Solution Brief

Data Center Modernization

Boost VMware Performance While Consolidating Servers and Reducing Cost

To drive maximum business advantage, enterprises must modernize both data center hardware and software together. Refreshing from vSAN 7 to vSAN 8 and from 2nd Gen to 5th Gen Intel® Xeon® Scalable processors at the same time delivers up to 22.4x higher IOPS performance[1] and up to a 12.5:1 server consolidation ratio.[2]

Authors:

Ewelina Kamyszek, Cloud Systems and Solutions Engineer in Intel Data Center and AI (DCAI)

Patryk Wolsza, Cloud Systems Architect in Intel Datacenter and AI (DCAI)

Christine McMonigal, Intel Hybrid Cloud and HCI Marketing

Four outcomes of modernization

Modernizing hardware and software infrastructure together enables consolidation of services, which can provide four significant outcomes:

- **Increased performance** highlights the value of data by enabling innovative business cases.
- **Decreased licensing fees** reduce costs by getting more work out of each compute node.
- **Smaller server footprint** saves on space, power and cooling, advancing sustainability goals.
- **Reduced infrastructure requirements** simplify operations and maximize value of data center investments.

Modernizing infrastructure helps IT organizations meet emerging business challenges with greater agility, lower cost and higher performance. In VMware environments, upgrading to the latest generation of VMware Cloud Foundation or VMware vSphere Foundation is a key path to reaching those goals. But refreshing software on its own misses the full benefit possible from modernizing hardware as well. Hyperconverged infrastructure (HCI) based on updated Intel® Xeon® processors delivers even higher performance and server consolidation ratios, while enabling more efficient hybrid and multi-cloud environments.

Advanced technologies from Intel and VMware provide operational efficiencies that streamline and reduce the cost of maintaining today's cloud environments. They improve workload efficiency by taking advantage of improvements to Intel Xeon processors, including built-in accelerators for the most common workloads running on VMware infrastructure. Higher performance and processor utilization make it possible to run more workloads per server, or to use fewer servers to run the same workloads. Making more efficient use of infrastructure reduces software license cost and capital expense, as well as operating costs for space, power and cooling.

Using outdated technologies can hinder business innovation, growth and worker productivity. This solution brief showcases the combined benefits of refreshing both hardware and software for VMware environments. These changes provide compelling business outcomes across a range of demanding workloads, helping drive greater value from data for better innovation and decision making, enhanced customer experiences and optimized total cost of ownership (TCO).

Modernized hybrid and multi-cloud solutions from VMware and Intel

Software building blocks from VMware provide the infrastructure to build high-performance hybrid and multi-cloud environments for demanding business workloads, optimized for Intel architecture. Upgrading to the latest version of VMware Cloud Foundation or VMware vSphere Foundation enables businesses to capture the full range of performance, scalability and efficiency advantages in their environments.

VMware Cloud Foundation is an enterprise-class private cloud platform. It provides full-stack infrastructure as a service, including software-defined compute, storage, networking, security and management. This hardened, self-service platform provides automation and orchestration to optimize agility, resilience, and scalability. VMware vSphere Foundation is a simplified, enterprise-grade workload platform for smaller and medium-sized customers.

Both VMware Cloud Foundation and VMware vSphere Foundation unlock the performance and efficiency capabilities of modern hardware, as they also simplify provisioning and operations.

vSAN 8 provides enterprise-class storage virtualization that helps reduce storage costs and complexity while supporting a robust path to hybrid/multi-cloud infrastructure. vSAN's optional new Express Storage Architecture (ESA) improves the efficiency of data processing in the storage stack, reducing write amplification and write latency compared to the older Original Storage Architecture (OSA).

Modernized hardware: Intel platform innovation

Servers based on the latest generation of Intel® architecture are optimized with the latest VMware software to provide the foundations for today's hybrid and multi-cloud environments. Customers often find that older servers are no longer capable of meeting today's workload demands, and the need to upgrade is especially compelling for enterprises still running 1st Gen Intel Xeon Scalable processors, support for which ended December 31, 2023.

