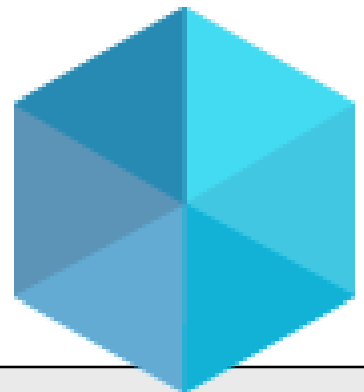


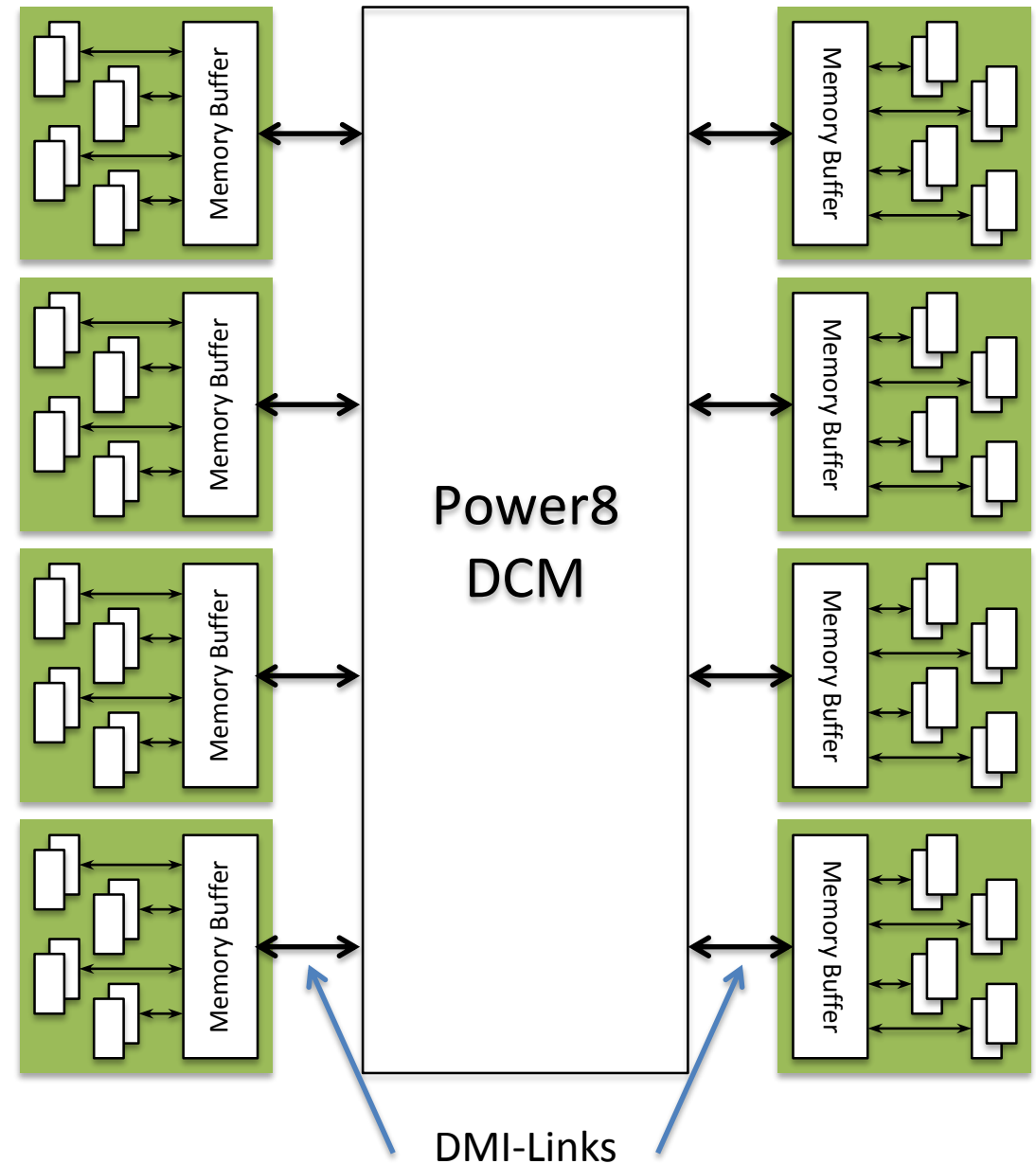


ConTutto - A flexible memory interface in the OpenPOWER ecosystem



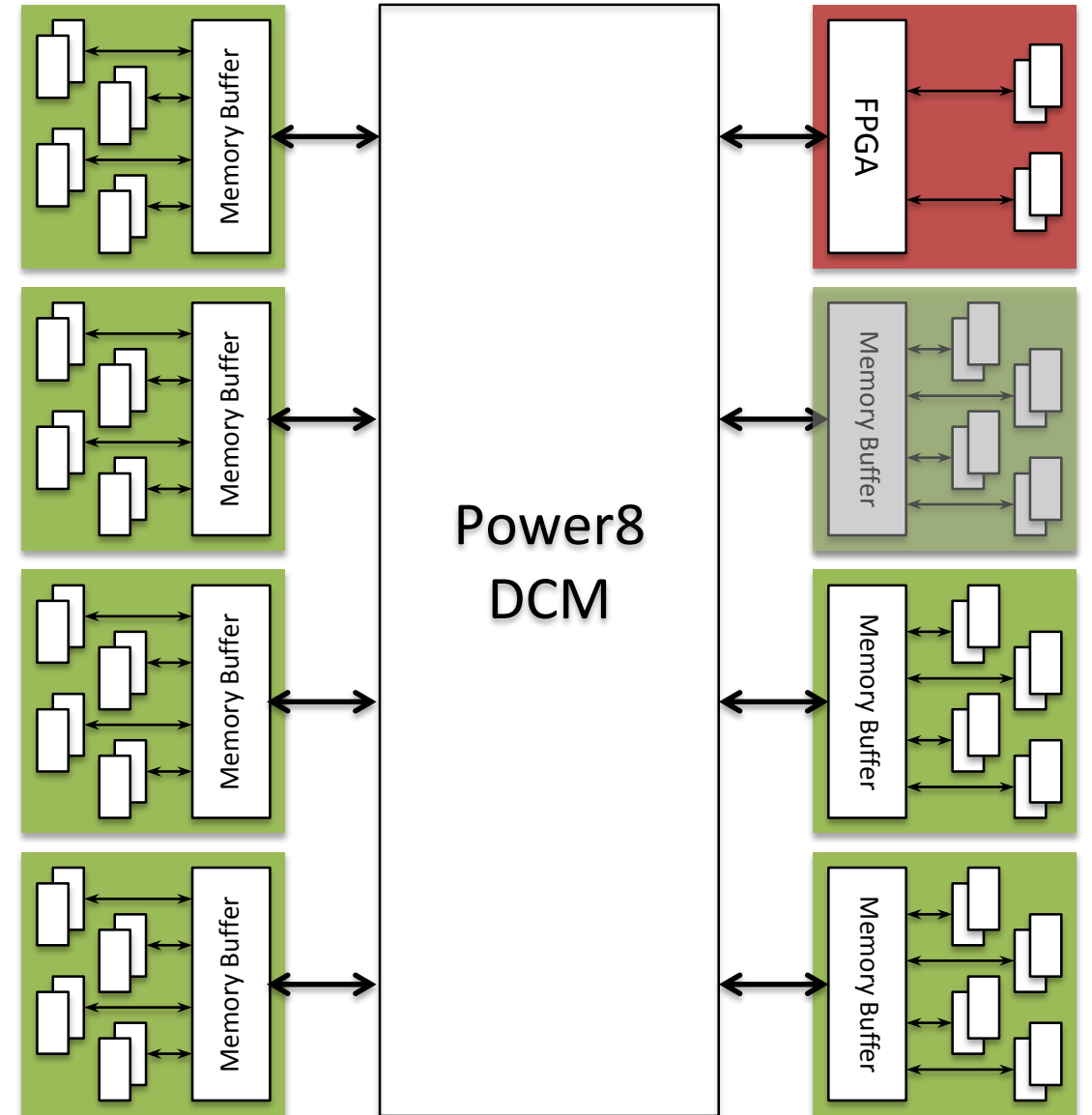
P8 Memory Sub-System

- 8 DMI links available on a P8 Dual-Chip-Module
- Differential Memory Interface (DMI) high-speed links connect to a memory buffer ASICs
- 4 memory controllers per memory buffer ASIC
- ASIC and DRAM chips are co-located on a custom DIMM (CDIMM)
- 32 memory controllers available to P8



