Linearized Track-Fitting on an FPGA

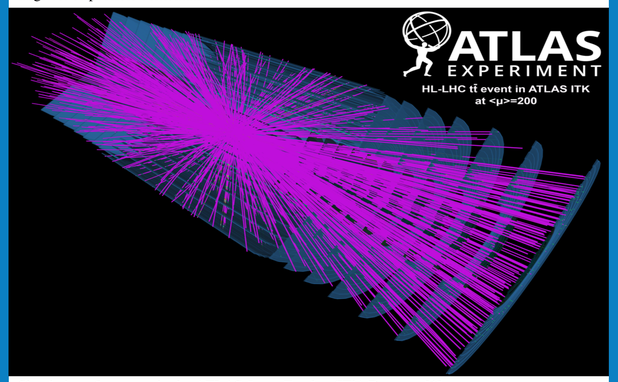
Joachim Zinßer, Dr. Sebastian Dittmeier, Prof. André Schöning
Physikalisches Institut Heidelberg

For the ATLAS experiment at the High-Luminosity LHC, a hardware-based track-trigger was originally envisioned, which performs pattern recognition via AM ASICs and track fitting on an FPGA. A linearized track fitting algorithm is implemented in the Track-Fitter that receives track candidates as well as corresponding fit-constants from a database and performs the χ^2-test of the track as well as calculates the helix-parameters. A prototype of the Track-Fitter has been set-up on a Intel Stratix 10 FPGA. Its firmware was tested in simulation-studies and verified on the hardware. The performance of the Track-Fitter has been evaluated in extensive simulation studies.

Problem

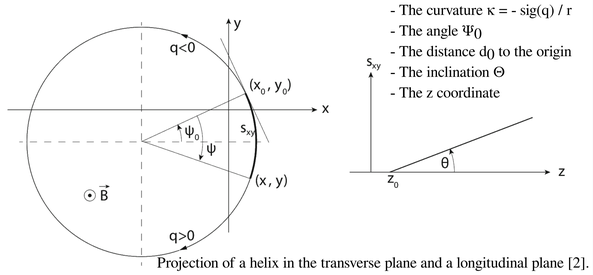
Track-Seeding Problem - Huge Combinatorical Background:

In the ATLAS Phase-II Upgrade for the High-Luminosity LHC, the Inner Tracker Detector (ITk) is in development. It is expected that the detector will observe a pileup of 200 proton-proton collisions. To gain information about the types of particles present in the event, each particle's helix track has to be reconstructed.
Each of the thirteen detector layers detects each particle independently. Track-Seeding describes the process of associating hits of different layers to a track candidate. That leads to an incredibly large set of possible combinations.



Simulation of an event in the ATLAS *Inner Tracker* (ITk) Detector [1].

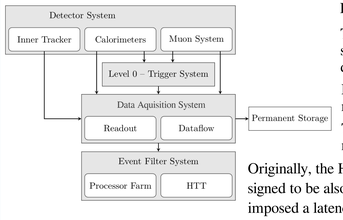
Second problem is the complexity of a non-linear five parameter fit:



- The curvature κ = - sig(q) / r
- The angle Ψ_0
- The distance d_0 to the origin
- The inclination Θ
- The z coordinate

Projection of a helix in the transverse plane and a longitudinal plane [2].

ATLAS-TDAQ dataflow and datarate



Level 0 Trigger rate of 1 MHz.

The detector is divided up into several regions and the data is distributed over 576 PRMs.

Each Track-Fitter gets a event rate of 270 KHz.

That corresponds to a rate of matched patterns of 80 MHz.

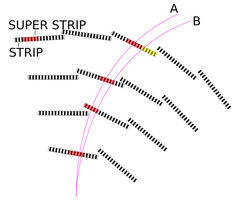
Originally, the Hardware-Track-Trigger (HTT) was designed to be also usable as part of a *Level 1 Trigger*. That imposed a latency requirement of ≈ 1 µs onto the PRM.

Ansatz

Pattern Recognition based Ansatz to Solve the Track Seeding Problem:

Offline Preparation:
1. Perform Monte Carlo Simulation of the detector.
2. Reduce the resolution of the detector. Combine neighbouring detector elements into so called super strips.
3. A pattern consists of one super strip per detector layer.
4. Store the most frequent patterns on AM ASICs.

Online:
1. Reduce the resolution of the incoming data.
2. Load the hit-data into the AM ASICs to compare it with the stored patterns..
3. Return the IDs of all patterns that have above a given threshold of their entries matched.



Tracks traversing strip-layers, organized into superstrips [3].

Linearized Ansatz to Solve the Track Fitting Problem:

Offline preparation:

1. Perform a helix fit on the hits of all patterns in the database.
2. Calculate the helix parameters of the fit.
3. Estimate goodness of the fit (χ^2-method).
4. Linearize the parameters and the χ^2 for small variations in the data.
5. Divide up the database into sectors for which the constants of the linear equations are similar.
6. For all sectors, save the constants in a database on the FPGA on an integrated High-Bandwidth-Memory (HBM).

