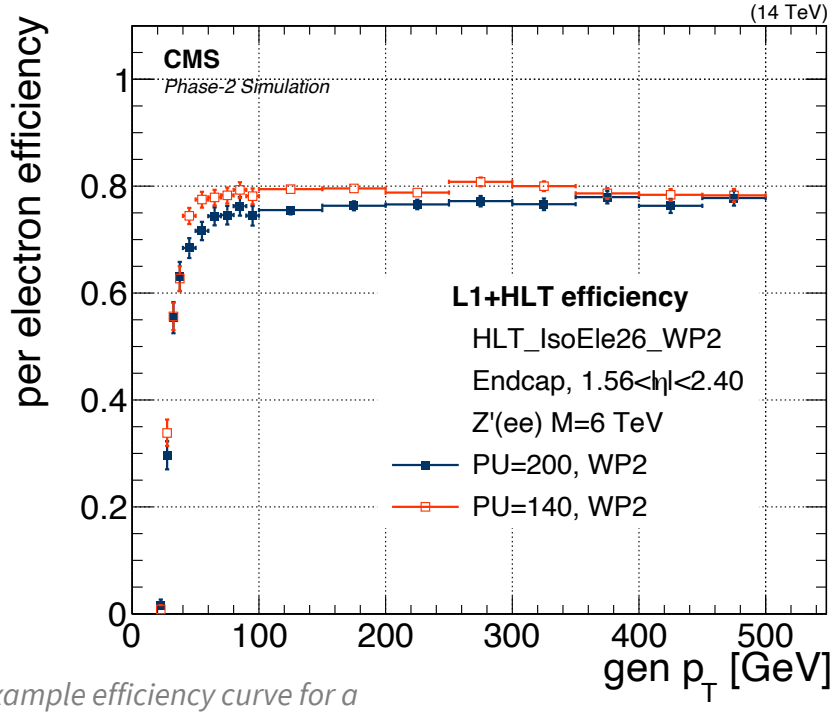
The CMS High-Level Trigger for LHC Run-3

