SNIA PERSISTENT MEMORY + COMPUTATIONAL STORAGE SUMMIT 2022

VIRTUAL EVENT • MAY 24-25, 2022

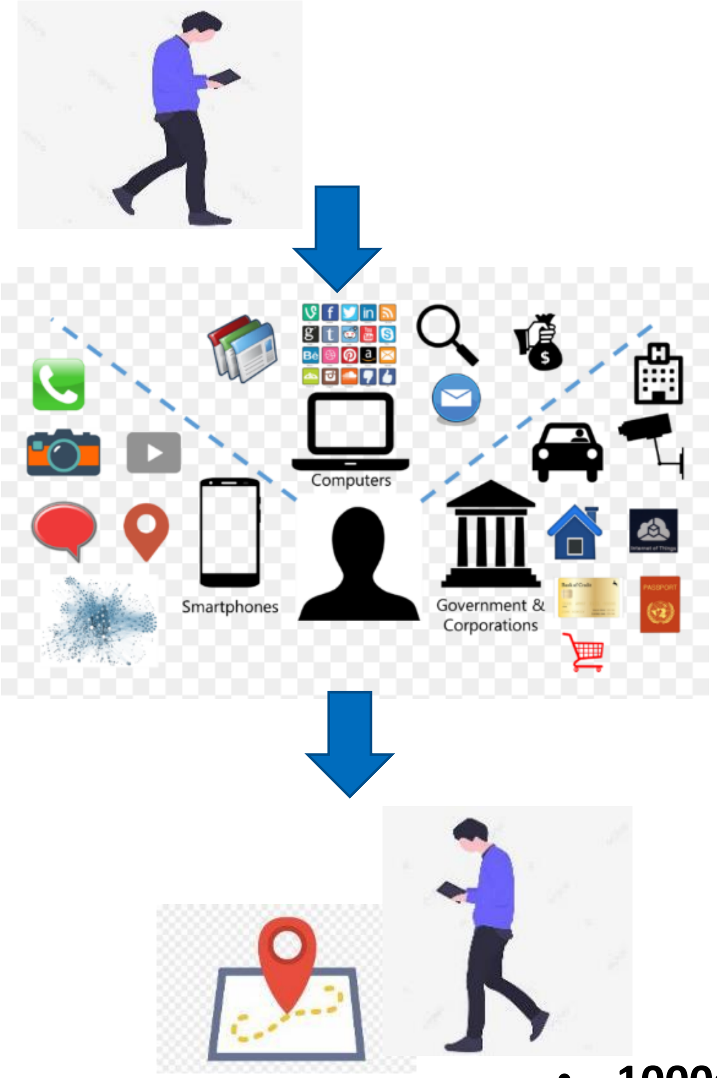
Computational Memory: Moving compute near the Data

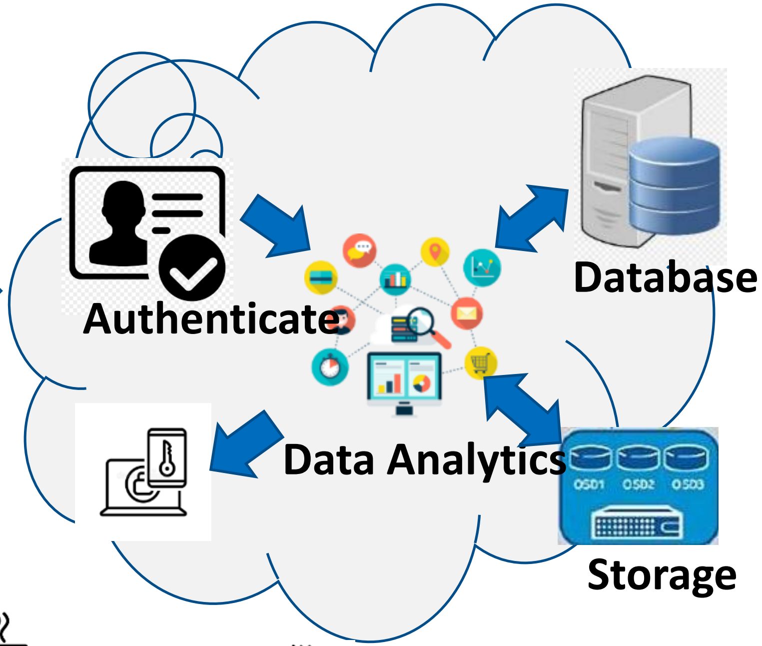
Presented by

Pekon Gupta <Pekon.Gupta@smartm.com>

Arthur Sainio <Arthur.Sainio@smartm.com>

Data Flow: Logical View





- **1000s of such queries generated per second.**
- **Result should be returned within few seconds.**

