



Сокращение совокупной стоимости владения и создание отказоустойчивых и быстро реагирующих баз данных с помощью твердотельных накопителей DC1500M Enterprise NVMe компании Kingston и гиперконвергентной инфраструктуры VMware vSAN

Автор: Hazem Awadallah, системный инженер, Kingston Technology

Проверено: Chris Selden, менеджер по разработке твердотельных накопителей, Kingston Technology



Содержание

- [Краткий обзор](#)
- [Типичные инфраструктурные проблемы, с которыми РСУБД сталкиваются сегодня в центрах обработки данных](#)
- [Решение: твердотельные накопители Kingston Technology Data Center DC1500M Enterprise NVMe](#)
- [Тестовая среда](#)
 - [I. Инфраструктура](#)
 - [II. Конфигурация базы данных](#)
 - [III. Производительность системы хранения данных vSAN](#)
- [Результаты испытаний](#)
 - [Тест 1, SQL server 2017 на виртуальной машине в vSAN на базе DC1500M 960GB с различными объемами памяти DRAM](#)
 - [Результаты теста 1, SQL server 2017 на виртуальной машине в vSAN на базе DC1500M 960GB с различными объемами памяти DRAM](#)
 - [Тест 2: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе Kingston DC500M SATA SSD, Micron 5200 eco SATA SSD и DC1500M NVMe SSD](#)
 - [Результаты теста 2: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе Kingston DC500M SATA SSD, Micron 5200 eco SATA SSD и DC1500M NVMe SSD](#)
 - [Тест 3: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе DC1500M NVMe и Micron 5200 eco SATA, более крупная схема и увеличенная продолжительность испытаний](#)
 - [Результаты теста 3: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе DC1500M NVMe и Micron 5200 eco SATA, более крупная схема и увеличенная продолжительность испытаний](#)
 - [Тест 4: сравнение производительности SQL Server 2017, производительность операций резервного копирования и восстановления, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA](#)
 - [Тест 4: Результаты: сравнение производительности SQL Server 2017, производительность операций резервного копирования и восстановления, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA](#)
 - [Тест 5: сравнение производительности SQL Server 2017, тест Noisy Neighbor, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA](#)
 - [Результаты теста 5: сравнение производительности SQL Server 2017, тест Noisy Neighbor, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA](#)
- [Заключение](#)

Краткий обзор

В последние годы внедрение NVMe произвело революцию в области хранения данных, существенно увеличив максимальную производительность флеш-памяти NAND и используя выгоды обширного функционала, низкой стоимости, высокой пропускной способности и перспективного стандарта шины расширения PCI Express. 5^е поколение, PCIe Gen5 обеспечивает скорость передачи данных до 8 ГБ/с на дорожку, устраняя связанные с шиной расширения узкие места в стеке хранения и поддерживая инновации и развитие не только контроллеров SSD и флеш-памяти NAND, но и всего аппаратного стека. Процессоры, конструкции шасси, системные платы и топологии аппаратных подсистем ввода-вывода постоянно совершенствуются для поддержки дополнительной пропускной способности. В центрах обработки данных сетевые топологии претерпевают серьезные изменения, чтобы обеспечить возможность внедрения NVMe. После появления спецификации NVMe-OF сетевые интерфейсы, коммутаторы и транспортные протоколы изменились и продолжают совершенствоваться для поддержки увеличенной полосы пропускания при сохранении качества обслуживания и передачи пакетов без потерь.

Но как внедрение NVMe влияет на производительность приложений? Можно ли уменьшить физические размеры системы хранения, одновременно повысив пропускную способность и сократив время отклика транзакций? Можно ли значительно сократить время резервного копирования базы данных, чтобы частично устранить проблему «шумных соседей» и свести к минимуму ее влияние на производственную среду? В этой статье мы попытаемся ответить на эти вопросы, изучив типичные рабочие нагрузки OLTP, определенные в спецификации TPCC, и приведем ряд практических сравнений, чтобы показать влияние NVMe на производительность транзакций в реальных сценариях.

Типичные инфраструктурные проблемы, с которыми РСУБД сталкиваются сегодня в центрах обработки данных

Стоимость, планирование емкости и масштабируемость

На фоне колоссального увеличения пропускной способности Интернета, скоростей обработки и интереса к аналитике данных за последние 2 десятилетия производственные базы данных OLTP быстро растут. И часто намного быстрее, чем планировалось архитекторами приложений и инфраструктуры. Базовая архитектура системы хранения данных и сети должна быть построена с нуля с учетом возможности масштабирования, чтобы соответствовать росту требований с течением времени и обеспечивать хороший баланс между стоимостью, простотой управления и производительностью. Выбор между созданием приложений в локальных центрах обработки данных или использованием облачных сервисов IaaS/PaaS представляет собой сложный проектный выбор. Работа приложений в локальных центрах обработки данных дает архитекторам решений полный контроль над масштабируемостью, безопасностью, отказоустойчивостью и производительностью, но требует тщательного планирования и иногда сопряжена с высокими первоначальными затратами. Использование облачных сервисов IaaS/PaaS ускоряет развертывание и упрощает масштабирование, но обеспечивает меньший контроль над производительностью, отказоустойчивостью и его стоимость может быстро расти по мере масштабирования приложения. Некоторые организации предпочитают гибридный подход, при котором более важные приложения уровня 1 могут размещаться в локальных центрах обработки данных, а приложения уровня 2 и старые приложения переносятся в облако. Для приложений, которые выполняются в локальных системах, хороший баланс стоимости, простоты, производительности и простоты масштабирования обеспечивается с помощью решений гиперконвергентной инфраструктуры, таких как VMware vSAN с группами дисков на основе только флеш-памяти (All-Flash).

Отказоустойчивость

Приложения уровня 1 должны создаваться или переноситься в инфраструктуру, способную выдержать более одного аппаратного сбоя во всем аппаратном стеке. При неправильном планировании сбоя

оборудования в центрах обработки данных могут привести к значительным денежным потерям из-за перебоев в обслуживании или, в худшем случае, к безвозвратной потере данных. В средах с общей системой хранения данных необходимо тщательно планировать построение базовой инфраструктуры таким образом, чтобы выдерживать сбои системы хранения и избыточные нагрузки компонентов.

Например, при использовании vSAN приложения уровня 1 должны иметь минимальное допустимое количество отказов (FTT), равное 1, при включенной функции высокой доступности (HA) vSphere. Это гарантирует, что виртуальные машины приложения и базы данных защищены по крайней мере от отказа одного компьютера, сети или системы хранения. Кроме того, можно включить службу vSphere Distributed Resource Scheduler (DRS) для балансировки нагрузки на ресурсы процессора и памяти между физическими серверами в кластере.

Различные требования к производительности

Требования к более высокой скорости и меньшей задержке транзакций продолжает расти по мере масштабирования приложений OLTP, поскольку рост числа пользователей ведет к росту транзакционной нагрузки на серверную базу данных. Архитекторы приложений должны планировать инфраструктуру хранения данных, которая может адаптироваться к этому росту требований и является достаточно гибкой, чтобы данные можно было переносить между различными уровнями системы хранения. Например, базы данных SQL, которые размещаются на виртуальных дисках, предоставленных из массивов хранения SAN, можно перенести в хранилище данных vSAN на основе флеш-памяти NVMe, которое обеспечивает более быстродействующие уровни хранения, такие как NVMe, с помощью функции VMotion системы хранения VMware.

Дилемма «шумных соседей»

Крайне важно спроектировать инфраструктуру, которая сможет предоставлять необходимые ресурсы для ключевых рабочих нагрузок. В среде с общей системой хранения данных, в которой выполняется несколько рабочих нагрузок, производительность может стать непредсказуемой, а нестандартные рабочие нагрузки могут создавать проблемы для основных производственных рабочих нагрузок. Это определение проблемы «шумных соседей». Далее в этом документе мы рассматриваем такой пример, как незапланированные операции резервного копирования базы данных на одном сервере, потребляющие ресурсы системы хранения и сети и влияющие на производительность и задержку других серверов, использующих те же ресурсы.

Твердотельные накопители Kingston DC1500M Enterprise NVMe

Kingston DC1500M — это новейшее предложение накопителей Enterprise U.2 PCIe 3.0x4 NVMe от компании Kingston с вариантами емкости от 960 до 7680 ГБ. Этот накопитель, оснащенный 16-канальным контроллером и памятью 3D TLC NAND, разработан с учетом строгих требований к качеству обслуживания (QoS) для обеспечения стабильно высокой производительности и согласованности корпоративных рабочих нагрузок при минимальной задержке. Встроенное ПО, ориентированное на корпоративные задачи, поддерживает такие функции, как избыточное выделение ресурсов, несколько пространств имен (с поддержкой до 64 пространств имен), а также более сложные алгоритмы обнаружения и исправления ошибок (ECC) для обеспечения надежности рабочих нагрузок предприятия в течение всего срока службы накопителя.

Поскольку твердотельные накопители SATA по-прежнему наиболее распространены в центрах обработки данных, в этой статье мы стремимся показать, что миграция или создание инфраструктуры хранения данных на основе твердотельных накопителей Enterprise NVMe, таких как Kingston DC1500M NVMe, поможет решить ряд указанных выше проблем.

В ходе нашего внутреннего тестирования отдельный твердотельный накопитель Kingston DC1500M NVMe обеспечивал до 6,5 раз большую пропускную способность и снижение задержки в 5,6 раз (рис. b ниже) по сравнению с отдельным твердотельным накопителем Micron 5200 eco Enterprise SATA, практически при аналогичной стоимости.

Такой уровень производительности в гиперконвергентной среде обеспечивает более высокую пропускную способность и меньшую задержку транзакций для баз данных SQL Server. Это также позволяет создать систему хранения меньшего размера и с меньшим энергопотреблением. В этом примере требуется 6 накопителей Micron 5200 eco, чтобы обеспечить такую же пропускную способность, как 1 накопитель DC1500M. Позже мы увидим, как эта производительность проявляется в реалистичных рабочих нагрузках SQL OLTP в инфраструктуре VMware vSAN.

Значительное повышение производительности твердотельных накопителей NVMe, таких как DC1500M, по сравнению с твердотельными накопителями SATA также означает, что их внедрение в совместно используемые гиперконвергентные среды также позволяет снизить влияние проблемы «шумного соседа» на приложения уровня 1. Твердотельные накопители NVMe корпоративного уровня, такие как DC1500M, позволяют выполнять непредвиденные рабочие нагрузки, такие как операции резервного копирования и восстановления в рабочее время, гораздо быстрее, сохраняя при этом низкую задержку и высокую пропускную способность транзакций для критически важных производственных рабочих нагрузок уровня 1. Мы продемонстрируем это в тестах Noisy Neighbor далее в этом документе.

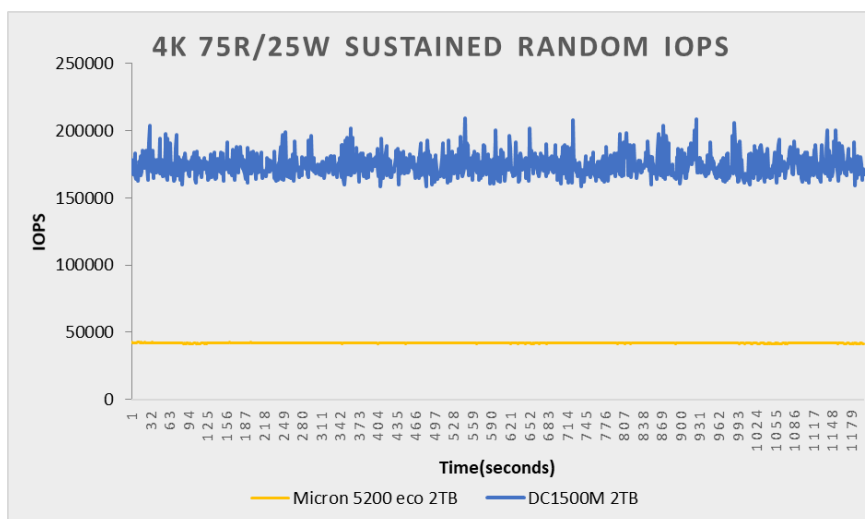


Рис. а. Посекундное сравнение числа операций ввода-вывода в секунду (IOPS) накопителей DC1500M 1920G и Micron 5200 ECO 1920G SATA SSD. Тестирование проводилось с одним физическим накопителем, подключенным в качестве вторичного к системе Linux с fio v3.17 после достижения твердотельными накопителями состояния стабильной производительности. При размере блока 4к, процентном количестве операций чтения 75% и глубине очереди 32.

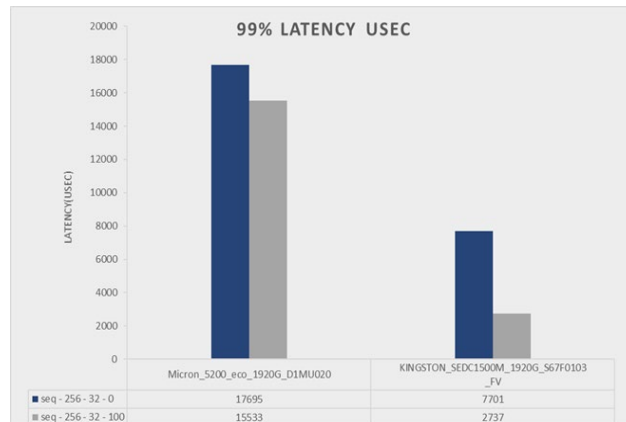
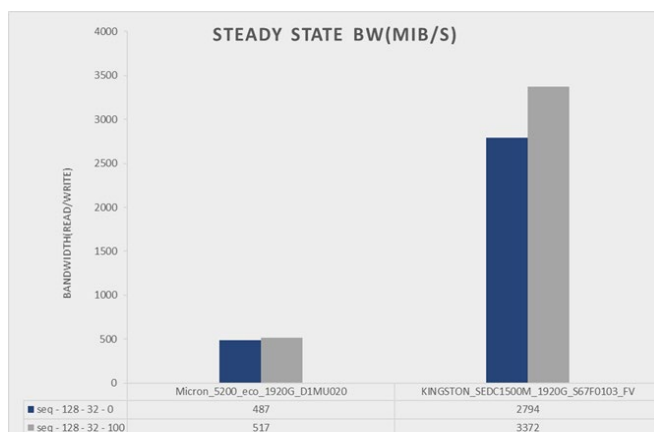


Рис. б. Сравнение пропускной способности (МБ/с) и задержки (мсек) при последовательном чтении/записи для накопителей DC1500M 1920G и Micron 5200 ECO 1920G SATA SSD. Тестирование проводилось с одним физическим накопителем, подключенным в качестве вторичного к системе Linux с fio v3.17 после достижения твердотельными накопителями состояния стабильной производительности. При размере блока 256к и глубине очереди 32.

Тестовая среда

I. Инфраструктура

Наши тестовые среды показаны на рисунках 1.1 и 1.2 ниже. Мы выбрали VMware vSAN в качестве гиперконвергентной инфраструктуры, поскольку это высокомасштабируемый, отказоустойчивый, централизованный и экономичный вариант системы хранения для гиперконвергентных виртуализированных сред.

VMware vSAN позволяет пользователям объединять локальные устройства хранения с нескольких серверов в единое хранилище данных, совместно используемое всеми хост-системами в кластере vSAN. Физические накопители каждого сервера разбиваются на группы. При этом 1 накопитель в группе используется в качестве устройства кэширования, а до 7 накопителей в группе — в качестве устройств для хранения. Сервер может иметь до 5 дисковых групп. Следовательно, в кластер vSAN может входить не более 35 устройств для хранения на каждый сервер. Дисковые группы всех хостов ESXi в кластере vSAN объединяются для создания хранилища данных vSAN. При этом трафик между хостами и хранилищем данных vSAN изолируется через выделенную сеть для vSAN (для vSAN на основе флеш-памяти требуется скорость 10 Гбит/с и выше). Это дает администраторам возможность начать с хранилища небольшого объема и добавлять узлы хранения для увеличения емкости (до 64 узлов на кластер) по мере необходимости. Также обеспечивается относительно простой способ контроля требований к производительности для конкретных виртуальных машин.

