

Additional studies during workshop period

B. Abi, R. Wastie, G. Barr (Oxford)

Outline:

- 1 - Future questions
- 2 - DPDK
- 3 - Vitis embedded flow
- 4 - Summary

1. Kostas and Rob have outlined investigations to read Ethernet and are targeting (largely) VUP devices with 100G:

- Goal is to be prepared for the PDR to be sure we have a way and some demonstrations of handling UDP packets.
- Complementary to the work with the FLX-712 with the horizontal drift and the Versal FELIX development



We are enthusiastic to explore **Vitis HLS**. This is Xilinx's more abstract firmware and is still rather new.

2. Next steps after that will include:

- Possible convergence of the designs for the different detectors (HD, VD-top, photonDetectors, (ND) etc.)
- Need for all streams to look the same by the time they reach the readout software.
- In Oxford, gain more familiarity with the firmware; we need firmware platform to test the coming FPGA boards

... and investigate UDP packet handling in **DPDK**

- See dpdk.org
- This is a software development kit. Has a core that is very useful.
- 10 years old now, started by Intel. Lots of new things added. Recently has more things about IOMMU.
- This is why we bought Intel NICs for the protoDUNE machines :) However it has user-mode drivers for many cards.
- Notably Xilinx is not a member?, but AMD is



- 2.2. Environment Abstraction Layer
- 2.3. Core Components
 - 2.3.1. Ring Manager (librte_ring)
 - 2.3.2. Memory Pool Manager (librte_mempool)
 - 2.3.3. Network Packet Buffer Management (librte_mbuf)
 - 2.3.4. Timer Manager (librte_timer)
- 2.4. Ethernet* Poll Mode Driver Architecture
- 2.5. Packet Forwarding Algorithm Support

Our plans for summer

- Make a UDP packet emitter and a packet absorber using 10G Intel NIC cards and DPDK poll mode user drivers
- Set up the emitter as way of generating packets for FW tests ***High priority***
- Play around with DPDK KNI (Kernel-NIC interface) for out-of-band data. Feed out-of-band data to Linux kernel.
- With readout SW authors, understand how DPDK can follow FELIX drivers/flxcard methodology to interface with DUNE flxlibs.
- Look at

github.com/Xilinx/dma_ip_drivers DPDK



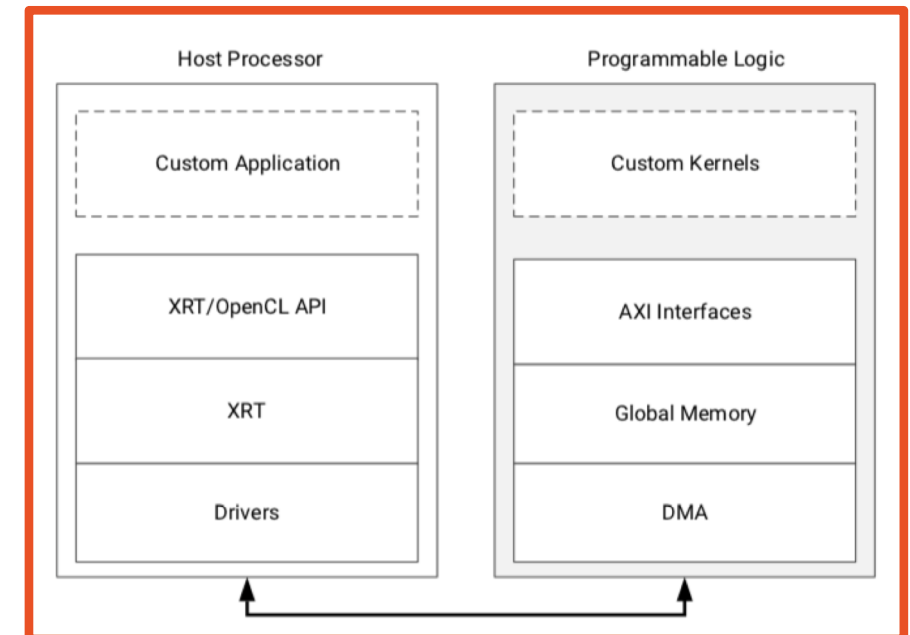
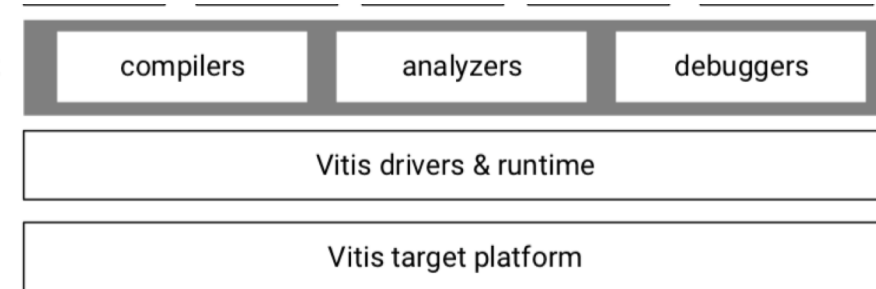
- 2.2. Environment Abstraction Layer
- 2.3. Core Components
 - 2.3.1. Ring Manager (librte_ring)
 - 2.3.2. Memory Pool Manager (librte_mempool)
 - 2.3.3. Network Packet Buffer Management (librte_mbuf)
 - 2.3.4. Timer Manager (librte_timer)
- 2.4. Ethernet* Poll Mode Driver Architecture
- 2.5. Packet Forwarding Algorithm Support

