Novel triggering strategies at High Luminosity LHC

Marco Zanetti

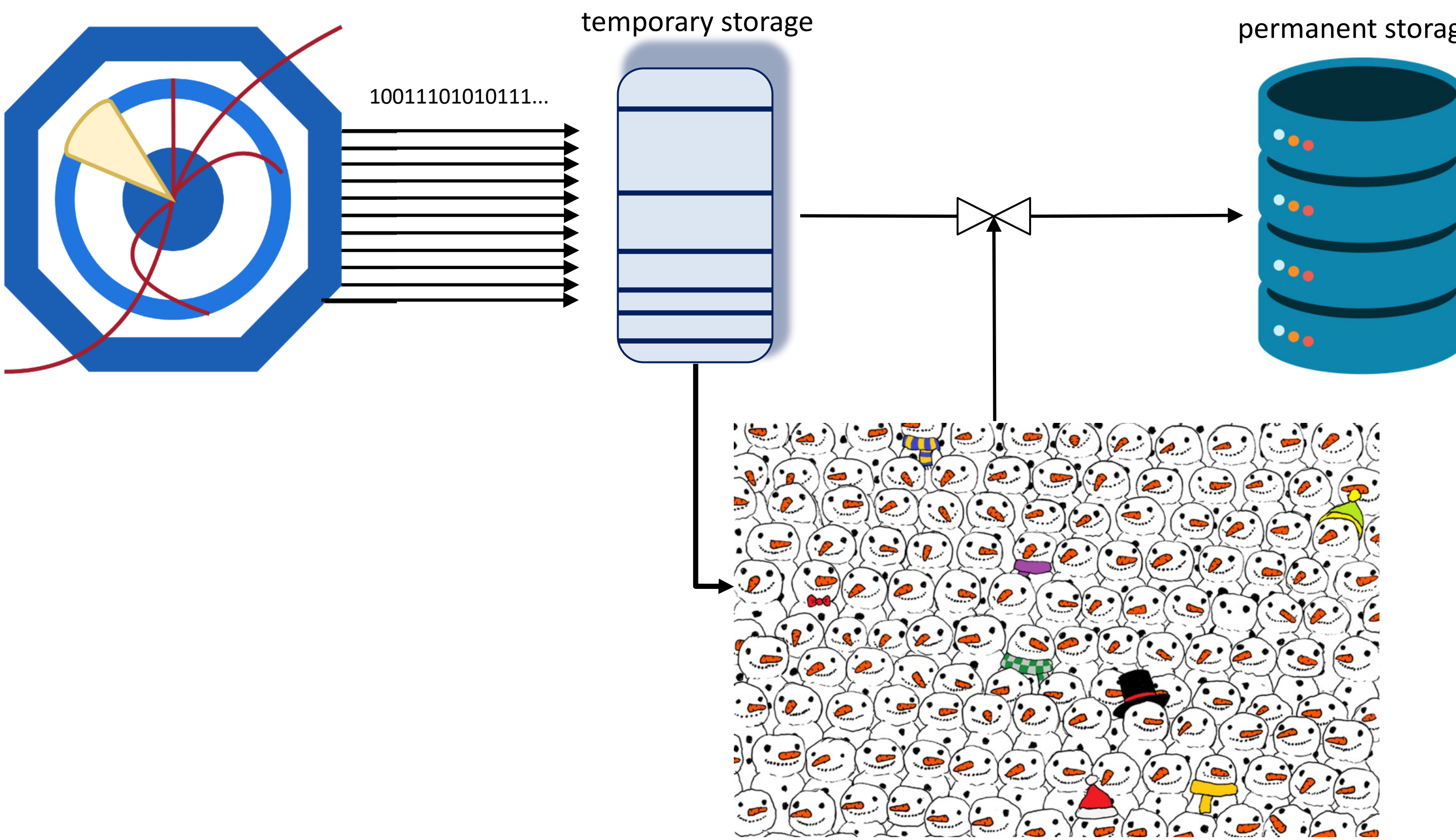
University of Padova and INFN

1222 · 2022
800 ANNI

UNIVERSITÀ DEGLI STUDI DI PADOVA

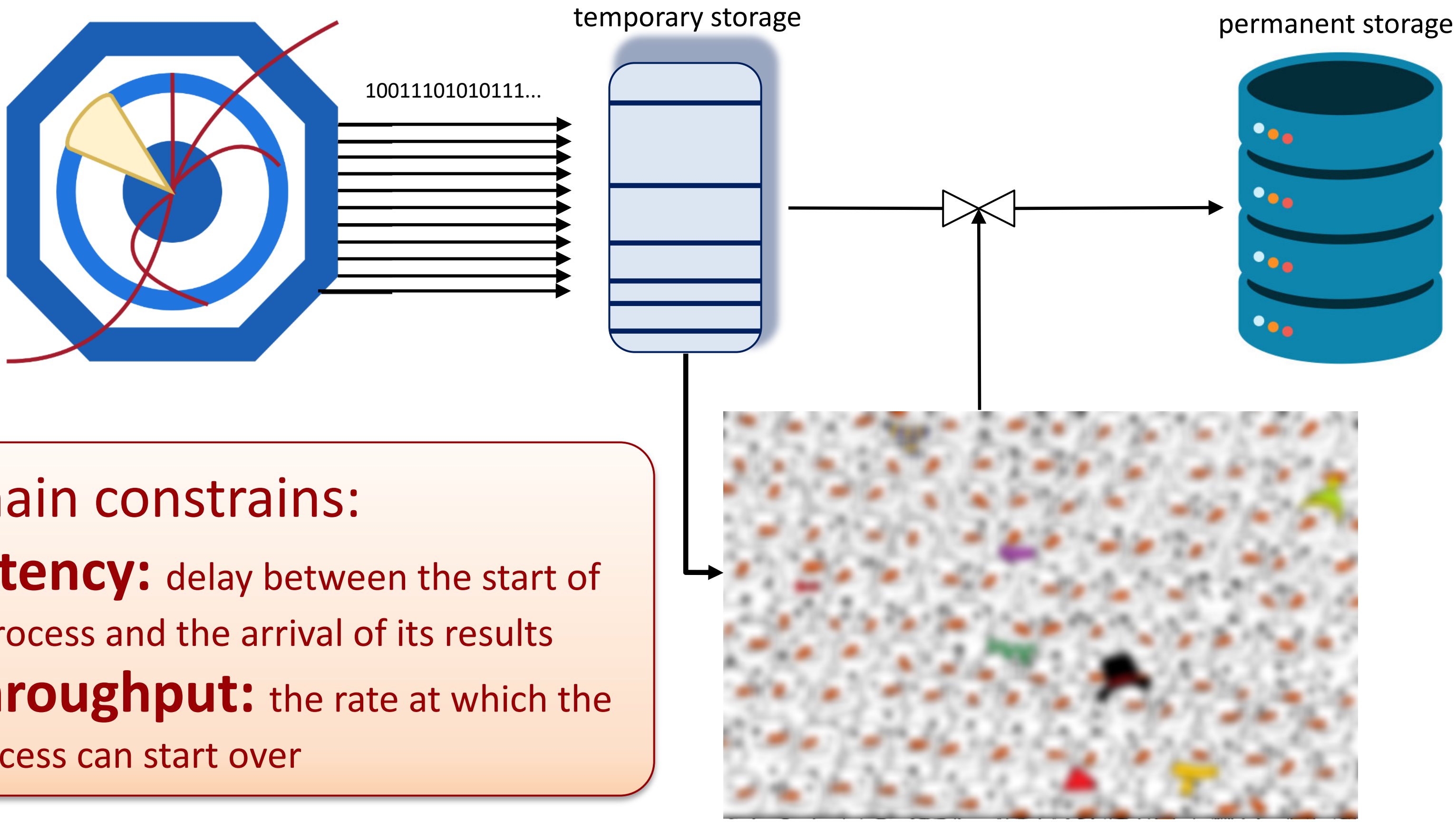
“Classical” Trigger systems

Target a set of predefined signals/topologies, run online algorithms that select them, reducing the event rate to disk to an amount manageable by offline computing and analysis



