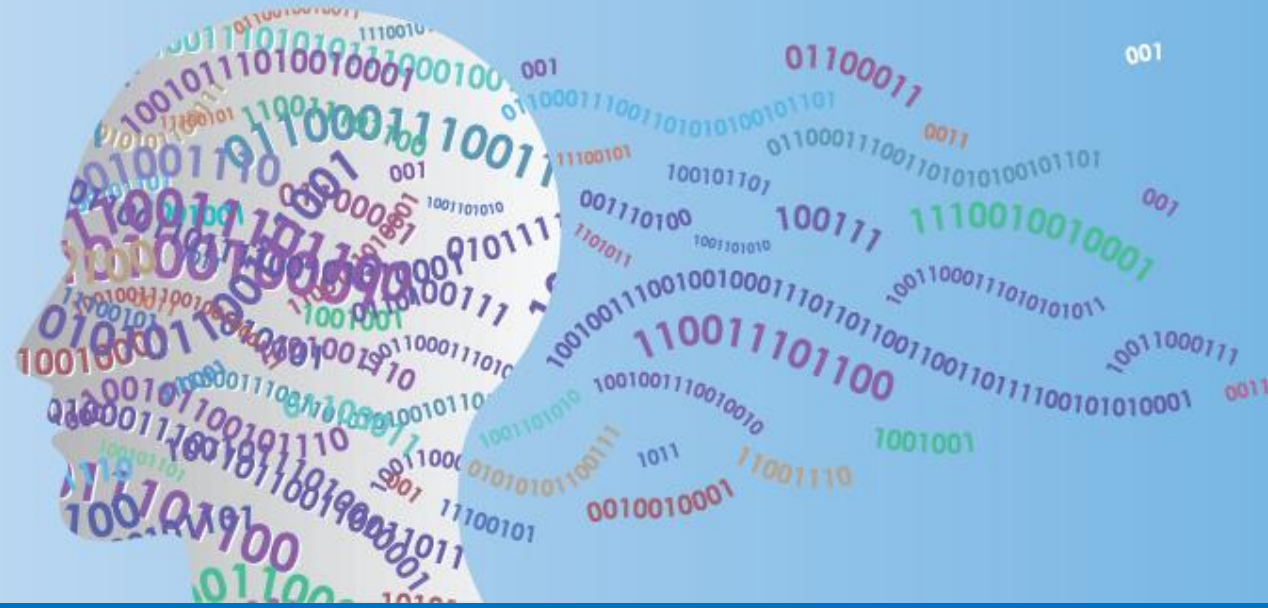
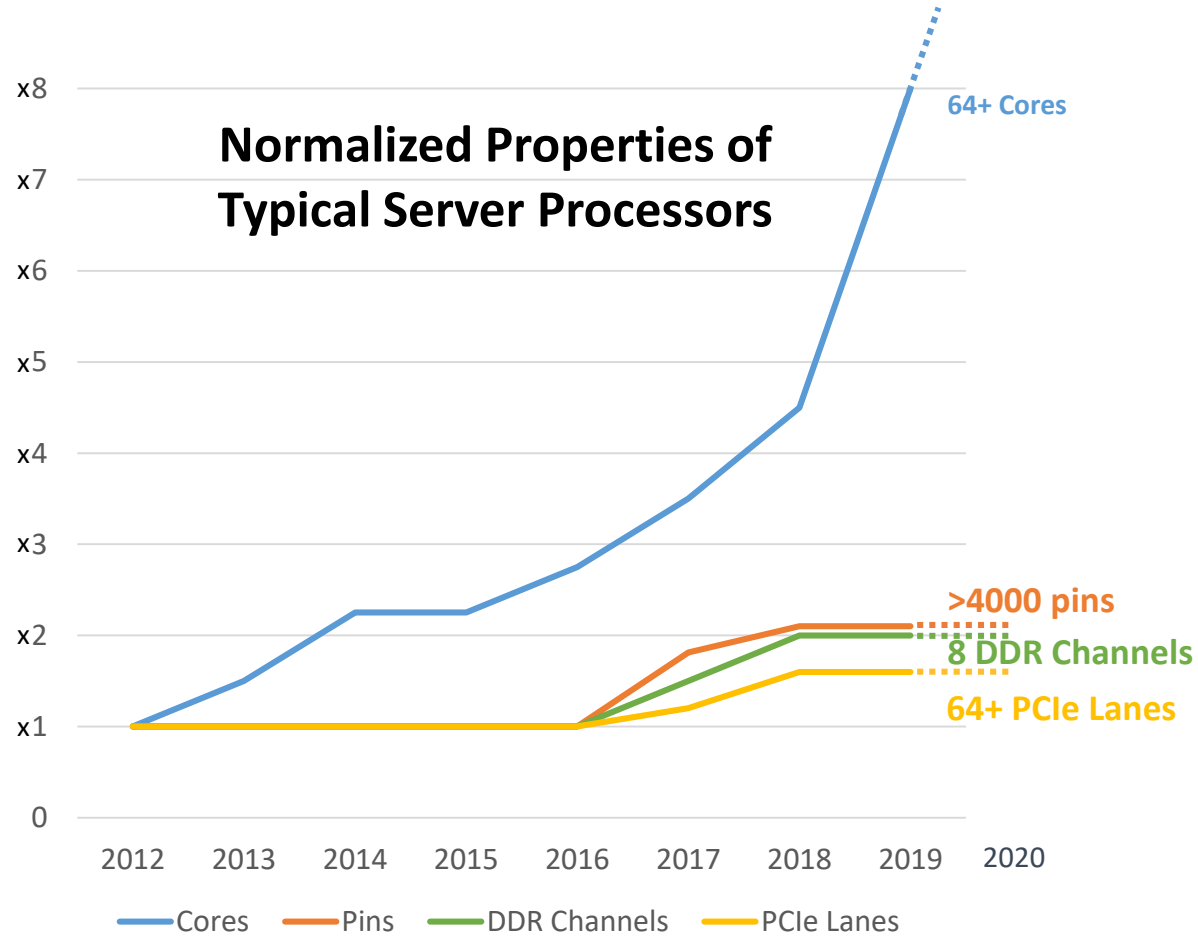


