An AM Approach L1 Tracking Trigger for CMS Phase 2 Upgrade
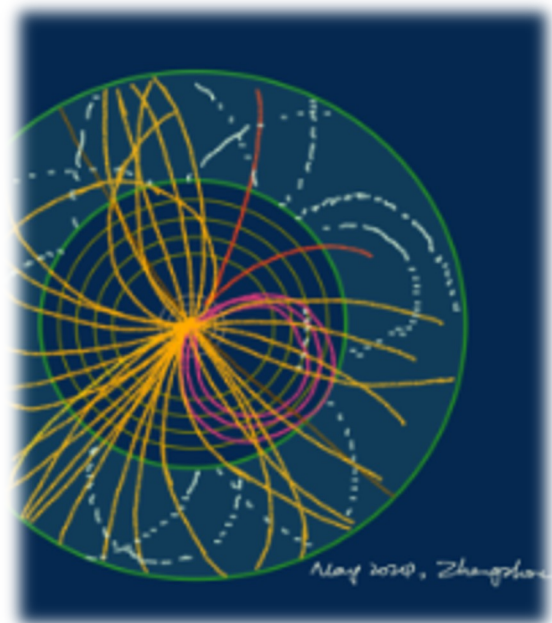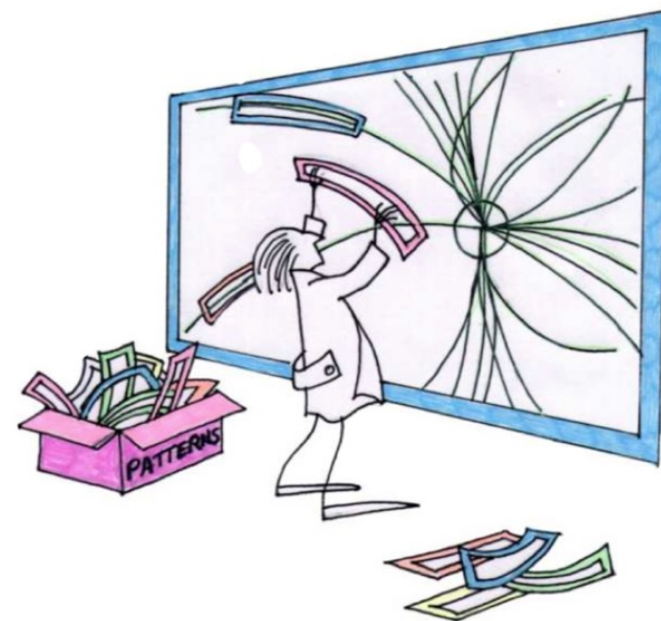
清華大学
Tsinghua University Zhen Hu

Workshop of Tracking
in Particle Physics Experiments

Zhengzhou University

May 18, 2024

Contents

- Motivation

- System overview
  - Multiplex in space and time
  - Associative Memory (AM) introduction

- System demonstration
  - Hardware introduction
  - Demonstrator at Fermilab
  - Performance

- Summary

Physics Motivation: HL-LHC Trigger Upgrade

- **High Luminosity LHC**
  - 40 MHz bunch crossing
  - Up to 200 pileup (high occupancy)
  - Tons of data

- **Current Level-1 Trigger will not work**
  - Current maximum bandwidth
    - only 100 kHz (L1 output rate)
  - For HL-LHC, current trigger system would give
    - EG rate @25 GeV → 100 kHz
    - Overall Trigger Rate → > > 1000 kHz (unsustainable) to reach physics goals
  - Increasing trigger threshold → lose the opportunity of new physics with low threshold

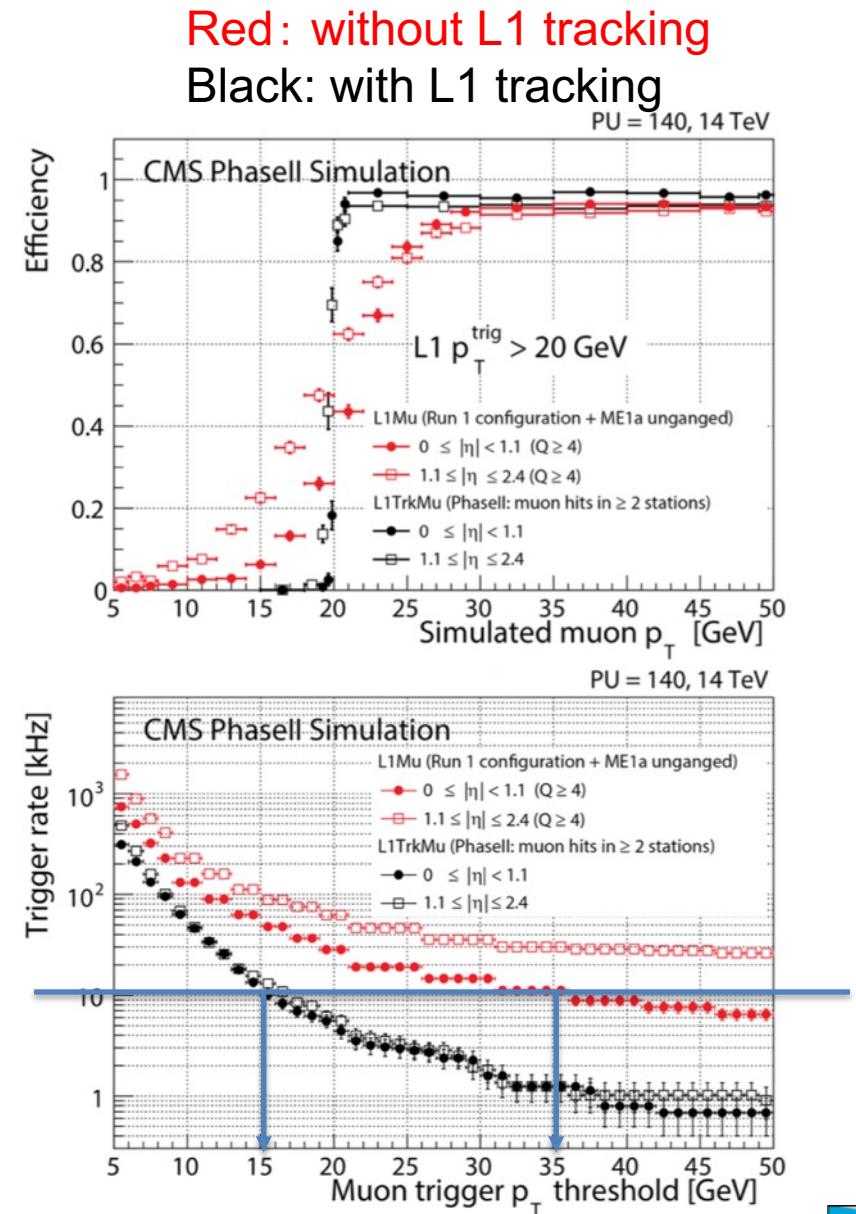Simulated Event Display at 140 PU (102 Vertices)
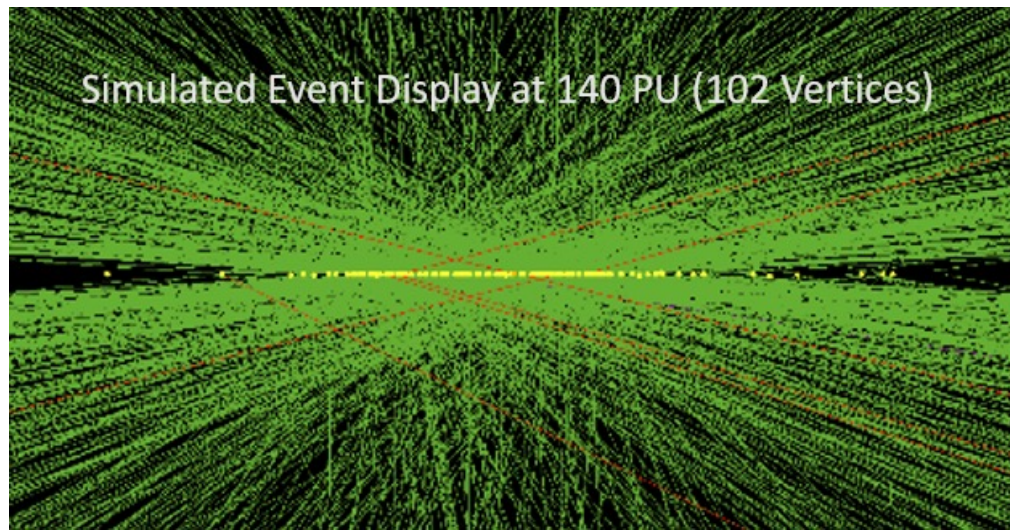
- **Upgrade trigger system**
  - Must increase total bandwidth
  - Must increase trigger capabilities
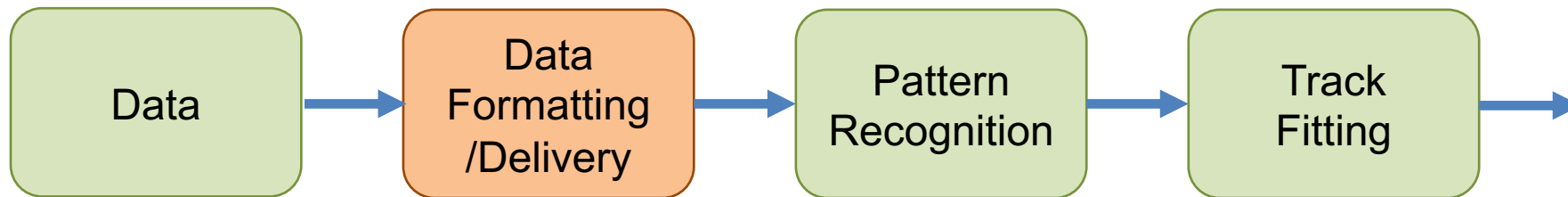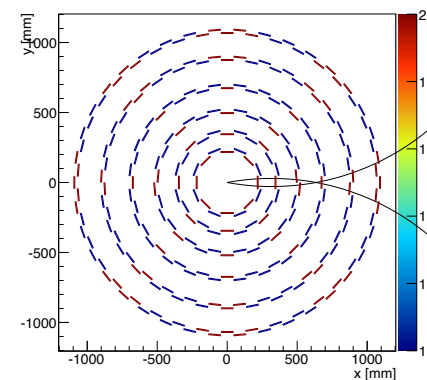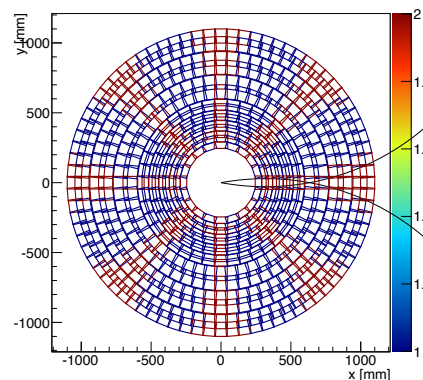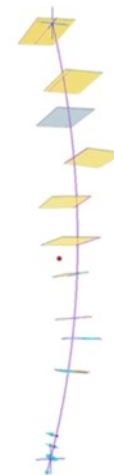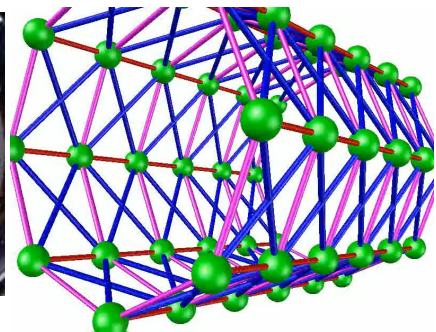 - ❖ **Level-1 Tracking is a completely NEW handle**

- **Silicon based tracking trigger is crucial for CMS Phase2 upgrade**
  - Sharp turn-on efficiency curve
  - Background rate reduction → allows for low object threshold

- **Huge challenges**
  - How to handle readout of the entire tracker?
    - 260 M channel, 40 MHz, 100 Tbps data (after on-detector suppression using pT-modules), 2e4 hits
  - 4 μs latency:
    - Data distribution, track reconstruction, track fitting …
 - Silicon-based L1 tracking trigger has never been realized under these conditions

