Open Source NVME IP with AI Acceleration

CHIPS Alliance Fall Workshop, 2021-10-12

Karol Gugala, kgugala@antmicro.com
Anand Kulkarni, anand.kulkarni@wdc.com

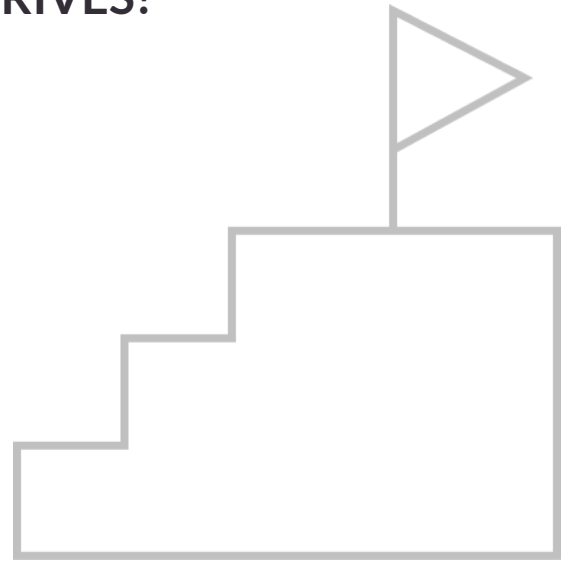
antmicro

WHAT EXACTLY IS NVME?

- Family of specifications for communication with non-volatile memory
- Supports various form factors and transport protocols
- M.2 and PCIe is the single most commonly used combination
- Specifications are freely available from NVM Express website

WHY DO WE NEED ACCELERATORS IN NVMe DRIVES?

- Machine Learning usually operates on large amounts of data
- Transferring data back and forth generates bottlenecks and costs
- NVMe accelerators reside close to stored data
- They allow us to process the data on the fly, or perform computation on already stored data, detect interesting patterns
- Data can be processed directly without consuming compute resources / spinning up machines



OUR GOALS

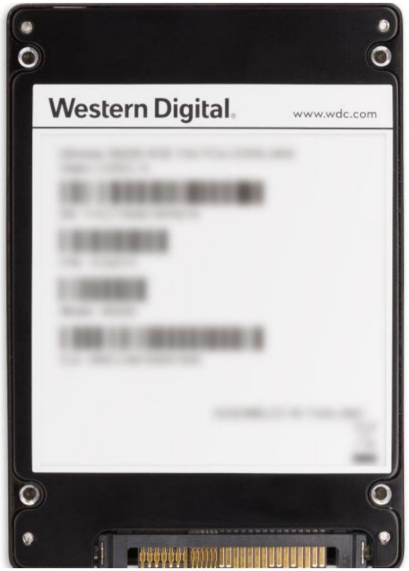
- Build an open source platform for NVMe accelerators development based on Xilinx US+ MPSoC
- Create an open source NVMe FPGA core
- Prepare firmware that handles essential NVMe operations
- Expand initial NVMe implementation with custom accelerator-related extensions

