Western Digital®

Toward a Memory-centric Architecture

Martin Fink
EVP & Chief Technology Officer
Western Digital Corporation

August 8, 2017

SAFE HARBOR | DISCLAIMERS

Forward-Looking Statements

This presentation contains forward-looking statements that involve risks and uncertainties, including, but not limited to, statements regarding our market positioning, product development efforts, growth opportunities, business strategy and market trends. 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.

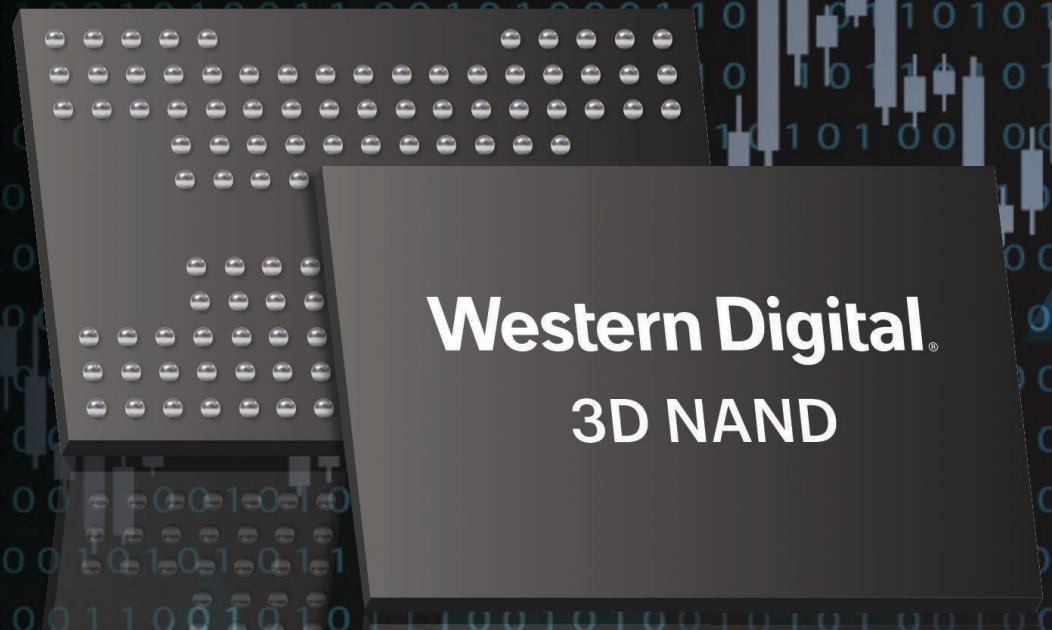
Additional key risks and uncertainties include the impact of continued uncertainty and volatility in global economic conditions; actions by competitors; difficulties associated with go-to-market capabilities; business conditions; growth in our markets; and pricing trends and fluctuations in average selling 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.