On the FPGA:

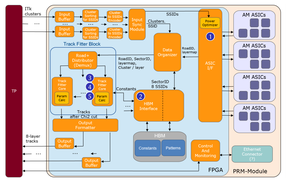
1. Retrieve the goodness-constants $\vec{S}_i$ and h_i for the sector of the hits.
2. Calculate the goodness of the track $\vec{x}$ for I degrees of freedom with:

$$\chi^2 = \sum_{i=0}^{I} (\vec{S}_i \cdot \vec{x} + h_i)^2 \ .$$

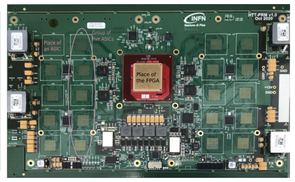
3. For good tracks, retrieve the fit-constants $\vec{C}_j$ and q_j.
4. Calculate the five helix parameters with:

$$p_j = \vec{C}_j \cdot \vec{x} + q_j \ .$$

The Pattern-Recognition-Mezzanine (PRM):



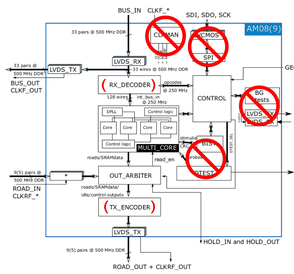
Firmware Diagram of the PRM [4].



Photograph of the HTT-PRM prototype board [5].

Implementation

The Associative-Memory (AM) ASIC and its Emulator:



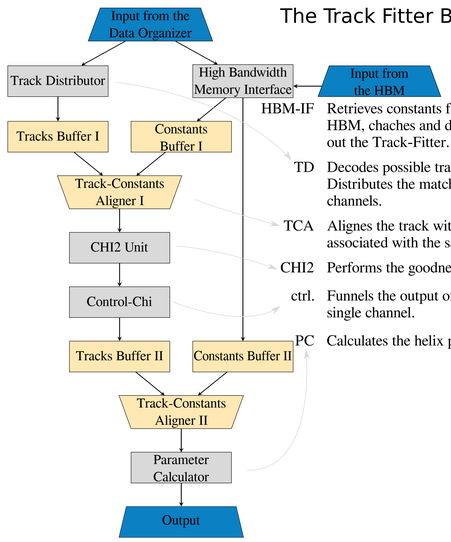
Functional Blocks of the AM-ASIC:

RX-Decoder	Parallelizes and separates commands from data.
Control	Reads commands.
Core	Compares Data with storage and reads Road-IDs.
Out-Arbiter	Multiplexes Roads, Control-Info and Input from the daisychain.
TX-Encoder	Serializes and prepares output.

Funktional block diagram of the AM ASIC [6].

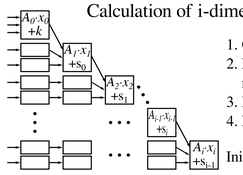
For the testing setups, the AM-ASICs are not used. An emulator that lives on the FPGA was designed to follow the algorithm of the AM-ASIC closely although with a smaller memory, and an extra set of in-ports to write patterns into the memory.

The Track Fitter Block:



HBM-IF	Retrieves constants from the Sectors of the HBM, chaches and distributes them throughout the Track-Fitter.
TD	Decodes possible tracks from the pattern. Distributes the matched patterns into four channels.
TCA	Aligines the track with the constants that are associated with the same Sector-ID.
CHI2	Performs the goodness-test.
ctrl.	Funnels the output of four CHI2 units into a single channel.
PC	Calculates the helix parameters.

Calculation of i-dimensional Scalar Product $\vec{A} \cdot \vec{x} + k$ in Hardware:



1. Chain of i DSPs. Outputs connected to addition-inputs.
2. Pipeline of i registers in front of the multiplication-inputs of the ith DSP.
3. Inputs A_0, x_0 and k are fed to the first DSP.
4. Further inputs A_i and x_i are fed to the registers.

Initial latency of i clock-cycles. Does not limit throughput.

Conclusion

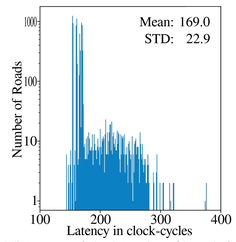
Results and Current State of the Project:

There are two simulation setups: one for the entire PRM with the AM-ASIC emulators and an HBM simulation model, the other one just for the Track-Fitter with an HBM emulator.

The simulation is running with a frequency of 250 MHz. The long-latency regime of the histogram is dominated by the HBM.

The Track-Fitter simulation was shown to be compliant with a road-rate of 80 MHz.

The Track-Fitter simulation has a mean latency of 680 ns. That would fulfills the latency requirement to be used as part of HTT in the Level 1 trigger set-up.



Histogram of the latencies of Roads from the input to the output of the Track-Fitter.

The PRM successfully runs on a prototype Intel Stratix 10 FPGA, although no tests have been performed with a frequency exceeding 100 MHz.

To achieve the same speed on the prototype hardware, further optimization would be necessary.

Intel Stratix 10 MX Development Kit [7].

Summary and Outlook:

- The linearized track-fitting approach was proven to be viable in simulation and on prototype hard-ware.
- Even though, the HTT project was cancelled, the design presented here is a starting point for further developments and can possibly be used in a modified form as part of the ATLAS event filter system.
- Investigations on the performance of the Track-Fitter without the HBM are currently conducted.
- A publication about the development of the PRM firmware is in preparation.

References:

[1] THE ATLAS COLLABORATION: *Event Displays from Upgrade Physics Simulated Data*; 2018.
[2] HERMANN KOLANOSKI & NORBERT WERMES: *Teilchendetektoren*; 2016.
[3] THE ATLAS COLLABORATION: *Technical Design Report for the Phase-II Upgrade of the ATLAS Trigger and Data Acquisition System*; 2018.
[4] THE ATLAS COLLABORATION: *Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC*; 2017.
[5] PAOLO FRANCAVILLA: *PRM HW Status*; 2021.
[6] THE AM DESIGN TEAM: *Phase-II Associative memory ASIC Specifications and Technical/ Scientific Report*; 2019.
[7] Intel Website: `www.intel.com/content/www/us/en/products/details/fpga/development-kits/stratix/10-mx.html`; 2022.