Trigger performance (including data scouting and GPU) at CMS and ATLAS

LHCP 2024 – Boston

Silvio Donato (INFN Pisa)
on behalf of the CMS and ATLAS collaboration

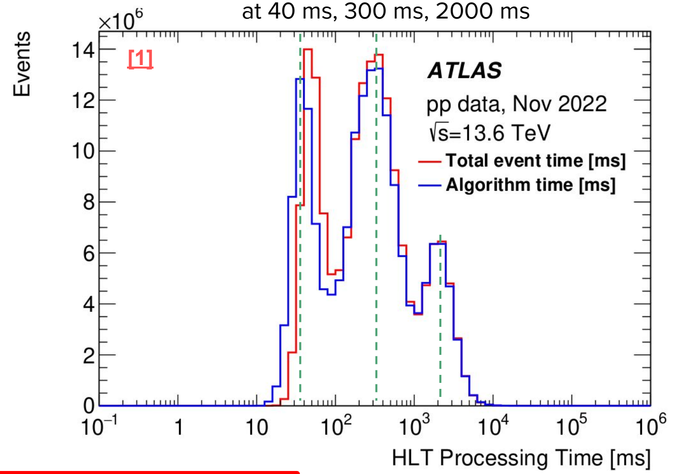
Outline

- Introduction: the ATLAS and CMS trigger system
- Delayed reconstruction/parking
- Trigger-level analysis/scouting
- Objects performance
  - Muons, electrons, photons, jets, MET, b-tagging, tau, tracking, long-lived particle
- Multithreading and GPU reconstruction
- Conclusions

The CMS and ATLAS trigger system

- **LHC** bunch crossing rate: **~30 MHz**
- **Hardware** trigger (**L1**): ~30 MHz ➔ ~**100 kHz**
  - simplified readout (no tracker), small latency (<4 µs).
- **Software** trigger (**HLT**): ~100 kHz ➔ **few kHz**.
  - full event readout available (~1 MB/event);
  - HLT farm with ~50k threads ➔ ~500ms/event on average
  - Events are rejected in the early stage of the reconstruction
- **Storage and offline reconstruction**

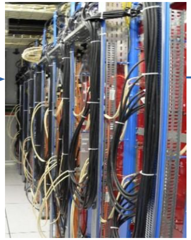
HLT processing time peaks
at 40 ms, 300 ms, 2000 ms



Simplified readout

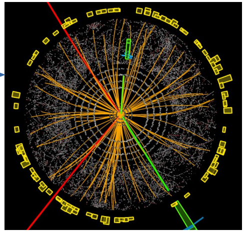
~30 MHz

L1 trigger



~100 kHz
(<4 µs)

Full readout



~100 kHz
~100 GB/s

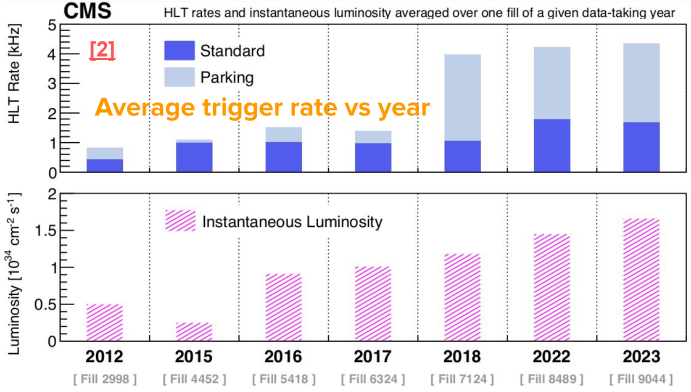
HLT

few kHz
(~500 ms)
~5 GB/s

Data storage

Delayed reconstruction/parking

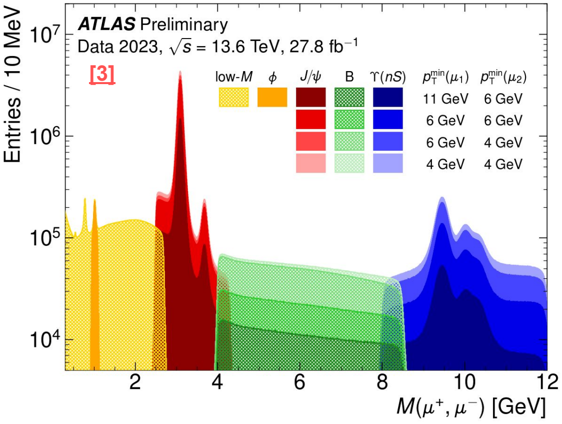
- The "standard" trigger cross section ~**100 nb**
  - eg. Lumi **1×10^{34}cm^{-2}s^{-1}** → Rate: **1 kHz**.
- Main bottleneck of rate: prompt offline reco.
- **Delayed reconstruction** (ATLAS) and **parking** (CMS) to bypass the rate limit.
- In 2018 CMS collected 10B events of displaced single muon
- Expanded strategy for Run-3 in ATLAS and CMS
  - larger trigger rate;
  - many different final states covered.



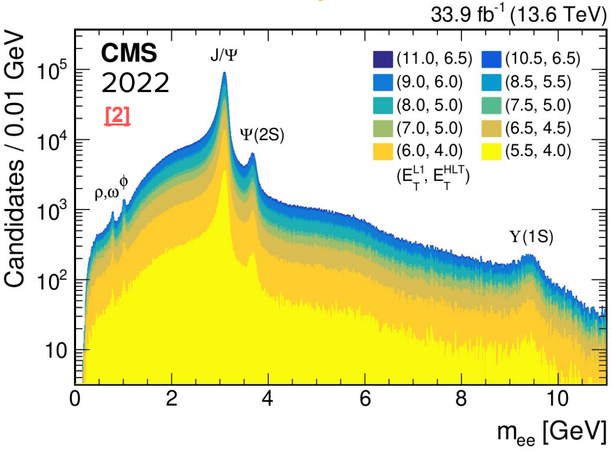
Trigger selection	ATLAS rate [1] (at 1.8×10^{34} cm^{-2}s^{-1})	CMS rate [2] (at 2.0×10^{34} cm^{-2}s^{-1})
2 muons (B-physics)	40 Hz	1.6 kHz
2 electrons (B-physics)	170 Hz	1.3 kHz (only 2022)
Vector-boson fusion	270 Hz	1.2 kHz (since 2023)
HH (2 jets + 2 b-jets)	160 Hz	180 Hz (since 2023)
6 jets	140 Hz	-
5 jets + 1 b-jet	50 Hz	-
Long-lived particle	-	150 Hz (since 2023)

