



利用金士顿 DC1500M 企业级 NVMe 固态硬盘和 VMware vSAN HCI 降低总拥有成本并构建富有弹性、响应极快的数据库

作者：Hazem Awadallah，Kingston Technology 系统工程师

审阅者：Chris Selden，Kingston Technology 固态硬盘产品工程经理



目录

- [执行摘要](#)
- [当今数据中心中 RDBMS 面临的常见基础架构挑战](#)
- [解决方案：金士顿科技数据中心 DC1500M 企业级 NVMe 固态硬盘简介](#)
- [测试环境](#)
 - [I. 基础架构](#)
 - [II. 数据库配置](#)
 - [III. vSAN 存储性能](#)
- [测试结果](#)
 - [测试 1：DC1500M 960GB vSAN SQL server 2017 虚拟机，配置不同大小的 DRAM](#)
 - [测试 1 结果：DC1500M 960GB vSAN SQL server 2017 虚拟机，配置不同大小的 DRAM](#)
 - [测试 2：金士顿 DC500M SATA 固态硬盘、Micron 5200 eco SATA 固态硬盘和 DC1500M NVMe 固态硬盘三种 vSAN 数据存储的 SQL Server 2017 性能比较](#)
 - [测试 2 结果：金士顿 DC500M SATA 固态硬盘、Micron 5200 eco SATA 固态硬盘和 DC1500M NVMe 固态硬盘三种 vSAN 数据存储的 SQL Server 2017 性能比较](#)
 - [测试 3：DC1500M NVMe 和 Micron 5200 eco SATA 两种 vSAN 数据存储的 SQL Server 2017 性能比较，更大的架构尺寸和更长的测试时长](#)
 - [测试 3 结果：DC1500M NVMe 和 Micron 5200 eco SATA 两种 vSAN 数据存储的 SQL Server 2017 性能比较，更大的架构尺寸和更长的测试时长](#)
 - [测试 4：SQL Server 2017 性能比较，备份与还原性能，DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN](#)
 - [测试 4：结果：SQL Server 2017 性能比较，备份与还原性能，DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN](#)
 - [测试 5：SQL Server 2017 性能比较，近邻干扰测试，DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN](#)
 - [测试 5 结果：SQL Server 2017 性能比较，近邻干扰测试，DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN](#)
- [结论](#)

执行摘要

近年来，NVMe 的问世彻底改变了数据存储领域，在最大限度提高 NAND 闪存性能并利用功能丰富、低成本、高带宽和面向未来的扩展总线标准 PCI express 方面实现重大飞跃。作为第 5 代即最新一代 PCIe 标准，PCIe Gen5 支持高达每通道 8GB/秒的传输速率，不仅能消除存储堆栈的扩展总线瓶颈，还推动了固态硬盘控制器和 NAND 闪存乃至整个硬件堆栈的创新与发展。处理器、机箱设计、主板和硬件 IO 拓扑结构不断发展，为更高的带宽提供支持。在数据中心，网络拓扑结构经历着重大调整以适应 NVMe；随着 NVMe-OF 规范的出现，网络接口、交换机和传输协议出现变化并持续改善，从而支持更高的带宽，同时维持 QoS 和无损包传送。

但是，NVMe 的问世如何影响应用性能？您可以减少存储占用空间，同时提高事务吞吐量和缩短事务响应时间吗？我们可以大幅缩短数据库备份时间，以规避近邻干扰问题并最大限度降低其对生产环境的影响吗？在本文中，我们尝试通过分析由 TPCC 规范定义的典型 OLTP 工作负载解答这些问题，并提供多项实际对比来展示 NVMe 在现实情景中对事务性能的影响。

当今数据中心中 RDBMS 面临的常见基础架构挑战

成本、容量规划与可扩展性

随着二十年来互联网带宽、处理速度和数据分析的蓬勃发展，生产 OLTP 数据库迅速增长，常常比应用和基础架构架构师的规划快得多。底层的存储和网络架构的构建必须从一开始就支持扩展，以随着时间的推移满足更高的需求，并在成本、易管理性和性能之间实现良好平衡。是选择在本地数据中心构建应用，还是使用 IaaS/PaaS 服务构建，这将成为困难的设计决定。在本地数据中心运行应用，可让解决方案架构师全面控制可扩展性、安全性、弹性和性能，但这需要精心规划，有时要投入相当大的前期费用。利用 IaaS/PaaS 云服务，可加速部署并简化可扩展性，但会降低对性能和弹性的控制，并可能随着应用的扩展而变得成本高昂。一些组织会选择混合方法，将更重要的第 1 层应用放在本地数据中心，并将第 2 层应用和旧应用迁移到云中。对于留在组织内部的应用，采用全闪存磁盘组的 VMware vSAN 等超融合基础架构 (HCI) 解决方案在成本、简易性、性能和易扩展性方面能够实现良好平衡。

弹性

第 1 层应用必须构建到或迁移到一个能经受住整个硬件堆栈范围内多次硬件故障的基础架构中。如果规划不当，数据中心的设备故障可能造成服务中断，在最糟情况下会导致数据永久丢失，从而带来重大金钱损失。在共享式存储环境下，必须进行周密规划，确保底层基础架构能够经受存储故障和组件性能超载。

例如，对于 vSAN，第 1 层应用的最小允许故障数 (FTT) 应为 1，在启用 vSphere High Availability (HA) 的情况下，可确保应用和数据库虚拟机能够防范至少一次计算机、网络或存储故障。此外，还可随后启用 vSphere Distributed Resource Scheduler (DRS)，以在集群内的所有物理服务器之间加载平衡的 CPU/内存资源。

不断改变的性能预期

随着 OLTP 应用持续向上扩展、更多用户将更高事务负载放在后端数据库，对更高事务处理速度和更低延迟的需求持续增长。应用架构师必须制定规划，让存储基础架构能够进行调整来支持这种更高的需求，并足够灵活地在不同的存储分层之间进行迁移。例如，使用 VMware 的 Storage VMotion，从 SAN 存储阵列设置的虚拟磁盘中的 SQL 数据库可以迁移到采用 NVMe 等更快存储分层的 NVMe 全闪存 vSAN 数据存储。

近邻干扰困境

至关重要的是要设计可让关键工作负载拥有执行所需资源的基础架构。在包含多个工作负载的共享式存储环境中，性能可能变得无法预测，异常的工作负载可能给关键生产工作负载造成问题。这就是近邻干扰问题的定义。正如我们在本白皮书稍后部分谈到的一个例子，一台服务器上的计划外数据库备份操作会消耗存储和网络资源，并影响到使用相同资源的其他服务器的性能和延迟。

金士顿 DC1500M 企业级 NVMe 固态硬盘简介

[金士顿 DC1500M](#) 是金士顿推出的最新企业级 U.2 PCIe 3.0x4 NVMe，提供 960GB-7680G 不等的存储容量。它采用一个 16 通道控制器和 3D TLC NAND，按照严格的服务质量 (QoS) 要求设计，可确保企业级工作负载具备持续的高性能和一致性，同时维持最低延迟。它的固件以企业为中心，支持多项特性，例如预留空间、多命名空间（支持多达 64 个命名空间）以及更先进的 ECC 算法，可在整个硬盘使用寿命期间确保企业级工作负载的可靠性。

考虑到 SATA 固态硬盘仍是数据中心最主流的固态硬盘，我们在本文中旨在展示，通过迁移到或构建基于金士顿 DC1500M NVMe 等企业级 NVMe 固态硬盘的存储基础架构，将有助于缓解上文提到的一些问题。

在我们的内部测试中，一个金士顿 DC1500M NVMe 固态硬盘的吞吐量和延迟性能分别是一个 Micron 5200 eco 企业级 SATA 固态硬盘的 6.5 倍和 5.6 倍，而两者成本相差无几。

在超融合环境中，这种性能水平意味着可让 SQL Server 数据库提高事务吞吐量和降低延迟。这还可以减小存储占用空间并降低功耗。在本例中，需要 6 个 micron 5200 eco 硬盘才抵得上 1 个 DC1500M 硬盘的吞吐量。我们稍后将了解这种性能如何影响 VMware vSAN 中的实际 SQL OLTP 工作负载。

DC1500M 等 NVMe 固态硬盘的性能较 SATA 固态硬盘大幅提升，这也意味着将其引入共享式超融合环境，还有助于降低近邻干扰对第 1 层应用的影响。DC1500M 等企业级 NVMe 固态硬盘能以快得多的速度处理意外的工作负载，例如生产期间的备份/还原操作，同时让第 1 层任务关键型生产工作负载维持低延迟和高事务吞吐量，如下文的近邻干扰测试所示。

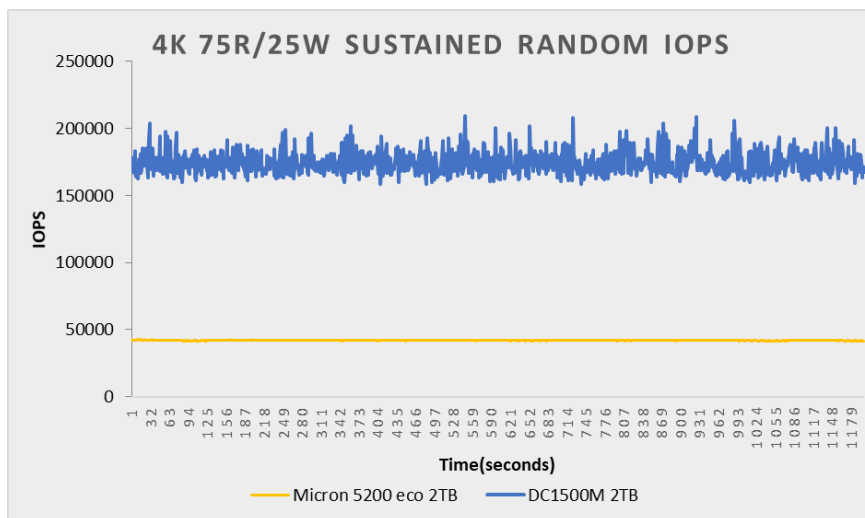


图 a) DC1500M 1920G 与 Micron 5200 ECO 1920G SATA 固态硬盘的逐秒 IOPS 比较。将一个物理硬盘作为辅助硬盘连接到安装 fio v3.17 的 linux 系统，待固态硬盘达到稳定的性能状态后进行了测试。采用 4K 块大小、75% 的已读百分比和 32 的队列深度

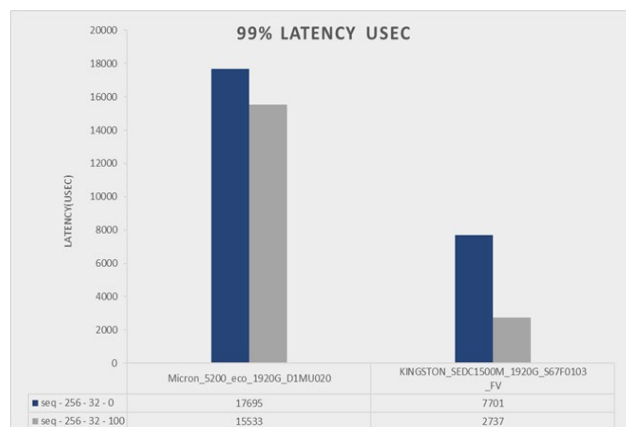
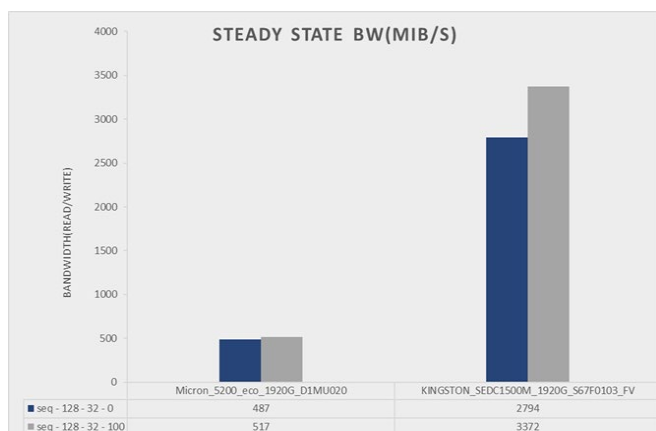


图 b) DC1500M 1920G 与 Micron 5200 ECO 1920G SATA 固态硬盘的顺序带宽 (MB/秒) 读取/写入和延迟 (usec) 比较。将一个物理硬盘作为辅助硬盘连接到安装 fio v3.17 的 linux 系统，待固态硬盘达到稳定的性能状态后进行了测试。采用 256k 块大小和 32 的队列深度

测试环境

1. 基础架构

我们的测试环境显示在下文的图 1.1 和 1.2 中。我们当然选择 VMware vSAN 作为我们的 HCI，因为它是适合超融合虚拟化环境的具备高度可扩展性、弹性、集中性和成本效益的存储选项。

利用 VMware vSAN，用户可将来自多个服务器的本地存储设备聚合到由 vSAN 集群中的所有主机共享的单个数据存储。来自每个服务器的物理磁盘被放置到磁盘组中，其中每个磁盘组有 1 个硬盘用作缓存设备，且每个磁盘组有多达 7 个硬盘用作容量设备。一台服务器最多可有 5 个磁盘组，因此，每台服务器最多 35 个容量设备用于 vSAN 集群。一个 vSAN 集群中所有 ESXi 主机的磁盘组合并组成一个 vSAN 数据存储，其中主机与 vSAN 数据存储之间的流量仅流进面向 vSAN 的专用网络（要求全闪存 vSAN 为 10Gbps+）。这不仅让管理员可以一开始采用较小的占用空间，并根据要求添加存储节点来向上扩展容量（每集群最多 64 个节点），还提供了相对简单的方式来控制特定虚拟机的性能要求。

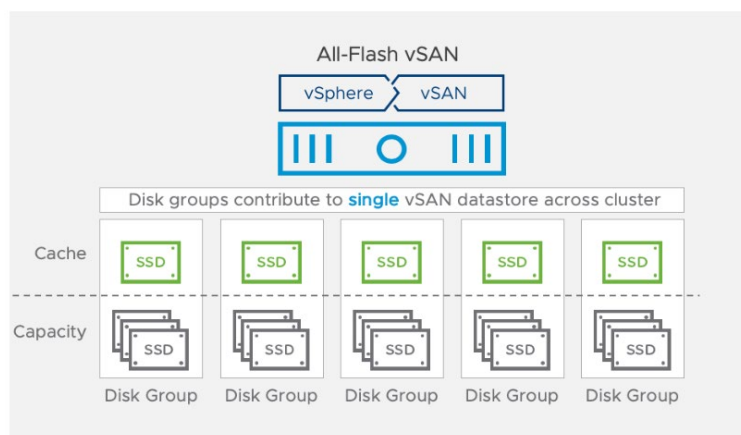


图1 全闪存vSAN 架构

vSAN 使用存储策略来决定特定虚拟磁盘的保护和条带化水平。利用默认的存储策略，vSAN 映射从 vSAN 数据存储设置的所有对象，但也支持管理员精细地控制从 vSAN 数据存储给虚拟机设置的虚拟磁盘的保护水平。例如，要让 SQL 数据驱动器 VMDK 经受至少一次来自集群的故障（整个服务器、磁盘或网络接口），我们可以指定基础允许故障数 (FTT) 为 1。然后，使用 vSAN 集群中一个主机上的一个副本组件和另一个主机上的另一个副本组件来创建 VMDK 的 RAID-1 镜

像。同样地，如果我们希望备份驱动器 VMDK 的弹性为零、性能最高，管理员可以指定 RAID 0（仅条带化）存储策略并将 FTT 设为 0；若使用 SQL AlwaysOn Failover Clustering，或如果数据库通过 Commvault 或 NetBackup 等常见备份解决方案定期进行备份，那么虚拟机高度可用。

