Online Track Reconstruction on GPUs for the Mu3e and LHCb experiments

IEEE Real Time 2018, Williamsburg, VA

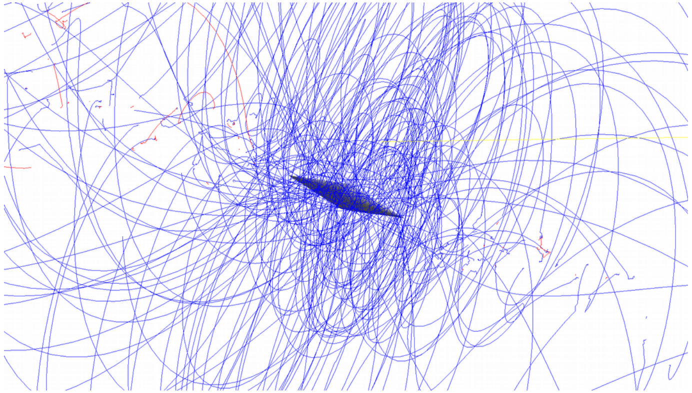
Dorothea vom Bruch

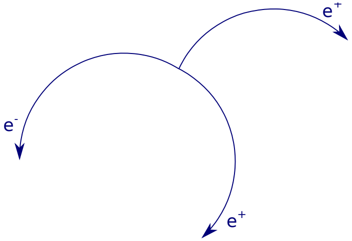
For the Mu3e and LHCb Collaborations

June 11th, 2018

Why Reconstruct Tracks Online?

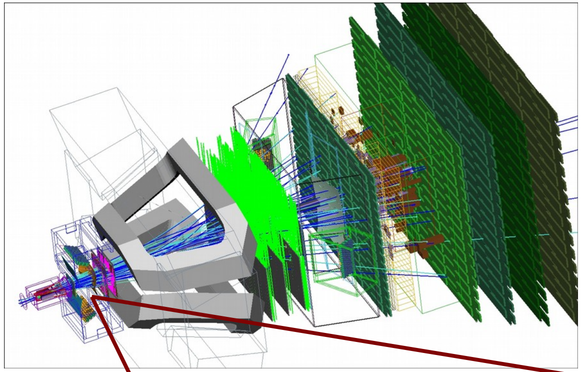
Mu3e: $\mu^+ \rightarrow e^+ e^- e^+$

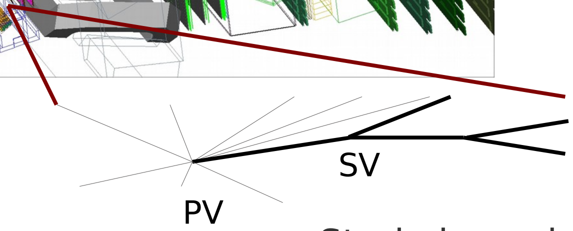




Continuous muon beam

LHCb: Specific decays





Study b- and c-hadrons

LHC bunch crossing at 40 MHz

Find specific track patterns

Challenges

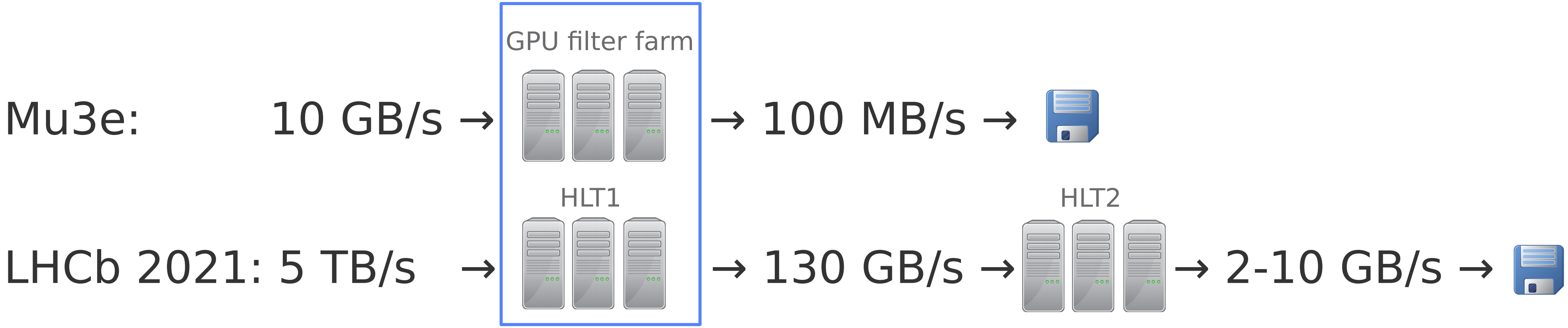
- No “easy” pattern to be used by hardware triggers
- Reconstruct all tracks at MHz rates
- Reduce data stream:

Mu3e: 10 GB/s →  → 100 MB/s → 

LHCb 2021: 5 TB/s →  → 130 GB/s →  → 2-10 GB/s → 

Challenges

- No “easy” pattern to be used by hardware triggers
- Reconstruct all tracks at MHz rates
- Reduce data stream:



Track reconstruction

Why Consider GPUs?

- Need many FPs/s at low cost
- Care about high throughput
- Market changing rapidly → be flexible with design choice

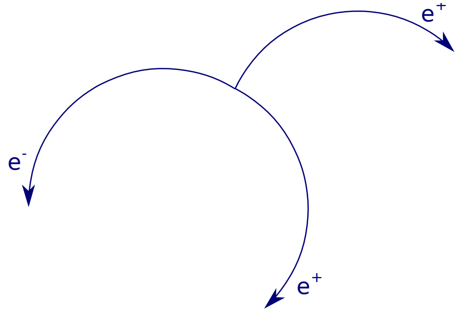
→ Go massively parallel on GPUs

Image sources: http://images.anandtech.com, http://images.nvidia.com/pascal

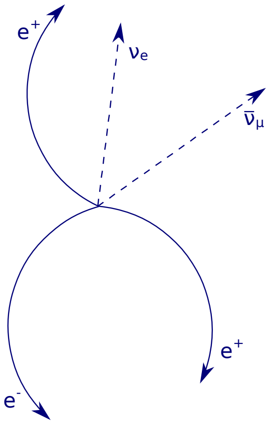
Mu3e

Mu3e Phase I

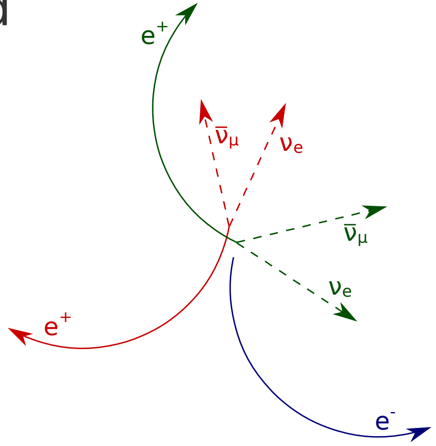
- Search for charged lepton flavour-violating decay $\mu^+ \rightarrow e^+e^-e^+$ with a sensitivity in branching ratio better than $2{\cdot}10^{-15}$
- In Standard Model: suppressed to BR $< 10^{-54}$
- Any hint of signal → new physics
- Located at Paul-Scherrer Institute (PSI), Switzerland