“Classical” Trigger systems

Target a set of predefined signals/topologies, run online algorithms that select them, reducing the event rate to disk to an amount manageable by offline computing and analysis

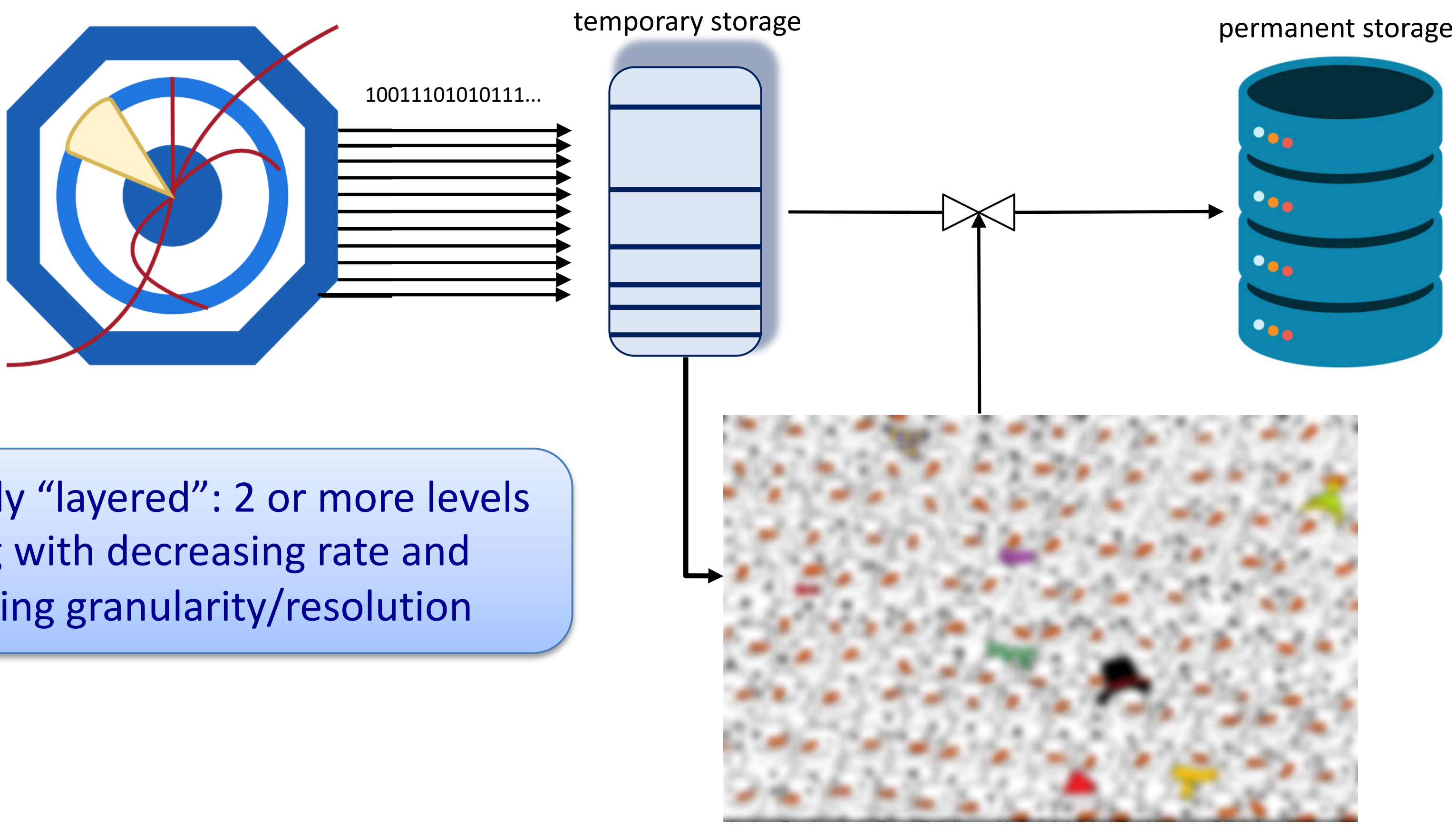


The main constrains:

- **Latency:** delay between the start of a process and the arrival of its results
- **Throughput:** the rate at which the process can start over

“Classical” Trigger systems

Target a set of predefined signals/topologies, run online algorithms that select them, reducing the event rate to disk to an amount manageable by offline computing and analysis



Typically “layered”: 2 or more levels dealing with decreasing rate and increasing granularity/resolution

Leveraging Technology Advancements

Trigger systems have been and will be profiting from sensational advancements in Telecom, Computing, Electronics, Algorithms

NETWORKING

high speed – high capacity optical links of hundreds of Gbps

ALGORITHMS

complex Machine Learning models for classification, data compression, anomaly detection

MEMORY

fast access – high density memories allowing deep buffering (longer latencies)

USER FRIENDLINESS

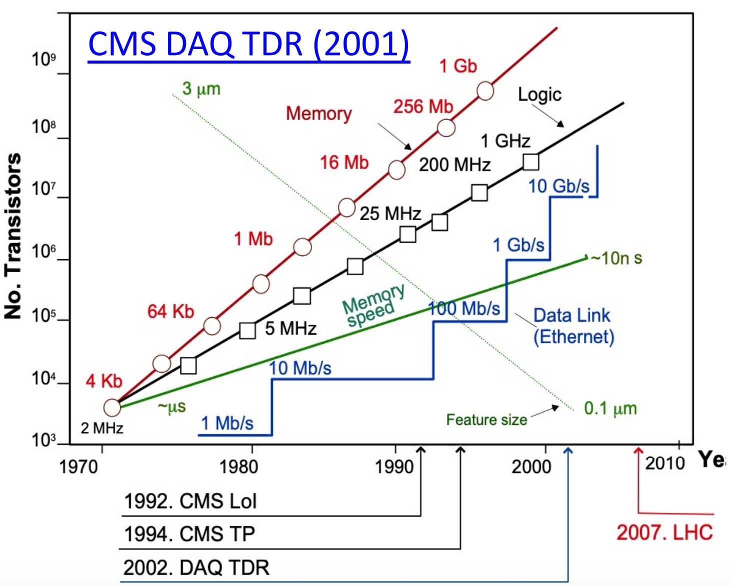
convenient tools, large communities, common standards: enabling development and deployment