Delayed reconstruction/parking

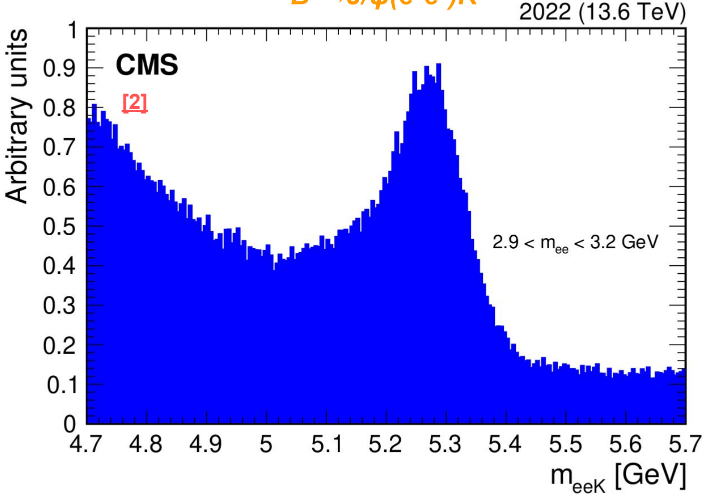
Dimuon spectrum



Dielectron spectrum

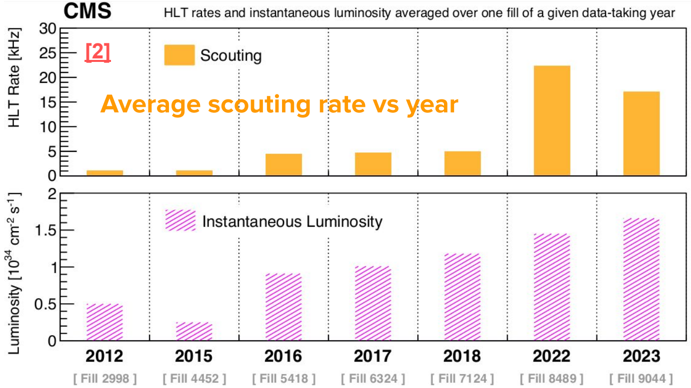


$B^+ \rightarrow J/\psi(e^+e^-)K^+$



Trigger-level analysis/scouting

- Trigger-level analysis (ATLAS) or scouting (CMS) strategy: save directly **trigger objects**
  - Event size around **10 kB/event** instead of ~1 MB/event,
- Important evolution since Run-1:
  - **Rate** increased is to **8**-**20** **kHz**:
    - Multijet, muons, electron/photons, …
  - **All main physics objects** reconstructed:
    - Photons, jets, tracks, b-tag (ATLAS), muons, electrons, PF candidates (CMS)
    - Multiple collections stored in the same event.
  - Different or same event content for different streams.

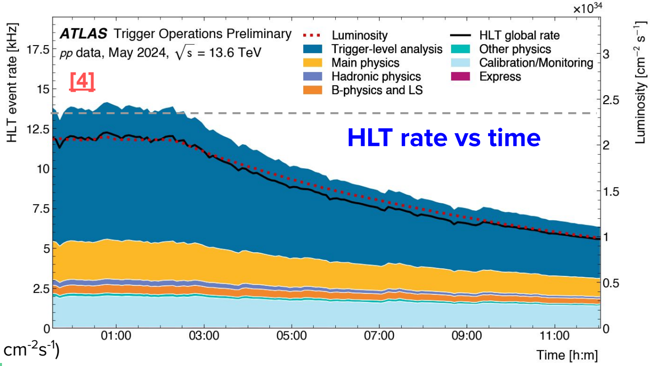


TLA:
8.5 kHz

Main:
2.5 kHz

Delayed:
1 kHz

Total:
~14 kHz
(at $1.7 \cdot 10^{34}$ cm^{-2}s^{-1})

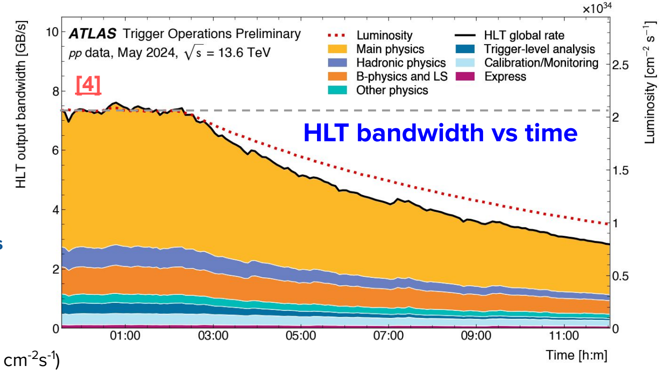


Main:
5.0 GB/s

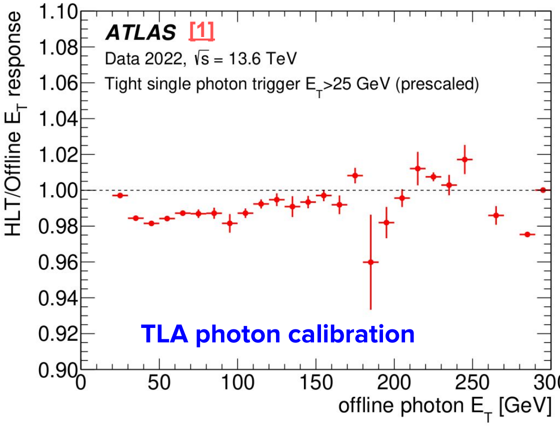
Delayed:
1.5 GB/s

TLA:
0.03 GB/s

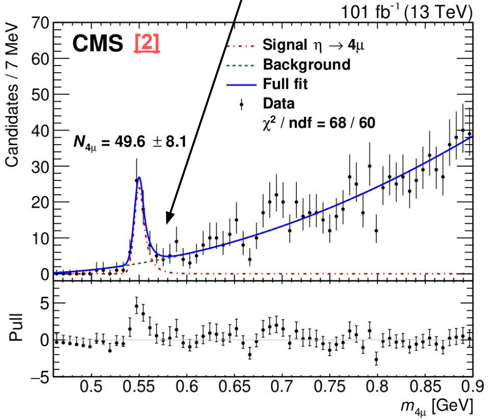
Total:
~7.5 GB/s
(at $2.1 \cdot 10^{34}$ cm^{-2}s^{-1})

