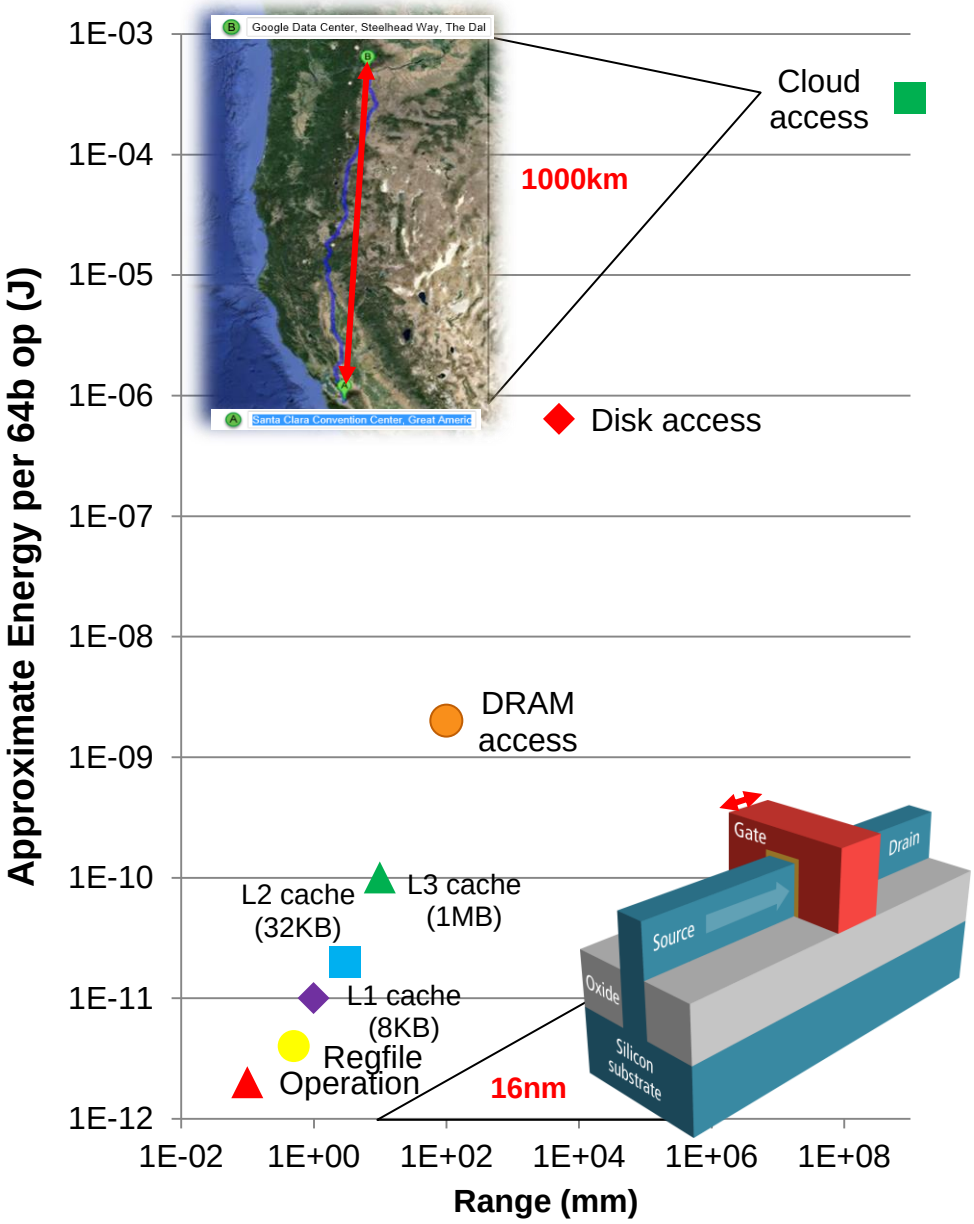
Flexible Architectures for Scalable Flash Controllers

Chris Rowen
CTO - IP Group

August 2014

cadence®

Where's My Data?



