Technical Research Report

PROWESS

For Peak Performance, Bare Metal with NVMe® RAID Comes Out on Top

Prowess Consulting testing showed that bare-metal servers with NVMe RAID controllers offer an ideal solution for performance-hungry workloads, such as analytics, AI, ML, and HPC.

Executive Summary

VMware vSAN™ has become a reliable workhorse platform for many businesses. The hyperconverged infrastructure (HCI) solution lets organizations more easily manage their workloads while helping ensure reliability and availability for their customers and for their own internal needs. But there are situations when businesses might find themselves needing higher levels of performance that aren't easily reached using VMware vSAN.

With VMware vSAN, the hypervisor and management software layers add latency to the overall platform, creating performance gaps that are difficult to overcome by upgrading servers, memory, or other hardware components. Non-HCI solutions with RAID offer an alternative for ensuring availability, but these platforms typically do not offer any performance improvements because they rely on Serial Attached SCSI (SAS)–based RAID controllers.

Newer NVM Express® (NVMe®)-based RAID controllers provide a higher-performance alternative to VMware vSAN because the NVMe protocol can overcome the latency gaps caused by the slower SAS protocol. To determine the viability of using platforms with NVMe RAID controllers for demanding workloads, Prowess Consulting ran comparison tests of this configuration against a VMware vSAN platform. Results showed that Dell™ PowerEdge™ R750 servers, built with newer Dell™ PowerEdge RAID Controller 11 (PERC 11) NVMe RAID controllers, significantly boosted performance for critical workloads while still ensuring the high levels of reliability that businesses require. For example, Microsoft® SQL Server® workloads on a single PowerEdge R750 server with RAID showed 13.9x more transactions in HammerDB TPROC-C database benchmark testing than did a four-server cluster built with VMware vSAN.[1]

Availability, Reliability, and Performance: Can You Have It All?

VMware vSAN is one of the most popular HCI solutions in the industry, with a 43-percent market share.[2] There's good reason for this popularity: the solution offers businesses a reliable platform for running traditional enterprise applications and services with high availability. And because the platform is so widespread, there is no shortage of IT professionals with vast VMware experience.

Indeed, for many workloads, VMware vSAN can be an ideal platform to deploy. However, there are some workloads that push the performance boundaries of what VMware vSAN can provide. Even with hardware and network performance tuning, organizations might find themselves or their customers growing frustrated while waiting for results. This situation can be particularly apparent when running many of the most common emerging artificial intelligence (AI), machine learning (ML), analytics, and database-dependent workloads.

The problem is exacerbated when business applications are dependent on large and ever-increasing quantities of data. For example, a business hosting a customer-accessed website needs to be able to process large amounts of rapidly changing data quickly or risk losing out to the competition. Retail companies also rely on up-to-date sales, inventory, and customer data for processing orders and gathering critical insights. Healthcare organizations process massive quantities of patient data, prescriptions, scheduling information, and other medical records on a daily basis.

For time-critical workloads like these, businesses need a reliable way to ramp up performance. Adding more VMware vSAN clusters can help to a degree but might be prohibitive in terms of both cost and physical footprint in the data center. In addition, expanding VMware vSAN clusters still might not go far enough to achieve the levels of performance needed for critical workloads. Layering hypervisors, virtual machines (VMs), and operating systems on top of hardware will always add overhead that contributes to latency.

For performance-hungry workloads, organizations need another option that can boost performance without sacrificing the required levels of availability and reliability for apps and data.

The Game-Changing Shift from SAS to NVMe RAID

One option for significantly reducing latency is to eliminate the overhead caused by virtualization by deploying workloads directly on bare-metal servers. But bare-metal deployments traditionally come with tradeoffs, particularly when it comes to maintaining availability and integrity of data.

RAID is an obvious choice to address availability concerns, but traditional SAS-based RAID controllers and drives can't provide significant (if any) performance gains over a VMware vSAN configuration. However, those constraints can now be eliminated by newer RAID controllers. Recent RAID controllers from Dell Technologies, for example, support the NVMe interface specification for connecting over a fast PCIe® bus. RAID and storage built on NVMe and PCIe can significantly improve bandwidth and reduce latency over traditional SAS-based or Serial ATA (SATA®)-based controllers.

Prowess Consulting decided to put this possibility to the test. Could a bare-metal server with fast NVMe RAID controllers and drives offer a higher-performance alternative to VMware vSAN deployments for critical workloads?

Putting NVMe RAID to the Test

Prowess testing pitted a configuration built with VMware vSAN against a bare-metal PowerEdge R750 server built with Dell PERC 11 RAID controllers that support the NVMe protocol.

We used SQL Server for testing, as this application is commonly used with data-intensive workloads that require top performance for businesses and their customers.

The system configurations for this test are shown in Table 1.

