Particle Track Fitting in Hardware for the ATLAS Phase-II Upgrade

Joachim Zinßer

Physikalisches Institut Heidelberg

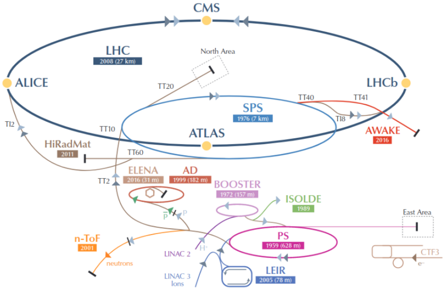
16th June 2021

The Large Hadron Collider (LHC) at CERN

- Largest (27 km circumference) and most powerful (13 TeV collision energy) particle accelerator built so far.
- Development:

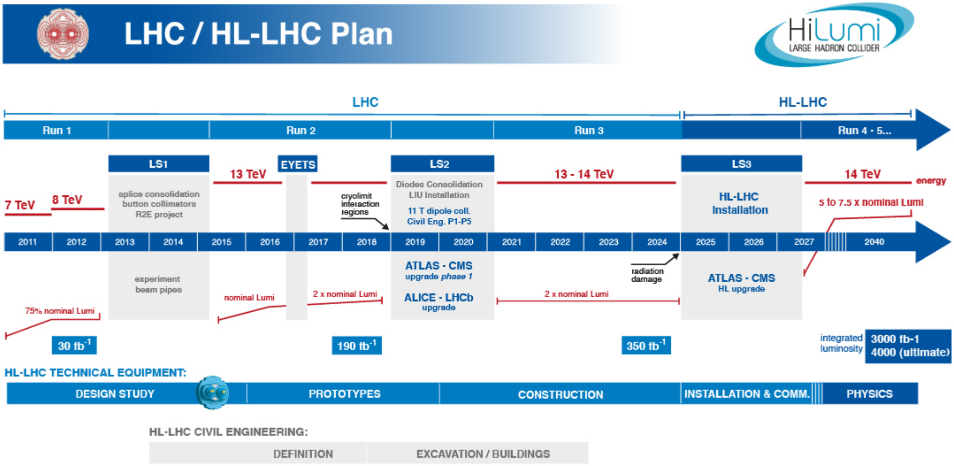
80ies	Envisioned the design
1996	Construction started
2008	Bringing to service

- Seven independent experiments at LHC: ALICE, ATLAS, CMS, LHCb, LHCf, TOTEM, MoEDAL



Map of the accelerators at CERN [CERN 2018].

The High Luminosity LHC – Overview

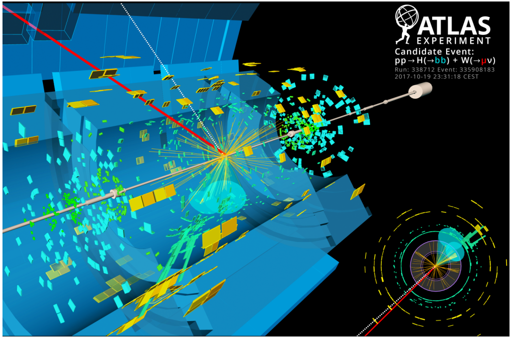


Timetable of the LHC program [CERN-Website].

The High Luminosity LHC – Scientific Aims

- Explore the *Electroweak Symmetry Breaking*.
- Measure properties of the *Higgs Boson*.
- Measure parameters of the *Standard Model* with high precision.
- Try to look *beyond* the Standard Model.
- Flavor Physics.
- Study of *Heavy Ion Collisions*.

[ATLAS 2018, p. 9,10]



The ATLAS detector observing a Higgs Boson decay into bottom quarks [CERN-Website].