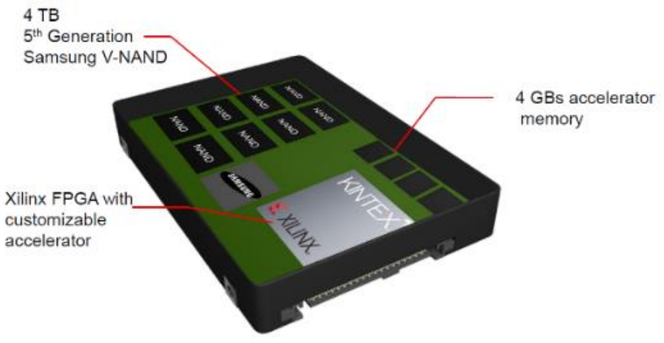
SNIA PERSISTENT MEMORY + SUMMIT 2022 COMPUTATIONAL STORAGE

Types of Hardware Accelerators

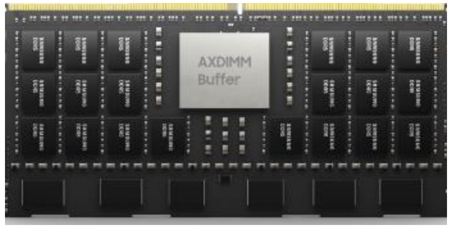
Network Accelerators (Smart NIC [1a/1b])

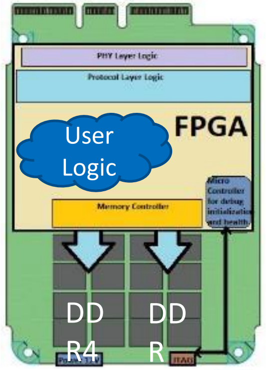
Storage Accelerators (Computational Storage[2a,2b])

SmartSSD® CSD



Memory Accelerators (Computational Memory[3a])



