SNIA NSF | NETWORKING STORAGE

xPU Accelerator Offload Functions

Live Webcast

June 29, 2022

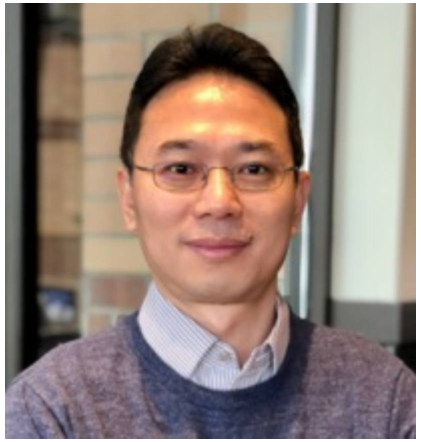
11:00 am PT / 2:00 pm ET

Today's Presenters

David McIntyre
Director, Product Planning
Samsung

Mario Baldi
Fellow
Pensando Systems

John Kim
SNIA NSF Chair
NVIDIA



Yadong Li
Principal Engineer
Intel

Dr. Joseph L. White
Fellow at Dell Technologies

SNIA NSF | NETWORKING STORAGE

SNIA-at-a-Glance

180
industry leading organizations

2,500
active contributing members

50,000
IT end users & storage pros worldwide

Learn more: **snia.org/technical** **@SNIA**

Ethernet, Fibre Channel, InfiniBand®

iSCSI, NVMe-oF™, NFS, SMB

Virtualized, HCI, Software-defined Storage

Storage Protocols (block, file, object)

Securing Data

Technologies We Cover

SNIA. NSF | NETWORKING STORAGE

SNIA Legal Notice

- The material contained in this presentation is copyrighted by the SNIA unless otherwise noted.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced in their entirety without modification
  - The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA.
- Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

  NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.

This is a 3-Part Series!

- Our first webcast: “SmartNICs to xPUs: Why is the Use of Accelerators Accelerating?”
- We covered:
  - What is an xPU?
  - Trends and workloads
  - Deployment and solutions
  - Market landscape
- Watch on demand at: https://bit.ly/SNIAxPU1

SNIA. NSF | NETWORKING STORAGE

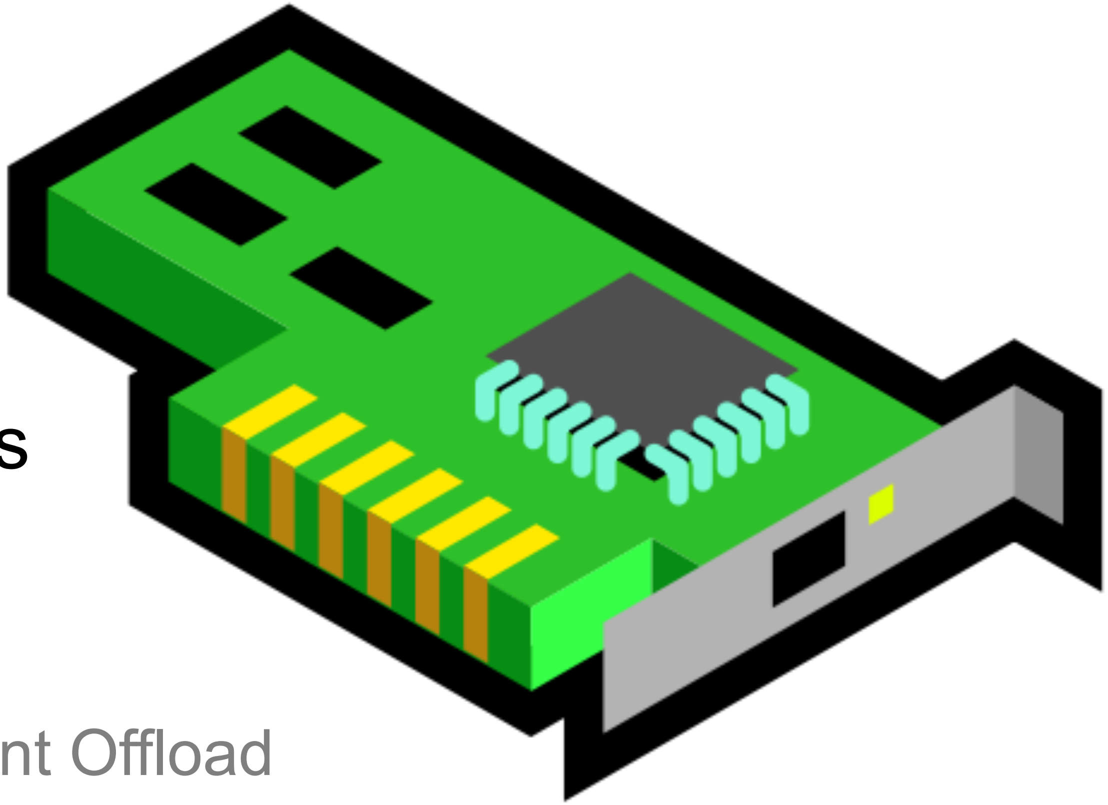
Agenda

- Network Offloads
- Security Offloads
- Storage Offloads
- Compute Offloads

Network Offloads

Mario Baldi, Fellow, Pensando Systems

NIC Functionality



- Host interfacing – driver
- Sending and receiving Ethernet frames
- TSO/LSO – LRO
  - TCP Segmentation Offload/Large Segment Offload
    Large Send Offload
  - Large Receive Offload
- QoS: multiple priorities and send/receive queues

Support for Virtualized Environments

Switching

- Virtual switch offload
  - SR-IOV support
- Possibly hardware switch

QoS

- Multiple priorities/queues
  - Remote traffic
  - Local traffic
- Hardware support for complex scheduling not available in host