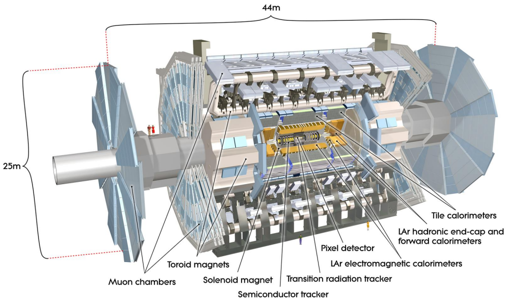
The High Luminosity LHC – Comparison

Parameter	Nominal LHC	HL-LHC
Beam energy in TeV	7	7
Protons per bunch	1.15×10^{11}	2.2×10^{11}
Events per crossing	27	140 – 200
Instantaneous luminosity in $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$	2×10^{34}	up to 7.5×10^{34}
Data collected in $\mathrm{PB\,a}^{-1}$	≈ 30	≈ 300

[ATLAS 2018, CERN 2018, CERN-Website]

The ATLAS Experiment

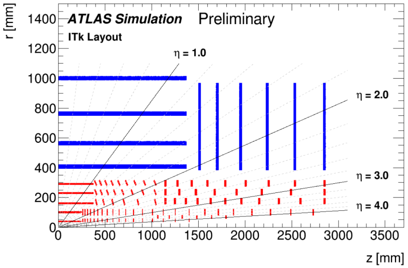
- Largest detector at the LHC.
- General-purpose detector.
- Can observe Proton-Proton and Lead-Ion collisions.
- Layered design of several different detectors.



Sketch of the ATLAS Detector [ATLAS-Website].

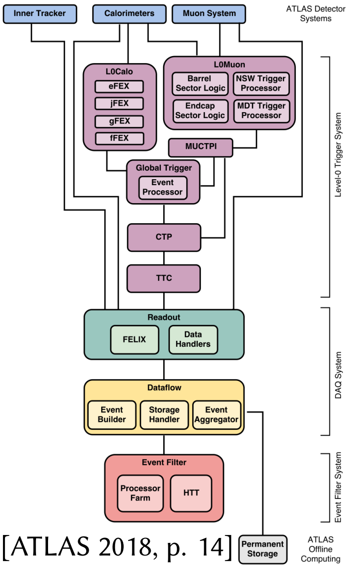
ATLAS – Phase II Upgrade

- Detectors degrade due to radiation.
- Opportunity for Upgrade:
  Installation of the new *Inner Tracker* silicon detector.
- Upgrades of the calorimeters and spectrometers.
- Increase in collisions per crossing from ≈ 40 to ≈ 200
  $\rightarrow$ Necessity to filter interesting events more efficiently
  $\rightarrow$ Trigger system Update



Sketch of the Inner Tracker (ITk) [Moening 2019, p. 5]

The Trigger System of ATLAS Phase II



[ATLAS 2018, p. 14]

- Level-0 Trigger System
  - Calorimeter Trigger
  - Muon Trigger
  - Global Trigger
  - Central Trigger

  Decides if the detectors are to be read out.
- DAQ System
  Bundles data from the detectors to events when triggered by Level-0.
- Event Filter System
  - Processor Farm
  - Hardware Tracking Trigger

  Programmable conditions.
  Accepted Events are sent to permanent storage.

Motivation for Hardware Track-Fitting

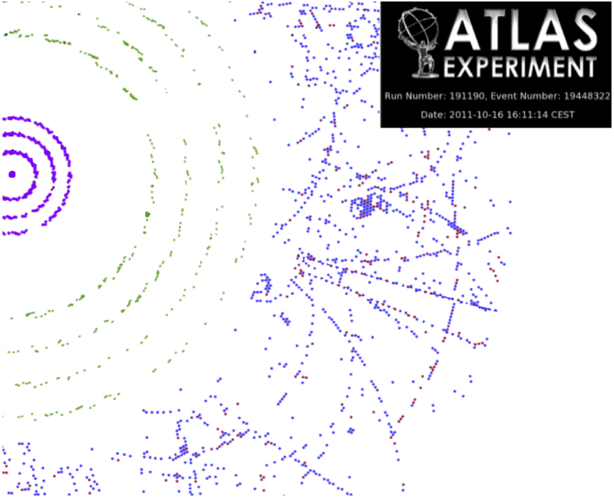
- High pileup conditions (200 interactions / bunch crossing) [ATLAS 2018, p. 39].
- Analyze up to 13 detector layers simultaneously [Dittmeier 2020, p. 15,23].
  $\rightarrow$ Very easy to parallelize hardware.
- High frequency 1 MHz Level-0 trigger rate [ATLAS 2018, p. 13]. $\rightarrow$ Custom hardware can be the fastest.
- Custom hardware is expensive to develop but cheap to produce.
- Custom hardware saves energy-costs.