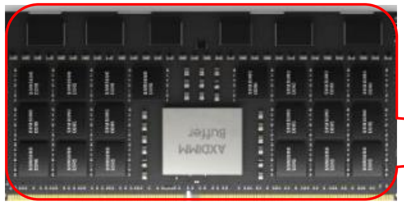


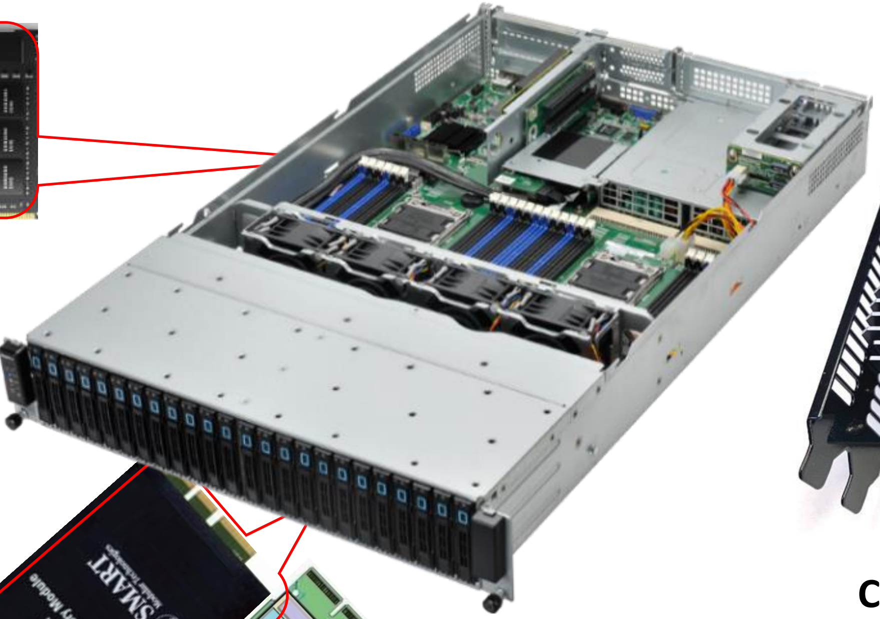
Processing in Memory

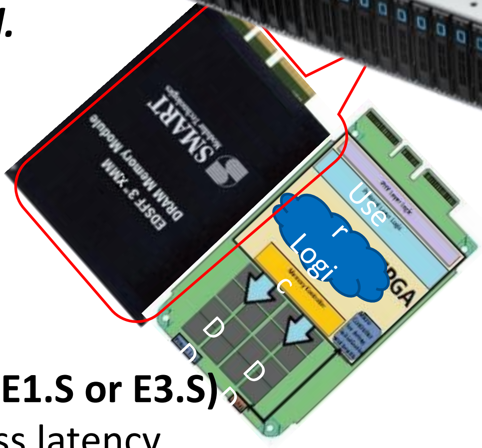
[1a] https://blocksandfiles.com/2020/10/28/amd-xilinx-smartnic-data-centre/
[1b] https://www.servethehome.com/intel-fpga-pac-d5005-high-end-drop-in-accelerator-launched
[2a] https://www.servethehome.com/xilinx-samsung-smartssd-computational-storage-drive-launched/
[2b] https://www.servethehome.com/intel-fpga-pac-d5005-high-end-drop-in-accelerator-launched/

[3a] https://tekdeeps.com/samsung-also-sees-the-future-in-memories-that-also-perform-calculations/

Form-factors for Computational Memory Devices (CMD)







CMD in DIMM

- ✓ Low latency as directly attached to DDR bus
- ❖ Memory capacity limited by Thermal and PCB.
- ❖ PCB real-estate design limits the acceleration.
- ❖ ***May bring down the speed of entire DDR channel if not running at maximum speed.***

CMD in EDSFF (E1.S or E3.S)

- ❖ Higher access latency
- ✓ High Memory capacity.
- ✓ **Scalable and Hot pluggable**
- ✓ Sustain up to 40W TDP for E3.S 2T

CMD in PCIe Card Electro Mechanical (CEM)

- ❖ High access latency.
- ✓ Very High Memory capacity. (In Tera Bytes)
- ✓ Scalable. Allows user to choose Memory capacity and select DIMM.
- ✓ Offload Heavy functions. Can scale up to 100s of Watt of TDP.
- ✓ **CPU and Platform Agnostic**

