



Kingston DC1500M エンタープライズ NVMe ソリッドステートドライブ (SSD) および VMware vSAN HCI を用いた 総保有コスト (TCO) の削減と、 レジリエンスおよび応答性の高い データベースの構築

著者 : Hazem Awadallah、システムエンジニア、Kingston Technology

レビュー : Chris Selden、SSD 製品エンジニアリングマネージャー、
Kingston Technology



目次

- [要旨](#)
- [現在のデータセンターで RDBMS が直面する共通のインフラ課題](#)
- [ソリューション：Kingston Technology データセンター DC1500M エンタープライズ NVMe SSD のご紹介](#)
- [テスト環境](#)
 - [I. インフラ](#)
 - [II. データベース構成](#)
 - [III. vSAN ストレージパフォーマンス](#)
- [テスト結果](#)
 - [テスト 1: 各種容量の DRAM を用いた DC1500M 960GB vSAN SQL Server 2017 VM](#)
 - [テスト 1 結果:各種容量の DRAM を用いた DC1500M 960GB vSAN SQL Server 2017 VM](#)
 - [テスト 2:Kingston DC500M SATA SSD、Micron 5200 eco SATA SSD および DC1500M NVMe SSD vSAN データストアでの SQL Server 2017 のパフォーマンス比較](#)
 - [テスト 2 結果:Kingston DC500M SATA SSD、Micron 5200 eco SATA SSD および DC1500M NVMe SSD vSAN データストアでの SQL Server 2017 のパフォーマンス比較](#)
 - [テスト 3:DC1500M NVMe と Micron 5200 eco SATA vSAN データストアの SQL Server 2017 のパフォーマンス比較、スキーマサイズを拡大し、テスト時間を延長](#)
 - [テスト 3 結果:DC1500M NVMe と Micron 5200 eco SATA vSAN データストアの SQL Server 2017 のパフォーマンス比較、スキーマサイズを拡大し、テスト時間を延長](#)
 - [テスト 4:SQL Server 2017 のパフォーマンス比較、バックアップおよびリストアのパフォーマンス、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較](#)
 - [テスト 4:結果:SQL Server 2017 のパフォーマンス比較、バックアップおよびリストアのパフォーマンス、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較](#)
 - [テスト 5:SQL Server 2017 のパフォーマンス比較、ノイジーネイバーテスト、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較](#)
 - [テスト 5 結果:SQL Server 2017 のパフォーマンス比較、ノイジーネイバーテスト、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較](#)
- [結論](#)

要旨

ここ数年、NVMe の導入によってデータストレージの分野が革新され、NAND フラッシュのパフォーマンスが最大に高まり、機能豊富で低コストかつ高帯域幅の、将来の変化に対応した拡張バス規格である PCI Express の活用などで大きな進歩がありました。現在第 5 世代の PCIe Gen5 では、レーンあたり最大 8GB/秒の転送速度が可能で、ストレージスタックにおける拡張バスのボトルネックを解消し、SSD コントローラーや NAND フラッシュだけでなく、ハードウェアスタック全体のイノベーションと進化を促進します。プロセッサ、筐体設計、マザーボード、ハードウェアの入出力ポロジは、帯域幅を増大するために絶えず進化しています。データセンターでは、NVMe に対応するため、ネットワークポロジが大きく変化しています。NVMe-OF 仕様によって、ネットワークインターフェイス、スイッチ、伝送プロトコルが変更され、帯域幅を増大しつつ、QoS およびロスレスパケット転送を維持するために継続的に改善されています。

しかし NVMe の導入はアプリケーションのパフォーマンスにどのような影響を与えたのでしょうか？ トランザクションスループットを改善したりトランザクション応答時間を短縮しながら、ストレージの設置面積を縮小することは可能でしょうか？ データベースのバックアップ時間を大幅に短縮して、ノイジーネイバーの問題を軽減し、本番環境での影響を最小限に抑えることができますか？ 本記事では、これらの疑問に答えるため、標準的な OLTP ワークロード (TPCC 仕様によって定義) を調査し、実践的な比較をいくつか行って、実際の状況でトランザクションパフォーマンスに NVMe が与える影響についてまとめました。

現在のデータセンターで RDBMS が直面する共通のインフラ課題

コスト、容量のプランニング、拡張性

過去 20 年間に起きたインターネット帯域幅の増大、処理速度の向上、およびデータ分析に対する人気の高まりで、本番 OLTP データベースは急速に成長しています。しばし、アプリケーションアーキテクトやインフラアーキテクトが策定したプランよりも、はるかにハイペースで進行しています。基盤となるストレージとネットワークアーキテクチャは、ゼロから構築し、需要の増大に対応して徐々に拡張できるようにすることで、コスト/管理しやすさ/パフォーマンスのバランスを取れるようにする必要があります。設計上の判断において、ローカルデータセンターにアプリケーションを構築するか、IaaS/PaaS クラウドサービスを利用するかを選択することは、困難になっています。ローカルデータセンターでアプリケーションを実行し続けると、ソリューションアーキテクトは拡張性、セキュリティ、レジリエンス、パフォーマンスなどを完全に管理できませんが、綿密な計画が必要で、多額の初期費用がかかることもあります。IaaS/PaaS クラウドサービスを利用すると、短期間で展開でき、拡張も簡素化されますが、パフォーマンスやレジリエンスを管理しにくくなり、アプリケーションの拡張にしたがって急速に高コストになります。中にはハイブリッドな手法を選択する組織もあり、重要性の高い Tier 1 アプリケーションをローカルデータセンターに残し、Tier 2 およびレガシーアプリケーションをクラウドに移行しています。社内に残すアプリケーションの場合、オールフラッシュのディスクグループを用いた VMware vSAN HCI のようなハイパーコンバージドインフラソリューションにより、コスト、簡索性、パフォーマンス、拡張しやすさのバランスが取れます。

レジリエンス

Tier 1 アプリケーションは、ハードウェアスタック全体でひとつ以上のハードウェア障害が発生しても耐久可能なインフラを構築するか、そのようなインフラに移行する必要があります。計画が不適切な場合、データセンター内で機器の障害が発生すると、サービス中断や多額の損失を被ることがあり、最悪な場合はデータが永久に失われます。共有ストレージ環境では、ストレージ障害やコンポーネントパフォーマンスの過負荷に耐えられる基盤インフラを構築するために、慎重な計画を立てる必要があります。

たとえば vSAN の場合、Tier 1 アプリケーションの最小 FTT (許容できる障害数) は 1 でなければならず、vSphere の高可用性(HA) が有効な場合、少なくともひとつのコンピュータ/ネットワーク/ストレージに起因する障害からアプリケーションとデータベースの VM が保護されます。さらに、vSphere 分散リソーススケジューラー (DRS) を有効にして、クラスタ内の物理サーバー間で CPU/メモリリソースの負荷を分散できます。

さまざまなパフォーマンスのニーズ

OLTP アプリケーションが継続的に拡張されるにつれて、トランザクションの高速化とレイテンシの低下の需要も増大し、ユーザーがバックエンドデータベースにさらに多くのトランザクション負荷をかけるようになります。アプリケーションアーキテクトは、このように増大する需要に対応し、ストレージの異なる階層 (Tier) に移行できるだけの柔軟性を備えたストレージインフラ計画を策定する必要があります。たとえば、SAN ストレージアレイからプロビジョニングされる仮想ディスクに存在する SQL データベースは、VMware のストレージ VMotion を使用する NVMe などを備えた高速なストレージ階層を持つ、NVMe オールフラッシュ vSAN データストアに移行できます。

ノイジーネイバーのジレンマ

インフラの設計では、主要なワークロードの実行に必要なリソースを確保できるようにしなければなりません。複数のワークロードを実行する共有ストレージ環境では、パフォーマンスが予測不能になり、異常なワークロードによって主要な本番ワークロードに問題が発生します。これがノイジーネイバー問題の定義です。たとえば、このホワイトペーパーで後述するように、あるサーバーで予定にないデータベースバックアップ操作を行うと、ストレージとネットワークのリソースを消費し、同じリソースを使用する他のサーバーのパフォーマンスとレイテンシに影響を与えます。

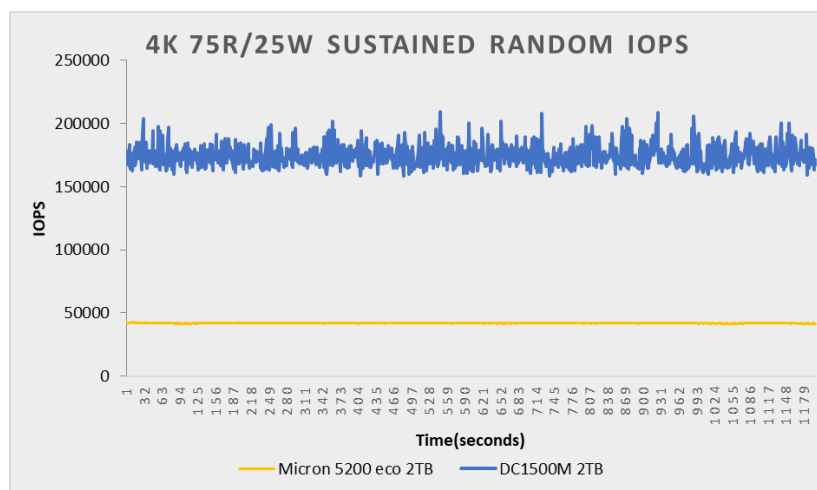
Kingston DC1500M エンタープライズ NVMe SSD の紹介

[Kingston DC1500M](#) は、Kingston が提供する最新のエンタープライズ U.2 PCIe 3.0x4 NVMe で、容量は 960GB~7680G です。16 チャンネルのコントローラと 3D TLC NAND を搭載しており、エンタープライズワークロードのパフォーマンスと一貫性を高く保ちながら、レイテンシを最低限に抑えるため、厳正なサービス品質 (QoS) 要件に沿って設計されています。そのエンタープライズ用ファームウェアは、オーバプロビジョニングや複数の名前空間 (最大 64 個の名前空間をサポート) の他、ドライブの寿命全体にわたってエンタープライズワークロードの信頼性を確保する高度な ECC アルゴリズムなどの機能をサポートしています。

データセンターでもっとも普及している SSD は SATA SSD ですが、このホワイトペーパーの目的は、Kingston DC1500M NVMe などのエンタープライズ NVMe SSD 上にストレージインフラを移行または構築すれば、上記の問題を軽減できることを示すことです。

当社の内部テストでは、単一の Kingston DC1500M NVMe SSD は、Micron 5200 eco Enterprise SATA SSD 1 台と比較して、スループットが最大 6.5 倍、レイテンシが 5.6 倍向上し (下図 b)、コストの差はほとんどありません。

ハイパーコンバージド環境でこのレベルのパフォーマンスを達成すると、SQL サーバーデータベースのトランザクションスループットが増大し、レイテンシが低下します。また、ストレージ設置面積を縮小し、消費電力を削減できます。この例では、DC1500M ドライブ 1 台と同じスループットを達成するには、micron 5200 eco ドライブ 6 台が必要です。このパフォーマンスが、VMware vSAN 上の実際の SQL OLTP ワークロードに与える影響については、後ほど説明します。



DC1500M などの NVMe SSD の導入によって、SATA SSD と比較して劇的にパフォーマンスが改善するのであれば、共有ハイパーコンバージド環境でも、

Tier 1 アプリケーションでのノイズネイバー課題の影響を軽減できます。DC1500M などのエンタープライズ NVMe SSD では、本番稼働中のバックアップ/リストア操作などの想定外のワークロードを非常に迅速に完了でき、そのうえで Tier 1 のミッションクリティカルな本番ワークロードのレイテンシを低く抑え、トランザクションスループットを高く維持します。これについては、後述のノイズネイバーテストで説明します。

図 a) DC1500M 1920G および Micron 5200 ECO 1920G SATA SSD の毎秒の IOPS の比較。Fio v3.17 を用いて Linux システムにセカンダリーとして取り付けられた単一の物理ドライブにおいて、SSD のパフォーマンスが安定した状態になった後に、テストを実施しました。ブロックサイズ 4K、読み取り率 75%、キューの深さ 32 とします。

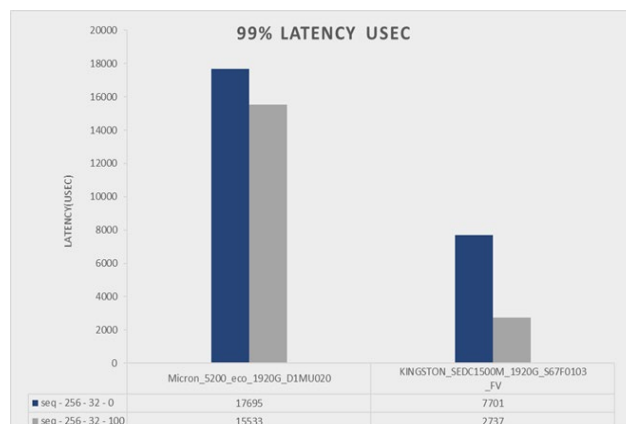
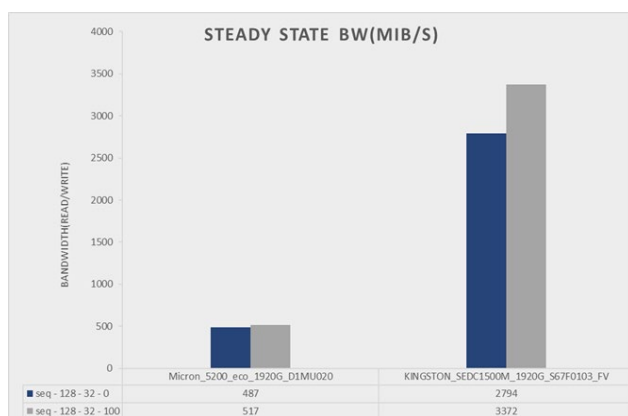


図 b) DC1500M 1920G と Micron 5200 ECO 1920G SATA SSD の順次 BW (MB/秒) 読み取り/書き込みおよびレイテンシ (μ 秒) の比較。Fio v3.17 を用いて Linux システムにセカンダリーとして取り付けられた単一の物理ドライブにおいて、SSD のパフォーマンスが安定した状態になった後に、テストを実施しました。ブロックサイズ 256k、キューの深さ 32 とします。

テスト環境

1. インフラ

当社のテスト環境を、下図 1.1 および 1.2 に示します。ハイパーコンバージド仮想環境向けのストレージオプションとして、拡張性、レジリエンス、集中管理、費用対効果の高い VMware vSAN を HCI として選択しました。

VMware vSAN を使用すると、ユーザーは複数のサーバーのローカルストレージデバイスを、vSAN クラスタ内のすべてのホスト間で共有される単一のデータストアに集約できます。各サーバーの物理ディスクがディスクグループにまとめられ、1つのドライブ/ディスクグループがキャッシュデバイスとして使用され、最大7つのドライブ/ディスクグループが容量デバイスとして使用されます。サーバーには最大で5つのディスクグループを設定できるため、vSAN クラスタに付けられるキャパシティデバイス/サーバーの数は最大35です。vSAN クラスタ中のすべての ESXi ホストのディスクグループを統合して、vSAN データストアが作られており、ホストと vSAN データストアの間のトラフィックは、vSAN 専用のネットワークを通じて分離されています(オールフラッシュ vSAN には 10Gbps 以上が必要です)。そのため、管理者は設置面積の小さいストレージから始め、必要に応じてストレージノードを追加して容量を拡張(最大64ノード/クラスタ)するため、特定の VM に合わせて比較的容易にパフォーマンス要件を管理できます。

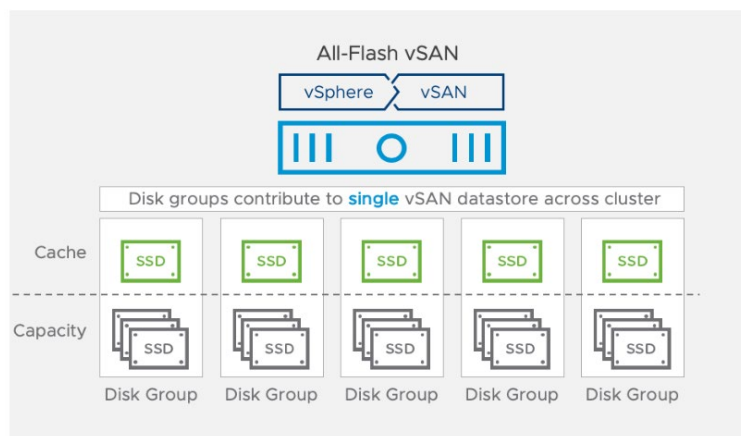


図1. オールフラッシュ vSAN アーキテクチャ

vSAN では、特定の仮想ディスクの保護とストライピングのレベルを記述するために、ストレージポリシーを使用します。デフォルトのストレージポリシーを使用すると、vSAN は、vSAN データストアからプロビジョニングされたすべてのオブジェクトをミラーリングしますが、vSAN データストアから VM にプロビジョニングされた仮想ディスクの保護のレベルについて、管理者がきめ細かく管理できるようにもなります。たとえば、クラスタ(サーバー全体、ディスク、またはネットワークイン

