AMD 400G Adaptive SmartNIC SoC

Technology preview

Jaideep Dastidar

Co-authors: David Riddoch, Jason Moore, Steve Pope, Jim Wesselkamper

Hot Chips 2022

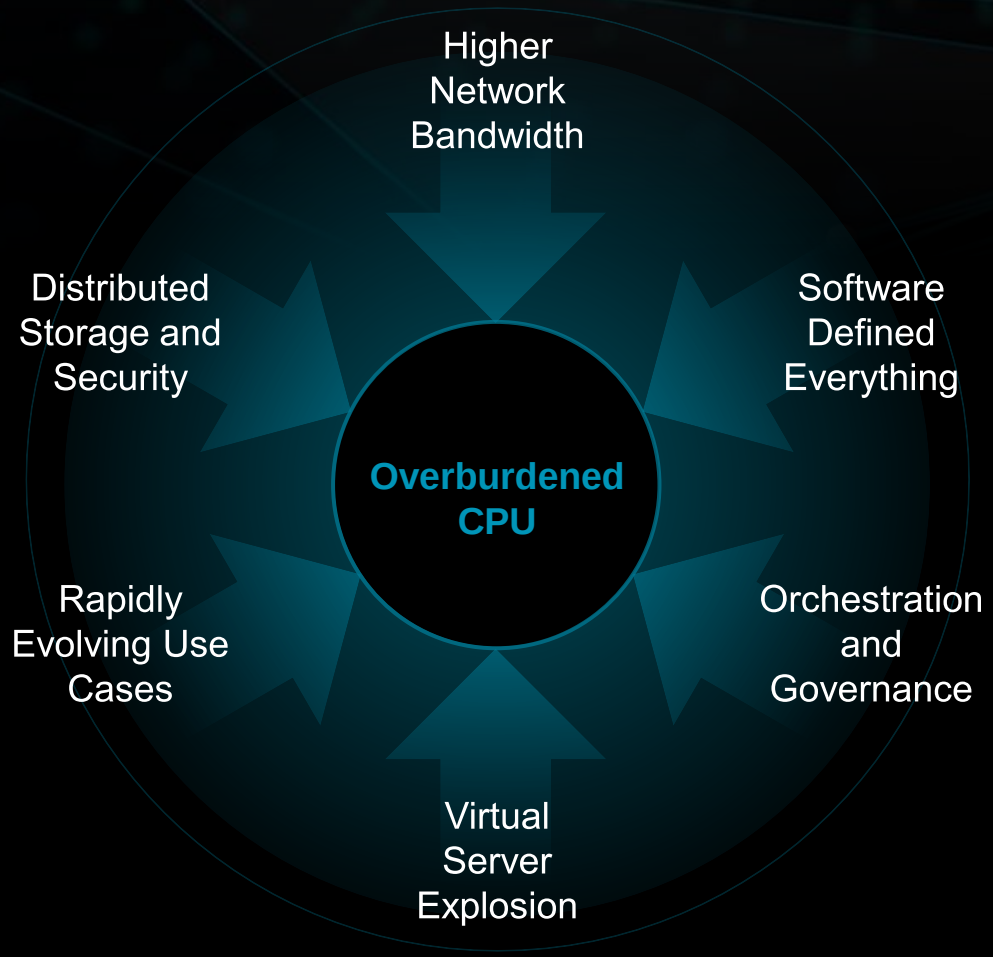
Agenda

Overview	• Motivating factors behind SmartNICs • AMD 400G Adaptive SmartNIC SoC Objectives
Solution – AMD 400G Adaptive SmartNIC SoC	• Offload Acceleration Blocks for Host, Network, and Embedded Processing • Adaptive Extensions • Pervasive Security and Confidential Computing
Use Case Examples	• OVS Offload • VirtIO Net Offload • Host-Attached Adaptive Network Offload Accelerator • Network Attached Adaptive NVMe Storage Endpoint
Summary	
Q&A	

AMD

Overview

Motivations for Adaptable, Intelligent Infrastructure



New Trends Expose Current Infrastructure Challenges

Intelligent Infrastructure for Next Gen Applications

Next Generation Bandwidth
400Gb Data Processing

Infrastructure Acceleration
Maximize CPU revenue potential

High Security
Multi-tenant Isolation and Protection

SmartNIC Solution Space

Proprietary

Customization

Global Hyperscalers

Other CSPs

SR-IOV OVS

HW Virtio.net / blk

IPsec Overlays SPDK

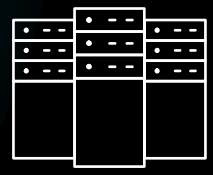
DPDK kTLS vDPA

Orchestration

Evolving Standards

Zero-touch

Telco/CDN



Enterprise

We are at an Architectural Inflection Point

Tug-of-War between

ASIC-like Fixed Function Logic
Frequent vs occasional functions

Embedded Processor Cores
Workload flexibility vs processing offload

FPGAs or Programmable Logic
Adapting new and existing functions

AMD 400G Adaptive SmartNIC SoC Objectives

Balanced Architecture

- Architect an adaptive SoC for the SmartNIC domain
- Strike the right balance between the heterogeneous elements
- Present a unified software view of the heterogeneous element SoC

Performance and Adaptability

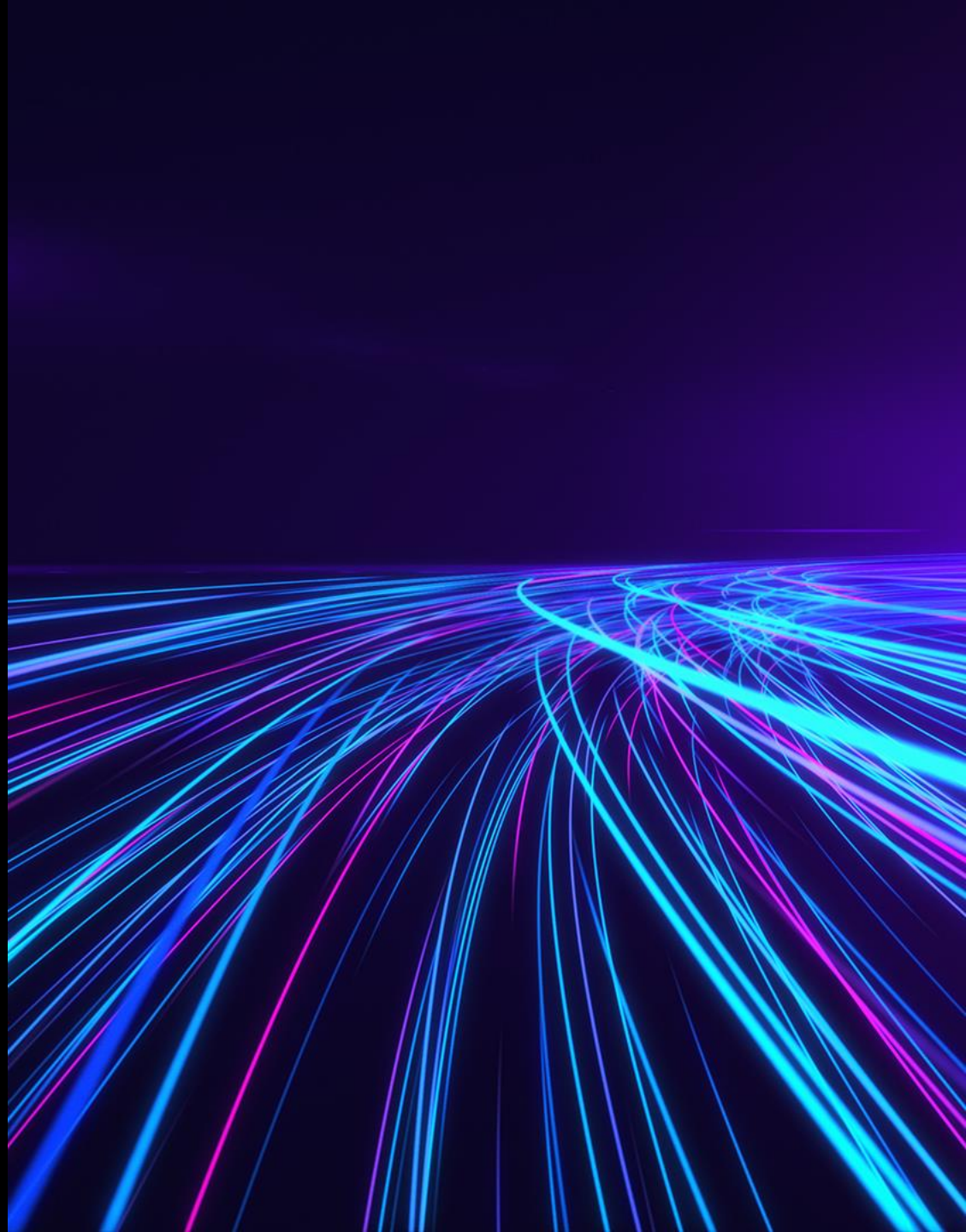
- Latest Interfaces/Speeds: 112G Ethernet, PCIe Gen5, CXL 2.0, LPDDR5/DDR5
- Adaptive Interface APIs for evolving workloads and customization