SNIA PERSISTENT MEMORY + SUMMIT 2022 COMPUTATIONAL STORAGE

Optane based Memory accelerator

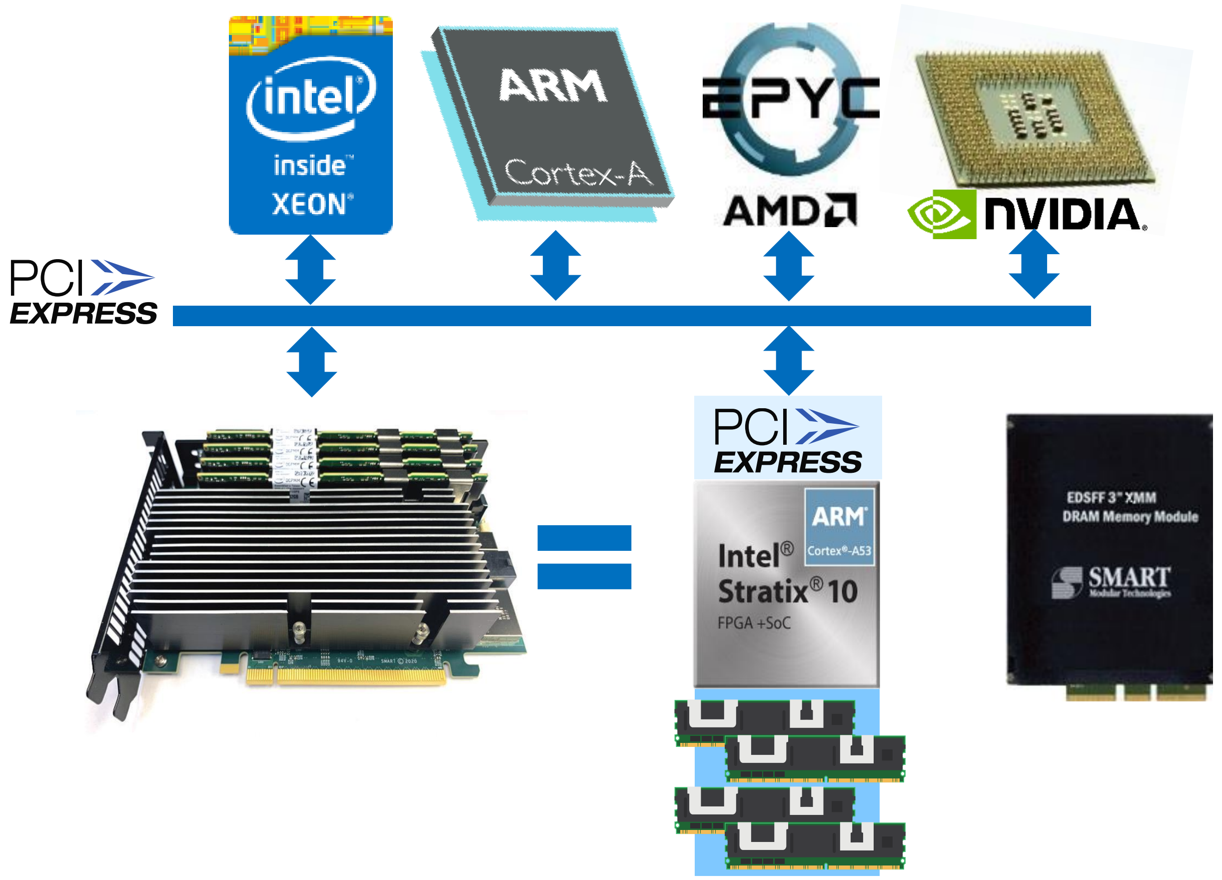
Optane Memory Expansion and Memory Accelerator

Wow! **Computational Memory Device can**
- Offload Host CPU
- Run analytics on parallel nodes
- Offload without any custom stack

- System Integrator

Feature	Description
Host Interface	• PCIe Gen4 x16
Memory	• Four DDR4 DIMM Slots (2 DIMMs Per Ch) • Up to 2TB (4 x 512GB) Intel Optane Persistent Memory DIMM or • Up to 512GB (2 x 256GB) DDR4 RDIMM
Form Factor	• FHHL (Full-Height Half-Length) Dual Slot
Acceleration	• Hardened Quad core ARM A53 with dedicated memory. • Hardware engines for compression, encryption and search can be implemented on Intel Stratix-10 DX FPGA
Power	• Less than 150W TDP for Memory expansion and Acceleration.
Cooling	• Passive cooling. Airflow requirement up to 45 CFM

