New storage Paradigm

Johann Lombardi, Principal Engineer
Extreme Storage Architecture & Development, Intel

notices and disclaimers

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration.

No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. For more complete information about performance and benchmark results, visit http://www.intel.com/benchmarks .

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 complete information visit http://www.intel.com/benchmarks .

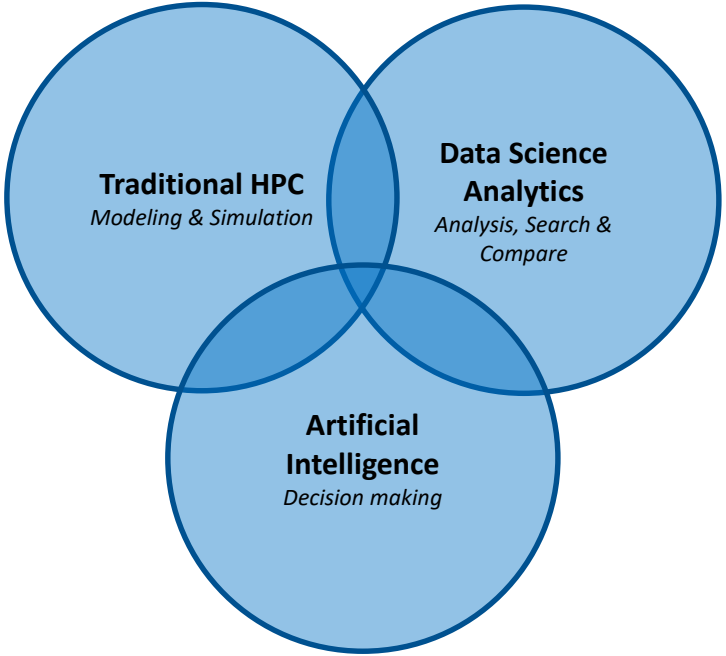
Intel® Advanced Vector Extensions (Intel® AVX)* provides higher throughput to certain processor operations. Due to varying processor power characteristics, utilizing AVX instructions may cause a) some parts to operate at less than the rated frequency and b) some parts with Intel® Turbo Boost Technology 2.0 to not achieve any or maximum turbo frequencies. Performance varies depending on hardware, software, and system configuration and you can learn more at http://www.intel.com/go/turbo.

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.
Intel, the Intel logo, and Intel Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.
*Other names and brands may be claimed as property of others.