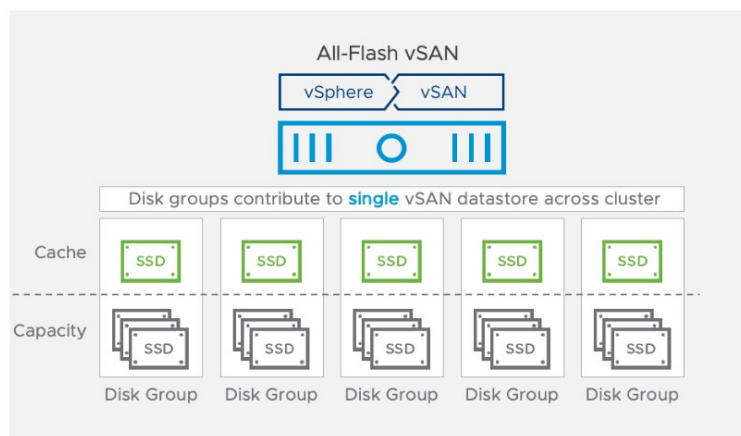


Рис. 1. Архитектура vSAN на основе флеш-памяти

vSAN использует политики хранения, чтобы определять уровень защиты и чередования для конкретных виртуальных дисков. Используя политику хранения по умолчанию, vSAN зеркально копирует все объекты, предоставленные из хранилища данных vSAN, а также предоставляет администраторам детальный контроль над уровнем защиты виртуальных дисков, предоставленных виртуальным машинам из хранилища данных vSAN. Например, чтобы диск с данными SQL VMDK мог выдержать по крайней мере один сбой в кластере (весь сервер, диск или сетевой интерфейс), мы можем указать первичный уровень

допустимого количества отказов (FTT), равный 1. Затем будет создана зеркальная копия RAID-1 объекта VMDK: с одной копией компонента на одном хосте и другой копией компонента — на другом хосте в кластере vSAN. Точно так же администраторы могут указать политику хранения RAID 0 (только чередование) с FTT, равным 0. В этом случае диск резервного копирования VMDK не будет обладать отказоустойчивостью и максимальной производительностью, а виртуальная машина будет обладать высокой доступностью благодаря SQL AlwaysOn Failover Clustering, или если резервное копирование базы данных регулярно выполняется с помощью стандартных решений для резервного копирования, таких как Commvault или NetBackup.

В лаборатории тестирования и проверки твердотельных накопителей Kingston Technology и для целей данного документа мы использовали 3 [сервера PowerEdge R740xD](#), поддерживающих 8 отсеков для 2,5-дюймовых накопителей NVMe и 16 отсеков для 2,5-дюймовых накопителей SATA/SAS на сервер, а выделенная сеть 10 Гбит/с поддерживалась 2 [коммутаторами Cisco Nexus 5k](#) для трафика vSAN, необходимого для тестирования твердотельных накопителей SATA. Мы использовали 4-узловой суперсервер Big Twin Supermicro [SYS-2029BT-HNR](#) с выделенной сетью 40 Гбит/с, поддерживаемой 1 [коммутатором Cisco 9k](#), для трафика vSAN, необходимого для тестирования накопителей NVMe. В ходе тестирования мы использовали настраиваемую политику хранения (FTT = 0), назначенную виртуальному диску гостевой виртуальной машины, чтобы обеспечить максимальную производительность блочного

хранилища во всех тестах, описанных в этом документе. Для различных проведенных нами тестов использовались разные твердотельные накопители, которые задокументированы в начале результатов каждого теста ниже, но в качестве стандарта мы использовали 3 физических накопителя с одинаковой емкостью на каждую группу дисков для тестирования как SATA, так и NVMe. Для сравнительного тестирования мы выбрали популярные накопители Micron 5200 e90 SATA SSD. Для управления и трафика VMotion была использована сеть 1 Гбит/с, поддерживаемая 1 24-портовым управляемым коммутатором Netgear JGS524PE.

Среда для тестирования NVMe (аппаратное обеспечение)	Среда для тестирования SATA/SAS/HYBRID (аппаратное обеспечение)
4-узловой кластер Supermicro SYS-2029BT-HNR с 6 отсеками для 2,5-дюймовых накопителей NVMe с возможностью «горячей» замены на сервер	3-узловой кластер PowerEdge Dell R740xD, поддерживающий 8 отсеков для 2,5-дюймовых накопителей NVMe и 16 отсеков для 2,5-дюймовых накопителей SATA/SAS на сервер
8 процессоров Intel(R) Xeon(R) Gold 6252 (48c/96t) с тактовой частотой 2,10 ГГц	8 процессоров Intel(R) Xeon(R) Silver 4114 (10c/20t) с тактовой частотой 2,20 ГГц
64 модуля памяти Kingston DDR4-2933 2Rx4 ECC REG DIMM емкостью 32 ГБ каждый (16x32 ГБ на узел), 512 ГБ на узел, 2048 ГБ на кластер	768 ГБ (24x32 ГБ) двухканальной памяти с ECC от Kingston, 2400 МГц на узел, 2304 ГБ на кластер
Два 20-портовых коммутатора Cisco Nexus N5K-C5010 10Gbe класса ЦОД для сетевого трафика vSAN	Один 32-портовый коммутатор Cisco Nexus 9332PQ Switch 40Gbe класса ЦОД, выделенный для сетевого трафика vSAN
	PERC H740P, настроенный в сквозном режиме адаптера главной шины (HBA)

Рис. 1.1. Аппаратное обеспечение, использованное во время тестов

Среда для тестирования NVMe (ОС и программное обеспечение)	Среда для тестирования SATA (ОС и программное обеспечение)
Гипервизор: VMware ESXi, 7.0.3, 19193900	Гипервизор: 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)
Гостевая ОС: Windows Server 2019 Data Center, v1809	Гостевая ОС: 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

Рис. 1.2. ОС и программное обеспечение

II. Конфигурация базы данных

В проведенных тестах мы использовали гостевую виртуальную машину Server 2019 с SQL Server 2017 и отдельный диск VMDK, предоставленный из хранилища данных vSAN для данных, журнала и резервного копирования. HammerDB — приложение с открытым исходным кодом для нагрузочного тестирования базы данных, поддерживающее выполнение теста TPCC для приложений OLTP и теста TPC-H для рабочей нагрузки по анализу данных. В различных тестах, описанных в этой статье, была выбрана эталонная спецификация TPCC для моделирования транзакционных рабочих нагрузок OLTP и обеспечения единообразия и надежности результатов тестирования.

Тест TPCC (формальное определение доступно на сайте tpc.org (домашняя страница TPCC)) — это хорошо известный стандартный отраслевой тест OLTP, который реализует компьютерную систему для выполнения заказов клиентов на поставку продуктов компании. Компания продает 100 000 товарных позиций и хранит запасы на складах. Каждый склад имеет 10 торговых районов, каждый из которых обслуживает 3000 клиентов. Клиенты звонят в компанию, операторы которой принимают заказ; каждый заказ содержит несколько позиций; затем заказы обычно выполняются с местного склада. Однако некоторых позиций в данный момент нет на складе, и они поставляются с альтернативного склада. Важно отметить, что размер компании не является фиксированным, и склады и торговые районы могут добавляться по мере роста компании. Поэтому тестовая схема может быть настолько маленькой или большой, насколько вы захотите. При этом более крупная схема приводит к увеличению базы данных TPC-C и требует более мощной компьютерной системы для обработки повышенного уровня транзакций (HammerDB).

Для этой статьи мы провели различные тесты; количество складов (размер схемы) и количество виртуальных пользователей задокументировано в начале каждого теста и объясняется в результатах теста. Во время всех тестовых прогонов мы записывали результаты Hammer DB для каждого прогона, одновременно собирая статистику по процессору, сети, памяти и дискам с помощью монитора производительности Windows (Perfmon) (с встроенным модулем Get-counter в Windows PowerShell) и монитора производительности vSAN, доступного на сервере vCenter.

III. Производительность системы хранения данных vSAN

Мы протестировали производительность хранилища данных vSAN для конфигураций, рассматриваемых в этом документе, перед выполнением тестов SQL, чтобы оценить уровень производительности, который можно ожидать от хранилища данных vSAN на основе накопителей DC1500M NVMe и Micron 5200 e9c SATA SSD. Мы использовали рекомендованное VMware средство для эталонного тестирования хранилища данных vSAN — [HCI Bench v2.5.3](#). Это набор инструментов для автоматизации, который развертывает несколько виртуальных машин, распределенных по всем узлам в кластере vSAN, при параллельном выполнении определенных рабочих нагрузок с использованием Vdbench на всех гостевых виртуальных машинах. Мы представляем ряд результатов прогона с 6 виртуальными машинами в кластере vSAN на основе DC1500M NVMe и на основе Micron 5200 e9c SATA.

На рис. 1.3 и 1.4 показаны результаты для смешанной рабочей нагрузки при устойчивой случайной рабочей нагрузке с 70 % операций чтения и 30 % операций записи с блоками различного размера в течение 30 минут для хранилища данных vSAN на основе DC1500M NVMe и хранилища данных vSAN на основе Micron 5200 e9c SATA SSD. При размере блока 4 КБ хранилище данных vSAN на основе DC1500M NVMe может обеспечить в 2 раза больше операций ввода-вывода в секунду при 70 % операций чтения/30 % операций записи (355 тыс. по сравнению с 178 тыс.), чем хранилище данных vSAN на основе SATA SSD, при этом каждая операция ввода-вывода выполняется на 33 % быстрее (0,4 мс по сравнению с 0,6 мс для vSAN на основе SATA SSD). Преимущество NVMe по производительности становится очевидным по мере увеличения размера передаваемых блоков в операциях ввода-вывода. Если рассмотреть случайную рабочую нагрузку с блоками размером 64 КБ и 70% операций чтения и 30% операций записи, хранилище данных vSAN на основе NVMe может обеспечить в 3 раза больше операций ввода-вывода в секунду (121240 по сравнению с 31756) и на 66% меньшую задержку на каждую операцию ввода-вывода (2,1 мс по сравнению с 6,4 мс для vSAN на основе SATA SSD).

На рисунках 1.5 и 1.6 показано сравнение устойчивой пропускной способности при чтении и записи и задержек для хранилища данных vSAN на основе DC1500M NVMe и на основе Micron 5200 e9c SATA SSD с различными размерами блоков. Мы смогли поддерживать пропускную способность 17,8 ГБ/с (размер блока 128 КБ) для хранилища данных на основе DC1500M NVMe. Это в 6,3 раза выше пропускной способности для операций чтения, обеспечиваемой хранилищем данных vSAN на основе SATA SSD (2,79 ГБ/с). Задержка при этом была в 5 раз ниже (0,9 мс по сравнению с 4,4 мс для vSAN на основе SATA). Для операций записи vSAN на основе DC1500M поддерживает пропускную способность 6,7 ГБ/с (размер блоков 128 КБ), что также в 5,9 раза выше по сравнению с vSAN на основе SATA. При этом задержка меньше в 5 раз.

Насколько эта разница в производительности для необработанных данных между хранилищем данных vSAN на основе NVMe и на основе SATA масштабируется, когда речь идет о производительности SQL?

Оправдываются ли затраты преимуществом NVMe по производительности? Будут ли операции резервного копирования или восстановления SQL выполняться быстрее, чтобы снизить влияние на критически важные рабочие нагрузки? В следующих разделах мы попытаемся ответить на этот вопрос, проведя несколько экспериментов.

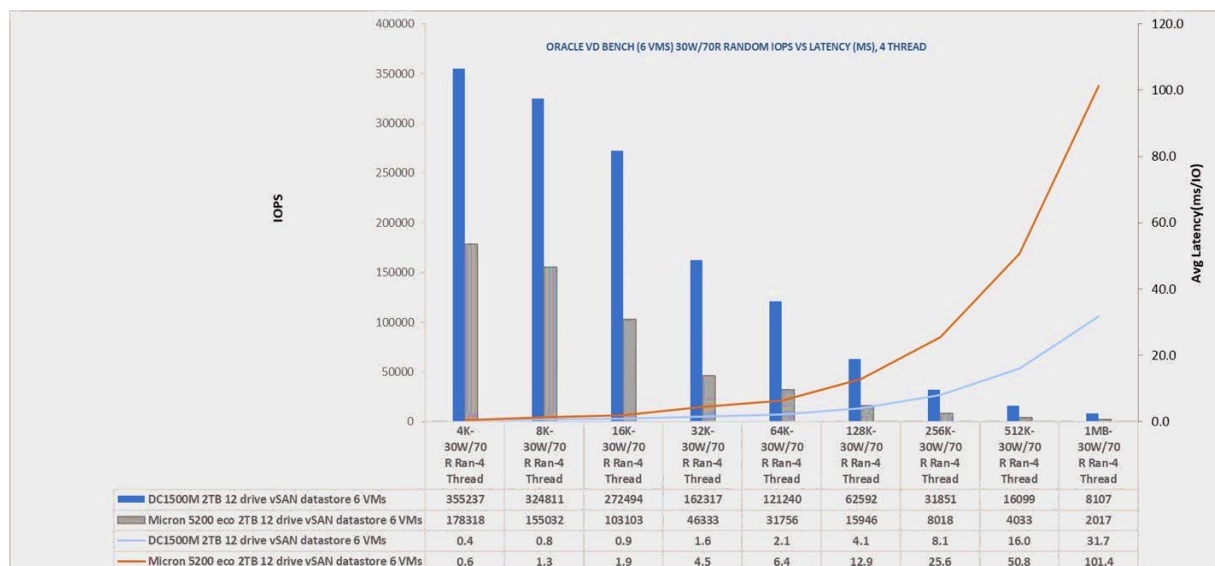


Рис. 1.3. Сравнение хранилища данных vSAN на основе DC1500M и хранилища данных vSAN на основе Micron 5200 eco, размер блока 4к, 70% операций чтения/30% операций записи, случайная рабочая нагрузка, QD=8, число потоков=4, 6 виртуальных машин HClBench, число операций ввода-вывода в секунду (IOPS) по сравнению со средней задержкой (мс)

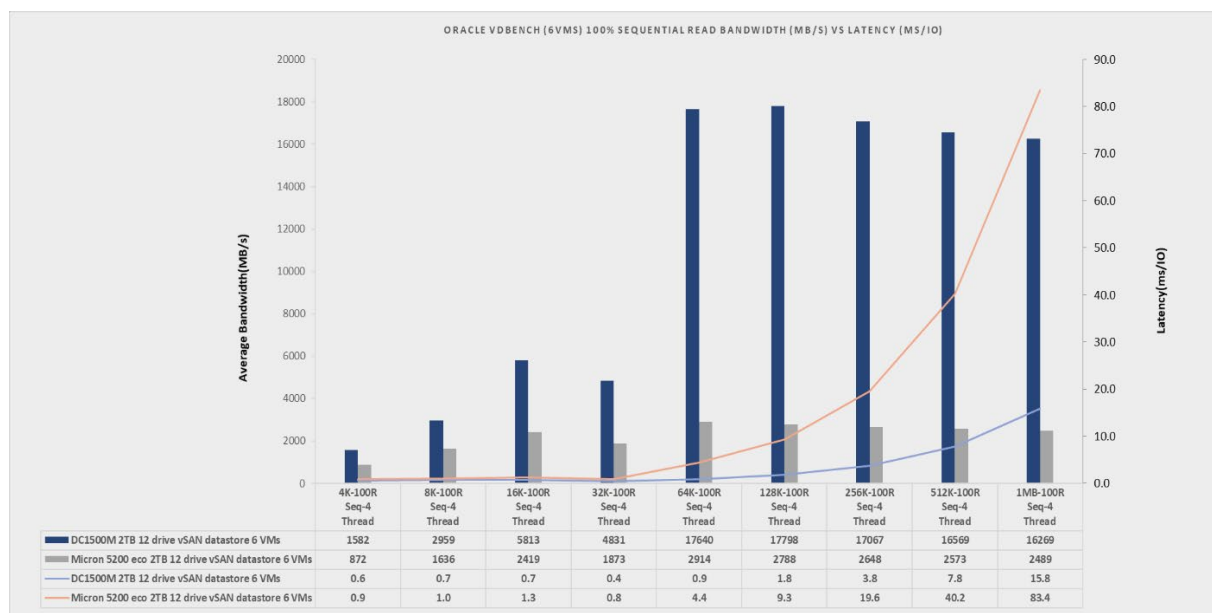


Рис. 1.4. Сравнение хранилища данных vSAN на основе DC1500M и хранилища данных vSAN на основе Micron 5200 eco, 100% операций чтения/0% операций записи, последовательные операции, QD=8, число потоков=4, 6 виртуальных машин HClBench, пропускная способность (МБ/с) и средняя задержка (мс на операцию ввода-вывода) для операций чтения

