Flash Memory Summit

Annual Flash Controller Update

David McIntyre

DSMcIntyreConsulting@gmail.com

Text **FMS** to (408) 772-7044

Overview

- Data Center Drivers
- Memory Hierarchy Drivers
- Current Flash Controller Challenges
- Supporting Technologies

Flash Memory Summit

Data Center Trends

- **Hyper Converged Infrastructure**
  - Integrated Compute/Storage/Networking
  - Massive interconnectivity
  - Good for Exchange, Oracle, SQL databases
  - Software managed virtualized resources

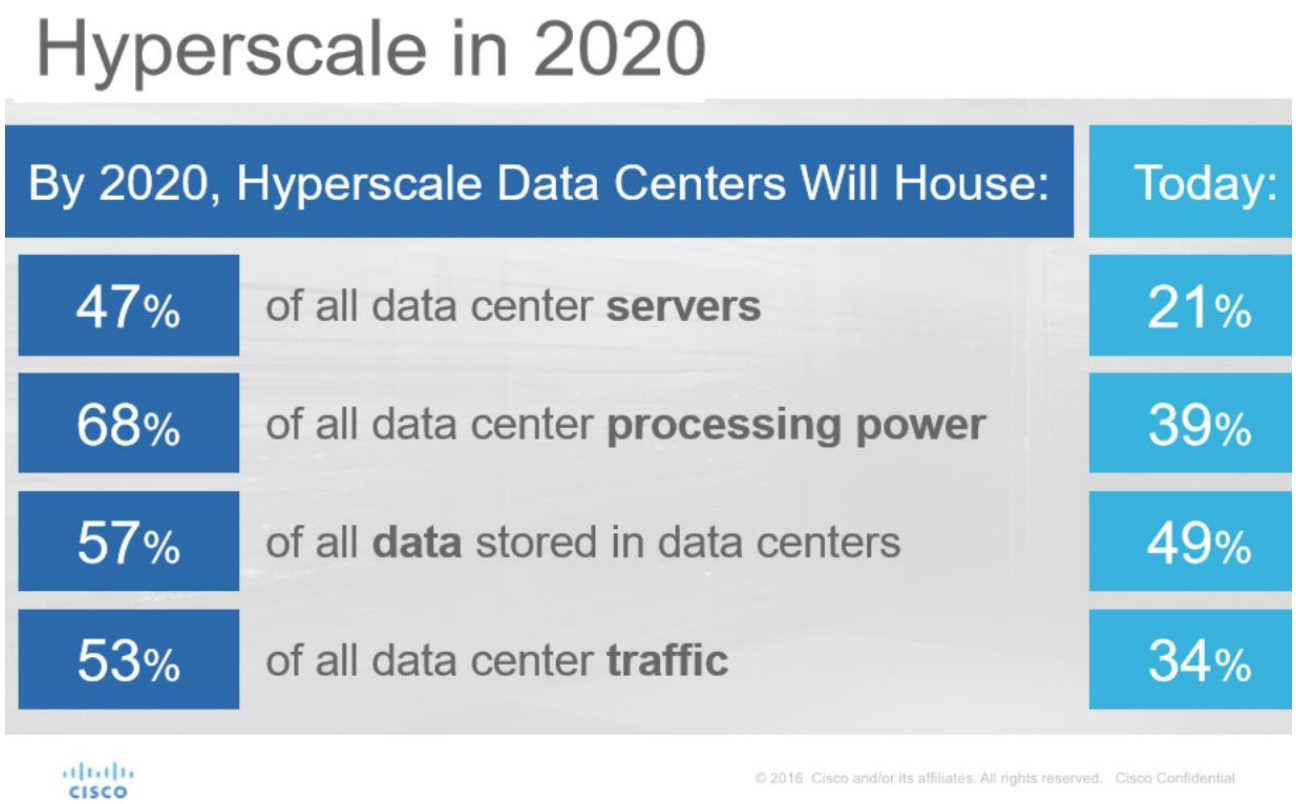
- **Hyper Scale**
  - Independent scaling of compute and storage resources
  - Good for elastic workloads, e.g. Hadoop, NoSQL
  - Also software managed

Flash Memory Summit

Data Center Trends

- **Storage**
  - Convergence of RAM/cache and SCM**.**
  - All flash and hybrid arrays
  - Persistent memory cache
- **Compute**
  - GPU, TPU and FPGA accelerators
- **Networking**
  - Low latency, high performance RDMA networks
- **Hybrid Cloud**
  - For lease and on premises-equipment
  - Deployment Options, e.g. OpenStack and Docker

