Triggers in LHCb
past, present, future

Triggering Discoveries in High Energy Physics II
Feb 1, 2018 ✦ Puebla City, MX

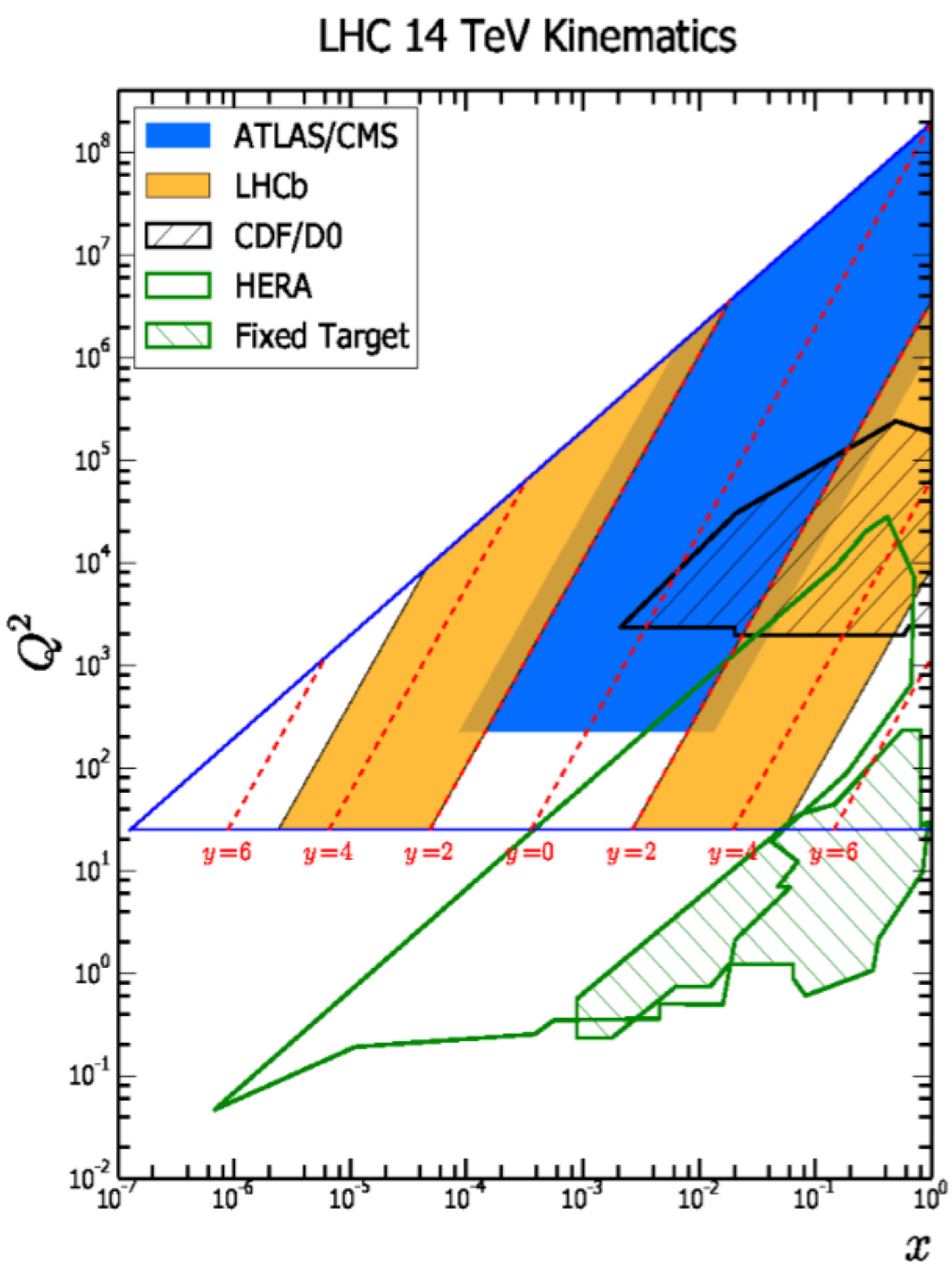
Fedor Ratnikov
on behalf of LHCb Collaboration

SCHOOL OF DATA ANALYSIS

NRU Higher School of Economics,
Yandex School of Data Analysis

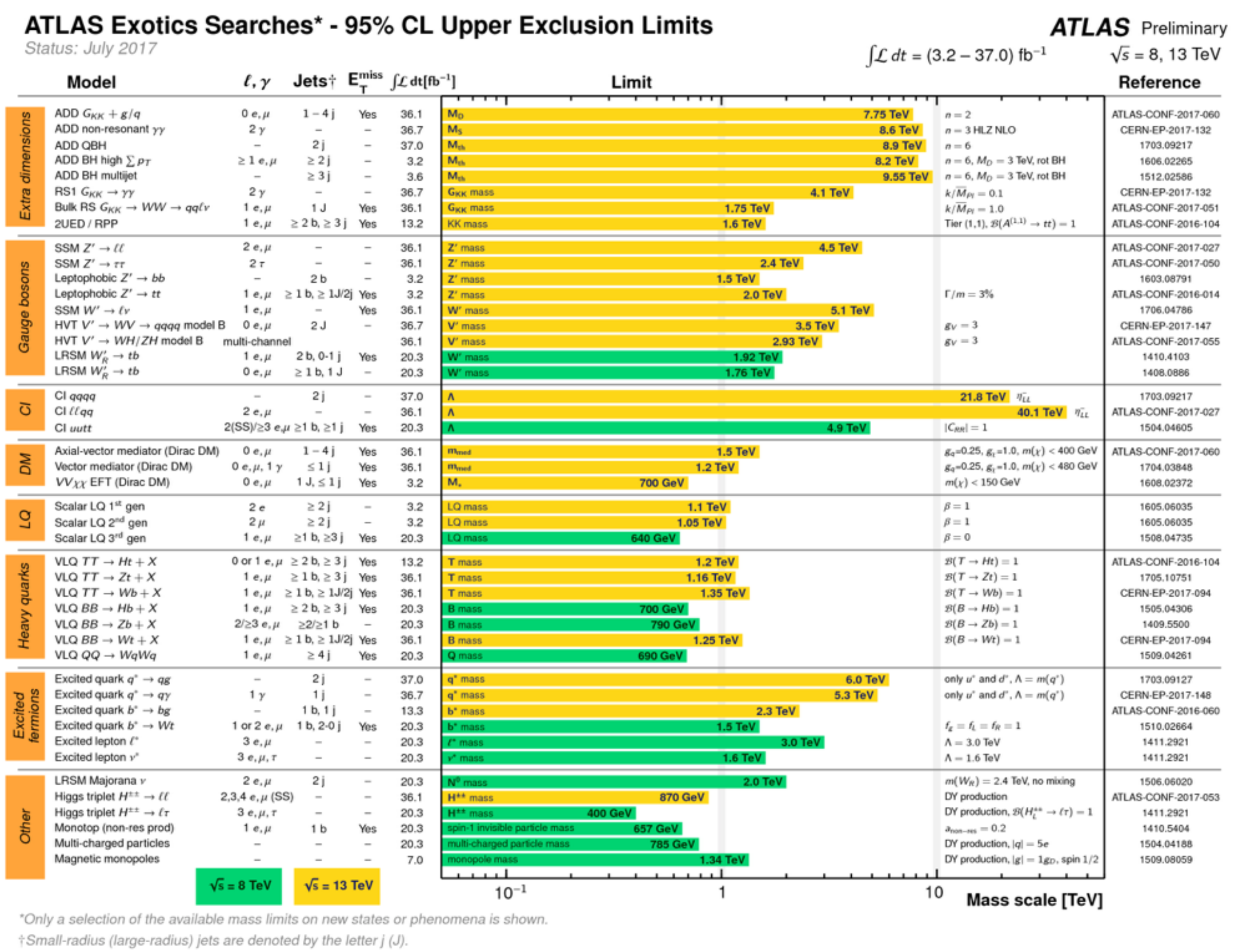
LHCb Physics Program

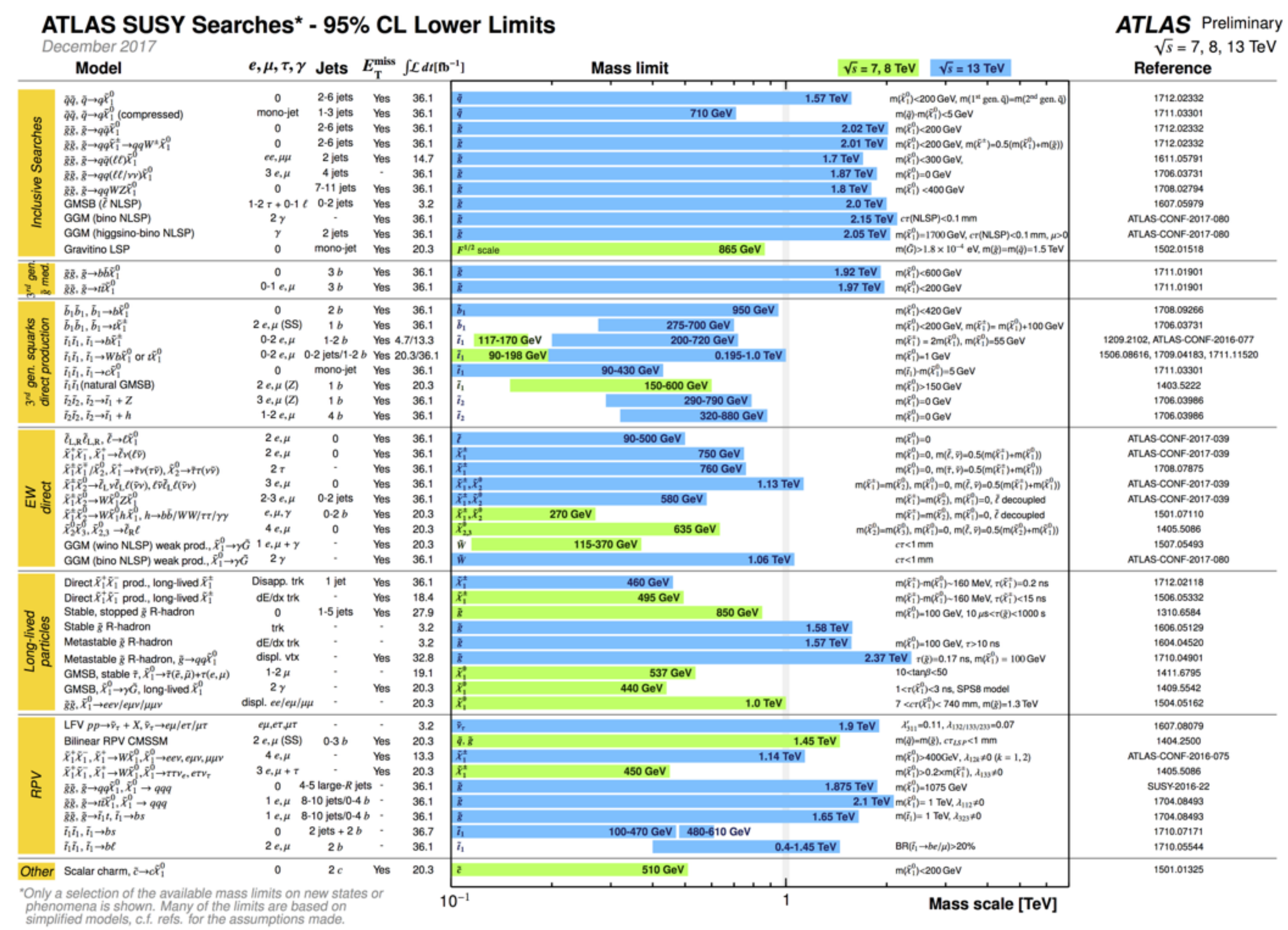
- LHCb – General purpose detector in forward direction
  - Physics program is focused on Heavy Flavour physics
- But LHCb trigger enables and has to enable a wide range of physics
  - Unique rapidity coverage at hadron collider
- Rich program of:
  - Electroweak physics
  - Production and spectroscopy
  - Heavy Ion and fixed target physics
  - Strange physics



New Physics Searches at LHC

- Direct observation
  - comprehensive searches by Atlas and CMS
  - no hints for New Physics yet

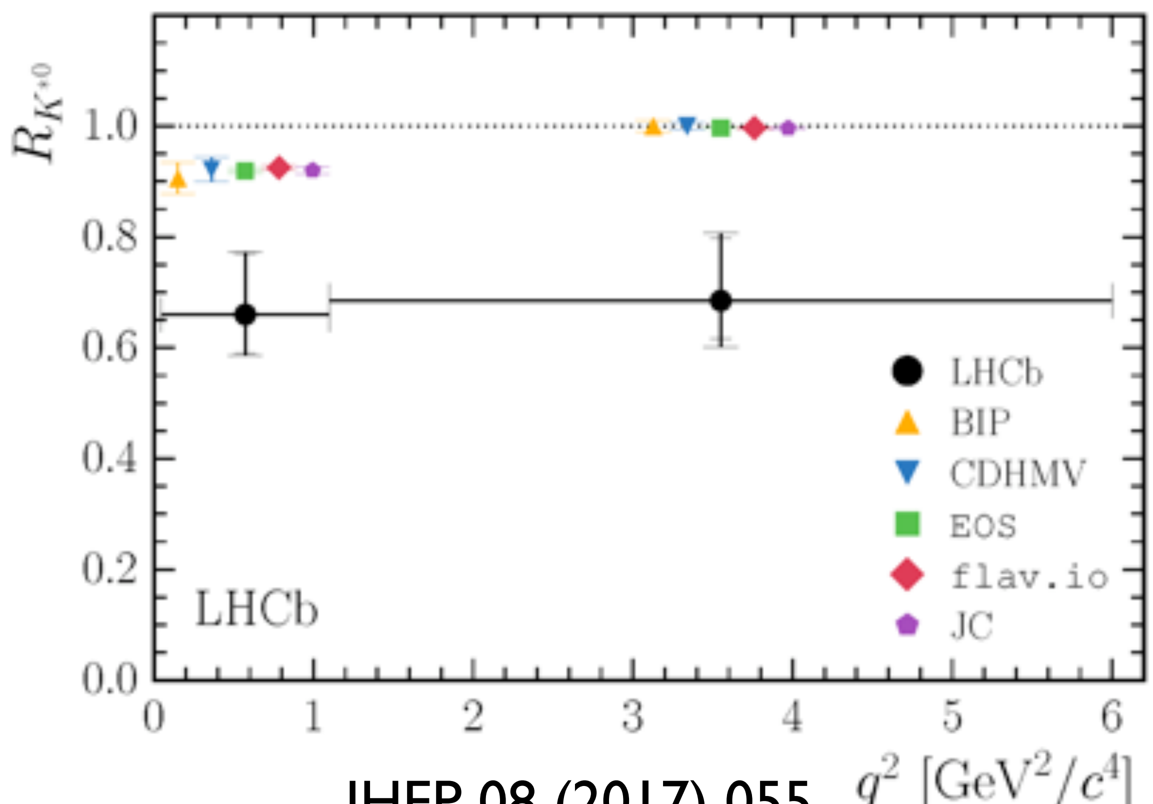




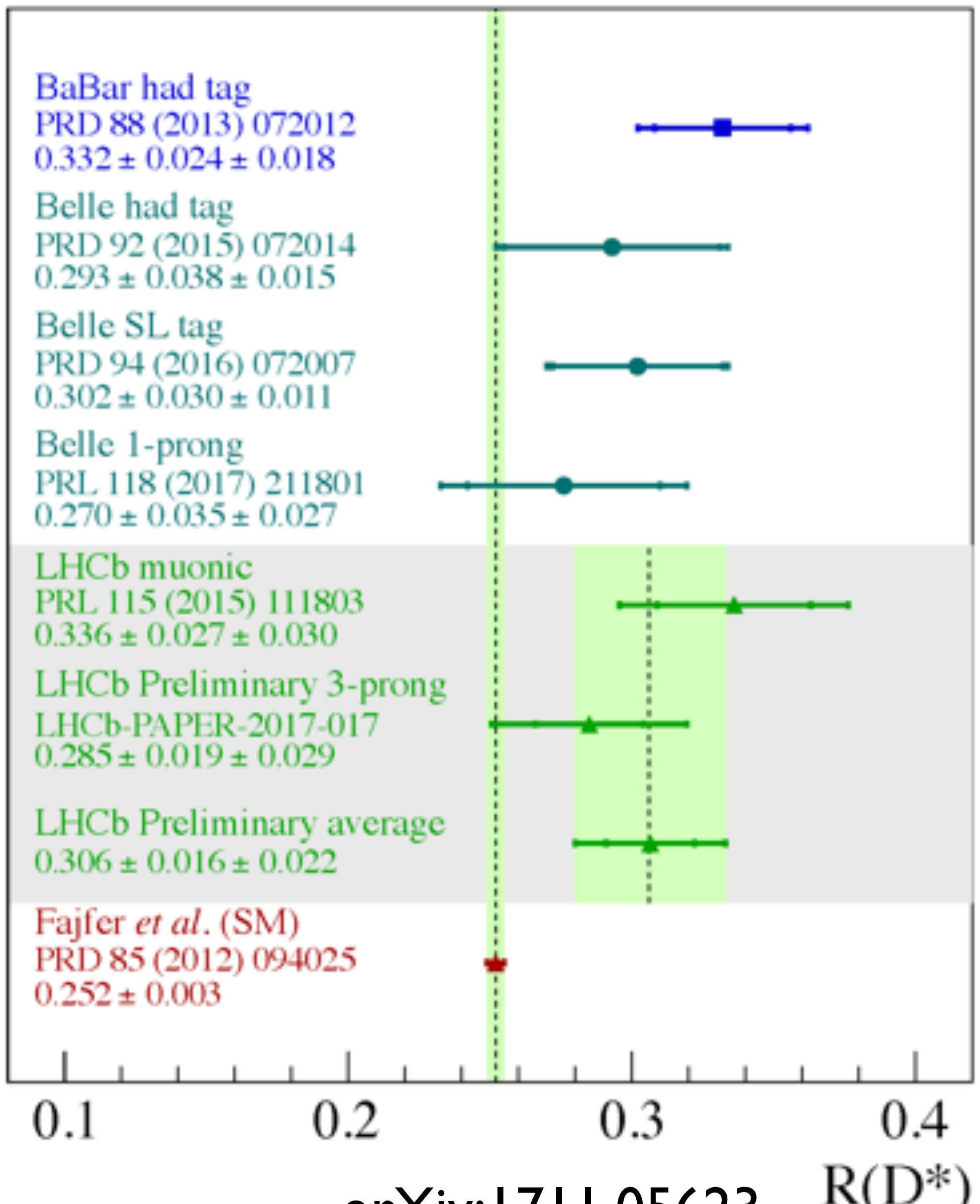
New Physics Searches at LHC

◇ Indirect searches

- ◇ NP driven fine effects in heavy flavour hadrons production and decays
- ◇ Hints are here, need more data for conclusive statements