Goal of the Regional *Hardware Track Trigger* – for ATLAS Phase-II Upgrade

The Level-0 trigger has a trigger rate of about 1 MHz.
$\rightarrow$ Need to suppress background of energetic but scientifically uninteresting events.

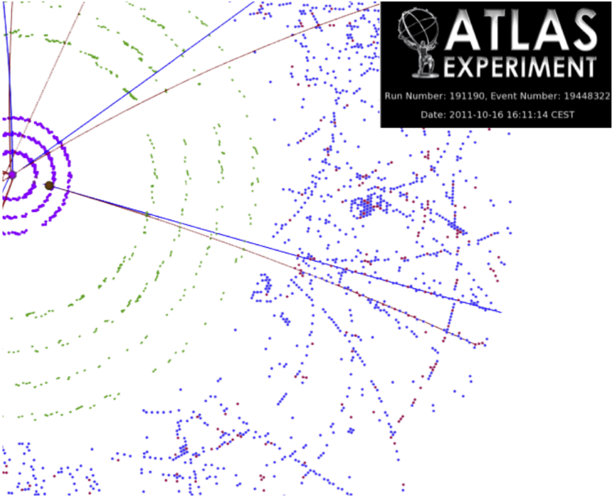
1. Identify which hits in the detector belong to the same particle.
2. Fit the tracks of each particles.
3. Estimate helix parameters.
4. Calculate momentum vectors.

Difficulty of Track-Fitting



A close-up view perpendicular to the beam direction of an ATLAS event [ATLAS 2017].

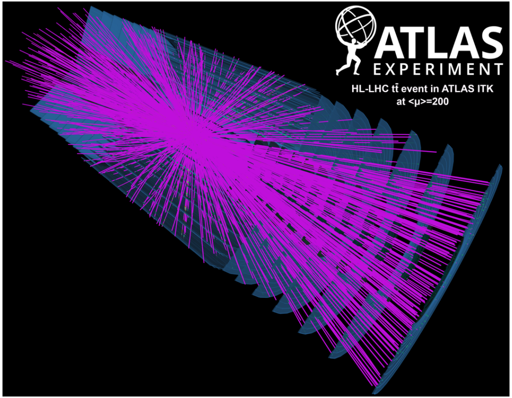
Difficulty of Track-Fitting



A close-up view perpendicular to the beam direction of an ATLAS event [ATLAS 2017].

Possible Options for Track Reconstruction

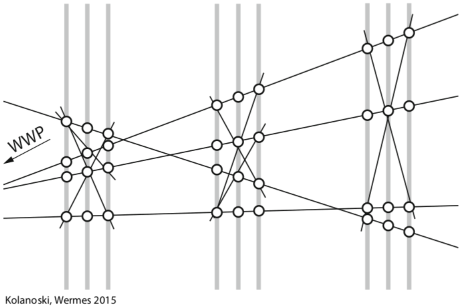
- Fit all permutations of hits and check the qualities of the fits?
- Have to reduce the combinations somehow.
- For example group hits to triplets and only check hits that are close to the line.
- Alternative: Kalman-Filter-Algorithmus
- A goal is to recognize all tracks together (like an eye!).
  - Machine learning
  - Hough-Transformations
  - Templates



Simulation of an Event in the ATLAS *Inner Tracker* (ITk) Detector [CERN-Website].

Possible Options for Track Reconstruction

- Fit all permutations of hits and check the qualities of the fits?
- Have to reduce the combinations somehow.
- For example group hits to triplets and only check hits that are close to the line.
- Alternative: Kalman-Filter-Algorithmus
- A goal is to recognize all tracks together (like an eye!).
  - Machine learning
  - Hough-Transformations
  - **Templates**

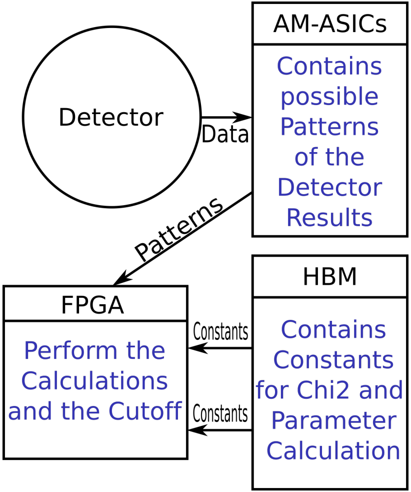


Grouping of 9 layers in 3 super-layers to reduce combinatorics [Kolanoski 2017].