3D NAND Technology Leadership

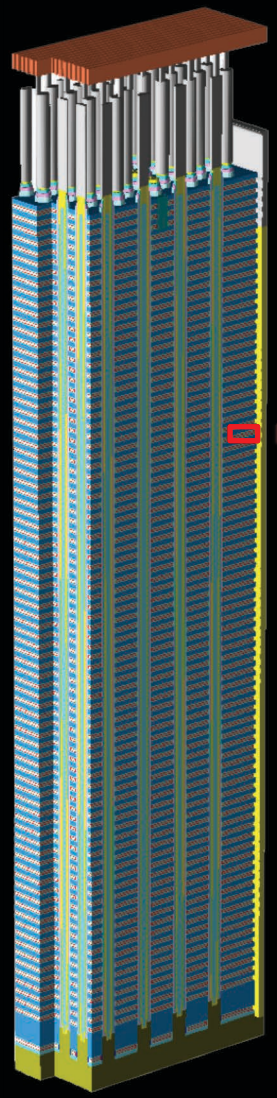
Recent developments

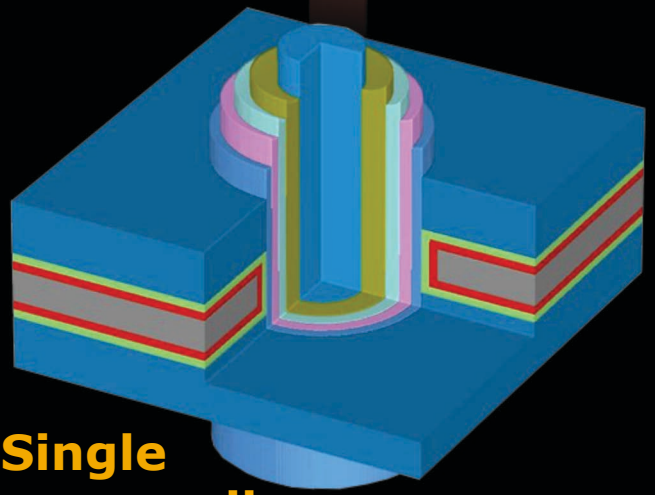
BiCS4
96-layer 3D NAND technology

X4 technology
Four-bits-per-cell flash memory architecture on 64-layer 3D NAND



Four-Bits-Per-Cell (X4) Technology

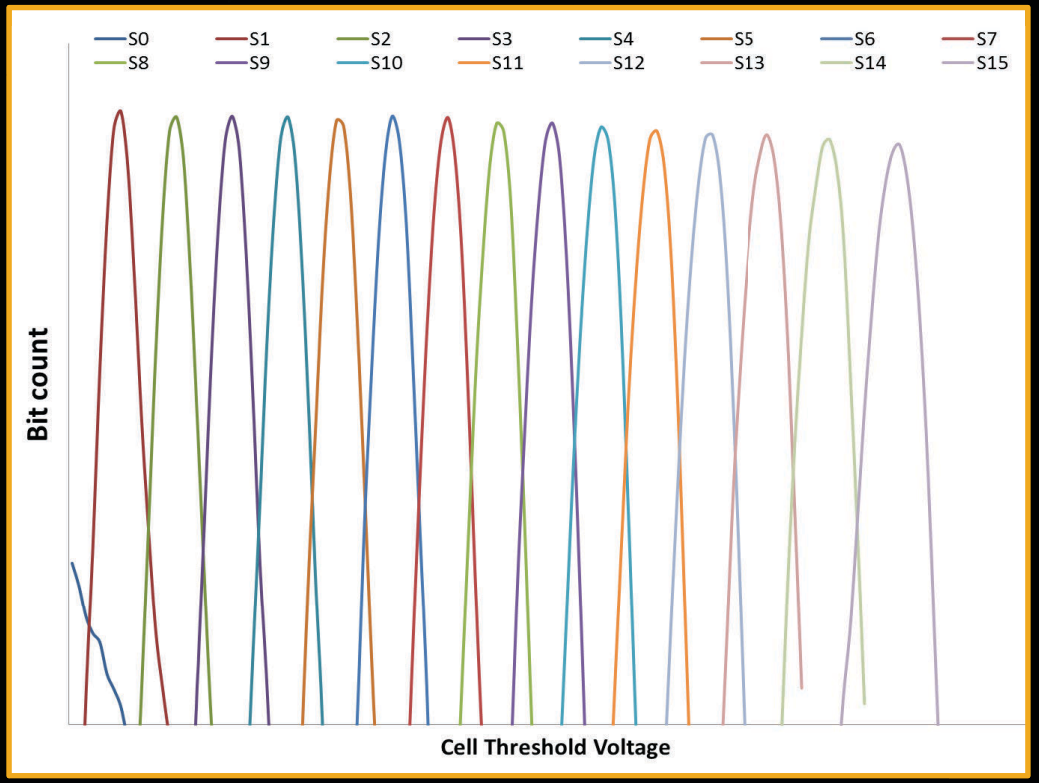




Single memory cell

Industry-leading 64-layer 3D NAND

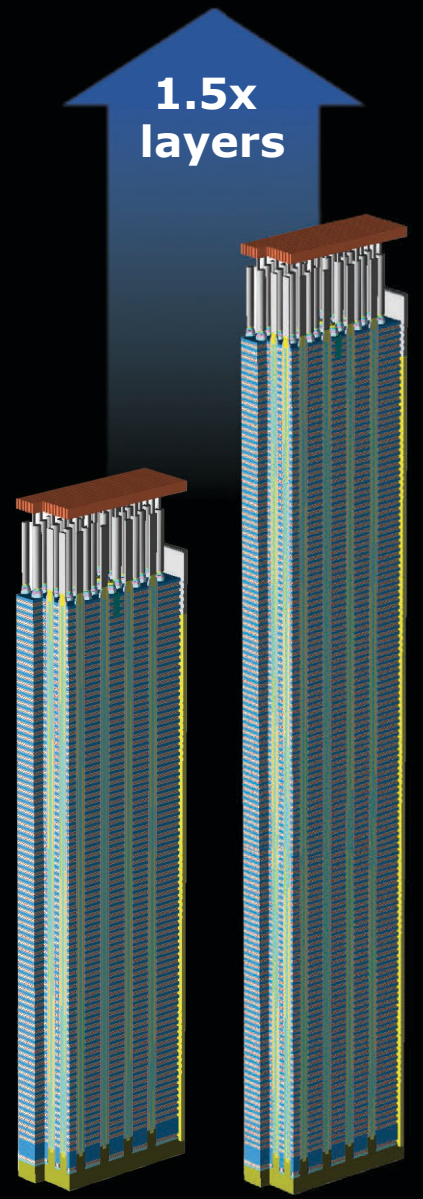
16 data levels per cell

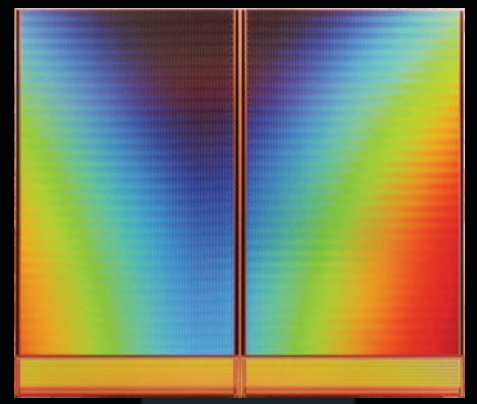


- SSDs based on 768Gb X4 die showcased at FMS
- Enables expanded choice of storage solutions for customers

