

# Trigger for LHC Run II and beyond

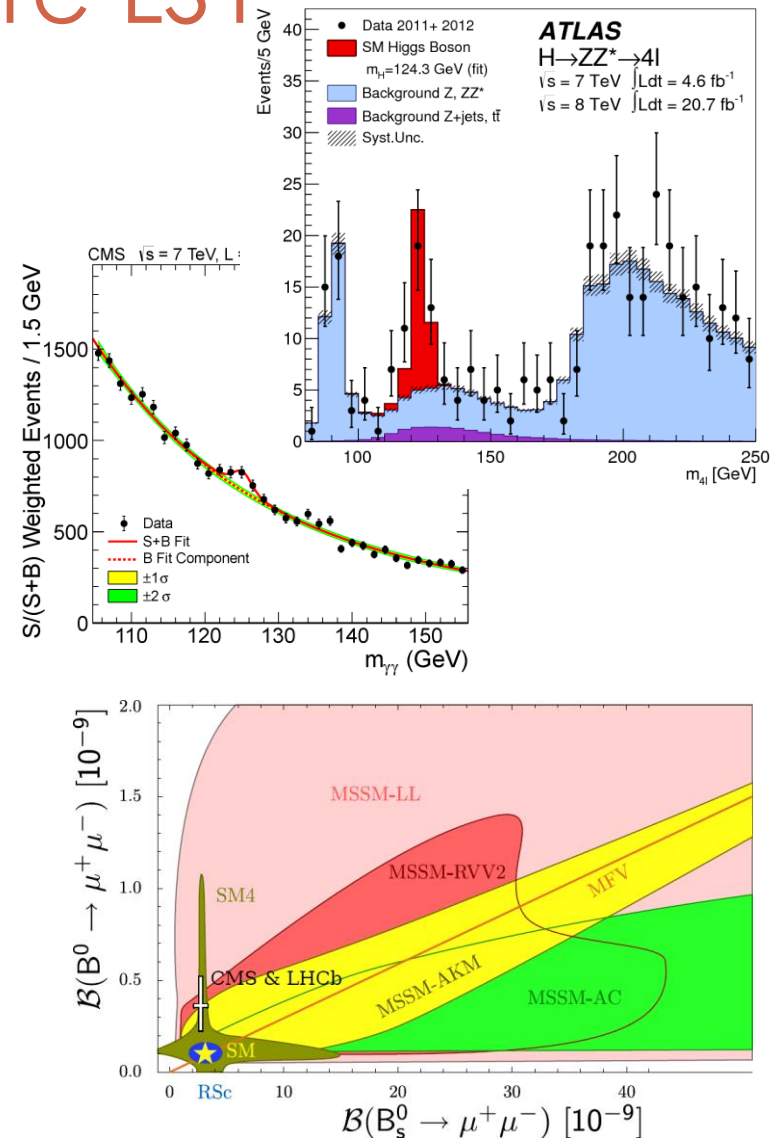
G. Volpi

Univ and INFN Pisa (FTK IAPP)



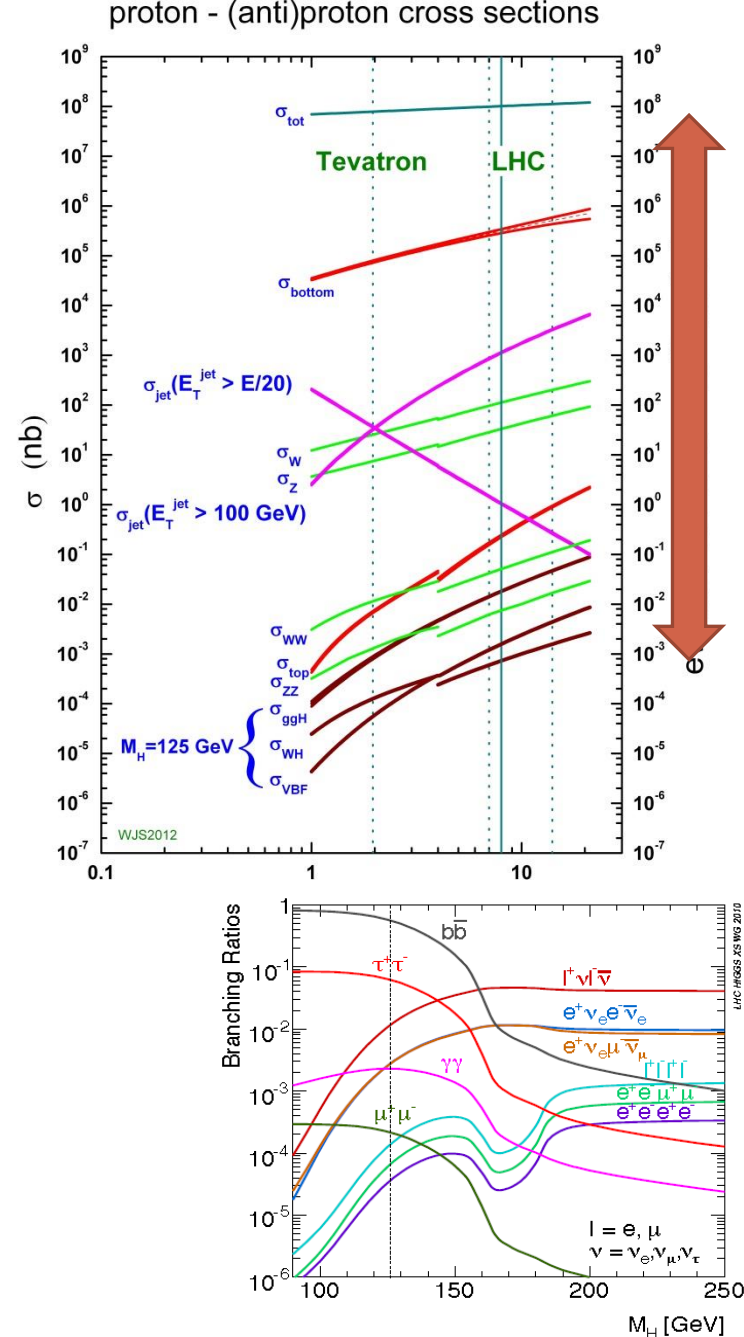
# LHC Experiments before LS1

- The LHC Run I was a tremendous success
  - The Higgs boson has been found
  - Beyond standard model parameters scanned with unprecedented power
  - World best results in challenging environment as flavor physics
- Data collection was extremely successful
  - All the experiments had a very a efficiency, better than 90%
- The experiments exceeded the design in many parameters
  - 50 ns bunch spacing made trigger selections for all the experiments more difficult than expected during the design



# The role of the trigger

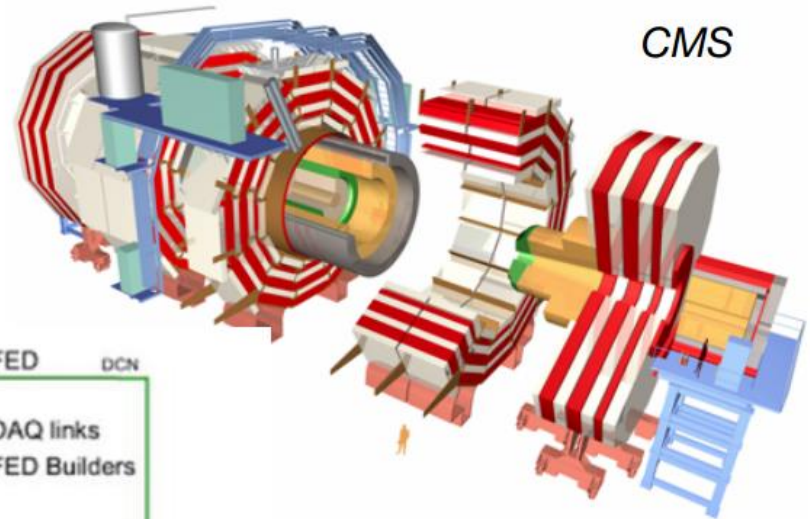
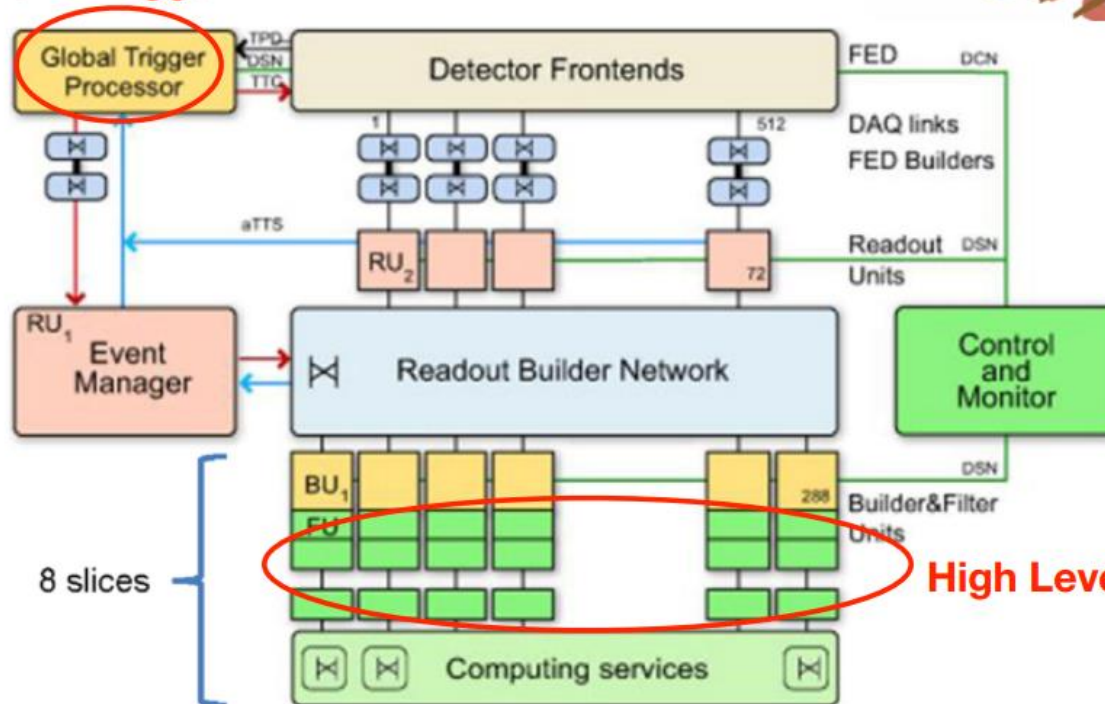
- High efficiency for rare final states
  - The trigger selects  $1/10^5$  events for permanent storage
  - Interesting events are buried in an enormous background of well-known physics
- Selections are complicated by the presence of multiple interactions per bunch-crossing
- Important to have a flexible DAQ system
  - All the experiments are designed by multi-level systems
  - Combination of custom hardware and commercial computing device



# CMS Trigger

*Two stage trigger architecture*

## Level 1 Trigger

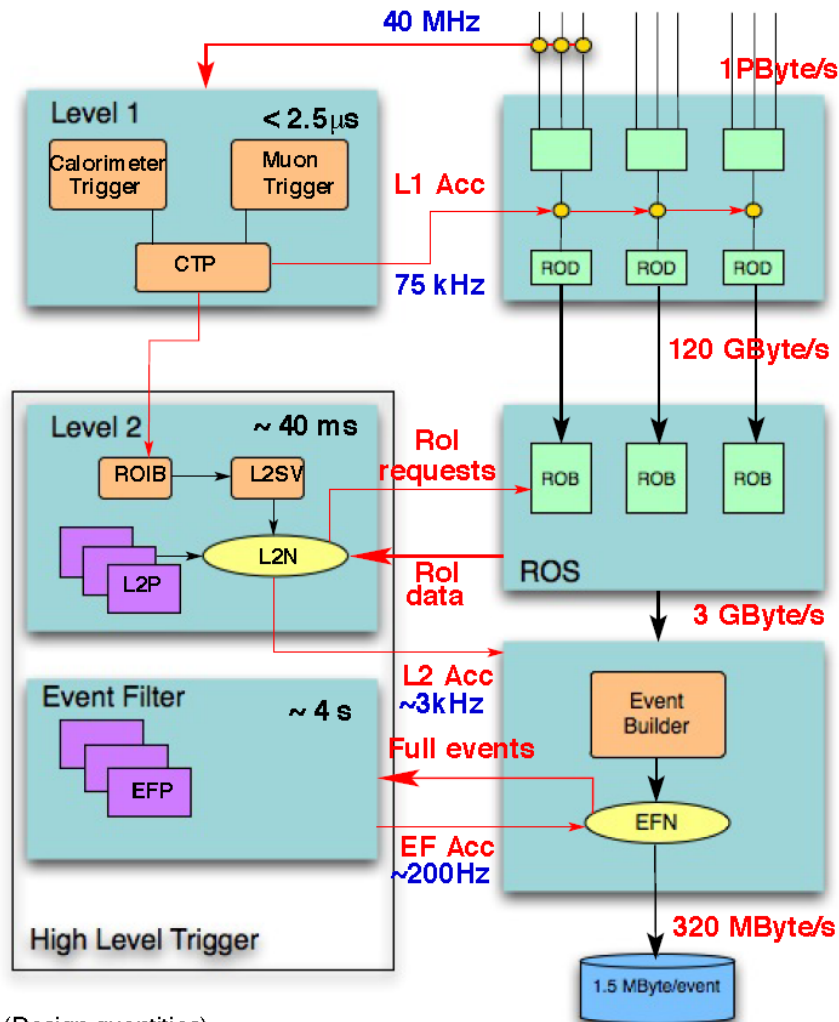


CMS

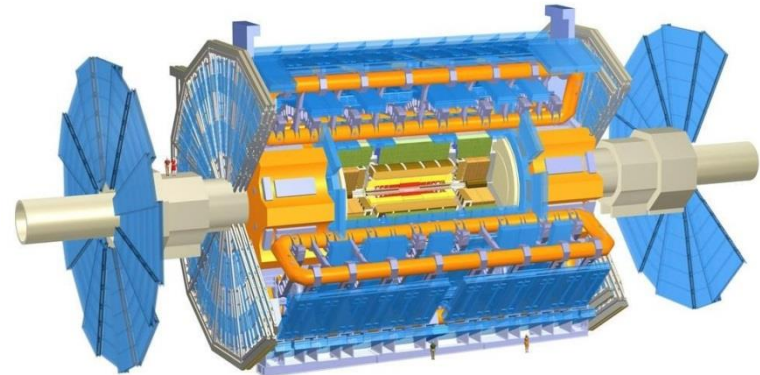
CMS has hardware Level-1 designed to work at low latency. The final selection is based on the use of a commercial CPU farm.

