Cherenkov angle reconstruction with FPGA compute accelerators

Christian Färber
CERN Openlab Fellow
LHCb Online group

On behalf of the LHCb Online group and the HTC Collaboration

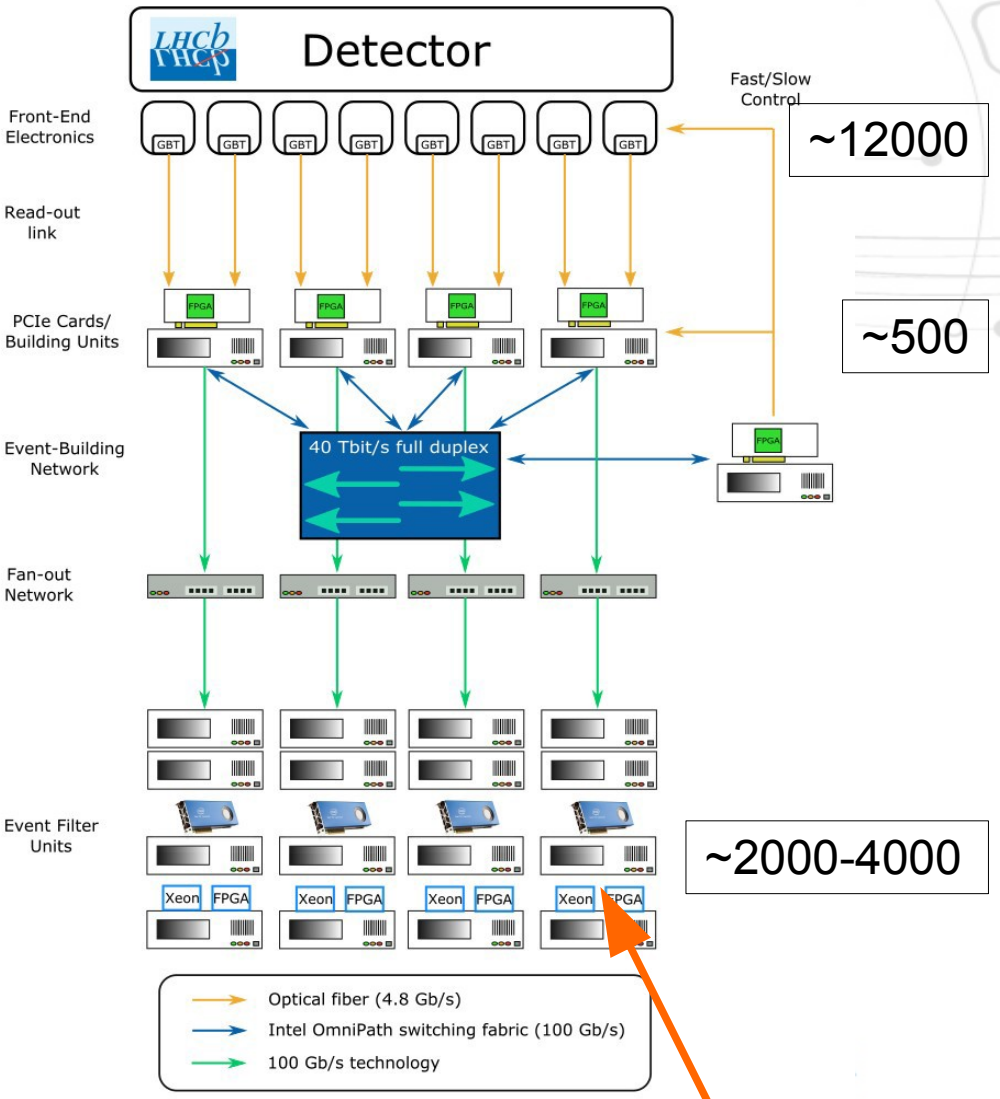
7th LHCb Computing Workshop
02.06.2016

Upgrade Readout Schematic

- Raw data input ~ 40 Tbit/s
- EFF needs fast processing of trigger algorithms, different technologies are explored.

