NexTGen
Next Generation Triggers

L1 scouting for HL-LHC

Matteo Migliorini on behalf of NGT 3.5 team

Next Generation Triggers 1st Technical Workshop
CERN-November 25-27, 2024

Introduction

The WP3.5 team

Rocco Ardino, Cristina Botta, Cecile Sarah Caillol, Valentina Camagni, Gianluca Cerminara, Emilio Meschi, Lukasz Michalski, Matteo Migliorini, Giovanni Petrucciani, Leah-Louisa Sieder, Giovanni Zago

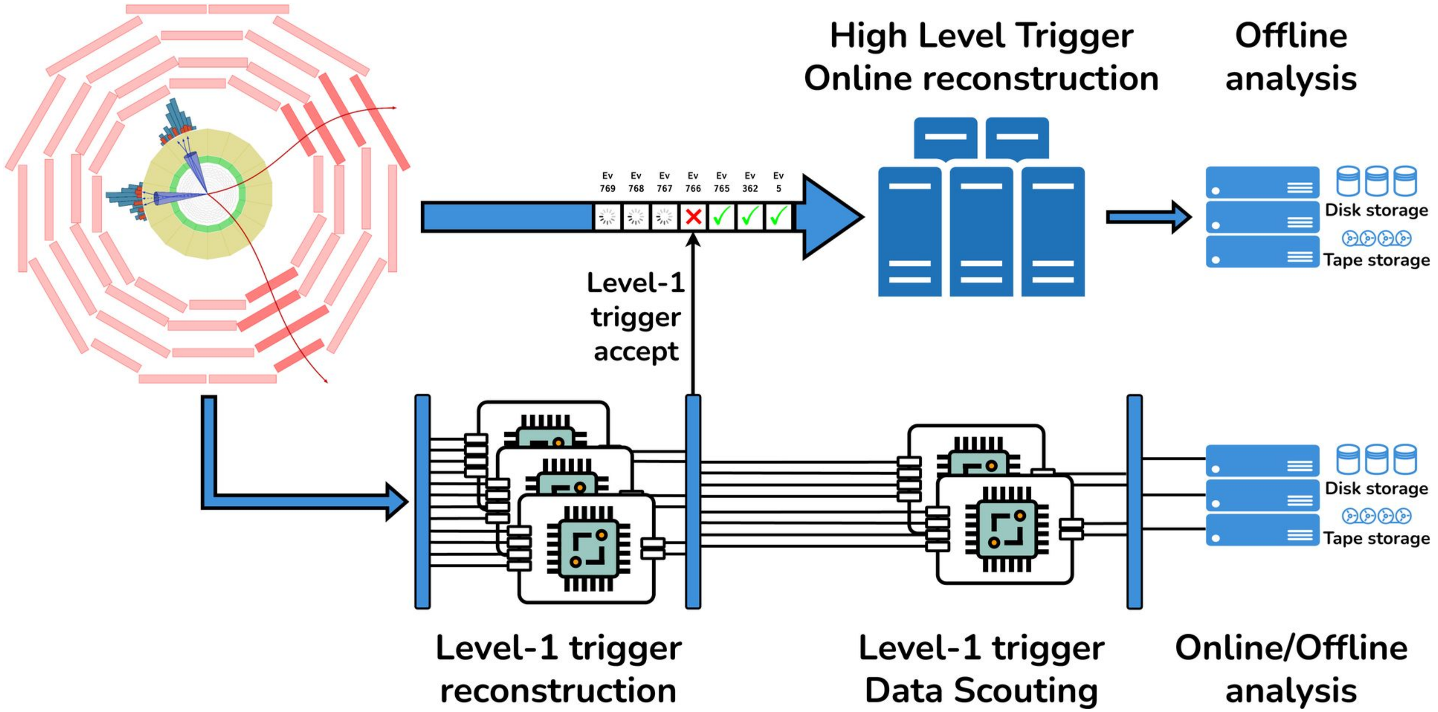
- ⇒ Joined with NextGen Triggers!

The Goal of NGT WP 3.5: Enhance the capabilities of the CMS L1T scouting system for HL-LHC

- Fully exploit the potentialities of the CMS Phase-2 upgrade to extend its physic reach
- Start from existing work carried out during Run-3

… what is L1T Scouting?