针对本白皮书，在金士顿科技固态硬盘测试和验证实验室中，我们在 SATA 固态硬盘测试中使用了 3 台 [PowerEdge R740xD 服务器](#)（支持 8 个 2.5" NVMe 和 16 个 2.5" SATA/SAS 驱动器槽/服务器），以及一个用于 vSAN 流量的专用 10Gb 网络（由 2 台 [Cisco Nexus 5k 交换机](#) 提供支持）。我们为 NVMe 测试使用了 4 节点 Big Twin Supermicro [SYS-2029BT-HNR super server](#)，以及一个用于 vSAN 流量的专用 40Gb 网络（由 1 台 [Cisco 9k 交换机](#) 提供支持）。在测试中，我们使用了分配给客户机虚拟机虚拟磁盘的定制存储策略 (FTT=0)，以在本白皮书执行的所有测试中最大限度提高块存储性能。对于我们执行的各种测试，我们使用了不同的固态硬盘（如下文各测试结果的开头部分所述），但作为标准，我们为 SATA 和 NVMe 测试均使用了 3 个物理硬盘，每个磁盘组的容量相同。我们选择了热门的 Micron 5200 eco SATA 固态硬盘用于比较测试。对于管理和 VMotion 流量，我们使用了由 1 台 Netgear JGS524PE 24 口管理型交换机支持的 1Gb 网络。

NVMe 测试环境（硬件）	SATA/SAS/混合型测试环境（硬件）
Supermicro SYS-2029BT-HNR 4 节点集群，包含 6 个热交换 2.5" NVMe 驱动器槽/服务器	PowerEdge Dell R740xD 3 节点集群，支持 8 个 2.5" NVMe 和 16 个 2.5" SATA/SAS 驱动器槽/服务器
Intel(R) Xeon(R) Gold 6252 CPU (48c/96t) @ 2.10GHz X 8	Intel(R) Xeon(R) Silver 4114 CPU (10c/20t) @ 2.20GHz x8
64x32GB 金士顿 DDR4-2933 2Rx4 ECC REG DIMM（每节点 16x32GB），512GB/节点，2048GB/集群	768 GB 24x32GB 金士顿双列 ECC 内存 @ 2400MHz/节点，2304GB/集群
2 台 Cisco nexus N5K-C5010 20 口 10Gbe 数据中心级交换机，用于 vSAN 网络流量	1 台 Cisco Nexus 9332PQ 32 口 40Gbe 数据中心级交换机，专用于 vSAN 网络流量
	PERC H740P，配置为 HBA 通过模式

图 1.1 我们在测试期间使用的硬件

NVMe 测试环境（操作系统和软件）	SATA 测试环境（操作系统和软件）
虚拟机监控程序：VMware ESXi, 7.0.3, 19193900	虚拟机监控程序：VMware ESXi, 7.0.3, 19193900
vSAN 7U3c (VMware ESXi, 7.0.3, 19193900 + VMware VirtualCenter 7.0.3 build-19234570)	vSAN 7U3c (VMware ESXi, 7.0.3, 19193900 + VMware VirtualCenter 7.0.3 build-19234570)
客户机操作系统：Windows Server 2019 Data center, v1809	客户机操作系统：Windows Server 2019 Datacenter, v1809
Microsoft SQL Server 2017 (RTM) - 14.0.1000.169 (X64)	Microsoft SQL Server 2017 (RTM) - 14.0.1000.169 (X64)
HammerDB-v3.2	HammerDB-v3.2
HCIBench 2.5.3	HCIBench 2.5.3

图 1.2：操作系统和软件

II. 数据库配置

在这里执行的测试中，我们使用了一个 Server 2019 客户机虚拟机，其中搭载 SQL server 2017 和一个从 vSAN 数据存储设置的独立 VMDK，用于数据、日子和备份。HammerDB 是一款开源的数据库负载测试应用，支持运行面向 OLTP 应用的 TPCC 基准测试和面向数据分析工作负载的 TPC-H 基准测试。在本白皮书的不同测试中，我们选用 TPCC 基准测试规范来模拟 OLTP 事务型工作负载，并确保测试结果的一致性与可靠性。

TPCC 基准测试（正式定义可从 tpc.org（TPCC 主页）找到）是众所周知的行业标准 OLTP 基准测试，可实现一个计算机系统，用于执行来自客户的订单并提供来自公司的产品。公司销售 100,000 个物品，并将存货保存在仓库中。每个仓库有 10 销售区域，每个销售区域服务 3000 名客户。客户致电公司，公司操作人员接受订单，每个订单包含多个物品，然后通常从本地仓库履行订单。不过，有时一些物品会缺货，就会由替代仓库提供。需要注意的是，公司规模不是固定不变，可能会随着公司的发展而增加仓库和销售区域。由于这个原因，您的测试架构根据需要可大可小，

较大的架构会产生较大的 TPC-C 数据库，并需要更强大的计算机系统处理更高水平的事务量 (HammerDB)。

在本文中，我们对仓库数量（架构大小）和虚拟用户数量运行了各种测试，这些数据被记录在每项测试的开始部分，并在测试结果部分做了解释。在所有测试运行中，我们会记录每项测试运行的 Hammer DB 结果，同时利用 Windows 性能监测程序 (Perfmon)、Windows PowerShell 的本机模块 Get-counter 以及 vCenter server 中可用的 vSAN 性能监测程序记录 CPU、网络、内存和磁盘统计数据。

III. vSAN 存储性能

在运行我们的 SQL 测试前，我们针对本白皮书关注的配置测试了 vSAN 数据存储的性能，以评估可从 DC1500M NVMe 和 Micron 5200 eco SATA 固态硬盘两种 vSAN 数据存储获取的性能水平。我们使用了 VMware 推荐的 vSAN 数据存储基准测试工具-[HCIBench v2.5.3](#)，这是一个自动化工具箱，可跨 vSAN 集群中的所有主机部署多个虚拟机，同时所有客户机虚拟机中使用 Vdbench 并行运行特定工作负载。我们呈现了利用 DC1500M NVMe vSAN 集群和 Micron 5200 eco SATA 集群上的 6 个虚拟机运行测试所得的一些结果。

图 1.3 和 1.4 显示了 DC1500M NVMe vSAN 数据存储和 Micron 5200 eco SATA 固态硬盘 vSAN 数据存储持续 70% 读取、30% 写入随机工作负载下的混合工作负载结果，其中采用各种不同的块大小、测试持续 30 分钟。块大小为 4k 时，DC1500M NVMe vSAN 数据存储提供的 70% 读/30% 写 IOPS（355k 和 178K）是 SATA 固态硬盘 vSAN 数据存储的 2 倍，每次 IO 完成速度快 33%（0.4ms 与 SATA 固态硬盘 vSAN 的 0.6ms）。随着 IO 传输大小的增加，NVMe 的性能优势变得显而易见。对于 64k 70% 读取、30% 写入随机工作负载，NVMe vSAN 数据存储的 IOPS 提高 3 倍（121240 与 31756），每 IO 的延迟缩短 66%（2.1ms 与 SATA 固态硬盘 vSAN 的 6.4 ms）。

图 1.5 和 1.6 显示了采用不同块大小时 DC1500M NVMe 和 Micron 5200 eco SATA 固态硬盘两种 vSAN 数据存储的 HCIBench 持续读取和写入吞吐量及延迟对比情况。我们可以利用 DC1500M NVMe 数据存储维持 17.8GB/秒 (128k) 的吞吐量，其读取吞吐量是 SATA 固态硬盘 vSAN 数据存储的 6.3 倍（2.79GB/秒），延迟是后者 1/5（0.9ms 与 SATA 固态硬盘 vSAN 的 4.4ms）。写入方面，DC1500M vSAN 维持了 6.7GB/秒的写入吞吐量 (128k)，是 SATA vSAN 的 5.9 倍，延迟是后者的 1/5。

如果转换为 SQL 性能，NVMe 和 SATA 两种 vSAN 数据存储之间的这种原始性能差距会有什么变化？NVMe 的性能优势能证明成本的合理性吗？SQL 备份或还原操作能否更快地完成，以规避对任务关键型工作负载造成的影响？在后续章节中，我们寻求通过实施一些试验回答这些问题。

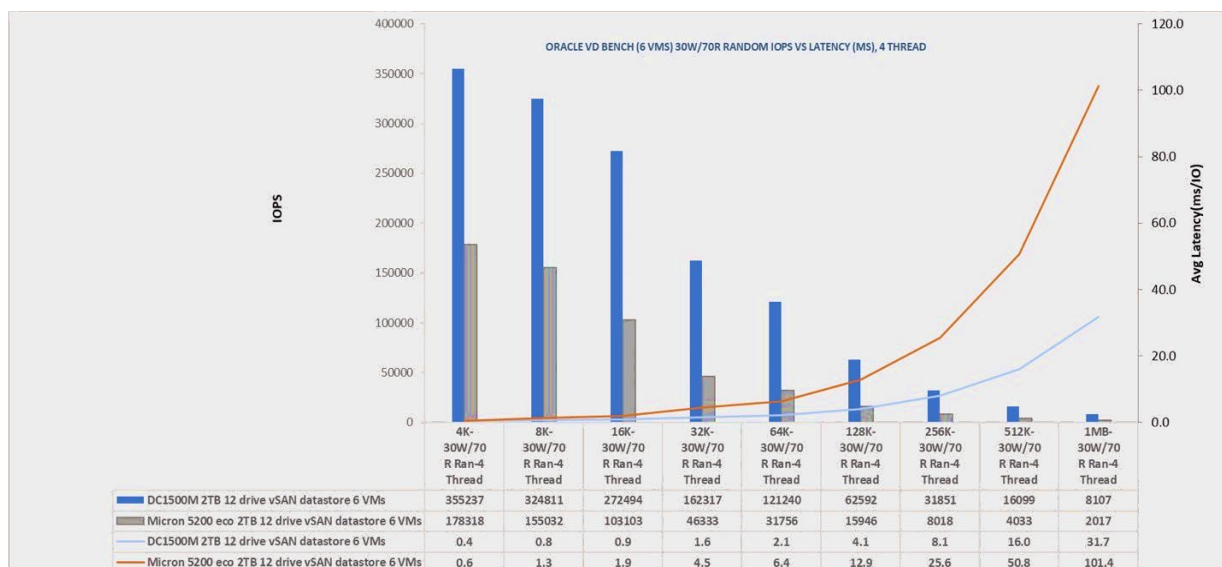


图1.3 DC1500M vSAN 数据存储与 Micron 5200 eco vSAN 数据存储, 4k 70 读取/30 写入, 随机, QD=8, 线程=4, 6 个虚拟机, HClBench, IOPS 与平均延迟 (ms)

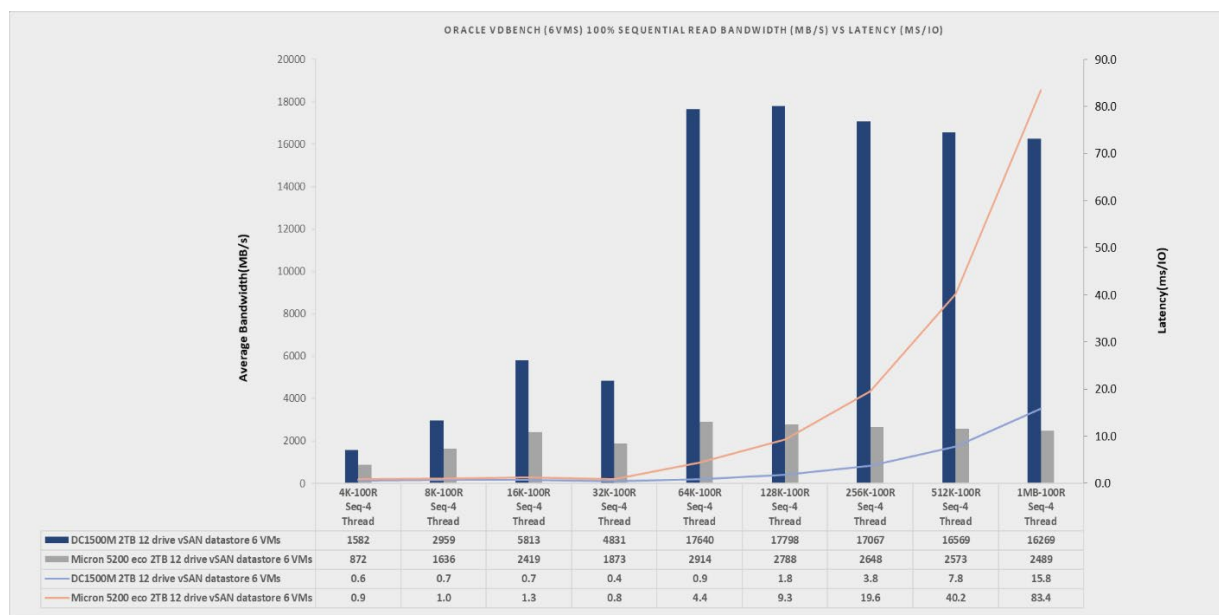


图1.4 DC1500M vSAN 数据存储与 Micron 5200 eco vSAN 数据存储, 100 读取/0 写入, 顺序, QD=8, 线程=4, HClBench, 6 个虚拟机, 读取吞吐量 (MB/秒) 和平均读取延迟(ms/IO)

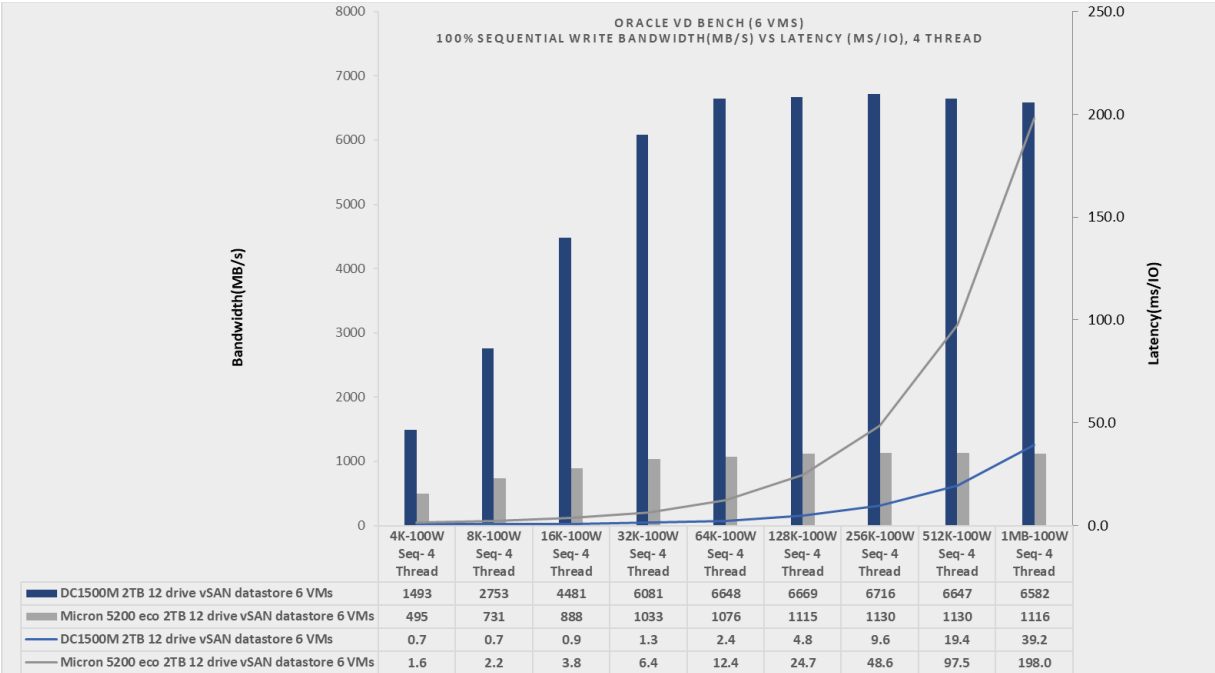


图 1.5 DC1500M vSAN 数据存储与 Micron 5200 eco vSAN 数据存储，100 写入/0 读取，顺序，QD=8，线程=4，HClBench，6 个虚拟机，读取吞吐量（MB/秒）和平均读取延迟(ms/IO)

