



Contribution ID: 212

Type: **not specified**

Open source FPGA NVMe accelerator platform for BPF driven ML processing with Linux/Zephyr

Wednesday, September 14, 2022 10:25 AM (30 minutes)

The talk will describe an open source NVMe development platform developed by Western Digital and Antmicro for server-based AI applications. The system combines an FPGA SoC with programmable logic and an AMP CPU, running Zephyr on the Corex-R cores handling NVMe transactions and Linux on Cortex-A in an openAMP setup.

The system utilizes Zephyr RTOS to perform time critical tasks including handling the base set of the NVMe commands, while all the custom commands are passed to the Linux system for further processing. The Linux system runs an uBPF virtual machine allowing users to upload and execute custom software processing the data stored on the NVMe drive.

The platform (custom hardware from Western Digital and open source software and FPGA firmware from Antmicro) was designed to enable users running ML pipelines designed in Tensorflow. To make it possible, the uBPF virtual machine has been extended to include functionalities to delegate certain processing to external, native, libraries interfacing the BPF code with hardware ML accelerators.

The platform includes an example showcasing a TensorFlow AI pipeline executed via the uBPF framework accelerating the AI model inference on an accelerator implemented in FPGA using TVM/VTA.

The platform intends to be a development platform for edge, near data processing acceleration research

I agree to abide by the anti-harassment policy

Yes

Primary author: GUGALA, Karol (Antmicro)

Presenter: GUGALA, Karol (Antmicro)

Session Classification: IoTs a 4-Letter Word MC

Track Classification: LPC Microconference: IoTs a 4-Letter Word MC