Status and prospects for technical solutions for real-time analysis for Run 3 and later

Miguel Ramos Pernas

on behalf of the ALICE, ATLAS, CMS and LHCb collaborations

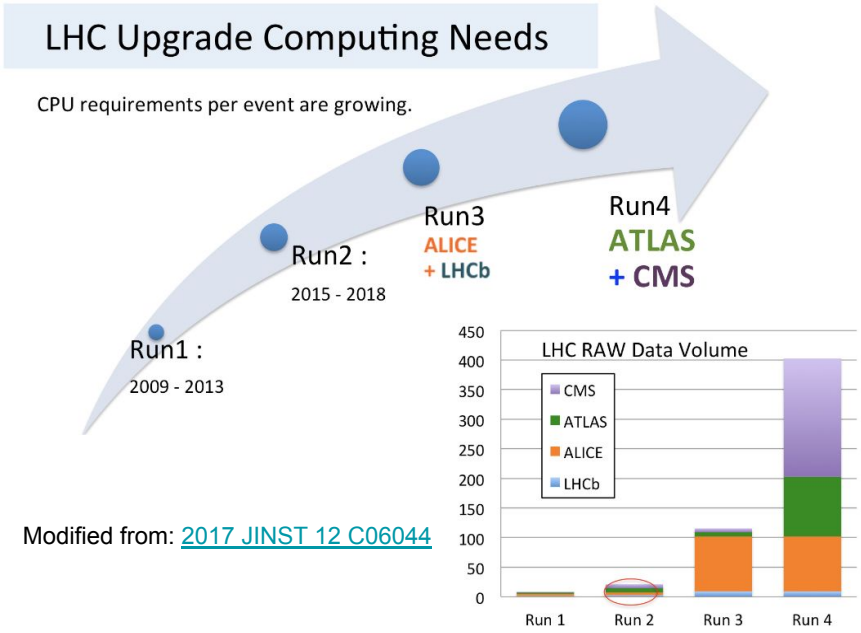
University of Warwick

miguel.ramos.pernas@cern.ch

LHCP 2021

Real-time analysis

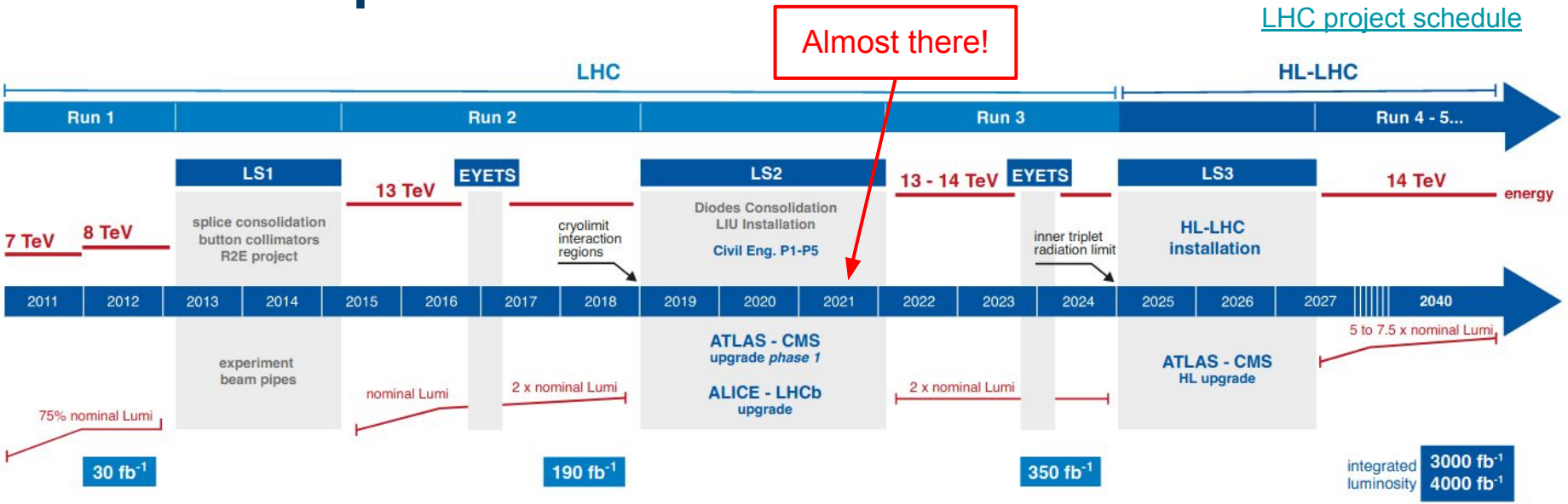
- The LHC experiments have grown interest on accumulating large data samples
  - Increase rate for known signatures
  - Search for rare or undiscovered processes
- The experiments must carefully filter uninteresting events during data-taking
- Bandwidth limitations due to the higher luminosity: must also reduce the event size
- Do as much work as we can at the trigger level
- When to keep and what to keep are crucial and completely analysis-dependent