Data Scouting at Level-1 Trigger



Implemented L1T Scouting demonstrator during Run-3

- Limited to the current L1T reconstruction

L1T scouting in a nutshell: Capture objects reconstructed by the L1-Trigger at 40MHz and run physics analysis on them

Why? study signatures evading the standard trigger chain

- Large "irreducible" backgrounds
- Identification can't fit the L1 fixed latency

How? Dedicated FPGA-based boards collecting L1T objects via optical links

- Transfer to online processing farm

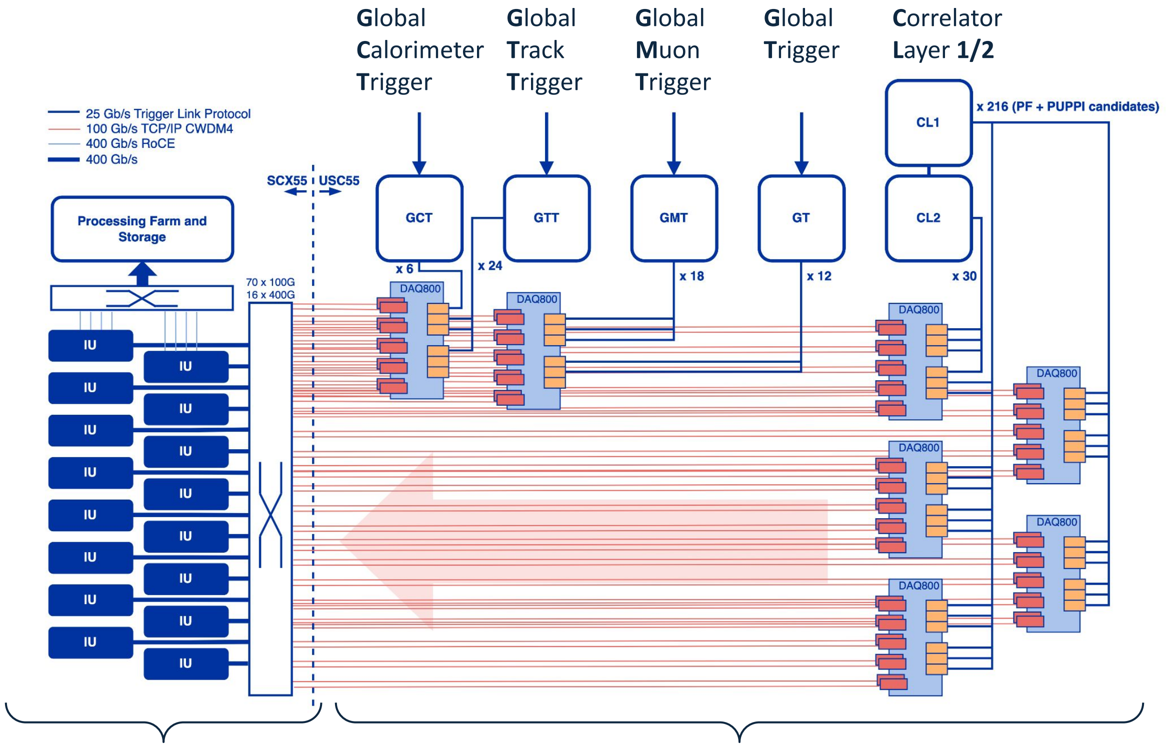
Scouting the CMS Phase-2 L1 trigger

No full detector readout as in HLT/Offline

- **Leverage improved L1T object reconstruction** quality
- Tracking, Particle Flow (PF), PileUp Per Particle Identification (PUPPI), …

Baseline system captures

- Standalone muon and calorimetric objects
- tracker objects from the GTT
- PUPPI objects from the CL2
- Final decision objects from the GT
- (Optional) PF candidates with and without PUPPI pileup subtraction from CL1



Ingestion Units aggregate and buffer data for processing

Capture trigger data and transmit to surface cluster using DAQ800 (CMS DAQ board)

Prototype physics analyses

Preliminary studies performed with L1T objects simulation

- Rare decays of W/Higgs bosons and Bs meson
- $W \to 3\pi$, $H \to \phi\phi \to 4K$, $B_s \to \tau\tau \to (3\pi)(3\pi)$, …
- First results described in a DP note and presented at CHEP2024

