

ケーススタディー

インテル® ソフトウェア開発ツール

インテル® コンパイラー、インテル® VTune™ Amplifier、インテル® MPI ライブラリー (インテル® MPI)、インテル® マス・カーネル・ライブラリー (インテル® MKL)、インテル® Parallel Studio XE Cluster Edition、インテル® Xeon Phi™ コプロセッサ

ウォーカー分子動力学研究室の バイオメディカル・ソフトウェアの最適化

インテル® ソフトウェア開発ツールによるアプリケーションのパフォーマンスと生産性の向上

「インテル® MKL とインテル® MPI ライブラリーを実装したことにより、開発者は最小限の労力で大幅な速度向上を達成できました。エンドユーザーにとってもこの投資は非常に価値のあるものです。」

– カリフォルニア大学サンディエゴ校 (UCSD)
サンディエゴ・スーパーコンピューター・センター (SDSC)
ウォーカー分子動力学研究室
Ross Walker 准教授

はじめに

サンディエゴ・スーパーコンピューター・センター (SDSC) のウォーカー分子動力学研究室では、高度な古典分子力学 (MM) と複合量子/分子力学 (QM/MM) シミュレーションの研究を行っています。創薬、生体触媒、酵素工学、高度なバイオマテリアル開発を含む、生体分子系研究のシミュレーションを可能にする、ウォーカーの Amber* 分子動力学ソフトウェアは、学術界および産業界の多くの科学研究者が利用しています。ウォーカー研究室は、多くの研究開発結果を Amber* ソフトウェアに適用し、分子動力学シミュレーションの最先端の手法をユーザーに提供しています。

ウォーカー研究室の研究対象は広範で、計算化学、分子生物学、ハイパフォーマンス・コンピューティングの分野を含み、特に古典 MM と複合 QM/MM 手法の並列計算および高速計算用の効率的なアルゴリズムの開発に力を入れています。研究では、酵素の物理的・化学的性質の決定における MM および QM/MM 分子動力学アルゴリズムの使用にも注目しています。

Amber* が研究者に提供するもの:

1. 生体分子シミュレーション用分子動力学力場のセット。これらはパブリックドメインで、さまざまなシミュレーション・プログラムで使用されています。
2. ソースコードとデモを含む、分子シミュレーション・プログラムのパッケージ。

ソフトウェアと力場は、酵素、脂質膜、高度なバイオマテリアル、触媒のシミュレーションに使用されます。脂質結合タンパク質との相互作用に加えて、薬物結合の親和性を予測し、セルの薬物透過性を正確に判定する能力が提供されるため、創薬に効果的です。

「**Intel® VTune™ Amplifier** は、コードの最適化で hotspot の特定に役立つ非常に優れたツールです。ユーザー・インターフェイスは使いやすく、詳細な情報を基に開発を迅速に進めることができます。Intel® VTune™ Amplifier の行単位で表示されるパフォーマンス・カウンターがなければ、精度が混在したコードがオリジナルの倍精度コードより遅くなっていた理由を特定することはできなかつたでしょう。」

– カリフォルニア大学サンディエゴ校 (UCSD)
サンディエゴ・スーパーコンピューター・センター (SDSC)
ウォーカー分子動力学研究室
Perri Needham 博士 研究員

課題

ウォーカー研究室は、Amber* の高度に最適化された分子動力学エンジン (PMEMD) 内の精度混在モデルのベクトル化を進めることで、Intel® アーキテクチャー上でさらに高いパフォーマンスを引き出そうとしていました。並列スケーラビリティを高め、強固な基盤を構築するため、Intel® Xeon Phi™ コプロセッサのオフロード・アクセラレーションをサポートすることにしました。さらに、ユーザーの要件を満たすため Amber* を Intel® メニー・インテグレートッド・コア (Intel® MIC) アーキテクチャー対応にすることにしました。

ウォーカー研究室は、ソフトウェアを最新の Intel® プロセッサで実行できるように、以前の経験に基づいて Amber* を変更しました。しかし、そのパフォーマンスは予想を大幅に下回るもので、どうすればパフォーマンスを向上できるのか特定することは困難でした。コード内の単精度から倍精度へのキャスト動作において、最初の仮定が正しくないことが分かりましたが、エンジニアはなぜ予想したパフォーマンスの向上が得られないのか原因が分からず困っていました。

Amber* は、計算負荷の高いアプリケーションです。そのため、ユーザーは、ソフトウェアを最先端の HPC コンピューター・アーキテクチャーで実行することを要求します。ウォーカー研究室は Intel と協力して、Intel® Parallel Computing Center で、Amber* を Intel® Xeon Phi™ コプロセッサで実行できるようにし、その後 Intel® Xeon® プロセッサと Intel® Xeon Phi™ コプロセッサ両方のアーキテクチャーでパフォーマンスを最適化するプロジェクトを開始しました。