Security

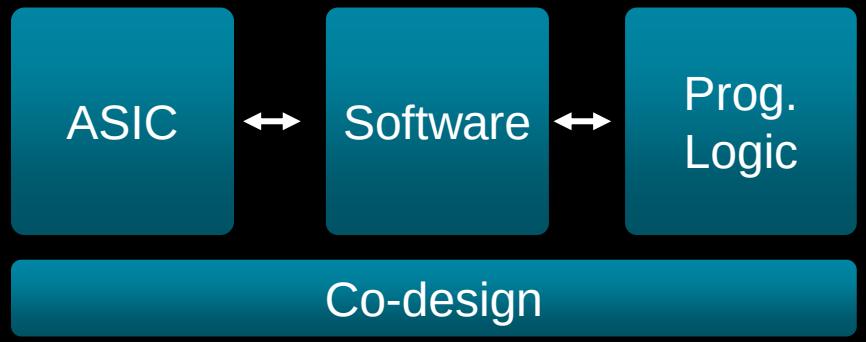
- Latest security standards: PCIe, OCP and NIST
- Symmetric and asymmetric crypto offload for data encryption and key exchange
- Physical and logical isolation for confidential compute

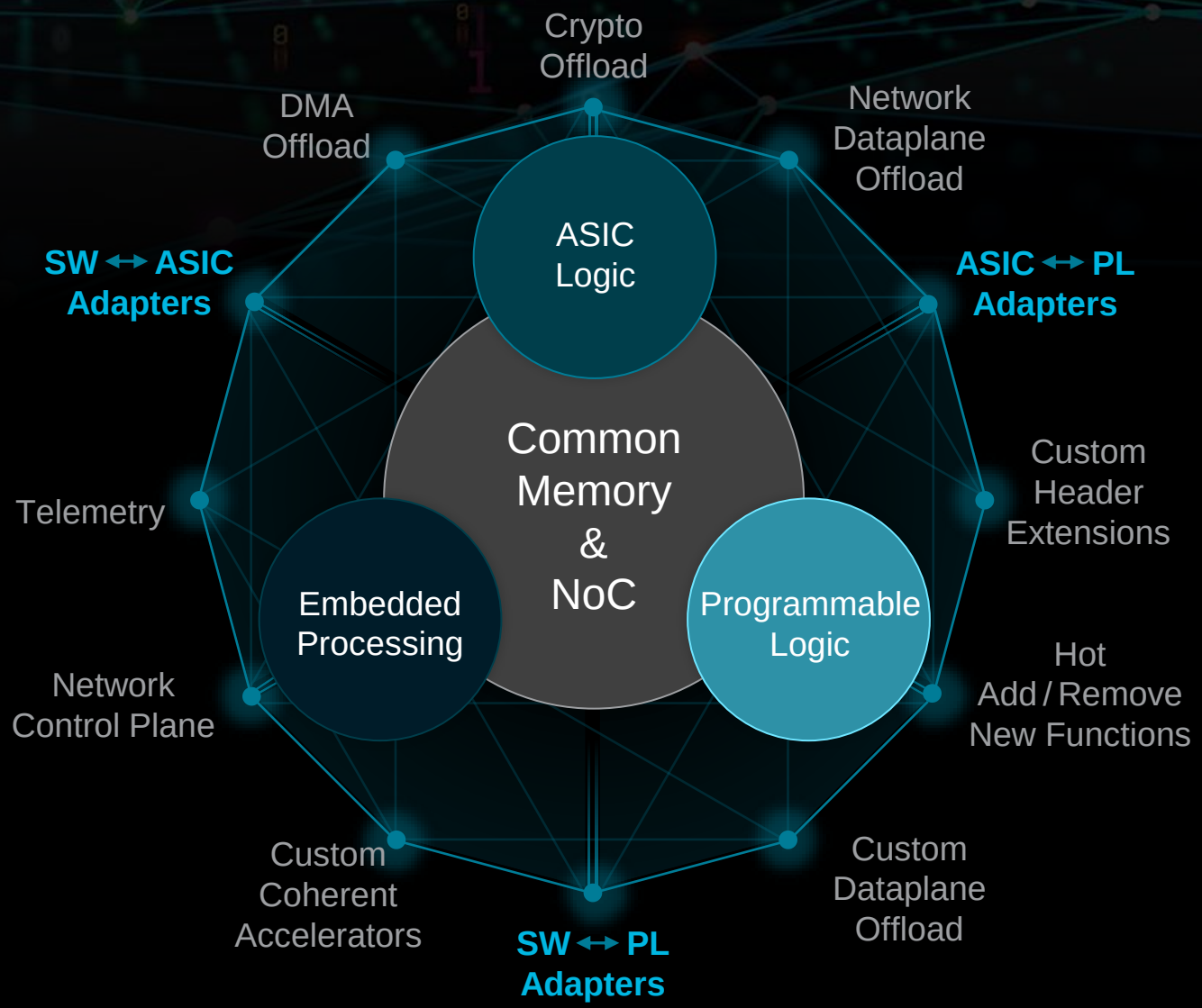
AMD 400G Adaptive SmartNIC SoC

