

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年6月4日現在

機関番号：12601

研究種目：基盤研究（B）

研究期間：2009～2011

課題番号：21300007

研究課題名（和文） メニーコア・超並列時代に向けた自動チューニング記述言語の方式開発

研究課題名（英文） Development of Auto-tuning Specification Language Towards Manycore and Massively Parallel Processing Era

研究代表者

片桐 孝洋（KATAGIRI TAKAHIRO）

東京大学・情報基盤センター・准教授

研究者番号：40345434

研究成果の概要（和文）：本研究では、多様な計算機環境において高性能を達成するため、自動チューニング（AT）を実現する以下の研究開発を行った。（1）マルチコア・超並列環境に適用できる AT 言語 ABCLibScript の機能拡張；（2）マルチコア CPU および GPU での AT 効果の検証；（3）いくつかのアプリケーションソフトウェアに新規 ABCLibScript 機能を適用し有効性を検証；（4）新規 ABCLibScript 処理系を、フリーソフトウェアとしてインターネット上に公開。

研究成果の概要（英文）：In this research, the following development is made to establish auto-tuning (AT) facility for high performance execution on several computer environments. (1) Function extension to an AT language, named ABCLibScript, for multicore and massively parallel environment; (2) Evaluation of the AT facility with multicore CPUs and GPUs; (3) Evaluation of effectiveness of the AT facility on ABCLibScript by adapting several application software; (4) Open the codes of preprocessor for the developed ABCLibScript as free software via the internet.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	7,600,000	2,280,000	9,880,000
2010年度	3,200,000	960,000	4,160,000
2011年度	3,200,000	960,000	4,160,000
年度			
年度			
総計	14,000,000	4,200,000	18,200,000

研究分野：総合領域

科研費の分科・細目：情報学・ソフトウェア

キーワード：高性能計算、並列化、マルチコア、GPU、ソフトウェア自動チューニング、ABCLibScript、コード自動生成、高生産性

1. 研究開始当初の背景

(1) 情報社会の基盤を支える計算機が高度に並列化し多様化している。サーバ系高性能計算機、クライアント系パーソナルコンピュータがマルチコア化や命令レベルでベクトル化され、高性能アーキテクチャの並列化・多様化が進展している。最先端の知を支えるス

ーパーコンピュータでさえも、多様な計算機アーキテクチャをもち、数千コアを有する超並列計算機である。近い将来、PC においては数十コア、スーパーコンピュータに至っては数十万コアという、メニーコア・超並列環境となる。計算機アーキテクチャ的に多様な並列計算機環境になることは疑う余地がない。こ

のような複雑化された並列計算機環境に高度に適合する、コンパイラを含む自動高速化の手段を我々は持っていない。

(2)高度に並列化され複雑化された計算機環境がもたらすチューニング作業の困難性の爆発的増大から、ソフトウェア性能を人手を介さずに最大限に引き出す仕組みが求められている。とりわけ、数値計算ソフトウェアの共通基盤となる数値計算ライブラリにおいては、個々のソフトウェアに特化された入力データに対しても高性能に動作することが強く要求される。このような状況のもと、アプリケーションの性能パラメータを対象計算機のキャッシュサイズ、コア数、通信性能などの計算機アーキテクチャの特性に加え、数値アルゴリズムの選択に至る広範な要因を自動的にチューニングするソフトウェア自動チューニング技術が、国内外から注目を集めている。

2. 研究の目的

以上の先行研究の成果と問題点を考慮し、本研究の目的は以下のようにまとめられる。

(i)将来利用されるメニーコア・超並列環境でも自動チューニングを達成する方式を研究する。

(ii) (i)を達成する自動チューニング記述用言語処理系のプロトタイプを行う。

3. 研究の方法

研究目的を達成するため、以下の項目の研究を進める。

●ABCLibScript の高度化：マルチコア環境に適用できる AT 機能の拡張を行う。特に、近年普及した GPU への対応のため、GPU とマルチコア CPU のプログラムに対しての AT 機能を持つ、ABCLibScript の設計と基本機能の開発を行う。

●マルチコア CPU および GPU 環境での AT 効果検証：上記 1 の AT 新機能の検証を、以下のアプリケーションで行う：(1) 行列-行列積、(2) ヤコビ反復解法、(3) 姫野ベンチマーク。10,000 コアを超える大規模な分散並列型計算機上で動作する固有値ソルバを自動チューニングする際に必要となるチューニング項目とパラメータの特定を実施する。並列処理における自動チューニングのための数理手法について研究し、知見を集積する。

