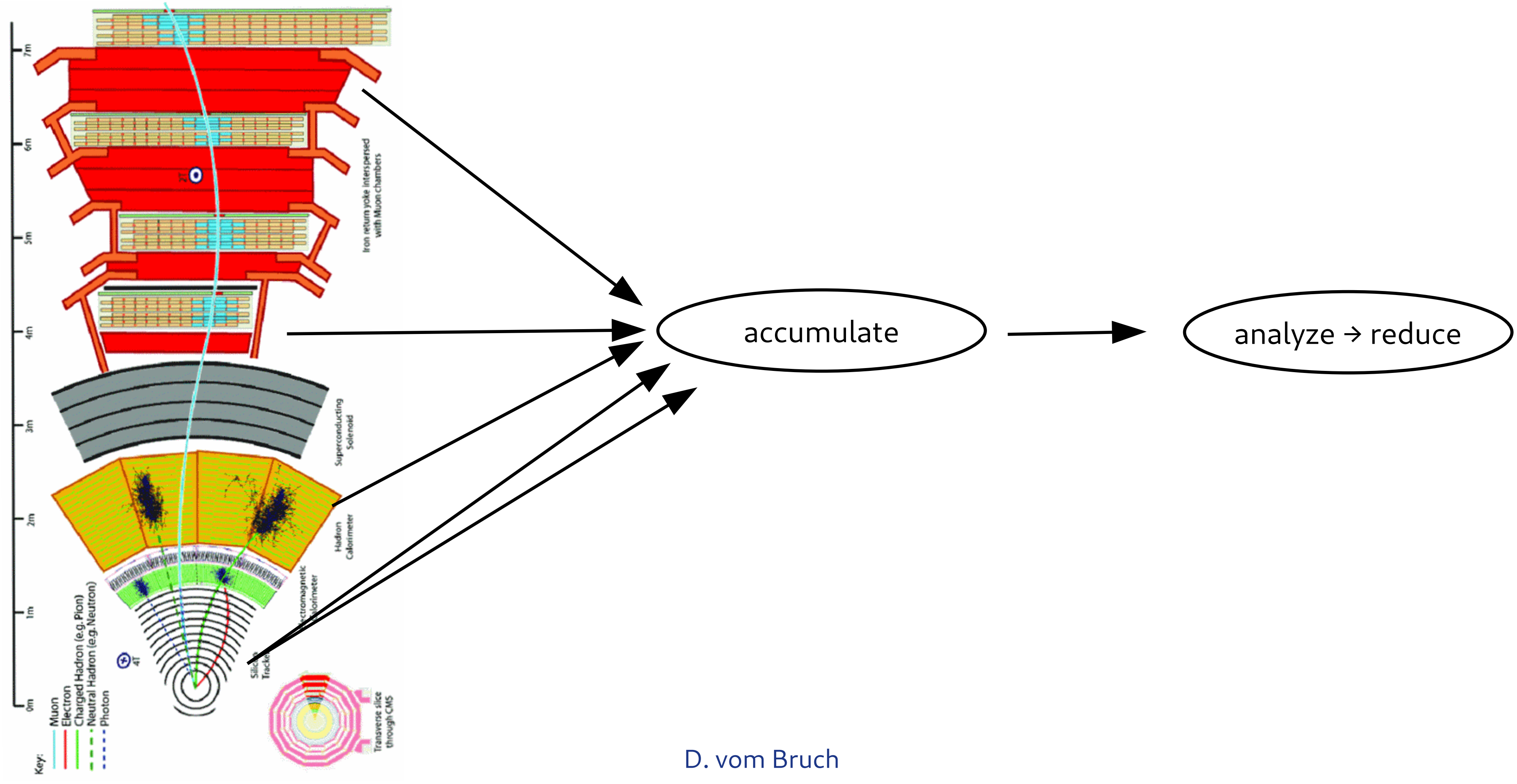
A heterogeneous software-only trigger for the upgraded LHCb experiment

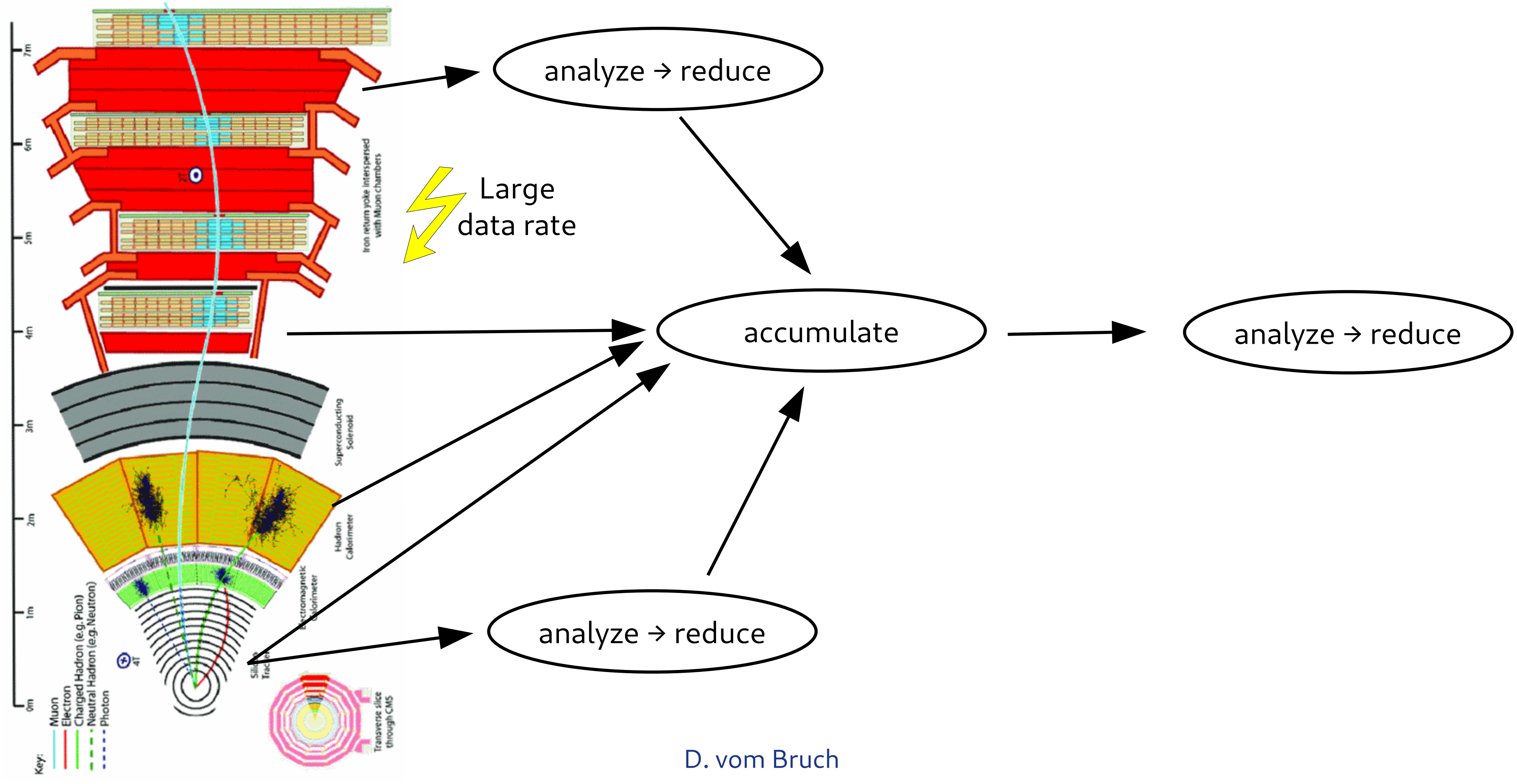
Dorothea vom Bruch
Center for Particle Physics Marseille (CPPM), Aix-Marseille University, IN2P3 / CNRS

September 12th 2022
Vistas on Detector Physics, Heidelberg

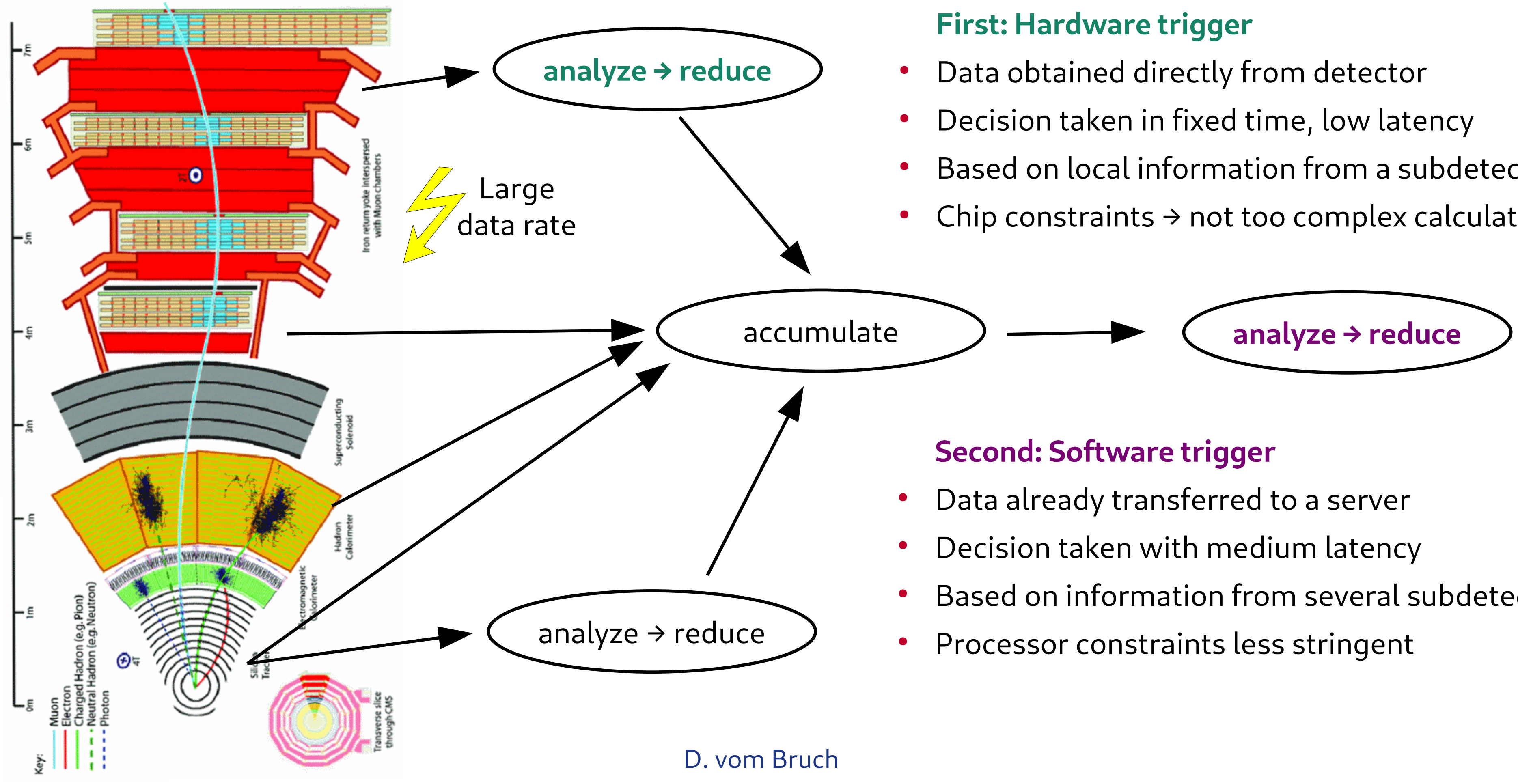
“Trigger”: Real-time data analysis and reduction



“Trigger”: Real-time data analysis and reduction



“Trigger”: Real-time data analysis and reduction



First: Hardware trigger

- Data obtained directly from detector
- Decision taken in fixed time, low latency
- Based on local information from a subdetector
- Chip constraints → not too complex calculations

Second: Software trigger

- Data already transferred to a server
- Decision taken with medium latency
- Based on information from several subdetectors
- Processor constraints less stringent

Match trigger to hardware

First: Hardware trigger

- Data obtained directly from detector
- Decision taken in fixed time, low latency
- Based on local information from a subdetector
- Chip constraints → not too complex calculations

Second: Software trigger

- Data already transferred to a server
- Decision taken with medium latency
- Based on information from several subdetectors
- Processor constraints less stringent

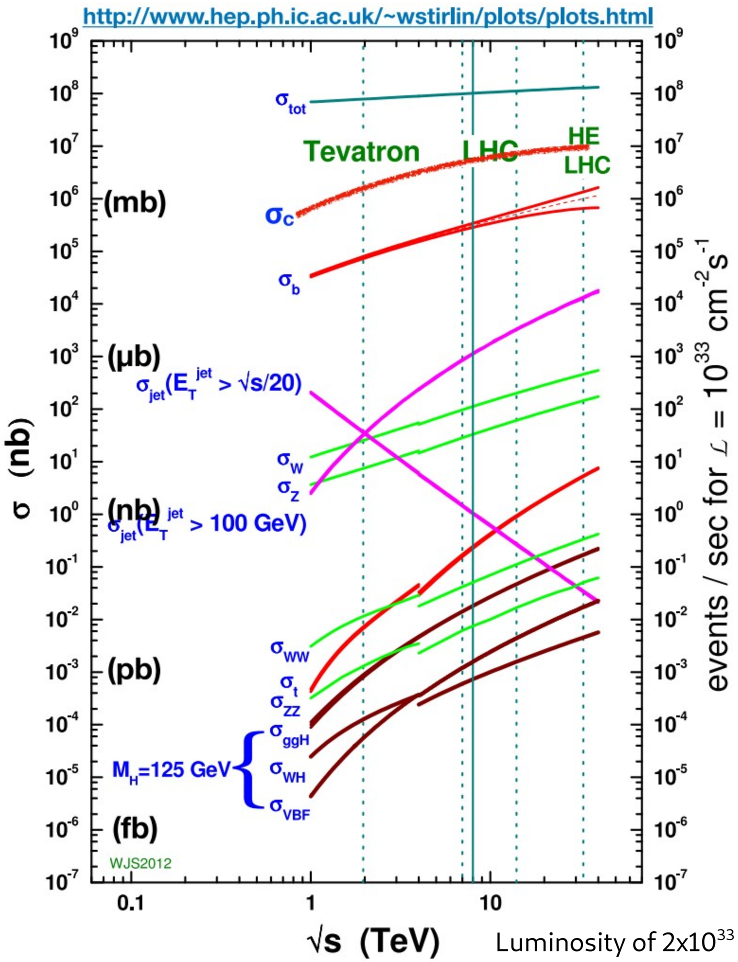
Field Programmable Gate Arrays (FPGAs)

- Low & deterministic latency
- Connectivity to any data source → high bandwidth
- Intermediate floating point performance