ソリューション

Intel のアプリケーション・エンジニア Ashraf Bhuiyan と共同作業を行うにあたり、ウォーカー研究室チームは Intel® VTune™ Amplifier を導入し、Intel® MPI ライブラリー (Intel® MPI) と Intel® マス・カーネル・ライブラリー (Intel® MKL) の最新機能を利用することにしました。Intel® MPI ライブラリーをサポートすることにより、並列パフォーマンスがすぐに 10 ~ 20 パーセント (MPI プロセス数に依存) 向上しました。また、Intel® MKL 内の Intel の最適化 FFT ルーチンを利用することで、ウォーカー研究室が元々使用していた pubFFT の修正バージョンのパフォーマンスが向上しました。

さらに、Intel® VTune™ Amplifier を利用することにより、ウォーカー研究室チームはコードを詳しく調査して、多くのパフォーマンス・ボトルネックを特定することができました。特に、倍精度から単精度への過度なキャストが行われていたため、精度混在モデルで予想した速度向上を妨げていたコードの場所を特定することができました。ウォーカー研究室は、Intel® VTune™ Amplifier を利用してこの問題が発生しているコード行を特定し、Intel® コンパイラーのエンジニアに連絡してコンパイラーが不要なキャストを行っていた原因の調査を依頼しました。Intel® VTune™ Amplifier のタイミングを行単位に分割して提供する機能がなければ、このパフォーマンス hotspot の場所を特定することは不可能だったでしょう。

Ashraf は、最適なパフォーマンスが得られるように Intel® MPI ライブラリーをチューニングするだけでなく、Amber* の hotspot を特定するため Intel® VTune™ Amplifier 環境のセットアップに協力し、現在使用しているオフロードコードも開発しました。この共同作業は現在も続けられています。Amber* v14 は、ビルドオプションを使用して Intel® MPI ライブラリーをサポートしました。さらに、Ashraf は、Intel® MPI ライブラリーをサポートするための正しいパスと環境の構築、SDSC の Intel® Xeon Phi™ コプロセッサ開発ハードウェアおよび関連ソフトウェアの準備も支援してくれました。

結果

パフォーマンスの向上

ウォーカー研究室は、インテル® VTune™ Amplifier でチューニング作業を行い、インテル® MPI ライブラリーおよびインテル® MKL を使用するようにシミュレーションを変更しました。インテル® MPI ライブラリー、インテル® コンパイラー、インテル® MKL を使用することにより、陽溶媒 PME (Particle Mesh Ewald) シミュレーションの並列スケーリングのパフォーマンスは約 20 パーセント、シリアル領域のパフォーマンスは (シミュレーションに応じて) 10 ~ 15 パーセント向上しました。また、暗溶媒 GB (Generalized Born) シミュレーションのパフォーマンスは 2.7 倍から 4.0 倍と大幅に向上しました。次の表は、サンプル・データ・セットのパフォーマンス向上の詳細を示しています。インテル® Xeon Phi™ コプロセッサのオフロードサポートに対応することで、陽溶媒シミュレーション (原子数 > 400,000) のパフォーマンスはさらに 10 パーセント向上しました。インテル® Xeon Phi™ コプロセッサ・アーキテクチャーからより長期的なパフォーマンス・ゲインを得るため、インテル® VTune™ Amplifier によるチューニング作業は現在も続けられています。

ソフトウェアと力場は、酵素、脂質膜、高度なバイオマテリアル、触媒のシミュレーションに使用されます。脂質結合タンパク質との相互作用に加えて、薬物結合の親和性を予測し、セルの薬物透過性を正確に判定する能力が提供されるため、創薬に効果的です。

モデル	システム	システムサイズ (原子数)	MPI/コンパイラー		スピードアップ (インテル/GNU*)
			GNU* - Red Hat* 4.4.7-3/ MPICH2 v1.5	インテル (2013/ impi 4.1.1.036)	
Generalized Born	TRPCAGE	304	63.57	176.12	2.77x
	MYOGLOBIN	2,492	1.17	4.43	3.79x
	NUCLEOSOME	25,095	0.01	0.04	4.00x
Particle Mesh Ewald-NVE アンサンプル	JAC - 4fs	23,558	13.09	15.36	1.17x
	JAC -2fs	23,558	6.88	8.01	1.16x

生産性

インテル® VTune™ Amplifier は、コードの hotspot を素早く特定できるようにして、開発期間を短縮しました。具体的には、不要なキャストによりパフォーマンスが低下している単精度/倍精度の精度混在モデルの領域を特定しました。インテル® VTune™ Amplifier により、ウォーカー研究室は、なぜコンパイラーがそのような方法で最適化を行っていたのか、異なるコードレイアウトで実験して調査することができました。また、異なるカットオフサイズや異なる FFT 格子サイズのような、異なる入力設定を用いて、コードの重要な hotspot を特定することができました。

