

# Heterogeneous Computing developments at the LHC experiments

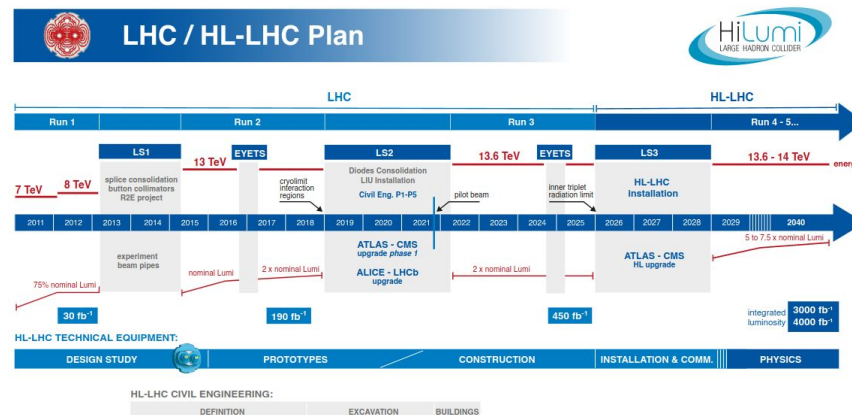
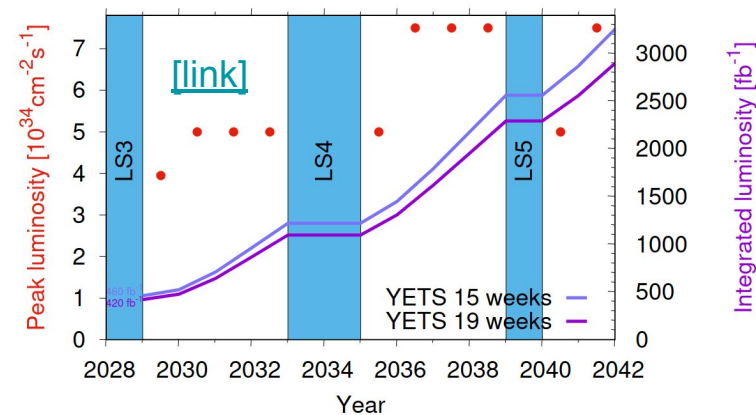
---

Wahid Redjeb<sup>1,2</sup>  
wahid.redjeb@cern.ch

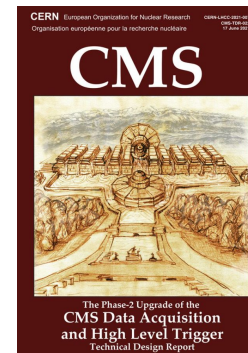
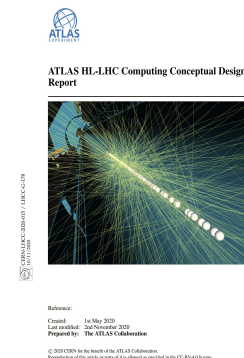
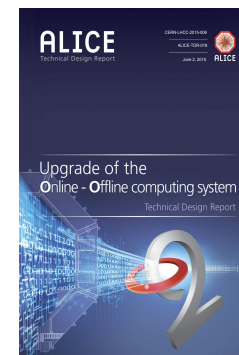
<sup>1</sup>CERN, European Organization for Nuclear Research, Meyrin, Switzerland  
<sup>2</sup>RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany,



- **New challenges for HEP computing, already started with Run-3**
  - ALICE: continuous readout Pb-Pb@50kHz
  - LHCb: 30MHz input rate at Software Trigger
- **Challenges starting from Run-4 for ATLAS-CMS**
  - 3x instantaneous luminosity
  - 10x simulated data
  - >3x trigger rate
  - 60 → 140-200 PU
- **Need of getting HEP software ready**
  - Management of exabyte-scale data
  - Exploit new available hardware and HPC centers
  - Performance Portability for sustainable computing