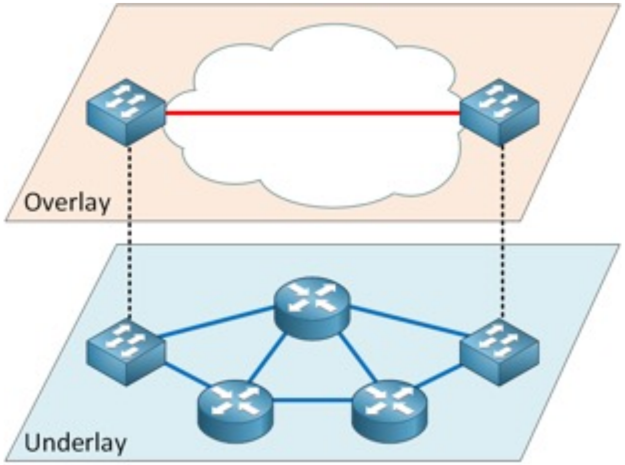
VLANs

Microsegmentation

Policies

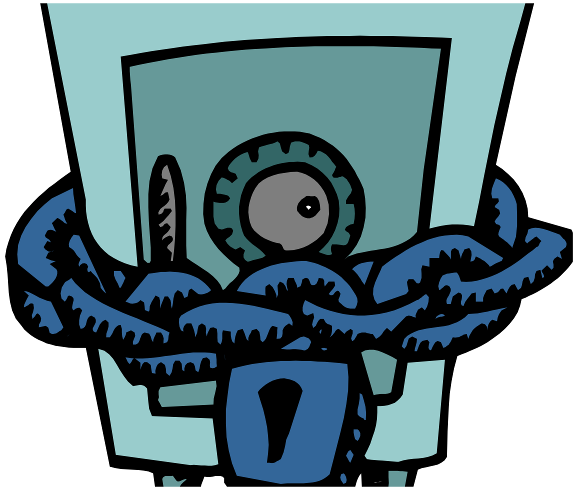
Gateway Functionality

Full-fledged Layer 3 endpoint



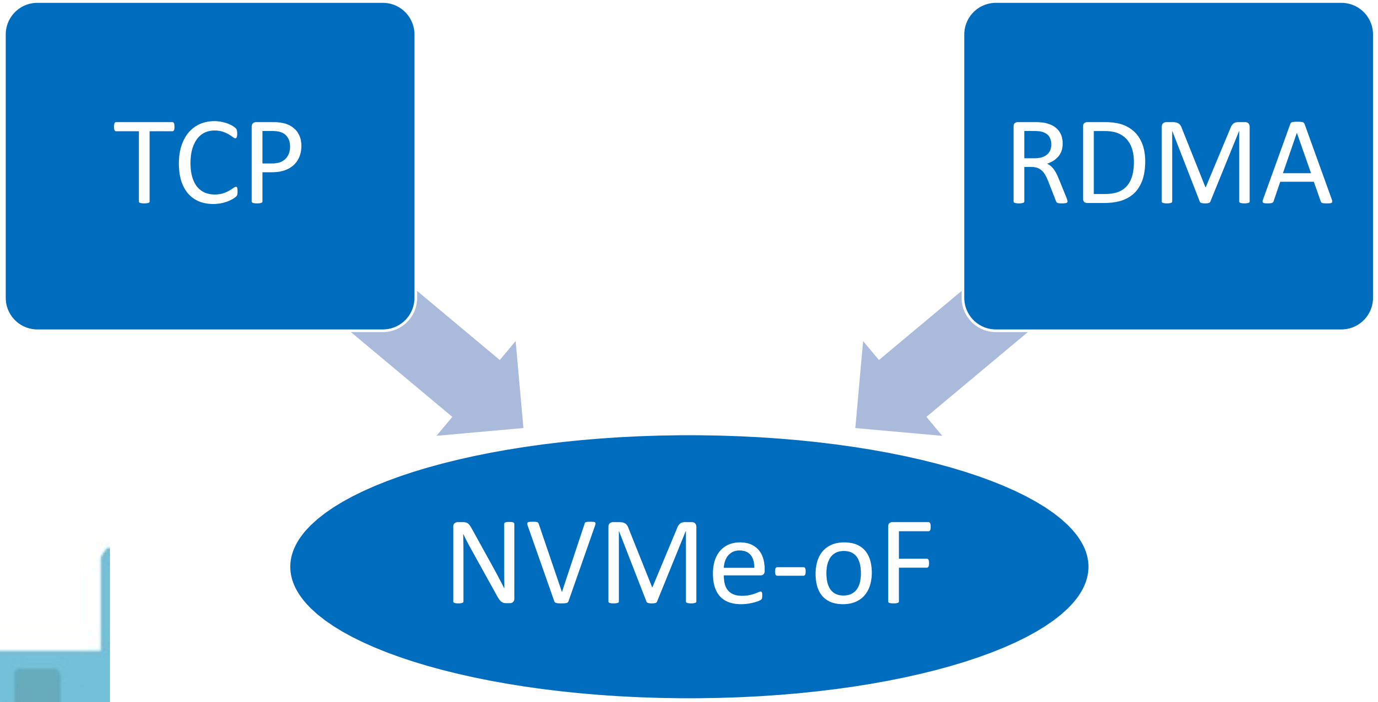
Overlay

IPsec



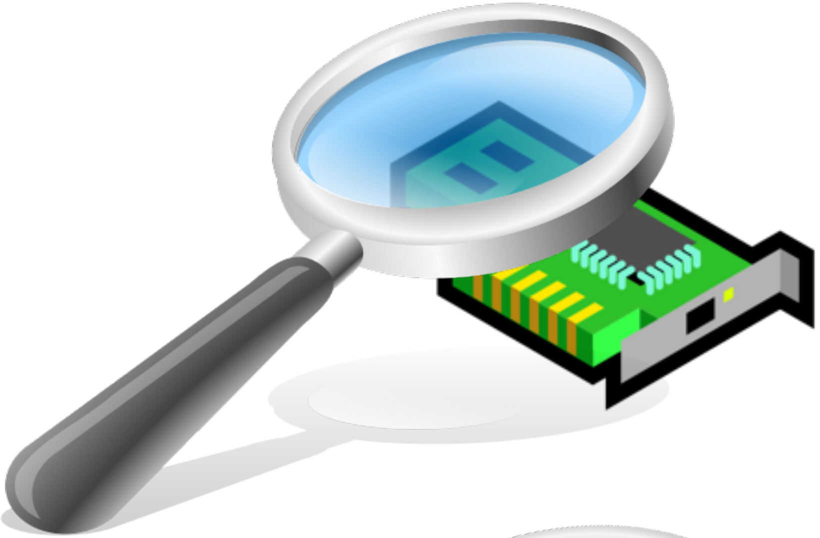
TLS Offload

Full Stack Offload



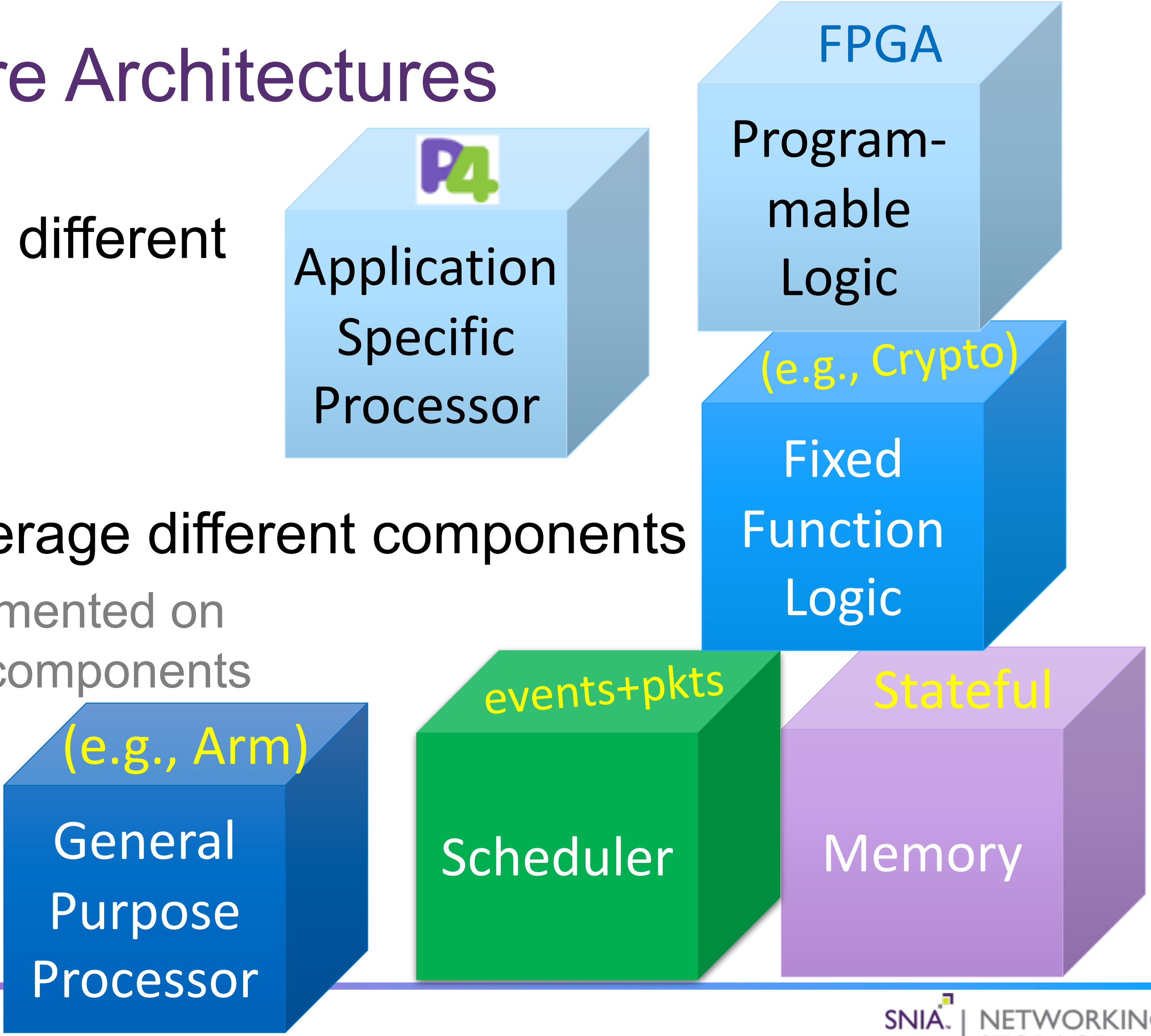
Observability

- Traditional and streaming telemetry
  - Flow based telemetry
    - NetFlow/IPFIX (IP Flow Information Export)
  - Packet level streaming telemetry
    - INT (In-band Network Telemetry)/In-situ OAM
- Packet capture and mirroring
  - ERSPAN
- Local monitoring, aggregation, and alarm generation