Signal

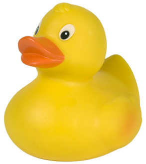


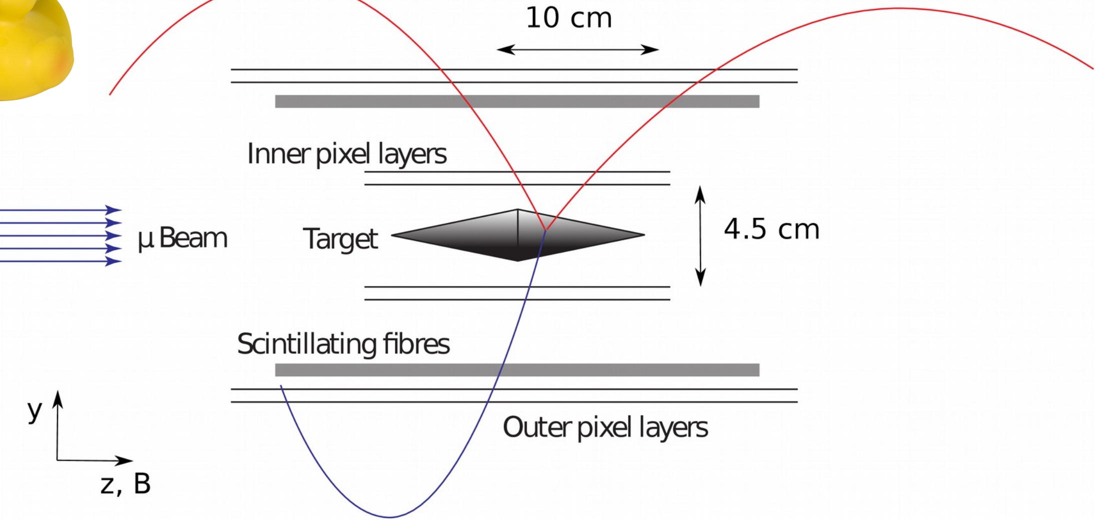
Internal conversion

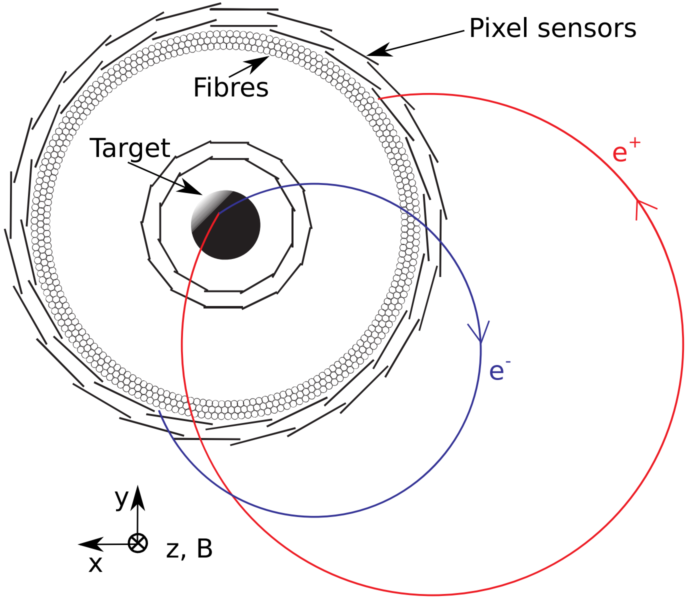


Combinatorics

Mu3e Detector, Central Region

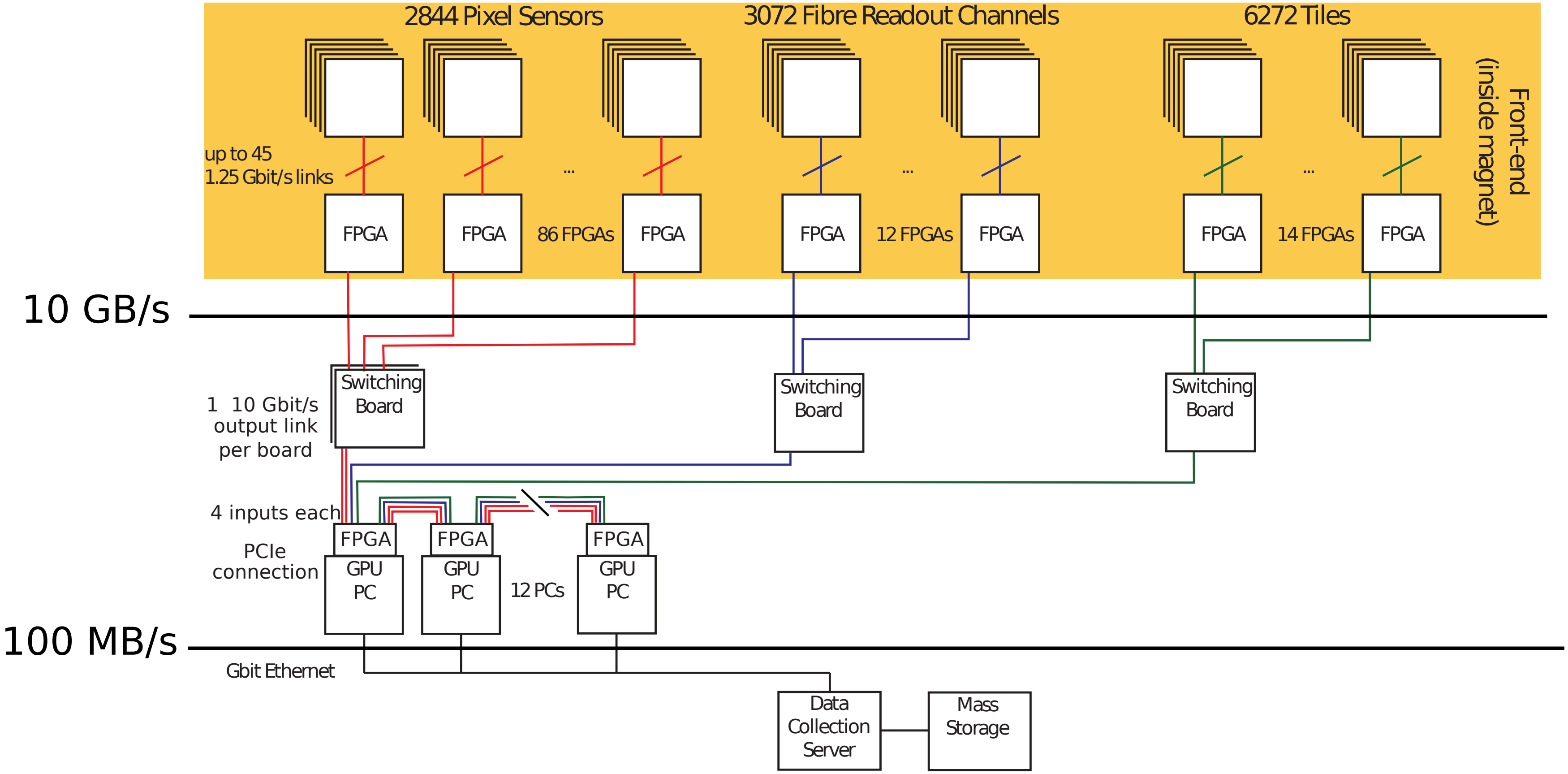




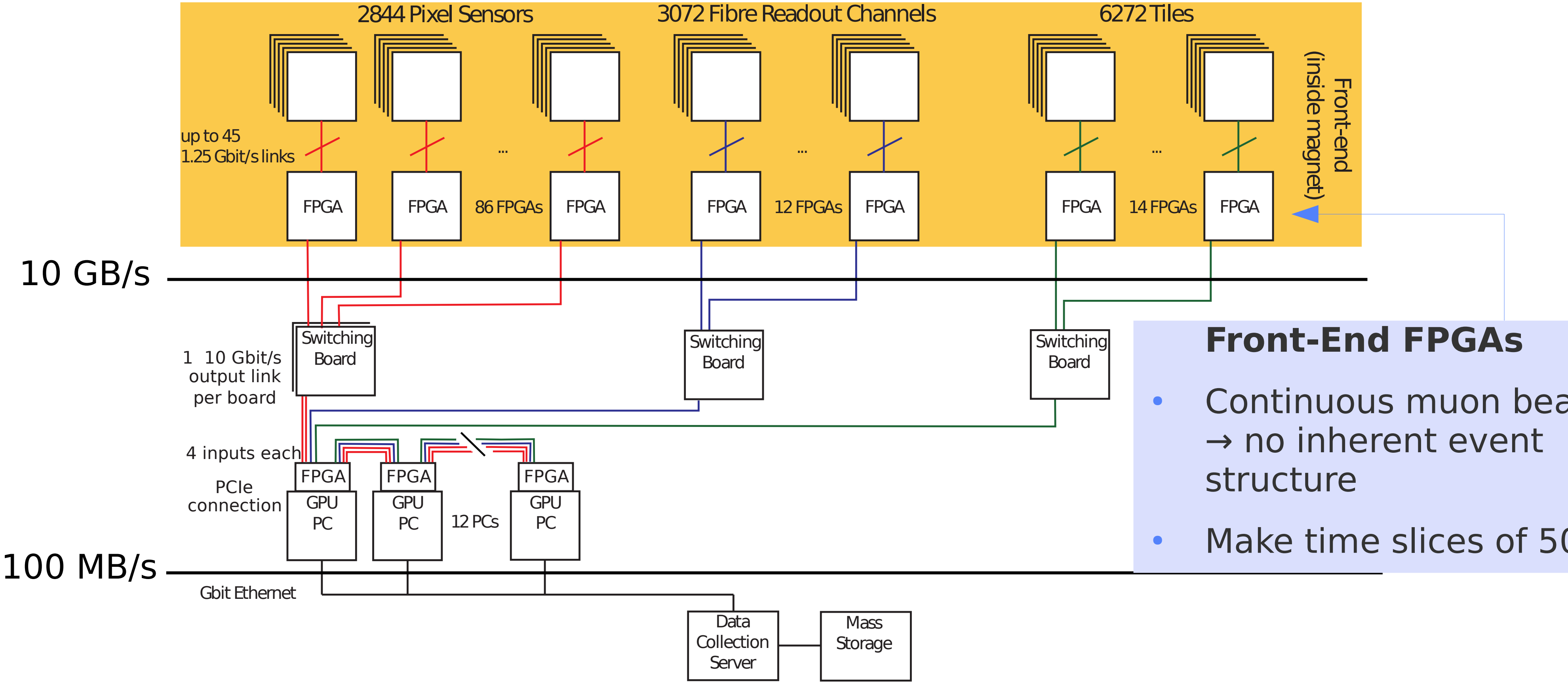


10^8 μ/s stopped on target, in solenoidal magnetic field

Mu3e Readout Scheme



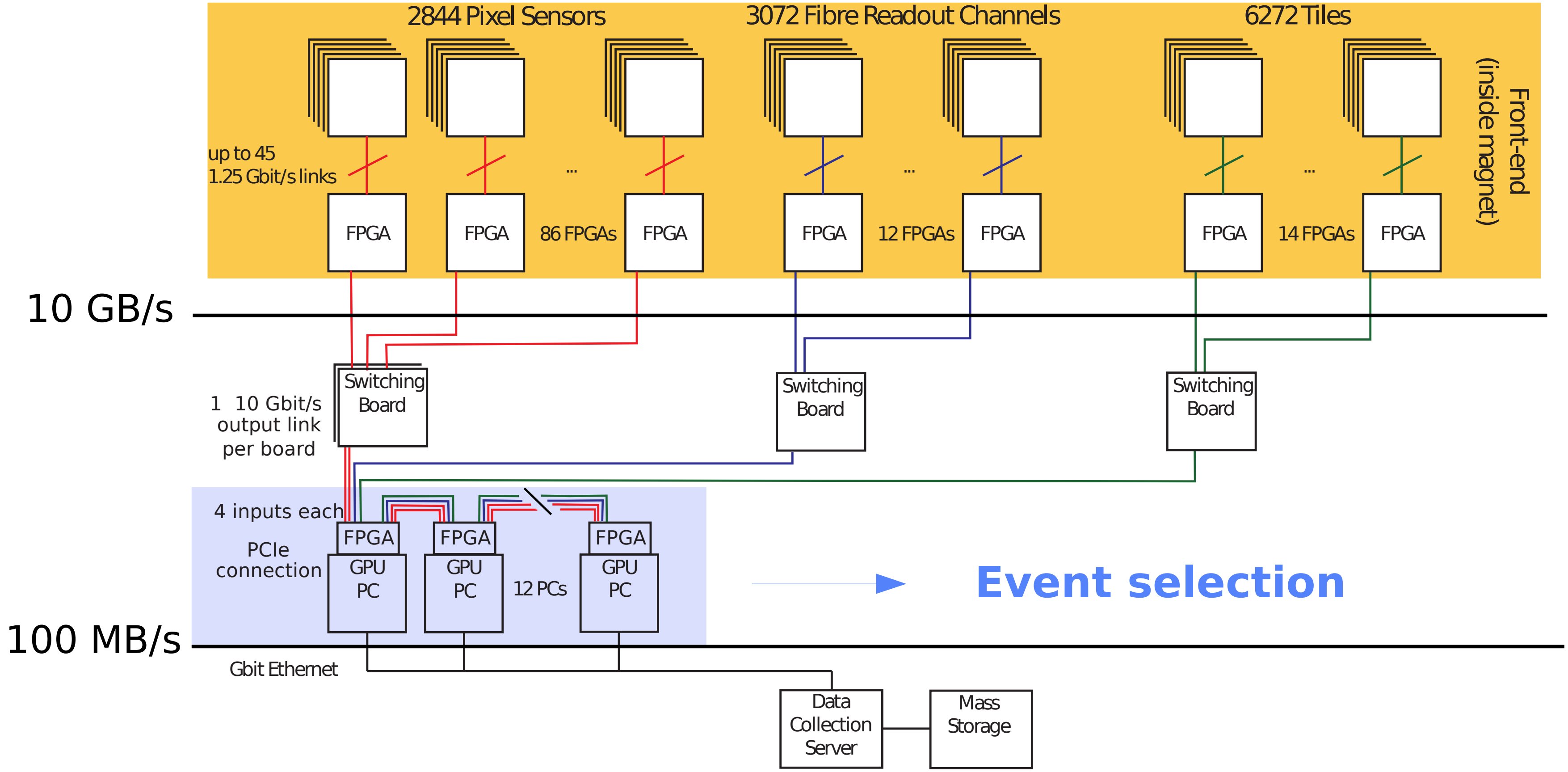
Mu3e Readout Scheme



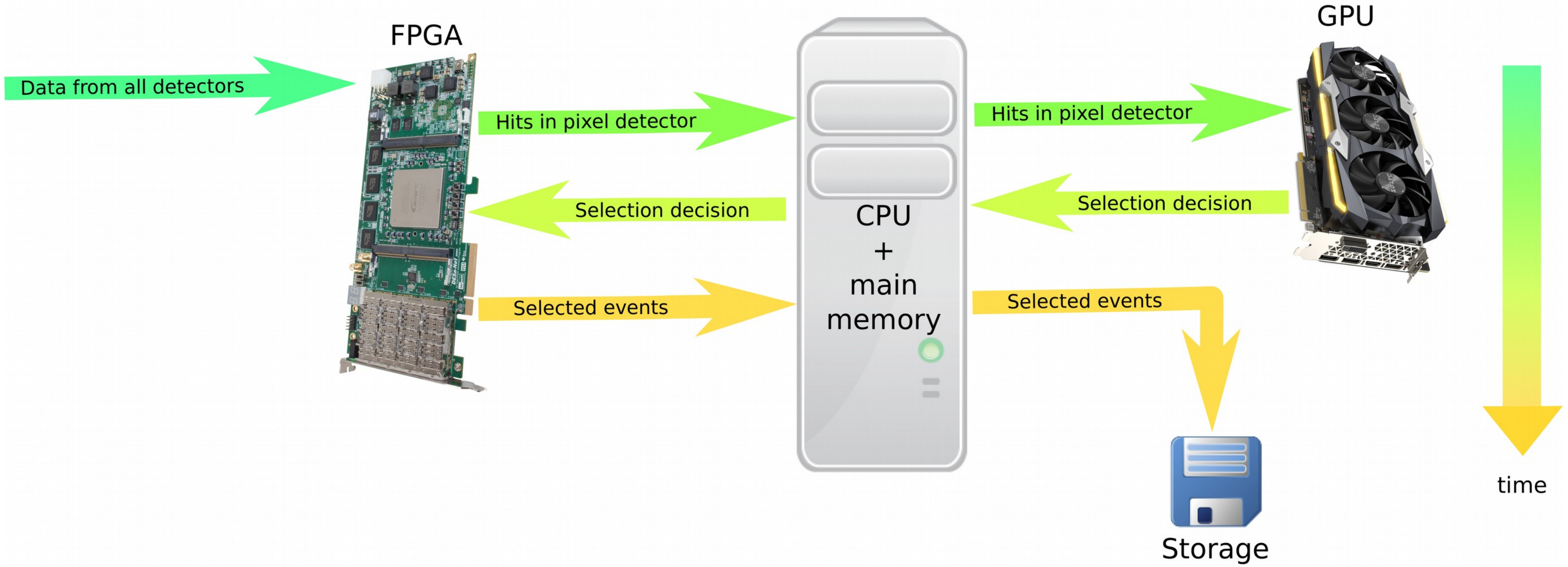
Front-End FPGAs

- Continuous muon beam → no inherent event structure
- Make time slices of 50 ns

Mu3e Readout Scheme



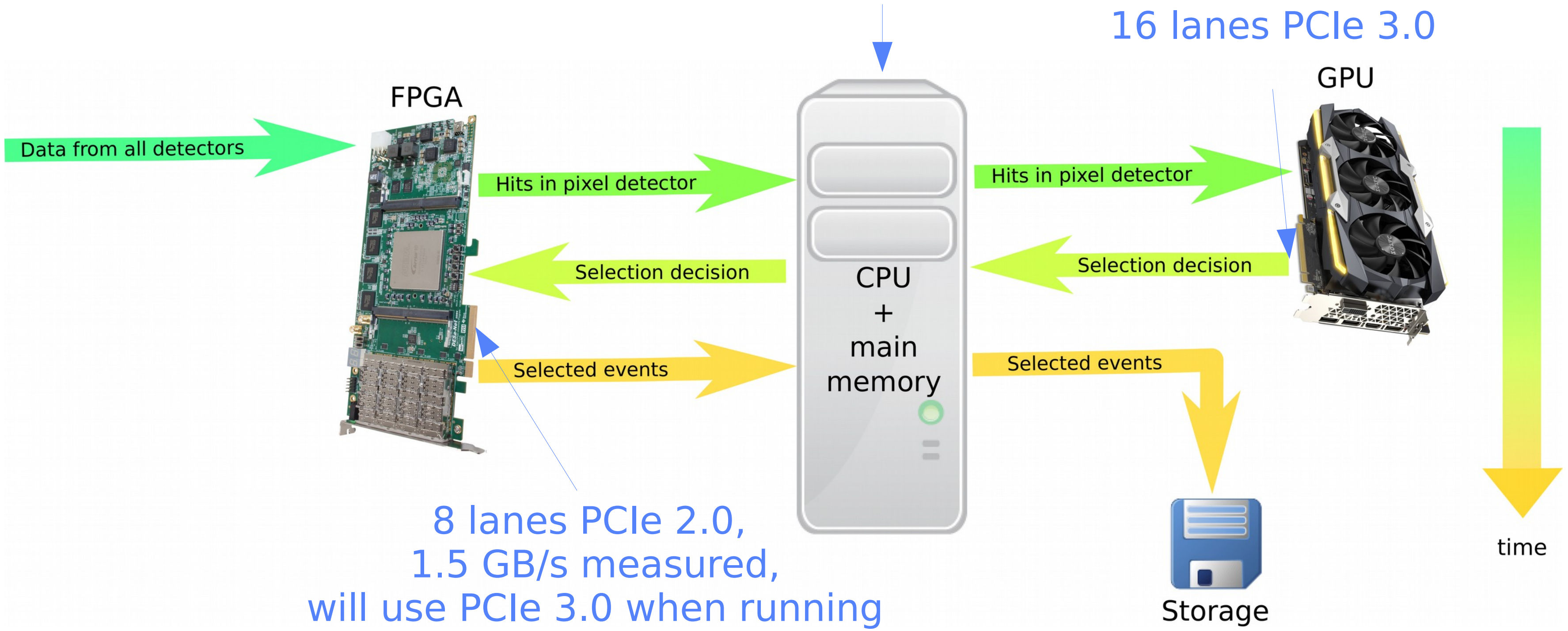
Data Flow



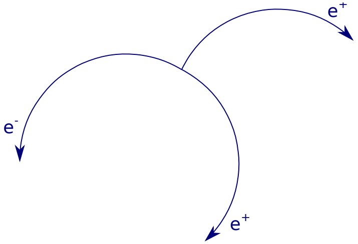
FPGA
Data from all detectors
Hits in pixel detector
Selection decision
Selected events
CPU
+
main
memory
Hits in pixel detector
Selection decision
Selected events
GPU
Storage
time

Data Flow

Direct Memory Acces (DMA) to and from one memory buffer in main memory
→ No extra copy



Mu3e: Find 3 Signal Tracks



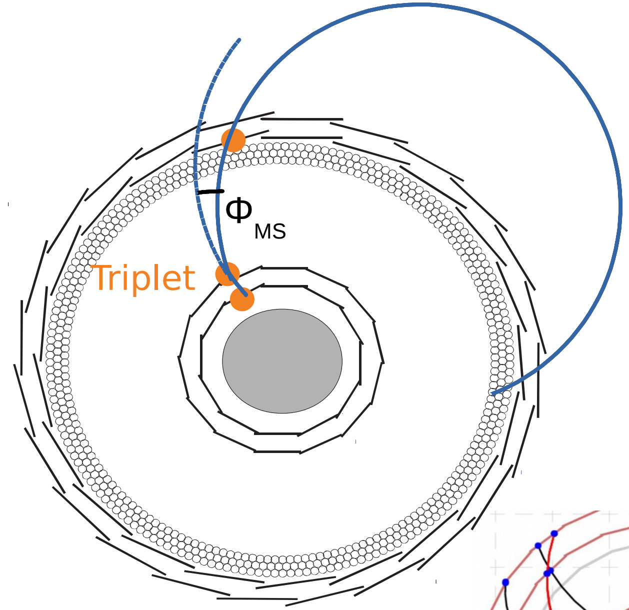
FPGA	GPU
Find combinations of hits in first 3 layers	Fit tracks with fit developed for multiple scattering dominated resolution
Transfer these + hits in 4th layer to GPU	
Reduce 3-hit combinations by factor 70	

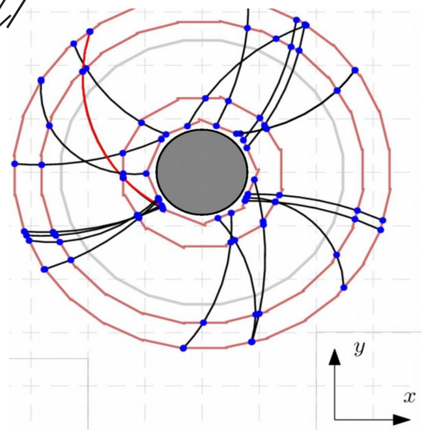
Currently emulated on GPU

Mu3e: Multiple Scattering Fit

