CompassUT: study of a GPU track reconstruction for LHCb upgrades

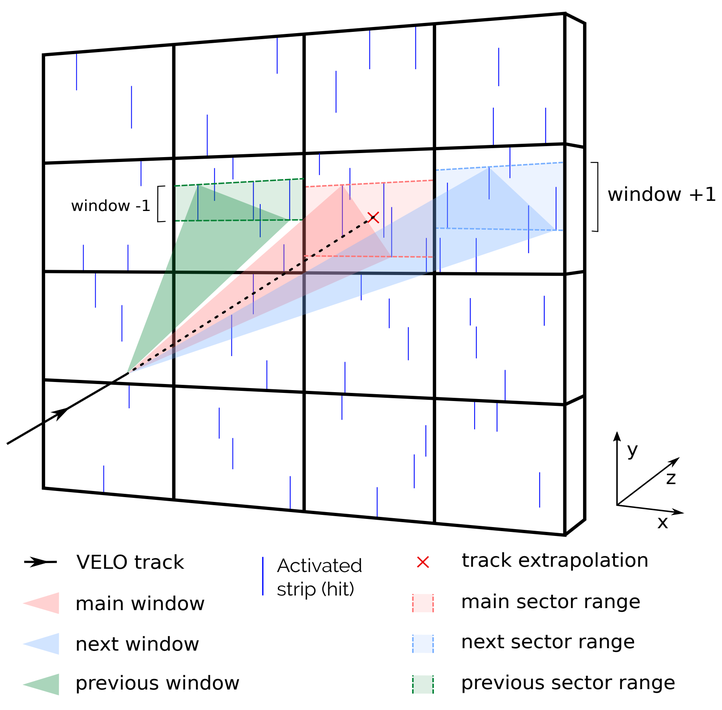
Placido Fernandez Declara on behalf of the LHCb collaboration
University Carlos III of Madrid and CERN

Abstract

We present a fast, data-oriented GPU tracking algorithm, CompassUT, as a potential option to cope with the expected throughput of 40Tbit/s for LHCb upgrade.

- The raw input data is decoded in parallel, optimized for GPU architecture.
- Hits are sorted by X and Y into sector groups
- The tracking is implemented for minimal memory footprint, reduced branching, and data oriented.
- The resulting performance increase is shown for 3 different GPUs with the impact on physics performance for different configurations.

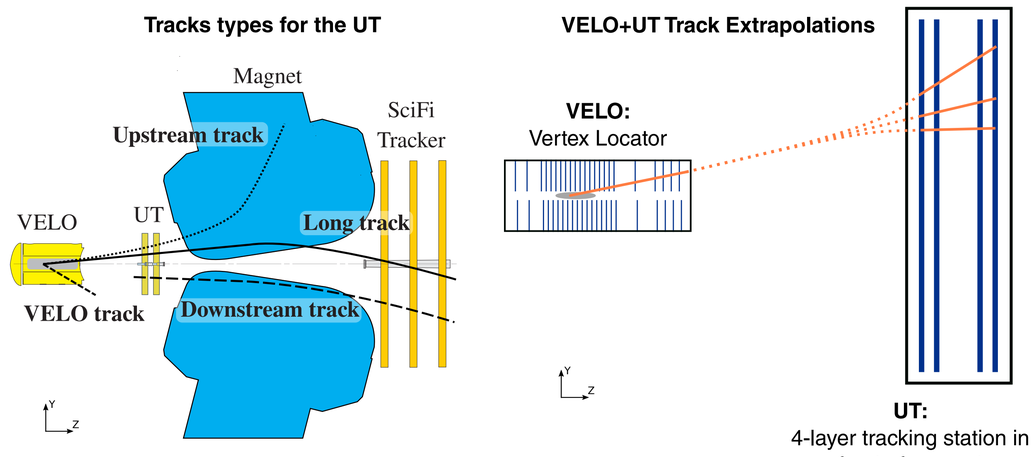
UT plane window ranges



UT tracking

UT tracking:

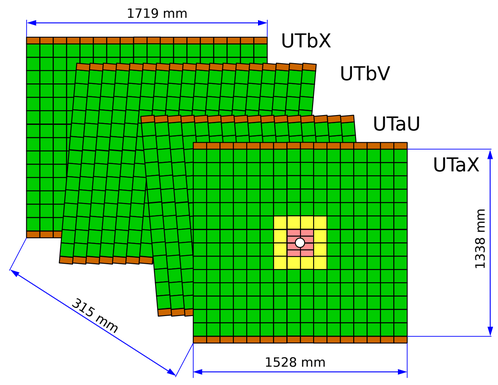
- Reconstruct charged particle trajectories that cross the UT
- Match a VELO track with UT hits to create a VELO+UT track
- Tracks are slightly bent by the magnetic field, introducing more multiplicity to the calculations.



