

**科学研究費助成事業 研究成果報告書**

平成 28 年 6 月 1 日現在

機関番号：11301

研究種目：基盤研究(B) (一般)

研究期間：2012～2015

課題番号：24300013

研究課題名(和文) 組み込みHPC向けカスタムアクセラレータ統合型ヘテロジニアス計算基盤技術の開拓

研究課題名(英文) Development of Heterogeneous-Computing Platform with Custom Accelerators for Embedded HPC Applications

研究代表者

張山 昌論 (Hariyama, Masanori)

東北大学・情報科学研究科・准教授

研究者番号：10292260

交付決定額(研究期間全体)：(直接経費) 12,100,000円

研究成果の概要(和文)：本研究では、従来のCPUコア・GPUアクセラレータコア構成に加えて、FPGAをカスタムアクセラレータとして用いる構成を対象とし、データ転送ボトルネックを解消するアーキテクチャ、および高水準ユーザーコードに対して性能を最大限引き出す手法を研究する。近年のHPC計算は様々なアルゴリズムは複雑化しておりCPUや固定的なデータバスをもつGPUだけではエネルギー効率のよい処理が難しい。そこでFPGAにより「応用にカスタマイズされた計算機」を構築し、CPU・GPUのプログラムと共に、広義のプログラムであるFPGA回路構成情報を自動生成するシステムコンパイラを開発する。

研究成果の概要(英文)：This research is aiming at design technologies for a heterogeneous computing platform consisting of CPU cores, GPUs, and FPGAs. FPGAs allows us to build the application-specific accelerators since it's datapath can be reconfigured by programming after fabrication. One major difficulty in designing the heterogeneous computing platform is to find task allocation maximizing the performance and to design the FPGA datapath as easily as possible. In order to solve this problem, we develop high-level synthesis technologies for the heterogeneous computing platform.

研究分野：リコンフィギャラブルコンピューティング

キーワード：リコンフィギャラブルコンピューティング FPGA カスタムアクセラレータ ビッグデータ処理 高性能計算

### 1. 研究開始当初の背景

高性能計算(HPC)応用が天文、流体、生命科学、環境工学、金融、医療分野など拡大し、従来高性能計算の対象であった同一の演算を高並列に行う処理だけでなく、複雑な演算を含む処理、部分的に規則性が損なわれる処理、固定小数点演算を並列に行う処理など HPC 計算が多様化している。

多様な計算に対応できる最も直接的な構成は CPU を用いたホモジニアス構成であるが、エネルギー効率が低いという問題がある。そこで GPU をアクセラレータとして用いた構成が注目されているが、GPU 向きの処理はデータ並列かつ規則的な数値演算処理に限定されており、前述の HPC 分野の多様化により GPU では高効率を実現できない次のような処理が多く存在する。

- 1) 条件分岐などの制御フローが不規則・複雑な処理
- 2) 低精度演算で十分な処理/倍精度以上の高精度演算処理: GPU は単精度/倍精度にデータパスが固定されているため、それ以外での語長ではエネルギー効率が低下 (例: 遺伝子解析における塩基配列照合)
- 3) 演算間の依存関係が複雑な処理: GPU では共有メモリを介したデータの受け渡しにより性能が低下。

本研究では、CPU・GPUに加え、FPGA(プログラムにより任意の処理アーキテクチャを構成できる再構成可能集積回路)をアクセラレータとして用いて計算ノードを構成する。FPGAでは、プログラム変更により、応用毎にデータパスを最適化できる。最近のFPGAでは、回路規模の飛躍的な増大、最新のメモリインタフェース、メモリバンド幅の増大、GPU並みの高速浮動小数点演算ユニットなど、FPGAベンダーもHPC分野を重視している。

ヘテロジニアス構成の最も大きな問題は、性能を引き出すプログラミングが一般ユーザーには難しい点である。これは、アクセラレータの演算器数・メモリ容量・演算の種類・コア間のデータ転送量などの多数の制約を考慮しなければならないためである。さらに、本研究で対象とするFPGAを用いる場合には、データパスの自由度が極めて高いためさらに最適化が難しくなる。

