ICSC SPOKE 2

The Milano-Bicocca FPGA Cluster

Francesco Brivio[1]
Paolo Dini[1]
Simone Gennai[1]

[1]INFN Milano-Bicocca

Spoke2 Annual Meeting

18-20 Dicembre 2023
CINECA

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

INFN MILANO BICOCCA

Involved People and Working Packages

WP2 Flagship: "Development of ultra-fast algorithms running of FPGAs" (definition link)

Project in between WP 2 and WP 4

WP 2
Tools and algorithms for Experimental Collider Physics

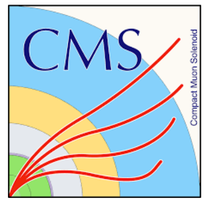
Find/Propose use cases and develop specific HEP algorithms

WP 4
Tools for porting/optimization on new architectures

Develop FPGA Cluster for the needs of HEP use cases

Person power involved in Milano-Bicocca

- Valentina Camagni (PhD student)
- Francesco Brivio (Tecnologo TD)
- Paolo Dini (Ricercercatore INFN)
- Pietro Govoni (Professore UniMib)
- Simone Gennai (Primo Ric. INFN)



- Maurizio Martinelli (Professore UniMib)
  + possibily 1 PhD student (?)

+ Researchers from other institutes (INFN + Universities) with FPGA experience, interested in exploiting the FPGA cluster

The (future) Milano-Bicocca FPGA Cluster

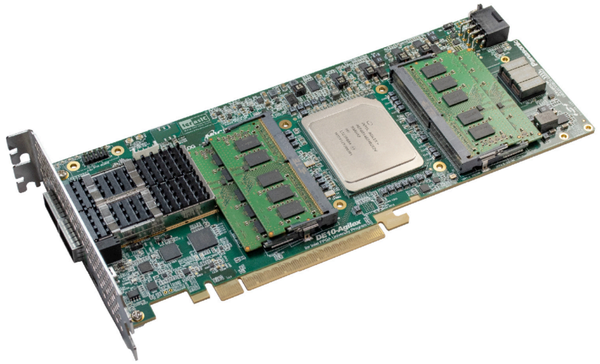
Hardware Setup and Cluster Topology

- **Servers**
  - Total of 16 boards equipped with FPGA
    - 8 Xilinx U55C Alveo boards
    - 8 Intel Agilex Terasic boards
- **Storage**
  - 10 TB storage for each board type
- **Connection**
  - Xilinx boards with 2 QSFP28 ports (100G)
  - Intel boards with 2 QSFP-DD ports (200G)

Exploiting the double-ports it's possible to implement "*full mesh topology*" to link 8 boards using breakout cables (inspired by LHCb's RETINA system)