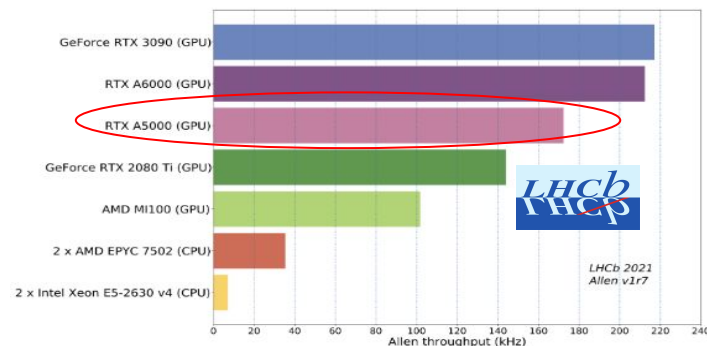
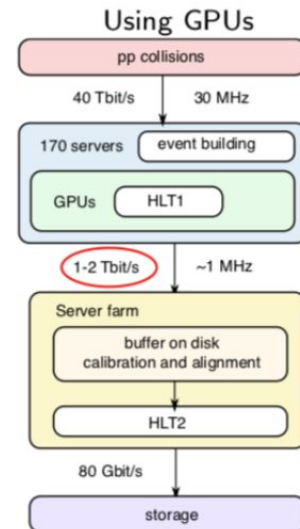
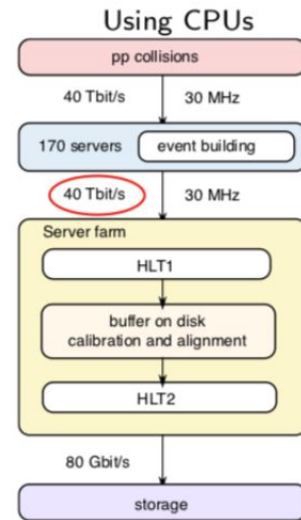


- **ALICE** and **LHCb** implemented their TDRs for **Run-3 and Run-4**
  - Upgrade of O<sup>2</sup> facility
  - Allen: HLT1 on GPU
- **ATLAS Phase-2 CDR out in 2020 while CMS CDR expected next year**
  - Solutions to stay in the computing budgets
    - R&D needed: simulations, accelerators, usage of HPC centers
- **New challenges with HL-LHC**
  - Unprecedented read-out rates and complexity
- **And computing technologies rapidly changing**
  - Leverage state-of-the-art technologies is mandatory
    - Need training and more synergies with computing experts



- New detector upgrades and new Trigger System
  - **~5x instantaneous Luminosity**
  - **1MHz → 30MHz input rate to software trigger**
  - **Full software-based trigger** (HLT1 + HLT2)
    - **FPGAs**-based clustering for Silicon Pixel detector
    - HLT1: **GPU** based reconstruction
      - Simplified and Faster reconstruction
      - Reduces output rate by a factor 30-60
    - HLT2: **CPU**-based full reconstruction
      - Offline-Quality reconstruction
    - Alignment and Calibrations performed on buffered data from HLT1 on CPU
- **Achieving 30MHz with less than 200 GPUs!**

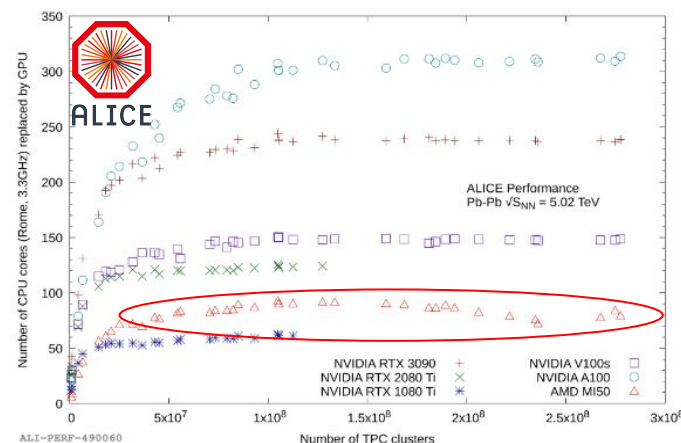
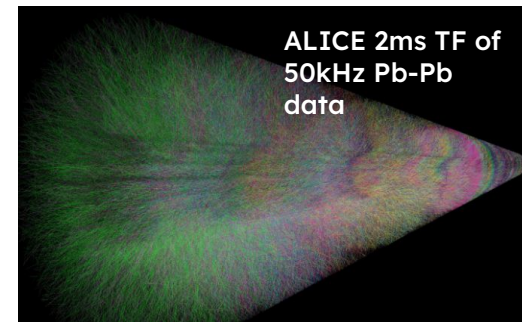
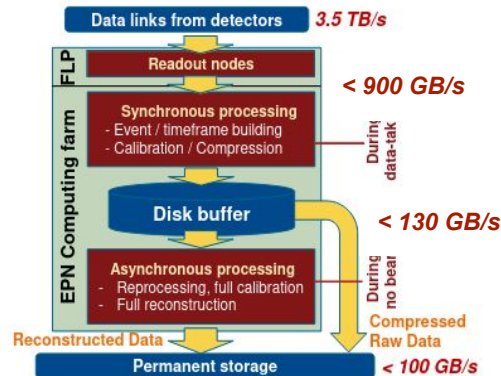
- [LHCb High-Level-Trigger TDR](#)
- [LHCb Computing Model TDR](#)
- [Commissioning LHCb's GPU high level trigger](#)



