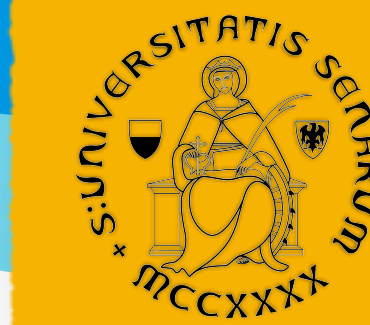
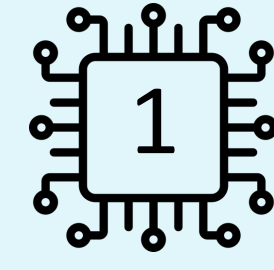


Demonstration of FPGA-based track reconstruction on live LHCb data

Francesco Terzuoli
Università di Siena & INFN - Pisa

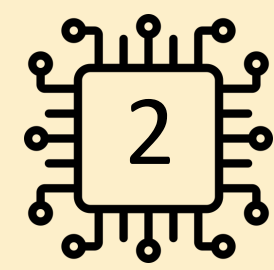


Why FPGAs for track reconstruction?

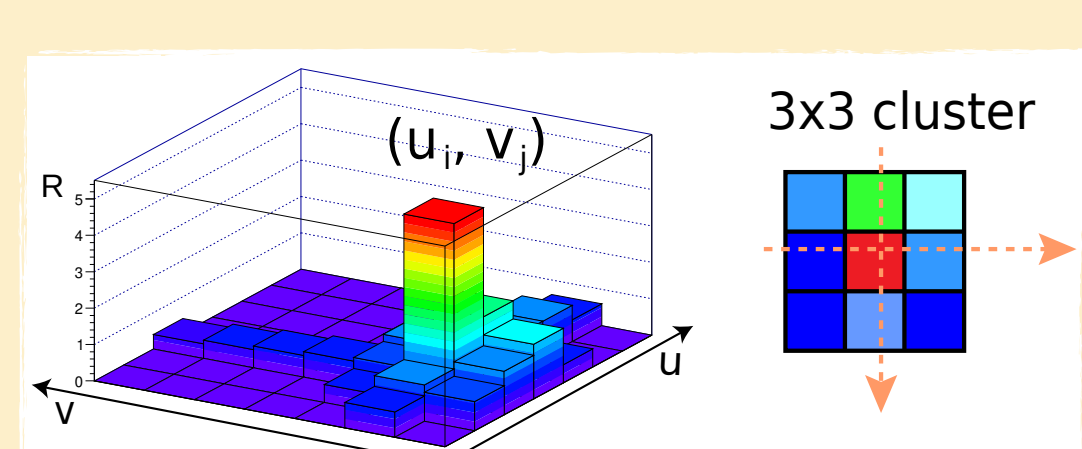
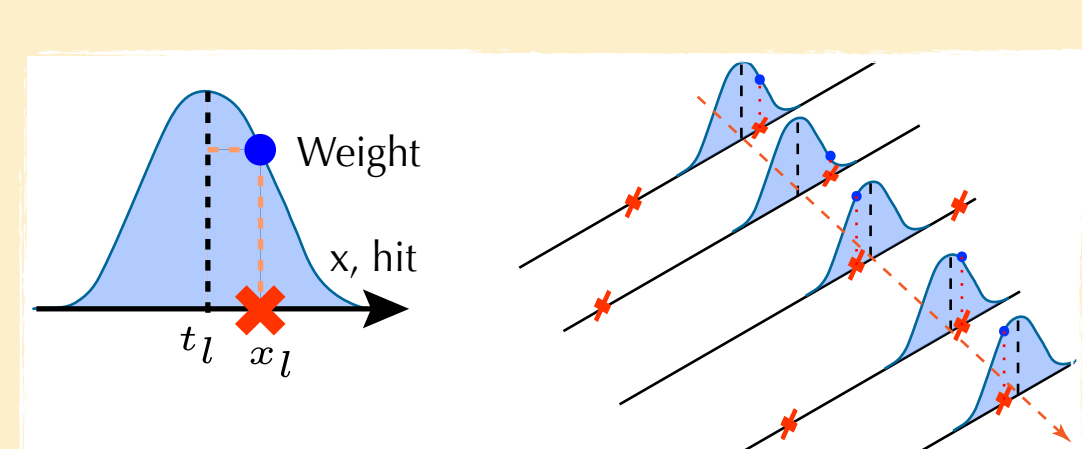
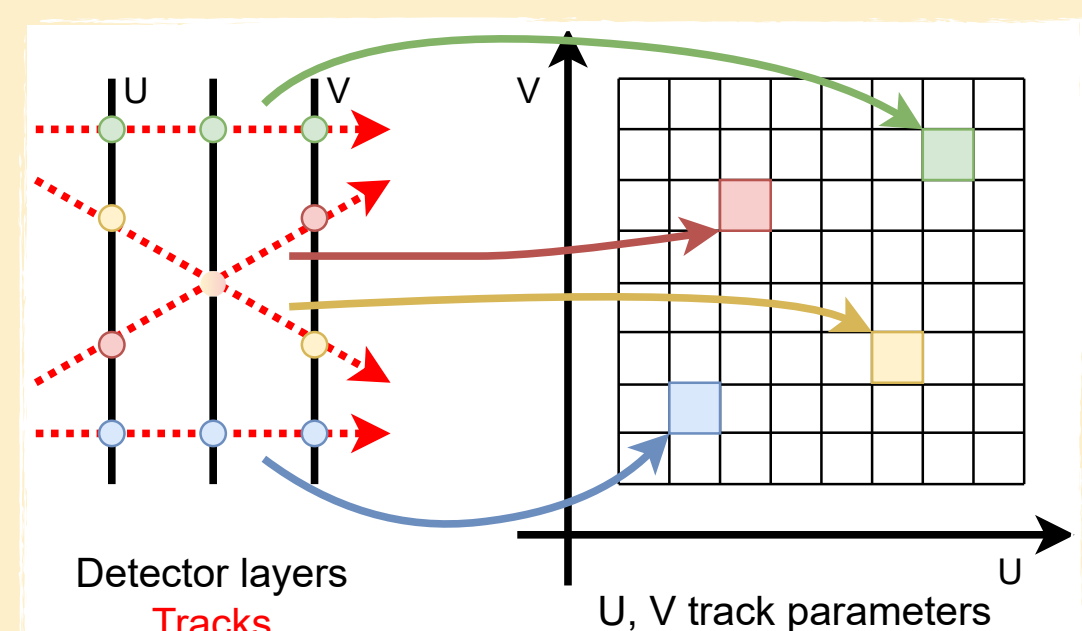


