The LHCb Trigger

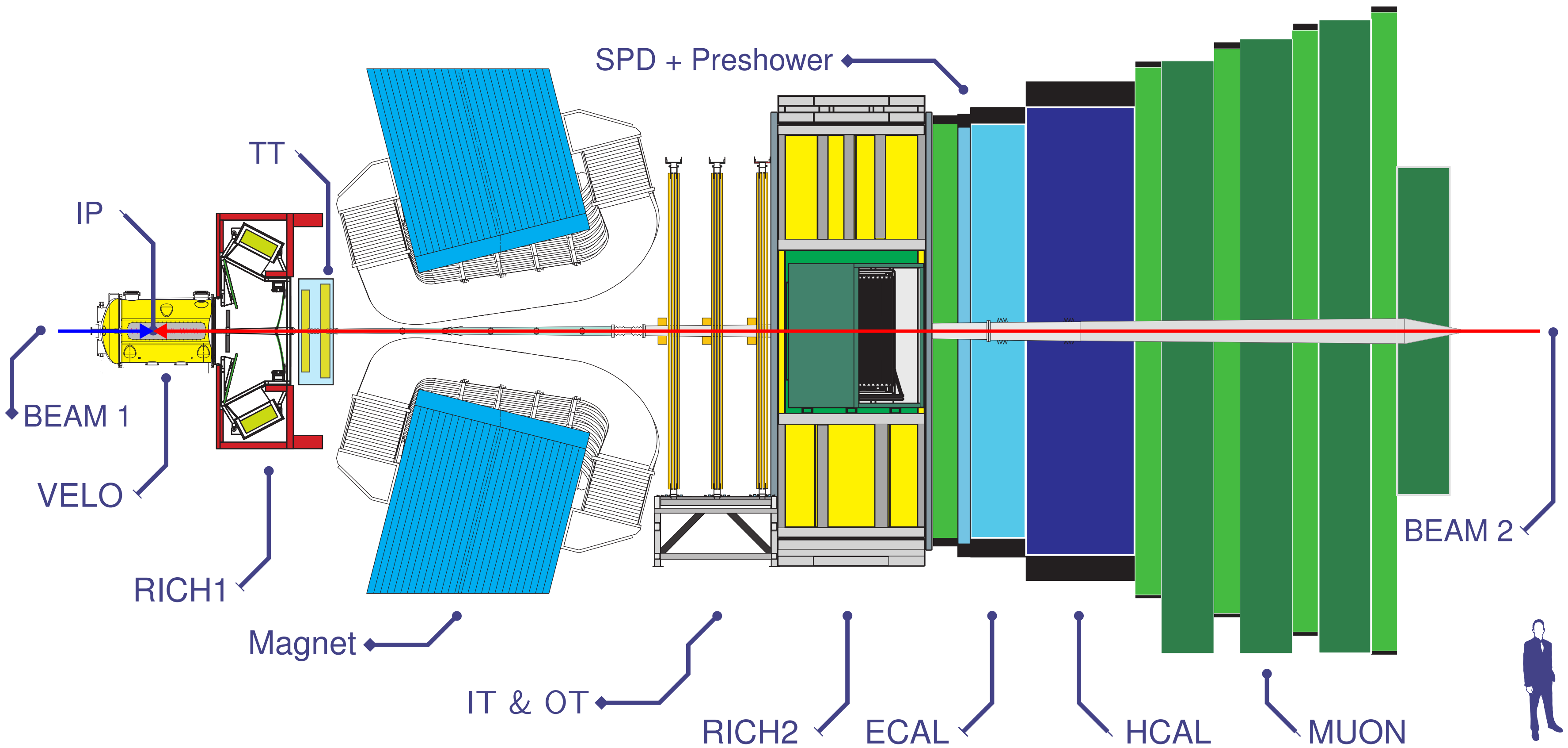
and its upgrade

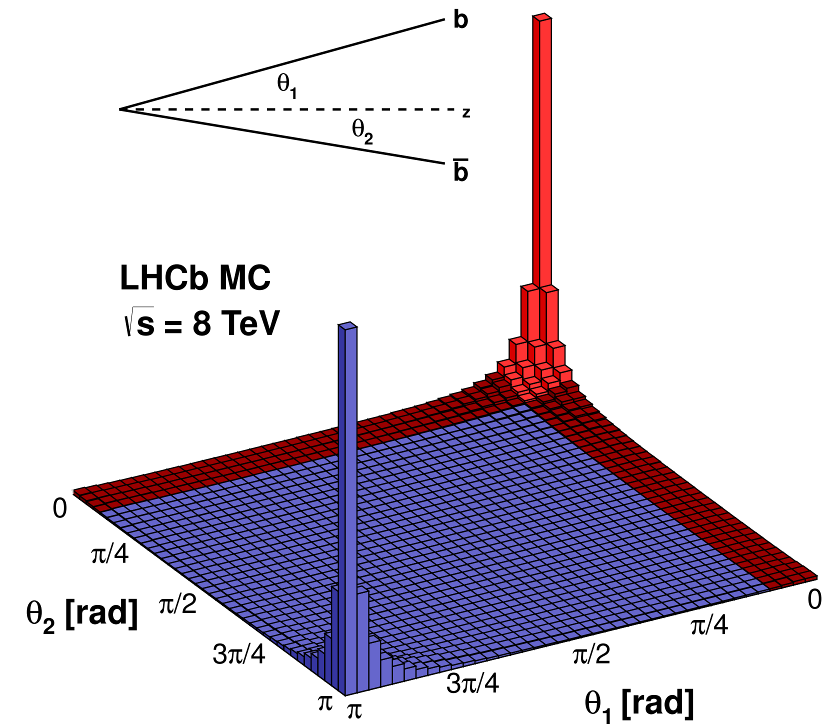
Past Performance:	Run 1 : - 2012
The Immediate Future:	Run 2 : 2015 - 2018
The Upgrade:	Run 3 : 2020 -

ACAT 2014
– bridging disciplines
1-5 September 2014
Faculty of Civil Engineering
Europe/Prague timezone

Gerhard Raven
VU VRIJE UNIVERSITEIT AMSTERDAM NIKHEF LHCb
On behalf of the LHCb HLT Project

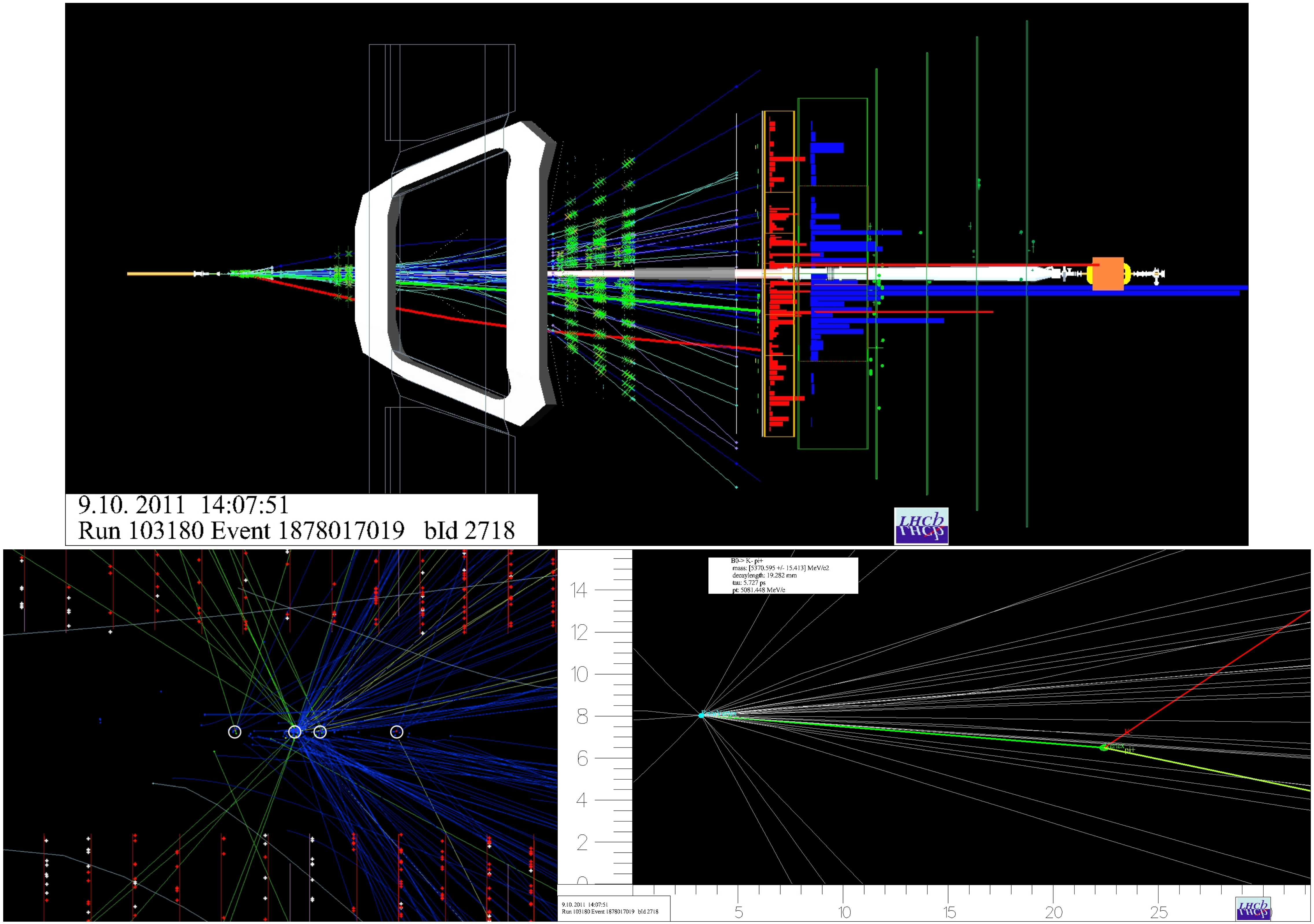
The LHCb Experiment





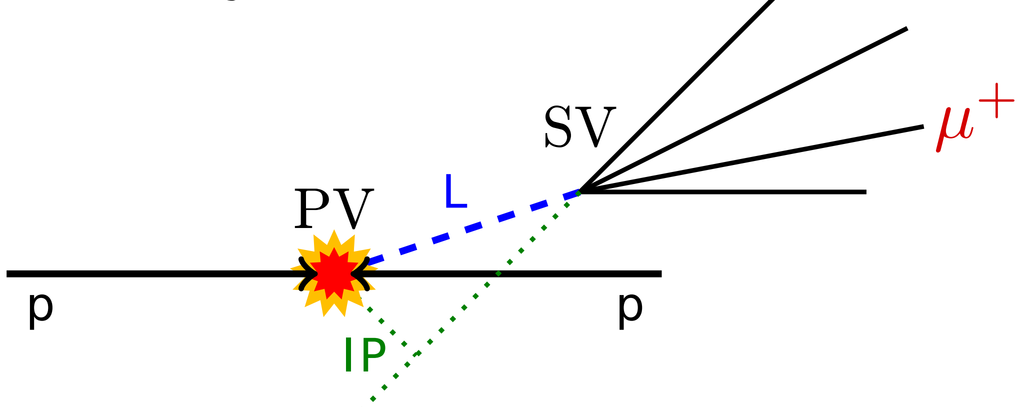
- Single-arm spectrometer ($2 < \eta < 5$) @ LHC
- Run 1 (2010-2012): $L = 4 \cdot 10^{32} cm^{-2} s^{-1}$ (2x design),
  - ~1.6 visible interactions/crossing
  - ⇒ 30 kHz $b\bar{b}$, 600 kHz $c\bar{c}$ within the acceptance
- Precision flavour physics: CP violation, rare decays, …

An LHCb Event



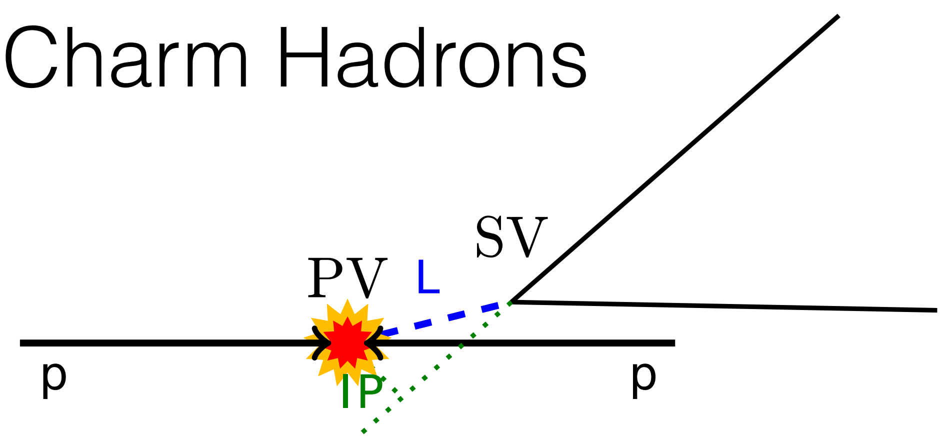
Typical Signatures

Beauty Hadrons



- B mass ~5.28 GeV/c^2, daughter P_T ~1 GeV/c
- τ_B~1.6 ps, flight distance ~ 1 cm
- Important signature: detached muons from from eg. B→J/ψX, J/ψ→$\mu^+\mu^-$

Charm Hadrons



- D^0 mass ~1.86 GeV/c^2, appreciable daughter P_T
- τ_D~0.4 ps, flight distance ~ 0.4 cm
- Also produced as 'secondary' charm from B decays.

Trigger Strategy:

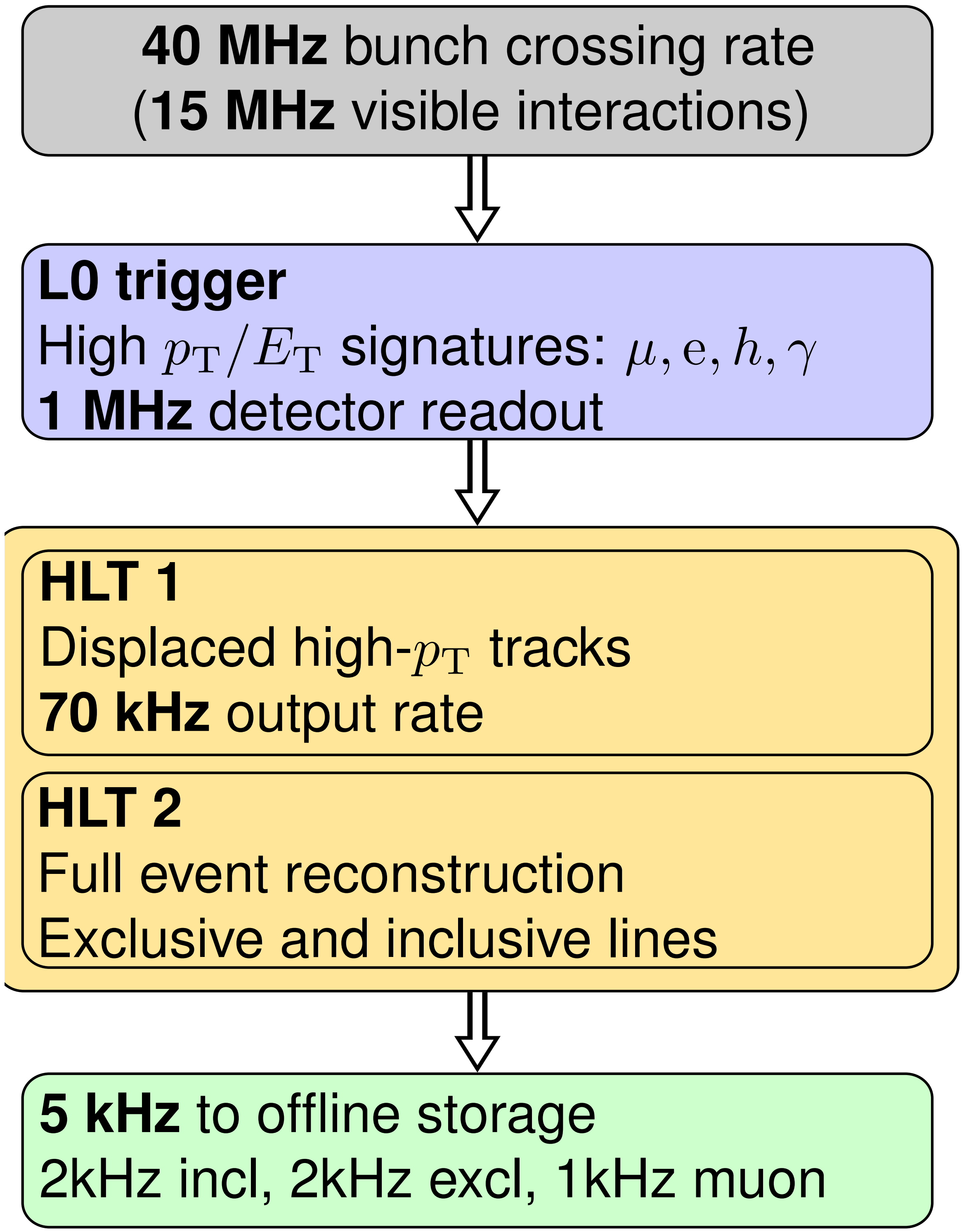
- Inclusive triggers on muons, displaced vertices with high PT tracks
- Exclusive triggers for anything else