Modified from: 2017 JINST 12 C06044

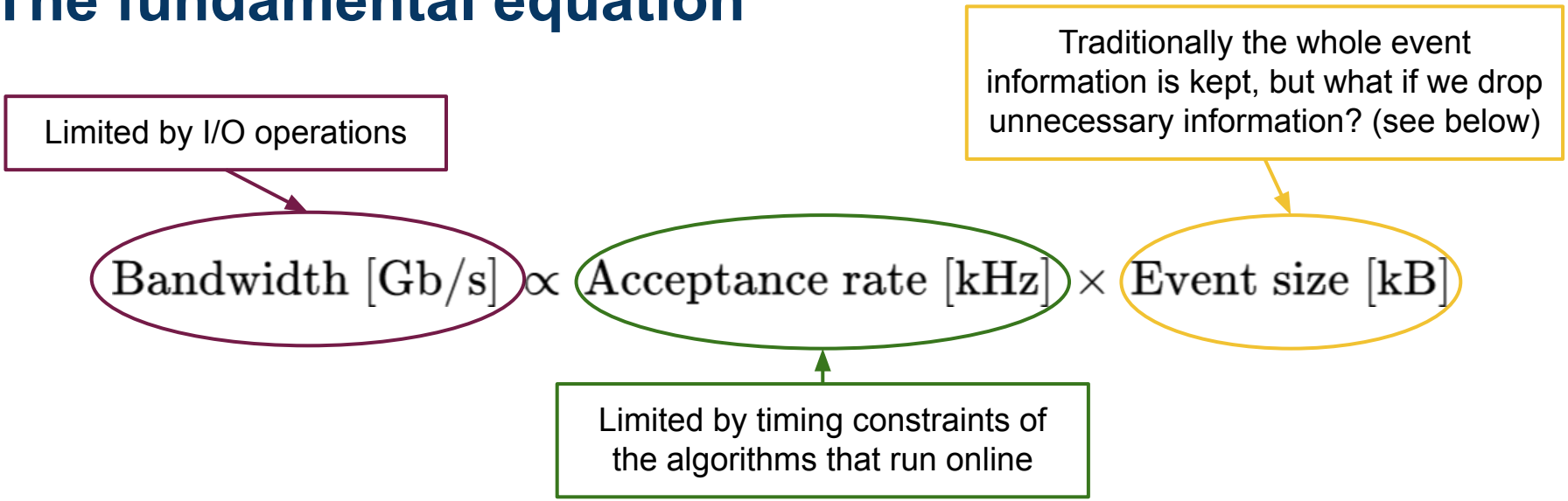
Real Time Analysis (RTA) is crucial for the future of the experiments!

The roadmap of the LHC



- A lot of developments have already been done to prepare the experiment for Run 3
- Many of these ideas are also prototypes/early designs for a HL-LHC.

The fundamental equation



- Experiments should try to find harmony given the resources and physics program
- The timing of the online algorithms can be reduced using new acceleration techniques (GPUs, FPGAs, … see the talks by Thomas Boettcher and Marten Ole Schmidt)
- The event size can be reduced by carefully dropping information that is not needed

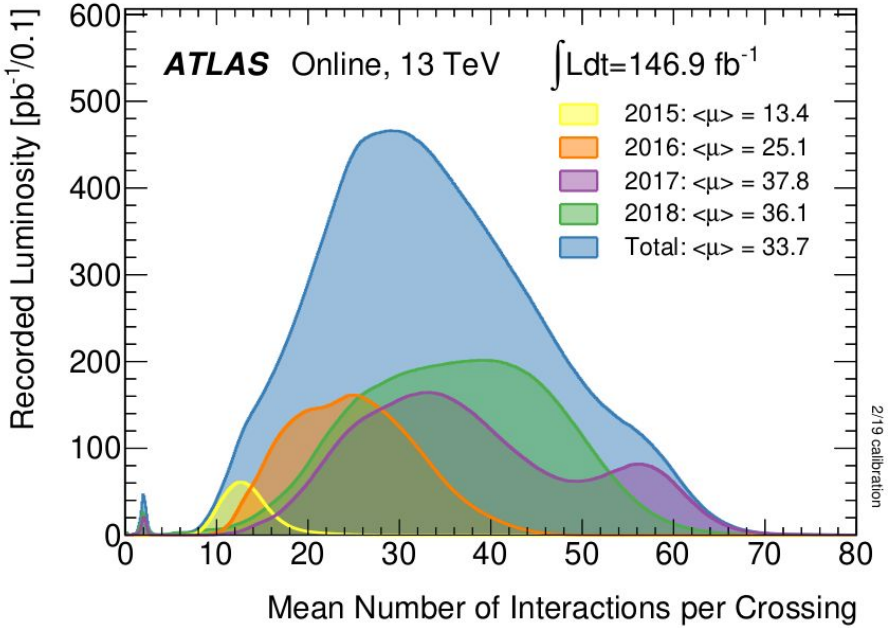
The pile-up problem

- Signal and background do not scale by the same factor with the luminosity
  - Combinations of particles in groups of two for growing *N*

