

Reduce TCO and Build Resilient, Responsive Databases with Kingston's DC1500M Enterprise NVMe Solid State Drives and VMware vSAN HCI

Written by: Hazem Awadallah, Systems Engineer, Kingston Technology

Reviewed by: Chris Selden, SSD Product Engineering Manager, Kingston Technology



Contents

- <u>Executive Summary</u>
- <u>Common infrastructure challenges that face RDBMS in the data center today</u>
- <u>Solution: Introducing Kingston Technology Data Center DC1500M Enterprise NVMe SSDs</u>
- <u>Testing Environment</u>
 - o <u>I. Infrastructure</u>
 - o <u>II. Database configuration</u>
 - o <u>III. vSAN storage performance</u>
- Testing Results
 - o <u>Test 1, DC1500M 960GB vSAN SQL server 2017 VM with varying amounts of DRAM</u>
 - Test 1 Results, DC1500M 960GB vSAN SQL server 2017 VM with varying amounts of DRAM
 - <u>Test 2: Comparing SQL Server 2017 performance on Kingston DC500M SATA SSD, Micron</u> <u>5200 eco SATA SSD and DC1500M NVMe SSD vSAN datastore</u>
 - <u>Test 2 Results: Comparing SQL Server 2017 performance on Kingston DC500M SATA SSD</u>, <u>Micron 5200 eco SATA SSD and DC1500M NVMe SSD vSAN datastore</u>
 - <u>Test 3: SQL Server 2017 performance comparison DC1500M NVMe vs Micron 5200 eco</u> <u>SATA vSAN datastore, larger schema size and longer test duration</u>
 - Test 3 results: SQL Server 2017 performance comparison DC1500M NVMe vs Micron 5200 eco SATA vSAN datastore, larger schema size and longer test duration
 - <u>Test 4: SQL Server 2017 performance comparison, backup and restore performance,</u> <u>DC1500M NVMe vs Micron 5200 eco SATA vSAN</u>
 - <u>Test 4: Results: SQL Server 2017 performance comparison, backup and restore</u> performance, DC1500M NVMe vs Micron 5200 eco SATA vSAN
 - Test 5: SQL Server 2017 performance comparison, The Noisy Neighbor test, DC1500M
 NVMe vs Micron 5200 eco SATA vSAN
 - <u>Test 5 results: SQL Server 2017 performance comparison, The Noisy Neighbor test,</u> <u>DC1500M NVMe vs Micron 5200 eco SATA vSAN</u>
- <u>Conclusion</u>



Executive summary

In recent years, the introduction of NVMe has revolutionized the field of data storage, taking a big leap forward in maximizing the performance of NAND flash and capitalizing on the feature rich, low cost, high bandwidth, and future proof expansion bus standard, PCI express. Currently on its 5th Generation, PCIe Gen5 allows for transfer rates of up to 8GB/s per lane, removing the expansion bus bottleneck from the storage stack and giving headway for innovation and evolution not only in SSD controllers and NAND flash, but throughout the hardware stack. Processors, chassis designs, motherboards, and hardware IO topologies are constantly evolving to support the added bandwidth. In the data center, network topologies are undergoing major changes to accommodate for NVMe; with the NVMe-OF spec, network interfaces, switches and transport protocols have changed and continue to improve to support the increased bandwidth while maintaining QoS and lossless packet transport.

But how does the introduction of NVMe impact application performance? Can you reduce your storage footprint while improving transaction throughput and reducing transaction response times? Can we significantly reduce database backup times to mitigate the Noisy Neighbor problem and minimize its impact in a production environment? In this article, we attempt to answer these questions by inspecting typical OLTP workloads, as defined by the TPCC specification, and offer a few practical comparisons to show the impact of NVMe on transaction performance in realistic scenarios.

Common infrastructure challenges that face RDBMS in the data center today

Cost, Capacity Planning and Scalability

With the tremendous increase in internet bandwidth, processing speeds and the data analytics boom that occurred over the last 2 decades, production OLTP databases are growing quickly, often a lot faster than planned for by application and infrastructure architects. The underlying storage and network architecture must be built to scale from the ground up to match that increased demand over time and offer a good balance between cost, ease of management and performance. It becomes a difficult design decision to choose to build the application in local data centers or using IaaS/PaaS cloud services. Keeping the application running in local data centers gives solution architects full control of scalability, security, resilience, and performance but requires meticulous planning and sometimes comes at a hefty upfront cost. Using IaaS/PaaS cloud services speeds up deployments and simplifies scalability but offers less control over performance, resilience and can get expensive quickly as the application scales. Some organizations prefer a hybrid approach, where more important tier 1 applications that are kept in-house, hyperconverged infrastructure solutions like VMware vSAN with All-Flash disk groups offer a good balance between cost, simplicity, performance, and ease of scalability.



Resiliency

Tier 1 applications must be built or migrated to infrastructure that can withstand more than one hardware failure throughout the entire hardware stack. If not planned for correctly, equipment failures in data centers can cause significant monetary loss via service disruptions or in worst-case scenarios, permanent data loss. In shared storage environments, careful planning must be made to ensure that the underlying infrastructure is built to withstand storage failures and component performance overload.

With vSAN, for example, tier 1 applications should have a minimum Failure to Tolerate (FTT) of 1, with vSphere High Availability (HA) enabled, to ensure that the application and database VMs are protected from at least one compute, network or storage failure. Additionally, vSphere Distributed Resource Scheduler (DRS) can then be enabled to load balance CPU/memory resources across the physical servers in the cluster.

Varying Performance Expectations

The demand for higher transaction speed and lower latency continues to increase as OLTP applications continue to scale up, with more users placing a higher transactional load on the backend database. Application architects must plan for storage infrastructure that can adapt to support this increased demand and is flexible enough to be migrated between different tiers of storage. For example, SQL databases residing on virtual disks provisioned from SAN storage arrays can be migrated to an NVMe all flash vSAN datastore with faster tiers of storage like NVMe using VMware's storage VMotion.

The Noisy Neighbor Dilemma

It is imperative to design infrastructure that allows key workloads to have the resources they need to execute. In a shared storage environment with multiple workloads, performance can become unpredictable and aberrant workloads can cause problems for key production workloads. This is a definition of the Noisy Neighbor problem. An example, as we see later in this paper, would be unscheduled database backup operations on one server, consuming storage and network resources and affecting the performance and latency of other servers using the same resources.

Introducing Kington DC1500M Enterprise NVMe SSDs

Kingston DC1500M is the latest Enterprise U.2 PCIe 3.0x4 NVMe offering by Kingston with capacities ranging from (960GB-7680G). Equipped with a 16-channel controller and 3D TLC NAND, it was designed with strict Quality of Service (QoS) requirements to ensure sustained high performance and consistency of enterprise workloads while maintaining the lowest latency. Its enterprise-focused firmware supports features like overprovisioning, multiple namespaces (supporting up to 64 namespaces) as well as more sophisticated ECC algorithms to ensure reliability of enterprise workloads within the entire lifetime of the drive.

With SATA SSDs still the most prevalent SSDs in the data center, in this paper, we aim to show that migrating or building your storage infrastructure on Enterprise NVMe SSDs, like Kingston's DC1500M NVMe, will help ease some problems mentioned above.



In our internal testing, a single Kingston DC1500M NVMe SSD offers up to 6.5x the throughput and 5.6x latency improvement (Figure b below) comparing it to 1 Micron 5200 eco Enterprise SATA SSD, with little to no cost parity.

This level of performance in a hyperconverged environment translates to higher transaction throughput and lower latency for SQL Server databases. It also translates to a lower storage footprint and lower power consumption. In this example, you need 6 micron 5200 eco drives to match the throughput of 1 DC1500M drive. We will see later how this performance translates in realistic SQL OLTP workloads on VMware vSAN.

The dramatic performance improvements NVMe SSDs like DC1500M introduced vs SATA SSDs also means that introducing them in

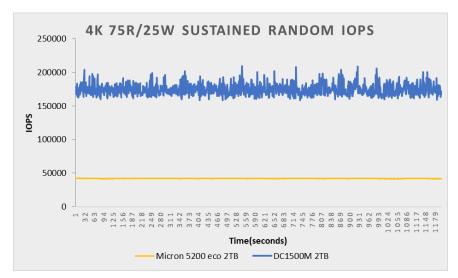


Figure a) Second-by-second IOPS comparison of DC1500M 1920G and Micron 5200 ECO 1920G SATA SSD. Tested on a single physical drive attached as a secondary to a linux system with fio v3.17 once the SSDs have reached a steady state of performance. Based on a block size of 4k, Read percentage of 75% and a Queue depth of 32

shared hyperconverged environments can also help reduce the impact of the Noisy Neighbor problem on tier 1 applications. Enterprise NVMe SSDs like DC1500M can complete unexpected workloads, like backup/restore operations during production hours at a much faster rate while still maintaining low latency and high transaction throughput for tier 1 mission critical production workloads as we show in the Noisy Neighbor tests later in this paper.

