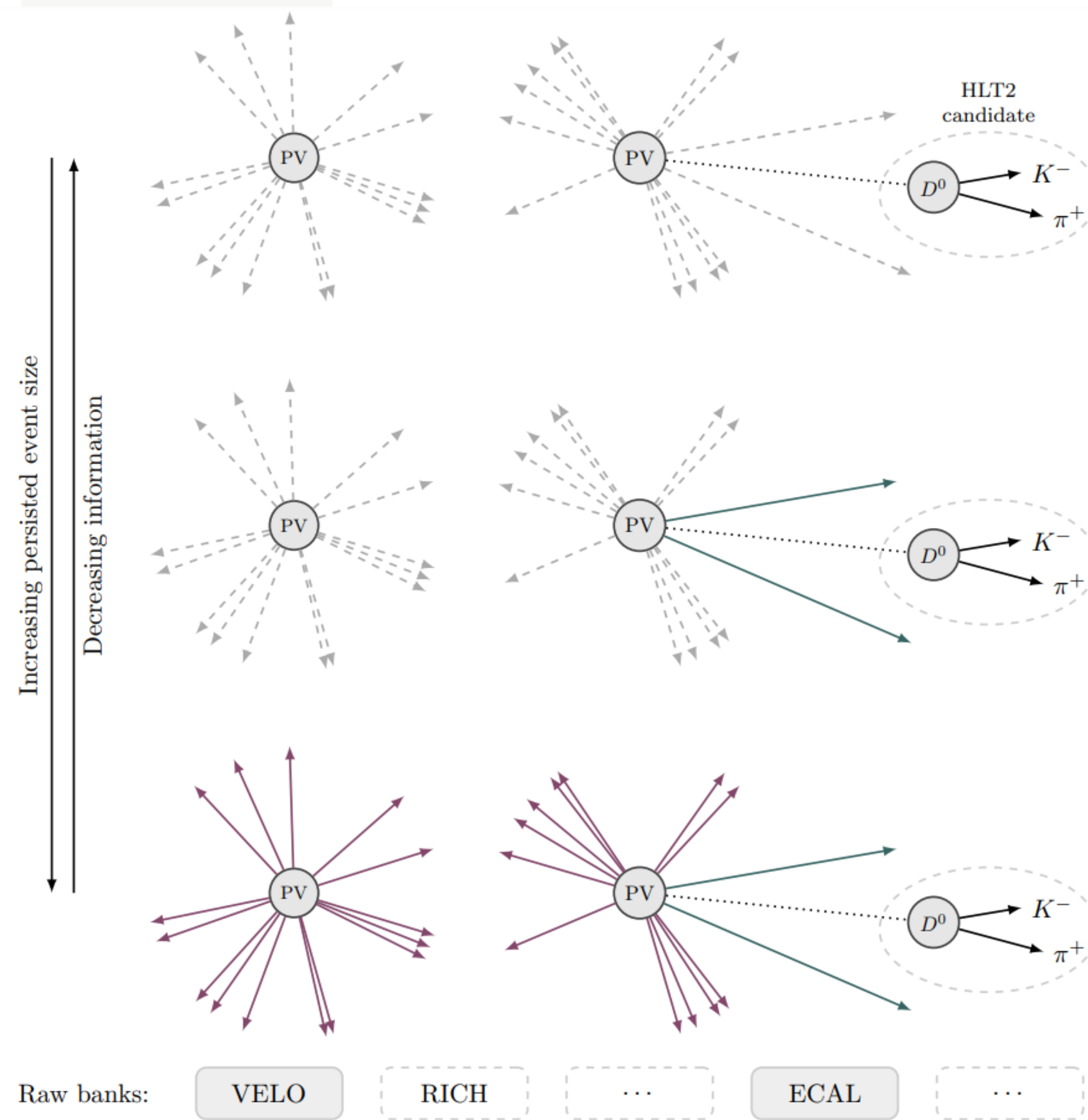
Amount of data in HEP



- Red: triggered
- Blue: recorded

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

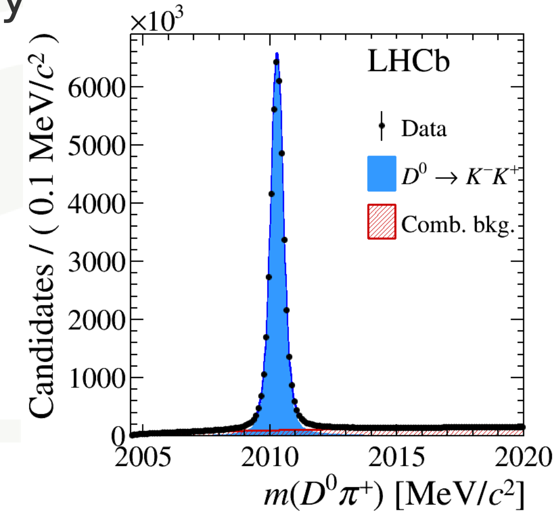
- Extensively used during the Run2 at LHCb
  - Around 30 % of the trigger rate is Turbo – almost all Charm physics
  - But only about 10 % of the bandwidth!
  - Approximately 2/3 lines keep raw detector information (Turbo SP)
- Significant reduction of data size → more events at same bandwidth

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

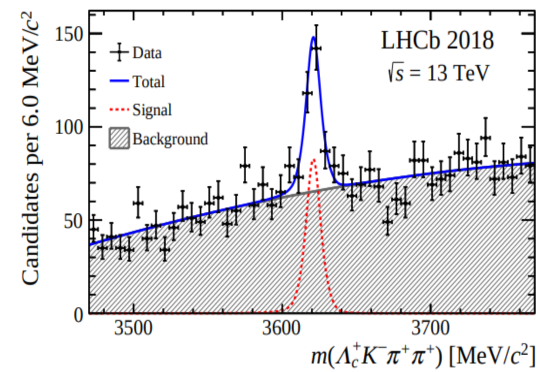
- Turbo stream relies on full detector alignment and calibration within the trigger phase

Accomplishments of Turbo

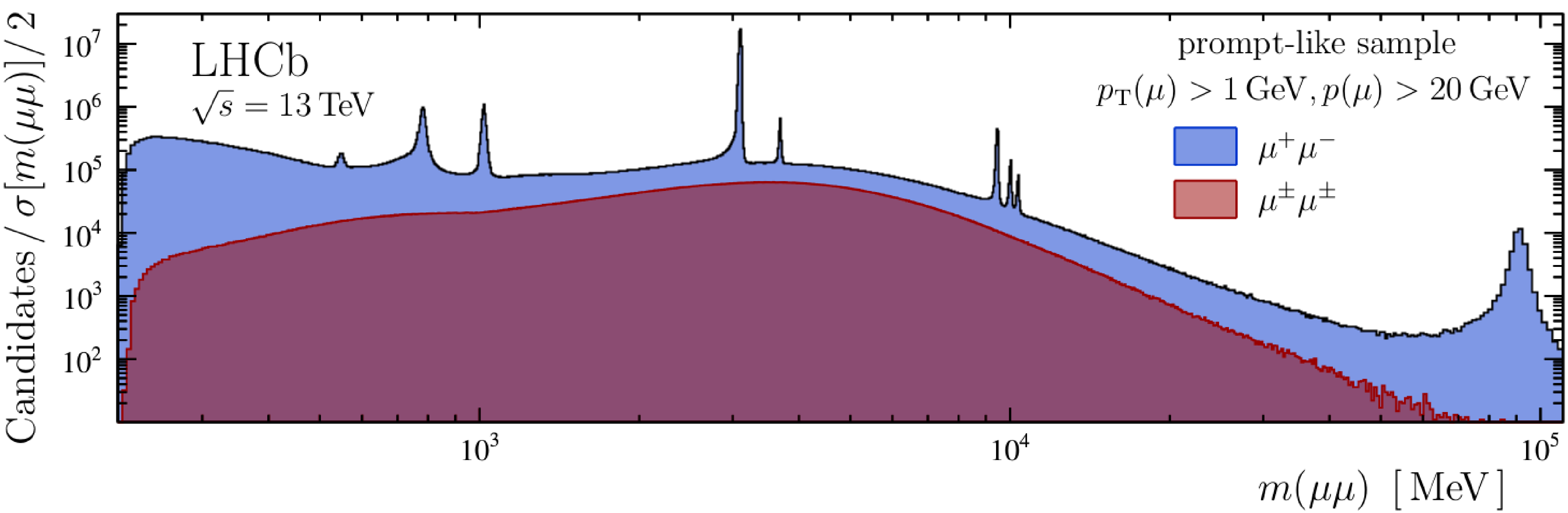
- Turbo lines proved to be necessary for keeping a significant charm program during Run II
- Suitable for a broad range of physics – from high to low rate
- One of the key ideas of Upgrade



PRL 122 (2019) 211803



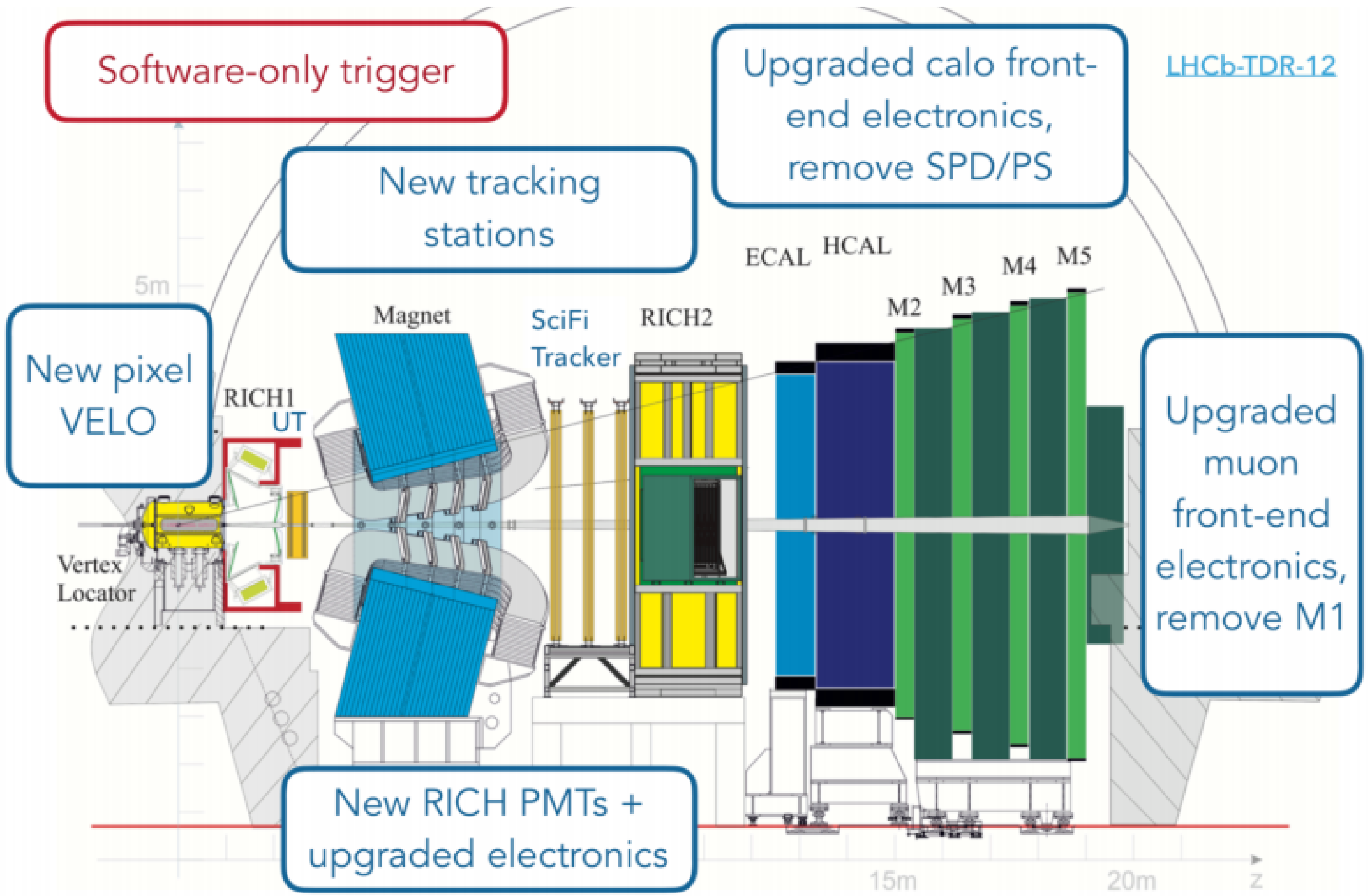
Sci. China-Phys. Mech. Astron. 63, 221062 (2020)



