CMS Level-1 trigger Data Scouting for online trigger-less data processing

ICSC Spoke 2 Annual Meeting
19 December 2023

Sabrina Giorgetti
Ph.D. student @ UNIPD
sabrina.giorgetti@student.unipd.it

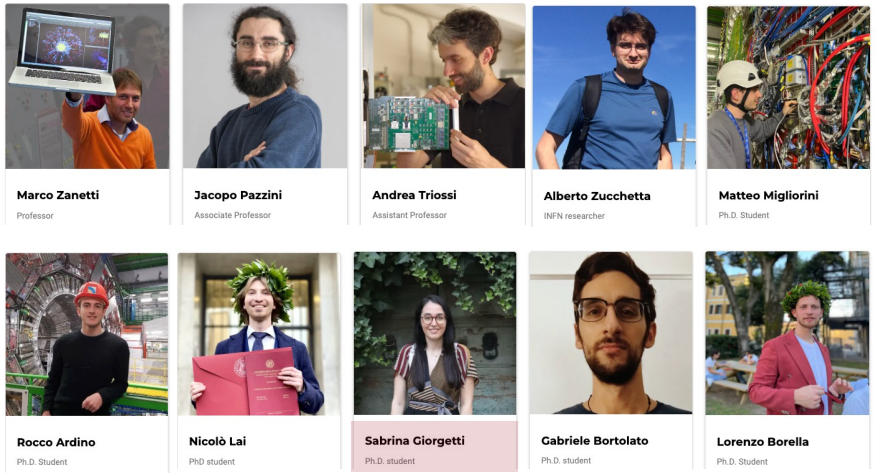
"This work is partially supported by ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union – NextGenerationEU"

Introduction

Research group

- Our team is mainly active in the CMS experiment at the Physics and Astronomy Department of the University of Padova
- Closely collaboration with CMS DAQ and Level-1 trigger groups

BOOSTLab

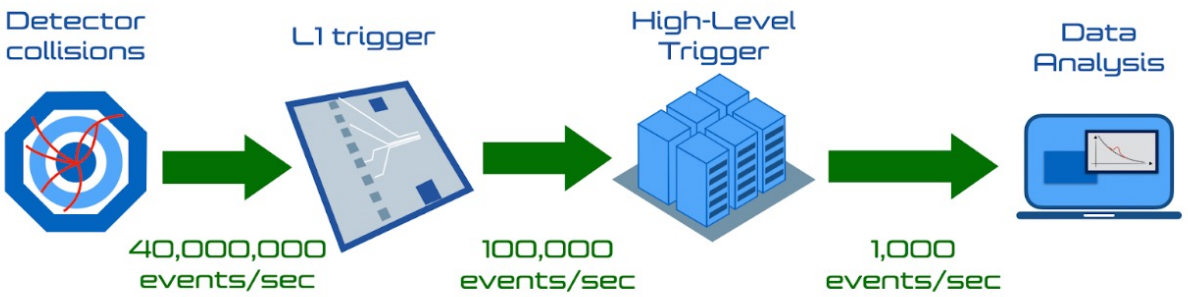


Involvement in ICSC Spoke2

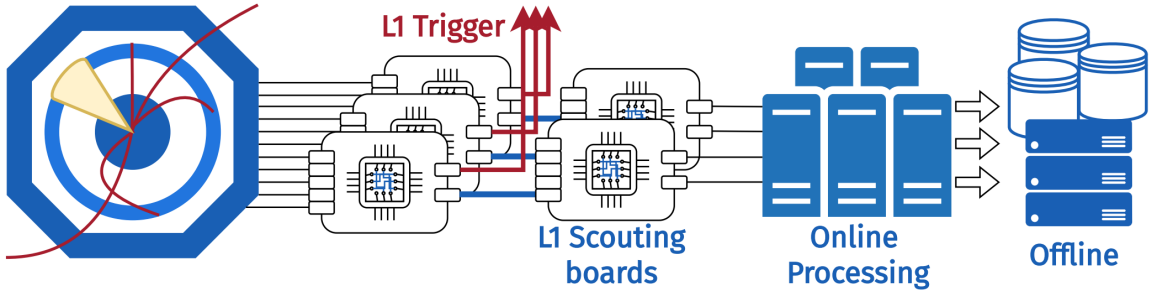
- Activities concerning FPGAs at various level within the **WP4**
- Training courses: Introductory course to HLS FPGA programming & Porting of GPUs of code an algorithms
- Flagship use case UC2.2.3: Ultra-fast algorithms running on FPGA (FAST_FPGA)
  - **Scouting and processing of Level-1 trigger data using FPGA to run on-the-fly momentum object calibration with ML-based algorithms**
  - Online reconstruction algorithms for data scouting
  - Development and testing of RDMA over converged ethernet (ROCE) on FPGA for data transfer from detectors' front-end to computing servers
  - Development of quantum-inspired Tree Tensor Networks for classification in Trigger on FPGA

