

様式 C-7-1

平成29年度科学研究費助成事業（科学研究費補助金）実績報告書（研究実績報告書）

所属研究機関名称		奈良先端科学技術大学院大学	機関番号	14603
研究代表者	部局	情報科学研究科		
	職	教授		
	氏名	中島 康彦		

1. 研究種目名 基盤研究(A)(一般) 2. 課題番号 17H00730

3. 研究課題名 エッジコンピューティングを牽引するニアメモリ高効率計算基盤

4. 研究期間 平成29年度～平成32年度 5. 領域番号・区分 -

6. 研究実績の概要

無数のセンサが一次データを全てクラウドに送信する分業方式では通信インフラの渋滞や応答速度の低下を招く。このため、センサやエッジにおいて二次データに加工圧縮後送信する分業方式が求められている。しかし、センサやエッジでは電源や冷却等の制約から潤沢な計算能力を確保することが難しい。そこでプログラマビリティと引き換えに電力効率の良い計算基盤が多数提案される状況にある。具体的には、GPU混載プロセッサによる自動運転の実現等、エッジコンピューティングと人工知能の融合が進んでいる。しかし、様々な場面に応用するには、広域分散ステンシル計算やベクトル長の短い畳み込み演算をGPUより遥かに低コストかつ低電力に実行できる計算基盤が必要である。本研究では、メモリとして扱うことができるシストリックリング型アクセラレータ設計と評価を行った。初年度の成果は次の通りである。（1）列方向マルチスレディング技術の考案；（2）チップ内ローカルリカレントメモリ（RMM）の導入；（3）ホストに対してメモリとして見せるインタフェースの考案；により、【1】エッジ高機能化に資するメモリ主導型CGRAの探索と体系化；【2】大規模専用ハードウェアを模擬できるCGRA仮想化技術の探索；に関して大きな成果を挙げる事ができた。また【2】プログラマビリティの飛躍的改善と性能チューニング手法の探索；についても、開発済CGRAとの機能互換を達成できたことにより、既存ツールチェーンの積み上げによる無駄の無い環境構築ができた。特に、ARM-SoC上に4列64行の機能を実装し、実機動作の確認まで完了できた。また、計画を前倒して28nmテクノロジーによる評価も完了した。以上、予定を大幅に上回るペースにて研究が進捗している。

7. キーワード

エッジコンピューティング シストリックリング マルチスレディング

8. 現在までの進捗状況

区分 (1) 当初の計画以上に進展している。

理由
当初の予定は次の通りであった。
【研究項目1】エッジ高機能化に資するメモリ主導型CGRAの探索と体系化：本格設計の前に、不規則な主記憶参照を伴う6重ルーブ画像処理カーネルを評価し、プロトタイプを設計する。見通しの良い性能予測のため、演算に干渉しない主記憶データ供給パスを設け、統合制御により、主記憶メモリ、メモリ間CGRA演算、メモリ主記憶を完全にオーバーラップさせ、稼働率100%を目指す。
【研究項目2】プログラマビリティの飛躍的改善と性能チューニング手法の探索：プログラマビリティの飛躍的改善に必要な基盤技術に関して網羅的に探索を行い、有望なアイデアを組み込んだツールチェーンの評価用プロトタイプを作成する。
【研究項目3】大規模専用ハードウェアを模擬できるCGRA仮想化技術の探索：専用回路において生じるデータ依存由来の無駄な動作時間をCGRA上に重畳実装することにより、CGRA資源の最大限活用と、CGRAの弱点である配線的大幅削減を同時に達成する。具体的な実装方法について詳細検討を開始する。
これに対し、今年度の実績は以下の通りである。
（1）列方向マルチスレディング技術の考案；（2）チップ内ローカルリカレントメモリ（RMM）の導入；（3）ホストに対してメモリとして見せるインタフェースの考案；により、【1】エッジ高機能化に資するメモリ主導型CGRAの探索と体系化；【2】大規模専用ハードウェアを模擬できるCGRA仮想化技術の探索；に関して大きな成果を挙げる事ができた。また【2】プログラマビリティの飛躍的改善と性能チューニング手法の探索；についても、開発済CGRAとの機能互換を達成できたことにより、既存ツールチェーンの積み上げによる無駄の無い環境構築ができた。特に、ARM-SoC上に4列64行の機能を実装し、実機動作の確認まで完了できた。また、計画を前倒して28nmテクノロジーによる評価も完了した。

