BoF: Reconfigurable Computing Architectures for HPC & HPDA

Acceleration of High Energy Physics Algorithms used for the LHCb Upgrade with the new Intel® Xeon®+FPGA

Dr. Christian Färber
CERN Openlab Fellow
LHCb Online group

On behalf of the LHCb Online group and the HTC Collaboration

ISC 2017, Frankfurt
20.06.2017
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® KNL computing accelerator
- Intel® Omni-Path Architecture 100 Gbit/s network
- Intel® Xeon®+FPGA computing accelerator
Detector Example: LHCb

- Single-arm spectrometer designed to search new physics through measuring CP violation and rare decays of heavy flavour mesons.
- 40 MHz proton-proton collisions
- Trigger with 1 MHz, upgrade to 40 MHz
- Bandwidth after upgrade up to 40 TBit/s
Future Challenges

- Higher luminosity from LHC
- Upgraded sub-detector Front-Ends
- Removal of hardware trigger
- Software trigger has to handle:
  - Larger event size (50KB to 100KB)
  - Larger event rate (1MHz to 40MHz)
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 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?
Intel® Xeon®+FPGA with Stratix® V FPGA

- Two socket system:
  First: Intel® Xeon® E5-2680 v2
  Second: Intel® 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 and OpenCL
LHCb Calorimeter Raw Data Decoding

• Raw data decoding needs many bit shifts and bit manipulations to convert compact ADC values
• On FPGAs this can be realized more efficiently
• Measured acceleration with Stratix® V: x25

![Graph showing performance comparison between Xeon® single thread, Xeon® hyper-threading, and Intel® Xeon®+FPGA]

Tested with 20,000 events
- Xeon Single Thread => 2585 ms
- Xeon all cores => 517 ms
- Xeon + FPGA => 98 ms
Intel® Xeon®+FPGA with Arria® 10 FPGA

- Multi-chip package including:
  - Intel® Xeon® E5-2600 v4
  - Intel® Arria® 10 GX 1150 FPGA
    - 427'200 ALMs, 1'708'800 Registers, 1'518 DSPs
- Hardened floating point add/mult blocks (HFB)
- Host Interface: Bandwidth 5x higher than Stratix® V version
- Memory: Cache-coherent access to main memory
- Programming model: Verilog soon also OpenCL
Test Case: RICH PID Algorithm

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

Calculations:
- solve quartic equation
- cube root
- complex square root
- rotation matrix
- scalar/cross products

Reference: LHCb Note LHCb-98-040
Intel® Xeon®+FPGA Results

Compare runtime for Cherenkov angle reconstruction with Intel® Xeon® CPU and Intel® Xeon®+FPGA

- Acceleration of up to factor 35 with Intel® Xeon®+FPGA
- Theoretical limit of photon pipeline: a factor 64 with respect to single Intel® Xeon® thread, for Arria® 10 a factor ~ 300
- Bottleneck: Data transfer bandwidth to FPGA, caching can improve this, tests ongoing
Compare Verilog - OpenCL

- Development time
  - 2.5 months -- 2 weeks
  - 3400 lines Verilog -- 250 lines C

- Performance
  - Cube root: x35 -- x30
  - RICH: x35 -- x26

- FPGA resource usage Stratix® V

<table>
<thead>
<tr>
<th>RICH Kernel</th>
<th>Verilog RTL</th>
<th>OpenCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPGA Resource Type</td>
<td>FPGA Resources used [%]</td>
<td>FPGA Resources used [%]</td>
</tr>
<tr>
<td>ALMs</td>
<td>88</td>
<td>63</td>
</tr>
<tr>
<td>DSPs</td>
<td>67</td>
<td>82</td>
</tr>
<tr>
<td>Registers</td>
<td>48</td>
<td>24</td>
</tr>
</tbody>
</table>
Compare PCIe – QPI Interconnect

- Nallatech 385 PCIe vs. Intel® Xeon®+FPGA QPI
- Both Intel® Stratix® V A7 FPGA with 256 DSPs
- Programming model: OpenCL
- Reconstruct 1,000,000 photons

**RICH Cherenkov photon reconstruction (OpenCL)**

![Graph showing comparison between Xeon®, PCIe Stratix V, and QPI Stratix V](image)

<table>
<thead>
<tr>
<th>Acceleration Factor</th>
<th>Xeon®</th>
<th>PCIe Stratix V</th>
<th>QPI Stratix V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>9 x3</td>
<td>26</td>
</tr>
</tbody>
</table>

**RICH Kernel**
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
- Programming model with OpenCL very attractive and convenient for HEP field
- Also other experiments want to test the usage of the Intel® Xeon®+FPGA with Arria10!
- High bandwidth interconnect coupled with Arria® 10 FPGA suggests excellent performance per Joule for HEP algorithms! Don’t forget Stratix® 10!
Backup
Implementation of Cherenkov Angle reconstruction Arria® 10

- 259 clock cycle long pipeline written in Verilog
  - Stratix® V blocks ported using HFB: complex square root, rot. matrix, cross/scalar product, ...

- Pipeline running with 200MHz → 5ns per photon
  - With Arria® 10 GT FPGA 400 MHz possible

- FPGA resources:

<table>
<thead>
<tr>
<th>FPGA Resource Type</th>
<th>FPGA Resources used [%]</th>
<th>For Interface used [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALMs</td>
<td>32</td>
<td>18</td>
</tr>
<tr>
<td>DSPs (HFBs)</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Registers</td>
<td>12</td>
<td>5</td>
</tr>
</tbody>
</table>
Implementation of Cherenkov Angle reconstruction Stratix® V

- 748 clock cycle long pipeline written in Verilog
  - Additional blocks developed: cube 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:

<table>
<thead>
<tr>
<th>FPGA Resource Type</th>
<th>FPGA Resources used [%]</th>
<th>For Interface used [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALMs</td>
<td>88</td>
<td>30</td>
</tr>
<tr>
<td>DSPs</td>
<td>67</td>
<td>0</td>
</tr>
<tr>
<td>Registers</td>
<td>48</td>
<td>5</td>
</tr>
</tbody>
</table>
Nallatech 385 Board

- FPGA: Intel® 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)