- Electrons: 12 – 53 MeV/c
- Resolution dominated by multiple Coulomb scattering
- Ignore hit uncertainty
- Three consecutive hits: “triplet”
- Assume multiple scattering at middle hit, minimize x^2

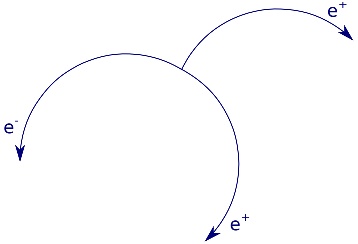
$$\chi^2 = \frac{\Phi_{MS}^2}{\sigma_{MS,\Phi}^2} + \frac{\theta_{MS}^2}{\sigma_{MS,\theta}^2}$$





N. Berger, A. Kozlinskiy, M. Kiehn, A. Schöning, NIM A, 2017, arXiv: 1606.04990

Mu3e: Find 3 Signal Tracks



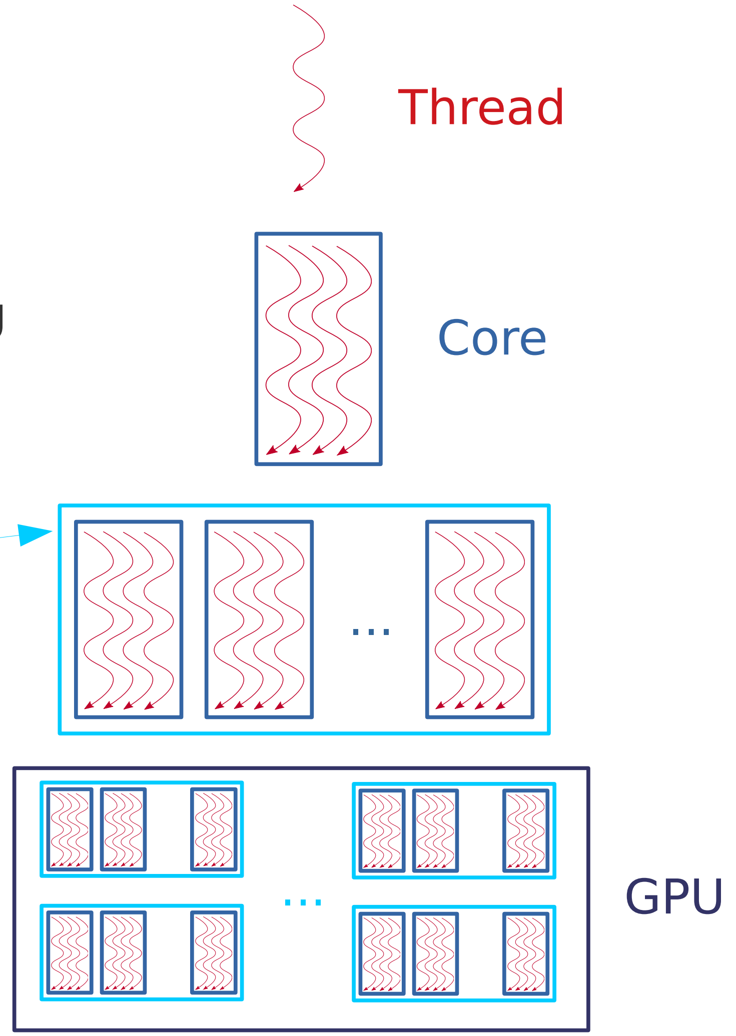
FPGA	GPU
Find combinations of hits in first 3 layers	Fit tracks with fit developed for multiple scattering dominated resolution
Transfer these + hits in 4^{th} layer to GPU	Positive tracks → Negative tracks → Select combinations of 2 positive, 1 negative track from one vertex, based on circle intersections
Reduce 3-hit combinations by factor 70	Reduce # of time slices by factor 140

Currently emulated on GPU

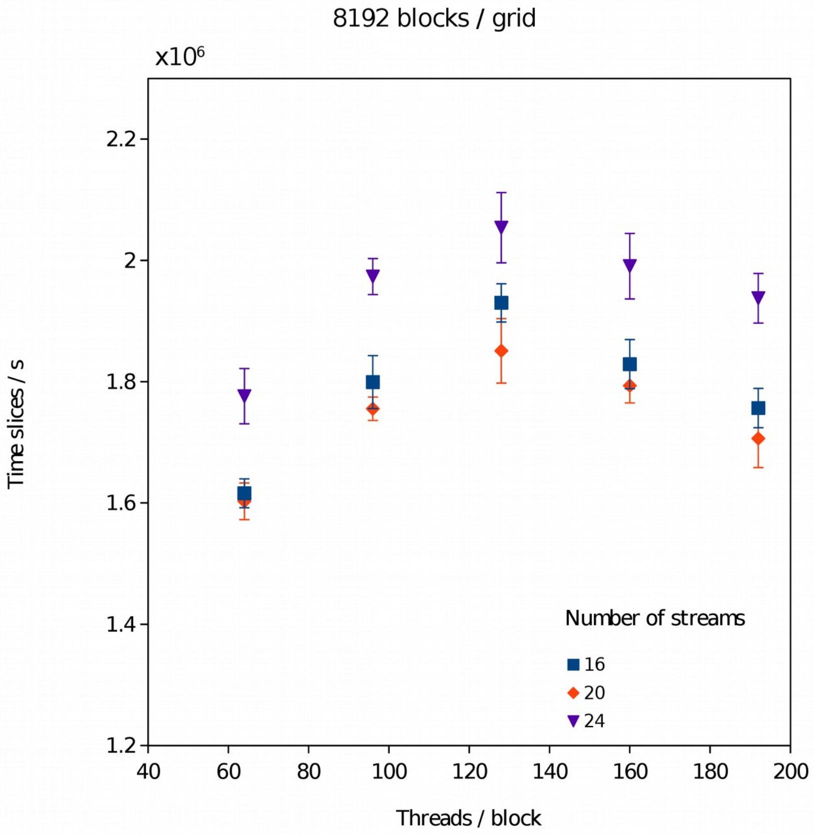
Signal selection efficiency: 98%

GPU Architecture

- Lower frequency, higher cache latency than CPU
- Thousands of threads → hide latency
- SIMD: Single Instruction Multiple Data
- Threads in lockstep: Run on one multiprocessor unit in parallel, execute the same instruction
- Thread count per multiprocessor typically multiple of ALU count
- Threads on different cores / multiprocessors: completely independent of each other