Red: without L1 tracking
Black: with L1 tracking

TP for Phase-II Upgrade

Solution : divide and conquer

Data → Data Formatting /Delivery → Pattern Recognition → Track Fitting →
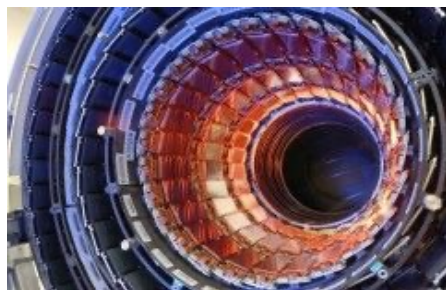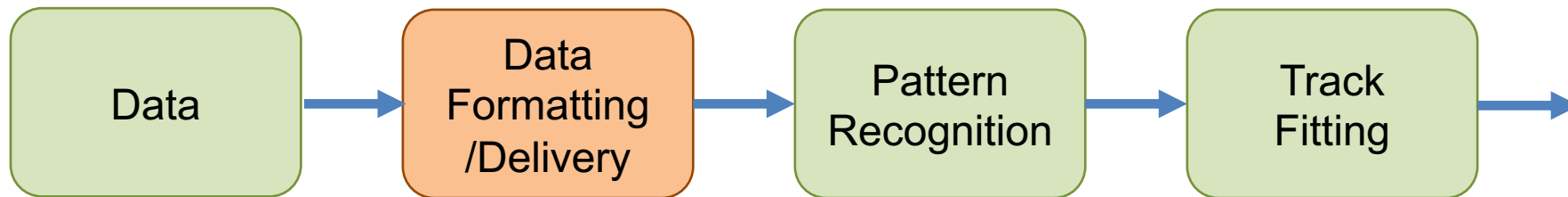
CMS Tracker

- Space parallel
  - 6*8 trigger tower
  - 100 Tbps → ~2 Tbps per tower

Solution : divide and conquer

```
[Data] → [Data Formatting /Delivery] → [Pattern Recognition] → [Track Fitting] →
```
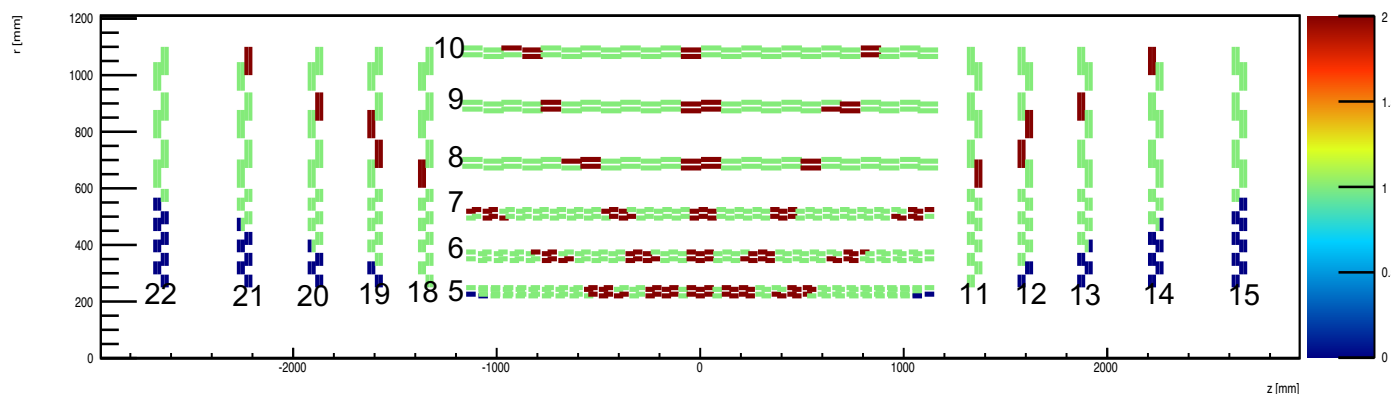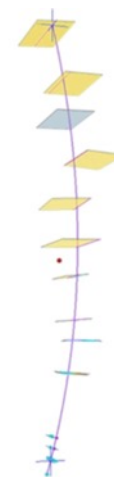
CMS Tracker
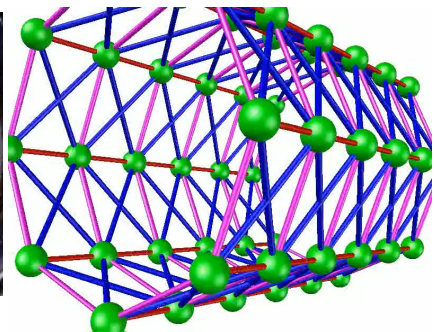
Associative Memory,
Hough Transform,
Tracklet,
… …

- ## Space parallel
  - 6*8 trigger tower
  - 100 Tbps → ~2 Tbps per tower

- ## Time parallel
  - 8x time multiplexing
  - 25 ns → 200 ns
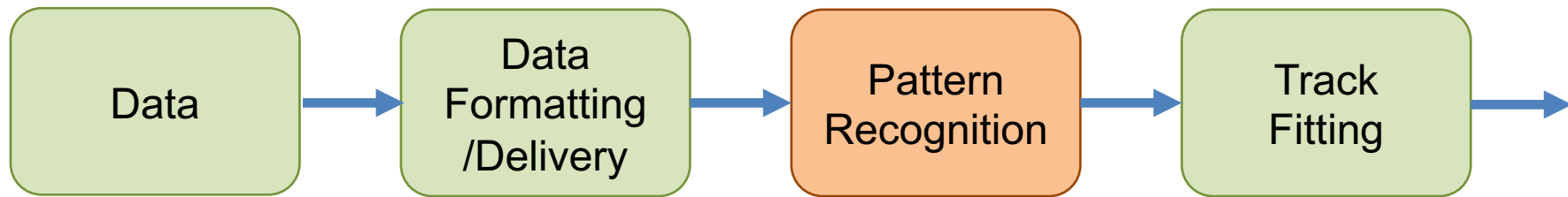
Conventional   ATCA full mesh

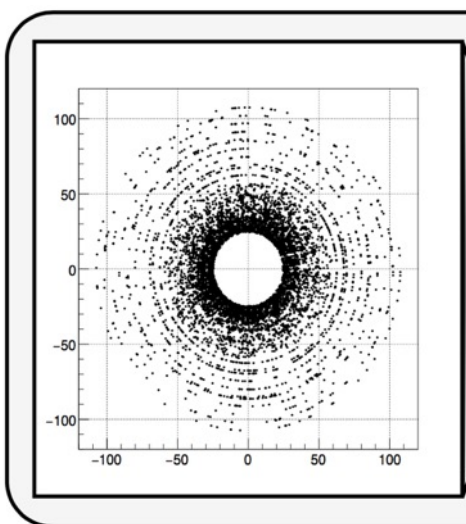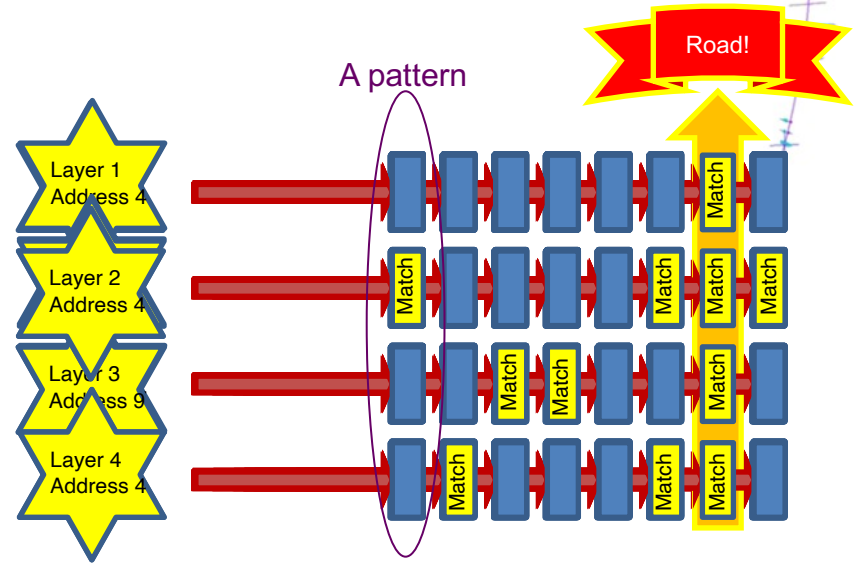*Huge amount of cabling work without ATCA*

Solution : Associative Memory

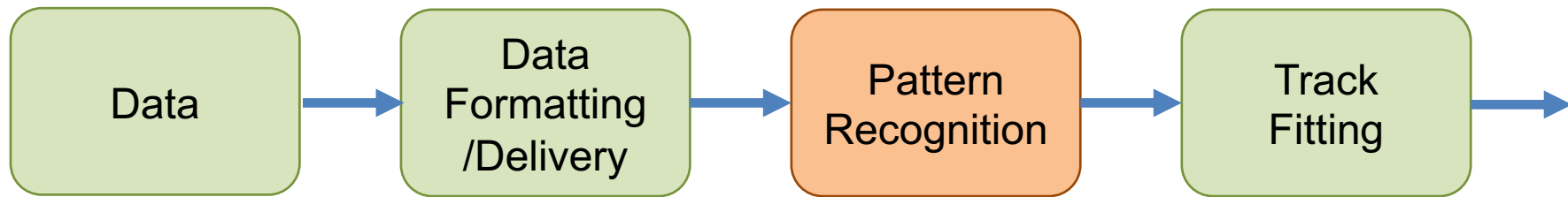Data → Data Formatting /Delivery → Pattern Recognition → Track Fitting

- **Associative Memory**
  - Based on *CAM cells to match* and *majority logic* to associate hits with a set of pre-determined hit patterns that represent allowed track trajectories (massively parallel, *finishes right after all hits arrive*)

Solution : Associative Memory

Data → Data Formatting /Delivery → **Pattern Recognition** → Track Fitting →

- **Associative Memory (AM)**
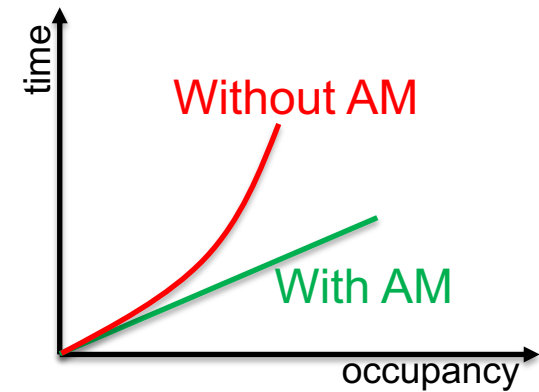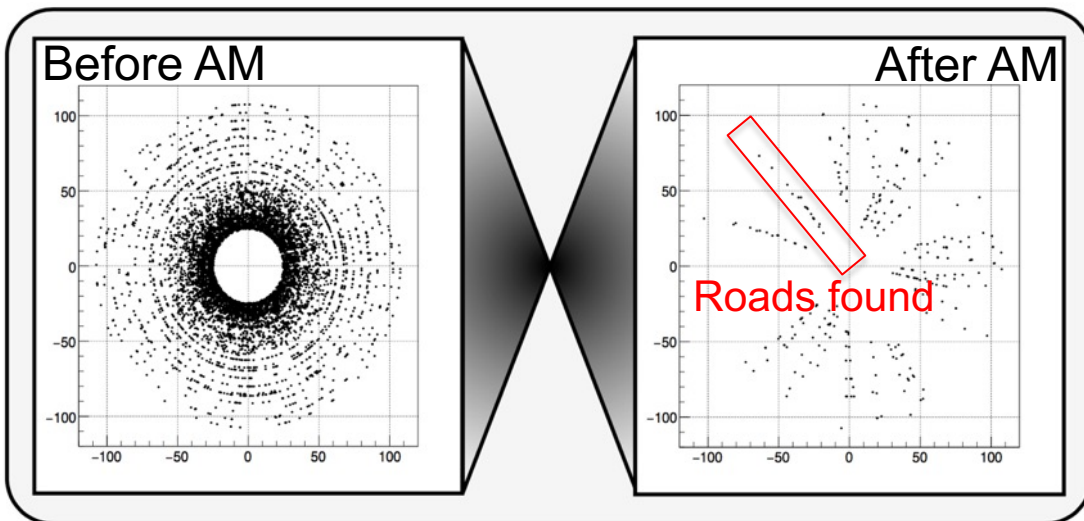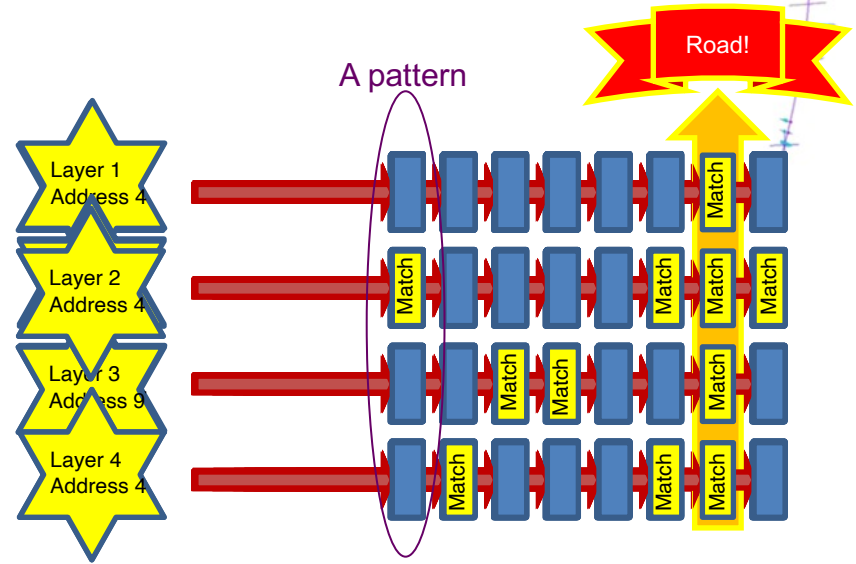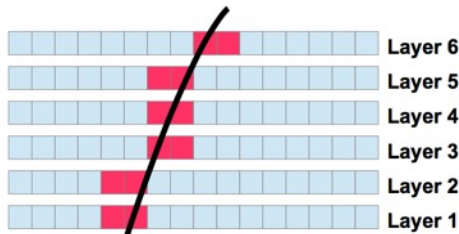 - Based on *CAM cells to match* and *majority logic* to associate hits with a set of pre-determined hit patterns that represent allowed track trajectories (massively parallel, *finishes right after all hits arrive*)
- **Perfect tool to mediate the high pileup effect**
  - Avoiding the typical power law dependence of execution time on occupancy

A pattern

Road!