CPUs and GPUs

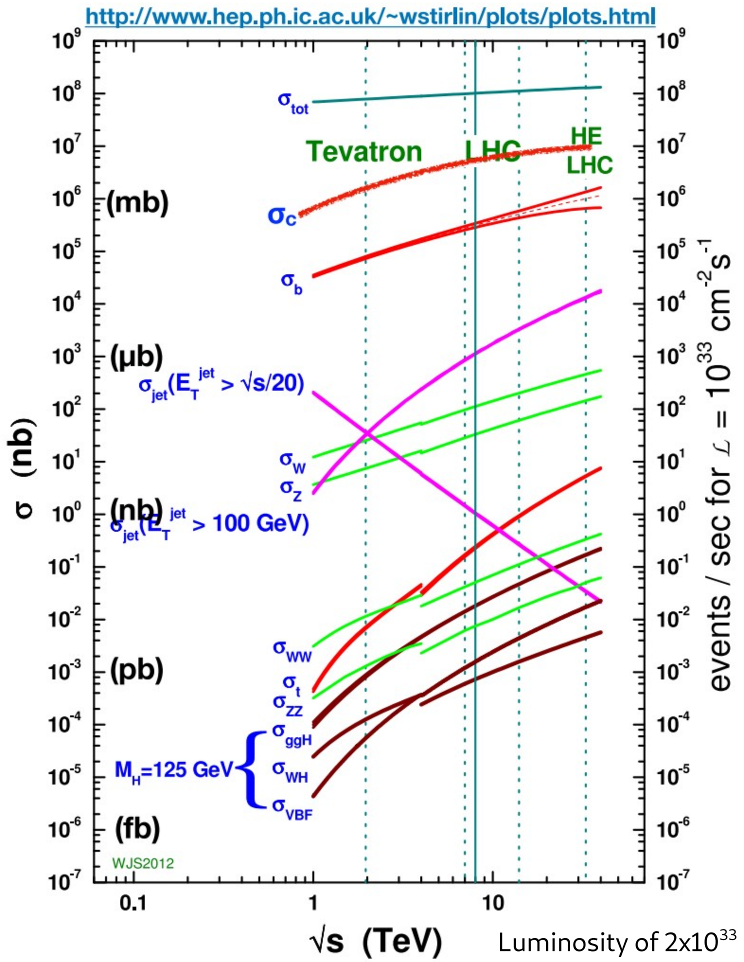
- Higher latency
- Very good floating point performance
- Connected to server (via PCIe connection for GPU)

Efficient signal selection



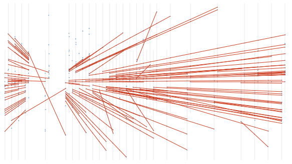
Luminosity of 2×10^{33} $cm^{-2}s^{-1}$

Efficient signal selection



Luminosity of $2\text{x}10^{33}\ cm^{-2}s^{-1}$

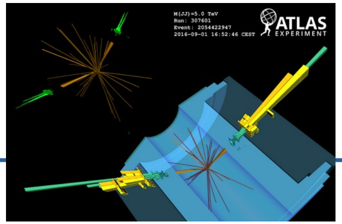
LHCb: Mainly beauty and charm physics



- Signal rates at MHz level
- Signal characteristics: Displaced vertices, momentum, particle type
- → No optimal local criteria for selection

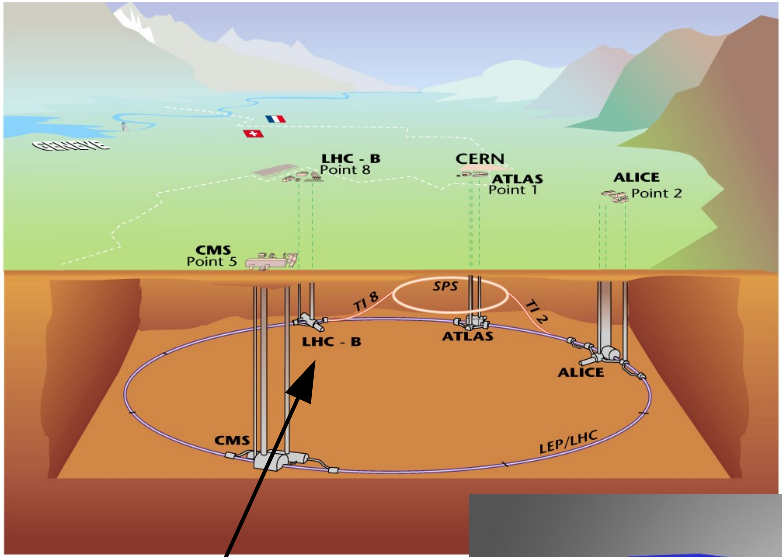
ATLAS & CMS: Mainly Higgs properties, high p_T new phenomena

- Signal rates up to hundreds of kHz
- Signal characteristics: high pT / transverse energy
- → Local criteria for selection possible

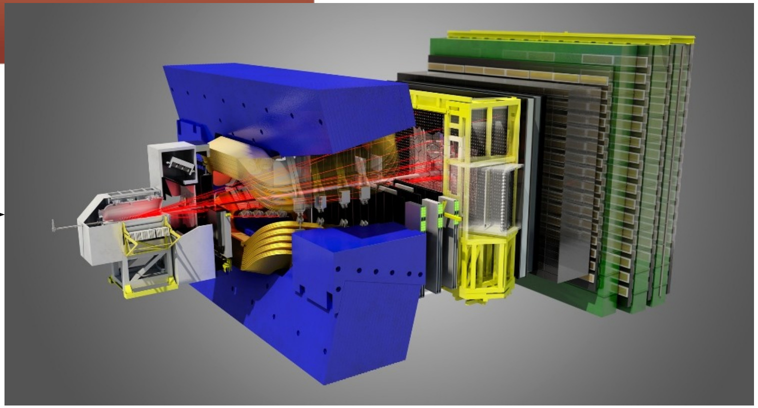


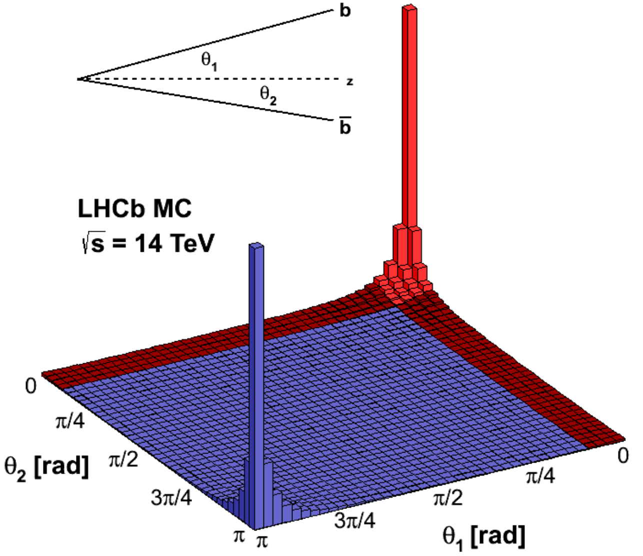
The LHCb experiment at CERN

LHC @ CERN



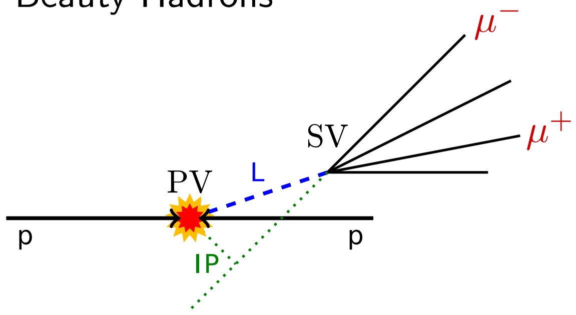
General purpose detector in the forward region specialized in beauty and charm physics





Beauty and charm decays

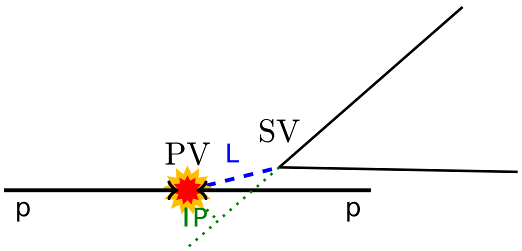
Beauty Hadrons



- $B^{\pm/0}$ mass ~5.3 GeV

  → Daughter p_T O(1 GeV)
- τ ~1.6 ps → flight distance ~1cm
- Detached muons from B→J/ΨX, J/Ψ → $\mu^+\mu^-$
- Displaced tracks with high p_T

Charm Hadrons

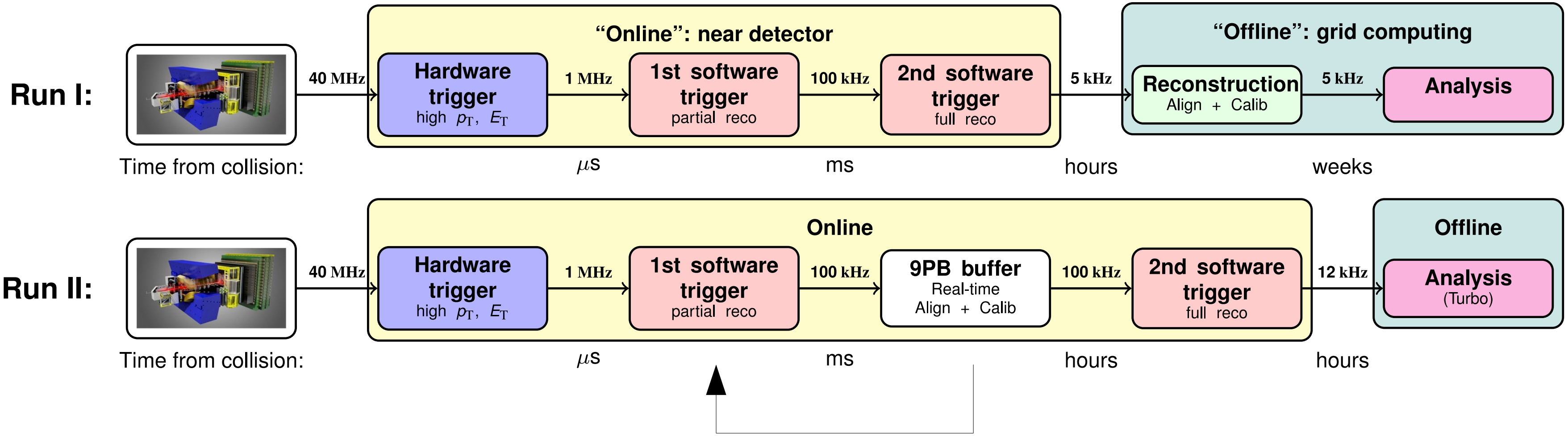


- $D^{\pm/0}$ mass ~1.9 GeV

  → Daughter p_T $\mathcal{O}$(700 MeV)
- τ ~0.4 ps → flight distance ~4mm
- Also produced from B decays

PV: Primary vertex
SV: Secondary vertex
IP: Impact parameter: distance between point of closest approach of a track and a PV

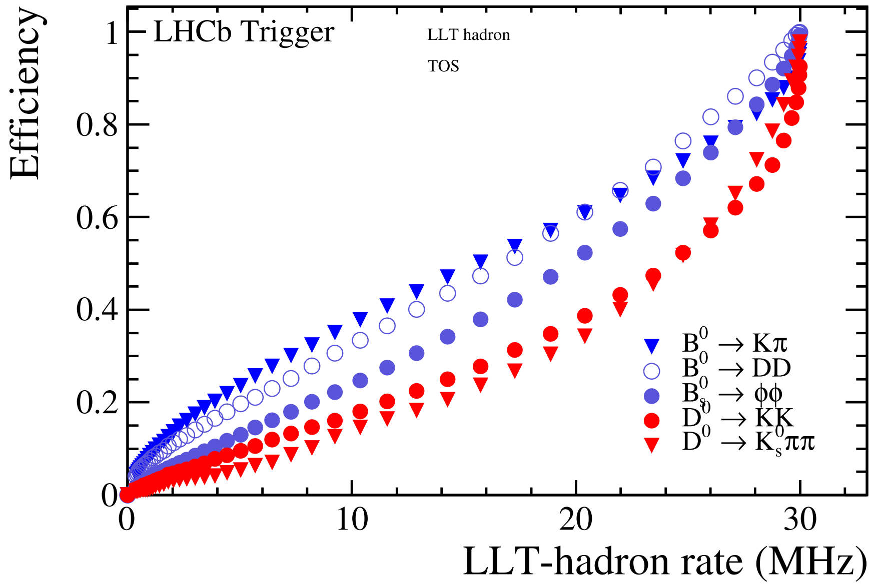
LHCb Run 1 & 2 trigger



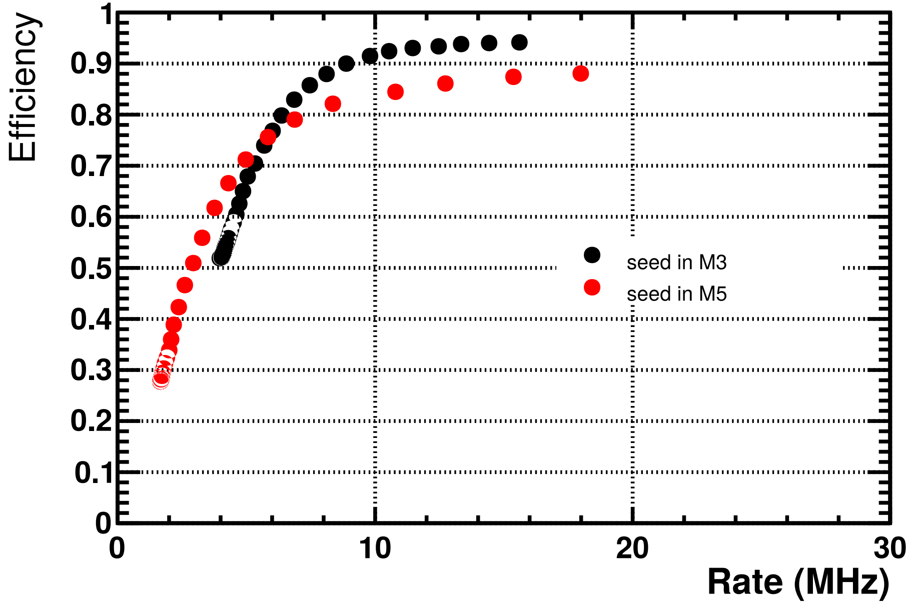
Update alignment & calibration once available

Why no low level trigger for LHCb in Run 3?

