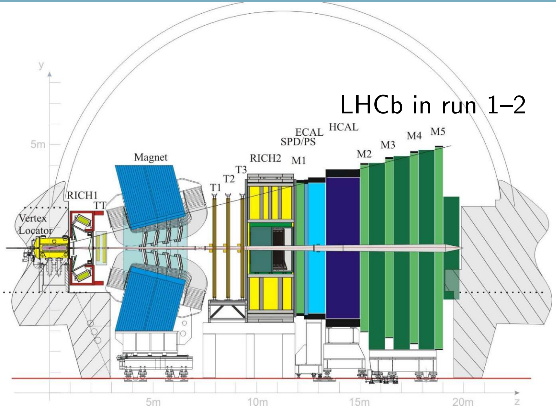
Software trigger for upgraded LHCb detector

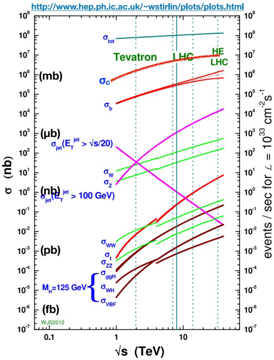
Anton Poluektov
on behalf of LHCb collaboration

Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France

25 February 2020

LHCb experiment



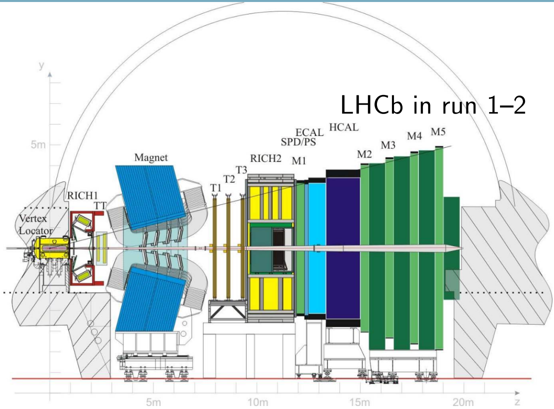


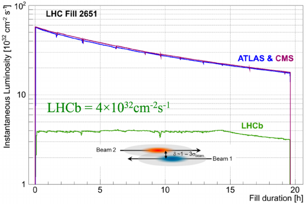
Forward spectrometer, optimised for b and c decays. $2 < \eta < 5$

- Excellent vertex resolution (weak decays)
- High-precision tracking before and after the magnet
- PID in broad range of momenta $3 < p < 150\,\mathrm{GeV}$
- Efficient trigger, including fully-hadronic final states

LHCb experiment

LHCb in run 1–2

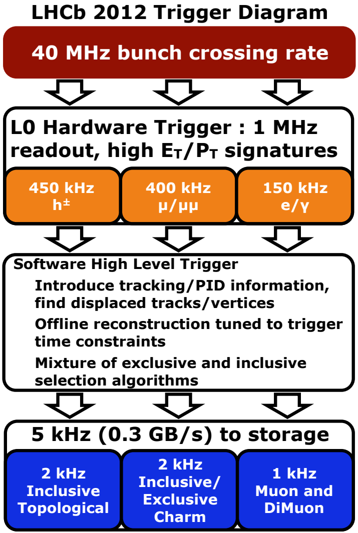




Forward spectrometer, optimised for *b* and *c* decays. $2 < \eta < 5$

- Excellent vertex resolution (weak decays)
- High-precision tracking before and after the magnet
- PID in broad range of momenta $3 < p < 150\,\mathrm{GeV}$
- Efficient trigger, including fully-hadronic final states

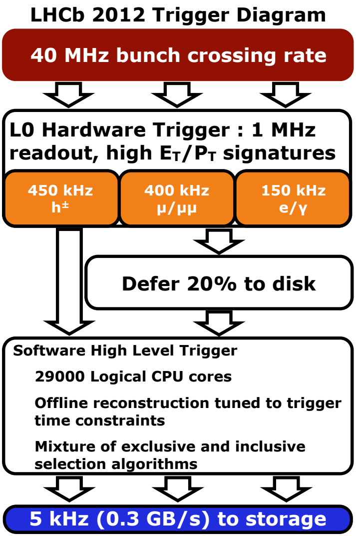
LHCb DAQ and trigger in Run 1–2



Trigger was continuously improved during Run 1-2 operation

- **Start of operation:**
  Storage bandwidth 5 kHz wrt 2 kHz in the design (additional b/w for charm)