- Test FPGA compute accelerators for the usage in:
  - Event building
    - Decompressing and re-formatting packed binary data from detector
  - Event filtering
    - Tracking
    - Particle identification

- Compare with: GPUs, Intel® Xeon/Phi and other compute accelerators



Which technologies?

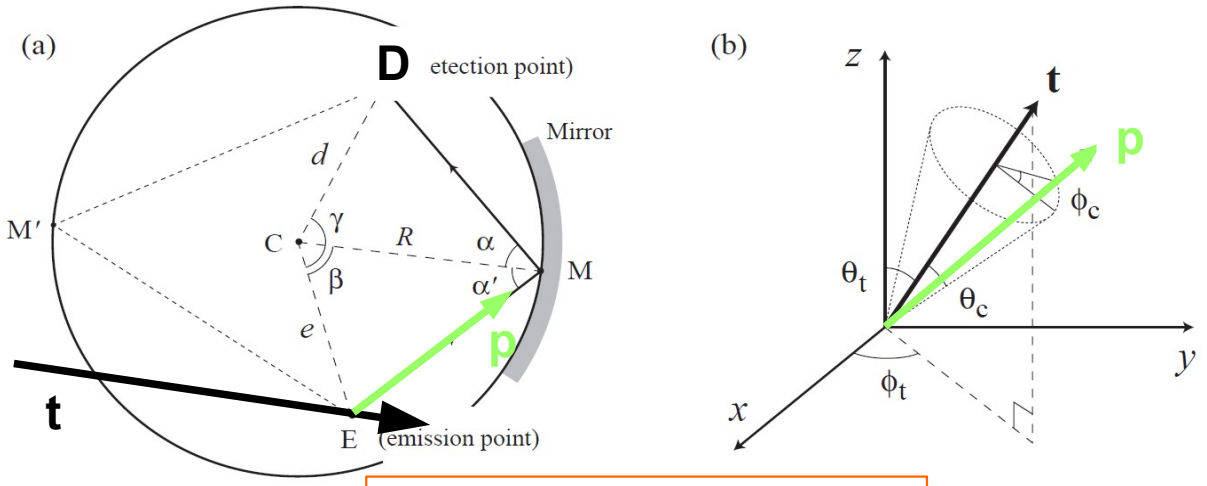
FPGAs as Compute accelerators

- Microsoft Catapult and Bing
  - Improve performance, reduce power consumption

- LHCb: Test for future usage in upgraded HLT farm:
  - Event building
  - Track fitting, pattern recognition, PID algorithms
- Current Test Devices in LHCb
  - Nallatech PCIe with OpenCL
  - Intel® Xeon/FPGA

Test case: RICH PID Algorithm

- Calculate Cherenkov angle Θ_c for each track **t** and detection point **D**
- RICH PID is not processed for every event, processing time too long!

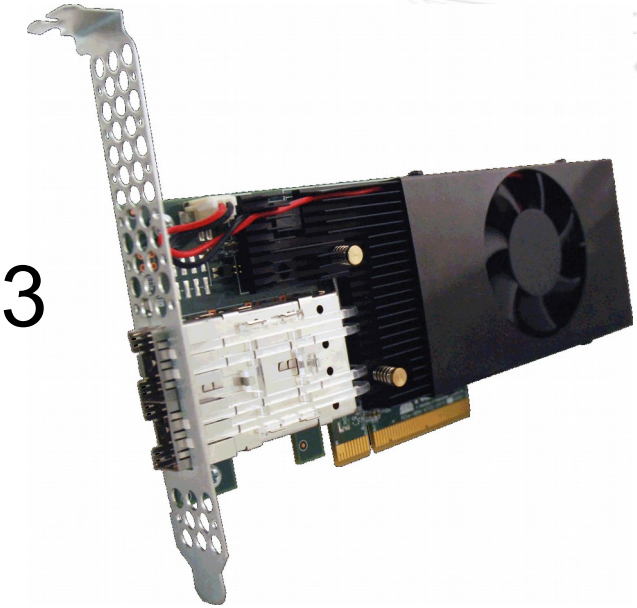


Reference: LHCb Note LHCb-98-040

SVN:
Rec/trunk/Rich/RichRecPhotonTools/src/RichRecPhotonTools/QuarticSolver.h@183637

