PROCEEDINGS OF SCIENCE

A comprehensive real-time analysis model in Run 2 at the LHCb experiment

Biljana Mitreska*, **on behalf of LHCb the collaboration**
The University of Manchester
E-mail: biljana.mitreska@cern.ch

A real-time data processing strategy is proposed for high-energy physics experiments, and its implementation at the LHCb experiment is presented. The reduced event model allows the signal candidate firing the trigger to be persisted, along with an arbitrary set of other reconstructed or raw objects from the event. This allows for higher trigger rates for a given output data bandwidth and reduction of the storage used/needed, when compared to the traditional model of saving the full raw detector data for each trigger, whilst accommodating inclusive triggers and preserving data mining capabilities. The gains in physics reach and savings in computing resources already made possible by the model during Run 2 of the experiment are discussed.

European Physical Society Conference on High Energy Physics - EPS-HEP2019 -
10-17 July, 2019
Ghent, Belgium

PoS(EPS-HEP2019)113

---

[*] Speaker.

1. Introduction

The LHCb experiment [1] [2] allows to probe flavour physics at the LHC by focusing on particles that contain b and c quarks, helps understanding the Standard Model(SM) and searches for new particles and couplings. As the production of the b-mesons is concentrated in the forward direction the LHCb detector was constructed as a single arm forward spectrometer which covers the pseudorapidity region $2 < \eta < 5$. Its tracking system includes the Vertex Locator (VELO) which is situated in the region of the proton-proton interaction. The VELO is a silicon detector that provides reconstruction of particle trajectories passing through, allowing to distinguish primary (PV) from secondary vertices (SV) [1] at the same time. The rest of the tracking system consists of the silicon-strip detector (TT) located upstream of the dipole magnet, three stations of silicon-strip detectors (IT) located closer to the beam pipe and straw drift tubes (OT) in the outer region, placed downstream of the magnet. The tracking system reconstructs the trajectories of the particles (tracks) and determines their momentum. The RICH (Ring Imaging Cherenkov) detectors provide particle identification using Cherenkov radiation. A large dipole magnet curves the paths of charged particles and enables calculation of their momenta. Their energy is measured by an electromagnetic (ECAL) and hadronic calorimeter (HCAL). The muon system detects muons in each event and measures their properties. The experiment has a trigger system that decides whether to accept or reject a given event for further use offline.

The trigger and data processing framework was redesigned during 2013-2015 to enable real-time detector alignment and calibration procedure (not being the case for Run 1 (2009-2013)). The main goals to confront here were to have offline quality reconstruction to avoid offline processing and have best quality data out of the alignment and calibration. LHCb has extended its physics programme during Run 2 (2015-2018) to electroweak, soft QCD and heavy-ion physics and this was made possible due to the trigger system and real-time reconstruction [3], which is responsible for reducing the rate of collisions saved for offline analysis. With the periodically increasing centre-of-mass energy of the proton-proton beams and instantaneous luminosity the challenge of providing enough computational resources for storing data for offline analysis persists. The disk space required for saving the events by the experiment is given by the product of the running time of the experiment and the trigger output bandwidth which is defined as

$$\text{Bandwidth [MB/s]} \propto \text{Trigger output rate [kHz]} \times \text{Average event size [kB]}.$$

Since there is no possibility of reducing the size of the raw event information a flexible, reduced event format is needed. The reduced event size contains a subset of reconstructed information that is permanently saved. Following the redesign of the LHCb trigger physics analyses can use information directly from the trigger. With the start of Run 2 all this has been implemented during data taking which resulted in the new real-time analysis model for LHCb.

2. The Run 2 trigger model

The LHC provides proton-proton collisions at a rate of 30 MHz. In order to be capable to filter out that many events and to place the large amount of output data in the available computing resources, a three-stage trigger system is used as shown on Fig. 1 [4]. The level-0 hardware trigger

---

[1] Primary vertex referring to the proton-proton interaction vertex and secondary as the vertex of decay particle.

