Sustainability of real time analysis at 5 TB/s data rate

Arantza Oyanguren (IFIC – Valencia)

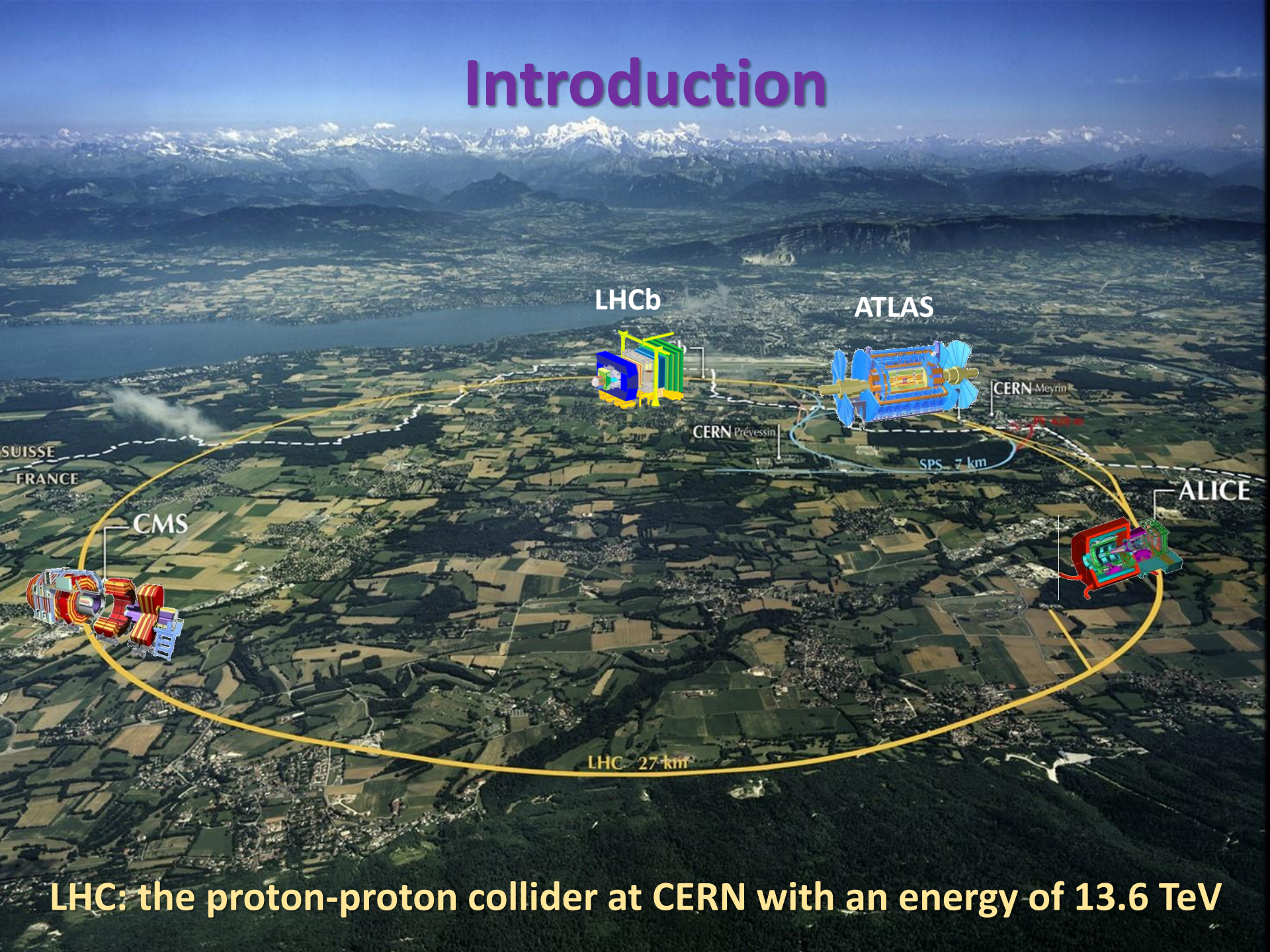
Alvaro Fernández, Luca Fiorini, Brij Kishor Jashal, Valerii Kholoimov, Volodymyr Svintozelskyi, Alberto Valero, Jiahui Zhuo

IFIC INSTITUT DE FÍSICA CORPUSCULAR
Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
Plan de Recuperación, Transformación y Resiliencia
GENERALITAT VALENCIANA Conselleria de Educación, Universidades y Empleo
GVANEXT Fondos Next Generation en la Comunitat Valenciana
CERN