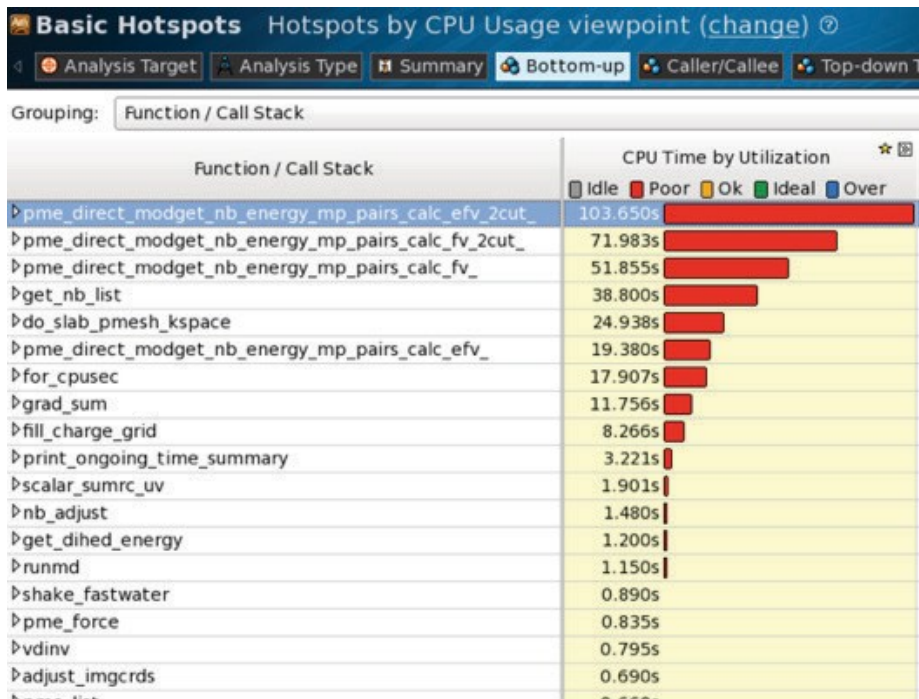


図 1. 特定のコードの hotspot の特定

このスクリーンショットから、pme_direct_modget_nb_energy には 4 つのバリエーションがあることがわかります。ウォーカー研究室は、インテル® VTune™ Amplifier を利用して、各関数のコスト追跡と最適化の優先度付けを行いながら、1 つのルーチンを 4 つの機能別のルーチンに分割して、内部ループで if 文を回避することができました。

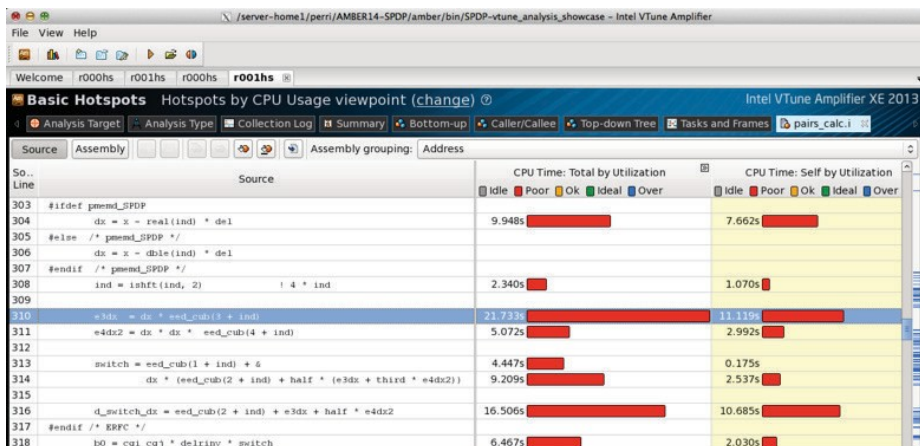


図 2. CPU の hotspot の特定

インテル® VTune™ Amplifier は、予想よりも多くの CPU 時間がかかっていたコード行の特定に大いに役立ちました。ウォーカー研究室でこれらの行に対応するコンパイル済みマシンコードをさらに調査した結果、コンパイラーが単精度から倍精度へのキャストを行い、再度逆のキャストを行っていたことがわかりました。コンパイル済み実行ファイルの hotspot の場所を特定した後、ウォーカー研究所はインテル® コンパイラーのエンジニアに問題の調査を依頼しました。

MPI パフォーマンスの向上により、シミュレーションの時間が短縮され、(より重要なことに) シミュレーションをより長く実行できるようになりました。ユーザーは、より広範な創薬ターゲットのような、異なる系のより多くのシミュレーションを実行できるようになりました。また、個々のシミュレーションの収束が改善され、まれに発生するイベントやより長いタイムスケールの動きを見ることができるようになりました。これにより、科学的な発見の可能性が高まることが期待されます。