4 版

9. 今後の研究の推進方策

当初予定の40nmテクノロジーによる大規模LSI試作は、もはや国内では必要なCADを保有している引き受け手がないことが判明した。また、十分な絶対性能に必要な高速I/Oが予算内では調達困難であることも判明した。CADレンタル費および高速I/O購入費を本予算で賄うことはできないため、LSI化の技術的目処が立った現状にて試作はスキップし、LSI開発については別途競争的資金に応募して継続を試みることにした。本年度は、LSI開発費削減のための技術開発を含む、以下3つの発展的課題に取り組む。

【4】エッジコンピューティング向け省面積CGRAのマルチチップ構成に関する検討と性能評価・分析：研究項目【1】と【3】を統合し、省面積LSIの特長を最大限生かす方向として、マルチチップ化による容易な拡張を可能とする構成の検討を開始する。このために、ARM-SoCを拡張するVU440を追加購入し、マルチチップ構成評価環境を整備する。

【5】エッジコンピューティング向けの共有CNN方式と分割推論モデルの構築と評価：研究項目【2】の発展として動画認識によるデータ圧縮をエッジコンピューティングのキラーアプリケーションとし、分散機械学習による複数エッジのグループ化、重みの共有、中間結果の圧縮による、エッジとデバイスを含む最適化技術の探索を行う。

【6】エッジコンピューティング向けのアナログアプロキシメイト演算方式の検討と性能評価・分析：半導体微細化の終焉により急速に注目を集めているアナログ近似計算機構を成果の出ているCGRAと組み合わせるハイブリッド構成について探索を開始する。

10. 研究発表（平成29年度の研究成果）

〔雑誌論文〕 計6件（うち査読付論文 6件 / うち国際共著論文 0件 / うちオープンアクセス 0件）

1. 著者名 Hoang Gia Vu, Shinya Takamaeda-Yamazaki, Takashi Nakada, Yasuhiko Nakashima	4. 巻 FCCM2017
2. 論文標題 CPRring: A Structure-aware Ring-based Checkpointing Architecture for FPGA Computing	5. 発行年 2017年
3. 雑誌名 FCCM2017	6. 最初と最後の頁 192-192
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/FCCM.2017.60	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Renyuan Zhang, Takashi Nakada and Yasuhiko Nakashima	4. 巻 CANDAR'17
2. 論文標題 A Feasibility Study of Programmable Analog Calculation Unit for Approximate Computing	5. 発行年 2017年
3. 雑誌名 CANDAR'17	6. 最初と最後の頁 180-186
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hoang-Gia VU, Shinya TAKAMAEDA-YAMAZAKI, Takashi NAKADA, and Yasuhiko NAKASHIMA	4. 巻 Vol. E101-D No.2
2. 論文標題 A Tree-based Checkpointing Architecture for the Dependability of FPGA Computing	5. 発行年 2018年
3. 雑誌名 IEICE TRANSACTIONS on Information and Systems	6. 最初と最後の頁 288-302
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hoang Gia Vu, Takashi Nakada, and Yasuhiko Nakashima	4. 巻 ARC2018
2. 論文標題 Efficient Multitasking on FPGA Using HDL-based Checkpointing	5. 発行年 2018年
3. 雑誌名 ARC2018	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takahiro Ichikura, Ryusuke Yamano, Yuma Kikutani, Renyuan Zhang, and Yasuhiko Nakashima	4. 巻 COOLCHIPS2018
2. 論文標題 EMAXVR: A Programmable Accelerator Employing Near ALU Utilization to DSA	5. 発行年 2018年
3. 雑誌名 COOLCHIPS2018	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Noriyuki Uetake, Renyuan Zhang, Takashi Nakada, and Yasuhiko Nakashima	4. 巻 COOLCHIPS2018
2. 論文標題 A Programmable Analog Calculation Unit for Vector Computations	5. 発行年 2018年
3. 雑誌名 COOLCHIPS2018	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計10件 (うち招待講演 3件 / うち国際学会 0件)

