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

It's no surprise that bare-metal servers often provide the highest level of performance for demanding workloads. Organizations might be tempted to use bare-metal deployments to maximize performance and value from their platforms, especially for use with demanding workloads such as analytics, artificial intelligence (AI), machine learning (ML), and high-performance computing (HPC). Until recently, this approach has been blocked due to concerns around high availability. For example, although RAID arrays could be used with bare-metal deployments to ensure availability, this option hasn't been viable in the past due to performance constraints of Serial Attached SCSI (SAS)–based RAID controllers.

Newer NVM Express® (NVMe®)-based RAID controllers remove that constraint. The NVMe protocol in newer RAID controllers, such as Dell™ PowerEdge™ RAID Controller 11 (PERC 11), 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 tests to measure transactional database performance using a Dell™ PowerEdge™ R750 server, built with Dell PERC 11 NVMe RAID controllers. We compared performance to a Dell PowerEdge R740xd, configured with Dell PERC 10 controllers using SATA SSDs. In our testing, the PowerEdge R750 with NVMe RAID demonstrated an impressive 2.92x increase in database transaction performance, demonstrating that the newer platform offers a compelling option for businesses looking for high performance with the reliability and availability offered by PERC 11 RAID.

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

The proliferation of high-performance workloads has pushed the performance boundaries of what many traditional deployments 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 AI, 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 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 overhead by deploying workloads directly on bare-metal servers instead of virtualized environments. 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 or Serial ATA (SATA®)-based RAID controllers and SSDs can't provide significant (if any) performance gains over traditional deployments. 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 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 high-performance option for critical workloads? And if so, how much of a performance boost would even newer PERC 11 RAID controllers in Dell PowerEdge R750 servers offer over previous-generation Dell PowerEdge R740xd servers built with PERC 10 controllers?

Putting NVMe RAID to the Test

Prowess testing looked at transactional database performance for a bare-metal PowerEdge R750 server built with Dell PERC 11 RAID controllers that support the NVMe protocol. We then ran comparable tests on a PowerEdge R740xd server built with Dell PERC 10 RAID controllers.

We used Microsoft® 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 bare-metal/RAID performance

Configuration	Dell™ PowerEdge™ Server Platform	
Server	1 x Dell™ PowerEdge™ R740xd	1 x Dell™ PowerEdge™ R750
Processor	Intel® Xeon® Gold 6238 processor	Intel® Xeon® Gold 6338 processor
Number of CPUs	2	2
Cores/threads per CPU	22/44	32/64
Cores/threads total	44/88	64/128
Frequency (base/SCT/MCT)	2.10 GHz	2.00 GHz
Storage controller 01	Dell™ PowerEdge™ RAID Controller (PERC) H740P Front Broadcom® LSI	Dell™ PowerEdge™ RAID Controller (PERC) H755N Front Broadcom® LSI

Disk	960 GB Samsung MZ7LH960HBJR0D3 SATA SSD	3.2 TB KIOXIA® Dell™ CM6 MU NVMe® SSD (KCM6XVUL3T20)
Number of disks	16	8
Storage controller 02	BOSS-S1	Marvell® Technology Group Ltd. Dell™ BOSS-S2
Disk	240 GB Intel SSDSCKKB240G8R M.2 SATA	3.2 TB KIOXIA® Dell™ CM6 MU NVMe® SSD (KCM6XVUL3T20)
Number of disks	2	8
Storage controller 03	-	Dell™ PERC H755N Front NVMe® Broadcom® controller
Disk	-	480 GB Micron® MTFDDAV480TDS
Number of disks	-	2
Installed memory	256 GB	256 GB
Memory DIMM	Hynix® HMA84GR7JJR4N-WM DDR4	Hynix® HMA82GR7DJR8N-XN 16 GB dual-rank DDR4
Memory speed	2,933 megatransfers per second (MT/s)	3,200 megatransfers per second (MT/s)
Number of memory DIMMs	8 x 32 GB	16 x 16 GB
Network	BCM57416 NetXtreme-E Dual-Media 10 Gb RDMA Ethernet Controller NetXtreme® BCM5720 Gb Ethernet PCIe® Ethernet 25G 2P E810-XXV Adapter	1 x Broadcom® Advanced Dual 25 Gb Ethernet
Operating system (OS)	Red Hat® Enterprise Linux® 8.5 (Ootpa)	Red Hat® Enterprise Linux® 8.8 (Ootpa)
OS kernel	Linux® 4.18.0-348.20.1.el8_5.x86_64	Linux® 4.18.0-477.15.1.el8_8.x86_64
BIOS version	2.13.3	1.10.2
OS performance profile	Tuned for Microsoft® SQL Server®	Tuned for Microsoft® SQL Server®
Microsoft® SQL Server® version	Microsoft® SQL Server® 2019 (RTM-CU15) (KB5008996)—15.0.4198.2 (X64)	Microsoft® SQL Server® 2022 (RTM-CU6) (KB5027505)—16.0.4055.4 (X64)