Outline

- Introduction
- Real Time Analysis at LHCb
- The HIGH-LOW project at IFIC - Valencia
- Hardware sustainability
- Software sustainability
- Conclusions & prospects

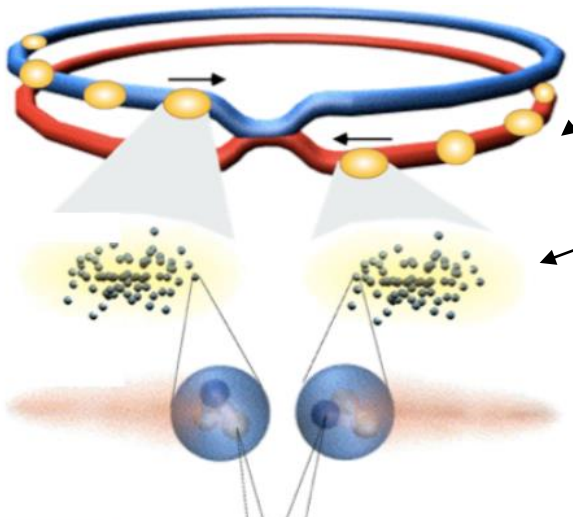
Introduction



LHC: the proton-proton collider at CERN with an energy of 13.6 TeV

Introduction

Proton-proton collision



2028 bunches of protons per beam

Beam energy of 7 TeV

Luminosity 10^{34} cm^{-2} s^{-1}

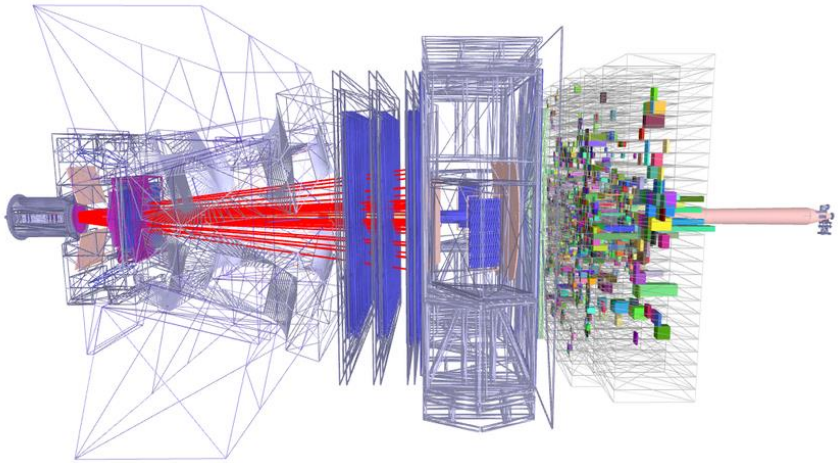
10^{11} protons per bunch

Crossing rate 40 MHz, i.e. 40 M collisions/s

and ~ 1 MB data from detectors

→ 5 - 40 TB/s

We need to select the events to be saved!