# What's going on in Run-3? - ALICE



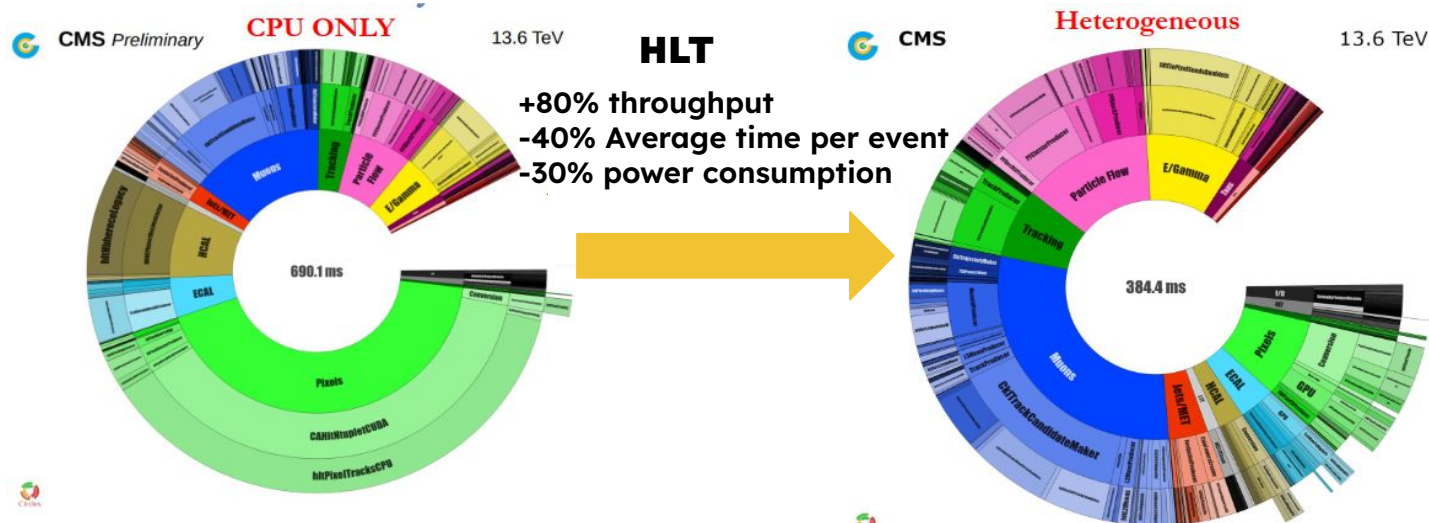
- **Continuous readout at 50kHz for Pb-Pb**
  - No trigger, no event rejection: **1kHz** → **50kHz**
  - Time-frame (TF) of 2.5-20ms instead of event acquisition
  - 100x more data to **process** and **store**
- **Upgrade in O<sup>2</sup> facility**
  - **First Level Processing (FLP)**:
    - Readout + **FPGA** corrections
  - **Event Processing Node (EPN)**
    - Fully equipped with **AMD GPUs**
    - **Synchronous Processing** (online)
      - TF building + Calibration + Compression
    - **Asynchronous Processing**
      - Full calibration + Full reconstruction
- **Replacing 80 AMD CPU 3.3GHz Cores with a single AMD-Mi50 GPU**



ALI-PERP-490060  
THE O2 SOFTWARE FRAMEWORK AND GPU USAGE IN ALICE ON AND OFF-LINE RECONSTRUCTION IN RUN3 - CHEP2023 - David Rohr, Giulio Eulisse

[ALICE Upgrade of the Online-Offline computing system](#)