結論

Amber* は現在 GPU、そしてインテル® Xeon Phi™ コプロセッサ・アーキテクチャーをサポートしています。ウォーカー研究室は、インテル® Xeon Phi™ コプロセッサの柔軟性により、ユーザーがネイティブモードで多くのシミュレーションを同時に実行したり、オフロードモードで Amber* GPU エンジンと同等のスピードアップが得られると予想しています。

ウォーカー研究室は、優れた最適化とパフォーマンスがもたらされると確信しています。特に、最新のインテル® Xeon Phi™ コプロセッサに焦点を当てたことで、より高速なシミュレーションが可能になるでしょう。Amber* のユーザーは、より長いシミュレーションを実行でき、より速いペースで科学研究を行うことができることから、多くの利点が得られるでしょう。また、デスクトップで対話型の計算や計算の試行が可能なシミュレーションからリアルタイムのフィードバックを行うような、科学研究の新しいアプローチを切り拓くことを期待しています。

インテル® ソフトウェア開発ツールについて

インテルは、25 年以上にわたって、ハイパフォーマンス・コンピューティング業界の開発者向けに標準準拠のツールを提供しています。業界最先端のソフトウェア・ツールには、インテル® Parallel Studio XE Composer Edition、Professional Edition、Cluster Edition が含まれます。

詳細:

<https://www.isus.jp/products-list/>

概要

カリフォルニア大学サンディエゴ校サンディエゴ・スーパーコンピューター・センターのウォーカー分子動力学研究室の研究対象には、計算化学、分子生物学、ハイパフォーマンス・コンピューティング (HPC) の分野が含まれます。研究所は、古典 MM 動的シミュレーションの計算効率と精度の向上に加えて、特に量子力学と複合量子/分子力学 (QM/MM) の並列計算および GPU 高速計算の効率的なアルゴリズムの開発に取り組んでいます。また、タンパク質ベース系の物理的・化学的性質の決定における MM および QM/MM 分子動力学アルゴリズムの使用にも注目しています。

ウォーカー研究室は、米国国立科学財団 (NSF)、米国国立衛生研究所 (NIH)、カリフォルニア大学、インテル、NVIDIA、Microsoft、英国外務・英連邦省、ビジネス・イノベーション・職業技能省から研究助成金を受けています。研究所で開発され実用化された技術には、次世代のウィルス阻害剤、バイオエタノール生産の改良、創薬、スーパーコンピューター上のハイパフォーマンス科学アプリケーション用の高度なアルゴリズムなどがあります。

詳細: <http://www.wmd-lab.org/> (英語)

インテル® ソフトウェア開発ツールの詳細は、
<https://www.isus.jp/products-list/> を参照してください。



性能に関するテストに使用されるソフトウェアとワークロードは、性能がインテル® マイクロプロセッサ用に最適化されていることがあります。SYSmark* や MobileMark* などの性能テストは、特定のコンピューター・システム、コンポーネント、ソフトウェア、操作、機能に基づいて行われたものです。結果はこれらの要因によって異なります。製品の購入を検討される場合は、他の製品と組み合わせた場合の本製品の性能など、ほかの情報や性能テストも参考にして、パフォーマンスを総合的に評価することをお勧めします。

詳細については、<http://www.intel.com/performance/> (英語) を参照してください。

インテル® ソフトウェア製品のパフォーマンスおよび最適化に関する注意事項については、<http://software.intel.com/en-us/articles/optimization-notice/#opt-jp> を参照してください。

最適化に関する注意事項: インテル® コンパイラーでは、インテル® マイクロプロセッサに限定されない最適化に関して、他社製マイクロプロセッサ用に同等の最適化を行えないことがあります。これには、インテル® ストリーミング SIMD 拡張命令 2、インテル® ストリーミング SIMD 拡張命令 3、インテル® ストリーミング SIMD 拡張命令 3 補足命令などの最適化が該当します。インテルは、他社製マイクロプロセッサに関して、いかなる最適化の利用、機能、または効果も保証いたしません。本製品のマイクロプロセッサ依存の最適化は、インテル® マイクロプロセッサでの使用を前提としています。インテル® マイクロアーキテクチャーに限定されない最適化のなかにも、インテル® マイクロプロセッサ用のものがあります。この注意事項で言及した命令セットの詳細については、該当する製品のユーザー・リファレンス・ガイドを参照してください。

注意事項の改訂 #20110804

© 2017 Intel Corporation. 無断での引用、転載を禁じます。Intel、インテル、Intel logo、Xeon、Intel Xeon Phi、VTune は、アメリカ合衆国および / またはその他の国における Intel Corporation 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。

JPN/1707/PDF/XL/SSG/TT