Hardware expected end of Q2 of 2024
Payed by the Terabit project

Intel Agilex Terasic Board



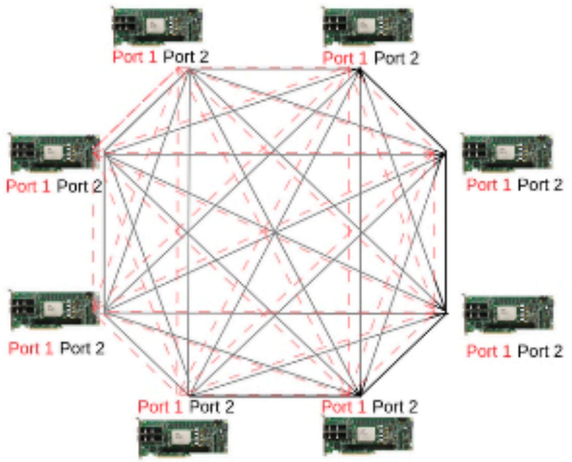


Figura 1. Esempio di rete full-mesh per le board Terasic

Xilinx U55C Alveo Board

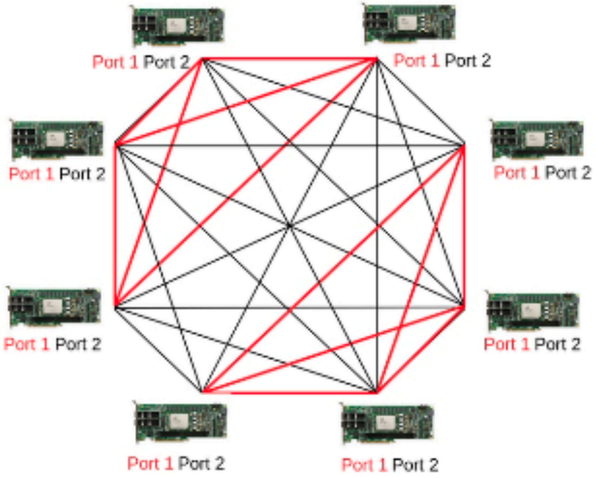


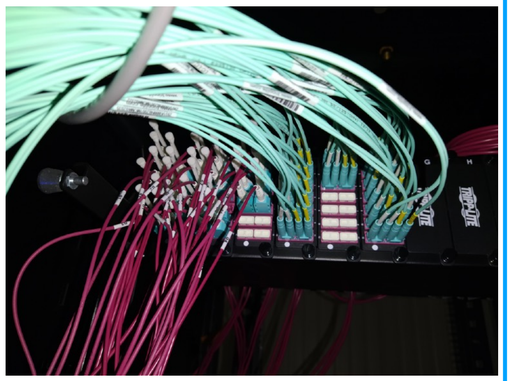
Figura 2. Esempio di rete full-mesh per board Xilinx U55C

FPGA Cluster Use Cases

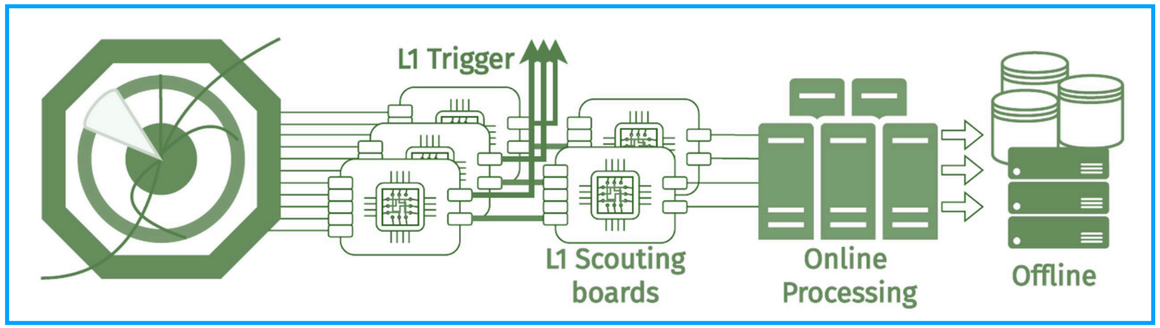
Cluster intended as shared resource available to all Spoke2 projects (with special attention to High Energy Physics use cases)

- **LHCb**
  - Extend and commission (a larger) RETINA system for "online" tracking at L1
- **CMS**
  - CMS Phase-2 scouting analysis @ 40 MHz
  - Tau regression for low p_T trigger taus
  - "Fast Inference As Service" for ML
- **ATLAS**
  - L1 and HLT triggering ML algorithms
    - Ultra fast CNN for muon triggering

Expect to have the cluster resources also available via cloud



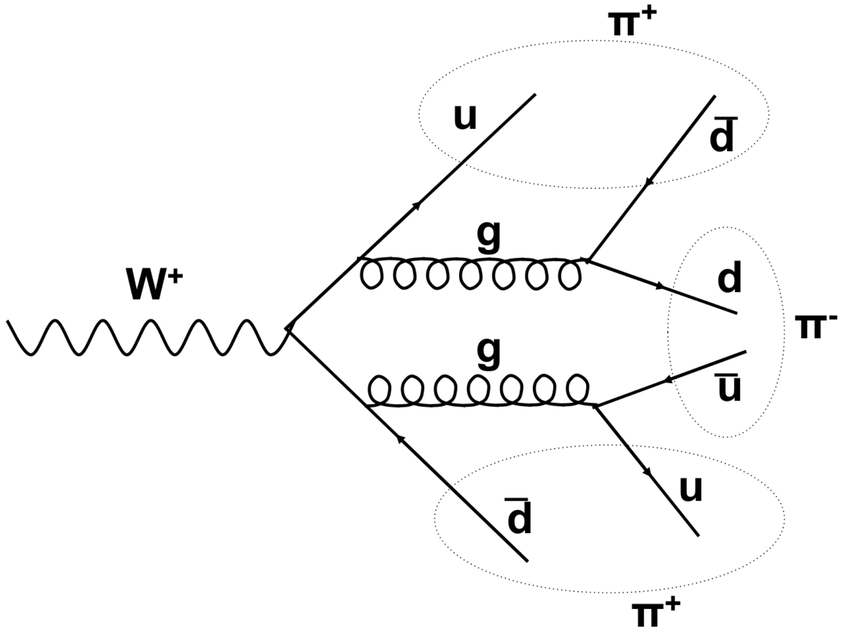
Photos from the RETINA system in the LHCb Control Room (CERN Point 8)

