intel®

Restoring the Balance Between Bandwidth and Latency

Bandwidth increases have historically greatly outpaced latency reductions. Until now.

Frank T. Hady, Ph.D.
Intel Fellow
Chief Optane Systems Architect

Intel Non-Volatile Memory
Solutions Group

With the ever-increasing size of datasets, data center workloads demand increasing levels of performance and capacity from both memory and storage. At the same time that more data must be processed per unit time, the components that make up the computing system are increasing in performance. Performance is a multifaceted topic where some measures (ex. bandwidth) increase at a greater rate than others (ex. latency).

The computer architect must navigate the intersection of these increasing data sets, and the relative performance increases of available technologies, to create a computing system that completes the job quickly. This brief explores the historical march of relevant technologies and the latest addition, Intel® Optane technology. This new technology delivers a needed new resource with latency and bandwidth that fill a traditional sweet-spot within the computing system to speed applications.

Memory and Storage – A (Very) Brief History

DRAM is a very high bandwidth low latency data store, but relatively expensive per bit. Increases in dataset size can be addressed by increasing the amount of DRAM in the system, but that is prohibitively expensive. Ten years ago when the only other available data store in many systems was a slow hard disk drive (HDD), there was often little choice. Accesses to high latency HDDs simply wasted too many processor cycles to wait for data.

The arrival of NAND solid state drives (SSDs) offered another place to store data, and speed access to more of the dataset. As a result, NAND-based SSDs have been widely adopted in the market. Now, even fast NAND SSDs are no longer adequate for today's data-driven applications that need to access and process data in real time or near real time. That's because like the HDDs of 10 years ago, these SSDs require the processor to wait too long for data, adding latency that can hold back systems from achieving the performance levels that modern CPUs are capable of delivering. As CPU performance has increased over time, storage latency has not kept pace, becoming a drag on overall system performance gains.

Today, more than 10 years since NAND SSDs were introduced, Intel® Optane™ SSDs significantly decrease latency for storage, down 10 microseconds, again offering a storage resource at an historical balance between bandwidth and latency and allowing the processor to get to more of its data set faster.
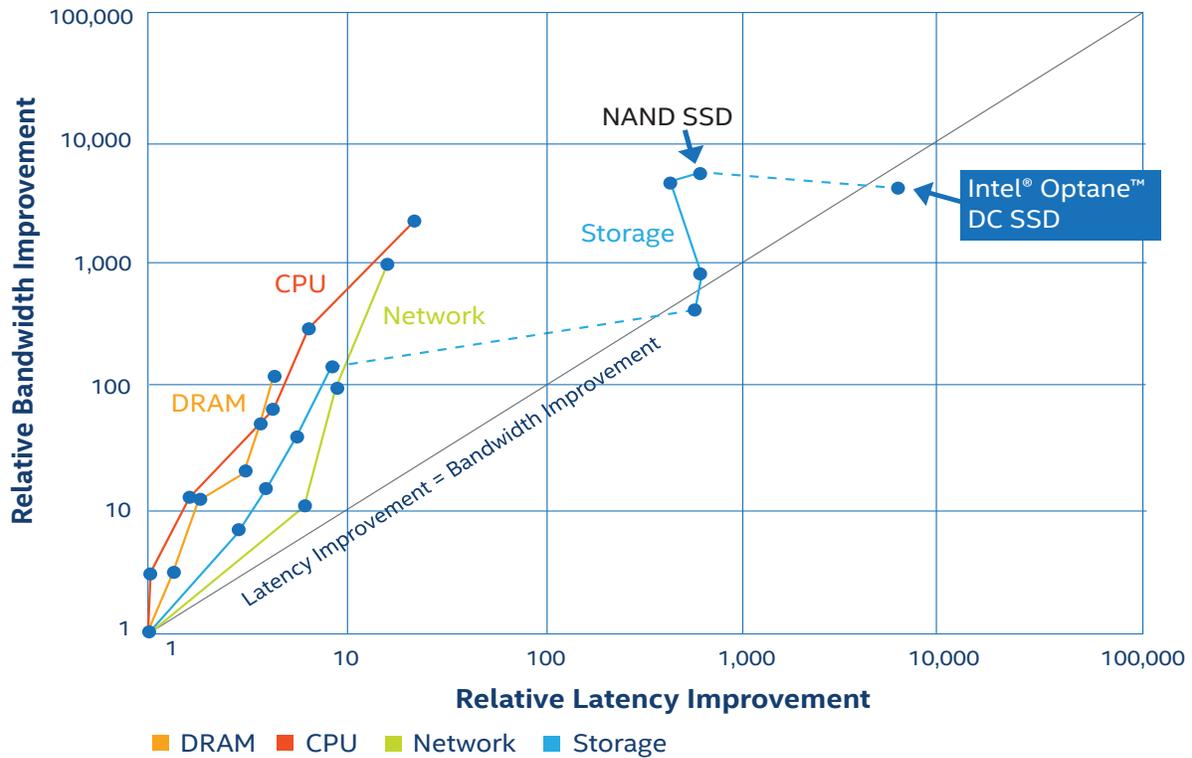
Memory and Storage Technical Series