ターフェース)内の最低ひとつの障害に SQL データドライブ VMDK が耐えられるようにするために、プライマリレベルの FTT (許容できる障害数) に 1 を指定できます。すると、VMDK オブジェクトの RAID-1 ミラーが作成されます。vSAN クラスタで、あるホストに複製コンポーネントが、別のホストにもうひとつの複製コンポーネントが作成されます。同じように管理者は、バックアップドライブ VMDK にレジリエンスが不要でパフォーマンスを最大にしたい場合、FTT 0 の RAID 0 (ストライピングのみ) ストレージポリシーも指定できます。これは、SQL AlwaysOn フェイルオーバークラスタリングによって VM の可用性が高い場合や、Commvault や NetBackup などの一般的なバックアップソリューションを介してデータベースを定期的にバックアップしている場合に適用されます。

Kingston Technology の SSD テストおよび検証ラボと、このホワイトペーパーでは、SATA SSD テスト用に、8 2.5" NVMe および 16 2.5" SATA/SAS ドライブベイ/サーバー対応の、3つの [PowerEdge R740xD サーバー](#) を使用し、vSAN トラフィックには2つの [Cisco Nexus 5k スイッチ](#) で

サポートされた専用 10Gb ネットワークを使用しました。NVMe テストには、4 ノード Big Twin Supermicro [SYS-2029BT-HNR スーパーサーバー](#)を使用し、vSAN トラフィックには 1 台の [Cisco 9k スイッチ](#)でサポートされた専用 40Gb ネットワークを使用しました。当社のテストでは、このホワイトペーパーで実施したすべてのテストのブロックストレージパフォーマンスを最大化するために、ゲスト VM 仮想ディスクにカスタムストレージポリシー (FTT=0) を割り当てました。当社の実施した各種テストでは、下記の各テスト結果の冒頭に記載されている通り、さまざまな SSD を使用しました。標準構成としては、SATA および NVMe テストの両方で、ディスクグループあたりの容量が同じ 3 つの物理ドライブを使用しました。比較テストには、普及型の Micron 5200 eco SATA SSD を選択しました。管理と VMotion トラフィックには、Netgear JGS524PE 24 ポートマネージドスイッチ 1 台でサポートされた 1Gb ネットワークを使用しました。

NVMe テスト環境 (ハードウェア)	SATA/SAS/HYBRID テスト環境 (ハードウェア)
ホットスワップ 2.5" NVMe ドライブベイ/サーバーを 6 台搭載した Supermicro SYS-2029BT-HNR 4 ノードクラスタ	2.5" NVMe 8 台および 2.5" SATA/SAS ドライブベイ/サーバー 16 台に対応した PowerEdge Dell R740xD 3 ノードクラスタ
Intel(R) Xeon(R) Gold 6252 CPU (48c/96t) @ 2.10GHz 8 台	Intel(R) Xeon(R) Silver 4114 CPU (10c/20t) @ 2.20GHz 8 台
64x32GB Kingston DDR4-2933 2Rx4 ECC REG DIMM (ノードあたり 16x32GB)、512GB/ノード、2048GB/クラスタ	768 GB 24x32GB Kingston Dual Rank ECC メモリ @ 2400MHz/ノード、2304GB/クラスタ
vSAN ネットワークトラフィック用の Cisco Nexus N5K-C5010 20 ポート 10Gbe データセンタークラススイッチ 2 台	vSAN ネットワークトラフィック専用の Cisco Nexus 9332PQ スイッチ 32 ポート 40Gbe データセンタークラススイッチ 1 台
	HBA パススルーモードで構成された PERC H740P

図1.1 当社テストで使用したハードウェア

NVMe テスト環境 (OS およびソフトウェア)	SATA テスト環境 (OS およびソフトウェア)
ハイパーバイザ:VMware ESXi、7.0.3、19193900	ハイパーバイザ:VMware ESXi、7.0.3、19193900
vSAN 7U3c (VMware ESXi、7.0.3、19193900 + VMware VirtualCenter 7.0.3 ビルド-19234570)	vSAN 7U3c (VMware ESXi、7.0.3、19193900 + VMware VirtualCenter 7.0.3 ビルド-19234570)
ゲスト OS:Windows Server 2019 Datacenter、v1809	ゲスト OS: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

図1.2 : OS およびソフトウェア

II. データベース構成

ここで実施したテストでは、SQL Server 2017 を搭載した Server 2019 ゲスト VM を使用し、データ、ログおよびバックアップ用の vSAN データストアからプロビジョニングされた個別の VMDK を使用しました。オープンソースデータベースのロードテスト用アプリケーションである

HammerDB を使用し、OLTP アプリケーションには TPCC ベンチマークを、データ分析ワークロードには TPC-H ベンチマークを実行します。このホワイトペーパーの各種テスト全体を通じて、OLTP トランザクションワークロードをシミュレートしてテスト結果の適合性と信頼性を確保するため、TPCC ベンチマーク仕様が選択されています。

TPCC ベンチマーク (正式な仕様は tpc.org (TPCC 公式サイト) で入手可能) はよく知られた業界標準 OLTP ベンチマークで、顧客からの受注を処理して製品を供給する会社のコンピュータシステムを模しています。その会社は 100,000 点の商品を販売し、倉庫に在庫を保管しています。各倉庫には 10 の販売地域があり、各販売地域で 3000 の顧客にサービスを提供しています。顧客がその会社に電話すると、オペレーターが注文を受けます。各注文内容には数点の商品が含まれ、通常はその地域の倉庫から出荷されます。しかし、受注時点で在庫切れの商品が数点あり、別の倉庫から供給されます。会社の規模は固定されておらず、会社の成長に応じて倉庫と販売地域を追加できることに注意してください。このため、テストスキーマの大きさを自由に変わることができ、大きなスキームを使用すると大規模な TPC-C データベースとなり、増大したトランザクションを処理するために強力なコンピュータシステムが必要です (HammerDB)。

この記事では、さまざまな倉庫数 (スキーマ) と仮想ユーザー数を用いて各種テストを実行します。この条件は各テストの冒頭に記載され、テスト結果で説明されています。実行したすべてのテストについて、各テストの HammerDB の結果を記録すると同時に、Windows パフォーマンスモニタ (Perfmon) と、Windows PowerShell のネイティブモジュールの Get-counter や、vCenter サーバーで利用可能な vSAN パフォーマンスモニタを使用して CPU、ネットワーク、メモリ、ディスクの統計も取得します。

III. vSAN ストレージパフォーマンス

SQL テストを実行する前に、このホワイトペーパーで重視する構成に基づいて、vSAN データストアのパフォーマンスをテストし、DC1500M NVMe および Micron 5200 eco SATA SSD vSAN データストアに期待できるパフォーマンスのレベルを評価しました。vSAN データストアのベンチマーキングには、VMware の推奨ツールの [HCIBench v2.5.3](#) を使用しました。この自動化ツールキットは、vSAN クラスタ内の全ホストに分散している複数の VM を展開しながら、並行して全ゲスト VM 上の Vdbench を使用し特定のワークロードを実行します。DC1500M NVMe vSAN クラスタと Micron 5200 eco SATA クラスタで 6 つの VM を実行した結果をいくつか示します。

図 1.3 と 1.4 では、DC1500M NVMe vSAN データストアと Micron 5200 eco SATA SSD vSAN データストアについて、さまざまなブロックサイズの場合の 30 分間の混合ワークロード (持続的に読み取り 70%、書き込み 30% の割合のランダムワークロード) の結果を示します。ブロックサイズが 4k の時、DC1500M NVMe vSAN データストアでの 70% 読み取り/30% 書き込みの IOPS は、SATA SSD vSAN データストアの 2 倍 (NVMe vSAN の 355k に対して、SATA SSD vSAN は 178K) になり、各 IO の完了時間は 33% 短縮します (NVMe vSAN の 0.4ms に対して、SATA SSD vSAN は 0.6ms)。IO 転送量が増加するにつれ、NVMe のパフォーマンスの優位性が明確になります。ブロックサイズが 64k の時、70% 読み取り、30% 書き込みの割合のランダムワークロードでは、NVMe vSAN データストアの IOPS が 3 倍になり (NVMe vSAN の 121240 に対して、SATA SSD vSAN は 31756)、IO あたりのレイテンシが 66% 向上します (NVMe vSAN の 2.1ms に対して、SATA SSD vSAN は 6.4 ms)。

図 1.5 と 1.6 では、さまざまなブロックサイズの DC1500M NVMe および Micron 5200 eco SATA SSD vSAN データストアについて、HCIBench の持続的な読み取りおよび書き込みスループットおよびレイテンシの比較を示します。DC1500M NVMe データストアからのスループットは 17.8GB/秒 (128k) で持続でき、SATA SSD vSAN データストアからのスループット (2.79GB/秒) の

6.3 倍、レイテンシは 5 分の 1 (0.9ms に対して、SATA vSAN は 4.4ms) でした。書き込みでは、DC1500M vSAN の持続的な書き込みスループット (128k) は 6.7GB/秒で、SATA vSAN の 5.9 倍であり、レイテンシは 5 分の 1 でした。

NVMe および SATA vSAN データストアのこの基本的なパフォーマンスの差が、SQL パフォーマンスの場合ではどの程度拡大するのでしょうか？ NVMe のパフォーマンス優位性は高価格を正当化できるのでしょうか？ SQL バックアップまたはリストア操作が短時間で完了し、ミッションクリティカルなワークロードへの影響を軽減するのでしょうか？ この後のセクションでは、いくつかの実験を行い、この疑問の答えを探します。

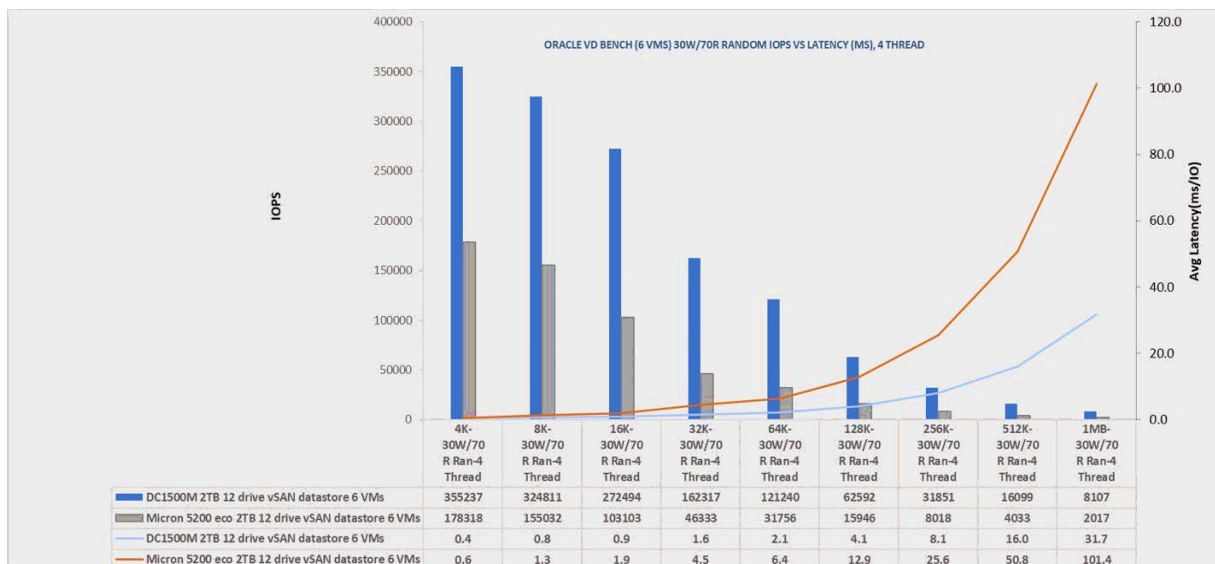


図 1.3 DC1500M vSAN データストア vs Micron 5200 eco vSAN データストア、4k 70R/30W、ランダム、QD=8、スレッド=4、6 VM HCIbench IOPS vs 平均レイテンシ(ms)

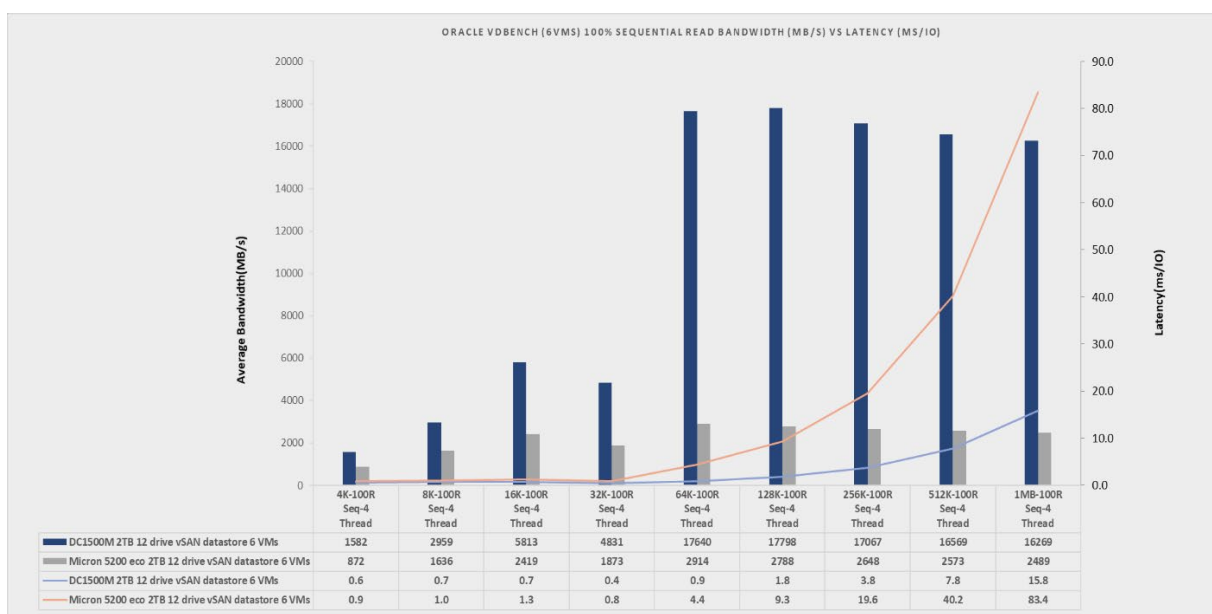


図 1.4 DC1500M vSAN データストア vs Micron 5200 eco vSAN データストア、100R/0W、順次、QD=8、スレッド=4、HCIbench 6 VM 読み取りスループット(MB/秒) および平均読み取りレイテンシ(ms/I/O)

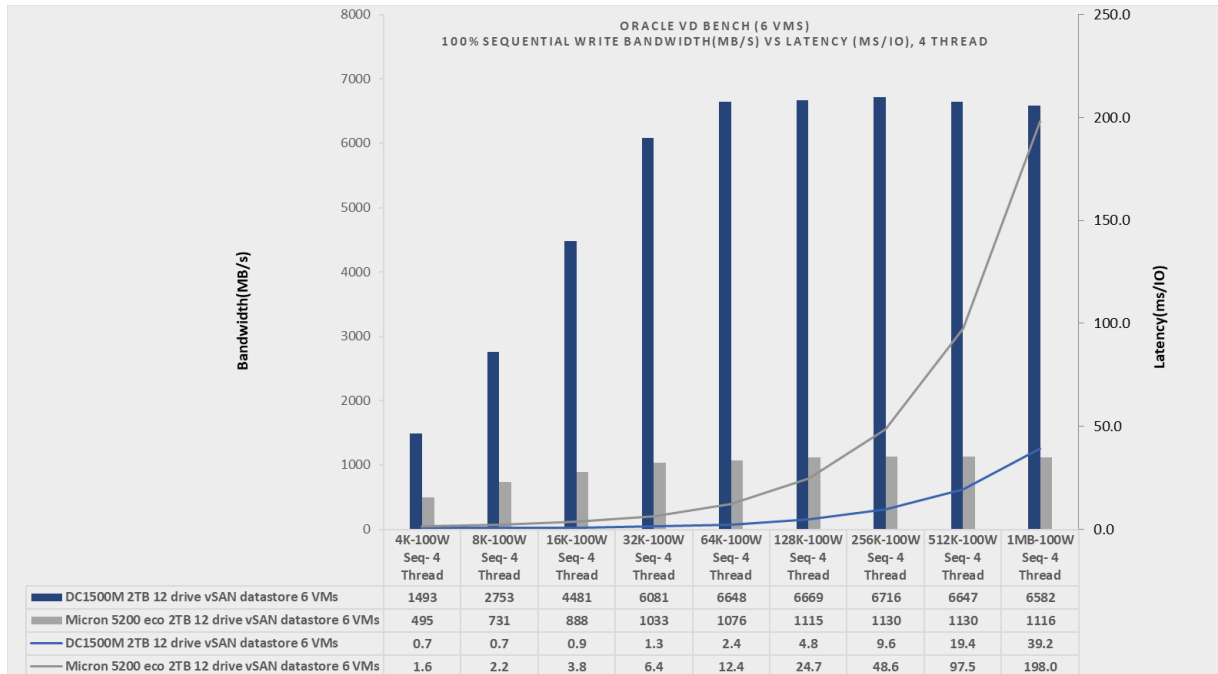


図 1.5 DC1500M vSAN データストア vs Micron 5200 eco vSAN データストア、100W/0R、順次、QD=8、スレッド=4、HCIBench 6 VM 読み取りスループット (MB/秒) および平均読み取りレイテンシ (ms/IO)

