

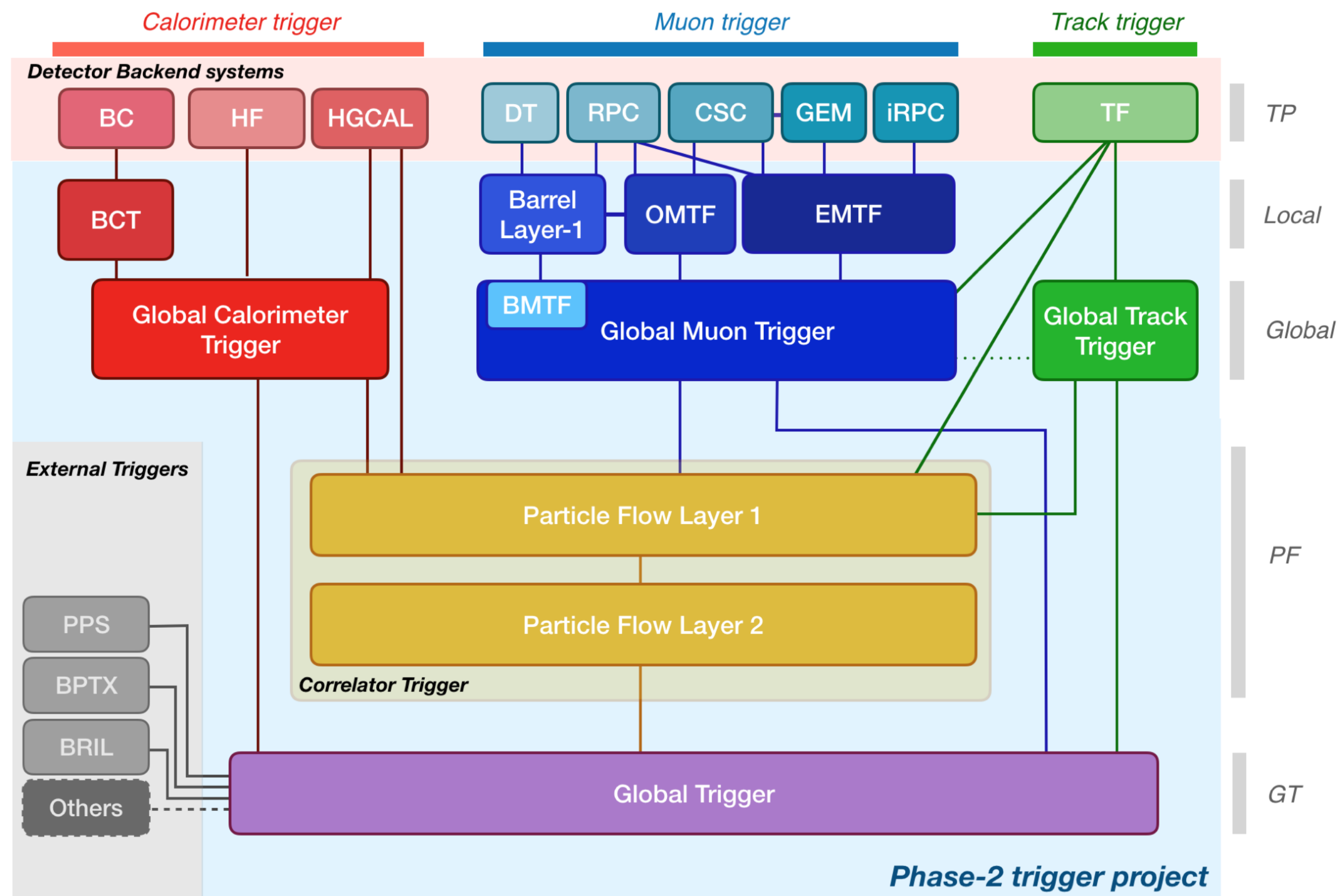
Development and demonstration of the CMS Phase-2 Level-1 trigger Data Scouting baseline system for HL-LHC

TWEPP2025, 10 October 2025, Rethymno

Rocco Ardino¹ for the CMS Collaboration

¹CERN, Geneva, Switzerland

HL-LHC: CMS Level-1 Trigger upgrade



Keep and/or improve physics acceptance at PU200

- $\sim 12.5 \mu\text{s}$ latency
- $\sim 750 \text{ kHz}$ of L1 output rate

Advanced object reconstruction on FPGA

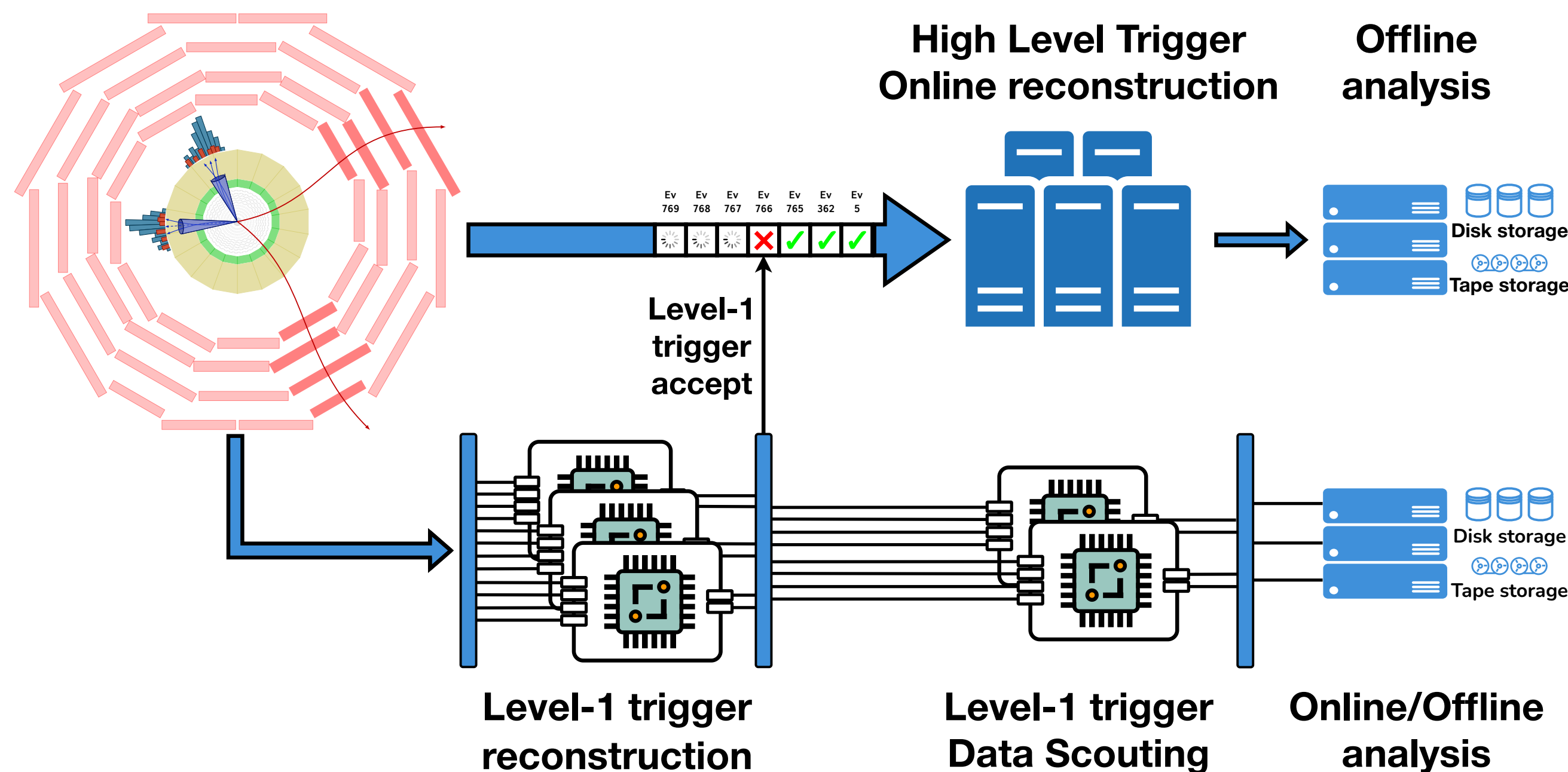
- **Global Calorimeter Trigger** (GCT) and **Global Muon Trigger** (GMT) (higher granularity)
- **Global Track Trigger** (GTT) (vertex finding)
- **Correlator Trigger** (CT) (Particle Flow)
- **Global Trigger** (GT) (more complex cut-based and machine learning algos)

L1 objects resolution comparable to offline one \Rightarrow Suitable for physics analysis, not only triggering!

HL-LHC: CMS Level-1 Trigger Scouting

Make sure we are not missing anything interesting in the data produced by HL-LHC

- Capture intermediate data reconstructed by the L1 trigger at the 40 MHz bunch crossing rate
- **Perform online analysis on the collected data, bypassing the L1 selection rejection**
- **Complement CMS standard physics program for HL-LHC**



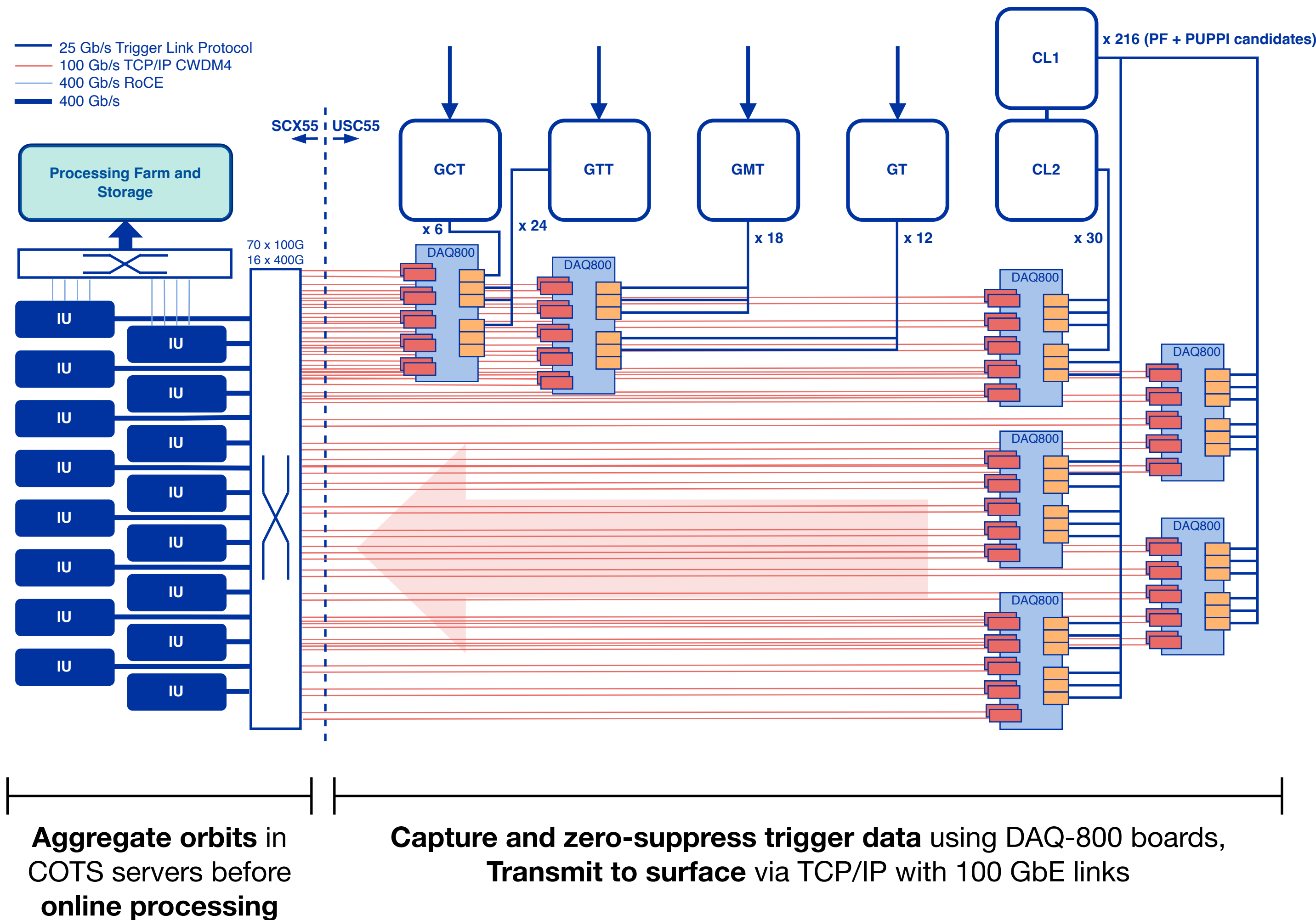
L1 scouting as powerful and versatile tool

- Monitor at 40 MHz rate
- Analyze topologies too frequent for selection or requiring too high latency for identification at L1
- Analyze inter-BX correlations

Already producing physics result in Run-3

- Run-3 L1 Scouting **demonstrator**
 - [PhD thesis](#) by R. Ardino
 - [PhD thesis](#) by M. Migliorini
- [CMS-PAS-EXO-25-010](#)

CMS Phase-2 L1 Scouting: architecture



Readout trigger links with 7×DAQ-800

- Phase-2 cDAQ readout board
 - [Talk](#) by P. Tzannis
- Input:** 48×25 Gb/s trigger links
- Zero suppression and concentration
- Output:** TCP/IP streams via 10×100 Gb Ethernet links to surface

Buffer data in set of COTS servers

- Ingestion Units**
- Receiver software reformats the data and produces fragments for processing

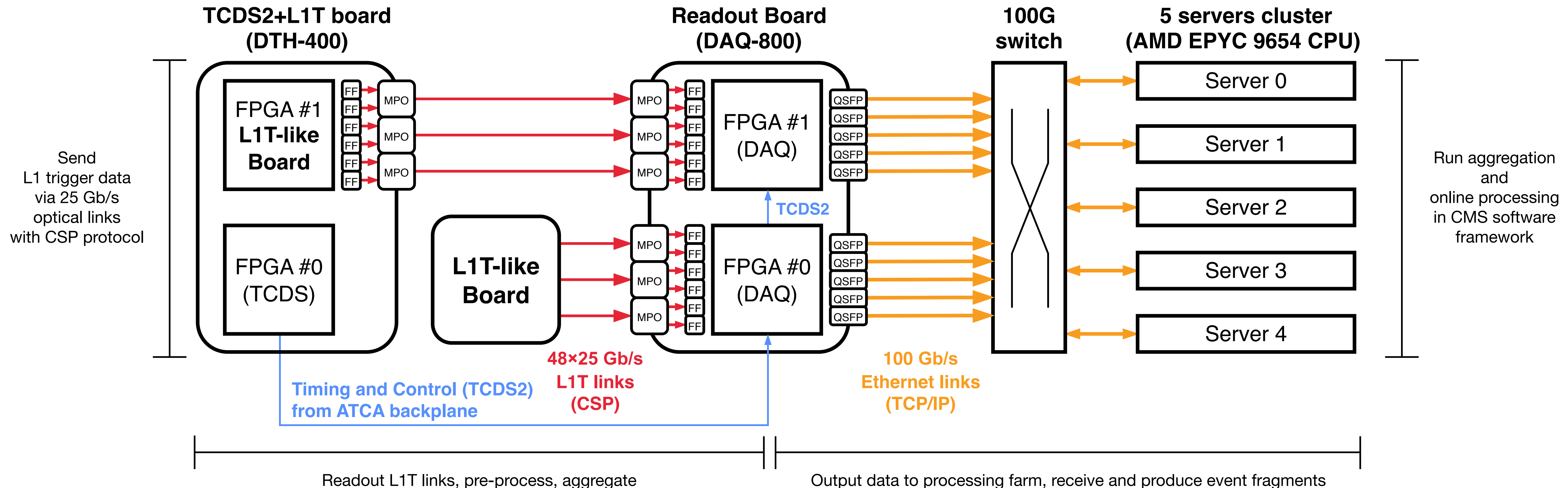
Inject data to processing farm

- Event aggregation + Online analyses in CMS software framework (CMSSW)**
- Throughput to Tier-0 around 10% of the cDAQ one