- Adopted **GPUs** at the HLT (200 nodes, 2CPU - 2 GPU NVIDIA T4)
  - **HCAL**, **ECAL**, **pixel local** reconstruction and **pixel tracking**.
    - Significant part of the HLT
- Lot of R&D in Performance Portability
  - **Alpaka** integration foreseen for 2023 Data Taking
- Aiming to offload **10% of** (Run-3 and Phase-2) **offline reconstruction** by end of 2023



- [The CMS heterogeneous reconstruction - CHEP2023, F.Pantaleo](#)
- [Run-3 Commissioning of CMS Online HLT reconstruction using GPUs - CHEP2023, G. Parida](#)

Online reconstruction time measured under realistic conditions, on 64000 proton-proton events with an average pileup of 56 collisions, collected on October 7th 2022 (run 359998, luminosity sections 242-243), on a full Run-3 HLT node (2x AMD Milan 7633 + 2 NVIDIA T4)

- **Adapt new multi-threaded framework AthenaMT**

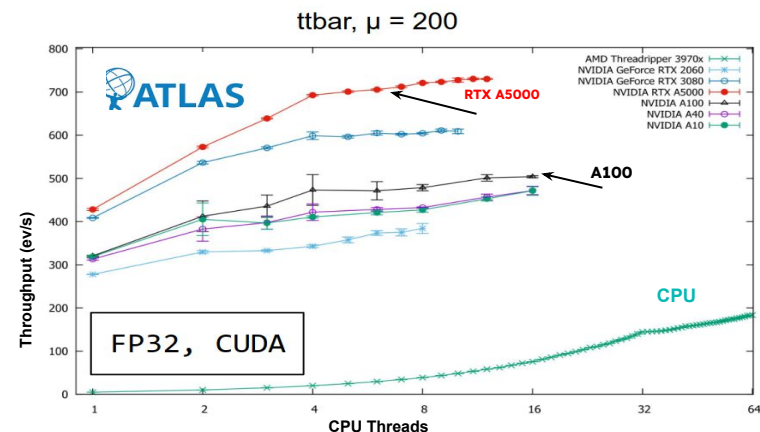
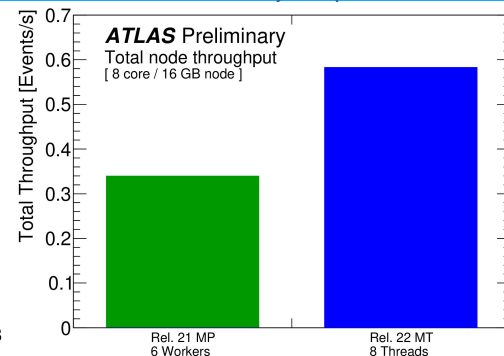
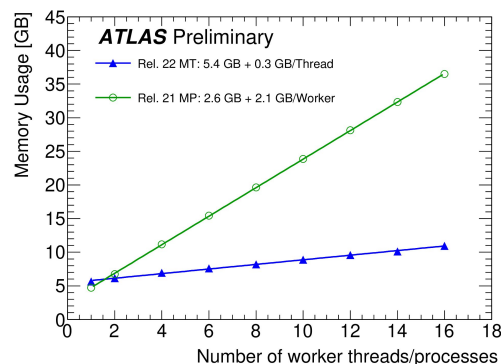
- Exploit Gaudi TBB scheduler
- Inter-event
  - Multiple-event in parallel
- Intra-event
  - Multiple-algorithm in parallel
- Intra-algorithm
  - Parallelism in the algorithm

- **Several improvements in L1Calo and HLT Algorithms**

- Usage of **FPGAs** at L1 for features extraction
- Adapt HLT to Multi-Threaded framework

- **ATLAS: several R&D projects for using accelerators**

- GPU Tracking with ACTS
- FastCaloSim on GPU

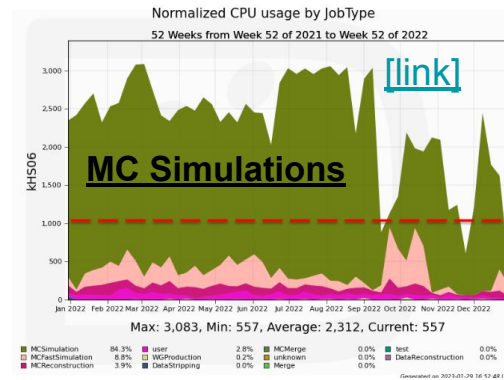


[Performance of Multi-threaded Reconstruction in ATLAS](#)

