Too much of a good thing

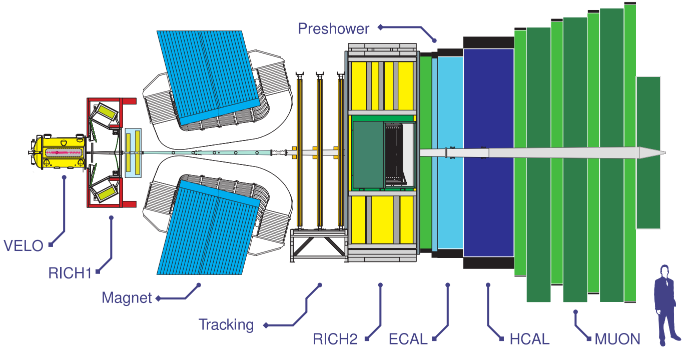
How to trigger in a signal-rich environment

Conor Fitzpatrick

CERN EP-IT Data Science seminar

LHCb: The precision flavour experiment

- LHCb was built to study beauty and charm at the LHC:



- Precise particle identification (RICH + MUON)
- Excellent decay time resolution: $\sim$45fs (VELO)
- High purity + Efficiency with flexible trigger

C. Fitzpatrick

December 13, 2017

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Question: What is a trigger?

- Trivial sounding question, but worth asking
  - 'Something that decides what events are interesting?'
  - 'Something that reduces rate?'
  - 'Something that complicates analysis'

Why do we need to trigger?

- Simply put, a trigger 'throws stuff away'
- If you're suboptimal in the trigger there is no turning back
- So why do we trigger when we could just select offline?

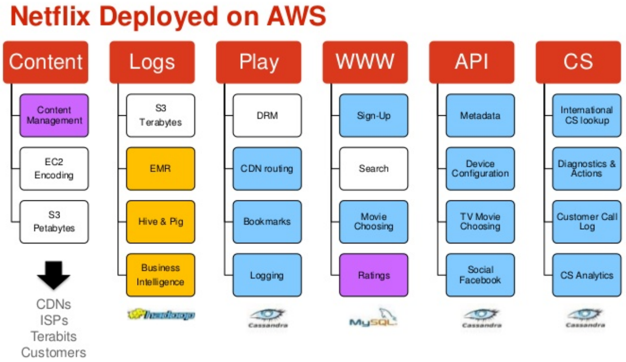
- Two reasons, both cost related

C. Fitzpatrick

December 13, 2017

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

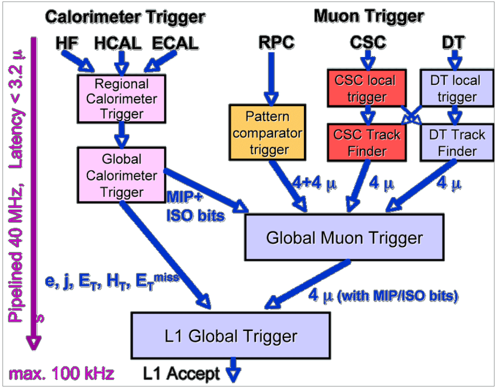
Storage is expensive



- LHC crossing rate: 30 MHz,
  LHC event sizes: $\sim 0.1 - 1$MB
  - If we kept everything: 150000 PB/year
  - Entire NETFLIX movie catalog: 40 PB[1]
- Data storage is expensive and we are not a *Fortune 500* company!

[1] Structure Data 2016

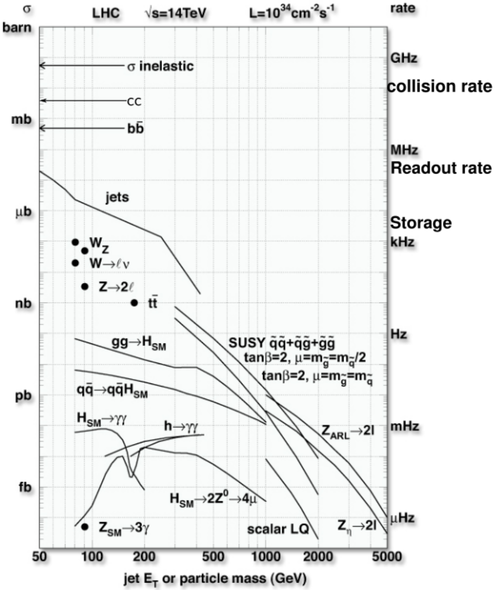
Reading out a detector is expensive



- All of the LHC experiments presently reduce the rate before detector readout
  - CMS for example: Run 1 & 2 readout operates at 100kHz
- Using limited local (muon, calorimeter) information buys a factor of 300 reduction

