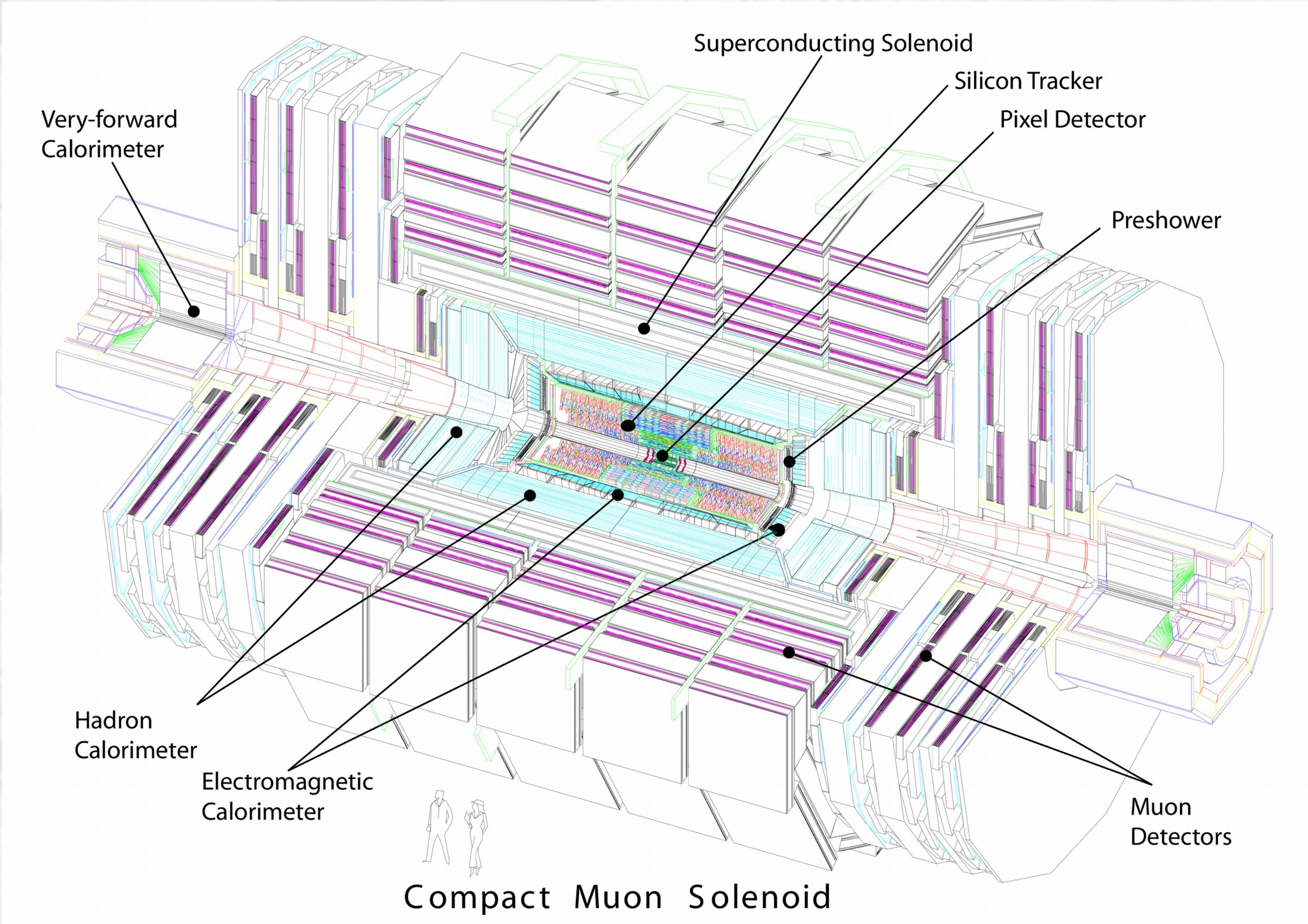
From the Physical Model to the Electronic System - OMTF Trigger for CMS

Presenter:
Wojciech M. Zabołotny
Institute of Electronic Systems
Warsaw University of Technology
&
Faculty of Physics
University of Warsaw

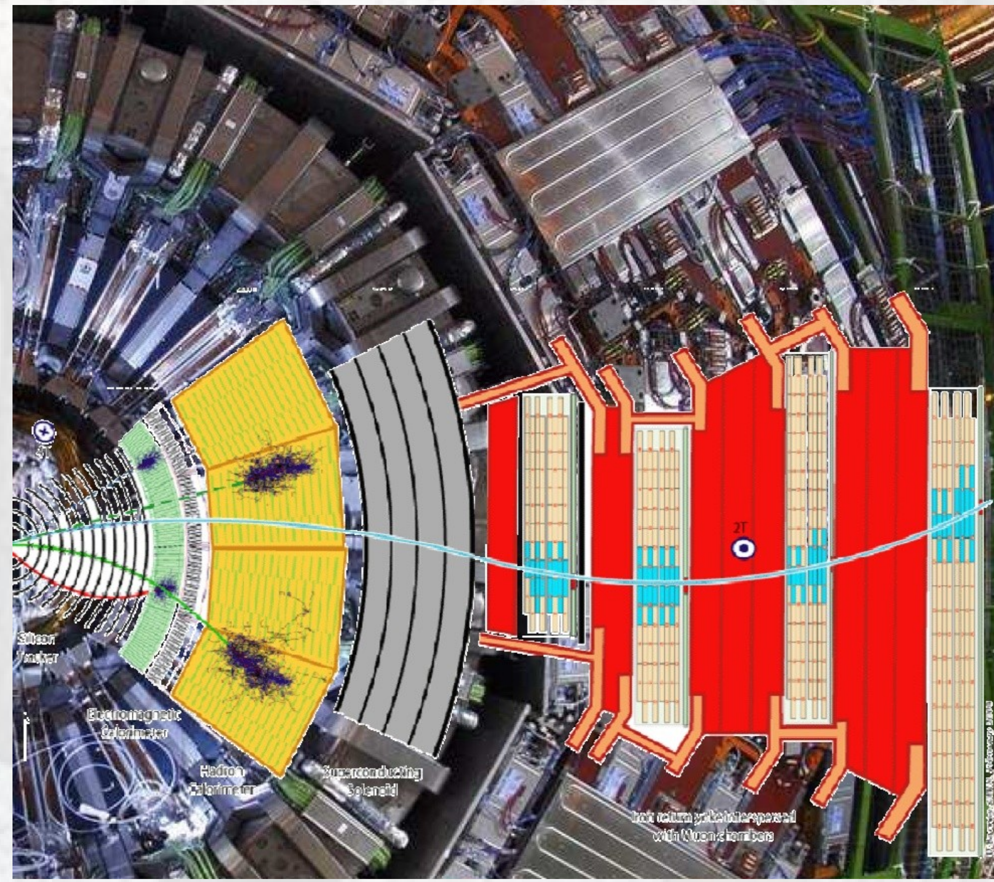
The material presented is the result of the work of the whole Warsaw CMS group