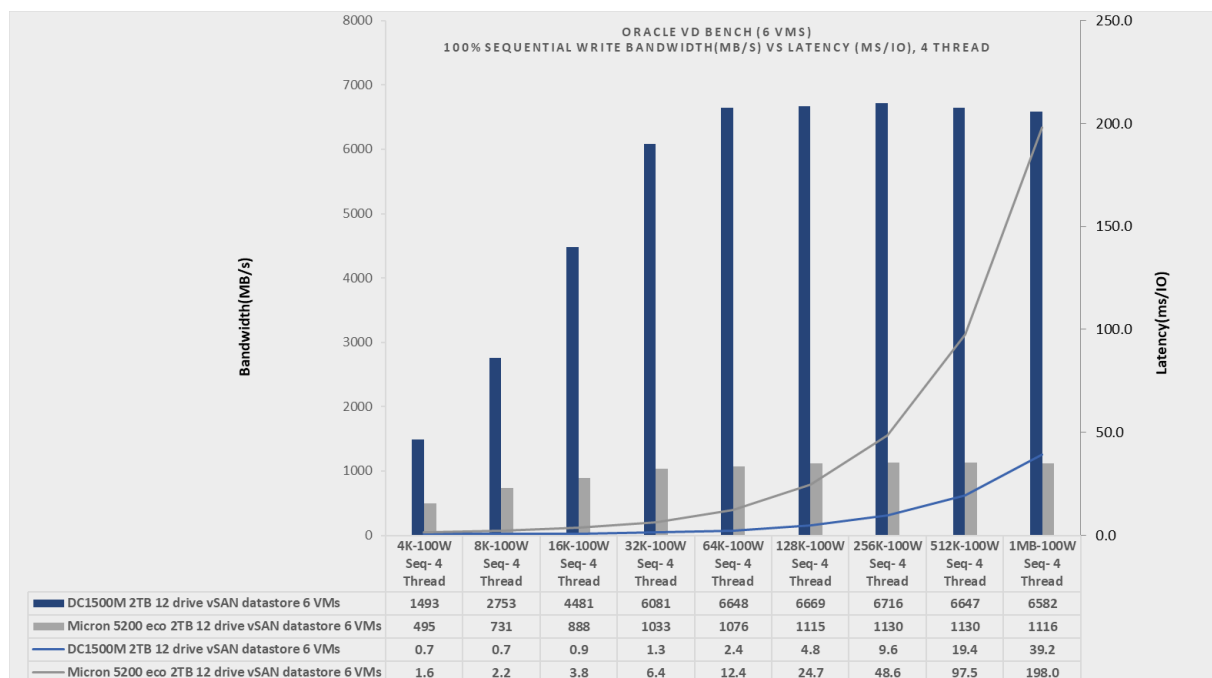


Рис. 1.5. Сравнение хранилища данных vSAN на основе DC1500M и хранилища данных vSAN на основе Micron 5200 eco, 100% операций записи/0% операций чтения, последовательные операции, QD=8, число потоков=4, 6 виртуальных машин HClBench, пропускная способность (МБ/с) и средняя задержка (мс на операцию ввода-вывода) для операций чтения

Результаты испытаний

Тест 1, SQL server 2017 на виртуальной машине в vSAN на базе DC1500M 960GB с различными объемами памяти DRAM

Конфигурация хранилища данных vSAN: 3 накопителя DC1500M 960G FW S67F0103 на дисковую группу, всего 4 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе NVMe. SQL Server 2017 с гостевой ОС Server 2019 Datacenter.		
Описание теста 1a	Описание теста 1b	Описание теста 1c
Виртуальный диск предоставлен из хранилища данных vSAN на основе DC1500M в среде для тестирования NVMe. Была выбрана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 128 ГБ ОЗУ. Другая виртуальная машина vSAN с 16 ядрами и 128 ГБ ОЗУ была выделена для работы в качестве сервера генерации нагрузки, который должен отправлять транзакции в тестируемую систему. Была создана следующая последовательность виртуальных пользователей: 1,2,3,5,8,13,21,34,55,89. Были выбраны время разгона 2 минуты и длительность тестовой последовательности 5 мин на пользователя.	Аналогично тесту 1a; но объем DRAM, выделенный для гостевой виртуальной машины, был уменьшен до 32 ГБ, чтобы увеличить число операций ввода-вывода для отдельной области данных. Сервер удаленной генерации нагрузки по-прежнему использовался для отправки транзакций в тестируемую систему, но выделенный для него объем DRAM также был уменьшен до 32 ГБ.	Аналогично тесту 1a; но объем DRAM, выделенный для гостевой виртуальной машины, был уменьшен до 32 ГБ, чтобы увеличить число операций ввода-вывода для отдельной области данных, а сам тест выполнялся локально на виртуальной машине тестируемой системы, чтобы устранить «узкие» места сети.

Рис. 2.1. Тест 1: Различные конфигурации DRAM хранилища данных vSAN на основе DC1500M

Цель теста 1 состояла в том, чтобы получить ожидаемый базовый уровень производительности с помощью эталонного теста TPCC для SQL Server 2017 в VMware vSAN с хранилищем данных vSAN на основе только флеш-памяти DC1500M NVMe с различными объемами памяти, выделенными для SQL-сервера. Смысл изменения объема DRAM, выделенного для тестируемой системы SQL (SUT), основан на следующих концепциях:

- Сокращение ОЗУ, выделенного для виртуальной машины базы данных SQL-сервера, приведет к увеличению числа операций ввода-вывода в область данных и подчеркнет производительность ввода-вывода базы данных, содержащей данную схему (база данных OLTP на диске).
- Если в виртуальной машине базы данных SQL-сервера достаточно DRAM, большая часть данных будет кэшироваться во время теста OLTP, а ввод-вывод в область данных будет минимальным (тест OLTP в памяти).

Мы создали схему размером 1200 складов, в результате чего размер базы данных TPCC составил примерно 100 ГБ. В первом тесте мы выделили для тестируемой системы 128 ГБ DRAM, поэтому вся схема может поместиться в оперативной памяти. Затем мы запустили последовательность виртуальных пользователей на удаленном сервере генерации нагрузки (LGS) для имитации пользователей, отправляющих транзакции в базу данных, с масштабированием от 1 до 89 пользователей, чтобы соответствовать размеру схемы и количеству ресурсов процессора и памяти, выделенных для виртуальной машины SQL-сервера. После завершения теста мы восстановили базу данных TPCC, затем уменьшили выделенный объем DRAM до 32 ГБ на тестируемую систему и сервер удаленной генерации нагрузки и повторно провели тот же тест с той же последовательностью пользователей. Наконец, мы выполнили тот же тест локально на виртуальной машине тестируемой системы, чтобы устранить любые «узкие» места в сети, вызванные сервером удаленной генерации нагрузки.

Результаты теста 1, SQL server 2017 на виртуальной машине в vSAN на базе DC1500M 960GB с различными объемами памяти DRAM

На рисунках 2.2 и 2.3 показаны количество транзакций в минуту (TPM) и количество новых заказов в минуту (NOPM), полученные в тестах 1a, 1b, 1c с использованием хранилища данных vSAN на основе DC1500M. Во всех тестовых прогонах мы наблюдаем увеличение показателей TPM и NOPM по мере увеличения числа виртуальных пользователей. При 89 виртуальных пользователях виртуальная машина SQL Server 2017 с базой данных OLTP, которая в основном находится в памяти, может достичь 1 113 300 TPM при 259 631 NOPM. Когда мы уменьшили объем DRAM, выделенный для виртуальной машины тестируемой системы и сервера удаленной генерации нагрузки, до 32 ГБ, мы смогли достичь показателей в 958 338 TPM и 208 311 NOPM, но когда мы запустили тест локально на виртуальной машине тестируемой системы, то достигли феноменальных результатов: 1 463 290 TPM и 318 092 NOPM!

Именно здесь мы видим преимущество твердотельных накопителей NVMe корпоративного класса в действии. Это означает, что при выделении недостаточной памяти для кэширования схемы, когда количество транзакций растет и база данных SQL-сервера должна записывать данные из памяти в файл журнала транзакций, виртуальный диск NVMe может реагировать достаточно быстро, чтобы поддерживать более высокую пропускную способность транзакций и увеличивать ее до тех пор, пока процессор не станет узким местом. На рис. 2.4 в тесте 1c видно, что даже при 89 виртуальных пользователях каждый пользователь может обрабатывать 16 441 транзакций в минуту. Основываясь на этих эмпирических результатах, мы можем сделать вывод, что построение базы данных на базе гиперконвергентной инфраструктуры на основе накопителей NVMe позволяет снизить расходы на дополнительную память DRAM, выделяемую для SQL Server 2017.

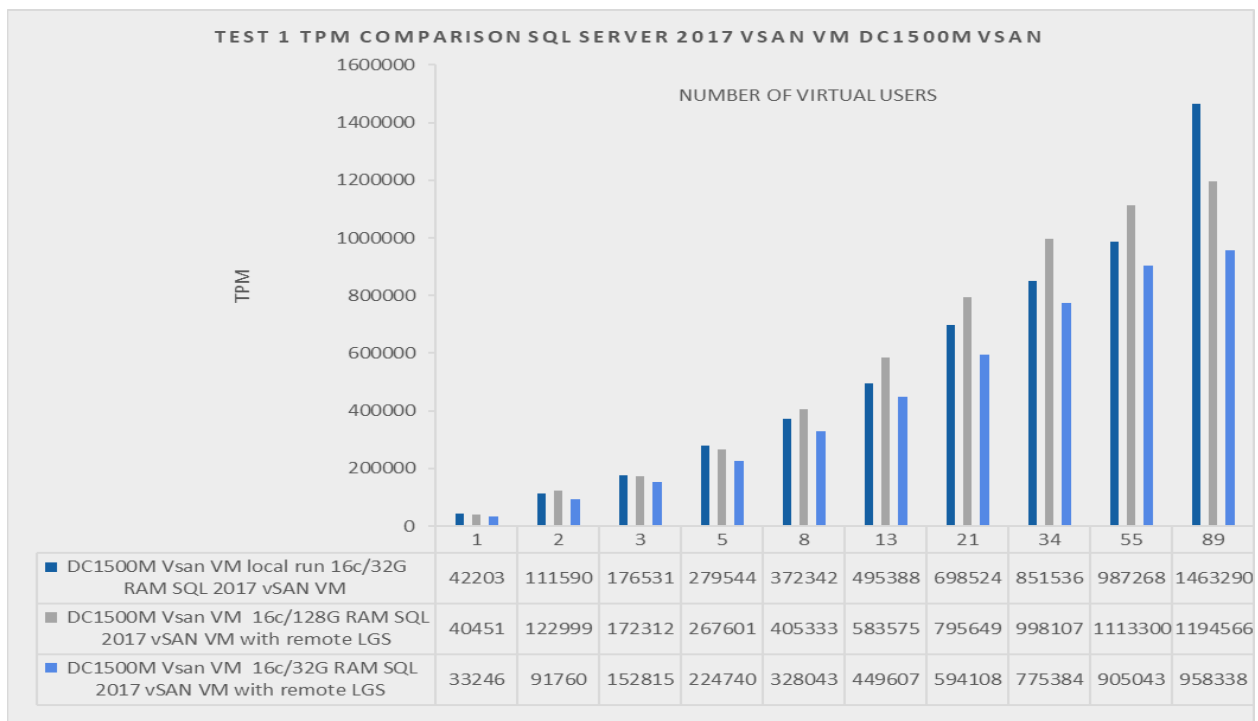


Рис. 2.2. Тесты 1а,б,с: сравнение TPM хранилища данных vSAN на основе DC1500M при различных объемах DRAM

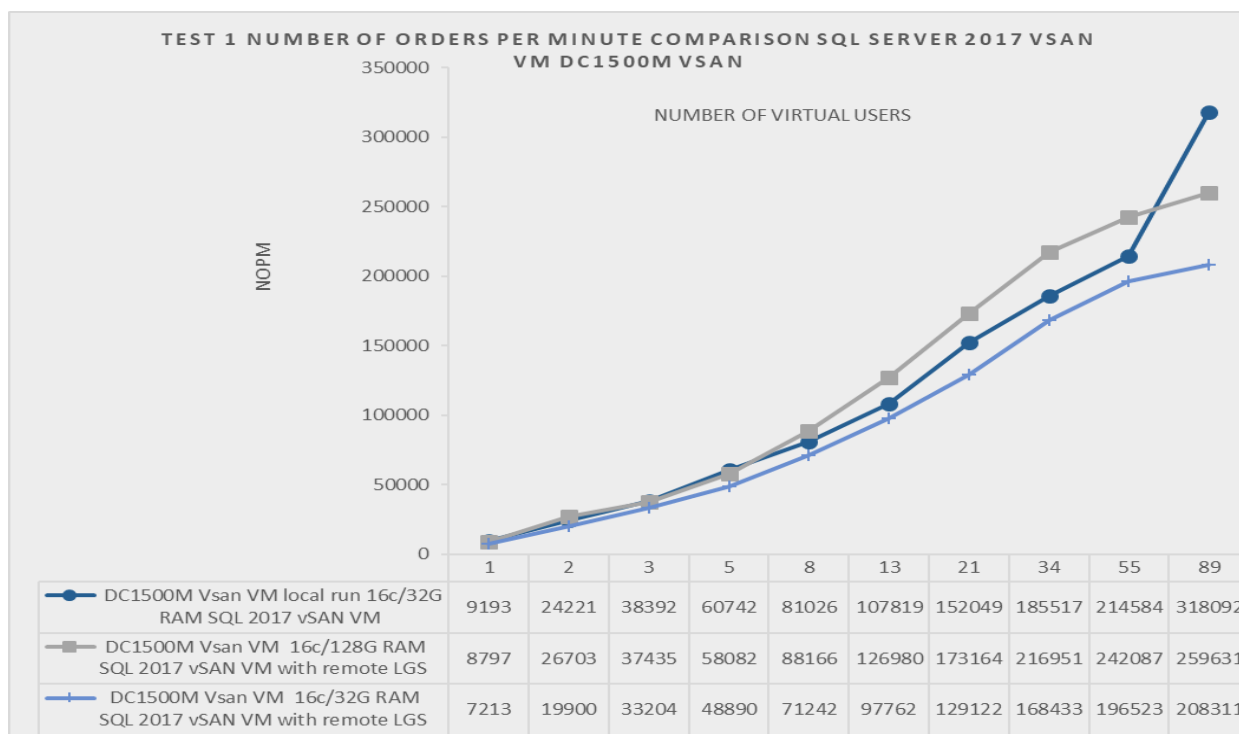


Рис. 2.3. Тесты 1а,б,с: сравнение NOPM хранилища данных vSAN на основе DC1500M при различных объемах DRAM

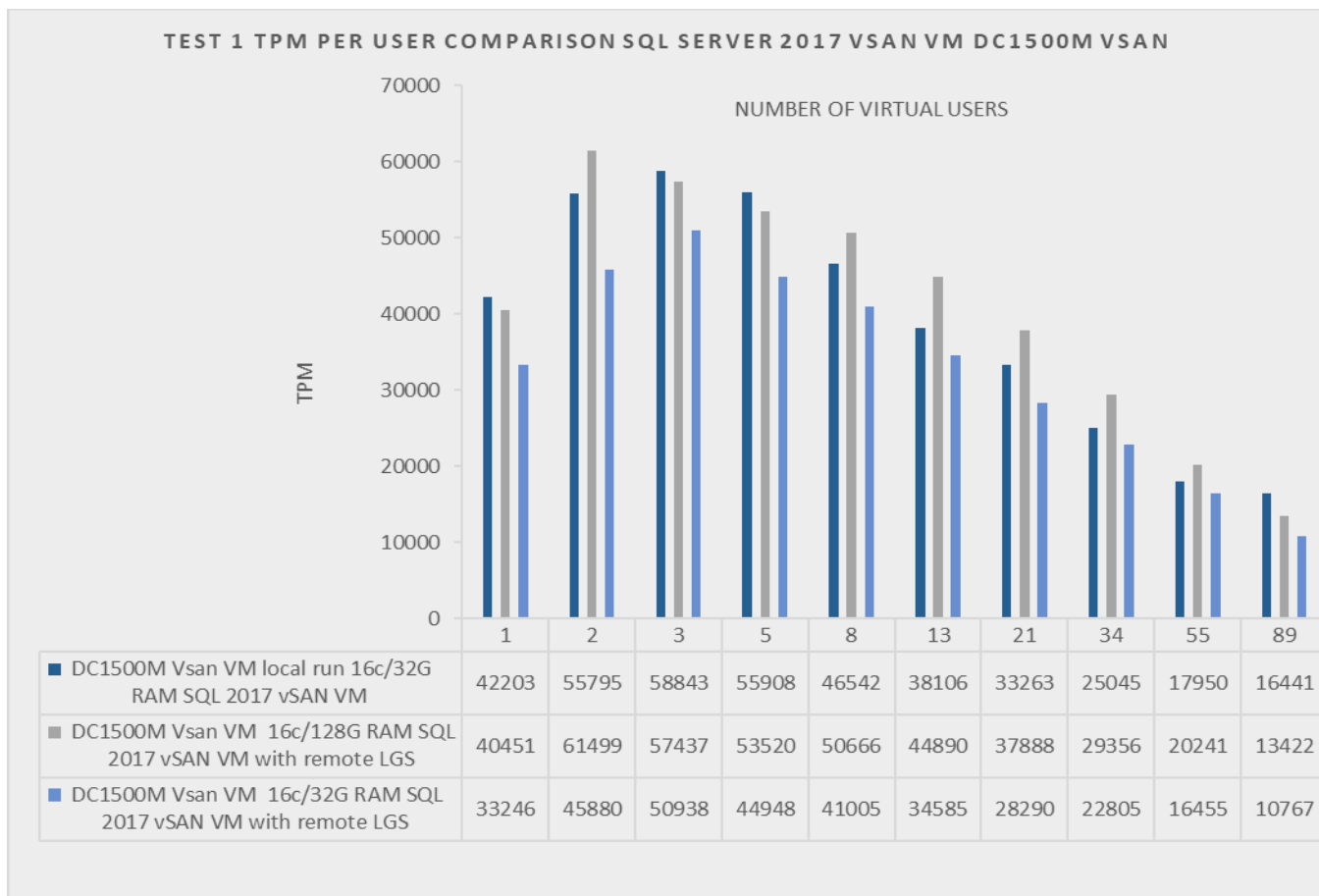


Рис. 2.4. Тесты 1а,б,с: сравнение TPM хранилища данных vSAN на основе DC1500M при различных объемах DRAM

Тест 2: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе Kingston DC500M SATA SSD, Micron 5200 eco SATA SSD и DC1500M NVMe SSD

• Конфигурация хранилища данных vSAN на основе NVMe для теста 1a: 3 накопителя DC1500M 960G FW S67F0103 на дисковую группу, всего 4 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе NVMe. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 1a) • Конфигурация хранилища данных vSAN на основе SATA для теста 1b: 3 накопителя DC500M 1920G FW SCEJK2.8 на дисковую группу, всего 3 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе SATA. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 1b) • Конфигурация хранилища данных vSAN на основе SATA для теста 1c: 3 накопителя Micron 5200 ECO 1920G FW D1MU004 на дисковую группу, всего 3 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе SATA. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 1b)		
Описание теста 2a	Описание теста 2b	Описание теста 2c
Виртуальный диск предоставлен из хранилища данных vSAN на основе DC1500M в среде для тестирования NVMe. Была выбрана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Была создана следующая последовательность виртуальных пользователей: 1,2,3,5,8,13,21,34,55,89. Были выбраны время разгона 2 минуты и длительность тестовой последовательности 5 мин на пользователя. Тест выполнялся локально на виртуальной машине тестируемой системы.	Виртуальный диск предоставлен из хранилища данных vSAN на основе D500M в среде для тестирования SATA. Была выбрана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Была создана следующая последовательность виртуальных пользователей: 1,2,3,5,8,13,21,34,55,89. Были выбраны время разгона 2 минуты и длительность тестовой последовательности 5 мин на пользователя. Тест выполнялся локально на виртуальной машине тестируемой системы.	Виртуальный диск предоставлен из хранилища данных vSAN на основе Micron 5200 eco в среде для тестирования SATA. Была выбрана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Была создана следующая последовательность виртуальных пользователей: 1,2,3,5,8,13,21,34,55,89. Были выбраны время разгона 2 минуты и длительность тестовой последовательности 5 мин на пользователя. Тест выполнялся локально на виртуальной машине тестируемой системы.

