Use Cases for CXL-based Active Memory Tiering and Near Memory Accelerators

Presenter: Divya Vijayaraghavan (Altera)

Co-authors: Tom Schulte and Pekon Gupta (Altera)

FMS the Future of Memory and Storage

Two Prominent CXL Use Cases

1. Active Memory Tiering
   - Local and remote memory tiers, migration of hot and cold pages between tiers
2. Near Memory Compute Acceleration
   - Remote memory tiers accelerate or process data near memory elements

Active Memory Tiering	Near Memory Compute Acceleration
High Performance Compute	SmartNIC, IPU
	Financial Services

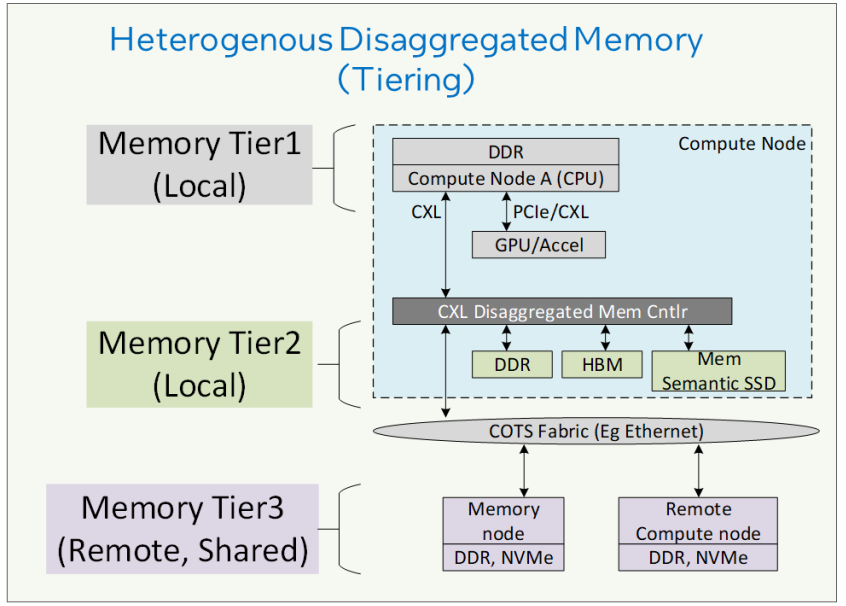
Architecture \ Attributes	Memory Expansion	1 Memory Disaggregation	2 Acceleration
Use Cases	Capacity Expansion Bandwidth Expansion Software assisted Tiering	Hardware assisted Tiering Differentiated Memory pooling Multi-host management	Inline acceleration Look-aside acceleration (QAT)
Cost Sensitivity	High	Moderate	Moderate
Bandwidth	80% of line-rate	~ 80% of line-rate	~ TBD
Form-factors	EDSFF (E3.S, E1.S)	PCIe CEM, Blade, Custom	PCIe CEM, Blade, OCP, custom
Latency (round trip)	< 100ns	~200ns to 350ns	~300ns to 500ns
Media	DRAM	DRAM DDR4/5, NAND, Emerging persistent Memory	DDR4/5 DRAM, NAND
Power	Low: 50% ~ 90% of DDR5	TBD	TBD

1 Active Memory Tiering Considerations

Approaches

- S/W driven: Kernel scans memory allocation, identifies local vs. remote memory references (1)
- H/W-based hot page detection: Identifies most frequently accessed physical pages in Tier 2 memory
- Hardware-assisted application-transparent memory tiering management (2)

Challenges	Accuracy of "hotness" classification	**Mitigation**	Increased offload of hot page detection to hardware
	Page migration latency		Provide enhanced memory access monitoring/reporting capabilities on cxl.mem HDM interface
	Understanding of workload characteristics		Identify frequently used Host Physical Addresses
	Hardware vs. software partitioning		