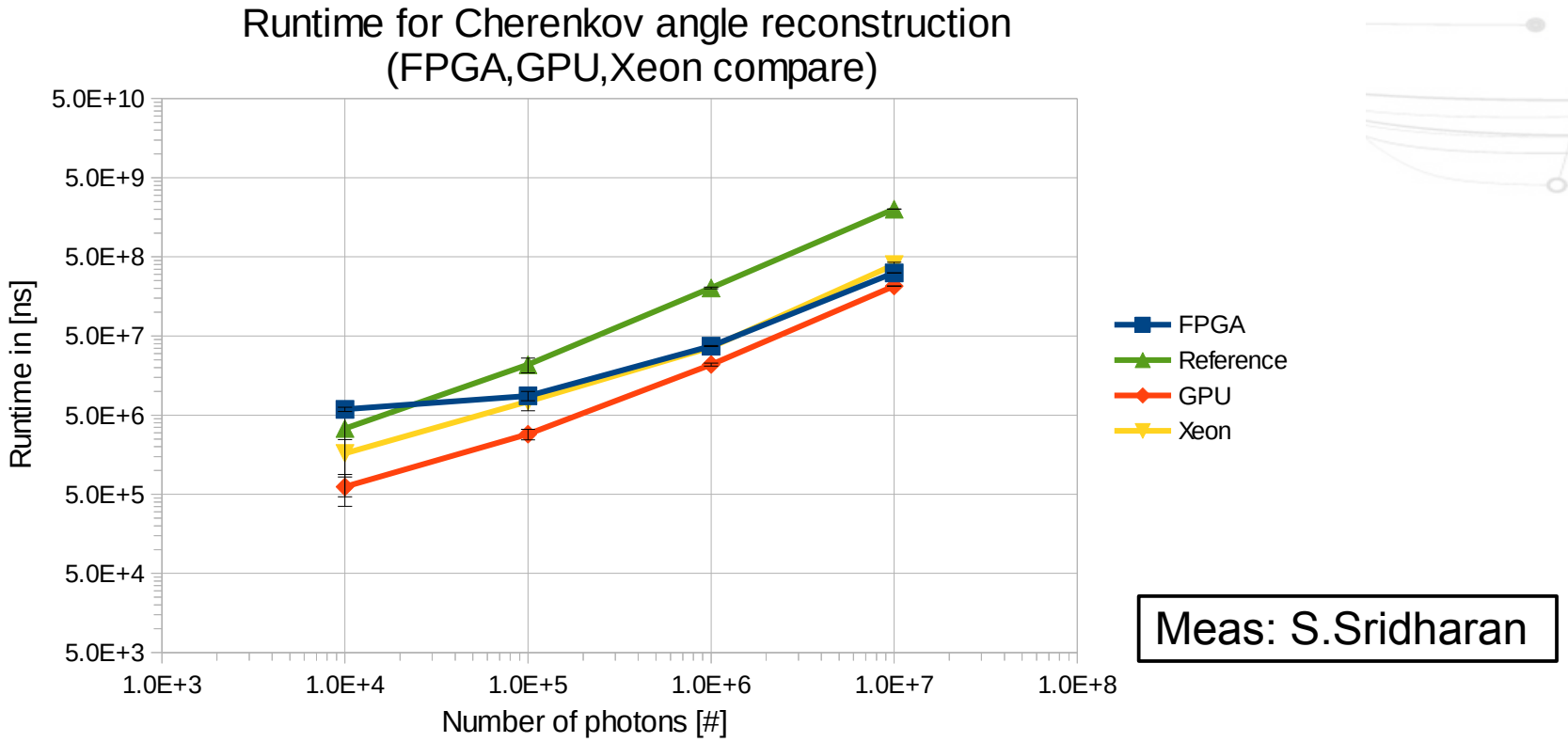
Nallatech 385 Board

- FPGA: Altera Stratix V GX A7
  - 234'720 ALMs, 940'000 Registers
  - 256 DSPs
- Programming model : OpenCL
- Host Interface: 8-lane PCIe Gen3
  - Up to 7.5GB/s
- Memory: 8GB DDR3 SDRAM
- Network Enabled with (2) SFP+ 10GbE ports
- Power usage: ≤ 25W (GPU up to 300W)