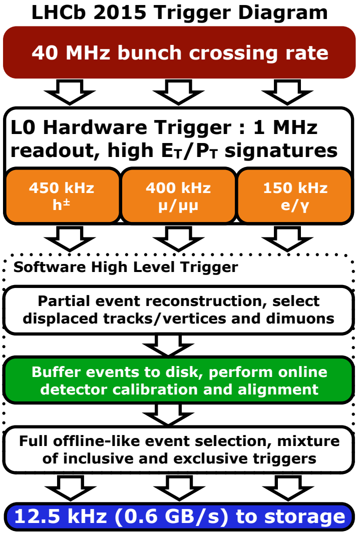
LHCb DAQ and trigger in Run 1–2



Trigger was continuously improved during Run 1-2 operation

- **Start of operation:** Storage bandwidth 5 kHz wrt 2 kHz in the design (additional b/w for charm)
- **2012:** Deferred trigger. Buffer 20% of bandwidth before HLT to disks (use interfill time)

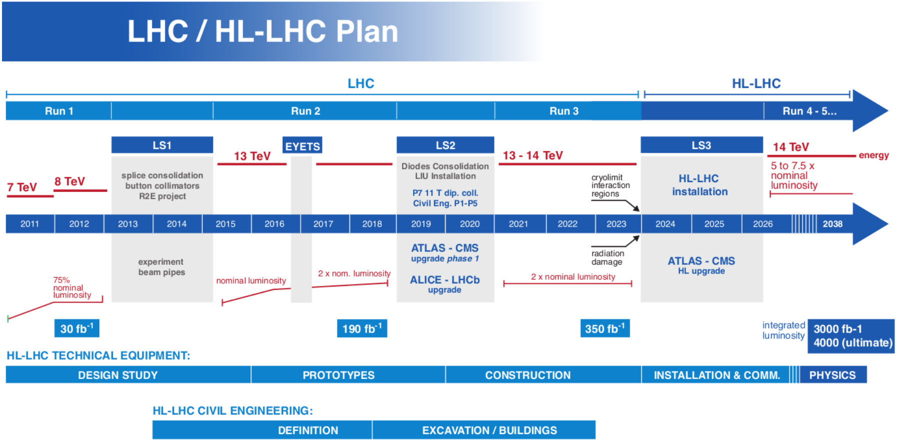
LHCb DAQ and trigger in Run 1–2



Trigger was continuously improved during Run 1-2 operation

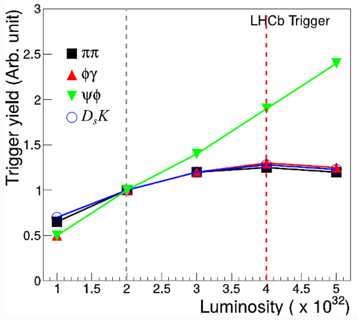
- **Start of operation:** Storage bandwidth 5 kHz wrt 2 kHz in the design (additional b/w for charm)
- **2012:** Deferred trigger. Buffer 20% of bandwidth before HLT to disks (use interfill time)
- **Run 2 (2015-2018):** Split HLT.
  - Buffer all HLT1 output to disk.
  - Run calibration and alignment.
  - Offilne-quality selections at the last stage of HLT.
  - Can run analyses on HLT2 output (`Turbo` stream)

LHC timeline



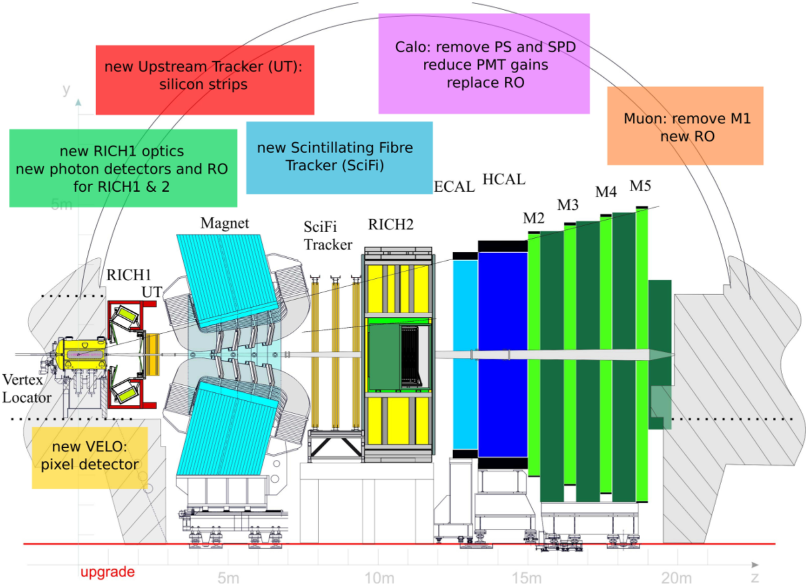
- LHC Run 2 finished in 2018
  - LHCb: $\int \mathcal{L} dt = 9\,\mathrm{fb}^{-1}$ collected in 2010-2018
- Long shutdown until 2021: upgrade of the machine and detectors
  - LHCb: major upgrade/replacement of the subsystems and readout