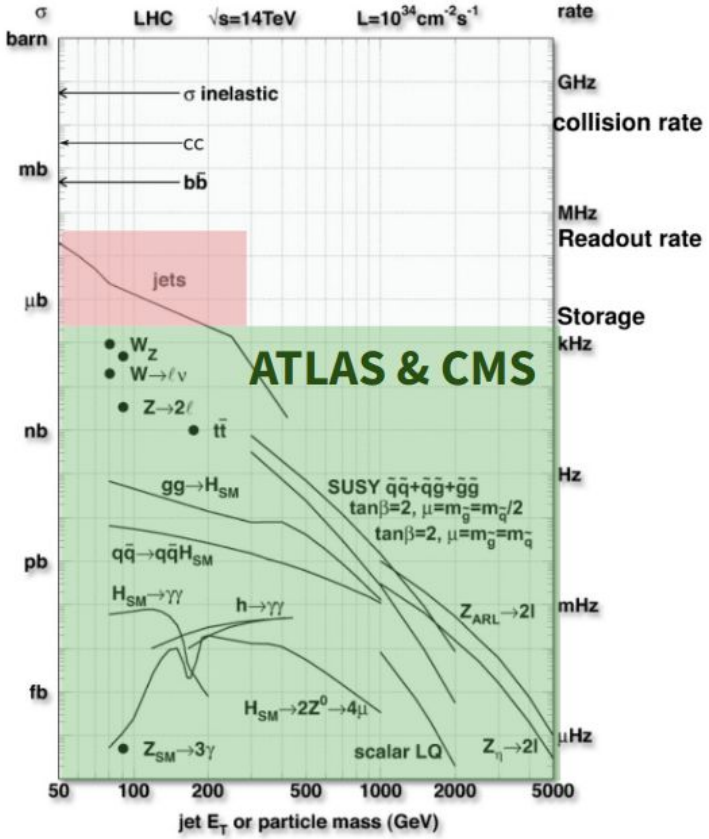
$$\binom{4}{2} = 6 \qquad \binom{8}{2} = 28$$

  - Raw data bandwidth scales quadratically with the luminosity
- Must improve the removal of fake/ghost tracks and background

ATLAS Run 2 public results



Triggering on high-mass signatures

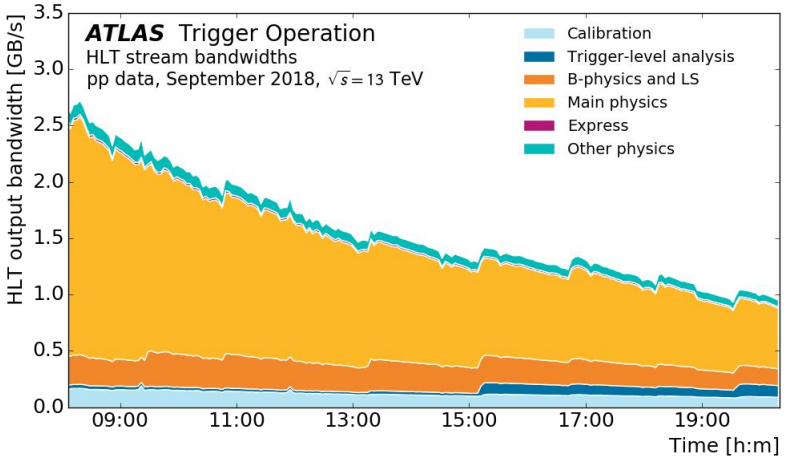


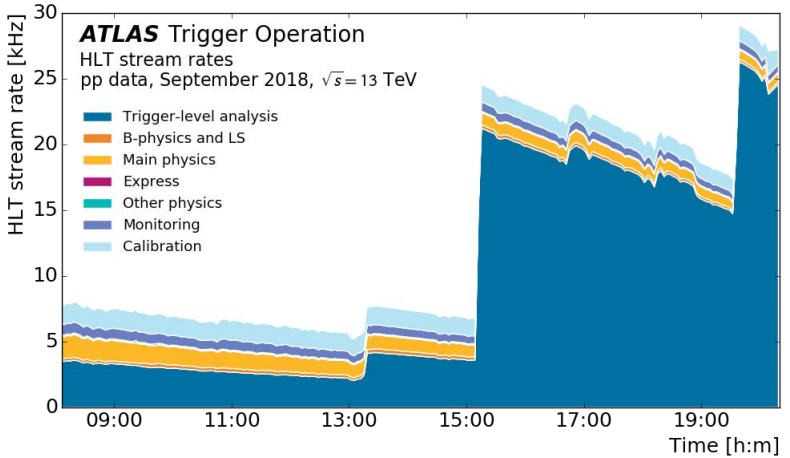
- General purpose detectors (ATLAS/CMS) mainly trigger on signatures produced on kHz
  - Triggers based on transverse momentum and missing transverse energy are enough
  - (Quite) simple topologies
- Current configurations are sufficiently efficient
- What about jets?