CMS Phase-2 L1 Scouting demonstrator

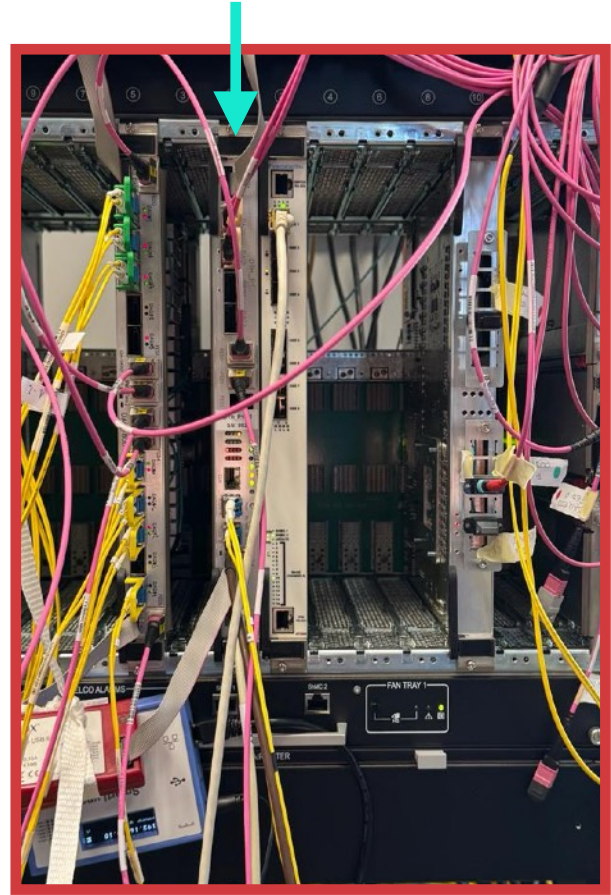
Demonstrator assembled in CMS DAQ laboratory at CERN to develop the system for HL-LHC

- First version summarized in [CMS-DP-2024-096](#) presented at [CHEP2024](#)
- Extended version summarized in [CMS-DP-2025-058](#) presented at [ACAT2025](#)

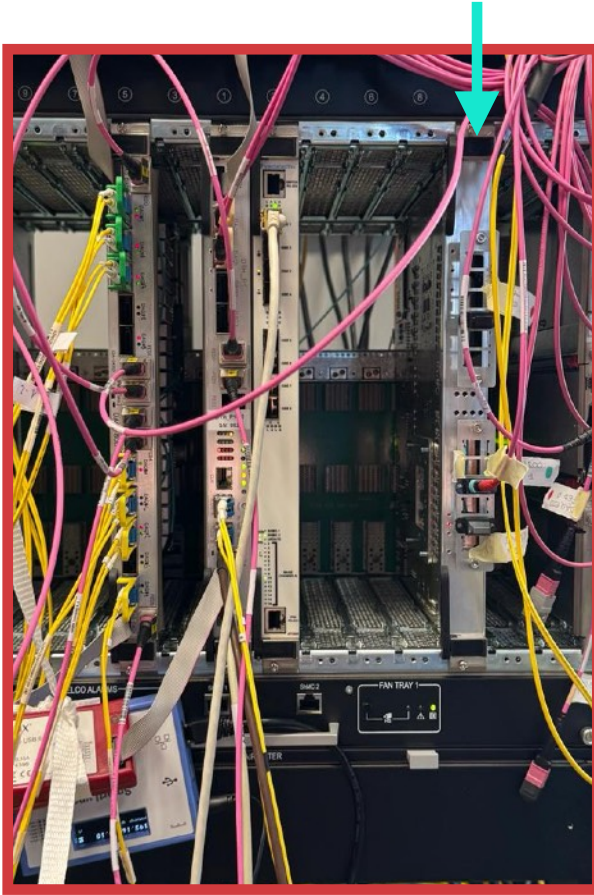


L1T readout demonstration

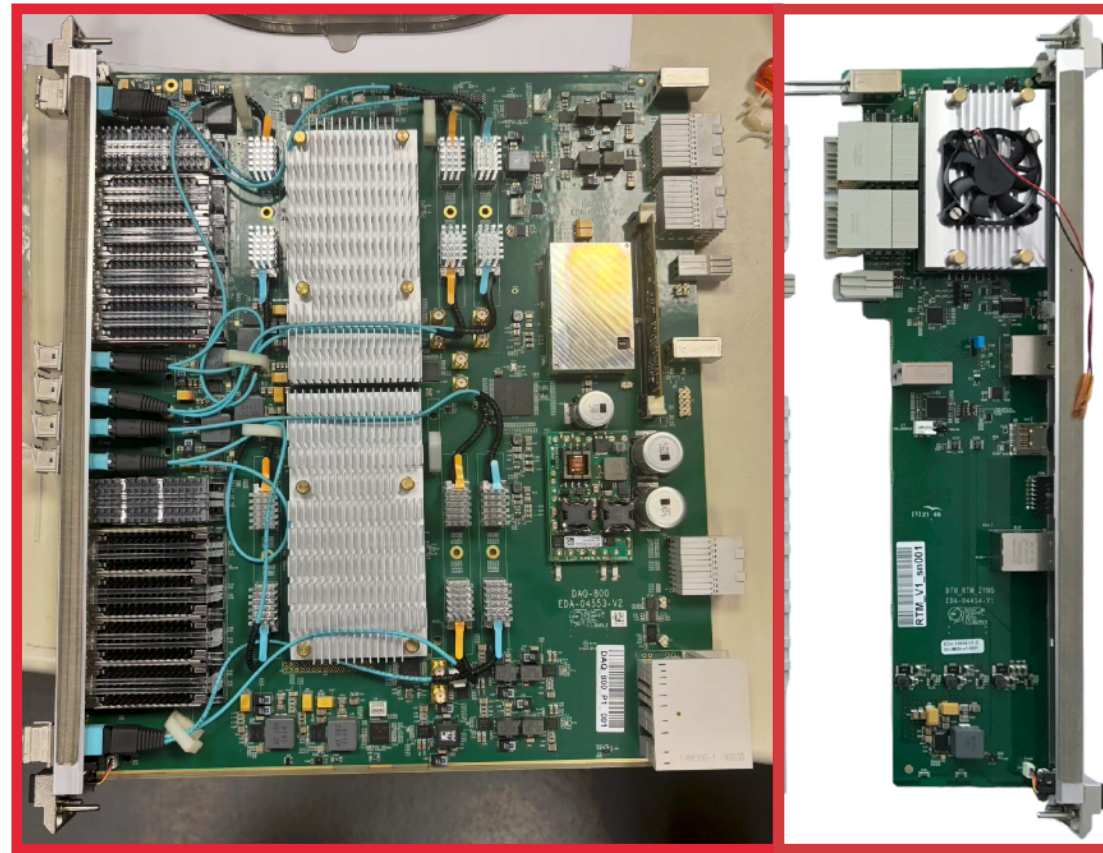
DTH TCDS + L1 source



Serenity (L1 source)

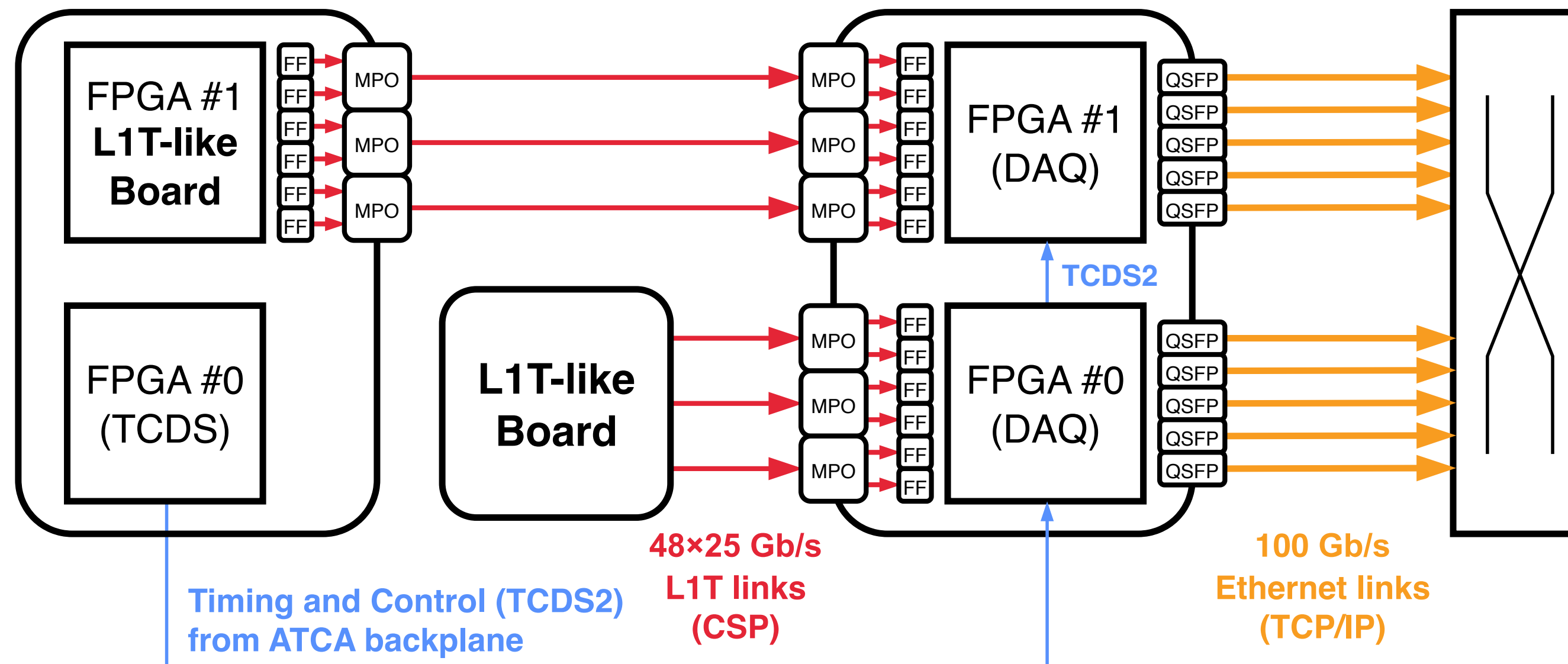


DAQ-800 (+ RTM)

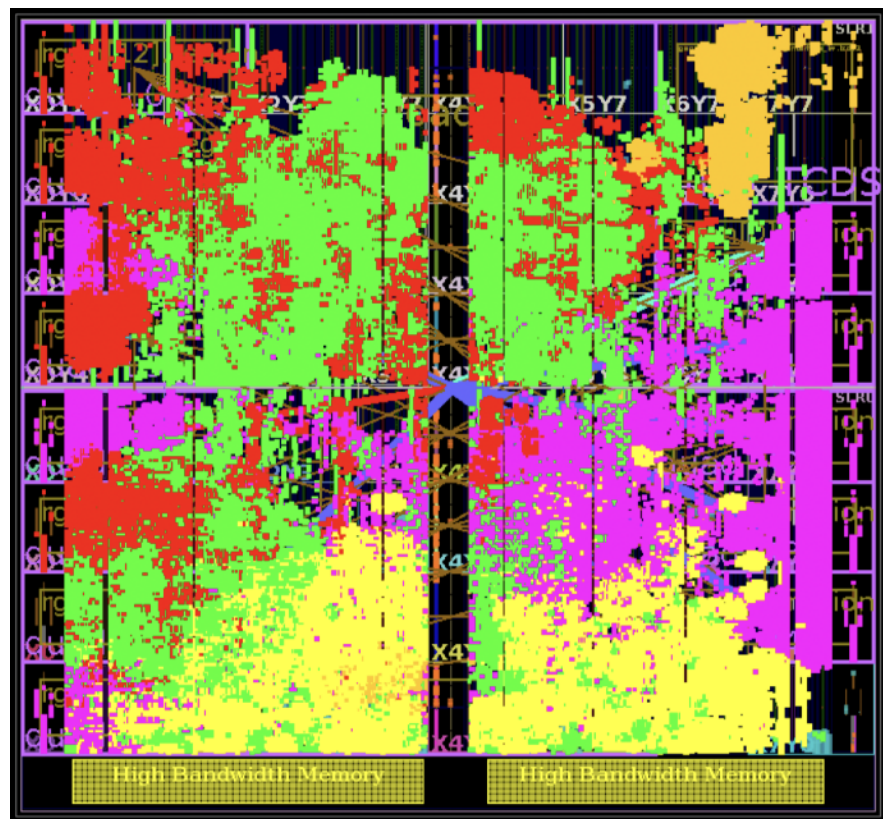
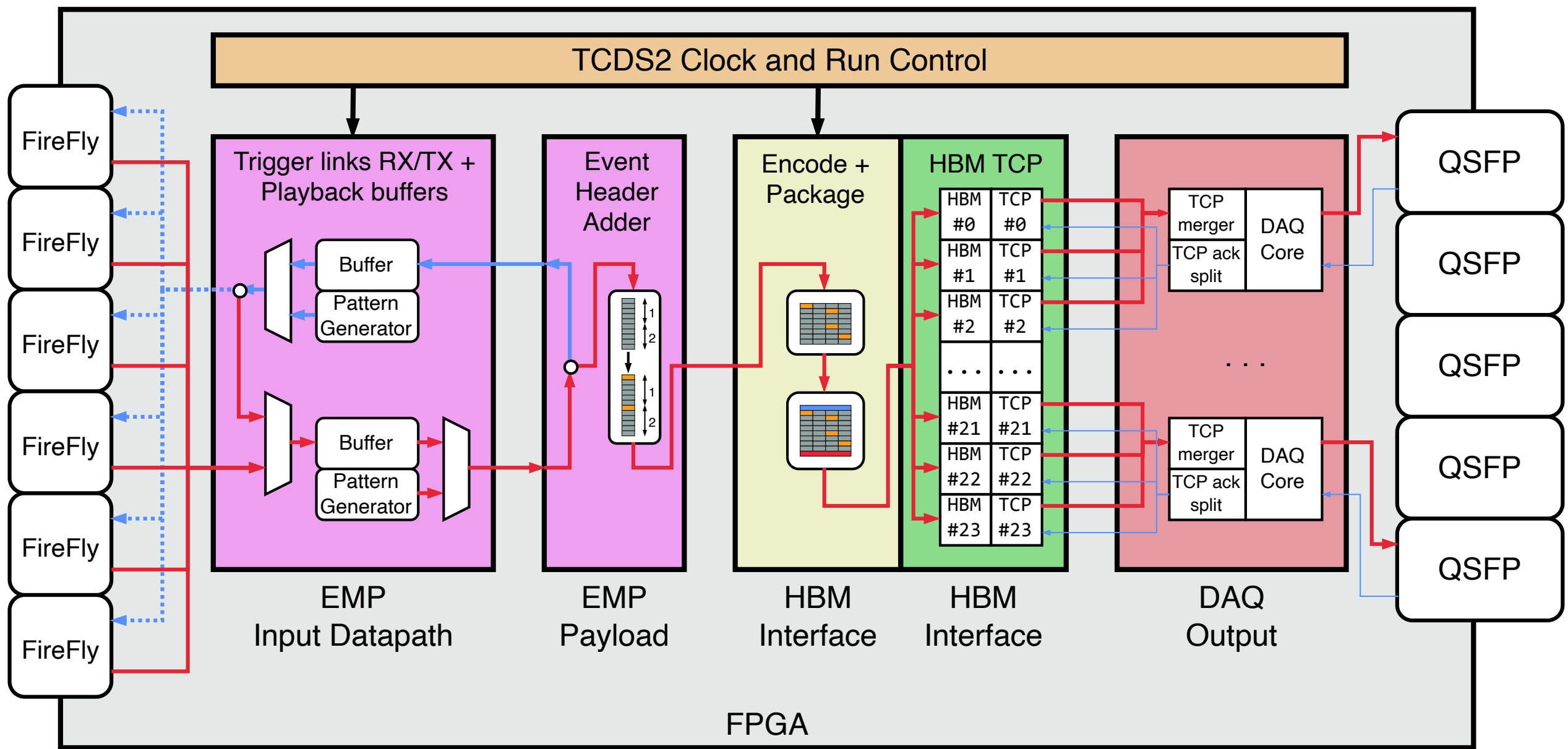