1. Energy per operation scales with distance: primitive op to cloud access
2. $>10^{10}$ energy range drives system architecture: Move computation closer to data
3. Emergence of Dataplane Processor Units (DPUs)
4. Cloud data aggregation [disk farm] vs personal data aggregation [SSD]
5. Flash has unique role in capacity vs. energy
6. Flash potential not yet exploited
   - Bandwidth
   - Transaction rate

Why? Need faster "smarts"

cadence

Processing My Data!
Proliferation of Dataplane Processor Units (DPUs)

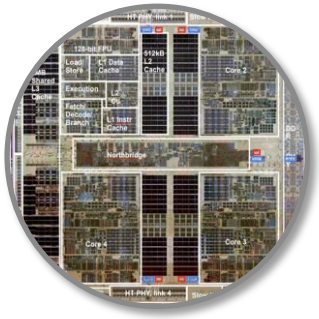
- CPU : **Central** Processing Unit
- GPU : **Graphics** Processing Unit
- NPU : **Network** Processing Unit
- Various DSPs
  - APU : Audio Processing Unit
  - VPU : Vision Processing Unit
  - WPU : Wireless Processing Units

SPU : **Storage** Processing Unit

Dataplane Processor Units (DPUs)

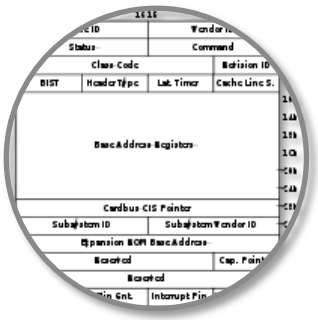
What are the Ingredients for Storage Success?

Solutions to fit customer needs



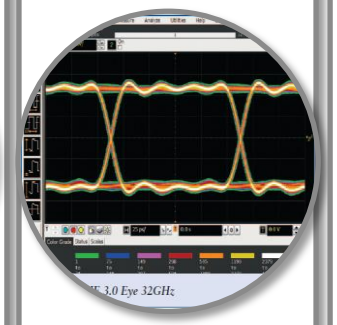
Storage Processors

- Tensilica dataplane processors
- Fully C programmable
- Optimized instruction set and interface capability
- Widely used architecture in storage systems



Controllers

- Async
- ONFI 1, 2, 3, 4 Toggle 1, 2, 3
- BCH ECC
- Advanced Features



PHYs

- ONFi 1, 2, 3, 4 and Toggle 1, 2, 3
- Soft Deliverable
- Supports all Fabs and technologies
- Scalable from 1 to >20 lanes

Software

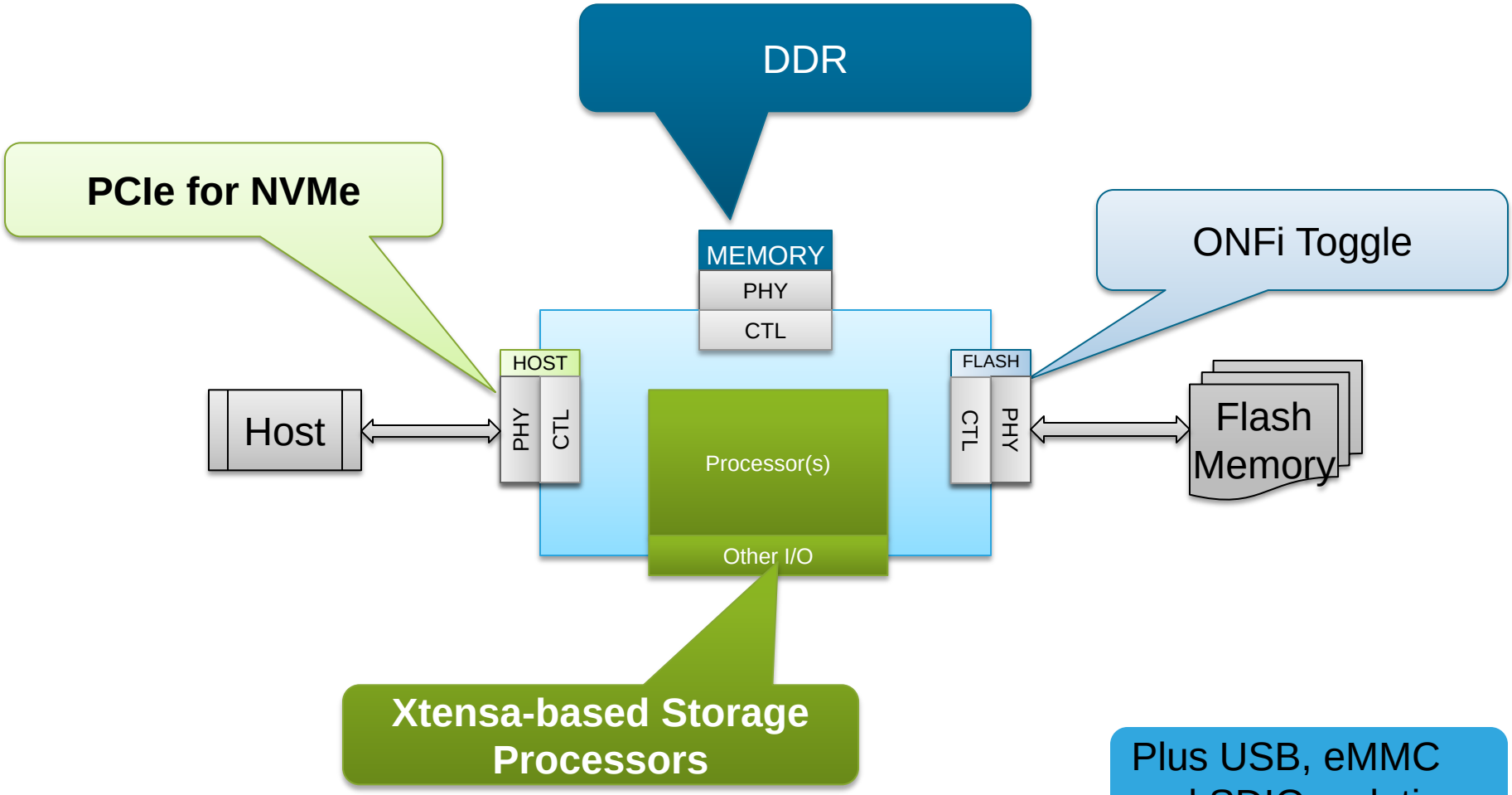
- Core Firmware Driver
- Reference Linux Driver
- Virtual Reference Platform