Università degli Studi di Padova
INFN
Fondazione ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani Piano Nazionale di Ripresa e Resilienza

40MHz Scouting for online trigger-less data processing



CMS data flow



Level-1 Scouting data flow

The CMS data flow

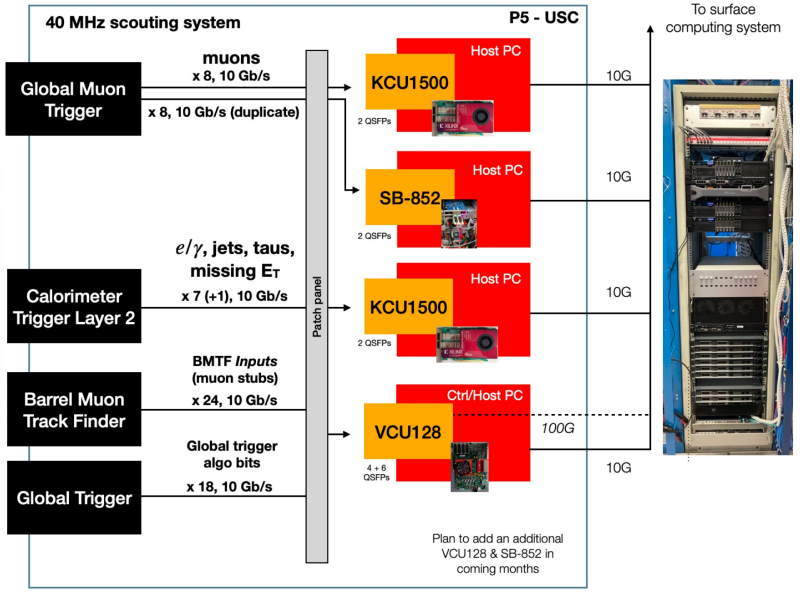
- Massive data flow (TB/s) from pp collisions at the LHC
- CMS employs a two-level trigger system to select events of physics interest to store for analysis
- Events rejected at the trigger levels are lost: what if new physics was there?
  - ↠ Reduce the bias of new physics searches with a triggerless infrastructure

Phase-2 Level-1 Data Scouting (L1DS)

- Scouting: acquisition of reduced-event-content data at considerably higher rates than the standard accept rate
- The L1DS is a novel data acquisition system under development for CMS Phase-2 at the HL-LHC
- L1DS read out the L1 trigger objects at the full bunch crossing (BX) rate of 40MHz with the aim of performing online physics analysis
- Potential ranges from improving trigger diagnostic to enabling the study of otherwise inaccessible signatures

Università degli Studi di Padova
INFN
Fondazione ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani Piano Nazionale di Ripresa e Resilienza