CMS detector



Muon detectors in CMS

- The muon trigger in the CMS detektor uses three types of muon detectors:
  - Cathode Strip Chambers (CSC)
  - Resistive Plate Chambers (RPC)
  - Drift Tubes (DT)
- The muon crossing the detector chamber generates a signal („hit” - shown in light blue in figure on the right)

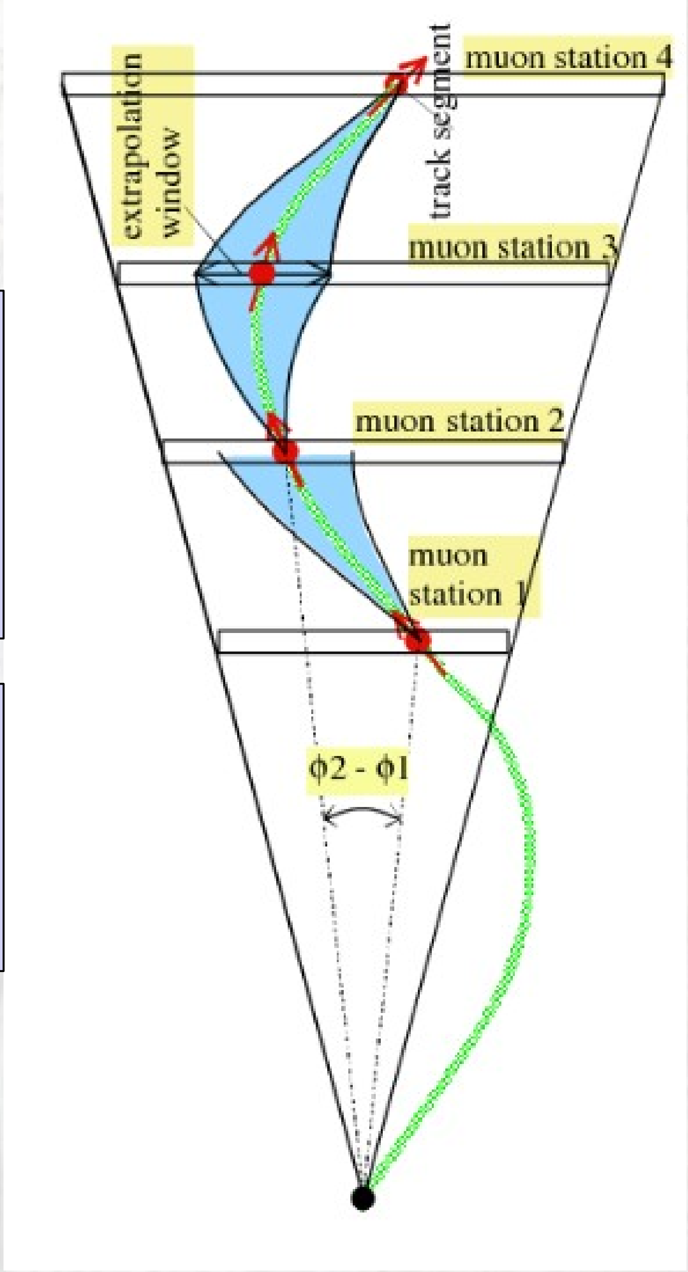


DT/CSC Track Finders

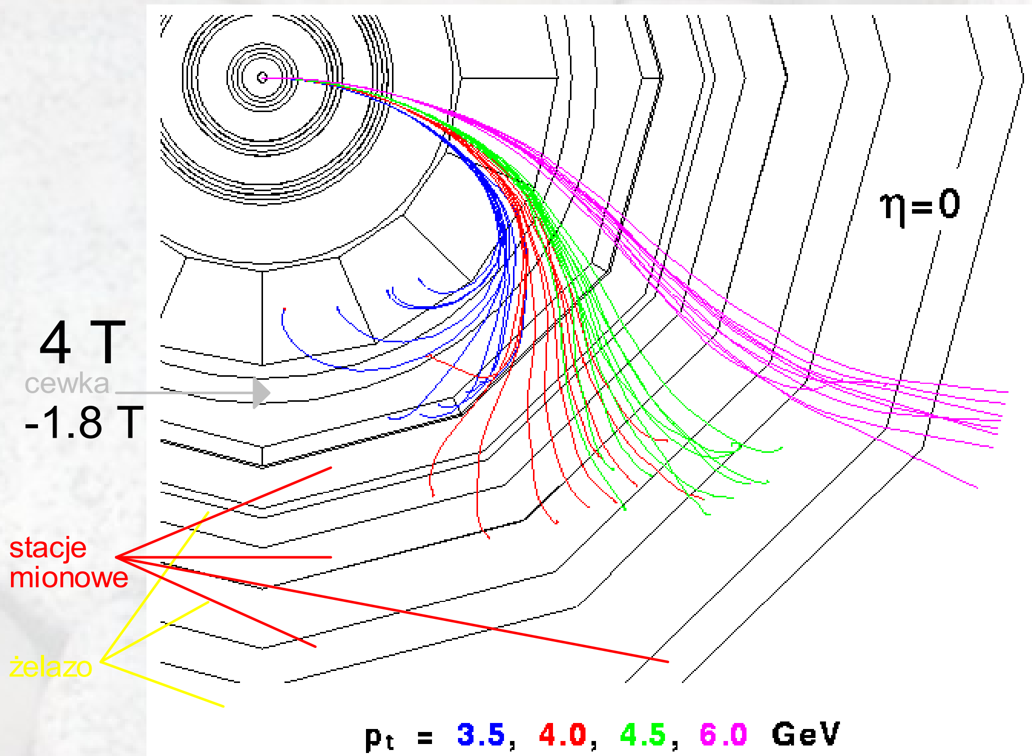
Track segments from DT and CSC combined by corresponding Track Finders (exchange at $|\eta|\approx 1$).

DTTF: The local position and bending are used to build track candidate from matched segments (extrapolation based on build-in LUTs). Momentum defined by two available lowest segments (assigned by LUTs). Output 4 candidats after sorting.

CSCTF: Pairs of segments are checked against hypothesis of common track . Momentum defined by memory LUTs, addressed by track η and φ difference in up to 3 stations. Output 4 candidates after sorting



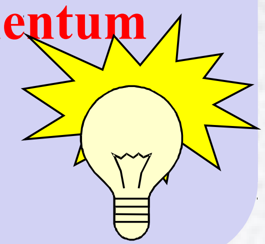
PACT – principle



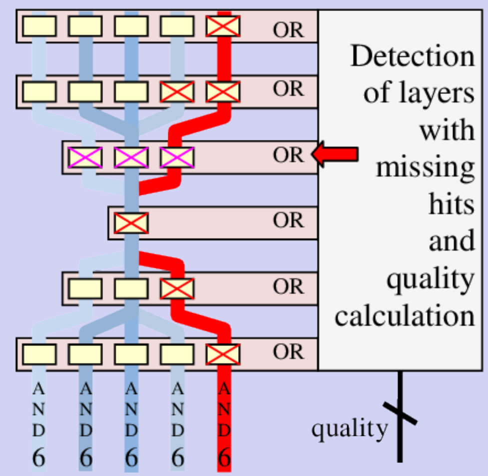
The muon track is bent in 3.8T magnetic field inside the solenoid and un-bent in 1.8T in yoke

curvature ↔ transverse momentum

track spread due to stochastic effects: multiple scattering and energy losses.



