
PowerEdge R720 を活用した SQL Server ストレージ性能の最適化

最高のデータベース性能は、最適なストレージソリューションの選択から

ソリューション性能分析グループ

Luis Acosta

アドバンスドストレージエンジニアリング

Joe Noyola



**本書は、情報提供のみを目的に執筆されており、誤字脱字や技術上の誤りには責任を負いません。
本書の内容は執筆時現在のものであり、明示的、暗示的を問わず、いかなる内容も保証いたしません。**

© 2012 Dell Inc. ©2012 デル株式会社 All rights reserved. (著作権所有)

デルとその関連会社は、誤字、脱字、誤植や、図、写真の誤りや不備について一切の責任を負いません。Dell、DELL のロゴマーク、PowerEdge は、米国 Dell Inc. の商標です。Intel、インテル、Xeon は、アメリカ合衆国およびその他の国におけるインテルコーポレーションおよび子会社の登録商標または商標です。Microsoft、Windows、Windows Server、SQL Server は、米国やその他の国々における Microsoft Corporation の登録商標または商標です。本書では、マークや名前を届け出た実在のもの、もしくは、その製品のいずれかを参照するため、その他の商標、商号を使用している可能性があります。デルは、その他のマークや名称について、商標上の利権に対する要求に一切に応じません。

2012 年 2 月 | Rev 1.0

目次

要旨	4
はじめに.....	4
テスト構成	5
ベンチマーク方法	7
テスト結果	7
データベース性能: TPS (1 秒あたりのトランザクション処理件数)	7
ストレージ性能: IOPS (1 秒あたりの I/O 処理件数)	9
データベース性能: レイテンシ	10
まとめ	10
付録: テスト構成	11

表

表 1. テストに使用したストレージ構成	7
----------------------------	---

図

図 1. PowerEdge R720xd の正面図	6
図 2. PowerEdge R720xd の背面図	6
図 3. PowerEdge R720 の構成.....	6
図 4. データベース性能比較 (TPS)	9
図 5. ストレージ性能 (IOPS).....	9
図 6. データベース性能: レイテンシ (ms)	10

要旨

新しい Dell™ PowerEdge™ R720 サーバは、高性能・高スケーラビリティを重視した設計となっており、中～大規模データベースのアプリケーションサーバに理想的です。デルでは今回、本サーバのリリースに合わせて、2つの新しいストレージソリューションを追加しました。そのうち1つは、「PowerEdge Express Flash PCIe SSD (ソリッドステートドライブ)」デバイス、そしてもう1つは、「Dell PERC H710P」です。この PERC H710P は、デュアルコアの RAID コントローラで、6 Gb/s の「Enterprise SAS SSD」をサポートします。このように PowerEdge R720 は、より高速な Enterprise SAS SSD と Dell PowerEdge Express Flash PCIe SSD を提供することで、これまで限界に来ていた OLTP データベースのランダム I/O 性能を飛躍的に高めることに成功しました。デルは、PERC H710P コントローラと各種のドライブ、および、Express Flash PCIe-SSD の性能を評価するため、Microsoft® SQL Server® の中規模 OLTP データベース上で、TPC-E タイプの I/O ワークロードを発生させるシミュレーションテストを実施しました。本書は、そのテスト内容と結果についてご報告致します。

はじめに

今回、ストレージソリューションに PERC H710P RAID コントローラ、Enterprise 6 Gb/s SAS SSD、Express Flash (PCIe SSD) が加わったことで、お客様の選択肢がさらに広がっており、データベースアプリケーション要件に合わせてストレージ性能をきめ細かく調整することができます。従来型のハードディスクドライブ (HDD) は、低コスト+大容量という強みを兼ね備えているものの、回転式の機械装置のため、ランダム I/O 性能には限界があります。オンライントランザクション処理 (OLTP) のデータアクセスパターンは 100% ランダム I/O となるため、一台の HDD の処理量には限度があり、すぐに I/O 量が飽和してレイテンシの大幅な増大を招いてしまいます。このランダム I/O が引き起こす深刻なレイテンシボトルネックの回避策は、できるだけ多くのドライブを使用して、データを分散させることです。しかし、この方法でレイテンシは改善できても、ドライブ数が増えればその分、使い切れずに残る容量も増えてしまいます。

