ICHEP 2018, Seoul
July 4-11 2018

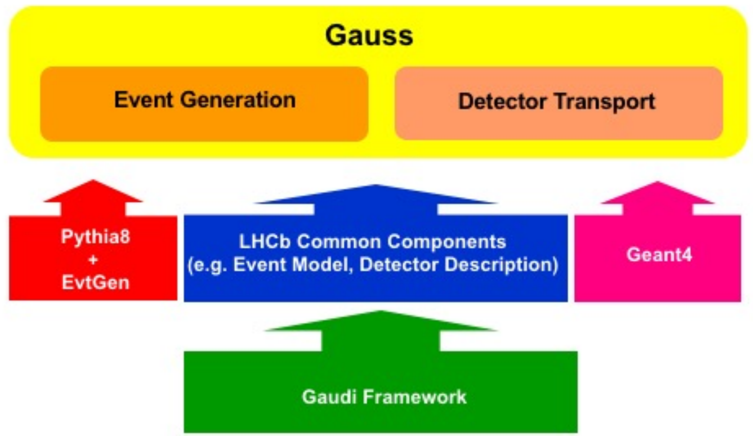
A Palette of Fast Simulations in LHCb

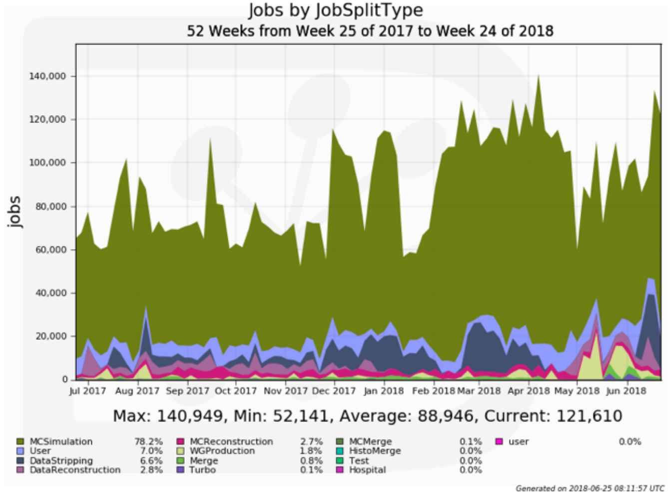
Mark Whitehead,
on behalf of the LHCb Collaboration

The LHCb simulation

- The simulation application for the LHCb experiment is *Gauss*
  - Particle generation and transport in the detector
  - Based on the Gaudi framework
  - Depends on a number of external libraries, including Geant4 for particle transport
  - A separate application, Boole, takes care of the digitized detector’s response

- Simulation takes most of the LHCb CPU resources





***Gaudi**: G. Barrand et al., Comput. Phys. Commun. 140 (2001) 45*
***Geant4**: S. Agostinelli et al., Nucl. Instrum. Methods Phys. Res., Sect. A 506, 250 (2003).*
***Gauss**: M. Clemencic et al., J. Phys. Conf. Ser. 331, 032023 (2011)*