- **The chamber hits (fired strips) in each event are compared with the predefined set of patterns every 25 ns (BX).** In case of matching the PACT muon candidate is set.
- Each pattern has assigned a transverse momentum which is then assign to reconstructed muon candidate



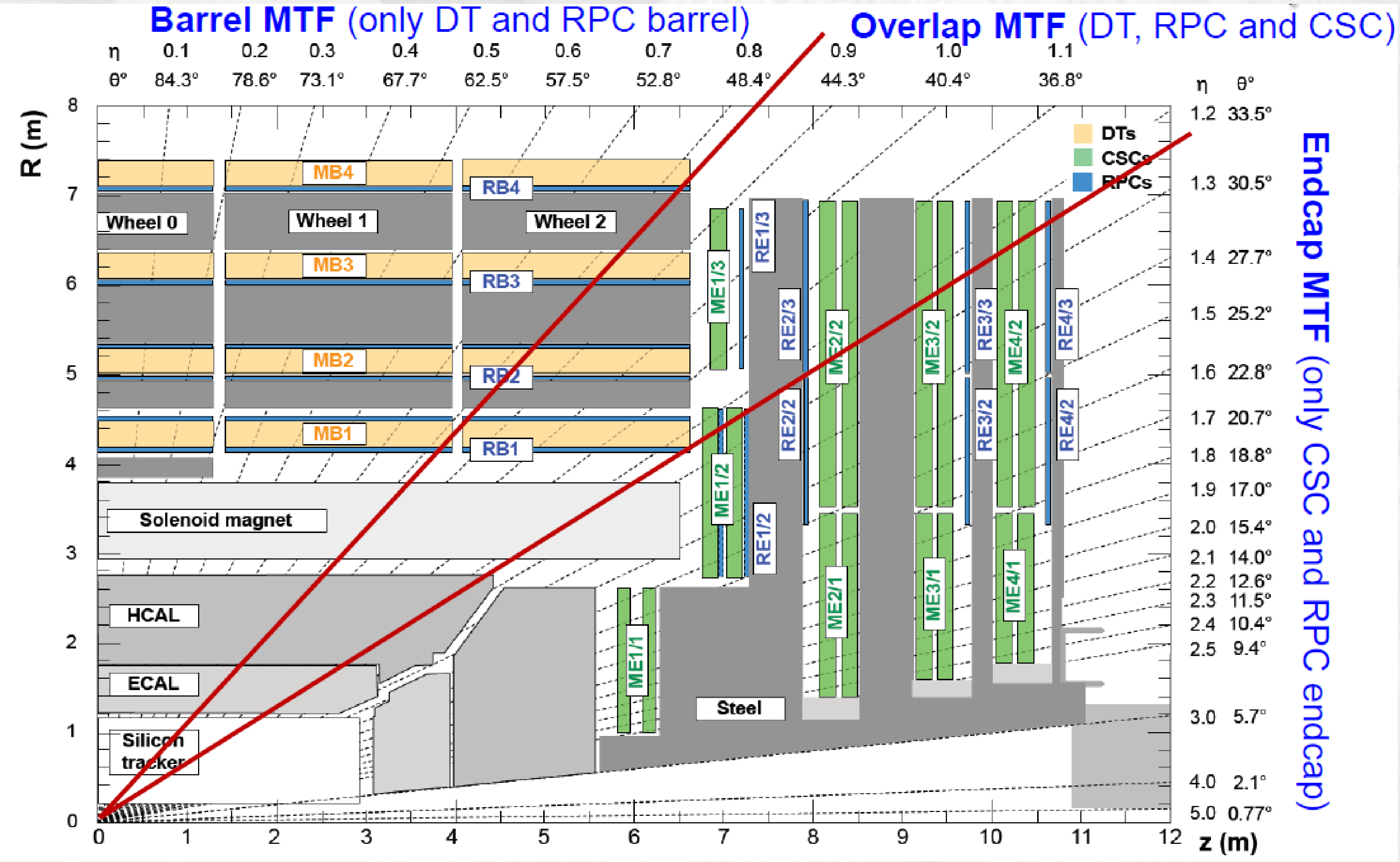
- Advanced algorithms are used to create the patterns from the simulated chamber hits, assign the p_T, and then select optimal set of patterns.

The goal is to achieve best possible trigger efficiency and purity with a patterns set that can fit into the PAC FPGA resources.

Trigger upgrade after LHC Run 1

- Planned increase of LHC luminosity (to ca.$2\times10^{34}cm^{-2}s^{-1}$) and center of mass energy (to 14 TeV) requires better selection of events in the L1 trigger, to keep L1 rate at 100 kHz.
- It is necessary to upgrade the electronic systems due to technology changes (depleting stocks of spare parts).
- The Warsaw CMS group is responsible for implementation of the Muon Trigger in the Overlap area of the detector

Specific features of the Overlap region



Common representation of hits

- We describe the CMS detector as a set of „layers", where each layer contain the muon detector of certain type (CSC, DT or RPC)
- The hit in each detector may be described by corresponding azimuthal angle φ.
- The „naive" approach could be to find the values of the azimuthal angle φ and the transversal momentum p_T, which provides the lowest distance between the measured hits and the theoretical track
- Problems:
  - There is no single deterministic muon track for particular p_T, and Φ, due to stochactic effects (energy losses, multiple scattering)
  - The sequential optimization of Φ and p_T, would be too time consuming
  - We need to handle events with more than one muon

Golden Pattern algorithm

- We want to measure „how good the detected hits match the average track of the muon with certain transversal momentum"
- Therefore for each detected range of transversal momenta we generate the pattern which will be compared with received hits
- A few questions still must be answered:
  - What should be the contents of the pattern?
  - How we can generate the patterns?
  - How the matching between hits and the patterns should be performed to ensure satisfactory performance?

Golden Pattern algorithm cont.