40MHz Scouting for online trigger-less data processing



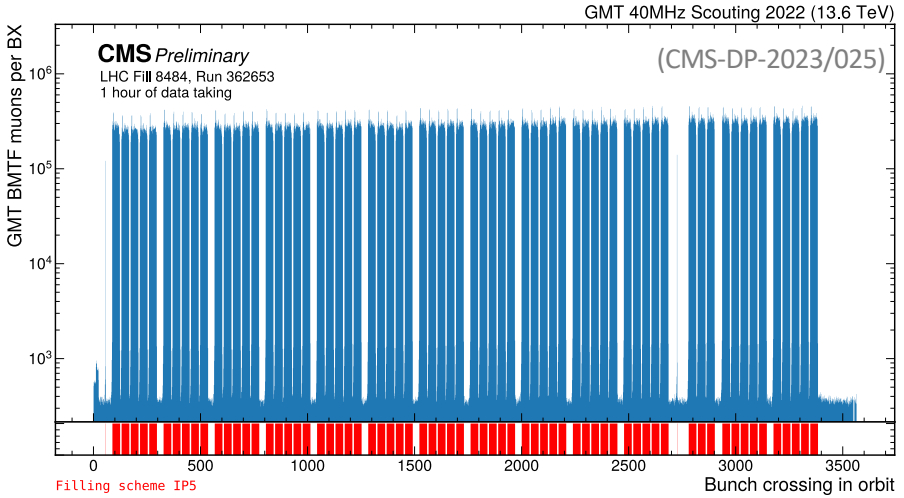
40MHz Scouting demonstrator for LHC Run-3 (2022-2025)

- Demonstrator tested as proof of concept with three types of FPGA-based receiving boards
- Acquisition of L1 trigger primitives at the full BX rate from multiple sources:
  - Global Muon Trigger (GMT): **up to 8 GMT muons**
  - Calo Layer 2 (DeMux): **Up to 12 e/γ, jets, τ and missing E_T**
  - Barrel Muon Track Finder (BMTF): **BMTF inputs**
  - Global trigger (GT): **GT algorithm bits**

Studies for the 40MHz Scouting demonstrator

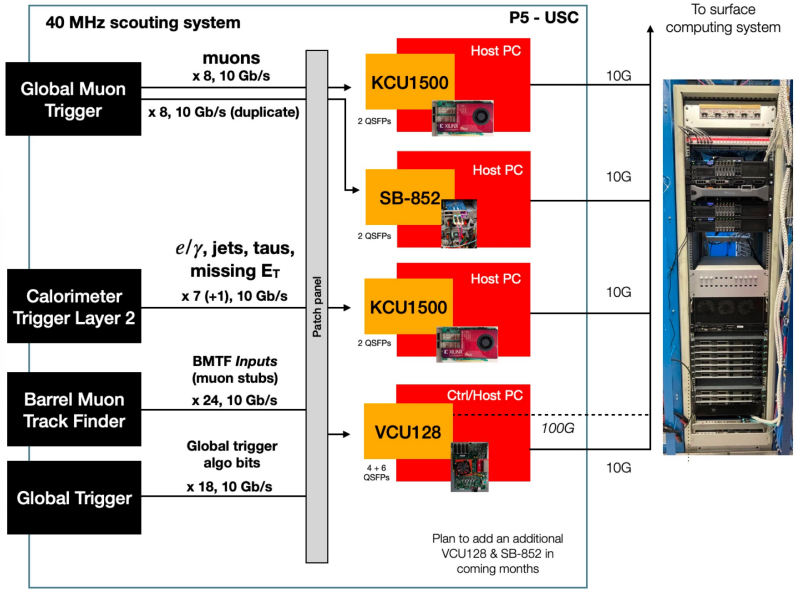
- Characterization of data collected by the system
- Computing models for online/offline analysis
- Exploration of distributed frameworks for data analysis:

  1 min of L1DS ⟹ 2.5GB of unpacked GMT data (parquet)
  - CERN Spark cluster within the CERN SWAN platform
  - Distributed RDataFrame with Spark/Dask back-end



Università degli Studi di Padova
INFN
Fondazione ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani Piano Nazionale di Ripresa e Resilienza

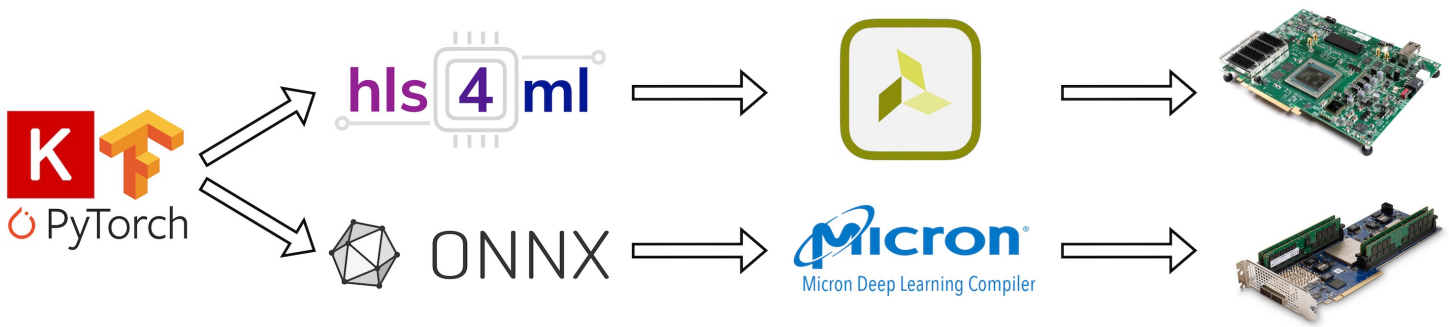
40MHz Scouting for online trigger-less data processing