PRL 120 (2018) 061801

LHCb Upgrade I (Run 3)

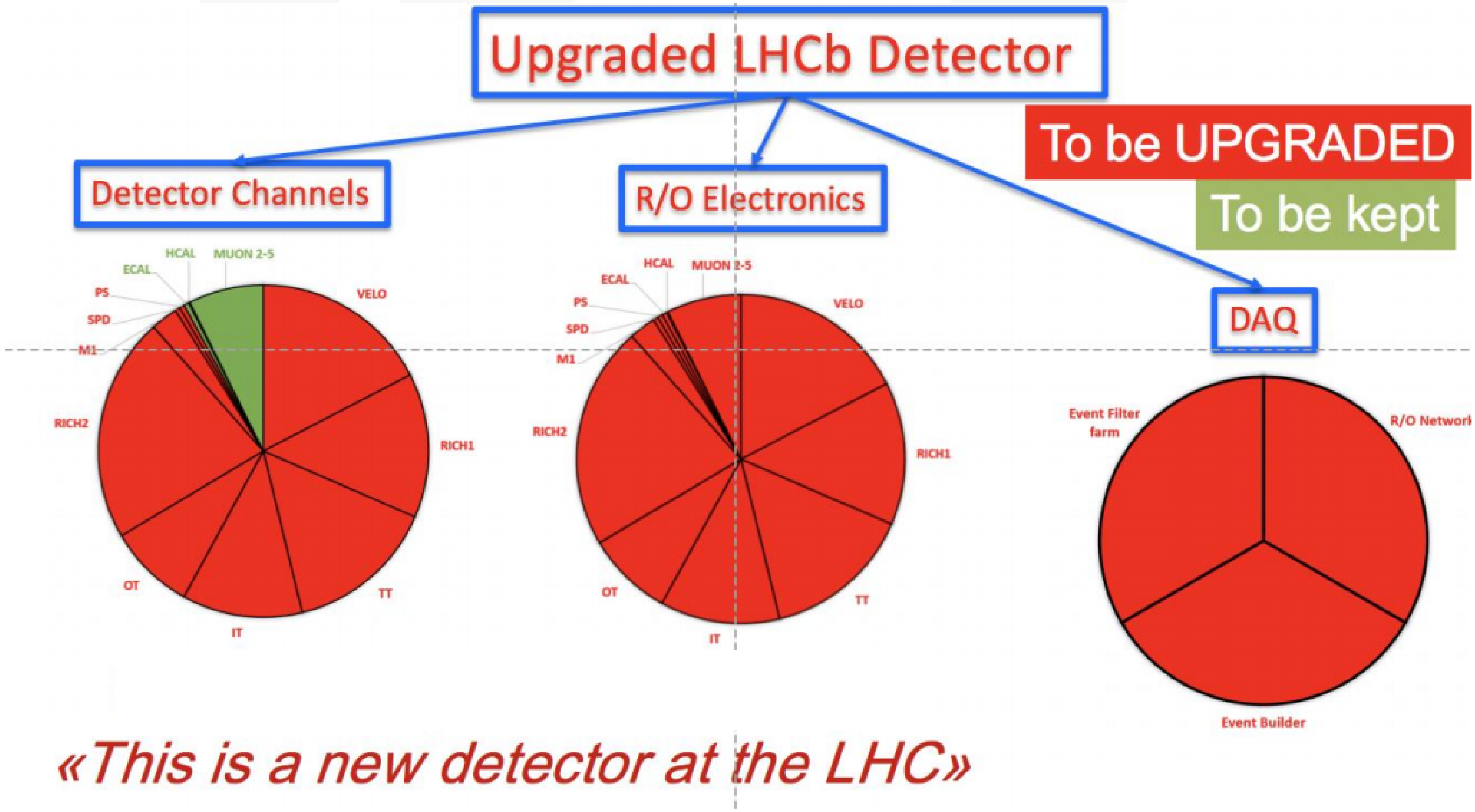
- Luminosity will increase 5x times at same collision energy of 13 TeV
- Aim is to maintain the same performance as during Run2 (more than 90 %)



LHCb upgrade overview: LI Yiming, Monday 9:50

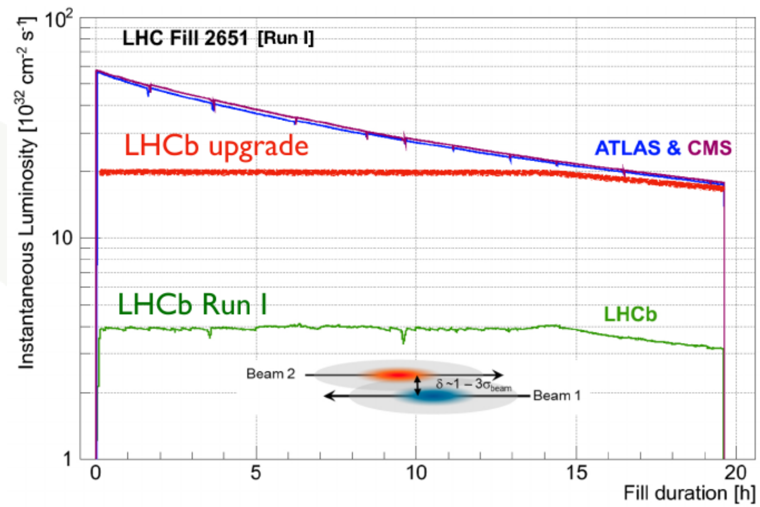
LHCb Upgrade I (Run 3)

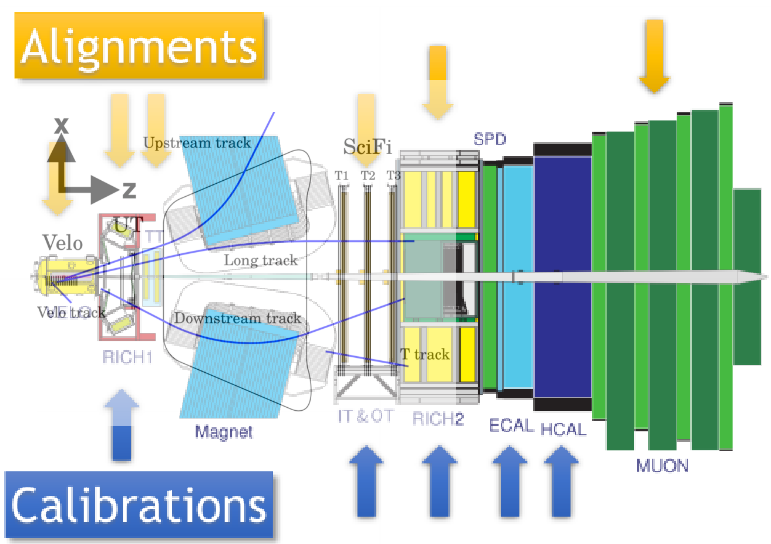
- Luminosity will increase 5x times at same collision energy of 13 TeV
- Aim is to maintain the same performance as during Run2 (more than 90 %)
- A large scale Upgrade!

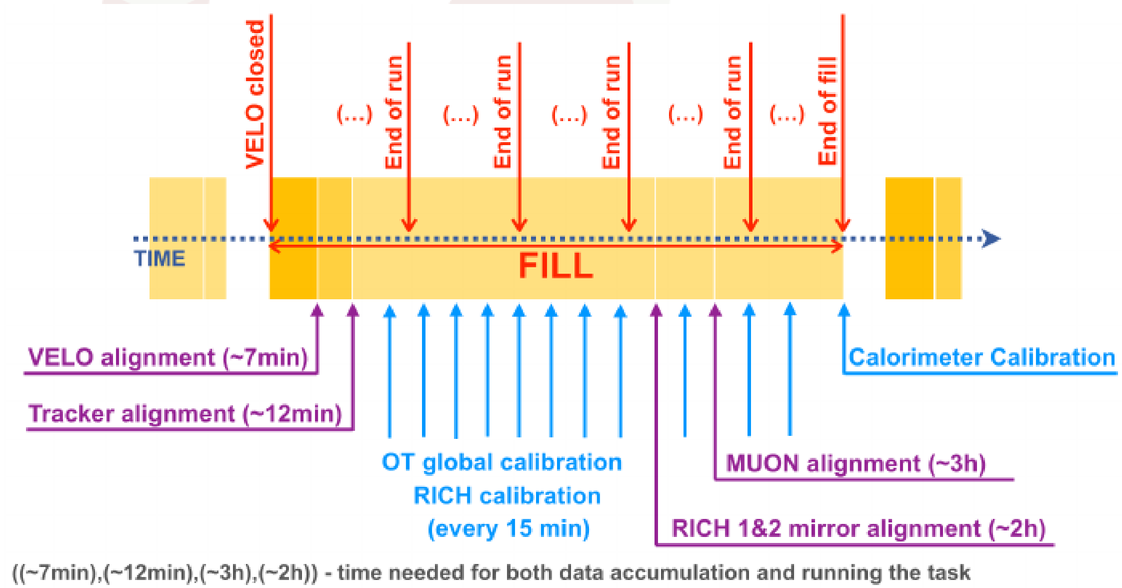


LHCb Upgrade I (Run 3)