THIAGO R. F. P. TOMEI
FOR THE CMS COLLABORATION

SPRACE

The Triple Challenge of the HLT



Example efficiency curve for a single electron trigger

Efficiency

- ❑ Select the events of interest
- ❑ Generalist vs. specialized triggers

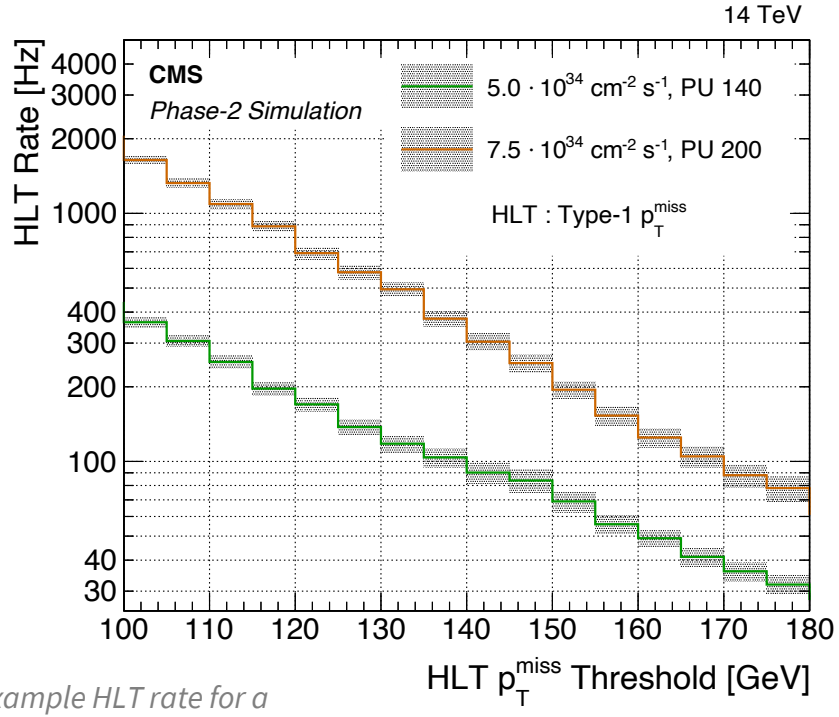
Rate

- ❑ Discard uninteresting events
- ❑ Output rate / bandwidth envelope

Timing

- ❑ Quasi-real time analysis
- ❑ Dependent on HLT farm size

The Triple Challenge of the HLT



Example HLT rate for a missing p_T trigger

Efficiency

- Select the events of interest
- Generalist vs. specialized triggers

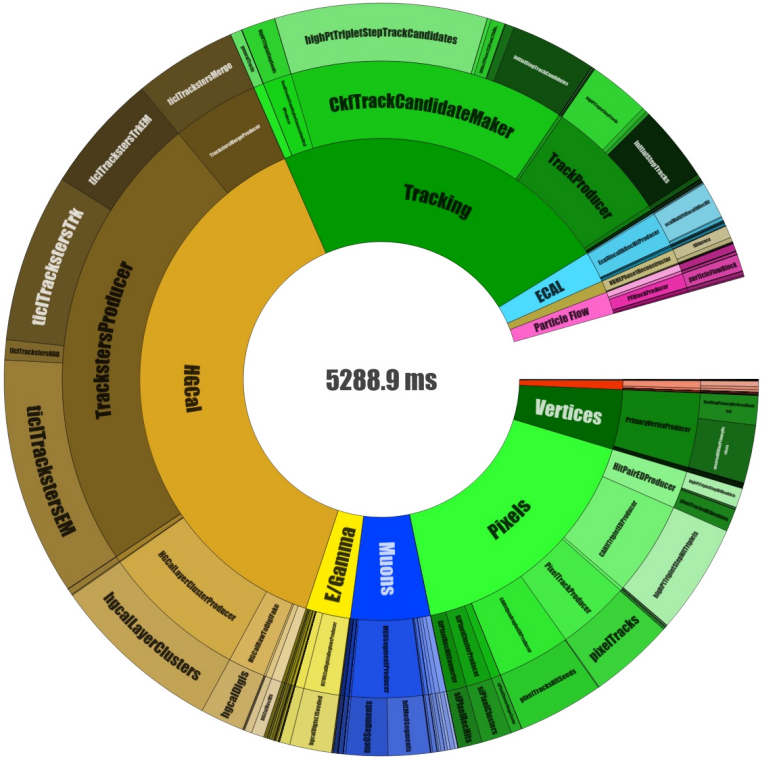
Rate

- Discard uninteresting events
- Output rate / bandwidth envelope

Timing

- Quasi-real time analysis
- Dependent on HLT farm size

The Triple Challenge of the HLT



Efficiency

- ❑ Select the events of interest
- ❑ Generalist vs. specialized triggers

Rate

- ❑ Discard uninteresting events
- ❑ Output rate / bandwidth envelope

Timing

- ❑ Quasi-real time analysis
- ❑ Dependent on HLT farm size

LHC Run 3 (2022) status

Delivered 42.0 fb^{-1}

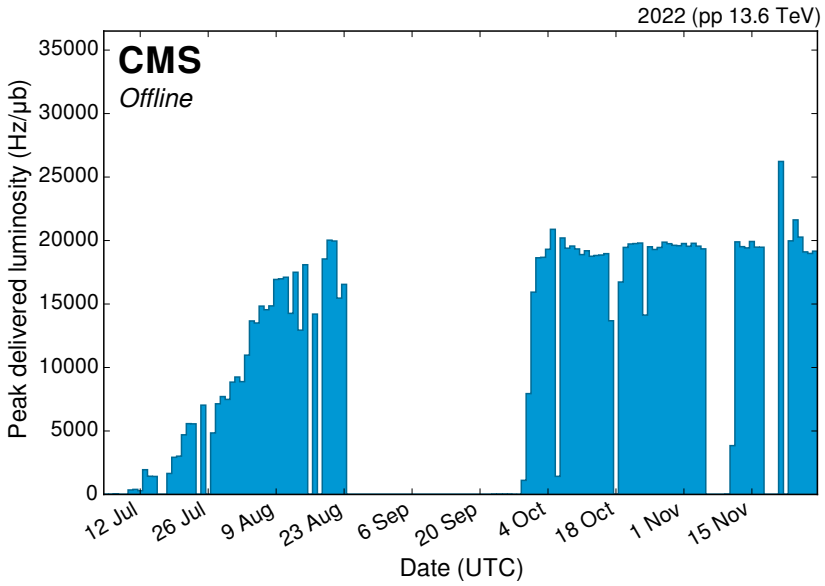
- CMS recorded: 38.5 fb^{-1}
- Regular running at 20 Hz/nb

Price: even higher pileup

- Run2: $\langle\mu\rangle = 34$
- 2022: $\langle\mu\rangle = 46$ (+35% increase!)

Different modus operandi

- Beta* levelling in first 5-6 hours of every fill.



LHC Run 3 (2022) status

Delivered 42.0 fb^{-1}

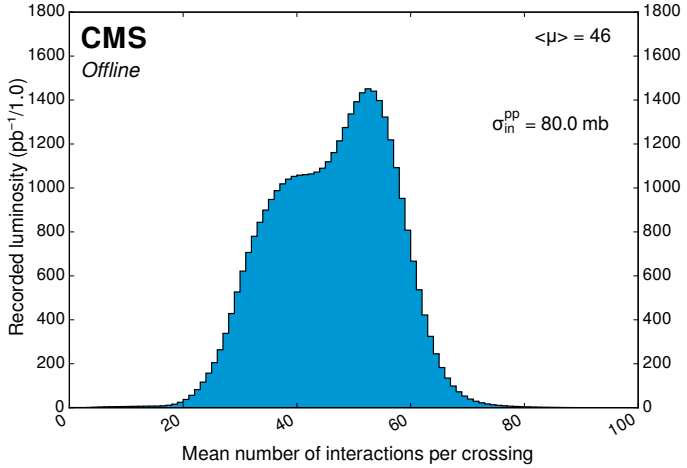
- CMS recorded: 38.5 fb^{-1}
- Regular running at 20 Hz/nb

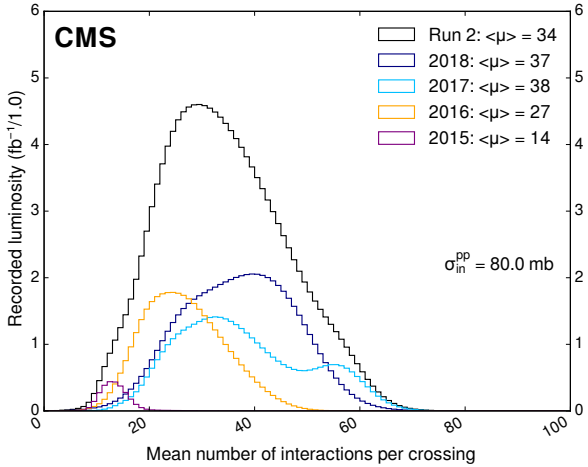
Price: even higher pileup

- Run2: $\langle\mu\rangle = 34$
- 2022: $\langle\mu\rangle = 46$ (+35% increase!)

Different modus operandi

- Beta* levelling in first 5-6 hours of every fill.





LHC Run 3 (2022) status

Delivered 42.0 fb^{-1}

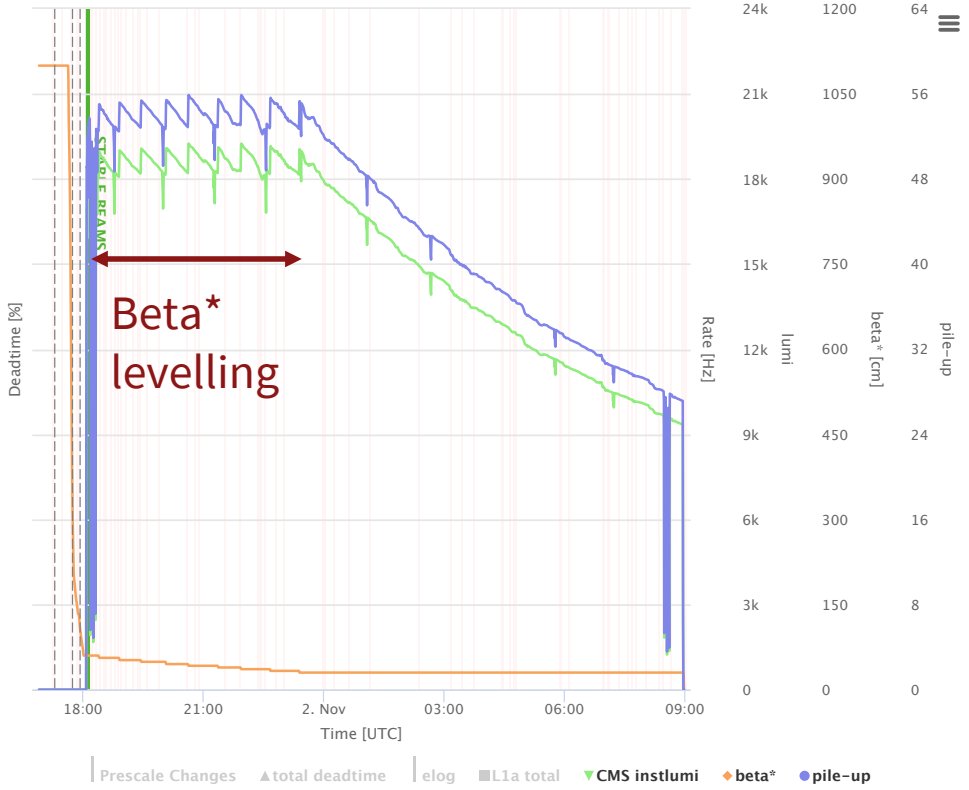
- CMS recorded: 38.5 fb^{-1}
- Regular running at 20 Hz/nb