●アプリケーションへの適用評価：AT 言語の実用性を検証するため、高精度行列-行列積プログラムに ABCLibScript 適用評価し有効性の検証を行う。疎行列を係数行列に持つ線形方程式に対する反復解法の前処理付きアルゴリズム、および、高周波電磁場解析で一般的 3次元 FDTD 法のキャッシュヒット

率を向上させる時空間タイリングによる実装法について AT の観点から効果検証をする。

●開発ソフトウェアの公開：開発した ABCLibScript の処理系のコードを、フリーソフトウェアとしてインターネット上に公開する。

4. 研究成果

以上の研究の方法の項目それぞれについて、本研究グループの研究分担者が行った研究成果を以下にまとめる。

(1) 研究代表者：片桐孝洋

●ABCLibScript の高度化：複雑化された並列計算機環境に高度に適合する自動チューニング (AT) を達成するため、将来利用されるメニーコア・超並列環境で通用する AT 方式を研究した。特に、「性能」を意識したソフトウェア工学の実現を目指すことを中心に、ソフトウェア開発コスト削減のため、専用 AT 言語 ABCLibScript の高度化を行った。

具体的には、ABCLibScript 用言語処理系のプロトタイプを中心に、以下の研究開発を行った：(i) C 言語版の ABCLibScript の新規開発；(ii) AT 言語で実装可能な AT 方式の開発；(iii) マルチコア・超並列計算機での AT 効果の有効性検証；

●マルチコア CPU および GPU 環境での AT 効果検証：ABCLibScript による最適化機能を、将来のメニーコア・超並列環境で用いるためには、現在主流となっている GPU (Graphics Processing Unit) と、マルチコア CPU との混合構成の計算機環境 (いわゆる、ハイブリッド計算機環境) における AT 機能の実現が不可欠である。そこで本研究では、マルチコア CPU と GPU の双方に対し適切に処理が分割できる機能を持つ ABCLibScript の新 AT 機能を開発することに成功した。この機能を持つ ABCLibScript を HxABCLibScript (Heterogeneous eXtended ABCLibScript) という。

●アプリケーションへの適用評価：OpenMP 記述から、GPU で直接起動する計算機言語 CUDA に変換可能なコンパイラ OMPCUDA と連携することで、より効率のよい AT 機能を実現することが出来た。また ABCLibScript による AT 機能は、姫野ベンチマークや、早稲田大学で開発された高精度な行列-行列積アルゴリズム (尾崎の補法) の計算へも適用し、十分な AT による速度向上の効果があることが確認できた。

アプリケーションで必要となる AT 方式の開発事例として、固有値ソルバ ABCLib_DRSSD、および、疎行列反復解法で必須となる疎行列-ベクトル積 SpMV (Sparse Matrix-vector Multiplication) における新 AT 方式の開発に成功した。このことで、近年

のハイブリッド計算機環境でも適用できる AT 方式へ展開ができた。

●開発ソフトウェアの公開：以上の成果を凝縮した ABCLibScript のソースコードを、フリーソフトウェアとして整備し、WEB で公開を行った。

(3) 研究分担者：今村俊幸

●マルチコア CPU および GPU 環境での AT 効果検証：10000 コアを超える大規模な分散並列型計算機上で動作する固有値ソルバを自動チューニングする際に必要となるチューニング項目やパラメタ（ブロックサイズのみならずアルゴリズム選択も含む）の特定を実施した。

また、固有値ソルバ（もしくは一般の線形計算プログラム）には冗長計算と集団通信を併用するが、複数プロセスグループ間での通信を削減する実装方法ならびに、集団通信のアルゴリズム選択機構について学生と協力研究を開始した。

さらに、d-Spline 関数と呼ばれる離散型の spline 関数を利用して、MPI の集団通信のコストを自動的に推定する手順として、逐次サンプリングデータ追加型の最適値探索アルゴリズムと、逐次サンプリングデータ追加型の関数形状近似アルゴリズムの研究を進めた。

(4) 研究分担者：須田礼仁

●ABCLibScript の高度化：本研究では、並列処理における自動チューニングのための数理手法について研究した。まず、自動チューニングと並列処理の関係について論じ、並列処理全体に影響する単一のチューニングパラメタを最適化する大域的チューニングと、各プロセッサの処理に影響するプロセッサごとのチューニングパラメタを最適化する局所的チューニングがあることを指摘した。そして、局所的チューニングに関する以下の手法を提案した。