## High Level Trigger

# ATLAS Trigger in 2012

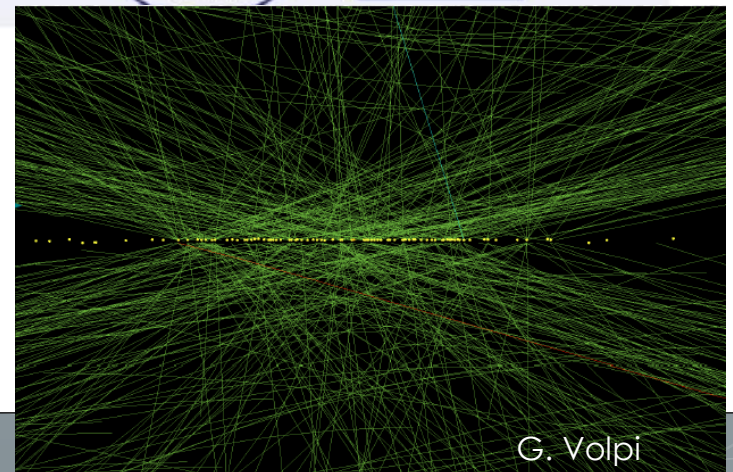
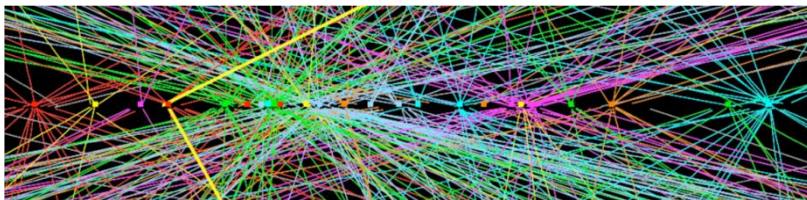
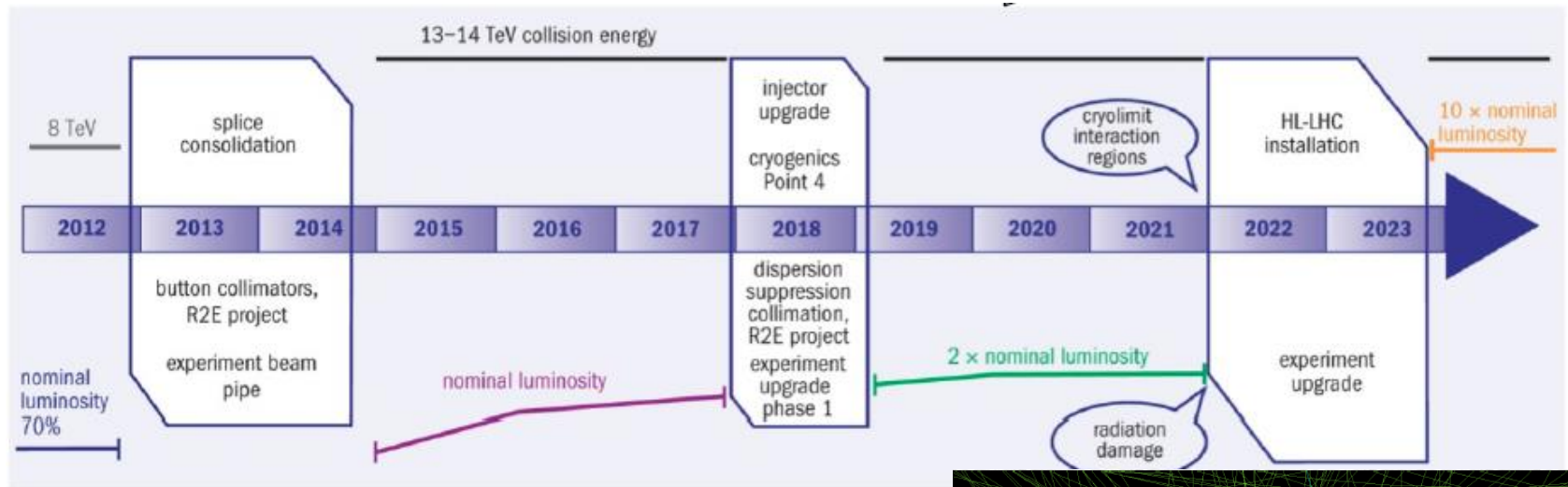
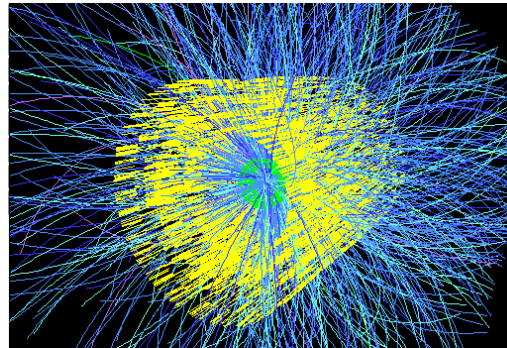


(Design quantities)



- ATLAS used a 3-level trigger scheme
- Level1 based on custom hardware
- Level2 and EF based on commercial CPU farms
  - The balance between the two farms was dynamic





# Why the upgrade? ATLAS example

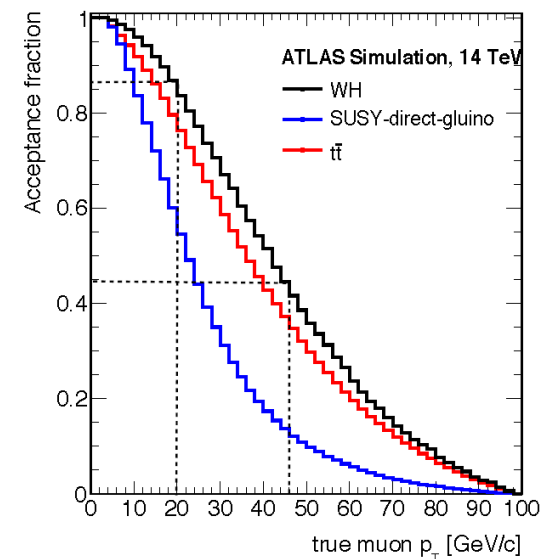
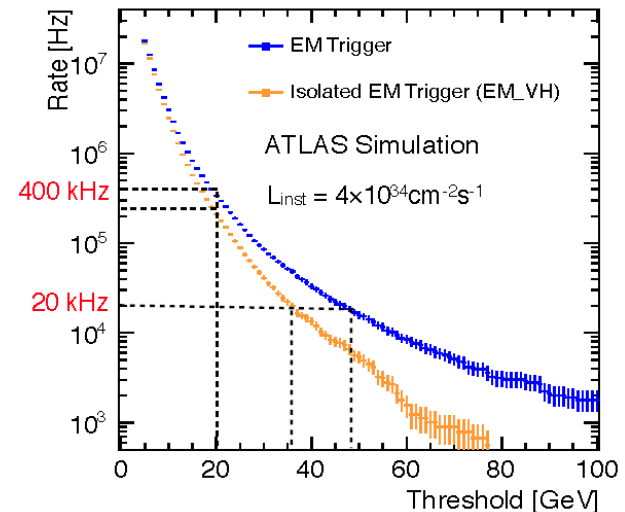
Run I menu @ $2 \times 10^{34}$			Run II menu			Run III menu		
	Offline Threshold $p_T$ [GeV]	Rate [kHz]		Offline Threshold $p_T$ [GeV]	Rate [kHz]		Offline Threshold $p_T$ [GeV]	Rate [kHz]
EM18VH	25	130	EM30VHI	38	14	EM25VHR	32	14
EM30	37	61	EM80	100	2.5	EM80	100	2.5
2EM10	$2 \times 17$	168	2EM15VHI	$2 \times 22$	2.9	2EM12VHR	$2 \times 19$	5.0
<b>EM Tot</b>		<b>270</b>	<b>EM Tot</b>		<b>18</b>	<b>EM Tot</b>		<b>20</b>
MU15	25	150	MU20	25	28	MU20	25	15
2MU10	$2 \times 12$	14	2MU11	$2 \times 12$	4.0	2MU11	$2 \times 12$	4.0
<b>Muon Total</b>		<b>164</b>	<b>Muon Total</b>		<b>32</b>	<b>Muon Total</b>		<b>19</b>

From ATLAS TDAQ Phase 1 Upgrade TDR

- ATLAS has projected the 2012 rates in the Run II conditions
  - The most used triggers would exceed the rate limits already at Level-1
  - New selections and hardware is required to mitigate the effect on the most important trigger selection

# High Luminosity LHC (>2022)

- The conditions will be even worse for the HL-LHC period
  - The number **of multiple collisions will be 140 or more**
- Single object selection are at serious risk but are still important
  - Increase the energy thresholds will not be a good solution
    - Precision measurements and beyond SM scenarios less effective
- The detectors and DAQ systems will need to mitigate the effect of multiple interaction
  - New detectors and strategies need to cope with extremely dense events

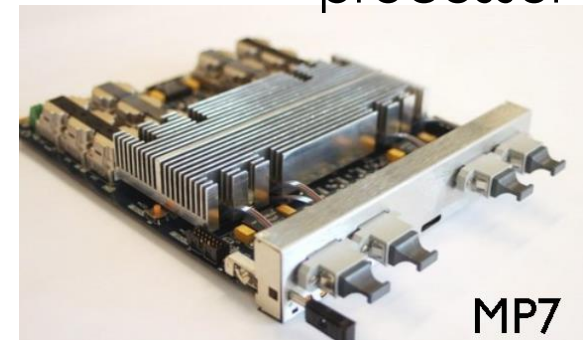




# LHC experiments upgrades in a nutshell

- All LHC experiments will be improving the DAQ infrastructures
  - Impossible to cover all the cases
- Improvements offered by more modern devices and standards
  - Newer CPUs, better FPGAs and improved VLSI fabrication processes
    - More power can be bought at reasonable price
  - **Difficult to cover all the improvements**
- In some cases a completely different approach is necessary
  - Trigger tracking processors gaining interest for Level-1 and 2

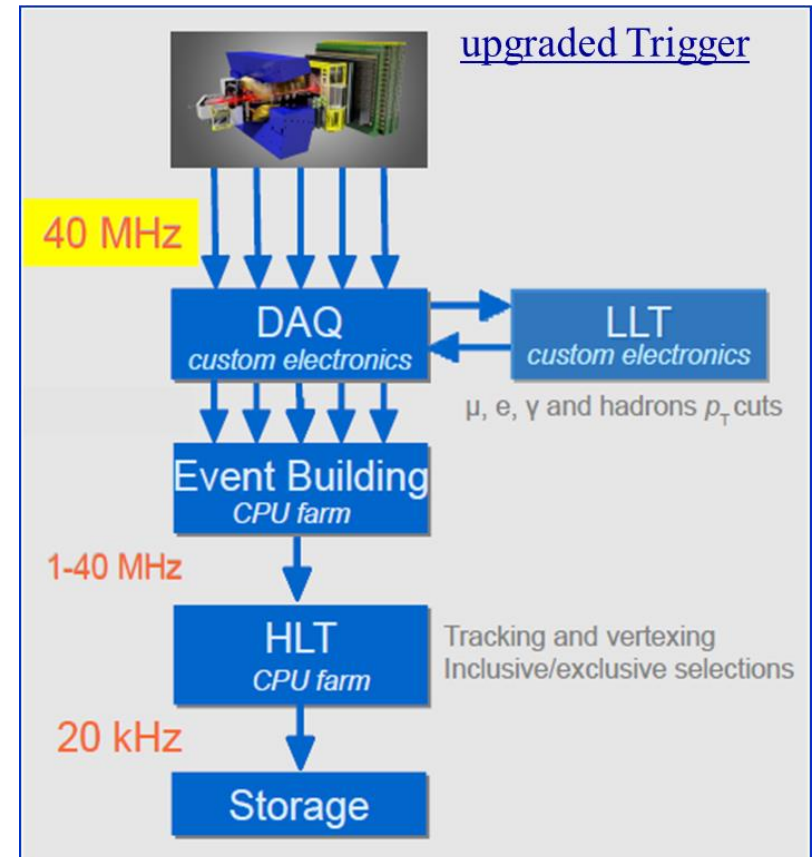
Example of  
CMS new  
trigger  
processor



MP7

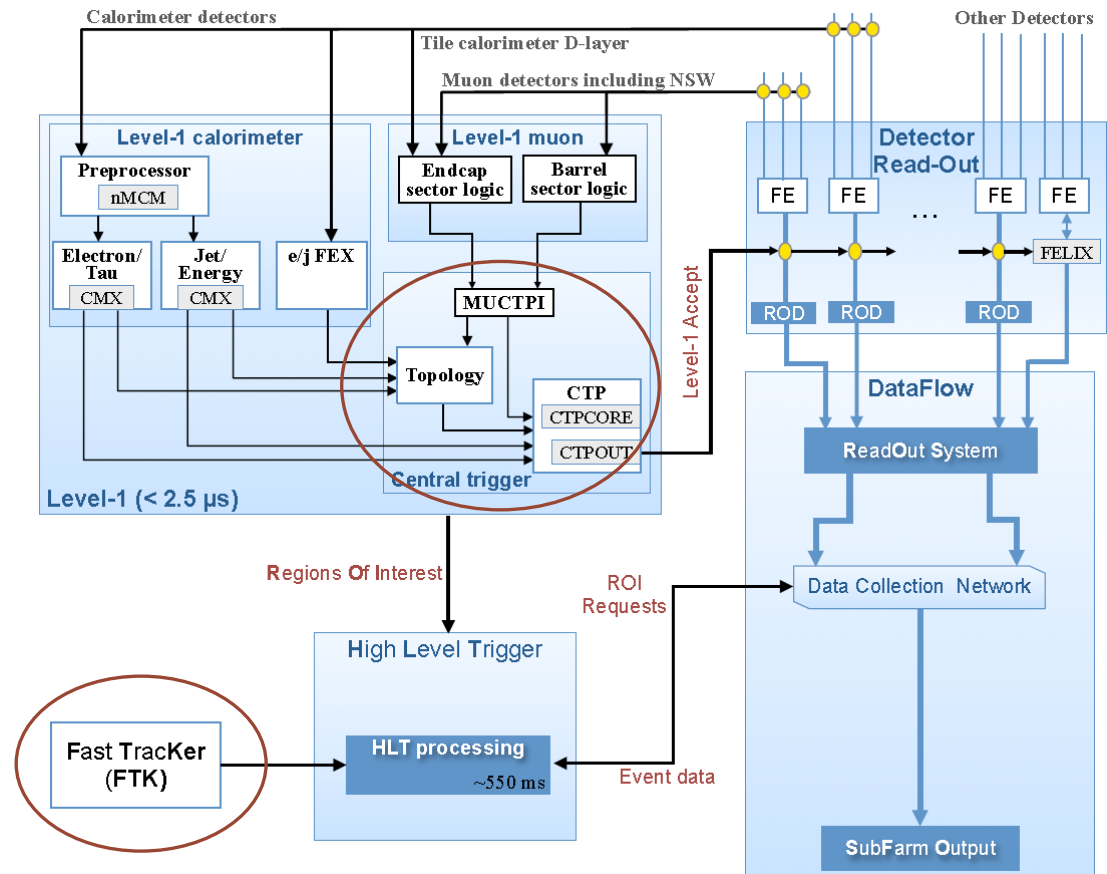
# LHCb upgrade: trigger-less

- All the detector should be upgraded to read data at 40 MHz
  - Data send on surface through optical link using GBT
- Data received in common back-end readout boards TELL40
  - LLT can potentially regulate the EB frequency
- Completely based on commercial computing solutions