Low level trigger on E_T from the calorimeter



Low level trigger on muon p_T, $B \to K^*\mu\mu$

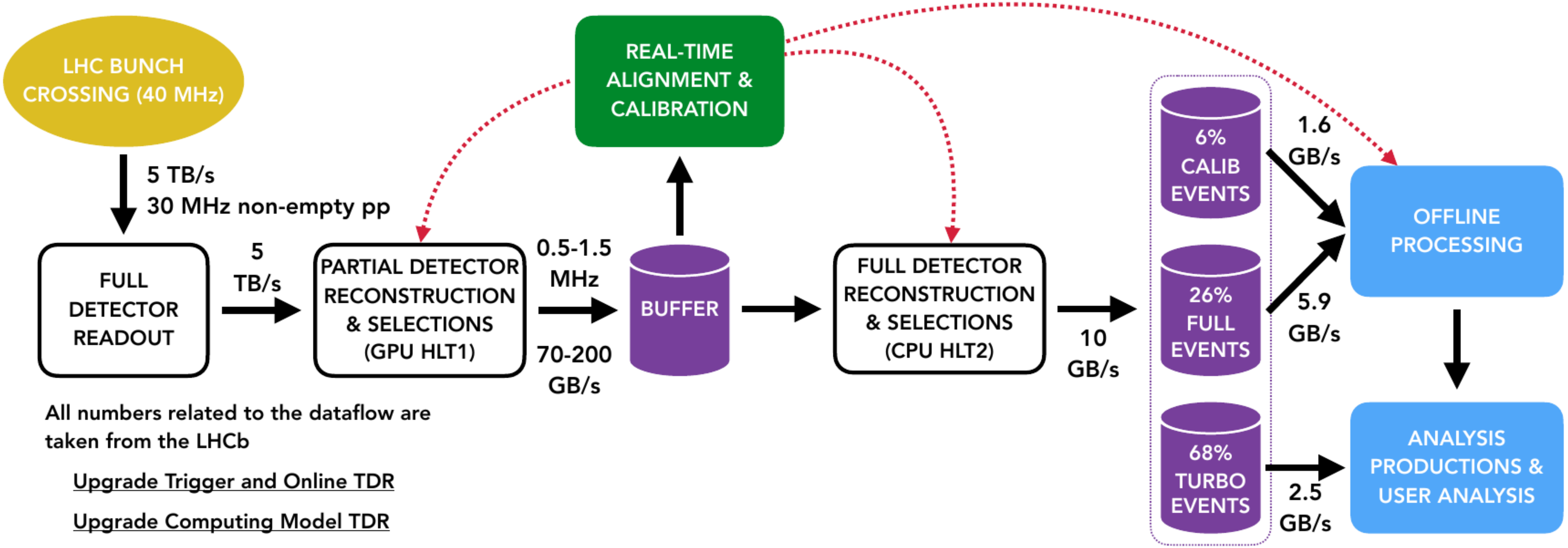


Need track reconstruction at first trigger stage

Change in trigger paradigm

Access as much information about the collision as early as possible

LHCb data processing in Run 3

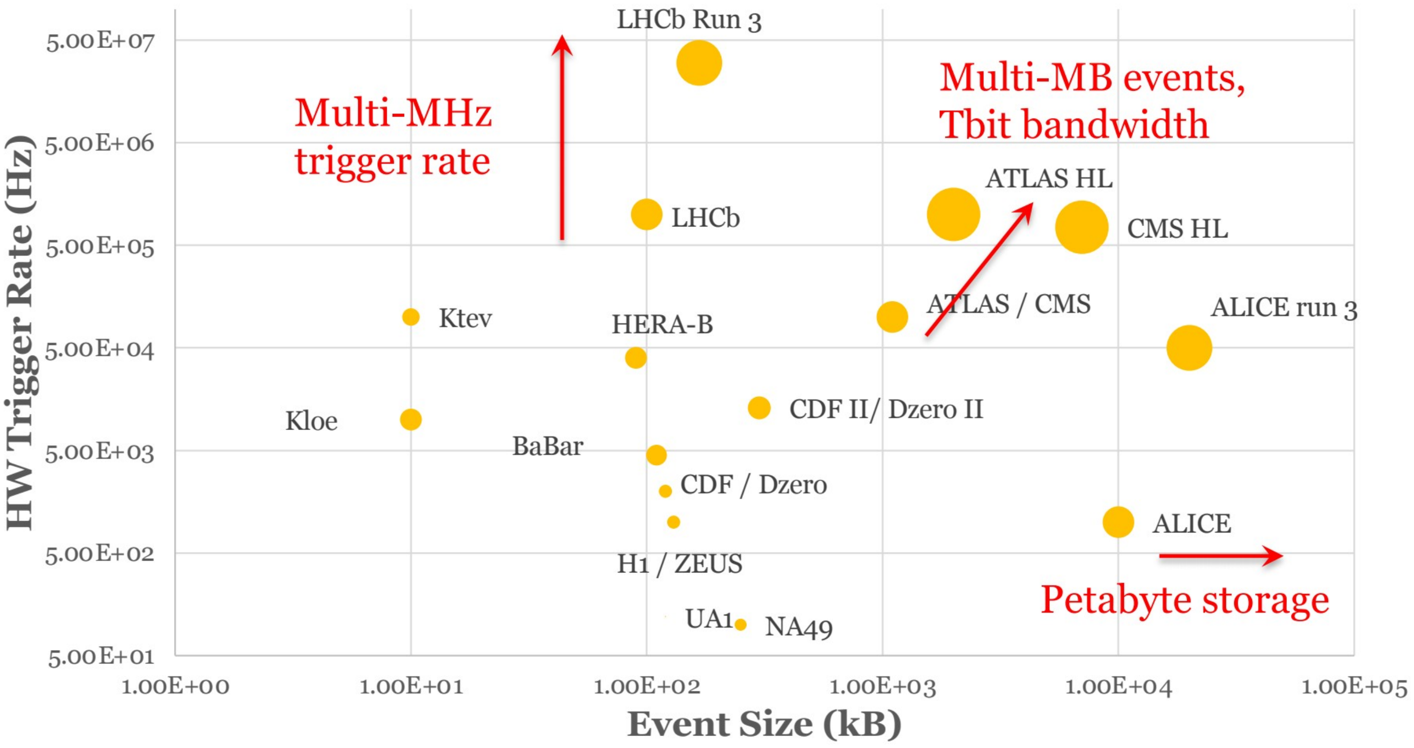


All numbers related to the dataflow are taken from the LHCb

Upgrade Trigger and Online TDR

Upgrade Computing Model TDR

Real-time software challenges in HEP



Courtesy Alex Cerri, LHCP 2022

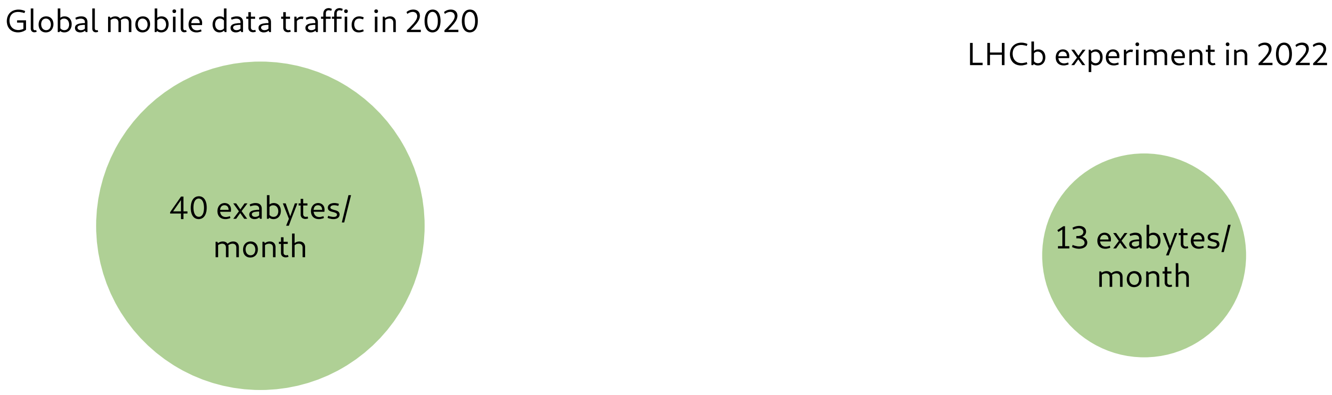
LHC Run 3 (2022)
LHCb: pp collisions at 30 MHz,
→ 5 TB/s processed in software

ALICE: PbPb collisions at 50 kHz
→ 3.5 TB/s processed in software

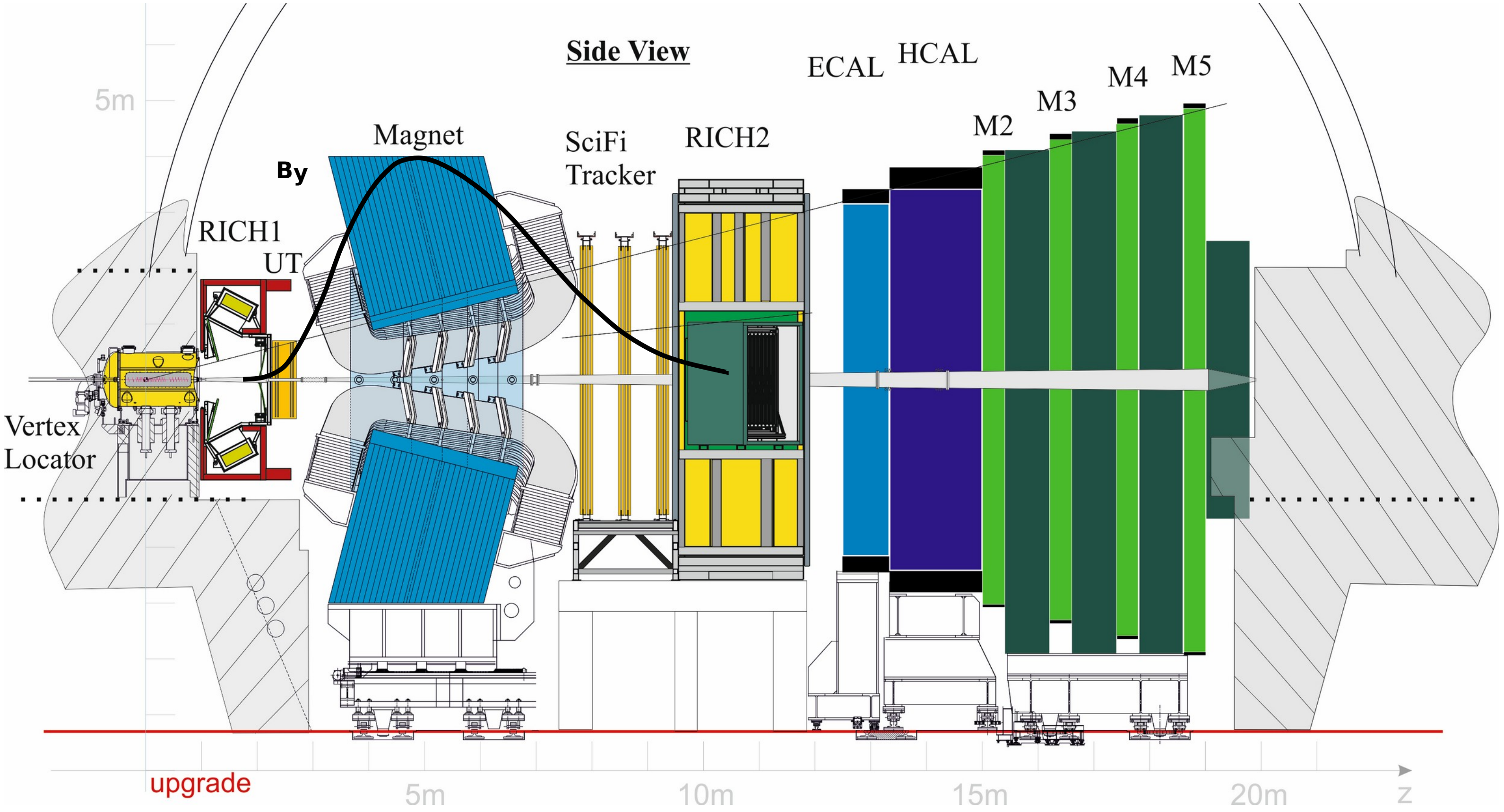
LHC Run 4 (~2029)
CMS & ATLAS
pp collisions at 40 MHz,
Hardware trigger rate increased:
100 kHz → 1 MHz
→ 6 TB/s processed in software

LHC Run 5 (~2035)
LHCb undergoes Upgrade II
25 TB/s processed in software

… in the global context

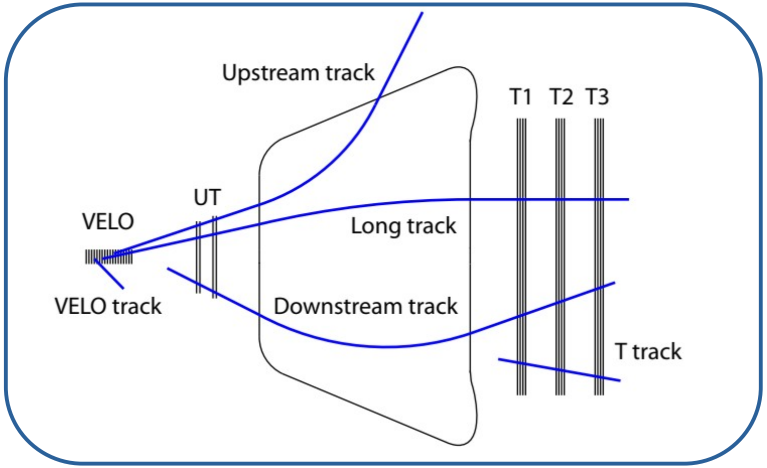


A closer look at LHCb

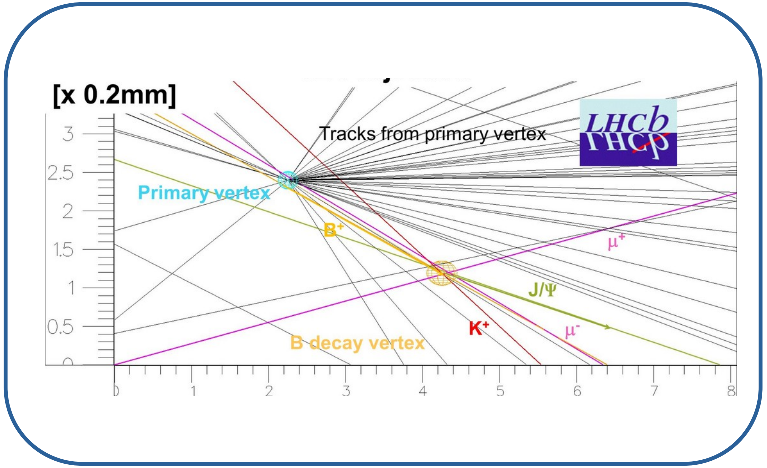


What do we reconstruct at LHCb?

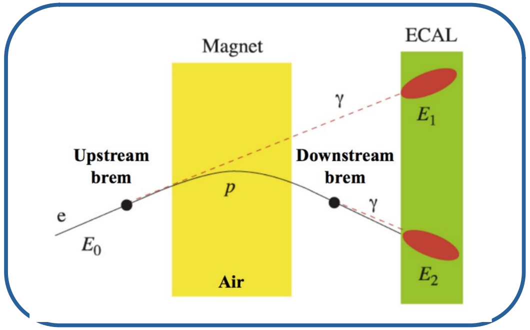
Tracks



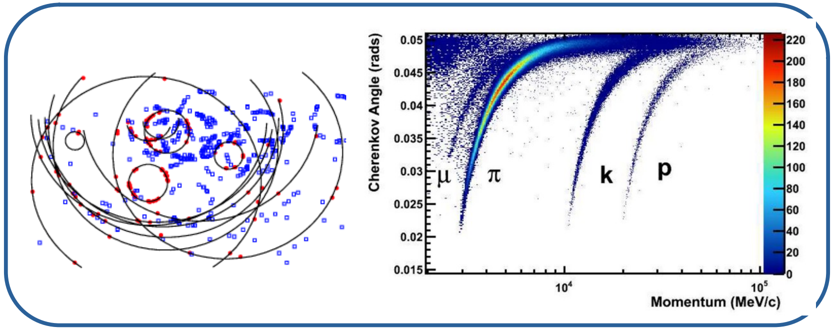
Vertices



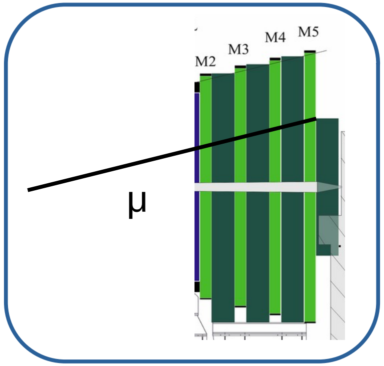
Electrons



Cherenkov rings

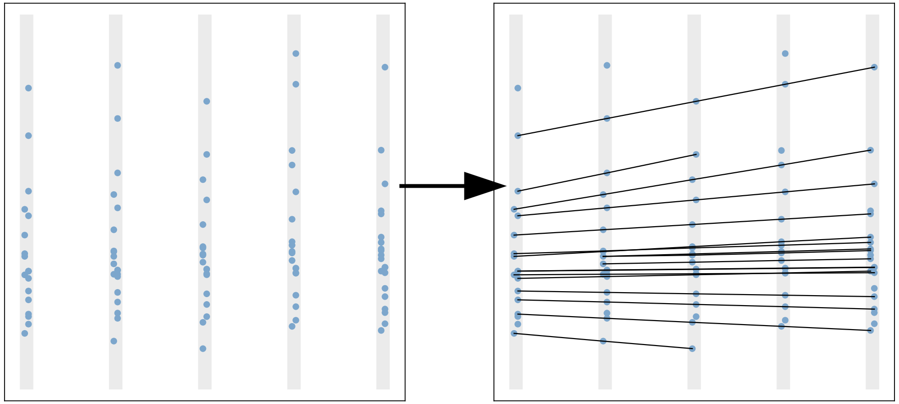


Muons

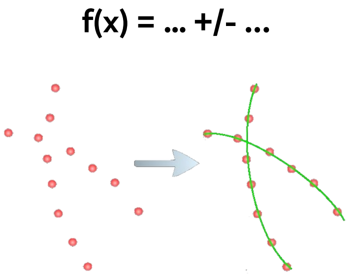


What does track reconstruction imply?