Рис. 3.1. Описание теста 2: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на основе SATA и на основе DC1500M NVMe

В тесте 2 сравнивается производительность эталонного теста TPCC для виртуальной машины тестируемой системы SQL Server 2017 при локальном выполнении в 3 различных хранилищах данных: хранилища данных vSAN на основе Kingston DC1500M Enterprise NVMe, на основе [Kingston DC500M](#) и на основе Micron 5200 eco SATA SSD. Тест 2 выполнялся локально на виртуальной машине тестируемой системы SQL Server 2017, чтобы увеличить число операций ввода-вывода в область данных и обратить внимание на производительность ввода-вывода базы данных, содержащей схему, а также проверить последовательность пользователей, масштабируемую от 1 до 89 пользователей, и согласовать размер схемы и количество выделенных ресурсов процессора и памяти для виртуальной машины SQL-сервера.

Результаты теста 2: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе Kingston DC500M SATA SSD, Micron 5200 eco SATA SSD и DC1500M NVMe SSD

На рисунках 3.2 и 3.3 показаны количество транзакций в минуту (TPM) и новых заказов в минуту (NOPM), полученные в тестах 2a, 2b и 2c. Во всех тестовых прогонах мы наблюдаем рост показателей TPM и NOPM по мере увеличения числа виртуальных пользователей, но масштабирование существенно отличается для NVMe и SATA. При 89 виртуальных пользователях виртуальная машина

SQL Server 2017 в хранилище данных vSAN на основе DC1500M может достичь 1 463 290 TPM при 318 092 NOPM. Для сравнения, мы достигли 738 067 TPM/160 410 NOPM для виртуальной машины SQL Server в vSAN на основе DC500M и 628 499 TPM/136 436 NOPM для хранилища данных vSAN на основе Micron 5200 есо. Это означает, что, используя одно и то же количество накопителей DC1500M NVMe в хранилище данных vSAN на основе NVMe можно эффективно удвоить пропускную способность транзакций и количество заказов в минуту по сравнению с хранилищем данных vSAN на основе SATA с таким же количеством твердотельных накопителей. В бизнес-контексте это означает, что если у вас есть 89 пользователей, одновременно отправляющих транзакции в базу данных, то каждый из них может обрабатывать на 235 % больше транзакций (что означает большее количество заказов в минуту) (рис. 3.4), если перейти к инфраструктуре VMware на основе накопителей NVMe корпоративного класса, таких как DC1500M.

На рис. 3.5 показано среднее время простоя процессора в зависимости от числа виртуальных пользователей в тестах 2а, б и с. Это хорошая мера эффективности работы виртуального диска — насколько быстро виртуальный диск может реагировать на увеличение количества транзакций, когда база данных SQL Server должна записывать данные из памяти в файл журнала транзакций. При 89 виртуальных пользователях время простоя процессора (iowait) для виртуальной машины vSAN на основе DC1500M NVMe составляет 15,5 % по сравнению с 37,8 % для виртуальной машины на основе DC500M и 44,2 % для виртуальной машины на основе Micron 5200. Это означает, что виртуальный диск на основе NVMe намного быстрее отвечает на запросы ввода-вывода, предотвращая простои процессора в ожидании завершения операции ввода-вывода и позволяя обрабатывать больше транзакций. В бизнес-контексте переход к инфраструктуре VMware на основе NVMe позволяет более эффективно использовать назначенные виртуальные ядра для виртуальной машины SQL Server, чтобы повысить пропускную способность транзакций и снизить затраты за счет удаления ненужных ядер из прежних виртуальных машин SQL, работающих на более медленных уровнях хранения.

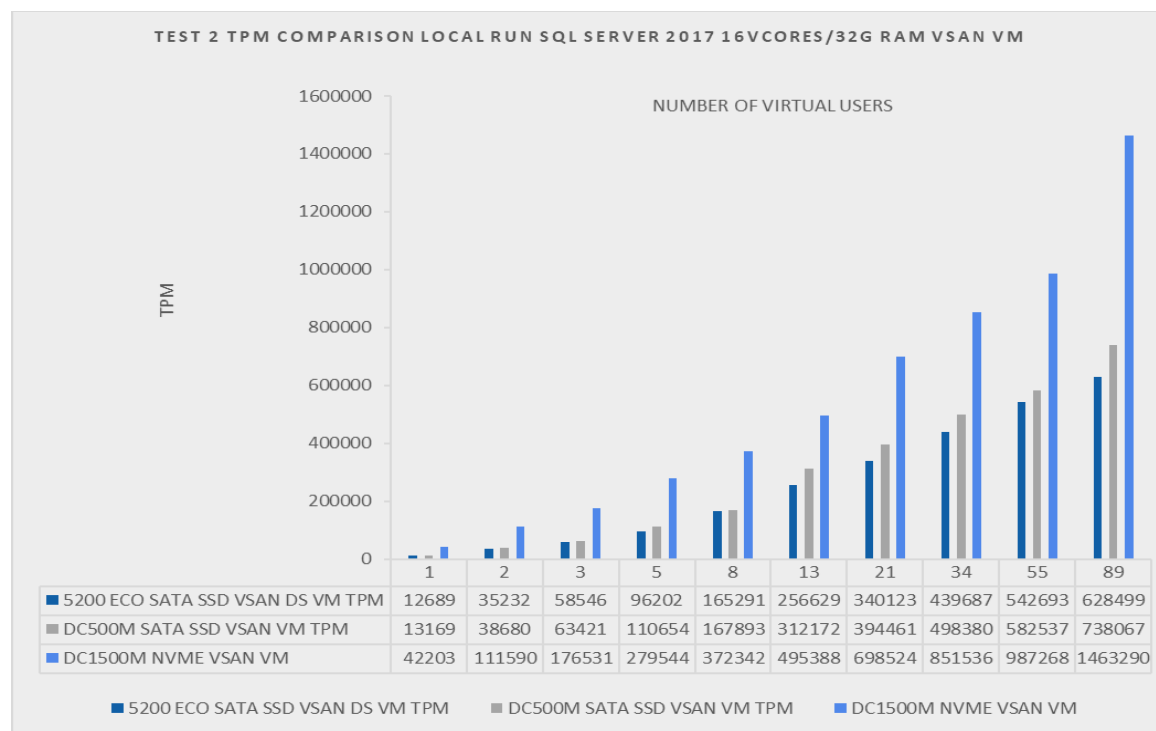


Рис. 3.2. Тест 2: сравнение TPM для хранилищ данных vSAN на основе NVMe и SATA



Рис. 3.3. Тест 2: сравнение NOPM для хранилищ данных VSAN на основе NVME и SATA

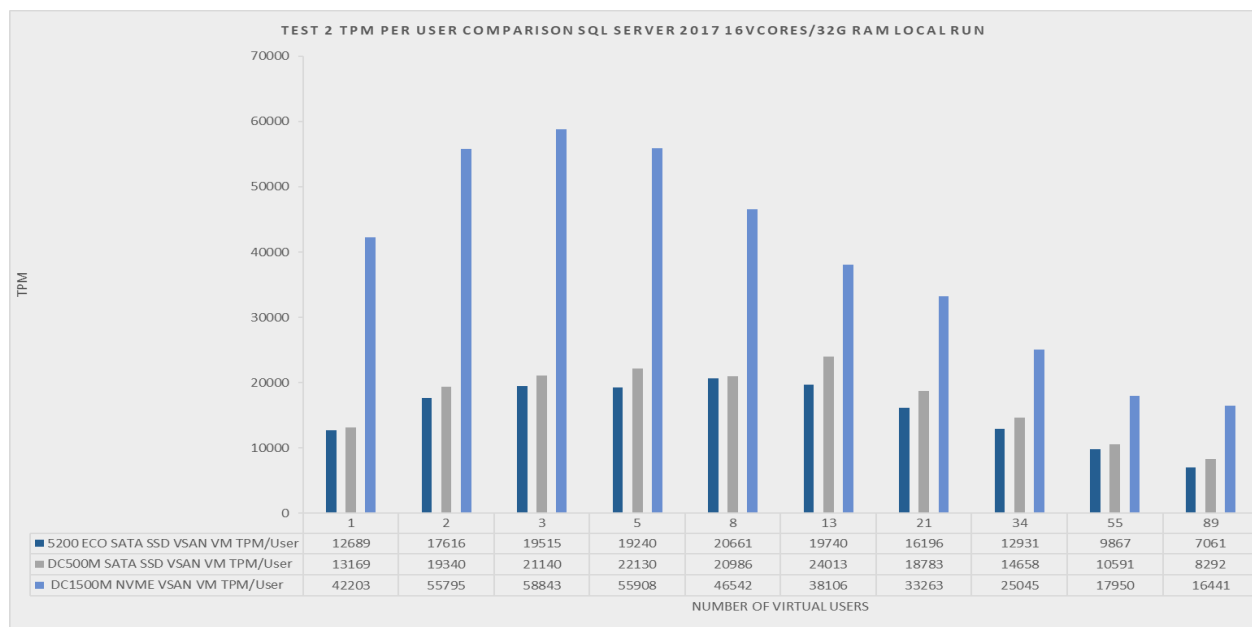


Рис. 3.4. Тест 2: сравнение TPM на пользователя для хранилищ данных VSAN на основе NVME и SATA

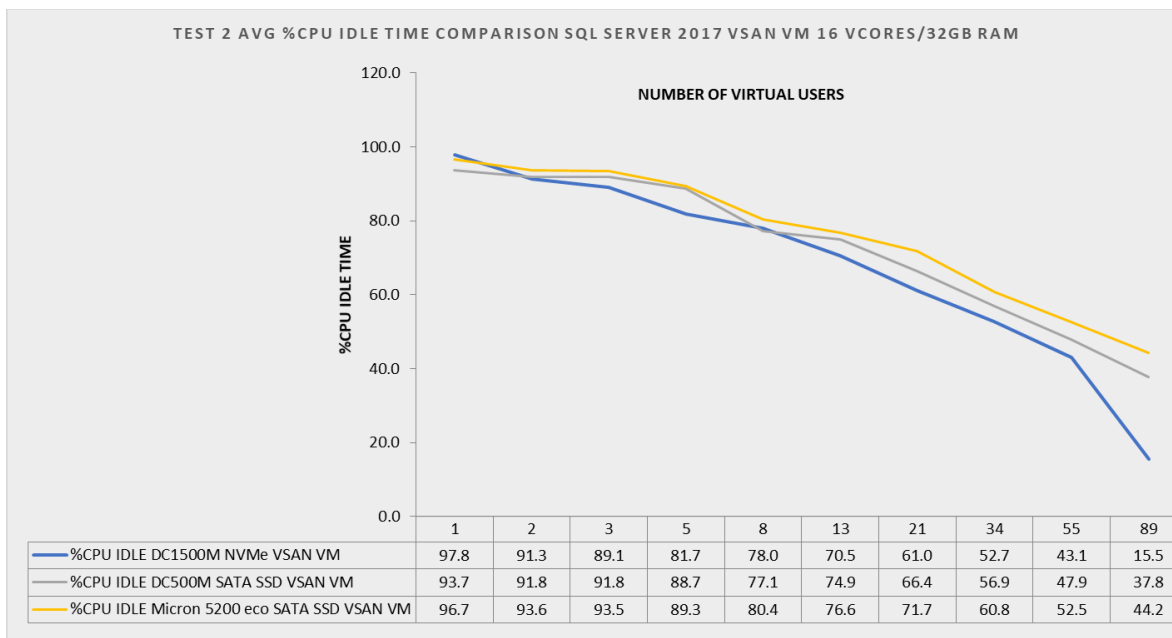


Рис. 3.5. Тест 2: сравнение % проста процессора для хранилищ данных VSAN на основе NVMe и SATA

Тест 3: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе DC1500M NVMe и Micron 5200 eco SATA, более крупная схема и увеличенная продолжительность испытаний

- Конфигурация хранилища данных vSAN на основе NVMe для теста 3a: 3 накопителя DC1500M 960G FW S67F0103 на дисковую группу, всего 4 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе NVMe. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 3a) - Конфигурация хранилища данных vSAN на основе SATA для теста 3b: 3 накопителя Micron 5200 ECO 1920G FW D1MU004 на дисковую группу, всего 3 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе SATA. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 3a)	
Описание теста 3a	Описание теста 2b
Виртуальный диск предоставлен из хранилища данных vSAN на основе DC1500M в среде для тестирования NVMe. Была выбрана схема базы данных с 2000 складами, представляющая базу данных размером 157 Гб. Виртуальной машине тестируемой системы (SUT) было выделено 40 виртуальных ядер и 32 Гб ОЗУ. Была создана следующая последовательность виртуальных пользователей: 1,2,4,8,16,32,64,89,128. Были выбраны время разгона 10 минут и длительность тестовой последовательности 20 минут на пользователя. Тест выполнялся локально на виртуальной машине тестируемой системы.	Виртуальный диск предоставлен из хранилища данных vSAN на основе Micron 5200 eco в среде для тестирования SATA. Была выбрана схема базы данных с 2000 складами, представляющая базу данных размером 157 Гб. Виртуальной машине тестируемой системы (SUT) было выделено 40 виртуальных ядер и 32 Гб ОЗУ. Была создана следующая последовательность виртуальных пользователей: 1,2,4,8,16,32,64,89,128. Были выбраны время разгона 10 минут и длительность тестовой последовательности 20 минут на пользователя. Тест выполнялся локально на виртуальной машине тестируемой системы.

Рис. 4.1. Описание теста 3: нагрузочный тест базы данных SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSD

Этот тест был разработан как более длительный нагрузочный тест с большим размером схемы базы данных для проверки предыдущих результатов и сравнения производительности в эталонном тесте TPCC для виртуальной машины тестируемой системы SQL Server 2017 при локальном выполнении на 2 разных хранилищах данных: vSAN на основе Kingston DC1500M Enterprise NVMe и vSAN на основе Micron 5200 eco SATA SSD. На этот раз мы выбрали схему, содержащую 2000 складов, в результате чего размер базы данных TPC-C составил 157 ГБ. Мы использовали 40 виртуальных ядер для каждой виртуальной машины SQL Server, чтобы выделить достаточно ресурсов процессора для генерации большого числа транзакций и насыщения пропускной способности транзакций, но выделили только 32 ГБ ОЗУ, чтобы ограничить тестовый ввод-вывод. Мы немного изменили последовательность виртуальных пользователей, чтобы увеличить количество пользователей с 1 до 128, и позволили каждой последовательности виртуальных пользователей работать в течение гораздо более длительного времени (20 минут с 10-минутным временем разгона). Это позволило нам собирать показатели задержки работы диска в течение всего тестового прогона.