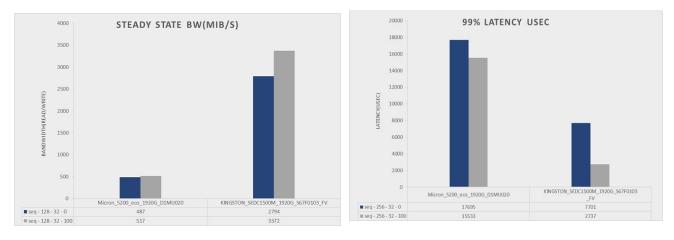


Figure b) Sequential BW(MB/s) Read/Write and Latency(usec) comparison DC1500M 1920G and Micron 5200 ECO 1920G SATA SSD. Tested on a single physical drive attached as a secondary to a linux system with fio v3.17 once the SSDs have reached a steady state of performnace. Based on a block size of 256k and a Queue depth of 32



Testing Environment

I. Infrastructure

Our testing environments are shown in Figure 1.1 and 1.2 below. We used VMware vSAN as our HCl of choice since it's a highly scalable, resilient, centralized, and cost-effective storage option for hyperconverged, virtualized environments.

VMware vSAN allows users to aggregate local storage devices from multiple servers into a single datastore shared between all hosts in the vSAN cluster. Physical disks from each server are placed into disk groups with 1 drive/disk group used as a cache device and up to 7 drives/disk group used as capacity devices. At most, a server can have up to 5 disk groups, so a max of 35 capacity devices/server contributing to the vSAN cluster. The disk groups of all ESXi hosts in a vSAN cluster are combined to create a vSAN datastore, with traffic between the hosts and the vSAN datastore isolated through a dedicated network for vSAN (10Gbps+ for all flash vSAN is a requirement). It allows administrators to start with a small storage footprint and add storage nodes to scale up capacity (up to 64 Nodes/cluster) as required and provides a relatively easy way to control performance requirements for specific VMs.

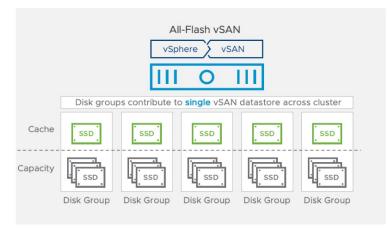


Figure 1 All flash vSAN architecture

vSAN uses storage policies to dictate the level of protection and striping for specific virtual disks. Using the default storage policy, vSAN mirrors all objects provisioned from the vSAN datastore, but it also provides administrators with granular control over the level of protection of the virtual disks provisioned to the VMs from the vSAN datastore. For example, to allow the SQL data drive VMDK to tolerate at least one failure in the cluster (entire server, disk, or network interface), we can specify a primary level of failure to

tolerate (FTT) of 1. A RAID-1 mirror of the VMDK object would then be created with one replica component on one host and another replica component on another host in the vSAN cluster. Likewise, administrators can specify a RAID 0 (striping only) storage policy with an FTT of 0 if we want the backup drive VMdk to have no resiliency and maximum performance; where the VM is highly available via SQL AlwaysOn Failover Clustering or if the database is regularly backed up via common backup solutions like Commvault or NetBackup.

In our Kingston Technology SSD testing and validation lab and for this paper, we used 3 <u>PowerEdge</u> <u>R740xD servers</u> supporting 8 2.5" NVMe and 16 2.5" SATA/SAS drive bays/server, with a dedicated 10Gb network supported by 2 <u>Cisco Nexus 5k switches</u> for vSAN traffic for SATA SSD testing. We used the 4 node Big Twin Supermicro <u>SYS-2029BT-HNR super server</u> with a dedicated 40Gb network supported by 1 <u>Cisco 9k switch</u> for vSAN traffic, for NVMe testing. In our testing, we used a custom storage policy (FTT=0) assigned to the Guest VM virtual disk to maximize block storage performance for all tests



conducted in this paper. For the various tests we conducted, we used different SSDs that are documented at the beginning of each test result below, but as a standard we used 3 physical drives with the same capacity per disk group for both SATA and NVMe testing. We selected the popular Micron 5200 eco SATA SSD for comparison testing. For management and VMotion traffic, we used a 1Gb network, supported by 1 Netgear JGS524PE 24-port managed switch.

NVMe testing environment (Hardware)	SATA/SAS/HYBRID testing environment (Hardware)	
Supermicro SYS-2029BT-HNR 4 Node Cluster with	PowerEdge Dell R740xD 3 Node Cluster	
6 Hot-swap 2.5" NVMe drive bays/server	supporting with 8 2.5" NVMe and 16 2.5" SATA/SAS drive bays/server	
Intel(R) Xeon(R) Gold 6252 CPU (48c/96t) @	Intel(R) Xeon(R) Silver 4114 CPU (10c/20t) @	
2.10GHz X 8	2.20GHz x8	
64x32GB Kingston DDR4-2933 2Rx4 ECC REG	768 GB 24x32GB Kingston Dual Rank ECC	
DIMM (16x32GB per node), 512GB/Node, 2048GB/cluster	Memory @ 2400MHz/Node, 2304GB/cluster	
2xCisco nexus N5K-C5010 20 port 10Gbe data center class switches for vSAN network traffic	1xCisco Nexus 9332PQ Switch 32 port 40Gbe data center class switch dedicated for vSAN network traffic	
	PERC H740P configured in HBA passthru mode	

Figure 1.1 Hardware used during our tests

NVMe testing environment (OS and Software)	SATA testing environment (OS and Software)
Hypervisor: VMware ESXi, 7.0.3, 19193900	Hypervisor: VMware ESXi, 7.0.3, 19193900
vSAN 7U3c (VMware ESXi, 7.0.3, 19193900 +	vSAN 7U3c (VMware ESXi, 7.0.3, 19193900 +
VMware VirtualCenter 7.0.3 build-19234570)	VMware VirtualCenter 7.0.3 build-19234570)
Guest OS: Windows Server 2019 Data center,	Guest OS: Windows Server 2019 Datacenter,
v1809	v1809
Microsoft SQL Server 2017 (RTM) - 14.0.1000.169	Microsoft SQL Server 2017 (RTM) - 14.0.1000.169
(X64)	(X64)
HammerDB-v3.2	HammerDB-v3.2
HCIBench 2.5.3	HCIBench 2.5.3

Figure 1.2: OS and Software

II. Database Configuration

In the tests conducted here, we used a Server 2019 Guest VM with SQL server 2017 and a separate VMDK provisioned from the vSAN datastore for Data, Log and Backup. HammerDB, which is an open source database load testing application that supports running the TPCC benchmark for OLTP applications and TPC-H benchmark for data analytics workload. Throughout the various tests in this paper, the TPCC benchmark specification is chosen here to simulate OLTP transactional workloads and ensure conformance and reliability of testing results.

The TPCC benchmark (formal definition available on tpc.org (TPCC home)), is a well-known industry standard OLTP benchmark that implements a computer system to fulfill orders from customers to



supply products from a company. The company sells 100,000 items and keeps its stock in warehouses. Each warehouse has 10 sales districts and each district serves 3000 customers. The customers call the company whose operators take the order, each order containing several items, then orders are usually satisfied from the local warehouse. However, a few items are not in stock at a point in time and are supplied by an alternative warehouse. It is important to note that the size of the company is not fixed and can add warehouses and sales districts as the company grows. For this reason, your test schema can be as small or large as you wish, with a larger schema resulting in a larger TPC-C database and requiring a more powerful computer system to process the increased level of transactions (HammerDB).

For this article, we run various tests with the number of warehouses (schema size) and number of virtual users documented at the beginning of each test and explained in the test results. Throughout all test runs, we record the Hammer DB results from each test run while simultaneously capturing CPU, network, memory, and disk statistics using Windows performance monitor (Perfmon), with the native module Get-counter in Windows PowerShell, and vSAN performance monitor available on vCenter server.

III. vSAN Storage Performance

We tested the performance of the vSAN datastore for the configurations that we focus on in this paper prior to running our SQL tests to assess the level of performance we can expect from the DC1500M NVMe and Micron 5200 eco SATA SSD vSAN datastore. We used VMware's recommended tool for benchmarking the vSAN datastore-<u>HCIBench v2.5.3</u>, which is an automation toolkit that deploys multiple VMs spread across all the hosts in the vSAN cluster while running specific workloads using Vdbench on all guest VMs in parallel. We present a few results from our run with 6 VMs on the DC1500M NVMe vSAN cluster and the Micron 5200 eco SATA cluster.

