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: \( \approx 30 \text{ MHz} \)
- **Hardware** trigger (L1): \( \approx 30 \text{ MHz} \rightarrow \approx 100 \text{ kHz} \)
  - simplified readout (no tracker), small latency (<4 \( \mu \text{s} \)).
- **Software** trigger (HLT): \( \approx 100 \text{ kHz} \rightarrow \text{few kHz} \)
  - full event readout available (~1 MB/event);
  - HLT farm with \( \approx 50k \) threads \( \rightarrow \approx 500\text{ms/event} \) on average
  - Events are rejected in the early stage of the reconstruction
- **Storage and offline reconstruction**
Delayed reconstruction/parking

- The “standard” trigger cross section $\sim 100$ nb
  - eg. Lumi $1 \times 10^{34} \text{cm}^{-2}\text{s}^{-1} \rightarrow$ 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.

### Average trigger rate vs year

<table>
<thead>
<tr>
<th>Year</th>
<th>ATLAS Rate</th>
<th>CMS Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2023</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trigger selection</th>
<th>ATLAS Rate [1] (at $1.8 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$)</th>
<th>CMS Rate [2] (at $2.0 \times 10^{34} \text{cm}^{-2}\text{s}^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 muons (B-physics)</td>
<td>40 Hz</td>
<td>1.6 kHz</td>
</tr>
<tr>
<td>2 electrons (B-physics)</td>
<td>170 Hz</td>
<td>1.3 kHz (only 2022)</td>
</tr>
<tr>
<td>Vector-boson fusion</td>
<td>270 Hz</td>
<td>1.2 kHz (since 2023)</td>
</tr>
<tr>
<td>HH (2 jets + 2 b-jets)</td>
<td>160 Hz</td>
<td>180 Hz (since 2023)</td>
</tr>
<tr>
<td>6 jets</td>
<td>140 Hz</td>
<td>-</td>
</tr>
<tr>
<td>5 jets + 1 b-jet</td>
<td>50 Hz</td>
<td>-</td>
</tr>
<tr>
<td>Long-lived particle</td>
<td>-</td>
<td>150 Hz (since 2023)</td>
</tr>
</tbody>
</table>
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.
Observation of $\eta \rightarrow 4\mu$ (Run-2 scouting)

TLA photon calibration

Scouting dimuon mass resolution (Run-2 scouting)

Scouting dielectron mass resolution
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

- 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

---

**CMS and ATLAS trigger performance - Silvio Donato (INFN Pisa)**

**CMS**
- $32.0\ \text{fb}^{-1}$ (13 TeV) and $17.6\ \text{fb}^{-1}$ (13.6 TeV)
- Position of pixel detector layers: 29-68-109-160 mm

**ATLAS Simulation**
- $\sqrt{s} = 13\ \text{TeV}$
- Offline selection:
  - Large-R PFlow jet $p_T > 460\ \text{GeV}$
  - Large-R jet $p_T > 200\ \text{GeV}$, $|\eta| < 1.8$

**Emerging jet efficiency vs $p_T$**
- $m_{\tau} = 1500\ \text{GeV}$
- $m_{\tau} = 20\ \text{GeV}$
- $c_t = 50\ \text{mm}$

**Delayed/trackless jet efficiency vs $p_T$**
- CMS Simulation
- H→XX→4b ($m_{\tau} = 1000\ \text{GeV}$, $m_{\tau} = 450\ \text{GeV}$, $c_t = 10\ \text{m}$)
- Delayed jet trigger with $H_T > 430\ \text{GeV}$ and jet timing $> 1\ \text{ns}$
- Trackless delayed jet trigger with $H_T > 430\ \text{GeV}$ and jet timing $> 1\ \text{ns}$
- Run 2 $H_T > 1950\ \text{GeV}$ trigger

---

**Scouting muon vs displacement**

**LLP jet efficiency vs $p_T$ for different HCAL time offset**

---

[1] [18] [2]
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 $1.8 \times 10^{34}$ cm$^{-2}$ s$^{-1}$ and $\sqrt{s} = 13.6$ TeV. Rates are quoted for the Main, Delayed and TLA streams subtracting off the contributions from the less inclusive streams.

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

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 $1.8 \times 10^{34}$ cm$^{-2}$ s$^{-1}$ and $\sqrt{s} = 13.6$ TeV.

<table>
<thead>
<tr>
<th>Trigger</th>
<th>$p_T$ threshold [GeV]</th>
<th>Rate [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBF di-jet</td>
<td>1000</td>
<td>270</td>
</tr>
<tr>
<td>Two jets, two $b$-jets ($e = 77%$)</td>
<td>80, 55, 28, 20</td>
<td>160</td>
</tr>
<tr>
<td>Six jets</td>
<td>6 $\times$ 35</td>
<td>140</td>
</tr>
<tr>
<td>Five jets, one $b$-jet</td>
<td>5 $\times$ 35, 25</td>
<td>50</td>
</tr>
<tr>
<td>$B$-physics di-muon</td>
<td>11, 6</td>
<td>40</td>
</tr>
<tr>
<td>$B \rightarrow K^*ee$</td>
<td>5, 5</td>
<td>170</td>
</tr>
</tbody>
</table>
### 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 $\mathcal{L}_{\text{inst}}$ 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.

<table>
<thead>
<tr>
<th>Year</th>
<th>$\mathcal{L}_{\text{inst}}$ [cm$^{-2}$s$^{-1}$]</th>
<th>PU</th>
<th>Standard rate [Hz]</th>
<th>Parking rate [Hz]</th>
<th>Scouting rate [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>$0.5 \times 10^{34}$</td>
<td>28</td>
<td>420</td>
<td>400</td>
<td>1000</td>
</tr>
<tr>
<td>2016</td>
<td>$0.9 \times 10^{34}$</td>
<td>35</td>
<td>1000</td>
<td>500</td>
<td>4500</td>
</tr>
<tr>
<td>2017</td>
<td>$1.0 \times 10^{34}$</td>
<td>43</td>
<td>1000</td>
<td>400</td>
<td>4500</td>
</tr>
<tr>
<td>2018</td>
<td>$1.2 \times 10^{34}$</td>
<td>38</td>
<td>1000</td>
<td>3000</td>
<td>5000</td>
</tr>
</tbody>
</table>

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 $\mathcal{L}_{\text{inst}}$ 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.

<table>
<thead>
<tr>
<th>Year</th>
<th>$\mathcal{L}_{\text{inst}}$ [cm$^{-2}$s$^{-1}$]</th>
<th>PU</th>
<th>Standard rate [Hz]</th>
<th>Parking rate [Hz]</th>
<th>Scouting rate [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>$1.2 \times 10^{34}$</td>
<td>38</td>
<td>1000</td>
<td>3000</td>
<td>5000</td>
</tr>
<tr>
<td>2022</td>
<td>$1.5 \times 10^{34}$</td>
<td>46</td>
<td>1800</td>
<td>2440</td>
<td>22000</td>
</tr>
<tr>
<td>2023</td>
<td>$1.7 \times 10^{34}$</td>
<td>48</td>
<td>1700</td>
<td>2660</td>
<td>17000</td>
</tr>
</tbody>
</table>