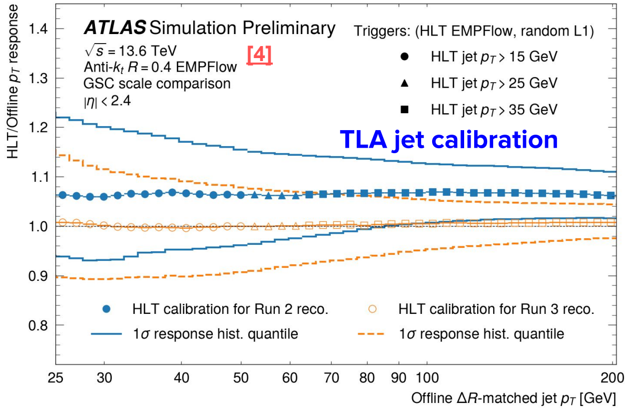


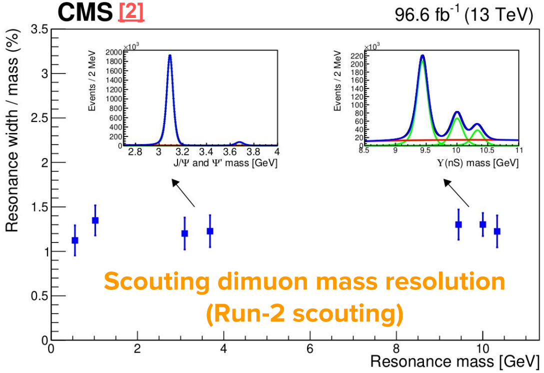
Trigger-level analysis/scouting

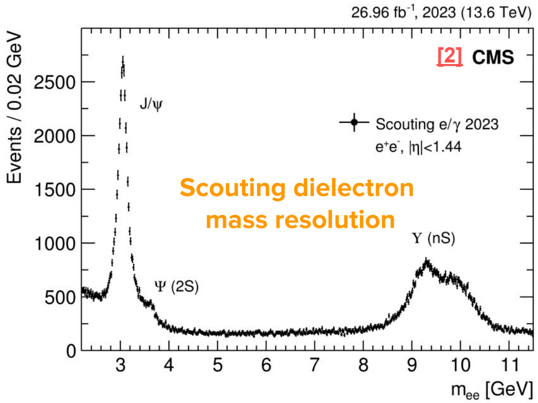


Observation of $\eta \to 4\mu$ (Run-2 scouting)



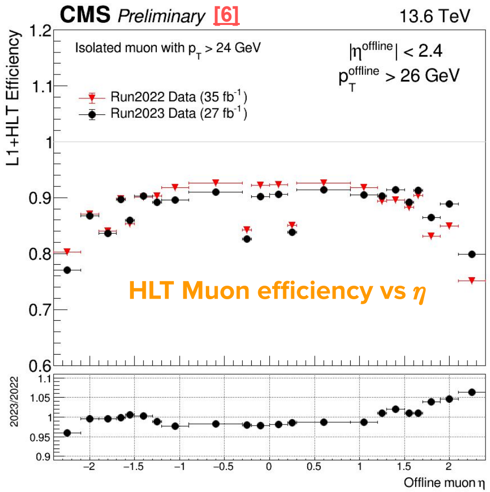


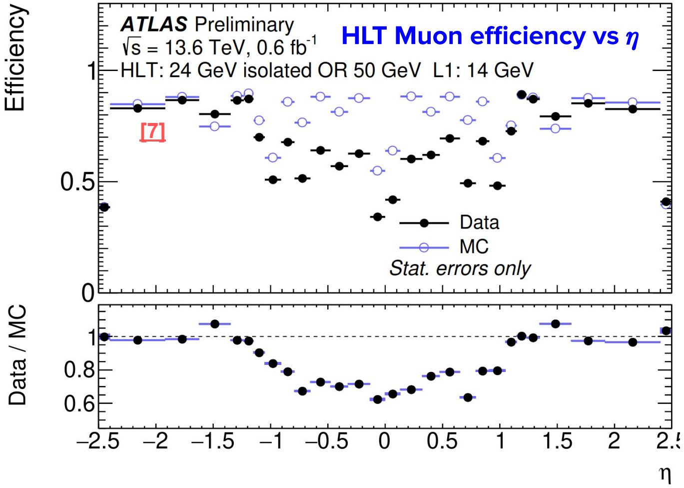


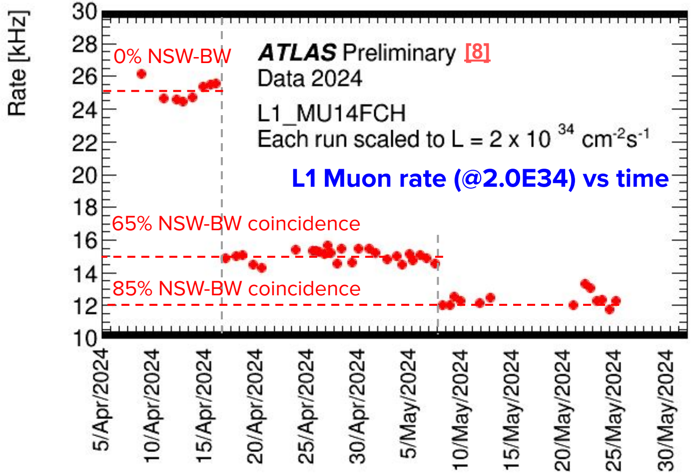


Muons

- Muon efficiency dominated by L1 trigger and isolation cut.
- L1 muon chamber inefficiency recovered during data taking.
- New Small Wheels (ATLAS) improved efficiency/rate ratio in the forward region.
    - Rate reduction: > -50% (13 kHz), with ~98% efficiency.