2011-2012 trigger

Three stage architecture:

- Level 0 (L0) — near detector hardware, fixed 4 μs latency
- Higher Level Trigger (HLT1 and HLT2) — software running on 29,000 cores

JINST 8 (2013) P04022, and arXiv:1310.8544



L0 trigger: Muons

Reconstruct muon segments

- $\Delta P/P \sim 20$ %

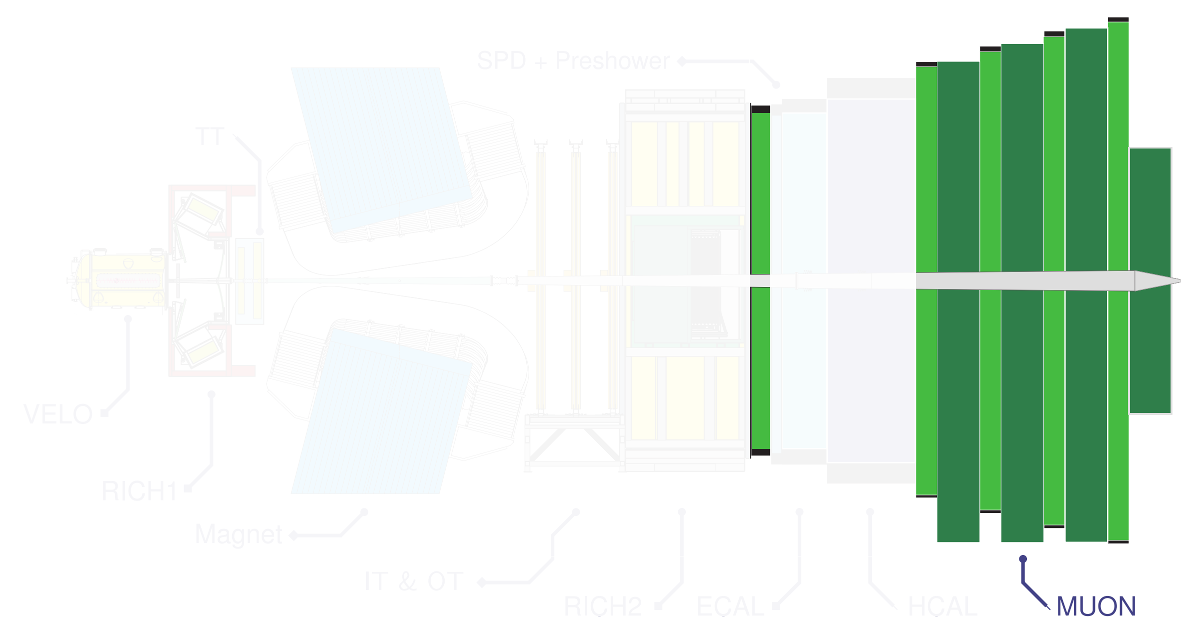
L0 SingleMuon:

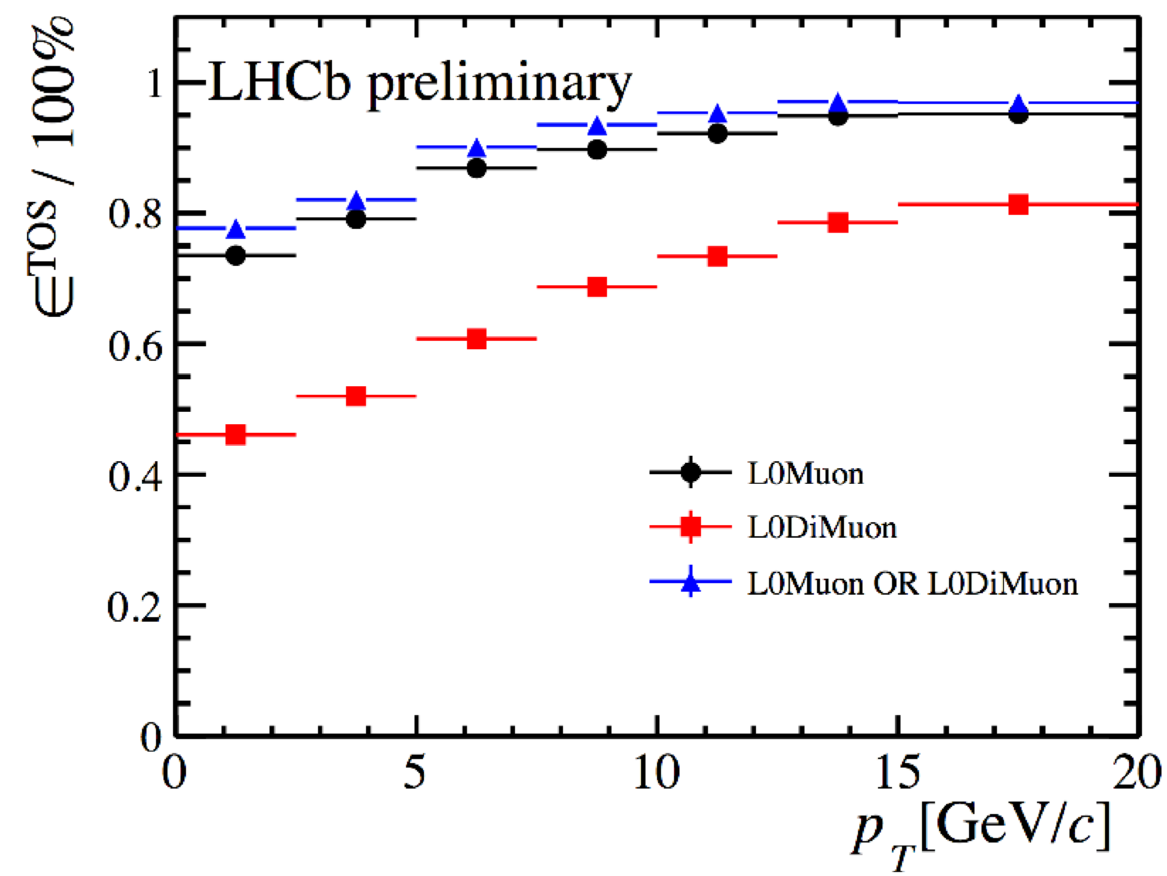
- $P_T > 1.76$ GeV/c

L0 DiMuon:

- $P_{T1} \times P_{T2} > (1.6\ \mathrm{GeV}/c)^2$

Typically over 90% efficient





L0 trigger: Calorimeter

Select hadrons, electrons and photons with large E_T

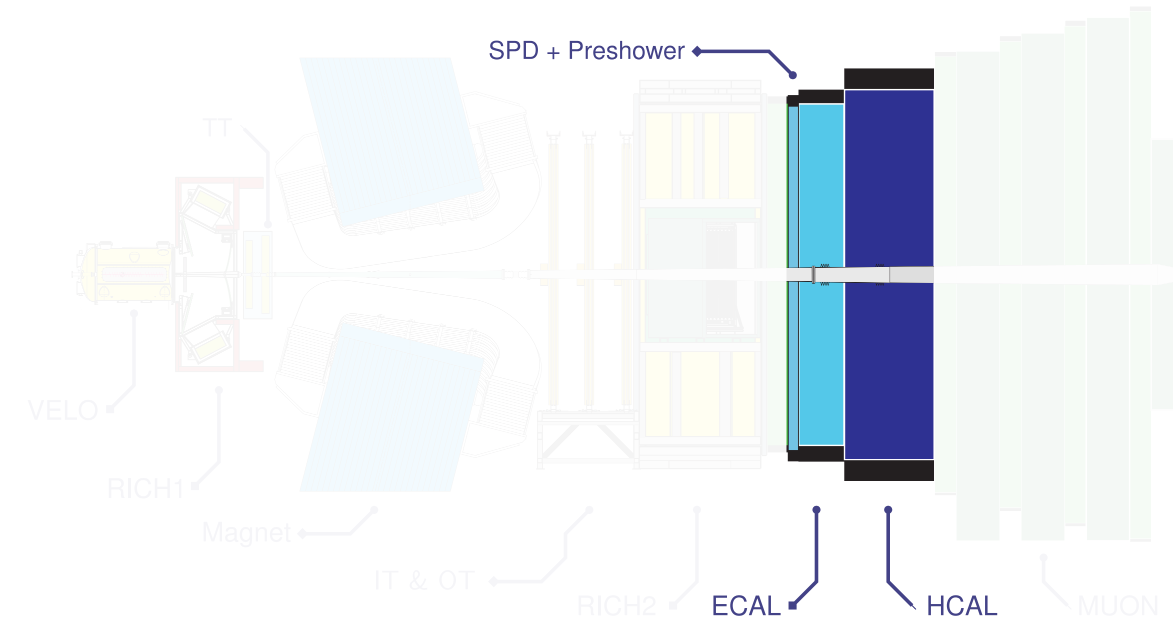
L0 hadron:

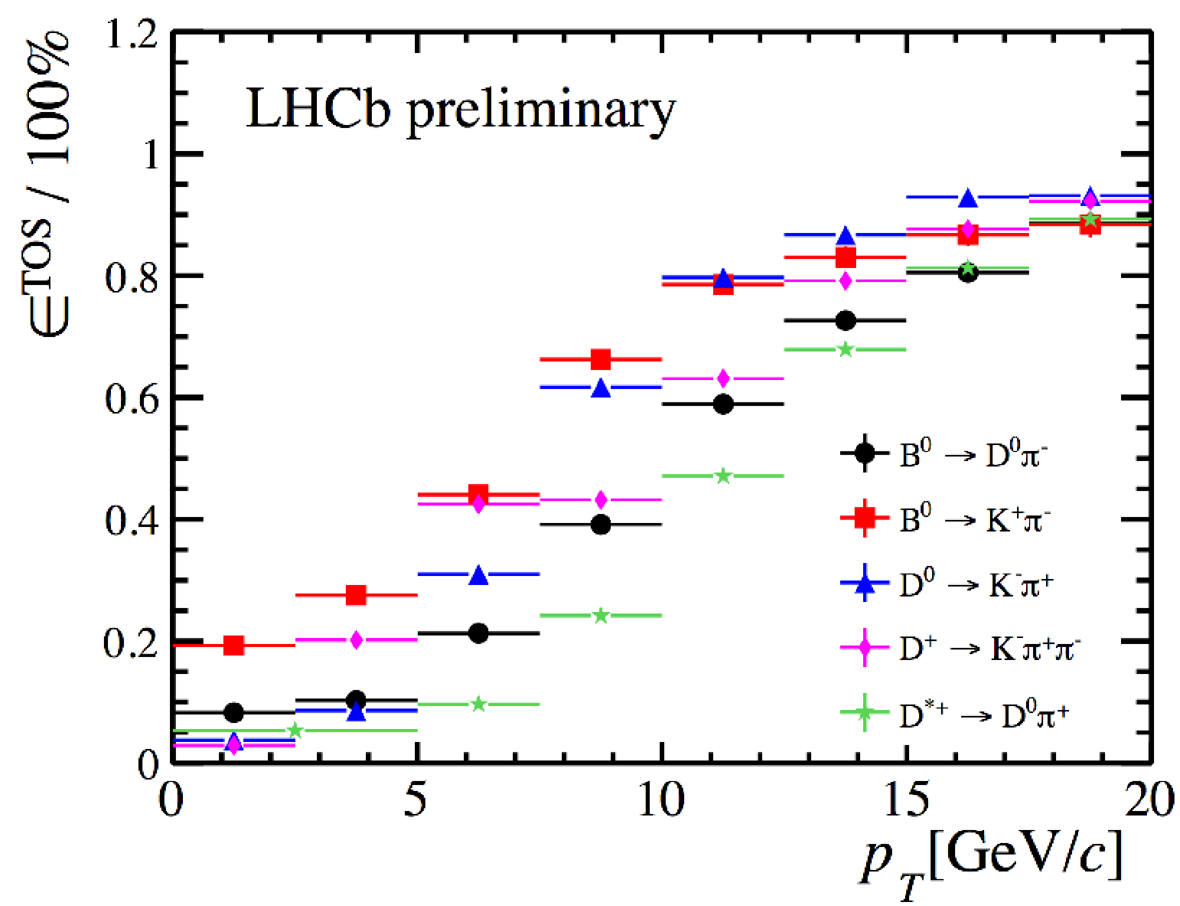
- $E_T >$ 3.6 GeV
- Rate ~490 kHz

L0 electron / photon

- SPD+Preshower discriminate between electrons and photons
- $E_T >$ 3 GeV
- Rate ~ 150 kHz
- ~80 % efficient for $B \to X\gamma$

Total L0 Rate: ~1 MHz





2012: Deferred Trigger

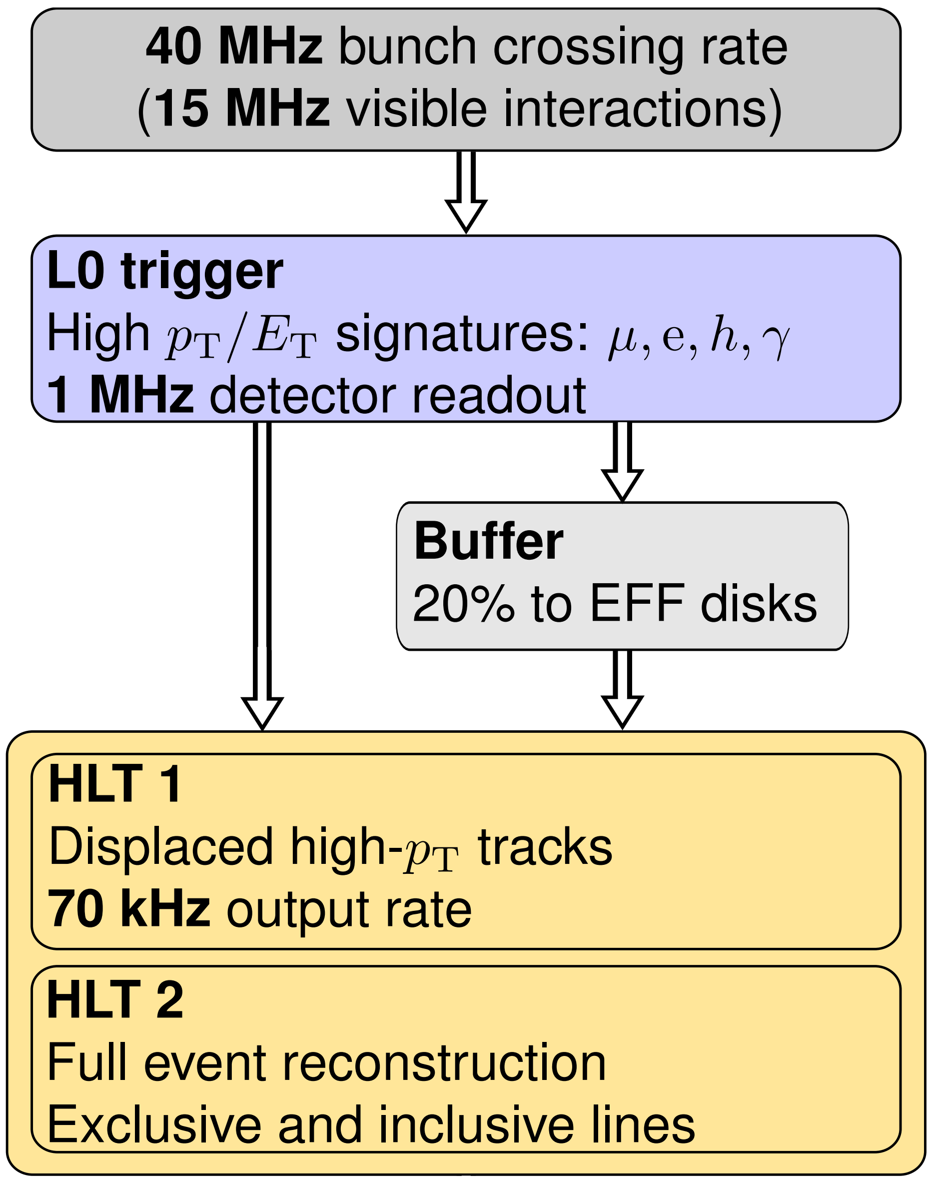
- LHC “only” delivers collisions ~35% of the time
  - trigger farm idle ~65% of the time!