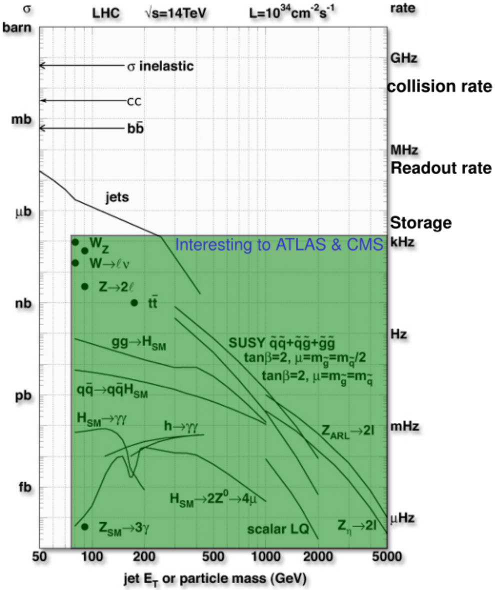
So triggers are important

- A trigger is needed to reduce storage and readout costs
- A *good* trigger does so by keeping more signal than background



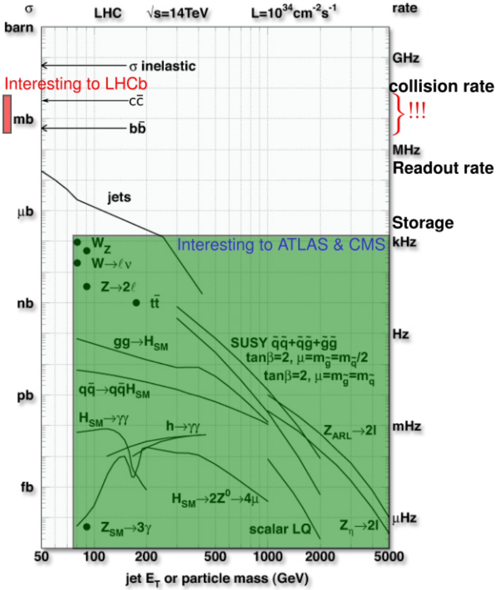
So triggers are important

- A trigger is needed to reduce storage and readout costs
- A *good* trigger does so by keeping more signal than background
- ATLAS and CMS are interested in signatures in the kHz region
  - Readout at 100kHz is efficient with reasonably straightforward E_T requirements



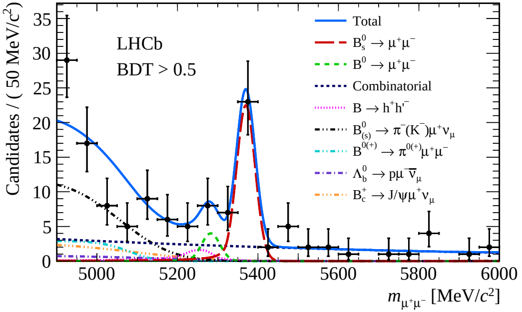
So triggers are important

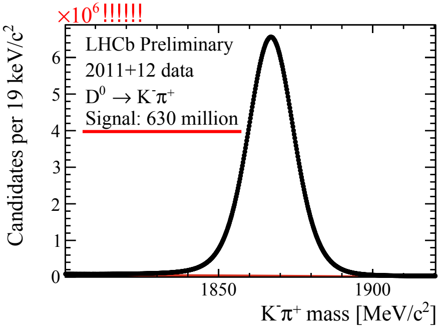
- A trigger is needed to reduce storage and readout costs
- A *good* trigger does so by keeping more signal than background
- ATLAS and CMS are interested in signatures in the kHz region
  - Readout at 100kHz is efficient with reasonably straightforward E_T requirements
- LHCb operates at $\mathfrak{L} = 4 \times 10^{32}\text{cm}^{-2}\,\text{s}^{-1}$ in Run 2
  - 45kHz of $b\bar{b}$, $\sim$1MHz of $c\bar{c}$
  - 1MHz readout is needed to stay efficient for beauty signals



The LHCb Run 2 trigger in two plots

- The LHCb trigger has to cover extremes of data taking:

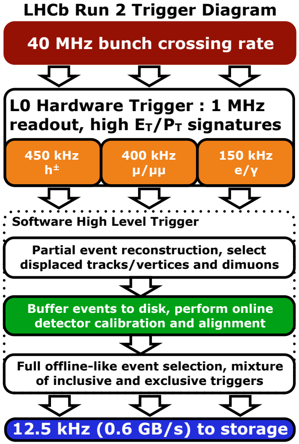




- High efficiency to collect rare decays like $B_s^0 \to \mu\mu$[2]
- High purity for enormous charm signals like $D^0 \to K\pi$[3]
- Must be flexible to operate in both extremes simultaneously: After readout, HLT has access to 100% of event in software