PERC H710P コントローラに SAS SSD 接続する場合は、二種類の構成法をサポートし、いずれも HDD より高いランダム I/O 性能を発揮します。その1つ目は、SAS SSD を CacheCade™ として構成する方法です。同ソフトウェアがデュアルコア PERC H710P 上で稼働し、HDD とデータベースアプリケーション間の読み取り専用キャッシュとして補完レイヤを形成します。頻繁に読まれるデータが特定されると、この SSD キャッシュに入れられ、以降、同じデータの読み取り要求には、HDD からではなくキャッシュから提供されるため、アクセス時間が短縮でき、ランダム I/O 性能が向上します。

SAS SSD のもう1つの構成法は、プライマリデータストレージとして活用する方法です。この場合、読み取りだけでなく書き込み性能も飛躍的に向上しますし、高性能を得るのにキャッシュのような前段階を踏む必要もありません。また、ストレージ全体の容量がフル活用でき、従来のハードドライブに見られたようなランダム I/O 性能の低下も発生しません。

一方、Dell PowerEdge Express Flash PCIe SSD は、I/O 機能の提供に SAS インタフェースを使わず、PCIe バスを使うため、NAND 型ソリッドステートテクノロジーの低レイテンシ能力が存分に引き出せます。この新しいストレージは、PCIe SSD の特徴である高性能に加え、極めて高い柔軟性も兼ね備えています。たとえば、4 台の 2.5 インチドライブはホットスワップ可能なうえ、R720 シャーシの前面から保守できるため、サーバシャーシ内に収められていた従来の PCIe フラッシュストレージよりも管理性が向上しています。

主なテスト結果

下記に、主なテスト結果を抜粋します (詳細は後述)。

- **OLTP データベース性能：1 秒あたりのトランザクション処理件数 (TPS)**
 - 1 台の SSD を CacheCade ボリュームとして使用した場合、ハードドライブのみの構成に比べ、データベース性能が 3 倍向上
 - 4 台の SAS SSD は、HDD のみのストレージより、データベース性能が 7.5 倍向上
 - 1 台の Express Flash ドライブは、HDD ストレージより、10.5 倍のデータベース TPS を達成¹
- **ストレージ性能：1 秒あたりの I/O 処理件数 (IOPS)**
 - CacheCade を使用すると、HDD のみのストレージより IOPS が 3 倍向上
 - SAS SSD ストレージは、HDD ストレージより 8 倍の IOPS を達成
 - 16 台の HDD と比較した場合、1 台の Express Flash ドライブは、10.5 倍の IOPS を達成
- **平均応答時間：ミリ秒 (ms)**
 - ハードドライブの応答時間 = 227ms
 - 1 台の Enterprise 6 Gb/s SAS SSD ドライブを CacheCade ボリュームとして稼働させたところ、応答時間が 108ms に短縮
 - 4 台の 6 Gb/s SAS SSD ドライブを使用した場合、応答時間がさらに向上し 52ms を達成
 - 1 台の Express Flash ドライブの応答時間は、最速で 26ms を達成。つまりレイテンシは 4 台の SAS SSD ドライブに比べ 1/4、16 台の従来型 HDD に比べ 1/10 まで短縮可能¹

テスト構成

PowerEdge R720 サーバは各種のストレージ構成をサポートしており、2U のシャーシ内に従来型の回転式記録媒体 (HDD) のほか、CacheCade、Enterprise 6 Gb/s SAS SSD、Express Flash PCIe SSD も搭載できます。ハードドライブのみで構成するテストでは、データベース性能を最大限に引き出すため、最高 26 台の 2.5 インチドライブをサポートする PowerEdge R720xd を使用しました。