Price: even higher pileup

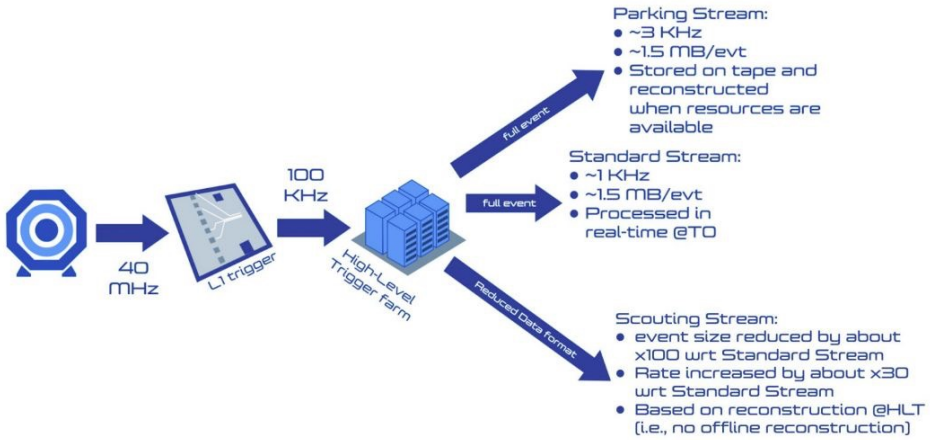
- Run2: $\langle\mu\rangle$ = 34
- 2022: $\langle\mu\rangle$ = 46 (+35% increase!)

Different modus operandi

- Beta* levelling in first 5-6 hours of every fill.



CMS Run 3 Trigger Strategy



Standard Physics

- ❑ Offline reconstruction in 48 hours
- ❑ Average 2 kHz

Data Parking

- ❑ Offline reconstruction dependent on computing resources availability
- ❑ Average 3 kHz

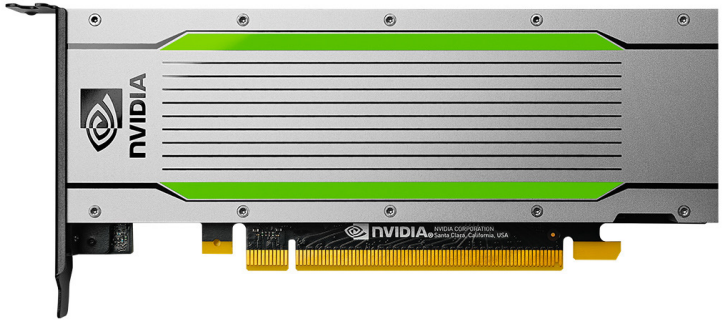
Data Scouting

- ❑ No offline reconstruction – no RAW!
  - Analysis with HLT information
  - 7 kB/event vs 1 MB/event
- ❑ Average 20 kHz

HLT Farm 2022

200 nodes – AMD EPYC 7763 “Milan”

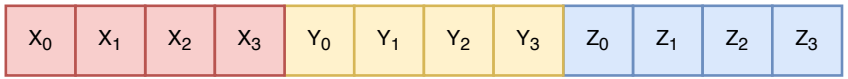
- 64 cores/processor, 2 processors/node
- 256 GB RAM/node
- **Two NVIDIA T4 GPUs**, each with:
  - 40 multiprocessors
  - 2560 CUDA “threads”
- Measured **3224 HS/node**
  - HLT performance known to be ~linear with HS number.



Heterogeneous Computing at the HLT

```
hltPixelVerticesSoA = SwitchProducerCUDA(
        cpu = cms.EDAlias(
                hltPixelVerticesCPU = cms.VPSet(
                        type = cms.string( "*" )
                )
        ),
        cuda = cms.EDAlias(
                hltPixelVerticesFromGPU = cms.VPSet(
                        type = cms.string( "*" )
                )
        )
)
```

Array of Structures



Structure of Arrays

CMSSW architecture

- ❑ Fully multithreaded
- ❑ Offload to GPUs – new for Run3

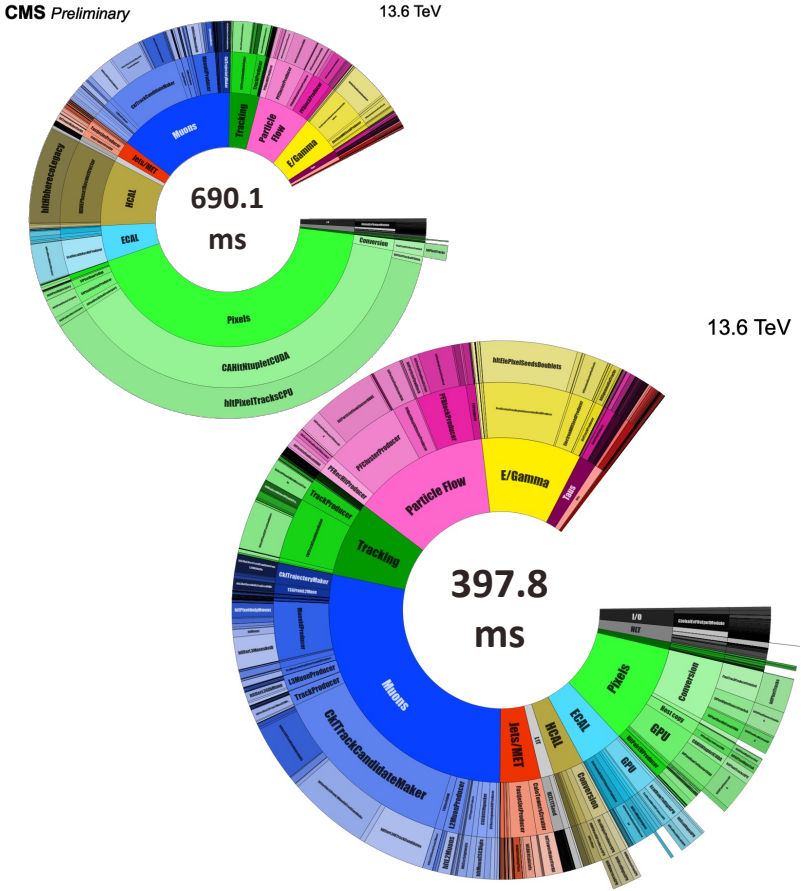
First application: HLT

- ❑ Dual implementation: CPU-only, CPU+GPU
- ❑ HLT in GPU ~ **40% faster**
- ❑ Reproducibility + validation of results