Integrated Solutions

- Controller and PHY integrated and verified

Photo courtesy of PC World

Storage Processors Unify Subsystem Design
Complements Cadence Storage Interface IP



cadence

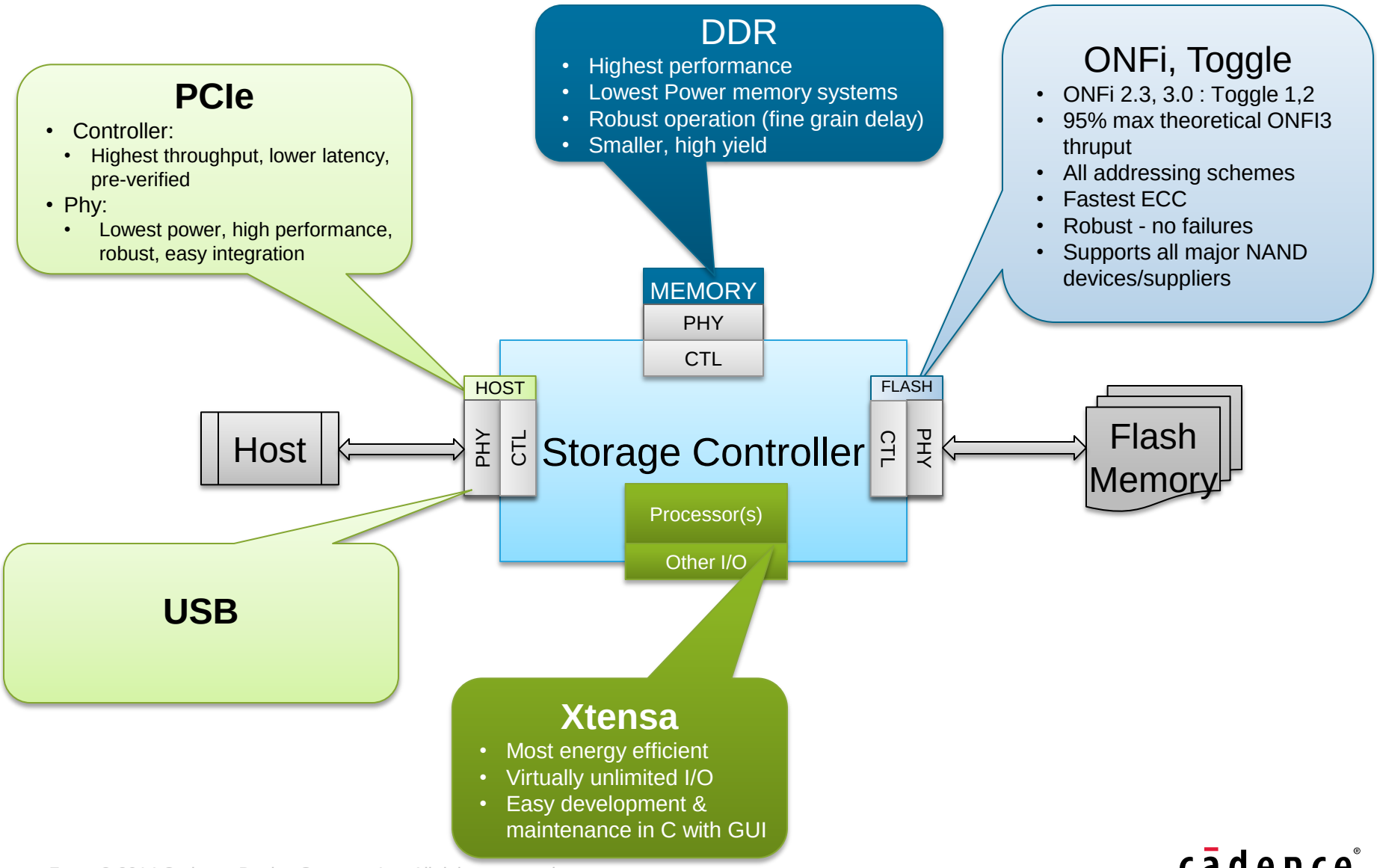
Issues in Storage System Design

Needs
Flexibility to support complex, evolving storage algorithms
Scalability in capacity, bandwidth and transaction rate
Robustness and data integrity
Silicon cost
Time to market for differentiated platform