図 1 と図 2 は、ドライブと CacheCade を搭載し、OLTP 系 SQL Server 2008 R2 データベース性能を測定したときに使用した R720xd 構成です。

¹ これは、Dell PowerEdge R720 サーバに Dell PowerEdge Express Flash を搭載し、SQL データベースアプリケーションを実行したときの結果です。Express Flash は、16 ドライブ構成の R720 でのみサポートします。

図 1. PowerEdge R720xd の正面図

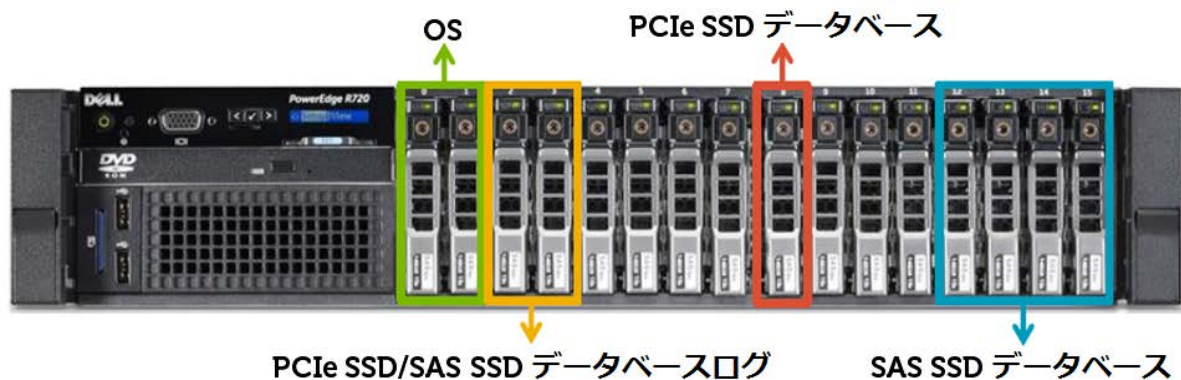


図 2. PowerEdge R720xd の背面図



Enterprise 6 Gb/s SAS SSD (SAS SSD) と Express Flash PCIe SSD の性能測定では、16 台の 2.5 インチドライブで構成した PowerEdge R720 を使用しました。図 3 に、このテスト構成を示します。

図 3. PowerEdge R720 の構成



R720 は、お客様のニーズに応じて、16 台の 2.5 インチドライブ、8 台の 3.5 インチドライブ、12 台の 3.5 インチドライブ用バックプレーン構成もご注文可能です。Express Flash は、16 ドライブ構成でのみサポートします。

表 1. テストに使用したストレージ構成

データベースストレージ構成	RAID 構成	ドライブの種類/スピード
ハードドライブ (基準性能)	16 ドライブで RAID 10 を構成	SAS HDD (15,000 回転 6 Gb)
Enterprise 6 Gb SAS SSD (SAS SSD)	4 ドライブで RAID 10 を構成	SAS SSD (6 Gb)
ハードドライブ + 1 SSD (CacheCade ボリュームとして構成)	16 ドライブの RAID 10 + 1 SAS SSD	SAS HDD (15,000 回転 6 Gb) と 1 台の SAS SSD (6 Gb)
Express Flash (PCIe SSD)	1 台の Express Flash ドライブ (RAID なし)	Express Flash

ベンチマーク方法

今回、データベース性能の測定に使用したのは、データベースビルド&ワークロードシミュレーションツールの「Quest® Benchmark Factory® (BMF) for Databases」です。また、テスト対象の各ストレージソリューション (ハードドライブ、SSD、Express Flash PCIe SSD) には、業界標準の TPC-E ベンチマークに即した 100GB の SQL Server 2008 データベースをそれぞれ構築しました。各テストのストレージ構成は、図 1~3 と表 1 を参照してください。