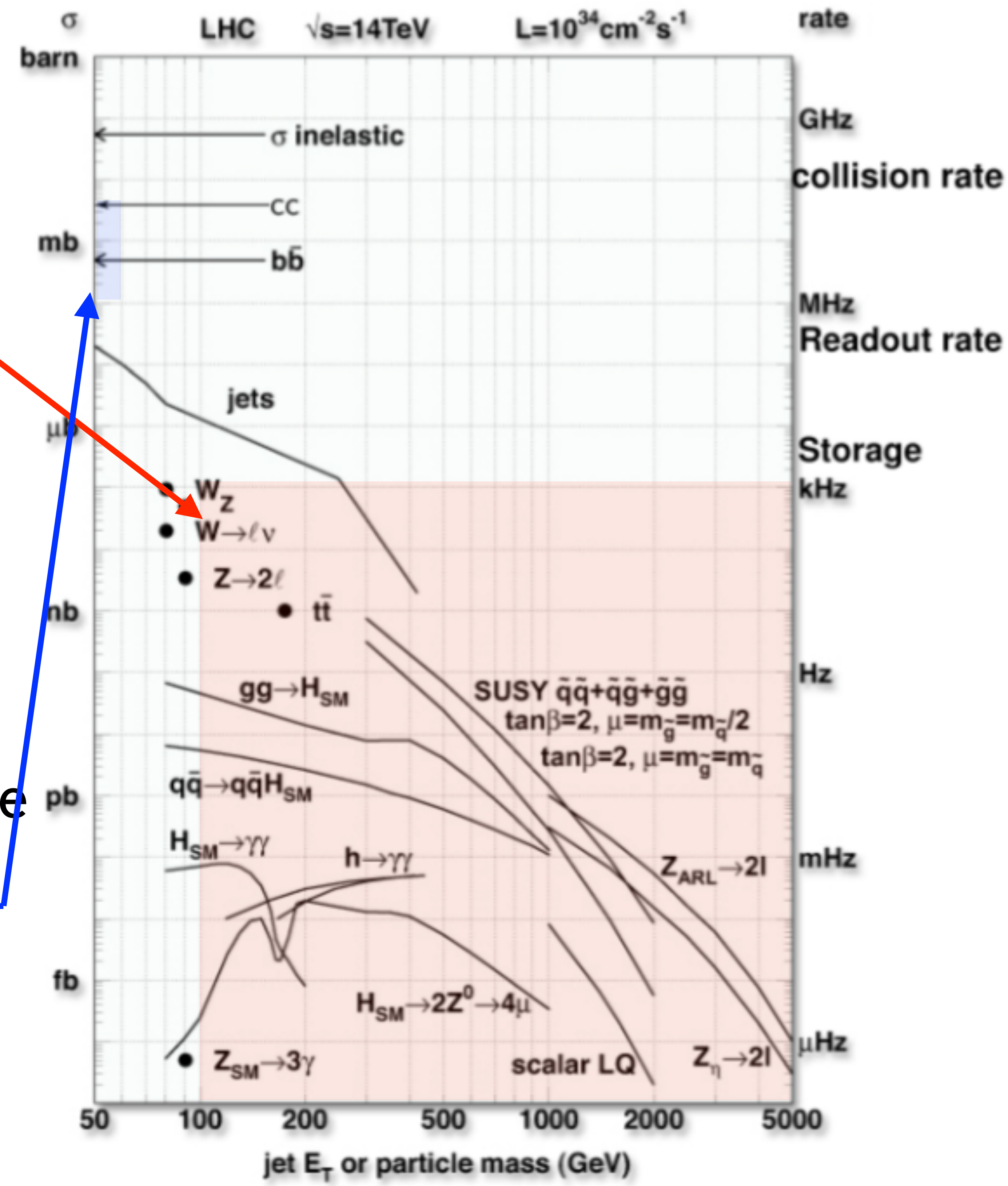
JHEP 08 (2017) 055



arXiv:1711.05623

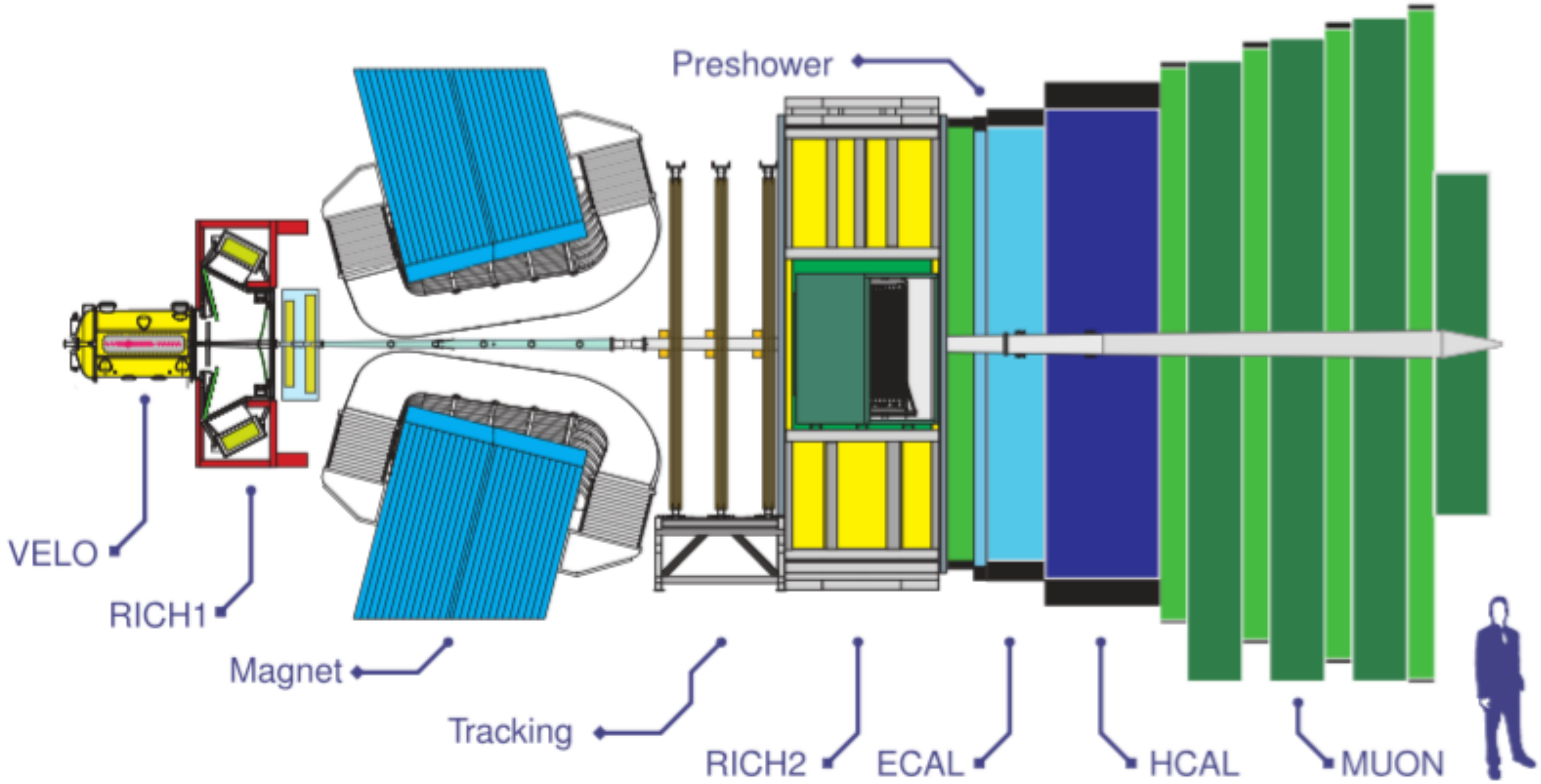
Challenge

- Direct observation of New Physics requires signal rate at level kHz and below in high-p_T region
  - regular trigger concept is applicable
  - CMS / Atlas use case
- Indirect observation of New Physics in parameters of heavy quarks decays requires signal rate at level of MHz in low-pT region
  - new trigger concepts are necessary
  - LHCb use case



LHCb Detector

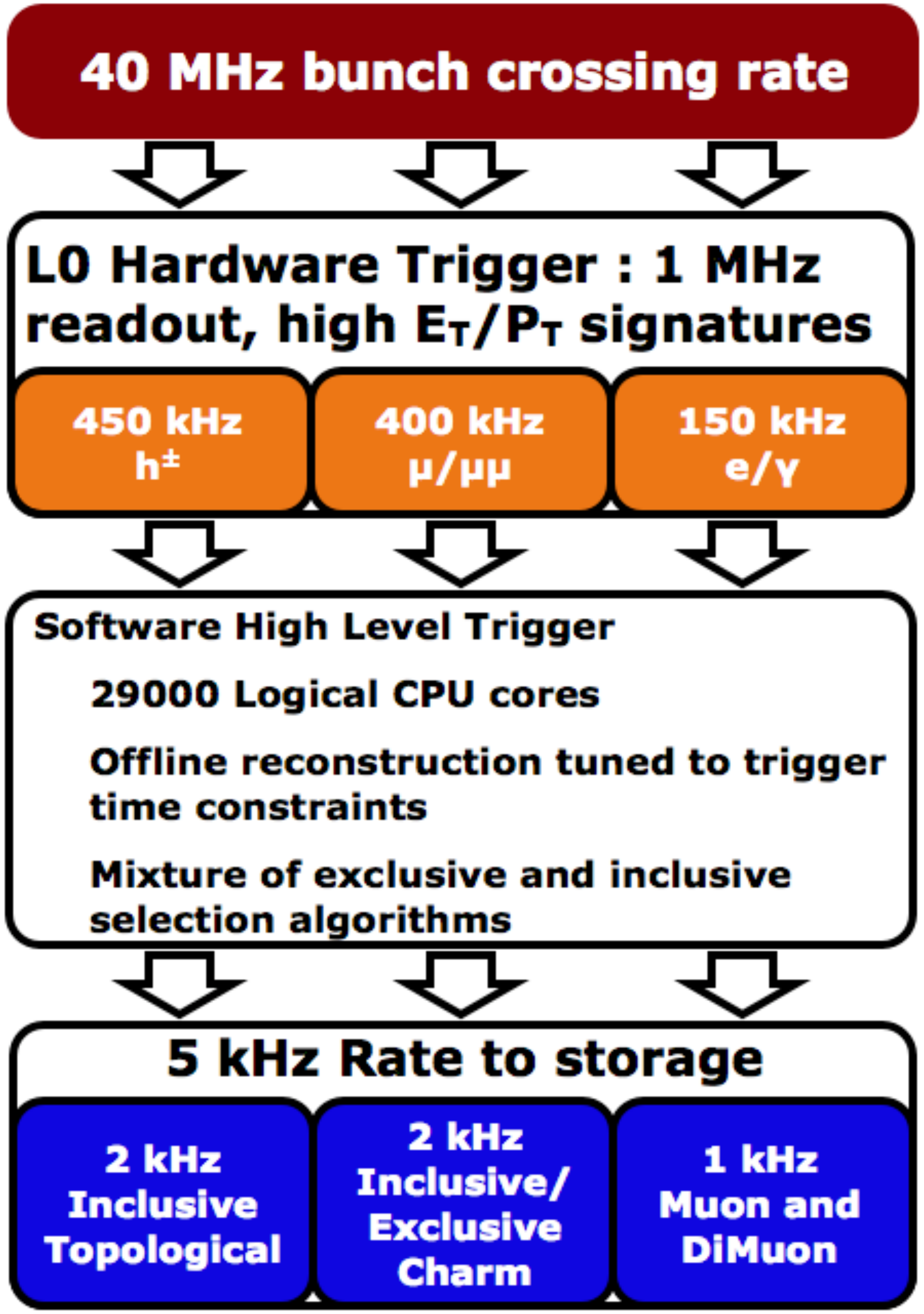
JINST 3 (2008) S08005, Int. J. Mod. Phys. A30 (2015) 1530022



- LHCb was built to study beauty and charm at the LHC
  - Precise particle identification (RICH + MUON)
  - Excellent decay time resolution: ~ 45fs (VELO)
  - High purity + Efficiency with flexible trigger

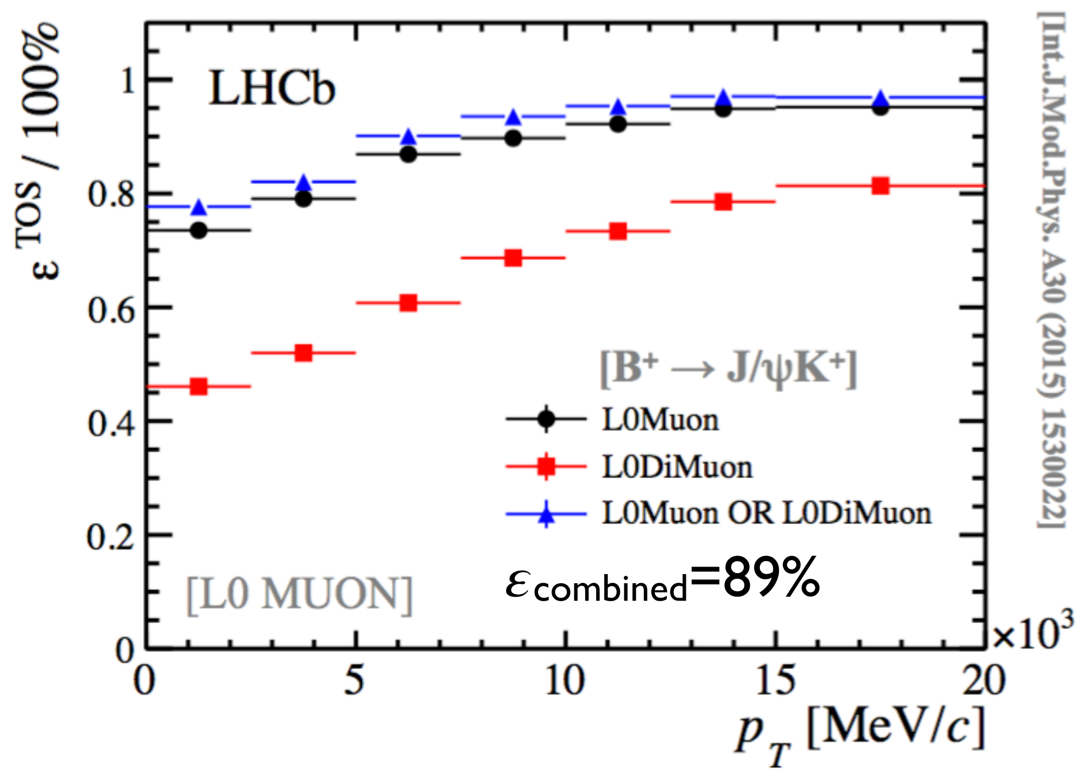
Run 1

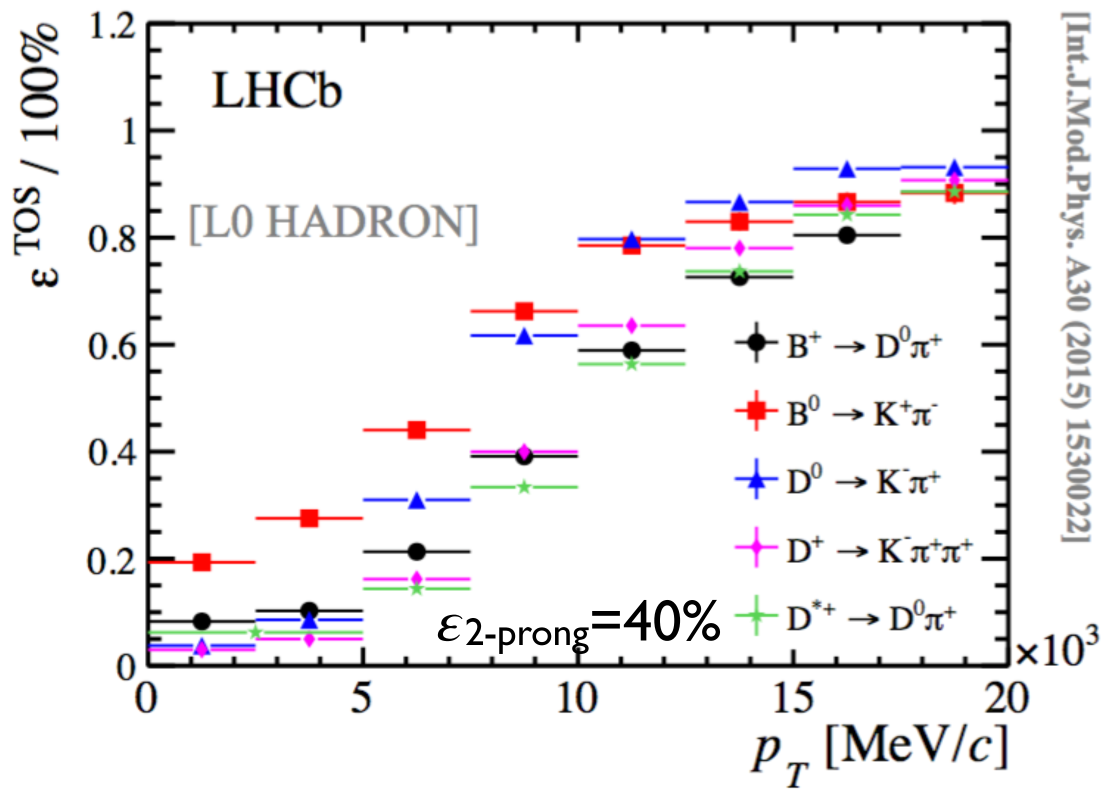
2010-2012, LHCb ran at 7 and 8 TeV @ LHCb $\mathscr{L}_{inst} \sim 4 \cdot 10^{32}$ /cm^2/s

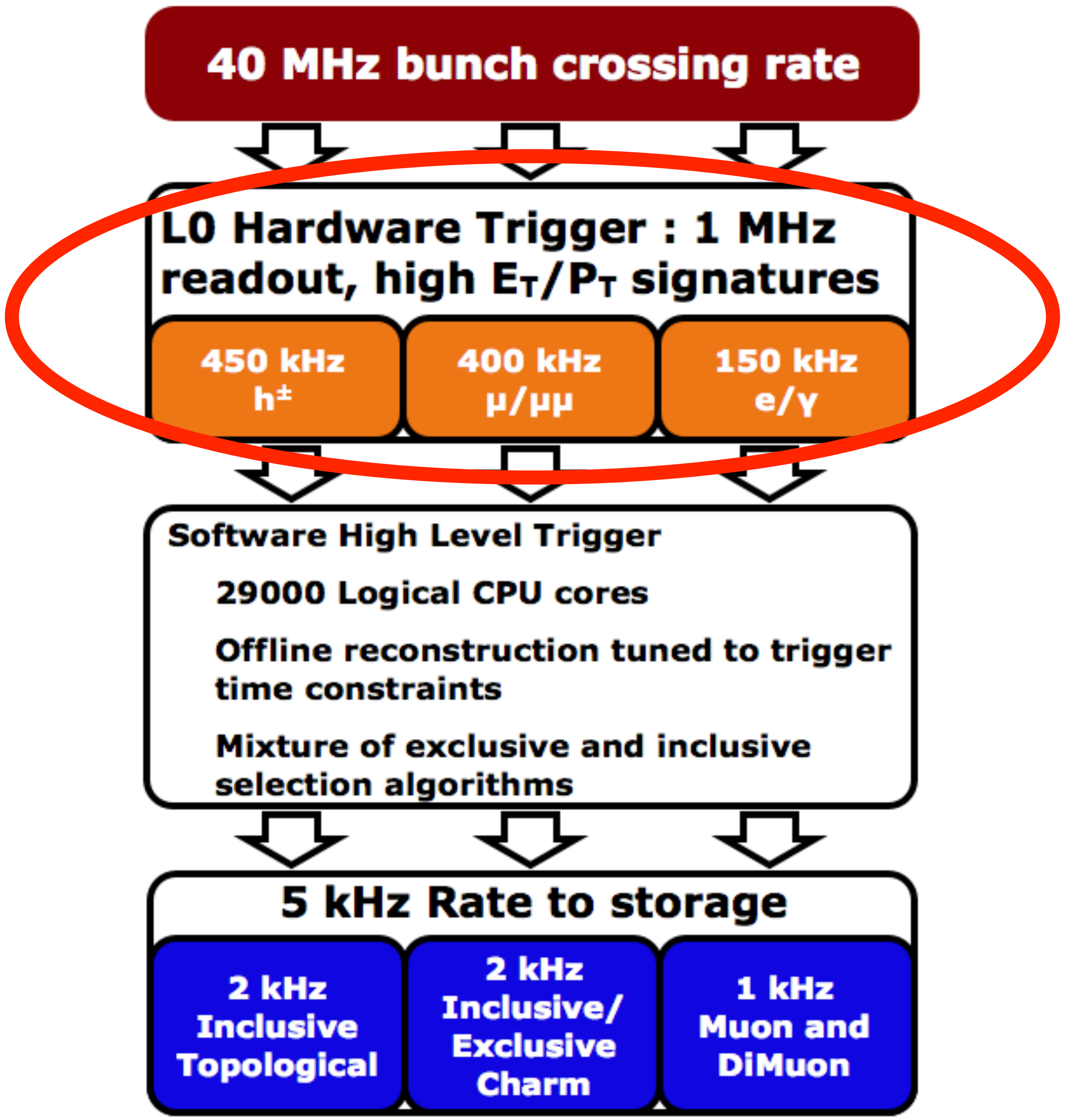

40 MHz bunch crossing rate
L0 Hardware Trigger : 1 MHz readout, high E_T/P_T signatures
450 kHz h±
400 kHz μ/μμ
150 kHz e/γ
Software High Level Trigger
29000 Logical CPU cores
Offline reconstruction tuned to trigger time constraints
Mixture of exclusive and inclusive selection algorithms
5 kHz Rate to storage
2 kHz Inclusive Topological
2 kHz Inclusive/ Exclusive Charm
1 kHz Muon and DiMuon

Hardware Trigger L0

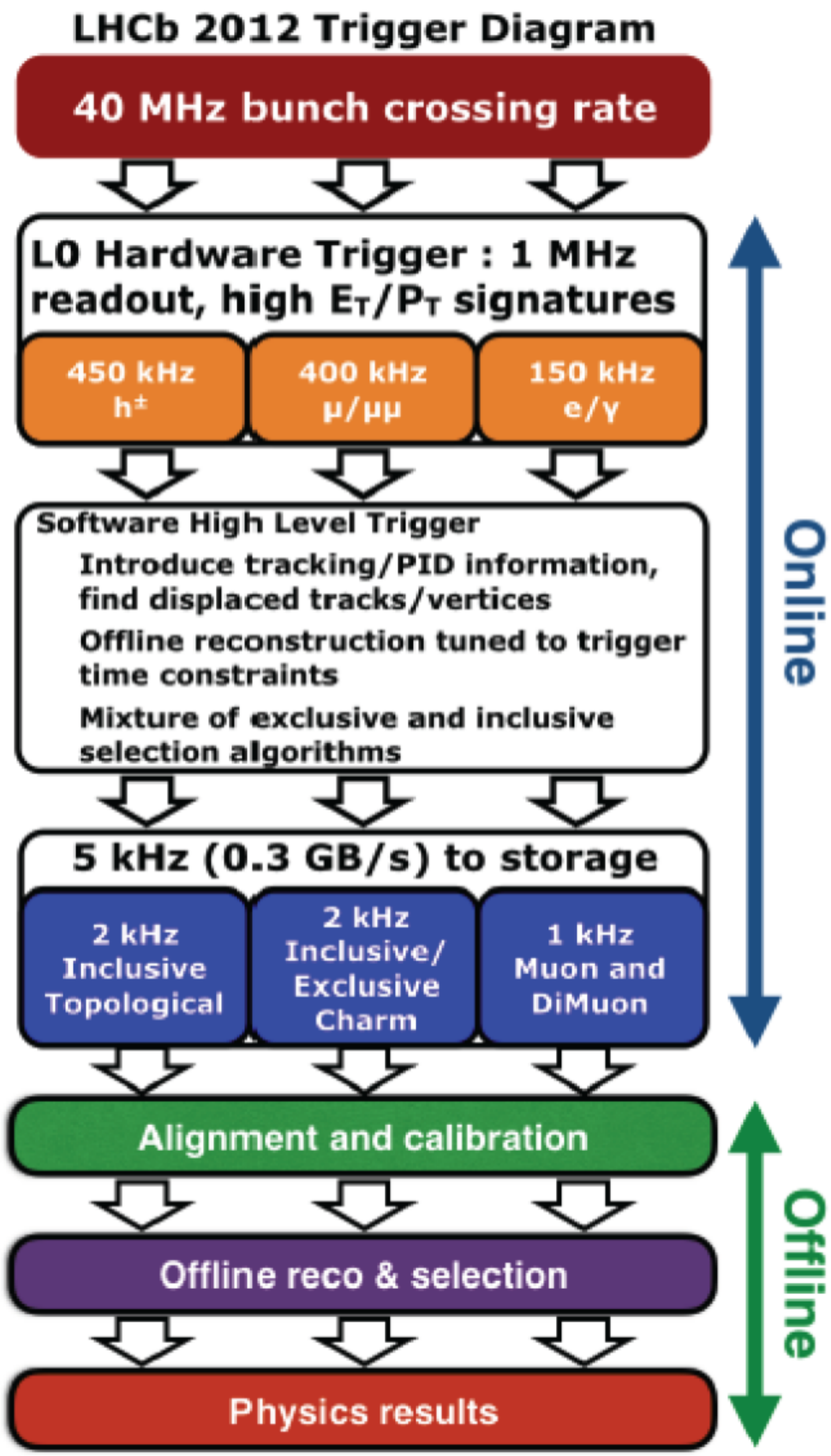
- Reduce from bunch crossing rate to ~1 MHz