Pattern recognition

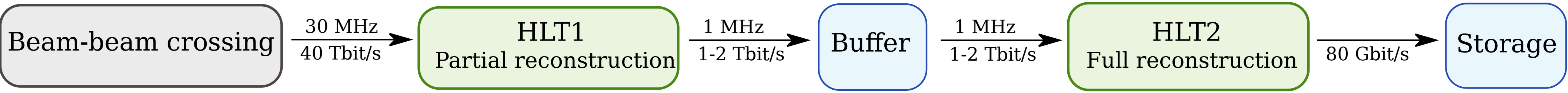


Track fit



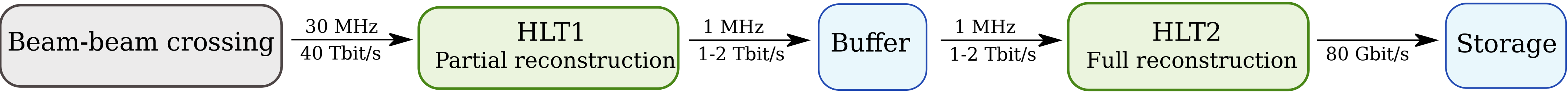
Huge computing challenge for 10^9 – 10^{10} tracks / second

Two stages of High Level Trigger (HLT)



- **High Level Trigger 1 (HLT1):**
  - Full charged particle track and vertex reconstruction
  - Electron and muon identification
  - Few inclusive single and two-track selections
- **High Level Trigger 2 (HLT2):**
  - Aligned and calibrated detector
  - Offline-quality pattern recognition
  - Full particle identification, including RICH reconstruction
  - Full track fit, requires detailed magnetic field and detector description

Two stages of High Level Trigger (HLT)

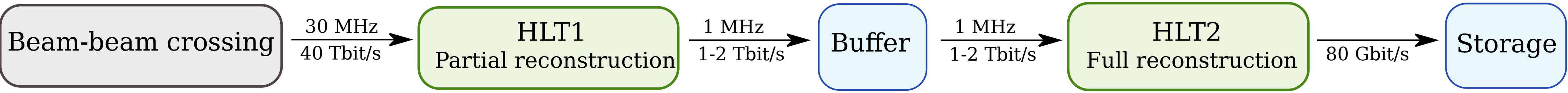


- **High Level Trigger 1 (HLT1):**
  - Full charged particle track and vertex reconstruction
  - Electron and muon identification
  - Few inclusive single and two-track selections
- **High Level Trigger 2 (HLT2):**
  - Aligned and calibrated detector
  - Offline-quality pattern recognition
  - Full particle identification, including RICH reconstruction
  - Full track fit, requires detailed magnetic field and detector description

- Manageable amount of algorithms
- Highly parallel tasks
- No detailed knowledge of magnetic field & detector required

- Exclusive selections using full PID information
- Best knowledge of alignment & calibration
- Reconstruction algorithms optimized for different track types
- Full track fit

Two stages of High Level Trigger (HLT)



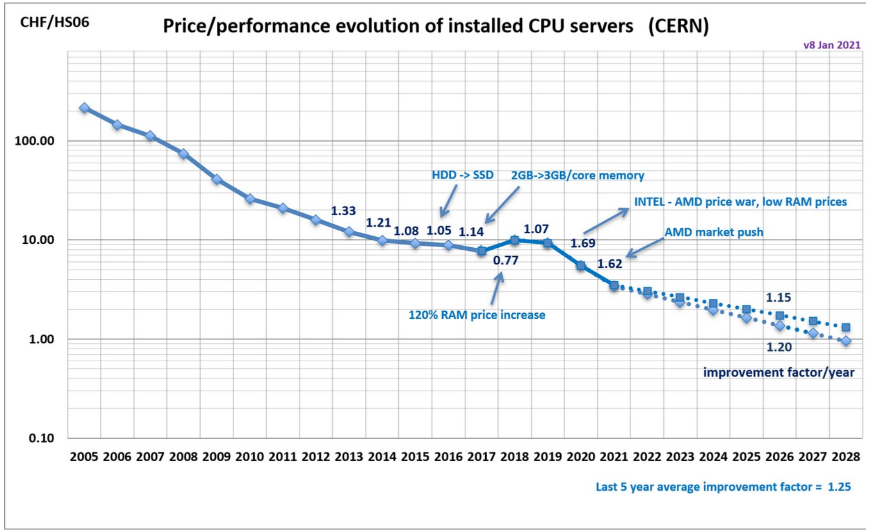
- **High Level Trigger 1 (HLT1):**
  - Full charged particle track and vertex reconstruction
  - Electron and muon identification
  - Few inclusive single and two-track selecti
- **High Level Trigger 2 (HLT2):**
  - Aligned and calib
  - Offline-
  - Full partic g RICH
    reconstruc
  - Full track fit, equires detailed magnetic field and detector description

- Ma
- etic field & detector required

Huge computing challenge

- Exclusive selections using full PID information
- Best knowledge of alignment & calibration
- Reconstruction algorithms optimized for different track types
- Full track fit

Computing performance challenge @ CERN



Courtesy Dr. Bernd Panzer-Steindel (CERN/IT, CTO)

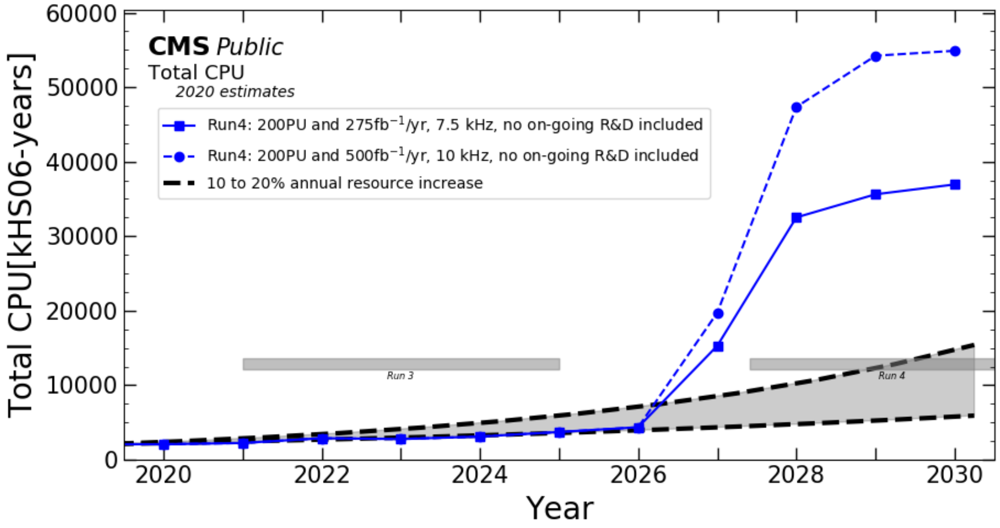


Image source

- Estimated improvement increase: 10-15% per year for the same budget
- Computing needs are not met

Trend towards heterogeneous solutions: TOP500

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, **Fujitsu** RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, **IBM** DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, **IBM / NVIDIA / Mellanox** DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, **NRCPC** National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	**Perlmutter** - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, **HPE** DOE/SC/LBNL/NERSC United States	761,856	70,870.0	93,750.0	2,589
6	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, **Nvidia** NVIDIA Corporation United States	555,520	63,460.0	79,215.0	2,646
7	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, **NUDT** National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
8	**JUWELS Booster Module** - Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, **Atos** Forschungszentrum Juelich (FZJ) Germany	449,280	44,120.0	70,980.0	1,764
9	**HPC5** - PowerEdge C4140, Xeon Gold 6252 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR Infiniband, **DELL EMC** Eni S.p.A. Italy	669,760	35,450.0	51,720.8	2,252
10	**Voyager-EUS2** - ND96amsr_A100_v4, AMD EPYC 7V12 48C 2.45GHz, NVIDIA A100 80GB, Mellanox HDR Infiniband, **Microsoft Azure** Azure East US 2 United States	253,440	30,050.0	39,531.2	

ARM processors

IBM Power9

Manycore processor: Sunway / NUDT

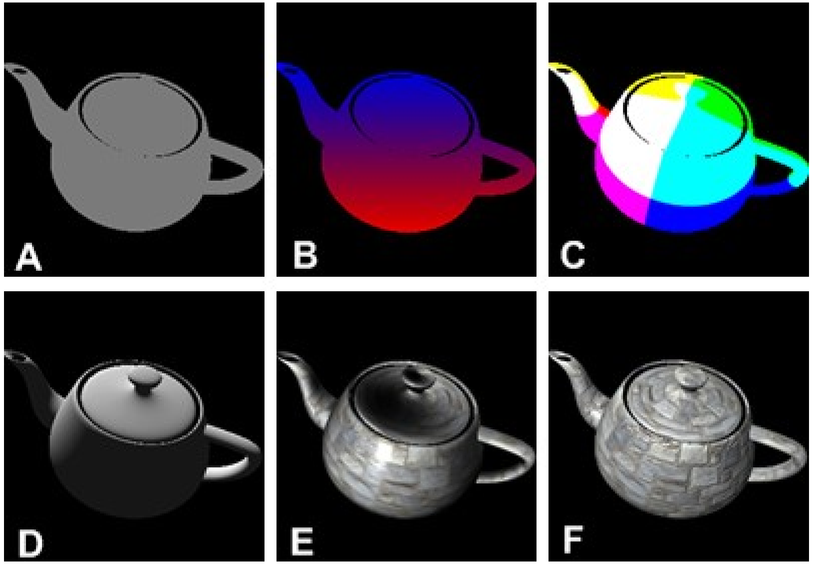
Nvidia GPUs: In 7/10 HPC centers

AMD & Intel CPUs

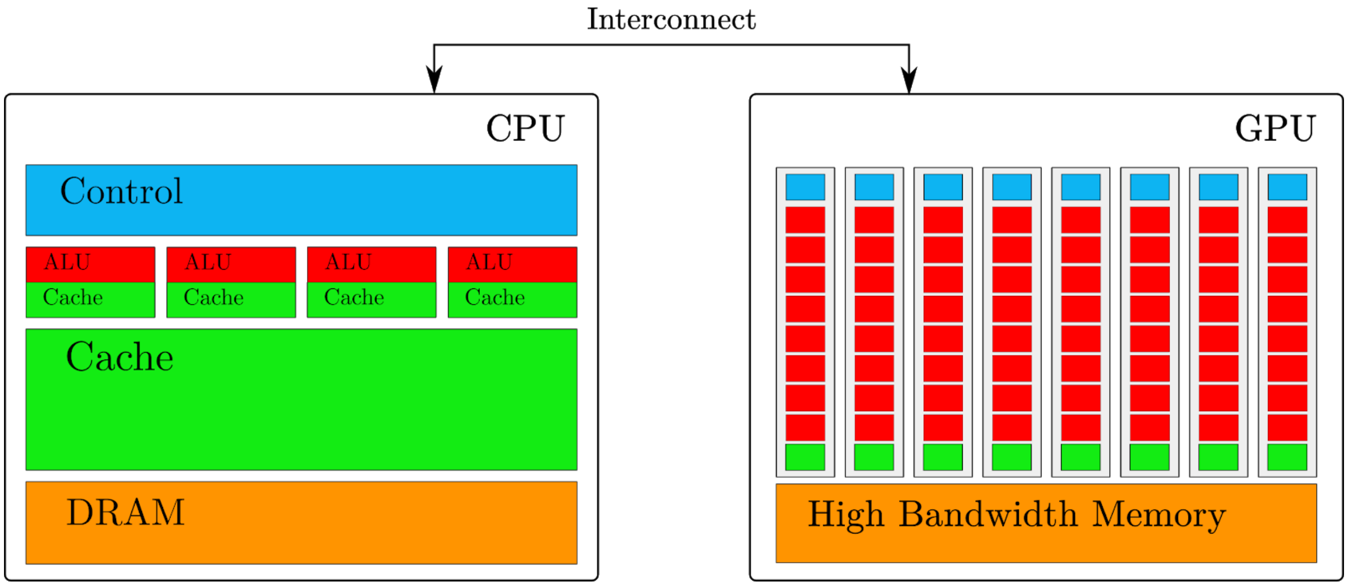
https://www.top500.org/lists/top500/2021/11/

Graphics Processing Unit (GPU)

Developed for graphics-oriented workloads



GPU compared to CPU



Low core count / powerful ALU

Complex control unit

Large chaches

→ **Latency optimized**

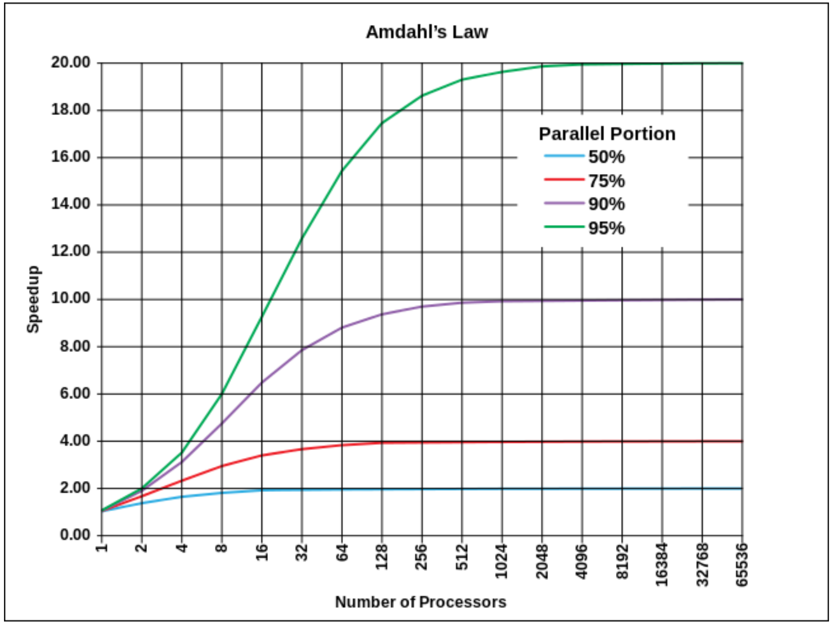
High core count

No complex control unit

Small chaches

→ **Throughput optimized**

When to go parallel? Amdahl's law



Speedup in latency = 1 / (S + P/N)

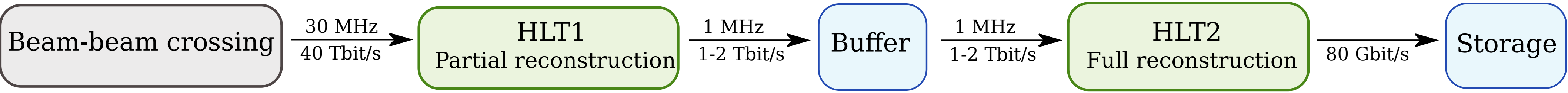
- S: sequential part of program
- P: parallel part of program
- N: number of processors

Parallel

Sequential

Consider how much of the problem can actually be parallelized!