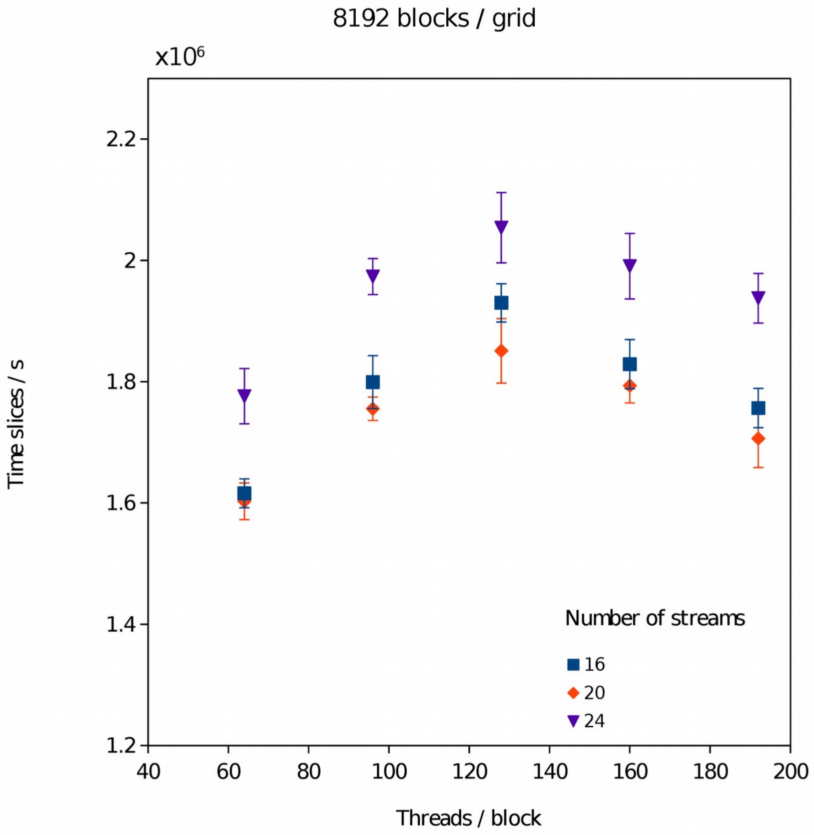


Mu3e Online Selection on GPUs



- Use 24 streams → memory copy & computations concurrently
- Optimize grid: launch 8192 time slices in one grid, with 128 threads / block
- Optimized memory layout, prepare it on FPGA (3-hit pre-selection)
- Use single precision

Mu3e Online Selection on GPUs



- Use 24 streams → memory copy & computations concurrently
- Optimize grid: launch 8192 time slices in one grid, with 128 threads / block
- Optimized memory layout, prepare it on FPGA (3-hit pre-selection)
- Use single precision

Process $2 \cdot 10^6$ time slices / s on one GTX1080Ti

→ 12 PCs enough, meet requirements :-)

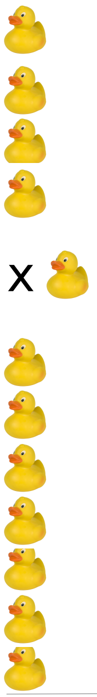
D. vom Bruch, PhD thesis, 2017, Heidelberg University

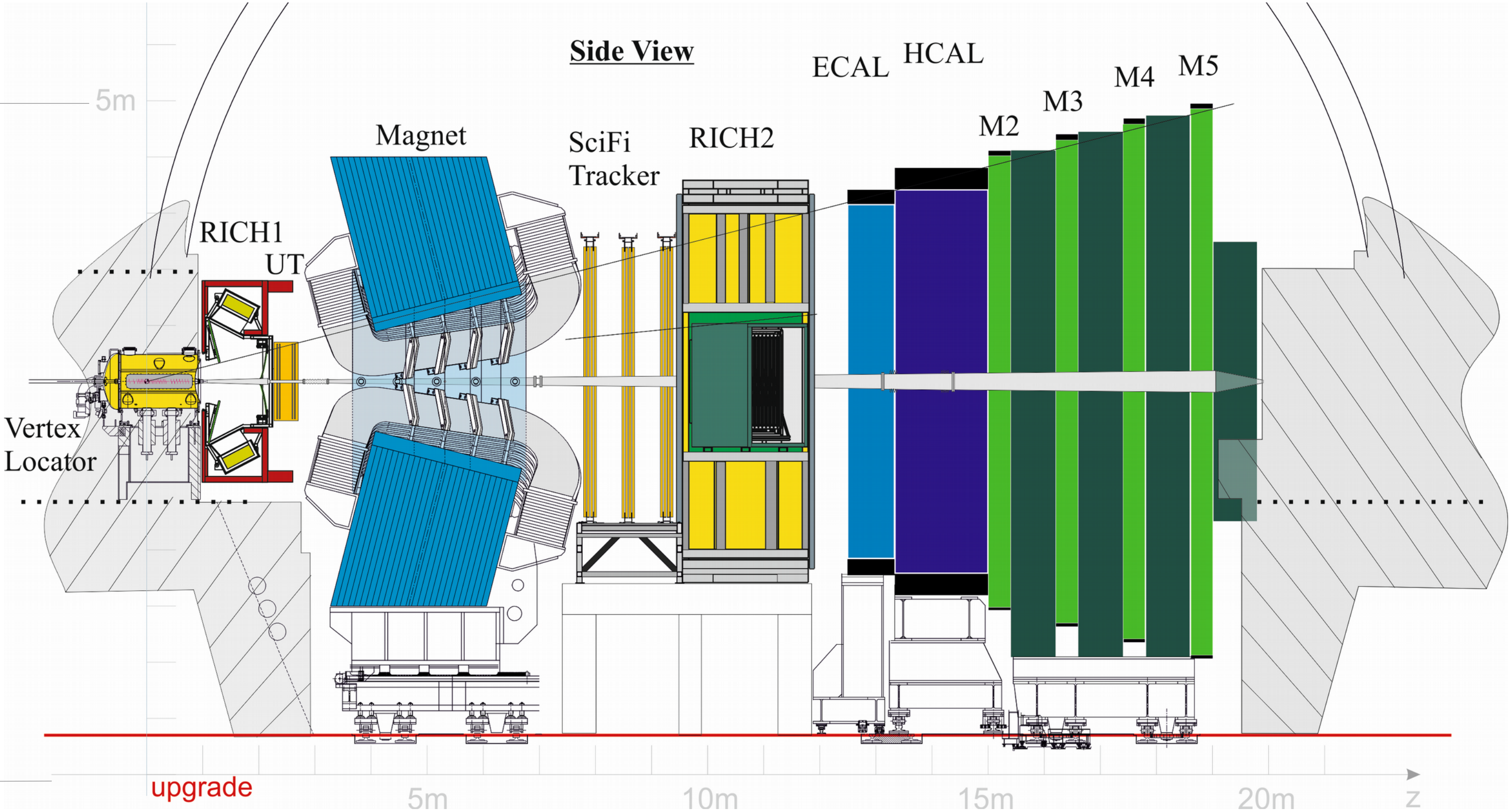
LHCb

LHCb Detector (2021+)

