

# 最新のメモリー・サブシステムをデータベース・コード、線形代数コード、ビッグデータ、エンタープライズ・ストレージに使用する利点

この記事は、インテル® デベロッパー・ゾーンに公開されている「[Modern Memory Subsystems Benefits for Data Base Codes, Linear Algebra Codes, Big Data, and Enterprise Storage](#)」の日本語参考訳です。

マルチチャンネル DRAM (MCDRAM (英語)) と高帯域幅メモリー (HBM) の利点は次のとおりです。

- オフパッケージのダイナミック・ランダムアクセス・メモリー (DRAM) よりも高い帯域幅をインテル® Xeon Phi™ 製品のコアに提供
- オフパッケージ・メモリーのキャッシュに使用可能

パフォーマンスについて、将来の 3D XPoint™ メモリーメディアは次のような利点をもたらします。

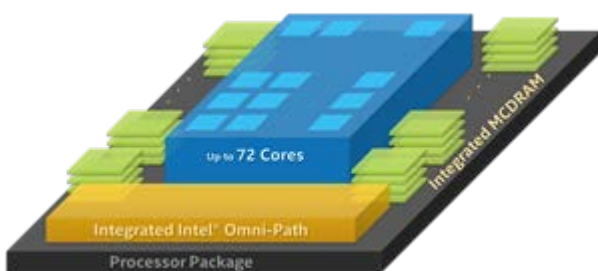
- 密度: テラバイトのメモリーをプロセッサの近くに配置
- 持続性: 電源障害が発生してもデータを保持

所有コストについて、将来の 3D XPoint™ テクノロジーは次のような利点をもたらします。

- より少ないプロセッサとマザーボードで必要なメモリーを供給
- ビットあたりのコストが低い
- ビットあたりのアイドル時の消費電力が低い

インテル® Omni-Path ファブリック (インテル® OP ファブリック) の利点は次のとおりです。

- ノード間のデータ移動に特別なコードが不要
- 既存のコードを実行可能、ノード全体にわたってアクセス数が少ない場合はより効率良く実行



以降のセクションでは、以下の分野における利点を述べます。

- [データベース・コード](#)
- [線形代数コード](#)
- [ビッグデータ \(Graph 500、ソーシャルメディア、Apache\\* Hadoop\\*、および生化学アプリケーション\)](#)
- [エンタープライズおよびクラウド・ビッグ・ストレージ](#)

## データベース・コード

データベース・システムは次の要件を満たしている必要があります。

- ユーザー要求から特定のアルゴリズムへの変換
- アルゴリズムの実行に必要なデータの取得
- アルゴリズムの実行
- データへの継続的な変更

最新のメモリー・サブシステムの優れた容量、パフォーマンス、持続性は、これらの処理を向上できます。

インテル® OP ファブリック、将来の 3D XPoint™ テクノロジーおよび同様の不揮発性デュアル・インラインメモリー・モジュール (NVDIMM (英語)) が永続データに与える影響に注目して説明します。

データベース・サーバーは、安定性を実現するため、永続ストレージにデータを複製します。通常、少なくとも 2 つの遠く離れたメモリー・サブシステムにデータを送ることで、災害時に両方のデータが破損するのを防ぎます。ノード数、互いの距離、データ保持に使用するストレージは、予測される脅威と安定性要件に依存します。例えば、電源供給源の異なる近くのマシンの DRAM に複製するだけで十分な場合もあれば、別の大陸への複製が必要になる場合もあります。

### 3D XPoint™ メモリーおよびストレージメディア

将来の 3D XPoint™ テクノロジーは、以下の理由から永続メディアへの複製を高速化できます。

- 最速の長距離インターコネクットの速度に匹敵する
- ローカルでは、3D XPoint™ テクノロジー・ベースのデュアル・インラインメモリー・モジュール (DIMM) の永続領域や永続ストレージ領域への書き込みのほうが、3D Xpoint™ テクノロジー・ベースのソリッド・ステート・ドライブ (SSD) への書き込みよりも高速であり、つまり、ほかの SSD やハード・ディスク・ドライブ (HDD) よりも高速である

データサーバー実行時には、より多くのデータをプロセッサの近くに保持できるため、大きなテーブルの読み取りで遅延が軽減されます。さらに、より大きなメモリーにより多くのスレッドがデータを保持できるため、データの待機によって一部のスレッドが遅延しても、より多くの保留中の作業が利用可能になります。

ノードの起動時には、リモートノードからフェッチする代わりに、3D XPoint™ メモリーからデータをチェックするだけで済みます。そのため、起動時間を大幅に短縮できます。

### インテル® OP ファブリック

インテル® OP ファブリックは、同じコードで部屋または大陸を超えて効率良くデータを分散することができます。最小限のプログラミングでリモートサイトからローカルキャッシュへのデータの高速度フェッチが可能です。同様に、インテル® OP ファブリックは、複数のリモートサイトへデータを簡単に素早く複製することもできます。

