平成 19 年度
実績報告

「情報システムの超低消費電力化を目指した技術革新と統合化技術」
平成 19 年度採択研究代表者

松岡　聡

東京工業大学学術国際情報センター・教授

ULP-HPC: 次世代テクノロジのモデル化・最適化による超低消費電力ハイパフォーマンスコンピューティング

1. 研究実施の概要

平成 19 年度は計算アクセラレータを中心に、メモリ・ネットワークなど、それぞれのスーパコンピュータの構成コンポーネンツに関して消費電力を考察した性能モデルの構築と、数倍～数十倍におよぶ高性能化や・大幅な省電力化を行うアーキテクチャおよびアルゴリズムを数多く要素技術的に提案し、さらにそれらの容易なプログラミングや自動最適化へ向けた基礎的な成果を数多くあげることができた。特に、従来は比較的密結合に限られていた GPU による計算加速において、流体計算において最適なエネルギー消費量では従来とあまり変わらない形で、数倍～数十倍の性能向上が得られ、結果として数倍～数十倍のエネルギー性能比向上が得られた。密結合問題においても、TSUBAME で低消費電力アクセラレータ ClearSpeed を有効に用いるアルゴリズムを研究開発し、CPU のみに比較して 2 倍近い性能向上を得、結果としてエネルギー効率を 2 倍向上させ、TSUBAME が二年連続で日本一のスパコンとなる原動力となった。その他、自動最適化においても今まではアドホックな探索で行われていたものに対し、数理的に精緻な結果を世界で初めて提示し、高い注目を浴びた。

2. 研究実施内容

（文中にある参照番号は４．(1)に対応する）

「研究代表者・松岡」グループ（東京工業大学）

①　行列積や FFT などの数値計算を GPU で効率よく実行する手法の研究を推進した。まず、GPU 上での FFT 演算の高速化を行った。特に FFT では GPU とホスト間のデータ転送がボトルネックとなりやすいが、GPU 上のデバイスメモリへのアクセスを最適化することによって転

送時間を含んだ場合でも最新の quad core CPU の 2 倍以上の性能を達成した[7]。また、GPU と CPU の併用、さらには複数の GPU を用いる並列処理手法の研究を推進した。まず、CPU と GPU へのタスク割り振り率の決定を自動化するため、性能モデルを構築した[1,3,5]。これにより、2D-FFT の実行時間を 15%以内の誤差で予測できることを示した。その際、CPU 単体と比較して 1.5 倍の性能向上が得られた。また複数の性能が異なる GPU を用いた行列積の計算では、セルフスケジューリングによる動的タスク分配を用い、静的に割り振った場合よりも良い性能が得られることを確認した。

② 省電力型アクセラレータによる HPC の大規模加速実験およびモデリングを行った。大規模加速実験として、東京工業大学 TSUBAME システム上で並列 Linpack の実験を、本年度導入の ClearSpeed アクセラレータを併用して行った[4]。10000 コア以上の汎用 CPU と 600 枚以上のアクセラレータにカーネル演算を適切に割り振ることにより、56.43TFlops の性能を達成した。この結果は Top500 スパコンランキングにおいて、日本 1 位、世界 16 位にランクされた。また、モデリングの研究として、汎用 CPU とアクセラレータが混在する HPC システムを想定し、性能モデルとジョブスケジューリングアルゴリズムの提案を行った。様々な加速特性を持つジョブ群をシミュレートし、make-span および ED 積の評価を行い、電力を考慮しない単純な手法に対する優位性を確認した。

③ 次世代低電力メモリを有効利用するシステムの提案・評価を行った。提案システムでは電力コストが大きい DRAM の搭載容量を削減するために、メインメモリの一部を MRAM に置き換え、スワップデバイスとして FLASH メモリを使用する。そしてメモリアクセスを高速な MRAM に集中させるような省電力ページング方式を提案した。シミュレーションにより、DRAM 容量を適切に削減した場合に、アプリケーションベンチマークの性能低下を 12%に抑えつつ、メモリモジュールの消費エネルギーを 26%に削減できることを示した[2,6]。

