

Oracle Exadata X8M の性能を検証し PMEM キャッシュで 2 倍の効果を確認 最新の Oracle Exadata X9M の可能性も探る

Oracle Database の構築や運用を手掛ける株式会社アシストの中村嘉宏氏と笠原克俊氏をお招きし、Oracle Exadata X8M の実機ベンチマークで明らかになったインテル® Optane™ パーシステント・メモリーの性能向上効果と、第3世代インテル® Xeon® スケーラブル・プロセッサ・ファミリーを搭載した最新の Exadata X9M の可能性についてお話を伺いました。



株式会社アシスト

●住所：〒102-8109 東京都千代田区九段北 4-2-1 市ヶ谷スクエアビル ●設立：1972年 ●事業内容：ただ単にソフトウェアを提供するのではなく、ソフトウェアで課題を解決する「パッケージ・インテグレーター」として活動

<https://www.ashisuto.co.jp/exadata/>



株式会社アシスト
ビジネスインフラ技術本部
データベース技術統括部
技術4部 課長

中村 嘉宏 氏



株式会社アシスト
ビジネスインフラ技術本部
データベース技術統括部
技術4部 主任

笠原 克俊 氏



インテル株式会社
インテル Optane グループ・
セールス
テクニカル・セールス・
スペシャリスト

松村 浩



インテル株式会社
インテル Optane グループ・
セールス
テクニカル・セールス・
スペシャリスト

石橋 史康



インテル株式会社
インダストリー事業本部
ソリューション技術統括部
シニア・データセンター・
アーキテクト

小佐原 大輔

Oracle Exadata X8M の実機を用いて PMEM キャッシュの効果をベンチマークを検証

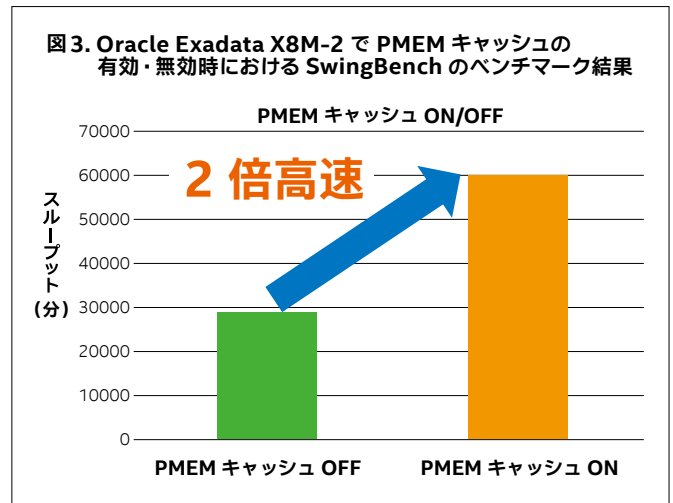
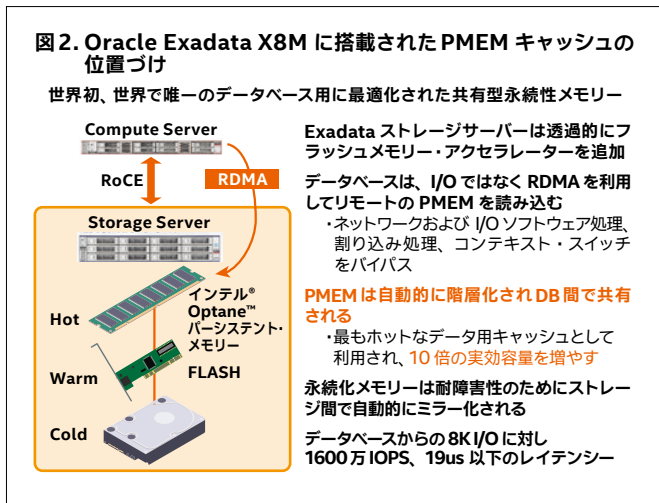
石橋 (インテル): 企業が扱うデータ量は増大の一途を辿っていて、データベース・サーバーにも一層の大容量化と高性能化が求められています。そうした中、アシスト様は 2021年1月に、Oracle のデータベース・プラットフォームである Oracle Exadata (エクサデータ) シリーズの性能を検証した「Exadata はどこまで速くなったか 最新 Exadata X8M を徹底検証!」という技術情報¹⁾を自社のブログサイト「テックノート」で公開しました。Exadata X8M²⁾に搭載されているインテル® Optane™ パーシステント・メモリー³⁾の性能向上効果が実際のワークロードで検証されていて、たいへん興味深い内容となっています。

