The Journal's name will be set by the publisher
DOI: will be set by the publisher


# LHCb trigger upgrade

Renato **Quagliani**[1,a] on behalf of the LHCb Collaboration

[1] *LPNHE/CNRS 4 Place Jussieu Tour 22, Paris, France*

**Abstract.** LHCb foresees a major upgrade for the 2021 LHC restart. At that time, LHCb will operate at five times larger luminosity than during Run II (2015-2018) data taking, reaching $\mathcal{L} = 2 \times 10^{33}\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ colliding $pp$ with a center of mass energy of 14 TeV. The designed bunch spacing of the colliding protons is 25 ns leading to a 40 MHz collision rate. The hardware based trigger strategy used during Run I (2010-2012) and Run II (2015-2018) will be completely removed and a fully software based trigger strategy will be adopted. Software applications performing event reconstruction and trigger selection will be executed in the online farm at the same rate of the visible interaction rate expected, *i.e.* 30 MHz. This document presents the recent developments for the online track reconstruction executed at collision rate and the current status of the trigger strategy for the LHCb upgrade.

## The LHCb experiment

LHCb is a high precision experiment at CERN taking advantage of the large cross section of $b$ and $c$ quark production in $pp$ collisions. Due to the $b$ hadrons production mechanism, LHCb is designed as a single arm forward spectrometer covering a pseudo-rapidity range between 2 and 5. A sketch of the current LHCb detector highlighting the various sub-detectors names and their functionalities is shown in Fig. 1.

During Run I LHCb has collected an integrated luminosity of $3\,\mathrm{fb}^{-1}$ at a center-of-mass energy of 7 (2011) and 8 (2012) TeV. In 2015 data taking was restarted for the Run II and will be continued until the end of 2018. After that, during the Long Shutdown 2 (LS2), the whole tracking and DAQ systems will be upgraded to enable running at a five times larger instantaneous luminosity. The Run III restart is foreseen in 2021.

LHCb's key objectives are the study of charge-partity (CP) violation effects and the study of the underlying New Physics (NP) phenomena. Although LHCb has been able to efficiently collect data in Run I and Run II measuring a wide range of processes in heavy flavor decays, both the granularity of the tracking system and the efficiency of the hardware trigger limit. The maximum instantaneous luminosity at which the detector can usefully operate. The critical limitation from the trigger is shown in Fig. 2, which illustrates the effective signal yield collected per year as a function of instantaneous luminosity of the current LHCb experiment. It can be seen that for non-muonic final states, the hardware trigger rapidly loses discriminating power above around $3{\times}10^{32}\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, plateauing the experiment's physics reach.

The LHCb upgrade key features are:

[a] e-mail: rquaglia@cern.ch

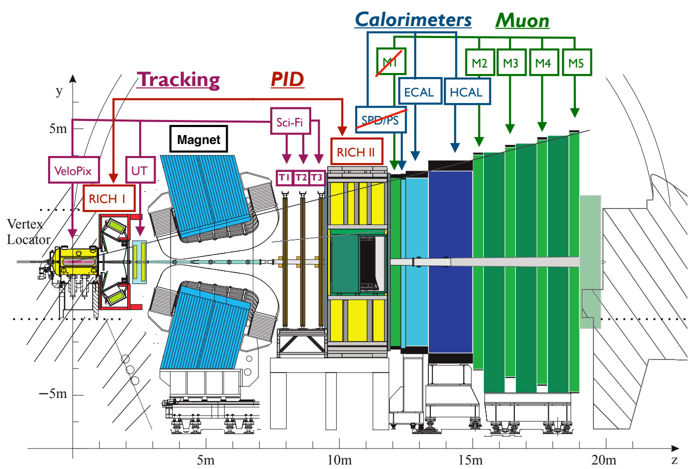


Figure 1: Sketch of the LHCb detector upgrade. In violet the tracking sub-detectors present in LHCb: VErtex LOcator (VeloPix), Upstream Tracker (UT) and the T-stations (SciFi). A dipole magnet with a bending power of 4 Tm is used to bend the charged particles originated from the $pp$ collision in the $x-z$ plane. The boxes with red slashed indicate the sub-detectors which are currently present for the Run I and RunII and will be removed for the upgrade.