Technology preview



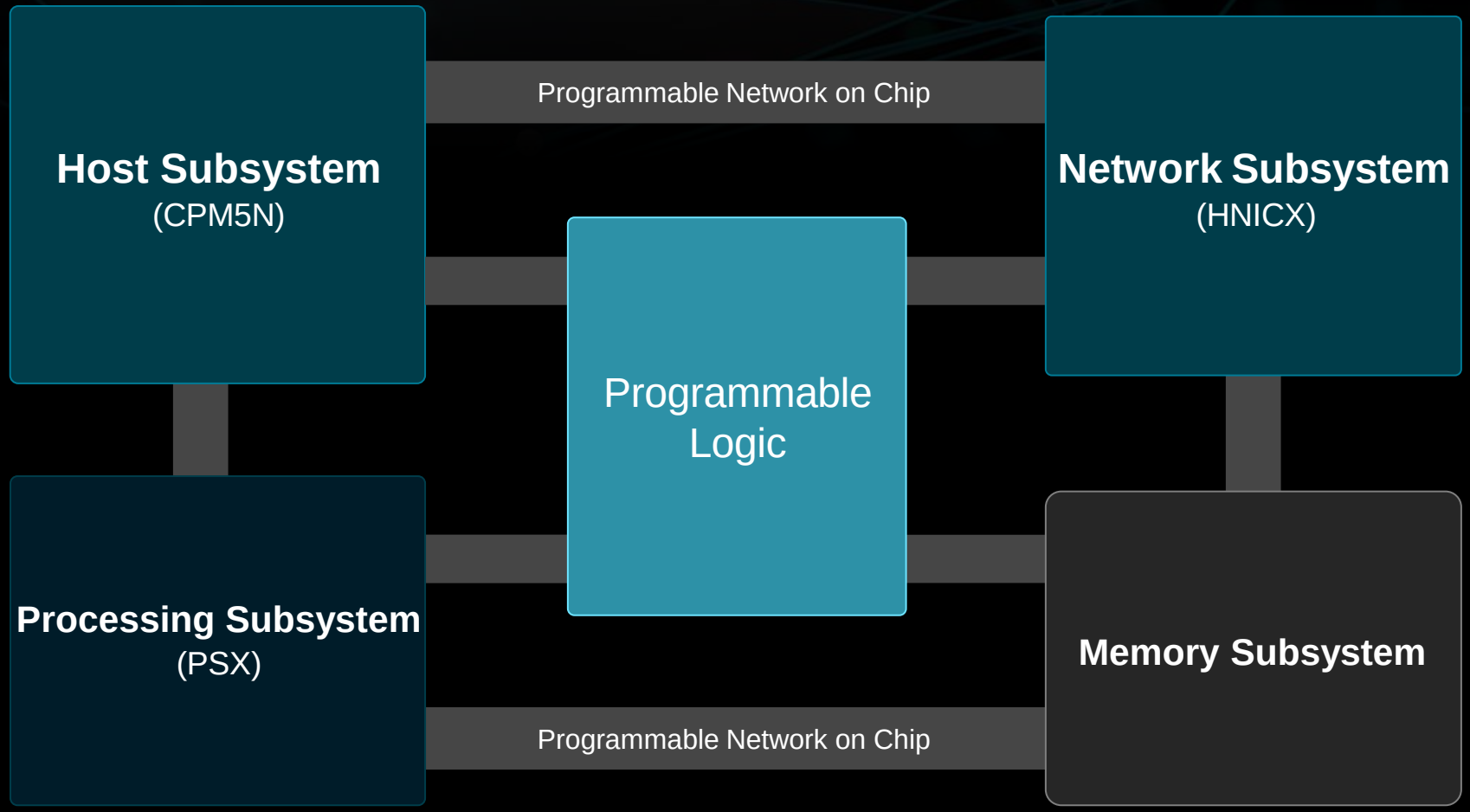
Solution





AMD

Tightly Coupled SmartNIC Blocks

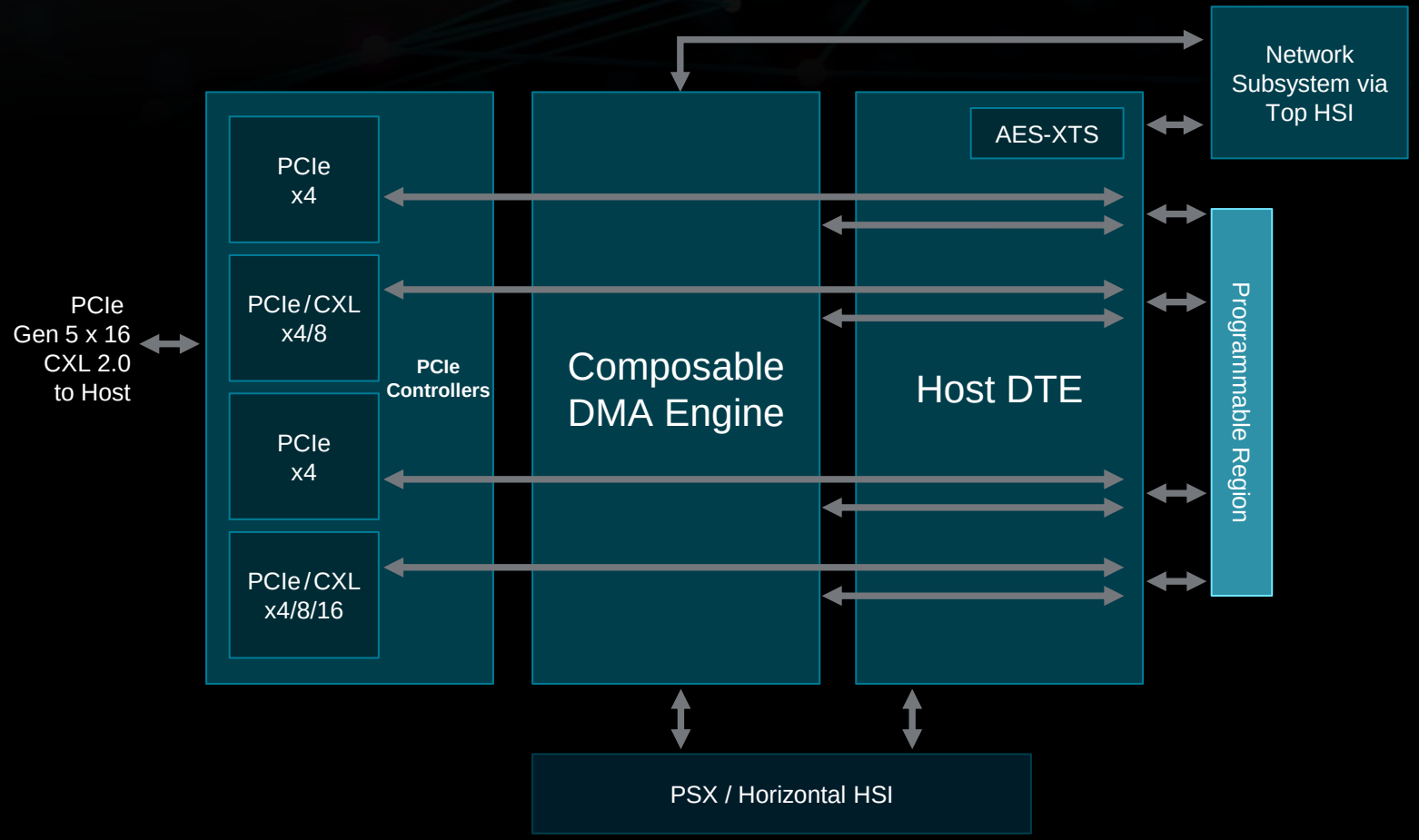


CPM5N: Host Subsystem Overview

PCIe Gen 5 or CXL 2.0 host connection

Composable DMA (CDx)

Host DTE (Data Transform Engine)