Evolving Storage Needs



REARCHITECTING THE MEMORY/STORAGE HIERARCHY

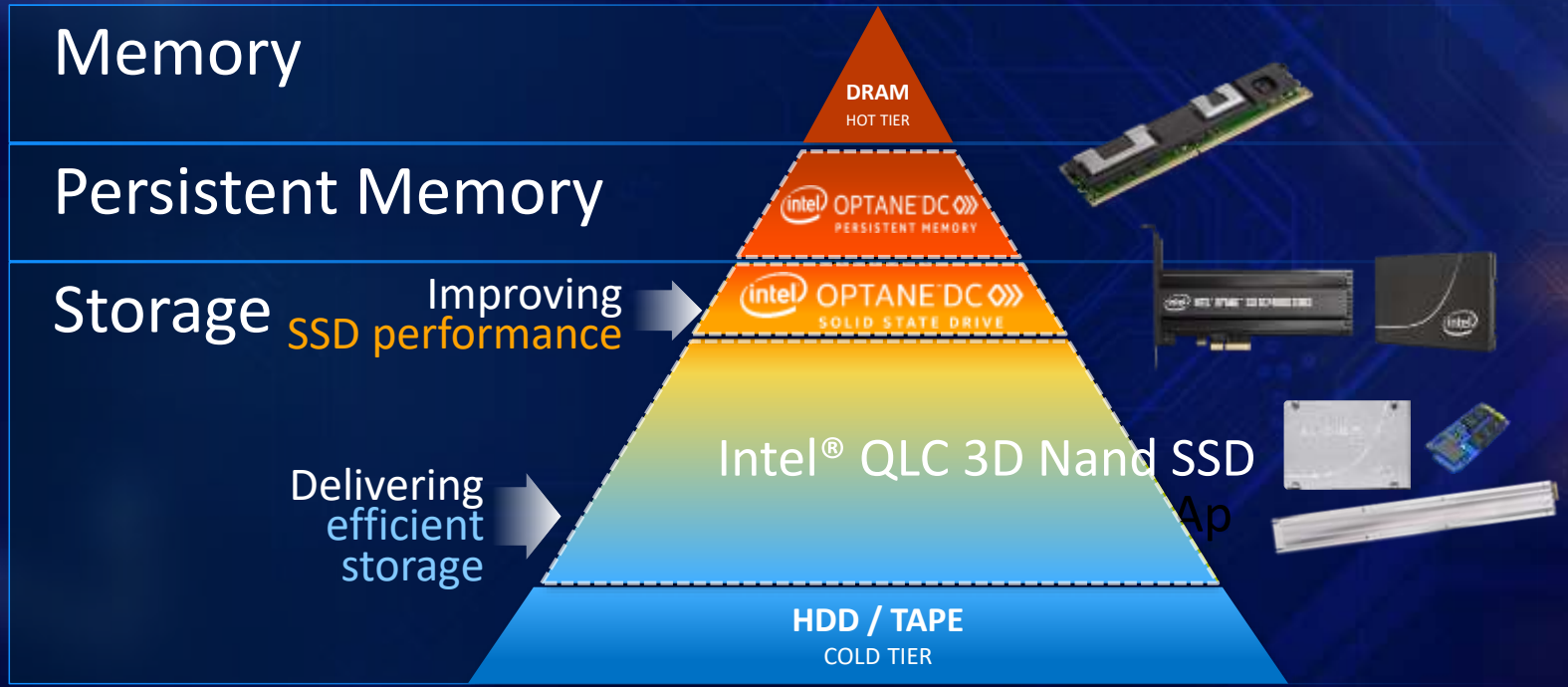
Memory

Persistent Memory

Storage

Improving SSD performance

Delivering efficient storage



End-to-End Performance Hardware

High msg rate & bandwidth
HW: Intel® Omni-Path
Ethernet

High IOPS & bandwidth
HW: Intel® Optane DC Persistent Memory
Intel® Optane DC SSDs
Intel® 3D-NAND SSDs

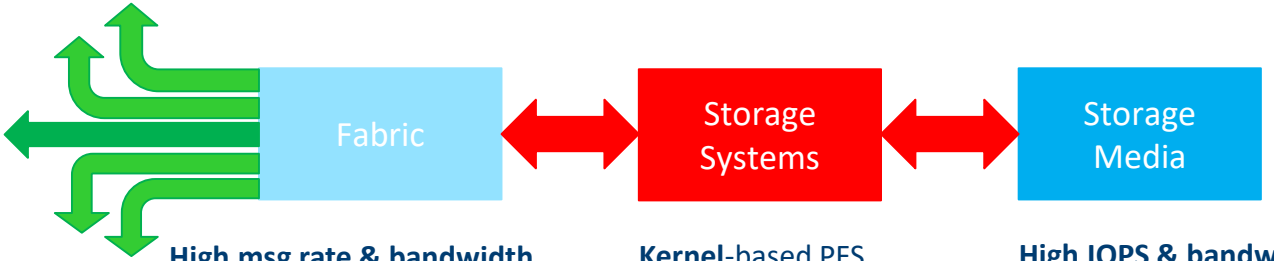
End-to-End Performance Software Building Blocks

High msg rate & bandwidth
HW: Intel® Omni-Path
Ethernet

SW: libfabric, Mercury, …
in **userspace**

High IOPS & bandwidth
HW: Intel® Optane DC Persistent Memory
Intel® Optane DC SSDs
Intel® 3D-NAND SSDs
SW: PMDK, SPDK, …
in **userspace**

