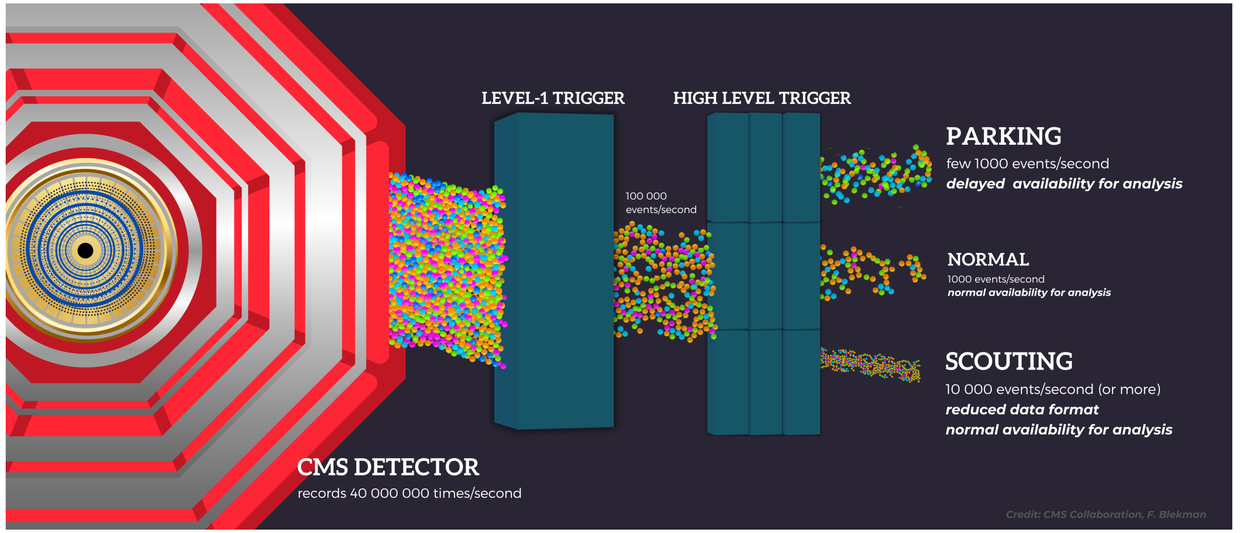
Probing new physics with dedicated data streams at CMS

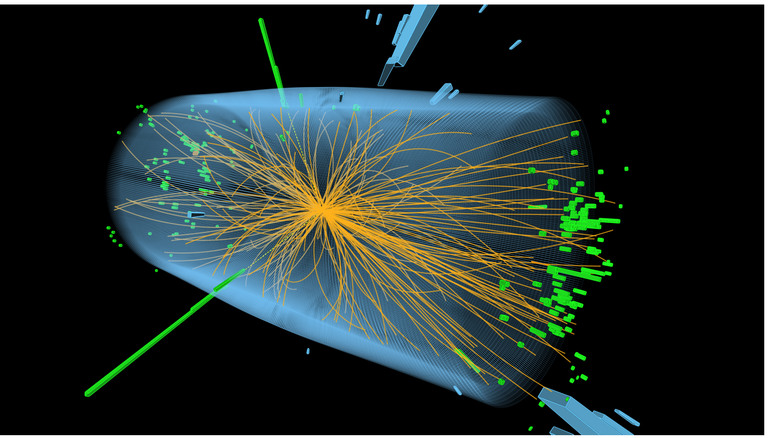
Daniel Diaz on behalf of the CMS Collaboration

ICHEP 2024

20.07.2024



Standard DAQ @CMS



Global traffic of the internet
~100 EB/month

Zettabyte

~ 1.4 ZB = 1400 EB = 1.4*10^6 PB *

Exabyte

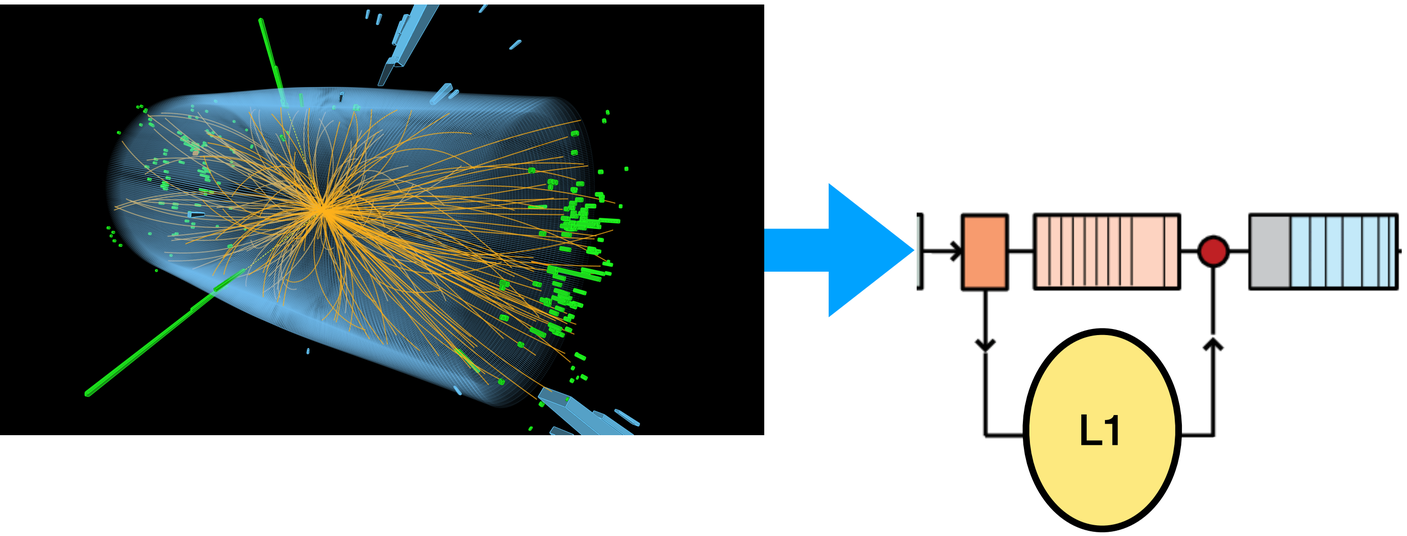
All known created and recorded data ~100 ZB