测试结果

测试 1：DC1500M 960GB vSAN SQL server 2017 虚拟机，配置不同大小的 DRAM

vSAN 数据存储的存储配置：3 个 DC1500M 960G FW S67F0103/磁盘组，共 4 个磁盘组（每个服务器 1 个），NVMe vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。		
测试 1a 说明	测试 1b 说明	测试 1c 说明
NVMe 测试环境中从 DC1500M vSAN 数据存储配置的虚拟磁盘。 选择了一个包含 1200 个仓库的数据库架构，用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 128GB RAM 还配置了另一个 16 核/128GB RAM 的 vSAN 虚拟机，用作负载生成服务器向 SUT 发送事务。 创建的虚拟用户序列为 1、2、3、5、8、13、21、34、55、89。 选择了 2 分钟提升时间和 5 分钟/用户的序列测试时长。	类似测试 1a；但为客户机虚拟机分配的 DRAM 降低至 32 GB，以提高到数据区的 IO。仍然使用了一个远程负载生成服务器 (LGS) 将事务发送到 SUT，但为此 LGS 分配的 DRAM 也降低为 32GB。	类似测试 1a；但为客户机虚拟机分配的 DRAM 降低至 32 GB，以提高到数据区的 IO，并且本测试在 SUT 虚拟机本地运行以消除任何网络瓶颈。

图 2.1 测试 1：DC1500M vSAN 数据存储不同的 DRAM 配置

测试 1 的目标是对采用全闪存 DC1500M NVMe vSAN 数据存储的 VMware vSAN 上的 SQL Server 2017 进行 TPCC 基准测试，并获取其性能水平的基准数据，其中为 SQL server 分配了不同的内存大小。改变为 SQL 受测系统 (SUT) 分配的 DRAM 大小，背后的想法基于以下概念：

- 通过降低向 SQL server 数据库虚拟机分配的 RAM，可提高数据区的 IO，并更加重视包含架构的数据库的 I/O 性能（磁盘上 OLTP 数据库）
- 如果 SQL server 数据库虚拟机拥有足够 DRAM，那么多数数据会在 OLTP 测试期间得到缓存，并且到数据区的 I/O 会最小化（内存中 OLTP 数据库）

我们创建了一个包含 1200 个仓库的架构，这生成了约 100GB 的 tpcc 数据库。在第一项测试中，我们向 SUT 分配了 128GB DRAM，因此整个架构可以放入内存中。然后，我们在远程负载生成服务器 (LGS) 上运行虚拟用户序列，以模拟用户向数据库发送事务，从 1 名用户扩展到 89 名，以匹配我们的架构大小以及分配给 SQL server 虚拟机的 CPU/内存资源量。测试完成后，我们还还原了 TPCC 数据库，然后将 SUT 和 LGS 上分配的 DRAM 降低至 32GB，并使用相同的用户序列重新运行了相同的测试。最后，我们在受测系统虚拟机上本地运行了相同测试，以消除远程负载生成服务器带来的任何网络瓶颈。

测试 1 结果：DC1500M 960GB vSAN SQL server 2017 虚拟机，配置不同大小的 DRAM

图 2.2 和 2.3 显示了我们在测试 1a、1b、1c 使用 DC1500M vSAN 数据存储实现的每分钟事务量 (TPM) 和每分钟新订单量 (NOPM)。在所有测试运行中，我们观测到 TPM 和 NOPM 随着虚拟用户数量的增加而向上扩展。89 名虚拟用户时，主要采用内存中 OLTP 数据库的 SQL Server 2017 虚拟机可以实现 1,113,300 TPM 和 259,631 NOPM。当我们将 SUT 和 LGS 虚拟机上分配的 DRAM 降低为 32GB 时，我们可以实现 958,338 TPM 和 208,311 NOPM，但如果我们在 SUT 虚拟机上本地运行测试，可以实现惊人的 1,463,290 TPM 和 318,092 NOPM！

这正是我们看到企业级 NVMe 固态硬盘发挥延迟优势的地方。这意味着，当分配不足的内存来缓存架构时，随着事务量的增加和 SQL server 数据库需要将数据从内存写入事务日志文件，NVMe 虚拟磁盘可以足够快地做出响应，以维持更高的事务吞吐量并向上扩展，直到 CPU 成为瓶颈。图 2.4 显示，在测试 1c 中，我们可以看到，即便虚拟用户数量达到 89 名，每名用户都可以每分钟处理 16,441 项事务。根据这些实证结果，我们可以得出结论：通过在 NVMe 超融合基础架构上构建数据库，可以节省向 SQL Server 2017 分配更多 DRAM 而产生的费用。

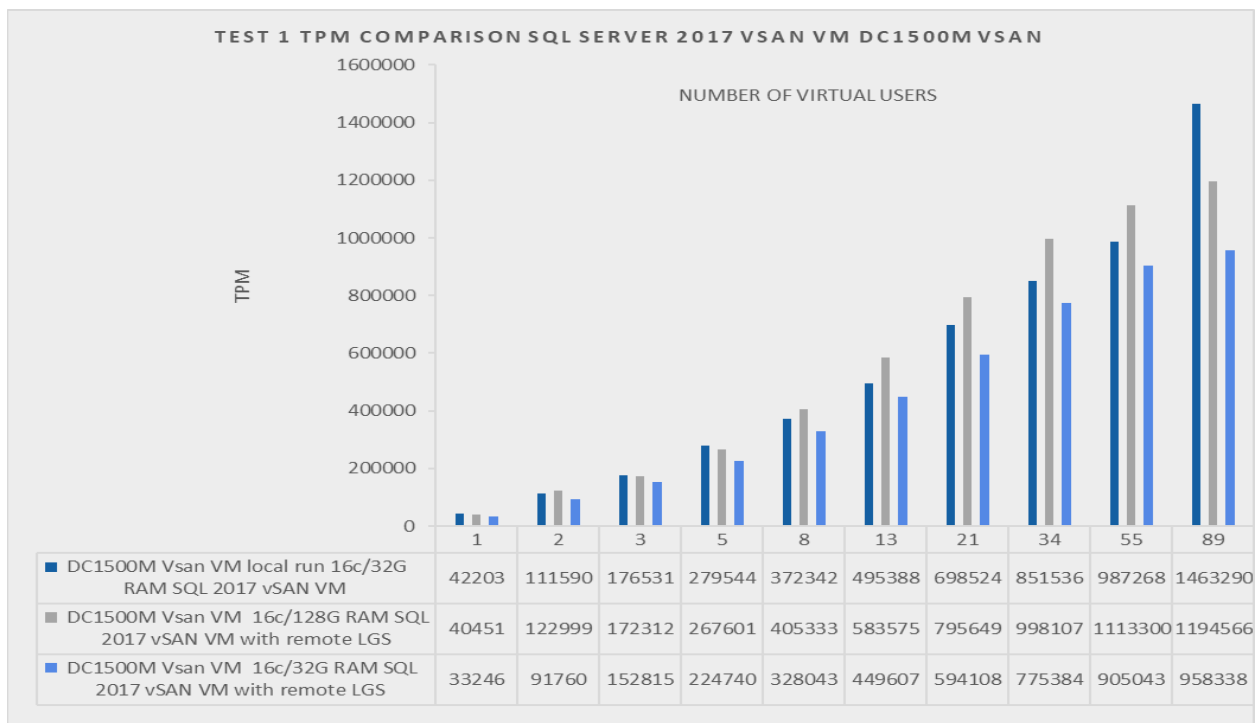


图2.2 测试1a、b、c: DC1500M vSAN 数据存储TPM 比较, 不同的DRAM 大小

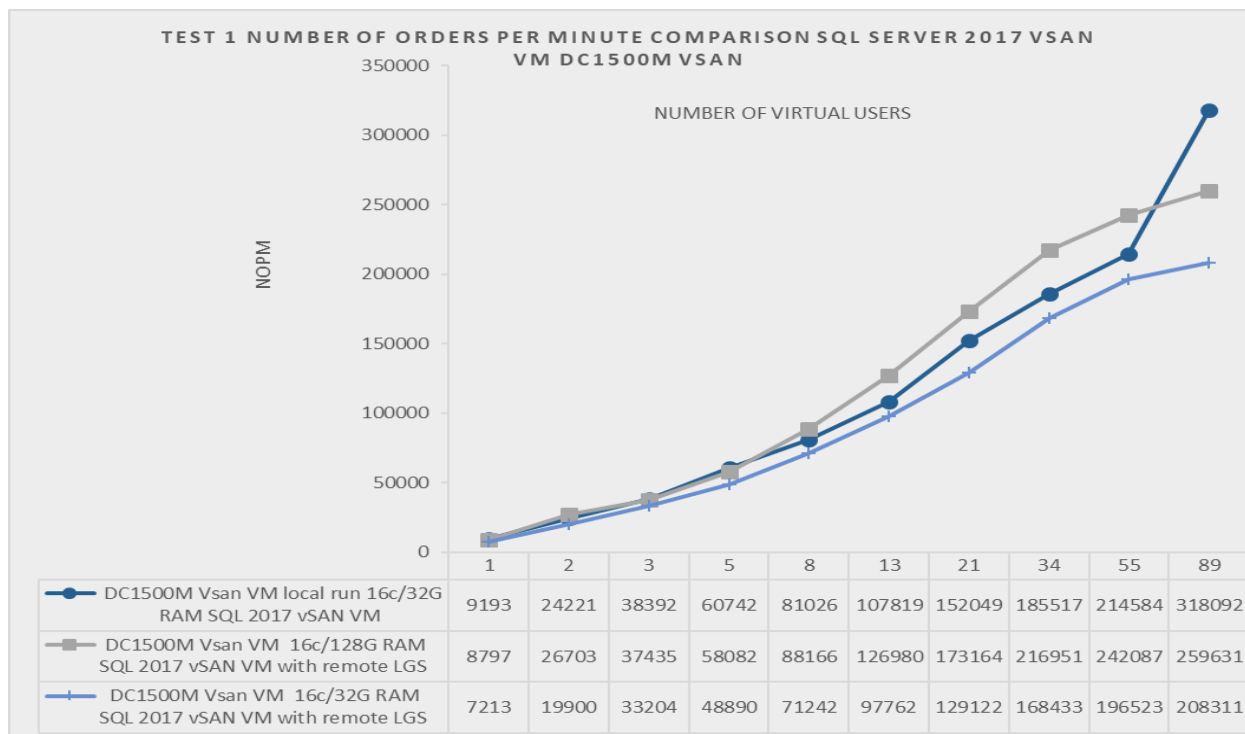


图2.3 测试1a、b、c: DC1500M vSAN 数据存储NOPM 比较, 不同的DRAM 大小

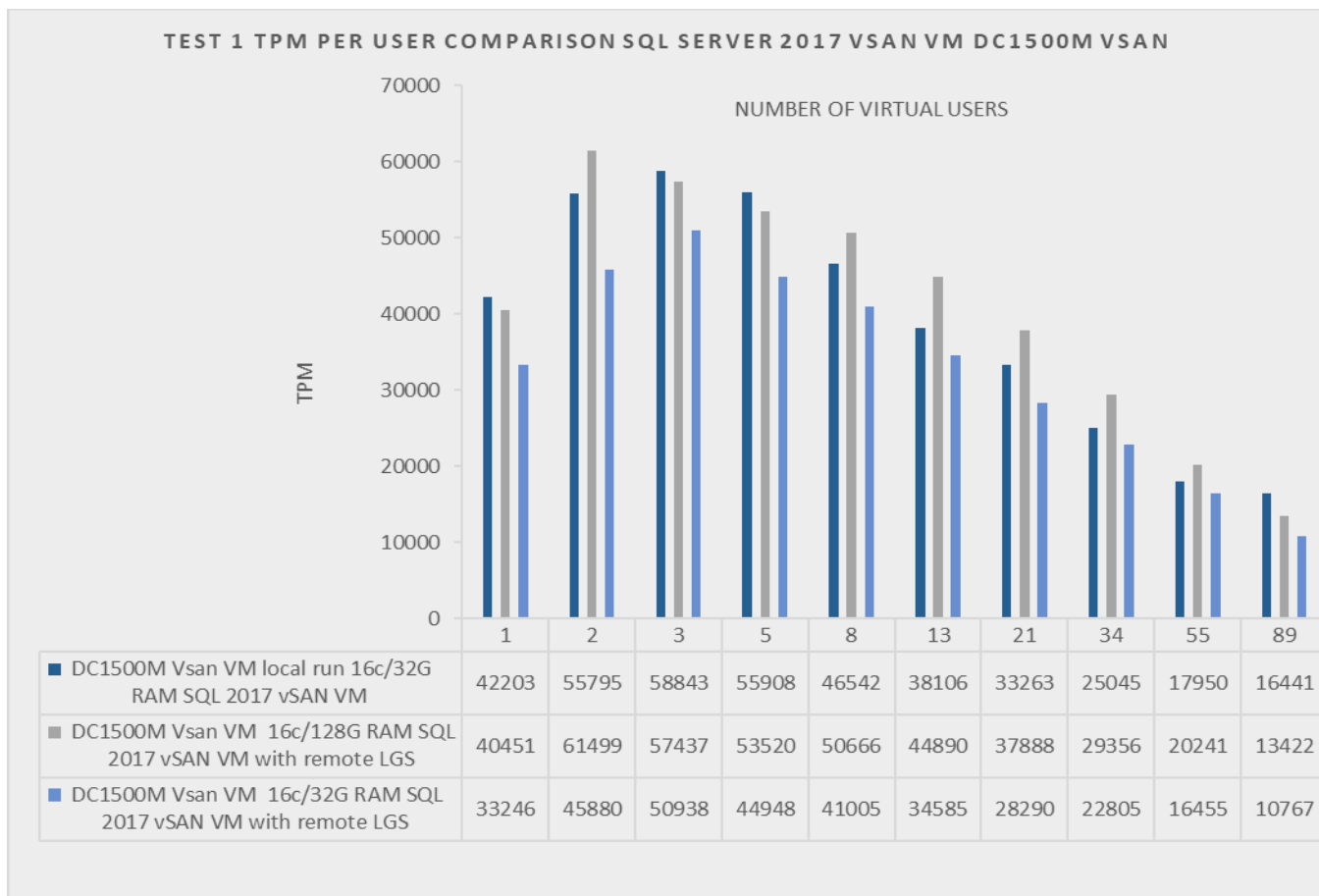


图2.4 测试1a、b、c：DC1500M vSAN 数据存储TPM 比较，不同的DRAM 大小

测试 2：金士顿 DC500M SATA 固态硬盘、Micron 5200 eco SATA 固态硬盘和 DC1500M NVMe 固态硬盘三种 vSAN 数据存储的 SQL Server 2017 性能比较

- 测试 1a 的 NVMe vSAN 数据存储的存储配置：3 个 DC1500M 960G FW S67F0103/磁盘组，共 4 个磁盘组（每个服务器 1 个），NVMe vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。（测试 1a） - 测试 1b 的 SATA vSAN 数据存储的存储配置：3 个 DC500M 1920G FW SCEJK2.8/磁盘组，共 3 个磁盘组（每个服务器 1 个），SATA vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。（测试 1b） - 测试 1c 的 SATA vSAN 数据存储的存储配置：3 个 Micron 5200 ECO 1920G FW D1MU004/磁盘组，共 3 个磁盘组（每个服务器 1 个），SATA vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。（测试 1b）		
测试 2a 说明	测试 2b 说明	测试 2c 说明
NVMe 测试环境中从 DC1500M vSAN 数据存储配置的虚拟磁盘。 选择了一个包含 1200 个仓库的数据库架构，用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 创建的虚拟用户序列为 1、2、3、5、8、13、21、34、55、89。 选择了 2 分钟提升时间和 5 分钟/用户的序列测试时长。 测试在 SUT 虚拟机上本地运行。	SATA 测试环境中从 D500M vSAN 数据存储配置的虚拟磁盘。 选择了一个包含 1200 个仓库的数据库架构，用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 创建的虚拟用户序列为 1、2、3、5、8、13、21、34、55、89。 选择了 2 分钟提升时间和 5 分钟/用户的序列测试时长。 测试在 SUT 虚拟机上本地运行。	SATA 测试环境中从 Micron 5200 eco vSAN 数据存储配置的虚拟磁盘。 选择了一个包含 1200 个仓库的数据库架构，用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 创建的虚拟用户序列为 1、2、3、5、8、13、21、34、55、89。 选择了 2 分钟提升时间和 5 分钟/用户的序列测试时长。 测试在 SUT 虚拟机上本地运行。

