Datacenter processors with OmniXtend interfaces

Paul Loewenstein,
Dejan Vucinic
Western Digital Corporation

March 13, 2019
Forward-Looking Statements

Safe Harbor | Disclaimers

This presentation contains certain forward-looking statements that involve risks and uncertainties, including, but not limited to, statements regarding: the RISC-V Foundation and its initiatives; our contributions to and investments in the RISC-V ecosystem; the transition of our devices, platforms and systems to RISC-V architectures; shipments of RISC-V processor cores; our business strategy, growth opportunities and technology development efforts; market trends and data growth and its drivers. Forward-looking statements should not be read as a guarantee of future performance or results, and will not necessarily be accurate indications of the times at, or by, which such performance or results will be achieved, if at all. Forward-looking statements are subject to risks and uncertainties that could cause actual performance or results to differ materially from those expressed in or suggested by the forward-looking statements.

Additional key risks and uncertainties include the impact of continued uncertainty and volatility in global economic conditions; actions by competitors; business conditions; growth in our markets; and pricing trends and fluctuations in average selling prices. More information about the other risks and uncertainties that could affect our business are listed in our filings with the Securities and Exchange Commission (the “SEC”) and available on the SEC’s website at www.sec.gov, including our most recently filed periodic report, to which your attention is directed. We do not undertake any obligation to publicly update or revise any forward-looking statement, whether as a result of new information, future developments or otherwise, except as otherwise required by law.
Agenda

• Enterprise datacenter RISC-V® CPU vision – all about open interfaces:
  – RISC-V® multi-core
  – NVDIMM-P memory interfaces
  – Accelerator interfaces (PCIe®, OpenCAPI™)
  – OmniXtend™ – memory protocol interface enabling memory centric architectures

• Planned open source contributions

• Western Digital first core SweRV™:
  – Microarchitecture introduction
  – Performance benchmarks

• OmniXtend™ protocol:
Vision of RISC-V open architecture datacenter CPU

It is all about open interfaces
Vision for future datacenter CPU architecture

- Multi-threaded, multi-core CPU:
  1. Medium performance, OOO RISC-V Core for general purpose OS and software applications
  2. Standardized and open JEDEC interface architecture (NVDIMM-P) for high density emerging non-volatile memories
  3. Support for high bandwidth and low latency accelerator interfaces:
     - Supporting machine learning and inference engine accelerators
  4. Support for standardized memory protocol fabric – e.g. OmniXtend - Tilelink over 802.3:
     - Allowing coherent scale-out for memory-centric architectures
Memory-centric architecture with OmniXtend

- Allows large numbers of RISC-V compute nodes to connect to universally shared memory (NUMA) – standardized and open coherence protocols
- Enables memory appliance, aggregation/disaggregation
Memory-centric architecture with OmniXtend

- Allows large numbers of RISC-V compute nodes to connect to universally shared memory (NUMA) – standardized and open coherence protocols
- Enables memory appliance, aggregation/disaggregation
OmniXtend™: direct to caches over commodity fabric

Dejan Vucinic
Emergence of memory fabric

- Memory fabric may mean different things to different people:
  - Page fault trap leading to RDMA request (incurs context switch and SW overhead)
  - Global address translation management in SW, leading to LD/ST across global memory fabric
  - Coherence protocol scaled out, global page management and no context switching
Memory-centric architecture with OmniXtend

- Allows large numbers of RISC-V compute nodes to connect to universally shared memory (NUMA) – standardized and open coherence protocols
- Enables memory appliance, aggregation/disaggregation
Memory-centric architecture with OmniXtend

- Allows large numbers of RISC-V compute nodes to connect to universally shared memory (NUMA) – standardized and open coherence protocols
- Enables memory appliance, aggregation/disaggregation
OmniXtend memory-centric fabric architecture

- Replaces Ethernet L2 with serialized TileLink messages
  - Keeps standard 802.3 L1 frame, interoperates with Barefoot Tofino and future OTS Ethernet switches
  - Custom frames are parsed and processed in P4 language
  - Enables stateful message processing inside the switching fabric
  - Supports innovation required for RAS
  - FPGA or ASIC switch; not limited to 802.3

- Protocol translation and modification inside fabric:
  - Requires no new silicon

- 100 Gb/s is available today
  - Clear roadmap to 200 and 400 with 56Gb PAM4 and x8
OmniXtend 1.0: the first ever two-socket RISC-V

- 50 MHz SiFive U54 on Xilinx UltraScale+ FPGA boards (VCU118)
- Live demo at RISC-V Summit in December 2018
- Binaries on github
OmniXtend 1.0 direct connect latency

![Graph showing average cache line access latency with test size (bytes) on the x-axis and latency (CPU cycles) on the y-axis. The graph compares NUMA local memory, NUMA remote memory, and single socket performance.]
Next: routing OmniXtend through programmable switch

- Barefoot Tofino® ASIC:
  - 64-port 100 GigE switch, 6.4 Tbit/s aggregate throughput, < 400 ns latency
  - Supports P4 HDL, successor to OpenFlow enabling protocol innovation
  - Describe TileLink message format in P4
  - Match-Action Pipeline (a.k.a. “flow tables”) enables line-rate performance
  - Modifications to coherence domains, protocols require no new silicon

- FPGA switch for more exotic experiments (Xilinx SDNet™, P4FPGA, etc.)
  - Eventually custom silicon for optimized latency/radix
Memory fabric protocol OmniXtend innovation platform