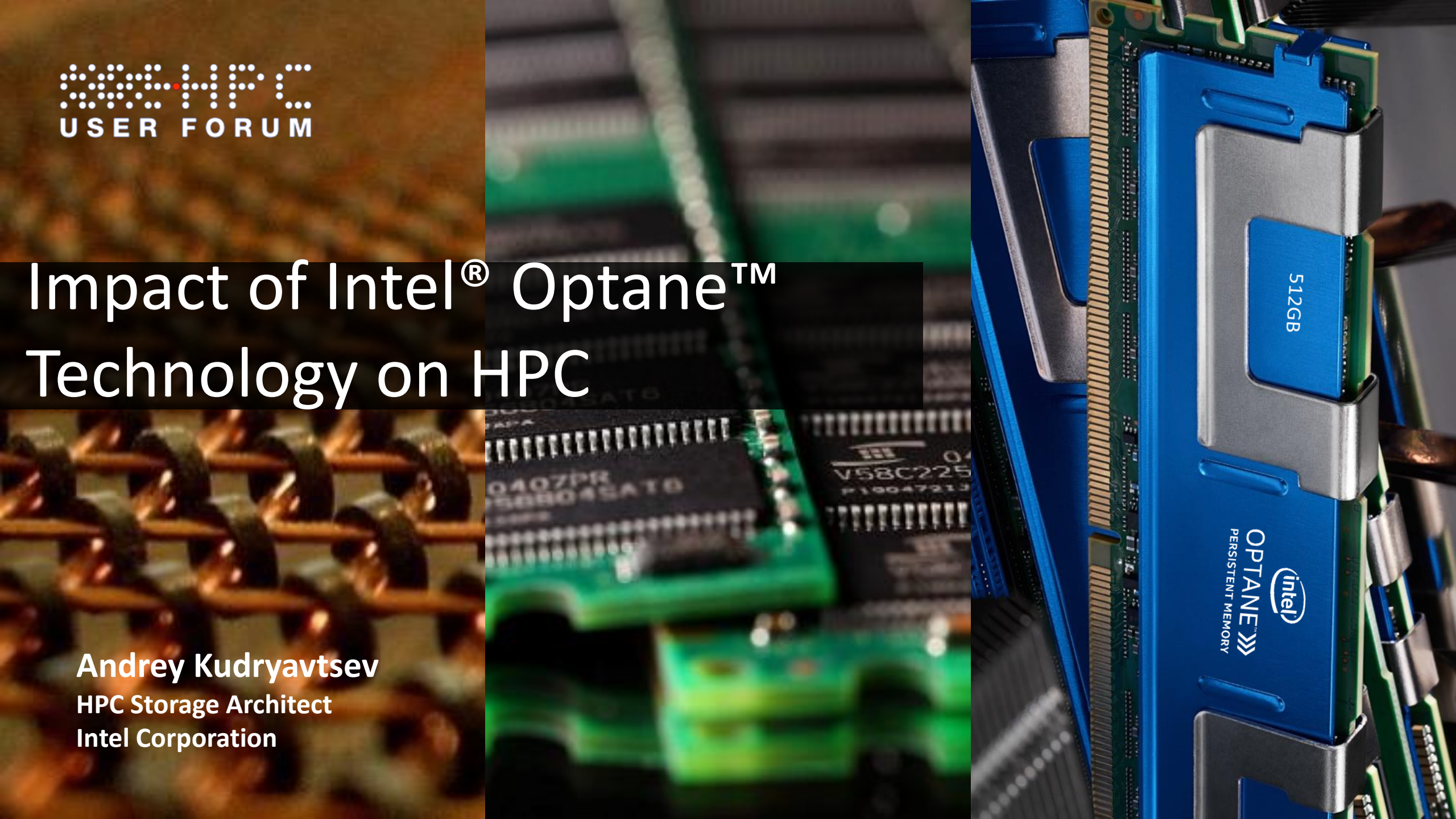
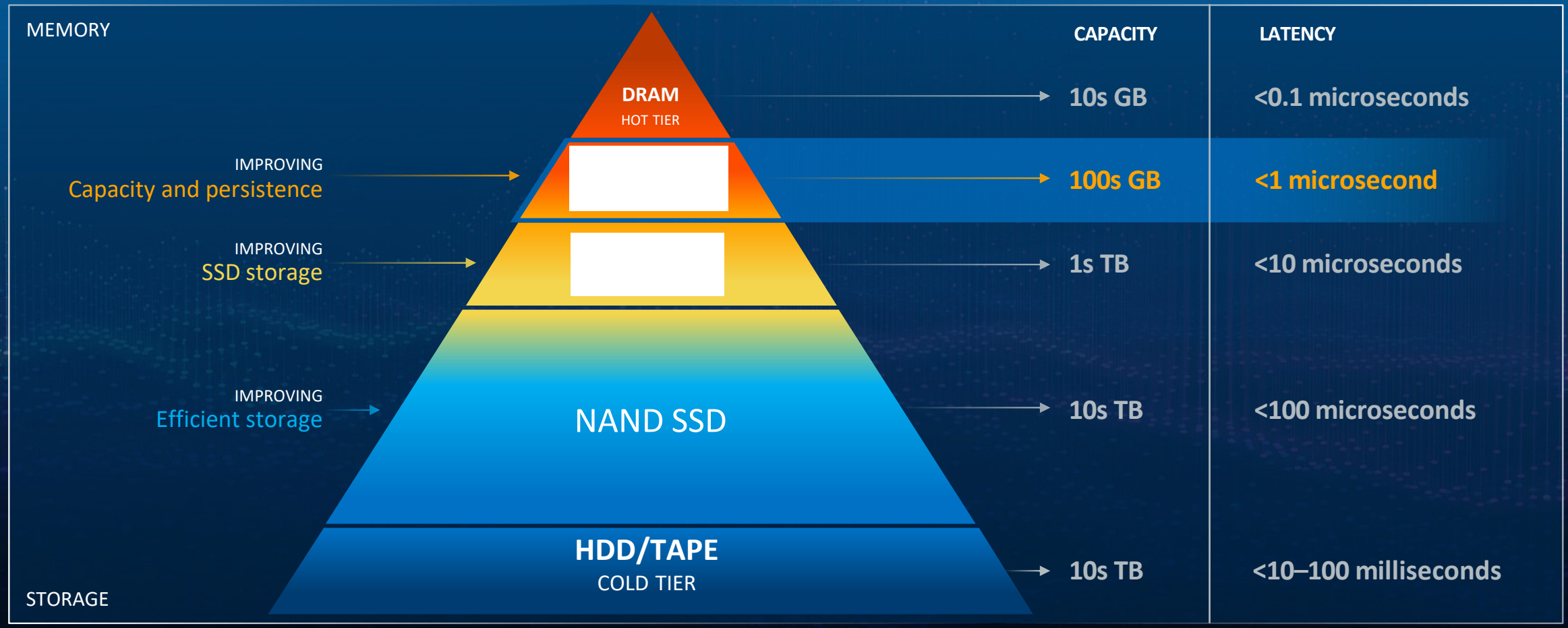