→ Individual calculations of particle Tracks as described will take to much processing time.

Concept of Track Reconstruction in the HTT

1. Compare data with a prepared collection of possible tracks (Pattern Recognition).
2. Calculate the fit quality with linear equations using constants from the χ^2-test.
3. Remove ill-fitting tracks.
4. Calculate the helix-parameters with linear equations and a respective set of constants.



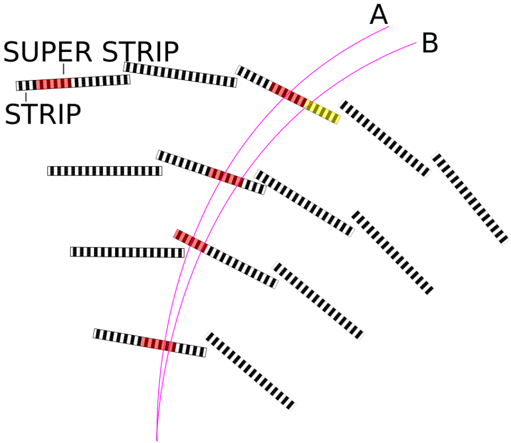
Preparations for Track-Fitting

1. Generate a pattern bank.
2. Perform a fit of the helix parameters on the patterns.
3. Estimate the quality of the fits (χ^2-method).
4. Formulate linear dependency of the χ^2 and the helix-parameters (for small sectors of the detector).
5. Save the constants for the linear equations in a database.

The Pattern Bank

- Generate a *pattern bank*:
  - Generated using full simulations of Monte-Carlo muons.
  - Different binning of pixels to superstrips.
  - Data is compressed (don't cares, wildcards).
  - Usage-count of the patterns is also saved.

  [ATLAS 2018, p. 350]
- The detector is separated into independent regions.
- 4 Million patterns prepared for each region [ATLAS 2018, p. 354].



Tracks traversing strip-layers organized into superstrips [ATLAS 2018, p. 351].

Preparation of Helix Parameters

- Combine similar patterns from the bank into *Sectors*.
- Each sector has a simulated set of n values y_i (at the positions x_i) and their covariance matrix V_x.
- Utilize a fit-function $f(x, \vec{\theta}) = \sum_{j=1}^{m} \theta_j f_j(x)$ (the $f_j(x)$ do not need to be linear).
- With the expected values $\eta_i = f(x_i, \vec{\theta}) = \sum_{j=1}^{m} \theta_j f_j(x_i) = \sum_{j=1}^{m} H_{ij}\theta_j$ you get the $n \times m$-Matrix H.
- Perform the *least-square-method* $S = \sum_k^n \sum_l^m (y_k - \eta_k)\, V_{ij}^{-1} (y_l - \eta_l)$ or in the matrix form $S = (\vec{y} - H\vec{\theta})^{\mathsf{T}}\, V^{-1}\, (\vec{y} - H\vec{\theta})$.
- Solve this system of linear equations for the parameters θ_j: $\vec{\theta} = (H^{\mathsf{T}} V^{-1} H)^{-1}\, H^{\mathsf{T}} V^{-1}\, \vec{y} = C\vec{y}$
- Save the m times n entries of C as constants for the Sector.

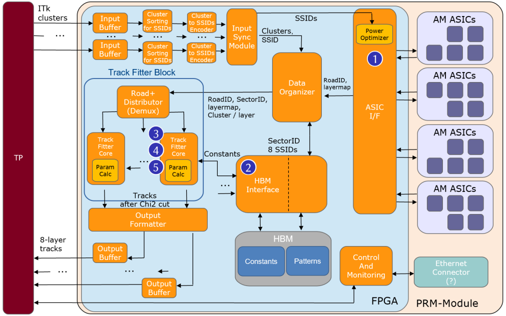
[Kolanoski 2017]

Memory Requirements for the Constants

- There are up to 65 000 Sectors in a region.
- Desired resolution of the number-encoding of the constants is 32 bit.
- Planing for the extreme case of four pixel and four strip layers in the ITk:
  - 3 kbit for a set of Chi2 constants
  - 2 kbit for a set of PC constants
- Together 325 Mbit of constants.
- Expected maximum Road-Rate at the Track-Fitter is 135 MHz.
- Although large, the necessary bandwidth to the memory is not 370 Gbit/s. Several roads will require the same constants.

