Flash Memory Summit

Differentiated Memory Pooling with CXL

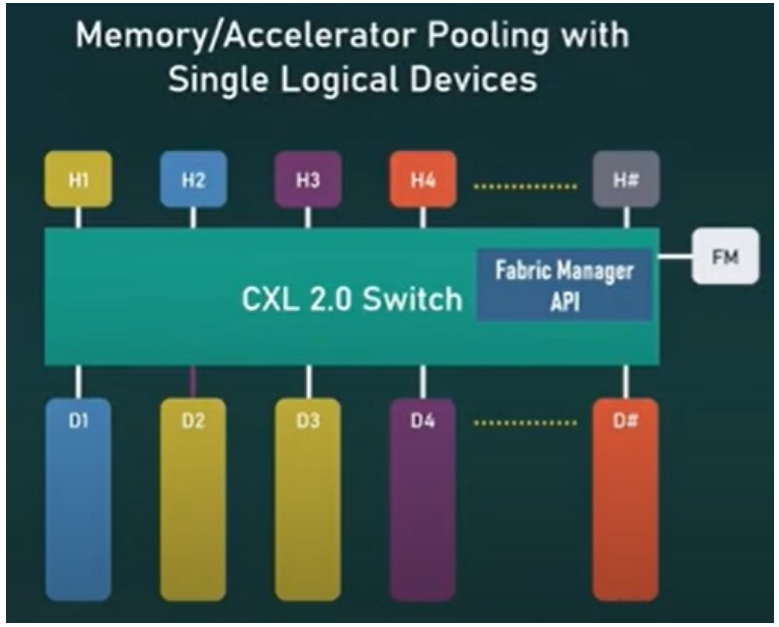
Presenter: Divya Vijayaraghavan (Intel)

Co-Authors: Pekon Gupta, Shawn Slockers, and Tom Schulte (Intel)

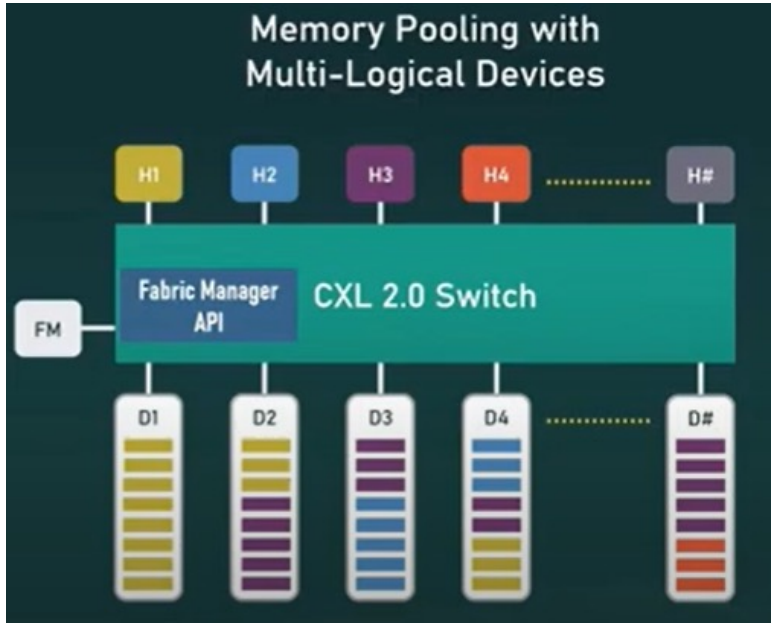
Memory Pooling Today (CXL 2.0)

- Efficient utilization of hardware
- Dynamic management and allocation of CXL-attached resources
- Software (OS, Fabric Manager), Hardware (Platforms, Switches, Memory Devices), Protocol (CXL) enhancements