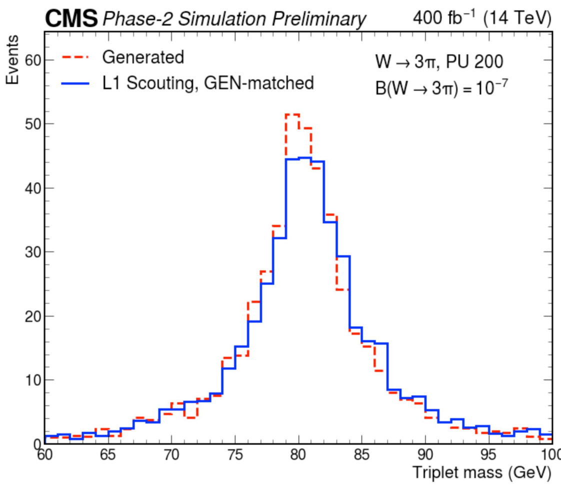
Signatures well suited for L1 scouting

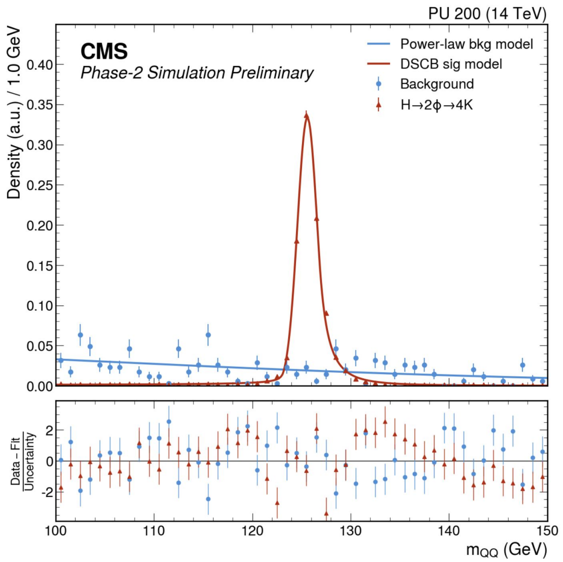
- Soft final state, not selected efficiently by the L1 Trigger
- Small signal yield
- Visible as narrow peak over smooth background

Working to extend the set of analyses

- Final states with multiple soft objects (Example), resonances in jj, ee, μμ, ττ

NGT⇒ accelerate baseline physics studies and explore potential for going beyond with more advanced algorithms and more L1T input





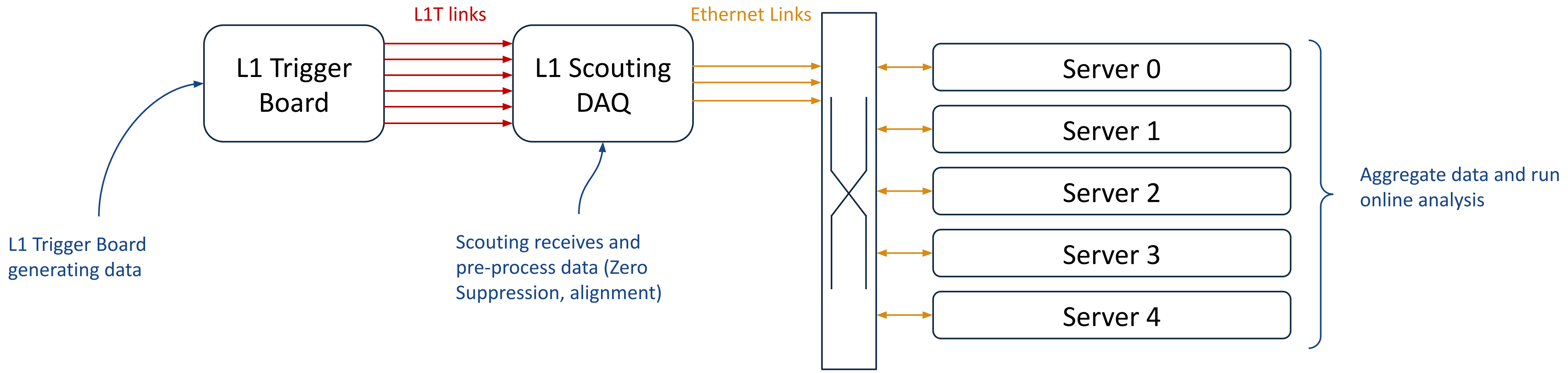
Demonstrator system

Develop data acquisition and real-time processing components for the **baseline** scouting system