- The matching between the pattern and the hits may be measured by „the probability, that the measured track was generated by the muon with transversal momentum from the certain range”
- Therefore the pattern may contain probabilities, that the hit in a certain range of azimuthal angle was generated by the muon with the transversal momentum from the range covered by that pattern
- In case of hits in multiple layers, this probability is the product of probabilities calculated for each layer.
- Those probabilities may be found from physical simulations (in CMSSW, from the number of simulated muons crosing the particular layer in the particular Φ range)
- Problems:
  - Probabilities have high „dynamic range”
  - How to handle events in which the muon have not generated any hit in a particular layer (or a few layers)
  - Multiplication is a relatively expensive operation
  - How to handle multi-muon events

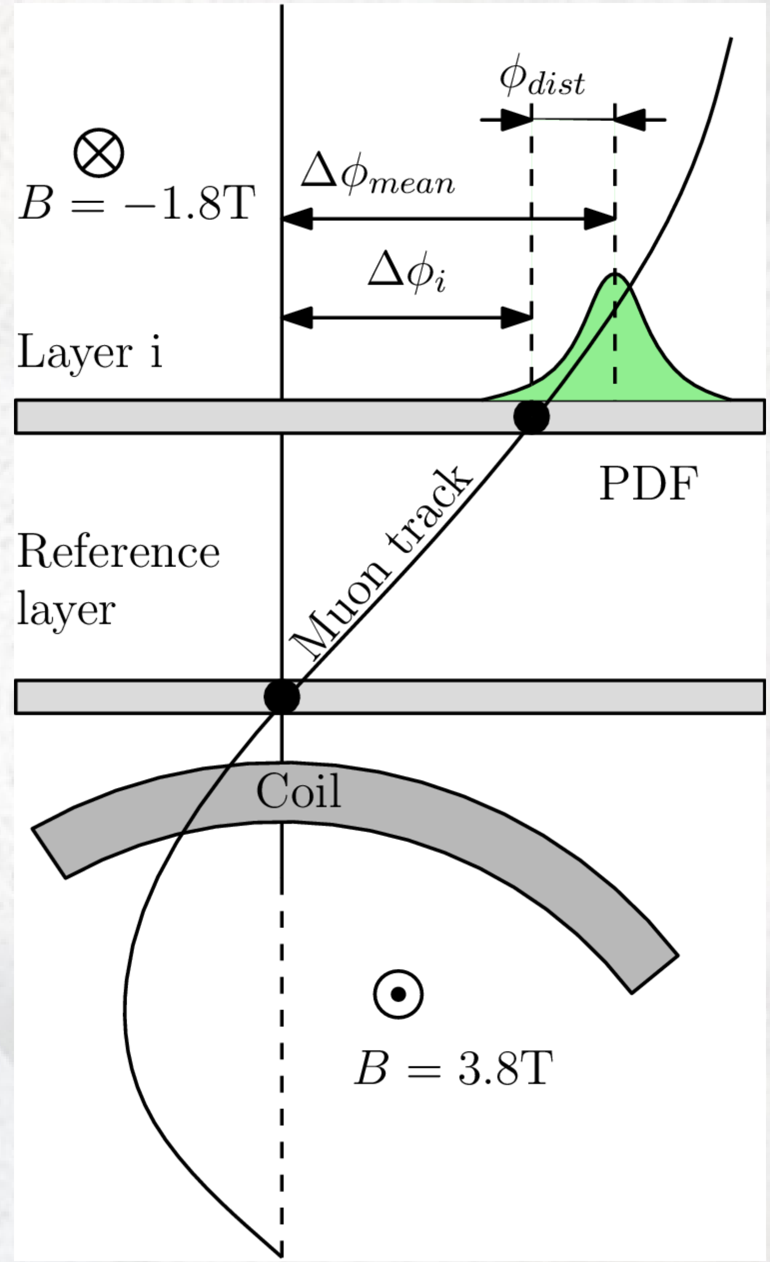
Golden Pattern algorithm – implementable version

- Instead of taking **directly the probability**, the algorithm uses values roughly corresponding to the **logarithm of this probability** (with certain offset and scaling, and limited to 0 for low probablities)

  This values will be denoted as PDFs (probability density functions)
- Thanks to that change the **multiplication may be replaced with summing.**
- The best matched pattern is found:
  - First: as the one with the highest number of layers giving non-zero PDF
  - Next: as the one with the highest sum of PDFs from all layers

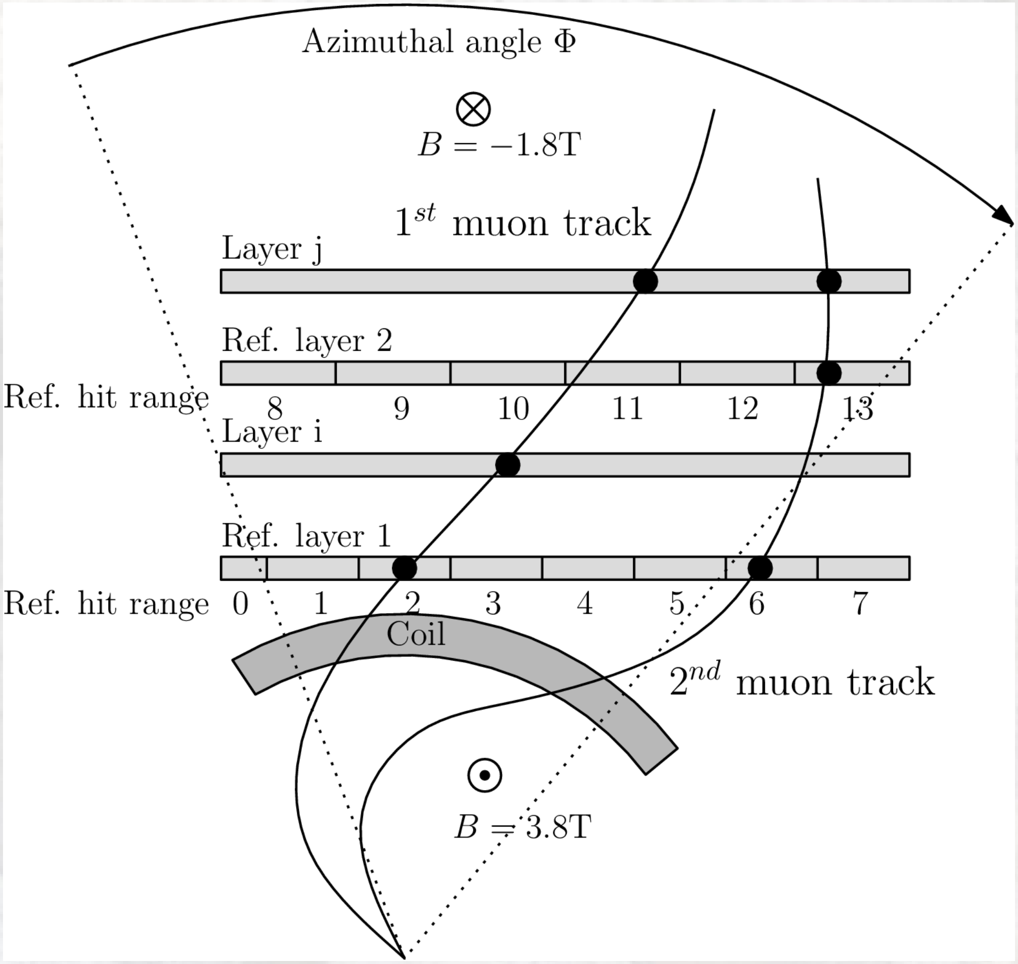
OMTF algorithm – calculation of PDF

- Utilizing the cylindrical symmetry of the detector, we perform all calculations for φ-φ_{ref}, where φ_{ref} is the azimuthal angle of the hit in certain „Reference layer”
- To further reduce complexity of the algorithm, the PDF values are precalculated and stored in LUTs
- To minimize LUT capacity, the position of PDF maximum is stored independently, and only non-zero values surrounding this maximum must be stored