Layer 1 Address 4
Layer 2 Address 4
Layer 3 Address 9
Layer 4 Address 4

Match

Before AM

After AM

Roads found
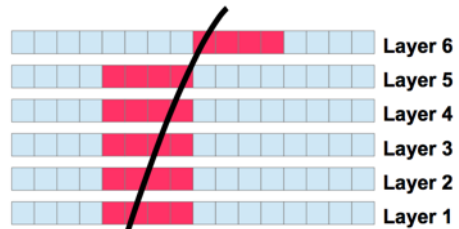
time

Without AM

With AM

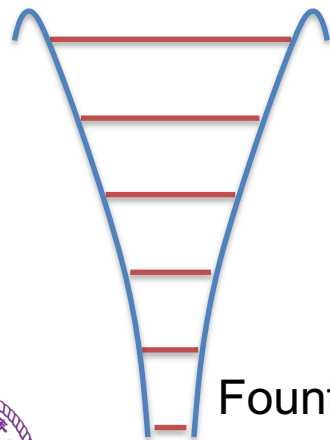occupancy

AM Pattern and Bank

- A pattern is a low resolution track
    - Made of 1 superstrip (SS) per layer
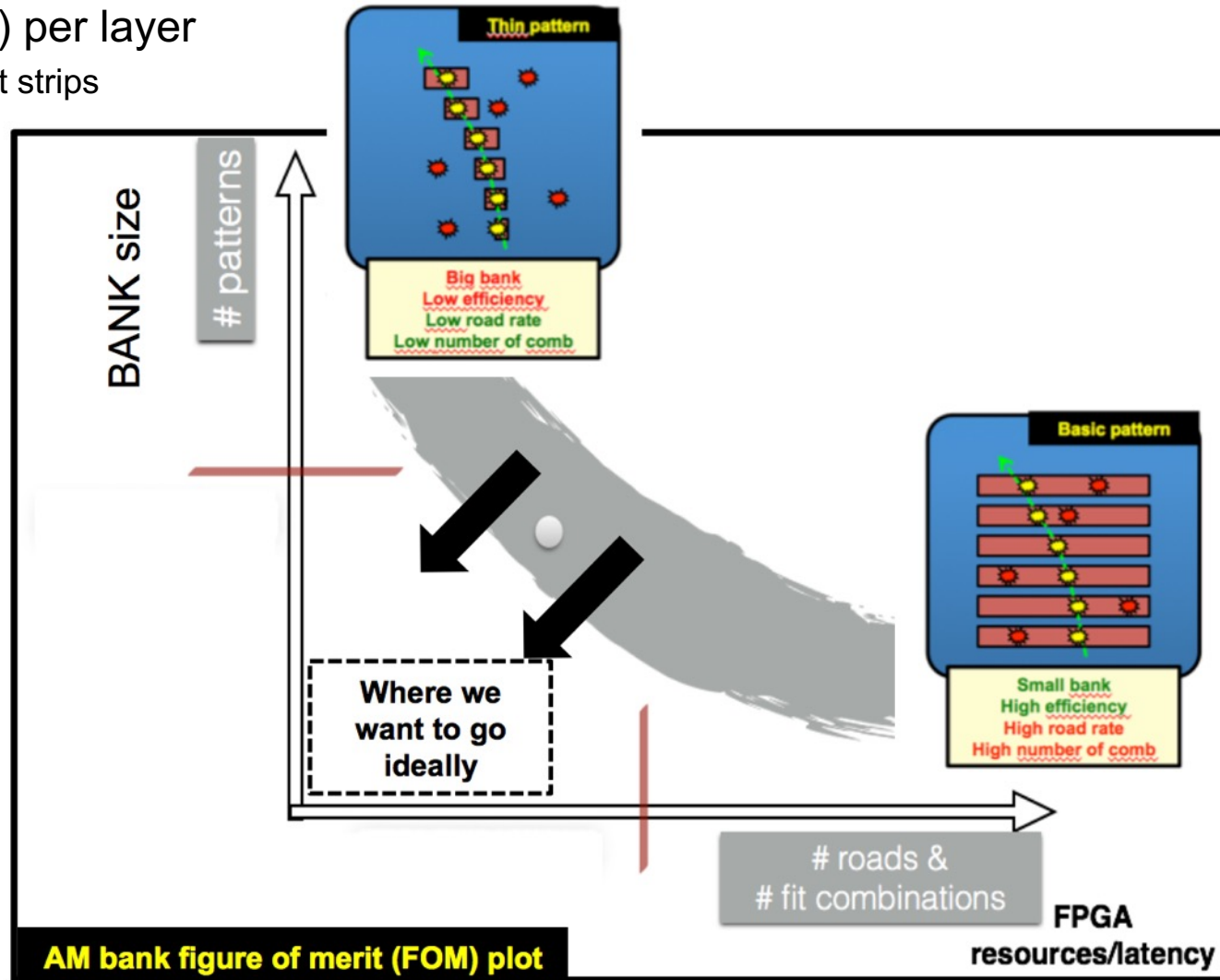   - A SS is a group of adjacent strips

Ex : supertrip size = 2 strips

Ex : supertrip size = 4 strips

Fountain SS

Thin pattern

Big bank
Low efficiency
Low road rate
Low number of comb

BANK size

patterns

Basic pattern

Small bank
High efficiency
High road rate
High number of comb

Where we want to go ideally

roads &
fit combinations
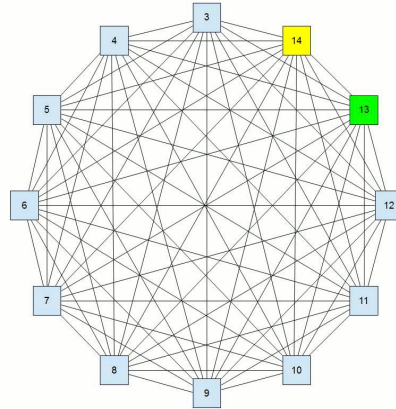
FPGA resources/latency

AM bank figure of merit (FOM) plot

More powerful AM => less demand on the FPGA
More powerful FPGA => less demand on the AM

FNAL based hardware R&D

ATCA shelf

The full mesh backplane interconnections effectively blur the distinction between FPGAs, and allow data sharing in both space and time
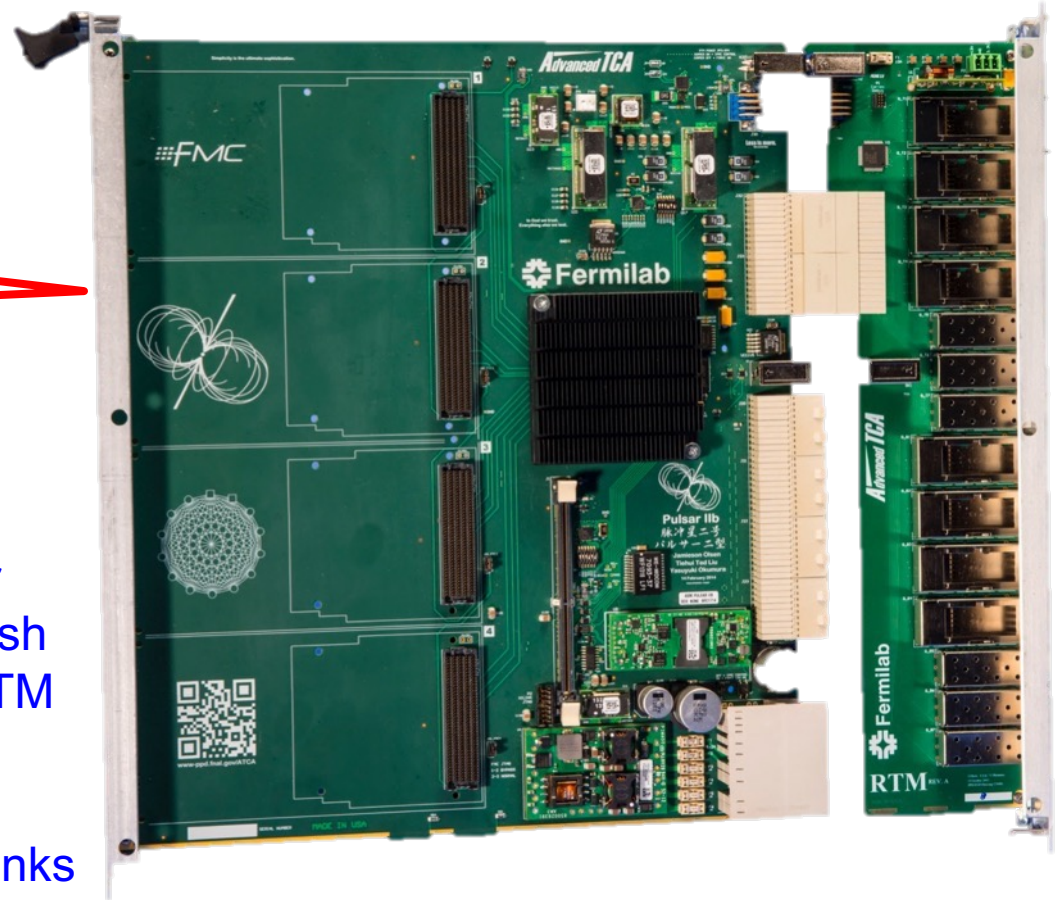
FNAL based hardware R&D

ATCA shelf

Pulsar2b

Pulsar2b RTM (Rear Transition Module)

We tested the data transfer performance for the full mesh back plan, Pulsar2b and RTM (10 Gbps)

100 Tbps / 48 tower / 400 links = 5 Gbps

- A custom ATCA full mesh enabled FPGA-based processor board
- Designed with the goal of creating a scalable architecture abundant in flexible, non-blocking, high bandwidth interconnections
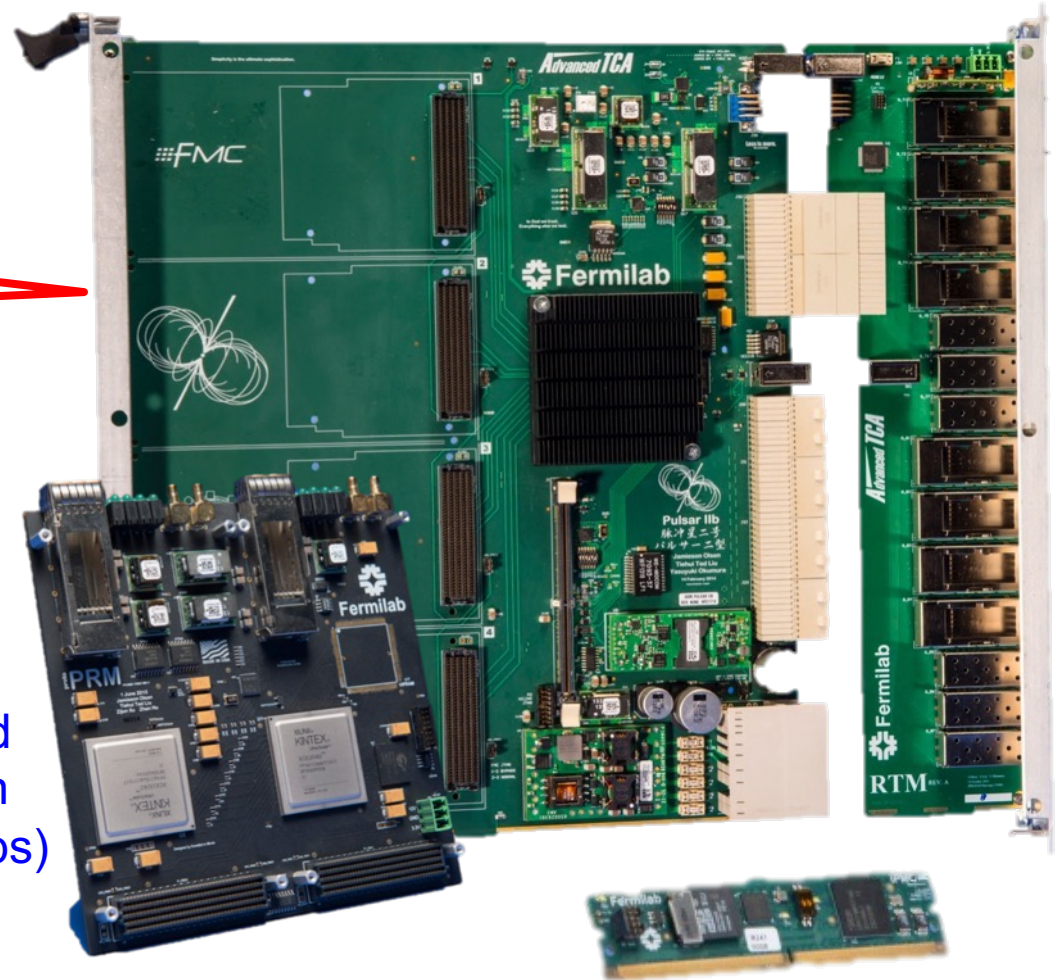
FNAL based hardware R&D

