

研究報告書

「大量ストリームデータのリアルタイム処理に向けた柔軟なアーキテクチャ探索と設計環境構築」

研究期間：2018年4月～2020年3月

研究者番号：50211

研究者：原 祐子

1. 研究のねらい

近年、Internet of Things (IoT)デバイスが急速に普及し、多くのデバイスがネットワークに接続されることで、膨大なデータがクラウド上のコンピュータで処理されている。このようなクラウドコンピューティングでは、IoT デバイスやデータ量が増えるにつれネットワーク帯域を切迫し、IoT エッジデバイスへの応答遅延が重大な問題を引き起こす。特に、ヘルスケアやファクトリーオートメーション(FA)における異常検出などにおいては、リアルタイム性は重要な要求である。この解決する方法として、エッジサーバやエッジデバイス上におけるコンピューティング技術、すなわちエッジコンピューティングに適したアーキテクチャ設計技術の確立が喫緊の課題である。

エッジコンピューティング向けアーキテクチャ設計には、リアルタイム性のほかにも、生産・管理コスト軽減のため小規模(小回路面積)かつ省エネルギー(バッテリ長寿命)であること、サービス(アプリケーション、アルゴリズム)の多様化・頻繁な仕様変更を見越した柔軟性(高い再利用性)を持つこと、などの要求が挙げられる。さらに、アーキテクチャ開発だけでなく、実用化、更なる研究促進、新たな技術・課題創出のために、考案したアーキテクチャとそのハードウェア・ソフトウェアの自動設計環境も含めた研究開発が必要不可欠である。

以上の要求を鑑みて、今後の IoT エッジコンピューティングの普及と IoT のセンシングデータの多くはストリームデータであることを踏まえ、本研究は「大量ストリームデータを効率的・柔軟に処理する省エネルギーなアーキテクチャとその設計環境」の構築に取り組む。本研究では、主に軽量演算から成るアプリケーションドメインを対象とした、新たな IoT エッジコンピューティング向けドメインスペシフィックアーキテクチャを提案する。大規模・複雑化し続けている組込みプロセッサの傾向とは逆転の発想に基づき、計算資源を極限まで削減した超小型プロセッサコアを用いたマルチコアプロセッサを実現し、ターゲットアプリケーションに応じて「設計柔軟性・リアルタイム性・省エネルギー」という 3 点の設計要求を全て満足させる。これにより、今後の IoT 社会の課題や在り方を見据えた本課題の一連の研究成果により、アーキテクチャだけでなく、サービス(アプリケーション)やセンサ技術等の周辺技術の研究開発も活性化し、今後の日本の情報社会の発展・成長に繋げることを狙う。

2. 研究成果

(1)概要

複数の演算器やプロセッサコアを利用してリアルタイム処理(高スループット処理)を実現する計算機システムとして、ベクトルプロセッサや高性能マルチコア(メニーコア)プロセッサなど

が数多く研究されてきた。しかし、各演算器やプロセッサコアの機能が過剰であるため、バッテリ容量によるエネルギー制約が厳しく、かつ、常時稼働することを要求されるヘルスケアやファクトリーオートメーションの異常検出を行う IoT デバイスには、既存のプロセッサは使用できない。