OMTF algorithm – multimuon event

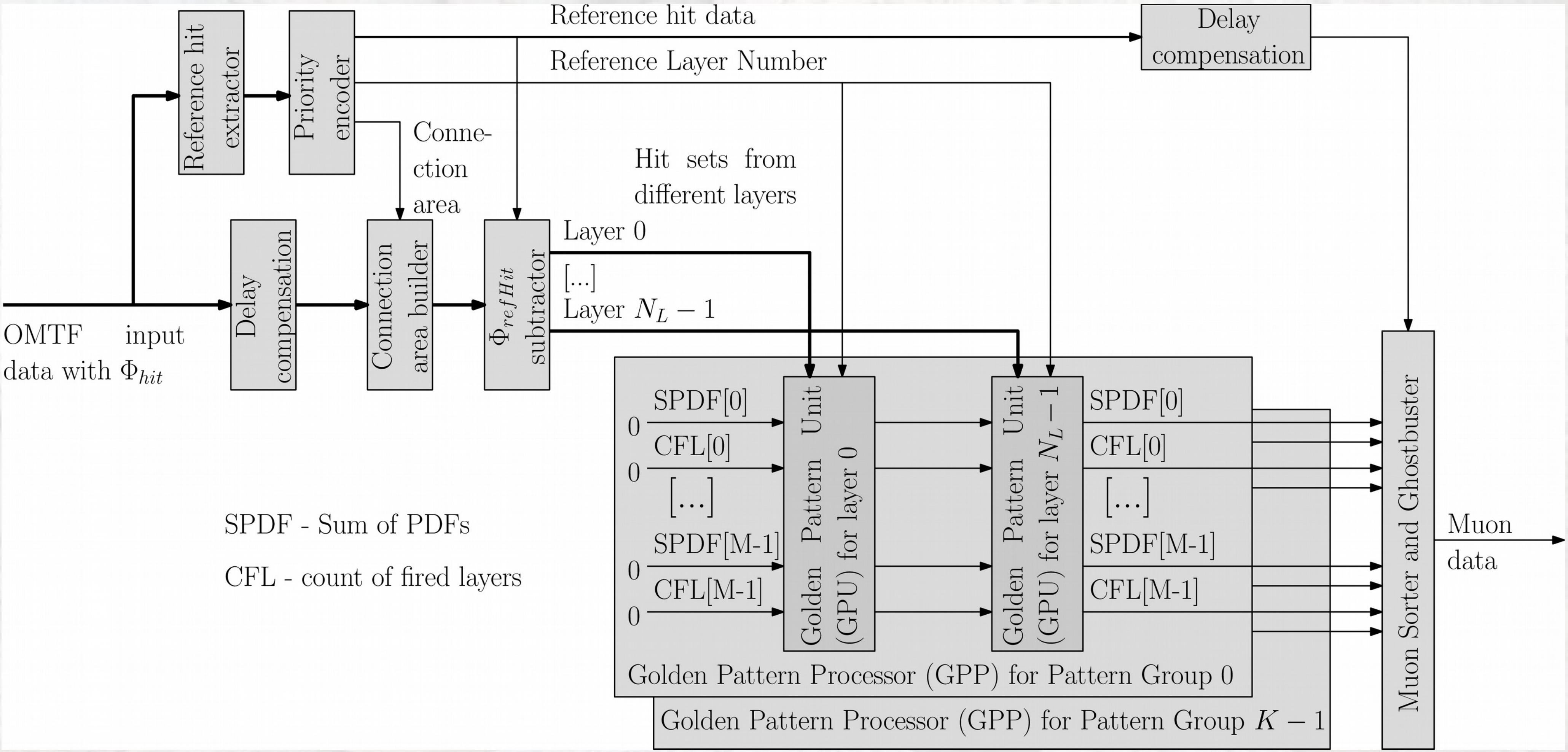
- To properly handle multimuon events, a few „reference hits" may be used as a basis for muon identification for each event (in each BX)
- The „reference hit ranges" are ordered in such a way, that taking the next one we maximize probability that in case of multi-muon event they will be associated with different muons



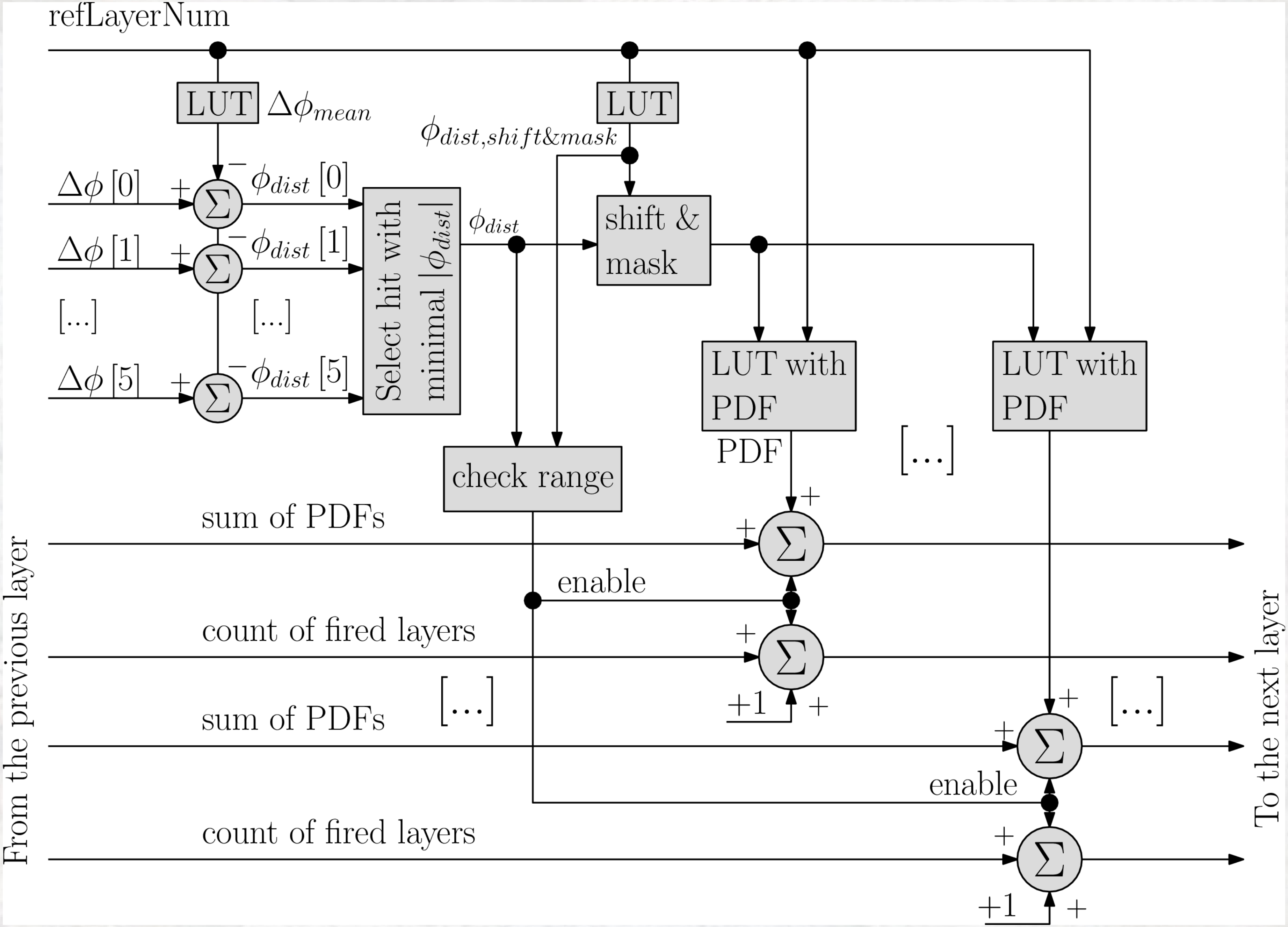
OMTF implementation – hardware architecture