Hardware for L1T readout demonstration

- **Serenity (L1-like data source, 24 links)**
- **DTH-400 prototype**
 - Lower FPGA: **TCDS2 signals** for 40 MHz LHC clock, OC1, run start/stop
 - Upper FPGA: **L1-like data source (24 links)**
- **DAQ-800 prototype**
 - Receive 24 × L1T links per FPGA unit via FireFly
 - Buffer data in 8 GB High-Bandwidth Memory on FPGA
 - Output via multiple TCP/IP streams over 5 × 100 Gb Ethernet links per FPGA
- **Rear Transition Module (RTM)**
 - Trenz System-on-Module running AlmaLinux9
 - Communicate with DAQ-800 board via AXI
 - Board control and monitoring software running on SoC
- **100 GbE switch**



DAQ-800 firmware



Resource	Usage	Usage (%)
FF	585k	33.6
LUT	319k	36.6
BRAM18	302	11.2
BRAM36	458	34.0
URAM	96	15.0
DSP	8	0.1

Floorplan and resource usage for 24xL1T inputs, 24xHBM/TCP streams, 4x100 GbE cores

DAQ-800 firmware in EMP framework

- **Input Datapath + Payload (360 MHz)**
 - Get data from L1T links or load collision patterns
 - Add header to BX record to encode number of objects
- HBM interface (at 250 MHz)
 - Encode data in 256-bit frames (HBM alignment)
 - Package in orbits (3564 BXs) with header and trailer
 - Buffer to HBM (256 MB / TCP) before TCP logic
- **DAQ Output**
 - Merge TCP streams to 100 Gb Ethernet link
- **TCDS2 signals for 40 MHz clk and run start/stop**

Firmware configurations

- General config: 1 trigger input → 1 TCP stream
- Possible to aggregate up to 4 L1T inputs → 1 TCP
- Resource usage < 50% in “most expensive” case

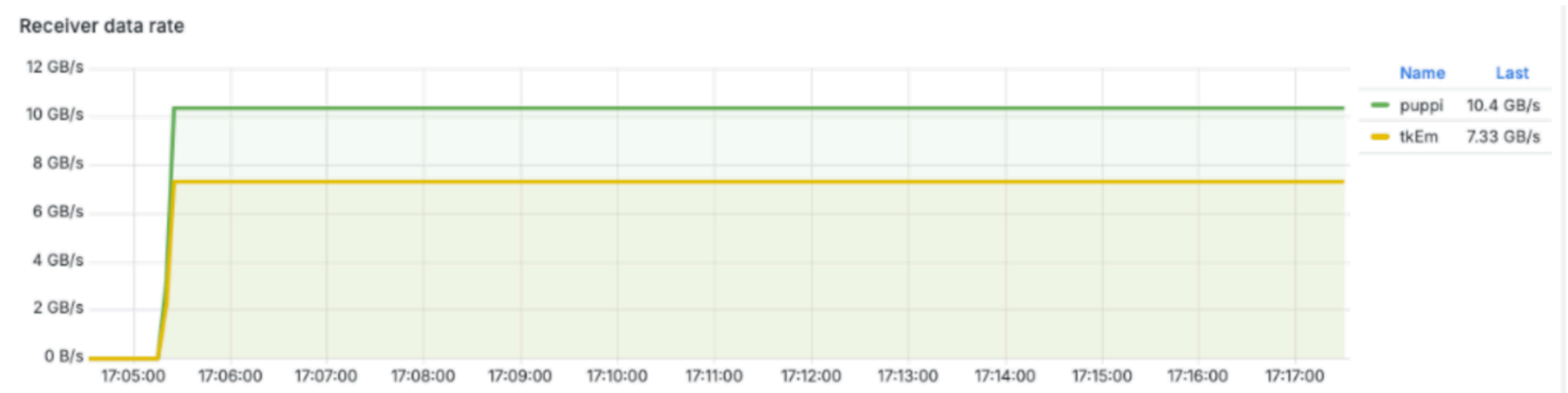
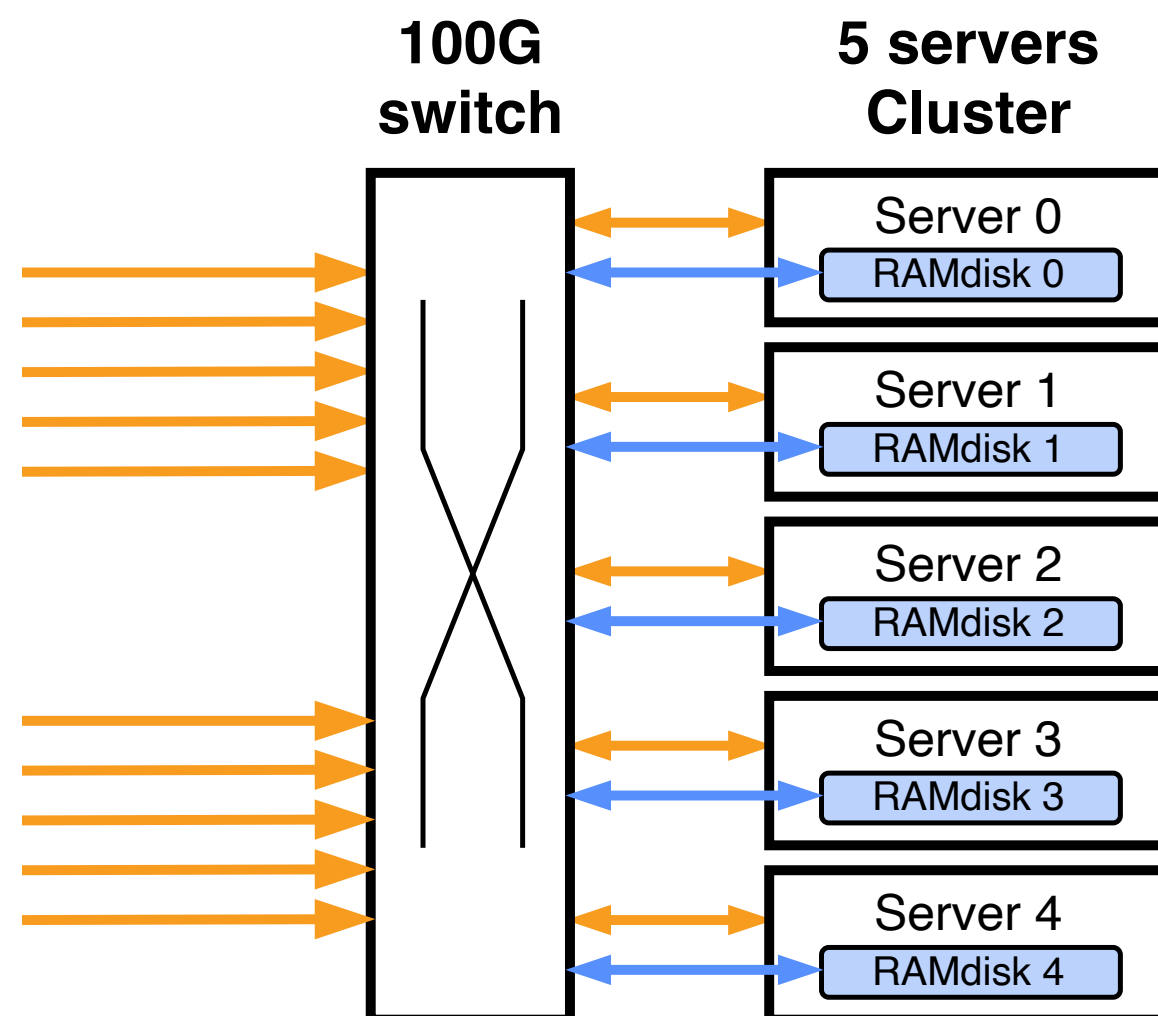
Data Acquisition demonstration



[Processing cluster](#) (5 × servers)

Processing running on cluster of 5 servers provided by [NGT](#) funds

- AMD EPYC 9654, 768 GB of RAM and 2x100G NIC
- Application receiving data (simulated PU200 collisions) from TCP/IP streams
 - 6 streams of L1 pileup-subtracted PF candidates (~1.7 GB/s each)
 - 6 streams of standalone e/γ candidates (~1.2 GB/s each)
- Producing event fragments to be aggregated
- Writing data on shared ramdisk over NFS
- Software services on cluster orchestrated by Kubernetes



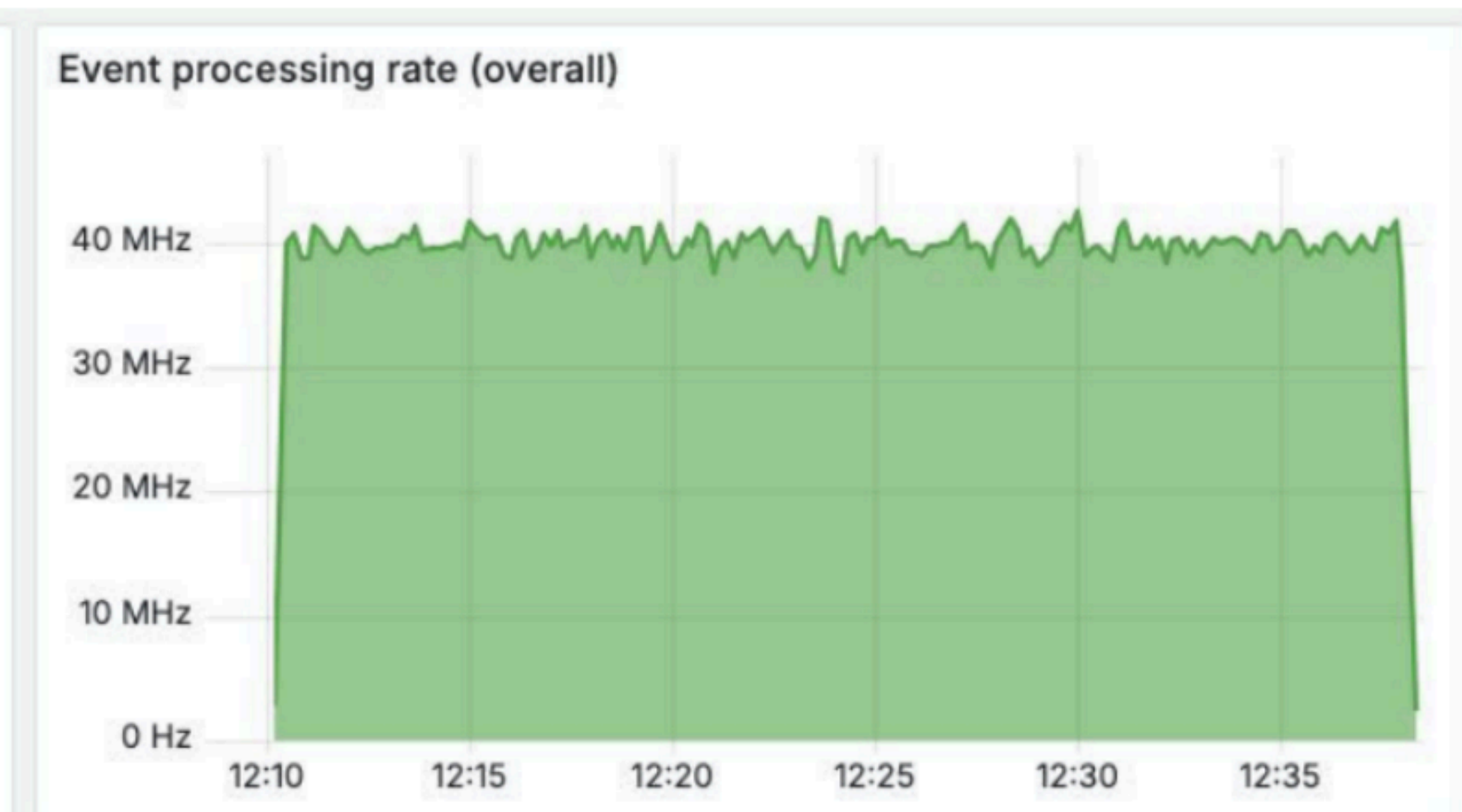
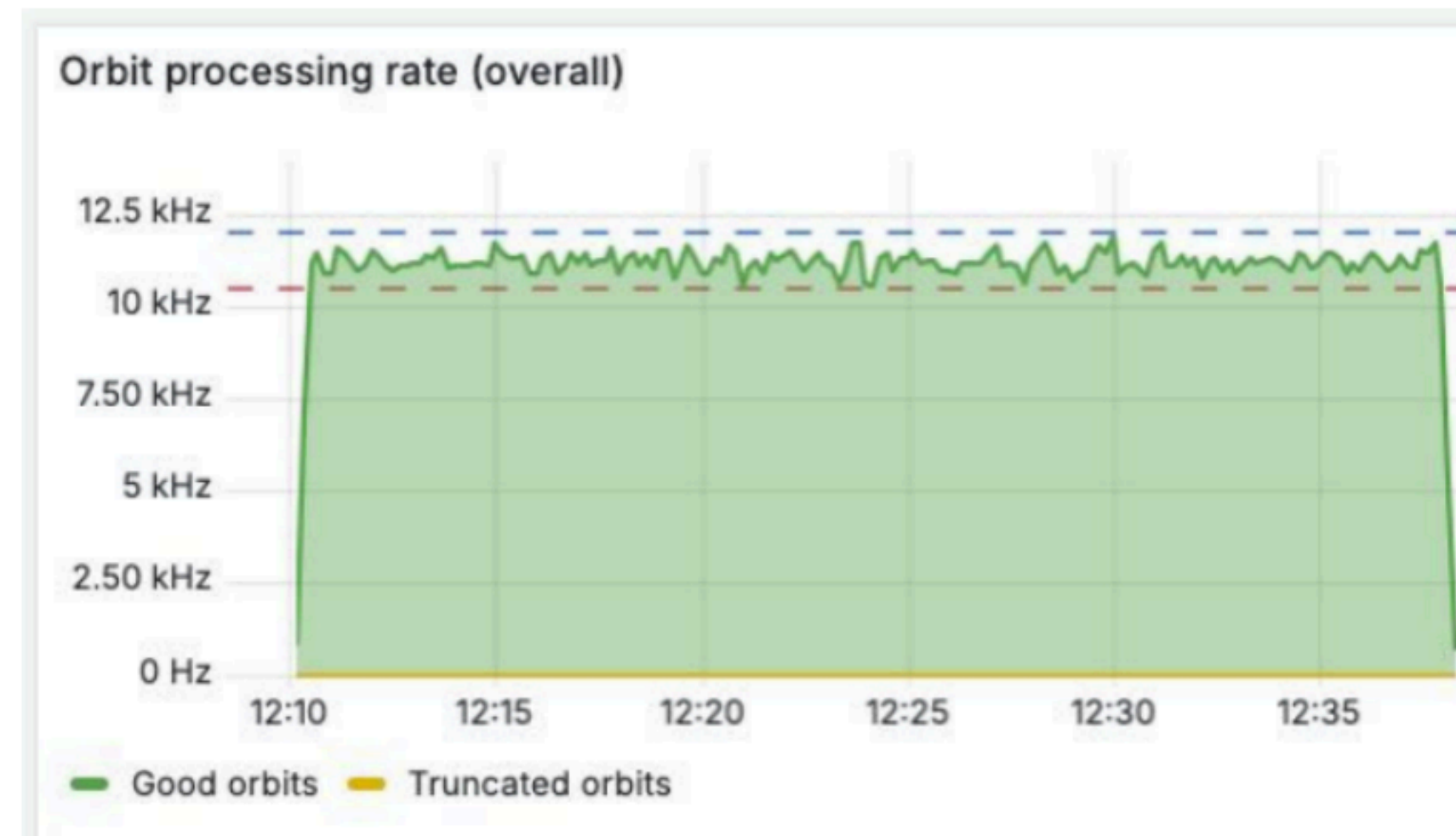
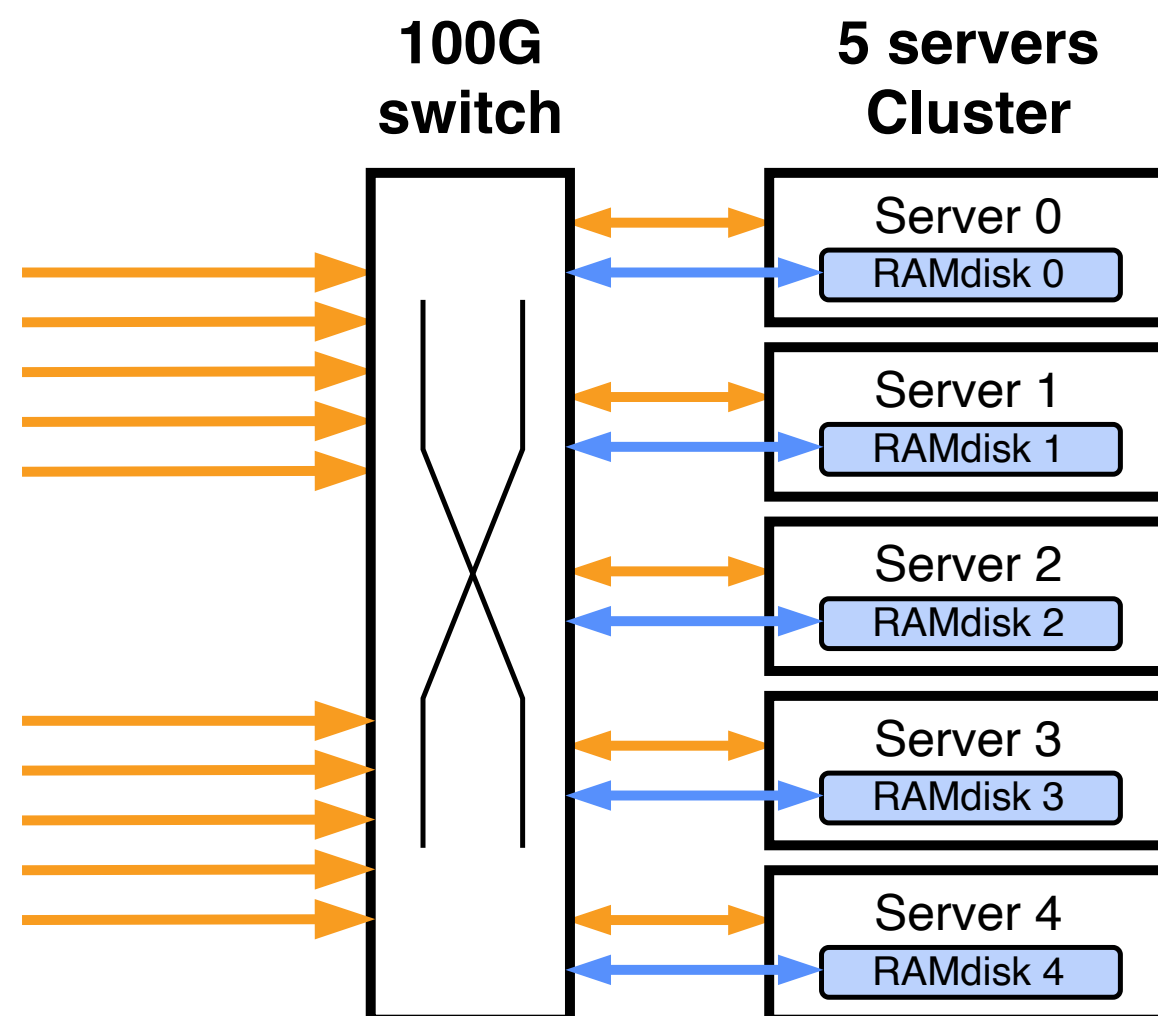
Online processing demonstration