Hardware and Software Architectures

- Different xPUs may include different components
  - Design choice
  - Addressed use cases
- Different solutions may leverage different components
  - Same function can be implemented on (a combination of) different components
    - Design choice
    - Performance
    - Resource optimization

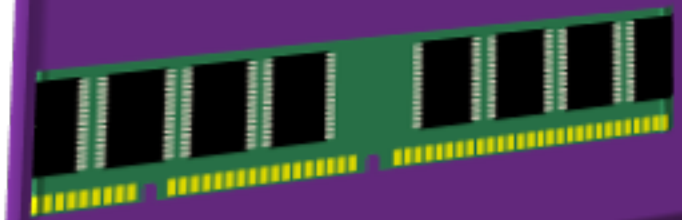


Network Offload Use Cases

Multiple functions

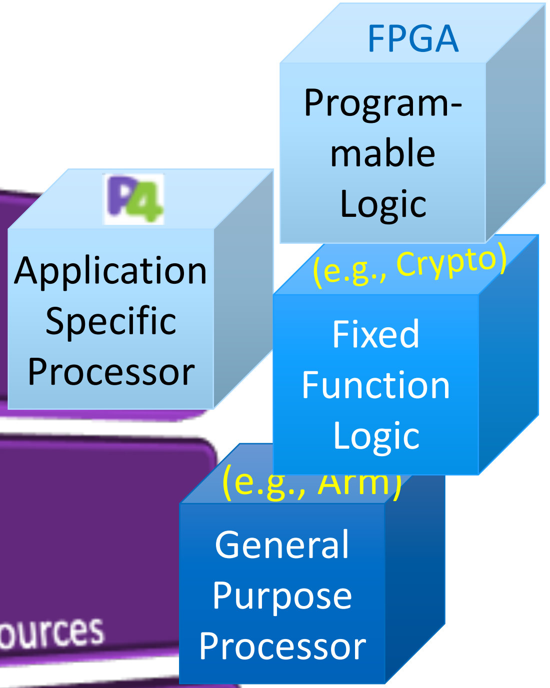
xPU deployment is a journey

- Lift&shift
- Move to increase performance and free resources



Performance at scale → DDR

- Stateful services
- Writeable data structures (e.g., tables)



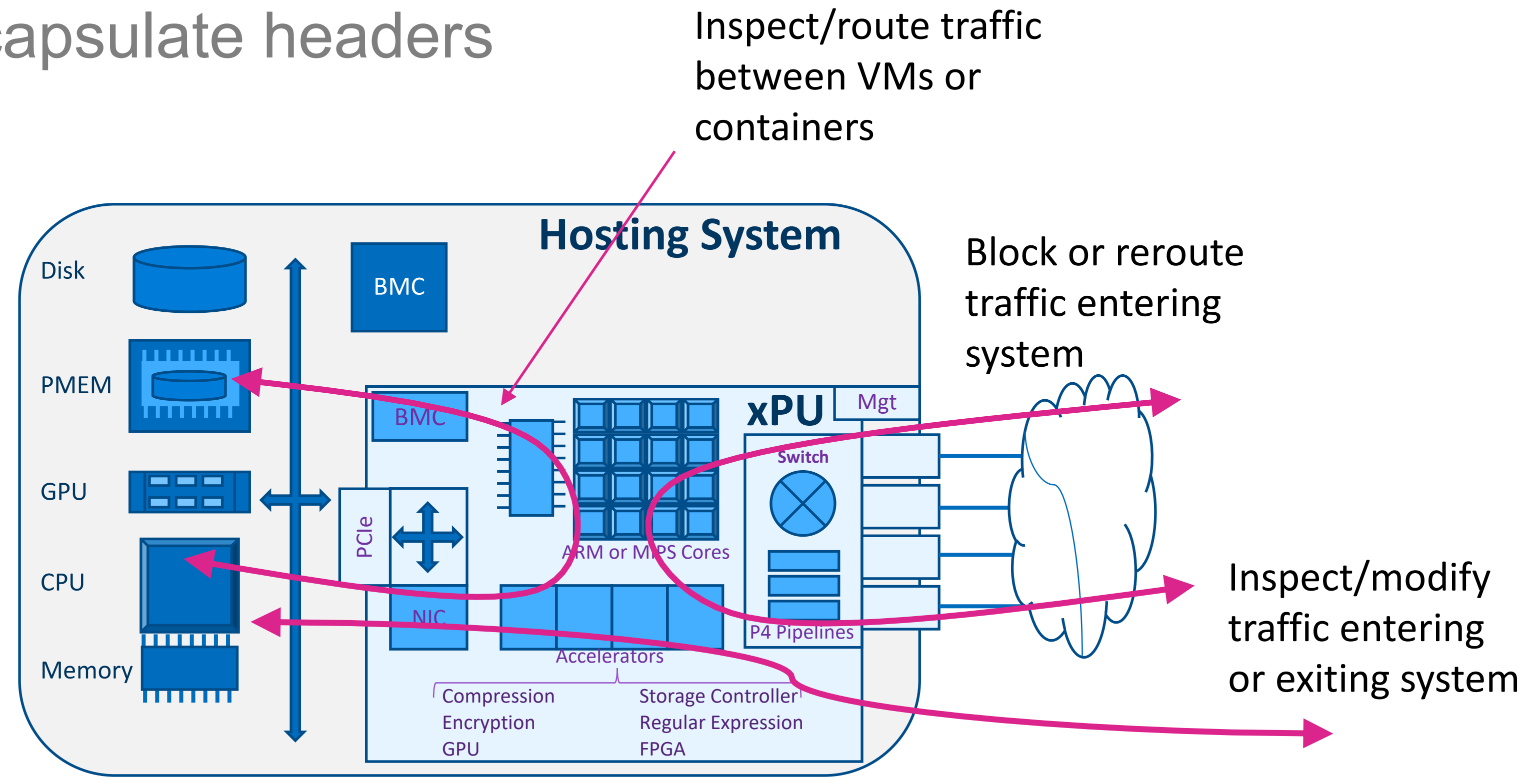
SNIA. NSF | NETWORKING STORAGE

Security Offloads

John Kim, Director of Storage Marketing, NVIDIA

Network Traffic Management Enables Security

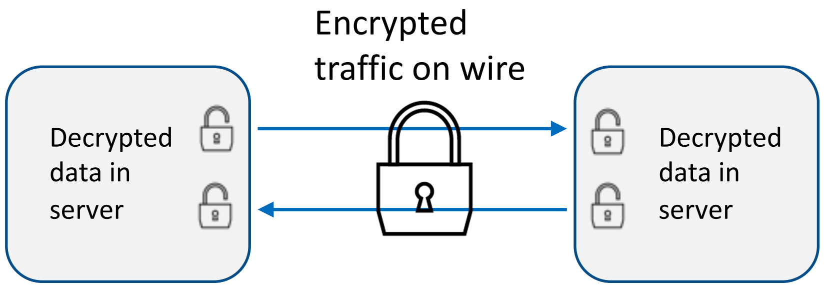
- Route, inspect, and/or modify packets for security reasons
  - Report, drop or redirect packets
  - Rewrite or encapsulate/decapsulate headers
- Support security software
  - Next generation firewall
  - Load balancer/WAF
  - IDS/IPS
  - Microsegmentation

