Part II

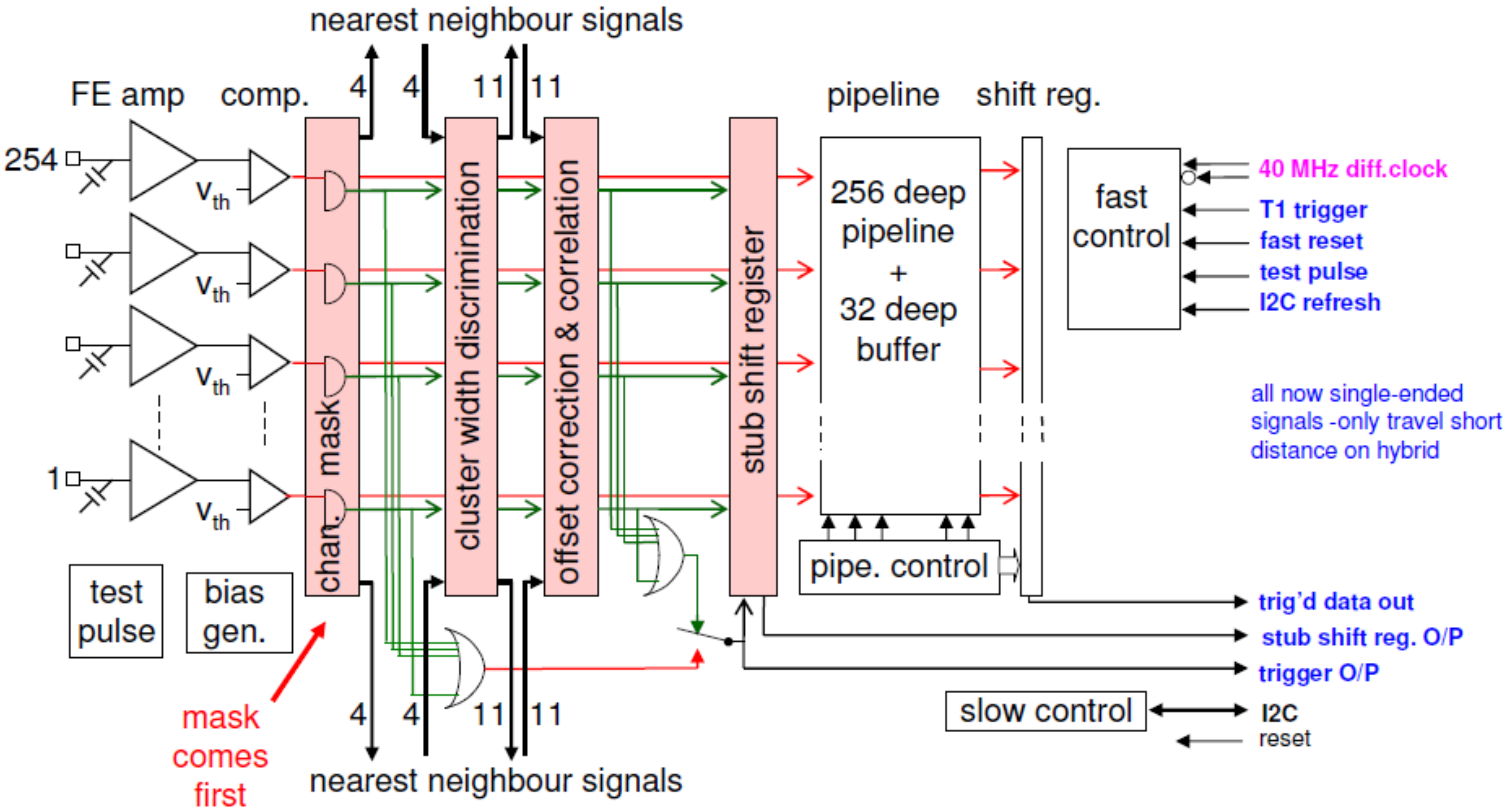
Intelligent Trackers for Triggers

Fabrizio Palla
INFN Pisa

III Seminario Nazionale Rivelatori Innovativi
Florence, 4-8 June 2012

Correlation electronics

CBC2 Architecture



blocks associated with Pt stub generation

- **channel mask**: block noisy channels (but not from pipeline)
- **cluster width discrimination**: exclude wide clusters
- **offset correction and correlation**: correct for phi offset across module and correlate between layers
- **stub shift register**: test feature - shift out result of correlation operation at 40 MHz
- **fast OR at comp. O/P and correlation O/P**: - can select either to transmit off-chip
  for normal operation choose correlation O/P

8

3

neighbour chip signals - CWD O/Ps

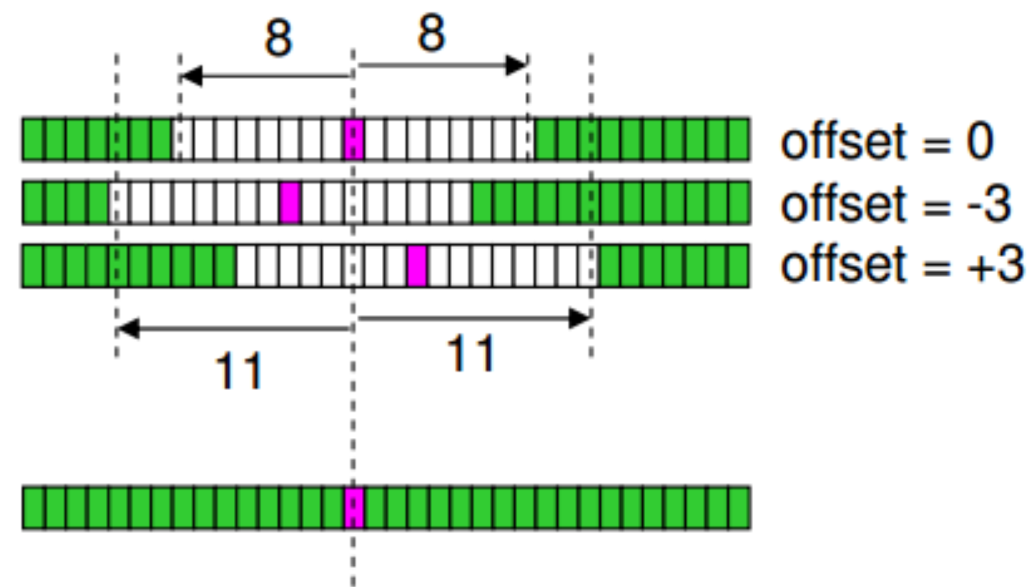
need programmability of **offset** and **window** width for upper layer channels to correlate with hit in inner layer

window defines Pt cut
    width programmable up to **+/- 8** channels

offset defines lateral displacement of window across chip
    programmable up to **+/- 3** channels

=> 11 signals to transmit to neighbouring chip
    11 to receive from neighbouring chip

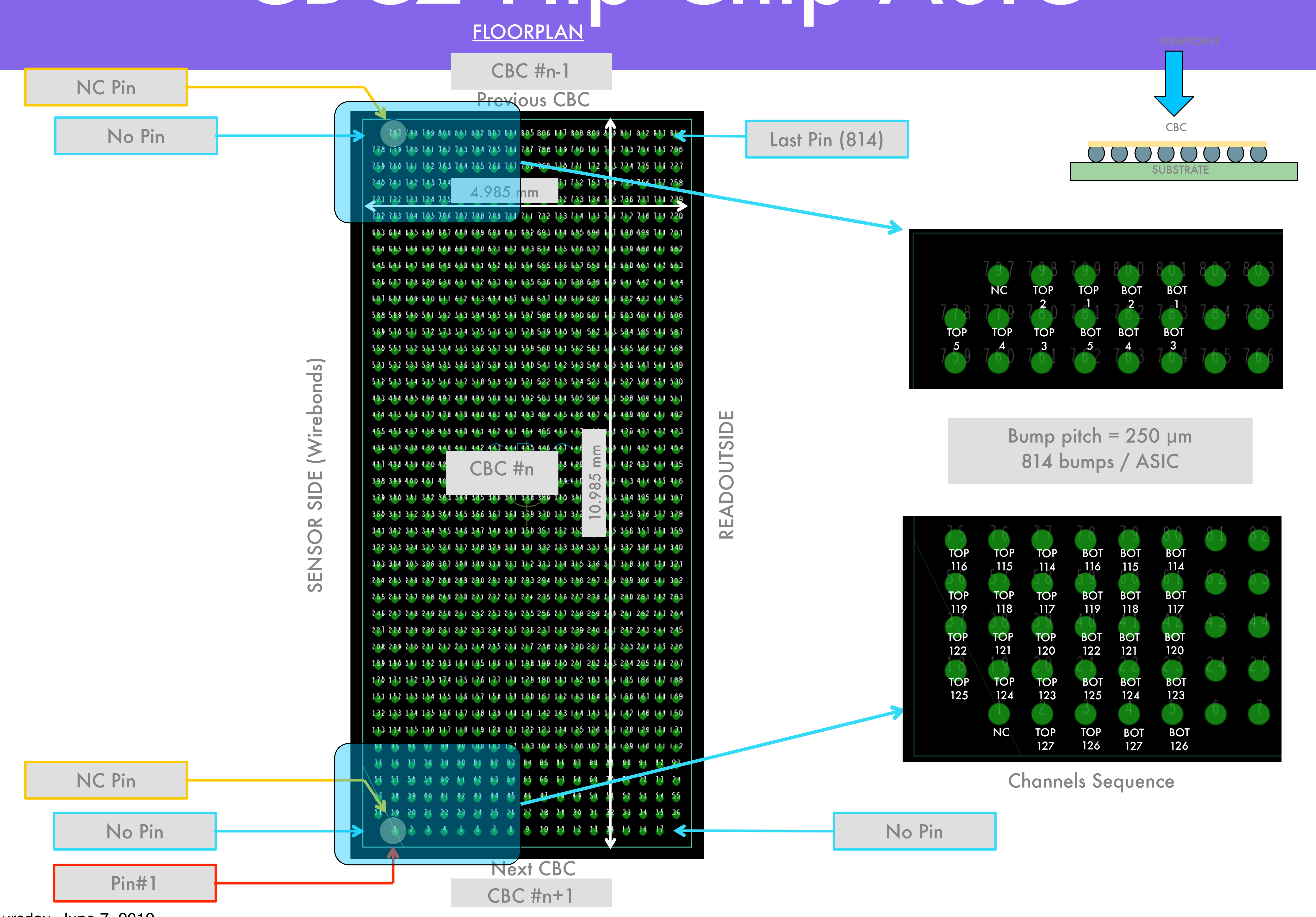
  = 22 signals



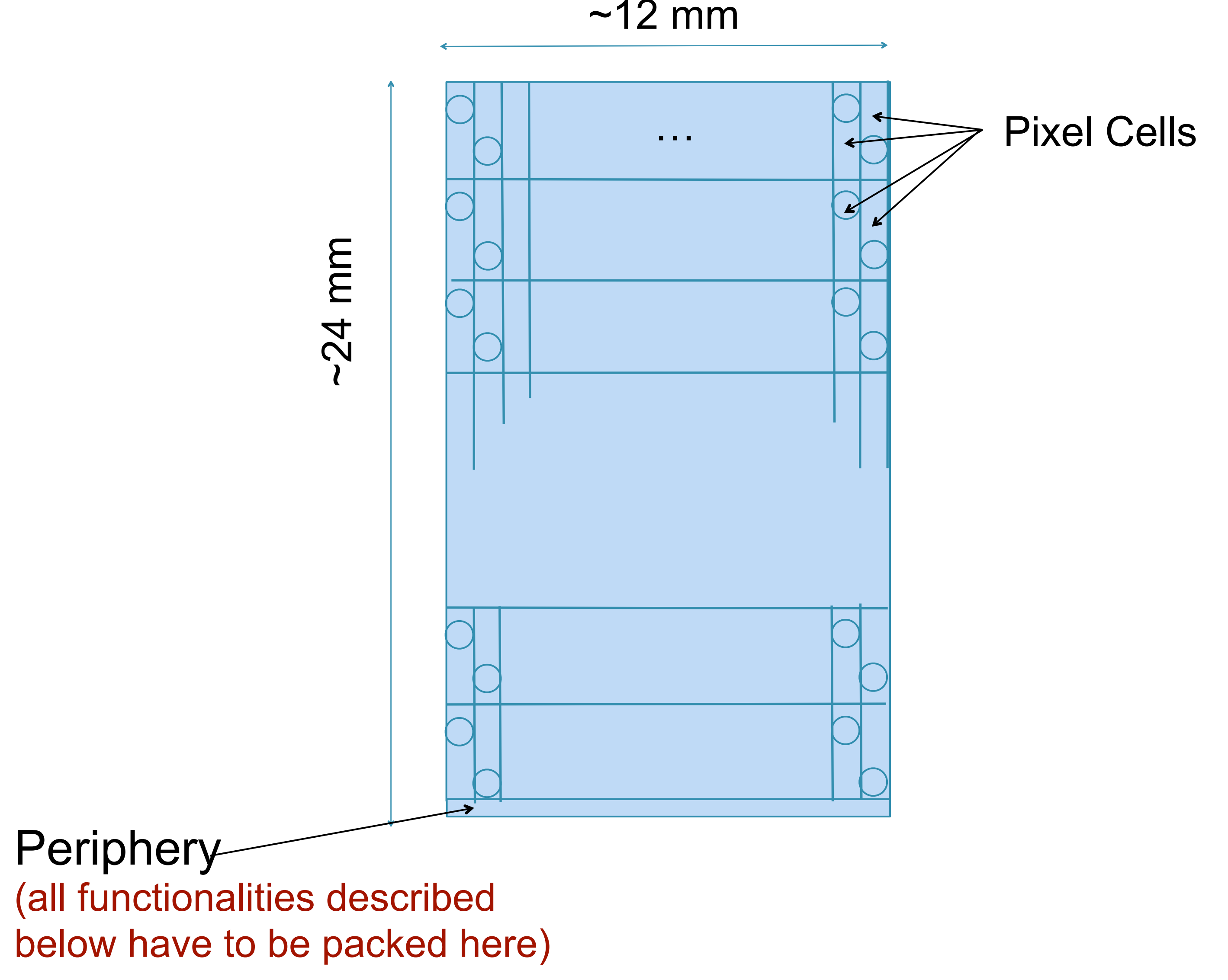
zero offset
offset
window width
outer layer CWD O/Ps
inner layer CWD O/Ps
r
ϕ
beam

adding comp O/Ps -> 30 signals altogether, top and bottom of chip

CBC2 Flip Chip ASIC

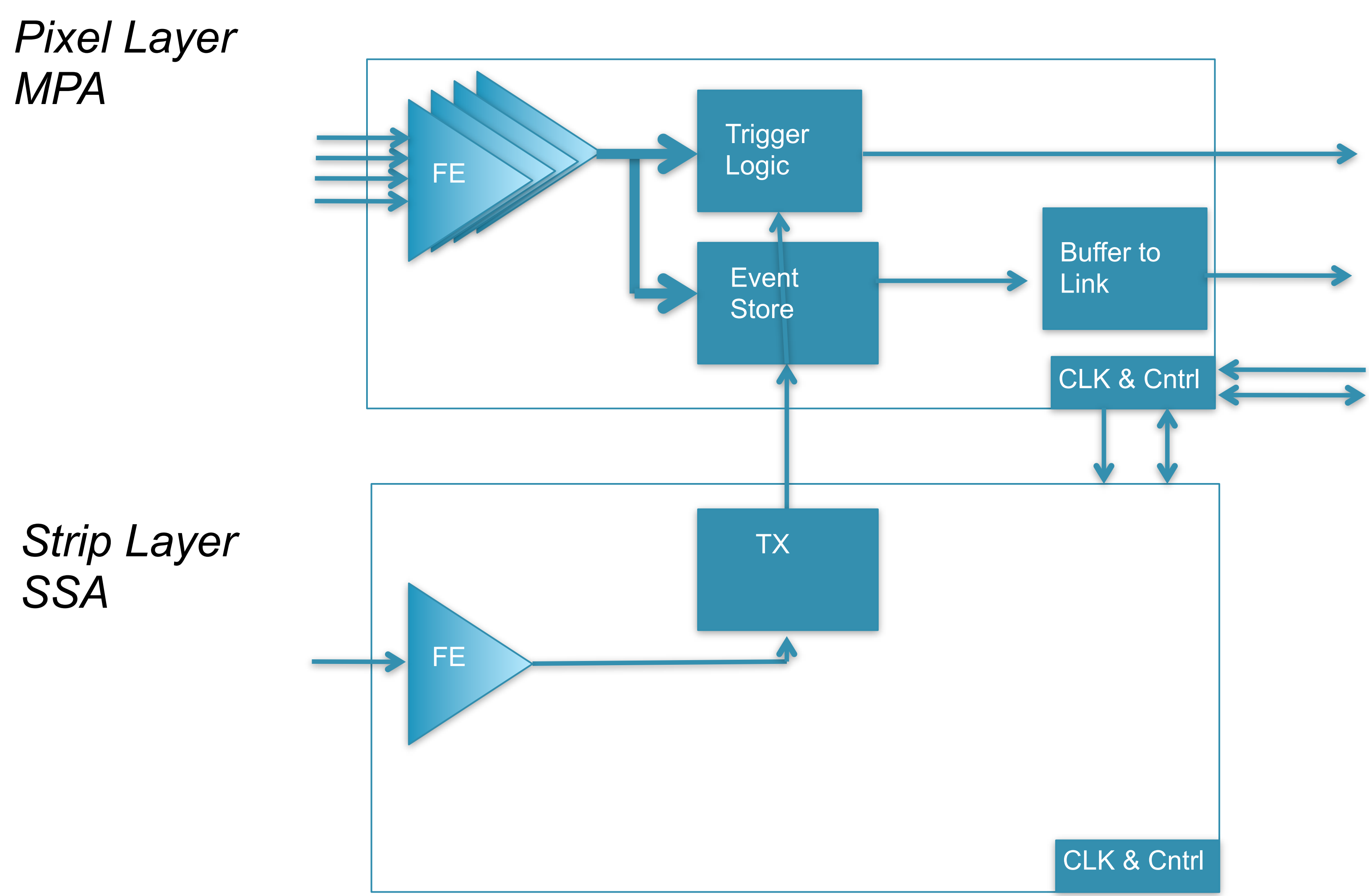


Macro pixel ASIC floorplan

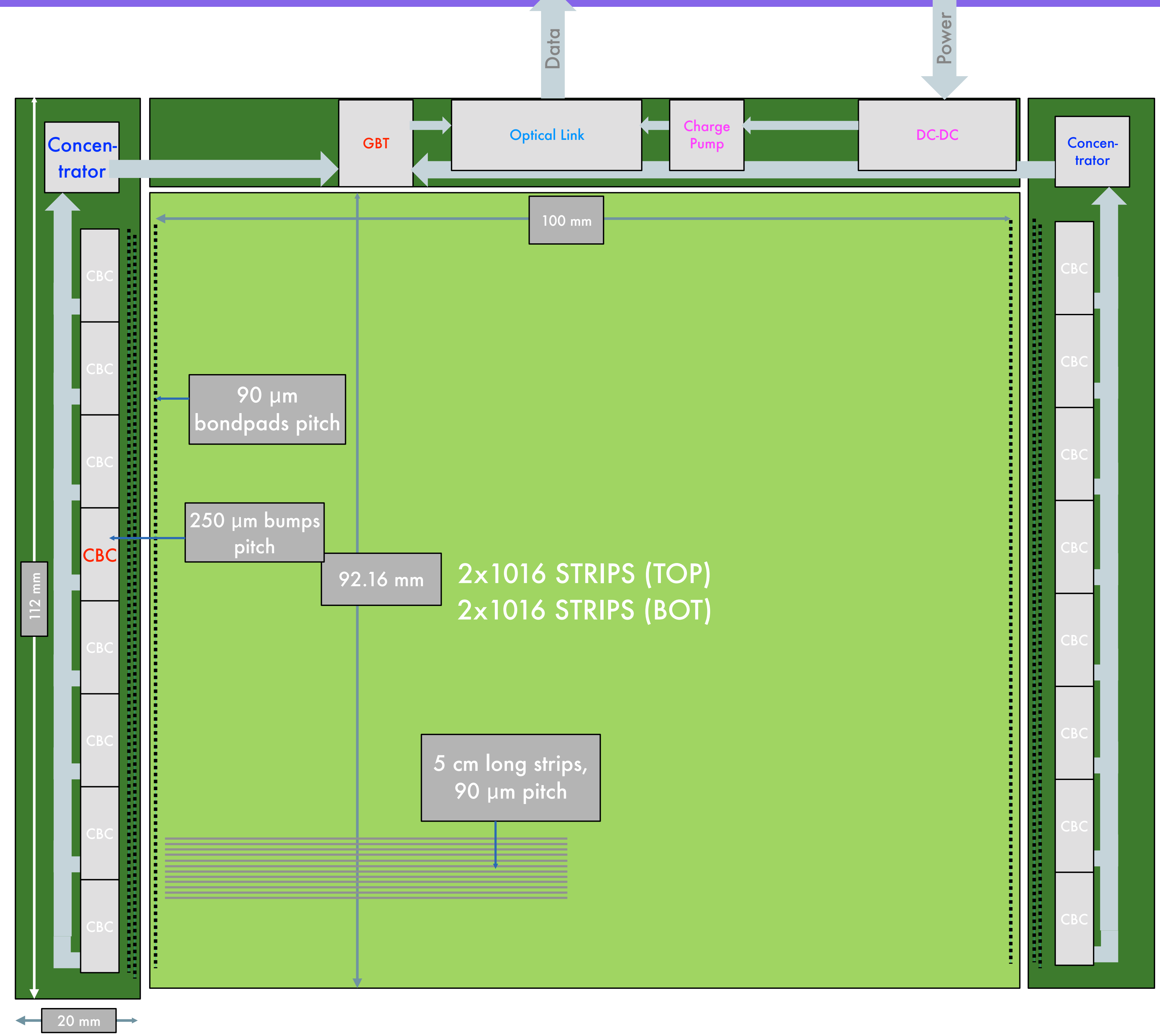


- MPixel ASIC size:
  - ~ 12 x 24 mm2
  - Pixel size: ~ 100 μm x 1500 μm
  - # pixels: 128 x 16 = 2048
  - Readout on one edge only