Current offloadings

- ❑ ECAL local RECO
- ❑ HCAL local RECO
- ❑ Pixel tracking and vertexing

Heterogeneous Computing at the HLT



CMMSW architecture

- ❑ Fully multithreaded
- ❑ Offload to GPUs – new for Run3

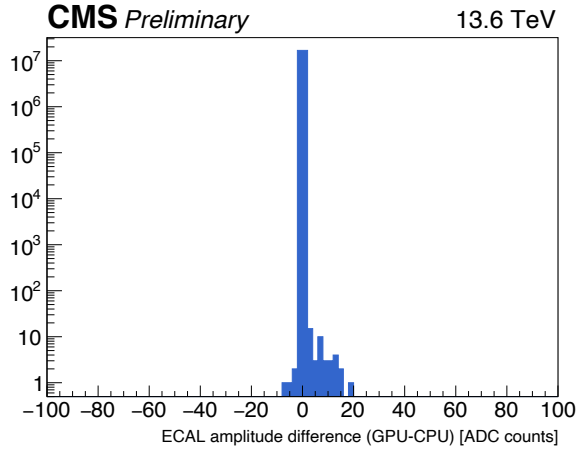
First application: HLT

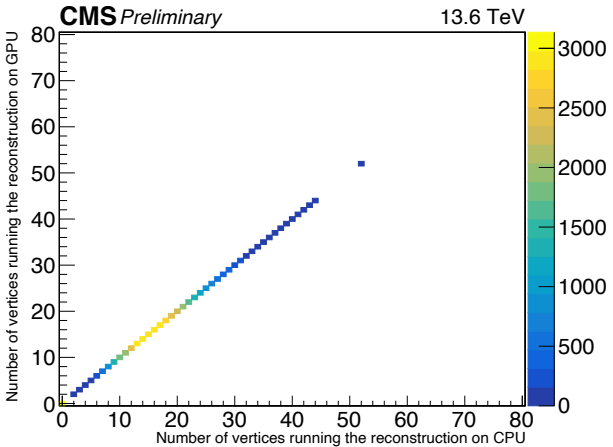
- ❑ Dual implementation: CPU-only, CPU+GPU
- ❑ HLT in GPU ~ **40% faster**
- ❑ Reproducibility + validation of results

Current offloadings

- ❑ ECAL local RECO
- ❑ HCAL local RECO
- ❑ Pixel tracking and vertexing

Heterogeneous Computing at the HLT





CMMSW architecture

- Fully multithreaded
- Offload to GPUs – new for Run3

First application: HLT

- Dual implementation: CPU-only, CPU+GPU
- HLT in GPU ~ **40% faster**
- Reproducibility + validation of results

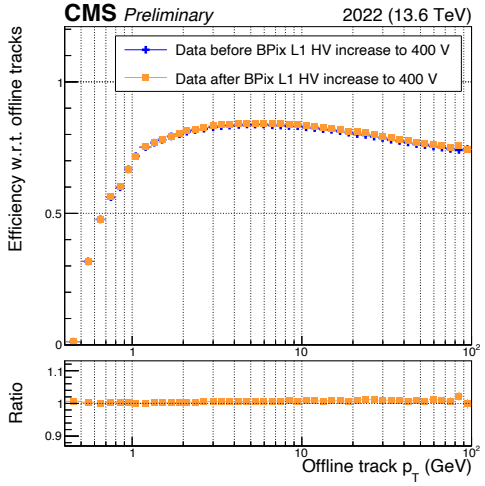
Current offloadings

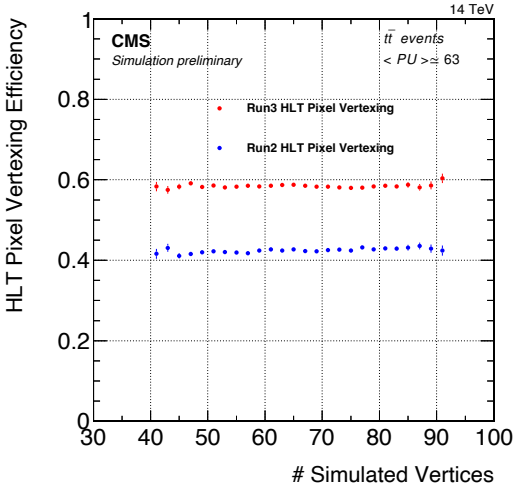
- ECAL local RECO
- HCAL local RECO
- Pixel tracking and vertexing

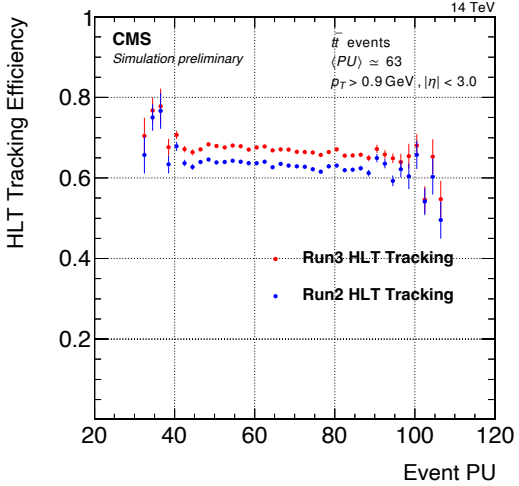
Tracking at HLT

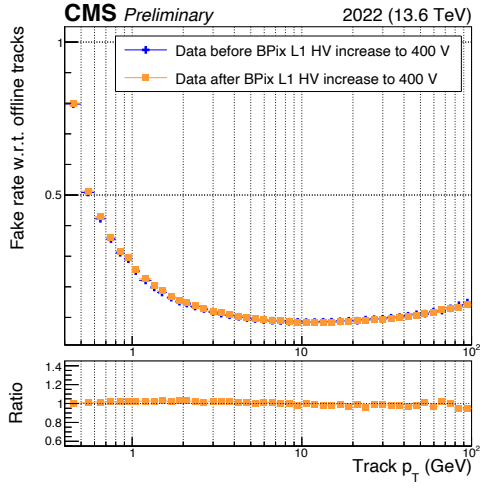
Single Combinatorial Kalman Filter iteration

- Seeds: Patatrack pixel tracks
  - ≥ 3 pixel hits
  - $p_T > 0.3$ GeV
  - Consistent with leading pixel vertex