「主たる共同研究者①・須田」グループ（東京大学）

須田グループでは，超省電力 HPC のための自動チューニング技術の開発を開始した．自動チューニングはソフトウェアに組み込まれたパラメタを変更しながら実環境で動作させて実効性能を評価することにより，最適な性能を与えるパラメタを自動的に選択する機能である．エネルギーという視点では，自動チューニングにより性能が低いパラメタで実行することによる無駄なエネルギー消費を抑えることができるが，最適なパラメタを探索するために行う試行が多すぎるとそこで消費されるエネルギーが無視できなくなる．この問題を解決するため，我々は実際の利用時にパラメタ探索を行う「オンライン自動チューニング」の概念を提案した [1]．さらに，オンライン自動チューニングのための数理基盤として Bayes 統計を用いることを提案し，Bayes 統計を用いたオンライン自動チューニングのための準最適アルゴリズムを提案，行列積を例題として評価を行い，その有効性を実証した [1]．また，超省電力 HPC のためのハードウェアとして期待される GPU を用いた数値計算に本手法を適用し，良好な結果を得た．また，超省電力 HPC に向けて最適化された

数値計算ライブラリの構築をめざし，GMRES 法のリスタート周期やヘテロクラスタにおける通信アルゴリズムなどの最適化を自動的に行うソフトウェアを実装し，Windows CCS クラスタ上でその有効性を実証した.

「主たる共同研究者②・青木」グループ（東京工業大学）

電力当たりのピーク性能が通常の CPU より 1 桁以上高い GPU を用い、高精度流体計算[1]の最も負荷の高い圧力の Poisson 方程式の計算を行った。これまで粒子モデルの計算[2,4]に GPU を使った報告はあったが、格子計算に対して GPU を適用するための多くの知見が得られた。格子系の計算ではメモリアクセスに対する 1 格子点上の計算量が少なく、G80 アーキテクチャの場合は shared メモリをキャッシュ的に使い、ビデオメモリへのアクセス回数を低減させるアルゴリズムを開発した。この計算アルゴリズムを用い、Point Jacobi 法による Poisson 方程式の計算で nVIDIA GeForce 8800 Ultra を使うことにより約 30 倍、4 GPU では 94 倍の加速が得られた。これは NEC SX-8 の 15 CPU に相当する性能である。また、高精度流体計算スキーム[1]は低次精度スキームと比較して 1 格子点当たりの計算量が多いため、GPU のような SIMD 型アクセラレータに向いていることも明らかになった。

GPU を搭載した PC に対して、電流・電圧測定ユニットに DC/AC クランプメータを組み合わせて、50 マイクロ秒の解像度でチャネル当たり 16 百万サンプル（800 秒）の測定が可能なデジタルオシロスコープ（日置電機 8855）を用いて電力測定を行った。1 台の PC において無負荷定常時、CPU のみを使う計算、ディスク入出力等の測定を行い、電力測定に関わる基礎的な知見を得た。DC 側の消費電力変動が時定数 10 ミリ秒程度の遅延で AC 側の消費電力変動として現れることを確認し、数十ミリ秒以上の消費電力変動であれば AC 側の測定から DC 側の消費電力変動を過渡的な現象の影響が大きくなく評価できることが明らかとなった。

上記 Poisson 方程式の計算を CPU のみで実行した場合と GPU（nVIDIA GeForce 8800 GT）を利用して実行した場合の消費電力測定を行った。GPU を利用した場合は、消費電力が CPU のみを利用する場合よりも平均で 7.5%大きくなるが、実行時間が 1/20 で済むために消費エネルギーは 5.3%で済むという結果が得られた。

「主たる共同研究者③・本多」グループ（電機通信大学）

本多グループにおいては、省電力化に有効な SIMD 型アクセラレータの有効活用を目指しその特徴である分散メモリに対応する統一的並列プログラミングインタフェースの暫定仕様を考察し、評価した。具体的には、SIMD 型アクセラレータ固有のプログラミングインタフェースが持つ性能を可能な限り維持しつつ、統一的なプログラミングインタフェースを提供することを目指した。このために解決する必要がある問題点は現状のプログラミングインタフェースにおいて明示的に計算とメモリを結びつけて確保できない点にあることを明らかにし、この問題に対応するために計算とメモリ