### 2. 研究の目的

本研究では、高速・低消費電力を両立できるヘテロジニアスアーキテクチャにおいて、従来のCPUコア・GPUアクセラレータコア構成に加えて、FPGAをカスタムアクセラレータとして用いる構成を対象とし、データ転送ボトルネックを解消するアーキテクチャ、および高水準ユーザーコードに対して性能を最大限引き出すソフトウェアを研究する。近

年のHPCアプリケーションは様々なアルゴリズムが組み合わされておりCPUや固定的なデータパスをもつGPUだけではエネルギー効率のよい処理が難しい。そこでFPGAにより「応用にカスタマイズされた計算機」を構築し、CPU・GPUのプログラムと共に、広義のプログラムであるFPGA回路構成情報を自動生成するシステムコンパイラを開発する。

### 3. 研究の方法

CPU・アクセラレータ(GPUやFPGA)から構成されるヘテロジニアス計算ノードを構築し、単一のノード内の各計算リソースに対するタスク分割手法、FPGAカスタムアクセラレータの最適データパス合成手法の開発・評価を行う。また、アクセラレータと外部メモリのバンド幅を拡大する手法として、メモリに圧縮されたデータを記憶し、演算時に伸張する手法を開発する。さらに、ノードを並列に結合したクラスタを対象とし、タスク分割の最適化手法、データ転送スケジュールの自動最適化手法を開発する。さらに、ノード間のデータ転送を高速化するために、FPGA内部に種々の応用に対応できるデータ圧縮・伸張モジュールを構築し、自動最適化に反映する。最後に電磁界解析、遺伝子解析などの応用で総合評価を行う。ハードウェアの全体構成は、多数の計算ノード、CPUを介した従来型ノード間結合、FPGA間高速結合網からなる。FPGA間の直接接続により、複数ノード上のFPGAを単一のノード上のように接続でき、大規模クラスタを効率よく構築できる。

研究方法の詳細を以下に示す。

- 1) **ヘテロジニアス計算ノードの構築:** CPUとアクセラレータ(GPU, FPGA)から構成される単一の計算ノードを構築する。応用に関する情報提供者から電磁界解析、生命科学(タンパク質の構造推定)、医療(臓器認識の画像処理など)などのユーザーコードを提供してもらい、それをヘテロジニアス計算ノードに、手でマッピングすることにより、タスク分割に有用なタスク情報、有用なアクセラレータアーキテクチャに関する知見を得る。
- 2) **アーキテクチャテンプレートデータベースの構築とデータパス最適化手法の確立:** 高性能計算アプリケーションにおいて有用となるアーキテクチャのテンプレートを集めたデータベースを構築する。このアーキテクチャテンプレートでは演算器数・語長、メモリ構成、それらを結合する相互結合網などパラメータ化されている。さらに、アーキテクチャテンプレートを用いて、演算器数・演算器語長・メモリサイ

ズなどを最適化するように、リソース制約での処理時間・消費電力最適化を達成するデータパス合成法を構築する。

- 3) **単一のヘテロジニアス計算ノードに対する評価:** 応用プログラムを単一のヘテロジニアス計算ノードに対してマッピングすることにより、タスク分割手法、スケジューリング手法などの評価を行い、必要に応じて改善を計画する。
- 4) **最適タスク分割手法の確立:** 処理の依存関係・並列性などの処理の性質に基づいて、アプリケーションをCPU, GPU, FPGA の異なる計算リソースに適するタスクに分割する。タスク分割の際には、全体として最適になるように、タスクの並列性、タスク間のデータ転送量、タスク内の演算の種類・語長、要求されるメモリ容量、予測性能・予測消費電力などを考慮してタスクを分割することが重要となる。
- 5) **複数ノード間の高速通信のためのFPGAによるデータ圧縮・伸張機構の開:** FPGA内にデータ圧縮/伸張モジュールを構築し、大量のデータを高速に転送できるようにする。典型的なHPC応用で表れるデータの性質に着目し、それらを効率よく転送できるデータ圧縮/伸張モジュールライブラリを構築し、応用に併せて最適なものを選択できるようにする。
- 6) **クラスタを用いた総合評価クラスタを用い、種々の高性能計算応用に対して性能評価を行う。**