图 3.1 测试 2 说明：比较 SQL Server 2017 在 SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的性能

测试 2 比较了在以下 3 个不同数据存储上本地运行时，SQL Server 2017 受测系统虚拟机的 TPCC 基准测试性能：金士顿 DC1500M 企业级 NVMe vSAN 数据存储、[金士顿 DC500M](#) SATA 固态硬盘 vSAN 数据存储和 Micron 5200 eco SATA 固态硬盘 vSAN 数据存储。在测试 2 中，我们在 SQL Server 2017 虚拟机受测系统上本地运行了测试，以提高到数据区的 I/O 并强调包含架构的数据库的 IO 性能，并在测试中将用户序列从 1 名用户扩展为 89 名，以匹配我们的架构大小以及分配给 SQL server 虚拟机的 CPU/内存资源量。

测试 2 结果：金士顿 DC500M SATA 固态硬盘、Micron 5200 eco SATA 固态硬盘和 DC1500M NVMe 固态硬盘三种 vSAN 数据存储的 SQL Server 2017 性能比较

图 3.2 和 3.3 显示了我们在测试 2a、2b 和 2c 中实现的每分钟事务量 (TPM) 和每分钟新订单量 (NOPM)。在所有测试运行中，我们观察到 TPM 和 NOPM 随着虚拟用户数量的增加而向上扩展，但 NVMe 与 SATA 的扩展情况截然不同。89 名虚拟用户时，采用 DC1500M 的 vSAN 数据存储 SQL Server 2017 虚拟机可以实现 1,463,290 TPM 和 318,092 NOPM。相比而言，我们为 DC500M SQL server vSAN 虚拟机实现了 738,067 TPM/160,410 NOPM，而为 Micron 5200 eco vSAN 数据存储实现了 628499 TPM/136436 NOPM。这意味着，相比采用 SATA 的 vSAN 数据存储，采用相同数量的 DC1500M NVMe 硬盘，基于 NVMe 的 vSAN 数据存储实际上可让事务吞吐量和每分钟订单量增加

一倍。在企业环境中，如果将 VMware 基础架构升级为采用 DC1500M 等企业级 NVMe 解决方案，当有 89 名用户同时向数据库发送事务时，每名用户可以处理多 235% 的事务量（即更多的每分钟订单量）（图 3.4）。

图 3.5 显示了测试 2a、b 和 c 的平均 CPU 空闲时间与虚拟用户数量。这是衡量虚拟磁盘的有效指标 – 随着事务量的增加和 SQL server 数据库需要将数据从内存写入事务日志文件，虚拟磁盘可以多快地做出响应。89 名虚拟用户时，采用 DC1500M NVMe 的 vSAN 虚拟机的 CPU 空闲时间 (iowait) 为 15.5%，采用 DC500M 的虚拟机为 37.8%，而采用 Micron 5200 的虚拟机为 44.2%。这意味着，NVMe 虚拟磁盘响应 IO 请求的速度快得多，可避免 CPU 因等待 IO 完成而出现空闲，从而处理更多的事务。在企业环境中，利用 NVMe 升级 VMware 基础架构，可以更高效地利用分配给 SQL server 虚拟机的虚拟内核，从而提高事务吞吐量，并消除在更缓慢的存储分层上运行的旧 SQL 虚拟机的非必要内核，进而降低成本。

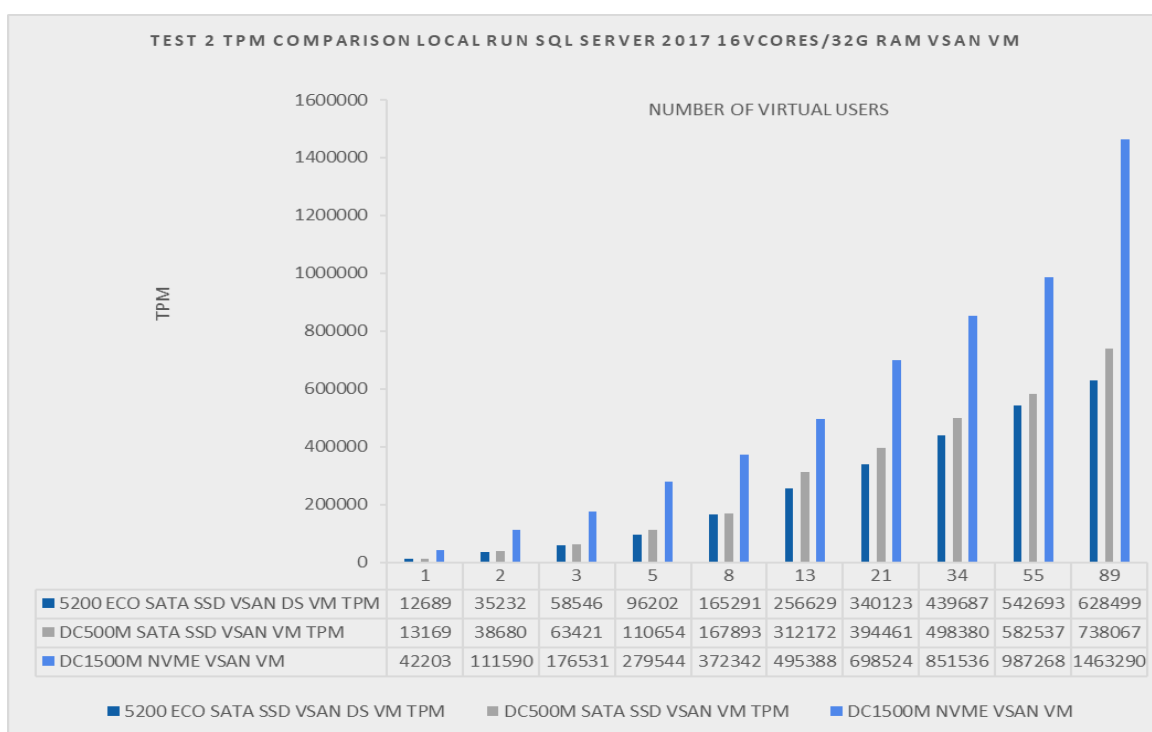


图 3.2 测试 2：TPM 比较，NVMe 与 SATA 两种 vSAN 数据存储



图3.3 测试2: NOPM 比较, NVME 与 SATA 两种 VSAN 数据存储

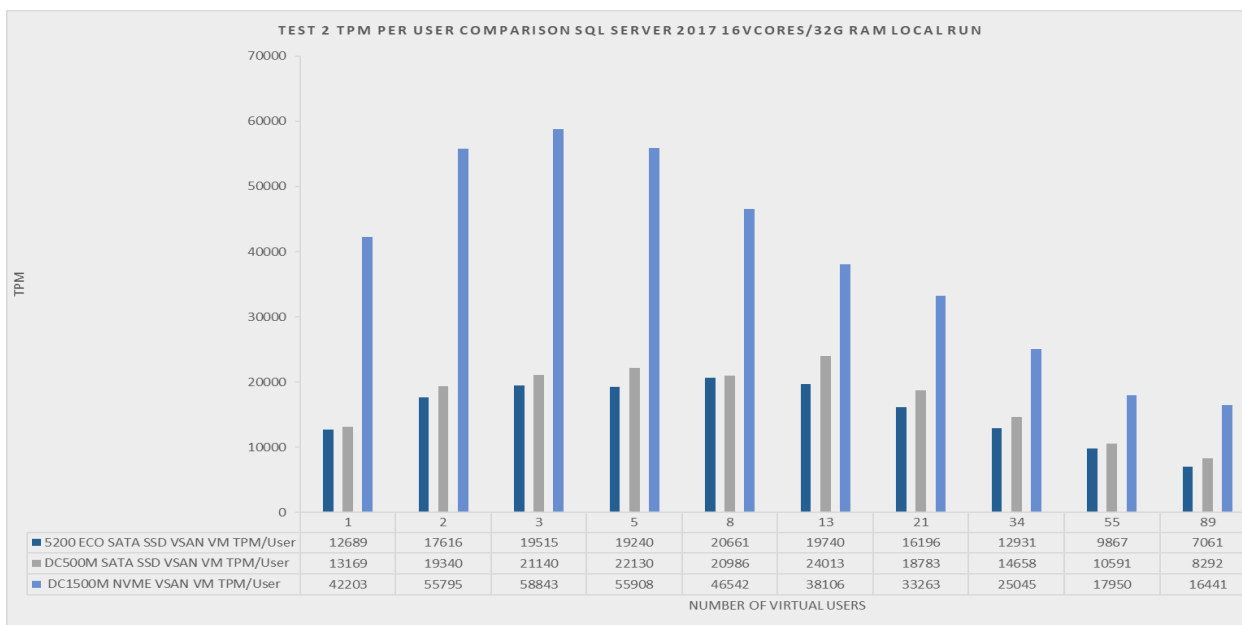


图3.4 测试2: 每用户TPM, NVME 与 SATA 两种 VSAN 数据存储

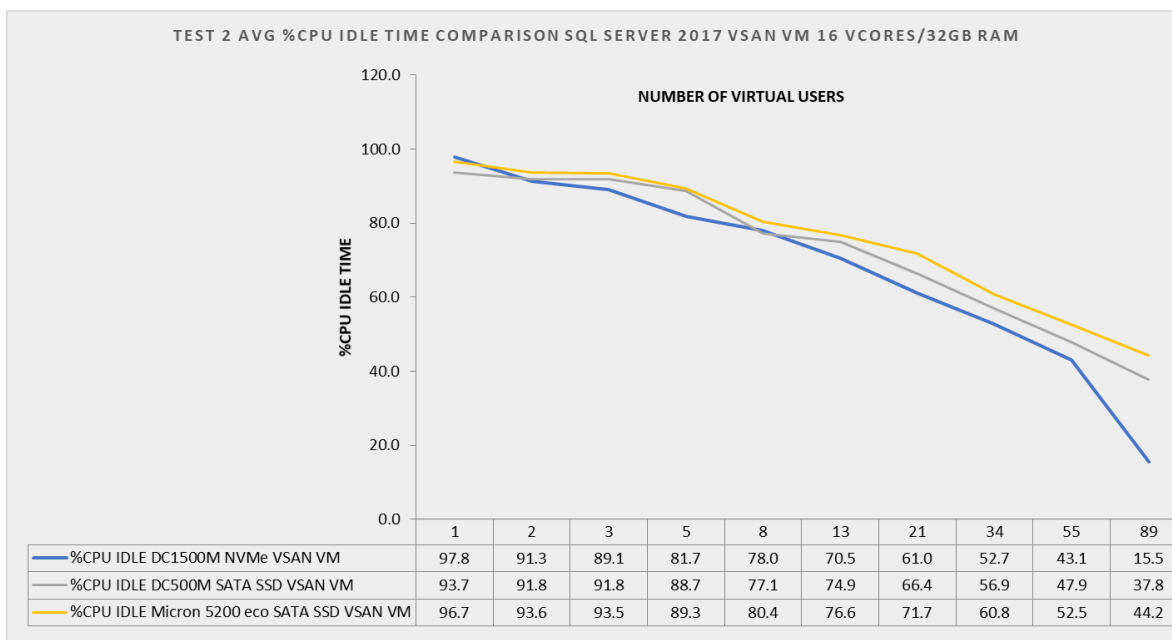


图3.5 测试2：%CPU 空闲比较，NVME 与 SATA 两种 VSAN 数据存储

测试 3：DC1500M NVMe 和 Micron 5200 eco SATA 两种 vSAN 数据存储的 SQL Server 2017 性能比较，更大的架构尺寸和更长的测试时长

- 测试 3a 的 NVMe vSAN 数据存储的存储配置：3 个 DC1500M 960G FW S67F0103/磁盘组，共 4 个磁盘组（每个服务器 1 个），NVMe vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。（测试 3a） - 测试 3b 的 SATA vSAN 数据存储的存储配置：3 个 Micron 5200 ECO 1920G FW D1MU004/磁盘组，共 3 个磁盘组（每个服务器 1 个），SATA vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。（测试 3b）	
测试 3a 说明	测试 2b 说明
NVMe 测试环境中从 DC1500M vSAN 数据存储配置的虚拟磁盘。 选择了一个包含 2000 个仓库的数据库架构，用于代表一个 157GB 的数据库。受测系统虚拟机 (SUT) 被分配了 40 个虚拟内核和 32GB RAM 创建的虚拟用户序列为 1、2、4、8、16、32、64、89、128 选择了 10 分钟提升时间和 20 分钟/用户的序列测试时长。 测试在 SUT 虚拟机上本地运行。	SATA 测试环境中从 Micron 5200 eco vSAN 数据存储配置的虚拟磁盘。 选择了一个包含 2000 个仓库的数据库架构，用于代表一个 157GB 的数据库。受测系统虚拟机 (SUT) 被分配了 40 个虚拟内核和 32GB RAM 创建的虚拟用户序列为 1、2、4、8、16、32、64、89、128 选择了 10 分钟提升时间和 20 分钟/用户的序列测试时长。 测试在 SUT 虚拟机上本地运行。

图4.1 测试3 说明：Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的 SQL Server 2017 数据库压力测试

本测试旨在利用更大的数据库架构运行更长时间的压力测试，以验证我们之前的结果，并比较在以下 2 个不同数据存储上本地运行时，SQL Server 2017 受测系统虚拟机的 TPCC 基准测试性能：金士顿 DC1500M 企业级 NVMe vSAN 数据存储和 Micron 5200 eco SATA 固态硬盘 vSAN 数据存储。这次我们选择了一个包含 2000 个仓库的架构，由此产生了一个 157 GB 的 TPC-C 数据库。我们为每个 SQL server 虚拟机使用了 40 个虚拟内核，以分配足够多的 CPU 资源来生成更多事务并实现最大事务吞吐量，但仅分配了 32 GB 的 RAM，让测试 IO 受到限制。我们微调了虚拟用户序列，从 1 名用户向上扩展为 128 名，并允许每个虚拟用户序列运行长得多的时间（20 分钟，包含 10 分钟提升时间）。这让我们可以在整个测试运行期间收集磁盘延迟指标。

测试 3 结果：DC1500M NVMe 和 Micron 5200 eco SATA 两种 vSAN 数据存储的 SQL Server 2017 性能比较，更大的架构尺寸和更长的测试时长

图 4.2 和 4.3 显示了我们在测试 3a 和 3b 中实现的每分钟事务量 (TPM) 和每分钟新订单量 (NOPM)。即便运行时间较长，采用 NVMe 固态硬盘和 SATA 固态硬盘的 SQL server 2017 虚拟机都可以随着虚拟用户数量增至 128 而向上扩展，但 NVMe 对应的扩展坡度大的多。89 名用户时，我们实现了 184 万 TPM 和 361743 NOPM，而采用 SATA 固态硬盘的 vSAN SQL 虚拟机则为 96 万 TPM 和 184451 NOPM。虚拟内核数量和已分配 DRAM 相同的情况下，采用 DC1500M NVMe 的 vSAN 数据存储的 TPM/NOPM 是采用 Micron 5200 eco 的 vSAN 虚拟机的 200%。