- 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

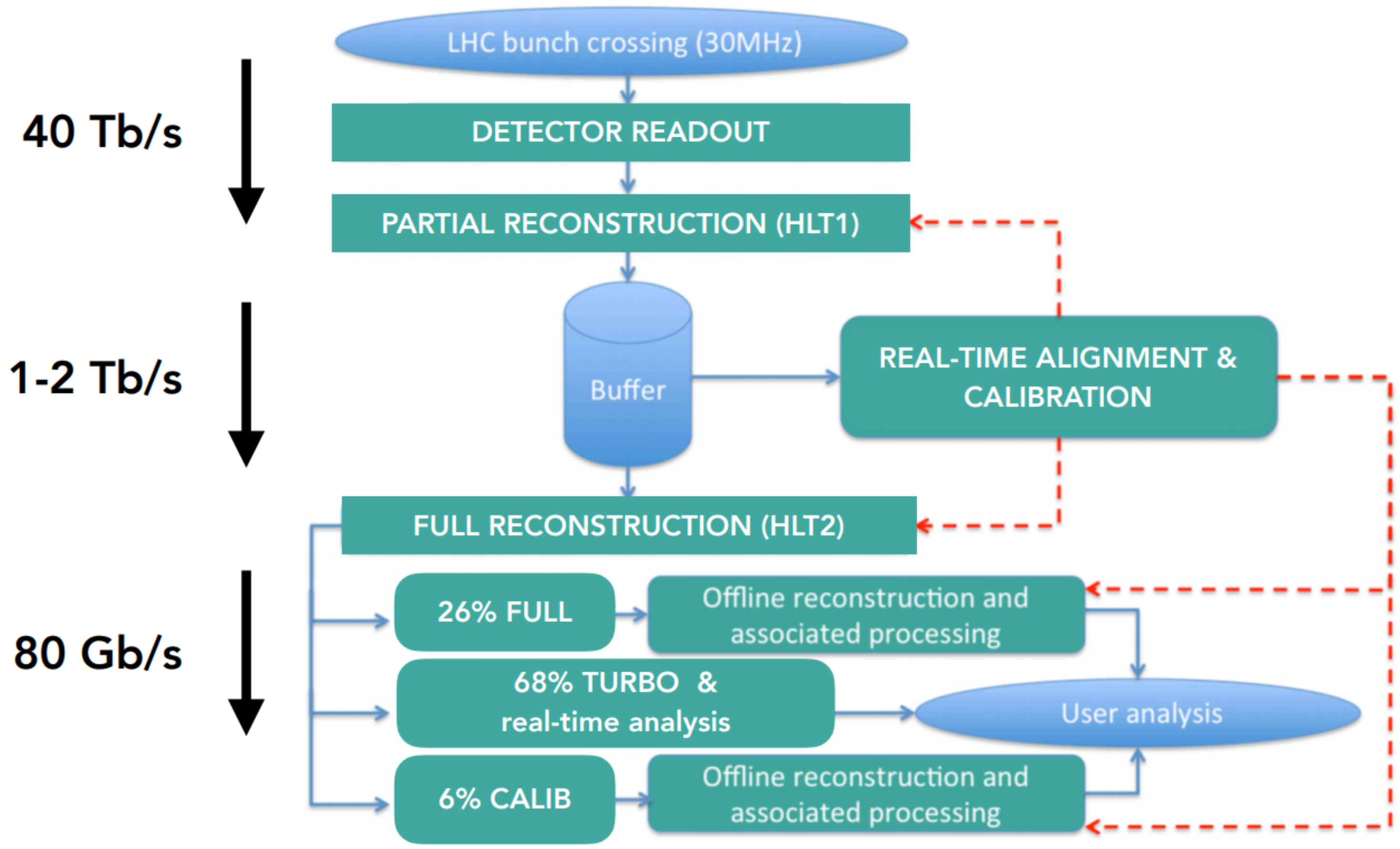






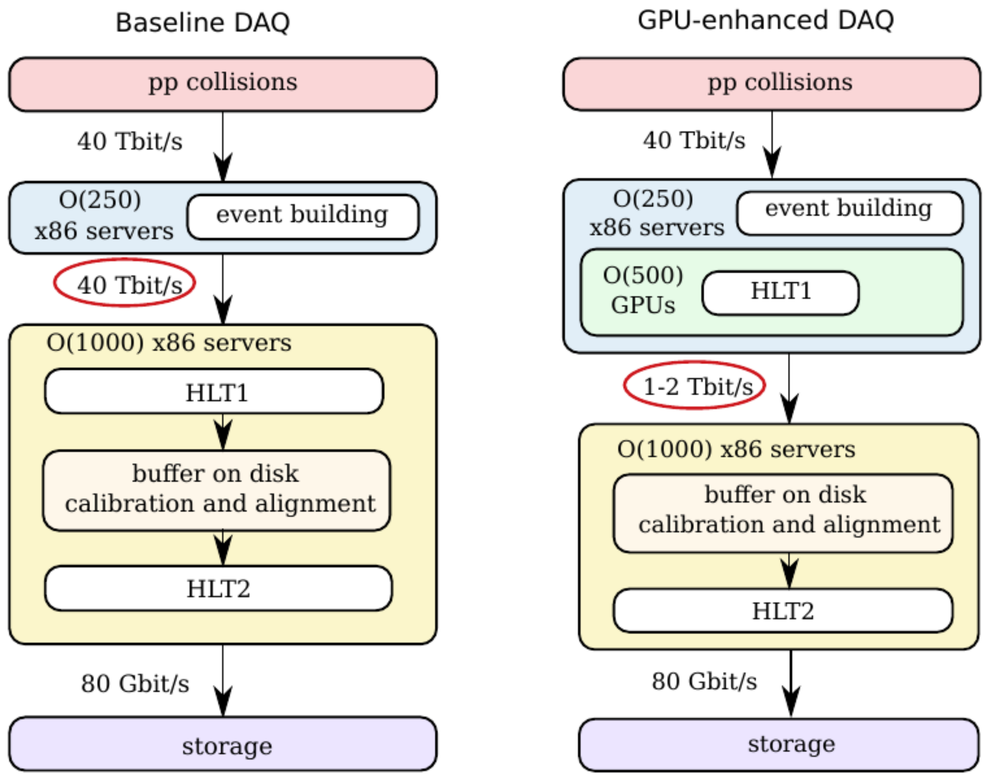
Idea of Real-Time Analysis

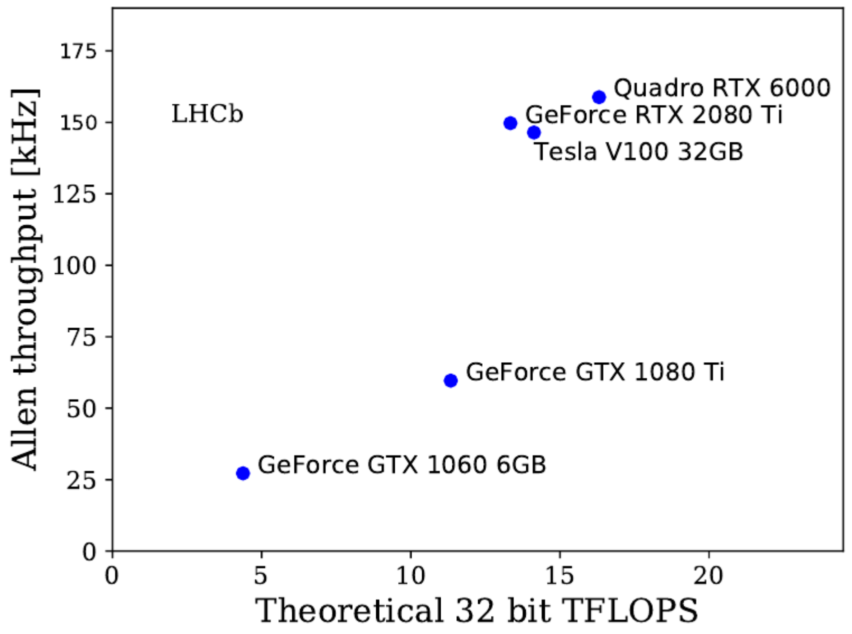
- Real Time Analysis – efficient decision about data in the full online mode
- Keeping only a signal and suppress any unnecessary information about event
- Continuous readout with full software trigger on 30 MHz (~ 40 Tb/s)



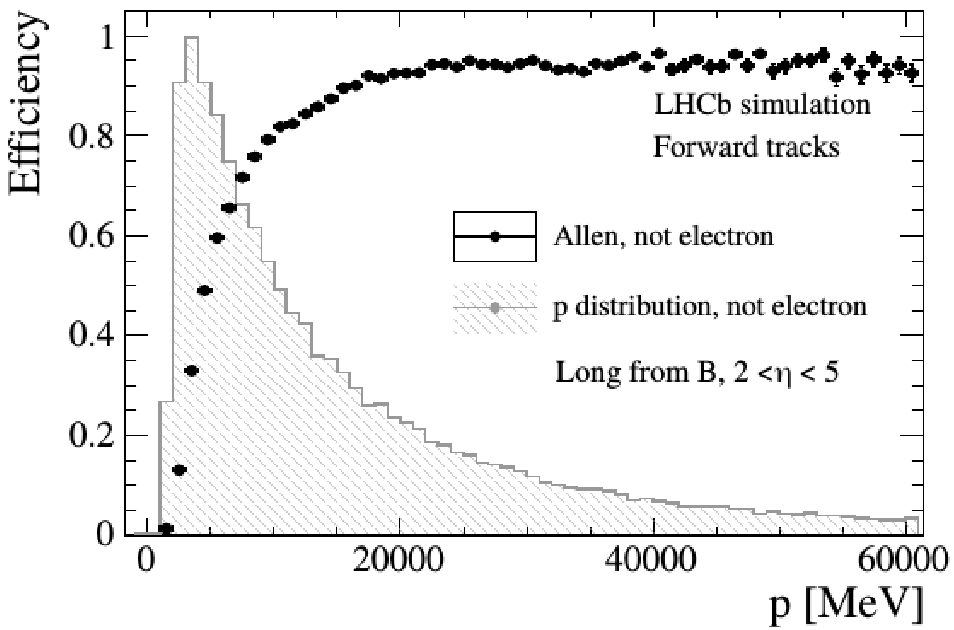
GPU-based HLT1

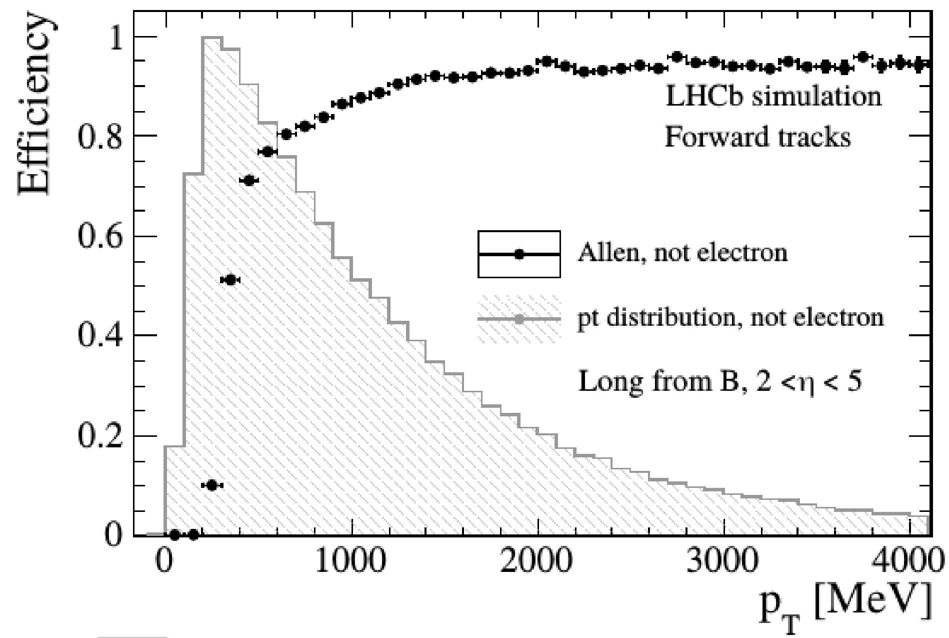
- HLT1 is by definition a parallel system with huge computation load
- Each raw event is relatively small (~ 100 kB)
- Highly parallel computation – a perfect match with a modern GPUs
- Usage of GPUs in HLT1 → The Allen project (Comput Softw Big Sci 4, 7 (2020), gitlab)
- LHCb solution will use around 500 GPUs
- First GPU-only HLT1 solution at hadron collider