Algorithm

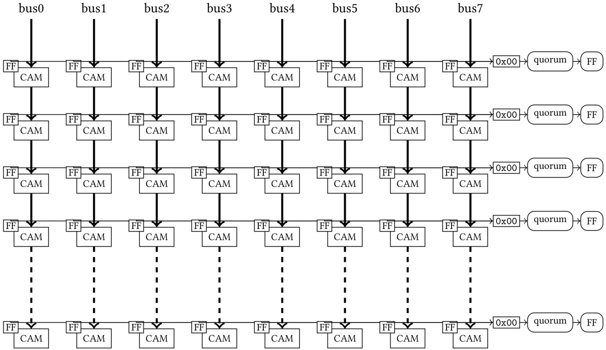
1. Compare data $\vec{x}$ with pattern-bank $\rightarrow$ receive ID of the matched sector.
2. Request constants for the Sector from the HBM.
3. Calculate the quality of the match with $\chi^2 = \sum_{i=1}^{7} \left(\vec{S}_i \cdot \vec{x} + h_i\right)^2$.
4. If χ^2 is below a threshold $\rightarrow$ request further constants.
5. Calculate the 5 helix parameters $p_j = \vec{C}_j \cdot \vec{x} + q_j$.



Firmware Diagram of the PRM-Module of the HTT [ATLAS 2017].

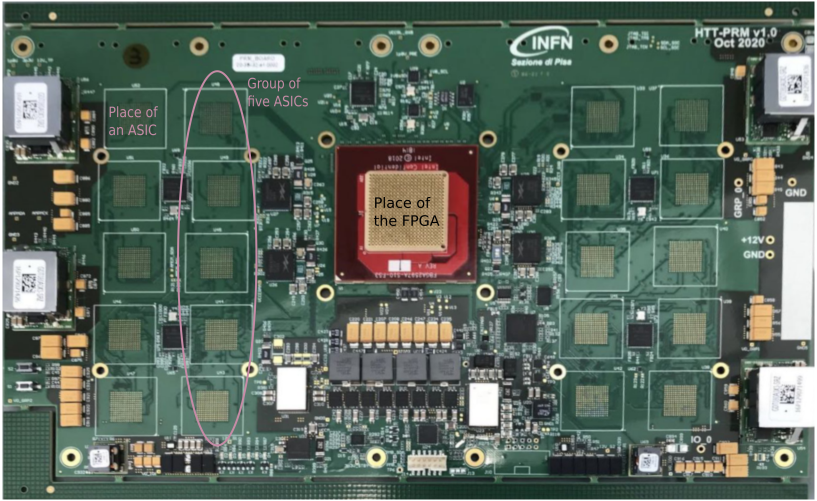
The *Associative Memory* (AM) ASICs

- Incoming data arrives at all CAM-units simultaneously.
- Each CAM-unit compares the data to its content. *rightarrow* this CAM-unit gets flagged.
- At the end of an event each line counts its flags.
- The lines represent patterns and if a pattern has enough flags, its ID is read out.



Sketch of the Pattern-Recognition firmware [Crescioli 2018].

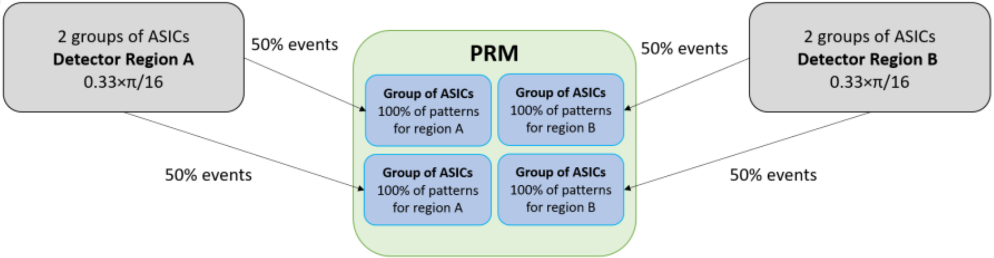
The *Pattern Recognition Mezzanine* (PRM)



Photograph of HTT-PRM prototype board [Francavilla 2021].