[Processing cluster](#) (5 × servers)

Aggregation and online processing in CMS software framework

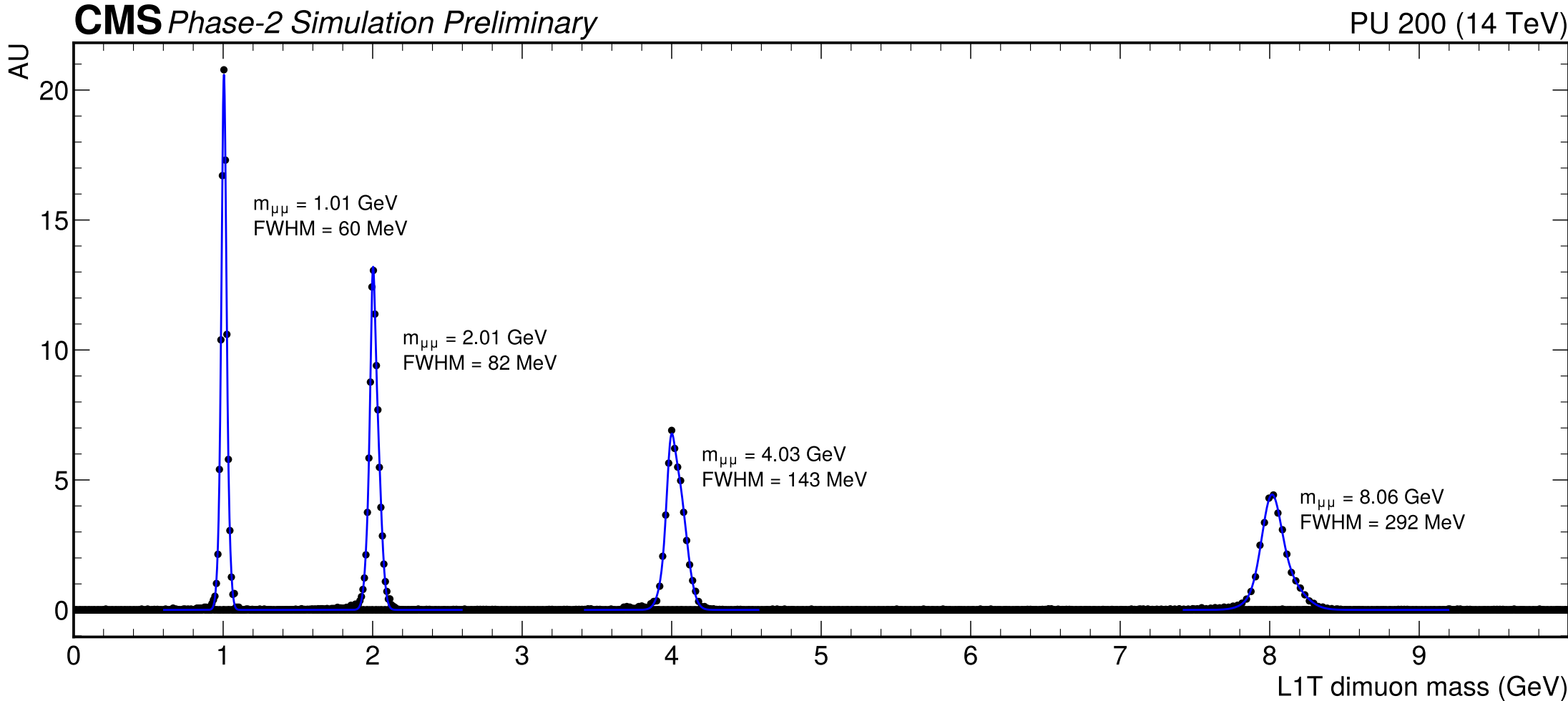
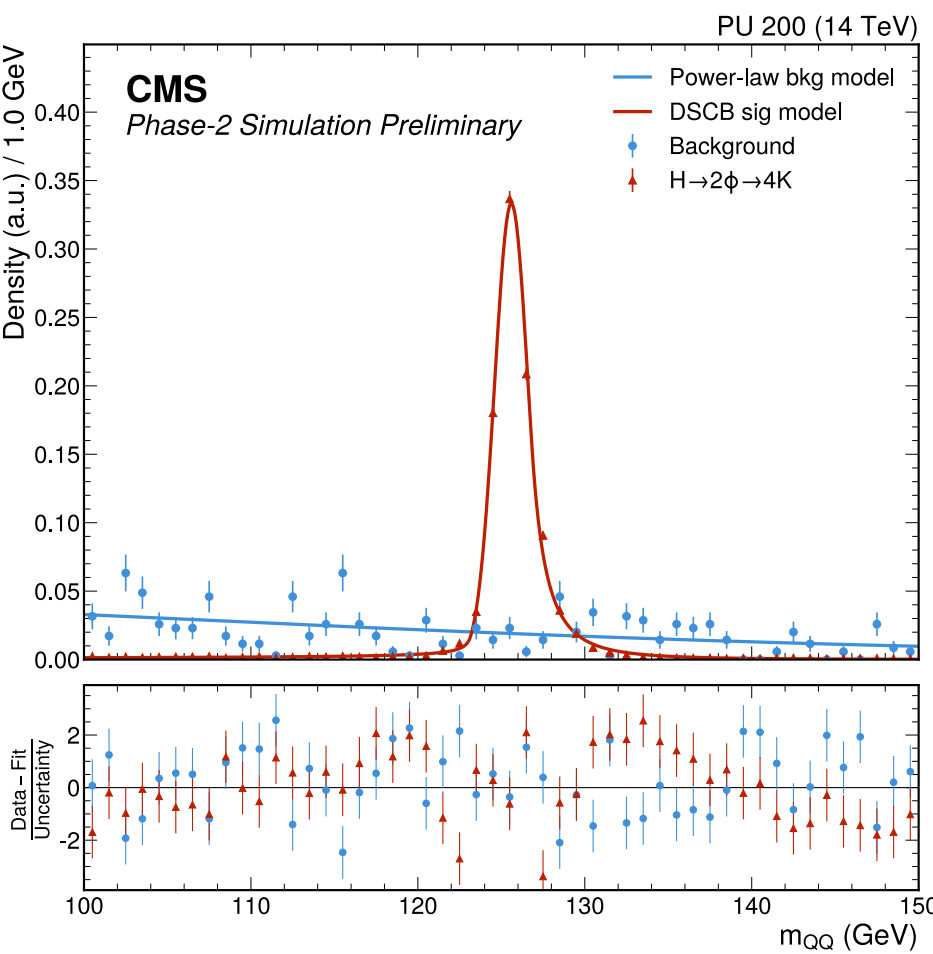
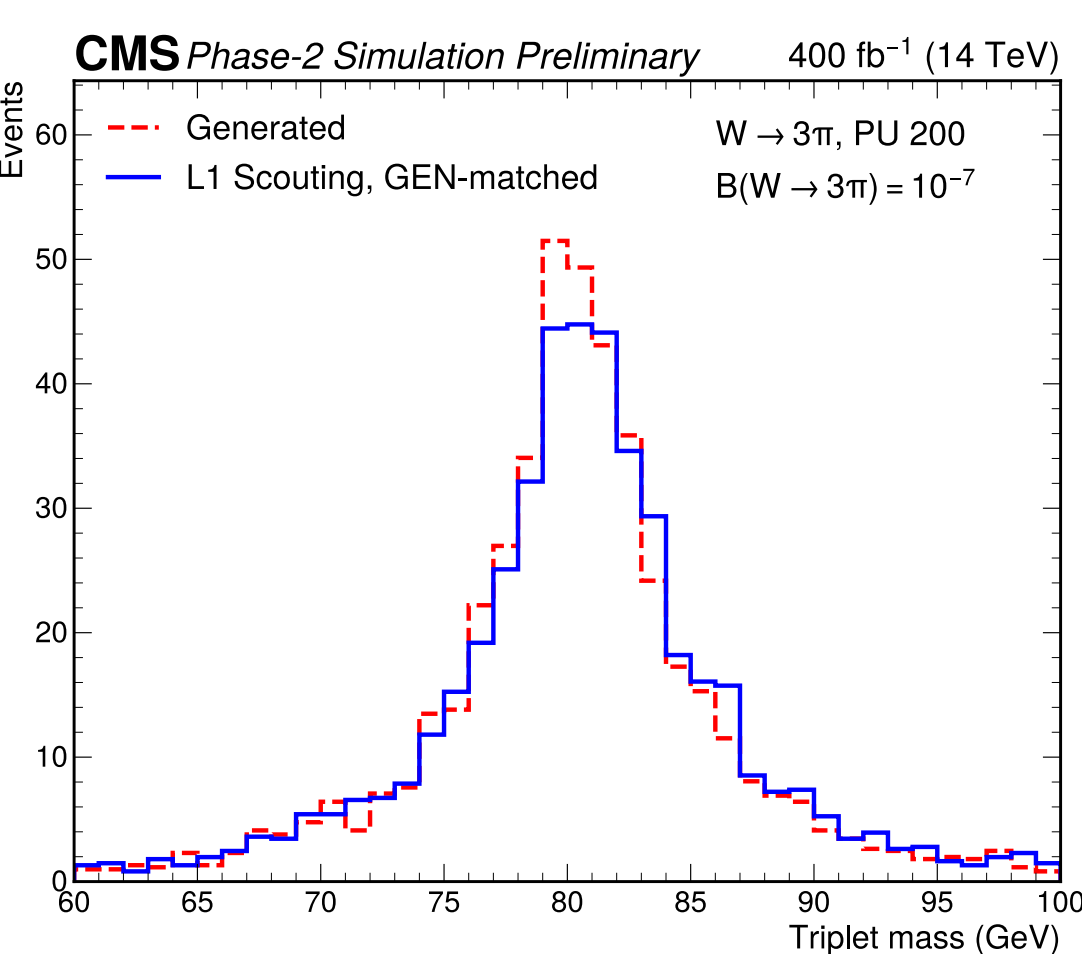
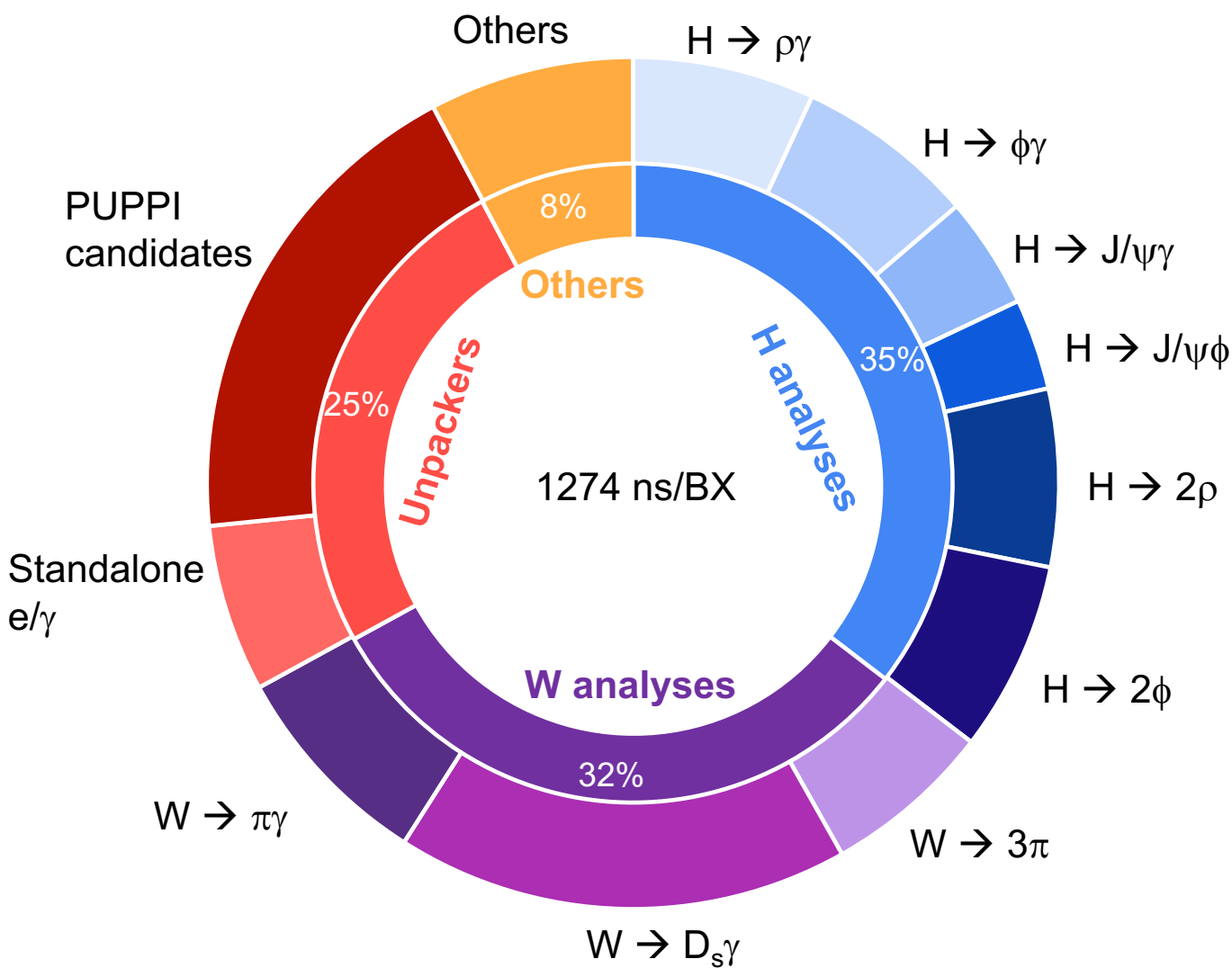
- Implementation taken from Run-3 L1 Scouting demonstrator
- File broker distributing RAW files to CMSSW processes
- Unpacking RAW data in dedicated [Orbit Collection](#) data format
- Processing in batches of 1 LHC orbits (~3564 bunch crossings)
- Spawning of CMSSW processes orchestrated by Kubernetes