Результаты теста 3: сравнение производительности SQL Server 2017 в хранилищах данных vSAN на базе DC1500M NVMe и Micron 5200 eco SATA, более крупная схема и увеличенная продолжительность испытаний

На рисунках 4.2 и 4.3 показаны количество транзакций в минуту (TPM) и новых заказов в минуту (NOPM), полученные в тестах 3a и 3b. Даже при большей продолжительности виртуальные машины SQL Server 2017 на основе твердотельных накопителей NVMe и SATA могут масштабироваться по мере увеличения числа виртуальных пользователей до 128, но градиент увеличения для NVMe намного выше. Для 89 пользователей мы достигли 1,84 млн TPM по сравнению с 0,96 млн TPM и 361 743 NOPM по сравнению с 184 451 NOPM для виртуальной машины SQL vSAN на основе SATA SSD. Это на 200 % больший показатель TPM/NOPM для хранилища данных vSAN на основе DC1500M NVMe по сравнению с хранилищем данных vSAN на основе Micron 5200 eco при том же количестве виртуальных ядер и выделенном объеме DRAM.

На рисунках 4.4 и 4.5 показана средняя задержка виртуального диска и задержка виртуального диска при выполнении 99 % операций при различном числе пользователей. Данные получены с помощью Windows Perfmon на виртуальных машинах SQL в хранилище данных vSAN на основе NVMe и на основе SATA SSD. Для каждой последовательности виртуальных пользователей виртуальный диск на основе DC1500M может поддерживать среднюю задержку менее 1 мс, даже если количество пользователей продолжает увеличиваться. При 89 виртуальных пользователях виртуальный диск на основе DC1500M демонстрировал среднюю задержку ввода-вывода 0,92 мс по сравнению с 2,36 мс для виртуального диска на основе твердотельных накопителей SATA, что на 256 % выше по сравнению с NVMe. Более интересна задержка QoS 99%: при 89 пользователях виртуальный диск на основе DC1500M мог выполнить 99% всех операций ввода-вывода за 1,61 мс, но виртуальный диск на основе SATA SSD выполнял 99% всех операций ввода-вывода за 7,05 мс, что на 437 % выше по сравнению с NVMe. Здесь подчеркивается разница в задержке между NVMe и SATA, и, поскольку накопитель DC1500M спроектирован так, чтобы поддерживать предсказуемую задержку QoS при выполнении устойчивых рабочих нагрузок OLTP, мы не наблюдаем внезапных всплесков задержки, даже когда число виртуальных пользователей увеличивается, и растет количество параллельных операций ввода-вывода на блочном уровне. С точки зрения бизнеса это означает, что переход от инфраструктуры VMware на основе накопителей SATA к инфраструктуре на основе накопителей NVMe корпоративного уровня, таких как DC1500M, позволяет увеличивать число транзакций и значительно снизить задержку при их выполнении, обеспечивая возможность быстрого масштабирования приложений и снижения затрат с течением времени.

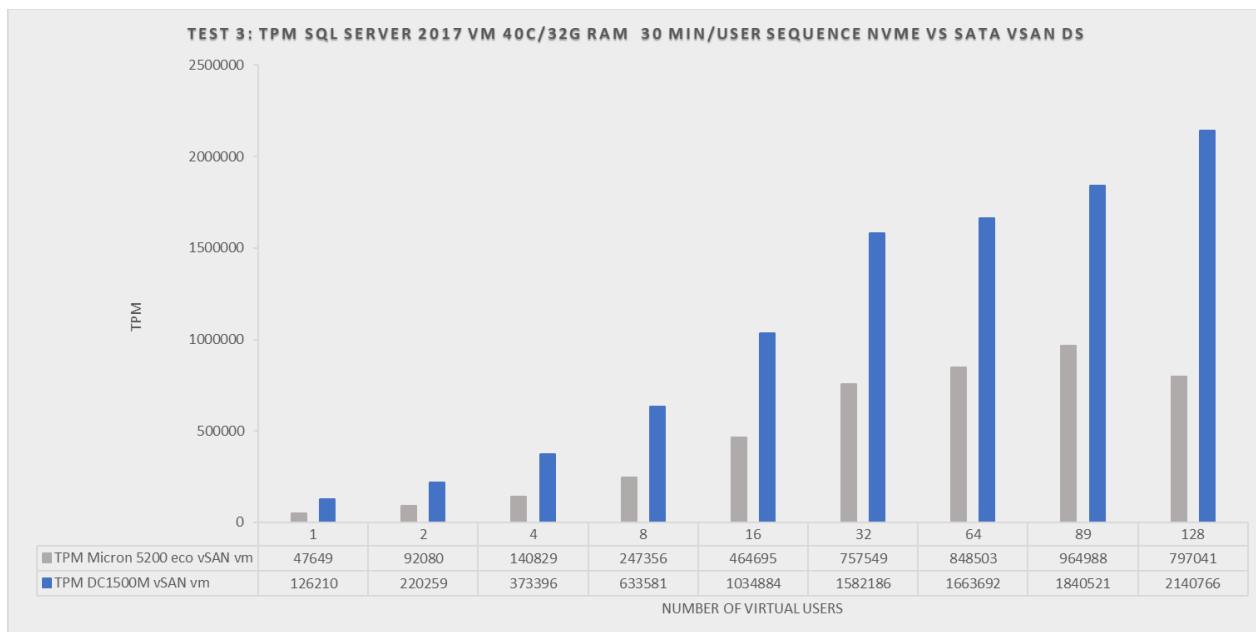


Рис. 4.2. Тест 3: сравнение TPM в нагрузочном тесте базы данных SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSD

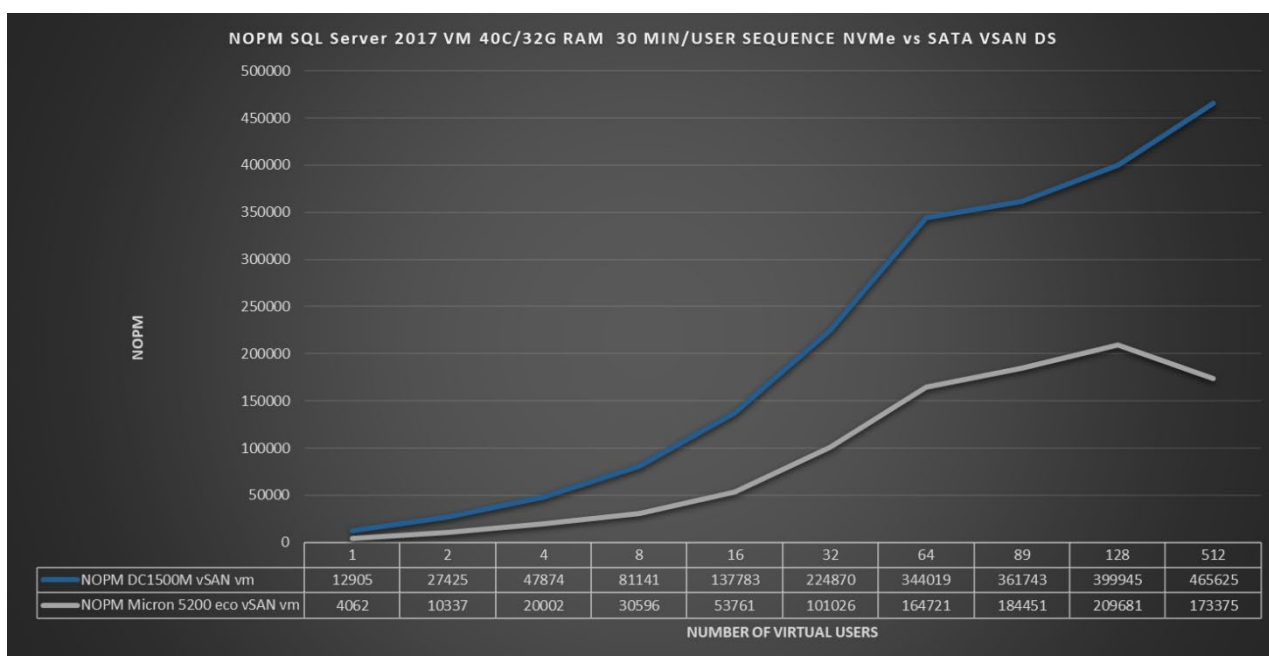


Рис. 4.3. Тест 3: сравнение TPM в нагрузочном тесте базы данных SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSD

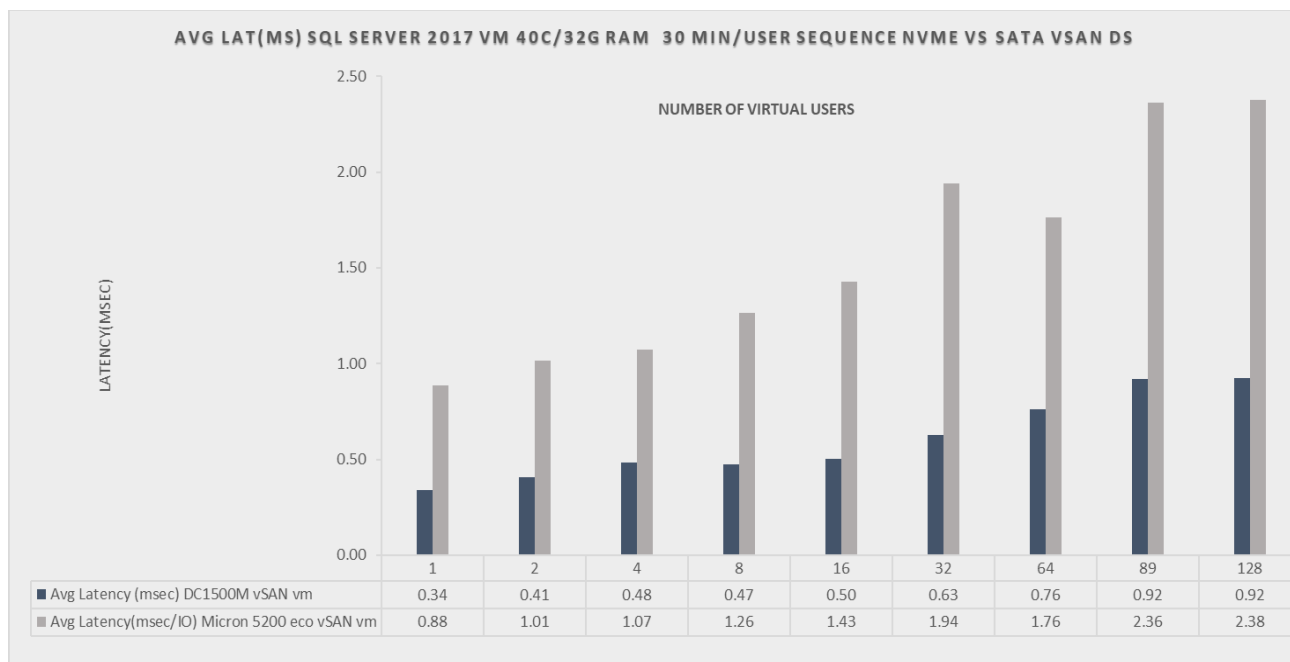


Рис. 4.4. Тест 3: сравнение средней задержки (мс) в нагрузочном тесте базы данных SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSD

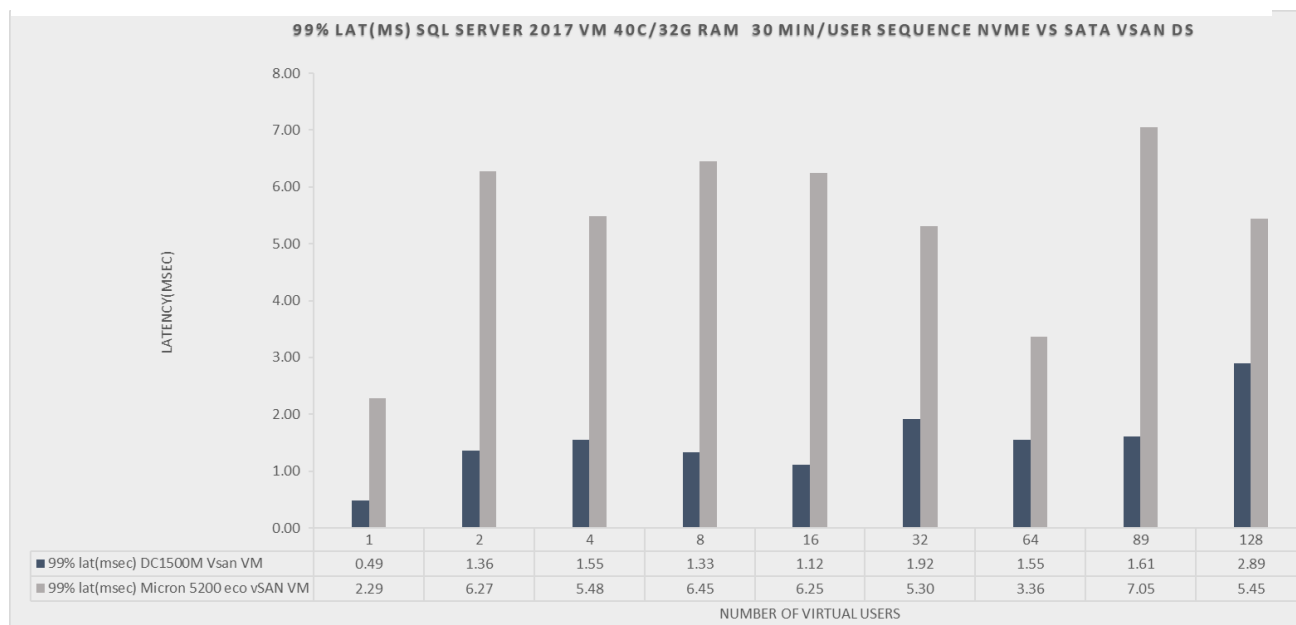


Рис. 4.5. Тест 3: сравнение задержки (мс) при выполнении 99% операций в нагрузочном тесте базы данных SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSD

Тест 4: сравнение производительности SQL Server 2017, производительность операций резервного копирования и восстановления, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA

- Конфигурация хранилища данных vSAN на основе NVMe для теста 3a: 3 накопителя DC1500M 960G FW S67F0103 на дисковую группу, всего 4 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе NVMe. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 4a) - Конфигурация хранилища данных vSAN на основе SATA для теста 3b: 3 накопителя Micron 5200 ECO 1920G FW D1MU004 на дисковую группу, всего 3 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе SATA. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 4b)	
Описание теста 4a	Описание теста 4b
Виртуальный диск предоставлен из хранилища данных vSAN на основе DC1500M в среде для тестирования NVMe. В тестируемой системе была создана схема базы данных с 2000 складами, представляющая базу данных размером 157 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Было запущено 3 цикла сценария резервного копирования/восстановления для резервного копирования и восстановления базы данных TPCC и записи показателей производительности с помощью монитора производительности Windows. Тест выполнялся локально на виртуальной машине тестируемой системы.	Виртуальный диск предоставлен из хранилища данных vSAN на основе Micron 5200 eco в среде для тестирования SATA. В тестируемой системе была создана схема базы данных с 1200 складами, представляющая базу данных размером 157 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Было запущено 3 цикла сценария резервного копирования/восстановления для резервного копирования и восстановления базы данных TPCC и записи показателей производительности с помощью монитора производительности Windows. Тест выполнялся локально на виртуальной машине тестируемой системы.

Рис. 5.1. Описание теста 4: сравнение производительности резервного копирования/восстановления SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSD

Операции резервного копирования и восстановления базы данных SQL — хороший способ измерить пропускную способность и задержку базового виртуального диска. Мы хотели установить базовый уровень пропускной способности и задержки для отдельной виртуальной машины в хранилище данных vSAN на основе NVMe и на основе SATA, фиксируя показатели виртуального диска с помощью монитора производительности Windows при запуске операций резервного копирования/восстановления TPC-C.

Тест 4: Результаты: сравнение производительности SQL Server 2017, производительность операций резервного копирования и восстановления, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA

На рисунках 5.2–5.4 показаны посекундные данные о пропускной способности и задержке, собранные в нашем сценарии монитора производительности Windows для одного из циклов резервного копирования/восстановления в тестах 4a) и 4b). Виртуальная машина SQL Server в хранилище данных vSAN на основе DC1500M NVMe завершила операцию резервного копирования базы данных TPCC за 265 секунд, достигнув средней пропускной способности 593 МБ/с и средней задержки 1,46 мс на операцию ввода-вывода. Операция восстановления базы данных TPCC завершилась за 129 секунд со средней пропускной способностью 1,4 ГБ/с и средней задержкой 2,65 мс на операцию ввода-вывода. По сравнению с виртуальной машиной в хранилище данных vSAN на основе Micron 5200 eco, операция резервного копирования была выполнена в 1,5 раза быстрее, а операция восстановления — в 2,15 раза быстрее на виртуальной машине SQL в хранилище данных vSAN на основе NVMe.