- Detector upgrade: the whole detector will be read-out at 30 MHz and all detector information will be directly available in the trigger system. Key components such as VErtex LOcator (VELO), tracking stations and RICH will undergo substantial upgrades to maintain or improve the detector performances in an environment with significantly increased pile-up and track multiplicity.
- Trigger: the L0 hardware based trigger, currently reducing the input rate of 40 MHz down to the maximal read-out limit of the tracking sub-detectors of 1.1 MHz, will be completely removed. An asynchronous real-time data processing trigger executed on off-the-shelf commercial server architecture will be used.

## Towards the LHCb upgrade trigger strategy definition

The optimal performance of the LHCb detector relies on the capability of performing analysis in a clean environment, maximize the trigger selection efficiencies and fully exploit the luminosity delivered by the LHC.

The switch to a fully software based trigger, with access to all subdetector information at 40 MHz, enables these goals to be achieved [1]. To cope with the significantly increased data rates a major speed up of entire software, including data processing and reconstruction algorithms is required.

The upgrade trigger strategy is taking advantage of the experience gained in Run II which will be described briefly in the next paragraph highlighting the differences with the one foreseen for the Run III.

### Detector upgrade

The current VELO will be substituted by an hybrid pixel sensor detector (VeloPix) surrounding the $pp$ interaction region [2] and having a radial distance between the first active sensor and the beam-pipe

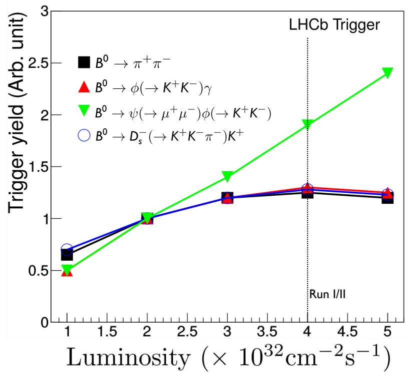


Figure 2: Evaluation of the trigger yields as a function of the instantaneous luminosity at LHCb for selected decay modes. The green triangles represents the trigger yields scaling as a function of the luminosity for the $B_s \to J/\psi\phi$ for which the muon L0 hardware trigger is used. For all the other modes, the hadronic L0 trigger selection is used. It is clear that the hardware (HW) based L0 trigger for non-muonic decays efficiency flattens out at higher luminosity.

of 5.1 mm. From tracking point of view, the single hybrid pixel sensor has the benefit of providing directly an $(x, y, z)$ measurement, leading to a more efficient, faster and simpler track reconstruction in the VELO. The tracks coming out from the interaction region and detected by the VELO can be described in good approximation by a straight line model since the magnetic field is negligible in the VELO region. The reconstructed VELO tracks are then used as input for other pattern recognition algorithms. The trackers up and downstream the magnet will be substituted [3]. The new Upstream Tracker (UT) will consists of four layers of large area silicon detector with a higher granularity improving the position resolution by a factor two and having a larger acceptance in the central region closer to the beam-pipe. This allows to extend the VELO tracks adding matching hits from the UT without a loss in performance due to the acceptance and to reconstruct an intermediate track with a momentum resolution of 20% which is a crucial information which allows to define small search windows when projecting the track in the tracker placed downstream the magnet.

The downstream tracker, currently made of 3 stations (4 layers each for a total of 12 layers) of composite technologies will be fully replaced by a homogeneous scintillating fiber detector (SciFi) made of 2.5 m long vertically oriented fibers read-out by silicon photomultipliers.[1]. This technology solution guarantees a fast read-out of the detector at 40 MHz, a single hit position resolution in all the SciFi acceptance of $\sigma_{x-z} \sim 100\ \mu\mathrm{m}$ guaranteeing the same final momentum resolution on tracks of the current tracker ($\frac{\sigma_p}{p} \simeq 0.6[\%] \cdot p\%$) but coping with the higher expected occupancy and allowing, thanks to the homogeneity of the detector, a simpler and faster detector treatment from a tracking point of view.