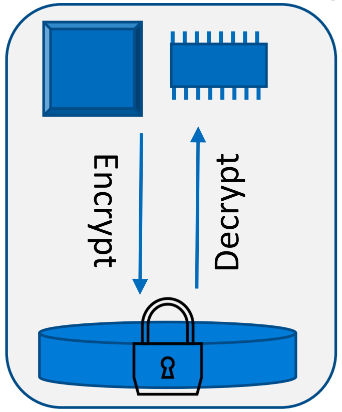


SNIA. NSF | NETWORKING STORAGE

Encryption and Key Management

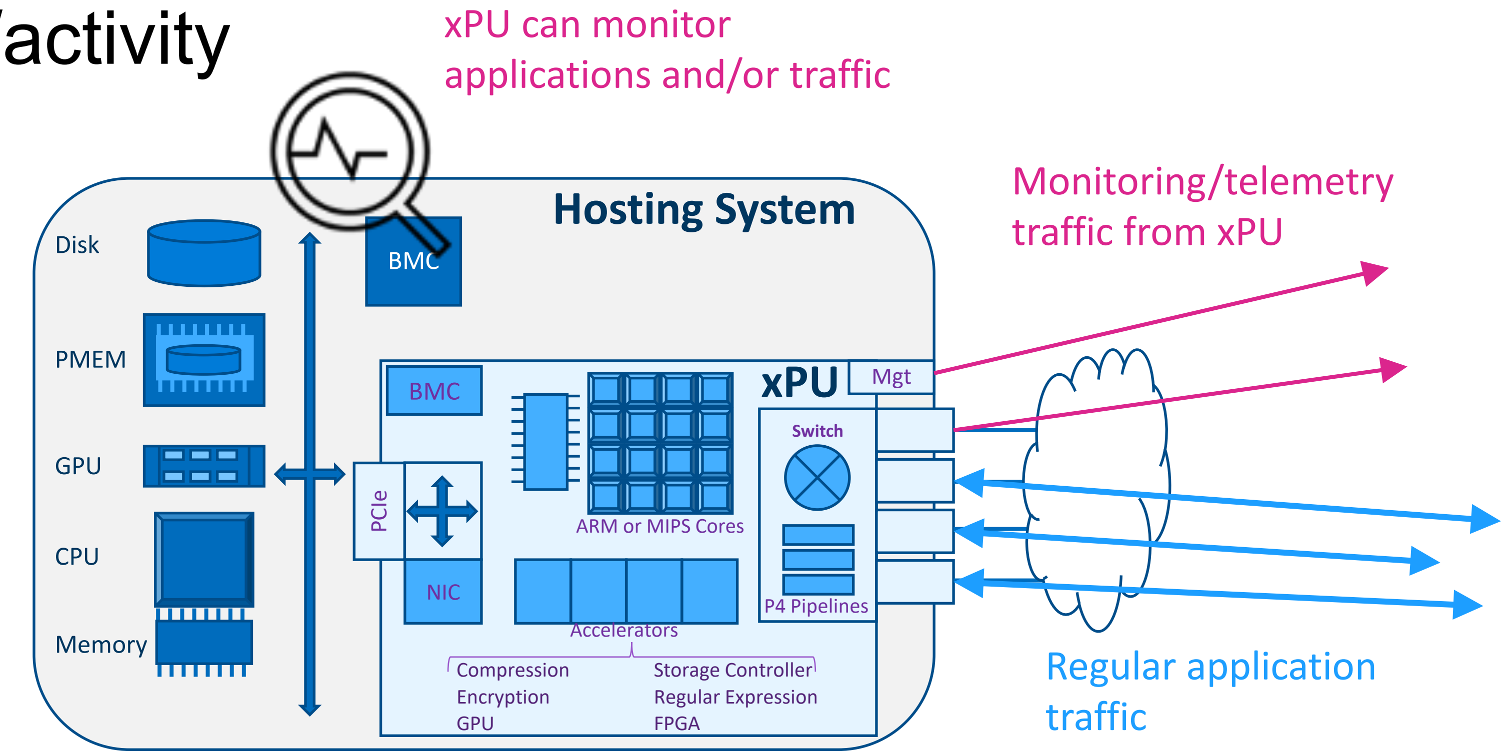
- Encrypt and decrypt network traffic
  - Faster – allows greater use of encryption
  - Different levels: MACsec, IPsec, TLS
  - Decryption Enables inspection and other security functions
- Encrypt/decrypt storage
- Key Management
  - Accelerate PKI
  - Random number generator





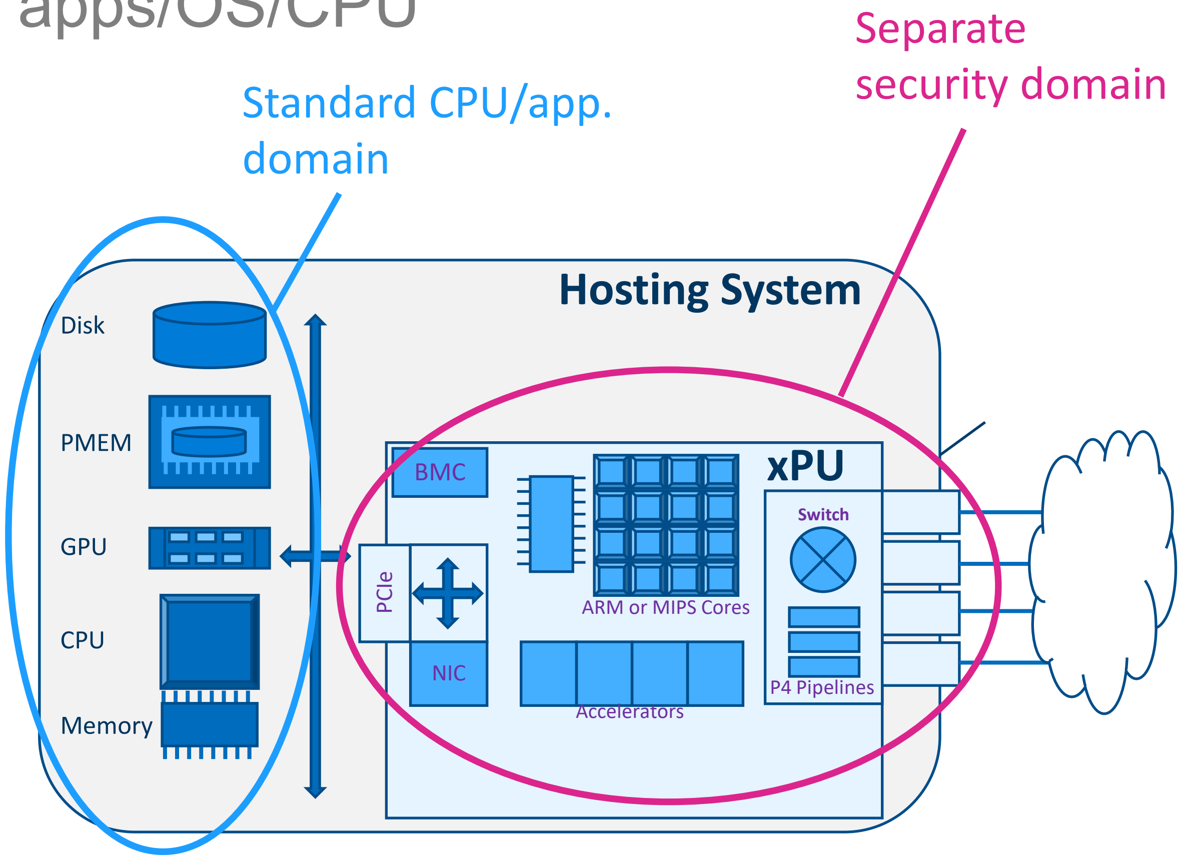
Monitoring and Telemetry

- Observability – as discussed by Mario
- Inspect traffic or server activity
- Report suspicious traffic/activity
- Act as telemetry agent



General Security Isolation and Other Functions

- General isolation from CPU/application domain
  - Run security functions separately from apps/OS/CPU
  - Security without VMs or agents
- Specialized security functions
  - Secure boot
  - Time sync
  - Application inspection/verification
  - Key exchange, RNG



Storage Offloads

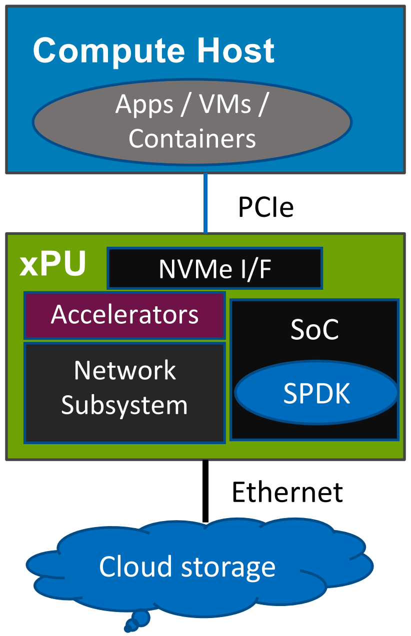
Yadong Li, Principal Engineer, Intel

Topics

- xPU Storage Usages and Accelerations
- Compression and Data-at-Rest Encryption
- NVMe-oF Offload
- Storage Stack Offloads

