L1 track trigger for the CMS HL-LHC upgrade using AM chips + FPGA

CTD/WIT 2017

6-9 March 2017

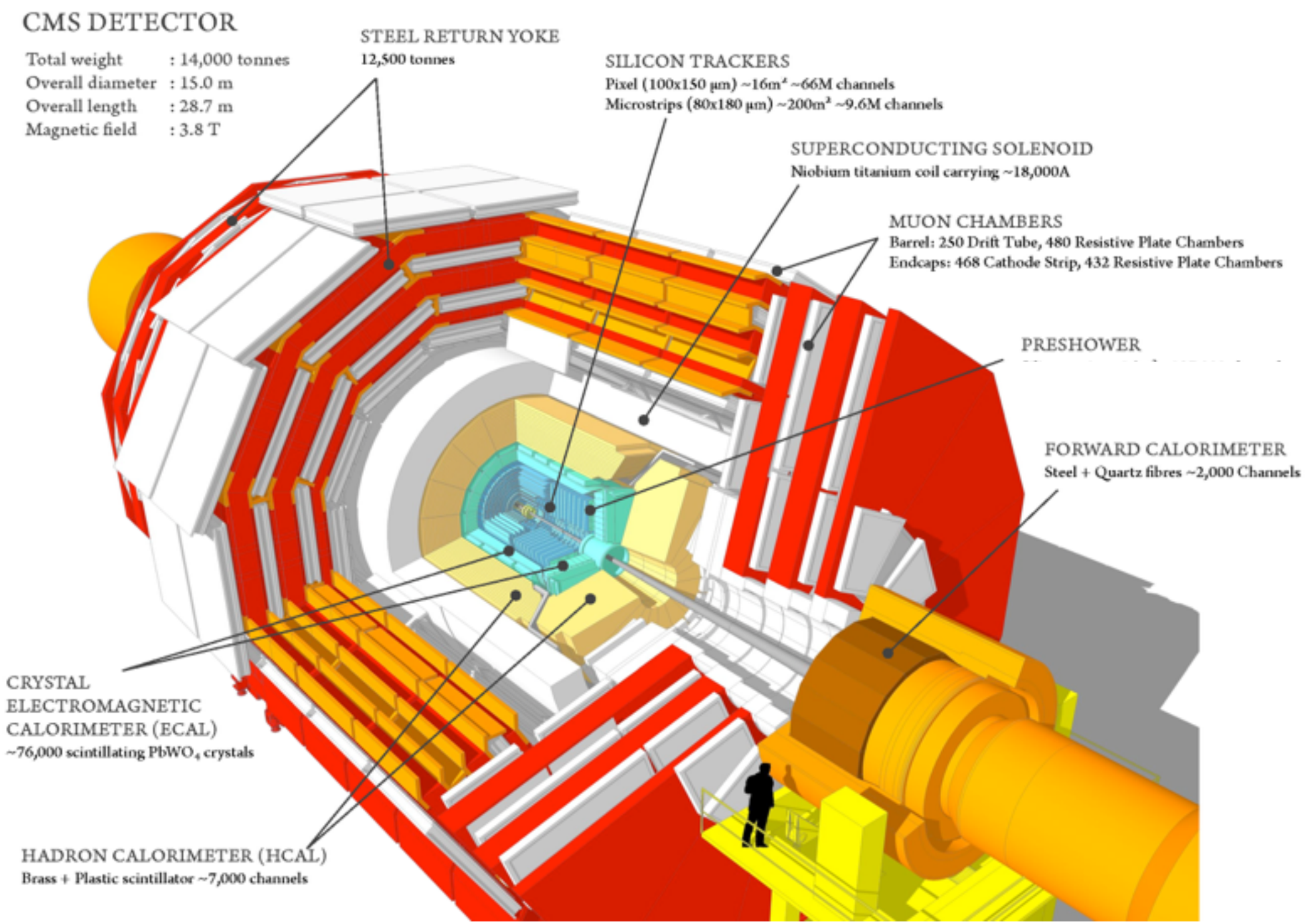
Giacomo Fedi

(Università di Pisa & INFN Pisa)

On behalf of CMS collaboration

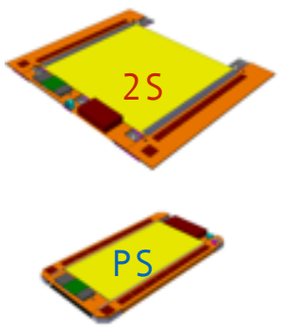
This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 317446, and from PRIN MIUR for H-TEAM project under grant agreement n° 957

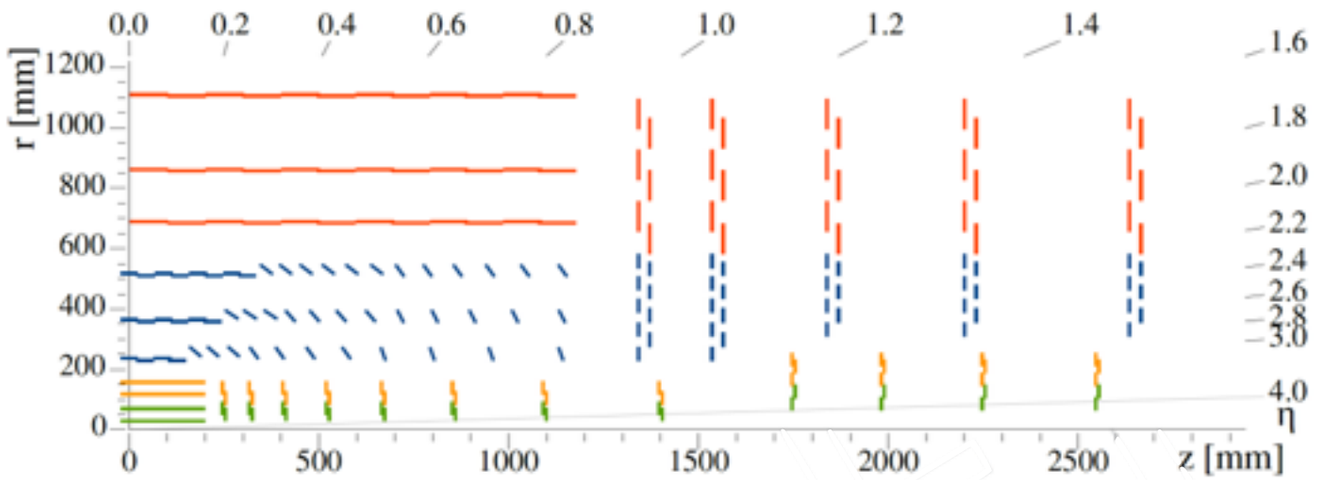
Introduction: CMS detector

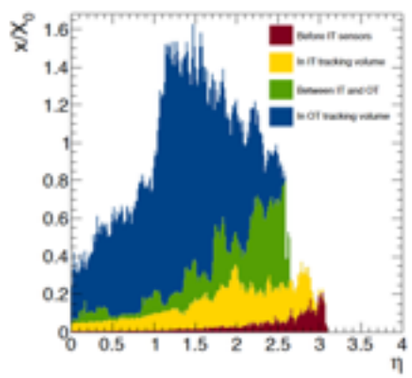


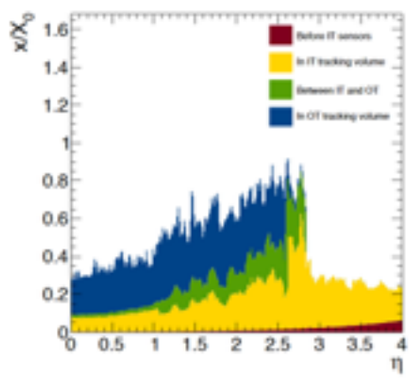
Introduction: tracking

HL-LHC phase: possible CMS tracker layout (Tilted Geometry)





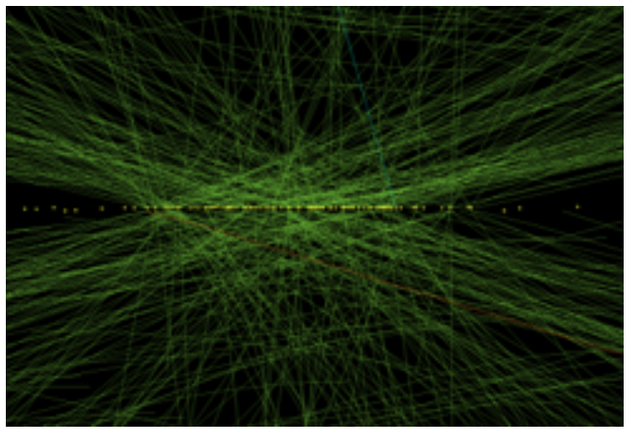




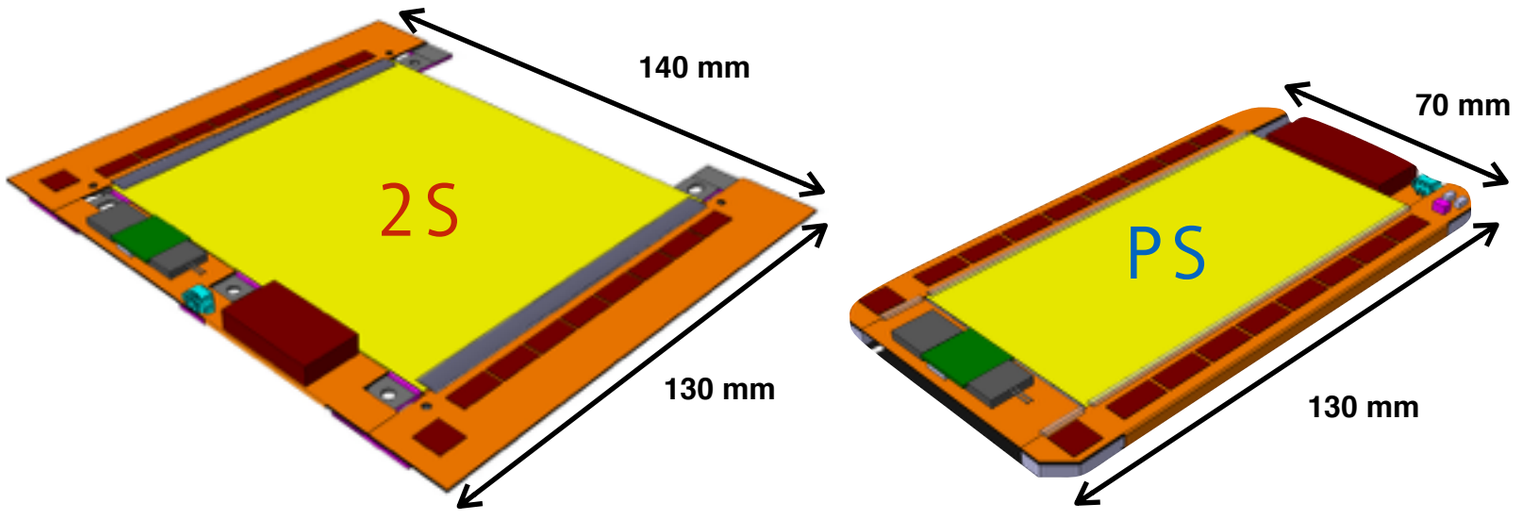
For this study we used the Flat Geometry

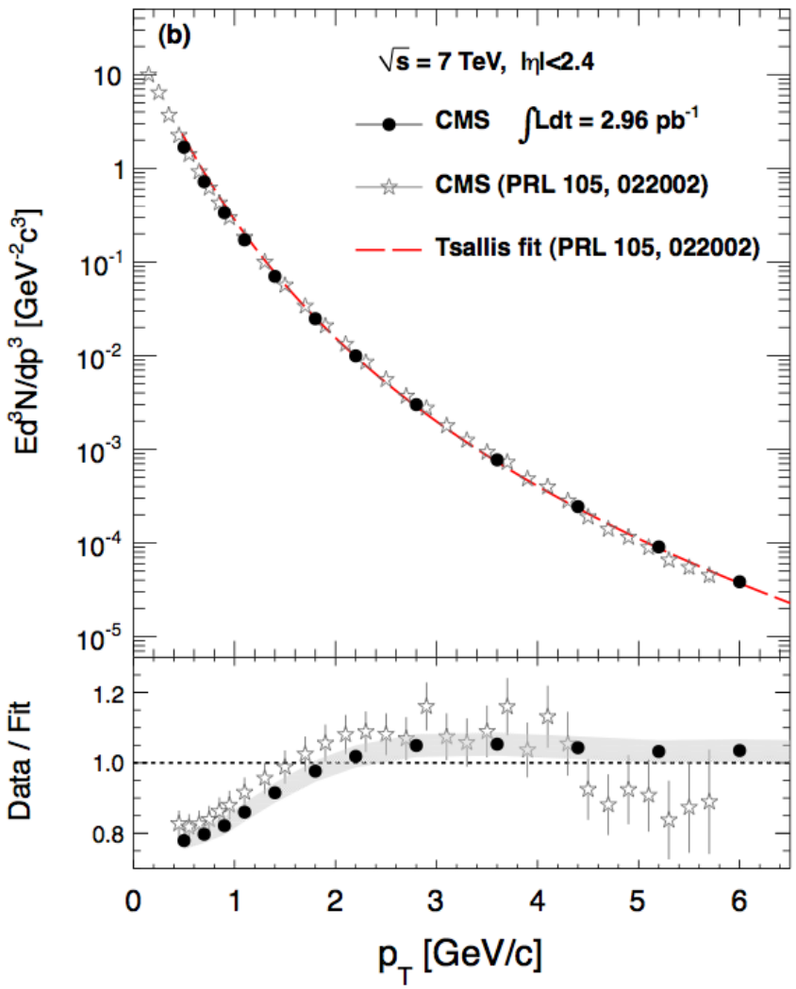
Challenges of tracking at the HL-LHC (2026):

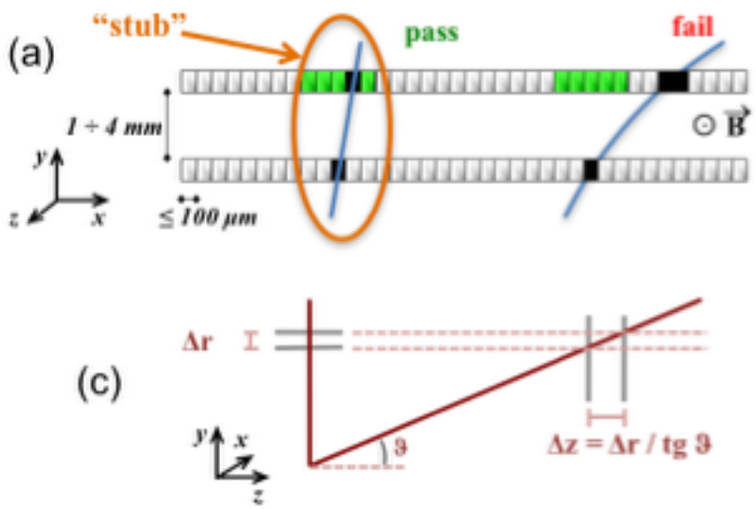
- **140-200 piled up p-p collisions**
- **A single bunch crossing generates thousands of particles**
- **Reconstruction of particle trajectories is challenging even offline**



In situ data reduction







With two superimposed sensors per module we are able to filter low p_T tracks (>2 GeV). Big data transfer reduction O(10)

INFN Istituto Nazionale di Fisica Nucleare

Goal

Development of a fast track reconstructor system, which can be used in the L1 trigger of CMS at the HL-LHC with the final latency of ~4 μs

Short term goal (End of 2016):

- Develop a small prototype (demonstrator) with the current technologies
- The prototype will demonstrate that scaling the technologies and the system we will be able to reconstruct track at L1 within 4 μs

Long term goal (After 2016):

- R&D on the system
- Construction of the system for the Phasell of LHC (~2026)

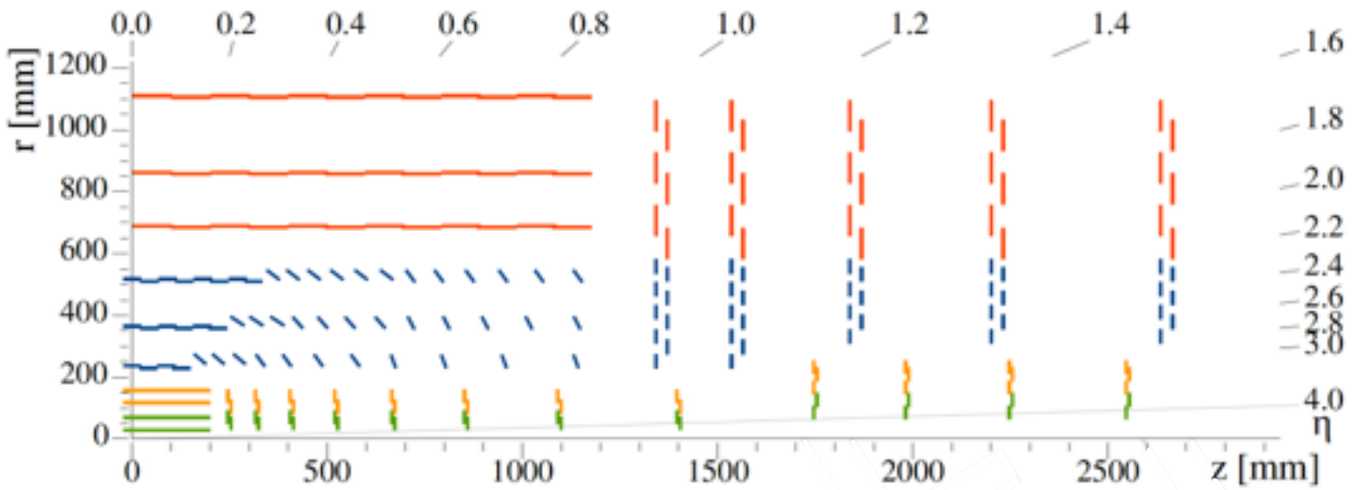
Main aspects of this approach:

- Usage of the associative memory (AM) ASICs
- Usage of latest generation FPGAs
- Usage of the Advanced TCA (ATCA) based boards for high bandwidth data processing
- Usage of fast serial links

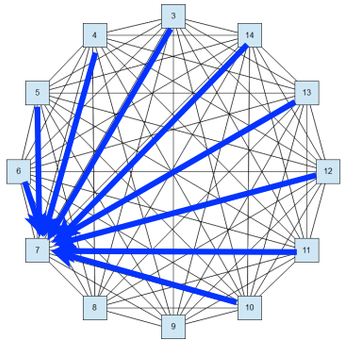
Overall project

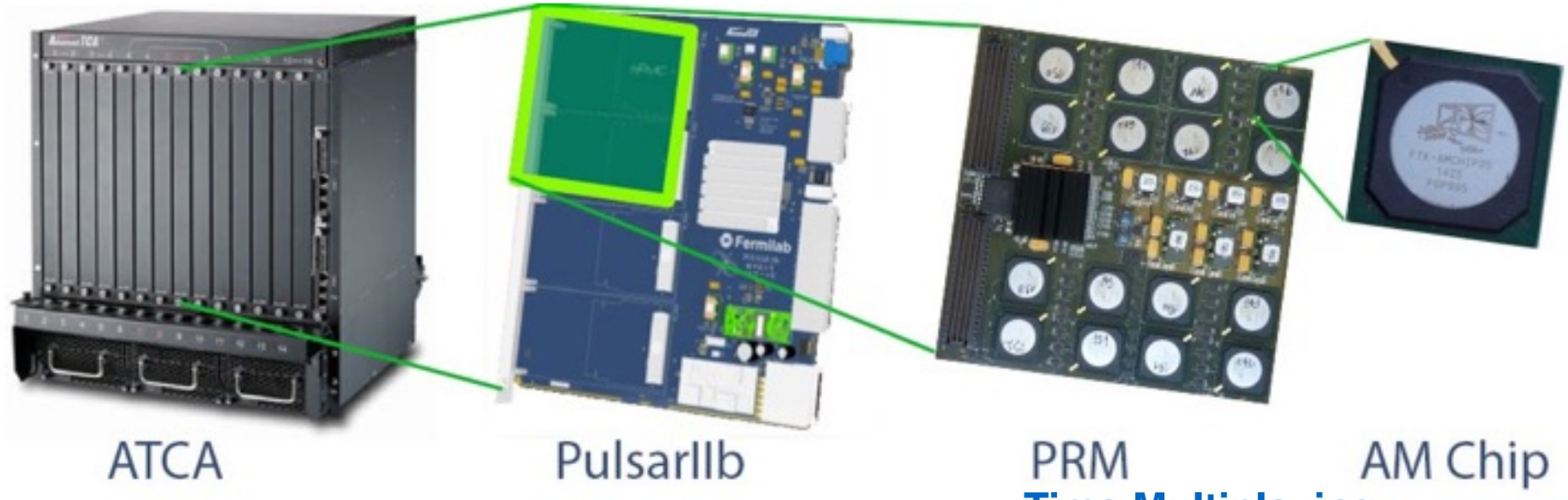
Hits from 48 trigger towers: 6 (η) x 8 (ϕ)

Each trigger tower inputs data: ~300 fibers ~ 2 Tbps



1 shelf per trigger tower in case of the demo, 1 shelf shared between 2 TT in case of final system





40G full-mesh backplane

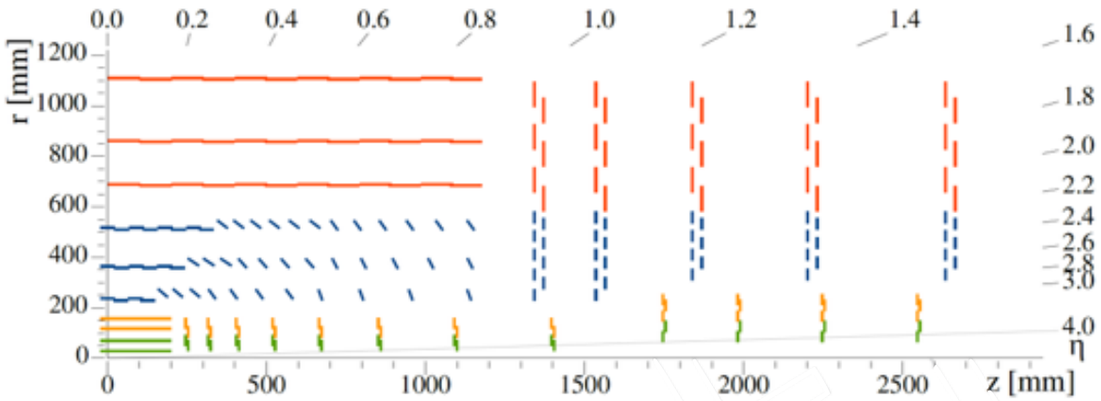
Developed by FNAL
Virtex7 FPGA
80 GTH up to 13.1 Gbps:

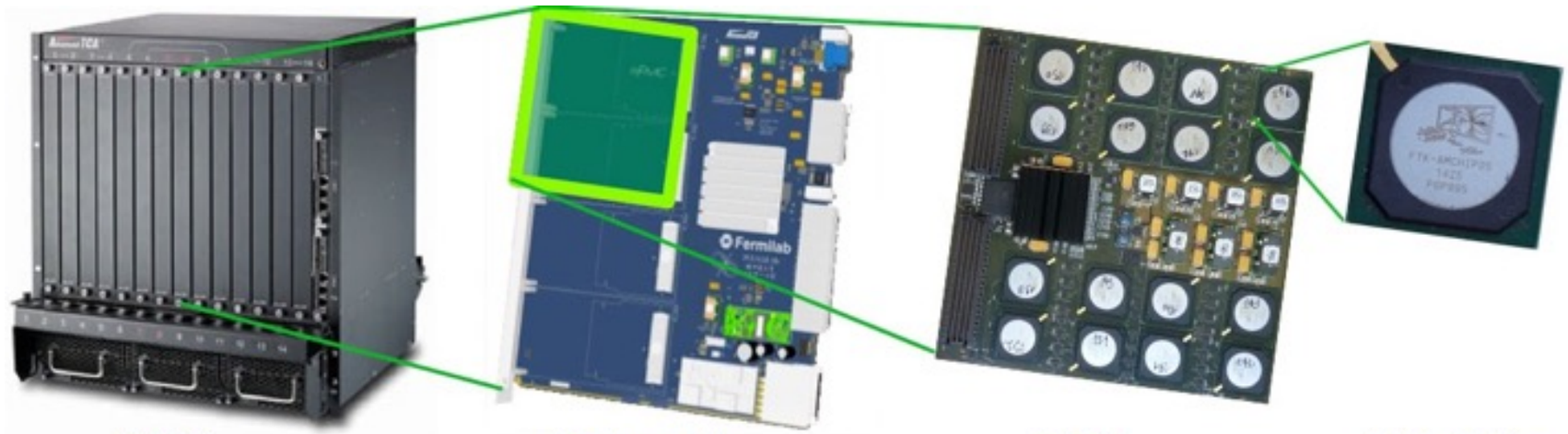
- **40 to the RTM**
- **28 to backplane**
- **12 to FMCs**

Time Multiplexing: each PRM can process one bunch crossing every 250 ns (x10) or 500 ns (x20)

Latency definition

Tracker front-ends





Data delivery latency

PRM ↔ **PRM-AM**

Pattern matching

Track fitting

PulsarIIb ↔ **PRM-FPGA**

INFN Istituto Nazionale di Fisica Nucleare

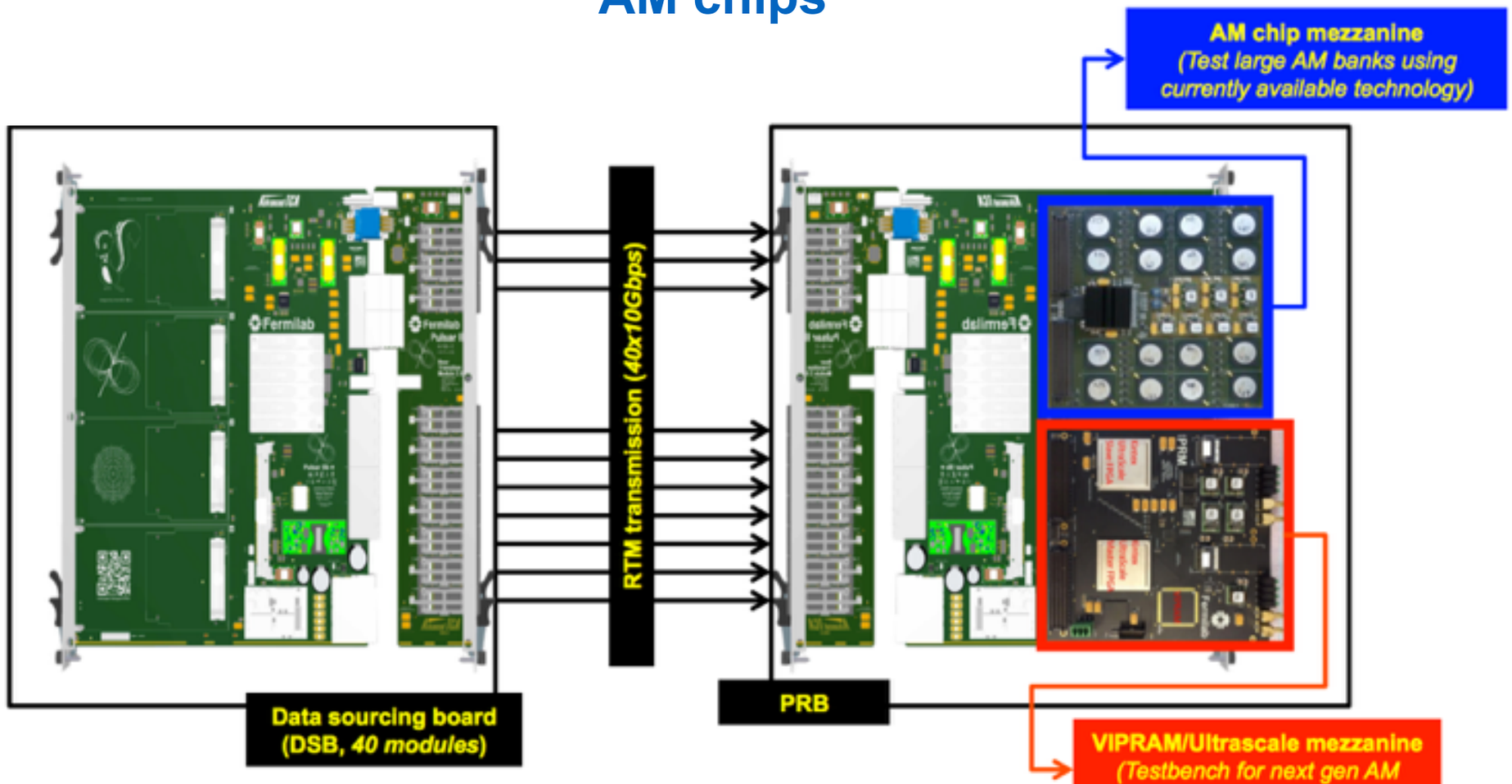
Demonstrator

Scaled to the current technology and to a single trigger tracker tower

1st ATCA shelf:
10 Pulsar IIb boards used as data source

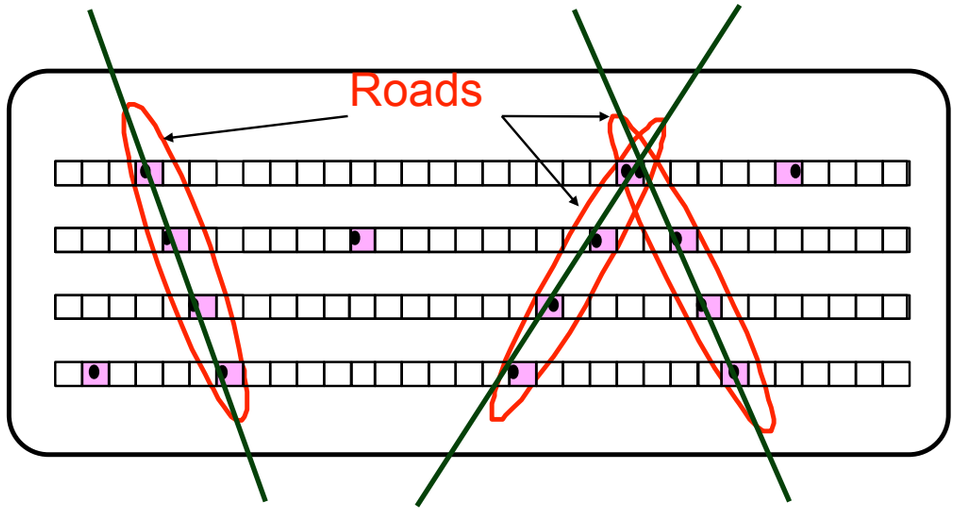
2nd ATCA shelf:
10 Pulsar IIb boards used as system test bench

2016: full prototype built with two ATCA shelves and 20 Pulsar IIb boards equipped with PRM mezzanines and AM chips



Associative Memories

Main feature: avoid stub combinatorics

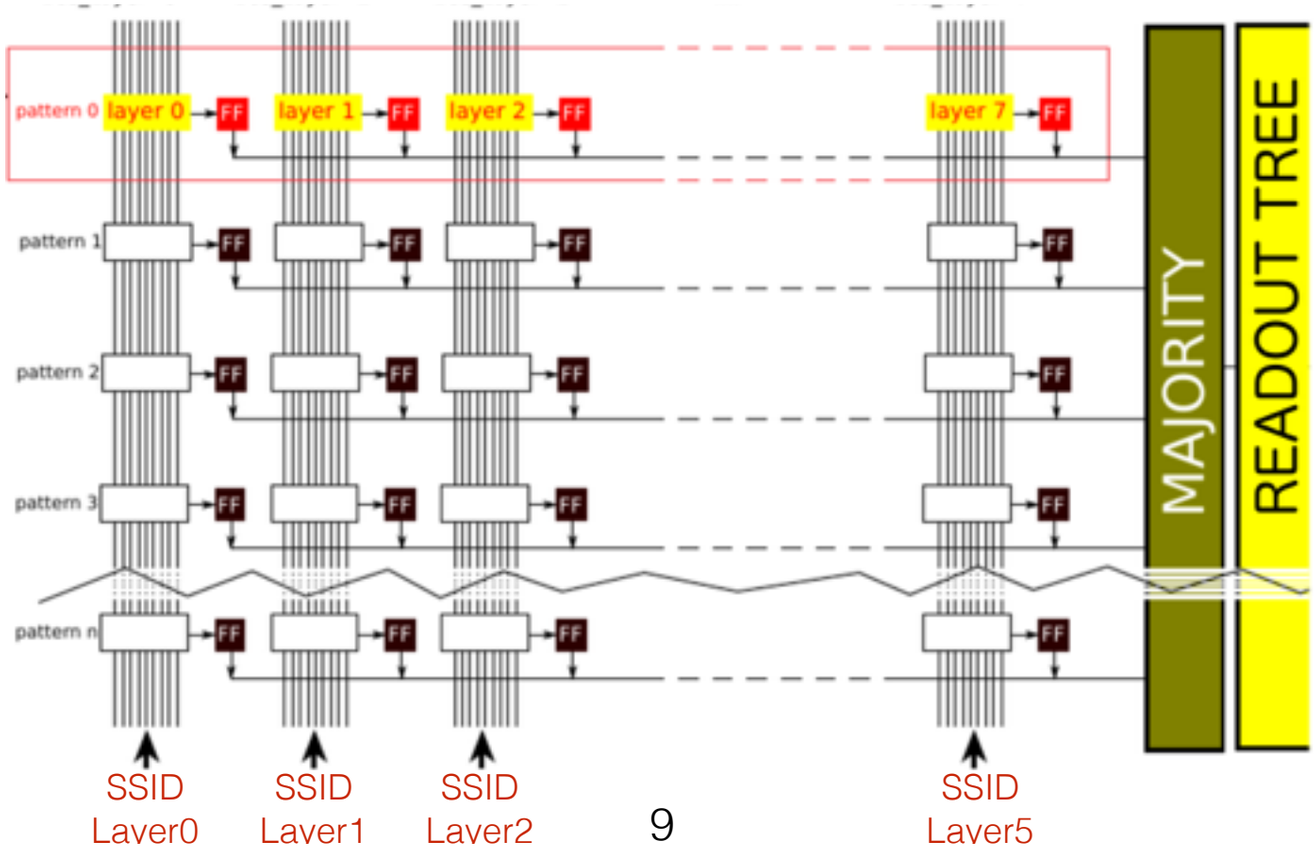


Configuration

- The combinations of coarse position stubs (SSIDs) generated by the tracks are stored inside the AM chip (**Pattern bank**)

Operation

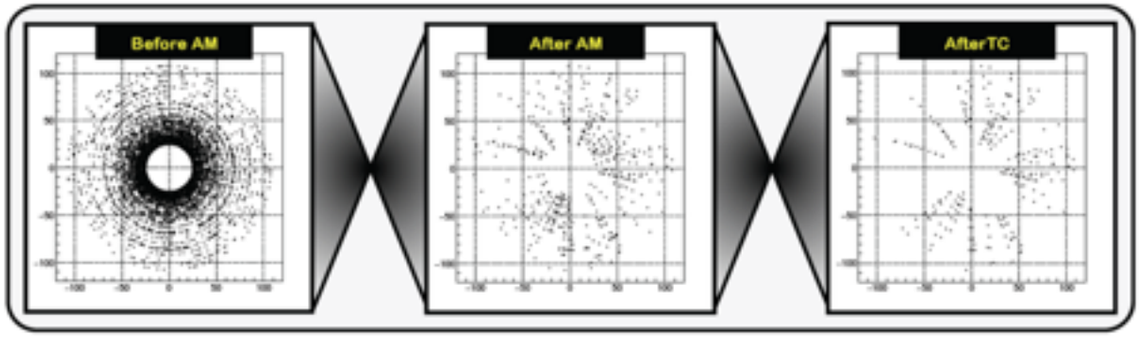
- The SSIDs of each tracker layers are sent to the AM input buses
- The matched road addresses are sent as AM chip output

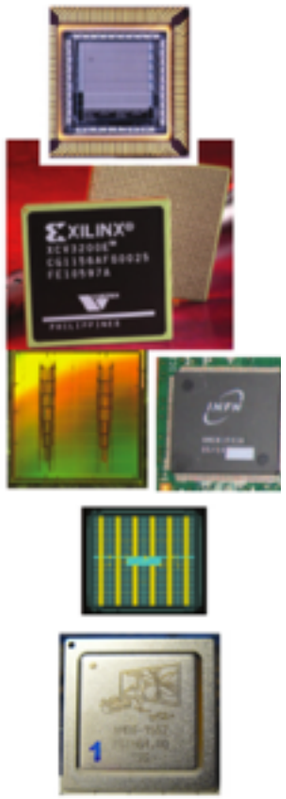


Features

- Majority logic
- DC bits/layer
- Ordering of the output road
- 8 available layers, 6 used in CMS

Planar AM chips

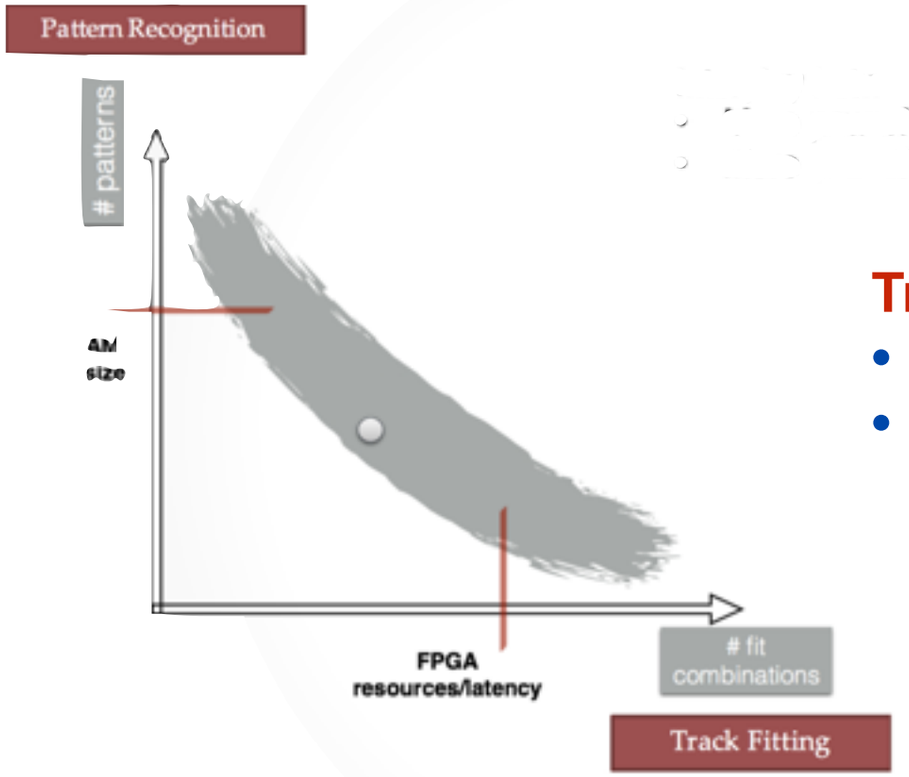




Version	Year	Patterns	Working Frequency (MHz)	Power (W)	Package	Technology	Area (mm^2)
AM01	1992	128			QFP	700 nm	
AM02	1998	128			QFP	350 nm	
AM03	2004	5000	40	1.26	QFP	180 nm	100
AM04	2012	8000	100	3.7	QFP	65 nm	14
AM05	2014	3000	100		BGA	65 nm	12
AM06	2015	128k	100	3	BGA	65 nm	150
AM07	*2016*	*16k*	*200*	*0.1-0.2*	*BGA*	*28 nm*	*10*
AM08	*2018*	*16k*	*250-400*	*0.1-0.2*	*BGA*	*28 nm*	*10*

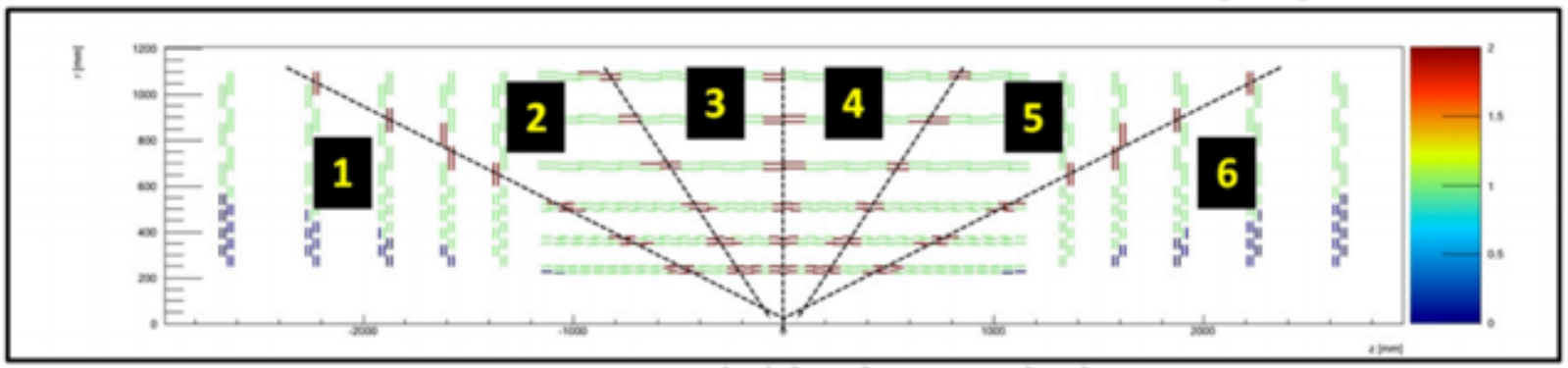
AM05: technology testing chip
AM06: production chip, used in FTK ATLAS track trigger
AM07: 28 nm technology testing chip, tests are ongoing
AM08: testing chip in designing phase, improved data link

Pattern bank optimisation



Trade-off between AM size and FPGA resources/latency:

- **More patterns in the AM chip -> less demand on FPGA**
- **More powerful FPGA -> less patterns in the AM chip**



AM chip pattern bank size per tower:

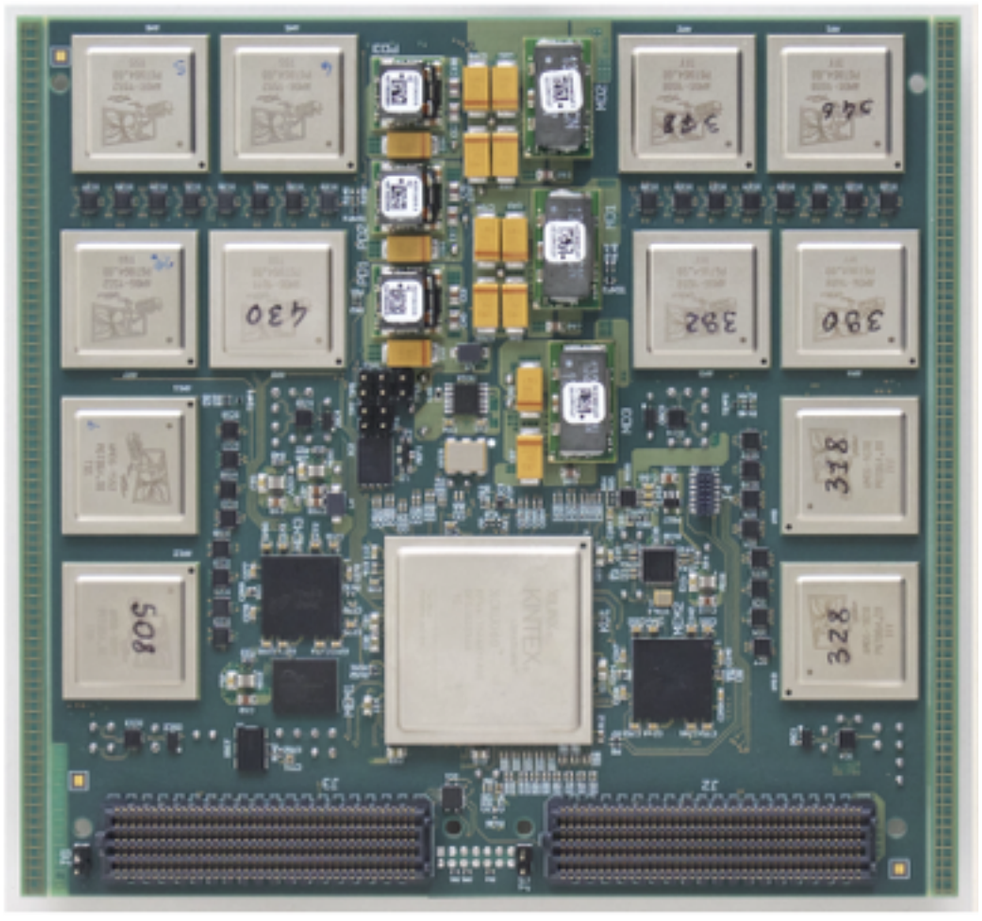
- **0.5M patterns for barrel**
- **1M patterns for hybrid**
- **0.5M patterns for endcap**

Number of strips in a SSID per layer

INFN Istituto Nazionale di Fisica Nucleare

INFN PRM

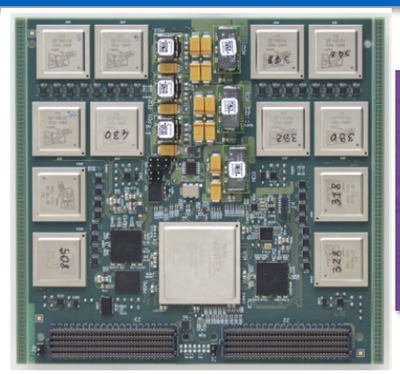
INFN PRM06



- 3 PRMs available and tested
- 12 PRMs in production for CMS/ATLAS groups
- FPGA: Kintex Ultrascale 060
- 12 AM06: total of 1.5 Mpatterns
- GTH maximum speed: 12.5 Gpbs
- Double RLD3RAM 1 Meg x 36 x 16 Banks, 1066 MHz DDR operation
- Flash memory

PRM06: designed to be used in the demonstrator, with its 1.5 million pattern bank can cover a full trigger tower pattern bank (0.5-1M patterns)