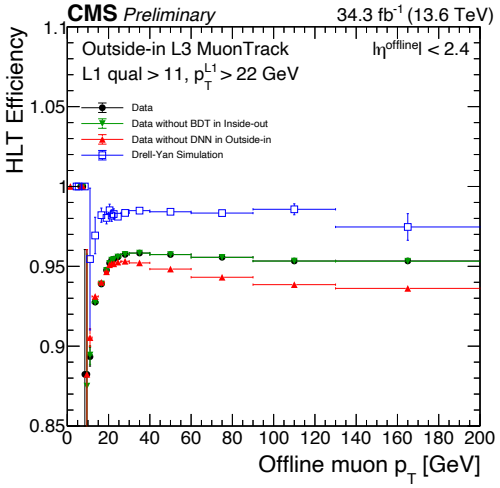


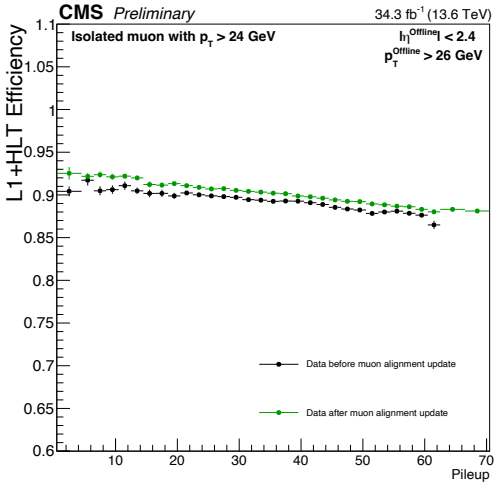


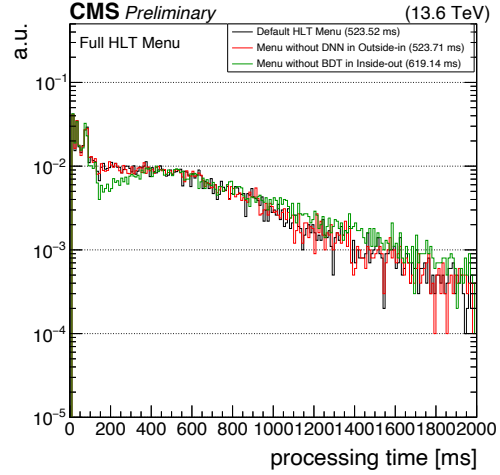
Muons at HLT

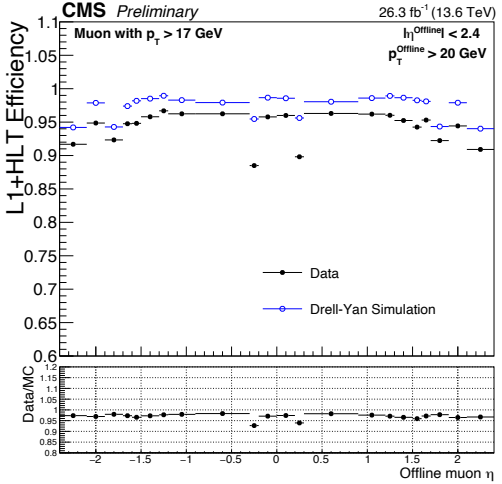
ML-based Improvements

- DNN: higher efficiency
  - Improved track seeding in outside-in reconstruction.
- BDT: reduced timing
  - Improved seed choice in inside-out reconstruction.

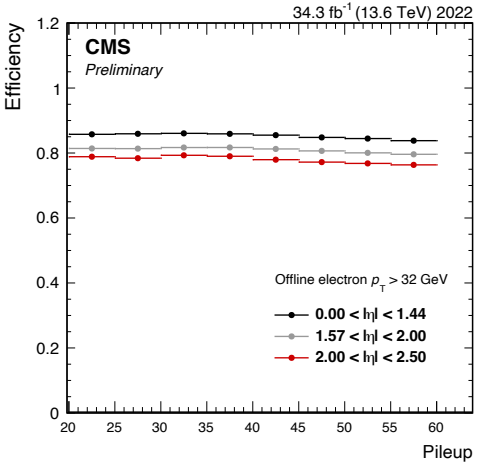


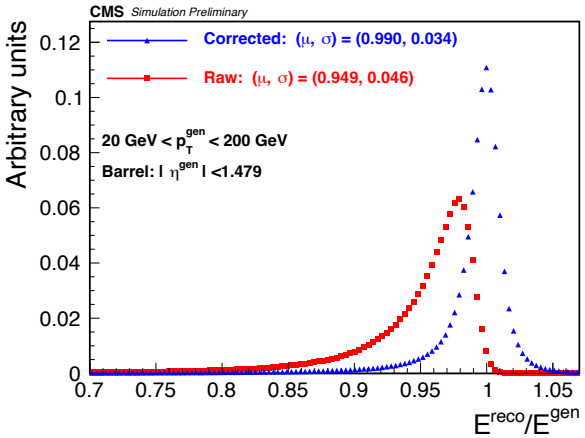






Electrons at HLT





Focus ID on high-purity

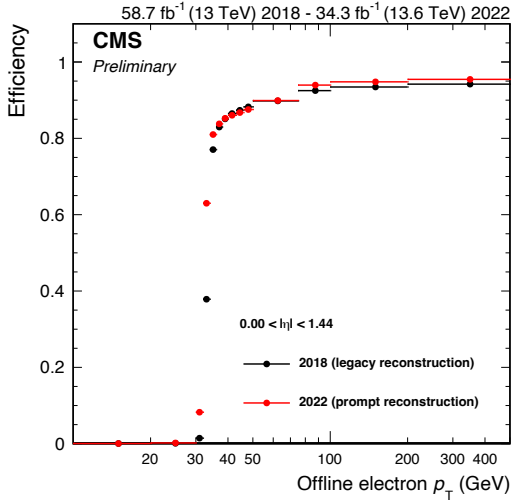
- Tracker-ECAL matching
- ECAL deposit shape
- Reject HCAL deposits