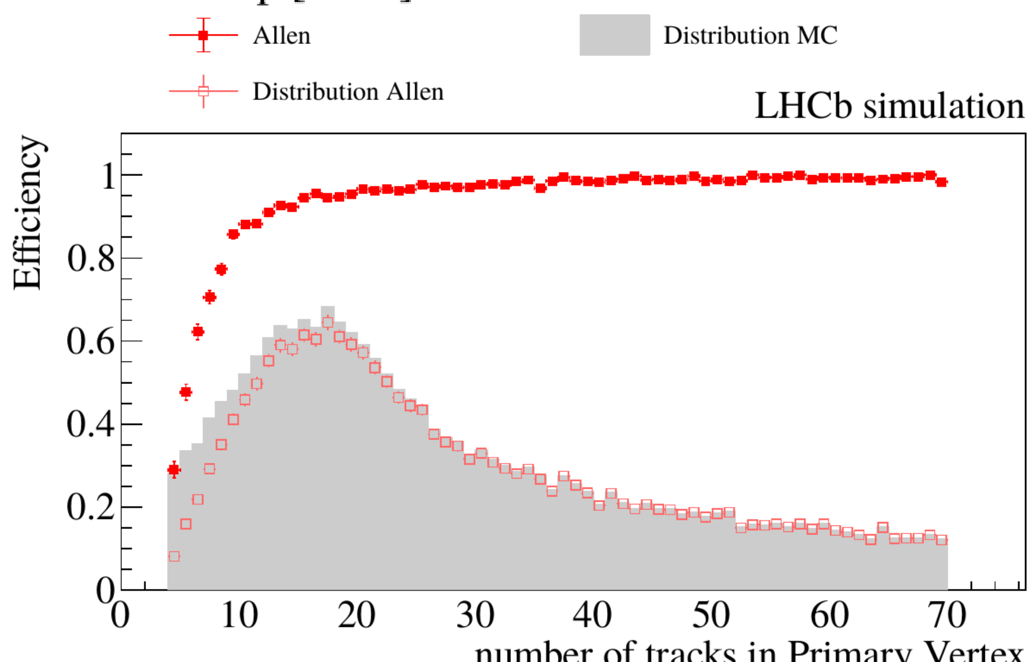




GPU-based HLT1

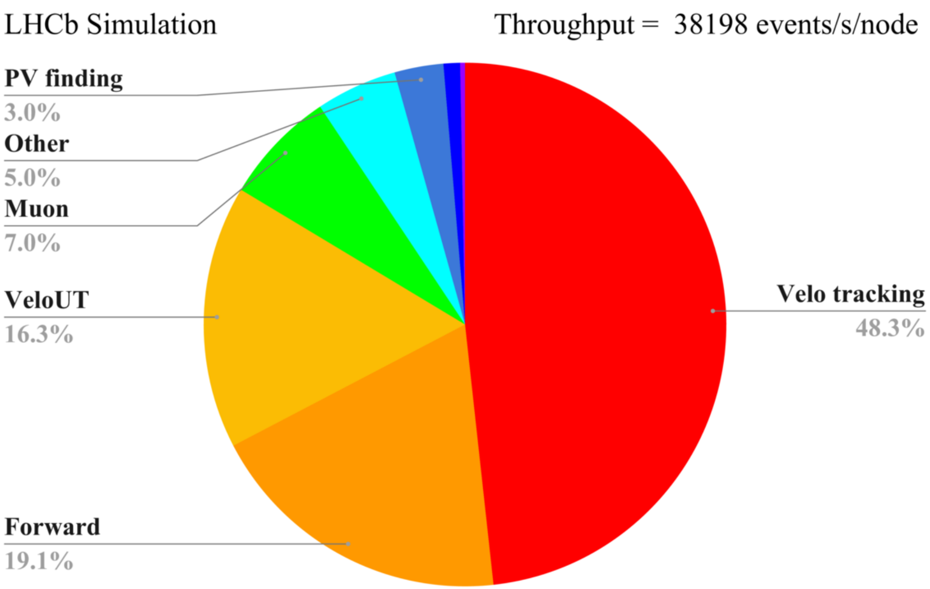


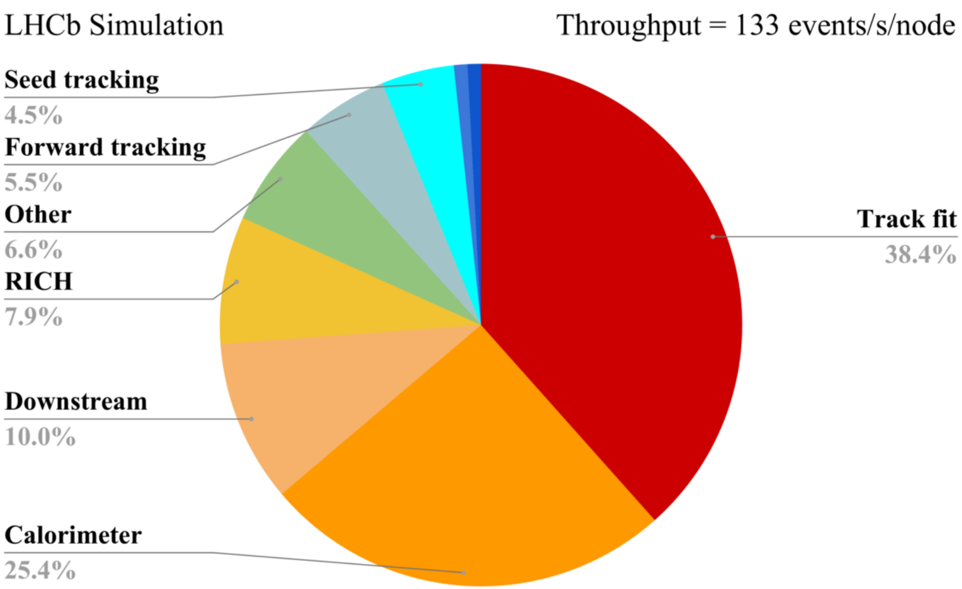




Breakdown of HLT1/2 throughput

- Status of the HLT1 (left) and HLT2 (right) reconstruction throughput as on 2020/04
- Breakdown based on old (CPU) reference system
- Nominal upgrade data taking conditions, automatic nightly test of throughput
- HLT2 is based on simulated minimum bias sample passing HLT1 selection

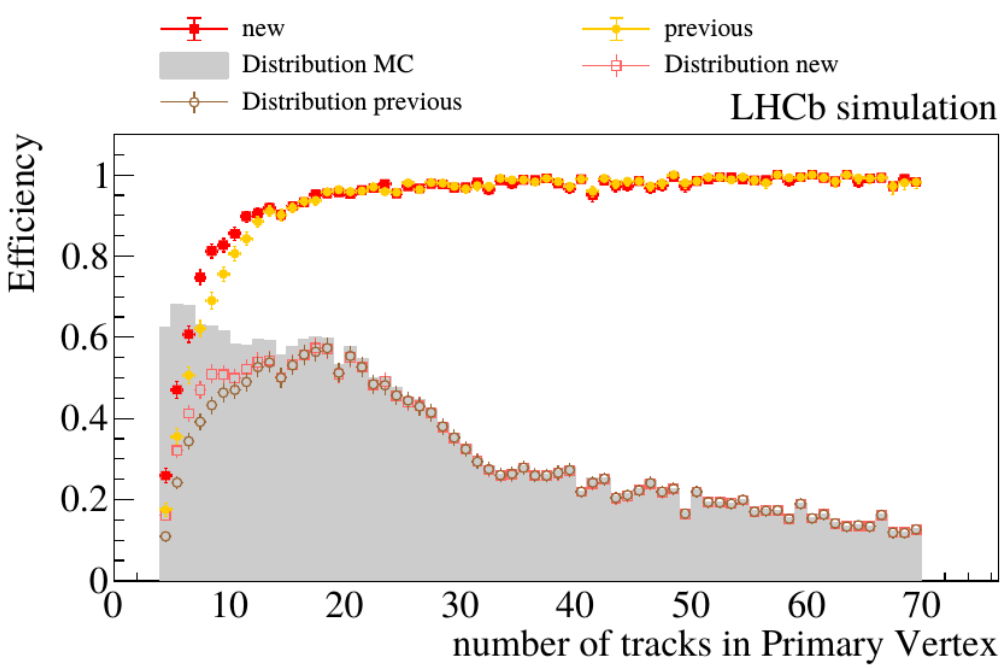


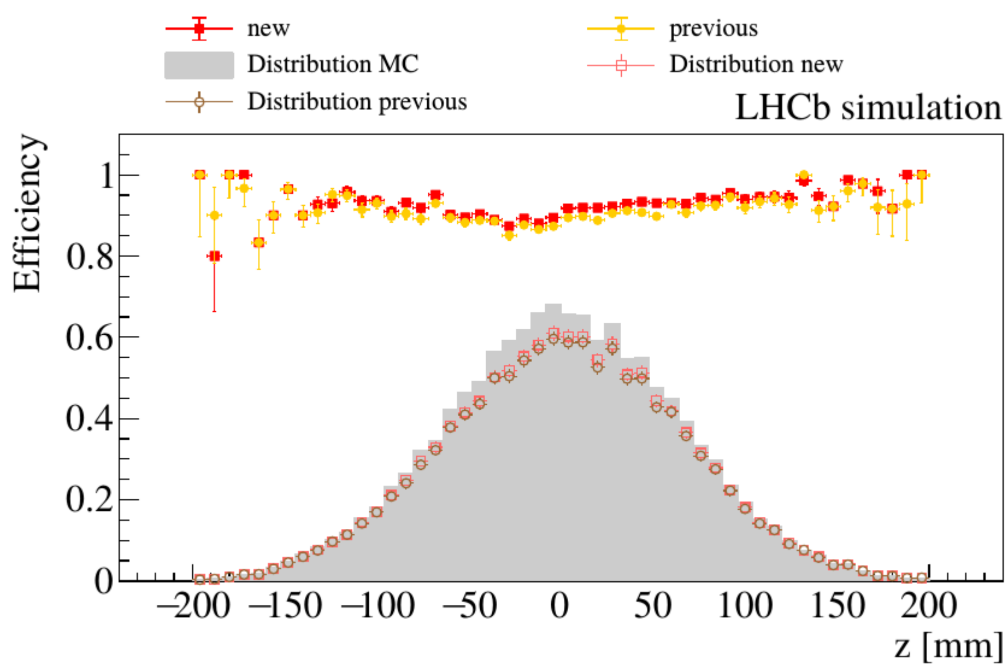


LHCb-FIGURE-2020-007

New PV reconstruction algorithm

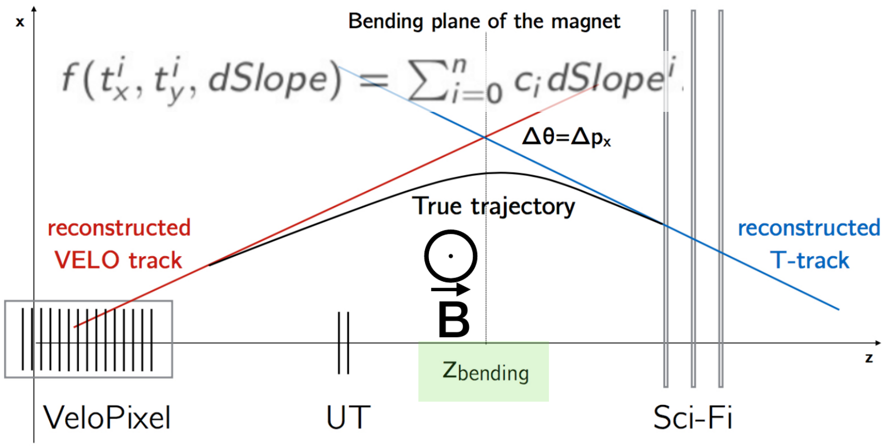
- Fast and precise reconstruction of Primary Vertex (PV) is the core of tracking system
- To speed up the HLT1 stage a new PV reconstruction algorithm was developed
- Approach is based on a peak search of z-position of VELO tracks at the point of their point of closet approach to the beamline – this is done without looking on x-y plane
- New algorithm is implemented in both baseline CPU and Allen based HLT1
- New reconstruction is fully compatible with RTA requirements and offers same or better efficiency when compared to Run2 model