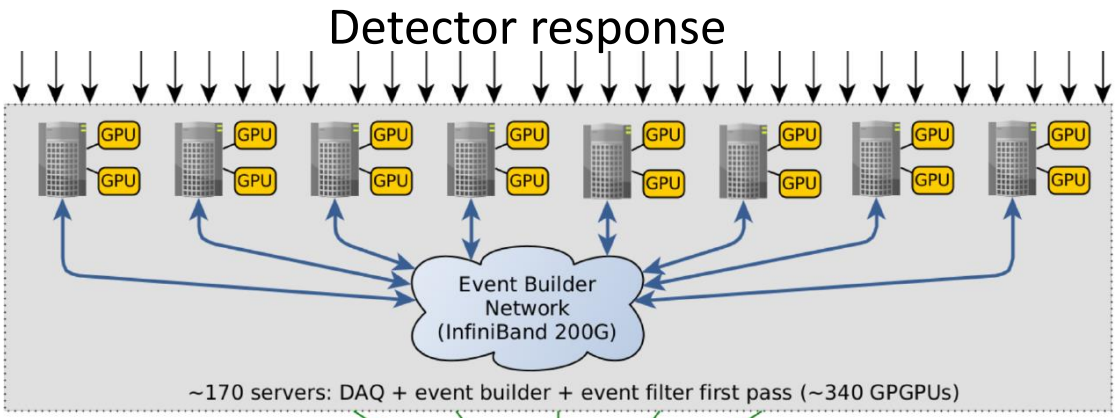


Google+ +

~ 0.3 TB/s

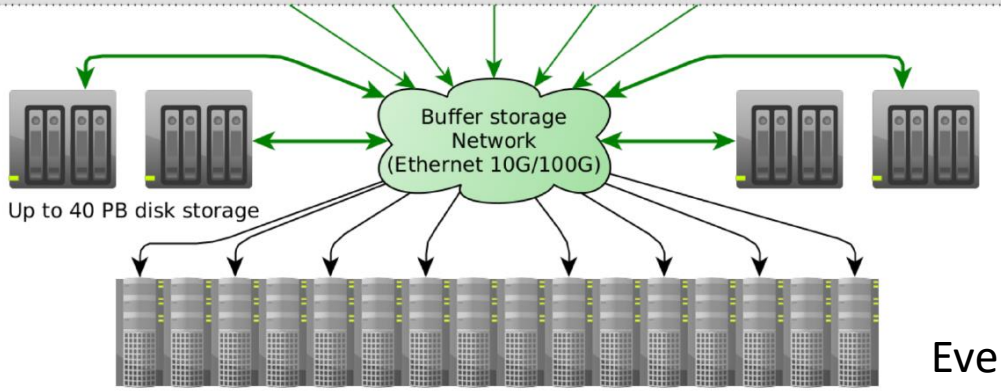
Real Time Analisis at LHCb

30 MHz
(non-empty pp)
5 TB/s



Event builder
& event filter 1

1 MHz
100 GB/s



100 KHz
10 GB/s

Event filter 2

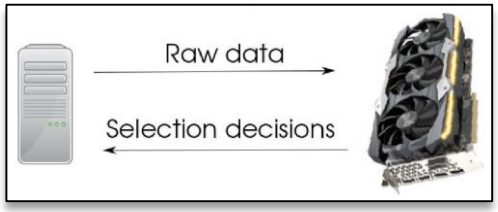
Event filter second pass (up to 4000 servers)