- We are planning FW studies to evaluate Vitis for the future. Exploratory work by Roy + Babak already
- Vitis is a new workflow provided by Xilinx, part targeted at using accelerator cards for e.g. DL etc.
- For Alevo cards, there is a 'Vitis target platform' predefined.
- For other cards, there is a Vivado flow to create a Vitis target platform. So we could make our own target platform for the FELIX cards, with customizations for e.g. receiving data

DUNE algorithms as 'Kernels'

Vitis core development kit

XRT



- Our plan is to make a Vitis target platform for FELIX card. Includes libraries interfacing our IP. (later: porting of existing DUNE FW)
- Use IP from e.g. open source projects (openNIC etc, see later), attempt to integrate in platform.
- Test interworking with OpenCL and Host/FW-kernel model
- Test interworking with DPDK
- Think about interworking with readout software (with readout software group help)

DUNE algorithms as 'Kernels'

Vitis core
development
kit

XRT

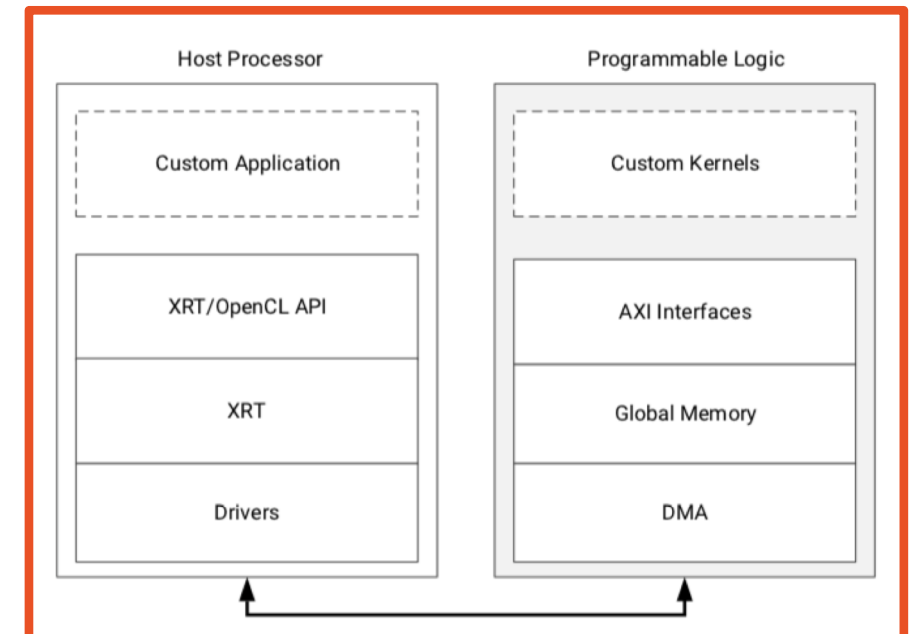
compilers

analyzers

debuggers

Vitis drivers & runtime

Vitis target platform



1. Alveo platform (U250,U280) ZYNQU+

- i. Xilinx 16nm UltraScale
- ii. Closed hardware

2. VERSAL based accelerators

1. VMK180; versal VM1802

Prototyping and reference design for VERSAL Prime

Higher performance compare to ZYNQU+(Alveo)

2. VCK5000 versal VC1902

Xilinx ACAP architecture for GPU replacement

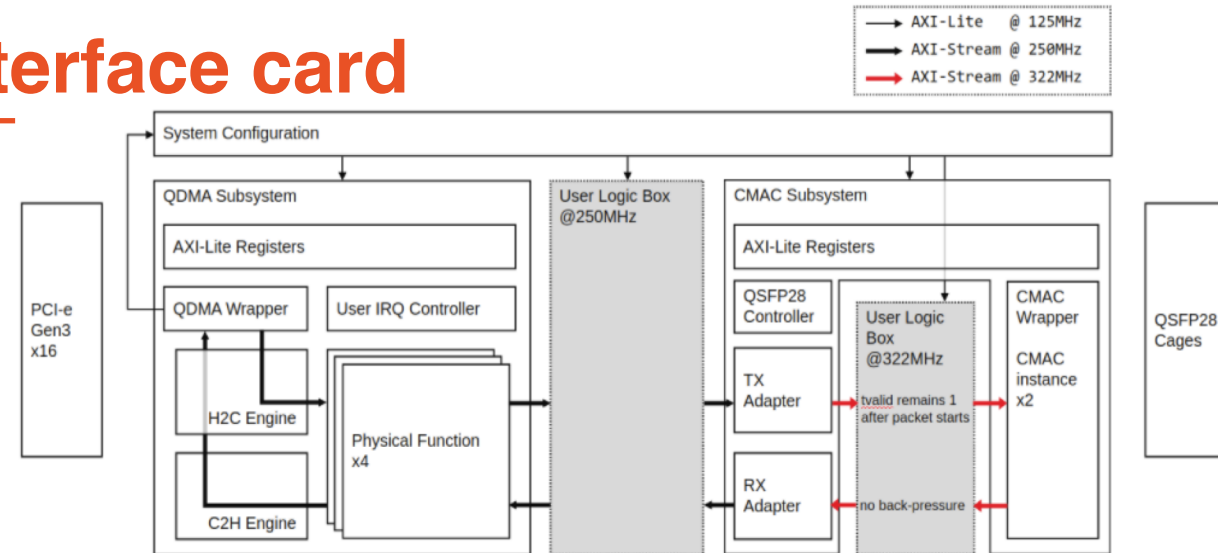
3. Custom design- FELIX (Open hardware)

1. Hardware concept is **Xilinx reference design** (VMK180) so firmware-wise it would be compatible with FELIX ATLAS firmware as well as other Versal cards, so aim is to make it as easy to use in Vitis as an accelerator.
2. Hardware expert opinion : VERSAL is significantly better than Ultrascale+ architecture (NOC and DSP and power per process ,....)

	Network Interfaces	Off-chip Memory Capacity		
Alveo250	2x QSFP28	64GB	Gen3x16	225W
Alveo280	2x QSFP28	32GB	Gen4x8	225W
VMK180	QSFP28/2x SFP28	8GB+DDR 4	Gen4x8	NA
VCK5000	2x QSFP28	16 GB	Gen3x16/Gen4x8	225W
Custom FELIX	20 Firefly	32/64GB	2xGen4x8	~100-120W

