Memory Challenges and CXL™ Solutions

Chris Petersen
CXL Director and Hardware Systems Technologist
Facebook

Prakash Chauhan
CXL Director and Systems Architect
Google
Today’s Presenters

Chris Petersen
CXL™ Consortium Director
Hardware Systems Technologist, Facebook

Prakash Chauhan
CXL™ Consortium Director
Systems Architect, Google
Agenda

- Memory Trends
- Memory Challenges
- Emerging memories for cost reduction
- CXL Enabled Memory Solutions
- CXL.Mem Flows for Type 3 memory devices
- Summary
Memory Trends: Increasing Costs

- Memory an increasing fraction of System Cost
  - Memory Price (cost/bit) flat due to scaling challenges
- Increasing core counts driving Memory Demand
  - Increased Capacity
  - Increased Bandwidth

Data Source: De Dios & Associates
Challenge: DRAM (1T-1C) Cell No Longer Scaling

- Cell Scaling is challenged by the following trends
  - Worsening Aspect ratio > 50:1
    - Burj Khaleefa A/R is 9:1
  - Mechanical stability
  - Layer to Layer registration
  - Reducing Capacitance value
  - Increasing leakage (High K dielectrics)
  - Increasing Cell-Cell interference (coupling capacitance)
  - Variable retention time
Memory Trends: BW vs. CPU Core Count

Source: De Dios & Associates

Compute Express Link™ and CXL Consortium™ are trademarks of the Compute Express Link Consortium.
Challenge: Adding Memory Bandwidth is Expensive

- Add channels to CPU
  - Large Sockets
    - Cost, Reliability
  - PCB Layer Count
    - Additional layer per channel
  - Board form-factor
    - Difficulty fitting in standard widths

- Increase Data Rates
  - PCB technology
    - Back-drill, SMT connectors, blind vias
    - Faster cross-socket links require ultra low loss materials (balanced bandwidth system)
  - Equalization circuits
    - Complexity, cost added to both ends
  - 1DPC
    - Capacity/Granularity Issues
    - Exotic DIMMs (LR, 3DS)
Emerging Memories for Cost Reduction

- Fast
- Expensive
- Low capacity
- Not persistent

- Slow
- Cheap
- High capacity
- Persistent

Emerging Persistent Memory devices

Load-store w/ various HW attach methods - e.g. DDR bus, CXL

Compute Express Link™ and CXL™ are trademarks of the Compute Express Link Consortium.
Emerging Cross-Point Memories Scale Well

- Column Decoder
- Row Decoder
- Sense Amps
- Base Planar Layout
- Peripheral Circuits Under Cell Array (CuA)
- Stacked Cell Arrays

$F - F^2 - 4F^2$

Compute Express Link™ and CXL Consortium™ are trademarks of the Compute Express Link Consortium.
Emerging Memory Attachment Challenge

- Emerging memories are often transactional
  - Non-Deterministic Timing
  - Asymmetric Read/Write timing

- Not ideal for sharing DDR Bus with DRAM
Problems with DIMM-only Based Memory Solutions

- Transactional DIMMs have problems coexisting with regular DDR DIMMs
  - Lowers Bus frequency (2DPC or lost channel B/W)
  - Arbitration between DRAM and Transactional DIMMs impacts efficiency
  - QoS impacts and latency increases for DRAM
    - Slowing access to high performance memory

- Only homogeneous DIMM types are possible
  - Same generation of DDR
    - Memory controller + PHY limitation
  - Same speed grades and timing
    - For bus efficiency and controller implementation
  - Same device geometry
    - To allow interleave across all channels
  - Same power and thermal envelopes
CXL Opportunities

1. CXL is a memory agnostic, but coherent interface

2. CXL can address system design challenges

3. CXL enables new compute and memory architectures
CXL Device Types

Type 1: Caching Devices / Accelerators
- Usages:
  - PGAS NIC
  - NIC atomics
- Protocols:
  - CXL.io
  - CXL.cache

Type 2: Accelerators with Memory
- Usages:
  - GPU
  - Dense Computation
- Protocols:
  - CXL.io
  - CXL.cache
  - CXL.memory