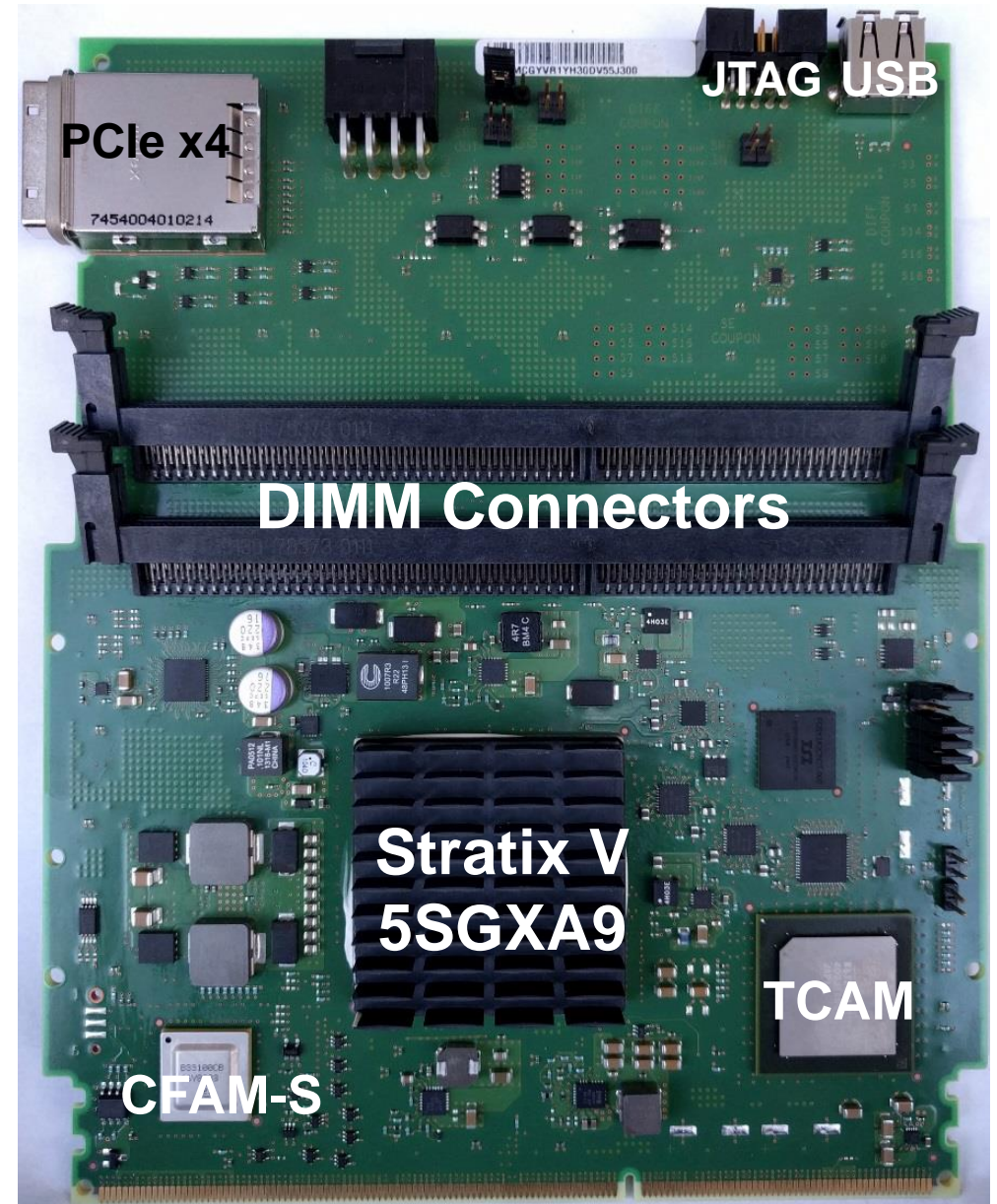
P8 Memory Sub-System with ConTutto

- Built an FPGA-based card that plugs into the DMI slot
- Enables regular system operation with any mix of CDIMMs and ConTutto cards populated
- Full compatibility with DMI protocol
- Memory controllers implemented in fabric logic and independent of DMI protocol logic
- Flexible system architecture enables easy implementation of additional features