The Direct Connection to Intel Fellows and Principal Engineers

This paper is part of a series designed to help system architects, engineers, and IT administrators understand the technological limitations of traditional memory and storage, how those limitations have led to performance and capacity gaps in the data center, and how Intel® Optane™ technology helps fill those gaps with a new industry-disrupting architecture.

The Memory and Storage series examines several topics that affect memory and storage performance and capacity, including bandwidth, latency, queue depth, quality of service (QoS), and reliability.

Figure 1. Relative bandwidth improvement versus relative latency improvement over time for memory, processors, HDDs, and SSDs.[1]

Maintaining Latency and Bandwidth Balance as Technologies March Forward

To illustrate the march of technologies it is useful to compare relative bandwidth performance versus latency improvement over time for various storage media. Building on a key study by David Patterson, Figure 1 adds SSD data points to Patterson's "latency lags bandwidth" chart.[1] Patterson showed that bandwidth has historically improved at a much faster rate than latency. Transistors have steadily increased in number according to Moore's law,[2] while multicore architectures have continued to evolve.

Those improvements have allowed processors to process more instructions, and therefore more data, in the same or less time than previous generation processors. But as CPU processing times have dropped, the time to get data from HDDs —the drive latency—did not drop correspondingly. That led to storage technology becoming the bottleneck in overall performance. For memory and storage technologies, bandwidth can be increased through parallelism but the time to access the technology is relatively constant. Only the introduction of a new technology delivers a lower latency.

To understand why this matters, consider what happens when latency decreases and bandwidth increases. In general, for memory and storage resources, a single unit data access is not enough to fill the pipe from memory to the processor. Put differently, bandwidth multiplied by the latency (the bandwidth

delay product) is larger than the access size. When possible, to use the full bandwidth of the resource, software is explicitly written to ask for bigger or more chunks of data in parallel. As the bandwidth delay product grows, fewer and fewer algorithms are able to ask for enough data in parallel to cover the latency. In cases where they cannot, system bandwidth and performance suffer. At the simplest level, this is why having a balanced bandwidth/latency ratio matters.

Referring back to Figure 1, the introduction of NAND-based SSDs provided a balanced bandwidth/latency solution for a time, and they brought much lower latencies than HDDs. Base access times dropped from milliseconds (ms) for HDDs, to less than 100 microseconds (μs) for NAND SSDs, meaning fewer CPU cycles spent waiting for data. With many applications able to often access the full bandwidth of the NAND SSDs, this sped up processing in ways that users noticed. Over time, bandwidth continued to increase while latency remained relatively constant, stranding bandwidth and again, putting the system out of balance.

The following example demonstrates how Intel Optane technology—deployed as low latency Intel® Optane™ DC SSDs—can increase performance and capacity for hyperconverged infrastructure solutions, like VMware vSAN.

Intel® Optane™ Technology Pushes vSAN Performance and Capacity to New Levels

Enterprise businesses and cloud service providers can use Intel Optane technology to affordably improve performance for applications running on virtual servers. An analysis performed by Evaluator Group found that Intel® Xeon® Scalable processors, combined with Intel Optane technology and Intel® 3D NAND SSDs with NVM Express (NVMe), can deliver better performance for a variety of common workloads running on hyper-converged systems using VMware vSAN.[3]

As shown in Figure 2, systems running VMware vSAN 6.7 that are built with Intel Xeon Scalable processors and Intel Optane DC SSDs can provide significant performance improvements compared to systems running with NAND SSD storage media. The systems built with Intel Optane technology and Intel 3D NAND SSDs support up to 1.6x more virtual machines (VMs) while still maintaining the same service level agreement for each VM.[3]

This is equivalent to supporting 60% more users per system, important to the bottom line and business growth. The result is a clear cost benefit driven by increased VM density and lower cost of infrastructure provided by Intel Xeon Scalable processors, VMware vSAN 6.7, and the combined use of efficient Intel 3D NAND SSDs with Intel Optane DC SSDs.[3]

The study concluded that performance was slower on older systems because the older storage technologies could not keep up with the input/output (I/O) demands of VMs. Essentially, the intense I/O workloads driven by multiple active VMs caused the NAND SSDs to back up with outstanding work, increasing latency to data until the required service level agreement required by the VMs was no longer maintained.