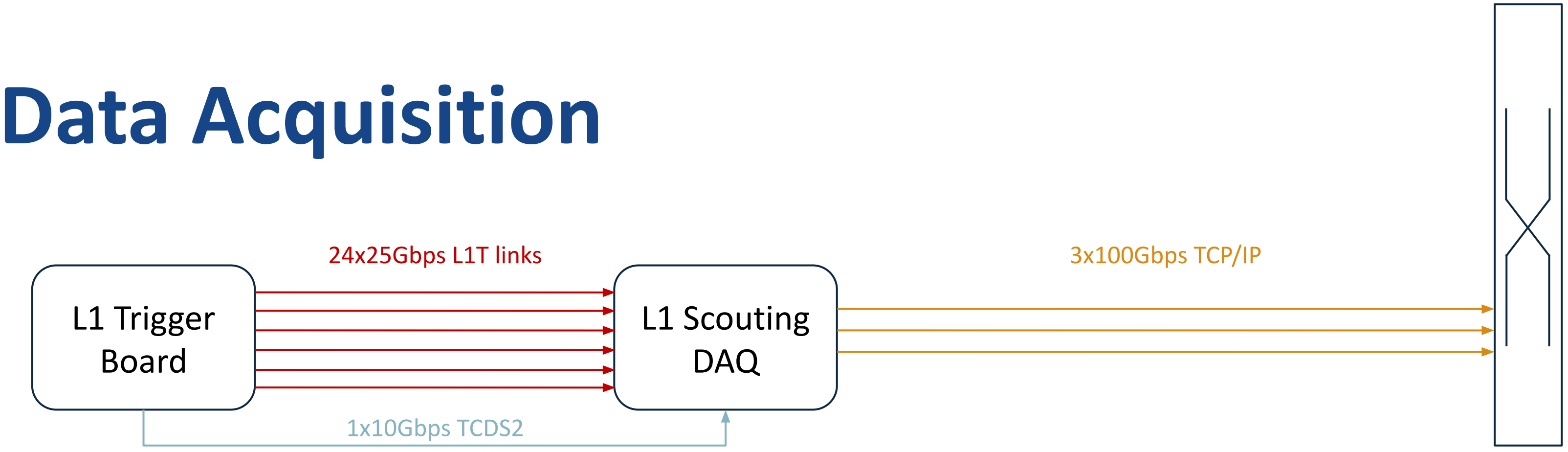
Set up a baseline demonstrator, and then go beyond it:

- Test alternative approaches for processing on baseline hardware
- Test beyond-baseline hardware for data acquisition

NGT!

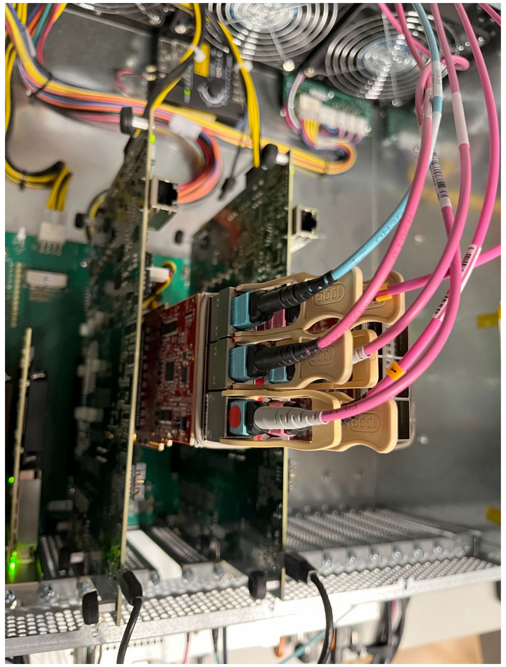


Demonstrator: Data Acquisition



DTH prototype ATCA board

- **D**AQ and **T**iming **H**ub
- Clock & TCDS2 source
- L1T-like data source
- Generate 6 data streams to simulate L1 PUPPI and e/γ objects
  - 24 links @ 25Gbps with L1 Protocol