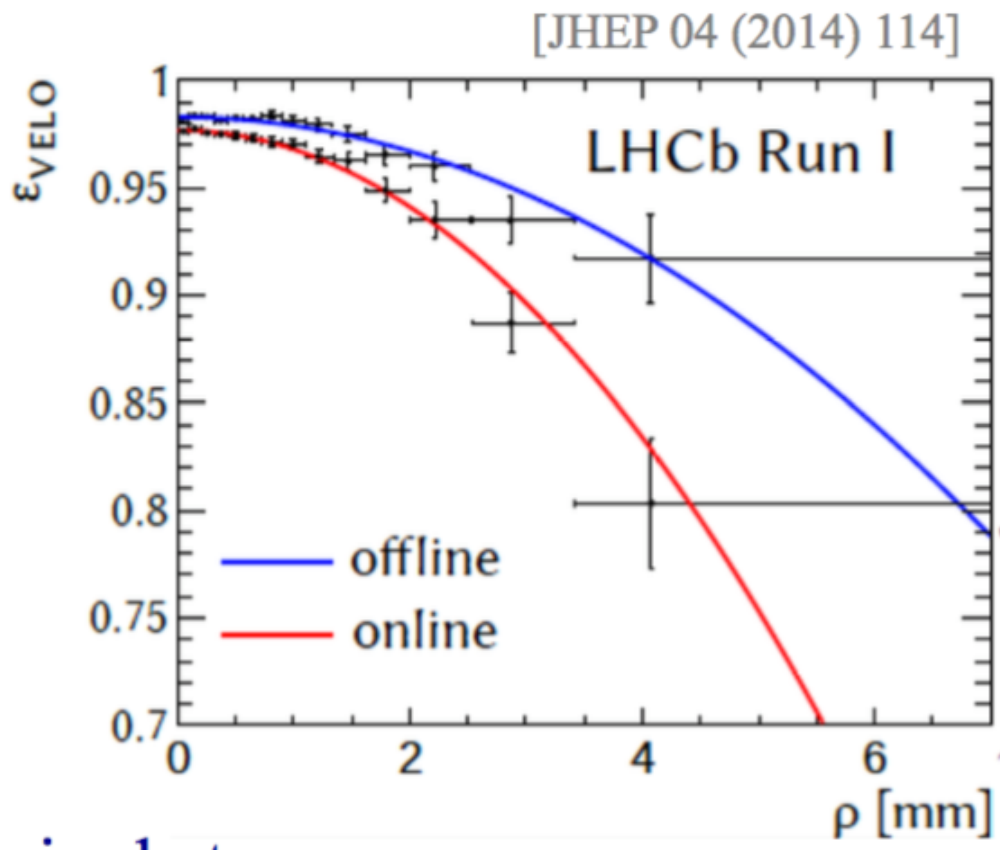
Online vs Offline



- Online: compromise between performance and stringent timing requirements
- Offline: best available performance without stringent timing requirements
- Differences:
  pattern recognition, detector alignment and calibration, candidate selection

Online vs Offline Difference Implication

- Necessary to store full event for the offline processing
  - limits the total number of stored events
    - charm streams are especially affected
- Extra systematics due to differences between online and offline selections

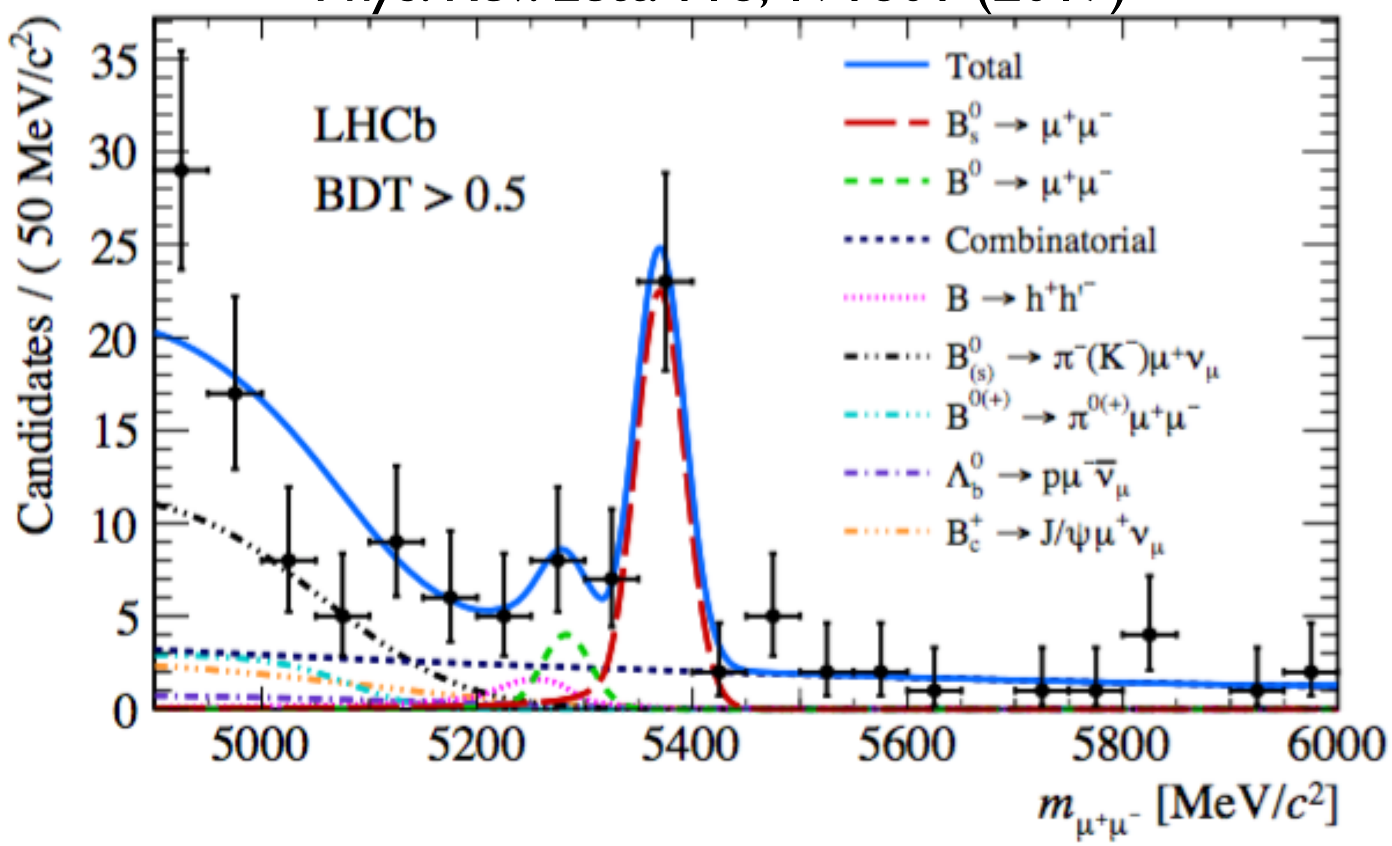


RUN 2

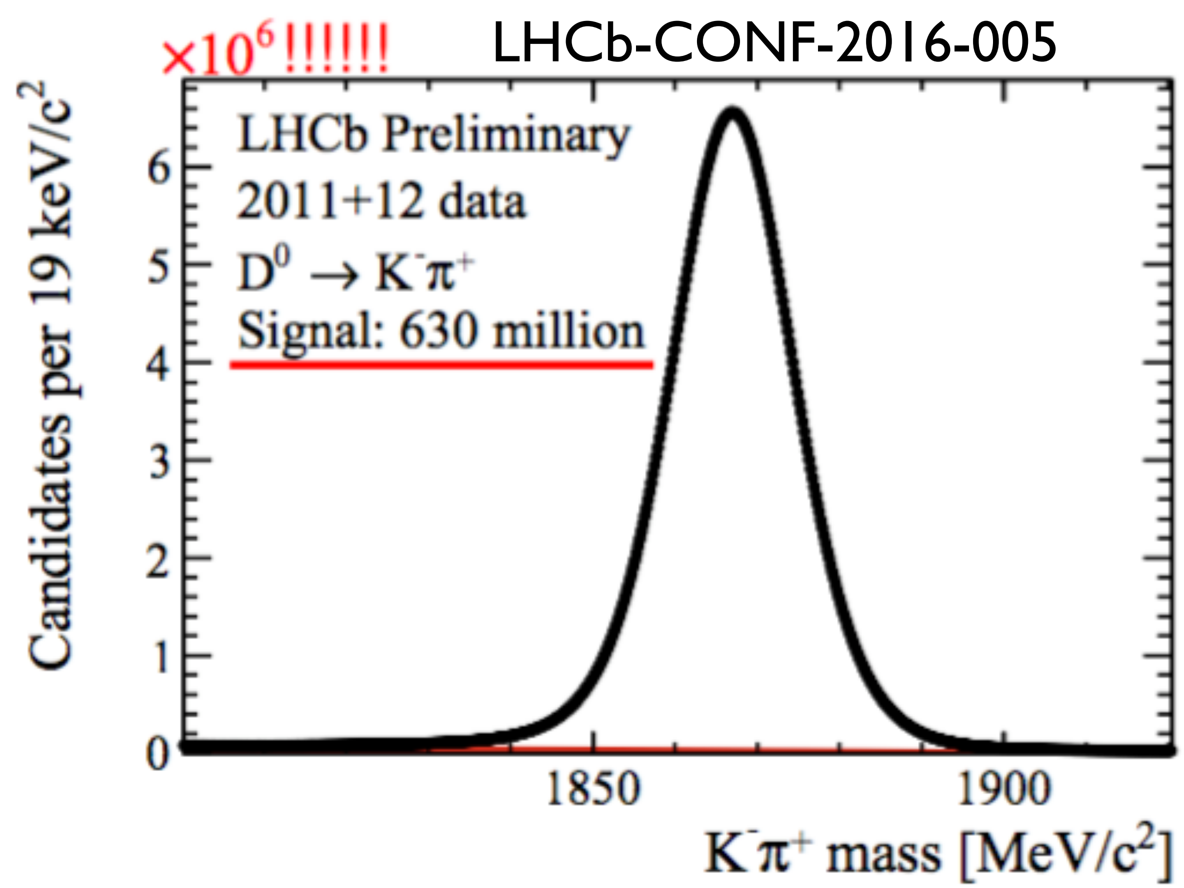
2015-2018, LHCb runs at 13 TeV @ LHCb $\mathscr{L}_{inst} \sim 4 \cdot 10^{32}$ /cm^2/s

Challenges

Phys. Rev. Lett. 118, 191801 (2017)

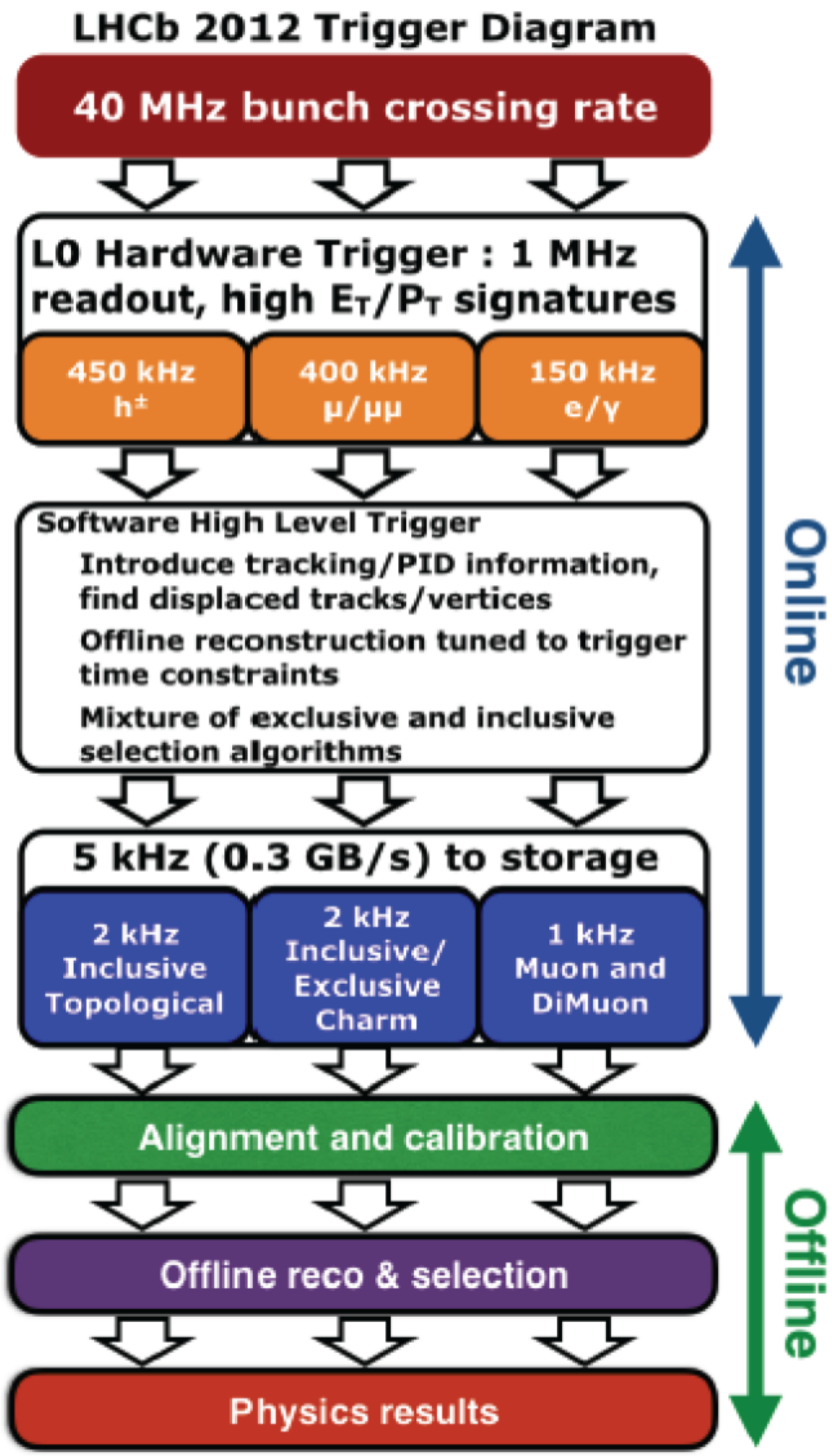


LHCb-CONF-2016-005

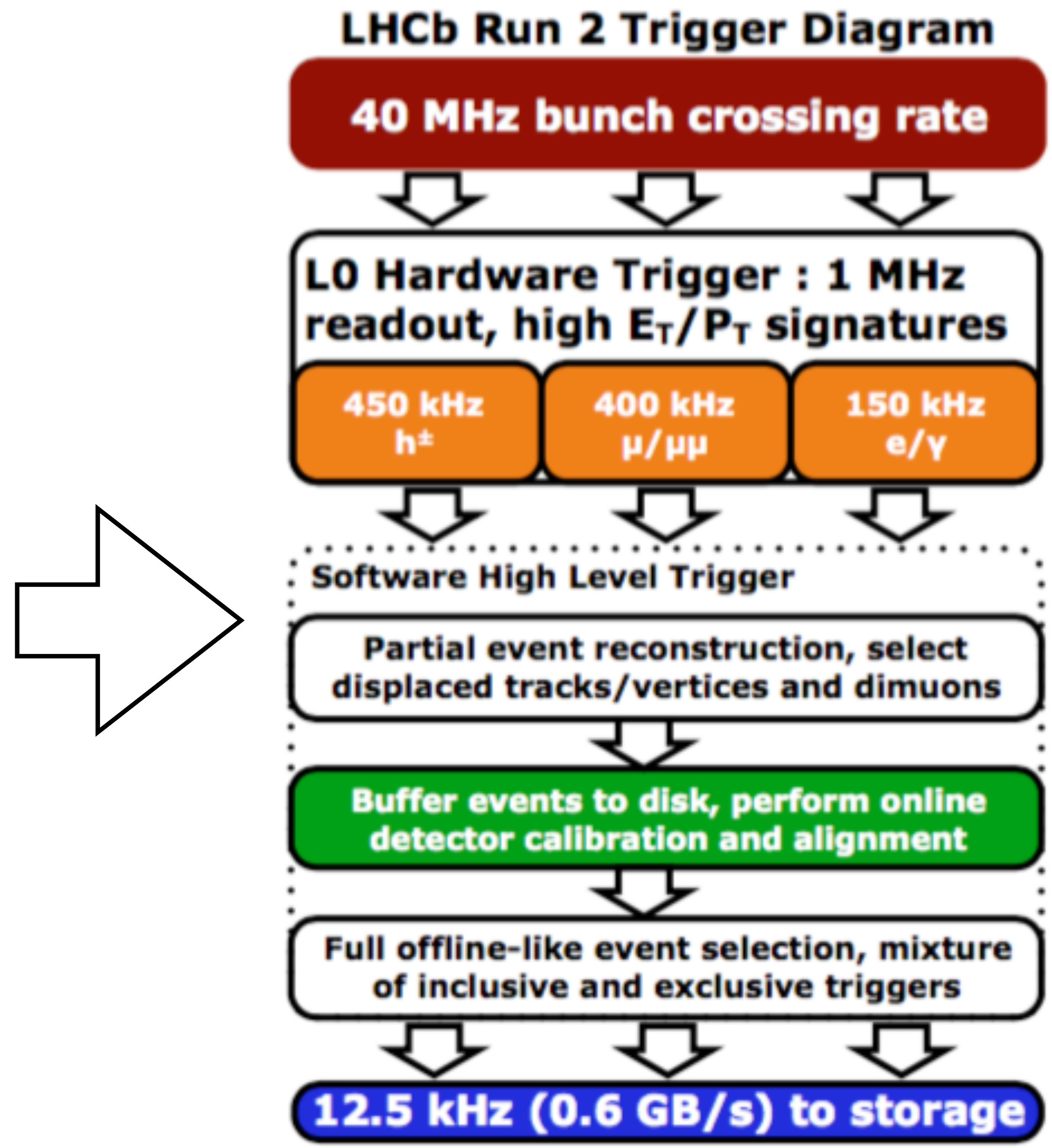


- The LHCb trigger has to cover extremes in data taking
  - High efficiency to collect rare decays like $B^0_s \to \mu^+\mu^-$
  - High purity for enormous charm signals like $D^0 \to K^-\pi^+$
  - Must be flexible to operate in both extremes simultaneously
    - After readout, HLT has access to 100% of event in software