中村 (アシスト): Oracle Exadata シリーズは、Oracle Database に最適化された検証済み (Engineered Systems) のアプライアンス・システムで、バッチ処理、OLTP、データウェア・ハウスなど、ミッション・クリティカルを含む幅広いデータベース・アプリケーションで活用されています。当社ではハードウェアの提供のほか構築や運用を手掛けています。

笠原 (アシスト): 2019年に発売されたモデルが Exadata X8M です (図1)。最大の特徴は、インテル® Optane™ パーシステント・メモリーを使った「PMEMキャッシュ」が Oracle Exadata シリーズとしては初めて搭載されたことです (図2)。それまでの最上位モデルだった Exadata X8 に対してアーキテクチャーが変わったこともあって、更改や新規導入を検討されているお客様にリアルな性能データを示したいと考え、Exadata X8M の実機を使ってベンチマーク・テストを行いました。

図1. Oracle Exadata X8M-2 の概要

| | | |
|--|--|---|
| | スケールアウト型 2 ソケット DB サーバー | 2 ソケット インテル® Xeon® プロセッサ 48 コア (サーバーあたり) 384GB-1.5TB DRAM |
| | 100Gb RDMA over Converged Ethernet (RoCE) ネットワーク・ファブリック | 100Gb/s RoCE ネットワーク・ファブリック 25/10GigE 外部ネットワーク |
| | スケールアウト型でインテリジェントな 2 ソケット・ストレージサーバー | ハイ キャパシティー 168TB ディスク容量 25.6TB PCI NVMe Flash 1.5TB 永続性メモリー 32 コア (SQL オフロード用) |
| | エクストリーム・フラッシュ | 51.2TB PCI NVMe Flash 1.5TB 永続性メモリー 32 コア (SQL オフロード用) |



松村 (インテル): Exadata X8M で PMEM キャッシュとして採用されているインテル® Optane™ パーシステント・メモリーは、DRAM で構成される主メモリーと、フラッシュメモリーなどで構成されるストレージとの性能ギャップを埋めるために開発されたテクノロジーです。フォームファクターは DDR4 DRAM DIMM と同じで、第 2 世代以降のインテル® Xeon® スケーラブル・プロセッサー・ファミリーの DRAM チャンネルに実装して使用します。パーシステント (永続的) の名前のとおり、電源を切ってもメモリーの内容が保持される不揮発性も特徴です。

笠原: Exadata X8M を構成するデータベース・サーバーとストレージサーバーのうち、PMEM キャッシュはストレージサーバー側に搭載されていて、お話のあったように DRAM と NVMe フラッシュの間に配置され、主にホットデータのキャッシュとして使われています。また、ストレージサーバーの PMEM キャッシュ内のデータを、双方の OS を介さずに、データベース・サーバーに直接転送する RDMA (リモート DMA) 機能が、RoCE (RDMA over Converged Ethernet) プロトコルとして実装されています。

**PMEM キャッシュでおおよそ 2 倍の性能向上を確認
 RDMA や Smart Scan も性能に寄与**

石橋: 「Exadata はどこまで速くなったか 最新 Exadata X8M を徹底検証!」というブログ記事にまとめられているベンチマークの結果について、ポイントをあらためて説明していただけませんか?

笠原: ベンチマークには、2 台のデータベース・サーバーと 3 台のストレージサーバーで構成された、Oracle Exadata X8M-2 (2 ソケット)、1/8 ラック、EF タイプを使用しました。PMEM キャッシュの容量はストレージサーバー 1 台あたり 1.5TB で、全体では 4.5TB です。

Oracle Exadata X8M の最大の特徴である PMEM キャッシュの性能向上効果を明らかにしたいと考え、PMEM キャッシュを有効 (デフォルト) にしたときと無効にしたときとで性能比較を行いました。テストにはデータベース・ベンチマークとして一般的な SwingBench⁴ を用い、オンライン系システムを模擬するために、顧客登録、製品表示、製品発注、注文処理、注文状況照会という 5 種類のトランザクションを複数セッションでランダムに発生させています。