複数候補の並列評価（「並列実験」と呼ぶ）は、並列計算機の多数のプロセッサに異なる候補を割り当てて、多数の候補の性能を並列に評価する手法である。ベイズモデルに基づき、準最適な実験計画法を示した。試行専用プロセッサによる並列試行（「並列試行」と呼ぶ）は、一定数プロセッサが試行を行い、よい候補が見つければ他の実働プロセッサにそれを知らせる方法である。実行の中断が本質的であり、試行専用プロセッサは 1 台がよいことを示した。

これらは並列自動チューニングの足掛かりになる研究成果であり、今後はこれらを融合した高度な手法を開発してゆく。

(5) 研究分担者：黒田久泰

●マルチコア CPU および GPU 環境での AT 効果検証：現在、GPU を利用して大規模な計算を行うことが頻繁に行われるようになってきた。しかしながら、現在の GPU は、ストリーミングプロセッサ数、メモリ容量、メモリ帯域などが異なる様々な機種が発売されており、それぞれの GPU が持つ性能を最大限に引き出すような単一のプログラムを作成することは困難になってきている。

通常、GPU に搭載されているメモリ容量はパソコン本体に搭載できるメモリ容量に比べて小さいため、頻繁に CPU と GPU との間でデータの転送を行うと実行性能が大きく落ちることになる。そのため、データ転送量、計算量、演算性能などの情報を基に CPU と GPU がうまく協調して計算を行う必要がある。

さらに、Intel や AMD の CPU の場合、アセンブラレベルで SSE 命令を駆使するような高度な最適化を行わない限り、単精度実数と倍精度実数の演算性能にほとんど違いはないが、GPU では機種毎に大きな違いがあることを考慮しておく必要がある。

●アプリケーションへの適用評価：本研究では、粒子間熱伝導問題を題材として、単精度実数と倍精度実数を利用した場合の CPU と GPU の実行性能の違い、及び、密行列と疎行列のそれぞれの格納形式での実行性能の違いを複数の GPU を用いて評価を行った。この評価結果は、今後、自動チューニングを行う際に必要となる実行性能値の予測をより正確に行うことに繋がるものである。

(6) 研究分担者：伊藤祥司

●アプリケーションへの適用評価：自動チューニング(AT)機能付きソフトウェアの中で、AT 機能付き数値計算ライブラリの求解アルゴリズム、特に疎行列を係数行列に持つ線形方程式に対する反復解法の前処理付きアルゴリズムの設計に関する研究を実施した。アルゴリズム記述上の演算量、所要メモリサイズ、求解性能の面から有用性の高い“前処理付き BiCGStab(PBiCGStab)”について分析し、従来から用いられているアルゴリズムの前処理変換では数理構造上の非合理的な問題点があることを指摘し、それを改善したアルゴリズムを提案した。

BiCGStab と併用する前処理演算を色々取り替えながら、50 種類近くのテスト問題に対し、体系的な性能評価を実施し、数値的にも本提案が有効であることを示した。

当アルゴリズムは、自動チューニング機能付き数値計算ライブラリ Xabclib にも実装されている。

(7) 研究分担者：岩下武史

●アプリケーションへの適用評価：研究分担者である岩下は研究代表者、他の研究分担

者と連携し、以下のような研究を行った。高周波電磁場解析の分野で最も一般的な解析手法である 3 次元 FDTD 法において、キャッシュヒット率を向上させる時空間タイリングによる実装法を考案した。さらに、同手法では計算において冗長な部分が存在したが、冗長な計算を排除する新たな実装法を考案し、さらなる性能改善を行った。

次に、これらの 3 次元 FDTD 法における重要な性能パラメータであるタイルサイズ（形状）、タイル内時間ステップ数を自動チューニングする手順を設計し、岩下が別途参画している研究プロジェクトで開発されたジョブ管理スクリプト言語 Xcrypt による実装、性能評価を行った。

(8) 研究分担者：滝沢寛之

●ABCLibScript の高度化: 本研究では、CPU と GPU を混載する複合型計算システム向け高水準プログラミング環境として研究開発されてきた Stream Programming with Runtime Auto-Tuning (SPRAT) を対象とし、SPRAT コンパイラから自動生成される CUDA プログラムのスレッドブロックサイズを実行時に自動チューニングする機構を開発した。