をそれぞれ確保し、これらを関連付けることができる統一的プログラミングインタフェースを提案した。また、SIMD 型アクセラレータとして有効活用できる GPU に対応するため MPI、OpenMP などの各種既存のプログラミングインタフェースの GPU 上での実装方法を明らかにし、複数の階層にわたる分散メモリ環境に対応するためのメッセージ通信インタフェースの適応報告を行った。また、分散メモリ型マルチコアプロセッサに対応するマルチスレッド型プログラミングインタフェースの実装、評価を行った。

「主たる共同研究者④・鯉渕」グループ（国立情報学研究所）

省電力インターコネクトの研究開発

HPC システムにおいてネットワークの連結性を保障する最低限のリンクのみ常にアクティブにし、その他のリンクのみ柔軟にアクティベイション制御を行う手法の検討を行った。この方法により、バーストトラフィックが発生した場合にもアクティベイト制御のオーバヘッドによるパケット遅延の発生を抑え、かつ、最低限必要な帯域を提供することができる。さらに、チップ間通信のみならず、チップ内通信についても検討を行った。通信遅延を抑えることで全体としての電力を抑える手法[1]、トラフィックパターンを事前に解析し、使用率の低い複数のリンクで１つのスイッチの入力ポートを共有する仕組みについて検討を行った。その結果、スループットの低下なしにハードウェア量を 55%削減できる場合があることがわかった。また、HPC 分野で使用される並列ベンチマーク・アプリケーションを、典型的な SIMD 型アクセラレータである ClearSpeed 製の CSX600 プロセッサに実装することで、アクセラレータボードの性能要因を通信性能から解析を行った。

「主たる共同研究者⑤・日向寺」グループ（東海大学）

「巨大分子量子化学計算における超省電力 HPC システムの性能評価」

平成１９年度、東海大グループでは、高電力コスト計算の実例として、生体分子などの大規模な分子サイズを対象とした高精度量子化学計算の入力データを構築した。大規模量子化学計算のコストを大幅に削減することが可能な手法として期待されているフラグメント分子軌道法を採用し、計算精度としては、電子相関を取り入れない Hartree-Fock 法と、摂動論によって電子相関を考慮するが電力コストの高い MP2 法に対して、複数の基底関数を使用して検証計算を行った。その結果、今回、検証を行った分子構造（抗体－糖鎖のモデル構造）では、MP2 法による計算は、電力コストが高いが、生体分子間相互作用に関してより良い結果を与えることが分かった。

3. 研究実施体制

(1)「研究代表者・松岡」グループ

①研究分担グループ長:松岡　聡(東京工業大学、教授)

②研究項目

次世代 HPC システムにて超省電力・高性能を達成するハードウェア・ソフトウェア統合システムの研究開発

(2)「主たる共同研究者①・須田」グループ

①研究分担グループ長:須田　礼仁(東京大学、准教授)

②研究項目

超省電力 HPC ソフトウェアのための自動チューニングの研究開発

(3)「主たる共同研究者②・青木」グループ

①研究分担グループ長:青木尊之(東京工業大学、教授)

②研究項目

超省電力型の HPC アプリケーション及びアルゴリズムの研究開発

(4)「主たる共同研究者③・本多」グループ

①研究分担グループ長:本多　弘樹(電気通信大学、教授)

②研究項目

超省電力化 SIMD アクセラレータのための汎用プログラミング環境

(5)「主たる共同研究者④・鯉渕」グループ

①研究分担グループ長:鯉渕　道紘(国立情報学研究所、助教)

②研究項目

省電力インターコネクトの研究開発

(6)「主たる共同研究者⑤・日向寺」グループ

①研究分担グループ長:合田(日向寺)　祥子(東海大学、講師)

②研究項目

巨大分子量子化学計算における超省電力 HPC システムの性能評価

4. 研究成果の発表等

(1) 論文発表(原著論文)

「研究代表者・松岡」グループ(東京工業大学)

[1] 尾形泰彦,丸山直也,遠藤敏夫,松岡聡."性能モデルに基づく CPU および GPU を併用する効率的な FFT ライブラリ".2008 年ハイパフォーマンスコンピューティングと計算科学シンポジウム論文集(HPCS2008), pp. 107-114, Jan 2008.