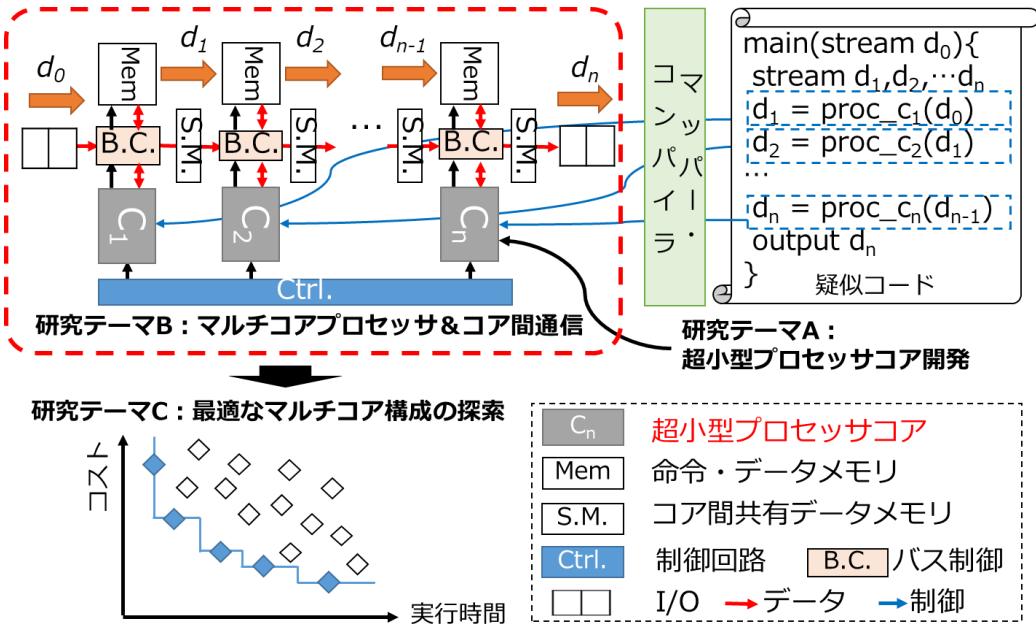
本研究は、IoT エッジデバイスのセンサから得られたストリームデータに対して、データセンタに転送する前に前処理、あるいは、リアルタイム処理によるサービスを提供するエッジコンピューティングのための新たなプロセッサを提案した。具体的には、従来のプロセッサより小型ながら、高性能かつ省エネルギーで対象アプリケーションを処理可能なプロセッサコア、および、それを応用したマルチコアプロセッサを開発した。プロセッサコアでは、大規模化し続ける従来の設計方針とは逆行したアプローチを取り、軽量アルゴリズムを対象ドメインとして、敢えて機能は必要最低限に抑えた超小型プロセッサの設計を行うことで、回路資源や電力消費の無駄を削減した。マルチコアプロセッサでは、この超小型プロセッサを複数コア用いることで、データをパイプライン式に処理し、エネルギー消費量の無駄を抑えつつ、高スループット処理を可能にした。さらに、アプリケーションや設計制約に応じてコア数やコア間通信トポロジを変更可能な構造にすることで、柔軟性・再利用性も得られる仕組みを確立し、その最適なマルチコアプロセッサ構成(コア数および通信トポロジ)を探索する手法を提案した。本課題では、ストリームデータを対象に行う軽量アルゴリズム数種類と実アプリケーション(ヘルスケアモニタリングにおける異常検出)を想定したケーススタディによって、既存プロセッサ(MIPS, ARM Cortex-M0, RISC-V)と本研究成果の比較を行った。超小型プロセッサコアは、回路面積・電力をともに既存プロセッサ Cortex-M0 の半分以下、RISC-V の 1/3 以下に抑えつつ、より高性能であり、実アプリケーションで要求される時間内に処理を完了できる(リアルタイム処理ができる)ことを示した。同様に、マルチコアプロセッサは複数のマルチコア構成で、既存プロセッサ MIPS よりエネルギー消費量を削減できることを示した。

(2) 詳細

新たなプロセッサを実用化へ繋げるためには、ハードウェア(アーキテクチャ)開発のみだけでなく、その効率的な利活用も重要な課題である。したがって、本研究では、図1に示す通り、小型プロセッサコア設計(研究テーマ A)、様々なマルチコアプロセッサの開発(研究テーマ B)、および、マルチコアプロセッサの最適な構成(コア数およびコア間通信トポロジ)の設計探索(研究テーマ C)という3つの課題に細分化して同時並行で進めた。

● 研究テーマ A「超小型プロセッサ SubRISC+の開発」(論文 1、論文 2、特許出願 1 件)