Impact of Intel® Optane™ Technology on HPC

Andrey Kudryavtsev
HPC Storage Architect
Intel Corporation

Re-architecting the Memory/Storage Hierarchy

Intel® Optane™ persistent memory (PMem) 200 series fills the DRAM gap



Accelerate Insight and Innovation

Intel® Optane™ persistent memory (PMem) 200 series fills the DRAM gap

intel
OPTANE™
PERSISTENT MEMORY

Extract more from larger datasets. Expand your memory pool in persistent memory and support near-real-time data analysis; deliver deep insights, improve operations, or create new revenue streams.

Lower overall total cost of ownership (TCO).[3] Do more with each server—increase CPU utilization, in-memory database capacity, throughput, virtual machine (VM) density, and services for users on a consolidated footprint.

Protect data automatically. Help secure all data at rest in persistent memory with application-transparent AES-256 encryption, which enhances security without code changes.

High performance. High capacity. Hardware-enhanced security.

[3] Supports for more than 1.2x more VMs than the previous generation. Based on testing by Intel as of April 27, 2020 (Baseline) and March 31, 2020 (New). For details, see Endnotes slides.

persistent memory: in a league of its own

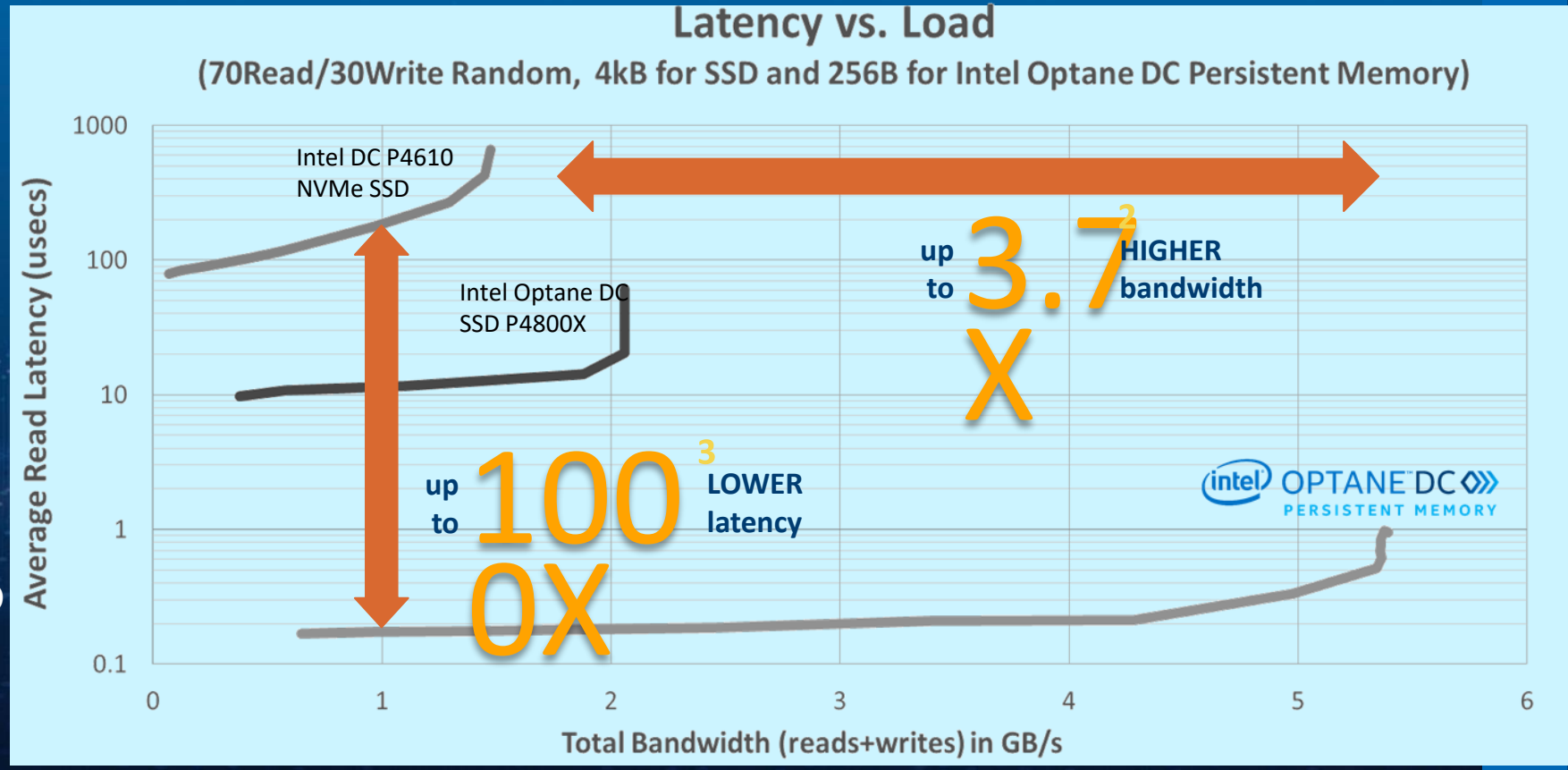
HIGHER BANDWIDTH

Up to **3.7X** read/write bandwidth vs NVMe SSDs, with one module; more with multiple modules

LOWER LATENCY

Orders of magnitude lower latency than NVMe SSDs

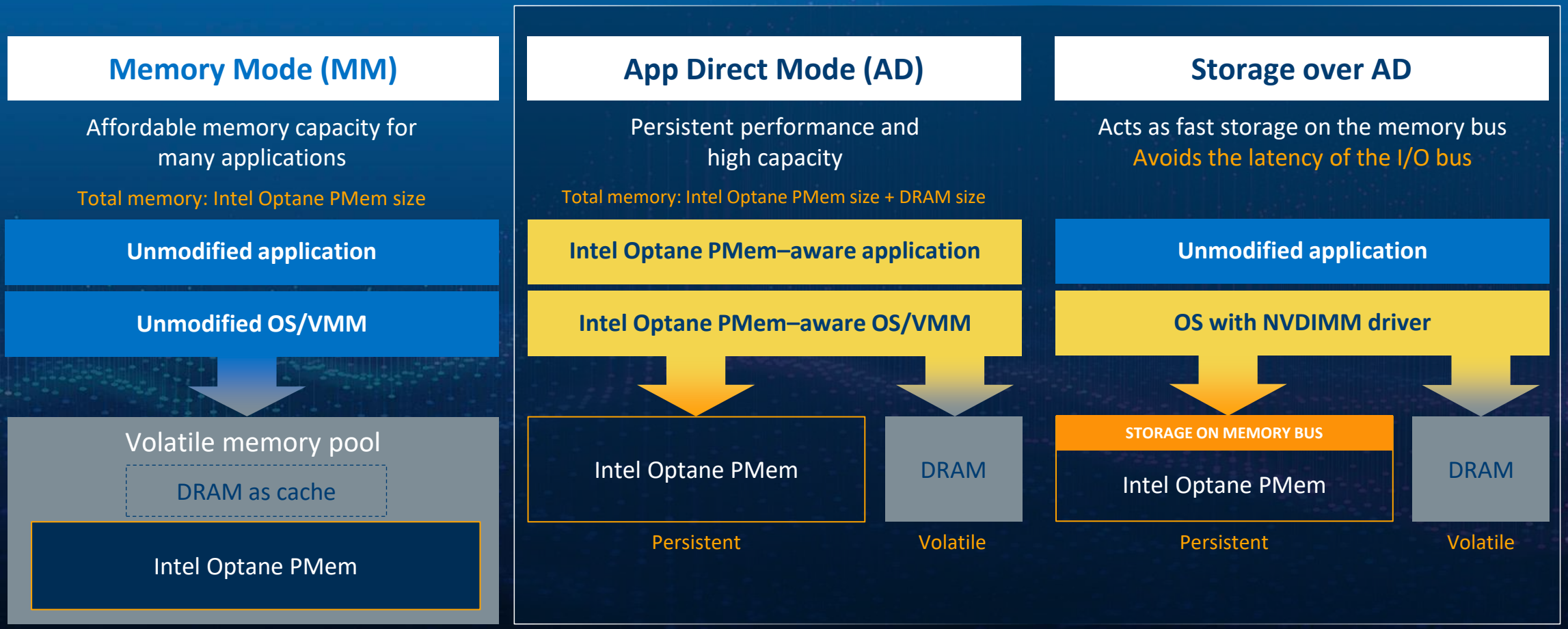
- **1000X** lower latency than NAND NVMe SSD at 1GB/s