Western Digital®

antmicro

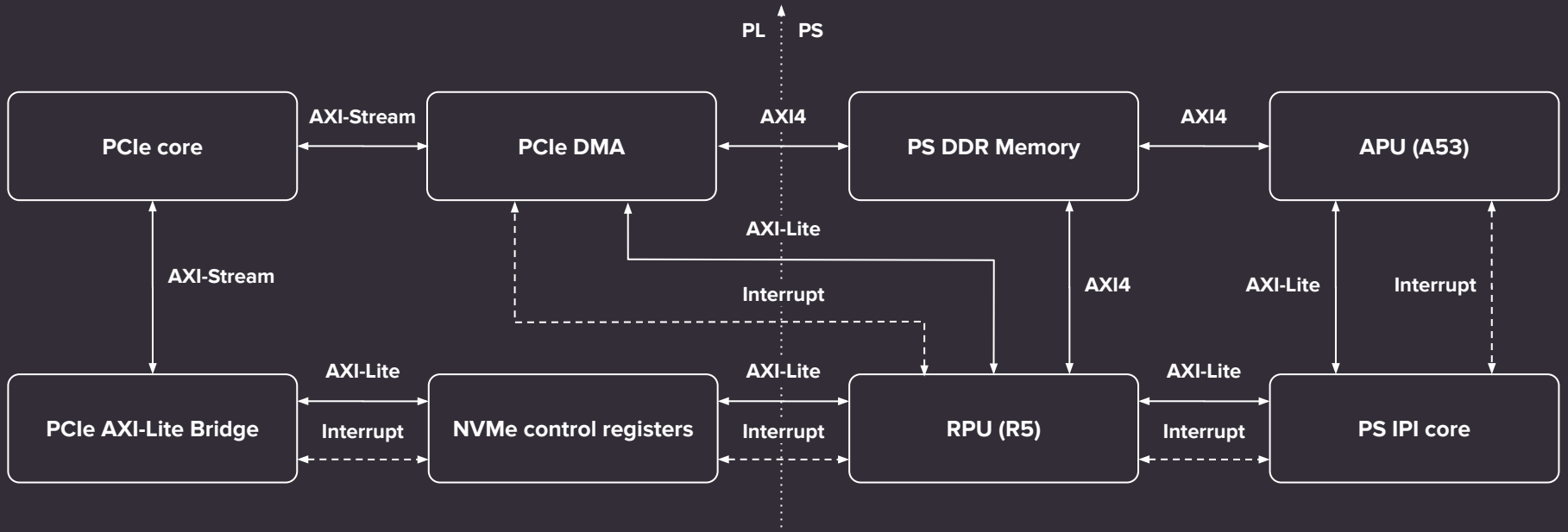
HW Platform

- FPGA Based PCIe ML/AI Accelerator Device in U.2 Formfactor
- Xilinx Ultrascale+ MPSoC XCZU7EV
- 4GB DDR
- Gen3 x4 PCIe 2.5” SFF
- 25W Max Power



antmicro

SYSTEM OVERVIEW



PCIE CORE

verilog-pcie:

- Part of the Corundum project
- Host PC memory access via AXI DMAs
- NVMe register file access via AXI Bridges
- Uses Xilinx PCIe Hard IP underneath

NVME REGISTER FILE

- Implemented with Chisel
- Implements registers for NVMe spec 1.4
- Register definitions are automatically generated from the NVMe specification document
- Two AXI-Lite interfaces
- Generates interrupts both for host PC and firmware
- Reuses parts of another open source Chisel IP - FastVDMA

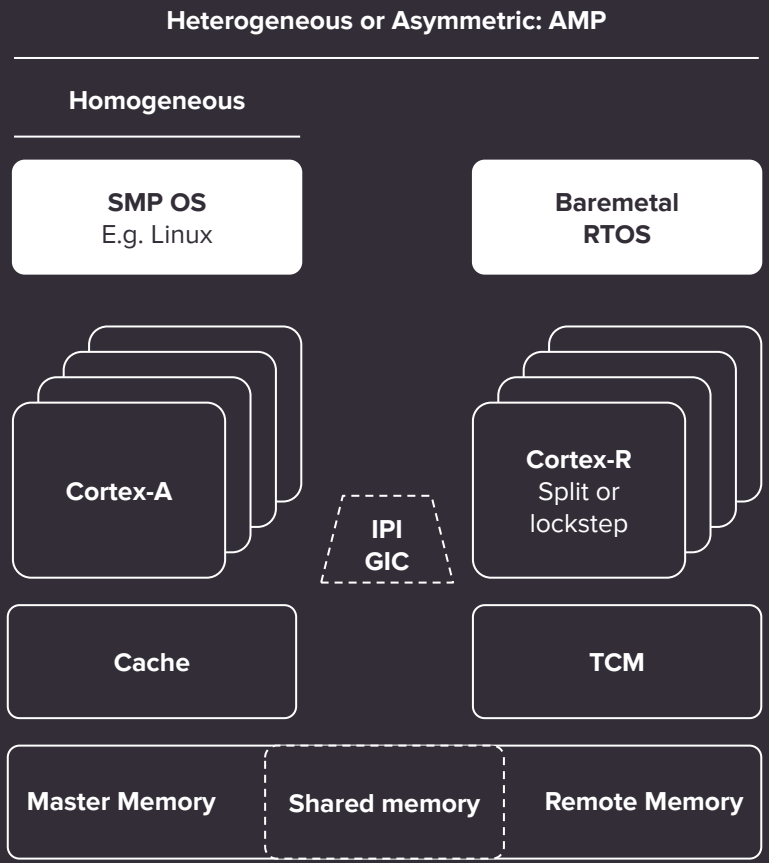
REAL-TIME SOFTWARE - ZEPHYR

- Open Source RTOS
- Easy to use
- Comes with all necessary building blocks
- Ported to the Cortex-R5 RPU with OpenAMP support
- Zephyr application is responsible for handling standard NVMe commands, custom commands are forwarded to Linux application

antmicro

OPENAMP

- Framework for systems with asymmetric multiprocessing
- Provides easy method of communication between CPUs in AMP system
- RPU side runs Zephyr and is controlled from Linux application (using openAMP)
- Linux application implements openAMP communication and interfaces NVMe blocks with eBPF virtual machine



ACCELERATOR FIRMWARE

- eBPF bytecode
- The firmware can be loaded using a custom NVMe command
- Executed inside BPF virtual machine
- Originally used for packet filtering
- Clang generates bytecode from C source
- eBPF virtual machine has been extended with Tensor Flow bindings

TensorFlow Lite

ENHANCING TENSORFLOW LITE WITH VTA

- VTA
  - Programmable accelerator IP core
  - Written in Chisel
  - Part of the TVM framework
- TFLite can utilize accelerators using delegates
- https://github.com/apache/tvm-vta

THANK YOU
FOR YOUR ATTENTION!