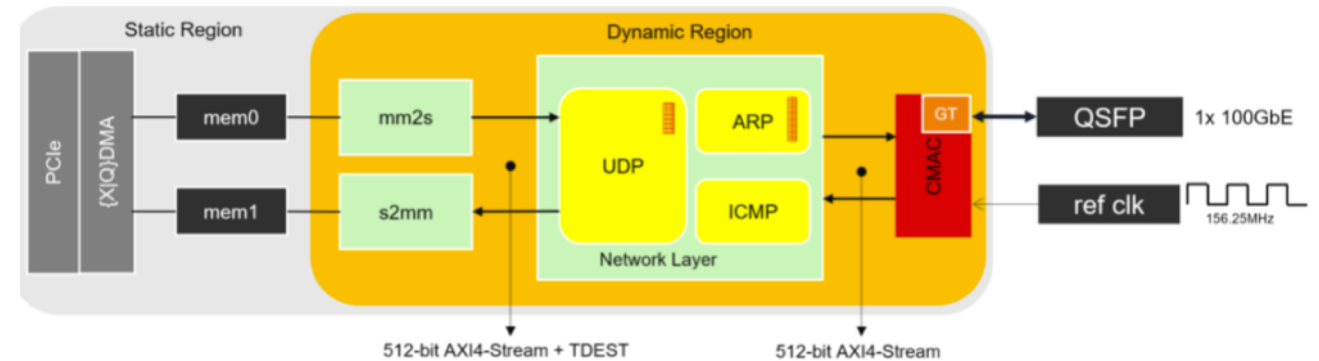
Xilinx Open-NIC Open network interface card

- We do not invent the wheel, already Xilinx has provided a complete reference code
 - All components , include the PCIe DMA and driver and Network stacks MAC ARM and ..
 - Linux PCIe engine and driver
 - One example is **OpenNIC** that is an open source project, focused on easing the integration of user logic for networking functions. Supported by Xilinx labs and big forum. Very similar to Xilinx smartNIC (backup slide)
 - OpenNIC carefully designed so that it hides many details and only exposes simple data and control interfaces to user logic.
- Repo : <https://github.com/Xilinx/open-nic>
- **Superior** to **Corundum** and **NetFPGA** (However we might explore **Corundum** too)



Xilinx out of the box network stacks for VITIS

- There are several VITIS HLS C code adapted open source frameworks for connecting FPGA-accelerated applications directly to networks without CPU intervention.
- Two of them offers separately provide good support for UDP/IP and TCP/IP network layer. Both are targeted at users developing applications with Xilinx Vitis, who need "out of the box" 100Gb/s networking. OpenNIC, on the other hand, is intended for networking research and experimentation with new FPGA-based networking components.
- Excellent features like :
 - CMAC subsystem can be configured to support jumbo frames for VD UDP packets
 - Network Layer kernel
 - The network layer kernel is a collection of HLS modules to provide basic network functionality. It exposes two 512-bit (with 16-bit TDEST) AXI4-Stream to the application,
 - The ARP table is readable from the host side, and the UDP table is configurable from the host as well.



DUNE FW group: Kostas

- Presented resource requirements for FELIX FW and Hit Finding processors
- Options for the VD Ethernet extension with an estimated resource utilization for a 100G solution targeting the VUP family
- DUNE Trigger Primitive Generation is ongoing, resources increase slightly in near future (mostly LUT, LUTRAMs and FFs)
- The FPGA choice will also affect both final resources and actual system cost
- MAC/PHY IP cores not always free
- Depending on final design choice, may not be enough hard cores; soft cores means extra resources

STFC UKRI: Rob Halsell

Now to December:

- Demo Aveo U50 with 1x 100G UDP to PCIe to server RAM
- Study trigger logic integration
- Meetings, documentation etc.