Table 5. LHC availability 2012

Mode	% of scheduled time
Access	14%
Setup	28%
Beam in	15%
Ramp and squeeze	8%
Stable beams	36%

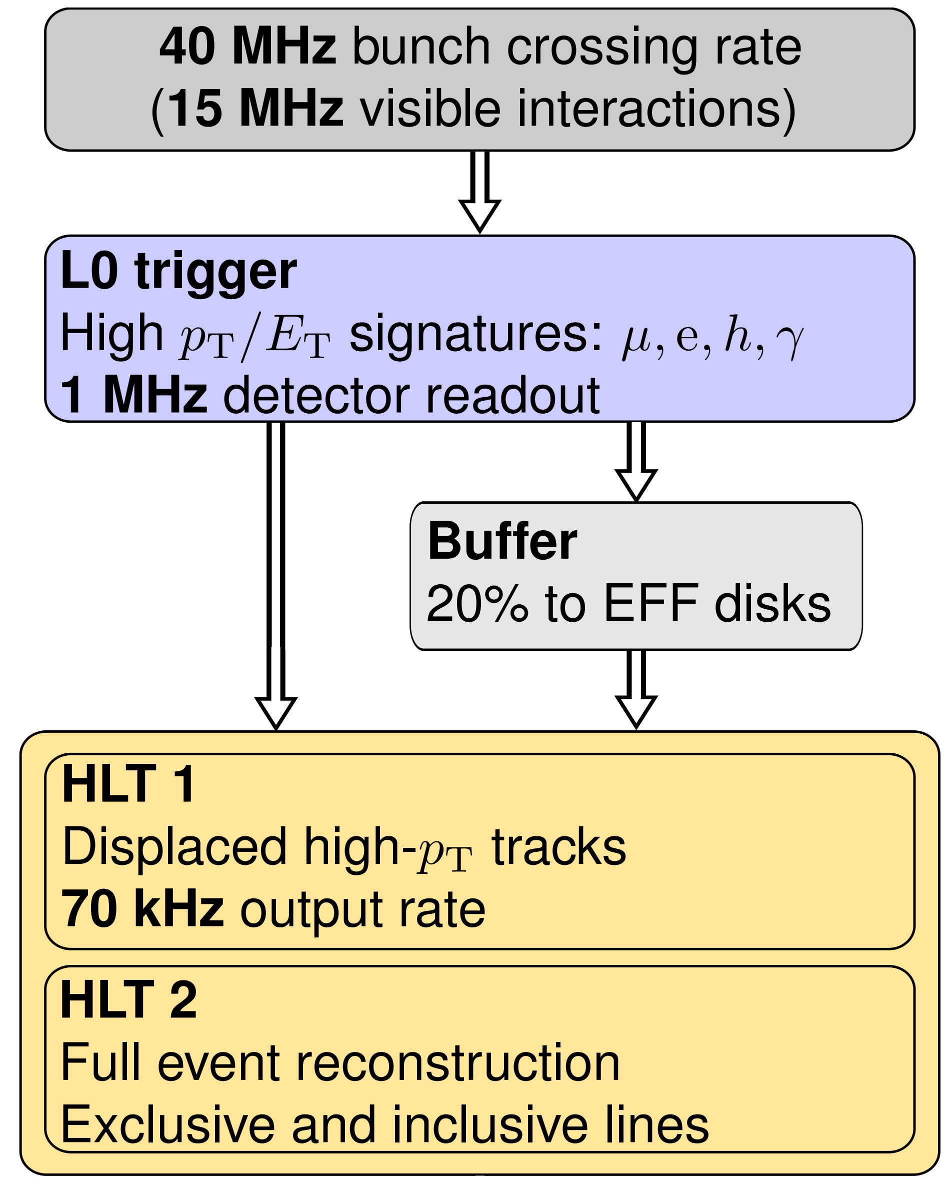
2012: Deferred Trigger

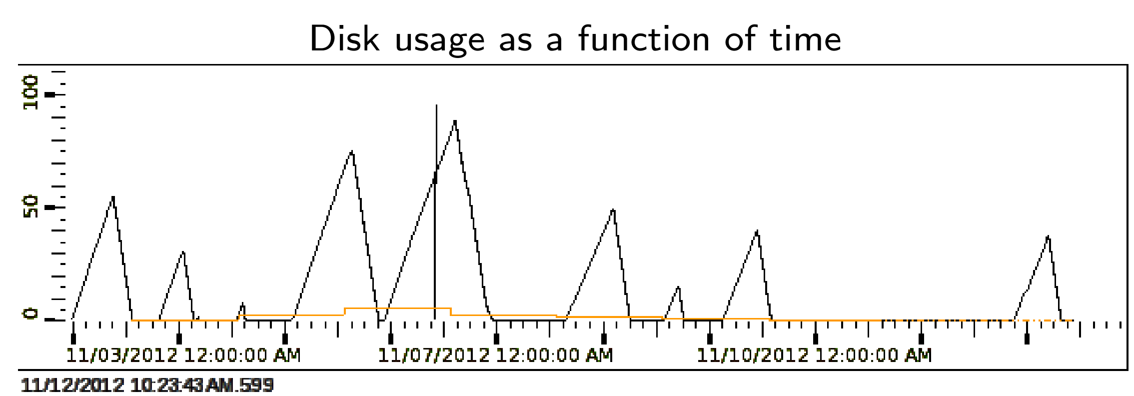
- LHC "only" delivers collisions ~35% of the time
    - trigger farm idle ~65% of the time!
- "Over commit" CPU resources, buffer overflow to local disk & catch up in between fills
    - 20% of L0 triggers are "deferred"
    - 25% extra CPU capacity!
    - allows decrease of Hlt2 tracking thresholds P_T> 500 MeV/c→ P_T>300 MeV/c



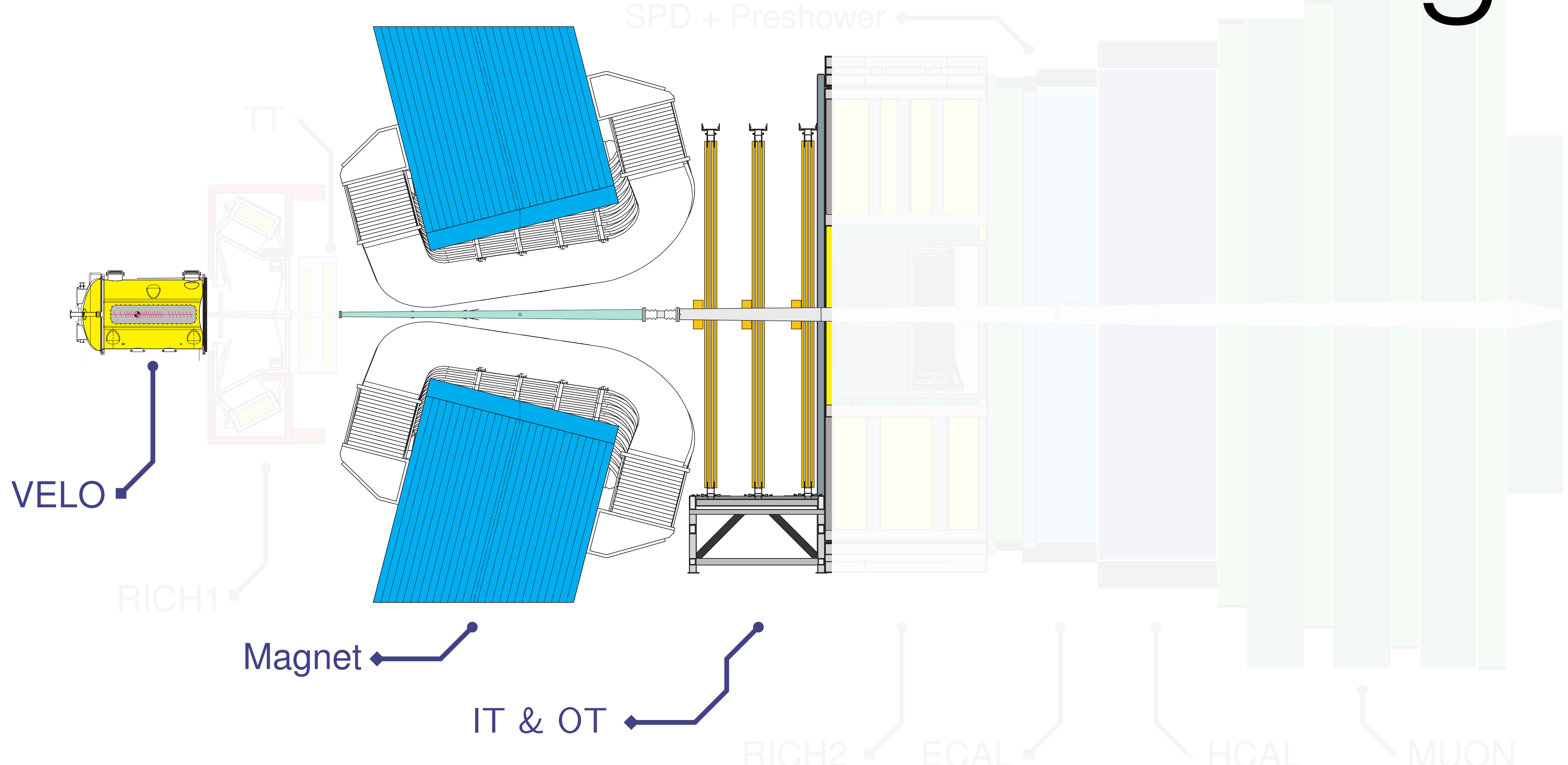
2012: Deferred Trigger

- LHC "only" delivers collisions ~35% of the time
  - trigger farm idle ~65% of the time!
- "Over commit" CPU resources, buffer overflow to local disk & catch up in between fills
  - 20% of L0 triggers are "deferred"
  - 25% extra CPU capacity!
  - allows decrease of Hlt2 tracking thresholds P_T> 500 MeV/c→ P_T>300 MeV/c
  - Peak disk usage in 2012: 88%





HLT1: Partial Tracking



- HLT1 adds tracking in VErtex LOcator (VELO) and primary vertex reconstruction

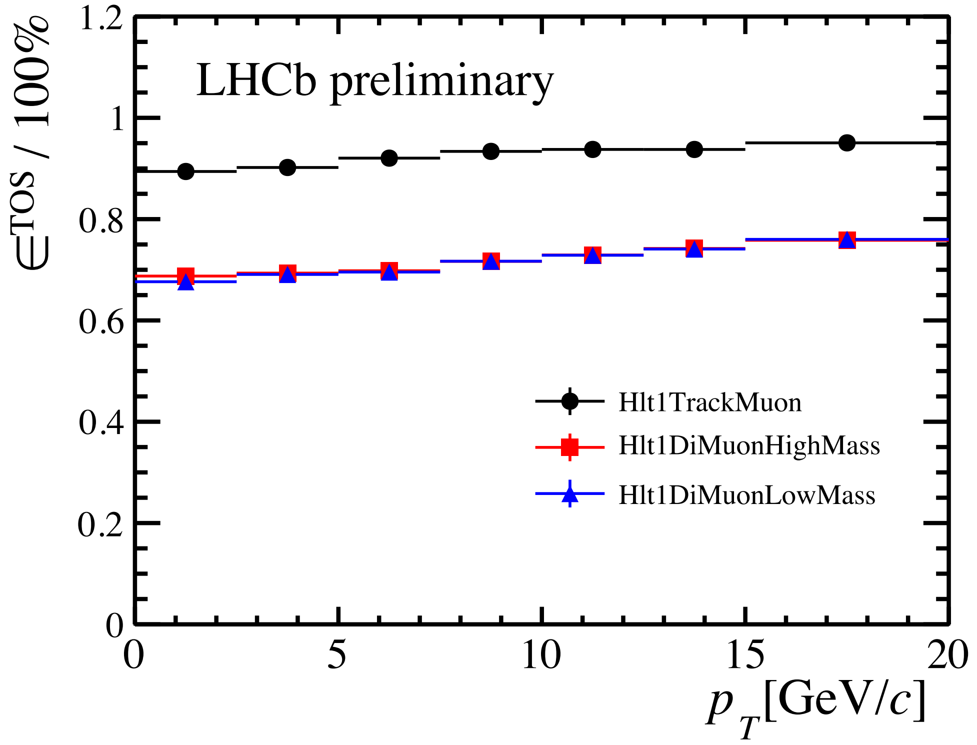
- VELO tracks, either matched to muon hits, or with large IP are extended through the magnet

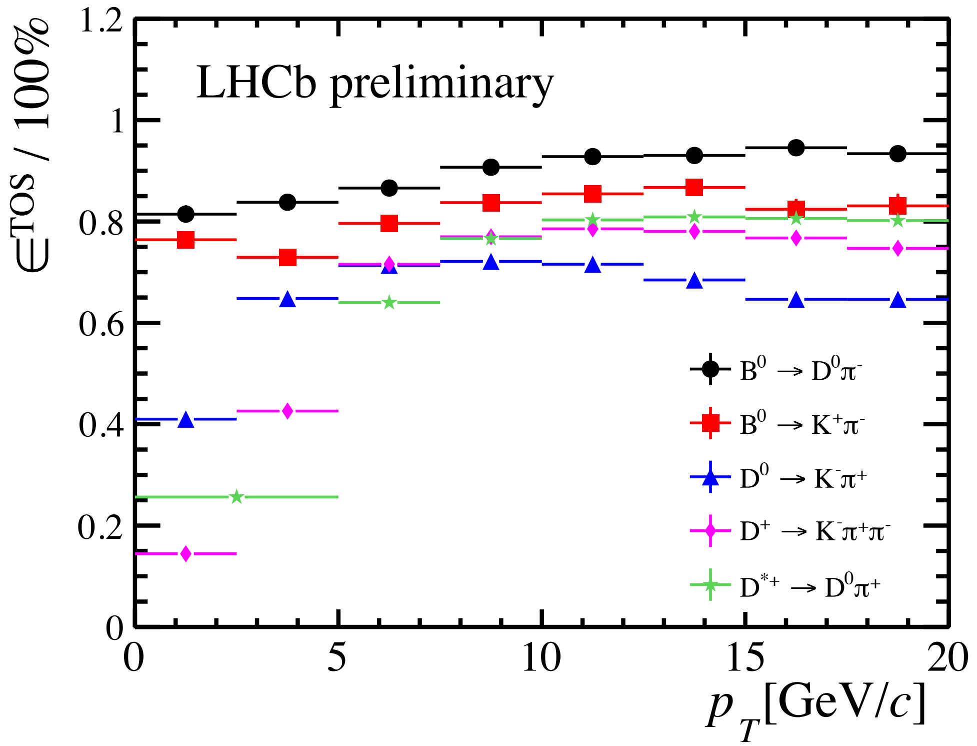
  - P_T dependent search windows:

track	μ	$\mu\ \mu$	other
min. p_T [GeV]	1.0	0.5	1.6

HLT1: Performance

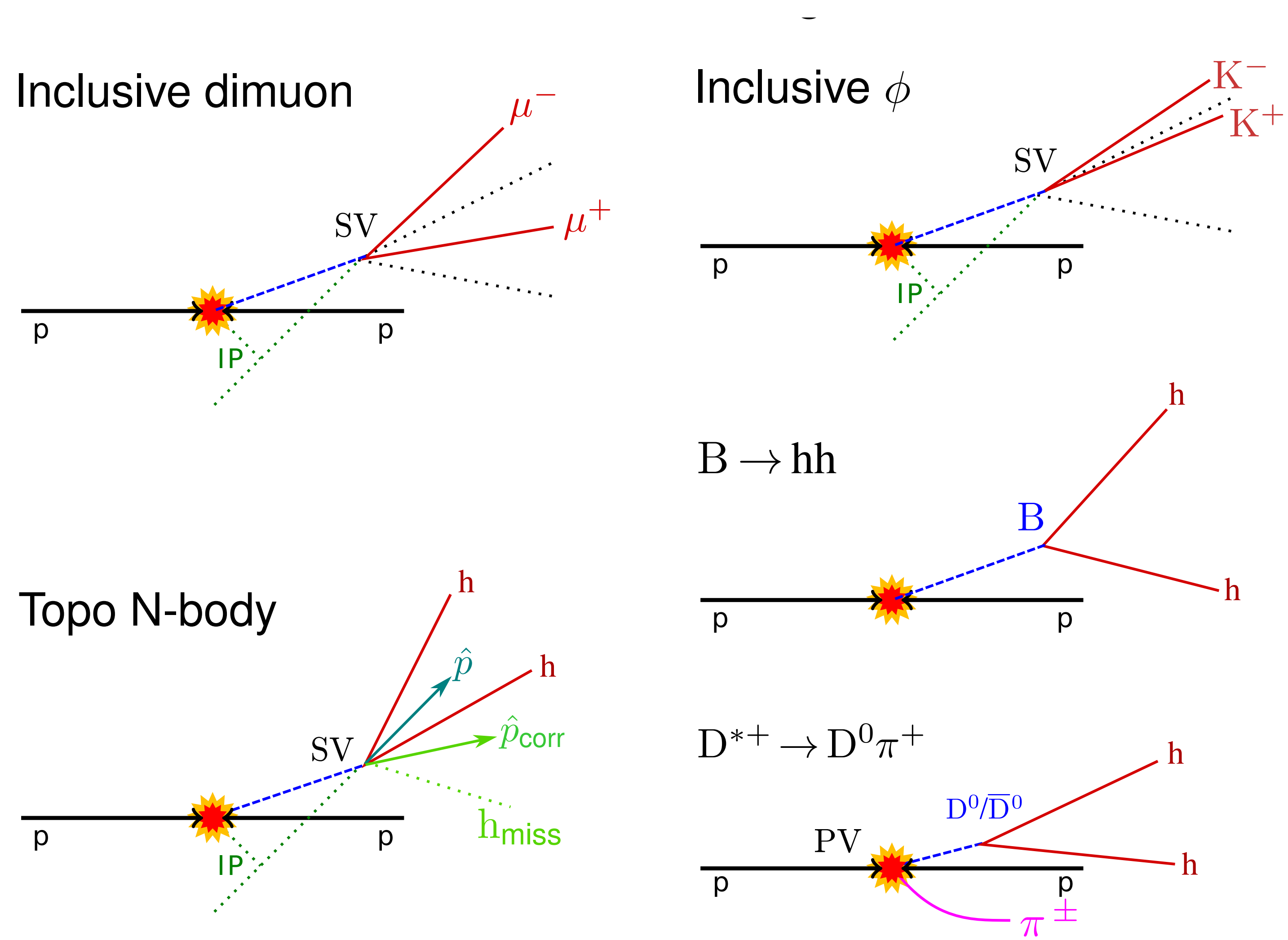
- Muon lines
  - Track matched to muon hits
  - Either high P_T or large IP
  - ~ 14 kHz
- Inclusive lines
  - Single track with large IP and high P_T
  - ~56 kHz
- Total ~ 70 kHz
  - Tuned to maximise HLT2 CPU usage





HLT2: Full Reconstruction

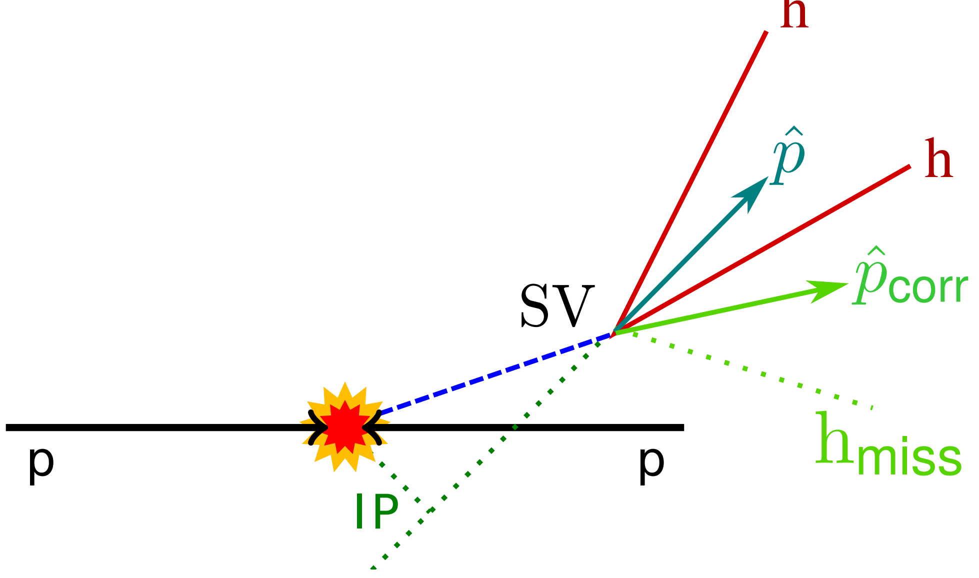
- Tuned versions of offline reconstruction algorithms
  - eg. P_T > 300 MeV/c
- Combination of inclusive and exclusive trigger decisions
- Flexible software environment
  - supports eg. dedicated MVA-based selections

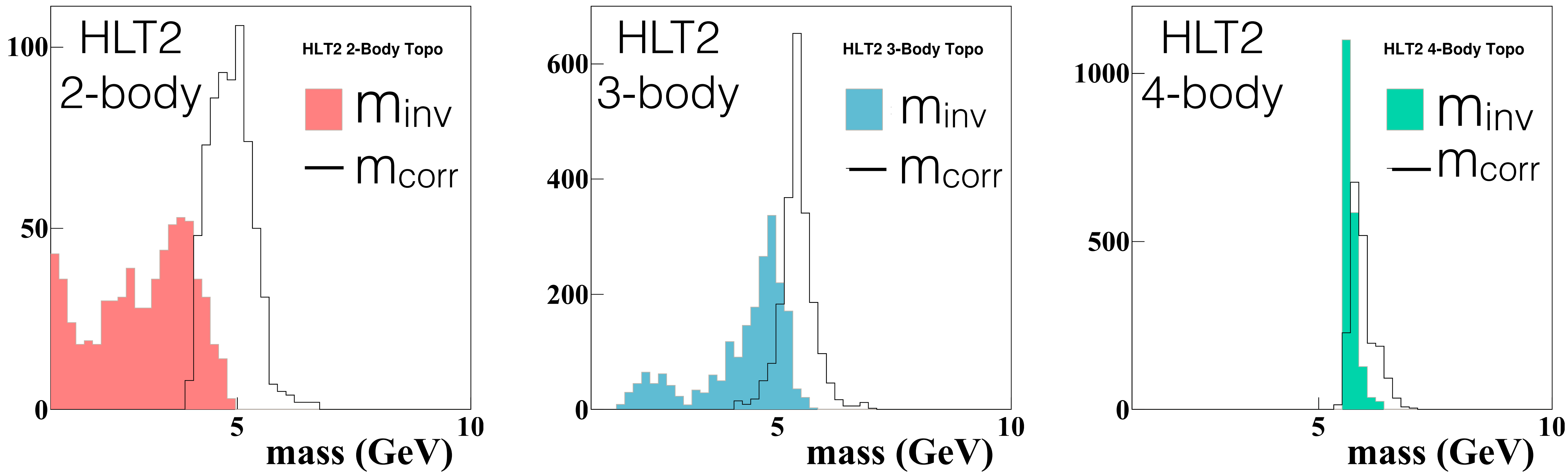


Topological N-body Triggers

- Utilizes excellent vertex and momentum resolution to compute:

$$m_{\rm corr} \equiv \sqrt{m_{\rm inv}^2 + |P_{T\,\rm miss}|^2} + |P_{T\,\rm miss}|$$





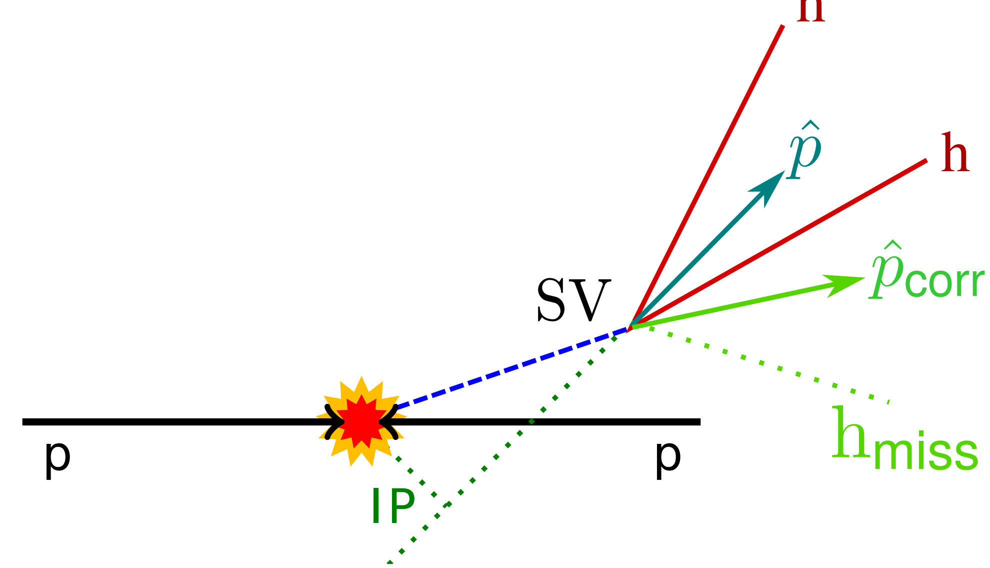
Example: 4-body B decay, $m_{\rm inv}$ and $m_{\rm corr}$ for 2, 3 and 4 body selections

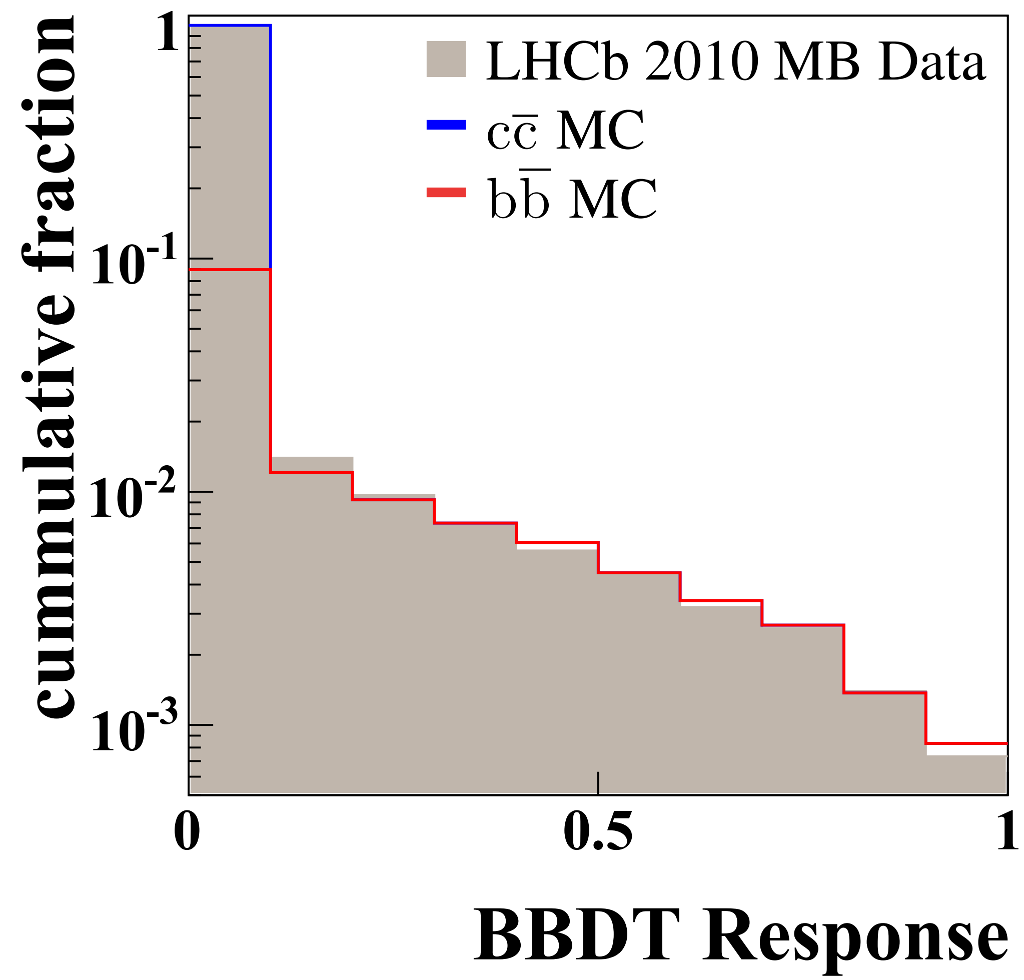
Topological N-body Triggers

- Utilizes excellent vertex and momentum resolution to compute:

$$m_{\mathrm{corr}} \equiv \sqrt{m_{\mathrm{inv}}^2 + |P_{T\,\mathrm{miss}}|^2} + |P_{T\,\mathrm{miss}}|$$

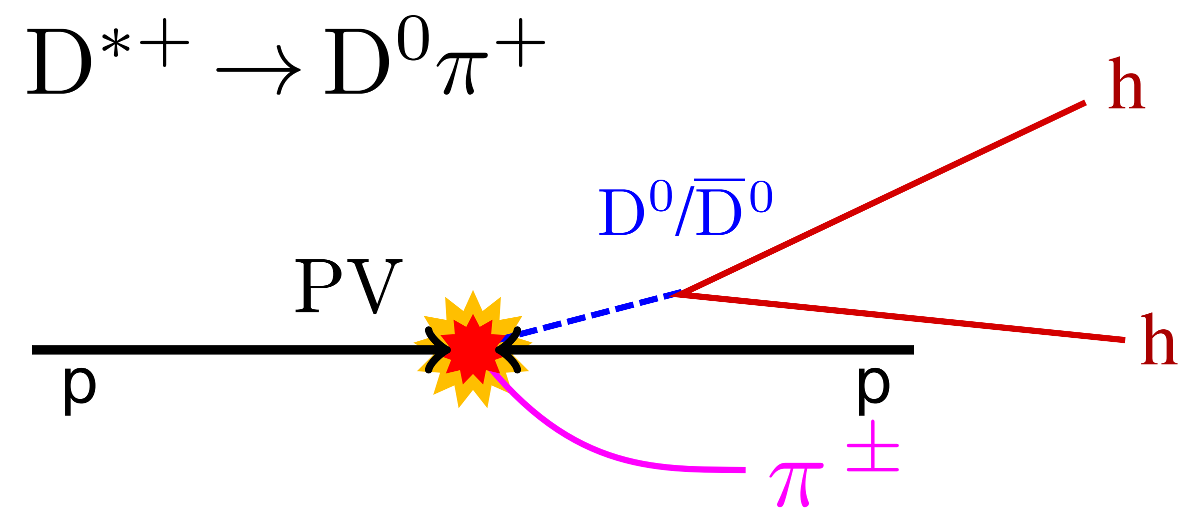
- Uses a dedicated "Bonzai" Boosted Decision Tree [JINST 8 (2013) P02013] with
    - P_T, IPχ^2, FDχ^2, m_{inv}, m_{corr}
- Capable of filling its allotted bandwidth with ~100% pure generic $b\bar{b}$ events