@ LHC, CERN, Switzerland

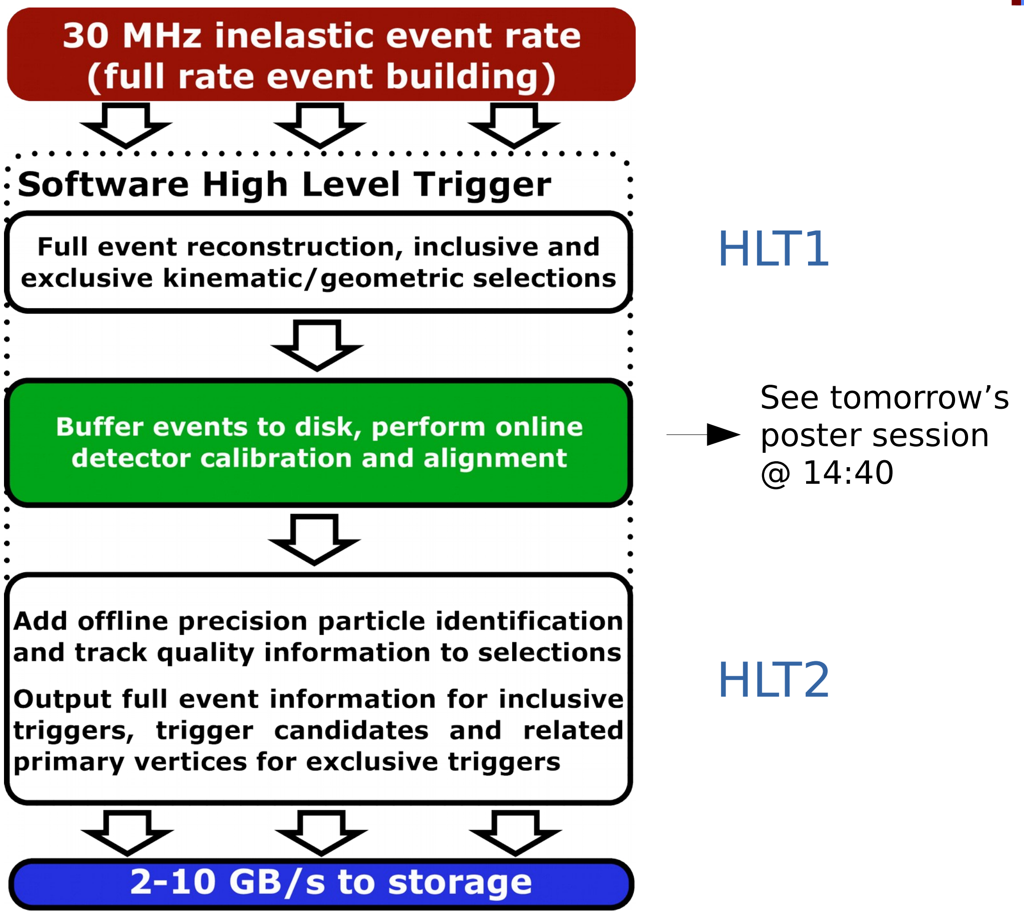
62 x





LHCb Upgrade TDR, LHCb Tracker, 2015, CERN-LHCC-2014-001

LHCb Software Trigger

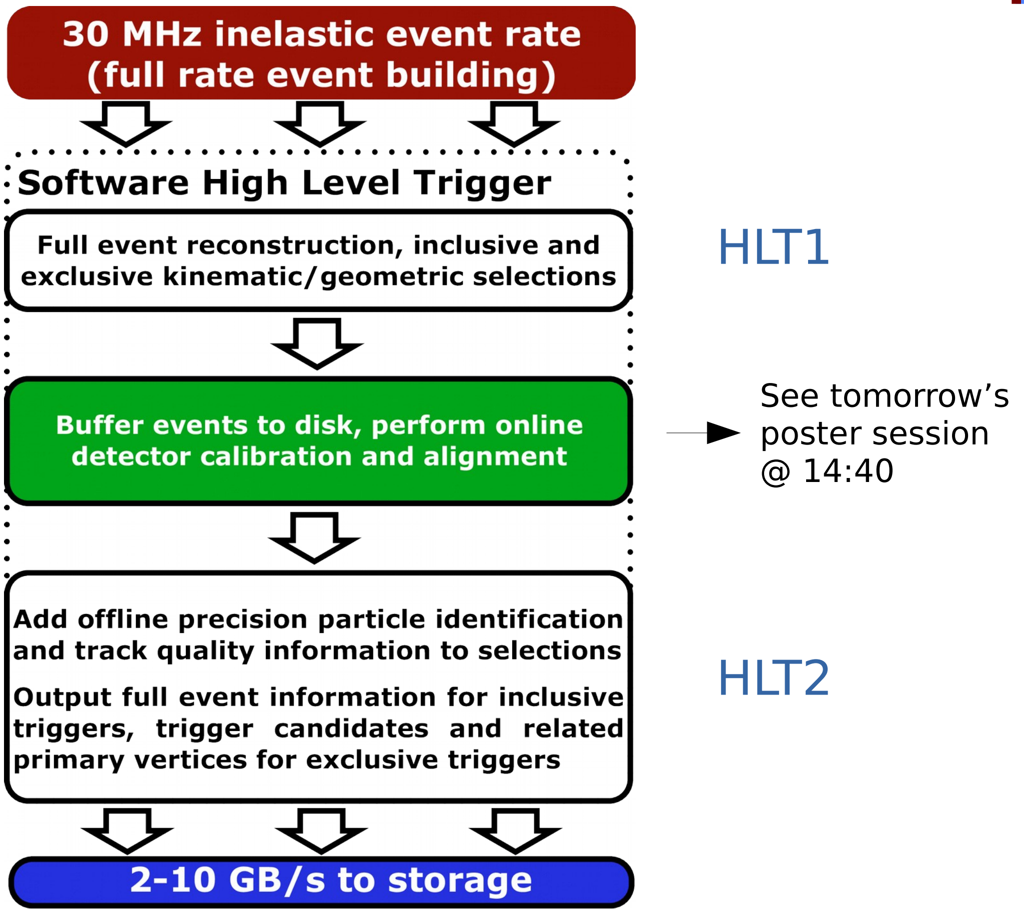


LHCb Software Trigger

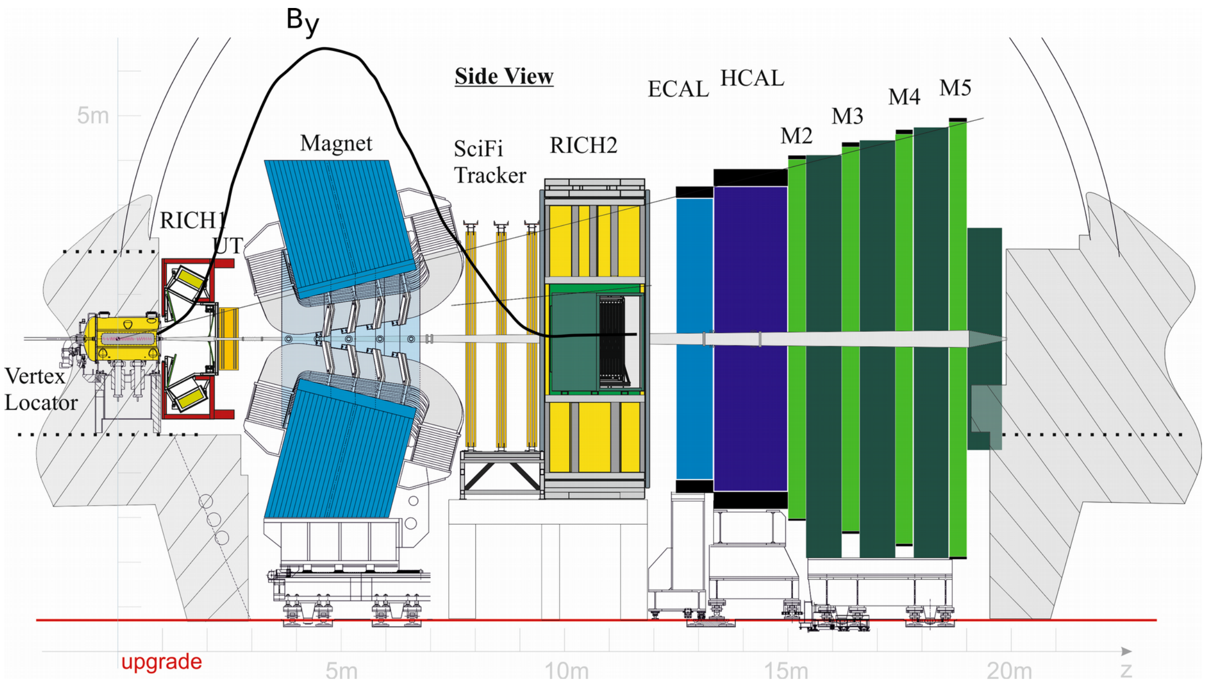
- 27 GB/s beauty hadrons available → secondary vertices alone do not suffice

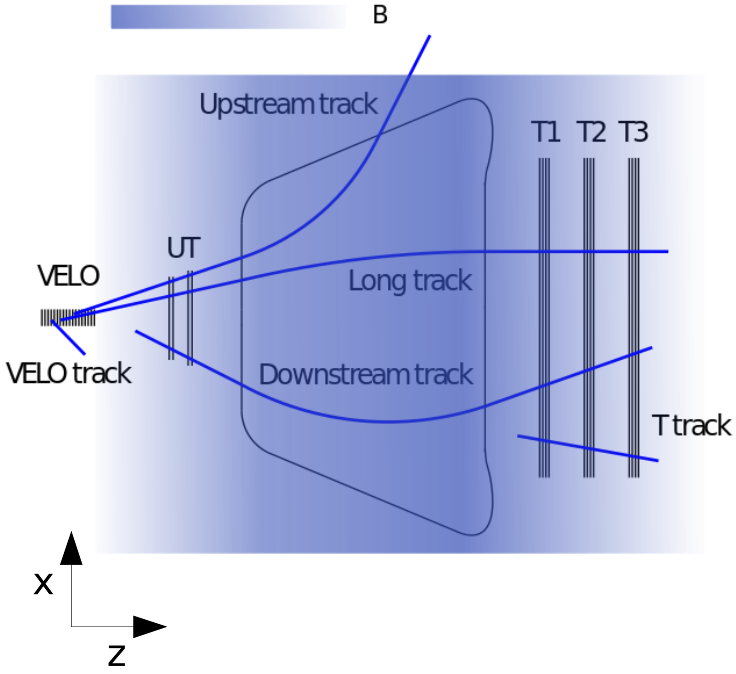
  → need further selection
- Computing challenge
- Study different architectures

  → choose the one with best physics performance / $
- One option: run HLT1 (or parts of it) on GPUs

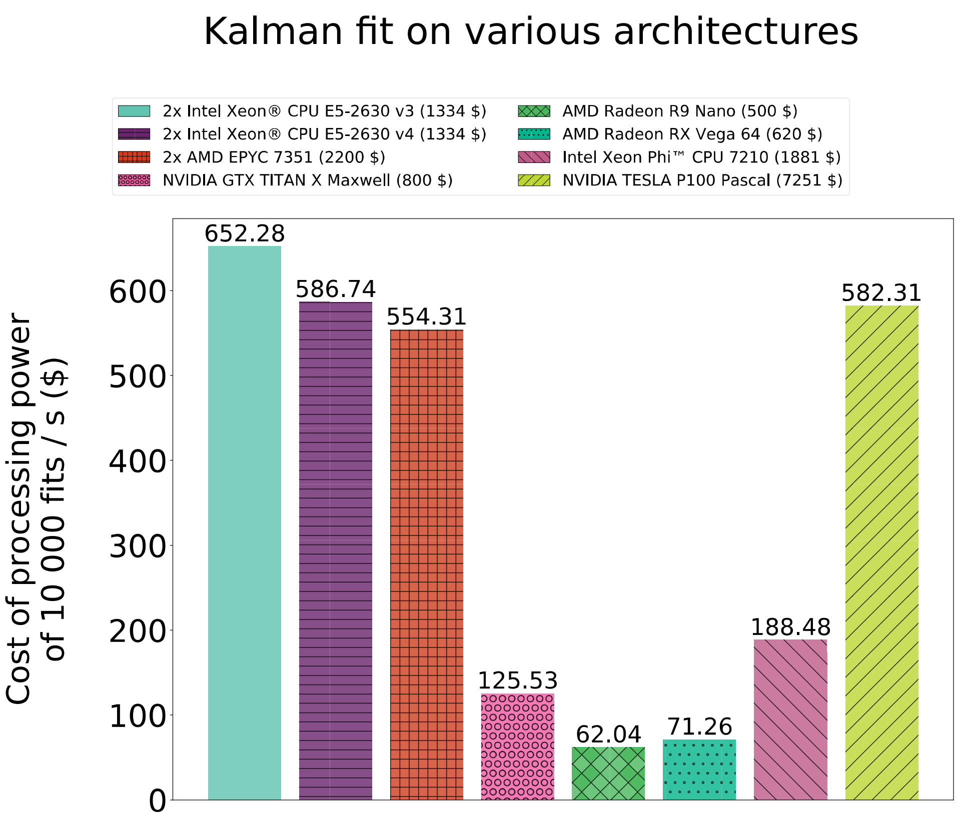


LHCb Track Reconstruction





LHCb Track Reconstruction on GPUs

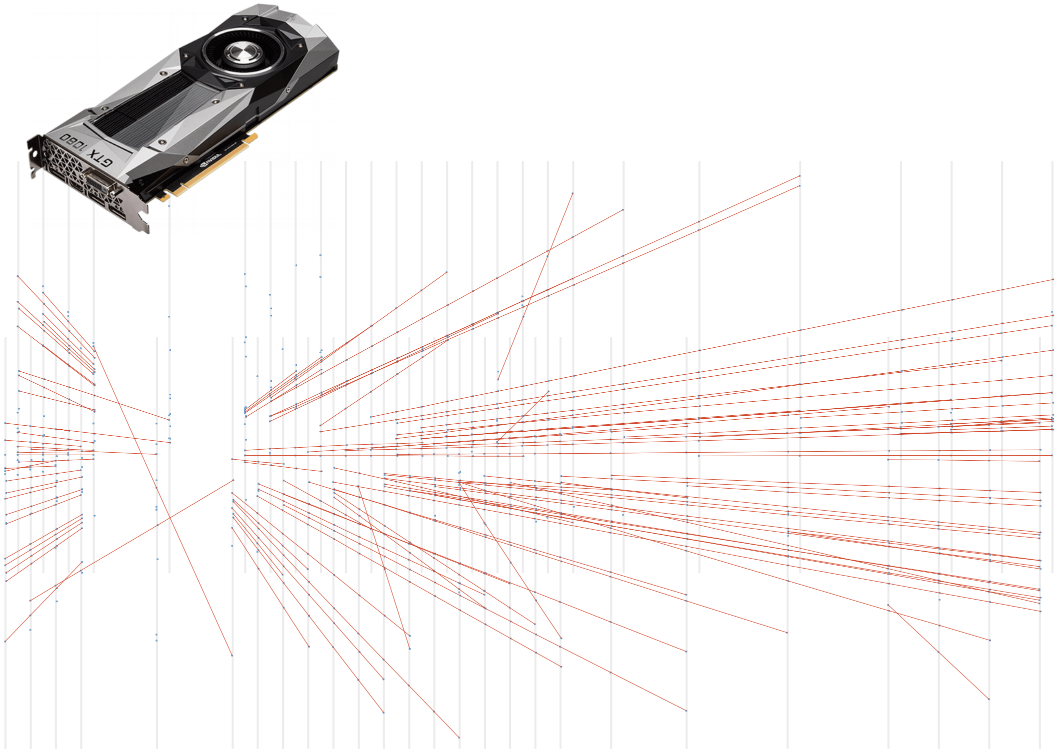


D. Cámpora, O. Awile, https://doi.org/10.1002/cpe.4483, 2018

GPUs have shown to be price-performant for track fitting in LHCb

→ motivates us to study the performance of the whole sequence on GPUs

LHCb Velo Tracks on GPUs



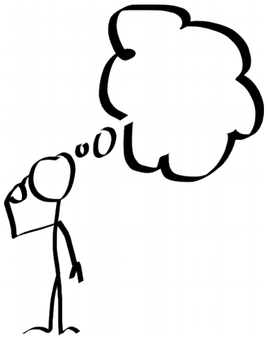
A. Badalov, D. Cámpora, N. Neufeld, X. Vilasís-Cardona,
JINST, 2016, 10.1088/1748-0221/11/01/P01001

- Tracks in VELO: straight lines
- Implemented on GPU
- Use as input for Velo → UT tracking

- Aim for similar setup as in Mu3e farm: place GPU in filter farm PC, transfer data via CPU
- GPU work: preparation for event building

LHCb: GPU Outlook

- Work in progress: VELO → UT tracking → first momentum estimate
- Later this year: add SciFi tracking and muon stations
  → full track reconstruction
- At the end of the year: Is the GPU solution realistic?