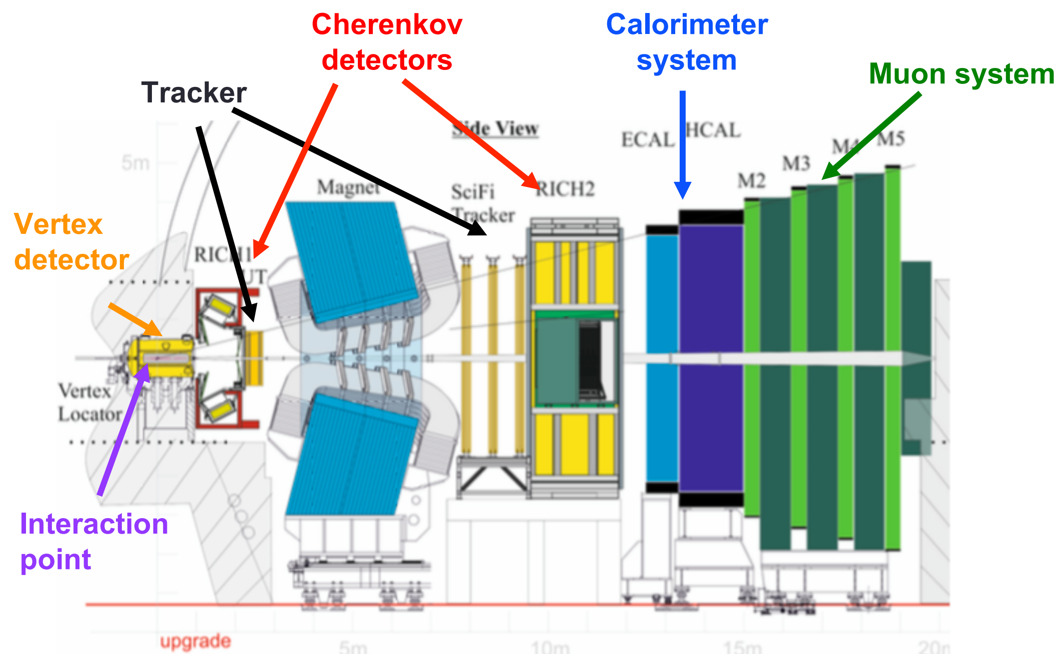
Setting the scene for the future

LHCb Upgrade for Run III

L_{inst}: 4 10^{32} → 2 10^{33} cm^{-2} s^{-1}
μ: 1.1 → 7.6
L_{int}: 3(Run1) + 5(Run2) fb^{-1} → 50fb^{-1}

- full software trigger with high signal purity
- analysis directly on trigger output

For more details on the upgrade, see talks by myself, Giovanni Passaleva, Stefano De Capua and Michele Piero Blago



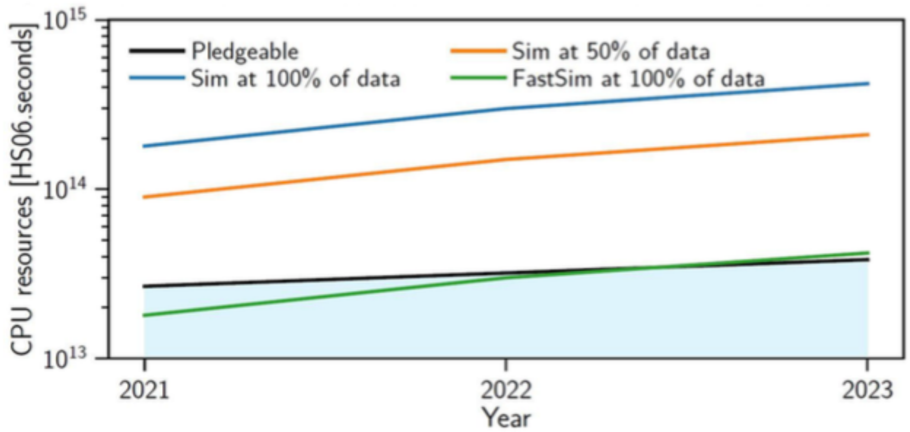
LHCb collaboration, A. A. Alves Jr. et al., The LHCb detector at the LHC, JINST 3 (2008) S08005.