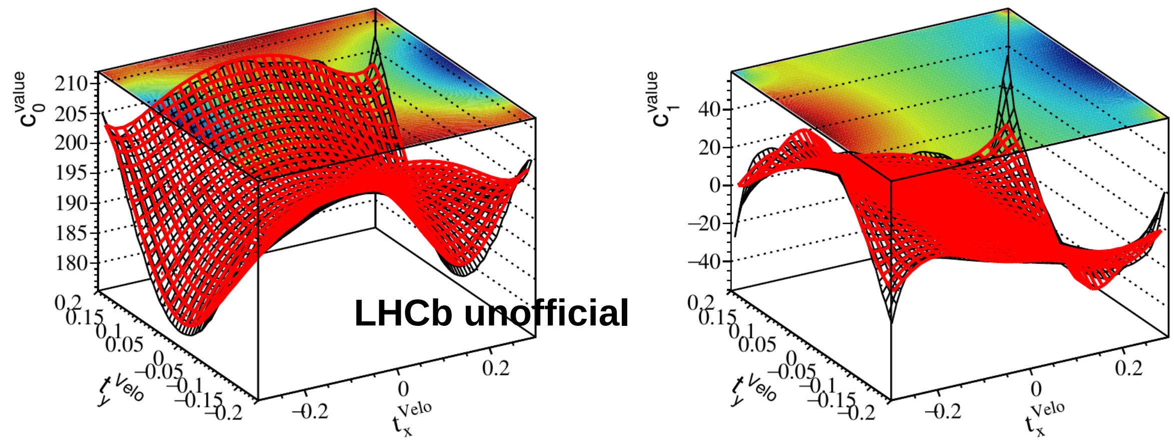




Allen contribution

- Momentum resolution in Allen is bellow the old CPU-based HLT1 resolution
- Improvement by using better parametrization of magnetic field and material interaction
- Example: evaluation of *q/p* from the reconstructed tracks
- Assumption: only B field effects



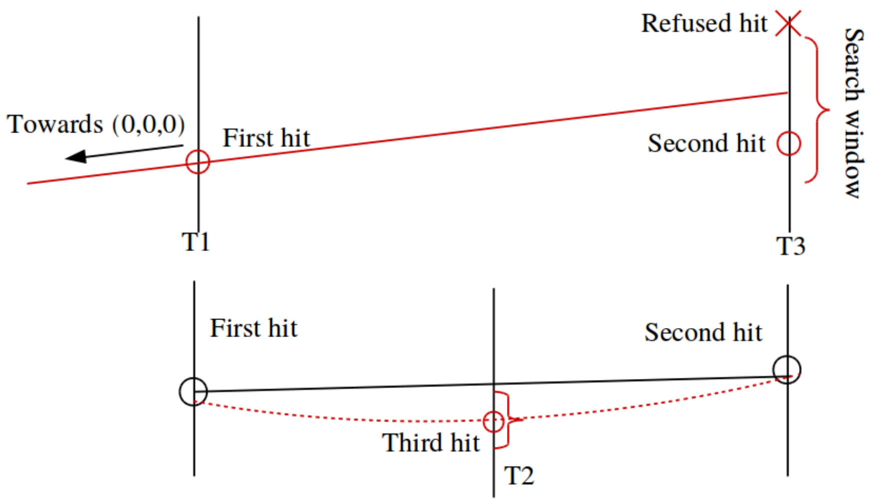


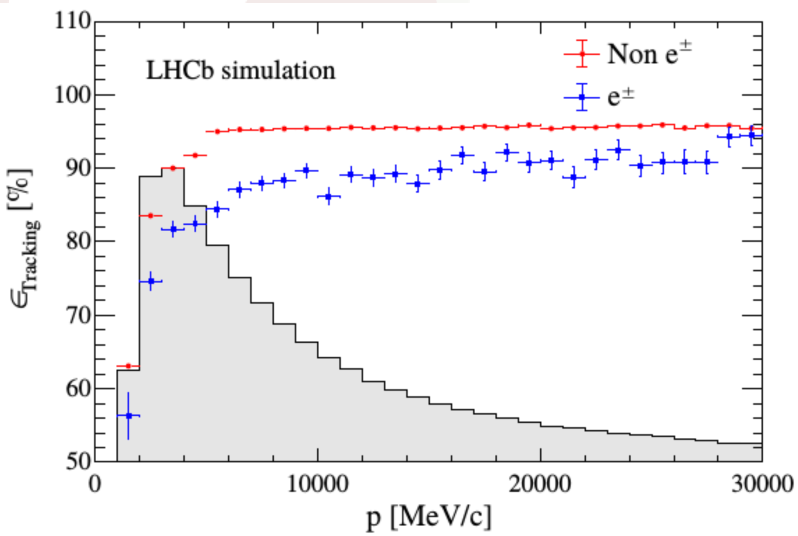
- Distribution *f* is fitted in 2D, resulting residual and injected back to pattern recognition
- Momentum resolution is improved with slight throughput loss ∼ 3 %
- This is to be pursued further with possibility to add effects from material interactions

- Additionally: SMOG studies and integration of Retina (FPGAs) project
- Presented studies done at Wuhan University by BIAN Lingzhu (边苓竹)

HybridSeeding for SciFi

- SciFi tracker is downstream of the magnet and designed for high occupancies
- Improved resolution of downstream tracks can allow more precise studies in channels with long-lived particles as KS or Λ
- HybridSeeding is new stand-alone tracking algorithm preparing track segments in SciFi
- Though designed mainly for Downstream tracks, HS can be used to form Long tracks
- Tracklets are first built in the non-bending plane in *x-z* plane
- The y-z patterns are built via relation: $x^{u,v}_{measured} - x_z = y_z \times \tan(\alpha)$
- Contribution from THU group by WANG Mengzhen (王梦臻)

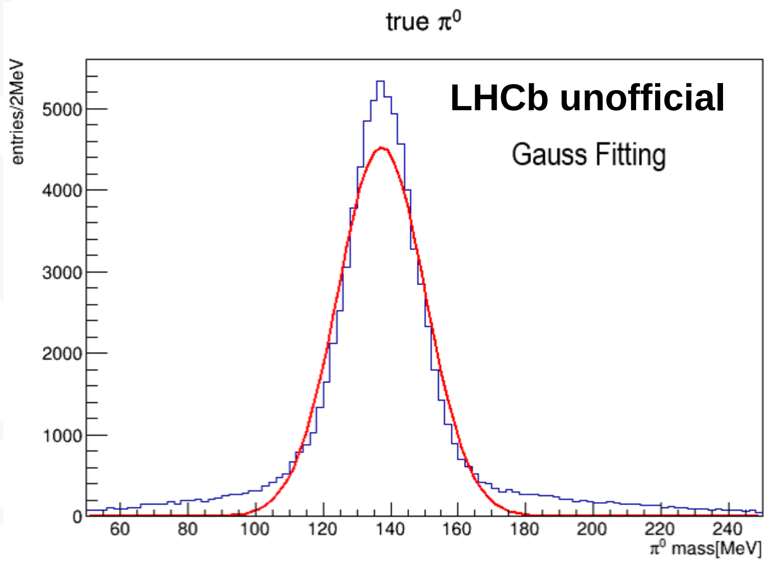


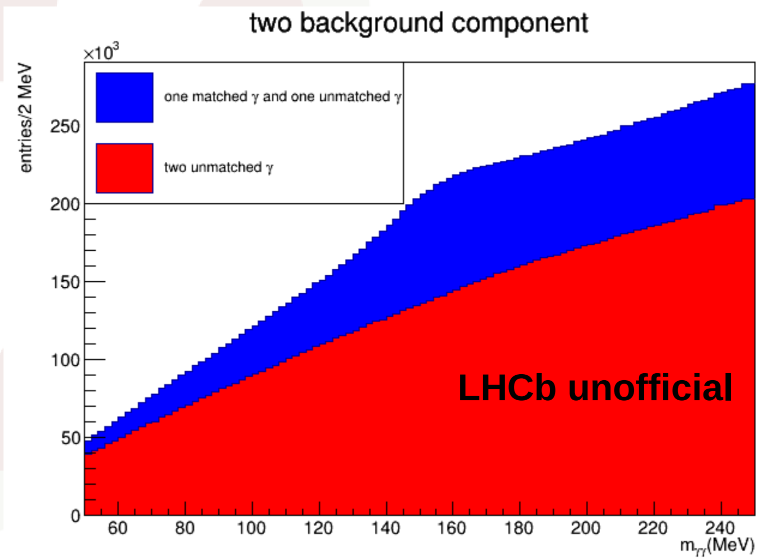


2007.02591

ECAL calibration

- ECAL calibration using $\pi^0 \to \gamma\gamma$ channel
- Well-calibrated ECAL is crucial for many studies
- At Run2 signal was fitted using Gaussian and 2nd order polynomial fit
- Run3 conditions require more complex approach
  - Long tails
  - Significant background
- Work at CCNU focus on creating of correct model for Run3 by GAO Yang (高扬) and ZHANG Dongliang (张东亮)
- ECAL studies relevant to neutral PID and Upgrade2 are ongoing at PKU by XU Zehua (许泽华) and THU by MU Hongjie (牟宏杰)





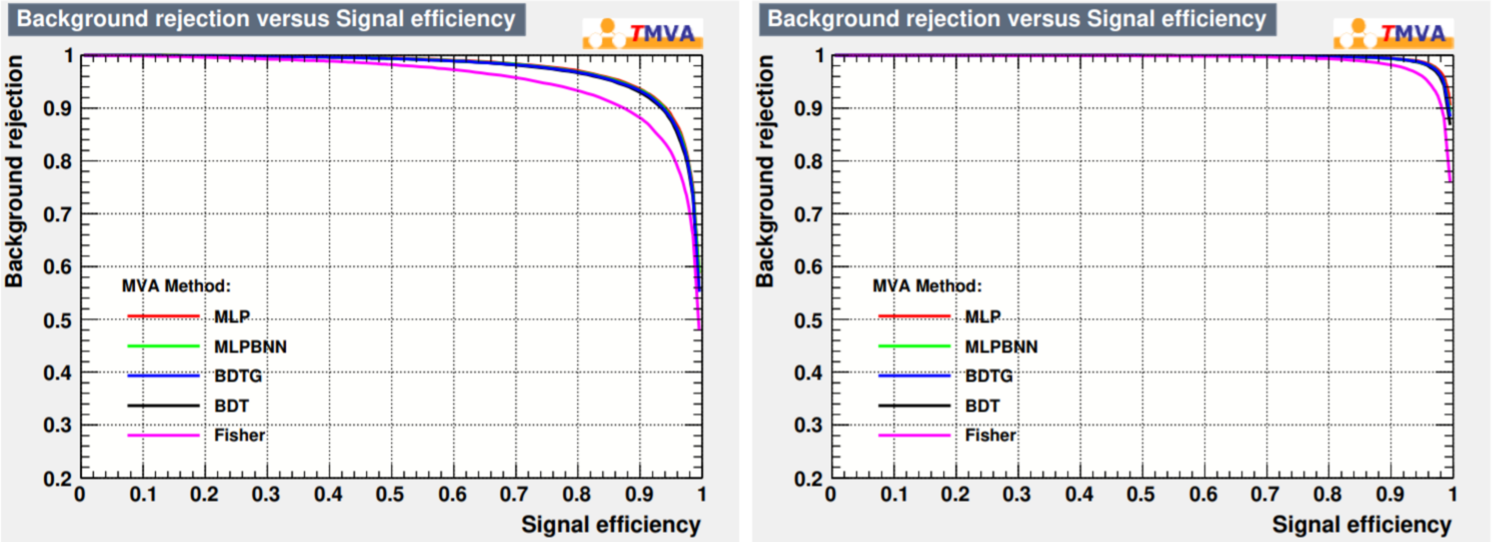
HLT2 lines

- HLT2 lines define exact processes which are to be reconstructed and recorded
- Turbo model will be a default model used for physics analyses in Run 3
- Changes due to new trigger system must be accordingly propagated to HLT2 lines
- Main objects:
  - To define which objects are to be persisted
  - Collecting and serialising those objects in the required format when writing, and deserialising when reading.
- All of this must be done within the strict time constrains (speed), with high efficiency and robustness
- One of the main difference between Run2 and Run3 configuration is CPU cost of persistency
- At Run2 CPU cost did not need much of optimizations it may be become an issue at Run 3
- Studies are ongoing at PKU by XU Ao (许傲) and at THU by MU Hongjie (牟宏杰)