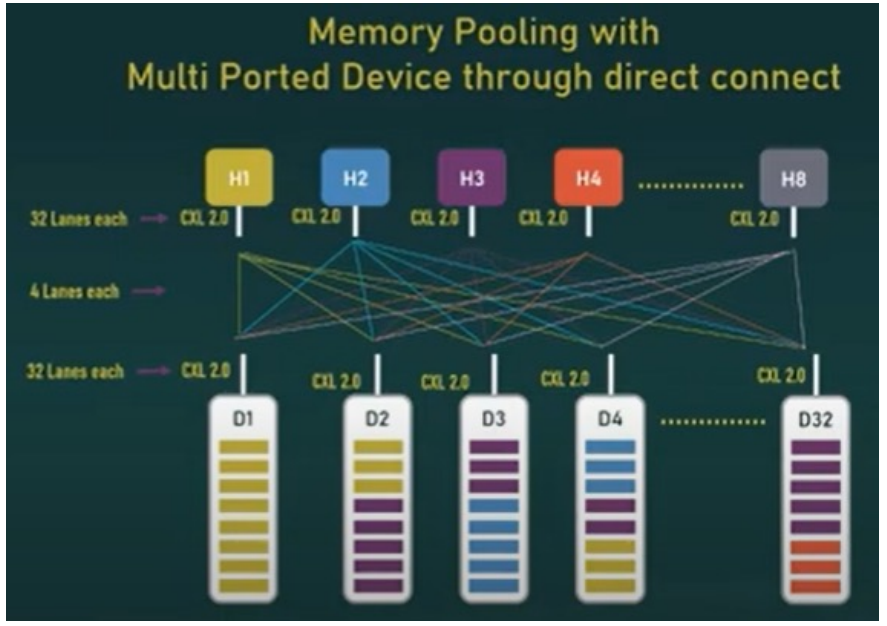
Benefits
Reduces memory stranding
Total cost of ownership (TCO) savings
Reduces memory overprovisioning



(1)



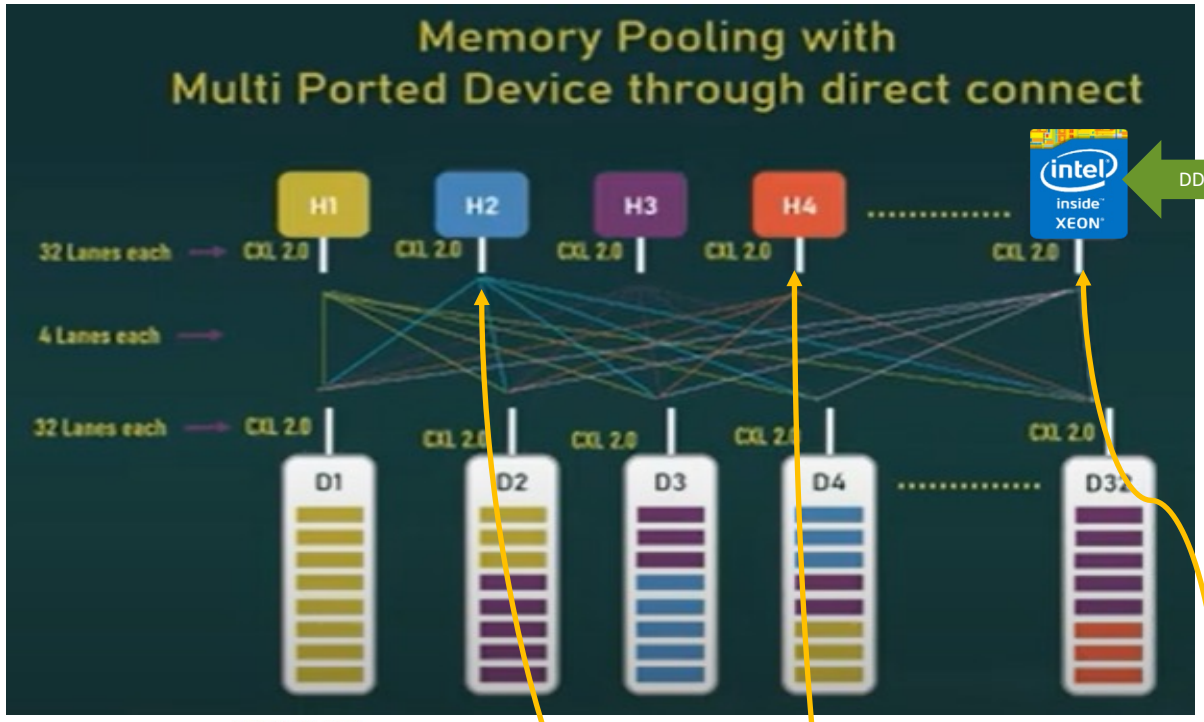
(1)