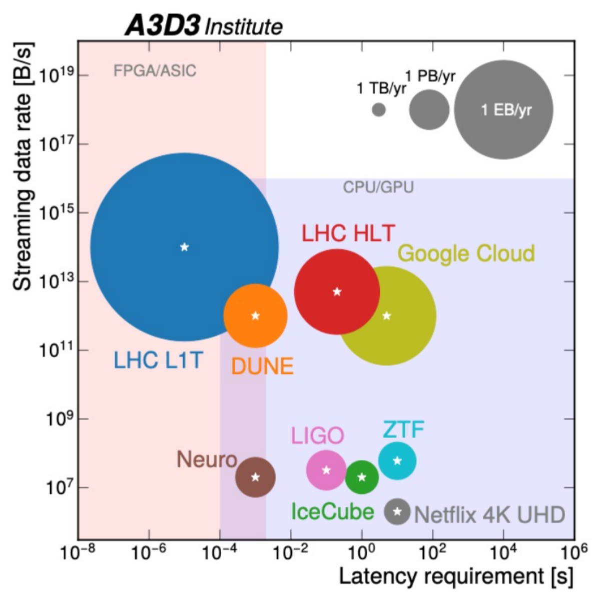
PROCESSING

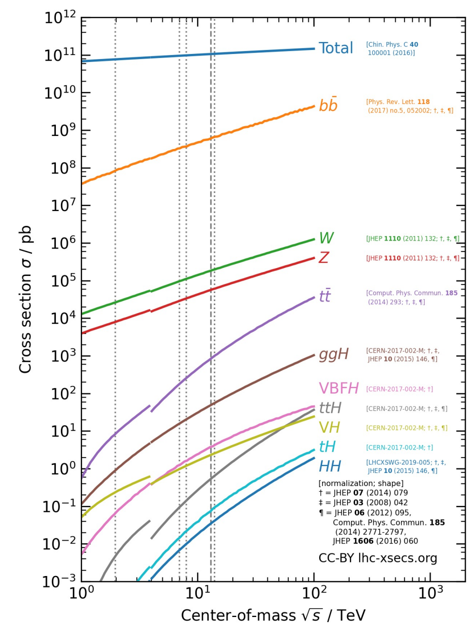
still improving CPUs; GPUs as boosted by AI revolution; FPGAs as new standard for high performance accelerator/processing unit; distributed computing

A word of caution

- 10+ years of LHC running ahead, may projections lead to reconsider the need of a trigger system?
- Trigger will evolve, but at least for omni-purposes experiments, it will still be playing a crucial roles
  - Among the biggest players (~50 TB/s), also considering commercial activities
  - We know a large portion of events are little interesting
  - Accurate Monte Carlo samples (with weight <1) corresponding to full stat impossible to produce







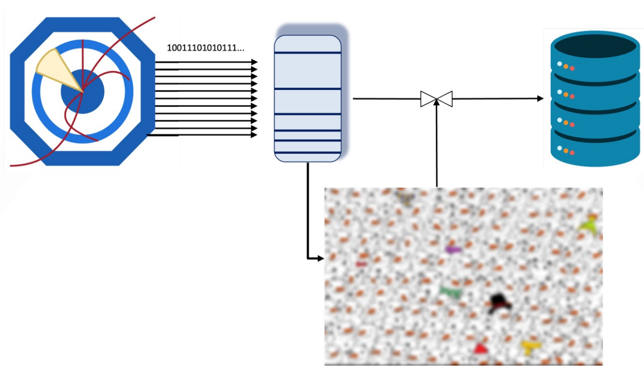
Objectives

New Trigger strategies are key to empower physics reach of experiments. That will especially the case for HL-LHC

More efficient selection for target signals/topologies

Extend phase space for interesting topologies (low p_T)

Reduce the bias! Deal with the unexpected (anomalies)



Strategies

Get a better "classical" trigger system

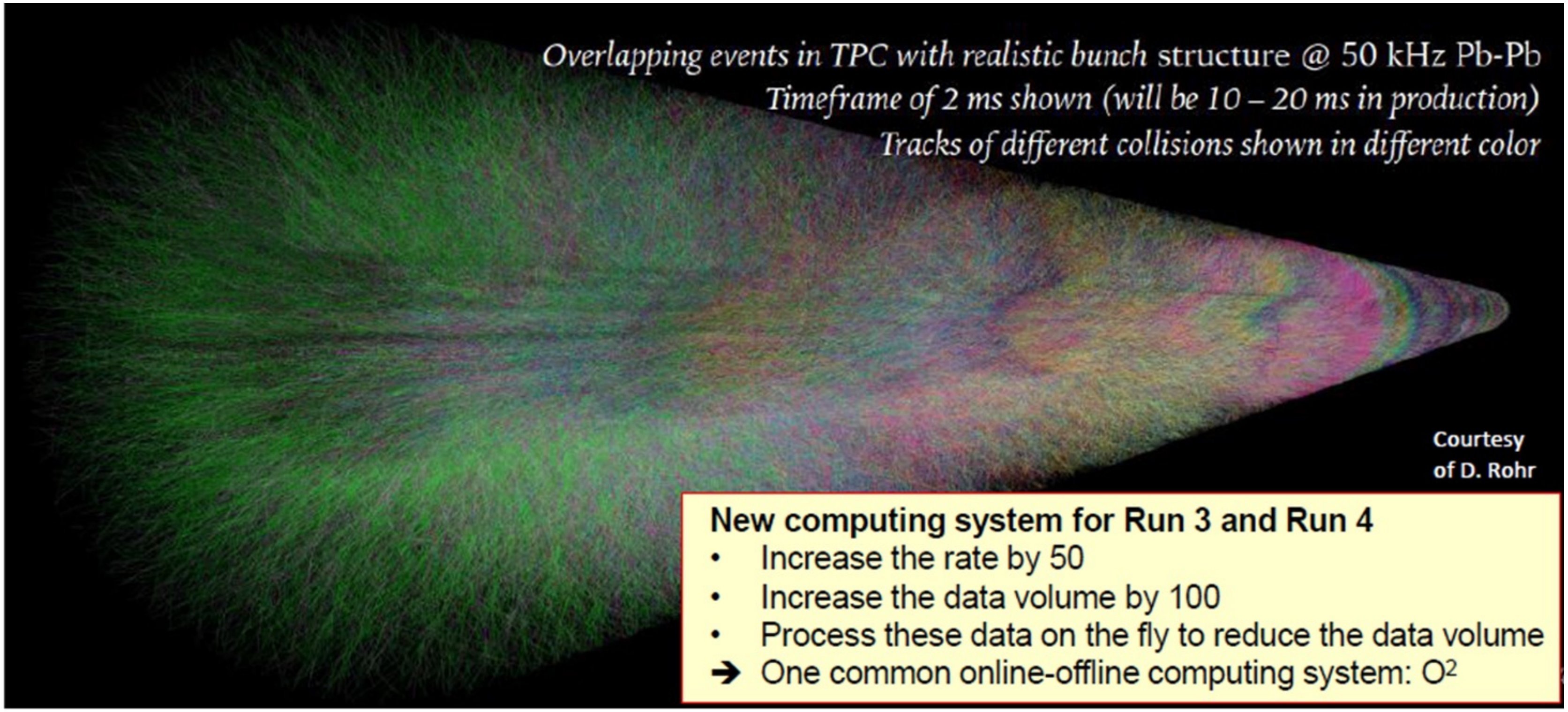
- More detectors participating to the trigger
- Increase latencies
- More complex offline-like algo, both in hardware and software: Machine Learning!
- Get rid of the hardware level
- Higher rate (of raw data) on disk

Profit from the event reconstruction / data reduction done by the trigger system

- Permanent store reconstructed objects, trading event size with event rate
- Enabled by fast online calibration procedures
- "Real Time Analysis" both online (producing aggregated results) and offline

ALICE continuos reading

- Heavy-Ions program entails very low S/B, online selection is essentially impossible → need large statistics!
- LHC will provide 50 kHz of min-bias Pb-Pb collisions (x100 w.r.t. run 2)
- Continuous read-out of all detector data, with reduced data volume to disk