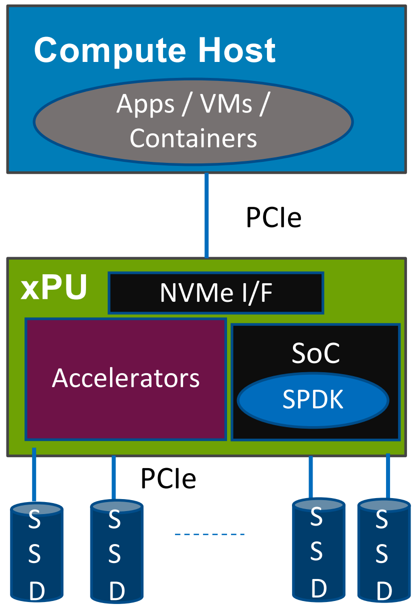
SNIA. NSF | NETWORKING STORAGE

xPU Storage Usages and Accelerations



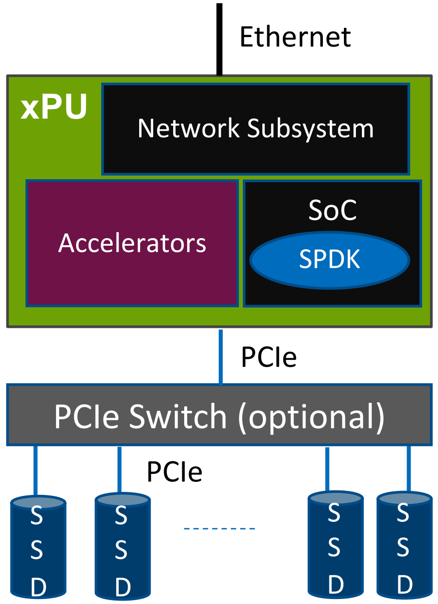
NVMe-oF Initiator Usage

- Storage Disaggregation in Cloud
- Bare-metal hosting
- **Compression + Crypto + HASH/CRC**



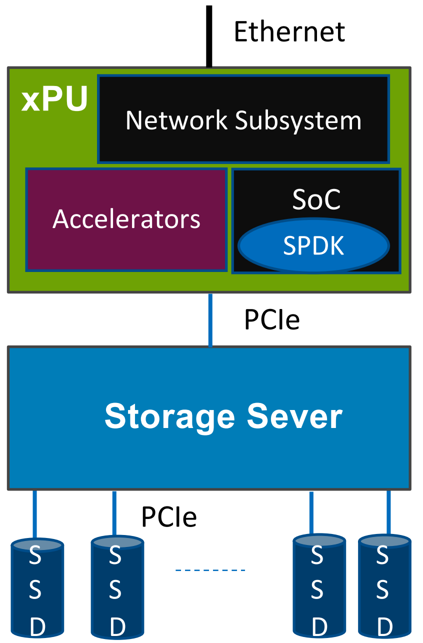
Local Storage Disaggregation

- Nitro SSD like design
- xPU provides NVMe virtualization
- **Inline Crypto + DIF/CRC**
- **Compression (?)**



xPU based JBOF Design

- Storage datapath acceleration
- Optional PCIe switch for fanout
- **Compression, HASH/dedup**
- **Crypto, CRC, Erasure Coding**



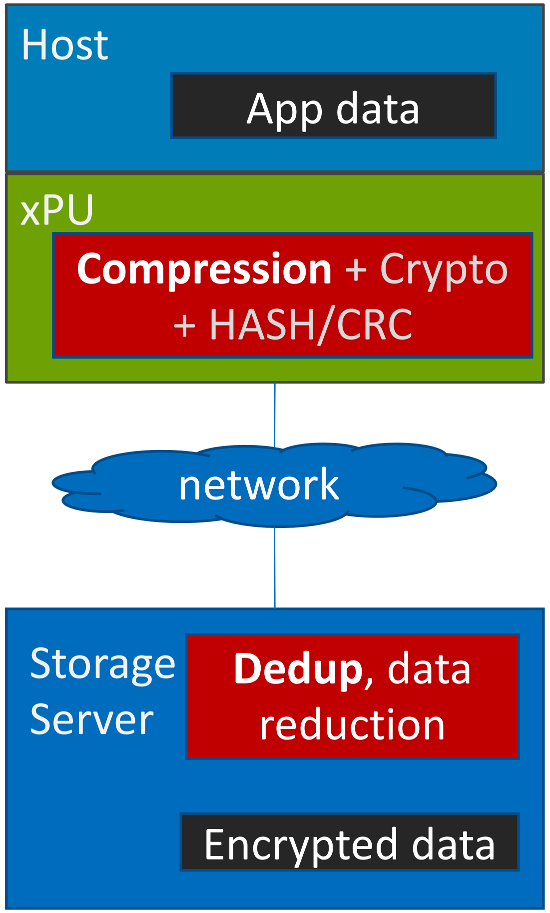
xPU + Storage Server

- *Also covers the HCI use case
- Storage datapath acceleration
- **Compression, HASH/dedup**
- **Crypto, CRC, Erasure Coding**

SNIA. NSF | NETWORKING STORAGE

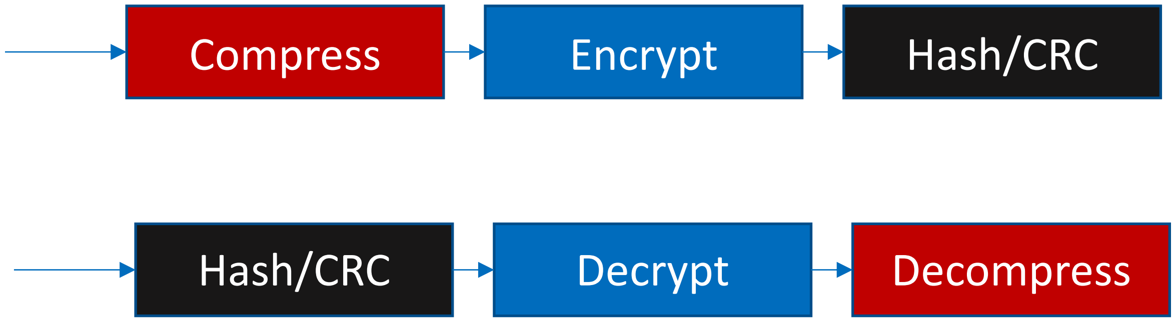
Compression Offload on Host/Initiator

- Cloud usages drive the data-at-rest encryption (per tenant) to Host/Initiator, reduces the compression and dedup opportunities on the backend storage servers.
- As a result, we also move the compression support onto the Host/Initiator.
- Key requirements for Compression+Crypto offload on Host/Initiator:
  - Data is compressed and encrypted on Host/Initiator.
  - No encryption keys shared among host/initiator and backend storage servers.
  - Compressed data still can enable dedup at the storage server side.
  - Identical data blocks produce identical ciphertext blocks to enable dedup (for a namespace using a single encryption key).
  - An efficient Compression+Crypto+HASH/CRC chained ops is required to reduce memory bandwidth consumption on xPU.



