

研究報告書

「高性能ストリームコンピューティング環境の構築」

研究タイプ: 通常型

研究期間: 平成 22 年 10 月～平成 26 年 3 月

研究者: 山際 伸一

1. 研究のねらい

マイクロマシン技術の発展により、加速度センサや、地磁気センサ、高精細な映像を取り込むイメージセンサなど、ヒトの行動に周囲の環境の影響を加味した知識の創生が可能になっている。センサからのデータは、途切れることなく、決められた時間間隔で生成され続ける。このようなデータストリームを扱う際、現代の情報機器の基盤となるフォンノイマン型の計算手法ではメモリにそのデータを一時的に保存して、プロセッサがそのデータを読み書きすることで処理を進めるため破綻をもたらす。より高精細なセンサ性能開発が進む一方で、処理能力との”いたちごっこ”は顕著に表れることが予想され新しいデータ処理方式を考案することが望まれる。データストリームを滞りなく高速処理する方法として、データパイプラインを専用ハードウェアで構成することが考えられる。専用ハードウェアを設計する際に、パイプラインを構成する演算器のイネーブルを出すタイミングは設計者により緻密に計算され、パイプラインが滞ることなく処理が継続されるタイミングにアサートされる必要がある。ハードウェア記述言語での設計が近年の主流であるが、演算器のイネーブルタイミングを決定するにはハードウェアの規模や動作スピードとの組み合わせを設計者が考慮しなければならず、不具合の誘発をおこしやすく、ハードウェア規模と動作性能が良くバランスした回路を構成するのは難しい。このようなデータパイプライン設計上の困難を排除するための新しいハードウェア設計技法の開発が望まれる。本研究は上記の問題点を解決するために、データストリームを入力とし、その各要素に対する計算を定義するプログラムが、順次対応する結果を出力するストリームコンピューティングの概念を取り入れたプログラミング環境を開発する。特に最近の性能向上が顕著なアクセラレータでのプログラム実行が可能な環境の構築を主軸とする。さらに、ストリームコンピューティング環境で実行されるプログラムの単位をハードウェア記述言語に変換するためのコンパイラを開発する。以上から、現代のプロセッサでは処理できない程度の大量なデータストリームに対する処理装置開発のための基盤を開発することをねらいとする。

2. 研究成果

(1) 概要

本研究課題における成果として、(1)アクセラレータプログラム向けの高生産性プログラミング環境の構築、(2)パイプライン実行の実行順序付けと並列性抽出による高速実行手法、そして、(3)コンパクトなパイプラインハードウェアを出力するコンパイラの開発、(4)ストリームデータ圧縮ハードウェアの開発があげられる。これらの成果により、研究の狙いである知を創生する高性能ストリームコンピューティングシステムの設計・実装を支援する開発システムを実現できる基礎技術を確立できた。(1)(2)の成果に関しては、システムの動作を確認するシミュレーション部分として利用でき、近年、注目される GPU や超並列アクセラレータを使って、実現するシステムの動作の確認を短期に完了できる手法の基盤技術として利用できる。ストリームを扱うシステムに対し、ソフトウェアレベルでの動作の確認が実現できたところで、(3)の成果であるハードウェアコンパイラを用いて、パイプラインハードウェアの合成ができる基盤を構築できた。アクセラレータで実行されるプログラムはデータストリームを扱うように記述されるという特徴を利用し、演算コンポーネントを実行タイミングを調整しながら接続することでパイプラインハードウェアを合成するコンパイラを開発した。以上のハードウェア合成技術に加え、データストリームを高速伝送する技術として、(4)のデータ圧縮ハードウェアを開発した。この圧縮技術では、実装に必要となるハードウェアのリソース量と圧縮率をコントロールし、さらに、伝送路の限界を超えるスループットを実現するリアルタイム伝送を可能にした。以上の4つの成果から、大量のデータストリームを扱うシステムをアクセラレータを使って高速にシミュレーションし、ハードウェアを合成するコンパイラを介して、ハードウェアを合成でき、さらには、その合成されたハードウェア間で移動するデータストリームを高速伝送することが可能になった。