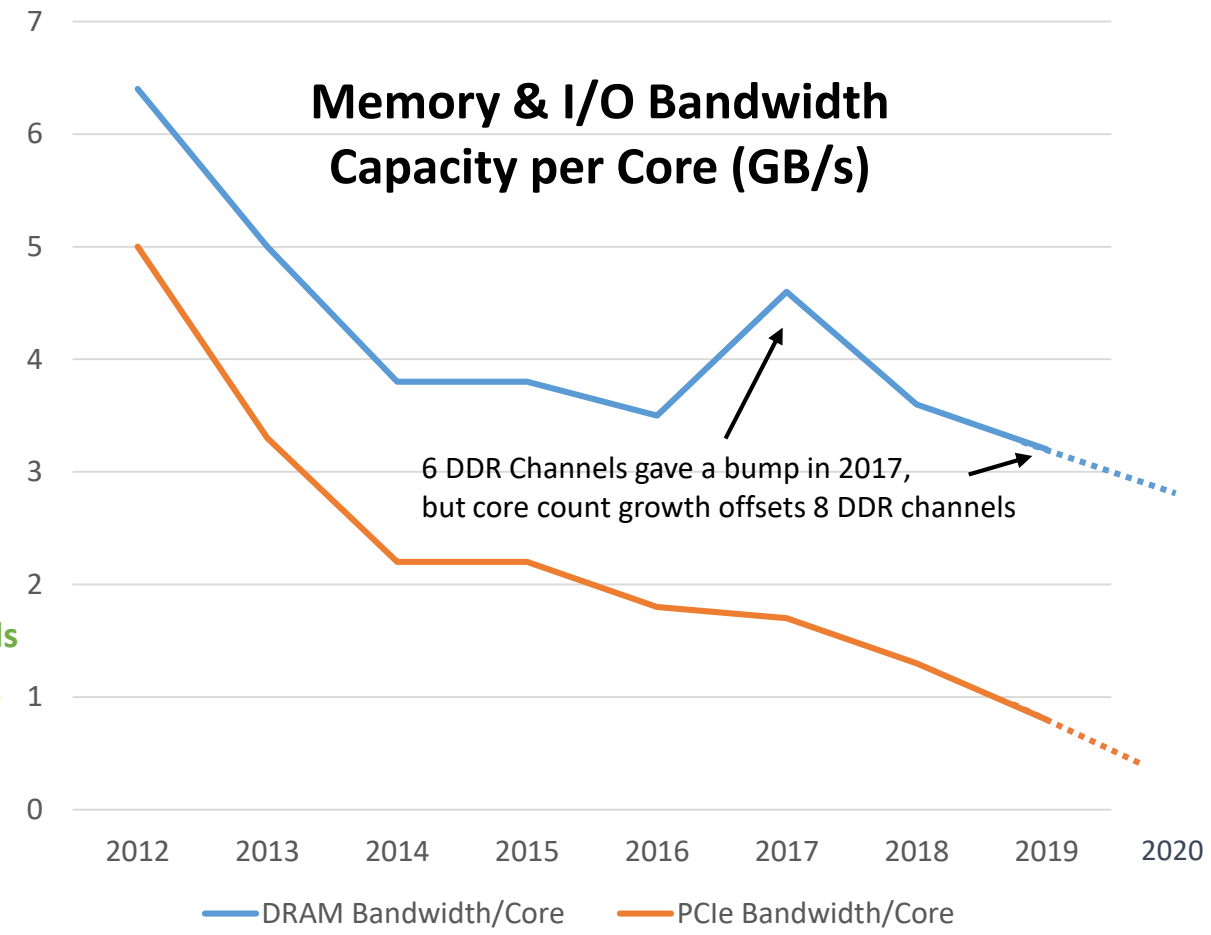
Future of Persistent Memory, DRAM and SSD Form Factors Aligned with New System Architectures

