Massively Parallel Processing Pattern Recognition in Extra Dimension

Innovation to go far beyond Moore's Law
for future triggering at high luminosity LHC

Thomas Klonowski
Fermi National Accelerator Laboratory
Quarknet
August 3rd, 2012

Fermilab

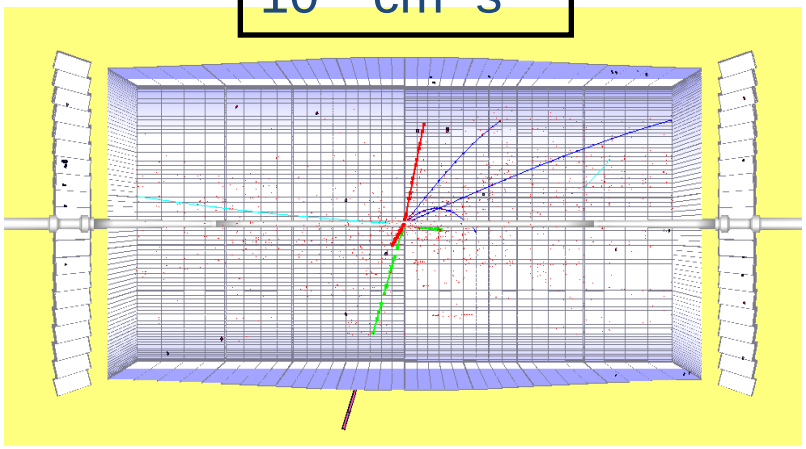
The What:

- A method for capturing and reconstructing particle tracks in real time
- Content Addressable Memory (CAM)
- Associative Memory for pattern recognition
- Very high speeds and pattern density
- 3D technology is the key

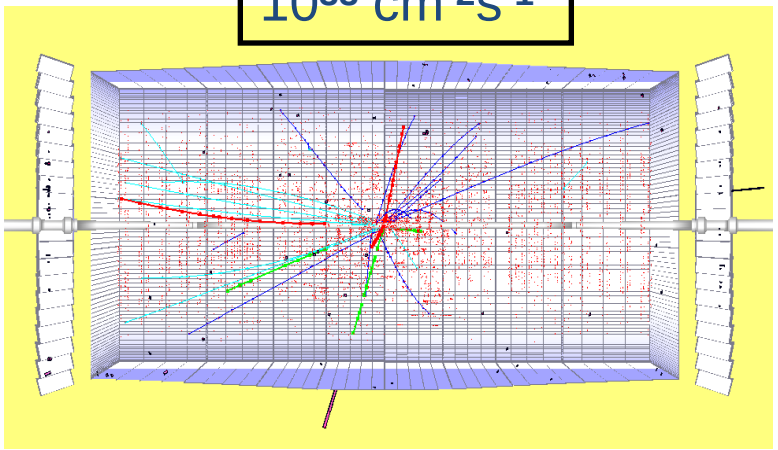
The Why:

The LHC will need much higher trigger performance in the future at higher luminosity