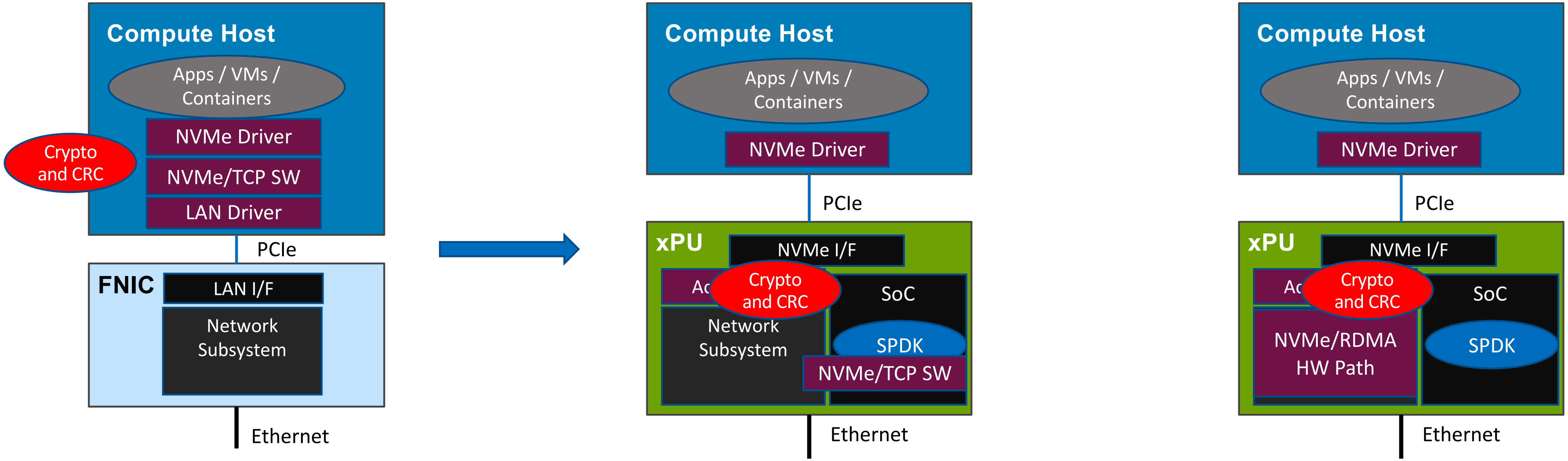
Compression and Encryption Algorithms for Storage Usages

- Compression/Decompression algorithms
  - Zstandard, Deflate, Snappy, …
- Data-at-Rest encryption/decryption
  - AES-XTS 256b, AES-XTS 512b
- CRC algorithms
  - CRC16, CRC32C, CRC64
- HASH algorithms
  - HMAC, SHA1/2/3, …
- Examples of chained ops
  - Compress + Crypto + HMAC
  - Crypto + CRC
  - CRC + Crypto + CRC



NVMe-oF Initiator Offload to xPU

- Fully leverages on the high performance NVMe interface and the integrated accelerators from xPU.
- The functions of NVMe-oF implementation can be split across HW and SW in any combination.



Standard NIC (FNIC) model

- Full NVMe/TCP SW stack on Host
- **Crypto and CRC on Host**

xPU based NVMe/TCP Initiator

- TCP stack on xPU cores, or in a HW path.
- **Inline Crypto and CRC offloads**

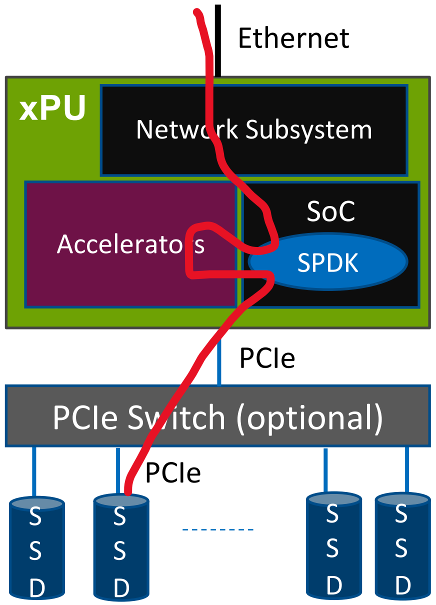
xPU based NVMe/RDMA Initiator

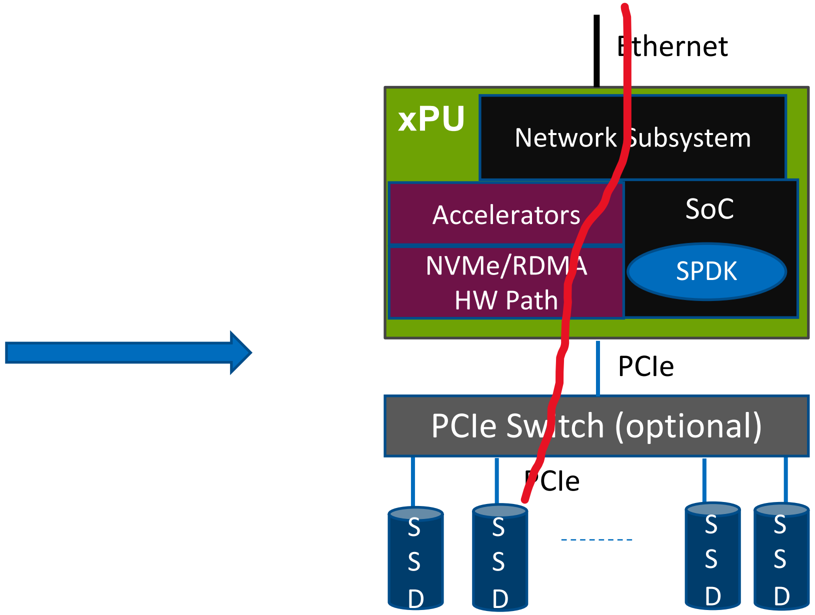
- HW or SW based NVMe/RDMA path
- **Inline Crypto and CRC offloads**

 |

NVMe-oF Target Offload to xPU

- The functions of NVMe-oF implementation can be split across HW and SW in any combination.
- Key challenges: Inline compression and inline crypto+CRC offloads on the data path.
- Compression+Crypto+CRC chained ops is necessary to reduce memory bandwidth consumption.





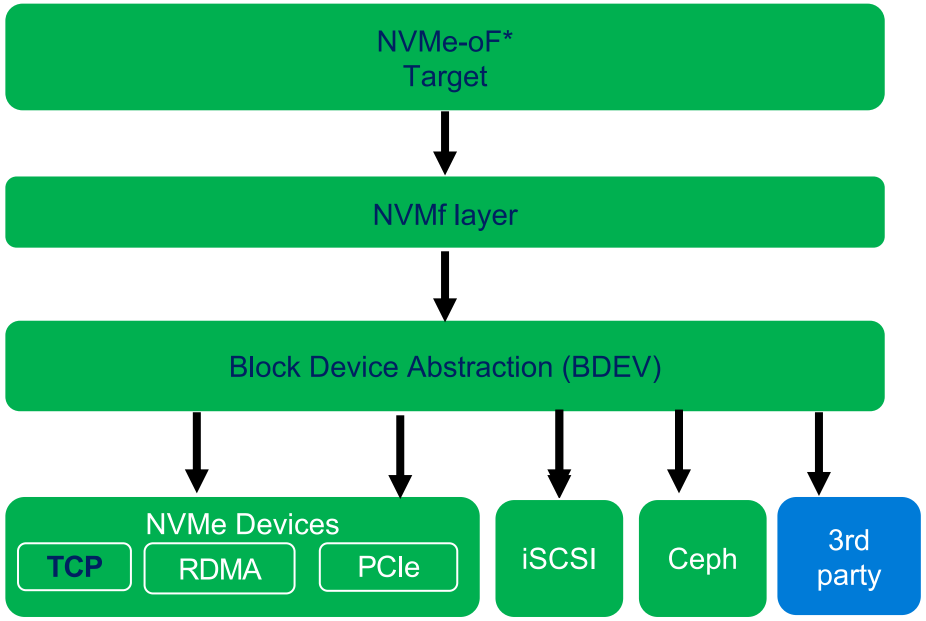
- SW path uses lookaside offload model
- **Full flexibility but less performing**

- HW path uses inline offload model
- **Low latency, high performing, but less flexible**

Storage Stack Offload

- Storage Performance Development Kit (SPDK) provides a set of tools and libs for writing high performance, scalable, user-mode storage applications.
- SPDK is an ideal framework used for xPU based storage solutions.
- SPDK integration and enhancements for xPU support:
  - NVMe/TCP initiator offload: e.g. add NPI transport on top of SPDK NVMf layer.
  - NVMe/RDMA initiator & target offload: add enhancements for RDMA QP hand-off to HW path.