Post Dec

- Switch Testing,
- Out of order packet re-ordering
- DPDK PCIe driver and server setup
- Trigger logic Integration
- System Design for an all Ethernet COTS solution

Oxford: This talk

Next 3-4 months:

Create simplified Vitis platform with reusable components:

- I. Ethernet Subsystem. (Adaptation to VERSAL) and Network layer (UDP and ARP/ICMP)
- II. PCIe DMA to Host PC (already some work done, in progress),
- III. Test of Host/Kernel model

DPDK packet generator/absorber, discussion with readout group, KNI

Understand how Ethernet data stream feeds to readout SW

Later:

Implementing the trigger primitive generator

Further integration studies

Questions about the UDP input specification

Back-Up Slides

Alveo SN1000 SmartNIC Accelerator Card



- The Xilinx Alveo SN1000
 - SmartNIC offering software-defined hardware acceleration for all function offloads in a single platform.
 - directly offload CPU-intensive tasks to optimize networking performance, with an architecture that can accelerate a broad range of custom offloads at line rate
 - Direct Offloading on NVMe™ over TCP

Networking	
Stateless Offloads	Yes
Tunneling Offloads	VXLAN / NVGRE / Custom
SR-IOV	Yes
Advanced Packet Filtering	Yes
Acceleration	TCPDirect – TCP/UDP, Open Virtual Switch (OVS), Virtio-net, vDPA, DPDK, Onload®, Virtio-blk, Ceph RBD Client Offload
Manageability	
PMCI Protocols	NC-SI, PLDM Monitoring and Control, PLDM MCTP
PMCI Transports	MCTP SMBus, MCTP PCIe VDM

PCI Express	PCIe Gen 4 x8 or Gen 3 x16
Network Interfaces	2x100G QSFP28 DA copper or optical transceiver
Link Speeds	100GbE
Arm Processor	Discrete 16-core Cortex-A72 Processor
DRAM Memory	
DDR Format	-1x 4GB x 72 DDR4-2400 (Arm® Processor) -2x 4GB x 72 DDR4-2400 (FPGA)
Performance	
Full Duplex Throughput	200Gbps
Packet Rate	100Mpps
TCP Throughput	100Gbps
Latency (1/2 RTT)	<3us