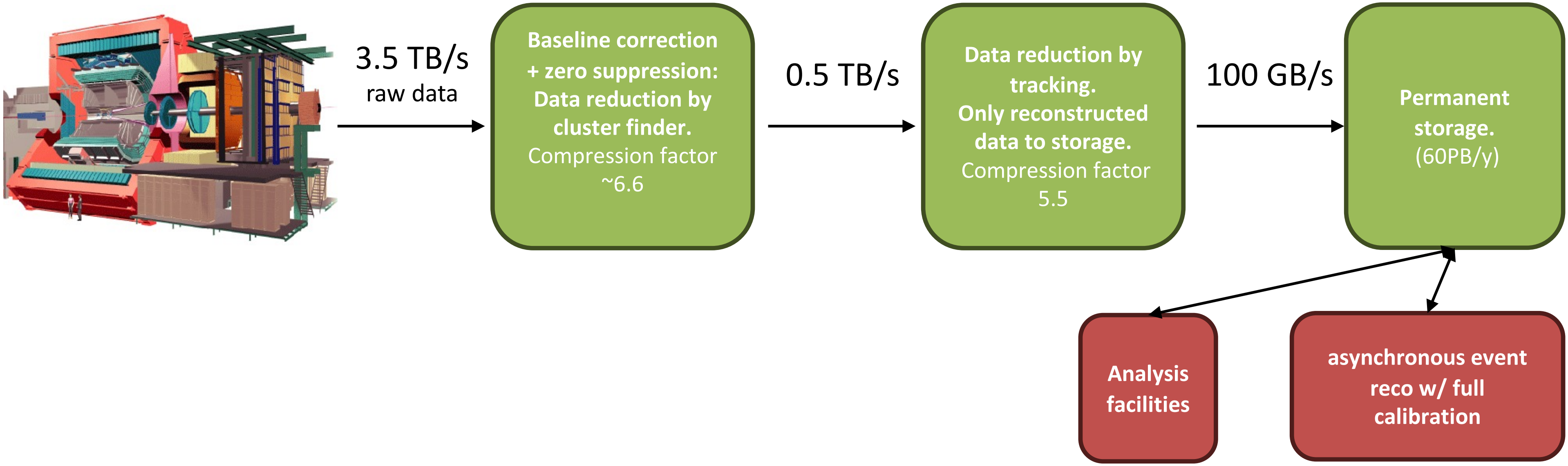


ALICE continuos reading

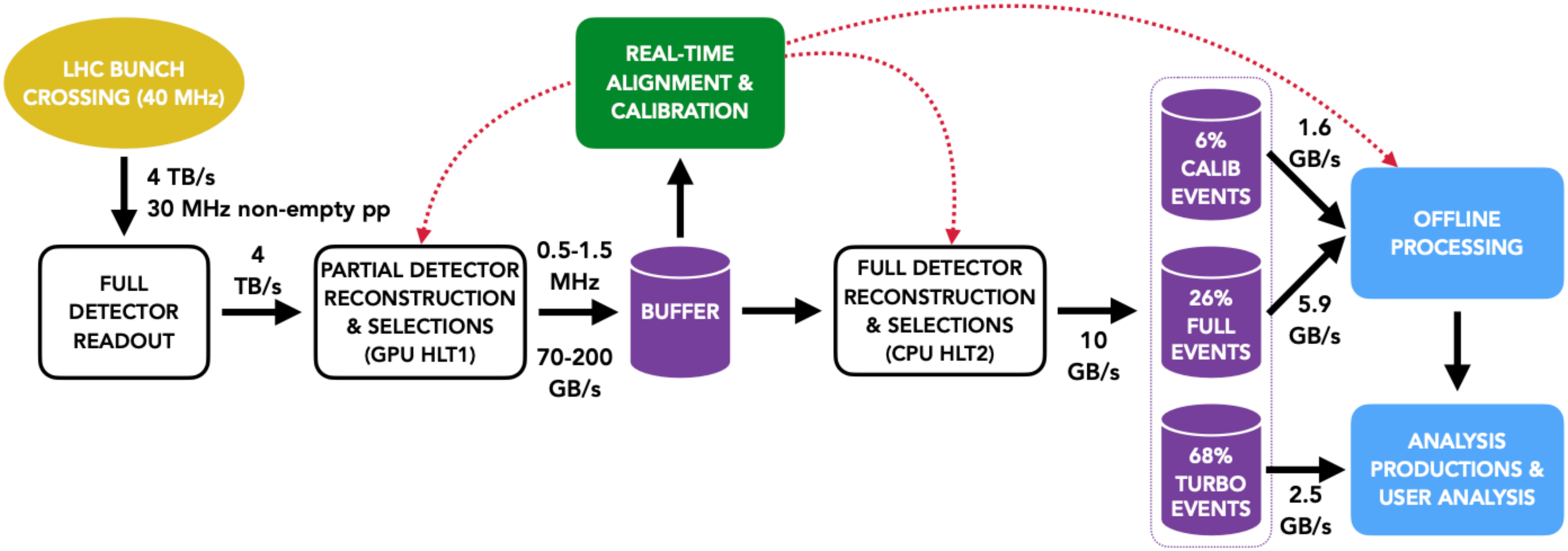
- Heavy-Ions program entails very low S/B, online selection is essentially impossible → need large statistics!
- LHC will provide 50 kHz of min-bias Pb-Pb collisions (x100 w.r.t. run 2)
- Continuous read-out of all detector data, with reduced data volume to disk

ALICE O^2 system

(common online-offline system)



LHCb Real Time Analysis

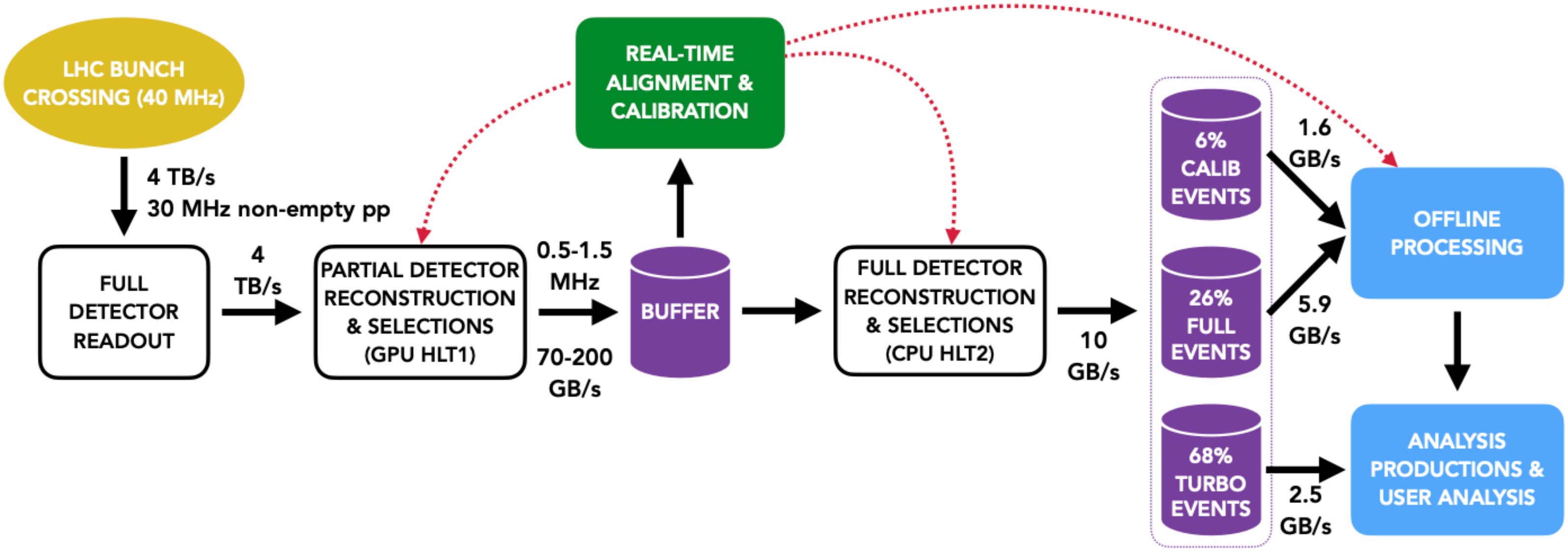


Hardware-based L0 would not allow profiting from higher luminosity → full detector readout (4 TB/s) + 2 levels of software-based trigger