Run 2 LHCb Trigger



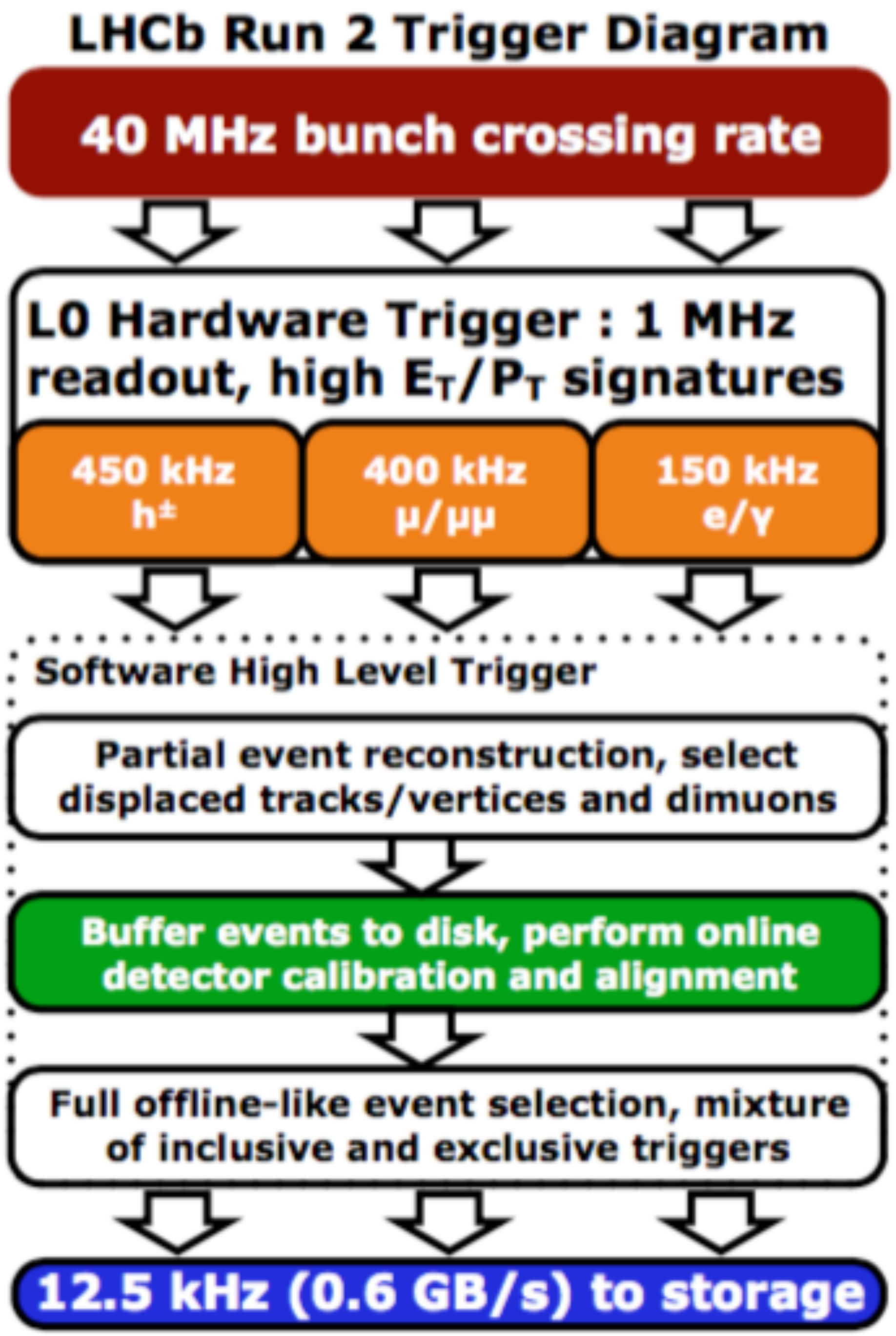
Run I



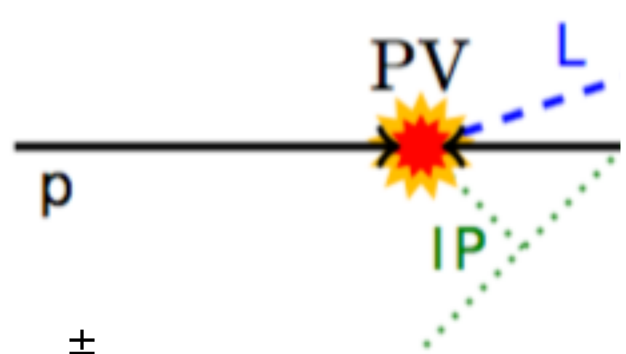
Run 2

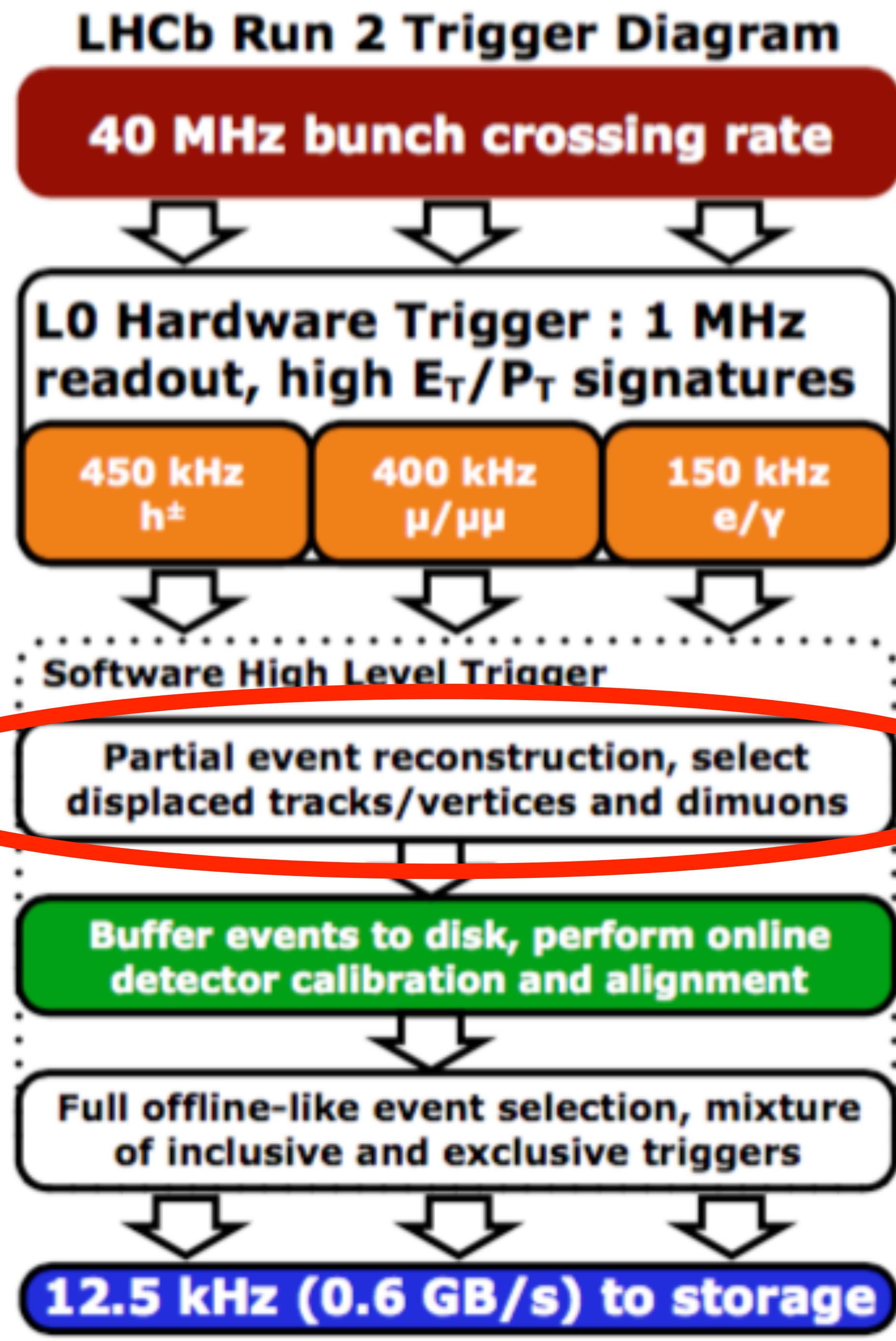
Run 2 LHCb Trigger

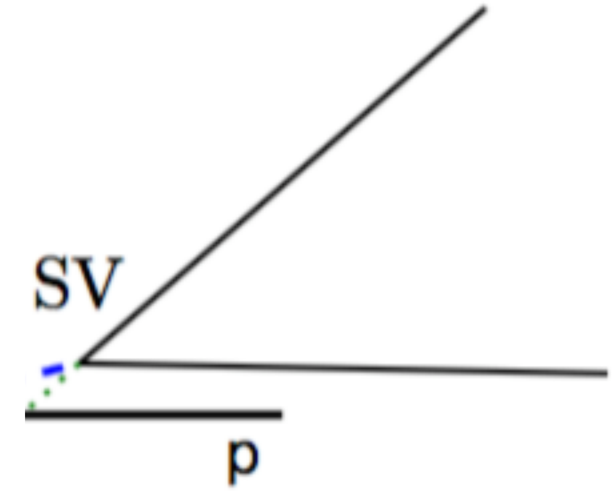
- Three levels, each has more time than the last:
  - Level-0 trigger buys time to readout the detector with Calo, Muon p_T thresholds: 40MHz→1MHz
  - Events built at 1MHz, sent to HLT farm (~27000 physical cores)
  - HLT1 has 40× more time, fast tracking followed by inclusive selections 1MHz → 100kHz
  - HLT2 has 400× more time than L0: Full event reconstruction, inclusive + exclusive selections using whole detector →12.5kHz
- Flexibility comes from software-centric HLT design



HLT1 Topological Trigger



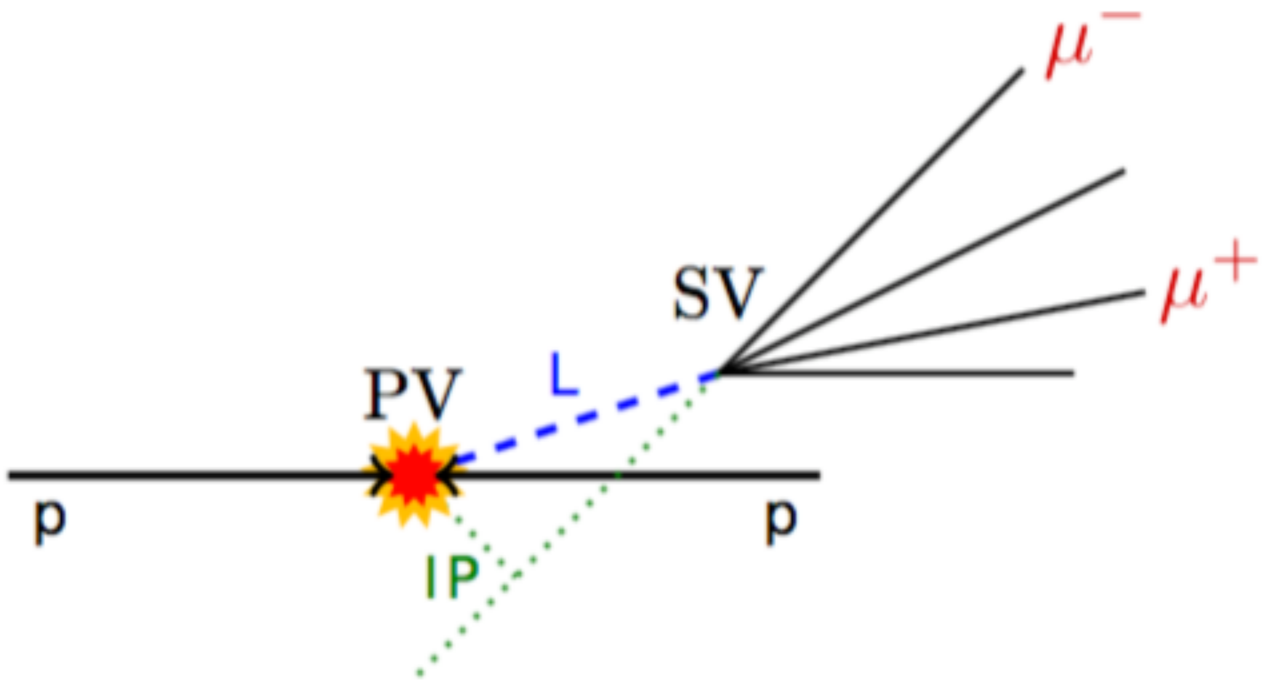


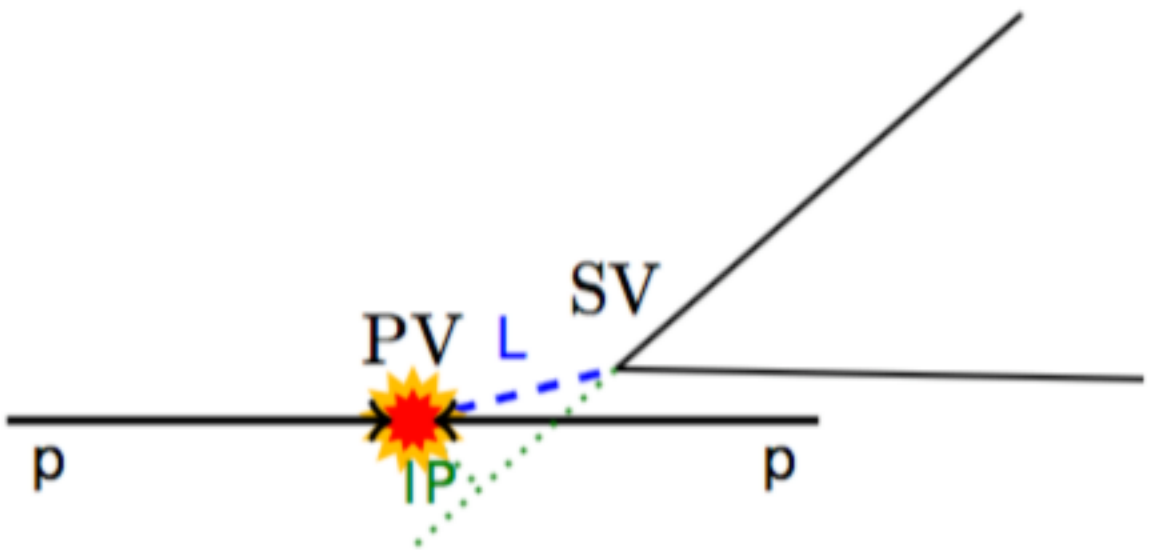


- $B^{\pm}$ mass ~5.28 GeV, dau … appreciable GeV)
- τ ~1.6 ps, Flight distanc … ance ~4 mm
- Important signature: De … condary' charm from B → J/ψX , J/ψ …

- Underlying HLT1 st
  - Fast reconstructio … nal Muon ID
  - Inclusive triggerin
    - ~ 100kHz outpu

HLT1 Topological Trigger



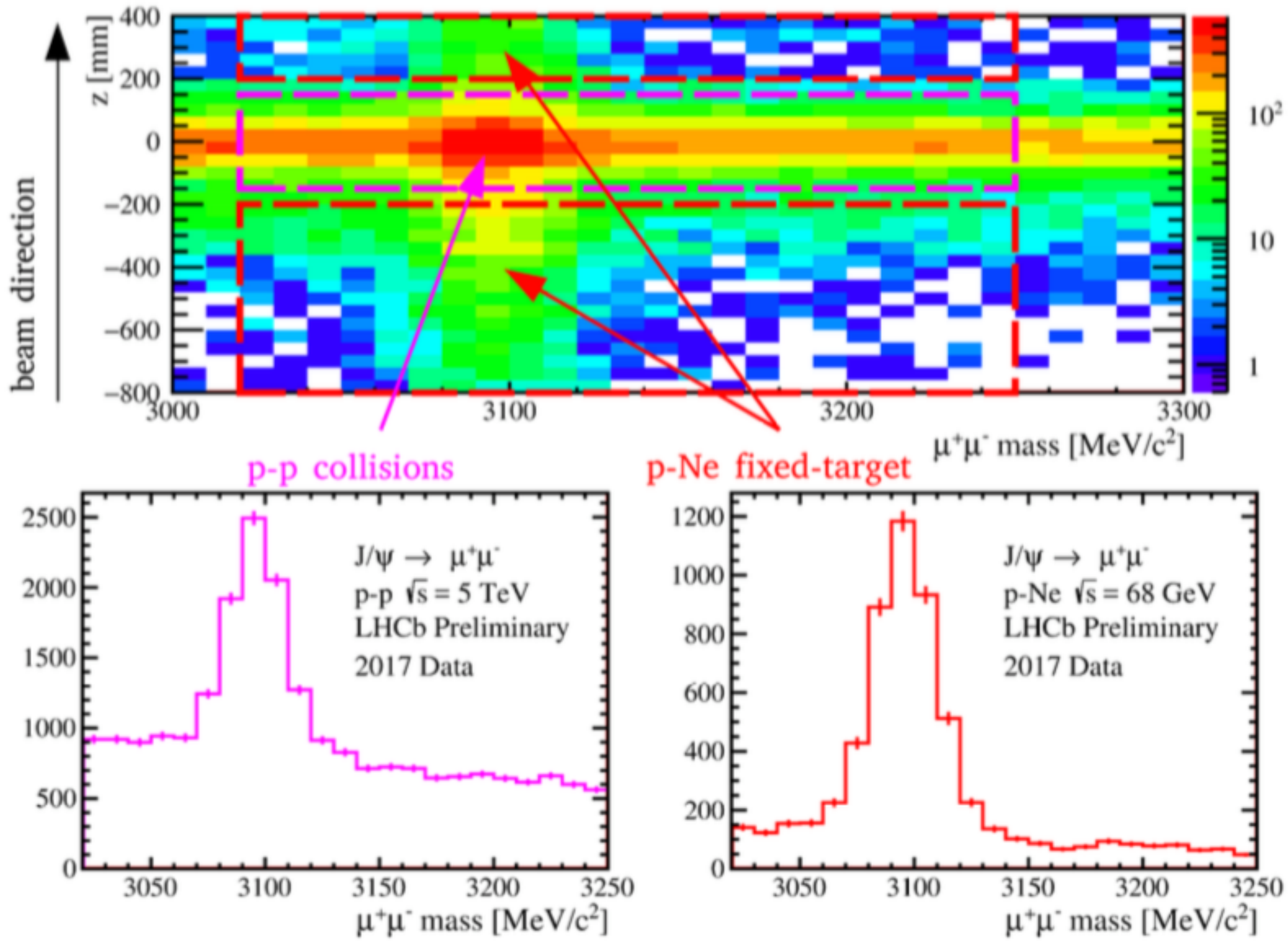


- $B^{\pm}$ mass ~5.28 GeV, daughter p_T O (1 GeV)
- τ ~1.6 ps, Flight distance ~1 cm
- Important signature: Detached muons from $B \to J/\psi X$, $J/\psi \to \mu\mu$

- D^0 mass ~ 1.86 GeV, appreciable daughter p_T
- τ ~0.4 ps, Flight distance ~4 mm
- Also produced as 'secondary' charm from B decays.

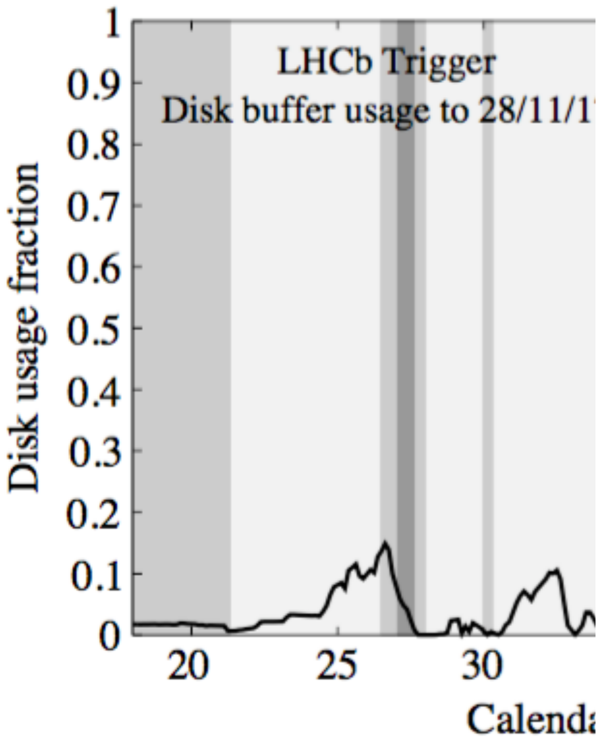
- Underlying HLT1 strategy:
  - Fast reconstruction: Primary Vertices, High p_T tracks, optional Muon ID
  - Inclusive triggering using MVAs on 1&2-track signatures:
    - ~ 100kHz output rate

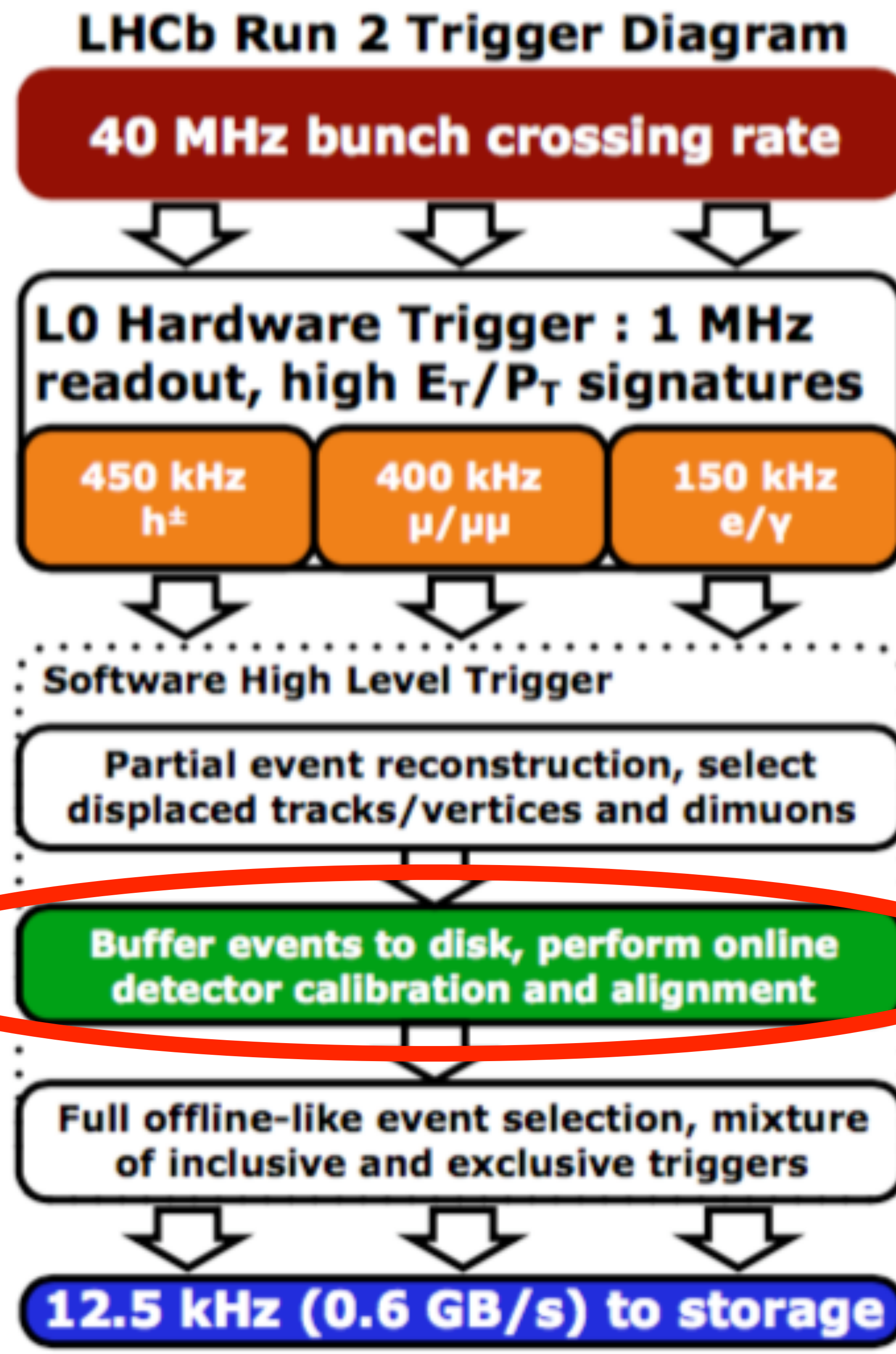
Flexibility on Software HLT1

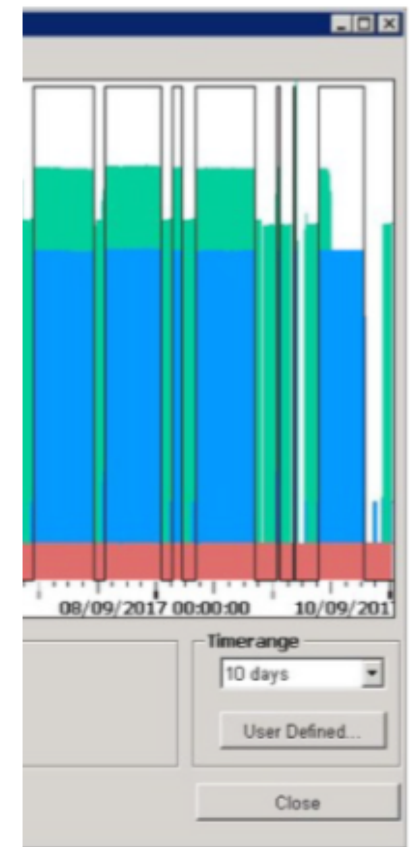


- 5 TeV LHC run in the end of 2017
  - both p-p collision and p-Ne fixed target at the same time
    - clear J/ψ signals in both pp and p-Ne at first software trigger stage

