

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 27 日現在

機関番号：11301

研究種目：基盤研究(B) (一般)

研究期間：2013～2015

課題番号：25280041

研究課題名(和文) シナジー効果を加速するソフトウェアとハードウェアの協調設計基盤

研究課題名(英文) Infrastructures for accelerating the synergy effect of software-hardware co-design

研究代表者

滝沢 寛之 (Hiroyuki, Takizawa)

東北大学・情報科学研究科・准教授

研究者番号：70323996

交付決定額(研究期間全体)：(直接経費) 13,900,000円

研究成果の概要(和文)：標準プログラミング環境としてOpenCLを想定し、より多様なアクセラレータアーキテクチャを利用するために足りない機能を指摘し、OpenCLの拡張を検討した。また、OpenCLはハードウェア記述にも使われるようになってきたが、そのカーネル部分を記述するための言語としてOpenCL C言語が必ずしも効率的とは限らない点を問題視し、画像処理や高性能計算で多用される処理を記述するための高生産性言語を設計、実装した。さらには、アクセラレータごとに適切な値の異なるパラメータを自動設定する手法を提案し、その実装と評価を行った。

研究成果の概要(英文)：Assuming OpenCL as a standard environment for accelerator programming, we have pointed out some missing features for supporting more various accelerator architectures, and proposed OpenCL extensions. Although OpenCL has gradually become to be used for hardware description, OpenCL C is not necessarily appropriate for describing OpenCL kernels. Hence, we have designed and implemented high productivity languages for typical computations in the fields of image processing and high performance computing. In addition, we have proposed an automatic tuning method for performance parameters, which need to be adjusted for individual accelerators. The proposed method has been implemented for evaluating its performance impacts.

研究分野：高性能計算

キーワード：アクセラレータ GPU FPGA

1. 研究開始当初の背景

汎用プロセッサと特定用途向けのアクセラレータとを混載する計算システムでは、それを制御するプログラムの構成に多くの類似点が見られる。多くのプログラムでは、汎用プロセッサがアクセラレータを初期化し、アクセラレータ上で行われる処理(カーネル)を設定し、アクセラレータへデータを転送し、アクセラレータ上でカーネルを実行し、カーネル実行完了を待って、アクセラレータからデータを受け取るといった一連の手順が記述される。このように汎用プロセッサからアクセラレータを制御する手順には明らかな共通性があるものの、それを記述するための共通基盤となるプログラミングモデルおよびその開発環境は確立されていない。

2. 研究の目的

本研究の目的は、特定用途向けのアクセラレータのハードウェアと、それを利用するソフトウェアの開発との共通インタフェースを定義し、今後の計算機工学のあらゆる階層、あらゆる場面で強く求められるハードウェアとソフトウェアの協調設計を支援することである。独自のハードウェアを開発する際、その実用的な性能および機能の評価のためにはアプリケーションプログラムを記述しなければならない。独自のプログラミングモデルの設計からソフトウェアスタックの構築までの全てを、ハードウェア設計のたびに行うのは困難である。またアプリケーションプログラム自体も高い価値を持つ資産であり、ハードウェアが変わるたびに全面的に開発しなおすことは通常許されない。このためハードウェアとソフトウェアとの境界となるインタフェースを多様なアクセラレータ間で共通化できる形式で定義することで、両者の協調設計を支援することを考える。

3. 研究の方法

本研究では、特定用途向けのアクセラレータのハードウェアと、そのアクセラレータ上での処理を記述するためのプログラミング環境を協調して設計することを考える。アクセラレータを利用するためのプログラミング環境は、汎用プロセッサからのアクセラレータの制御を記述するためのアクセラレータ制御用言語(あるいは API)と、アクセラレータ上での処理であるカーネルを記述するためのカーネル記述用言語から構成されることができると考えることができる。そのようなプログラミング環境の設計において、アクセラレータ制御用言語をアプリケーション間、アクセラレータ間で極力共通化する。一方、カーネル記述言語については、各アクセラレータに適切なものを切替えて使えるように設計する。このような構成に分けて考え、それぞれのプログラミングモデルを適切に設計することによって、アプリケーションプログラムの大部分をアクセラレータ間で共通化でき、アク