Simulation will be dominating the CPU needs, even more than today

C. Bozzi, "Challenges for LHCb trigger", WLCG & HSF Workshop, Naples, March 2018
R. Matev, "Real-time analyses - the LHCb case" , WLCG & HSF Workshop, Naples, March 2018

Need of a faster detector simulation

Collecting more interesting events in Run III – and further – will require more events to be simulated



Legend:
"Sim at 50% of data" = FullSim sample is 50% the datasize, FastSim sample is 50% the datasize

FastSim speed assumed to be 10 that of FullSim

➡ **Need to shift towards a scenario where a significant fraction of LHCb MC events is fast-simulated**

Solutions to face disk storage limitations have been deployed in past years

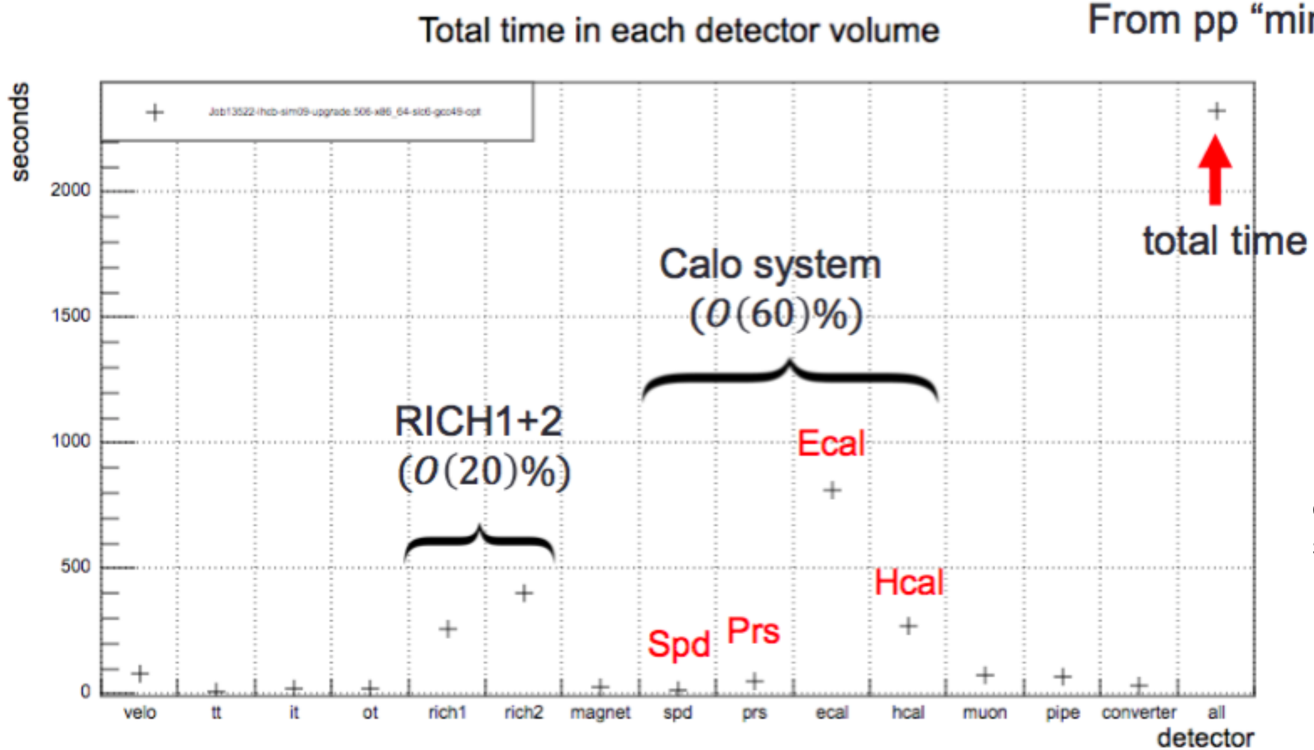
Fast simulation options

- Simplified detector simulation
  - Reduced detector: RICH-less or tracker-only. In production
  - Calorimeter showers fast simulation. Under development
  - Muon low energy background, used with full muon detector simulation. In production
- Simulation of partial events
  - Simulate only particles from signal decay. In production
  - ReDecay, e.g. use N-times the non-signal decay part of the event. In production
- Fully parametric simulation
  - Parametrized tracking, calorimeter and particleID objects with a DELPHES-based infrastructure. Under development

No single solution for all needs, but different options organized under the unique *Gauss* framework