Coincidence and data handling in the pixel ASIC



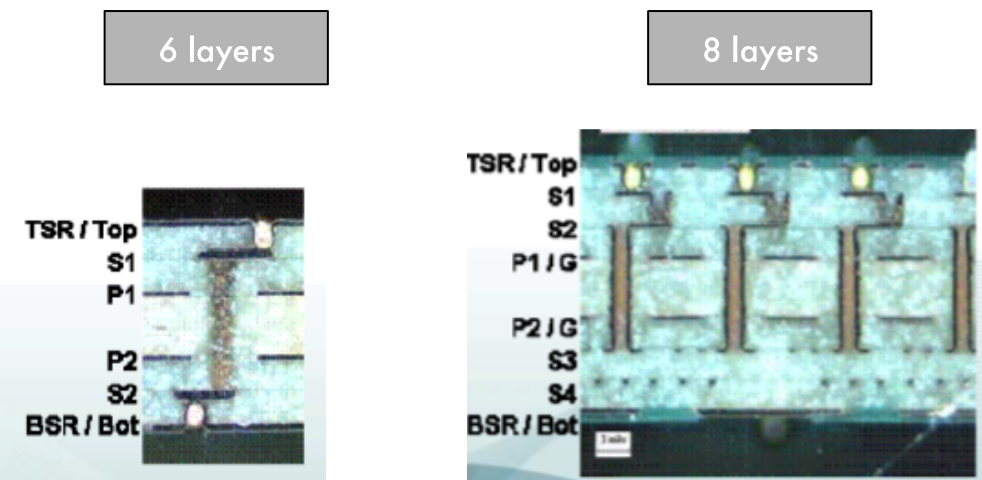
2S-pT Module: Hybrid Topology

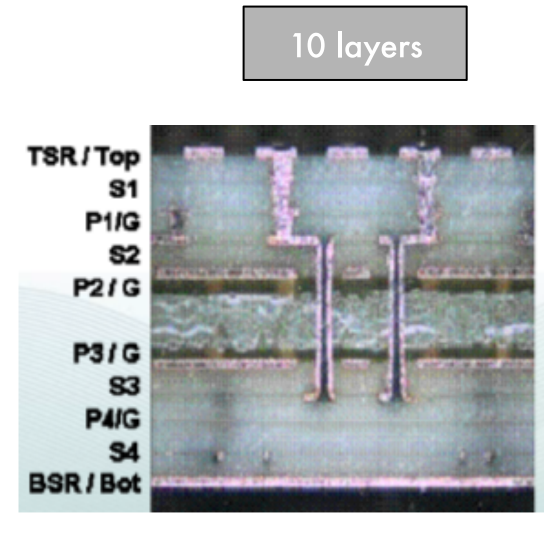


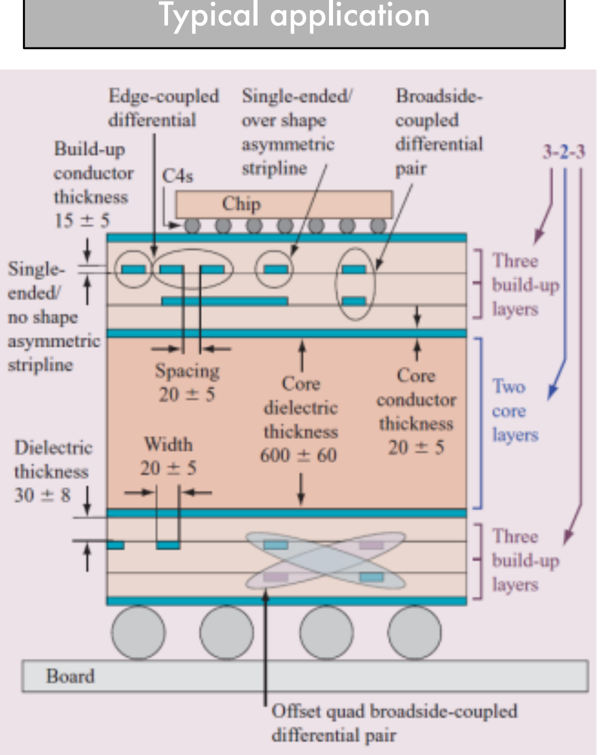
Rigid substrates

- **Build-up substrates are commonly used for chip packaging.**

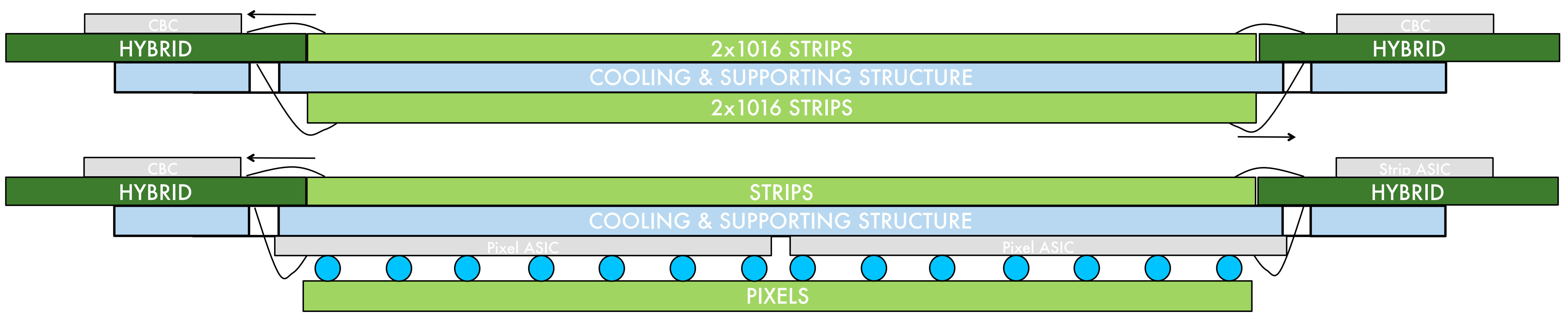
- Core layer provides:
  - Power/Ground planes
  - Rigid core material.
  - Mid density routing and through hole vias.
- Build-up layers are laminated on top and bottom of core:
  - Very high density interconnections on constrained areas.
  - Microvias to connect build up layers to core external layers.
  - No through hole vias..

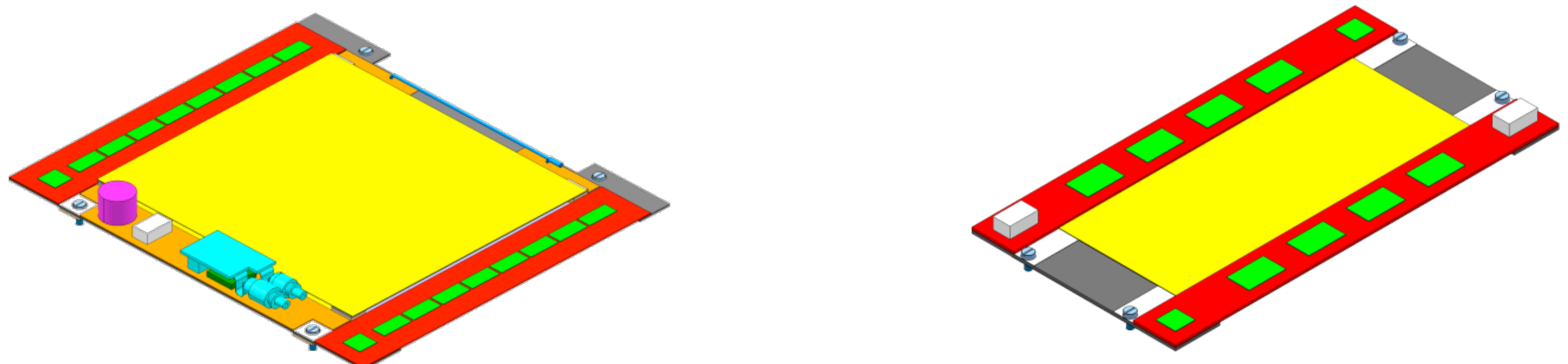






Build-up substrates applied to CMS Tracker modules





Rigid, organic build up substrates offer a standard baseline construction for the 2S and PS modules.

The routability has been confirmed using a 1-4-1 build up structure.

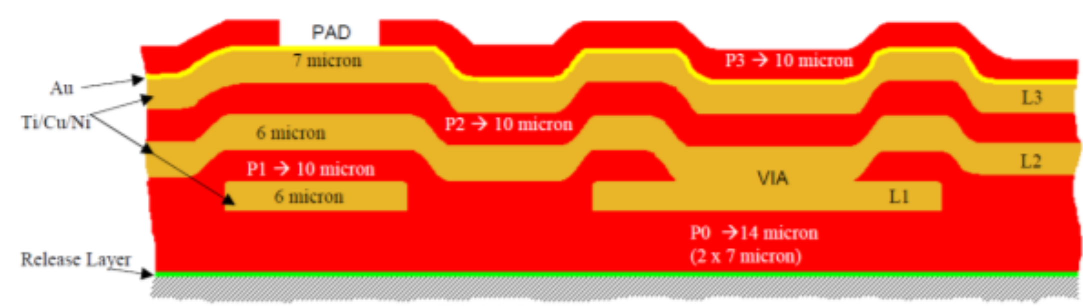
Non negligible mass, but power distribution is adequate to feed the ASICs.

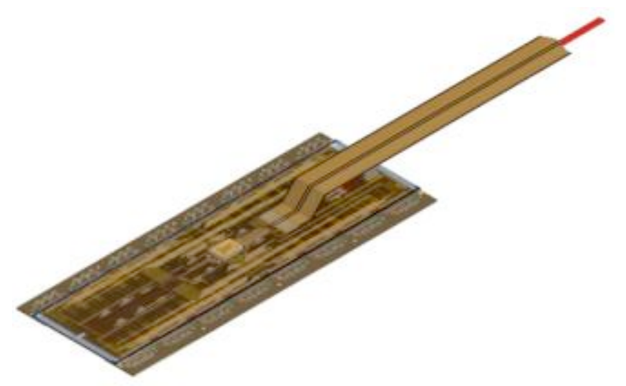
Mechanical integration to be studied: glueing on cooling structure, interconnection with the service board, flatness for wirebonding and bump bonding, wirebonding through groove for bottom sensor.

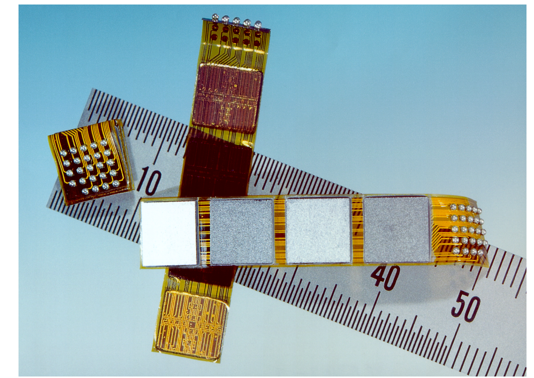
Flexible substrates

- **Flexible polyimide is a quickly emerging technology.**

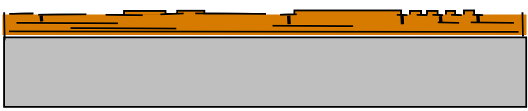
- Thin film flex technology made of spinned liquid polyimide on square panels.
    - Very high density layouts: Tracks w/s = 20 μm, microvias = 30 μm.
    - Silicon matching CTE = 3 ppm/K.
    - Very low mass: Cu thickness < 7 μm, film thickness ≈ 10 μm.
    - However: 4 layers maximum, no copper on base layer, limited power delivery capabilities.

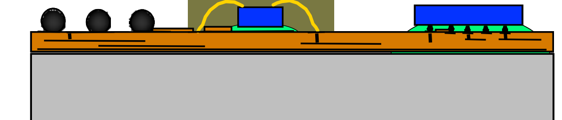




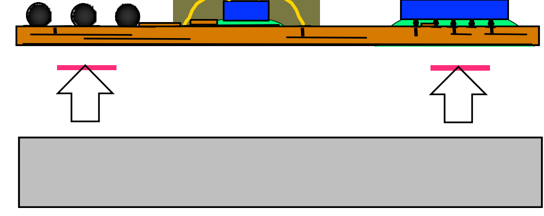


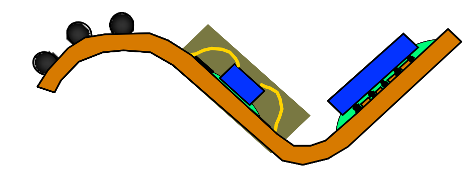
Fabrication of Multilayer Structure on Rigid Carrier Substrate





Separation of Multilayer from Rigid Substrate Reuse of Carrier

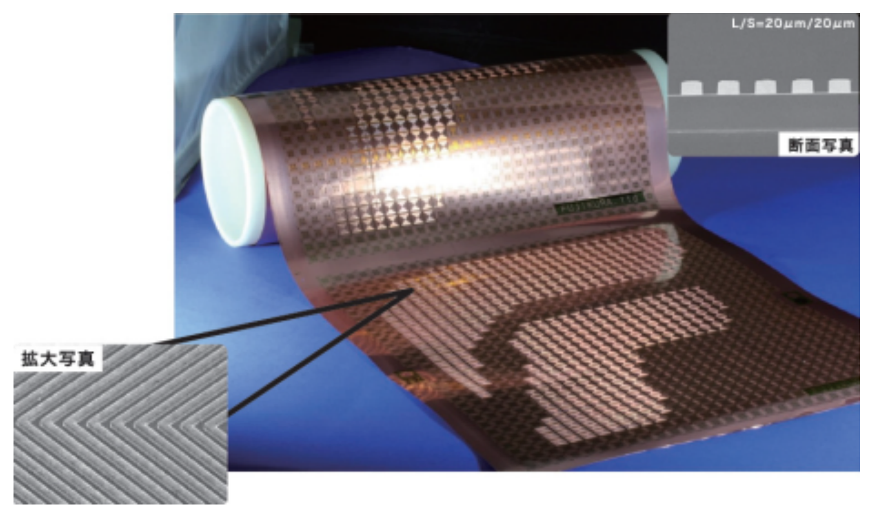




Flexible substrates

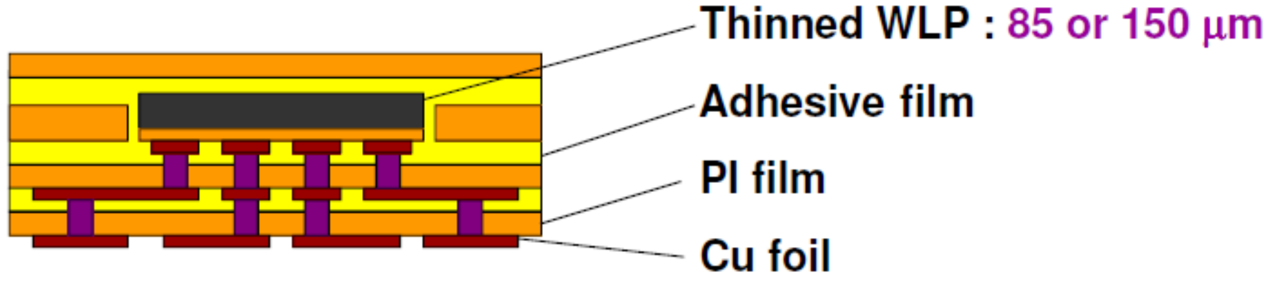
- **Packaging industry is adopting this technology for large volume and integration.**

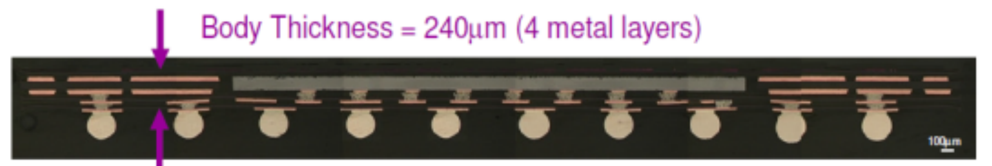
- Several suppliers are today available for panelized flex films.
  - They all provide very high density, small microvias, thin foils on limited number of layers.
  - Flip chip compatible, wirebonding compatibility to be evaluated.
- Trend is now to use this technique for:
  - Roll to roll lamination of flex circuits for very large volume productions.
  - Embedding of dies into multilayer system in package overmolded structures.



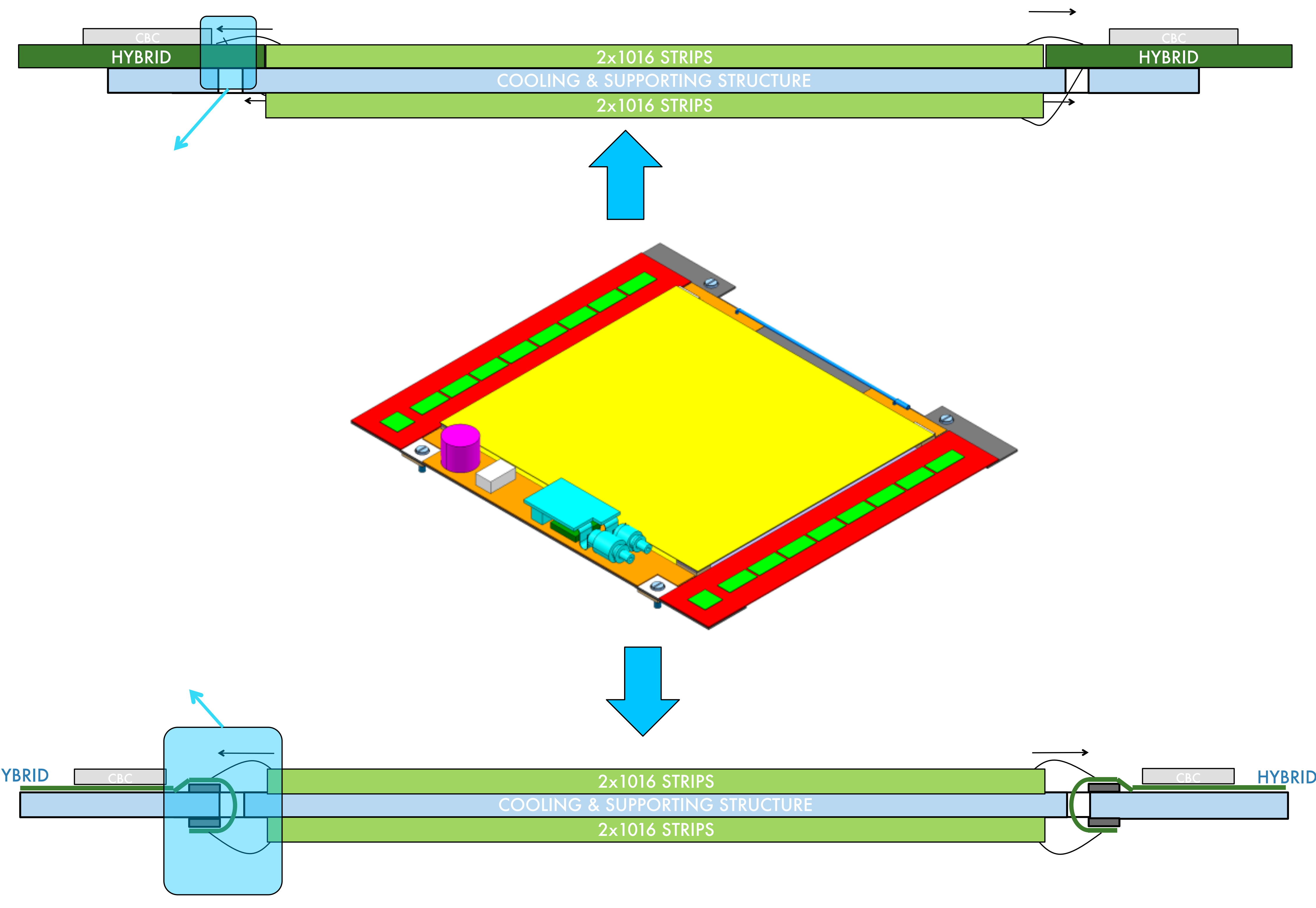
IMAPS MINAPAD Forum
Grenoble, April 2012.

Ultra slim package due to the thinned WLP die and multilayer flex based structure.

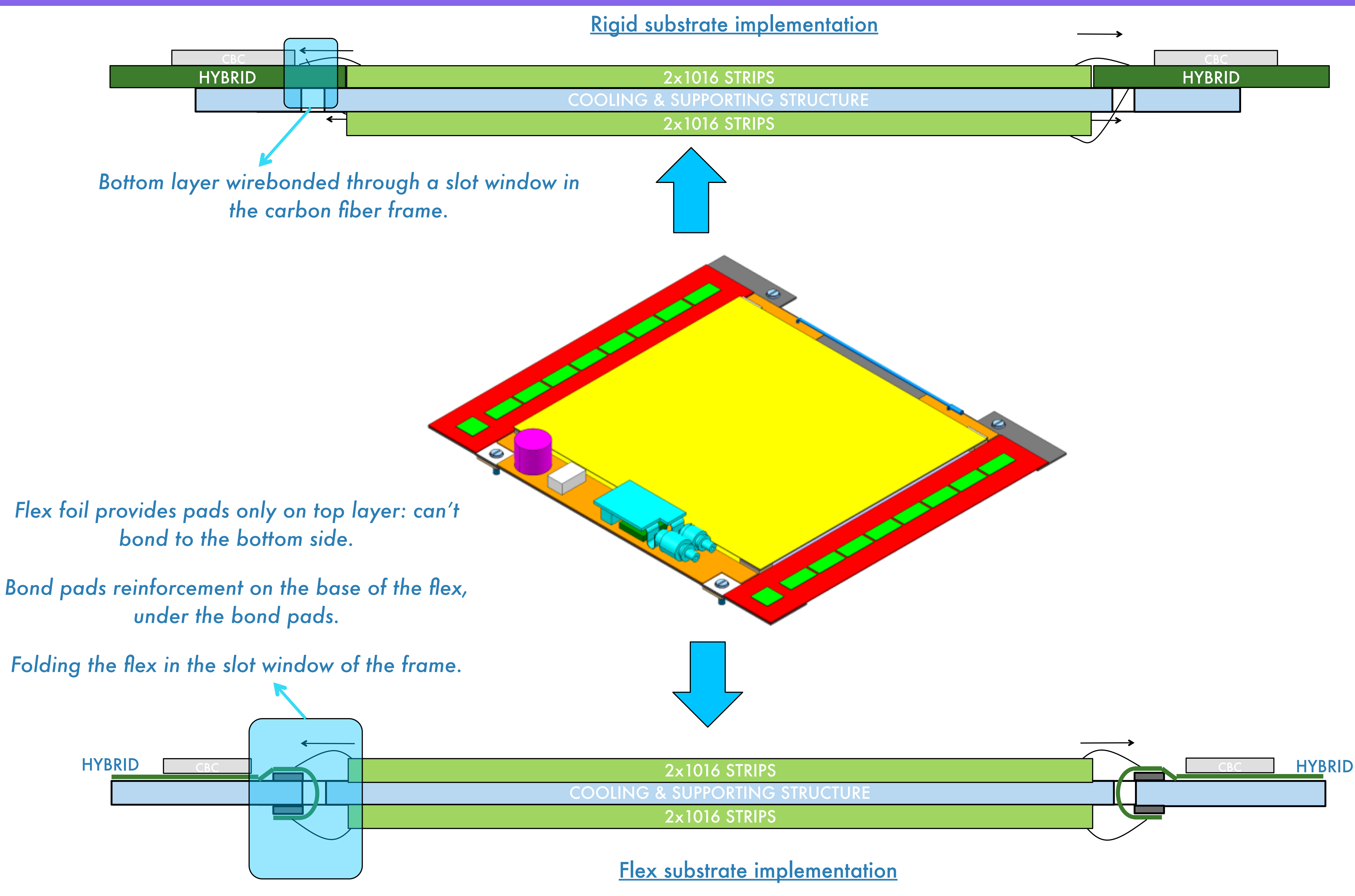




Flexible substrates for the CMS tracker modules

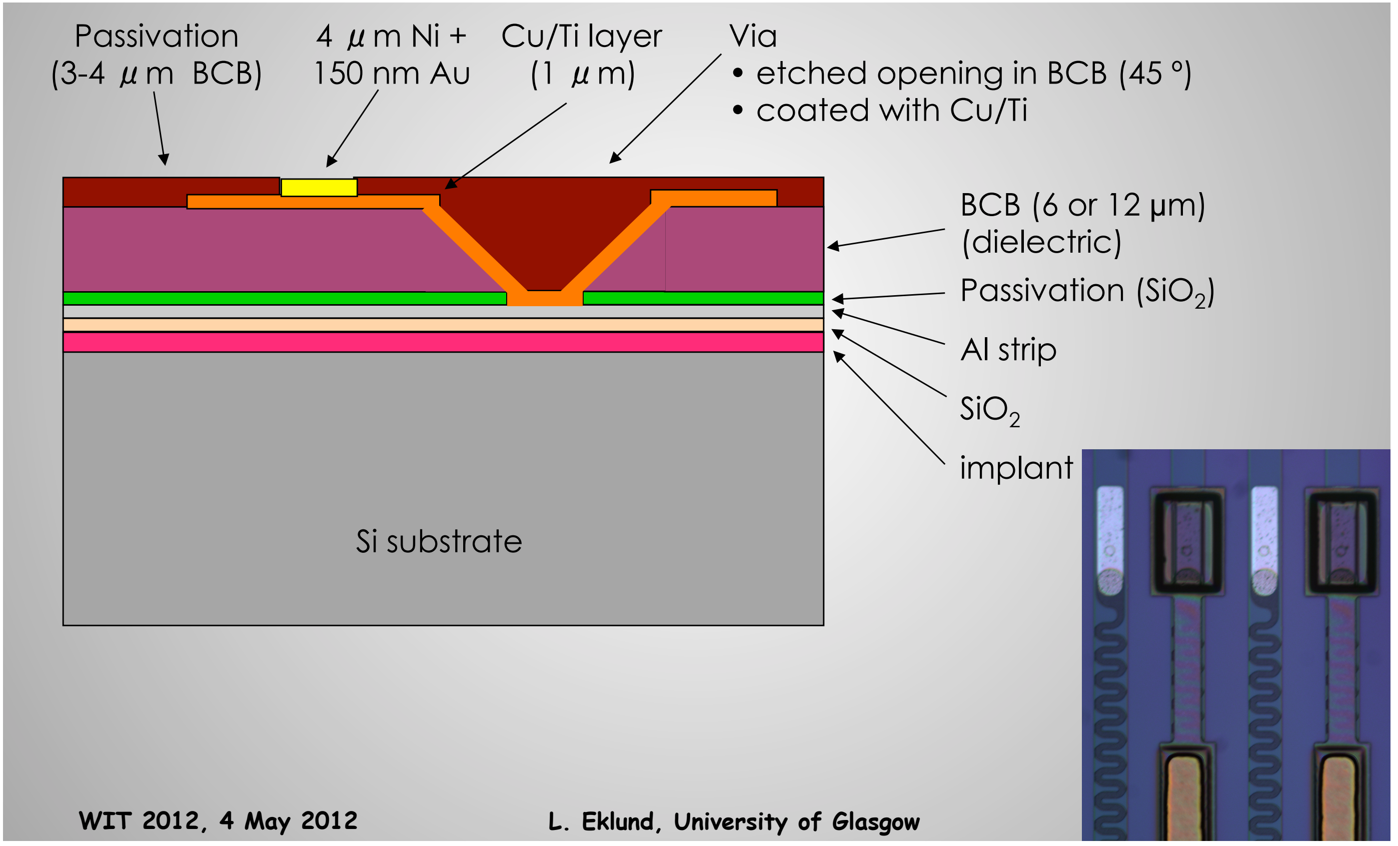


Flexible substrates for the CMS tracker modules



Multi Chip Module-Deposited (MCM-D) to build FE directly on the silicon sensor

- Traditional silicon module build (electrical parts)
  - Sensors,flex circuit,substrate,pitch adaptor, wire bonds, FE-chips, passive components



