New software technologies in the LHCb Simulation

Gloria Corti[1], **Adam Davis**[2], **Michal Kreps**[3], **Michał Mazurek**[1], **Dmitry Popov**[1,4], **Benedetto Gianluca Siddi**[5]

[1]CERN, Espl. des Particules 1, 1211 Meyrin, Switzerland
[2]University of Manchester, Oxford Rd., M13 9PL Manchester, United Kingdom
[3]University of Warwick, Coventry CV4 7AL, United Kingdom
[4]University of Chinese Academy of Sciences, 52 Sanlihe Rd., Xicheng District, Beijing, China
[5]Università degli studi di Ferrara, Via Savonarola 9, 44121 Ferrara, Italy

E-mail: michal.mazurek@cern.ch

Abstract. Monte Carlo simulations are key to the design and commissioning of new detectors as well as the interpretation of physics measurements. The amount of simulated samples required for the Run 3 physics program of the LHCb experiment will increase significantly from 2022 onward to match the increase in the amount of data collected with respect to Run 1 and 2 operation. A new version of the LHCb Gauss simulation framework has been developed to better accommodate new simulation techniques and software technologies to produce the necessary samples within the computing resources allocated for the next few years. It provides the LHCb specific functionality while the generic simulation infrastructure has been encapsulated in an experiment-independent framework, Gaussino. The latter combines the Gaudi core software framework and the Geant4 simulation toolkit and fully exploits their multi-threading capabilities. Fast simulation interface is the latest feature being developed in Gaussino to interact with Geant4 giving the possibility of replacing its detailed description of physics processes with an extensive palette of fast simulation models for a specific LHCb sub-detector. A facility to ease the production of training datasets for fast simulations models has also been introduced.

1. Introduction

The LHCb experiment [1, 2] at the Large Hadron Collider (LHC) was originally designed for the study of particles containing b and c quarks. The detector is a single-arm spectrometer with a forward angular coverage. The experiment is undergoing its first major upgrade. It will resume operation in Run 3 with higher luminosity and trigger rates than in previous runs with the aim of collecting a higher amount of physics data. Processing events in these new conditions will be very challenging for software and computing systems. New software is being deployed for all data processing of the experiment, including in the underlying LHCb core software framework, Gaudi [3, 4]. One of the major improvements introduced in Gaudi is an inter-event-based parallelism. The LHCb simulation framework, Gauss, had to be adapted accordingly, with the additional constraint that it also relies on external simulation libraries that have their own multi-threading implementation, as in the case of the Geant4 [5] toolkit. An additional challenge for the Gauss framework arises from the need to support at the same time the production of samples for Run 1 and Run 2 physics analyses as well as those for Run 3. Different technologies are used for the geometry description of the detector in the different

Journal of Physics: Conference Series **2438** (2023) 012108 doi:10.1088/1742-6596/2438/1/012108

periods, and legacy applications are used for the simulated data processing after it is produced with Gauss.

The LHCb simulation team decided to take the occasion to consolidate the Gauss software and extract all generic components into a new core simulation framework, called Gaussino. This allows easier prototyping and testing of new technologies where only the core elements are affected. The new version of Gauss exploits the Gaussino infrastructure and will provide all needed and specific functionality for LHCb. The multi-threaded approach in Gauss-on-Gaussino, as we refer to for the version of Gauss based on Gaussino, has already been shown to be able to simulate more events in time with respect to the current version [6].

