The ALICE TRD Global Tracking Unit

An FPGA-based High-speed, Low-latency Processing System for High-Energy Physics

Stefan Kirsch
University of Frankfurt
Frankfurt Institute for Advanced Science

FPL2010 - Milan
September, 2nd 2010
A Large Ion Collider Experiment

- LHC at CERN
  - Proton - Proton collisions @ 14 TeV
  - Heavy-Ion collisions (PbPb) @ 1150 TeV
A Large Ion Collider Experiment

- LHC at CERN
  - Proton - Proton collisions @ 14 TeV
  - Heavy-Ion collisions (PbPb) @ 1150 TeV

- ALICE Detector
  - Research Quark-Gluon-Plasma in PbPb collisions
  - Assembly of many detectors to cover a wide range in particle momentum and PID
Transition Radiation Detector

540 drift chambers
6 layers
90 stacks
18 super-modules
$|\eta| \leq 0.9$

Global Tracking Unit (GTU)
Task of the TRD

- High multiplicities in PbPb: few thousand charged particle tracks in acceptance per event
- Fast trigger detector: Provide L1 trigger after 6.2μs
  - Online reconstruction of high-$p_t$ tracks, calculation of $p_t$
  - Various trigger schemes
- Barrel tracking detector: Raw event data for offline analysis
  - Raw data buffering and transmission to data acquisition system
  - Support interlaced trigger sequences and multi-event buffering
TRD Data Chain

- 2-stage data processing
- On-Detector Front-End Electronics
  - 65,000 ASICs with 260,000 CPUs
- Global Tracking Unit
  - 109 FPGAs in 3-stage hierarchy
Front-End Tracklets

- Massively parallel calculations: Hit calculation and straight line fit
- Tracklets: Parametrization of stiff track segment (y-pos, z-pos, deflection and PID)
- Up to 20000 32-bit tracklet words
- Shipped to GTU via 1080 optical fibres at 2.5Gb/s approx. 4.5µs after interaction

Drift Region

Radiator

Tracklet

Particle Tracks

y

x

3 cm

Drift Volume

Radiator Material

Particle track
Global Tracking

- 3D track reconstruction for up to 20000 tracks in < 1.8µs
- Massively parallel implementation
- Projection of tracklets to virtual plane
- Intelligent sliding window algorithm Δy, Δα_{vertex}, Δz
- Estimation of particle transverse momentum
GTU Structure

Tracklets • Raw data • 240GB/s

Processing Node

Processing Node

x18

TRD Super-module

Concentrator Node

Concentrator Node

Tier 0

Event Buffering
Online Track Reconstruction

Tier 1

Trigger Sequence Reception
Event Buffer & Readout Control
Momentum Reconstruction
Trigger Stage 0

Tier 2

Trigger Stage 1

DAQ 3.35 GB/s

High p_t tracks

Raw data

High p_t tracks +
trigger information

Raw data

L1 trigger contribution
to ALICE CTP
One GTU Segment

GTU segment for one TRD supermodule

Patch panel with 60 fibres for one TRD supermodule
Processing Node

- 6U CompactPCI card
- 14 layer PCB
- Virtex-4 FX100 FPGAs
- Tier-specific assembly & add-on cards
- 2 Embedded PowerPC cores
- 64 MB DRAM
- 4 MB SRAM
- SDCard and ethernet connectors
- LVDS backplane interface
Concentrator Node

- LVDS Link (120 MHz DDR, 4 Bit) to tier 2, 0.9 Gbit/s
- 5 LVDS Links (240 MHz DDR, 8 Bit) from tier 0 nodes 3.8 Gbit/s each

- SDCard Slot
  - 4 GByte SDHC Card

- DDR2 SDRAM
  - 64 MByte

- Interface to ALICE trigger system

- Interface to ALICE DAQ system

- 4 SFP modules
  - 1000Base-SX to switches

- Custom LVDS I/O
  - 72 Pairs

- JTAG

- DCS board

- SIU add-on card
### Tier 0 Design, Rev. 1712

<table>
<thead>
<tr>
<th>Res.</th>
<th>Event Buffering</th>
<th>Tracking</th>
<th>PPCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF</td>
<td>10.945</td>
<td>8.863</td>
<td>3.733</td>
</tr>
<tr>
<td>LUT</td>
<td>5.925</td>
<td>23.463</td>
<td>4.086</td>
</tr>
<tr>
<td>CY / DSP</td>
<td>2505</td>
<td>8,507 / 9</td>
<td>346</td>
</tr>
<tr>
<td>BRAM</td>
<td>14</td>
<td>132</td>
<td>88</td>
</tr>
<tr>
<td>Dist. mem</td>
<td>32</td>
<td>1221</td>
<td>187</td>
</tr>
</tbody>
</table>