Figure 1.3 and 1.4 show the mixed workload results in a sustained 70% Read, 30% Write random workload with various block sizes for a duration of 30 minutes for the DC1500M NVMe vSAN datastore and the Micron 5200 eco SATA SSD vSAN datastore. At a block size of 4k, The DC1500M NVMe vSAN datastore could deliver 2X as many 70%R/30%W IOPS (355k vs 178K) as the SATA SSD vSAN datastore with each IO completing 33% faster (0.4ms vs 0.6ms for the SATA SSD vSAN). The NVMe performance advantage becomes clear as the IO transfer size increases. If you look at the 64k 70% Read, 30% Write random workload, the NVMe vSAN datastore could deliver 3x as much IOPS (121240 vs 31756) with 66% better latency per IO (2.1ms vs 6.4 ms for the SATA SSD vSAN).

Figure 1.5 and 1.6 show a comparison for the HCIBench sustained read and write throughput and latencies for the DC1500M NVMe and Micron 5200 eco SATA SSD vSAN datastore with various block sizes. We could sustain a throughput of 17.8GB/s (128k) from the DC1500M NVMe datastore, 6.3x the read throughput from the SATA SSD vSAN datastore (2.79GB/s) and 5x lower latency (0.9ms vs 4.4ms for SATA vSAN). For writes, DC1500M vSAN sustained a throughput of 6.7GB/s write (128k), also 5.9x higher than the SATA vSAN with 5x lower latency.

How much does this raw performance difference between the NVMe and SATA vSAN datastore scale when it comes to SQL performance? Does the performance advantage of NVMe justify the expense? Will SQL backup or restore operations complete faster to mitigate the impact on mission critical workloads? In the upcoming sections, we seek to answer this question by conducting a few experiments.



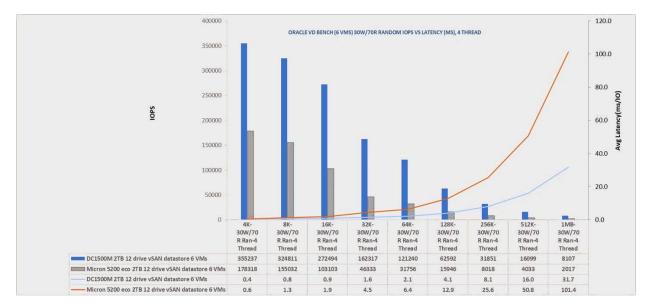


Figure 1.3 DC1500M vSAN datastore vs Micron 5200 eco vSAN datastore, 4k 70R/30W, Random, QD=8, threads=4, 6 VMs HCIBench IOPS vs average latency(ms)

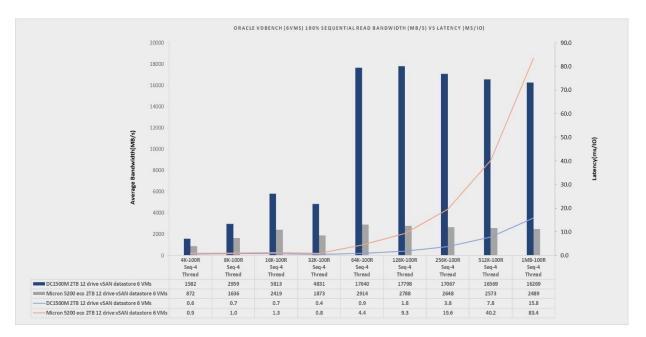


Figure 1.4 DC1500M vSAN datastore vs Micron 5200 eco vSAN datastore, 100R/0W, Sequential, QD=8, threads=4, HCIBench 6 VMs Read Throughput(MB/s) and Average Read Latency(ms/IO)



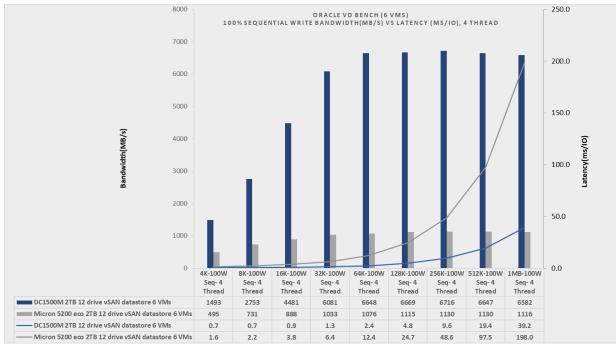


Figure 1.5 DC1500M vSAN datastore vs Micron 5200 eco vSAN datastore, 100W/0R, Sequential, QD=8, threads=4, HCIBench 6 VMs Read Throughput(MB/s) and Average Read Latency(ms/IO)

Testing Results

Test 1, DC1500M 960GB vSAN SQL server 2017 VM with varying amounts of DRAM

vSAN Datastore storage configuration: 3 DC1500M 960G FW S67F0103/disk group, 4 total disk groups (1 per server), NVMe vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS.

001		
Test 1a Description	Test 1b Description	Test 1c Description
Virtual disk provisioned from DC1500M vSAN datastore on the NVMe test environment. A 1200 warehouse database schema representing a 100GB database was chosen. System under test VM (SUT) was assigned 16vCores and 128GB RAM Another vSAN VM with 16c/128GB RAM was provisioned to act as a load generation server to send transactions to the SUT. Virtual User sequence created was 1,2,3,5,8,13,21,34,55,89. 2-minute ramp up time and 5 min/user sequence test duration were chosen.	Like Test 1a; but the allocated DRAM for the Guest VM was reduced to 32 GB to increase IO to the data area. A remote load generation server was still used to send transactions to the SUT, but the allocated DRAM for the LGS was also reduced to 32GB.	Like Test 1a; but the allocated DRAM for the Guest VM was reduced to 32 GB to increase IO to the data area and this test was run locally on the SUT VM to eliminate any network bottlenecks.

Figure 2.1 Test 1: DC1500M vSAN datastore different DRAM configurations



Our goal for test 1 was to get a baseline on the level of performance expected with the TPCC benchmark on SQL Server 2017 on VMware vSAN with an all flash DC1500M NVMe vSAN datastore, with varying amounts of memory allocated to SQL server. The idea behind varying the amount of DRAM allocated to the SQL System under test (SUT) is based on these concepts:

- Reductions in allocated RAM to the SQL server database VM will increase IO to the data area and place more emphasis on the I/O performance of the database containing the schema (ondisk OLTP database)
- If the SQL server database VM has enough DRAM, most of the data will be cached during an OLTP test and I/O to the data area will be minimal (in-memory OLTP test)

We created a schema size of 1200 warehouses, which resulted in a tpcc database size of ~100GB. In the first test, we allocated 128GB DRAM to the SUT, so the entire schema can fit in memory. Then we ran the virtual user sequence on a remote load generation server (LGS) to simulate users sending transactions to the database, scaling from 1-89 users to match our schema size and the amount of allocated CPU/memory resources to the SQL server VM. After the test completed, we restored the TPCC database, then reduced the allocated DRAM to 32GB on the SUT and the LGS and reran the same test with the same user sequence. Finally, we ran the same test locally on the system under test VM to eliminate any network bottlenecks introduced by the remote load generation server.

Test 1 Results, DC1500M 960GB vSAN SQL server 2017 VM with varying amounts of DRAM

Figure 2.2 and 2.3 shows the Transactions Per Minute (TPM) and New Orders Per Minute (NOPM) we achieved for Test 1a, 1b, 1c using the DC1500M vSAN datastore. For all test runs, we observe the TPM and NOPM scale up as the number of virtual users increases. At 89 virtual users, the SQL Server 2017 VM with a mostly in-memory OLTP database could achieve 1,113,300 TPM with 259,631 NOPM. When we reduced the DRAM allocated to 32GB on the SUT and LGS VM, we could achieve 958,338 TPM and 208311 NOPM, but when we ran the test locally on the SUT VM, we achieved a phenomenal 1,463,290 TPM and 318092 NOPM!

This is where we see the latency advantage of Enterprise NVMe SSDs in action. What this means is, when allocated insufficient memory to cache the schema, as the number of transactions increase and the SQL server database needs to write data from memory to the transaction log file, the NVMe virtual disk can respond fast enough to sustain the higher transaction throughput and scale up until the CPU becomes the bottleneck. From Figure 2.4, in test 1c, we can see that even at 89 virtual users, each user can process 16,441 transactions per minute. Based on these empirical results, we can conclude that building your database on NVMe hyperconverged infrastructure allows you to save cost on extra DRAM allocated to SQL Server 2017.