ConTutto Card

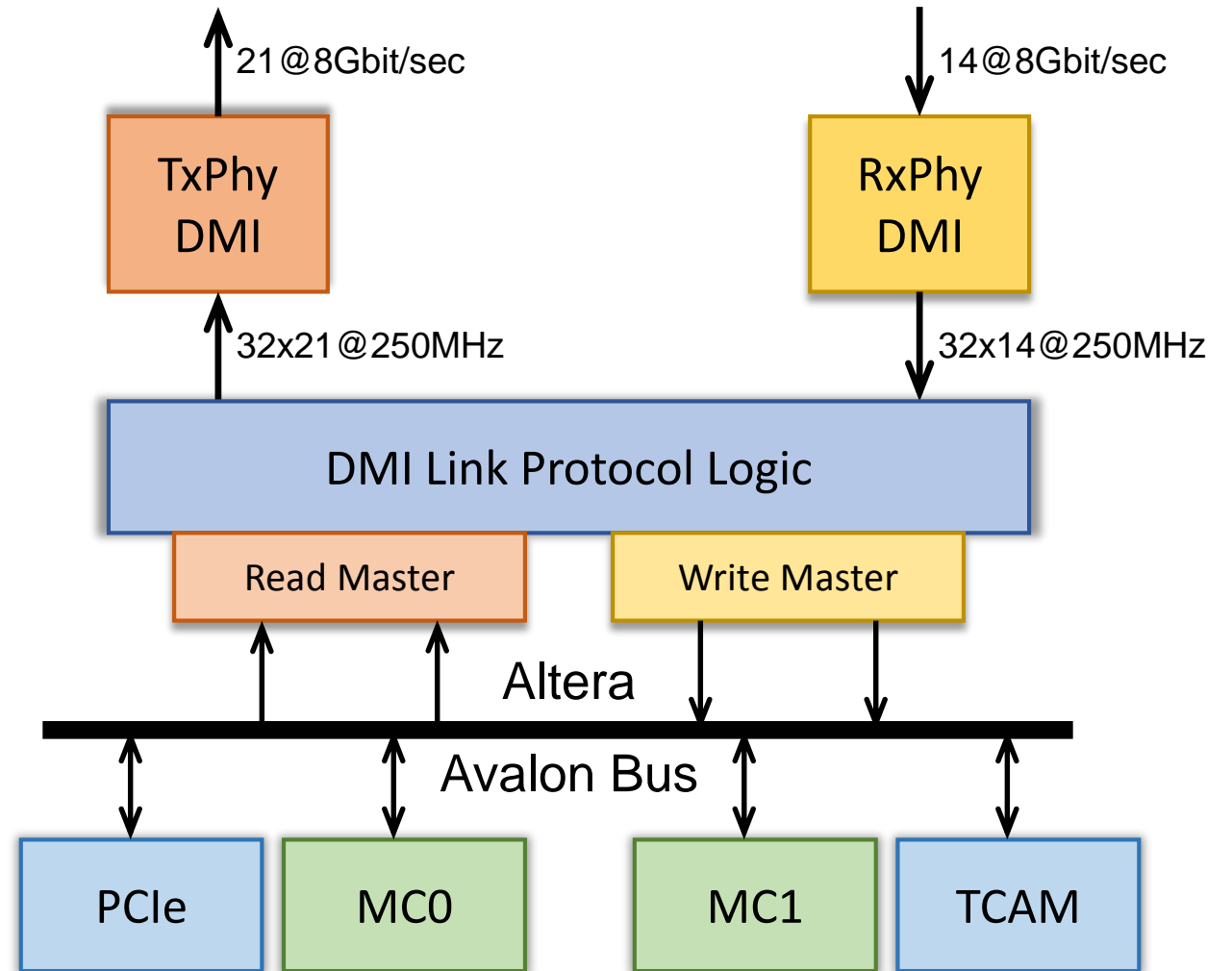
- Intended to be an experimentation and proto-typing vehicle
- Card characteristics
 - 10 signal layers
 - 10 power/ground layers
- Plug compatible with CDIMM, but 2.5" higher-- and DIMMs add width
- Large Altera FPGA with capacity for additional function incorporated
- CFAM-S (connection to service processor) enables system integration