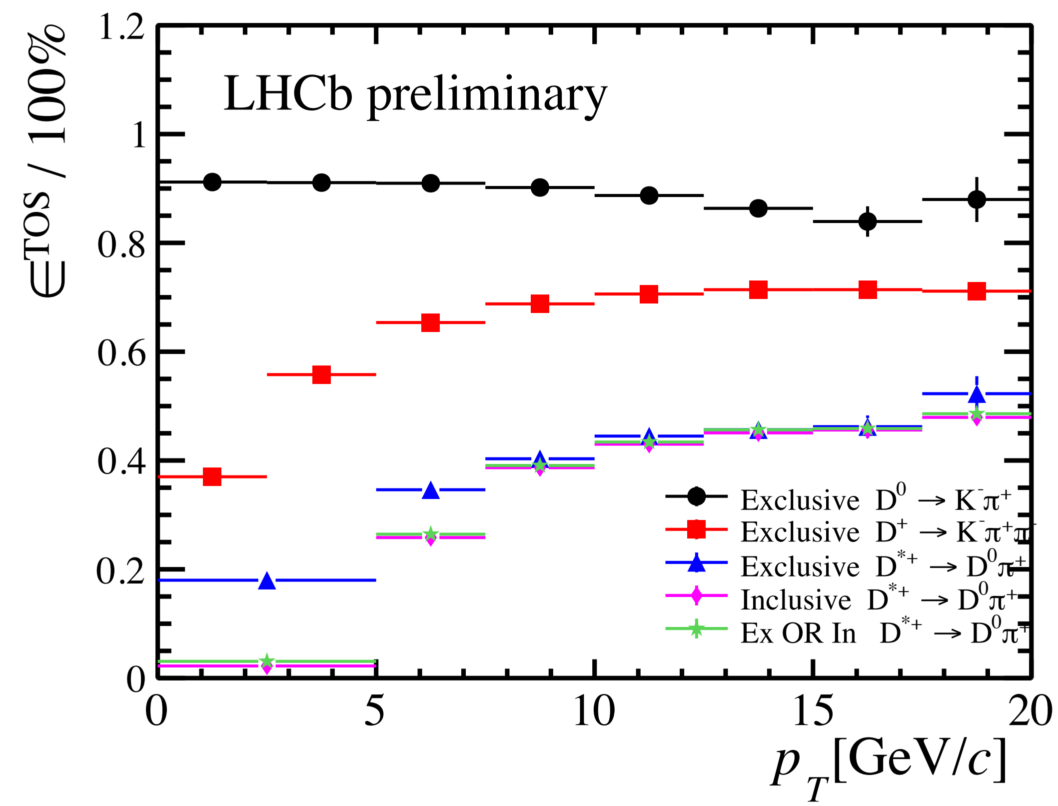




Charm Triggers

- Charm important part of LHCb physics:
  - Observation of D^0-D^0bar oscillations [PRL 110 (2013) 101802]
  - Measurement of D^0-D^0bar mixing parameters [PRL 111 (2013) 251801]
- High production rate, 600 kHz in 2012, requires *exclusive* selections.
- Exception: $D^{*+} \rightarrow D^0 \pi^+$
  - use D^*-D^0 mass difference to select $D^0 \rightarrow h^+ h'^-$
  - Cabibbo favored ($D^0 \rightarrow K^- \pi^+$) rate is 300x suppressed rate ($D^0 \rightarrow \pi^- K^+$)





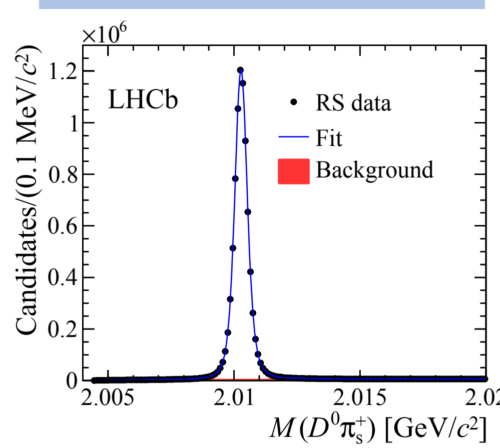
Run1 Performance

- Trigger efficiencies for selected channels

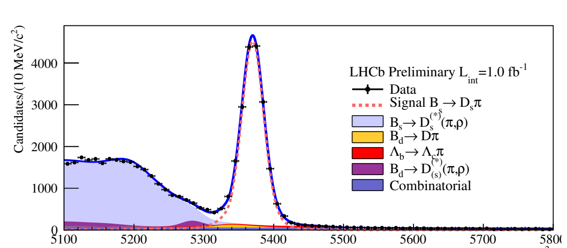
	Hadronic		Dimuon	Radiative
Mode	$D \to hhh$	$B \to hh$	$B^+ \to J/\psi K^+$	$B^0 \to K^*\gamma$
ϵ(L0) [%]	27	62	93	85
ϵ(HLT | L0) [%]	42	85	92	67
ϵ(HLT $\times$ L0) [%]	11	52	84	57

- Very pure samples after offline selection

$D^* \to D^0\pi$ [1211.1230]

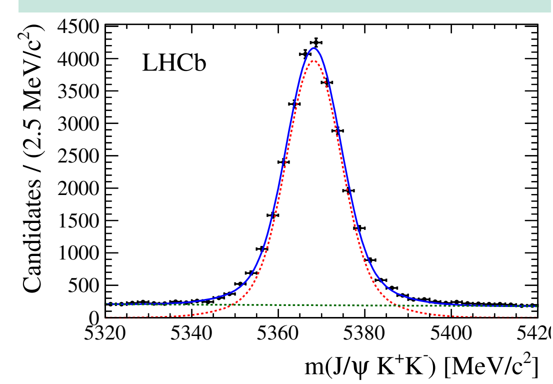


$B_s^0 \to D_s\pi$ [1304.4741]

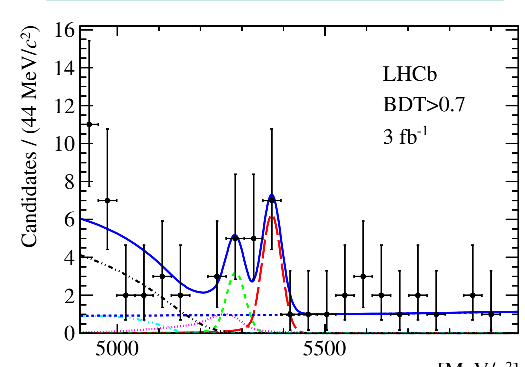


$D_s\pi$

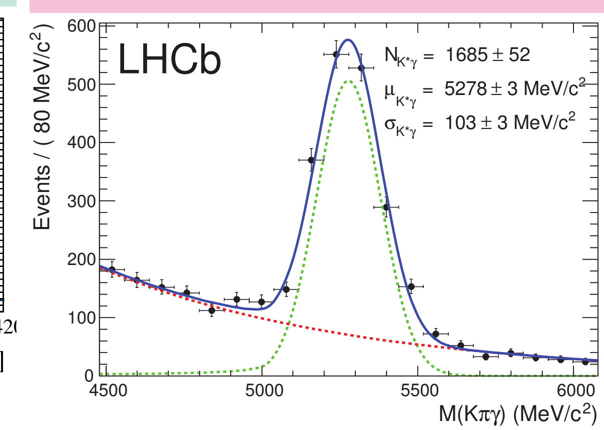
$B_s^0 \to J/\psi\phi$ [1304.2600v3]

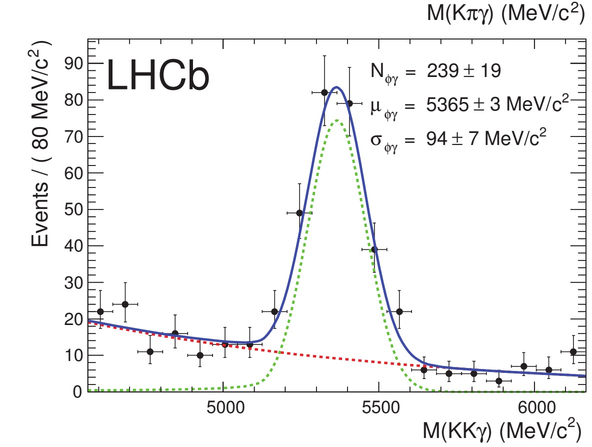


$B_s^0 \to \mu\mu$ [1211.2674]



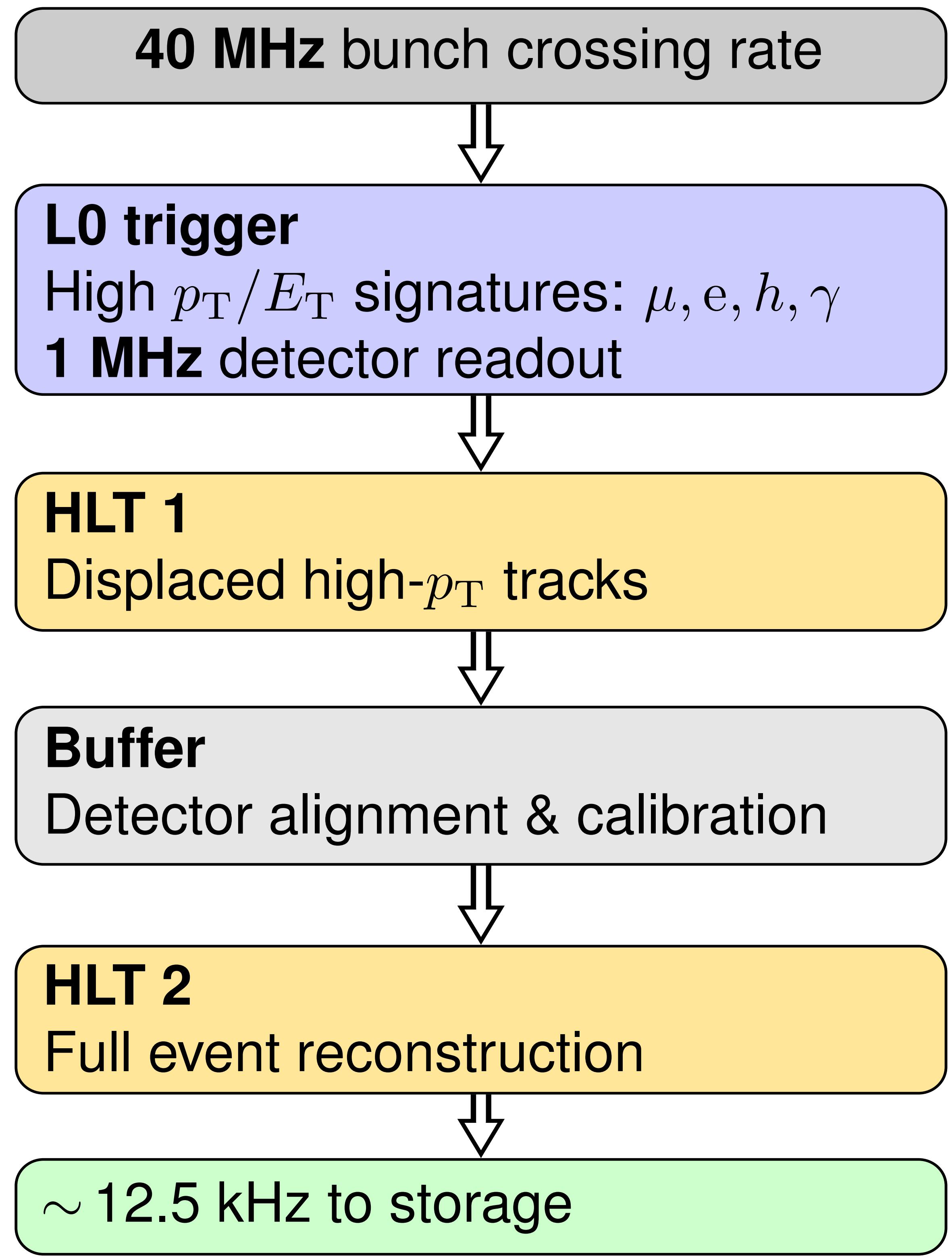
$B_d^0 \to K^*\gamma, B_s^0 \to \phi\gamma$ [1202.6267]





Run2 Prospects

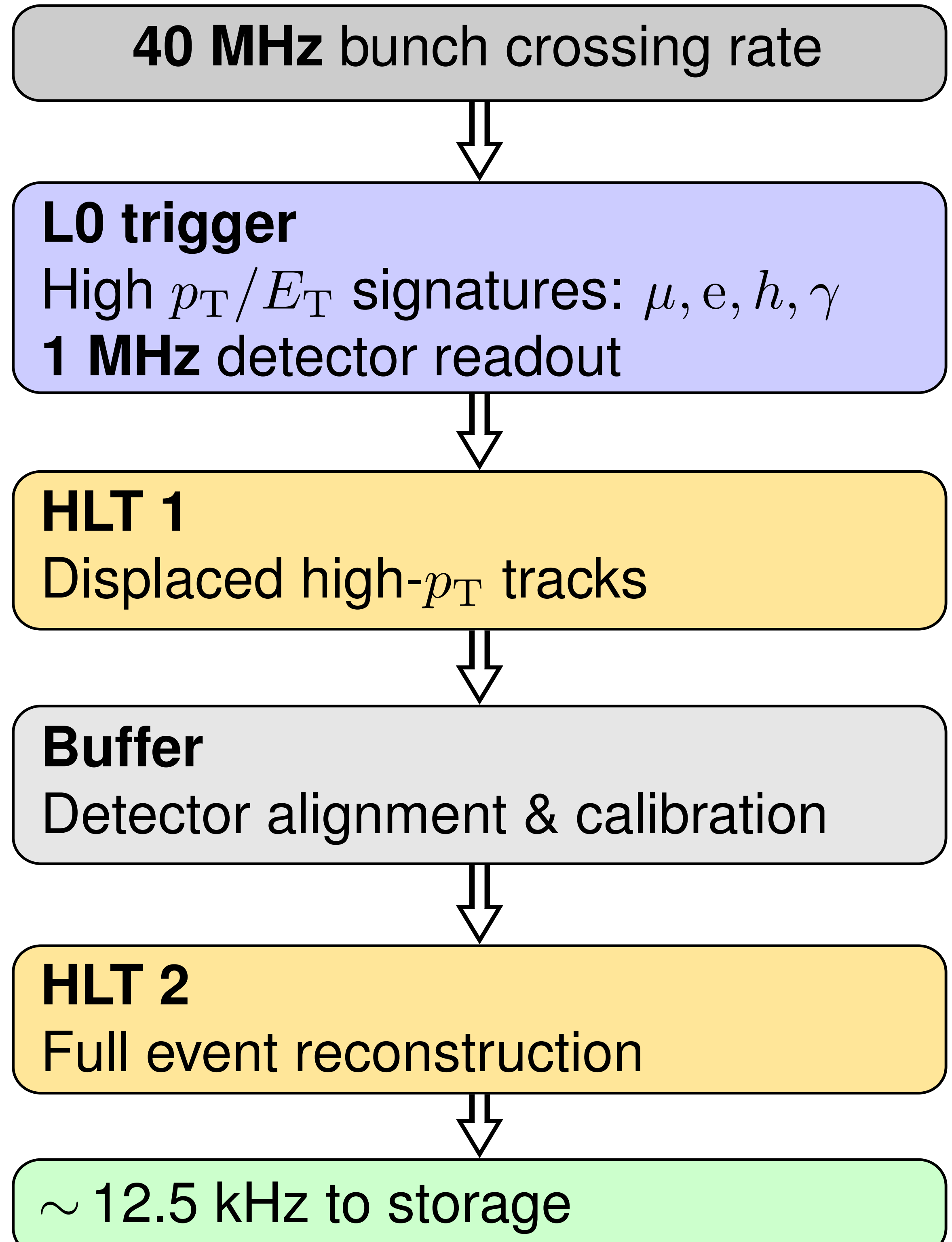
- Energy: 8 TeV → 13 TeV
  - $\sigma_{inelastic}$: x 1.15
  - σ_{bb}: x1.6
  - multiplicity: x1.2
- Bunch spacing: 50 ns → 25 ns
  - pileup: / 2
  - but still 1 MHz L0 limit!