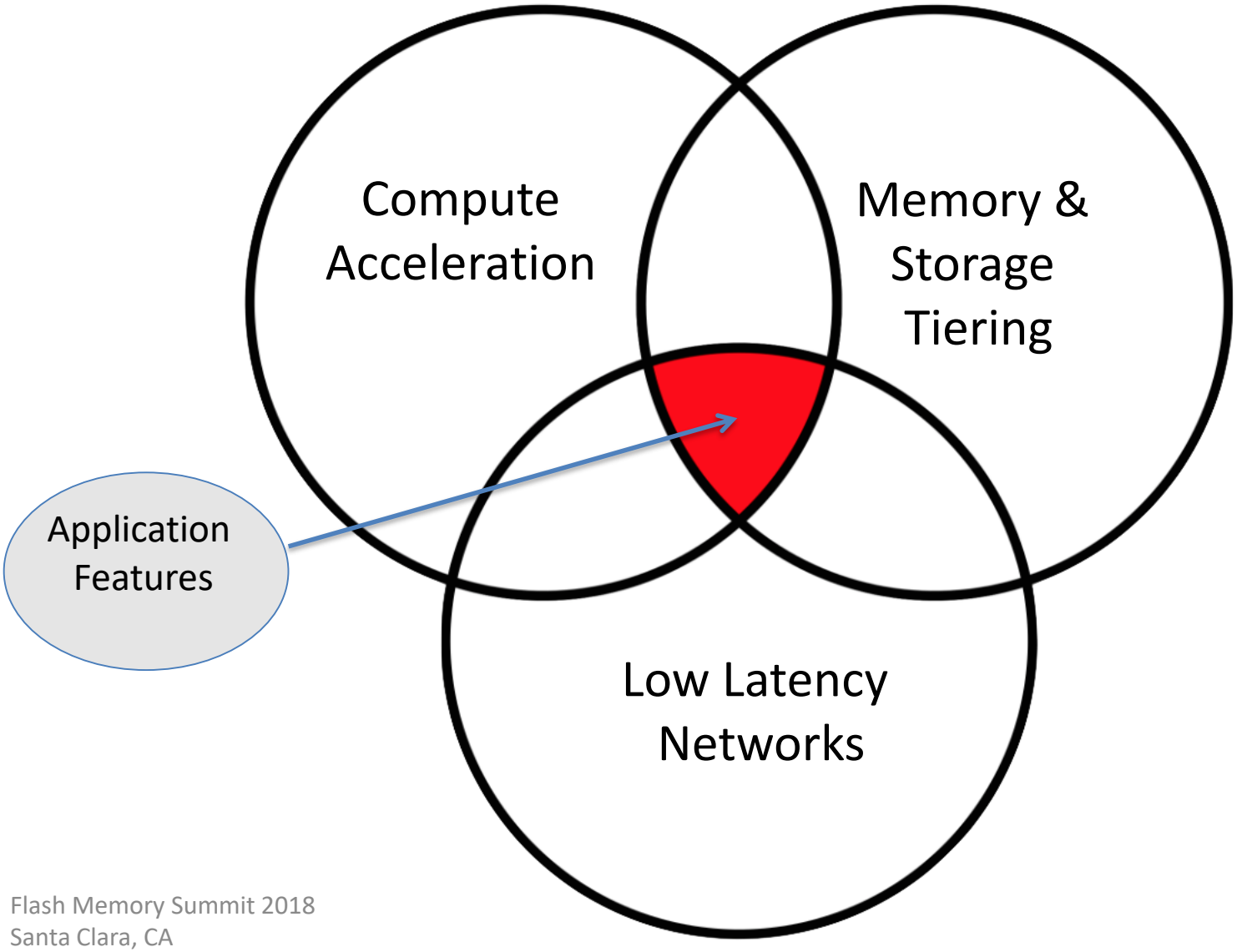
Hyperscaler Priority



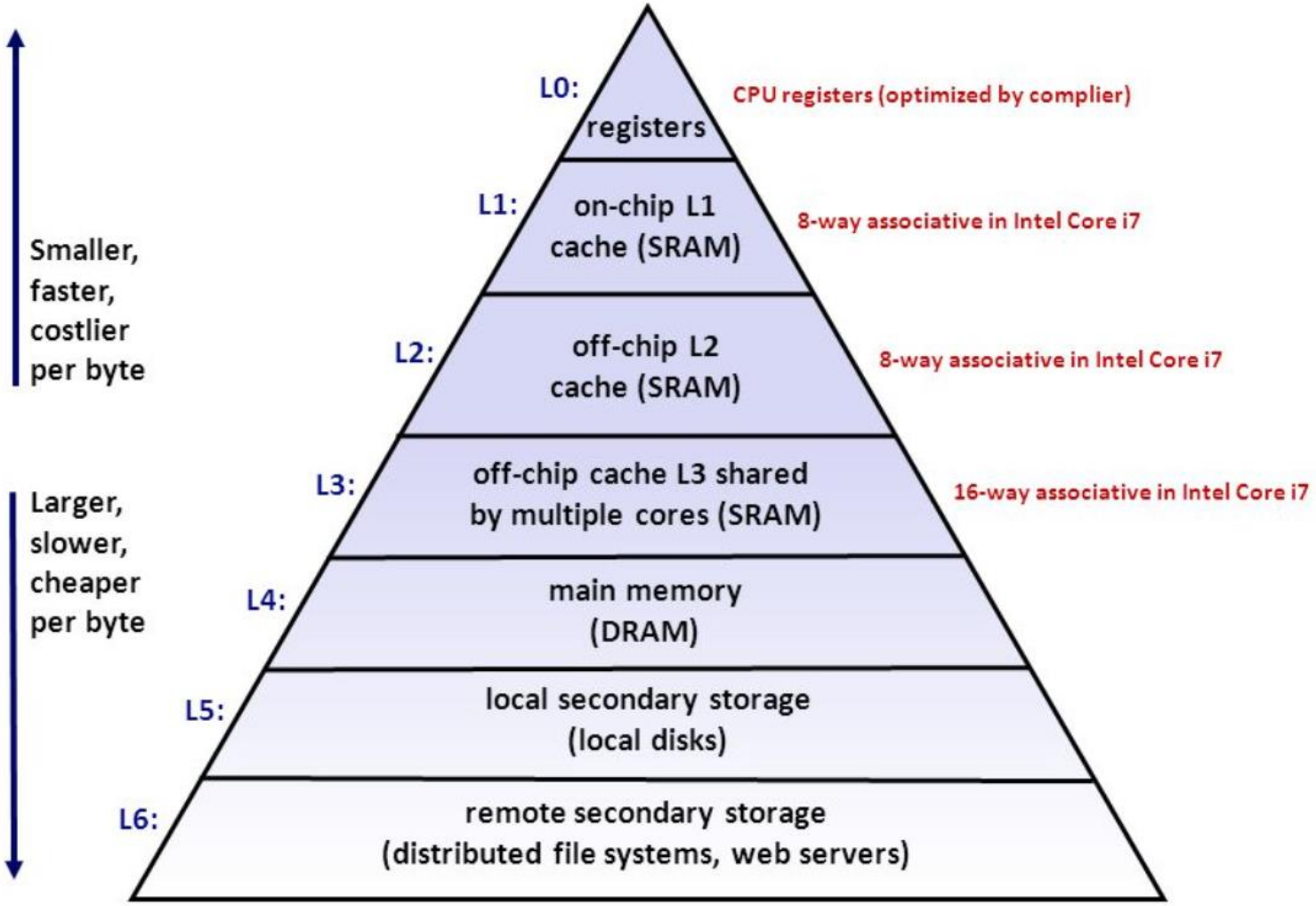
- Flash controllers must support hyperscale requirements (latency, performance/watt, endurance, reliability)