LHCb data
center at Pit 8

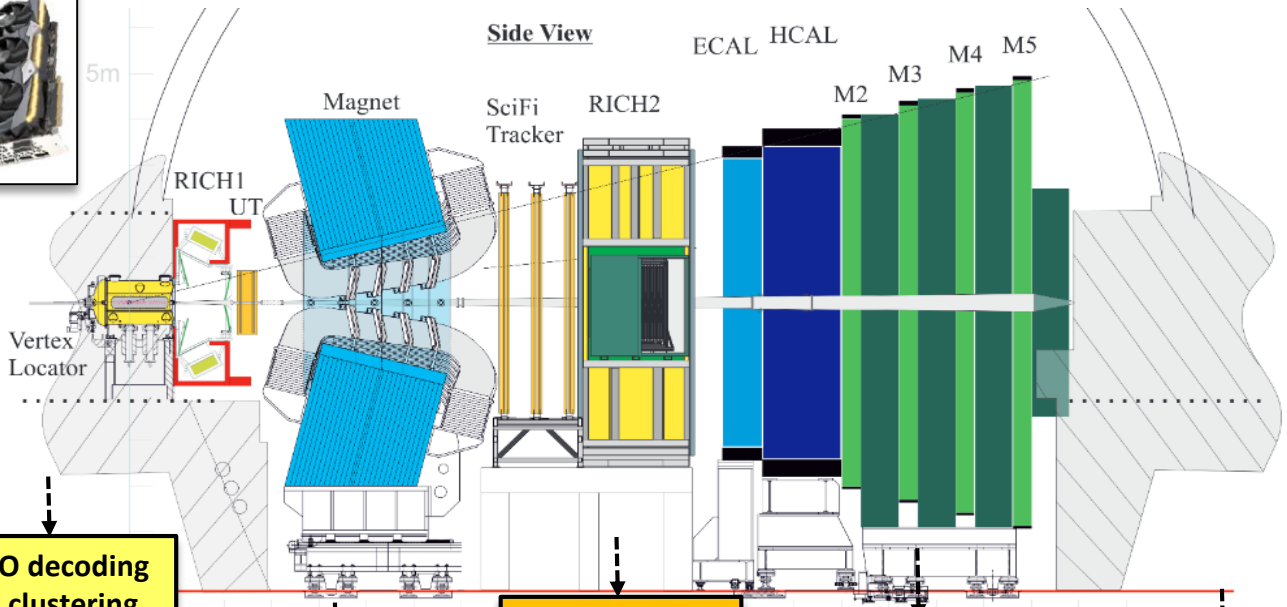
Real Time Analysis at LHCb

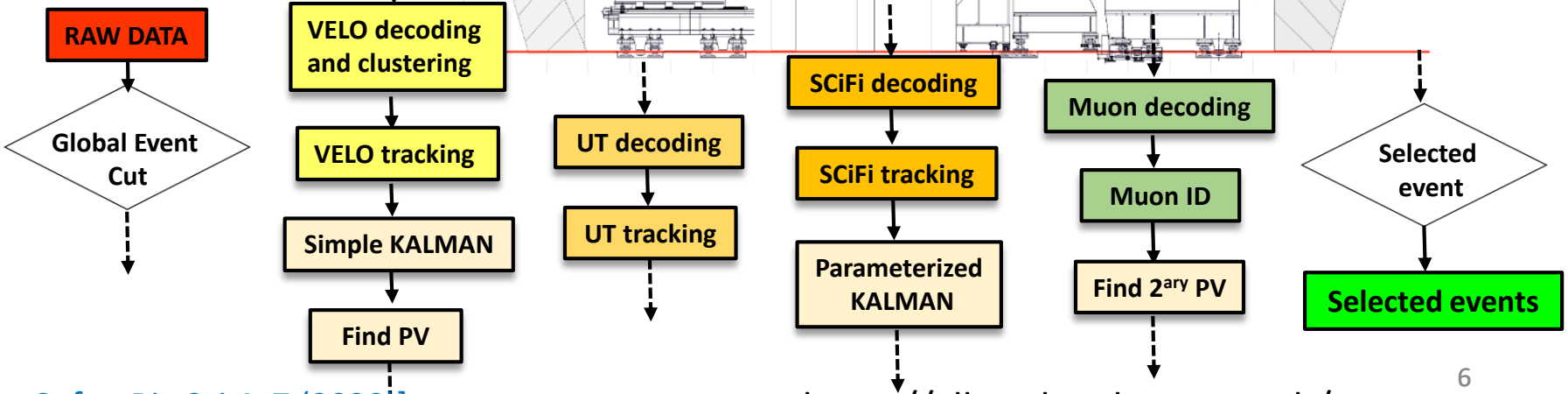
Allen: the LHCb high-level trigger 1 (HLT1) application on GPUs.

[LHCB-TDR-021] → Fast detector reconstruction in O(300) Nvidia RTX A5000



More than 300 algorithms with a high level of parallelization





[Com. Softw Big Sci 4, 7 (2020)]

https://allen-doc.docs.cern.ch/

The HIGH-LOW project at Valencia

"DESIGN OF HIGH PERFORMANCE ALGORITHMS FOR LOW POWER SUSTAINABLE HARDWARE FOR LHC EXPERIMENTS AND THEIR UPGRADES"

- Transversal project: ATLAS and LHCb experiments
- PIs: Luca Fiorini (ATLAS), & Arantza Oyanguren (LHCb)
- Funded by the Spanish Ministry of Science and Innovation (TED2021-130852B-I00)
- About 10 people (physicists + engineers + students)

Aim:

Benchmarking new hardware architectures and developing fast and high efficient algorithms with reduced power consumption

HL hardware:

- Rack K RETEX LOGIC-2 A600 42U F1000 PH
- APC Metered Rack PDU ZeroU 2G AP8
- SWITCH D-LINK DXS-1210-28T 24x 10GB