セラレータの機能を最大限に活用しなければならない重要なカーネル部分のみを適切なカーネル記述言語を選択して柔軟に記述できる。その結果、アプリケーションプログラムの共通化と各種アクセラレータ向けの特殊化とを両立するプログラミング環境を構築できることを明らかにする。

4. 研究成果

平成 25 年度には、アクセラレータ制御用言語として OpenCL を想定し、カーネル記述言語の設計を行った。ストリーム処理専用プロセッサを対象としてカーネル記述言語を設計し、FPGA 上に同プロセッサを実装する際のハードウェア量削減に効果があることが示された。

また、より多様なアクセラレータに対応するために、現在の OpenCL では足りない機能を指摘し、その機能を補う拡張を提案した。

その一方で、特定用途向けのアクセラレータのハードウェア設計も進め、ストリーム処理専用プロセッサやマルチメディア処理専用プロセッサの設計と評価も行った。

さらには、画像処理の高速化におけるアクセラレータの有効性を示すために、現在最も広く使われている GPU を用いて画像処理プログラムの実装を行い、その性能向上を評価した。

平成 26 年度には、高性能計算から画像処理まで規模の異なる様々なシステム構成で利用可能なプログラミング環境を実現するために、ファイルシステムを含むメモリ階層全体を考慮してデータ管理やデータ移動を設計できる必要があることを考え、特にファイル内のデータとアクセラレータのメモリ間のデータ転送の効率化を目的とし、OpenCL と MPI とファイル I/O 関数との組み合わせによって効率的なデータ転送を実現する研究を行った。OpenCL を拡張し、そのイベント管理機構をノード間通信やファイルアクセスにも利用することにより、データ間の依存関係を意識しつつ適切な組み合わせでデータ転送を実現できることが示された。

また、ステンシル計算を主な対象と考え、それを高効率に実行するハードウェアの設計のためのプログラミング環境を検討した。ステンシル計算のように規則性の高い処理のハードウェアを設計する場合、その構成要素には類似性が見られる。しかしながら通常のハードウェア記述言語では低い抽象度での記述が求められ、開発の生産性が低くなってしまふ。この問題に対してステンシル計算用アクセラレータを高い抽象度で記述するための言語を設計し、その言語処理系を開発した。その結果として、アクセラレータ設計の多くの部分を自動化、半自動化することができた。

さらには並列性やメモリアクセスパターンに基づいて処理を分類し、それぞれの対応するプログラムの構成要素を事前に用意す

ることで、多様なアクセラレータで実行されるカーネルを高い抽象度で記述することを目指し、そのためのドメイン特化型言語の設計も行った。プログラムの部品をソフトウェアとして事前準備するのと同様に、ハードウェア側でも必要な構成要素を準備することで、それらの組み合わせによって各アプリケーションに対して適切なアクセラレータを構成する可能性を検討した。

平成 27 年度には、ステンシル計算を対象として性能への影響のあるパラメータを自動設定する研究を行い、そのためのプログラミング環境を試作した。実システム上での性能解析によらずに性能パラメータを設定することにより、性能パラメータ数が多い場合にも実用的な時間内にパラメータの値を設定できる可能性を示した。

FPGA を使ったストリーム計算ハードウェアに関しても、時間方向と空間方向の並列性に関する設計空間の探索を可能とする専用言語を提案することができた。

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 19 件)

K. Sano, “DSL-based Design Space Exploration for Temporal and Spatial Parallelism of Custom Stream Computing,” 2nd International Symposium on FPGAs for Software Programmers, pp. 29–34, 2015, 査読有。
M.A. Amrizal, S. Hirasawa, H. Takizawa, and H. Kobayashi, “Automatic Parameter Tuning of Hierarchical Checkpointing,” Lecture Notes in Computer Sciences, vol.8969, pp. 298–309, 2015, 査読有。

DOI:0.1007/978-3-391-17353-5_25.