PRM06 HW validation

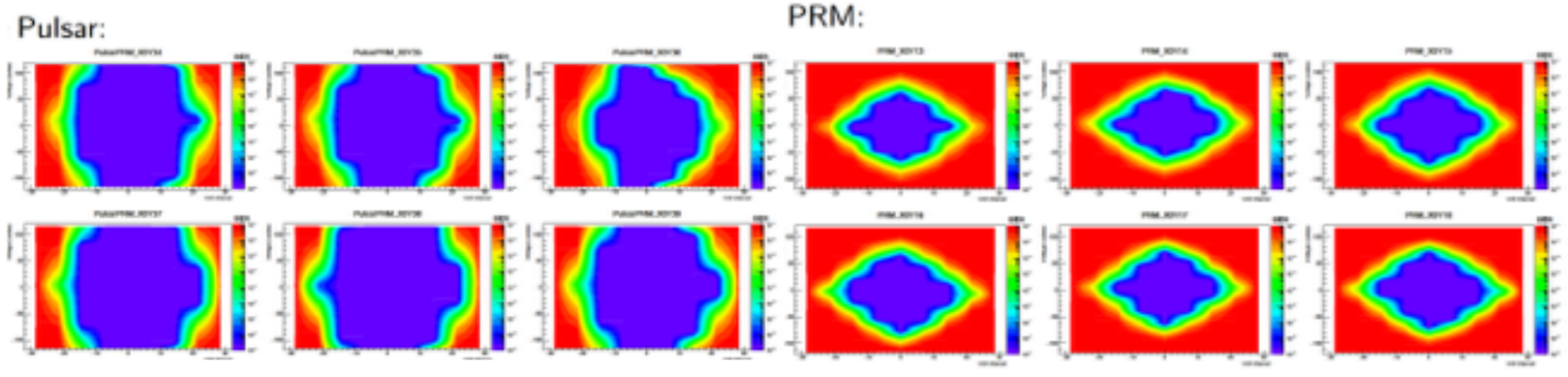


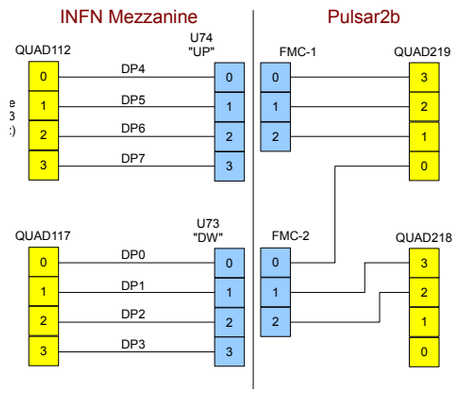
Prototype delivery date:	June 2016
Prototype validation date:	July 2016
2 additional PRM06s:	July 2016
Pulsar-PRM06 testing date:	September 2016

PRM06 hardware tested and working as expected

- **GTH links: IBERT PRBS7 on all links**
- **AM06 communication and configuration: JTAG communication (each AM06 tested before to be mounted on PCB: serdes' and bank memory with built-in test)**
- **Serdes links: PRBS on all links**
- **LVDS links going to FMC connector: loopback on evaluation board, static test**
  - **All the links with the Pulsar and they are all ok**
- **RLD3RAM: using Xilinx tools, checked the reading and writing**
- **External flash memory: using OpenCores IP we tried basic communication**
- **Test of the GTH links between PulsarIIb and PRM06**

IBERT eyescans @12.5Gbps



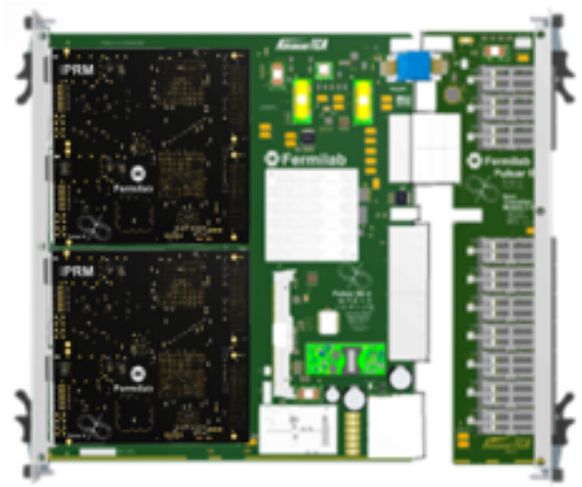


Data distr. in the demonstrator

Two ATCA shelves fully loaded with Pulsar IIb boards (20)

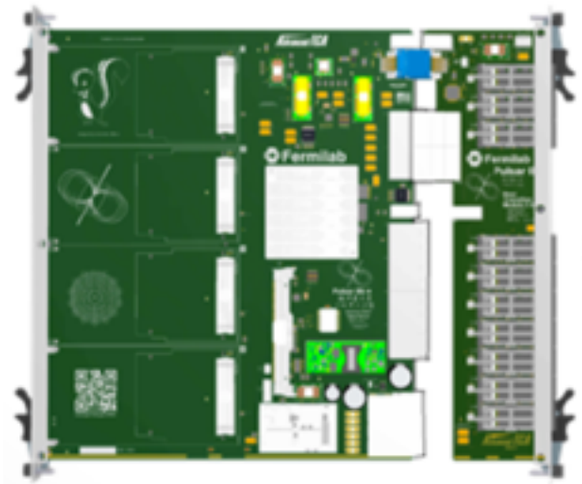
Pattern Recognition Boards (PRB) shelf

- One trigger tower
- 10 Pulsar IIb
- Some boards equipped with PRM



Data Source Boards

- Emulated the output of ~400 tracker modules
- 10 Pulsar IIb
- 100 QSFP+ optical links



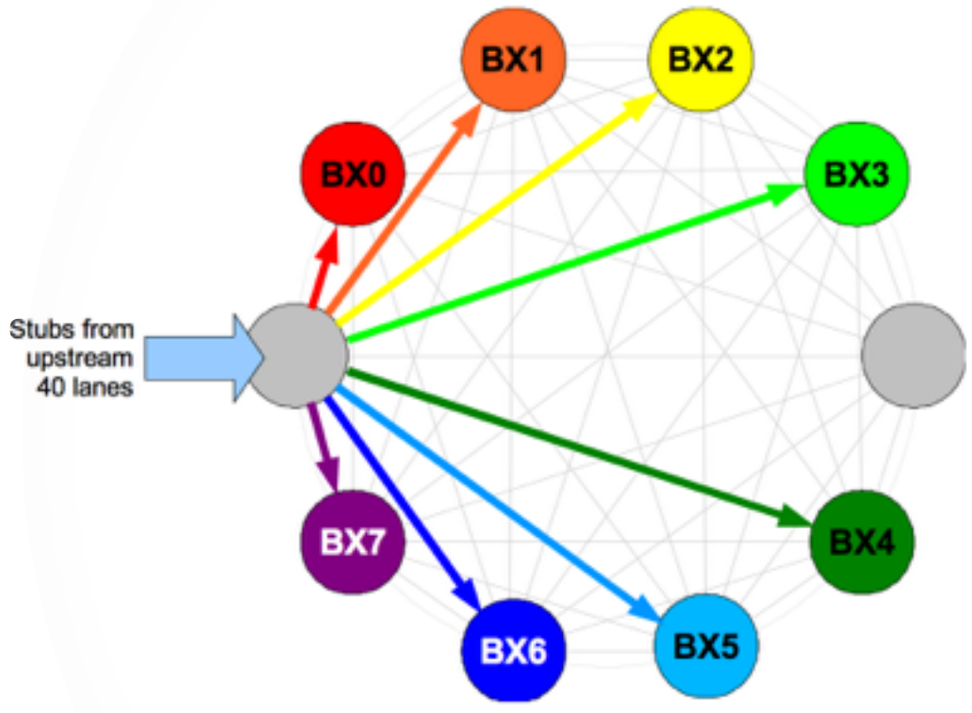
Rear view

Total of 12x40=480 channels, 4.8 Tbps, extensively tested running in sync, 64/66 as link protocol

Close collaboration: FNAL, Northwestern, U. Florida, Texas A&M, SPRACE/UERJ and Peking with support of LPC

Data distr. on full mesh backplane

We exploit the full meshed architecture of the ATCA shelf for data delivery

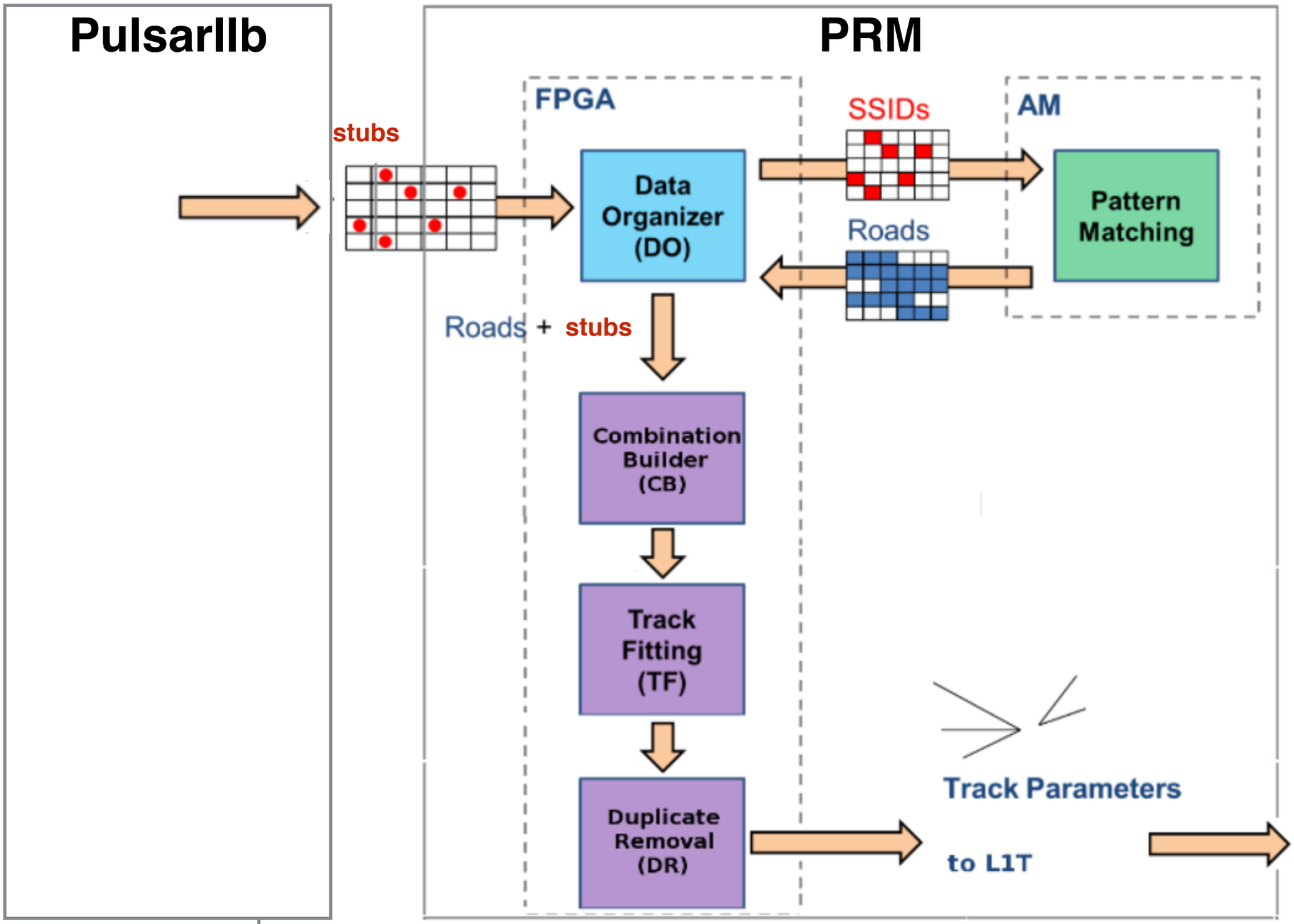


- Each Pulsar IIb receives stubs on 40 optical links from the rear module
- Stubs arrive in a “train” which contains stubs for up to 8 bunch crossings (BX)
- New train every 200 ns
- Pulsar IIb FPGA sorta the stubs BX and sends to 7 or 8 neighbors over the backplane
- Backplane transfer must complete in 200 ns
- Each board can send up to ~100 stubs to each neighbor
- Full mesh channels are 2x10 Gbps

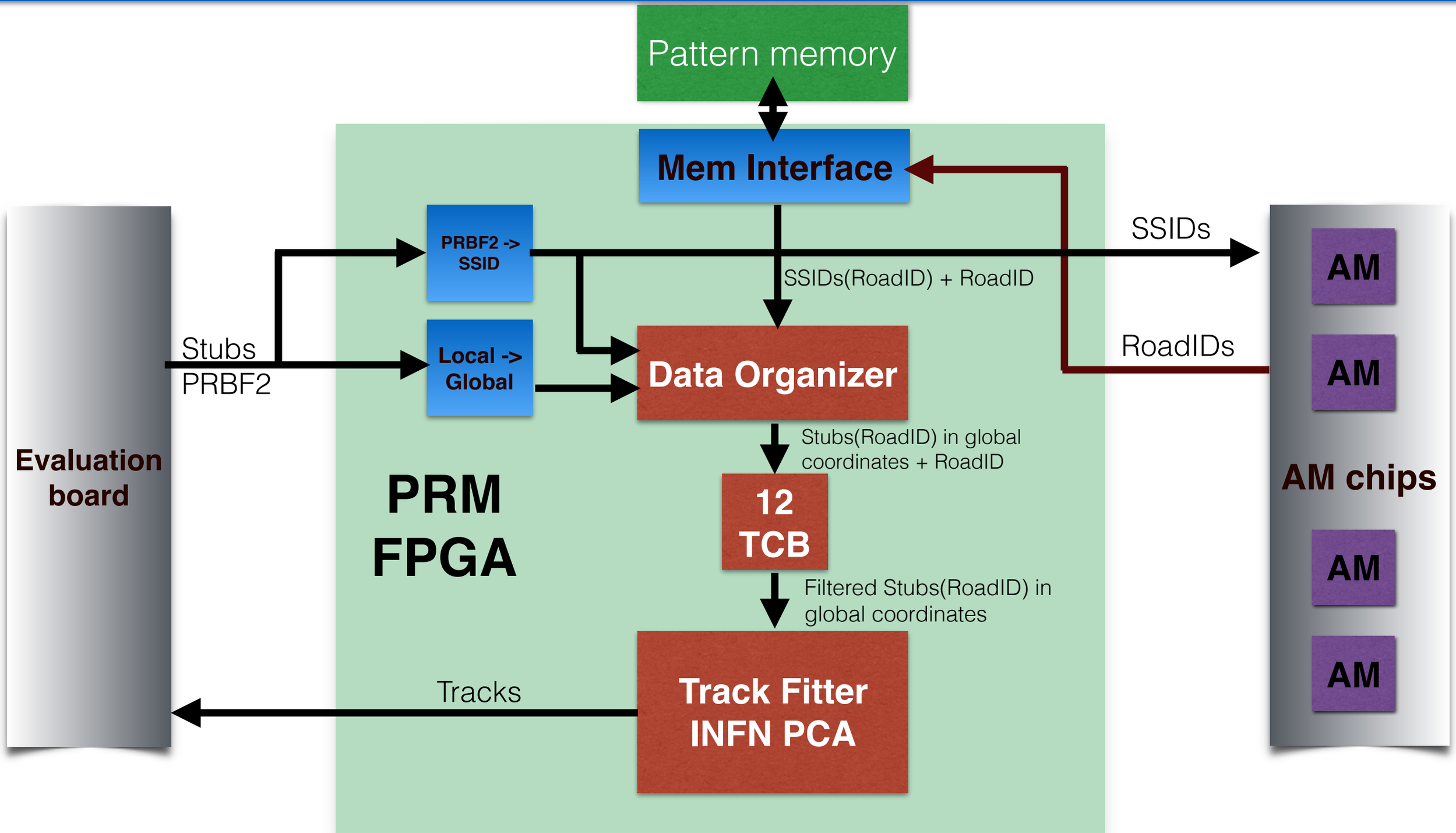
Data delivery latency:

- First stub to PRM at 1.2 μs
- Last stub to PRM at max 1.7 μs

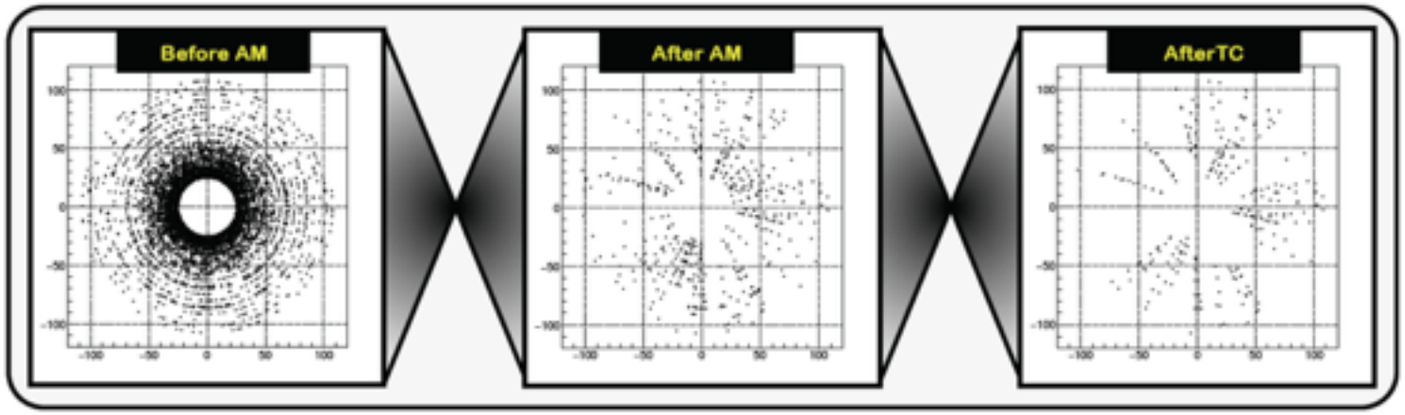
Data flow in the PRM



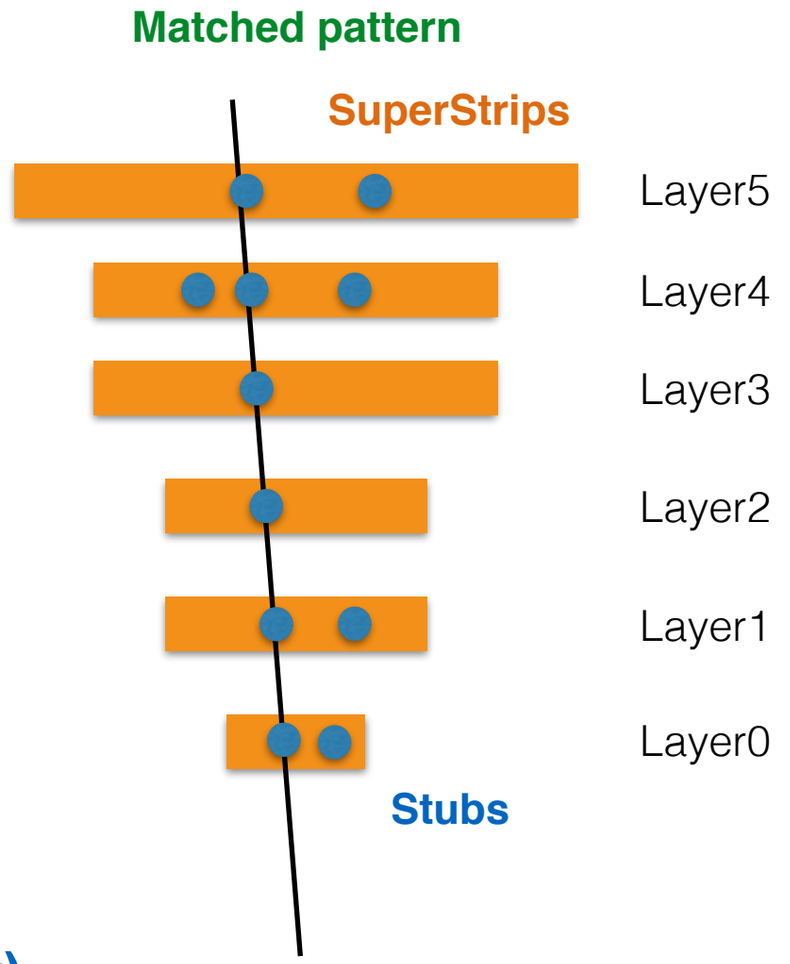
PRM06 FW data flow



Dealing with residual combinatorics

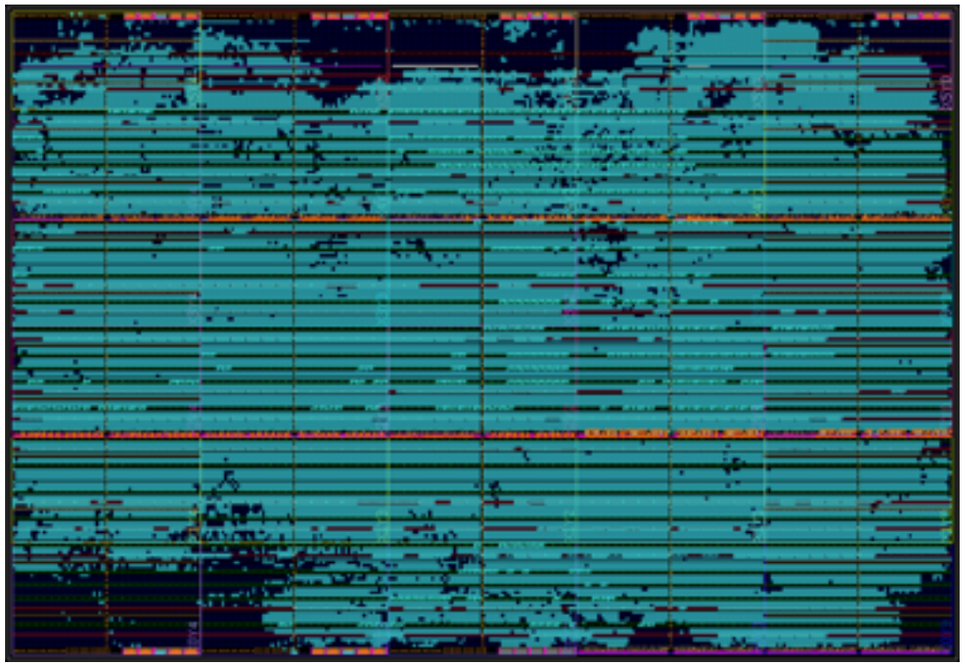


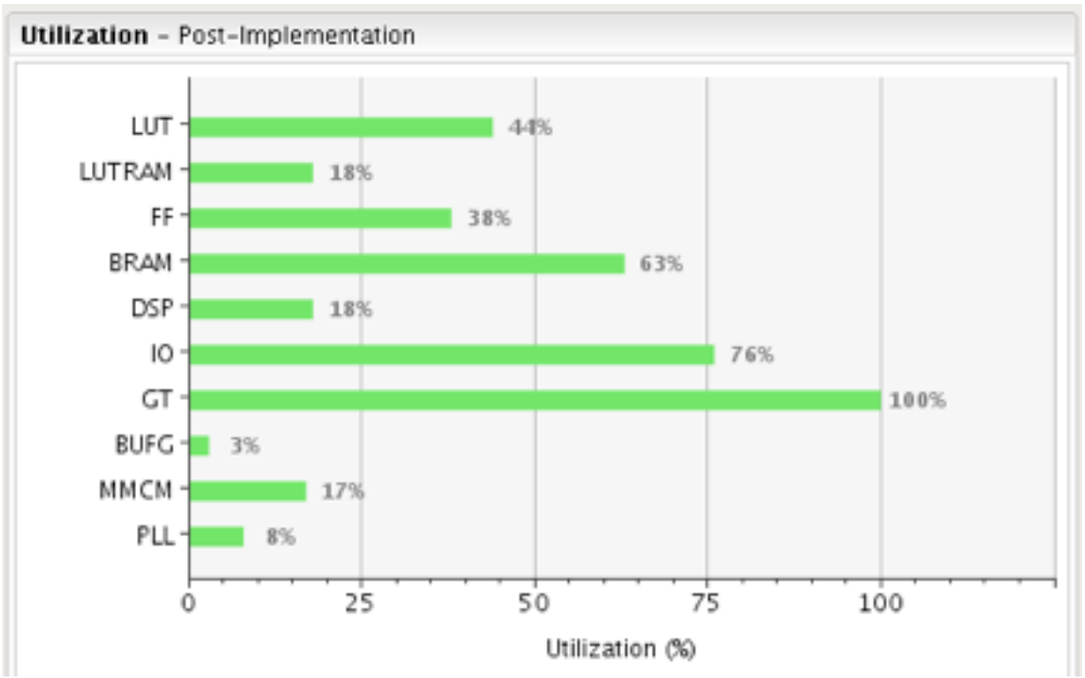
- **Combination builder (CB)**
  - Returns all the combinations with one stub per layer
- **All combination builder (ACB)**
  - Returns all the combinations with one stub per layer adding the permutations of 5 stubs out of 6 to recover tracking efficiency
- **Track candidate builder (TCB)**
  - Returns one track candidate (5 or 6 stubs, one per layer)
  - Builds the track candidate using pairs of stubs from the first three as seeds, and checking the compatibility of the other layer stubs. The best combination of stubs is send to the track fitter
  - Pros: less fakes/duplicates, low latency in case of high stub combinatorics
  - Cons: suboptimal in case of low stub combinators