(2) 詳細

本研究は以下の4つの大きなテーマでの成果を得られた。

研究テーマ A 「アクセラレータプログラム向けの高生産性プログラミング環境の構築」

システムの設計者は、データストリームを扱う計算アルゴリズムを flow-model と呼ばれるモデルにパックし、そのモデルをつなぎ合わせパイプラインを構成する。このモデルは、上記アクセラレータ向けプログラムと、その入出力ストリーム定義を含む。各モデルはアクセラレータで実行される必要があるが、従来の方法では、そのホストプロセッサがモデルをアクセラレータにマップし、さらに、入出力データをアクセラレータ側と交換する手続きを明に記述していくため、2種類のプログラムを同時に書く必要があり、生産性が悪い。本研究テーマでは、アクセラレータのプログラムをコマンドラインから直接実行できるメカニズムを実現し、パイプライン処理を構成する、個々のモデルの実行順序と入出力関係を指示することで、自動的に実行できる仕組みを実現できた。この実行を可能とするシステムに CarSh と名付け、flow-model や、flow-model とその I/O の関係性を記述したバッチファイルをコマンドラインに与えると、CarSh が flow-model をアクセラレータに自動的にマップしながら、データストリームを伝搬させ、パイプライン処理を実行していく[発表論文4]。図1に上記をまとめた全体の図を示す。

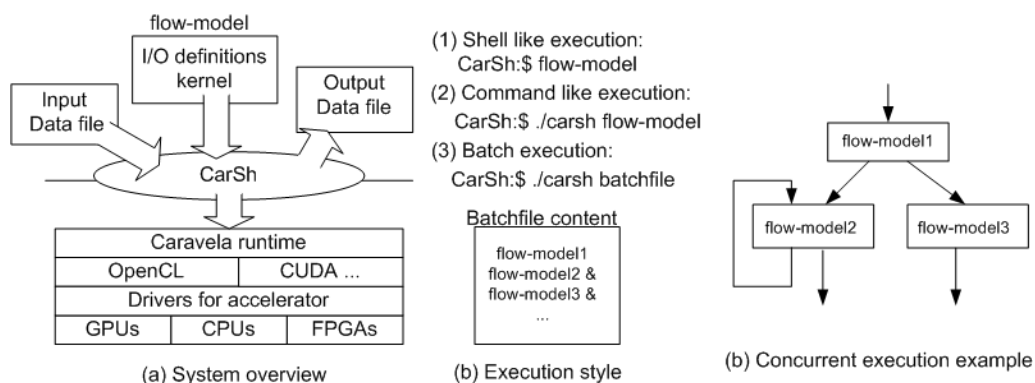


図1: CarSh によるプログラミング手法

研究テーマ B 「パイプライン実行の実行順序付けと並列性抽出による高速実行手法」
 アクセラレータが複数存在する環境において、構成したパイプラインから実行順序を特定し、さらに、パイプラインとしての並列性を抽出する手法を開発した[発表論文2]。この手法では、上述の CarSh の実行対象となるパイプライン定義を使って、最初の実行可能な flow-model をルートとする Spanning Tree を構成することで、唯一の実行順序を特定するとともに、自動的に並列性を抜き出し、パイプライン処理の並列化ができる。さらに、ホストプロセッサとアクセラレータ間のデータ移動がもたらすコピー操作のオーバヘッドにも着目し、Scenario-based Execution と呼ばれる新しいアクセラレータプログラム実行方式を確立した[発表論文3]。この実行方式の概略を図2に示す。