#### 4. 研究成果

##### 1) DNA配列解析のためのヘテロジニアスカスタムアクセラレータの開発

CPUとFPGAを用いたヘテロジニアスノードを構築し、ビッグデータアクセラレータの例として、遺伝子配列解析のためのアクセラレータを開発した。近年、遺伝子配列解析のためには計算量を削減できる高度なアルゴリズムが様々提案されている。その中で、本開発では、BWT変換に基づくアルゴリズムを採用した。このアルゴリズムでは、遺伝子配列に対してBW変換と呼ばれる前処理を施すことにより、遺伝子の長さに依存しない検索処理が可能となる。CPUでの実装に適する部分とFPGAでの実装に適する部分を見出し、ヘテロジニアスな処理をすることにより計算時間を大幅に削減している。また、FPGA実装においては、メモリのアクセスパターンに着目し、データアクセスの局所性を最大限に活用できるメモリアーキテクチャを開発した。このような工夫により、従来のPCベースの処理と比較して、80倍程度の高速化2を達成

できることを明らかにした。

##### 2) 大規模グラフ解析処理アクセラレータの開発

ビッグデータ応用の重要な分野の一つである、大規模なグラフ処理に対するアクセラレータを開発した。例として、ネットワーク解析、関係解析、交通流シミュレーションなど様々な応用がある、最短経路問題を取り上げた。グラフ表現に適する「簡潔データ表現」、演算に必要なノードのみをFPGA内のローカルメモリで動的に管理する効率の良いメモリアーキテクチャとSIMD型データ並列処理アーキテクチャにより、PCベースの処理と比較して、60枚程度の高速化を達成できる見通しを得た。

##### 3) パーティクルフィルタプラットフォームの開発

パーティクルフィルタは、システムの動的状態推定手法である。応用は多岐にわたり、動画像中の物体追跡、ロボットの自己位置推定、データ同化シミュレーションなどが挙げられる。本開発では、パーティクルフィルタを並列処理する際に問題となるリサンプリング処理に対して、パイプライン処理に適した手法を提案し、OpenCL高位設計環境を用いて実装した。OpenCLではC言語を用いてFPGA回路を設計できるため、様々な応用に対して、システム方程式などを柔軟に変更できる。提案のパーティクルフィルタプラットフォームを、いくつかの応用に対して用いて、その開発容易性を明らかにした。また、性能としても、PCベースの処理と比較して、700倍の高速化を達成した。

##### 4) アーキテクチャテンプレートと最適データパス構築技術の開発

カスタムアクセラレータの開発事例をもとに、アーキテクチャテンプレートとして、電磁界解析のためのFDTD計算を一般化した、ステンシル計算のためのアーキテクチャテンプレートとその最適設計法を確立した。ステンシル計算は、FDTD計算のみならず、流体計算、熱力学、音響計算など様々な応用がある、代表的な高性能計算である。本研究では、外部メモリとFPGA(再構成可能集積回路)間のデータ転送ボトルネックの解消のために、複数のタイムステップをパイプラインで一括処理することにより、中間結果を内部メモリに蓄えるアーキテクチャモデルを提案した。提案アーキテクチャをOpenCL高位設計環境を用いて構成する手法を確立した。2次元FDTD処理をGPUとFPGAを用いた提案手法で実装し比較したところ、3倍以上の高性能化を達成した。

##### 5) OpenCL設計環境を用いた画像処理アクセラレータ設計

OpenCLを用いて、画像間の対応付けをサブピクセル精度で行うことができる位相限

定相関のためのアクセラレータを始めとする種々の画像処理アクセラレータを開発した。パイプライン処理と空間的並列処理を最適に融合する設計手法により、従来の GPU による実装では消費電力が高く、モバイル機器への組み込みは難しかった。本研究では、処理の性質に基づき処理を要素処理に分割し空間的並列アーキテクチャとパイプラインの組み合わせを最適化した。その結果、GPU と同等性能で、消費電力を 1/10 程度に減らすことを可能とした。位相限定相関に基づくアクセラレータは、ロボティクス、バイオメトリクス認証、高安全自動車など、種々の応用が期待されている。