Performance results are based on testing as of Feb 22, 2019 and may not reflect all publicly available security updates. No product or component can be absolutely secure.
Results have been estimated based on tests conducted on pre-production systems, and provided to you for informational purposes. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to www.intel.com/benchmarks.

[2, 3] Based on internal testing by Intel . Configuration: see slides 38 and 39.

Intel® Optane™ Persistent Memory (PMem) 200 Series Operating Modes



Intel® Optane™ Persistent Memory (PMem) at a Glance

intel OPTANE PERSISTENT MEMORY

UP TO 25% higher memory bandwidth over the previous generation[3]

AES 256-BIT encryption

Secure Erase

UP TO 512 GB modules

12–15 W thermal design power

Intel Optane PMem 100 series

intel XEON PLATINUM

2nd Generation Intel® Xeon® Scalable processors
SP on 2S/4S/8S platform

6 channels memory	**2,666 MT/s** DDR4
3 TB Intel Optane PMem per socket*	**18 W** top sustained

Intel Optane PMem 200 series

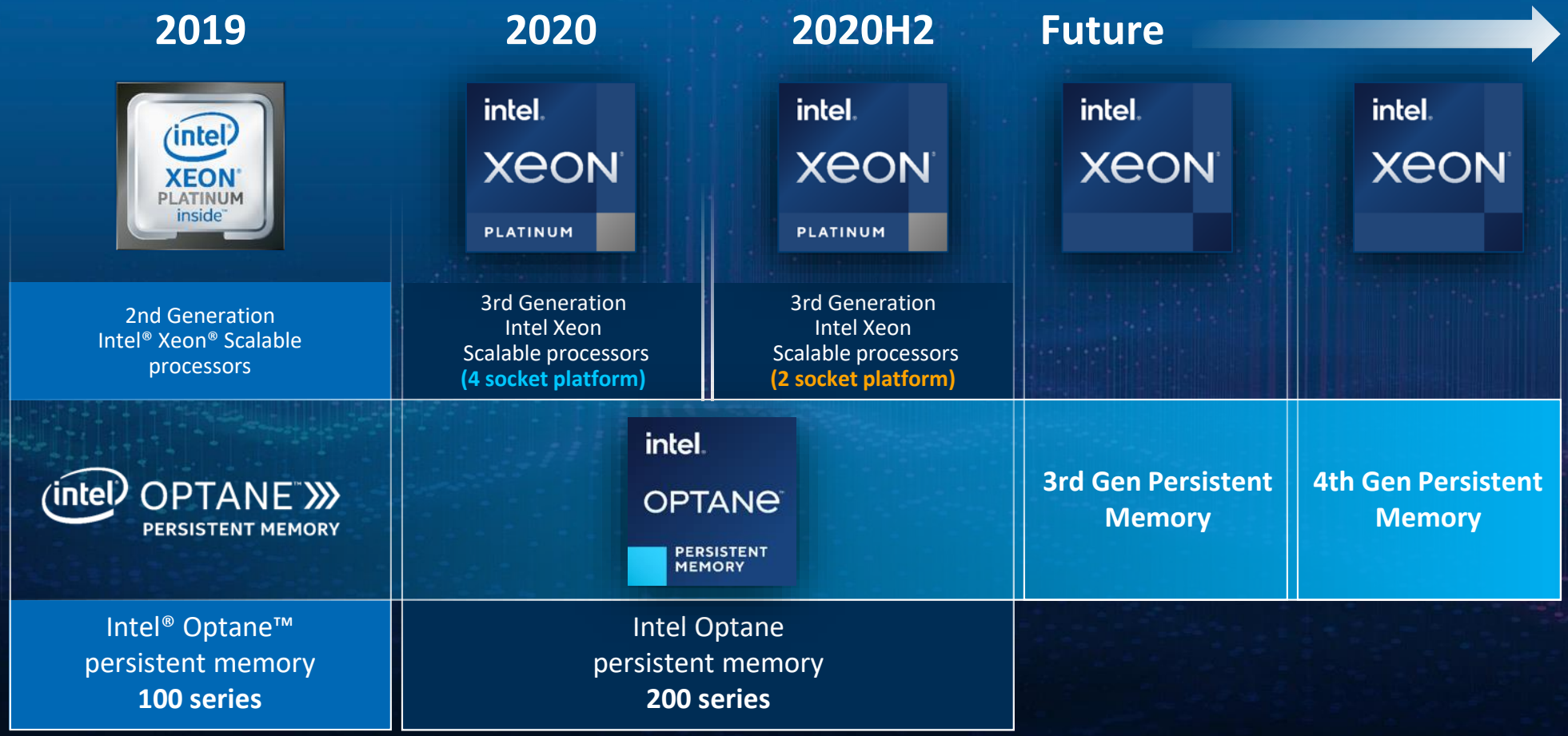
intel XEON PLATINUM

3rd Generation Intel Xeon Scalable processors
SP on 4S platform

16–28 cores	**6 channels** memory
3 TB Intel Optane PMem per socket*	**2,666 MT/s** DDR4