[2] Phys. Rev. Lett. 118, 191801 (2017)
[3] LHCb-CONF-2016-005

The Run 2 LHCb Trigger

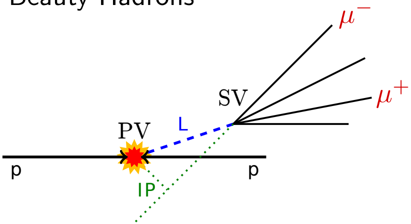


- On the surface, Run 2 trigger similar to that of other experiments:
- Three levels, each has more time than the last:
  - Level-0 trigger buys time to readout the detector with Calo, Muon p_T thresholds: $40 \rightarrow 1$MHz
  - Events built at 1MHz, sent to HLT farm ($\sim$27000 physical cores)
  - HLT1 has $40\times$ more time, fast tracking followed by inclusive selections $1\text{MHz} \rightarrow 100\text{kHz}$
  - HLT2 has $400\times$ more time than L0: Full event reconstruction, inclusive + exclusive selections using whole detector
- Flexibility comes from software-centric HLT design

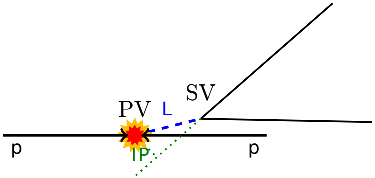
HLT1

- Beauty and charm hadron typical decay topologies:

Beauty Hadrons



Charm Hadrons



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

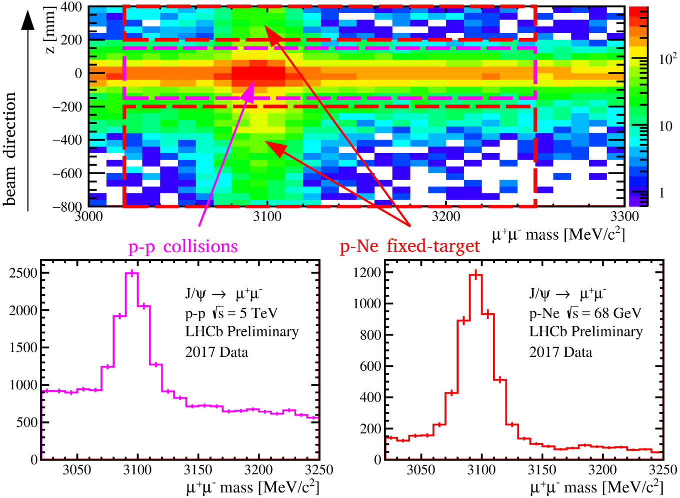
- D^0 mass ~ 1.86 GeV, appreciable daughter p_T
- $\tau \sim 0.4$ ps, Flight distance ~ 4 mm
- Also produced as 'secondary' charm from B decays.

Underlying HLT1 strategy:

- Fast reconstruction: Primary Vertices, High p_T tracks, optional Muon ID
- Inclusive triggering using MVAs on 1&2-track signatures: $\sim$100kHz output rate

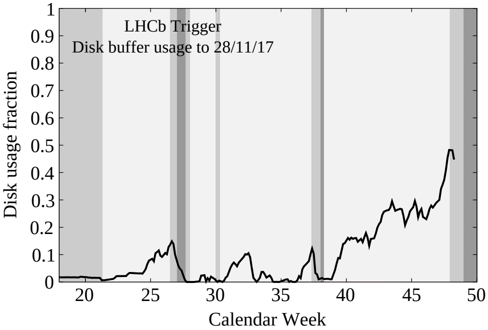
Aside: Software flexibility

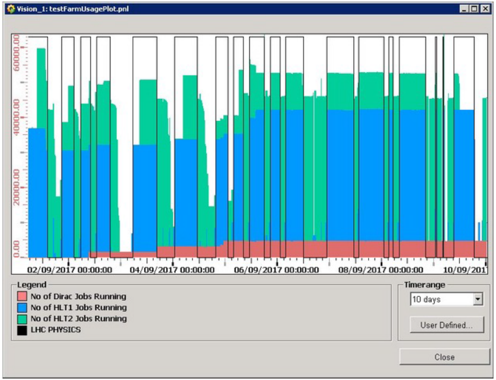
- Example: For the 5 TeV data taking period LHCb took fixed target p-Ne data
- Able to quickly deploy custom reconstruction to *simultaneously* collect pp data



