The CMS Heterogeneous Computing Revolution

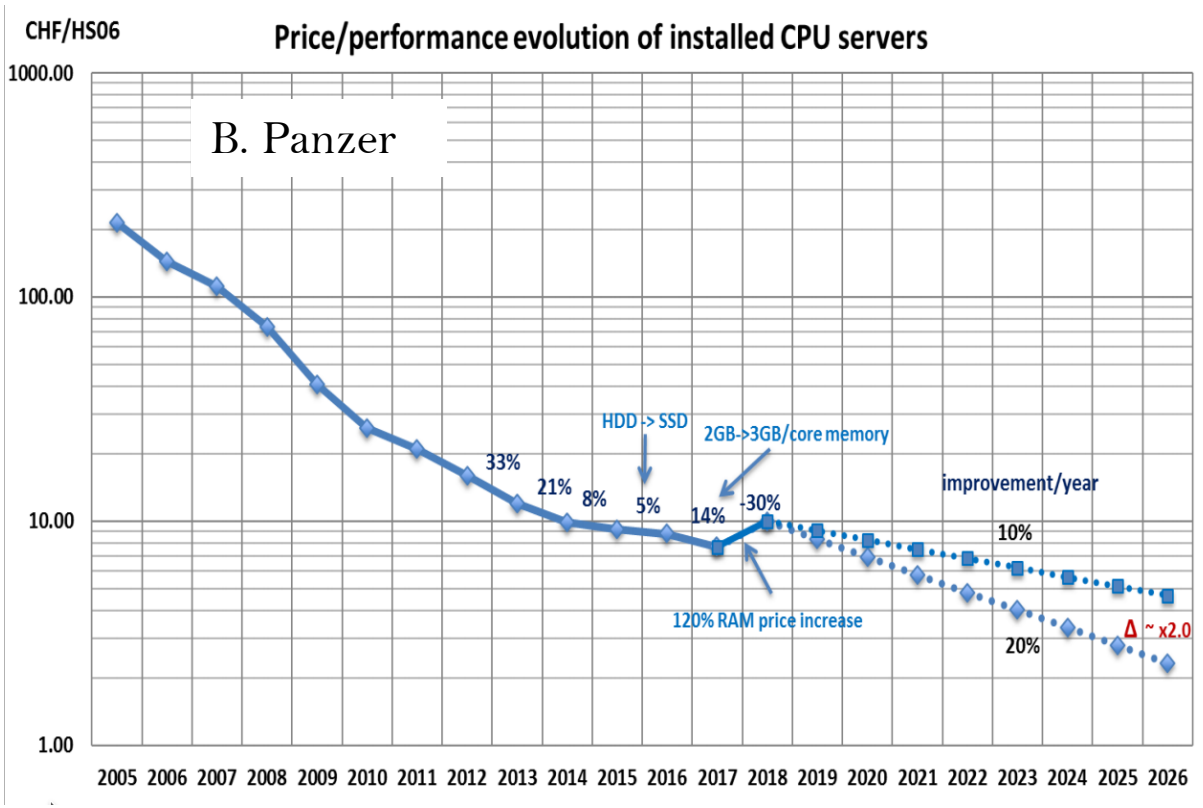
Felice Pantaleo: for the Patatrack Team
for the CMS Experiment

felice@cern.ch

Patatrack

- Patatrack is a software R&D incubator
- Born in 2016 by a very small group of passionate people
- Interests: algorithms, HPC, heterogeneous computing, machine learning, software engineering
- Lay the foundations of the online/offline heterogeneous reconstruction starting from 2020s