Typical datacenter size e.g., Facebook/Amazon/Microsoft

- Event Rate ~40MHz

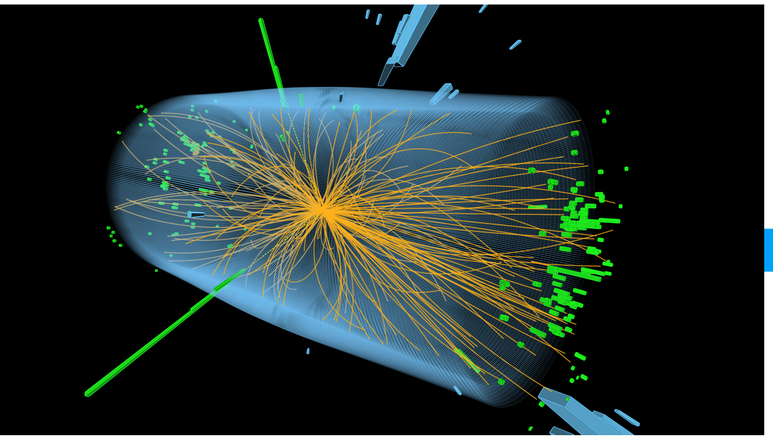
*Based on Run2 RAW data size

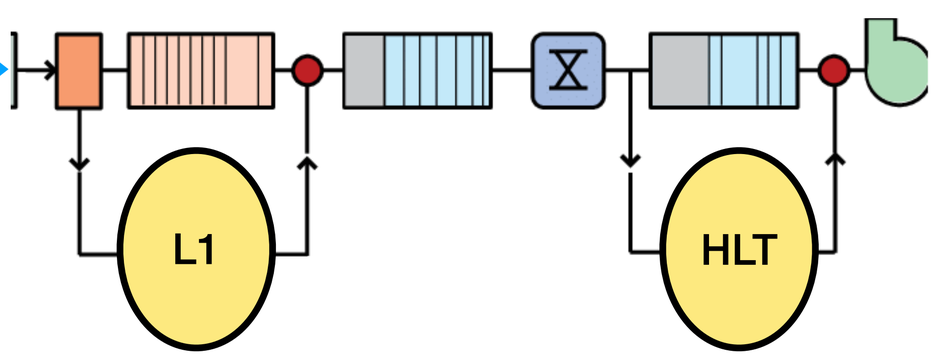
Standard DAQ @CMS



- First level event filter
- **Reduces event rate to ~100kHz** (hard limit based on detector design)
- Decision based on fast readout (FPGAs) of detector electronics.
- Limited event information, mostly calorimeter and muon detector data.

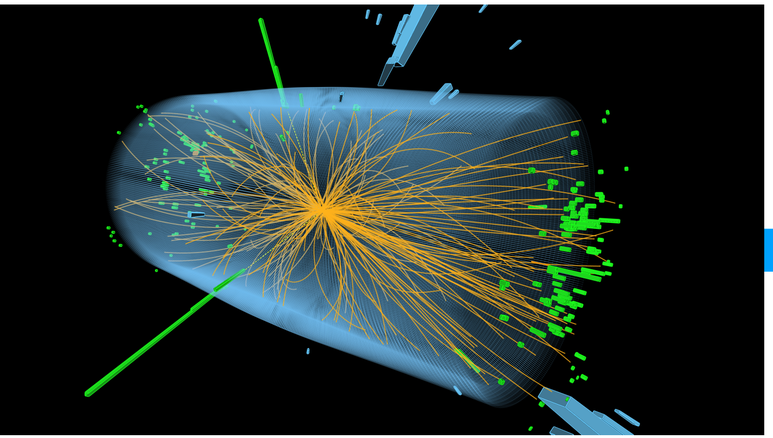
Standard DAQ @CMS

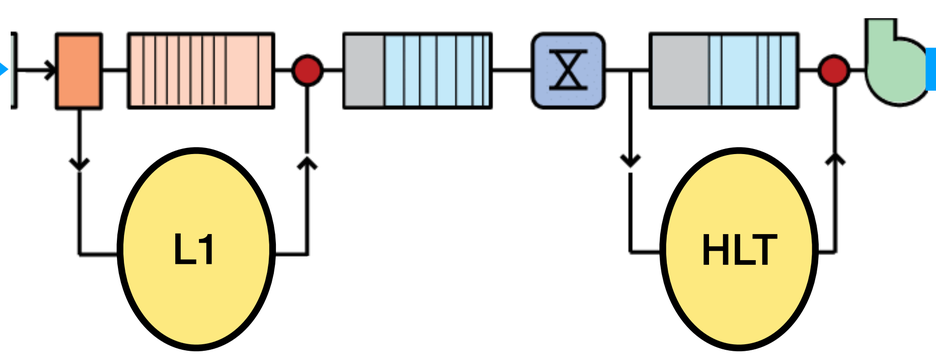




- Final online event filter
- **Reduces event rate to ~1kHz**
- Software based filter running on farm of CPUs and GPUs.
- Complete detector readout available.