Processing in Memory (PIM): Moving compute near the Data

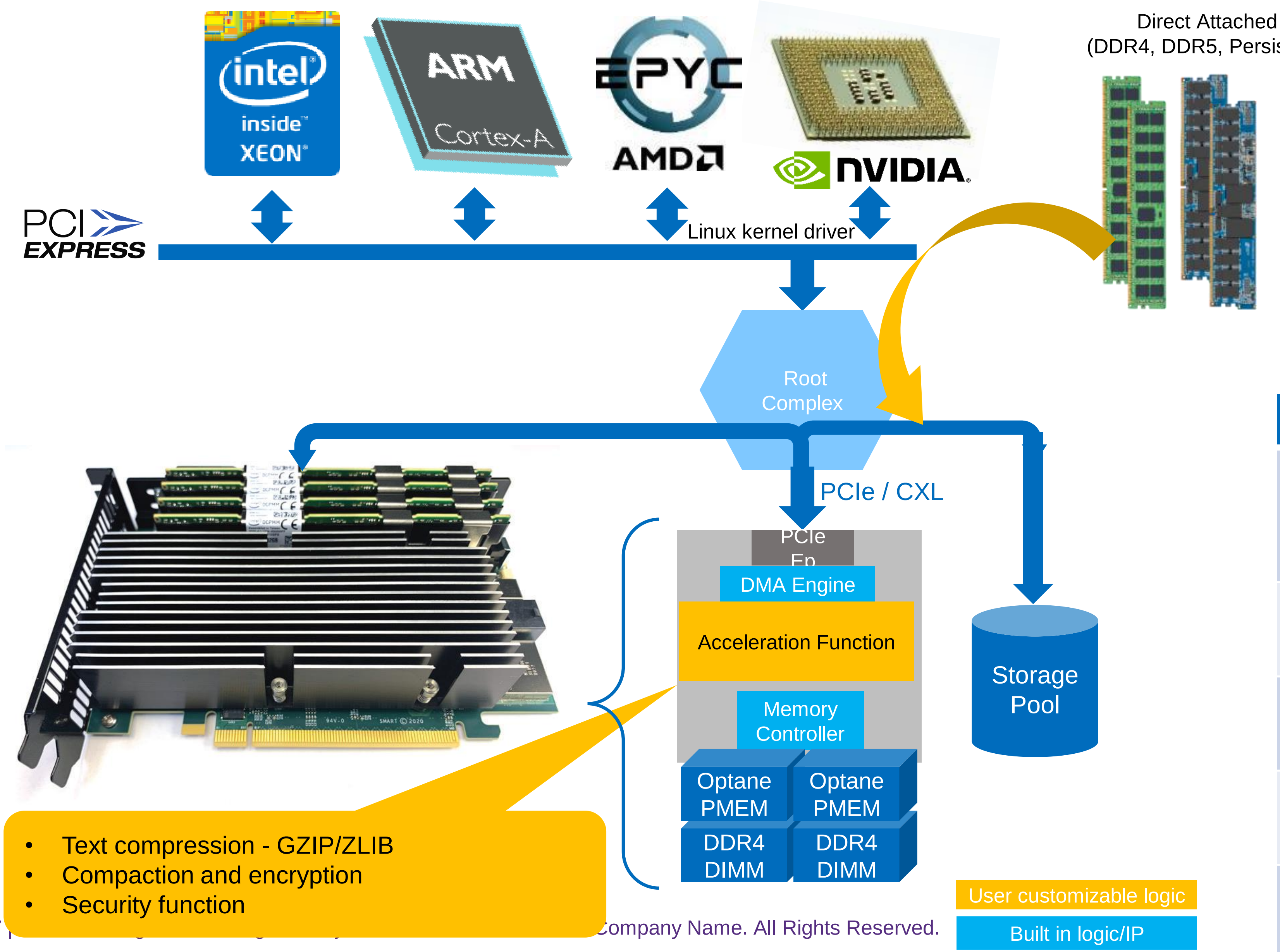


CPU Agnostic
- Abstraction of Memory technology (DDR4, DDR5, LPDDR-X or any Persistent Memory) from CPU architecture
- Memory can be upgraded independent of CPU architecture

Improve Utilization, Reduce Copy and discard.
- Pre filter, search and organize the data before copying.
- Processing on multiple data-sets in parallel.
- Free up CPU from mundane tasks.

Scalability
- PCIe CEM allow standard off-the-shelf Memory DIMMs for both Memory acceleration and expansion
- EDSFF devices are Hot Pluggable.