- LHCb reconstructs events at **30 MHz** in Run 3 ($\mathcal{L} = 2 \times 10^{33} \text{cm}^2\text{s}^{-1}$)
 - High Level Trigger: Level 1 (HLT1) on **GPUs** + Level 2 (HLT2) on **CPUs**
- Expected **factor x5-10** in luminosity with **Upgrade II**^[1] presents a challenge
- HEP experiments are seeking **heterogeneous computing solutions** in view of **increasing luminosity** with Moore's law slowing down^[2]
- Modern FPGAs can perform **highly parallel processing** with high throughputs and low latencies
- FPGAs as **greener solution**: less power-hungry than CPUs and GPUs

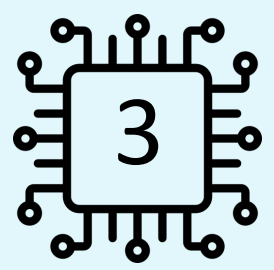
The "Artificial Retina" architecture



- "Artificial Retina" architecture^[3] is a fast implementation of a computation resembling the **Hough transform**^[4] and deployable on FPGAs^[5]
 - Generalised approach (not only straight lines as track model)
 - Numerical evaluation: using a **preset of reference tracks**
 - Continuous** (non binary) **response** quantifying agreement with reference tracks and subsequent **interpolation**
- STEP 1**
 - Track parameter space is discretised as a matrix of **cells**
 - Reference tracks are produced for each cell and their intercepts with detector layers computed (**receptors**)
- STEP 2**
 - For each hit within a **search distance** from each receptor, a weight is computed depending on their distance
 - The weights of receptors belonging to the same cell are summed over
- STEP 3**
 - Tracks as **local maxima** in the cell matrix
 - Interpolating** responses from nearby cells for estimating **real track parameters**

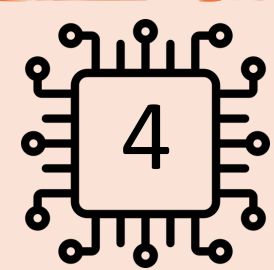


The system implementation



- Each cell of the matrix implements an **engine** for computing the weights
- Engines work in **fully parallel** way, thus FPGA-friendly
- System can be spread over multiple FPGAs boards
 - Distribution network for spreading data across the system
 - Switch for distributing hits to cells where their weight is significant
 - Fast optical network for inter-board communication^[6]

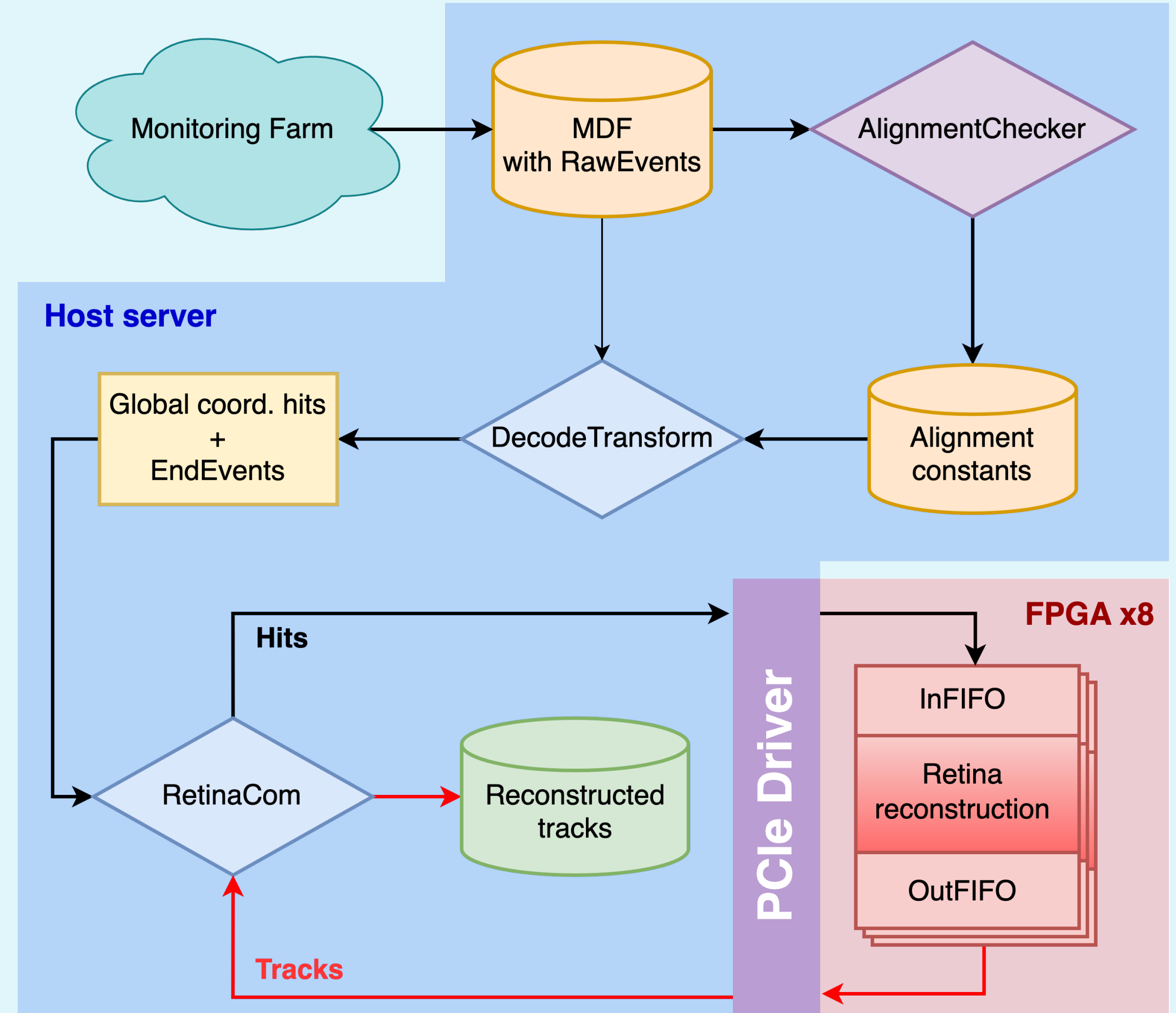
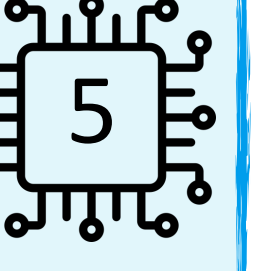
Demonstrator