Goals of the LHCb upgrade

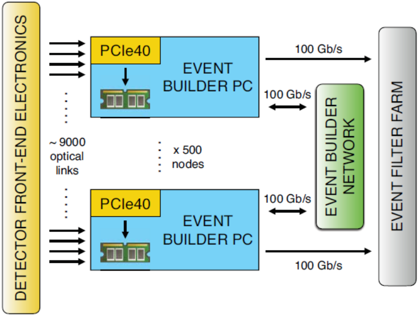


- Instantaneous luminosity: 4×10^{32} (Run 2) $\rightarrow$ $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
- L0 rate limit of 1 MHz saturates fully hadronic modes already in Run 2 (higher rate $\Rightarrow$ higher p_{T} thresholds)
  - The only solution: read full event at bunch-crossing rate and apply track reconstruction/IP selections.
- Upgrade/replace subsystems:
  - Cope with higher occupancy.
  - Faster/higher precision tracking
- Fully replace DAQ and trigger.

LHCb upgrade



Upgraded LHCb DAQ

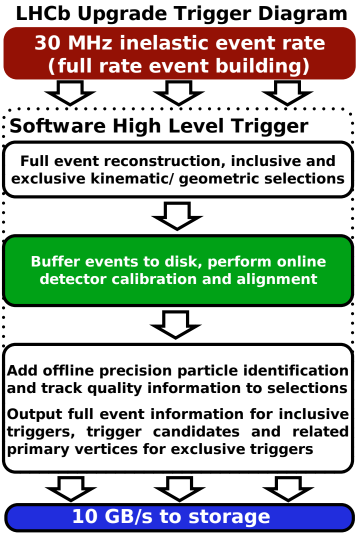


©CPPM

- New PCIe40 readout boards
  - 24 optical inputs, PCIe interface
- Event builder network using commercial technology
  - HDR InfiniBand© with remote direct memory access

- Event rate: 30 MHz non-empty bunch crossing
- Event size: ~ 100 kB
- Input bandwidth: 40 Tbit/s

Upgraded LHCb trigger



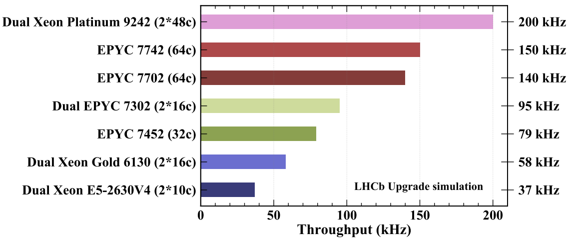
HLT1: [LHCb upgrade computing TDR]

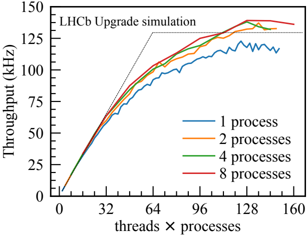
- Subdetector reconstruction:
  - VELO: clustering, tracking, vertex reconstruction
  - UT, SciFi: tracking
  - Muon: Hit-track matching
- Global event reconstruction:
  - Track fit (Kalman filter)
  - Reconstruction of secondary vertices
- Selections: [LHCb-PUB-2019-013]
  - Single displaced tracks
  - Two-track displaced vertices
  - Single displaced muons
  - Low-mass displaced two-muon vertices
  - High-mass dimuons