5th Gen Intel Xeon processors deliver advances across the balanced platform to enable high throughput, workload density and power efficiency. They provide up to 64 cores per socket with higher per-core performance than predecessors, supported by up to 320 MB last-level cache, DDR5 memory operating at up to 5600 MT/s and up to 80 lanes of PCIe. The processors provide advanced functionality for modern workloads using built-in hardware accelerators to offload critical tasks, that include the following:

- **Intel® Advanced Matrix Extensions (Intel® AMX)** accelerate AI inference and training.
- **Intel® QuickAssist Technology (Intel® QAT)** accelerates encryption and compression operations.
- **Intel® Software Guard Extensions (Intel® SGX)** isolate sensitive data in-use using hardware-protected memory.

Intel® Ethernet network controllers and interface cards extend the performance advantages of 5th Gen Xeon processors using up to 100 Gb Ethernet enabled with intelligent offloads to improve VM performance and scalability.

Performance and server consolidation results

At multiple block sizes — common for workloads running on vSAN — modernizing server hardware along with VMware software delivers significant performance increases. To quantify that benefit, Intel performance engineers compared I/O performance outcomes between 2nd Gen Intel Xeon processors running vSAN 7 OSA and 5th Gen Xeon CPUs running vSAN 8 ESA, as reflected in Figure 1.[1]