- J/ψ signals in both pp and pNe at first software trigger stage

Disk Buffer

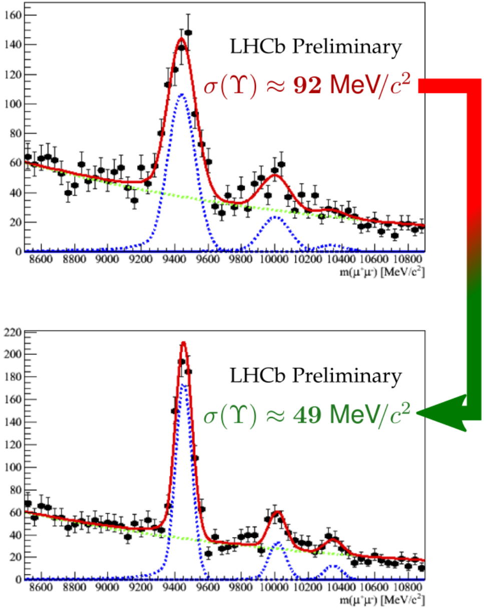




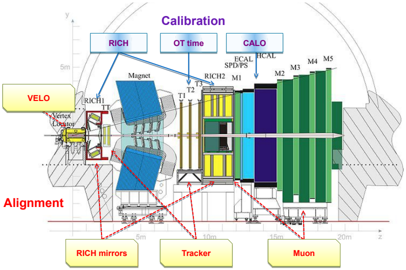
- HLT Farm is off-the shelf servers: Considerable (11PB) disk capacity
- HLT1 accepted events written to the disk in-fill at 100kHz: 2 week contingency
- HLT2 throughput in-fill is 30kHz, out of fill 90kHz when HLT1 isn't running
- Effectively doubles trigger CPU capacity, Farm is used twice for HLT, excess used for simulation
- Asynchronous HLT has another big advantage though...

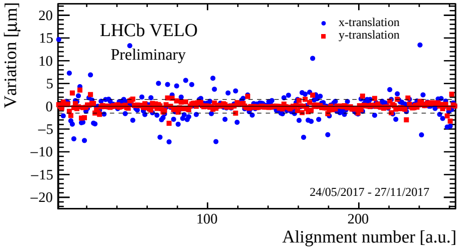
Real-time Alignment + Calibration

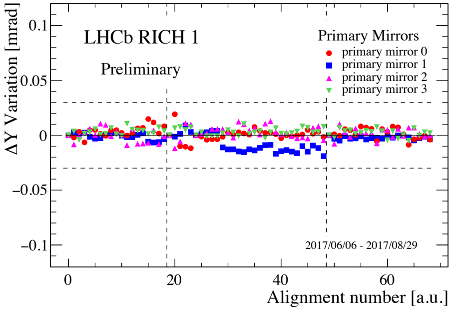
- With Run 2 signal rates, efficient & pure output requires full reconstruction at HLT2
  - Online selections $\rightarrow$ offline selections
  - Reduces systematic uncertainties and workload for analysts
- Alignment and calibration of full detector in the trigger needed
- While HLT1 is written to disk,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



A fully aligned detector





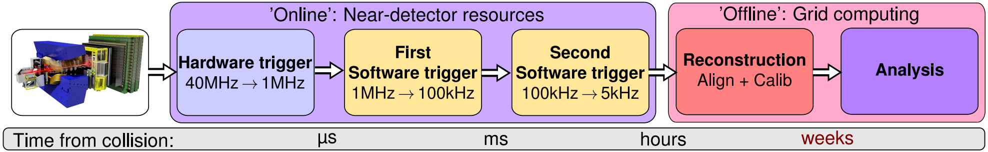


- All detectors are aligned & calibrated in-situ using the 100kHz HLT1 output rate
- Updates applied automatically if needed prior to HLT2 starting

HLT2

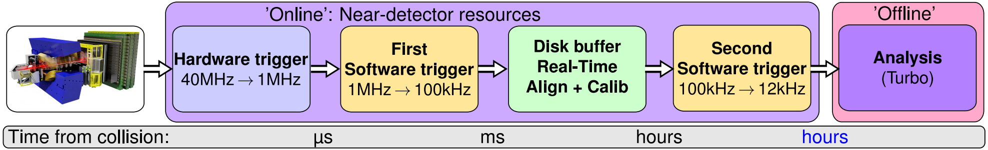
- HLT2 performs full event reconstruction using aligned and calibrated detector information
- Reconstructed objects in HLT identical to those produced offline
- Selections of arbitrary complexity on the entire event possible
- Combination of inclusive & exclusive trigger selections
  - Main B physics trigger: Inclusive, topology-based MVA
- 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...