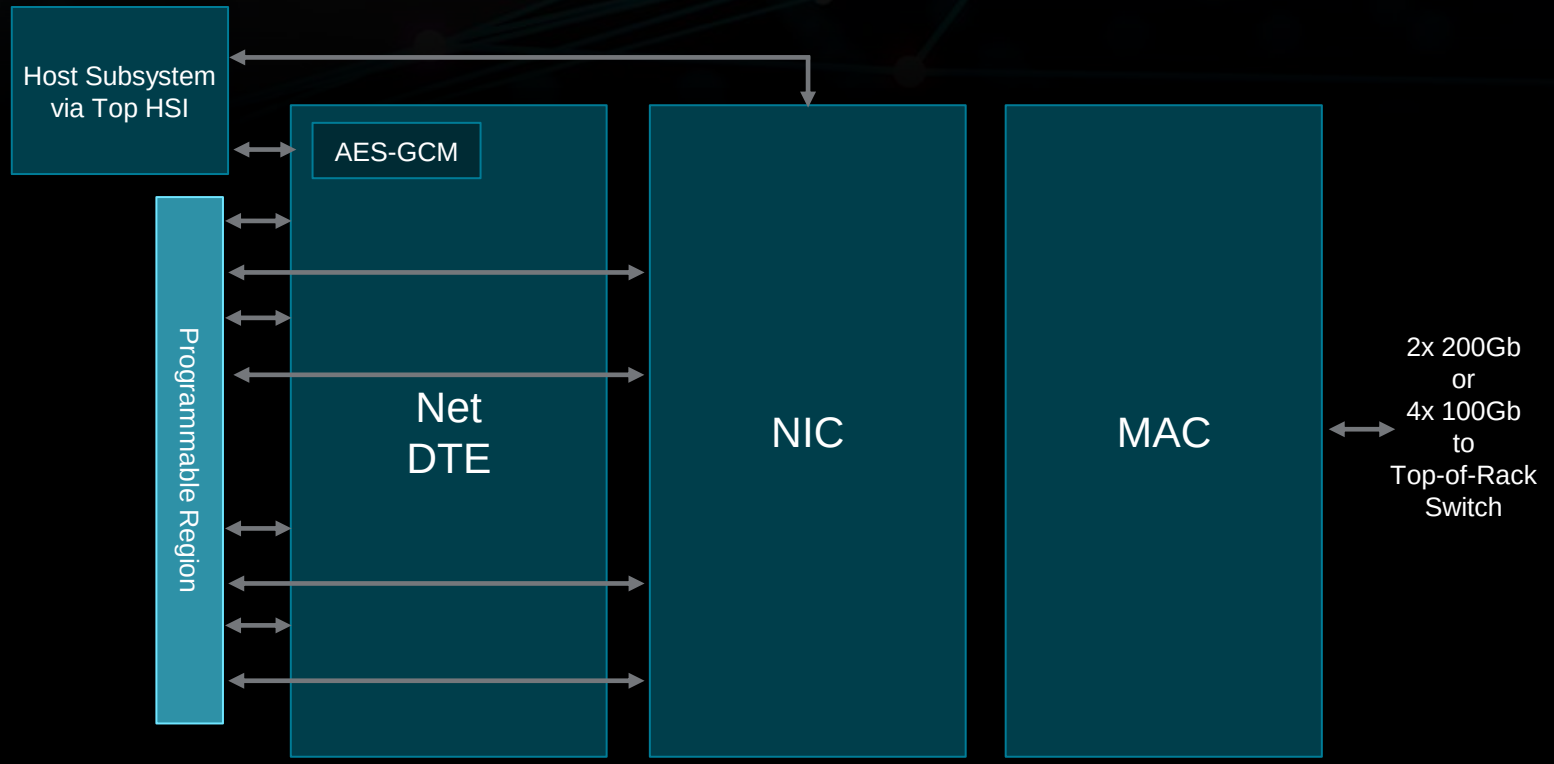
AMD

HNICX: Networking Subsystem Overview

Hardened NIC (HNIC)

Network DTE (Data Transform Engine)

MAC



Memory Subsystem Overview

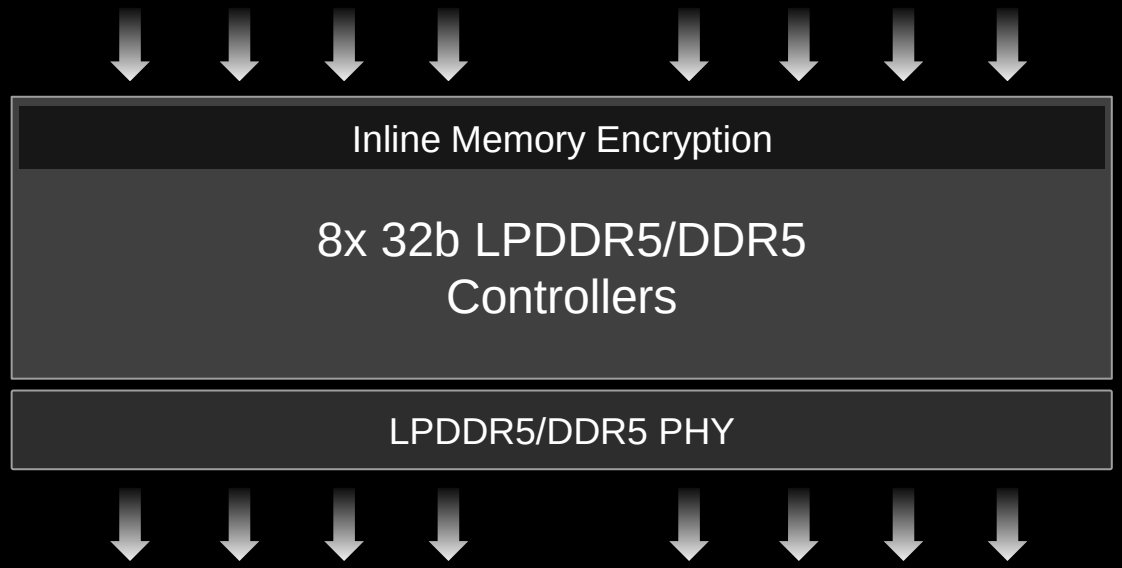
8x LPDDR5 / DDR5 Controllers

- 32-bit DDR5-5600 with ECC or
- 16-bit, 32-bit, or dual 160-bit LPDDR-6400 with ECC

- Inline Memory Encryption
  - AES-GCM / AES-XTS with DPA countermeasures

DDR Controller Bypass Mode for SCM

- PHY also supports other standards such as MIPI
- Custom PL Controllers for Storage Class Memory



AMD

Processing Subsystem (PSX) Overview

4 real-time cores arranged as Two 2-core Clusters w/ Tightly-Coupled Memory

LPD

Cortex-R52 32KB I/D$ 128KB TCM	**Cortex-R52** 32KB I/D$ 128KB TCM
LS/Split	LS/Split
Cortex-R52 32KB I/D$ 128KB TCM	**Cortex-R52** 32KB I/D$ 128KB TCM

Interconnect

1MB OCM	**DMA**
GbE (2)	**CANFD (2)**
USB2 (2)	**UART (2)**
SPI (2)	**I2C+I3C (2)**
MIO	**Timers**
PSM PS Mgmt	**Debug** PS CoreSight