Deploy solutions when mature for physics

CPU time in the LHCb Geant4-based simulation



From pp “min bias” events

For a recent version of the Gauss simulation for Run II

Most of the CPU time in Geant4-simulated events is spent in the calorimeter system

Simulating without RICH physics

In production

- Switch off generation of optical photons in the RICH
- RICH material still include in detector geometry and fully simulated → no impact on downstream detectors

✔ Roughly a 30% speed up

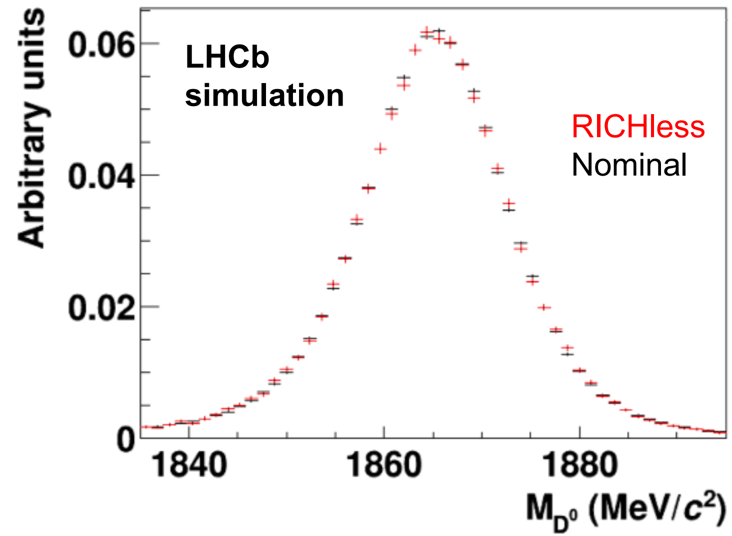
✔ Slight reduction in disk usage

✗ No hadron particle ID in the sample

but not all analysis use it

✗ Requires a special trigger and offline selection

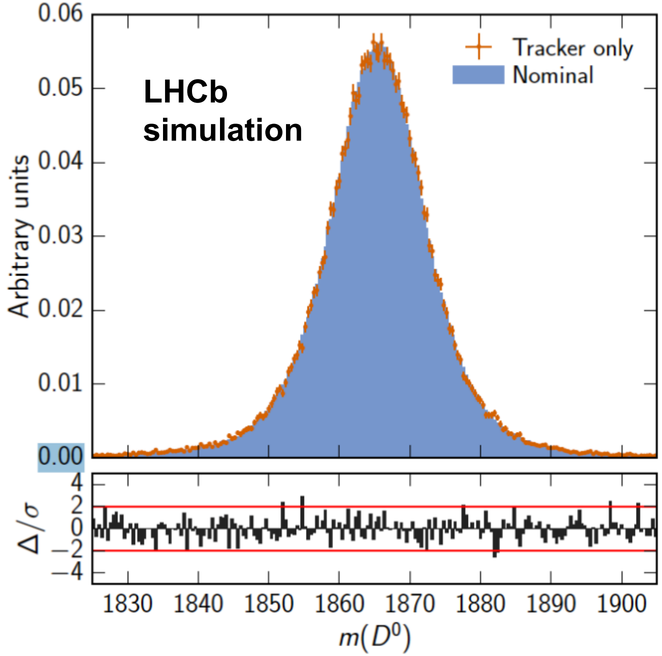
Validated and used in the R(D) analysis*



Tracker-only simulation

In production

Validated with charm samples



- Switch off generation of optical photons in the RICH, but RICH geometry included
- Calorimeter and Muon Systems removed from detector geometry