ATCA shelf

ATCA Processing Blade: Pulsar2b

Pulsar2b RTM (Rear Transition Module)

We tested the data transfer performance between PRM and Pulsar2b (10 Gbps) and the interconnection between two FPGAs in PRM (16 Gbps)
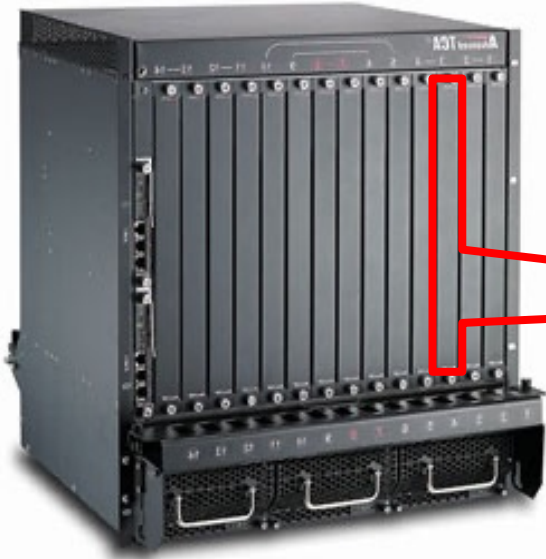
PRM (Pattern Recognition Mezzanine Card)

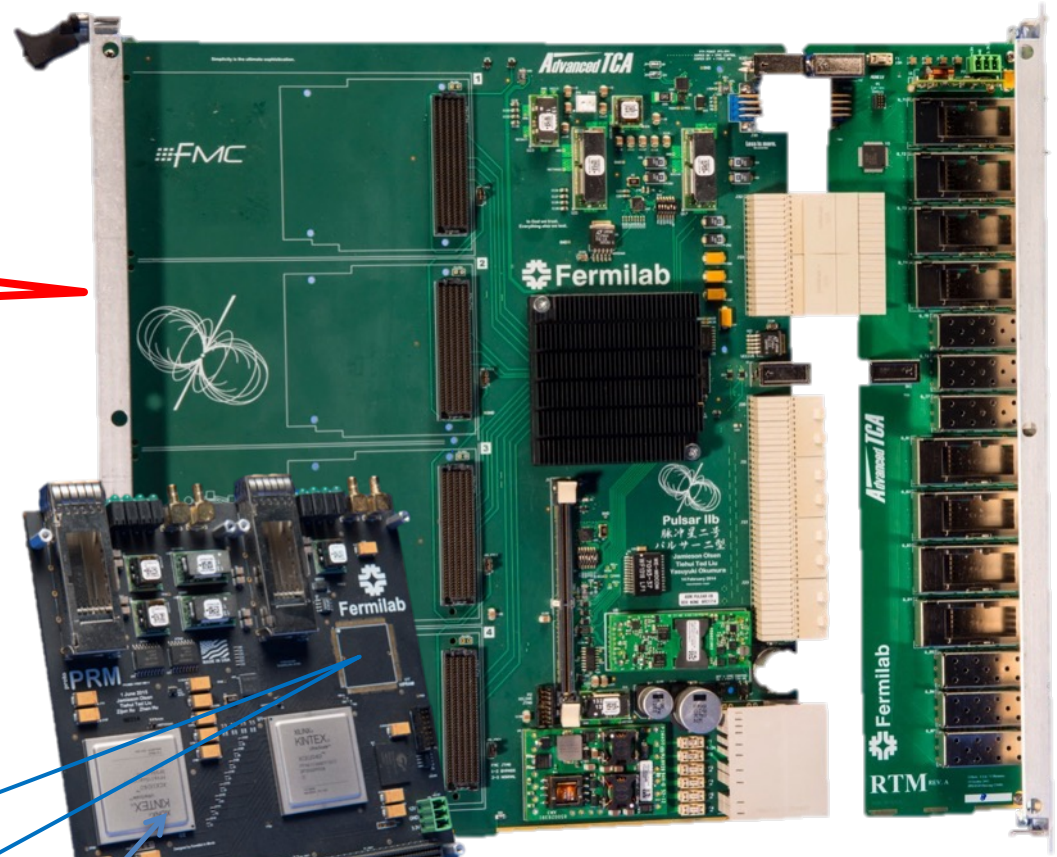IPMC (Intelligent platform management controller)
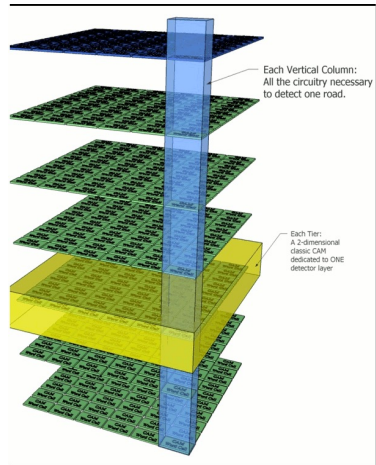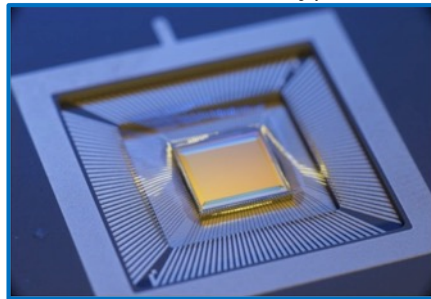
FNAL based hardware R&D

ATCA shelf

ATCA Processing Blade: Pulsar2b

Pulsar2b RTM (Rear Transition Module)

VIPRAM ASIC
(Vertically Integrated Pattern Recognition Associative Memory)

PRM (Pattern Recognition Mezzanine Card)

IPMC (Intelligent platform management controller)

AM can be emulated with FPGA

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 Pattern Modules
    - I/O Tier -> fired roads serialization and output
  - 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

- CAM tier logic is optimized for 7-Series/UltraScale FPGA architecture
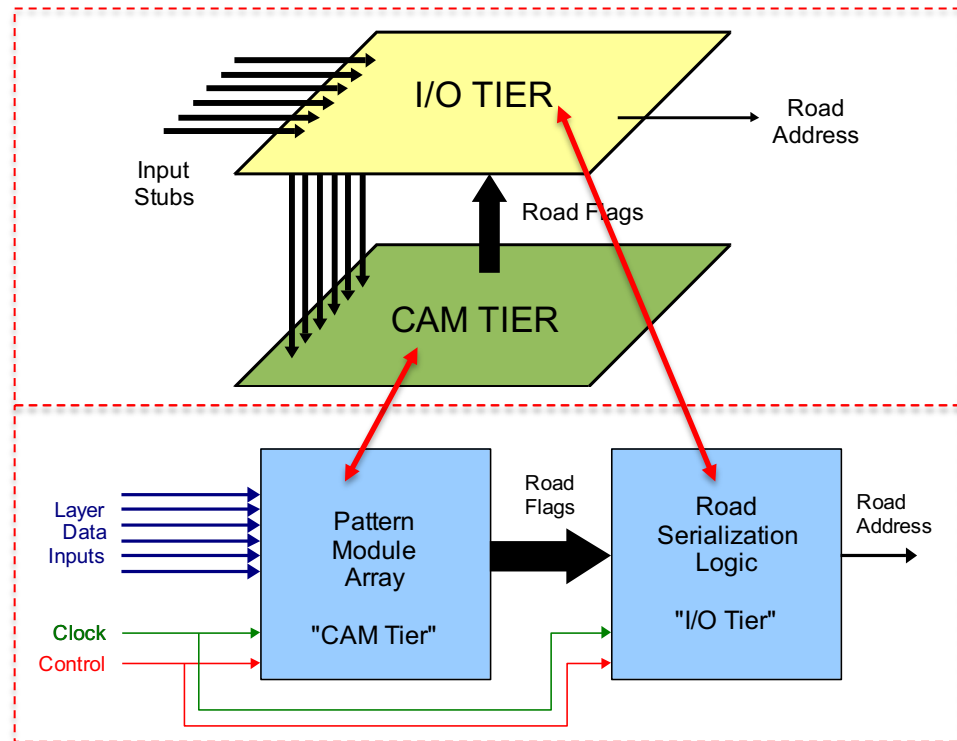
Excellent hardware performance

- ATCA shelf
  - 10 blades for parallel processing
  - Full mesh backplane is a natural solution for time multiplexing
    - All of the 56 bidirectional links among 8 Pulsar2b boards were tested at 10Gbps

- Rear Transition Module
  - 10 QSFP bidirectional links
    - 10 Gbps per link achieved