## Current and upgrade LHCb trigger strategy

The current LHCb trigger strategy is divided in 3 steps:

[1] Fibers are mirrored at $y = 0$, where y is the direction of the $\overrightarrow{B}$ field lines and z is the beam pipe direction.

- L0 hardware trigger:

  the 40 MHz input rate is reduced to 1.1 MHz based on the high transverse energy deposit found in the calorimeters or the high transverse momentum recorded in the muon system which are information available at the hardware level. The selection thresholds are tuned to reduce the *pp* collision rate to 1.1 MHz which is the rate at which the other sub-system (including the tracking system) are able to be read-out. In order to achieve the same rate reduction at higher luminosity, the selection thresholds has to be increased (especially for the the hadronic trigger) significantly limiting the physics capabilities of the experiment as shown in Fig. 2. For this reason the L0 hardware trigger will be completely removed in the upgrade trigger strategy.

- The High Level Trigger (HLT) event reconstruction is executed in three steps:

  1. HLT1: a partial event reconstruction is performed aiming for signatures of high momentum tracks and vertices in the event. Events are selected according to 1-track or 2-track criteria [4]. The reconstruction algorithms executed in the HLT1 are required to fit within the timing budget driven by the available computing resources. In the upgrade the HLT1 reconstruction will process the full LHC collision rate, but the reconstruction and selection algorithms will operate according to the same general principles as today.

  2. Events passing the selection are buffered on disk while the output of specific HLT1 selections is used to perform the real-time alignment and calibration of the entire detector [5] which has been introduced during Run II data taking and it will be kept for the upgrade.

  3. HLT2: full event reconstruction is performed, benefiting from the real-time detector alignment and calibration to deliver the best possible performance. The HLT2 aims at finding all the tracks in the event with the highest possible performance. In this stage inclusive and exclusive trigger selections are used.

- Persistency on disk: during Run II, LHCb has introduced the possibility to save directly high level reconstructed objects to disk which can be used to perform data analysis in real-time using the direct output of the online trigger selection and reconstruct. This has become possible thanks to the real-time alignment and detector calibration which guarantees to perform offline-quality selections on reconstructed objects directly in the online trigger system.

A sketch comparing the trigger strategy used for Run II and the one foreseen for the upgrade is shown in Fig. 3.

## LHCb upgrade track reconstruction

Efficient event classification based on the signal content requires the employment of a trigger strongly aligned to the offline selection requirement. The reconstruction sequence for the LHCb upgrade starts with preparing the data from the tracking sub-detectors. This step is crucial for tracking purposes: data format and data ordering affects in a crucial way the execution time of pattern recognition algorithm, thus the real-time software event reconstruction used for triggering purposes. The bandwidth and the available resources for the upgraded sub-detectors put stringent limits on what can be achieved at the read-out level. As a consequence, part of the sub-detector data pre-processing is perfomed at the software level:

- The hits used for tracking from the VeloPix detector are obtained after a software clustering finding and merging neighboring fired pixels (*VPClustering* algorithm).

- The hits used for tracking from the Upstream Tracker are currently assumed to come from the FPGA, their sorting and conversion from binary format to geometrical information is achieved at software level (*PrepareUTHits* algorithm).
- The hits used for tracking from the SciFi will arrive directly from the FPGA ordered in the same way tracking algorithms expects them to be (*PrepareFTHits* algorithm).

The amount of hits present in each tracking sub-detector for simulated events at the LHCb upgrade conditions are shown in Fig. 4. A total of around 11.000 hits (on average) has to be prepared to perform pattern recognition and track reconstruction.

Global Event Cuts (GEC) are applied to the combination of UT and SciFi hits since the amount of clusters (*i.e* hits) present in the sub-detectors are known a priori. The baseline of upgrade reconstruction sequence is configured to reject 10% (same rejection rate used also for RunI and RunII) of the most busy events, since the gain in physics performance from including them is not proportional to the resources required to reconstruct them.

The first algorithm executed in the HLT1 upgrade reconstruction sequence is dedicated to the reconstruction of straight line segments (VELO tracks) in the VeloPix (*PixelTracking*). VELO tracks are used to identify the Primary Vertices (PVs) in the event with the *PVFinding* algorithm and they are propagated to the Upstream Tracker.