Conclusions

Mu3e

- Can reduce data rate by factor 140
- Keep 98 % of signal decays
- Process $2 \cdot 10^6$ time slices / s on one GTX1080Ti
- → Can run on the planned 12 DAQ PCs

LHCb

- Promising first results
- Velo tracking implemented on GPU
- Velo → UT tracking work in progress
- Plan: Implement full reconstruction chain

Backup

LHCb: Per Event Yields

For generator-level Monte-Carlo

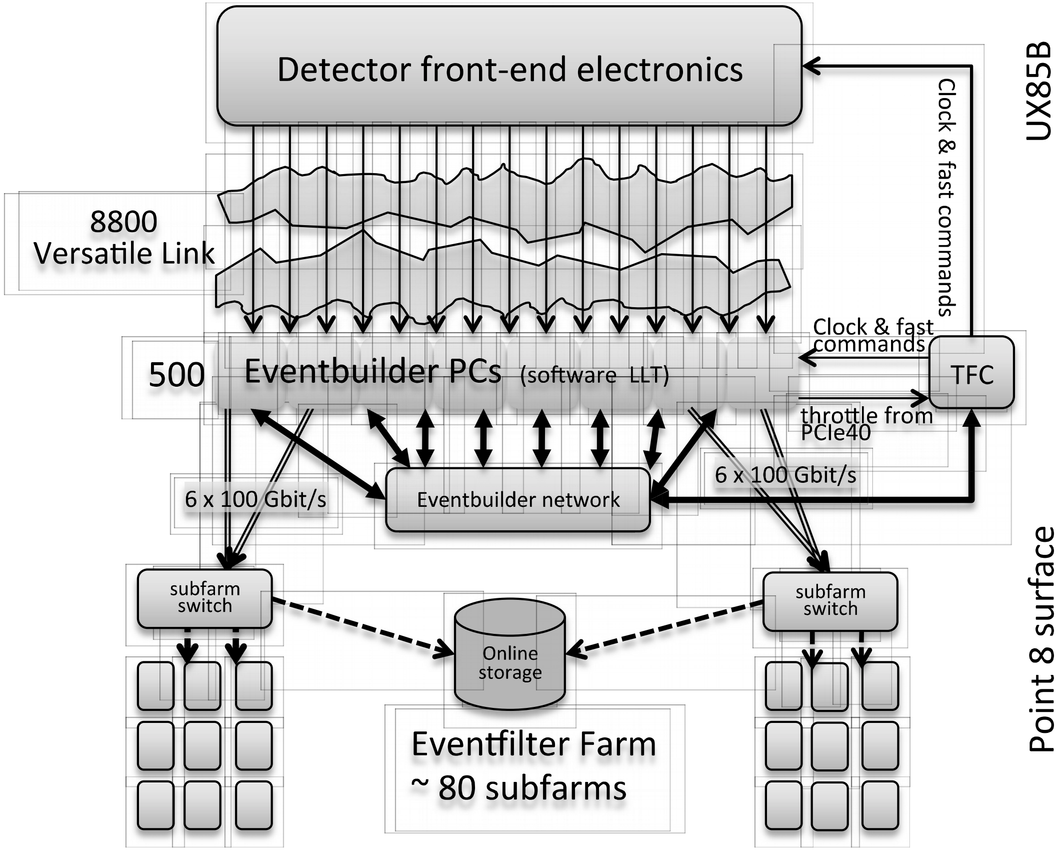
ε(VELO): efficiency for candidates having at least two tracks traversing at least three modules in the VELO
ε(LHCb): efficiency for candidates having all child tracks contained in the LHCb acceptance

Run 1, Original VELO geometry			
Category	Yield in 4π	ϵ(VELO)	ϵ(VELO) $\times$ ϵ(LHCb)
b-hadrons	0.0258 ± 0.0004	$30.5 \pm 0.6\%$	$11.1 \pm 0.4\%$
c-hadrons	0.297 ± 0.001	$21.9 \pm 0.2\%$	$14.2 \pm 0.1\%$
light, long-lived hadrons	8.04 ± 0.01	$6.67 \pm 0.02\%$	$6.35 \pm 0.02\%$

Upgrade, nominal luminosity, VELO pixel geometry			
Category	Yield in 4π	ϵ(VELO)	ϵ(VELO) $\times$ ϵ(LHCb)
b-hadrons	0.1572 ± 0.0004	$34.9 \pm 0.1\%$	$11.9 \pm 0.1\%$
c-hadrons	1.422 ± 0.001	$24.73 \pm 0.04\%$	$15.12 \pm 0.03\%$
light, long-lived hadrons	33.291 ± 0.006	$7.022 \pm 0.004\%$	$6.257 \pm 0.004\%$

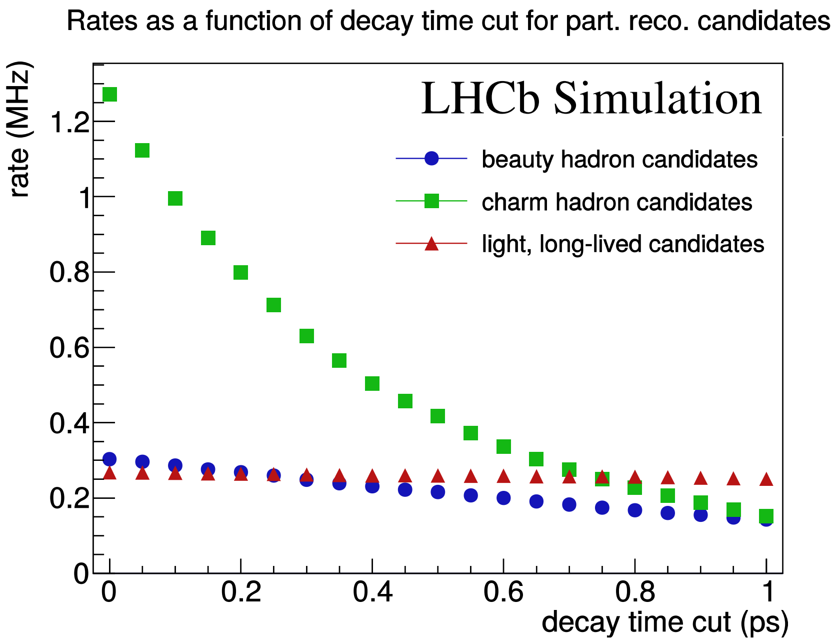
Source: LHCb Upgrade TDR, Trigger and Online, CERN/LHCC 2014-016

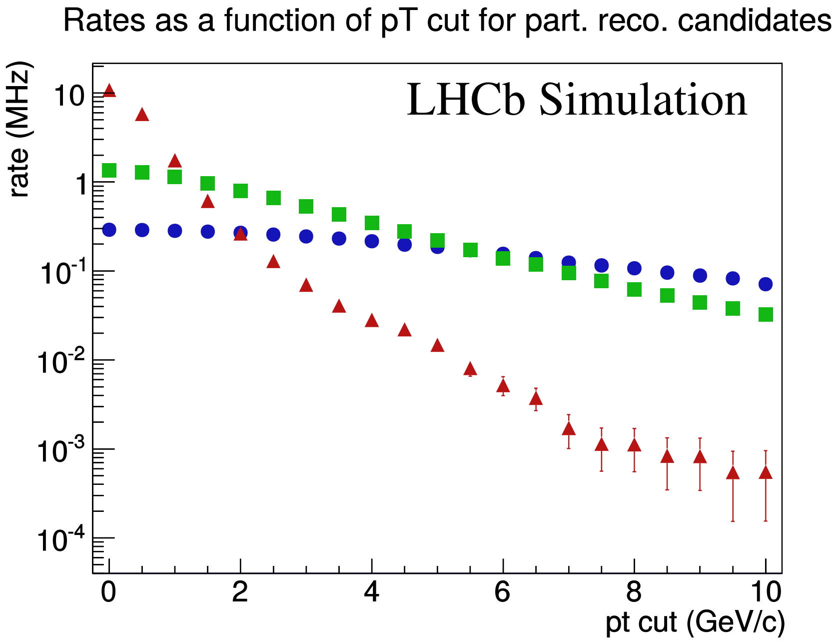
LHCb Upgrade Readout Scheme



LHCb Upgrade TDR, Trigger and Online, CERN/LHCC 2014-016

LHCb: Secondary Vertices



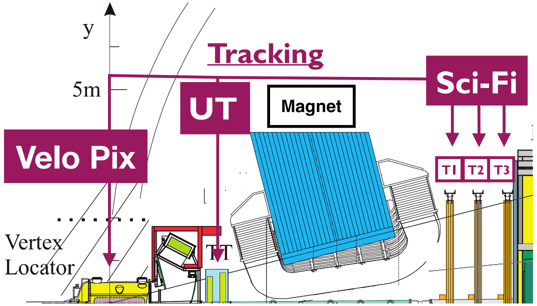


Source: LHCb Upgrade TDR, Trigger and Online, CERN/LHCC 2014-016

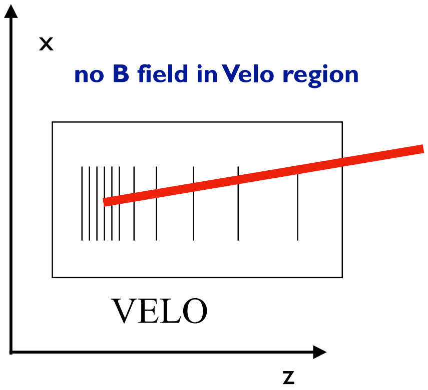
Partial reconstruction sequence : Velo tracking and PV finding

Velo Pix

- **2-D pixel sensor** @-20°C: direct **x-y-z** measurement.
- **Tracking on raw data.**
- **5.1 mm distance** from interaction point (8.2 mm current VELO)



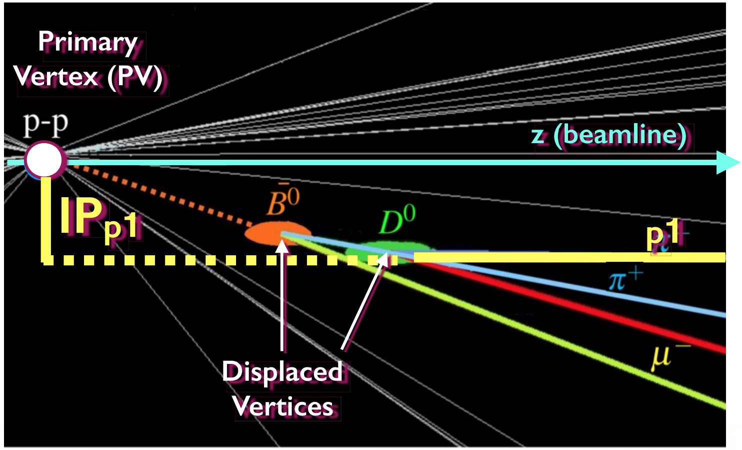
Velo tracking and PV reconstruction



Velo tracking

Velo tracks used to find the PVs.