Disk Buffer

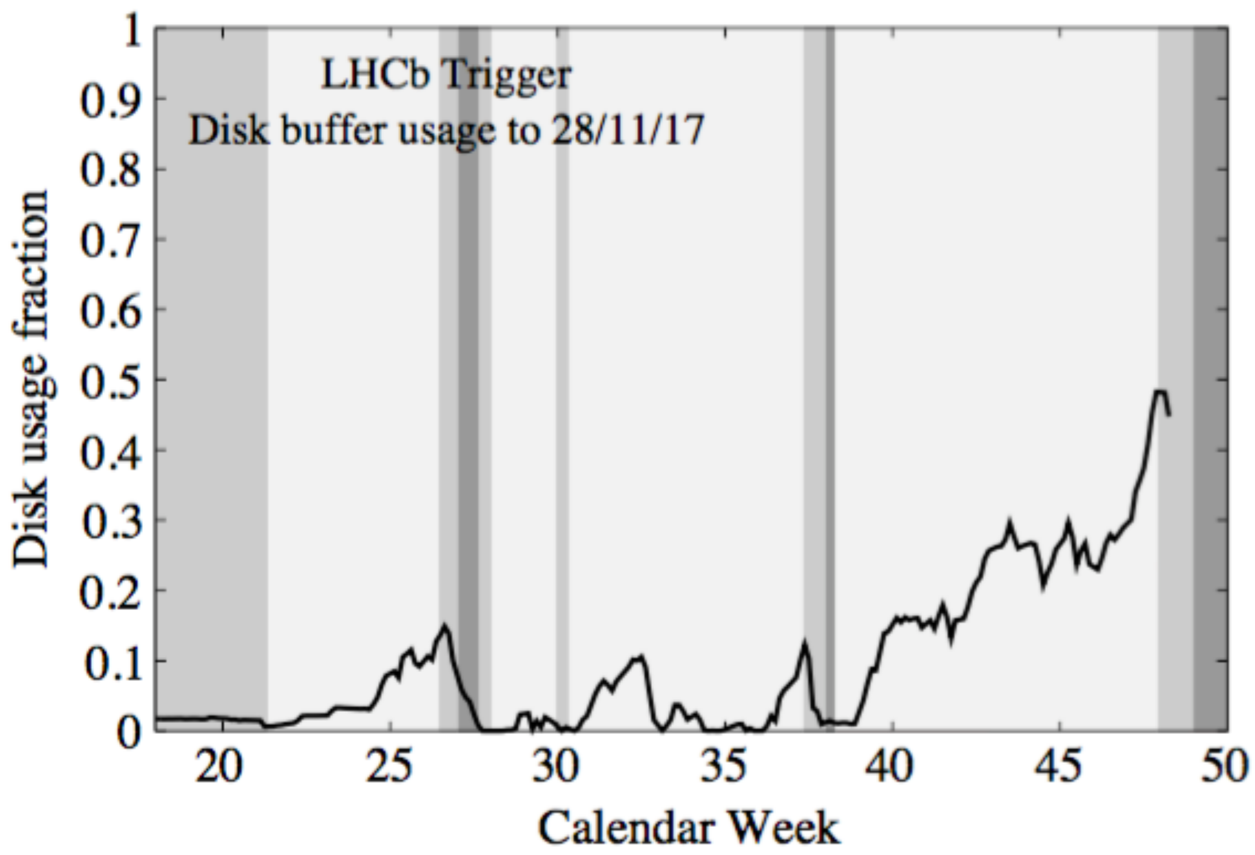


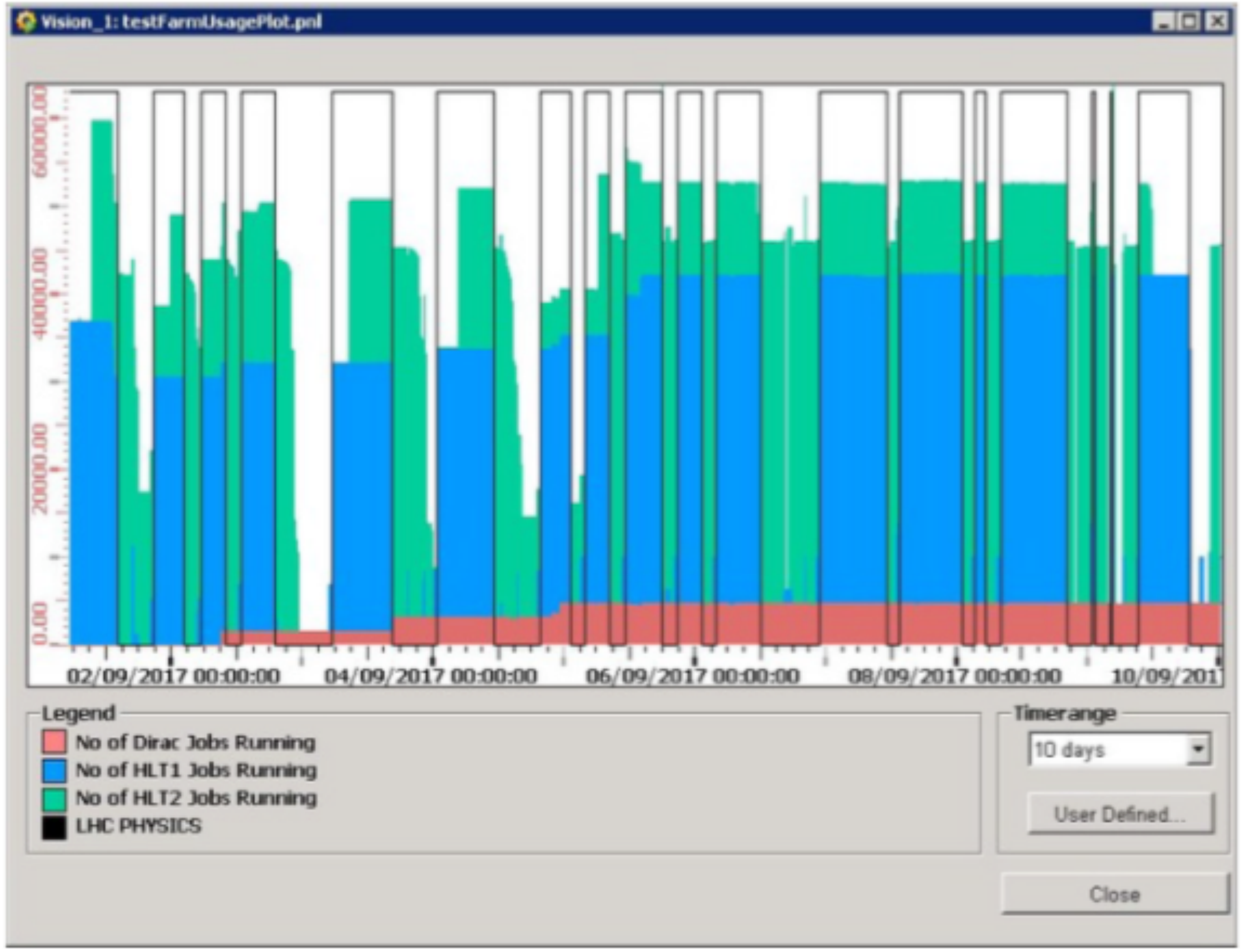




- HLT Farm is off-th y
- HLT1 accepted eve contingency
- HLT2 throughput i running
- Effectively doubles , excess used for simulation
- Asynchronous HLT

◇ This needs man

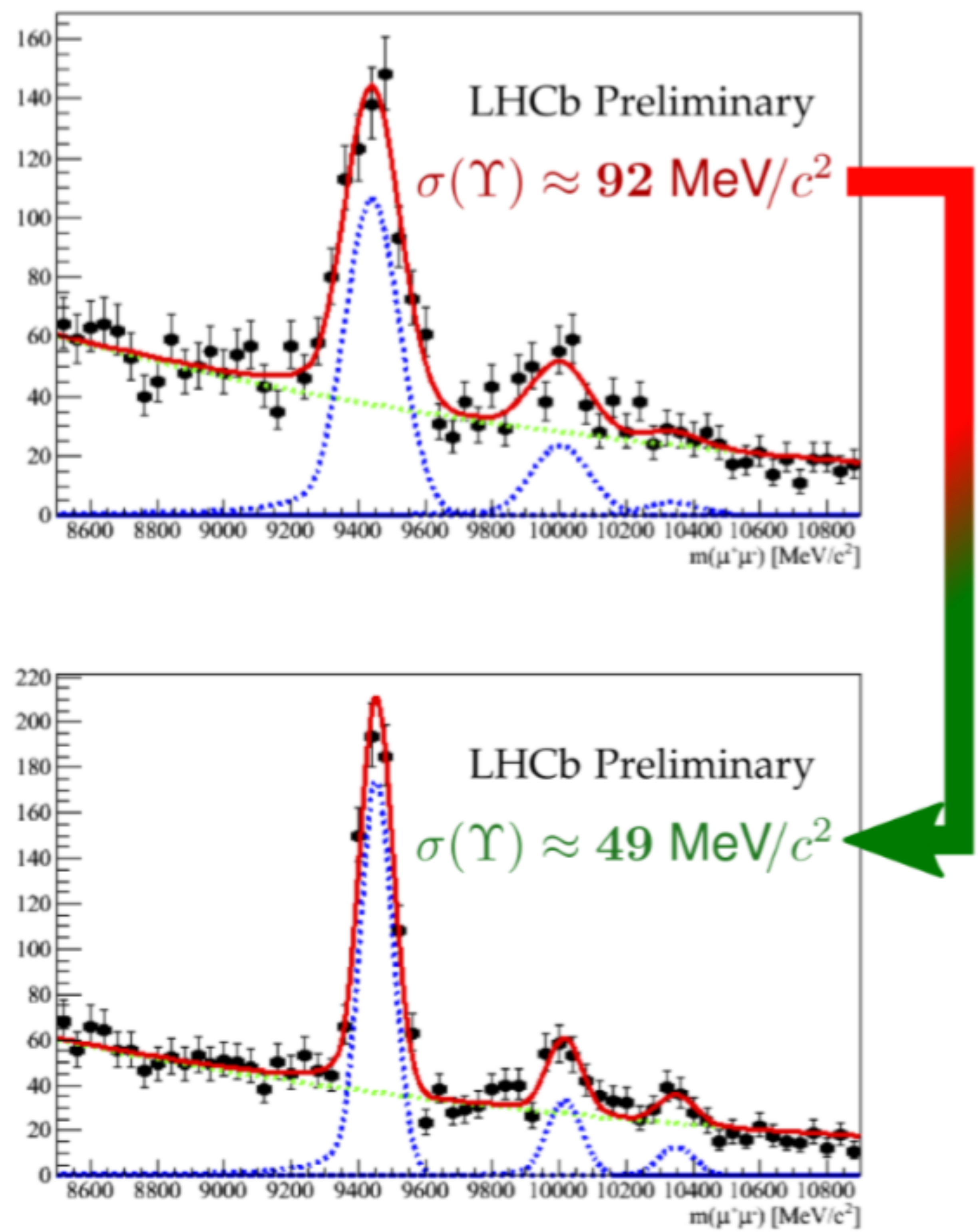




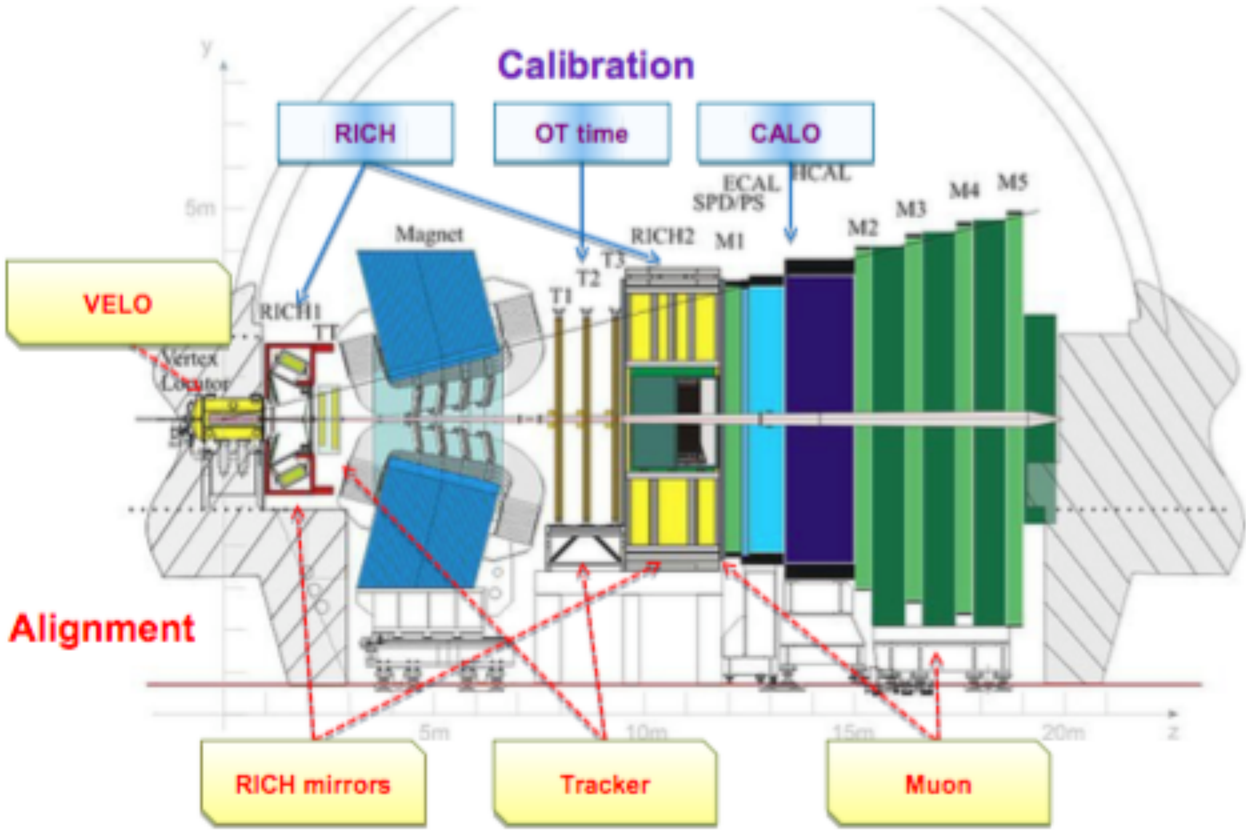
- Off-the shelf servers: 11PB disk capacity
- Accepts HLT1 events written to the disk in-fill at 100kHz: 2 week contingency
- HLT2 throughput in-fill is 30kHz out of fill 90kHz when HLT1 isn't running
- Effectively doubles trigger CPU capacity: farm is used twice for HLT, excess used for simulation

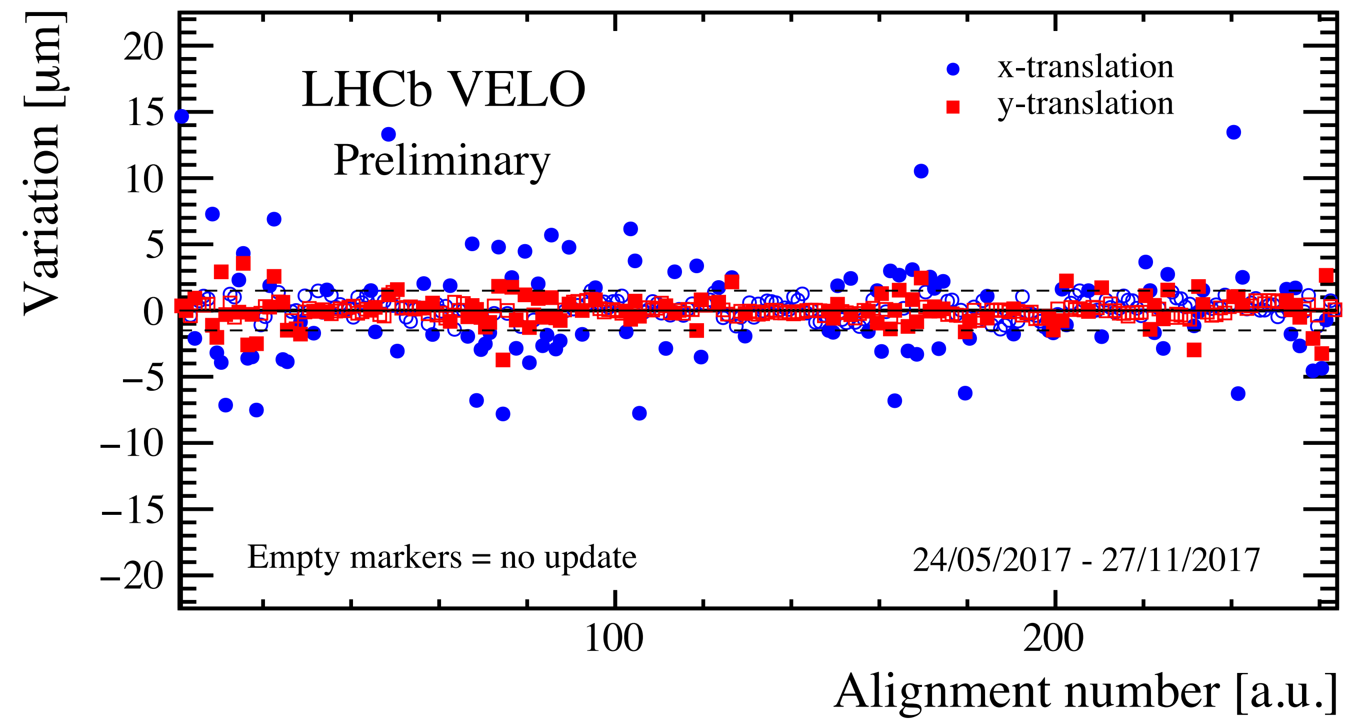
Real-time Alignment & Calibration

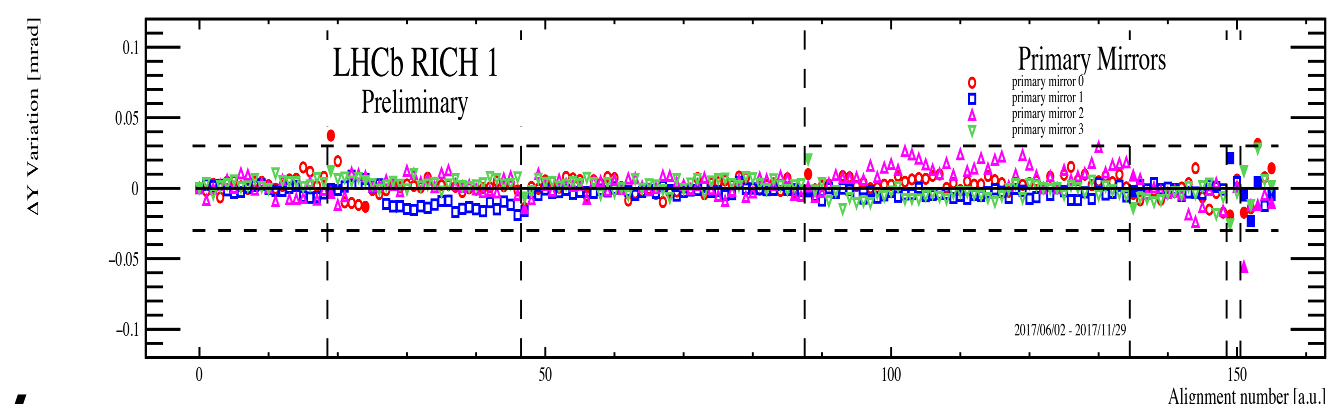
- With Run 2 signal rates, efficient & pure output requires full reconstruction at HLT2
  - online selections ≡ offline selections
    - Reduces systematic uncertainties and workload for analysts
  - Alignment and calibration of full detector in the trigger needed
- While HLT1 is written to disk, alignment & calibration tasks run



Detector Aligned and Ready for HLT2



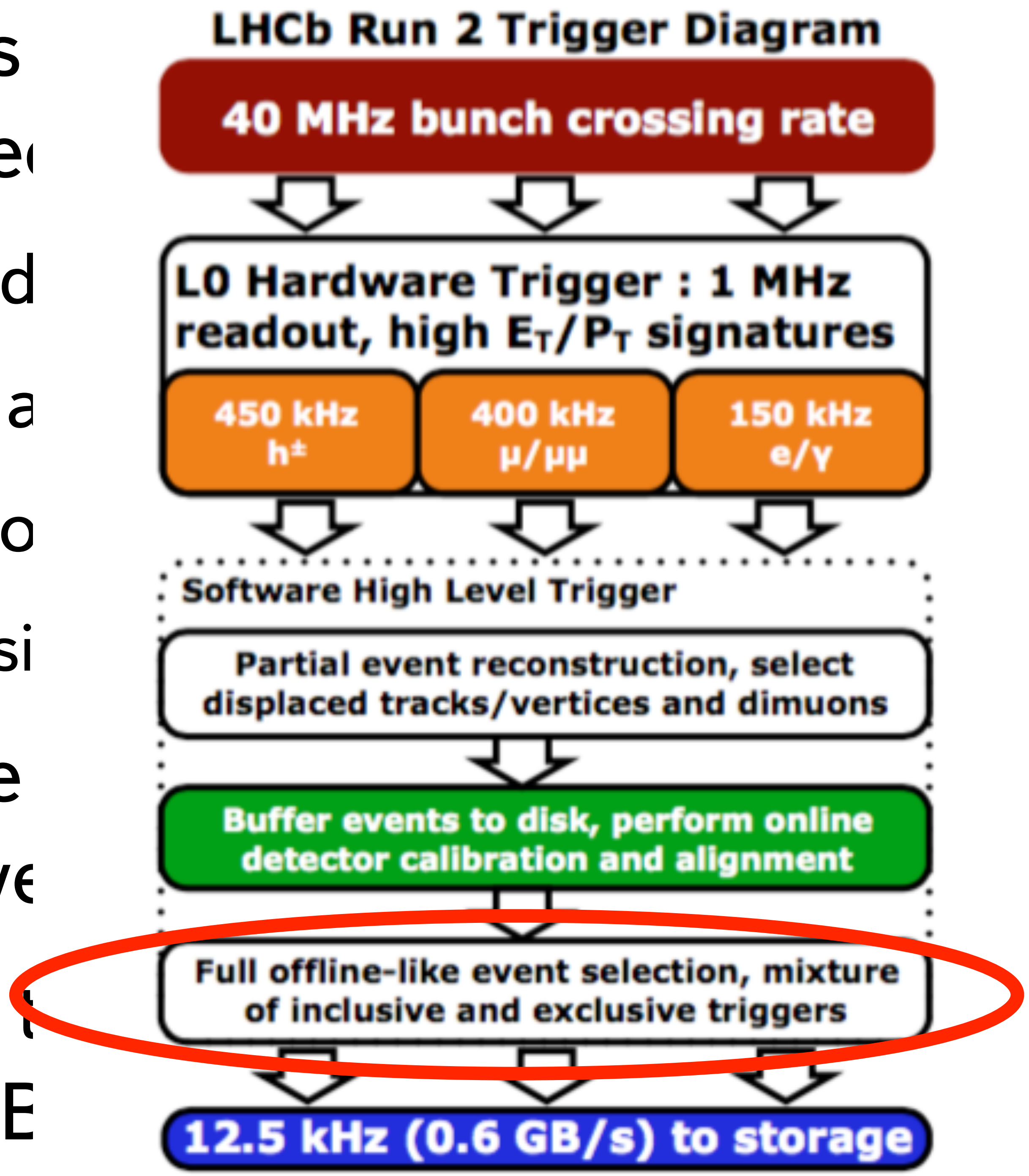




- All detectors are aligned & calibrated in-situ using the 100kHz HLT1 output rate
- Updates applied automatically if needed prior to HLT2 starting

HLT2

- HLT2 performs … aligned and calibrated dete…
  - Reconstructed … roduced offline
  - Selections of a… vent possible
  - Combination o… ctions
    - Main B physi… MVA
- Offline storage … assuming a nominal LHC ye…
- Even in Run 2, t… ciency losses for charm at 100kE…



HLT2

- HLT2 performs full event reconstruction using aligned and calibrated detector information
  - Reconstructed objects in HLT identical to those produced offline
  - Selections of arbitrary complexity on the entire event possible
  - Combination of inclusive & exclusive trigger selections
    - Main B physics trigger: Inclusive, topology-based MVA
- However...
  - Offline storage capacity limits us to 700MB/s assuming a nominal LHC year
  - Even in Run 2, this would mean significant efficiency losses for charm at 100kB/event...