图 4.4 和 4.5 对比了采用 NVMe 与采用 SATA 固态硬盘的 vSAN SQL 虚拟机的平均虚拟磁盘延迟以及 99% 虚拟磁盘延迟与用户数量，这些数据是使用 Windows perfmon 收集的。对于每个虚拟用户序列，采用 DC1500M 的虚拟磁盘可以维持 <1ms 的平均延迟，即使用户数量持续增长也是如此。89 名虚拟用户时，采用 DC1500M 的虚拟磁盘的平均延迟为 0.92ms/IO，而采用 SATA 固态硬盘的虚拟磁盘为 2.36ms/IO，平均延迟比 NVMe 高 256%。更有趣的是 QoS 99% 延迟；89 名用户时，DC1500M 虚拟磁盘可在 1.61ms 内完成 99% 的全部 IO，而采用 SATA 固态硬盘的虚拟磁盘在 7.05ms 内完成 99% 的全部 IO，比 NVMe 多 437%。这里凸显了 NVMe 与 SATA 的延迟差距，而且，由于 DC1500M 设计为在整个持续 OLTP 工作负载内维持可预测的 QoS 延迟，我们没有看到延迟出现任何突然的激增，即便是虚拟用户的数量出现增长，并进而对块层提出更多并行的 IO 请求，也不例外。从企业的角度看，这意味着，将 VMware 基础架构从 SATA 固态硬盘升级为 DC1500M 等 NVMe 硬盘，可让您向上扩展事务并大幅降低事务延迟，从而让应用可以快速扩展并随着时间的推移降低成本。

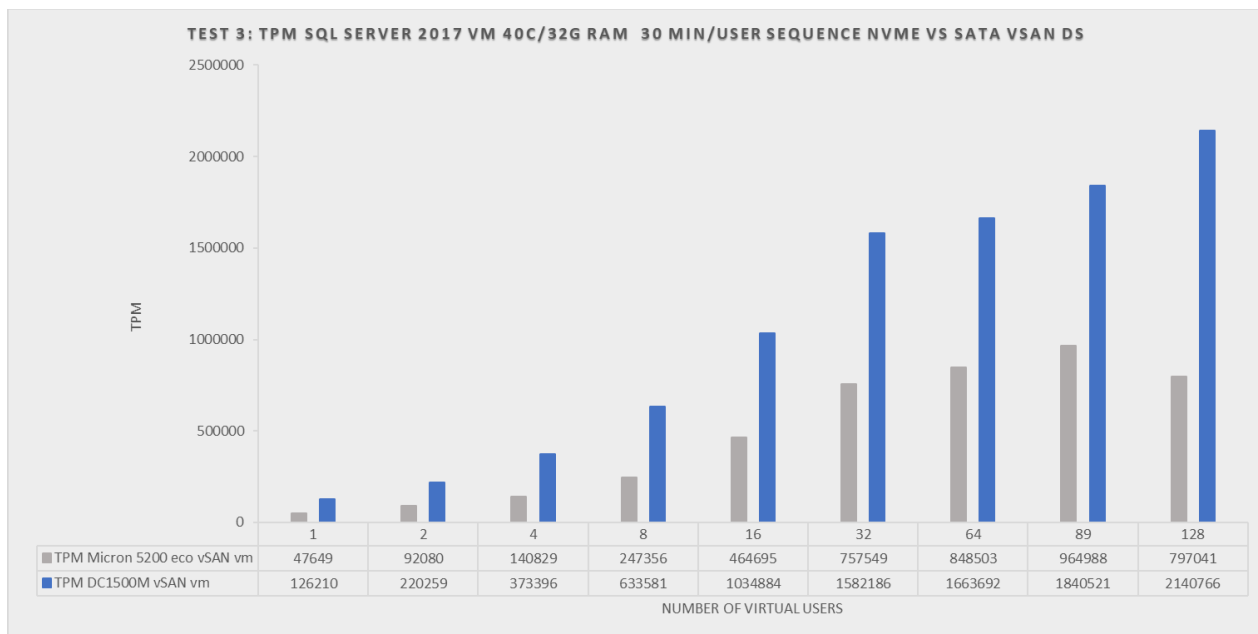


图 4.2 测试 3 TPM 比较, Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的 SQL Server 2017 数据库压力测试

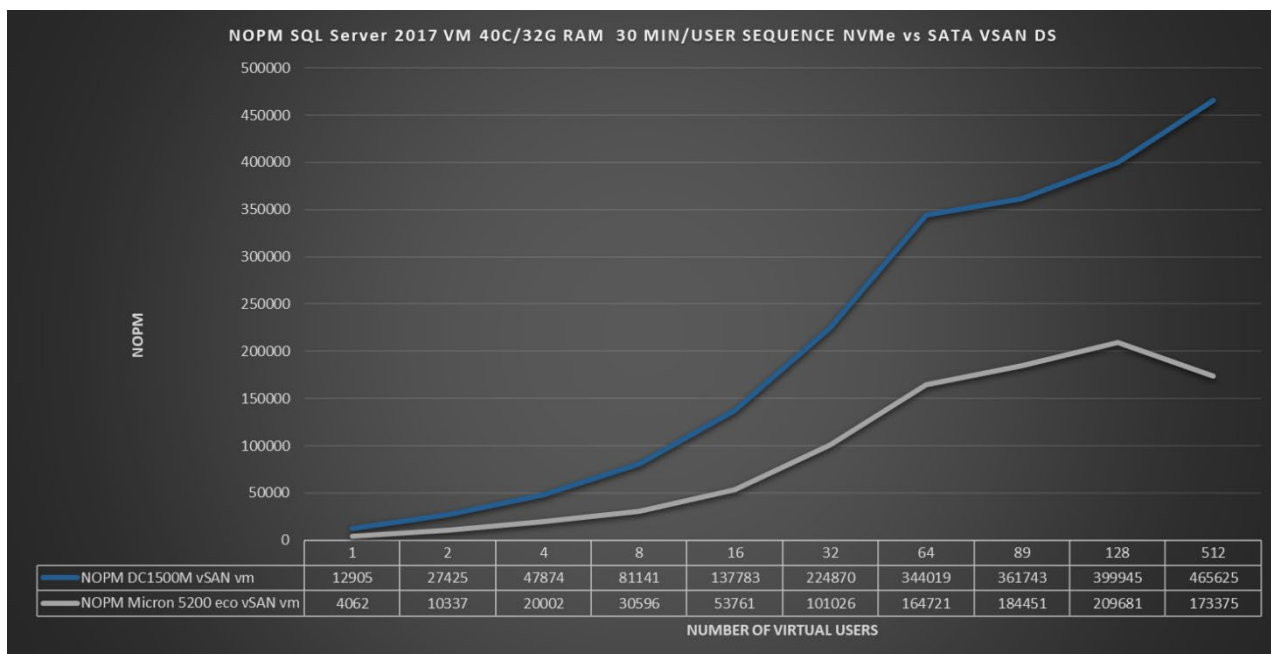


图 4.3 测试 3 TPM 比较, Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的 SQL Server 2017 数据库压力测试

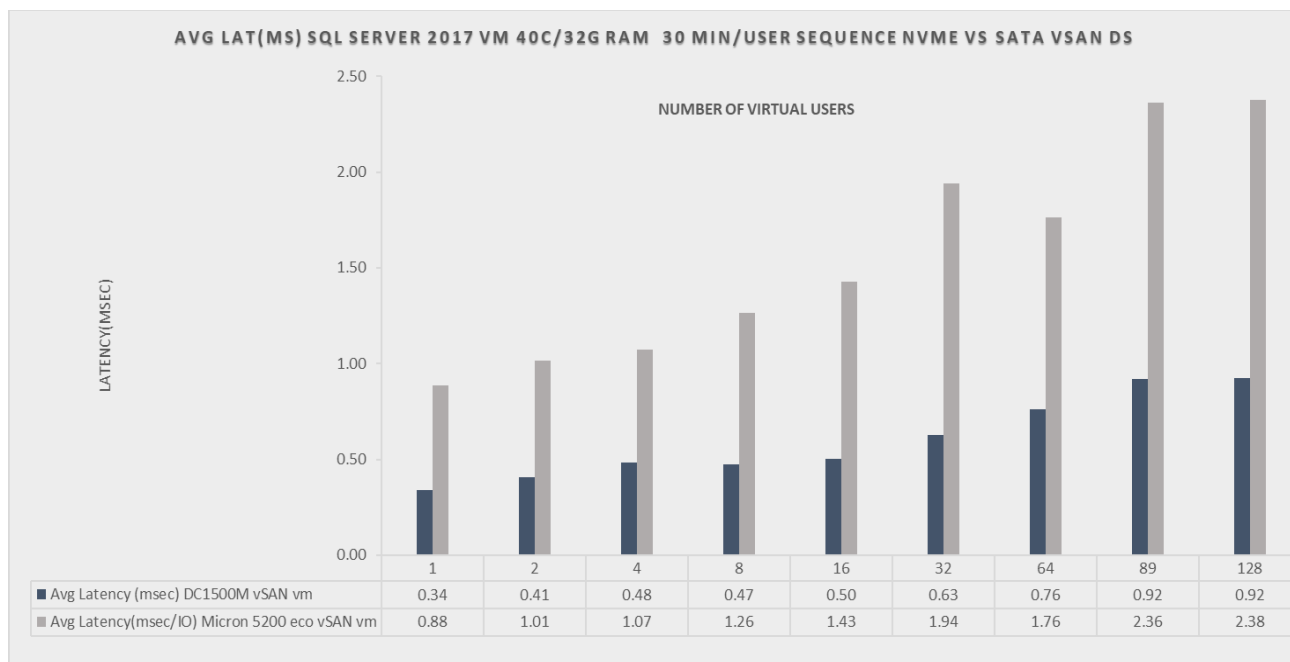


图 4.4 测试 3 平均延迟 (ms) 比较, Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的 SQL Server 2017 数据库压力测试

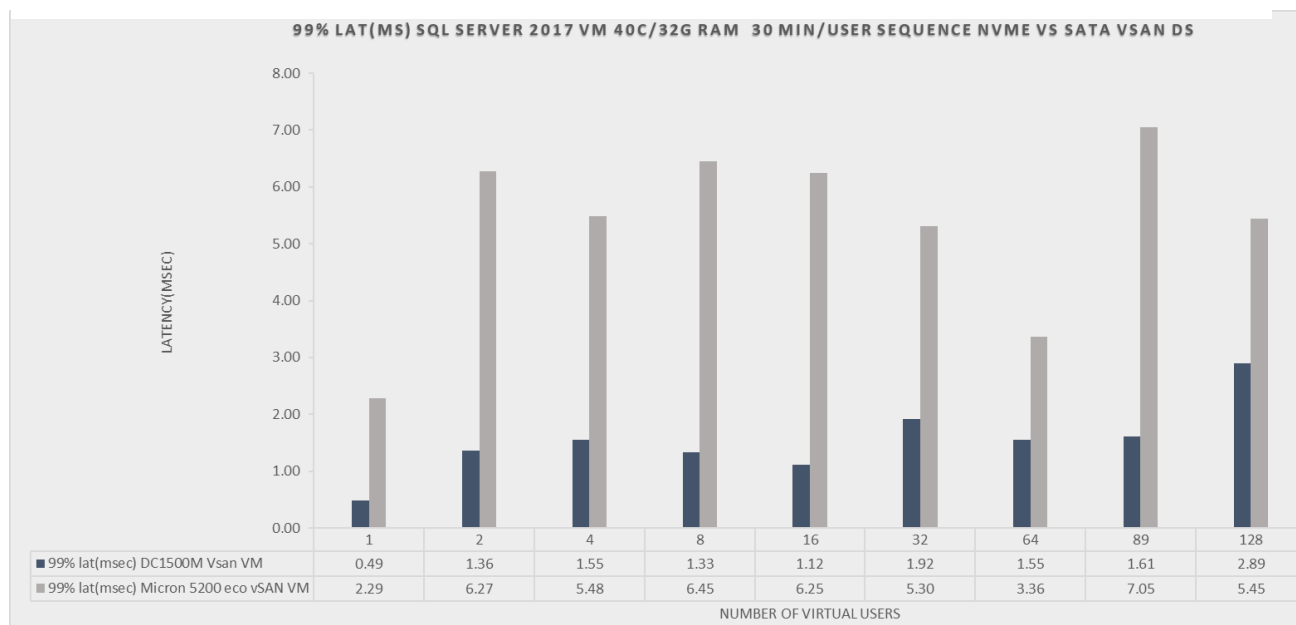


图 4.5 测试 3 99th % 延迟 (ms) 比较, Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的 SQL Server 2017 数据库压力测试

测试 4: SQL Server 2017 性能比较, 备份与还原性能, DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN

- 测试 3a 的 NVMe vSAN 数据存储的存储配置: 3 个 DC1500M 960G FW S67F0103/磁盘组, 共 4 个磁盘组 (每个服务器 1 个), NVMe vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。(测试 4a) - 测试 3b 的 SATA vSAN 数据存储的存储配置: 3 个 Micron 5200 ECO 1920G FW D1MU004/磁盘组, 共 3 个磁盘组 (每个服务器 1 个), SATA vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。(测试 4b)	
测试 4a 说明	测试 4b 说明
NVMe 测试环境中从 DC1500M vSAN 数据存储配置的虚拟磁盘。 在 SUT 上创建了一个包含 2000 个仓库的数据库架构, 用于代表一个 157GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 一个备份/还原脚本的 3 个循环被触发, 以备份和还原 tpcc 数据库和通过 Windows 性能监控程序录制的性能指标 测试在 SUT 虚拟机上本地运行。	SATA 测试环境中从 Micron 5200 eco vSAN 数据存储配置的虚拟磁盘。 在 SUT 上创建了一个包含 1200 个仓库的数据库架构, 用于代表一个 157GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 一个备份/还原脚本的 3 个循环被触发, 以备份和还原 tpcc 数据库和通过 Windows 性能监控程序录制的性能指标 测试在 SUT 虚拟机上本地运行。

图 5.1 测试 4 说明: Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的 SQL Server 2017 备份/还原性能比较

SQL 数据备份和还原操作是衡量底层虚拟磁盘吞吐量与延迟的好方法。我们希望在 TPC-C 备份/还原操作被触发时使用 Windows 性能监控程序收集虚拟磁盘指标, 从而对采用 NVMe 的 vSAN 虚拟机和采用 SATA 的 vSAN 虚拟机建立吞吐量和延迟性能基准。

测试 4: 结果: SQL Server 2017 性能比较, 备份与还原性能, DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN

图 5.2-5.4 显示了我们使用 Windows 性能监控程序脚本, 针对测试 4a) 和测试 4b) 中某个备份/还原循环收集的逐秒吞吐量和延迟。采用 DC1500M NVMe vSAN 数据存储的 SQL server 虚拟机在 265 秒内完成了 TPCC 数据库备份操作, 实现了 593MB/秒的平均吞吐量和 1.46ms/IO 的平均延迟。TPCC 数据库还原操作在 129 秒内完成, 其中平均带宽为 1.4GB/秒, 平均延迟为 2.65ms/IO。相比采用 Micron 5200 eco vSAN 的虚拟机, 采用 NVMe vSAN 的 SQL 虚拟机完成备份操作和还原操作的速度分别快 1.5 倍和 2.15 倍。

通常情况下, 备份和还原操作在非工作时间完成, 避免对生产虚拟机造成任何影响。不过, 凡事都有例外。如果 SQL 备份或还原操作在工作高峰时间完成, 您会希望操作尽管完成, 避免对在共享相同 vSAN 数据存储的第 1 层应用上执行事务的用户造成延迟影响。通过将您的 SQL 数据库迁移到采用 NVMe 的 vSAN 数据存储, 您可以削弱这种影响。即便备份/还原操作在非工作时间完成, 也可以更快地完成操作, 让共享相同资源的第 1 层数据库的停机时间缩短。