Two stages of High Level Trigger (HLT)

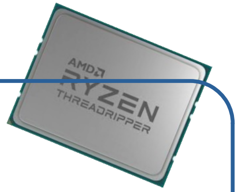


- **High Level Trigger 1 (HLT1):**
  - Full charged particle track and vertex reconstruction
  - Electron and muon identification
  - Few inclusive single and two-track selections
- **High Level Trigger 2 (HLT2):**
  - Aligned and calibrated detector
  - Offline-quality pattern recognition
  - Full particle identification, including RICH reconstruction
  - Full track fit, requires detailed magnetic field and detector description

Graphics Processing Units (GPUs)

- Manageable amount of algorithms
- Highly parallel tasks
- No detailed knowledge of magnetic field & detector required

Central Processing Units (CPUs)



- Exclusive selections using full PID information
- Best knowledge of alignment & calibration
- Reconstruction algorithms optimized for different track types
- Full track fit

How does HLT1 map to GPUs?

Characteristics of LHCb HLT1	**Characteristics of GPUs**
Intrinsically parallel problem: - Run events in parallel - Reconstruct tracks in parallel	Good for - Data-intensive parallelizable applications - High throughput applications
Huge compute load	Many TFLOPS
Full data stream from all detectors is read out → no stringent latency requirements	Higher latency than CPUs, not as predictable as FPGAs
Small raw event data (~100 kB)	Connection via PCIe → limited I/O bandwidth
Small event raw data (~100 kB)	Thousands of events fit into O(10) GB of memory

Perfect fit!

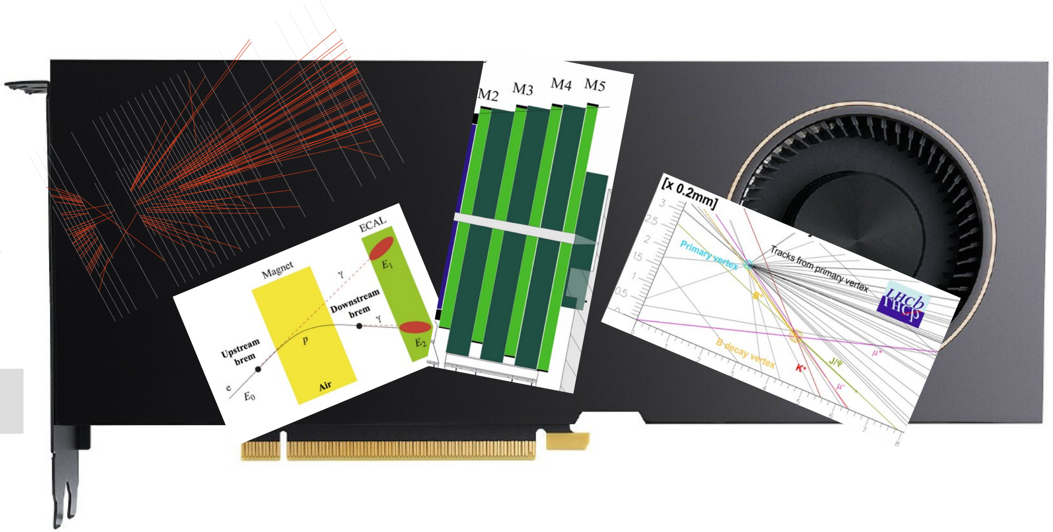
Minimize copies to / from GPU

Server

Raw data

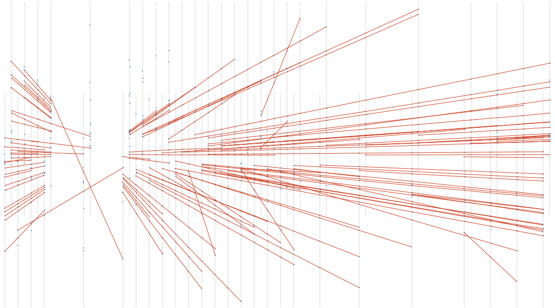
Selection decisions

GPU

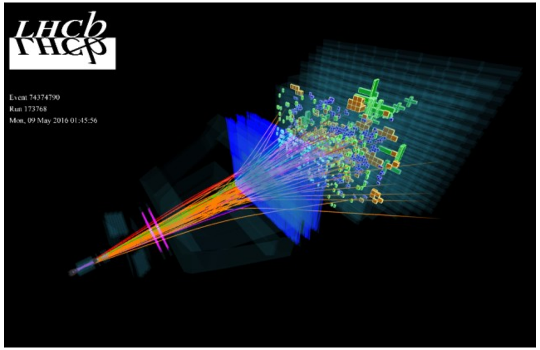


Three levels of parallelization

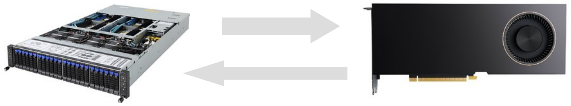
Intra-event: Tracks, vertices, …



Events



Event batches



The Allen software project

- Named after Frances E. Allen

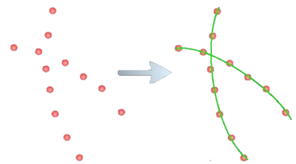
- Fully standalone software project: https://gitlab.cern.ch/lhcb/Allen, Sphinx documentation
- Framework developed for processing LHCb´s HLT1 on GPUs

- Cross-architecture compatibility via macros & few coding guide lines
    - GPU code written in CUDA, runs on CPUs, Nvidia GPUs (CUDA), AMD GPUs (HIP)
- Algorithm sequences defined in python and generated at run-time
- Multi-event processing with dedicated scheduler
- Memory manager allocates large chunk of GPU memory at start-up
- Reconstruction algorithms re-designed for parallelism and low memory usage: O(MB) per core

Common intra-event parallelization techniques

Raw data decoding

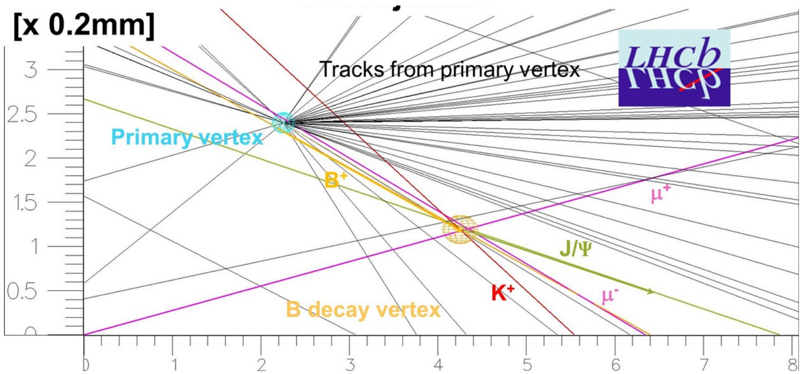
- Transform binary payload from subdetector raw banks into collections of hits (x,y,z) in global coordinate system
- Parallelize over all readout units



f(x) = … +/- …

Track reconstruction

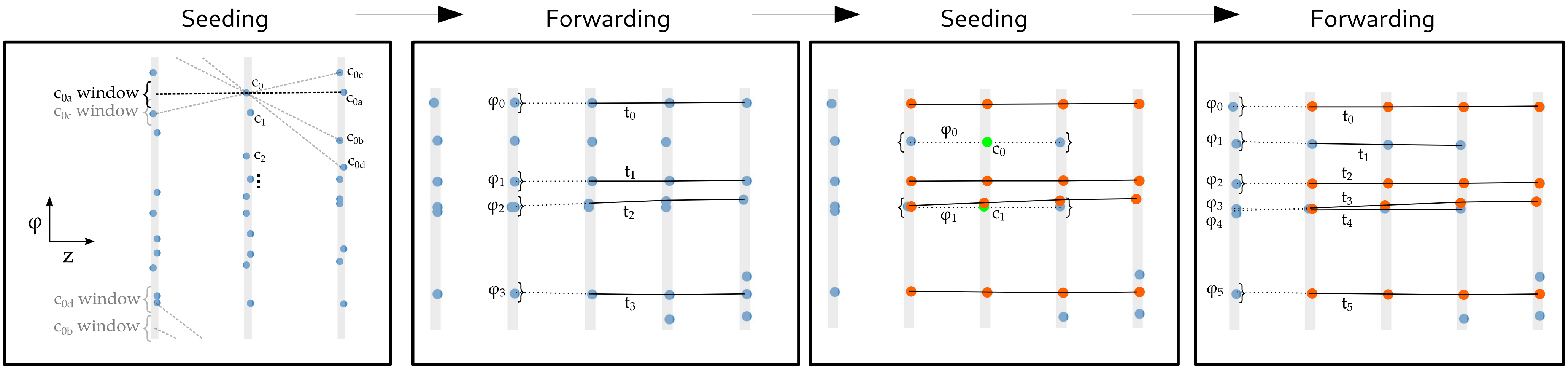
- Consists of two steps:
  - Pattern recognition: Parallelize across hit combinations
  - Track fitting: Parallelize across track candidates



Vertex finding

- Reconstruct primary and secondary vertices
- Parallelize across combinations of tracks and vertex seeds

Example algorithm: "Triplet" finder

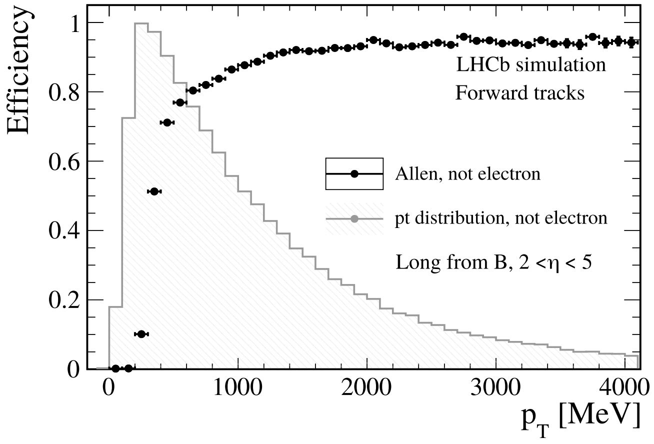


- Build "triplets" of three hits on consecutive layers → parallelization
- Choose them based on alignment in phi
- Hits sorted by phi → memory accesses as contiguous as possible: data locality
- Extend triplets to next layer → parallelization

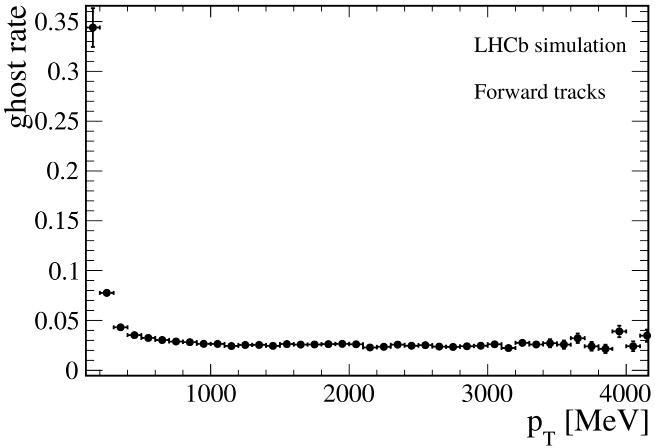
D. Campora, N. Neufeld, A. Riscos Núñez: "A fast local algorithm for track reconstruction on parallel architectures", IPDPSW 2019