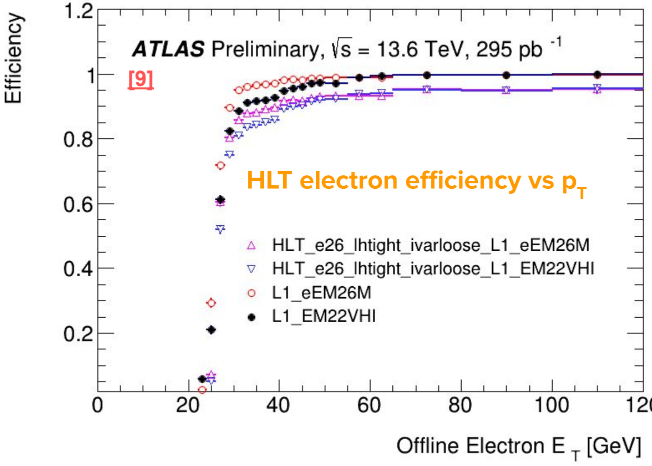


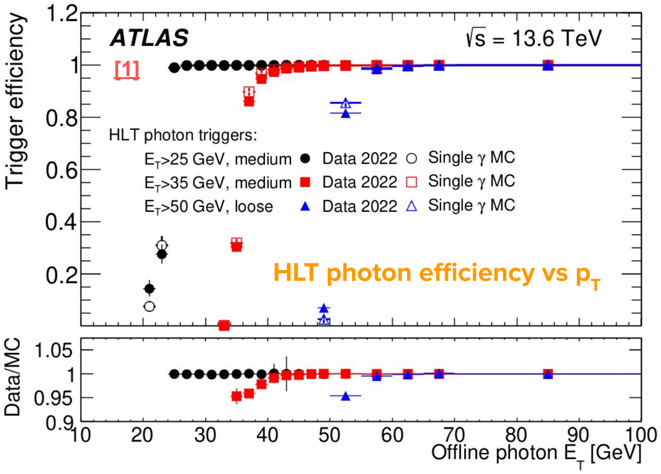


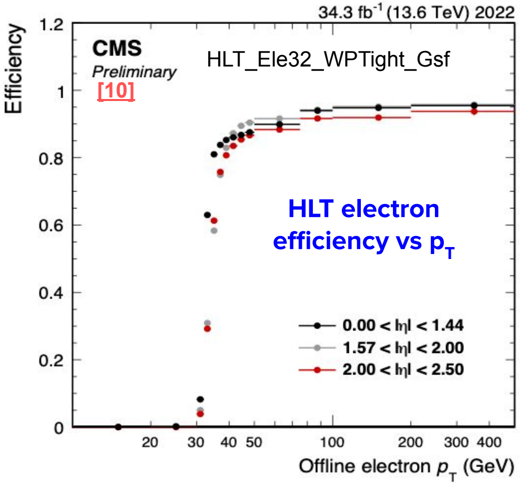


Electrons and photons

- Excellent performance
- New Phase-1 algorithm in ATLAS in L1 trigger ➔ better efficiency

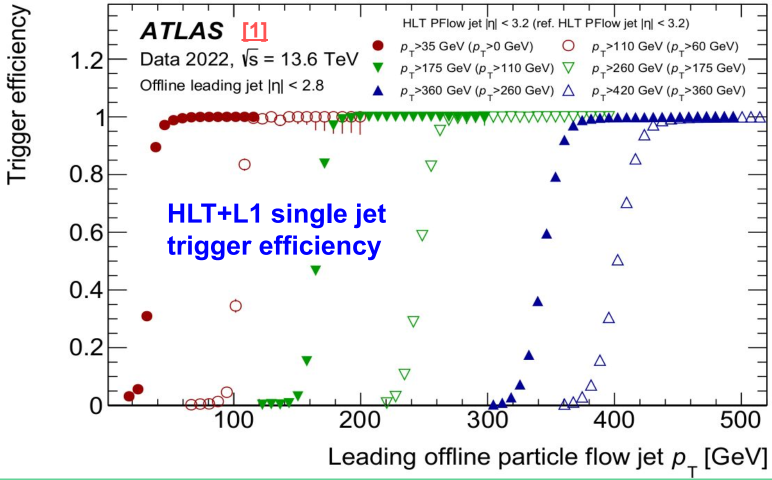


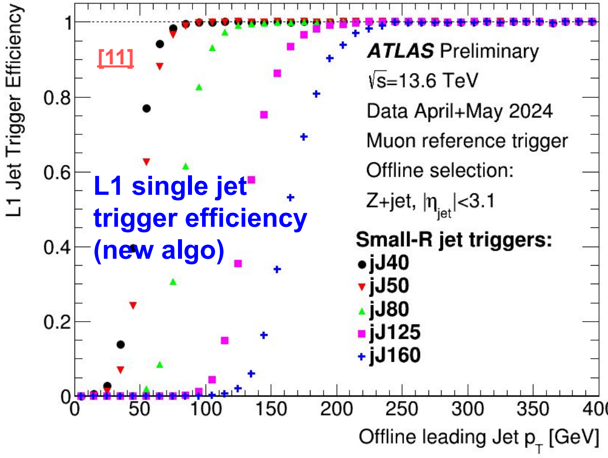


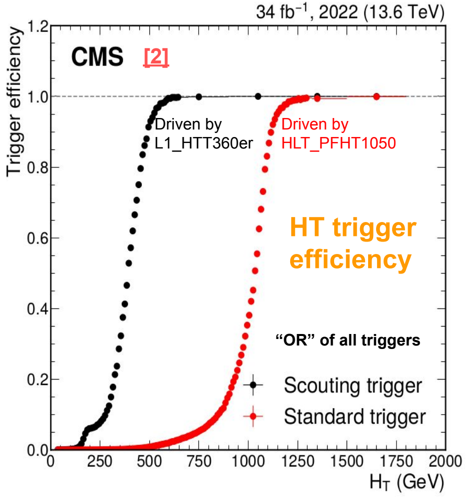


Jet and HT

- Good Jet/HT performance.
- Scouting/TLA allows a large gain in trigger acceptance.
  - Larger gain with the activation of L1_HTT280er in 2023
- New Phase-I jet triggers
  - jet Feature Extractor (jFEX) applies a more refined jet calibration than the legacy L1 jets received