Table 1 | System configurations used for testing VMware vSAN™ vs. bare-metal/RAID performance

Configuration	Single-Server RAID	Four-Server VMware vSAN™ Cluster
Server	1 x Dell™ PowerEdge™ R750	4 x Dell™ PowerEdge™ R750
Processor	Intel® Xeon® Gold 6338 processor	Intel® Xeon® Gold 6330 processor
Number of CPUs	2	2
Cores/threads per CPU	32/64	28/56
Cores/threads total	64/128	56/112
Frequency (base/SCT/MCT)	2.00 GHz	2.00 GHz
Storage controller 01	Dell™ PowerEdge™ RAID Controller (PERC) H755N Front NVM Express® (NVMe®) Broadcom® LSI	Broadcom® LSI Dell™ HBA355i Front
Disk	3.2 TB KIOXIA® Dell™ NVMe® CM6 MU	1 x 480 GB Intel® SSDSC2KG480GZR 7 x 1.2 TB Seagate® ST1200MM0099
Number of disks	8	8
Storage controller 02	Dell™ PERC H755N Front NVMe® Broadcom® controller	Marvell® Technology Group Ltd. Dell™ Boot Optimized Server Storage (BOSS)-S2
Disk	3.2 TB KIOXIA® Dell™ NVMe® CM6 MU (KCM6XVUL3t20)	240 GB Micron® MTFDDAV240TDU
Number of disks	8	2
Storage controller 03	Marvell® Technology Group Ltd. Dell™ BOSS-S2	Not applicable (N/A)
Disk	480 GB Micron® MTFDDAV480TDS	N/A
Number of disks	2	N/A
Installed memory	256 GB	128 GB
Memory DIMM	Hynix® HMA82GR7DJR8N-XN 16 GB dual-rank DDR4	Hynix® HMA82GR7DJR8N-XN 16 GB dual-rank DDR4
Memory speed	3,200 megatransfers per second (MT/s)	2,933 MT/s (max 3,200 MT/s)
Number of memory DIMMs	16 x 16 GB	8 x 16 GB
Operating system (OS)	Red Hat® Enterprise Linux® 8.6 (Ootpa)	VMware ESXi™ 7.0.3
OS kernel	Linux® 4.18.0-372.26.1.el8_6.x86_64	20328353
Microsoft® SQL Server® version	Microsoft® SQL Server® Enterprise Evaluation (64-bit) 16.0.950.9 Microsoft® SQL Server® 2022 (RC1)	
BIOS version	1.6.5	1.7.5
OS performance profile	Tuned for Microsoft® SQL Server®	

One VMmark® 3 tile and a single Red Hat® Enterprise Linux® 8.7 VM with SQL Server 2022 (RC1) was deployed on the VMware vSAN cluster.

Measuring Database Performance

The goal of this testing was to generate performance data showing the new orders per minute (NOPM) performance of a SQL Server database running on each PowerEdge server platform. Prowess engineers also collected other performance data to verify that the two systems were operating as intended with comparable configurations.

For these tests, HammerDB was run using the TPROC-C benchmark to assess online transaction processing (OLTP) performance against the MySQL database. Specifically, the benchmark was run against a 640-warehouse database with 32, 64, 96, 128, 160, and 192 virtual users. We selected the TPROC-C benchmark because it is derived from the industry-standard TPC-C® workload, and it offers a useful, repeatable test with results presented in NOPM.

As Figure 1 shows, our testing revealed significant performance gains when running SQL Server workloads on the bare-metal PowerEdge R750 server with an NVMe RAID array, compared to running those workloads on a cluster of four PowerEdge R750 servers in a VMware vSAN configuration.