ATLAS

- Similar structure as in Run 1 and Run 2
- Trigger composed by L1 (hardware) and HLT (software)
  - Saving up to 25 kHz at the HLT (2018)
- Increase the rate at the end-of-fill using lower thresholds for RTA
- Include jet reconstruction at the trigger level
  - Jet calibration is still a challenge
  - Coarse algorithms and thus poor resolution
  - More robust implementations expected for the HL-LHC era

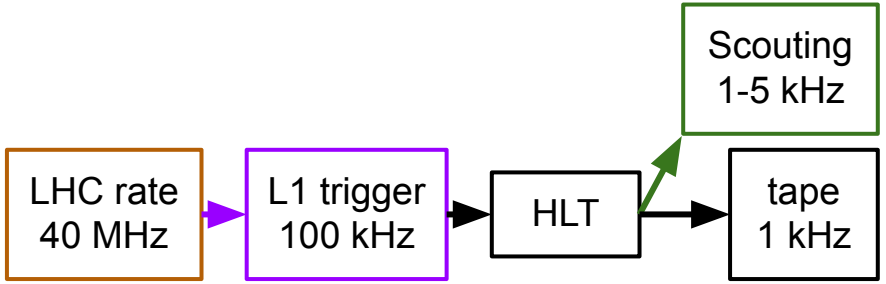
ATLAS public results



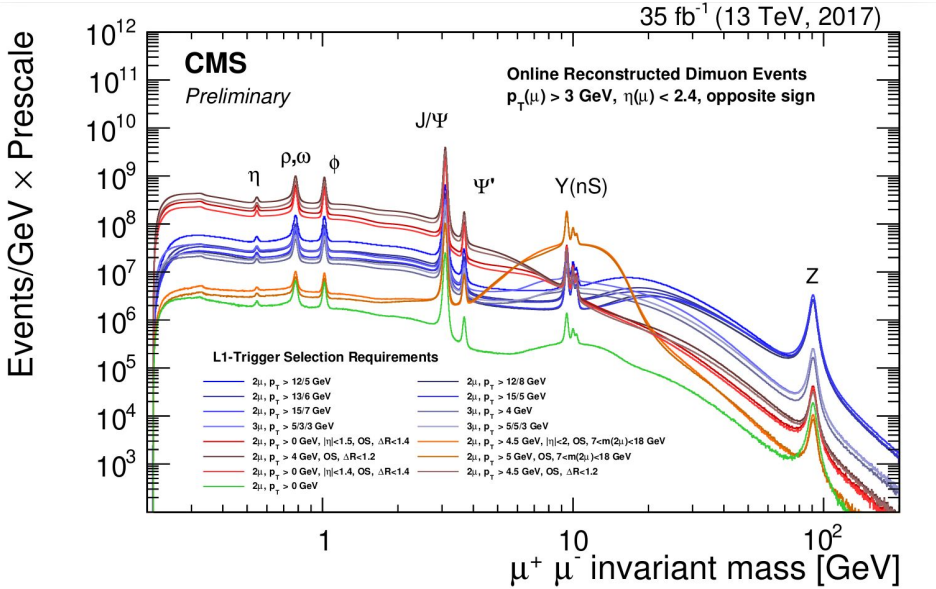


CMS

- Improved Scouting data stream at the trigger level (mainly for new searches).
- Particle-Flow (PF) based
  - vertices, PF muons, PF candidates, PF jets, MET
  - CPU-limited: 10 kB/event, 330 MB/s
- Calo scouting
  - vertices, muons, Calo jets and MET
  - L1-limited: 1.5 kB/event, 30 MB/s
- Target for HL-LHC: data scouting at L1



CMS public results



ALICE

Up to 1 MHz of p-p interaction rate

Synchronous processing:

- Online calibration and data compression
- TPC tracking and space point calibration

Data bufferend for 2 weeks; apply detector calibration

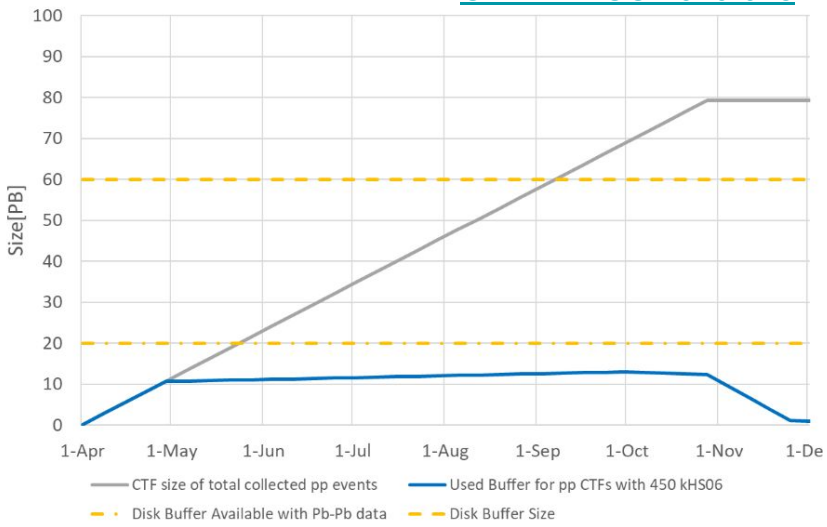
Asynchronous processing:

- Physics grade reconstruction
- AOD creation and selection of the events of interest (< 0.1% of the total)

Expecting to collect 200 pb^{-1} of p-p data (3000 x Run 2 luminosity for the Minimum Bias stream)