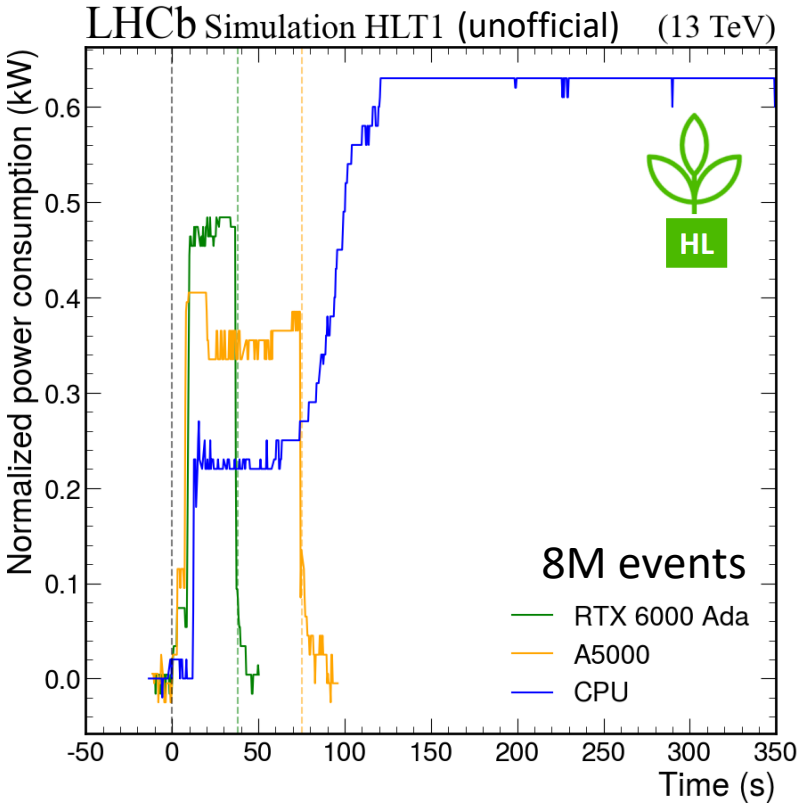
- T10G Dual Xeon Scalable HPC 10xGPU PCIe
  - 2 x Intel® Xeon® Gold 5318Y 2,10GHz 24 Cores 3.40 GHz
  - 8 x 32GB DDR4 3200MHz ECC REG
  - 1 x SSD Samsung 990 PRO 2TB M.2 NVMe 2280 PCIe 4.
  - 1 x Controladora BROADCOM MegaRAID 9560-16i PCIe
  - 8 x HD 10TB SAS 12Gb/s 7.200 rpm
  - 1 x NVIDIA® RTX™ A5000 24GB GB GDDR6 ECC
  - 1 x NVIDIA® RTX™ A6000 Ada Generation 48GB GB GDDR6 ECC
  - 1 x HBA Broadcom N210GBT Dual 10GbE RJ45

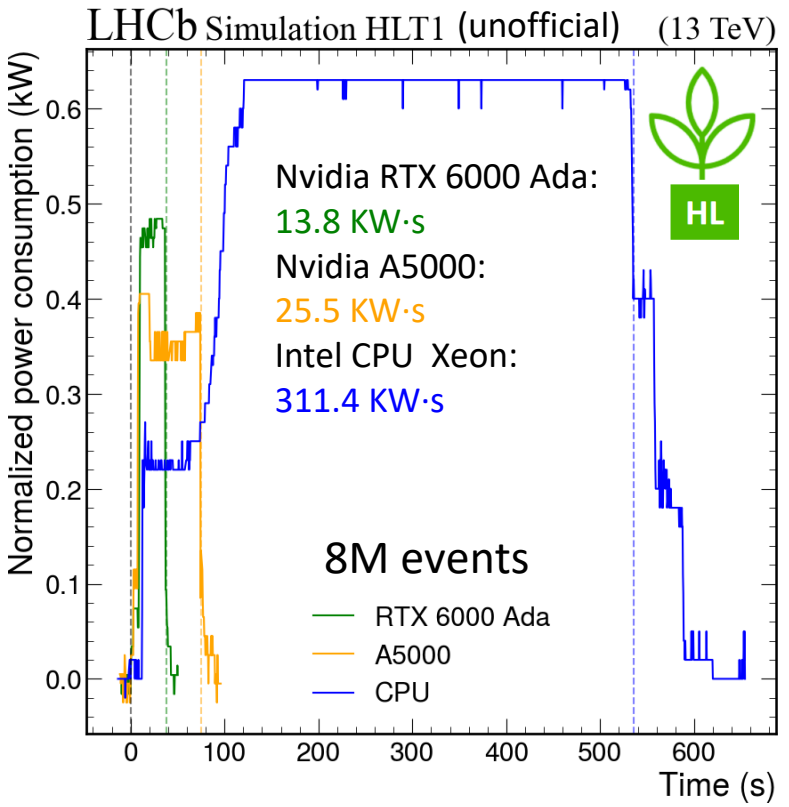
The HIGH-LOW project at Valencia

Hardware sustainability

1) Using the best available hardware:

Ex: LHCb Allen framework in RTX 6000 Ada and RTX A5000 (used at LHCb) .





The LHCb decision to run HLT1 on GPUs may have saved a factor 10 in energy consumption!