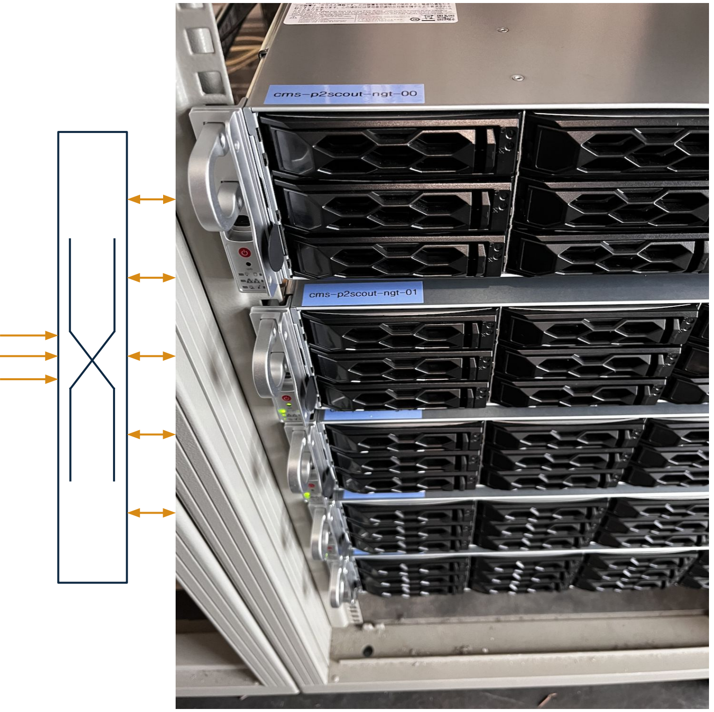


Xilinx VCU128 devkit

- 1 VU37P FPGA (~half DAQ800)
- 8G HBM for buffering
- Receive + Zero Suppression
- Transmission via 100Gbps TCP/IP (x3 links)

100Gbps network switch connecting VCU128 and processing servers

Demonstrator: Software online processing



5 servers with AMD EPYC 9654, 768 GB of RAM and 100G NIC

Receivers writing data on ramdisk shared over NFS

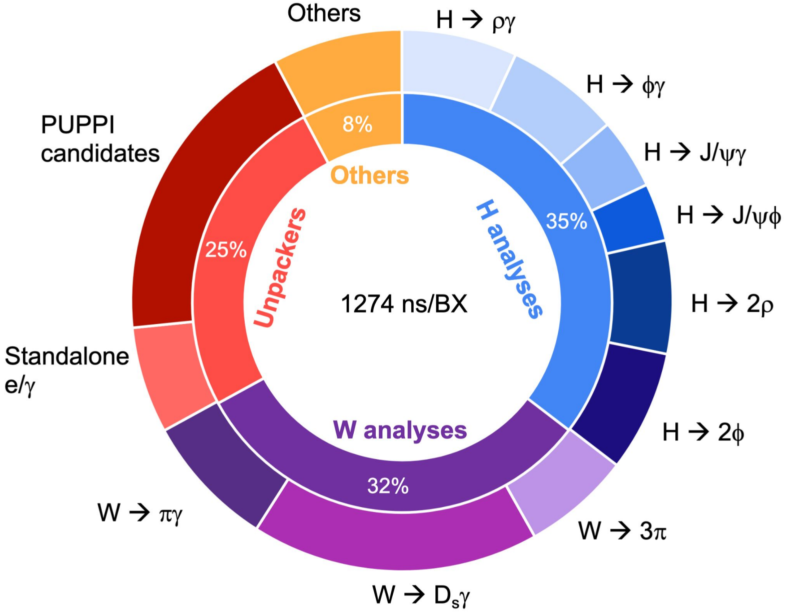
- Aggregation of the stream and processing implemented in CMSSW
- Standard software framework in CMS

Implemented unpackers and 9 analyses for rare Higgs and W decays

- Processing batches of events (1 orbit), on average ~1.3us per BX (CPU time)
- Preliminary implementation, room for improvement!

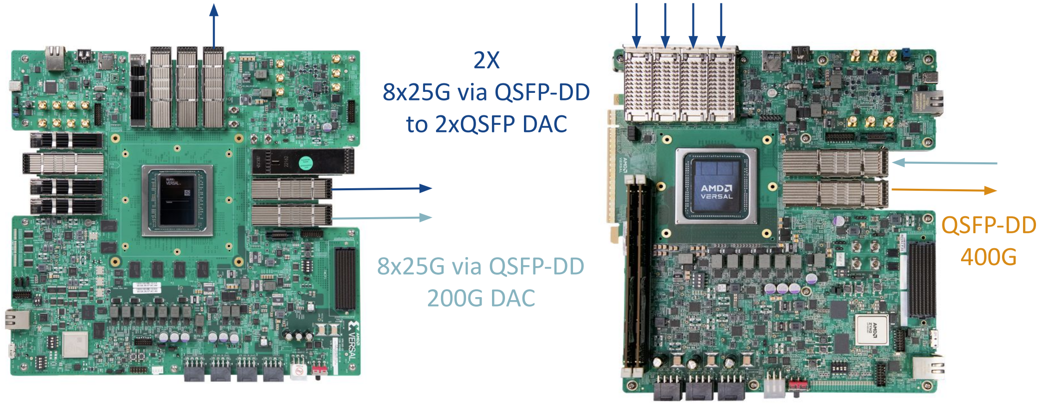
⇒Successfully demonstrated receiving and processing of 4 stream of events (~27MHz) of L1 PUPPI and e/γ