テスト結果

テスト 1: 各種容量の DRAM を用いた DC1500M 960GB vSAN SQL Server 2017 VM

vSAN データストアストレージ構成:DC1500M 960G FW S67F0103/ディスクグループ 3 つ、合計ディスクグループ 4 つ (サーバーあたり 1 つ)、NVMe vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。

テスト 1a の説明	テスト 1b の説明	テスト 1c の説明
<p>NVMe テスト環境の DC1500M vSAN データストアからプロビジョニングした仮想ディスク。</p> <p>倉庫の数が 1200 のデータベーススキーマは、100GB のデータベースが選択されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 128GB RAM が割り当てられました。</p> <p>もうひとつの vSAN VM には 16c/128GB RAM が割り当てられ、SUT にトランザクションを送る負荷生成サーバーとして機能するように、プロビジョニングされました。</p> <p>作成された仮想ユーザーシーケンスは、1、2、3、5、8、13、21、34、55、89 です。</p> <p>立ち上がり時間は 2 分、ユーザーシーケンスのテスト時間は 5 分を選択しました。</p>	<p>テスト 1a と同様ですが、データ領域への IO を増やすため、ゲスト VM に割り当てられる DRAM は 32 GB へ減らしました。</p> <p>今回もリモート負荷生成サーバーを SUT へのトランザクション送信に使用しましたが、LGS に割り当てられた DRAM も 32GB へ減らしました。</p>	<p>テスト 1a と同様ですが、データ領域への IO を増やすため、ゲスト VM に割り当てられる DRAM は 32 GB へ減らしました。</p> <p>また、このテストは、ネットワークのボトルネックを解消するために SUT VM 上でローカルで実行されました。</p>

図 2.1 テスト 1:DC1500M vSAN データストアの異なる DRAM 構成

テスト1の目標は、SQL サーバーにさまざまな容量のメモリを割り当てて、オールフラッシュ DC1500M NVMe vSAN データストアを搭載した VMware vSAN 上の SQL Server 2017 で TPCC ベンチマークに期待されるパフォーマンスのレベルのベースラインを取得することです。SQL テスト対象システム (SUT) に割り当てる各種 DRAM の容量は、次のコンセプトに基づいて決定します。

- SQL サーバーデータベース VM に割り当てる RAM を減らすと、データ領域への I/O が増大し、(ディスク上 OLTP データベース) スキーマを含むデータベースの I/O パフォーマンスを重点的に処理できます
- SQL サーバーデータベース VM に十分な DRAM がある場合、データの大部分は OLTP テスト中にキャッシュされ、データ領域への I/O が最低限になります (メモリ内 OLTP テスト)

倉庫の数が 1200 のスキーマを作成し、TPCC データベースのサイズが 100GB 程度になりました。最初のテストでは、スキーマ全体がメモリに収まるように 128GB の DRAM を SUT に割り当てましたその後、リモート負荷生成サーバー (LGS) で仮想ユーザーシーケンスを実行し、ユーザーによるデータベースへのトランザクション送信のシミュレーションを行いました。当社のスキーマサイズと、SQL サーバー VM に割り当てられた CPU/メモリリソースの容量に合わせてユーザーの数は 1~89 人としました。テスト完了後、TPCC データベースをリストアし、SUT と LGS に割り当てられた DRAM を 32GB に減少し、同じテストを同じユーザーシーケンスで再度実行しました。最後に、リモート負荷生成サーバーに伴うネットワークのボトルネックをなくすため、テスト対象システムの VM で同じテストをローカルで実行しました。

テスト1結果:各種容量の DRAM を用いた DC1500M 960GB vSAN SQL Server 2017 VM

図 2.2 および 2.3 は、DC1500M vSAN データストアを使用してテスト 1a、1b、1c で達成した 1 分あたりのトランザクション (TPM) と 1 分あたりの新規発注 (NOPM) を示しています。すべてのテスト実行で、仮想ユーザー数の増加に従って、TPM および NOPM の拡大が見られました。仮想ユーザーの数が 89 に達したときに、ほぼメモリ内の OLTP データベースを使用する SQL Server 2017 VM が、1,113,300 TPM と 259,631 NOPM を達成することができました。SUT および LGS VM で DRAM を 32GB に減少すると、958,338 TPM と 208311 NOPM を達成できましたが、SUT VM でローカルにテストを実行した場合、驚異的な 1,463,290 TPM および 318092 NOPM を達成しました。

この点でエンタープライズ NVMe SSD のレイテンシ優位性が実際に機能していることがわかります。つまり、トランザクションの件数が増えて、スキーマをキャッシュするのに十分なメモリが割り当てられていない場合や、SQL サーバーデータベースがメモリからトランザクションログファイルにデータを書き込む必要がある場合に、NVMe 仮想ディスクは十分な速度で応答して、高いトランザクションスループットを維持でき、また、CPU がボトルネックになるまで拡張できます。図 2.4 を見ると、テスト 1c で仮想ユーザーの数が 89 人に達しても、各ユーザーが 1 分あたり 16,441 件のトランザクションを処理できることがわかります。これらの経験的結果に基づいて、NVMe ハイパーコンバージドインフラを構築すると、SQL Server 2017 に割り当てられた余分な DRAM のコストを節減できると結論づけることができます。

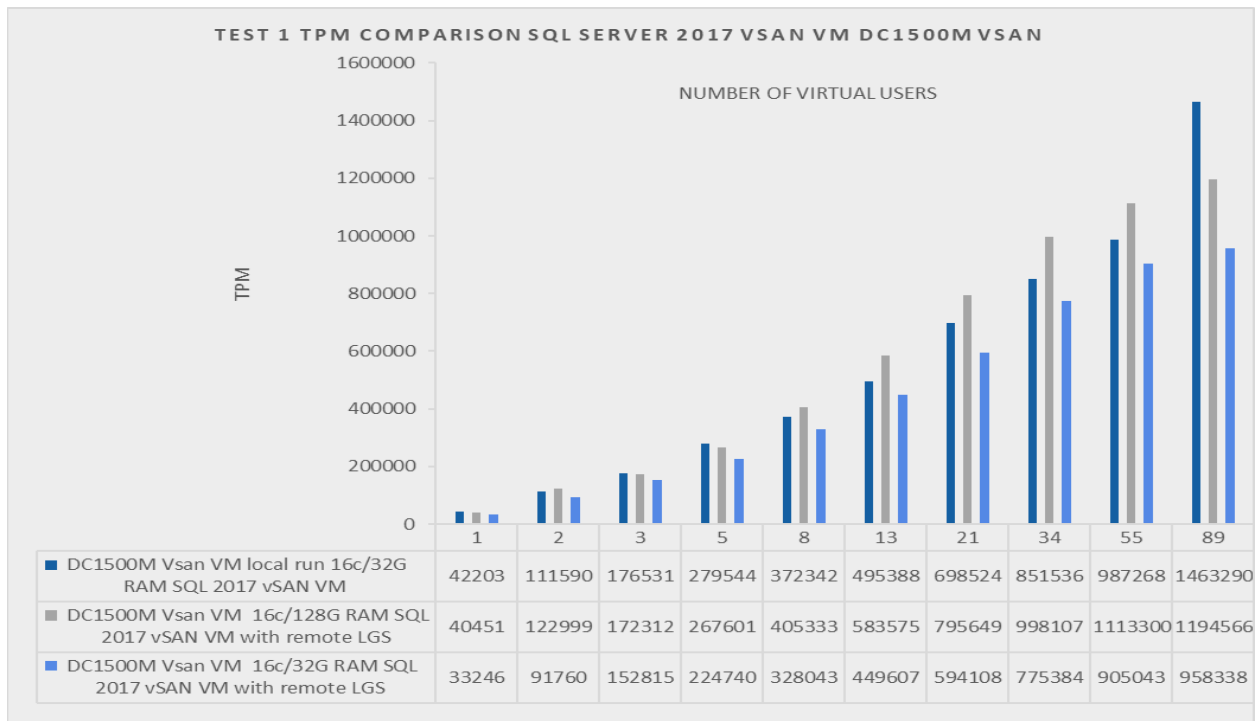


図2.2 テスト1a、b、c:異なるDRAM容量のDC1500M vSAN データストアTPMの比較

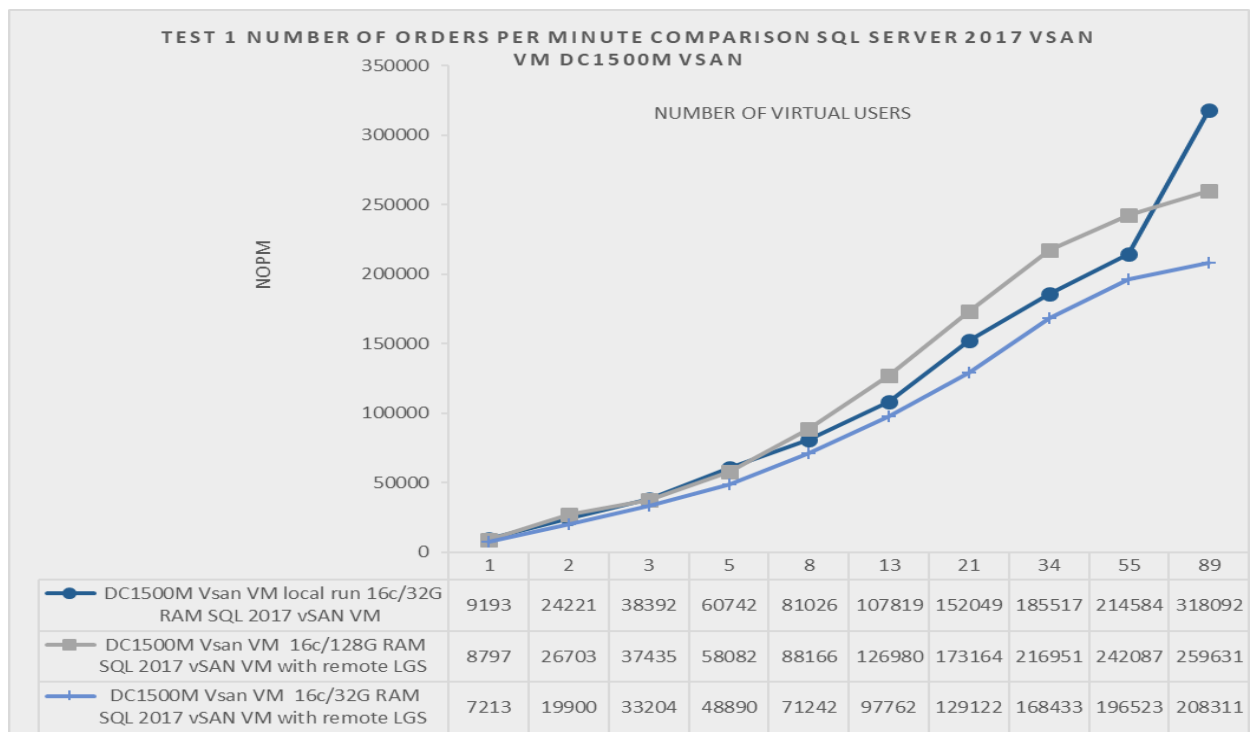


図2.3 テスト1a、b、c:異なるDRAM容量のDC1500M vSAN データストアNOPMの比較

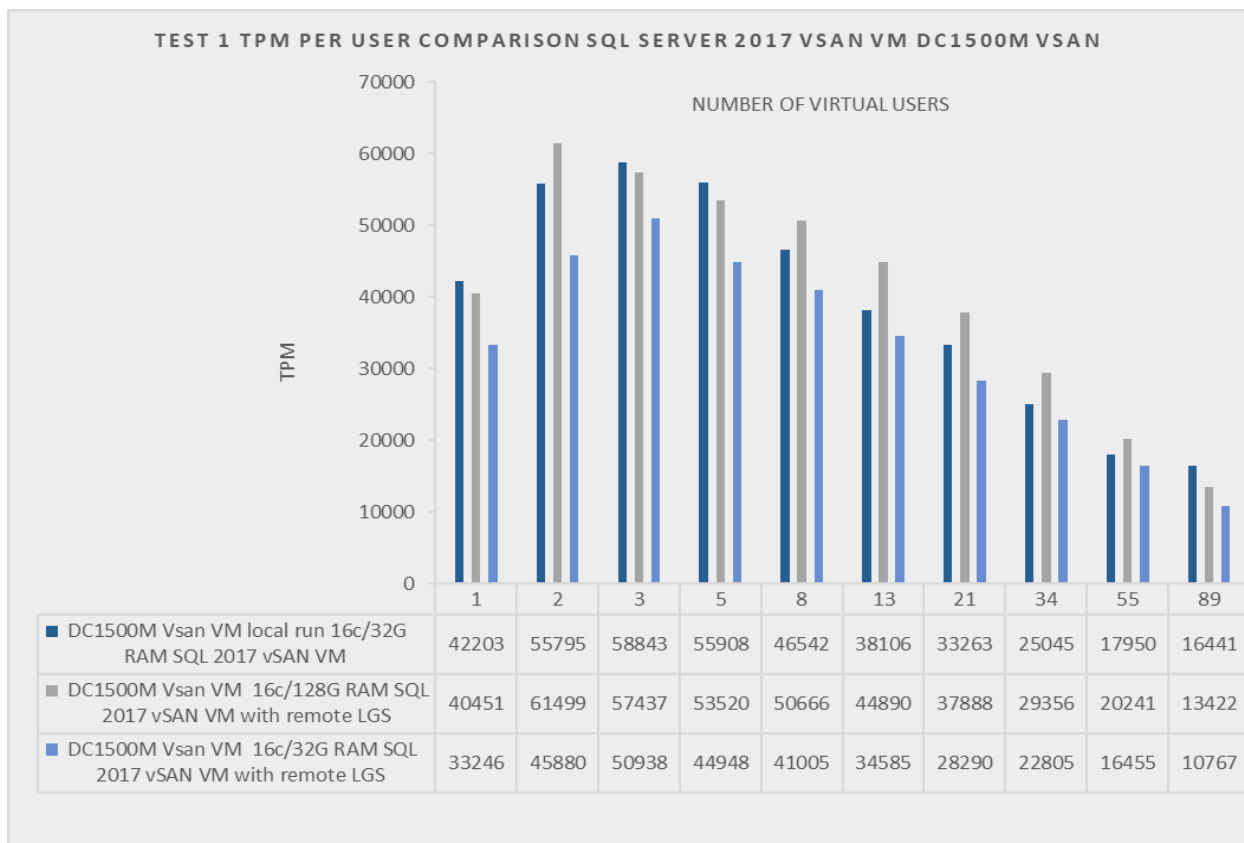


図2.4 テスト1a、b、c:異なるDRAM容量のDC1500M vSAN データストアTPMの比較

テスト 2: Kingston DC500M SATA SSD、Micron 5200 eco SATA SSD および DC1500M NVMe SSD vSAN データストアでの SQL Server 2017 のパフォーマンス比較