96-Layer 3D NAND Technology

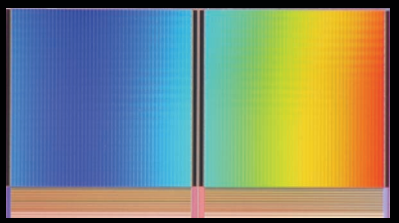
- Denser storage, improved scalability at an attractive cost
- Enables up to 1 terabit of storage per die
- Faster IO speed using low power, low voltage IO design





512G, 64L BiCS3

35%
die
shrink



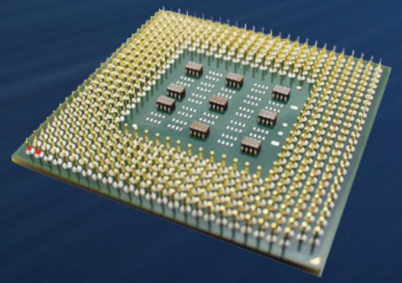
512G, 96L BiCS4

NAND Driving Technology Advancement

~170 Billion*
per BiCS NAND die

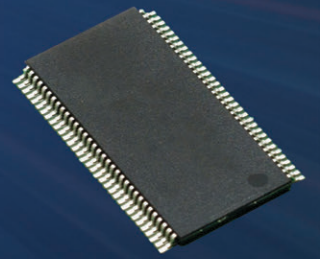
~10-20 Billion

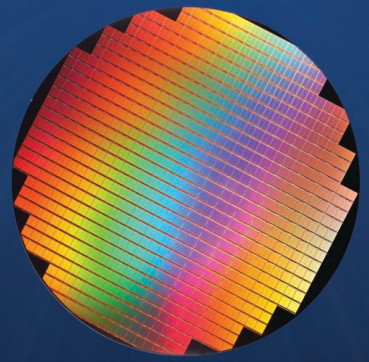
per CPU die 



~16 Billion

 per DRAM die





Devices

*512 Gb x3 BICS die

Agenda

1	**The Evolving Role of Data** • The data-driven economy • Big Data, Fast Data applications • Purpose-built architectures
2	**NVM System Opportunities** • Multiple NVM technologies • Memory not storage • Programming models
3	**Memory-centric Compute** • Motivation and opportunity • Industry enablers
4	**Technology Predictions**

Agenda

1. The Evolving Role of Data
   - The data-driven economy
   - Big Data, Fast Data applications
   - Purpose-built architectures
2. NVM System Opportunities
   - Multiple NVM technologies
   - Memory not storage
   - Programming models
3. Memory-centric Compute
   - Motivation and opportunity
   - Industry enablers
4. Technology Predictions