- Work in progress to extend it to 6



Extending the baseline: DAQ

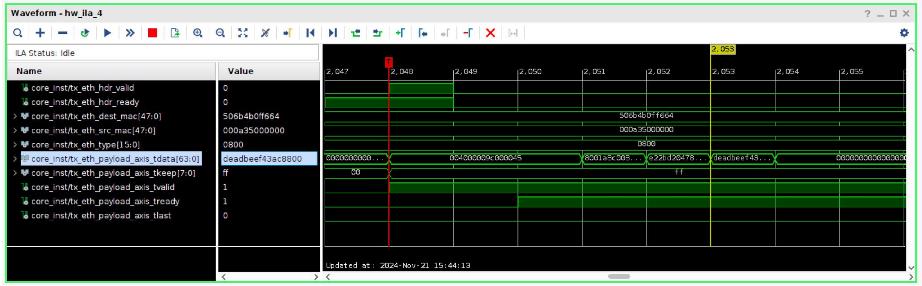
Building a new DAQ system/board based on newer technologies: AMD Versal, 400Gbps networking, …



Explore new protocol for the transmission, e.g. RoCEv2

- Working 100G implementation on VCU118
- Started porting to VHK158, to be extended to 400G?

Status: working UDP layer @10/25G



Generate the 24 trigger links from the VPK180 and receive on the VHK158

Extending the baseline: Accelerators

Baseline system processing on CPU

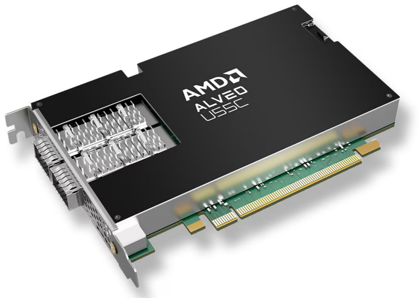
- Investigate acceleration of physics analysis and data processing
- Explore AMD AI Engines, FPGA accelerator cards and GPU

Implement benchmark physics analysis, e.g. W⟶3π

- Unpack PUPPI objects from raw data
- First level filtering, compute particles isolation, …
- Combinatorics to identify the correct triplet