✔ Roughly a speed up of 10 times

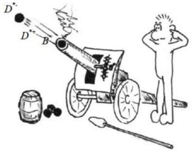
✔ Reduction in disk usage of ~ 4 times

✗ No particle ID in the sample

✗ Not possible to run hardware trigger emulator

Particle Guns

- Generate and simulate only the signal particles, no remainder of the event



- Primary vertex smearing applied in the reconstruction phase

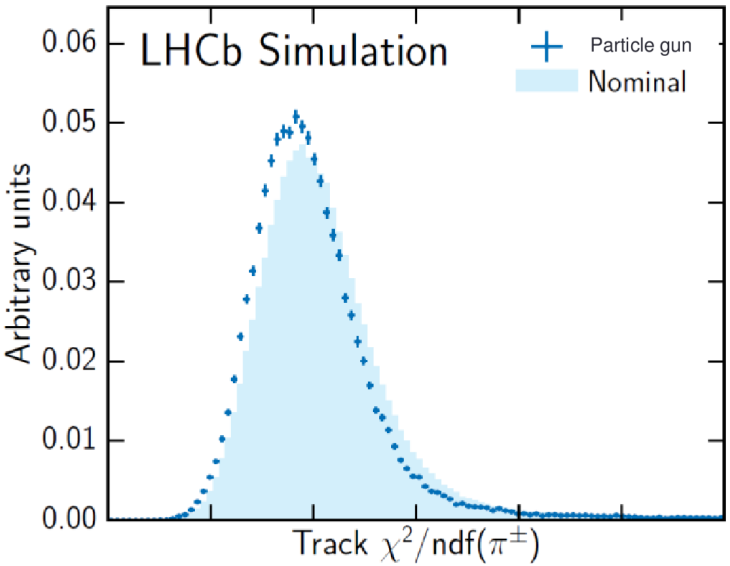
✔ Roughly 50-100 times faster

✔ About 100 times less disk usage

✗ Does not describe global variables

✗ Variables affected by local occupancy not well described

In production



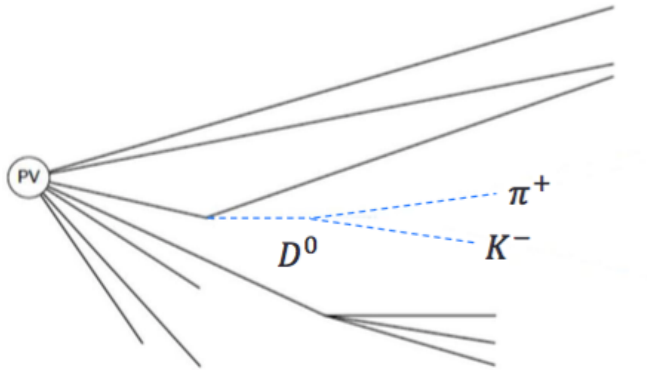
Ideal for CEP-like analysis

Very useful for specific studies

Signal ReDecay

In production

- Combines full simulation and particle gun approach
- Generates a full event but reuses it *N* times, each time replacing the signal decay
  - Hadronization stays the same, Geant4 detector simulation only done once per full event
  - origin and **p** of the redecayed particle stay the same to keep correct correlation with the underlying event...
  - ... but its flight length and final state kinematic change

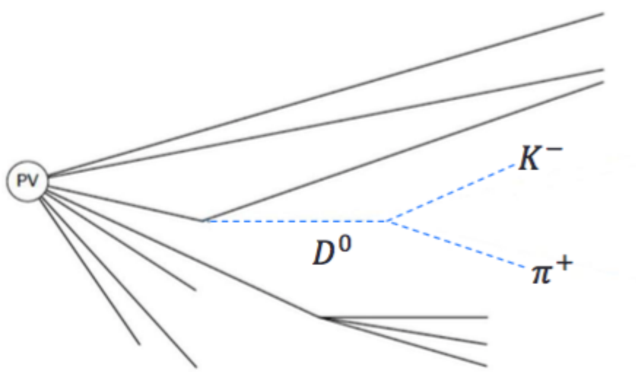


D. Müller, ReDecay, 2017 IEEE NSS and MIC

Signal ReDecay

In production

- Combines full simulation and particle gun approach
- Generates a full event but reuses it *N* times, each time replacing the signal decay
  - Hadronization stays the same, Geant4 detector simulation only done once per full event
  - origin and **p** of the redecayed particle stay the same to keep correct correlation with the underlying event...
  - ... but its flight length and final state kinematic change