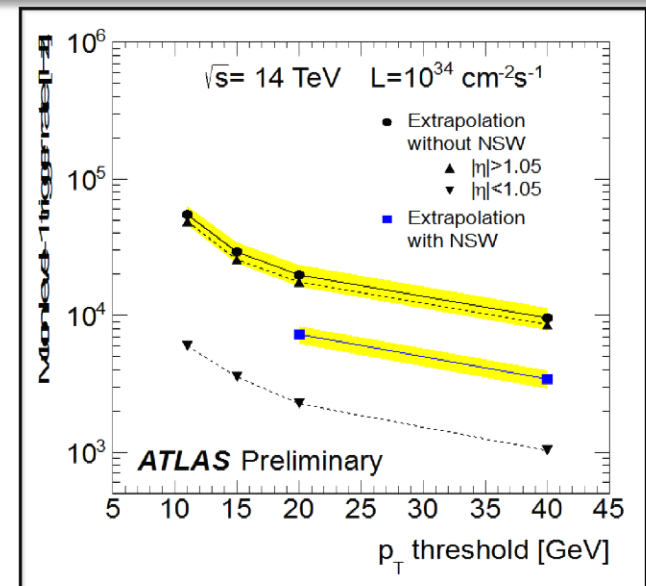
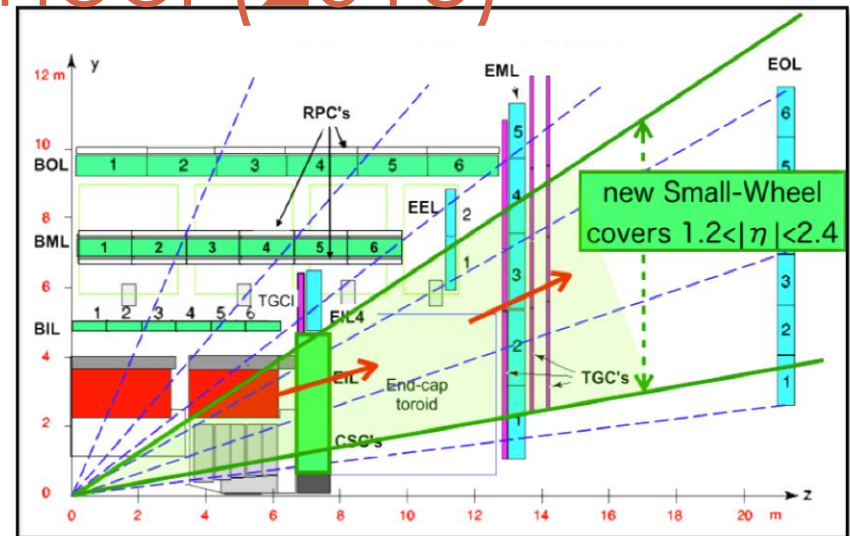
# ATLAS Upgrades during LS1 (>2015)

- The ATLAS experiment has started to redesign the DAQ
- The HLT has been uniformed in a **single farm**
- The electronic is going to be upgraded and new systems added
  - **L1 Topological processor** and the **Fast Tracker**
- New detectors will be installed to reduce the high occupancy issue in the muon system



# ATLAS New Small Wheel (2018)

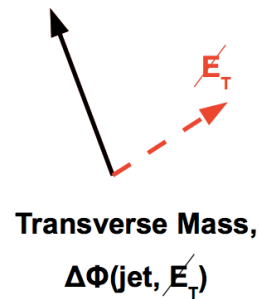
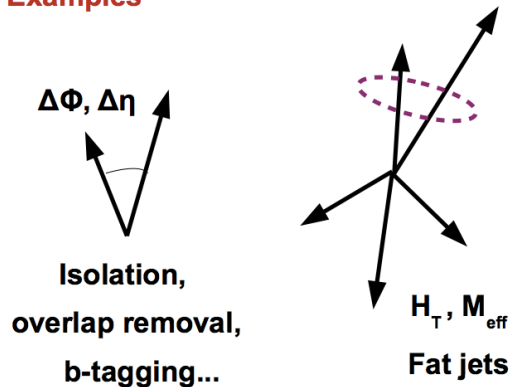
- Consequences of luminosity rising beyond design values for forward muon wheels
  - degradation of the tracking performance (efficiency / resolution)
  - L1 muon trigger bandwidth exceeded unless thresholds are raised
- Replace Muon Small Wheels with **New Muon Small Wheels**
  - improved tracking and trigger capabilities
  - position resolution  $< 100 \mu\text{m}$
  - IP-pointing segment in NSW with  $\sigma_\theta \sim 1 \text{ mrad}$
  - Meets Phase-II requirements
    - compatible with  $\langle \mu \rangle = 200$ , up to  $L \sim 7 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
  - Technology: MicroMegas and sTGCs



L1MU threshold (GeV)	Level-1 rate (kHz)
$p_T > 20$	$60 \pm 11$
$p_T > 40$	$29 \pm 5$
$p_T > 20$ barrel only	$7 \pm 1$
$p_T > 20$ with NSW	$22 \pm 3$
$p_T > 20$ with NSW and EIL4	$17 \pm 2$

# ATLAS L1 Topological processor

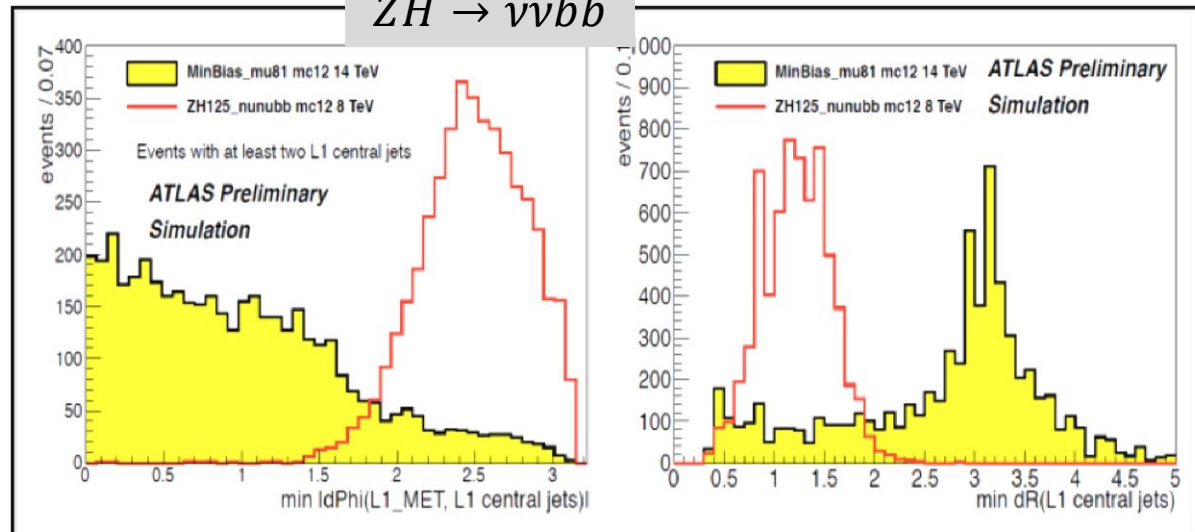
## Examples



- The ATLAS trigger Central Trigger Processor had a limited elaboration power
- Only logic operations were allowed, with a limited number of triggers
- The RoI information had additional information that is not used

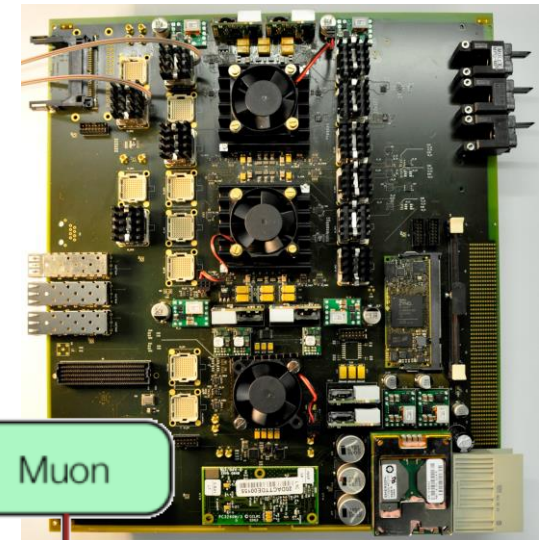
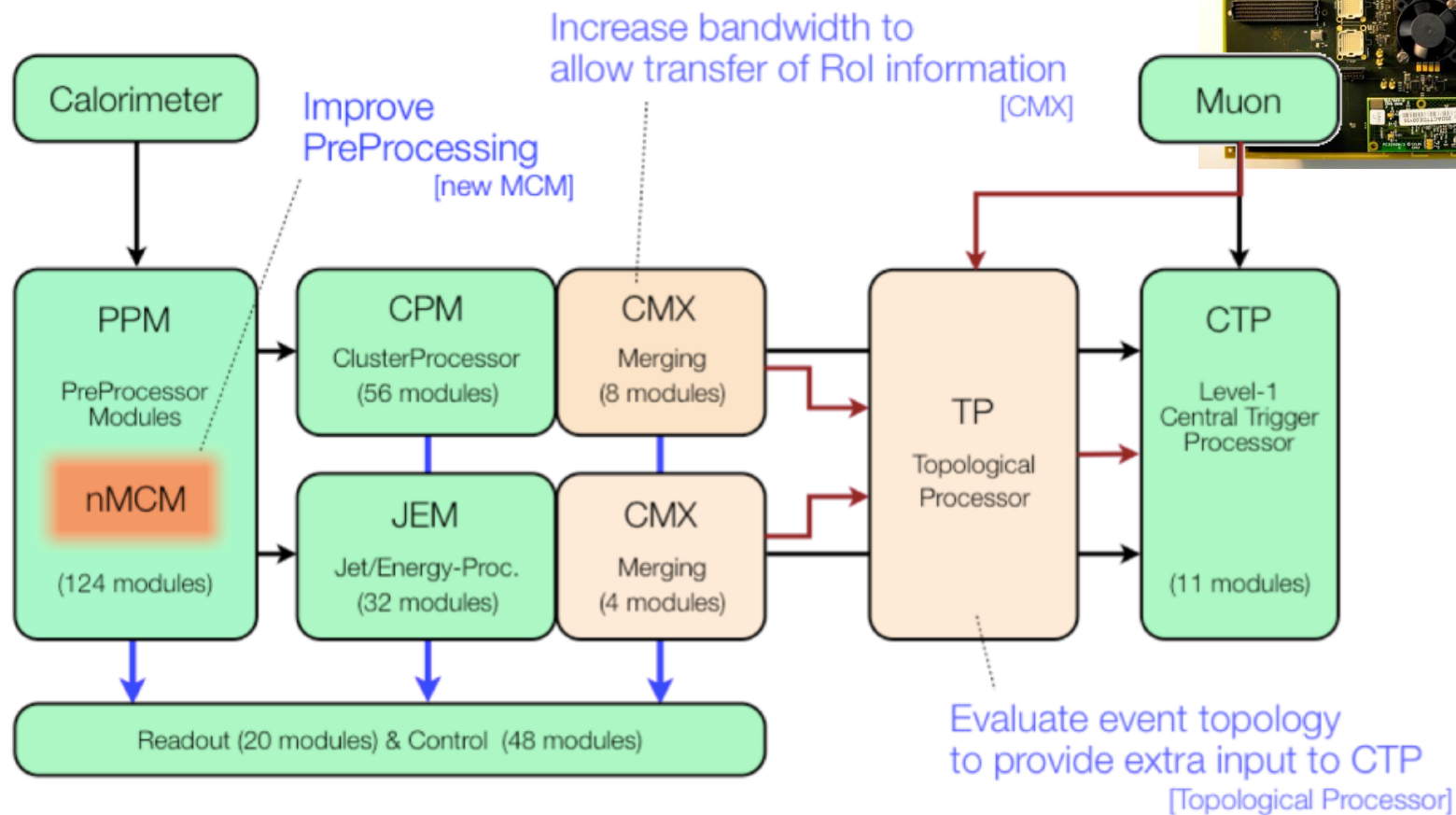
- The L1Topo calculates kinematic variables among L1 objects
- The additional information at Level-1 allows to build smarter selections
- The L1 rate can be kept under control limiting the impact on the signal efficiency

$ZH \rightarrow \nu\nu b\bar{b}$





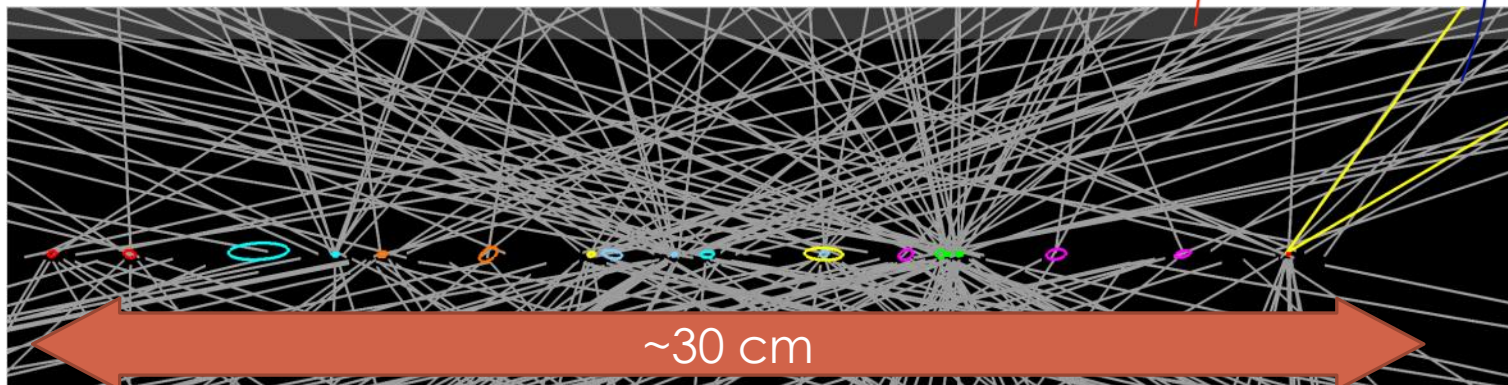
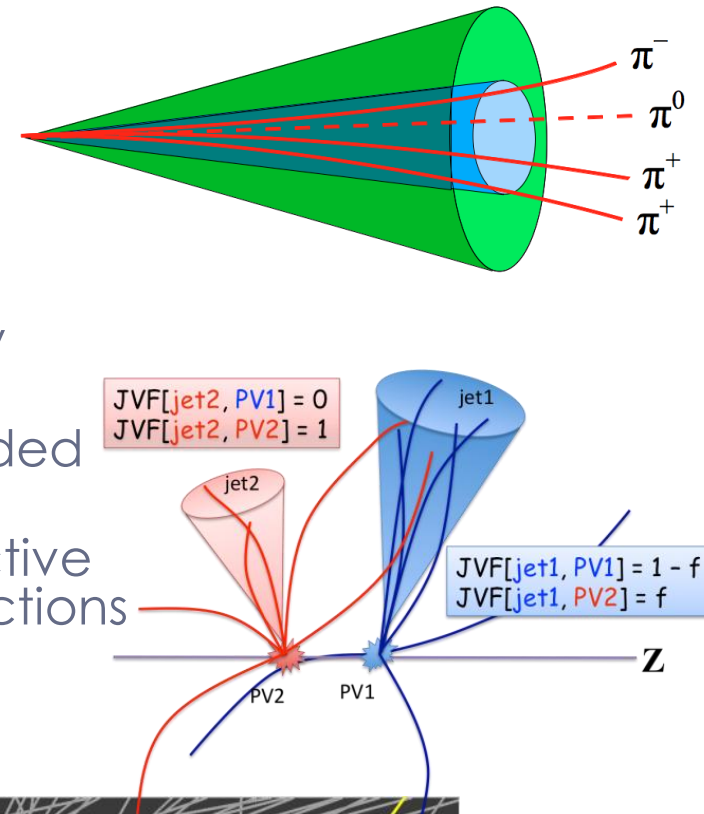
# L1Topo Hardware