#### 6)大規模データ検索のためのデータ圧縮に基づくハードウェア向き簡潔データ表現

遺伝子配列解析処理のアーキテクチャをテンプレートとして一般化し、ストレージから演算部へのデータ転送の問題を解決するために、検索に適してかつ少容量で記憶できる簡潔データ表現に着目した。従来は CPU での処理を前提としていたため、計算量の問題からデータ圧縮を積極的に導入することが難しかった。本研究では、FPGA での並列処理を活用して処理時間オーバーヘッドを削減できる簡潔データ表現とそのためのアーキテクチャを開発した。本技術は、計算ノード間でのデータ通信にも拡張を検討している。

#### 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 5 件)

Hasitha Muthumala Waidyasooriya, Daisuke Ono and Masanori Hariyama, "Hardware-Oriented Succinct- Data-Structure for Text Processing Based on Block-Size-Constrained Compression", International Journal of Computer Information Systems and Industrial Management Applications, Vol. 8, pp.1-11(2016)URL:

[http://www.mirlabs.net/ijcisim/regular\\_papers\\_2016/IJCISIM\\_1.pdf](http://www.mirlabs.net/ijcisim/regular_papers_2016/IJCISIM_1.pdf), 査読有り

Yasuhiro Takei, Hasitha Muthumala Waidyasooriya, Masanori Hariyama and Michitaka Kameyama, "Data-Transfer-Aware Design of an FPGA- Based Heterogeneous Multicore Platform with Custom Accelerators", IEICE Transaction on Fundamentals, Vol. E98-A, No. 12, pp.2658- 2669(2015) DOI: 10.1587/ transfun.E98.A.2658, 査読有り

Hasitha Muthumala Waidyasooriya, Masanori Hariyama, Yasuhiro Takei, and Michitaka Kameyama, "FDTD

Acceleration for Cylindrical Resonator Design Based on the Hybrid of Single and Double Precision Floating-Point Computation", Journal of Computational Engineering, pp.1-8(2014) DOI: 10.1155/ 2014/ 634269, 査読有り

Yoshiya Komatsu, Masanori Hariyama, Michitaka Kameyama, "Architecture of an Asynchronous FPGA for Handshake-Component- Based Design", IEICE Transaction on Information and Systems, Vol.E96- D, No.8, pp.1632-1644(2013) DOI: 10.1587/transfun.E96.A.2576, 査読有り

Yasuhiro Takei, Hasitha Muthumala Waidyasooriya, Masanori Hariyama, Michitaka Kameyama, "Evaluation of an FPGA- Based Heterogeneous Multicore Platform with SIMD/ MIMD Custom Accelerators", IEICE Transaction on Information and Systems, Vol. E96-A, No.12, pp.2576- 2586 (2013) DOI: 10.1587/transfun.E96.A.2576, 査読有り

##### [学会発表](計 17 件)

Hasitha Muthumala Waidyasooriya, Masanori Hariyama and Yasuo Ohtera, "FPGA Architecture for 3-D FDTD Acceleration Using OpenCL", Progress in Electro-magnetics Research Symposium(PIERS), 2016 年 08 月 08 日~ 2016 年 08 月 16 日, Shanghai, CHINA (To be presented)

Yasuhiro Takei, Masanori Hariyama, Michitaka Kameyama, "Evaluation of an FPGA- Based Shortest- Path-Search Accelerator", International Conference on Parallel and Distributed Processing Techniques and Applications(PDPTA)(国際学会), 2015 年 07 月 29 日, Las Vegas, USA

Shunsuke Tatsumi, Masanori Hariyama, Mamoru Miura, Koichi Ito, Takafumi Aoki, "OpenCL- based Design of an FPGA Accelerator for Phase- Based Correspondence Matching", International Conference on Parallel and Distributed Processing Techniques and Applications(PDPTA), 2015 年 07 月 28 日, Las Vegas, USA

Yasuhiro Takei, Hasitha Muthumala Waidyasooriya, Masanori Hariyama, Michitaka Kameyama, "FPGA- Oriented Design of an FDTD Accelerator Based on Overlapped Tiling", International Conference on Parallel and Distributed Processing Techniques and Applications(PDPTA), 2015 年 07 月 29 日, Las Vegas, USA

Hasitha Muthumala Waidyasooriya, Daisuke Ono, Masanori Hariyama, Michitaka Kameyama, "An FPGA Architecture for Text Search Using a Wavelet-Tree-Based Succinct-Data-Structure", International Conference on Parallel and Distributed Processing Techniques and Applications(PDPTA)(国際学会), 2015年07月28日, Las Vegas, USA