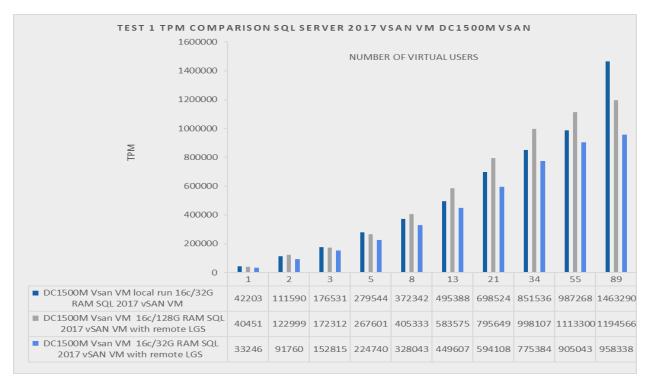


Figure 2.2 Test 1a,b,c: DC1500M vSAN datastore TPM comparison different DRAM size

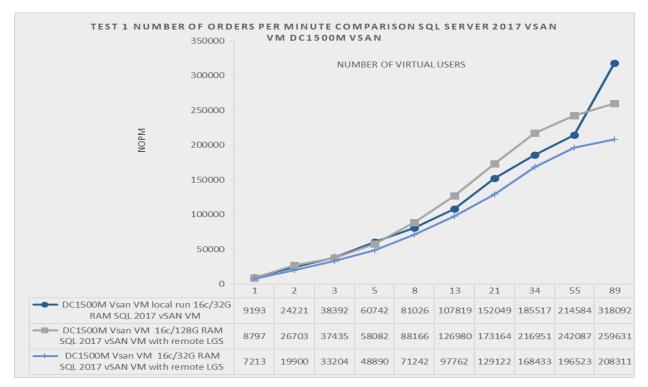


Figure 2.3 Test 1a,b,c: DC1500M vSAN datastore NOPM comparison different DRAM size



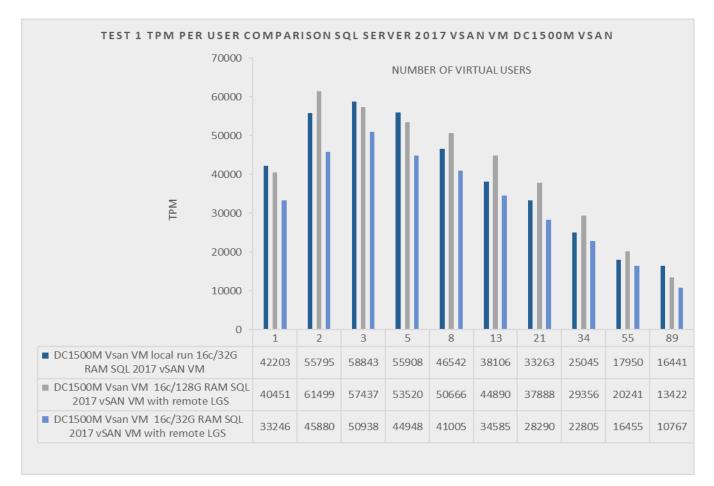


Figure 2.4 Test 1a,b,c: DC1500M vSAN datastore TPM comparison different DRAM size



Test 2: Comparing SQL Server 2017 performance on Kingston DC500M SATA SSD, Micron 5200 eco SATA SSD and DC1500M NVMe SSD vSAN datastore

- NVMe vSAN Datastore storage configuration for test 1a: 3 DC1500M 960G FW S67F0103/disk group, 4 total disk groups (1 per server), NVMe vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 1a)
- SATA vSAN Datastore storage configuration for test 1b: 3 DC500M 1920G FW SCEJK2.8/disk group, 3 total disk groups (1 per server), SATA vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 1b)
- SATA vSAN Datastore storage configuration for test 1c: 3 Micron 5200 ECO 1920G FW D1MU004/disk group, 3 total disk groups (1 per server), SATA vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 1b)

Test 2a Description	Test 2b Description	Test 2c Description
Virtual disk provisioned from DC1500M	Virtual disk provisioned from D500M	Virtual disk provisioned from Micron
vSAN datastore on the NVMe test	vSAN datastore on the SATA test	5200 eco vSAN datastore on the SATA
environment.	environment.	test environment.
A 1200 warehouse database schema	A 1200 warehouse database schema	A 1200 warehouse database schema
representing a 100GB database was	representing a 100GB database was	representing a 100GB database was
chosen. System under test VM (SUT)	chosen. System under test VM (SUT)	chosen. System under test VM (SUT)
was assigned 16vCores and 32GB RAM	was assigned 16vCores and 32GB RAM	was assigned 16vCores and 32GB RAM
Virtual User sequence created was	Virtual User sequence created was	Virtual User sequence created was
1,2,3,5,8,13,21,34,55,89.	1,2,3,5,8,13,21,34,55,89.	1,2,3,5,8,13,21,34,55,89.
2-minute ramp up time and 5 min/user	2-minute ramp up time and 5 min/user	2-minute ramp up time and 5 min/user
sequence test duration were chosen.	sequence test duration were chosen.	sequence test duration were chosen.
Test was run locally on the SUT VM.	Test was run locally on the SUT VM.	Test was run locally on the SUT VM.

Figure 3.1 Test 2 description: Comparing SQL Server 2017 performance on SATA and DC1500M NVMe SSD vSAN datastore

Test 2 compares the performance of the TPCC benchmark for SQL Server 2017 system under test VM, when run locally on 3 different datastores, Kingston DC1500M enterprise NVMe vSAN datastore, Kingston DC500M and Micron 5200 eco SATA SSD vSAN datastores. In test 2 we ran locally on the SQL Server 2017 VM System Under test, to increase I/O to the data area and emphasize IO performance of the database containing the schema, and a test a user sequence to scale from 1-89 users to match our schema size and the amount of allocated CPU/memory resources to the SQL server VM.

Test 2 Results: Comparing SQL Server 2017 performance on Kingston DC500M SATA SSD, Micron 5200 eco SATA SSD and DC1500M NVMe SSD vSAN datastore

Figure 3.2 and 3.3 shows the Transactions per minute (TPM) and New orders per minute (NOPM) we achieved for Test 2a, 2b and 2c. For all test runs, we observe the TPM and NOPM scale up as the number of virtual users increases, but the scaling is dramatically different for NVMe vs SATA. At 89 virtual users, the DC1500M backed vSAN datastore SQL Server 2017 VM could achieve 1,463,290 TPM with 318,092 NOPM. Comparatively, we achieved 738,067 TPM/160,410 NOPM for the DC500M SQL server vSAN VM and 628499 TPM/136436 NOPM for the Micron 5200 eco vSAN datastore. This means that using the same number of DC1500M NVMe drives, on an NVMe backed vSAN datastore, you can effectively double your transaction throughput and orders per minute when compared to a SATA backed vSAN datastore with the same number of SSDs. In a business context, if you have 89 users sending transactions to the database simultaneously, each user can process 235% more transactions (translating



into more orders per minute) (Figure 3.4) if you upgrade your VMware infrastructure to be backed by Enterprise NVMe solutions like DC1500M.

Figure 3.5 shows the average CPU idle time vs number of virtual users for test 2a, b, and c. This is an effective measure of the efficiency of the virtual disk— how fast the virtual disk can respond as the number of transactions increase and the SQL server database needs to write data from memory to the transaction log file. At 89 virtual users, our CPU idle time (iowait) for our DC1500M NVMe backed vSAN VM is 15.5% compared to 37.8% for the DC500M backed VM and 44.2% for Micron 5200 backed VM. This means that our NVMe virtual disk responds much faster to IO requests preventing the CPU from idling waiting for the IO to be completed and allowing for more transactions to be processed. In a business context, upgrading your VMware infrastructure to NVMe allows for more efficiently use assigned virtual cores to your SQL server VM to drive transaction throughput up and reduce cost by removing unnecessary cores from legacy SQL VMs running on slower storage tiers.