SPDK Components for NVMe-oF solutions.
For details, please refer to **https://spdk.io/**



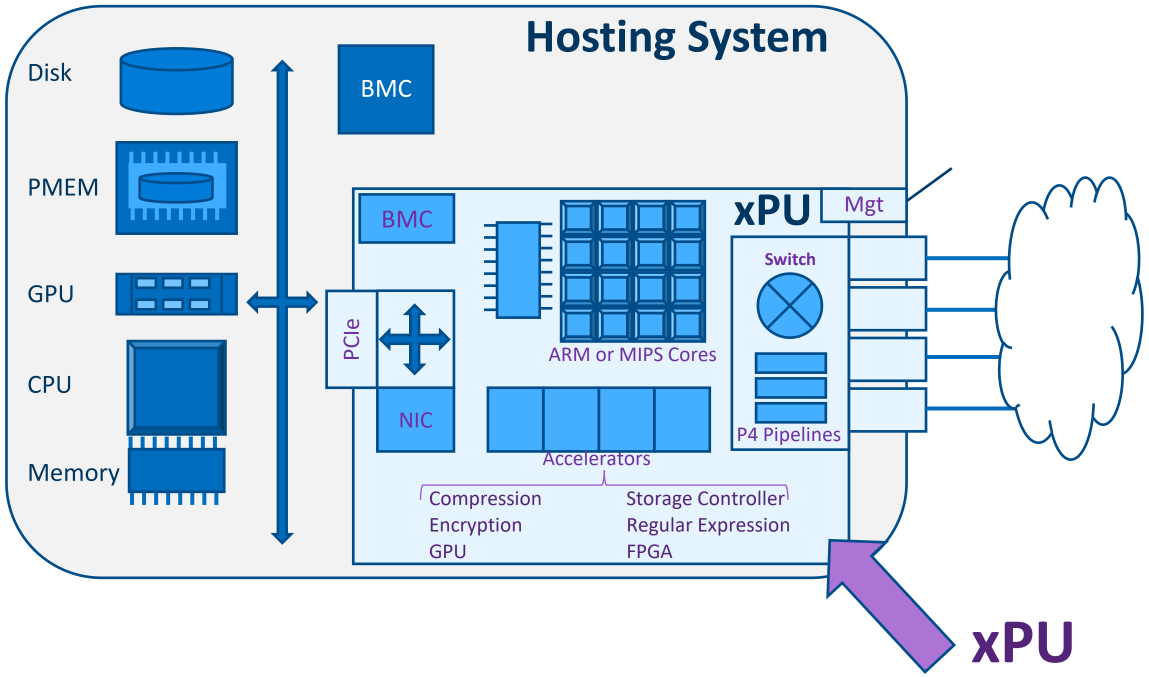
Compute Offloads

Dr Joseph L White, Fellow, Dell

xPUs and Computation

Remember our definition:

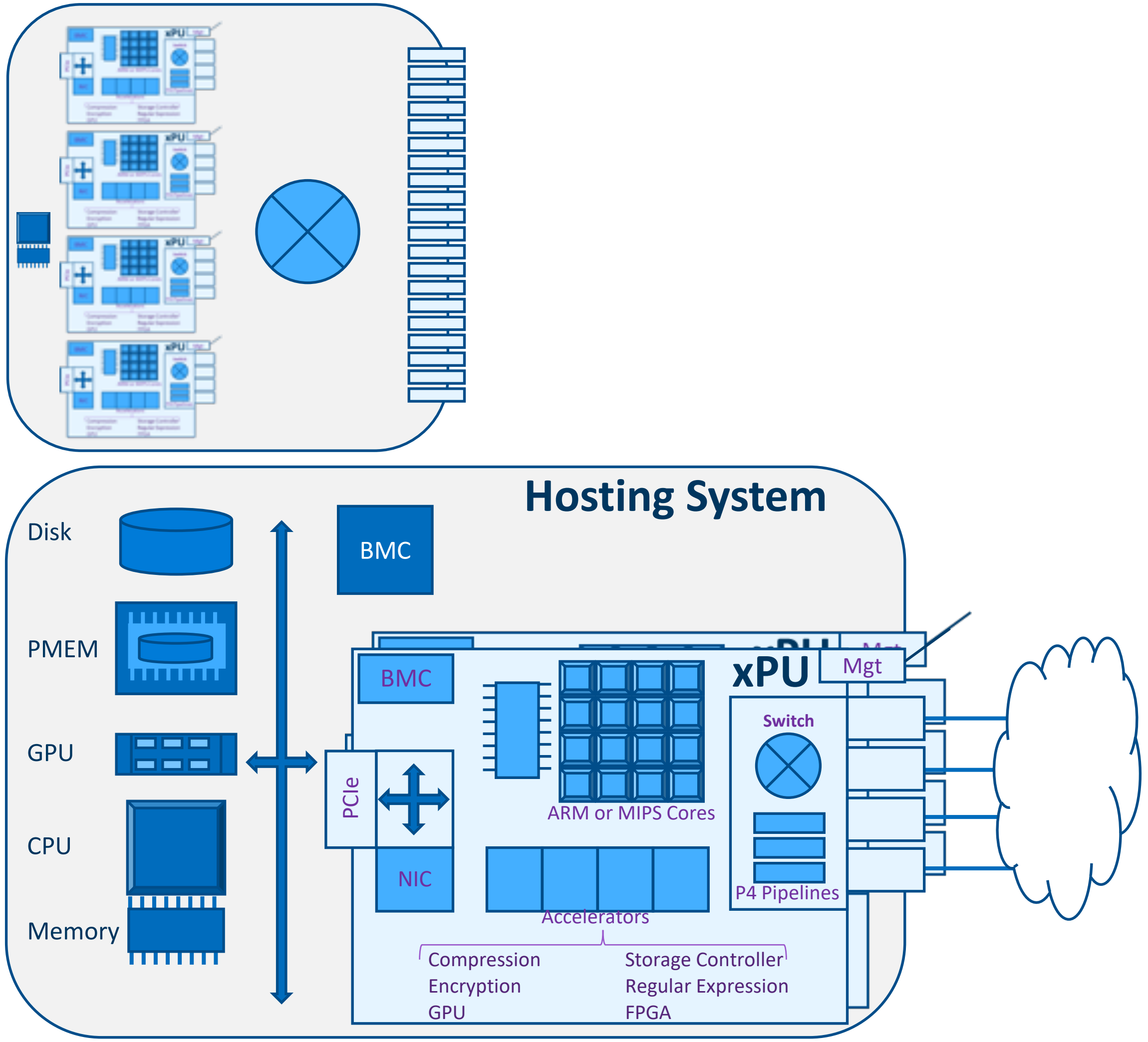
An xPU is effectively a micro-server optimized for dataflow and packet processing providing accelerators, offload engines, and local services.



Topics:

- Architecture
- Application Execution
- Host Application Support
- Resource Virtualization

