Enabling Persistent Memory

Kurtis Bowman – Gen-Z Consortium President
What Does It Mean To Be Highly Performant Remote Memory?

- Remote is a POV
What Does It Mean To Be Highly Performant Remote Memory?

- Interconnect requirements
  - Memory Semantics (No Stack)
  - Efficient Protocol
  - High Bandwidth
  - Remote Memory Controller

- It’s all about latency!
What Makes Persistent Memory Special

- Start with the obvious: It persists!
  - That means it is storage and we need to treat it as a storage device
- Availability becomes a key requirement
  - Who needs access?
  - Is it OK if it is on an island?
  - Does your application require multiple paths?
- It’s all about RAS!
Feeding Compute Cores

- Modern compute cores are hungry
  - Feeding the beasts require advanced caching strategies, multiple memory channels, and tiered memory
- Cores are steadily increasing and improving their IPC… they’re getting hungrier
- Adding DDR memory channels requires lots of pins on devices that are already pin constrained
- It’s all about Bandwidth!
Disaggregation of Persistent Memory

- Memory is expensive
  - Customers desire a pay-as-you-grow model
  - Reallocation of unused resources is a must
- Workloads require different memory characteristics
  - They may benefit from different characteristics from each memory tier
- Heterogeneous compute environments will use common memory pools
- It’s all about composability!
Other Considerations for Remote PM

- Security
  - Access protection
  - Encryption

- Scalability
  - Scales to multiple memory types
  - Scales to multiple hosts
  - Scales to multiple Terabytes, even Petabytes

- Open and Interoperable
  - Avoid lock-in and encourage innovation
It’s all about finding balance!!!
Gen-Z Delivers The Characteristics Remote PM Requires

- **High Performance**
  - High Bandwidth, Low Latency, Scalable
  - Eliminates protocol translation cost / complexity / latency
  - Eliminates software complexity / overhead / latency

- **Reliable**
  - No stranded resources or single-point-of-failures
  - Transparent bypass path and component failure
  - Enables highly-resilient data (e.g., RAID / erasure codes)

- **Secure**
  - Provides strong hardware-enforced isolation and security

- **Flexible**
  - Multiple topologies, component types, etc.
  - Supports multiple use cases using simple to robust designs
  - Thorough yet easily extensible architecture

- **Compatible**
  - Use existing physical layers, no OS modifications required

- **Economic**
  - Lowers CAPEX / OPEX, unlocks / accelerates innovation
Adding Memory Bandwidth – Feeding The Cores

8 Ch DDR4
16 Dimms
200GB/s

Future CPU

x16 GenZ
200GB/s BW

2+ x16 PCIe

Future CPU

8 Ch DDR4
16 Dimms
200GB/s

Gen-Z – PM Bridge

Gen-Z – PM Bridge

PM can be added without sacrificing main DDRx BW/Latency

DDR4 data transfer rates:
- DDR4 2133: 17 GB/s
- DDR4 2400: 19.2 GB/s
- DDR4 2666: 21.3 GB/s
- DDR4 3200: 25.6 GB/s

DDR5 data transfer rates:
- DDR5 4266: 34.1 GB/s
- DDR5 4800: 38.4 GB/s
- DDR5 5200: 41.6 GB/s
- DDR5 6400: 51.2 GB/s

Gen-Z enables platform level optimizations – swapping PCIe for Gen-Z lanes

PM can be added without sacrificing main DDRx BW/Latency
Advanced Global Shared Memory – Bring Compute to Data

Gen-Z Switch

Gen-Z – PM Bridge

Gen-Z Media Box
Advanced Global Shared Memory – Bring Compute to Data

Low Latency, High Bandwidth Fabric

Gen-Z Switch

Gen-Z Media Box

Gen-Z – PM Bridge
Advanced Global Shared Memory – Bring Compute to Data

Gen-Z Switch

Gen-Z Media Box

Gen-Z – PM Bridge

Redundant Paths
Advanced Global Shared Memory – Bring Compute to Data

Gen-Z Switch

Gen-Z Media Box

Gen-Z – PM Bridge

Fully composable
Advanced Global Shared Memory – Bring Compute to Data

Easily Scalable
Highly Secure

Advanced Global Shared Memory – Bring Compute to Data

Gen-Z Switch

Gen-Z Media Box

Gen-Z – PM Bridge
Gen-Z Enables the Balance PM Requires
Thank You