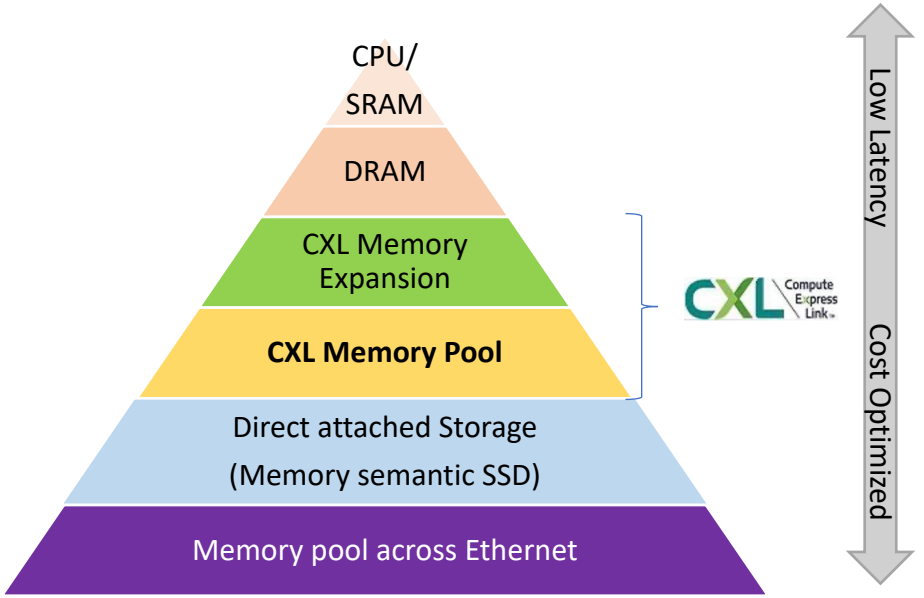


(1)

Source: Footnote (1)

Memory Pooling with Acceleration





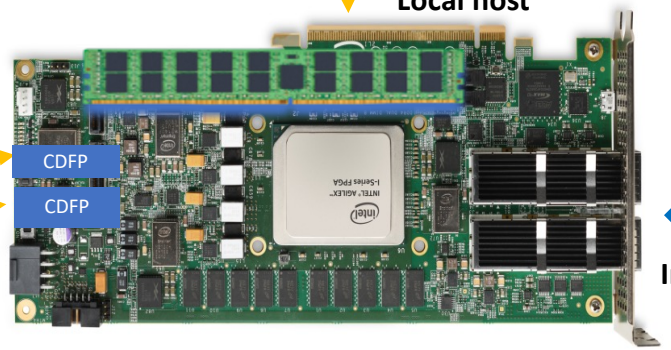
CXL Compute Express Link

Remote host

Deterministic latency

CXL Compute Express Link

Local host



Ethernet

Indeterministic Latency

RDMA,
NVMeoF or
Proprietary protocols like
CXL over Ethernet for
reliable Data sharing

Source: Footnote (2)

Memory Pooling Challenge #1

Flash Memory Summit

Challenge: Memory pooling moves data further from main compute, increasing access latency of pooled data

Mitigation	Provided by CXL 2.0 Specification	Mitigation	What a device could do
	QoS Telemetry to indicate current DevLoad (Section 3.3.2)		
	CXL.cache H2D performance monitoring (Section 3.2.3.5)		
	Multi Logical Device (MLD) to reduces hops (Section 2.4)		

Source: Footnote (3)

Memory Pooling Challenges #2 and #3

Flash Memory Summit

Challenge: Presents new challenges in system security			
Mitigation	**Provided by CXL 2.0 Specification**	**Mitigation**	**What a device could do**
	Effective Fabric Management of Virtual Hierarchies (Section 2.4)		Encrypt/decrypt with special key pre-negotiated with host/Virtual Hierarchy
			• Implement multiple data processing algorithms simultaneously • Compression, encryption

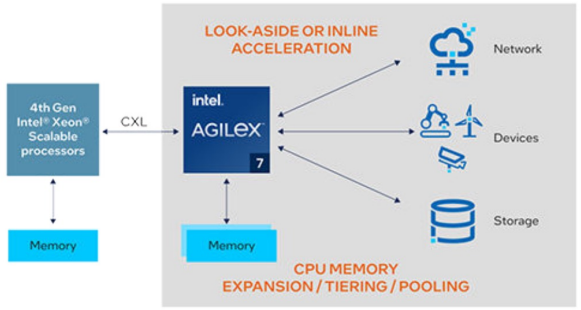
Challenge: Fault tolerance			
Mitigation	**Provided by CXL 2.0 Specification**	**Mitigation**	**What a device could do**
	Managed Hot Removal, Global Persistent Flush (Sections 9.9, 9.8)		• Use partial re-configuration to achieve fault isolation with flexibility • Aging/electrical fault or fix bugs on the fly