Model	Generation	Decay	Propagation	Migration to Gauss-on-Gaussino
ReDecay [7]	✓	✓	✓	done
ParticleGun [8]	✓	✓	✓	done
SplitSim [8]	✓	×	✓	done
RICHless [8]	×	×	✓	in progress
TrackerOnly [8]	×	×	✓	in progress
Lamarr [9]	×	×	✓	to be done
Point library [10]	×	×	✓	to be done
GANs [11]	×	×	✓	to be done

Table 1. Fast simulations available in LHCb

The improvement is nevertheless not sufficient to produce all necessary events within the computing resources allocated for the next few year [12, 13]. Further optimization of the time spent in the detector simulation, where the modelling of particles propagation through matter and of the physics processing occurring wherein, is needed. Many fast simulation techniques have already been implemented in Gauss to tackle the problem of producing simulation samples for Run 2 [8]. Additional models are being developed and validated. An overview on the status of their migration to Gauss-on-Gaussino is given in Table 1. The Point library [14] and a model based on generative adversarial networks (GANs) [11], designed to replace the Geant4 detailed description of the physics processes occurring in the LHCb calorimeters [15], require a dedicated FastSimulation interface in Gaussino in order to fully exploit the mechanisms available in Geant4 for this purpose. A description of this interface is given in the next section.

2. Fast Simulation Interface in Gauss-on-Gaussino

The aim of the FastSimulation interface in Gaussino is to minimize the work needed to implement fast simulation models in Gauss-on-Gaussino, and at the same time guarantee the integrity of the simulated data. Following the design already adopted elsewhere in Gaussino, the FastSimulation interface consists of factories ensuring that relevant Geant4 objects are configured properly and at the right moment in time when running the application.

The performance of the interface has been tested with two abstract fast simulation models, `ImmediateDeposit` and `ShowerDeposit`, in order to determine the lower bound for any further fast simulation model implementation. `ImmediateDeposit` gives useful information about the time needed for the infrastructure itself to call the fast simulation methods. The `ShowerDeposit` provides the minimum amount of time needed to generate a specific number of hits with no additional calculations. The time spent by each of these models on the simulation of a grid of 3328 evenly-spaced photons originating at the LHCb interaction point is shown in Figure 1. The time spent by the `ImmediateDeposit` model is comparable across different photon energies, while the time spent in the `ShowerDeposit` model increases with the energy of photons. Both models perform as expected and clearly show that the time spent in the infrastructure of the interface is negligible.

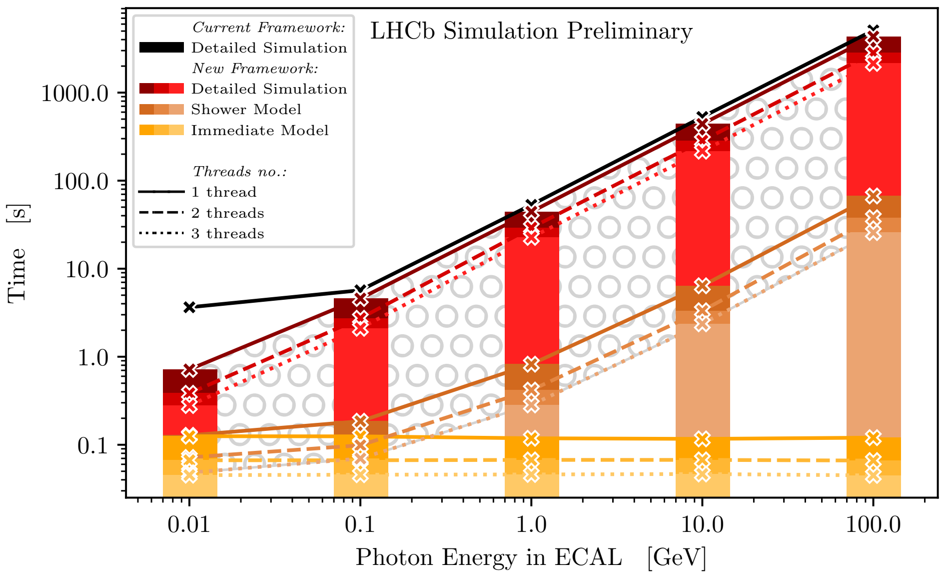

Figure 1. Comparison of the time spent by different fast simulation models, which indicate the performance of the infrastructure, and a detailed simulation with Geant4 in the electromagnetic calorimeter [16, 17]. In each of the models tested, a particle gun generates a grid of evenly-spaced photons of a given energy. For the detailed simulation the time of the current version of Gauss is also given as reference.

