Exploring CXL® Use Cases and Implementations

June 27, 2024

Presented by: Oren Benisty (UnifabriX), Ahmed Medhioub (Astera Labs), and Matthew Burns (Samtec)
Memory Wall

Source: https://www.youtube.com/watch?v=Slumdt2vLyo
Memory is Key to AI, HPC & Big Data Performance

- The emergence of Terabyte-Class Models pushes the limits of the infrastructure toward MEMORY POOL
- ML models grow exponentially
Meet MAX: World’s first Software-Defined Memory Pool

- Standard compliant: Works with any CXL 1.1/2.0 CPU and GPU
- Fits into any standard rack from factor (2U)
- Large memory capacity 4-32 TB
- Scalable memory capacity >256 TB with CXL 3.1 fabric
- Scalable memory bandwidth >512 GB/s
- Memory Pooling and Adaptive Sharing
- High-Performance NVMe® Storage
- PMEM/Optane replacement

Availability: We are ready, would you like to test the Memory Pool?
SuperScaling HPC & GenAI with **MAX-Memory**

UnifabriX MAX provides up to 256TB of GPU/CPU shared memory for the most demanding HPC & Generative AI workloads.

The introduction of CXL opens the door to new opportunities.

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium.
Unifabrix MAX - Acceleration across the board

MAX-Memory accelerates the highest demanding workloads

Use-case:
( benchmark)

- Generative AI inference (BERT)
- HPC (HPCG)
- Data crunching (FIO)
- Data analytics (TPC-H)
- AI Recommendation Engine (DIEN)

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium
Contact details:
oren.benisty@unifabrix.com
www.unifabrix.com
Memory Expansion with CXL Type 3 Devices

Ahmed Medhioub – PM, Astera Labs
CXL Type 3 Devices

**CXL.mem**
Processor access to device-attached memory
- Memory bandwidth expansion
- Memory capacity expansion
- Memory Pooling & Sharing

**CXL.io**
PCIe discovery, configuration, interrupts, etc.
- DMA
- Interrupts MSI-X
- DDR5 memory enablement
- Support for PCIe switches
CXL Type 3 Use Cases

Memory Expansion
- **Challenge:** Limited CPU memory channels
- **Solution:** Flexible Memory expansion for application specific workloads

Memory Pooling
- **Challenge:** Stranded memory due to over-provisioning
- **Solution:** Dynamic memory allocation

Memory Sharing
- **Challenge:** Inefficient data flow and memory copies
- **Solution:** Shared memory regions across hosts

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium
Lower TCO with CXL Memory Expansion

- Optimize CPU utilization and deliver more efficient memory expansion
- Expand memory with a flexible CXL fabric
- Lower CAPEX and OPEX for memory-intensive applications

Unprecedented Memory Capacity and Bandwidth Expansion with Low TCO

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium
Enabling Cloud-Scale Deployment with Broad CPU and Memory Support
Leo - A Holistic Solution

CXL Solution Requirements

- Application Performance
- RAS
- Security
- Inband System Management and Telemetry
- Out-of-band System Management and Telemetry
- CPU Interoperability
- DIMM Interoperability

Platform Solution Stack

- Application
- OS/Kernel/BMC
- COSMOS Platform APIs
- Device drivers
- Modules
- COSMOS Embedded Software
- Silicon

Astera Labs Provides a Holistic Solution That Enables the CXL Ecosystem

*Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium*
Comprehensive CXL Connectivity Solution

Local CXL-Attached

Medium Reach, CXL-Attached

Long Reach, CXL-Attached

Mem expansion within the chassis

Memory Appliance

Rack-scale Shared/Pooled Memory

Unlocking More Capabilities
CXL Over Optics – Why the Need?
Matthew Burns | Samtec | Director, Technical Marketing
**Evolving AI and disaggregated computing system topologies/architectures require more external cabling**

- Reach requirements vary depending on use case
  - 2m – Within the rack
  - 7m – Rack-to-rack
  - 10m+ – Larger clustering

- Several industry efforts underway to answer the challenge
  - Passive DACs
  - Re-timed AECs
  - New PCI-SIG® CopprLink™ Cable Specification
  - Optical Transceiver MSAs

- Is there another way?
The Case for Mid-Board Optics

Easier Layout, Higher Density, Better Signal Integrity, Lower Power

At Board Edge
On Board
On Package

FPP
Memory
Processor

Motherboard
Backplane

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium
Samtec Mid-Board Optical Transceivers

Data connection is taken “off board,” simplifying board layout and enhancing signal integrity from IC to faceplate

FEATURES

- Up to 32 Gbps per channel via optical cable for greater reach
- Industry leading miniature footprint allows for higher density close to the data source
- Protocol agnostic

- Simple to use system with easy insertion/removal and trace routing, no through-holes, and a SMT connector system
- Supports data center, HPC and FPGA Protocols, including 10/40/100 GbE Ethernet, InfiniBand™, Fibre Channel, PCIe®, CXL, and Aurora
PCIe®/CXL over Fiber Optical Transceivers

- Transmits PCIe 4.0 electrical signals through Samtec optical transceivers > 100 m
  - PCIe 5.0 under development

- Supports x4, x8 and x16 configurations

- Enabling high data throughput, coherency and low latency

- High-performance signal quality with BER better than e-12

- Supports CXL protocols:
  - CXL.memory, CXL.cache, and CXL.io
CXL of Over Optics PoC
CXL Over Optics Form Factors

• PCIe® CEM AIC supports PCIe 4.0
  • PCIe 5.0 under development

• Supports x4, x8 and x16 configurations
  • Single x16, x8
  • Dual x8, x4
  • Quad x4

• Reconfigurable for host or target operation

• Other form factors under development
  • OCP NIC 3.0, E3.S 2T, M.2, etc.
Optical Connectivity Case Study

Samtec Optical Transceiver
Key Takeaways

• CXL over optics a reality to support evolving AI and disaggregated computing system topologies/architectures

• Samtec offers the industry’s most comprehensive portfolio of mid-board optical modules ideally suited for next-gen system architectures

• Visit us to learn more:
  • www.samtec.com/optics
  • optics@samtec.com
Q&A