本研究では、超小型プロセッサの 1 つである SNG (Subtract and branch on NeGative with 4 operands)に着目し、必要最小限の機能拡張を施した小型プロセッサを提案した。SNG は、1 種類の命令のみを扱えるプロセッサ(One-Instruction Set Computer; OISC)の 1 つであり、プロセッサの初期に提唱されたものの 1 つである。チューリング完全であり、1 種類の命令を組み合わせて、いかなる演算も実現できる。プロセッサ構造が極めてシンプル、すなわち回路面積が非常に小さいというメリットがある一方、その非効率さ(命令数やメモリアクセス多)により実アプリケーションへの適用は困難である。本研究テーマでは、この SNG をベースに、超小型というメリットを活かしつつ、より演算効率が高い(すなわち高性能な)小型プロセッサ SubRISC+



【図1: 提案手法全体の概要】

TSMC 65nm LP	Cortex M0 (w/o Mult)	zero-riscy ^{*2}		micro-riscy ^{*3}	提案
		Fast Mult	Slow Mult		
面積 [K um ²]	20,934	34,319	32,226	24,094	10,356
ゲート数 [K GE ^{*1}]	14,538	23,833 (x1.64)	22,379 (x1.54)	16,732 (x1.15)	7,192 (x0.49)

*1 GE=Gate Equivalent (全面積um²を2入力NAND面積um²で除した値)

*2 レジスタファイル32エントリ, 16ビット乗除算有り

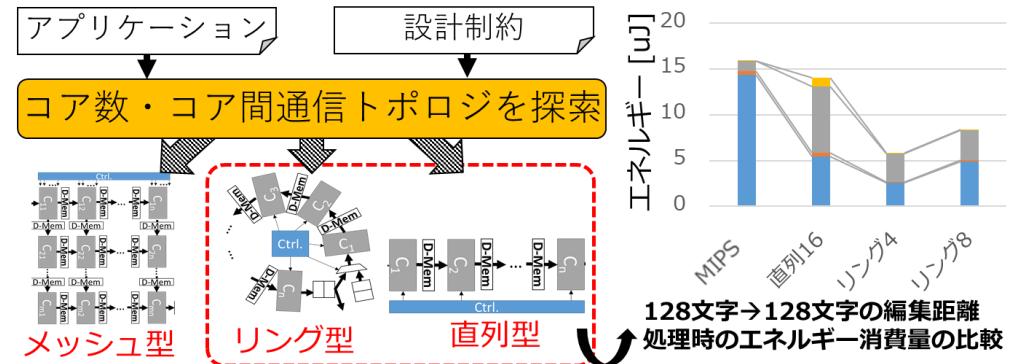
*3 レジスタファイル16エントリ, 乗除算無し

【表1: SubRISC+(提案)の回路実装評価】

実行時間 [ms]	M0 提案	正規化	エネルギー [uJ]	M0 提案	正規化
Sort	1.9 1.2	(x0.63)	Sort	8.6 5.4	(x0.63)
Motion	0.19 0.14	(x0.74)	Motion	0.81 0.59	(x0.73)
IntDCT	0.11 0.09	(x0.82)	IntDCT	0.5 0.38	(x0.76)
Histo	0.12 0.06	(x0.50)	Histo	0.57 0.28	(x0.49)
Edge	0.36 0.34	(x0.94)	Edge	1.62 1.25	(x0.77)
WordCnt	2.37 3.27	(x1.38)	WordCnt	10.46 13.54	(x1.29)

【表2: 軽量アルゴリズムにおける実行時間とエネルギー消費量の評価】

を開発した。具体的には、本研究が対象とする軽量アルゴリズム(ヘルスケアやファクトリーオートメーションにおける異常検出等)において頻出する演算の種類を解析し、少ないハードウェア資源で、かつ、多様な演算を効率良く実現できる4種の命令のみを命令セットとして選定した。既存の小型プロセッサ(ARM Cortex-M0, RISC-V)よりさらに小規模な回路実装を実現した(表1)。さらに Cortex-M0 に比べ、提案プロセッサは高速かつ省エネルギーに6種のアプリケーションを処理できることを示した(表2)。また、ヘルスケアモニタリングのケーススタディとして、100Hzサンプリング周波数の脳波から、てんかんの発作検出を行うアプリケーションに