本テストでは、4 種類のストレージ、つまり「ハードドライブのみ」「ハードドライブ + 1 台の SSD (CacheCade ボリューム用)」「4 台の Enterprise SAS SSD」「1 台の Express Flash ドライブ」にデータベースを構築して、それぞれの性能を測定しました。各性能テスト中は、BMF が TPC-E タイプの SQL トランザクション処理を行うユーザをシミュレーションするので、各データベースのストレージサブシステムにランダムな読み書き I/O 処理が発生します。BMF は、データベースユーザのシミュレーションを実行しながら、各タイプのストレージデバイスが達成した合計 IOPS 数、TPS 数、クエリーの平均応答時間をそれぞれ収集します。

今回、BMF のシミュレーションによってかけられた最大負荷は、200 仮想ユーザです。すべての I/O 要求が物理ストレージ上で確実に処理されるようにするため、SQL Server 2008 のデフォルトインスタンスの実行は、1GB RAM のみに制限しています。このように、SQL Server で利用できるメモリ量を最小限に抑えることで、ストレージサブシステムの I/O 能力を最大限に活用することができます。

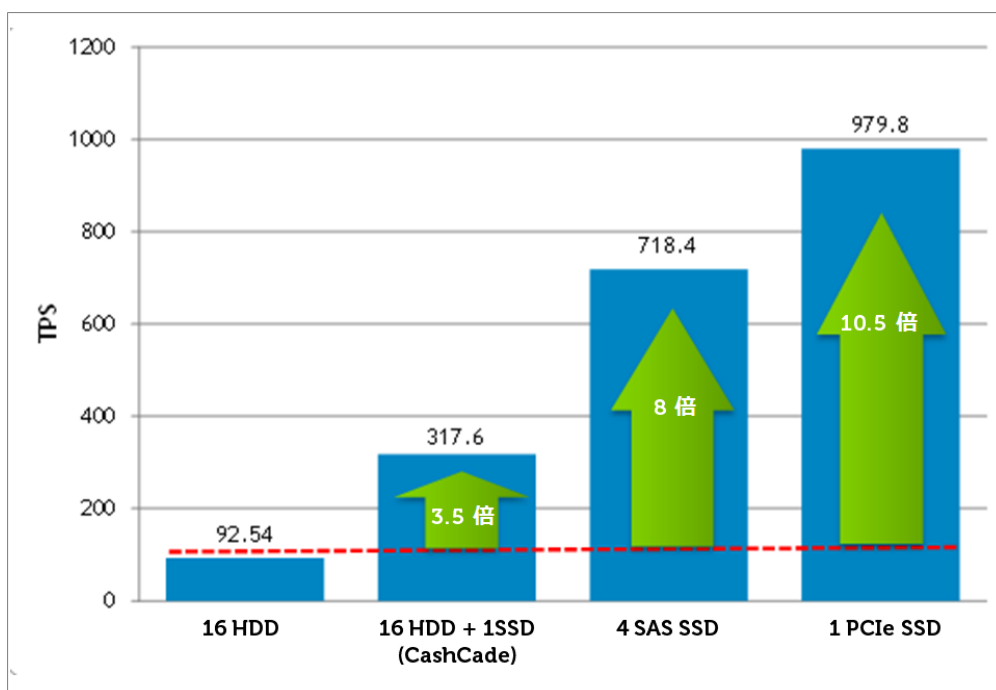
テスト結果

このセクションでは、先に挙げたテスト結果の詳細を述べます。

データベース性能: TPS (1 秒あたりのトランザクション処理件数)

今回、基準としたハードドライブデータベースの性能測定値は、92 TPS でした。次に、1 台の Enterprise SAS SSD を CacheCade ボリュームとして構成したところ、317 TPS を上回るデータベース性能が達成されました。これは、約 3.5 倍の性能向上となります。さらに、すべて Enterprise SAS SSD で構成したデータベースの性能は 718 TPS に達し、約 8 倍の向上となりました。最後に 1 台の Express Flash ドライブを使ったデータベースでは、979 TPS 前後の性能を達成し、HDD のみ使ったストレージ上のデータベースと比較して、約 10.5 倍の性能向上となりました。図 4 は、HDD、CacheCade、Enterprise SAS SSD、Express Flash で達成されたデータベース性能の比較グラフです。