Standard DAQ @CMS





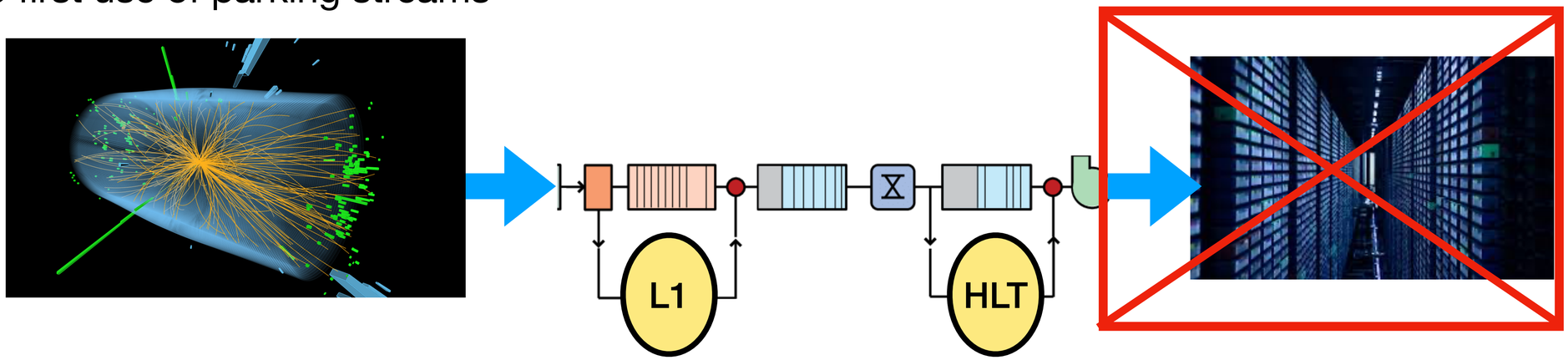
- RAW detector data stored for offline “prompt” reconstruction.
  - Data becomes available to analysts after ~48 hours.
- Typical Event size ~1MB
- **Data rate @1kHZ ~ 1GB/s**
- Data bandwidth limitation ~2GB/s (Run2)

Data Scouting

*Turbo is the analog in ATLAS & LHCb

- General concept: Reduce event content in order to increase event rate.
- Scouting streams cherry pick relevant HLT level data to store
  - First use in 2011 stored just jet four momenta.
- **No RAW detector data stored**
- **Benefit**: reducing the event content allows lowering of trigger thresholds
  - Reduce the detector readout bandwidth limitation
  - Access lower mass/lower pT objects standard streams cannot.
- **Downside**: Less information available to characterize the event.*

*Leads to first use of parking streams



Data Scouting

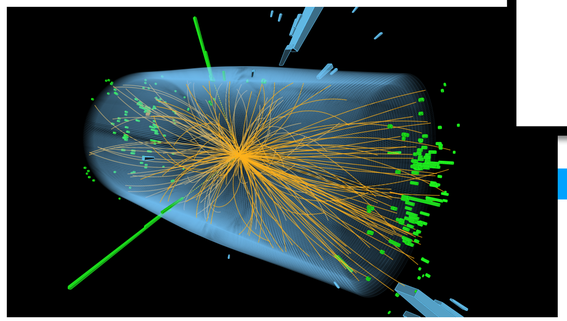
*Turbo is the analog in ATLAS & LHCb

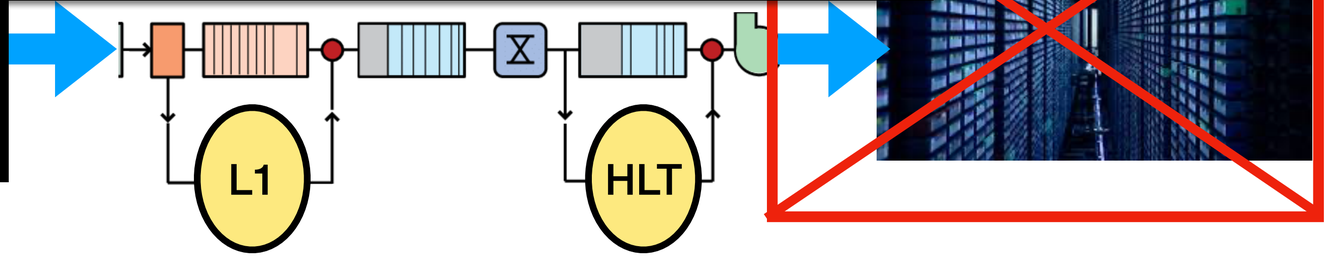
- General concept: Reduce event
- Scouting streams cherry pick re
  - First use in 2011 stored just jet
- **No RAW detector data stored**
- **Benefit**: reducing the event con
  - Reduce the detector readout ba
  - Access lower mass/lower pT ob
- **Downside**: Less information av

*Leads to first use of parking strea

~10,000 events/s*

***Standard stream: ~1000 events/s**

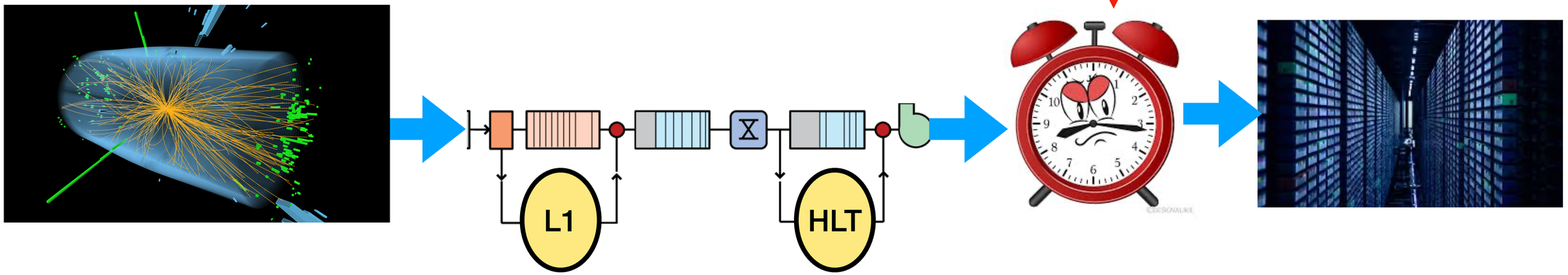




Data Parking

*Delayed Stream = the analog in ATLAS

- First use case of parking in CMS was to cross-check the early scouting streams.
- General concept: Store large amount of RAW detector data; reconstruct when computing resources more available.
- Parking streams store directly to TAPE the RAW detector data
  - NO prompt reconstruction performed
- **Benefits**: No prompt reco -> we can increase event throughput
  - Lower trigger thresholds, store additional interesting events.
- **Downside**: Reconstruction must **wait** for computing resources to become available.
  - Typically this means when the detector is not collecting data.
  - Data unavailable for examination by experimentalists.