石橋: ブログ記事によると、PMEM キャッシュの有効時と無効時とで、おおよそ 2 倍の性能差があったそうですね。

笠原: そのとおりです。PMEM キャッシュを無効にして従来モデルの Oracle Exadata X8 と同等構成にしたときの性能はおおよそ 29,000TPM (transaction per minute) でしたが、PMEM キャッシュ有効時には 2 倍強に相当するおおよそ 60,000TPM が得られました (図3)。詳しくは弊社テックノートの記事をお読みいただくとして、ユーザー I/O のトータル待ち時間の 40% 短縮や、OS 使用率の 10% 削減などの効果も明らかになりました。

先ほど説明したように、Oracle Exadata X8M では PMEM キャッシュ上のデータをサーバー間で直接やり取りする RDMA 機能がサポートされています。当社では直接は確認していませんが、Oracle の公表値によると、SQL のレイテンシーは Oracle Exadata X8 以前の NVMe フラッシュベースでは 250μs だったのが PMEM キャッシュと RDMA の組み合わせによって 19μs 以下と 1/10 以下に短縮されているそうです。また、クエリーの際にストレージサーバーからデータベース・サーバーに返すデータ量をおおよそ 1/10 に削減する Oracle Exadata Smart Scan 機能なども組み込まれています。こうした工夫がトランザクション性能の大幅向上に寄与しているのではないかと考えています。

松村: インテル® Optane™ パーシステント・メモリーを使えばサーバーの性能は一般には上がるのですが、実際のワークロードを模したベンチマークで 2 倍もの性能向上が得られることはなかなかありません。インテル® Optane™ パーシステント・メモリーから最高の性能を引き出そうと、Oracle が RDMA 機能などを含めさまざまなチューニングを行って低レイテンシーを実現した結果が、アシスト様のベンチマークに表れているのだと感じました。

石橋: インテル® Optane™ パーシステント・メモリーには電源オフ時にもデータが保持されるという特徴があるのですが、その点についてはいかがですか?

笠原: 仮に障害などで Oracle Exadata X8M の電源が遮断されたとしても、Exadata は Oracle Database と統合された開発とテストを通して RDBMS としてのデータ整合が担保された上で、通常時

には高い性能を得られることになりしますので、メリットの1つとしてお客様に訴求できていると思っています。

第3世代インテル® Xeon® スケーラブル・プロセッサファミリーを搭載した Oracle Exadata X9M が新たに登場

石橋：Oracle は 2021 年 9 月に、Oracle Exadata シリーズの最上位モデルとなる Oracle Exadata X9M を発表しました (図 4)。Ice Lake[†] と呼ばれる最新の第 3 世代インテル® Xeon® スケーラブル・プロセッサ・ファミリー⁵ が搭載されているということで、インテルとしても関心を寄せています。

笠原：Oracle Exadata X9M の特徴や変更点については「最新情報！ Exadata X9M からの変更点をピックアップ紹介」という弊社ブログサイト「アシスト現場ブログ」の記事⁶でもご紹介していますが、8KB 読み取りの I/O 性能は Exadata X8M の最大 1,600 万 IOPS に対して Oracle Exadata X9M では 1.7 倍に相当する最大 2,760 万 IOPS と謳われています。PMEM キャッシュの容量に変更はありませんが、プロセッサの変更や、内部の PCIe バスが Gen3 から Gen4 に上がって I/O 帯域が 2 倍になったことが、性能向上の要因ではないかと考えています。

小佐原 (インテル)：Oracle が公表している両機種のスペックを見ると、データベース・サーバーのプロセッサは、Oracle Exadata X8M は第 2 世代 (開発コード名: Cascade Lake) のインテル® Xeon® Platinum 8260 プロセッサ (2.40GHz / 24 コア) ですが、Oracle Exadata X9M では第 3 世代 (開発コード名: Ice Lake) のインテル® Xeon® 8358 プロセッサ (2.60GHz / 32 コア) に変更され、コア数が 24 から 32 に増強されています。