- Installed at the Coprocessor TestBed facility located at LHCb site
- 8 Intel Stratix 10 FPGA
- LHCb VELO** quadrant coverage
- Full-mesh network** for fast data exchange
 - 28 full-duplex links at 25.8 Gbps
- Engines on different boards cover different parameter space regions

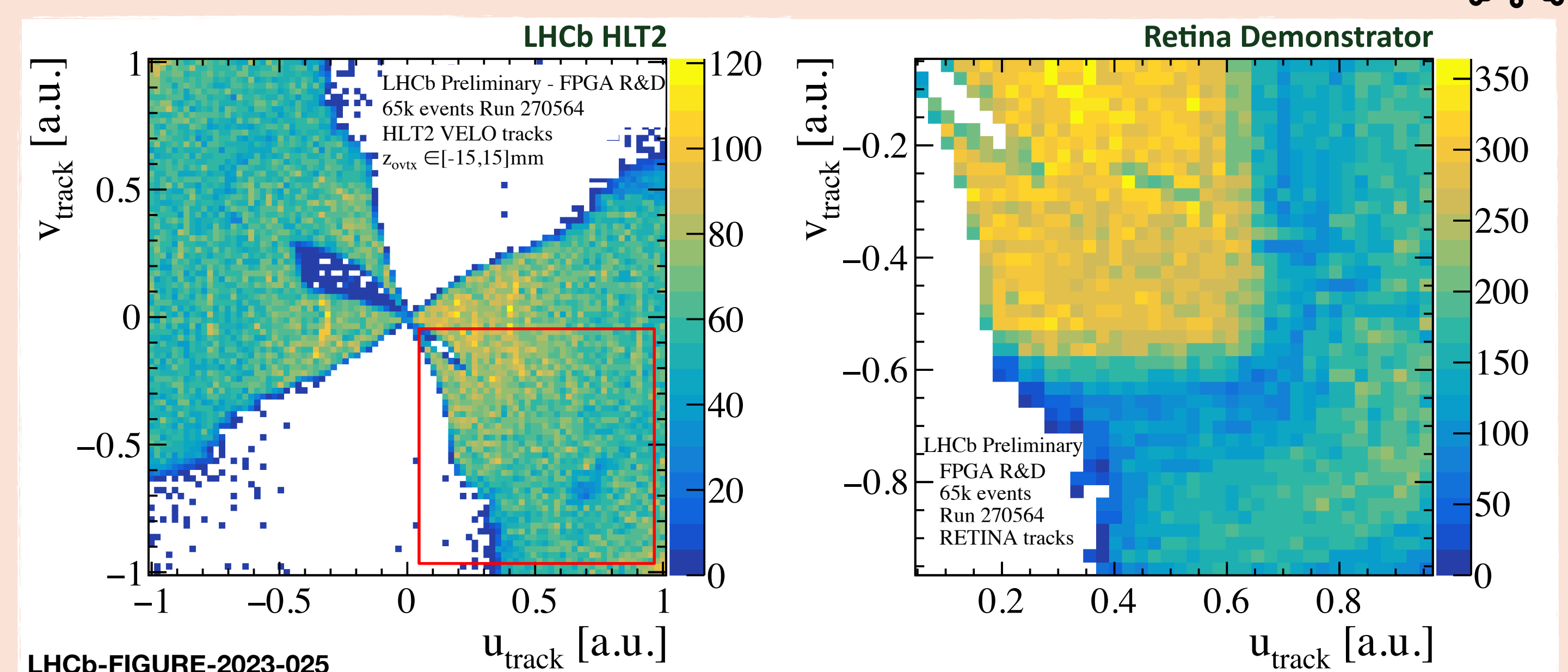
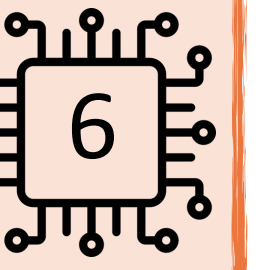