ATCA board equipped with Virtex 7 FPGAs

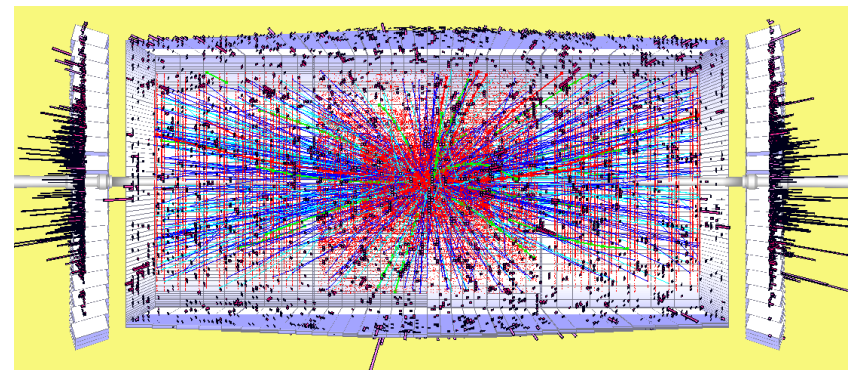
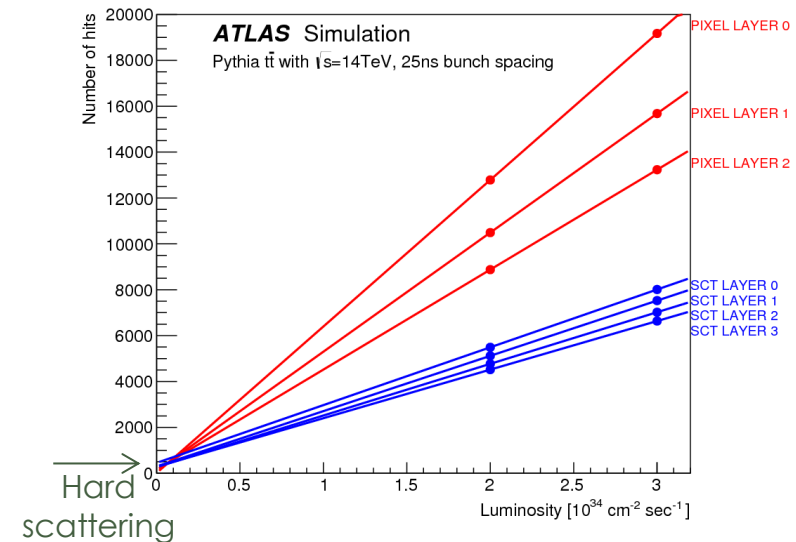
# Tracking & Pileup

- Tracking is powerful identifying special topologies: b-jet, tau-jet, isolated leptons
- Multiple interactions make difficult to disentangle the contributions of the many superimposed collisions
  - Calorimetric measurements are degraded by the pileup contribution
  - Multi-object selection can be less effective because can combine different interactions
- Tracking detectors have the precision to highlight the hard scattering

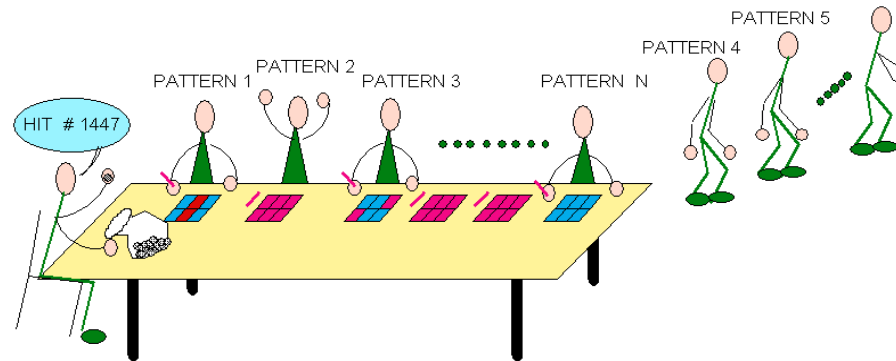
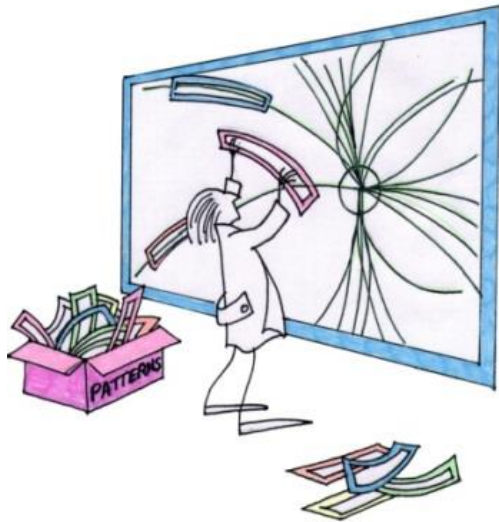


# Tracking in dense events

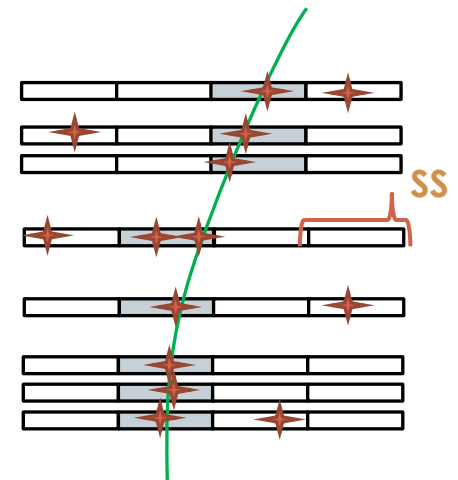
- Tracking algorithms are time consuming
  - Full tracking reconstruction done in the HLT computing farm at LHC
  - Large fraction of the resources are consumed by tracking
- Use the full tracking information at trigger level extremely challenging
  - 1000s of hits received by silicon sensors
    - x10 in HL-LHC
  - Use of specific algorithms and devices to reduce the combinatorial problem
- CDF-like associative memory approach is promising
  - ATLAS L2 RT tracking (Fast Tracker, FTK)
  - L1 track trigger for ATLAS and CMS



# Tracking reconstruction with Associative Memories



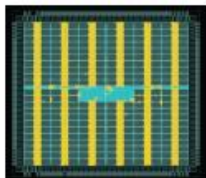
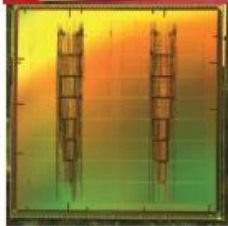
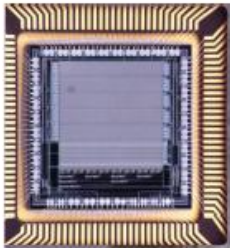
- Tracking can be divided into two sequential steps
  - Search for roads compatible with predetermined patterns
  - Fit of the track candidates within the found roads
- Possible patterns can be precalculated at coarse resolution
  - Clusters can be grouped in super-strips and full precision restored later
- The track's parameters are evaluated from the full resolution hits using a linear PCA algorithm (j.nima.2003.11.078)
  - Effective in FPGA





# Associative Memory History

128 patterns



128k patterns

- **90's Full custom VLSI chip – 0,7 mm** (INFN-Pisa) patterns, 6x12 bit words each (F. Morsani et al., The AM chip: a Full-custom MOS VLSI Associative memory for Pattern Recognition, IEEE Trans. on Nucl. Sci., vol. 39, pp. 795-797 (1992).)
- **1998 FPGA (Xilinx 5000)** for the same AMchip (P. Giannetti et al., A Programmable Associative Memory for Track Finding, Nucl. Instr. and Meth., vol. A 413/2-3, pp.367-373, (1998) ).
- **1999 first standard cell** project presented at LHCC
- **2006 AMChip 03 Standard Cell UMC 0,18 mm, 5k patterns in 100 mm<sup>2</sup>** for CDF SVT upgrade total: AM patterns (L. Sartori, A. Annovi et al., A VLSI Processor for Fast Track Finding Based on Content Addressable Memories, IEEE TNS, Vol 53, Issue 4, Part 2, Aug. 2006)
- **2012 AMchip04 (Full custom/Std cell) TSMC 65 nm LP technology, 8k patterns in 14mm<sup>2</sup>** Pattern density x12. First variable resolution implementation. (F. Alberti et al, 2013 JINST & C01040, doi:10.1088/1748-0221/8/01/C01040)
- **2013 AMchip05, 4k patterns in 12 mm<sup>2</sup>** a further step towards final AMchip version. **Serialized I/O buses** at 2 Gbs, further power reduction approach. BGA 23x23 package.
- **2014 AMchip06: 128k patterns in 180 mm<sup>2</sup>.** Final version of the AMchip for the ATLAS experiment.

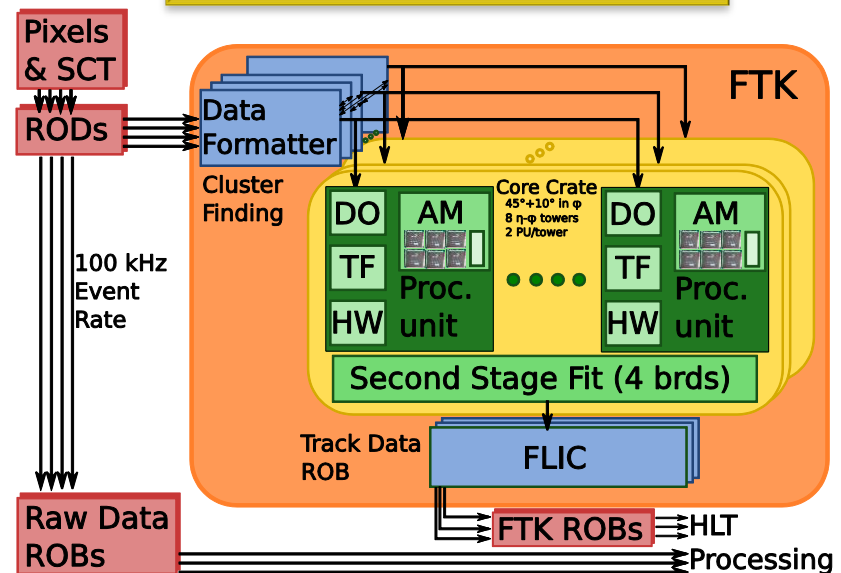


# Fast Tracker Processor at Level 2

- The FTK processor will provide full tracking for each Lvl1 accept (100 kHz)
  - Within a latency of 100  $\mu$ s
  - Targeting all tracks  $p_T > 1$  GeV
- Data are read from the Pixel and SCT RODs through dual output HOLAs
- To increase the bandwidth the data are distributed in **64  $\eta$ - $\phi$  towers**
- Tracks use up to 12 silicon layer and immediately available to HLT algorithms

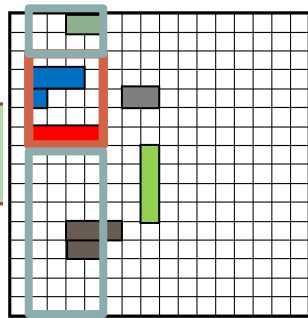
64 core crates

- 2 9U VME boards/towers
- 4 tracking engines each
- **8192 AM Chips**
- **+1000 FPGAs**
- **10000s of high-speed serial links**

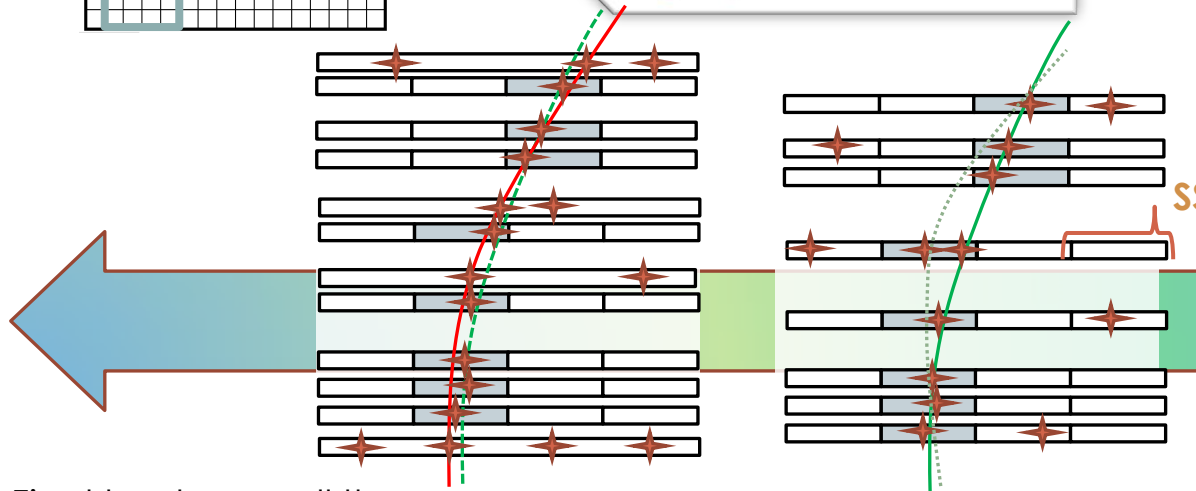
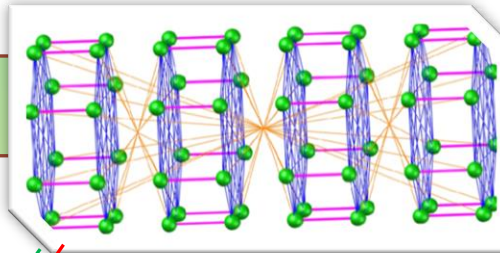


# FTK Main Algorithms Overview

FTK has a custom clustering algorithm, running on FPGAs



The data are geometrically distributed to the processing units and compared to existing track patterns.