【図 2: 様々なマルチコア構成(研究テーマ B)と最適な構成の自動探索(研究テーマ C)]について評価した結果、サンプリング周期より高速に発作検出可能なことを確認し、十分なリアルタイム性を実現できることを示した。

本研究成果は、国際論文誌および査読付き国際会議にて発表済みであり、特許出願を行った[論文 1、論文 2、特許出願 1]。また、ヘルスケアモニタリングのケーススタディについて、国際会議の University Booth で成果発表をした。また、本研究成果は ASIC チップ試作を行つた(3月末に試作チップ納品のため、4月以降に評価予定)。

● 研究テーマ B「超小型プロセッサコアを用いたマルチプロセッサ」(その他の成果 1 件)

小型プロセッサを複数コア用いて、ストリームデータをパイプライン式に処理することでさらに高性能・リアルタイム処理を実現するマルチコアプロセッサを開発した。アプリケーションや設計制約に応じて、コア数やコア間通信トポロジを柔軟に変更可能な仕組みを構築した。ACT-I 期間中(2016 年 12 月～2018 年 3 月)に既に直列型のマルチコアを実現し、ソーティングや文字列の編集距離(リーベンシュタイン距離)を算出するアプリケーションを用いて、既存プロセッサに対して性能およびエネルギー消費量の優位性を示した。ACT-I 加速フェーズでは、さらにリング型やメッシュ型のマルチコアプロセッサを実現した(図 2 左)。

本研究テーマは、研究テーマ A と並行して進めたため、研究テーマ A がベースとした SNG を用いてマルチコアの検討・実装を進め、コア間通信の改良点を探った。

図 2 右は、編集距離(128 文字)について、従来プロセッサの MIPS(シングルコア)、ACT-I 研究で開発した直列型マルチコア(16 コア)、および、ACT-I 加速研究で開発したリング型マルチコア(4 コア・8 コア)のエネルギー消費量を比較した結果である(結果の一部を抜粋)。研究テーマ A と研究テーマ B を統合前の段階で、既に既存プロセッサよりエネルギー効率を大幅に改善できた。研究テーマ A の成果と統合することで、更に性能・エネルギー効率を改善できると考えられる。

本研究成果は、国際会議にて発表済みである[その他の成果(国際会議招待講演)]。

● 研究テーマ C「最適なマルチコアプロセッサ構築のための設計空間探索」

図 2 に示した通り、研究テーマ B で開発した様々な構成のマルチコアプロセッサのうち、エネルギーを最小化する構成の探索は、IoT デバイスの開発において非常に重要な課題である。最適なエネルギー効率を実現するマルチコア構成は、アプリケーションに異なる。また、設計制約(回路面積等)により、採用不可能な構成を早期に候補から排除することも重要であ

る。本研究テーマでは、アプリケーションやコア間通信データ量に応じてマルチコア構成(コア数およびコア間通信トポロジ)をモデリングし、設計制約に応じて最適な構成を探索する手法を提案した。モデリングには、Timed Petri Nets という、時間情報を扱えるペトリネットを活用し、各コアの処理時間、コア間通信に要する時間、コア間の同期を抽象的な数学モデルとして表現した。

本研究テーマは、ACT-I 加速研究終了時までに完了できなかったが、コア間通信データ量は抽象的かつ概ね正確なモデルを構築できている。現在は、アプリケーション全体のエネルギー消費量(計算処理にかかるエネルギー消費量とコア間通信にかかるエネルギー消費量)のモデリングに取り組んでおり、2020 年度中にエネルギー効率を最適化するマルチコア構成を自動探索可能なフレームワークが完成する見込みである。

3. 今後の展開

本研究は、IoT エッジデバイス上で行うエッジコンピューティングに向けた、小型・省電力プロセッサーアーキテクチャの開発を中心に、ハードウェアアーキテクチャとソフトウェアアプリケーションの両面から幅広く研究に取り組んだ。さらに、開発技術の応用として、特にヘルスケアやファクトリーオートメーション等のアプリケーションドメインを対象に、新たなエッジコンピューティングを開拓する技術に注力してきた。これらのアプリケーションでの評価において、本研究成果は IoT デバイスの物理セキュリティの向上と高い親和性を持ち、IoT セキュリティの強化に発展・展開できる感触を得られた。今後は、ACT-I 加速研究成果を更に拡張・応用し、物理セキュリティを強化した、小型・省電力 IoT デバイスの開発に取り組む予定である。ネットワークに繋がる IoT 機器のセキュリティレベルを底上げすることで、IoT 全体のセキュリティを強化でき、産業・社会への波及効果は大きいと考えられる。