Nallatech 385 Board Results I

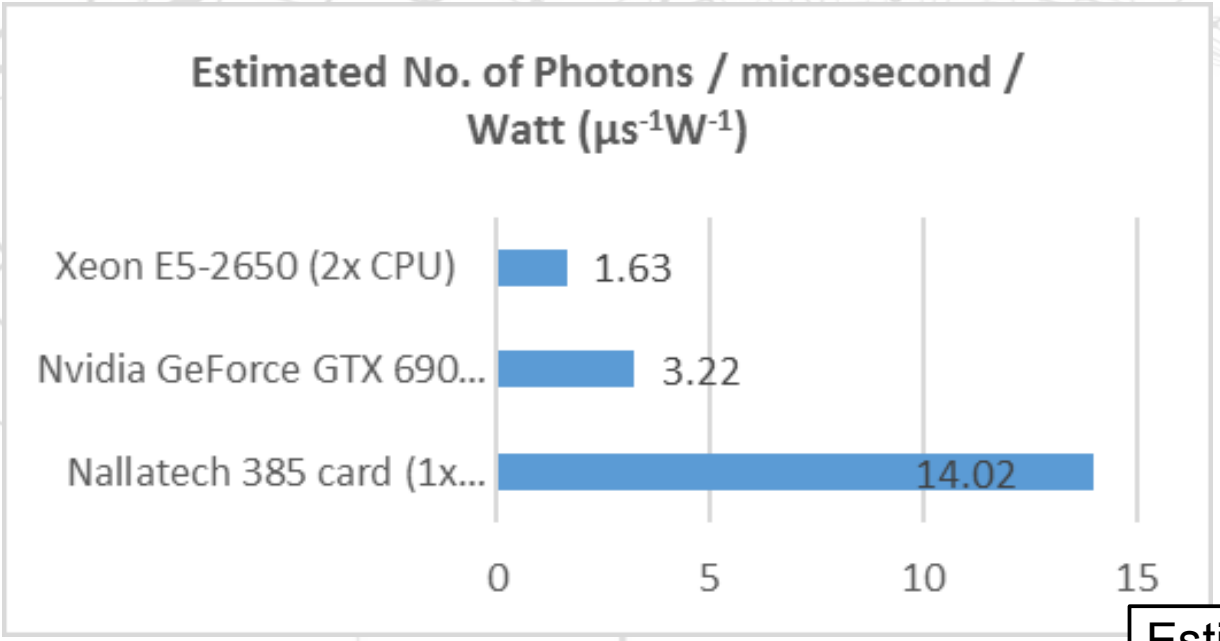
- Performance reference:
  - Intel Core i7-4770 CPU single thread vectorized



- Acceleration of factor up to 6 with Nallatech 385
- FPGA kernel faster, bottleneck data transfer