Data Parking

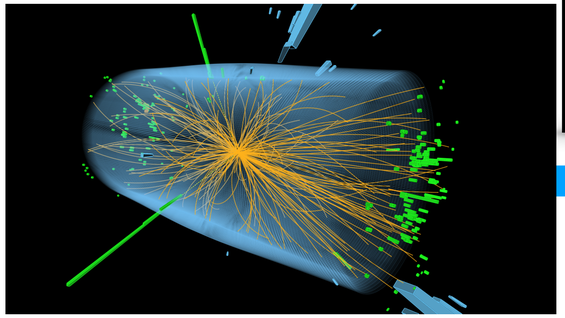
*Delayed Stream = the analog in ATLAS

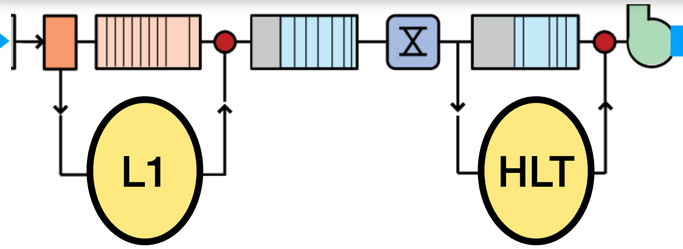
- First use case of parking in CMS w
- General concept: Store large amou mputing resources more available.
- Parking streams store directly to T
  - NO prompt reconstruction perform
- **Benefits**: No prompt reco -> we c
  - Lower trigger thresholds, store ad
- **Downside**: Reconstruction must w ble.
  - Typically this means when the dete
  - Data unavailable for examination b

~2 GB/s*
(2018)

Double the dataset size

***Standard stream: ~1GB/s**



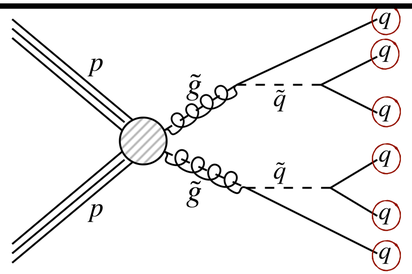


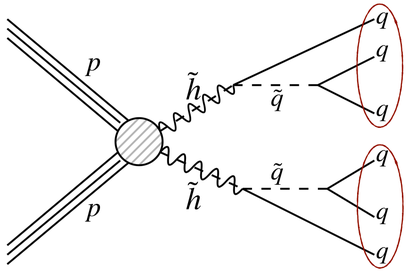
Scouting and Parking in Action

Submitted to PRL: Arxiv

Search for Mulitjet Resonances

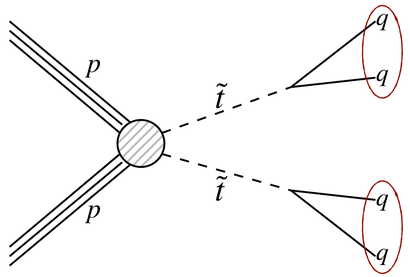
- Search for
  - Boosted dijet and trijet events
  - Resolved trijet events
- Context RPV SUSY
- First search involving hadronically decaying higgsinos.
- Dataset: PF scouting H_T 410
  - Enriches search by using low summed scalar jet transverse momentum requirement
  - Low-mass phase space accessible
- Perform bump hunt of multijet mass
- No new resonances observed