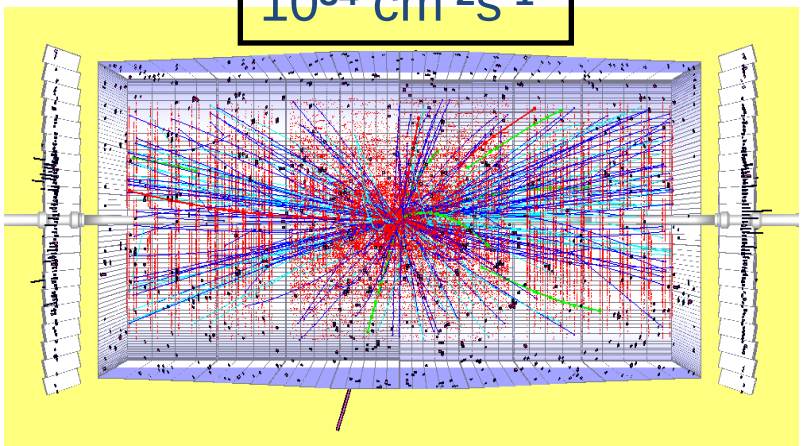
10^{32} cm^{-2}s^{-1}



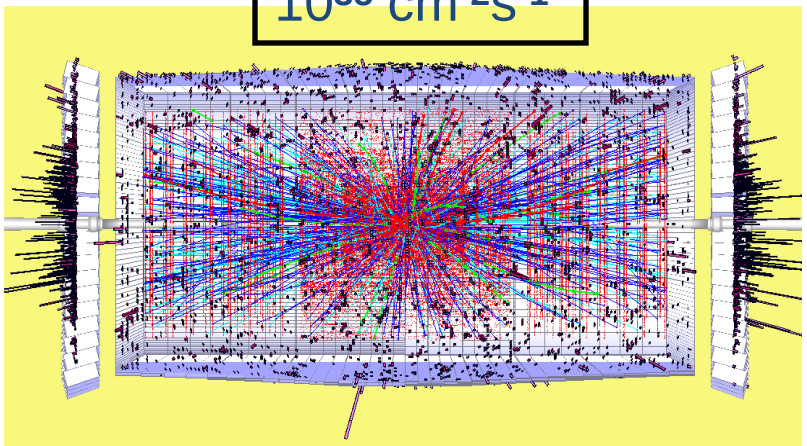
10^{33} cm^{-2}s^{-1}



10^{34} cm^{-2}s^{-1}

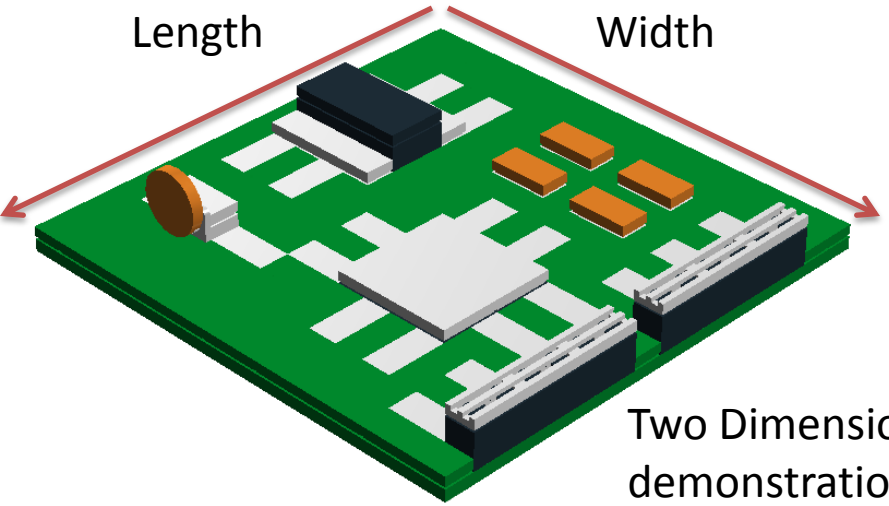


10^{35} cm^{-2}s^{-1}



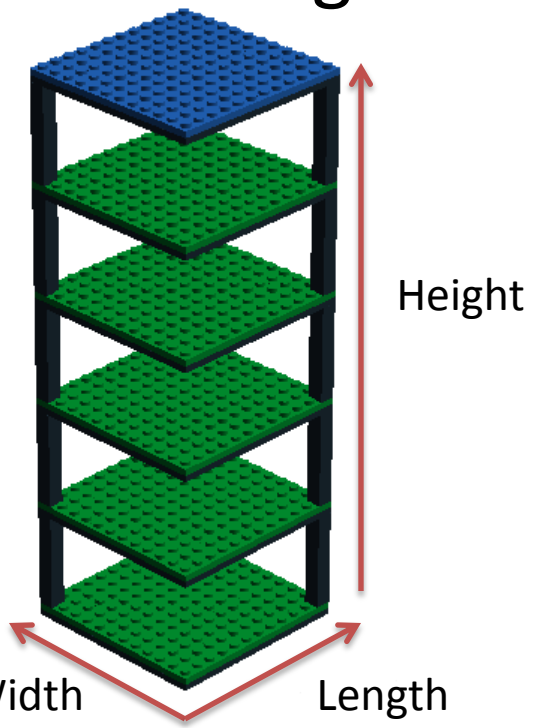
Extra Dimension?

- Traditional circuitry only built with length and width (two dimensional)
- New vertically-integrated design adds height to circuit (three dimensional)



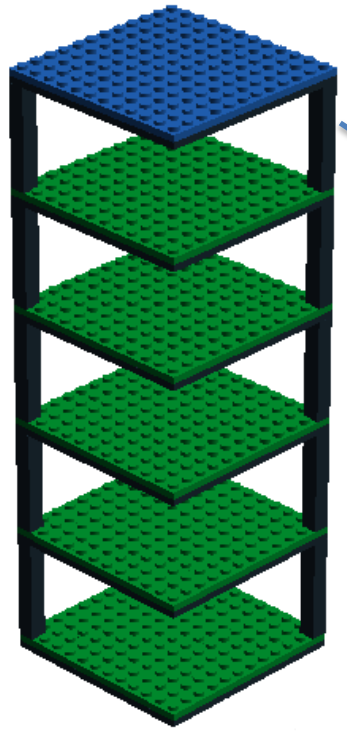
Two Dimensional Circuit demonstrational model

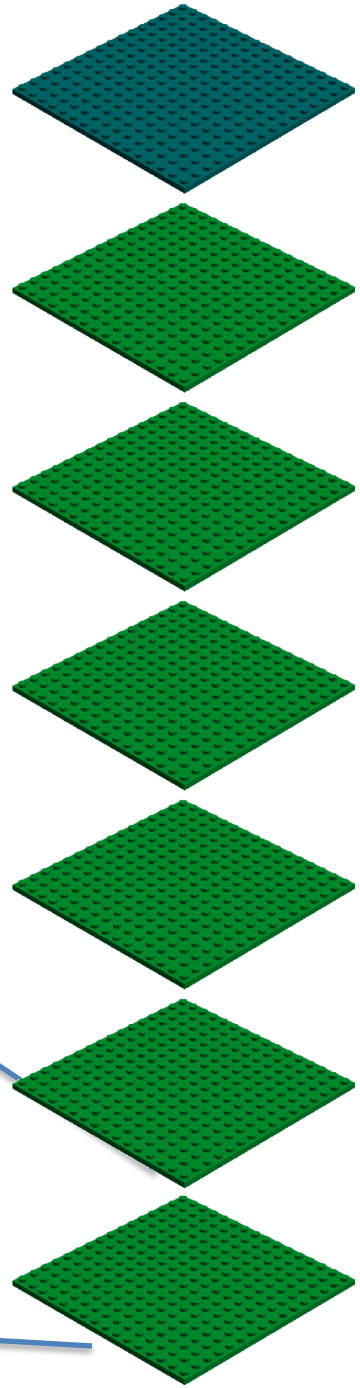
New vertically-integrated circuit concept model (expanded vertically)



Why Three Dimensional?