Online processing demonstration

Run set of online analyses on unpacked data

- Standard Model rare decays analyses
 - $W \rightarrow 3\pi, \pi\gamma, D_s\gamma$
 - $H \rightarrow \rho\gamma, \phi\gamma, J/\psi\gamma; H \rightarrow 2\rho, 2\phi, \phi J/\psi$
- $X \rightarrow \mu\mu, ee$ (not added in timing benchmark yet)
- Produce dataset with only selected BXs by the online analyses
- 1.2 μs / orbit on average running on simulated collision events



$W \rightarrow 3\pi$: $m_{3\pi}$ invariant mass distribution ([CMS-DP-2024-096](#))

$H \rightarrow 2\phi \rightarrow 4K$: m_{QQ} invariant mass distribution ([CMS-DP-2024-096](#))

$X \rightarrow \mu\mu$: $m_{\mu\mu}$ invariant mass distribution ([CMS-DP-2025-058](#))

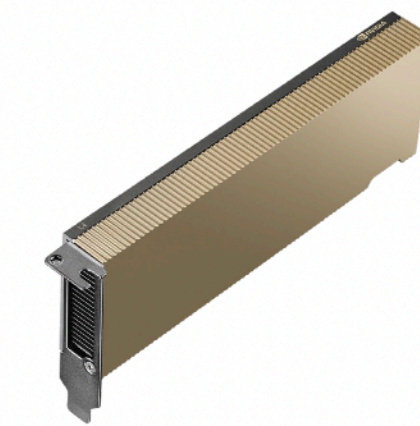
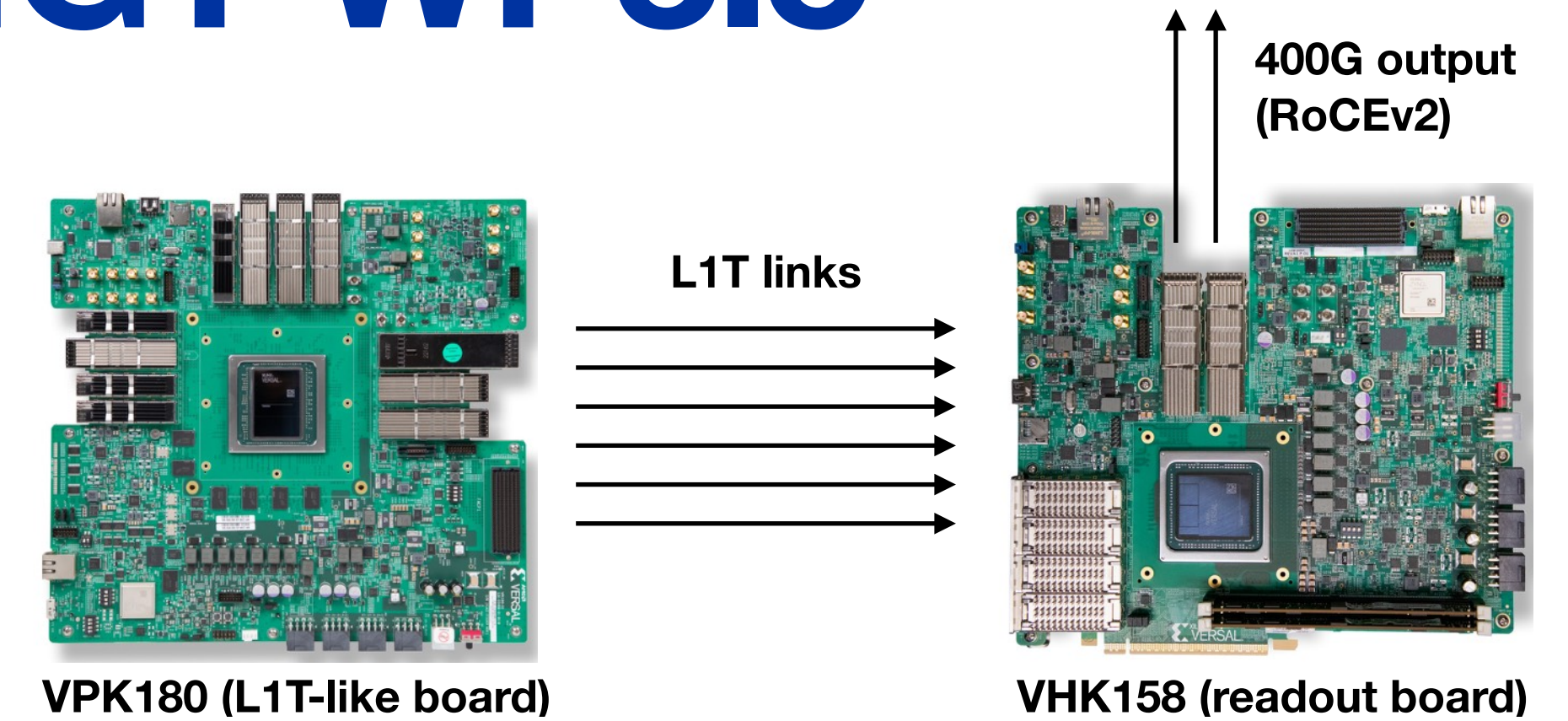
Going beyond the baseline: NGT WP3.5

NGT WP3.5 aims at extending the presented baseline

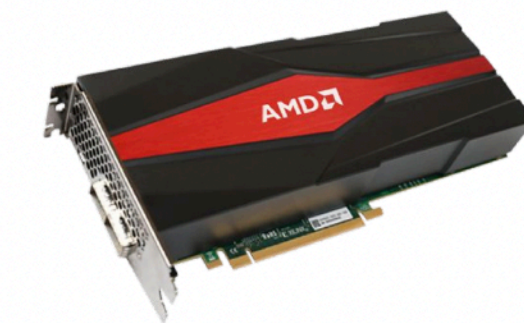
- Build new readout board based on newer technologies
- Versal ACAP, 400 GbE, AI engines, ...
- Explore modern network protocols (RoCEv2) for DAQ output
 - [Poster](#) by M. Migliorini

Explore use of accelerators for online processing

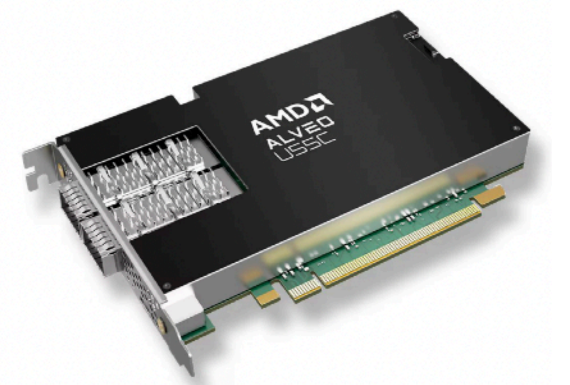
- GPUs, Alveo cards and Versal AI engines
 - [Thesis](#) by L. Sieder
 - [Thesis](#) by G. Zago
- $W \rightarrow 3\pi$ analysis taken as benchmark for accelerator studies
 - Unpack PUPPI objects
 - First level filtering and particles isolation
 - Combinatorics to identify the correct triplet
- Promising results from benchmarks on accelerators