<ul style="list-style-type: none"> • テスト 1a の NVMe vSAN データストアストレージ構成: DC1500M 960G FW S67F0103/ディスクグループ 3 つ、合計ディスクグループ 4 つ (サーバーあたり 1 つ)、NVMe vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。(テスト 1a) • テスト 1b の SATA vSAN データストアストレージ構成: DC500M 1920G FW SCEJK2.8/ディスクグループ 3 つ、合計ディスクグループ 3 つ (サーバーあたり 1 つ)、SATA vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。(テスト 1b) • テスト 1c の SATA vSAN データストアストレージ構成: Micron 5200 ECO 1920G FW D1MU004/ディスクグループ 3 つ、合計ディスクグループ 3 つ (サーバーあたり 1 つ)、SATA vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。(テスト 1b) 		
テスト 2a の説明	テスト 2b の説明	テスト 2c の説明
<p>NVMe テスト環境の DC1500M vSAN データストアからプロビジョニングした仮想ディスク。 倉庫の数が 1200 のデータベーススキーマは、100GB のデータベースが選択されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。 作成された仮想ユーザーシーケンスは、1、2、3、5、8、13、21、34、55、89 です。 立ち上がり時間は 2 分、ユーザーシーケンスのテスト時間は 5 分を選択しました。 テストは SUT VM 上でローカルで実行されました。</p>	<p>SATA テスト環境の D500M vSAN データストアからプロビジョニングした仮想ディスク。 倉庫の数が 1200 のデータベーススキーマは、100GB のデータベースが選択されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。 作成された仮想ユーザーシーケンスは、1、2、3、5、8、13、21、34、55、89 です。 立ち上がり時間は 2 分、ユーザーシーケンスのテスト時間は 5 分を選択しました。 テストは SUT VM 上でローカルで実行されました。</p>	<p>仮想ディスクは SATA テスト環境の Micron 5200 eco vSAN データストアからプロビジョニングされました。 倉庫の数が 1200 のデータベーススキーマは、100GB のデータベースが選択されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。 作成された仮想ユーザーシーケンスは、1、2、3、5、8、13、21、34、55、89 です。 立ち上がり時間は 2 分、ユーザーシーケンスのテスト時間は 5 分を選択しました。 テストは SUT VM 上でローカルで実行されました。</p>

図3.1 テスト 2 の説明: SATA と DC1500M NVMe SSD vSAN データストア上の SQL Server 2017 のパフォーマンスの比較

テスト 2 では、Kingston DC1500M エンタープライズ NVMe vSAN データストア、[Kingston DC500M](#) および Micron 5200 eco SATA SSD vSAN データストアの 3 つの異なるデータストアでローカルに実行した場合の、SQL Server 2017 テスト対象システム VM の TPCC ベンチマークのパフォーマンスを比較します。テスト 2 は、データ領域への I/O を増加し、スキーマを含むデータベースの IO パフォーマンスを強調するために、SQL Server 2017 VM テスト対象システム上でローカルで実行し、加えて、スキーマサイズと SQL サーバー VM に割り当てた CPU/メモリリソースに合わせてユーザー数を 1~89 に拡大するユーザーシーケンスをテストしました。

テスト 2 結果: Kingston DC500M SATA SSD、Micron 5200 eco SATA SSD および DC1500M NVMe SSD vSAN データストアでの SQL Server 2017 のパフォーマンス比較

図 3.2 および 3.3 は、テスト 2a、2b、2c で達成した 1 分あたりのトランザクション (TPM) と 1 分あたりの新規発注 (NOPM) を示しています。すべてのテスト実行で、仮想ユーザー数の増加に従って、TPM および NOPM の拡大が見られましたが、NVMe と SATA では拡大が劇的に異なります。仮想ユーザー数 89 の場合、DC1500M を用いた vSAN データストア SQL Server 2017 VM では、1,463,290 TPM と 318,092 NOPM を達成しました。これと比較して、DC500M SQL サー

バー vSAN VM では 738,067 TPM/160,410 NOPM、Micron 5200 eco vSAN データストアでは 628499 TPM/136436 NOPM を達成しました。つまり、NVMe を用いた vSAN データストア上で同じ数の DC1500M NVMe ドライブを使用すると、SSD の台数が同じで SATA を用いた vSAN データストアと比較して、トランザクションスループットと 1 分あたりの注文数を実質的に倍増できます。これを実際の業務に即して言い換えると、DC1500M のようなエンタープライズ NVMe ソリューションを使用して VMware インフラをアップグレードした場合、89 人のユーザーがデータベースにトランザクションを同時に送信した際に、各ユーザーで処理できるトランザクションの件数が 235% 増加します (つまり 1 分あたりに処理可能な注文数が増加します) (図 3.4)。

図 3.5 は、テスト 2a、b、および c の平均 CPU アイドル時間と仮想ユーザー数の比較を示しています。これは、仮想ディスク効率の効果的な測定方法であり、トランザクション件数が増加し、SQL サーバーデータベースがメモリからトランザクションログファイルにデータを書き込む必要がある場合に、仮想ディスクがどの程度の速さで応答可能かを示します。仮想ユーザー数が 89 の場合、DC1500M NVMe を用いた vSAN VM の CPU アイドル時間 (iowait) は 15.5% で、DC500M を用いた VM は 37.8%、Micron 5200 を用いた VM は 44.2% です。つまり、NVMe 仮想ディスクの応答時間は IO 要求よりもずっと速く、CPU が IO の完了を待機してアイドル状態にならないため、処理できるトランザクション件数が増加します。これを実際の業務に即して言い換えると、VMware インフラを NVMe にアップグレードすると、SQL サーバー VM に割り当てられた仮想コアを効率的に使用して、トランザクションスループットを増大できますので、速度の遅いストレージ Tier で実行されるレガシー SQL VM から不要なコアを除去して、コストを削減できます。

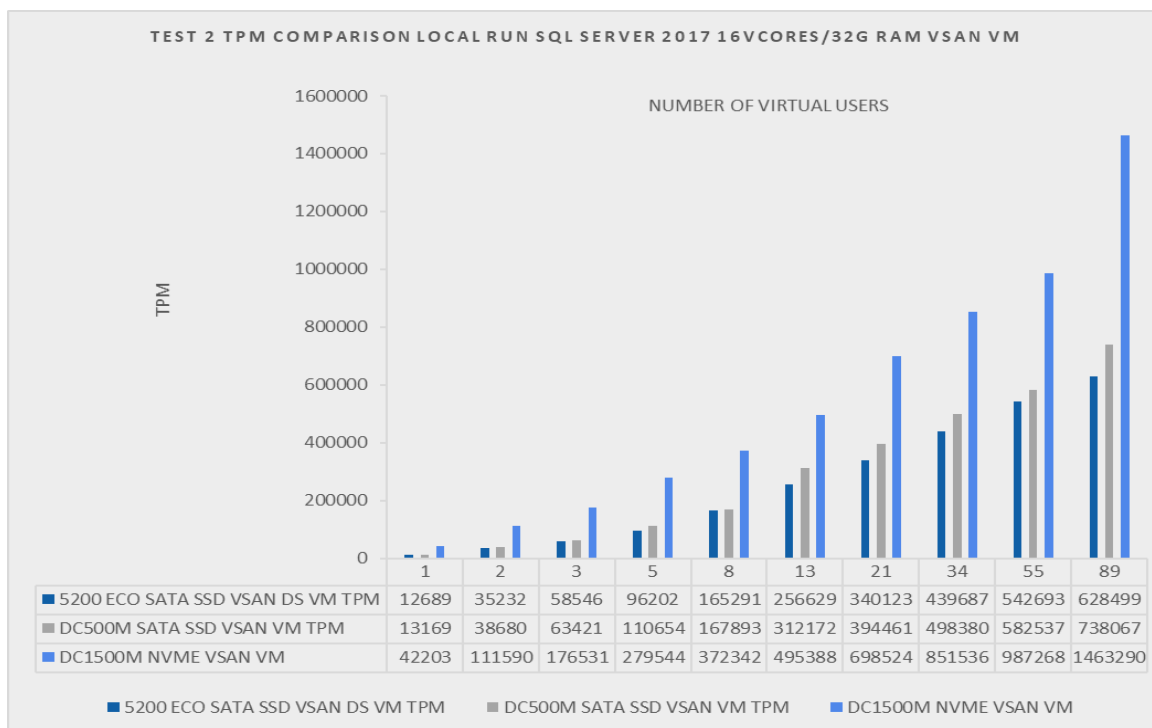


図 3.2 テスト 2: NVMe と SATA vSAN データストアの TPM の比較



図3.3 テスト2:NVMe と SATA vSAN データストアのNOPM の比較

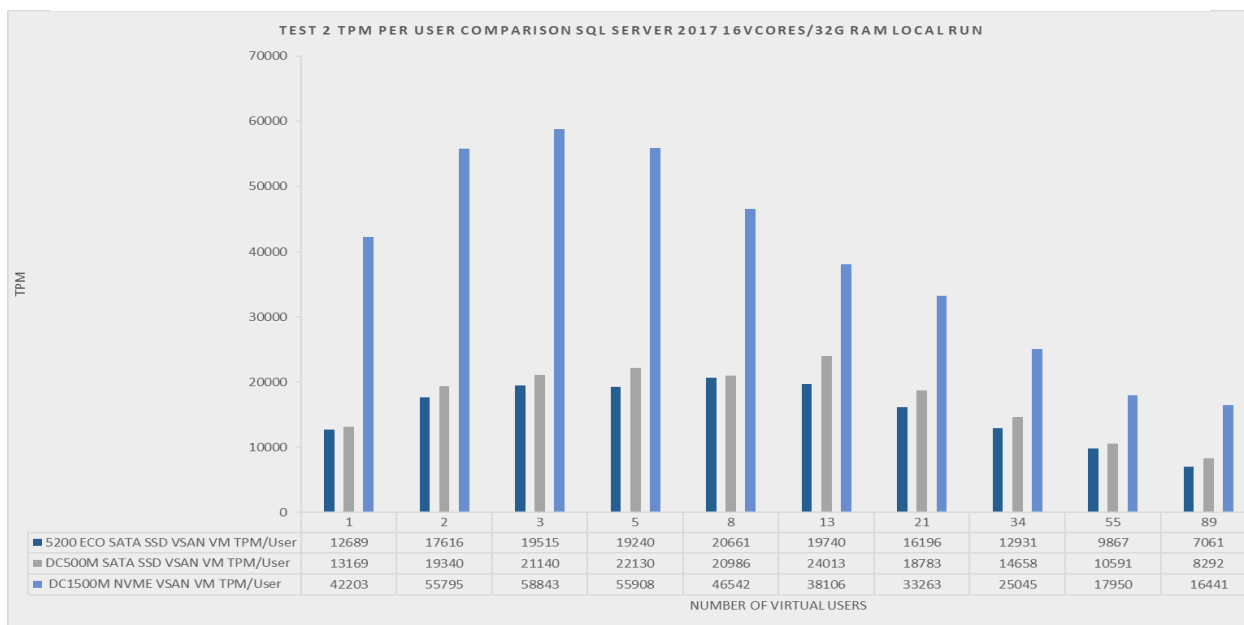


図3.4 テスト2:NVMe と SATA vSAN データストアのユーザーあたりのTPM の比較

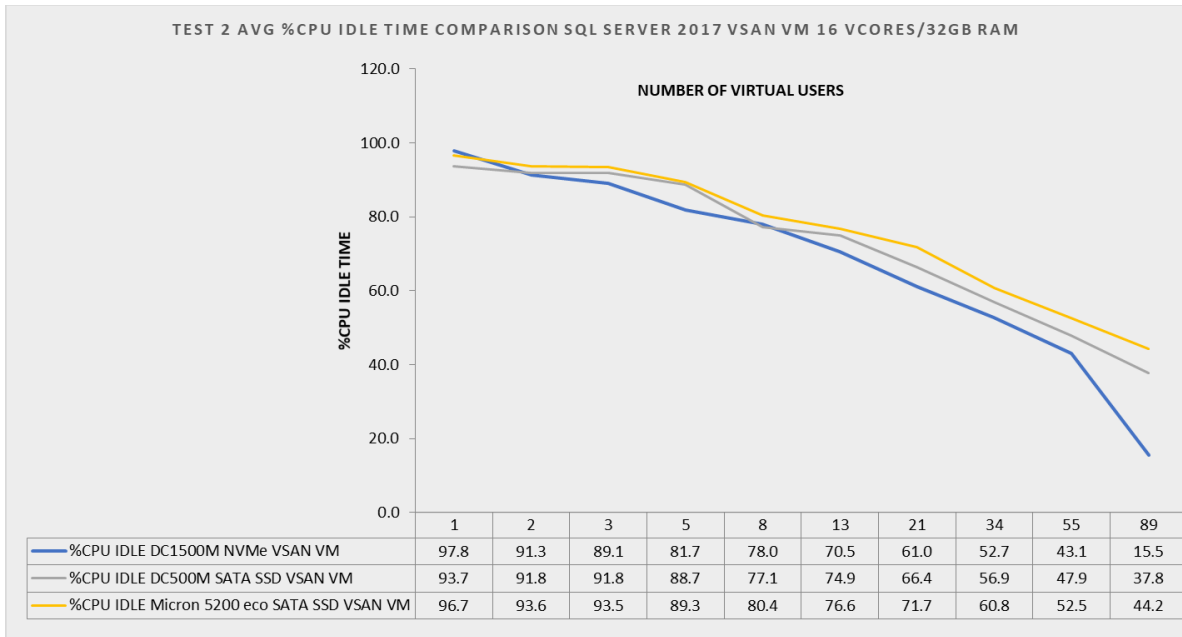


図3.5 テスト2:NVMe と SATA vSAN データストアのCPU アイドル率の比較

テスト3:DC1500M NVMe と Micron 5200 eco SATA vSAN データストアのSQL Server 2017のパフォーマンス比較、スキーマサイズを拡大し、テスト時間を延長

<ul style="list-style-type: none"> テスト3aのNVMe vSAN データストアストレージ構成:DC1500M 960G FW S67F0103/ディスクグループ3つ、合計ディスクグループ4つ(サーバーあたり1つ)、NVMe vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載したSQL Server 2017。(テスト3a) テスト3bのSATA vSAN データストアストレージ構成: Micron 5200 ECO 1920G FW D1MU004/ディスクグループ3つ、合計ディスクグループ3つ(サーバーあたり1つ)、SATA vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載したSQL Server 2017。(テスト3b) 	
<p>テスト3aの説明</p> <p>NVMe テスト環境のDC1500M vSAN データストアからプロビジョニングした仮想ディスク。倉庫の数が2000のデータベーススキーマは、157GBのデータベースが選択されていることを示します。テスト対象システムのVM(SUT)には、40のvCoreと32GB RAMが割り当てられました。作成された仮想ユーザーシーケンスは、1、2、4、8、16、32、64、89、128です。立ち上がり時間は10分、ユーザーシーケンスのテスト時間は20分を選択しました。テストはSUT VM上でローカルで実行されました。</p>	<p>テスト3bの説明</p> <p>仮想ディスクはSATA テスト環境のMicron 5200 eco vSAN データストアからプロビジョニングされました。倉庫の数が2000のデータベーススキーマは、157GBのデータベースが選択されていることを示します。テスト対象システムのVM(SUT)には、40のvCoreと32GB RAMが割り当てられました。作成された仮想ユーザーシーケンスは、1、2、4、8、16、32、64、89、128です。立ち上がり時間は10分、ユーザーシーケンスのテスト時間は20分を選択しました。テストはSUT VM上でローカルで実行されました。</p>

図4.1 テスト3の説明: Micron 5200 eco SATA およびDC1500M NVMe SSD vSAN データストアのSQL Server 2017のDBストレステスト

このテストは、SQL Server 2017 テスト対象システム VM を、Kingston DC1500M エンタープライズ NVMe vSAN データストアと Micron 5200 eco SATA SSD vSAN データストアの、2つの異なるデータストアでローカルに実行した場合に、TPCC ベンチマークのパフォーマンスを比較するため、およびそれ以前のテスト結果を検証するために、大規模なデータベーススキーマサイズの長時間のストレステストとして設計されました。今回はスキーマサイズとして倉庫数 2000 を選択し、TPC-C データベースのサイズが 157 GB になりました多くのトランザクションを生成し、トランザクションスループットを飽和状態にするのに十分な CPU リソースを割り当てるために各 SQL サーバー VM に 40 の仮想コアを使用しましたが、テスト IO バウンドの作成には 32 GB の RAM のみを割り当てました。仮想ユーザーシーケンスを微調整してユーザー数を 1~128 に拡大し、各仮想ユーザーシーケンスを実行可能な時間を長くしました (20 分。そのうち立ち上がり時間は 10 分)。これによって、テストの実行時間全体にわたって、ディスクのレイテンシの測定値を収集できました。

テスト 3 結果:DC1500M NVMe と Micron 5200 eco SATA vSAN データストア の SQL Server 2017 のパフォーマンス比較、スキーマサイズを拡大し、テスト時間を延長