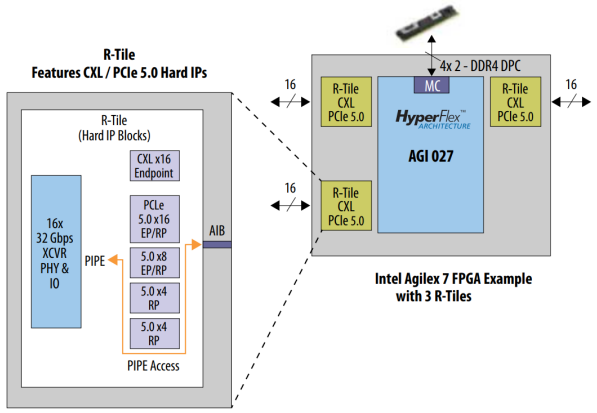
Memory Pooling: FPGA Opportunities

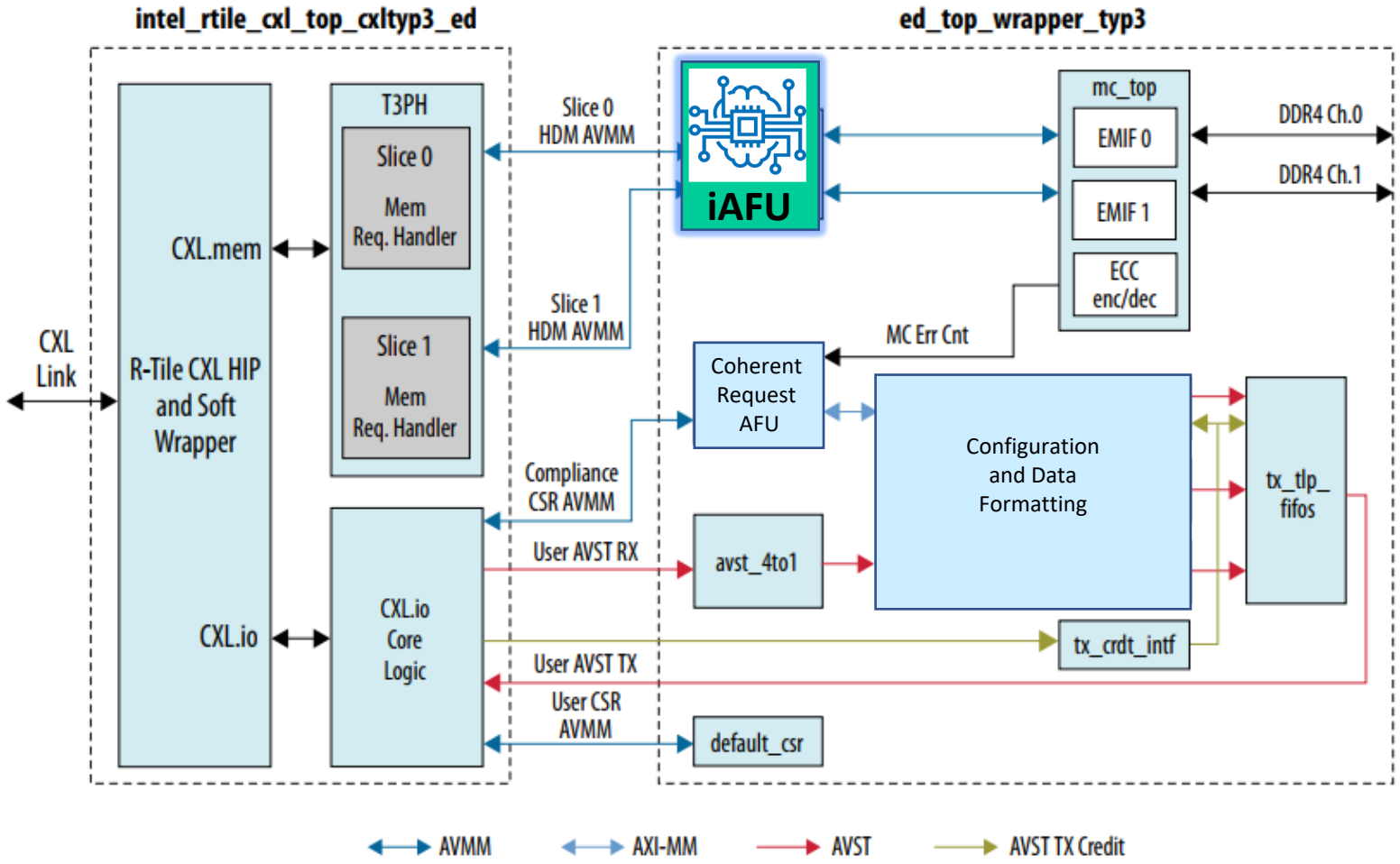
	FPGA as a Memory Pooled Device	FPGA ++
CXL Workload(s):	Pooling and switching (share) Tiering (page management)	+
Use Cases:	Dynamic platform pooling and tiering Rack pooling and tiering Multi-host management Data protection and redundancy	++
Cost Sensitivity:	Moderate (TCO comparison)	+
Bandwidth (Line Rate):	20% ~ 80%	++
Form Factors:	PCIe CEM, other	+
Latency (End Point):	~200ns to 500ns	
Memory Technologies:	DDR4/5, emerging and custom memory architectures, various speeds	+
Transceivers:	Connectivity to Ethernet, GenZ, OpenCAPI, custom protocols	++
Hard Processor System:	Fabric management, dynamic resource re-routing, offloading CPU workloads	+
Power:	Depends on implementation (TCO comparison)	+