Flash Memory Summit

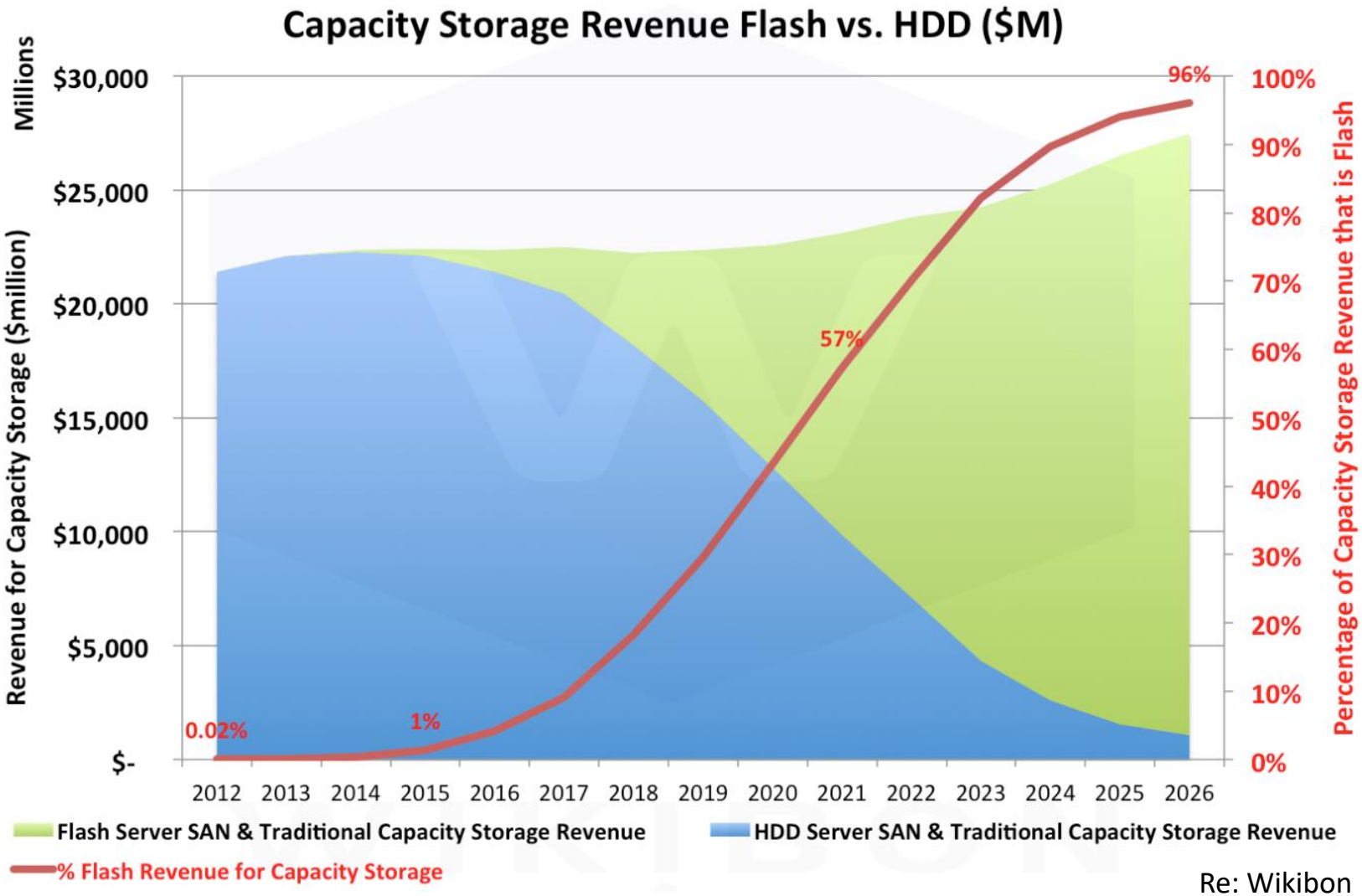
Supporting Infrastructure



Memory and Storage Tiering



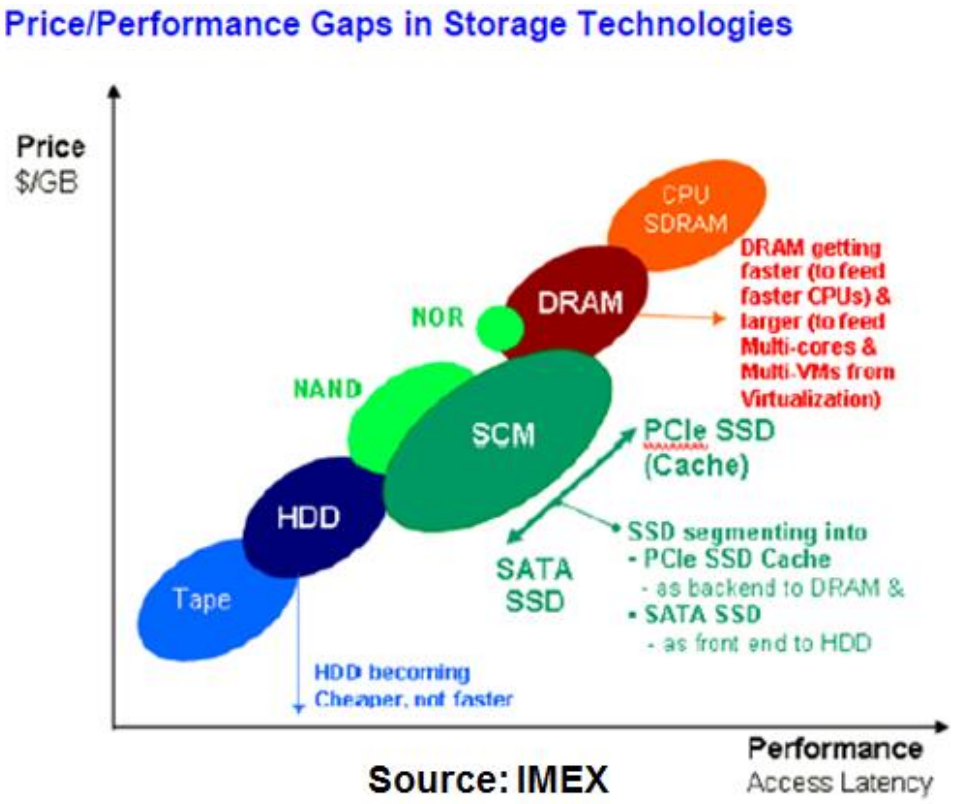
Changing of the Guard



Re: Wikibon

Flash System Challenges

- Error correction costs increasing
- Endurance limits
- Slow write speeds continue
- IO bottlenecking
- Emerging NV technologies
  (MRAM, PCM, RRAM)



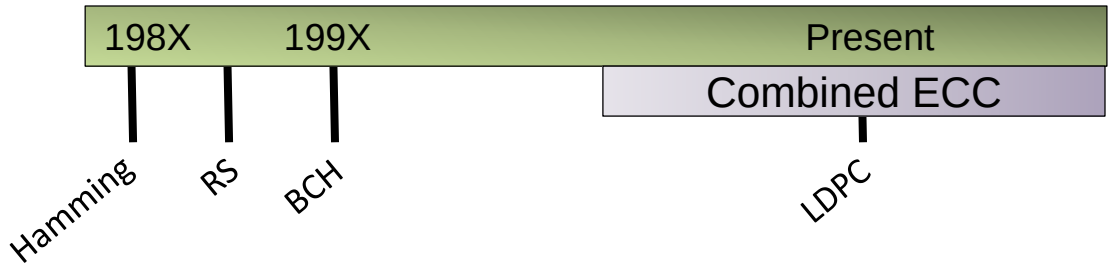
Error Correction Overview

Driving Factors for New ECC

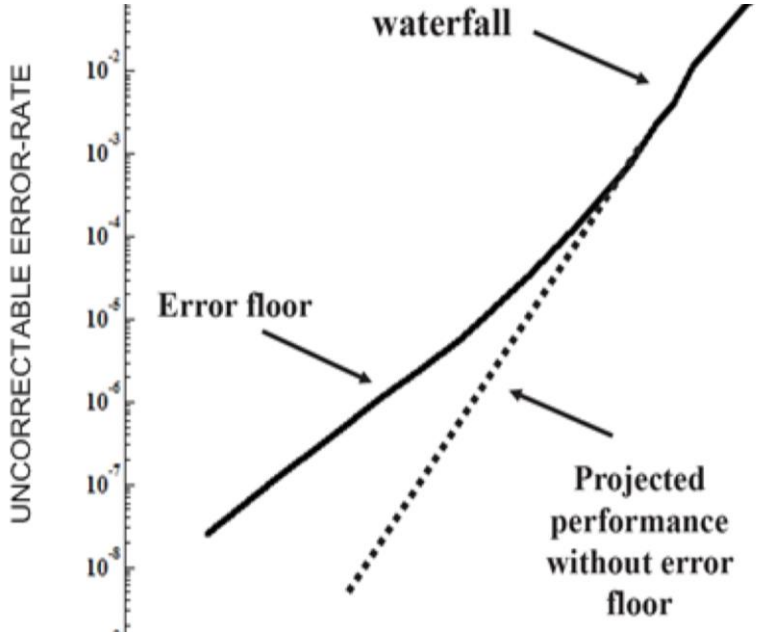
- Increasing Bit errors in NAND Flash
- Soft error occurrences
- Decrease in write cycles
- RS, BCH overhead for data and spare area
- Increase use of Metadata in file systems
- Correction Overhead
- Gate count
- Requirement for no data loss