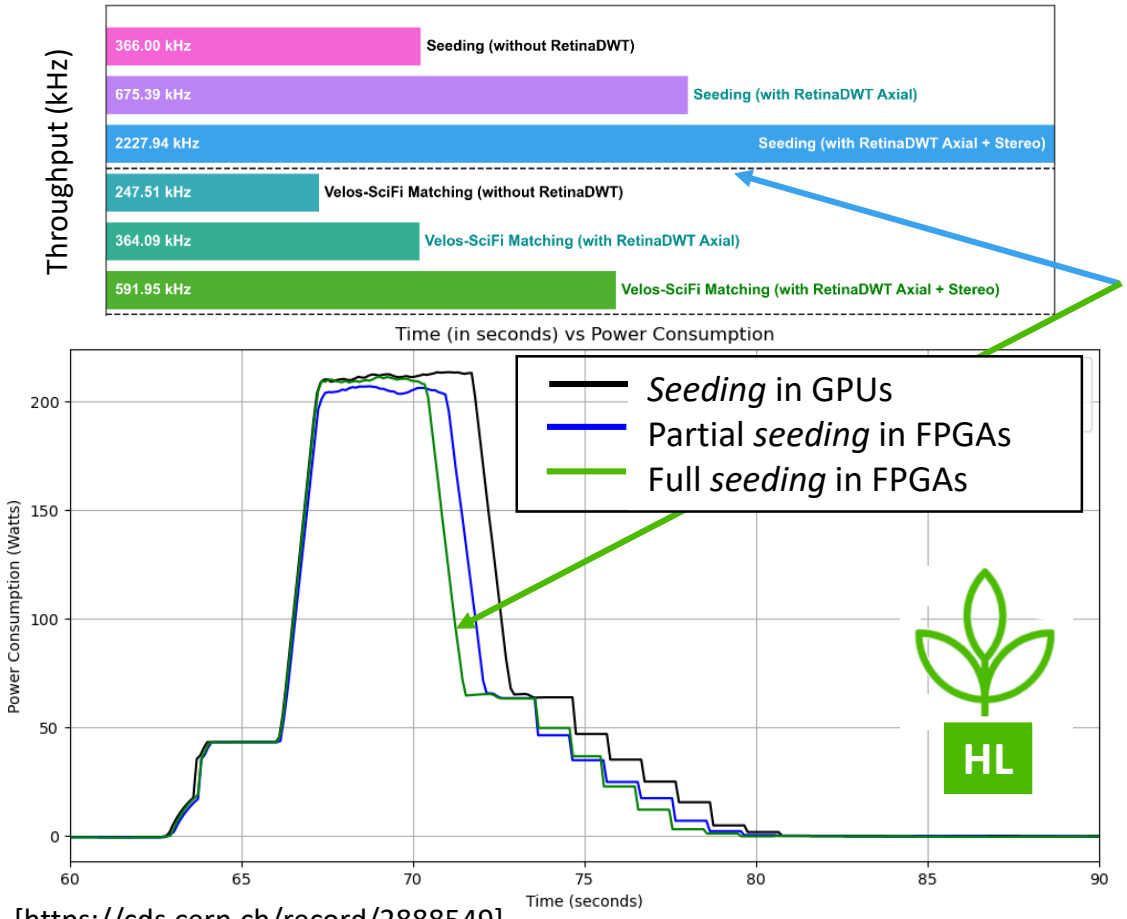
Comput Softw Big Sci **6**, 1 (2022). https://doi.org/10.1007/s41781-021-00070-2

Hardware sustainability

1) Using the best available hardware:

Offloading some tasks to FPGAs (RETINA framework – Pisa group) Stratix 10 (1SG280HN2)

[Framework TDR for the LHCb Upgrade II, CERN-LHCC-2021-012]



[https://cds.cern.ch/record/2888549]

Seeding algorithm for making tracklets in the last LHCb tracker (SciFi) in FPGAs:
throughput (#events processed/s) increases by 30%
→ Saving 6.2 mW·s/event