D. Müller, ReDecay, 2017 IEEE NSS and MIC

Signal ReDecay

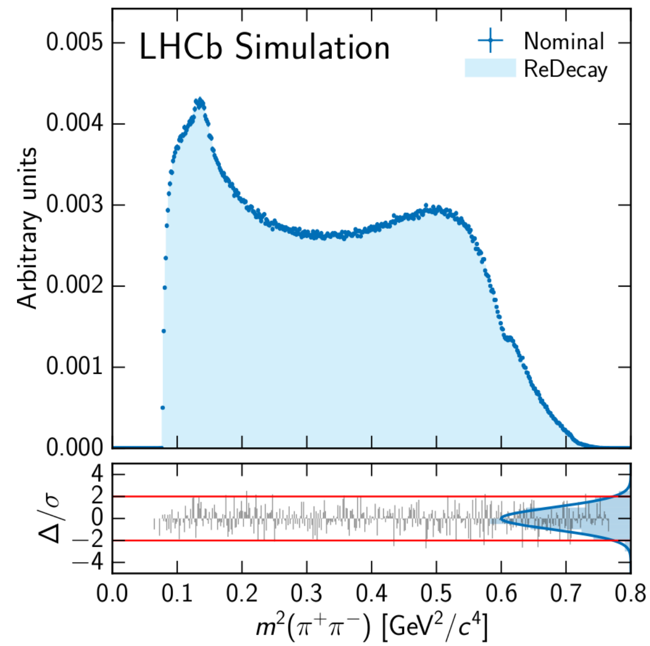
- Independent of the generator used
- Compatible with other fast simulation options
- Automatically benefits from improvements to detector simulation

✔ Roughly 10 to 50 times faster depending on the sample

✔ Complexity of fully simulated events

✗ Care needed when treating correlations for global variables

✗ No disk space saving



Very beneficial for extracting multi-dimensional efficiencies as a function of final-state quantities. Now used by a number of analysis

Fast simulation of the Calorimeter system

Under development

- A generic infrastructure has been developed in the Geant4 interface of the current Gauss framework to allow the Geant4 transport for a given particle to be interrupted to inject a given fast simulation counterpart

- Two fast parametrization solutions currently under development
  - Classic Frozen Shower Libraries
  - Hits generation based on Generative Adversial Networks (GAN)
  - … not necessarily mutually exclusive ☺. Could solve the Shower Library problem of a fast search in multi-dimensional phase space by reducing the dimensions with Machine Learning techniques, e.g. autoencoders

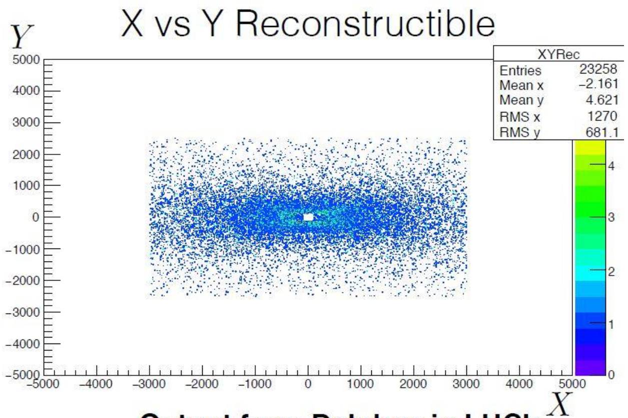
- Aim to speed up by factor 3 to 10 the simulation of the calorimeters