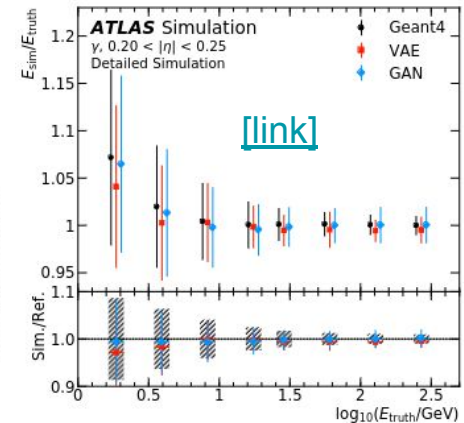
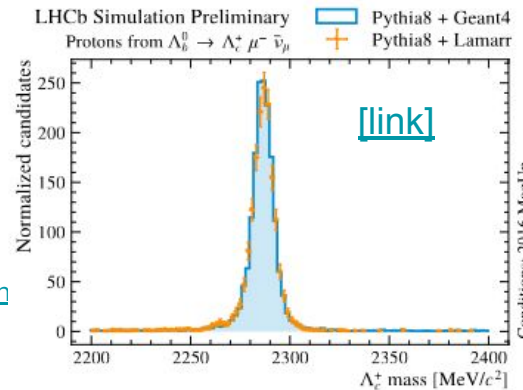
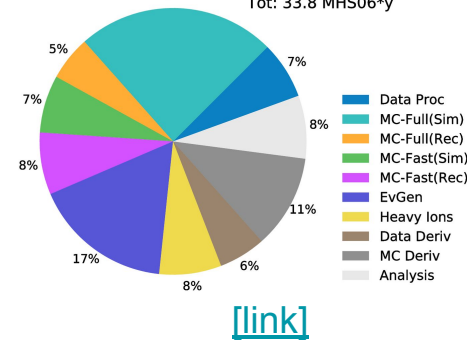
- [tracc - A \(Close To\) Single-Source Tracking Demonstrator on CPUs/GPUs - A. Krasznahorkay](#)

- **Simulations are dominating Run 3 CPU usage**
  - ALICE 50%, LHCb 90%, ATLAS 50%
- **Lot of R&D towards new techniques and use of accelerators**
  - AdePT and Celeritas for full GPU simulation
  - FastSim using parameterized simulations
  - FastSim using ML Techniques

- [Lamarr the LHCb Ultra-Fast Simulation Framework - L.Anderlini](#)
- [Deep generative models for fast photon shower simulation in ATLAS](#)

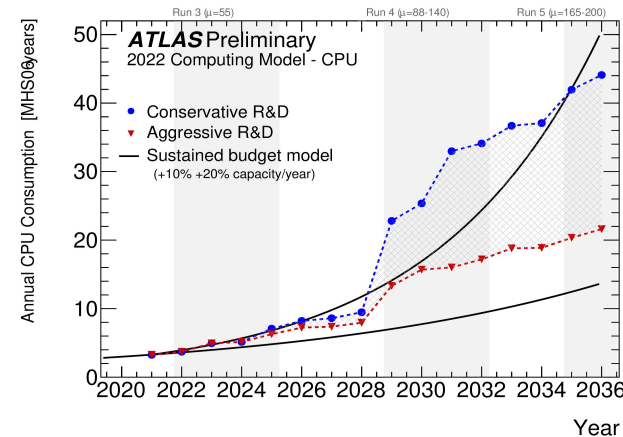
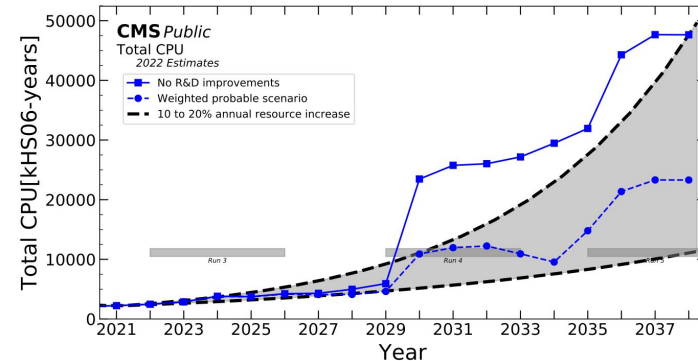