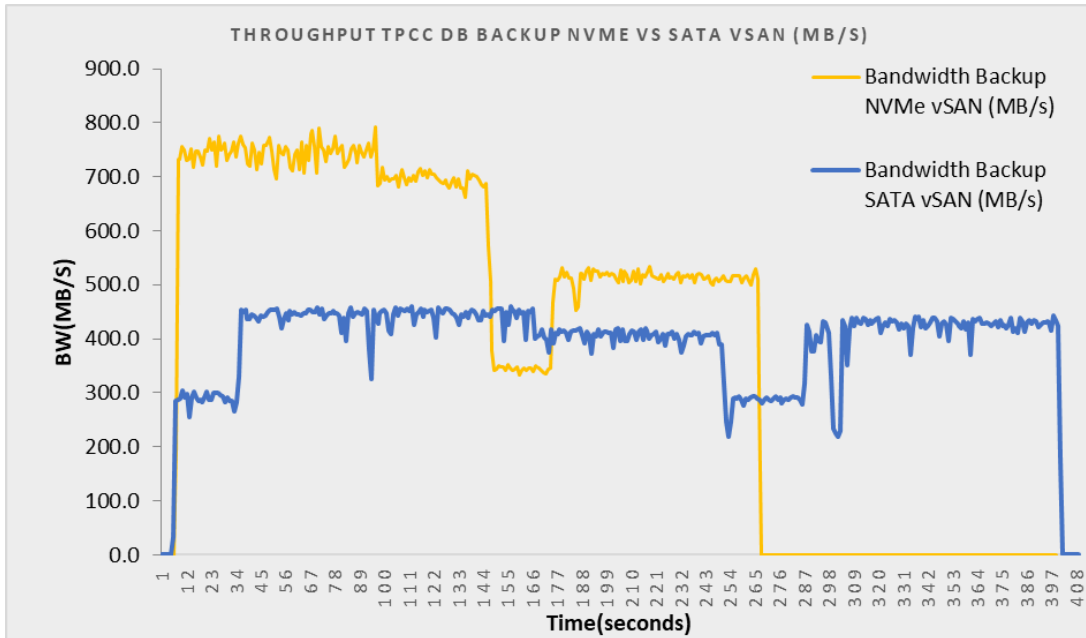


图5.2 吞吐量比较, SQL Server 2017 TPCC 数据库备份, Micron 5200 eco SATA 与 DC1500M NVMe 固态硬盘两种vSAN 数据存储 (MB/秒)

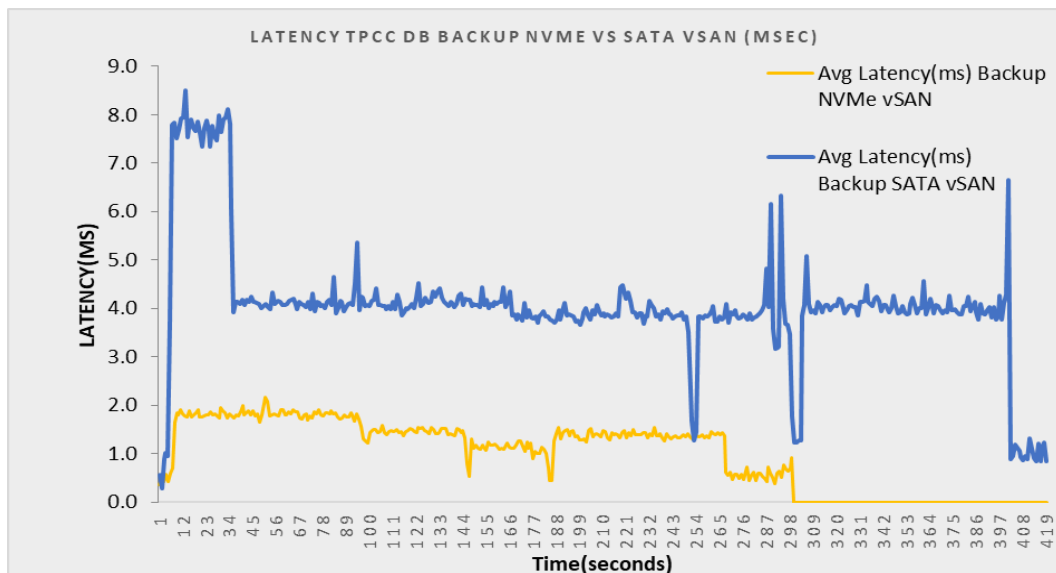


图5.3 平均延迟 (ms) 比较, SQL Server 2017 TPCC 数据库备份, Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种vSAN 数据存储

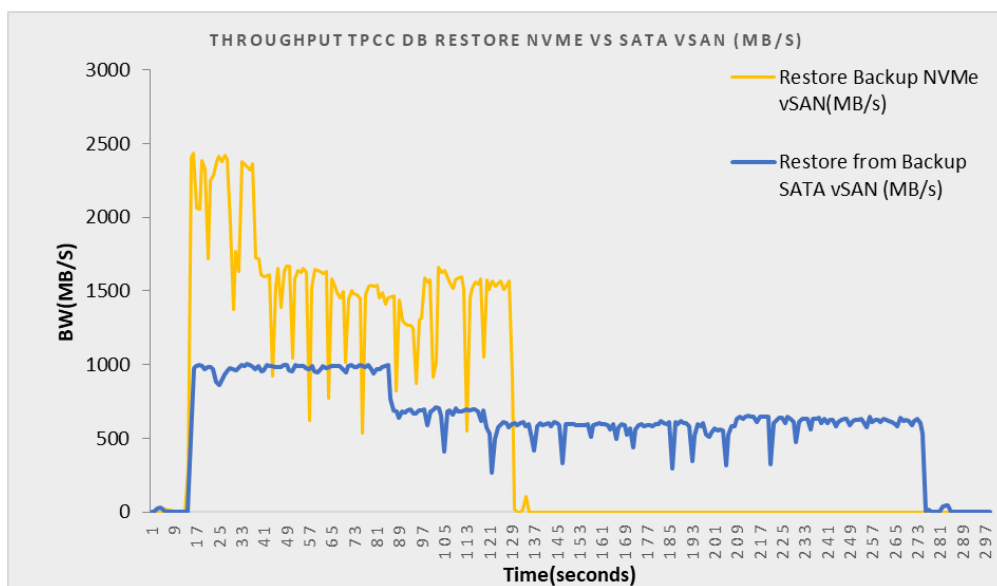


图5.4 吞吐量比较, SQL Server 2017 TPCC 数据库还原, Micron 5200 eco SATA 与 DC1500M NVMe 固态硬盘两种vSAN 数据存储 (MB/秒)

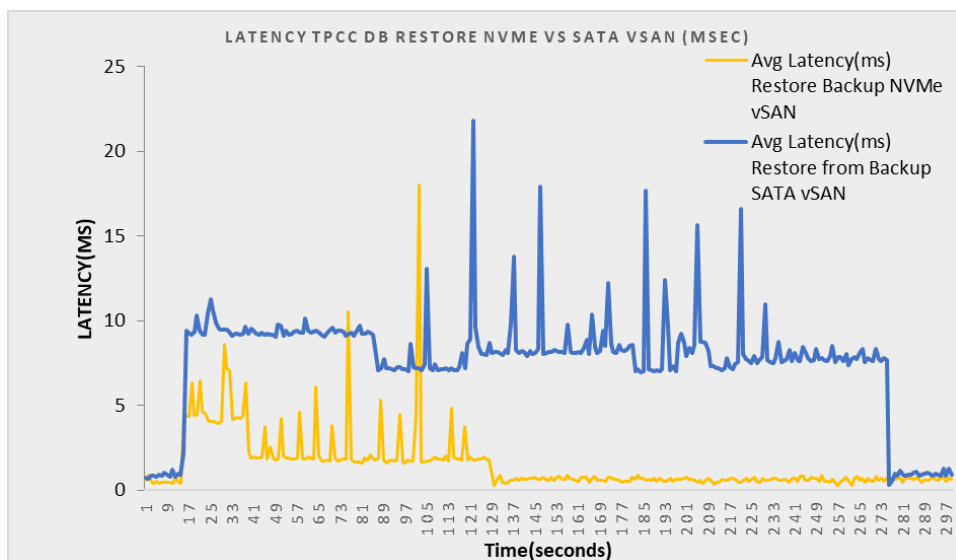


图5.5 延迟比较(ms), SQL Server 2017 TPCC 数据库还原, Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种vSAN 数据存储

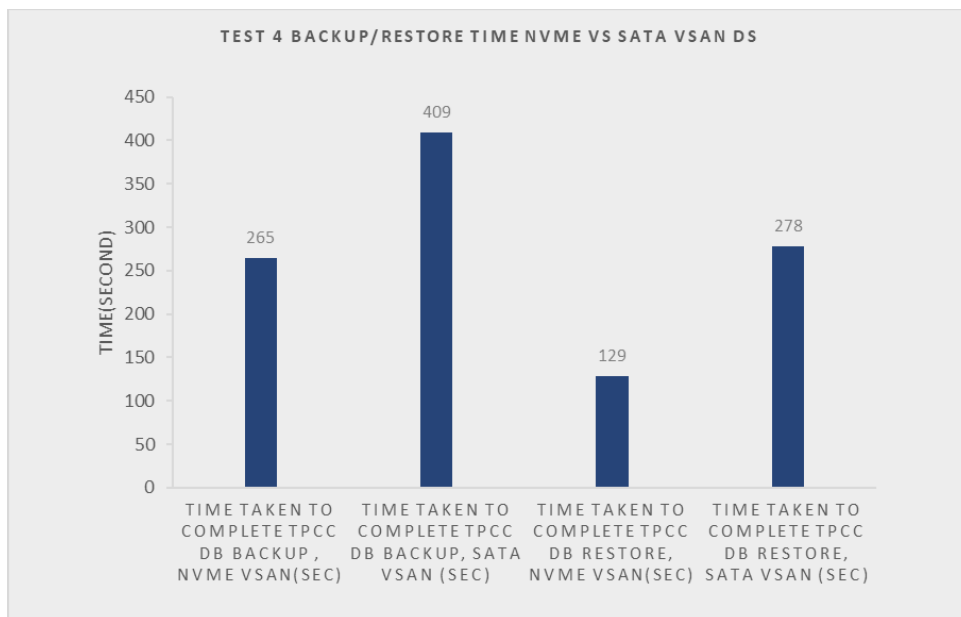


图5.6 完成SQL Server 2017 TPCC 数据库备份/还原操作花费的时间，Micron 5200 eco SATA 和DC1500M NVMe 固态硬盘两种vSAN 数据存储（秒）

测试 5: SQL Server 2017 性能比较, 近邻干扰测试, DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN

- 测试 3a 的 NVMe vSAN 数据存储的存储配置: 3 个 DC1500M 960G FW S67F0103/磁盘组, 共 4 个磁盘组 (每个服务器 1 个), NVMe vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。(测试 5a) - 测试 3b 的 SATA vSAN 数据存储的存储配置: 3 个 Micron 5200 ECO 1920G FW D1MU004/磁盘组, 共 3 个磁盘组 (每个服务器 1 个), SATA vSAN 测试环境。SQL Server 2017 与 Server 2019 Datacenter 客户机操作系统。(测试 5b)			
测试 5a 说明	测试 5b 说明	测试 5c 说明	测试 5d 说明
NVMe 测试环境中从 DC1500M vSAN 数据存储配置的 SQL 2017 虚拟机虚拟磁盘。 在 SUT 上创建了一个包含 1200 个仓库的数据库架构, 用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 受测 SUT 被克隆了 11 次, 并分配了 3 个 SUT 虚拟机/物理服务器 (共 12 个 SUT 虚拟机) 测试配置为运行 89 名虚拟用户, 在每个 SUT 虚拟机上选择了 30 分钟提升时间和 300 分钟测试时长 测试同时在全局 12 个 SUT 虚拟机上被触发	SATA 测试环境中从 Micron 5200 eco vSAN 数据存储配置的 SQL 2017 虚拟磁盘。 在 SUT 上创建了一个包含 1200 个仓库的数据库架构, 用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 受测 SUT 被克隆了 8 次, 并分配了 3 个 SUT 虚拟机/物理服务器 (共 9 个 SUT 虚拟机) 测试配置为运行 89 名虚拟用户, 在每个 SUT 虚拟机上选择了 30 分钟提升时间和 300 分钟测试时长 测试同时在全局 9 个 SUT 虚拟机上被触发	NVMe 测试环境中从 DC1500M vSAN 数据存储配置的 SQL 2017 虚拟机虚拟磁盘。 在 SUT 上创建了一个包含 1200 个仓库的数据库架构, 用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 受测 SUT 被克隆了 11 次, 并分配了 2 个虚拟机/物理服务器 (共 8 个 SUT 虚拟机) 来运行 HDB 工作负载。测试配置为运行 89 名虚拟用户, 在每个 SUT 虚拟机上选择了 30 分钟提升时间和 300 分钟测试时长。 1 虚拟机/物理服务器拥有一个包含 1200 个仓库的 tpcc 架构 (100GB), 备份脚本每 100 秒触发一次 (共 4 个虚拟机), 同时工作负载在其他 SUT 虚拟机上运行 10 个循环 8 个运行 HDB 工作负载的 SUT 虚拟机; 4 个运行备份脚本的虚拟机。 测试同时在全局 12 个虚拟机上被触发	SATA 测试环境中从 Micron 5200 eco vSAN 数据存储配置的 SQL 2017 虚拟机虚拟磁盘。 在 SUT 上创建了一个包含 1200 个仓库的数据库架构, 用于代表一个 100GB 的数据库。受测系统虚拟机 (SUT) 被分配了 16 个虚拟内核和 32GB RAM 受测 SUT 被克隆了 8 次, 并分配了 2 个虚拟机/物理服务器 (共 6 个 SUT 虚拟机) 来运行 HDB 工作负载。测试配置为运行 89 名虚拟用户, 在每个 SUT 虚拟机上选择了 30 分钟提升时间和 300 分钟测试时长。 1 虚拟机/物理服务器拥有一个包含 1200 个仓库的 tpcc 架构 (100GB), 备份脚本每 100 秒触发一次 (共 4 个虚拟机), 同时工作负载在 SUT 虚拟机上运行 6 个运行 HDB 工作负载的 SUT 虚拟机; 3 个运行备份脚本的虚拟机。 测试同时在全局 9 个虚拟机上被触发

图 6.1 测试 5 说明: Micron 5200 eco SATA 和 DC1500M NVMe 固态硬盘两种 vSAN 数据存储上的 SQL Server 2017 实际近邻干扰测试

我们在本测试中的目标是模拟一个现实情景，其中虚拟机中令人厌恶的工作负载（在本例中我们使用 TPCC 数据库备份操作）与运行生产工作负载（在试验中 TPCC 基本测试充当生产工作负载）的 SQL server 虚拟机共用相同的 vSAN 数据存储，并通过评估 TPCC 基准测试结果和分析通过 perfmon 和 vSAN 性能监控程序收集的重要存储指标对整体性能影响进行评估。

在测试 5a) 和 5b) 中，我们通过在所有虚拟机上并行运行 TPCC 基准测试来建立基准，其中未发生备份操作。我们在 NVMe vSAN 集群和 SATA vSAN 集群上每个服务器运行 3 个 SQL 虚拟机，使得总数达到 12 个面向 NVMe 的 SUT 虚拟机和 9 个面向 SATA 的 SUT 虚拟机。这项测试的架构大小为 1200 个仓库，相等于约 100GB 的 TPC-C 数据库大小，我们运行了包含 89 名用户的 TPCC 工作负载 300 分钟，包括 30 分钟的提升时间。

在测试 5c) 和 5d) 中，我们在所有 SUT 虚拟机上还原了 TPC-C 数据库。然后，我们触发了一个脚本，以在 4 个面向 NVMe 集群的虚拟机和 3 个面向 SATA 集群的虚拟机上执行 10 个 TPC-C 数据库备份循环，同时在其余 SUT 虚拟机上并行运行相同的 TPC-C 基准测试。这意味着，在 NVMe vSAN 集群中，同时有 8 个虚拟机运行 TPC-C 工作负载、4 个虚拟机运行备份工作负载。与此同时，在 SATA vSAN 集群中，同时有 6 个虚拟机运行 TPC-C 工作负载、3 个虚拟机运行 TPC-C 数据库备份工作负载。