Nallatech 385 Board Results II

- Energy efficiency comparison of three devices



Estimate: S.Sridharan

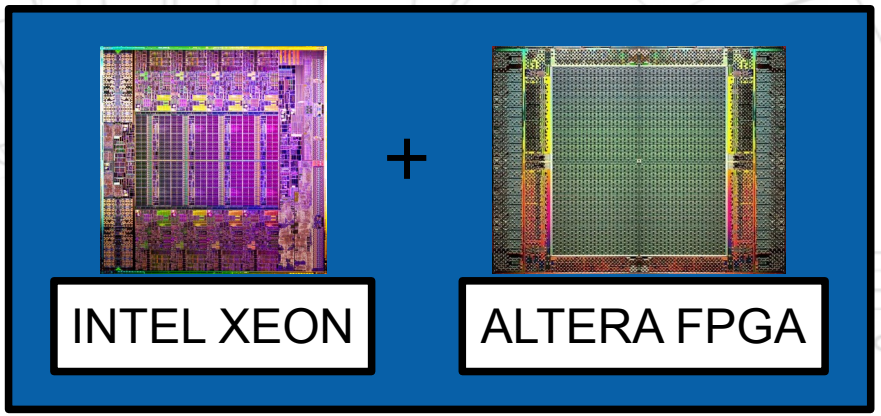
- It is estimated that the FPGA accelerator is a factor 4.3 more energy efficient than the GPU
  - Power measurements will follow!

Intel® Xeon/FPGA

- Two socket system:

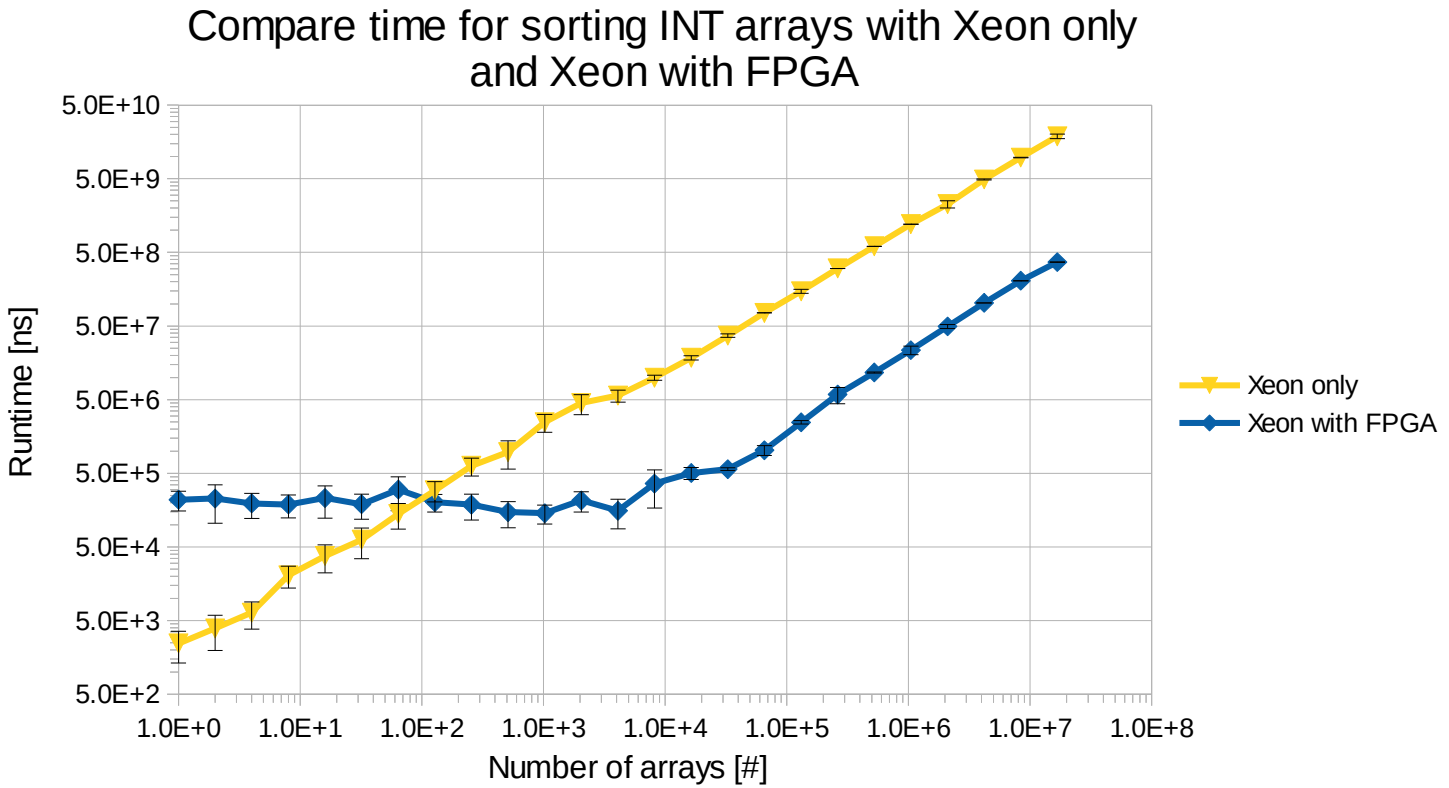
  First: Intel® Xeon®
  E5-2680 v2

  Second: Altera Stratix V GX A7 FPGA
    - 234'720 ALMs, 940'000 Registers, 256 DSPs
- Host Interface: high-bandwidth and low latency
- Memory: Cache-coherent access to main memory
- Programming model : Verilog now also OpenCL
- Power usage: Will be tested with new version



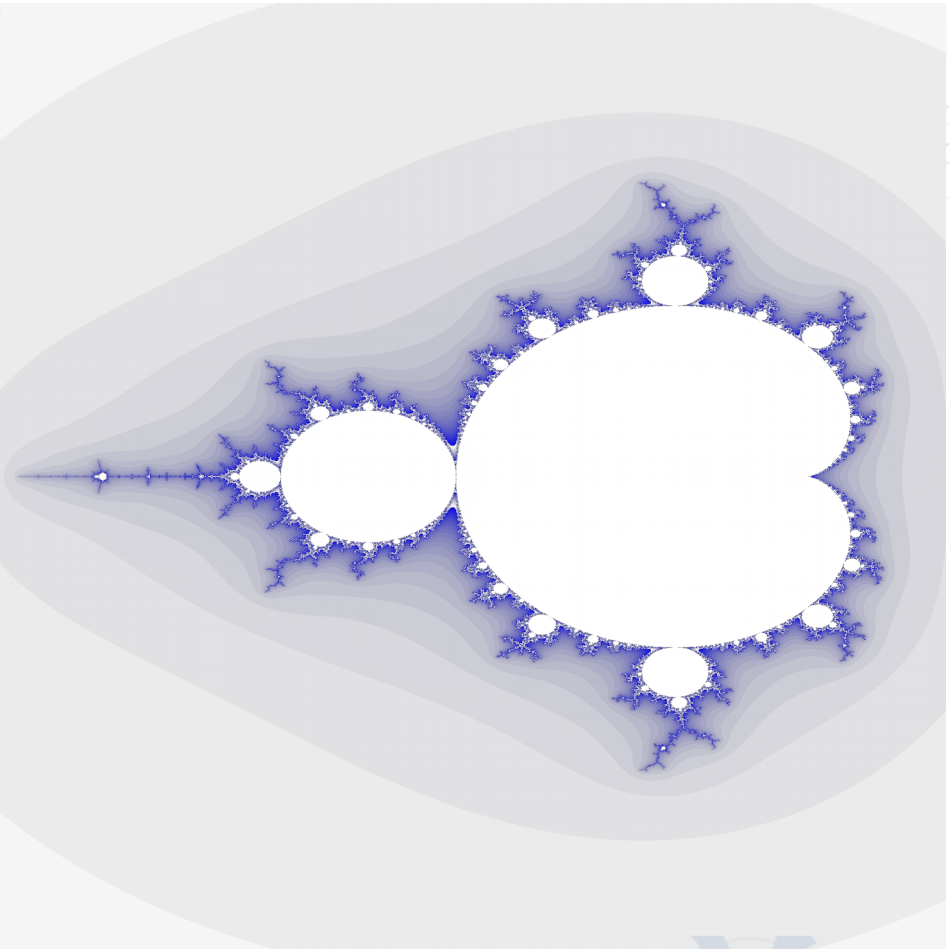
First results with Xeon/FPGA I

- Sorting of INT arrays with 32 elements
  - Implemented pipeline with 32 array stages
  - FPGA sort is x50 faster than single Intel® Xeon thread