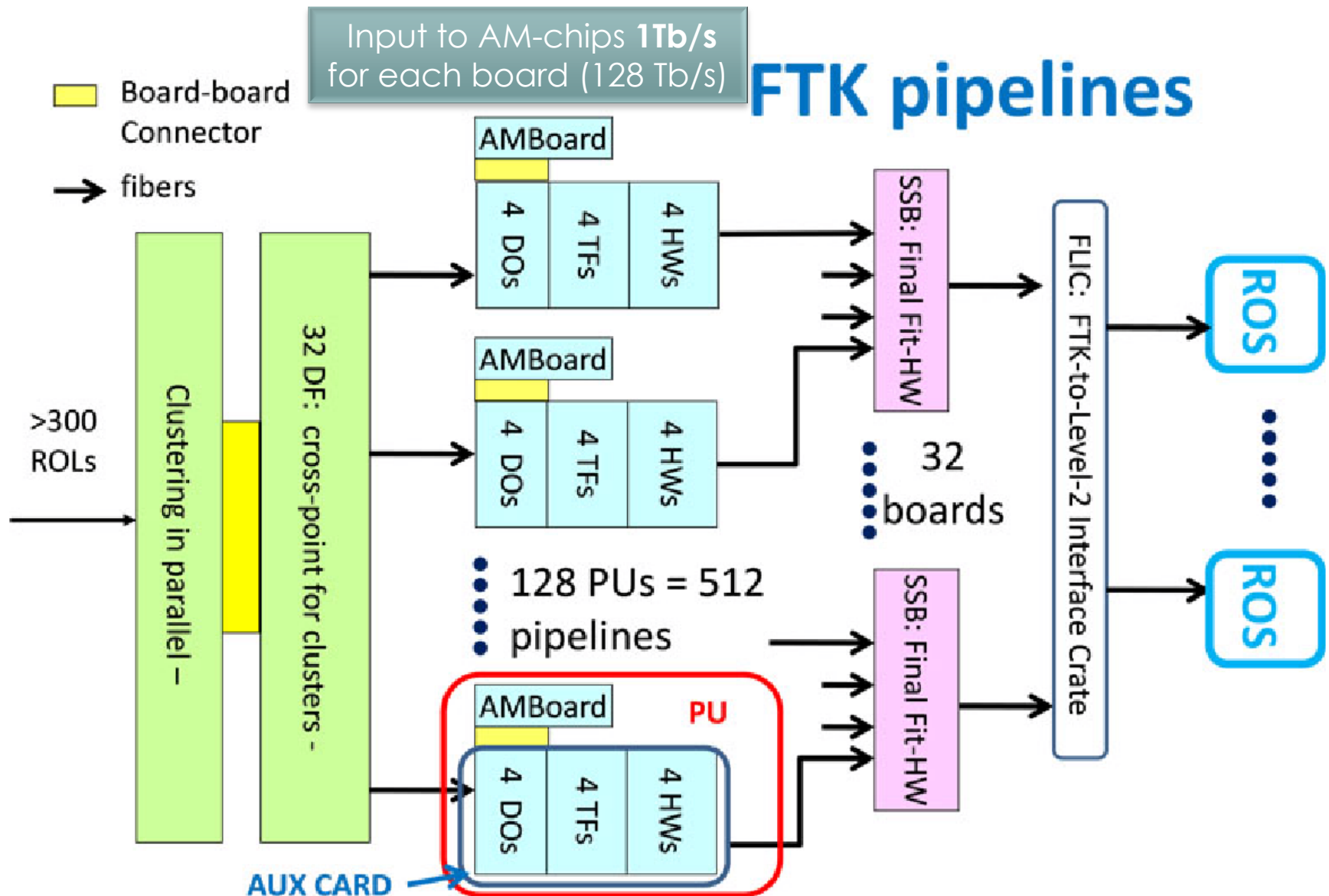


Final tracks use all the available layers

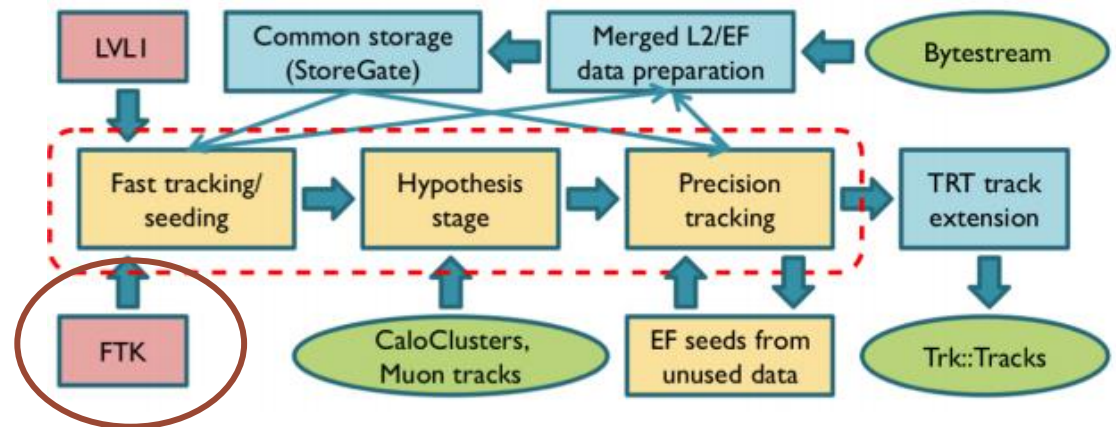
Pattern matching limited to 8 layers: 3 pixels + 5 SCTs. Hits compared at reduced resolution.

Full hits precision restored in good roads. Fits reduced to scalar products.

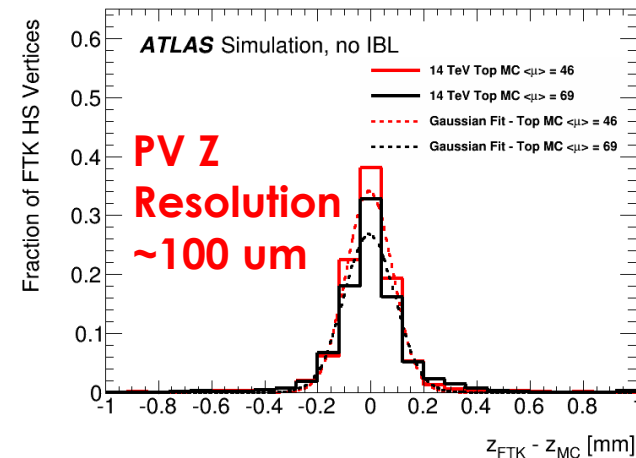
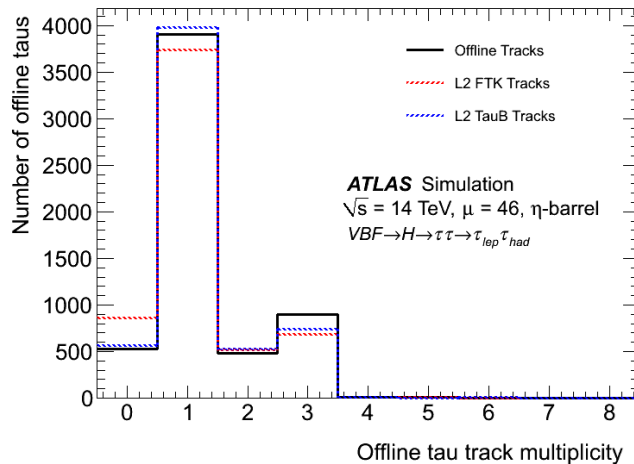
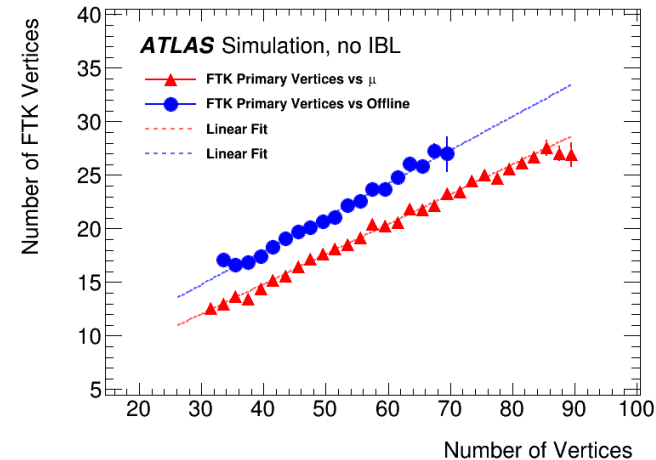
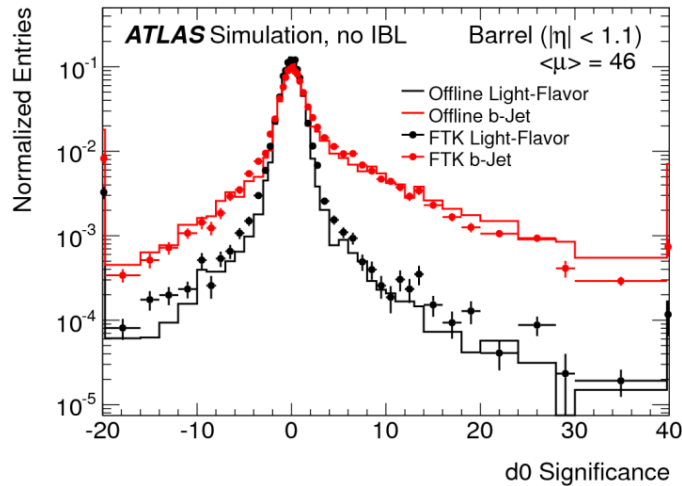
$$p_i = \sum_j C_{ij} \cdot x_j + q_i$$



# Integration with the HLT



- FTK tracks can be integrated with the HLT in different ways
  - Tracks can be directly used to take decision at the begin of HLT processing
    - Allows to run **not-Rol based** selection at full Level-1 rate
  - Tracks can be refitted using CPU-base algorithms
    - No information added, limit the effect of the linearization
  - Used as “jump-start” for global CPU tracking,
    - The pattern recognition step can be largely improved using existing FTK tracks as seeds
- Study on the use of the FTK tracks ongoing



- Efficiency and resolutions are comparable to offline
  - Allow to implement complex selections at the begin of HLT
- Unique possibility to search for primary vertexes



# Track Trigger at Level 1

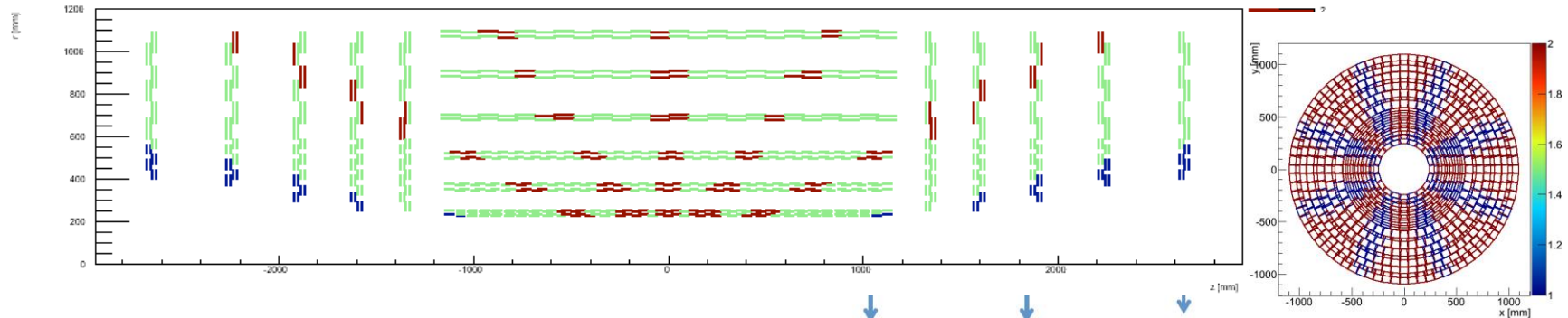
## ATLAS (pull path)

- ATLAS is planning to have L1 tracking “on-demand”
  - Track reconstruction in regions of interest
    - From muons and calorimeters
  - A Level-0 processor is required to drive the tracking
- Tracks used to reject fake EM and muon signatures

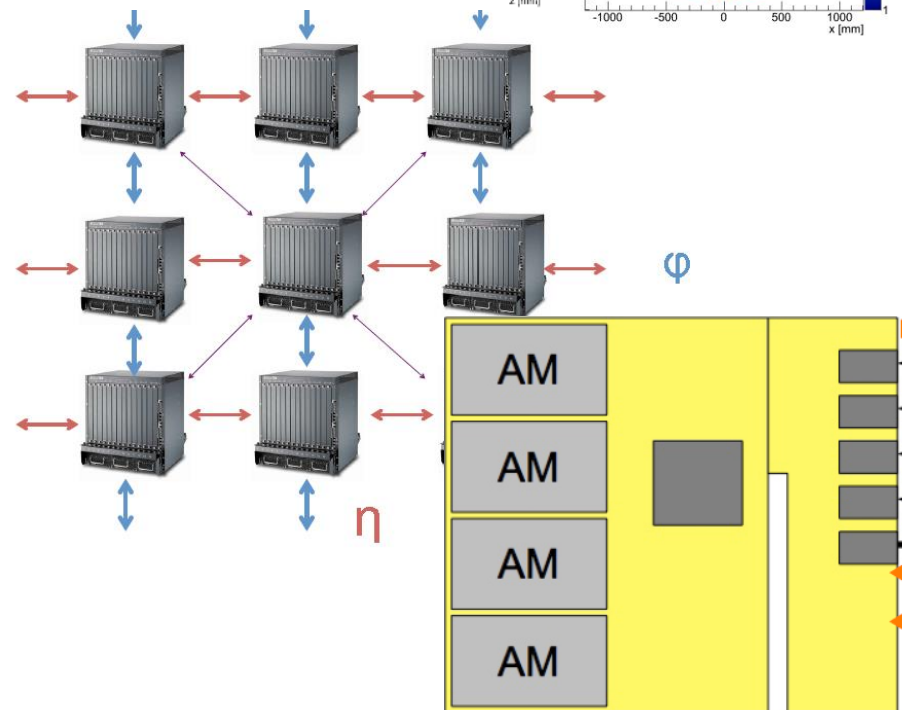
## CMS (push path)