Isolation

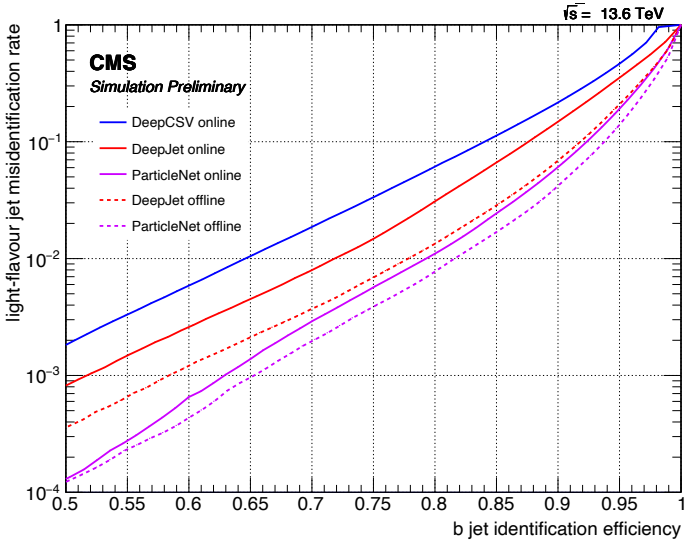
- Tracker, ECAL, HCAL

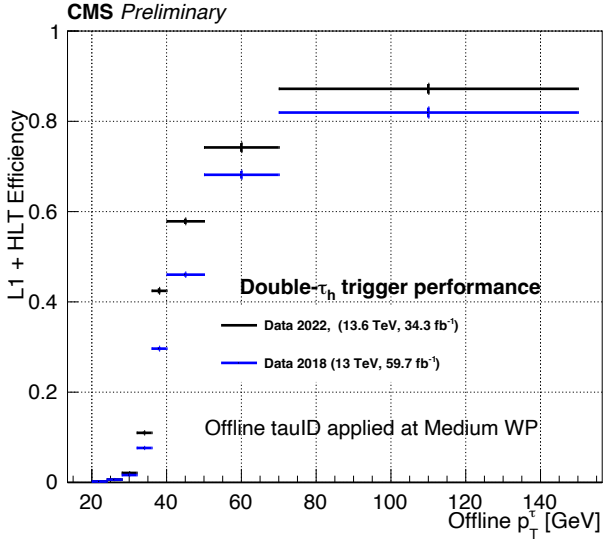
Energy regression

- BDT-based



B-tagging and Taus at HLT





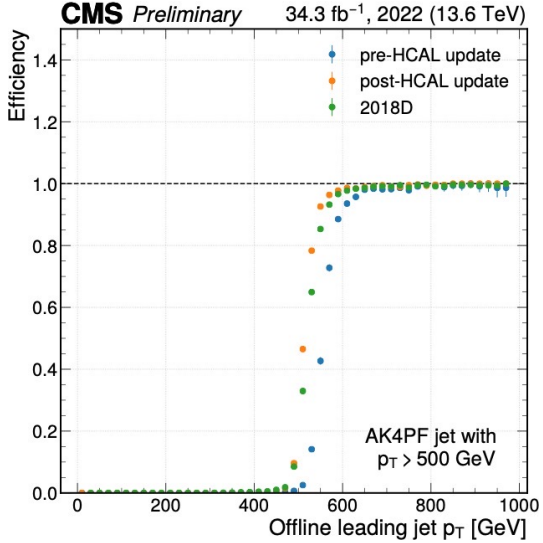
Increased PU in Run 3 → more advanced mitigations

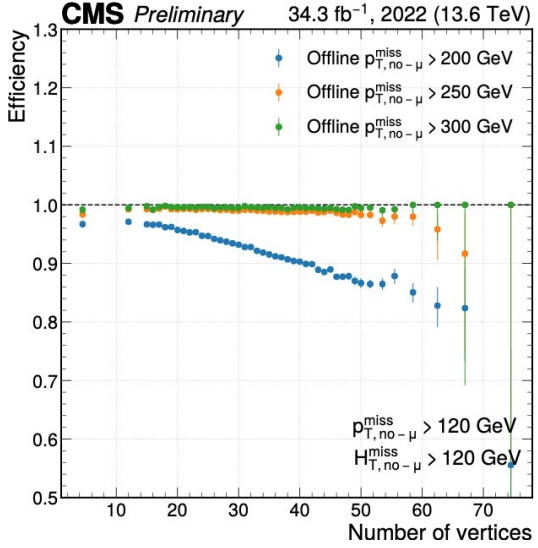
- B-tagging: ParticleNet tagger instead of DeepCSV and DeepJet
- Tau ID: DeepTau instead of Hadron-Plus-Strip + MVA isolation.

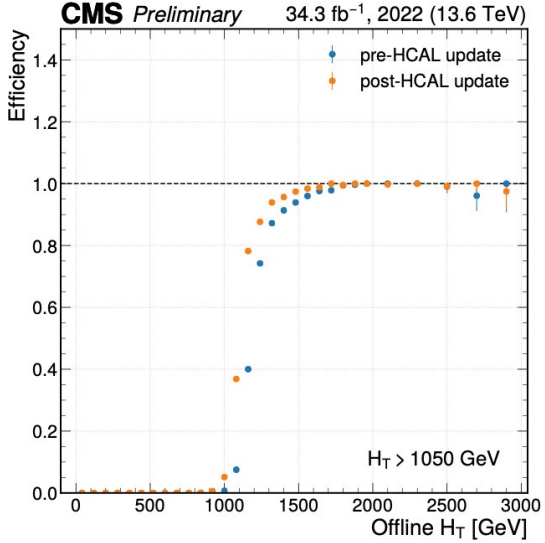
Jets and Energy Sums at HLT

Mitigations soon to be added to the HLT:

- ❑ Charged Hadron Subtraction (CHS), OR
- ❑ PileUp Per-Particle Identification (PUPPI)





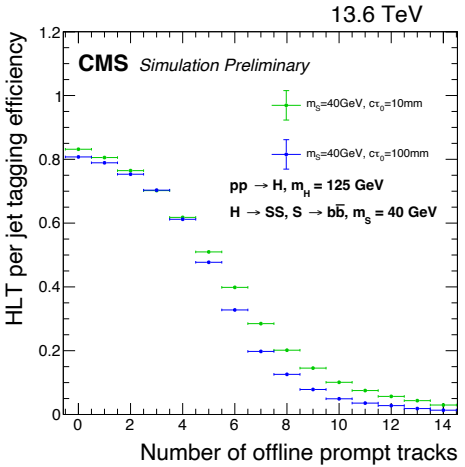


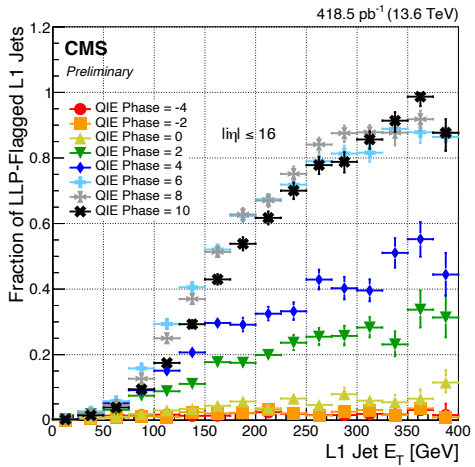
Long-Lived Particles (LLPs) at the HLT

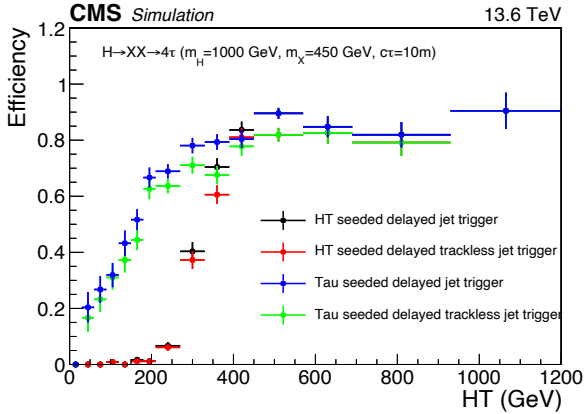
LLPs signatures may escape conventional triggers