HLT1: Track reconstruction performance

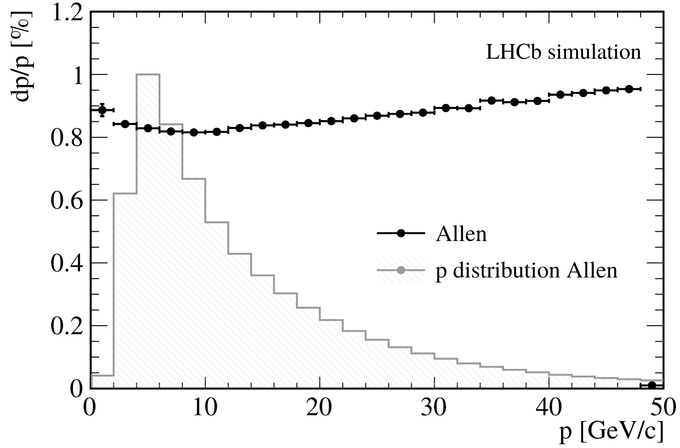
Track reconstruction efficiency



Fake rate

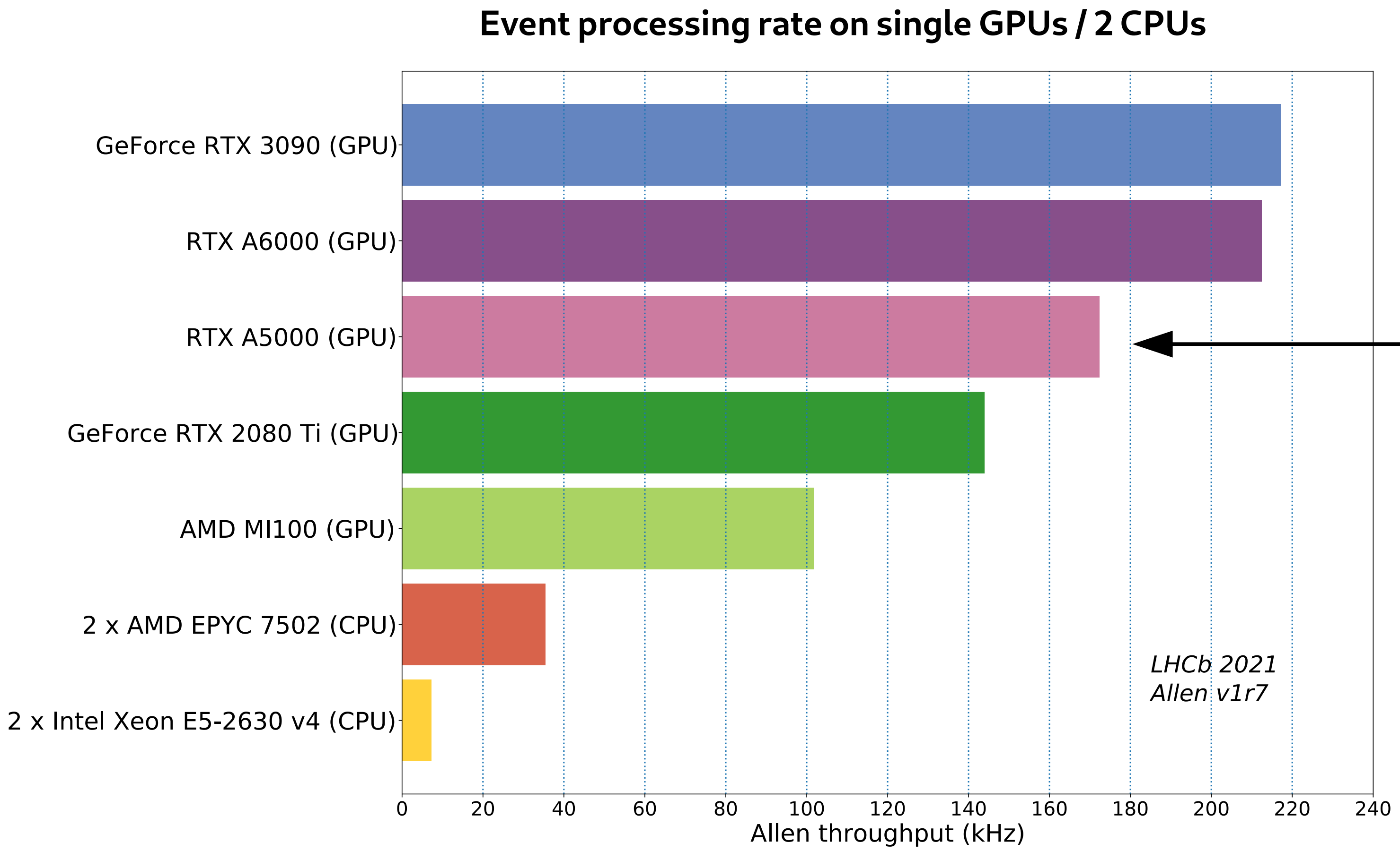


Momentum resolution



LHCb-FIGURE-2020-014

HLT1: Computing throughput

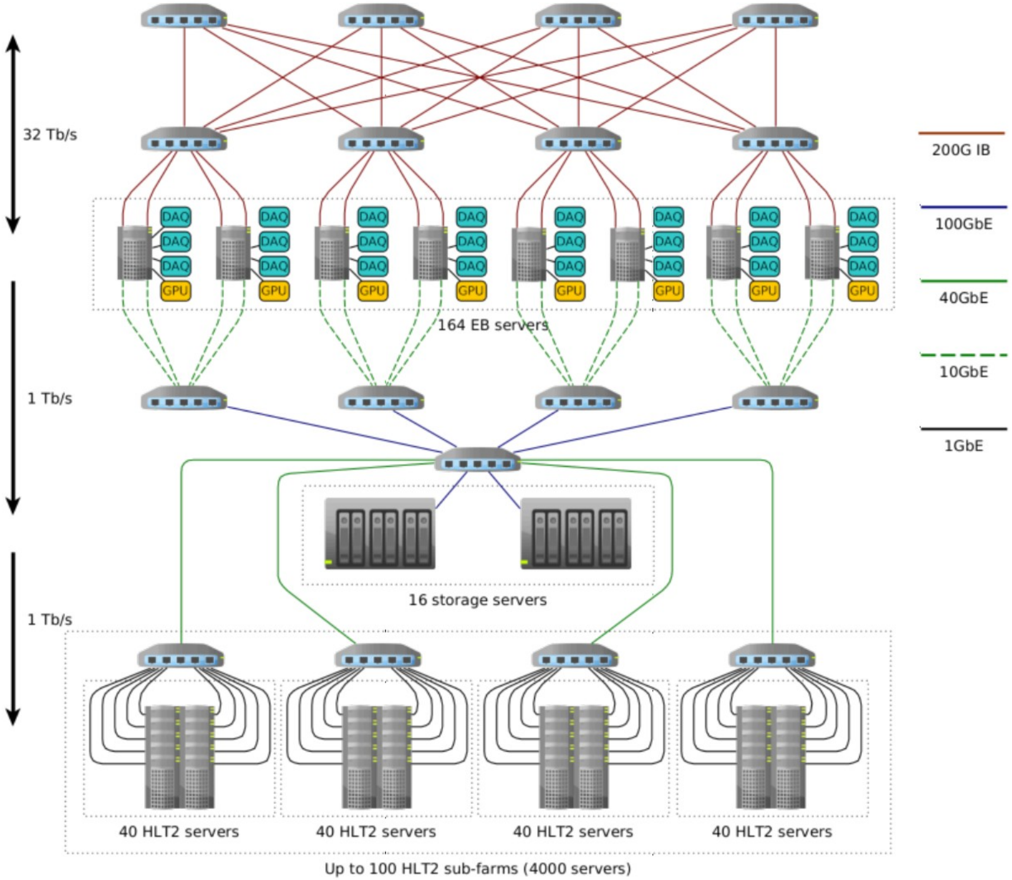


Chose RTX A5000 for 2022

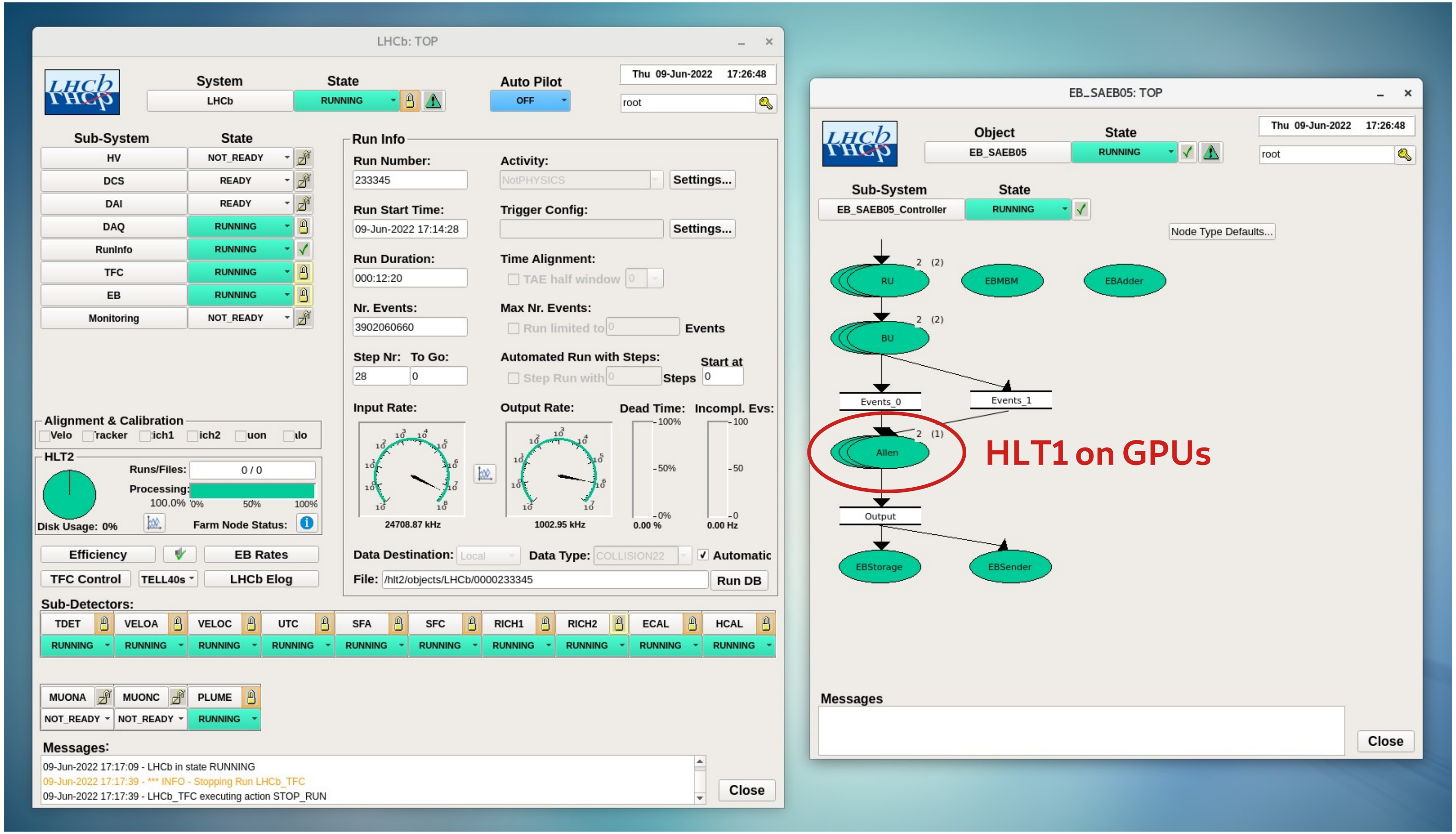
Need O(200) GPUs to process HLT1 at 30 MHz

Update of LHCb-FIGURE-2020-014

GPU HLT1 within data acquisition system

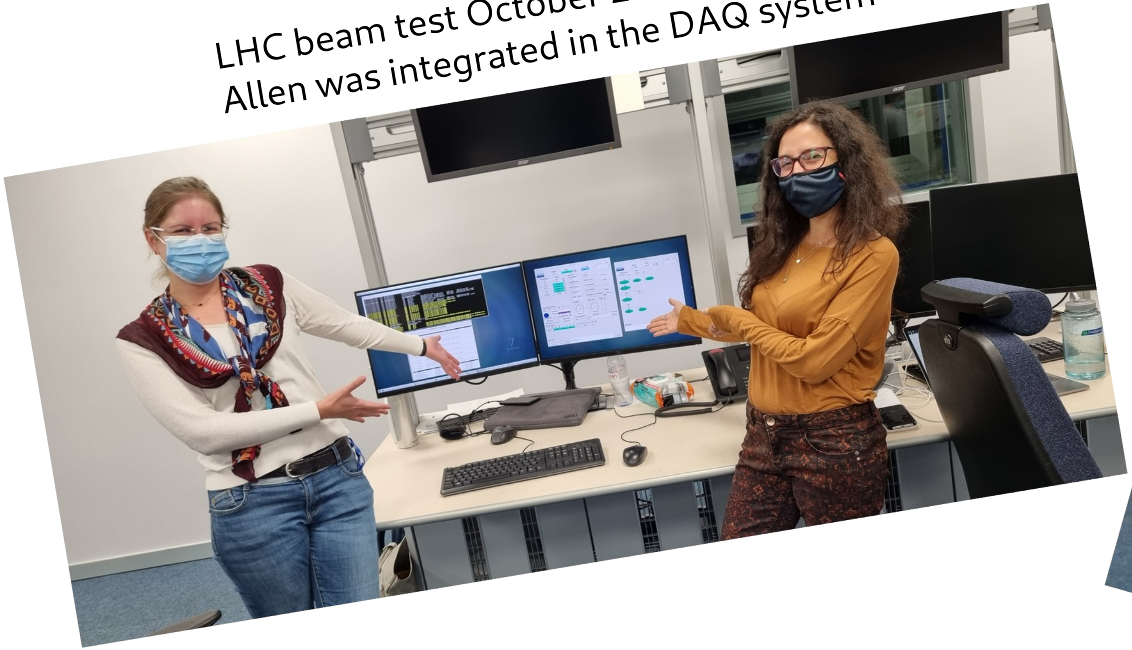


HLT1 commissioning: Allen within the DAQ system



HLT1 commissioning: Towards first collisions

LHC beam test October 2021: First time Allen was integrated in the DAQ system



May 2022: First time Allen ran at 25 MHz input rate

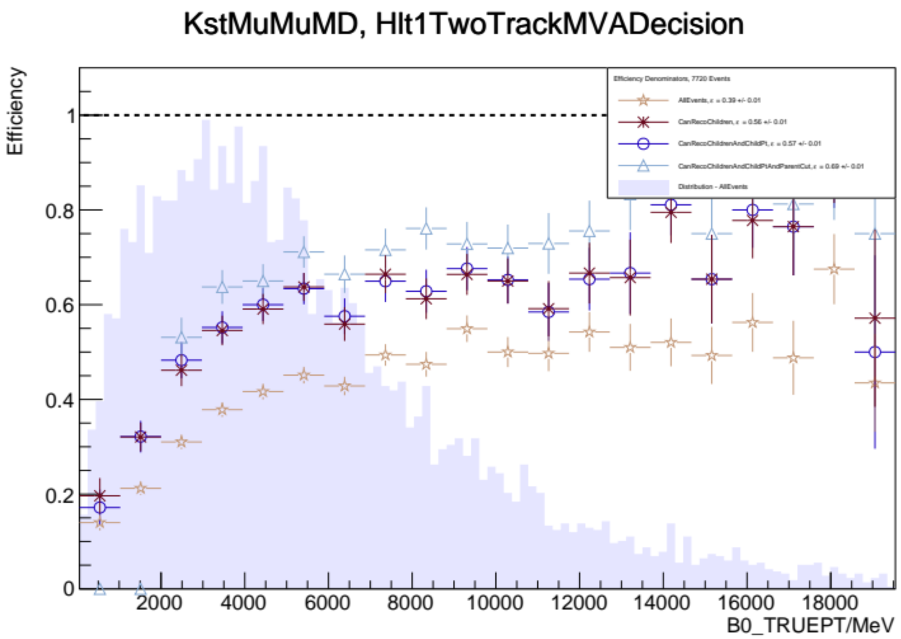
HLT1 commissioning: Towards first collisions

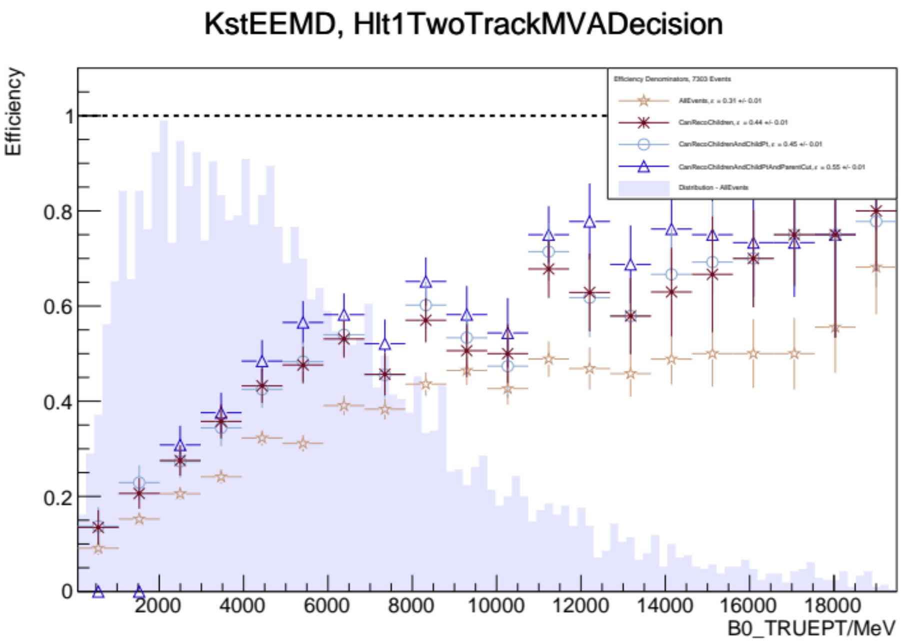
July 2022: First collisions @ 13.6 TeV at the LHC
Happy trigger commissioning team