FPGA Focus → Disaggregation and Acceleration

Intel's FPGA-based Type 3 Features to Support Memory Pooling *(Simplified Functionality)*







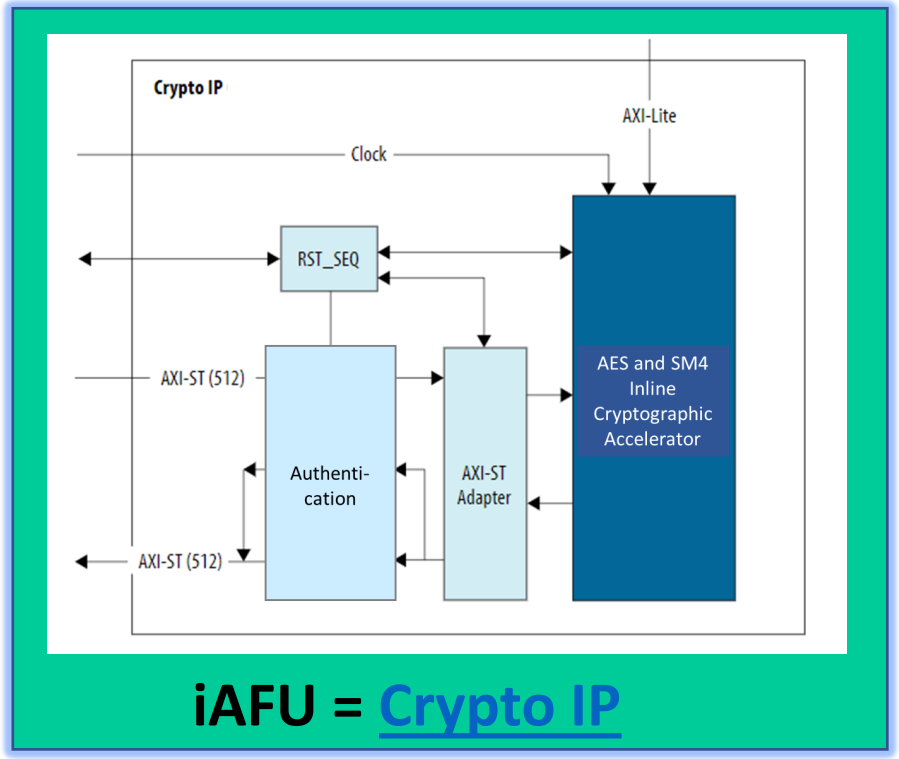
Source: Footnote (4)