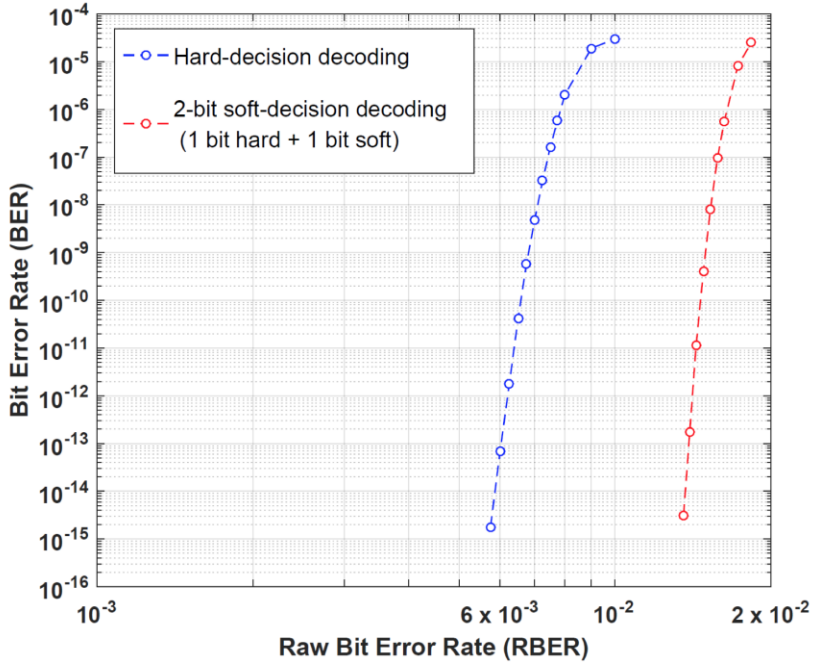
Comparing ECC Solutions

Features	BCH	LDPC
Gate Count	Low	Mid
Latency	Low	Medium
Tuneablity	low	high
Soft Data	No	Yes



Example Codelucida LDPC





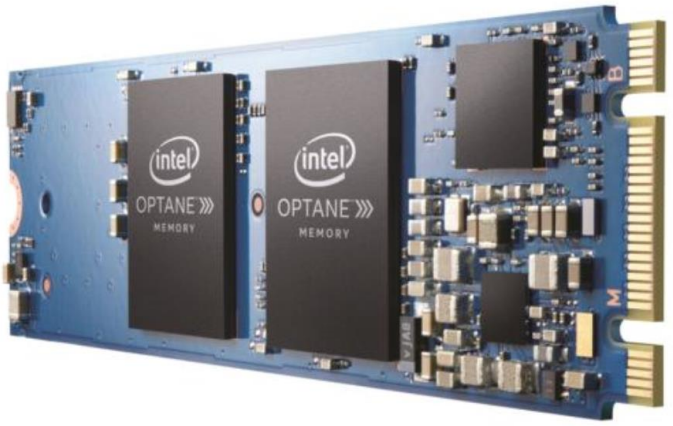
Efficient ECC with NVMe Performance for 3D NAND TLC/QLC

Flash Controller Support

IP	IO	Speed	Logic Density	Comments
ONFI 4.1	40 pins/ch	400 MTps	5KLE/ch	NAND flash control, wear leveling, garbage collection
Toggle Mode 2.x	40 pins/ch	400 MTps	5KLE/ch	Same
DDR4		3.2Gbps	10KLE	Flash control modes available for NVDIMM
PCM MRAM			5KLE 5KLE	PCM- Pending production $ MRAM- Persistent memory controller
BCH			<10KLE	Reference design
PCIe	G4x8	128Gbps	HIP	Flash Cache

Persistent Friend or Foe

- Intel/Micron Xpoint Claimed Attributes vs. NAND
  - Performance (10X)
  - Endurance (1000X)
  - Latency (1/1000X)
  - Byte addressable

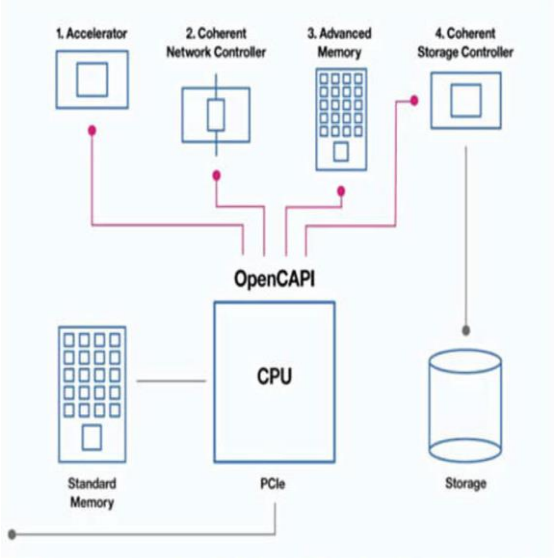


- Est Cost (2X+)

- Opportunity for NAND to support load/store-driven data center applications (NVDIMM-F and NVDIMM-P)

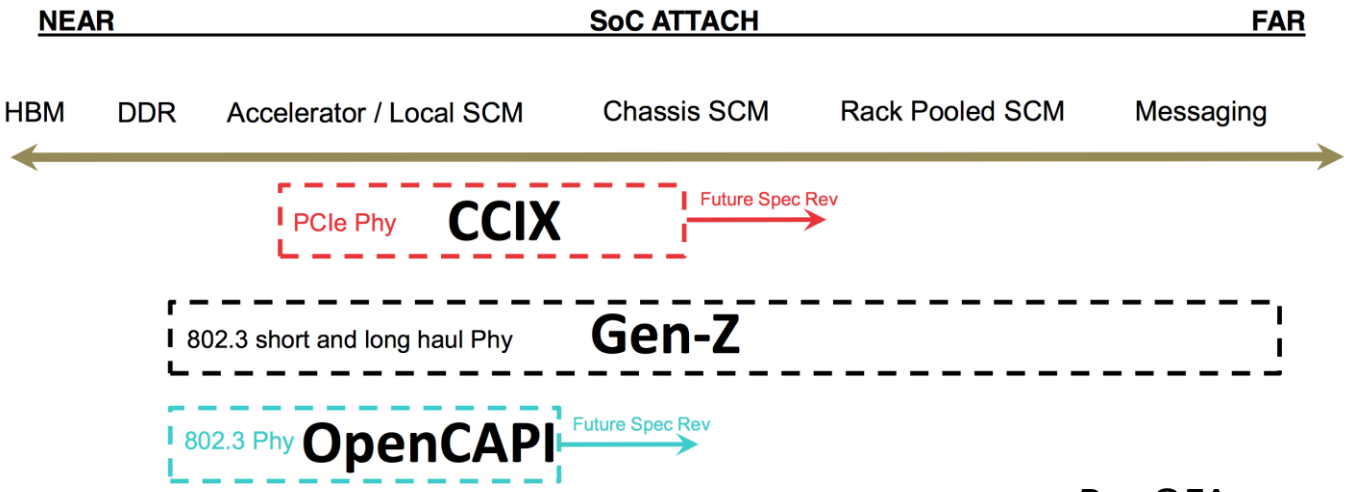
Flash Controller Challenges: Now

- **Host Interface IO**
  - Gen Z, CCIX, OpenCAPI
  - PCIe Gen 4
  - Open Channel
- **Application Requirements**
  - Deterministic latencies
  - Load/Store vs Block
  - Performance
  - Endurance
- **Hybrid Control**
  - 3D NAND, 2D NAND
  - Cache: 3DXpoint, MRAM



