

ケーススタディー

ライフサイエンスの
ハイパフォーマンス計算



Software

膨大な情報を 管理する

Lab7 Systems はインテル® Parallel Studio XE とインテル® C++ コンパイラーを使用した最適化により BioBuilds* ツールのパフォーマンスを向上

「インテル® コンパイラーに含まれる現代のマルチコア・プロセッサのパフォーマンスを引き出す自動ベクトル化や自動並列化などの最適化により、インテル® 64 アーキテクチャ上で BioBuilds* パッケージのパフォーマンスを大幅に向上することができました。」

—Lab7 Systems
主席ソフトウェア・アーキテクト
Cheng Lee 氏

新しいテクノロジーにより生成される大量のデータを効率良く管理する方法を見つけることは、多くの産業にとって重要な懸案事項です。特に、増え続ける膨大なデータに基づいて画期的な研究が行われているライフサイエンスの世界では、大きな課題となっています。

Lab7 Systems は、生命情報学者、科学者、IT チームの取り組みを支援するため、インテル® Parallel Studio XE を使用してオープンソースの BioBuilds* ツール・コレクションを最適化しています。

オープンソース・ソリューション

テキサス州オースチンにある Lab7 Systems は、新しいライフサイエンス・テクノロジーにより生成される大量のデータの管理作業を軽減し、絶えず進化しているハードウェアから最大限のパフォーマンスを引き出せるように支援するソフトウェアを開発しています。Lab7 Systems の主要商用製品である Lab7 Enterprise Science Platform* (ESP) は、データ量の多い研究所での操作向けに開発された、最初の完全な "Sample-to-Answer" 型ソフトウェア・プラットフォームです。

Lab7 Systems は、Linux* (x86 および IBM* Power Systems* プラットフォーム) 向けと OS X* 向けに、事前にビルドされたゲノム研究用のオープンソースの生命情報学ツールのコレクションである、BioBuilds* の管理、保守、およびサポートも行っています。サポートするライブラリーをすべて含めることにより、BioBuilds* は、これらのツールのカスタムバージョンを保守およびビルドする必要性を排除し、ほかのソフトウェア・パッケージの特定のバージョンに依存するソフトウェア・パッケージをインストールしている人々に「依存性地獄」から抜け出す道を提供します。BioBuilds* は、共同作業者が同じバージョンのツールを使用していることを保証する基準点を提供することにより、作業の再現性を実現します。BioBuilds* は、サポートするすべてのプラットフォームで一貫した計算環境を保証するため、conda パッケージ・マネージャーを使用しています。標準の BioBuilds* リリースは無料でダウンロードして使用できます。Lab7 Systems では、追加のサポートやカスタムパッケージが必要なユーザー向けに有料のオプションも用意しています。

パフォーマンスの最大化

Lab7 Systems は、インテル® 64 プラットフォームでパフォーマンスを最大限に引き出すため、インテル® Parallel Studio XE に含まれているツールを利用しました。このツールセットは、現在および次世代のプロセッサでスケーリングする、優れた C++、Fortran、Python* アプリケーション・パフォーマンスを実現できるように開発者を支援します。また、高速かつ安定した並列コードの作成プロセスを容易にします。

Lab7 Systems は、GNU* コンパイラー・コレクション (GCC) の代わりにインテル® コンパイラーの利用をサポートするようにアップストリームのビルドシステムを変更しました。BioBuilds* 2017.05 の一部のバイナリーのビルドにインテル® C++ コンパイラーを使用したところ、GCC でビルドしたバージョンと比較してパフォーマンスが大幅に向上しました。



ケーススタディー | 膨大な情報を管理する

Lab7 Systems は、次のような要因により、インテル® C++ コンパイラーを選択しました。

- 標準 C/C++ 言語のサポート
- GCC との互換性
- インテル® 64 向けの優れた最適化
- インテル® 64 向けの自動 SIMD ベクトル化および並列化
- プロシージャー間の最適化
- 並列コードの作成と向上に役立つプロファイラーおよび解析ツール

最適化の種類には、SSE4.2 アーキテクチャー向けの明示的な最適化 ("-x"), AVX, CORE-AVX-I, CORE-AVX2 アーキテクチャー向け代替コードバスの生成 ("-ax"), 最適化 ("-O3"), プロシージャー間の最適化 ("-ip" および "-ipo"), 自動並列化 ("-parallel") が含まれます。Lab7 Systems のいくつかのパッケージでは、コンパイルエラーやランタイムエラーを回避するため、通常はこれらのオプションを使用していませんでした。