## 線形代数コード

線形代数コードは、1 台のマシンで数秒間で実行できるものから、大規模システムで数カ月かかるものまでさまざまです。

### MCDRAM と HBM

インテル® Xeon Phi™ プロセッサのメニーコアと MCDRAM または HBM オンパッケージ・メモリーの組み合わせることで、多くのコアによってアクセスされる大規模データ構造を含むプログラミングを簡単に、そして効率良く行うことができます。通常、ほとんどのメモリアクセスがコアごとのキャッシュ内で行われるようにコードをタイル化し、遠く離れた DIMM に格納されている大規模データ構造は MCDRAM や HBM にキャッシュします。そうすることで、特に将来の 3D XPoint™ メモリーによってローカル DIMM のストレージが劇的に増加する場合、1 つのノードでより大きな問題に取り組むことができます。

### インテル® OP ファブリック

1 基のインテル® Xeon® プロセッサまたはインテル® Xeon Phi™ プロセッサで対応できない多くの計算処理またはデータを必要とする問題では、一般に MPI を使用してコードとデータを分散します。

インテル® OP ファブリックは、スレッドが 1 ノードで実行する場合も、または多数のノードで実行する場合も、効率良く動作するコードを使用することで、データの共有を容易にします。

### 3D XPoint™ テクノロジーとインテル® OP ファブリック

数時間または数日間実行するアプリケーションでは、1 台以上のマシンを意図的に停止したり、クラッシュや電源障害によってマシンがダウンしても、計算を続行することが求められます。

将来の 3D XPoint™ メモリーを使用することで、SSD よりも高速なチェックポイントを提供できます。

インテル® OP ファブリックは、ほかのノードへデータを素早く複製できるため、ノードをシャットダウンして、別のノードで作業を再開することが可能です。

## ビッグデータ (Graph 500、ソーシャルメディア、Apache\* Hadoop\*、および生化学アプリケーション)

これらのフレームワークとアプリケーションは、通常、データベース・コードと線形代数コードの多くの側面の併せ持ち、データはやや不規則で、アルゴリズムはそこまで特化したものではありません。

### 3D XPoint™ メモリーおよびストレージメディア

将来の 3D Xpoint™ メモリーの予想される容量とビットあたりのコストは、これまでになく多くのインデックスをローカルで利用できるようにし、より遠くのデータへのアクセス時間を軽減します。より大きなメモリーにより多くのスレッドがデータを保持できるため、データの待機によって一部のスレッドが遅延しても、より多くの保留中の作業が利用可能になります。

3D XPoint™ テクノロジーは、以下の理由からも永続メディアへの複製を高速化できます。

- 最速の長距離インターコネクトの速度に匹敵する
- ローカルでは、3D XPoint™ テクノロジーベースの DIMM の永続領域や永続ストレージ領域への書き込みのほうが、3D Xpoint™ テクノロジーベースの SSD への書き込みよりも高速であり、つまり、ほかの SSD やハードディスクドライブ (HDD) よりも高速である

ノードの起動時には、リモートノードからフェッチする代わりに、3D XPoint™ メモリーのデータをチェックするだけで済みます。そのため、起動時間を大幅に短縮できます。

## インテル® OP ファブリック

インテル® OP ファブリックは、同じコードで部屋または大陸を超えて効率良くデータを分散することができます。最小限のプログラミングでリモートサイトからローカルキャッシュへのデータの高速フェッチが可能です。同様に、インテル® OP ファブリックは、複数のリモートサイトへデータを簡単に素早く複製できます。

## エンタープライズおよびクラウド・ビッグ・ストレージ

これらのフレームワークとアプリケーションは、通常、データベース・コードと線形代数コードの多くの側面の併せ持ち、1つのシステム内で広範なアプリケーションを扱います。

幸いにも、将来の 3D XPoint™ テクノロジーとインテル® OP ファブリックは、さまざまなアプリケーションにおいて、困難な最適化の問題に取り組むインテル® Xeon Phi™ プロセッサから、ビジネスに重要なデータを保持する高度にスケラブルで信頼性の高いストレージ・サブシステムまで、あらゆるものを一元化するのに役立ちます。

## まとめ

以前の記事、[NUMA ハードウェアの対象ユーザー](#) (英語) では、ほとんどのアプリケーションが新しいメモリー・サブシステム・ハードウェア・テクノロジーの利点を得られることを示しました。特定のアプリケーション・スタイルによる利点についても詳しく説明しました。次の記事、[メモリー・パフォーマンスの概要](#)では、新しいハードウェアのパフォーマンスを最大限に引き出すために必要な情報を提供します。

## 著者紹介

Bevin Brett. インテル コーポレーションの主任エンジニアで、プログラマーとシステムユーザーがアプリケーション・パフォーマンスを向上できるように支援するツールに取り組んでいます。アプリケーションが多くの時間を費やしている場所を見つけてパフォーマンスを向上したり、より良いアルゴリズムを見つけてそれによってもたらされるパフォーマンス・ゲインを評価したり、利用可能なハードウェアでそれらを実装する方法を見つけることを楽しんでいます。

コンパイラーの最適化に関する詳細は、[最適化に関する注意事項](#)を参照してください。