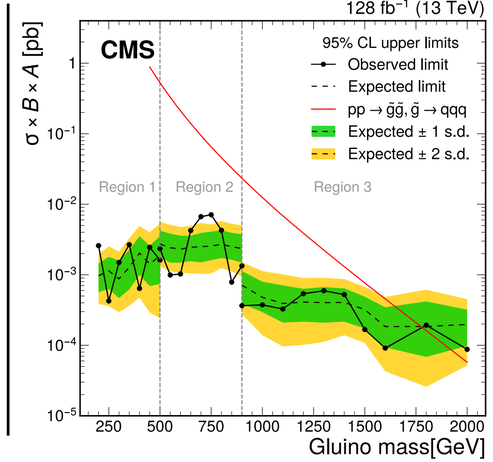
New

Low mass results

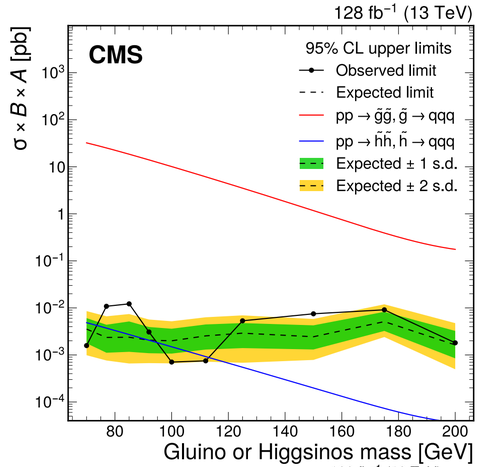


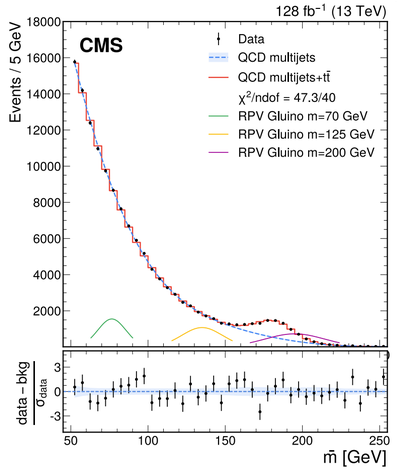
Improved limits

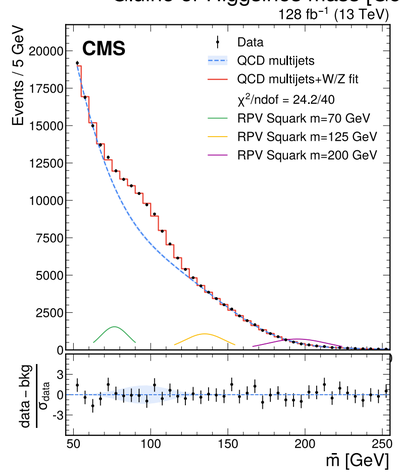
Resolved Regime

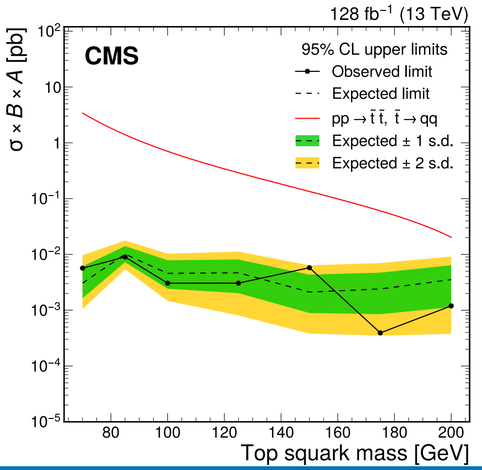


Boosted Regime







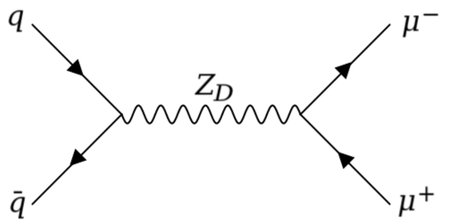


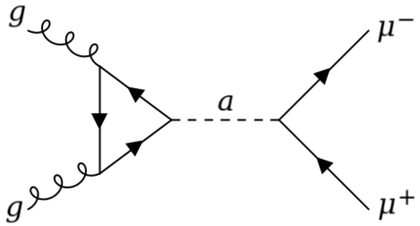
10.1007/JHEP12(2023)070

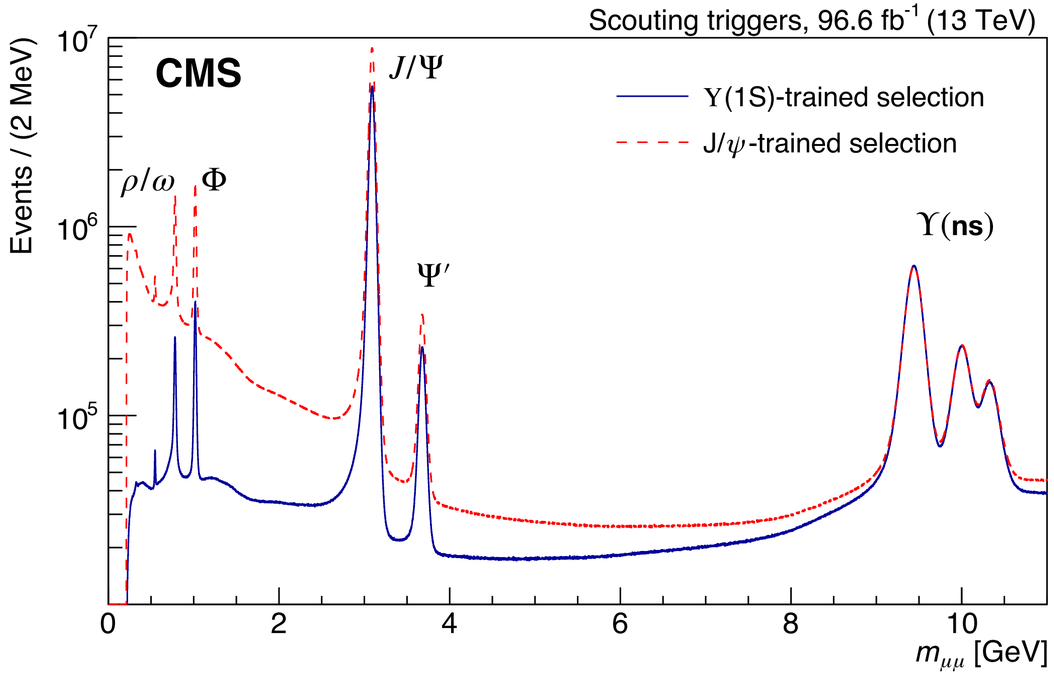
Low Mass $\mu\mu$ Resonance Search

- Search for direct production of dimuon resonances at low mass
- Dataset: 2017+2018 **muon scouting** streams.
- BSM mediator mass range: **1-8 GeV**
  - Analogous effort with standard data stream limited to mediator masses **> 45 GeV**.
- Two mass range bands considered
  - 1.1 - 2.6 GeV
  - 4.2 - 7.9 GeV