See backup for configuration details. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.
[3] Based on testing by Intel as of April 27, 2020 (Baseline) and March 31, 2020 (New). For details, see Endnotes slides.
* 3 TB = 6 x 512 GB Intel Optane PMem per socket
** 4 TB = 8 x 512 GB Intel Optane PMem per socket

A Strong Memory and Storage Future



For each processor above, Intel Optane PMem will be supported on select SKUs.

intel® optane™ technology breaks memory and storage bottlenecks

compute

acceleratin

keep more data in memory

DRAM

Intel® Optane™ Persistent Memory

Caching

copy of less frequently used data from DRAM

DRAM

Intel Optane Persistent Memory

Tiering

app-aware data placement

DRAM

Intel Optane Persistent Memory

storage

intel OPTANE SSD

acceleratin

Intel® Optane™ SSD

store data about data

Intel® 3D NAND SSD

Caching

Intel Optane SSD

temporarily copy or hold hottest data

Intel 3D NAND SSD

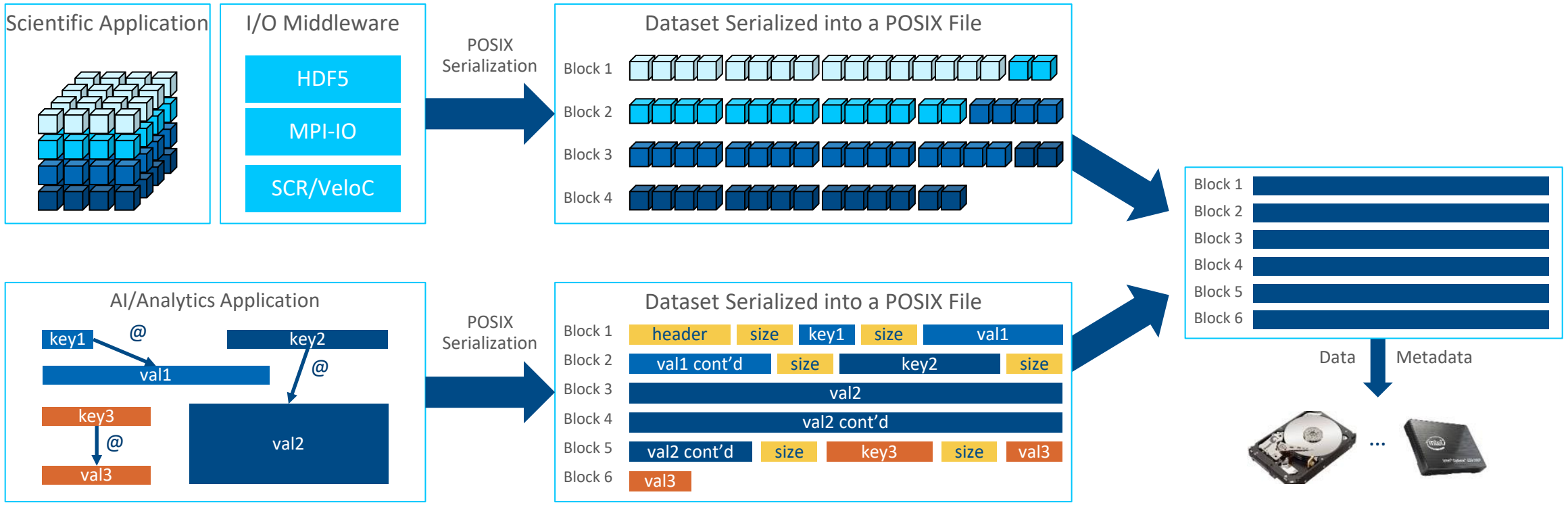
Tiering

intelligent data placement

Intel Optane SSD

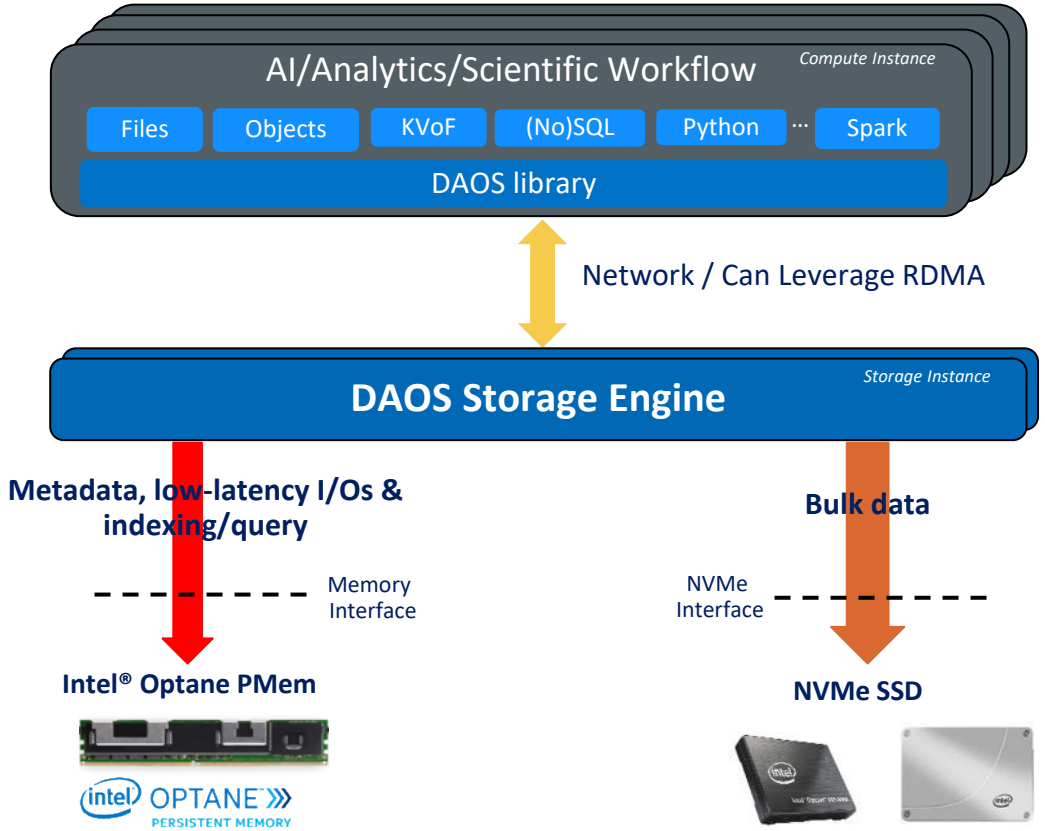
Intel 3D NAND SSD

The Problem with POSIX & Blocks (& Objects)



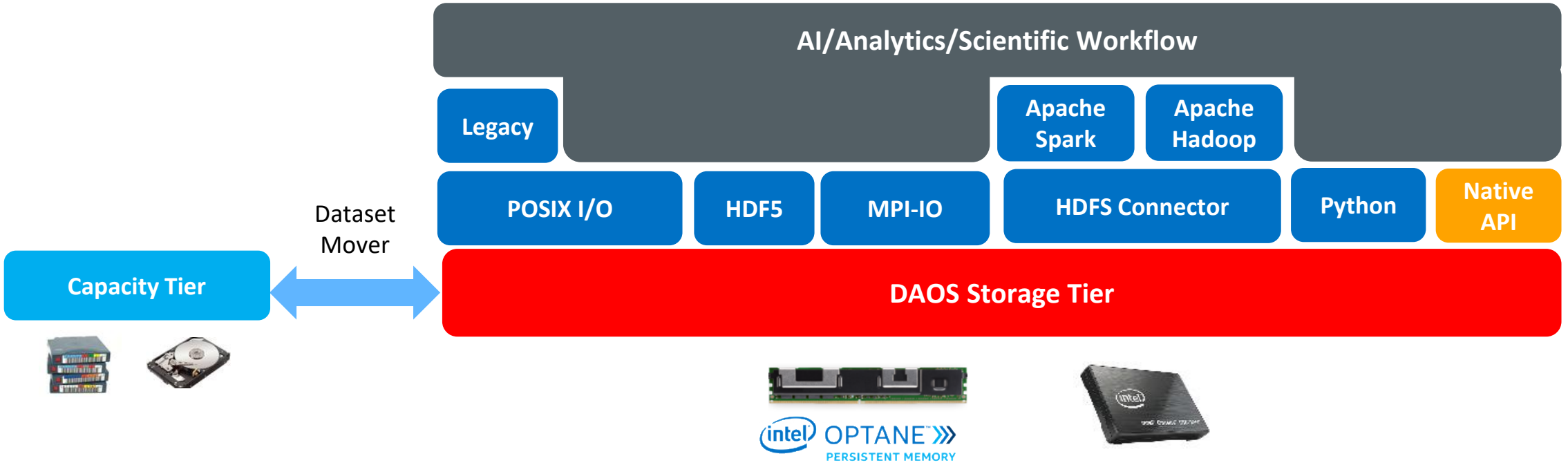
What is DAOS?