First results with Xeon/FPGA II

- Mandelbrot with floating point precision
  - Implemented 22 fpMandel pipelines running at 200MHz, each handles 16 pixels in parallel (total: 352 pixels).
  - FPGA is x12 faster as Xeon running 20 threads in parallel.
  - Used 72/256 DSPs
  - Reuse of data on FPGA high

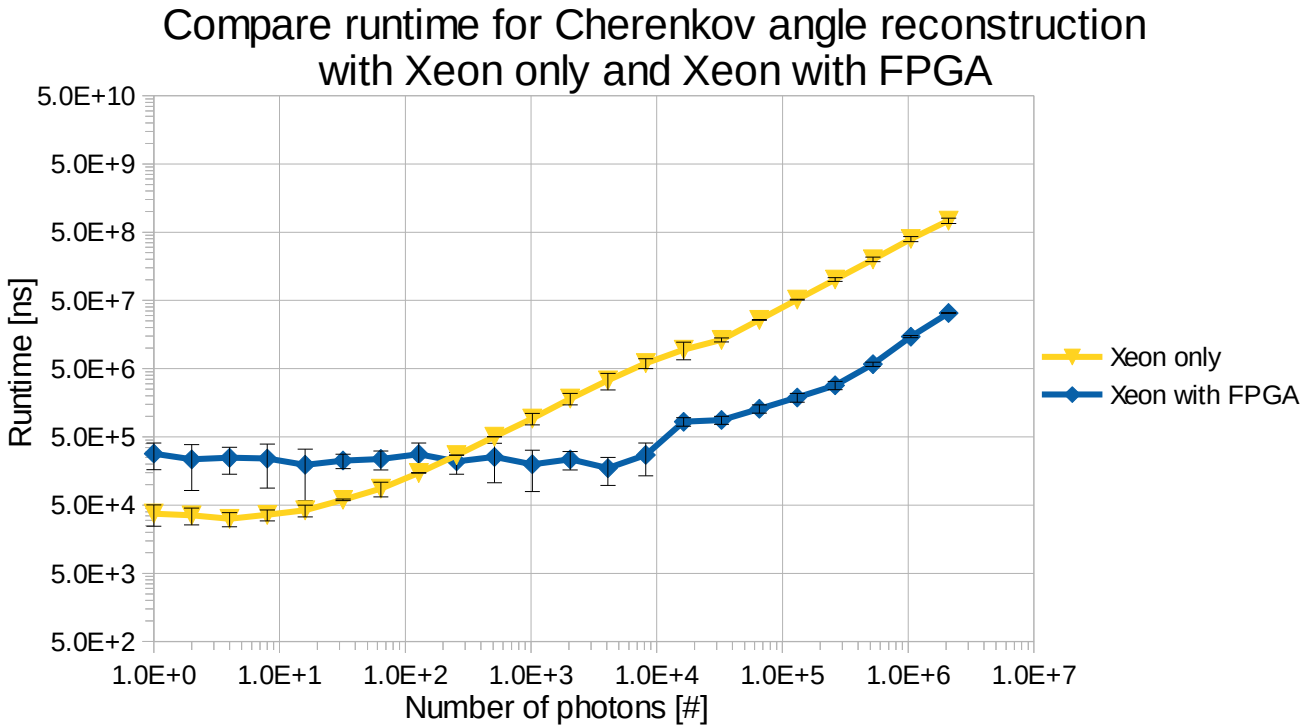


Implementation of Cherenkov Angle reconstruction

- 748 clock cycle long pipeline written in Verilog
  - Additional blocks developed: cubic root, complex square root, rot. matrix, cross/scalar product,...
  - Lengthy task in Verilog with all test benches (implementation took 2.5 months)
- Pipeline running with 200MHz → 5ns per photon
- FPGA resources:

FPGA Resource Type	FPGA Resources used [%]	For Interface used [%]
ALMs	88	30
DSPs	67	0
Registers	48	5