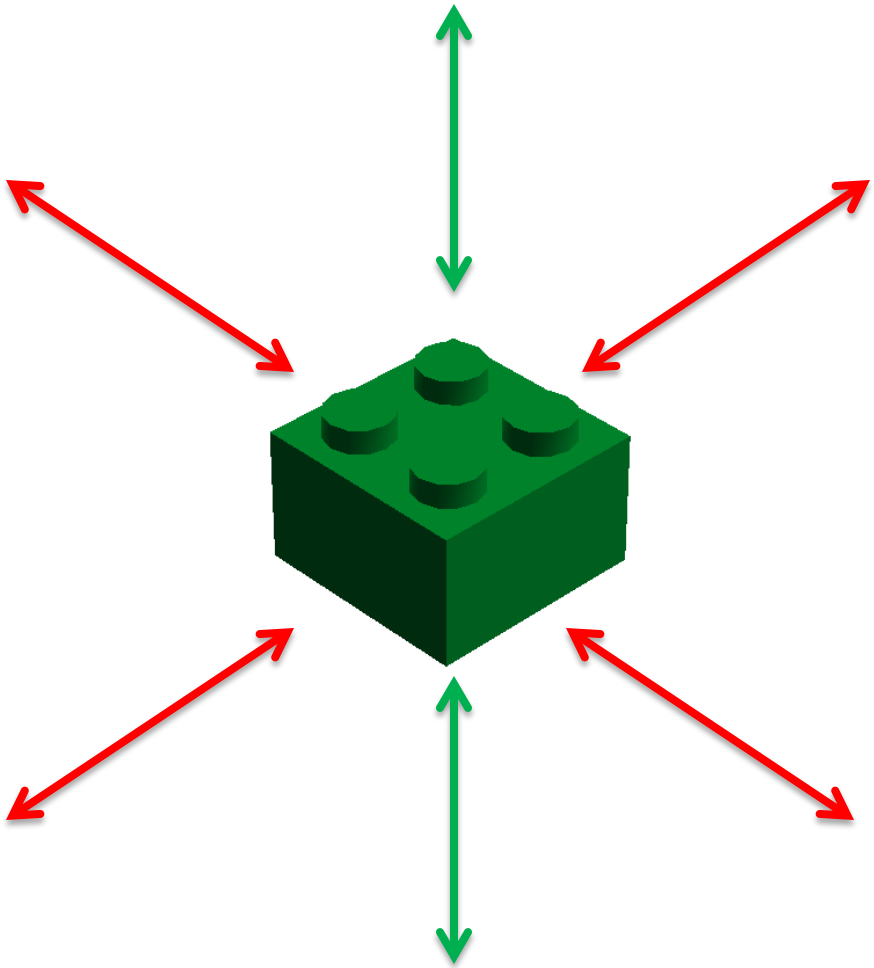
- Shorter interconnections
  - Lower power density
- Higher Transistor density



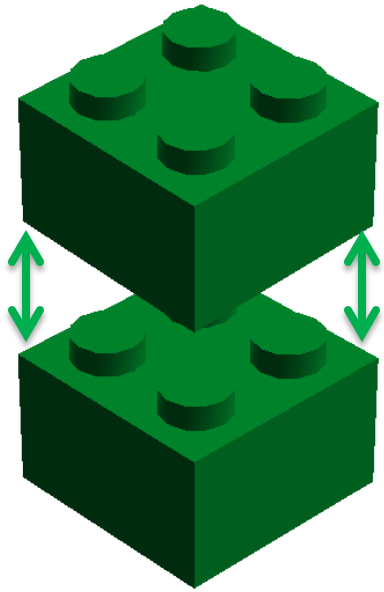


•No space in between tiers

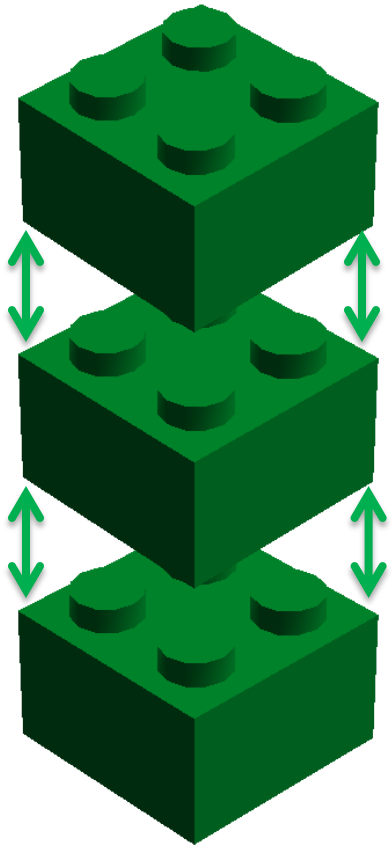
It’s Like This:



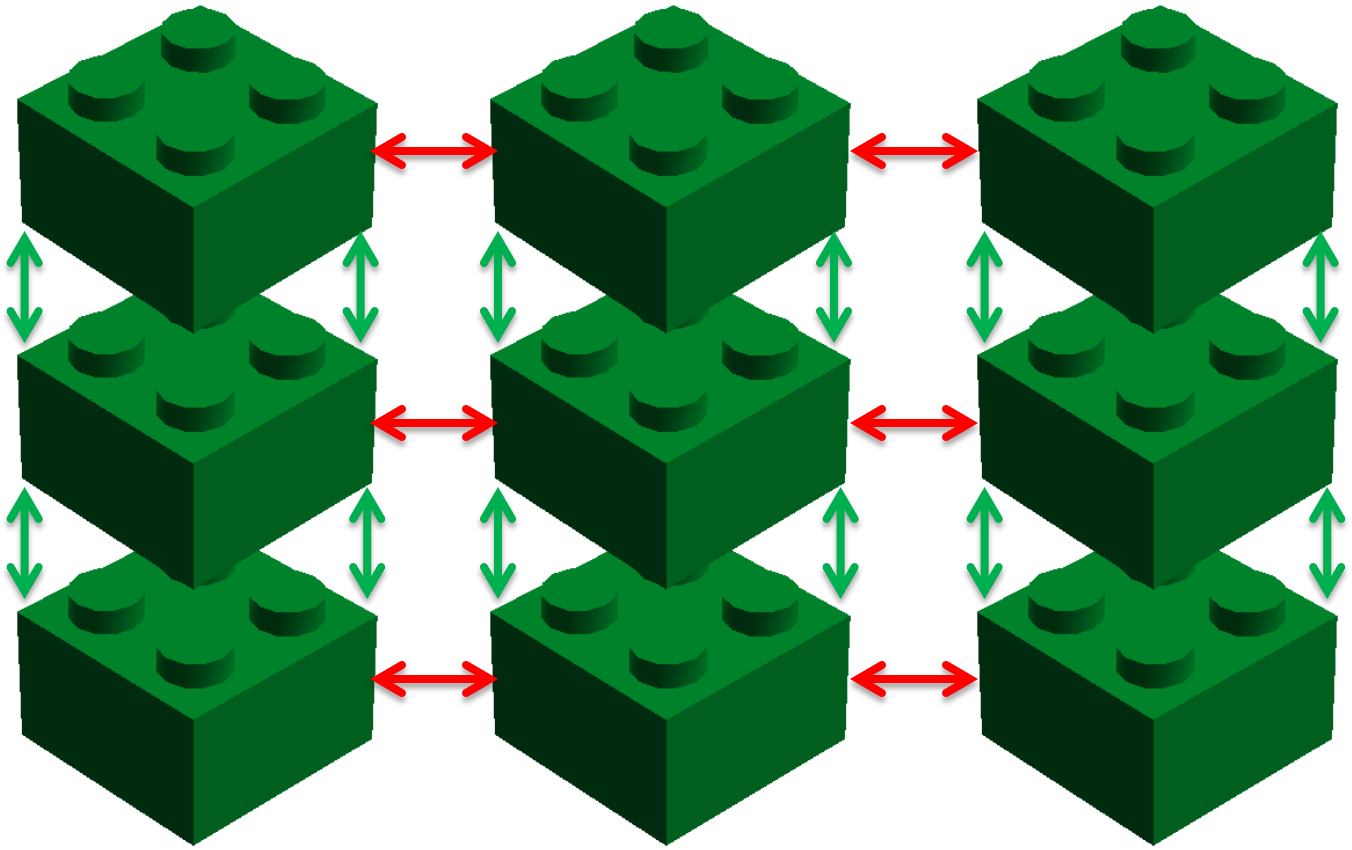
It’s Like This:



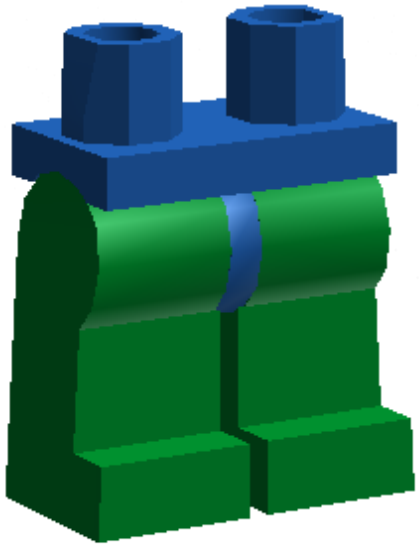
It’s Like This:



It’s Like This:



It’s Like This:



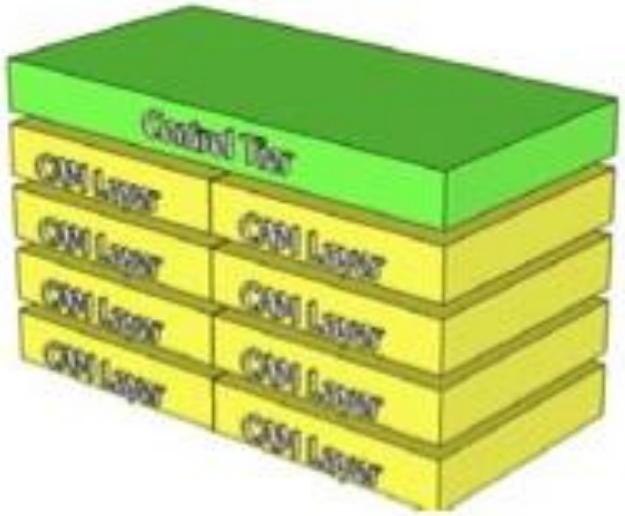
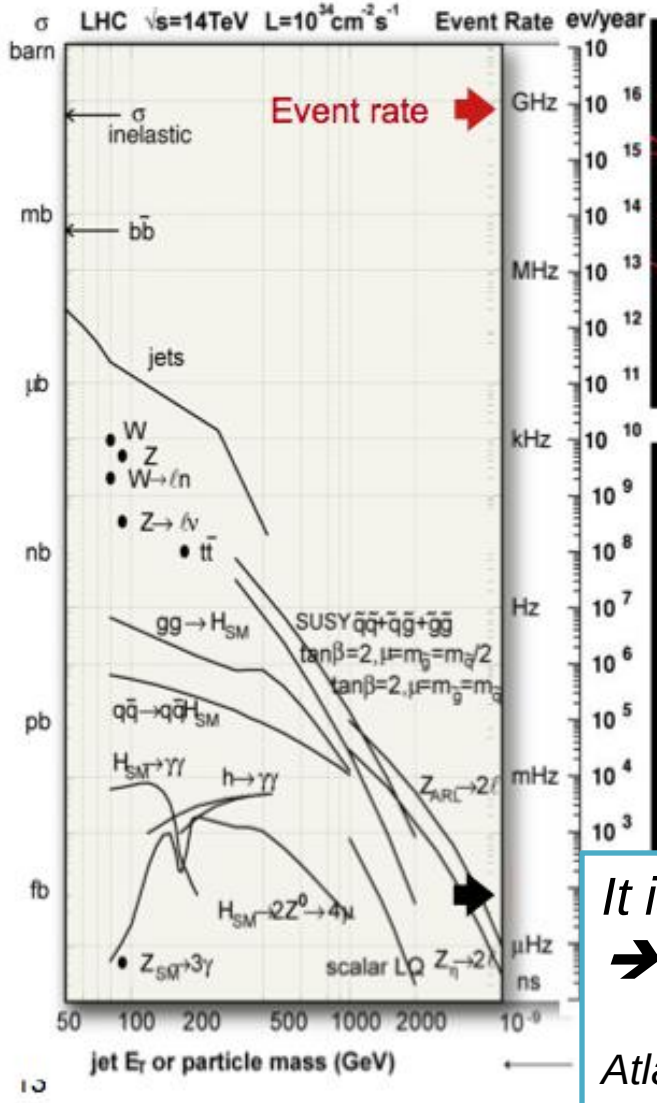
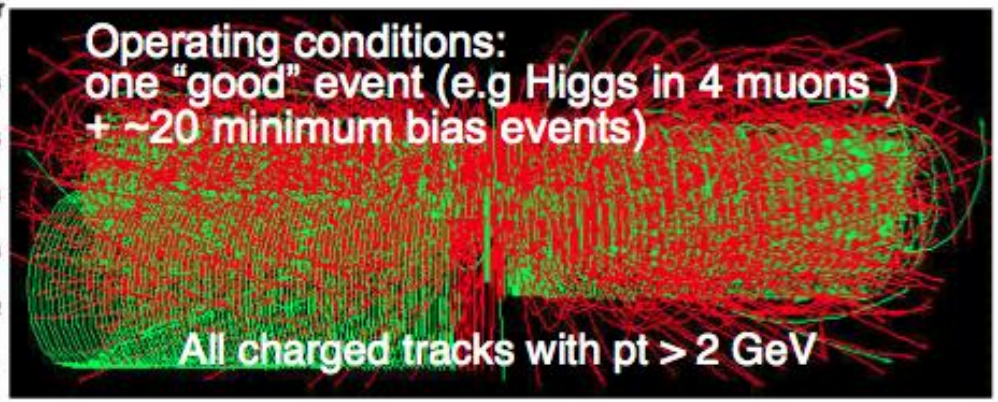
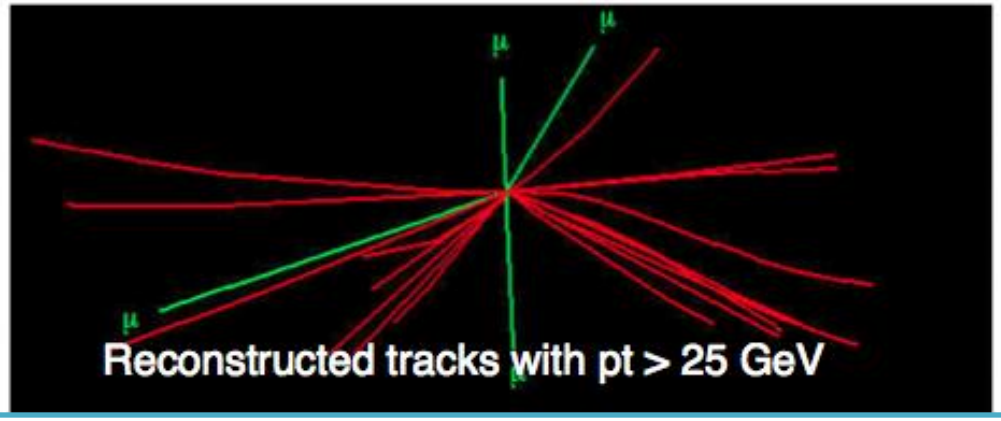