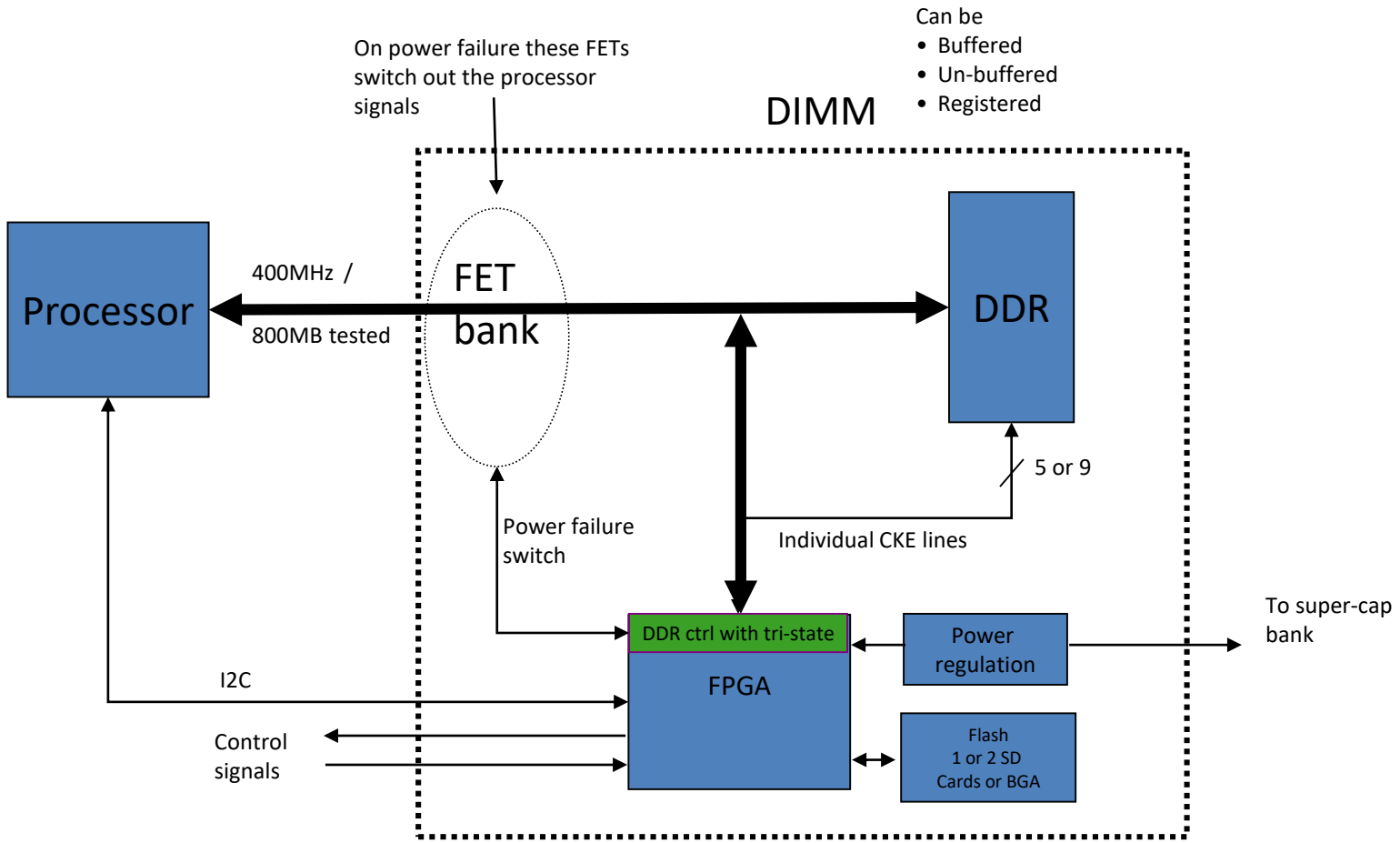
Coherent Networks Roadmap

- Cache coherency will continue to expand into SCM into SSD caches

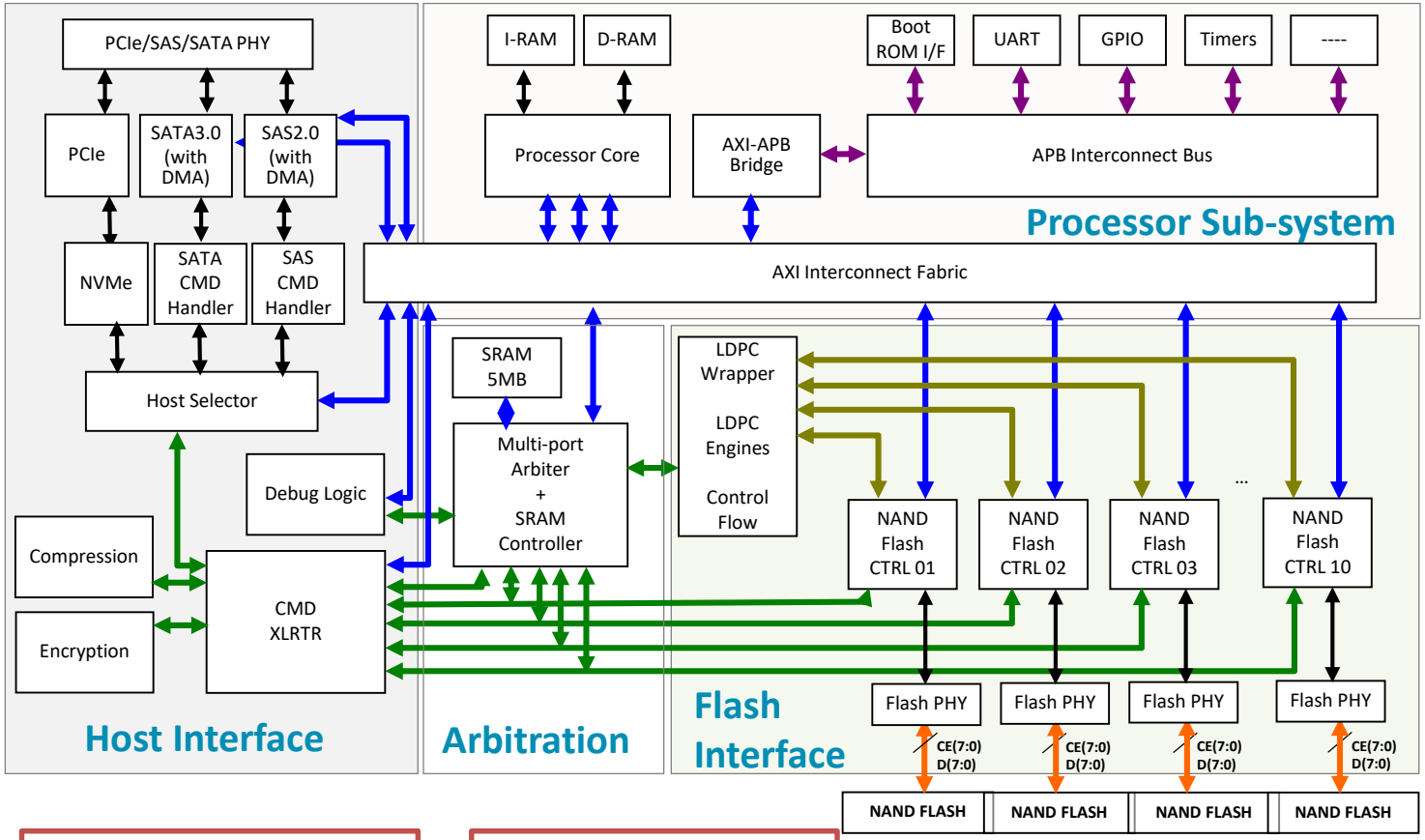


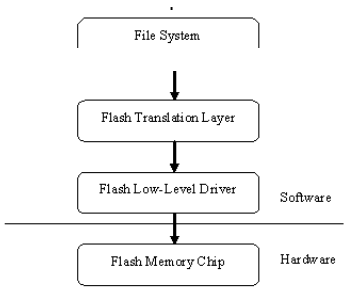
Re: OFA.org

NVDIMM Controller Architecture



Typical SSD Controller Architecture





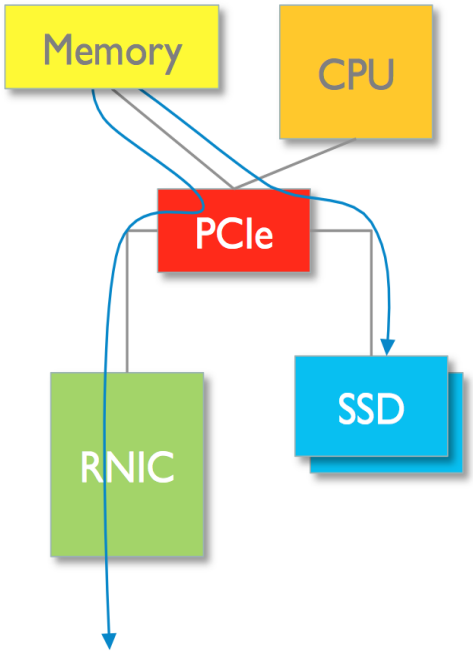
Typical Attributes
- Number of Ports 8 to 32
- Pin Count 250 to 1000+
- Power 1 to 3.5 Watts
- Internal, External RAM

Variations
- Number of CPU's
- Error Correction
- Interfaces
- Memory Type and Size

AXI Interconnect
APB Bus
AXI for Memory

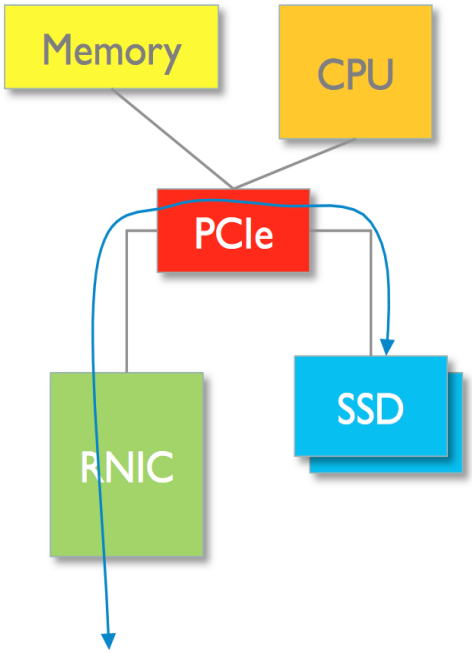
SNIA

NVMe Roadmap- a NVMeF Precursor



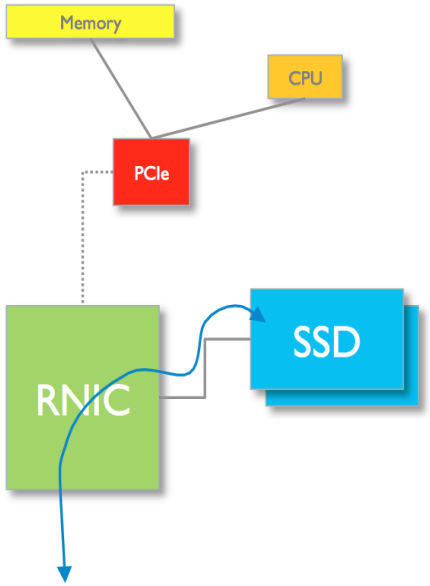
DMA to/from host memory
CPU handles command transfer