- The standard MTF7 board used in CERN trigger update is used
  - Two FPGAs: Kintex 7 XC7K70 and Virtex 7 XC7VX690T
  - OMTF is implemented in XC7VX690T
- To fullfill requirement of possibly **small** and **constant latency** (not more than 20 periods of ca. 40 MHz LHC clock) the pipelined architecture must be used

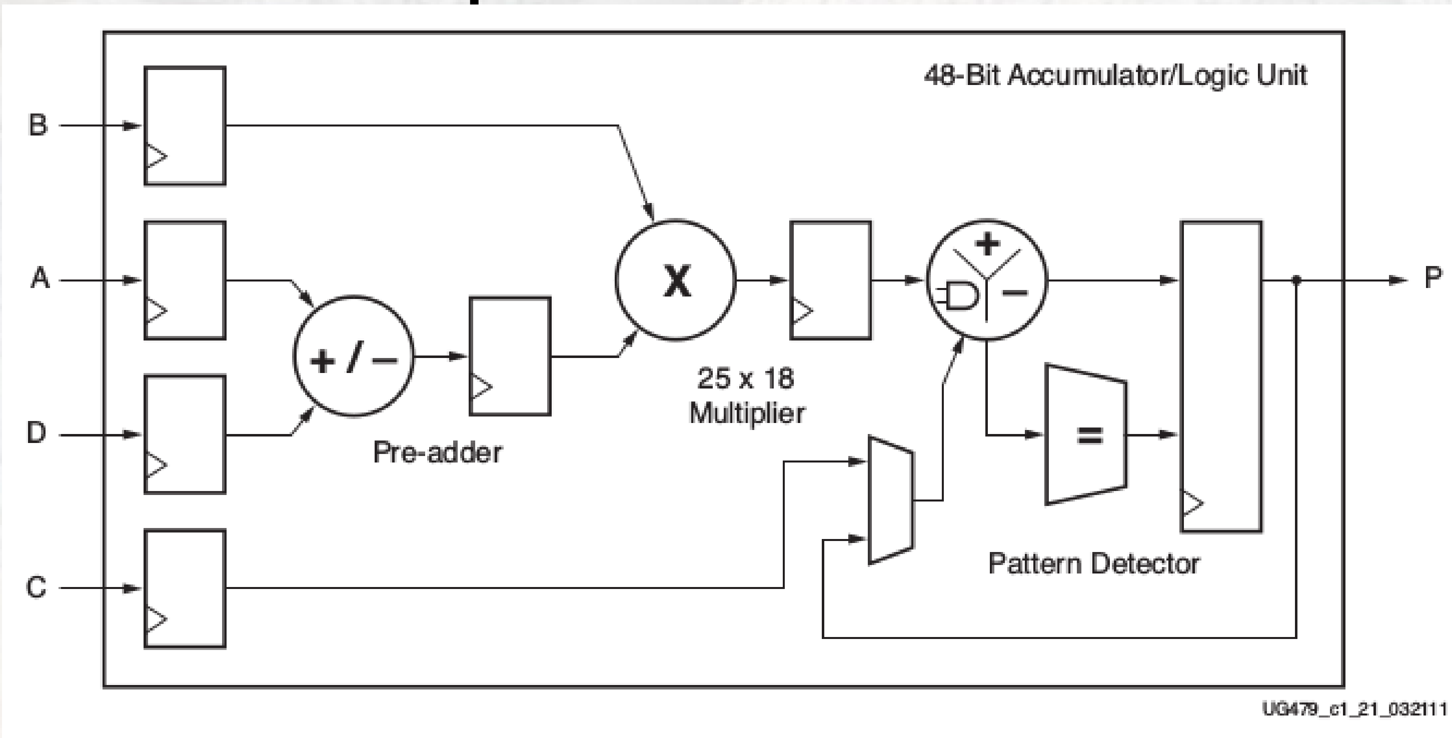
OMTF Implementation structure



GPU implementation



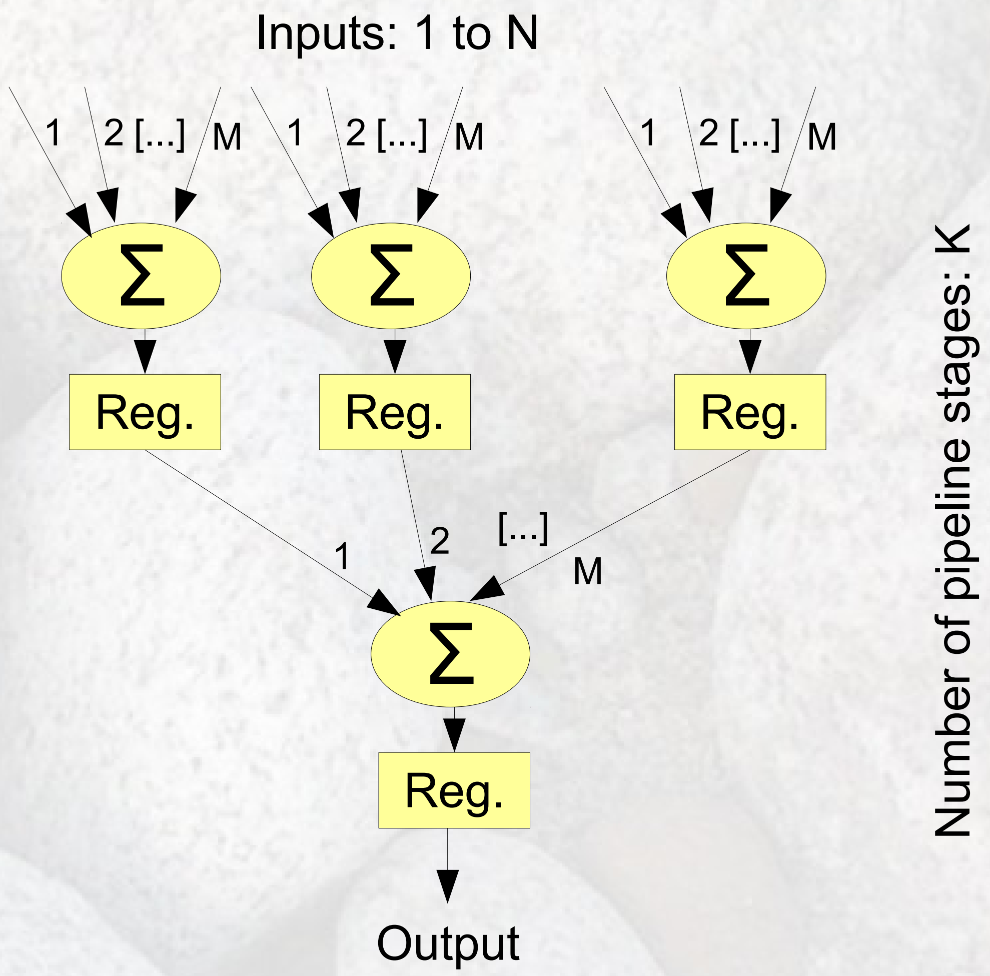
Hardware optimizations – DSP blocks



- The Virtex 7 FPGAs offer DSP48E1 blocks able to perform arithmetic operations on 48-bit values
- A special SIMD mode is available to perform calculations on shorter words (24-bits or 12-bits)