HLT1: performance on CPU

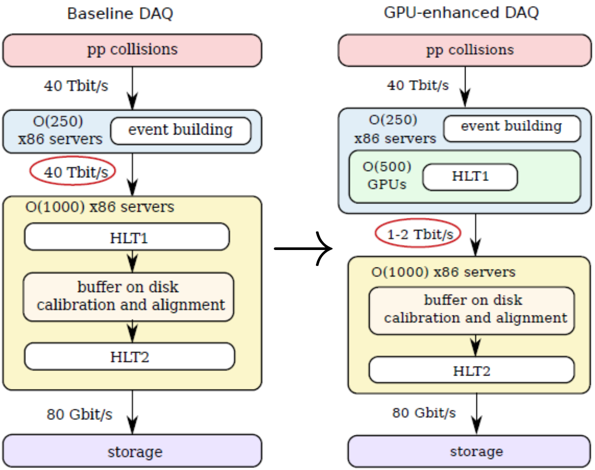
[LHCb-FIGURE-2020-001]

HLT1 track reconstruction performance is measured with various multicore/multithreaded CPU (Intel Xeon, AMD EPYC)





Name	Cores	Freq (GHz)	L3 (MB)	TDP (W)	#CPU, Boost?	HLT1 evts/s
EPYC 7302	16c/32t	3.0/3.3	128	155	2, no	95k
EPYC 7452	32c/64t	2.35/3.35	128	155	1, no	79k
EPYC 7702	64c/128t	2.0/3.35	256	200	1, no	140k
EPYC 7742	64c/128t	2.25/3.4	256	225	1, yes	150k
Xeon E5-2630V4	10c/20t	2.2/3.1	25	85	2, yes	37k
Xeon Gold 6130	16c/32t	2.1/3.7	22	125	2, yes	58k
Xeon Platinum 9242	48c/96t	2.3/3.8	77	350	2, yes	200k

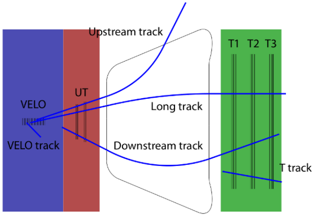


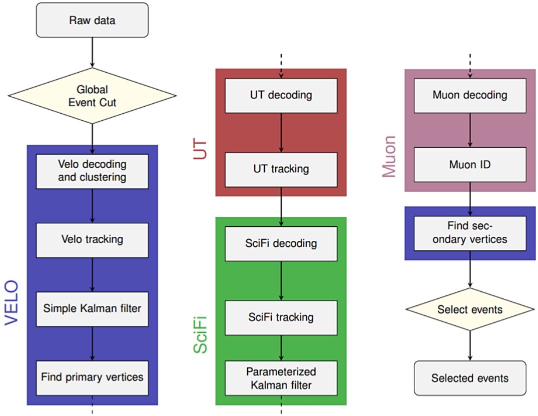
- Reduce network bandwidth between EB and filter farms
- Free up filter farm CPU for HLT2 only

Project currently under review, decision in the next months

- Framework for GPU-based execution of an algorithm sequence [GitLab repo]
- Based on `C++17, CUDA v10.2, boost, ZeroMQ`

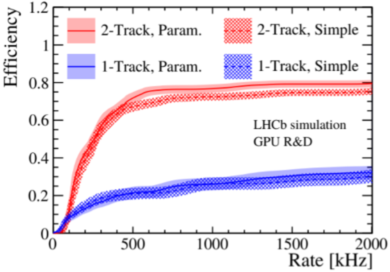
- 40 Tbit/ s input ⇒ ∼ 500 PCIe GPU cards to feed all data.
- Each GPU processes a slice of ∼ $O(1000)$ events in parallel





Trigger	Rate [kHz]
1-Track	215 ± 18
2-Track	659 ± 31
High-p_T muon	5 ± 3
Displaced dimuon	74 ± 10
High-mass dimuon	134 ± 14
Total	999 ± 38

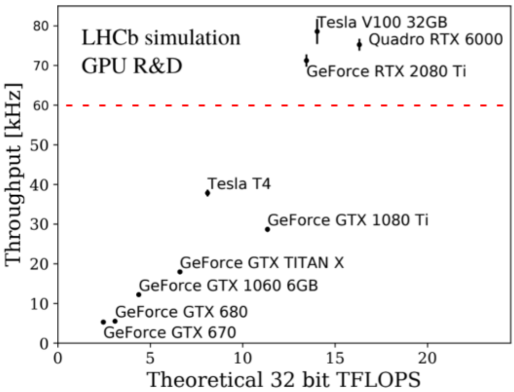
Rates of HLT1 lines on minimum bias events