This VMware vSAN example shows one way you can deploy Intel Optane DC SSDs to fill gaps in the data center memory and storage hierarchy. Check the Intel Optane technology web site often for new examples of how businesses are using Intel technology to better meet the demanding needs of the modern data center.

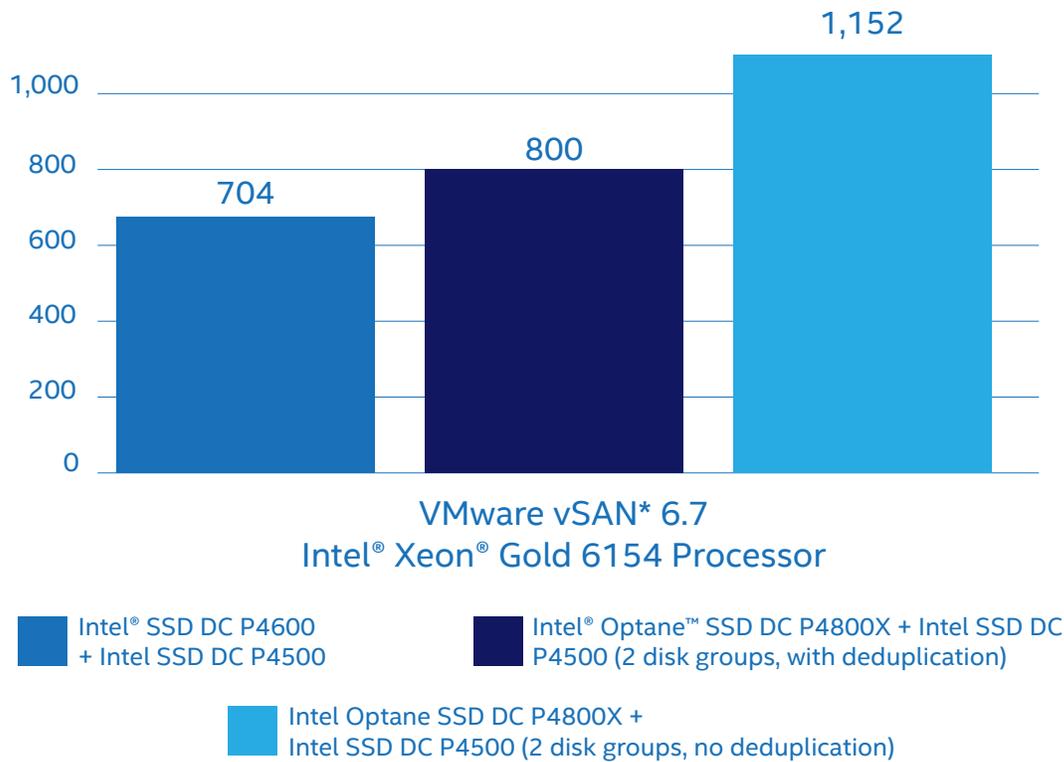## IOMARK-VM* PERFORMANCE: 4-NODE HYPER-CONVERGED CLUSTER

Figure 2. Newer VMware vSAN systems, built with Intel® Xeon® Scalable processors, Intel® 3D NAND SSDs, and Intel® Optane™ DC SSDs, offer up to 1.6x higher performance than systems built on Intel NAND SSDs alone.[3]

A New Architecture for Memory and Storage

Intel® Optane™ technology can be deployed in a variety of roles within the system. An Intel® Optane™ DC SSD can connect to systems using a standard PCIe NVMe interface with a bandwidth/latency balance to speed an important data center application, as shown in the previous example. In this form, idle average latency is about 10 microseconds, compared to more than 80 microseconds for NAND SSDs.[4] Figure 3 shows both system hardware and software latency. Intel Optane DC SSDs feature hardware latency roughly equal to the system-stack software latency, bringing another kind of balance to the system. Consistently low latency, even under heavy load, along with high endurance, makes these SSDs ideal for fast caching or tiering of hot data.