YESTERDAY



Peer-to-peer DMA through PCI bus
No host memory use

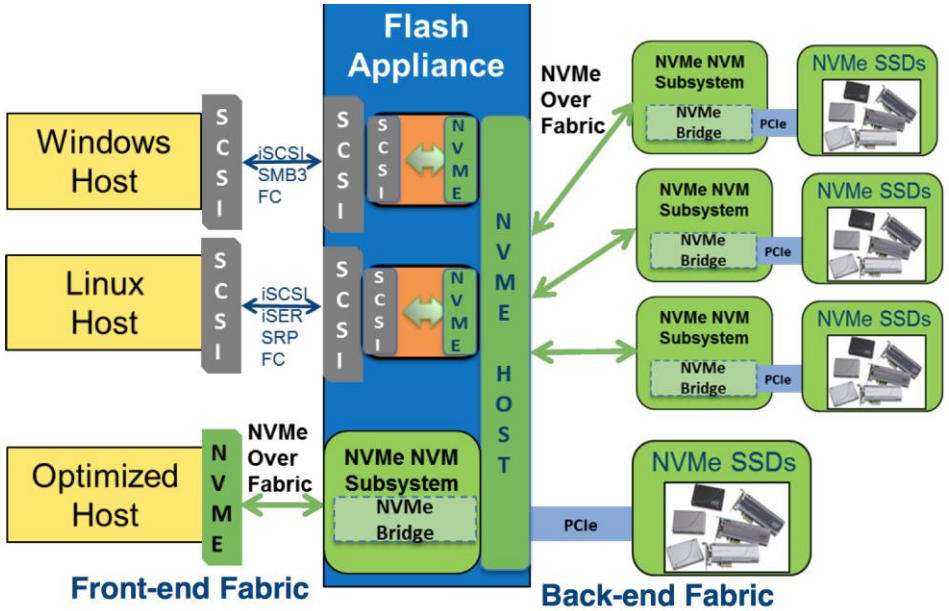
TODAY

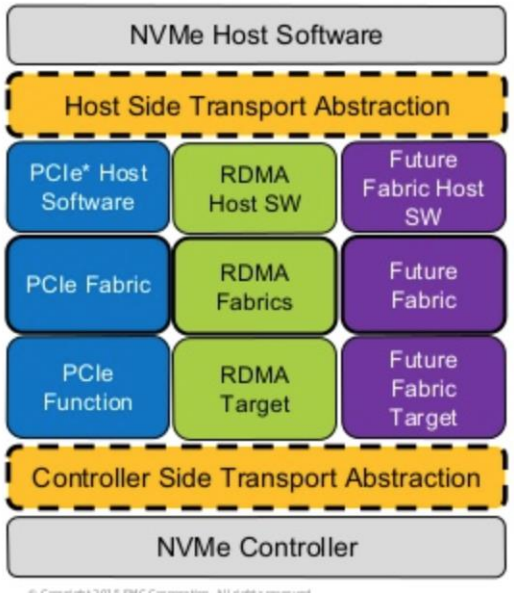


Network integrated NVMe storage
No CPU or memory involvement

TOMORROW

NVMeF- Key Value Points





OpenFabricsAlliance.org

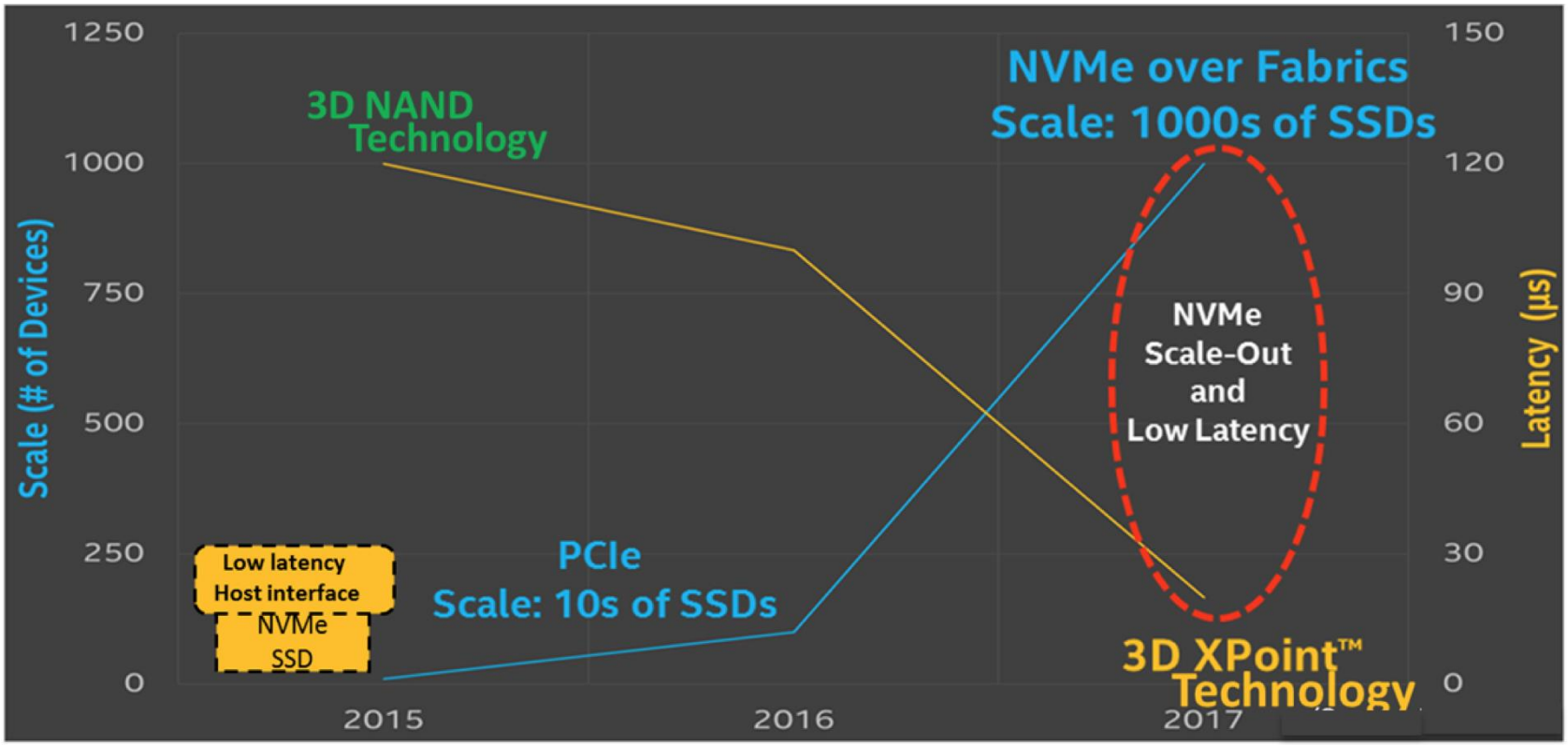
Encapsulated protocol transmitted over fabrics

RDMA (iWARP, RoCE)
FC, Infiniband, L1 Tunneling

The goal is to enable next-gen technologies to deliver a 4KB I/O in less than 10µs - about one thousandth of the latency of a 7200 RPM SATA drive

Flash Memory Summit