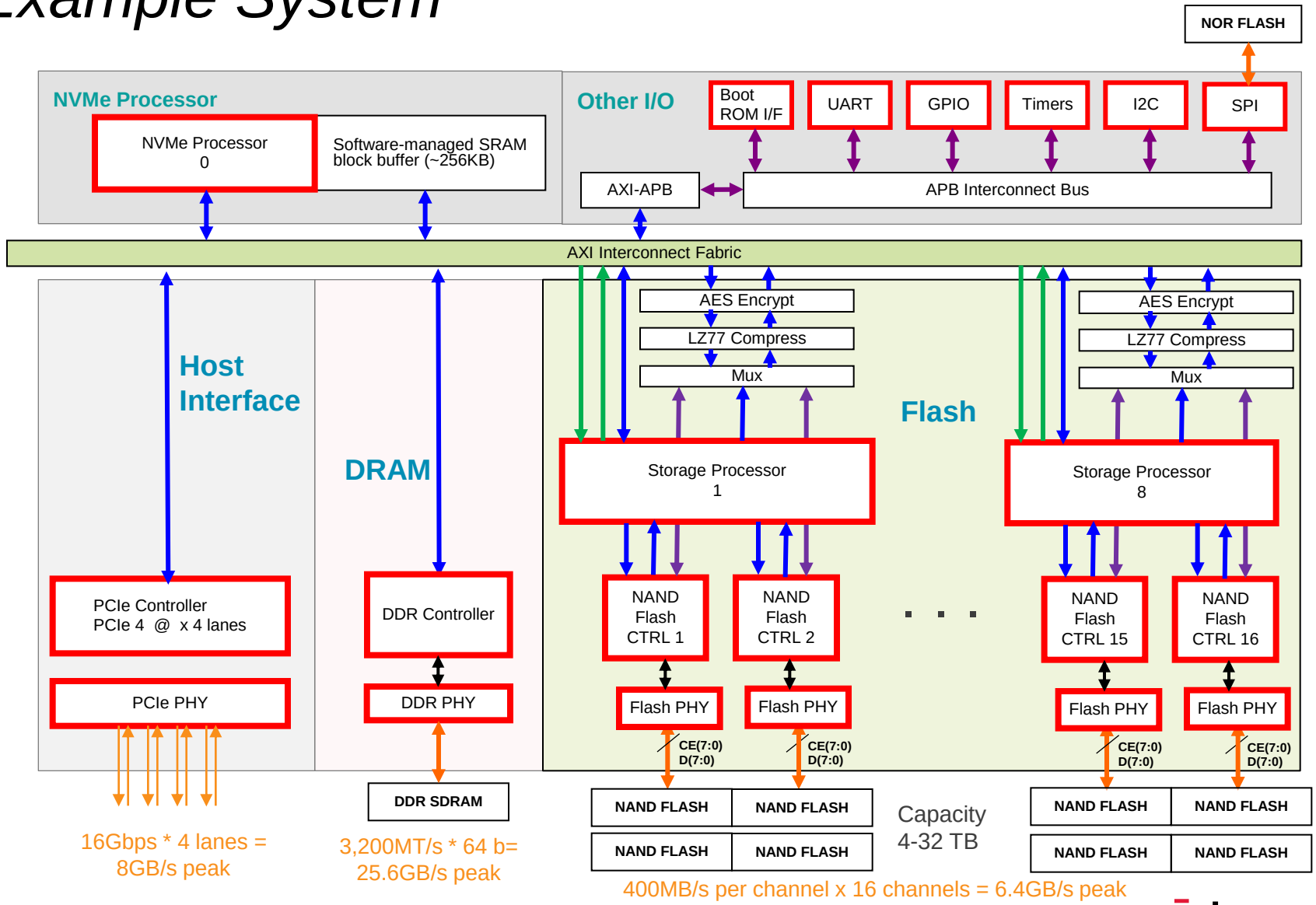
Key Questions

- How can we scale interface performance?
- How does processor handle complex protocols?
- How do you improve both time-to-market and differentiation?

Cadence Storage IP



Flash Systems with Smart Storage Processors
Example System



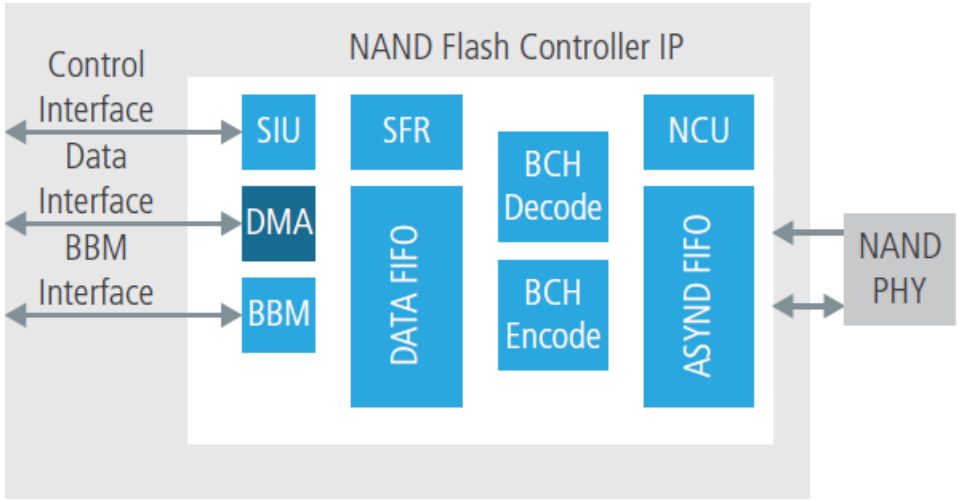
cadence

Example System Parameters

Parameter	Value
16 OnFi 3 channels @ 400MB/s	6.4GB per second
4KB logical block size	Up to 1.6M logical block requests/sec
Up to 4 devices per OnFi channel @ 4Tb per device	32TB capacity
Total logical blocks to map in system	8B blocks
Flash block organization	64 pages of ~2KB = 128K flash block
Processor clock	1000MHz
Flash page service time	2KB at 800MB/s = ~400 cycles/block

Cadence Complete Ingredients: ONFi NAND Flash Controller + PHY