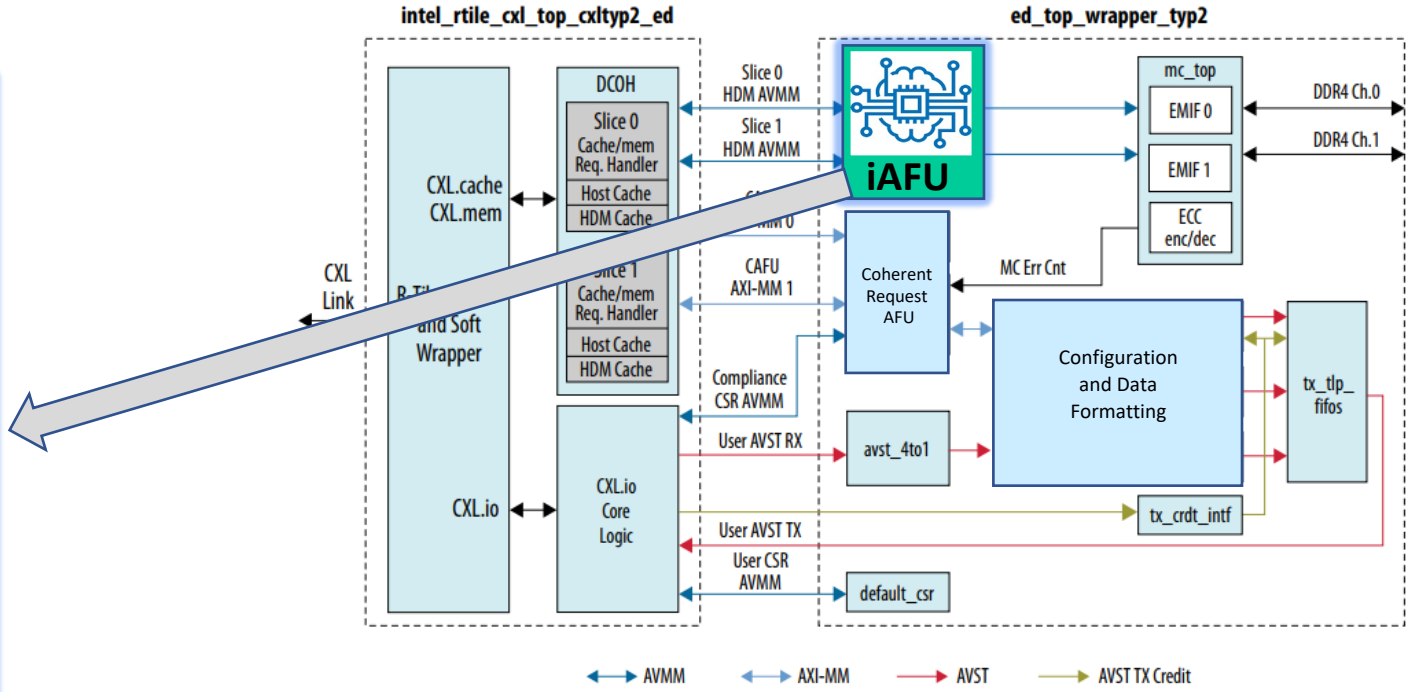
Possible Option: Crypto iAFU in FPGA

Flash Memory Summit

Benefits:

- End-to-end flexible and unique encryption/decryption capability
- Customized security-related algorithm
- Single FPGA implements many types of encryption





Key Takeaways

Flash Memory Summit

- CXL helps memory pooling implementation, but challenges remain

- Some challenges can be mitigated by FPGA-based solutions
  - CXL IP with value added features including ‘out-of-order’ HDM transaction support
  - Dynamic re-configurability
  - Hard Processor System
  - 10 – 800 GbE capability
  - PCIe 5.0/6.0 connectivity
  - Pre-built ‘accelerator’ functions (IP/Solutions)

Learn More or Get Started
(IP, boards, etc.) by contacting Intel FPGA team:

https://plan.seek.intel.com/psg_WW_dcaipsgloc_LPCS_EN_2022_IFTDContactUs

Open Discussion

If time is available, Questions & Answers with audience.

Other Contributors from Intel

- Eric Pham
- Seshan Sekariapuram
- Tom Kendzulak
- Sung San Choe
- Miguel Amador Alvarez
- Eliath Guzman Flores

Sources

(1) Compute Express Link 2.0 Specification: Memory Pooling - CXL Consortium

(2)

- Smart Modular XMM CXL E3S
- Intel Agilex I Series FPGA
- Samsung MS SSD

(3) Data Center Memory Systems at Meta - OCP Global Summit 2022

(4) Diagrams from Intel Agilex 7 R-Tile Compute Express Link (CXL) 1.1/2.0 IP User Guide (RDC # 763328)

(5) Redfish Support for CXL

Legal Disclaimers

Flash Memory Summit

Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex. Results may vary. Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications, and roadmaps.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No product or component can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com

Flash Memory Summit

Thank You

Presenter: Divya Vijayaraghavan (Intel)

Co-Authors: Pekon Gupta, Shawn Slockers, and Tom Schulte (Intel)