UT decoding and CompassUT tracking

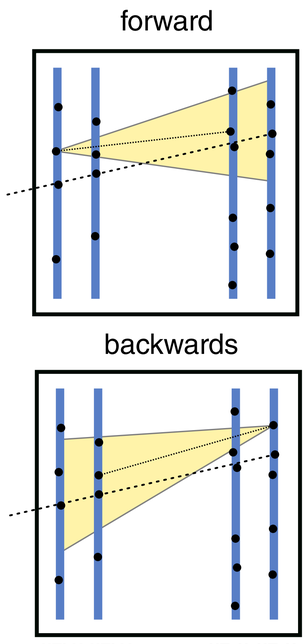
UT decoding (5 kernels):

1. Calculate number of hits to pre-allocate memory
2. Get offsets for efficient access using prefix sum
3. Sort hits into X regions defined by sectors
4. Calculate the permutations needed to sort by Y
5. Decode and sort by Y into sector groups

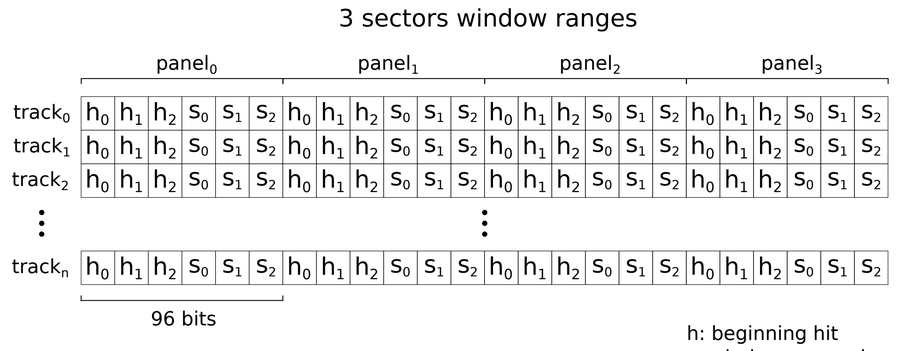


CompassUT tracking:

- Binary search the regions extrapolating VELO track
- Calculate all window ranges for N sectors
- Minimize branching, filter non-valid tracks with shared memory, cache window ranges into shared memory
- Search for triplet/quadruplet with combined forward-backward loop
- Look for best cluster based on configured parameter



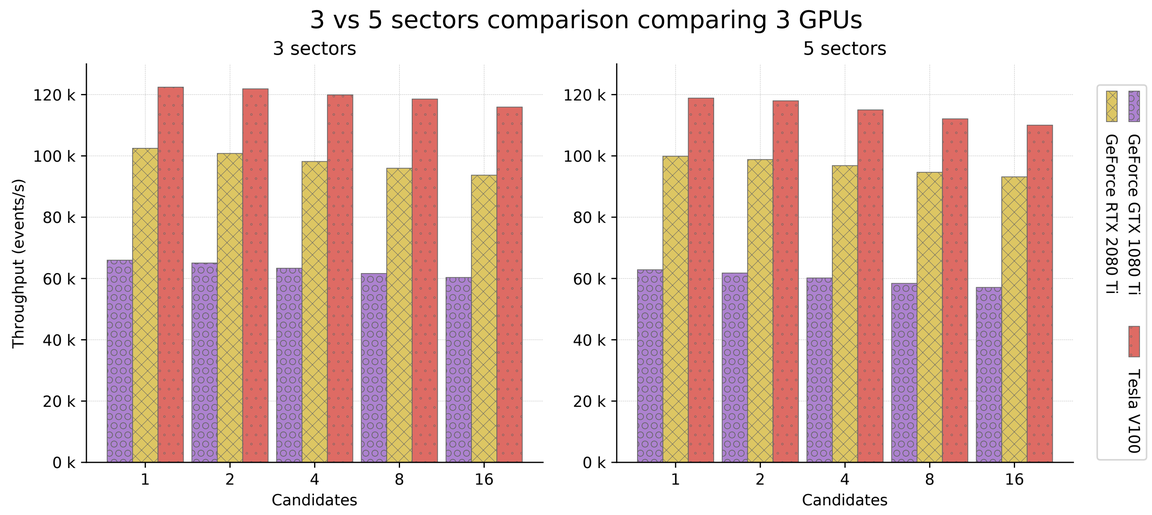
3 sectors window ranges



h: beginning hit
s: window range size

Physics and computing performance results

Number of sectors	Number of candidates	Long tracks reco. efficiency	Long tracks clone rate	Fake rate
1 sector	2	79.70%	0.35%	10.09%
	4	80.52%	0.32%	9.83%
	16	80.99%	0.34%	9.48%
3 sectors	2	93.67%	0.37%	9.89%
	4	94.66%	0.36%	9.59%
	16	95.21%	0.38%	9.22%
5 sectors	2	94.20%	0.37%	9.90%
	4	95.19%	0.36%	9.60%
	16	95.74%	0.37%	9.23%



The impact on the physics performance is shown here. Long tracks achieve > 95% reconstruction efficiency with low clone and fake rate. The version using just 1 sector is discarded due to low physics performance. The computing performance differences is greater between the consumer cards 1080 Ti and 2080 Ti, compared to the server-grade V100 despite their price differences. Several configurations between 3 and 5 sectors offer similar performance with comparable physics efficiency.

References

[1] LHCb collaboration, Framework TDR for the LHCb Upgrade: Technical Design Report, CERN-LHCC-2012-007. LHCb-TDR-12, 2012.
[2] E. Bowen, B. Storaci, VeloUT tracking for the LHCb Upgrade, CERN-LHCb-PUB-2013-023. LHCb-INT-2013-056, CERN, Geneva, 2014.
[3] LHCb Collaboration, Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era, Technical Report CERN-LHCC-2017-003, CERN, Geneva, 2017.