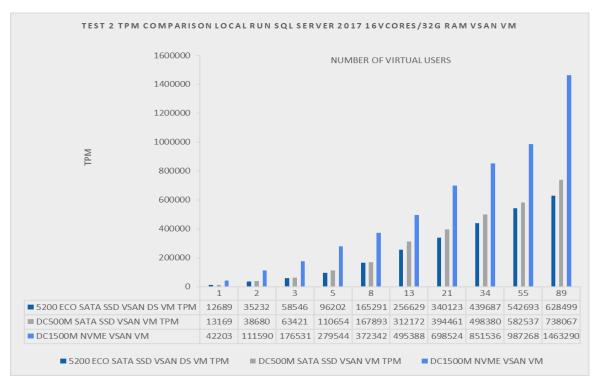


Figure 3.2 Test 2: TPM comparison NVME VS SATA VSAN datastore



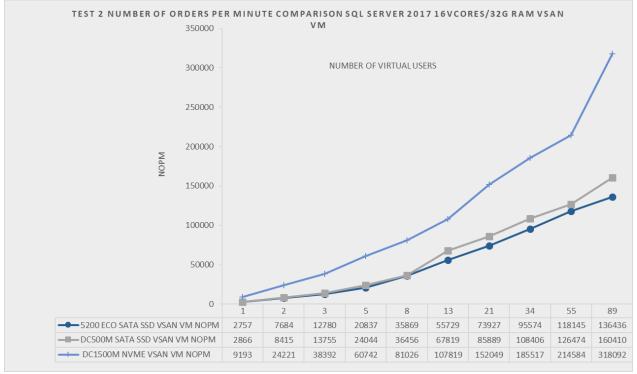


Figure 3.3 Test 2: NOPM comparison NVME VS SATA VSAN datastore

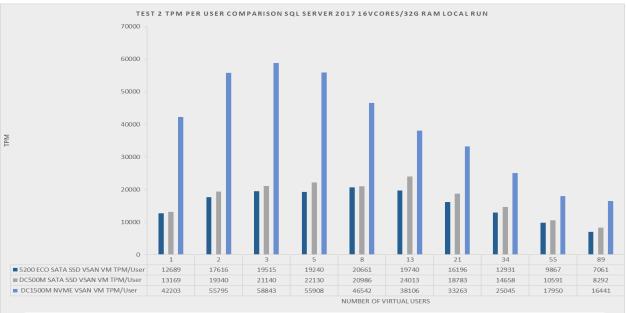


Figure 3.4 Test 2: TPM per user comparison NVME VS SATA VSAN datastore



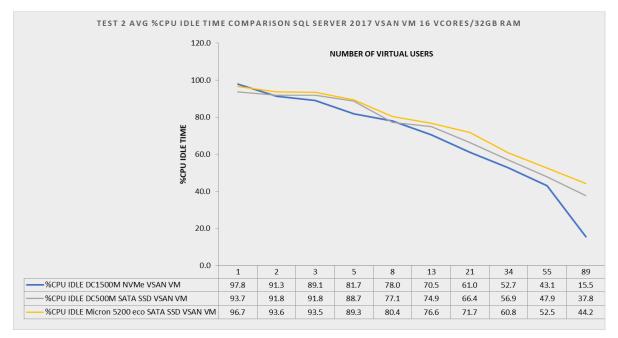


Figure 3.5 Test 2: %CPU idle comparison NVME VS SATA VSAN datastore

Test 3: SQL Server 2017 performance comparison DC1500M NVMe vs Micron 5200 eco SATA vSAN datastore, larger schema size and longer test duration

•	NVMe vSAN Datastore storage configuration for test 3a: 3 DC1500M 960G FW S67F0103/disk group, 4 total disk groups (1 per server), NVMe vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 3a)
•	SATA vSAN Datastore storage configuration for test 3b: 3 Micron 5200 ECO 1920G FW D1MU004/disk group, 3 total disk groups (1 per server), SATA vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 3b)

Test 3a Description	Test 2b Description	
Virtual disk provisioned from DC1500M vSAN datastore on	Virtual disk provisioned from Micron 5200 eco vSAN	
the NVMe test environment.	datastore on the SATA test environment.	
A 2000 warehouse database schema representing a 157GB	A 2000 warehouse database schema representing a 157GB	
database was chosen. System under test VM (SUT) was	database was chosen. System under test VM (SUT) was	
assigned 40vCores and 32GB RAM	assigned 40vCores and 32GB RAM	
Virtual User sequence created was 1,2,4,8,16,32,64,89,128	Virtual User sequence created was 1,2,4,8,16,32,64,89,128	
10-minute ramp up time and 20 min/user sequence test	10-minute ramp up time and 20 min/user sequence test	
duration were chosen.	duration were chosen.	
Test was run locally on the SUT VM.	Test was run locally on the SUT VM.	

Figure 4.1 Test 3 Description: SQL Server 2017 DB Stress test on Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore

This test was designed to be a longer duration stress test with a larger database schema size to validate our earlier results and compare the performance of the TPCC benchmark for SQL Server 2017 system under test VM, when run locally on 2 different datastores, Kingston DC1500M enterprise NVMe vSAN datastore and Micron 5200 eco SATA SSD vSAN datastore. This time we chose a schema size of 2000 warehouses which resulted in a TPC-C database size of 157 GB. We used 40 virtual cores for each SQL server VM to allocate enough CPU resources to generate more transactions and saturate the



transactional throughput, but only assigned 32 GB of RAM to make the test IO bound. We tuned the virtual user sequence slightly to scale up from 1-128 users and allowed each virtual user sequence to run for a much longer time (20 minutes, with a 10-minute ramp up time). This allowed us to collect disk latency metrics during the entire duration of the test run.

Test 3 results: SQL Server 2017 performance comparison DC1500M NVMe vs Micron 5200 eco SATA vSAN datastore, larger schema size and longer test duration

Figure 4.2 and 4.3 shows the Transactions Per Minute (TPM) and New Orders Per Minute (NOPM) we achieved for Test 3a and 3b. Even with a longer duration, both the SQL server 2017 VMs backed by NVMe and SATA SSDs could scale up as the number of virtual users increases to 128, but the gradient of scale is much higher for NVMe. At 89 users we achieved 1.84M TPM compared 0.96TPM and 361743 NOPM compared to 184451 NOPM for the SATA SSD backed vSAN SQL VM. This is a 200% increase in TPM/NOPM for the DC1500M NVMe backed vSAN datastore compared to the Micron 5200 eco vSAN backed VM, with the same number of vCores and allocated DRAM.

Figure 4.4 and 4.5 shows a comparison for the Avg virtual disk latency and 99% virtual disk latency vs number of users collected using windows perfmon on the SQL NVMe and SATA SSD backed vSAN VMs. For each virtual user sequence, the DC1500M backed virtual disk can maintain <1ms avg latency even as the number of users continued to scale up. At 89 virtual users the DC1500M backed virtual disk had an average latency of 0.92ms/IO compared to 2.36ms/IO for the SATA SSD backed vdisk-a 256% increase in average latency compared to NVMe. What's more interesting is the QoS 99% latency-at 89 users the DC1500M virtual disk could complete 99% of all IOs in 1.61ms but the SATA SSD backed vdisk completed 99% of all IOs in 7.05ms-a 437% increase compared to NVMe. The latency difference between NVMe and SATA is highlighted here, and because the DC1500M is engineered to maintain predictable QoS latency throughout sustained OLTP workloads, we do not see any sudden spikes in latency, even as the number of virtual users increases which translates into more parallel IO requests on the block layer. From a business standpoint, this means that upgrading your VMware infrastructure from SATA SSDs to Enterprise NVMe drives like DC1500M allows you to scale up transactions and lower transaction latency drastically, allowing applications to scale rapidly and reducing cost over time.



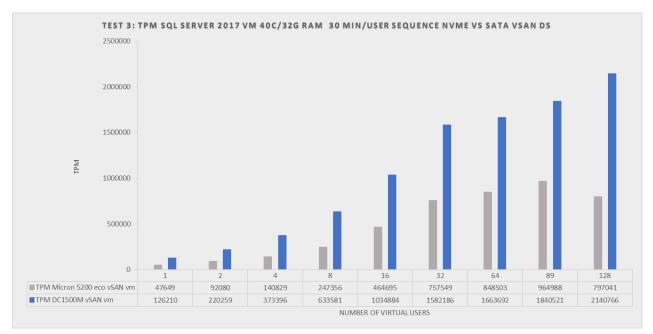


Figure 4.2 Test 3 TPM comparison SQL Server 2017 DB stress test on Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore

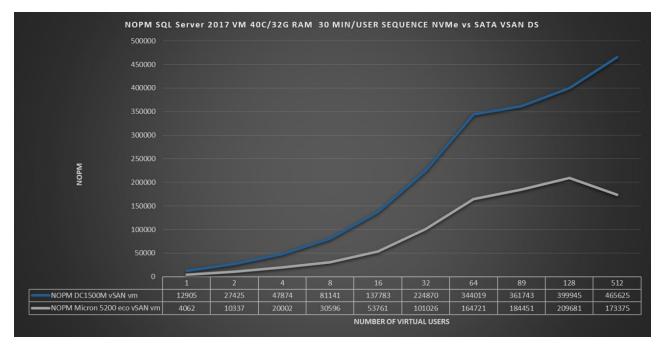


Figure 4.3 Test 3 NOPM comparison SQL Server 2017 DB stress test on Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore



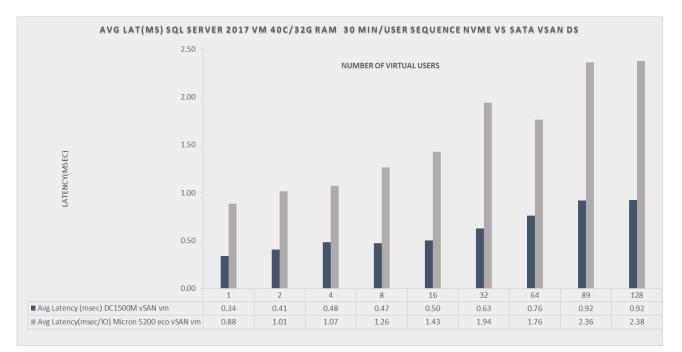


Figure 4.4 Test 3 Avg Latency(ms) comparison SQL Server 2017 DB stress test on Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore

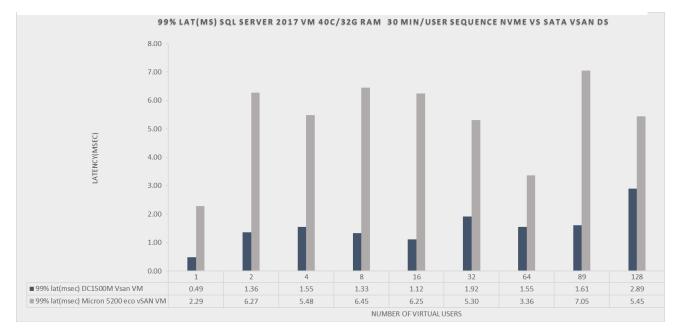


Figure 4.5 Test 3 99th % Latency(ms) comparison SQL Server 2017 DB stress test on Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore



Test 4: SQL Server 2017 performance comparison, backup and restore performance, DC1500M NVMe vs Micron 5200 eco SATA vSAN

- NVMe vSAN Datastore storage configuration for test 3a: 3 DC1500M 960G FW S67F0103/disk group, 4 total disk groups (1 per server), NVMe vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 4a)
- SATA vSAN Datastore storage configuration for test 3b: 3 Micron 5200 ECO 1920G FW D1MU004/disk group, 3 total disk groups (1 per server), SATA vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 4b)

Test 4a Description	Test 4b Description
Virtual disk provisioned from DC1500M vSAN datastore on	Virtual disk provisioned from Micron 5200 eco vSAN
the NVMe test environment.	datastore on the SATA test environment.
A 2000 warehouse database schema representing a 157GB	A 1200 warehouse database schema representing a 157GB
database was created on SUT. System under test VM (SUT) was assigned 16vCores and 32GB RAM	database was created on SUT. System under test VM (SUT) was assigned 16vCores and 32GB RAM
3 cycles of a backup/restore script were triggered to backup and restore the tpcc database and performance metrics recorded with windows performance monitor	3 cycles of a backup/restore script were triggered to backup and restore the tpcc database and performance metrics recorded with windows performance monitor
Test was run locally on the SUT VM.	Test was run locally on the SUT VM.

Figure 5.1 Test 4 description: SQL Server 2017 Backup/restore performance comparison on Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore

SQL Database backup and restore operations are a good way to measure the throughput and latency of the underlying virtual disk. We wanted to establish a baseline on throughput and latency metrics from a single NVMe backed and a SATA backed vSAN VM by capturing virtual disk metrics with windows performance monitor when the TPC-C backup/restore operations are triggered.

Test 4: Results: SQL Server 2017 performance comparison, backup and restore performance, DC1500M NVMe vs Micron 5200 eco SATA vSAN

Figure 5.2-5.4 shows the second-by-second throughput and latency collected by our windows performance monitor script for one of the backup/restore cycles for test 4a) and test 4b). The SQL server VM backed by DC1500M NVMe vSAN datastore completed the TPCC database backup operation in 265 seconds, achieving an average throughput of 593MB/s and an average latency of 1.46ms/IO. The TPCC database restore operation completed in 129 seconds, with an average BW of 1.4GB/s and an average latency of 2.65ms/IO. Comparing that to the Micron 5200 eco vSAN backed VM, the backup operation completed 1.5x faster and the restore operation completed 2.15x faster on the NVMe vSAN backed SQL VM.

Typically, backup and restore operations are done off hours to avoid any impact to production VMs. However, this is not always the case. If SQL backup or restore operations are done during peak business hours, you want them to complete as fast as possible to avoid the latency impact on users performing transactions on the tier 1 application sharing the same vSAN datastore. Migrating your SQL databases to NVMe backed vSAN datastores, allows you to absorb that impact. Even if the backup/restore operations are done off hours, completing them faster allows for less downtime for tier 1 databases sharing the same resources.



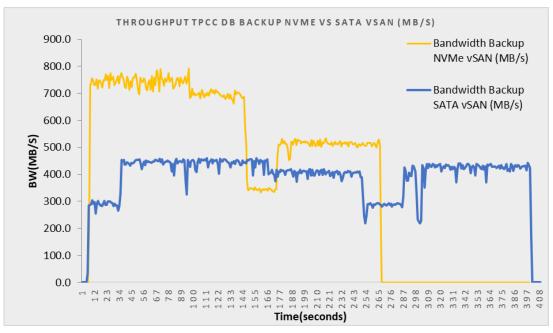


Figure 5.2 Throughput comparison SQL Server 2017 TPCC DB Backup Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore(MB/s)

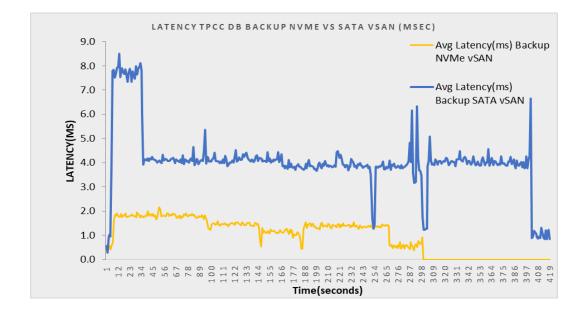


Figure 5.3 Avg Latency(ms) comparison SQL Server 2017 TPCC DB Backup Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore



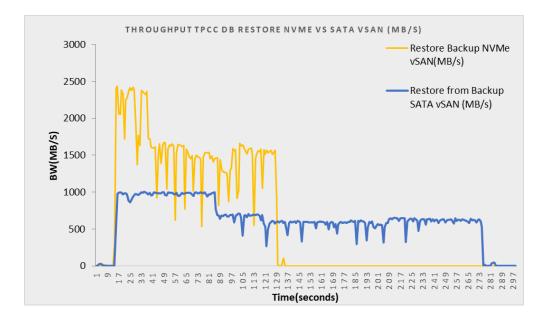


Figure 5.5 Latency(ms) comparison SQL Server 2017 TPCC DB Restore Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore

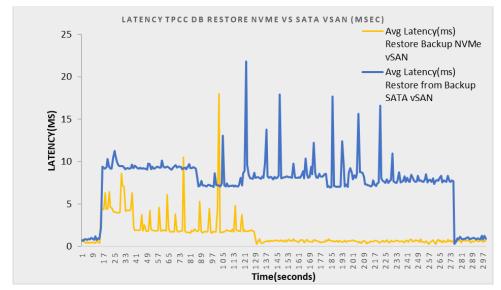


Figure 5.4 Throughput comparison SQL Server 2017 TPCC DB Restore Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore(MB/s)



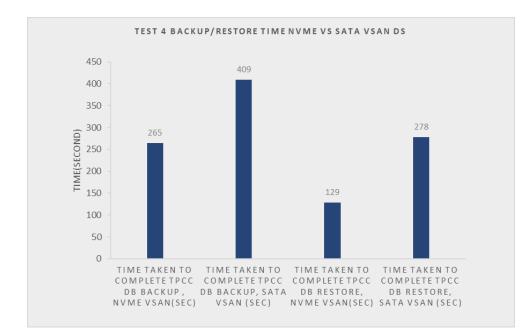


Figure 5.6 Time taken to complete SQL Server 2017 TPCC DB Backup/Restore operation Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore(sec)



Test 5: SQL Server 2017 performance comparison, The Noisy Neighbor test, DC1500M NVMe vs Micron 5200 eco SATA vSAN

- NVMe vSAN Datastore storage configuration for test 3a: 3 DC1500M 960G FW S67F0103/disk group, 4 total disk groups (1 per server), NVMe vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 5a)
- SATA vSAN Datastore storage configuration for test 3b: 3 Micron 5200 ECO 1920G FW D1MU004/disk group, 3 total disk groups (1 per server), SATA vSAN testing environment. SQL Server 2017 with Server 2019 Datacenter Guest OS. (Test 5b)