Indirect Searches Needs Huge Statistics

Phys. Rev. Lett. 116, 191601 (2016)

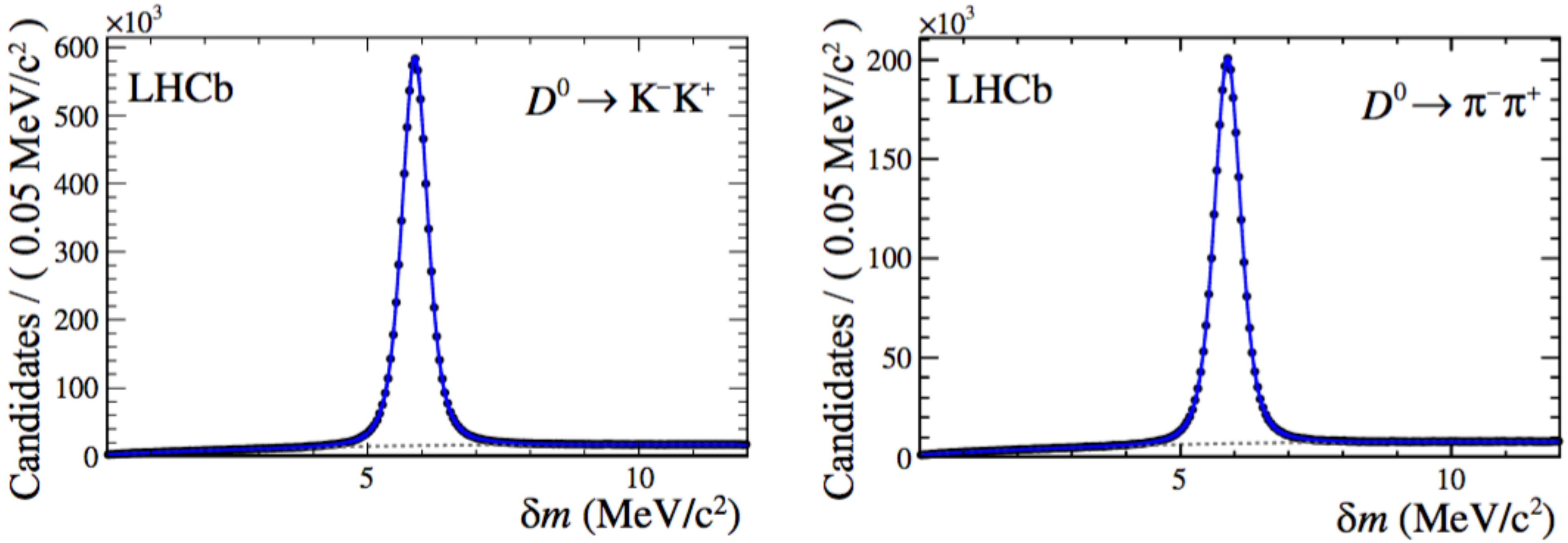
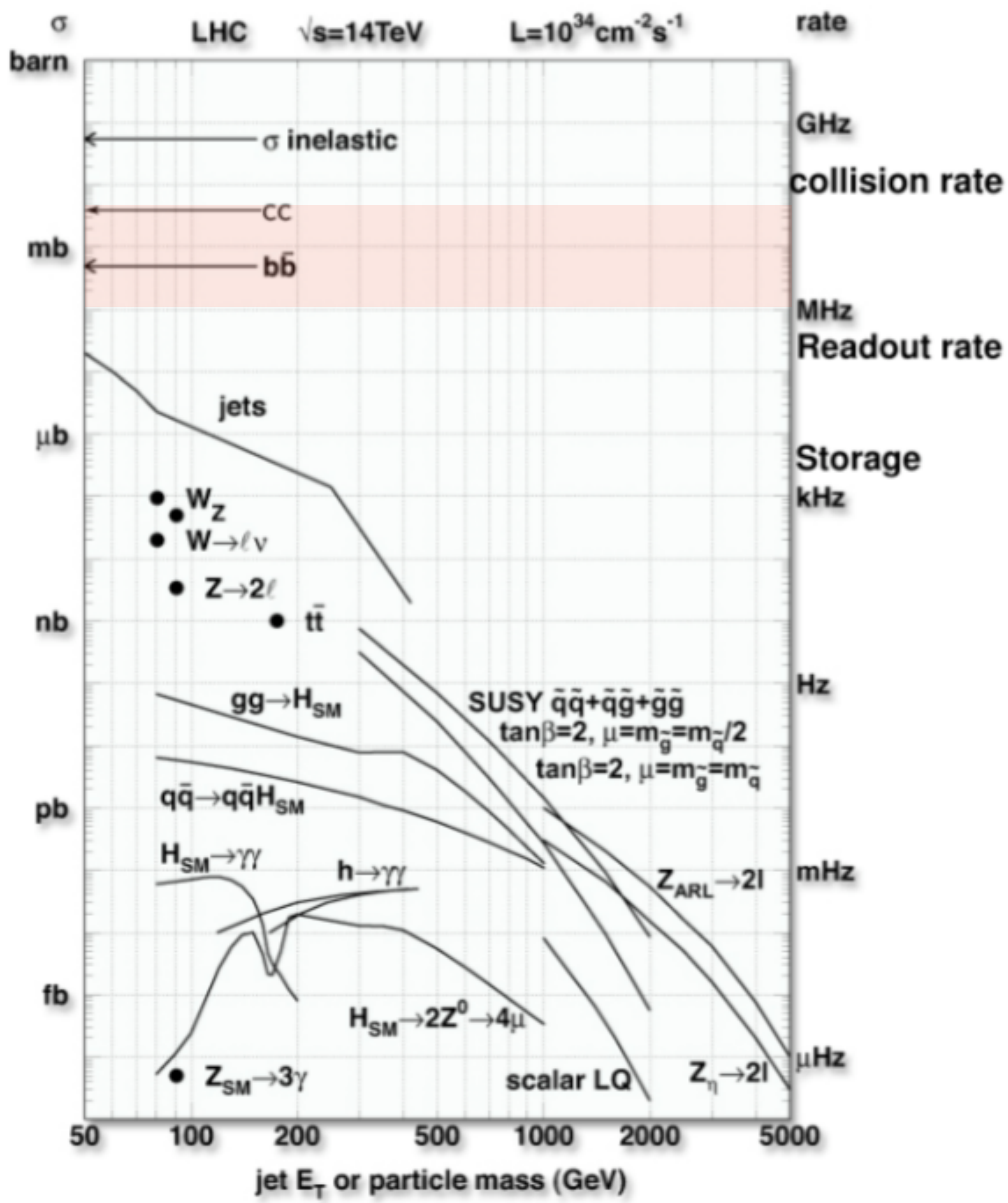


Figure 1: Fit to the δm spectra, where the D^0 is reconstructed in the final state (left) K^-K^+ and (right) $\pi^-\pi^+$. The dashed line corresponds to the background component in the fit.

$$\Delta A_{CP} = (-0.10 \pm 0.08\ (\text{stat}) \pm 0.03\ (\text{syst}))\,\%,$$

- For indirect studies of SM and New Physics huge signal datasets are necessary to achieve good sensitivity

Event Bandwidth Limitations



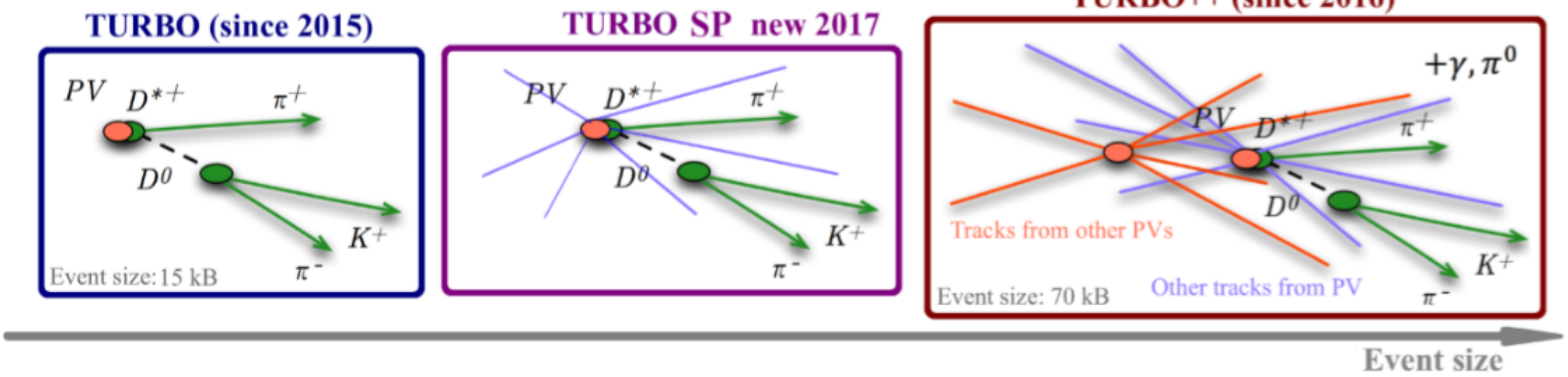
- 700Mb/s @ 100kB/events
  - Significant losses for charm and beauty analyses
    - charm suffers more

Limited Offline Storage Capacity

- Trade between event rate and event size for particular physics analyses
  - NB: events are fully reconstructed after HLT2
- Can drop RAW content, can drop low level objects, can drop unnecessary high level objects
  - keeping only high level objects necessary for the analysis
- Reverse HEP Physics analysis paradigm
  - from:
    - store generic data objects, then develop precise analysis techniques using any necessary objects
  - to:
    - thoughtfully develop analysis techniques in advance, then run analysis on trigger objects, and store all those objects, which are necessary to get solid analysis result
- Turbo stream @ LHCb
  - similar approaches: Data Scouting @ CMS, Trigger Level Analysis @ Atlas

Turbo

arXiv:1604.05596



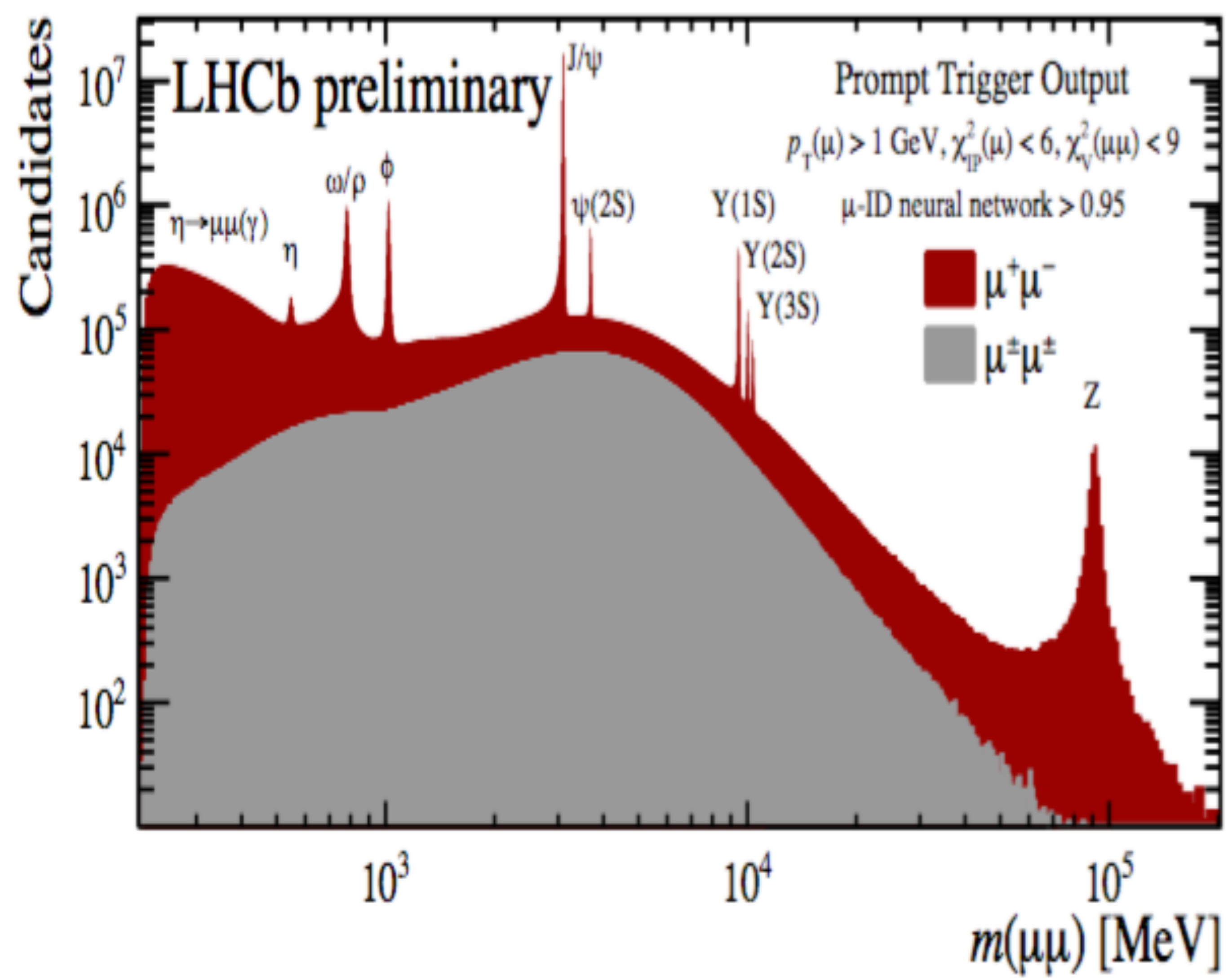
- Turbo is the LHCb paradigm for reduced event format data
- Dedicated high statistics calibration event samples are used to derive necessary detector performance parameters directly from data
- High degree of flexibility: Save only as much of the event as is needed for analysis
  - Keep all reconstructed objects, drop the raw event: 70kB
  - Keep only objects used to trigger: 15kB
  - 'Selective Persistence' objects used to trigger + user-defined selection: 15 → 70kB

Run 2 Experience

- In 2017:
  - 528 trigger lines at HLT2. 50% are Turbo
  - 25% of the trigger rate is Turbo but it counts for only 10% of the bandwidth
  - Many analyses would not be possible without Turbo

CERN-EP-2017-248
LHCb-PAPER-2017-038
October 5, 2017

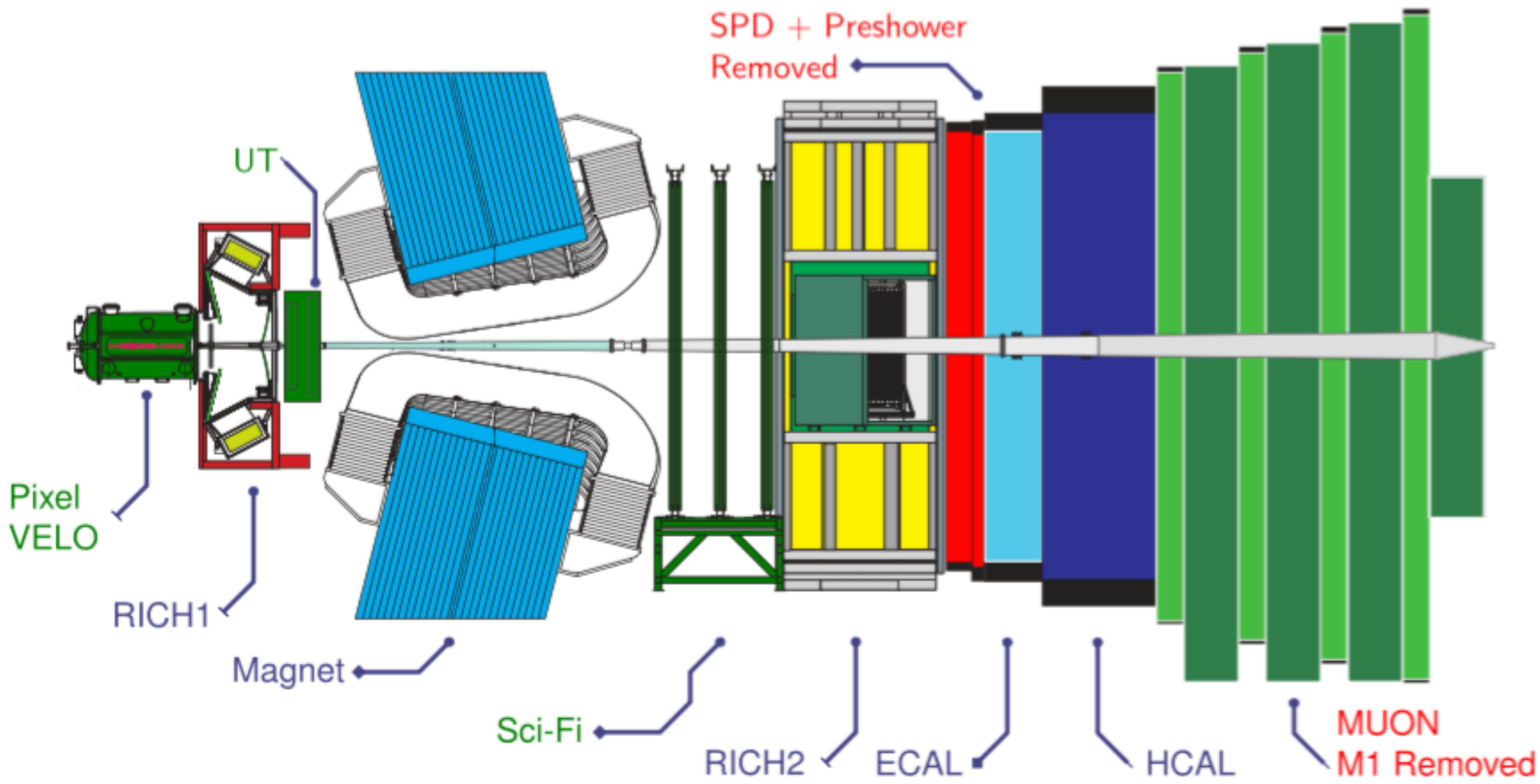
Search for dark photons produced in 13 TeV *pp* collisions



RUN 3

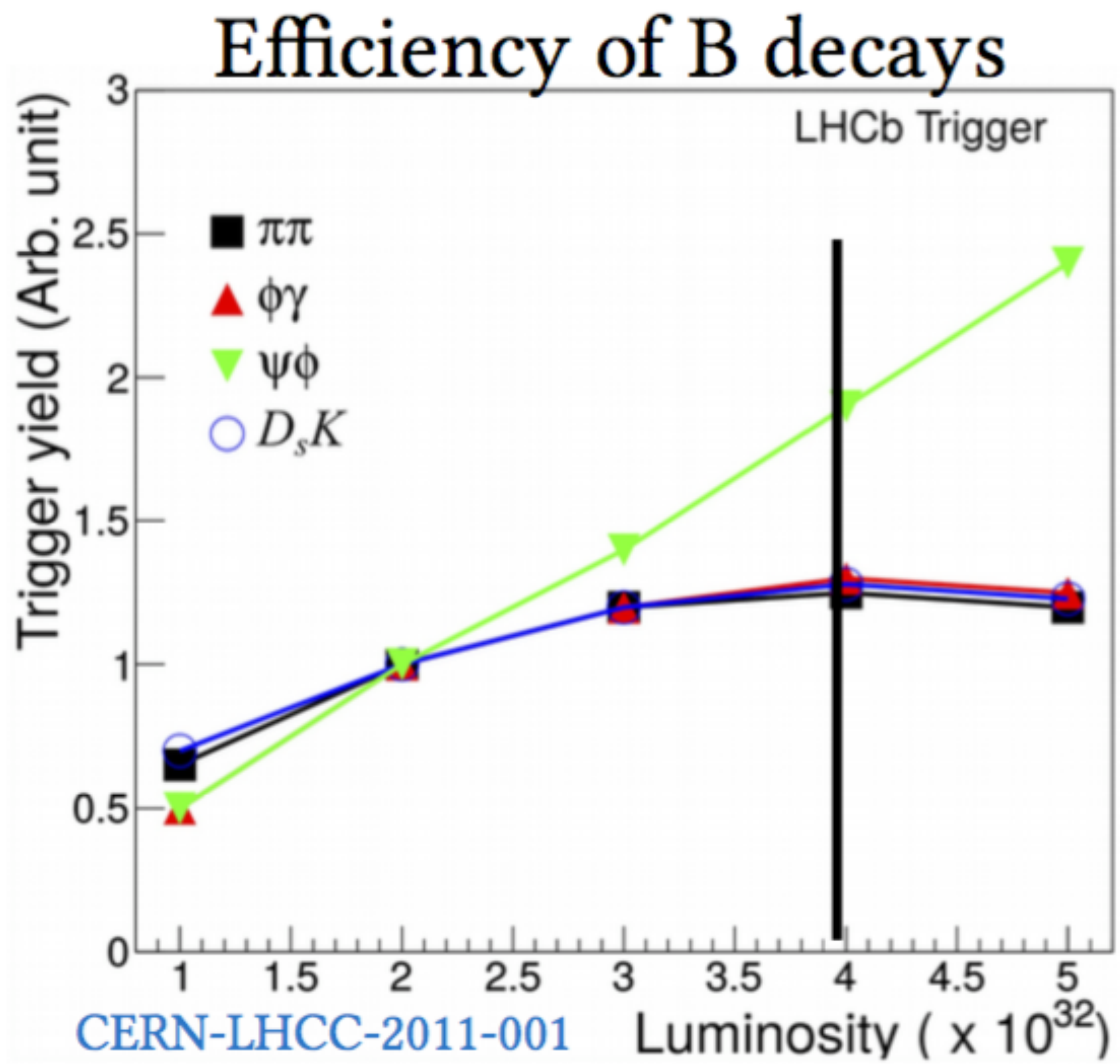
From 2021, LHCb will run at 14 TeV @ LHCb $\mathscr{L}_{inst} \sim 2 \cdot 10^{33}$ /cm^2/s