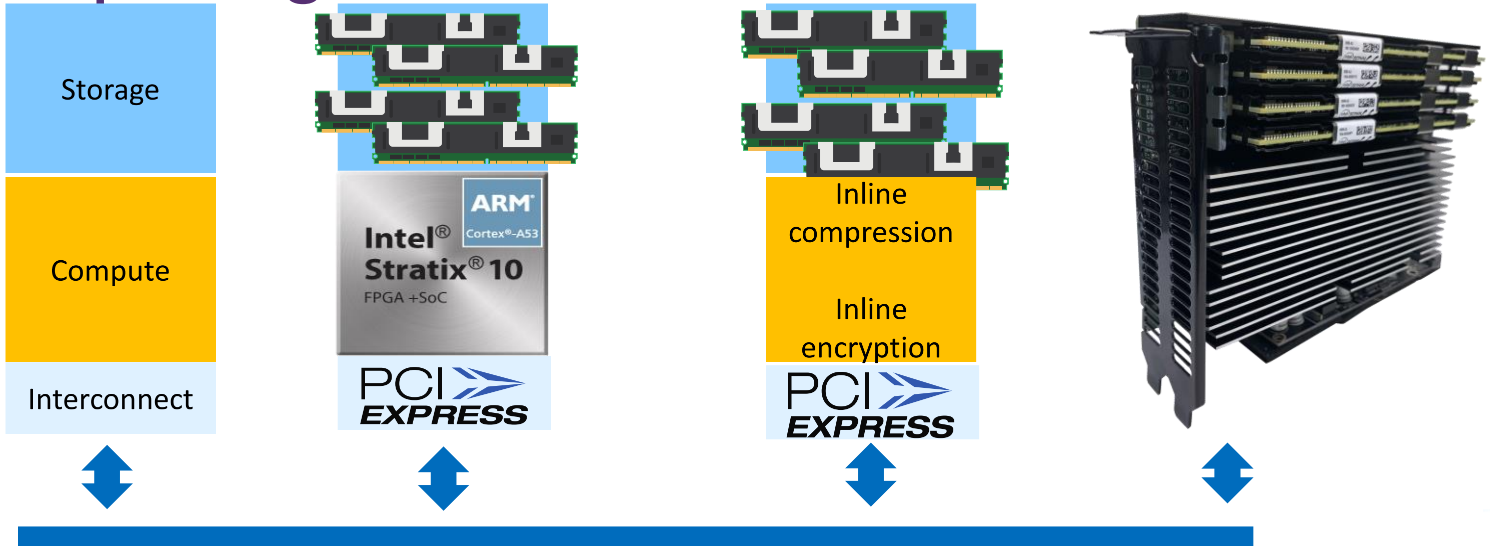
Computational Memory Device as Storage Cache



Benefits

- CPU can use 2TB of “persistent” Memory as “Storage Cache”.
- Byte addressable Memory Tier in between Storage and DDR4/DDR5 Memory DIMMs.
- Open source driver support and platform agnostic attachment.
- Pluggable in existing chassis with passive cooling.
- Offload in-line functions while read or writing data to storage.

Comparing the numbers on real hardware: Kestral



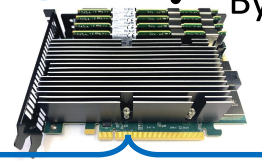
Lower latency than fastest SSD

Built in Acceleration

- In-line compression and encryption
- Key-Value /Content Addressable search

CPU agnostic

- Re-use existing PCIe infrastructure
- Byte level addressable.



Parameters	Samsung NVMe SSD	Intel Optane NVMe SSD	Kestral PCIe-Gen3-x16 (4 Apache Pass/2DPC)	Kestral PCIe-Gen4-x8 (4 Apache Pass/2DPC)
Capacity	2TB	2x0.8TB=1.6TB	2TB	2TB
Host Interface	2 x (PCIe Gen4 x4)	2 x (PCIe Gen4 x4)	PCIe Gen3 x16	PCIe Gen4 x8
Random Read Perf (GB/s)**	2x2.72=5.44	2x6.2=12.4	4.3	8.0***
Random Write Perf (GB/s)**	2x2.52=5.04	2x6.4=12.8	2.4	7.0***
Sequential Read Perf (GB/s)**	2x7.0=14	2x7.4=14.8	9.7	12.2
Sequential Write Perf (GB/s)**	2x5.1=10.2	2x7.4=14.8	9.7	9.9
Read Latency (QD=1)*	20us/(Z-NAND SSD)	<6us	TBD	< 0.3us
Write Latency (QD=1)*	16us(Z-NAND SSD)	<6us	TBD	< 0.5us