Reduced event formats



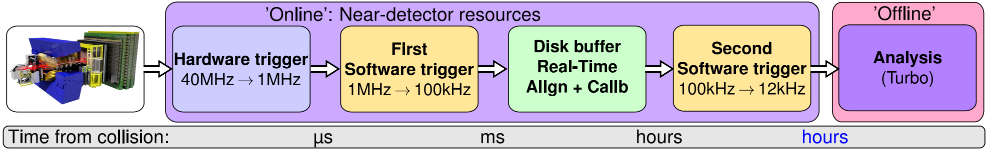
- Trigger *rates* aren't important, output *bandwidth* is
- Offline reprocessing previously needed to recover best quality

Reduced event formats



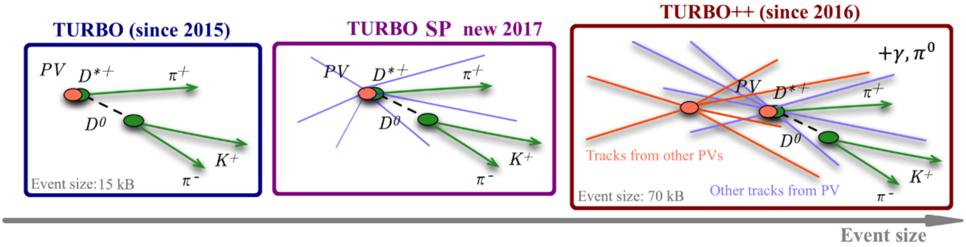
- Trigger *rates* aren't important, output *bandwidth* is
- Offline reprocessing previously needed to recover best quality
- If online == offline, why reprocess? Do analysis on trigger objects, write only the relevant objects offline
- Significant reduction in event size → higher rates for the same bandwidth

Reduced event formats



- Trigger *rates* aren't important, output *bandwidth* is
- Offline reprocessing previously needed to recover best quality
- If online == offline, why reprocess? Do analysis on trigger objects, write only the relevant objects offline
- Significant reduction in event size → higher rates for the same bandwidth
- Added bonus: offline CPU freed up for simulation.
- CMS, ATLAS, LHCb call this Data Scouting, Trigger Level Analysis, Turbo respectively

Turbo



- Turbo is the LHCb paradigm for reduced event format data[4]
- 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

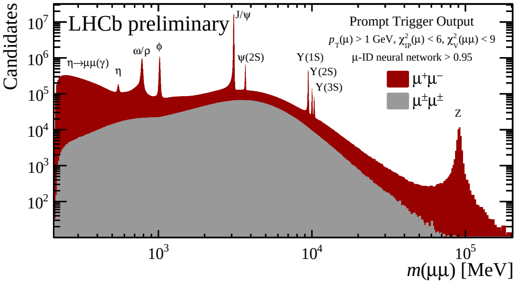
[4] arXiv:1604.05596

Turbo usage in Run 2

- 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[5]

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

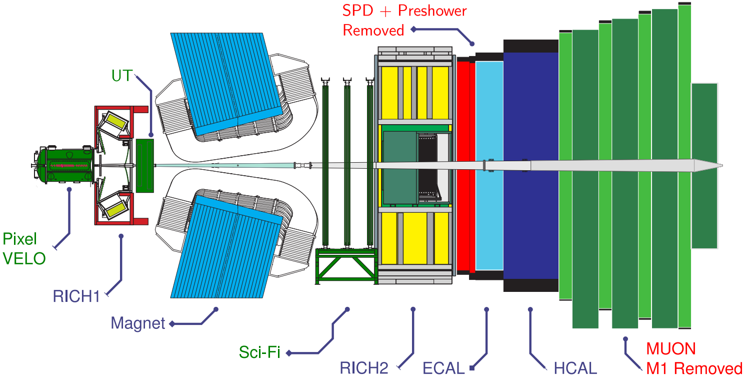
Search for dark photons produced in 13 TeV *pp* collisions



[5] LHCb-PAPER-2017-038

The first LHCb Upgrade

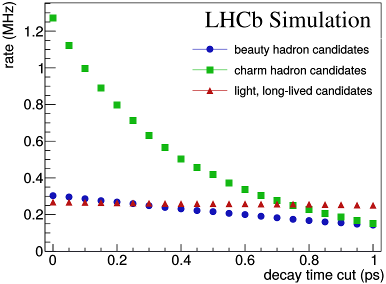
- From 2021, LHCb will run at $\mathcal{L} = 2 \times 10^{33}$ cm^{-2} s^{-1}