End-to-End Performance
Storage Software Stack



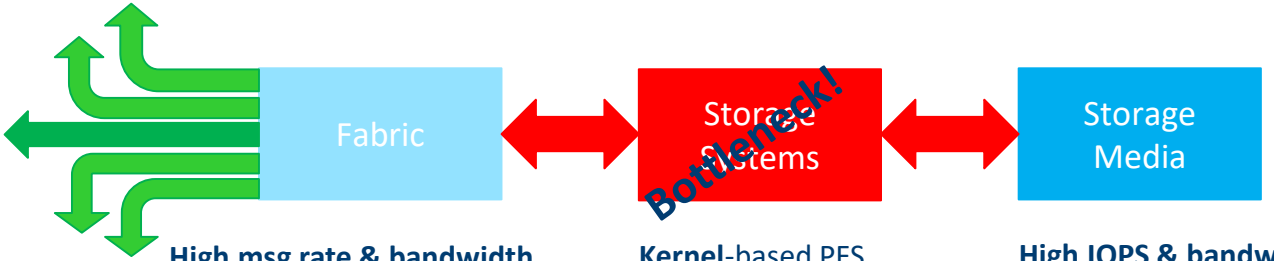
High msg rate & bandwidth
HW: Intel® Omni-Path
Ethernet

SW: libfabric, Mercury, …
in **userspace**

Kernel-based PFS,
disk-based fs, …

High IOPS & bandwidth
HW: Intel® Optane DC Persistent Memory
Intel® Optane DC SSDs
Intel® 3D-NAND SSDs
SW: PMDK, SPDK, …
in **userspace**

End-to-End Performance
Software Bottleneck



High msg rate & bandwidth
HW: Intel® Omni-Path
Ethernet

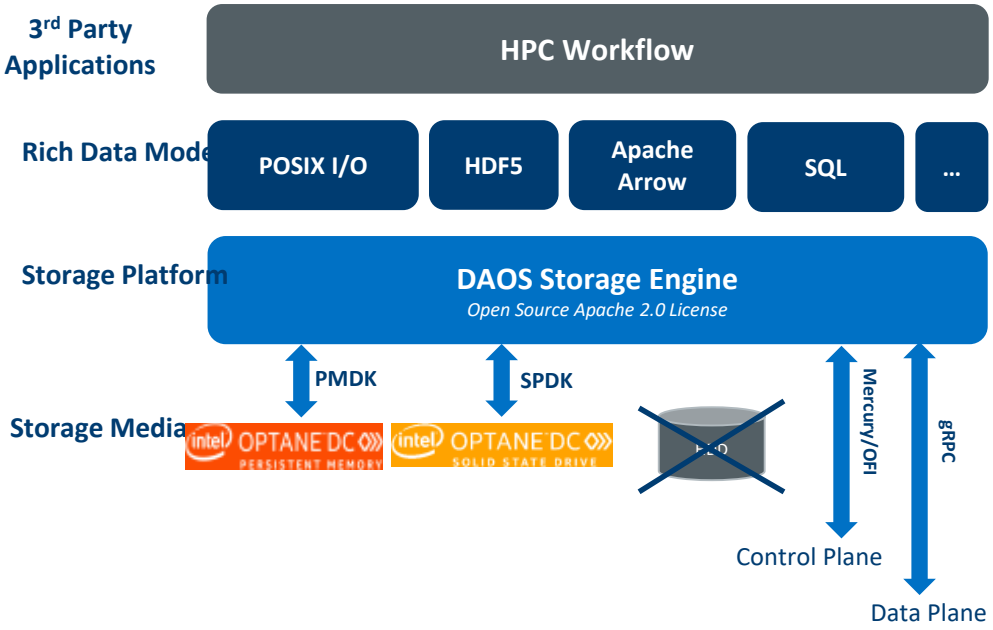
SW: libfabric, Mercury, …
in **userspace**

Kernel-based PFS,
disk-based fs, …

High IOPS & bandwidth
HW: Intel® Optane DC Persistent Memory
Intel® Optane DC SSDs
Intel® 3D-NAND SSDs
SW: PMDK, SPDK, …
in **userspace**

Traditional storage stack entirely **masks**
low latency & capabilities of **next-gen storage technologies!**