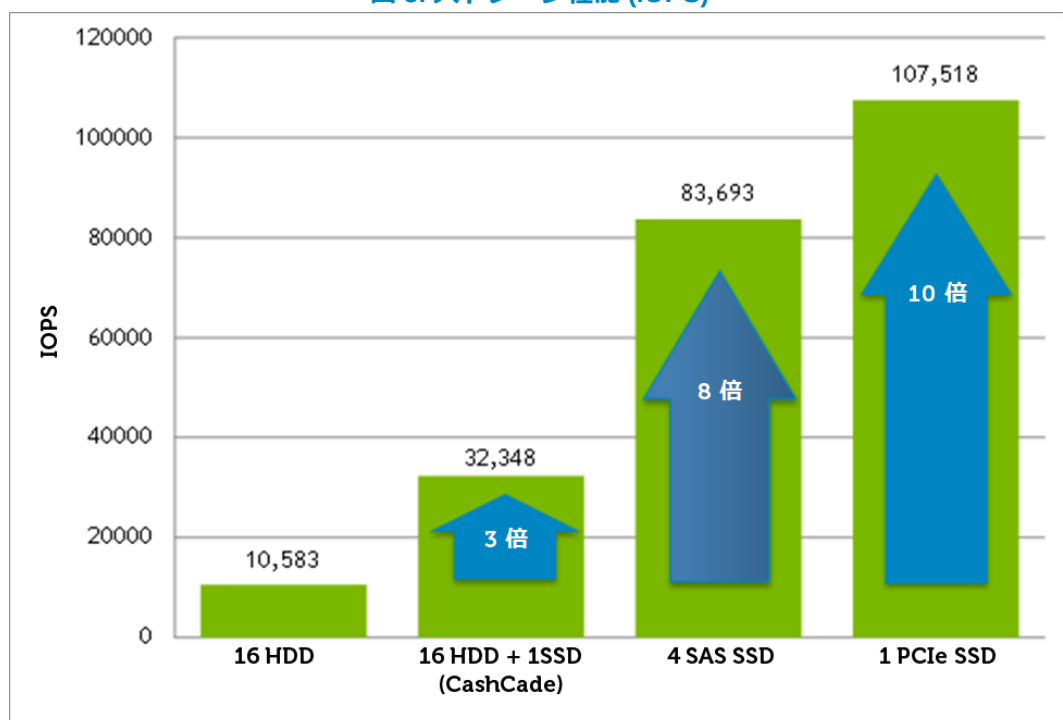
図 4. データベース性能比較 (TPS)



ストレージ性能: IOPS (1秒あたりのI/O処理件数)

200 ユーザがデータベースクエリーを同時実行したとき、16 台のハードドライブ構成では平均 IOPS = 10,583 を記録しましたが、これに 1 台の CacheCade ボリュームを加えると、スループットが約 3 倍向上し、32,348 IOPS になりました。次に、Enterprise SAS SSD ストレージを使用すると、平均 IOPS = 83,693 となり約 8 倍向上しました。さらに、1 台の Express Flash ドライブでは、平均 IOPS = 107,518 を達成し、HDD から約 10 倍の向上となりました。図 5 に、この結果をまとめます。