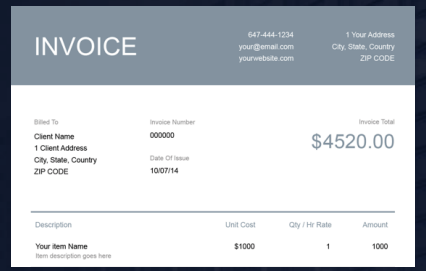
The Evolving Role of Data

Creating the data-driven economy

Richness

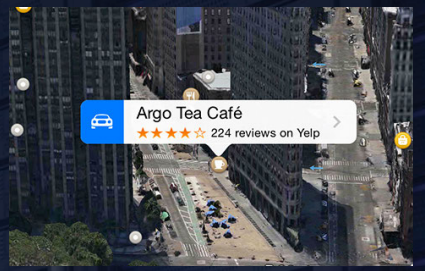
Value

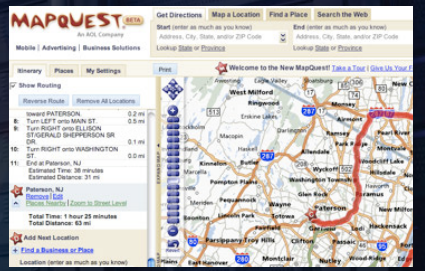
Data as a record



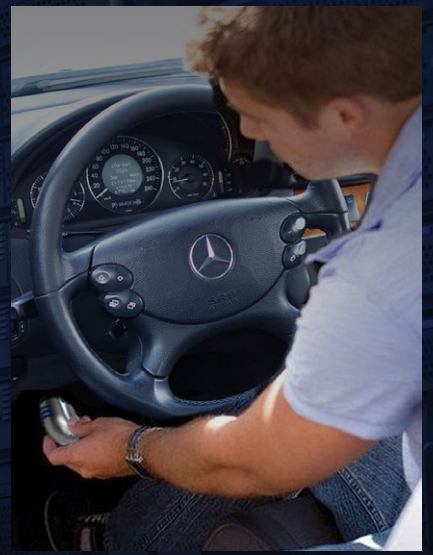
Data as communication

Data as efficiency





Data as currency



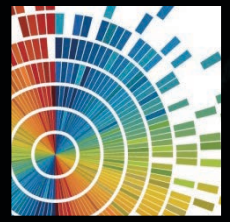