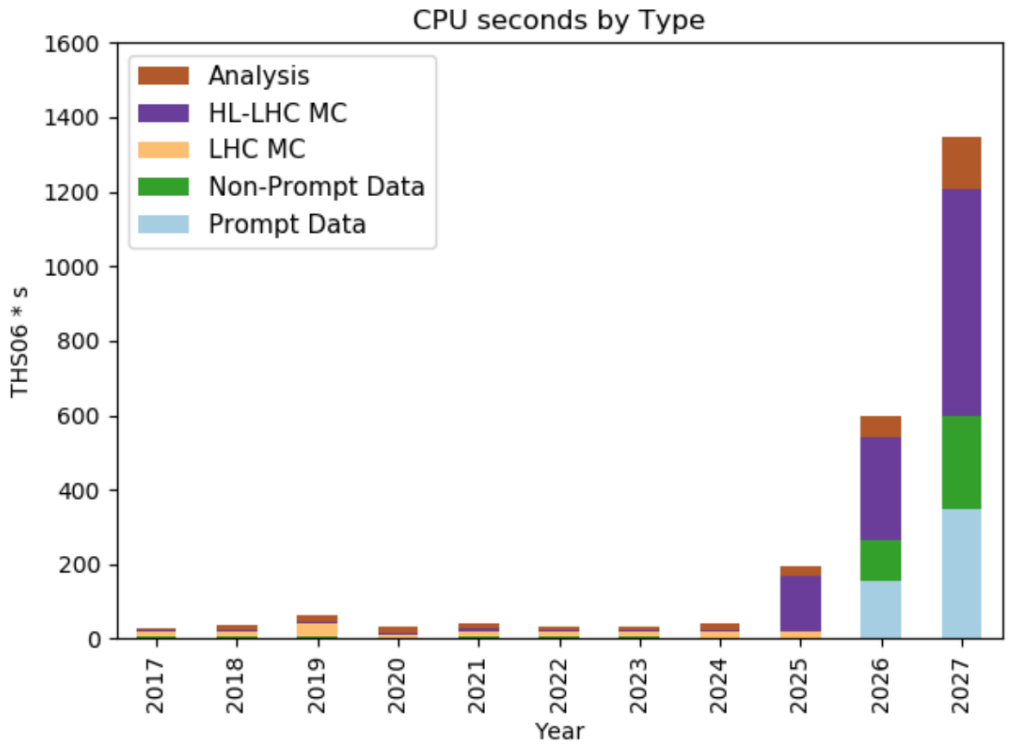
A reminder that…



CPU evolution is not able to cope with the increasing demand of performance

…CMS is in a computing emergency

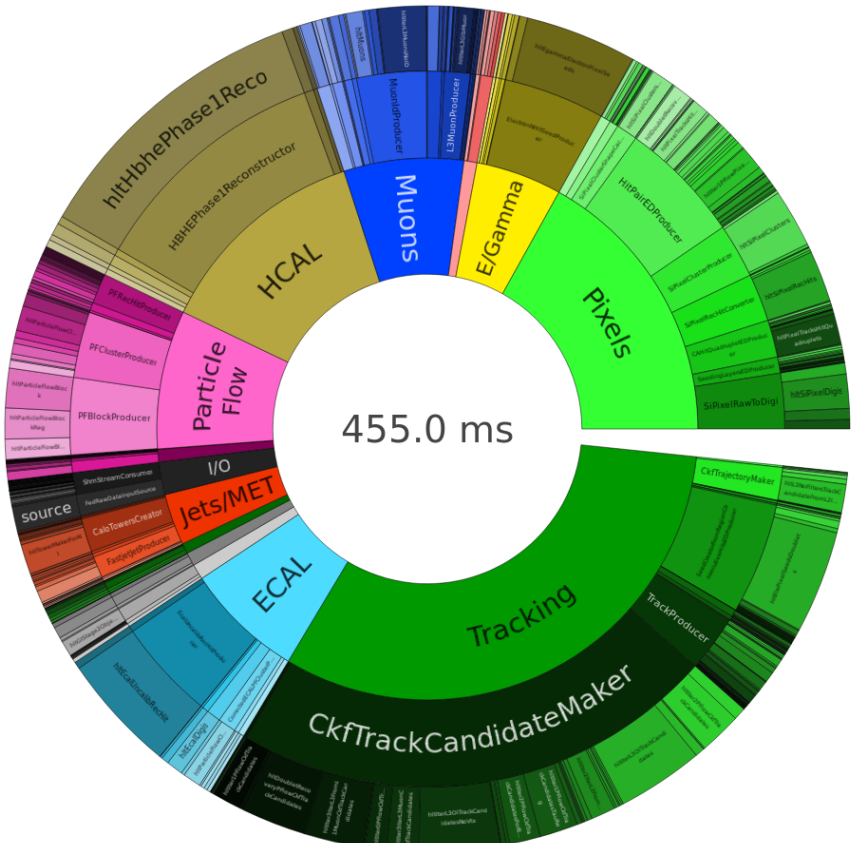
- Performance demand will increase substantially at HL-LHC
- 30x more CPU performance offline and online