PRM06 FW features

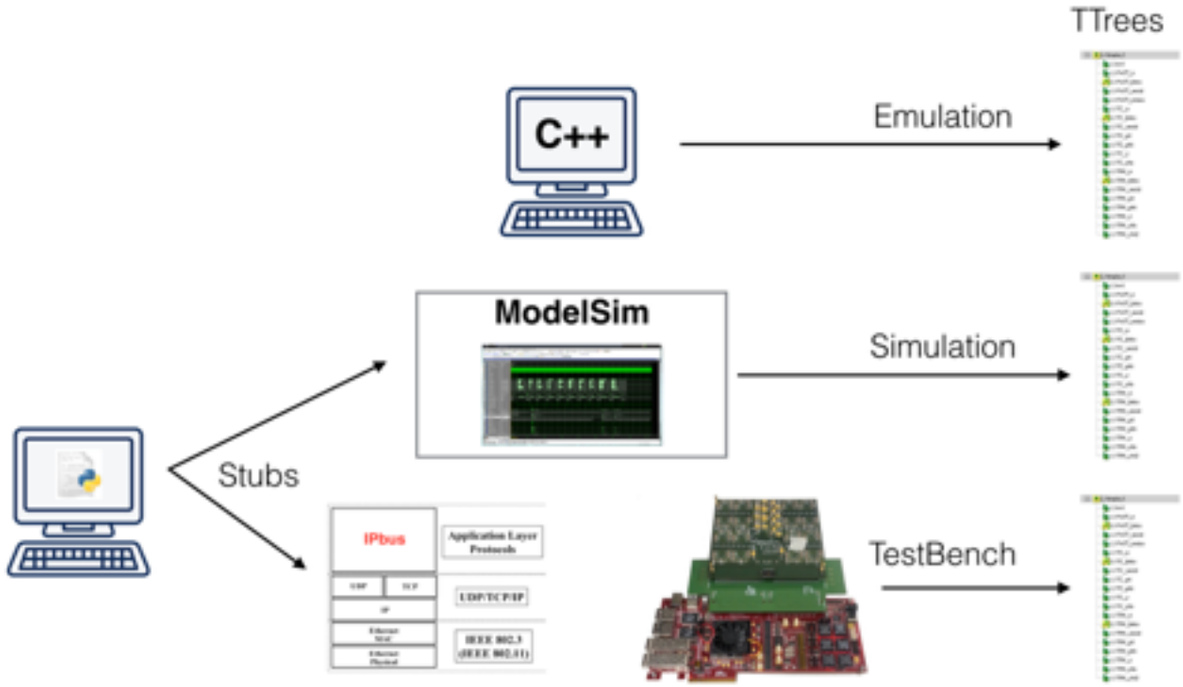
- **Features**
  - **Multiple clock domains: DO and TF @ 200 MHz, TCB @ 100 MHz**
  - **Per layer DC bits**
  - **Full trigger tower coverage**
  - **Multiple AM chip handling (up to 12)**
  - **Missing layer (5/6)**
- **Resources: about 50% of the Kintex KU060 resources have been used. It is possible to add 3 more TF and one more DO.**
- **Power consumption with the loaded FW and 12 AM chip configured: 37 W**



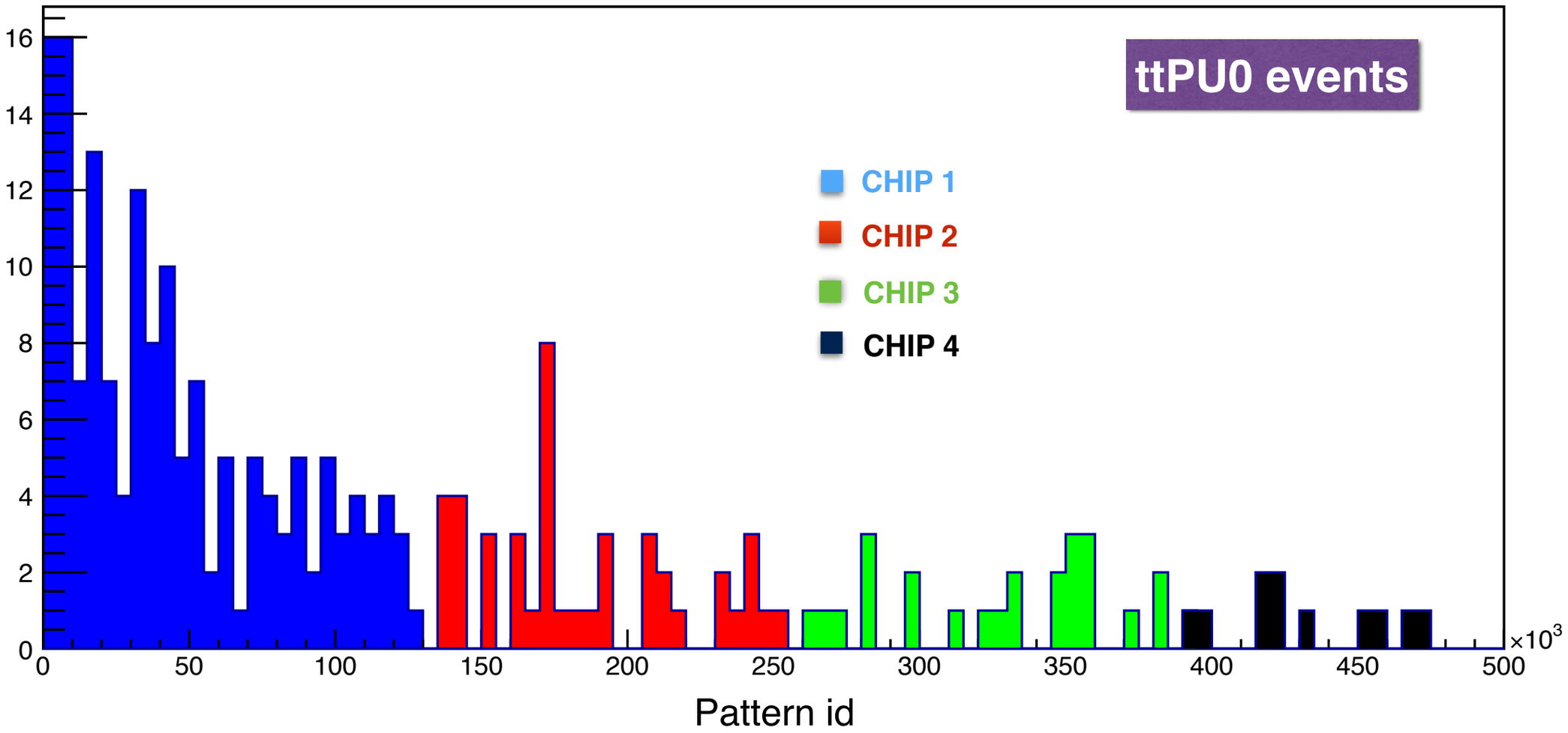


Event testing

- **Event testing is based on different methods, each method is compared with the others to have a full cross checked track reconstruction**
  - **Emulation: the full reconstruction chain is emulated using C++/Python code in CMSSW in float and in integer representation**
  - **Simulation: the vhdl/sv code that is used in the FPGA mezzanine is simulated using ModelSim. The RAM and AM chip latencies have been tuned according to the HW performance.**
  - **TestBench: using the evaluation board we send stubs and retrieve the fitted track parameters using GT links.**



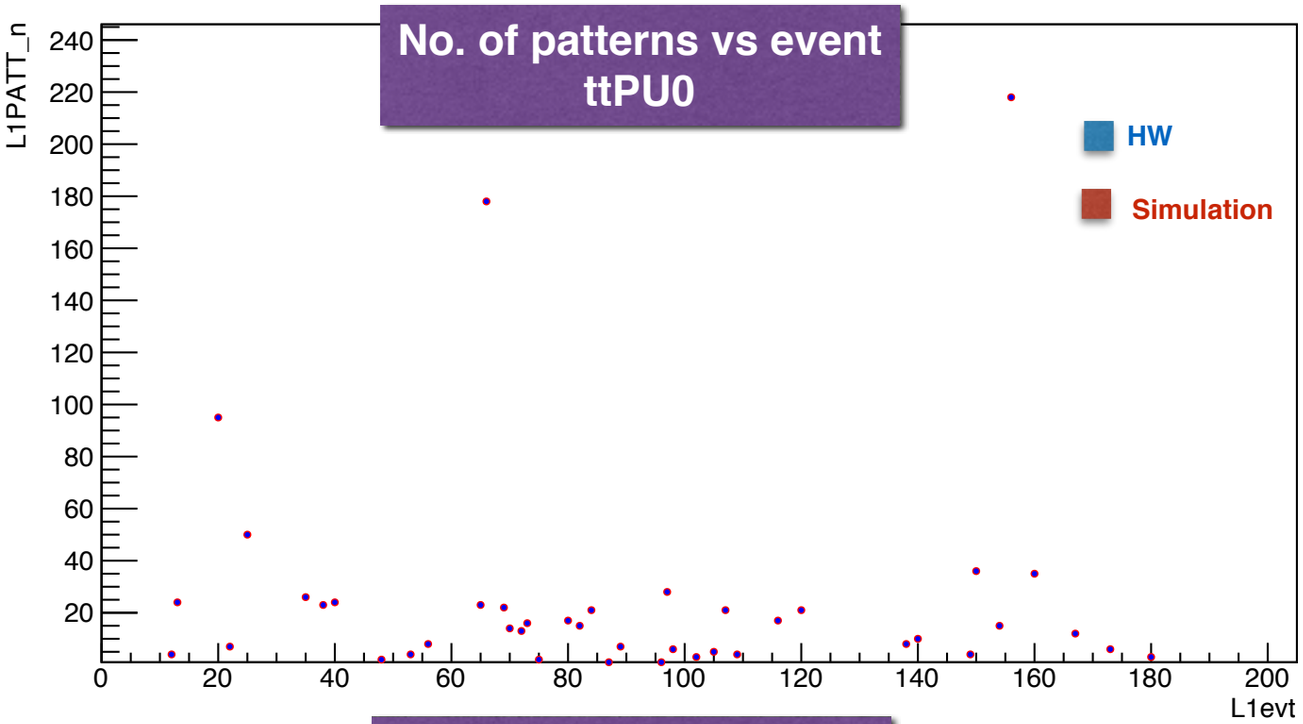
Pattern ID distribution from HW

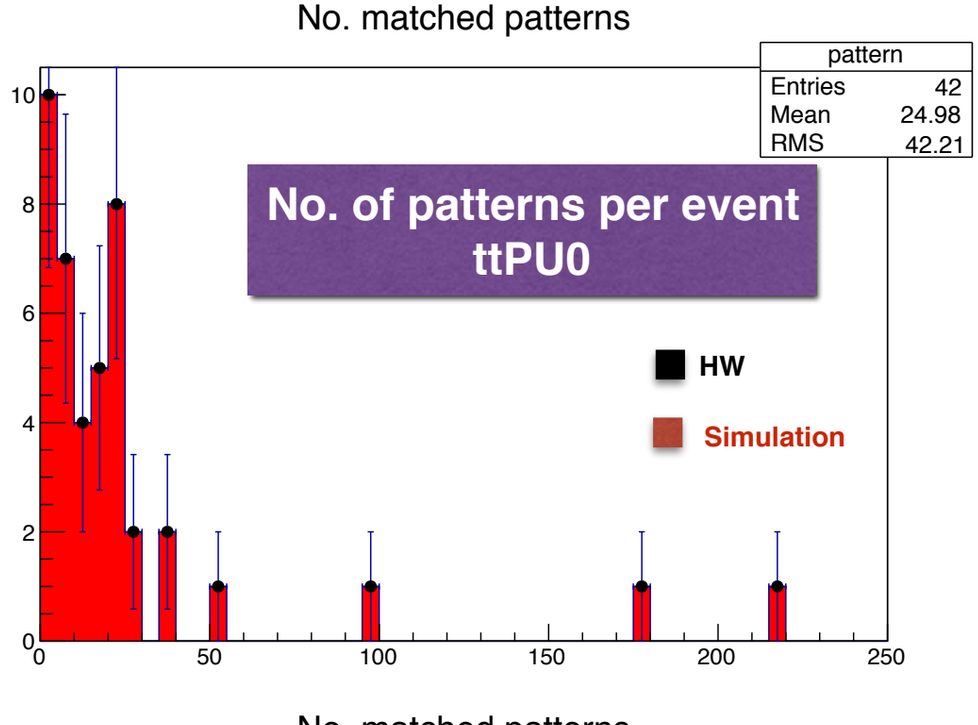


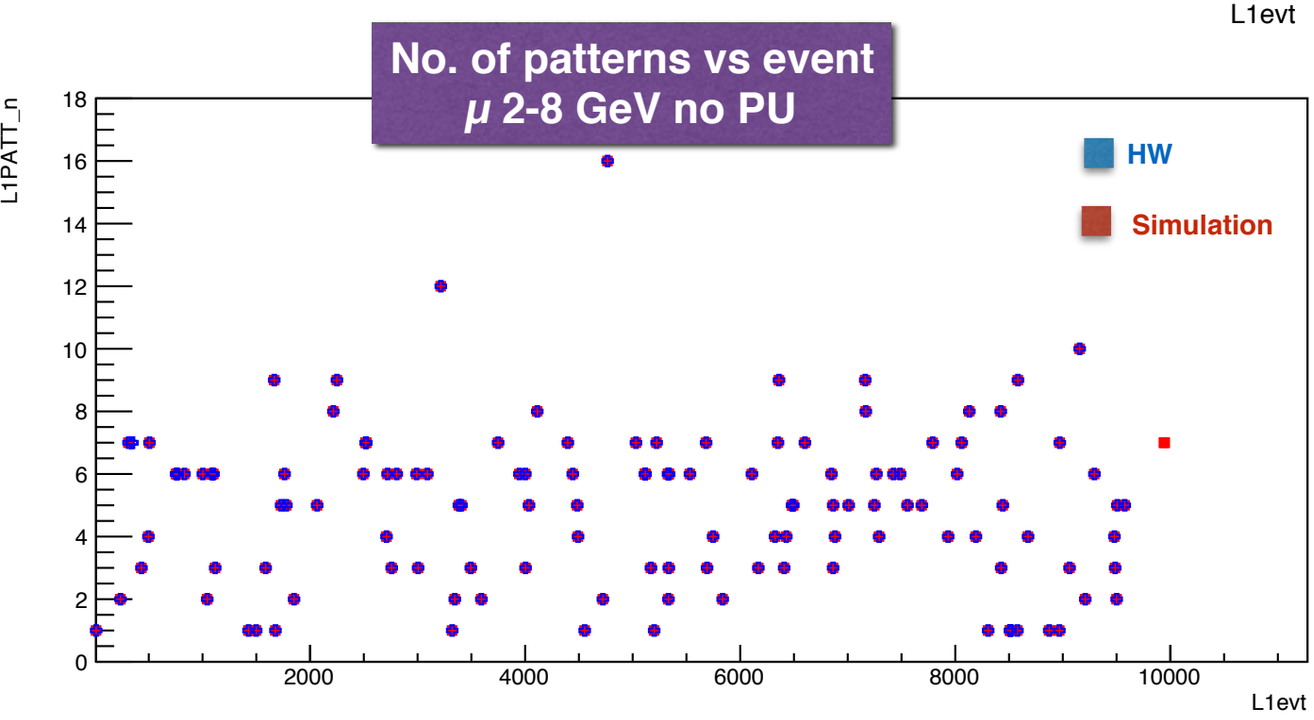
Patterns are loaded in the 4 AM chips in ‘popularity’ order (chip1 most popular patterns, chip4 least popular ones)

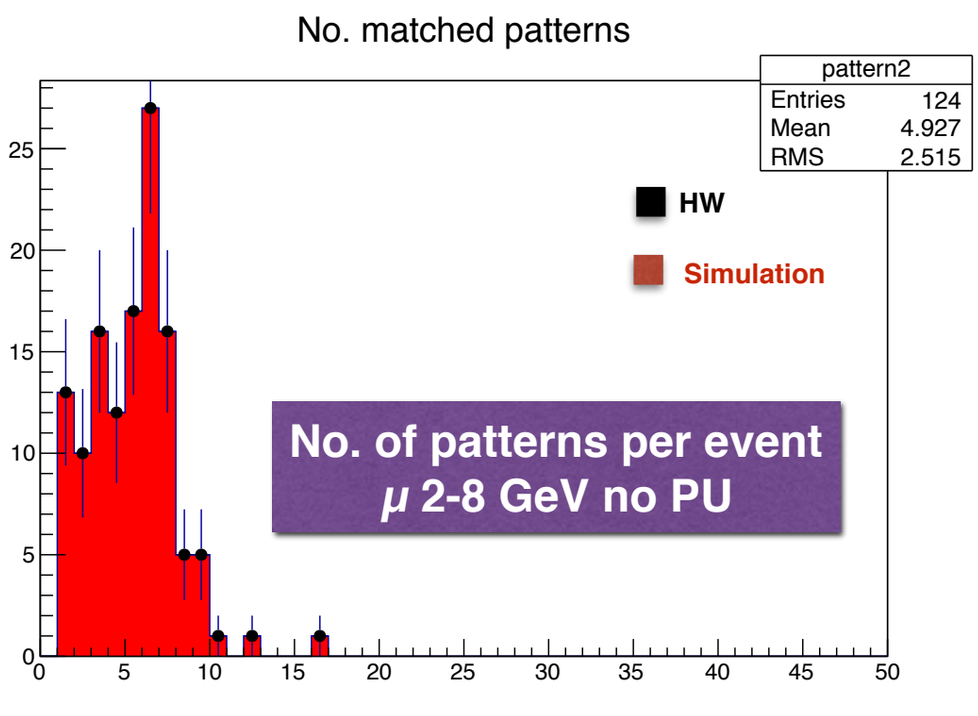
Matched pattern distribution in ttbar events reflects it: 1st chip has larger matches and 4th chip has fewer matches

Hardware output vs Simulation









All the patterns expected are firing - chips are behaving OK

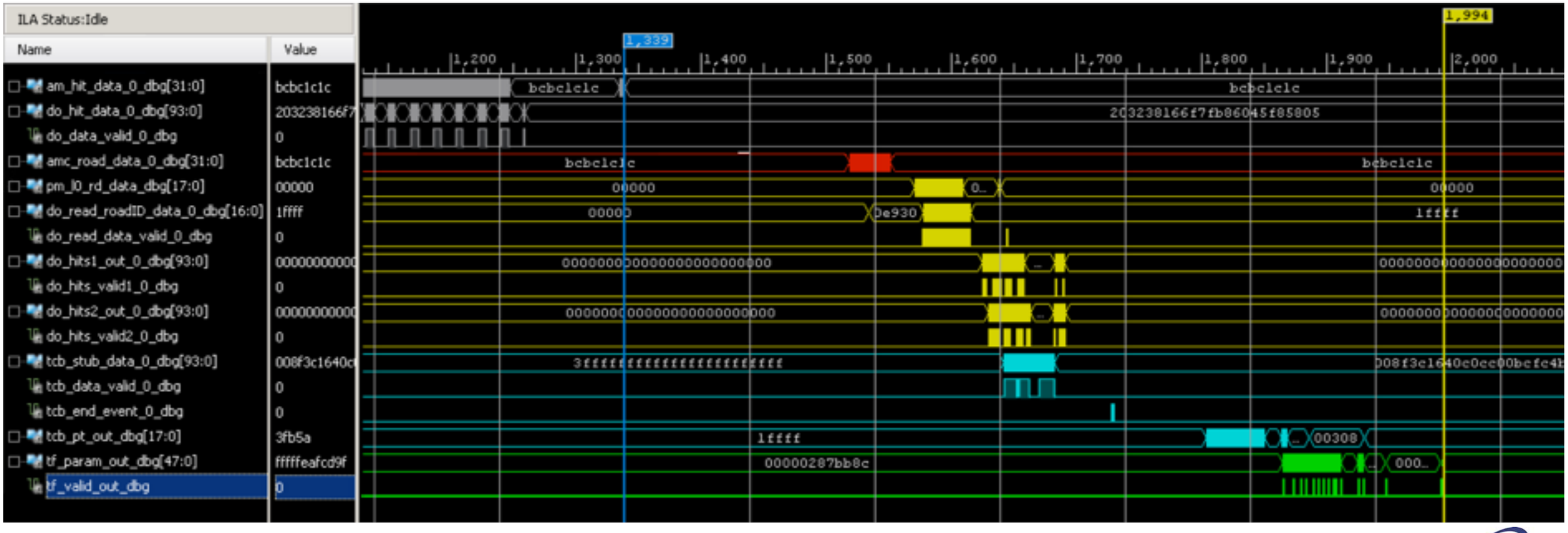
INFN Istituto Nazionale di Fisica Nucleare

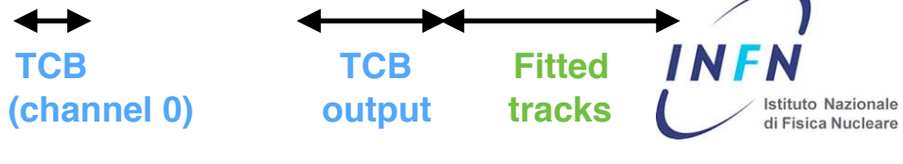
PRM06 timing in HW

- **Processing time measured for a complex event in the test stand sampling at 200 MHz**
- **Goal of the FW was to deal with big pattern banks and with multiple chips, processing time reduction is considered a secondary goal:**
  - **No target frequency: slowed down clocks for testing**
  - **No optimisation of resources: 12 TCB, relaxed delay parameters**
  - **Presence of special pattern to control the AM chip response**

Processing time from the end of stubs transmission to the AM06 chips up to the last fitted track in output: ~3.3 μs

1 busy $t\bar{t}$+PU140 event: max(stubs/layer)=138, 19 matched roads from multiple chips, 15 fitted tracks