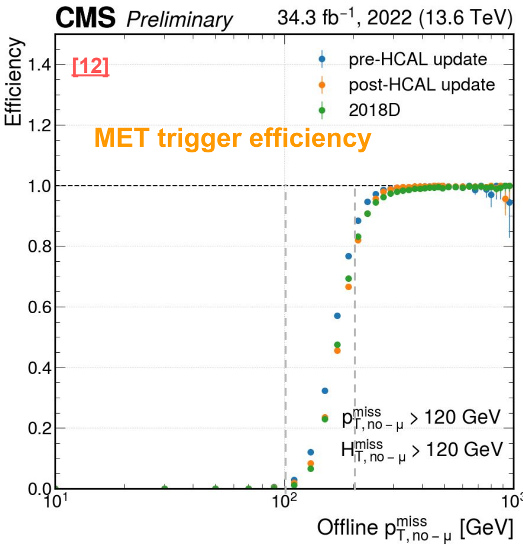


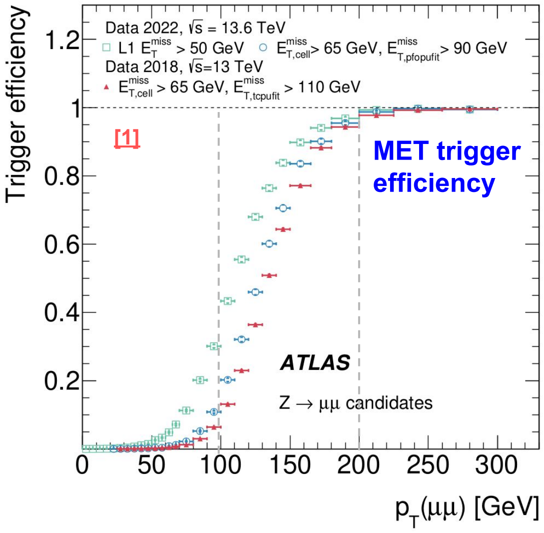


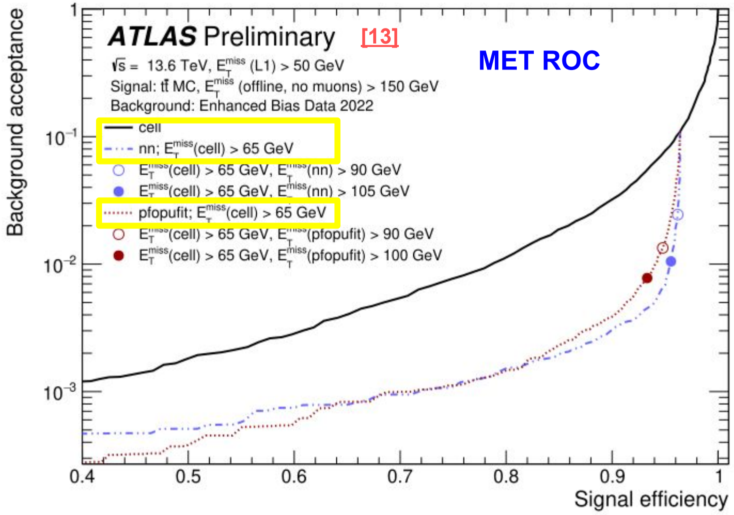


Missing transverse energy

- Missing transverse energy computed as the sum of particle flow candidate
- New method based on NN deployed by ATLAS in 2024
    - ➡ improved efficiency at fixed rate

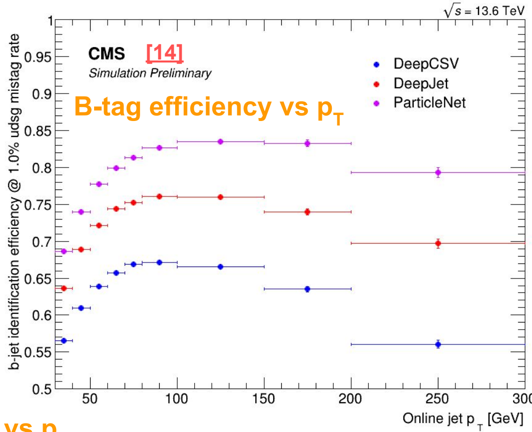




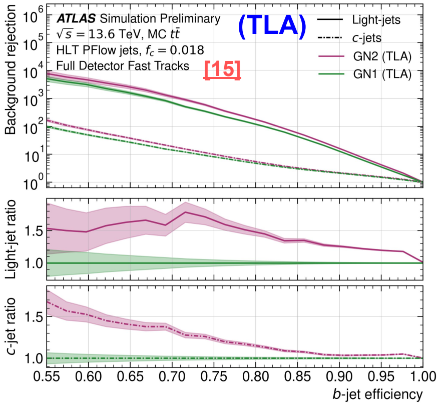


B-tagging and tau tagging

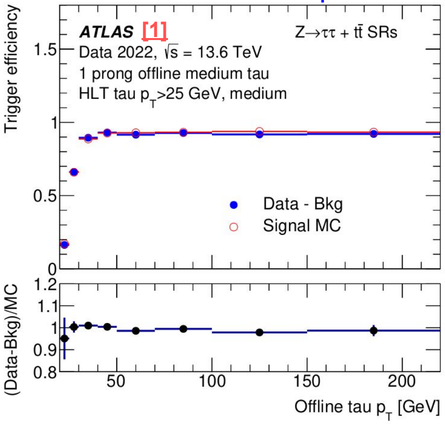
- Graph neural network used for b-tagging (ParticleNet, GN2)
- Large improvement in performance
- GN2 used in TLA