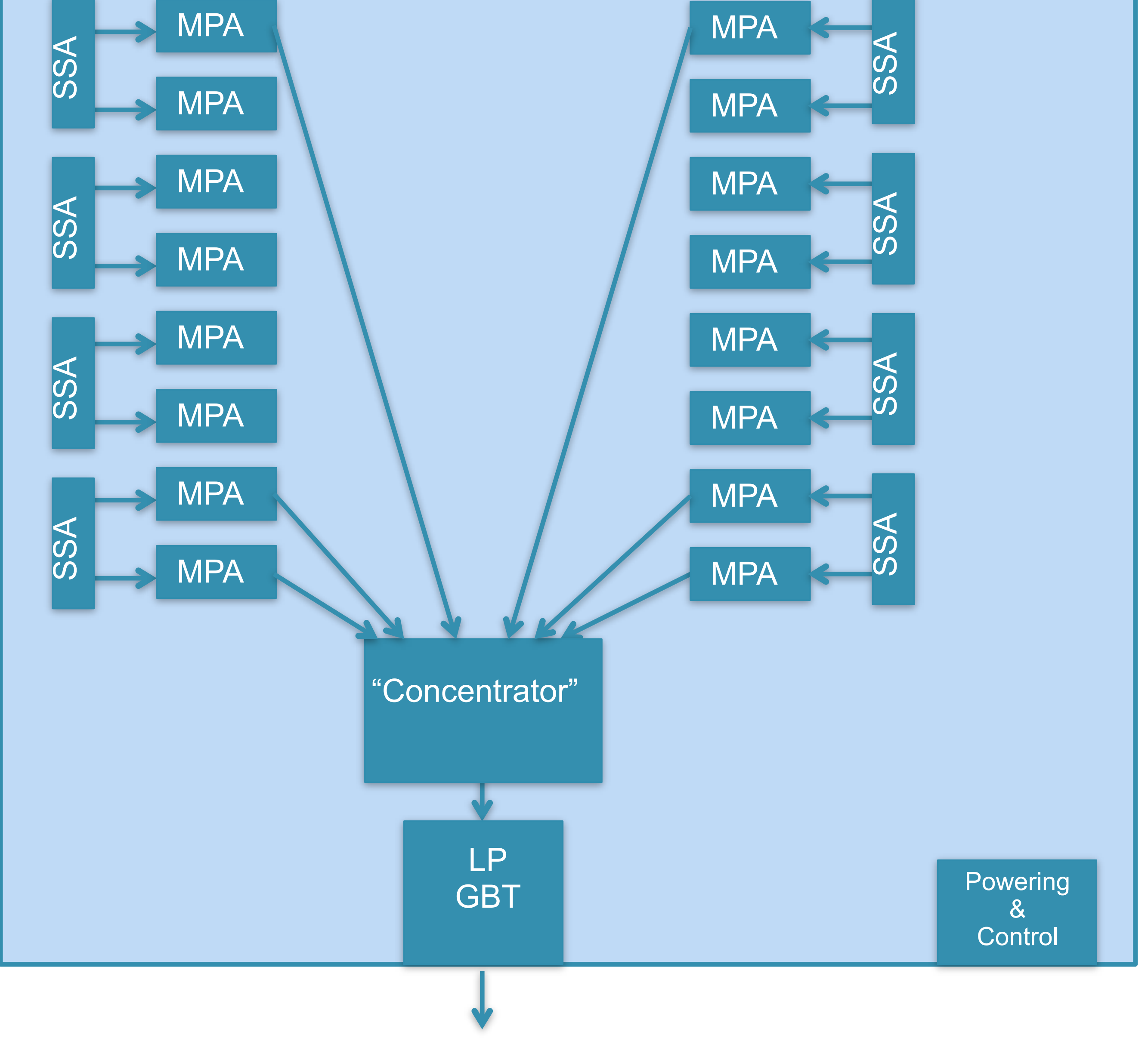
- MCM-D
  - Deposit dielectric and metal layers directly on thesilicon sensor Layout concepts similar to PCBs All-in-one: Sensor, hybrid, pitch-adaptor and strip connections

Module read out architectures

Option 1

SSA MPA MPA
SSA MPA MPA
SSA MPA MPA
SSA MPA MPA
MPA MPA SSA
MPA MPA SSA
MPA MPA SSA
MPA MPA SSA
LP GBT
Powering & Control

Option 2

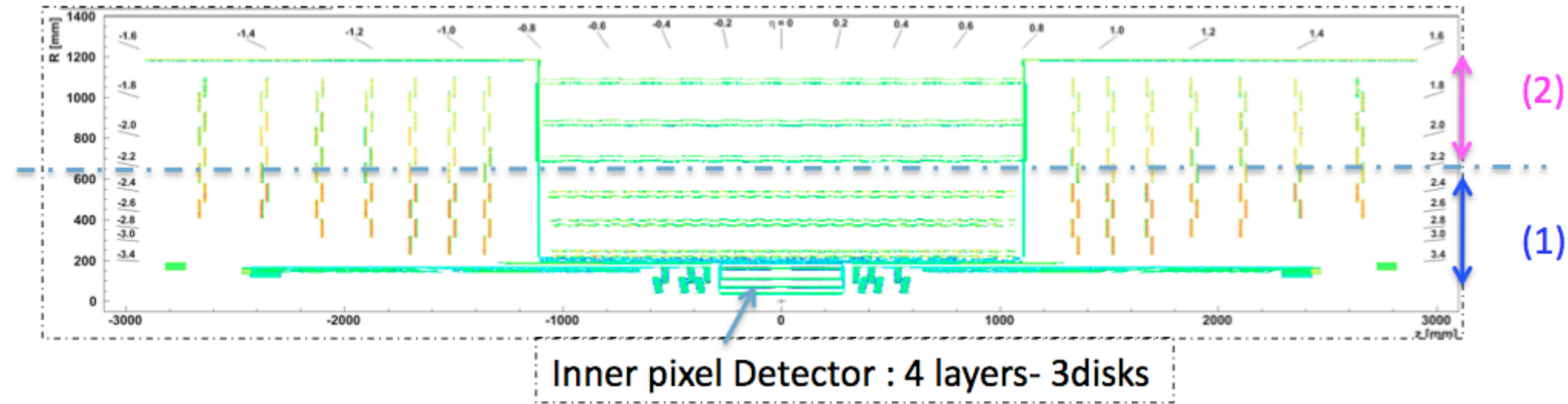


➢ Separate inputs for “trigger” and “readout”

Some simulation results

CMS - traditional geom.

The “Barrel-EndCap” design comprises 6 barrel layers and 7 endcap disks composed of rings.



The inner part (1) is populated by Pixel-Strip stacked (PS) modules .

The outer part (2) is populated by Strip stacked (2S) modules.

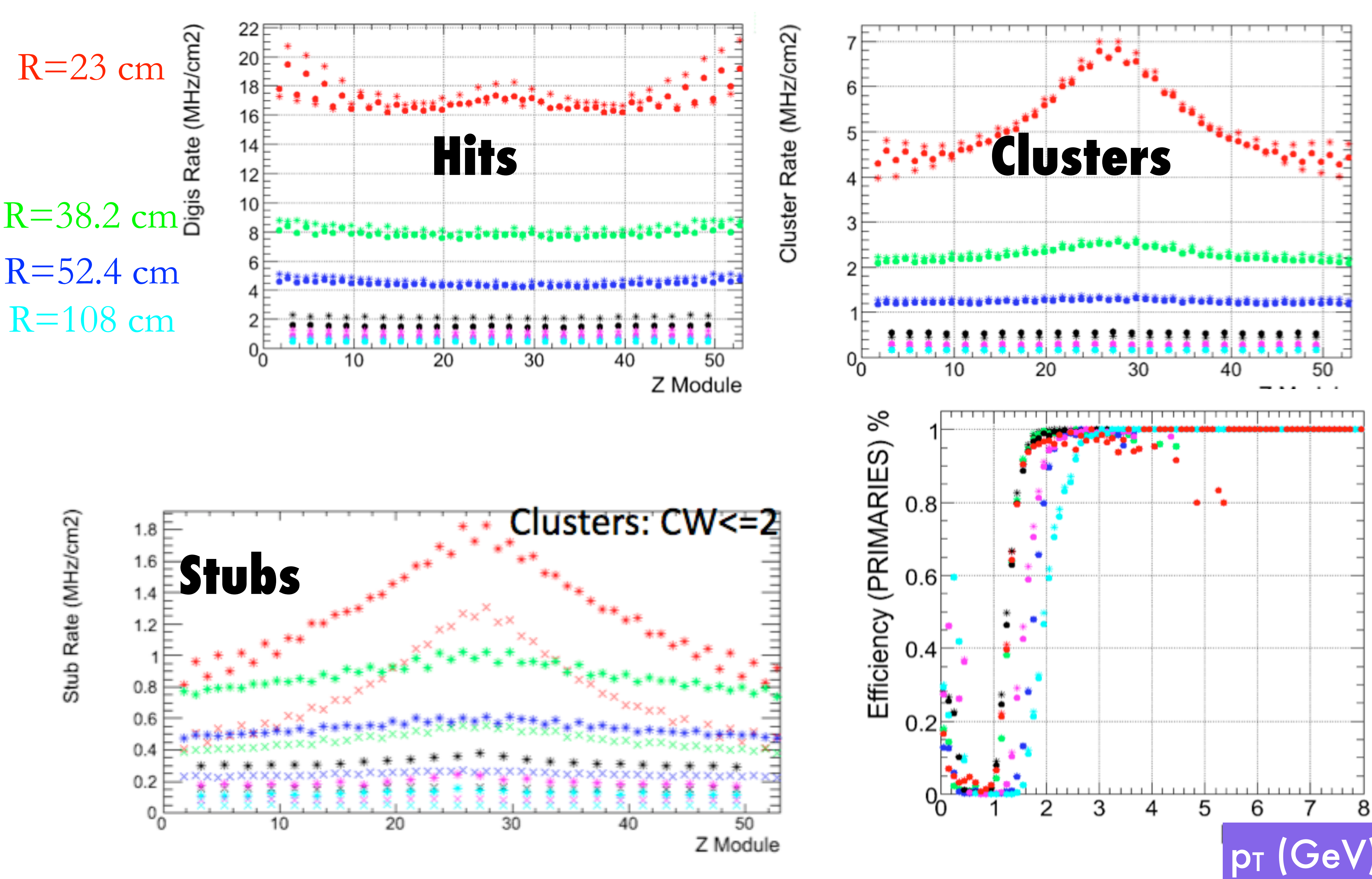
The number of endcap disk is optimized for tracking performance.

Different spacings between the two sensors of the Pt modules:
0.8mm in the outer barrel (2S)
1.6 and 2.6mm in the inner barrel (PS)
4.0, 2.6 and 1.2mm in the outer end-cap (2S)
4.0mm in the inner end-cap (PS)

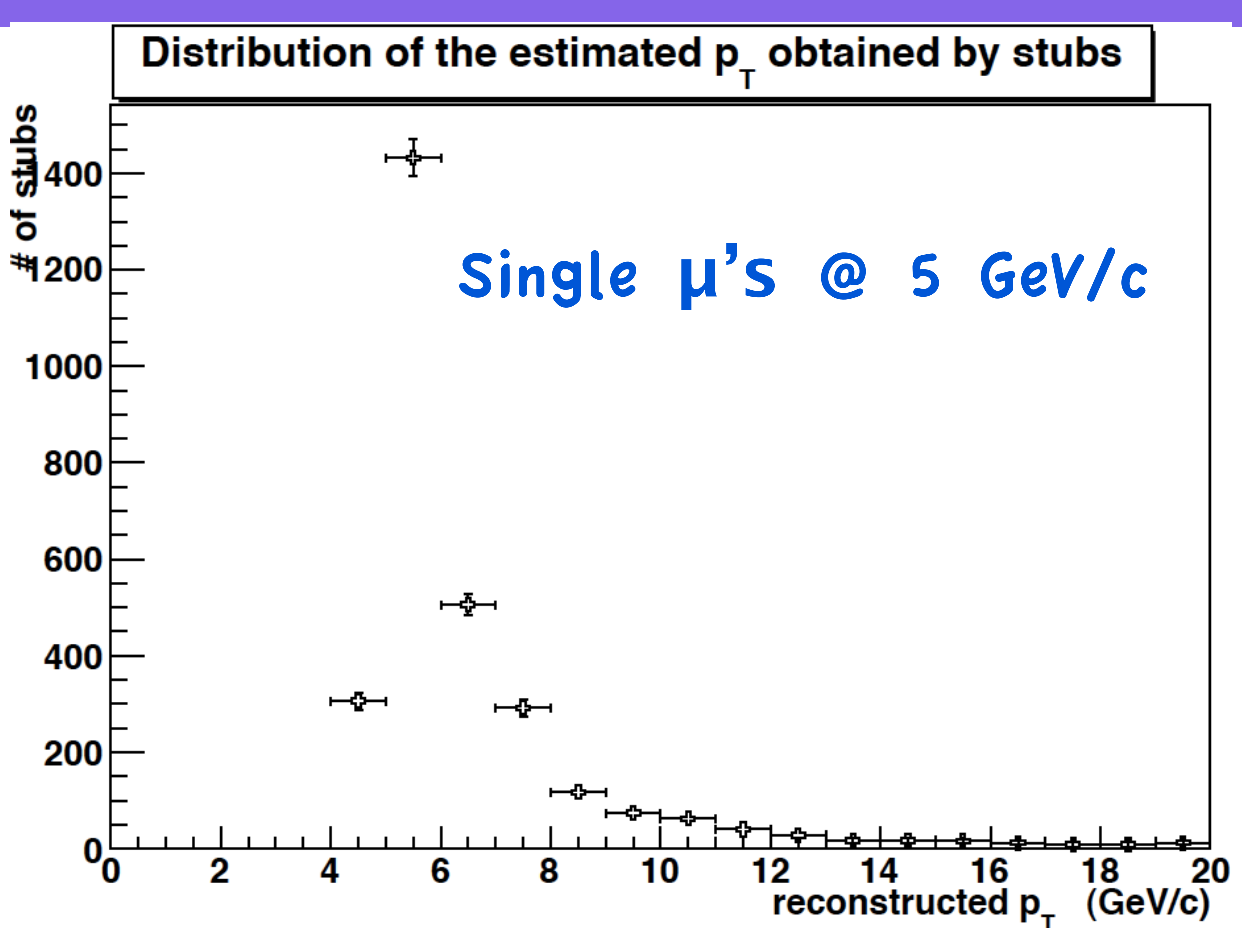
L1 tracking precision potential
- pT resolution 4% @ 10 GeV in forward

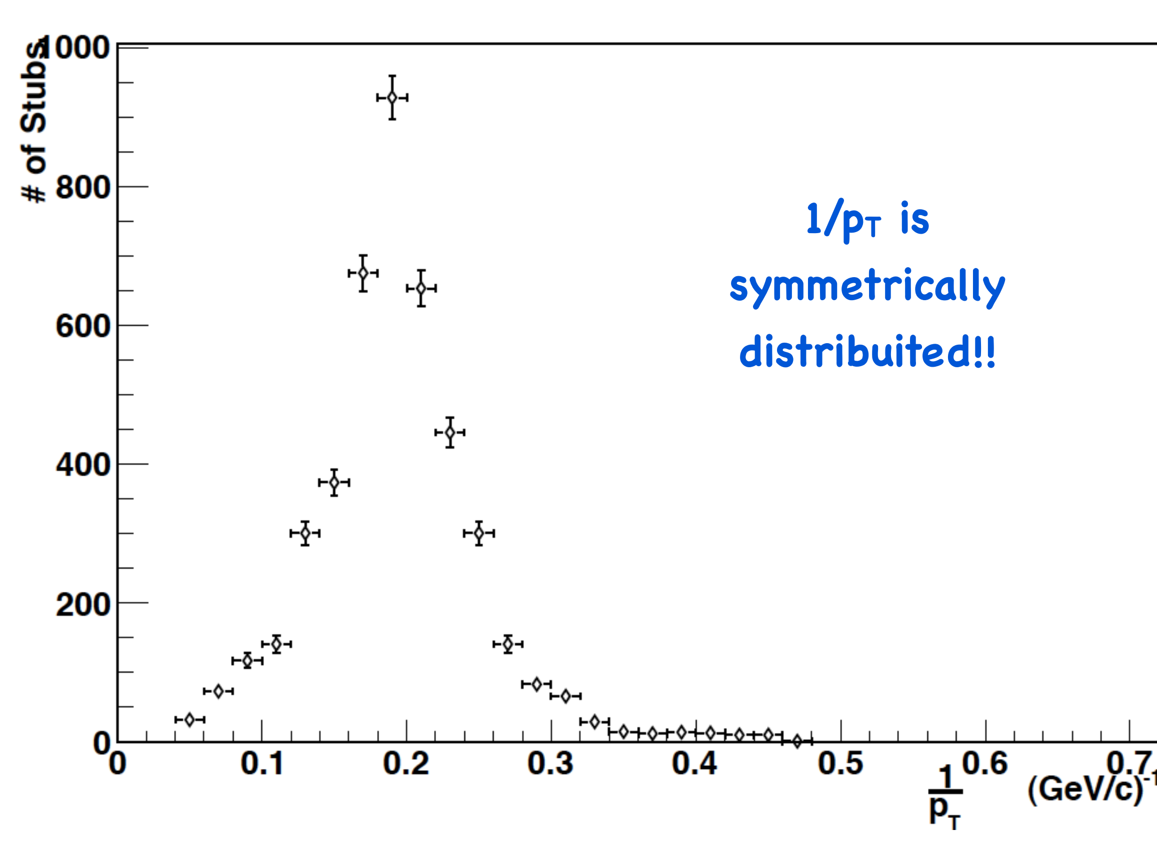
Tracking precision
- pT resolution 1.4% @ 10 GeV
- pT resolution 3% @ 100 GeV

Data reduction



Stub p_T Measurement





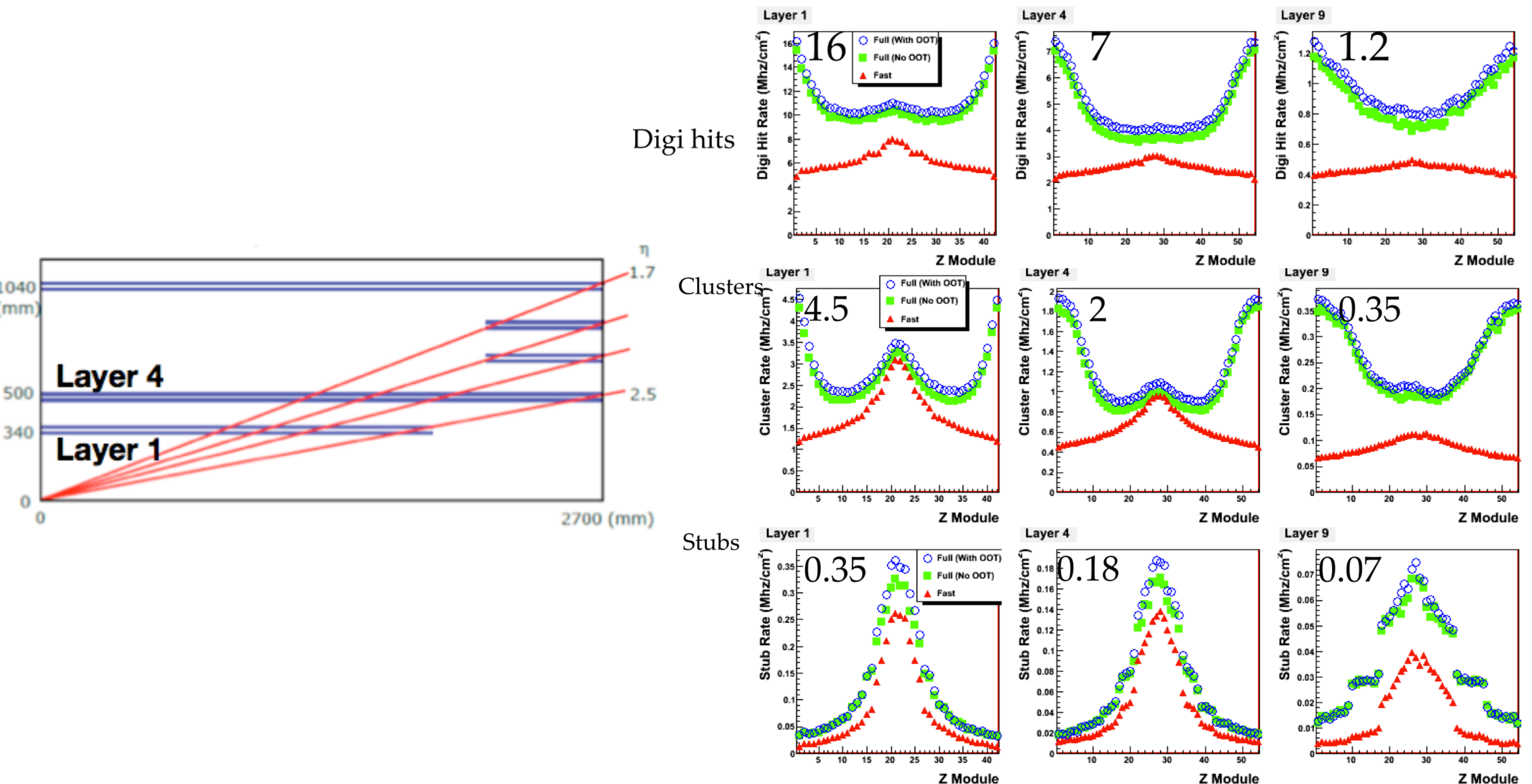
- μ-strip 98 μm pitch
- sensors separation ΔR=1 mm