LHC beam
Allen was i

Allen ran at 25 MHz input rate

Looking at the physics performance





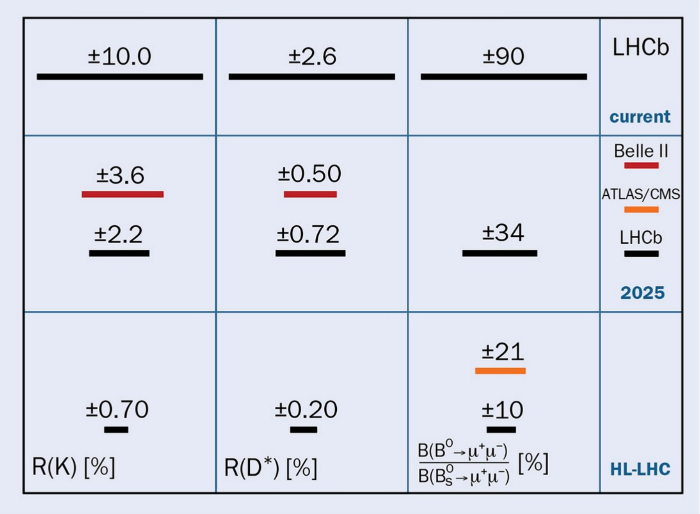
CERN-LHCC-2020-006

Selection efficiencies for electron and muon final states similar

In Run 2: Electron selection efficiency roughly factor two worse than muons due to hardware level trigger

Physics prospects with the all-software trigger

- Understand the current pattern of flavor anomalies
- Exploiting the higher statistics and larger phase space of electrons
- Precision measurements of rare decays with electrons: b → see, b → dee
  - Branching fractions, ratios of branching fractions to muon modes, angular analyses
- Semileptonic decays with electrons: b → ceν
  - Ratios of branching fractions to tauonic mode, angular analyses
- Exploit higher statistics at low momentum
  - Decays with multiple tracks in the final state
  - Charm decays
- Adding on to the trigger in the future
  - Reconstruct tracks of long-lived particles: K_s studies
  - Fill histograms directly in the trigger, for example for dark photon searches

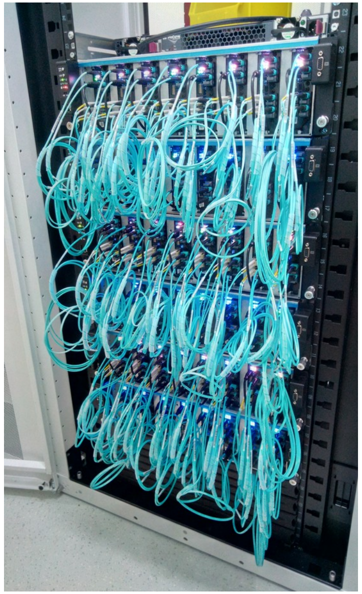


ArXiv 1808.08865

Summary

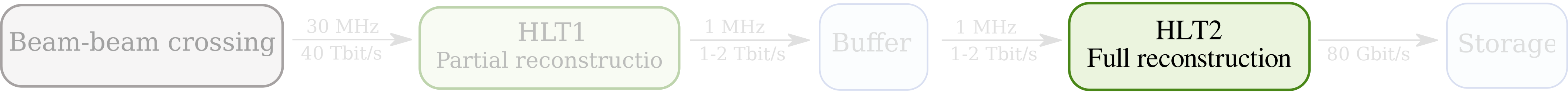
- HEP experiments real time analysis systems are entering the exascale computing era
- Need to exploit modern computing techonolgies to face this challenge
- LHCb is commissioning a fully software trigger for Run 3 (started in 2022)
- First full trigger stage entirely on GPUs @ 30 MHz → a first in HEP
- Developed Allen: heterogeneous software framework for multi-event processing
- Gain expertise in heterogeneous DAQ systems

  → Preparing to exploit emerging new architectures entering the market
- Physics performance opens new options for physics analyses
- In good position to prepare for LHCb Upgrade II with 400 Tbit/s of data rate

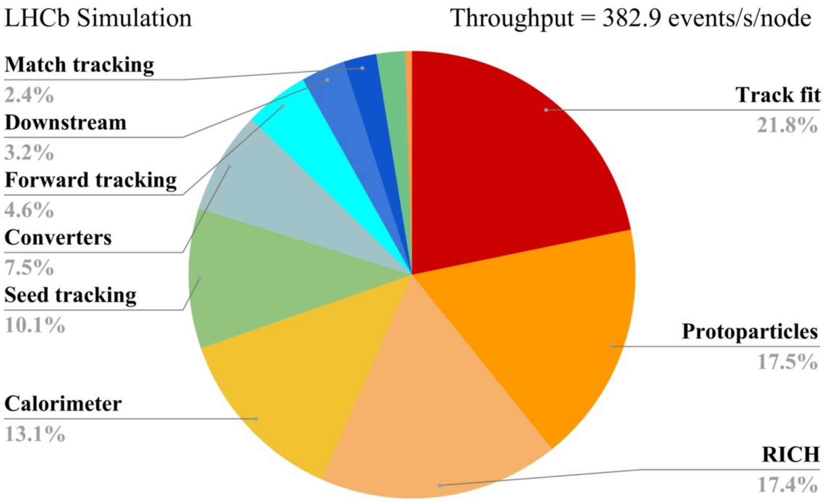


Backup

HLT2 on CPUs

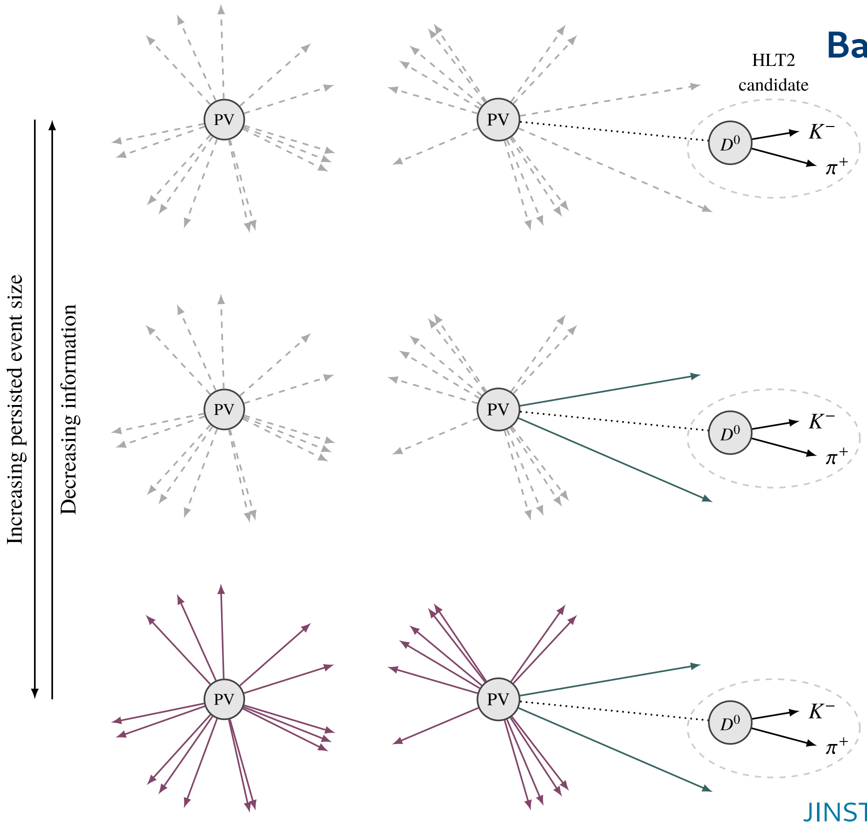


- Fully aligned & calibrated detector, offline quality track fit & particle identification @ 1MHz
- HLT2 throughput significantly improved over last years
- Hundreds of exclusive selections being written for specific analyses, using new multi-threaded framework



LHCb-FIGURE-2021-003

Selective persistency: "Turbo stream"



Bandwidth [MB/s] ~ Trigger output rate [kHz] x average event size [kB]

- Trigger bandwidth is crucial, not trigger rate
- Only store high-level objects reconstructed in real-time
- → Reduction of event size
- High degree of flexibility:
  - Only objects used in trigger selection
  - Objects used in trigger selection & user-defined selection
  - All reconstructed objects
- Raw data only stored in calibration stream

JINST 14 (2019) P04006

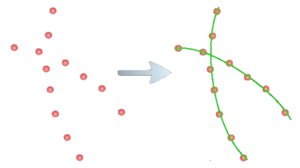
Recurrent tasks in real-time data analysis

Raw data decoding

- Transform binary payload from subdetector raw banks into collections of hits (x,y,z) in LHCb coordinate system

Track reconstruction

- Consists of two steps:
  - Pattern recognition: Which hits were produced by the same particle? → "Track"
  - Track fitting: Describe track with mathematical model



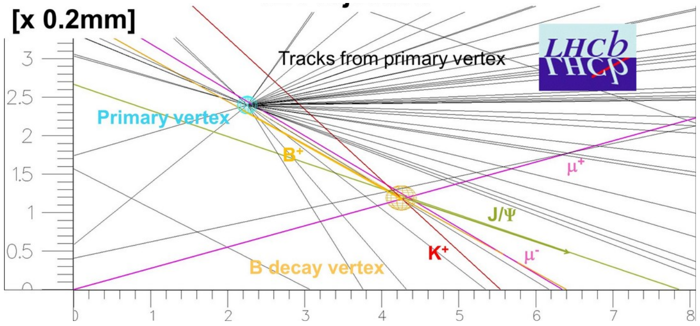
f(x) = … +/- …

Vertex finding

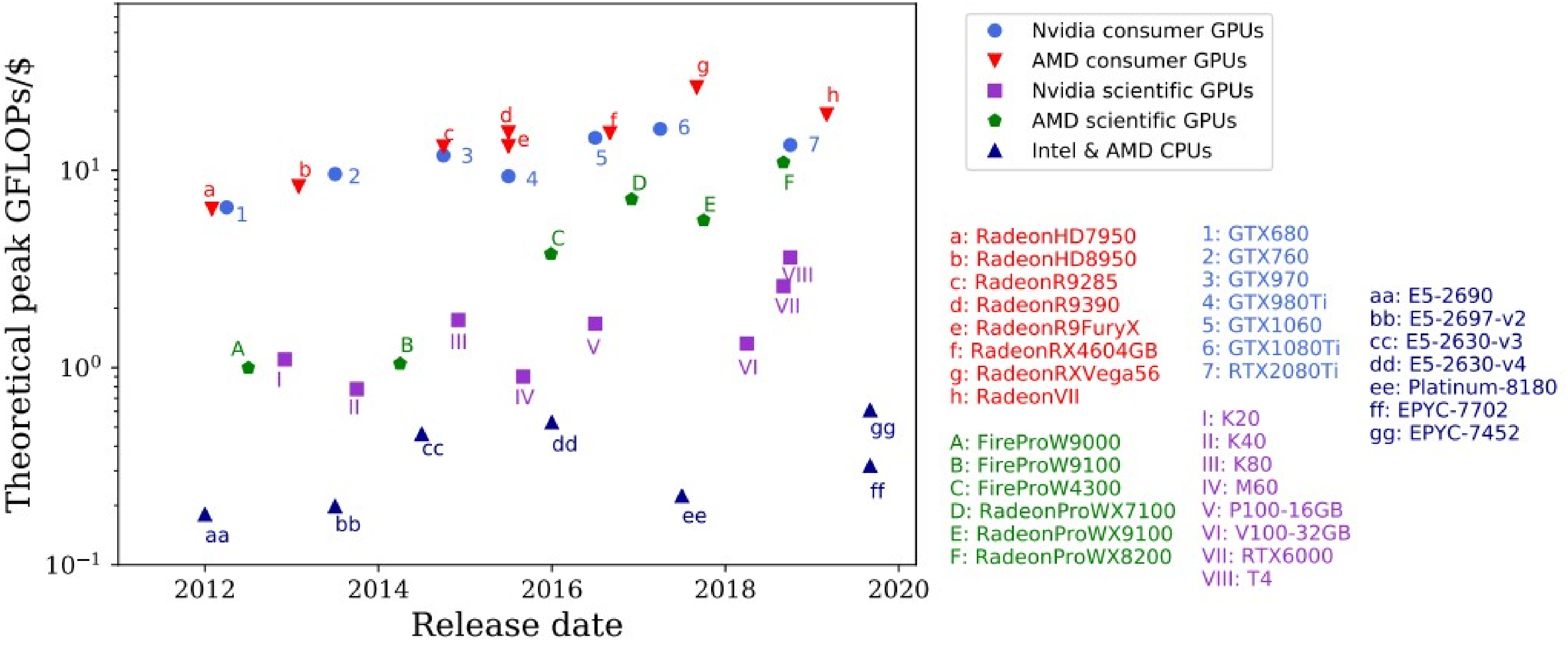
- Where did proton-proton collisions take place?
- Where did particles decay within the detector volume?

Particle identification

- Reconstruct clusters in the calorimeter / muon detectors
- Reconstruct rings in the RICH detectors
- Match tracks to clusters / RICH signals



What about the cost?



https://arxiv.org/pdf/2003.11491.pdf

Heterogeneous solutions & sustainability: Green500

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	301	**MN-3** - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Preferred Networks Japan	1,664	2,181.2	55	39.379
2	291	**SSC-21 Scalable Module** - Apollo 6500 Gen10 plus, AMD EPYC 7543 32C 2.8GHz, NVIDIA A100 80GB, Infiniband HDR200, HPE Samsung Electronics South Korea	16,704	2,274.1	103	33.983
3	295	**Tethys** - NVIDIA DGX A100 Liquid Cooled Prototype, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 80GB, Infiniband HDR, Nvidia NVIDIA Corporation United States	19,840	2,255.0	72	31.538
4	280	**Wilkes-3** - PowerEdge XE8545, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 80GB, Infiniband HDR200 dual rail, DELL EMC University of Cambridge United Kingdom	26,880	2,287.0	74	30.797
5	30	**HiPerGator AI** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Infiniband HDR, Nvidia University of Florida United States	138,880	17,200.0	583	29.521

- All top 5 Green500 use accelerators
- 4/5 use Nvidia GPUs combined with AMD Epyc
- MN-3 uses an accelerator optimized for matrix arithmetic
- Of the top 30 Green500:
  - 26 use Nvidia GPUs
  - 3 use A64FX vector-processors (ARM)
  - 1 uses a many-core microprocessor (PEZY-SC3)

https://www.top500.org/lists/green500/2021/11/

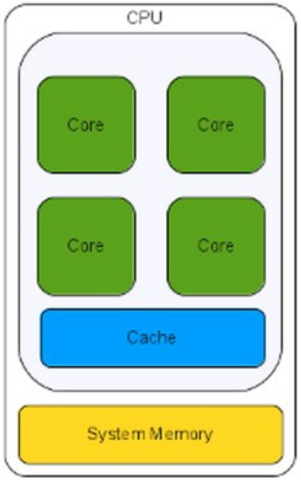
Multi-core versus many-core architecture

Multi-core

- O(10) cores
- Flexible: designed for both serial and parallel code
- Larger caches
- Emphasis on single thread performance