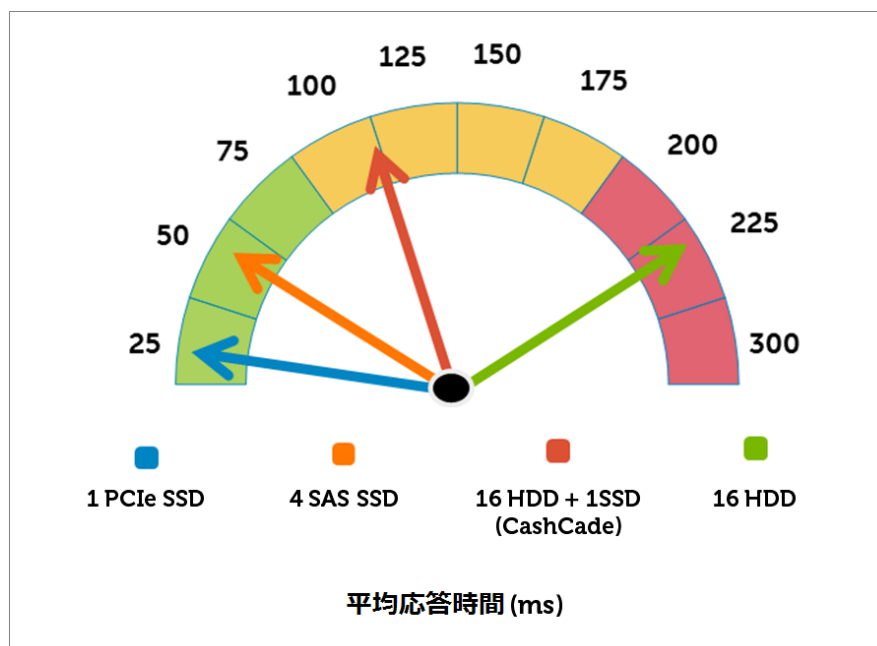
図 5. ストレージ性能 (IOPS)



データベース性能：レイテンシ

200 人の仮想ユーザが負荷をかけたとき、ハードドライブ構成の平均応答時間は、222ms でした。その後、単に CacheCade を有効にするだけで、平均応答時間が 108ms にまで短縮されました。これで、レイテンシ全般が 55% 削減されたことになります。次に Enterprise 6 Gb SAS SSD ストレージの平均応答時間を測定したところ、CacheCade のレイテンシをさらに凌ぎ、52ms となりました。Express Flash では 26ms を達成し、他のどのストレージソリューションよりもレイテンシが低減できています。図 6 に、本テスト仕様で記録されたレイテンシを示します。

図 6. データベース性能: レイテンシ (ms) - 低いほど優れる



まとめ

従来型のハードドライブは、依然、「データベースアプリケーション用ストレージの最も低コストなオプション」と考えられています。確かに、電気機械にまつわる制限を考えれば、ハードドライブのランダム I/O 性能は許容できるものです。しかし、フラッシュメモリに基づく SSD (ソリッドステートドライブ) テクノロジーなら、HDD の回転によるseekタイムの遅れを排除することができますし、データの取得後も即座に応答できるため、従来の HDD で発生していたランダム I/O のボトルネックを劇的に改善できます。1 台の Enterprise 6 GB SAS SSD を CacheCade ボリュームとして使用すると、ハードドライブの 3 倍もの性能が達成されたことから、多くのストレージを追加しなくても高いデータベース性能が得られるより良いソリューションとなります。

さらに SSD だけでデータベースを構築すると、従来の回転式記録媒体に比べて性能が飛躍的に伸ばせるため、ストレージ台数を抑えながら、データベース性能を大幅強化するのに極めて優れた選択となります。しかし、最高の性能と最小限のレイテンシを求めるデータベースであれば、今回テストした全ストレージソリューション中、最もハイパフォーマンスを達成した、1 台の Express Flash PCIe SSD ドライブ構成が向いています。

付録：テスト構成

以下に、本テストで使用したシステムの構成内容をまとめます。

付録 A：PowerEdge R720 の構成内容

サーバハードウェア	Dell PowerEdge R720
プロセッサ	デュアルインテル® Xeon® プロセッサ E5-2680、2.70GHz、20MB キャッシュ、8 コア
システム内の合計 RAM (GB)	64
メモリの種類	PC3-12800R
スピード (MT/s)	1600
RAM モジュール数	8
ランク構成	デュアルランク
ベンダとモデル番号	Dell M393B1K70DH0
システム内のディスク数	4
容量 (GB)	146
回転速度 (RPM)	15,000
種類	SAS 6 Gb
RAID の種類	RAID 1
コントローラ	PERC 内蔵 H710P、ファームウェア 21.0.1-0130、ドライバ 5.1.90.64
OS	Microsoft® Windows Server® 2008 R2 Enterprise
ファイルシステム	NTFS
NIC モデル	Broadcom® BCM5720 NetXtreme® II
種類	オンボード