Efficiency of 1-Track and 2-Track selections with $B_s^0 \to \phi\phi$ MC

Signal	GEC	TIS -OR- TOS	TOS	GEC × TOS
$B^0 \to K^{*0}\mu^+\mu^-$	89 ± 2	91 ± 2	89 ± 2	79 ± 3
$B^0 \to K^{*0}e^+e^-$	84 ± 3	69 ± 4	62 ± 4	52 ± 4
$B_s^0 \to \phi\phi$	83 ± 3	76 ± 3	69 ± 3	57 ± 3
$D_s^+ \to K^+K^-\pi^+$	82 ± 4	59 ± 5	43 ± 5	35 ± 4
$Z \to \mu^+\mu^-$	78 ± 1	99 ± 0	99 ± 0	77 ± 1

Efficiencies of HLT1 selection for benchmark signals



HLT1 throughput for various GPU cards.

60 kHz is the miminum requirement for 30 MHz imput rate and 500 GPU cards

HLT2 selections

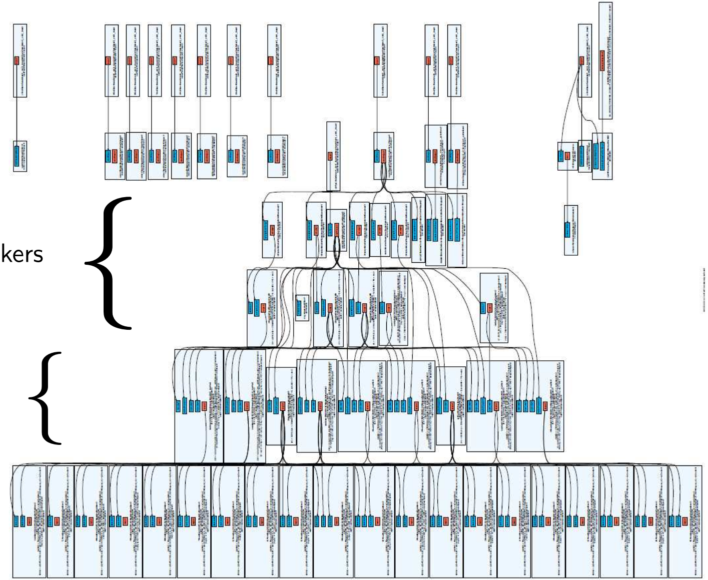
Exclusive HLT2 selections are being developed based on Run 2 selections.

Data producers

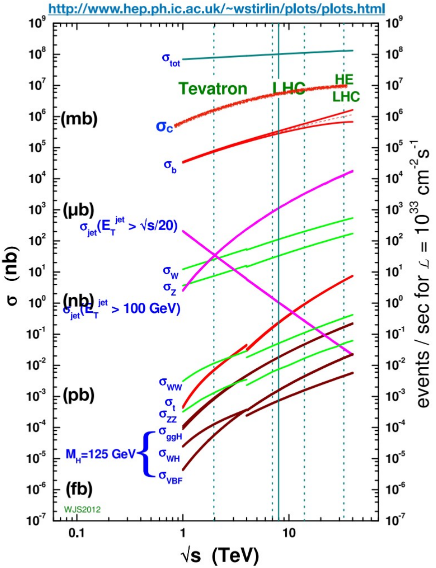
Basic particle makers

Shared combined makers/filters

Final signal makers/filters



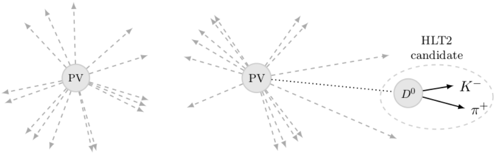
HLT2 output rate



- At $\mathcal{L} = 2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$:
  O(10) MHz charm +
  O(1) MHz beauty rate
- Output bandwidth limited to 10 GB/ s.
  Up to 100 kHz with full event size of 100 kB.
- Need to reduce the event size for higher rate

Persistency model [LHCb upgrade computing TDR]

Selective persistency: write out only the “interesting” part of the event.

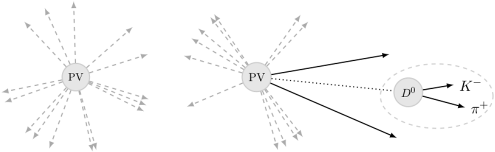


- `Turbo` stream:
  - Miminum output: only HLT2 signal candidates

Limitations: cannot refit tracks and PVs offline, rerun flavour tagging etc.
Advantage: Event size $O(10)$ smaller than RAW

Persistency model [LHCb upgrade computing TDR]

Selective persistency: write out only the "interesting" part of the event.