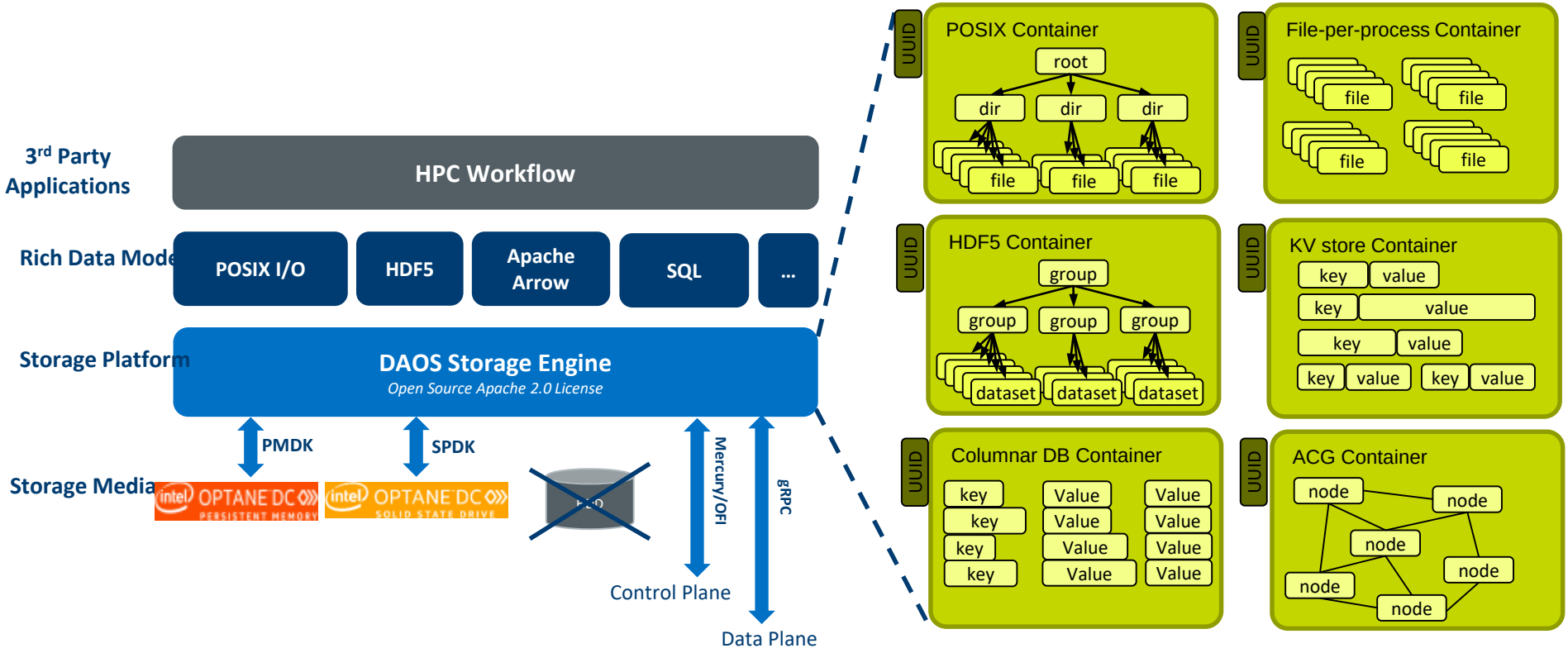
Distributed Async Object Storage



Benefits

- Built natively over **new userspace** PMEM/NVMe software stack
- **Rich** storage semantics
- High **throughput/IOPS @arbitrary** alignment/size
- **Ultra-fine grained** I/O
- **Scalable** communications & I/Os
- **Software**-managed **redundancy**
- **Microservices** architecture
- **Open source**

Distributed Async Object Storage



Lightweight I/O Stack

Mercury userspace function shipping

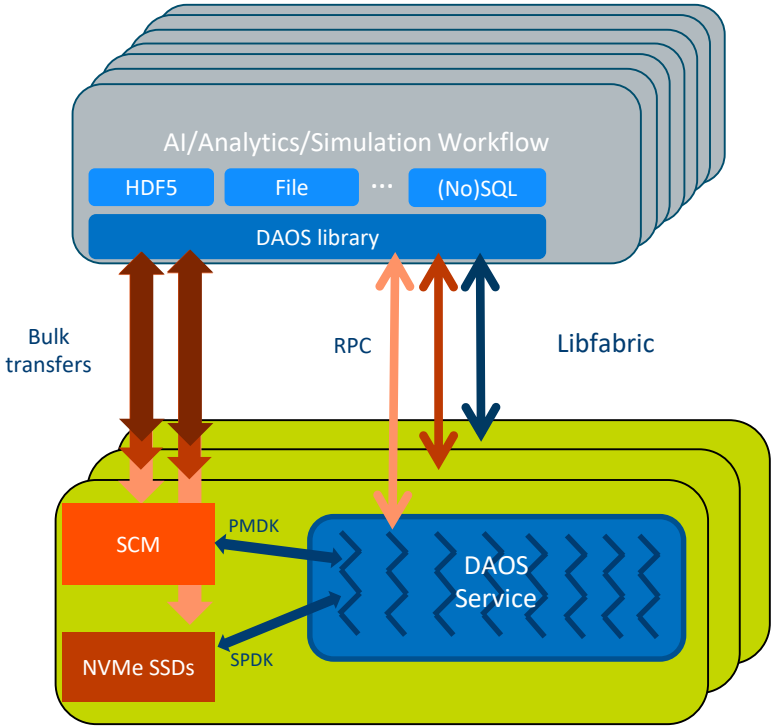
- **MPI** equivalent communications **latency**
- Built over libfabric