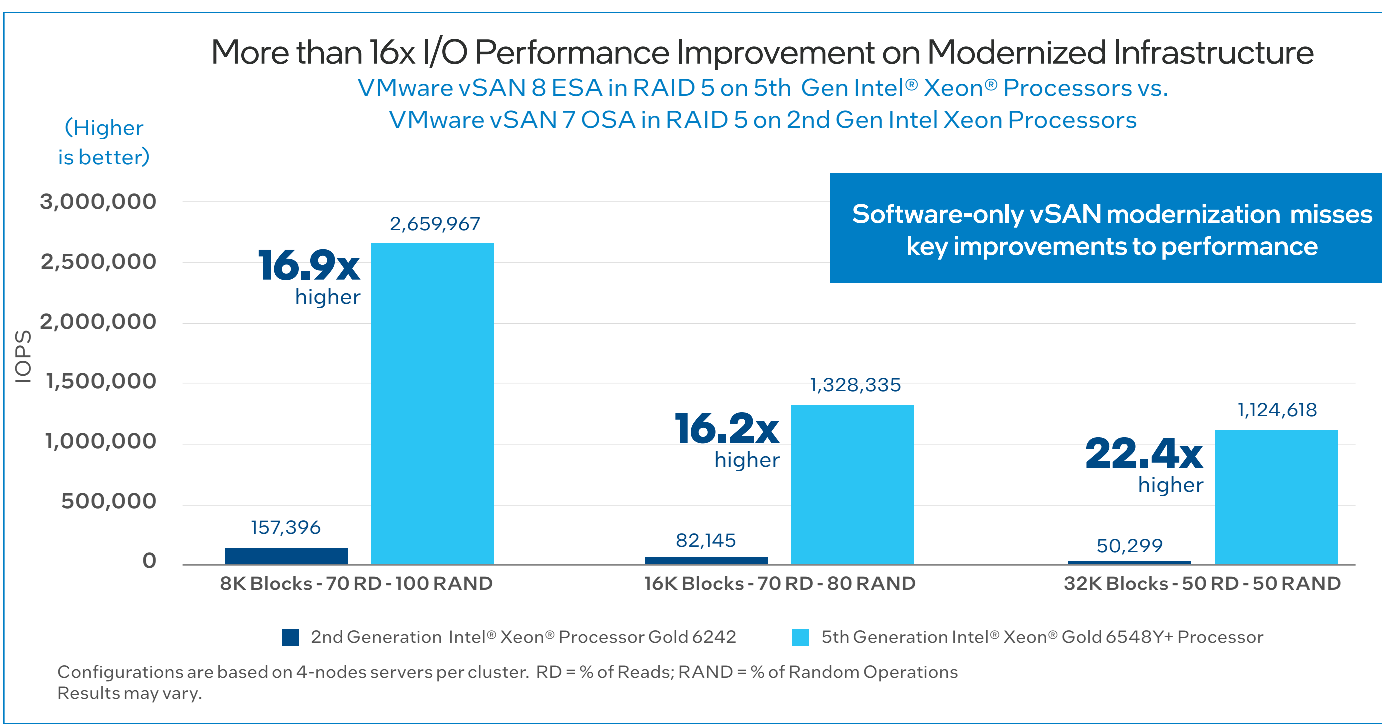


Figure 1. VMware vSAN performance gains, across block sizes, from modernizing hardware and software together.[1]

This comparison uses RAID 5 with both vSAN 7 and vSAN 8; RAID 5 was an option for vSAN 7, but not the out-of-box default until vSAN 8. The performance testing considers three workload scenarios with varying block sizes and finds substantial performance benefits for all three scenarios, which correspond to common workloads that run on vSAN HCI:

- **Scenario 1 (8K blocks, 70% read and 100% random)** simulates a typical filesystem or NoSQL database transaction profile, where refreshing both hardware and software together provides a 16.9x performance improvement.
- **Scenario 2 (16K blocks, 70% read and 80% random)** simulates another typical filesystem or Microsoft SQL database transaction profile, where refreshing both hardware and software together provides a 16.2x performance improvement.
- **Scenario 3 (32K blocks, 50% read and 50% random)** simulates another typical Microsoft SQL database transaction profile, where refreshing both hardware and software together provides a 22.4x performance improvement.

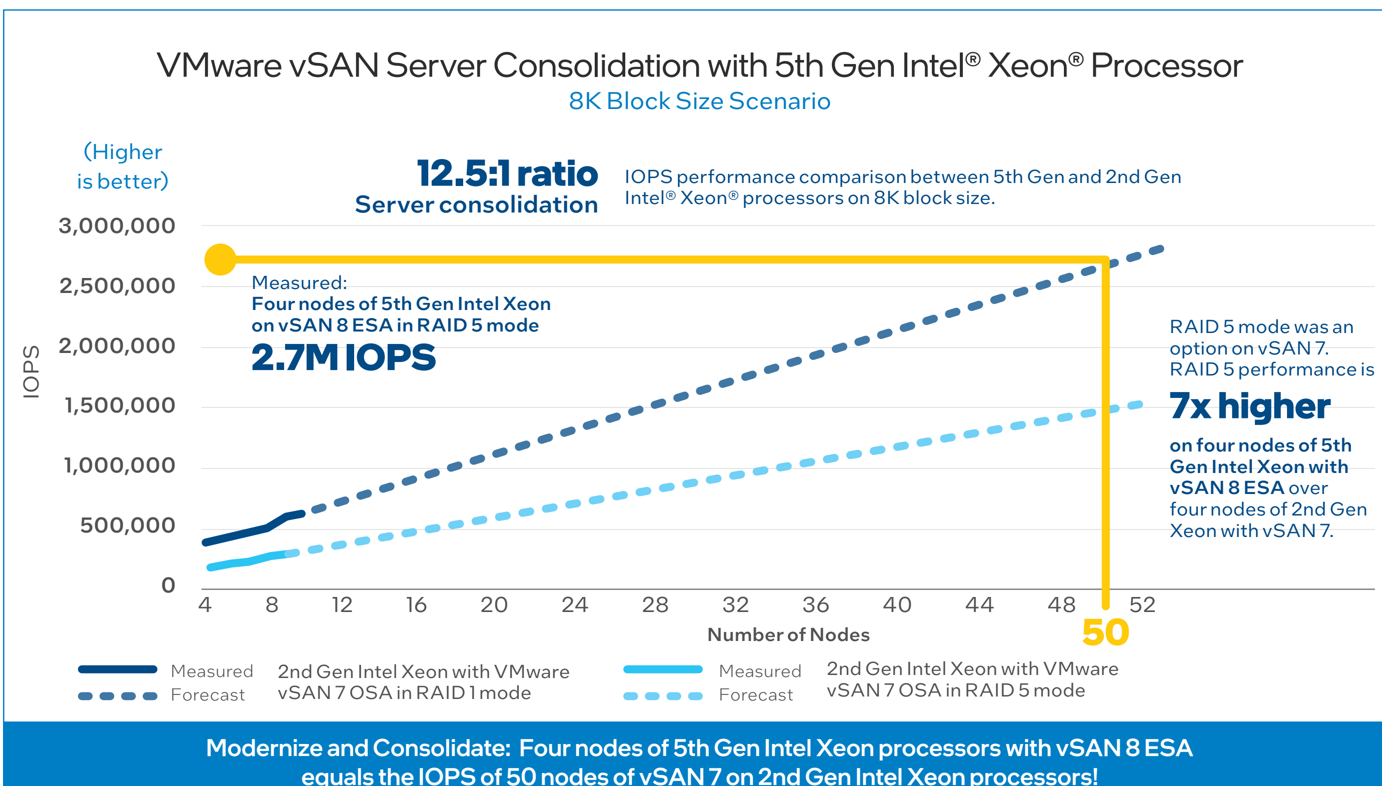
These performance improvements enable significant server consolidation as a result of hardware and software refresh, reducing the server count to support a given usage model. This smaller footprint promises TCO savings in data center resources as well as the staff requirements to deploy and maintain them. To demonstrate this benefit, Figure 2 expresses the IOPS measurements provided above for the 8K block size scenario in terms of server consolidation ratios.[2]