- PRM performance
  - Communication between Pulsar2b and PRM FPGAs
    - 10 Gbps achieved
  - Two latest generation of Xilinx FPGAs
    - Interconnection achieved 16.3 Gbps
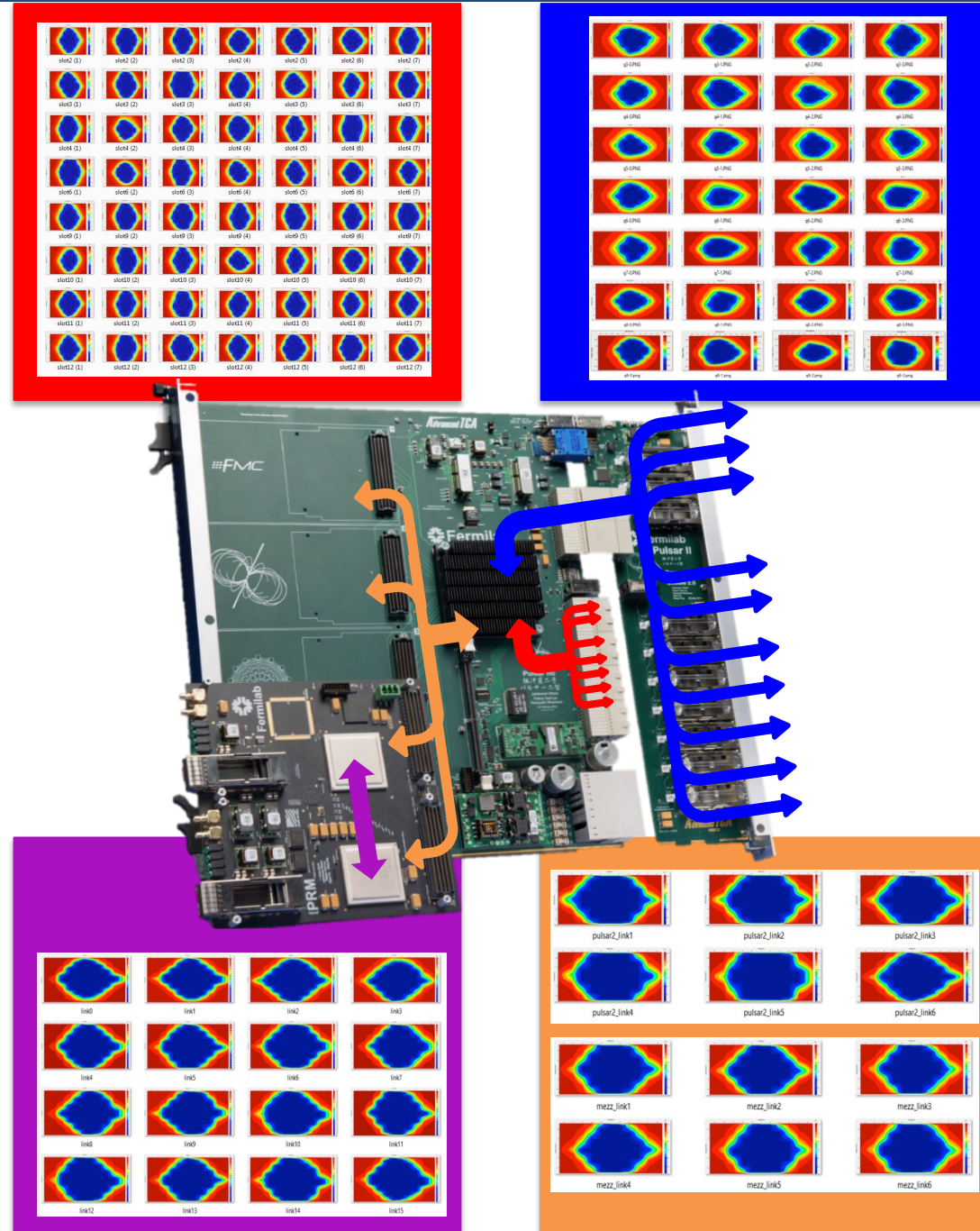
Full system demonstration
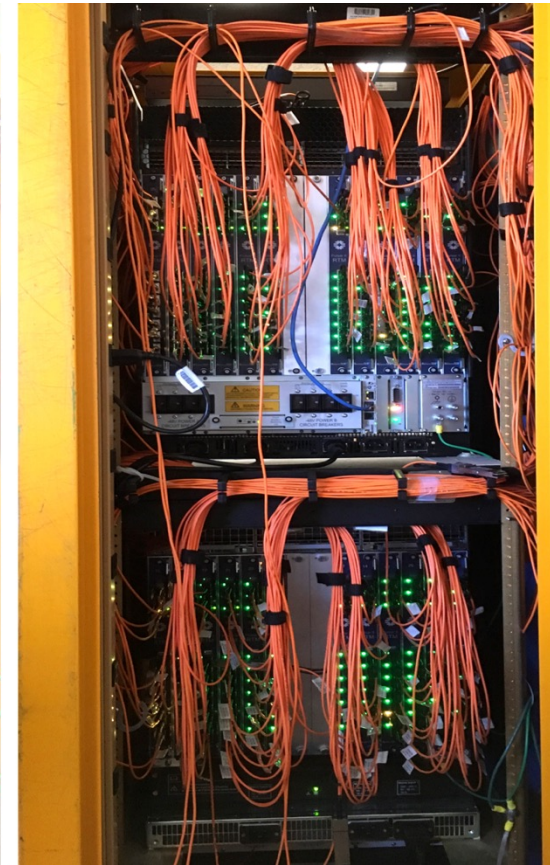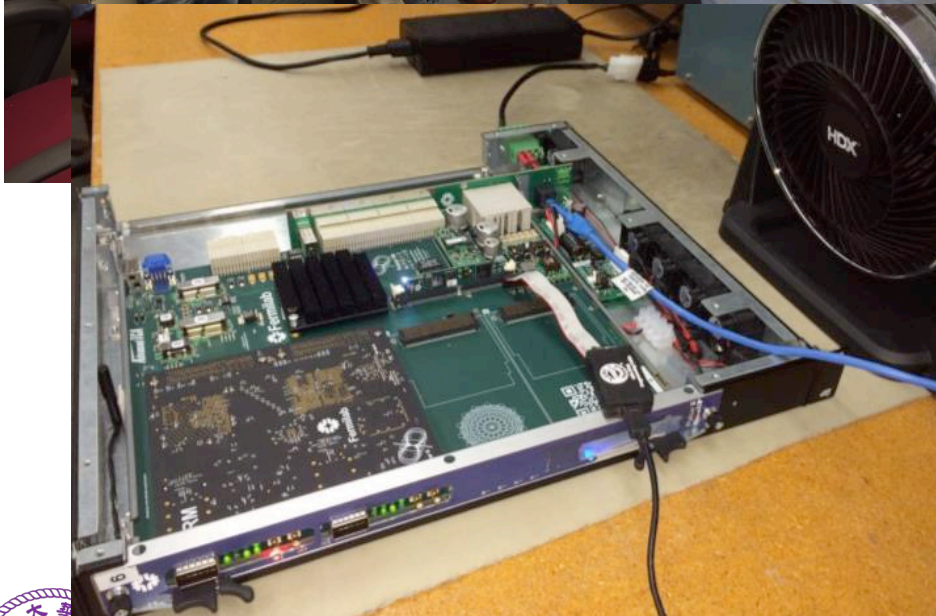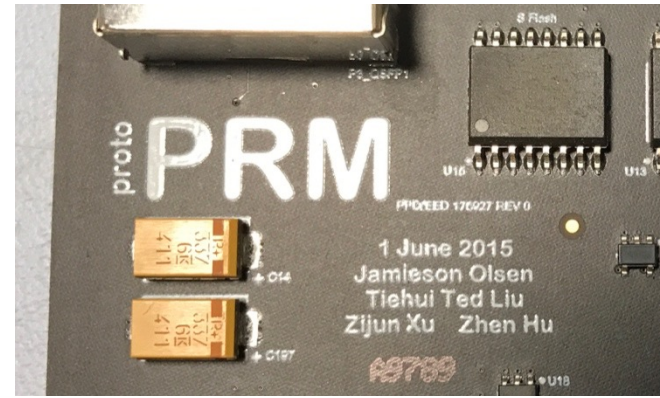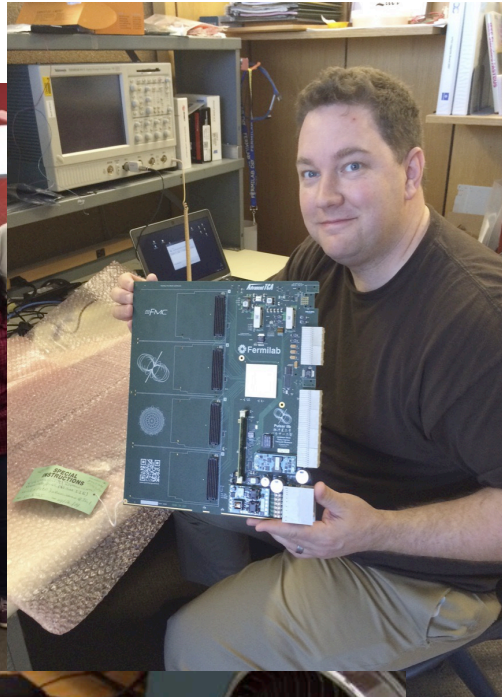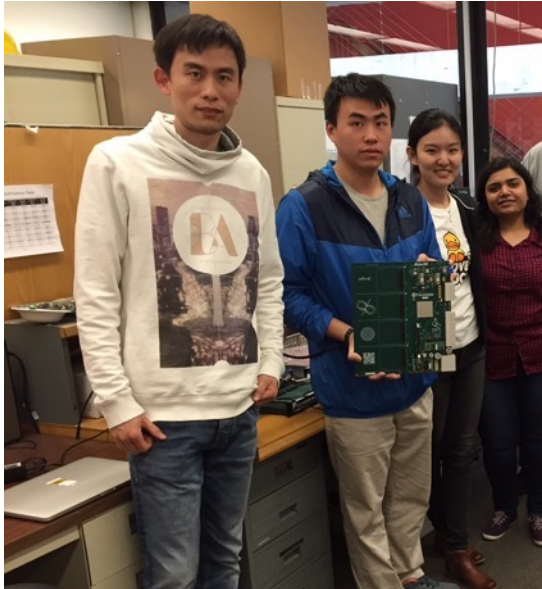
Full system demonstration
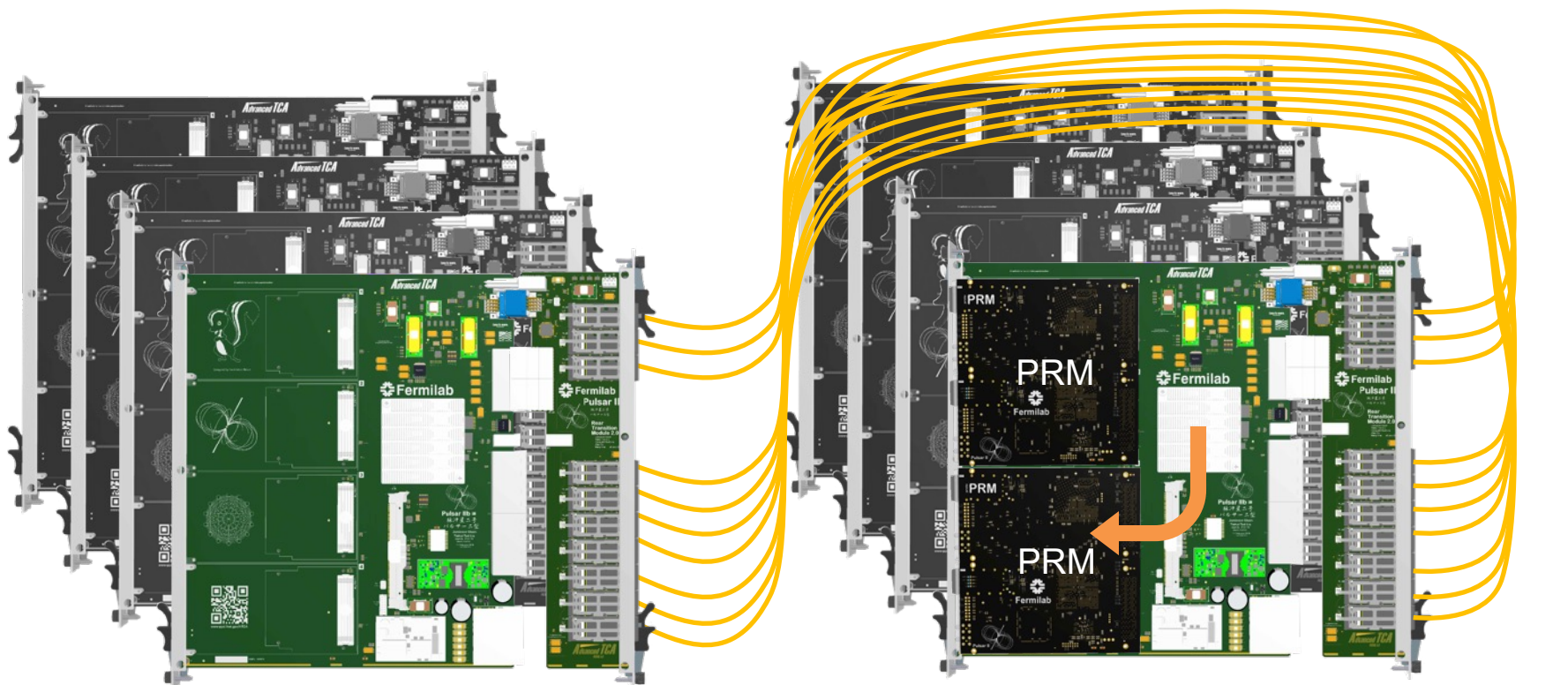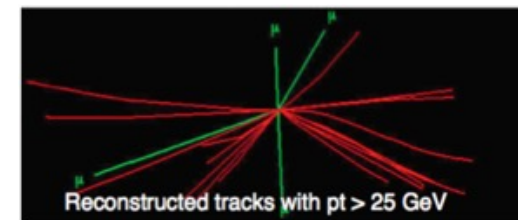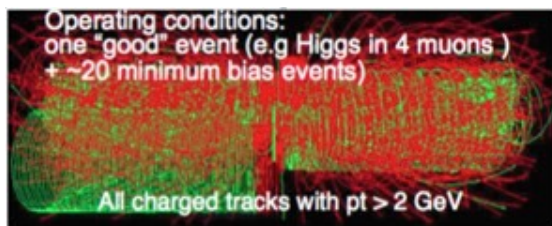
- Using the technology today to demonstrate track trigger feasibility