図 4.2 および 4.3 は、テスト 3a と 3b で達成した 1 分あたりのトランザクション (TPM) と 1 分あたりの新規発注 (NOPM) を示しています。時間が長くなっても、NVMe と SATA SSD を用いた SQL Server 2017 VM は、仮想ユーザー数が増大して 128 になるまで拡張できましたが、NVMe の方が急激に拡張しました。ユーザー数 89 の時点で 1.84M TPM、361743 NOPM を達成し、これと比較して SATA SSD を用いた vSAN SQL VM では 0.96TPM と 184451 NOPM でした。つまり、DC1500M NVMe を用いた vSAN データストア では、同じ数の vCore を搭載し、DRAM を割り当て、Micron 5200 eco vSAN を用いた VM に比べて TPM/NOPM が 200% 増大しました。

図 4.4 と 4.5 では、SQL NVMe および SATA SSD を用いた vSAN VM 上で、Windows の perfmon を使用して収集したユーザー数に対する、平均仮想ディスクレイテンシと、99% の仮想ディスクレイテンシの比較を示しています。それぞれの仮想ユーザーシーケンスで、DC1500M を用いた仮想ディスクは、ユーザー数の拡大が続いても、1ms 未満の平均レイテンシを維持できます。仮想ユーザー数 89 の時点で、DC1500M を用いた仮想ディスクの平均レイテンシは 0.92ms/IO で、これと比較して SATA SSD を用いた仮想ディスクでは 2.36ms/IO で、NVMe と比較した平均レイテンシの増加は 256% でした。さらに興味深いのは、ユーザー数 89 の時点の QoS 99% レイテンシで、DC1500M 仮想ディスクは全 IO の 99% を 1.61ms で完了できましたが、SATA SSD を用いた仮想ディスクは全 IO の 99% を 7.05ms で完了し、NVMe と比較して 437% の増加でした。ここでは NVMe と SATA のレイテンシの差が目を見ますが、DC1500M は持続的な OLTP ワークロード全体を通じて予測可能な QoS レイテンシを維持するように設計されているため、レイテンシが突然急上昇することはなく、仮想ユーザー数が増加しても、ブロック層で並行 IO 要求が増加します。これを実際の業務の視点から見ると、SATA SSD から DC1500M のようなエンタープライズ NVMe ドライブに VMware インフラをアップグレードすると、トランザクションを拡張し、トランザクションレイテンシを大幅に短縮でき、アプリケーションを迅速に拡張して、コストを徐々に削減できます。

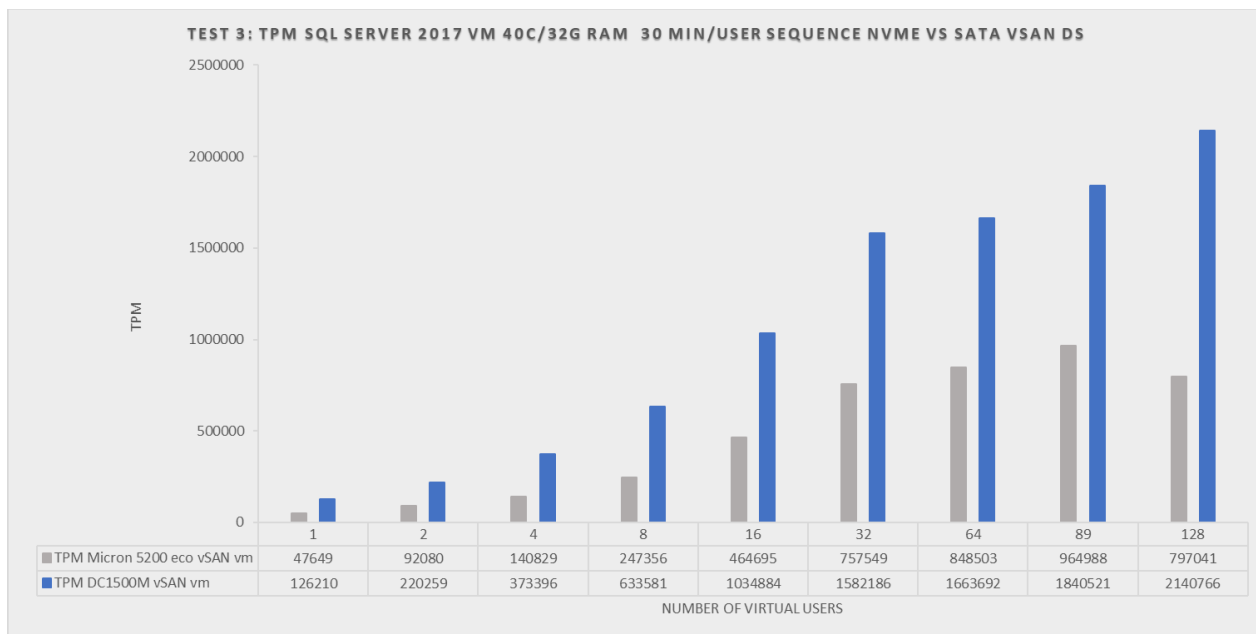


図4.2 テスト3 TPM の比較 Micron 5200 eco SATA およびDC1500M NVMe SSD vSAN データストアのSQL Server 2017 のDB ストレステスト

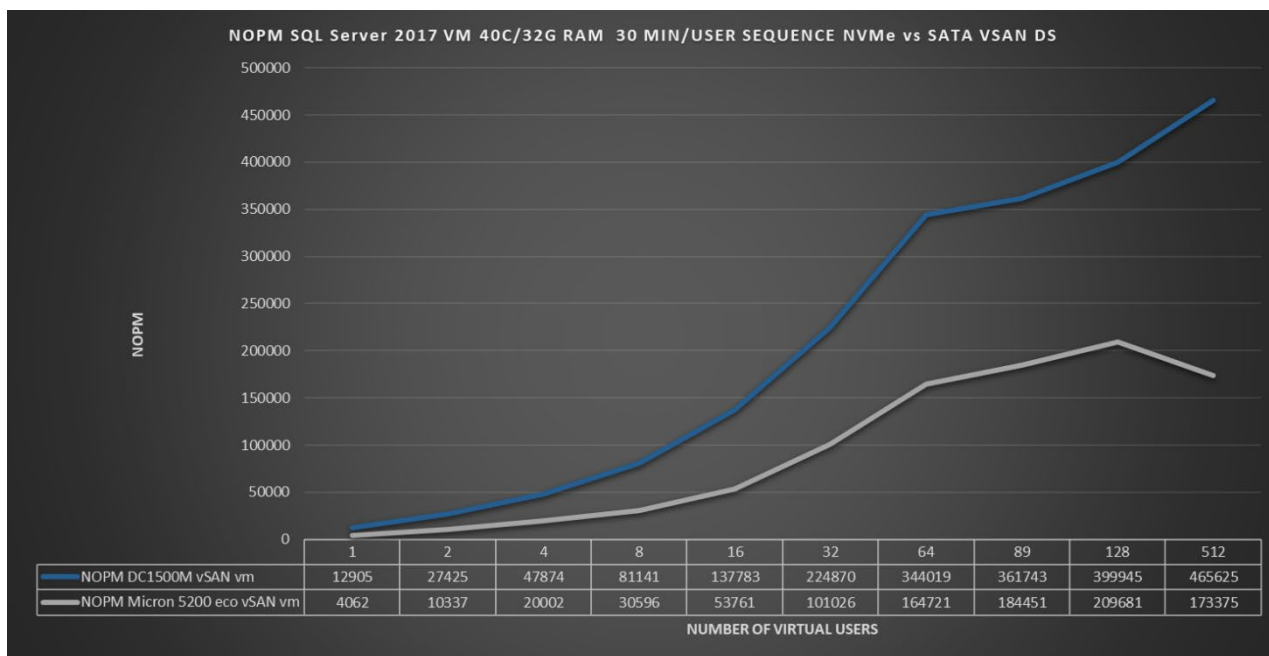


図4.3 テスト3 TPM の比較 Micron 5200 eco SATA およびDC1500M NVMe SSD vSAN データストアのSQL Server 2017 のDB ストレステスト

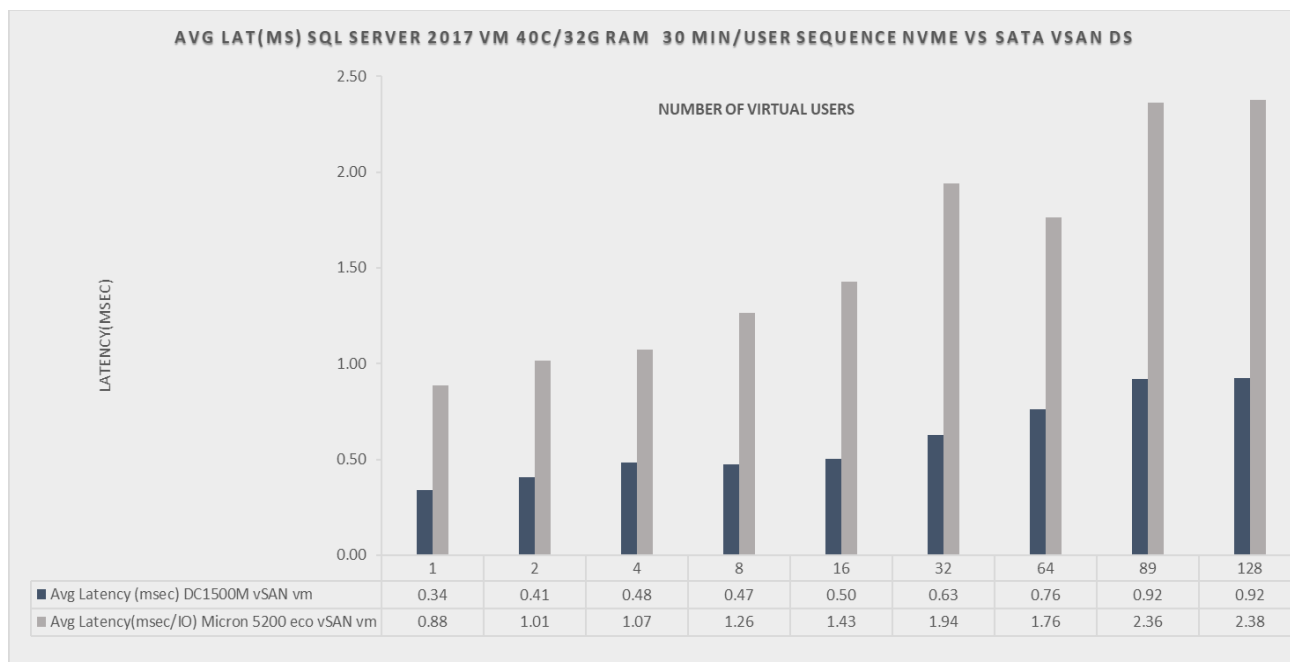


図 4.4 テスト 3 平均レイテンシ (ms) の比較 Micron 5200 eco SATA および DC1500M NVMe SSD vSAN データストアの SQL Server 2017 の DB ストレステスト

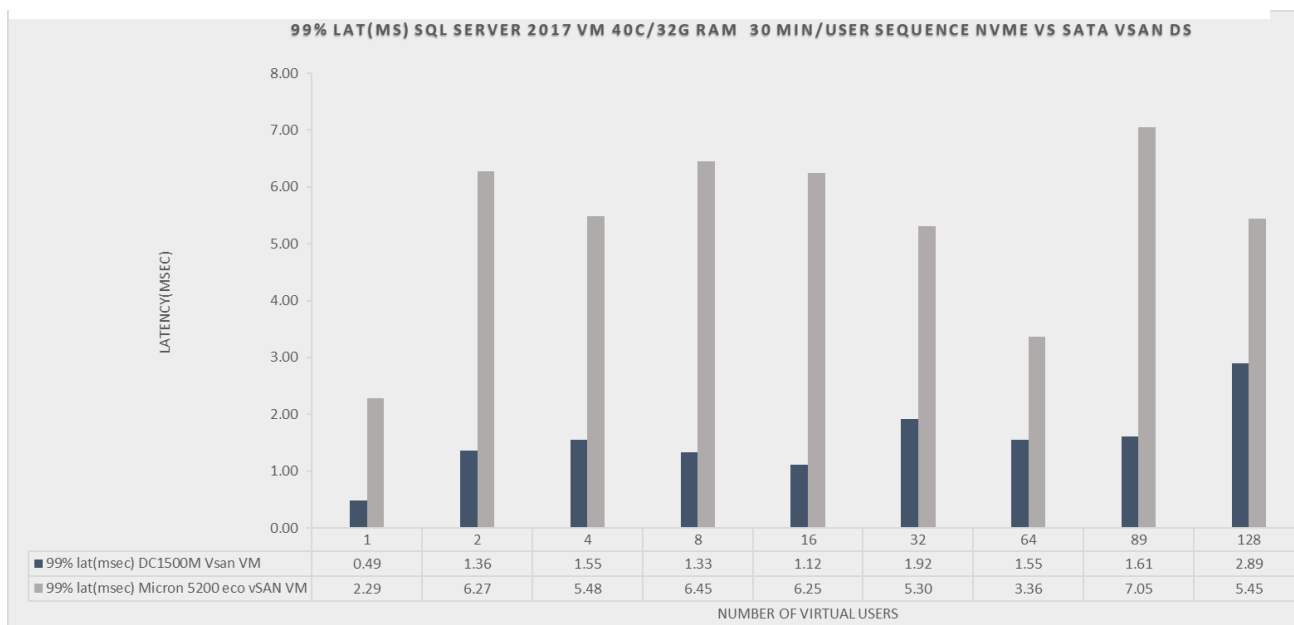


図 4.5 テスト 3 99%レイテンシ (ms) 比較 Micron 5200 eco SATA および DC1500M NVMe SSD vSAN データストアの SQL Server 2017 の DB ストレステスト

テスト 4:SQL Server 2017 のパフォーマンス比較、バックアップおよびリストアのパフォーマンス、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較

<ul style="list-style-type: none"> • テスト 3a の NVMe vSAN データストアストレージ構成:DC1500M 960G FW S67F0103/ディスクグループ 3 つ、合計ディスクグループ 4 つ (サーバーあたり 1 つ)、NVMe vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。(テスト 4a) • テスト 3b の SATA vSAN データストアストレージ構成: Micron 5200 ECO 1920G FW D1MU004/ディスクグループ 3 つ、合計ディスクグループ 3 つ (サーバーあたり 1 つ)、SATA vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。(テスト 4b) 	
テスト 4a の説明	テスト 4b の説明
<p>NVMe テスト環境の DC1500M vSAN データストアからプロビジョニングした仮想ディスク。 倉庫の数が 2000 のデータベーススキーマは、157GB のデータベース SUT 上に作成されていることを示します。 テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。 3 サイクルのバックアップ/リストアスクリプトのトリガーが発動し、Windows パフォーマンスモニタで記録した TPCC データベースとパフォーマンス測定値をバックアップおよびリストアしました テストは SUT VM 上でローカルで実行されました。</p>	<p>仮想ディスクは SATA テスト環境の Micron 5200 eco vSAN データストアからプロビジョニングされました。 倉庫の数が 1200 のデータベーススキーマは、157GB のデータベース SUT 上に作成されていることを示します。 テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。 3 サイクルのバックアップ/リストアスクリプトのトリガーが発動し、Windows パフォーマンスモニタで記録した TPCC データベースとパフォーマンス測定値をバックアップおよびリストアしました テストは SUT VM 上でローカルで実行されました。</p>

図 5.1 テスト 4 の説明: Micron 5200 eco SATA および DC1500M NVMe SSD vSAN データストアでの SQL Server 2017 のバックアップ/リストアのパフォーマンス比較

SQL データベースのバックアップおよびリストア操作は、基盤仮想ディスクのスループットとレイテンシを測定するのに適した方法です。当社では、TPC-C のバックアップ/リストア操作のトリガーが発動した際に、Windows パフォーマンスモニタで仮想ディスクの測定値を取得して、単一の NVMe および SATA を用いた vSAN VM からのスループットとレイテンシの測定値のベースラインを確立したいと考えました。

テスト 4:結果:SQL Server 2017 のパフォーマンス比較、バックアップおよびリストアのパフォーマンス、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較

図 5.2~5.4 は、テスト 4a) とテスト 4b) のバックアップ/リストアサイクルのひとつで、Windows パフォーマンスモニタスクリプトによって収集された毎秒のスループットとレイテンシを示しています。DC1500M NVMe vSAN データストアを用いた SQL サーバー VM では、TPCC データベースのバックアップ操作を 265 秒で完了し、平均スループットは 593MB/秒、平均レイテンシは 1.46ms/IO を達成しました。TPCC データベースのリストア操作は 129 秒で完了し、平均 BW は 1.4GB/秒で平均レイテンシは 2.65ms/IO でした。これを Micron 5200 eco vSAN を用いた VM と比較すると、NVMe vSAN を用いた SQL VM ではバックアップ操作が 1.5 倍の速さで完了し、リストア操作が 2.15 倍の速さで完了します。