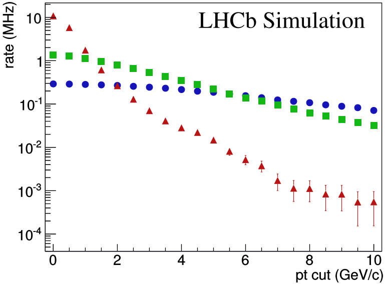


- VELO moves from r, ϕ strips to pixels: LHCb-TDR-013
- RICH replaces photon detectors, SPD, PRS, M1 removed: LHCb-TDR-014
- Trackers replaced: scintillating fibers + silicon microstrips: LHCb-TDR-015
- The readout & trigger gets upgraded: LHCb-TDR-016

The MHz signal era

- LHCb will take 5 $\times$ more collisions per second





- Readout becomes a bottleneck as signal rates $\rightarrow$ MHz even after simple trigger criteria [6]

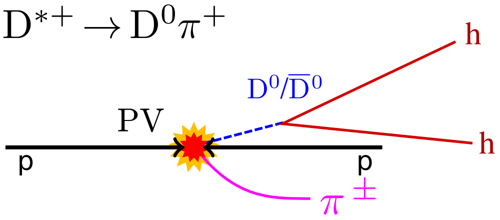
[6] LHCb-PUB-2014-027

So what 'stuff' can we throw away?

- The problem is no longer one of rejecting (trivial) background
- Fundamentally changes what it means to trigger

- Instead, we need to categorise different 'signals'
  - Run 2 showed us how, in Run 3 it's a necessity

Triggering with MHz signals

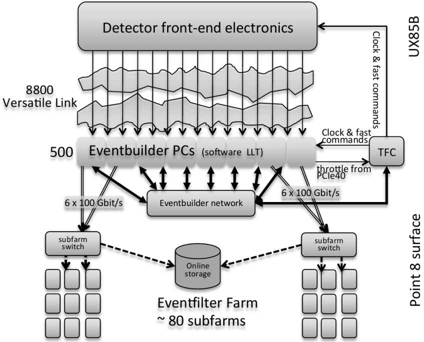


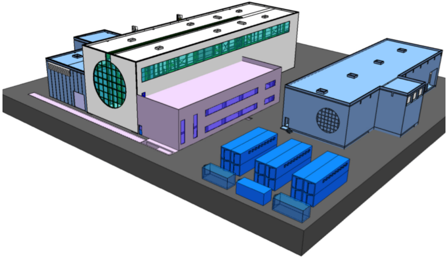
- Example: Charm mixing[7]
  - Cabbibo favoured $D^0 \to K^-\pi^+$ is $300\times$ more abundant than DCS $D^0 \to K^+\pi^-$
  - Want to keep 100% of the 'interesting' DCS mode, but prescale the CF mode
  - Cannot be done using simple 'trigger' criteria
  - Full reconstruction + Particle ID in the trigger needed to make this possible

[7] Phys. Rev. Lett. 111, 251801 (2013)

Reading out at 30MHz

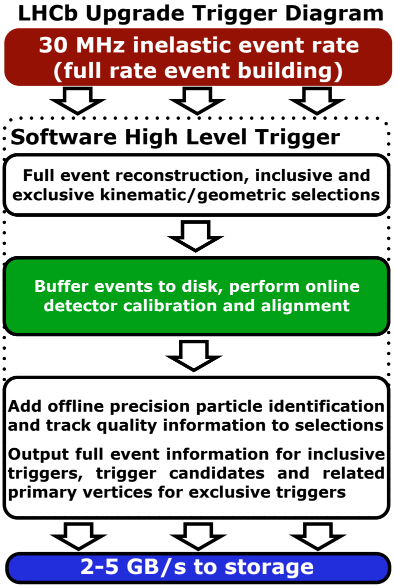
- Solution: Readout and reconstruct 30 MHz of collisions in software!





- LHCb Upgrade phase 1: Detector readout at the LHC bunch crossing frequency
- Event builder, trigger farm & disk buffer in containers above LHCb

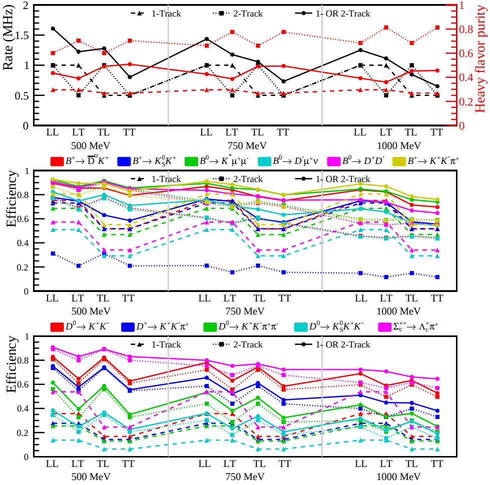
The Run 3 Trigger



- Run 2: has proven the strategy at 1MHz at a pileup of ~ 1
- Run 3: must now process full 30MHz at $5 \times$ the pileup
- Overall strategy similar, but:
  - HLT1 $\rightarrow$ first level trigger. Output 100kHz $\rightarrow \sim$1MHz
  - Disk buffer has contingency of O(days) instead of weeks
  - HLT2 $\rightarrow$ second level trigger. 2-5GB/s output