- HLT2 lines efficiency can be further affected by the ghost tracks, which must be kept on the lowest level
- Study is ongoing at CCNU by XU Menglin (徐梦琳)

HLT2 lines

- All the HLT2 lines must be ported from Run2 to Run3 model
- Ideally, lines should be fast and efficient (S/B ratio) to allow high HLT2 throughput
- In many cases this may be achieved by TMVA-based HLT2 lines
- An example from B2OC group using TMVA implementation of BDT:
  - Generic BDT for lines fired by hadron candidate
  - Specific BDT trained for Dh candidates
- BDT is trained using various Run3 specific MC samples
- Results show large efficiency > 90 %, allowing about 10 % reduce of recorded B mesons
- Contribution from PKU by ZHANG Shunan (张舒楠) and at THU by MU Hongjie (牟宏杰)



Rapid response test framework

- Full RTA system must be thoroughly tested before real running
- As the most of the data will be using Turbo lines, mistake in any part of data processing can be fatal
- Test conditions must be as close as possible to the expected situation during the real data-taking
- Process is iterative procedure following the progress of RTA project and Upgrade itself
- Goal is to prepare an integrated rapid response test framework for monitoring performance of high level quantities
- Dashboard monitoring physics performance of new reconstruction system is ready
  - Performance is evaluated every day as part of software nightly testing
- Experts responsible for specific tasks can then check effect of their development

- Additionally, this framework may help with Run3 early measurements preparation and give proponents the opportunity to test full analysis chain before real data available

- Contribution from UCAS by SAUR Miroslav

Conclusion

- Real Time analysis at LHCb is a novel approach for hadron collider experiments enable substantially increase the amount of recorded data
- Such a new approach is crucial for a future collider experiments on HL-LHC, SppC/FCC-hh, ...
- RTA development is ongoing and preparing for the Run3 data-taking

- Chinese groups are involved in many different aspects of RTA
  - Only selected contributions shown due to time constrains
- All these contributions are important for successful Run3 and onwards

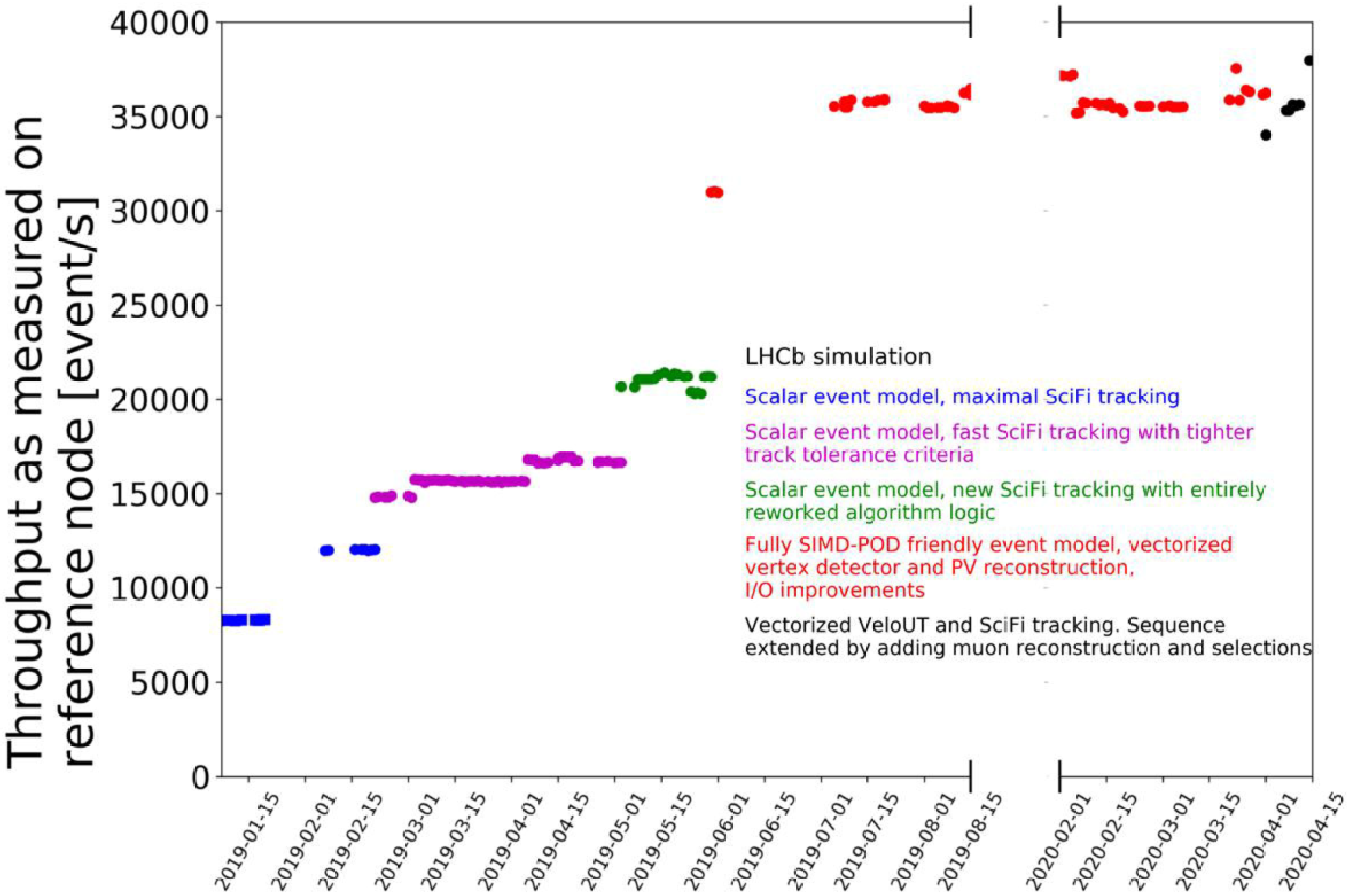
Thank you for your attention

谢谢大家

BACKUP Slides

HLT1 on 30 MHz

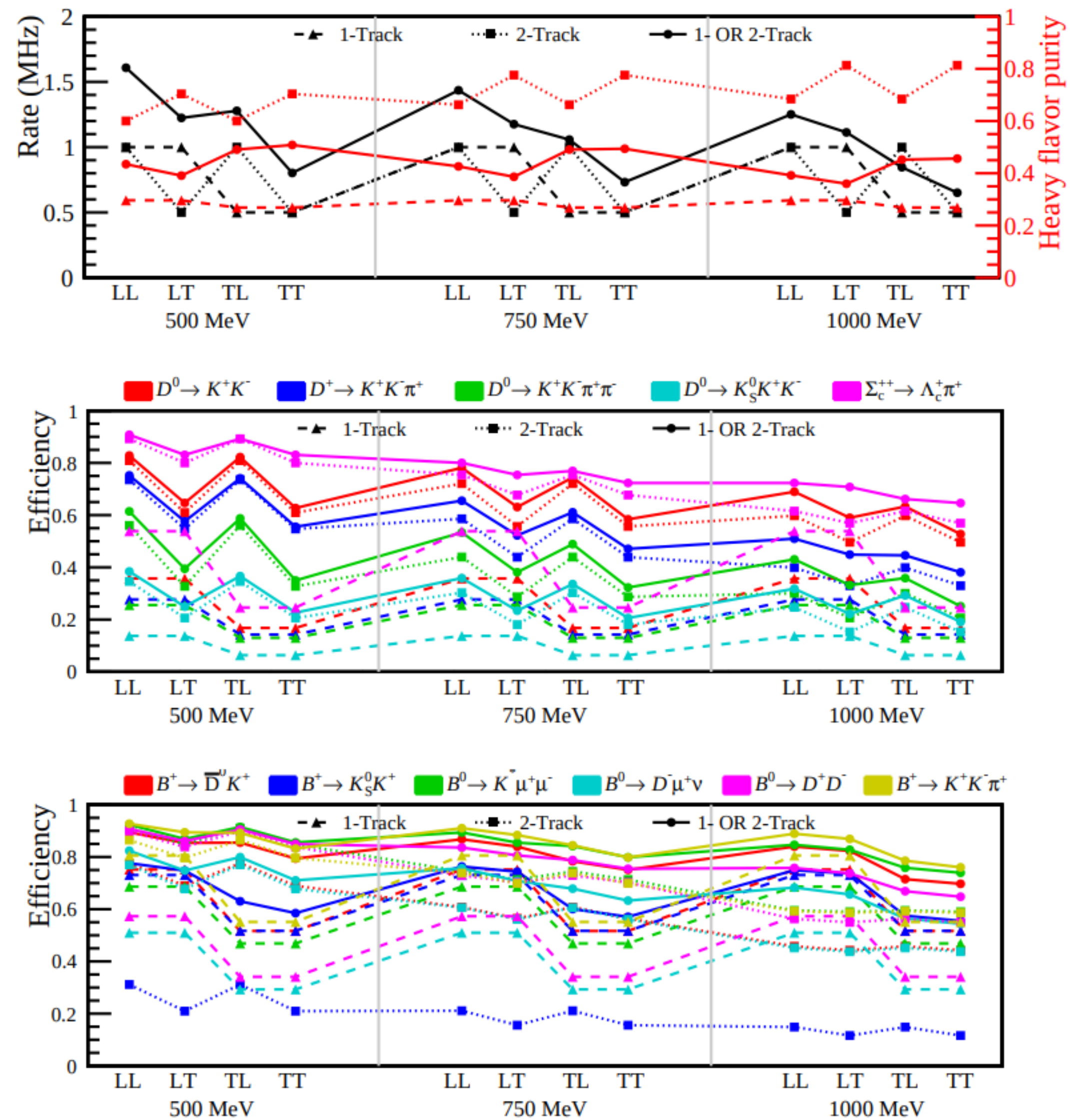
- HLT1 throughput evolution between 2019/01 and 2020/04
- Still using old (CPU) reference system