Figure 2: System-level Interconnect Architecture

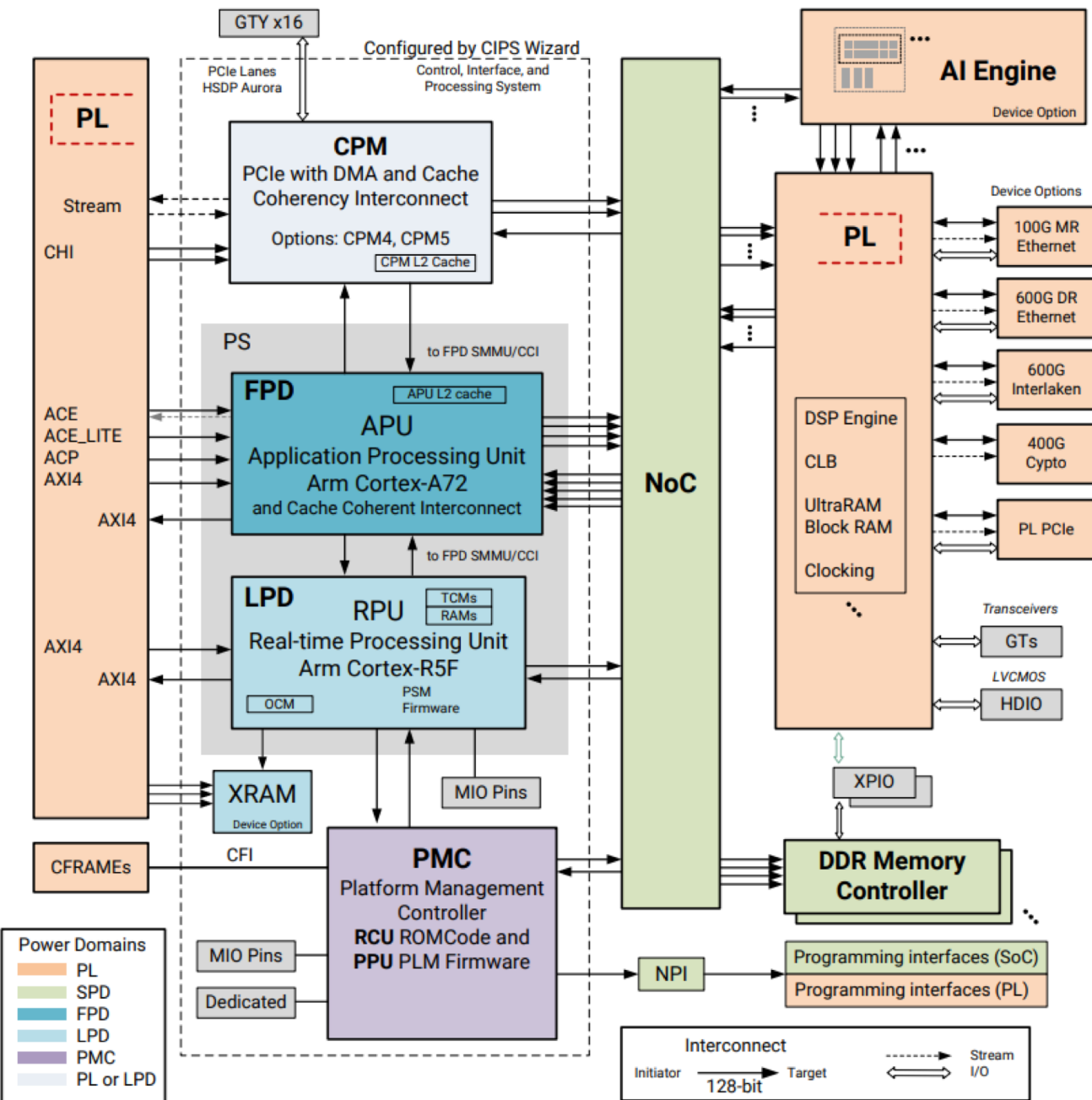
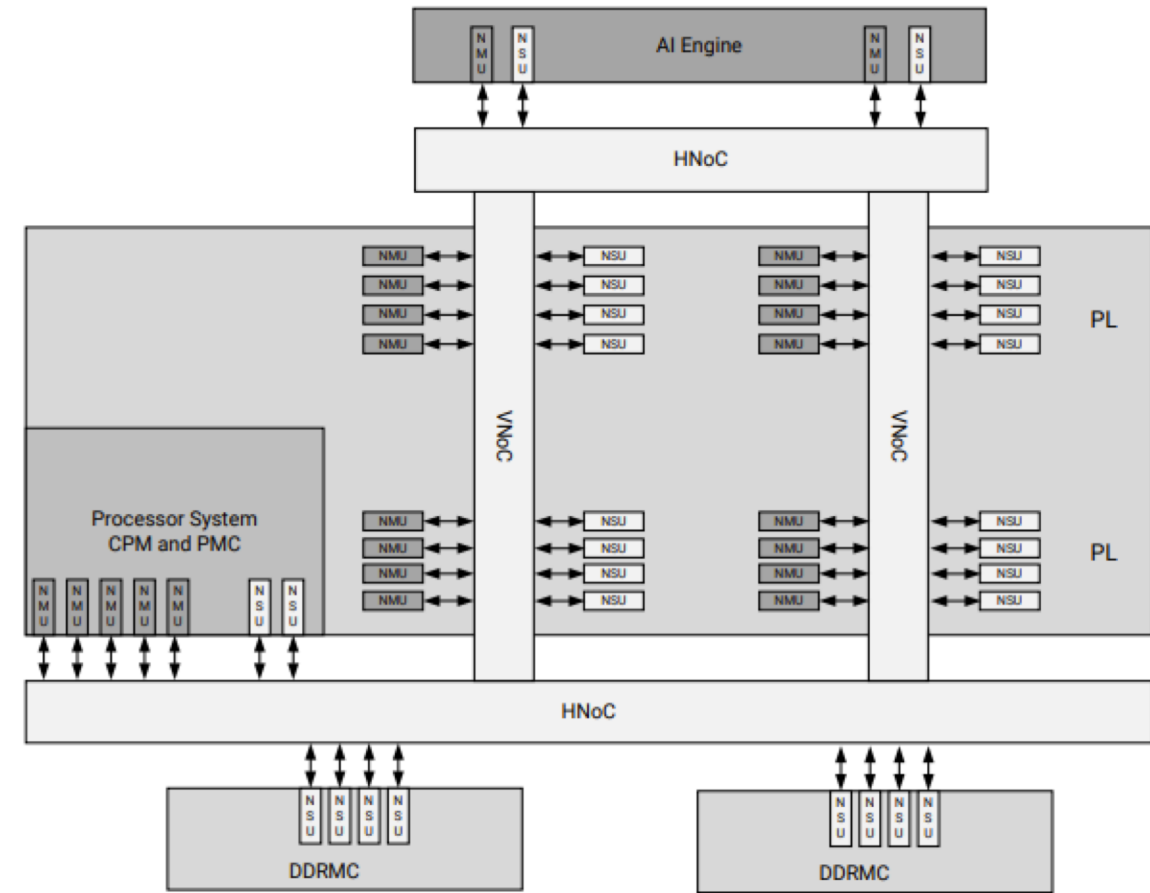