FPD

Cortex-A78AE 64KB I$/D$ 512KB L2$	**Cortex-A78AE** 64KB I$/D$ 512KB L2$	**Cortex-A78AE** 64KB I$/D$ 512KB L2$	**Cortex-A78AE** 64KB I$/D$ 512KB L2$
LS/Split	LS/Split	LS/Split	LS/Split
Cortex-A78AE 64KB I$/D$ 512KB L2$	**Cortex-A78AE** 64KB I$/D$ 512KB L2$	**Cortex-A78AE** 64KB I$/D$ 512KB L2$	**Cortex-A78AE** 64KB I$/D$ 512KB L2$
L3 Cache 2MB	L3 Cache 2MB	L3 Cache 2MB	L3 Cache 2MB

Coherency Mesh Network

LLC 16MB

Core Sight	**System MMU**	**Interconnect**	**GIC**	**PKI**

High Performance Domain

16 Application cores arranged as Four 4-core Clusters w/Crypto Extensions

Public Key Infrastructure
TLS 1.3
Offload Acceleration

AMD

Tight Coupling between CPM5N ↔ PSX

256Gbps Host MMIO to Embedded Cores Data Flow w/ Cache Injection via PCIe/CXL.io Endpoint

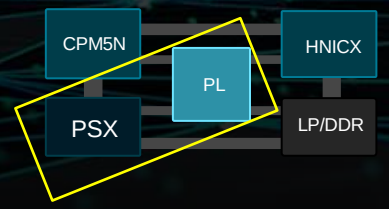
Host Subsystem
(CPM5N)

256Gbps Embedded Cores to Host Data/Doorbells

Processing Subsystem
(PSX)

AMD

Tight Coupling between Prog. Logic ↔ PSX



512Gbps Near-PL Fabric I/O Coherent access w/ Cache Injection and Maintenance

256Gbps Embedded Core MMIO access to Near-PL Memory

Programmable Logic

Processing Subsystem (PSX)