R 51 cm - mean $\sqrt{(1/p_T)}$: ~0.076

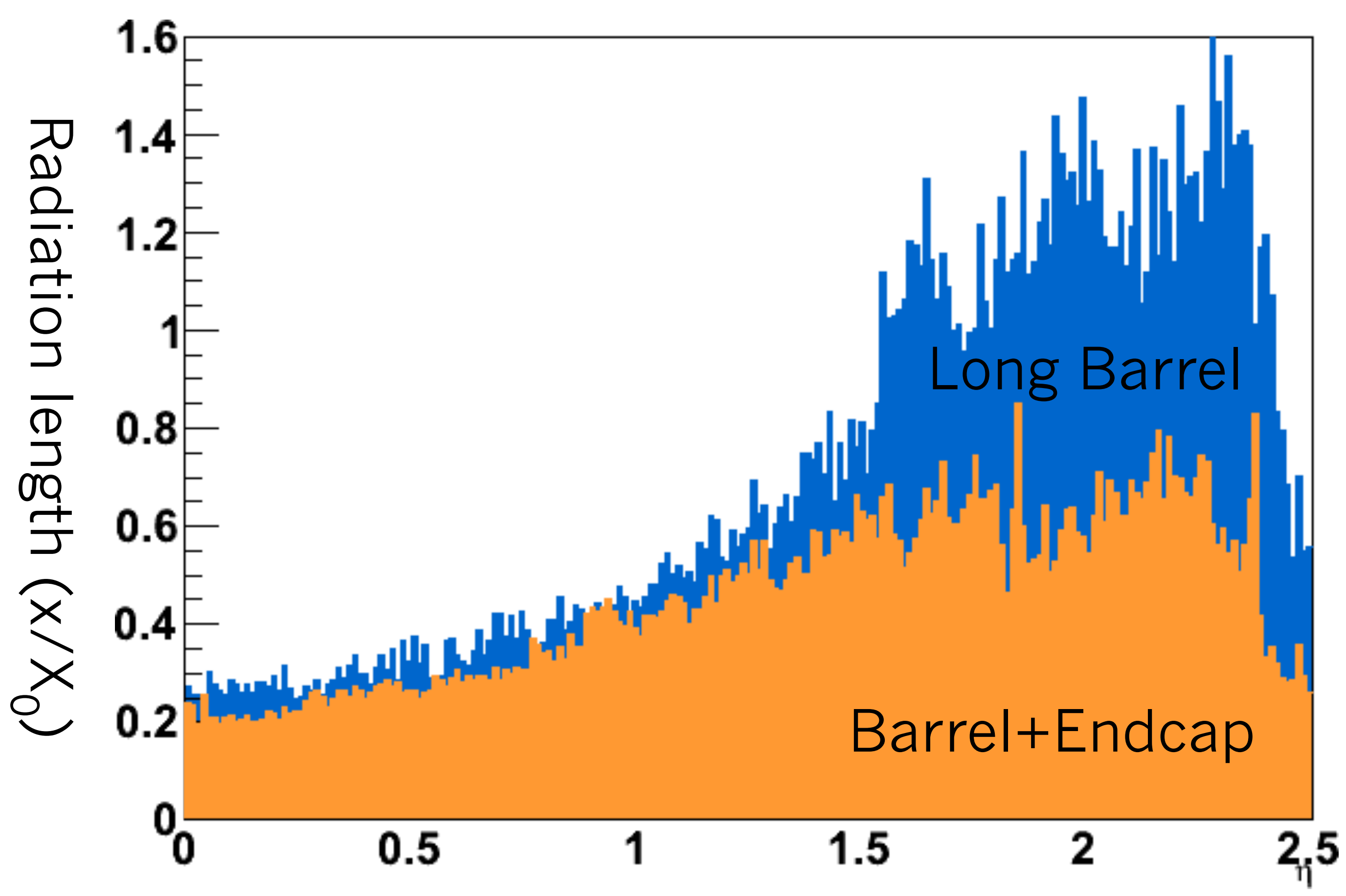
R 82 cm - mean $\sqrt{(1/p_T)}$: ~0.073

R 102 cm - mean $\sqrt{(1/p_T)}$: ~0.069

Long Barrel layout (CMS)

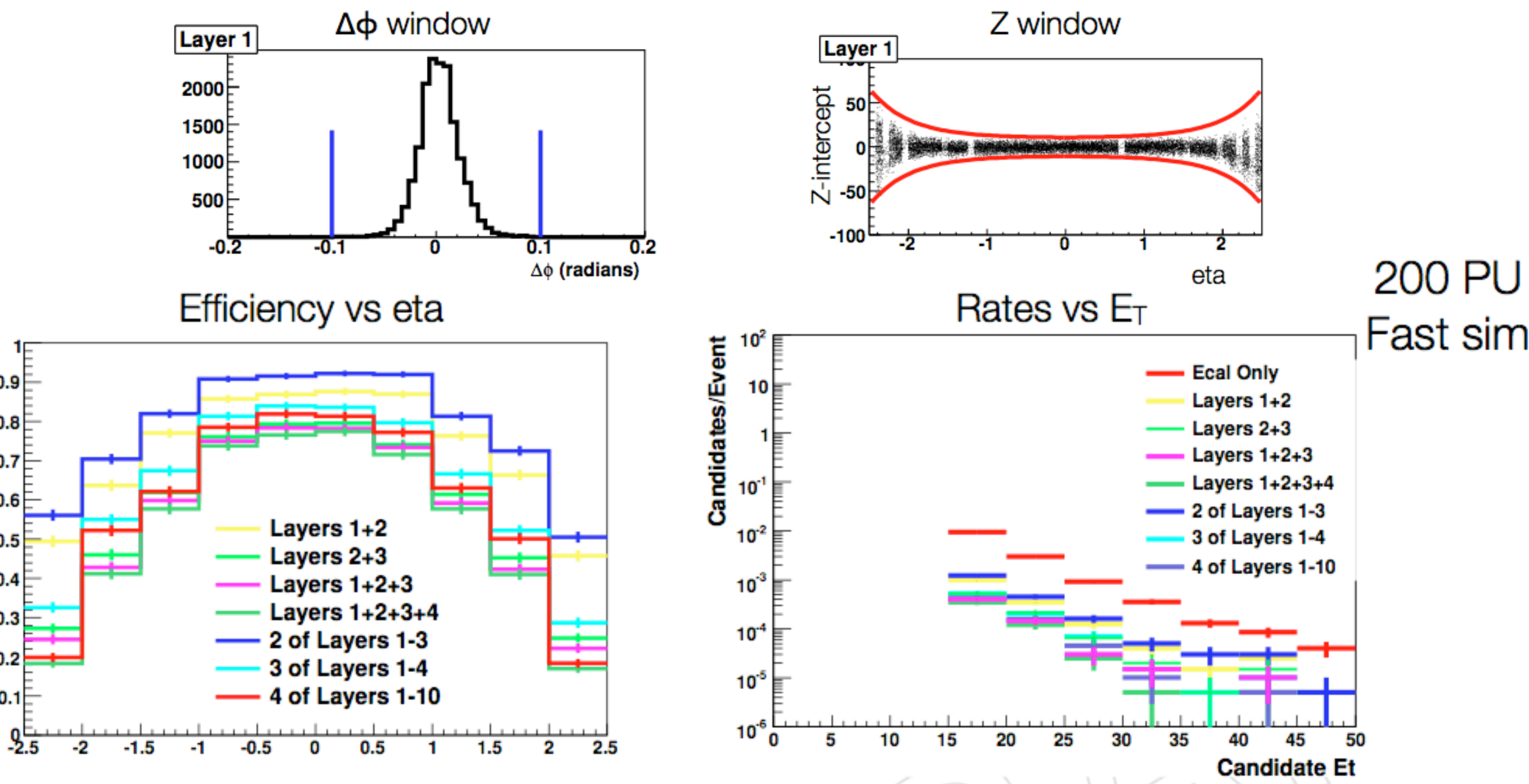


Geometry Comparison



Example: match ECAL+stub info

Matching between stub and projected electron trajectory:



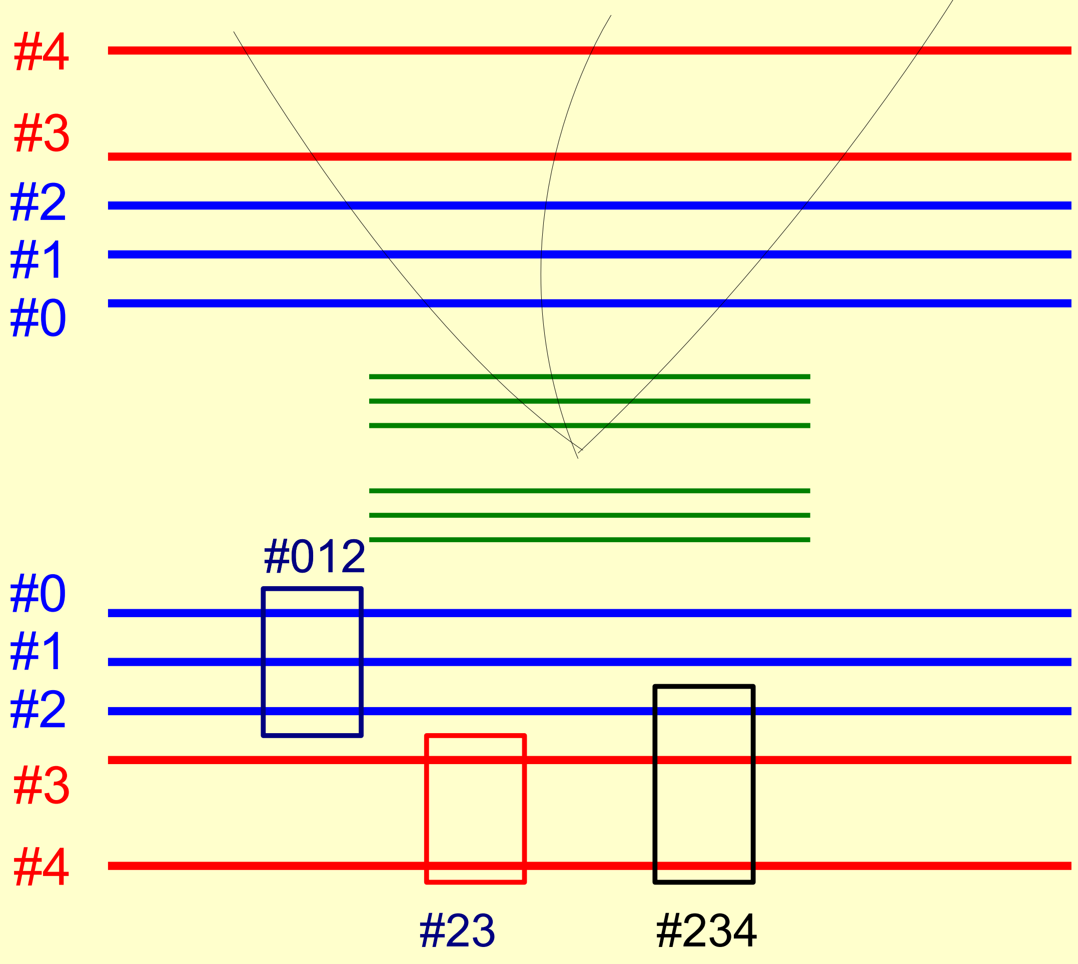
ATLAS

Pixel + Strip Sensor Layers

Long Strips (*Δz=10cm*)

Short Strips (*Δz=2.5cm*)

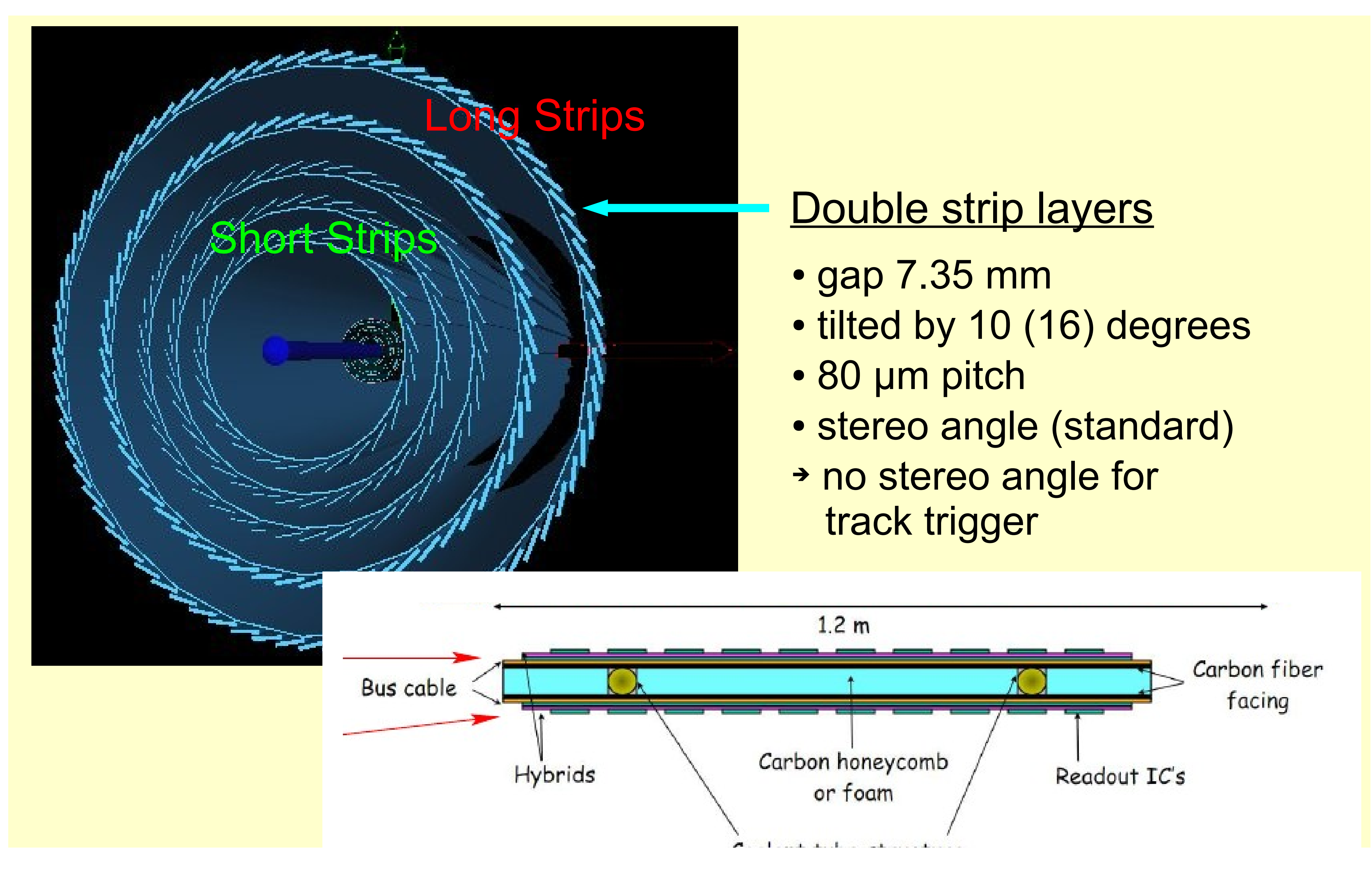
Pixel (not used)



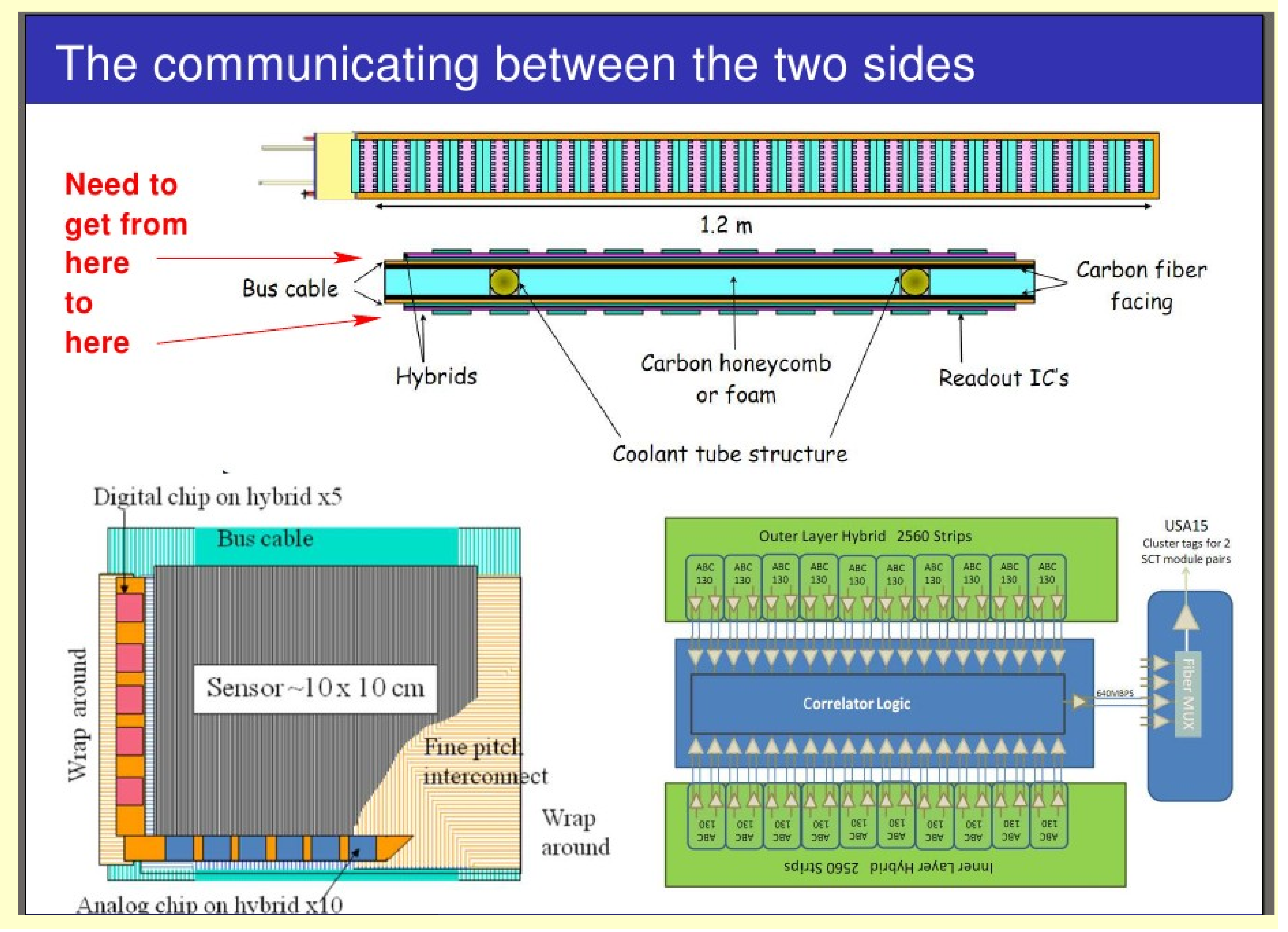
Layer combinations studied for track trigger:

- #0, #1, #2 (only short strips)
- #3, #4 (only long strips)
- #2, #3, #4 (mixed, outer layers)

ATLAS layout



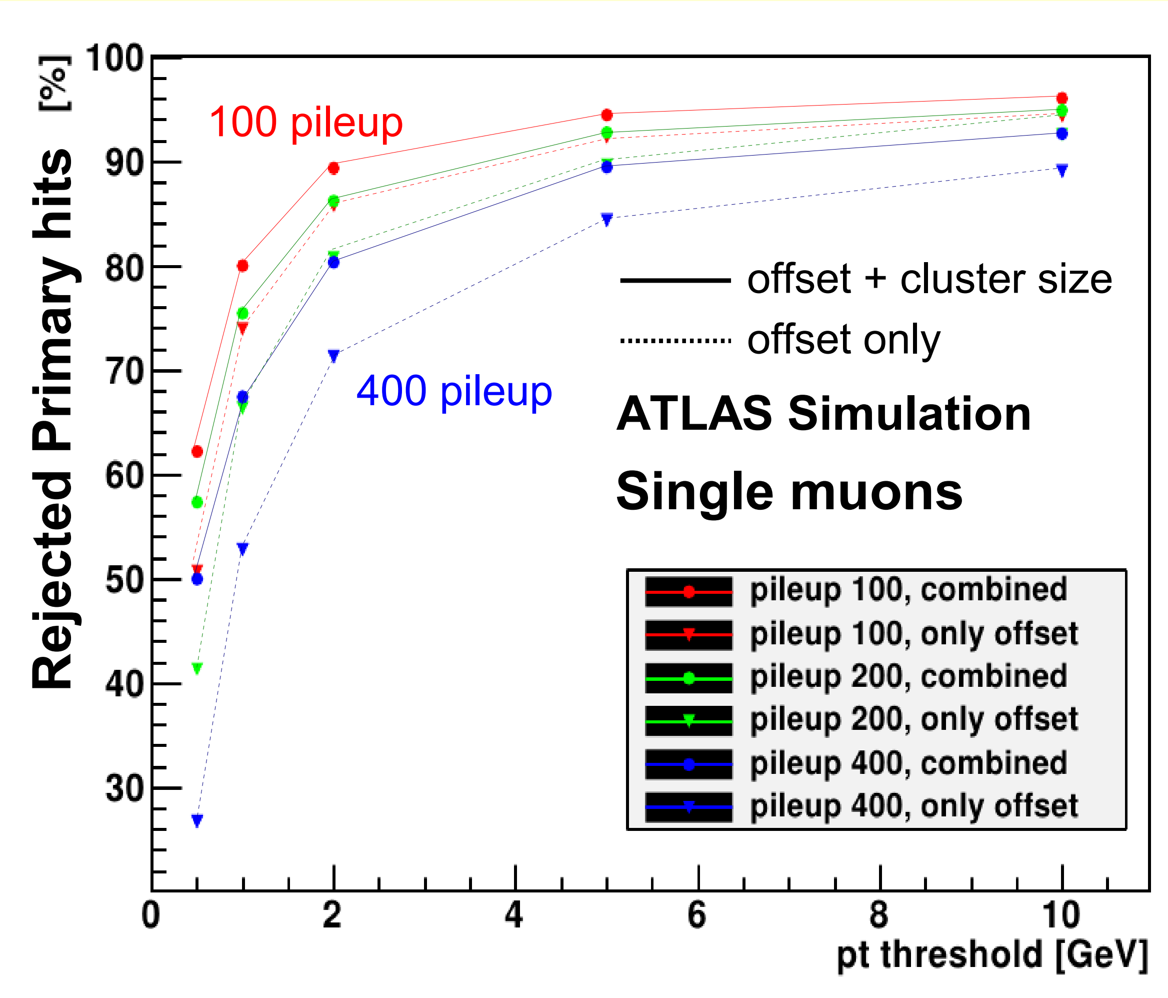
Fast clustering in ATLAS



Rejection as Function of p_T Threshold

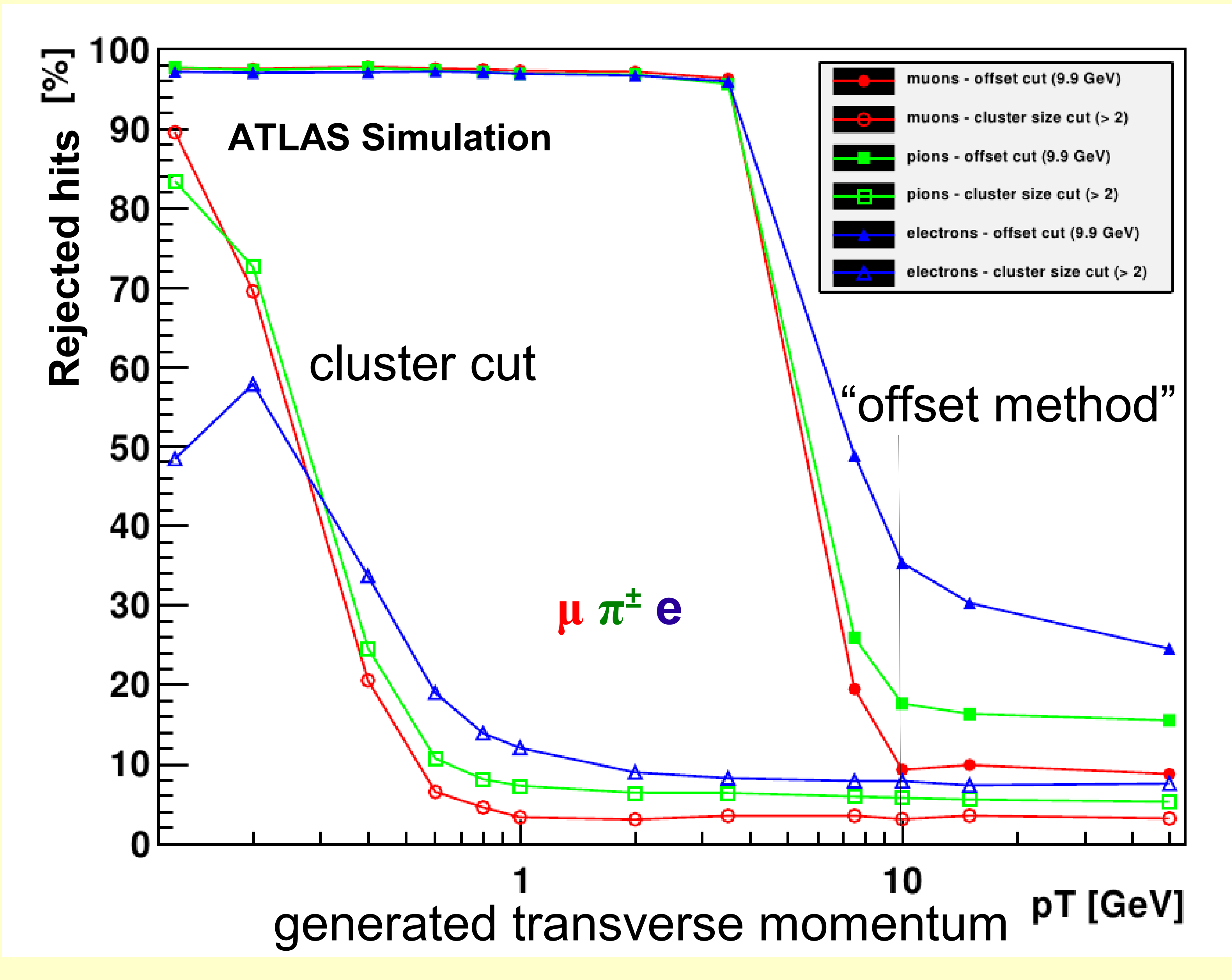
Rejection
of Primary Hits

Minimum Bias
Events



- **most tracks (at low p_T) are rejected already with a low p_T threshold**
- **rejection power higher if cluster size and offset cut are combined**
- **rejection power affected by high pileup**