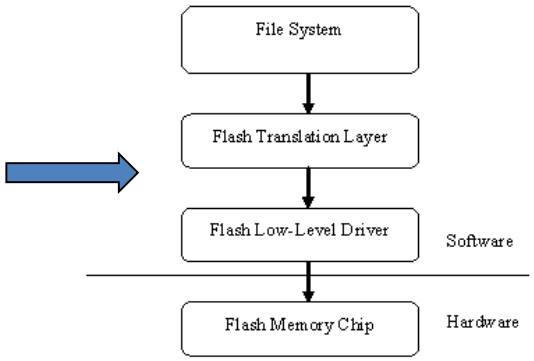
SSD Scale Out over Fabrics

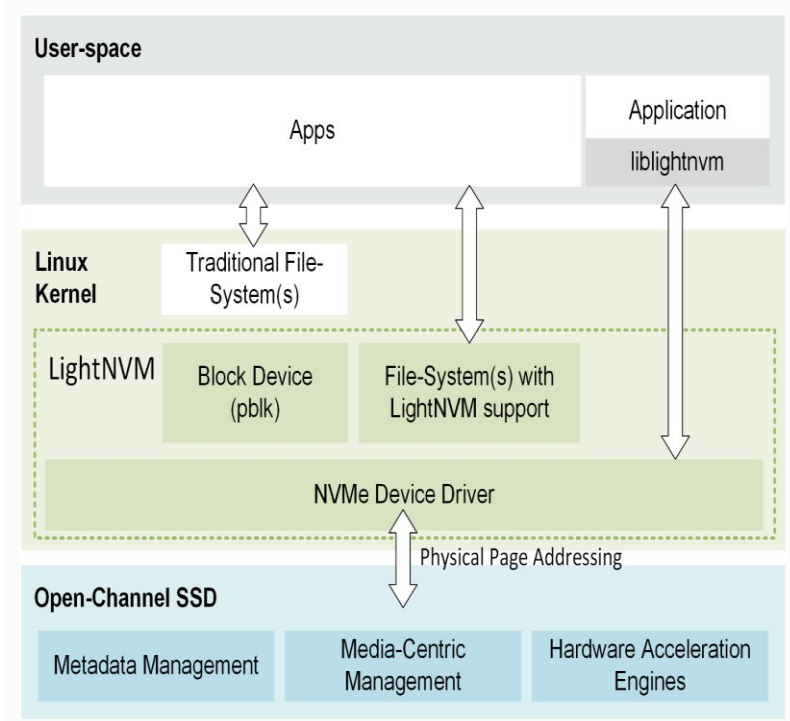


- NVMeF breaks through local NVMe barrier and supports low latency

Open Channel Pros and Challenges

- I/O Isolation and Determinism
- Software managed resources
- Application-centric
- Linux kernel support required
- Vendor-specific attributes





Re: CNEX

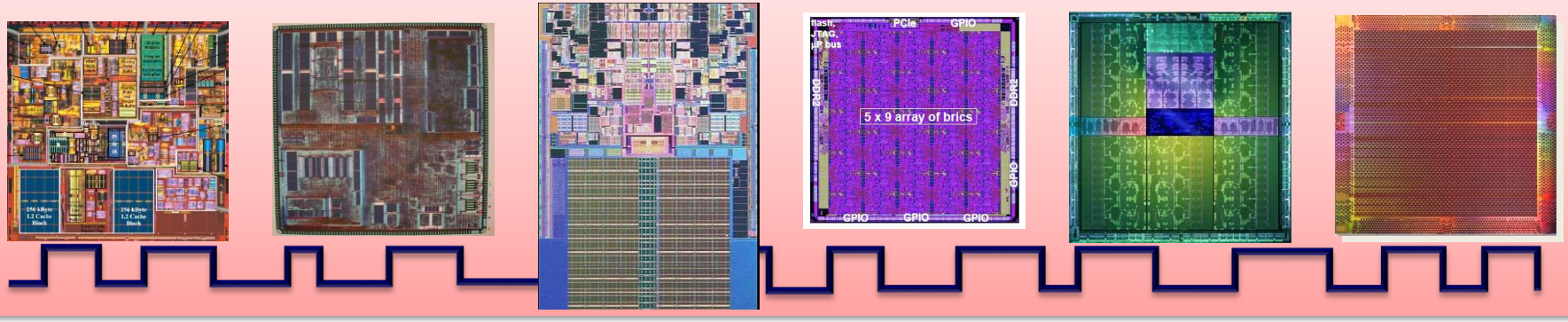
Flash Memory Summit

Differentiation Paradox

- Hyperscaler standardization
- Vendor competitive selling features
- Application requirements
  - I/O
  - Power Consumption
  - Capacity
  - Latency