[2] 細萱祐人,遠藤敏夫,松岡聡."省電力ページング方式を実装した次世代メモリアーキテクチャ上での並列プログラム".2008 年ハイパフォーマンスコンピューティングと計算科学シンポジウム論文集(HPCS2008), pp. 25-32, Jan 2008.

Accepted:

[3] 尾形泰彦,丸山直也,遠藤敏夫,松岡聡."性能モデルに基づく CPU 及び GPU を併用する効率的な FFT ライブラリ".情報処理学会論文誌, Vol.49, ACS22, 2008.

[4] Toshio Endo and Satoshi Matsuoka. Massive Supercomputing Coping with Heterogeneity of Modern Accelerators. In Proceedings of IEEE International Parallel & Distributed Processing Symposium (IPDPS 2008), April 2008.

[5] Yasuhiko Ogata, Toshio Endo, Naoya Maruyama and Satoshi Matsuoka. An Efficient, Model-Based CPU-GPU Heterogeneous FFT Library. In Proceedings of 17th International Heterogeneity in Computing Workshop (HCW '08), in conjunction with IPDPS 2008, April 2008.

[6] Yuto Hosogaya, Toshio Endo and Satoshi Matsuoka. Performance Evaluation of Parallel Applications on Next Generation Memory Architecture with Power-Aware Paging Method. In Proceedings of The Fourth Workshop on High-Performance, Power-Aware Computing (HPPAC), in conjunction with IPDPS 2008, April 2008.

[7] 額田彰,尾形泰彦,遠藤敏夫,松岡聡."CUDA 環境における高性能 3 次元 FFT". 先進的計算基盤システムシンポジウム SACSIS2008 論文集, June 2008.

「主たる共同研究者①・須田」グループ(東京大学)

[1] 須田礼仁(東京大学)、「オンライン自動チューニングのための Bayes 統計に基づく逐次実験計画法」、情報処理学会 2008 年ハイパフォーマンスコンピューティングと計算科学シンポジウム(HPCS2008)、東京工業大学 大岡山キャンパス、2008 年 1 月 17 日、pp. 73-80.

「主たる共同研究者②・青木」グループ(東京工業大学)

[1] Yohsuke Imai, Takayuki Aoki and Kenji Takizawa, Conservative form of interpolated differential operator scheme for compressible and incompressible fluid

dynamics, Journal of Computational Physics, Vol. 227, Issue 4, 2008, 2263-2285
[2] S. Moriguchi and T. Aoki: Simulation of Free Surface Flow Interacting with Moving Particles by Using Immersed Boundary Method, International Conference on Violent flows 2007, p301-303, 2007, Nov20-22, Fukuoka
[3] Kenta Sugihara, Takayuki Aoki: Accuracy study of the IDO-CF scheme by Fourier analysis, The Third Asian-Pacific Congress on Computational Mechanics and the Eleventh International Conference on the Enhancement and Promotion of Computational Methods in Engineering and Science (APCOM'07 in conjunction with EPMESC), P.89, 2007, Dec 3-6, Kyoto
[4] S. Moriguchi and T. Aoki: Numerical method for geomaterial based on fluid-particle interaction, The Third Asian-Pacific Congress on Computational Mechanics and the Eleventh International Conference on the Enhancement and Promotion of Computational Methods in Engineering and Science (APCOM'07-EPMESC), P.493, 2007, Dec 3-6, Kyoto
[5] Satoi Ogawa, Takayuki Aoki, Toru Tamagawa: Numerical Simulation for Vertical-Axis Wind Turbine by High-accurate Overset Grid method, The Third Asian-Pacific Congress on Computational Mechanics and the Eleventh International Conference on the Enhancement and Promotion of Computational Methods in Engineering and Science (APCOM'07-EPMESC XI), P.81, 2007, Dec 3-6, Kyoto

「主たる共同研究者④・鯉渕」グループ（国立情報学研究所）

[1] Michihiro Koibuchi(NII), Hiroki Matsutani(Keio U), Hideharu Amano(Keio U), Timothy M. Pinkston(U of Southern California), ``A Lightweight Fault-tolerant Mechanism for Network-on-chip'', Proc. of the 2nd ACM/IEEE International Symposium on Networks-on-Chip (NOCS'08), pp.13-22, Apr 2008.