Как правило, операции резервного копирования и восстановления выполняются в нерабочее время, чтобы избежать какого-либо воздействия на производственные виртуальные машины. Однако так бывает не всегда. Если операции резервного копирования или восстановления базы данных SQL выполняются в рабочие часы с пиковой нагрузкой, они должны быть завершены как можно быстрее,

чтобы избежать влияния задержек на пользователей, которые выполняют транзакции в приложениях уровня 1, совместно использующих то же хранилище данных vSAN. Перенос баз данных SQL в хранилища данных vSAN на основе NVMe позволяет снизить это влияние. Даже если операции резервного копирования/восстановления выполняются в нерабочее время, их ускорение позволяет сократить время простоя баз данных уровня 1, использующих те же ресурсы.

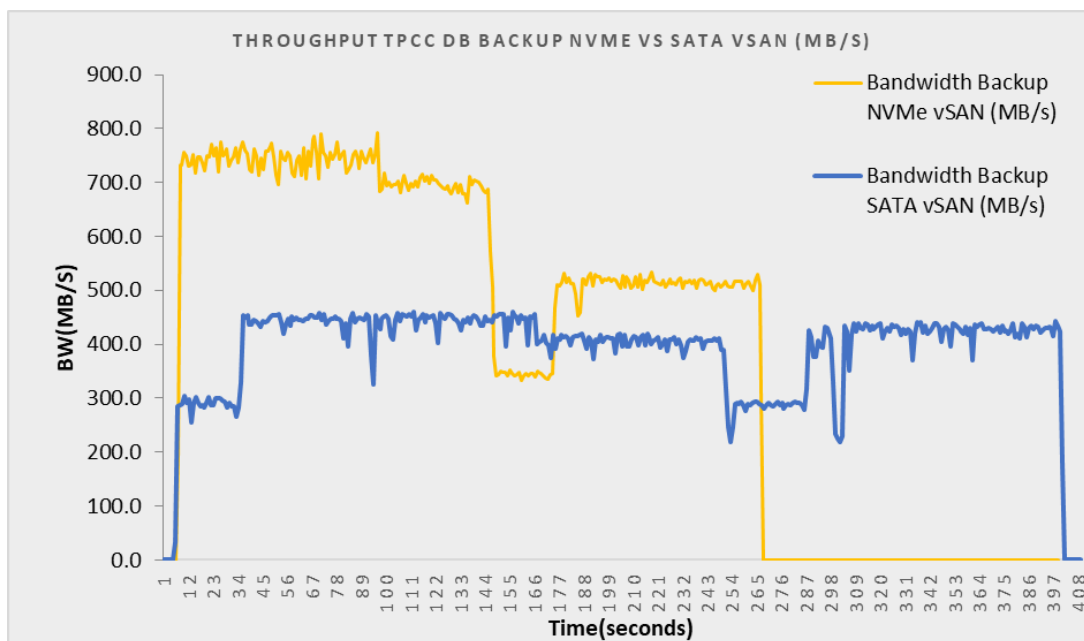


Рис. 5.2. Сравнение пропускной способности операций резервного копирования базы данных TPCC на SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 e9c SATA и на основе DC1500M NVMe SSD (МБ/с)

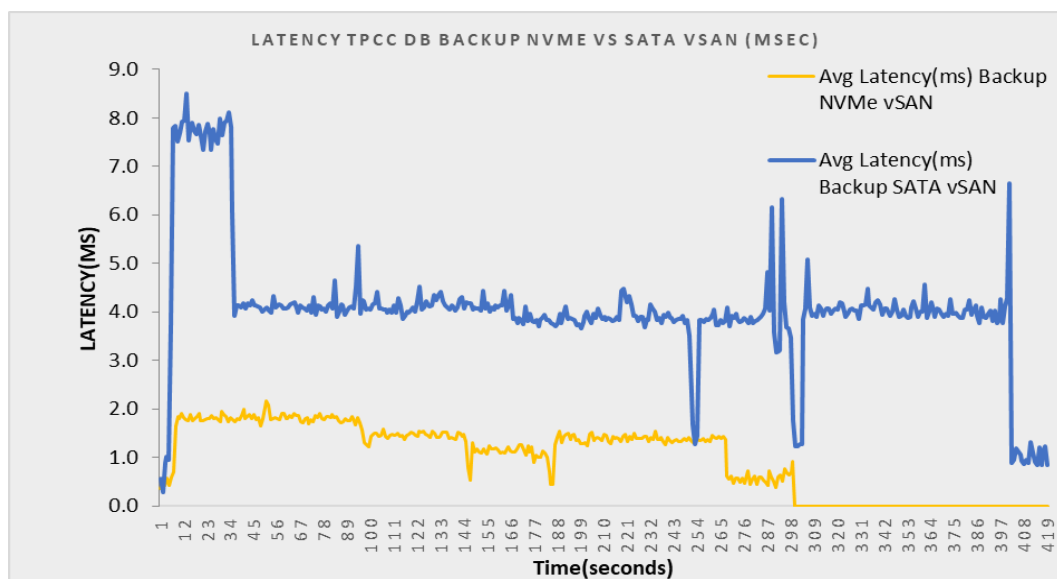


Рис. 5.3. Сравнение средней задержки (мс) для операций резервного копирования базы данных TPCC на SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 e9c SATA и на основе DC1500M NVMe SSD

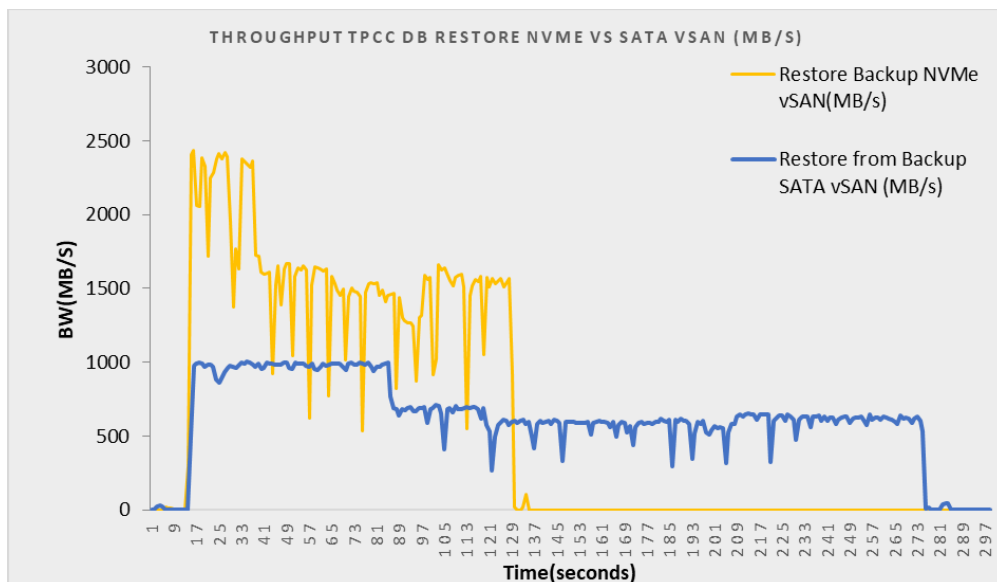


Рис. 5.4. Сравнение пропускной способности операций восстановления базы данных TPCC на SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 e90 SATA и на основе DC1500M NVMe SSD (МБ/с)

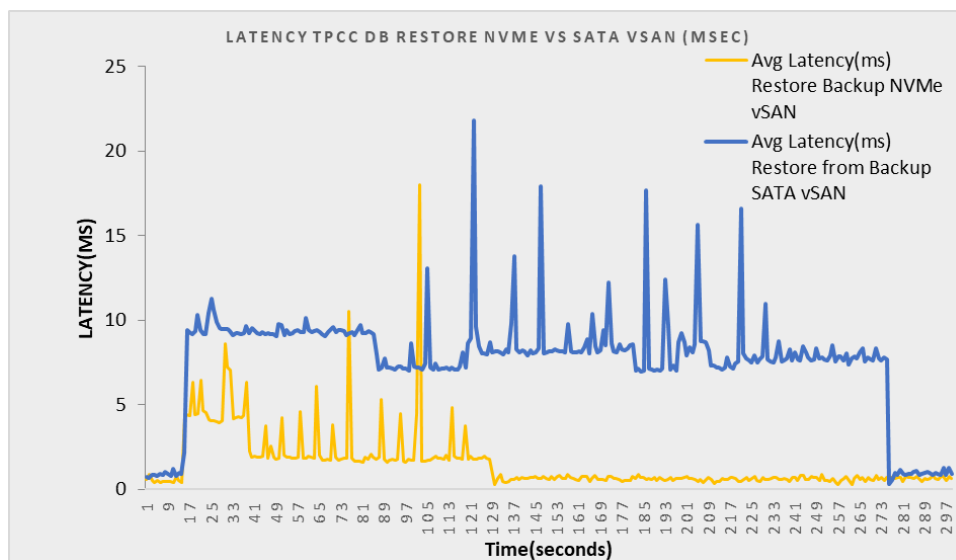


Рис. 5.5. Сравнение задержки (мс) для операций восстановления базы данных TPCC на SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 e90 SATA и на основе DC1500M NVMe SSD

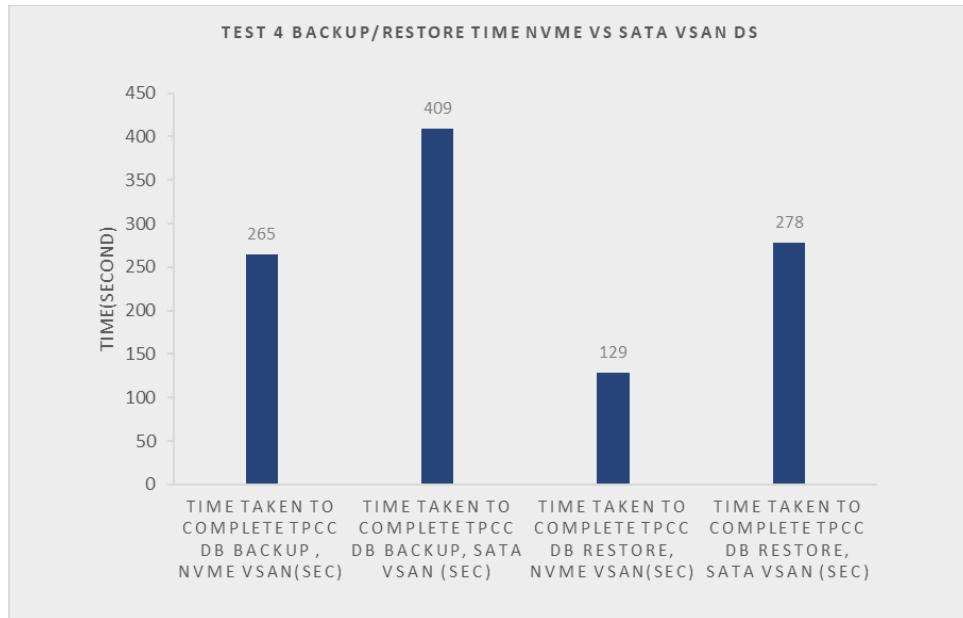


Рис. 5.6. Время выполнения операции резервного копирования/восстановления базы данных TPCC на SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSD (секунды)

Тест 5: сравнение производительности SQL Server 2017, тест Noisy Neighbor, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA

• Конфигурация хранилища данных vSAN на основе NVMe для теста 3a: 3 накопителя DC1500M 960G FW S67F0103 на дисковую группу, всего 4 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе NVMe. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 5a) • Конфигурация хранилища данных vSAN на основе SATA для теста 3b: 3 накопителя Micron 5200 ECO 1920G FW D1MU004 на дисковую группу, всего 3 дисковых группы (по 1 на сервер), среда для тестирования vSAN на основе SATA. SQL Server 2017 с гостевой ОС Server 2019 Datacenter. (тест 5b)			
Описание теста 5a	Описание теста 5b	Описание теста 5c	Описание теста 5d
Виртуальный диск виртуальной машины SQL 2017 предоставлен из хранилища данных vSAN на основе DC1500M в среде для тестирования NVMe. В тестируемой системе была создана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Тестируемая система была клонирована 11 раз, и на каждый физический сервер было назначено 3 виртуальных машины тестируемой системы (всего 12 виртуальных машин тестируемой системы). Тест был сконфигурирован для работы 89 виртуальных пользователей. На каждой виртуальной машине тестируемой системы были выбраны время разгона 30 минут и продолжительность теста 300 минут. Тест запускался параллельно на всех 12 виртуальных машинах тестируемой системы.	Виртуальный диск SQL 2017 был предоставлен из хранилища данных vSAN на основе Micron 5200 eco в среде для тестирования SATA. В тестируемой системе была создана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Тестируемая система была клонирована 8 раз, и на каждый физический сервер было назначено 3 виртуальных машины тестируемой системы (всего 9 виртуальных машин тестируемой системы). Тест был сконфигурирован для работы 89 виртуальных пользователей. На каждой виртуальной машине тестируемой системы были выбраны время разгона 30 минут и продолжительность теста 300 минут. Тест запускался параллельно на всех 9 виртуальных машинах тестируемой системы.	Виртуальный диск виртуальной машины SQL 2017 был предоставлен из хранилища данных vSAN на основе DC1500M в среде для тестирования NVMe. В тестируемой системе была создана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Тестируемая система была клонирована 11 раз, и на каждый физический сервер было назначено 2 виртуальных машины тестируемой системы (всего 8 виртуальных машин тестируемой системы) для выполнения рабочей нагрузки HDB. Тест был сконфигурирован для работы 89 виртуальных пользователей. На каждой виртуальной машине тестируемой системы были выбраны время разгона 30 минут и продолжительность теста 300 минут. 1 виртуальная машина на физический сервер имела размер схемы TPCC с 1200 складами (100 ГБ), и сценарий резервного копирования запускался каждые 100 секунд (всего 4 виртуальных машины), в то время как рабочая нагрузка выполнялась на других виртуальных машинах тестируемой системы в течение 10 циклов. На 8 виртуальных машинах тестируемой системы выполнялась рабочая нагрузка; на 4 виртуальных машинах выполнялся сценарий резервного копирования. Тест запускался параллельно на всех 12 виртуальных машинах.	Виртуальный диск виртуальной машины SQL 2017 был предоставлен из хранилища данных vSAN на основе Micron 5200 eco в среде для тестирования SATA. В тестируемой системе была создана схема базы данных с 1200 складами, представляющая базу данных размером 100 ГБ. Виртуальной машине тестируемой системы (SUT) было выделено 16 виртуальных ядер и 32 ГБ ОЗУ. Тестируемая система была клонирована 8 раз, и на каждый физический сервер было назначено 2 виртуальных машины тестируемой системы (всего 6 виртуальных машин тестируемой системы) для выполнения рабочей нагрузки HDB. Тест был сконфигурирован для работы 89 виртуальных пользователей. На каждой виртуальной машине тестируемой системы были выбраны время разгона 30 минут и продолжительность теста 300 минут. 1 виртуальная машина на физический сервер имела размер схемы TPCC с 1200 складами (100 ГБ), и сценарий резервного копирования запускался каждые 100 секунд (всего 4 виртуальных машины), в то время как рабочая нагрузка выполнялась на виртуальной машине тестируемой системы. На 6 виртуальных машинах тестируемой системы выполнялась рабочая нагрузка; на 3 виртуальных машинах выполнялся сценарий резервного копирования. Тест запускался параллельно на всех 9 виртуальных машинах.

Рис. 6.1. Описание теста 5: реалистичный тест Noisy Neighbor базы данных SQL Server 2017 в хранилище данных vSAN на основе Micron 5200 eco SATA и на основе DC1500M NVMe SSDz

Наша цель в этом тесте состояла в том, чтобы смоделировать реалистичный сценарий, в котором посторонние рабочие нагрузки (в этом случае мы используем операции резервного копирования базы данных TPCC) на виртуальных машинах используют то же хранилище данных vSAN, что и виртуальные машины SQL-сервера, выполняющие рабочие нагрузки (в этом эксперименте — эталонный тест TPCC в качестве рабочей нагрузки), а также оценить общее влияние на производительность по результатам тестов TPCC и анализу ключевых показателей хранилища данных, собранных с помощью Perfmon и монитора производительности vSAN.

В тестах 5a) и 5b) мы устанавливаем базовый уровень, запуская тест TPCC на всех виртуальных машинах параллельно, без операций резервного копирования. Мы используем по 3 виртуальные машины SQL на каждый физический сервер для работы в кластерах vSAN на основе NVMe и SATA, в результате чего общее количество виртуальных машин тестируемой системы достигает 12 для NVMe и 9 для SATA. Размер схемы для этого теста составлял 1200 складов, что соответствует размеру базы данных TPC-C примерно 100 ГБ. Рабочая нагрузка TPCC выполнялась с 89 пользователями в течение 300 минут, с временем разгона 30 минут.