- L1Track carried at each 40 MHz
  - Tracking with extremely low latency (fes  $\mu$ s)
- New tracker detectors designed to allow fast reconstruction
  - Sensors planned to find stubs with  $p_T > 2$  GeV.
    - Reducing the needed bandwidth for the track trigger

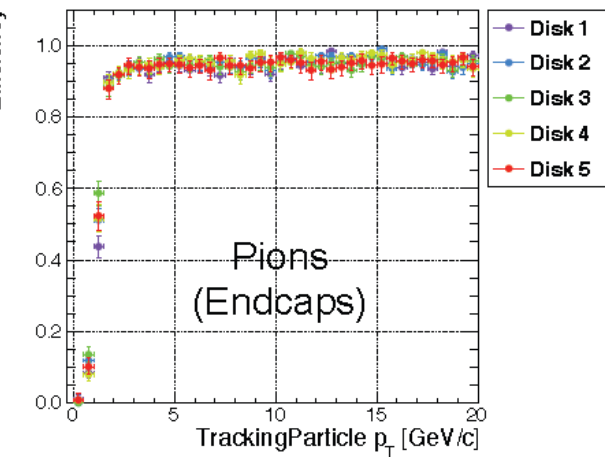
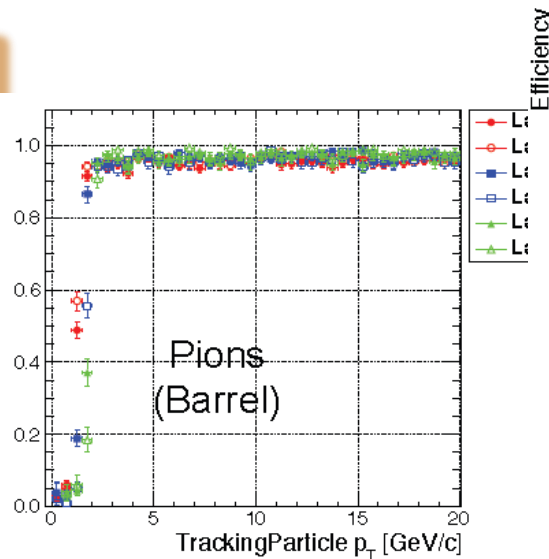
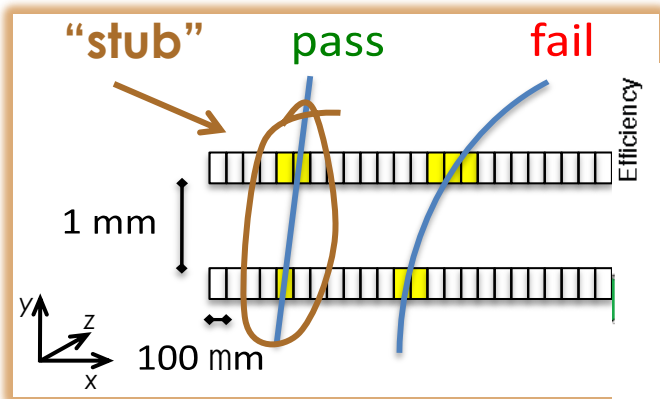
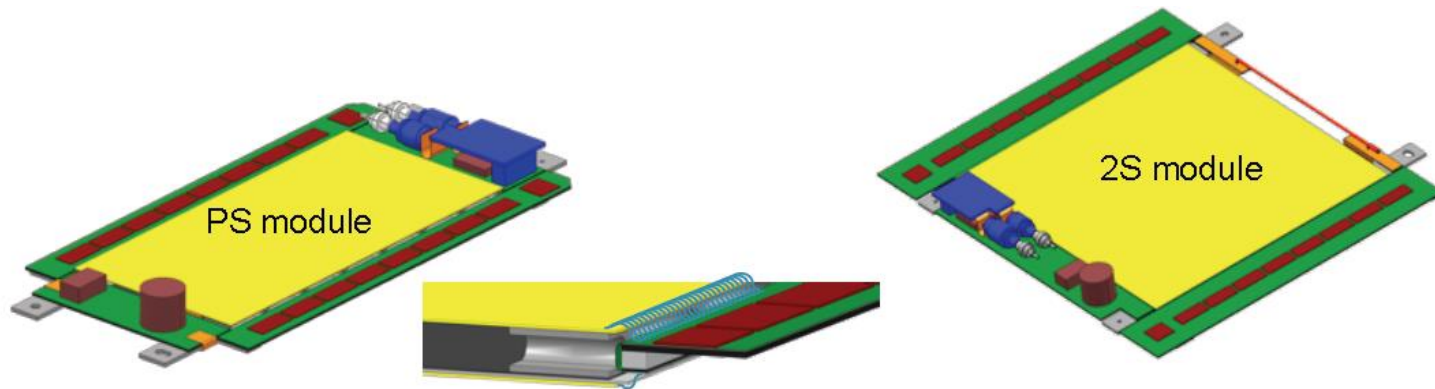
# CMS L1 Tracking Trigger idea



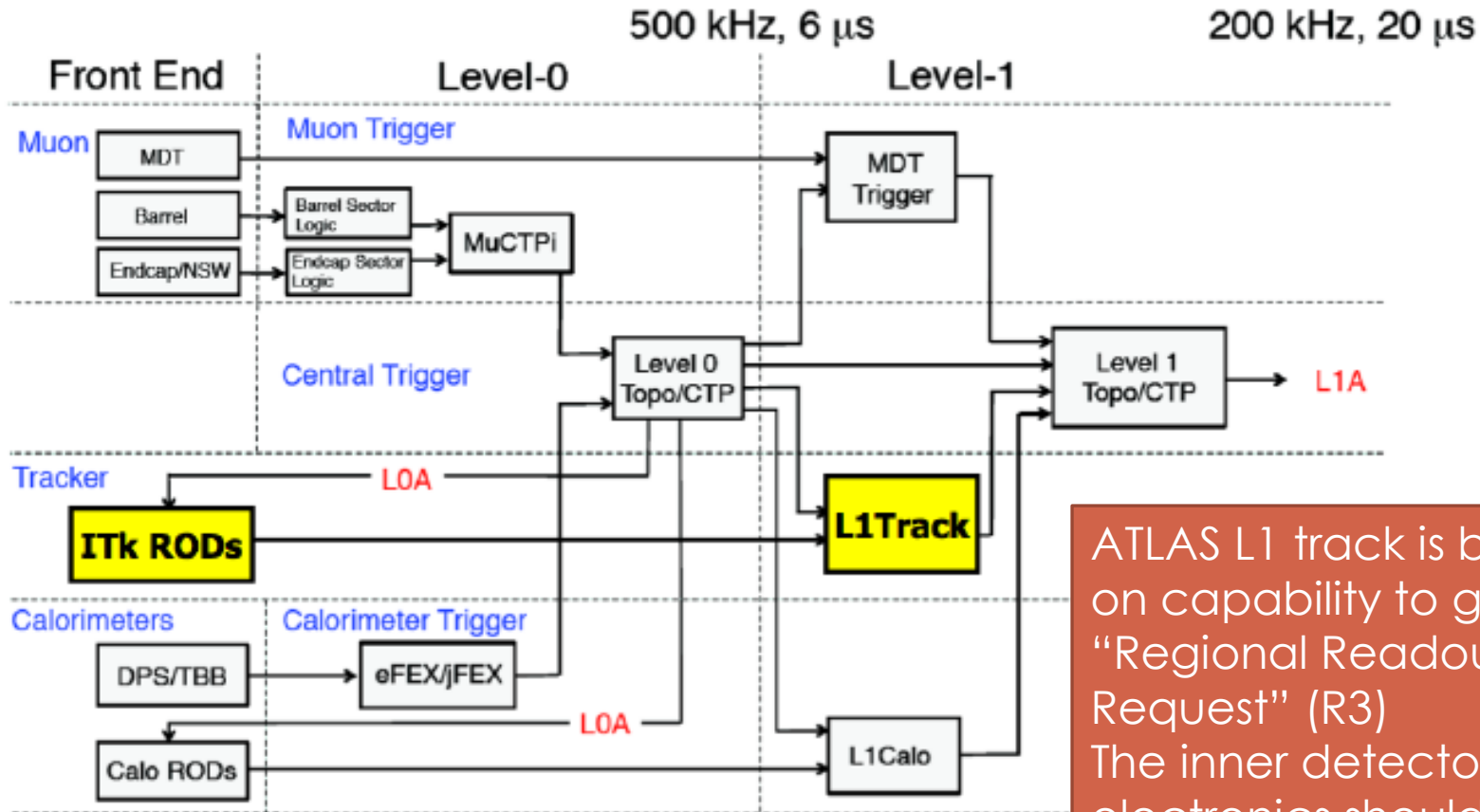
- Detectors can be divided in 48 eta-phi towers
  - Reduced overlap to avoid inefficiency at the boundaries
- Each tower will be controlled by an ATCA crate
  - Each crate will perform the tracking
  - Output composed by track candidates
- Net of interconnected crates
  - Structure derived from FTK Data Formatter



# CMS Outer tracking modules



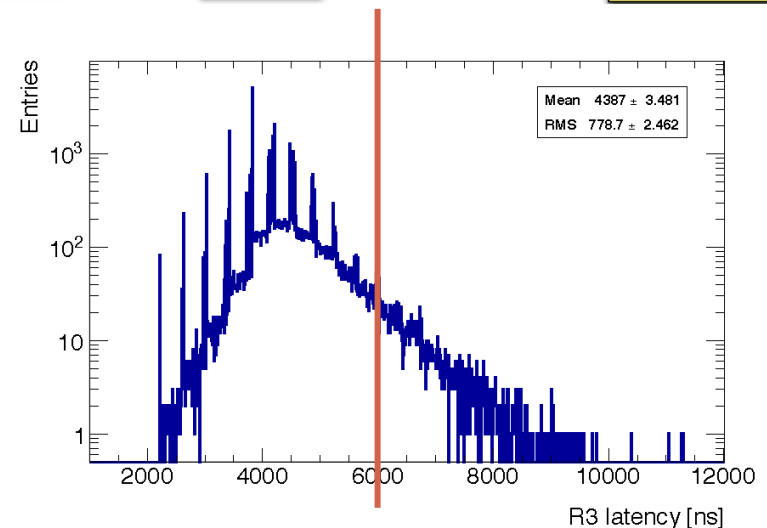
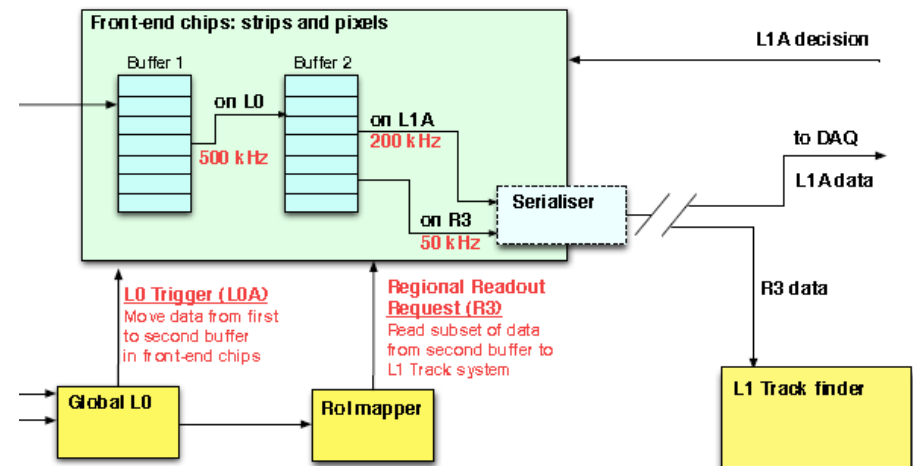
# ATLAS L1 tracking architecture



ATLAS L1 track is based on capability to generate "Regional Readout Request" (R3)  
The inner detector electronics should allow to read the silicon sensor in short time (about 6  $\mu$ s)

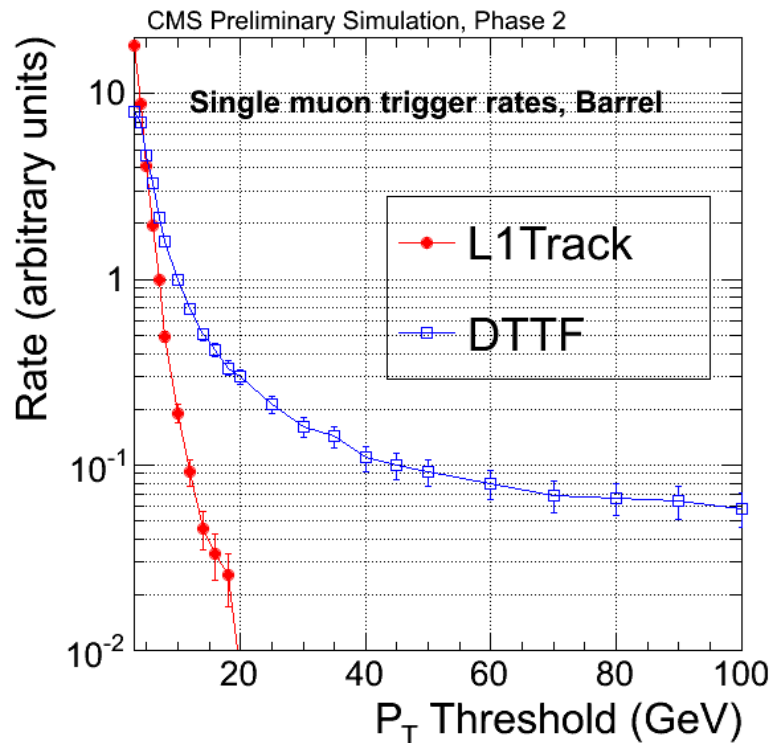
# Proposed Strip Tracker read-out

- The silicon read-out has to transfer data through 2 links
  - The R3 link send data to the track-finder
  - The L1 link send data to the Level-1 processor
  - R3 data are prioritized with respect to the L1 data
- The silicon FE is expected to send data to the R3 processor in  $6 \mu\text{s}$ 
  - Preliminary simulation show how 95% of the hits can be read in such time