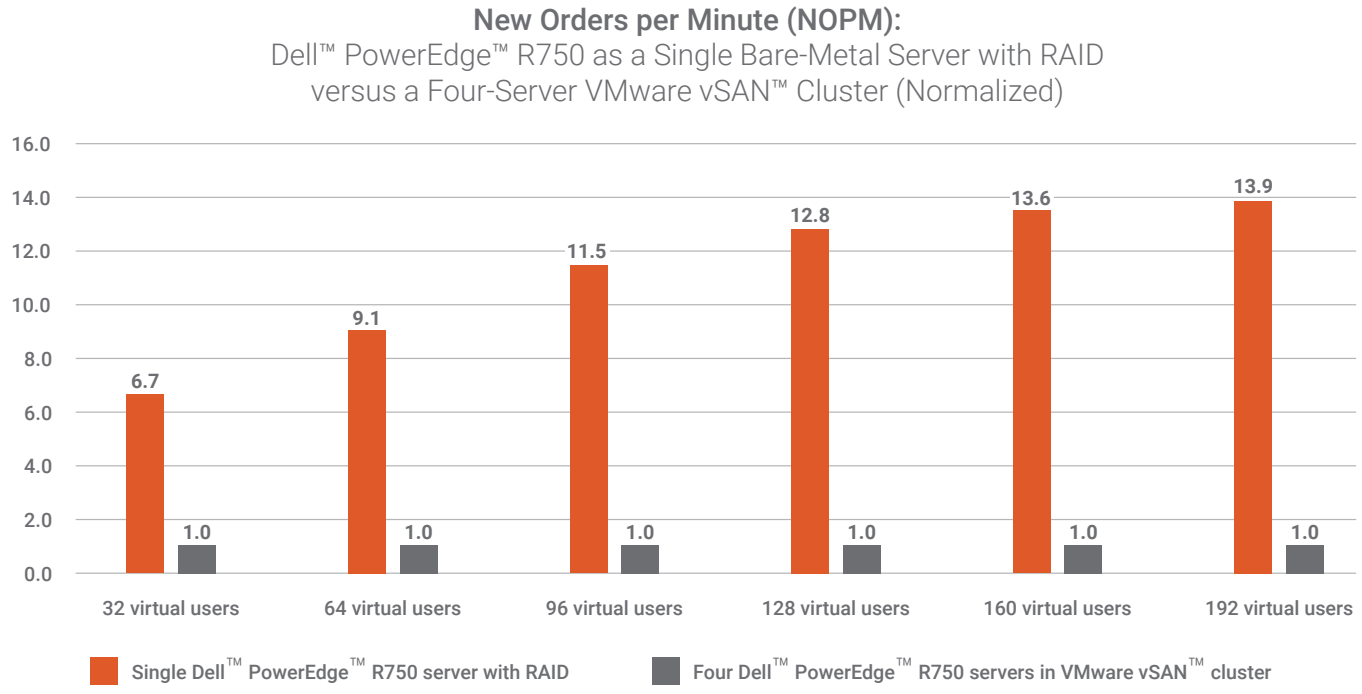

Figure 1. Database performance comparing a bare-metal Dell™ PowerEdge™ R750 server with RAID to four PowerEdge R750 servers in a VMware vSAN™ cluster

The test results were impressive. The bare-metal deployment with NVMe RAID controllers showed significantly higher NOPM across all tested virtual user configurations, from 6.7x more NOPM at 32 virtual users to 13.9x more NOPM at 192 virtual users. This represents a NOPM increase of up to 93 percent for the bare-metal configuration.

Behind the Results

As the results show, eliminating the overhead inherent in a hypervisor-based system resulted in significant performance gains. There's nothing earth-shattering in measuring higher performance for a bare-metal system alone. What stands out in this test scenario was that these performance gains were apparent even with the use of RAID drives, which traditionally have exhibited unacceptable levels of latency. By incorporating NVMe, the Dell PERC 11 controllers and drives in the tested PowerEdge R750 servers provided the required levels of availability for the platform without contributing measurable latency.

Using bare-metal systems with RAID can also provide other advantages over a VMware vSAN configuration, including the flexibility to fine-tune capacity versus performance. For example, with SQL Server, the transaction log acts as a buffer for the main database. That means administrators can optimize for performance by configuring the front RAID controllers to use RAID 10, which is the highest performing option. For the back-end database, which is not as performance sensitive, admins can choose to use RAID 5 (or any preferred level) to sacrifice some performance for increased capacity. VMware vSAN doesn't provide this same level of flexibility because it only lets users specify the cache and capacity levels, with other aspects of load-balancing and performance management handled internally by the VMware vSAN software.

For more details on how RAID with support for NVMe drives affects performance for SQL Server transactions, compared to SAS-based drives, see this **previous test report** from Prowess Consulting, which includes details on log disk reads and writes.

NVMe RAID Controllers

As discussed previously, NVMe adapters were the key to the observed performance gains. The PowerEdge R750 server used in the testing paired two Dell PERC 11 H755N Front NVMe adapters with all-NVMe solid-state drives (SSDs) to maximize storage bandwidth and throughput.

The Dell PERC 11 H755N Front NVMe adapter is based on the Broadcom® SAS3916 PCIe to SAS/SATA/PCIe RAID-on-Chip (RoC) controller. These are the first RAID controllers from Dell Technologies to offer both PCIe Gen4 host and PCIe Gen4 storage interfaces, which deliver double the bandwidth and 75 percent more input/output operations per second (IOPS), compared to previous generations.[3]

The Dell PERC 11 H755N Front NVMe adapter provides high PCIe (NVMe) storage-interface data-transfer rates of 16 gigatransfers per second (GT/s), 8 GT/s, 5 GT/s, and 2.5 GT/s per lane, in addition to reliability, high performance, and fault-tolerant disk subsystem management and support for RAID levels 0, 1, 5, 6, 10, 50, and 60.