Diverse and Connected Data Types

Tight coupling between Big Data and Fast Data

Big Data

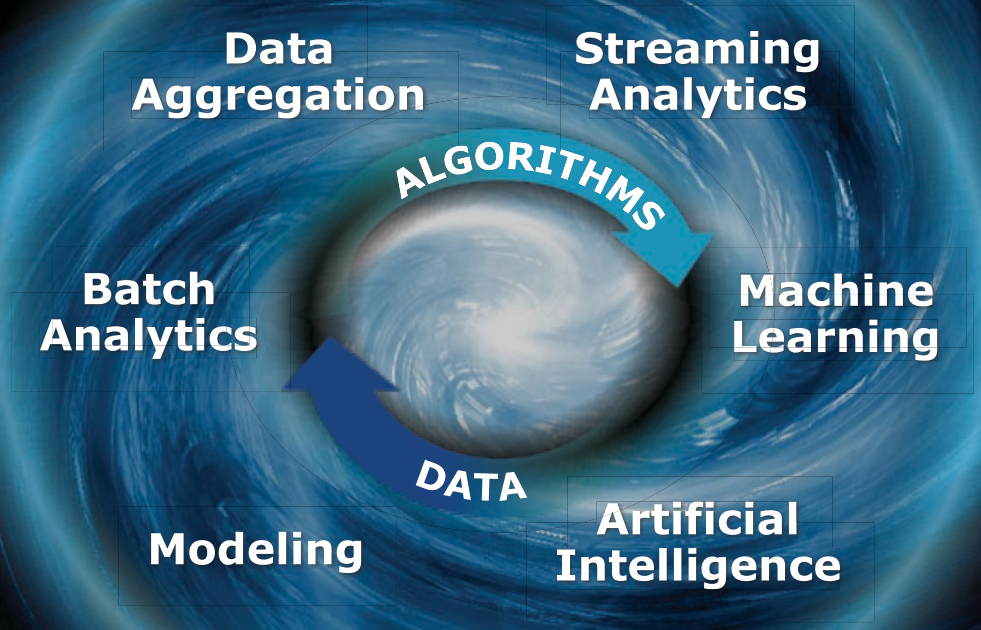
Insight

Prediction



Prescription

Scale



Fast Data

Mobility

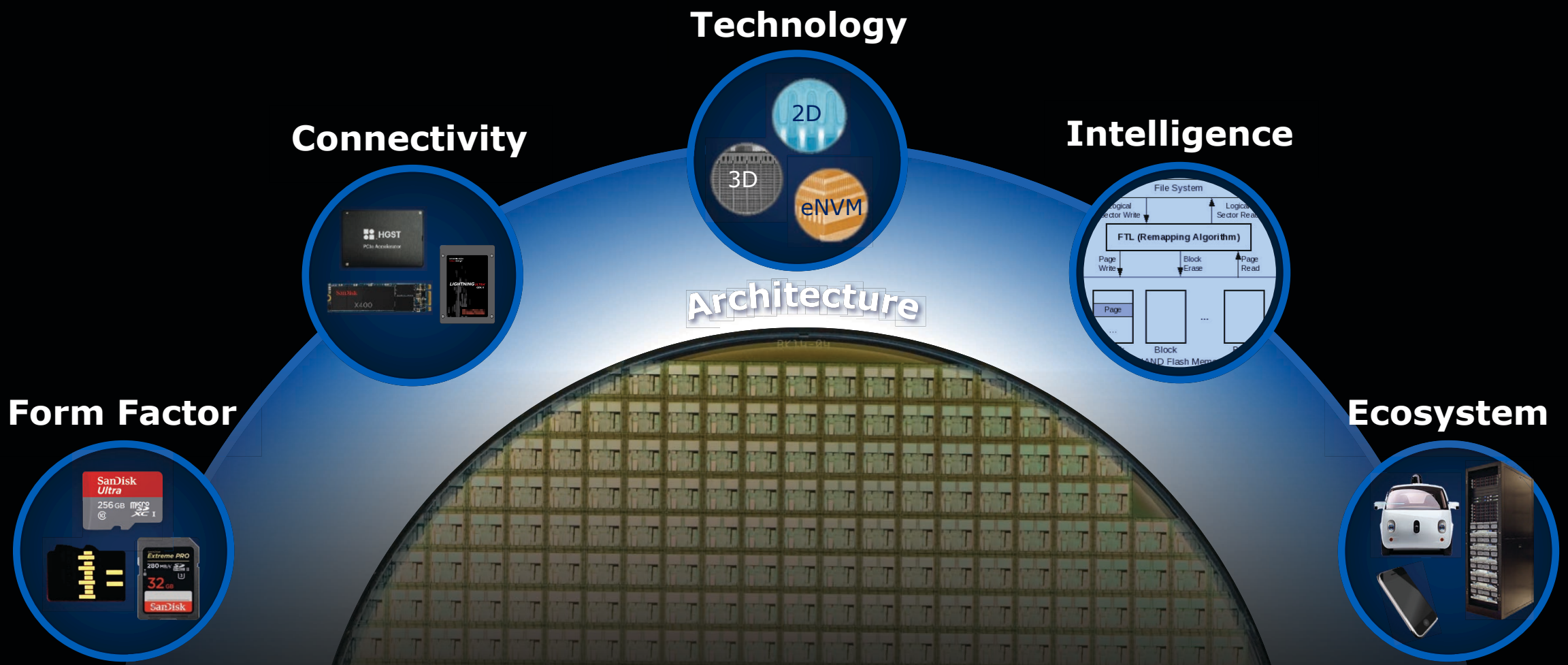
Real-time Results

Smart Machines

Performance

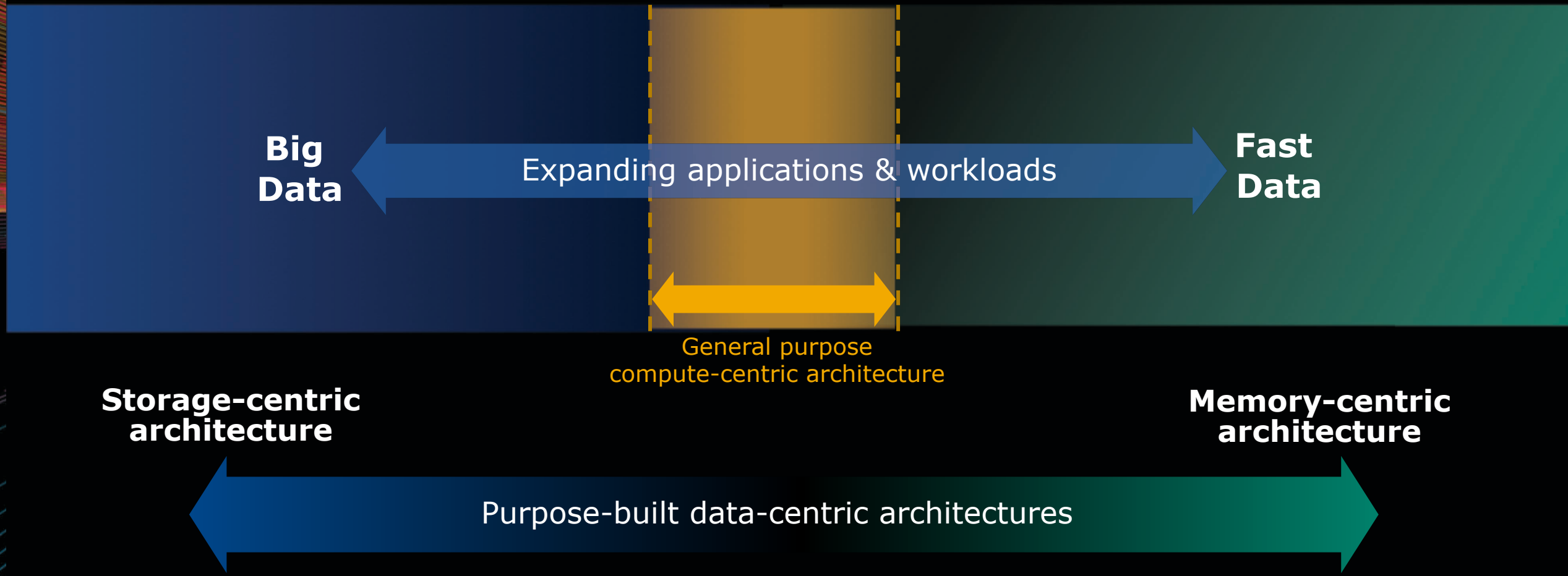
Vectors of Innovation Expanding and Accelerating