256Gbps Cache Coherent Near-PL Fabric access including Cache Coherent ATOMICS

AMD

Security and Confidential Compute Across Domain-specific Accelerators

Pervasive Security

Layers of Security

CHECK

Secure Boot, Key Exchange and Attestation

AMD keys, SmartNIC owner keys, and SmartNIC tenant keys

PROTECT

Memory and Peripheral Protection Units

SoC-wide hardware firewalls for tenants and accelerators

SHIELD

Secure Data in Flight, at Rest, in Use

Encryption for data, host/network interfaces, and DRAM

Secure monitors for external, internal, physical attacks

Crypto engines with DPA countermeasure protection

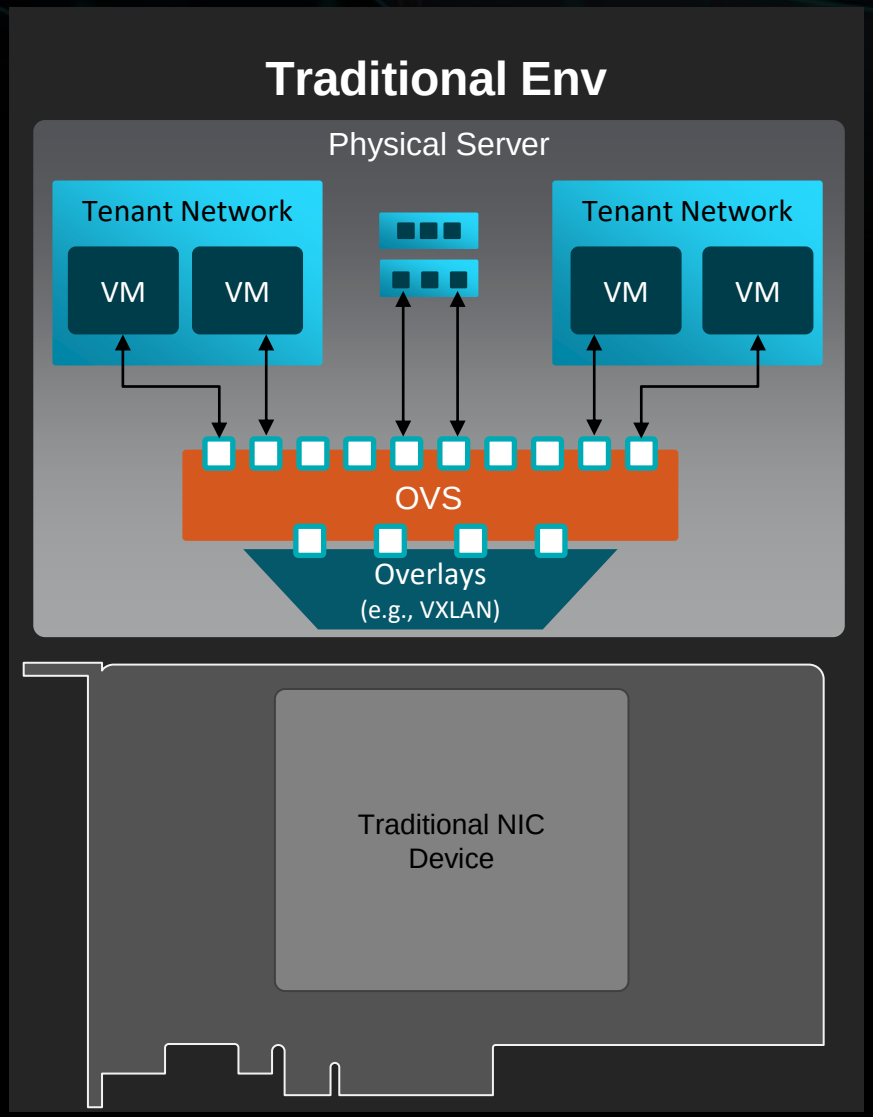
AMD

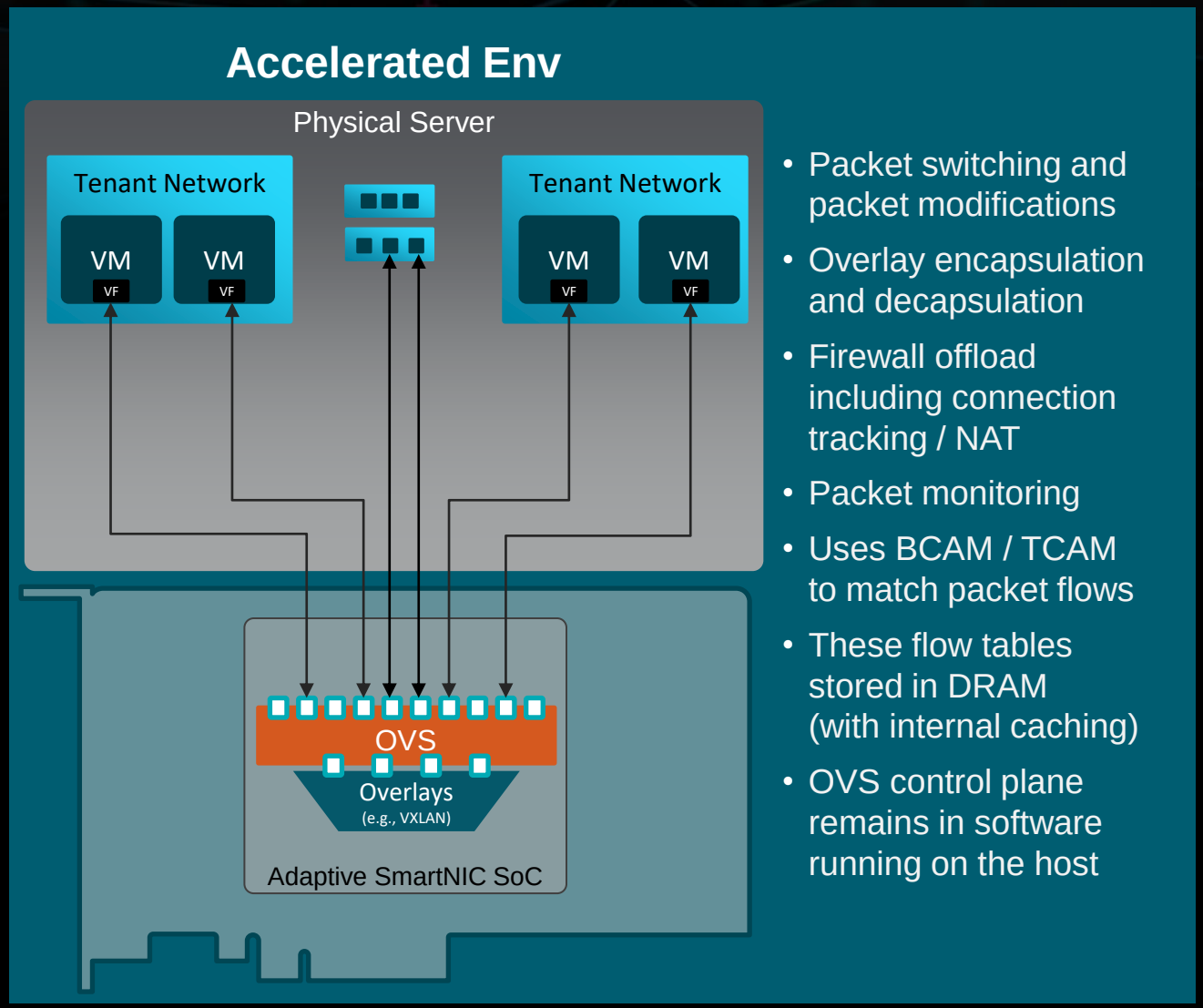
Use Case Examples

Use-cases


Network Attached
Host Attached
Host Attached
Storage Attached

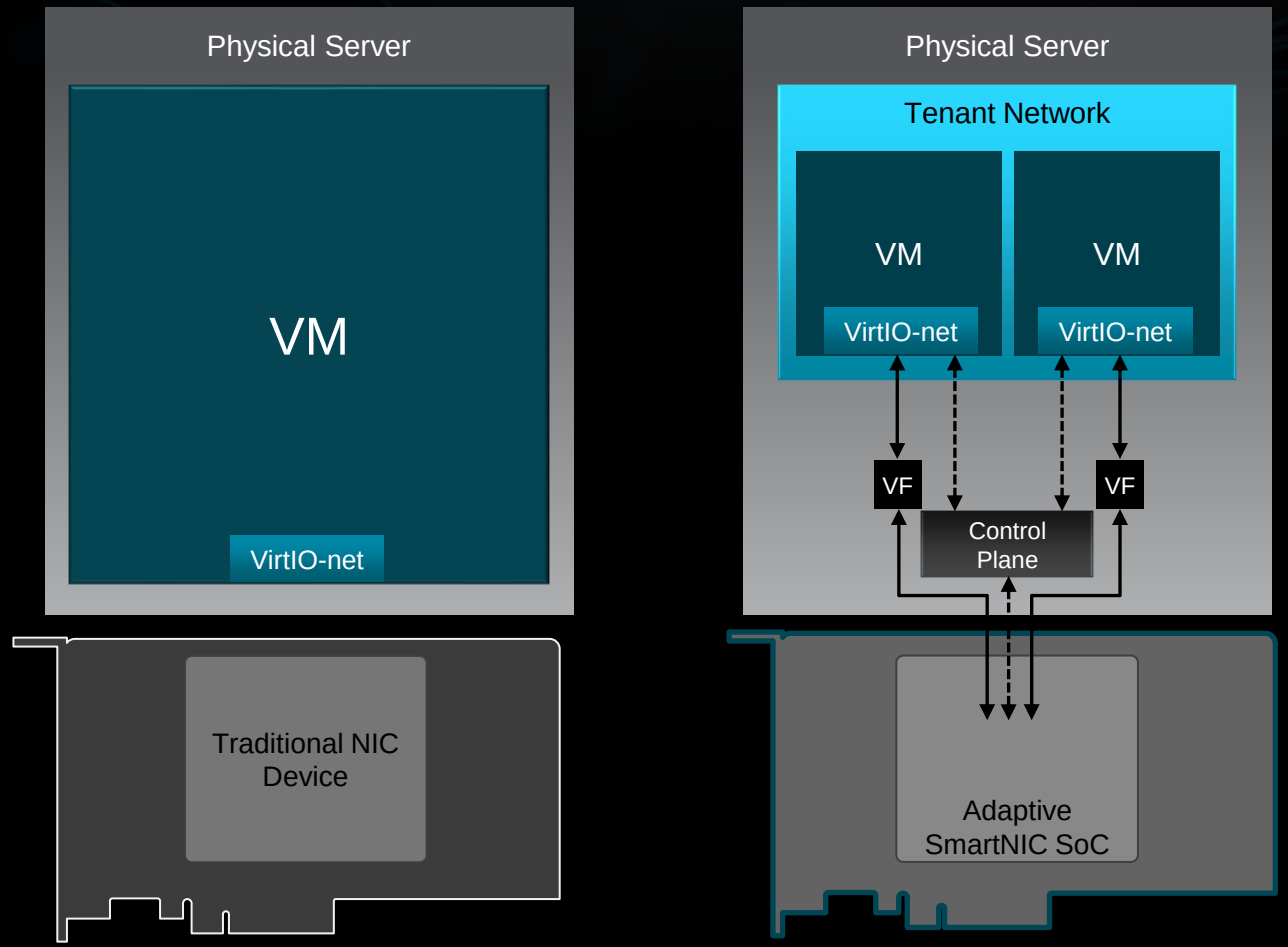
Full OVS Offload





- Packet switching and packet modifications
- Overlay encapsulation and decapsulation
- Firewall offload including connection tracking / NAT
- Packet monitoring
- Uses BCAM / TCAM to match packet flows
- These flow tables stored in DRAM (with internal caching)
- OVS control plane remains in software running on the host