Today

- High Level Trigger
  - readout of the whole detector with full granularity, based on the CMS software, running on 30,000 CPU cores
  - Maximum average latency is ~600ms with HT

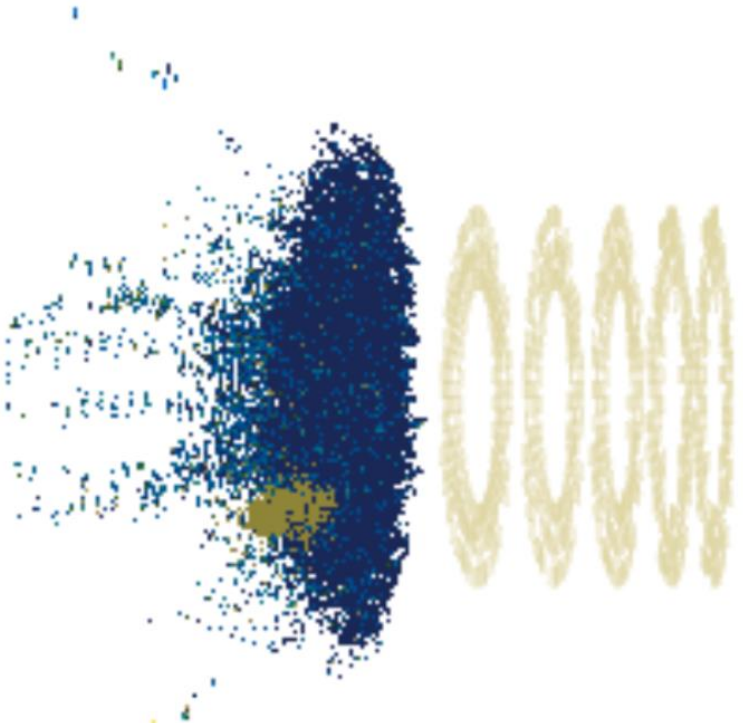
- Offline reconstruction takes ~25s/evt at PU50



Circles

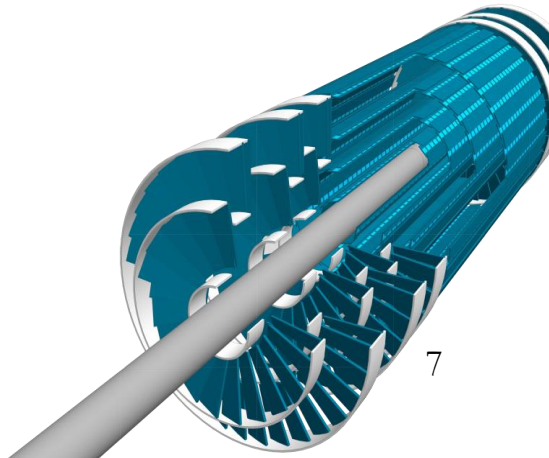
The CMS Trigger in Phase 2

- Level-1 Trigger output rate will increase to 750 kHz (7.5x)
- Pileup will increase by a factor 3x-4x
- The reconstruction of the new highly granular Calorimeter Endcap will contribute substantially to the required computing resources
- Missing an order of magnitude in computing performance



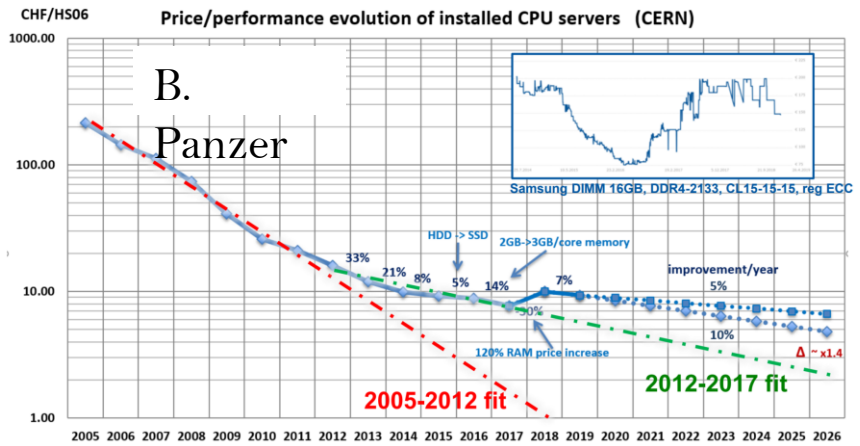
Heterogeneous Run3 HLT Farm

- We would like to exercise a Heterogeneous HLT farm well before Run4
- Targeted Pixel Tracks and Vertices, ECAL and HCAL recos
- What does CMS gain in the short term?
  - Better physics performance
  - Reconstruction able to run on Supercomputers
  - Expertise in Heterogeneous computing