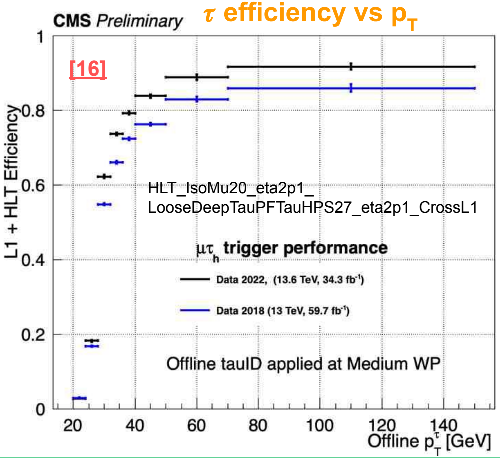
B-tag efficiency vs uds rejection



τ efficiency vs p_T



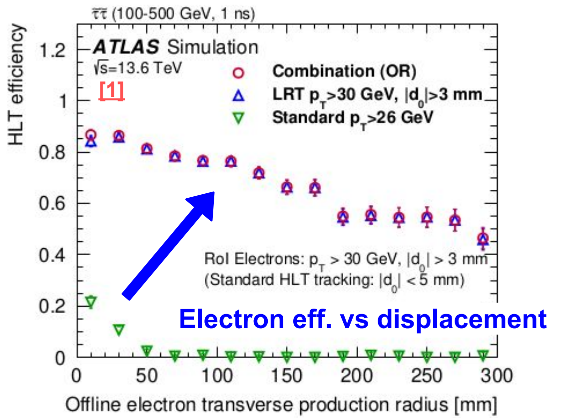
τ efficiency vs p_T

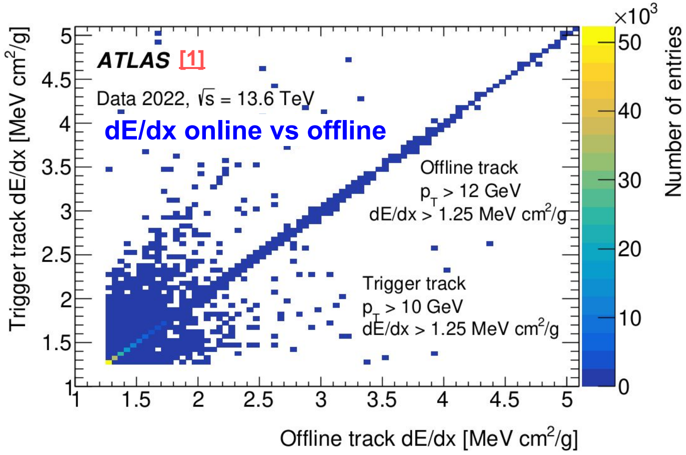


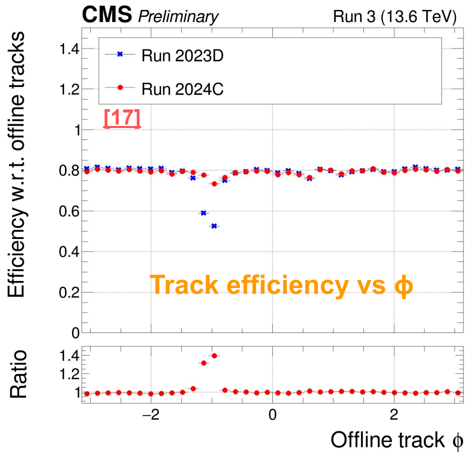
- Good performance in tau reconstruction
- Migration of tau reconstruction to ParticleNet in 2024 (CMS)

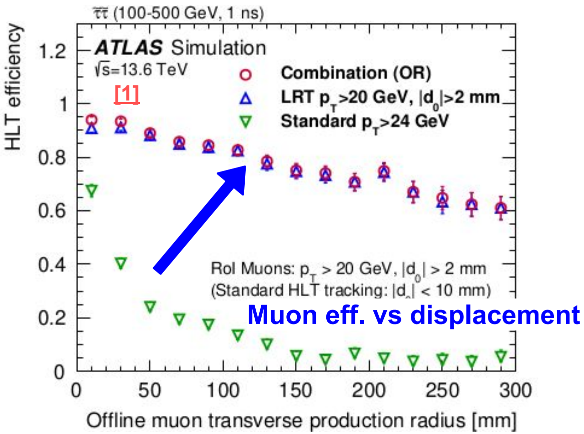
Tracking

- Excellent precision in dE/dX measurement
- Issues in few pixel modules in CMS after TS1 in 2023
  ➔ recovered using a doublet recovery in 2024
- Development of dedicated tracking for long-lived particles



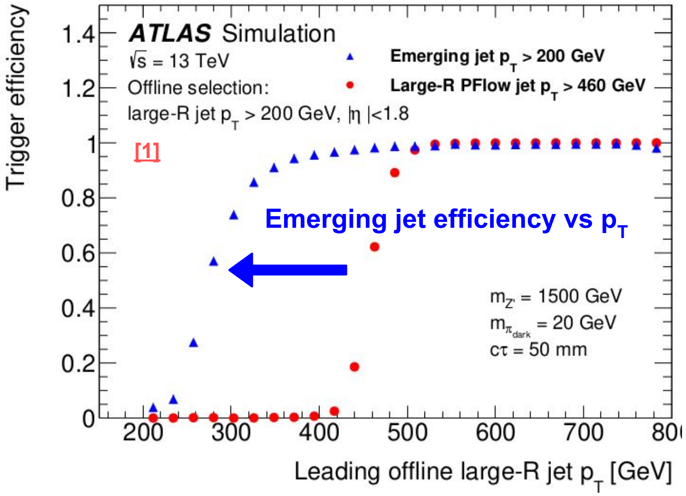


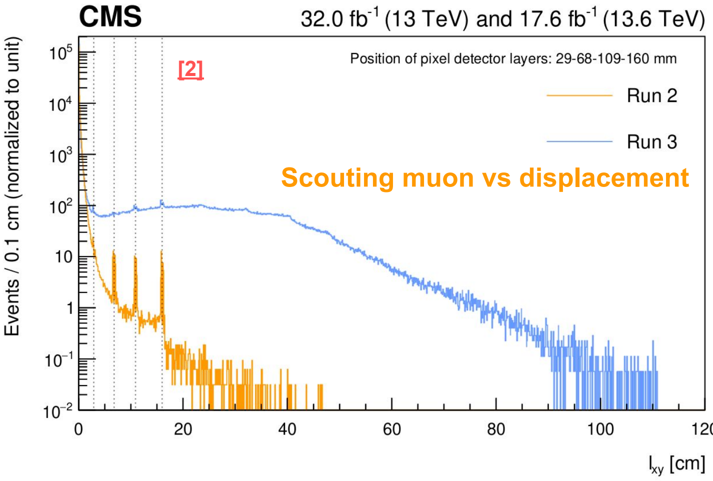




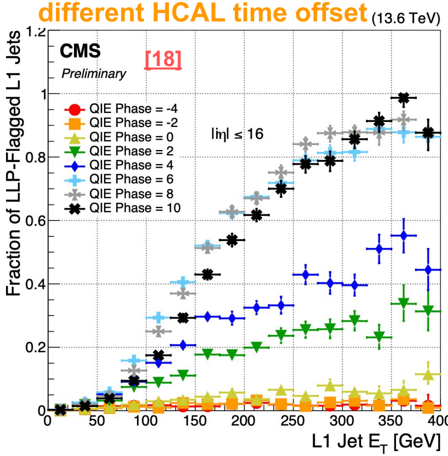
Long lived particles

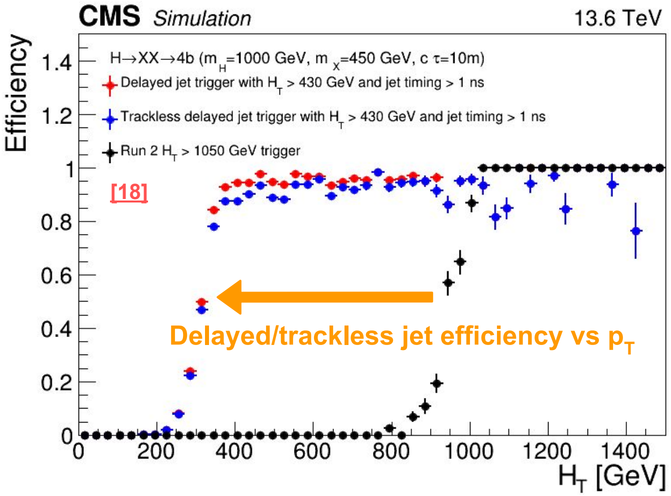
- New set of triggers targeting long-lived particles
  - Trackless or displaced jets;
  - Measurement of time delay in ECAL and HCAL;
  - Displaced muons
    - Dedicated L1 trigger
    - Included in scouting