Controller Options

Technology scaling favors programmability and parallelism



CPUs DSPs Multi-Cores Array GPGPUs FPGAs

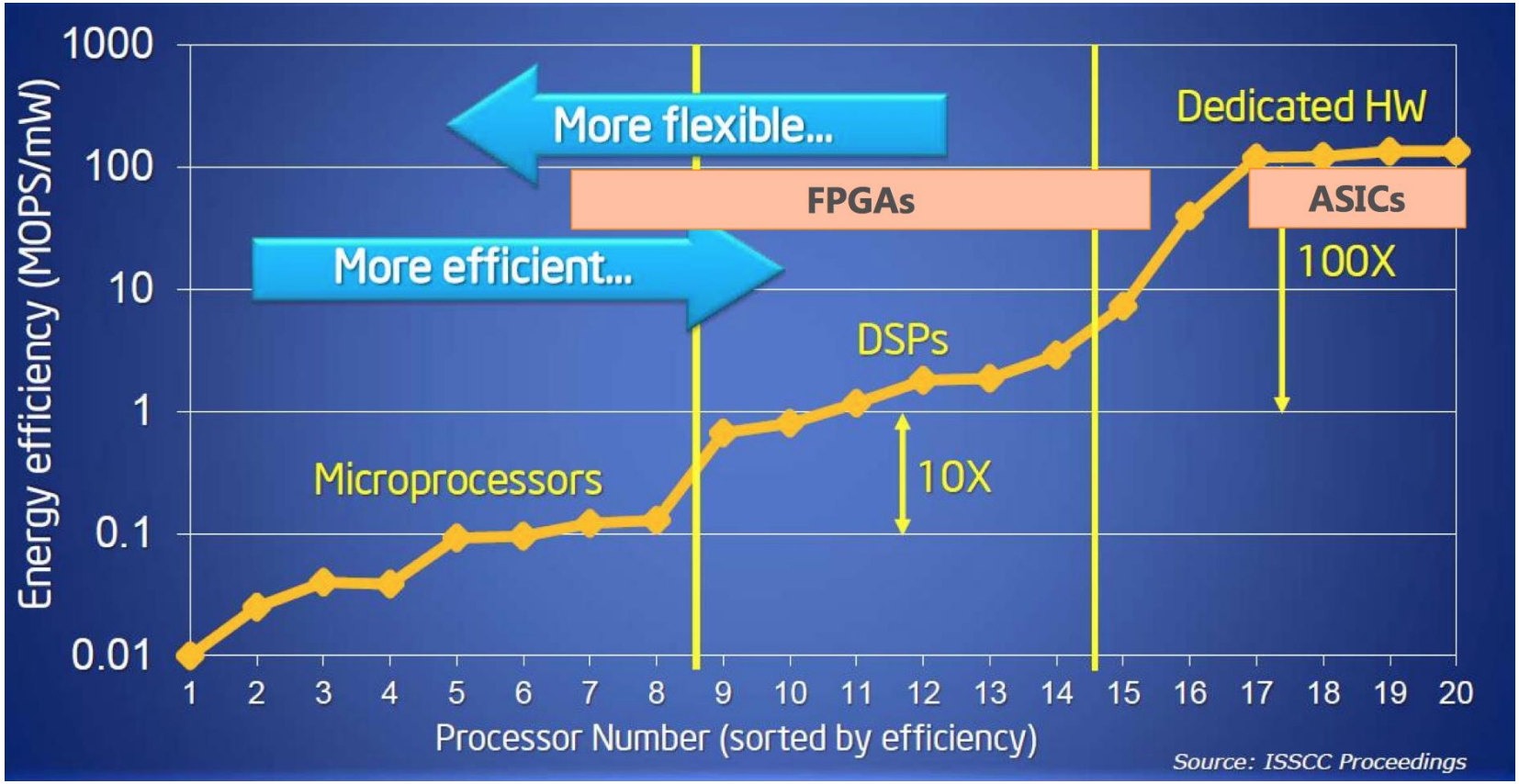
Single Cores

Multi-Cores
Coarse-Grained
CPUs and DSPs

Coarse-Grained
Massively
Parallel
Processor
Arrays

Fine-Grained
Massively
Parallel
Arrays

Flash Controller Technology Options



- Data center metric is performance/watt
- Performance, power efficiency and flexibility is required to support data center applications

Technology Comparison

Technology	Pros	Cons
CPU	Well established products	• Limited cores for parallel processing • Power consumption
FPGA	Heterogeneous parallel processing Performance/Watt Flexibility	• Rudimentary development environment • Inefficient per unit costing
GPU	Same task parallel processing Developer ecosystem	• Power consumption • Leading variable types
ASIC	Highest Performance	• High NRE Custom design
ASSP	Custom Performance	• Limited functionality

Flash Memory Summit

Summary

- Flash Control has extended into tiered subsystem management
  - Caching has extended into SCM, necessitating hybrid control
  - IO interfaces need to support fabric
  - Advancing geometries and process technologies require more and advanced error correction
  - Hyperscaler applications demand load/store performance with deterministic latency

Flash Memory Summit

Annual Flash Controller Update

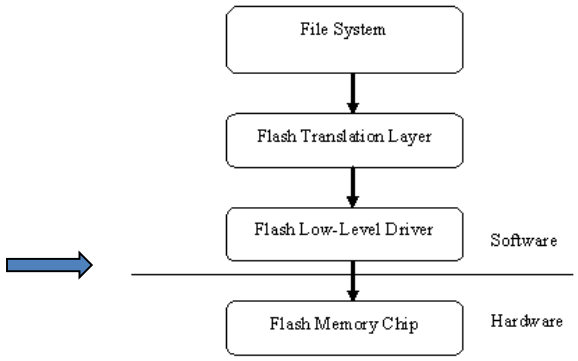
David McIntyre

DSMcIntyreConsulting@gmail.com

Text **FMS** to (408) 772-7044

Flash Controller Challenges: Then

- **Emerging memory types**
  - ONFI 4.0, Toggle Mode 2.x
  - PCM, MRAM
  - DDR4
- **Controller Performance Options**
  - Write back cache, queuing, interleaving, striping
- **ECC levels**
  - BCH, LDPC, Hybrid
- **FTL location- Host or companion**
- **Data transfer interface support**
  - PCI Express, SAS/SATA, FC, IB



Flash Controller Challenges: Now

- **I/O Performance**
  - Interchip coherency
  - Host Communications
  - Network
- **Latency**
  - HPC network latencies
- **Density**
  - 3D, HBM2
- **Heterogeneous flash memory types**
- **Reliability and Endurance**