- A new, innovative scale-out storage-as-a-service stack based on Intel Optane Persistent Memory and NVMe SSDs
- Globally accessible from many nodes
- Delivers exceptionally high bandwidth and IOPS on commodity servers
- Can be utilized either as a standalone file system, or as a performance tier integrated with existing storage systems

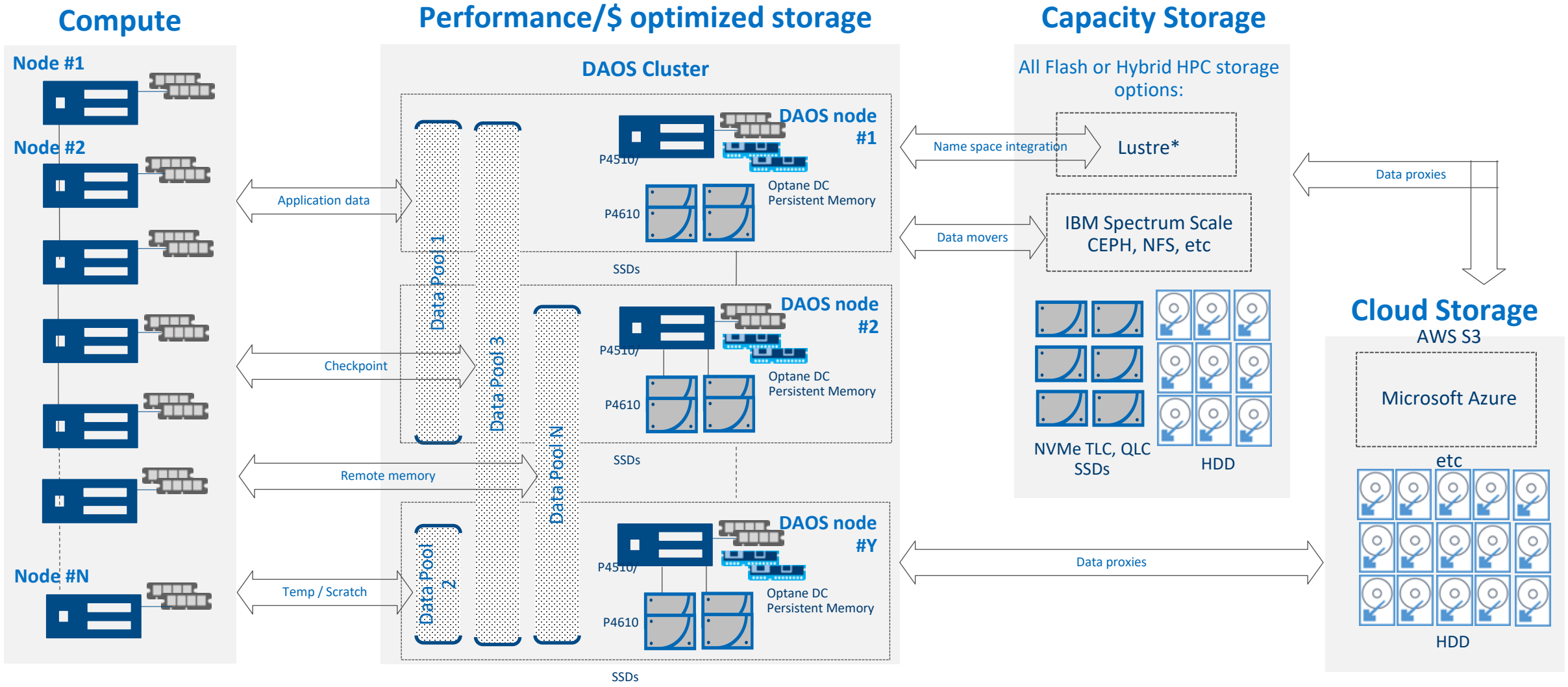


More IOPs and bandwidth per dollar

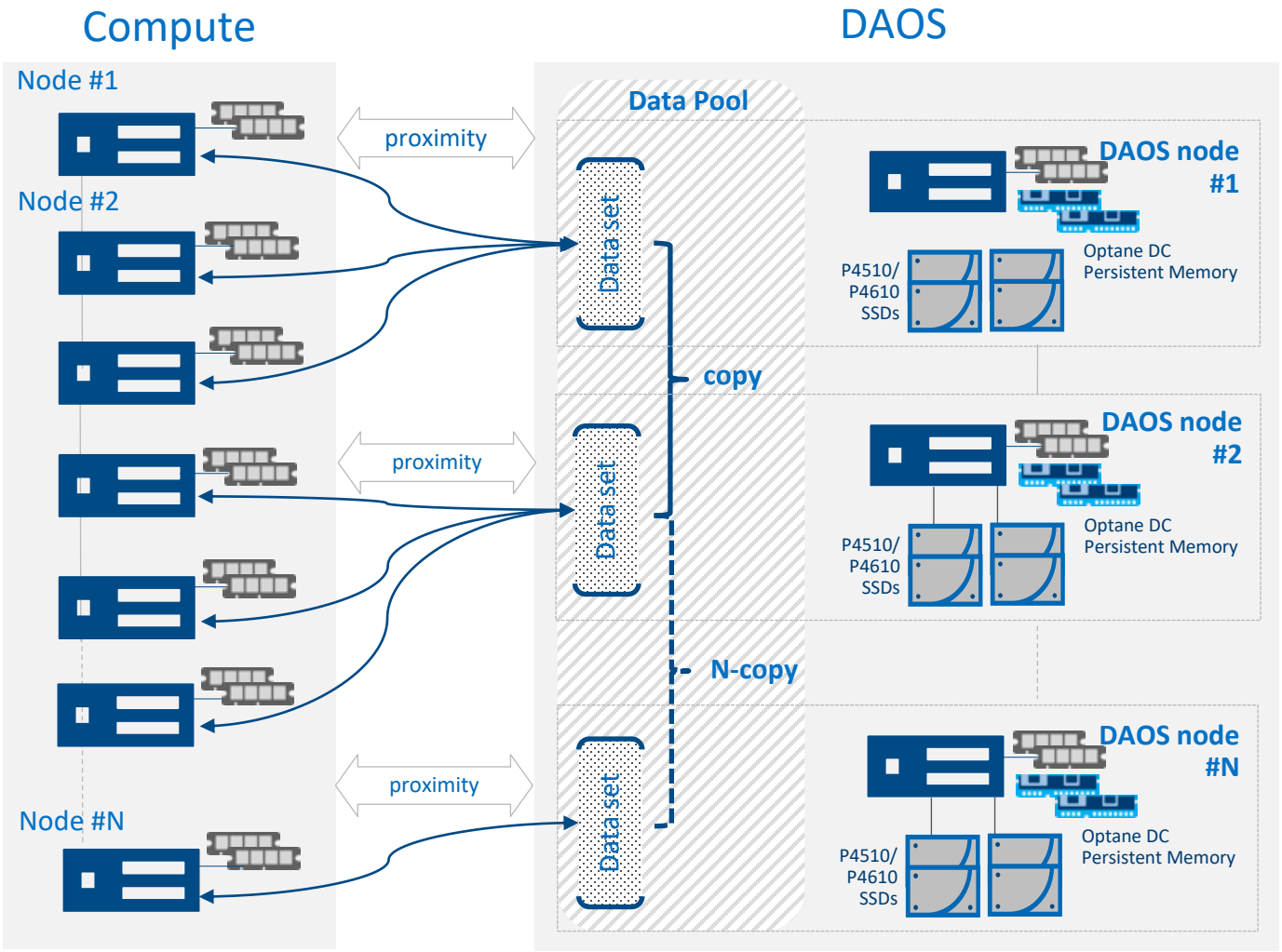
Application Interface



DAOS in the overall cluster architecture



DAOS and AI Best Play Together



AI training workloads dominate more reads than writes

Parallel throughput becomes critical at training stage for a shared data set

Traditional parallel filesystems are not able to keep up with performance demand

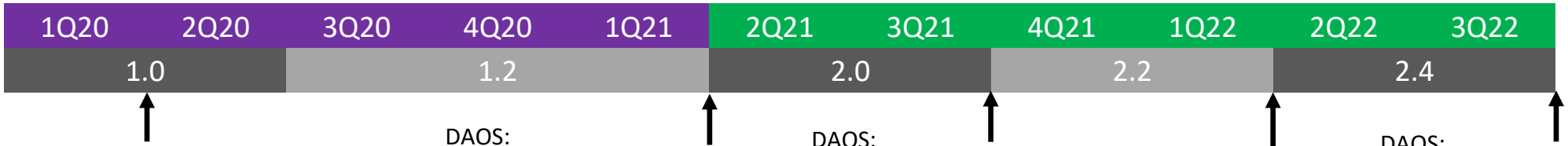
DAOS has a native capability to set the **replication factor per object**

This can be used to bring higher bandwidth for the given object and **increase parallelism and proximity** to the data.

- In addition to replication, DAOS has optimized object placement planned in the roadmap for version 2.2.

DAOS further accelerates All Flash Arrays already been used with GPU accelerated compute nodes.

DAOS Community Roadmap



Released on June 18