- Supports all major NAND devices & interfaces
- Fully integrated controller and PHY
- Integrated or system-side Data DMA
- FlexBCH: Full optimized hardware with flexible correction strength and codeword size selection
- Comand Processing:
  - Interleaving and queuing with context saving
  - Repeating commands through Super Sequence
  - Support for undocumented instructions
- Boot from flash preconfigured with pin strapping
- Hardware-assisted Bad Block Managment
- Full Software Deliverables:
  - Driver
  - Wrapper
  - Diags
  - Demo



High performance architecture enables 95% of maximum theoretical performance of ONFi3 interface

Cadence Complete Ingredients

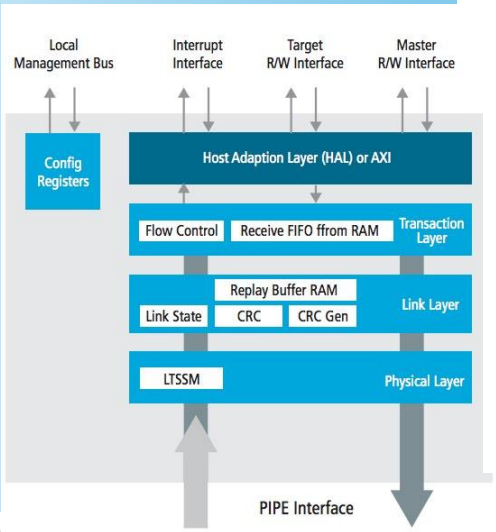
Cadence PCI Express® Controller IP + PHY