To arrive at this comparison, the team tested scale-out performance from four to eight nodes based on the older platform with both RAID 1 and RAID 5 modes on vSAN 7, then extrapolated those results to larger clusters using exponential smoothing forecasting. It then compared those projections with tested results using a four-node cluster based on 5th Gen Xeon CPUs with vSAN ESA with RAID 5.

Note: Exponential smoothing forecasting in Excel is based on the AAA version (additive error, additive trend and additive seasonality) of the Exponential Triple Smoothing (ETS) algorithm, which smooths out minor deviations in past data trends by detecting seasonality patterns and confidence intervals.

The dark blue line in the figure plots the performance in IOPS of vSAN 7 in RAID 1 mode (which was the default mode in vSAN 7) on 2nd Gen Xeon CPUs relative to the number of nodes in the cluster. The solid portion of the dark blue line represents the measured results on four to eight nodes, and the dashed portion represents the projection out to higher node counts using ETS.

The light blue line in the figure represents the performance of vSAN 7 in RAID 5 mode (an optional mode in vSAN 7) on the same 2nd Gen Xeon CPU-based hardware, measured and projected to higher node counts similarly to the RAID 1 case. RAID 5 performance is 7x higher on four 5th Gen Xeon CPU-based nodes with vSAN 8 ESA compared to four 2nd Gen Xeon CPU-based nodes with vSAN 7 OSA. As with the dark blue line, the solid portion of the light blue line represents measured results, and the dashed portion represents projected results.



The forecast uses measured values from 4 through 8 nodes to project higher node counts using the AAA version of the Exponential Smoothing (ETS) algorithm. Results may vary.