一方のストレージサーバーは、Oracle Exadata X8M は第 2 世代のインテル® Xeon® Gold 5218 プロセッサ (2.30GHz / 16 コア) ですが、Oracle Exadata X9M では第 3 世代のインテル® Xeon® 8352Y プロセッサ (2.20GHz / 32 コア) に変更され、インテル® スピード・セレクト・テクノロジー (インテル® SST) を用いて 32 コアのうち 16 コアのみを使って 2.60GHz で動作させています。Oracle はそのほうが全体性能が高くなると判断したのかもしれませんが。

石橋：メモリーのチャンネル数も、第 2 世代インテル® Xeon® スケーラブル・プロセッサ・ファミリーの 6 から、第 3 世代インテル® Xeon® スケーラブル・プロセッサ・ファミリーでは 8 に増え、なおかつメモリーバスの周波数も 3200MHz まで上がっています。これに対応したのが最新のインテル® Optane™ パーシステント・メモリー 200 シリーズで、Oracle Exadata X9M の PMEM キャッシュにもこちらが採用されていると思われます。

また、Oracle Exadata Systems Software のバージョン 21.2.0 のリリースノートには、暗号化テーブルや圧縮テーブルを対象にした Smart Scan 処理の性能を上げるために、第 3 世代インテル® Xeon® スケーラブル・プロセッサ・ファミリーで高速化されたインテル® アドバンスド・ベクトル・エクステンション 512 (インテル® AVX-512) 命令を活用していると書かれていました。Oracle Exadata X9M では今まで以上のチューニングが行われていると推測されます。

中村：これまでプロセッサ周りの細かい違いに着目することはあまりなかったのですが、今のお話はとても参考になりました。Oracle Exadata X9M の特徴の1つとしてお客様にお伝えしていきたいと思えます。

高度化するユーザーニーズに応える 高性能でスケーラブルな Oracle Exadata プラットフォーム

石橋：Exadata X8M や Exadata X9M は、高性能なデータベース・システムを必要とする顧客に、どのようなメリットをもたらすとお考えでしょうか？

中村：Engineered Systems である Oracle Exadata X8M や最新の Oracle Exadata X9M は、オンプレミス以外にも、パブリック・クラウドである Oracle Exadata Cloud Service (ExaCS)⁷ や、本体を顧客のデータセンターに設置して Oracle のクラウド・エキスパートが運用する Oracle Exadata Cloud@Customer としても導入可能です。また、1/8 ラックモデルからフルラックモデルまでが用意されるなどスケーラビリティにも優れていますので、極めて自由度の高いシステムといえます。バッチ処理や OLTP (オンライン・トランザクション処理) に加え、近年注目を集めている企業データの OLAP (オン

図 4. Oracle Exadata X9M の概要 (赤字は X8M からの変更)

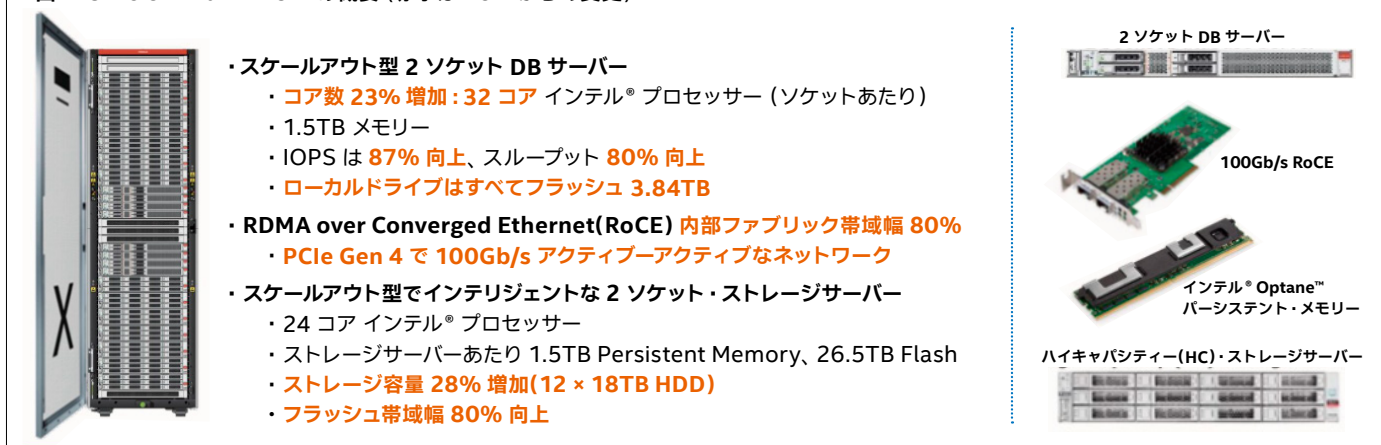


図 4. Oracle Exadata X9M の概要 (赤字は X8M からの変更)