Figure 3 - A Block layout of the protoPants.

Collisions (p-p) at LHC







It is crucial to have tracking at early trigger stages
➔ enormous challenges !

Atlas FTK is the first step in this direction at LHC, and uses all silicon tracker info at L2 (see FTK talk by Guido Volpi)

From "Triggers in HEP" by Ted Liu

Content Addressable Memory

- A special type of memory used in certain very high speed searching applications
- User supplies data word
- Searches entire memory to see if that data is stored anywhere in a single operation
- Returns list of addresses where word was found
- Hardware, not software

From Wikipedia:
en.wikipedia.org/wiki/content-addressable-memory

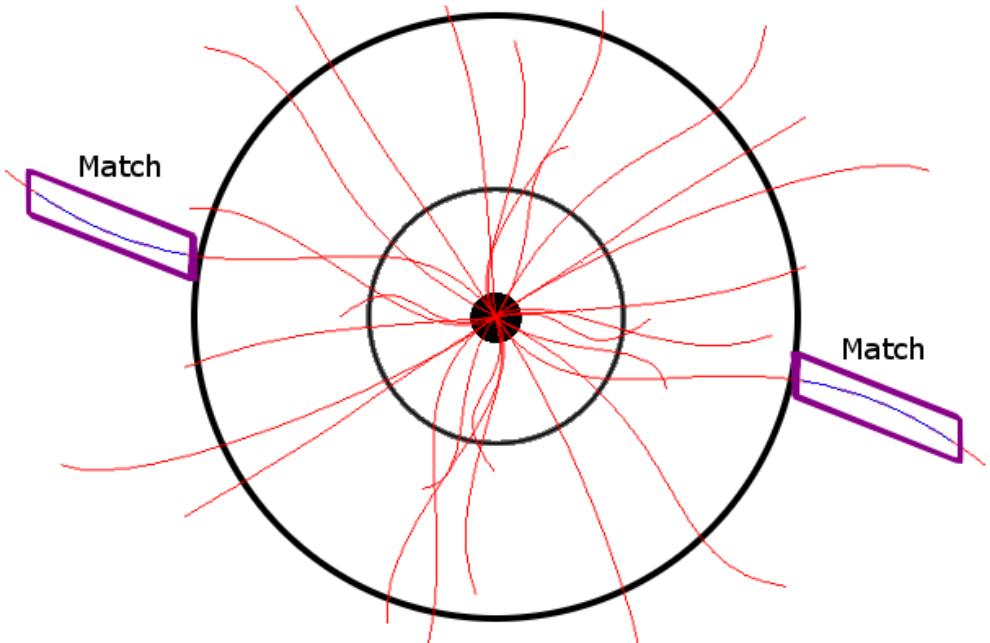
Content-Addressable (CAM) vs. Random Access (RAM)

- Searches entire memory in one operation
- User supplies data word
- Returns list of storage addresses
- Extremely fast

- Searches one address at a time
- User supplies address
- Returns data word at supplied address

Associative Memory

- Using CAM to match hits
- Majority logic to associate the matches of different layers into track patterns
- Massive parallel processing



Comments on Associative Memory

- Based on *CAM cells to match and majority logic to associate* hits in different detector layers to a set of pre-determined hit patterns
  - Performance fundamentally limited by Moore's Law

- *This is the main limitation of an otherwise very powerful and proven approach for its future applications within and beyond HEP.*