H. Takizawa, S. Hirasawa, M. Sugawara, I. Gelado, H. Kobayashi, and W-M.W. Hwu, “Optimized Data Transfer Based on the OpenCL Event Management Mechanism,” Scientific Programming Journal, ID 576498, pp. 1–16, 2015, 査読有。DOI: 10.1155/2015/576498.

S. Hirasawa, H. Takizawa, and H. Kobayashi, “A Light-weight Rollback Mechanism for Testing Kernel Variants in Auto-tuning,” IEICE Transactions on Information and Systems, Vol. E98-D(12), pp. 2178–2186, 2015, 査読有。

T. Yamada, S. Hirasawa, H. Takizawa, and H. Kobayashi, “A Case Study of User-defined Code Transformations for Data Layout Optimizations,” 3rd International Symposium on Computing and Networking, International

Workshop on Legacy HPC Application Migration (LHAM2015), pp. 535–541, 2015, 査読有。

S. Hirasawa, H. Takizawa, and H. Kobayashi, “A Verification Framework for Streamlining Empirical Auto-tuning,” 3rd International Symposium on Computing and Networking, International Workshop on Legacy HPC Application Migration (LHAM2015), pp. 515–520, 2015, 査読有。

C. Wang, S. Hirasawa, H. Takizawa, and H. Kobayashi, “Combining Code Refactoring and Auto-tuning to Improve Performance Portability of High-Performance Computing Applications,” 6th International Conference on Computational Logics, Algebras, Programming, Tools, and Benchmarking, 20–26, 2015, 査読有。

C. Wang, S. Hirasawa, H. Takizawa, and H. Kobayashi, “Identification and Elimination of Platform-Specific Code Smells in High Performance Computing Applications,” International Journal of Networking and Computing, vol. 5, 180–199, 2015, 査読有。

K. Sano, H. Suzuki, R. Ito, T. Ueno, S. Yamamoto, “Stream Processor Generator for HPC to Embedded Applications on FPGA-based System Platform,” 1st International Workshop on FPGAs for Software Programmers, 43–48, 2014, 査読有。

H. Takizawa, S. Hirasawa, Y. Hayashi, R. Egawa, and H. Kobayashi, “Xeolver: An XML-based code transformation framework for supporting HPC application migration,” IEEE International Conference on High Performance Computing, 1–11 (USB), 2014, 査読有。

X. Xiong, S. Hirasawa, H. Takizawa, and H. Kobayashi, “An Approach to Customization of Compiler Directives for Application-specific Code Transformations,” 8th International Conference on Embedded Multicore/manycore SoCs, 99–106, 2014, 査読有。

C. Wang, S. Hirasawa, H. Takizawa, and H. Kobayashi, “Platform-Specific Code Smell Alert System for High Performance Computing Applications,” 16th Workshop on Advances on Parallel and Distributed Processing Symposium, 652–661, 2014, 査読有。

K. Sano, Y. Kono, H. Suzuki, R. Chiba, R. Ito, T. Ueno, K. Koizumi, and S.

Yamamoto, "Efficient Custom Computing of Fully-Streamed Lattice Boltzmann Method on Tightly-Coupled FPGA Cluster," ACM SIGARCH Computer Architecture News, Vol. 41, pp. 47-52, 2014, 査読有.

K. Sano, R. Ito, H. Suzuki, Y. Kono, "Parallel and Scalable Custom Computing for Real-Time Field Simulation on a Cluster Node with Four Tightly-Coupled FPGAs," 23rd International Conference on Field-Programmable Logic and Applications, pp.1, 2014, 査読有.

K. Sano, R. Chiba, T. Ueno, H. Suzuki, R. Ito, and S. Yamamoto, "FPGA-based Custom Computing Architecture for Large-Scale Fluid Simulation with Building Cube Method," ACM SIGARCH Computer Architecture New, Vol.42, pp.45-50, 2014, 査読有.

伊藤涼, 鈴木隼人, 上野知洋, 佐野健太郎, 山本悟, "ストリームプロセッサ自動生成コンパイラの開発と格子ボルツマン法専用計算コアへの適用," 第28回数値流体力学シンポジウム講演論文集, E07-1, 1-4, 2014, 査読無.