Requirements to the AM-ASIC Interface

- Up to 8 million patterns can be stored.
  $\rightarrow$ only 4 million different patterns needed.
  $\rightarrow$ increase speed through parallelization.
- A data-word is simultaneously compared to all patterns in the database.
- Matches are stored and read out sequentially.



Schematic view of event processing and pattern distribution [ATLAS 2019] (evolved scenario).

The Intel Stratix 10 FPGA with embedded *High Bandwidth Memory*

- Almost 900 ports.
- Approximately 700 000 *Adaptive Logic Modules* (ALMs).
- Almost 4000 *Digital Signal Processor* (DSP) Blocks.
- 140 M bit Block-Memory.
- Two embedded 8 GB *High Bandwidth Memory* (HBM) units.
- Two HBMs provide a bandwidth of up to 270 Gbit/s.

Requirements to the *High Bandwidth Memory Interface* (HBM-IF)

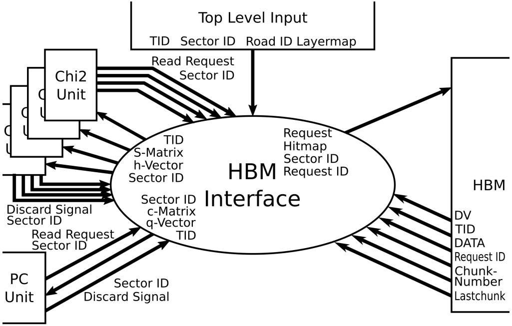
Desired resolution (32 bit number-encoding)

$\rightarrow$ 5 kbit for one set of constants $\vec{C}_j$, q_j, $\vec{S}_i$ and h_i

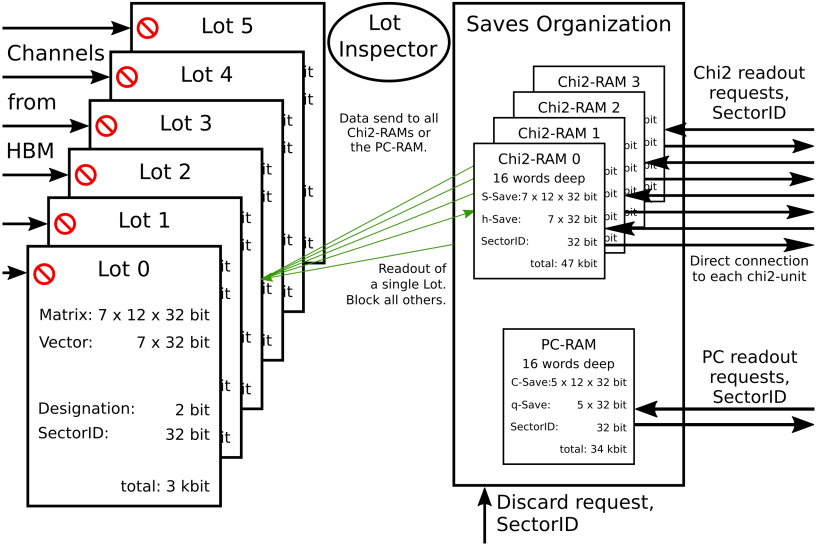
Have to be available each clock-cycle, possibly at several places in the FPGA!

Tasks of the *HBM-Interface*:

- Transmits requests.
- Stores incoming data until a set of constants is complete.
- Distributes constants between the various FPGA-entities.
- Suppresses repeated requests, temporarily storing constants $\rightarrow$ Reduces required bandwidth.

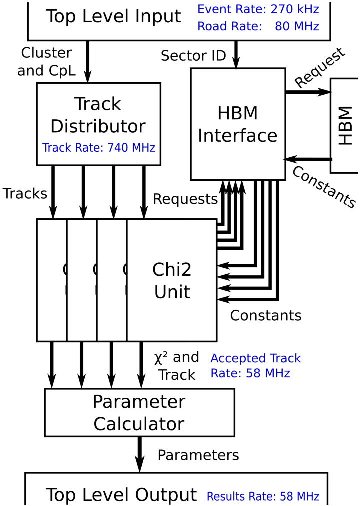


The Concept of the HBM Interface



The Track Fitter Block

- All entities are designed to not limit throughput.
- To save resources perform χ^2-test first.
- Less Parameter Calculator Units than Chi2-Units.