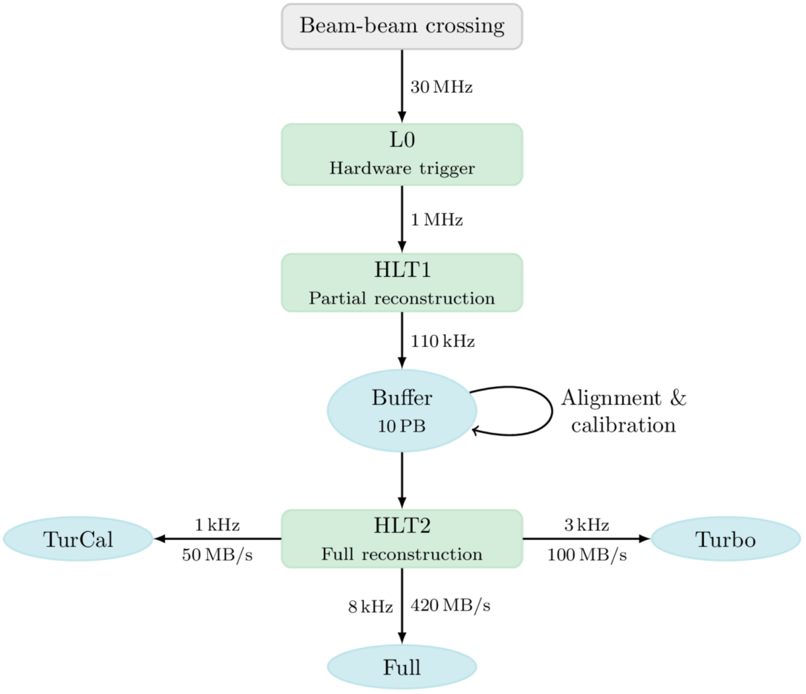

Figure 1: A schematic view of the LHCb trigger during Run 2.

(L0) uses information from the calorimeter and muon systems to select events at a rate of about 1 MHz. The selected events are sent to the two-stage High Level Trigger (HLT). The HLT1 is representing the first software stage of the trigger system. It uses tracking and calorimetry information to perform partial reconstruction of trajectories of charged particles with transverse momentum (p_T) larger than 500 MeV/c. A precise primary vertex reconstruction is also performed at this stage. HLT1 uses the muons in the LHCb detector as they can be precisely detected and allows their identification to be made at this stage as well.

After the partial reconstruction in HLT1 the raw information for each selected event is written on a 10 PB buffer at a rate of 110 kHz. Up to two weeks of consecutive HLT1 data taking is possible with a buffer of this size. The buffer allows to use the data written from HLT1 to select samples for aligning and calibrating the detector [5] [6]. The alignment procedure is based on the Kalman filter [7] method with a minimum χ^2 algorithm and runs automatically at the beginning of each fill. The procedure corrects for any misalignments and writes the new alignment constants (translations and rotations along/around x, y, z axis) for a set of detector elements in a new database by performing an update on the alignment constants if there are significant variations (Fig. 2). The alignment is evaluated for the full tracking system at LHCb (VELO, TT, IT, OT, Muon chambers) and the calibration is done for the RICH and the OT. The newly computed constants are stored for further use in the second stage of the software trigger (HLT2). The HLT2 performs a full event reconstruction using the newly calculated alignment and calibration constants. The full reconstruction consists of track reconstruction of charged particles, reconstruction of neutral particles and particle identification (PID). Using the selections in HLT2 the events ready for offline storage are written at a rate of 12.5 kHz. The trigger lines in HLT2 are assigned to one or more output streams which are sent

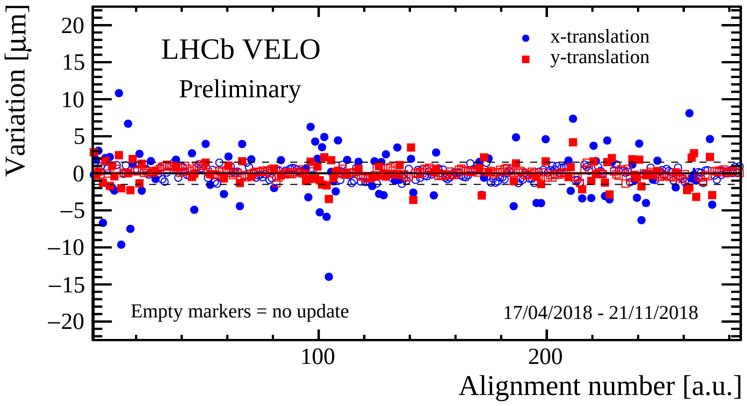

Figure 2: Stability of the alignment of the VELO halves during all Run 2 fills. Each point is obtained running the online alignment procedure and shows the difference between the initial alignment constants (the ones used in the previous fill) and the new ones computed by the alignment. The alignment is updated if the variations are above the horizontal lines at ± 2 μm.

offline. In the specific output stream a subset of the raw banks is created which has the information created by different sub-detectors and by the trigger itself. The set of streams used for persisting the selected events are the full stream, Turbo and the TurCal. Events that are assigned to the full stream are persisted with the complete set of raw banks. Events assigned to the Turbo stream persist the reconstructed objects used to make the trigger decision which is the reduced event format. The TurCal calibration stream keeps both the reduced and full formats and it is used for centralised evaluation of the track reconstruction and particle identification(PID) performance.