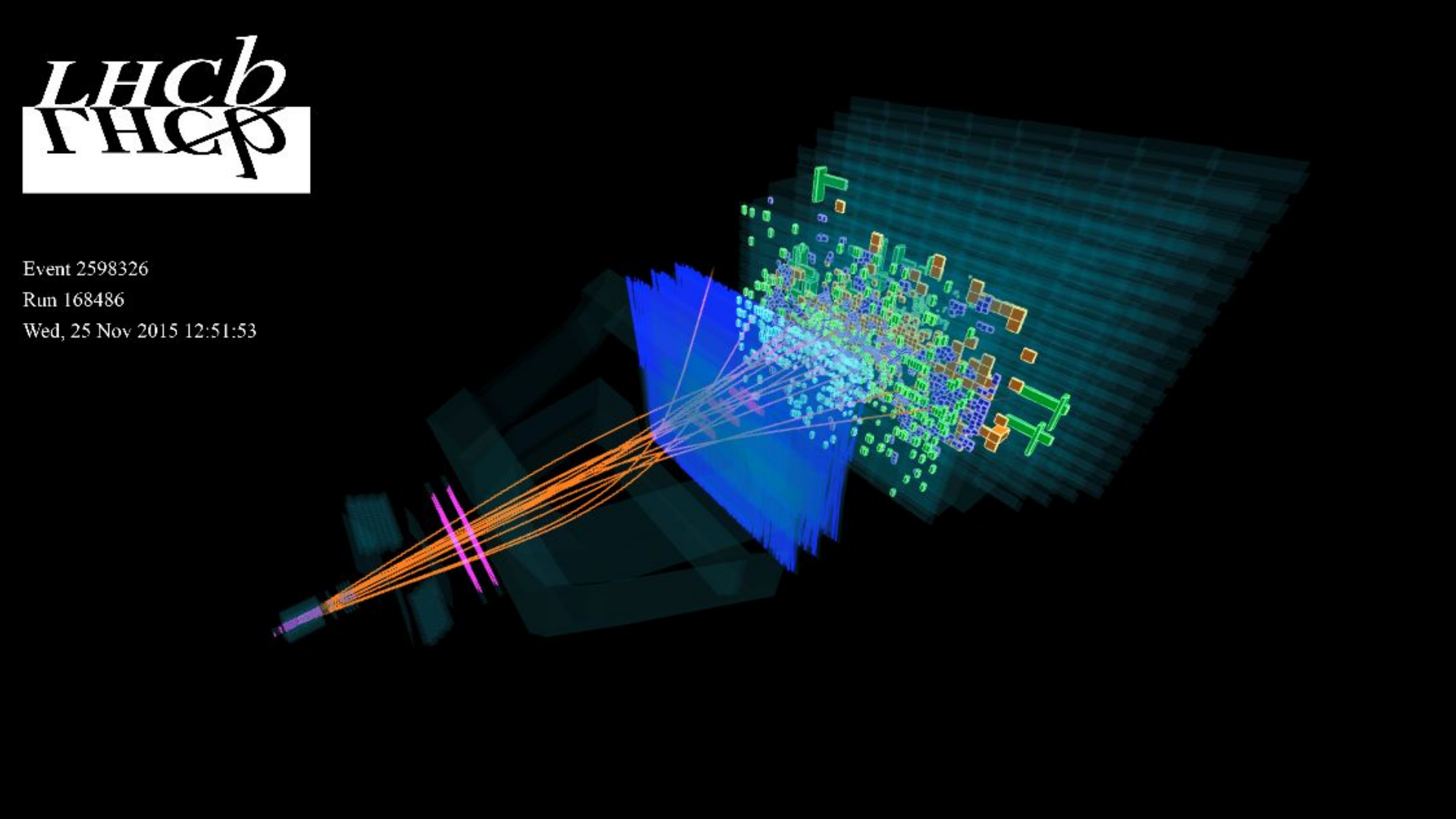
⇨ Hardware triggers are not efficient for the physical signatures; move to asynchronous analysis triggers