Intel® Xeon/FPGA Results



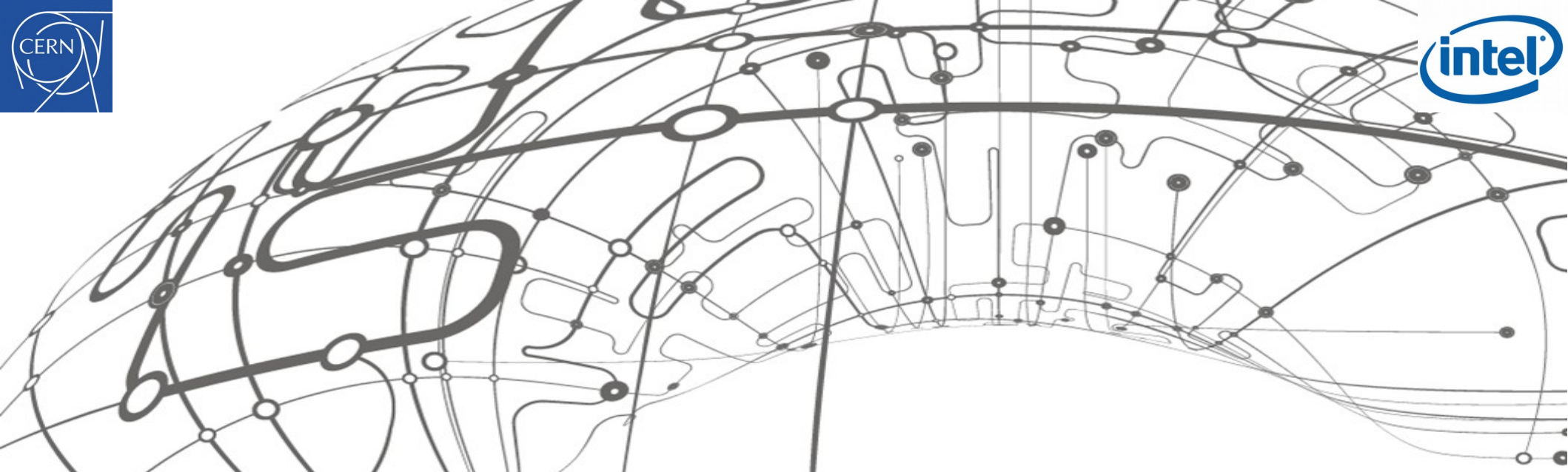
- Acceleration of factor up to 35 with Intel® Xeon/FPGA with respect to single Intel Xeon thread, using multi Xeon threads (OpenMP) the acceleration is a factor 11 (prelimintary)
- Theoretical limit of photon pipeline: a factor 64 with respect to single Intel Xeon thread
- Bottleneck: Data transfer bandwidth to FPGA

Future Tests

- Implement additional LHCb HLT algorithms
  - Tracking, decompressing and re-formatting packed binary data from detector, ...
- Compare performance with new Intel® Xeon/FPGA system with Arria 10 FPGA
  - Hardened floating point mult/accumulate blocks
- Test Nallatech CAPI (cache-coherent)
- Compare Verilog - OpenCL AFUs
- Power measurements
  - Compare with GPUs

Summary

- Results are very encouraging to use FPGA acceleration in the HEP field
- Intel® Xeon/FPGA accelerator performs better than the Nallatech PCIe board using the same FPGA
- FPGAs are strong in performance per Watt
- Programming model with OpenCL very attractive
  - Faster and easier algorithm implementation
- Test Intel® Xeon/FPGA with Arria 10
  - Larger, faster and higher bandwidth to FPGA



Thank you

Christian Färber,

Backup

HTCC

- High Throughput Computing Collaboration
- Members from Intel and CERN LHCb/IT
- Test Intel technology for the usage in trigger and data acquisition (TDAQ) systems
- Projects
  - Intel® Omni-Path 100 Gbit/s network
  - Intel® Xeon/Phi computing accelerator
  - Intel® Xeon/FPGA computing accelerator