New Run 3 strategies

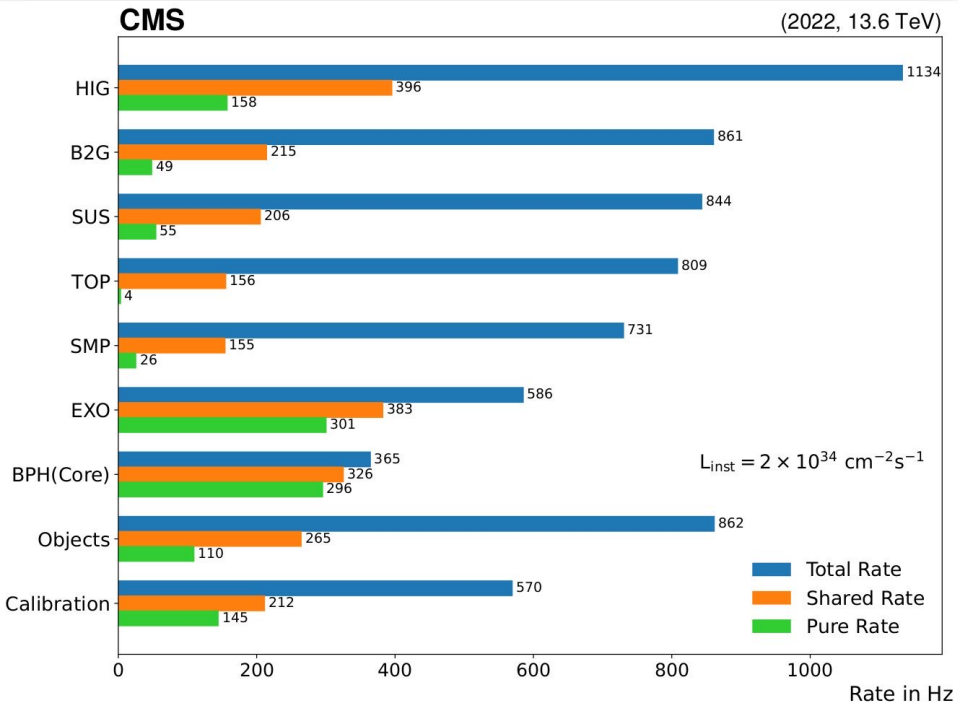
- Displaced jet triggers
  - Modified tracking for light LLPs
  - HCAL depth segmentation
- Delayed jet triggers
  - ECAL and HCAL timing
- CSC high multiplicity triggers
- Displaced muons at L1







The HLT Run 3 Menu



Standard Physics

(Rates scaled to 2E34 cm^{-2} s^{-1})

HLT algorithm	Rate
Isolated muon with $p_T > 24$ GeV	250 Hz
Isolated electron with $E_T > 32$ GeV	182 Hz
Particle flow (PF) based $p_T^{miss} > 110$ GeV	81 Hz
4 PF jets with $p_T > 70/50/40/35$ GeV with 2 b tag	57 Hz
Two isolated tau leptons with $p_T > 35$ GeV	54 Hz
Muon with $p_T > 50$ GeV	51 Hz
Two electrons with $E_T > 25$ GeV	21 Hz
AK4 PF jet with $p_T > 500$ GeV	16 Hz
Two same-sign muons with $p_T > 18/9$ GeV	10 Hz

~ 600 paths

Additional HLT Data Streams: Scouting

Only reconstructed HLT data saved

– no RAW data!

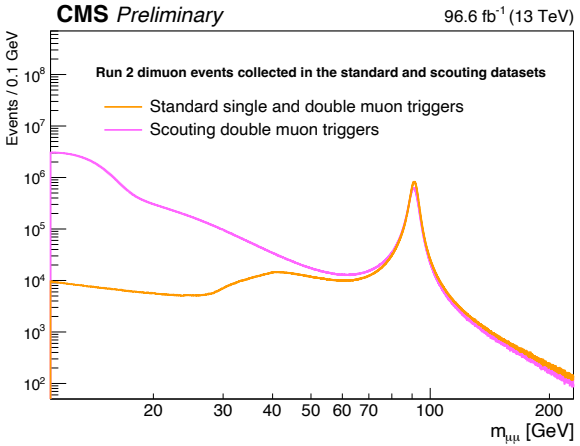
- Active in CMS since Run 1

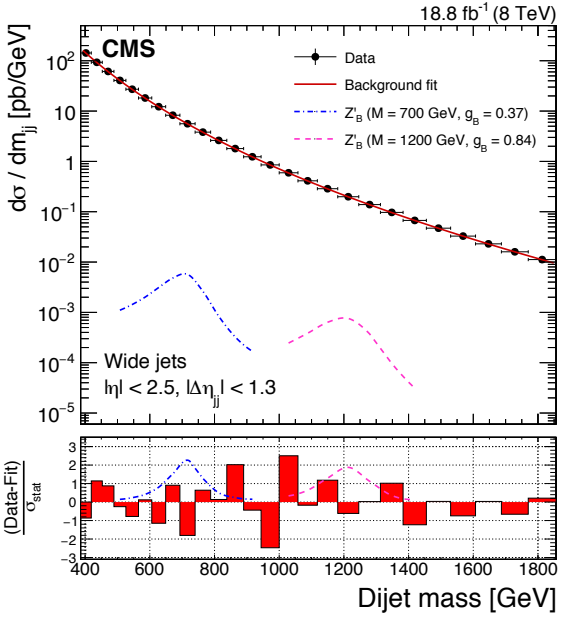
Run 3 scouting

- Dimuon
- Electrons
- Jets / missing p_T

Run 3 dimuon scouting

- Save OR of L1 dimuon seeds
- Pixel-only particle flow
- p_T resolution in 1—2% range





Additional HLT Data Streams: Scouting

Only reconstructed HLT data saved
– no RAW data!