通常、バックアップおよびリストア操作は、本番 VM に影響を与えないように時間外に実行されます。しかし、常にそうとは限りません。SQL のバックアップまたはリストア操作がピークの営業時間内に実行された場合、同じ vSAN データストアを共有する Tier 1 アプリケーションでトランザクションを実行するユーザーにレイテンシの影響を与えないように、できるだけ早く完了する必要があります。SQL データベースを移行して NVMe を用いた vSAN データストアにすると、その影響を緩和できます。バックアップ/リストア操作を時間外に実行する場合でも、短時間で完了すれば、同じリソースを共有する Tier 1 データベースのダウンタイムを短縮できます。

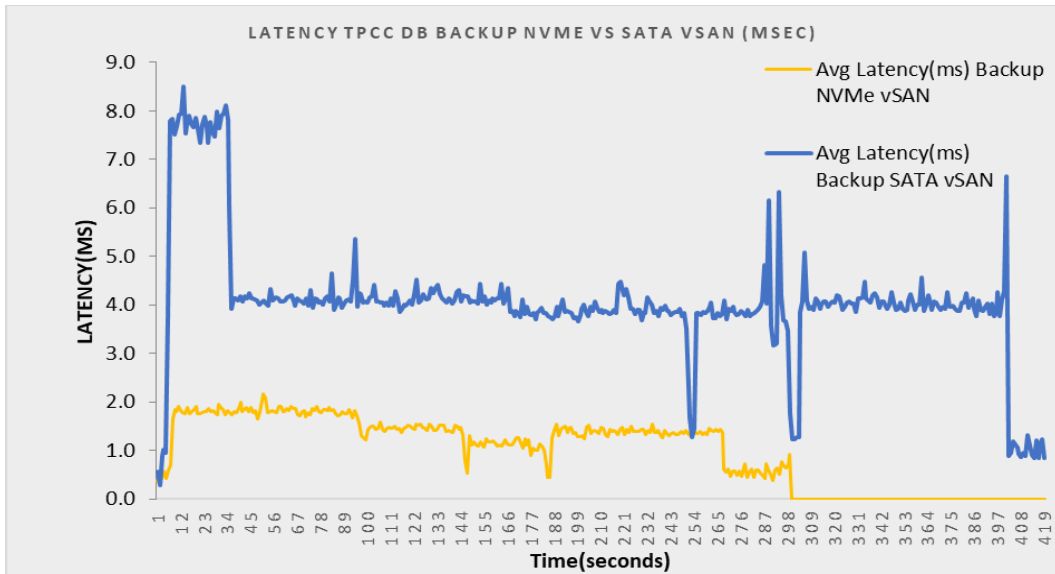


図5.2 スループットの比較SQL Server 2017 TPCC DB バックアップMicron 5200 eco SATA およびDC1500M NVMe SSD vSAN データストア (MB/秒)

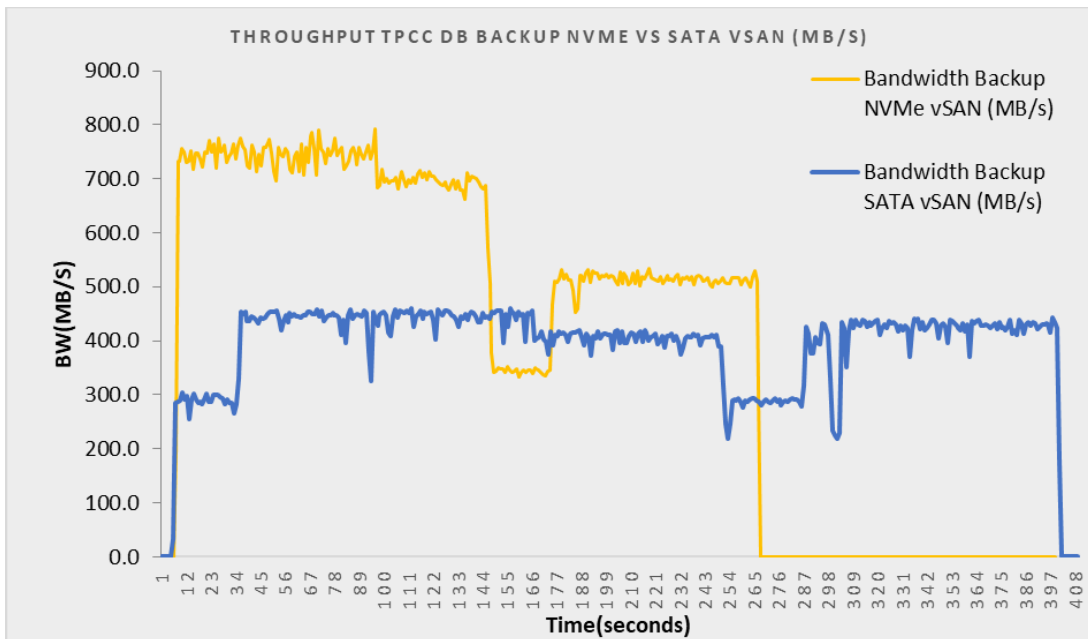


図5.3 平均レイテンシ(ms)の比較SQL Server 2017 TPCC DB バックアップMicron 5200 eco SATA およびDC1500M NVMe SSD vSAN データストア

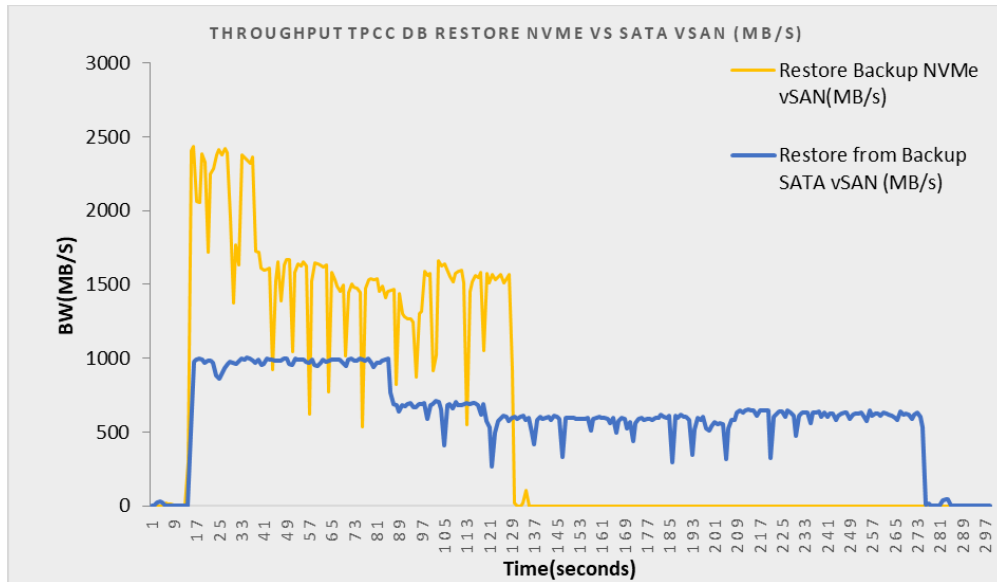


図5.4 スループットの比較SQL Server 2017 TPCC DB リストアMicron 5200 eco SATA およびDC1500M NVMe SSD vSAN データストア (MB/秒)

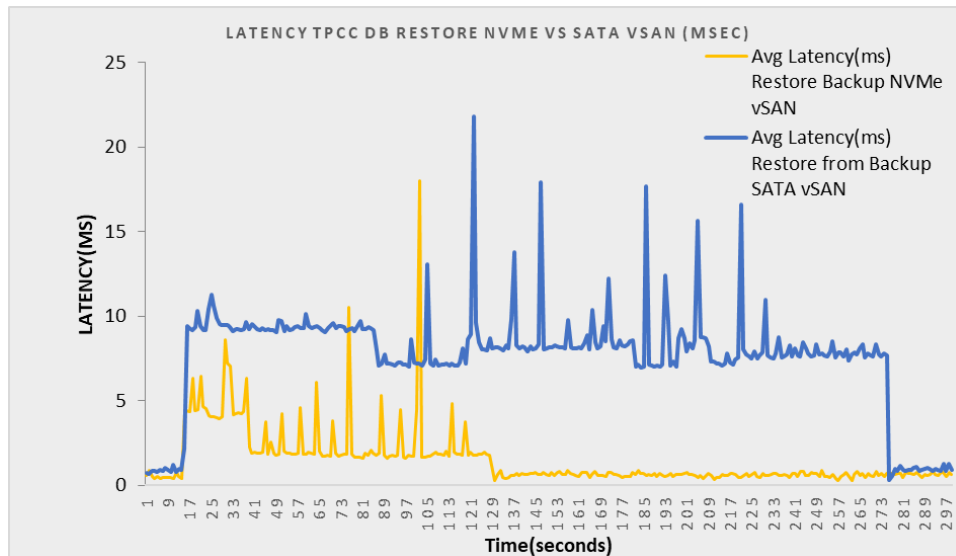


図5.5 レイテンシ(ms)の比較SQL Server 2017 TPCC DB リストアMicron 5200 eco SATA およびDC1500M NVMe SSD vSAN データストア

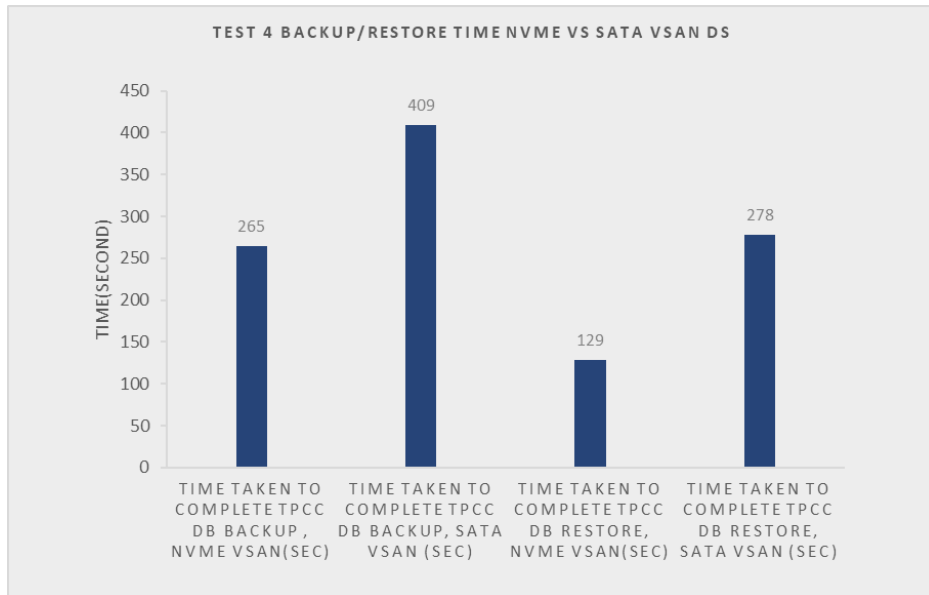


図 5.6 SQL Server 2017 TPCC DB バックアップ/リストア操作の完了までの所要時間
Micron 5200 eco SATA および DC1500M NVMe SSD vSAN データストア (秒)

テスト 5:SQL Server 2017 のパフォーマンス比較、ノイジーネイバーテスト、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較

- テスト 3a の NVMe vSAN データストアストレージ構成:DC1500M 960G FW S67F0103/ディスクグループ 3 つ、合計ディスクグループ 4 つ (サーバーあたり 1 つ)、NVMe vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。(テスト 5a)
- テスト 3b の SATA vSAN データストアストレージ構成:Micron 5200 ECO 1920G FW D1MU004/ディスクグループ 3 つ、合計ディスクグループ 3 つ (サーバーあたり 1 つ)、SATA vSAN テスト環境。Server 2019 Datacenter Guest OS を搭載した SQL Server 2017。(テスト 5b)

テスト 5a の説明	テスト 5b の説明	テスト 5c の説明	テスト 5d の説明
<p>SQL 2017 VM 仮想ディスクは NVMe テスト環境の DC1500M vSAN データストアからプロビジョニングされました。</p> <p>倉庫の数が 1200 のデータベーススキーマは、100GB のデータベース SUT 上に作成されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。</p> <p>テスト対象 SUT は 11 回クローンされ、物理サーバーあたり 3 つの SUT VM が割り当てられました (合計 12 の SUT VM)</p> <p>テストは、89 の仮想ユーザーを実行するために構成され</p> <p>各 SUT VM の立ち上がり時間は 30 分間、テスト時間は 300 分を選択しました</p> <p>テストは、12 の SUT VM で並行して起動されました</p>	<p>SATA テスト環境の Micron 5200 eco vSAN データストアからプロビジョニングした SQL 2017 仮想ディスク。倉庫の数が 1200 のデータベーススキーマは、100GB のデータベース SUT 上に作成されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。</p> <p>テスト対象 SUT は 8 回クローンされ、物理サーバーあたり 3 つの SUT VM が割り当てられました (合計 9 SUT VM)</p> <p>テストは、89 の仮想ユーザーを実行するために構成され</p> <p>各 SUT VM の立ち上がり時間は 30 分間、テスト時間は 300 分を選択しました</p> <p>テストは、9 つの SUT VM すべてで並行して起動されました</p>	<p>SQL 2017 VM 仮想ディスクは NVMe テスト環境の DC1500M vSAN データストアからプロビジョニングされました。</p> <p>倉庫の数が 1200 のデータベーススキーマは、100GB のデータベース SUT 上に作成されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。</p> <p>テスト対象 SUT は 11 回クローンされ、物理サーバーあたり 2 つの VM が割り当てられ (合計 8 SUT VM)、HDB ワークロードが実行されました。</p> <p>テストは、89 の仮想ユーザーを実行するために構成され</p> <p>各 SUT VM の立ち上がり時間は 30 分間、テスト時間は 300 分を選択しました。</p> <p>物理サーバーあたり 1 つの VM の TPCC スキーマサイズは倉庫数 1200 (100GB) で、バックアップスクリプトは 100 秒ごとに起動され (合計 4 VM)、一方でワークロードは他の SUT VM で 10 サイクル実行中でした。</p> <p>8 つの SUT VM が HDB ワークロードを実行、4 VM がバックアップスクリプトを実行。</p> <p>テストは、12 の VM すべてで並行して起動されました</p>	<p>SATA テスト環境の Micron 5200 eco vSAN データストアからプロビジョニングした SQL 2017 VM 仮想ディスク。</p> <p>倉庫の数が 1200 のデータベーススキーマは、100GB のデータベース SUT 上に作成されていることを示します。テスト対象システムの VM (SUT) には、16 の vCore と 32GB RAM が割り当てられました。</p> <p>テスト対象 SUT は 8 回クローンされ、物理サーバーあたり 2 つの VM が割り当てられ (合計 6 SUT VM)、HDB ワークロードが実行されました。</p> <p>テストは、89 の仮想ユーザーを実行するために構成され</p> <p>各 SUT VM の立ち上がり時間は 30 分間、テスト時間は 300 分を選択しました。</p> <p>物理サーバーあたり 1 つの VM の TPCC スキーマサイズは倉庫数 1200 (100GB) で、バックアップスクリプトは 100 秒ごとに起動され (合計 4 VM)、一方でワークロードは SUT VM で実行中でした。</p> <p>6 つの SUT VM が HDB ワークロードを実行、3 VM がバックアップスクリプトを実行。</p> <p>テストは、9 つの VM すべてで並行して起動されました</p>

図 6.1 テスト 5 の説明: Micron 5200 eco SATA および DC1500M NVMe SSD vSAN データストアの SQL Server 2017 の現実的なノイジーネイバーテスト

このテストの目標は、本番ワークロード（この実験では、TPCC ベンチマークが本番ワークロードとして機能）を実行する SQL サーバー VM と同じ vSAN データストアを共有する、VM 上の望ましくないワークロード（このケースでは、TPCC データベースのバックアップ操作を使用）の現実的なシナリオをシミュレーションし、perfmon と vSAN パフォーマンスモニタから収集した TPCC ベンチマークの結果の評価と主要なストレージメトリクスの分析によって、全体的なパフォーマンスの影響を評価することでした。