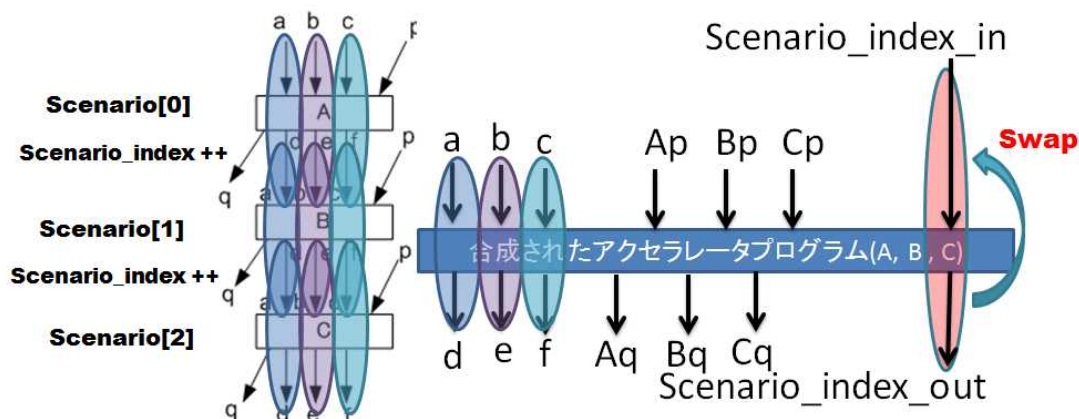


図2: Scenario-based Execution

研究テーマ C 「コンパクトなパイプラインハードウェアを出力するコンパイラの開発」
 上記、研究テーマ A, B の成果で動作確認をしたパイプライン処理を構成する flow-model を直接コンパイルすることでハードウェアを出力するストリームパイプラインコンパイラを開発した[発表論文5]。対象言語は OpenCL とし、そのアクセラレータで実行されるプログラムを中間モデル (HAM) に変換する。このとき、利用可能なコンポーネントを集めることで、実装可能な演算を合成する。HAM をハードウェア記述言語 (VHDL) に出力することで、パイプラインハードウェアを実装する。この VHDL で記述された回路を FPGA 等の実装向け EDA ツールに入力することで、ハードウェア合成が可能になった。

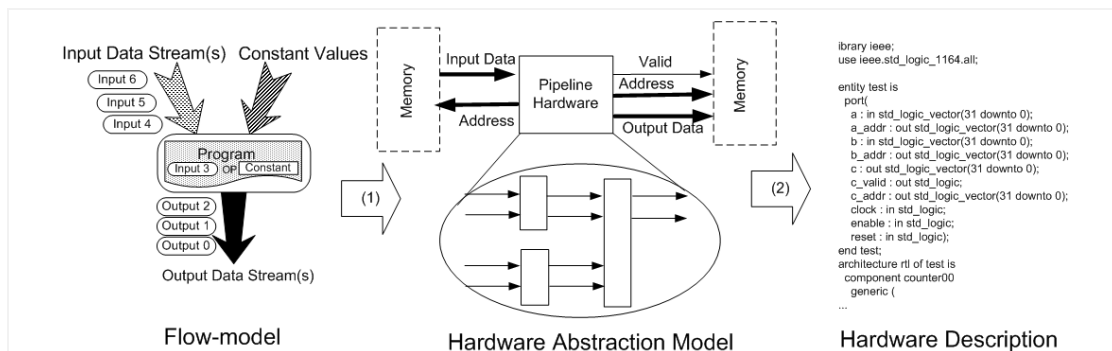


図3: ストリームパイプラインコンパイラ

研究テーマD 「ストリームデータ圧縮ハードウェアの開発」

データ伝送路のスピードは光速に近づきつつあり、このままだと実装技術の限界に達し、破綻してしまう恐れがある。そこで、データストリームの伝送路に対し、データの高度な圧縮が可能になれば、データ伝送路の実装限界の問題は緩和され、光の速度を超えるデータ伝送が期待できる。しかし、現代の圧縮回路は(1)圧縮のための処理時間が一定ではなく、(2)必要となるメモリサイズが予測できなく、さらに、(3)圧縮率とハードウェア量を制御できない。そこで、2期生の坂本比呂志氏とのコラボレーションにより、LCA-SLT と呼ばれる新たなストリーム圧縮手法を開発した[発表論文1]。この手法では、2シンボルから1シンボルに変換する機構をもち、テーブルを圧縮・解凍の双方で交換しなくてよい。さらに、複数の圧縮・解凍モジュールをスケラブルに接続でき、圧縮率とハードウェア量の調整が可能になっている。図4にこの手法の概要を示す。

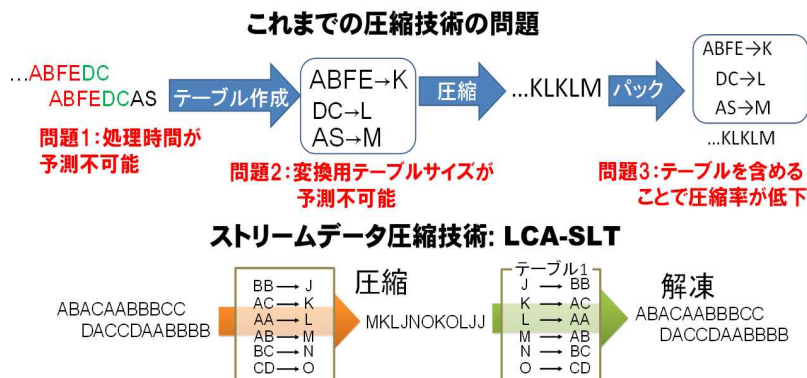


図4: ストリームデータ圧縮ハードウェア

3. 今後の展開

研究テーマから得られた成果を元に、ハードウェアを統合的に開発できる GUI 環境を構築し、実用化可能な段階へと進めることを目標とする。アクセラレータでのシミュレーション部分には、スーパーコンピュータを利用可能なインタフェースをターゲットとする。