LHCb-FIGURE-2020-007

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

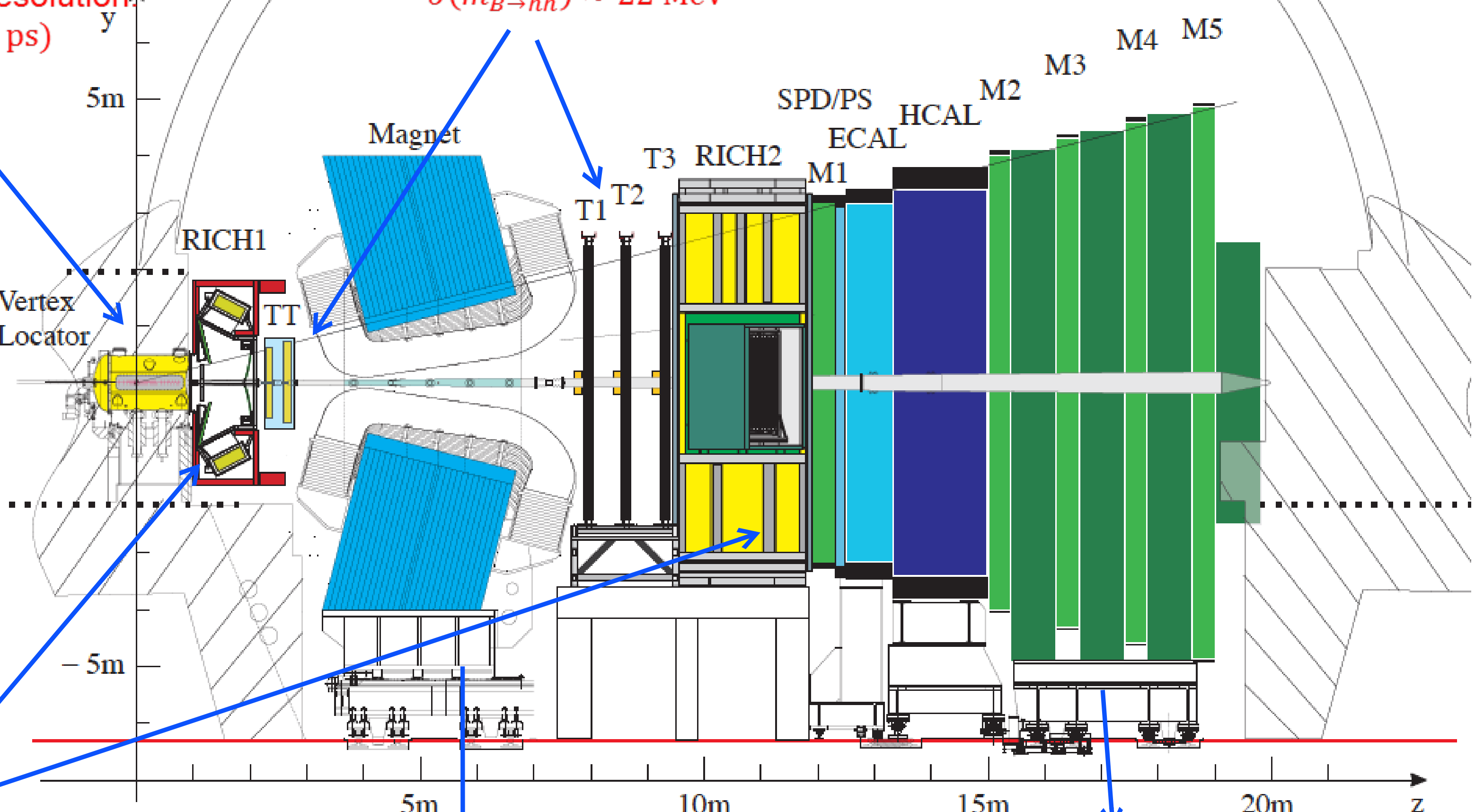
LHCb detector 2010-2018

- 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)

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

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



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

Magnet
- Bending power: 4 Tm

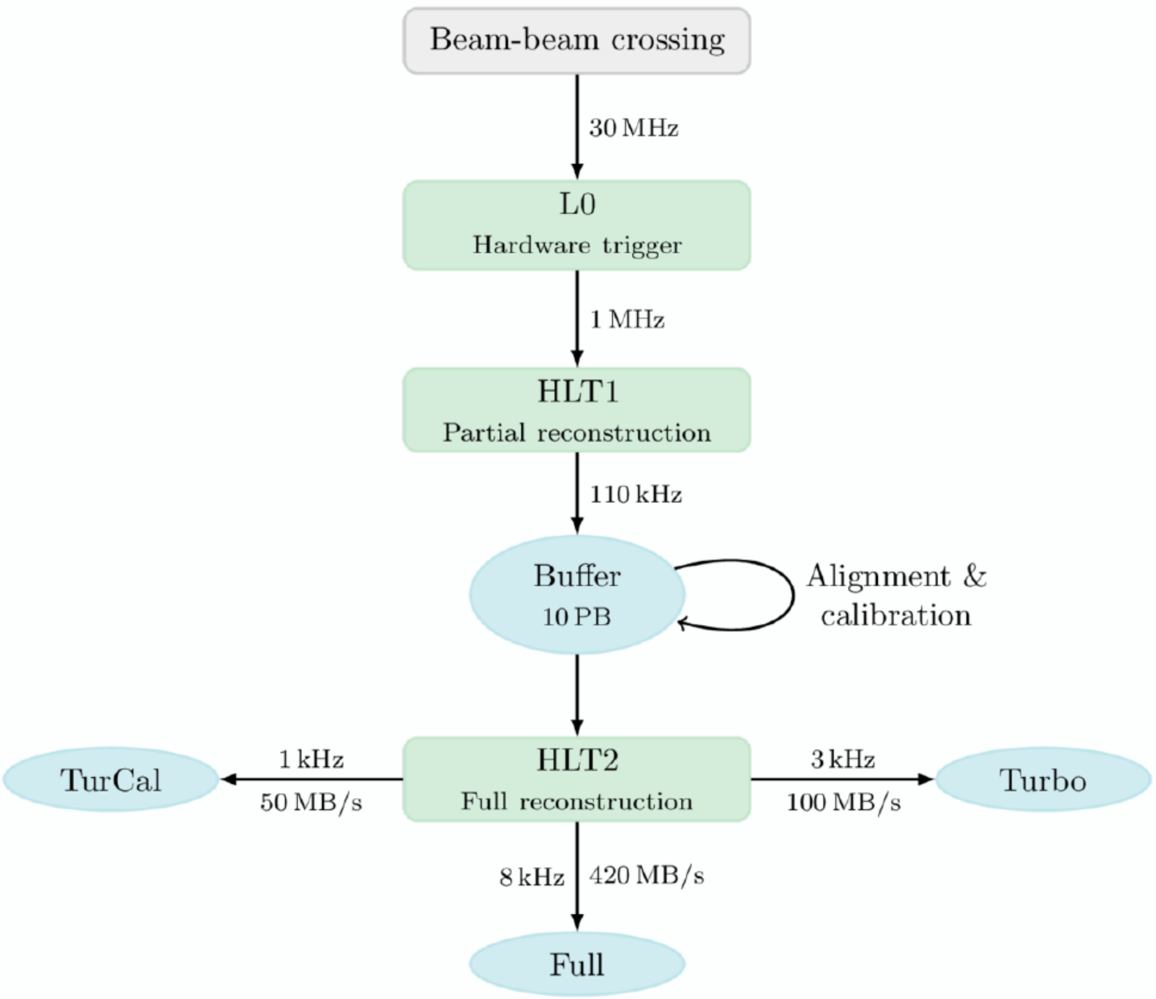
Muon system
- μ ID: $\varepsilon(\mu \to \mu)$ ~97%
- Mis-ID: $\varepsilon(\pi \to \mu)$ ~1-3%

Trigger during Run II

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

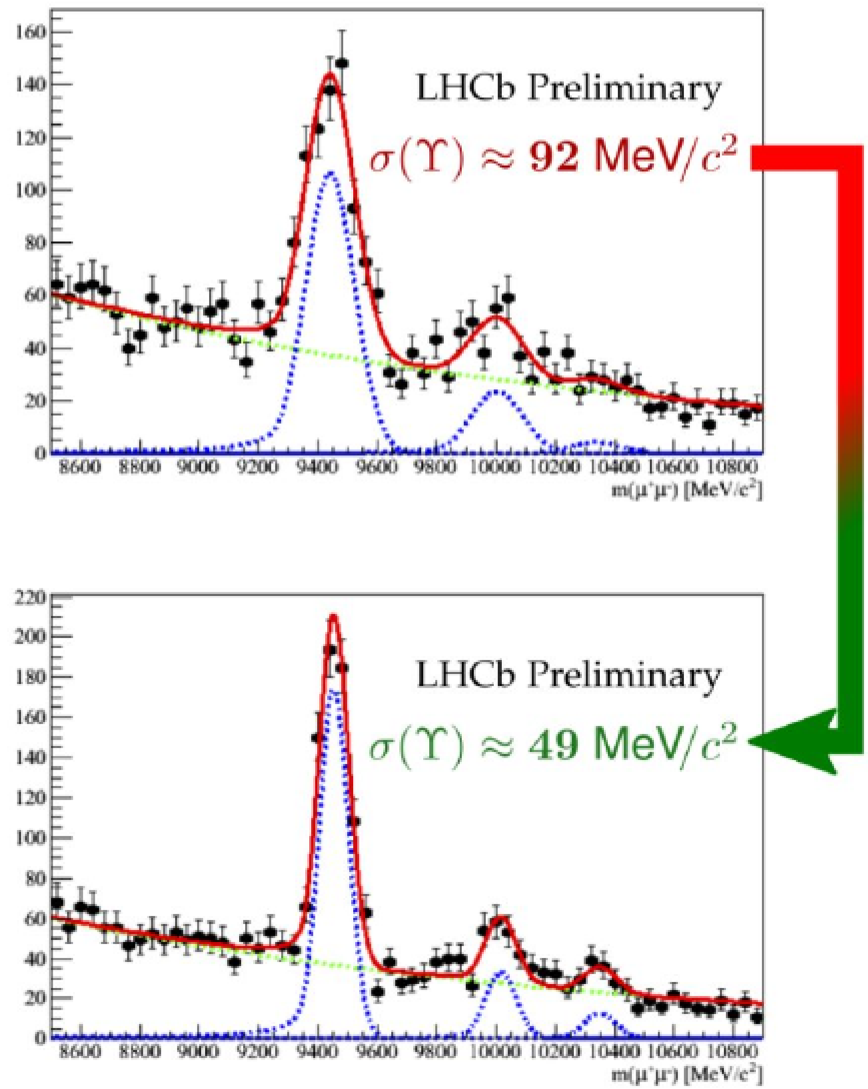
1) L0 Hardware trigger
   - fast detectors
   - CALO and MUON information
2) High Level Trigger 1
   - 10 PB disk buffer
   - Around two weeks of data taking
3) High Level Trigger 2