Figure 2. Server consolidation potential from modernizing hardware ad software together.[2]

The yellow dot in the figure represents the performance of a four-node vSAN 8 ESA cluster in RAID 5 (the default for vSAN 8 ESA) running on 5th Gen Xeon, delivering 2.7M IOPS. This result suggests that to deliver the same 2.7M IOPS would require 50 2nd Gen Xeon CPU-based nodes with vSAN 7, represented by the intersection of the dark blue and yellow lines. A modernization initiative that replaces 50 2nd Gen CPU-based nodes running vSAN 7 with just four 5th Gen CPU-based nodes would provide a server consolidation ratio of 12.5:1.

Note: Actual consolidation results will vary based on the specific workloads running on vSAN clusters.

Conclusion

Data center modernization is a strategic imperative for today's enterprises, supporting digital transformation and increased competitiveness. To capture the full performance and cost benefits of the technology opportunity, architects must consider upgrades to both hardware and software at the same time. The holistically modernized VMware and Intel environment enhances services and experiences for enhanced revenue as it reduces costs and technical debt.

Learn More

www.intel.com/vmware

Solution provided by:

[1] **Intel® Xeon® Gold 6242 (2nd Gen) – RAID 5:** Test by Intel as of 05/7/24. 4-node clusters, 2x Intel Xeon Gold 6242 CPU @ 2.80GHz, 16 cores, HT On, Turbo On, Total Memory 384GB (12x32GB DDR4 2666 MT/s [2666 MT/s]), BIOS 2.20.1, microcode 0x5003604, 2x I350 Gigabit Network Connection, 2x Ethernet Controller X710 for 10GbE SFP+, Boot: 2x 349.3G INTEL MDTPE21K375GA, Storage: 6x 1.8T INTEL SSDPE2KX020T8, OS/Software: VMware 7.0U3, 23794027, vSAN OSA – default policy (RAID 5, 2DG), using HCI Bench 2.8.3, FIO3.3. Throughput test 8K profile: I/O size 8K, Read percentage 70%, Random percentage 100%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB. Throughput test 16K profile: I/O size 16K, Read percentage 70%, Random percentage 80%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB. Throughput test 32K profile: I/O size 32k, Read percentage 50%, Random percentage 50%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB.

Intel Xeon Gold 6242 (2nd Gen) – RAID 1: Test by Intel as of 05/7/24. 4- node clusters, 2x Intel Xeon Gold 6242 CPU @ 2.80GHz, 16 cores, HT On, Turbo On, Total Memory 384GB (12x32GB DDR4 2666 MT/s [2666 MT/s]), BIOS 2.20.1, microcode 0x5003604, 2x I350 Gigabit Network Connection, 2x Ethernet Controller X710 for 10GbE SFP+, Boot: 2x 349.3G INTEL MDTPE21K375GA, Storage: 6x 1.8T INTEL SSDPE2KX020T8, OS/Software: VMware 7.0U3, 23794027, vSAN OSA – default policy (RAID 1, 2DG), using HCI Bench 2.8.3, FIO3.3. Throughput test 8K profile: I/O size 8K, Read percentage 70%, Random percentage 100%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB. Throughput test 16K profile: I/O size 16K, Read percentage 70%, Random percentage 80%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB. Throughput test 32K profile: I/O size 32k, Read percentage 50%, Random percentage 50%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB.

Intel Xeon Gold 6548Y+ (5th Gen): Test by Intel as of 05/7/24. 4-node cluster, 2x Intel Xeon Gold 6548Y+ CPU @ 2.50GHz, 32 cores, HT On, Turbo On, NUMA 2, Integrated Accelerators Available [used]: DLB 0 [0], DSA 2 [0], IAA 0 [0], QAT 0 [0], Total Memory 512GB (16x32GB DDR5 5600 MT/s [5200 MT/s]), BIOS 3B07.TEL2P1, microcode 0x21000200, 2x Ethernet Controller E810-C for QSFP, Boot: 2x 223.6G INTEL SSDSCKKB240GZ, Storage: 8x 2.9T KIOXIA KCD81PUG3T20, OS/Software: VMware ESXi 8.0.2, 23825572, vSAN ESA – Optimal default policy (RAID 5, flat), using HCI Bench 2.8.3, FIO3.3. Throughput test 8K profile: I/O size 8K, Read percentage 70%, Random percentage 100%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB. Throughput test 16K profile: I/O size 16K, Read percentage 70%, Random percentage 80%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB. Throughput test 32K profile: I/O size 32k, Read percentage 50%, Random percentage 50%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB.