xPUs and Computation: Architecture

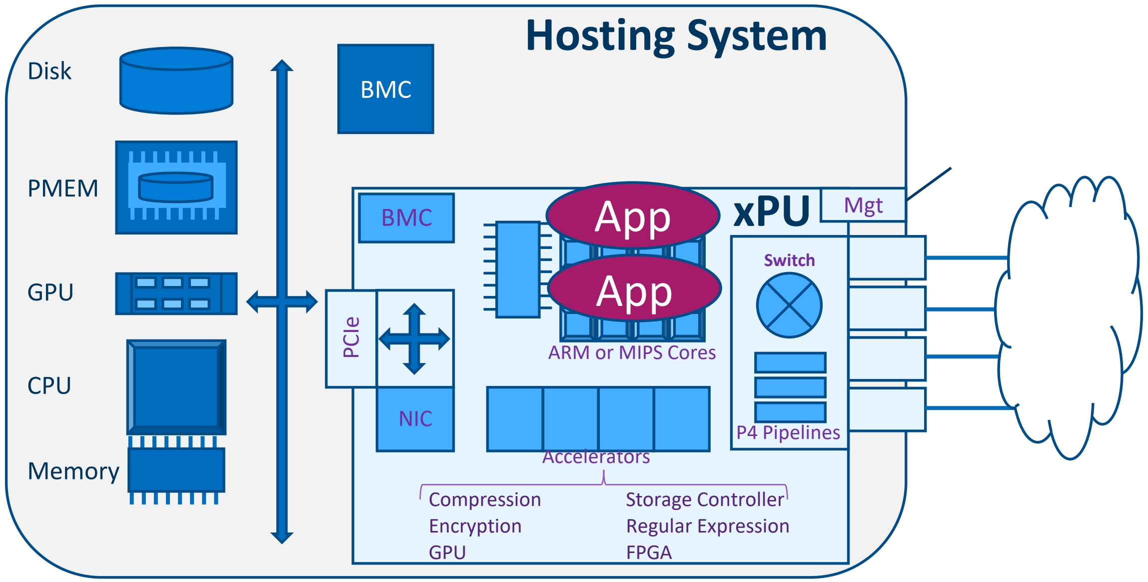


- DPU Internal Components
  - General Purpose CPU Cores with Memory
  - PCIe Interface
  - Network Interfaces (Data and Management)
  - Local Switching
  - Accelerators & Offloads
  - Programmable Pipelines
  - Embedded BMC
- Server Architecture
  - DPUs typically a built as a PCIe Card (>1 allowed)
  - Other instantiations like switch embedded or standalone possible
  - DPUs present conventional functions to hosting servers
  - DPUs can directly access PCIe Devices
- DPU Operating System
  - Linux (N flavors, Ubuntu/Debian is common)
  - VMware
  - proprietary
- Common Tool Chains Apply
  - System configuration and management
  - Network configuration and management
- K8s applicable for container installation and management

SNIA. NSF | NETWORKING STORAGE

xPUs and Computation: Applications and Containers

- Core Acceleration Functions
  - Networking
  - Storage including NVMe/TCP, NVMe/RoCE
  - Security (Firewall, DPI, Keys)



➢ Self Contained Applications and Services can run on the general purpose xPU resident cores.
  ➢ Leverage accelerators and offloads.
  ➢ Deploy as a container or process on the xPU OS

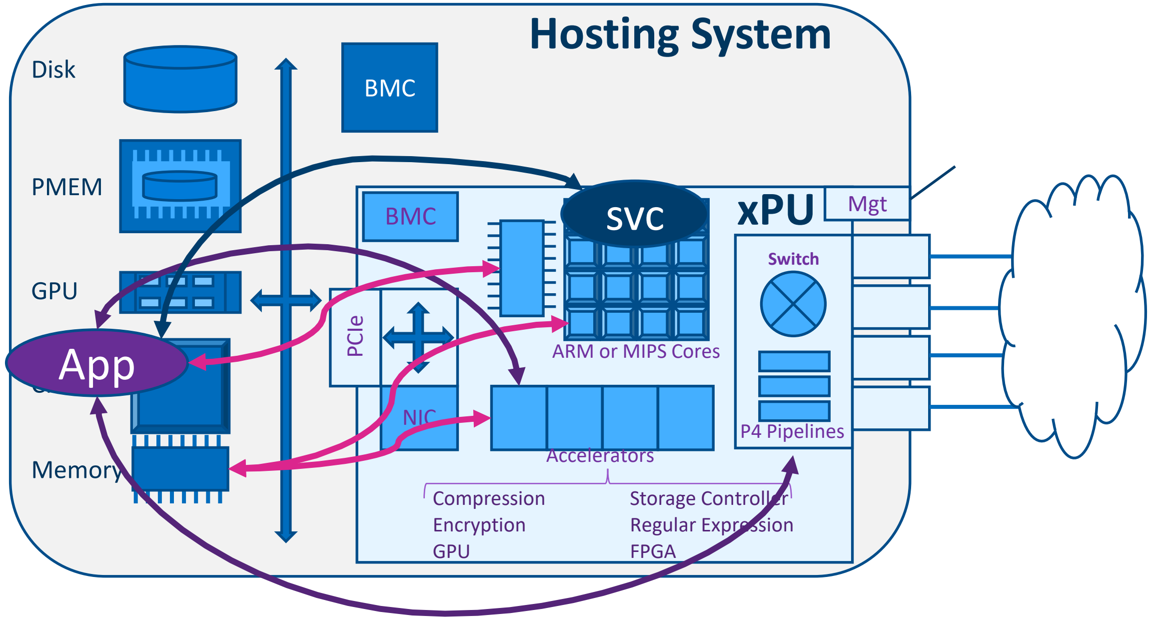
❖ Examples:
- Hypervisor offload (partial or complete)
- Control Plane Services
- AI/ML Model Evaluation
- Enhanced Telemetry Processing
- NFV (Container Network Functions)
- Firewalls, Key Management, DPI, Intrusion

 |

SNIA. NSF | NETWORKING STORAGE

xPUs and Computation: Host Application Support

- Core Acceleration Functions
  - Networking
  - Storage including NVMe/TCP, NVMe/RoCE
  - Security (Firewall, DPI, Keys)



- Leverage accelerators directly from Hosting System application: compression, encryption, RegEx, P4 Programs
- Treat the xPU as a coprocessor
- Interact with xPU resident application or service
- Directly access Hosting System or xPU Memory
  - Bi-directional DMA
- TCP & RoCE offload

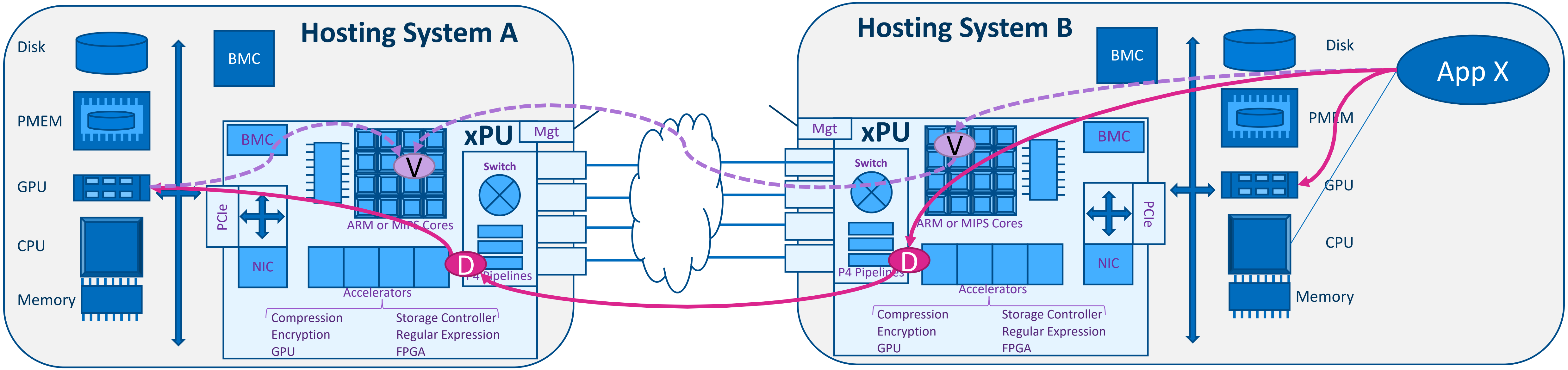
xPUs and Computation: Resource Virtualization

- Virtualize the Hosting System's Resources for use by other systems
  - GPU, Disk, PMEM, etc
- Build 'pools' of resources optimizing across a cluster

V Control plane cluster app to manage allocation of resources

D Data plane acceleration to mediate device access

Example: Suppose 'App X' on Hosting System B has run out of local GPU capacity but would benefit from adding more while Hosting system A has unused GPU capacity. App X could request from the common GPU pool (purple ovals labeled V). The resources granted would be accessed by a data plane acceleration component which handles the connection mapping and I/O data transfers transparent to App X.



Remember This is a Series!

- Our 3rd Session will be:

xPU Deployment and Solutions
Deep Dive

Follow us on Twitter @SNIANSF
For Date and Time

SNIA. NSF | NETWORKING STORAGE

After this Webcast

- Please rate this webcast and provide us with your feedback
- This webcast and a copy of the slides are available at the SNIA Educational Library https://www.snia.org/educational-library
- A Q&A from this webcast, including answers to questions we couldn't get to today, will be posted on our blog at https://sniansfblog.org/
- Follow us on Twitter @SNIANSF

Thank You