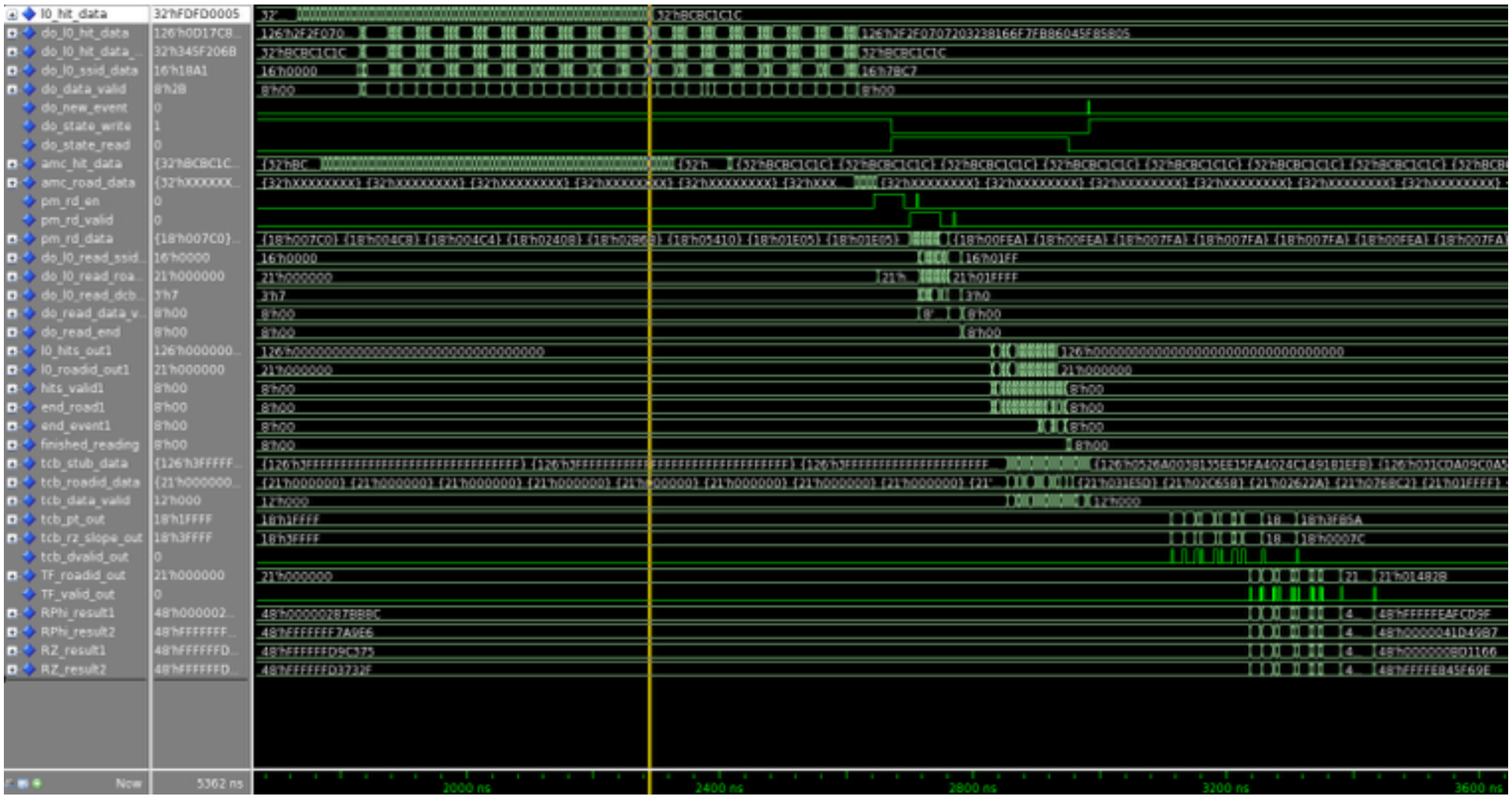


Extrapolation in simulations

- Processing time measured for a complex events in simulation (ModelSim) with the same banks and chip configuration
- Extrapolation of the processing time in case of faster clocks. Case study:
  - AM chip (AM08):250 MHz
  - Data Organizer: 400 MHz
  - TCB: 300 MHz
  - Track Fitter: 500 MHz
- The same event of previous slide has been sent to simulated FW/HW

Processing time from the end of stubs transmission to the AM06 chips up to the last fitted track in output: ~1.1 µs

1 busy $t\bar{t}$+PU140 event: max(stubs/layer)=138, 19 matched roads from multiple chips, 15 fitted tracks



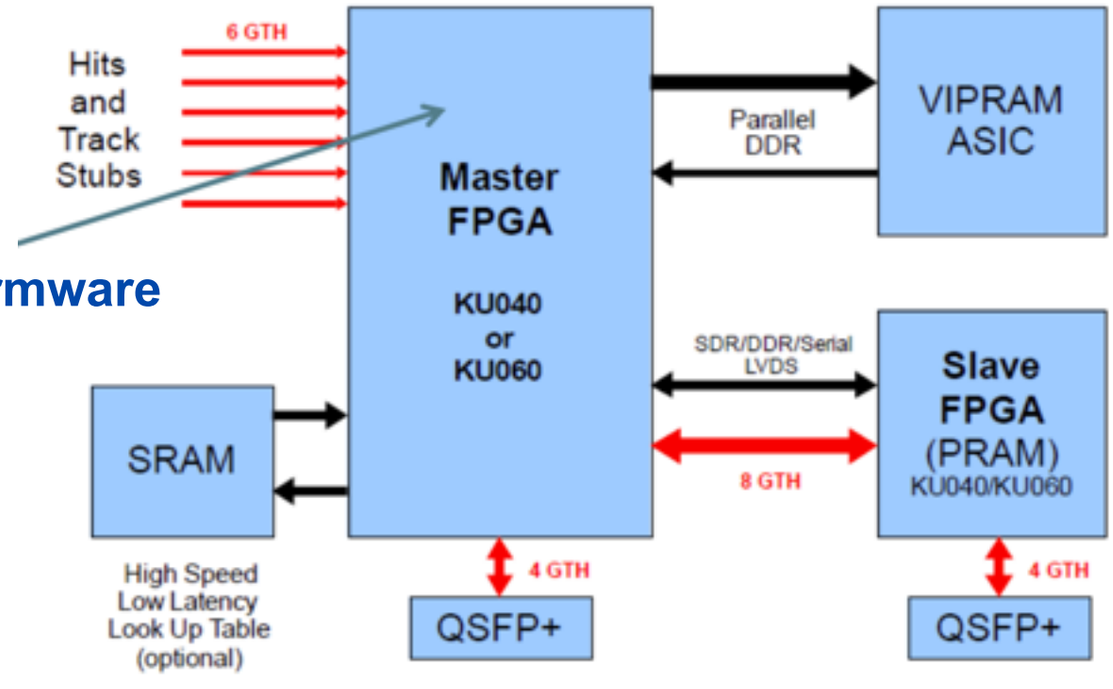
FNAL PRM

Developed to test vertically integrated associative memory (VIPRAM) technology

Currently used to push the latency low using AM emulation (FPGA) and fast links, it demonstrates one AM for 1/4 trigger tower

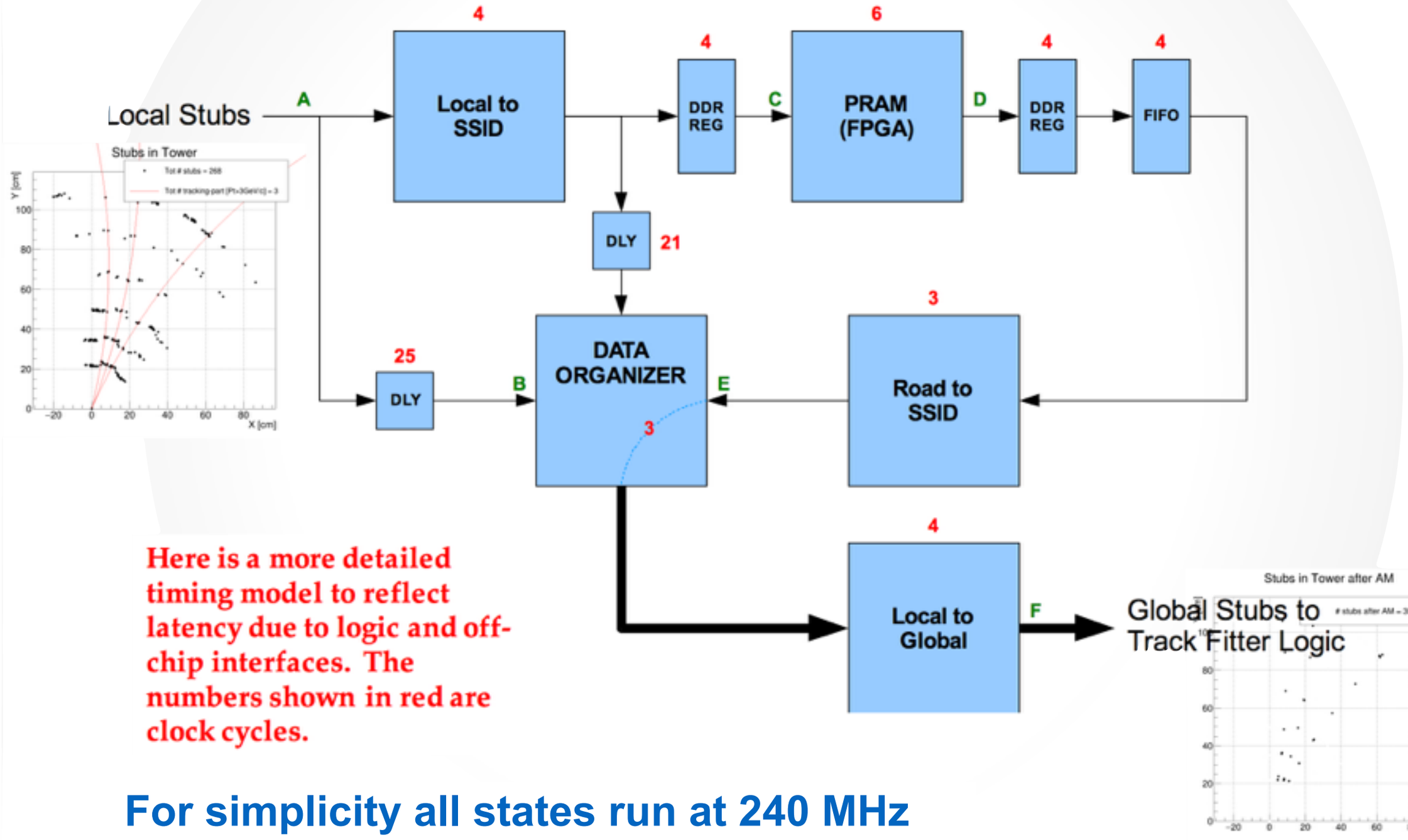
- **Dual Kintex Ultrascale FPGAs**
  - **Master FPGA: PRM firmware**
  - **Slave FPGA: AM emulation dealing with 1/4 of the trigger tower**
- **36 Mb low latency DDR II+ static RAM**

PRM Firmware



INFN Istituto Nazionale di Fisica Nucleare

FNAL FW PR latencies



Here is a more detailed timing model to reflect latency due to logic and off-chip interfaces. The numbers shown in red are clock cycles.

For simplicity all states run at 240 MHz

INFN Istituto Nazionale di Fisica Nucleare

FNAL FW track fitting latencies

- The stubs related to each matched road are combined (ACB)
- Each combination of stubs is fitted with a linearized fitter
  - 40 clock cycle of latency
  - Pipelined module: 1 clock cycle per fit
- Both modules running at 240 MHz
- With 4x3 instances of ACB+TF latency for PU250 events: 525 ns.
- Truncation at 480 combinations

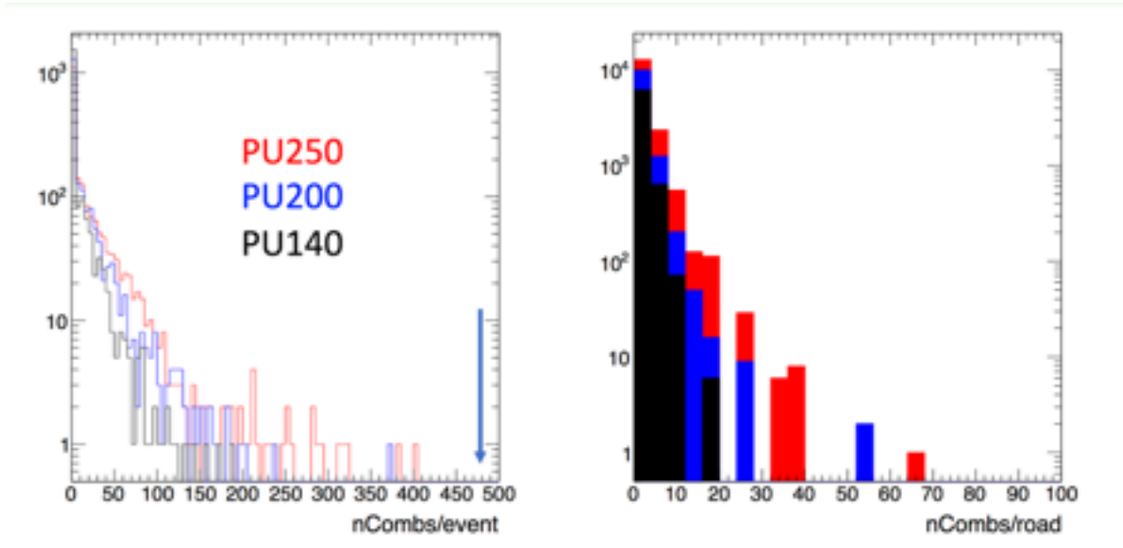
First order

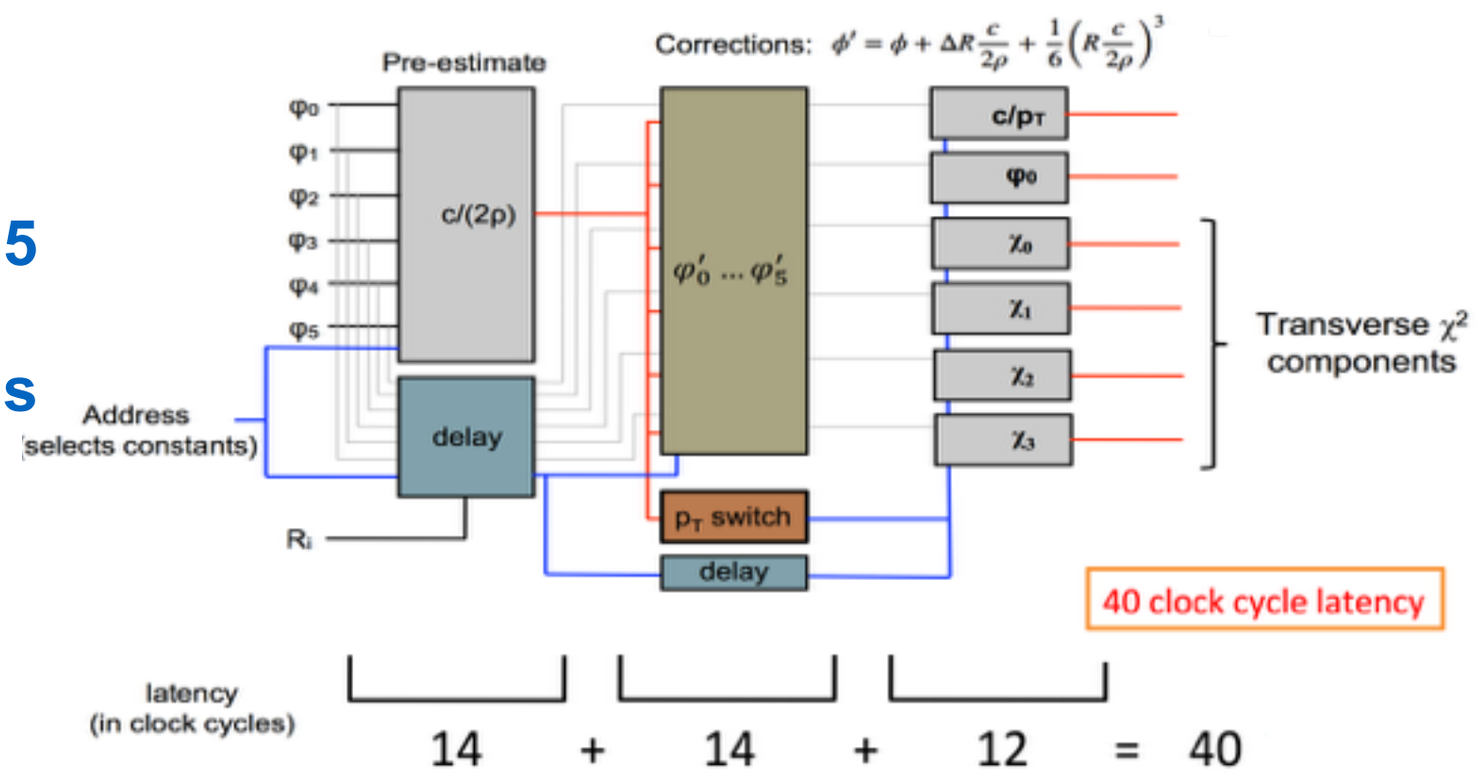
$$\phi' = \phi + \Delta R\frac{c}{2\rho} + \frac{1}{6}\left(R\frac{c}{2\rho}\right)^3$$

$$z' = z - \Delta R\cot\theta - \frac{1}{6}\cot\theta\left(\frac{c}{2\rho}\right)^2 R^3$$

Second order

Number of combination for 1/4 of trigger tower





INFN Istituto Nazionale di Fisica Nucleare

FNAL FW latency and performance

System latency demonstrated with the FNAL PRM and AM chip emulation

- Data delivery to the PRM starts/ends: @1.2 - 1.7 µs
- + Pattern recognition starts/ends: @1.85 - 2.3 µs
- + Track fitting starts/ends: @2.03 - 2.53 µs

Tested with events up to PU250

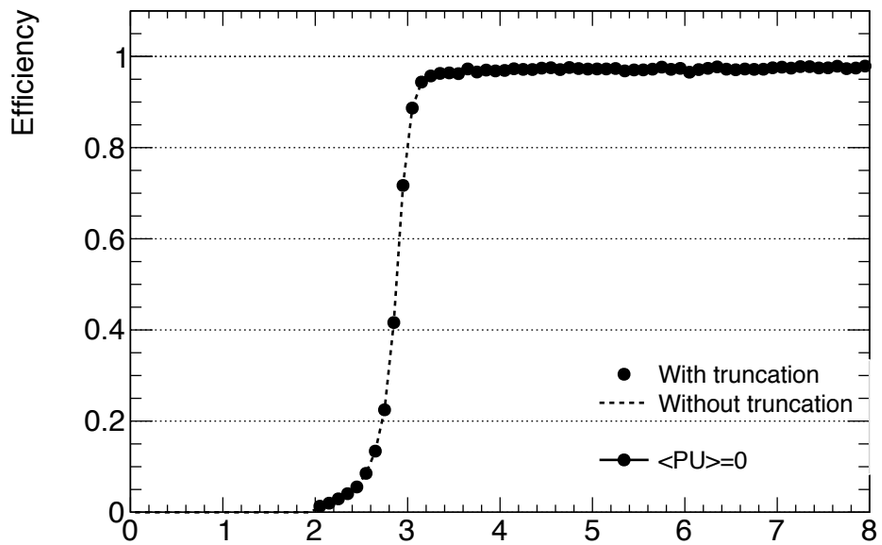
1.5 µs spare latency left to final processing (duplicate removal)

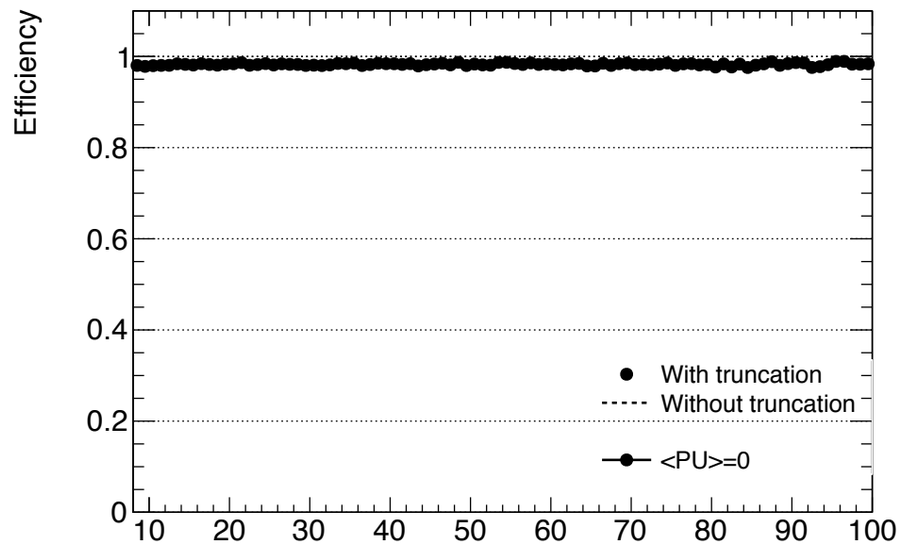
System performance

- Excellent based on today's demonstrator
- Room for additional improvements
  - Bank optimization
  - Dedicated banks (electrons, jets)
  - Combinatorics reduction