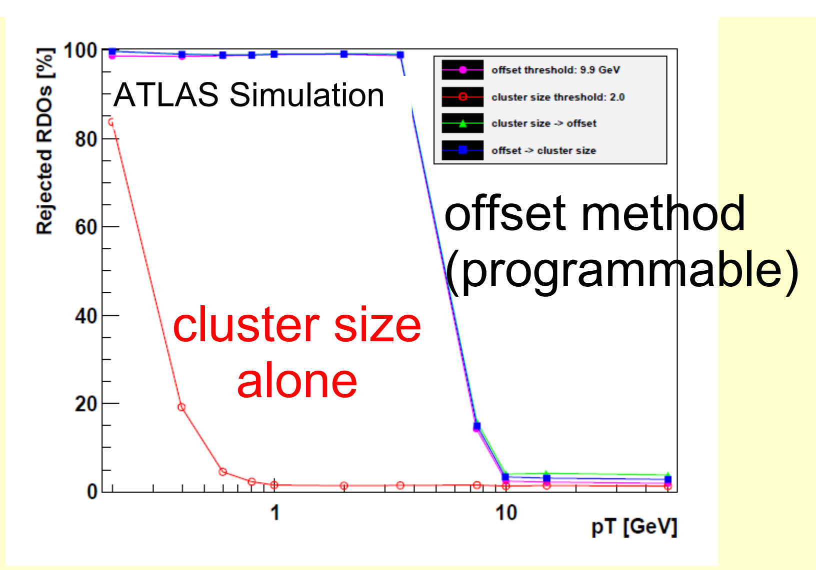
e, μ, π± Rejection (single particle)



hit reduction also above p_T threshold due to secondary IA

Performance in ATLAS

- pileup 100 minimum bias (Pythia)
- p_T>10 GeV (offset)



offset cut only

	# hits (layer)	27-153 degrees # hits (SS 3 accept.)	40-140 de # hits (LS 2 acc
SS 1:	6.4%	4.3%	2.8%
SS 2:	5.5%	4.7%	2.9%
SS 3:	5.1%	5.1%	3.4%
LS 1:	8.0%	8.0%	6.2%
LS 2:	6.5%	6.5%	6.5%

Hardware Implementation:

number of patterns $O(10^{10})$

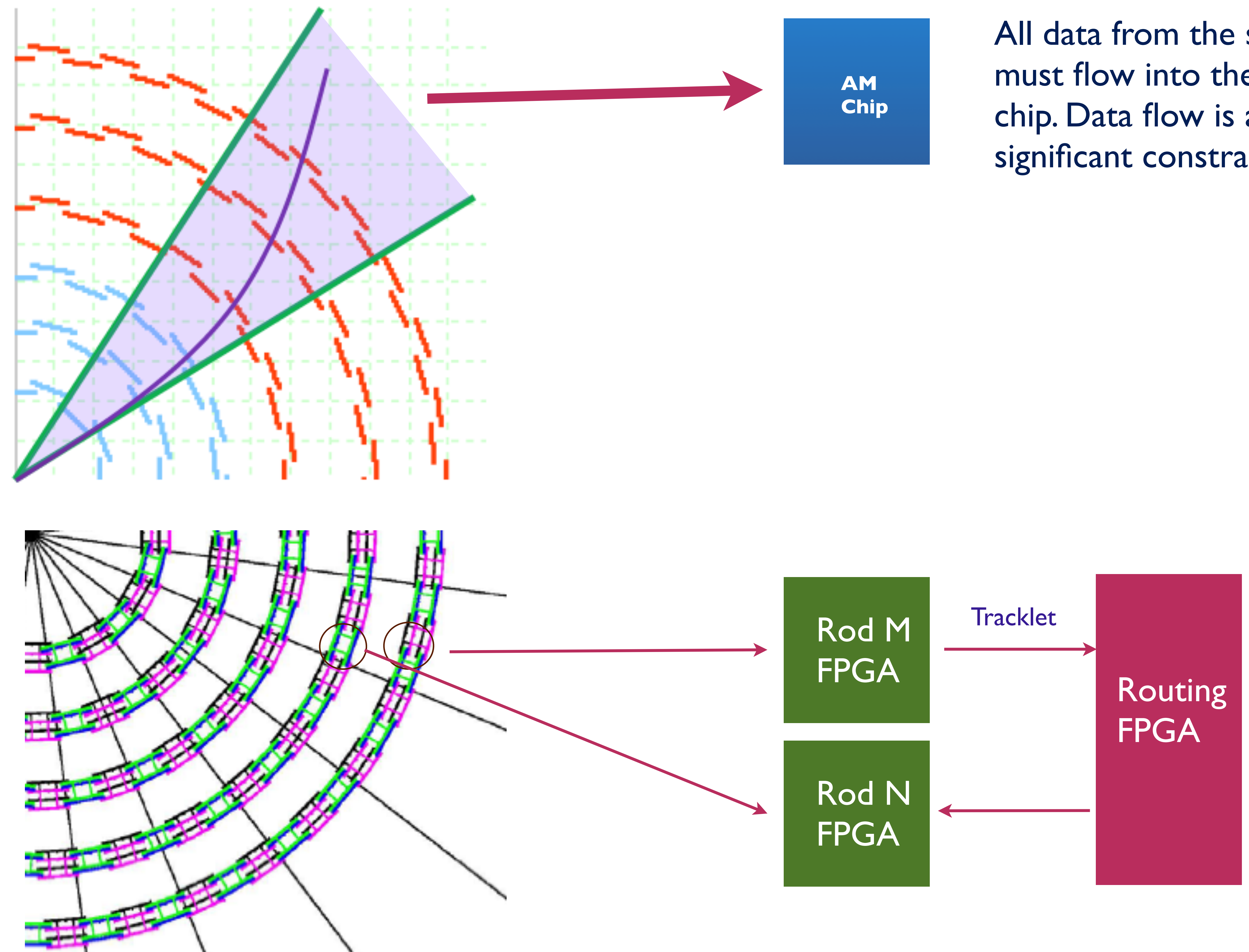
→ talk S.Schmitt (WIT2010

Reduction factors of: 15-30 on short strip layers

~15 on long strip layers

Pattern recognition

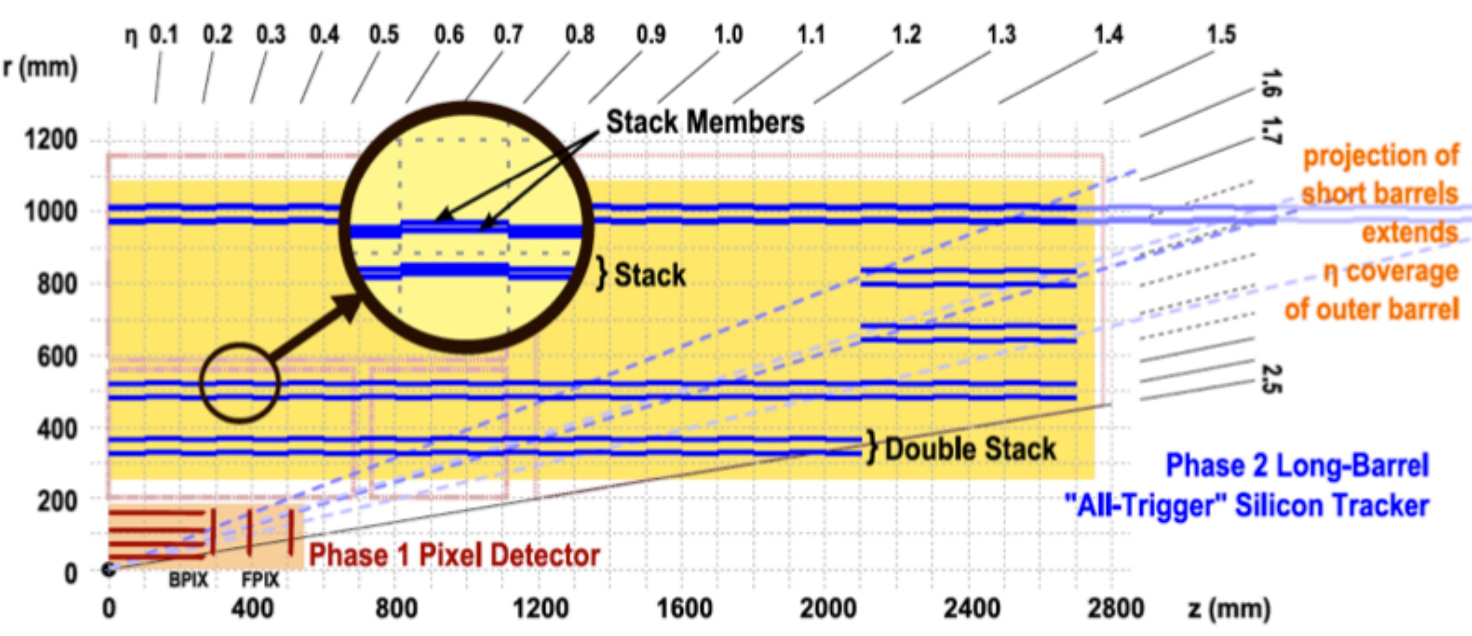
Local FPGA vs AM

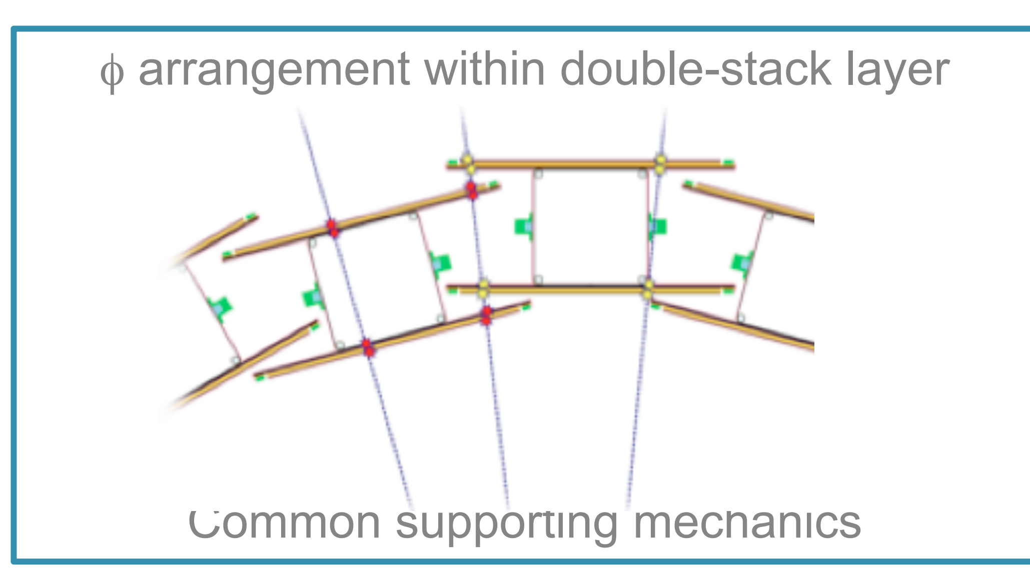


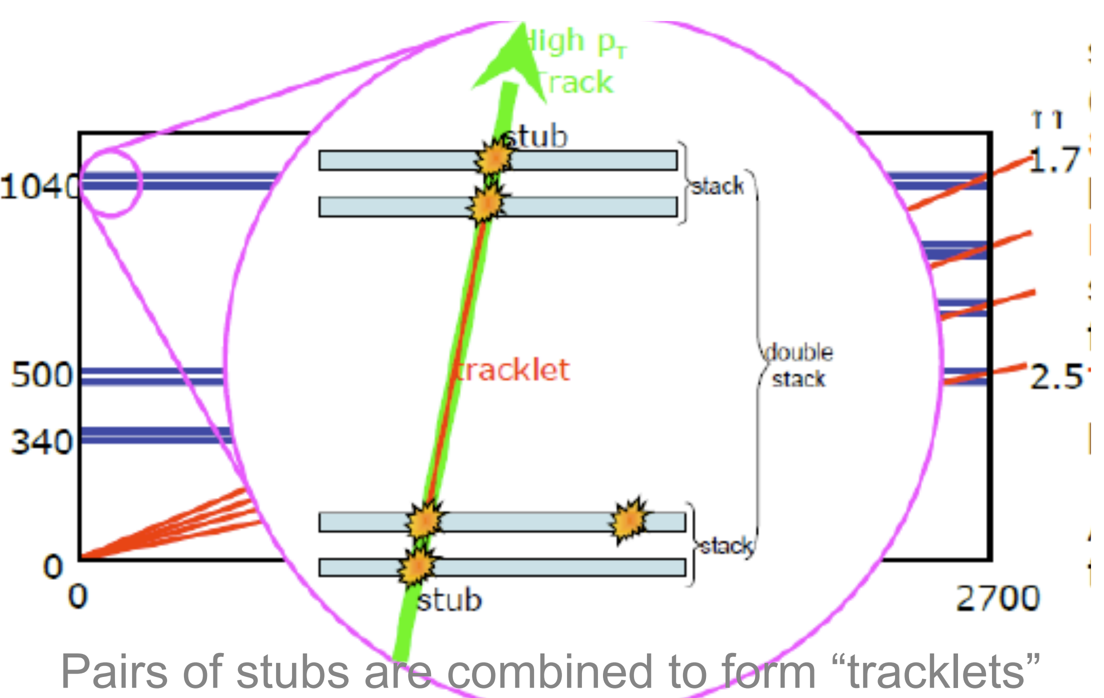
All data from the sector must flow into the AM chip. Data flow is a significant constraint

Long Barrel option in CMS

➤The "long-barrel" double-stack layout

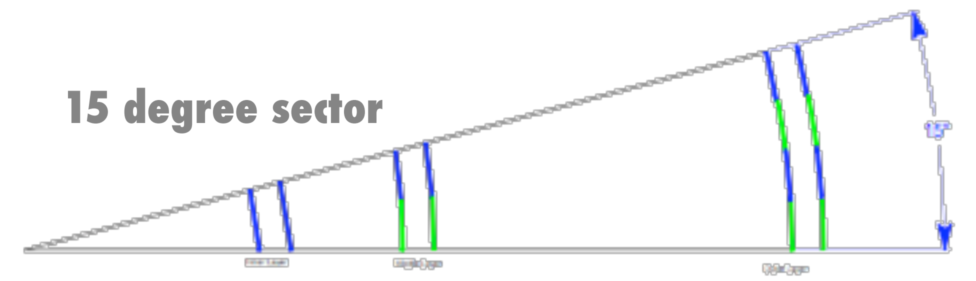






Pairs of stubs are combined to form "tracklets"

Self-contained ϕ sectors.
Each sector needs to be combined with the two neighbouring sectors (left and right) to "contain" ~2.5 GeV tracks.

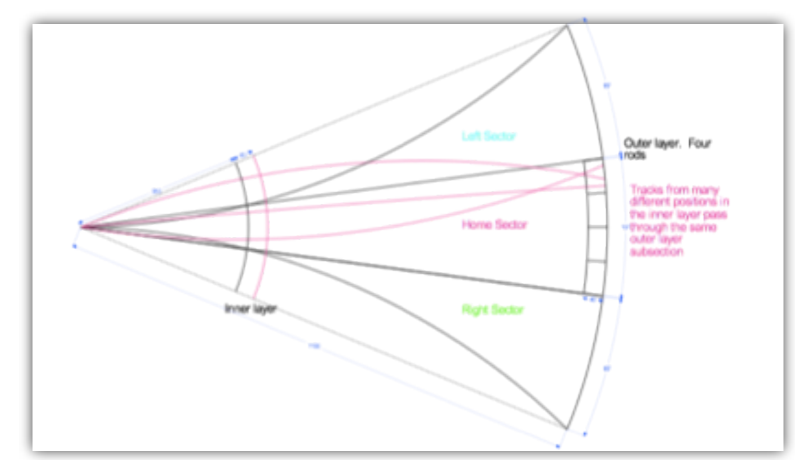


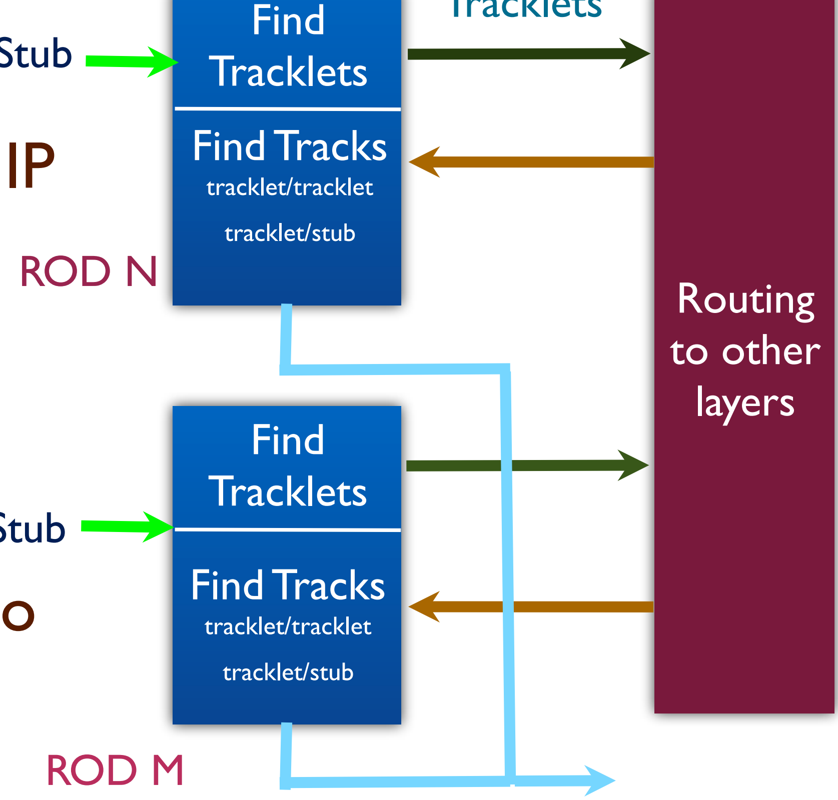
Off detector processing

Off-Detector Processing

The local design minimizes data transfer and interconnection complexity.

- Input FPGA finds tracklets within rod N by comparing stub phis within a window defined by the beam spot
- Check that DZ is consistent with IP
- Project to other layers using a table look-up
- Move tracklet N information to destination rod N
- Compare tracklets N projection to tracklets and stubs in rod M to form track candidates



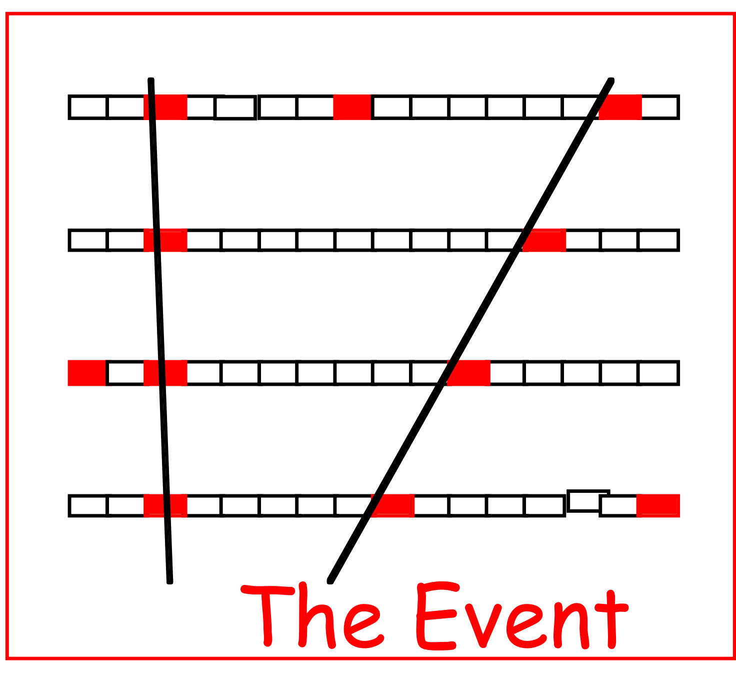


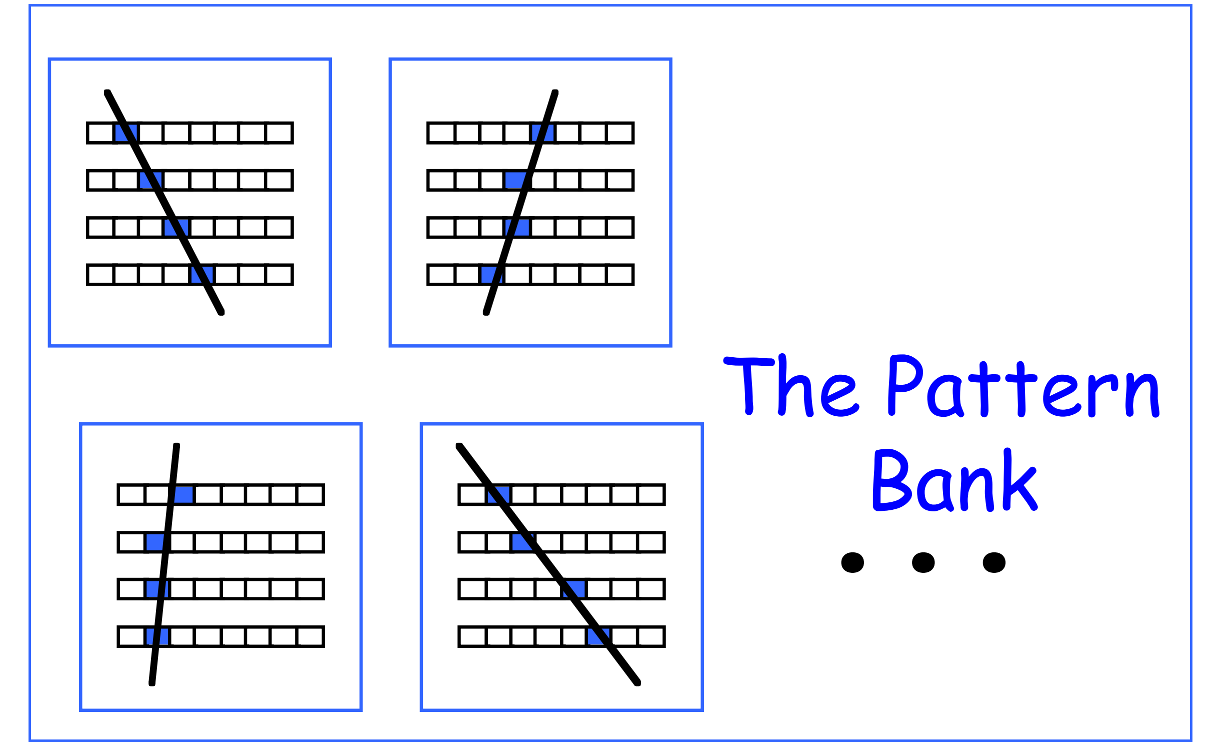
p_T (GeV)	$\sigma(p_T)/p_T$	$\sigma(z)$
3	~1%	1.5 mm
10	~1.5%	0.6 mm
30	~2.5%	0.5 mm

First latency estimate ~2μs

Pattern matching with AM

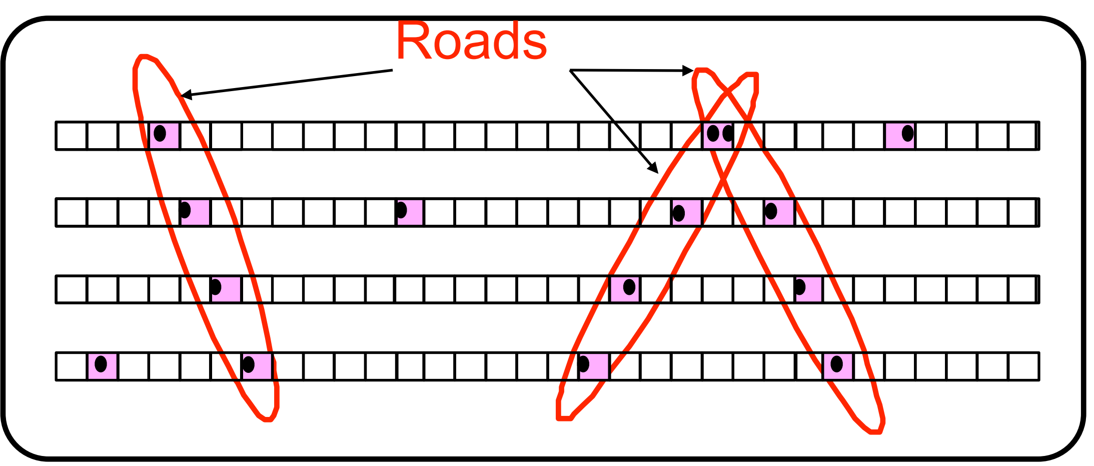
- The pattern bank is a set of pre-calculated patterns
  - can accommodate for alignment
  - changing detector conditions
  - beam displacements
- An Associative Memory holds different patterns banks and compares them with the current event pattern