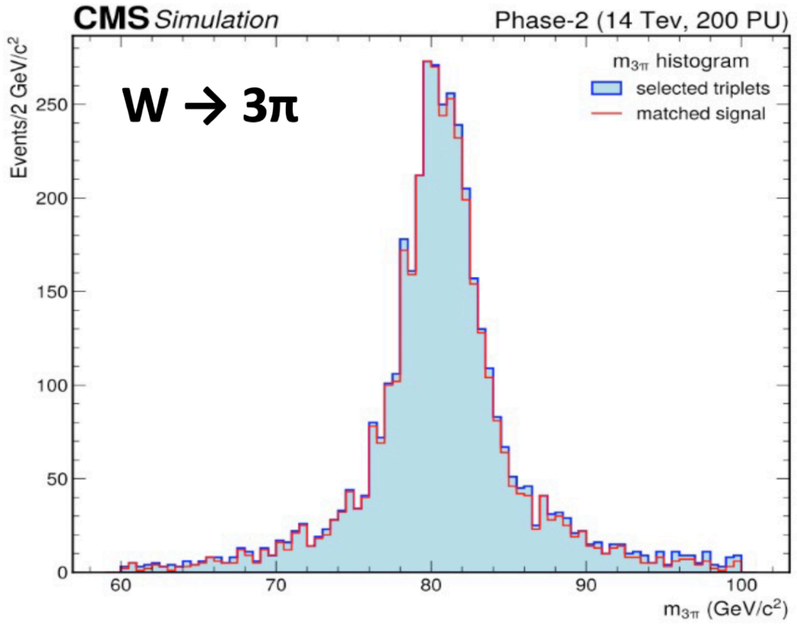


Physics use case: the W→3π analysis for CMS Phase-2

Analysis of the decay of a W boson to 3 charged pions

- **Extremely rare process**
  - BR (W→3π) < ~ 9 · 10^{-7}
    - Latest result by CMS in PhysRevLett.122.151802
- **Perfect use case for CMS Scouting at 40 MHz during Phase-2**
  - Analyze online all *pp* collision events
  - Requires extreme levels of background rejection...
    - ... and high signal selection efficiency!





- **Current status**
  - Analysis implemented in the "*c++ style*"
    - Processing "*offline*" using Phase-2 simulated data
  - Working on the implementation of a L1 trigger already now in Run 3 (2022-2025)
- **Prospects in the context of WP 2/4**
  - Exploit the FPGA Cluster to perform the analysis "*online*"
    - Physics interesting
    - Learning opportunities

W→3π on FPGA

Two main developments ongoing

Spoke 2 repository: https://github.com/ICSC-Spoke2-repo/W3Pi

- **Translate the current (cut based) analysis to HLS**
  - Synthesize a firmware to run on CMS L1T scouting boards, doing:
    - Data decoding
    - Data selection
    - Data pre-processing
      - ...for the next ML step!

- **Improve the current analysis with ML techniques**
  - Exploring different ML techniquest (NN, BDTs...) to improve the event selection and background rejection efficiency
    - "Cut-based" analysis: $N_S/N_B \sim 7.2 \cdot 10^{-6}$
    - "ML-based" analysis: $N_S/N_B \sim 8.8 \cdot 10^{-6}$
    - Obtaining similar signal purity (92%)
  - Implementation on FPGA with most recent ML-HLS tools
    - *hls4ml*
    - *Conifer*

hls4ml Conifer

Course on HLS programming organized by WP 2 + WP 4

https://agenda.infn.it/event/38191/

Profited from CMS ML@L1Trigger workshop at CERN

https://indico.cern.ch/event/1301711/

Summary

The Milano-Bicocca FPGA Cluster

- **Project shared between WP 2 and WP 4**
  - Flagship in WP 2
- **FPGA Cluster to be built in Milano-Bicocca and made available to INFN via cloud**
  - Mix of Intel and Xilinx boards, with possibility of "full-mesh topology"
  - Hardware expected end of Q2 of 2024
- **Use Cases**
  - Triggering, Scouting analysis, online tracking... → mostly involvement from LHC expertiment
    - But not only limited to this!
      - *e.g.* The BondMachine project (INFN and Università Perugia)
  - Currently focussing on
    - Porting the W→3π CMS scouting analysis to FPGA
      - Including deployment of NN on FPGA for fast inference
    - Several possible other use cases being investigated

Backup Slides

WP 2 Flagship

Spoke 2 use cases

Definition and workflow document

Ultra-fast algorithms running on FPGA

Spoke	2
WP	2, 4
Use case short name	FAST_FPGA
Use case ID	UC2.2.3
Expected Completion	31/8/2025