DMI Connector

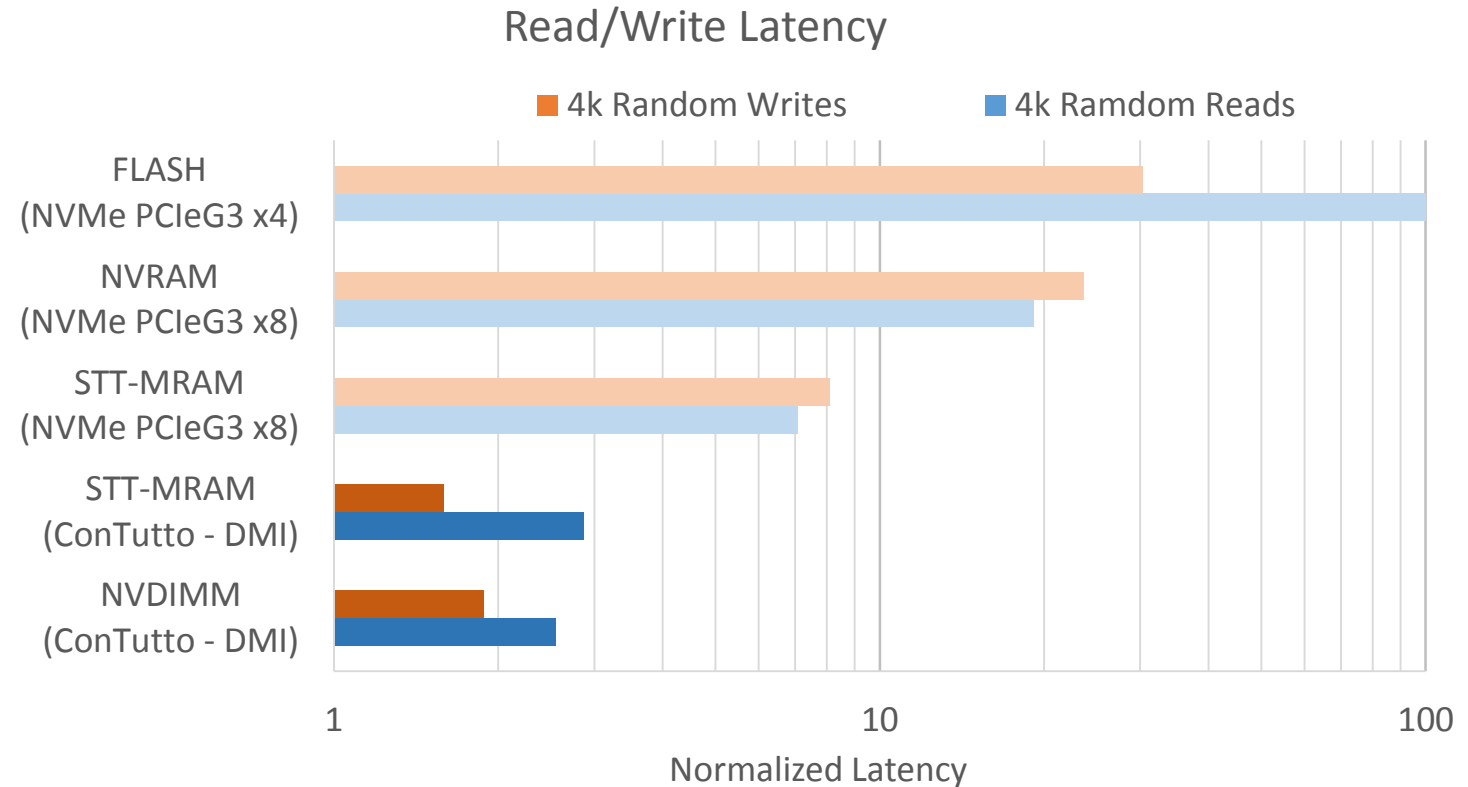
Logic design

- Support for all DMI commands implemented
- Some Memory Buffer ASIC features not available in initial version of ConTutto:
 - No DMI bus sparing or fail-over
 - No L4 cache
 - Several performance features not implemented
- 32:1 mux ratio between DMI link and FPGA fabric logic
- Using about 40% of available logic resources in Stratix V A9 device
- Avalon bus for design modularity, i.e. plug-and-play memory controller