40MHz Scouting demonstrator for LHC Run-3 (2022-2025)

- Demonstrator tested as proof of concept with three types of FPGA-based receiving boards
- Acquisition of L1 trigger primitives at the full BX rate from multiple sources:
  - Global Muon Trigger (GMT): **up to 8 GMT muons**
  - Calo Layer 2 (DeMux): **Up to 12 e/γ, jets, τ and missing E_T**
  - Barrel Muon Track Finder (BMTF): **BMTF inputs**
  - Global trigger (GT): **GT algorithm bits**

- For close-to-real-time analysis ML algorithms can be developed in hardware for the 40MHz Scouting pipeline



ML pipeline for the 40MHz Scouting system

- `hls4ml`: firmware implementations of machine learning algorithms using high level synthesis language (HLS)

- `Micron Deep Learning Accelerator` (MDLA): Micron proprietary inference engine to run ML models on FPGAs

Università degli Studi di Padova
INFN
Fondazione ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani Piano Nazionale di Ripresa e Resilienza

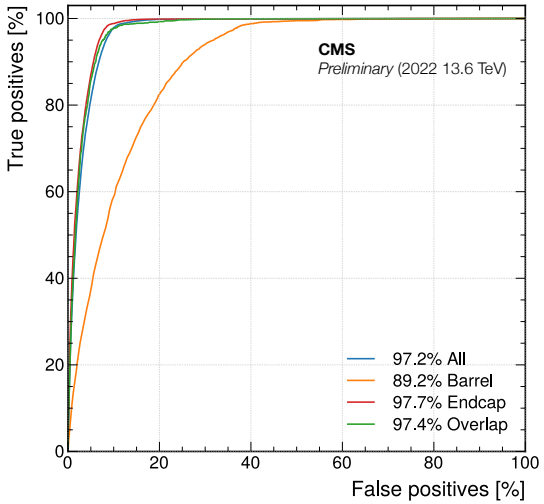
Deep learning on FPGAs for L1DS

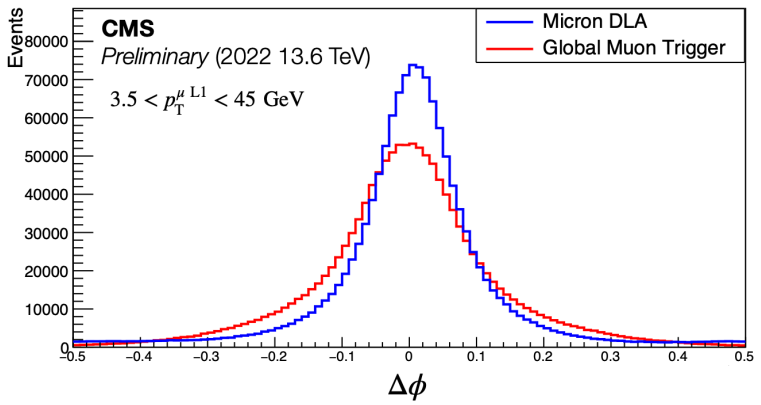
Recalibration of L1 muons

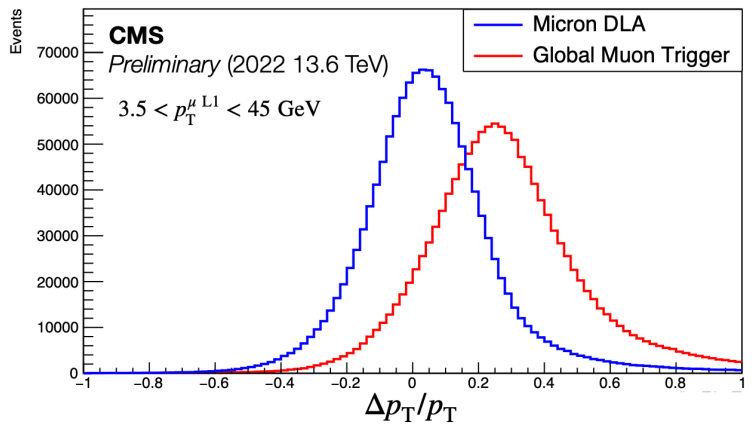
- Limited resolution of the trigger objects: no tracker information plus calibrated for a given efficiency at threshold
- Improve calibration for for a physics analysis
- Use the offline reconstructed (RECO) objects as the learning target of a neural network (NN)
- First implementation in a test setup with the MDLA hardware and collision data collected at the start of LHC Run-3

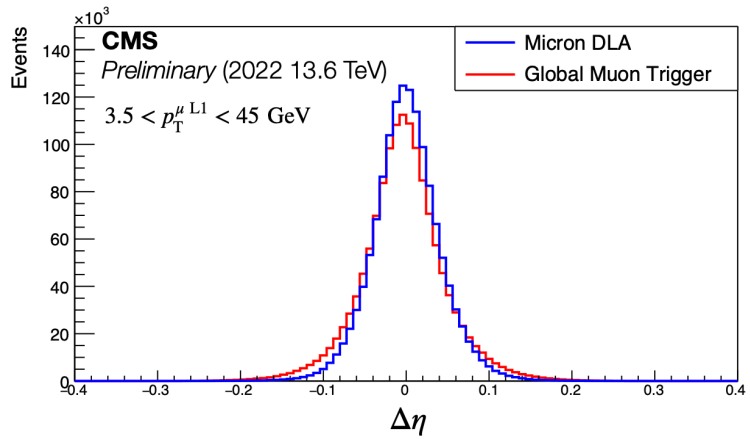
Muon pairs classifier

- Trained a NN for binary classification to distinguish real/fake L1 muon pairs
- Real (fake) → muon (not) matched to a RECO track
- Employ a classifier to select higher purity samples for study of di-muon events









The distribution of differences between the Micron Deep Learning Accelerator (MDLA) prediction (or GMT) values, and the offline reconstructed muon tracks.