Type 3: Memory Buffers
- Usages:
  - Memory BW expansion
  - Memory capacity expansion
  - Storage Class Memory
- Protocols:
  - CXL.io
  - CXL.mem

Compute Express Link™ and CXL™ are trademarks of the Compute Express Link Consortium.
CXL: A Common Memory Interface

• CXL provides a common, standard interface for many types of memory

• Media independence: could be used to connect DDR3/4/5, LPDDR3/4/5, Persistent memory, etc

• Enables flexibility for different media characteristics (persistence, latency, BW, endurance, etc)
Heterogeneous Memory Attach - DIMMs vs CXL

- DIMMs are not suitable for Heterogeneous Memory types

- CXL solves the problem
  - Enables a slow memory tier to be completely *isolated* from main tier
    - Minimal interference between CXL and direct attached DRAM DIMMs
  - Enables other memory types whose *bandwidth is additive* to existing platform memory bandwidth
    - E.g. DDR4 and DDR5 can coexist in the same platform
  - CXL *capacity additive* to platform memory capacity
    - With inevitable move towards 1 DIMM per DDR channel, CXL becomes a cost-effective path for capacity expansion
1. Power density
   • DIMM slots tend to be power limited to 15-18W (less in more dense platforms)
   • CXL enables:
     • Separation of DIMM slots and CXL Memory slots
     • Higher power CXL Memory devices (e.g. 25W+)

2. Memory channel count scaling
   • Parallel DDR* interfaces require 200+ pins
   • CXL enables:
     • Less pins per package = more channels or smaller packages
     • Lower mother board PCB layer counts
Scaling Compute and Memory Independently

- Adds 300GB/s of BW
- Enables flexibility to add a variety of memory without impacting DDR5 DIMMs
- 800+ less pins/signals vs. 10x DDR5 channels

Compute Express Link™ and CXL™ are trademarks of the Compute Express Link Consortium.
CXL Type 3 Devices

Memory Write Flow

Compute Express Link™ and CXL™ are trademarks of the Compute Express Link Consortium.
CXL Type 3 Devices

Memory Write Flow

1. Host
2. CXL Device Memory Controller
3. Memory Media

- MemRd
- MetaValue=NoOp
- Read
- Data + ECC + Meta=2
- Data MetaValue = 2
- No Meta Update Needed
CXL Type 3 Devices

Memory Invalidate Flow

MemInv used for read and writing MetaValue

New MetaValue

Old MetaValue

Compute Express Link™ and CXL™ are trademarks of the Compute Express Link Consortium.
CXL: A Scalable Solution for Memory

Fast
Expensive
Low capacity
Not persistent

Slow
Cheap
High capacity
Persistent

CPU
DRAM
NRAM/MRAM
3D-Xpt, ReRAM, PCM
PCIe Low Latency SSD
LLNAND, Optane
NAND SSD

CXL-attached

Compute Express Link™ and CXL™ are trademarks of the Compute Express Link Consortium.
CXL Memory: More Work Ahead

• Hardware
  • Ecosystem growth and interoperability
  • Form factor development and alignment
    • Power, thermal, mechanical and management interfaces
  • Evaluation and characterization of memory latency
    • Tools and Benchmarks

• Software
  • Non Uniform Memory
    • Job scheduling based on latency, bandwidth, capacity
  • Interleaving
    • Interleave sets for maximum parallelism without impacting performance
  • RAS
    • Unified error reporting and handling for heterogeneous memory types
  • Security
    • Firmware security, media security, data integrity and isolation with heterogeneous media controllers
  • QoS
    • At link and Media level
Summary

• CXL enables memory tiering with various media types including emerging memory technologies

• CXL enables media independence and solves system design challenges

• CXL will help us address the future memory roadmap by providing new opportunities to scale memory capacity and bandwidth
Call to Action

- To join the CXL Consortium, visit www.computeexpresslink.org/join
- Download an evaluation copy of the CXL 1.1 specification
- Engage with us on social media:
  - @ComputeExLink
  - CXL Consortium
  - CXL Consortium Channel
Q & A
Thank You!