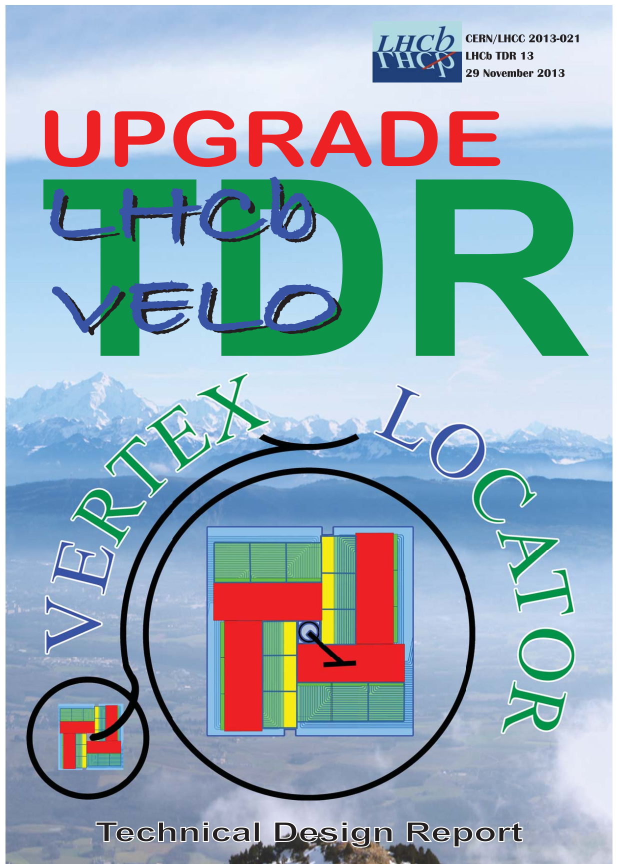
Run2 Prospects

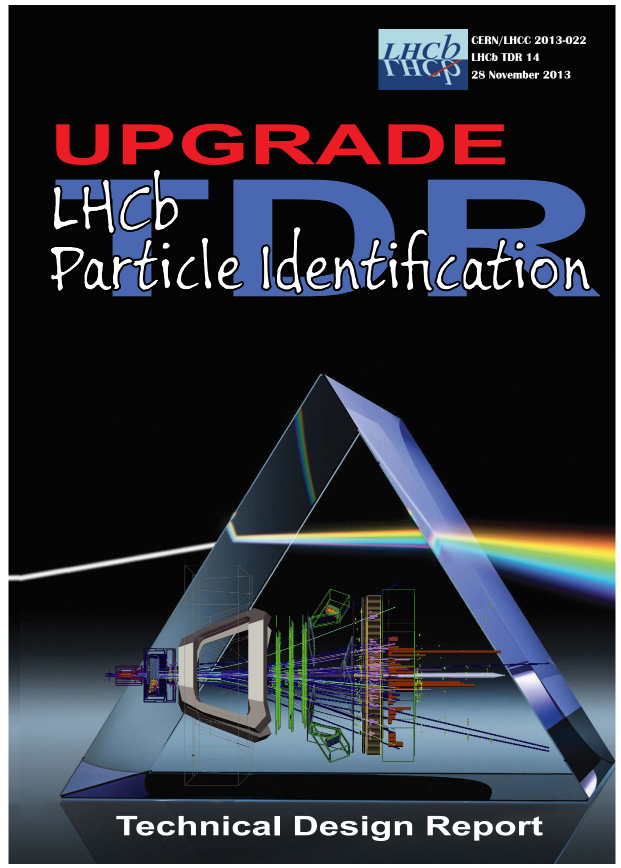
Changes in architecture:

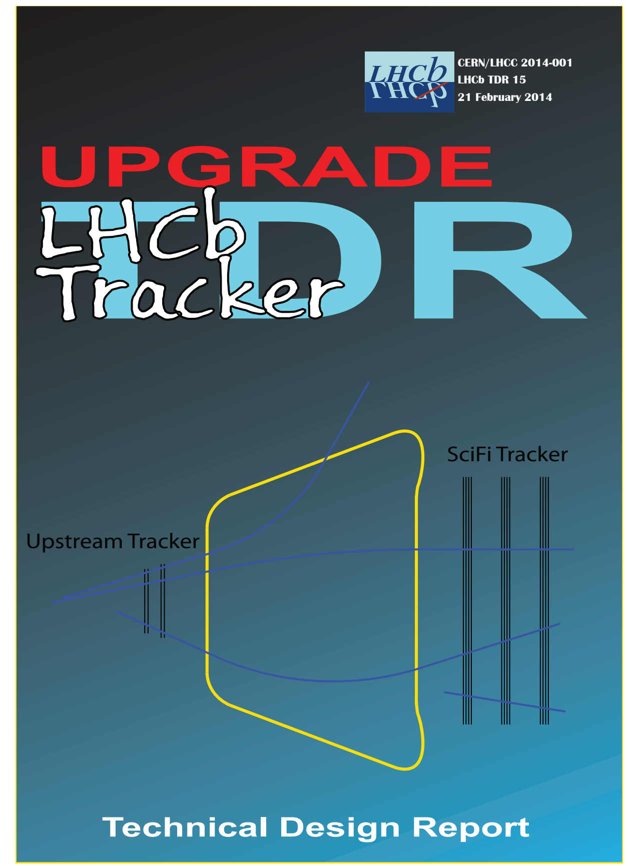
- ‘Split’ Hlt1 and Hlt2
  - Buffer data after HLT1, perform alignment & calibration, *prior* to Hlt2
  - Hlt2 now very close to offline reconstruction, *including* RICH PID
  - RICH PID allows pre-scaling of Cabibbo favored charm, whilst keeping the full suppressed rate.
- Increased output rate, add ‘parking’
- Investigate analysis directly on HLT output only (2.5 kHz without offline reconstruction)

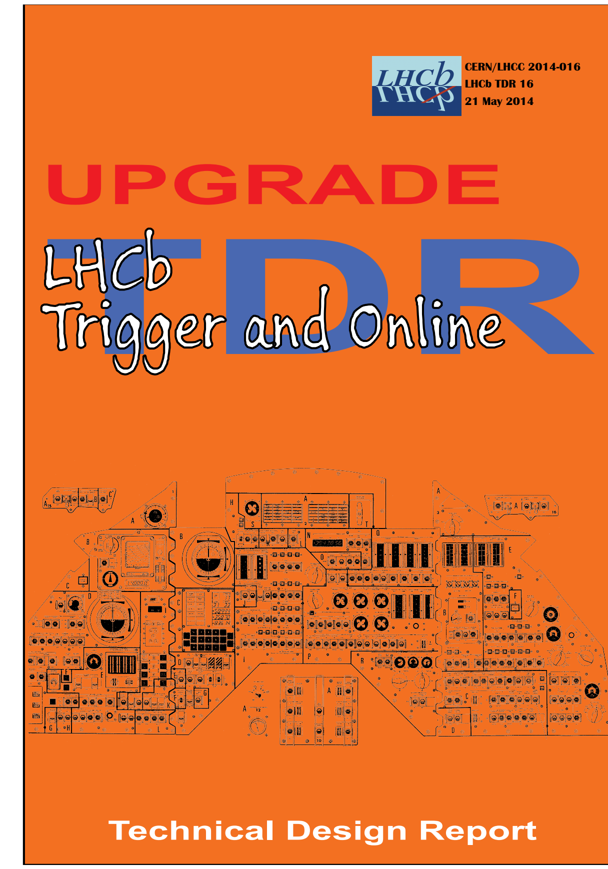


The LHCb Upgrade



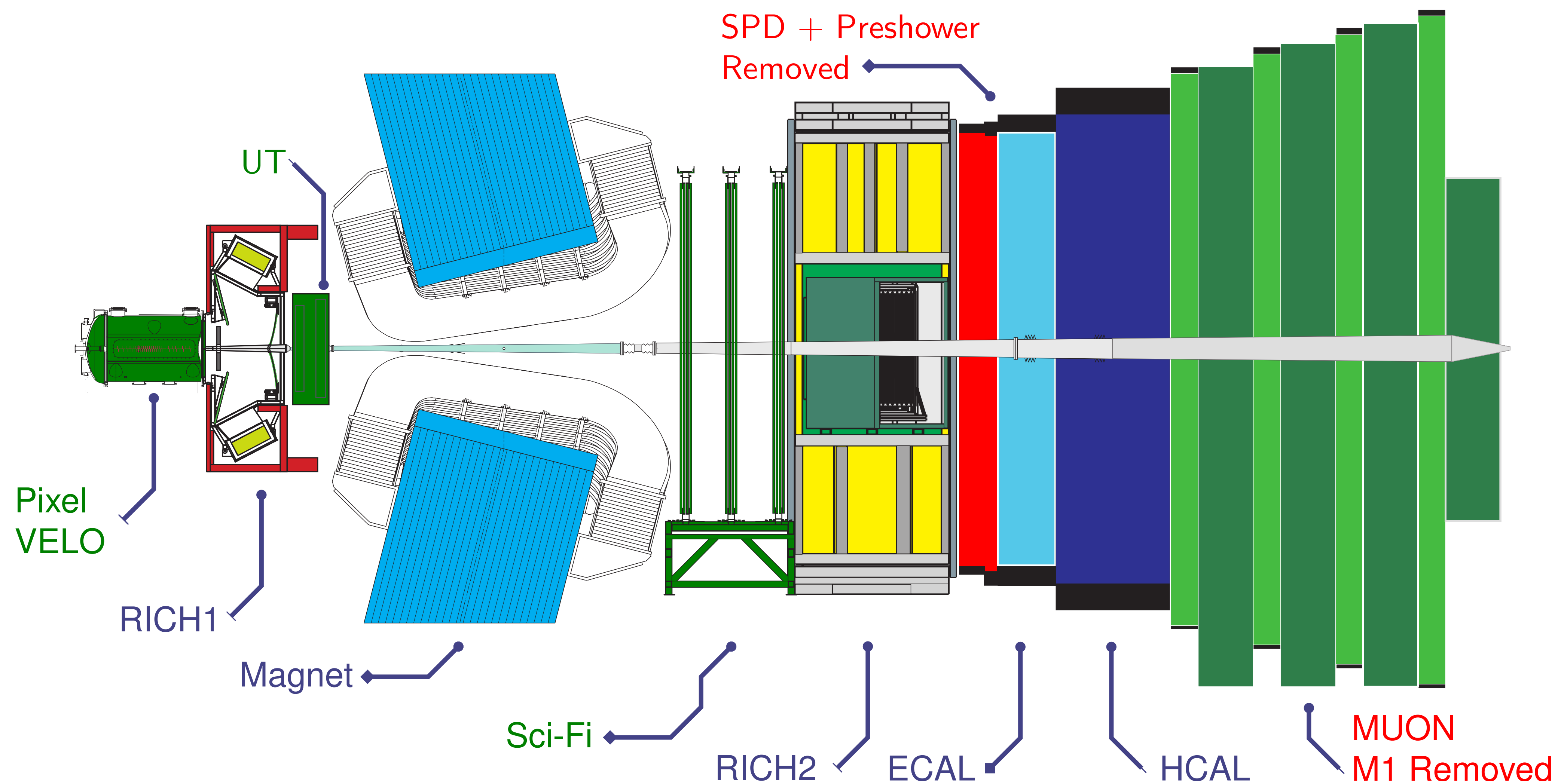






The LHCb Upgrade

- After LS2, LHCb will run at 5x higher luminosity: $L = 2 \cdot 10^{33}\ cm^{-2} s^{-1}$



- VELO: r,ϕ strips →pixels
- Trackers: strawtubes → scintillating fibers + silicon microstrips
- RICH: replace photon detectors; CALO: SPD, PRS removed; MUON: M1 removed

Upgrade Environment

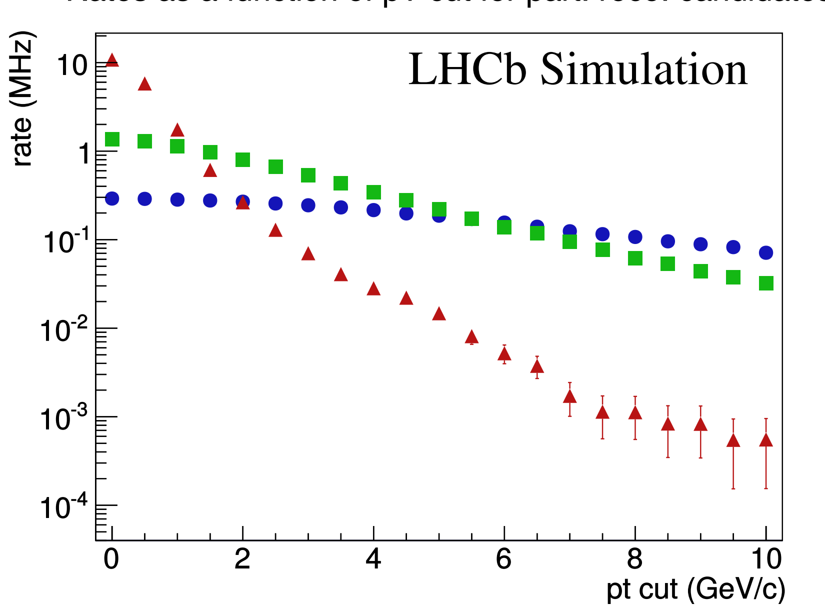
- Average pp collisions per bunch crossing: 2.0→7.6

Run I	Per event	with vertex in VELO	Rate [GB/s]
b-hadrons	0.0258 ± 0.0004	0.0029 ± 0.0001	0.9
c-hadrons	0.297 ± 0.001	0.0422 ± 0.0005	3.3
light, long-lived hadrons	8.04 ± 0.01	0.511 ± 0.002	1.1

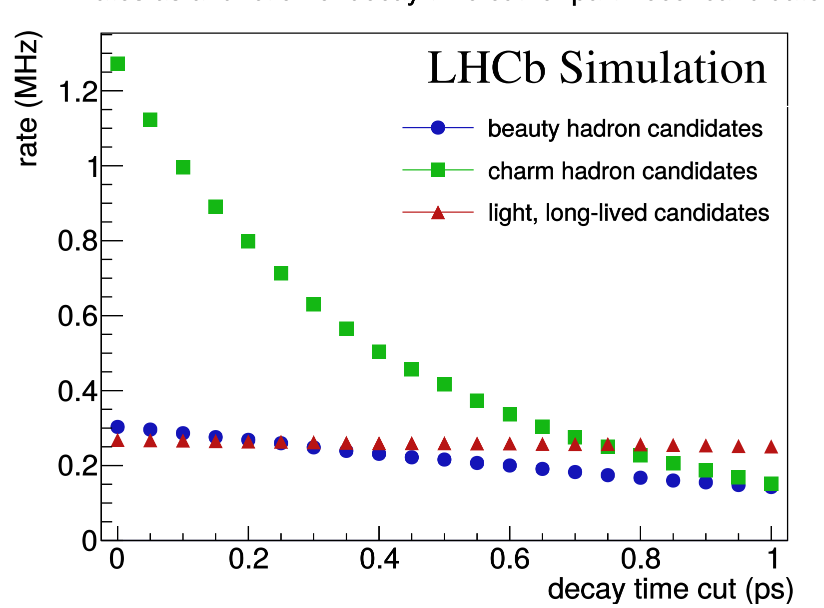
Upgrade	Per event	with vertex in VELO	Rate [GB/s]
b-hadrons	0.1572 ± 0.0004	0.01874 ± 0.0001	27
c-hadrons	1.422 ± 0.001	0.2138 ± 0.0005	80
light, long-lived hadrons	33.291 ± 0.006	2.084 ± 0.001	26

- Challenge: must go beyond rejecting background —- classify signal, and choose wisely…

Rates as a function of pT cut for part. reco. candidates



Rates as a function of decay time cut for part. reco. candidates

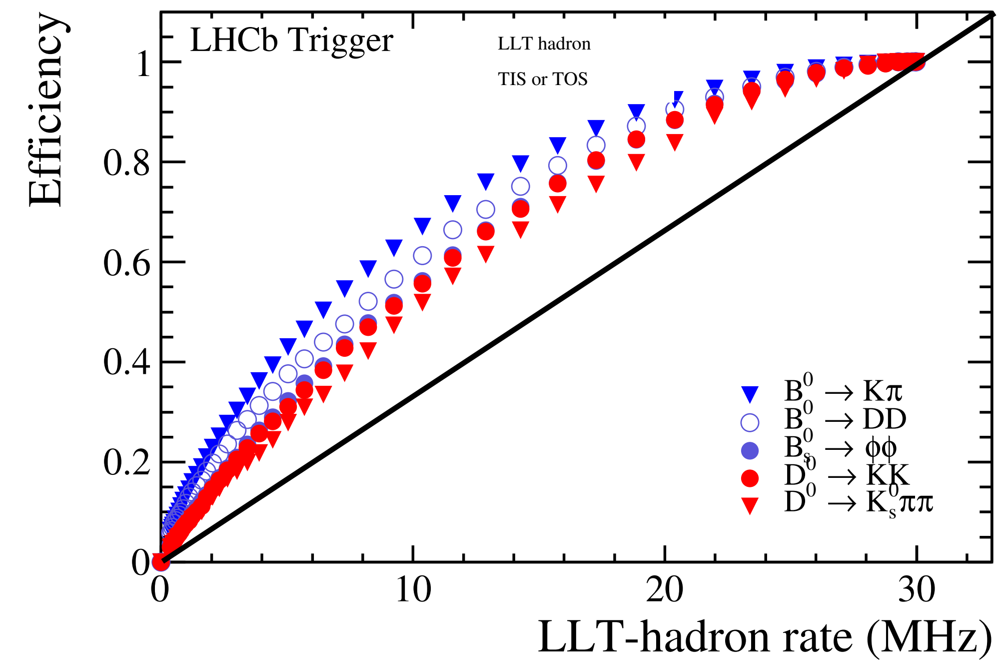


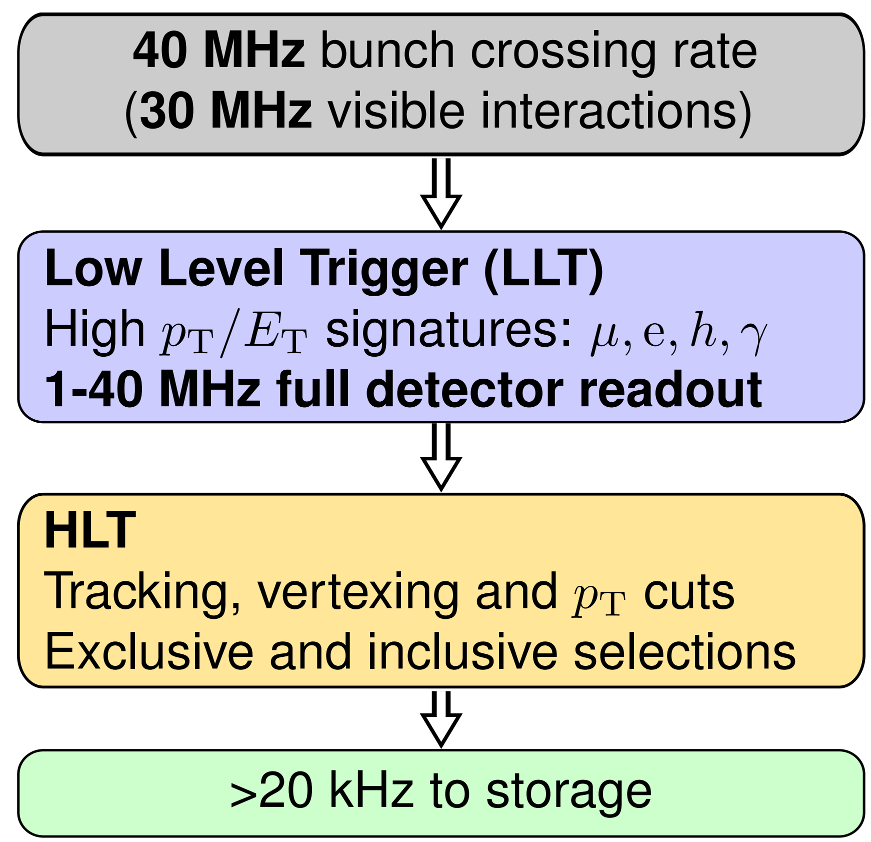
- P_T and IP alone not sufficient to reduce rate: requires *all* available detector information…