VirtIO-net Support



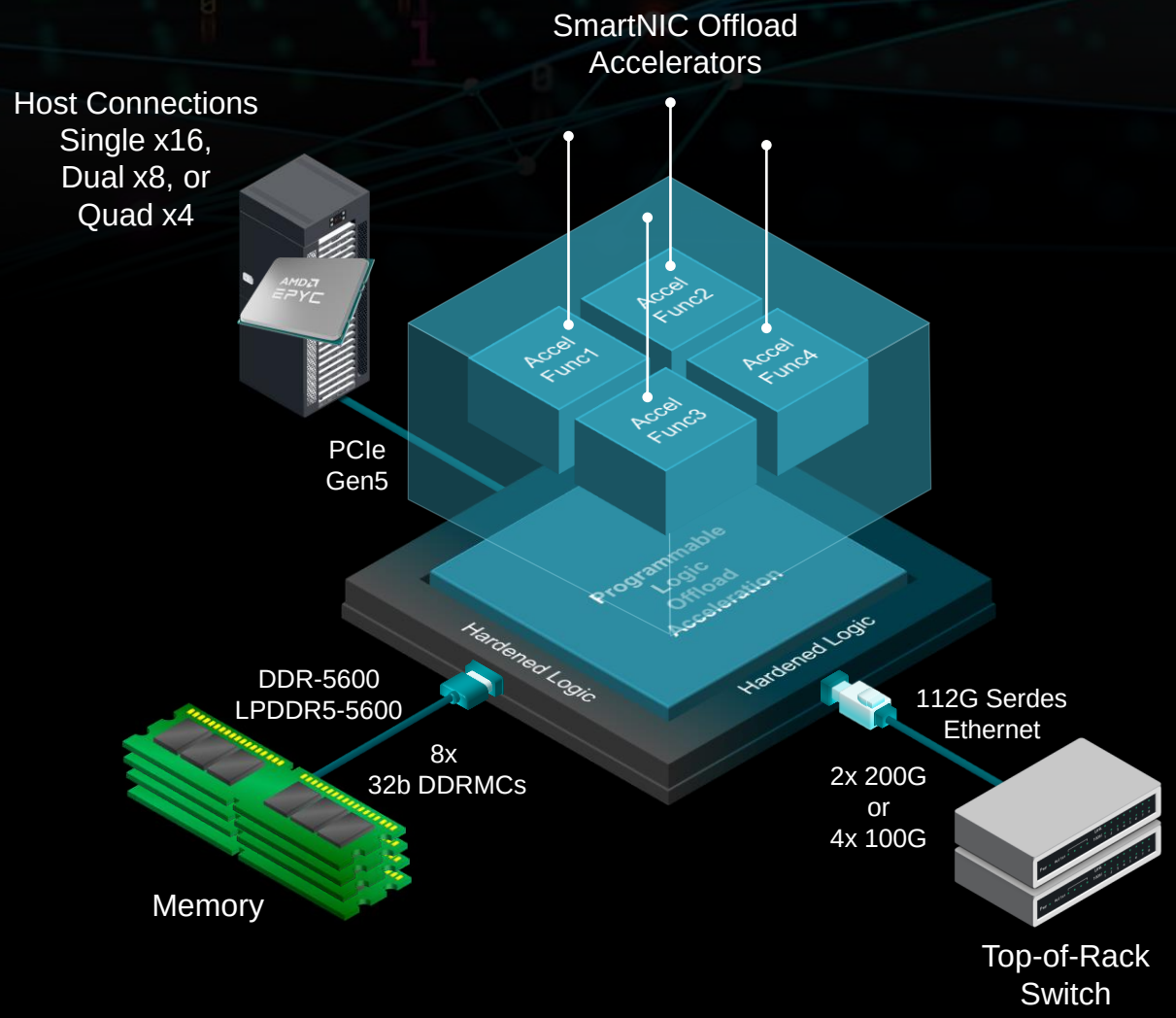
Device Natively Supports VirtIO Queues

Two Deployment Models

- Bare Metal VirtIO
- vDPA
  (Virtual Data Path Acceleration)

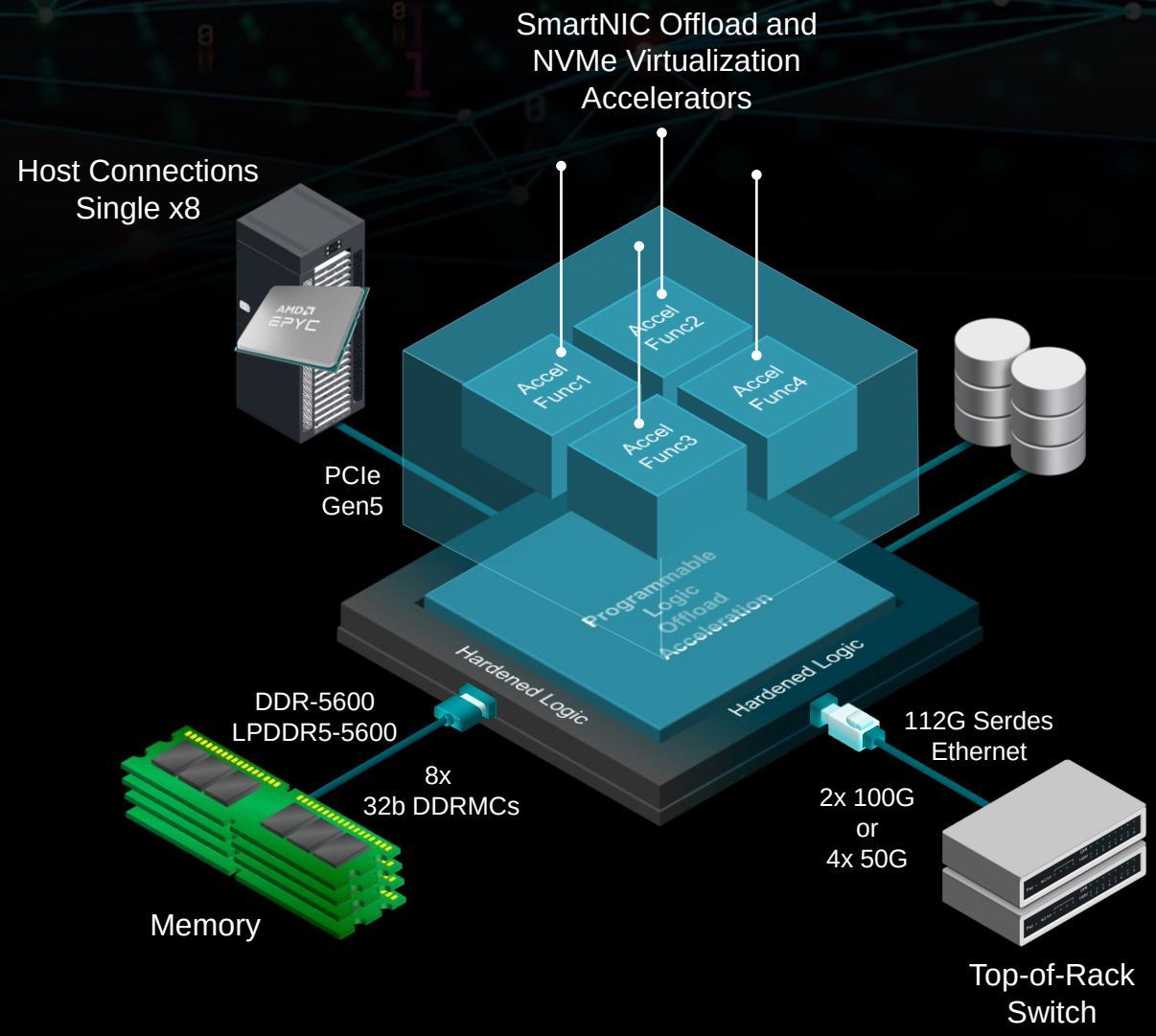
PCIe Endpoint 400G SmartNIC w/ with Host Storage and Network Offload

- Standard Ethernet Packet Flow
- Custom Header, Encapsulation, or Encryption Flow
- Host Connection with Advanced Switch Hot Plug