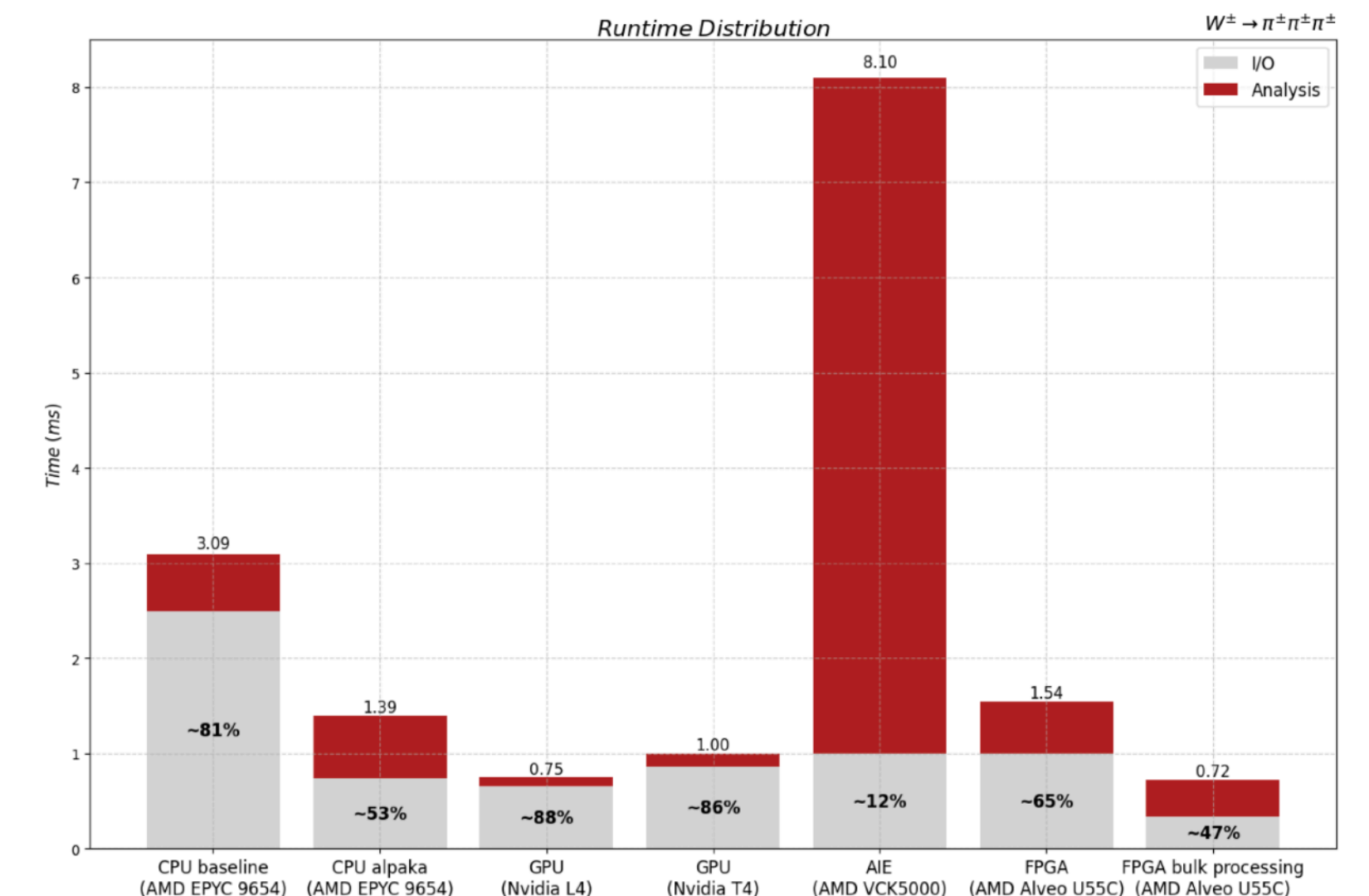
NVIDIA L4 GPU



AMD VCK5000



Alveo U55C



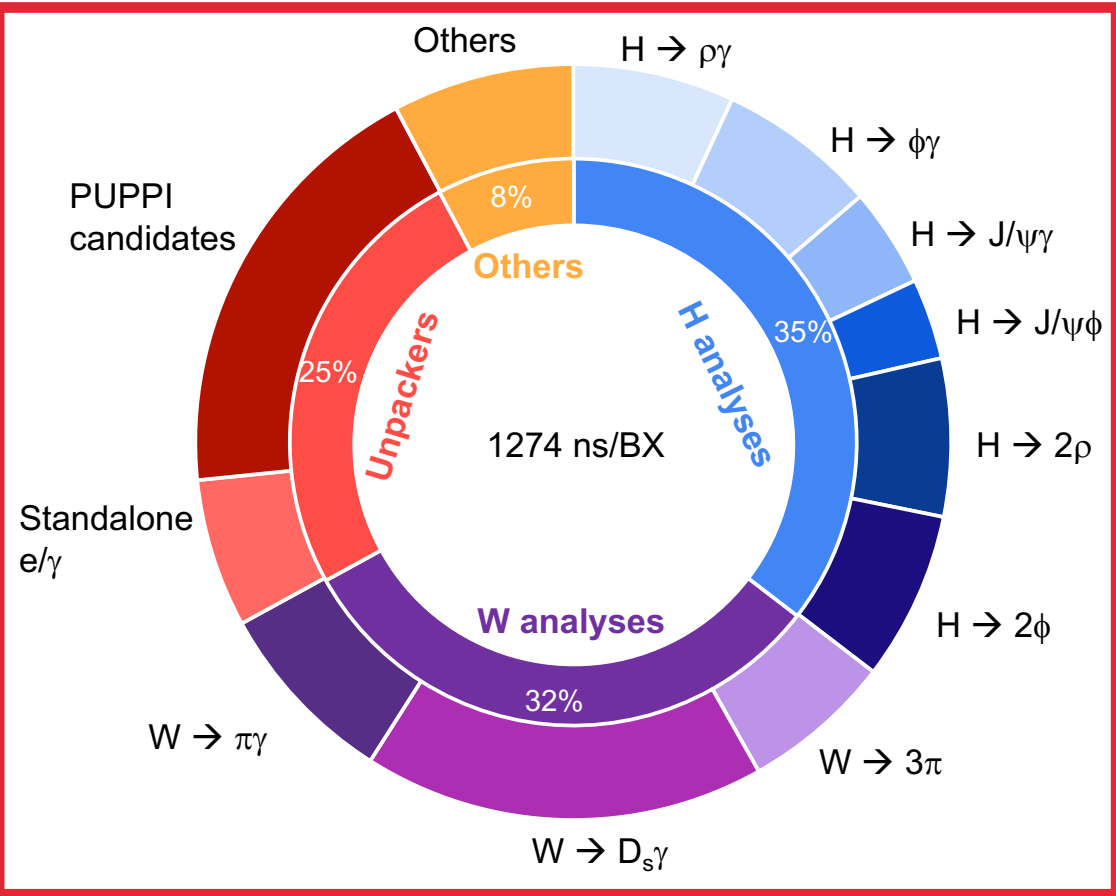
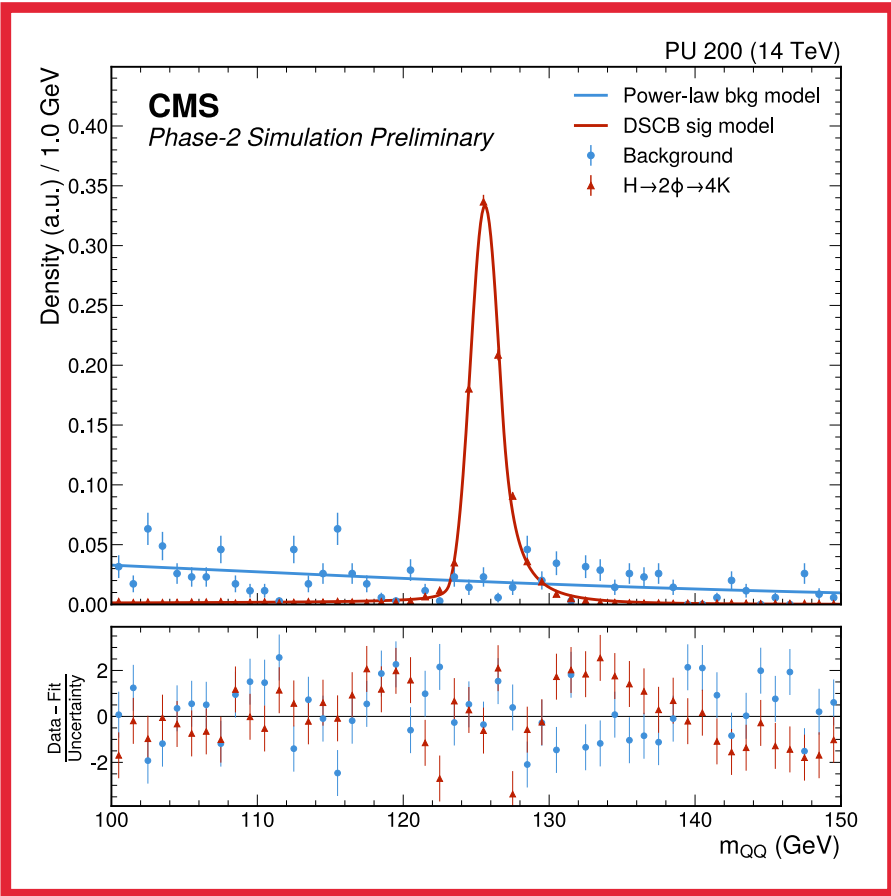
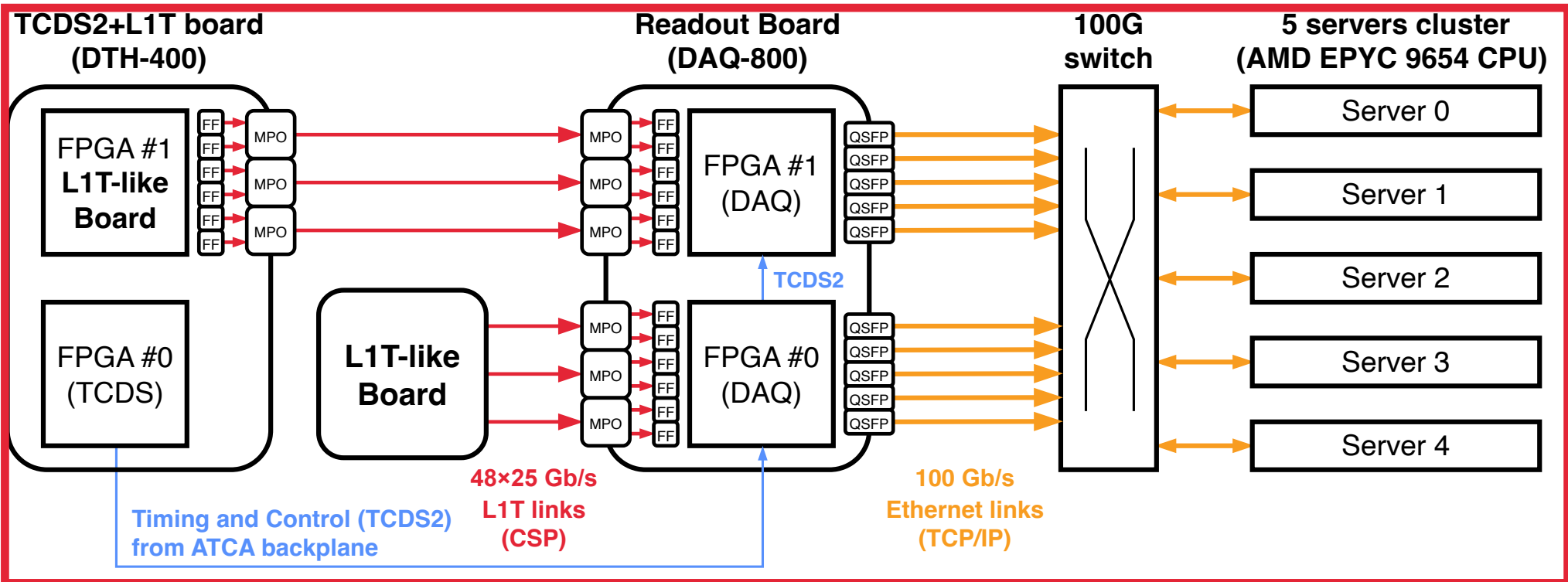
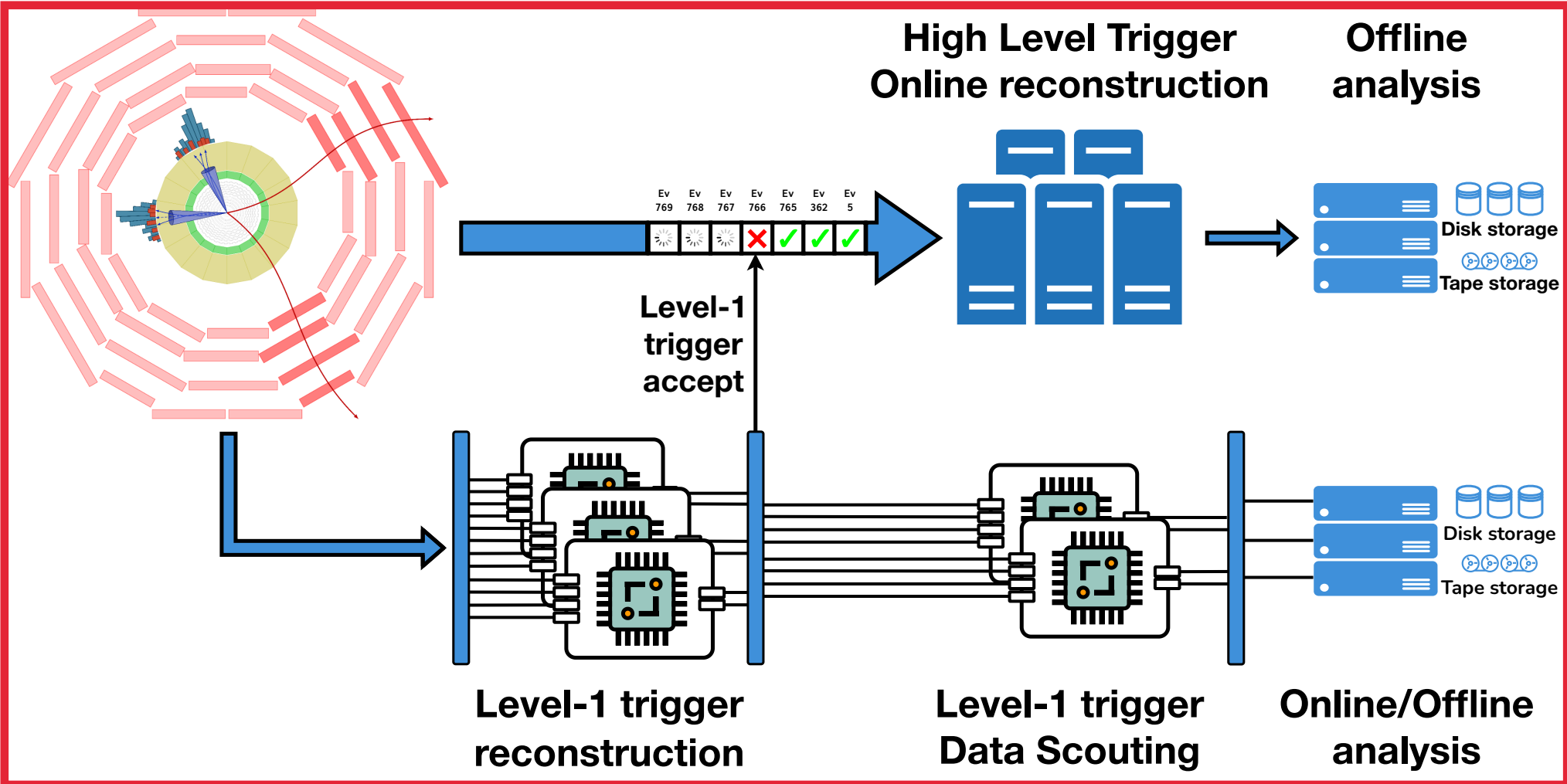
Conclusions

On-going L1 Scouting development for HL-LHC

- Baseline architecture demonstration is mature
 - **Run-3**: collecting real collision data and **producing physics results!**
 - **Phase-2**: capable of receiving 1/7th of the inputs of the full system
- Set of online analyses is expanding
 - Rare SM boson decays
 - Dark Matter searches with di-muon and di-electron final states
- Extend processing to PF candidates and L1 tracker tracks
 - Include more analyses: B physics ($B_S \rightarrow \tau\tau$), $X \rightarrow \tau\tau$

Going beyond the baseline

- [WP3.5](#) of Next-Generation Trigger initiative
- Accelerators to run analysis pipelines (GPUs, AI engines, ...)
- New network protocols (RoCEv2) and higher speed (400G)
- Newer generation of FPGA technology (Versal ACAP)