**ATLAS Preliminary**  
2022 Computing Model - CPU: 2031, Conservative R&D  
24% Tot: 33.8 MHS06\*y





- LHC Luminosity  $5\text{-}7.5 \times 10^{34}$ 
  - 3x events
  - Increasing event complexity: PU 140-200
- **CMS and ATLAS fully upgraded**
  - Increased Trigger Rate
    - ATLAS: 3.2 kHz  $\Rightarrow$  ~10 kHz
    - CMS: 2.6 kHz  $\Rightarrow$  ~7.5 kHz
- **Intensive Computing R&D needed to stay in the computing model budget**
  - Generators and detector simulation
  - Reconstruction algorithms
  - R&D needed to exploit new accelerators and HPC Centers



- **HPC centers offer unique opportunities**
  - Heterogeneity: CPU, GPU, FPGA
  - Different architectures (x86, ARM, Power9)
    - As well as for GPUs (Nvidia, AMD, Intel)
  - Usage of accelerators to offload compute-intensive tasks and fully exploit node capabilities
- **LHC workflows at HPC: not a trivial task**
  - Different connectivity requirements
  - Different hardware setups (RAM, local storage)
  - Authentication
- **Lot of expertise and efforts needed**
  - Adapt HTC to HPC
    - Infrastructures, policies
  - Performance Portability plays an important role to achieve flexibility

AMD + Nvidia



AMD

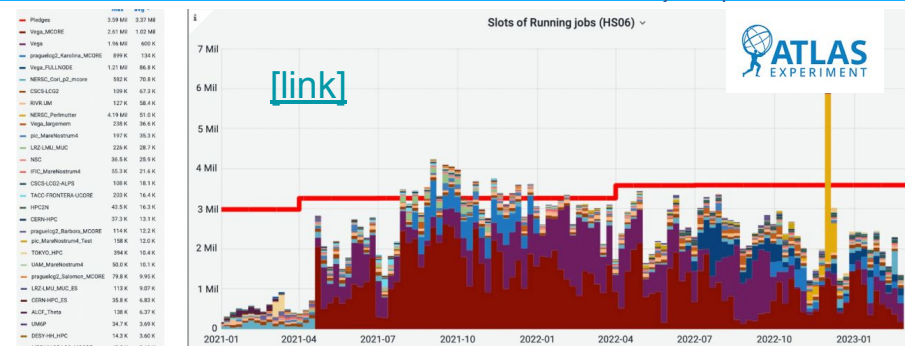


Intel



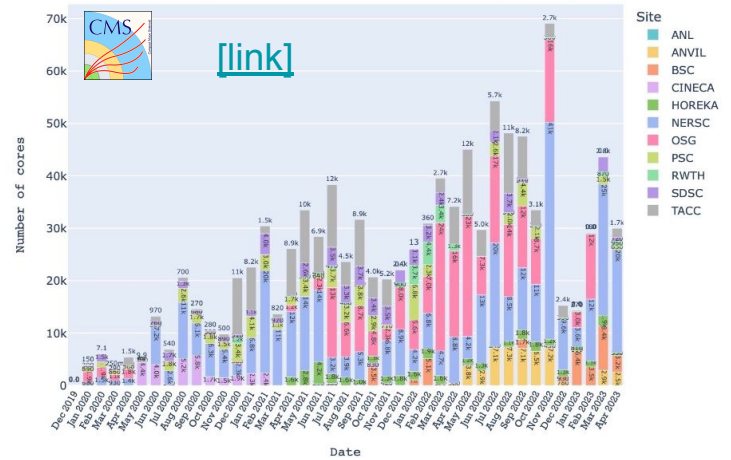
Intel + Nvidia

- **ALICE:** Marconi @ CINECA, Cori and Lawrenceium @ LBNL for O<sup>2</sup> MC Production
  - Also ported O<sup>2</sup> to ARM, under study
- **ATLAS:** Offload MC Production on many HPC centers: @ NERSC, NSF, CINECA, BCS, Vega, Karolina, ...
  - Recently crossed 1M simultaneous cores
- **CMS:** Offload Reconstruction and MC + digitization
  - Through HEPCloud or sites extension
  - up to 10% CPU capacity of CMS
- **LHCb:** Offloading MC Production
  - Efforts in developing tools to access HPC sites with different requirements