- Total slices: 38,910 (92%)
- Total LUTs: 53,406 (63%)
- Global clocks: 16 (50%)
- Block RAMs: 248 (65%)
PowerPC System

109 FPGA Virtex4 FX-100
→ 218 PowerPC cores
Low-level PowerPC

- PPC405 at 400 MHz
- 64-bit IOCM bus at 200MHz
- 32-bit DOCM bus at 200MHz
- Most data contained in DCache → Low interrupt latency (< 380ns)
- Prioritizing interrupt controller
- Data exchange memory for communication with high-level PowerPC
- Used for
  - Robust multi-event buffer control
  - Monitoring
  - Statistics collection / histogramming
  - Readout of non time-critical data (I²C)
Multi-Event Buffering

- **ALICE**: 3-stage trigger hierarchy with interleaved sequences
- Single-event buffering in 260,000 FEE chips
- Multi-event buffering in 90 GTU boards
- Decoupled data taking from stage-2 decision and readout → Significant dead time reduction
Multi-Event Buffering: Data integrity

• How to ensure data integrity?
  How to handle erroneous trigger sequences, etc.?

• Hardware-/Software co-design
  ‣ Support units
    ‣ Handle normal, time-critical operation
    ‣ Error detection
    ‣ Detailed logging of all relevant system events
  ‣ PowerPC
    ‣ Intervene in case of error
    ‣ Analyze system events to identify exact error scenario
    ‣ Report and recover
MEB on Tier 1: Readout control

- Reception of global ALICE trigger sequences
- Sequence validation
- Readout of five Tier0 event buffers
- Fast control logic for normal operation in fabric
- Transfer of control to low-latency PowerPC
- Complex error recovery in software
Statistics

• Complexity of detector, data and trigger chain → Collection of statistics information vital

• Collection subject to trigger sequences
  ‣ Several kHz (Stage0) to approx. 200Hz (Stage2)

• Statistics collection throughout designs

• Limited resources / large amount of statistics data
Statistics Collection Units

- Event-driven push of statistics data to dedicated buffers
- Statistic buffers request readout from low-level PowerPC
  - periodic
  - fill-level dependent
- Prioritizing interrupt controller
- Processing of statistics data in software
Statistics Example: DDL Monitoring

Statistics for last transmitted event:

<table>
<thead>
<tr>
<th>Event size [kB]</th>
<th>Net DDL bandwidth [MB/s]</th>
<th>Backpressure fraction [%]</th>
<th>Unused fraction [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.8</td>
<td>78.6</td>
<td>58.66</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Statistics for transmitted events during last 5 seconds:

- Processed events: 22544
- Rate: 4508.8 Hz
- Average net DDL bandwidth: 58.0 MB/s
- Average DAQ backpressure: 69.51%

Statistics for all transmitted events:

- Number of transmitted words: 0x01208644E
- Data transmitted: 1154 MB
- Average net DDL bandwidth: 56.0 MB/s
Design Verification

- We need to simulate the full system!
  - LHC restart in March 2010
    - Continuous data taking with only minor breaks
  - Limited availability of test setup
  - Large, complex designs with many interfaces

- Aspects
  - Minimize simulation overhead
  - Minimize simulation time
  - Stimuli generation
  - Build flow integration / “Nightly simulation”
  - Rerunning old simulations
Simulation Framework - Features

- Makefile-based automated flow
  - Project file to define whole simulation
  - Adjustable
    - Sources and libraries
    - Back-annotation paths
    - Start-up scripts and runtime commands
  - Subdirectories for simulation
    Libraries and framework file as subdirectory
    Out-of-tree build
  - Included stimulus generation / output checking included
Thank you!
Additional Slides
Global Tracking Unit

• GTU: Second Processing Stage
  • Three 19" racks outside L3 magnet
  • 109 custom PCAs with large FPGAs

• Level-1 Trigger Contribution
  • Detection & full 3D reconstruction of high-\(p_t\) tracks based on tracklets
  • Calculation of transverse momenta
  • Provides various trigger schemes: di-lepton decays (\(J/\psi, \Upsilon\)), jets, ...