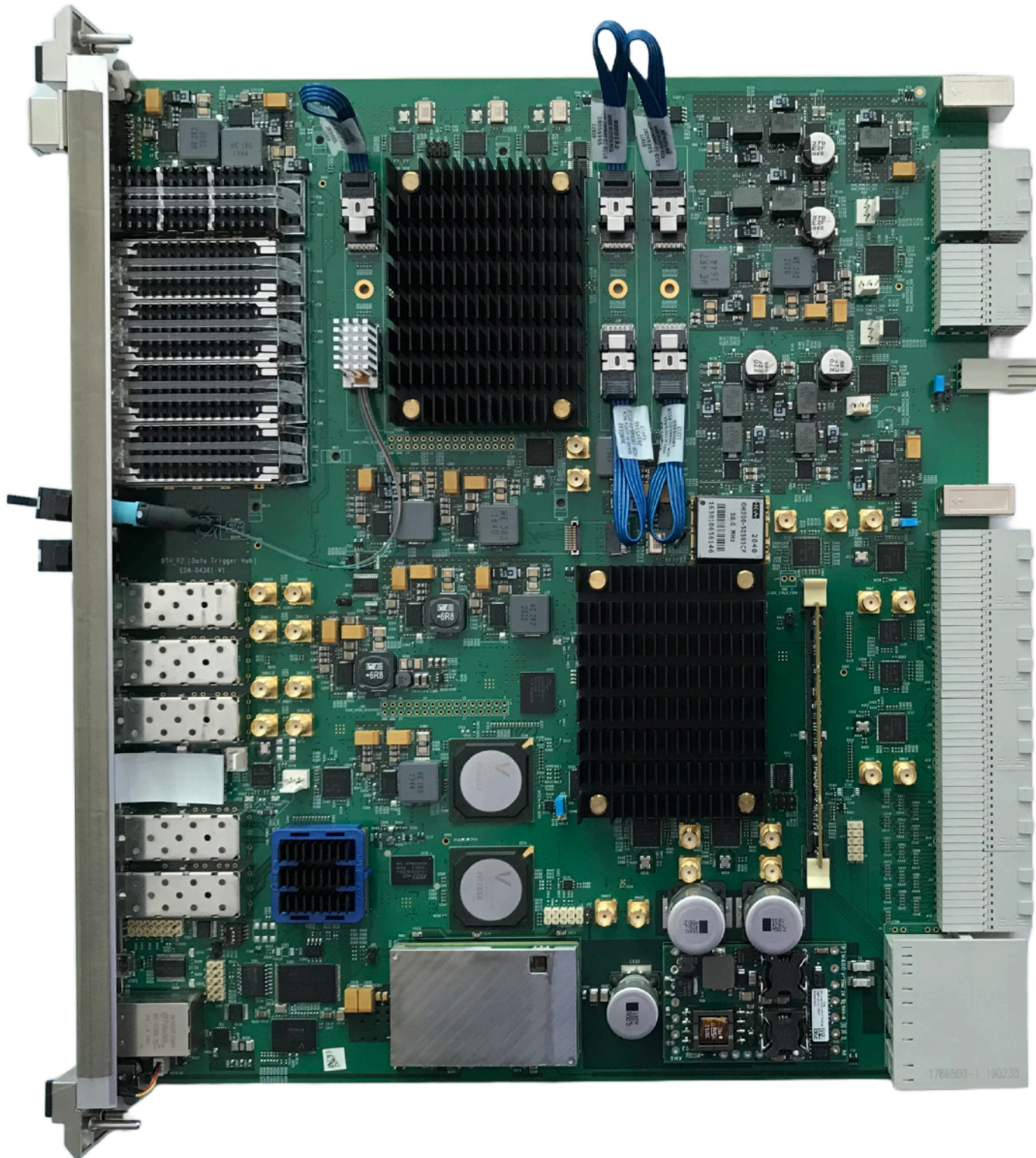


Backup

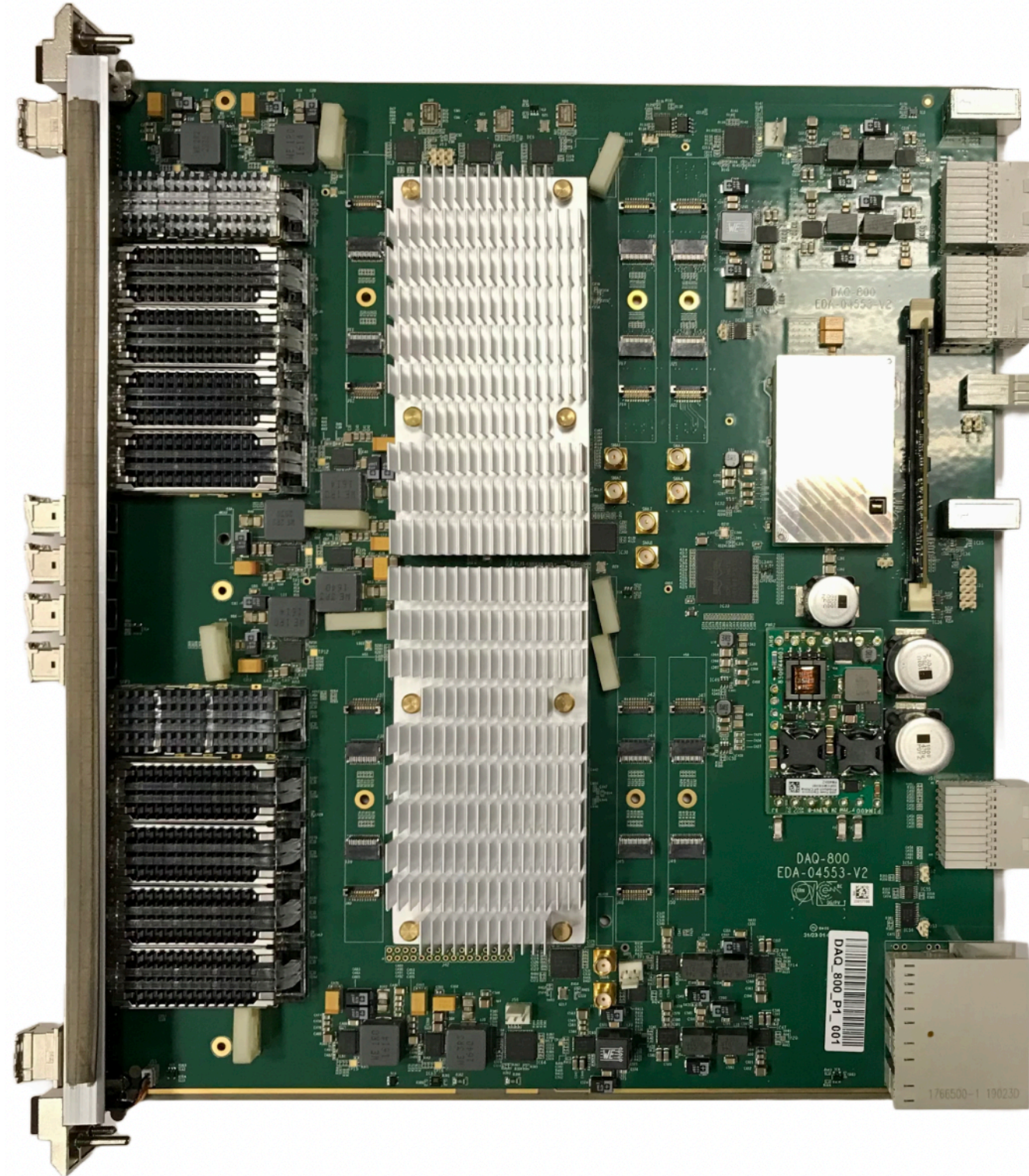
DAQ custom hardware for CMS Phase-2

2 (+RTM) custom ATCA boards developed to cover Phase-2 CMS DAQ and trigger control

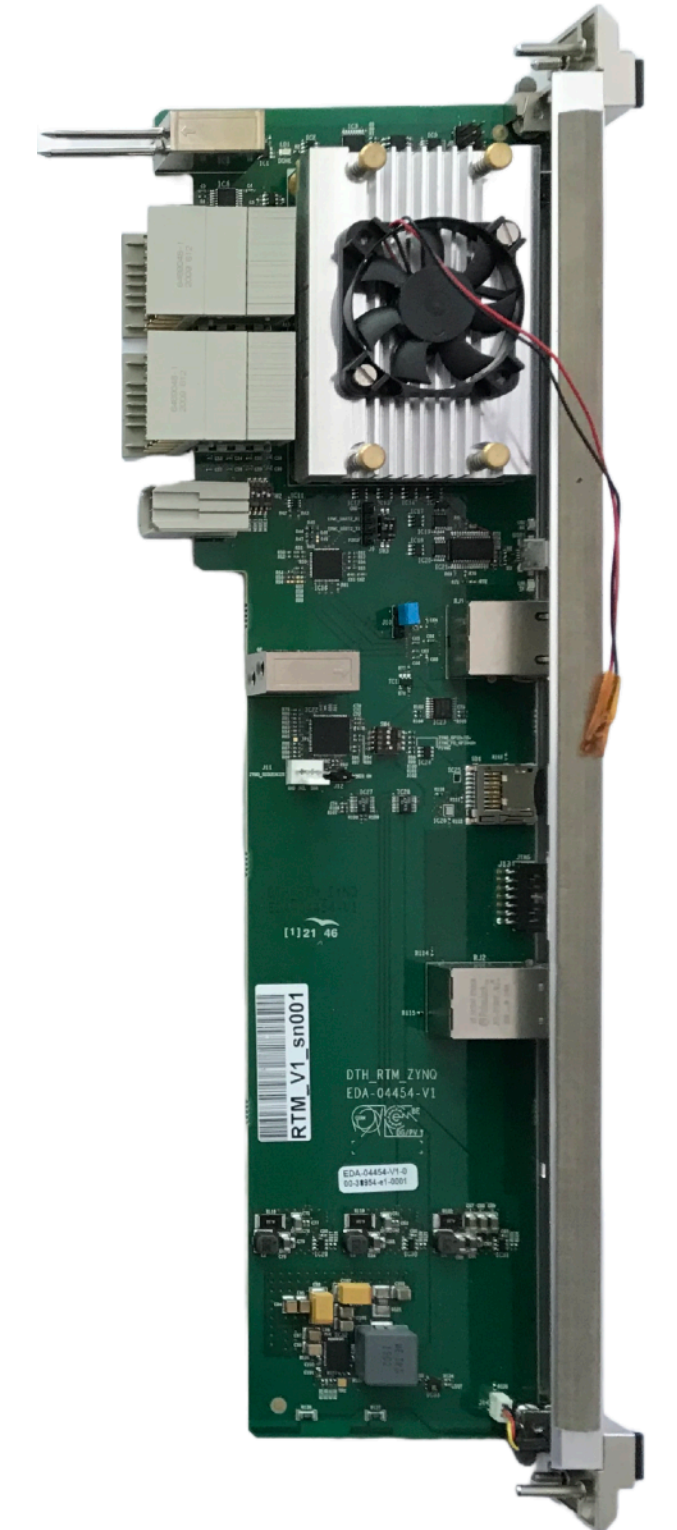
- DTH-400: interface subsystem back-ends to central timing, trigger and DAQ
- DAQ-800: throughput expander companion of the DTH-400



DTH-400 pre-series

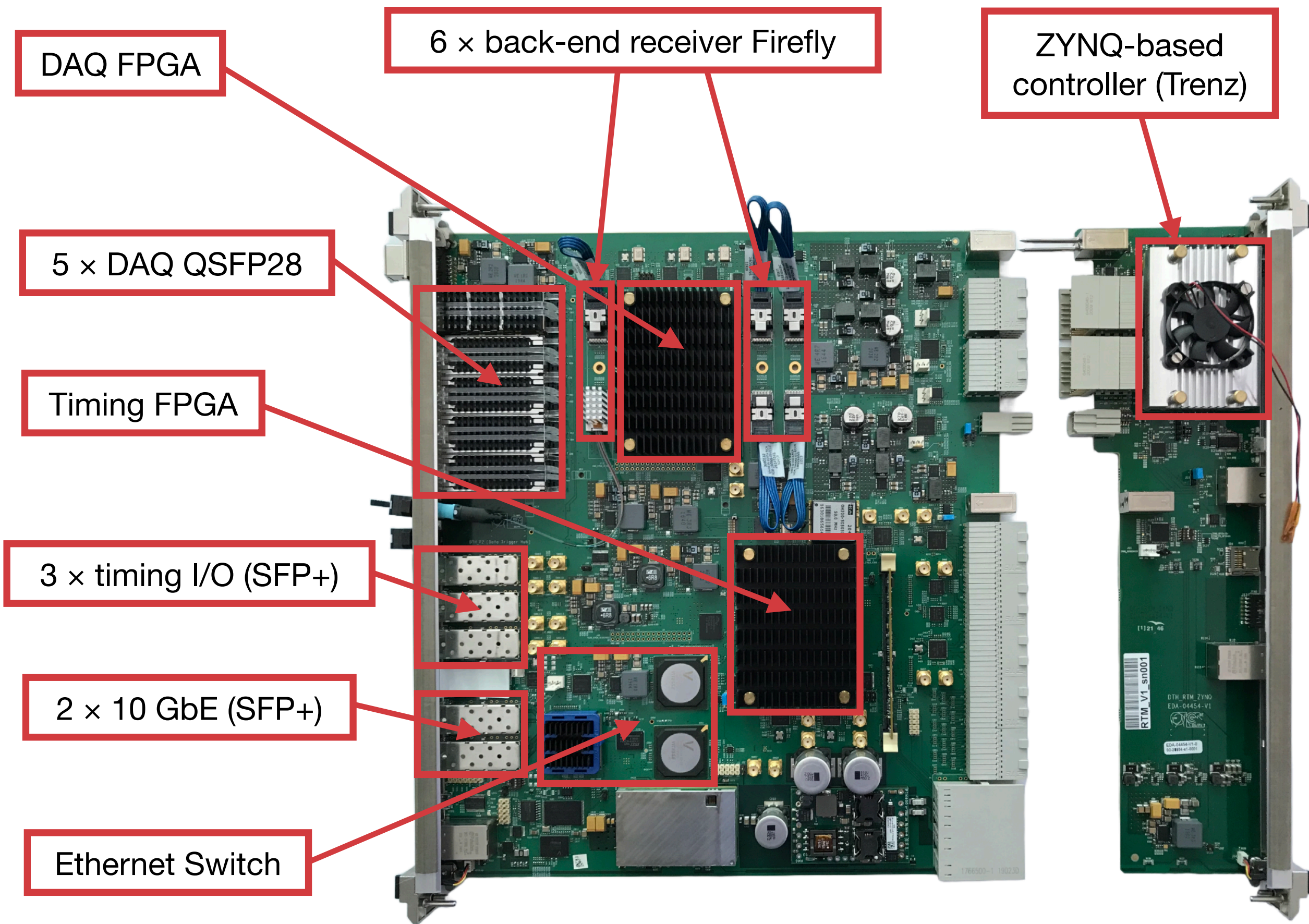


DAQ-800 prototype



RTM with Trenz module

The DTH400 (P2) board with Trenz RTM



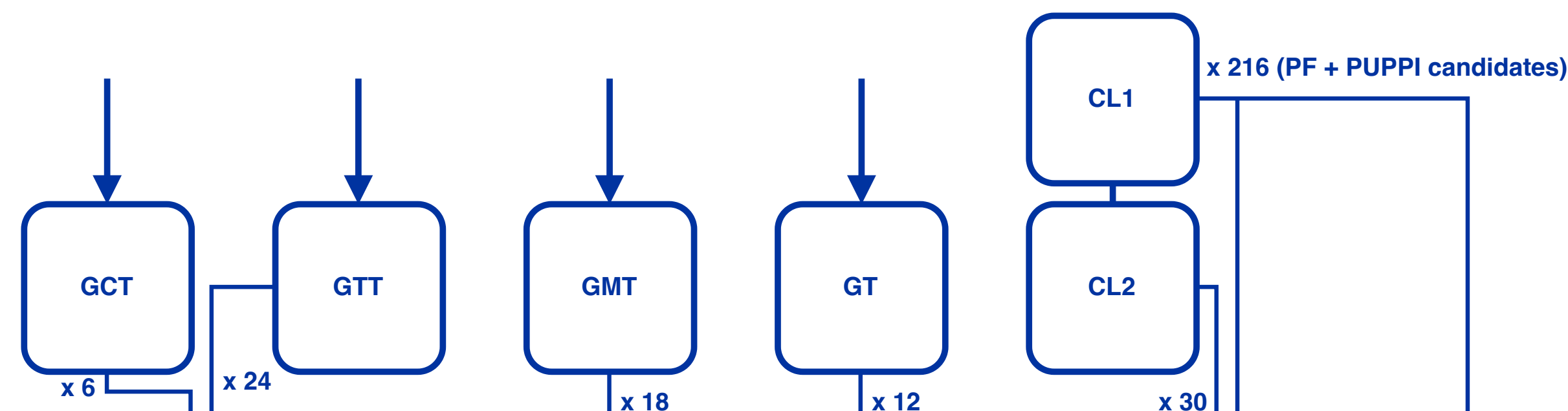
DTH-400 features

- HUB card (meant to be in central slot), but can also be used as leaf card
- 1 × DAQ FPGA (VU35P)
 - 6 × FireFly → 24 × 25 Gb/s optical inputs
 - 5 × 100 Gb/s outputs towards the DAQ
- 1 × TCDS FPGA (VU35P)
 - 5 SFP+ to TCDS FPGA
- CERN IPMC
- Shelf network switch (1 Gb/s to each leaf card)

Rear Transition Module (RTM)

- Trenz System-on-Module
- Control DAQ/TCDS FPGAs via AXI chip2chip

CMS Phase-2 L1 Scouting: data sources



Stageable system, baseline:

- Standalone muon and calorimetric objects
- Tracker objects from Global Track Trigger (GTT)
- PUPPI from Correlator Trigger Layer-2 (CL2)
- Global Trigger final decisions

Perform custom reconstruction, e.g.

- Jets, electrons, ... without latency limitations
- Full combinatorics with PUPPIs...

Extend with PF candidates / L1 tracks?