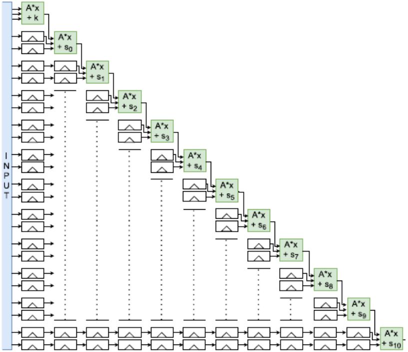
Flow-Diagram of one TF.

Scalar Product calculated by using Pipelined DSPs

How to calculate 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.

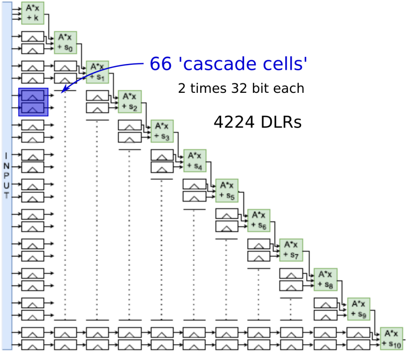


Sketch of the pipelined DSP cascade of an scalar product [Braagaard 2020].

Scalar Product calculated by using Pipelined DSPs

How to calculate 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.

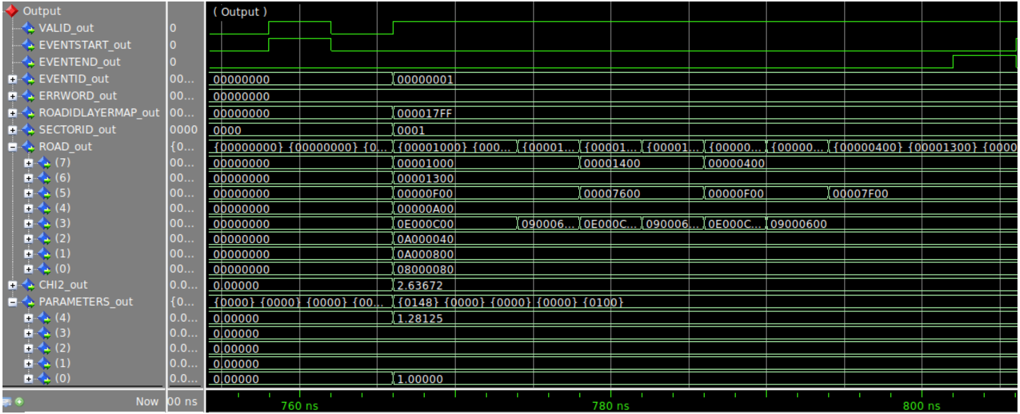


Compilation Hierarchy Node	Combinational ALUTs	Dedicated Logic Registers	Block Memory Bits	DSP Blocks
	53104 (84)	226528 (4420)	725376	424
|gen_chi_square:0:chi_square|	861 (19)	35090 (5003)	34304	91
|FIFO|	672 (0)	480 (0)	34304	0
|SQUARE_IP_0|	0 (0)	0 (0)	0	1
|GEN_INNER_SUMS:0:inner_sum|	0 (0)	4224 (4224)	0	12

Resource Usage of the Preliminary Track-Fitter Firmware

Compilation Hierarchy Node	Combinational ALUTs	Dedicated Logic Registers	Block Memory Bits	DSP Blocks
	53104 (84) 7.4%	226528 (4420) 8.0%	725376 0.5%	424 10.6%
|\gen_chi_square:0:chi_square| (x 4)	861 (19)	35090 (5003)	34304	91
|\gen_constants_bu...:buffer_constants| (x 4)	1748 (0)	1288 (0)	46848	0
|\gen_track_buffers:0:buffer_tracks| (x 4)	456 (0)	336 (0)	8576	0
|\gen_track_const_...align_track_const| (x 4)	32 (32)	3458 (3458)	0	0
|align_track_const_ii|	29 (29)	2642 (2642)	0	0
|buffer_chi2|	3073 (781)	2617 (849)	70912	0
|buffer_constants_ii|	1255 (0)	924 (0)	33536	0
|buffer_pc|	475 (0)	350 (0)	8832	0
|hbm_interface|	26859 (10)	22340 (17)	221312	0
|output_buffer|	2822 (167)	1253 (276)	14848	0
|parameter_calculator|	475 (0)	21643 (161)	8832	60
|track_distributor|	5686 (106)	9651 (1891)	8192	0

Output of the Track-Fitter



The Tracks of Road 000017 have parameters $P = (1, 0, 0, 0, 1.28)$.