From "Triggers in HEP" by Ted Liu

The Challenge of future AM design

Increase the patterns density by 2 orders of magnitude;

and

increase the speed by a factor of >~ 3,

while

keeping the power consumption more or less the same

Much higher Patten Density & higher Speed

Yet much less Power Density

almost too good to be true

New idea: could go to “extra dimension” to achieve this

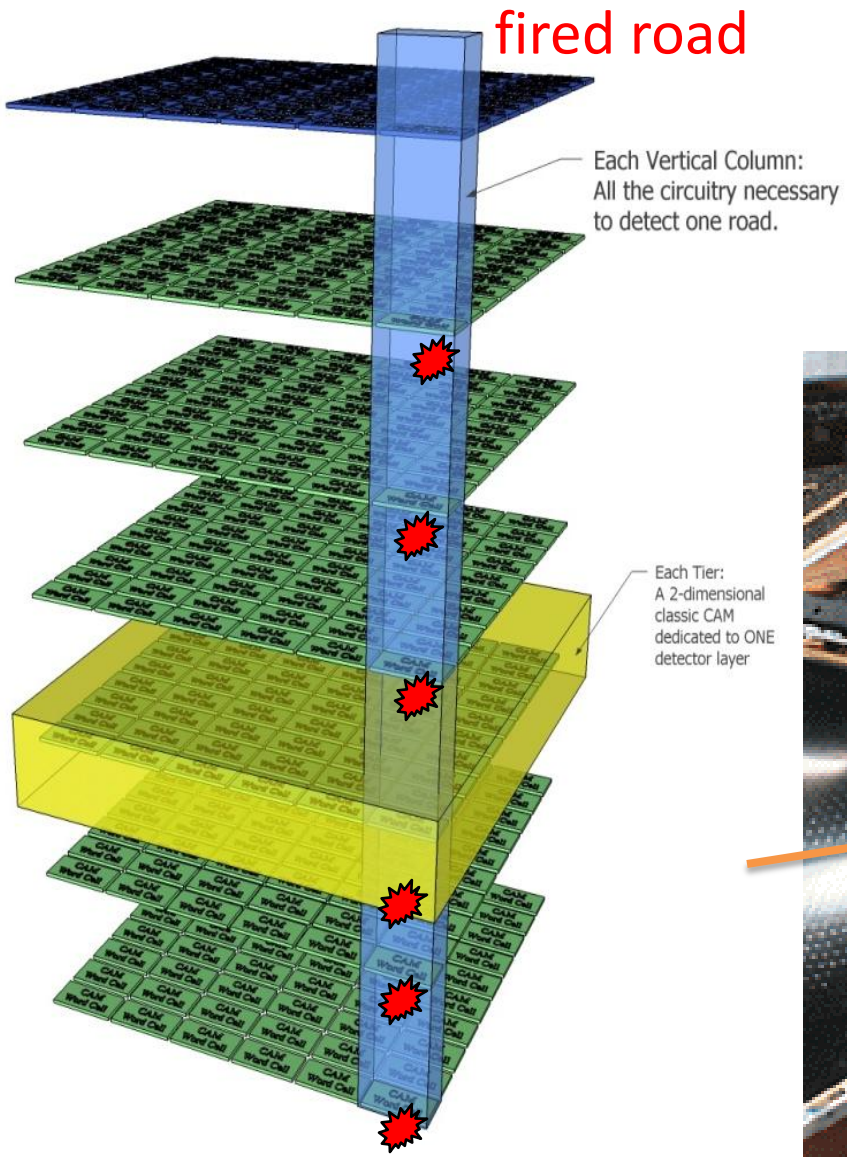
➔ generic R&D effort at Fermilab

From “Triggers in HEP” by Ted Liu

VIPRAM

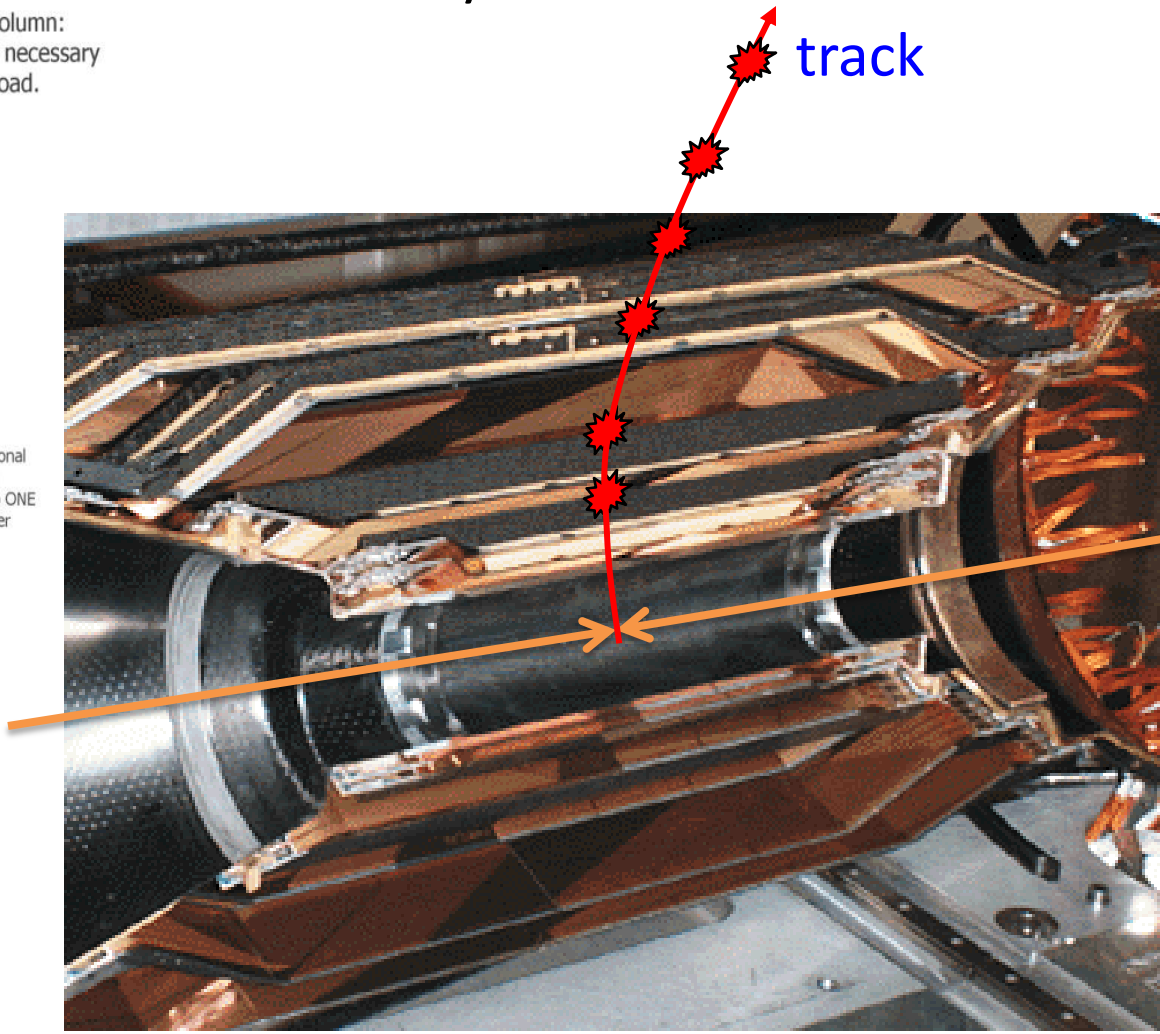
(Vertically Integrated Pattern Recognition Associative Memory)