Triggerless Readout @ 40 MHz

- At L= 2 ·10^{33}cm^{-2}s^{-1}, the 1 MHz readout limit becomes a bottleneck
  - Signal no longer easily identifiable
- Readout upgraded to 40 MHz
- ⇒ Ship *every* visible pp interaction

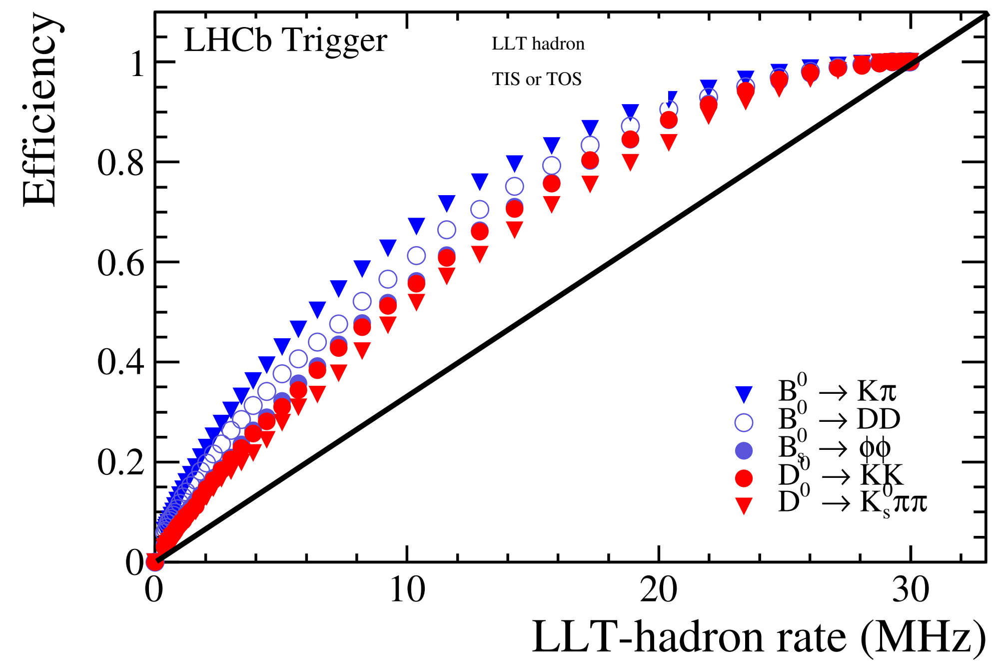
  (30 MHz) to a CPU farm running the Higher Level Trigger
- Low-Level Trigger (LLT) only as 'handbrake' during commissioning

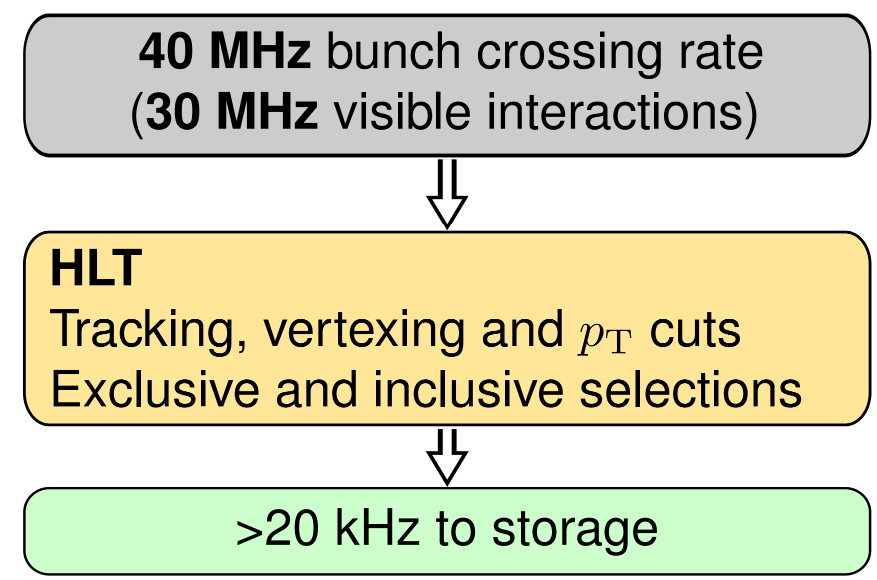




Triggerless Readout @ 40 MHz

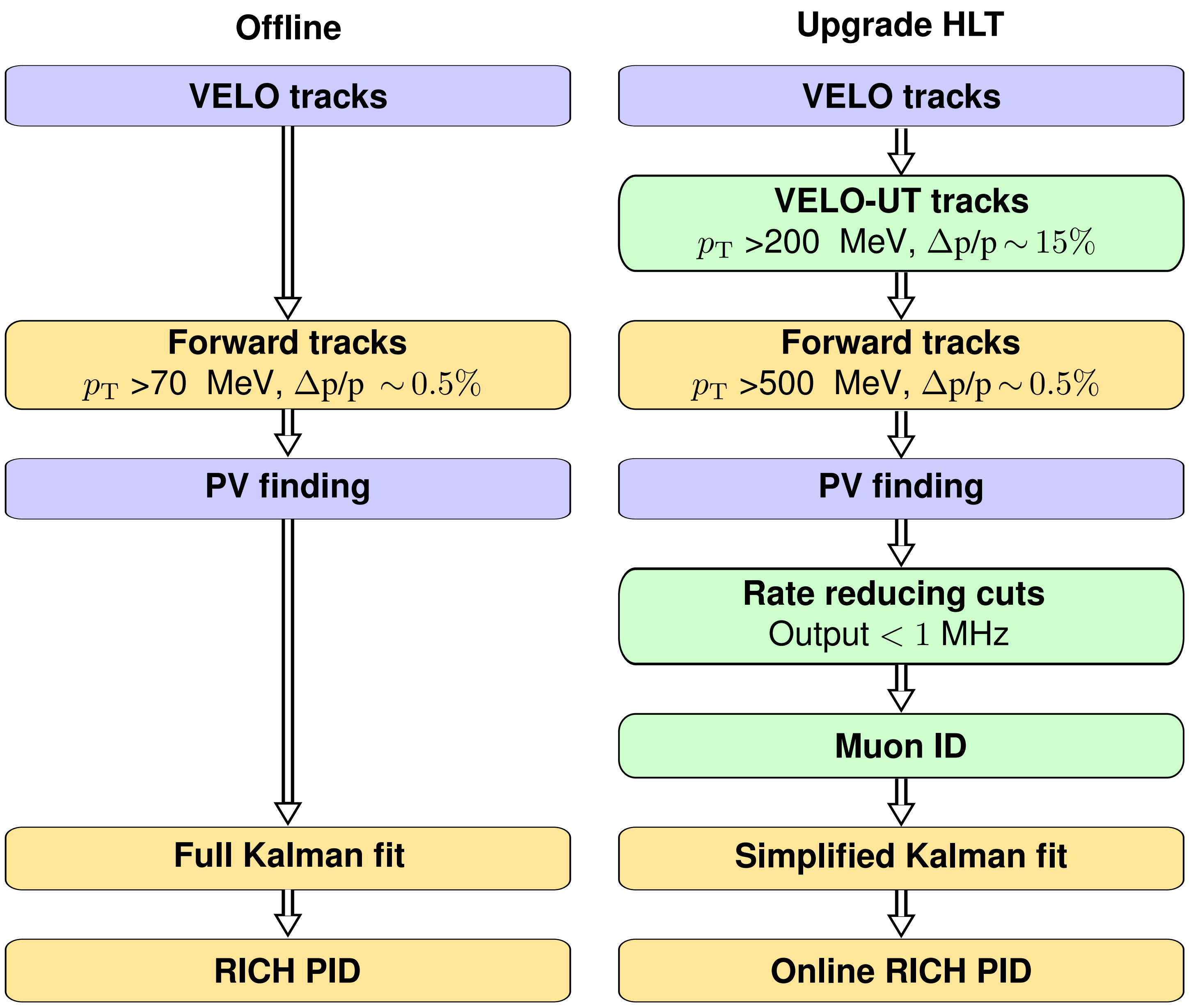
- At L= 2 ·10^{33}cm^{-2}s^{-1}, the 1 MHz readout limit becomes a bottleneck
  - Signal no longer easily identifiable
- Readout upgraded to 40 MHz
- ⇒ Ship *every* visible pp interaction (30 MHz) to a CPU farm running the Higher Level Trigger
- Low-Level Trigger (LLT) only as 'handbrake' during commissioning





Upgrade Trigger

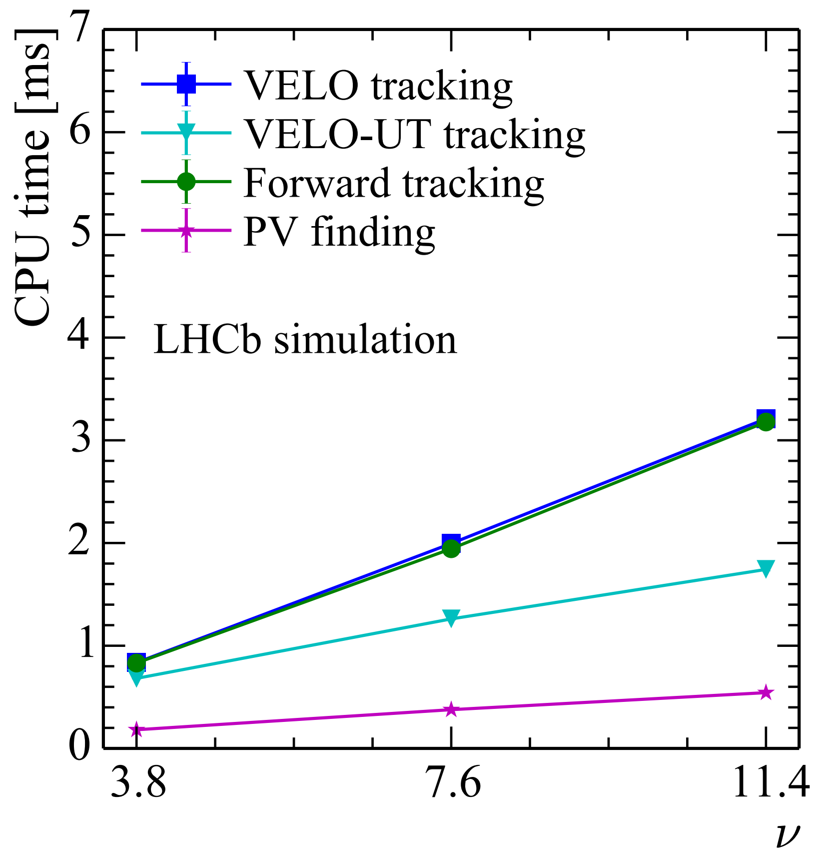
- Offline-quality tracking, *in software*, is possible @ 30 MHz
    - Estimated trigger farm: O(1000) servers, 2.8 MCHF.
    - Tracking requires 5.4 ms/event, out of an estimated budget of 13 ms/event @ 30 MHz (*)
    - Thanks to the upgraded vertex detector & tracker designs!
- Converge online and offline reconstruction.