HLT1

- performs a fast reconstruction to identify inclusive beauty and charm signatures as well as high p_T muons at 30 MHz
- partial event reconstruction (tracking) on GPUs hosted in the event-building servers
- select events with 2-tracks vertices, displaces single tracks, muons (configurable)
- Output rate ~1 MHz onto a ~30PB disk buffer (80h)

LHCb Real Time Analysis

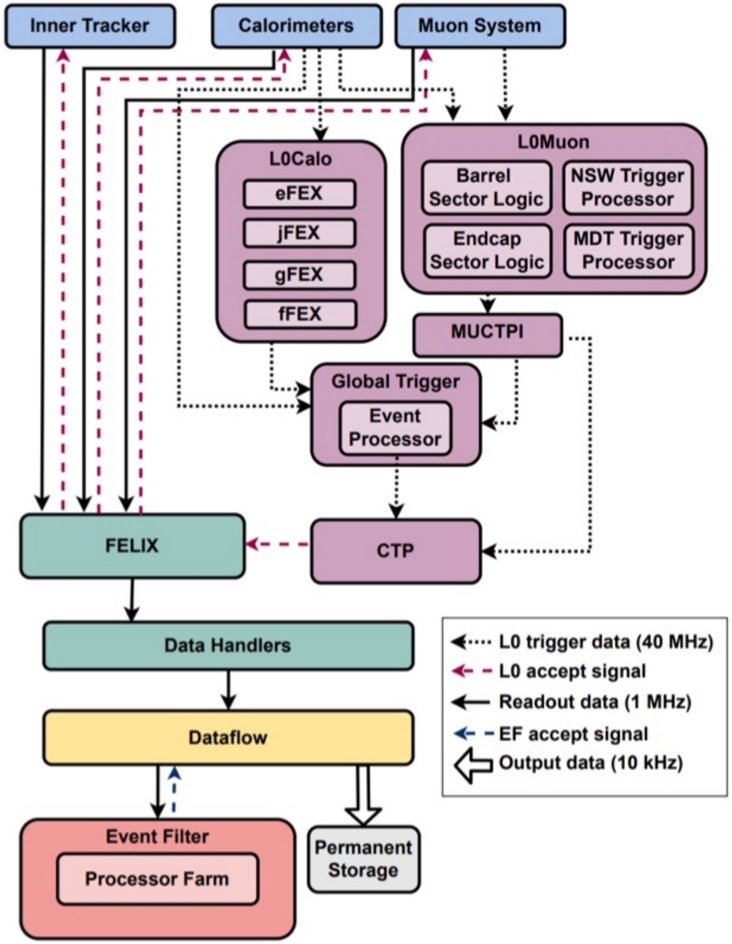


Hardware-based L0 would not allow profiting from higher luminosity → full detector readout (4 TB/s) + 2 levels of software-based trigger

HLT2

- provided with real-time computed calibration constants, performs full event reconstruction on CPU and possibly on accelerators
- run O(1000) algorithms selecting relevant signal events and a few inclusive topologies
- outputs 10 GB/s, split into dedicated streams (majority to “Turbo”, with high level event content)

ATLAS TDAQ for HL-LHC



L0

- 40MHz → 1MHz, latency of 10 μs
- Full granularity data from Calorimeters
- Data from all muon detectors
- Global Trigger: offline-like algorithms to combine/refine regional candidates running on farm of common HW

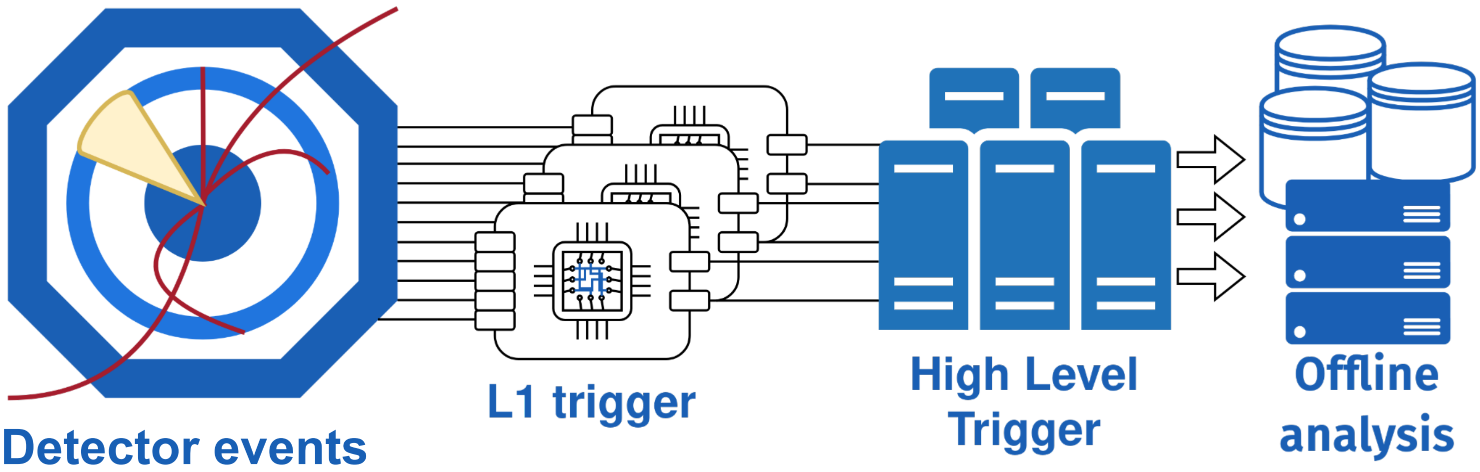
Event Filter Trigger

- 1MHz → 10kHz
- Full event building dealing with throughput of 5 TB/s
- Farm of heterogeneous commodities processors running offline-like algos on accelerators too (GPU/FPGA)