PCIe 1,2,3,4 Controller:

- Best in class performance, >95% Link utilization
- Scalable design: 1 to 16 lanes
- Proven Virtualization and Bifurcation capabilities

- **Efficient and Proven**
  - Energy efficient designs with low-power ECNs (L1 substates, OBFF, LTR, ASPM Optionality)
  - In Silicon with multiple PHYs, customers certified
- **PCI Specifications:** PCI Express v1.1, 2.1, 3.0, and 4.0 (Beta), Intel PIPE v3.0 and 4.x, SRIOV v1.1

- **IP Products:**
  - Root Ports, End Points, Dual mode
- **Integrated Solution**
  - PCIe Controller
  - Proven PHYs
  - Software: Core Driver/Linux Reference Driver



PCIe Gen 4-ready PHY with FinFET Realization

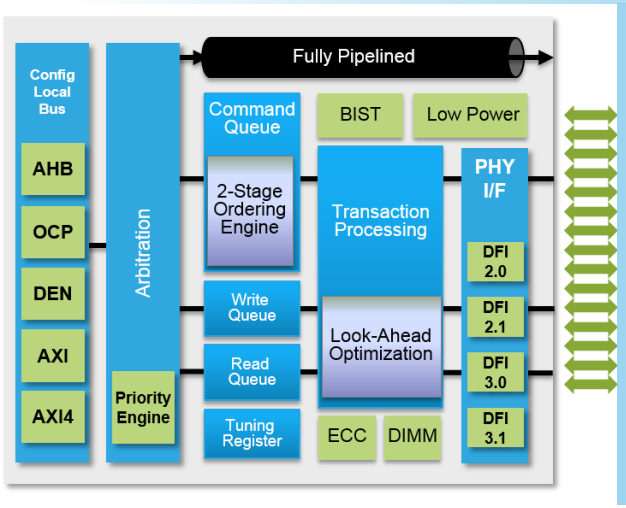
- **Less power, less jitter, better data recovery**
  - Power saving with leakage optimizations
  - Combined LC-VCO and RO PLLs
  - Supports >30 dB channel loss
- **Compliant Standards**
  - PCIe v1.1, 2.1,3, and 4; Intel PIPE 4.0
  - 10G-KR, XAUI, QSGMII/SGMII
- **Key Features**
  - 5 Tap DFE with adaptive CTLE and offset correction
  - Auto-calibration of on-chip termination
  - x1- x16 configurations
  - Scalable PIPE width w/ standard AMBA-APB interface
- **IP Readiness:** Test chip T/O done

cadence®

Cadence Complete Ingredients:
Cadence DDR Controller IP + PHY

Leading DDR Controller

- 1000+ configurations
  - Arbitration and priority engines
  - Supports multiple clients, bus masters & widths
  - Performance analysis and tuning
- Low Risk Solution
  - Verified with industry leading memory models
  - Interoperability leadership : DFI standard
  - Silicon validated with PHY and DRAM
  - First to market: DDR4, LPDDR3, and Wide IO



- Protocols
  - DDR 4/3/3L/2/1,
  - LPDDR 4/3/2/1,
  - Wide IO
  - DFI 4.0, 3.1, 3.0 and 2.1
  - AHB, OCP, DEN, AXI3, AXI4

Robust Silicon-Proven PHY

- First to market: DDR2400, WideIO, LPDDR4
- Complete jitter analysis and char reports
- Verified with leading memory models
- System Design-In Kit : PCB modeling: SI/PI analysis

- High Speed PHY - Silicon Proven
  - Scalable to 3200 performance
  - Per bit de-skewing, per rank leveling
  - DDR4/3/3L, LPDDR4/3/2 support
  - DDR2400 silicon characterization
- LP PHY- High volume production
  - Up to 1600 performance
  - Lowest power/area, integrated hard control
  - DDR3/3L/2 LPDDR3/2 support
- Fit IP to chip, not other way round
  - 8 bit hardened slices + IO + PLL integration
  - Flexible IO ordering, bump pitch, form factor
  - Custom, fully hardened PHY option

Track record of 1500+ deployments

cadence

Cadence ingredient that binds it all together:

Tensilica automated hardware/software synthesis

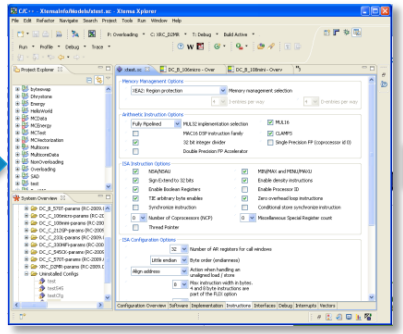
Architects' needs:

- **Faster command queue processing**
- **More memory bandwidth**
- **Hide latency to DDR**
- **Rapid search of huge mapping tables**
- **Continuous software upgrade of EVERYTHING**

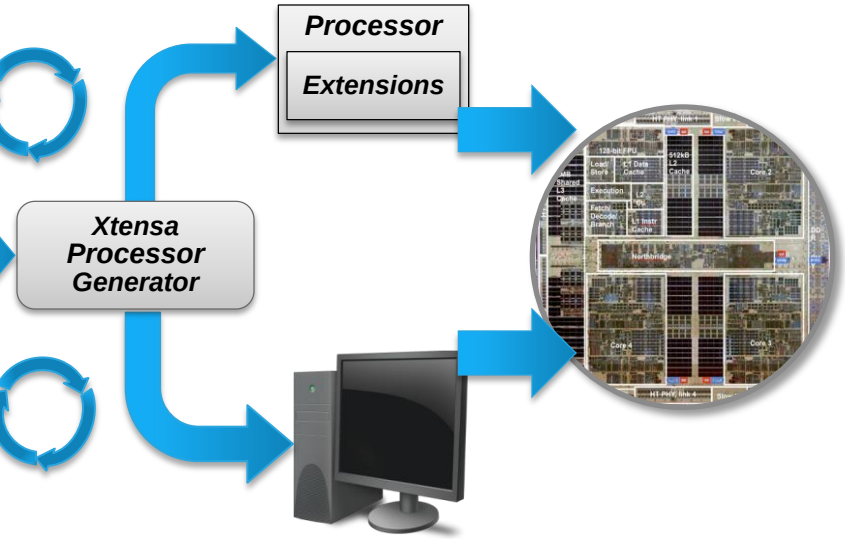
Full hardware design

Source pre-verified RTL, EDA scripts, test suite

Processor configuration



1. Select from menu
2. Explicit instruction description (TIE)



Customized software tools

C/C++ compiler Debuggers, Simulators, RTOSes

No manual intervention needed in hardware or software completeness

cadence

Increasing Computational Throughput

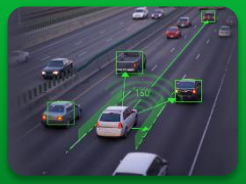
In Xtensa Processors – more choices

Add new instructions

- Raise overall performance
- Accelerate critical functions
- Often over 10x performance improvement

Improve register availability

- Add more registers for the compiler to use
- Set custom widths, up to 1024bits, to keep area as low as possible

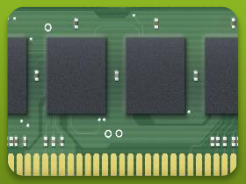


Execute multiple instructions at once

- Arbitrary new VLIW instruction bundles
- Automatically used by compiler
- A general purpose performance improvement

Increase local memory bandwidth

- Multiple Load/Store units
- Up to 1024bits total per cycle



DMA and direct-connect to system memories and logic

- Overcome bandwidth/latency bus bottlenecks
- For Command Queues/Code overlays etc.
- No processor cycles required to move data

Tasks for Flash Management Processing

Mapping

- For each logical block in file systems, find flash block
- Update usage statistics as accessed
- Large-scale hash tables for speed and scalability
- **Instruction Set: Fast Hash, Table search with >2^{32} entries**

Write management

- Cache or temporarily write data in available unwritten pages
- Migrate data to reduce fragmentation
- Reclaim deleted data
- **Instruction Set: Fast data movement**

Wear leveling

- Track lifetime writes per flash block and write less-used, blocks
- Track read-disturb – accumulate reads since last write
- Rewrite [move] data to refresh data
- **Instruction Set: Parse and update data-structures**