Run 3 first level trigger

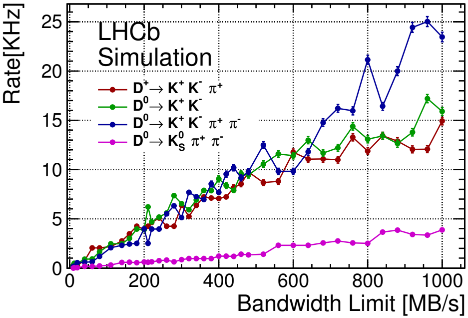
- 1- and 2- track performance under study[8]
  - MVA parameters for Loose and Tight configurations
  - Several tracking thresholds $500 \rightarrow 1000$ MeV
- Results with minimal changes from Run 2:
  - 1-track needs more work
  - 2-track performance is good already



[8] LHCb-PUB-2017-006

Run 3 second level trigger

- 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[9]:



[9] LHCb-PUB-2017-006

There's no turning back. . .

- 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

Throughput

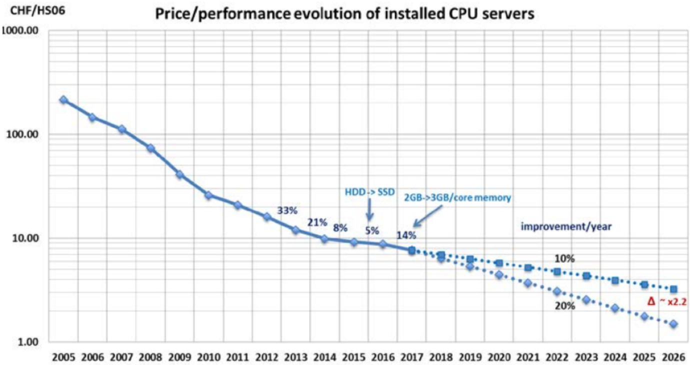
- Upgrade phase 1 starts taking data in 2021
- Upgrade farm budget: 1000 computing nodes
- Benchmark using today's CPUs and extrapolate

$$T = N \times t \times g^{\Delta y}$$

- Throughput T determined using Number of nodes, N, throughput on single node, t
- Growth factor per year at equal cost g, extrapolates growth in years until data taking, Δy
- Goal: $T > 30\text{MHz}$

CPUs are evolving

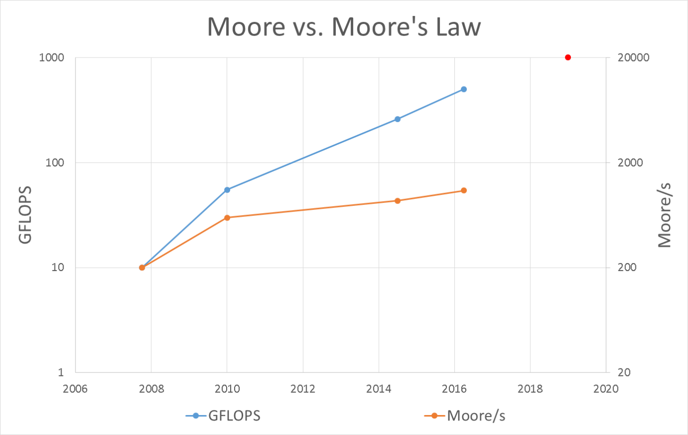
- Growth rate at equal cost is slowing down:



- Throughput extrapolated from 2012 hardware: 33MHz. 2017 hardware: 5MHz[10]