В тестах 5c) и 5d) мы восстановили базу данных TPC-C на всех виртуальных машинах тестируемой системы. Затем мы запустили сценарий для выполнения 10 циклов резервного копирования базы данных TPC-C на 4 виртуальных машинах для кластера NVMe и на 3 виртуальных машинах для кластера SATA, одновременно выполняя тот же эталонный тест TPC-C на оставшихся виртуальных машинах тестируемой системы. Это означает, что в кластере vSAN на основе NVMe 8 виртуальных машин выполняли рабочую нагрузку TPC-C, а 4 виртуальные машины параллельно выполняли рабочую нагрузку резервного копирования. А в кластере SATA vSAN 6 виртуальных машин выполняли рабочую нагрузку TPC-C, а 3 виртуальных машины параллельно выполняли рабочую нагрузку резервного копирования базы данных TPC-C.

Результаты теста 5: сравнение производительности SQL Server 2017, тест Noisy Neighbor, vSAN на базе DC1500M NVMe и Micron 5200 eco SATA

На рисунках 6.2 и 6.3 показаны количество транзакций в минуту (TPM) и новых заказов в минуту (NOPM), полученные в тестах 5a и 5b. С 89 виртуальными пользователями, работающими на каждой из 12 виртуальных машин SQL Server в хранилище данных vSAN на основе DC1500M NVMe, мы смогли достичь в среднем 523 516 TPM и 113 812 NOPM на каждую виртуальную машину по сравнению со средним значением 269 320 TPM и 58 544 NOPM на каждую из 9 виртуальных машин SQL в кластере на основе Micron 5200 eco SATA. При просмотре числа операций ввода-вывода в секунду и задержки, собранных с помощью монитора производительности vSAN (рисунки 6.4 и 6.5 ниже), результирующий ввод-вывод на блочном уровне преобразуется в 120 000 операций ввода-вывода в секунду при чтении и 60 000 — при записи в кластере NVMe с задержкой 800 мкс для операций чтения/записи. А в кластере vSAN на основе SATA были достигнуты показатели в 50 000 операций чтения и 20 000 операций записи со средней задержкой 3,8 мс для операций чтения и 5,5 мс для операций записи. Это еще раз подчеркивает разницу в производительности между NVMe и SATA и демонстрирует способность виртуального диска на основе DC1500M NVMe принимать параллельные запросы и обрабатывать их с гораздо меньшей задержкой в обоих направлениях.

На рисунках 6.5 и 6.6 показаны количество транзакций в минуту (TPM) и новых заказов в минуту (NOPM), полученные в тестах 5c и 5d. На каждой из 8 виртуальных машин SQL Server в хранилище данных vSAN на основе DC1500M NVMe работали 89 виртуальных пользователей, и параллельно на 4 виртуальных машинах запускалось резервное копирование. В этих условиях мы смогли достичь средних показателей 575 933 TPM и 125 206 NOPM. При этом на 6 виртуальных машинах SQL в хранилище данных vSAN на основе Micron 5200 eco SATA, выполняющих рабочую нагрузку TPCC, в то время как параллельно на 3 виртуальных машинах запускалось резервное копирование, были достигнуты средние значения 351 258 TPM и 76 355 NOPM. Для полноты картины мы должны проанализировать задержку и показатели хранилища данных для кластера vSAN как на основе SATA, так и на основе NVMe, а также проверить, насколько быстро резервное копирование выполняется в обоих кластерах.

На рисунках 6.8 и 6.9 приведены показатели IOPS и задержки vSAN, полученные в кластерах на основе NVMe и SATA с помощью монитора производительности vSAN в тестах 5c и 5d. Сценарий резервного копирования был настроен на запуск каждые 100 секунд в течение 10 циклов. Мы можем наблюдать влияние запущенных операций резервного копирования виртуальных машин на число операций ввода-вывода в секунду (IOPS) и задержку операций чтения и записи в кластерах vSAN на основе NVMe и SATA. Однако влияние на задержку отличается. Максимальная задержка ввода-вывода при чтении/записи в кластере NVMe увеличилась до 4 мс при среднем значении 2,5 мс на операцию ввода-вывода, в то время как для vSAN на основе SATA она увеличилась до 9 мс на операцию и в среднем составила 7,3 мс на операцию ввода-вывода при чтении и 4,9 мс — при записи. Это задержка, которую ощутит конечный пользователь, когда пытается отправить заказ, обновить свою корзину или просмотреть позиции с других складов.

На рисунке 6.11 показано время, которое потребовалось для завершения циклов резервного копирования на одной из виртуальных машин SQL Server в хранилище данных vSAN на основе DC1500M и на одной из виртуальных машин SQL в хранилище данных vSAN на основе Micron 5200 есо vSAN, исключая время ожидания между циклами резервного копирования. На выполнение 10 операций резервного копирования потребовалось 73 минуты, то есть в среднем 7 минут на одну операцию, для виртуальной машины SQL Server в хранилище vSAN на основе NVMe, и 122,15 минуты на выполнение 10 операций резервного копирования виртуальной машины SQL Server в хранилище данных vSAN на основе SATA SSD — в среднем 12 минут на одну операцию. Виртуальная машина в хранилище данных vSAN на основе DC1500M завершила циклы резервного копирования в 1,67 раза быстрее, чем виртуальная машина в хранилище данных vSAN на основе Micron 5200 есо. Это фактически доказывает, что обновление инфраструктуры VMware с использованием хранилищ данных на основе DC1500M NVMe помогает смягчить проблему «шумных соседей», позволяя выполнять нежелательные операции, такие как резервное копирование базы данных, намного быстрее, а благодаря огромному снижению задержки и увеличению пропускной способности NVMe может компенсировать влияние задержек, связанных с нежелательными рабочими нагрузками, на приложения уровня 1.

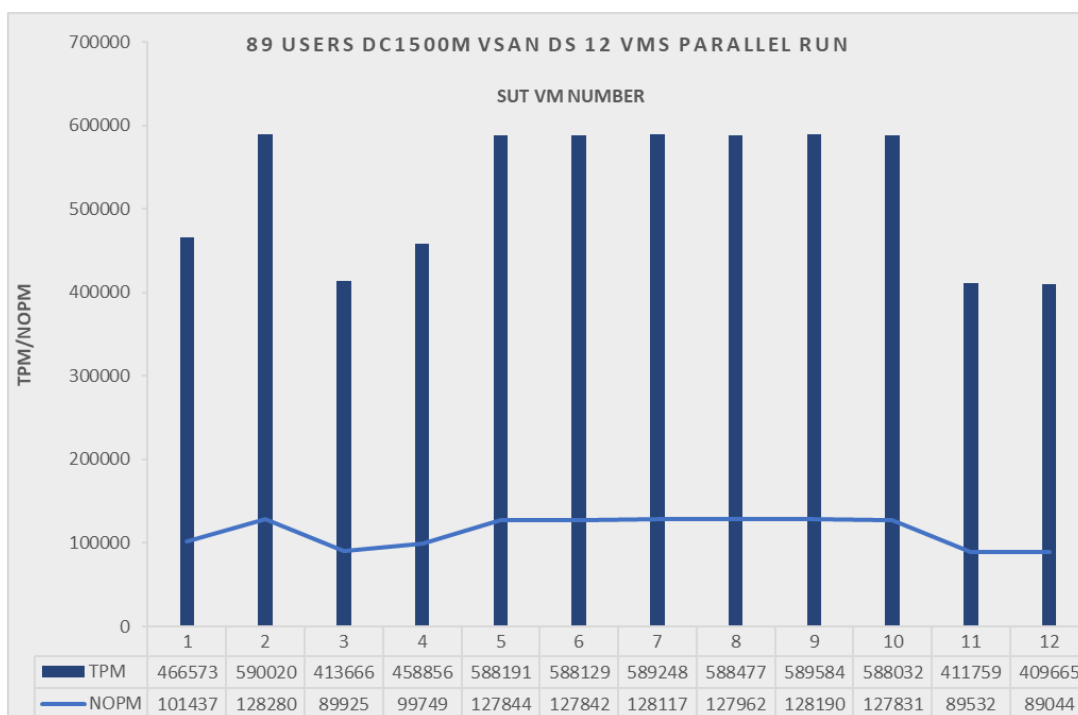


Рис. 6.2. Тест 5a: TPM для SQL Server 2017, 300 минут, параллельное выполнение 12 виртуальных машин, 89 виртуальных пользователей, хранилище данных vSAN на основе DC1500M NVMe SSD

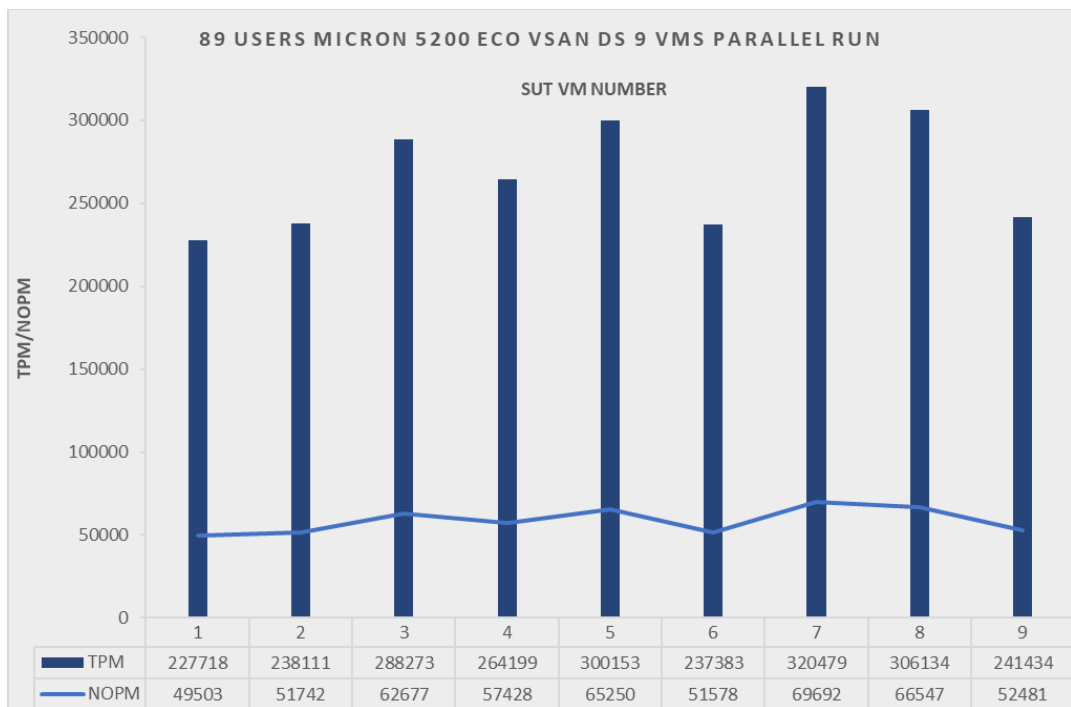


Рис. 6.3. Тест 5b: TPM для SQL Server 2017, 300 минут, параллельное выполнение 12 виртуальных машин, 89 виртуальных пользователей, хранилище данных vSAN на основе DC1500M NVMe SSD

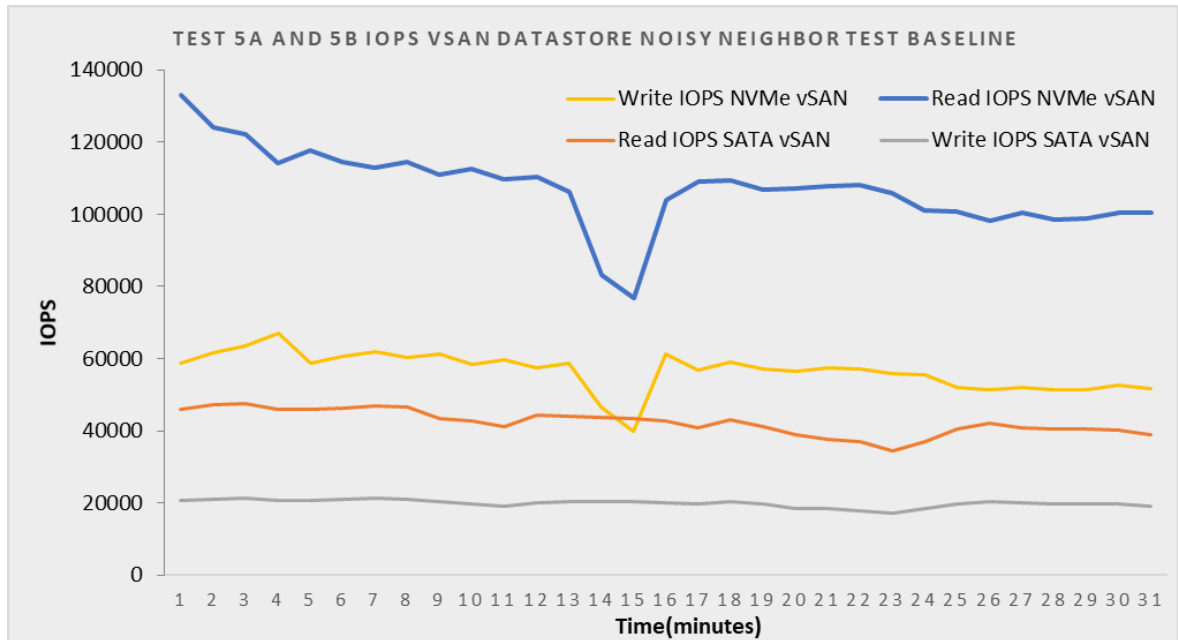


Рис. 6.4. Тесты 5a и 5b: IOPS в тесте Noisy Neighbor, хранилище данных vSAN на основе DC1500M NVMe и Micron 5200 eco

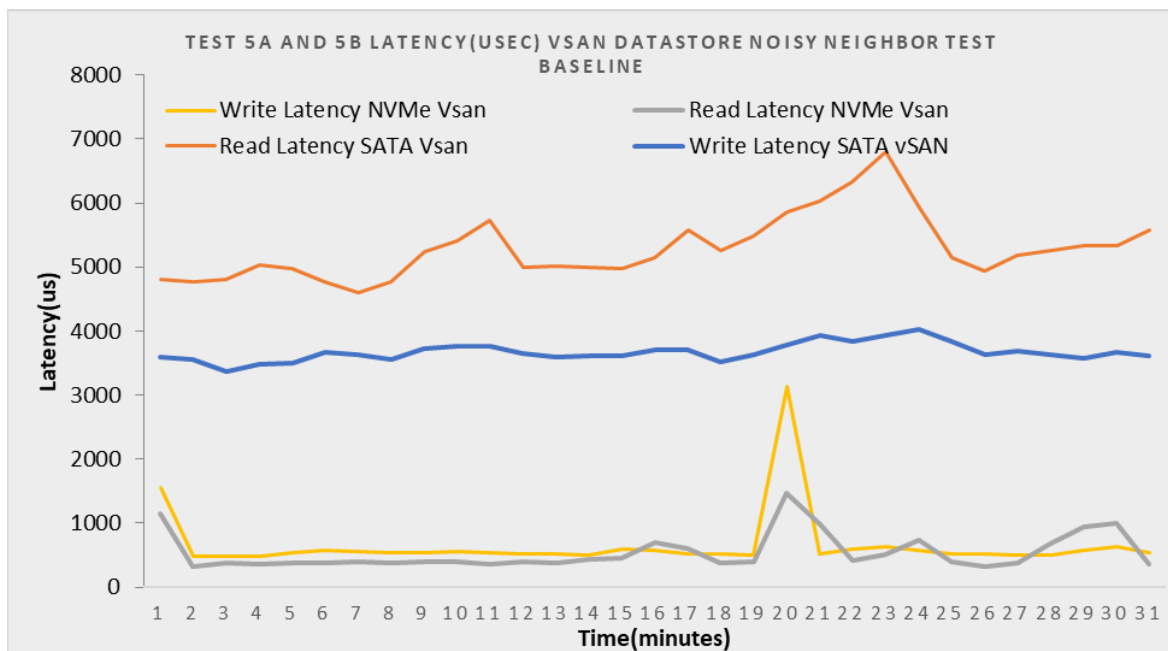


Рис. 6.5. Тесты 5a и 5b: задержка в тесте Noisy Neighbor, хранилище данных vSAN на основе DC1500M NVMe и Micron 5200 eco