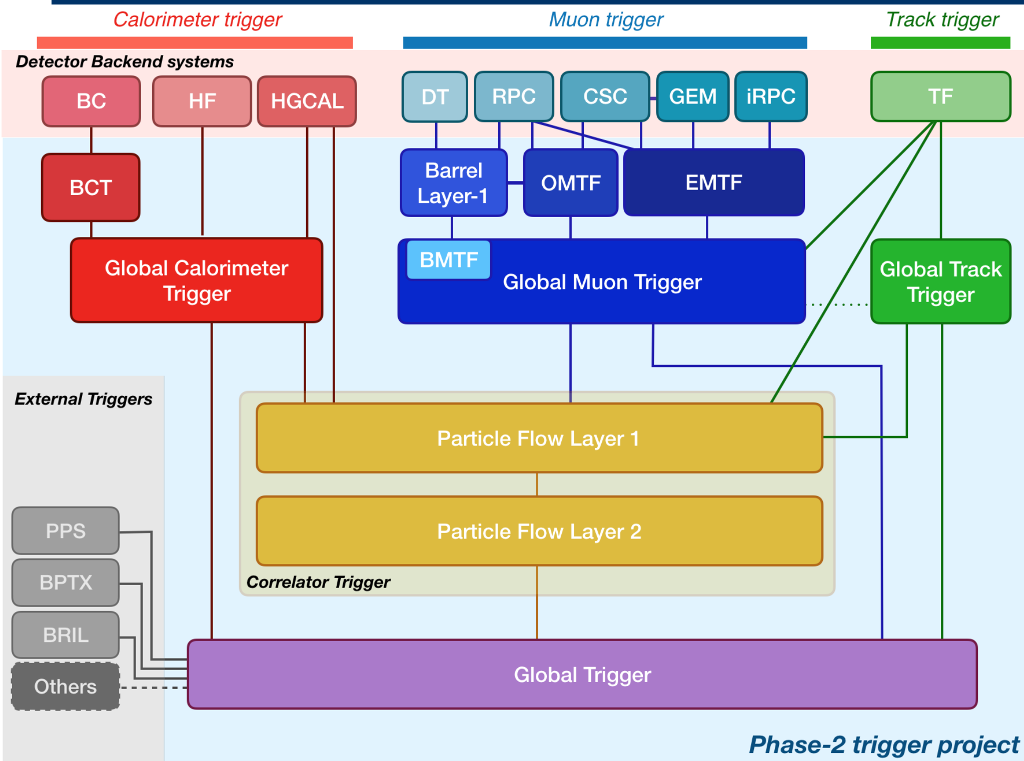
CMS Trigger for HL-LHC

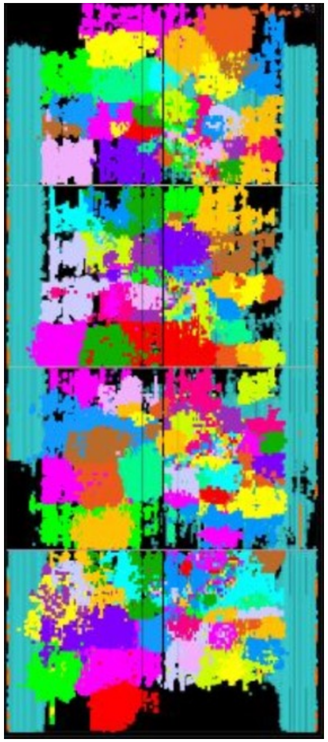
- Maintain the 2 levels structure:
  - L1: custom FPGA boards (ATCA)
  - HLT: farm of servers with CPUs and GPUs
- Updated figures for rates, throughput, latency

CMS detector	LHC Phase-1	HL-LHC Phase-2	
Peak ⟨PU⟩	60	140	200
L1 accept rate (maximum)	100 kHz	500 kHz	750 kHz
Event Size at HLT input	2.0 MB [a]	6.1 MB	8.4 MB
Event Network throughput	1.6 Tb/s	24 Tb/s	51 Tb/s
Event Network buffer (60 s)	12 TB	182 TB	379 TB
HLT accept rate	1 kHz	5 kHz	7.5 kHz
HLT computing power [b]	0.7 MHS06	17 MHS06	37 MHS06
Event Size at HLT output [c]	1.4 MB	4.3 MB	5.9 MB
Storage throughput [d]	2 GB/s	24 GB/s	51 GB/s
Storage throughput (Heavy-Ion)	12 GB/s	51 GB/s	51 GB/s
Storage capacity needed (1 day [e])	0.2 PB	1.6 PB	3.3 PB



CMS Trigger for HL-LHC



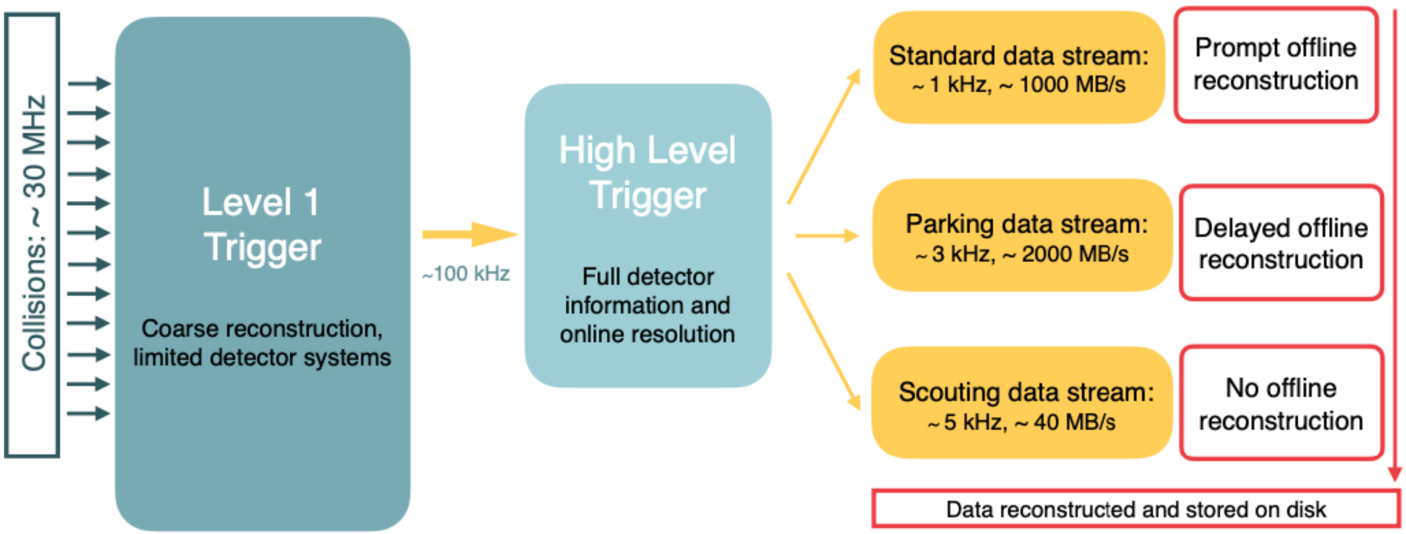


400 algos on 1 FPGA

- Major upgrade of the L1 system → all detectors (but silicon-pixel detector) participating in providing primitives!
- Information from subsystems combined in an offline-like fashion by the Trigger Correlator:
  - Particle Flow, PU mitigation; ubiquitous usage of ML
- Physics objects quality comparable to those of HLT/offline

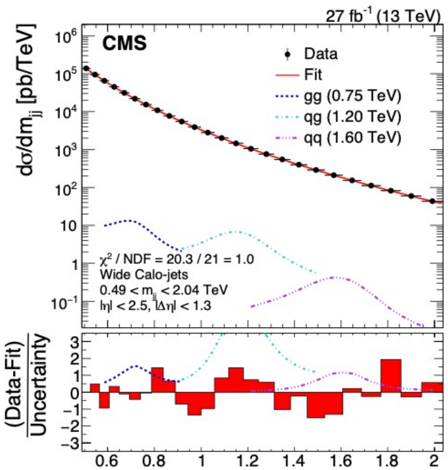
Enriching the Physics Program

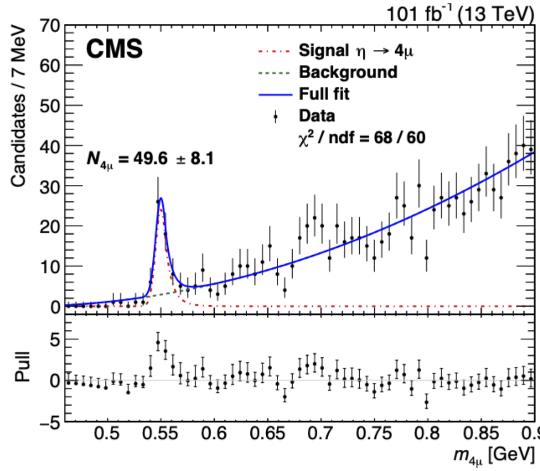
- Software-based trigger layers provide physics objects with accuracy comparable to offline → exploit those to extend phase space to regions filtered out by online selections
- Store dedicated data streams with high level info only
- Several measurements performed and published both by ATLAS and CMS