同機構では、小さなデータサイズでプロファイリングした結果に基づいて実運用での性能を予測することにより、自動チューニングに要するコスト軽減を実現している。また、SPRAT からシームレスに呼び出し可能な、実行時自動プロセッサ選択機能つきライブラリの試作も行った。それらのライブラリは SPRAT のプロセッサ選択機能と連携することにより、データサイズに依存して CPU と GPU の性能の高い方を実行時に自動選択できる。これらの成果は、GPU の存在を意識せずにその性能を利用するための高水準プログラミング環境の実現に向けて必要不可欠な機能を実現するものである。

さらに、研究期間の後半には、次世代標準プログラミング環境として整備が進んでいる OpenCL の自動チューニングのための研究も行った。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 13 件)

- 1) 伊藤祥司、片桐孝洋、櫻井隆雄、猪貝光祥、大島聡史、黒田久泰、直野健、BiCGStab 法の前処理付きアルゴリズムに対する改善、情報処理学会論文誌: ACS、査読有、Vol. 5、No. 3、2012、pp. 11-21

- 2) 伊藤祥司、片桐孝洋、櫻井隆雄、猪貝光祥、大島聡史、黒田久泰、直野健、BiCGStab 法の前処理付きアルゴリズムに対する改善、2012 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、査読有、2012、pp. 117-126
- 3) 南武志、高橋康人、岩下武史、中島浩、キャッシュメモリを考慮した 3 次元 FDTD カーネルの性能改善、情報処理学会論文誌: ACS、査読有、Vol. 4、No. 2、2012、pp. 70-83
- 4) 佐藤功人、小松一彦、滝沢寛之、小林広明、OpenCL におけるタスク並列化支援のための実行時依存関係解析手法、情報処理学会論文誌: ACS、査読有、Vol. 5、No. 1、2012
- 5) 片桐孝洋、佐藤雅彦、疎行列-ベクトル積におけるブロック化 BSS 法と高スレッド並列環境での性能評価、査読有、情報処理学会論文誌: ACS、Vol. 4、No. 3、2011、pp. 1-8
- 6) Katsuto Sato、Kazuhiko Komatsu、Hiroyuki Takizawa、and Hiroaki Kobayashi、A History-based Performance Prediction Model with Pro-file Data Classification for Automatic Task Allocation in Heterogeneous Computing Systems、In The 9th IEEE International Symposium on Parallel and Distributed Processing with Applications、査読有、2011
- 7) 片桐孝洋、ペタフロップス環境における小規模行列用対称密行列固有値ソルバに向けて一逆変換の改良、情報処理学会論文誌: ACS、査読有、Vol. 3、No. 2、2010、pp. 1-8

- 8) Takahiro Katagiri, Challenges of Run-time Auto-tuning for Sparse Iterative Solvers, Fifth International Workshop on Automatic Performance Tuning (iWAPT2010), Proceedings of VECPAR' 2010, 査読有、2010
- 9) 伊藤祥司, 杉原正顯, 姫野龍太郎, クリロフ部分空間法に対する前処理方式と収束判定について, 情報処理学会論文誌: ACS, 査読有、Vol. 3, No. 2, 2010, pp. 9-19
- 10) Imamura, T., Yamada, S. and Machida, M., Narrow-band reduction approach of a DRSM eigensolver on a multicore-based cluster system, International Parallel Computing conference 2009 (ParCo2009), Parallel Computing: From Multicores and GPU's to Petascale, 査読有、Vol. 19, 2010, pp. 91-98
- 11) Reiji Suda, Methods of Parallel Experimental Design of Online Automatic Tuning and their Application to Parallel Sparse Matrix Data Structure", Proc. iWAPT2010 (VECPAR' 10), 査読有、2010
- 12) Kazuhiko Komatsu, Katsuto Sato, Yusuke Arai, Kentaro Koyama, Hiroyuki Takizawa and Hiroaki Kobayashi, Evaluating Performance and Portability of OpenCL Programs, Proc. iWAPT2012 (VECPAR' 10), 査読有、2010
- 13) 今村俊幸, 超並列環境向きの固有値計算アルゴリズムと自動チューニング, 応用数理, 査読有、Vol. 20, No. 3, 2010
- 14) 片桐孝洋, 尾崎克久, 萩田武史, 大石進一, 高精度行列 - 行列積アルゴリズムのスレッド並列化と ABCLibScript への機能実装, 第 133 回 HPC 研究会, 情報処理学会研究報告 HPC-133, 査読無、2012
- 15) 近藤大貴, 吉田剛啓, 田村遼也, 今村俊幸, 自動チューニングによる通信最適化を施した固有値ソルバの開発について, 情報処理学会研究報告 HPC-133, 査読無、2012
- 16) 片桐孝洋, 疎行列反復解法ライブラリにおける自動チューニング機能の開発, 京都大学数理解析研究所, 研究集会, 科学技術計算アルゴリズムの数理的基盤と展開 (代表者: 大石進一 (早稲田大学)), 若手セッション, 京都大学数理解析研究所講究録 1733, 査読無、2011, pp. 40-46
- 17) Takahiro Katagiri, Development of C Language Version of ABCLibScript - The Impact To Auto-tuning Software, SIAM Conference on Computational Science and Engineering (CSE11), MS44, Advanced Performance Auto-tuning for Computer Systems and Numerical Libraries - Part I of II, 査読無、2011年3月1日, 米国ネバダ州レノ市
- 18) 片桐孝洋, 大島聡史, 平澤将一, 本多弘樹, HxABCLibScript: 非均質計算機向け自動チューニング記述言語拡張, 情報処理学会研究報告 HPC-129, 査読無、2011
- 19) Takahiro Katagiri, Towards Auto-tuning Description Language to Heterogeneous Computing Environment, Proceedings of Fifth International Workshop on High-level Parallel Programming and Applications (HLPP 2011) (Affiliated to ICFP 2011, Sponsored by ACM SIGPLAN), 査読有、