测试 5 结果：SQL Server 2017 性能比较，近邻干扰测试，DC1500M NVMe 与 Micron 5200 eco SATA 两种 vSAN

图 6.2 和 6.3 显示了我们在测试 5a 和 5b 中实现的每分钟事务量 (TPM) 和每分钟新订单量 (NOPM)。在 12 个采用 DC1500M NVMe vSAN 数据存储的 SQL server 虚拟机上（每个虚拟机 89 名用户），我们可以实现每个虚拟机平均 523,516 的 TPM 和平均 113,812 的 NOPM，而 9 个采用 Micron 5200 eco SATA 集群的 SQL 虚拟机则为每个虚拟机平均 269,320 TPM 和 58544 NOPM。从 vSAN 性能监控程序收集的 IOPS 和延迟指标（下文的图 6.4 和 6.5）显示，块层上实现的 IO 意味着，NVMe 集群具备 120,000 读取 IOPS 和 60,000 写入 IOPS，且其读/写操作的延迟为 800μs，而 SATA vSAN 集群为 50,000 读取/20,000 写入，其读取延迟平均为 3.8ms，写入平均延迟为 5.5ms。这再次凸显了 NVMe 和 SATA 之间的性能差距，并表明采用 DC1500M NVMe 的虚拟磁盘可以削弱并行请求的影响，并以快得多的往返延迟处理请求。

图 6.5 和 6.6 显示了我们在测试 5c 和 5d 中实现的每分钟事务量 (TPM) 和每分钟新订单量 (NOPM)。当虚拟用户运行在 8 个采用 DC1500M NVMe vSAN 数据存储的 SQL server 虚拟机（每个虚拟机 89 名用户），且同时虚拟机备份在 4 个虚拟机上并行触发时，我们可以实现平均 575,933 的 TPM 和平均 125,206 的 NOPM；而在采用 Micron 5200 eco SATA 的 SATA vSAN SQL 虚拟机上，当 6 个 SQL 虚拟机运行 TPCC 工作负载、且同时虚拟机备份在 3 个虚拟机上并行触发时，可以实现平均 351,258 的 TPM 和 76355 的 NOPM。要了解完整情况，我们必须分析 SATA vSAN 集群和 NVMe vSAN 集群的延迟和存储指标，并分析在这两个集群中完成备份的速度有多快。

图 6.8 和 6.9 显示了在测试 5c 和 5d 中使用 vSAN 性能监控程序从 NVMe 集群和 SATA 集群收集的 vSAN IOPS 和延迟指标。备份脚本配置为每 100 秒运行一次，运行 10 个循环。我们可以看到触发的虚拟机备份对 NVMe 集群和 SATA vSAN 集群的 IOPS 和读写延迟的影响。不过，对延迟的影响存在差异。NVMe 集群的最大读取/写入 IO 延迟达到 4ms/IO，同时维持平均 2.5 ms/IO 的读取/写入

操作延迟；而 SATA vSAN 高达 9ms/IO，并维持平均 7.3 ms/IO 和 4.9 ms/IO 的读取/写入 IO 延迟。当最终用户尝试提交订单、更新购物车或查看其它仓库的产品时，就会感受到这种延迟。

图 6.11 显示了一个采用 SQL Server DC1500M vSAN 的虚拟机和一个采用 Micron 5200 eco vSAN 的 SQL 虚拟机完成备份循环所花费的时间，其中不包含两个循环之间的等待时间。SQL server NVMe vSAN 虚拟机花了 73 分钟完成 10 次备份，平均每次备份 7 分钟；而采用 SQL server SATA 固态硬盘的 vSAN 虚拟机完成 10 次备份花了 122.15 分钟，平均每次备份 12 分钟。采用 DC1500M vSAN 的虚拟机完成备份循环的速度是采用 Micron 5200 eco vSAN 的虚拟机的 1.67 倍。这个实证证据表明，通过将 VMware 基础架构升级为采用 DC1500M NVMe 的数据存储，可以更快地完成数据库备份等不想要的操作，从而有助于规避近邻干扰问题；并且，由于卓越的延迟和吞吐量能力，NVMe 可以削弱这些令人厌恶的工作负载对第 1 层应用的延迟影响。

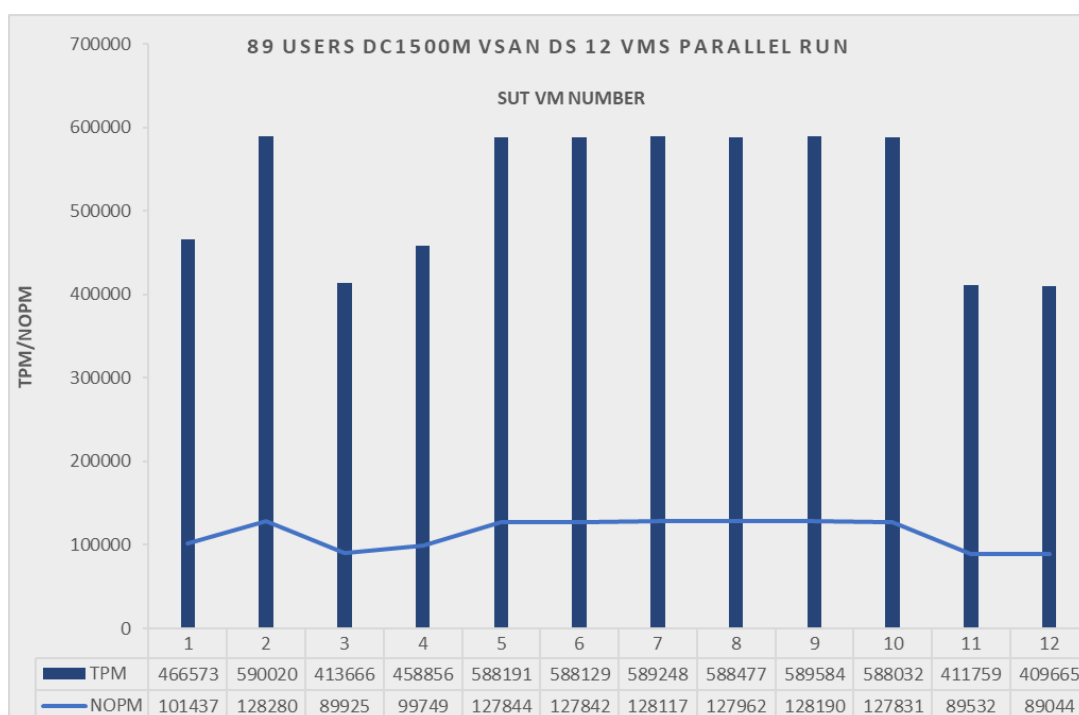


图6.2 测试 5a，TPM，SQL Server 2017 300 分钟，12 个虚拟机并行运行，89 名虚拟用户，DC1500M NVMe 固态硬盘 vSAN 数据存储

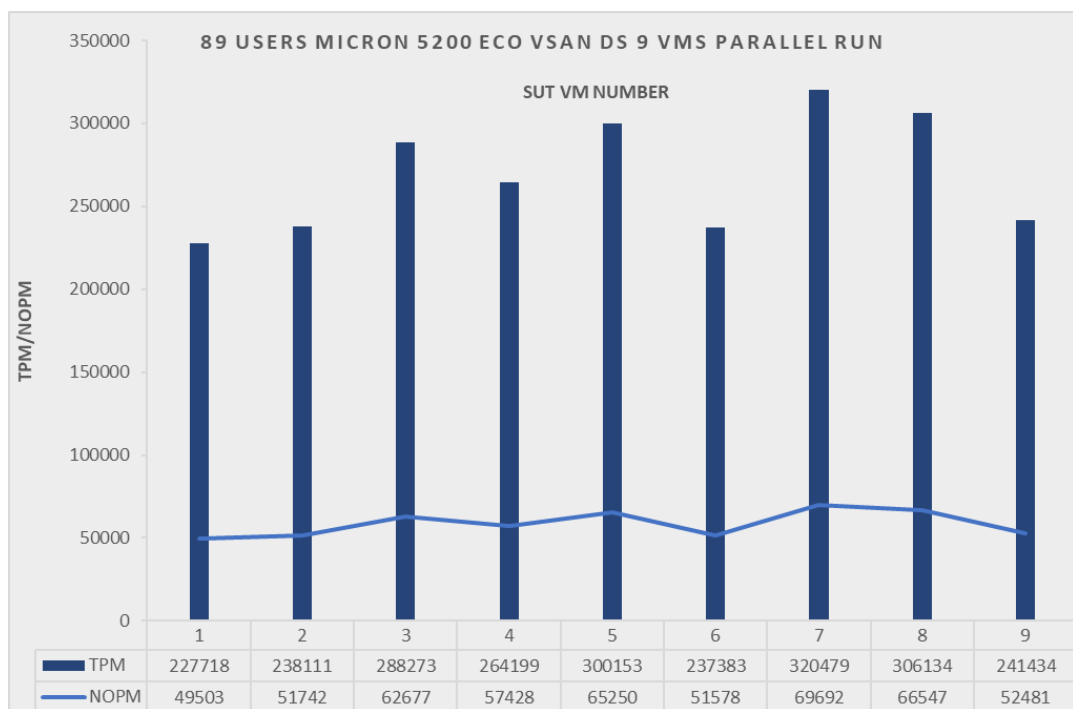


图6.3 测试5b, TPM, SQL Server 2017 300 分钟, 12 个虚拟机并行运行, 89 名虚拟用户, DC1500M NVMe 固态硬盘 vSAN 数据存储