Many-core

- O(100-1000) cores
- Designed for parallel code
- Small caches
- Simpler cores



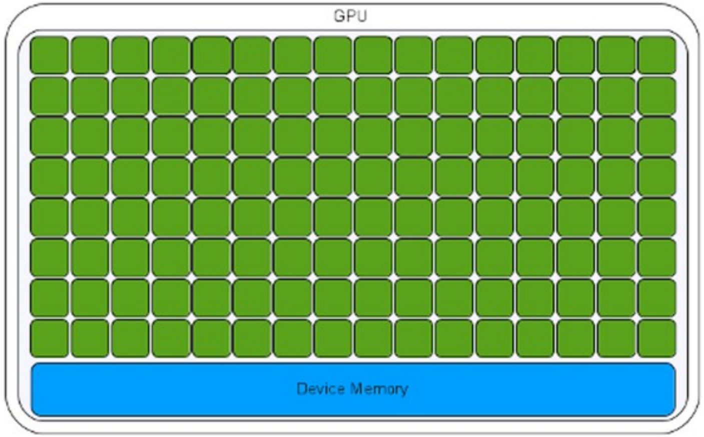


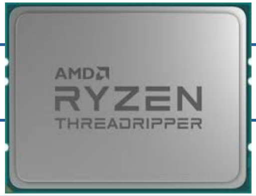
Image source

Types of GPUs

	Scientific GPUs	Gaming GPUs
Precision	~3 times more single precision TFLOPS than double precision → suited for double precision	~40 times more single precision TFLOPS than double precision → not well suited for double precision
Error correction	Available	Not available
Connection	NVLink & PCIe	Only PCIe
Price	~5-6 x the price of gaming GPUs	Several hundred dollars Depending on model (and year)

GPU vs. CPU: Specifications

	AMD Ryzen Threadripper 3990X	Nvidia A100
Core count	64 cores / 128 threads	6912 cores
Frequency	2.9 GHz	1.41 GHz
Peak Compute Performance	3.7 TFLOPs	19.5 TFLOPs (single precision)
Memory bandwidth	Max. 95 GB/s	1.6 TB/s
Memory capacity	Max O(1) TB	40/80 GB
Technology	7 nm	7 nm
Die size	717 mm^2	826 mm^2
Transistor count	3.8 billion	54.2 billion
Model	Minimize latency	Hide latency through parallelism



Connectivity with GPU: PCIe connection

PCIe generation	1 lane	16 lanes	Year of announcement
2.0	500 MB/s	8 GB/s	2007
3.0	985 MB/s	15.75 GB/s	2010
4.0	1.97 GB/s	31.5 GB/s	2011
5.0	3.94 GB/s	63 GB/s	2017
6.0	7.56 GB/s	121 GB/s	2019

https://en.wikipedia.org/wiki/PCI_Express

CPU – GPU - FPGA

	Latency	Connection	Engineering cost	FP performance	Serial / parallel	Memory	Backward compatibility
CPU	O(10) µs	Ethernet, USB, PCIe	Low entry level: Programmable with C++, pthon, etc.	O(1-10) TFLOPs	Optimized for serial, increasingly vector processing	O(100) GB RAM	Compatible, except for vector instruction sets
GPU	O(100) µs	PCIe, Nvlink	Low to medium entry level: Programmable with CUDA, OpenCL, etc.	O(10) TFLOPs	Optimized for parallel performance	O(10) GB	Compatible, exept for specific features
FPGA	Fixed O(100) ns	Any connection via PCB	High entry level: traditionally hardware description languages, Some high-level syntax available	Optimized for fixed point performance	Optimized for parallel performance	O(10) MB on the FPGA itself	Not easily backward compatible

https://arxiv.org/pdf/2003.11491.pdf

Overview of GPU usage in various HEP experiments

Experiment	Main tasks processed on GPU	Event / data rate	Number of GPUs	Deployment date
Mu3e	Track- & vertex reconstruction	20 MHz / 32 Gbit/s	O(10)	2023
CMS	Decoding, clustering, pattern recognition in pixel detector	100 kHz	O(400)	2022
ALICE	Track reconstruction in three sub-detectors	50 kHz Pb-Pb or < 5 MHz p-p / 30 Tbit/s	O(2000)	2022
LHCb	Decoding, clustering, track reconstruction in three sub-detectors, vertex reconstruction, muon ID, selections	30 MHz/ 40 Tbit/s	O(250)	2022

https://arxiv.org/pdf/2003.11491.pdf

Common characteristics of software frameworks

- Same code base compiled for various computing architectures: GPUs, x86,…
- Memory management system for GPU memory: avoid dynamic memory allocation
- Schedule pipelines of GPU (and CPU) algorithms → hide memory copies
- Integration into experiments' main software frameworks

Allen framework at LHCb

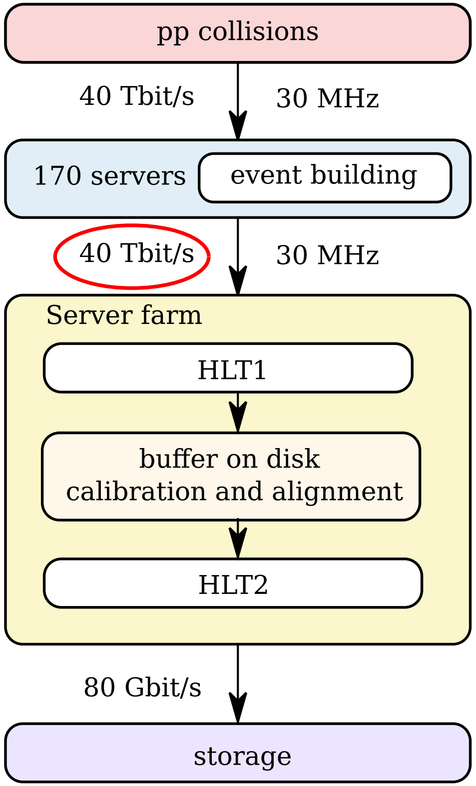
Patatrack at CMS

O2 at ALICE

History: HLT1 architecture choice

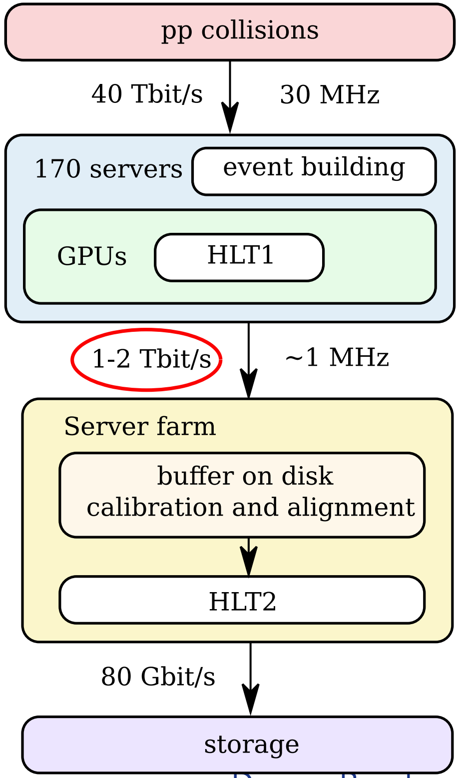
Proposal in TDR (2014)

CERN-LHCC-2014-016

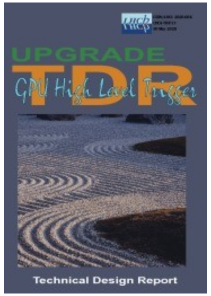


Updated strategy (as of 5/2020)

CERN-LHCC-2020-006



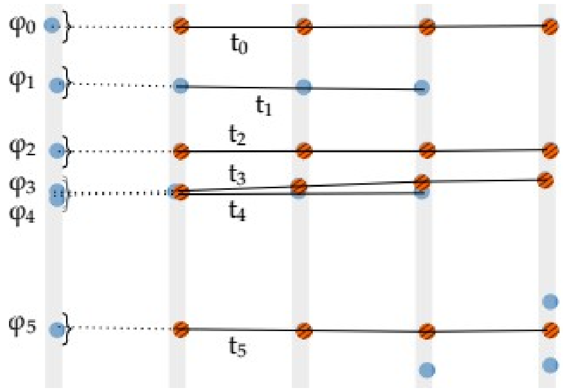
- Developed two solutions simultaneously
- Both the multi-threaded CPU & the GPU HLT1 fulfilled the requirements from the 2014 TDR
- Detailed cost benefit analysis (arXiv:2105.04031)
- GPU solution leads to cost savings on processors and the network
- Throughput headroom for additional features
- Decision: A GPU-based software trigger will allow LHCb to expand its physics reach in Run 3 and beyond.



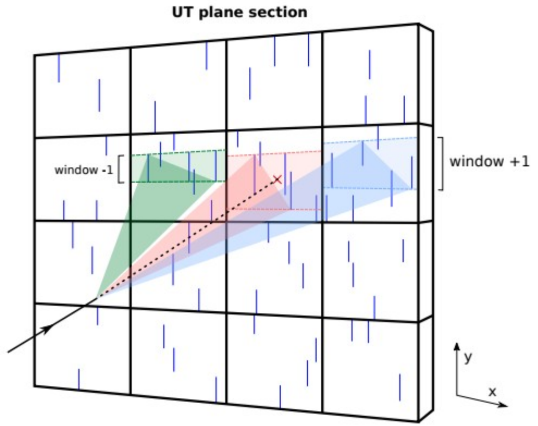
See also arXiv:2106.07701 on LHCb's energy efficiency with a CPU and GPU HLT1

Parallelization of reconstruction tasks

Search for combinations
of hits in parallel

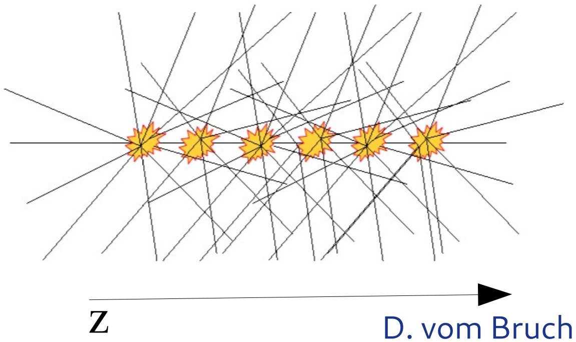


Store objects (for example hits)
In best suited memory layout

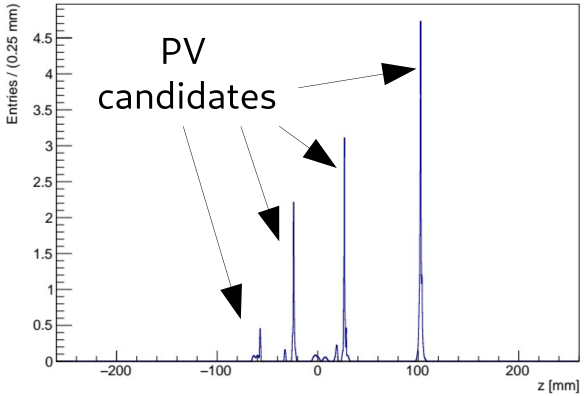


Split problem into
independent tasks

Example: primary vertex
(PV) reconstruction

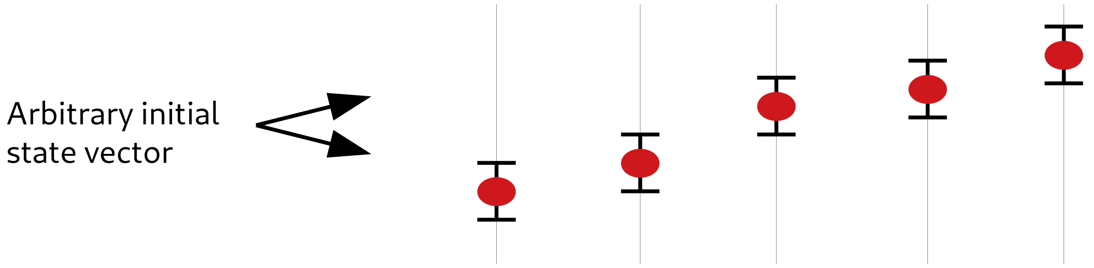


D. vom Bruch



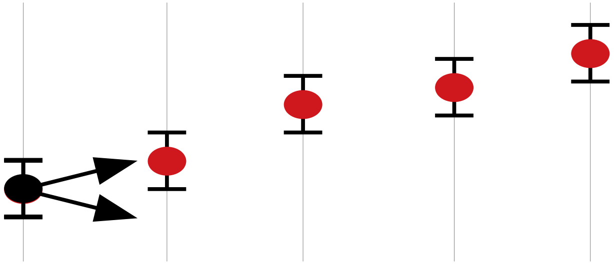
Kalman Filter

- One method for track fitting
- Subsequently iterates over all hits on a track
- For every hit, estimate the state of the track at that location:
  - First: predict it based on the previous state
  - Second: update it based on the measurement (hit)



Kalman Filter

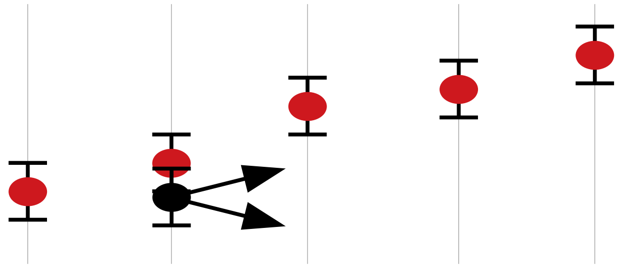
- One method for track fitting
- Subsequently iterates over all hits on a track
- For every hit, estimate the state of the track at that location:
    - First: predict it based on the previous state
    - Second: update it based on the measurement (hit)



First update

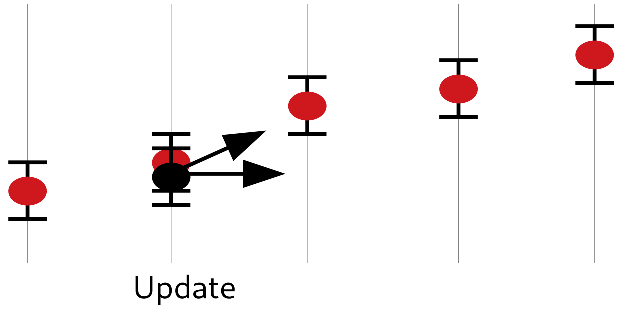
Kalman Filter

- One method for track fitting
- Subsequently iterates over all hits on a track
- For every hit, estimate the state of the track at that location:
  - First: predict it based on the previous state
  - Second: update it based on the measurement (hit)



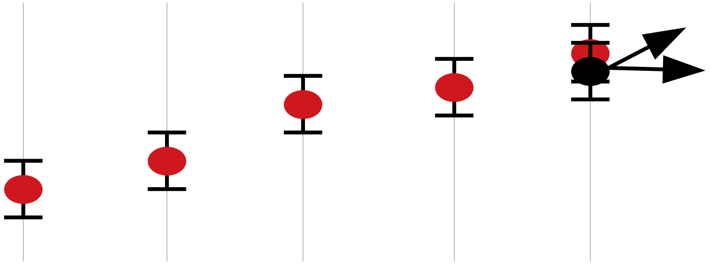
Kalman Filter

- One method for track fitting
- Subsequently iterates over all hits on a track
- For every hit, estimate the state of the track at that location:
    - First: predict it based on the previous state
    - Second: update it based on the measurement (hit)



Kalman Filter

- One method for track fitting
- Subsequently iterates over all hits on a track
- For every hit, estimate the state of the track at that location:
  - First: predict it based on the previous state
  - Second: update it based on the measurement (hit)
- At last plane: best linear estimator for track state



- Only parallelizable over all tracks in one event