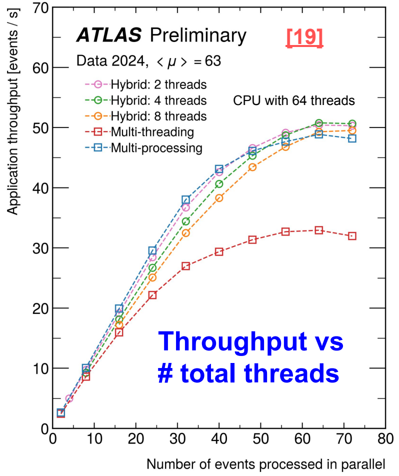
LLP jet efficiency vs p_T for different HCAL time offset

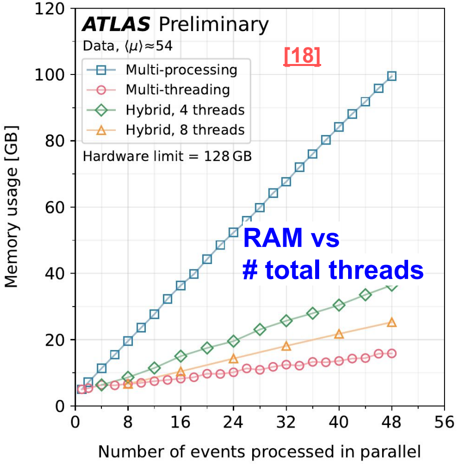


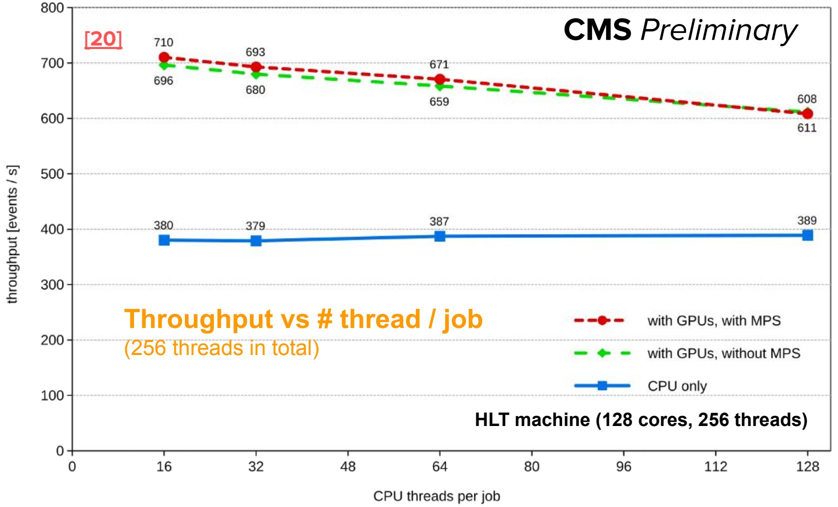


Multithreading and GPU

- Multithreading (MT) is key to fully exploit HLT farm computational power
  - inter-event, intra-event, in-algorithm parallelism;
  - usage of “data handles” to define the data dependency among modules;
  - lower memory usage.
  - AthenaMT online since 2022.
  - CMSSW support MT since 2015.
- CMS HLT farm heterogeneous since 2022 (AMD CPU + Nvidia T4):
  - **40%** of HLT reconstruction ported to **GPU** (CUDA)
    - Pixel local reconstruction
    - Pixel tracking and vertexis
    - ECAL local reconstruction
    - HCAL local reconstruction

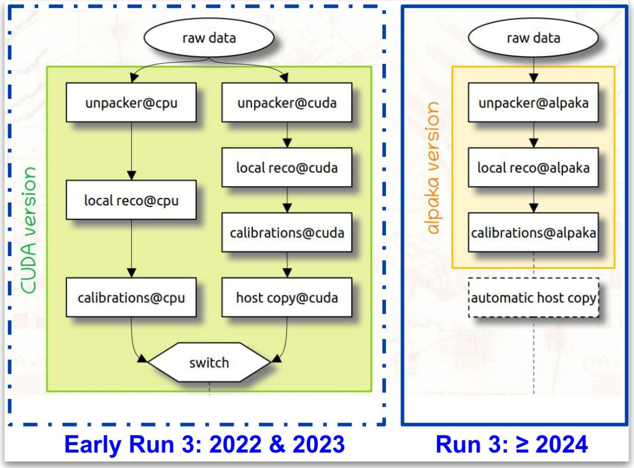


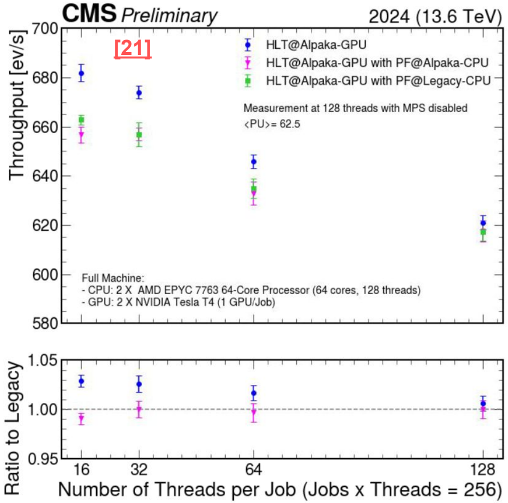


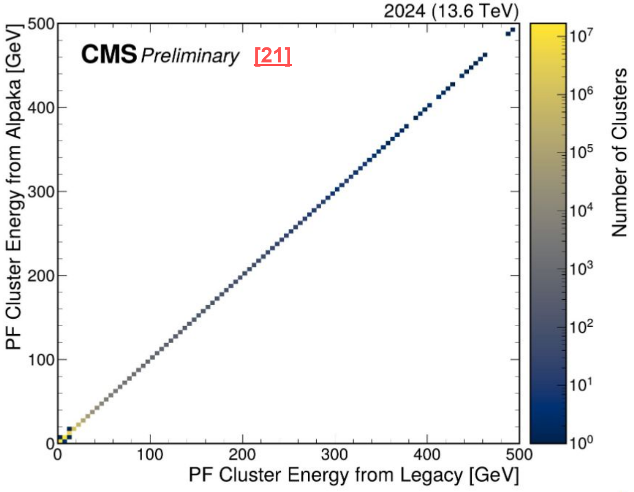


Migration to alpaka (CMS)

- **Alpaka** is a **portability** library. Same code able to run on
  - **multiple** hardware **vendors** (eg. AMD GPU, Intel GPU)
  - **multiple** kinds of **accelerators** (eg. GPU, FPGA)
- Pixel and ECAL code migrated from CUDA to Alpaka in 2024.
  - HCAL local reco migration in progress.
- Part of the Particle Flow recently ported directly to Alpaka from CPU-only.