The First LHCb Upgrade



- VELO moves from r , φ strips to pixels:
  - LHCb-TDR-013
- RICH replaces photon detectors, SPD, PRS, M1 removed:
  - LHCb-TDR-014
- Trackers replaced: scintillating fibers + silicon microstrips:
  - LHCb-TDR-015
- The readout & trigger gets upgraded:
  - LHCb-TDR-016

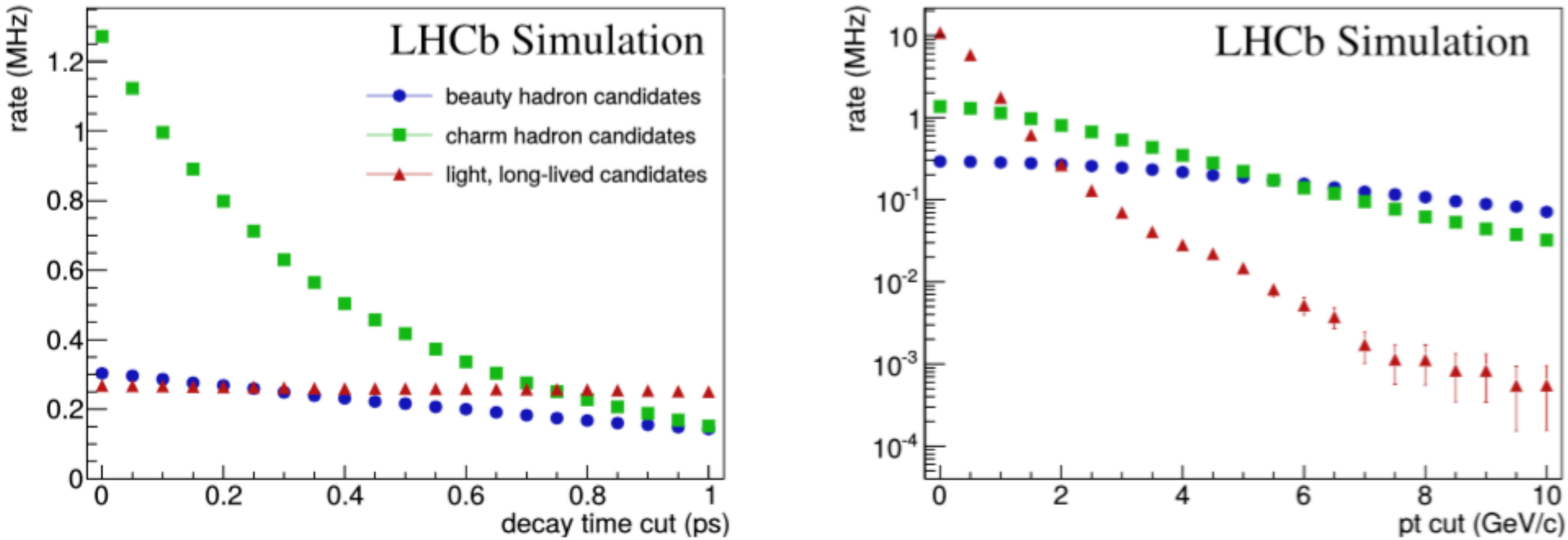
Hardware Trigger Limitations



- Hadronic signal trigger rates saturate for higher luminosity
  - consequence of total trigger rate throttling by E_T cut at L0

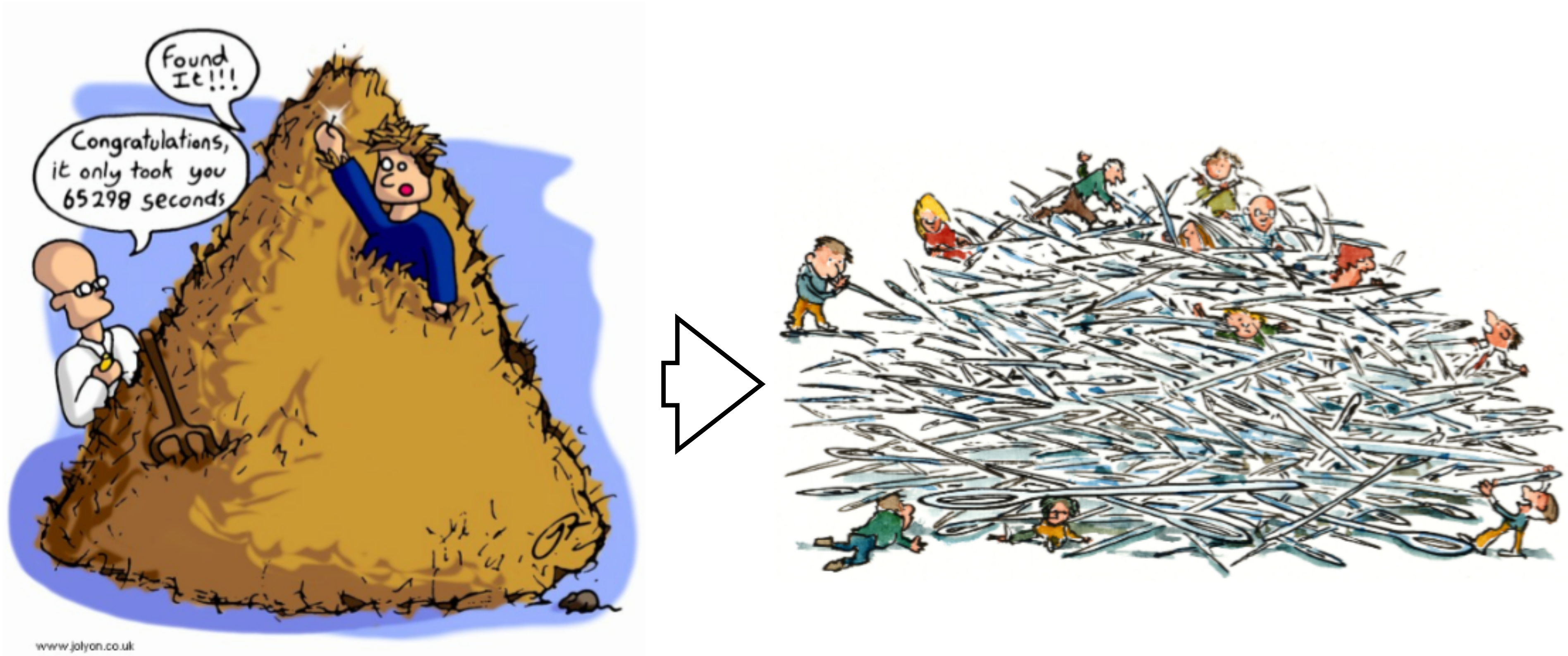
MHz Signal Era

LHCb-PUB-2014-027



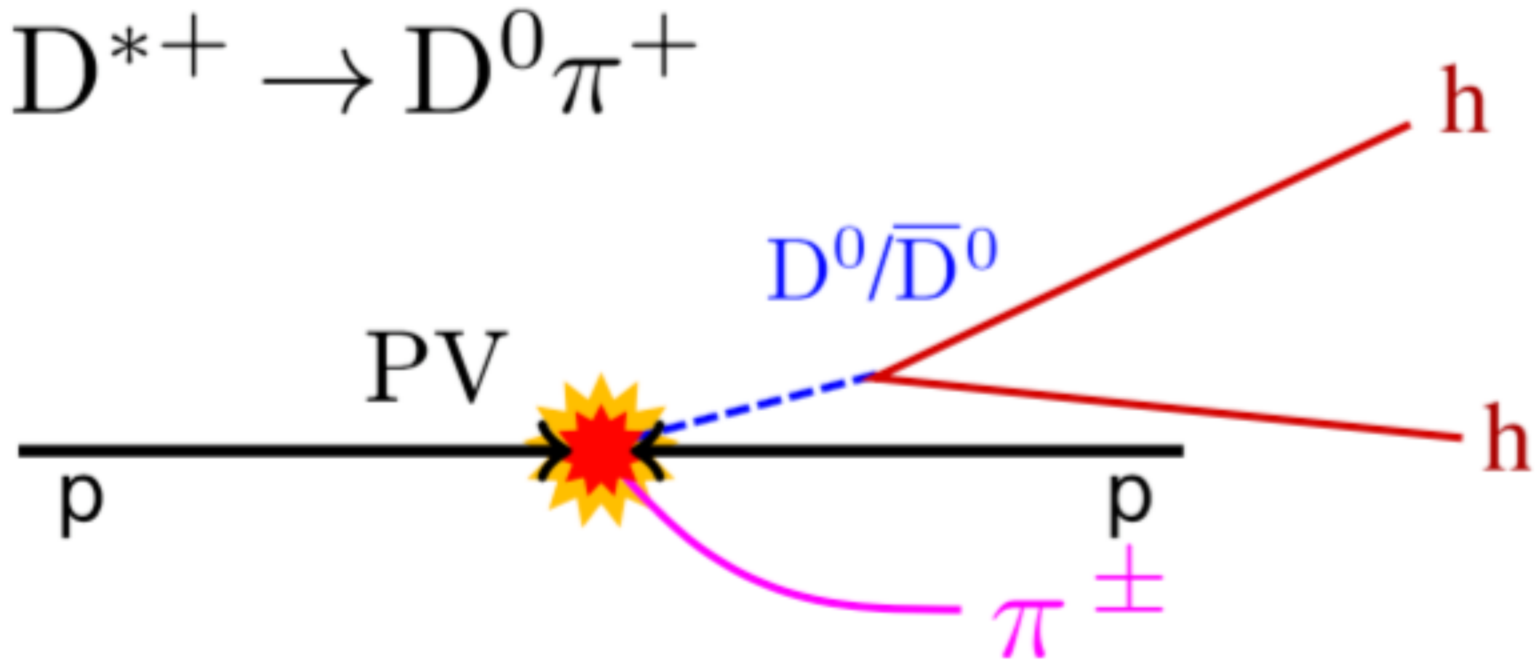
- LHCb will take 5× more collisions per second
- Readout becomes a bottleneck as signal rates → MHz even after simple trigger criteria
  - 24% (80GB/s) of events contain reconstructable charm hadron
  - 2% (27GB/c) of events contain a beauty hadron
  - 1GB/s for Run 2 offline readout
    - higher throughput is expected after upgrade

Change in Trigger Paradigm



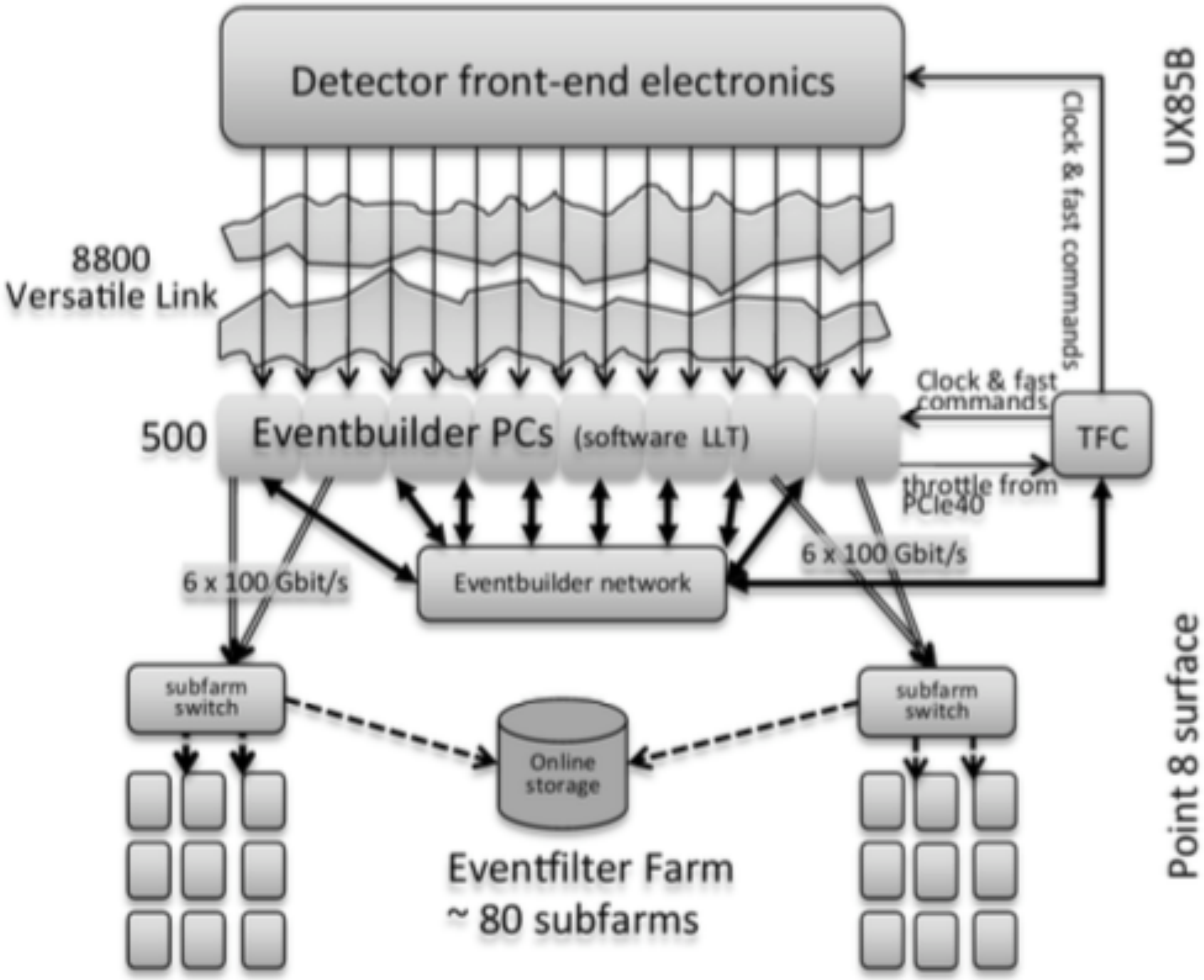
- From: which (trivial) backgrounds we could reject
  - categorise signals from backgrounds
- To: which signal we could afford to throw away
  - categorise different signals
- Most triggers must be exclusive

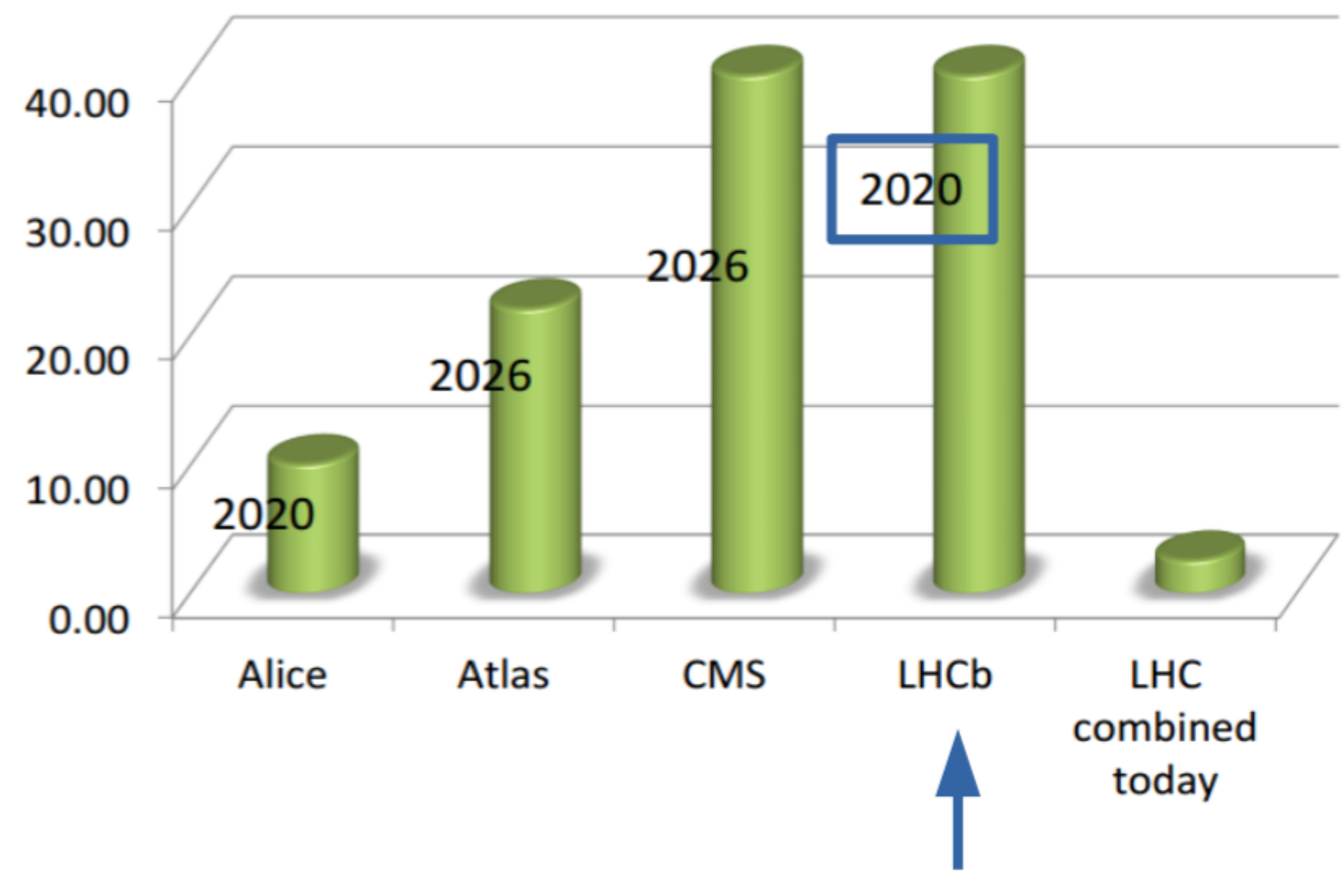
Example: Charm Mixing



- Cabibbo favoured $D^0 \to K^-\pi^+$ is 300× more abundant than DCS $D^0 \to K^+\pi^-$
- Want to keep 100% of the 'interesting' DCS mode
  - but prescale the CF mode
    - cannot be done using simple 'trigger' criteria
- Full reconstruction + Particle ID in the trigger needed to make this possible
  - Phys. Rev. Lett. 111, 251801 (2013)

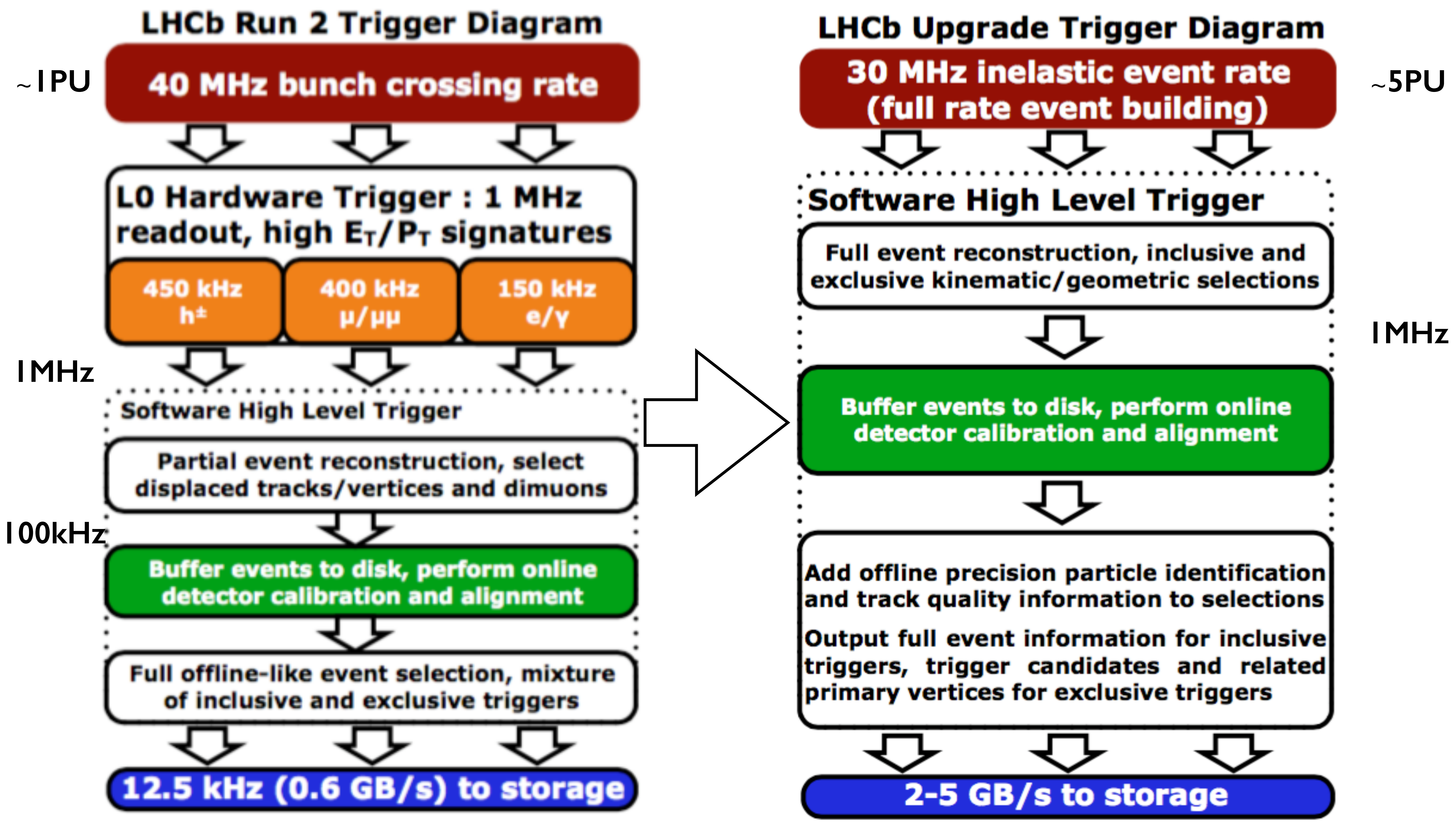
Solution: Readout Directly to Software





- Get rid of hardware trigger
- Detector readout at the LHC bunch crossing frequency
- Full reconstruction is available @HLT farm
  - 40 Tbit/s event building and filtering