Guest 65: (Test 56)			
Test 5a Description	Test 5b Description	Test 5c Description	Test 5d Description
SQL 2017 VM Virtual disk	SQL 2017 Virtual disk	SQL 2017 VM Virtual disk	SQL 2017 VM Virtual disk
provisioned from DC1500M	provisioned from Micron	provisioned from DC1500M	provisioned from Micron
vSAN datastore on the	5200 eco vSAN datastore on	vSAN datastore on the	5200 eco vSAN datastore on
NVMe test environment.	the SATA test environment.	NVMe test environment.	the SATA test environment.
A 1200 warehouse database	A 1200 warehouse database	A 1200 warehouse database	A 1200 warehouse database
schema representing a	schema representing a 100GB	schema representing a	schema representing a
100GB database was created	database was created on SUT.	100GB database was	100GB database was
on SUT. System under test	System under test VM (SUT)	created on SUT. System	created on SUT. System
VM (SUT) was assigned	was assigned 16vCores and	under test VM (SUT) was	under test VM (SUT) was
16vCores and 32GB RAM	32GB RAM	assigned 16vCores and	assigned 16vCores and
The SUT under test was	The SUT under test was	32GB RAM	32GB RAM
cloned 11 times, and 3 SUT	cloned 8 times, and 3 SUT	The SUT under test was	The SUT under test was
VMs/physical server were	VMs/physical server were	cloned 11 times, and 2	cloned 8 times, and 2
assigned (12 SUT VMs in	assigned (9 SUT VMs in total)	VMs/physical server were	VMs/physical server were
total)	Test was configured to run 89	assigned (8 SUT VMs in	assigned (6 SUT VMs in
Test was configured to run	Virtual User with a	total) to run the HDB	total) to run the HDB
89 Virtual User with a	30-minute ramp up time and	workload. Test was	workload. Test was
30-minute ramp up time and	300 min test duration were	configured to run 89 Virtual	configured to run 89 Virtual
300 min test duration were	chosen on each SUT VM	User with a	User with a
chosen on each SUT VM	The test was triggered on all 9	30-minute ramp up time	30-minute ramp up time
The test was triggered on all	SUT VMs in parallel	and 300 min test duration	and 300 min test duration
12 SUT VMs in parallel		were chosen on each SUT	were chosen on each SUT
		VM.	VM.
		1 VM/physical server had a	1 VM/physical server had a
		1200 warehouse tpcc	1200 warehouse tpcc
		schema size (100GB), and a	schema size (100GB), and a
		backup script was triggered	backup script was triggered
		every 100 seconds (4 VMs	every 100 seconds (4 VMs
		total) while the workload	total) while the workload
		was running on the other	was running on the SUT VM
		SUT VMs for 10 cycles	6 SUT VMs running HDB
		8 SUT VMs running HDB	workload; 3 VMs running
		workload; 4 VMs running	backup script.
		backup script.	
		The test was triggered on all	The test was triggered on all
		12 VMs in parallel	9 VMs in parallel

Figure 6.1 Test 5 description: SQL Server 2017 Realistic Noisy Neighbor test on Micron 5200 eco SATA and DC1500M NVMe SSD vSAN datastore

Our goal with this test was to simulate a realistic scenario where abhorrent workloads (in this case we use TPCC database backup operations) on VMs that are sharing the same vSAN datastore as SQL server



VMs running production workloads (In this experiment, the TPCC benchmark is acting as a production workload) and assess the overall performance impact by assessing the TPCC benchmark results and analyzing key storage metrics, collected from perfmon and the vSAN performance monitor.

In test 5a) and 5b) we establish a baseline by running the TPCC benchmark on all VMs in parallel, with no backup operations occurring. We use 3 SQL VMs per physical server to run on both the NVMe and SATA vSAN clusters, bringing the total to 12 SUT VMs for NVMe and 9 SUT VMs for SATA. Our schema size for this test was 1200 warehouses, translating to a TPC-C database size of ~100GB and we ran the TPCC workload with 89 users for 300 minutes and a 30-minute ramp up time.

In test 5c) and 5d) we restored the TPC-C database on all SUT VMs. Then we triggered a script to execute 10 backup cycles of the TPC-C database on 4 VMs for the NVMe cluster and 3 VMs for the SATA cluster, while simultaneously running the same TPC-C benchmark on the remaining SUT VMs. This means that on the NVMe vSAN cluster, 8 VMs were running the TPC-C workload and 4 VMs were running the backup workload in parallel. Meanwhile, on the SATA vSAN cluster, 6 VM were running the TPC-C workload and 3 VMs were running the TPC-C database backup workload in parallel.

Test 5 results: SQL Server 2017 performance comparison, The Noisy Neighbor test, DC1500M NVMe vs Micron 5200 eco SATA vSAN

Figure 6.2 and 6.3 show the Transactions per minute (TPM) and New orders per minute (NOPM) we achieved for Test 5a and 5b. With 89 virtual users running on each on 12 DC1500M NVMe vSAN datastore backed SQL server VMs, we could achieve an average of 523,516 TPM and an average NOPM of 113,812 per VM, compared to an average of 269,320TPM and 58544 NOPM per VM with 9 SQL VMs backed by the Micron 5200 eco SATA cluster. Looking the IOPS and Latency metrics collected from the vSAN performance monitor (Figure 6.4 and 6.5 below), the resulting IO on the block layer translated to 120,000 Read IOPS, 60,000 Write IOPS on the NVMe cluster, with a latency of 800µs for read/write operations, and 50,000R/20,000W on the SATA vSAN cluster, with read latency average of 3.8ms and a write latency average of 5.5ms. This again highlights the performance difference between NVMe and SATA and showcases the ability of DC1500M NVMe backed virtual disk to absorb parallel requests and process them in a much faster round trip latency.

Figure 6.5 and 6.6 show the Transactions per minute (TPM) and New orders per minute (NOPM) we achieved for Test 5c and 5d. With 89 virtual users running on each on 8 DC1500M NVMe vSAN datastore backed SQL server VMs, while VM backups were triggered in parallel on 4 VMs, we could achieve an average of 575,933 TPM and an average NOPM of 125,206, compared to an average of 351,258 TPM and 76355 NOPM with 6 SQL VMs running the TPCC workload, while VM backups were triggered in parallel on 3 VMs on the SATA vSAN SQL VMs backed by the Micron 5200 eco SATA. To tell the full story, we must analyze the latency and storage metrics from both the SATA and NVMe vSAN cluster, as well as look at how quickly the backups completed on both clusters.

Figure 6.8 and 6.9 shows the vSAN IOPS and latency metrics collected from the NVMe and SATA cluster using vSAN performance monitor for test 5c and 5d. The backup script was configured to run every 100 seconds for 10 cycles. We can see the impact the triggered VM backups have on the IOPS and read and write latency of both the NVMe and SATA vSAN cluster. However, the impact on latency varies. The NVMe cluster maximum read/write IO latency spiked to 4ms/IO, while sustaining an average of 2.5



ms/IO for read/write operation, while the SATA vSAN spiked to 9ms/IO, and sustained an average of 7.3 ms/IO for read and 4.9 ms/IO for write IO. This is the latency that the end user will feel when they're trying to submit an order, update their shopping cart, or view products from other warehouses.

Figure 6.11 shows the time taken to complete the backup cycles on one of the SQL Server DC1500M vSAN backed VMs, and one of the Micron 5200 eco vSAN backed SQL VM, excluding the wait time between backup cycles. It took 73 minutes to complete 10 backup-an average of 7 min/backup the SQL server NVMe vSAN VM and 122.15 minutes to complete 10 backups for SQL server SATA SSD backed vSAN VM-an average of 12 minutes/backup. The DC1500M vSAN backed VM completed the backup cycles 1.67x faster than the Micron 5200 eco vSAN backed VM. This is empirical evidence that upgrading your VMware infrastructure to DC1500M NVMe backed datastores helps mitigate the Noisy Neighbor problem by allowing for unwanted operations like database backups to complete much faster, and due to the tremendous latency and throughput capability, NVMe can absorb the latency impact these abhorrent workloads have on tier 1 applications.