⇒ Run full pipeline on the accelerator

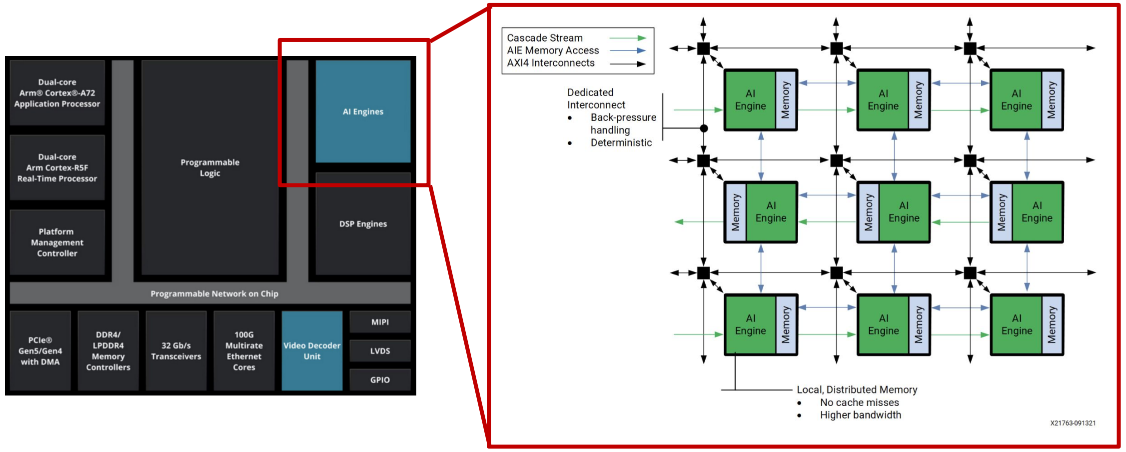
AMD VCK5000



Alveo U55C

NVIDIA L4 GPU

AI engines



VCK5000 contains Versal VC1902 ACAP

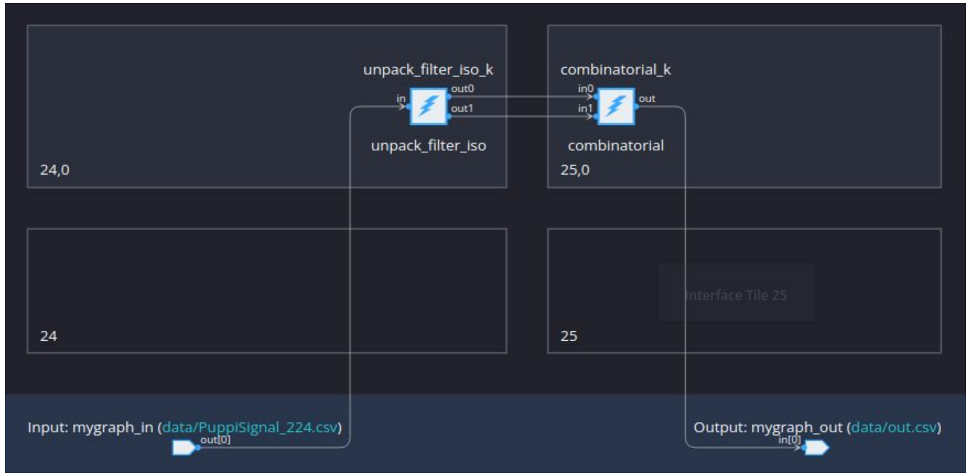
- Matrix of 400 AI Engine tiles @1.25 GHz
- multi-core capability, vector APIs to fully exploit the engine capabilities

Implemented the analysis in two kernels

- Mixed Scalar and Vectorized kernel in charge of unpacking, filtering and computing objects isolation
- Combinatorial kernel running on scalar processor
  - Acting on list of selected particles, much faster than vectorized implementation

AIE simulation

- Verify correctness of the algorithm
- Average latency of 7.5µs



Input: 224 PUPPI candidates

Output: selected triplet, invariant mass

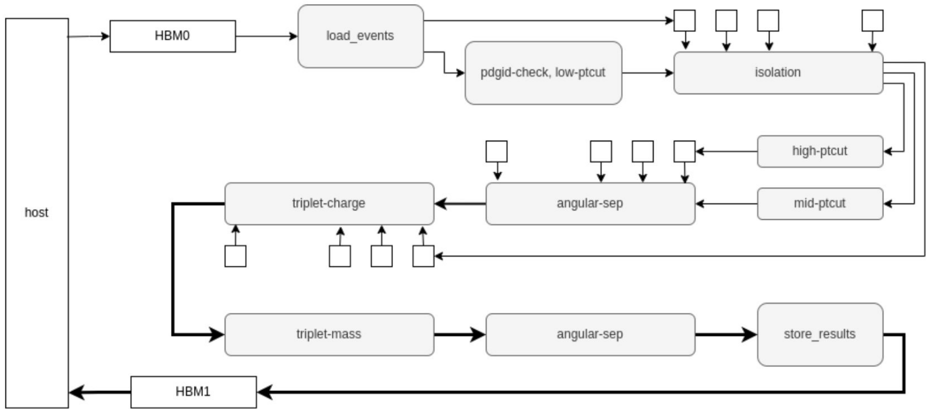
High Performance Compute Card

Implemented on AMD Alveo U55C (Virtex XCU55C UltraScale+ FPGA, 16G HBM)

- Allocated memory on the device and copy data using OpenCL

From the host data copied to the HBM

- Run sequence of modules progressively filtering the data
- Finally return to the host triplets passing the filters



Kernel execution time: ~11µs / event

Contained resource usage:

NAME	KERNEL	LUT (%)	REGISTER (%)	BRAM (%)	URAM (%)	DSP (%)
analysis_1	analysis	6.8 %	4.32 %	0.4 %	N/A	8.47 %

GPU

End-to-end pipeline implemented on GPU

- Adopted Alpaka, used in the CMSSW framework
- Portability across different back-ends (CPU, CUDA, HIP, ...)

Optimized memory layout to achieve best performance

- Converted Array Of Structures into Structure Of Arrays
  - memory alignment / cache hinting managed by CMSSW