Conclusions

- Many improvements in the ATLAS and CMS triggers have been deployed with Run-3 in the framework, algorithms, and trigger strategy
- The TLA/scouting have been deeply renewed with larger rate and new collections
- The delayed reconstruction/parking strategy have been expanded in B-physics (electrons or muons in the final state) and also to hadronic final states (eg. HH and VBF) and LLP.
- Good performance in the main objects reconstruction
  - Detector issues have been promptly addressed
  - Improved performance with respect to Run-2, despite the larger pileup.
- The ATLAS software framework is now multithreaded
- CMS offloading 40% of reconstruction time to GPU since 2022
  - The migration to the portability library Alpaka is almost completed
- Looking forward to collect more data in Run-3

  … and to face the Phase-2 upgrade challenge!

Thank you!

More infos at:

- “Enriching the physics program of CMS via data scouting and data parking”
- “The ATLAS Trigger System for LHC Run 3 and Trigger performance in 2022”
- Twiki: TriggerPublicResults (ATLAS) +
- Twiki: HighLevelTriggerRunIIIResults (CMS)

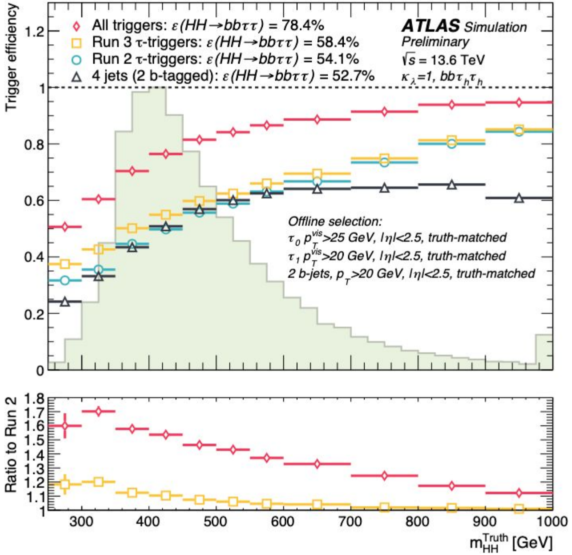
Related talks at LHCP + many posters:

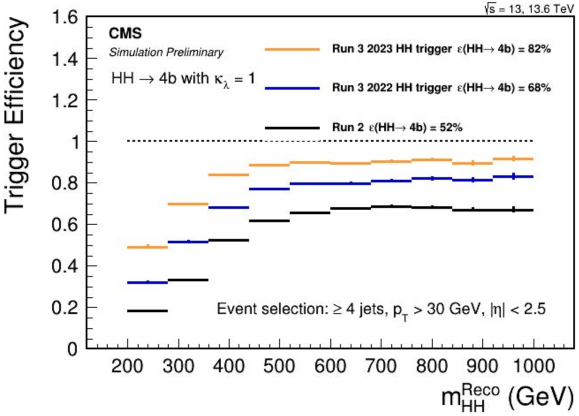
- Data processing techniques with focus on triggers, GPUs, etc. (plenary)
- Read-out developments for HL-LHC
- Novel triggering strategies (HW and SW) at the HL-LHC
- Anomaly detection in ATLAS
- Anomaly detection in CMS

Backup

HH trigger

- Dedicated trigger in delayed reconstruction for parking
- Better discriminator, more rate, large increase in acceptance





ATLAS trigger rate

Table 1: Example break-down of approximate total rates for physics triggers grouped by signature at luminosity of $1.8 \times 10^{34}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ and $\sqrt{s} = 13.6\ \mathrm{TeV}$. Rates are quoted for the Main, Delayed and TLA streams subtracting off the contributions from the less inclusive streams.

Signature	Rate per stream [Hz]		
	Main	Delayed	TLA
Electron	270		
Photon	120		
Muon	290		
Tau	160		
Missing transverse momentum	140		
Unconventional Tracking	40		
B-physics and light states		240	
Jet	490	460	5000
Jet with b-hadrons	190	160	
Combined	240	50	830

Table 2: Summary of selected triggers in the delayed streams. The VBF di-jet trigger applies Vector Boson Fusion selection requirements to the two-jet system with the highest mass. Rates are given at luminosity of $1.8 \times 10^{34}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ and $\sqrt{s} = 13.6\ \mathrm{TeV}$.

Trigger	p_T threshold [GeV]	Rate [Hz]
VBF di-jet	1000	270
Two jets, two b-jets ($\epsilon = 77\%$)	80, 55, 28, 20	160
Six jets	6×35	140
Five jets, one b-jet	5×35, 25	50
B-physics di-muon	11, 6	40
$B \to K^* ee$	5, 5	170

CMS trigger rate

Table 1: Comparison of the typical HLT trigger rates of the standard, parking, and scouting data streams during Run 1 and Run 2. The average $\mathcal{L}_{inst}$ over one typical fill of a given data-taking year and the average pileup (PU) are also reported, consistent with the scenarios reported in Fig. 2.

Year	$\mathcal{L}_{inst}$ [$cm^{-2}s^{-1}$]	PU	Standard rate [Hz]	Parking rate [Hz]	Scouting rate [Hz]
2012	0.5×10^{34}	28	420	400	1000
2016	0.9×10^{34}	35	1000	500	4500
2017	1.0×10^{34}	43	1000	400	4500
2018	1.2×10^{34}	38	1000	3000	5000

Table 7: Comparison of the typical HLT trigger rates of the standard, parking, and scouting data streams during 2018 (Run 2), 2022, and 2023 (Run 3). The average $\mathcal{L}_{inst}$ value over one typical fill of a given data-taking year and the average pileup (PU) are also reported, coherently with the scenarios reported in Fig. 2.

Year	$\mathcal{L}_{inst}$ [$cm^{-2}s^{-1}$]	PU	Standard rate [Hz]	Parking rate [Hz]	Scouting rate [Hz]
2018	1.2×10^{34}	38	1000	3000	5000
2022	1.5×10^{34}	46	1800	2440	22000
2023	1.7×10^{34}	48	1700	2660	17000

Configuration	2022	2023
Scouting path	Rate per path [kHz]	
1 e/γ	—	9.1
$\geq$2 e/γ	—	0.3
$\geq$2 muons	—	3.4
Jets or H_T	—	11.0
e/γ, $\geq$2 muons, jets or H_T	31.3	—