As first presented at SNIA Persistent Memory Summit 2021

Large Datasets Need Memory that Scales



Processor memory and I/O technologies …

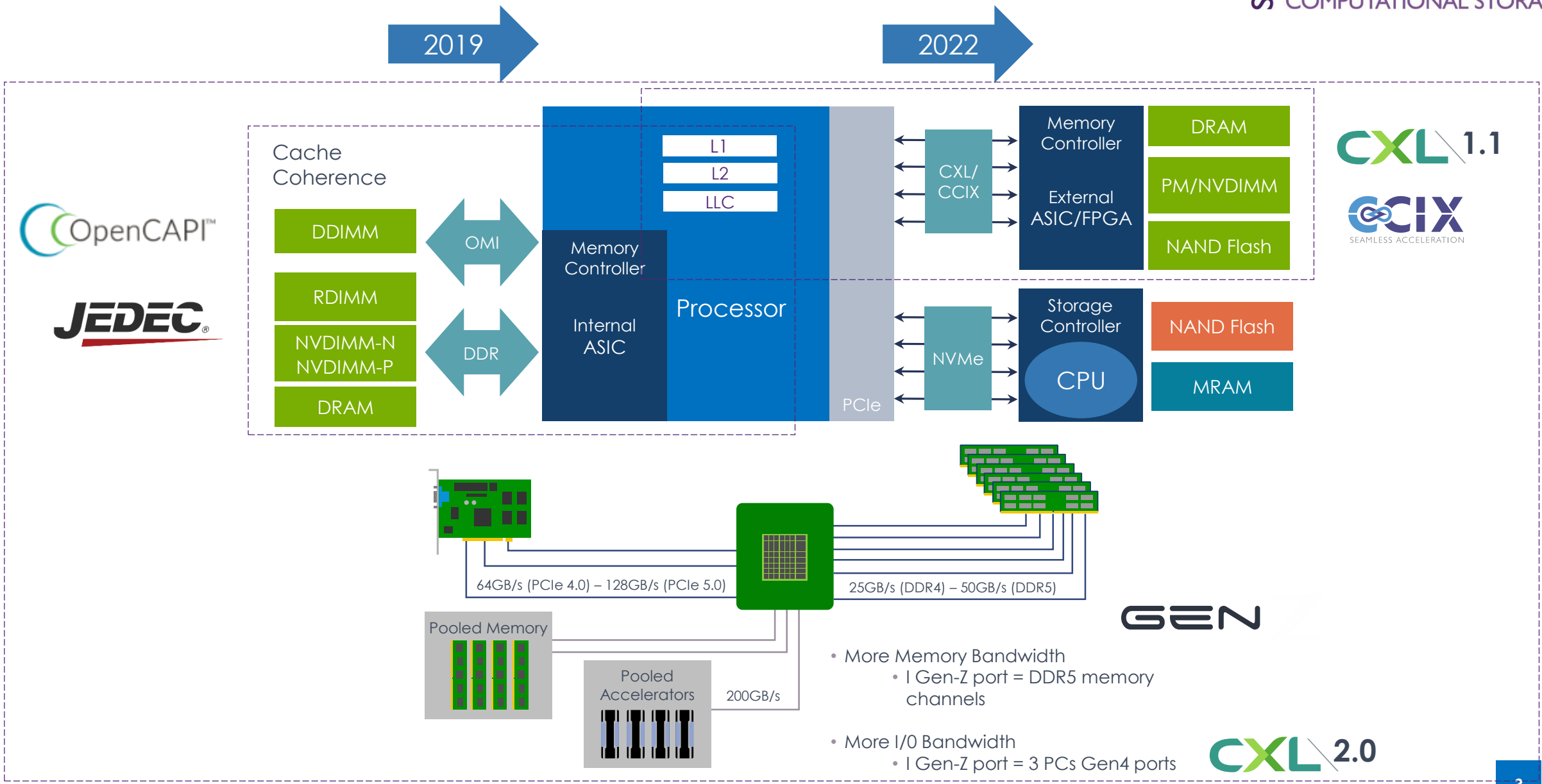


… are being stretched to their limits

More than 2X digital data will be created over the next five years compared to the combined amount since the advent of digital storage (Source; IDC, Mar 2020)

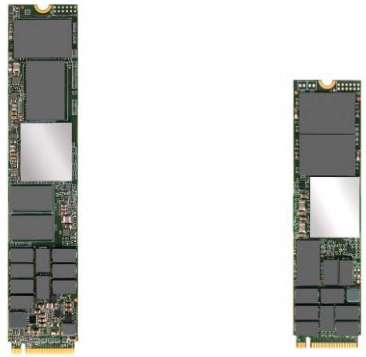
Source; Modified from Gen-Z Consortium

Memory and Accelerator Alignment with Fabrics



Source; Modified from Intuitive Cognition Consulting

Form Factor Migration



M.2

Performance Acceleration

E1.S

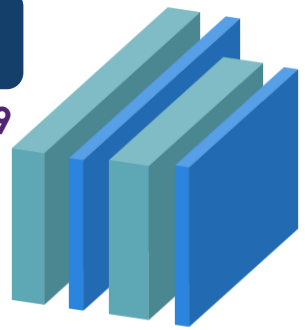
SFF-TA-1006

U.2

Memory Expansion

E3

SFF-TA-1009



E1.L

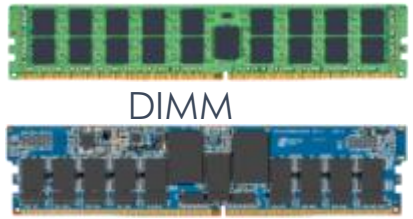
Source; Modified from J. Adiran, Microsoft SNIA, FMS, 2020

Performance

Direct attached (Parallel Bus) 100's of GB