テスト 5a) および 5b) では、すべての VM で TPCC ベンチマークを並行して実行することによって、ベースラインを確立しました。バックアップ操作は発生しませんでした。物理サーバーあたり 3 つの SQL VM を使用して、NVMe と SATA vSAN の両方のクラスタで実行しました。合計で NVMe は 12 の SUT VM、SATA は 9 つの SUT VM となりました。このテストのスキーマサイズは倉庫数 1200 で、TPC-C データベースサイズは 100GB 以下となり、ユーザー数 89 で 300 分間、立ち上がり時間 30 分で TPCC ワークロードを実行しました。

テスト 5c) と 5d) で、すべての SUT VM に TPC-C データベースをリストアしました。その後、スクリプトを起動し、NVMe クラスタでは 4 つの VM、SATA クラスタでは 3 つの VM で、TPC-C データベースのバックアップサイクルを 10 回実行し、同時に、残りの SUT VM で同じ TPC-C ベンチマークを実行しました。つまり、NVMe vSAN クラスタ上では、8 つの VM が TPC-C ワークロードを実行中で、並行して 4 つの VM でバックアップワークロードを実行中でした。その間、SATA vSAN クラスタ上では、6 つの VM が TPC-C ワークロードを実行中で、並行して 3 つの VM で TPC-C データベースのバックアップワークロードを実行中でした。

テスト 5 結果:SQL Server 2017 のパフォーマンス比較、ノイジーネイバーテスト、DC1500M NVMe と Micron 5200 eco SATA vSAN の比較

図 6.2 および 6.3 は、テスト 5a と 5b で達成した 1 分あたりのトランザクション (TPM) と 1 分あたりの新規発注 (NOPM) を示しています。DC1500M NVMe vSAN データストアを用いた SQL サーバー VM 12 個のそれぞれで 89 の仮想ユーザーを実行した場合、VM あたりの平均 TPM は 523,516、平均 NOPM は 113,812 となり、これと比較して Micron 5200 eco SATA クラスタを用いた SQL VM 9 個では、VM あたりの平均 TPM は 269,320、平均 NOPM は 58,544 となりました。vSAN パフォーマンスモニタから収集した IOPS とレイテンシの測定値（下図 6.4 と 6.5）を見ると、結果として、NVMe クラスタでのブロック層の上の IO は、読み取り IOPS が 120,000、書き込み IOPS が 60,000、読み取り/書き込み操作のレイテンシは 800 μ s になり、SATA vSAN クラスタでは 50,000 (読み取り)/20,000 (書き込み)、読み取りレイテンシの平均は 3.8ms、書き込みレイテンシの平均は 5.5ms になりました。これはまた、NVMe と SATA のパフォーマンスの差を明確に示し、DC1500M NVMe を用いた仮想ディスクが、並行した要求の影響を緩和して、はるかに速いラウンドトリップレイテンシで処理できることを示しています。

図 6.5 および 6.6 は、テスト 5c と 5d で達成した 1 分あたりのトランザクション (TPM) と 1 分あたりの新規発注 (NOPM) を示しています。DC1500M NVMe vSAN データストアを用いた SQL サーバー VM 8 個でそれぞれ 89 の仮想ユーザーを実行し、その間に VM のバックアップを並行して 4 個の VM で起動した場合、平均 TPM は 575,933、平均 NOPM は 125,206 となり、これと比較して Micron 5200 eco SATA を用いた SATA vSAN SQL VM で、SQL VM 6 個で TPCC ワークロードを実行し、その間に VM バックアップを並行して 3 個の VM で起動した場合、平均 TPM は 351,258、平均 NOPM は 76,355 となりました。さらに当社では、SATA と NVMe vSAN クラスタの両方でレイテンシとストレージの測定値を分析する必要がありましたが、両方のクラスタでバックアップが完了する速さも調べました。

図 6.8 と 6.9 は、テスト 5c と 5d で vSAN パフォーマンスモニタを使用して、NVMe と SATA クラスタから収集した vSAN IOPS とレイテンシの測定値を示しています。バックアップスクリプトは、100 秒ごとに 10 サイクル実行するように構成されました。VM バックアップが起動すると、NVMe と SATA vSAN クラスタの両方で、IOPS と読み取りおよび書き込みレイテンシに影響が見られます。しかし、レイテンシへの影響が異なります。NVMe クラスタの最大読み取り/書き込み IO レイテンシは 4ms/IO に急上昇しますが、読み取り/書き込み操作では平均 2.5 ms/IO を維持します。一方、SATA vSAN では、9ms/IO に急上昇し、読み取りでは平均 7.3 ms/IO、書き込み IO では平均 4.9 ms/IO を維持します。これは、エンドユーザが発注や、ショッピングカートの更新、他の倉庫の製品の表示などをしようとするときに感じるレイテンシです。

図 6.11 は、DC1500M vSAN を用いた SQL Server VM のひとつと、Micron 5200 eco vSAN を用いた SQL VM での、バックアップサイクルの完了までの所要時間を示します。バックアップサイクルの待ち時間は除外しています。NVMe vSAN を用いた SQL サーバー VM で、10 回のバックアップの完了にかかった時間は 73 分、バックアップ 1 回あたり平均 7 分で、SATA SSD vSAN を用いた SQL サーバー VM で、10 回のバックアップの完了にかかった時間は 122.15 分で、バックアップ 1 回あたり平均 12 分でした。DC1500M vSAN VM を用いた VM は、Micron 5200 eco vSAN を用いた VM と比べて、バックアップサイクルの速さが 1.67 倍でした。これは、VMware インフラを DC1500M NVMe を用いたデータストアにアップグレードすると、データベースのバックアップなどの望まれない操作を短時間で完了でき、レイテンシとスループットの性能にも優れているため、NVMe は Tier 1 アプリケーション上のこれらの迷惑なワークロードのレイテンシへの影響を緩和でき、ノイジーネイバーの課題の軽減に役立つという経験的な証拠です。

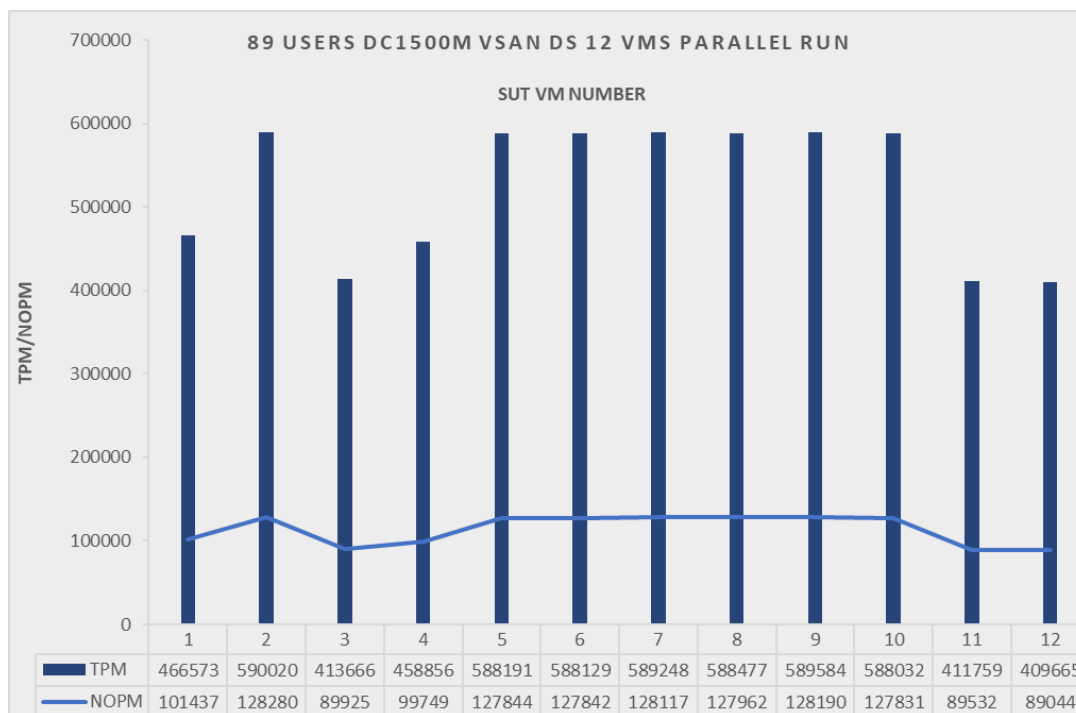


図 6.2 テスト 5a TPM SQL Server 2017 で 300 分間 12 個の VM を並行して実行、仮想ユーザー数 89、DC1500M NVMe SSD vSAN データストア

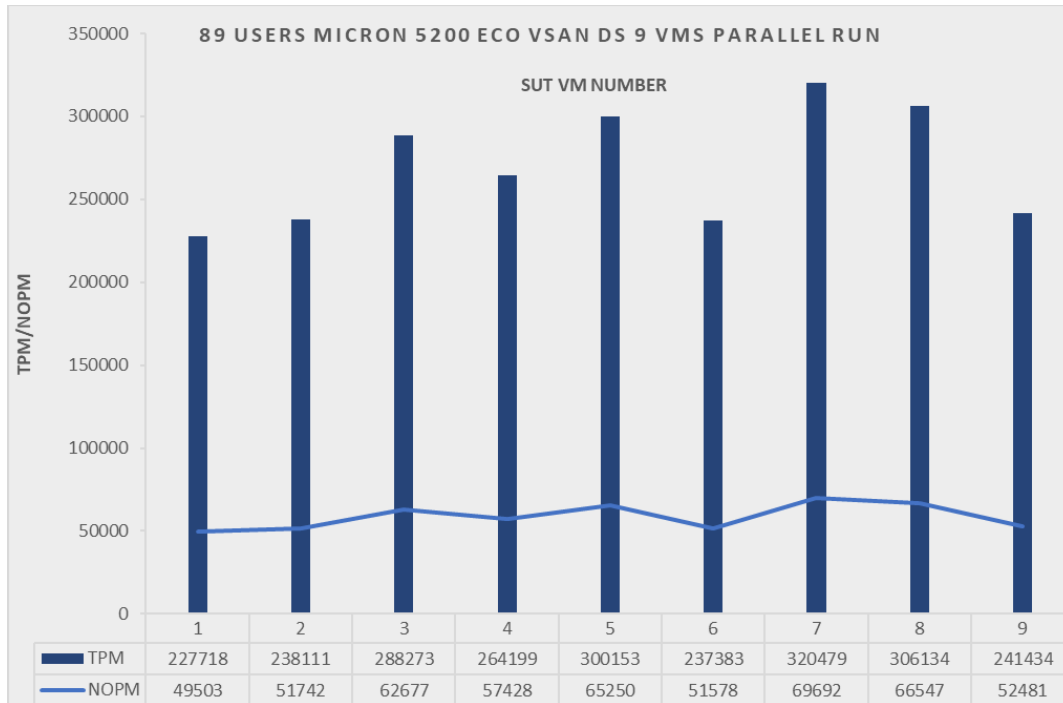


図6.3 テスト 5b TPM SQL Server 2017 で300 分間12 個の VM を並行して実行、仮想ユーザー数89、DC1500M NVMe SSD vSAN データストア