System performance - Single muons

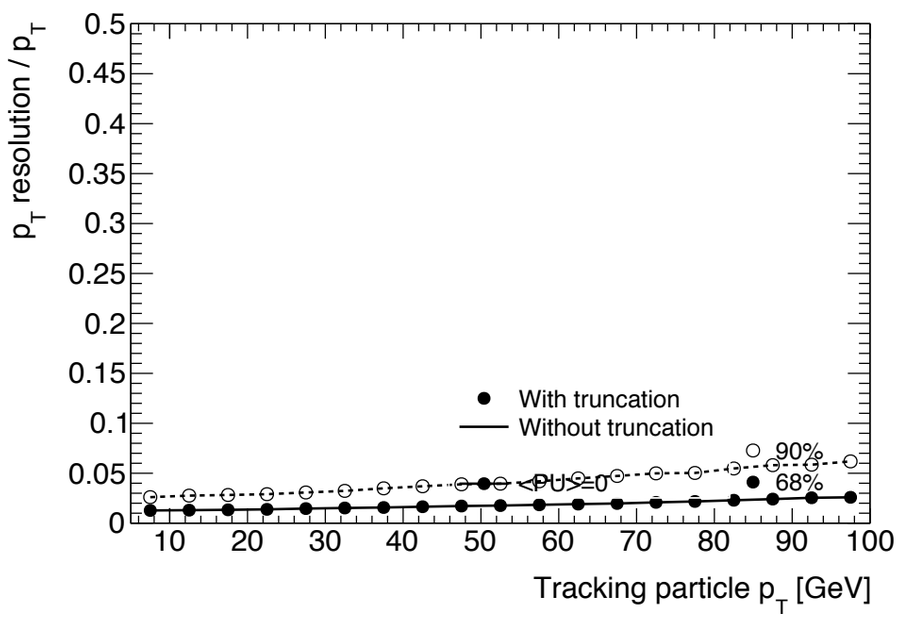
Efficiency vs pT: 2-8 GeV, 8-100 GeV





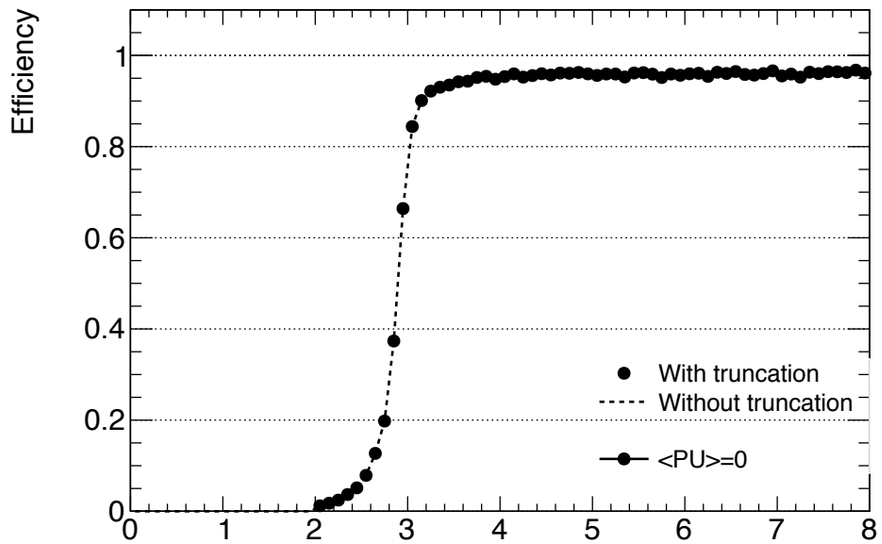
ptres/pt vs pt: 2-100 GeV

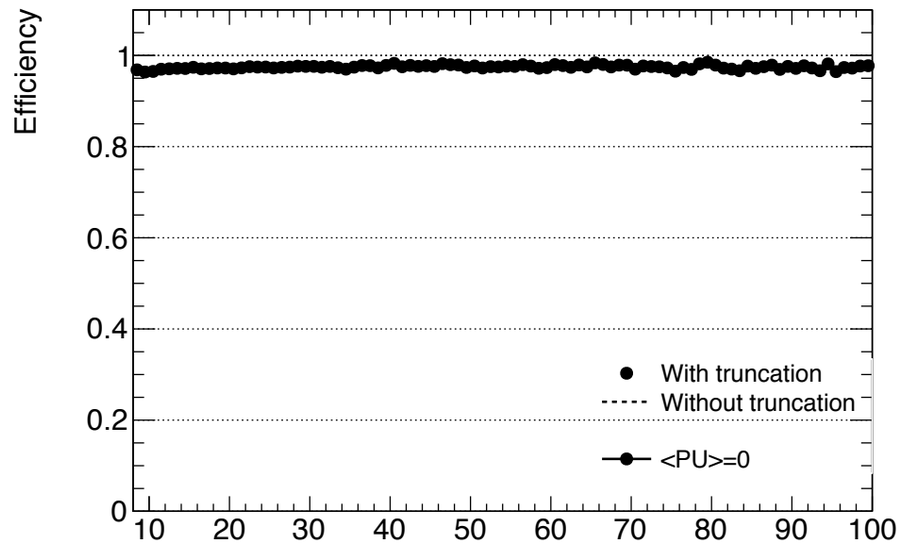
Demonstrator accepts stubs with pt>3 GeV



System perf. - Single pions

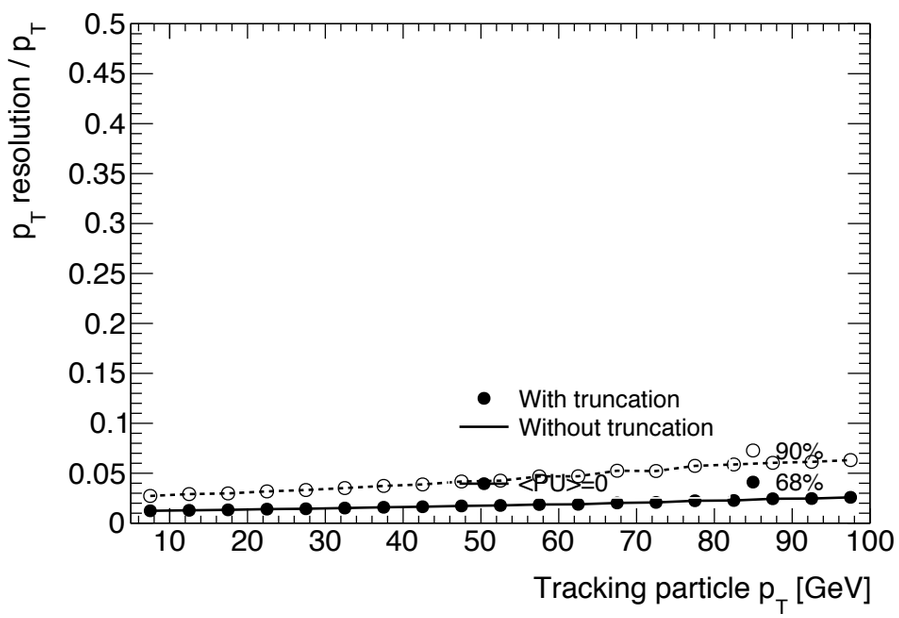
Efficiency vs pT: 2-8 GeV, 8-100 GeV





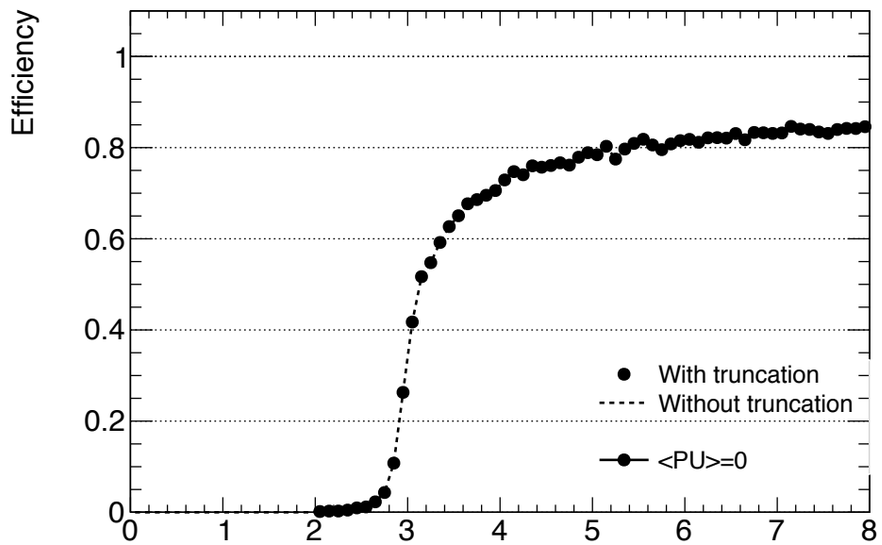
ptres/pt vs pt: 2-100 GeV

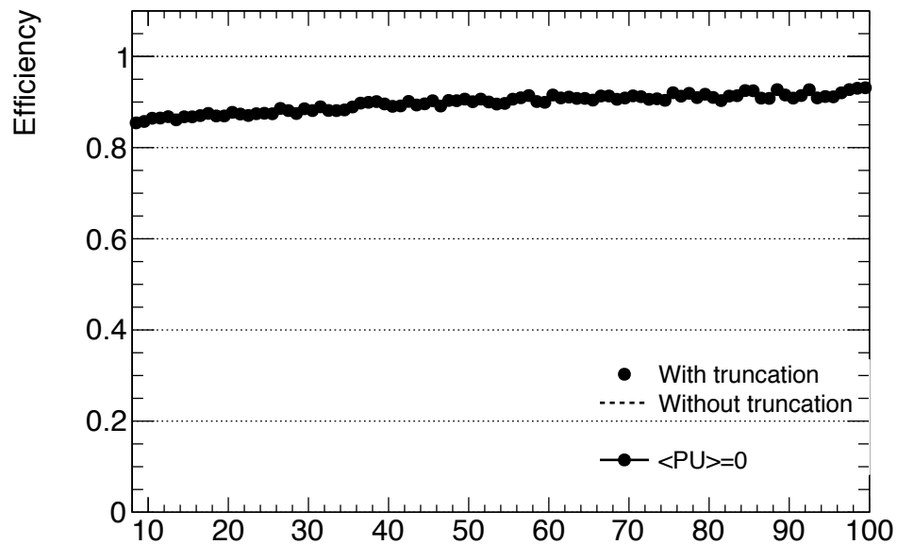
Demonstrator accepts stubs with pt>3 GeV



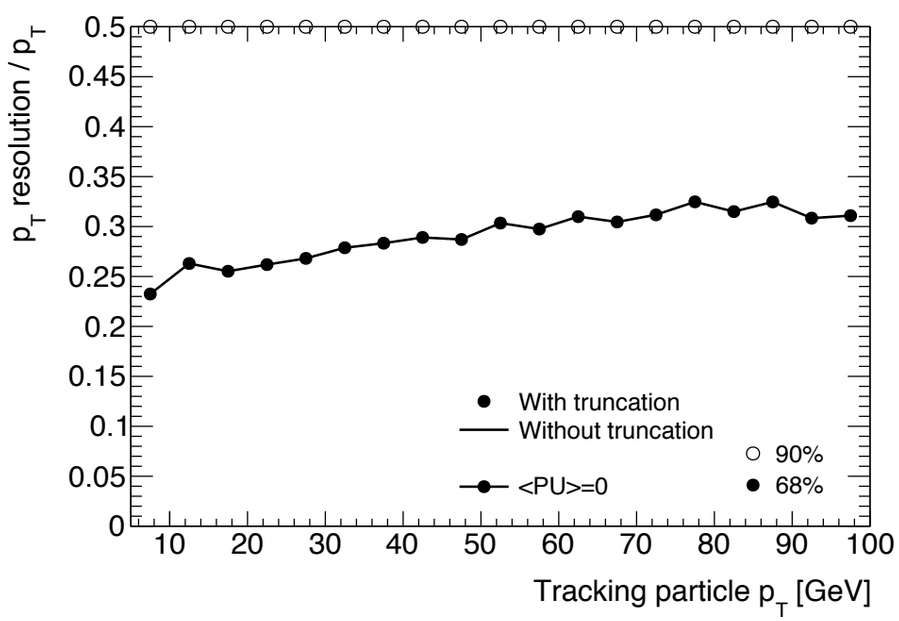
System perf. - Single electrons

Efficiency vs pT: 2-8 GeV, 8-100 GeV





ptres/pt vs pt: 2-100 GeV



Demonstrator accepts stubs with pt>3 GeV

INFN Istituto Nazionale di Fisica Nucleare

Conclusions

A low latency track reconstruction demonstrator for L1 trigger of CMS for HL-LHC has been built

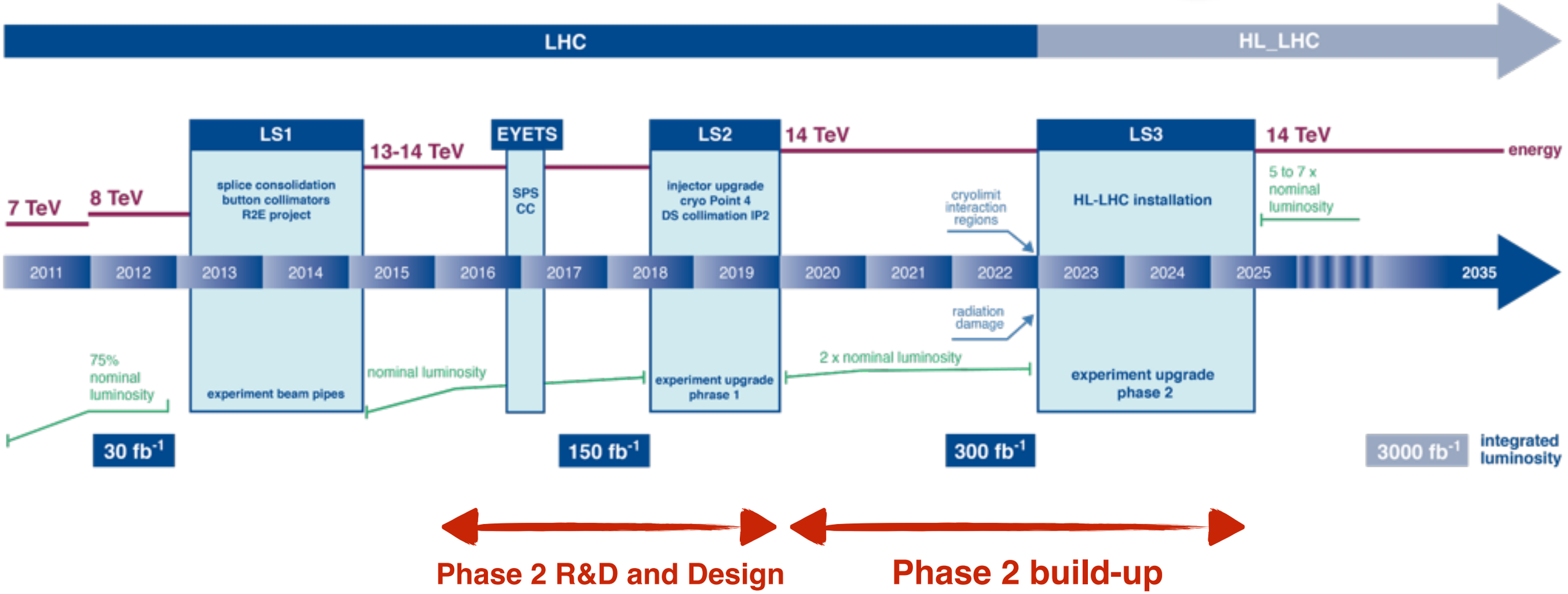
- **A system of 2 ATCA shelves demonstrated the data delivery from the tracker front-end electronics to the PRMs**
- **A fully equipped PRM designed with current available hardware validated the concept of pattern matching with multiple AM chips in parallel and the full process up to fitting**
- **A PRM with an AM chip emulation showed that the pattern matching and track fitting can be done with less than 4 μs with the future chips and technologies**
- **Excellent performance for muons and tracks with very good pt resolution up to event with PU200 and double ttbar events**

Extra

Timeline

LHC / HL-LHC Plan

High Luminosity LHC



Phase 2 R&D and Design **Phase 2 build-up**

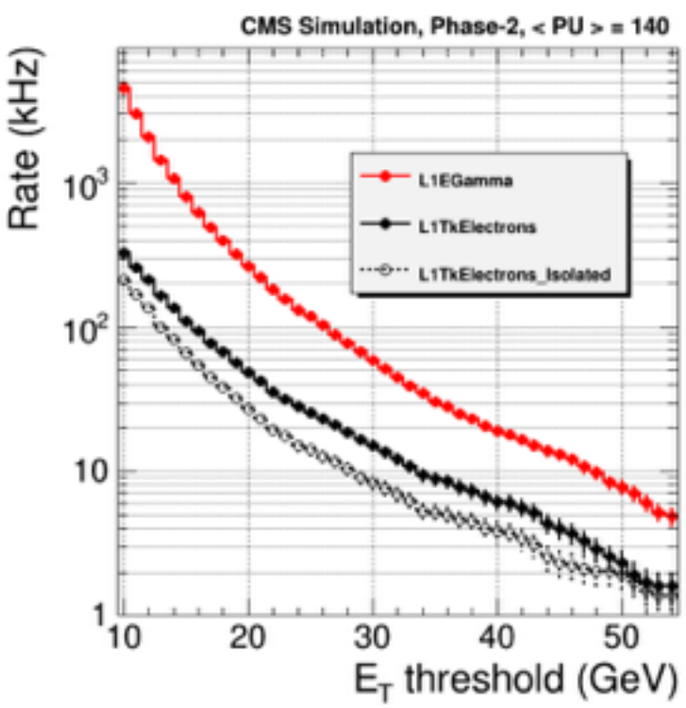
INFN Istituto Nazionale di Fisica Nucleare

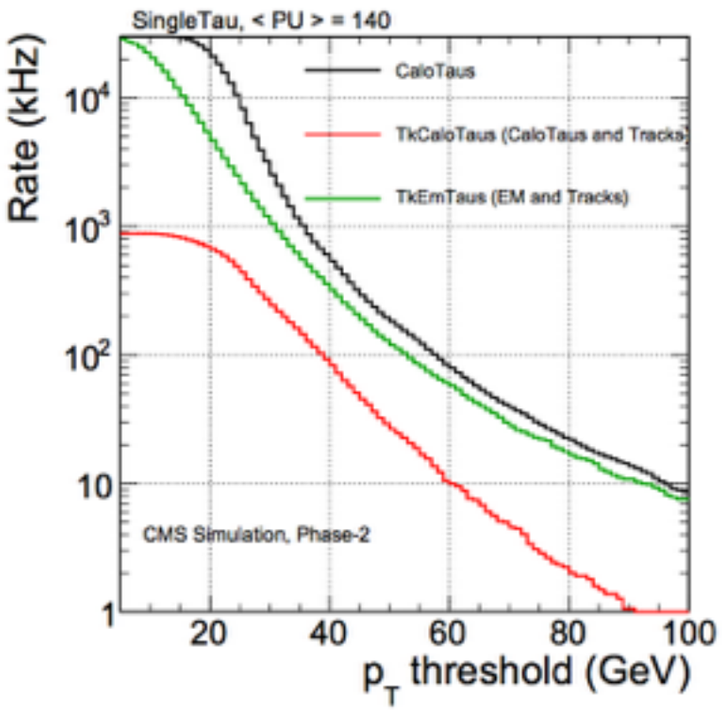
Motivation

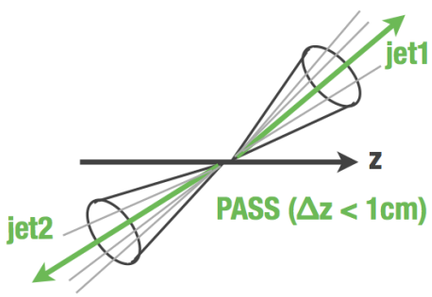
HL-LHC physics goals require excellent Trigger selectivity on basic objects (leptons, jets, b-jets, MET)

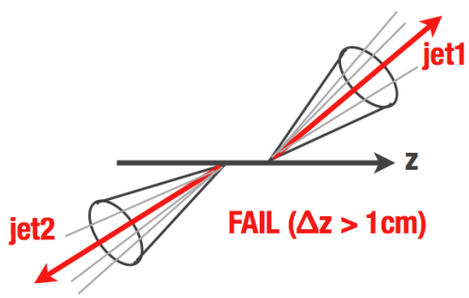
This might be jeopardized by the increased level of pileup events

- **Huge rate of μ from heavy flavors ⇨ use better p_T resolution from tracker**
- **Prompt electrons at L1 need to be separated from huge γ ⇨ Tracker tracks**
- **High E_T jets from (many) different primary vertices ⇨ jet-vertex association**
- **Photon isolation in Calorimeters compromised by large pileup ⇨ use tracks**

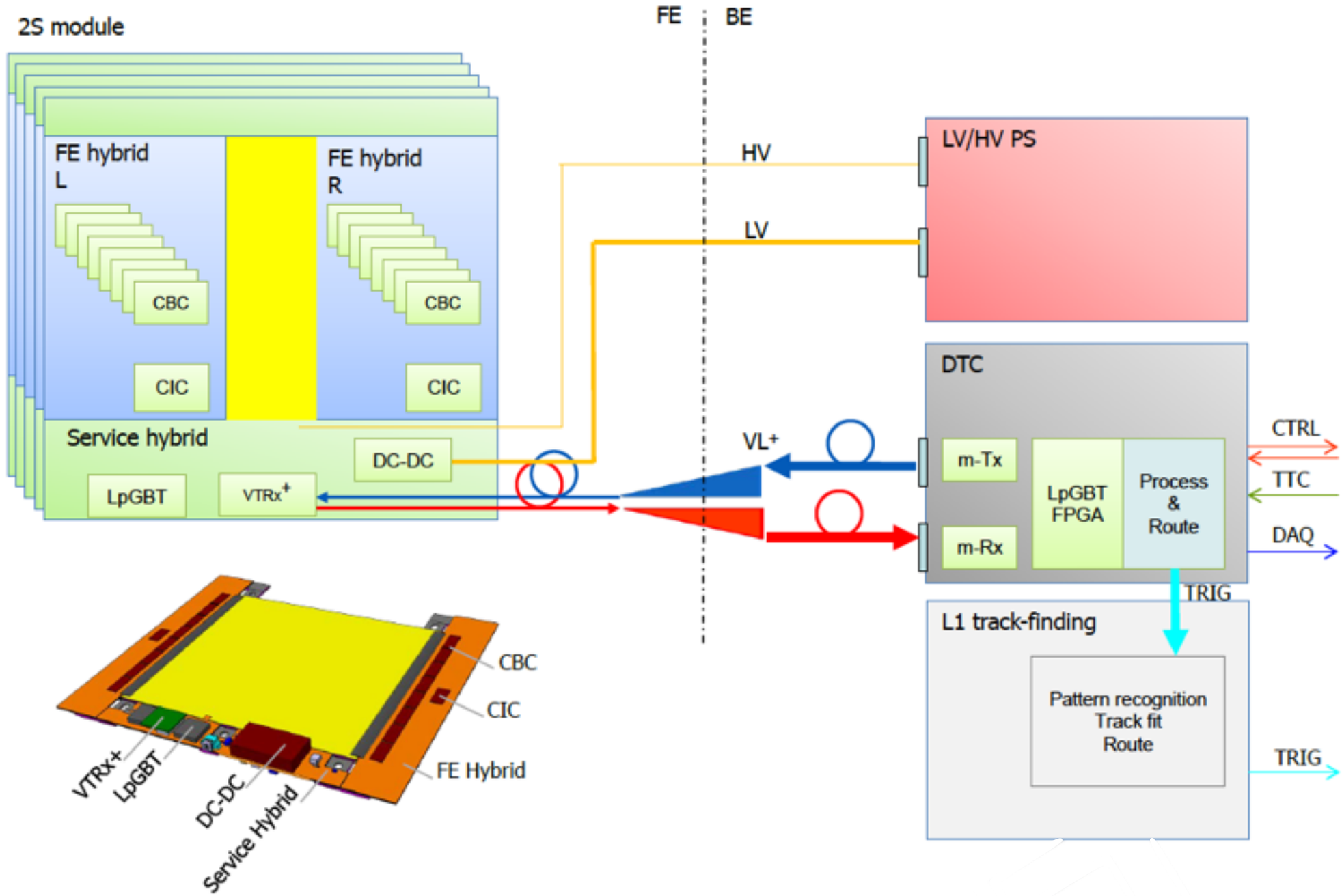








Module readout



Don’t care bits

In AMchip04 it was introduced the possibility to have ternary logic (don't care) bits to achieve variable resolution .