DAOS:
- NVMe & DCPMM support
- Per-pool ACL
- UNS in DAOS via dfuse
- Replication & self-healing (Preview)

Application Interface:
- MPI-IO Driver
- HDF5 DAOS Connector
- Basic POSIX I/O support

DAOS:
- End-to-end data integrity
- Per-container ACL
- Improved control plane
- Replication & self-healing
- Conditional updates
- Erasure Code (Preview)
- Online server addition
- Lustre/UNS integration

Application Interface:
- HDF5 vol w/o async
- POSIX I/O with conditional update support
- POSIX data mover
- Async HDF5 operations over DAOS
- Spark

DAOS:
- Erasure code
- Telemetry& per-job statistics
- Advanced control plane
- Distributed transactions
- Multi OFI provider support

Application Interface:
- POSIX I/O with distributed transaction support
- HDF5 data mover
- Container parking/serialization

DAOS:
- Progressive layout
- Placement optimizations
- Checksum scrubbing

DAOS:
- Catastrophic recovery tools

NOTE: All information provided in this roadmap is subject to change without notice.

Resources

- Github
  - https://github.com/daos-stack/daos
- DAOS online documentation
  - http://daos.io
- Community mailing list
  - https://daos.groups.io
- Annual DAOS User Group in November
  - https://wiki.hpdd.intel.com/display/DC/DUG20