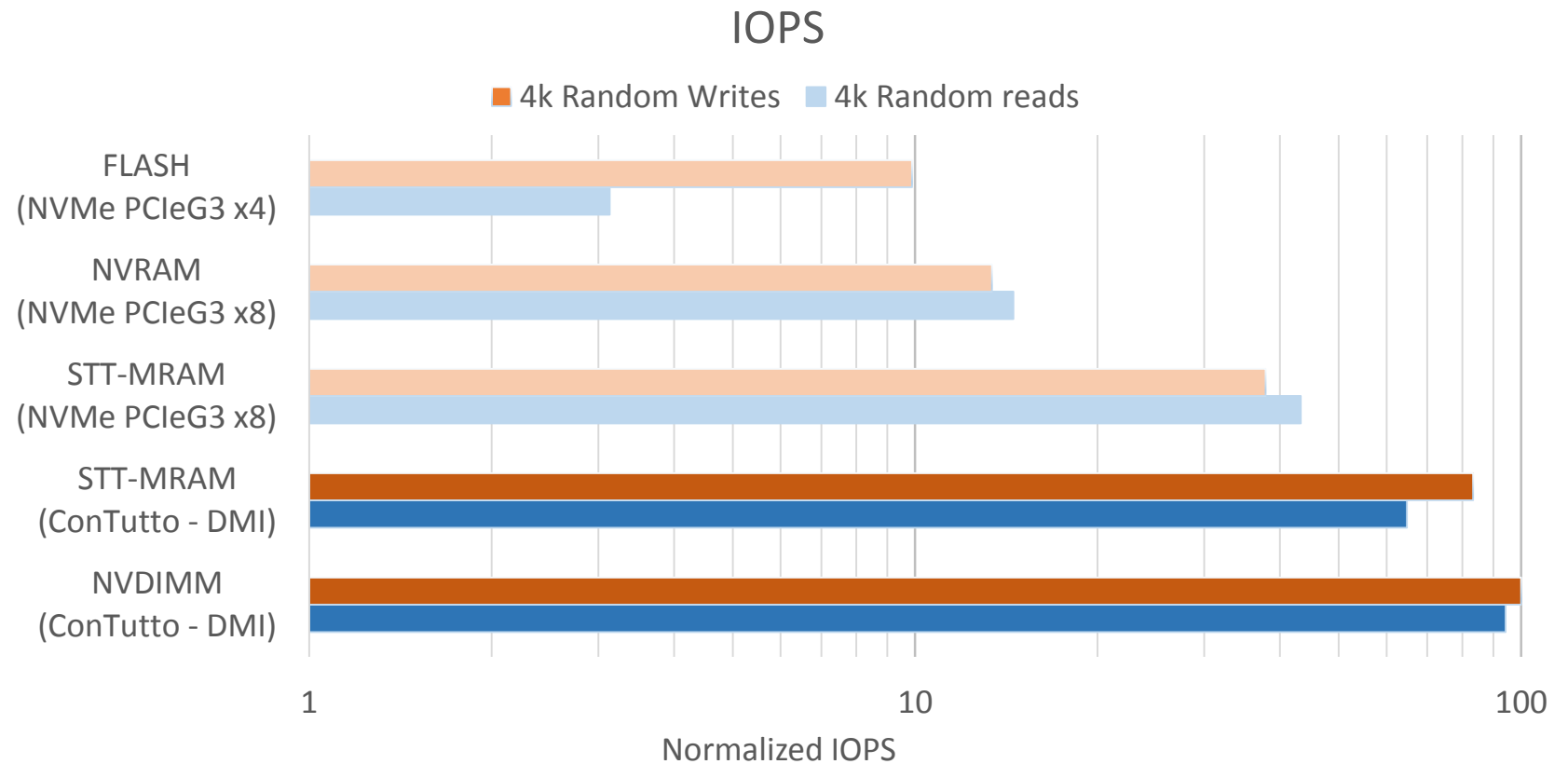
Early Test Results - Latency

- Early latency results using FIO with beta level Kernel Driver in POWER8 System
- NVDIMM-N and STT-MRAM show similar performance with ConTutto on DMI
- DMI/Memory bus is the lowest latency attach point
- FIO – Flexible IO Benchmarking