3. Fast Simulation training datasets in Gauss-on-Gaussino

Many parametric and machine learning models are currently being developed in the LHCb experiment. They require prior tuning or training on some input data in order to provide valid results. The information required to train these models is not always available in the standard output file, as Gauss stores only the information needed to digitise the response of the sub-detectors and what is required for physics analyses. Additional information about the simulated objects can be obtained by introducing new generic, sensitive sub-detectors. This functionality is provided by a new feature in Gaussino, called ExternalDetector. An example of the external plane-like detector, as seen by Geant4, is illustrated in Figure 2. The ExternalDetector can be extended to customize it for the training of a specific model.

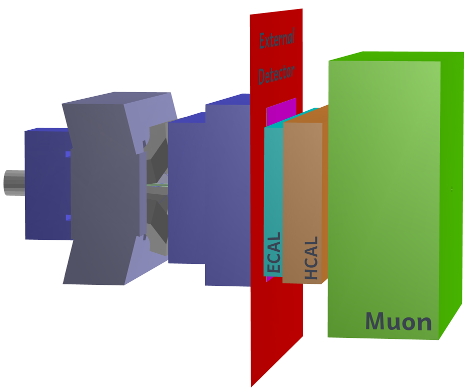

Figure 2. LHCb upgrade geometry as seen by the Geant4 toolkit with an example of a plane-like detector (red, thin plane), introduced by the ExternalDetector package [16, 17]. When used as a collector of particle information, it may provide the source of training information about incident particles for all sub-detectors placed downstream from it along the beamline: ECAL (cyan box), HCAL (orange box), or muon system (green box).

It might be the case that the external detector would need to be positioned where it could cause overlaps with other existing volumes in the geometry. In order to prevent that from happening, ParallelGeometry exploits an abstract concept introduced by Geant4 that allows for having multiple geometries in parallel, each of them performing the particle transport without interference from objects defined in other geometries. An example of a training dataset that can be used for the fast simulation studies involving the electromagnetic calorimeter is presented in Figure 3 where the same grid of evenly-spaced photons, as described in the previous section, has been used.

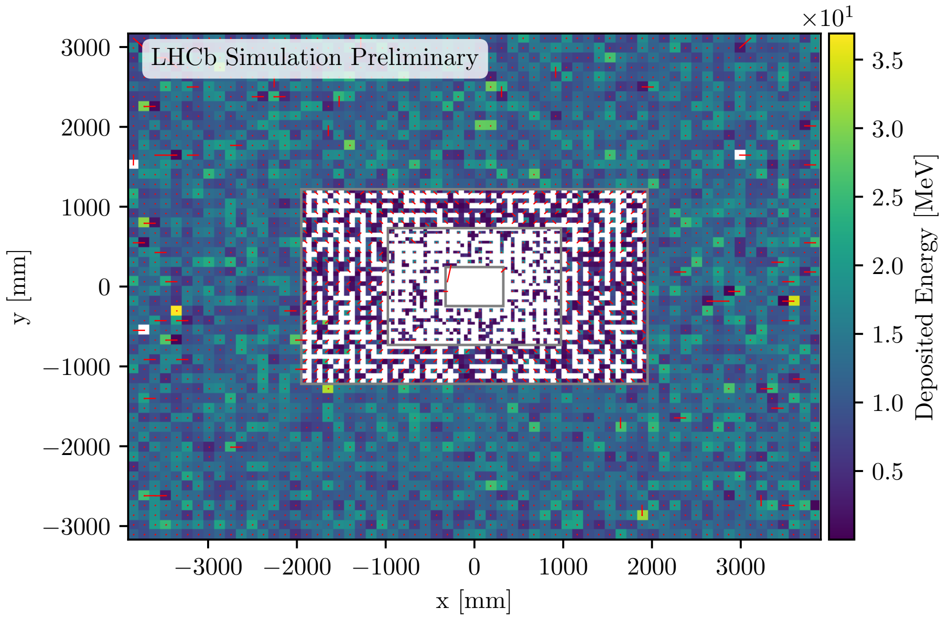

Figure 3. Visualization of the training dataset produced by placing a collector plane in front of the electromagnetic calorimeter [16, 17]. The image represents ECAL energy deposits (hits) projected onto an xy-plane.