Serial attached and PCIe attached 100's of GB to TB's

Network Attached TB's to PB's



DIMM

NVDIMM-N, NVDIMM-P (Persistent)

E1.S (x4)

E1.S (x8)

E3.S (x16)

PCIe AIC (x16)

OpenCAPI DDIMM

ZMM

	DDR DIMM	E1.S 1C (x4)	E1.S 2C (x8)	E3.S | AIC (x16)	E3 with x8 (2C)	Across network
Current Generation *	DDR4@3200 25.6GB/s	PCIe-Gen4-x4 7.8GB/s	PCIe-Gen4-x8 15.7GB/s	PCIe-Gen4-x16 31.5GB/s	OMI 25.6 GB/s 8 lanes	RDMA (Fabric and work load dependent)
Future Generation **	DDR5@4800 63.0GB/s	PCIe-Gen5-x4 15.7GB/s	PCIe-Gen5-x8 31.5GB/s	PCIe-Gen5-x16 63.0GB/s	DDR5 DDIMM (TBD) GB/s	Gen-Z, NVMe-oF (TBD)

* Source(s): https://en.wikipedia.org/wiki/PCI_Express#History_and_revisions
https://en.wikipedia.org/wiki/DDR4_SDRAM#JEDEC_standard_DDR4_module
https://www.smartm.com/media/press-releases/smart-modular-to-showcase-its-ddr4-differential-dimm-at-the-flash-memory-summit