[10] LHCb-PUB-2017-005

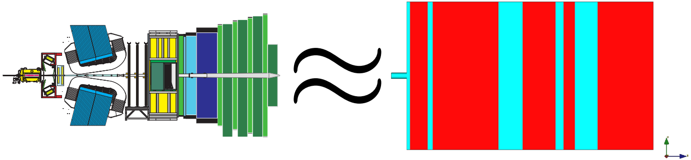
A multithreaded Trigger



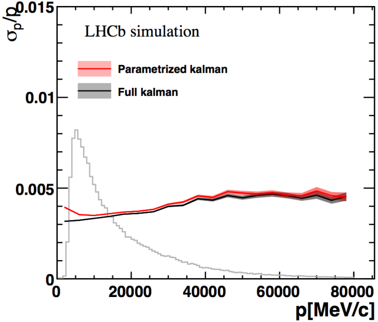
- Clock frequencies aren't increasing as fast, but the FLOPS are there
- Number of processors per CPU core are increasing (multi-threading)
- and more instructions per clock cycle (vectorisation)
- LHCb is moving from multiprocessing to a multithreading model

Using fewer cycles

- Track fit (Kalman Filter) uses a significant fraction of HLT1 budget
- Run1: Material lookup + B-field propagation
- Run2: Material map replaced with a simplification



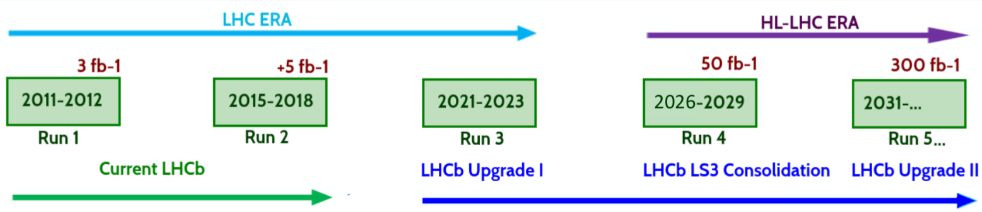
- For the upgrade, one step further: "Parameterised Kalman"
- Replace both material and B-field with analytic functions
- Much faster and already excellent performance



Conclusions

- LHCb signal rates in the Upgrade change the definition of a trigger:
  - 'Rejects background' $\rightarrow$ 'categorises signal'
  - 'Reduces rate' $\rightarrow$ '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 forward for Run 3
- Not without its challenges: Extensive upgrades to the software as well as the detector

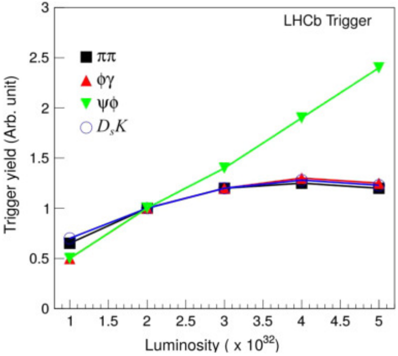
Upgrade timelines



- LHCb: 8fb^{-1} Run1 + Run 2
- 50fb^{-1} Run 3 + Run 4
- 300fb^{-1} Run 5 + ...

L0 limit

- L0 efficiency for hadronic final states degrades with increasing luminosity[11]:

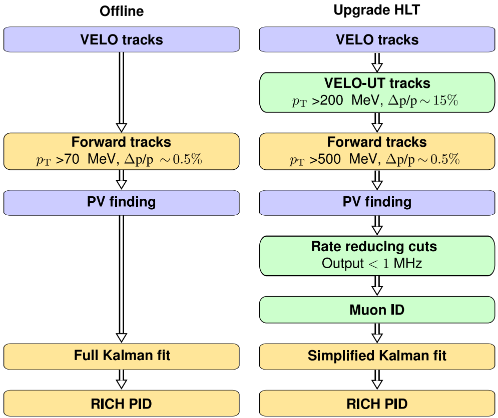


- LHCb Run 1+2: $4 \times 10^{32} \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
- LHCb Run 3: $2 \times 10^{33} \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
- LHCb Run 5: $2 \times 10^{34} \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$?

[11] CERN-LHCC-2011-001

Online == Offline

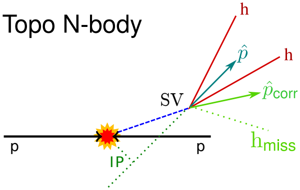
- Reconstruction in the trigger vs. offline
- In Run 2: Simplified Kalman used offline too

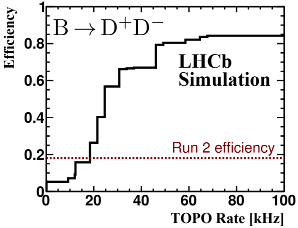


Inclusive triggering in Run3

- Inclusive topological trigger used in Run 2

Topo N-body





- Save event based on partial signal information, full reconstruction later
- Rejects 'obvious' backgrounds, looks for displaced n-track vertices
- 99% of output is b hadrons.

- Upgrade: Topological trigger will need to be much tighter[12]
- Exclusive triggers needed to stay efficient

[12] LHCb-PUB-2014-031