- For each layer: a ``bin'' is identified by a number with DC bits (X)
- Least significant bits of ``bin'' number can use 3 states (0, 1, X)
- The ``bin'' number is stored in the Associative Memory
- The DC bits can be used to OR neighborhood high-resolution bins, which differ by few bits, without increasing the number of patterns

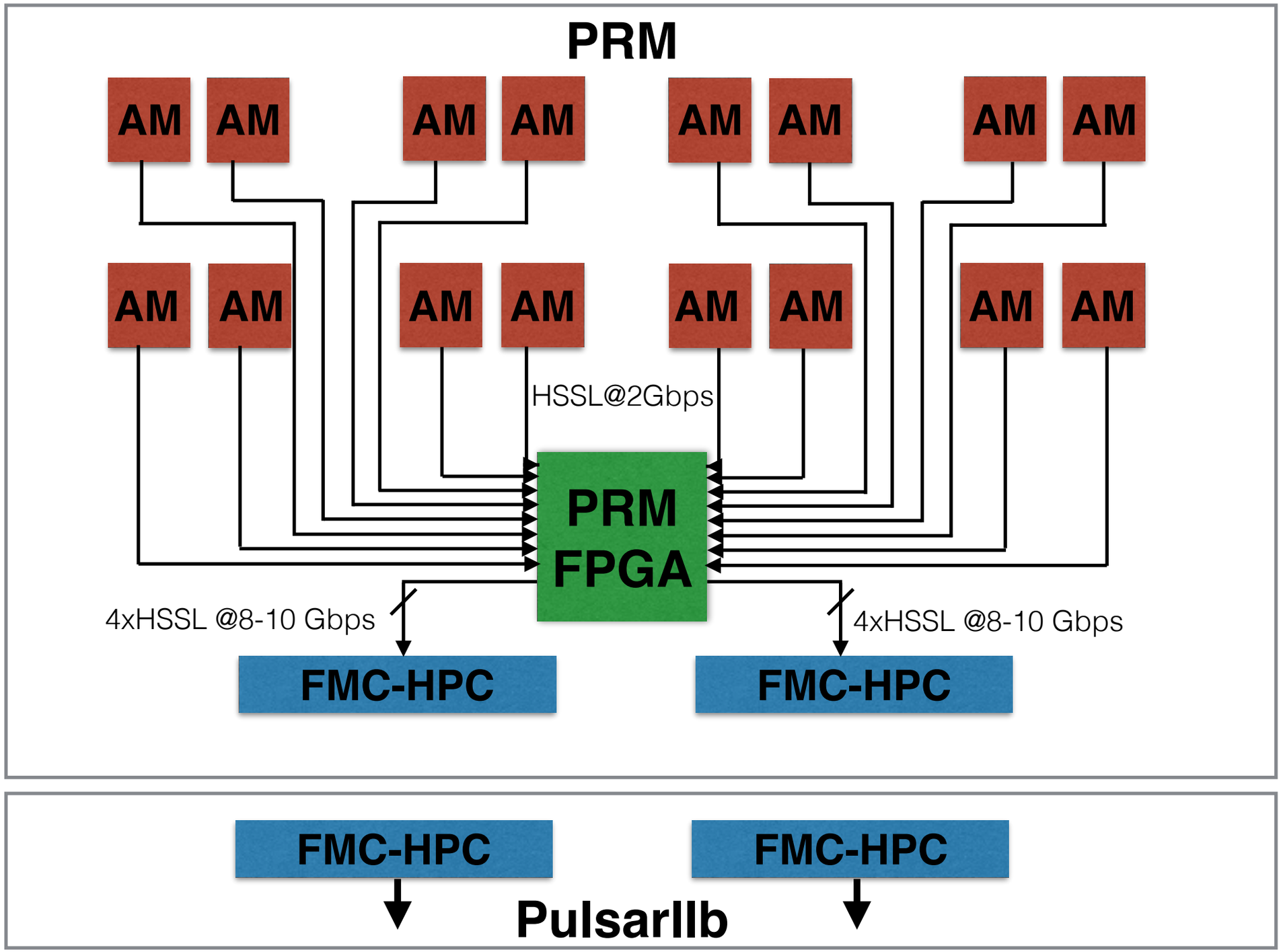
Pixels:

0	1	2	3	4	5	6	7
8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31

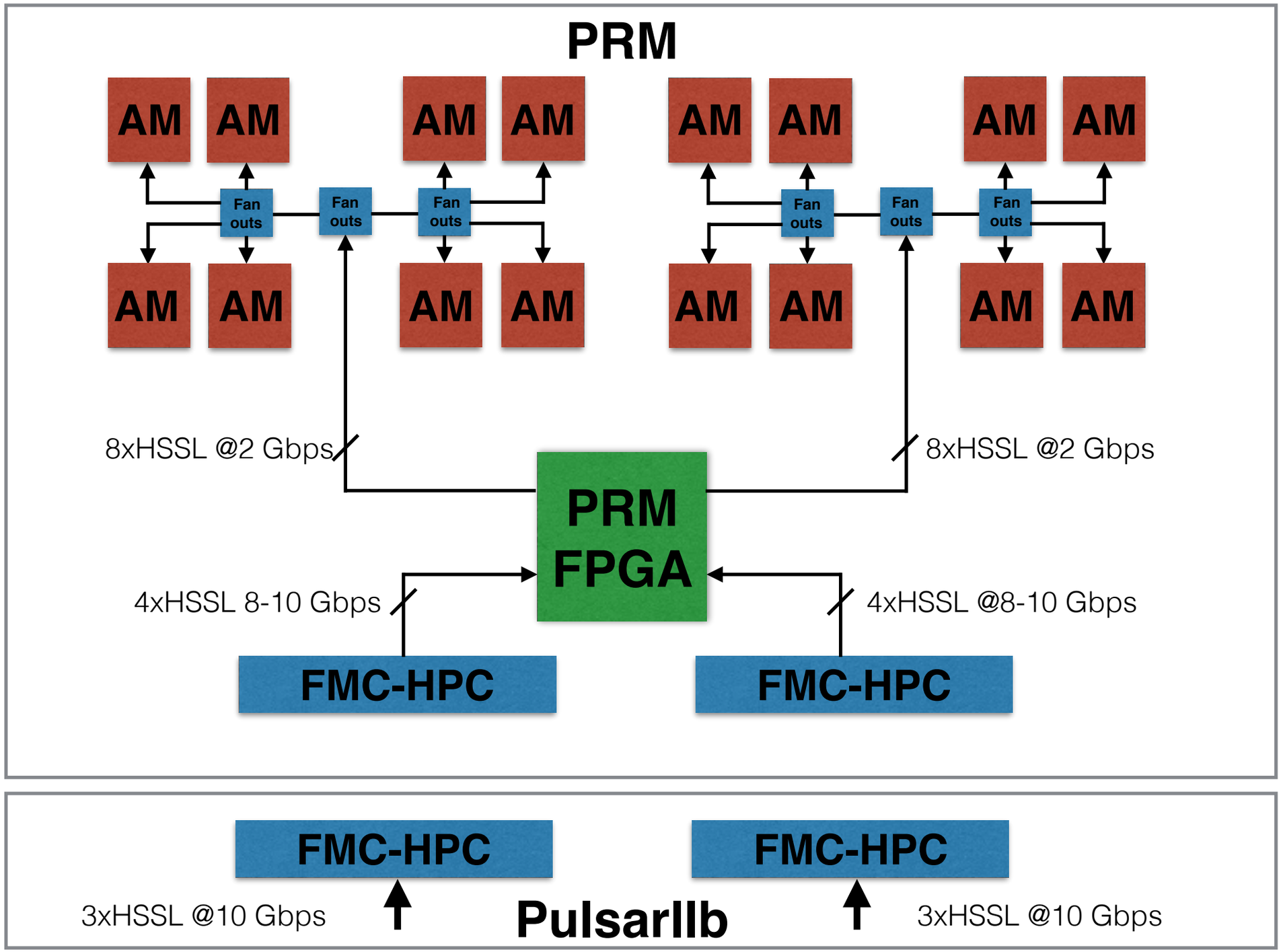
Using binary format
“01010” selects bin 10
“0001x” selects bins 2 or 3
“1x000” selects bins 16 or 24
“0x11x” selects bins 6,7,14, or 15
“111xx” selects bins 28 to 31

Francesco Crescioli’s slide

INFN PRM data flow (RoadID/Tracks)



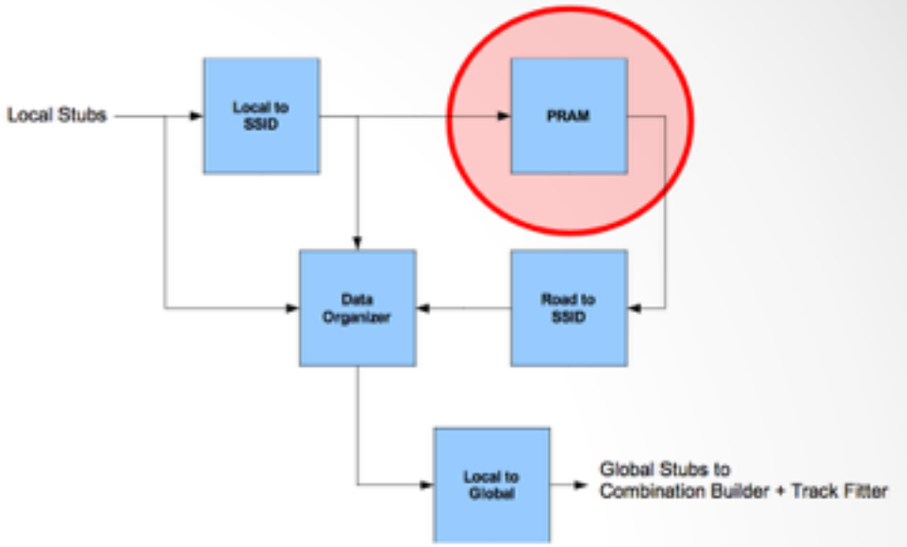
INFN PRM data flow (stubs/SSIDs)

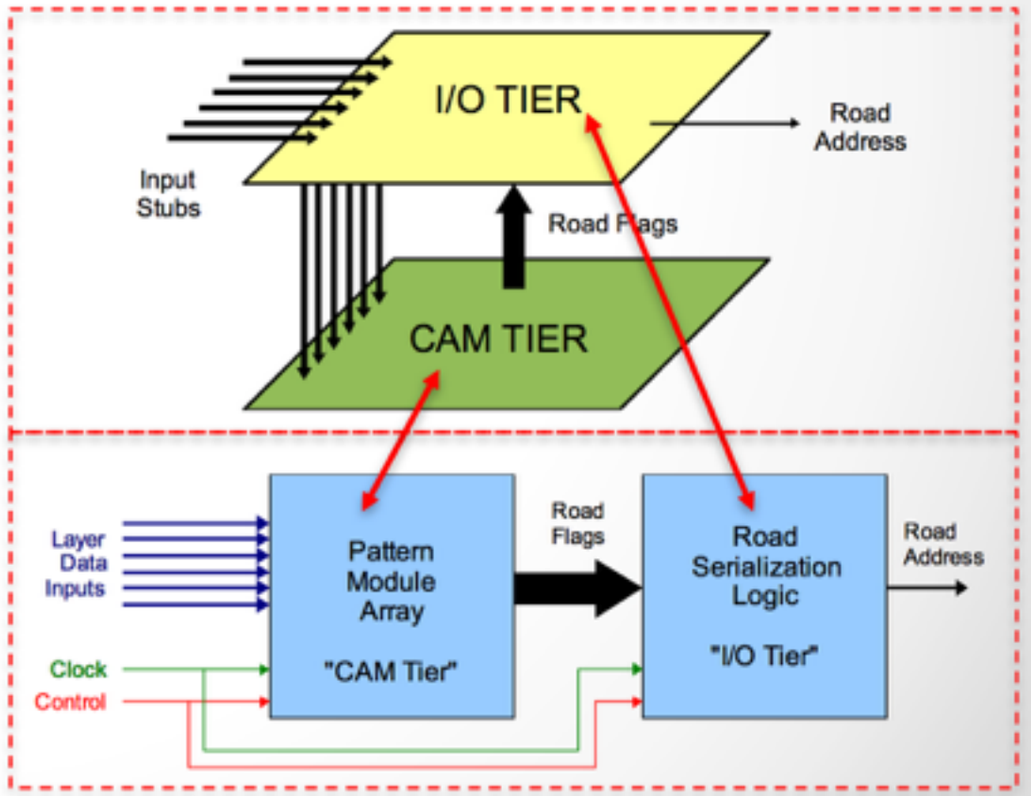


FNAL PRM - AM Emulation

AM in FPGA: Overview

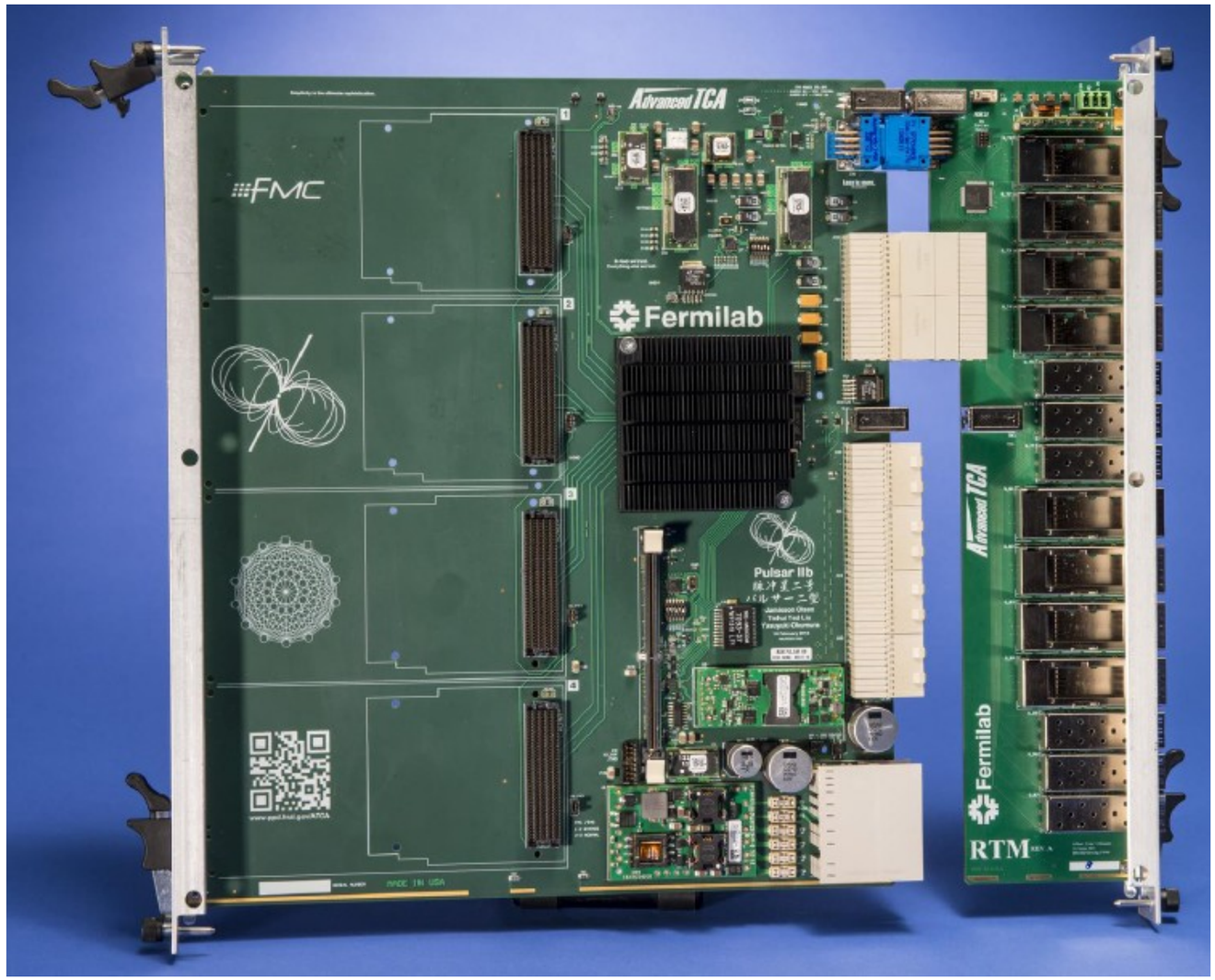
- AM in FPGA: very **closely follows** the AM ASIC (chip) design
  - **Match two silicon tiers** in ASIC with two modules in FPGA firmware
    - CAM Tier -> a 2D array of Patterns
    - I/O Tier -> input and output of fired roads
  - **Pipelined operation**
    - CAM tier: processes pattern matching with stubs for current event N
    - I/O tier: outputs road addresses for event N-1 at the same time





Pulsar IIb

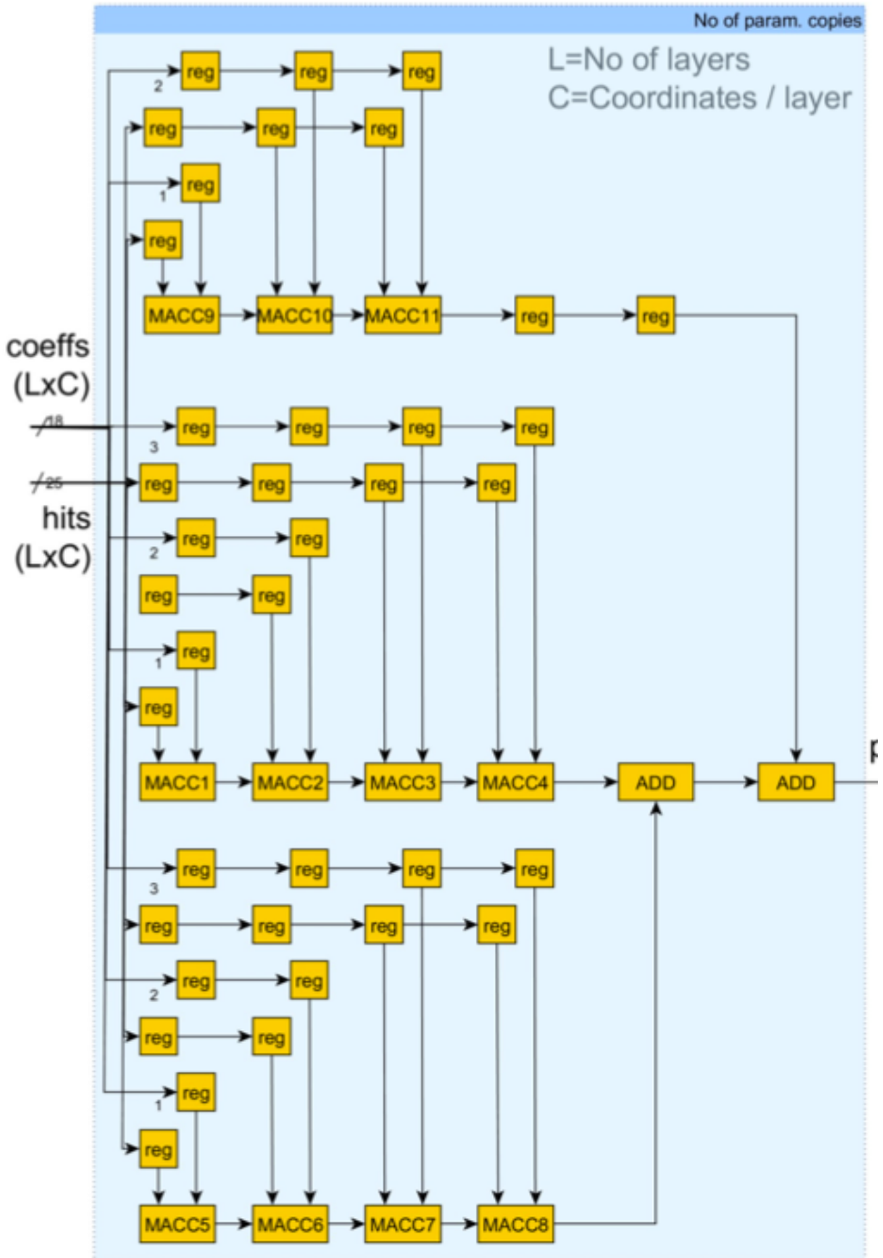
- FNAL developed ATCA board & Rear transition module
  - Pulsar-2b
    - Virtex 7 FPGA
    - 80 GTH @ up to 13.1 Gpbs
      - 40 to RTM
      - 28 to backplane
      - 12 to FMCs
  - RTM
    - V2.0 (not shown) completed in July 2014
    - 10 QSFP+ (400 Gpbs)



(slide from FNAL)

TF implementation

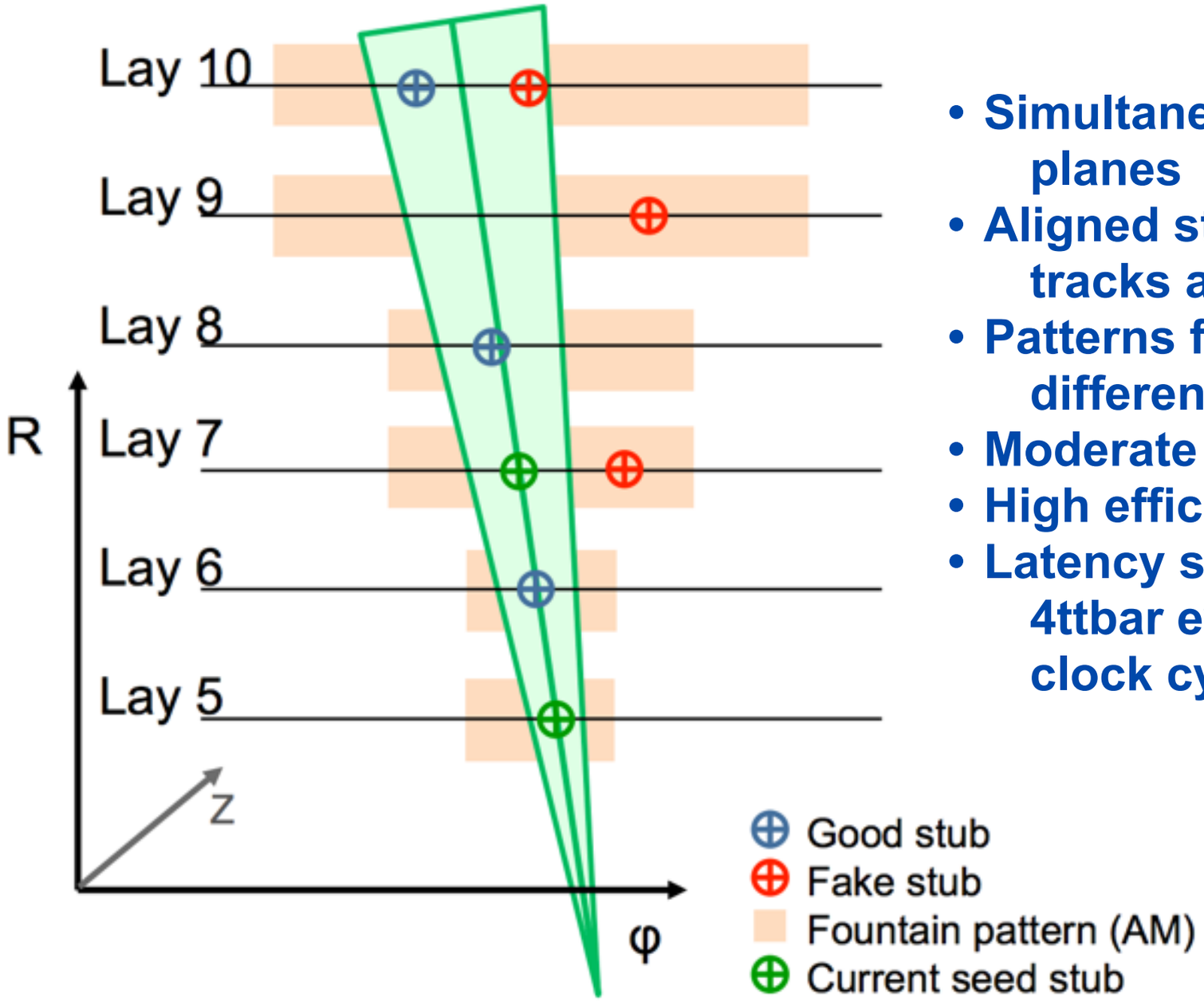
$$p_i = \sum_{j=1}^{N} a_{ij} x_j + b_i$$



- **Extensive use of DSP unit**
- **Frequency of 500MHz+**
  - **Initial delivery of the results 100ns latency**
  - **Any new result pops up every 2 ns**
- **Multiple units can run in parallel**
  - **4 units can fit in a mid-grade device giving 2GFits/sec**
- **Fit coefficients stored in BRAM memory resources**
  - **If we can do with a small number of coefficient sets, possible to fit more TF units in high grade devices**

Track candidate builder

From the identify pattern we select an unique set of the stubs compatible with stub pairs (seeds) from PS layers/strings