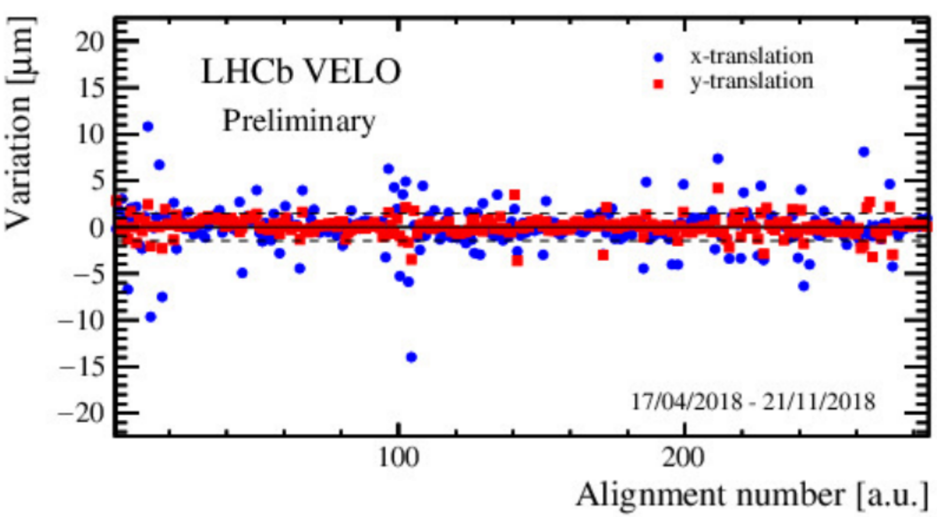
- Around 500 specific HLT lines (particular decays)
- Introduction of TESLA framework → Turbo stream



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



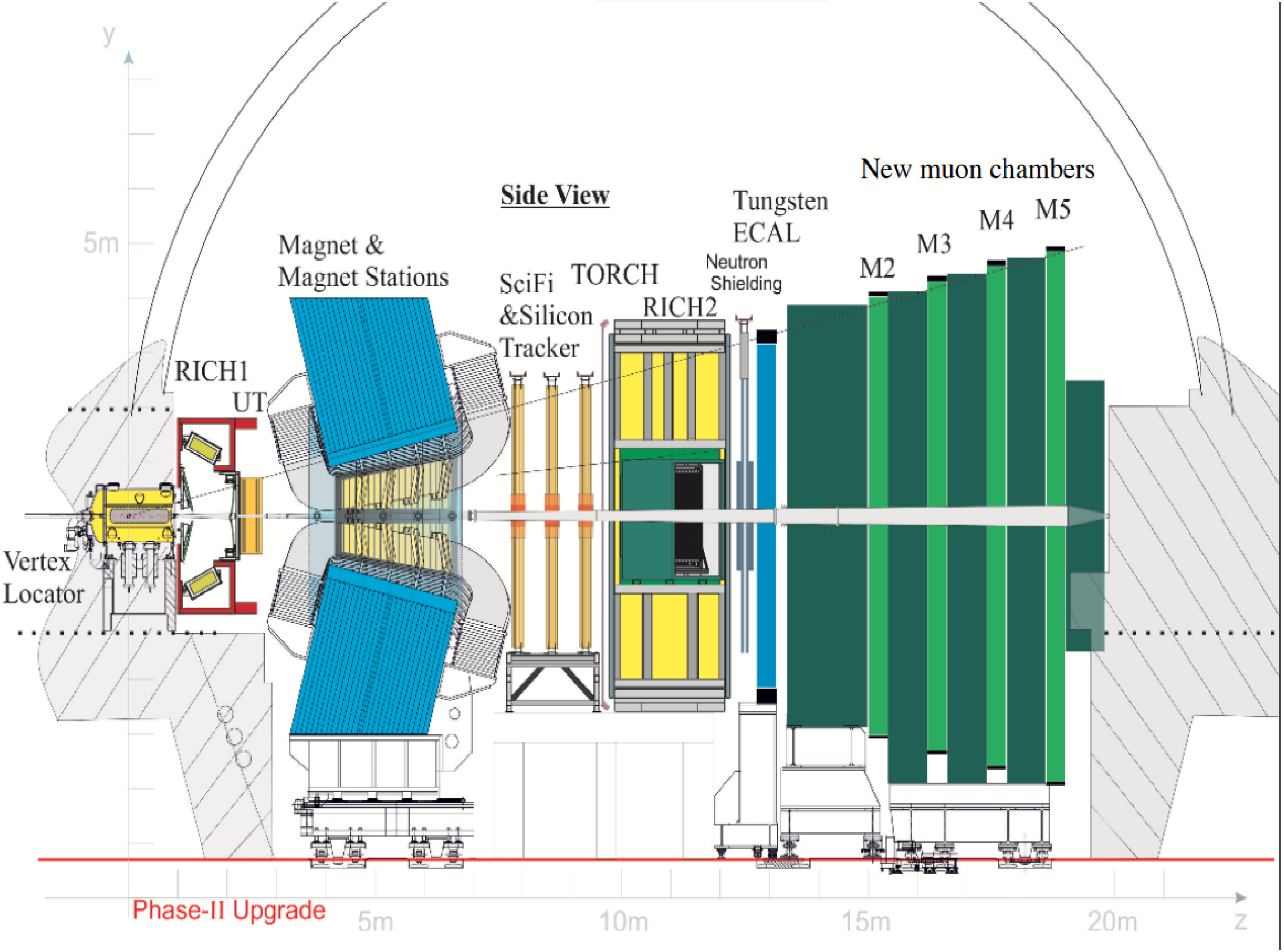


LHCb Upgrade I - Physics

Type	Observable	Current precision	LHCb 2018	**Upgrade** ($50\,\mathrm{fb}^{-1}$)	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
	$A_{\mathrm{fs}}(B_s^0)$	6.4×10^{-3} [18]	0.6×10^{-3}	0.2×10^{-3}	0.03×10^{-3}
Gluonic penguin	$2\beta_s^{\mathrm{eff}}(B_s^0 \to \phi\phi)$	–	0.17	0.03	0.02
	$2\beta_s^{\mathrm{eff}}(B_s^0 \to K^{*0}\bar{K}^{*0})$	–	0.13	0.02	< 0.02
	$2\beta^{\mathrm{eff}}(B^0 \to \phi K_S^0)$	0.17 [18]	0.30	0.05	0.02
Right-handed currents	$2\beta_s^{\mathrm{eff}}(B_s^0 \to \phi\gamma)$	–	0.09	0.02	< 0.01
	$\tau^{\mathrm{eff}}(B_s^0 \to \phi\gamma)/\tau_{B_s^0}$	–	5 %	1 %	0.2 %
Electroweak penguin	$S_3(B^0 \to K^{*0}\mu^+\mu^-; 1 < q^2 < 6\,\mathrm{GeV}^2/c^4)$	0.08 [14]	0.025	0.008	0.02
	$s_0\,A_{\mathrm{FB}}(B^0 \to K^{*0}\mu^+\mu^-)$	25 % [14]	6 %	2 %	7 %
	$A_{\mathrm{I}}(K\mu^+\mu^-; 1 < q^2 < 6\,\mathrm{GeV}^2/c^4)$	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	γ ($B \to D^{(*)}K^{(*)}$)	~ 10–$12°$ [19, 20]	4°	0.9°	negligible
	γ ($B_s^0 \to D_s K$)	–	11°	2.0°	negligible
	β ($B^0 \to J/\psi\,K_S^0$)	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

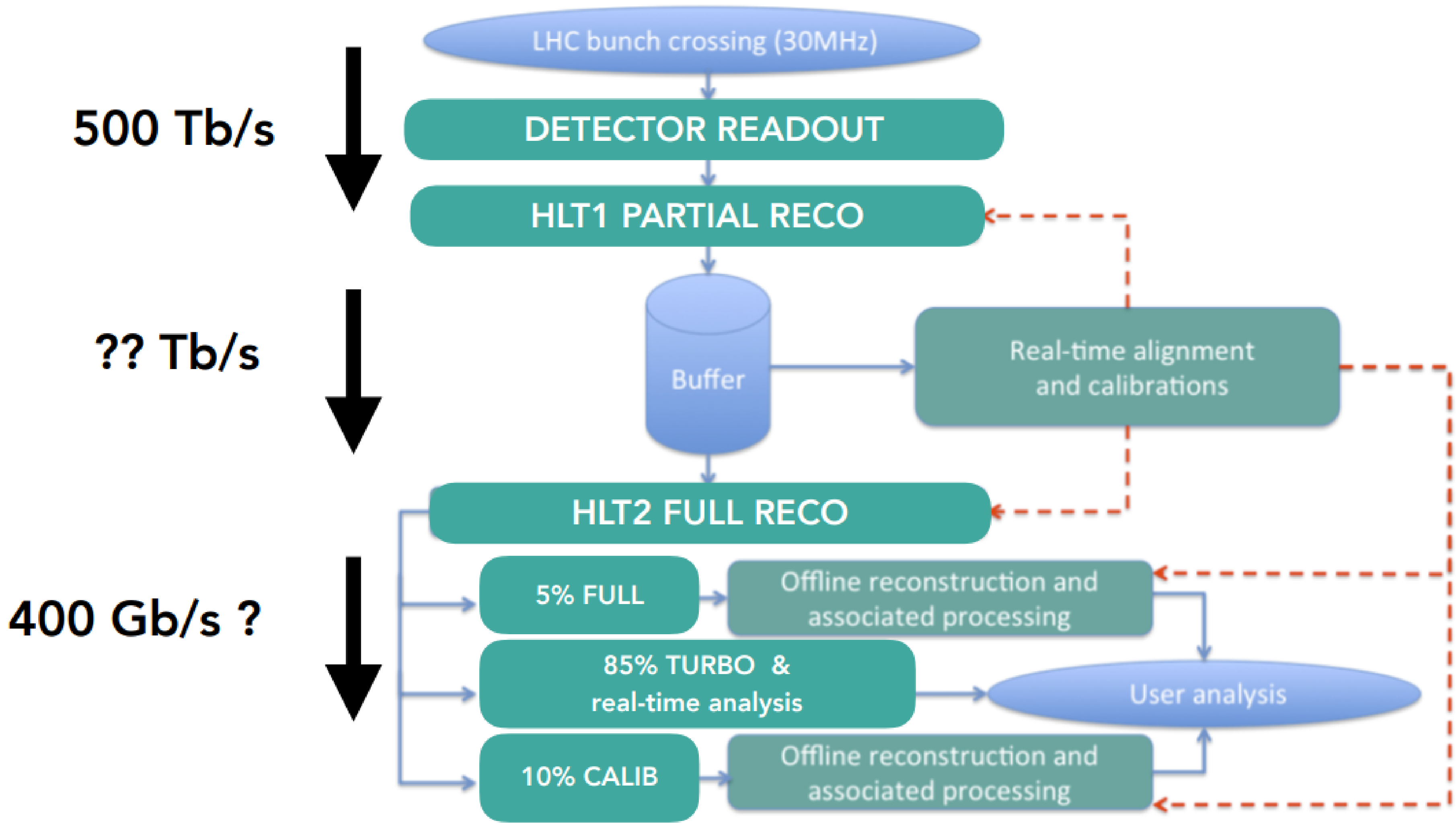
LHCb upgrade Phase II (Run V)



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 $\tau_{B_s^0 \to \mu^+\mu^-} \approx 0.03\,\mathrm{ps}$	
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
ϕ_s with $B_s^0 \to J/\psi K^+K^-$, $J/\psi\pi^+\pi^-$	Uncertainty on $\phi_s \approx 3\,\mathrm{mrad}$	in CP observables from Phase-II
$\phi_s^{s\bar{s}s}$ with $B_s^0 \to \phi\phi$	Uncertainty on $\phi_s^{s\bar{s}s} \approx 8\,\mathrm{mrad}$	ECAL and low-p tracking.
$\Delta\Gamma_d/\Gamma_d$	Uncertainty on $\Delta\Gamma_d/\Gamma_d \sim 10^{-3}$	Approach SM value.
Semileptonic asymmetries $a_{\mathrm{sl}}^{d,s}$	Uncertainties on $a_{\mathrm{sl}}^{d,s} \sim 10^{-4}$	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.