Intel Optane technology is now also available as Intel® Optane™ DC persistent memory modules that plug directly into DIMM slots. Unlike DRAM DIMMs, Intel Optane DC persistent memory offers persistence and larger memory capacity (up to 512 GB per module). As Figure 3 shows, latency for data access with Intel Optane DC persistent memory is much smaller than even Intel Optane DC SSDs.

Intel Optane DC persistent memory can be accessed directly from applications without involving the operating system storage stack, so the software overhead is removed. With persistent memory, idle average read latency drops to between 100 and 340 nanoseconds (ns).[5] Consider this low latency in terms of the bandwidth-delay product mentioned earlier. Because latency is low, this memory can be accessed with a small unit size, a single cache line, and still provide its full bandwidth. Intel Optane DC persistent memory is therefore a cache line-accessible, high performance, persistent store—a truly unique new resource.

Because of its high-performance and persistence, Intel Optane DC persistent memory forms another new data storage layer that can be used in a variety of ways to fill system gaps in capacity and performance. That flexibility allows businesses to architect data centers that can better meet the processing and memory needs of modern workloads. For example, Intel Optane DC persistent memory can be used to significantly increase capacity for in-memory databases. And because persistent memory is non-volatile, data does not need to be reloaded into memory after a database restart, which increases serviceability and system uptimes, and improves business continuity.
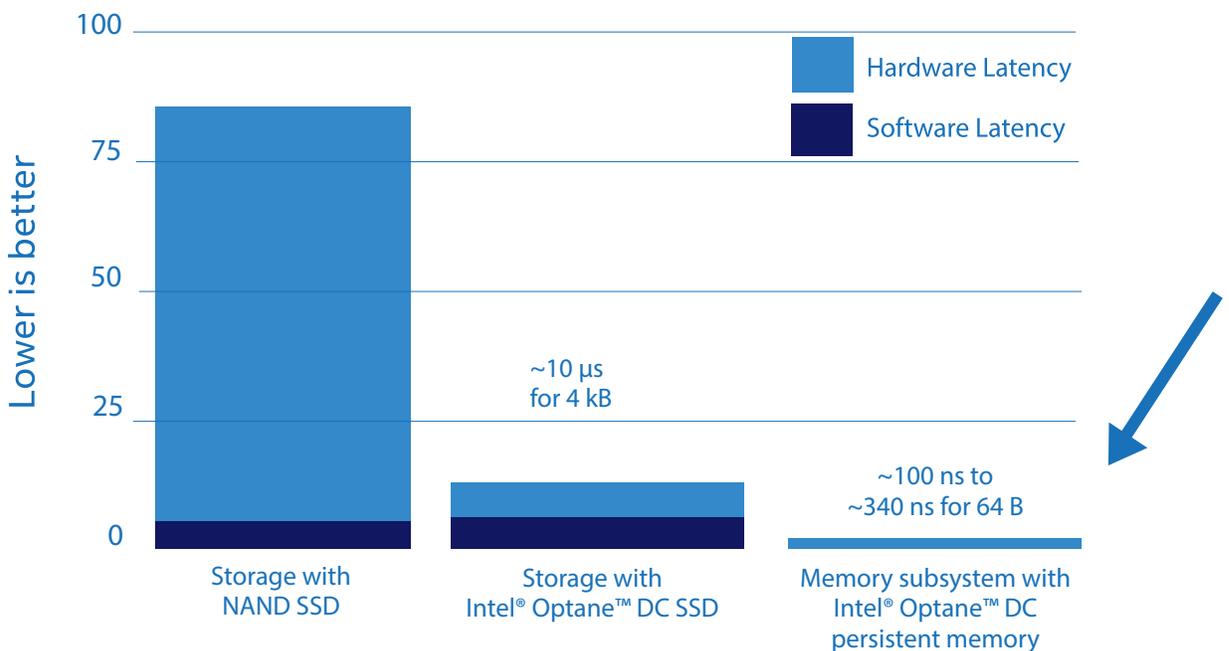
IDLE AVERAGE RANDOM READ LATENCY

Figure 3. Comparison of latency for NAND SSDs, Intel® Optane™ DC SSDs, and Intel Optane DC persistent memory.[4, 5]

Conclusion

In computing systems, the memory and storage hierarchy places data that is more frequently accessed closer to the processor while the preponderance of data is moved to less expensive memory further (higher latency) away from the processor. The inherent latency of memory and storage technologies tends to drop slowly over time, while processors increase in performance at a much faster rate. This effectively moves these memories farther from the processor—as a result, the processor has to waste more instruction cycles waiting for data. Only the introduction of new lower latency memory technologies and new, more tightly integrated system integration points brings the system back into balance.