Y. Gao, N. Shoji, R. Egawa, H. Takizawa, and H. Kobayashi, "Design and evaluation of a media-oriented vector processor with a multi-banked cache memory," 11th IEEE/ACM Symposium on Embedded Systems for Real-Time Multimedia, pp. 78-87, 2013, 査読有.
M. Miura, K. Fudano, K. Ito, T. Aoki, H. Takizawa, and H. Kobayashi, "Performance evaluation of phase-based correspondence matching on GPUs," SPIE Optical Engineering+ Applications. International Society for Optics and Photonics, volume 8856, pp. 885614-1-885614-9, 2013, 査読有.
M. Miura, K. Fudano, K. Ito, T. Aoki, H. Takizawa, and H. Kobayashi, "GPU implementation of phase-based stereo correspondence and its application," International Conference on Image Processing, pp. 1697-1700, 2012, 査読有.

[学会発表](計23件)

H. Takizawa, S. Hirasawa, and H. Kobayashi, "Performance Engineering of HPC Application Based on Pattern Matching," 23rd Workshop on Sustained Simulation Performance, 仙台, 2016年3月17日(招待講演).

H. Takizawa, S. Hirasawa, and H. Kobayashi, "Data Layout Optimization Using User-Defined Code

Transformations," 2016 Conference on Advanced Topics and Auto Tuning in High-Performance Scientific Computing, Taipei, 2016年2月19日(招待講演).

長洲航平, 佐野健太郎, 河野郁也, 中里直人, "ストリーム計算ハードウェアコンパイラ SPGen を用いた FPGA 津波シミュレータの開発," 信学会リコンフィギャラブルシステム研究会信学技法, vol.115, pp. 131--136, 慶大日吉キャンパス, 2016年1月20日.

上野友也, 田中大智, 佐野健太郎, 山本悟, "FPGA による流体専用並列計算ハードウェアの試作実装と性能評価," 第29回数値流体力学シンポジウム講演論文集, pp. 1-2, 福岡県春日市, 2015年12月17日.

H. Takizawa, S. Hirasawa, K. Komatsu, R. Egawa, and H. Kobayashi, "Expressing system-awareness as code transformations for performance portability across diverse HPC systems," Workshop Among HPC Architectures for Scientific Applications, Austin, USA, 2015年11月15日.

佐野健太郎, 伊藤涼, 菅原啓介, 山本悟, "階層のモジュール設計を可能とするストリーム計算コア高位合成コンパイラ," 信学会リコンフィギャラブルシステム研究会信学技法, vol.114, pp. 79-84, 京都, 2015年6月20日.

平澤将一, 滝沢寛之, 小林広明, "Xevolver による実アプリケーションの性能と保守性の両立," 計算工学研究会論文集, Vol. 20, E-2-3, 筑波, 2015年6月8日.

H. Takizawa, "A Framework for Separation of Concerns Between Application Requirements and System Requirements," 2015 SIAM Conference on Computational Science and Engineering, Salt Lake City, USA, 2015年3月18日.

H. Takizawa, "Autotuning with User-defined Code Transformations," 2015 Conference on Advanced Topics and Auto Tuning in High Performance Scientific Computing, Taipei, 2015年2月28日(招待講演).

H. Takizawa, "Xevolver; an extensible framework for user-defined code transformation," 20th workshop on sustained simulation performance, Stuttgart, Germany, 2014年12月15日(招待講演).

