



Lower TCO Through Expanding Memory Capacity in Virtualized Environments

Save time and money by achieving greater per-socket memory capacity, high data center efficiency, and lower TCO across a range of applications by expanding memory capacity

You might need Intel® Optane™ DC persistent memory if you have...

- App owners allocating too much memory to their VMs
- Challenges reclaiming memory from app owners
- Low active memory usage and low CPU usage
- Memory resources running out before CPU
- The desire to land larger VMs
- The need to increase VM density while saving on hardware costs

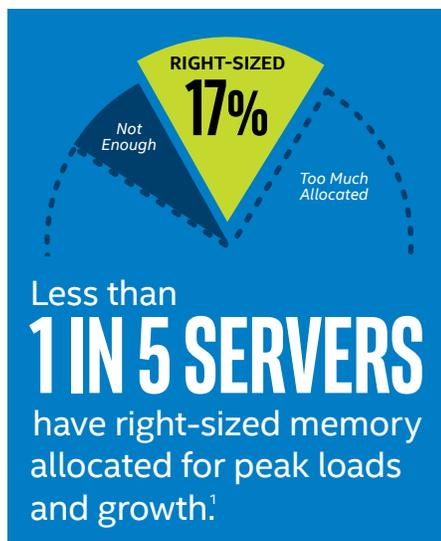
Industry Strategic Challenges

The rising tide of worldwide data growth, fueled by trends such as high-definition video surveillance, the IoT, AI, and 5G connectivity, dramatically increases workload sizes. This data inflation places a heavy emphasis on memory expectations. IT knows that the virtual machines (VMs) tasked with handling these workloads will experience periodic fluctuations and occasional peaks. Having insufficient memory will yield disk swaps that lead to performance bottlenecks, so IT often pursues a strategy of expanding memory capacity for peak loads, even though true utilization rarely reaches those levels.

However, provisioning more memory per VM than needed reduces the total number of VMs possible on a given server. This impairs total resource utilization (due to lower VM density) and infrastructure ROI. Unfortunately, IT staff often don't get to see the larger economic picture. They're doing their jobs, working to make sure that applications perform quickly and reliably. But they, and even executive management, may not grasp the true costs of provisioning more VM memory with DRAM than strictly needed.

A 2018 report from IT management solution provider ControlUp details a survey including 71,845 server VMs spanning over 900 organizations worldwide. The survey found that 77 percent of servers were over-provisioned. Moreover, across all servers examined, any given server was over-provisioned by an average of 55 percent.¹ All told, ControlUp pegged the amount overspent on this unneeded DRAM at USD 23.8 million. Additionally, overcommitment of DRAM leaves less memory available for adding more VMs, resulting in lower CPU utilization per physical server, and thus higher TCO.

Organizations need a way to achieve higher server memory capacities—and the high VM performance that can come from having right-sized memory allocation—without the potentially steep per-gigabyte costs of DRAM. IT can future-proof data centers by adding persistent memory—specifically, Intel® Optane™ DC persistent memory, which provides a path to affordably expanding system memory while maintaining the high-performance levels associated with running large in-memory workloads. Intel Optane DC persistent memory modules (DCPMMs) fit in the same DDR4 slots as conventional DRAM, making them easy to install on servers running 2nd Generation Intel® Xeon® Scalable processors while delivering capacities several times that of most DDR4 modules. Businesses can keep running DRAM for their most performance-sensitive “active”



data and deploy Intel Optane DC persistent memory to satisfy remaining memory needs, such as occasional peak loads or high seasonal cycles. By having more total system memory available, managers can make sure application owners retain all the memory they expect, which in turn will help save considerable time on handling resource requests.

With this new technology, organizations running Microsoft SQL Server can achieve up to a 36 percent increase in VM density while saving up to 25 percent in hardware costs per VM.² Alternatively, administrators could maintain the same number of VMs per server while reducing total server cost. Intel Optane technology opens multiple routes for improving efficiency and boosting TCO savings. In turn, those substantial savings can allow your budget to drive innovation and company-wide growth.

“As a society, we’re moving towards more data becoming important, and therefore more real-time analytics are needed. To help handle that, [Intel Optane DC persistent memory] is a very important milestone, because it offers something that you cannot do with flash, and it offers a cost advantage, as well. This is a beautiful technology that brings a fantastic set of characteristics to the table.”

— Michael Yang
Analyst, Informa Tech (previously IHS Markit)

Business Drivers and Desired Outcomes

A range of factors can make businesses feel the pinch of insufficient server memory. Infrastructures need modernization. Applications must deliver faster times to insight. The rising tide of data everywhere yields rapidly swelling workloads. These factors, because of their resource demands, can negatively impact TCO. Thus, there is constant pressure to lower costs and improve key performance metrics, including resource utilization. Intel Optane DC persistent memory is available only with 2nd Generation Intel® Xeon® Scalable processors and can offer a cost-effective path to an array of benefits:

- Run more VMs per server for optimized resource and rack space utilization
- Accommodate massive workloads, enabling new breakthroughs and business insights
- Align compute resources with forward-looking technologies, such as AI and analytics
- Reduce IT costs and platform TCO
- Spend less time negotiating with application owners about how much memory they really need
- Improved ability to meet quality of service (QoS) levels
- Future-proof virtualized infrastructure against scaling workloads and virtualization demands

Intel® Optane™ DC Persistent Memory Modes

Intel Optane DCPMMs can be configured in two principle modes: Memory Mode and App Direct Mode.