• Raw Data Buffering
  • 2.1 Tbit/s via 1080 links from detector
  • Multi-event buffering & interface to ALICE DAQ system
  • Interlaced trigger sequences & extended error handling
Tracking - Track Matching

- 3D track matching: find tracklets belonging to one track
- Processing time less than approx. 1.5 µs

- track bendings and tracklet misorientations exaggerated -
Tracking - Track Matching II

- Projection of tracklets to virtual transverse planes
- Intelligent sliding window algorithm: $\Delta y$, $\Delta \alpha_{\text{Vertex}}$, $\Delta z$
- Track: $\geq 4$ tracklets from different layers inside same window
• Linear fit on matching tracklets: line parameter a, sum of tracklet PID
• Primary vertex assumption
• Estimation of $p_t$ from $a$: $p_t = \frac{const}{a}$, $\Delta p_t/p_t < 1\%$
• Fast $p_t$ cut decision: $const \leq |p_{t,\text{min}} \cdot a|$
Tracking - Design Overview

- 18 matching units running in parallel, 9 track finders
- Fully pipelined data push architecture for minimal latency
- Only fast integer arithmetics, pre-computed look-up tables and DSP blocks used, 18 bits fixed point 9.7
- Critical path: deep combinatorial paths BRAM read data → read addr, 12 logic levels @ 60 MHz

Figure 7.5: The architecture of the TMU trigger design. The TMU receives the track segment data from a module stack, combines it to tracks using several processing stages and reconstructs the transverse momentum of the original particles. The data of the found tracks are transmitted to a superordinate trigger logic that combines the results of the different TMUs. Parametrized tracks are transmitted to the 18 SMUs. In the SMU, tracks from the associated TMUs are combined. The resulting information, which is transmitted to the TGU, as well as the trigger decision algorithm in the TGU itself depend on the chosen trigger scenario.

Each of the GTU's 90 identical TMUs independently processes the data of one detector stack. Since the distance between the stacks is relatively large compared to the curvature of the target particle tracks, and as the geometry of the detector modules is designed to be projective to the point of interaction, looking for tracks across several stacks of detector modules is not necessary. Figure 7.5 gives an overview of the architecture of the TMU trigger design.
Beam Test 2007 Results:

- Accelerator: CERN Proton Synchrotron
- Electrons, Pions with $p_t$ 0.5 – 6 GeV/c
- 8 days of continuous operation, few million events
- GTU algorithm: $\Delta p_t/p_t < 1\%$
- TRD total: $\Delta p_t/p_t < 3\%$
**Tracking - Processing Time**

- Minimum latency of about 550 ns
- Slow nearly linear rise with number of tracklets
- Total latency depending significantly on number of tracklets
Tracking - Latest Collisions at LHC

• 7 supermodules installed, data taking with collisions
• Tracklet tuning ongoing: resolution + availability time
• Preliminary GTU tracking results:

Supermodule 01, Stack 2:
Valid L0 sequences (341787 of 341789=99%):
Tracking in time (<6us): 99% (341654/341787)
Tracking duration: 0.6us: 339928 0.7us: 1723 0.9us: 3 4.5us: 133
Tracking done after L0: 5.5us: 186686 6.0us: 154620 6.5us: 348 8.0us: 133
Tracks (num/cnt): 0: 340842 1: 933 2: 12

Run 124886 2010-07-02 - 17m, 341787 events, 3.5 TeV p-p, Trigger 333 Hz
Tracks per Stack | STACK0 | STACK1 | STACK2 | STACK3 | STACK4
=======================================================================
SEGMENT 00 | 561  957  142  890  961
SEGMENT 01 | 615  863  760  636 1011
SEGMENT 07 | 488  576  262  96  504
SEGMENT 08 | 593  481  441  570  442
SEGMENT 09 | 475  607  328  692  817
SEGMENT 10 | 910  937  710  554  906
SEGMENT 17 | 865  373  540  192  552
=======================================================================
TRD total | 21307 tracks
=======================================================================

• 21,307 tracks in 341,787 events (6%)
  21 GTU tracks/s