- ・スケールアウト型 2 ソケット DB サーバー
 - ・ コア数 **23% 増加: 32 コア** インテル® プロセッサ (ソケットあたり)
 - ・ 1.5TB メモリー
 - ・ IOPS は **87% 向上**、スループット **80% 向上**
 - ・ ローカルドライブはすべて**フラッシュ 3.84TB**
- ・ RDMA over Converged Ethernet (RoCE) 内部ファブリック帯域幅 **80%**
 - ・ **PCIe Gen 4** で **100Gb/s** アクティブ-アクティブなネットワーク
- ・スケールアウト型でインテリジェントな 2 ソケット・ストレージサーバー
 - ・ 24 コア インテル® プロセッサ
 - ・ ストレージサーバーあたり 1.5TB Persistent Memory、26.5TB Flash
 - ・ **ストレージ容量 28% 増加 (12 × 18TB HDD)**
 - ・ **フラッシュ帯域幅 80% 向上**

2 ソケット DB サーバー

100Gb/s RoCE

インテル® Optane™ パーシステント・メモリー

ハイキャパシティー (HC) ストレージサーバー

ライン分析処理)にも最適で、高度化するお客様のニーズを具現化できるデータベース・プラットフォームであると考えています。

アシストは35年以上にわたって Oracle Database の構築や運用に携わってきました。クラウド環境への移行支援の実績も豊富です。培ってきた経験や知識を強みに、Oracle Exadata シリーズを活用しながら、お客様のデータ基盤の安定化や最適化に貢献できるようこれからも努めていきます。

石橋：アシスト様のベンチマークを通じてインテル® Optane™ パーシステント・メモリーが Oracle Exadata X8M の性能向上に大きく貢献していることが明らかになりましたし、第3世代インテル® Xeon® スケーラブル・プロセッサ・ファミリーが搭載された Oracle Exadata X9M では、さらに高い性能が期待できることが分かりました。インテルとしてはアシスト様のテックノートや本ホワイトペーパーを通じて、インテル® Optane™ パーシステント・メモリーおよび第3世代インテル® Xeon® スケーラブル・プロセッサ・ファミリーのメリットを広く紹介していきたいと考えています。本日はありがとうございました。

↑ 開発コード名

*1 https://www.ashisuto.co.jp/tech-note/article/20210113_oracle.html

*2 <https://www.oracle.com/jp/engineered-systems/exadata/>

*3 <https://www.intel.co.jp/content/www/jp/ja/architecture-and-technology/optane-dc-persistent-memory.html>

*4 <https://github.com/domgiles/swingbench-public/releases/tag/production>

*5 <https://www.intel.co.jp/content/www/jp/ja/products/platforms/details/ice-lake-sp.html>

*6 https://www.ashisuto.co.jp/db_blog/article/exadata-x9m-new-features.html

*7 Exadata Cloud Serviceは2022年2月時点でX8Mのみ提供



この文書は情報提供のみを目的としています。この文書は現状のまま提供され、いかなる保証もいたしません。ここにいう保証には、商品適格性、他者の権利の非侵害性、特定目的への適合性、また、あらゆる提案書、仕様書、見本から生じる保証を含みますが、これらに限定されるものではありません。インテルはこの仕様の情報の使用に関する財産権の侵害を含む、いかなる責任も負いません。また、明示されているか否かにかかわらず、また禁反言によるとよらずにかかわらず、いかなる知的財産権のライセンスも許諾するものではありません。

Intel、インテル、Intel ロゴ、その他のインテルの名称やロゴは、Intel Corporation またはその子会社の商標です。

その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

インテル株式会社

〒100-0005 東京都千代田区丸の内3-1-1

<http://www.intel.co.jp/>

©2022 Intel Corporation. 無断での引用、転載を禁じます。

2022年3月

350267-001JA
JPN/2203/PDF/CB/BCG/AY