Università degli Studi di Padova · INFN · Fondazione ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing · Finanziato dall'Unione europea NextGenerationEU · Ministero dell'Università e della Ricerca · Italiadomani Piano Nazionale di Ripresa e Resilienza

Conclusions and outlook

CMS Level-1 trigger Data Scouting for online trigger-less data processing

- Continuous efforts in advancing trigger-less data acquisition and processing techniques.
- Ongoing development addressing both firmware (receiving streams from L1 boards) and software aspects (online and offline computing models).
- First promising deployment of ML implementations on FPGA targeting L1 Muon objects:
  - NN models for calibration and fake detection
  - Exploring the potential to extend ML algorithms beyond L1 Muon objects
  - Current focus on transitioning to Xilinx VCU128 boards and `hls4ml`

More on the other activities of the group

Online reconstruction algorithms for data scouting

- Data scouting of new CMS Drift Tubes Front-End electronics with FPGA-based boards
- CHEP2023 – M.Migliorini et al *"Triggerless data acquisition pipeline for Machine Learning based statistical anomaly detection"*

Development and testing of RDMA over converged ethernet (RoCE) on FPGA for data transfer from detectors' front-end to computing servers (FERoCE)

- Dynamic simulation implemented and now working towards RoCEv2 with UDP/IP transport protocol targeting Xilinx VCU118
- TWEPP23 – G.Bortolato et al "Front-End RDMA Over Converged Ethernet, real-time firmware simulation"

Development of quantum-inspired Tree Tensor Networks for classification in Trigger on FPGA

- Project in its initial phase, synergy with colleagues from CN1 Spoke 10

Università degli Studi di Padova
INFN
Fondazione ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani Piano Nazionale di Ripresa e Resilienza