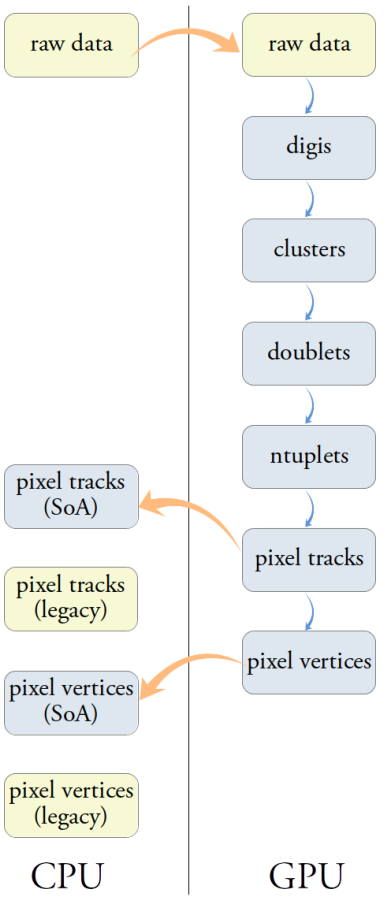
Why should our community care?

- Accelerators are becoming ubiquitous
  - Driven by more complex and deeper neural networks
  - Details hidden to the user by the FW
- Better Time-to-Solution, Energy-to-Solution, Cost-to-Solution
- Experiments are encouraged to run their software on Supercomputers
  - We are not using their GPUs
  - Summit: 190PFLOPS out of 200PFLOPS come from GPUs
- Training neural networks for production workflows is a negligible part
- Redesigning our algorithms and data structures to be well digested by a GPU can speed it up also when running on CPUs
- Maintaining and testing two codebases might not be the most sustainable solution in the medium/long term
- Not a showstopper at the moment, but will become one when we will transfer ownership of the code to the collaboration
- In the long term other accelerators might appear