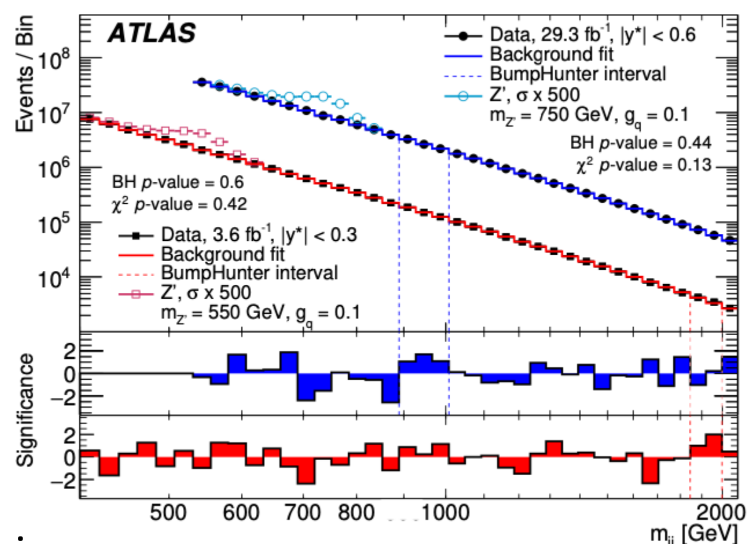
Enriching the Physics Program

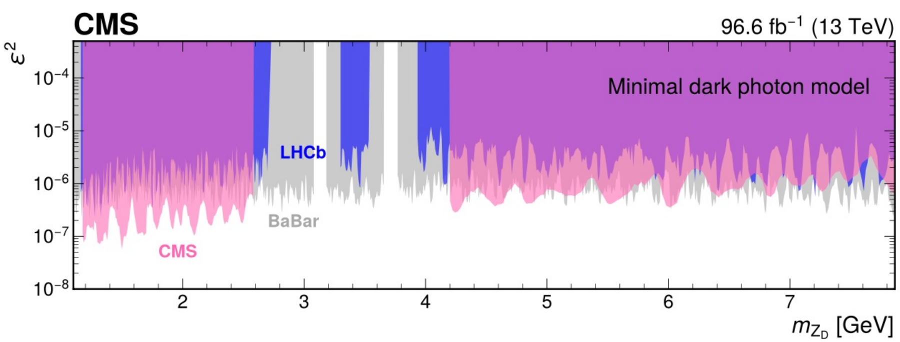
- Software-based trigger layers provide physics objects with accuracy comparable to offline → exploit those to extend phase space to regions filtered out by online selections
- Store dedicated data streams with high level info only
- Several measurements performed and published both by ATLAS and CMS





[Phys. Rev. Lett. 121, 081801](#)

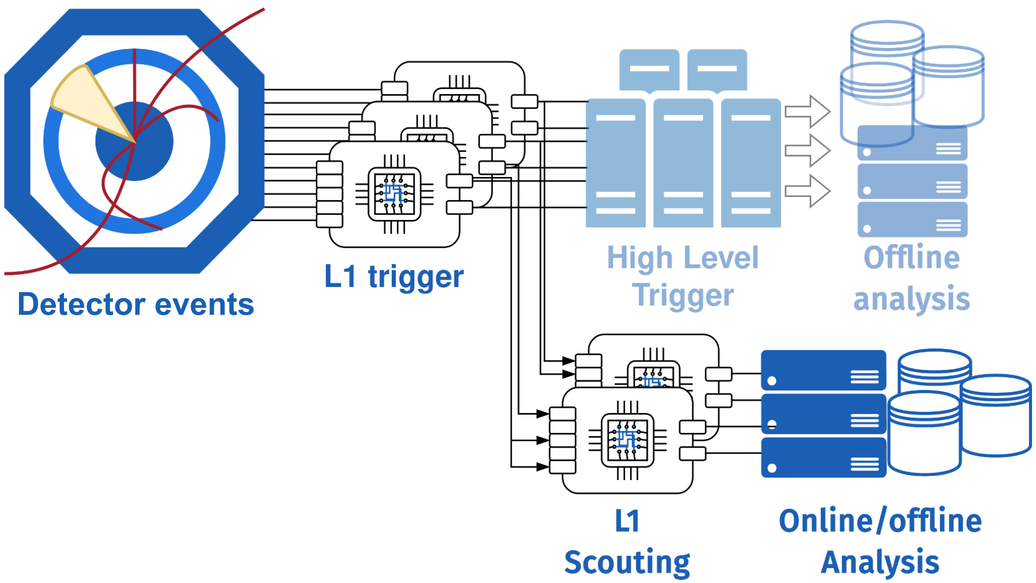




CMS review submitted to [Phys. Rept.](#)

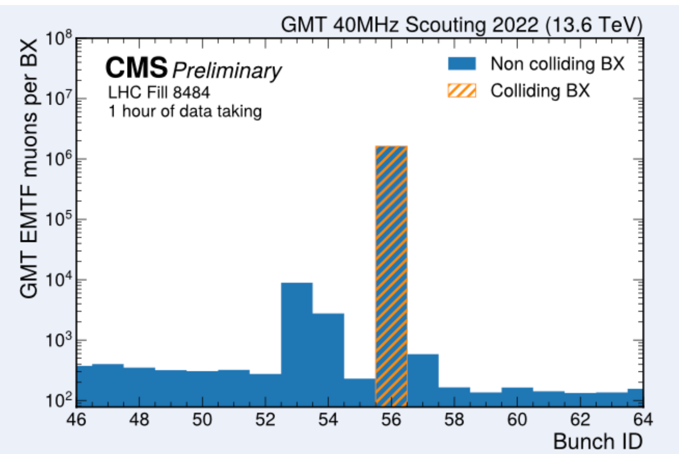
CMS L1 Scouting

- Push further this by gathering @ 40MHz L1 objects and use them for physics analysis → potentially unprescaled Zero Bias sample
- Tradeoff between rate and resolution. Proof of principle now in production
  - Interesting analyses could be addressed anyhow, e.g. dijet resonances
- Full potential of the system expected for HL-LHC w/ high resolution L1 objs
- Address both extensions to inaccessible topologies (e.g. $H \to \phi\phi \to 4K$, $W \to 3\pi$) and fully unbiased anomaly searches
  - Correlation with other bunch crossing also possible (looong lived particles)

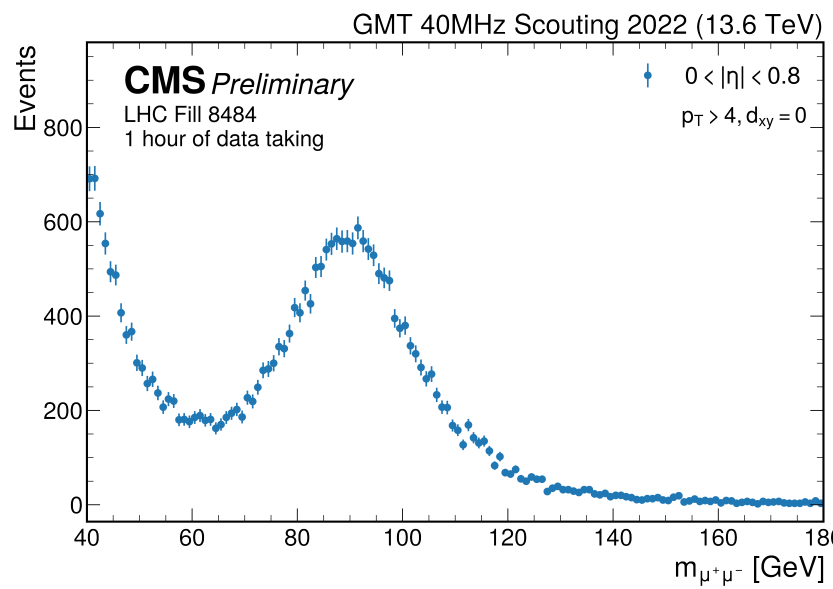


CMS L1 Scouting

- Push further this by gathering @ 40MHz L1 objects and use them for physics analysis → potentially unprescaled Zero Bias sample
- Tradeoff between rate and resolution. Proof of principle now in production
  - Interesting analyses could be addressed anyhow, e.g. dijet resonances
- Full potential of the system expected for HL-LHC w/ high resolution L1 objs
- Address both extensions to inaccessible topologies (e.g. $H \to \phi\phi \to 4K$, $W \to 3\pi$) and fully unbiased anomaly searches
  - Correlation with other bunch crossing also possible (looong lived particles)