Ethernet Controllers

For the purposes of this testing, we isolated the test platforms to remove network speed as a variable. However, the PowerEdge R750 server includes the Broadcom® NetXtreme® E-Series BCM57414 50G PCIe 3.0 Ethernet controller, based on Broadcom's scalable 10/25/50/100/200 gigabit Ethernet (GbE) controller architecture. This network card combines a high-bandwidth Ethernet controller with a unique set of highly optimized hardware-acceleration engines to enhance network performance and improve server efficiency for enterprise and cloud-scale networking and storage applications, including high-performance computing (HPC), telco, ML, storage disaggregation, and data analytics.

Dell™ PowerEdge™ R750 Server Advantages

The Dell™ PowerEdge™ R750 server is an enterprise server designed to deliver high performance for demanding workloads. Powered by 3rd Gen Intel® Xeon® Scalable processors, the PowerEdge R750 server is a dual-socket/2U rack server with support for eight channels of memory per CPU and up to 32 DDR4 DIMMs at 3,200 MT/s speeds. In addition, the PowerEdge R750 server supports PCIe® Gen4 and up to 24 NVM Express® (NVMe®) drives (with an option for four additional rear-mounted drives) with improved air-cooling features and optional Direct Liquid Cooling (DLC) to support increasing power and thermal requirements. This makes the PowerEdge R750 server a compelling option for a wide range of workloads, including database and analytics, high-performance computing (HPC), traditional corporate IT, virtual desktop infrastructure (VDI), and artificial intelligence (AI)/machine learning (ML) environments that require performance, extensive storage, and graphics processing unit (GPU) support. For more information, see the **PowerEdge R750 specification sheet** on the Dell Technologies website.

Bare Metal with RAID Delivers Top Performance with Availability

Compared to VMware vSAN, bare-metal deployments of SQL Server on platforms built with NVMe-based RAID offer a significant performance boost while still providing required levels of reliability for apps and data. Prowess Consulting testing measured gains of up to 13.9x for SQL Server transactions performed on a standalone server configured with Dell PERC 11 NVMe RAID controllers, compared to a four-server VMware vSAN cluster.

In addition, the bare-metal configuration shows the potential for reducing an organization's data center footprint, along with the associated power and cooling costs. The four-server cluster used in our testing was replaced with a single server, representing a 75-percent decrease in the server chassis space required.

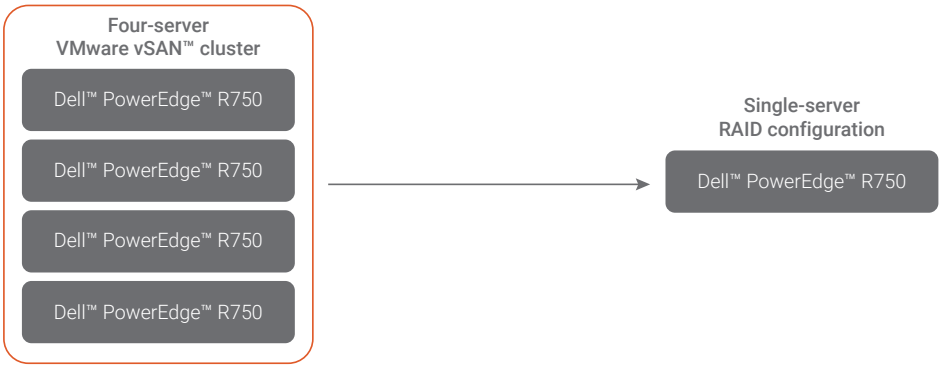

Figure 2. A standalone configuration with NVMe® RAID reduces infrastructure footprint, compared to a VMware vSAN™ cluster

There's no question that VMware vSAN is still useful for many workloads, including virtual desktop infrastructure (VDI) deployments, test/dev environments, and traditional business applications such as Microsoft® Exchange, accounting systems, and other services. But as Prowess Consulting testing showed, bare-metal servers with NVMe RAID controllers offer an ideal alternative for performance-hungry workloads, such as analytics, AI, ML, and HPC.

For detailed testing methodology and configurations used in this study, see "**Behind the Report: For Peak Performance, Bare Metal Wins**."

Learn more about the Dell PowerEdge servers discussed in this study:

Dell PowerEdge R750

[1] Based on testing by Prowess as of December 2022. For configuration details, see "Behind the Report: For Peak Performance, Bare Metal Wins." January 2023. **www.prowesscorp.com/project/dell-poweredge-r750-bare-metal-with-NVMe-raid-boosts-performance/**.

[2] VMware. "HCI Systems Powered by VMware vSAN Lead the Market in Q1 of 2022, According to IDC." July 2022. **https://blogs.vmware.com/virtualblocks/2022/07/28/hci-systems-powered-by-vsan-lead-market-q1-2022/**.

[3] Broadcom internal data, provided by Dell Technologies.