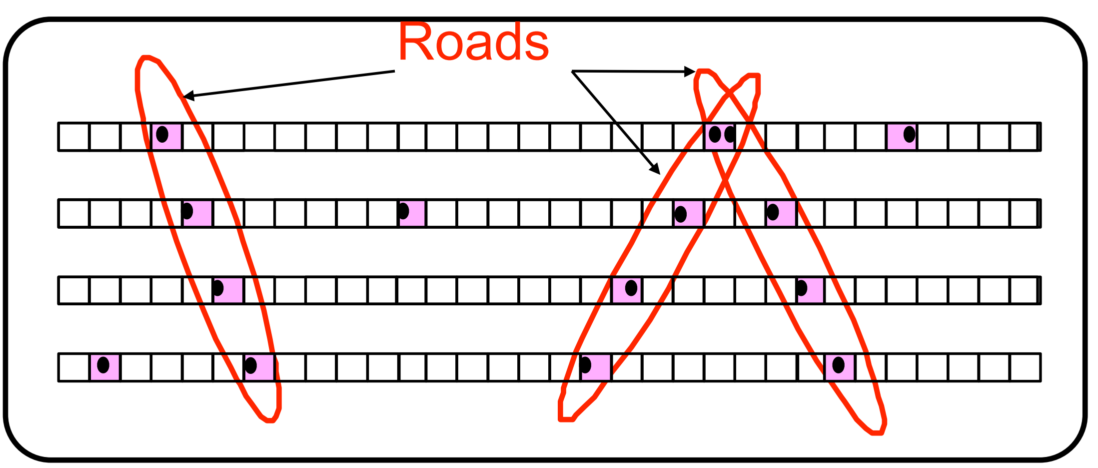
Too large AM? 2 step approach

1. Find low resolution track candidates called "roads". Solve most of the pattern recognition



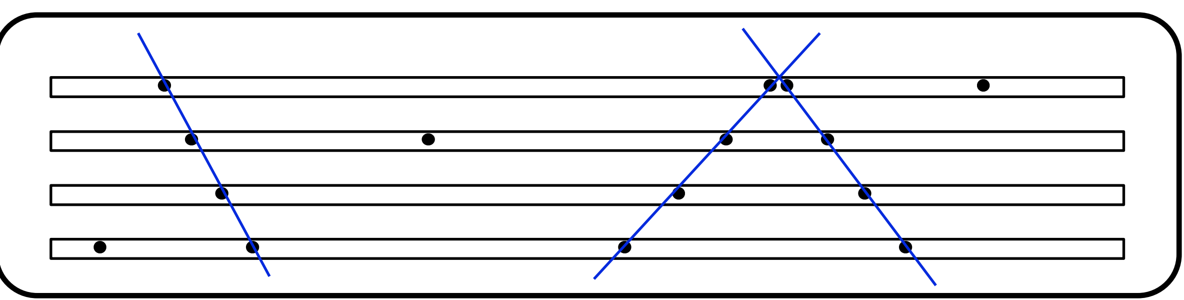
Too large AM? 2 step approach

1. Find low resolution track candidates called "roads". Solve most of the pattern recognition



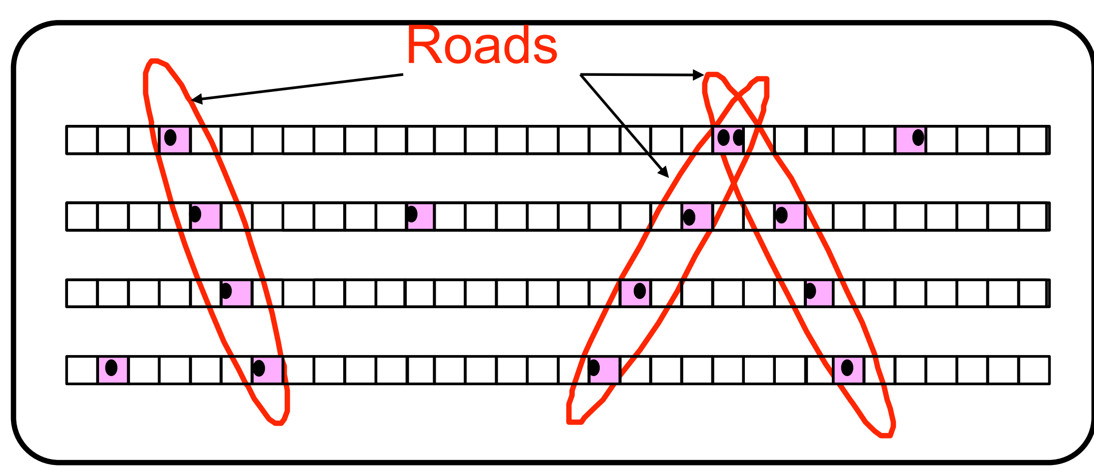
Super Bin (SB)

2. Then fit tracks inside roads. Thanks to 1st step it is much easier



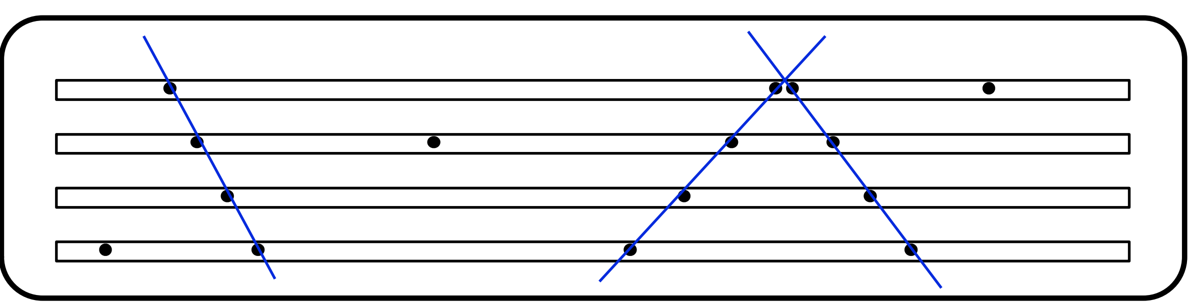
Too large AM? 2 step approach

1. Find low resolution track candidates called “roads”. Solve most of the pattern recognition



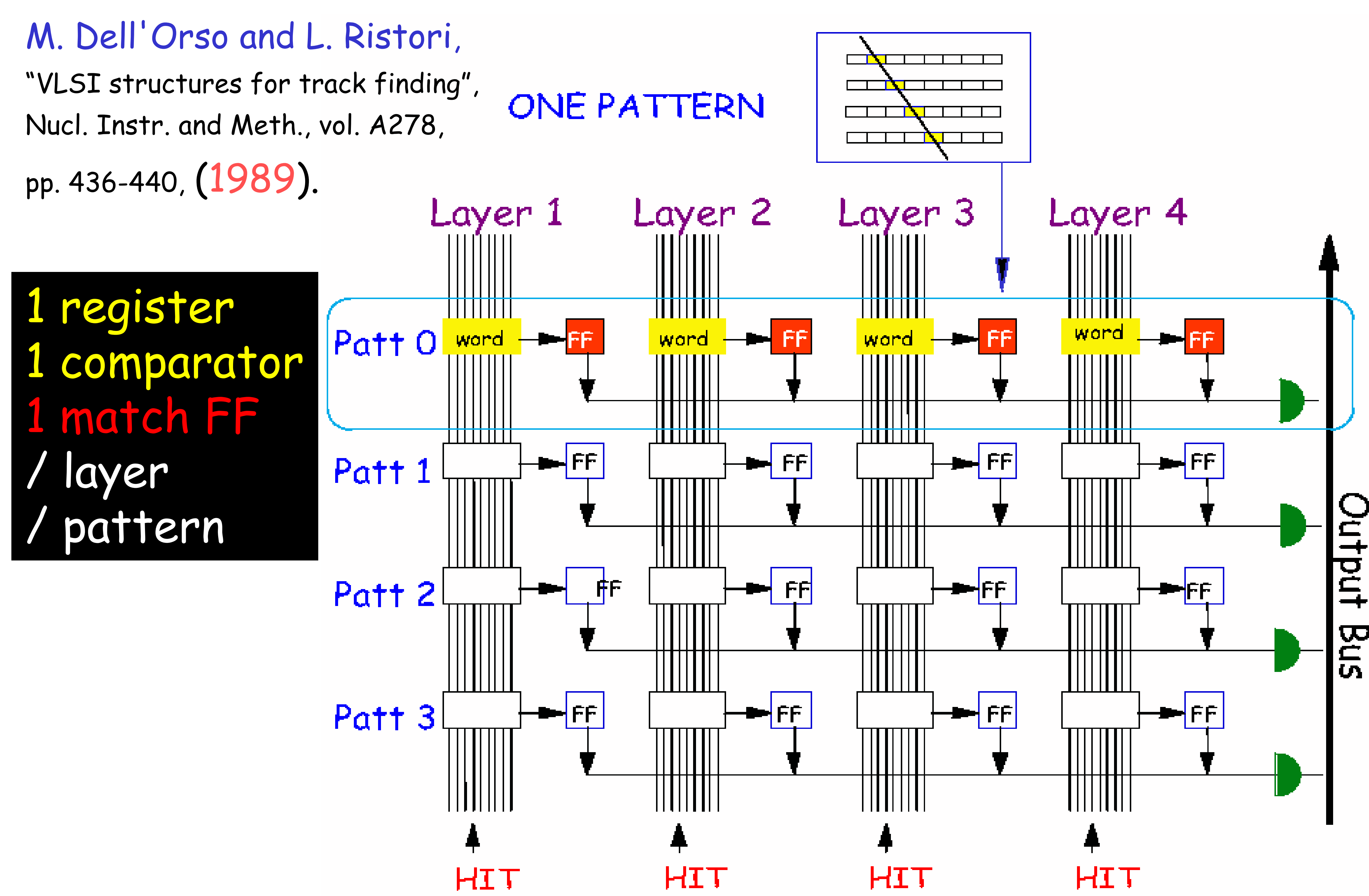
Super Bin (SB)

2. Then fit tracks inside roads. Thanks to 1st step it is much easier



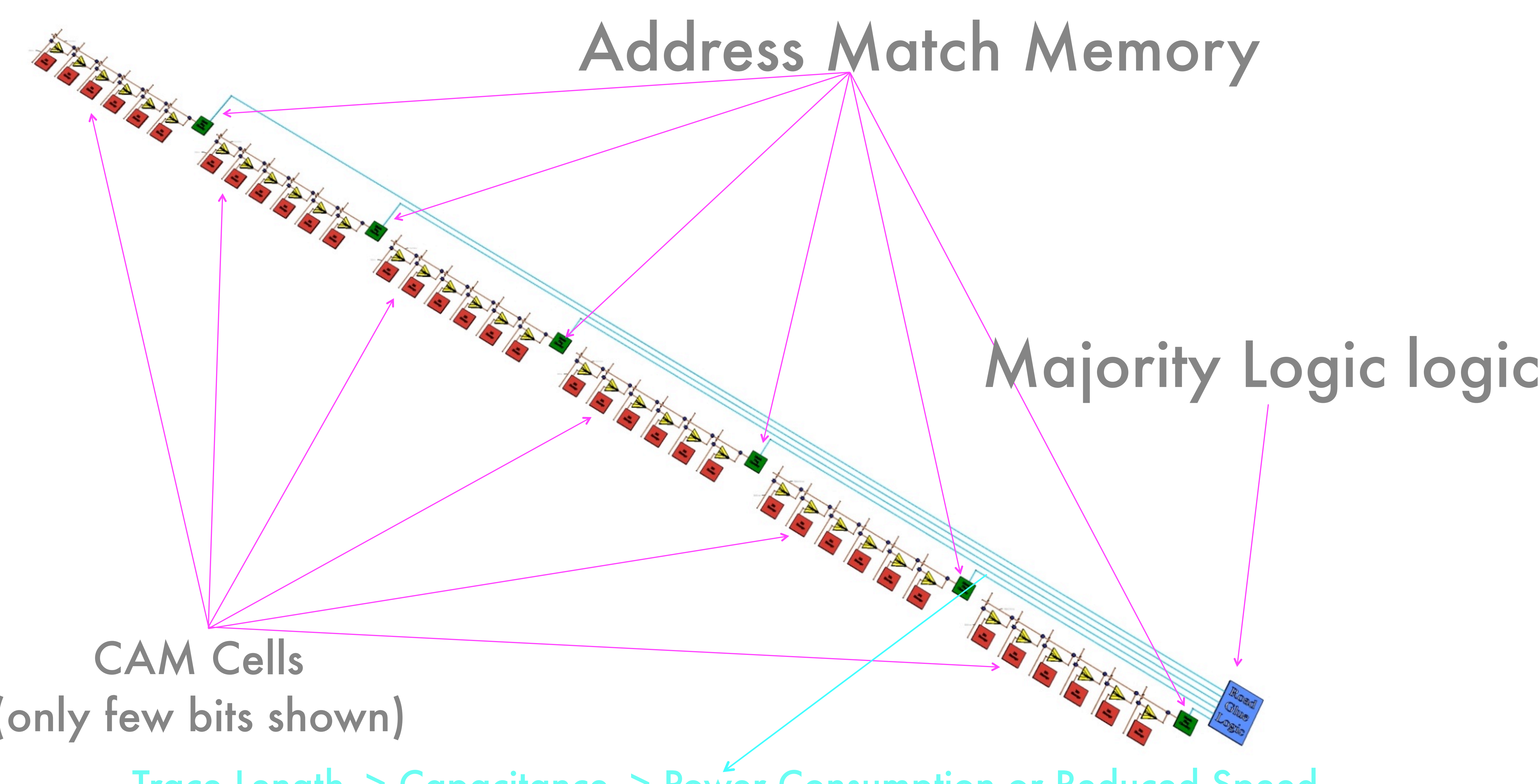
IFF smaller resolution wanted (probably not for Trigger)
OTHER functions are needed:
Hit Buffer + **Track fitter** + **Hit Finder**

Associative Memory for pattern matching



Anatomy of a PRAM

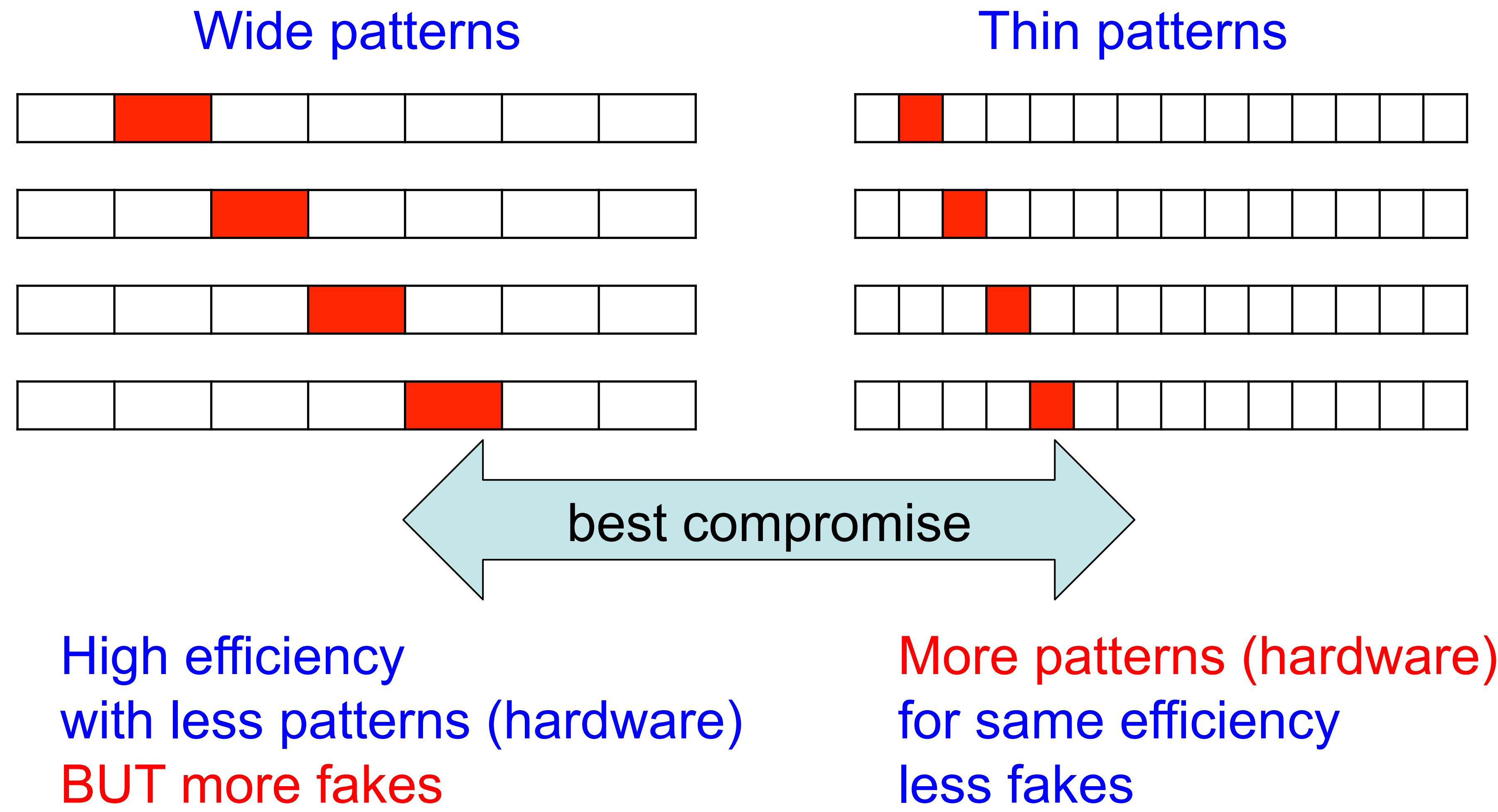
(Pattern Recognition Associative Memory)



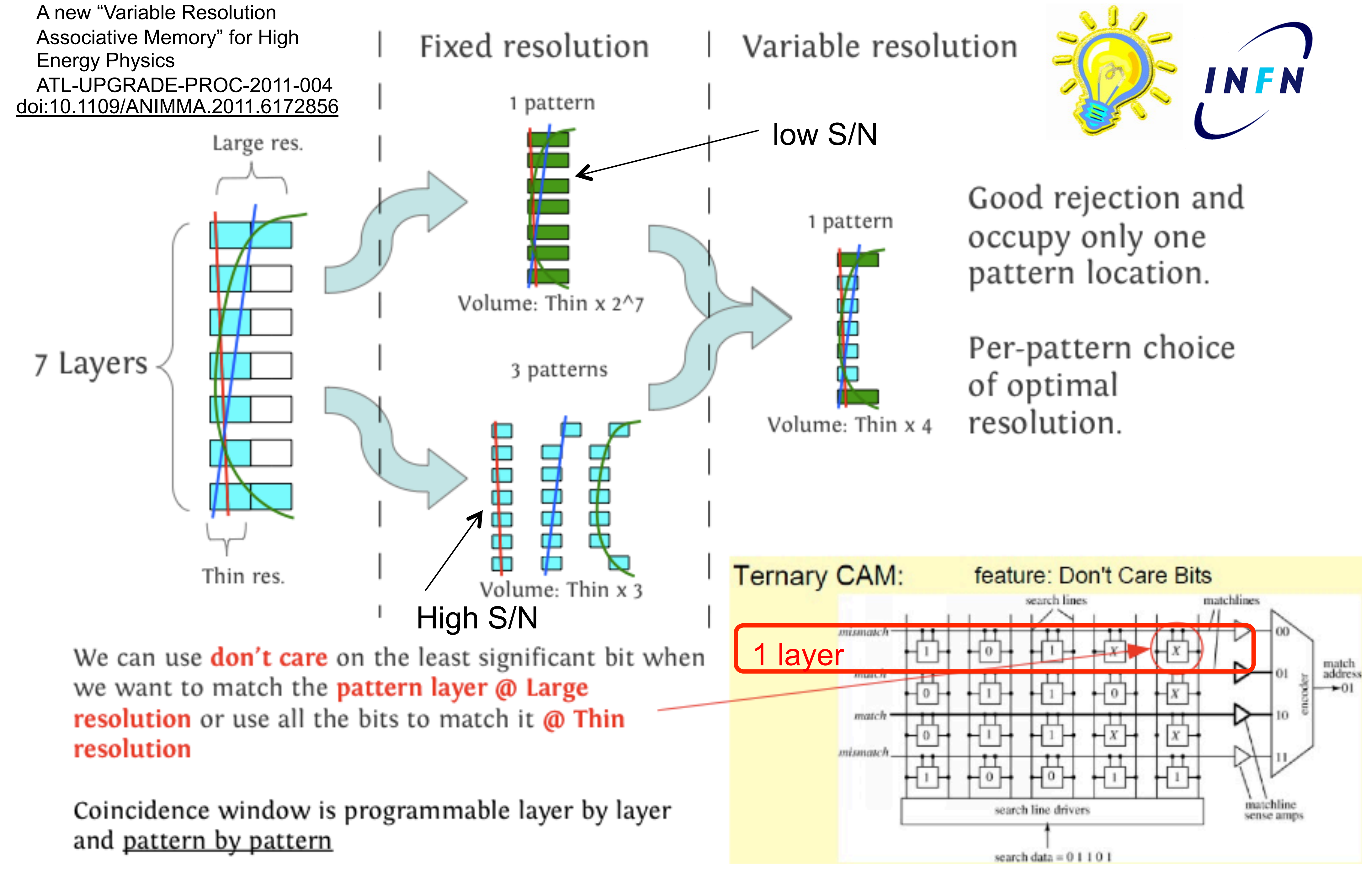
More detector layers, or more bits involved, design more spread out in 2D

→ less pattern density, higher power consumption ...