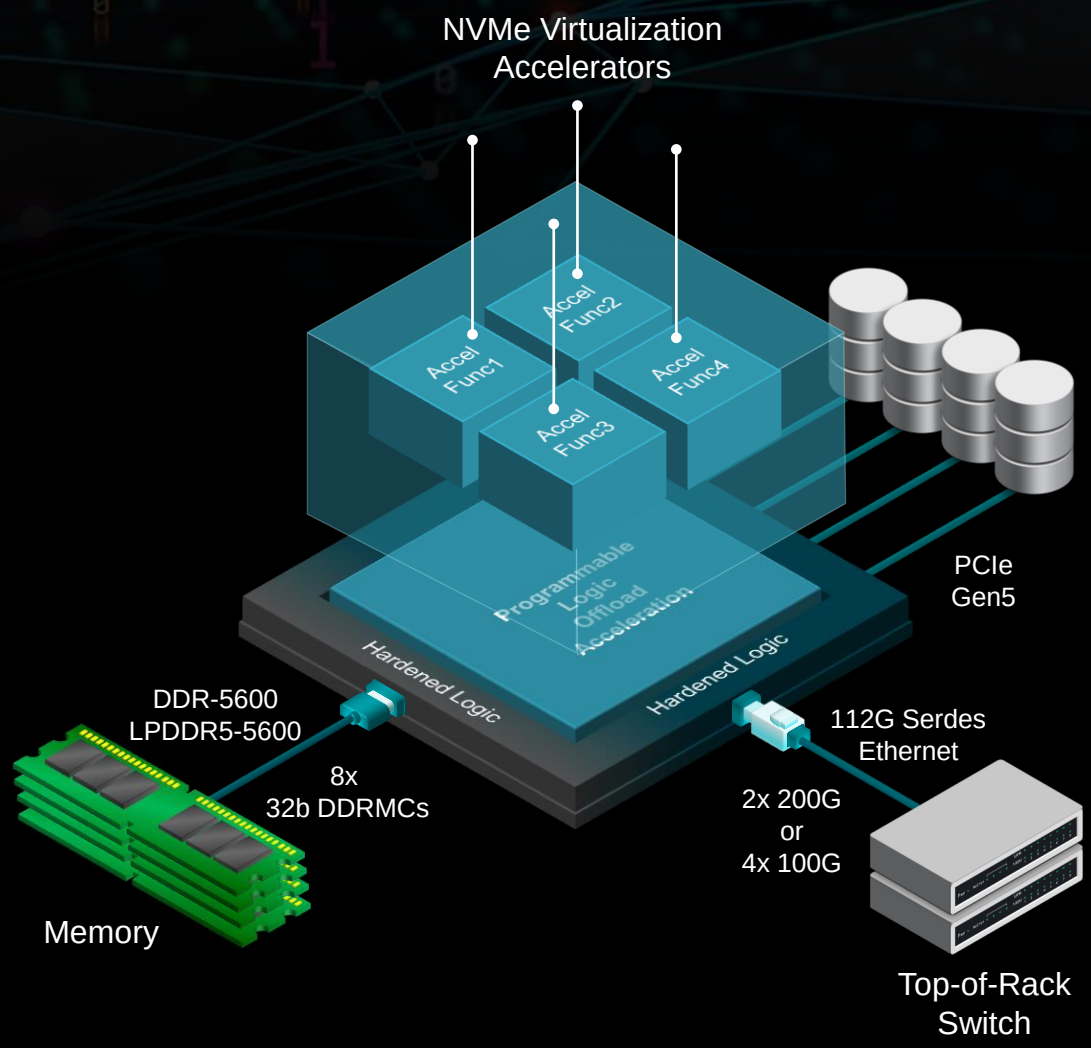
PCIe Endpoint 200G SmartNIC w/ 2 Gen5 x4 NVMe SSDs

- Host endpoint connection NVMe virtualization
- SmartNIC Root-Port connections for dual-NVMe SSDs w/ AES-XTS offload acceleration
- 2x 100G or 4x 50G network connection



PCIe Root Complex 400G NVMe Storage Node

- Network-attached storage appliance
- SmartNIC Root-Port connections for quad-NVMe SSDs
  w/ AES-XTS offload acceleration
- 2x 200G or 4x 100G remote storage connections



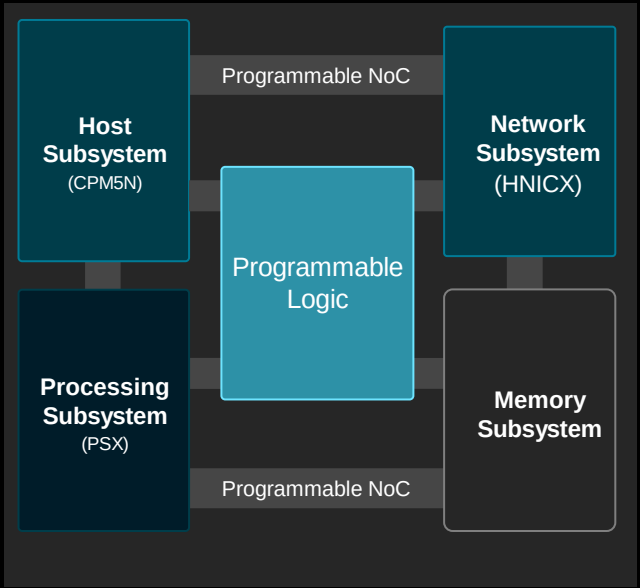
AMD

Summary

Performance & Resources

Manufacturing Technology

Process Technology	TSMC 7nm FF



Architectural Resources

Host subsystem	PCIe Gen5 x16 PCIe CMA/DOE/IDE CXL v2.0
Network subsystem	2x 200G Ethernet Virtual switch offload Stateless offloads
Application Processing Subsystem	16x A78-AE 24MB Shared Cache
Real-time Processing Subsystem	4x ARM R52
PKI (TLS1.3 Offload Engine)	Yes
DDR subsystem	8x 32b LPDDR5/DDR5 ECC Inline encryption 1.638Tb/s LPDDR5 BW

SmartNIC SoC Performance†

Programmable logic packet rate[1]	400Mpps Ingress + 400Mpps Egress
Host ⇔ Network packet rate[2]	400Mpps RX + 400Mpps TX
Full Virtio.NET offload BW[3]	400Gbps RX + 400Gbps TX
AES-XTS offload BW[4]	800Gb/s
AES-GCM offload BW[5]	800Gb/s

† See Endnotes

Thank You

To the worldwide AMD AECG team and all the other AMD teams who made the AMD 400G Adaptive SmartNIC possible

Their creativity, ingenuity, and unwavering dedication to the project has made for a unique Adaptive SmartNIC DSA

Endnotes

† SmartNIC SoC performance claims are based on simulation results of pre-silicon models with AMD traffic generators used to emulate external port traffic behavior as of August 23, 2022. Actual performance of production silicon may vary.

1. Interface packet rate between Programmable logic and the Network subsystem
2. Sustained TX and RX packet rate for 64B packets
3. Sustained TX and RX BW with 1500B payload per packet
4. Sustained BW for 1KB packets; Encrypt-only, Decrypt-only, or 400Gb/s Encrypt + 400Gb/s Decrypt with Network Interface configured to 2x200G Ethernet ports
5. Sustained BW for 1KB packets; 400Gb/s Encrypt + 400Gb/s Decrypt with Network Interface configured to 2x200G Ethernet ports
6. #4 and #5 were also measured concurrently enabled

Disclaimer and Attribution

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors.

The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

ATTRIBUTION

AMD

together we advance_

Thank You for Participating