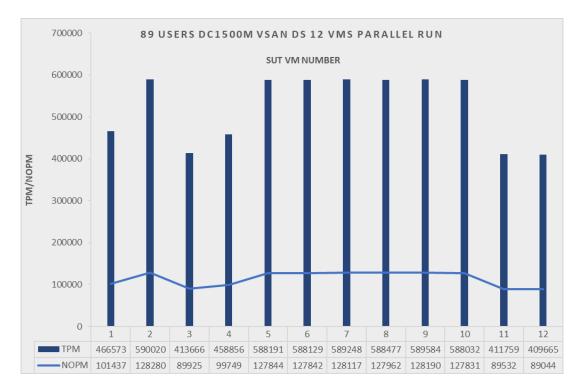


Figure 6.2 Test 5a TPM SQL Server 2017 300 min 12 VM parallel run, 89 virtual users, DC1500M NVMe SSD vSAN datastore



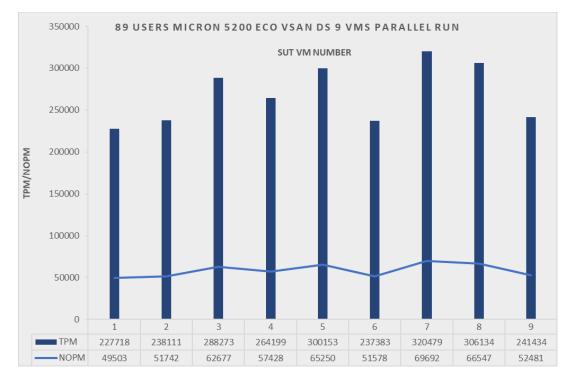


Figure 6.3 Test 5b TPM SQL Server 2017 300 min 12 VM parallel run, 89 virtual users, DC1500M NVMe SSD vSAN datastore

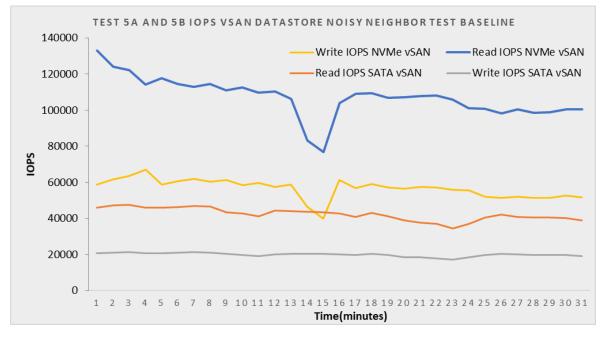


Figure 6.4 Test 5a and 5b Noisy Neighbor IOPS, DC1500M NVMe and Micron 5200 eco vSAN datastore



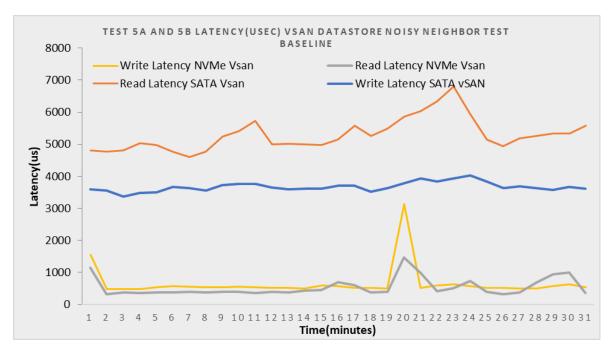


Figure 6.5 Test 5a and 5b Noisy Neighbor latency, DC1500M NVMe and Micron 5200 eco vSAN datastore

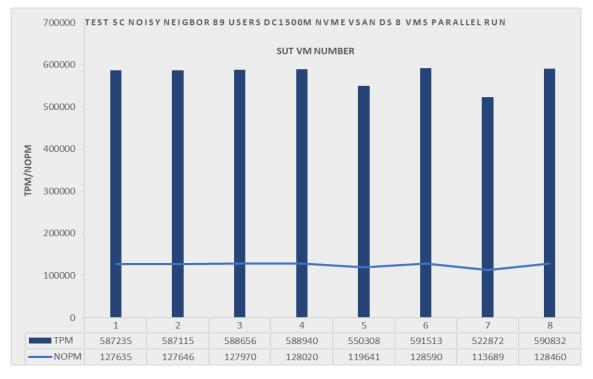


Figure 6.6 Test 5c TPM, Noisy Neighbor implementation 8 VM parallel run DC1500M NVMe vSAN datastore



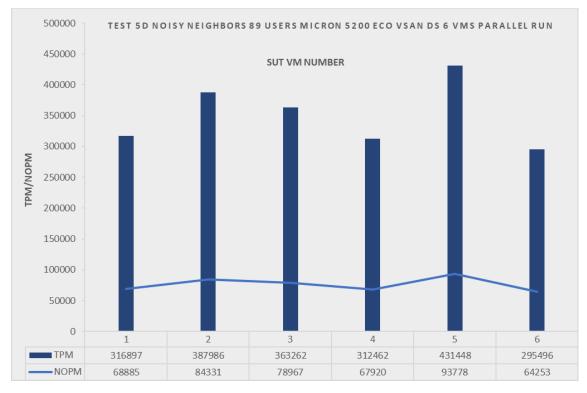


Figure 6.7 Test 5D TPM, Noisy Neighbor implementation 6 VM parallel run Micron 5200 eco vSAN datastore

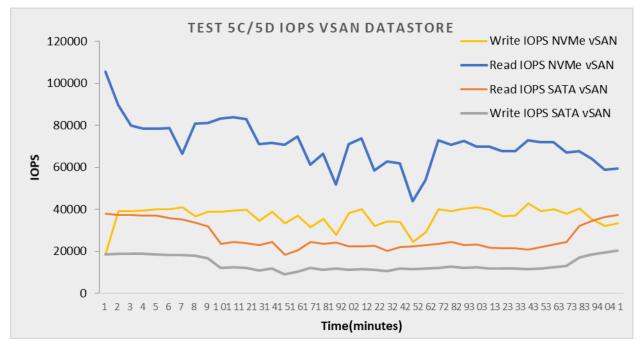


Figure 6.8 Test 5C/5D IOPS, Noisy Neighbor implementation NVMe vs SATA SSD vSAN datastore



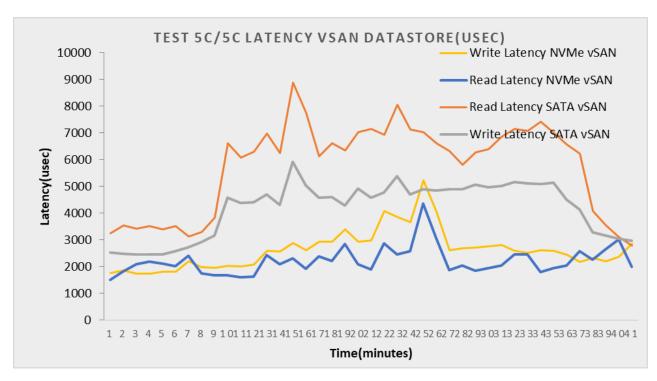


Figure 6.9 Test 5C/5D Latency, Noisy Neighbor implementation NVMe vs SATA SSD vSAN datastore

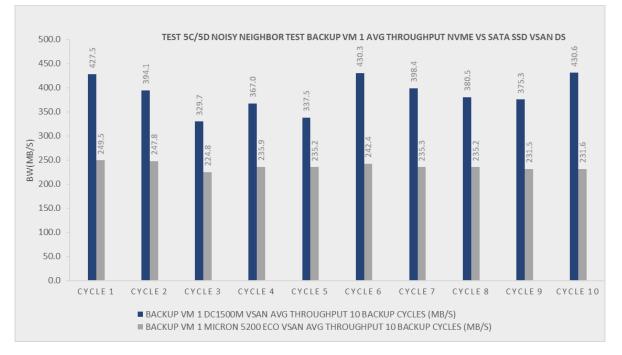


Figure 6.10 Test 5C/5D Backup VM throughput, Noisy Neighbor implementation NVMe vs SATA SSD vSAN datastore



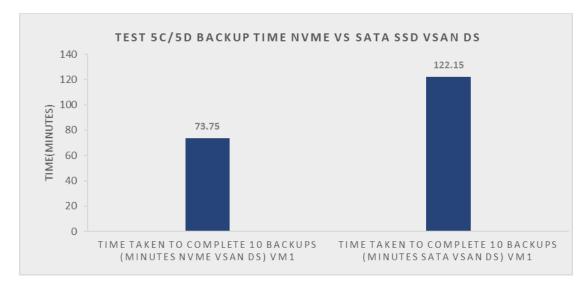


Figure 6.11 Test 5C/5D Backup VM time taken to complete backup, 10 cycles Noisy Neighbor implementation NVMe vs SATA SSD vSAN datastore

Conclusion

In this white paper, we showed how consolidating your database workloads to NVMe can help maximize existing hardware, due to its incredible efficiency and near 0 IO wait times, which allows you to use fewer CPU cores to achieve the same transactional throughput. We provided a few comparisons to Enterprise SATA SSDs and showed that by migrating your SQL workloads to an NVMe backed datastore, you can allow your applications to scale up as you double your transaction throughput while providing sub-msec latency. Then, we showed how NVMe can help mitigate impact to tier 1 applications by allowing unwanted workloads, like database backup/restore operations to complete faster.

Kingston's Enterprise NVMe SSDs, <u>DC1500M</u> paired with Kingston Server Memory (Server Premier) provides an excellent solution for users looking to virtualize their database infrastructure and maximize their workload efficiencies.

Visit <u>https://www.kingston.com/en/solutions/servers-data-centers</u> to learn more about Kingston's data center solutions

References

HammerDB. (n.d.). Understanding the TPCC workload. Retrieved from https://www.hammerdb.com/docs3.3/ch03s05.html

TPCC home. (n.d.). Retrieved from https://www.tpc.org/