CERN-LHCC-2020-018



Analysis framework: J. F. Grosse-Oetringhaus

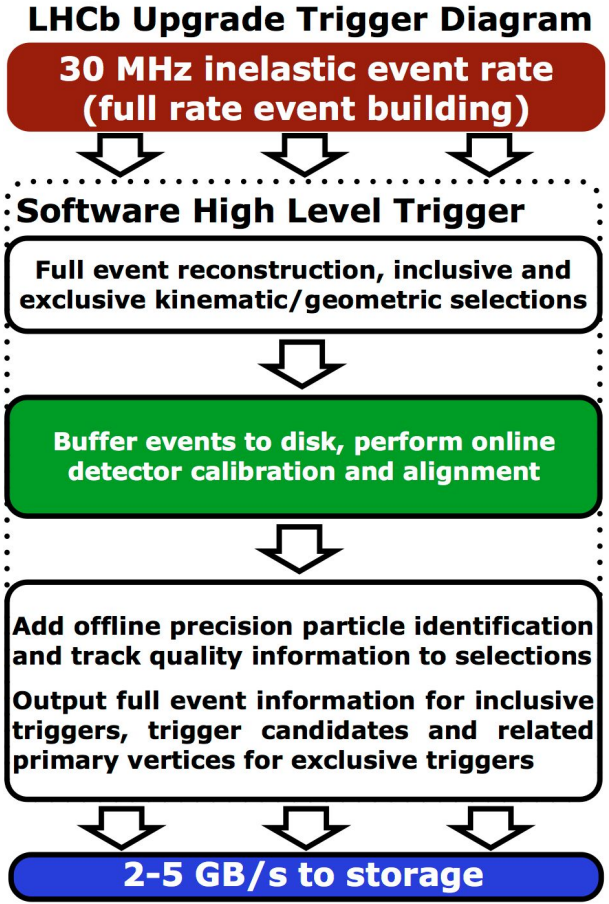
GPUs in ALICE: Marten Ole Schmidt


LHCb
Event 2598326
Run 168486
Wed, 25 Nov 2015 12:51:53

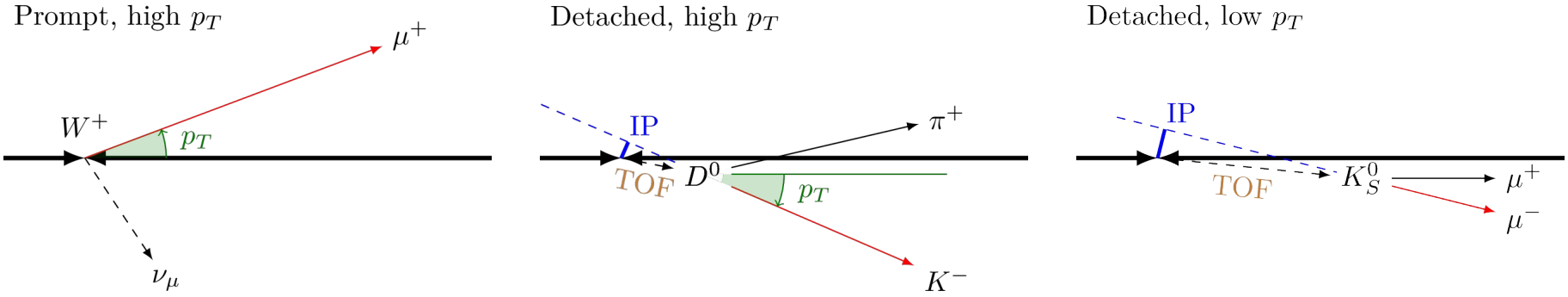
LHCb

LHCb public figures

- Removal of the L0 (hardware) trigger
- Trigger composed by two software triggers (HLT1 and HLT2)
- Implementation of a HLT1 fully on GPUs [Comput Softw Big Sci 4, 7 (2020)] (see talk by Thomas Boettcher)
- Asynchronous HLT2 with exclusive and inclusive selections
- Analysts can apply current offline selections at HLT2 including a selective persistence

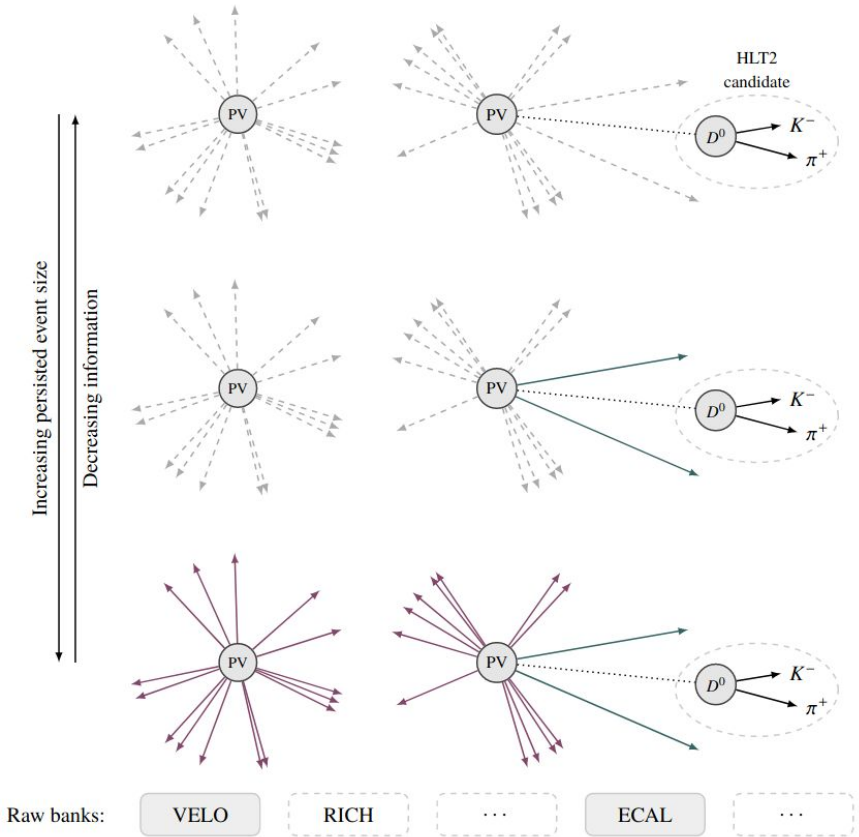


The LHCb trigger for Run 3



- Targeting many different topologies (often highly complex)
  - Time-dependent measurements: avoid selections based on the flight-distance, vertex position, …
  - Rare decays: specific requirements depending on the decay
  - Exotic searches (long-lived or high-mass): ALPs, dark photons, …
  - Spectroscopy, QCD, …
- Increase the usage of exclusive selections (in particular for well-known decays)