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 the Dell PowerEdge server platform. Prowess engineers also collected other performance data to verify that the system was operating as intended.

Prowess Consulting engineers generated performance data showing the NOPM performance of a SQL Server database running on each PowerEdge server platform. For this testing, we used BenchCraft®, a Microsoft benchmarking tool that processes data like a TPC-C® benchmark.[1]

As Figure 1 shows, our testing revealed a significant jump in performance when running a SQL Server workload on the bare-metal PowerEdge R750 server with an NVMe RAID array, compared to a similar workload on the PowerEdge R740xd with a SATA array.

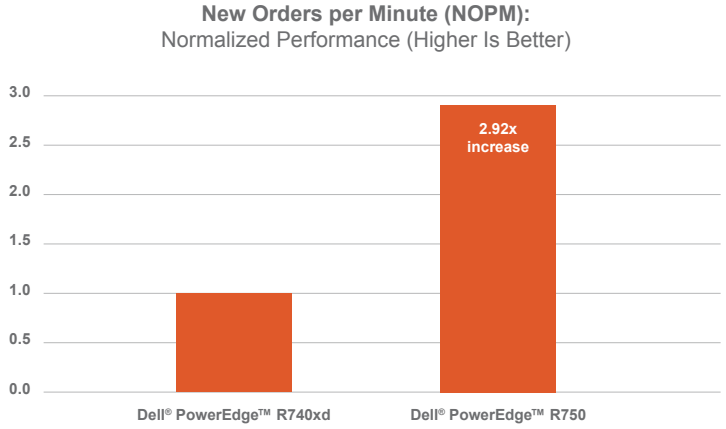

Figure 1. Database performance in new orders per minute (NOPM), comparing a bare-metal Dell™ PowerEdge™ Dell R740xd server with Dell® PERC 10 RAID to a Dell PowerEdge R750 server with Dell PERC 11 RAID

The test results showed that a bare-metal deployment with NVMe RAID controllers can provide exceptional results for users looking to maximize database performance on bare-metal servers, without giving up the availability benefits provided by a RAID array.

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 benefits. 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.

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, 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 were 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.[2]

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 that require performance, extensive storage, and graphics processing unit (GPU) support. These workloads include database and analytics, high-performance computing (HPC), traditional corporate IT, virtual desktop infrastructure (VDI), and artificial intelligence (AI)/machine learning (ML) environments. For more information, see the PowerEdge R750 specification sheet on the Dell Technologies website.

Bare Metal with RAID Delivers Top Performance with Availability

Bare-metal deployments of SQL Server on platforms built with NVMe-based RAID offer exceptional performance while still providing required levels of reliability for apps and data. As our testing showed, this configuration offers organizations an ideal option for deploying 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] Note that BenchCraft® does not conform to the TPC-C® testing standards, and results between BenchCraft and TPC-C are not comparable.

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

0823/220132