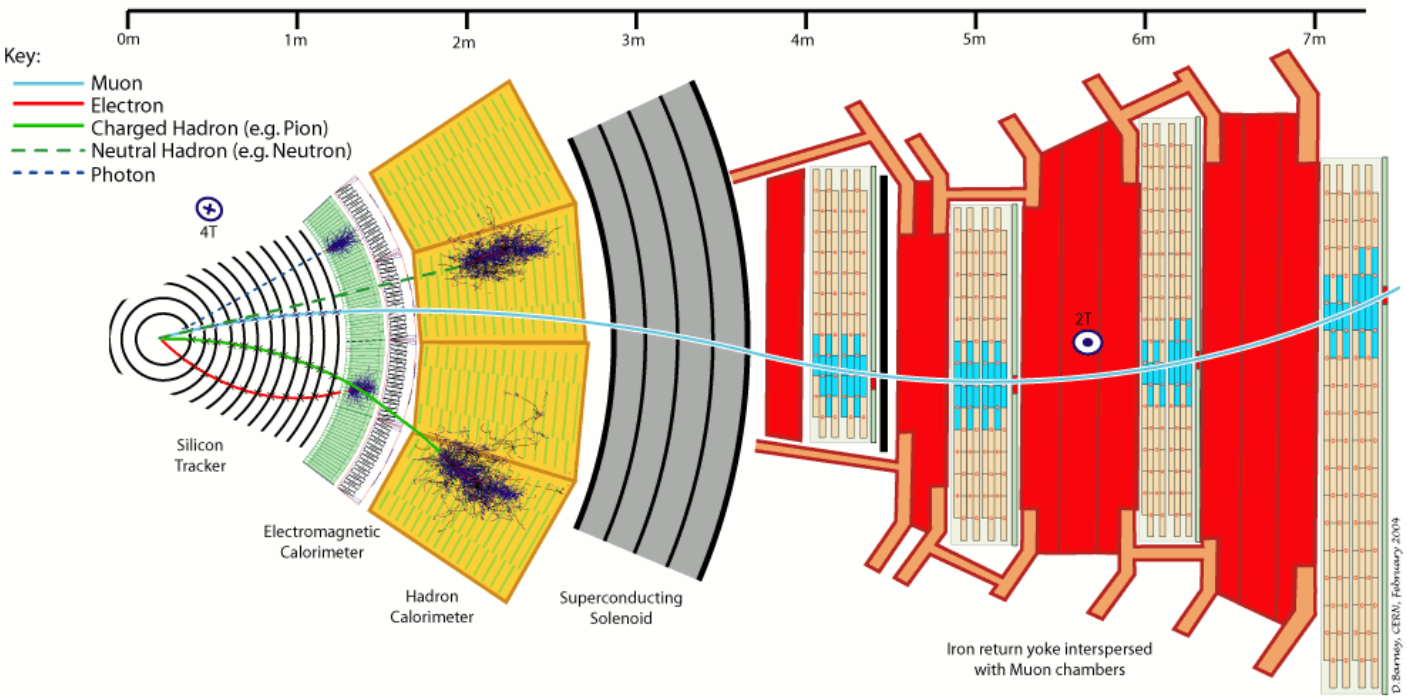
http://hep.uchicago.edu/~thliu/projects/VIPRAM/TIPP2011_VIPRAM_Paper.V11.preprint.pdf