Higher is better

Lower is better

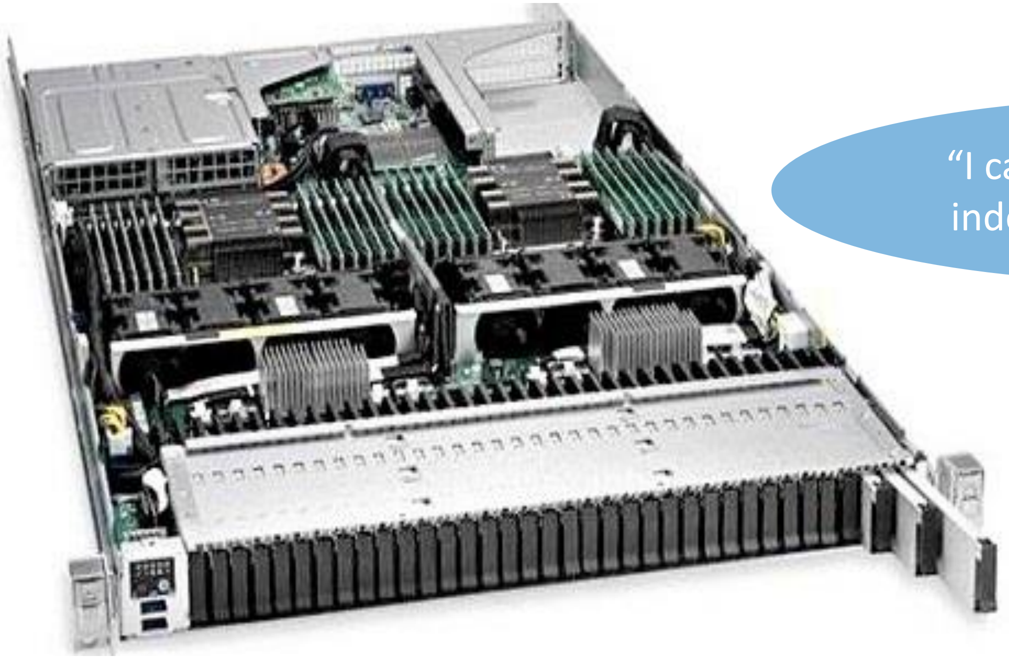
** Converting IOPS to GB/s assuming IO size = 4KB. to convert IOPS to GB/s. Also assuming two x4 SSD connected in parallel to compare with PCIe-Gen4-x8.

*** Projected numbers extrapolated from PCIe-Gen3-x15 and PCIe-Gen4-x8 testing.

- Samsung NVMe SSD: Samsung 980 Pro https://www.samung.com/us/computing/memory-storage/solid-state-drives/980-pro-pcie-4-0-nvme-ssd-2tb-mz-v8p2t0b-am/#specs
  Samsung Z-NAND https://semiconductor.samsung.com/resources/brochure/Ultra-Low%20Latency%20with%20Samsung%20Z-NAND%20SSD.pdf

Intel Optane NVMe SSD: https://www.intel.com/content/www/us/en/products/docs/memory-storage/solid-state-drives/data-center-ssds/optane-ssd-p5800x-p5801x-brief.html

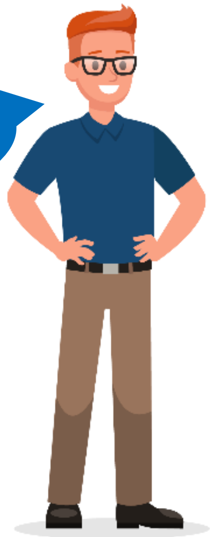
Summarizing benefits of Computational Memory architecture

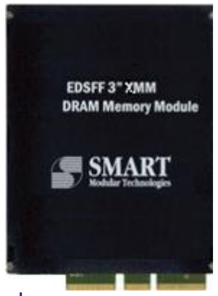


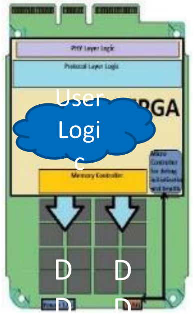
“I can upgrade the memory independent of Processor”

“I can do more with less”

“Frees CPU cycles and reduces wasteful copy and discards.”







Please take a moment to rate this session.

Your feedback is important to us.