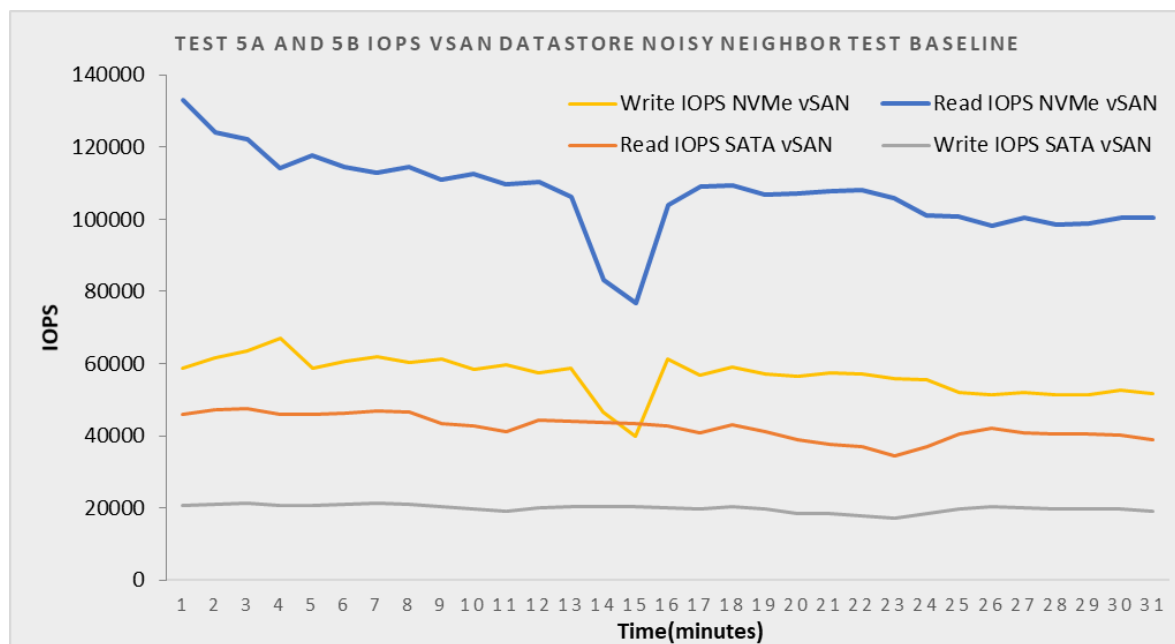


图6.4 测试5a 和5b, 近邻干扰, IOPS, DC1500M NVMe 和 Micron 5200 eco 两种vSAN 数据存储

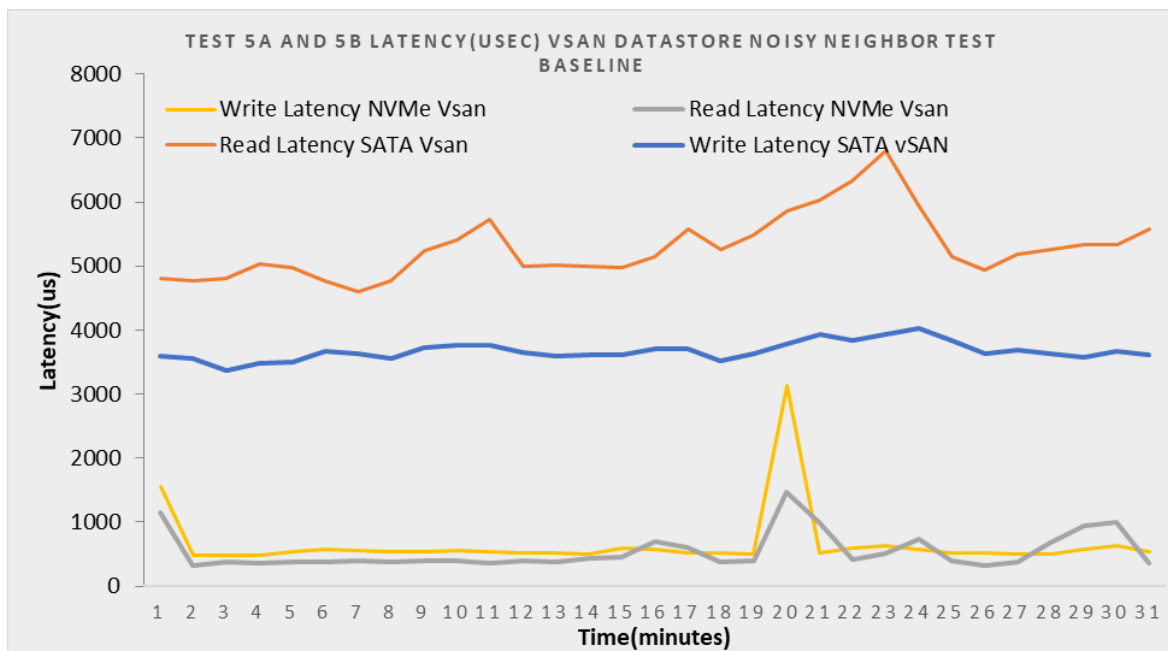


图6.5 测试5a 和5b, 近邻干扰, 延迟, DC1500M NVMe 和 Micron 5200 eco 两种vSAN 数据存储

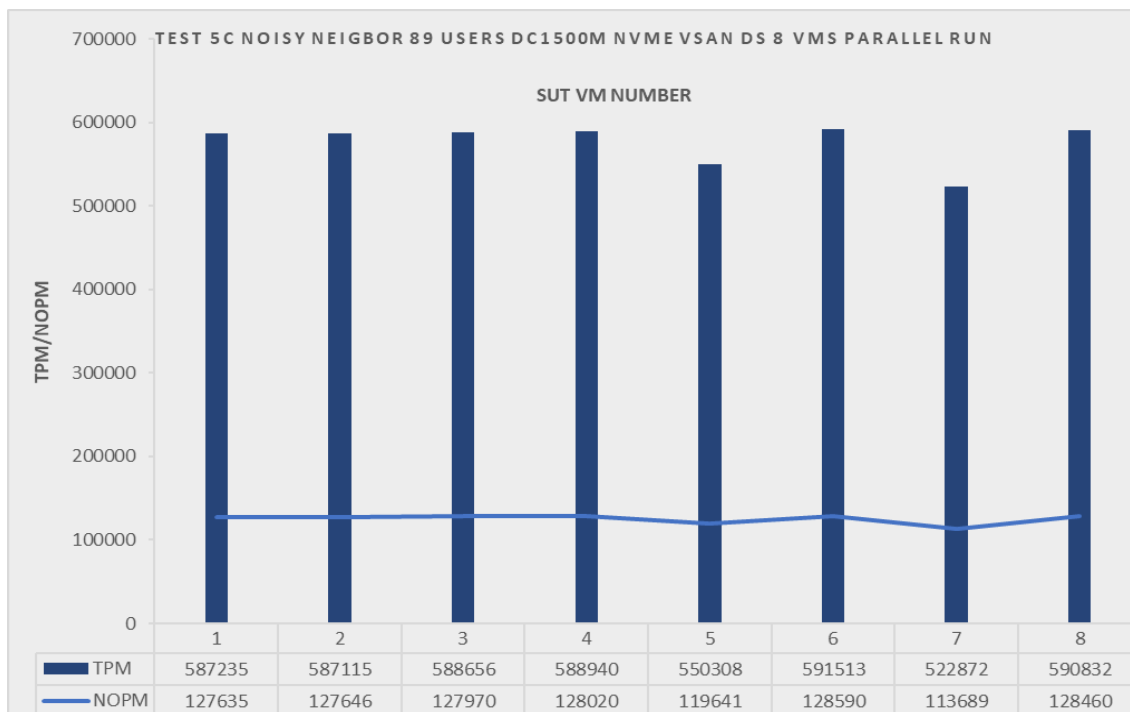


图6.6 测试5c, TPM, 近邻干扰实现, 8个虚拟机并行运行, DC1500M NVMe vSAN 数据存储

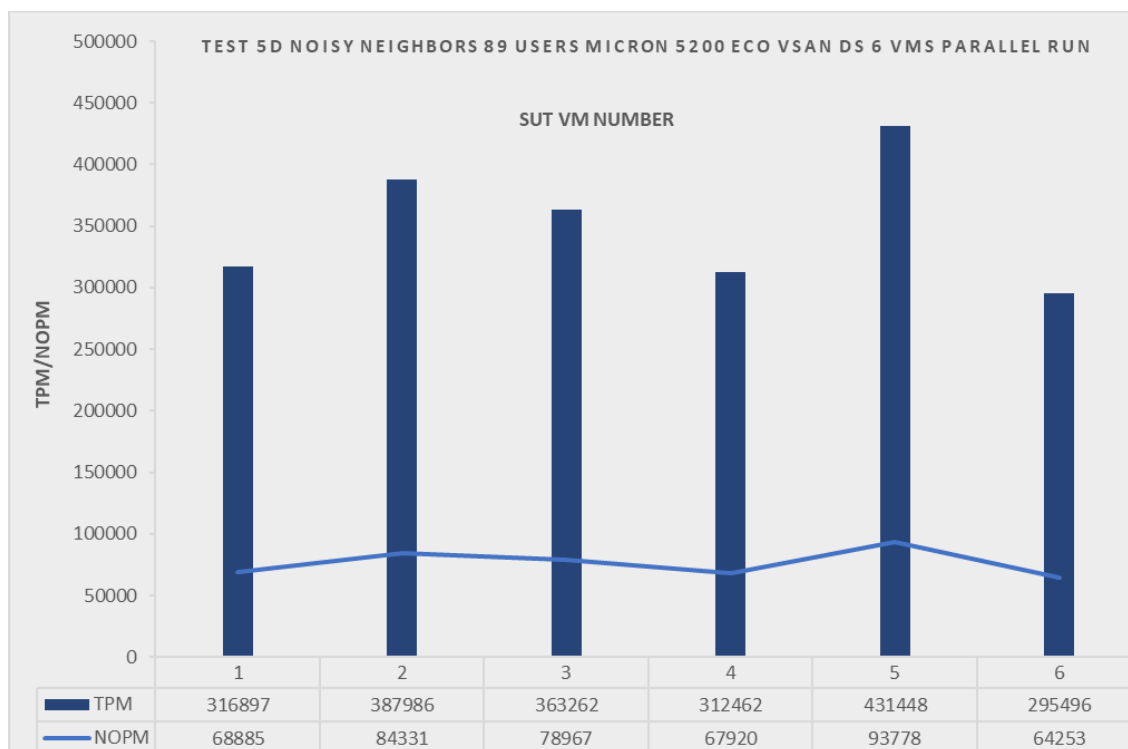


图6.7 测试5d，TPM，近邻干扰实现，6个虚拟机并行运行，Micron 5200 eco vSAN 数据存储

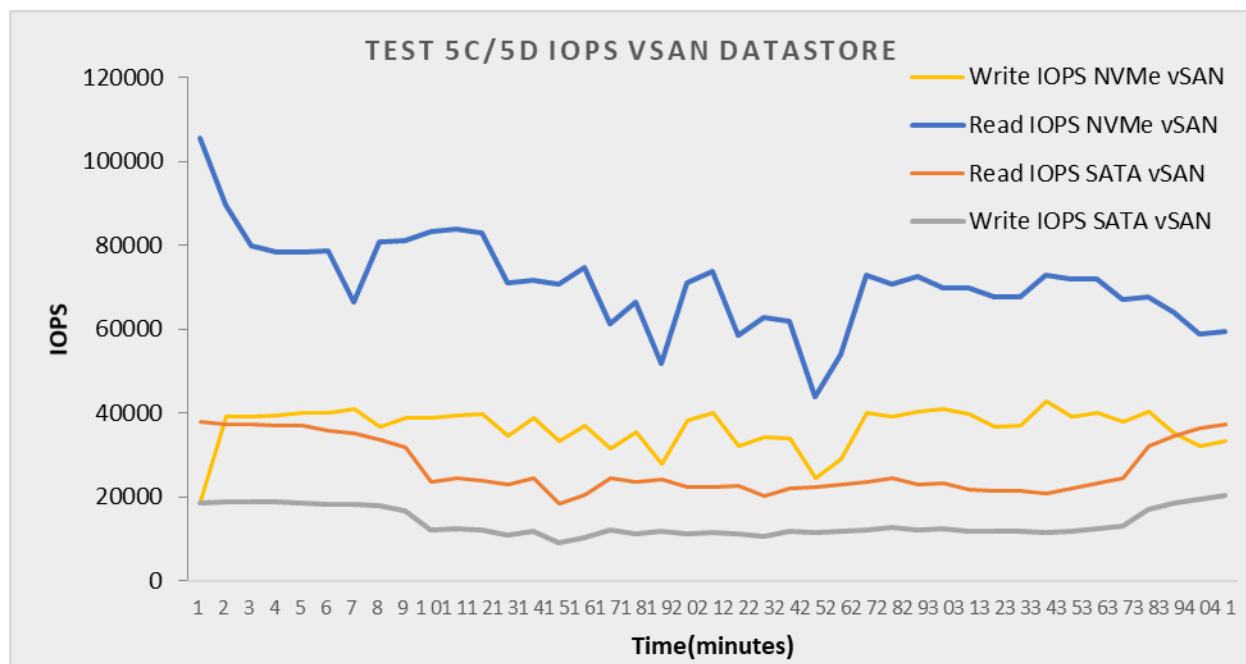


图6.8 测试5c/5d，IOPS，近邻干扰实现，NVMe 与 SATA 固态硬盘两种vSAN 数据存储

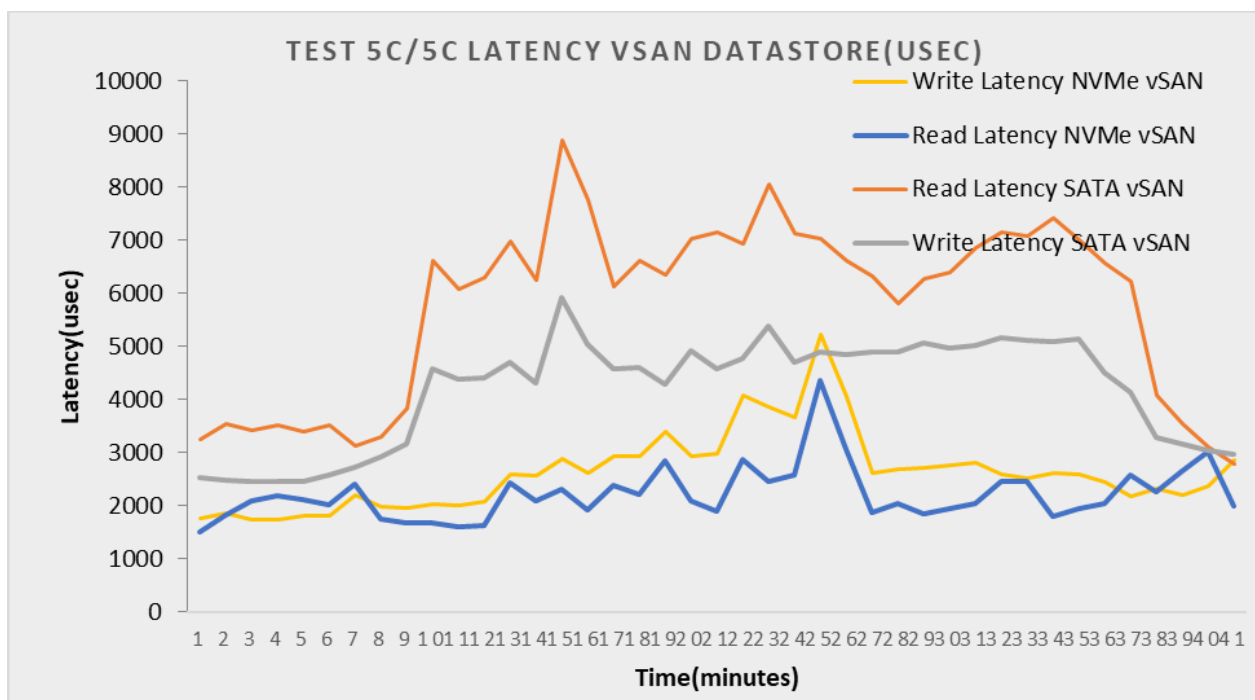


图6.9 测试5c/5d，延迟，近邻干扰实现，NVMe 与SATA 固态硬盘两种vSAN 数据存储

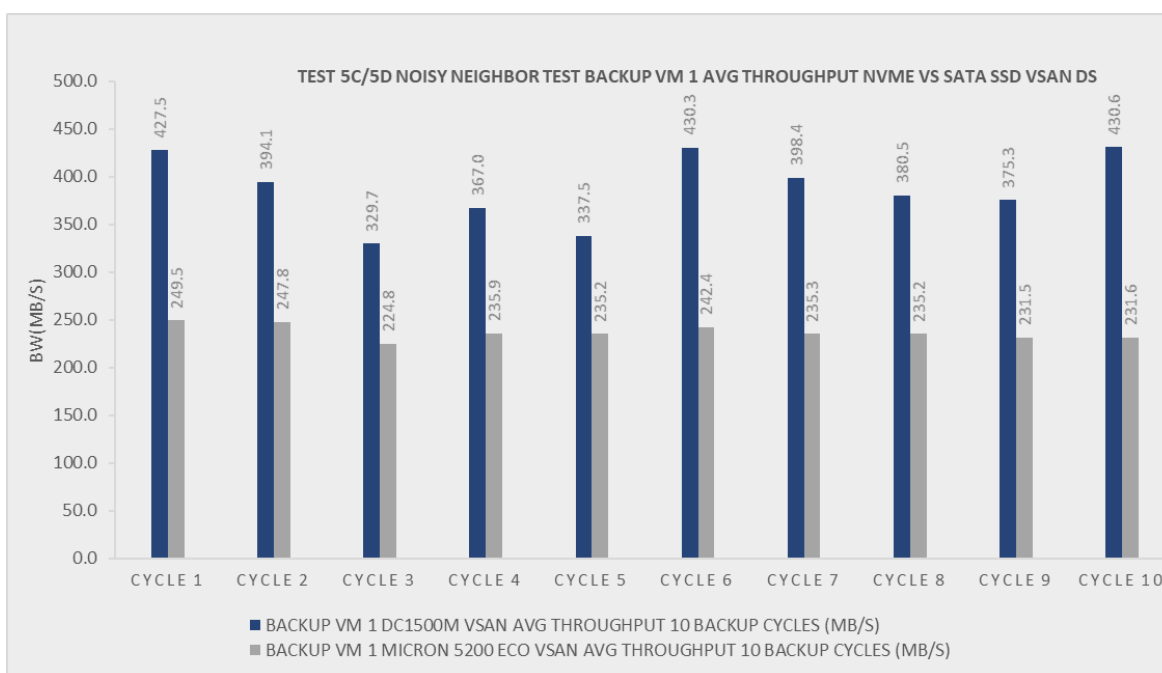


图6.10 测试5c/5d，备份虚拟机，吞吐量，近邻干扰实现，NVMe 与SATA 固态硬盘两种vSAN 数据存储

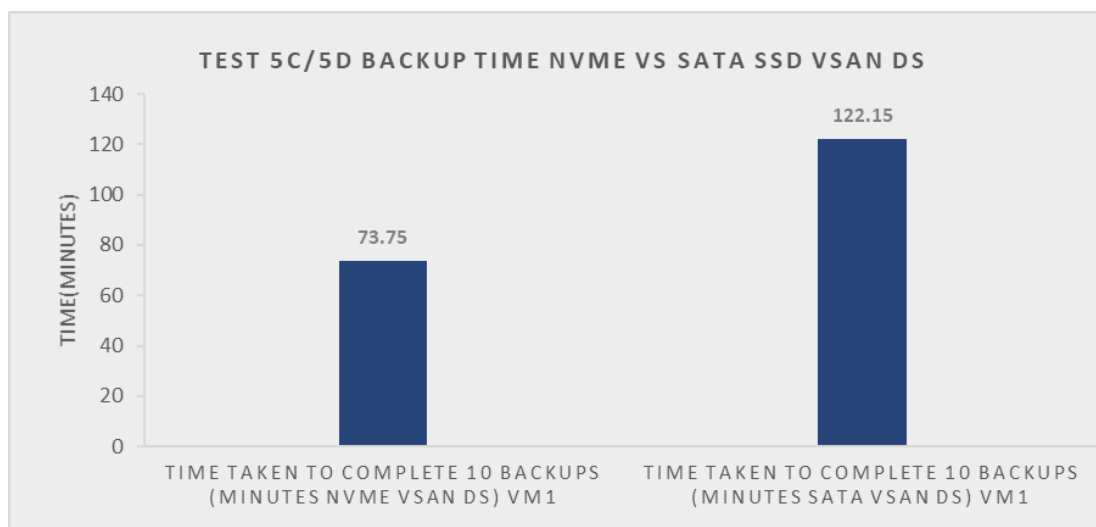


图6.11 测试5c/5d，备份虚拟机，完成备份所花时间，10次循环，近邻干扰实现，NVMe与SATA固态硬盘两种vSAN数据存储

结论

在本白皮书中，我们展示了将数据库工作负载整合到 NVMe 会有助于最大限度利用现有硬件，得益于 NVMe 的惊人效率和接近零的 IO 等待时间，可让您使用更少的 CPU 内核实现相同的事务吞吐量。我们将其与企业级 SATA 固态硬盘做了比较，结果显示通过将 SQL 工作负载迁移到采用 NVMe 的数据存储，您可以让应用向上扩展，将事务吞吐量提高一倍并实现亚毫秒级延迟。然后，我们展示了 NVMe 可以如何更快地完成数据库备份/还原操作等不必要的工作负载，从而帮助规避对第 1 层应用的影响。

金士顿的企业级 NVMe 固态硬盘 [DC1500M](#) 与金士顿服务器内存 (Server Premier) 搭配使用，可为寻求实现数据库基础架构虚拟化并最大限度提高工作效率的用户提供出色的解决方案。

访问 <https://www.kingston.com/en/solutions/servers-data-centers>，进一步了解金士顿的数据中心解决方案

参考资料

HammerDB。（日期不详）*Understanding the TPCC workload*。获取自 <https://www.hammerdb.com/docs3.3/ch03s05.html>

TPCC 主页。（日期不详）获取自 <https://www.tpc.org/>