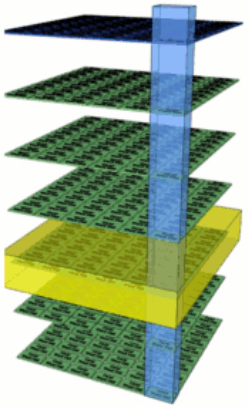


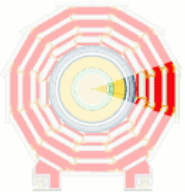
From “Triggers in HEP” by Ted Liu

Pattern recognition for tracking is naturally a task in 3D



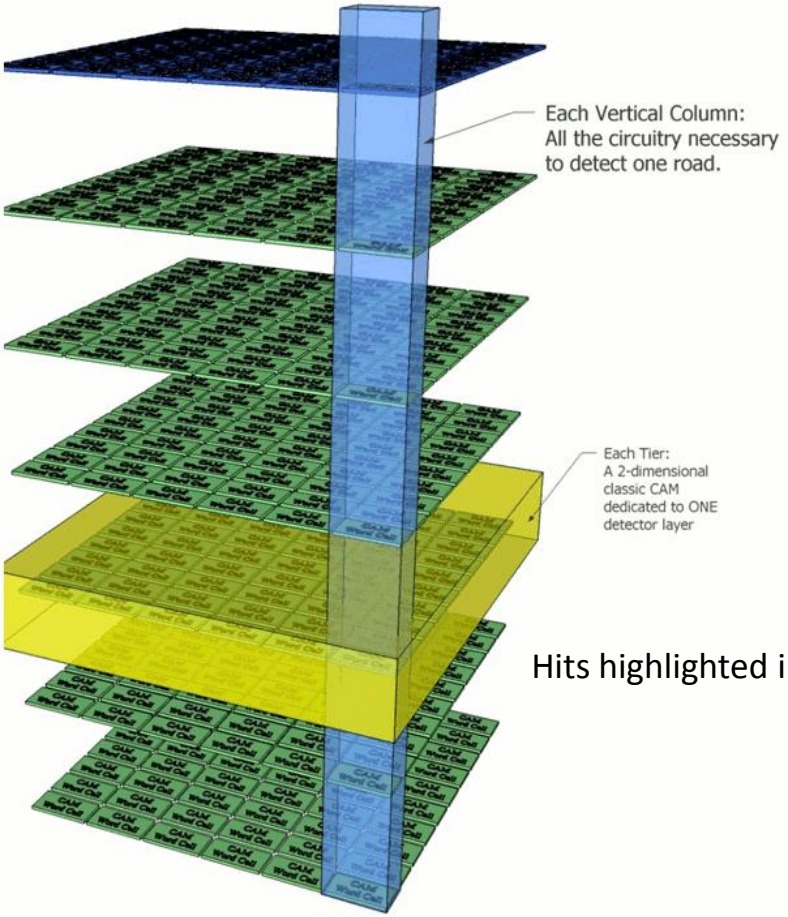






Transverse slice through CMS

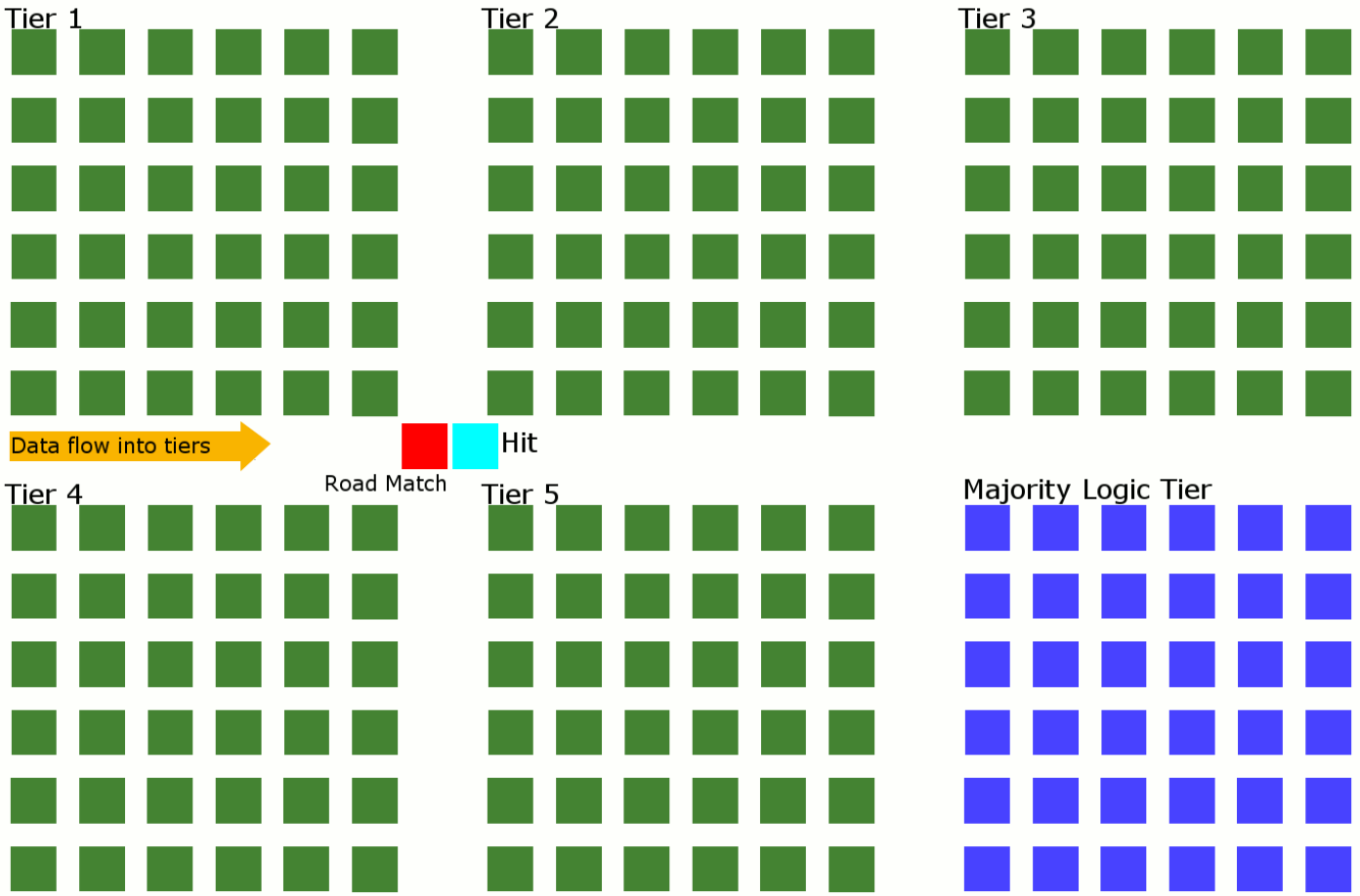
Inside the Chip



Hits highlighted in red

- Multiple tiers for pattern recognition for multiple detector layers
- Each tier 10 µm thick
- Hundreds of thousands of CAM cells per tier
  - Hundreds of thousands of patterns per chip
- Majority logic tier sends signal when match criteria is met

Inside the Chip



Internal activity on a 5-tier example chip

Improving Robustness

- 3D VIPRAM Architecture intrinsically fault-tolerant
  - Each tube/pattern is independent
  - New architecture allows simple implementation of redundancy for critical signals
- Eventually self-repairing or monitor circuit included

In Summary

- The LHC will need much higher trigger performance in the future at higher luminosity
- The current technology cannot be scaled in a simple manner to accommodate the demand
- Significant improvements or breakthroughs will be needed
- 3D technology is a promising way to go
  - VIPRAM is a good example

The Learning Experience

- Real-world experience
- Fantastic opportunity for college and post-college jobs
- Opportunity to work with scientists on a completely new technology

If You Have Reached This Slide

- The presentation is over

Backup

1. Side view of collision
2. Cutaway of collision
3. 3D model designs- expanded chip
4. Collider model design