Selective persistence (SP) at LHCb



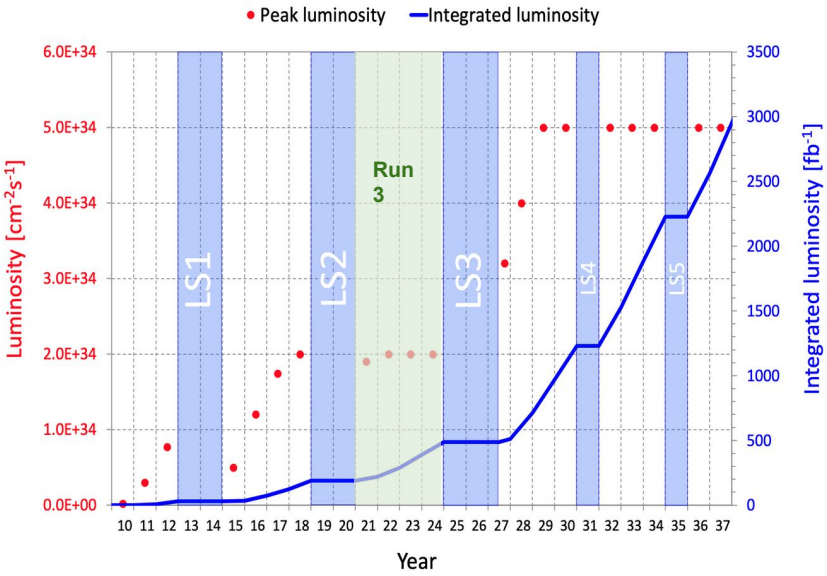
[2019 JINST 14 P04006]

- Analysts are encouraged to work and save specific information from the event
- Part of the Run 2 data was taking already using this implementation
- Run 2 numbers: 70 kB (full), 5 kB (SP) at 0.6 GB/s
- 528 HLT2 lines in Run 2, expected x 10 for Run 3
- A lot of work is being done to parallelize and boost the selections

The High-Luminosity LHC

With the increase of luminosity almost all the events will contain a signature of interest: doing RTA is no longer optional!

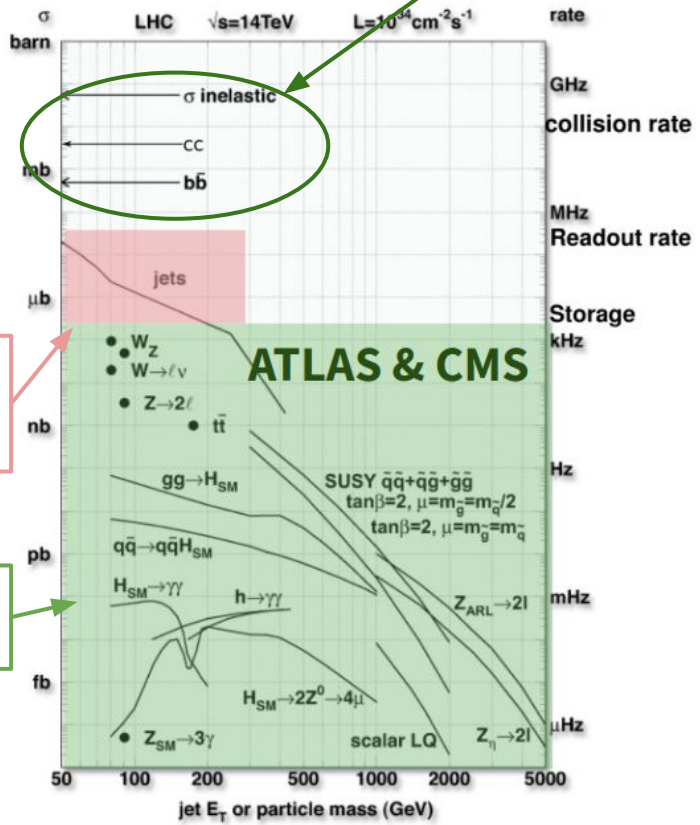
LHC commissioning schedule



Challenging scenario for ALICE and LHCb: almost every event will contain a charm or beauty candidate

Jet backgrounds will be even harder to control

Natural candidates for data scouting



Summary

- Experiments are finalizing the preparations for Run 3 data-taking
- Profit from the advances in computing to boost the execution of trigger selections and improve the throughput
- Fast and effective real-time alignment and calibration is crucial

Many ideas could be implemented in the middle of Run 3, in particular those aiming at working in the HL-LHC