CMS Public

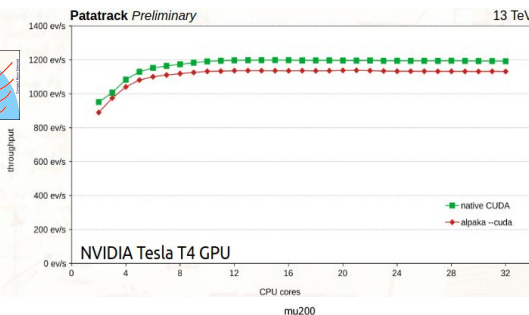
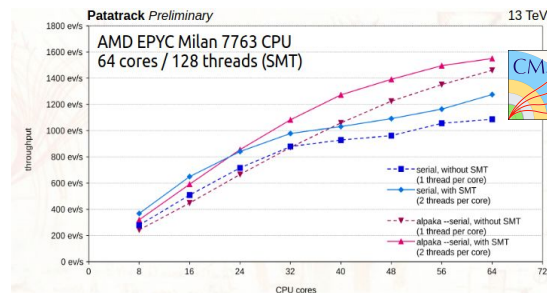
Number of Running CPU Cores on HPCs - Monthly Average



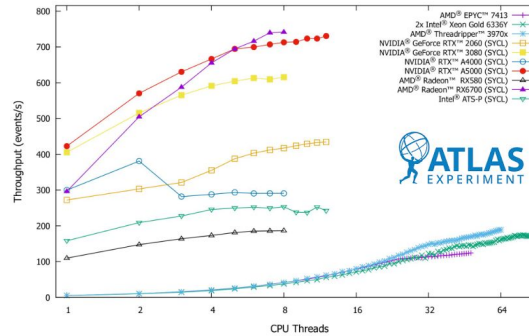
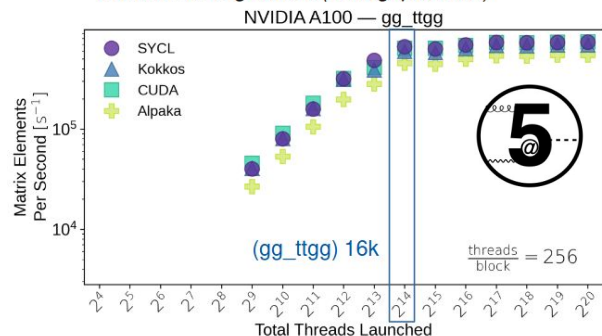
- **Several accelerators available on the market**
  - CPUs, GPUs, FPGAs, ASICs
- **And also several vendors**
  - Intel, AMD, ARM, Power, NVIDIA
  - Different programming libraries
- And we must be ready for new devices and solutions
- **Portable code is the solution**
  - Long-Term maintainability, and testability
  - Avoid code duplication
  - Same algorithms for different hardware
  - Support for new devices
- **Performance Portability Libraries**
  - Abstraction layer to hide backend implementation
  - Express parallelism on all the backends
  - Alpaka, Kokkos, SYCL



- **ALICE** opted for a custom portability layer
  - Generic C++ wrappers to support CUDA, HIP and CPU
- **CMS** adopted **Alpaka** as portability layer for Run-3
  - Migrating CUDA code to Alpaka
  - Close-to-native performance achieved with Alpaka
  - Evaluating other portability layers for Phase-2
- **ATLAS** and **LHCb** are evaluating Performance Portability Library
  - SYCL, Kokkos
- **R&D** for event generation with Madgraph
  - SYCL, Kokkos, Alpaka



Variable GPU-grid size (throughput scan)



- [Adoption of the alpaka performance portability library in the CMS software - A.Bocci](#)
- [Integrating oneAPI/SYCL in the ATLAS Software](#)
- [Speeding up Madgraph5\\_aMC@NLO through CPU vectorization and GPU offloading](#)

- Heterogeneous Architectures and multi-threaded platforms are fundamental to fully exploit the Run-3 and HL-LHC physics program
  - Demonstrated in Run-3 by all the experiments
- R&D needed **now** to face the challenges of HL-LHC
  - New developments in parallel computing and ML
  - Fully exploit HPCs
    - Performance Portability extremely important for flexible software
    - Maintain synergies with HPCs and experiments
- Need retain and train new computing experts
  - Hackathons are nice way to kickstart projects, to train and support people

*Special thanks to:*

*Zach Marshall (ATLAS), Concezio Bozzi, Michel De Cian (LHCb), Stefano Piano, David Rohr (ALICE), Danilo Piparo, Christoph Wissing, Vladimir Ivantchenko (CMS)  
for providing the material and the help in preparing this talk*