✳Chosen to exclude J/ψ, ψ(2S), and Υ(1S) resonances

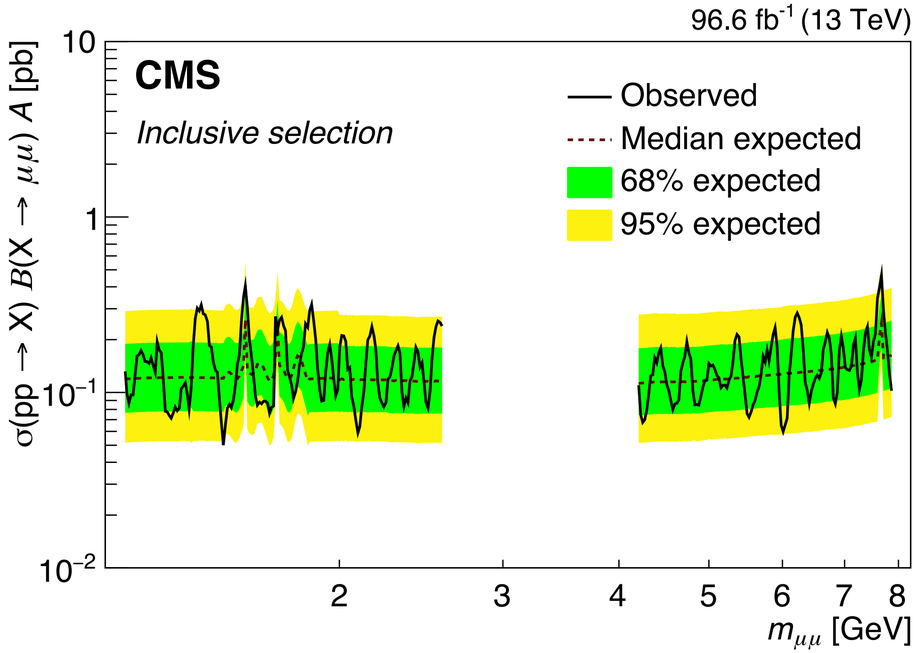






10.1007/JHEP12(2023)070

Low Mass $\mu\mu$ Resonance Search

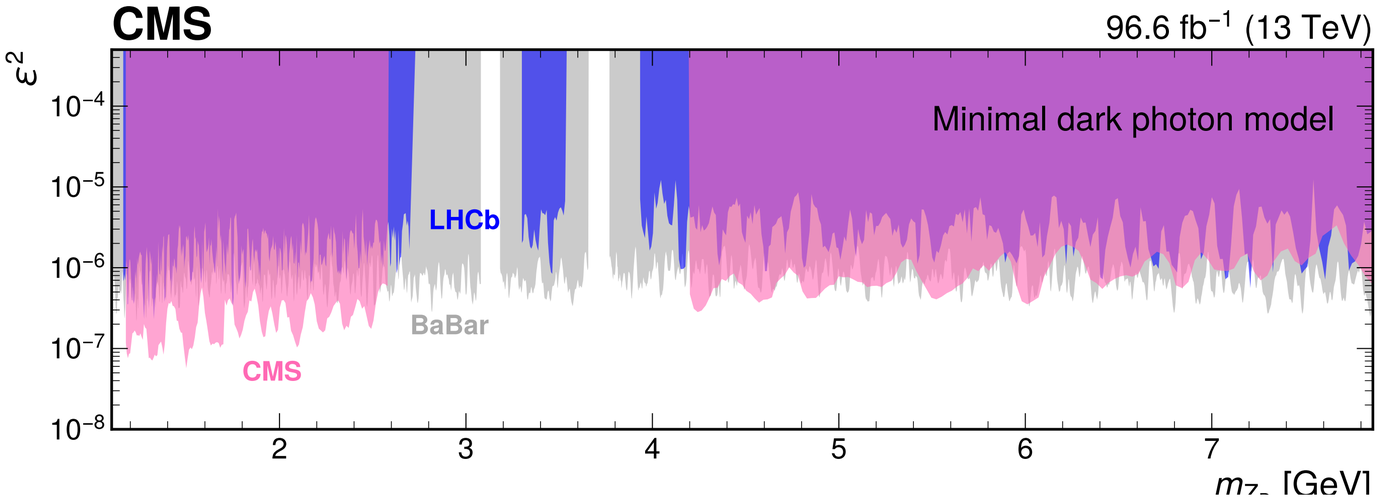


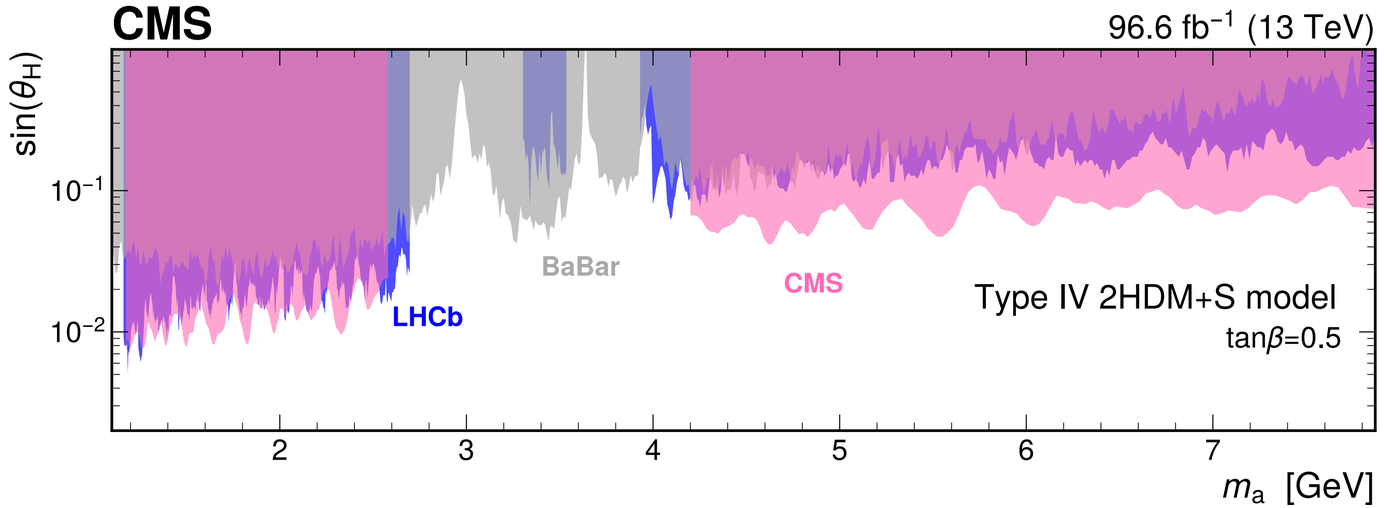
- No significant excess observed with respect to the SM prediction.
- 95% CL limits on the cross-section for a mediator decaying to a dimuon pair is set at 0.10 (0.15) pb for the 1.1-2.6 (4.2-7.9) GeV mass range.
- Results further interpreted for model specific interpretations
  - Dark photon
  - 2 Higgs Doublet + Scalar (2HDM+S)

10.1007/JHEP12(2023)070

Low Mass $\mu\mu$ Resonance Search

- Additional limits set for the
  - Dark photon coupling (ε)
  - 2HDMS mixing angle (sin(θ_H))