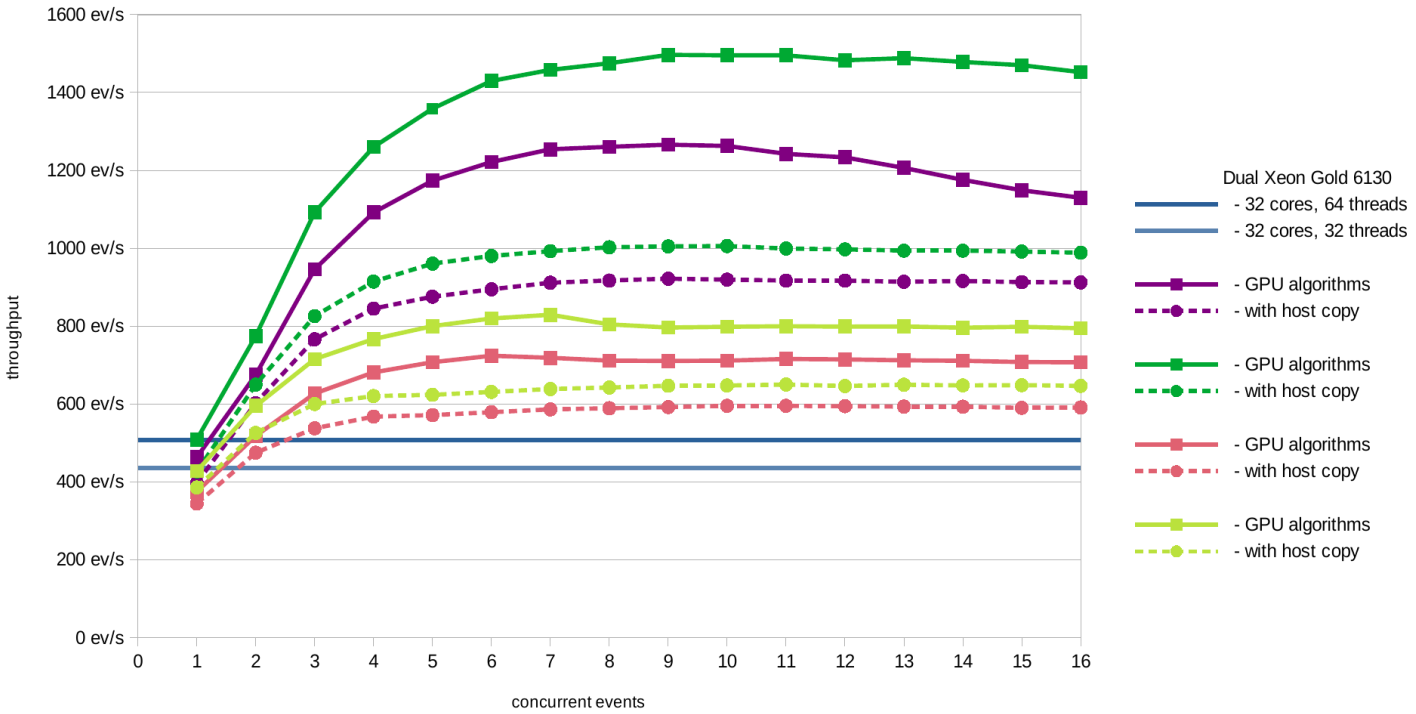


Patatrack Pixel Reconstruction Workflow

- Full Pixel Track reconstruction in CMSSW
  - from Raw data decoding to Primary Vertices determination
- Raw data for each event is transferred to the GPU initially (~250kB/event)
- At each step data can be transferred to CPU and used to populate "legacy" event data
- The standard validation is fully supported
- Integer results are identical