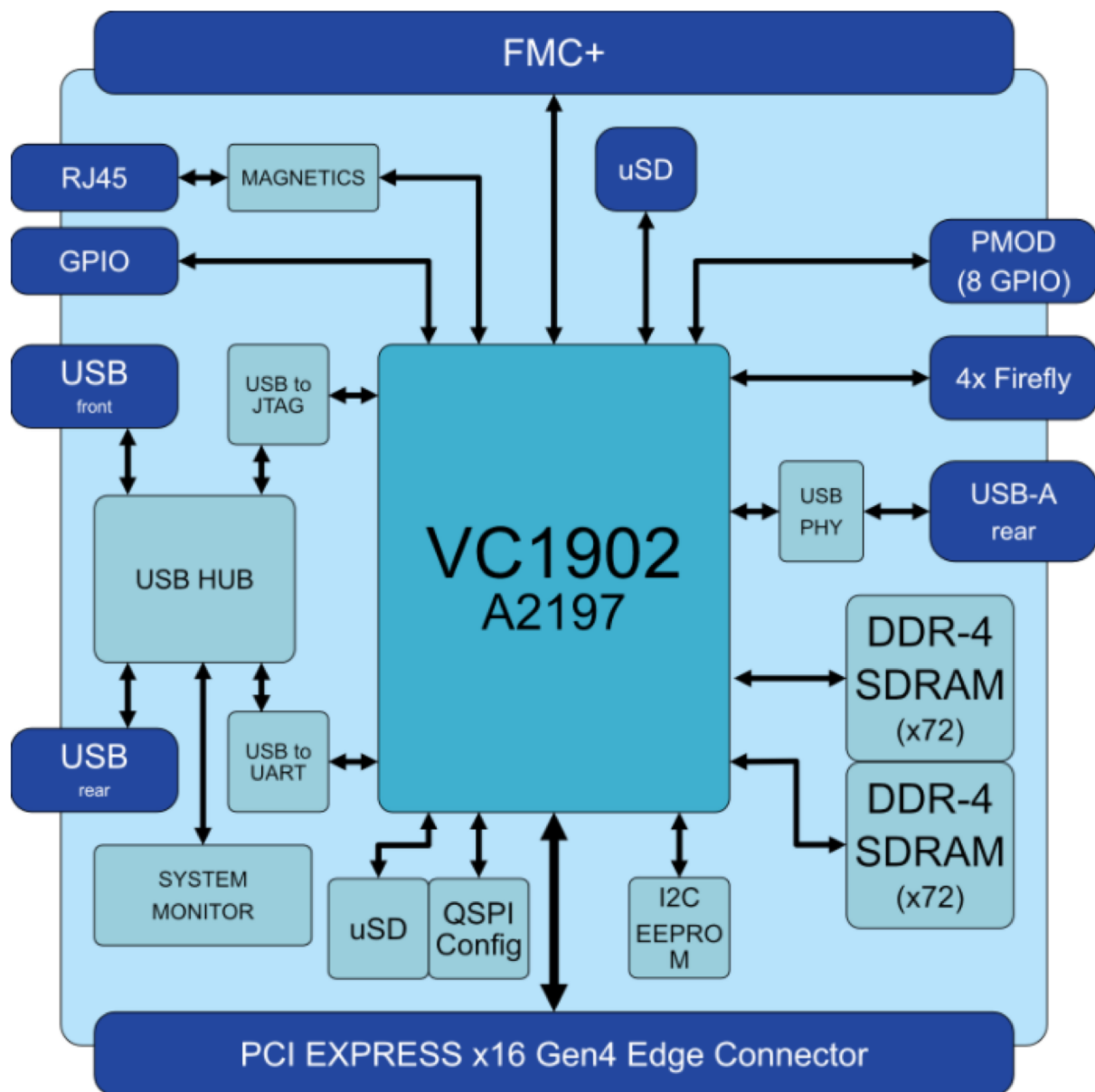