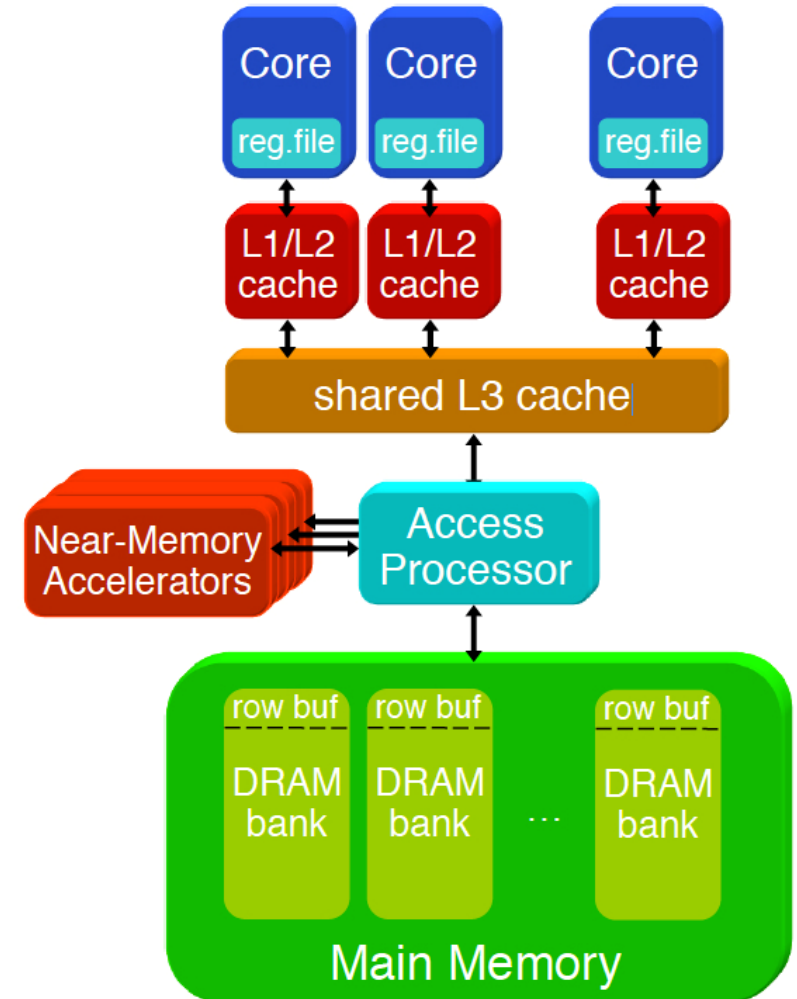
Early Test Results - IOPS

- Early IOPS results using FIO with our beta level Kernel Driver in POWER8 System
- Improved IOPS using DMI attach point
- Testing with higher iodepth and numjobs ongoing



Programmable Near-Memory Acceleration

- **Conventional computer architecture**
 - Memory system is a “slave” of the host processor
- **Novel approach**
 - Memory system **actively participates** to ensure that data is stored, accessed and transferred in the most (power-)efficient way resulting in the highest performance/Watt
 - Memory system integrates compute capabilities
- ▶ **Memory Controller → Access Processor**
- Novel **programmable** architecture
 - Enabling/differentiating technologies:
 - programmable state machine technology
 - programmable address mapping scheme
 - power-efficient “self-running” instructions
- Near-memory accelerators attach to Access Processor



Near-Memory Acceleration on ConTutto

- Ideal platform to investigate and experiment with Near-Memory Acceleration on a commercial OpenPOWER server, addressing multiple aspects:
 - design of near-memory accelerator devices
 - integration into computer system architecture
 - use of multiple devices to scale to larger storage and processing capabilities
 - programming of a hybrid system based on near-memory computing applications
- Demonstration of initial implementation of Programmable Near-Memory Accelerator concept on ConTutto for FFT computation
- Ongoing work
 - design space exploration covering device, system and application levels
 - development of near-memory computing tool set and ecosystem including compiler, debugger, performance analysis, and run-time optimization tools

Summary

ConTutto is an FPGA-based memory card that plugs into the DMI memory slot of an IBM Power8 server.

- Enables the use of different memory technologies in a Power system
 - DRAM
 - MRAM
 - NVDIMM
- Highest bandwidth & lowest latency FPGA attach point in any computer system
- Near-memory acceleration

Acknowledgments

- This work was partially funded by the Department of Energy under the FastForward-2 contract.
- Team members: Adam McPadden, Bharat Sukhwani, Kyu-hyoun Kim, Chuck Haymes, Dean Sanner, David Cadigan, Jan van Lunteren, Sameh Asaad, Dan Dreps, Hillery Hunter and many many more