Matching hits from the four detection layers in the UT are added to the input VELO tracks to form the Upstream tracks by the *Velo-UT* tracking algorithm. In this step one can pre-filter the input Upstream track by requiring those tracks to be displaced from the PV. If we assume that the impact of the integrated $B$ field between the VELO and the UT is negligible we can also require any input VELO track to have a minumum $p_T$ which determines the resulting field of interest search window in the UT ($\propto \frac{1}{min(p_T)}$). The preliminary event selection which is based on the presence in the event of at least one displaced high $p_T$ track. Therefore, the requirements on $p_T$ and displacement from the PV, if done in a early stage of the reconstruction sequence allow to achieve a large speed-up of the sequence

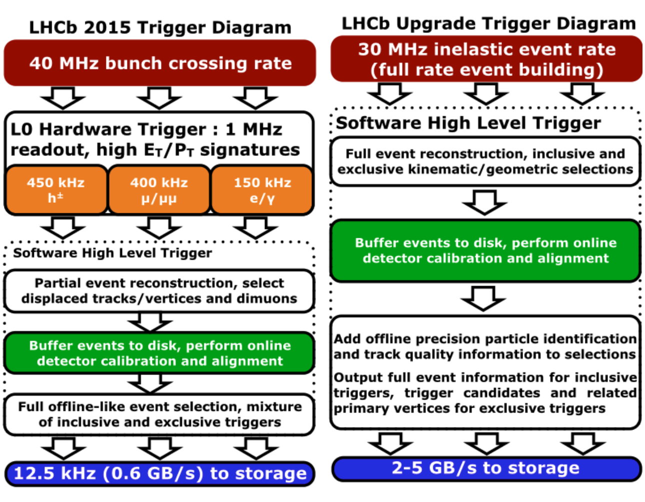


Figure 3: Comparison between the Run II and the upgrade trigger strategy.

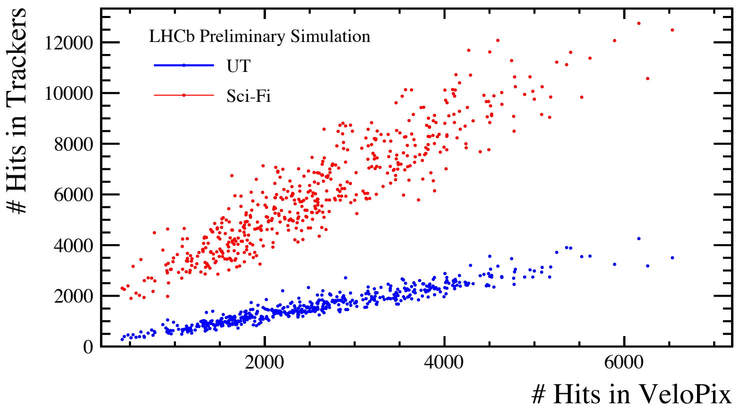


Figure 4: Upgrade condition detector occupancy differences between the various detectors. Each data point is an event, on the x-axis the number of hits present in the VeloPix detector, in the y axis the number of hits in the UT (blue) and the SciFi (red).

while keeping almost unchanged the trigger efficiencies for displaced track based physics selections (*i.e.* $b$ or $c$ hadrons trigger selection).

The momentum resolution using only the information from the VELO and the UT is 15-30 % depending on the track momentum. This preliminary information is used to pre-filter the Upstream tracks which are propagated to the tracker placed downstream the magnet.

An Hough Transform based tracking algorithm (*Forward Tracking*) is used to find extension of the Upstream track candidates in the SciFi. Tracks traveling from the UT to the SciFi experience the full integrated dipole magnetic field of LHCb: the search windows to open in the SciFi for each input Upstream track is proportional to size of the kick provided by the dipole magentic ($\propto \frac{q}{p_{\mathrm{T}}}$) of the track, where $q$ is the charge of the track and $p_{\mathrm{T}}$ its transverse momentum. High $p_{\mathrm{T}}$ tracks search implies small search windows significantly reducing the timing of the reconstruction sequence.