Imperative to add value beyond silicon

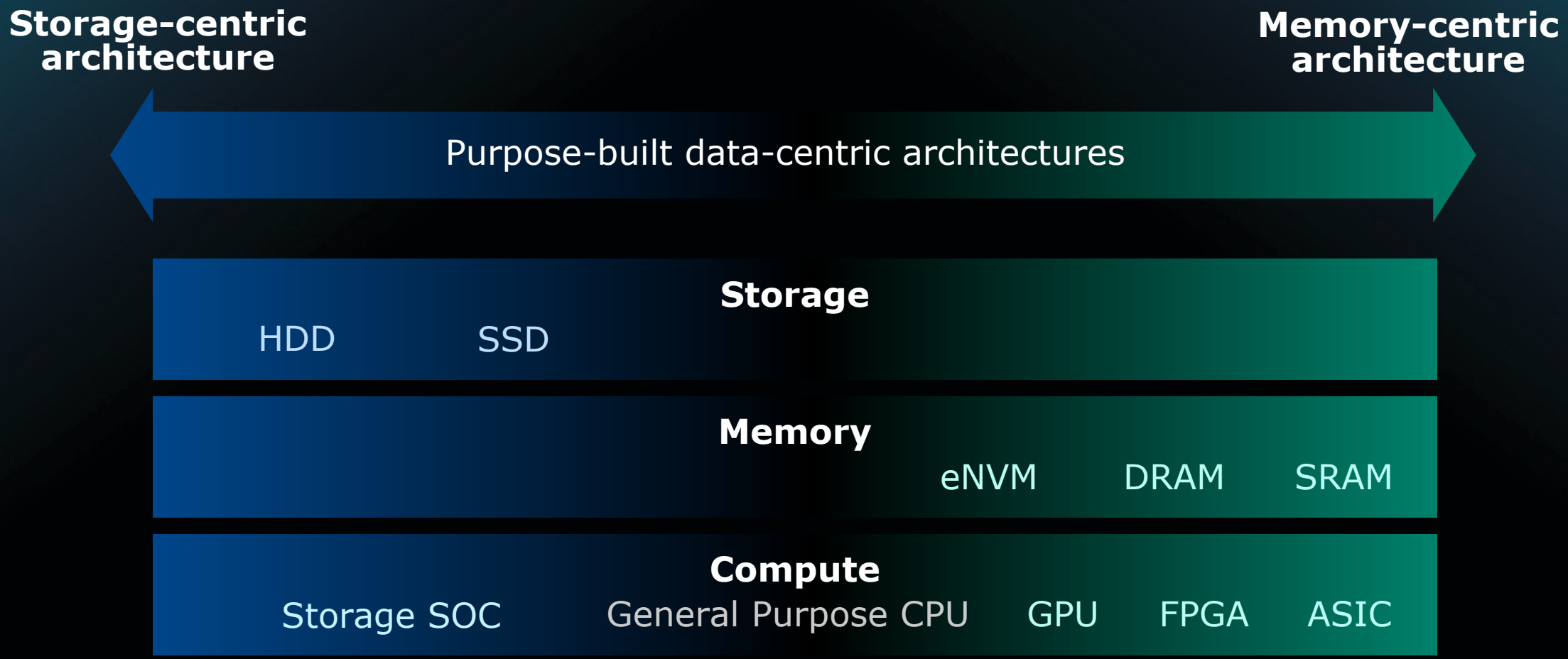


From General Purpose to Purpose-built

Architectures designed for Big Data, Fast Data applications



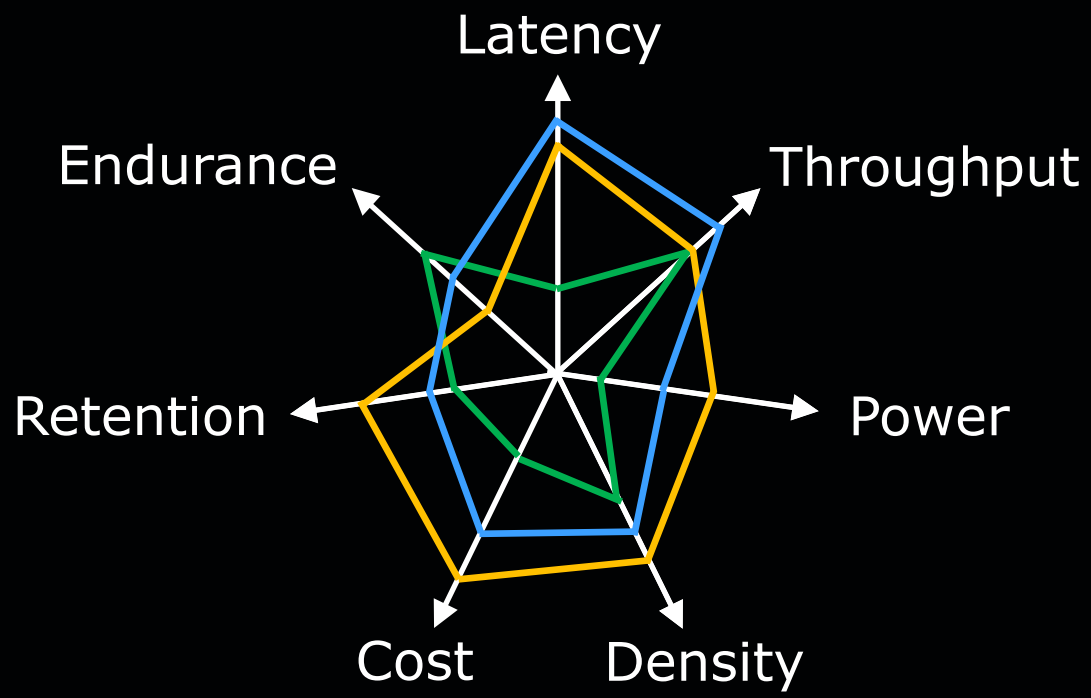
Workload Diversity Demands Diverse Technologies and Architectures



Agenda

1. The Evolving Role of Data
   - The data-driven economy
   - Big Data, Fast Data applications
   - Purpose-built architectures
2. NVM System Opportunities
   - Multiple NVM technologies
   - Memory not storage
   - Programming models
3. Memory-Centric Compute
   - Motivation and opportunity
   - Industry enablers
4. Technology Predictions

Multiple Emerging NV Memories

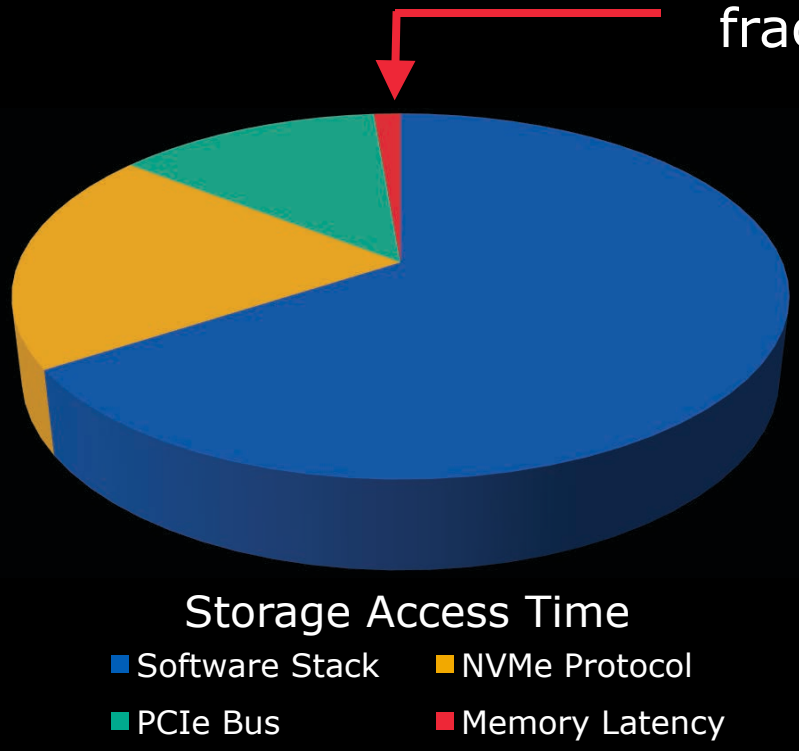


Multiple NV memories
for different applications:

- Analytics
- Real-time applications
- Mobility
- Artificial Intelligence
- IoT

NVM Value Proposition: Memory Not Storage

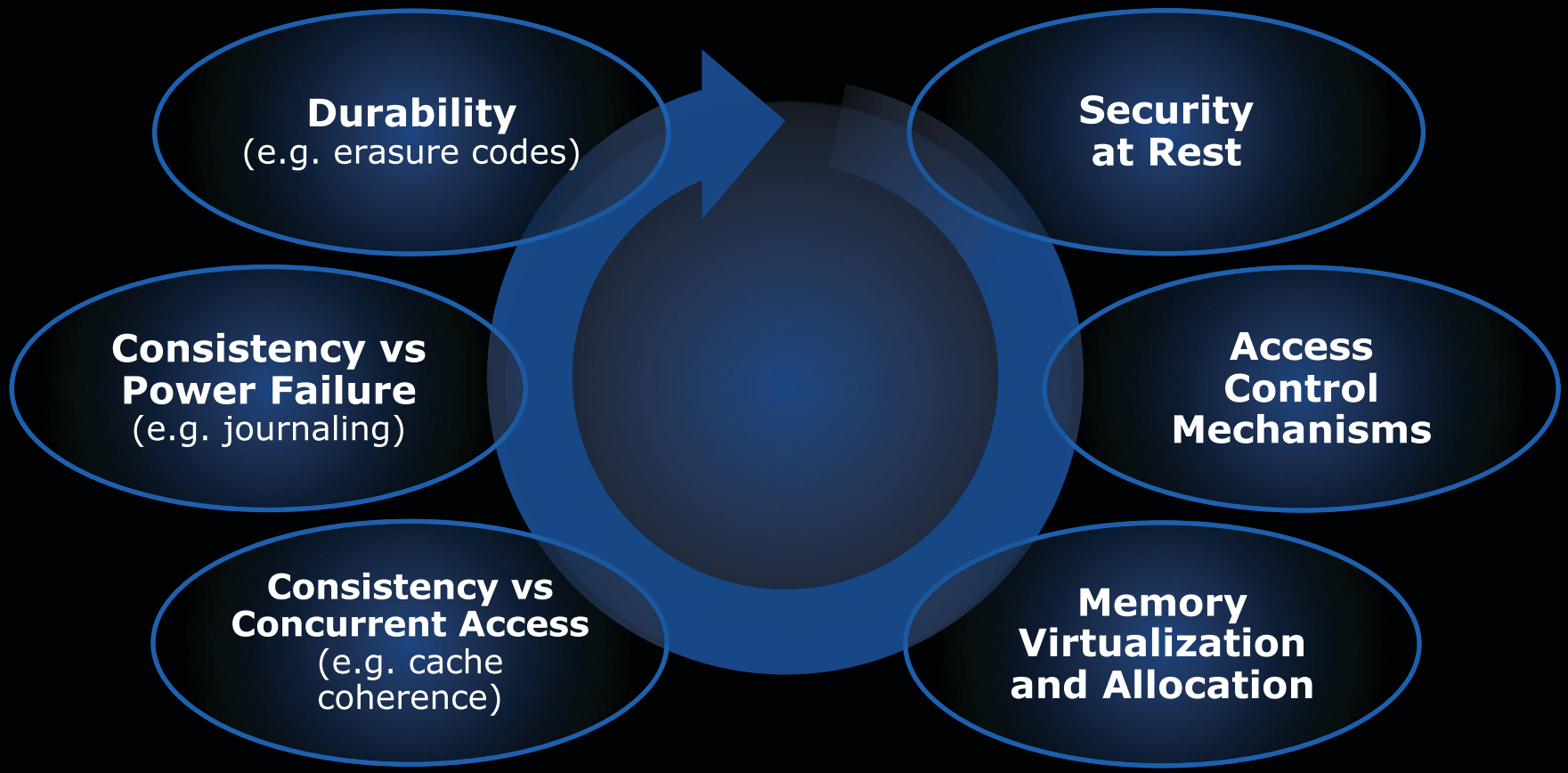
NV Memory latency is only a small fraction of the total storage IO latency



Storage Semantics	Memory Semantics
Block IO (4 KB)	Cache Line Load/Store (64-128 B)
Complex protocol, mostly in SW	Simpler protocol, mostly in HW
Designed for media with 10's of μs latencies	Designed for media with 50-150 ns latency

NV Memory Programming Models

What does it mean for memory to be used as main persistent storage?