Figure 1: NoC Block Diagram





Request a quote

Accelerator card

1. What we use VERSAL or older version-> definitely VERSAL based accelerator card are better but not in market yet ?! **VCK5000**
 - a) <https://www.alpha-data.com/product/adm-pa100/>
2. I think **RAL** and **Imperial** (Bristol?) have a card



Alveo U250 (Active)

\$6,995 per board

Lead Time: 6 weeks*

Part Number: A-U250-A64G-PQ-G

Add to Cart



Alveo U200 (Passive)

\$4,495 per board

Lead Time: 2 weeks*

Part Number: A-U200-P64G-PQ-G

Add to Cart




Alveo U200 (Active)

\$4,495 per board

Lead Time: 2 weeks*

Part Number: A-U200-A64G-PQ-G

Add to Cart



Add To BOM

Datasheet

Compare

A-U200-A64G-PQ-G

Xilinx

Xilinx U200 Alveo Data Center accelerator card: 18.6 INT8 T ... More

Avnet Manufacturer Part #: A-U200-A64G-PQ-G

1+

\$4,749

USD \$: 4,748.52

Min: 1 Mult: 1

-

1

+

ADD TO

Need More? Click To Quote

0 - In Stock

Additional Inventory

Factory lead time - 2 Weeks

Notify Me When In-Stock



Add To BOM

A-U250-P64G-PQ-G

Xilinx

Xilinx U250 Alveo Data Center accelerator card: 33.3 INT8 T ... More

1+

\$7,390

USD \$: 7,389.52

Min: 1 Mult: 1

-

1

+

ADD TO

To Quote


0 - In Stock

Additional Inventory

Factory lead time - 2 Weeks

Sort	Product Detail	Description	Quantity	Availability	Unit Price (GBP)	Ext. Price (GBP)	
------	----------------	-------------	----------	--------------	------------------	------------------	--

1



Mouser No:

Mfr. No:

Mfr.:

Customer No:

217-A-U250-A64G-PQ-G

A-U250-A64G-PQ-G

Xilinx

Customer No

Accelerator Cards

Alveo U250 Data Center Accelerator Card - Active

RoHS Compliant

QuickView


1

1 Dispatches Now

£6,585.38

£6,585.38

2



Mouser No:

Mfr. No:

Mfr.:

Customer No:

217-A-U200-A64G-PQ-G

A-U200-A64G-PQ-G

Xilinx

Customer No

Accelerator Cards

Alveo U200 Data Center Accelerator Card - Active

RoHS Compliant

QuickView

1

1 Dispatches Now

£4,231.77

£4,231.77

By submitting your order you agree to these terms and conditions.
For additional information on availability, click on the Mouser Part No.

Incoterms:DDP

All prices include duty and customs fees on select shipping methods.
Free shipping on most orders over £33 (GBP)

Merchandise:	£10,817.15
Shipping:	FedEx International Priority £0.00
	+ More Delivery Options
Subtotal:	£10,817.15