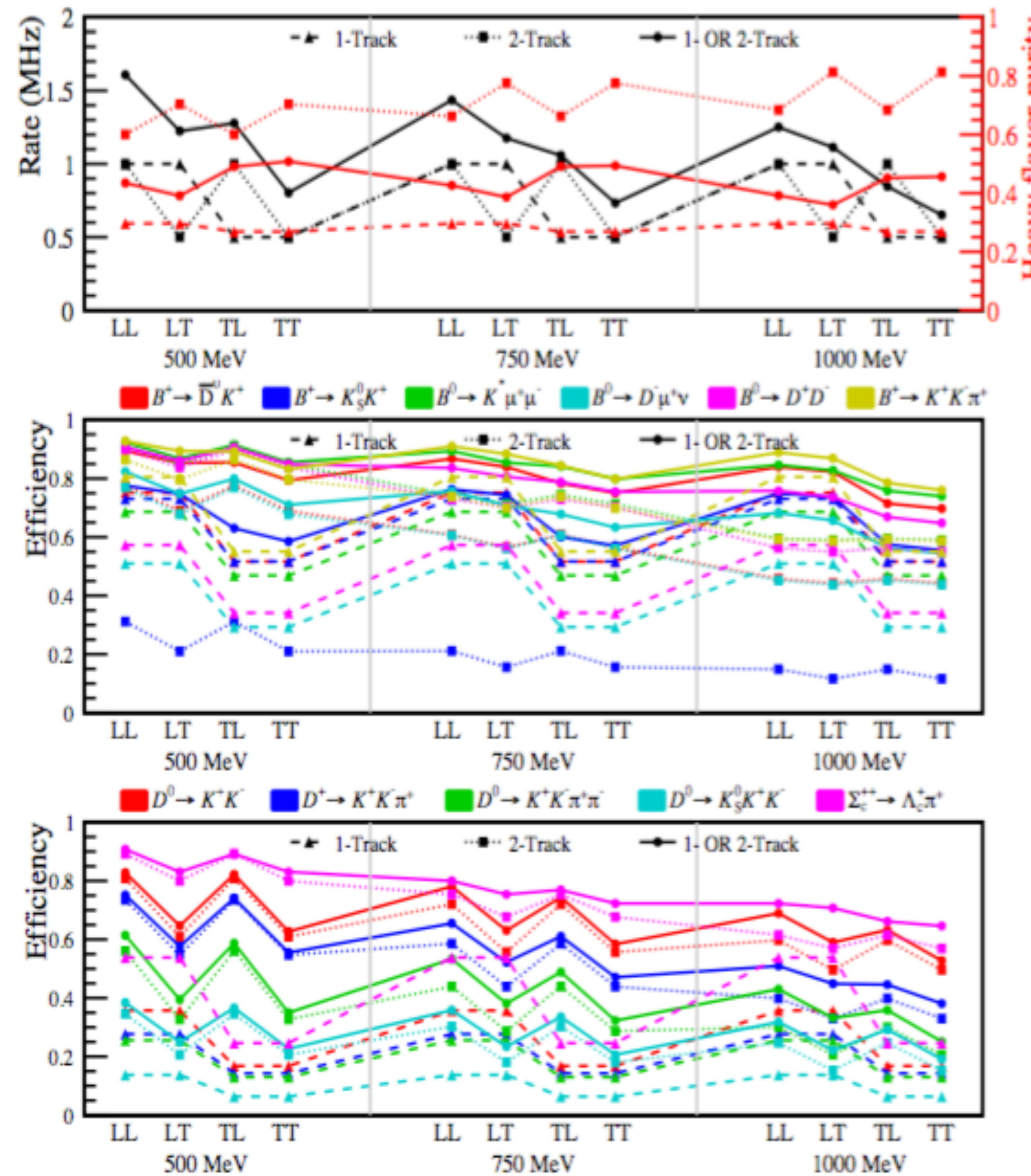
The Run 3 Trigger



Software “L1” trigger

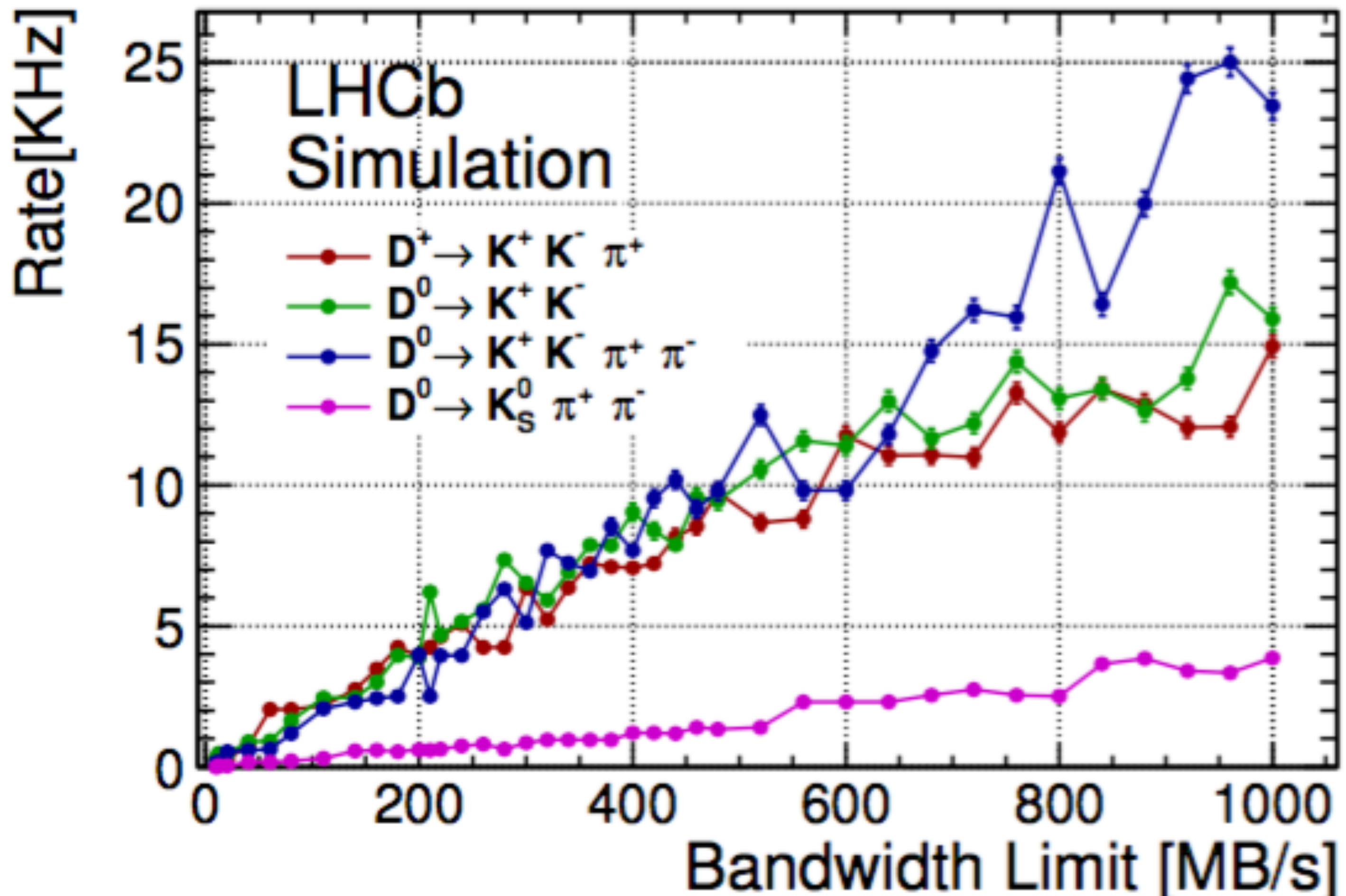
LHCb-PUB-2017-006

- 1- and 2- track performance under study
  - MVA parameters for Loose and Tight configurations
  - Several tracking thresholds 500 → 1000 MeV
- Results with minimal changes from Run 2:
  - 1-track needs more work
  - 2-track performance is good already



Software "L2" Trigger

LHCb-PUB-2017-006



- Turbo paradigm: More exclusive selections than in Run 2, with wide adoption of MVAs
- With many (> 500) trigger lines, sharing output bandwidth equitably is a challenge
- Genetic algorithm based procedure makes this easier, analysts decide between event size and output rate

No Free Cheese

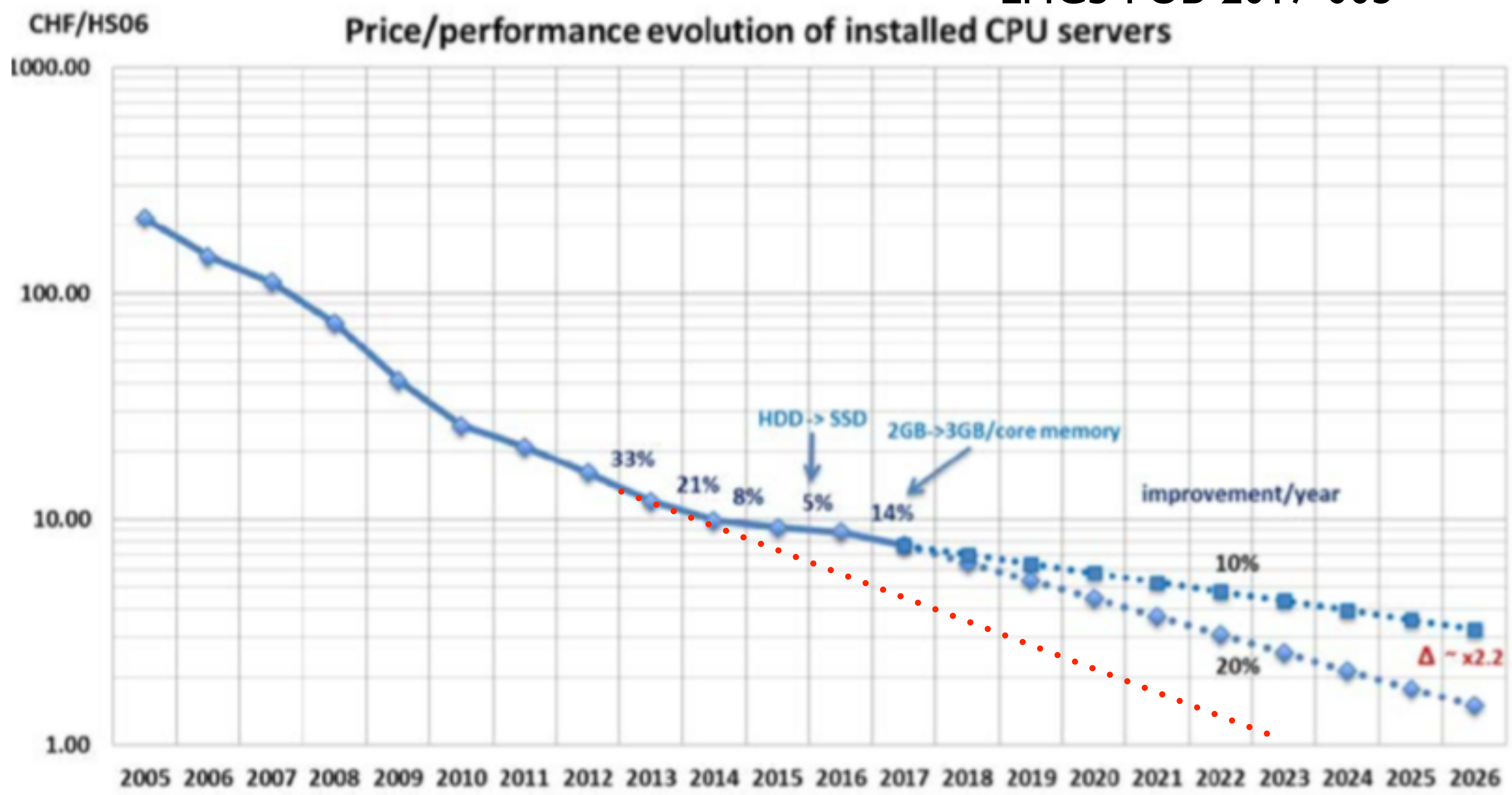
- Throwing away most of the event means care must be taken
- Turbo relies on never needing to reprocess:
  - Online monitoring & data quality are even more important
  - In Run 2 the disk buffer allows up to 2 weeks of safety margin
  - Not so in Run 3, where buffer will have O(days)
- Integration testing, real-time monitoring & robust procedures are critical components of the trigger

👍 In Run 2, we have never needed to reprocess thanks to these procedures

Computing Throughput

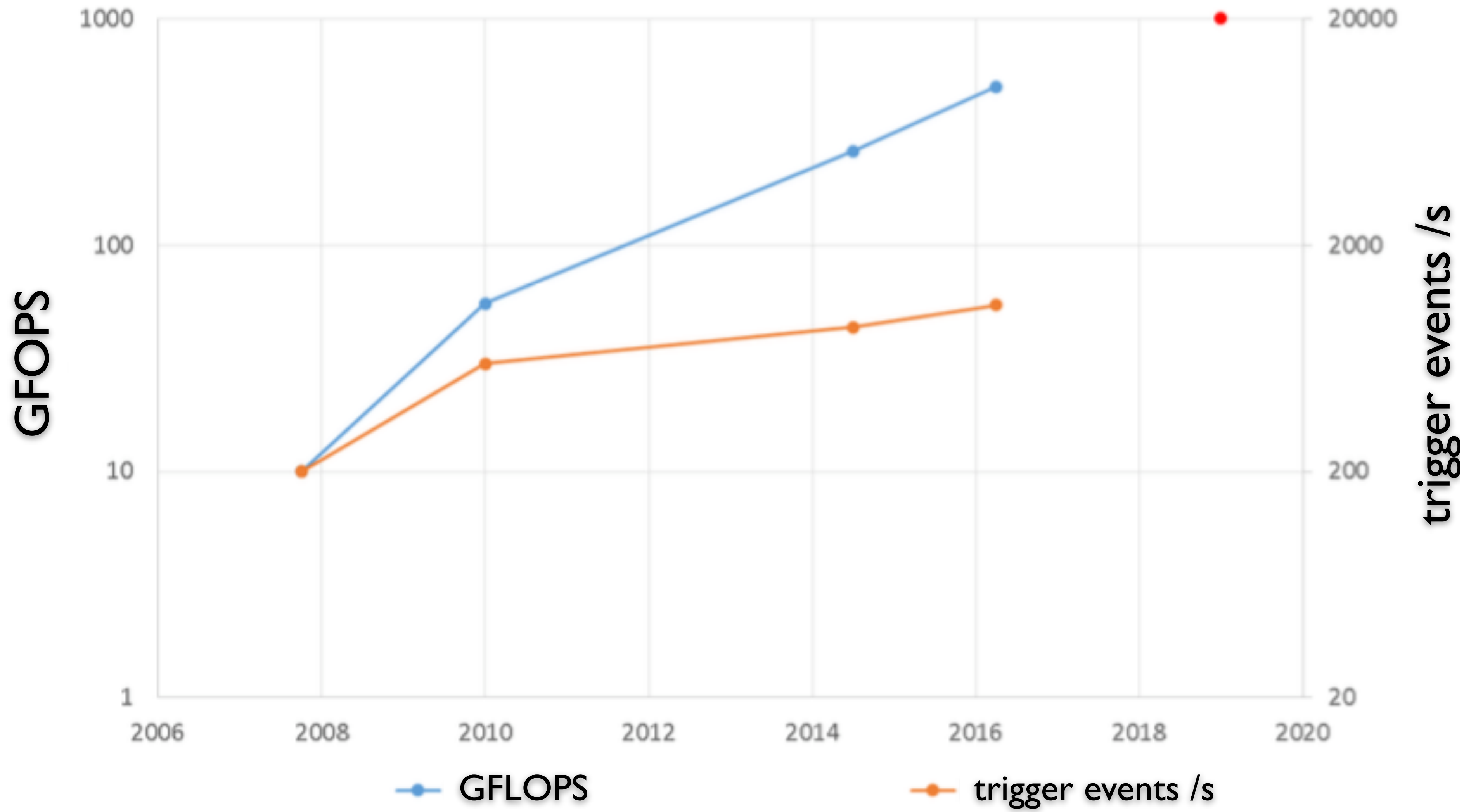
- Upgrade phase 1 starts taking data in 2021
- Upgrade farm budget: 1000 computing nodes
- Goal: Throughput > 30MHz

LHCb-PUB-2017-005



- Throughput extrapolated from 2012 hardware: 33MHz
  - from 2017 hardware: 5MHz

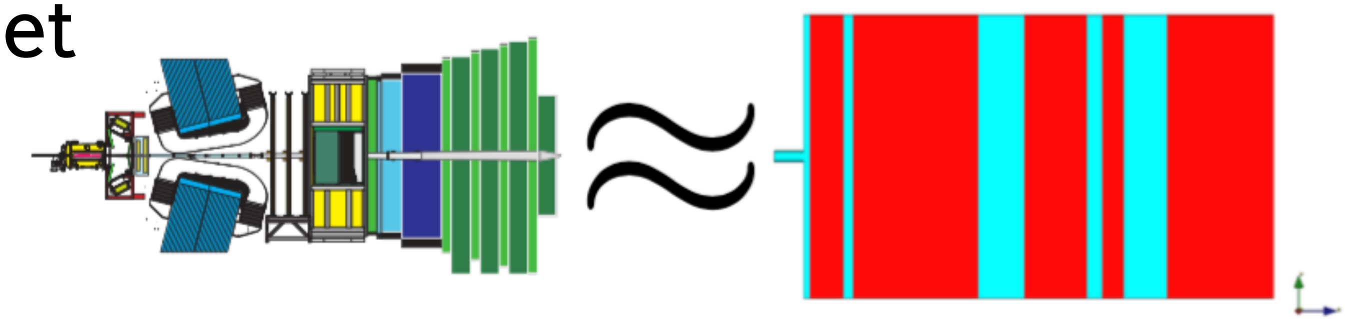
Multithreading



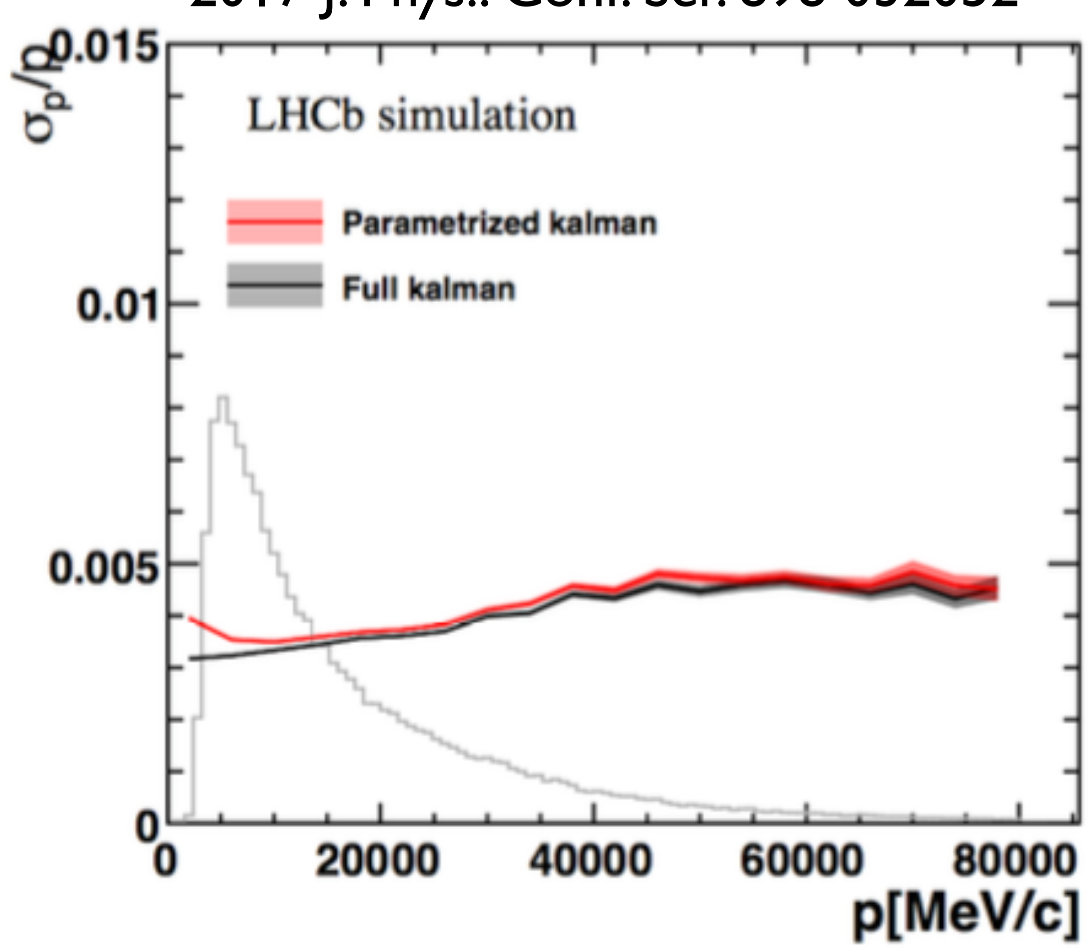
- Clock frequencies aren't increasing as fast, but the FLOPS are there
- Number of processors per CPU core are increasing (multi-threading)
- and more instructions per clock cycle (vectorisation)
- LHCb is moving from multiprocessing to a multithreading model

Algorithms Speed Up

- Many areas to improve
- E.g. track fit (Kalman Filter) uses a significant fraction of HLT1 budget
  - Run1: Material lookup + B-field propagation
  - Run2: Material map replaced with a simplification



2017 J. Phys.: Conf. Ser. 898 032052



- For the upgrade, one step further:
  - "Parameterised Kalman"
  - Replace both material and B-field with analytic functions
    - Much faster and already excellent performance

Conclusions

- LHCb signal rates in the Upgrade change the definition of a trigger:
  - ’Rejects background’ → ’categorises signal’
  - ’Reduces rate’ → ’Reduces bandwidth’
- In order to efficiently categorise MHz signals, LHCb will use a triggerless readout
- Offline quality selections mean only subset of the event has to be saved for analysis
  - Not only possible, necessary to keep high efficiency for signals
    Requires fully aligned & calibrated detector in the trigger
  - Run 2 has shown that this is the way to go for Run 3
- Not without its challenges: Extensive upgrades to the software as well as the detector