Data source board
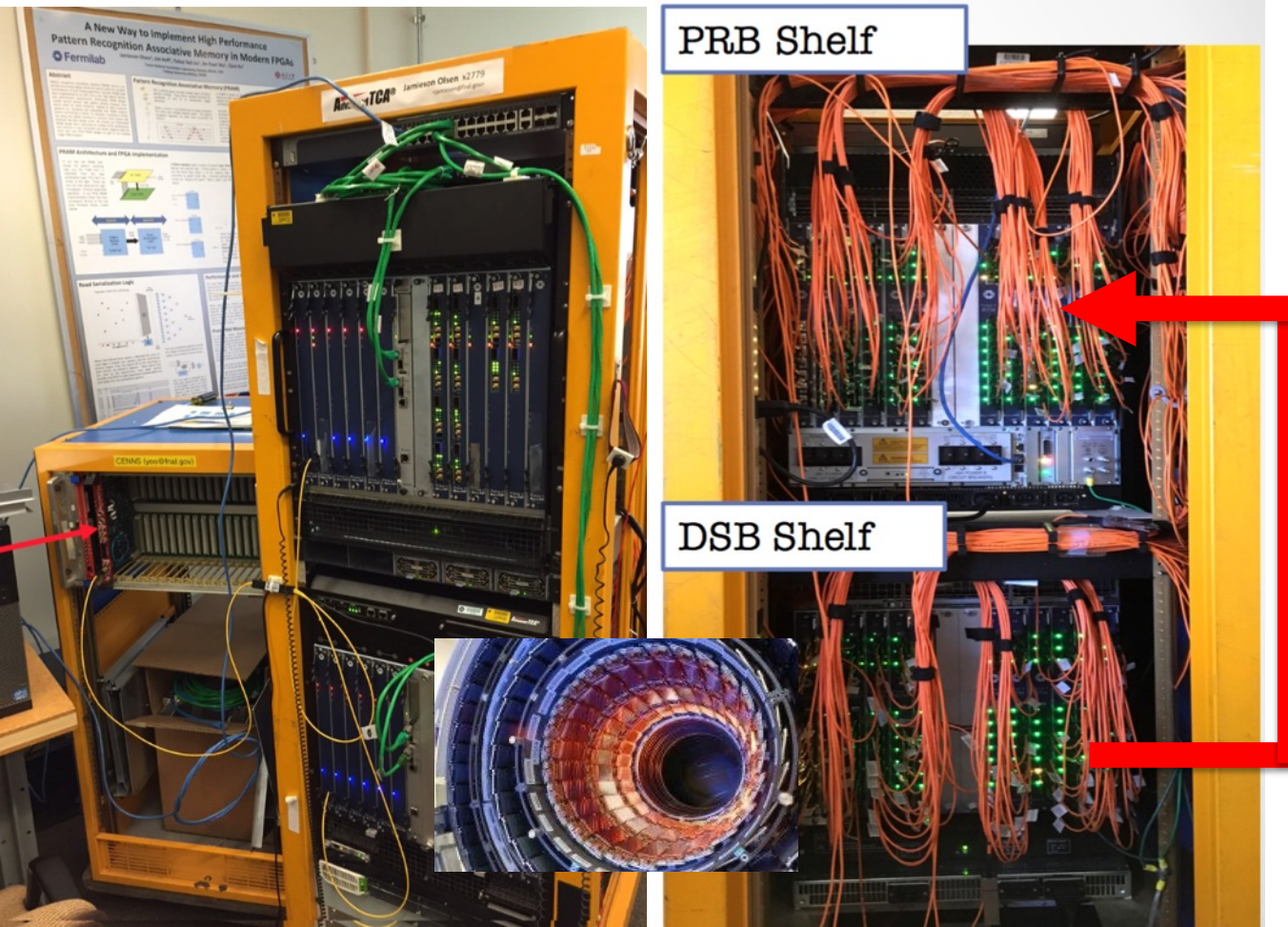
Pattern recognition board

PRM

PRM

Operating conditions:
one "good" event (e.g Higgs in 4 muons)
+ ~20 minimum bias events)

All charged tracks with pt > 2 GeV

Reconstructed tracks with pt > 25 GeV

Full system demonstration at Fermilab

- Using the technology today to demonstrate tracking trigger feasibility
- For one Trigger Tower: two shelves fully loaded with Pulsar2b boards

Front view

PRB Shelf

DSB Shelf

Back view

Pattern Recognition Board (PRB) shelf

- 12 Pulsar2b with PRM Mezzanines
- *Bandwidth between any pair of Pulsars is 20Gbps*

Data Source Board (DSB) shelf

- 12 Pulsar2b to Emulates the detector output of ~400 modules

120 QSFP+ fibers

- Each with 4 bidirectional lanes each running at 10Gbps
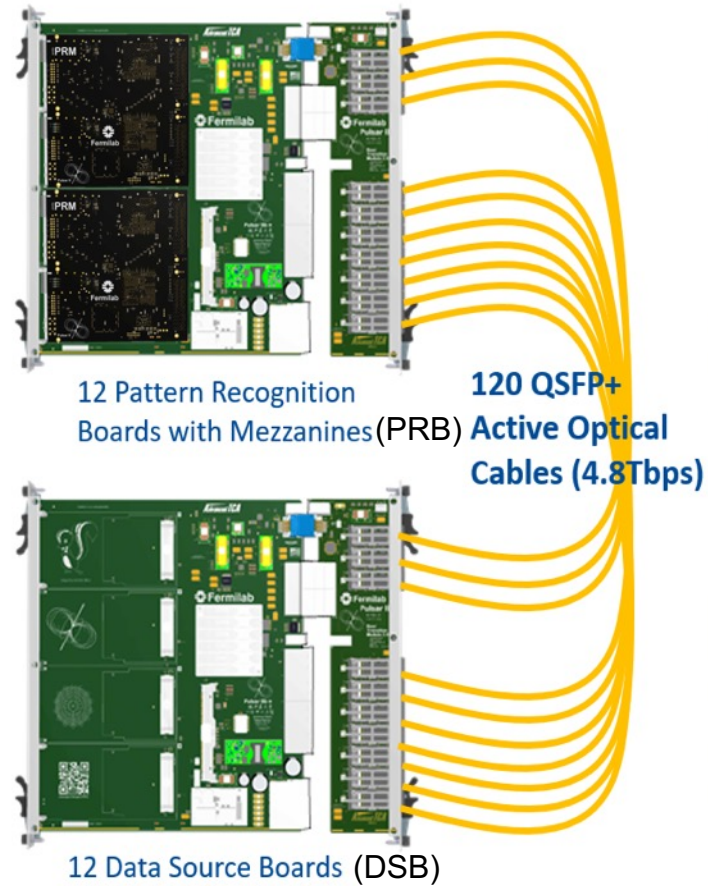- *Capable of sourcing up to 4.8 Tbps data with full shelf*

- All Pulsar2b boards in the system are synchronized to a common master clock optical link

1. Provided by a CERN TTCcx board that encodes a simulated LHC 40MHz bunch crossing clock and various control signals
2. This optical link is received by one Pulsar2b board in each shelf
3. From this Pulsar2b board the master clock and other control bits are distributed to other Pulsar2b boards in the shelf over dedicated clocks on the ATCA backplane
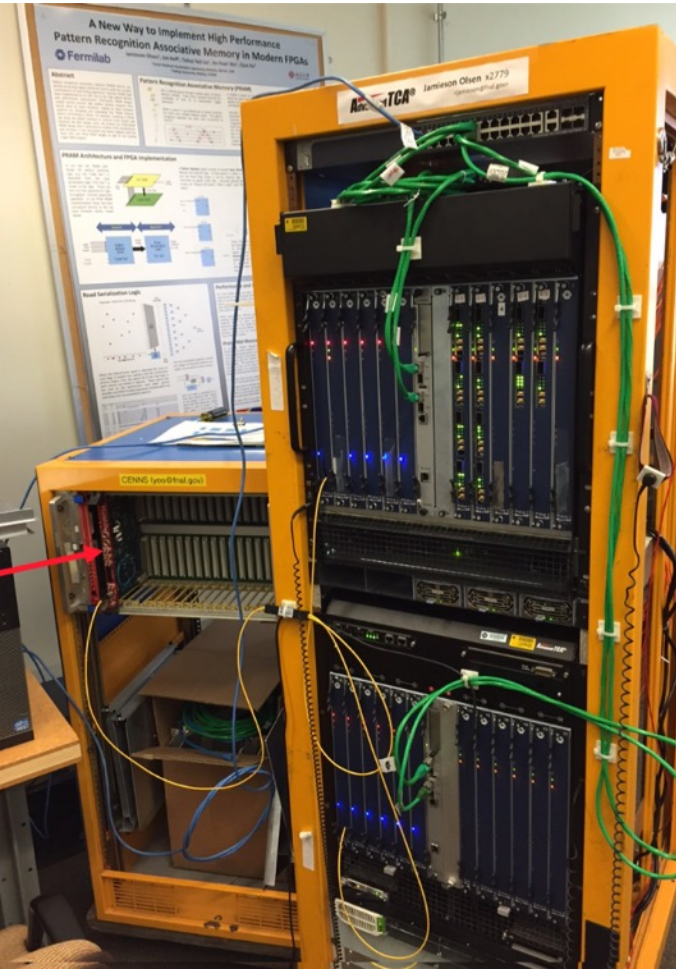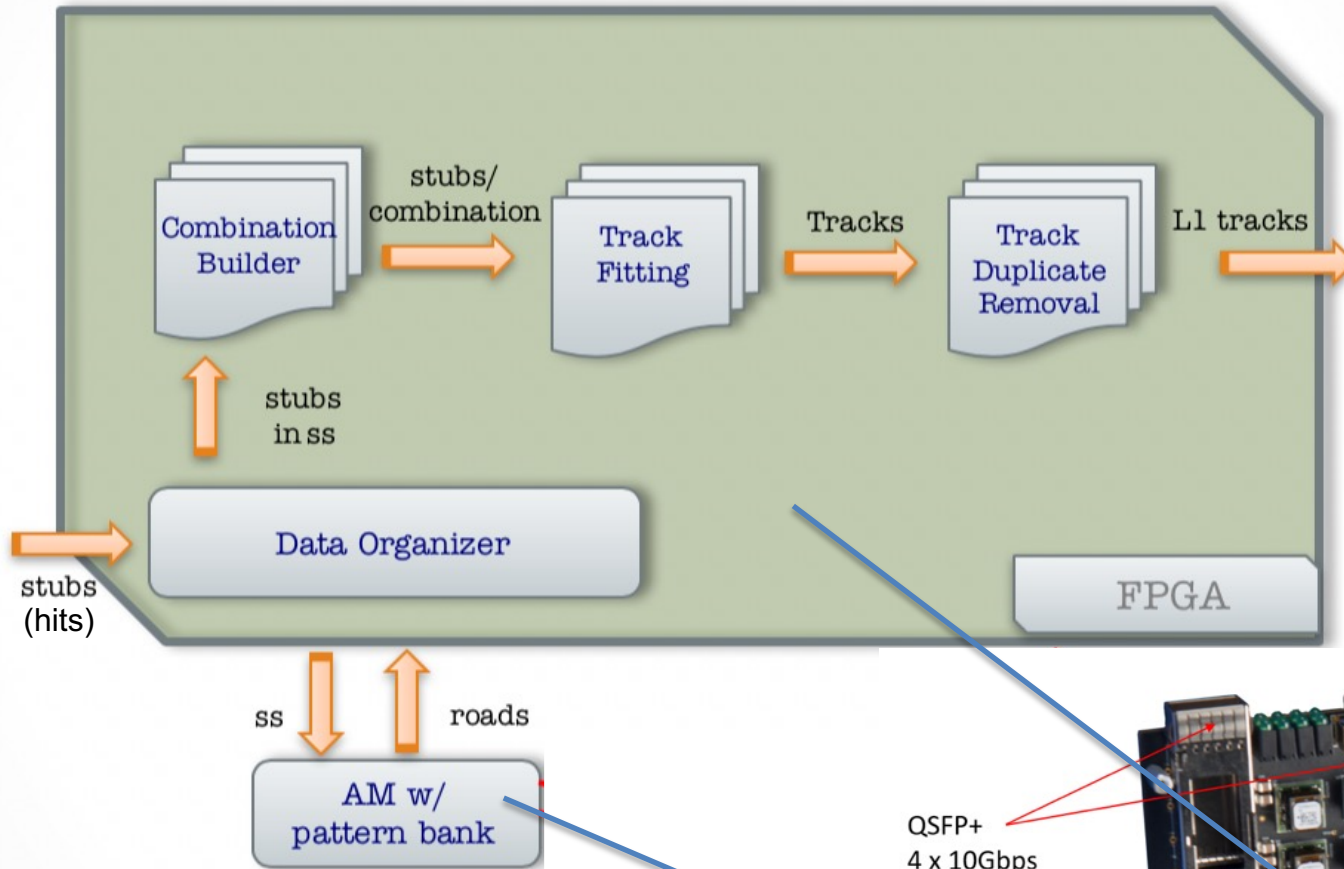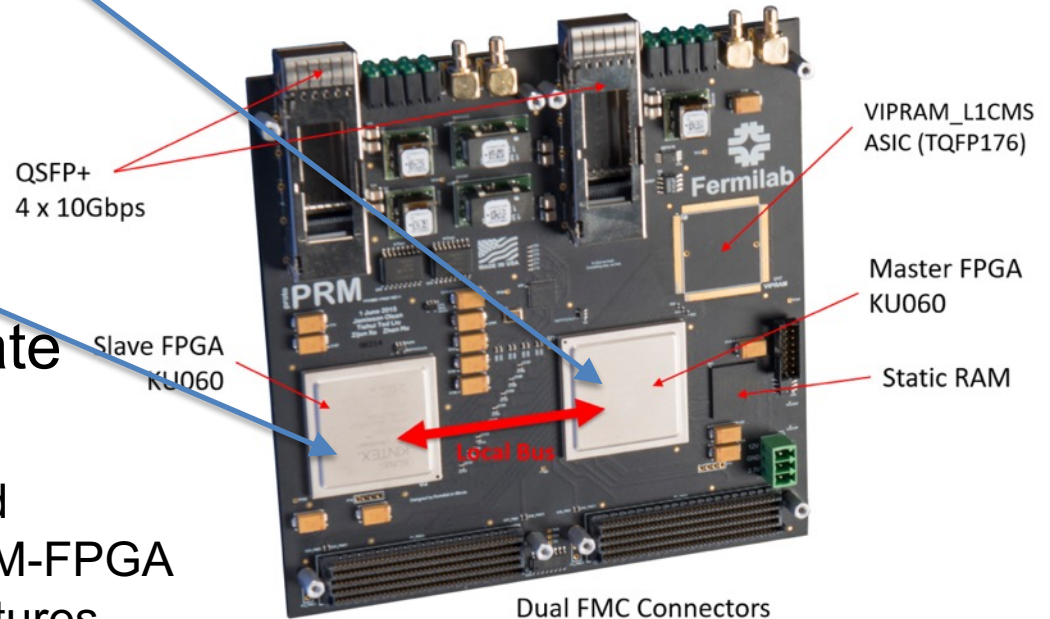
Full system demonstration at Fermilab