** Source(s): https://www.tomshardware.com/news/ddr5-6400-ram-benchmarks-major-performance-gains-ddr4

Latency

Direct attached (Parallel Bus) 100's of GB

DIMM

NVDIMM-N, NVDIMM-P

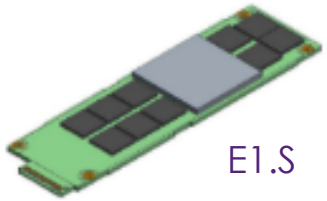
JEDEC

Serial attached and PCIe attached 100's of GB to TB's

DDIMM (40ns latency)

OpenCAPI™

Cache Coherent AIC



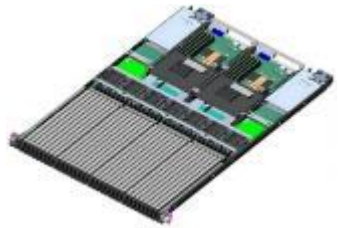
E1.S

CCIX SEAMLESS ACCELERATION

XMM, NV-XMM

CXL Compute Express Link

Network Attached TB's to PB's



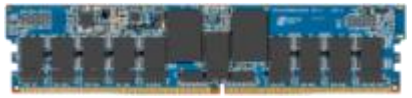
ZMM

GEN Z

Form Factors

Direct attached (Parallel Bus) 100's of GB

DIMM



NVDIMM-N, NVDIMM-P

Serial attached and PCIe attached 100's of GB to TB's

E1.S (x4)

E1.S (x8)

E3.S (x16)

E3.S (x8)

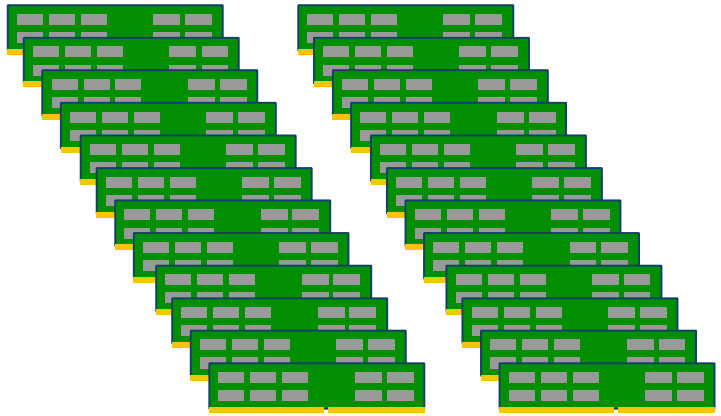
Network Attached TB's to PB's

	DDR4 DIMM	E1.S with x4 (1C)	E1.S with x8 (2C)	E3.S with x16 (4C)	E3 with x8 (2C)	Network Card
Pins	288 pins (64 data, 87 sideband, rest power)	56 pins (16 diff-data, 16 sideband, 24 power)	84 pins (32 diff-data, 18 sideband, 34 power)	140 pins (64 diff-data, 24 sideband, 52 power)	84 pins (32 diff-data, 18 sideband, 34 power)	Media and protocol specific
Connector (LxW)	142.0mm x 6.5mm	23.8mm x 6.0mm	35.6mm x 6.0mm	57.0mm x 6.0mm	35.6mm x 6.0mm	SFP/QSFP/…
Power	Input voltage=1.2V VPP 2.5	Input voltage 12V Vaux 3.3 (optional)	Input voltage 12V Vaux 3.3 (optional)	Input voltage 12V Vaux=3.3 (optional)	Input voltage 12V	Vendor specific