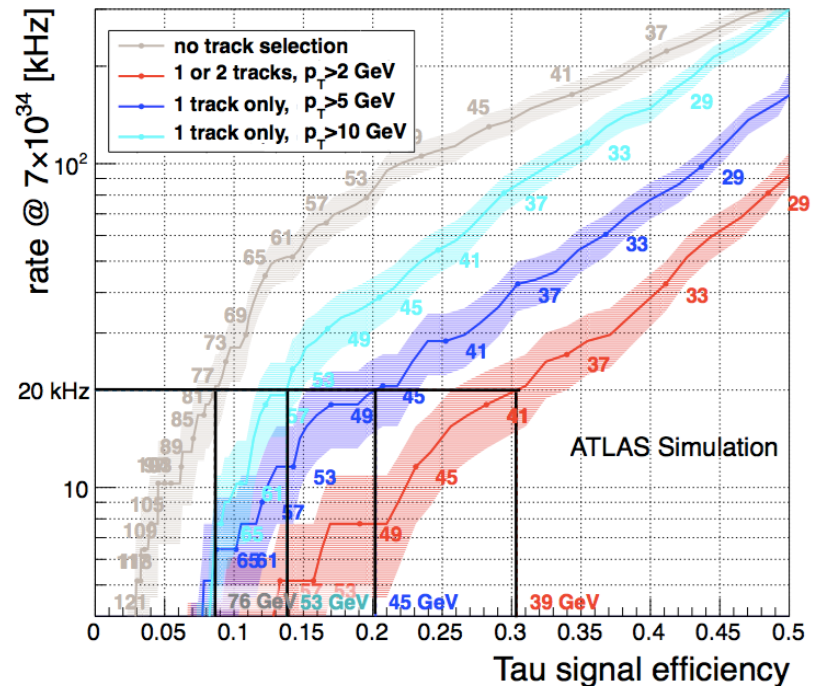




# L1 tracking expected benefit



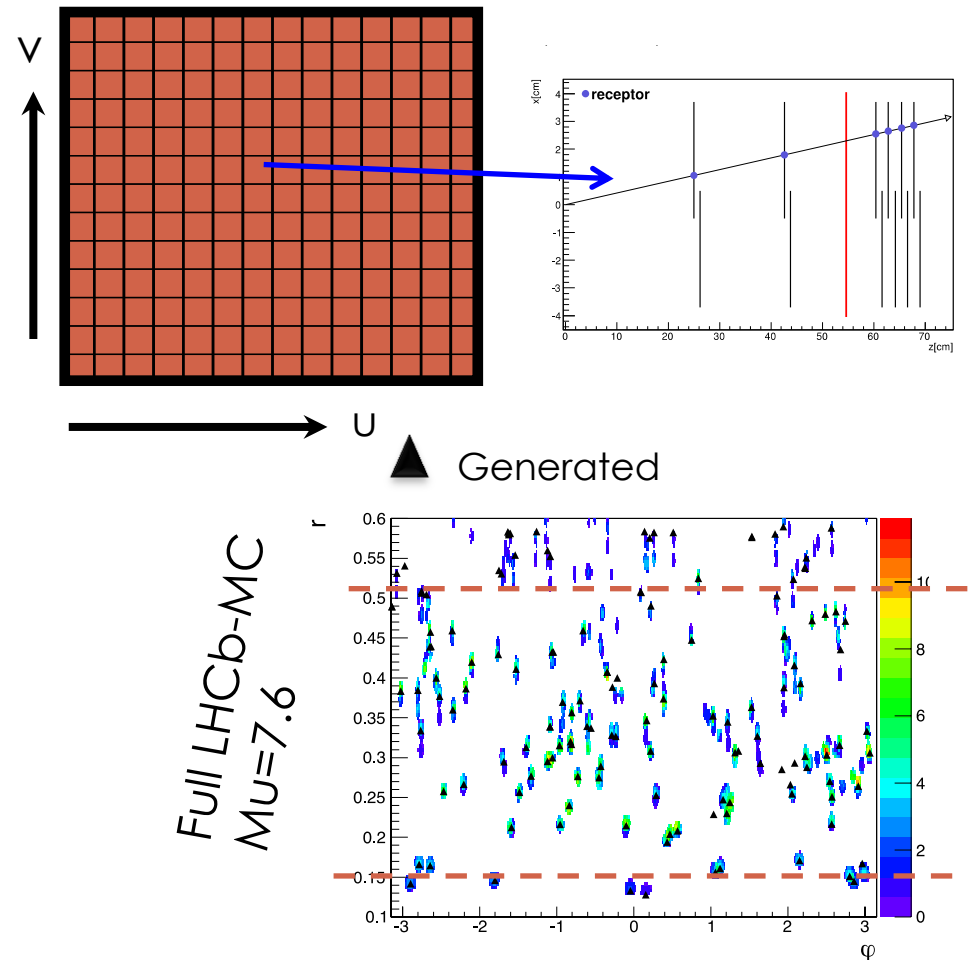
- At CMS the L1TT can fix the single-muon rate flattening
- Other benefits expected



- ATLAS expects to increase the efficiency on  $\tau$ s and reduce the rate from single muon and EM objects

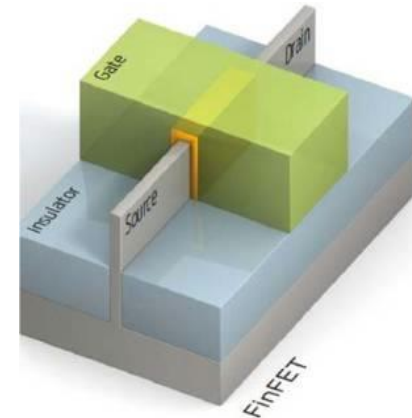
# Human retina inspired algorithm

- LHCb is studying a biological inspired algorithm
  - L. Ristori NIM.A 453 (2000) 425-429
- Track reconstructing using VELO and UT
  - Phase space can be quantized
    - Each set of parameter is assigned to a block of logic: an engine
- A single FPGA in a TELL40 can host 100s of engines
  - **Tracks can be found in  $<1\mu\text{s}$ , efficiency 95%**



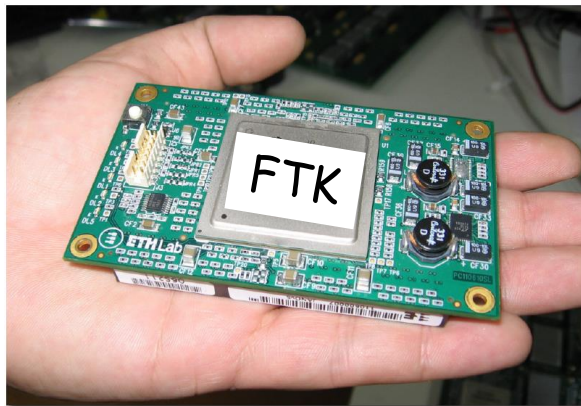
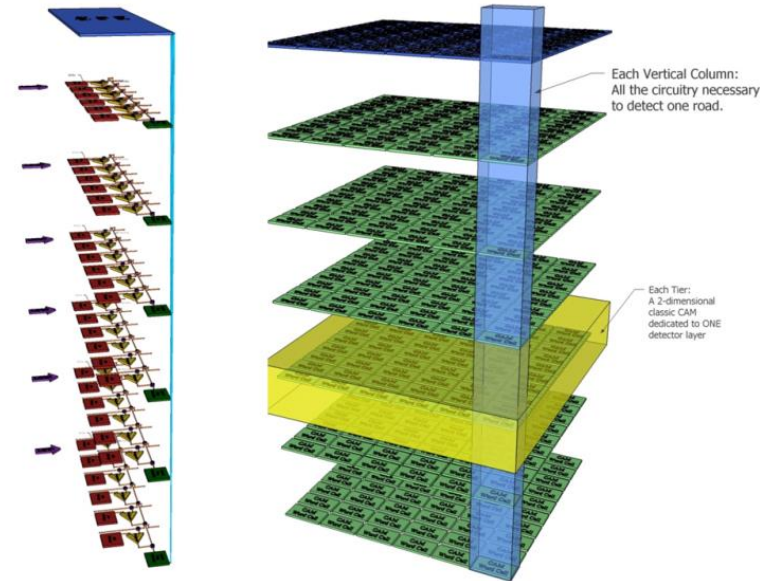
# New electronic standards

- Electronic components are exploiting new fabrication methods
  - ASIC are currently using 65 nm, next step 45 nm
  - Multi-packaging and 3D integration can give an additional boost to miniaturization
- New standard for power-consuming device and high connectivity
  - CMS is moving part of the boards to  $\mu$ TCA
  - ATLAS LS1 upgrades partially based on ATCA
- Intelligent detectors, embedding elaboration capabilities



# AM-based tracking system evolution

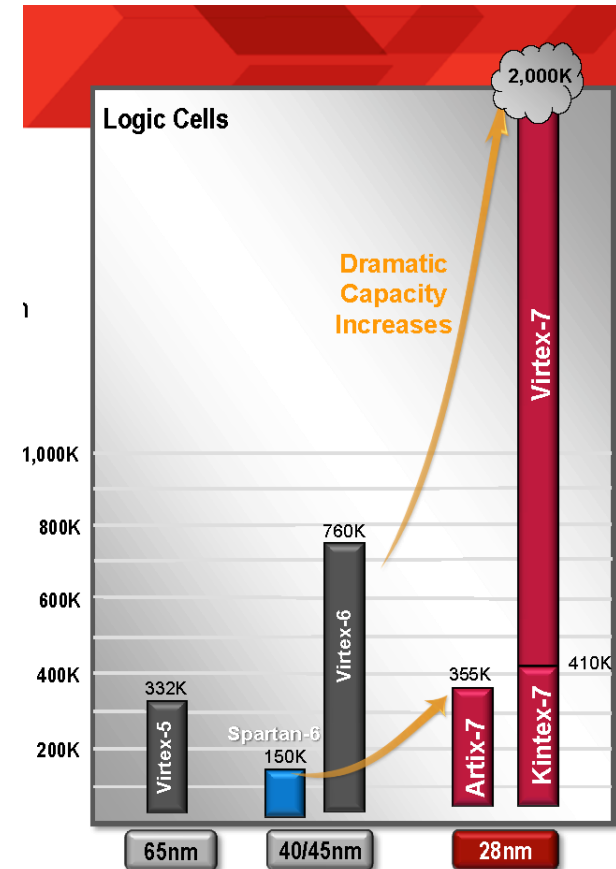
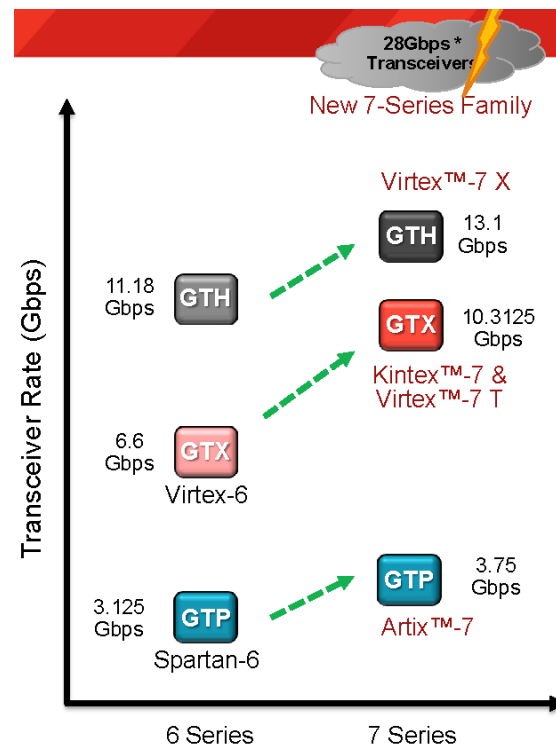
- The AM and track fitting is a candidate for tracking at L1
- The density of patterns and bandwidth requirements are more challenging
- System in a Packages (SiP): AMchip and FPGA in a single chip
  - Extreme miniaturization
  - Better connection of the 2 tracking stages
    - Fake roads rejected inside the chip by the fit stage



- 3D technology offers a further step ahead for logic integration
  - AM chip layers connected by vertical vias
  - Density of patterns increases
  - Power consumption benefits from the vertical integration
- Layers can have different role
  - i.e. FPGA core

# Evolution of FPGA

- Logic Cells
  - 28 nm: > 2X gains over 40 nm→
- On-Chip High Speed Serial Links:
  - Connect to new compact high density optical connectors (SNAP-12...)

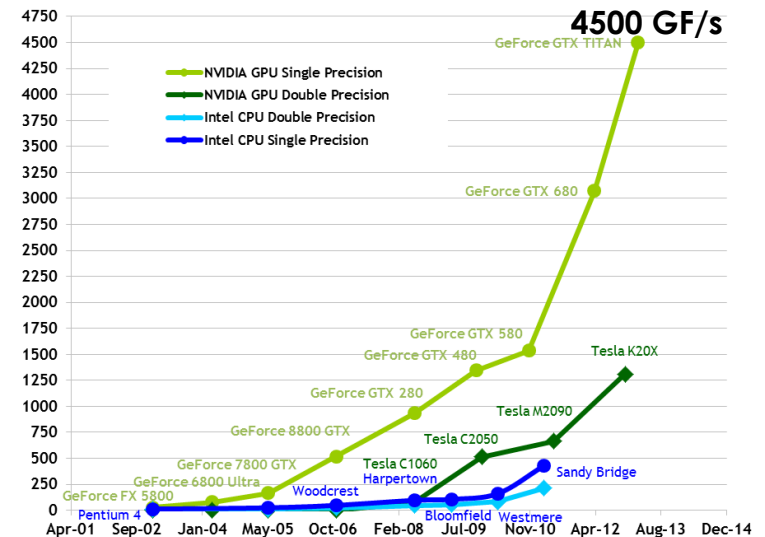




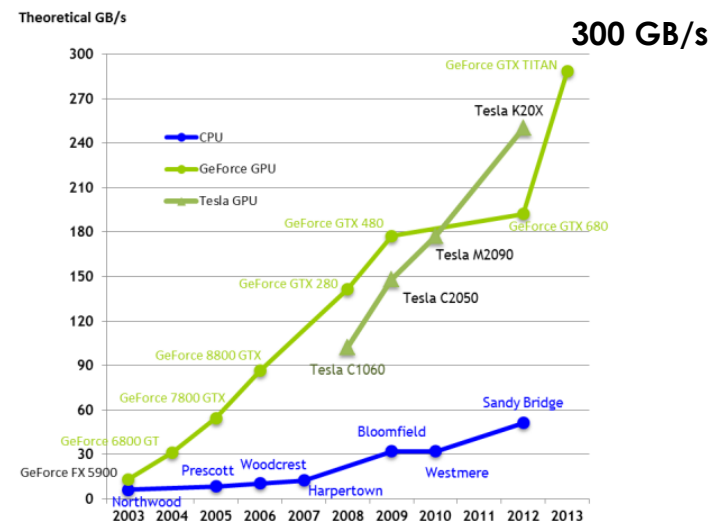
# CPU, GPU and new architectures

- The computer farms remain the workhorse
- Spotlight on new parallel architectures: GPU, Intel Phi
- GPUs are receiving most of the attention
  - Power effective
  - Support of well-known computing languages
    - C/C++, Fortran, Python
- Need to learn how to exploit the new architectures
- Large efforts in studying the use at HLT of GPUs
  - In ALICE, LHCb and NA62 GPUs will have a central role
  - ATLAS and CMS are studying improvements in HLT tracking performance