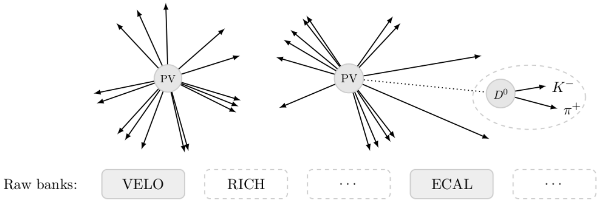
- `Turbo` stream:
  - Miminum output: only HLT2 signal candidates
  - Optionally: (parts of) *pp* vertex (*e.g.* "cone" around candidate for spectroscopy searches)

Limitations: cannot refit tracks and PVs offline, rerun flavour tagging etc.
Advantage: Event size $O(10)$ smaller than RAW

Persistency model

[LHCb upgrade computing TDR]

Selective persistency: write out only the “interesting” part of the event.

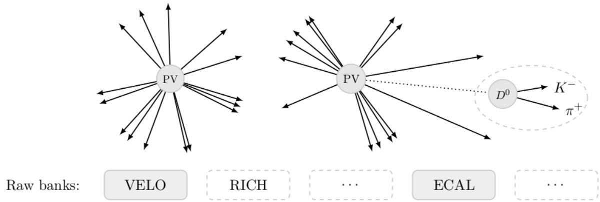


- Turbo stream:
  - Miminum output: only HLT2 signal candidates
  - Optionally: (parts of) *pp* vertex (*e.g.* "cone" around candidate for spectroscopy searches)

  Limitations: cannot refit tracks and PVs offline, rerun flavour tagging etc.
  Advantage: Event size $O(10)$ smaller than RAW
- FULL stream: all reconstructed objects in the event
  - + selected RAW banks

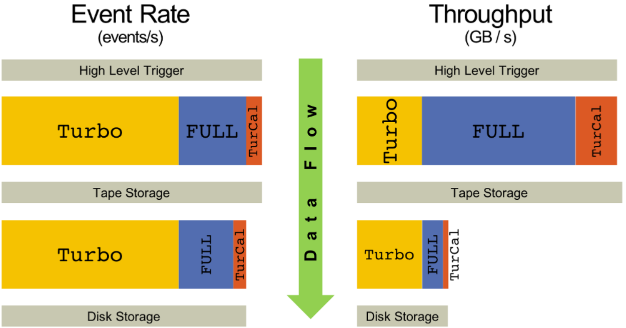
Persistency model [LHCb upgrade computing TDR]

Selective persistency: write out only the "interesting" part of the event.



- `Turbo` stream:
  - Miminum output: only HLT2 signal candidates
  - Optionally: (parts of) *pp* vertex (*e.g.* "cone" around candidate for spectroscopy searches)

  Limitations: cannot refit tracks and PVs offline, rerun flavour tagging etc.
  Advantage: Event size $O(10)$ smaller than RAW
- `FULL` stream: all reconstructed objects in the event
  - + selected RAW banks
- `TurCal` stream: HLT2 candidates and RAW banks
  Used for offline calibration and performance measurement



Rate and bandwidth to tape

stream	rate fraction	throughput (GB/s)	bandwidth fraction
FULL	26%	5.9	59%
Turbo	68%	2.5	25%
TurCal	6%	1.6	16%
total	100%	10.0	100%

Disk bandwidth

stream	throughput (GB/s)	bandwidth fraction
FULL	0.8	22%
Turbo	2.5	72%
TurCal	0.2	6%
total	3.5	100%

Summary

- LHCb upgrade ongoing, Run 3 to start in 2021
- Aim to increase instantaneous luminosity to $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ (5 times pre-upgrade).
  - Collect $50\,\mathrm{fb}^{-1}$ by 2023
- Major redesign of readout and trigger in LHCb upgrade.
- Remove hardware L0 stage, read out full detector at $30\,\mathrm{MHz}$ non-empty bunch crossing rate
- Readout and event builder farm: consolidation ongoing
- HLT filtering farm:
  - Architecture of split trigger with disk buffer, alignment and calibration $\Rightarrow$ offline-quality output.
  - Option to run HLT1 stage on GPU in the event builder farm under investigation.
  - HLT2 selections under development.
  - Increase physics output by moving most of bandwidth to `Turbo` stream (reduced size, no RAW information).