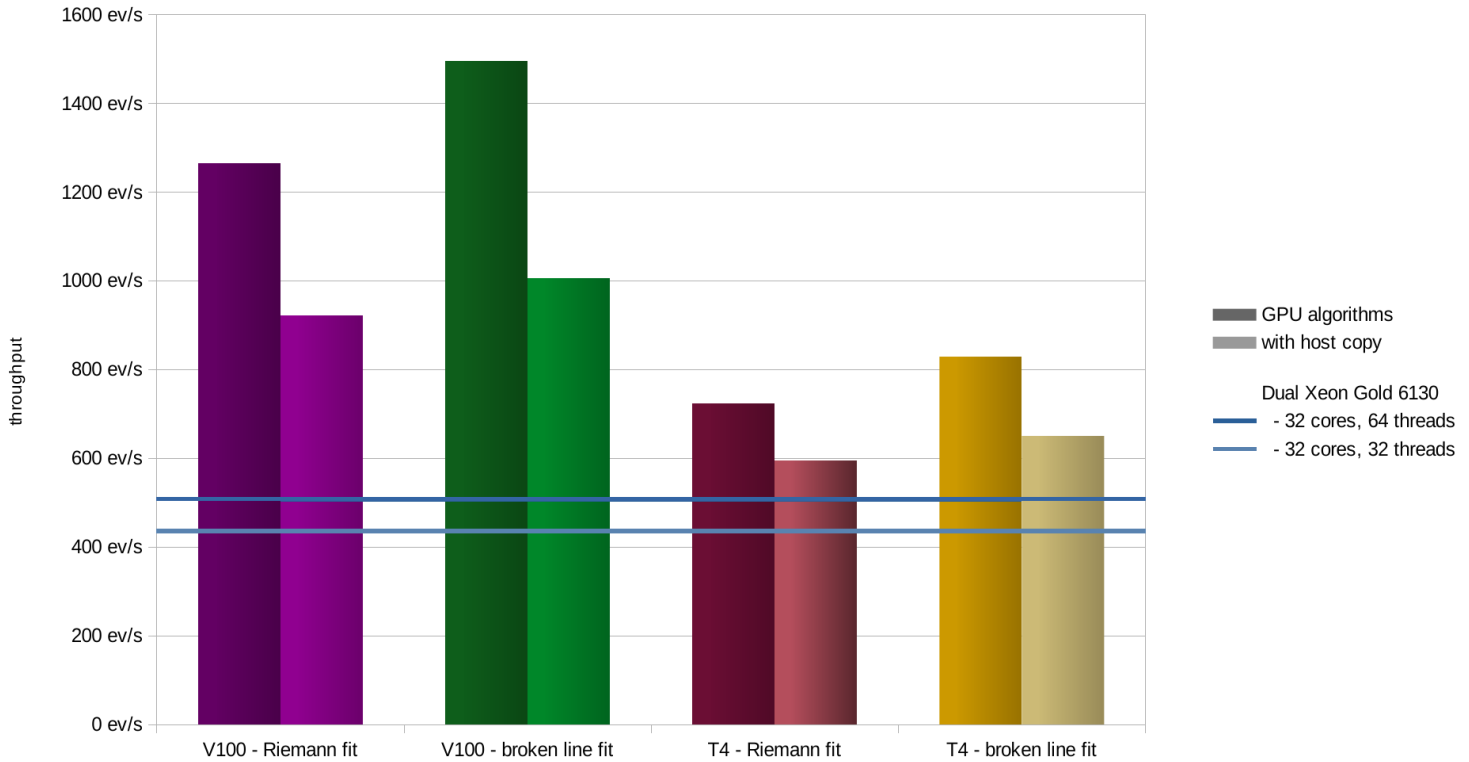


Computational Performance

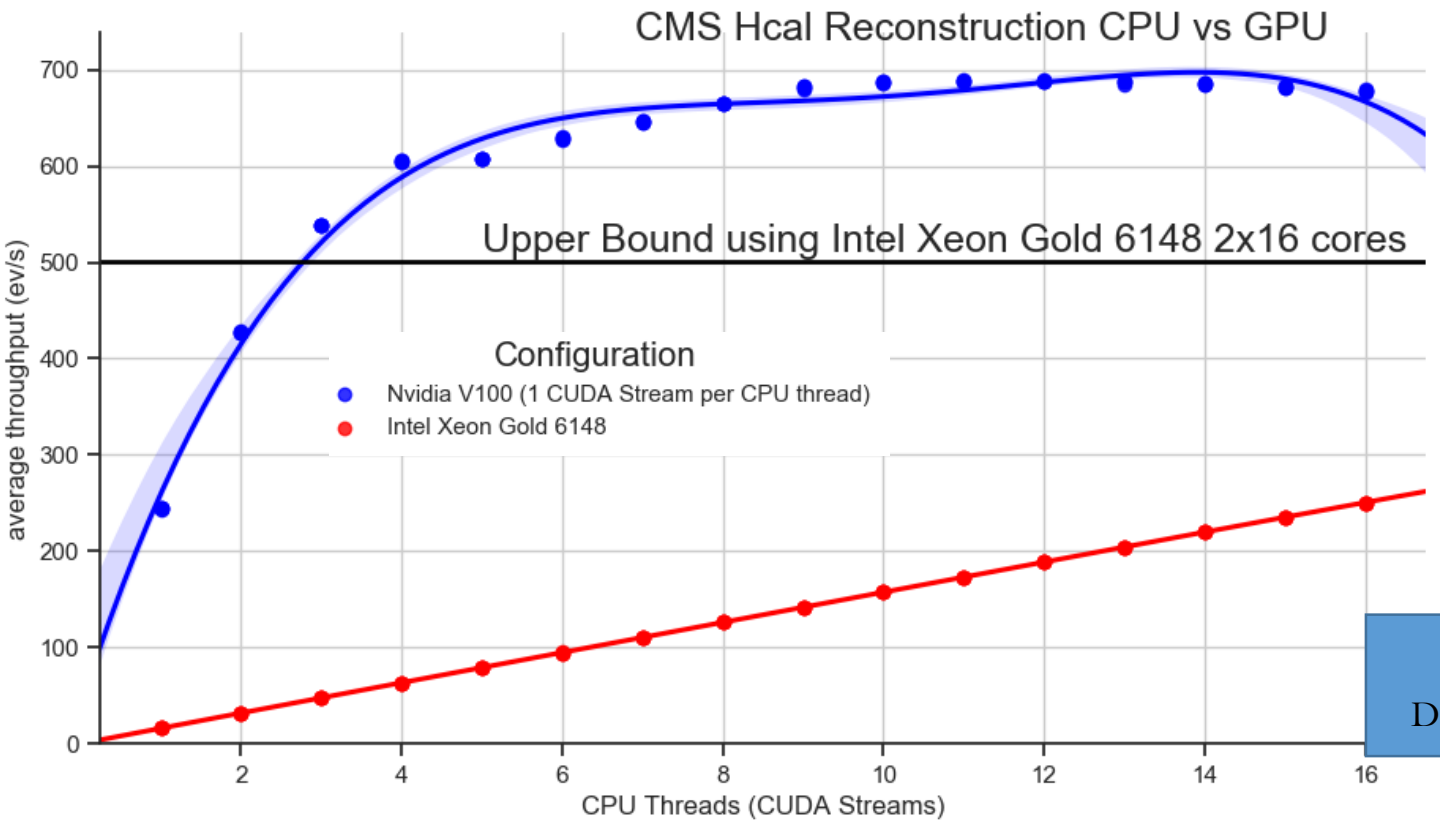


Pixel reconstruction consumers can either work directly on the GPU or ask for a copy of the tracks and vertices on the host

Computational Performance



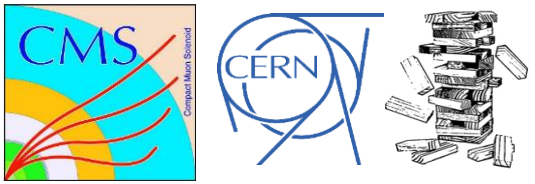
Pixel reconstruction consumers can either work directly on the GPU or ask for a copy of the tracks and vertices on the host