Bandwidth

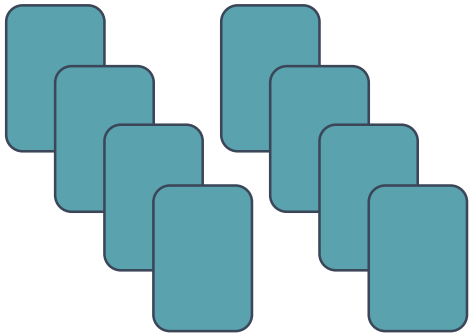
2022 CPU
96 cores

12 Channels
2 DIMMs/Channel
128GB DIMM
3TB Memory



- Theoretical maximum bandwidth of 921.6 GB/S
- 9.6 GB/s per core

8 x8 links
256GB E3.S
2TB Memory

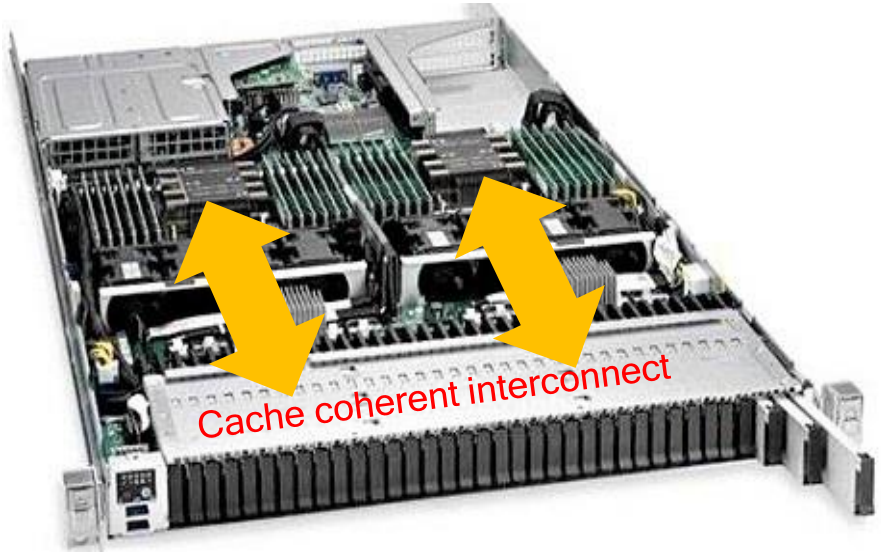


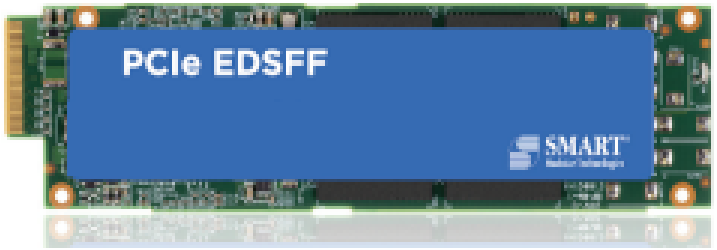
- Theoretical maximum bandwidth of 252.16 GB/s
- 2.6 GB/s per core.

Estimated higher bandwidth of 12.2 GB/s per core

Source; Modified from CXL Consortium Webinar Aug 2020

E1.S and E1.L for Memory Acceleration





Feature	Description
Host Interface	• Data: PCIe x4,x8 • Sideband: SMBus (I2C) • Wake-up, Low-power (PWRDIS), …
Memory	64-128GB with DDR4 or DDR5
Protocols	NVMe, CXL, CCIX, Gen-Z
Power	• Multiple profiles from 12,16, 20, 25W • Completely bus powered: 12V (main), 3.3V Aux • Supports low power modes (CLKREQ#, PWRDIS signaling)
Targeted Use Cases Memory Acceleration	Targeted for 1U Servers • 16 – 32 Slots per 1U Server • Improves performance by offloading fixed functions like encryption, compression or Key-Value semantics to Memory modules.

Source: https://www.anandtech.com/show/16248/edsff-form-factor-updates

E3.S and E3.L for Memory Expansion

Feature	Description
Host Interface	• Data: PCIe x16 • Sideband: SMBus (I2C) • Wake-up, Low-power (PWRDIS), …
Memory	Up to 256GB with DDR4 or DDR5 ** A Non-volatile persistent memory feature could be support on this form-factor using back-up and restore functionality like in NVDIMM-N.
Protocols	NVMe, CXL, CCIX, Gen-Z
Power	• 2 profiles 25W (thin) and 40W (thick) • Bus powered: 12V (main), 3.3V Aux • Supports low power modes (CLKREQ#, PWRDIS signaling)
Targeted Use Cases Memory Expansion	Targeted for 2U Server Enables 4TB – 8 TB of Memory expansion with 16 E3.S modules in single 2U server, achieving better throughput than direct attached DDR4 DIMM.