With the introduction of Intel Optane technology, Intel has delivered a new memory into the system to fill the gap between DRAM and NAND SSDs. As both SSD and persistent memory, the new Intel Optane technology enables computer architects to keep large persistent data structures closer to the processor, minimizing the wait time for data and speeding application execution. When system architects balance bandwidth demand with low latency, it unleashes the power of the CPU. With the balance between bandwidth and latency restored by Intel Optane technology, the CPU can now consume and process data quickly to achieve optimal system performance.

Intel Fellow Frank Hady

Frank Hady is an Intel Fellow and the Chief Optane Systems Architect in Intel's Non-Volatile Memory Solutions Group (NSG). Frank leads research and definition of Intel® Optane™ technology products and their integration into the computing system. Frank has served as Intel's lead platform I/O architect, delivered research foundational to Intel® QuickAssist Technology, and driven significant platform performance advances. He has authored or co-authored more than 30 published papers on topics related to networking, storage, and I/O innovation and presents often on memory and storage. He holds more than 30 U.S. patents. Frank received his bachelor's and master's degrees in electrical engineering from the University of Virginia, and his Ph.D. in electrical engineering from the University of Maryland.

Learn More

Low, consistent latency is just part of the Intel Optane technology story. Learn more about how Intel Optane DC persistent memory and Intel Optane DC SSDs are disrupting the memory and storage hierarchy in the data center by exploring other papers in the Memory and Storage Technical Series.

For more information on Intel Optane technology, visit **intel.com/optane** and read the technology brief "What is Intel Optane Technology?" at **http://intel.com/content/www/us/en/products/docs/memory-storage/optane-technology/what-is-optane-technology-brief.html**.

For more information on Intel Optane DC SSDs, visit **intel.com/content/www/us/en/products/memory-storage/solid-state-drives/data-center-ssds/optane-dc-p4800x-series.html**.

To learn more about Intel Optane DC persistent memory, visit **intel.com/content/www/us/en/architecture-and-technology/optane-dc-persistent-memory.html**.

1 David A. Patterson. "Latency Lags Bandwith." Communications of the ACM, Vol. 47 No. 10, October 2004. https://dl.acm.org/citation.cfm?id=1022596. Intel® 3D NAND and Intel® Optane™ DC SSD data points added by Intel based on product brief specifications for Intel® SSD X-25M, and Intel SSD DC S3700, Intel SSD DC P3700, Intel SSD DC P4600, and Intel Optane SSD DC P4800X. For more information, see: Intel. "Intel® SSD Data Center Family." intel.com/content/www/us/en/products/memory-storage/solid-state-drives/data-center-ssds.html.

2. Source - Intel.com. Moore's Law. In 1965, Gordon Moore made a prediction that would set the pace for our modern digital revolution. From careful observation of an emerging trend, Moore extrapolated that computing would dramatically increase in power, and decrease in relative cost, at an exponential pace. The insight, known as Moore's Law, became the golden rule for the electronics industry, and a springboard for innovation. As a co-founder, Gordon paved the path for Intel to make the ever faster, smaller, more affordable transistors that drive our modern tools and toys.

3. Evaluator Group. "Lab Insight: Latest Intel Technologies Power New Performance Levels on VMware vSAN – 2018 Update." October 2018. https://www.evaluatorgroup.com/document/lab-insight-latest-intel-technologies-power-new-performance-levels-vmware-vsan-2018-update/.

4. Source: Intel-tested: Average read latency measured at queue depth 1 during 4k random write workload. Measured using FIO 3.1. comparing Intel Reference platform with Optane™ SSD DC P4800X 375GB and Intel® SSD DC P4600 1.6TB compared to SSDs commercially available as of July 1, 2018. Performance results are based on testing as of July 24, 2018 and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.

5. Intel® Optane™ DC persistent memory: results based on Intel testing on February 20, 2019. Configuration: Intel® C620 Series chipset, 28-core Intel® Xeon® Scalable processor (QDF QQYZ), 2,666 megatransfers per second (MT/s), 256 GB, 18 W, 32 GB DDR4 DRAM (per socket), 128 GB Intel® Optane™ DC persistent memory (per socket), firmware: 5336, BIOS: 573.D10, WW08 BKC, running Linux OS4.20.4-200.fc29. Performance tuning quality of service (QoS) disabled, IODC=5(AD).