Generating the pattern bank



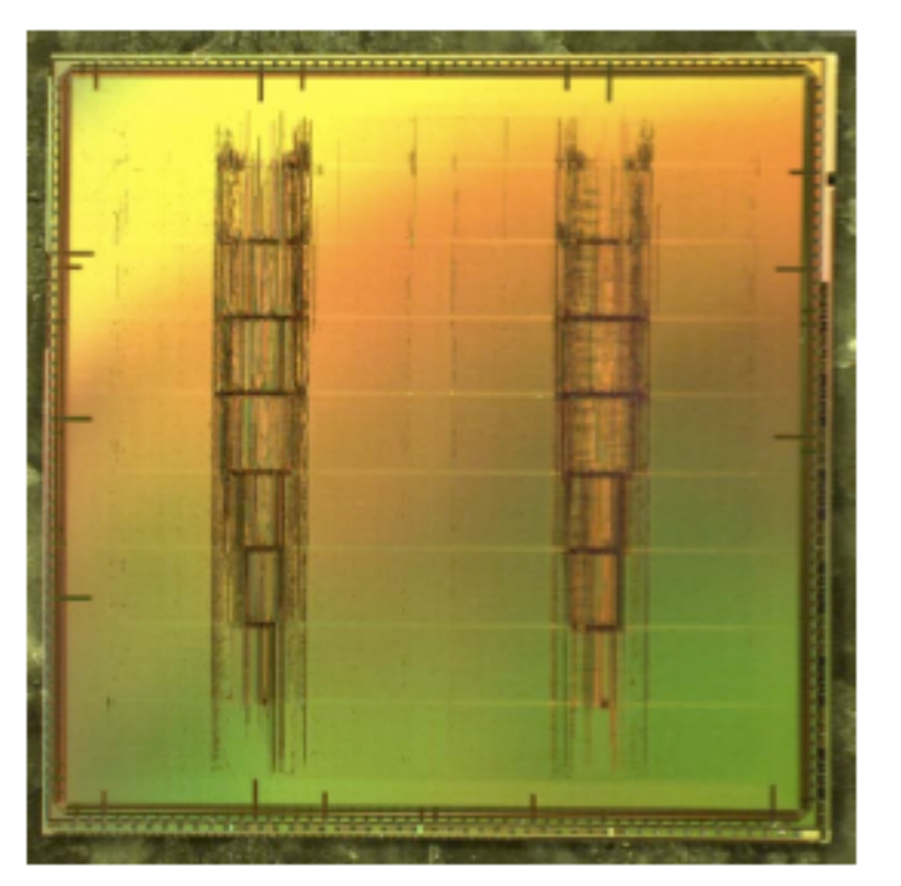
Increasing the pattern density



Associative memories evolution

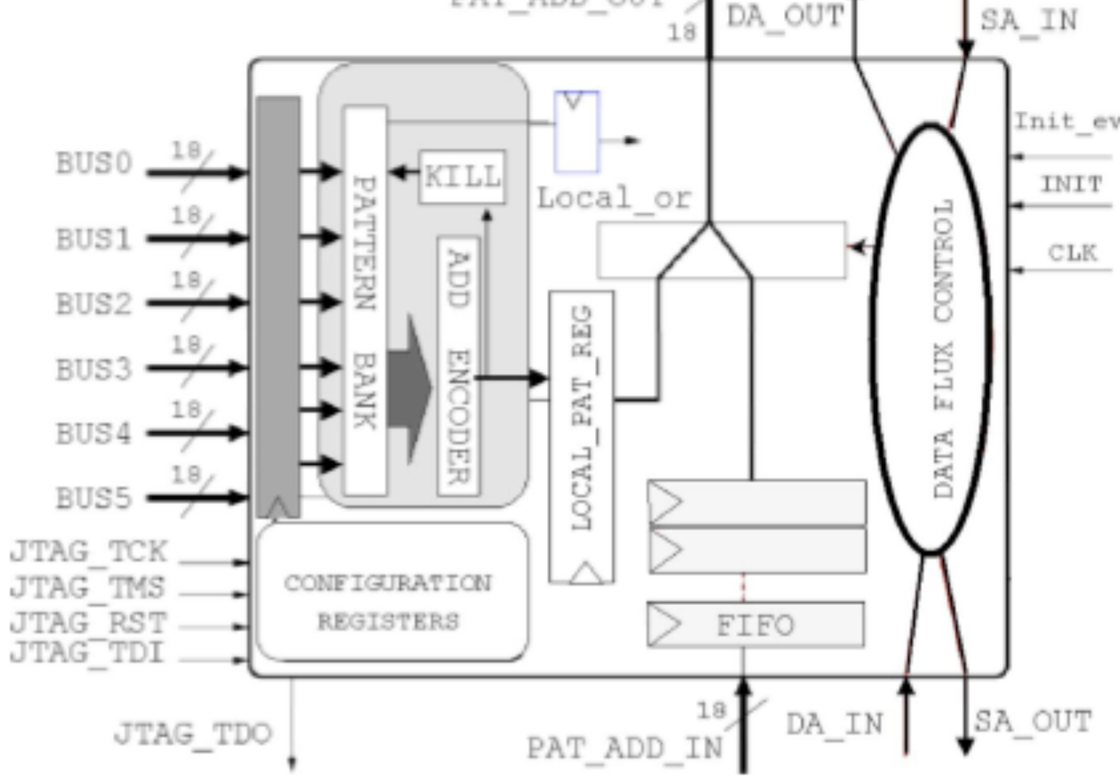
- Long history
  - 1990: Full custom VLSI chip - 0.7 μm (INFN-Pisa), 128 patterns/chip: high pattern density, not easy design
  - FPGA approach 1998: easier design but fewer density
  - A good compromise is the standard cell approach used for the SVT CDF upgrade: J. Adelman et al., Nuclear Science Symposium, 2005 IEEE, vol. 1, 2005, p. 603.
    - 0.18μm (INFN-Pisa), 5000 patterns/chip, 6 buses input lines, 50 MHz/bus, 18 bits/bus
      - produced by UMC (Taiwan) - design time ~8 months + 2 months production
  - All in all: allow to reach ~30K patterns/chip with 200 MHz/bus speed

AM03 chip details



9.8 mm

Fig. 2. Micrograph of the AMchip03 device. Four manually optimized columns of 1280 patterns each are visible. One on the left, one on the right and two in the middle. The two columns of lower density logic correspond to the interconnection and readout logic that was automatically placed. (Color version available online at http://ieeexplore.ieee.org.)



AM04 chip

	AMchip03	AMchip04	effect
Technology	180nm	65nm LP	X8 pattern density
Clock freq.	50MHz	100MHz	Faster, higher power cons.
Die size	$10x10mm^2$	$12x12mm^2$	X1.5 patterns (prototype $3.5x4mm^2$)
Core voltage	1.8V	1.2V	Lower power cons.
Core power	1.3W	2W	At 40MHz and 100MHz respec.
Full custom	No	Yes	X2 pattern density
Layers	6	8	¾ pattern density
Patterns/chip	5k	80k	8k in prototype
Ternary layers	N/A	3 to 6	Better S/N with variable resolution
Bits/layer	18	15	
Input hit b/w	4.3	12	Gbit/s
		2 event buf.	readout 1st, load 2nd event

Breaking news! Arrived and under test

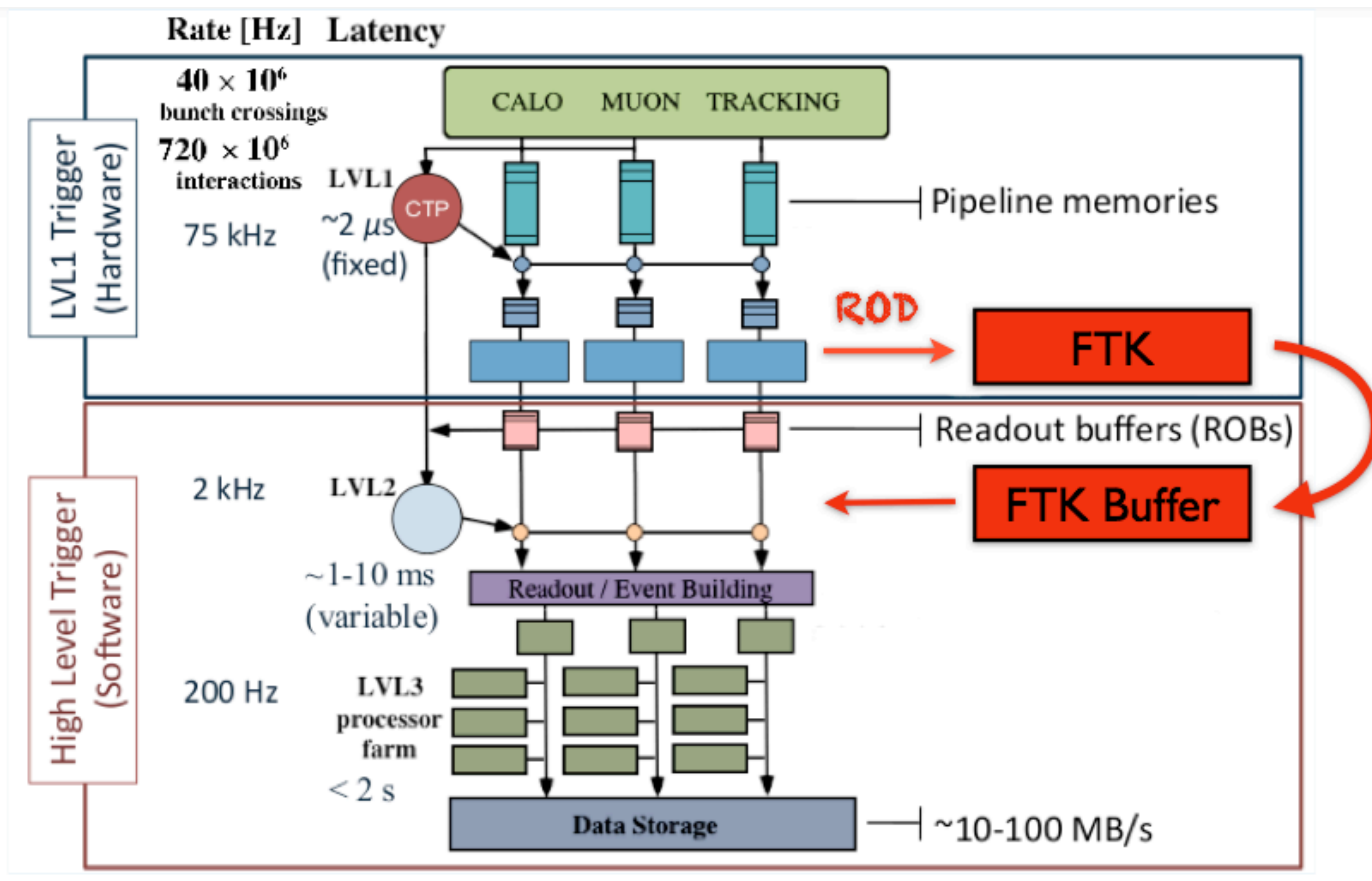
Evolution of the AM

65 nm technology provides a factor 8 → 20000 patterns/chip
Full custom cell provides at least a factor 2 → 40000 patterns/chip
8 layers instead of 12 provides a factor 1,5 → 60000 patterns/chip
1,2 x 1,2 cm^2 2D chip → 80000 patterns/chip
With a **2 D chip** we gain a factor **30**!

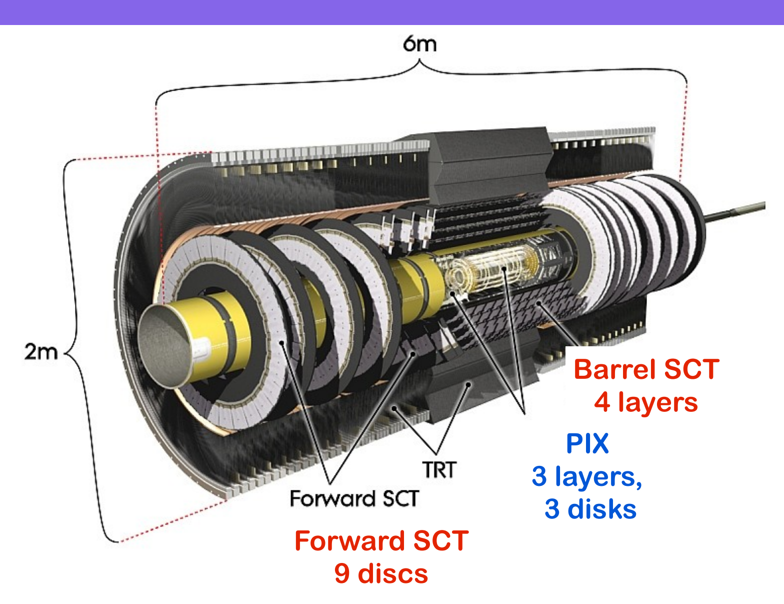
1 AMboard: 128 chips → ~10 Mpatterns per board
1 Crate: 16 AMboard → ~160 Mpatterns per crate

Current prototype under design:
65nm TSMC, 12mm^2 MPW run, 100 MHz running clock
8000 patterns/chip 8 layers each
Layer words of 12 bits + 3 ternary bits → variable resolution patterns

Usage in ATLAS @L1.5

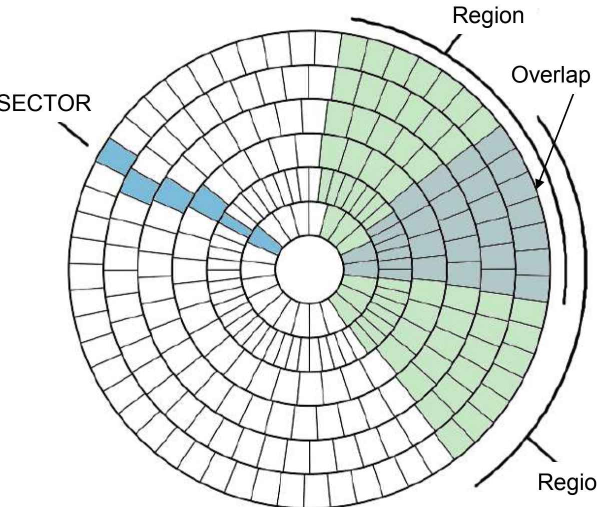


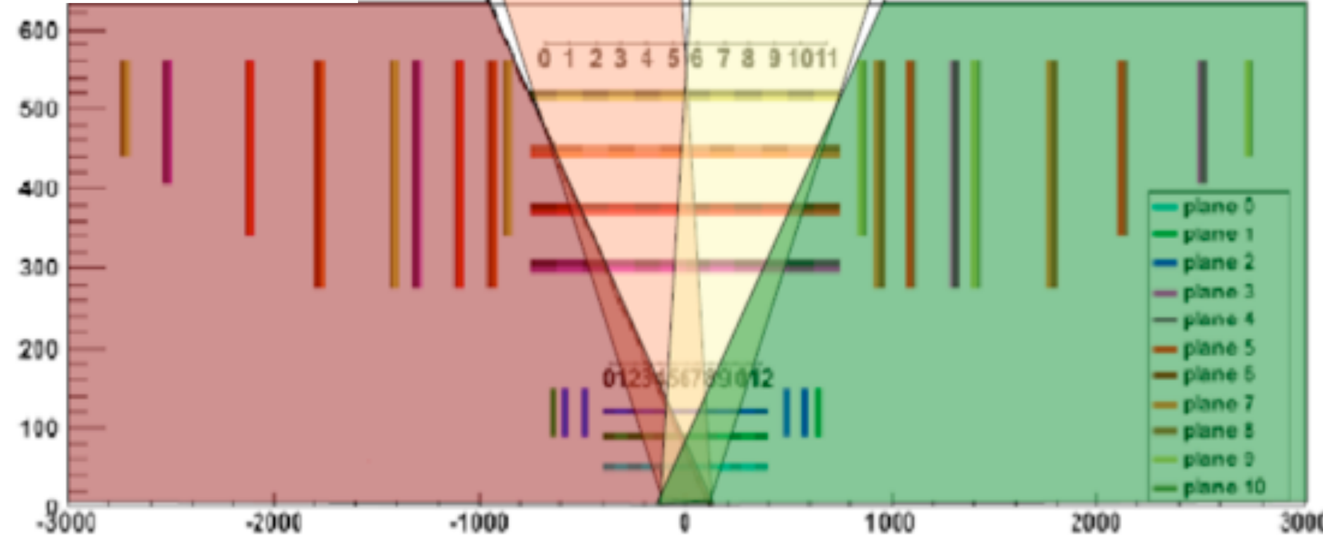
ATLAS FTK



- To deal with data flow designed as highly parallel system
  - 8 ‘core crate’ with own pattern recognition and track fitter
  - Detector subdivided in 64 trigger tower

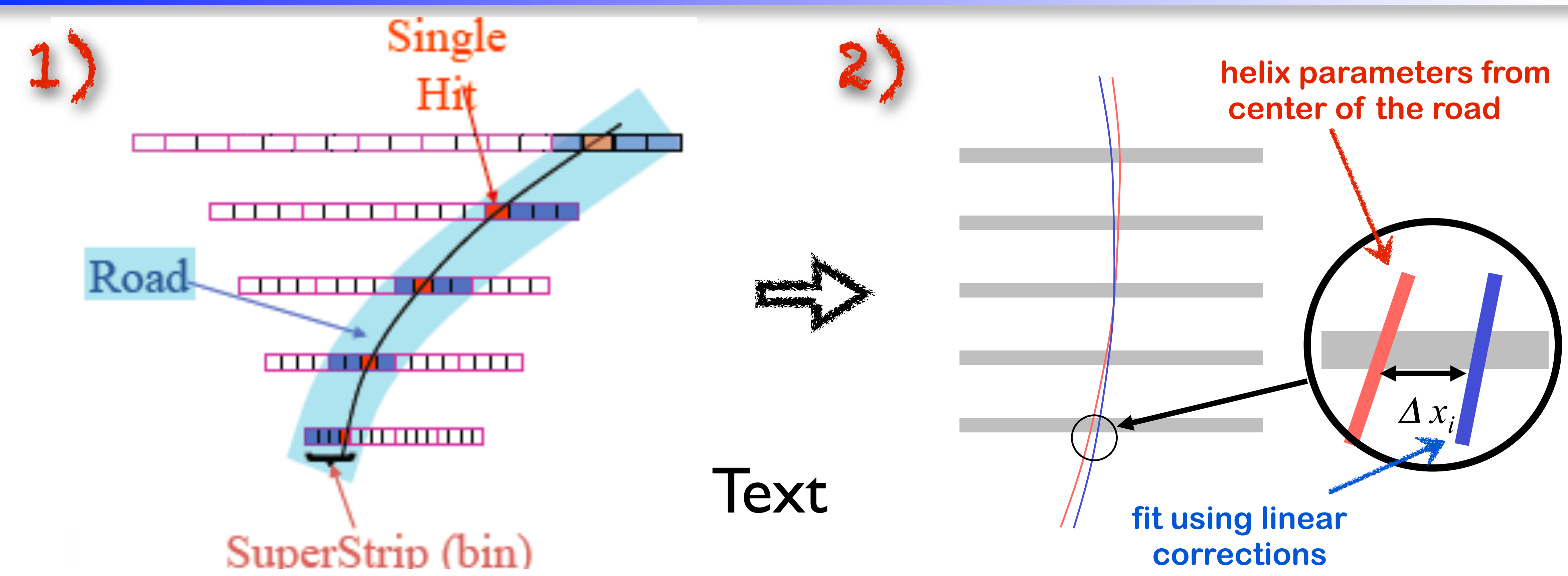
- PIX (3 layers) & SCT (4 double layers)
- Fit posses combinatorics problem, executed in two sequential steps:
  - Use 8 layer for patter recognition and 8 layer fit
  - Refit track found using all 11 layer





FTK working principle

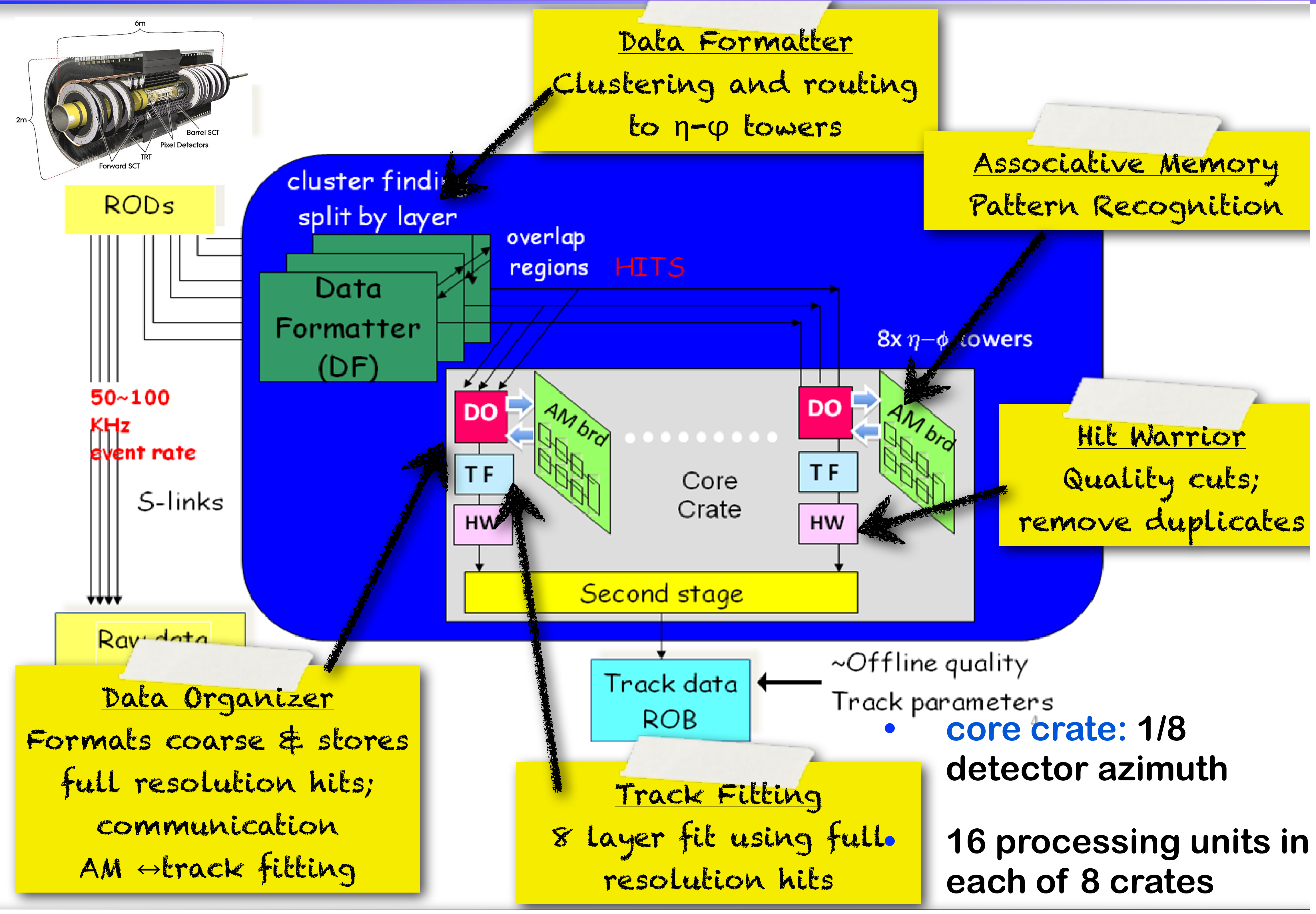
8 Layer Tracking Step



- Find low resolution track candidates (roads)
- PIX 3 + SCT 4 axial + (SCT 1 stereo || IBL) layer
- Use parallelism in Associative Memory chips

- Use full resolution hits in 8 layers
- Obtain high resolution helix parameter from road

System sketch



- **core crate:** 1/8 detector azimuth

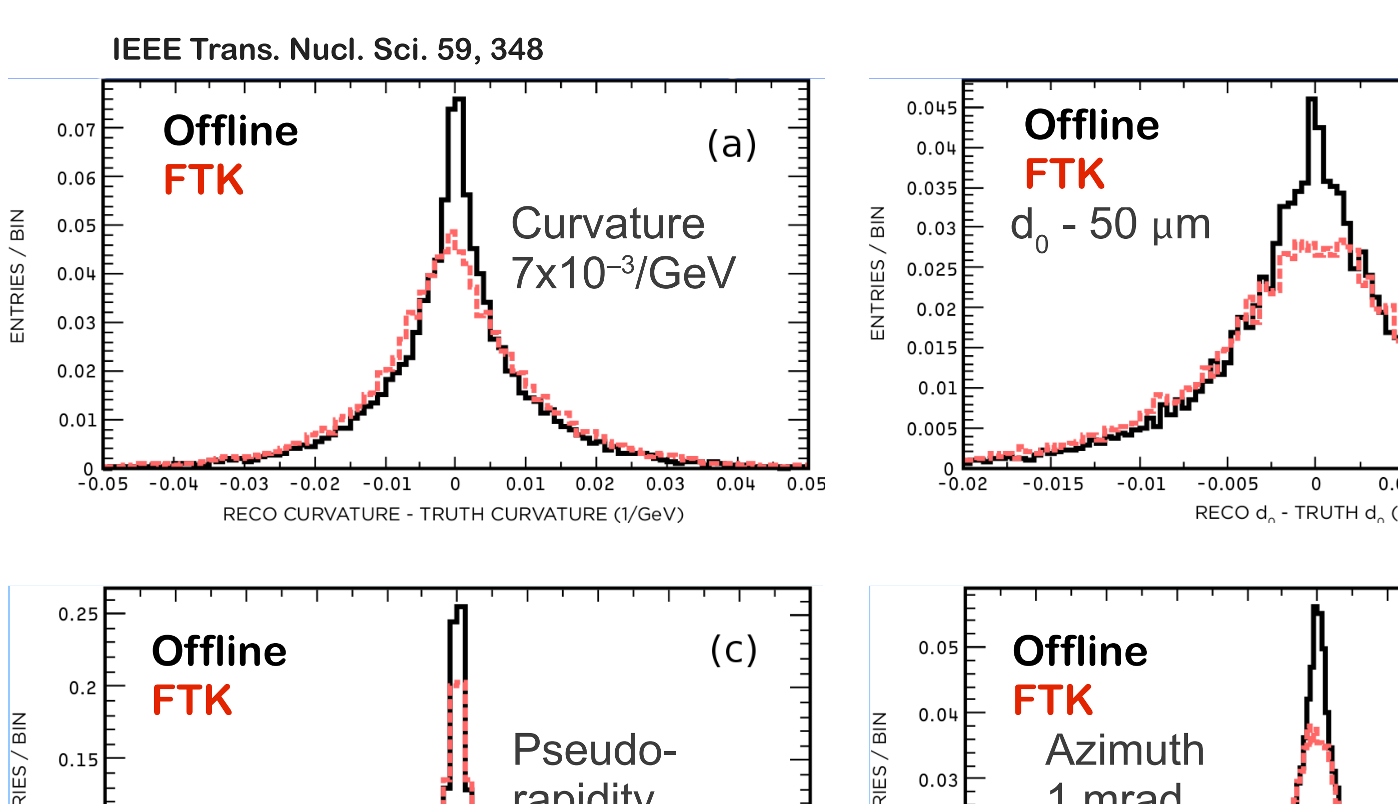
16 processing units in each of 8 crates

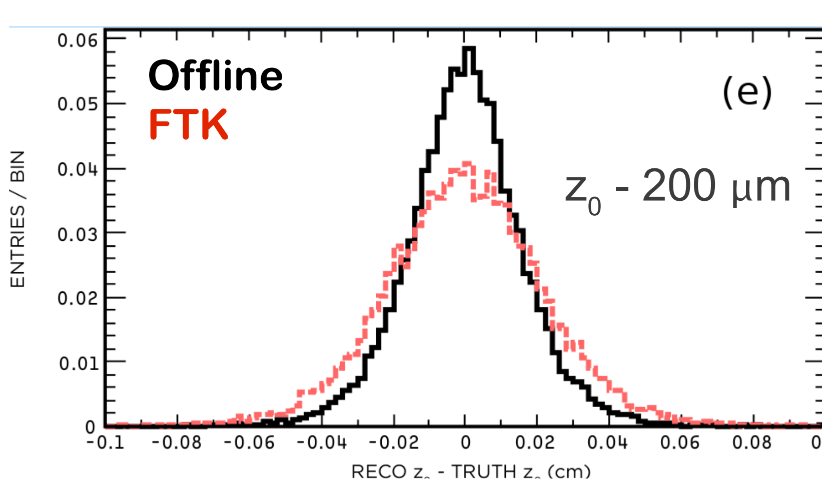
FTK system needs

- Predictions for 40-pileup events (~2015)
  - 1000-2000 clusters per core crate per layer
  - 20k roads per core crate using don't-care
  - 100k fit combinations per core crate