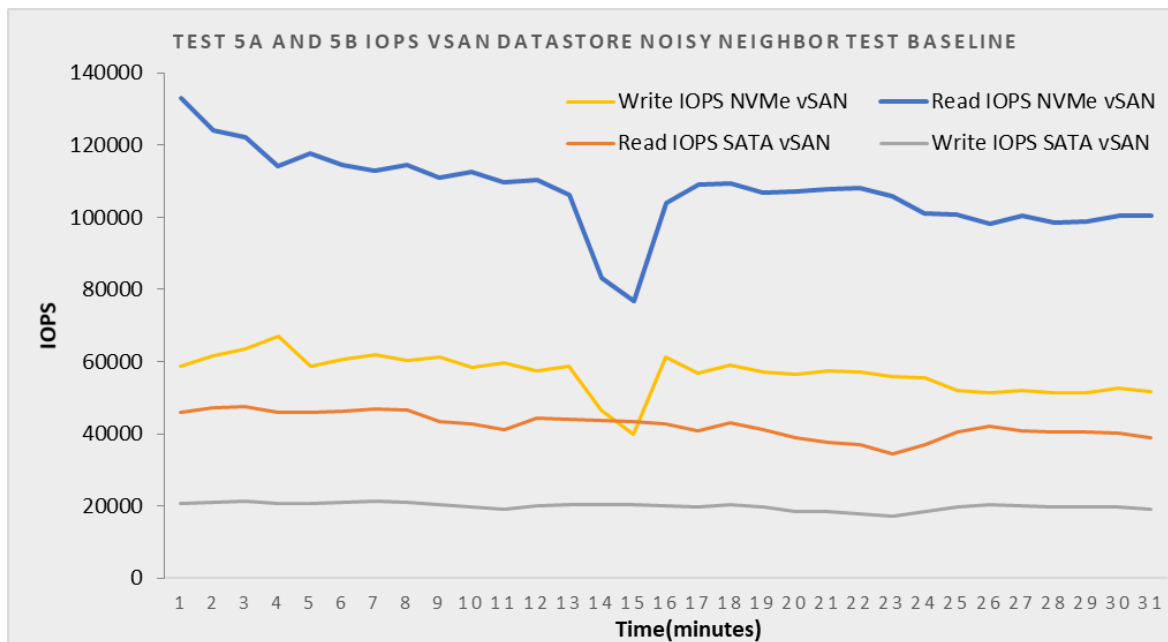


図6.4 テスト 5a および5b ノイジーネイバーIOPS、DC1500M NVMe およびMicron 5200 eco vSAN データストア

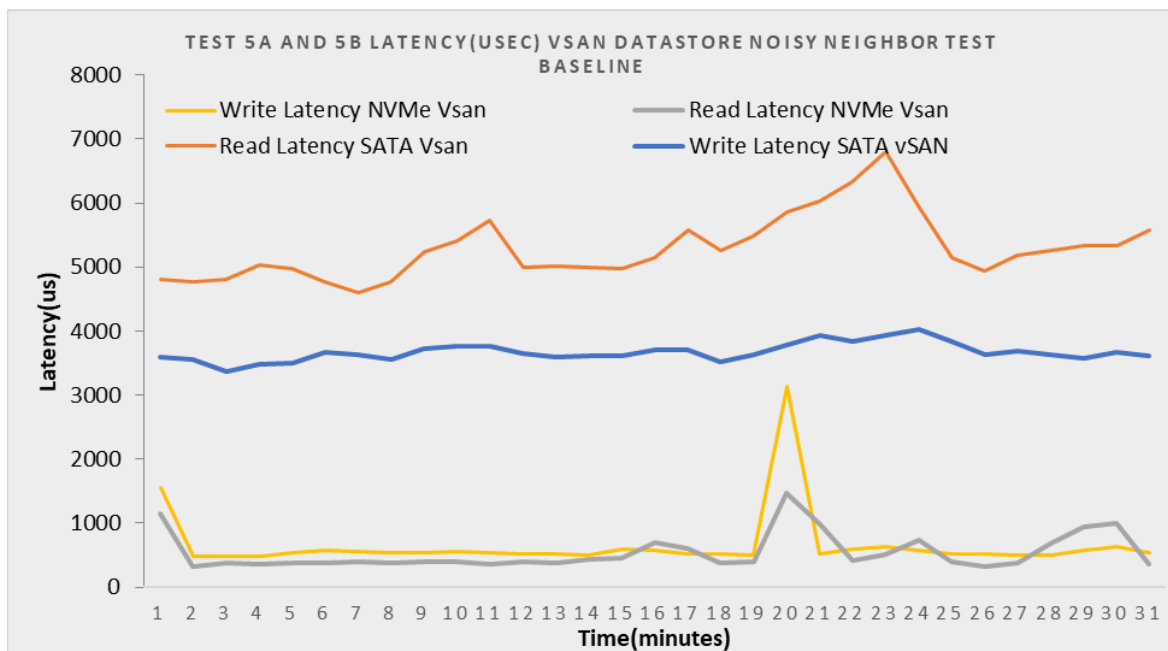


図6.5 テスト5a および5b ノイジーネイバー レイテンシ、DC1500M NVMe およびMicron 5200 eco vSAN データストア

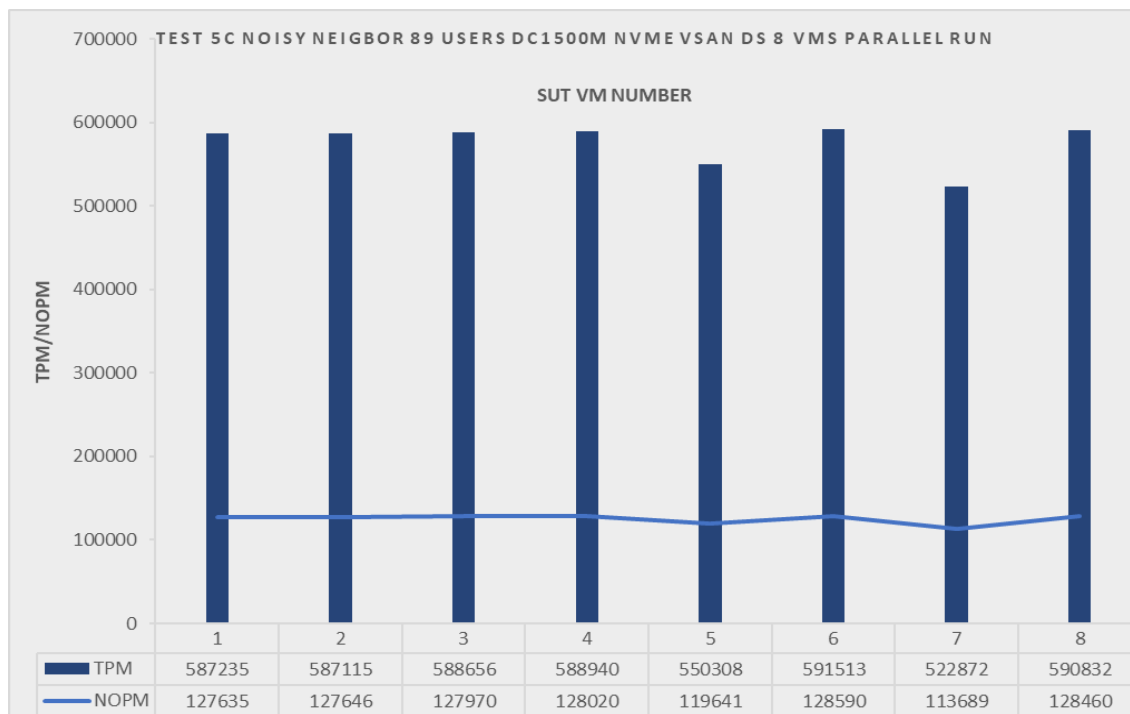


図6.6 テスト5c TPM、ノイジーネイバー実装8つのVMを並行して実行DC1500M NVMe vSAN データストア

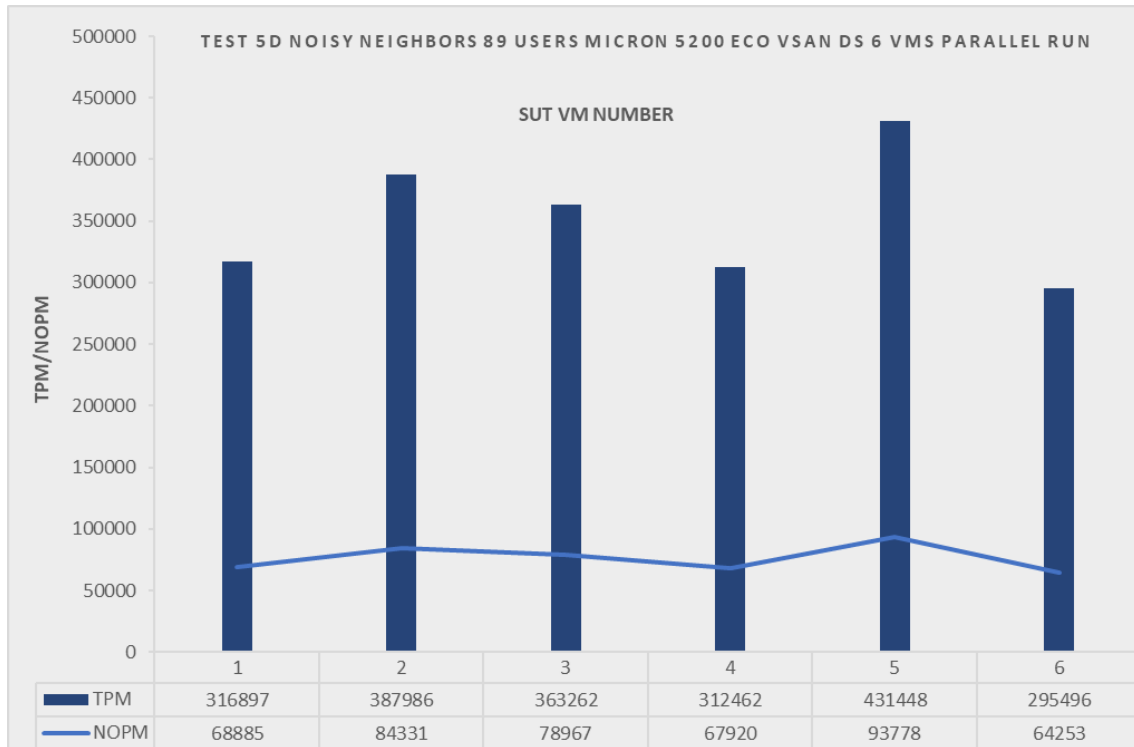


図6.7 テスト5D TPM、ノイジーネイバー実装6つのVMを並行して実行Micron 5200 eco vSAN データストア

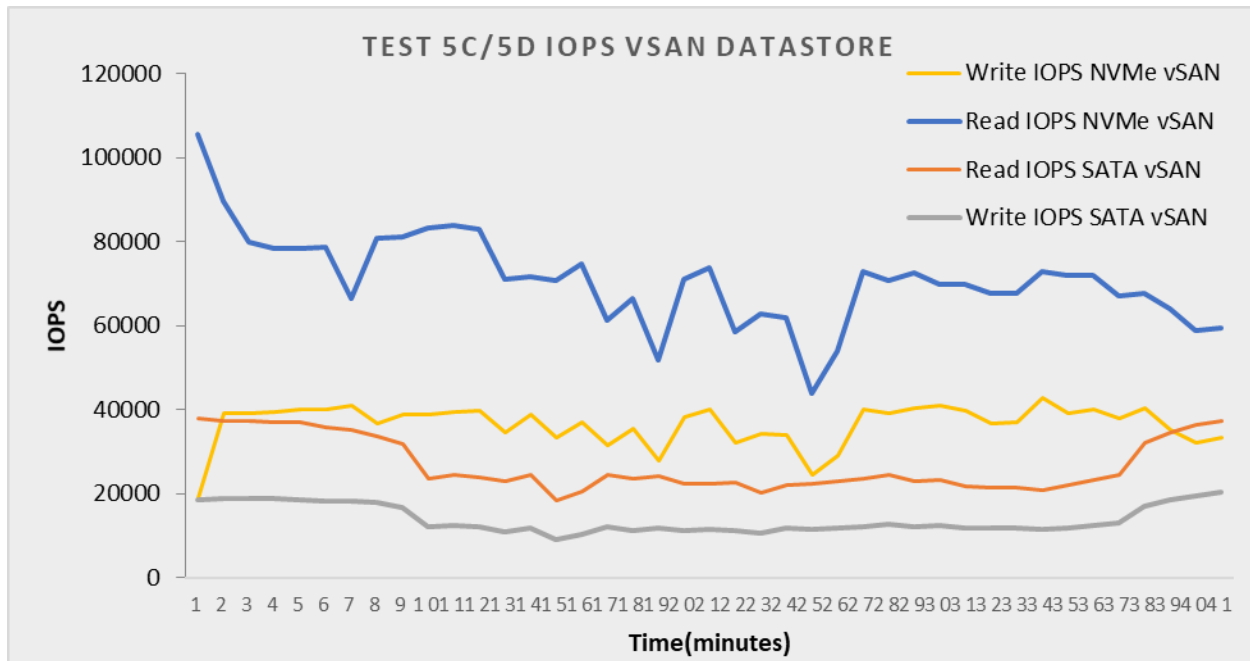


図6.8 テスト5C/5D IOPS、ノイジーネイバー実装NVMe vs SATA SSD vSAN データストア

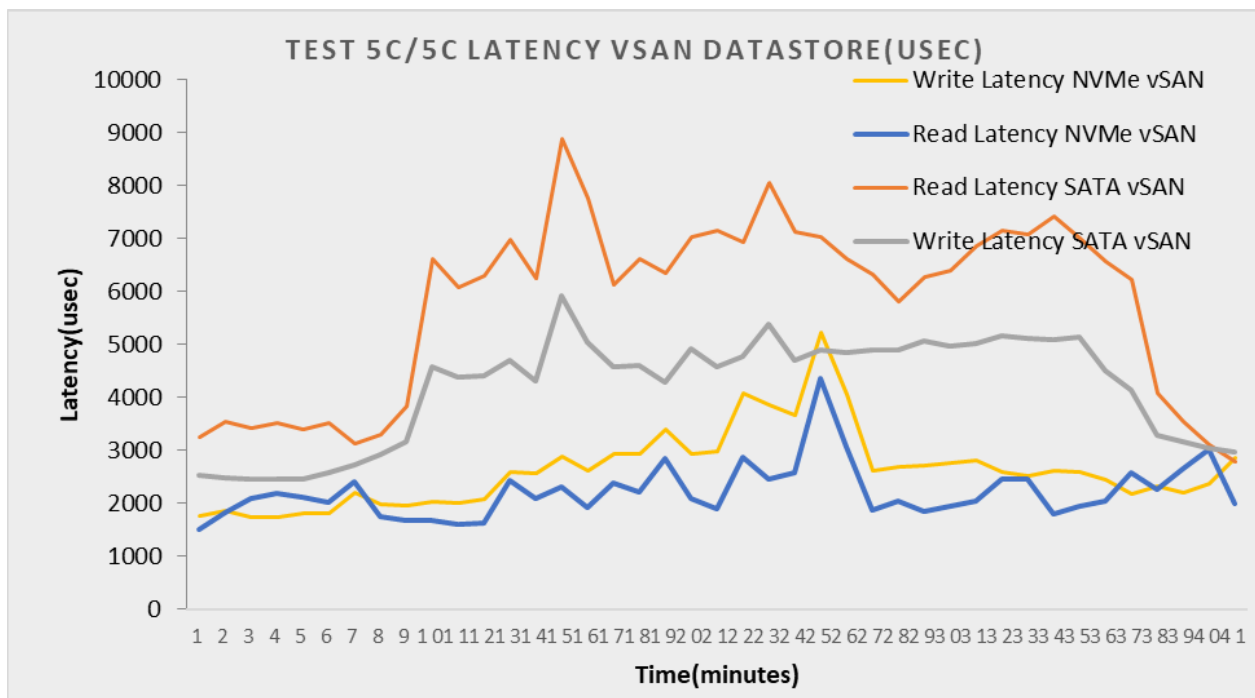


図6.9 テスト5C/5D レイテンシ、ノイジーネイバー実装 NVMe vs SATA SSD vSAN データストア

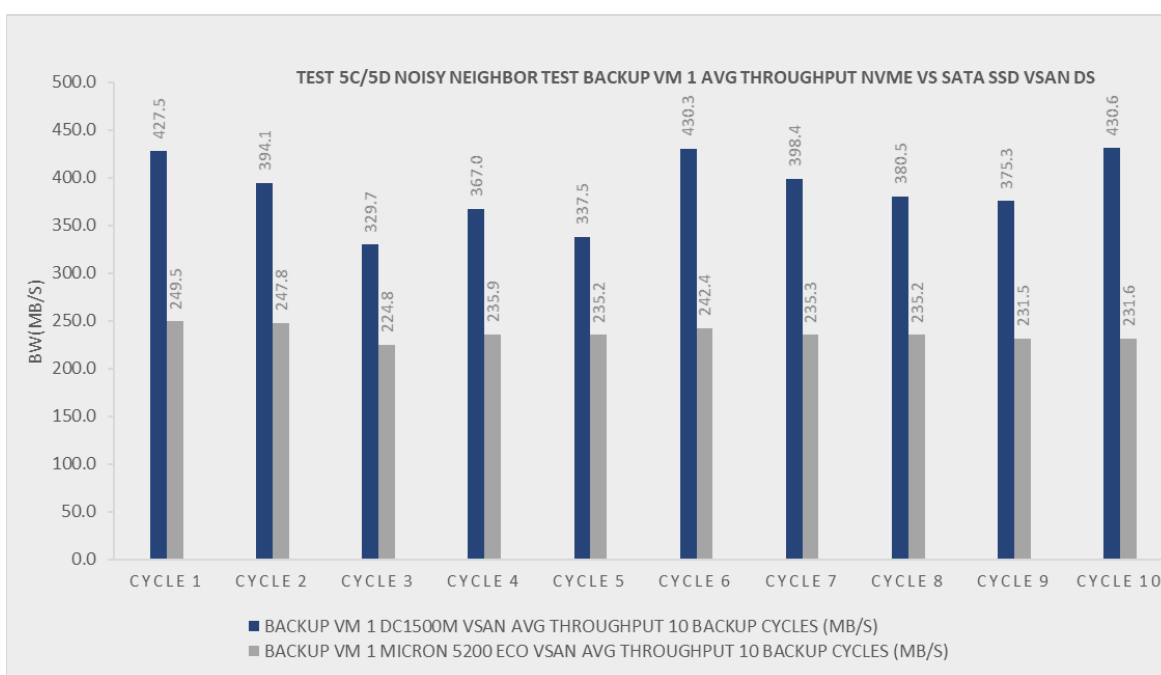


図6.10 テスト5C/5D バックアップVM スループット、ノイジーネイバー実装 NVMe vs SATA SSD vSAN データストア

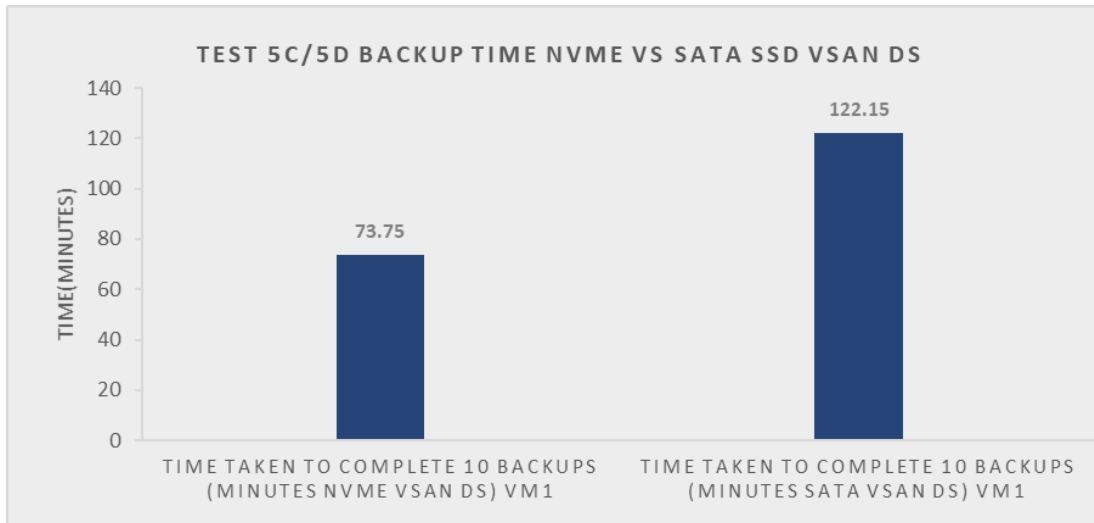


図6.11 テスト5C/5D バックアップVM バックアップ完了までの所要時間、10 サイクルノイズネイバー実装 NVMe vs SATA SSD vSAN データストア

結論

このホワイトペーパーでは、データベースワークロードを NVMe に統合すると、大幅に効率が向上して IO 待ち時間がほとんど 0 になり、使用する CPU コアを減らしても同じトランザクションスループットを達成できるため、既存のハードウェアを最大活用できることをご説明しました。エンタープライズ SATA SSD といくつかの比較を行い、SQL ワークロードを、NVMe を用いたデータストアに移行することで、トランザクションスループットを倍増し、レイテンシをサブミリ秒単位に抑えられるため、アプリケーションを拡張できることを示しました。その後、NVMe が、データベースのバックアップ/リストア操作などの望ましくないワークロードを迅速に完了することにより、Tier 1 アプリケーションへの影響の軽減に役立つことを示しました。

Kingston のエンタープライズ NVMe SSD、[DC1500M](#) を Kingston サーバーメモリ (Server Premier) を組み合わせれば、データベースインフラを仮想化してワークロード効率を最大化したいユーザーに、卓越したソリューションを提供します。

Kingston のデータセンターソリューションについて詳しくは、<https://www.kingston.com/en/solutions/servers-data-centers> をご覧ください

参照

HammerDB。 (n.d.)。 *TPCC ワークロードの概要*。
<https://www.hammerdb.com/docs3.3/ch03s05.html> から取得

TPCC ホーム。 (n.d.)。 <https://www.tpc.org/> から取得