V. Khristenko,

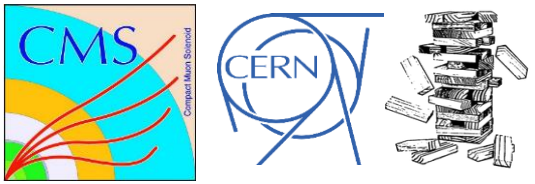
DEEP-EST, openlab

Back on the envelop calculation

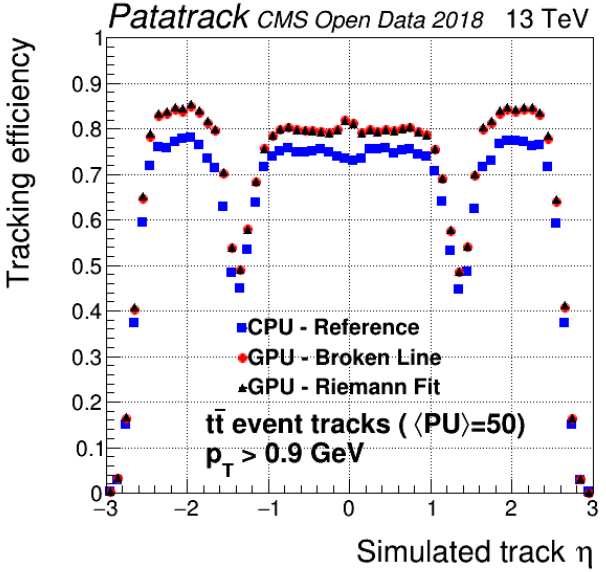
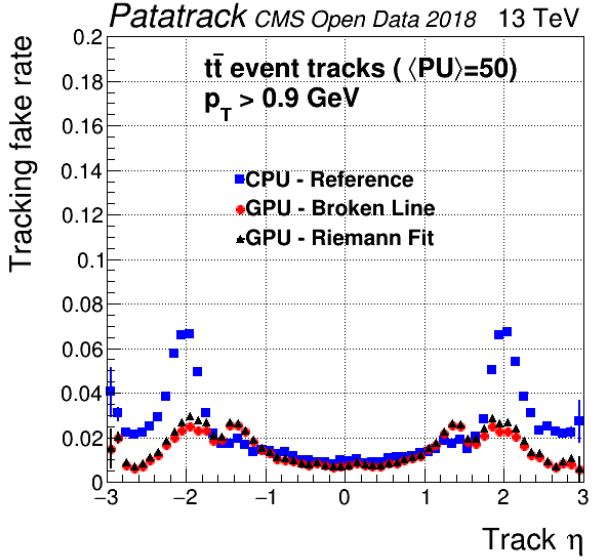
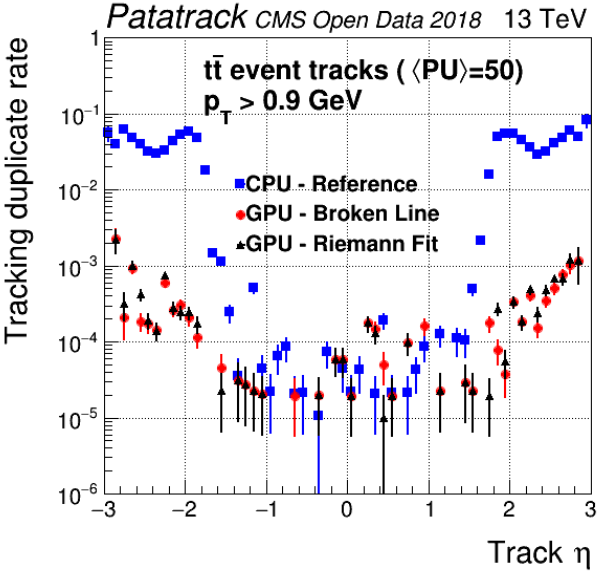
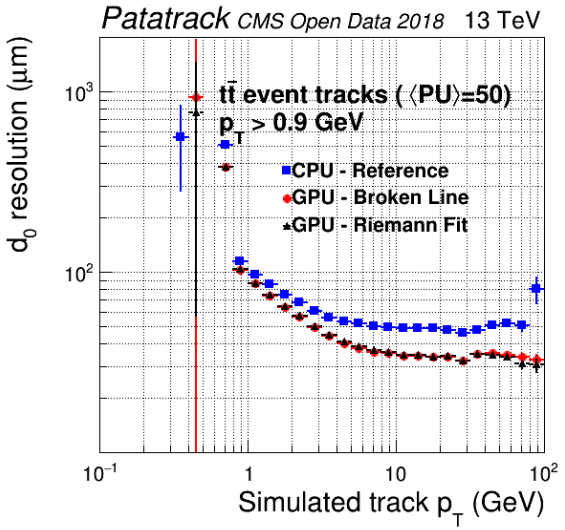
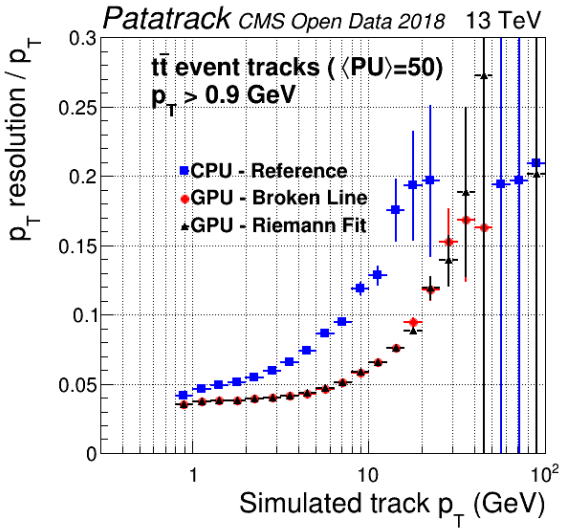


- If one node costs 7x 🍌
- The HLT filter farm farm would cost 7000 x 🍌
- HLT farm not running Pixel Tracking (on 10% of the events) would cost 700 x🍌 less
- A NVIDIA T4 costs 1.5x 🍌
- To run Patatrack Pixel Reconstruction on 100kHz we need 154 T4, that would cost 231x 🍌
- We can eat the other 469x🍌 and enjoy a way better Pixel Reconstruction

Conclusion



- A GPU-based full reconstruction of the Pixel detector from RAW data decoding to Pixel Tracks and Vertices determination, ECAL and HCAL local reconstruction has been implemented
- This reconstruction is fully integrated in the CMS Software
- Conversion to the legacy data formats and the standard validation can be run on demand
- Can achieve better physics performance, faster computational performance at a lower cost with respect to the baseline solution
- Working ongoing to develop the entire Phase-2 High Granularity Calorimeter Reconstruction directly with Heterogeneous Programming techniques
- Portable code is key for long-term maintainability, testability and support for new accelerator devices
    - Ongoing study and comparisons of solutions in Patatrack for CMS reconstruction
    - Starting from a CUDA code makes life **much** easier

backup

Physics Performance