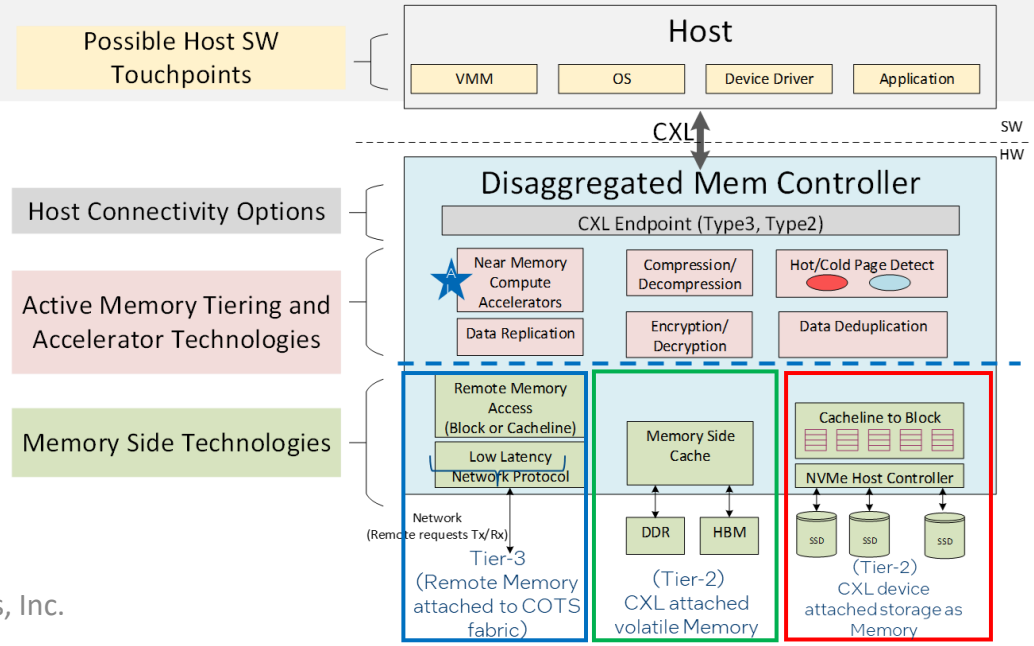


2 Near Memory Compute Acceleration Considerations

Approaches

- Near memory processing engine implemented in proximity to EMIF controller on CXL EP device (3)
- Memory tiering with computational memory devices and standard memory devices (4)

Challenges	Acceleration functions decision - analytics, matrix manipulation or other	**Mitigation**	FPGAs, flexibility, accelerate functions near memory
	SDK development (new accelerated function)		Combine Type 2 computational memory and Type 3 memory tiering for best TCO
	Computational memory is expensive		

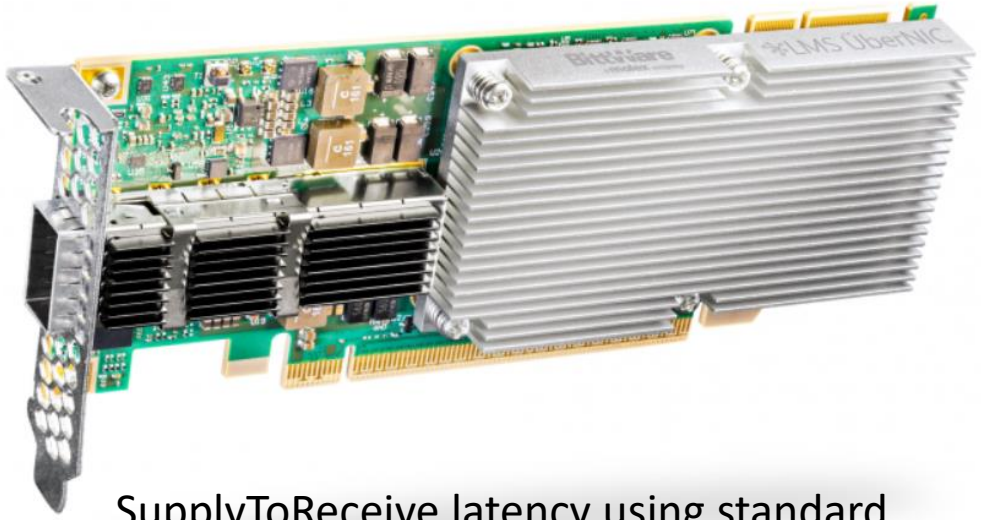


Performance Metrics

- Publicly disclosed data points emerging showing latency advantage of CXL Type 2 Near Memory Acceleration vs. traditional PCIe [5]

Posted June 25, 2024

STAC Report: LMS ÜberNIC CXL with 10GbE and 25GbE under STAC-N1



SupplyToReceive latency using standard 264-byte message sizes

New records from the first tests of a pure FPGA-based or CXL-based UDP stack.

Intel-UIUC ksm offload to CXL Type 2 device

- Kernel features increase tail latency of applications and consume CPU cycles
- Offloading the kernel features to CXL Type-2 device
  - 83% lower tail latency of application
  - 61% fewer CPU cycle consumption by `ksm`

Measured CXL.cache latency 68% lower than PCIe

Protocol	Tool	Initiator	Target	Design Used/Command
CXL.cache	Intel CXL Stress Tester	Device	Host	CXL Type2 ED/ Rdcurr
PCIe	MCDMA Driver	Device	Host	MCDMA ED / Mem Read

Takeaways

- CXL-based Memory Tiering and Near Memory Acceleration provide advantages
  - Reduction in system TCO
  - Offloading of processing from CPU
  - Reduction in processing latency for specific workloads
- Some challenges can be mitigated by FPGA-based solutions
  - CXL IP and design example with configurable pre-built accelerator functions
  - Dynamic reconfigurability

Contributors from Altera

- Bhushan Chitlur
- Sung San Choe
- Shawn Slockers
- Navneet Rao
- Zhongqian Yu
- Lingyan Li
- Xuan Zhao
- Jiwei He

References

(1) Meta, UMichigan - TPP: Transparent Page Placement for CXL-enabled Tiered-Memory: https://arxiv.org/abs/2206.02878

(2) Google - https://doi.org/10.1145/3582016.3582031

(3) SKHynix - https://www.youtube.com/watch?v=pbnTlY41h08

(4) SmartModular - https://www.youtube.com/watch?v=A_PML20fk-Y

(5) STAC - STAC Report: LMS ÜberNIC CXL with 10GbE and 25GbE under STAC-N1 | STAC - Insight for the Algorithmic Enterprise | STAC (stacresearch.com)

(6) Unifabrix - https://www.unifabrix.com/

Backup