Memory Mode provides a single volatile memory pool in which the DRAM installed in the system acts as a cache to deliver DRAM-like performance for DCPMM as the high-capacity main memory. This mode can help drive down IT costs, increase VM density, and improve platform scalability.

App Direct Mode makes both DRAM and DCPMMs visible to the host. However, applications enabled for App Direct Mode can decide where to place data within these two memory types, which can yield higher performance. Additionally, because DCPMMs can be non-volatile in App Direct Mode, data may remain persistent and provide much faster restart/reload times. App Direct Mode can deliver very large system memory capacities, improve application availability, and drive down data center server footprint, which can lower power consumption, infrastructure costs, and TCO.

Digital Transformation and Business Innovation

Affordably expanding memory capacity can meaningfully improve a data center’s efficiency. Higher resource utilization can lead to fewer physical servers, lower power consumption, a smaller infrastructure footprint, and lower licensing costs. Concurrently, being able to handle larger working data sets helps organizations delve into applications and workloads that might have been previously impossible. In today’s world, legacy architectures are no longer competitive. Not surprisingly, research from the Enterprise Strategy Group found that companies with newer infrastructures experience 2.5x improved competitiveness, 16x increased innovation, and 18x faster time to market.³ Intel Optane DC persistent memory can provide the capability organizations need to improve their agility and reduce infrastructure TCO while simultaneously realizing fresh value from enabling new business prospects in information processing.

VM Memory Expansions are a Good Fit...



LOW CPU UTILIZATIONS

When CPU utilizations are low, use untapped server potential



HIGH RATIOS

When read versus write ratios are high



COST MOTIVATION

When cost for performance is a primary motivation

Intel Optane Technology Foundation

The development of Intel Optane DC persistent memory technology, with its combination of low latency, high endurance, high QoS, and high throughput, provided a breakthrough to extending the value and functionality of memory and storage in modern computing platforms. When combined with 2nd Gen Intel Xeon Scalable processors, Intel Optane DC persistent memory can bring these attributes to supporting and enhancing a growing number of software applications from ISVs and in-house IT developers. Intel continues to work with hardware and software solution providers to ensure that Intel Optane technology can stand at the heart of tomorrow's IT transformation, enabling larger workloads, faster results, and the underpinnings of a versatile multi-cloud enterprise.

[Learn more about Intel Optane DC persistent memory.](#)

¹ ControlUp, "Resource Provisioning Report." August 2018. controlup.com/controlup-research/resource-provisioning-report

² Performance results are based on testing as of February 1, 2019 and may not reflect all publicly available security updates.

Baseline configuration (DRAM): 2x 2nd Gen Intel® Xeon® Platinum 8276 processors @ 28 cores/socket. Memory: 768 GB (24x 32 GB DDR4 @ 2666 MHz). Network: Intel® X520 SR2 (10 Gbps). Storage: 8x Samsung PM963M.2 960 GB, 4x Intel® SSDs S3600 (1.92 TB). BIOS: WW02'19. Microcode: 0x04000014. OS/VM: Windows Server 2019. WL Version: OLTP Cloud Benchmark (internal private customer confidential workload). Dataset/instance or workload size: 1.1 TB. Security mitigations: variants 1,2,3 enabled. Performance of 22 SQL VM instances.

New configuration (DCPMM): 2x 2nd Gen Intel® Xeon® Platinum 8276 processors @ 28 cores/socket. Memory: 1 TB (8x 128 GB Intel® Optane™ DC persistent memory in Memory Mode + 12x 16 GB DDR4 @ 2666 MHz). Network: Intel® X520 SR2 (10 Gbps). Storage: 8x Samsung PM963M.2 960 GB, 4x Intel® SSDs S3600 (1.92 TB). BIOS: WW02'19. Microcode: 0x04000014. OS/VM: Windows Server 2019. WL Version: OLTP Cloud Benchmark (internal private customer confidential workload). Dataset/instance or workload size: 1.5 TB. Security mitigations: variants 1,2,3 enabled. Performance of 30 SQL VM instances.

³ Enterprise Strategy Group, March 2018, "Research Proves IT Transformation's Persistent Link to Agility, Innovation, and Business Value." dell EMC.com/en-us/whitepaper/esg-it-transformation-maturity-report-agility-innovation-business-value.htm

Intel® Optane™ DC persistent memory pricing and DRAM pricing referenced in TCO calculations are provided for guidance and planning purposes only and do not constitute a final offer. Pricing guidance is subject to change and may revise up or down based on market dynamics. Please contact your OEM/distributor for actual pricing. Pricing guidance as of July 28, 2019.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps. Performance varies depending on system configuration. Check with your system manufacturer or retailer, or learn more at intel.com.

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'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.

Intel, the Intel logo, Xeon, and Optane are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

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

© Intel Corporation 1219/CWAN/KC/PDF 341247-001US