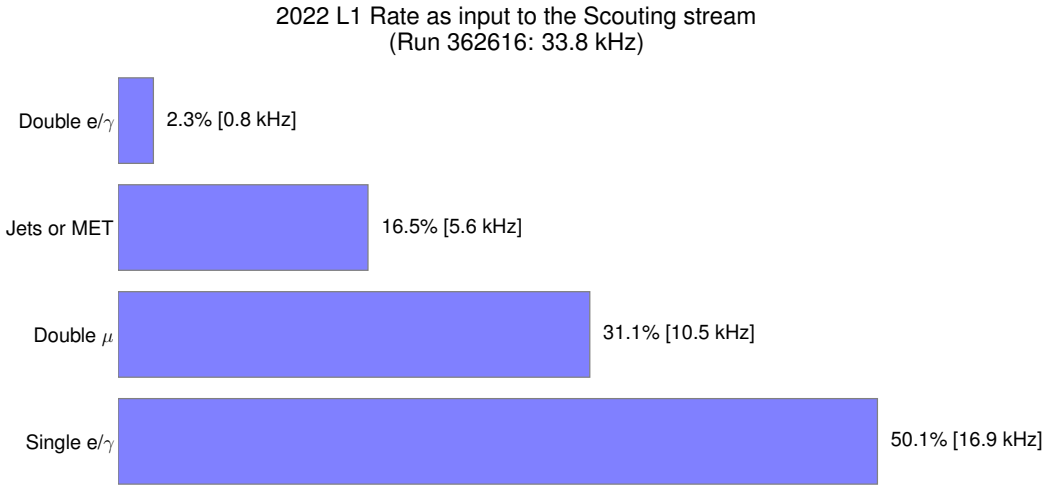
- Active in CMS since Run 1

Run 3 scouting

- Dimuon
- Electrons
- Jets / missing p_T

Run 3 dimuon scouting

- Save OR of L1 dimuon seeds
- Pixel-only particle flow
- p_T resolution in 1—2% range



Additional HLT Data Streams: Scouting

Only reconstructed HLT data saved – no RAW data!

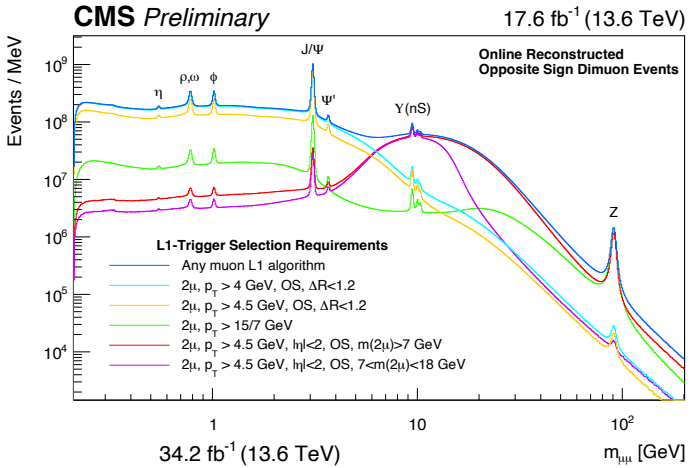
- Active in CMS since Run 1

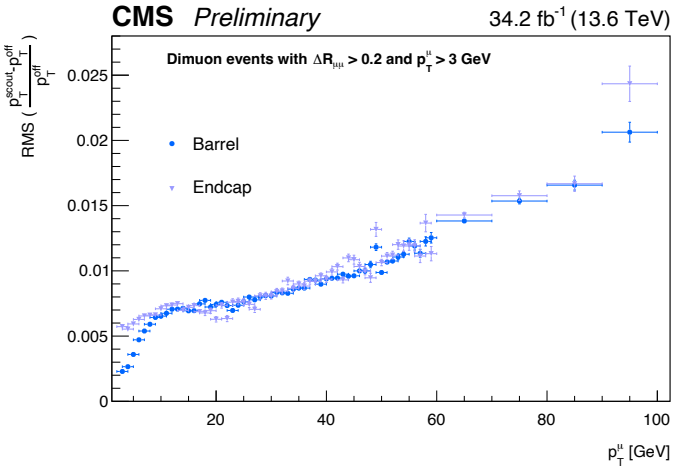
Run 3 scouting

- Dimuon
- Electrons
- Jets / missing p_T

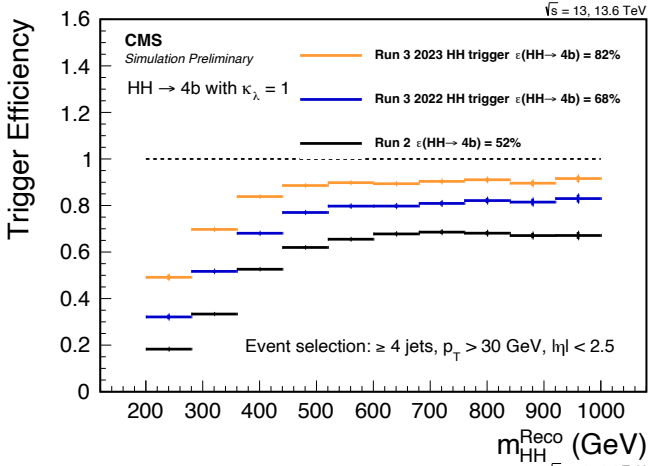
Run 3 dimuon scouting

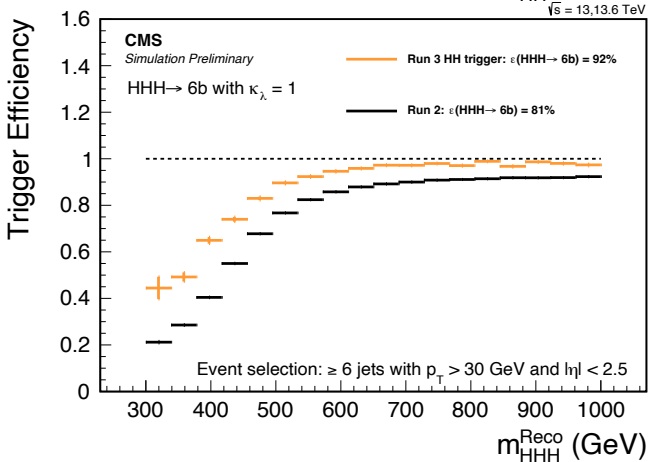
- Save OR of L1 dimuon seeds
- Pixel-only particle flow
- p_T resolution in 1—2% range





Additional HLT Data Streams: Parking





Trade prompt data reconstruction for increased rate and acceptance.

Run 3 example: Higgs parking

- Target HH and HHH signatures
  - Multi b-jets
- 2022 configuration: 60 Hz @ 2E34
  - 4 jets p_T > 70, 50, 40, 35 GeV
  - PNet@AK4 (mean 2 highest b-tag score) > 0.65
- 2023 configuration: 180 Hz @ 2E34
  - H_T > 280 GeV
  - 4 jets with p_T > 30 GeV
  - PNet@AK4 (mean 2 highest b-tag score) > 0.55

Conclusions

Run 3 LHC conditions much harsher than before.

- Higher pileup, sustained for longer time.
- Challenges in trigger and data acquisition.

Mitigation strategies at the CMS High-Level Trigger.

- GPU-equipped nodes + SoA data + task offloading.
- Advanced algorithms: Patatrack, PUPPI, ParticleNet, DeepTau.

Physics program extension.

- Long-lived particles.
- Alternative data streams: scouting, parking.
- Others to come...

CMS is ready and running for Run 3!