OpenCAPI High Bandwidth Memory - DDIMM

OpenCAPI™

Feature	Description
Host Interface	• OpenCAPI
Memory	• Up to 256GB
Protocols	• OMI – Open Memory Interface • The memory bus is defined with one read port and one write port per channel, each having eight unidirectional differential lanes
Performance	• DDR4-3200 • Latency 40ns • Data throughput rate of 25.6GB/s with 8 lanes • The DDIMM/OMI approach delivers up to 4TB of memory on a server at about 320GB/second or 512GB at up to 650GB/s sustained rates.
Targeted Use Cases	• Targeted for servers • High bandwidth, low latency serial connection for memory, accelerators, network, storage, and other devices like ASICs

NVDIMM for Persistent Memory

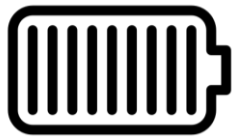
Key Features of DDR4 and DDR5 NVDIMM-N

- Operation like DRAM
- Fast recovery from system power loss
- Software overhead can be eliminated

JEDEC

SNIA

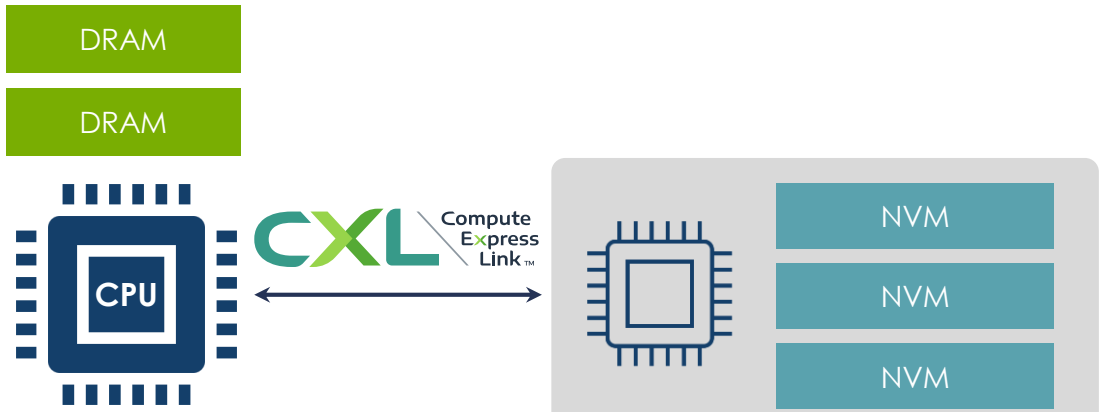
+

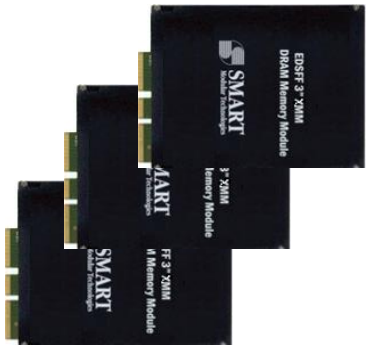


Backup Power

Feature	Description
Host Interface	• DDR
Memory	• DDR4 16GB, 32GB • DDR5 32GB, 64GB
Protocol	• JEDEC Compliant DDR4 / DDR5
Features	• Throughput of 25.6GB/s (DDR4) • Latency ~20ns • AES 256 bit Encryption
Targeted Use Cases	• All Flash Arrays, Storage Servers, HPC, AI Training Servers • Needed for very low latency tiering, caching, write buffering, metadata storage, checkpointing • Needed for AI/ML algorithm processing

CXL-based NVDIMM (NV-XMM)





Source: Modified from PIRL 2019, "Accelerate Everything", Stephen Bates, Eideticom

High-speed DRAM with built-in back-up power to back-up data to on-module Flash during power loss.

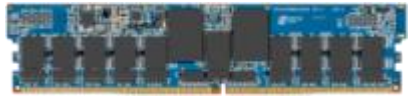
Source: PIRL 2019, "Accelerate Everything", Stephen Bates, Eideticom

Conclusions

Future of Persistent Memory, DRAM and SSD Form Factors

JEDEC

DIMM



NVDIMM-N, NVDIMM-P
(Persistent)

JEDEC

OpenCAPI™

PCIe AIC (x16)

OpenCAPI
DDIMM

Dependence
on CPU

CCIX
SEAMLESS ACCELERATION

CXL Compute Express Link™

E1.S (x4)

E1.S (x8)

E3.S (x16)

Scalability, Serviceability, Savings
Modularize the Technology

GEN Z

2020

2022

2025