For details see F. Ratnikov presentation at this ICHEP

Fully parametric fast simulation

Under development

- Work in progress on a fully parametric ultra-fast simulation based on the DELPHES package
  - Parametrizes not only the detector response but also the reconstruction
- Crucial to cope with large amount of simulated statistics needed for Run III and future Upgrade II. **Goal: 100-1000x faster then full simulation.**
- Functional prototype integrated in the current Gauss framework
  - **Tracking** efficiency and resolution
  - **Primary vertex** reconstruction
  - **Photon** calorimetric objects
  - Outputs **LHCb reconstructed high level objects,** compatible with the experiment analysis tools



Output from Delphes in LHCb acceptance

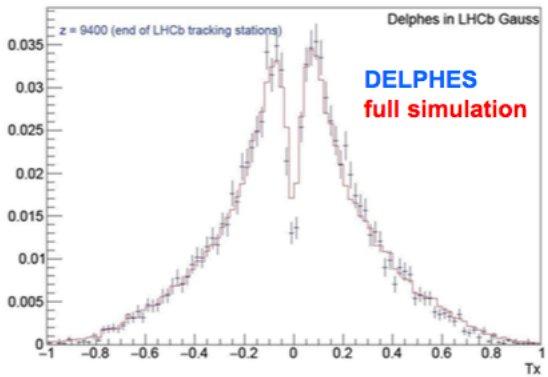
J. De Favereau et al., JHEP 02 (2014) 057
B. Siddi, "DELPHES talks", to be presented at CHEP 2018

Fully parametric fast simulation

Under development

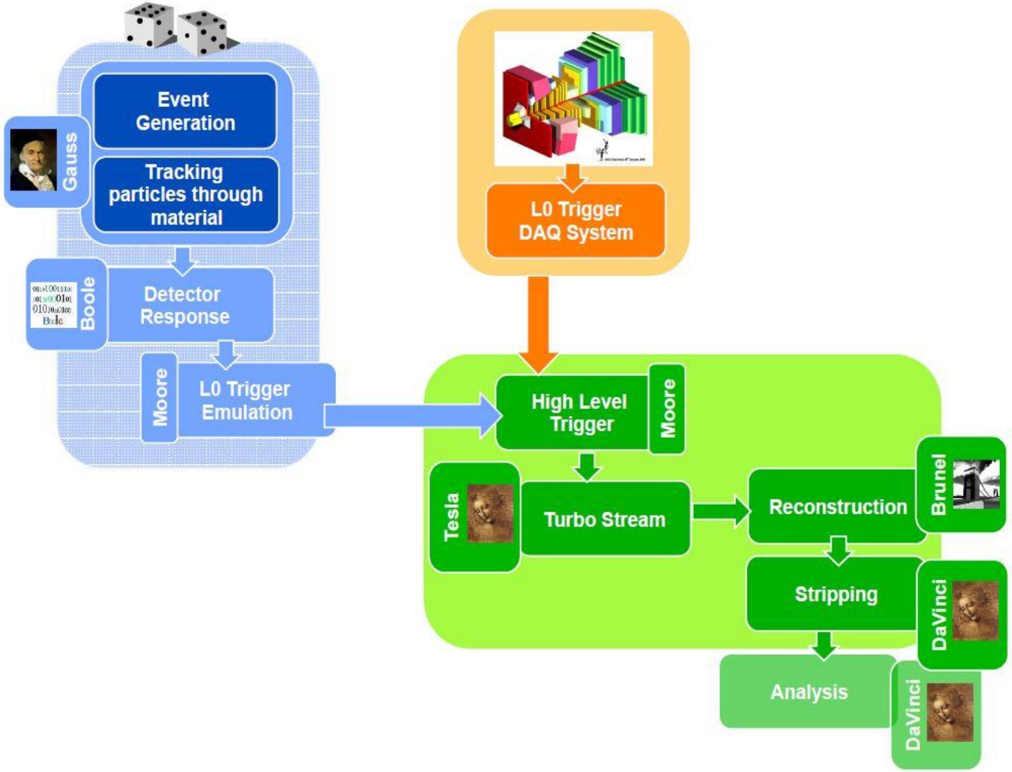
- Deployment in steps with a first beta release for users to try out with some physics analysis coming soon
- In parallel extensive development to obtain all relevant quantities and replace all applications between Gauss and physics analysis input
  - Include error covariance matrices in **charged particle tracking**
  - Parametric **calorimetric response** for all electromagnetic particles
  - **Particle identification hypothesis** to be integrated adapting particle ID calibration mechanism for data
- Parametrization is fully experiment dependent and specific to a given data taking period
  - To be properly calibrated for precision measurements it requires both extremely reliable full simulation and highly precise data-driven measurements of the detector performance