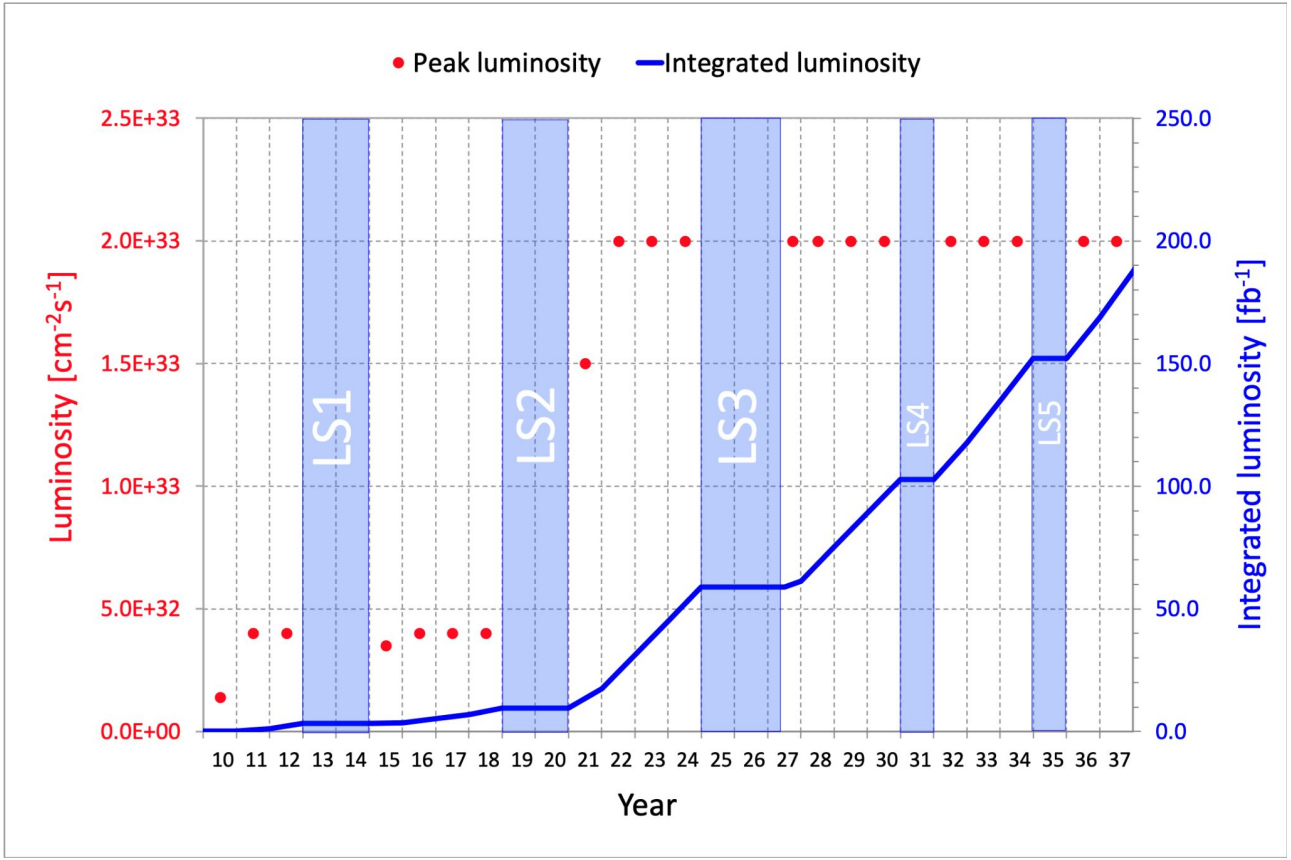
- The HL-LHC is a more challenging environment, with up to 7.5 x current luminosity
- Throughput limitations might be difficult to overcome with CPUs (RTA fully on GPUs/FPGAs?)
- We will definitely learn from Run 3


Thanks for your attention

Backup

LHCb luminosity

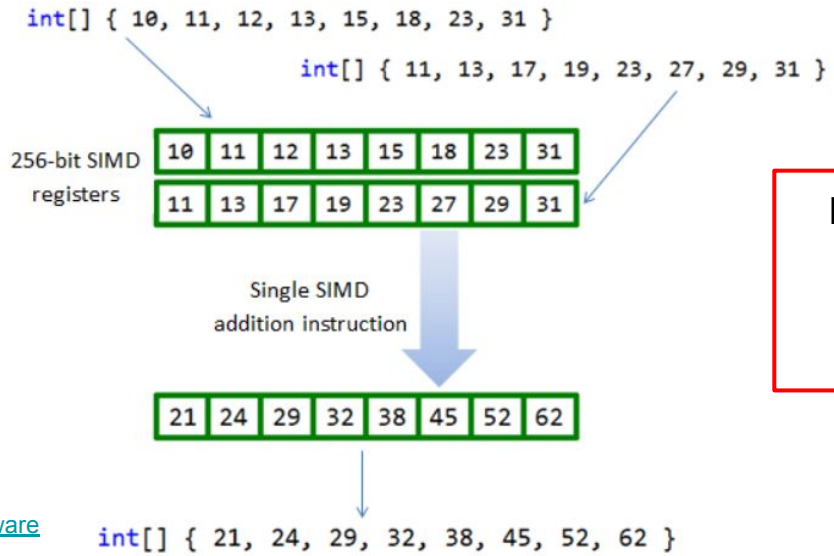
LHC commissioning schedule



LHCb computing developments at HLT2

- There are challenges to tackle at HLT2 (timing constraints, combinatorics problems, ….), which is based on CPUs
- Ongoing developments of SIMD combined with multithreading to boost the HLT2 processing (AVX, AVX2, SSE, NEON).

Can win a factor ~4 in timing



Credit: Instil Software

Preliminary studies look promising, but the implementation is not complete yet