4. 評価

(1) 自己評価

研究成果 A、B に関しては、研究のねらいの前半部分のハードウェアシミュレーションを高速に実施する点に一致し、研究成果 C が、研究の狙いのコンパイラ部分にあたる。これらの基本

理論は確立できたがデザインツールとして全体をまとめるまでは至らなかったが、研究の狙いの通り、研究は実施されたと評価する。また、データ伝送路のデータ圧縮技術に関しても、研究成果 D で、さきがけ研究者内において分野を超えた専門技術情報交換ができ、新たな技術開発、さらには、本研究課題の発展までに着手できたことは、驚きであり、高く自己評価する。

(2) 研究総括評価(本研究課題について、研究期間中に実施された、年2回の領域会議での評価フィードバックを踏まえつつ、以下の通り、事後評価を行った)。

ストリーム・コンピューティングのパラダイムを採用し、大量のデータフローを効果的に扱うための高性能でコンパクトなハードウェアシステムを構成できる、ハードウェア化まで一貫している基盤技術を開発するという研究である。

研究期間を通じ、アクセラレータプログラム向けの高生産性プログラミング環境の構築、パイプライン実行の実行順序付けと並列性抽出による高速実行手法、コンパクトなパイプラインハードウェアを出力するコンパイラの開発を行っている。これらによって、大量のデータストリームを扱うシステムをアクセラレータを使って高速にシミュレーションし、ハードウェアを合成するコンパイラを介して、ハードウェアを合成でき、さらに、その合成されたハードウェア間で移動するデータストリームを高速伝送することが可能になっており、目標を達成している。また、二期生の坂本比呂志研究者のデータ圧縮技術を取り入れ、当初予定外のストリームデータ圧縮ハードウェアも開発しており、製品化に向けた取り組みを行っており、評価する。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

- | |
|--|
| 1. Shinichi Yamagiwa and Hiroshi Sakamoto, A Reconfigurable Stream Compression Hardware based on Static Symbol-Lookup Table, Proceedings of IEEE BigData/BPOE 2013, pp. 86-93, IEEE, October 2013. |
| 2. Shinichi Yamagiwa, Ryo Jozaki, Shixun Zhang, Ryo Zaizen and Dewen Xu, "Exploiting Execution Order and Parallelism from Processing Flow Applying Pipeline-based Programming Method on Manycore Accelerators", Proc. ICPP, 708-717, IEEE, 2013. |
| 3. Shinichi Yamagiwa and Shixun Zhang, Scenario-based Execution Method for Massively Parallel Accelerators, The 11th IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA-13), pp. 1039-1048, IEEE, July 2013 |
| 4. Shinichi Yamagiwa and Shixun Zhang, CarSh: "CarSh: A Commandline Execution Support for Stream-based Acceleration Environment", Procedia Computer Science 18:601-610, Elsevier, 2013. |
| 5. Shinichi Yamagiwa, Ryoyu Watanabe and Koichi Wada, "Operation Synchronization Technique on Pipeline-based Hardware Synthesis Applying Stream-based Computing Framework", Proc. IEEE IPDPS/APDCM2013, 761-770, IEEE, 2013. |

(2) 特許出願

研究期間累積件数: 4件

1.

発明者：山際伸一
発明の名称：プログラム・及び情報処理装置
出願人：筑波大学
出願日：2013/7/10
出願番号：特願 2013-144661

2.

発明者：山際伸一、坂本比呂志
発明の名称：データ圧縮器及びデータ解凍器
出願人：筑波大学、九州工業大学
出願日：2013/6/4
出願番号：特願 2013-118356

3.

発明者：山際伸一
発明の名称：アクセラレータ処理実行装置、及びアクセラレータ処理実行プログラム
出願人：筑波大学
出願日：2013/5/28
出願番号：特願 2013-109741

4.

発明者：山際伸一
発明の名称：ハードウェア設計装置、及びハードウェア設計用プログラム
出願人：筑波大学
出願日：2013/5/17
出願番号：特願 2013-105024

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

<招待講演>

(1) GPU パワーを利用するアカデミックチャレンジ ～津波予測からスポーツ科学まで～,
NVIDIA GPU Technology Conference 2012, 2012 年 7 月 26 日.

(2) ストリームコンピューティング方式のプログラミングを利用したハードウェア合成”,
IEICE ソサエティ大会, 2011 年 9 月

<報道>

(1) NIKKEI BP Tech-On! 「GPGPU は EDA をどう変えるか」、札幌で電子情報通信学会ソ
サエティ大会の特別企画を聴講—その2, 2011 年 9 月 26 日

<http://techon.nikkeibp.co.jp/article/NEWS/20110926/198590/>