The Changing Landscape of Compute
with RISC-V Open Architecture and P4

Richard New, VP Research
Western Digital
Forward-Looking Statements

Safe Harbor | Disclaimers

This presentation contains forward-looking statements that involve risks and uncertainties, including, but not limited to, statements regarding our products and technologies, business strategies, product development efforts and growth opportunities, emerging storage and memory technologies, our investments in and contributions to the RISC-V ecosystem, industry and 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.

Key risks and uncertainties include volatility in global economic conditions; business conditions and growth in the storage ecosystem; impact of competitive products and pricing; market acceptance and cost of commodity materials and specialized product components; actions by competitors; unexpected advances in competing technologies; our development and introduction of products based on new technologies and expansion into new data storage markets; risks associated with acquisitions, mergers and joint ventures; difficulties or delays in manufacturing; and other risks and uncertainties listed in the company’s 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 required by law.
A Leading Data Infrastructure Company

PORTFOLIO BREADTH

TECHNOLOGY ENGINE

~14K active patents

GLOBAL SCALE

~62K employees worldwide

CUSTOMER VALUE

Western Digital. WD G-Technology SanDisk
Cache Coherence

Multi-Core CPU

PCIe Root Complex
- Core
- Cache
- Shared Cache

Home Agent and Memory Controller

Network Interface
- DRAM (DDR)
- CPU
- SSD
- HDD
- PCIe

System PCIe Root Complex

CPU Socket 1
- CPU 1
  - Core 1
    - Control
    - ALU
    - L1 Cache
  - Core 2
    - Control
    - ALU
    - L1 Cache
  - L2 Cache
- Coherency

CPU Socket 2
- CPU 2
  - Core 1
    - Control
    - ALU
    - L1 Cache
  - Core 2
    - Control
    - ALU
    - L1 Cache
- L2 Cache
- DRAM
Cache Coherence Protocols

• Enables multiple processors to meaningfully share the same memory
• Coherence protocol ensures that all the processors see synchronized copies of data
• Key building block for creating a useful memory consistency model
• Enables single operating system to run across multiple cores

MESI Protocol

- **Invalid (I)**: Cache line is invalid (unused).
- **Shared (S)**: Cache line may be stored in other caches and is clean.
- **Exclusive (E)**: Cache line present only in the current cache, and is clean.
- **Modified (M)**: Cache line present only in the current cache, and is dirty.

**States:**
- **PR** = processor read
- **PW** = processor write
- **BR** = observed bus read
- **BW** = observed bus write
- **S/~S** = shared or not shared

**Transitions:**
- **PR/S**
- **PR+PW**
- **BR+BW**
- **BW**
- **PW/~/S**
- **PW/~S**
Other Alternatives

**Memory Extension**
Page fault trap leading to RDMA request (incurs context switch and SW overhead) (e.g. mmap across network)

**Software Address Translation**
Global address translation management in SW, leading to LD/ST across global memory fabric (e.g. LD/ST to MMIO PCIe device that translates address to network destination).

**Distributed Software**
Abandon the concept of shared memory, and write distributed software which explicitly passes messages back and forth or copies data with RDMA between nodes.

Context switch cost comparable to remote memory access latency; wasted wall power

 Require software/kernel support and/or rewriting of applications
RISC-V Open Instruction Set Architecture

Open
- Completely open source ISA that can be applied for any modern computing devices

Free
- Develop independent IP
- RISC-V Foundation does not charge authorization fee
- BSD license allows developers to decide if they want to share their own work

Safe
- Clear separation between user and privileged ISA

Regulation
- Prevent fragmentation of RISC-V in China
- Small standard base ISA with only 40+ base instructions

Extensibility/Specialization
- Variable-length instruction encoding
- Vast opcode space available for instruction-set extensions

Stable
- Base and standard extensions are frozen
- Additions via optional extensions, not new versions
Western Digital SweRV Core™

• First production-grade open source RISC-V core
• 2-way, superscalar, in-order core with 9 stage pipeline
• 1 GHz operation @ 28nm
• Support for RV32IMC
• Ideal for high-performance embedded applications

Download it at: https://github.com/westerndigital corporation/swerv

Western Digital ships in excess of 1 Billion cores per year... and we expect to double that.
SoC with OmniXtend™ Support

Western Digital and SiFive collaborating on extension of TileLink cache coherency protocol over ethernet.

- Leverages ubiquitous Ethernet
- Routable and switchable fabric
- Scalable as Ethernet speeds increase

OmniXtend initial specification and FPGA bitstream are now live at https://github.com/westerndigital.corporation/omnixtend
OmniXtend Ethernet Implementation

- Replace L2 Ethernet with Tilelink on top of 802.3 L1
  - Data plane frame processing programmed in P4.
  - Supports innovation required for RAS
- Barefoot Tofino™ ASIC
  - 256-lane 25 GT/s Ethernet switch, 6.4 Tbit/s throughput
  - Supports P4 HDL, successor to OpenFlow
  - Match-Action Pipeline enables line-rate performance
OmniXtend Architectures

Two Servers Point to Point

Multiple Devices Through an Ethernet Switch

Looks like single machine with more CPUs
First OmniXtend 0.1.1 measurements

Average cache line access latency (50 MHz uncore)

- NUMA Tofino remote memory
- NUMA Tofino local memory
- NUMA direct remote memory
- NUMA direct local memory
- single socket

Latency (microseconds)

Latency (clock cycles)

Test size (bytes)

L1 Cache Hits

L2 Cache Hits
A Truly Open Memory Fabric -- OmniXtend

- Data is the center of the architecture
- No established hierarchy – CPU doesn’t ‘own’ the GPU or the Memory
- Cache Coherence preserved system-wide over the Fabric
Contacts

• Check out RISC-V as a candidate for embedded networking cores. Join RISC-V Foundation and Chips Alliance.
  https://riscv.org/
  https://chipsalliance.org/

• Western Digital SweRV core on github:
  https://github.com/westerndigitalcorporation/swerv

• If interested in memory-fabric applications of P4 and programmable switch, lease contact Western Digital, or see github site.
  https://github.com/westerndigitalcorporation/omnixtend
Thank You

Western Digital