Global Muon Trigger (GMT) muons occupancy per bunch crossing (BX) within an LHC orbit for muons reconstructed by the Endcap Muon Track Finder (EMTF) in the BX range [46,64]. The single isolated colliding bunch (BX 56) is highlighted by the orange transversal lines [CMS-DP-2023-025].



Anomaly Detection, Data Compression

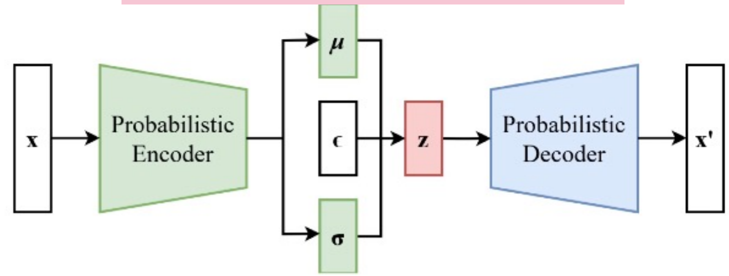
Courtesy of G. Grosso

Anomaly Detection (& Data Compression)

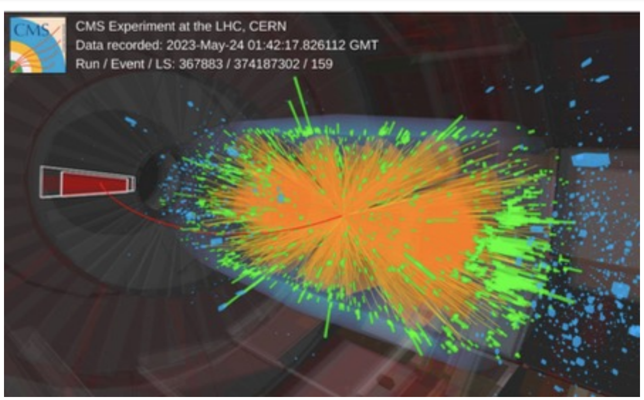
- Variational AutoEncoders used to "compress" input into a latent space; such representation can be used as metric to classify standard vs anomalous data
- AXOL1TL: use VAE to trigger on anomalous events

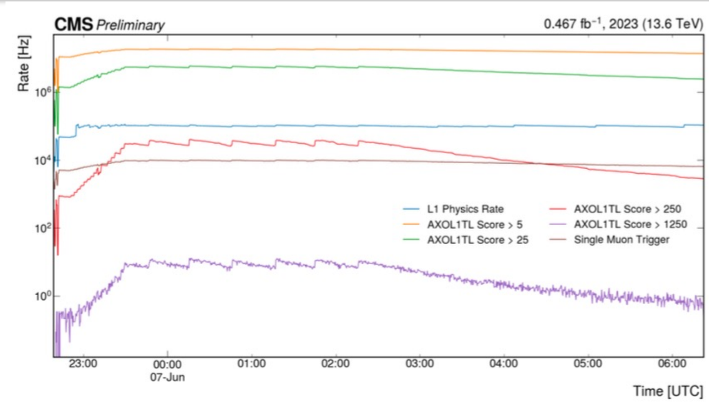
$$\mathscr{L} = (1-\beta)\,||x-\hat{x}||^2 + \beta\,\frac{1}{2}(\mu^2+\sigma^2-1-\log\sigma^2)$$

reconstruction latent representation



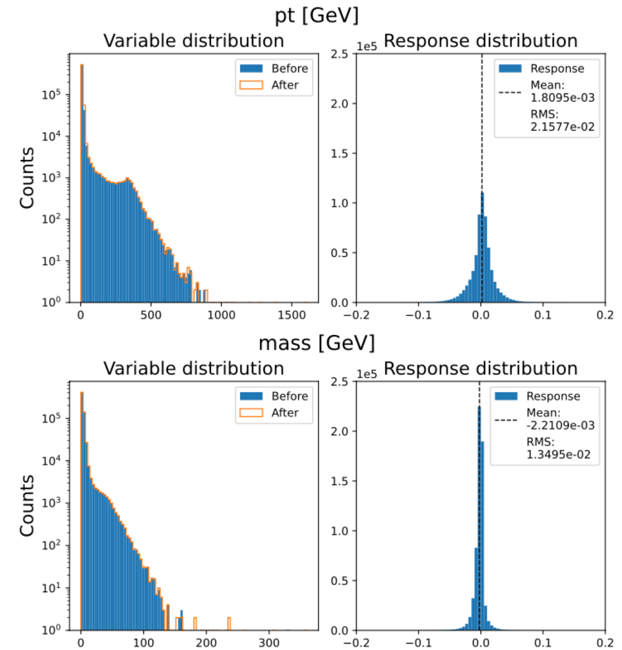
- Train the model on L1 objects in a large Zero Bias datasets
- It results that μ is the enough to discriminate anomalous events.
- In production at CMS since a couple of weeks!



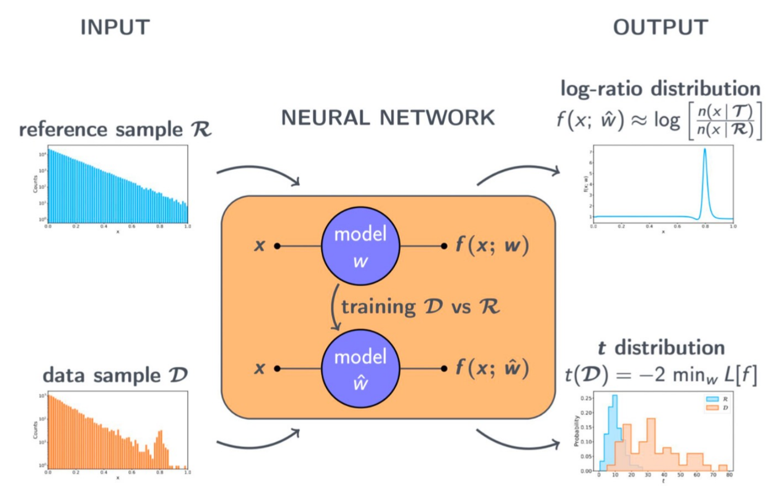


Anomaly Detection (& Data Compression)

- The (lossy) compression provided by VAE could be used as an efficient way to gain bandwidth to disk
- Attempts to evaluate faithfulness of the method are ongoing (e.g. Baler)



- In addition to seeking single anomalous events, statistical anomalies can be addressed
- Run anomaly detection algos on batches of data online (e.g. NPLM), storing statistics parameters to let potential anomalies building up



Summary

- Trigger systems are critical enablers of the LHC experiments' physics programs at the incoming High Luminosity phase
- Tremendous technological advancements in telecom, electronics, algoritms will allow processing huge data flows with offline-comparable accuracy
- Boost in sensitivity for most targeted signals, several ideas being developed for being prepared for the unexpected