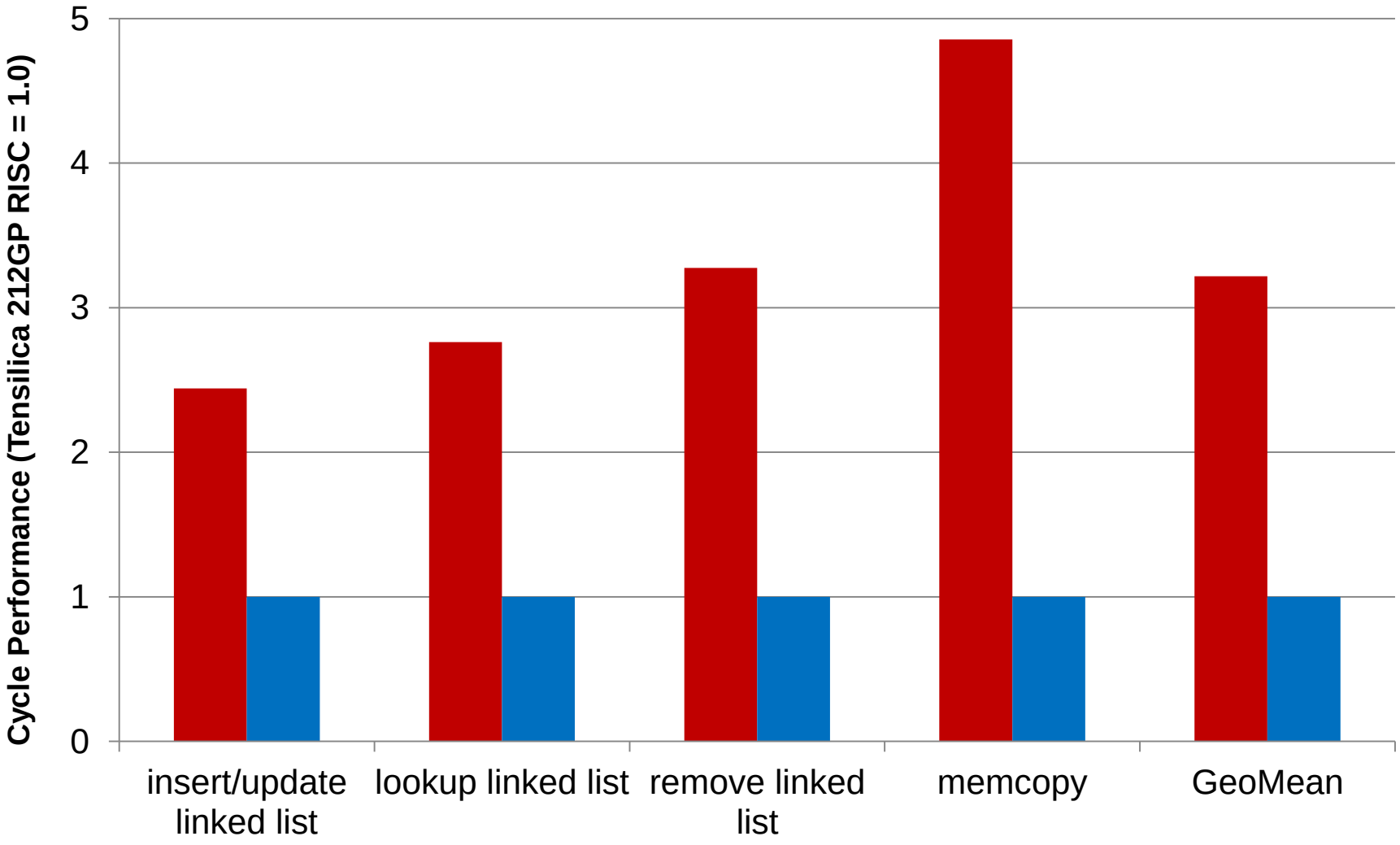
Command processing

- Requests: For each logical read/write, determine and schedule sequence of reads and writes required
- Response: On completion, reply to host with data/status
- **Instruction Set: Command parsing/dispatch, direct queue I/O**

Example:
Instruction Set for Storage

- Built on widely popular Xtensa processor family
- Implemented in Tensilica Instruction Extension (TIE) code
- Add:
  - 64b register file
  - 64b non-blocking read/write queue
  - 64b x N entry local tag cache
  - Parity protection on added register files and queues
  - 3-way instruction issue in 64b VLIW instruction width
- Operations
  - 64b load/store including address update load/store
  - “parsing” load and “packing” store operates on up to 3 32b, 16b, 8b fields
  - chaining loads checks tag match and null pointers
  - Non-blocking read and write requests to memory: 64b address, hashing
  - Tag cache management operations: lookup, find_victim, check_dirty
  - Multiple issue slots for bit field insertion and extraction

Generalized List Processing Instruction Set Gives Significant Performance Boost



Wrap-up

- All the key building blocks for high-performance intelligent storage
- Cadence has the state-of-the-art design ingredients: controllers and PHYs
  - ONFi
  - DDR
  - PCIe
- New category of DPUs - Storage Processor Unit raises performance by an order-of-magnitude
- Performance AND generality unleashes system architect creativity
- 40 leading companies already using Cadence flash-specific IP in major SOC platforms


cadence®