Theoretical GFLOP/s



Theoretical GB/s

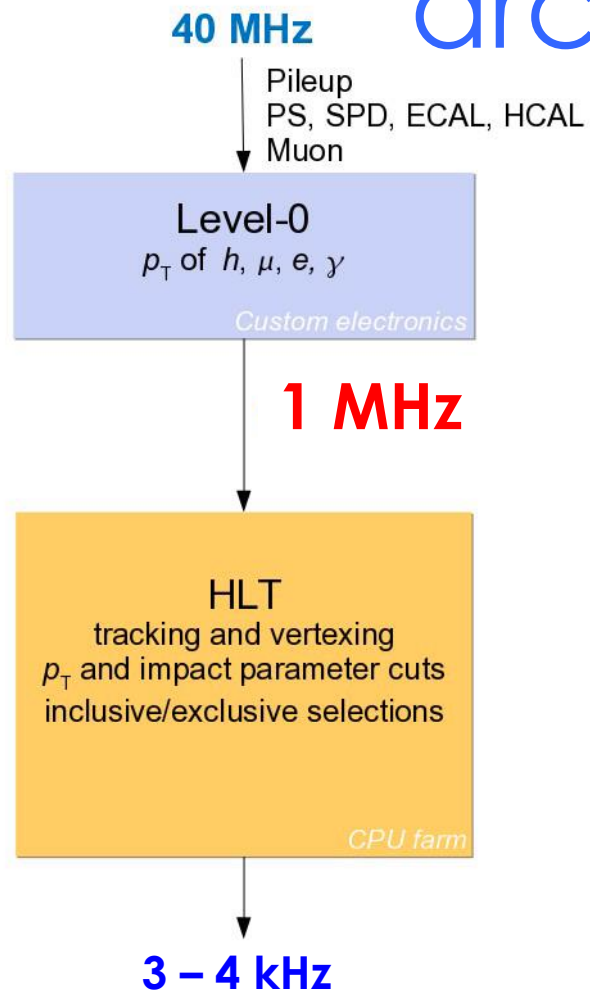


# Conclusions

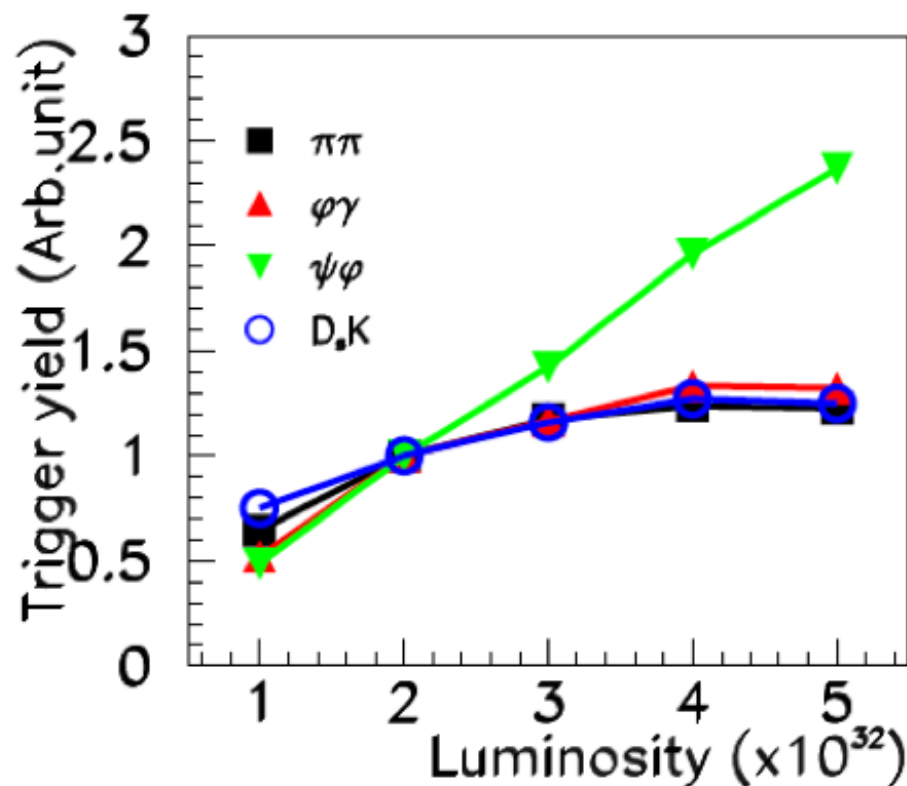
- Trigger and DAQ system were extremely effective for all the LHC experiments in challenging collisions
- The LHC consolidation will bring the experiments more data in even more challenging conditions
  - Experiments will upgrade TDAQ to cope with the new requirements
  - Early availability of tracking can help against the pileup
- Many more upgrades ongoing during LS1 or under design for the next LHC shutdowns
  - Better electronic will allow to extract more information from calorimeters and muon systems
  - Better computing infrastructures will increase elaboration power in the HLT farms
  - Apologize for the incomplete selection based on personal biases

Thank you

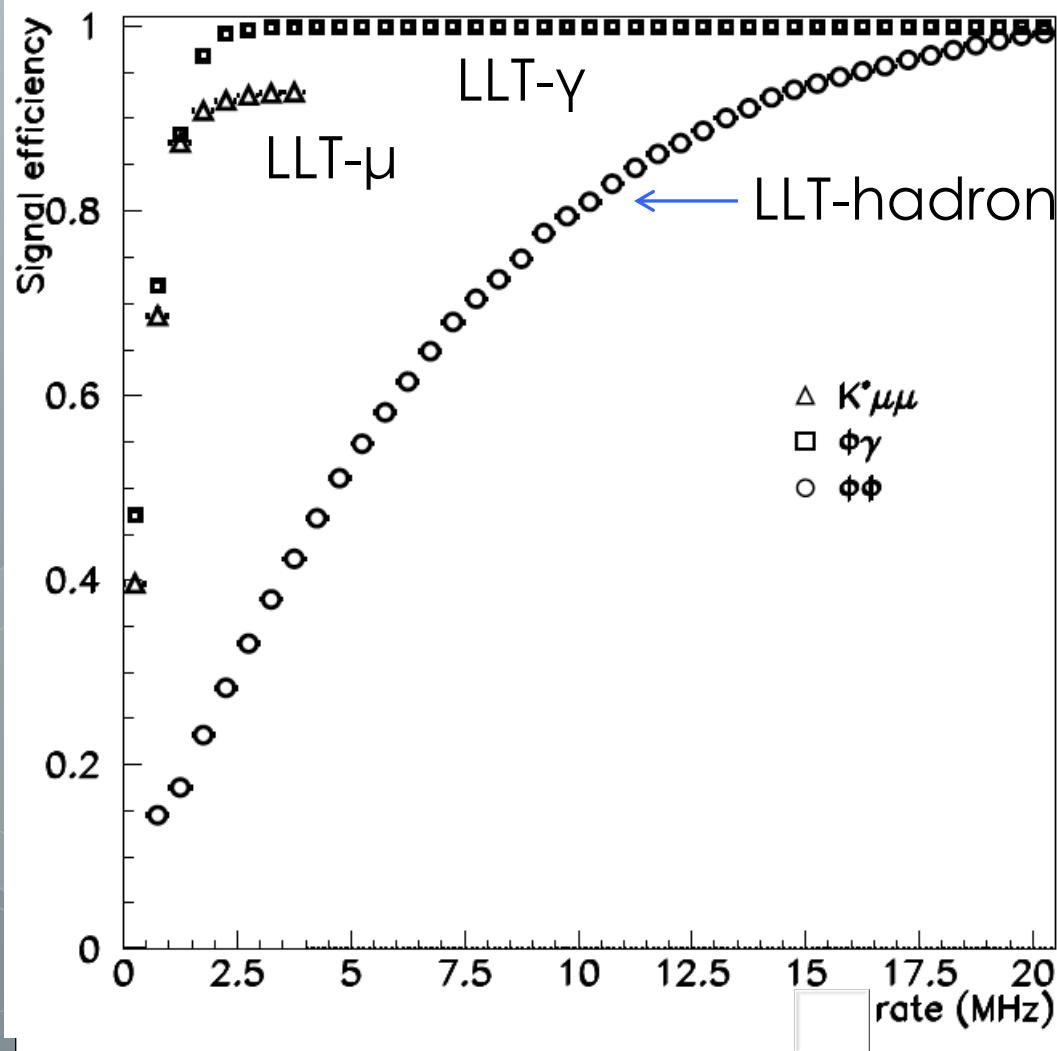
# The present L0 trigger architecture



## 1 MHz L0 trigger rate limitation



# LLT efficiency vs LLT output rate



LLT efficiency

LLT-rate (MHz)	1	5	10
$B_s \rightarrow \phi\phi$	0.12	0.51	0.82
$B^0 \rightarrow K^*\mu\mu$	0.36	0.89	0.97
$B_s \rightarrow \phi\gamma$	0.39	0.92	1.00

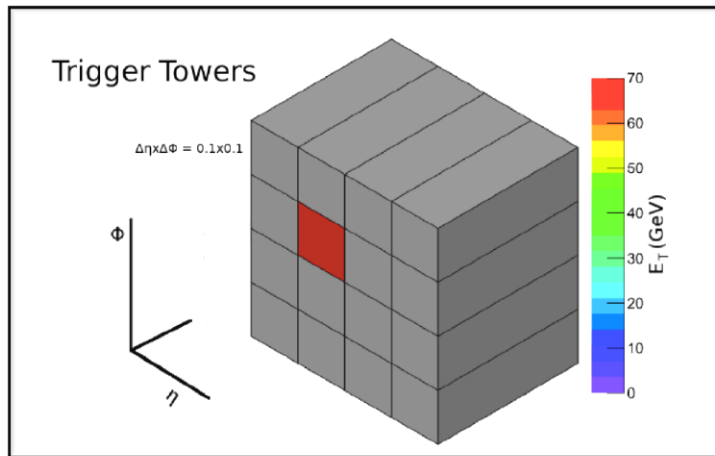


# Level-1 calorimeter trigger

Run-1 calorimeter trigger input:

Trigger Towers  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$

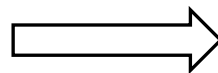
- Used to calculate core energy, isolation



Run-1 trigger menu  
at  $L_{\text{inst}} = 3 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$



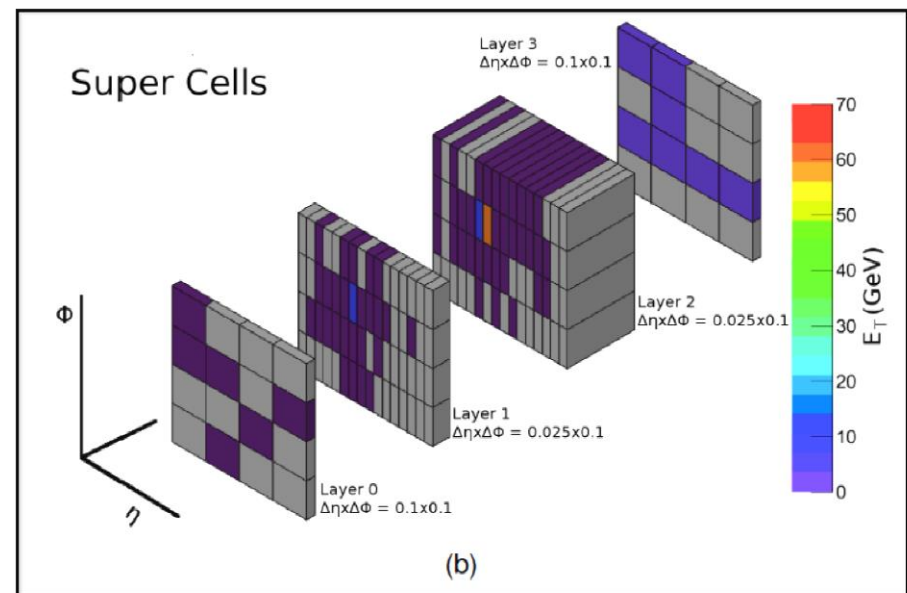
Total rate for EM triggers  
would be **270 kHz**!  
(Total L1 bandwidth is 100kHz)



maintain lower thresholds  
at an acceptable rate



Provide better granularity  
and better energy resolution



Complemented by new L1Calo  
trigger processors eFEX and jFEX

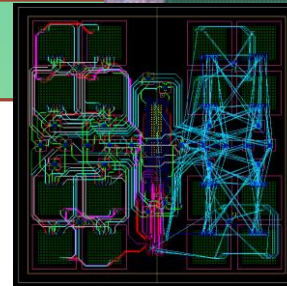
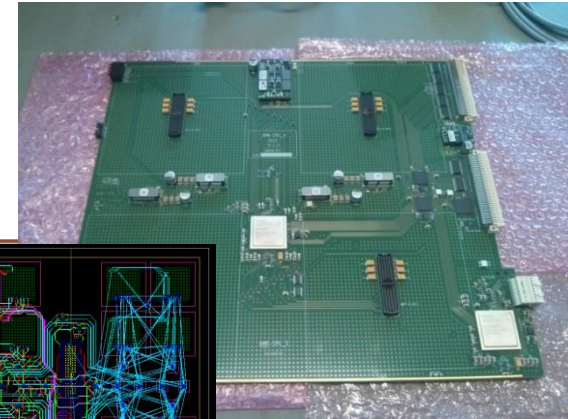
# Final design for the boards

AM board

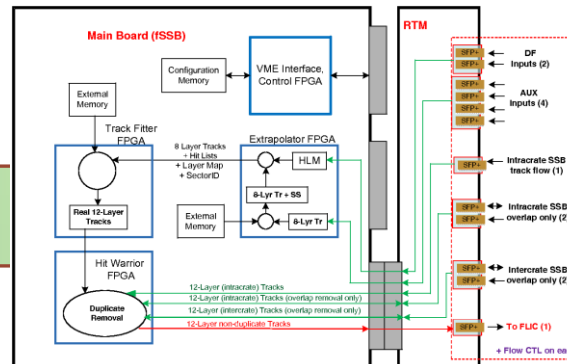
FTK\_IM



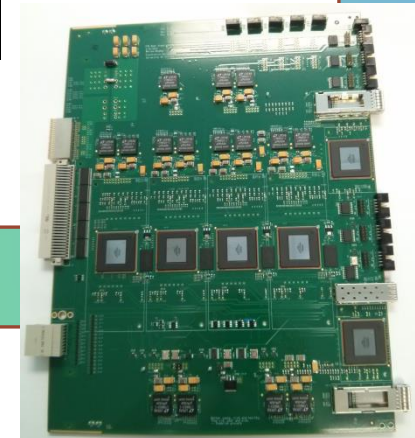
Data Formatter



FTK to Level2  
Interface



2<sup>nd</sup> stage board



AUX card