H. Takizawa, "Evolutionary Adaptation of HPC Applications to Revolutionary System Changes," 29th

- International Supercomputing Conference, Leipzig, Germany, 2014 年 6 月 23 日(招待講演).
H. Takizawa, "An Extensible Programming Framework for Custom Code Transformations," 2014 Conference on Advanced Topics and Auto Tuning in High Performance Scientific Computing, Taipei, 2014 年 3 月 15 日(招待講演).
 平井亮太, 平澤将一, 滝沢寛之, 小林広明, "アクセラレータのためのプログラム最適化とその性能評価," 第 12 回情報シナジー研究会, 東北大学, 2014 年 2 月 24 日.
 伊藤涼, 鈴木隼人, 千葉諒太郎, 佐野健太郎, 山本悟, "ストリーム計算のための高位合成コンパイラの設計と実装," 信学会リコンフィギャラブルシステム研究会信学技法, 慶應義塾大学日吉キャンパス, 2014 年 1 月 28 日.
 千葉諒太郎, 鈴木隼人, 伊藤涼, 佐野健太郎, 山本悟, "フラクショナルステップ法に基づく非圧縮性流体計算専用ストリームプロセッサの設計と評価," 第 27 回数値流体力学シンポジウム, 名古屋, 2013 年 12 月 18 日.
H. Takizawa, S. Hirasawa, and H. Kobayashi, "Xevolver: an XML-based Programming Framework for Software Evolution," Poster Presentation at Supercomputing Conference(SC13), Denver, USA, 2013 年 11 月 19 日.
K. Sano, "FPGA-based Scalable Custom Computing Accelerator for Computational Fluid Dynamics Based on Lattice Boltzmann Method," 18th Workshop on Sustainable Simulation Performance, Stuttgart, Germany, 2013 年 10 月 29 日.
 千葉諒太郎, 鈴木隼人, 伊藤涼, 佐野健太郎, 山本悟, "フラクショナルステップ法に基づく非圧縮性流体計算専用ストリームプロセッサの設計と評価," 信学会リコンフィギャラブルシステム研究会信学技法, 北陸先端科学技術大学院大学, 2013 年 9 月 18 日.
滝沢寛之, 平澤将一, 小松一彦, 小林広明, "OpenACC における性能チューニングとその効果," 日本応用数理学会 2013 年度年会, 福岡, 2013 年 9 月 10 日.
 X. Xiao, S. Hirasawa, H. Takizawa, and H. Kobayashi, "A Case Study of Performance Tuning with the POET Framework," 平成 25 年度電気関係学会東北支部連合大会, 会津大学, 2013 年 8 月 22 日.
- 21 C. Wang, S. Hirasawa, H. Takizawa, and H. Kobayashi, "Code Refactoring for High Performance Computing Applications," 平成 25 年度電気関係学会東北支部連合大会, 会津大学, 2013 年 8 月 22 日.
- 22 山尾創輔, 三浦衛, 酒井修二, 伊藤康二, 青木孝文, "デジタルカメラの移動撮影に基づく 3 次元形状計測システム," 第 16 回画像の認識・理解のシンポジウム, 東京, 2013 年 7 月 30 日.
- 23 三浦衛, 酒井修二, 石井純平, 伊藤康二, 青木孝文, "デジタルカメラを用いた簡便で高精度なステレオビジョンシステム," 第 19 回画像センシングシンポジウム, 横浜, 2013 年 6 月 13 日 ~ 14 日.
6. 研究組織
- (1)研究代表者
 滝沢 寛之 (Hiroyuki Takizawa)
 東北大学・情報科学研究科・准教授
 研究者番号: 70323996
- (2)研究分担者
 小林 広明 (Hiroaki Kobayashi)
 東北大学・サイバーサイエンスセンター・教授
 研究者番号: 40205480
- (3)研究分担者
 青木 孝文 (Takafumi Aoki)
 東北大学・情報科学研究科・教授
 研究者番号: 80241529
- (4)研究分担者
 佐野 健太郎 (Kentaro Sano)
 東北大学・情報科学研究科・准教授
 研究者番号: 00323048
- (5)連携研究者
 江川 隆輔 (Ryusuke Egawa)
 東北大学・サイバーサイエンスセンター・准教授
 研究者番号: 80374990
- (6)連携研究者
 多田 十兵衛 (Jube Tada)
 山形大学・理工学研究科・助教
 研究者番号: 30361273
- (7)連携研究者
 伊藤 康一 (Koichi Ito)
 東北大学・情報科学研究科・助教
 研究者番号: 70400299