Hardware (FPGA) Accelerated Deep Learning for Edge Intrusion Detection Systems

Ioannis Morianos

ML4ECS Workshop
HiPEAC 2026 Krakow

Dienekes

DIENEKES

- Cybersecurity
- Hardware Security
- Hardware Acceleration
- Computer Architecture

https://dienekes.eu/

Rescale Workshop in HiPEAC 2026 Krakow
Wednesday 28/01
11:00-18:30

https://rescale-project.eu/

Intrusion Detection Systems

- Monitor network traffic and devices for malicious activity or policy violations

- Detect malicious activity such as:
  - DDoS and DoS attacks
  - Injections
  - Backdoors

Detection Techniques

- **Ruled-based detection:** Matches known attack patterns (signatures) like SNORT®

- **Anomaly-based detection:** Identifies deviations from normal behavior

Rule-based vs DL-based IDS

- Rule-based IDS struggle with novel, sophisticated attacks

- DL-based IDS try to tackle that problem but:
  - they guess (make mistakes)
  - they need more computational resources.
  - they are slower

- Need for advanced, scalable solutions for acceleration with FPGA

Why FPGA?

- More specific hardware solutions than CPU-GPU, especially for high throughput use cases.
- Power-efficiency.
- Reconfigurability.

FPGA-Accelerated DL-based IDS

- Custom implementation with RTL (Verilog, VHDL)
- High-level frameworks like HLS4ML, Logicnets and FINN
- Even more software friendly platforms like Vitis-AI (DPU cores)

Industrial Intrusion Detection System (I^{2}DS)

- FPGA-Accelerated Deep Learning IDS
- Deep Learning for Anomaly Detection.
- FPGA for Real-Time Processing.

I. Morianos *et al.*, “I2DS: FPGA-based deep learning industrial intrusion detection system,” in *Embedded Computer Systems: Architectures, Modeling, and Simulation*, Springer, 2025, pp. 165–176.

Evaluation Board ZCU104

It’s a hybrid CPU-FPGA System on Chip (SoC)

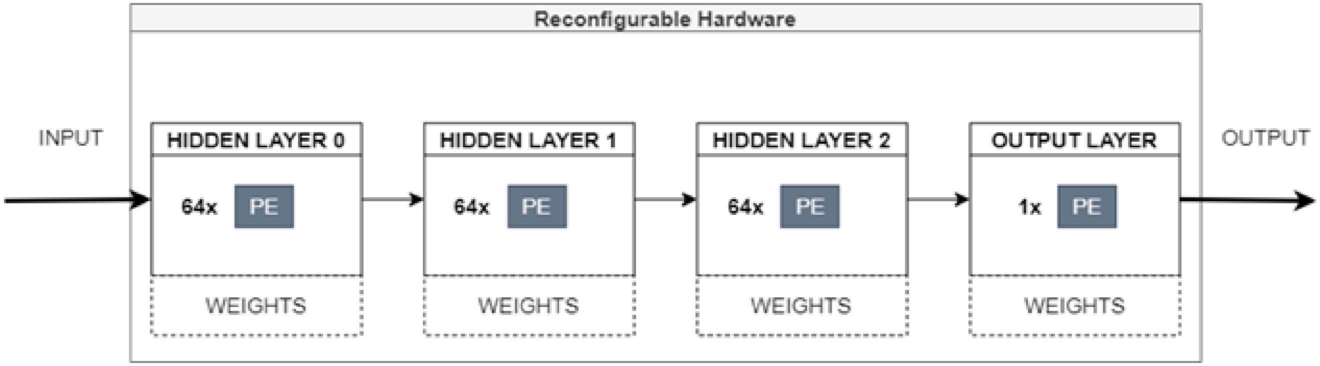
AMD

XILINX

ALL PROGRAMMABLE™

Creating a Hardware model

- Creation of binarized (quantized) datasets
- Quantization Aware Training (QAT) using 2-bit weights and activations.
- Synthesis of the hardware model.
- Generation of the bitstream.
- Implementation of a host (SW) that will communicate with the accelerator with efficiency.



Results

- Performance enhancement with acceleration
- Low resource utilization in the FPGA fabric
- High energy efficiency suitable for IIoT edge devices
- More suitable solution for IIoT compared to CPU and GPU
- Scalable Architecture (room for more parallelization)

Thank you!
Questions?