The multi-stage trigger increases the available event processing time from one stage to the next one and each stage requires more computing power than the previous one. The addition of a disk buffer between the two HLT stages allows the alignment and calibration to be evaluated in a few minutes time as compared with the few hours needed to align and calibrate the detector in Run 1. This whole procedure can be referred as real time with its exact definition as the interval between a collision occurring and the point at which the corresponding event must be either discarded or sent offline for permanent storage. With an offline-quality reconstruction in the final trigger stage (HLT2), it is no longer necessary to run another reconstruction offline. That allows the objects created by trigger selections to be written out to the permanent storage and physics measurements to be performed with them. This approach eliminates possible processing requirements offline and reduces the output bandwidth, if the relevant subset of the reconstruction is smaller than the raw event. In the next section the turbo data model that provides this will be reviewed.

2.1 The Turbo data processing model

In the proton-proton collisions at LHCb around 40 tracks on average are associated with a primary vertex. Usually not all the tracks are being used for analysis (2-6 to reconstruct a decay), so the persisted event size is employed by discarding reconstructed objects not needed in the offline analysis. During Run 2 three new developments to the Turbo processing model were introduced: the standard Turbo model, selective persistence and complete reconstruction persistence. Containers of physics objects are serialised per event into raw banks and an application developed by LHCb, TESLA [8], serves as a tool for transforming the HLT2 output into analysis usable format.

1. **The standard Turbo model** uses trigger lines which define exclusive selections where certain objects are saved. For each line there is a reconstructed decay for which the set of all tracks and neutral objects, calorimeter and PID information and decay vertices that form the candidate, together with all of the reconstructed primary vertices in the event, are saved.

2. **The selective reconstruction persistence** allows to specify additional information to be stored and avoid storing the complete reconstruction event.

3. **The complete reconstruction persistence** allows to persist information on the whole reconstructed event and drop the raw event information which saves an additional disk space.

An example that illustrates the different types of persistence is a trigger that reconstructs and selects $D^0 \to K^-\pi^+$ decay. Fig. 3 shows the objects that are stored when using the Turbo, the selective and

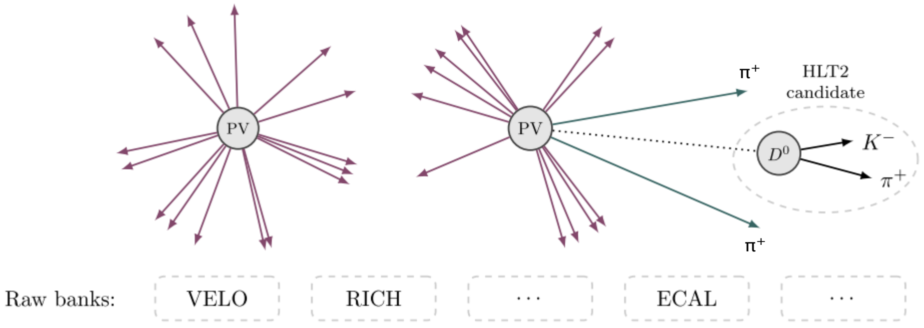

Figure 3: Sketch of selecting the $D^0 \to K^-\pi^+$ using different persistence methods: standard Turbo model(black) and the additionally persisted objects with the selective persistence (green) and complete persistence (red).

the complete reconstruction persistence. If complete reconstruction persistence is enabled for this line, the whole reconstructed event will be stored. With selective persistence, additional objects are instead specified explicitly, such as all charged pions that are associated to the same PV as the D^0 and form D^* candidate with discarding the D^* and saving the pion needed. This however allows for adding similar selections for other particles which is able to support the spectroscopy measurements at LHCb. Using data collected in 2018 the average event sizes for different persistence methods have been obtained (Table 1). The Turbo reduces the event size by 10 times as compared with the raw event.

An other implementation at LHCb since 2018 is the selective raw persistence. Here persistence is applied on the raw banks if they fulfill a condition of firing a trigger line. If the raw banks are not requested by any firing trigger line they are being discarded.

Persistence method	Average event size (kB)
Turbo	7
Selective persistence	16
Complete persistence	48
Raw event	69

Table 1: Average event sizes for different persistence methods, measured on data collected in 2018.