Working on the implementation of the analysis modules

- 10x speedup for properly written modules (e.g. particle isolation)
- Combinatorial part currently slower than CPU (focused on correctness)
  - Original algorithm optimized for single-thread execution
  - Lot of if/else → branch divergence
  - different computing scheme needed for heterogeneous programming

alpaka

```
using BxArray = edm::StdArray<uint16_t, constants::BX_ARRAY_SIZE>;
using OffsetsArray = edm::StdArray<uint32_t, constants::OFFSETS_ARRAY_SIZE>;

GENERATE_SOA_LAYOUT(PuppiSoALayout,
  SOA_SCALAR(BxArray, bx),
  SOA_SCALAR(OffsetsArray, offsets),
  SOA_COLUMN(float, pt),
  SOA_COLUMN(float, eta),
  SOA_COLUMN(float, phi),
  SOA_COLUMN(float, z0),
  SOA_COLUMN(float, dxy),
  SOA_COLUMN(float, puppiw),
  SOA_COLUMN(int16_t, pdgId),
  SOA_COLUMN(uint8_t, quality)
)
```

Portable Collection for PUPPI candidates in CMSSW

Conclusions and Outlook

Performed **first set of prototype physics analyses**, proving the potential of L1 scouting

- Ongoing work to include additional examples

Implemented demonstrator of the “baseline” scouting system

- From the collection of L1 trigger objects to running online analyses

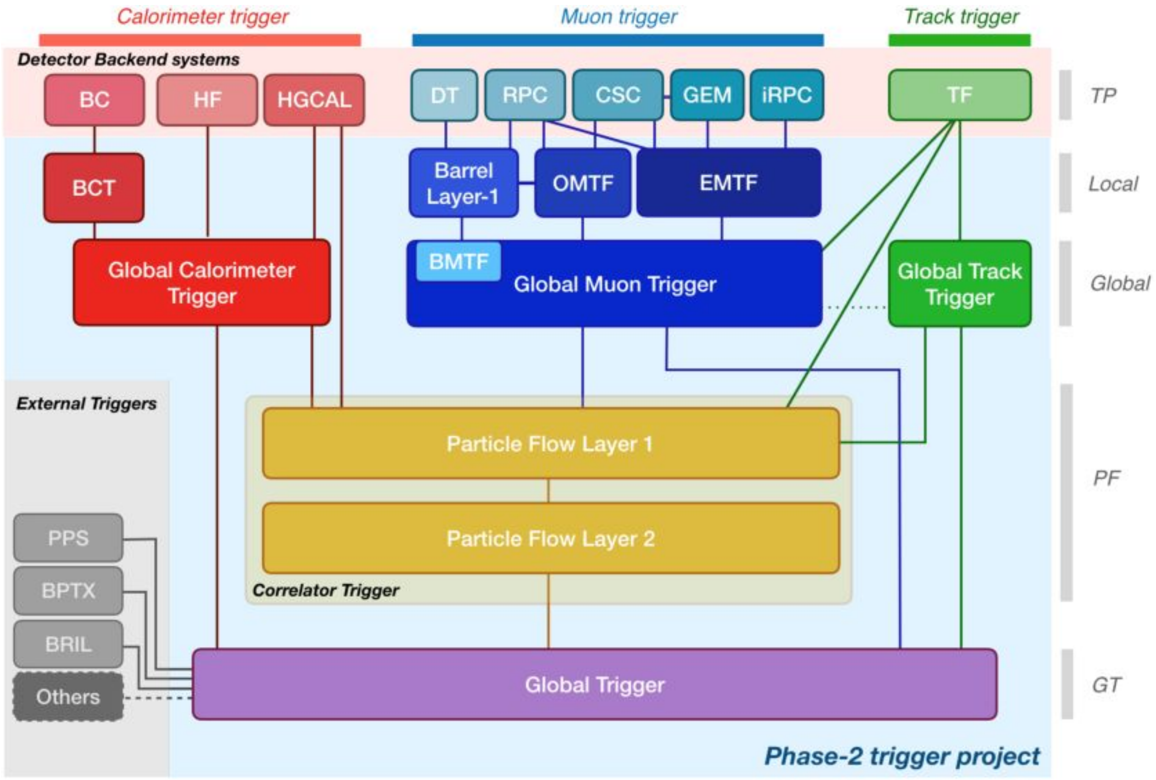
Leah’s poster at CHEP2024 and DP Note

Now **exploring the possibility of extending the system**

- Versal devkits capable of handling higher link count and speed, new data transmission protocols
    - Base for a new Versal-based DAQ board with higher bandwidth and link speed
- Working on accelerators for analysis / running complex algorithms
    - Familiarizing with the devices while showing promising results

Backup

CMS Phase-2 L1-Trigger



Selecting events which will be processed by HLT

- Hardware trigger working at 40MHz
  - 12.5us maximum latency
- Maximum rate of 750kHz

Complex algorithm running on FPGAs to reconstruct high level quantities

- Higher granularity with respect to Phase-1
- Tracking, Particle Flow, …