Source	Trigger input links (baseline)	Trigger input links (Upstream ZS)
GT	12	12
GTT	24	24 + 48 (Tracks ZS)
GCT	6	6
GMT	18	18
CL2	30	30 + 24 (PUPPI ZS)
CL1	216	84 (PF $ \eta \leq 3$ ZS)
Total	306	246

From L1 Scouting parallel talk at [ICHEP2024](#)

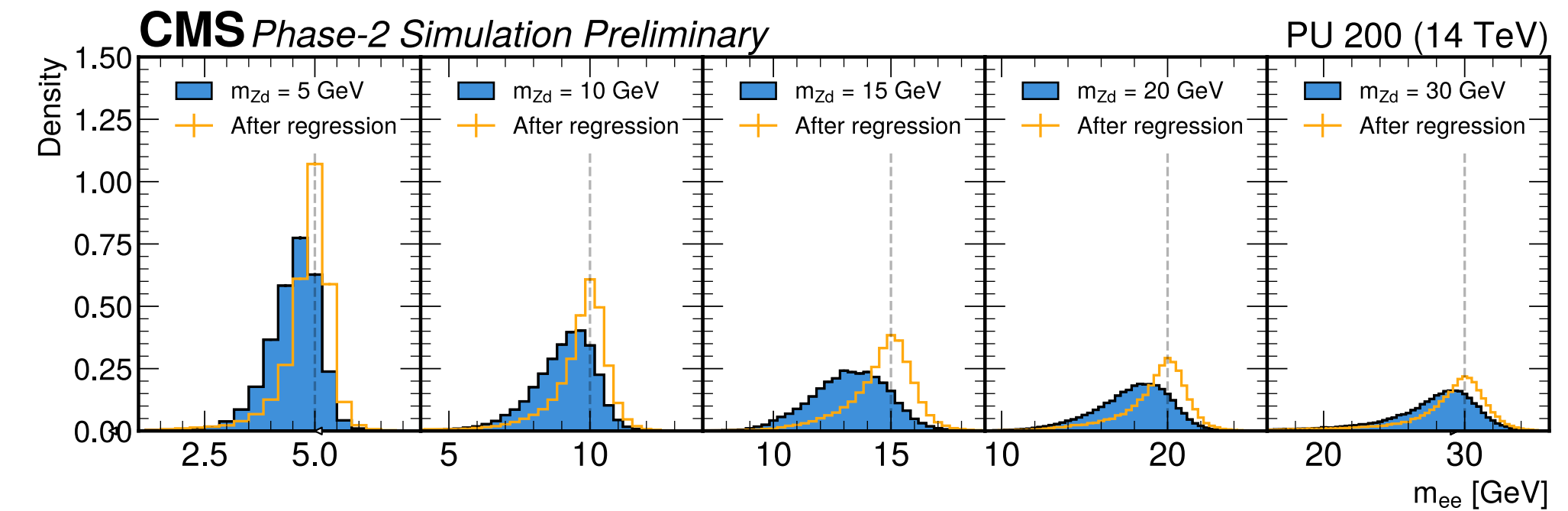
Baseline system in 1 ATCA crate

- 7 × DAQ-800
- 1 × DTH-400 to distribute timing and control signals (TCDS2)

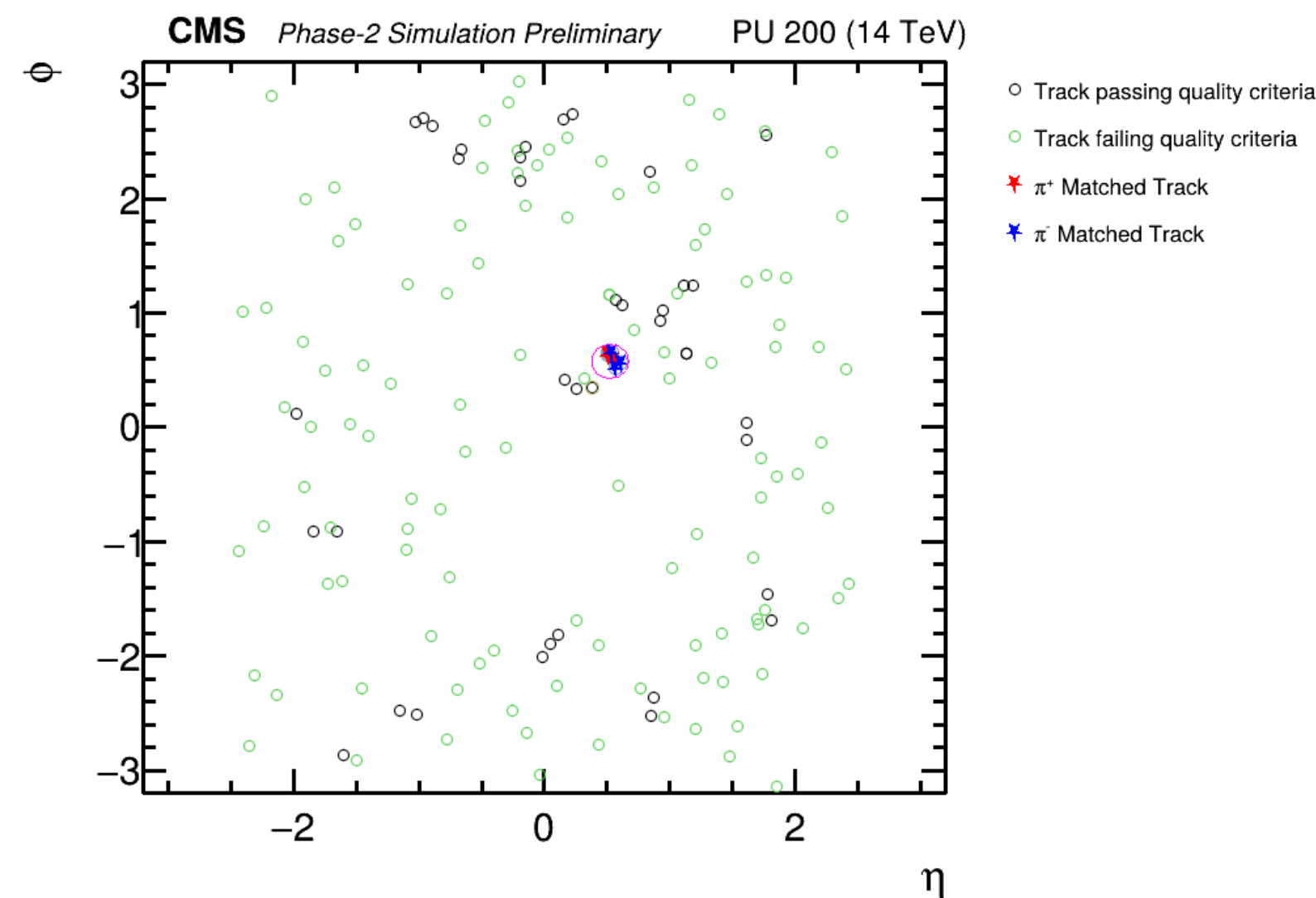
L1 Scouting Online processing extension

Extend demonstrator with more trigger inputs

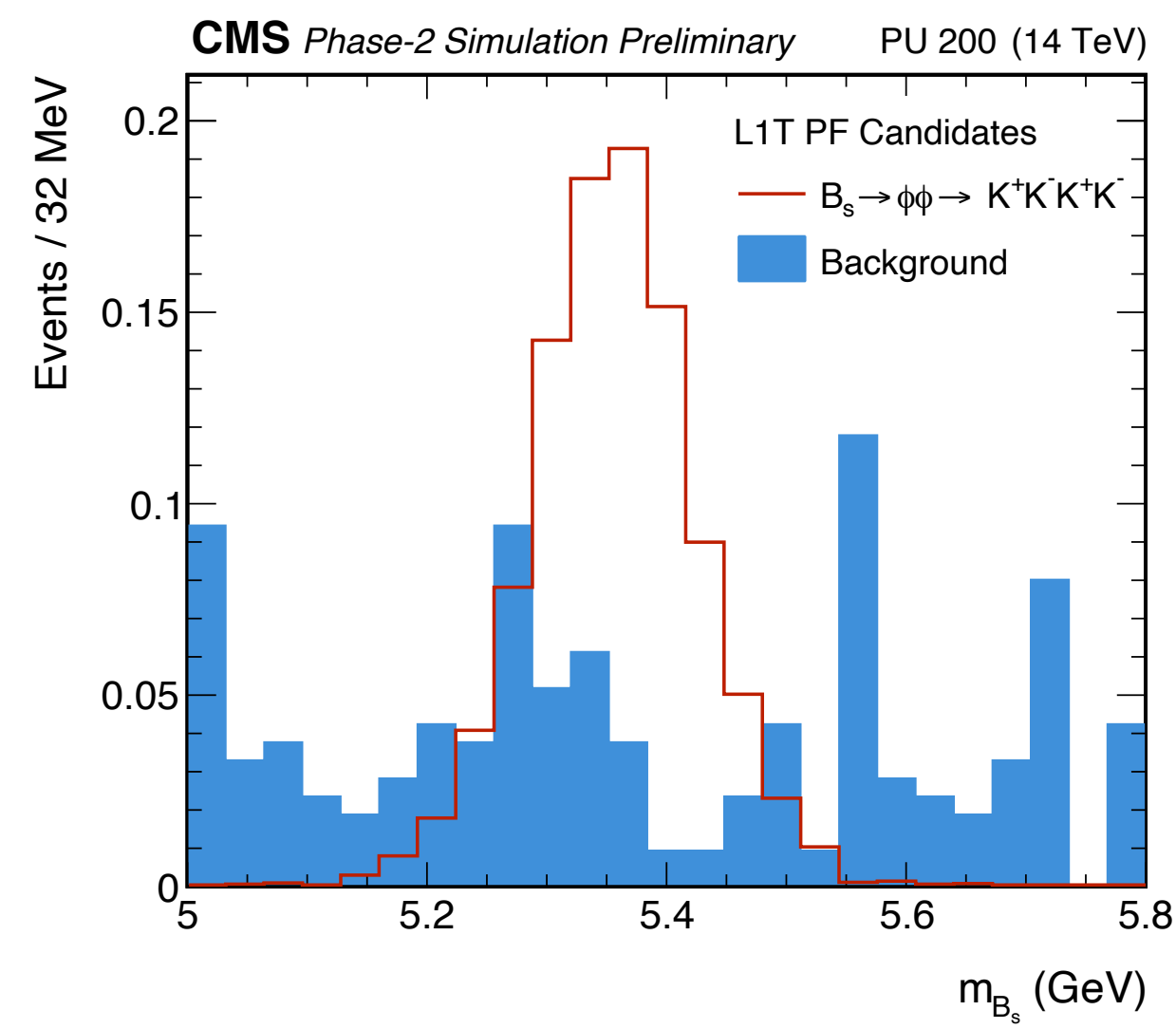
- Introduction of processing of PF candidates / L1 tracks?
- Recover efficiency for rare boson decay analyses
- Allow analyses like $B_s \rightarrow \tau\tau$, $B_s \rightarrow \phi\phi$, $H \rightarrow \tau\tau$, ...
- Clustering / ML applications? CLUE algorithm and p_T regression



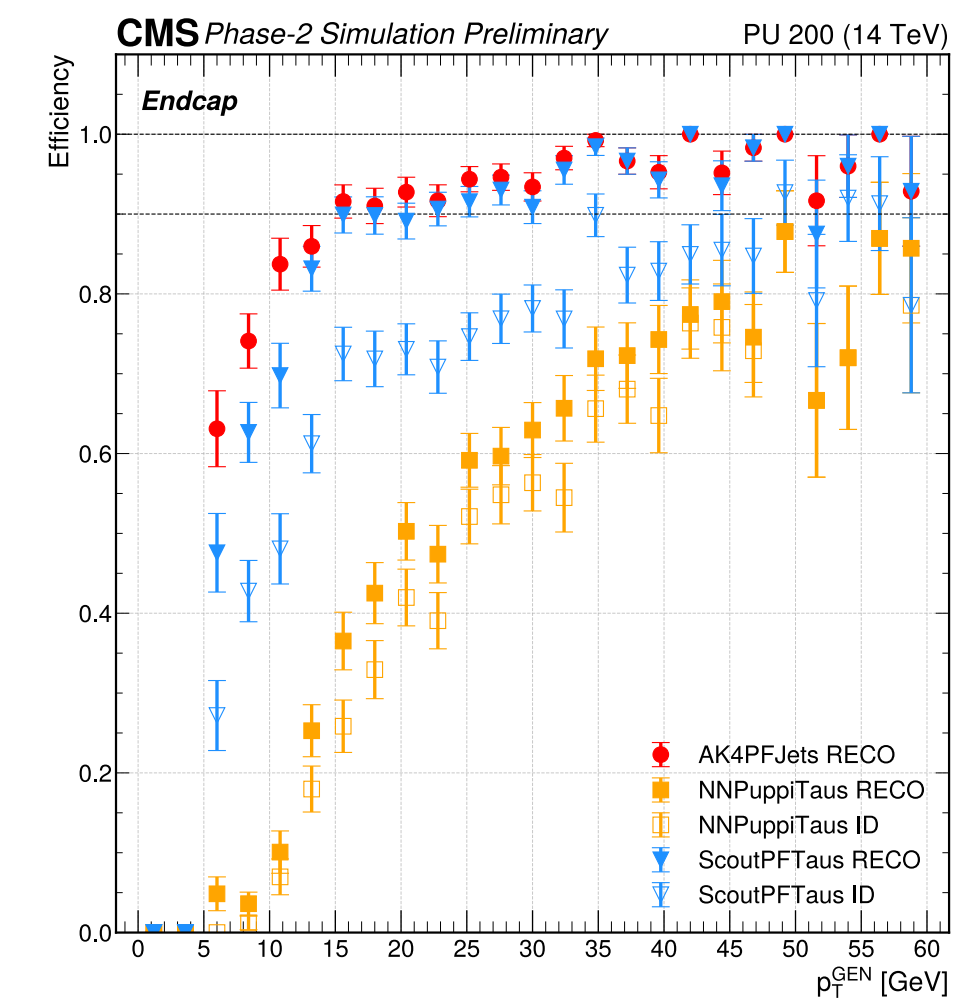
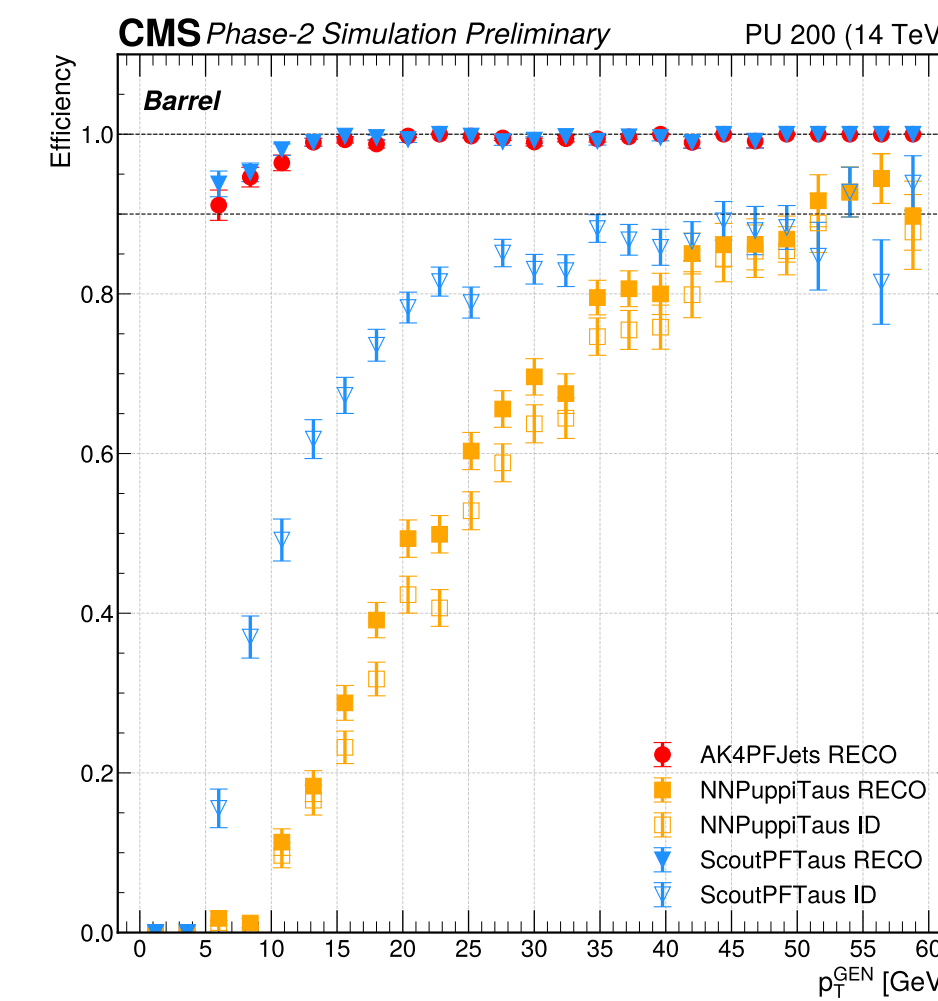
$Z_d \rightarrow ee$: m_{ee} invariant mass distribution after p_T regression ([CMS-DP-2025-058](#))



$B_s \rightarrow \tau\tau \rightarrow (3\pi)(3\pi)$ clustering studies ([CMS-DP-2024-096](#))



$B_s \rightarrow \phi\phi \rightarrow 4K$: m_{4K} invariant mass distribution ([CMS-DP-2025-058](#))



CLUE clustering: reconstruction of τ candidates for L1 scouting ([CMS-DP-2025-058](#))