1. Simulated event data is first loaded into the DSB FPGAs
2. When triggered by bunch crossing signal, data is transmitted over the 120 QSFP+ optical links at full speed to the PRBs in the upper shelf
3. The 12 PRBs receive the incoming data, perform sophisticated time multiplexed data transfers over the full mesh ATCA backplane
4. Finally the event data sent to the FMC Pattern Recognition Mezzanine (PRM) cards
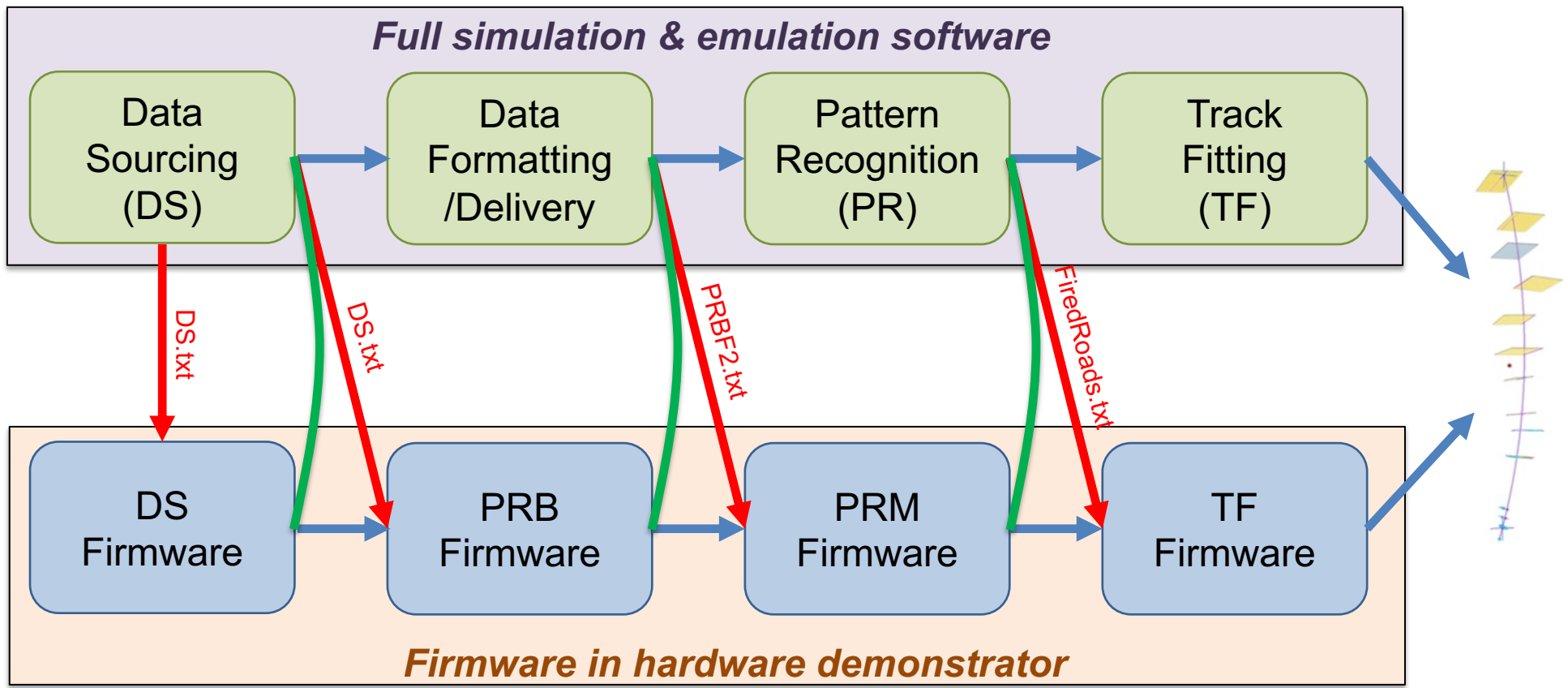

12 Pattern Recognition Boards with Mezzanines (PRB)

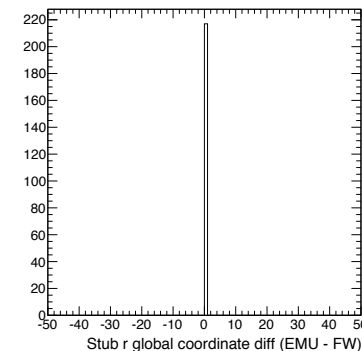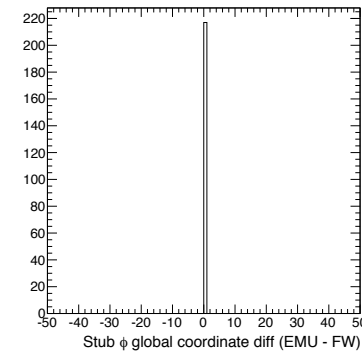120 QSFP+ Active Optical Cables (4.8Tbps)

12 Data Source Boards (DSB)

- ## Slave FPGA is used to emulate associative memory
  - Enables us to quickly evaluate and optimize the performance of the AM-FPGA interface and internal AM logic features
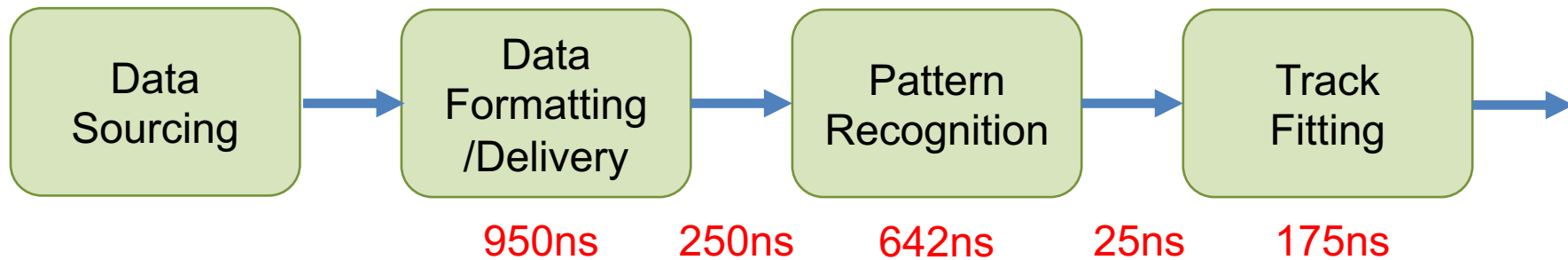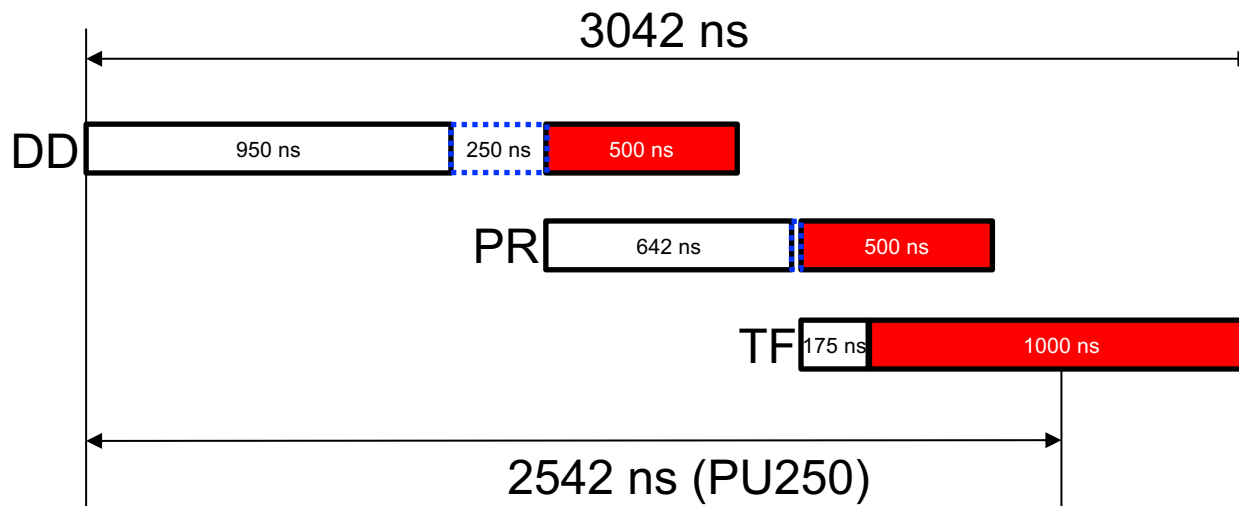
Demonstrator validation

Data Sourcing (DS) → Data Formatting /Delivery → Pattern Recognition (PR) → Track Fitting (TF)

DS.txt
DS.txt
PRBF2.txt
FiredRoads.txt

Firmware in hardware demonstrator

DS Firmware → PRB Firmware → PRM Firmware → TF Firmware

- Hardware and emulator perfectly matched
  - Output from each stage validated bit-by-bit
- With the full chain demonstrator, we have measured
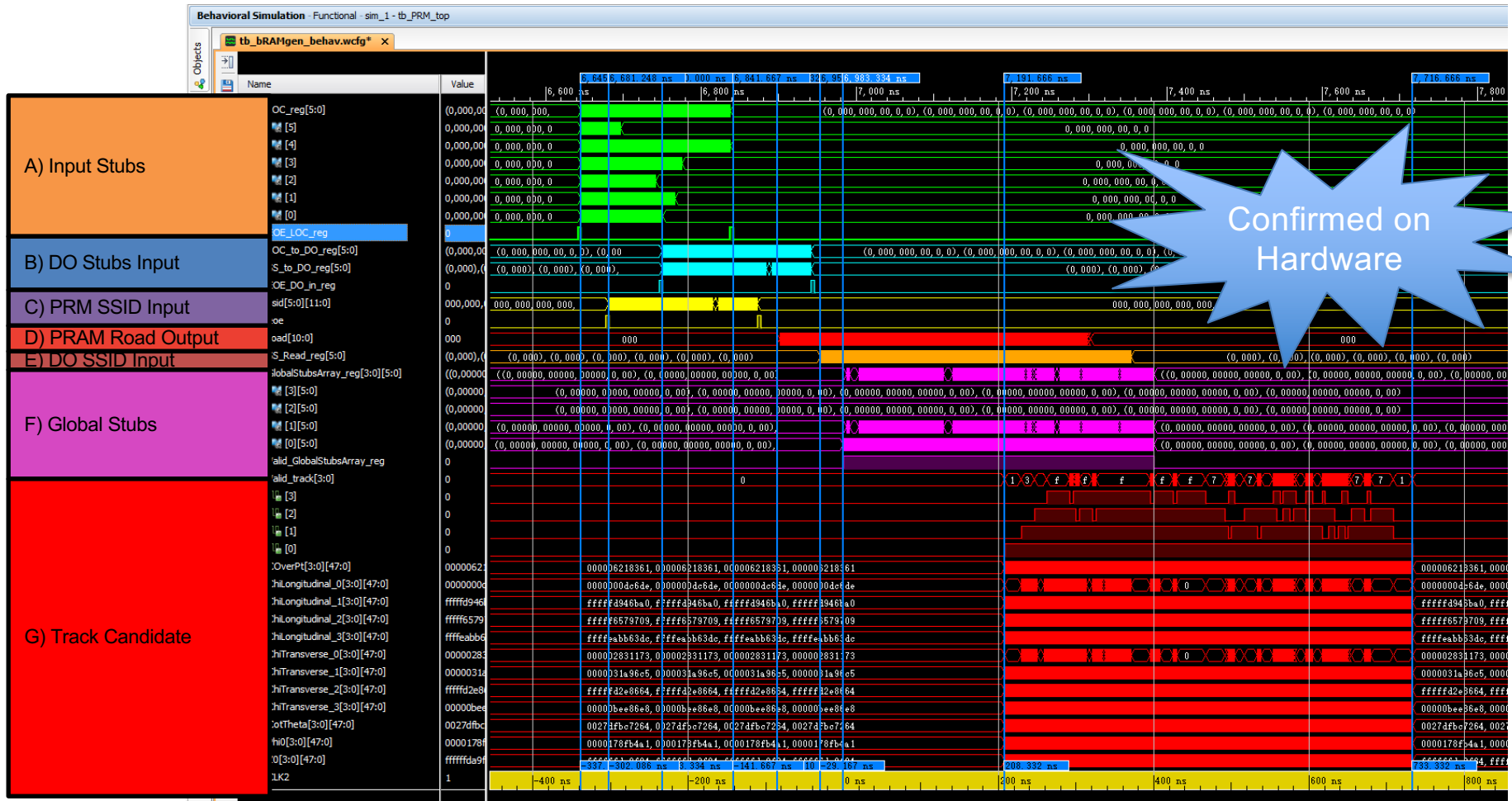 - Latency, FGPA resource usage, efficiency, resolution