付録 B：PowerEdge R720xd の構成内容

サーバハードウェア	Dell PowerEdge R720xd
プロセッサ	デュアルインテル® Xeon® プロセッサ E5-2680、2.70GHz、20MB キャッシュ、8 コア
システム内の合計 RAM (GB)	64
メモリの種類	PC3-12800R
スピード (MT/s)	1600
RAM モジュール数	8
ランク構成	デュアルランク
ベンダとモデル番号	Dell M393B1K70DH0
システム内のディスク数	20
容量 (GB)	146
回転速度 (RPM)	15,000
種類	SAS 6 Gb
RAID の種類	RAID 1 x 2
コントローラ	PERC 内蔵 H710P、ファームウェア 21.0.1-0130、ドライバ 5.1.90.64

OS	Microsoft Windows Server 2008 R2 Enterprise
ファイルシステム	NTFS
NIC モデル	Broadcom BCM5720 NetXtreme II
種類	オンボード

付録 C : ストレージの構成内容

ストレージ構成 : すべて HDD (基準)	Dell PowerEdge R720xd
利用可能な合計ドライブスロット数	26
インストールしたドライブ数	20
データベース用 : 16 ドライブの RAID 10 ボリューム	データベースログ用 : 2 ドライブの RAID 1 ボリューム
コントローラ	PERC 内蔵 H710P、デュアルコア、1GB RAM
PERC のキャッシュ設定	Write back (ライトバック)
	Adaptive read ahead (適応先読み)

ストレージ構成 : CacheCade	Dell PowerEdge R720xd
利用可能な合計ドライブスロット数	26
インストールしたドライブ数	20 + 1 Enterprise 6 Gb SAS SSD (CacheCade 用)
データベース用 : 16 ドライブの RAID 10 ボリューム	データベースログ用 : 2 ドライブの RAID 1 ボリューム
コントローラ	PERC 内蔵 H710P、デュアルコア、1 GB RAM
PERC のキャッシュ設定	Write back (ライトバック)
	Adaptive read ahead (適応先読み)

ストレージ構成 : メインストリーム パフォーマンス SAS SSD	Dell PowerEdge R720
利用可能な合計ドライブスロット数	16
インストールしたドライブ数	4 x 2.5 インチ 15,000 回転 SAS ドライブ + 4 x Enterprise 6 Gb SAS SSD ドライブ
データベース用 : 4 ドライブの RAID 10 ボリューム	データベースログ用 : 2 ドライブの RAID 1 ボリューム
コントローラ	PERC 内蔵 H710P、デュアルコア、1GB RAM
PERC のキャッシュ設定	Write back (ライトバック)
	Adaptive read ahead (適応先読み)

ストレージ構成 : Express Flash	Dell PowerEdge R720
利用可能な合計ドライブスロット数	4
インストールしたドライブ数	1 x Express Flash ドライブ
データベース用 : 1 ドライブ (RAID なし)	データベースログ用 : 2 ドライブの RAID 1 ボリューム
コントローラ	適用外

付録 D：データベースの構成内容

データベースソフトウェア	Microsoft SQL Server 2008 R2 Enterprise
メモリ構成	1GB
データベースコンテナサイズ	150GB
ログサイズ	50GB
データベースインスタンスのメモリ	1GB に限定 (すべての I/O をディスクで処理させるため)

付録 E：ベンチマークの構成内容

Quest Benchmark Factory for Databases	
バージョン	6.1.1
業界標準ベンチマーク	TPC-E
スケールファクタ	12
データベース全体のサイズ	126GB
仮想ユーザ数	無制限
データ収集	有効に設定。さらに BMF は、Windows パフォーマンスモニタから論理ディスクの測定値も収集