- Unfortunately the SIMD mode is not fully supported by synthesis tools. Therefore its usage significantly complicates the HDL description.
- Additionally in the pipelined architecture it is necessary to buffer data bypassing the DSP blocks, which finally resulted in not using of those blocks (situation may change when new version of Vivado tools is available)

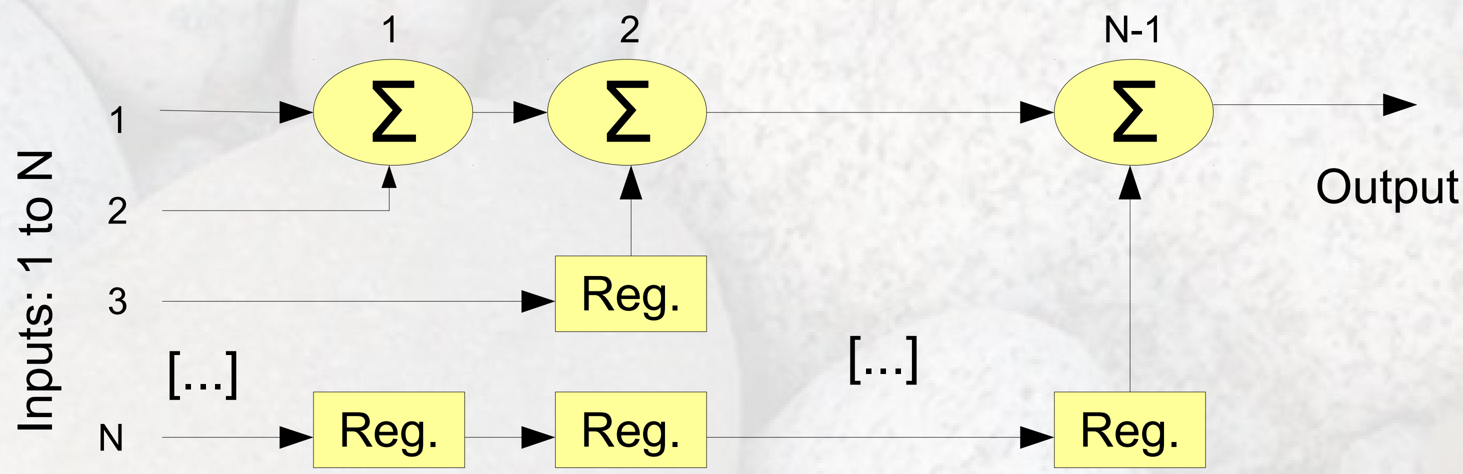
Parameterized processing blocks Pipelined tree adders

- The number of inputs „N” and the number of inputs handled in each elementary adder affects:
  - The complexity of combinational logic between pipeline stages (and therefore the maximum clock frequency)
  - The number of stages needed, and, therefore, the overall latency of the adder
- The optimum must be found.