A sketch of the default fast reconstruction sequence that is executed at collision rate is shown in Fig. 5.

## Sequence throughput

LHCb has developed a parallel, multi-threaded framework to fully take advantage of modern multi-core architectures. Algorithms have been ported to the new framework and in this process also a first round of optimization has been performed updating the results presented in Ref.[6]. The performance of the tracking sequence has been benchmarked in terms of throughput (amount of events processed per second) using one node of the LHCb HLT farm employed during the Run II data taking period. The node has 20 physical with 2-way hardware threading.

The throughput of each algorithm composing the default HLT1 tracking sequence has been significantly improved since 2017 as shown in Tab. 1.

The *displaced track reconstruction* scenario, which is defined requiring a minimum $p_T$ in Velo-UT of 800 MeV/$c$ and a minimum $p_T$ in the Forward tracking of 1000 MeV/$c$ together with requiring for

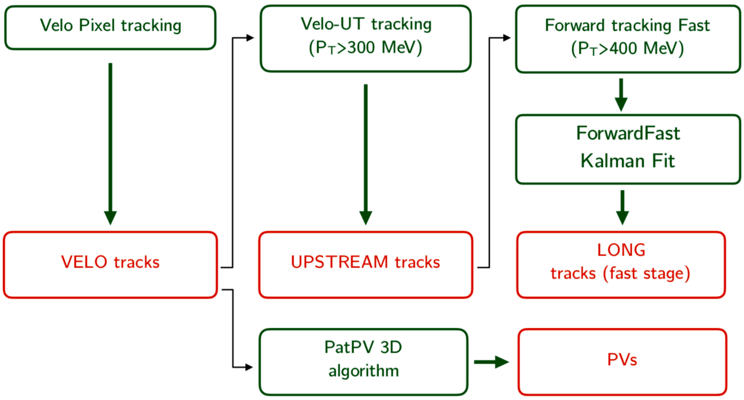


Figure 5: Schematic representation of the default HLT1 (fast) reconstruction sequence used for the LHCb upgrade. The VELO-UT tracking algorithm default search windows is based on the assumption that the input VELO tracks have $p_T$ >400 MeV/$c$, while the *Forward Tracking* algorithm requires for the opening search windows in SciFi the input tracks to have $p_T > 400$ MeV/$c$.

VELO tracks to propagate to the Upstream Tracker to be displaced from any primary vertex found in the event of at least 100 $\mu$m, has also been benchmarked [7].

The predicted throughput of the *displaced track reconstruction* sequence is shown in Fig. 6. The new multi-threaded framework allows to bring a 20 % gain with respect to the case in which the algorithms are not executed multi-threaded. The optimal performance is found when executing 2 processes with each process distributing the reconstruction sequence algorithm execution in 20 threads.

| HLT1 algorithm | Throughput (2017) | Throughput (2018) | Speed-up |
|---|---|---|---|
| Velo (Clustering + Tracking) | 10 kHz | 30 kHz | × 3 |
| Prepare UT data | 43 kHz | 88 kHz | × 2 |
| Prepare SciFi data | 22 kHz | 86 kHz | × 2 |
| Velo-UT tracking | 66 kHz | 146 kHz | ×2.2 |
| PV finding | 32 kHz | 91 kHz | ×2.8 |
| Forward tracking | 15 kHz | 19 kHz | ×1.2 |
| HLT1 (total) | 3,5 kHz | 7,5 kHz | × 2.1 |

Table 1: Comparison between the throughput measured for the biannual upgrade document review and the current state of the art of the default HLT1 reconstruction sequence.

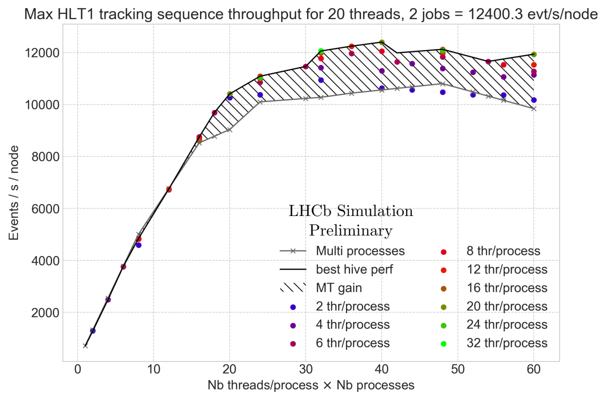


Figure 6: HLT1 tracking sequence throughput scan over different parallelization level as given in the legend. The throughput peak performance is found to be 12400 evt/s/node for running 2 jobs and 20 threads per job. This peak performance can be compared to the "Multi processes" scenario, which is the maximum performance which could be achieved without multi-threading.