3. Run 2 achievements and Run 3 prospects

The Turbo model has been very successful during Run 2. It has been used by many analyses and provided faster results from analysis. Physics analyses that use the Turbo model include charm and J/ψ cross section measurements [9] [10], the discovery of new ground-state and excited charm baryons [11] [12], the discovery of CP violation in charm [13], J/ψ production in jets [14] and the searches for dark photons [15].

During Run 3 (2021-2023) the LHCb experiment will be upgraded so that it will receive higher luminosities by a factor of 5. The detector will have a full detector readout at 40 MHz with a full software trigger system. This means that most of the analyses will be oriented to use the Turbo model but with a caution not to lose valuable information when reducing the event size. Therefore a safe decision is to discard objects that are not related to the signal trigger object such as objects originating from other primary vertices. On the other hand primary vertices from which associated information should not be persisted can be identified using a minimum impact parameter cut. Very challenging in terms of reducing the event size are the inclusive trigger selections. Several selections can help in reducing the event size in the inclusively selected events such as rejecting a track from a poor quality vertex or rejection of objects that have been identified by a multivariate algorithm trained to distinguish uninteresting objects from those associated to the signal in the inclusive selection.

4. Conclusions

The real-time alignment and calibration procedure between the two software trigger stages introduced in 2015, the speeding up in the reconstruction software and availability of larger computing resources allowed to develop the Turbo model for the LHCb experiment. With this, offline-quality signal candidates are persisted directly from the trigger for later analysis. The updated Turbo model has already provided a 50 % reduction in bandwidth in comparison with saving the raw event. The model is now capable of supporting the entirety of the experiment's broad research programme. Given the large increase in instantaneous luminosity and trigger efficiency expected in Run 3, this model can use everything that has been developed during Run 2 and gives support in continuation of the future physics measurements at LHCb.

References

[1] LHCb Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.

[2] LHCb Collaboration, *LHCb detector performance*, *Int. J. Mod. Phys.* **A30** (2015) 1530022 [hep-ex/1412.6352].

[3] LHCb Collaboration, *A comprehensive real-time analysis model at the LHCb experiment*, *JINST* **14** (2019) P04006.

[4] LHCb collaboration, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, *JINST* **14** (2019) no.04, P04013.

[5] A. Dzuirda, L. Grillo, F. Polci and M. Sokoloff, *Real-time alignment and reconstruction: performance and recent developments at the LHCb experiment*, *IOP Journal of Physics: Conf. Series* **1085** (2018) 042003.

[6] S. Borghi, *Novel real-time alignment and calibration of the LHCb detector and its performance*, *Nucl. Instrum. Methods Phys. Res. A* **A 845** (2017) 560-564.

[7] W.D. Hulsbergen, *The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment*, *Nucl. Instrum. Methods Phys. Res. A* **600** (2009) 471-477.

[8] LHCb Collaboration, *Tesla: an application for real-time data analysis in High Energy Physics*, *Comput. Phys. Commun.* **208** (2016) 35 [hep-ex/1604.05596].

[9] LHCb Collaboration, *Measurement of forward J/ψ production cross sections*, *JHEP* **10** (2015) 172, Erratum ibid. 05 (2017) 063 [hep-ex/1509.00771].

[10] LHCb Collaboration, *Measurements of prompt charm production cross sections in pp collisions at $\sqrt{s}$ = 13 TeV*, *JHEP* **03** (2016) 159, Erratum ibid. 09 (2016) 013, Erratum ibid. 05 (2017) 074 [hep-ex/1510.01707].

[11] LHCb Collaboration, *Observation of five new narrow Ω_c^0 states decaying to $\Xi^+ K^-$*, *Phys. Rev. Lett.* **118** (2017) 182001, [hep-ex/1703.04639].

[12] LHCb Collaboration, *Observation of the doubly charmed baryon Ω_{cc}^{++}*, *Phys. Rev. Lett.* **118** (2017) 112001, [hep-ex/1707.01621].

[13] LHCb Collaboration, *Observation of CP Violation in Charm Decays*, *Phys. Rev. Lett.* **122** (2019) 211803, [hep-ex/1903.08726].

[14] LHCb Collaboration, *Study of J/ψ production in jets*, *Phys. Rev. Lett.* **118** (2017) 192001, [hep-ex/1701.05116].

[15] LHCb Collaboration, *Search for dark photons produced in 13 TeV pp collisions*, *Phys. Rev. Lett.* **120** (2018) 061801, [hep-ex/1710.02867].