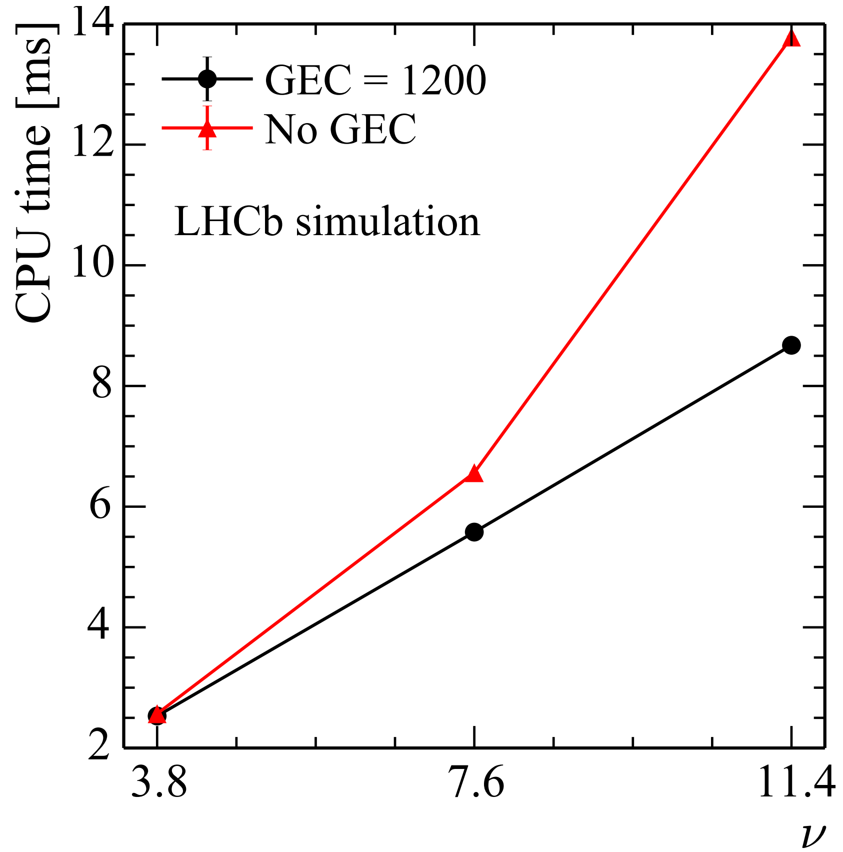


(*) on our 2011 reference machine: Intel X5650 (Westmere) @ 2.67 GHz

Robustness of benchmarks

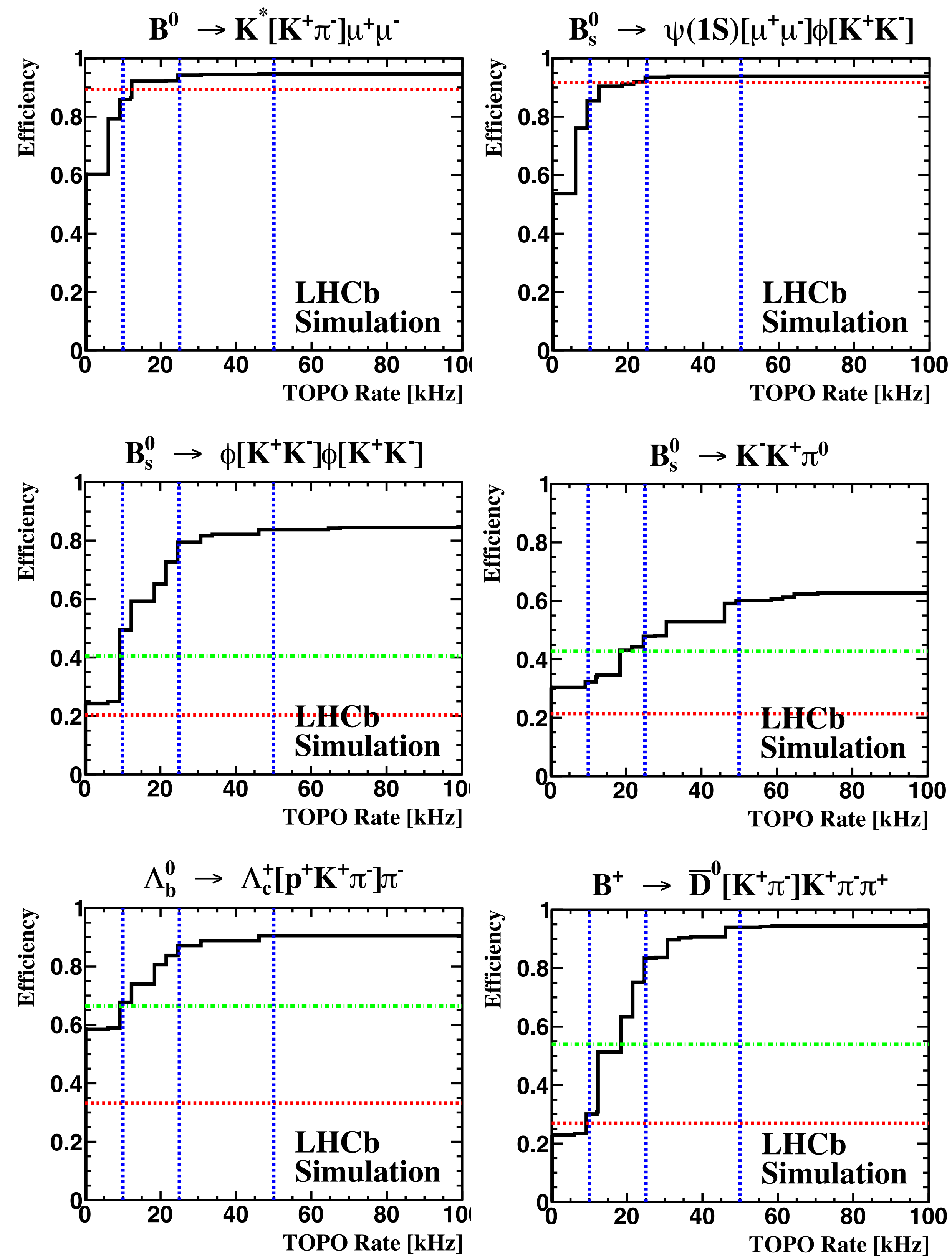
- Timing of tracking studied at three working points:
  - $L = 1 \cdot 10^{33}\ cm^{-2}s^{-1}$ [ν=3.8]
  - $L = 2 \cdot 10^{33}\ cm^{-2}s^{-1}$[$\nu$=7.6, nominal]
  - $L = 3 \cdot 10^{33}\ cm^{-2}s^{-1}$[$\nu$=11.4]
- Several optimizations can be made to improve timing and efficiency for different working points
  - Global Event Cut (GEC) on event multiplicity also used in Run 1:
  - “Crowded” events contain more background, use disproportionate amount of resources





Upgrade Topological Trigger

- Same principle as Run 1 : preselect displaced tracks with $\sum P_T$, followed by BBDT
- Timing: <0.1 ms (*)
- At 25-50 kHz output rate, large efficiency gains over Run 1
  - red: run 1 efficiency
  - green: 2x run 1 efficiency
- LHCb-PUB-2014-031



(*) on our 2011 reference machine: Intel X5650 (Westmere) @ 2.67 GHz

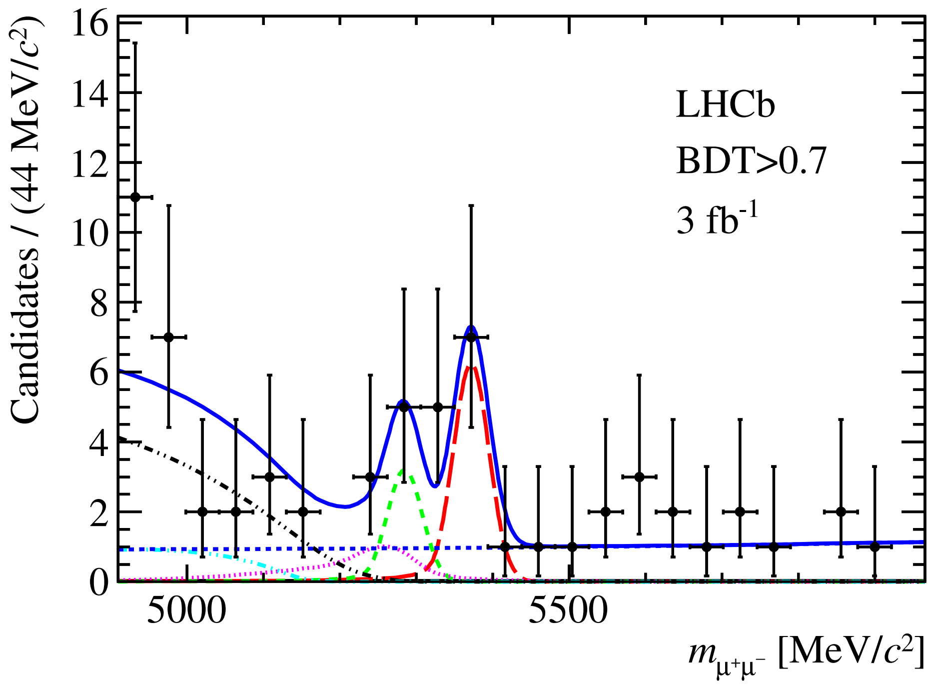
Upgrade Sensitivities

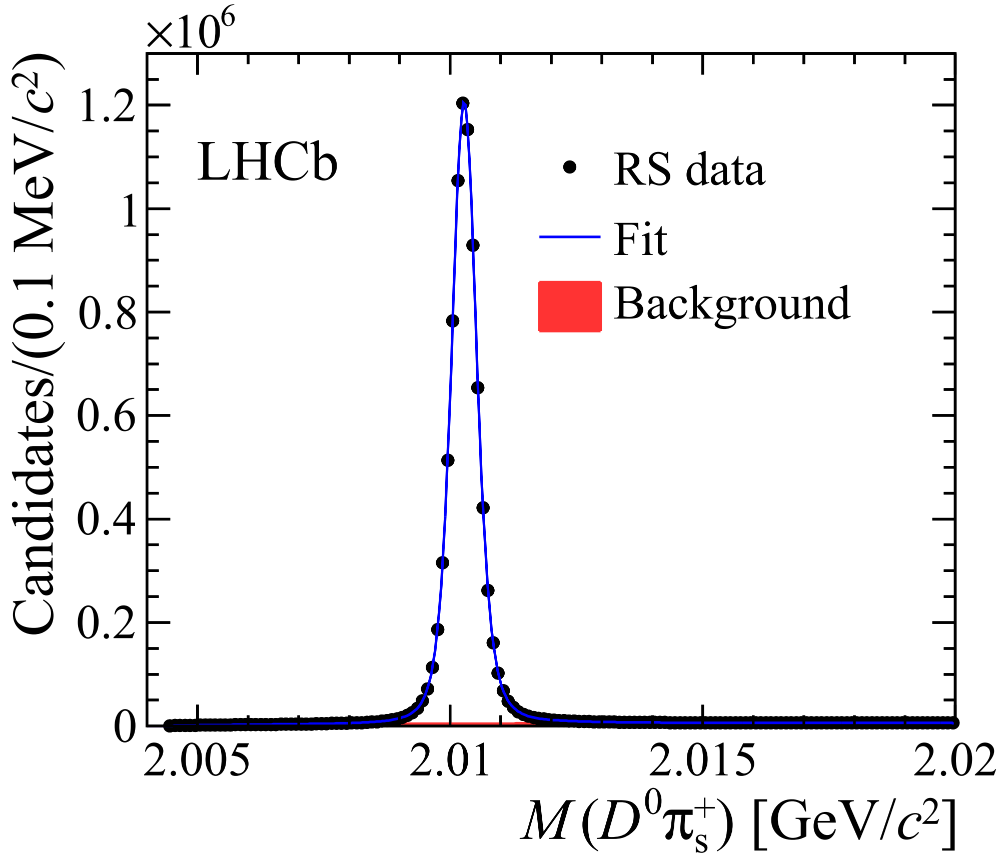
- Implications of LHCb measurements and future prospects [EPJC 73 (2013) 2373]
- LHCb will need to trigger on a very broad range of physics processes:

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 [138]	0.025	0.008	~ 0.003
	$2\beta_s$ ($B_s^0 \to J/\psi\ f_0(980)$)	0.17 [214]	0.045	0.014	~ 0.01
	a_{sl}^s	6.4×10^{-3} [43]	0.6×10^{-3}	0.2×10^{-3}	0.03×10^{-3}
Gluonic penguins	$2\beta_s^{\text{eff}}(B_s^0 \to \phi\phi)$	–	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 [43]	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 penguins	$S_3(B^0 \to K^{*0}\mu^+\mu^-; 1 < q^2 < 6\,\text{GeV}^2/c^4)$	0.08 [67]	0.025	0.008	0.02
	$s_0\ A_{\text{FB}}(B^0 \to K^{*0}\mu^+\mu^-)$	25 % [67]	6 %	2 %	7 %
	$A_{\text{I}}(K\mu^+\mu^-; 1 < q^2 < 6\,\text{GeV}^2/c^4)$	0.25 [76]	0.08	0.025	~ 0.02
	$\mathcal{B}(B^+ \to \pi^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$	25 % [85]	8 %	2.5 %	~ 10 %
Higgs penguins	$\mathcal{B}(B_s^0 \to \mu^+\mu^-)$	1.5×10^{-9} [13]	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–$12°$ [244, 258]	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° [43]	0.6°	0.2°	negligible
Charm CP violation	A_Γ	2.3×10^{-3} [43]	0.40×10^{-3}	0.07×10^{-3}	–
	$\Delta\mathcal{A}_{CP}$	2.1×10^{-3} [18]	0.65×10^{-3}	0.12×10^{-3}	–

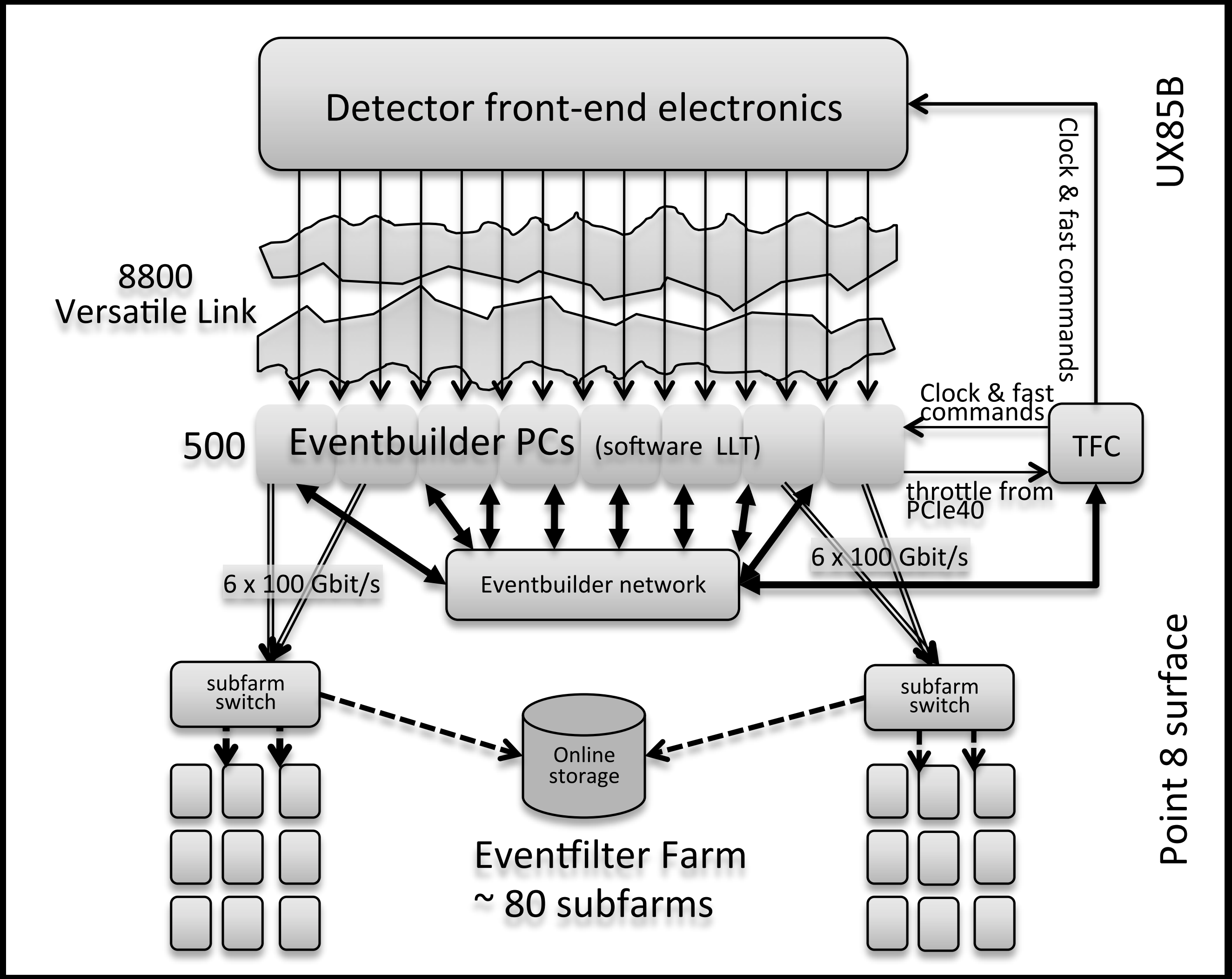
Summary

- During Run 1, LHCb trigger covered a wide range of requirements in a challenging environment
    - high efficiency for the rarest B decay
    - high purity for the largest charm samples
- Run 2 will introduce calibration and alignment ‘in between’ Hlt1 and Hlt2
- The upgrade trigger builds on the Run 1 experience:
    - Readout at full collision rate
    - Full software trigger at 30 MHz
    - Expect doubling of many signal efficiencies

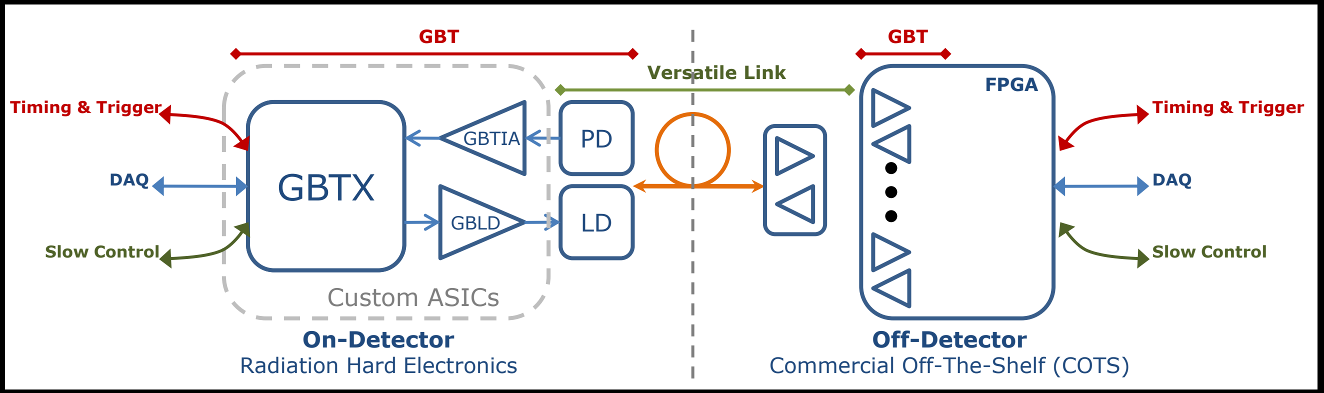




Event Building @ 40 MHz



- 32 Tbit/s
- "All data to the surface"
- Decouple front-end electronics from event builder network
    - Frontend→GBT→PCIe
    - GBT: Rad-hard, integrated into front-end, so no commodity solution possible…
- Buffering in PC memory



Event Building @ 40 MHz

- “COTS” as soon as possible
- O(500) servers for event building
- “Data Center” (“thin” switch, Infiniband/Ethernet/…) instead of “Telecom” (ATCA, “fat” switch)
- Event Filter: O(1000) servers