Processing live LHCb data - Data flow



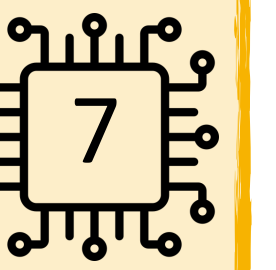
- Deployed custom chain for **feeding data from LHCb DAQ in real-time** to the demonstrator and apply **most recent alignment at every new run**
- Communication with FPGAs via **PCIe**

Processing live LHCb data - Results



- Run more than 30 days of *pp*-collisions data without hiccups with data delivered to the TestBed facility at the event rate of 1 kHz
- Real-time reconstructed tracks** appear **sensible** when their distribution is qualitatively compared with LHCb HLT2 reconstruction
- Retina Architecture demonstrator is able to reconstruct a portion of a detector in real time** with real data coming from the detector DAQ

Additional tests and future prospects



- Official LHCb Monte Carlo simulated** events, with Run 3 conditions ($\mathcal{L} = 2 \times 10^{33} \text{cm}^2\text{s}^{-1}$), are injected in the demonstrator internal RAMs and read in loop in order to maximise input rate
- Comparison between tracks reconstructed by the **hardware** and **custom deployed C++ emulator** shows **exact bitwise adherence** between the two
- About **10 days of running without errors** (much higher than bunch life)

The event throughput has now reached an unprecedented rate of 19 MHz
A final rate of ~ 31 MHz is estimated when all ongoing tunings will be completed

Proposal for a FPGA-based downstream tracking in Run 4 is currently under review by the LHCb collaboration

These results are achieved thanks to the funding from Italian INFN and the kind support of the RTA and Online groups in LHCb

[1] LHCb Collaboration, "Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era", CERN-LHCC-2017-003 (2017)

[2] T. Theis and H. Wong, "The End of Moore's Law: A New Beginning for Information Technology", Computing In Science and Engineering, 19, 41-50 (2017)

[3] L. Ristori, "An artificial retina for fast track finding", NIMA 453, 1, 425-429 (2000)

[4] P. Hough, "Analysis Of Bubble Chamber Pictures", Proc. Int. Conf. High Energy Accelerators and Instrumentation C590914, 554-558 (1959)

[5] G. Tuci and G. Punzi, "Reconstruction of track candidates at the LHC crossing rate using FPGAs", EPI Web Conf. 245, 10001 (2020)

[6] F. Lazzari et al., "FPGA-based real-time data processing for accelerating reconstruction at LHCb", Journal of Instrumentation 17, 4, C04011, (2022)