[2] **Intel® Xeon® Gold 6242 (2nd Gen) – RAID 5:** Test by Intel as of 05/7/24. 4- to 8-node clusters, 2x Intel Xeon Gold 6242 CPU @ 2.80GHz, 16 cores, HT On, Turbo On, Total Memory 384GB (12x32GB DDR4 2666 MT/s [2666 MT/s]), BIOS 2.20.1, microcode 0x5003604, 2x I350 Gigabit Network Connection, 2x Ethernet Controller X710 for 10GbE SFP+, Boot: 2x 349.3G INTEL MDTPE21K375GA, Storage: 6x 1.8T INTEL SSDPE2KX020T8, OS/Software: VMware 7.0U3, 23794027, vSAN OSA – default policy (RAID 5, 2DG), using HCI Bench 2.8.3, FIO3.3. Throughput test 8k profile: I/O size 8k, Read percentage 70%, Random percentage 100%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB.

Intel Xeon Gold 6242 (2nd Gen) – RAID 1: Test by Intel as of 05/7/24. 4- to 8-node clusters, 2x Intel Xeon Gold 6242 CPU @ 2.80GHz, 16 cores, HT On, Turbo On, Total Memory 384GB (12x32GB DDR4 2666 MT/s [2666 MT/s]), BIOS 2.20.1, microcode 0x5003604, 2x I350 Gigabit Network Connection, 2x Ethernet Controller X710 for 10GbE SFP+, Boot: 2x 349.3G INTEL MDTPE21K375GA, Storage: 6x 1.8T INTEL SSDPE2KX020T8, OS/Software: VMware 7.0U3, 23794027, vSAN OSA – default policy (RAID 1, 2DG), using HCI Bench 2.8.3, FIO3.3. Throughput test 8k profile: I/O size 8k, Read percentage 70%, Random percentage 100%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB.

For both RAID 5 and RAID 1 testing on Intel Xeon Gold 6242 CPUs), performance was measured at 4 to 8 nodes (five data points), and then projected to higher node counts using exponential smoothing forecasting. Exponential smoothing forecasting in Excel is based on the AAA version (additive error, additive trend and additive seasonality) of the Exponential Triple Smoothing (ETS) algorithm, which smooths out minor deviations in past data trends by detecting seasonality patterns and confidence intervals.

Intel Xeon Gold 6548Y+ (5th Gen): Test by Intel as of 05/7/24. 4-node cluster, 2x Intel Xeon Gold 6548Y+ @ 2.50GHz, 32 cores, HT On, Turbo On, NUMA 2, Integrated Accelerators Available [used]: DLB 0 [0], DSA 2 [0], IAA 0 [0], QAT 0 [0], Total Memory 512GB (16x32GB DDR5 5600 MT/s [5200 MT/s]), BIOS 3B07.TEL2P1, microcode 0x21000200, 2x Ethernet Controller E810-C for QSFP, Boot: 2x 223.6G INTEL SSDSCKKB240GZ, Storage: 8x 2.9T KIOXIA KCD81PUG3T20, OS/Software: VMware ESXi 8.0.2, 23825572, vSAN ESA – Optimal default policy (RAID 5, flat), using HCI Bench 2.8.3, FIO3.3. Throughput test 8k profile: I/O size 8k, Read percentage 70%, Random percentage 100%, latency target mode<10ms, #VMs per cluster 16, vCPU 4, vRAM 8, # data disks per VM 4, size of disk 50GB.

0824/CM/MESH/356889-001US