Applications link directly with DAOS lib

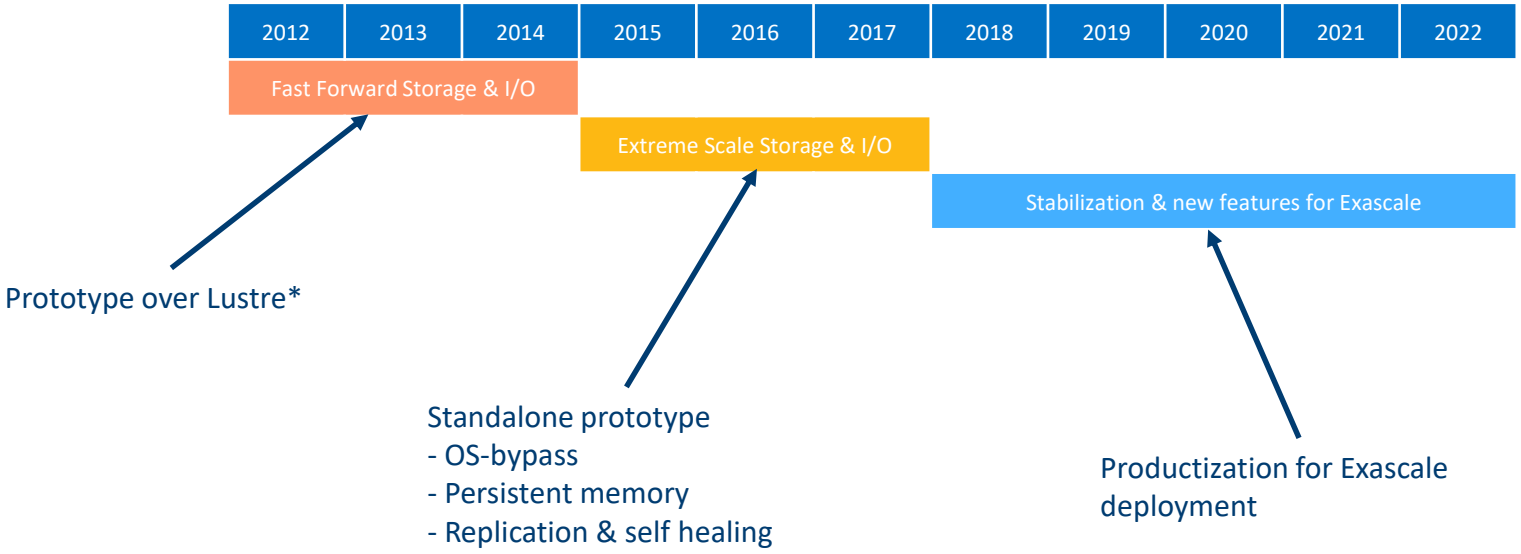
- Direct call, no context switch
- **Small** memory **footprint**
- No locking, caching or data copy

Userspace DAOS server

- Mmap non-volatile memory via PMDK
- NVMe access through SPDK/Blobstore



DAOS Development & Exascale HPC



*Other names and brands may be claimed as the property of others.

Resources

Source code on GitHub

- https://github.com/daos-stack/daos

Community mailing list on Groups.io

- daos@daos.groups.io
- https://daos.groups.io/g/daos

Wiki

- http://daos.io

Contact

- johann.lombardi@intel.com


intel
experience
what's inside™