Foundational changes to programming required

Agenda

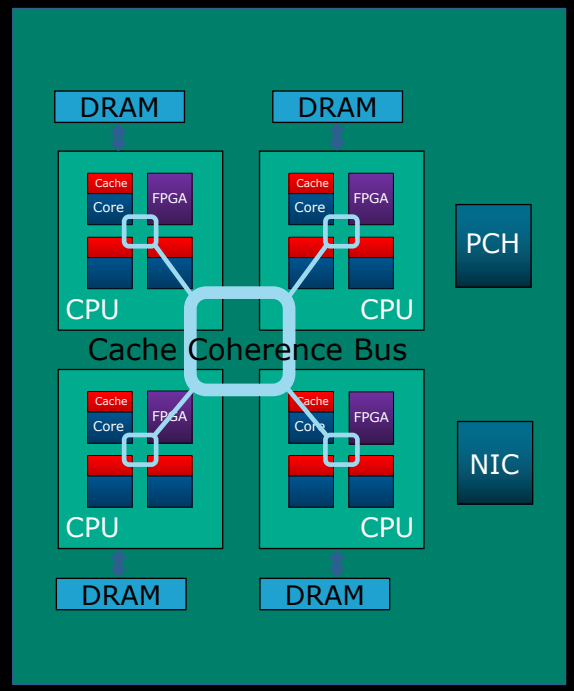
1. The Evolving Role of Data
   - The data-driven economy
   - Big Data, Fast Data applications
   - Purpose-built architectures
2. NVM System Opportunities
   - Multiple NVM technologies
   - Memory not storage
   - Programming models
3. Memory-Centric Compute
   - Motivation and opportunity
   - Industry enablers
4. Technology Predictions

Memory-centric Computing

For many emerging challenges, memory capacity, memory access latency and memory bandwidth are more constrained than compute resources

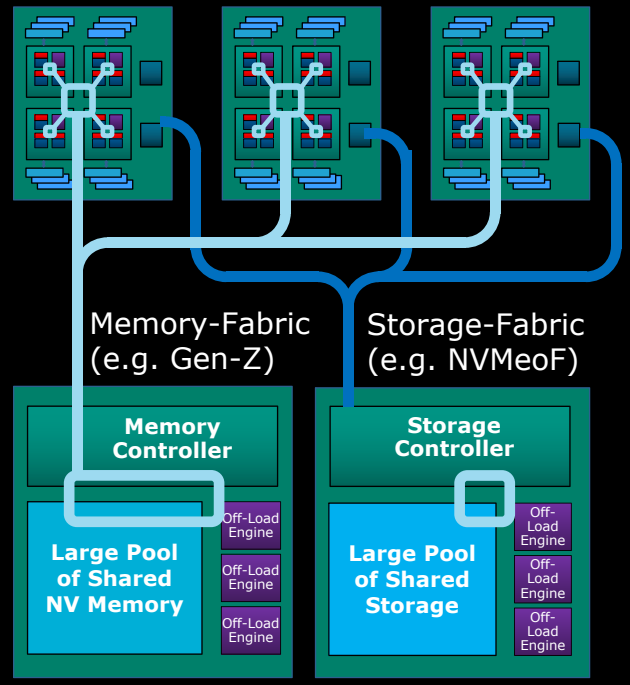
- **Memory Disaggregation**
  Remove memory from behind the processor
- **Memory Pooling & Sharing**
  Enable efficient use of memory. Address new class of problems with large memory footprint
- **Heterogeneous Compute**
  Enable multi-vendor heterogeneous compute (e.g. ML accelerators)

2018



CPU-centric architecture

2020



Data-centric architecture

Compute @ Data

Too much work moving the data around!

Ship compute to data for IO bound workloads.

Power
Reduction in data movement count and distance

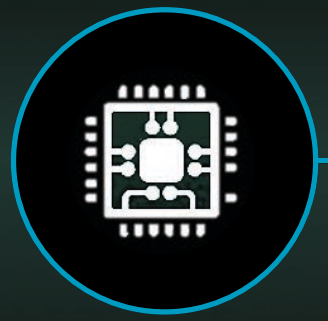
Performance
Parallelism, bandwidth, and latency expose full memory aggregate bandwidth

Cost
Low gate count embedded cores with future open ISA and tools

Industry Enablers

Open industry standards are key enablers for memory-centric compute

Memory Interface



- Standards-based interface to CPU e.g. NVDIMM-P
- Asynchronous R/W
- R/W asymmetry

Low-Latency Fabrics

- Low-latency memory-semantic fabrics e.g. Gen-Z
- Few hundred ns total memory access latency

Cache Coherency

- Open and widely-adopted cache coherency protocols
- Enable heterogeneous compute

Programing Models

- Programming models for persistent memory
- NUMA-aware software stack

Expect innovation and investment to accelerate

Agenda

1	The Evolving Role of Data • The data-driven economy • Big Data, Fast Data applications • Purpose-built architectures
2	NVM System Opportunities • Multiple NVM technologies • Memory not storage • Programming models
3	Memory-centric Compute • Motivation and opportunity • Industry enablers
4	Technology Predictions

Technology Predictions

1. From general purpose to purpose built
2. Open architectures and interfaces
3. Industry standard memory-semantic fabric
4. Use case-driven persistent memory technologies
5. Memory-centric computing