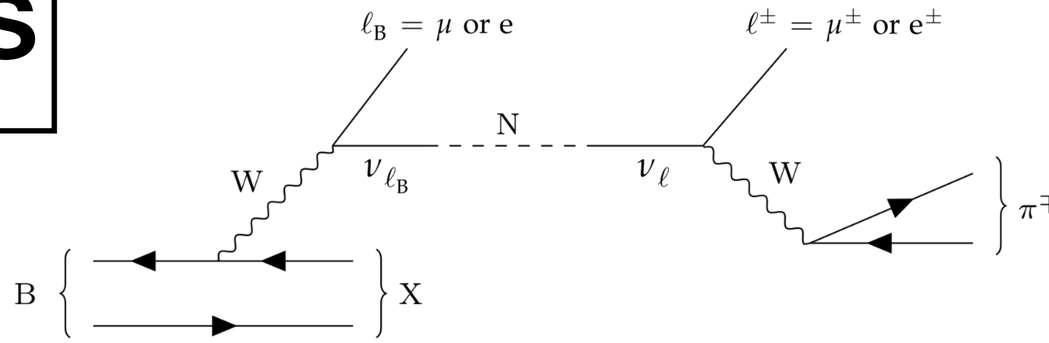


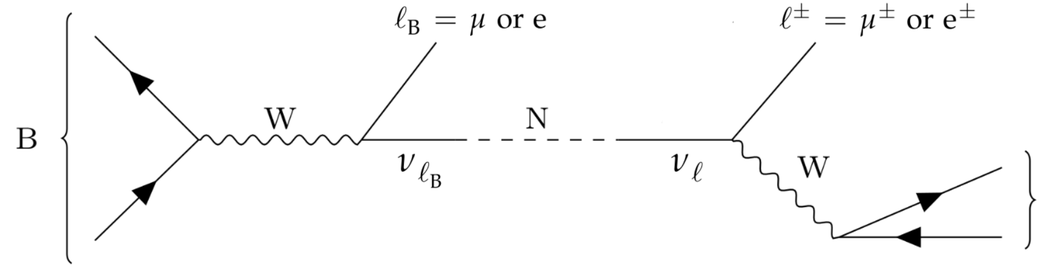


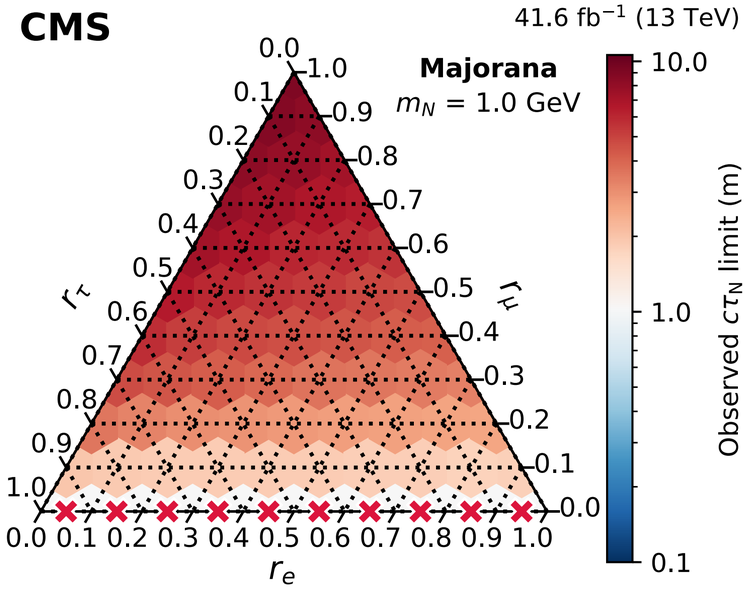
Search for long-lived HNLs

10.1007/JHEP06(2024)183

- First time B initiated HNLs searched in CMS.
- Dataset: 2018 **B-Parking** dataset (~10^{10} bbar decays)
- **B meson** as production mode
  - More neutrinos (compared to W production mode)
  - Lower mass (than W)=> lower signal pT (more signal acceptance)
- For more info, see Martin's talk.