Stub φ global coordinate diff (EMU - FW)

Stub r global coordinate diff (EMU - FW)

Latency well within specs

Data Sourcing → Data Formatting /Delivery → Pattern Recognition → Track Fitting

950ns　　250ns　　642ns　　25ns　　175ns

Data delivered to PRM starts/ends: @1.20 – 1.70 μs
Pattern Recognition output starts/ends: @1.84 – 2.34 μs
Track Fitting output starts/ends: @2.04 – 3.04 μs

3042 ns

DD | 950 ns | 250 ns | 500 ns

PR | 642 ns | 500 ns

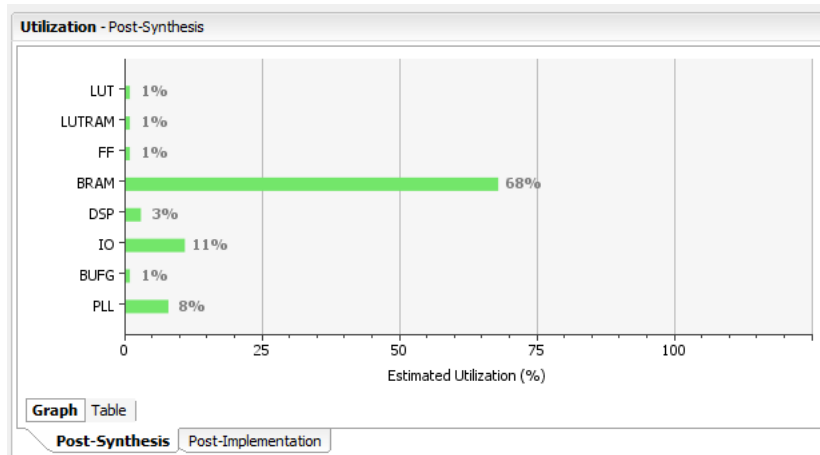TF | 175 ns | 1000 ns

2542 ns (PU250)

An example event in Vivado

4xTF, 7191.666 ns to 7716.666ns = 525ns, 126 clk @240MHz
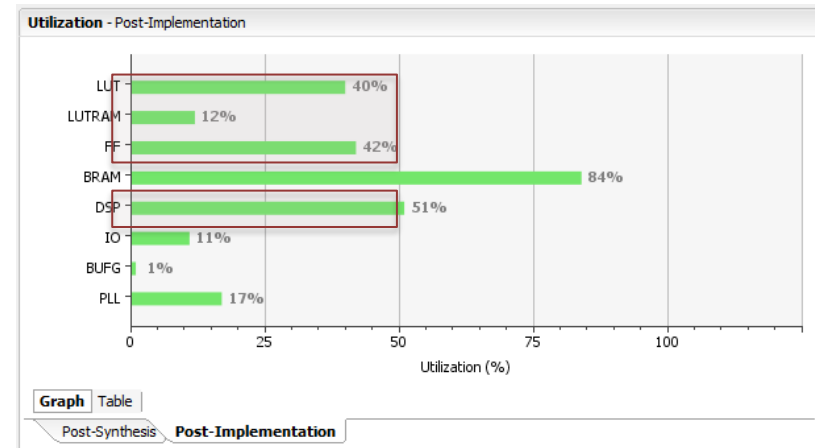This is with a selected ttbar + PU200 event with high tail of combinations

Data Organizer only

Data Organizer and 8 Tracker Fitters

- Very light weighted design

- BlockRAM mainly used for DO
  - TF does not increase BlockRAM usage, leaving enough room for TF

- TF: modest increase in registers and DSP blocks
  - Plenty of room for parallel copies of the fitter

- **System is robust against higher luminosity or increase in hit occupancy**
  - We demonstrate that the system reconstructs all tracks for events with PU250 within 2.5 μs (no truncation needed)
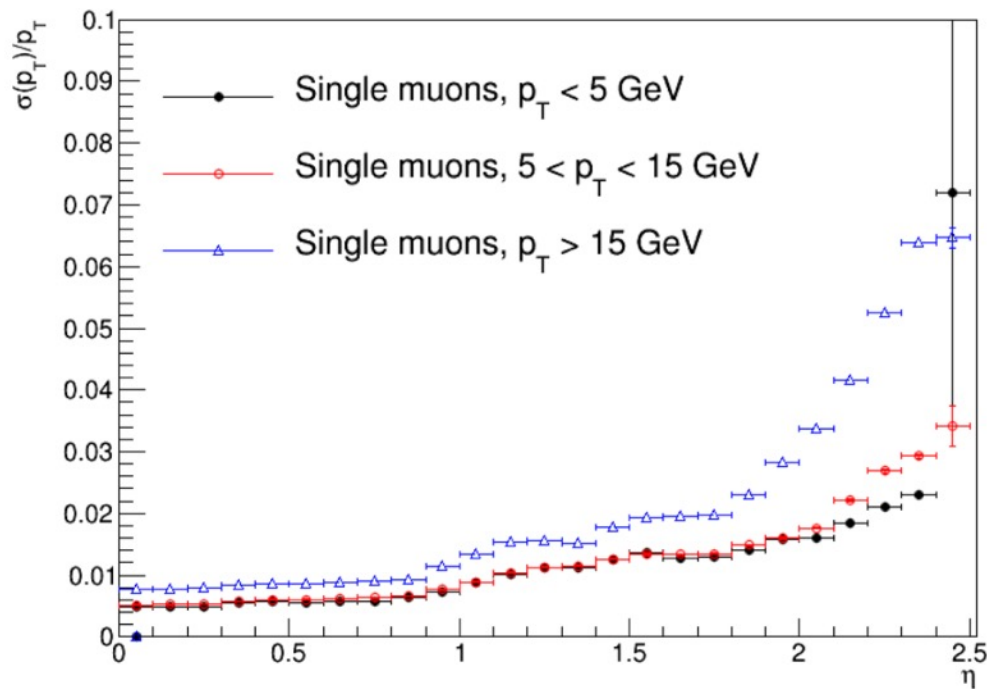 - Only for very high pT jet, truncation needed to meet the pipeline window
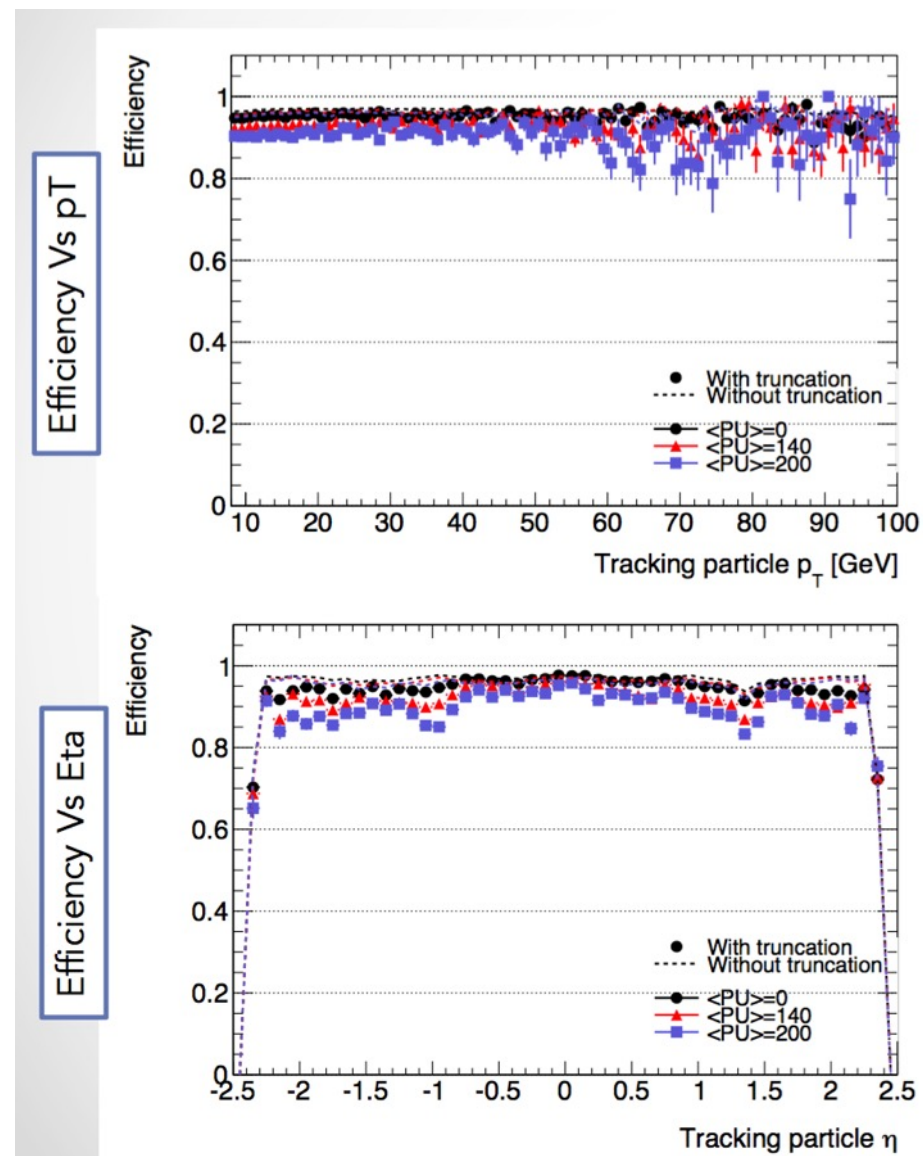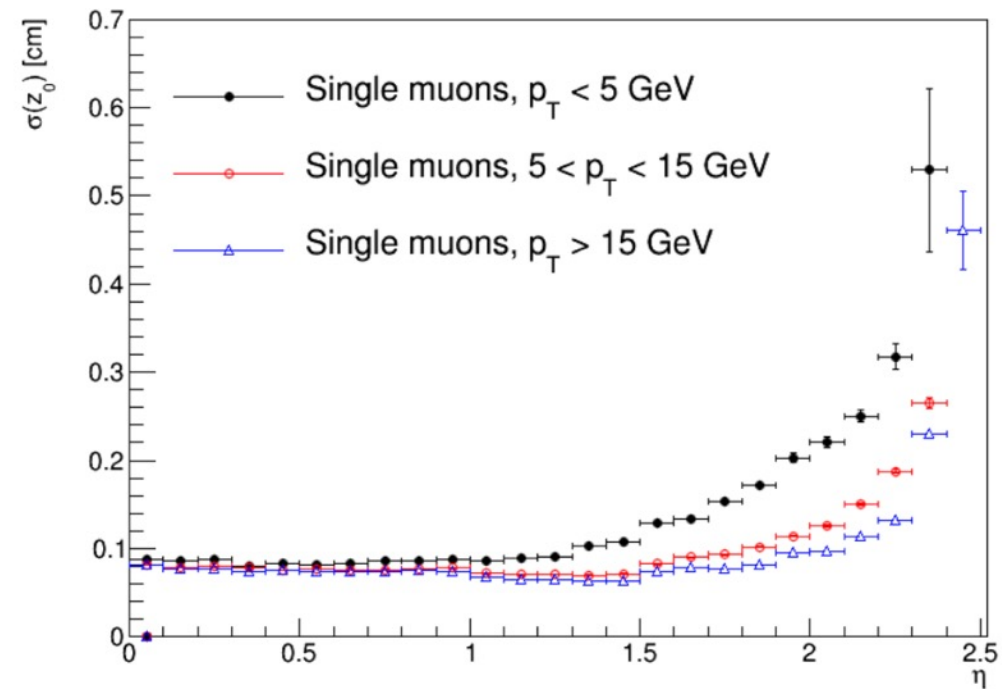
Tracking efficiency in ttbar+PU jets

- ## Excellent performance for L1 trigger application

p_T resolution

z_0 resolution

Summary

- ## Demonstrated with a vertical slice
  - Achieved excellent performance in terms of tracking efficiency and momentum resolution
  - Very low total latency (2.5 μs): data dispatch to the trigger towers, pattern recognition, track fitting

- ## The success of the demonstration system
  - An existence proof of fast data delivery, fast pattern recognition and track fitting implemented using the full mesh ATCA and associative memory approach
  - Technology advancements could lead to reduced size of the system in the future

1. "Charged Particle Tracking in Real-Time Using a Full-Mesh Data Delivery Architecture and Associative Memory Techniques", *JINST* 17, P12002 (2022)
2. "A High-performance Track Fitter for Use in Ultra-fast Electronics", *Nucl. Instrum. Methods Phys. Res. A* 935, 95-102 (2019)

Backup

Close collaboration:
FNAL, Northwestern, U. Florida, Texas A&M,
Brazil(SPRACE/UERJ) and China (Peking).
With FULL support of LPC

T. Liu, AM + FPGA Overview

Zhen Hu

May 18, 2018