Hasitha Muthumala Waidyasooriya, Daisuke Ono and Masanori Hariyama, "Hardware-Oriented Succinct-Data-Structure based on the Block-Size-Constrained Compression", Soft computing and Pattern Recognition(SoCPaR), 2015年11月15日, Fukuoka, Japan

張山昌論, "大規模計算応用のためのカスタムアクセラレータの展望 ~ 応用・アーキテクチャ・回路 ~", 電子情報通信学会技術報告(信学技報)集積回路研究会(招待講演), 2014年06月11日, 仙台

Hasitha Muthumala Waidyasooriya, Masanori Hariyama and Michitaka Kameyama, "FPGA-Accelerator for DNA Sequence Alignment Based on an Efficient Data-Dependent Memory Access Scheme", International Symposium on Highly-Efficient Accelerators and Reconfigurable Technologies(HEART), 2014年06月10日 Sendai, Japan

Yasuhiro Takei, Hasitha Muthumala Waidyasooriya, Masanori Hariyama, "Design of an FPGA-Based FDTD Accelerator Using OpenCL", International Conference on Parallel and Distributed Processing Techniques and Applications(PDPTA) 2014年07月23日, Las Vegas, USA

Yasuhiro Takei, Hasitha Muthumala Waidyasooriya, Masanori Hariyama, "An SIMD Architecture for Shortest-Path Search and Its FPGA Implementation", International Conference on Parallel and Distributed Processing Techniques and Applications(PDPTA), 2014年07月24日, Las Vegas, USA

Daisuke Ono, Hasitha Muthumala Waidyasooriya, Masanori Hariyama, "Efficient Data Transfer Scheme Using Word-Pair-Encoding-Based Compression for Large-Scale Text-Data Processing, IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), 2014年11月19日 Okinawa, Japan

張山昌論, "人にやさしい応用を拓く「計算」技術", IEEE CASS Kansai

Chapter 技術講演会(招待講演), 2013年10月03日, 神戸大学, 神戸, 日本

武井康浩, ハシタムトゥマラウイシディスーリヤ, 張山昌論, 亀山充隆, "OpenCLを用いたFPGAベースFDTDアクセラレータの設計", 電子情報通信学会技術報告(信学技報) Electronic Simulation Technology(EST), 2013年10月24日, 東北大学, 仙台, 日本

Yoshiya KOMATSU, Masanori HARIYAMA and Michitaka KAMEYAMA, "An Area-Efficient Asynchronous FPGA Architecture for Handshake-Component-Based Design", Proc. International Conference on Engineering of Reconfigurable Systems and Algorithms (ERSA), 2013年07月24日, Las Vegas, USA

Hasitha Muthumala Waidyasooriya, Hirokazu Takahashi, Yasuhiro Takei, Masanori Hariyama and Michitaka Kameyama Reducing Floating-Point Error Based on Residue-Preservation and Its Evaluation on an FPGA, Proc. International Conference on Engineering of Reconfigurable Systems and Algorithms (ERSA) 2013年07月24日 Las Vegas, Nevada, USA

Yasuhiro Takei, Hasitha Muthumala Waidyasooriya, Masanori Hariyama and Michitaka Kameyama, "Heterogeneous Multicore Platform with Accelerator Templates and Its Implementation on an FPGA with Hardcore CPUs", Proc. International Conference on Engineering of Reconfigurable Systems and Algorithms (ERSA), 2013年07月24日, Las Vegas, USA

Hasitha Muthumala Waidyasooriya, Masanori Hariyama and Michitaka Kameyama, "Implementation of a Custom Hardware-Accelerator for Short-read Mapping Using Burrows-Wheeler Alignment", 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), 2013年07月04日, Osaka, Japan

## 6. 研究組織

### (1) 研究代表者

張山昌論 (Masanori Hariyama)  
東北大学・情報科学研究科・准教授  
研究者番号: 10292260

### (2) 連携研究者

ハシタムトゥマラウイシディスーリヤ (Hasitha Muthumala Waidyasooriya)  
東北大学・情報科学研究科・助教  
研究者番号: 60723533