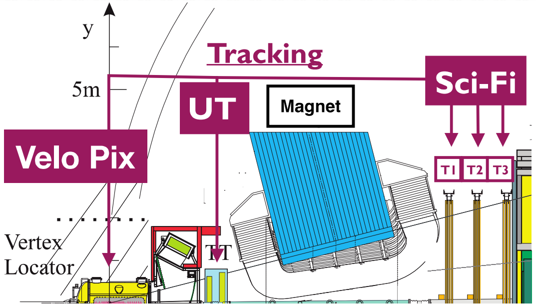
PV position used to identify displaced tracks in the event.



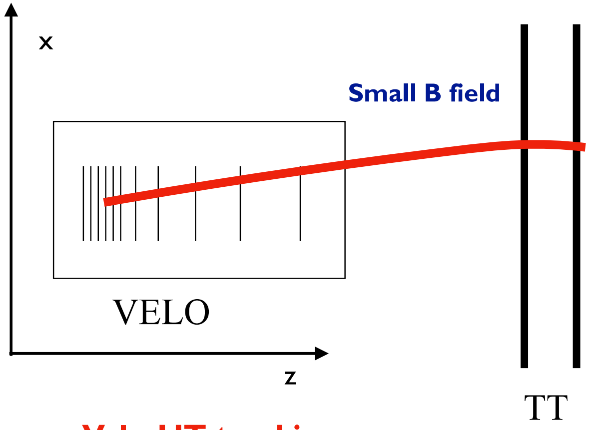
Partial reconstruction sequence : Velo-UT tracking

Upstream tracker (UT)

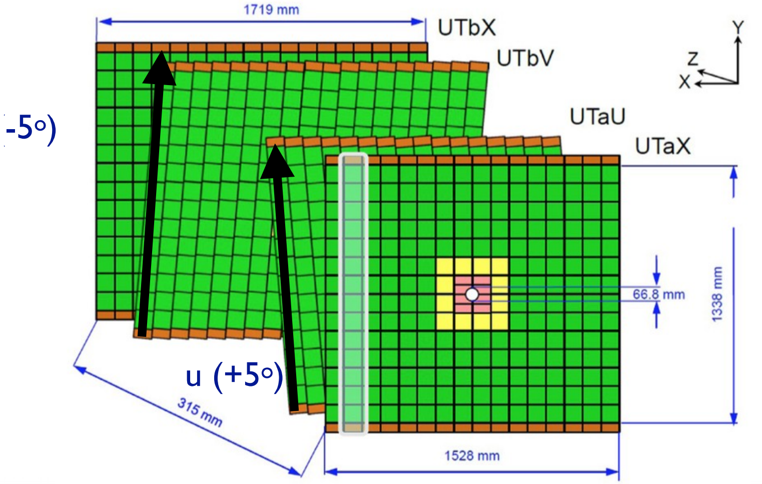
- ☞**Larger acceptance** in central region
- ☞**Reduced thickness**
- ☞**Improved $\sigma_{x-z} \sim 50$ µm**
- ☞**Achievable Δp/p** ~ 15-30%



CERN-LHCC-2013-021



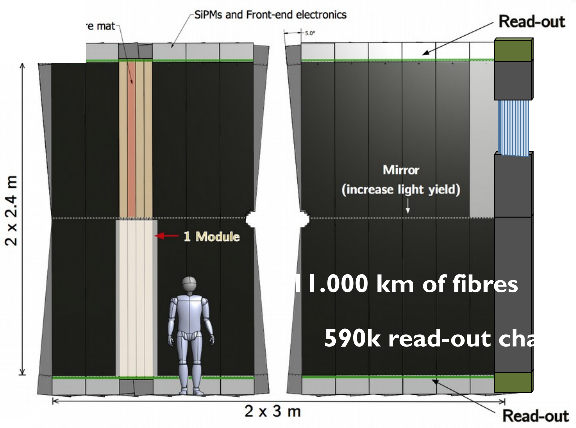
Velo-UT tracking



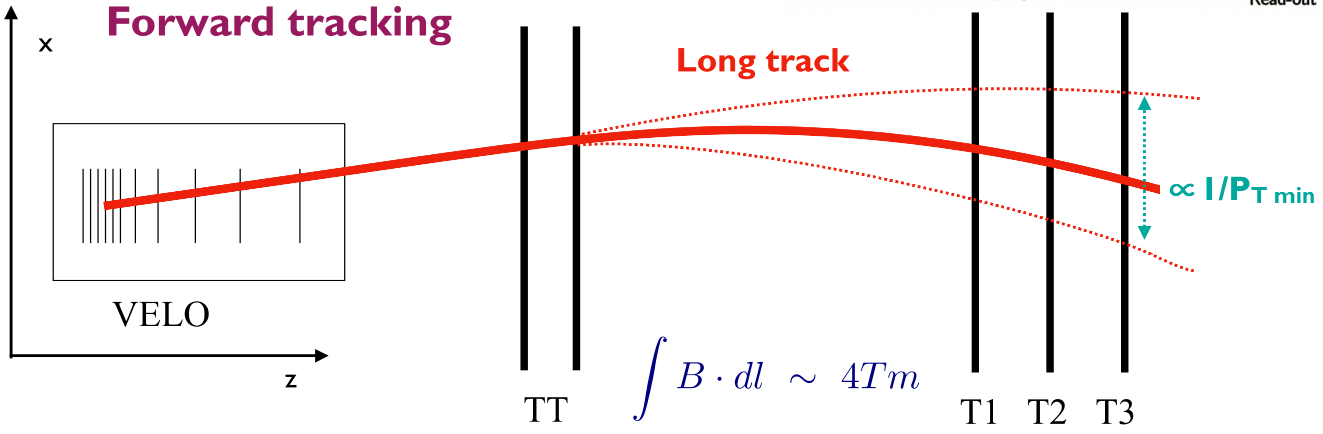
Partial reconstruction sequence : Forward tracking

Scintillating Fibre Tracker (Sci-Fi)

- ☞3 stations x 4 planes (*x*/*u*/*v*/*x*) of 6 stacked 2.4 m long **scintillating fibres (ϕ=**250 μm)
- ☞Read-out by Silicon-Photon multipliers (250 μm channel pitch)