上記の応用・展開には、組込システムの設計レイヤの縦の連携や、セキュリティの数理モデル等の横の連携が欠かせない。縦・横の異分野連携を通して、新たな学術分野を開拓するとともに、積極的に各分野の研究者の交流を図り、学生の教育にも力を入れていく。

4. 自己評価

● 研究目的の達成状況

本研究は、益々普及していくIoTデバイスのセンサから収集され、増加し続けるストリームデータを IoT エッジデバイス上で処理することを想定した、計算機(プロセッサ)およびその開発環境を構築することを目的として、ACT-I 研究から継続して取り組んできた。本研究課題は、アーキテクチャ設計とソフトウェア開発環境の両課題から成るため、3 つの研究テーマにブレイクダウンして同時並行に進めることで、テーマ間で問題を共有しに相互フィードバックすることで、研究促進を図ることができた。広く使用されている商用プロセッサ(ARM Cortex-M0 や RISC-V micro-risky)より、小型・高速・省エネルギーなプロセッサコア設計、および、その応用(マルチコア化)を示すことができた。特に、インパクトファクターの高い国際論文誌(IF=1.5)である IEEE Embedded Systems Letters に論文が採択されたことは、組込みシステムの研究コミュニティから国際的に高い評価を得られたことを実証している。また、本研究成果を論文執筆だけにとどめることなく、チップ試作や特許出願にも挑戦できた点は、一研究者として大きく成長できた点であると考えている。対象とすべきドメインやアプリケーションの特徴を整理し、狙いをクリアにできた点も意義が高い。

一方で、ACT-I 加速フェーズの 2 年間をもってしても、3 つの研究テーマすべてを完結させ、統合することができなかった。特に研究テーマ C については、未だ成果は得られていないが、手応えを感じ始めている段階であるため、ACT-I 研究終了後も継続して取り組んでいく。

ACT-I 加速フェーズ中に得られた成果と知見を元に、今後は AIP 加速課題として IoT のセキュリティを強化する方向で、本研究成果をさらに拡張・展開していく。

● 研究の進め方

現在のみならず今後確実に増加すると考えられるアプリケーションの傾向を調査するとともに、エッジコンピューティングの社会的ニーズを明らかにした。それにより、対象とすべきアプリケーションドメインを絞り、サポートすべき機能(命令セット)を明らかにした。提案マルチコアプロセッサを開発するにあたり、(1)コア設計、(2)マルチコア設計(特にコア間通信のメカニズム)、(3)設計空間探索フレームワークの 3 点に課題をブレイクダウンし、同時並行で進めた。適宜、各技術・実装で必要な点を情報共有・フィードバックすることで、設計の早い段階から実装の完成度を高めることができた。

回路設計用のソフトウェア(特にシミュレーション)を使用するためにハイパフォーマンスな計算機環境(サーバ等)を充実させる必要があり、本研究費を有効に執行できたと考えている。さらに、開発したプロセッサの特許出願、および、ASIC チップ試作にチャレンジすることができた。これらの成果を起点に、今後も学術研究だけでなく社会実装の可能性を検討していく。

● 研究成果の科学技術及び学術・産業・社会・文化への波及効果

本研究成果は、IoT センサ等に搭載し、そのデバイス上で行うエッジコンピューティングに利用することを想定している。現在のエッジコンピューティングは、比較的高性能なエッジサーバにおける計算処理を行うことを想定している。バッテリ制約が厳しく、かつ、長時間・常時動作することを要求されるセンサノードのような IoT デバイスへ、既存プロセッサを適用することは困難である。