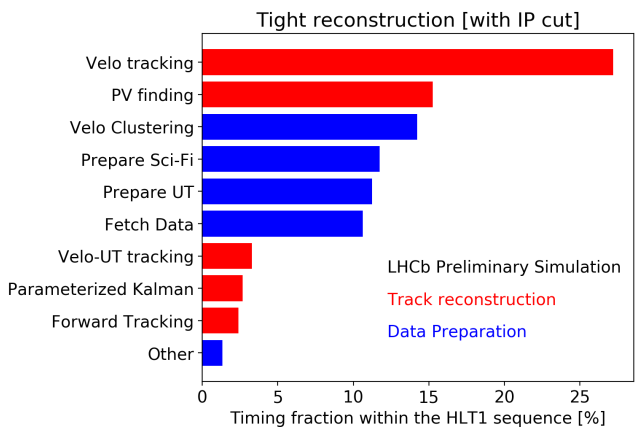


Figure 7: Timing division of the HLT1 track reconstruction sequence for the peak throughput configuration. In red the timing fraction of pattern recognition algorithms, in blue the timing fraction of data preparation algorithms.

The maximal achieved throughput is 12.4 kHz for the benchmarking machine, corresponding to roughly 12.4 MHz for the assumed (with the current resources for the LHCb upgrade computing) EFF in total.

The breakdown in resource usage by the reconstruction algorithm at the predicted peak throughput performance is shown in Fig. 7 for the full reconstruction sequence.

As can be seen, around 50% of the overall reconstruction resources are spent in the data preparation of the sub-detectors performing decoding, clustering, and sorting the raw hits in the tracking sub-detectors [2].

The maximum achievable throughput as a function of the track displacement and transverse momentum criteria applied in the reconstruction sequence is shown in Fig. 8.

There is a stronger dependence on the transverse momentum than on the displacement criteria, although it should be kept in mind that since this sequence is slower than the 30 MHz target, even small slowdowns on this plot correspond to large slowdowns with respect to the desired sequence. The throughput of the sequence and the timing cost of the individual algorithms is also studied as a function of the detector occupancy (see Figs. 9). The overall non-linear evolution of the sequence throughput with occupancy is driven by the highly non-linear behavior of the Forward Tracking. On the other hand it is encouraging that all the other algorithms behave linearly.

## Conclusions

In this document, the state of the art of the HLT1 reconstruction sequence for the LHCb upgrade is presented. LHCb has introduced a new multi threaded framework in the last years and for the first time we here present the results showing that an around 20 % throughput gain is achieved when executing the HLT1 sequence multi-threaded with respect to single thread. Furthermore, an update to the results

[2]In the current (Run II) HLT1 reconstruction sequence, the data preparation step takes under 10% of the total HLT1 resources but with a 30 times smaller input rate with respect to the upgrade conditions.