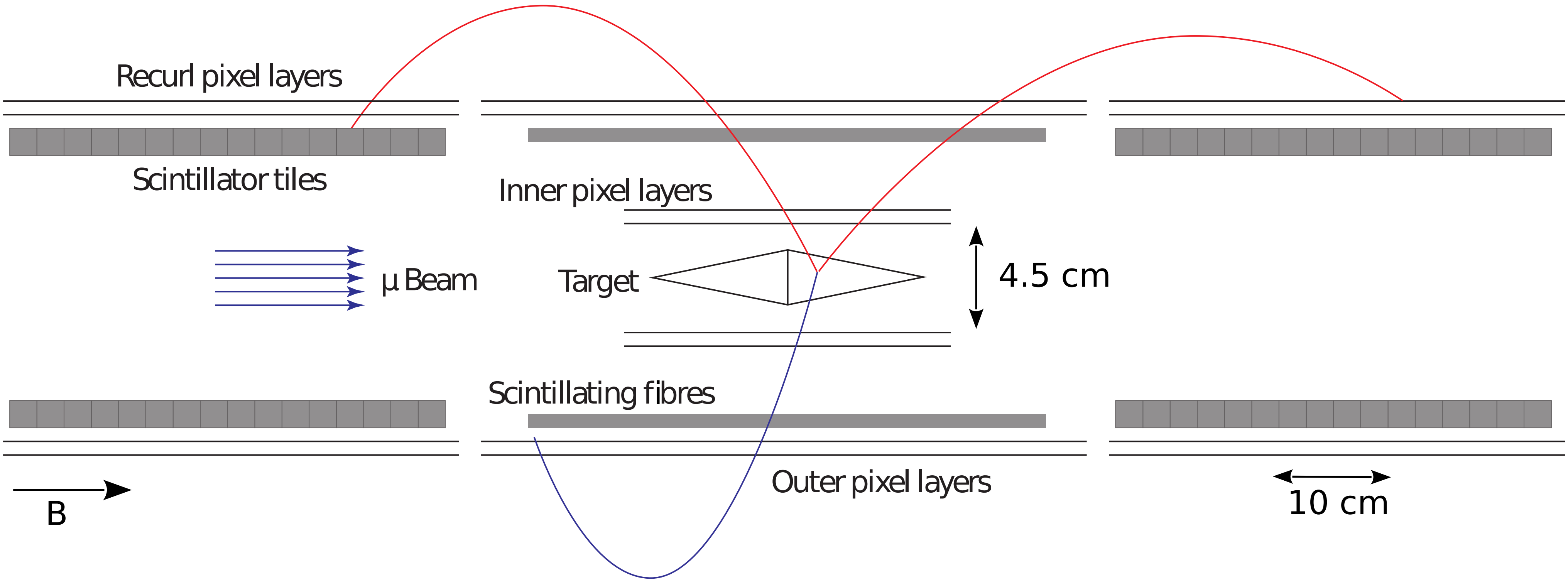


Forward tracking

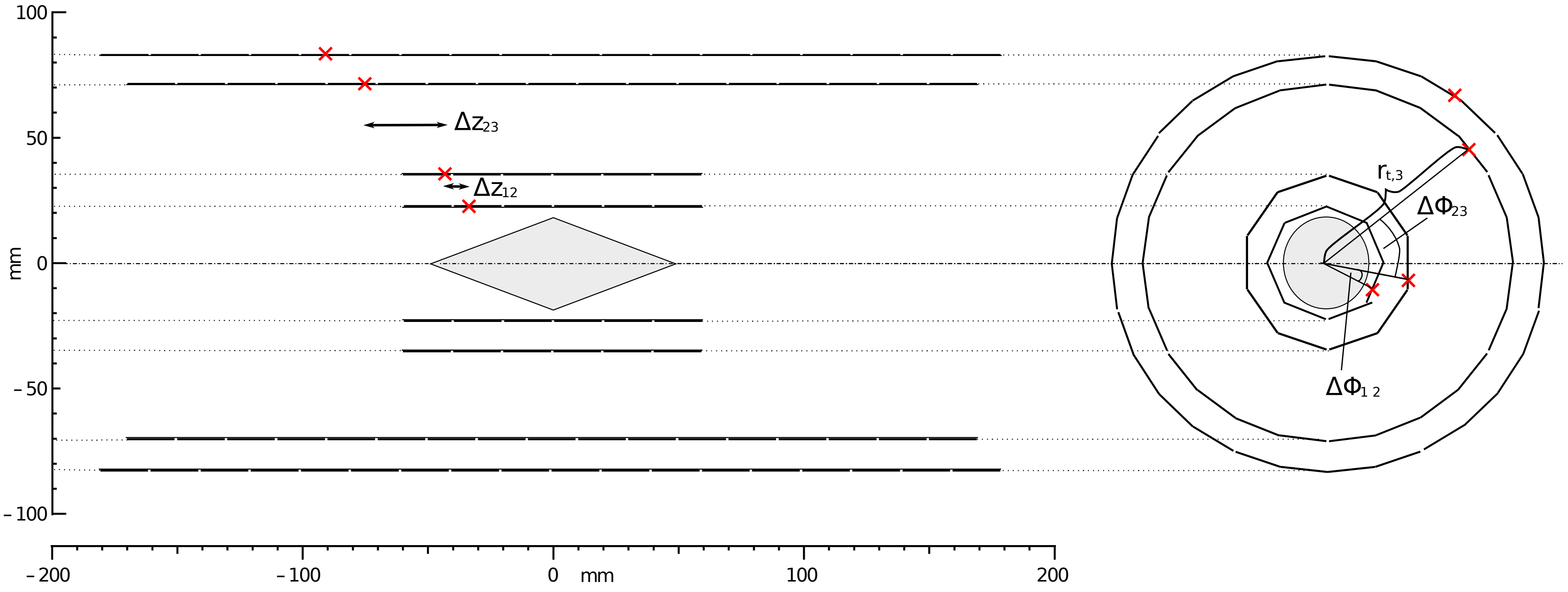


Find matching segments in SciFi according to transverse momentum tolerances

Mu3e: Full Detector

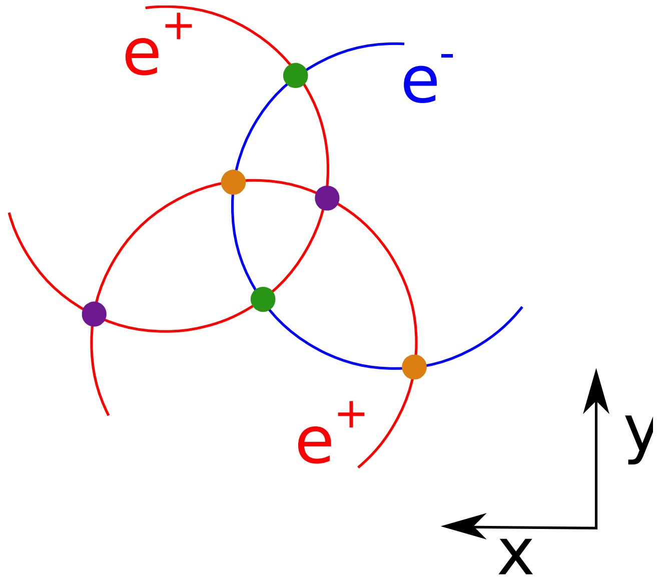


Mu3e: Geometrical selection cuts



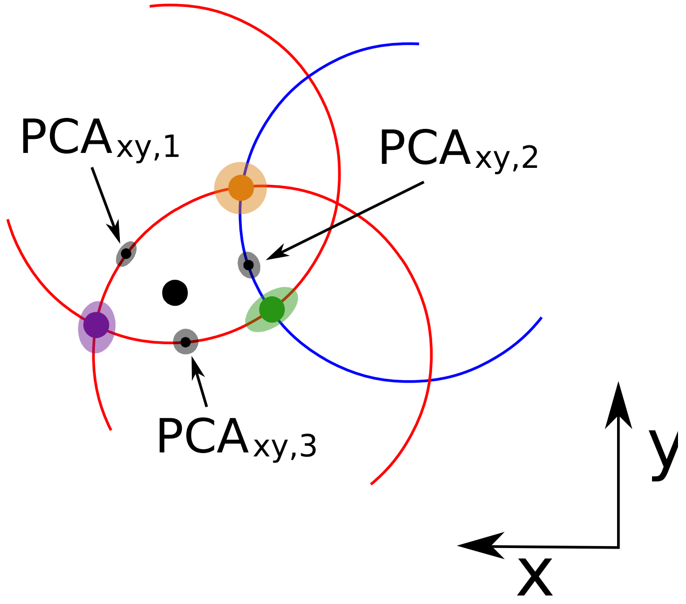
Mu3e: Vertex Estimate

- Study each combination of two e^+, one e^-
- In xy-plane: find intersections of track circles
- Calculate weights of intersections based on uncertainties due to
  - multiple scattering
  - pixel size

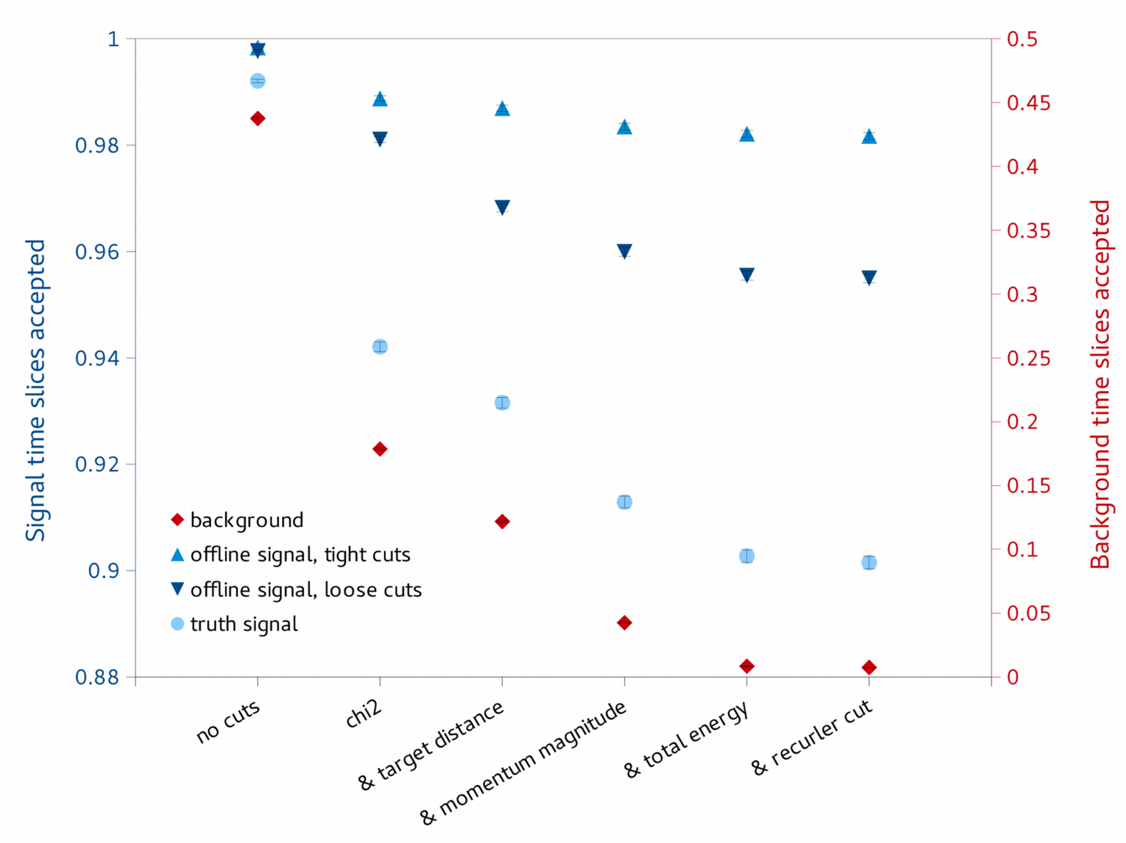


Mu3e: Vertex Estimate

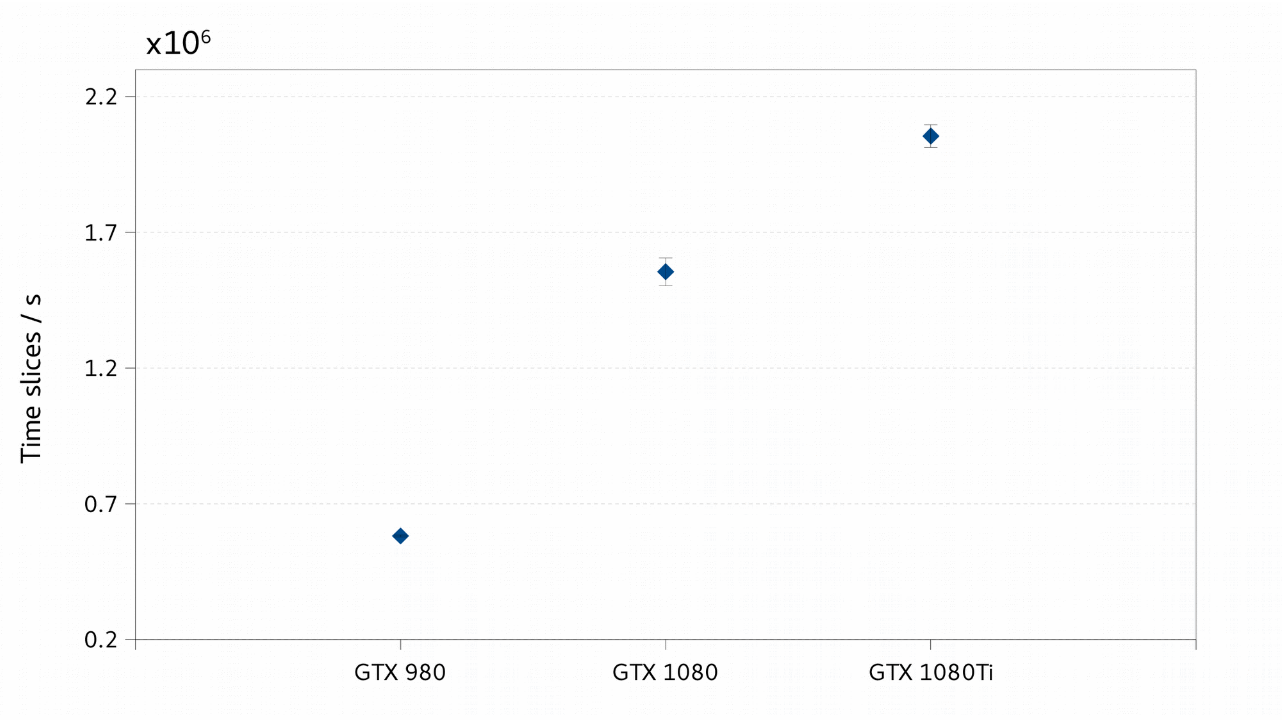
- Calculate weighted mean of intersections from three different tracks
- Find point of closest approach (PCA_{xy}) to weighted mean in xy-plane on each track
- Calculate z-position PCA_z and weight at PCA_{xy}
- Find weighted mean in z-coordinate
- Achieve vertex resolution of ~400 µm sigma



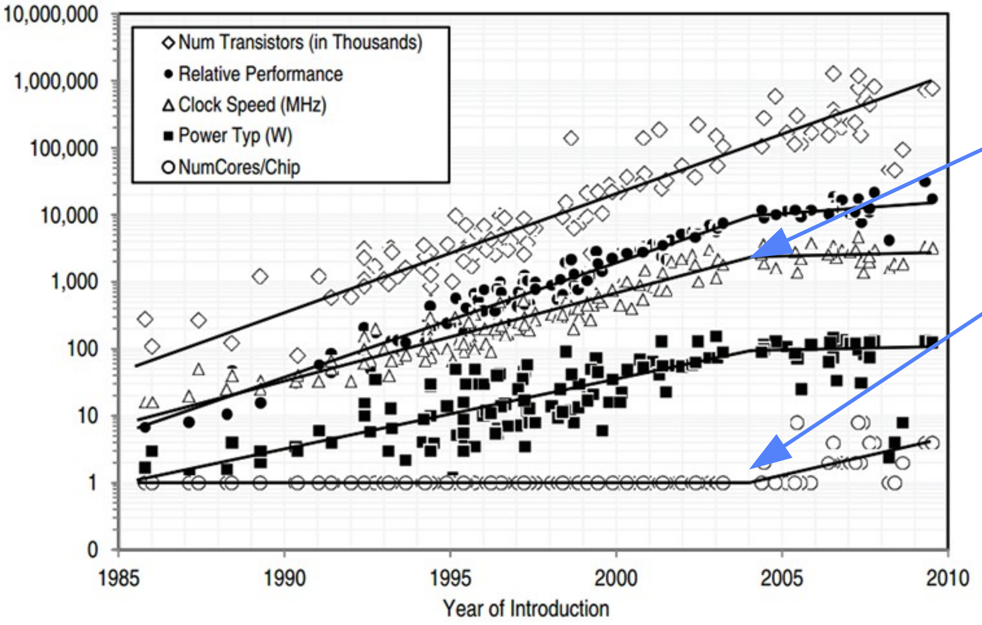
Mu3e: Signal Selection



Mu3e: GPU Generations



Clock speed no longer increasing



https://hackadaycom.files.wordpress.com/2015/09/numtransistors.png

- 2004: clock speed stopped increasing due to heat limit
- → Multiple core processors
  - (Intel i7: 4 cores)
- Next: quantum mechanics limit: O(10 nm)