FTK performance

Performance



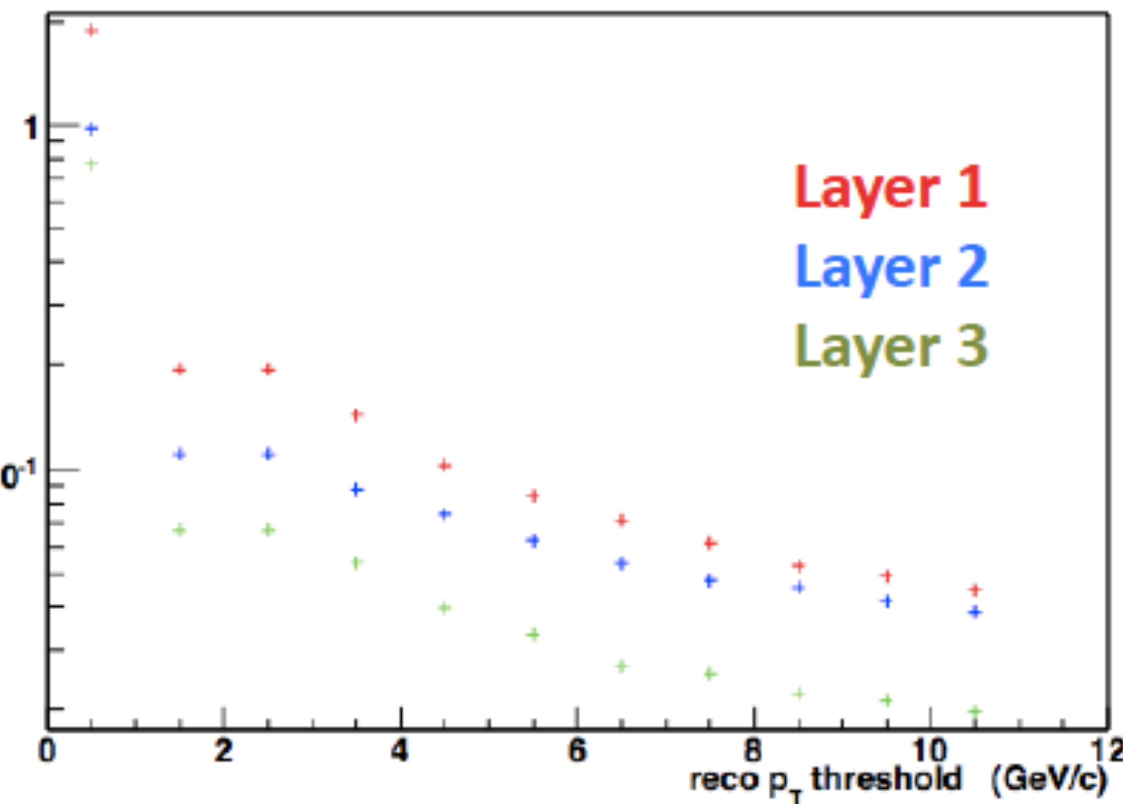


Timing to reconstruct all tracks with p_T>1 GeV: ~20 msec/ROI

- **Tracks reconstructed in the barrel p_T >1GeV**
- **Based on single muon events**
- **~80-90% efficiency with almost offline resolution**

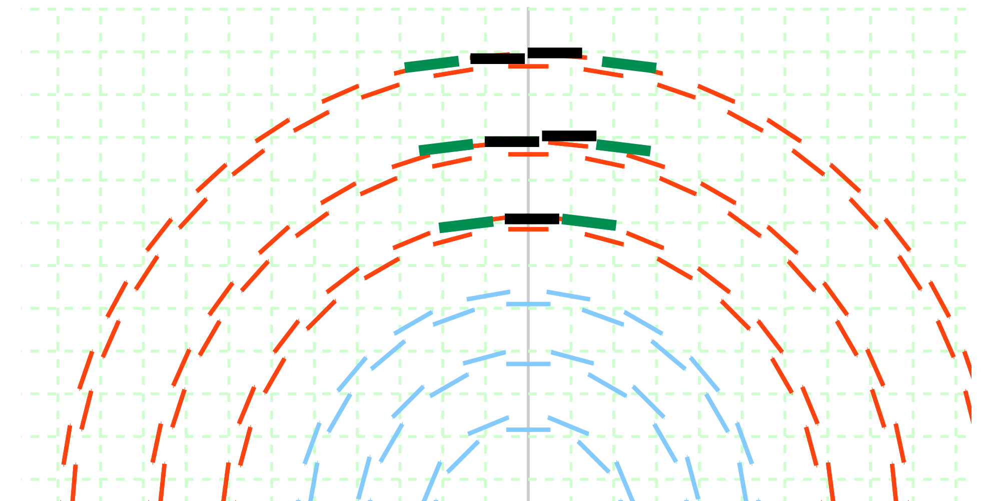
CMS L1 Track Trigger

- **Stub/ladder data rate (layers R>50 cm)**
- Outer layers with 2S pT-modules: (stubs p_T>2 GeV) ~100-200 kHz/cm^2
- Module area: 86.64 cm^2 : ~8.664 MHz/40 MHz = 0.22-0.45 stubs/module/bx - on average
- 12 modules per ladder: layer data rate~100-200 MHz (2.5-5 stubs/ladder/bx)
- stub info used for patter recognition:14 bits (6 bits r-ϕ (1 mm resolution), 1 bit r-z, 4 module #, 3 time stamp). Eventually full granularity sent to the AM board, but use it smartly. (see next transparencies)
  - If 20 bits sent for full resolution, expected data rate per layer~up to 4 Gbps
  - Given that a AM chip inputs at 16 bits x 100 MHz/ layer and that has intrinsic latency, a single AM-board cannot process in time a single event. Hence need a switch that distributes events to several AM-boards in parallel.

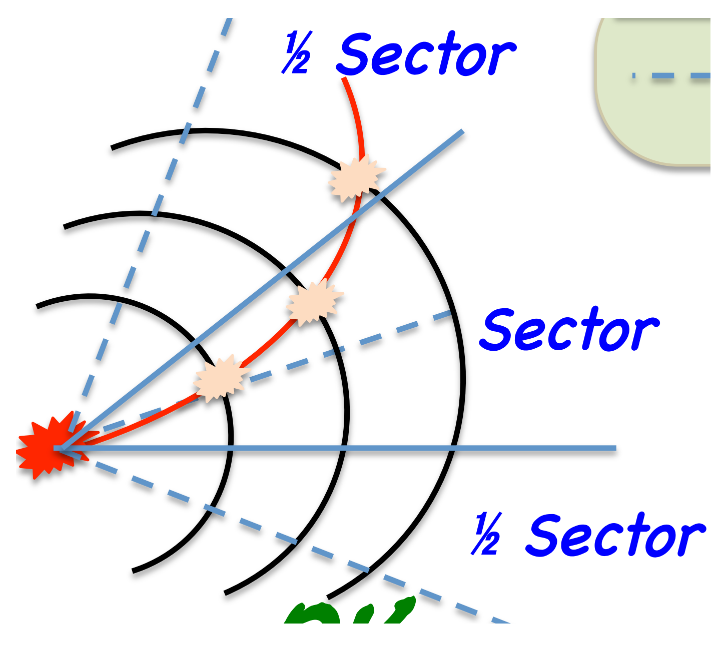


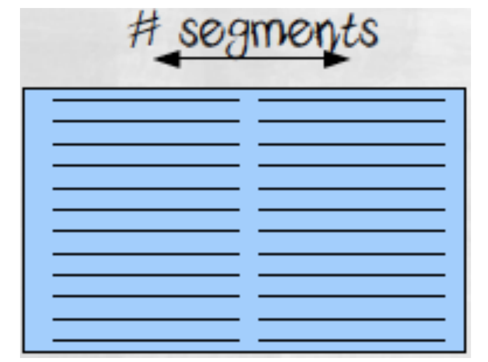
Trigger sectors dimensioning

- **Full GEANT4 simulation**
  - Using 3 layers from 50 to 108 cm radius
    - 39 azimuthal trigger sectors allow ~2 GeV p_T cutoff
      - +z and -z sides lumped together
      - adjacent (half) sectors sent to the same AM



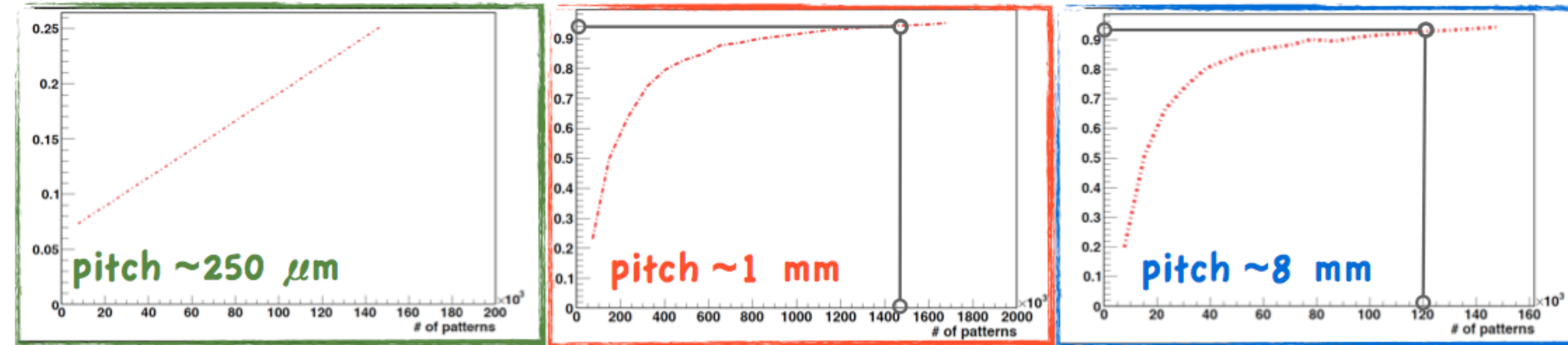
- **Super-Strip**
  - **contains the information of a stub in a sector**
  - **15 bits**
    - **5 (z module position) +1 (phi module position)+ 6 (cluster position with a precision of ~1 mm) +1 (strip segment) + 2 (p_T range (2,5,10,infty) from the stub)**





Encoding

- **Pattern**
  - for a N-layers layout a word of N-Super Strips information
    - SS_k Super Strip information for layer k. Pattern = <SS_1,SS_2,SS_3>
- **Bank Coverage**
  - No. stored patterns for which the bank reaches 90% efficiency of reconstructed patterns
  - Generate single muons with $2<p_T<60$ GeV in the barrel acceptance
    - studied with different pitch resolution (250 µm, 1 mm and 8 mm) ↔(17, 15, 12 bits SS_k)
      - ➨ Almost linear relationship with pitch size: 1.5 M patterns/sector for 1 mm resolution, 120K patterns for 8 mm pitch.
      - ➨ Needs a compromise between the number of stored patterns and fake rate.



Preliminary results

Single μ efficiency

Roads/bunch crossing
200 pileup Minimum Bias events

pattern recognition through AM (Pattern Bank)$_{(2\ GeV/c)}$ matching Efficiencies & Fake Rates Vs. p_T for MinBias tracks

hadron efficiency (Δ)
and fake rate (O) vs track p_T
200 pileup Minimum Bias events

The fake rate is mainly due to the cluster splitting. Will improve with a better clusterizer

Advanced AM

- Main limitations of AM approach for L1 track trigger
  - pattern bank density
  - latency limitations

VIPRAM (Vertically Integrated Pattern Recognition Memory)

VIPRAM concept (developed at Fermilab):
http://hep.uchicago.edu/~thliu/projects/VIPRAM/TIPP2011_VIPRAM_Paper.V11.preprint.pdf

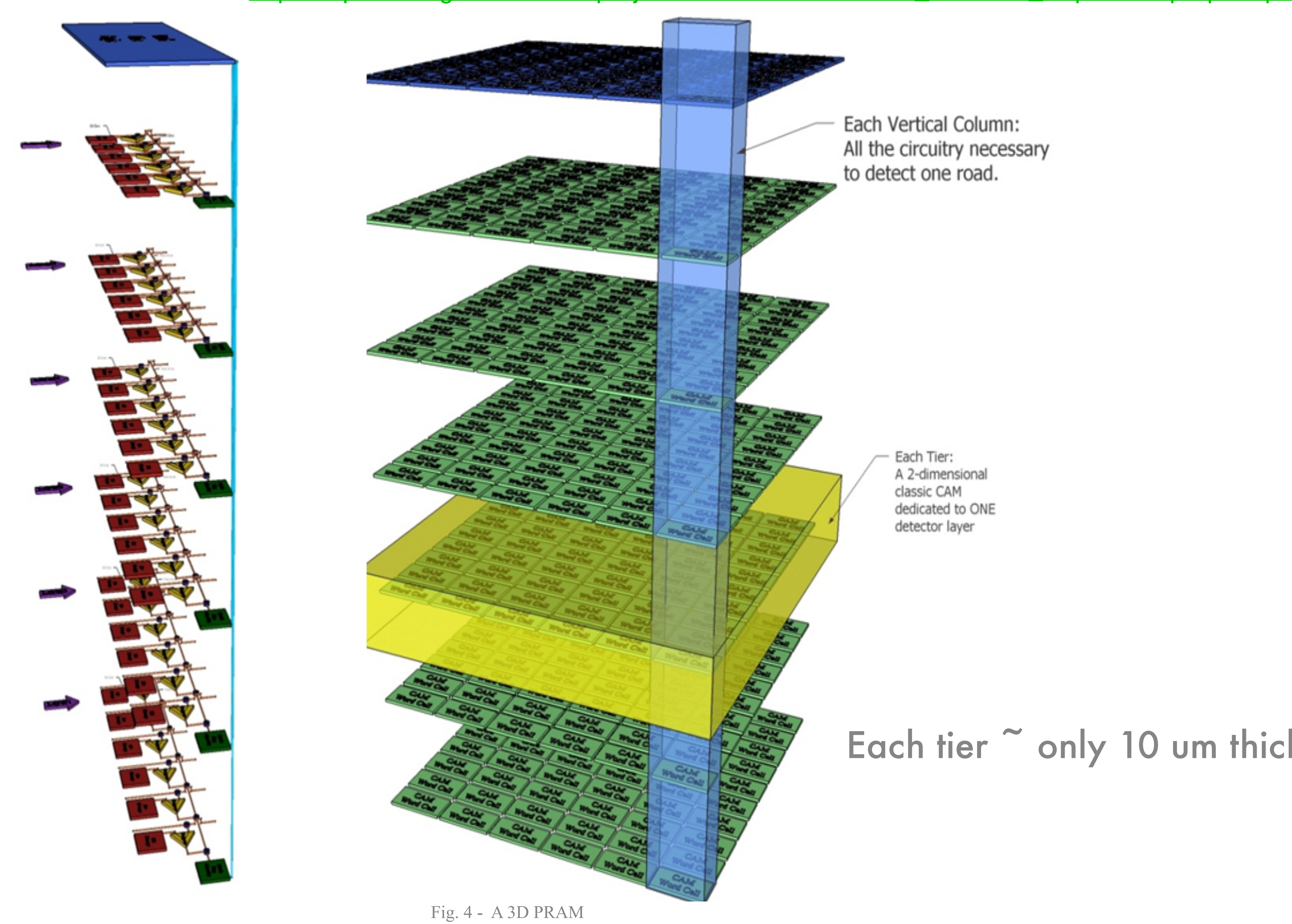
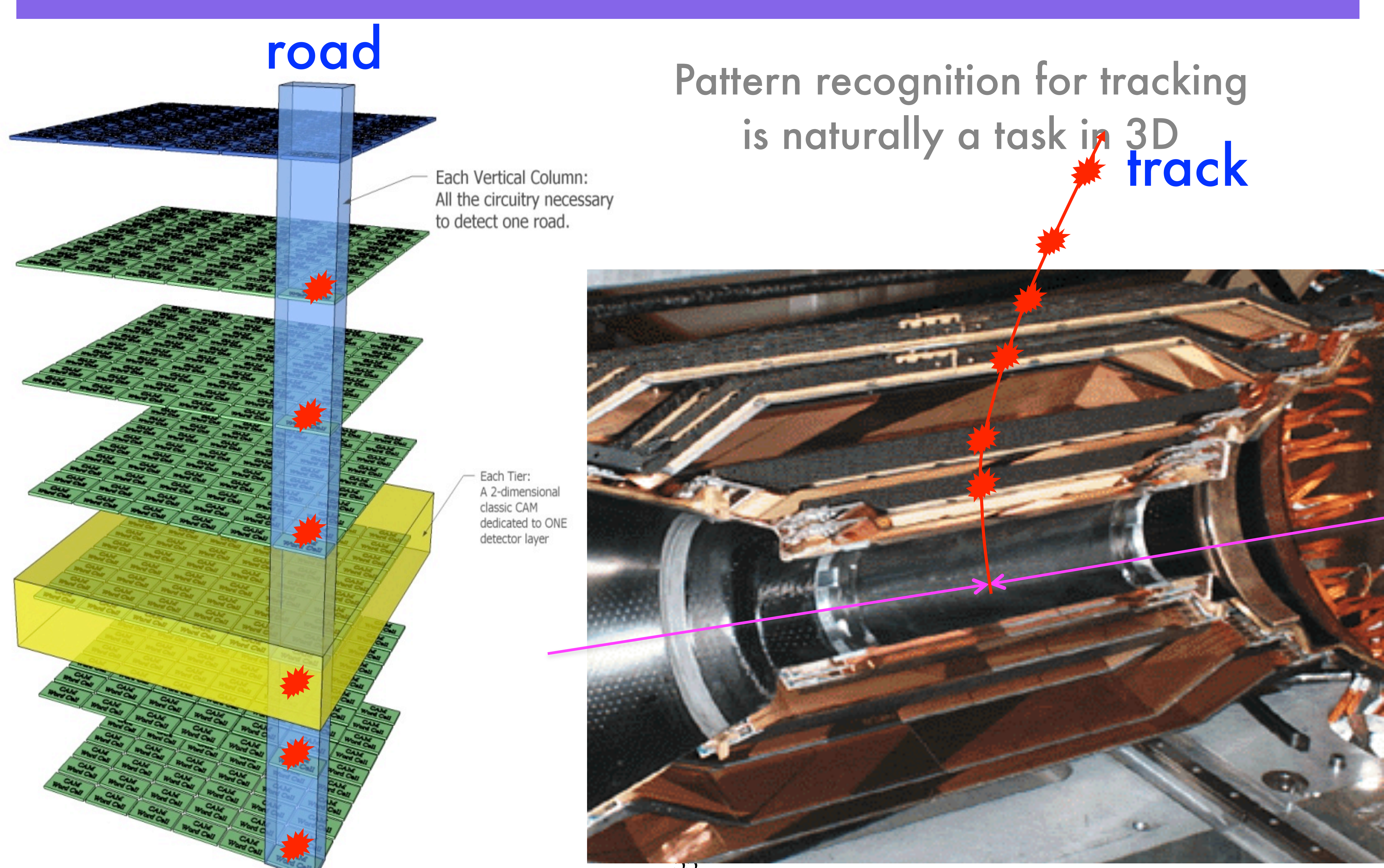


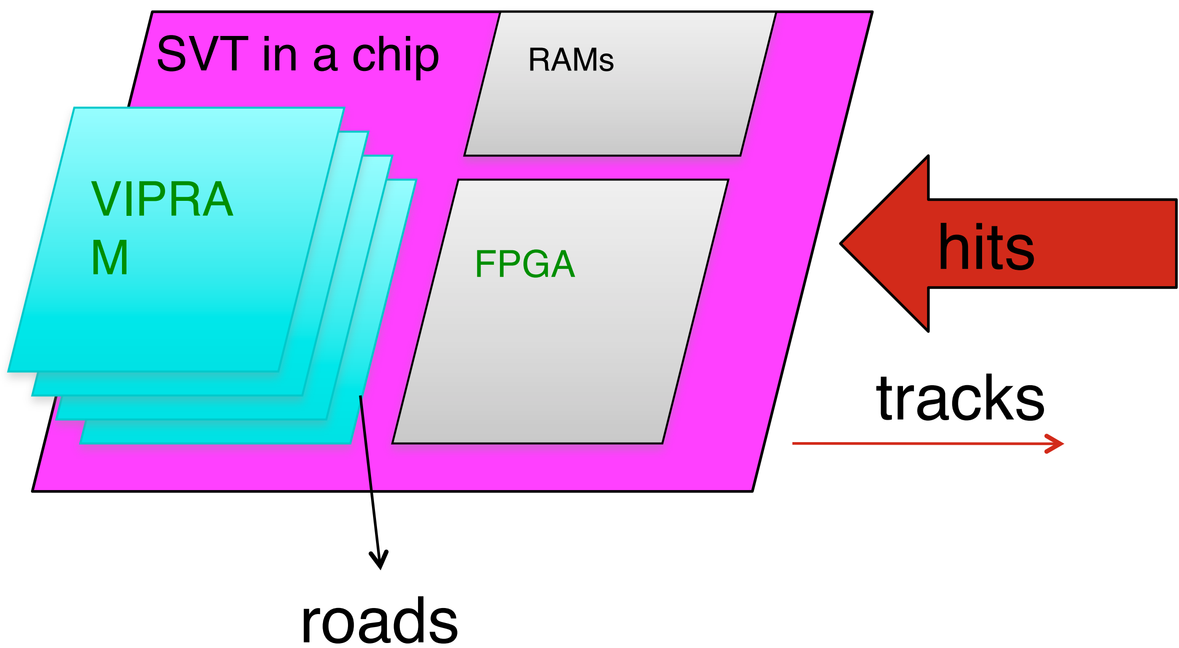
Fig. 4 - A 3D PRAM

VIPRAM



Further evolution of AM

- ~ 500K patterns/cm **2
- Running with > 100 MHz input rate
- N CAM tiers + Control tier
- integrated with FPGA/RAM
  (*general purpose pattern recognition*)



Feasible by 2020?

SVT in one chip?: 2nd phase of VIPRAM project

RAMs

AM

FPGA

2 meters

0,1 fan-out fan-in 10,11

2,3 4,5 6,7 8,9

Original SVT system had 384K patterns total
Aim to reach ~500K per cm**2 for VIPRAM …

References

- Where to find this (and more) material
  - WIT2010 and WIT2012 Workshops
  - ACES Workshops
  - TIPP Conference series
  - TWEPP Conference series