Lab7 Systems は、いくつかのアプリケーションで、インテル® Parallel Studio XE でビルドしたことにより向上したパフォーマンスを計測するベンチマーク・テストを行っています。

「インテル® コンパイラーに含まれる現代のマルチコア・プロセッサのパフォーマンスを引き出す自動ベクトル化や自動並列化などの最適化により、インテル® 64 アーキテクチャー上で BioBuilds* パッケージのパフォーマンスを大幅に向上することができました。さまざまなハードウェアで最適化されたパフォーマンスが保証される、複数の自動ディスパッチ・コードを含むバイナリーを生成できることも大きなポイントです。」と Lab7 Systems の主席ソフトウェア・アーキテクト、Cheng Lee 氏は述べています。

継続的なサポート

もう 1 つの大きな要因は、このオープンソース・プロジェクトにおける (資金およびインテル® Parallel Studio XE ライセンスの提供を含む) インテルのサポートでした。

BioBuilds* のリリースポリシーに従って、Lab7 Systems は、インテル® コンパイラーを使用して最適化した生命情報学ツールを無料で提供しています。「これは、Lab7 Systems が生命情報学におけるオープンソースのサポートに深く関与していることを示すものです。ユーザーは、通常は利用できない事前にビルドされたハイパフォーマンス・バイナリーを利用することができます。」(Lee 氏)。

さらに、Lab7 Systems では、特定のインテル® 64 プロセッサ向けに追加の最適化が必要なユーザーやライセンスの制限などの問題によりパブリック (無料) リリースに含めることができない最適化パッケージを利用したいユーザー向けに有料のサポート契約も提供しています。

今後の予定

Lab7 Systems は、2017.11 リリースのパフォーマンス・チューニング・パッケージのビルドにもインテル® コンパイラーを引き続き使用する予定です。また、インテル® Parallel Studio XE のコンポーネントを使用して特定のアプリケーションのパフォーマンスをさらに向上することも計画しています。

- **インテル® VTune™ Amplifier:** アプリケーション・パフォーマンス、スケーラビリティ、メモリアクセスのチューニングに役立つ高度なパフォーマンス・プロファイラー。C、C++、Fortran、Python*、Go*、アセンブリー、Java* に加えて、これらの言語が混在したコードを正確にプロファイルできます。
- **インテル® Advisor:** ベクトル化の最適化とスレッドのプロトタイプ生成に役立つ C、C++、C#、Fortran アプリケーション向けのベクトル化/スレッド化アドバイザー・ツール。

関連資料

[インテル® Parallel Studio XE](#)

[インテル® C++ コンパイラー](#)

[BioBuilds* \(英語\)](#)



インテル® テクノロジーの機能と利点はシステム構成によって異なり、対応するハードウェアやソフトウェア、またはサービスの有効化が必要となる場合があります。

実際の性能はシステム構成によって異なります。絶対的なセキュリティを提供できるコンピューター・システムはありません。詳細については、各システムメーカーまたは販売店にお問い合わせいただくか、<http://www.intel.co.jp/> を参照してください。

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

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、機能に基づいて行ったものです。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、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、<https://www.intel.com/performance> (英語) を参照してください。

インテルは、本資料で参照しているサードパーティーのベンチマーク・データまたは Web サイトの設計や実装について管理や監査を行っていません。本資料で参照している Web サイトまたは類似の性能ベンチマーク・データが報告されているほかの Web サイトも参照して、本資料で参照しているベンチマーク・データが購入可能なシステムの性能を正確に表しているかを確認されるようお勧めします。

この文書および情報は、インテルのお客様向けの参考情報として記載されているものであり、現状のまま提供され、明示されているか否かにかかわらず、いかなる保証もいたしません。ここにいう保証には、商品適合性、特定目的への適合性、知的財産権の非侵害性への保証を含みますが、これらに限定されるものではありません。本資料は、本資料に記述、表示、または記載された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。インテル製品は、医療、救命、延命措置、重要な制御または安全システム、核施設などの目的に使用することを前提としたものではありません。

© 2017 Intel Corporation. 無断での引用、転載を禁じます。Intel、インテル、Intel ロゴ、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/1710/PDF/XL/SSG/TT