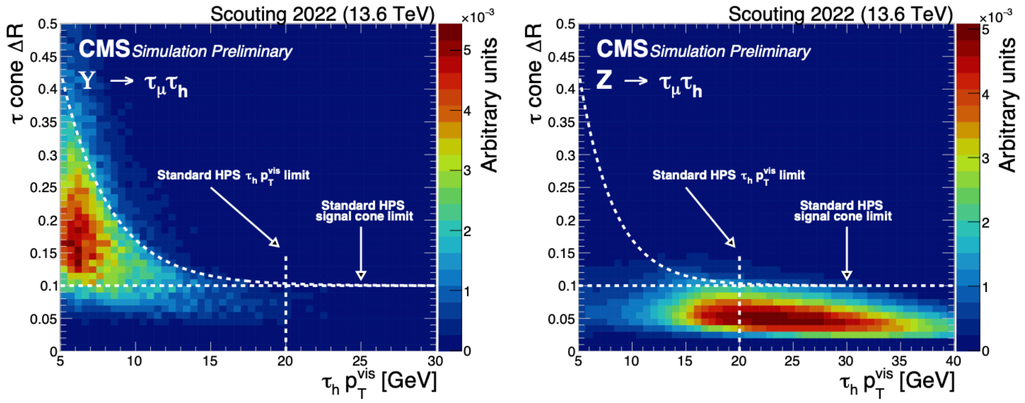


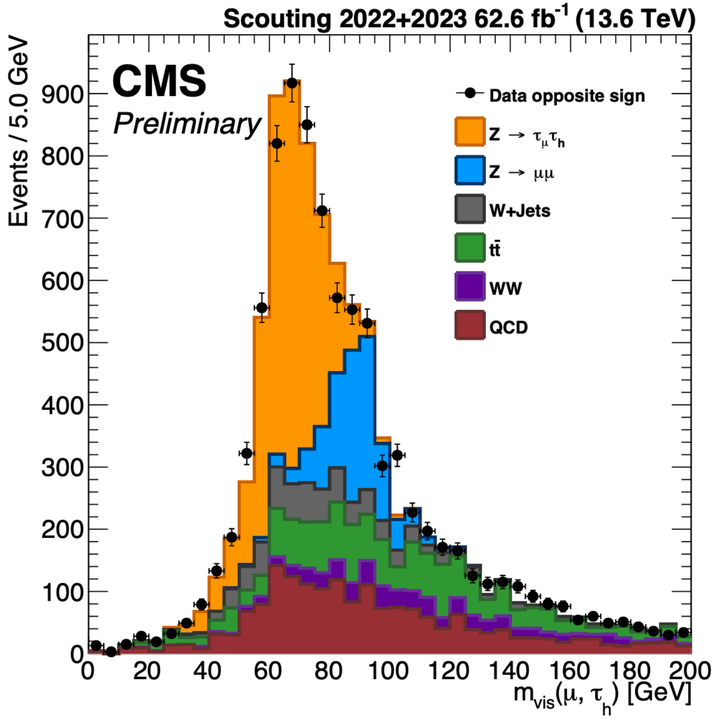
Low-p_T τ_h Reconstruction

CMS NOTE-2024/006

NEW

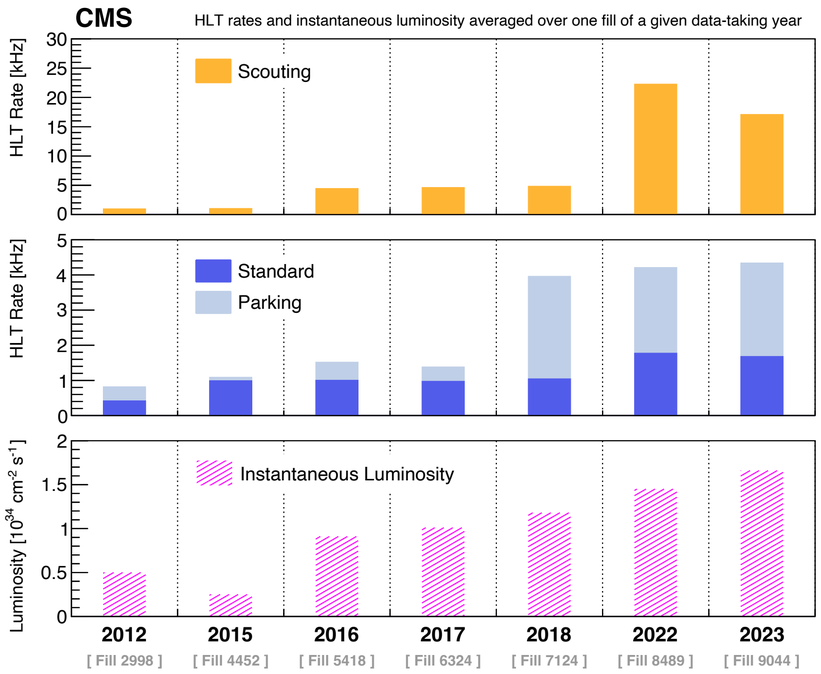
- New object added to scouting stream.
- Based on HPS algorithm used in standard stream.
- Features an expanded jet cone wrt to standard method.
  - Allows algorithm to extend to low-p_T τ_h.
- Uses OR of HT EGamma scouting streams.
  - e/γ, K_L^0, $\pi^{\pm}$, and $\mu^{\pm}$ candidate objects exist in scouting sets.
- **Developed new object at L1 for τ_h.**
  - Lowered p_T threshold to reconstruct τ_h from **20 to 5** GeV.
  - Modified HPS algorithm used in standard stream.
  - Successfully reconstruct $Z \rightarrow \tau_\mu \tau_h$ peak.





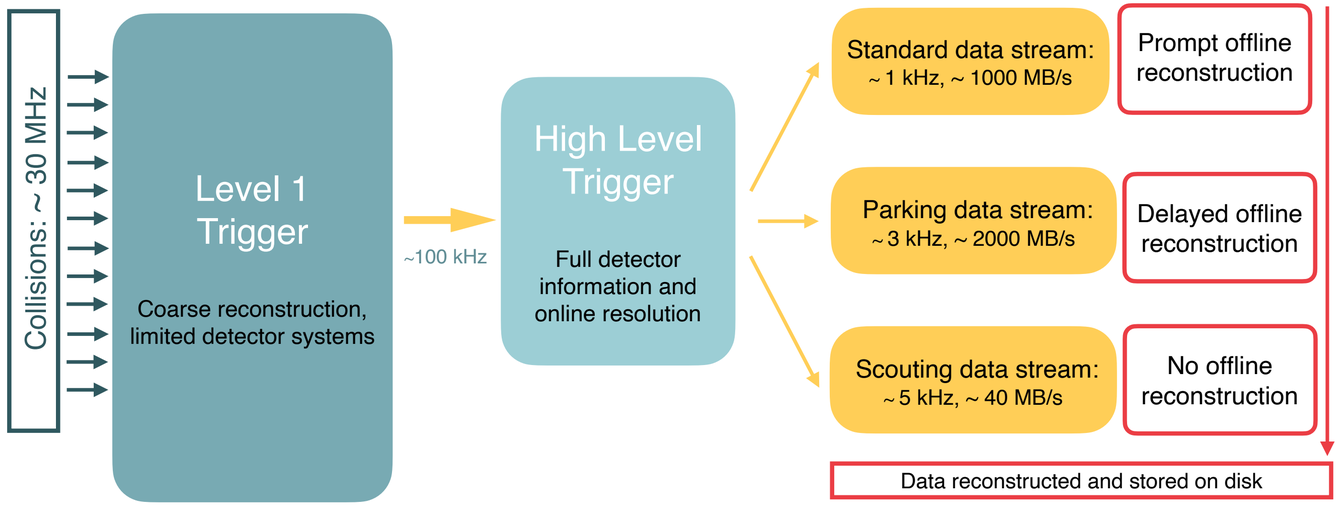
Scouting and Parking Summary Paper

Submitted to Physics Reports: Arxiv



- Comprehensive paper detailing the history and use of the scouting and parking datastreams within CMS.
- Physics results from Run 2.
- Parking/scouting strategies for Run 3.
- Much more detail about the streams that cannot be covered in this talk.

Data flow for a typical 2018 data-taking scenario



Summary

- Although the CMS detector is reaching physical limitations in the storing and accepting of data from LHC collisions, we have adapted our workflow to enable specialized data streams circumventing these limitations.
- The **Scouting** data stream enables analysts to select data using much lower trigger thresholds than the standard data stream.
  - Lower thresholds enable additional low mass/low pT searches
- The **Parking** data stream bypasses the computational limit faced when reconstructing data, doing so enables CMS to store much more events.
  - The additional events present unique opportunities to explore a dataset with objects that would otherwise be excluded in the standard data stream.
- CMS has utilized these data streams to produce multiple leading results during Run 2 and the scouting/parking program continues to grow in Run 3.

End