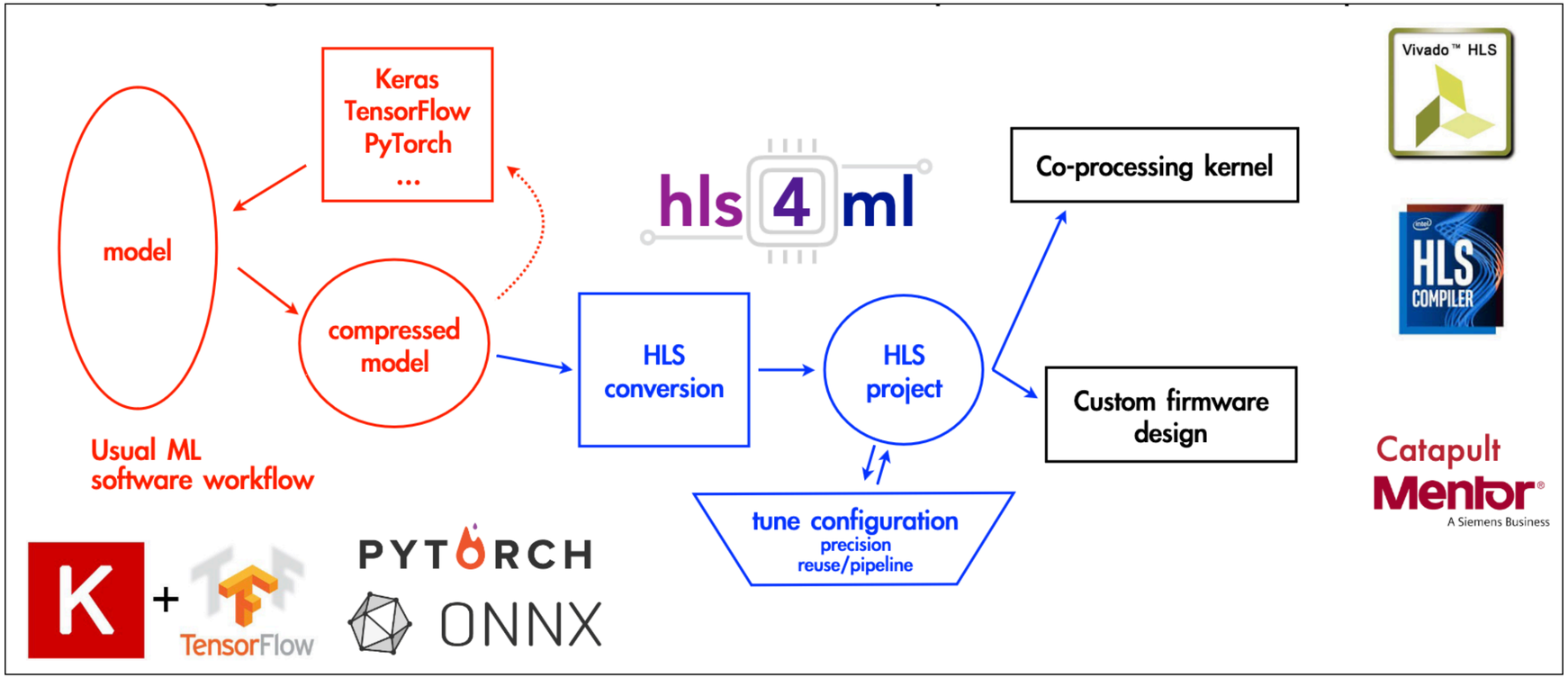
Flagshop Definition Document

https://docs.google.com/document/d/1i0qke_GiJ9VOyZIDAtc-HnnLANsSBORTJt9qw_nEnqA/edit#heading=h.rnaep8ewc7o

hls4ml and *Conifer*

- Repository
  https://fastmachinelearning.org/hls4ml/
  https://github.com/thesps/conifer

- Introduction and Tutorial
  https://agenda.infn.it/event/38191/timetable/?view=standard#b-34263-hls4ml-conifer



The BondMachine project

- Repository
  https://github.com/BondMachineHQ

- Introduction and Tutorial
  https://agenda.infn.it/event/38191/timetable/?view=standard#b-34265-the-bond-machine

The BondMachine is a software ecosystem for the dynamic generation of computer architectures that:

- Are composed by many, possibly hundreds, computing cores.
- Have very small cores and not necessarily of the same type (different ISA and ABI).
- Have a not fixed way of interconnecting cores.
- May have some elements shared among cores (for example channels and shared memories).

M. Mariotti (INFN Perugia)