$$M^{K-1} < N \leqslant M^{K}$$

Parameterized processing blocks Simple pipelined adders

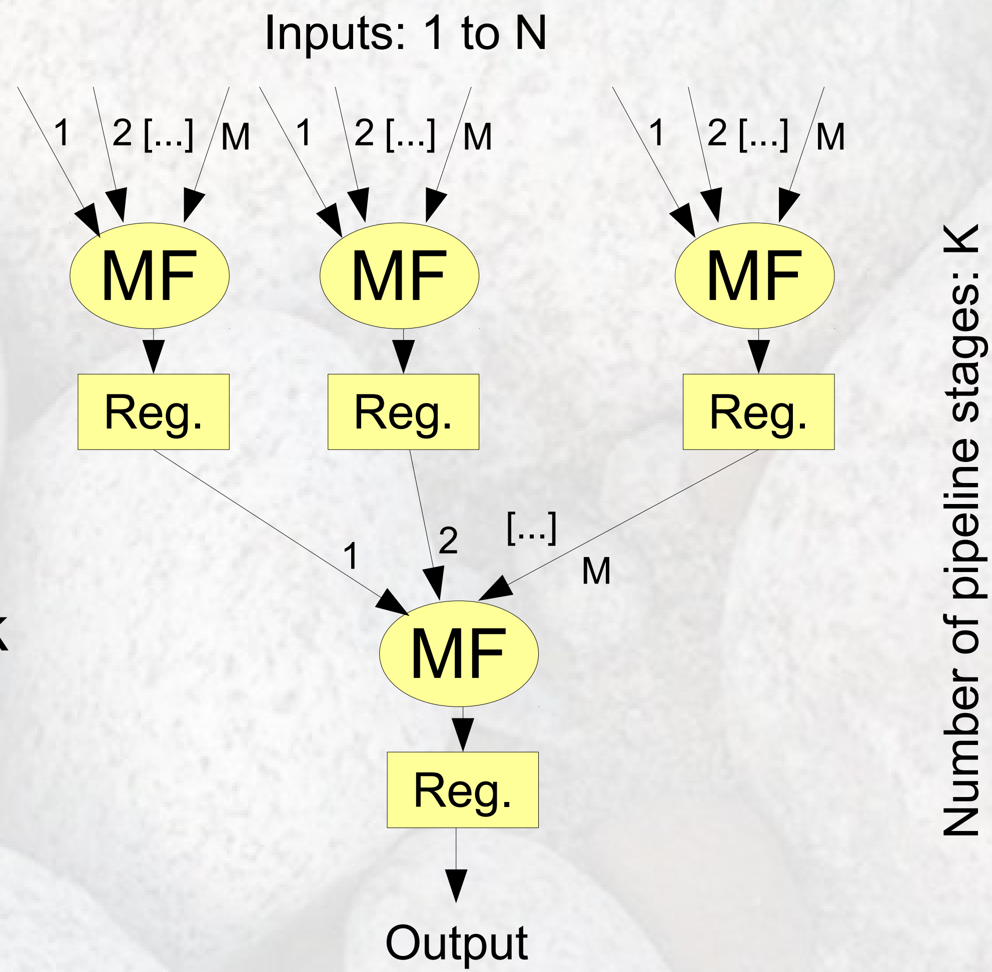


Number of pipeline stages: K=N-1

- This design provides very high clock frequency, but the number of stages and number is proportional to the number of inputs „N"
- The number of needed registers is proportional for to N*N
- Therefore this design is appropriate only for smaller values of N

Parameterized processing blocks
Pipelined tree maximum or minimum finders

- The number of inputs „N” and the number of inputs handled in each elementary maximum (or minimum) finder:
  - The complexity of combinational logic between pipeline stages (and therefore the maximum clock frequency)
  - The number of stages needed, and, therefore, the overall latency of the finder
- The optimum must be found.



$$M^{K-1} < N \leqslant M^K$$

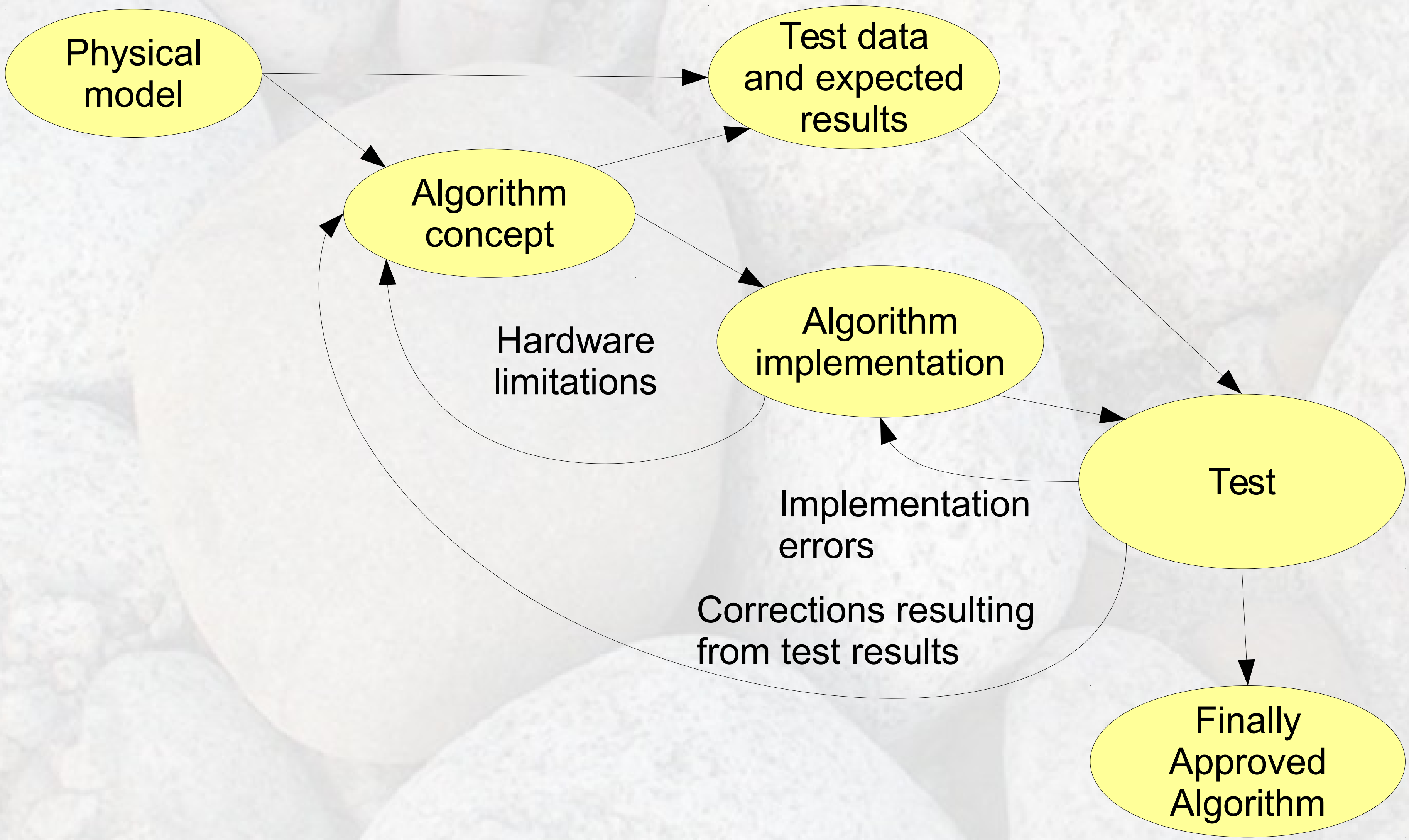
Synthesis of the algorithm

- The OMTF algorithm together with the surrounding infrastructure has been successfully synthesized with Xilinx Vivado tools
- Resource consumption:
  - Slices used: 94.48% (102319 from 108300 available)
  - Slice LUTs: 64.88% (281040 from 433200 available)
  - Slice Registers: 32.87% (284752 from 866400 available)
  - Block RAM tiles: 51.22% (753 from 1470 available)

Testing and optimization of the algorithm

- Testing and optimization of the algorithm is performed basing on the results of analysis of the test data generated in physical simulations (in CMSSW)
- Tests were performed both in **HDL simulations** and **in the real hardware** (with artificially generated signals)**.**
- Basing on those results we were able to detect some hidden bugs (especially in the handling of rare special cases), and also to optimize some parameters.
- For the above activity, it was essential, that the OMTF algorithm is implemented with the **high level, structured and parameterized behavioral VHDL description**
- Many optimizations change the latency of certain paths in the OMTF implementation. To avoid problems with data incorrectly aligned in time a special method for automatic latency equalization has been developed and will be used in future versions of OMTF

Development process



Results and conclusions

- The OMTF algorithm is still being tested, but is already functional (RPC muons are visible in uGMT)
- The work at implementation and testing of the OMTF has shown that it is a significant multidisciplinary problem requiring the cooperation of professionals from both areas: physics and electronics.
- In the development of the current form of the algorithm especially valuable was participation of persons who combined knowledge about physical phenomena with understanding of possibilities and limitations of the electronic hardware.


Thank you for your attention!