[学会発表] (計 56 件)

2011, pp.1-2

DOI:10.1145/2034751.2034753

- 20) Shoji Itoh, Takahiro Katagiri, Takao Sakurai, Mitsuyoshi Igai, Satoshi Ohshima, Hisayasu Kuroda, Ken Naono, Kengo Nakajima, An improvement in preconditioned BiCGStab method、2011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、査読有、2011、pp.62-62 (Poster)
- 21) 南武志、岩下武史、中島浩、3次元FDTDカーネルのキャッシュメモリを考慮した性能改善の実装と性能評価、情報処理学会研究報告HPC-130、査読無、2011
- 22) 今村俊幸、山田進、町田昌彦、ポストペタスケール時代の超並列固有値ソルバー開発、日本応用数学会 2011 年度年会講演予稿集、査読無、2011
- 23) 片山祐、黒田久泰、GPU を利用した線形方程式の反復解法の実装と性能評価、情報処理学会第 73 回全国大会講演論文集 (1)、査読無、2011、pp.203-204

[図書] (計 3 件)

- 24) Takahiro Katagiri、Springer、ABCLibScript: A Computer Language for Automatic Performance Tuning、Software Automatic Tuning, Naono, K.; Teranishi, K.; Cavazos, J.; Suda, R. (Eds.)、2010、pp.295-314
- 25) Shoji Ito, Masaaki Sugihara、Springer、Systematic performance evaluation of linear solvers using quality control techniques、Software Automatic Tuning, Naono, K.; Teranishi, K.; Cavazos, J.; Suda, R. (Eds.)、2010、pp.135-152
- 26) Katsuto Sato、Hiroyuki Takizawa、Kazuhiko Komatsu、Hiroaki Kobayashi、Automatic Tuning of CUDA Execution

Parameters for Stencil Processing、Software Automatic Tuning, Naono, K.; Teranishi, K.; Cavazos, J.; Suda, R. (Eds.)、2010、pp.209-228

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

Xabclib ホームページ:

<http://www.abc-lib.org/>

6. 研究組織

(1) 研究代表者

片桐 孝洋 (KATAGIRI TAKAHIRO)
東京大学・情報基盤センター・准教授
研究者番号: 40345434

(2) 研究分担者

今村 俊幸 (IMAMURA TOSHIYUKI)
電気通信大学・情報理工学部・准教授
研究者番号: 60361838

須田 礼仁 (SUDA REIJI)
東京大学・情報理工学系研究科・教授
研究者番号: 40251392

黒田 久泰 (KURODA HISAYASU)
愛媛大学・理工学研究科・准教授
研究者番号: 60323507

伊藤 祥司 (ITOH SHOJI)
東京大学・情報基盤センター・特任准教授
研究者番号: 70333482

岩下 武史 (IWASHITA TAKESHI)
京都大学・学術情報メディアセンター・
准教授
研究者番号: 30324685

滝沢 寛之 (TAKIZAWA HIROYUKI)
東北大学・情報科学研究科・准教授
研究者番号: 70323996

(3) 連携研究者: なし