px/pz distribution



Software infrastructure

- In the process of building an integrated simulation framework
  - possibility to mix different simulation options for different particles
  - configuration as much simple and transparent as possible to the user



- Task complicated by the fact that Gauss is not the only application involved: some configuration choices affect also subsequent steps of the simulated data processing

Summary

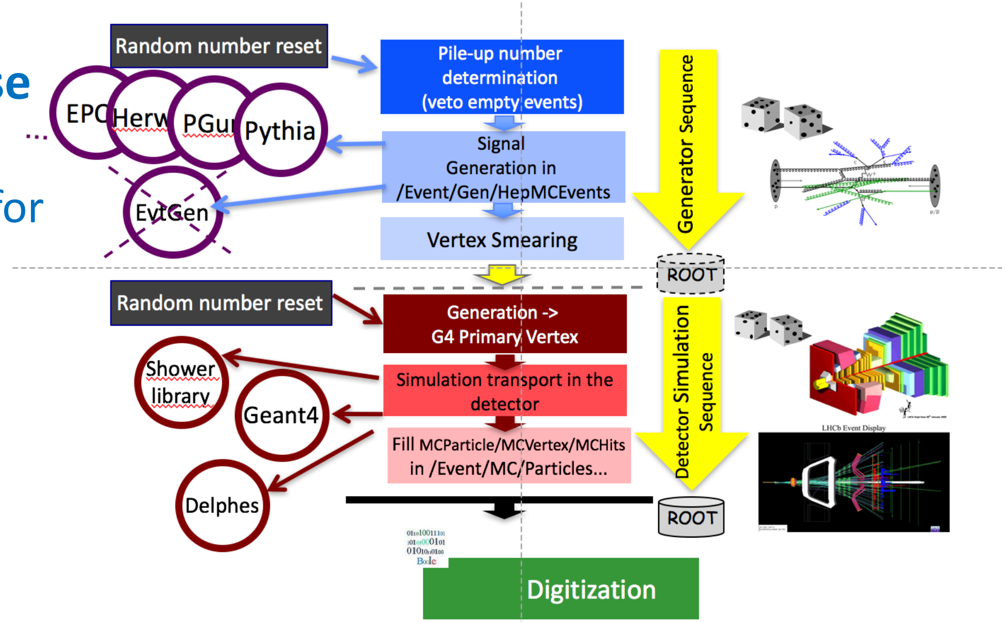
- The role of fast simulation in LHCb is becoming more and more important to face the limitation of computing resources vs the experimental needs for larger samples
- Significant increase of development activity in last two years
- Not a single fast simulation, but a number of complementary options
  - Some already used in production, others under development
  - All integrated in the existing simulation/reconstruction framework
- Working towards a flexible framework which allows users to select full/fast simulation modes for different particle types and subdetectors, depending on their needs

BACKUP

Integrated simulation framework

No single solution for all needs but different simulation options organized under the Gauss unique framework

Building an integrated, easy to use and **safe** simulation framework
- Working fast simulation solutions for the whole event
- Prototypes allowing a mix of simulation flavors for different detectors and particles
- Output compatible with what expected by later processing



Benchmark and physics performance measurements to choose baseline production setting