Results with Respect to the Extended Requirements

- The HTT can currently be build with a maximum frequency of 170 MHz.
  Fewer than the expected 250 MHz ([?]) but probably achievable with optimization.
- The Track Fitter can handle the expected Track Rate of up to 450 MHz ([?]) but not a Road-Rate of 135 MHz.
  Probably necessary to reorganize the Track-Distributor.
- It is expected that only 9 MHz of Tracks pass the Chi2-test.
  Parameter-Calculator can handle that.

Summary and Outlook

- The Track-Fitter is now working in Simulation.
- Combined firmware of the HTT-PRM works in simulation.
- The HTT is compliant with the baseline of the ATLAS Phase-II upgrade.
- The HTT is not yet compliant with the extended requirements.
- Some parts of the PRM are already verified to work on hardware.
- In the next weeks, the Track-Fitter will be tested on hardware.

CERN:
"LHC Season 2 – Facts & Figures",
Informative Brochure, July 2018.

CERN, Hi-Lumi Project:
"HL-LHC Industry – Project Schedule",
project-hl-lhc-industry.web.cern.ch/content/project-schedule, 21st April 2020.

CERN:
"Long-sought Decay of Higgs Boson Observed",
home.cern/news/press-release/physics/long-sought-decay-higgs-boson-observed, 22nd April 2020.

Roman copy of Greek original, Photographer: Lalupa:
"Statua romana di Atlante – Collezione Farnese – Museo Archeologico Nazional di Napoli",
en.wikipedia.org/wiki/Atlas_(mythology), 21st April 2020.

CERN, The ATLAS Experiment:
"Detector & Technology",
atlas.cern/discover/detector, 17th April 2020.

The ATLAS Experiment:
"Event Display from Upgrade Physics Simulated Data",
CERN, 2021.

The ATLAS Collaboration, Klaus Moening:
"Expected Tracking Performance of the ATLAS Inner Tracker at the HL-LHC",
ATL-PHYSI, PUB-2019-014, March 2019.

The ATLAS Collaboration:
"Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties",
Journal of Instrumentation, Volume 12, May 2017.

The ATLAS Collaboration:

”Technical Design Report for the Phase-II Upgrade of the ATLAS Trigger and Data Acquisition System”,
CERN-LHCC-2017-020, June 2018.

The ATLAS Collaboration:
”HTT ATLAS Electronics Specification for Pattern Recognition Mezzanine (PRM)”,
ATL-COM, DAQ-2018-039, December 2019.

The AM Design Team:
”Phase-II Associative Memory ASIC Specifications and Technical/Scientific Report”,
CERN-OPEN-2018-003, November 2019.

Francesco Crescioli:
”Development of Associative Memory ASIC in 28nm technology”,
Joueneés VLSI - FPGA - PCB et Outils CAO de l’IN2P3, May 2018.

Paolo Francavilla:
”PRM HW Status”,
PRM Meeting, Presentation, March 2021.

Hermann Kolanoski, Norbert Wermes:
”9 Spurrekonstruktion und Impulsmessung” in ”Teilchendetektoren”,
Springer-Verlag, Berlin und Heidelberg, 2016.

Mathias Braagaard:
”Implementation of a linearized χ^2-Square [sic] method for the ATLAS Pattern Recognition Mezzanine”,
DTU Kongens Lyngby, early 2020.

Magnus Høegh Jensen:
”Implementation of Particle Track Parameter Calculations for the ATLAS Pattern Recognition Mezzanine”,
DTU Kongens Lyngby, early 2020.

Sebastian Dittmeier:
”The ATLAS Hardware Tracking for the Trigger System”,

HEP Group Meeting, February 2020.