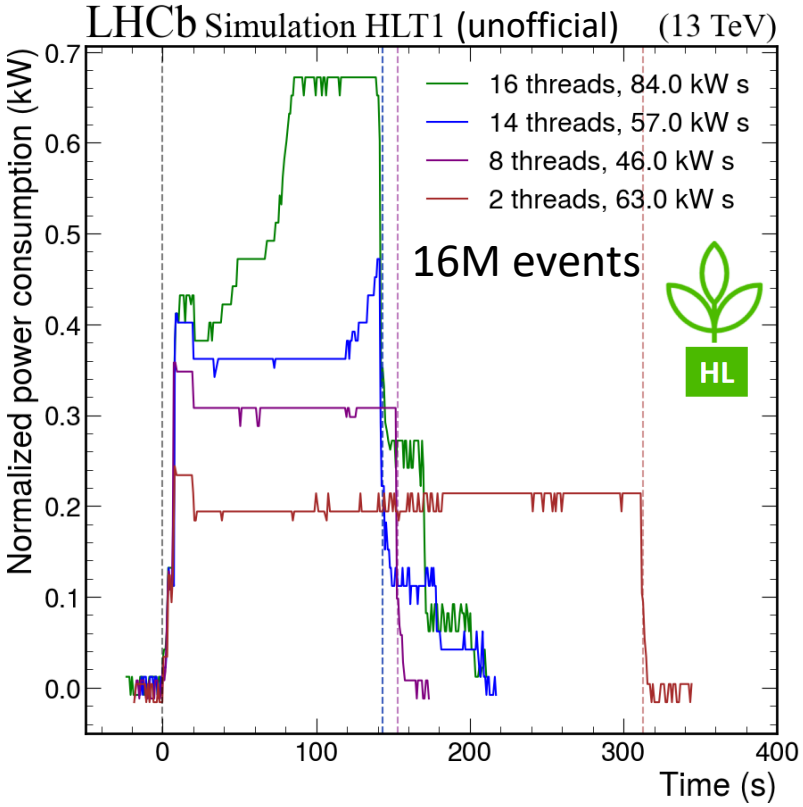
(for 30 MHz rate: 186kW/s)

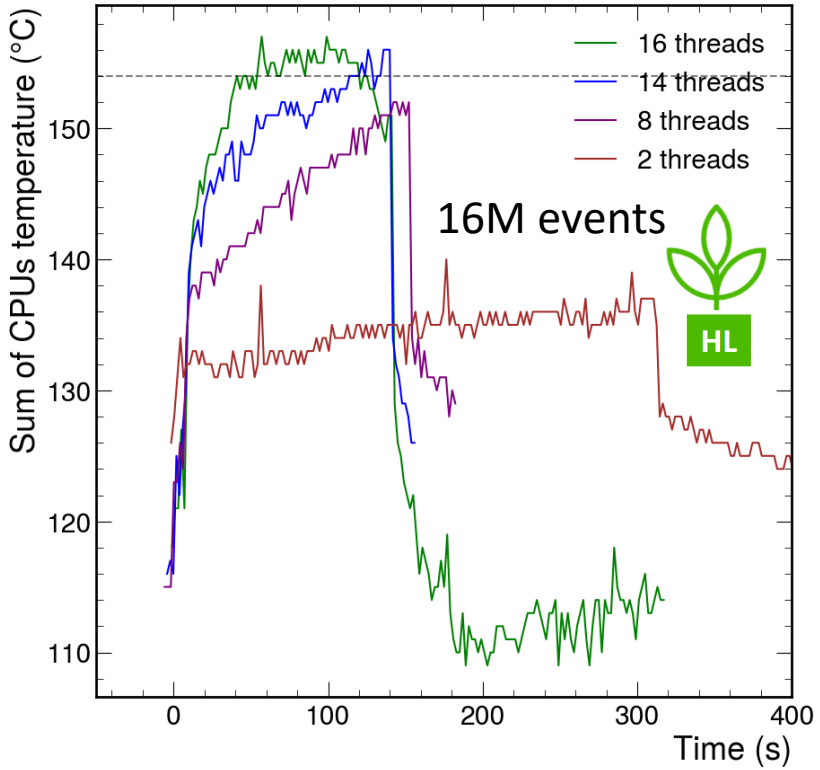
→ Use hybrid systems (FPGAs + GPUs) to take benefits of each one

Hardware sustainability

2) Optimizing the hardware utilization:

Dependence with the GPU parameters:
Ex: number of used threads



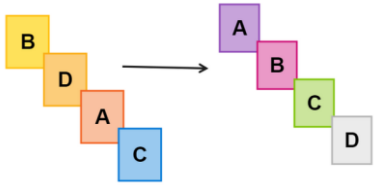


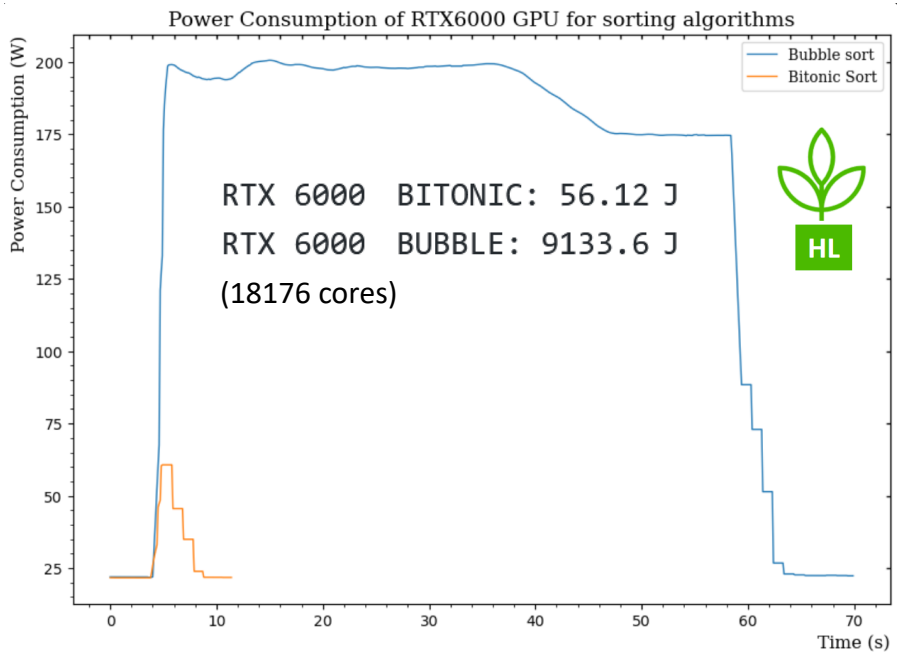
→ **Key point: best configuration to avoid systems become too hot**

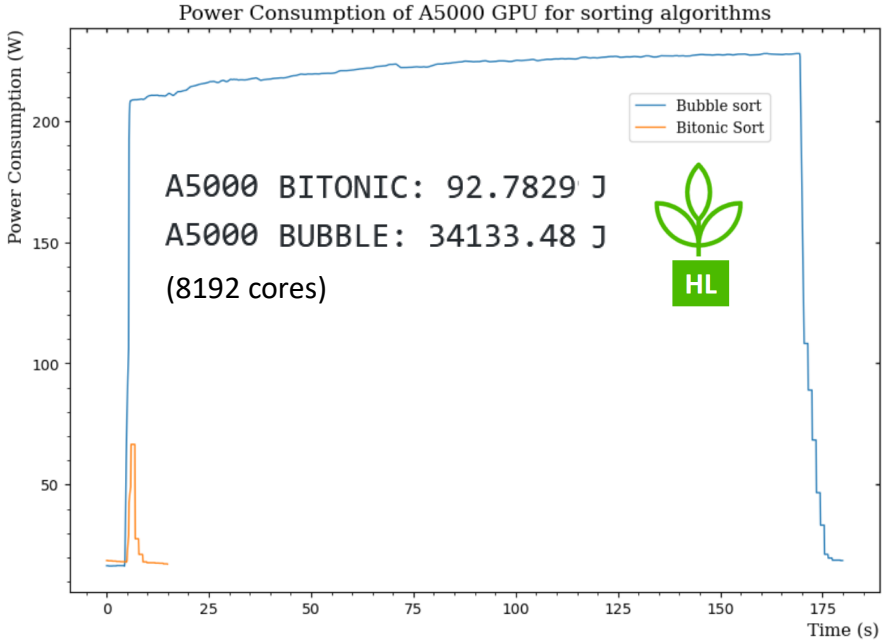
Software sustainability

2) Optimizing the software design:

A basic known example:
Sorting 4M elements with ***Bubble*** and ***Bitonic*** algorithms:







→ **Two orders of magnitude difference in energy consumption, with a high dependence of the hardware utilization!**

Software sustainability

2) Optimizing the software design:

HIGH THROUGPUT ↔ LOW POWER CONSUMPTION

- Proper parallelization
- Memory usage
- Balanced work distribution
- Instruction-work organization

….

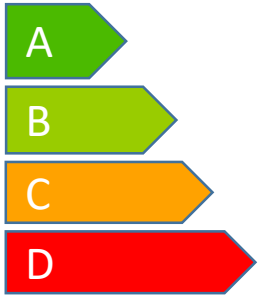
To take into account:

- *Performance vs energy*

Ex: which is the cost of 0.01% gain in efficiency for a tracking system/clustering algorithm? Per track? Per event? Per year? Vs final physics performance?

• We are developing a software tool to check the power consumption of an algorithm in a specific hardware

Greenific (in progress)



Conclusions & prospects

HL

- Use the best and more efficient available hardware (vs €)
- Optimize the utilization of the hardware
- Optimize the software design
- Take advantages of the correlations among them

Thank you!

Hardware sustainability