4. Conclusions

Simulation software plays a key role in High Energy experiments. In the last few years, the LHCb simulation software Gauss has been redesigned in order to provide support for inter-event parallelism and fast simulation models in Geant4. In this paper, the performance of the FastSimulation interface in the new experiment-independent core simulation framework, called Gaussino, was presented. The time introduced by this infrastructure is negligible. A new way of producing training datasets in Gaussino has also been described, with an example of how this functionality can be used for the models targeting the electromagnetic calorimeter. The functionalities shown in this paper will facilitate integration of the current and future fast simulation models developed by the LHCb collaboration into Gauss-on-Gaussino framework.

References

[1] LHCb Collaboration 2008 *JINST* **3** S08005
[2] LHCb Collaboration 2015 *Int. J. Mod. Phys.* **A30** 1530022 (*Preprint* 1412.6352)
[3] Barrand G, Belyaev I, Binko P, Cattaneo M, Chytracek R, Corti G, Frank M, Gracia G, Harvey J, Herwijnen E, Maley P, Mato P, Probst S and Ranjard F 2001 *Computer Physics Communications* **140** 45–55 ISSN 0010-4655 cHEP2000

[4] Clemencic M, Degaudenzi H, Mato P, Binet S, Lavrijsen W, Leggett C and Belyaev I 2010 *Journal of Physics: Conference Series* **219** 042006
[5] Agostinelli S *et al.* (Geant4 collaboration) 2003 *Nucl. Instrum. Meth.* **A506** 250
[6] LHCb Collaboration 2019 Performance of a multithreaded prototype for the future LHCb simulation framework (Gauss-on-Gaussino) https://cds.cern.ch/record/2694003 LHCB-FIGURE-2019-012
[7] Müller D, Clemencic M, Corti G and Gersabeck M 2018 *The European Physical Journal C* **78** URL https://doi.org/10.1140/epjc/s10052-018-6469-6
[8] Whitehead M P 2018 A Palette of Fast Simulations in LHCb https://cds.cern.ch/record/2630475 LHCb-TALK-2018-302
[9] LHCb Collaboration 2019 Performance of the Lamarr Prototype: the ultra-fast simulation option integrated in the LHCb simulation framework https://cds.cern.ch/record/2696310 LHCb-FIGURE-2019-017
[10] Rama M 2020 Fast calorimeter simulation in the LHCb Gauss framework https://cds.cern.ch/record/2725640 LHCb-TALK-2020-108
[11] Ratnikov F 2019 Generative Adversarial Networks for LHCb Fast Simulation https://cds.cern.ch/record/2699549 LHCb-TALK-2019-403
[12] LHCb Collaboration 2019 LHCb CPU usage forecast https://cds.cern.ch/record/2696552h LHCB-FIGURE-2019-018
[13] LHCb collaboration 2018 Computing Model of the Upgrade LHCb experiment http://cds.cern.ch/record/2319756 CERN-LHCC-2018-014, LHCb-TDR-018
[14] Rama M and Vitali G 2019 *EPJ Web Conf.* **214** 02040
[15] Abellan Beteta C *et al.* 2020 Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC (*Preprint* 2008.11556)
[16] LHCb Collaboration 2021 Performance of the fast simulation interface in Gauss-on-Gaussino https://cds.cern.ch/record/2781378 LHCB-FIGURE-2021-004
[17] Mazurek M, Corti G and Müller D 2021 *COMPUTING AND INFORMATICS* **40** 815–832