本研究成果は、今後社会実装されるべきエッジコンピューティングの在り方を考え直す、

一つの切欠になると考えらえる。すなわち、データはどの程度エッジ端末上で処理するのか・保存するべきなのか、どのような処理・データをクラウド側・エッジ側で扱うのか等について、これまで検討する余地がなかったが、IoT エッジデバイスで効率良く処理する技術があることを示すことで、再考されるものと期待される。現在のニーズだけでなく、将来のニーズを見据えた技術開発を行っている点で、本研究課題の社会的影響・波及効果は高いと考える。

学術研究において、プロセッサの命令セットや小型化に取り組む研究はこれまで多数あるが、その多くは「どんなアプリケーションに対しても」高性能であることを目指してきた。一方、近年 IRDS も提唱しているように、今後はアプリケーションドメインに特化した「ドメインスペシフィックアーキテクチャ (DSA)」が重要なアプローチとなる。DSA では、どのようなドメインを切り出し、対象にすべきかが、重要な視点となる。本研究は、その先駆けとして、多くの研究がドメインとしてケアしてこなかったが今後確実に増えるアプリケーションの特徴と粒度をドメインとしてフォーカスし、さらにそのドメインに特化した効率性を追求したアプローチを実施した。この研究アプローチと成果は、今後の学術研究の取り組みに大きく寄与すると考える。

● 研究課題の独創性・挑戦性

本研究課題の技術は、IoT エッジデバイスのセンサから得られたデータに対して、複雑なアプリケーション(算術演算等)の前処理として、軽量処理な(異常検出等)を省電力・省エネルギーに行うためのものである。一般的な前処理としては、算術演算を多く含むフィルタ等があり、これまでにも数多く・多様な DSP 設計が行われてきた。しかし、本研究では、今後は IoT センサノードのような小型 IoT デバイスにおけるエッジコンピューティングの需要の高まり、および、エッジコンピューティングにおいて DSP とは異なる性質を持つ前処理(データの比較・探索、差分検出等から成る前処理)が重要となることを見据え、独創的な着眼点によって、これまでの対象とは異なる前処理にフォーカスしている。

一般的には柔軟性(再利用性)と効率的(無駄がない)設計は相反する性質を持ち、両立困難である。本研究では、アーキテクチャ設計において、アーキテクチャのどの部分をアプリケーションに応じてコンフィギュレーション(柔軟にカスタマイズ)可能にするか、および、柔軟性より効率を優先してハードウェイア設計(カスタマイズ不可)にするかについて、定性的かつ定量的に検討し、上記の相反する課題を両立しようしている点で、挑戦的である。

上述の独創性および挑戦性の点から、本研究課題は、今後のアーキテクチャ設計における1つのアプローチを開拓するものであると考えられる。

5. 主な研究成果リスト

(1)論文(原著論文)発表

1. Kaoru Saso and Yuko Hara-Azumi, "Revisiting Simple and Energy-Efficient Embedded Processor Designs Towards the Edge Computing," IEEE Embedded Systems Letters, 2019, 卷号・頁 未定 (Early Access 版公開済み)
2. Kaoru Saso and Yuko Hara-Azumi, "Simple Instruction-Set Computer for Area and Energy-Sensitive IoT Edge Devices," International Conference on Application-specific Systems, Architectures and Processors (ASAP), pp.93–96, Milan, Italy, Jul. 10th, 2018.

(2)特許出願

研究期間累積件数:1 件

1.

発明者: 原 祐子、佐宗 馨、楊 明宇

発明の名称 :少命令セット組込みプロセッサ

出願人: 東京工業大学

出願番号: 特願 2019-172638

(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)

- 国際会議招待講演

Yuko Hara-Azumi, "Highly-Scalable and Flexible Multicore Processors with Limited ISAs," 18th International Forum on MPSOC for Software-defined Hardware (MPSOC), Aug. 2nd, 2018.

- 国際会議デモンストレーション

Mingyu Yang and Yuko Hara-Azumi, "Implementation and Evaluation of an Embedded Processor for Lightweight IoT eHealth," University Booth of Design, Automation & Test in Europe (DATE), Grenoble, France, Mar. 11–12th 2020.