- Simultaneously operated on (R,φ) and (r,z) planes
- Aligned stubs are added into temporary tracks and the best one outcome the TCB
- Patterns from the AM chip can be send to different TCB instances
- Moderate FPGA resource occupancy
- High efficiency (>98%)
- Latency studied in case of 1000 140PU + 4ttbar events, 10 TCB instances ~ 100 clock cycle on average

INFN Track fitter PCA

Principal Components Analysis (PCA)

- **Linearization of the fit problem**

$$p_i = \sum_{j=1}^{N} a_{ij} x_j + b_i$$

p_i = helix parameter

x_i = stub coordinate

a_{ij} and b_j = constants, from simulations

Factorizing the 3D problem into 2 views

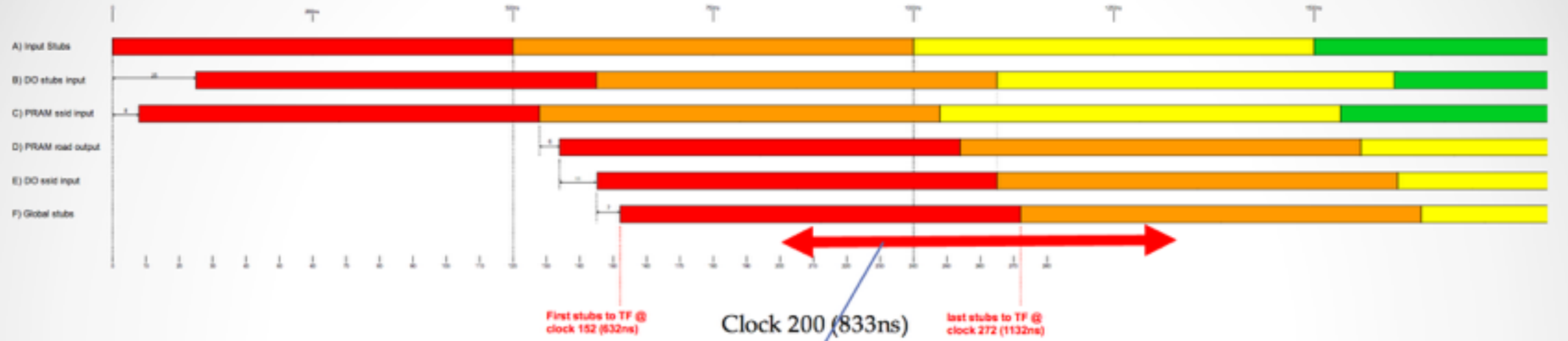
- **R-z: 20 bins in eta (0.05 wide)**
- **R-φ: p_T bins x2 (charge)**
- **Rationale is to keep small the number of constants stored in BRAM memory and to allow more fitters implemented in the FPGA**

Implemented both in floating point and in integer representation (C++)

- **Coordinated represented with 18 bits**
- **Constants represented with 24 bits**
- **Perfect agreement between the floating point and integer representation fitter in C++**

FNAL FW latency

Summary of Pattern Recognition and CB/TF latency



At 2533 ns, track fitting ends
At 1200 ns+ 833ns = 2033ns, track fitting starts

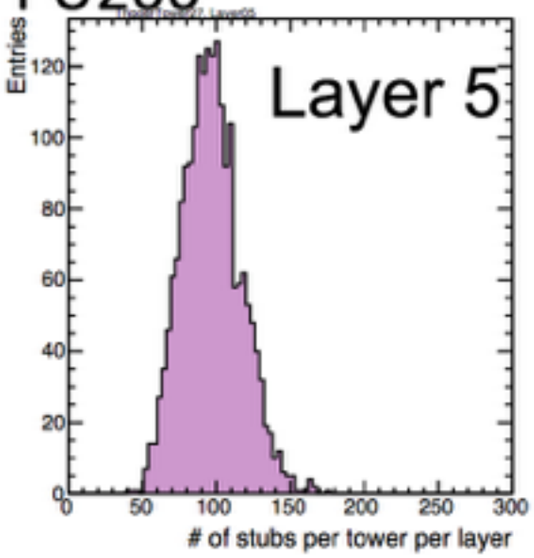
At 1200 ns + 650ns ~ 1850 ns, the first road with full resolution stubs are ready for downstream CB/TF …
At 1233 ns, the first stubs/SS arrives at AM input
At 1200 ns, first stubs arrive at PRM input from Pulsar

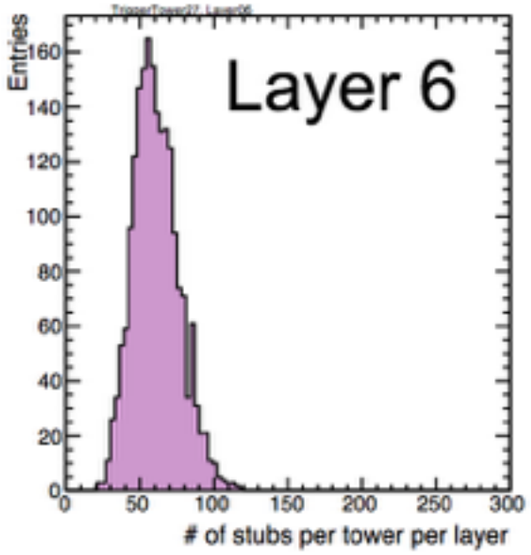
Works for PU140, PU200, PU250 events, and most of the ttbar + PU200 events (except the very high tail end)

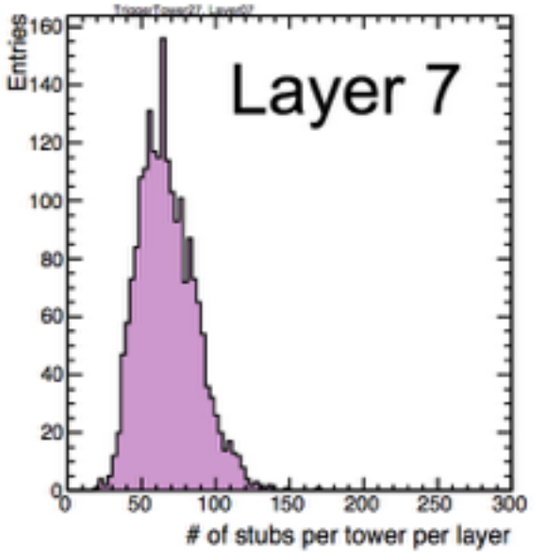
Tracker data

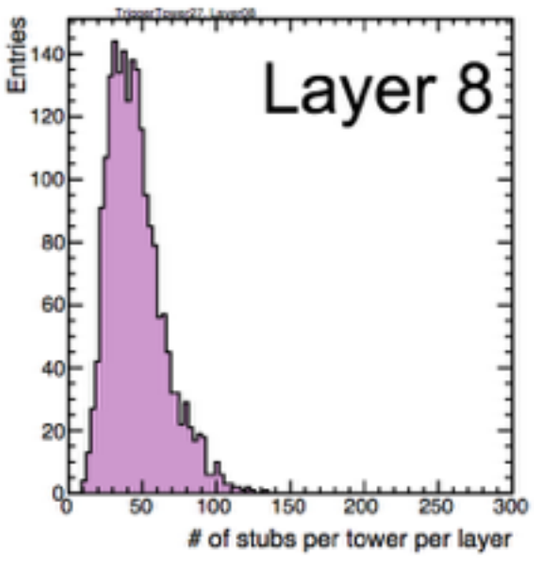
Data arriving to PRM (full barrel tower)

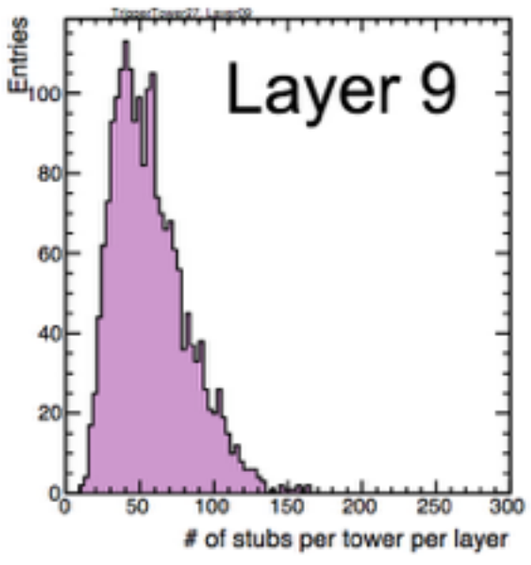
PU200

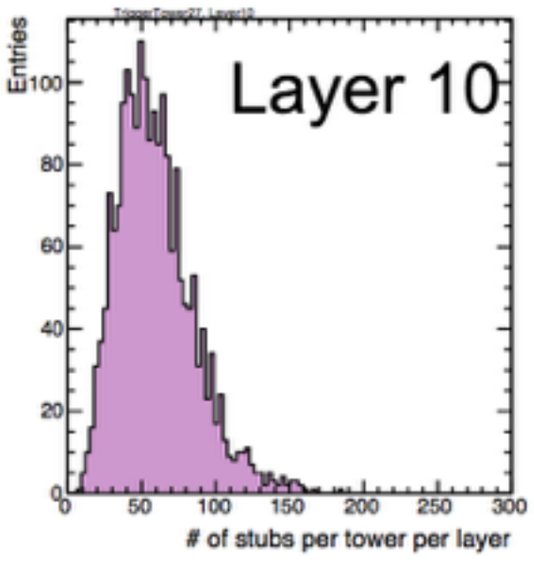




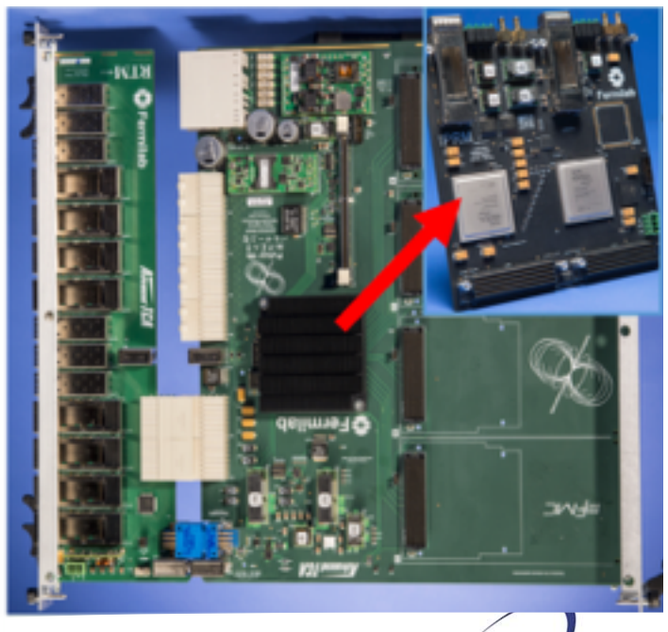








Data flow goes from PRB to PRM via 6 10 Gbps input lane, each lane for a layer

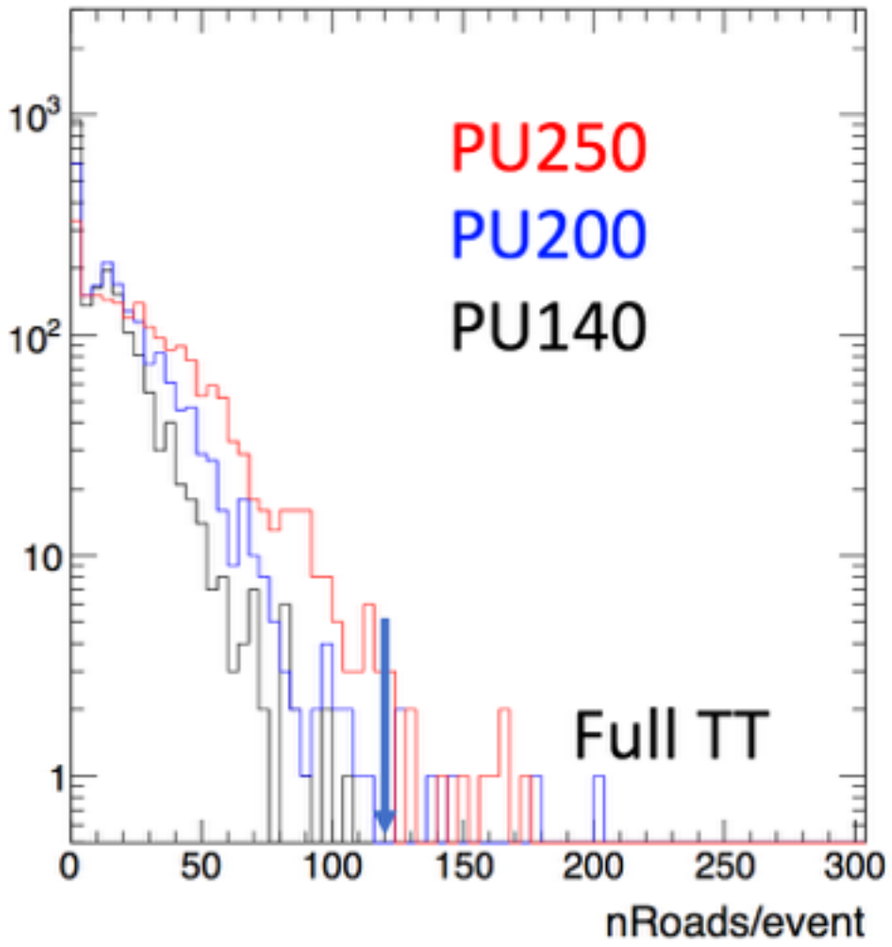


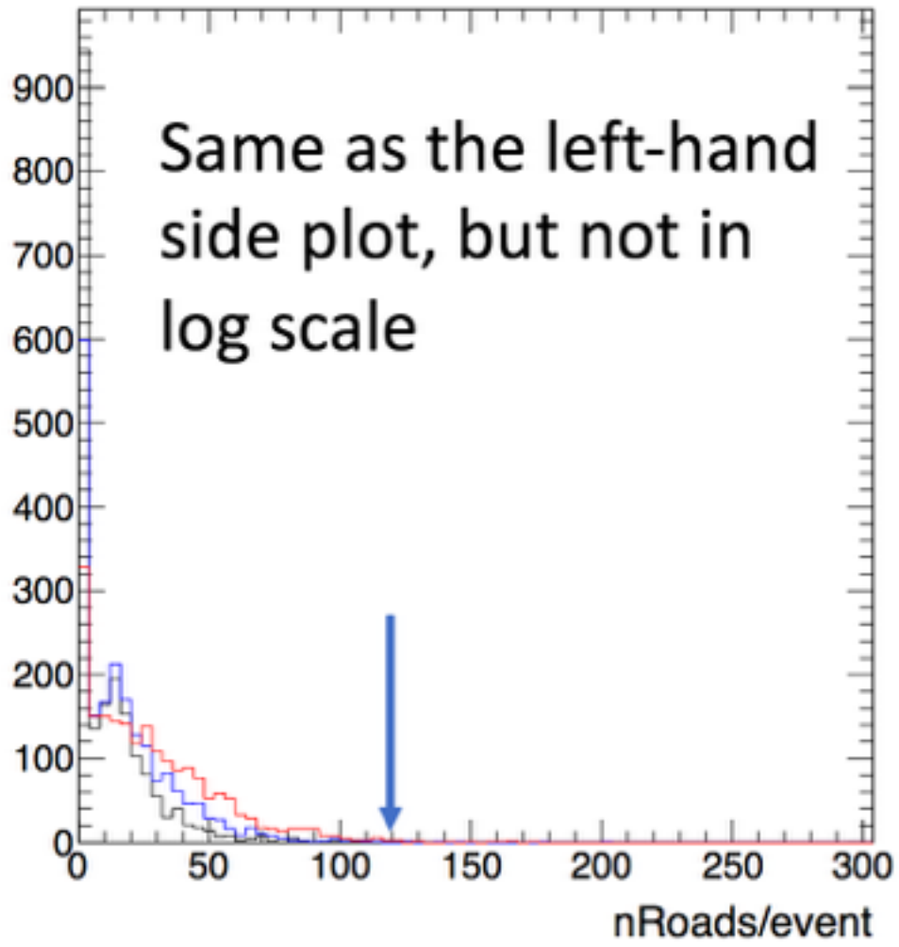
Official deltaS cut (3 GeV inner layers, 2 GeV outer layers)

INFN Istituto Nazionale di Fisica Nucleare

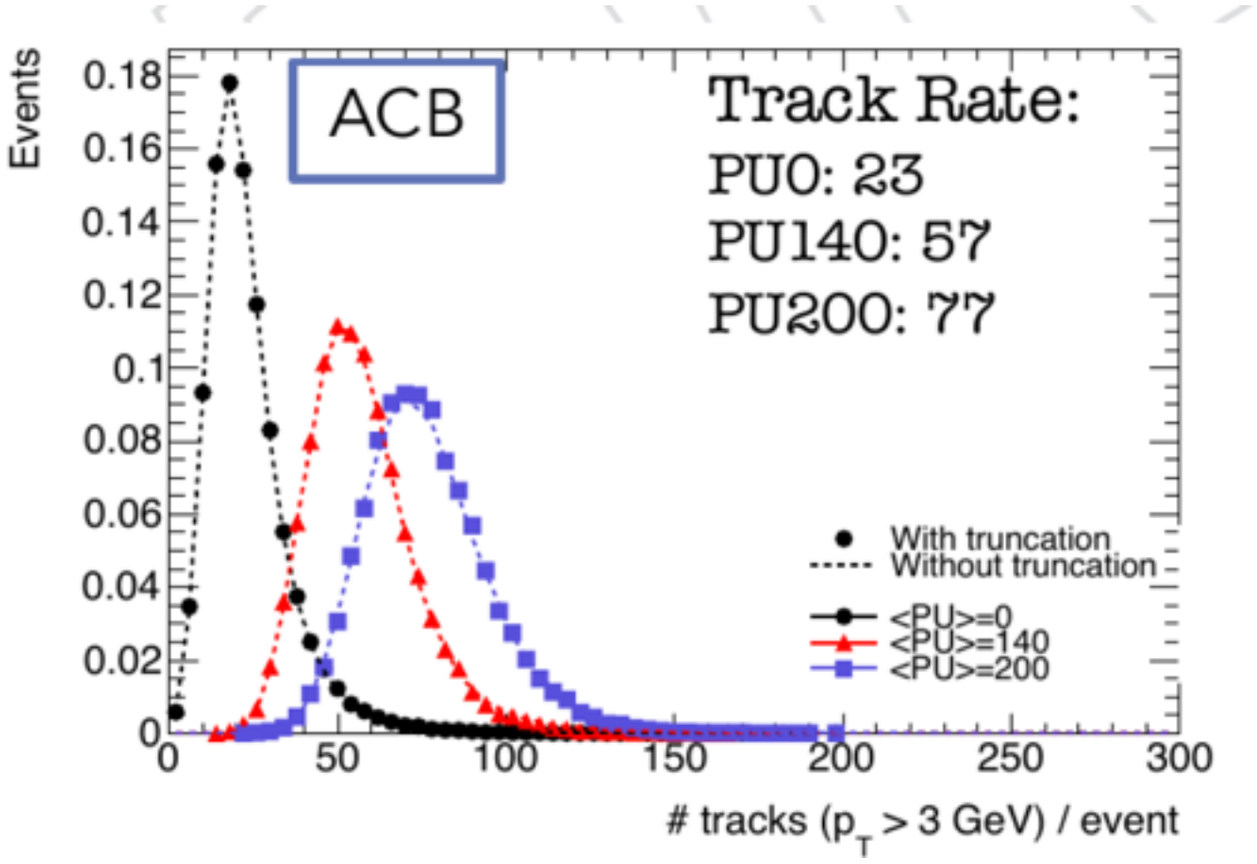
Road per trigger tower

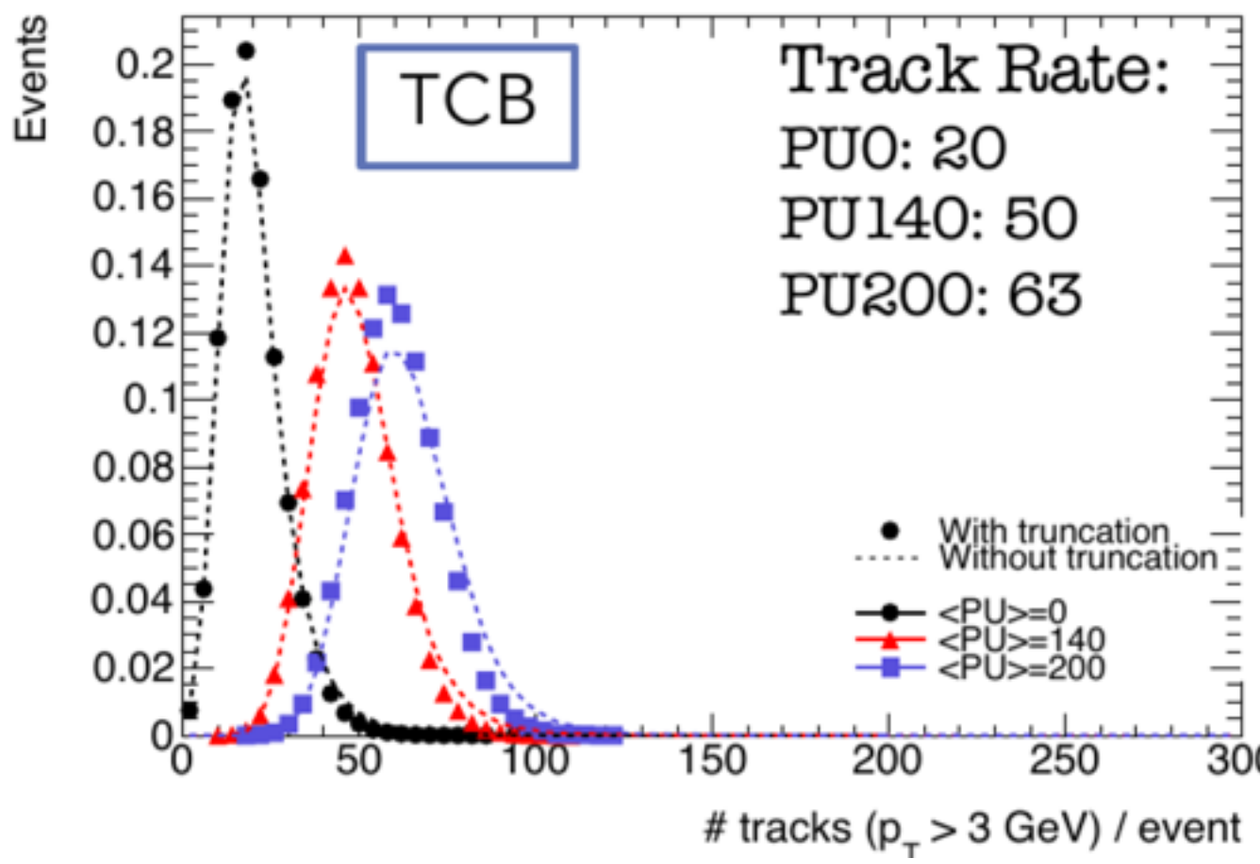
Number of roads per evt (full TT)





Track rate


ACB
Track Rate:
PU0: 23
PU140: 57
PU200: 77
With truncation
Without truncation
<PU>=0
<PU>=140
<PU>=200
Events
tracks (p_T > 3 GeV) / event


TCB
Track Rate:
PU0: 20
PU140: 50
PU200: 63
With truncation
Without truncation
<PU>=0
<PU>=140
<PU>=200
Events
tracks (p_T > 3 GeV) / event