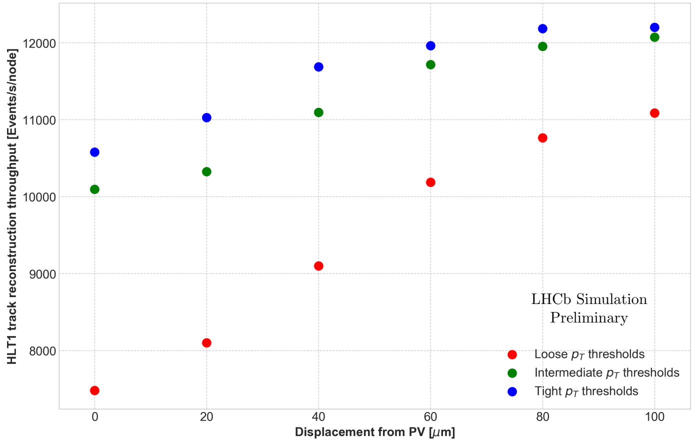


Figure 8: HLT1 tracking sequence throughput depending on the track reconstruction transverse momentum requirements and the filtering of the VELO tracks propagated to the Upstream Tracker based on their displacement from the primary vertices. The red points correspond to the throughput scan over the displacement from PVs requirements for the default tracking reconstruction configuration for which the min $p_T$ required in the Velo-UT (Forward tracking) algorithm is 300 (400) MeV/$c$. The green points correspond to a min $p_T$ in Velo-UT (Forward tracking) algorithm to be 600 (800) MeV/$c$ while the blue ones correspond the a min $p_T$ in the Velo-UT (Forward tracking) algorithm of 800 (1000) MeV/$c$.

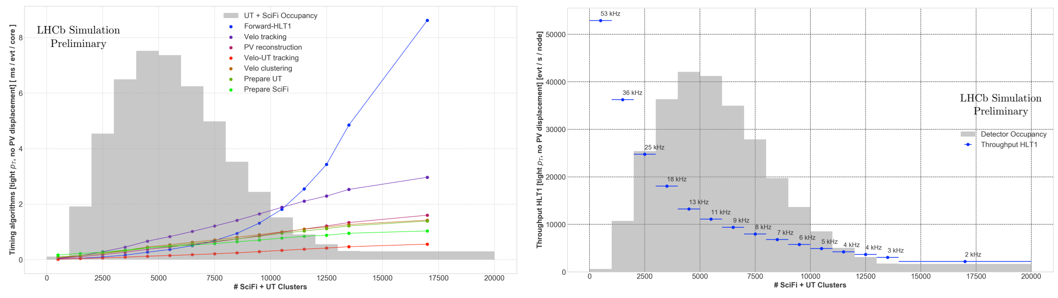


Figure 9: (Left) Timing of various algorithms composing the HLT1 sequence as a function of the detector occupancy. (Right) Throughput dependence of the HLT1 sequence as a function of the detector occupancy. Events having more than 11000 hits in the SciFi and UT are rejected by the Global Event Cut.

shown in [6] is presented. Since then, a factor 2 in throughput has been achieved without any loss in physics performance. We also present here the *displaced track* reconstruction sequence for the upgrade trigger, which allows extra factors in terms of throughput to be achieved. This sequence has some implications on the physics program of the LHCb experiment, expecially for the prompt physics programme. Further optimizations are ongoing with the goal of reaching the 30 MHz throughput.

## Acknowledgements

Renato Quagliani acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 724777 "RECEPT".

## References

[1] Tech. Rep. CERN-LHCC-2011-001. LHCC-I-018, CERN, Geneva (2011), `https://cds.cern.ch/record/1333091`

[2] *LHCb VELO Upgrade Technical Design Report* (2013), lHCb-TDR-013

[3] *LHCb Tracker Upgrade Technical Design Report* (2014), lHCb-TDR-015

[4] J. Albrecht, V.V. Gligorov, G. Raven, S. Tolk, the Lhcb Hlt project, Journal of Physics: Conference Series **513**, 012001 (2014)

[5] R. Quagliani, L. Collaboration, Journal of Physics: Conference Series **762**, 012046 (2016)

[6] R. Aaij, J. Albrecht, B. Couturier, S. Esen, M. De Cian, J.A. De Vries, A. Dziurda, C. Fitzpatrick, M. Fontana, L. Grillo et al., Tech. Rep. LHCb-PUB-2017-005. CERN-LHCb-PUB-2017-005, CERN, Geneva (2017), `https://cds.cern.ch/record/2244312`

[7] M. De Cian, A. Dziurda, V. Gligorov, C. Hasse, W. Hulsbergen, T.E. Latham, S. Ponce, R. Quagliani, H.F. Schreiner, S.B. Stemmle et al., Tech. Rep. LHCb-PUB-2018-003. CERN-LHCb-PUB-2018-003, CERN, Geneva (2018), `https://cds.cern.ch/record/2309972`