1. 発表者名 Keisuke Fujimoto, Takashi Nakada, Shinya Takamaeda-Yamazaki, Yasuhiko Nakashima
2. 発表標題 A Multi-Level Power-Capping Mechanism for FPGAs
3. 学会等名 xSIG2017
4. 発表年 2017年

4版

1. 発表者名 平賀由利亜, 三谷剛正, 福岡久和, 中田尚, 中島康彦
2. 発表標題 エッジコンピューティングによる分散ニューラルネットワークの構想
3. 学会等名 電子情報通信学会コンピュータシステム研究会
4. 発表年 2017年

1. 発表者名 山野龍佑, 中島康彦
2. 発表標題 時分割多重実行によるシストリックリングの面積効率向上手法
3. 学会等名 電子情報通信学会コンピュータシステム研究会
4. 発表年 2017年

1. 発表者名 福岡久和, 山野龍佑, 中島康彦
2. 発表標題 各種FPGAによる畳み込み演算向けシストリックリングの実装と評価
3. 学会等名 電子情報通信学会コンピュータシステム研究会
4. 発表年 2017年

1. 発表者名 中島康彦
2. 発表標題 GoogleのTPUにも使われたシストリックアレイアーキテクチャとDeep Learningについて
3. 学会等名 富士通研究所技術講演会(招待講演)
4. 発表年 2017年

1. 発表者名 中島康彦
2. 発表標題 Deep Learningに向けたApproximate Computingとシストリックアレイアーキテクチャ
3. 学会等名 革新的コンピューティングの研究開発戦略検討会（招待講演）
4. 発表年 2017年

1. 発表者名 Renyuan Zhang, Takashi Nakada, Yasuhiko Nakashima
2. 発表標題 A Programmable Analog Calculation Unit based on Support Vector Regression
3. 学会等名 電子情報通信学会コンピュータシステム研究会
4. 発表年 2017年

1. 発表者名 中島康彦
2. 発表標題 Approximate Computingとシストリックアレイ
3. 学会等名 ジスクソフト技術講演会（招待講演）
4. 発表年 2017年

1. 発表者名 菊谷雄真, 山野龍佑, 一倉孝宏, 中島康彦
2. 発表標題 時分割多重実行型シストリックリングの実装と評価
3. 学会等名 電子情報通信学会コンピュータシステム研究会
4. 発表年 2017年

4版

1. 発表者名 菊谷雄真, 山野龍佑, 一倉孝宏, 中島康彦
2. 発表標題 エッジコンピューティング向けアクセラレータの実装と評価
3. 学会等名 電子情報通信学会関西支部第23回研究発表講演会
4. 発表年 2017年

〔図書〕 計0件

1 1. 研究成果による産業財産権の出願・取得状況

〔出願〕 計2件

産業財産権の名称 データ処理装置（高効率アクセラレータ構成方法）	発明者 中島康彦	権利者 同左
産業財産権の種類、番号 特許、特願2017-96061	出願年 2017年	国内・外国の別 国内

産業財産権の名称 データ処理装置（高効率アクセラレータ構成方法）	発明者 中島康彦	権利者 同左
産業財産権の種類、番号 特許、PCT/JP2018手続中	出願年 2018年	国内・外国の別 外国

〔取得〕 計0件

1 2. 科研費を使用して開催した国際研究集会

計0件

1 3. 本研究に関連して実施した国際共同研究の実施状況

-

1 4. 備考

奈良先端科学技術大学院大学コンピューティングアーキテクチャ研究室
<http://arch.naist.jp>