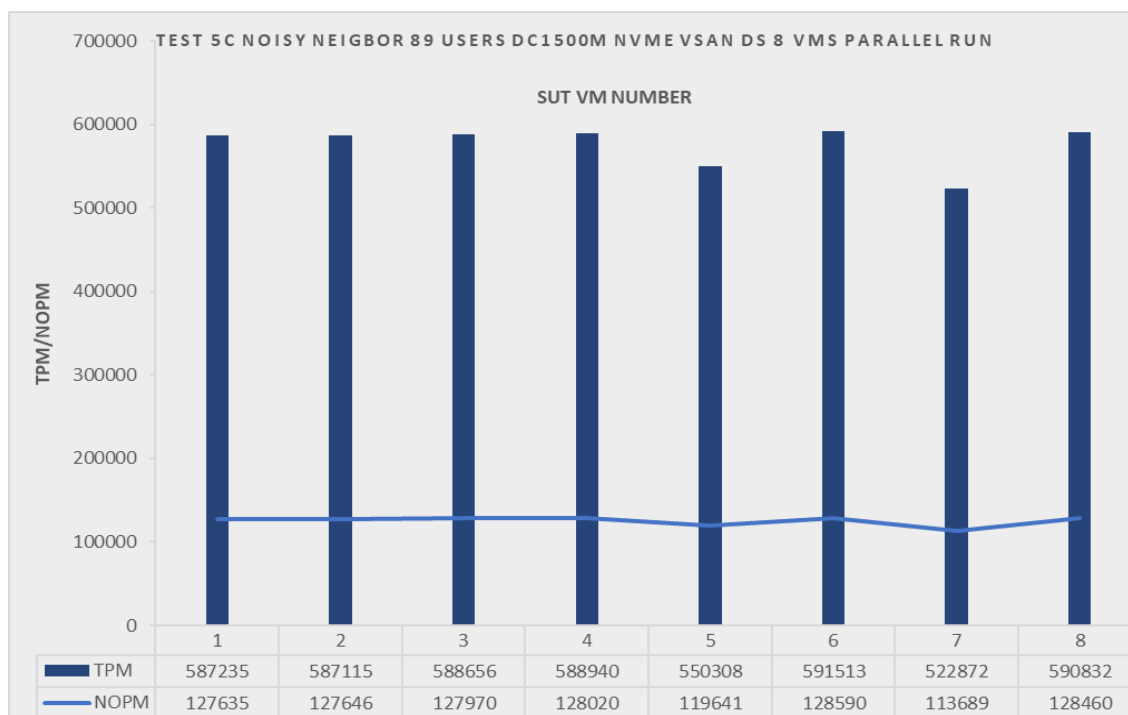


Рис. 6.6. Тест 5c: TPM, реализация «шумного соседа», параллельное выполнение 8 виртуальных машин, хранилище данных vSAN на основе DC1500M NVMe

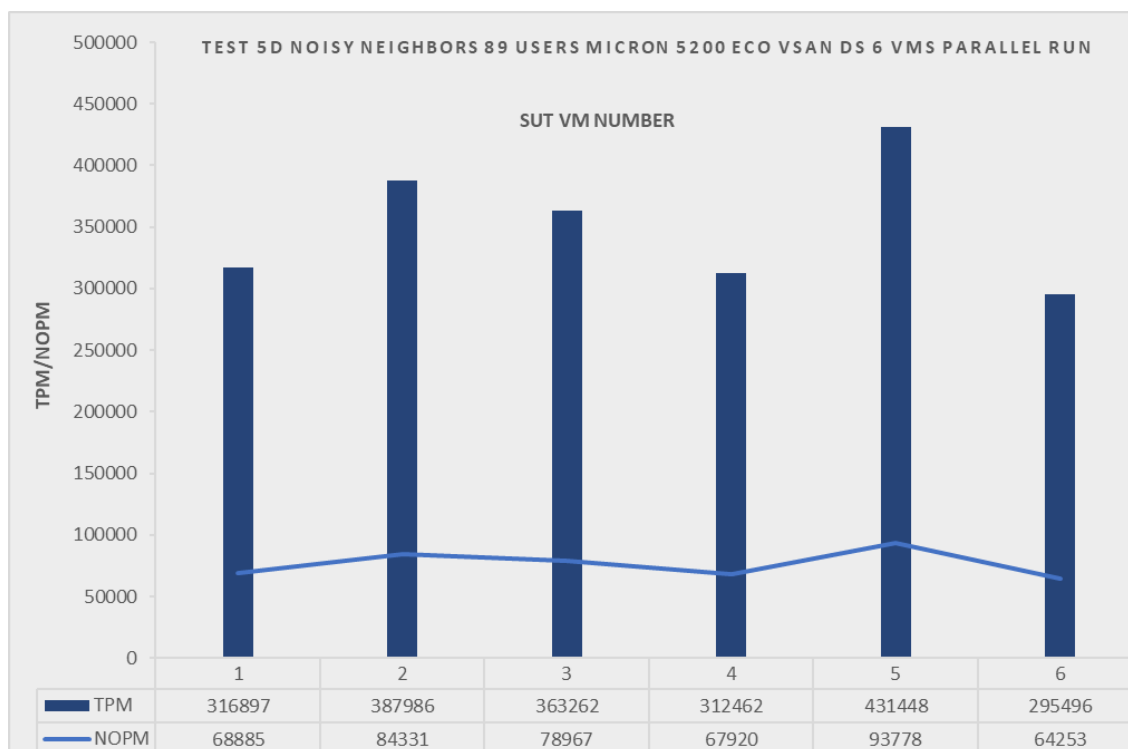


Рис. 6.7. Тест 5с: TPM, реализация «шумного соседа», параллельное выполнение 6 виртуальных машин, хранилище данных vSAN на основе Micron 5200 eco

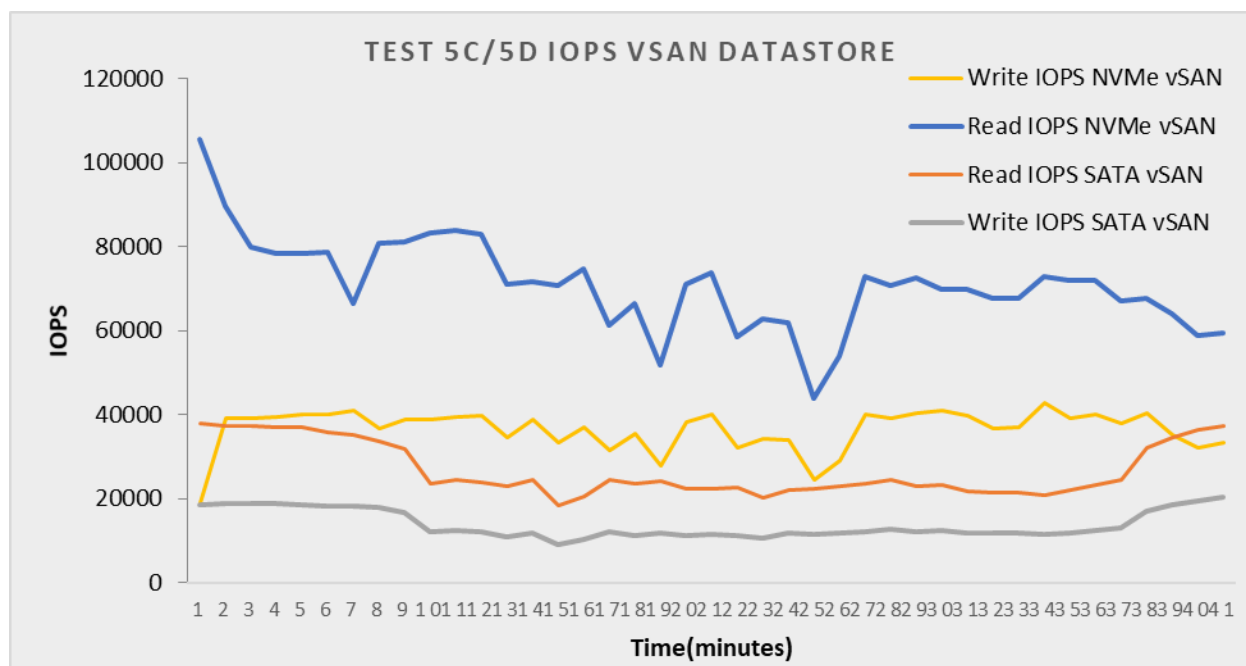


Рис. 6.8. Тест 5с/5d: IOPS, реализация «шумного соседа», сравнение хранилищ данных vSAN на основе NVMe и на основе SATA SSD

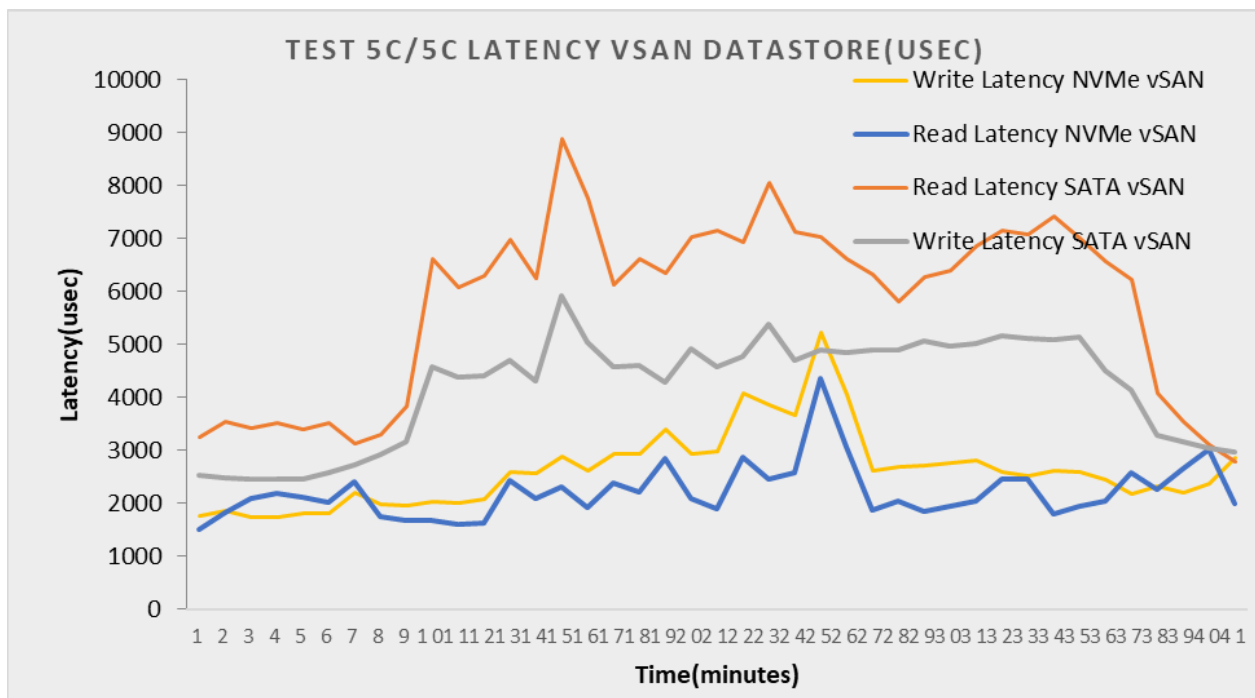


Рис. 6.9. Тест 5с/5d: задержка, реализация «шумного соседа», сравнение хранилищ данных vSAN на основе NVMe и на основе SATA SSD

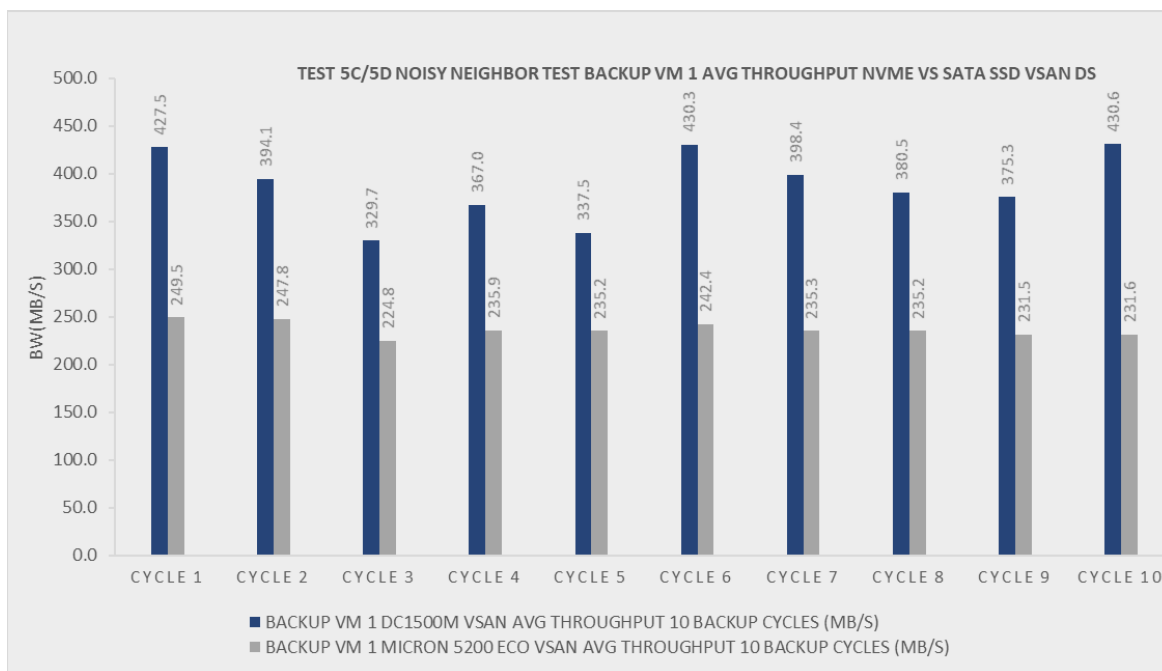


Рис. 6.10. Тест 5с/5d: пропускная способность виртуальных машин в операциях резервного копирования, реализация «шумного соседа», сравнение хранилищ данных vSAN на основе NVMe и на основе SATA SSD

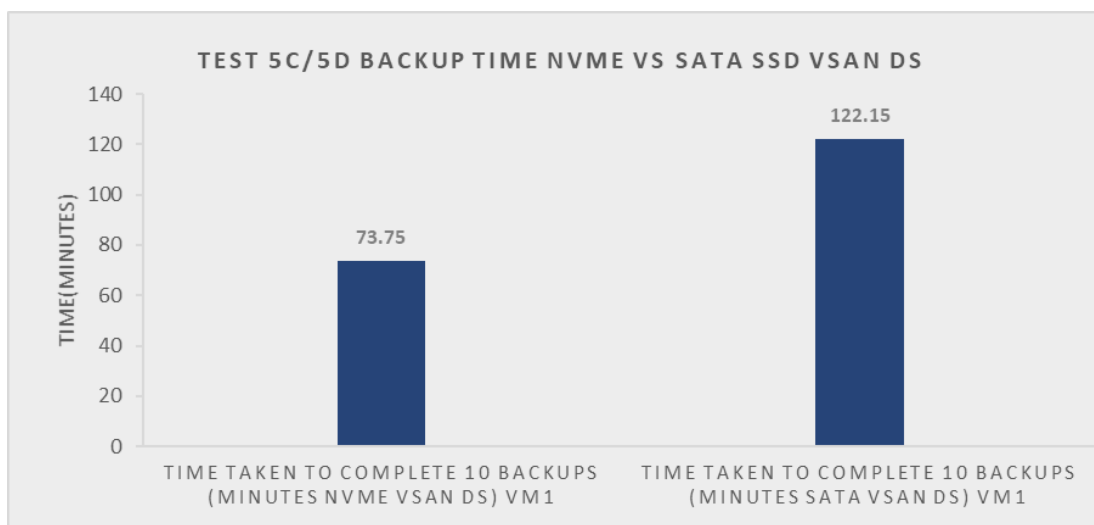


Рис. 6.11. Тест 5c/5d: время, затраченное на резервное копирование виртуальных машин, реализация «шумного соседа», сравнение хранилищ данных vSAN на основе NVMe и на основе SATA SSD

Заключение

В этом техническом документе мы показали, как консолидация рабочих нагрузок базы данных на накопителях NVMe может помочь в максимальной степени использовать существующее оборудование благодаря его невероятной эффективности и почти нулевому времени ожидания операций ввода-вывода, что позволяет использовать меньше ядер процессора для достижения той же пропускной способности транзакций. Мы провели несколько сравнений с твердотельными накопителями SATA корпоративного класса и показали, что перенос рабочих нагрузок SQL в хранилище данных на основе NVMe позволяет обеспечить масштабирование приложений, удвоив пропускную способность транзакций и обеспечивая при этом задержку менее миллисекунды. Затем мы показали, как NVMe позволяет смягчить влияние на приложения уровня 1, позволяя быстрее выполнять нежелательные рабочие нагрузки, такие как операции резервного копирования/восстановления базы данных.

Твердотельные накопители NVMe корпоративного класса от компании Kingston, такие как [DC1500M](#), в сочетании с серверной памятью от Kingston (Server Premier) представляют собой отличное решение для тех, кто хочет виртуализировать инфраструктуру базы данных и достичь максимальной эффективности при выполнении рабочих нагрузок.

Посетите веб-сайт <https://www.kingston.com/en/solutions/servers-data-centers>, чтобы узнать подробнее о решениях Kingston для центров обработки данных.

Ссылки

HammerDB. (дата не указана). *Понимание рабочей нагрузки TPCC*. Получено с сайта <https://www.hammerdb.com/docs3.3/ch03s05.html>

Домашняя страница TPCC. (дата не указана). Получено с сайта <https://www.tpc.org/>