

## 科学研究費助成事業 研究成果報告書

平成 29 年 6 月 1 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2013～2016

課題番号：25730028

研究課題名(和文)新アーキテクチャによる高効率プロセッサコアおよびそのマルチコア構成の研究

研究課題名(英文)A Study of Novel Computer Architecture for High-efficient Processor Cores and its Multicore Structure

研究代表者

入江 英嗣(Hidetsugu, Irie)

東京大学・大学院情報理工学系研究科・准教授

研究者番号：50422407

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：コンピュータ性能の持続的成長を可能とする新しい基本方式の開発を行った。本研究はCPUが自動的にプログラムを効率良く実行する方式を改良するもので、メニーコアやSIMDといった、プログラムによる指示を効率化の前提とする方式と相乗あるいは補完する。研究初年度では上書きがないことを特徴とする独自命令セットを提案し、2年目はこの特徴が従来CPUの弱点を解消することを明らかにした。3年目はコンパイラを開発し、最終年度ではHWおよびSW性能向上技術により新方式の強化を行った。これらの成果より、新方式の実行効率は従来の数倍が見込まれ、本研究はシステム設計や人工知能対応への拡張など、次の段階に移行している。

研究成果の概要(英文)：In this research, we developed a novel computer architecture which enables sustainable improvement of the computing performance. By improving the technology in that CPU automatically optimizes the execution at runtime, this study complements the technology that optimizes the execution based on the given instructions by the programmers such as many-core or SIMD technologies. In the first year we proposed an original instruction set which is characterized by eliminating register overwrites, and in the second year we clarified that this feature resolves the weak point of conventional CPUs. In the third year, dedicated compiler is developed, and in the final year our architecture is strengthened by novel HW/SW optimizing technology. These results reveal that the execution efficiency of our architecture is expected to be several times higher than conventional, and this research has shifted to the next stage such as system design and extension to artificial intelligence correspondence.

研究分野：コンピュータ・システム

キーワード：コンピュータ・アーキテクチャ プロセッサ コンパイラ HDL マイクロアーキテクチャ キャッシュ  
・メモリ 命令レベル並列実行

### 1. 研究開始当初の背景

研究開始当初、多くの汎用プロセッサがマルチコア・アーキテクチャを採用するようになり、主要な性能向上は、より多くのプロセッサを1チップに集積させ、それらを並列に動作させることによって実現されていた。しかし、1チップ上のプロセッサ数が数十に達してくると i) プロセッサの負荷の偏りに起因する性能効率減少 ii) 全てのプロセッサ稼働に必要な電力供給の不足、の影響が顕著となり、単純なマルチコア・アーキテクチャの限界が指摘されるようになった。

このため、使える資源量や目的とするアプリケーションに応じて、異なる規模・用途のプロセッサを組み合わせるヘテロジニアス・マルチコア構成が注目されるようになり、数多くの小規模コアを集積するメニー・コア、シングルスレッド実行用の大きいコアとスループット実行(あるいは省電力)用の小さいコアで混成するビッグ・リトル、また GPU や FPGA の混載など、様々な構成が研究・議論されていた。

### 2. 研究の目的

コアの組み合わせの議論だけでは、コンピュータにおける処理の基本である、1スレッドのプログラム実行を効率化・高速化することはできず、汎用的に使えるチップの性能を劇的に向上させることは難しい。そこで本研究ではマルチコアを構成する個々のプロセッサについて、新しい実行方式により性能と電力消費の双方を改善することを目的とした。

このような根本的な改善を可能とするアイデアとして、デバイス技術の傾向に着目し、トランジスタ数を原資として、熱の原因となるクリティカルパスやホットスポット構造を排除することを基本方針とした。当時革新が急速に進んでいた分散キー・バリューストア技術に着目し、プロセッサのレジスタからバージョン管理を排して単純なキー・バリューストアとすれば、現在の高性能プロセッサ内でやりとりされている多くの制御信号を不要としながら同等の性能が達成できると着想し、この方式の開発を進める形で研究を進めることとした。

### 3. 研究の方法

提案する新実行方式「STRAIGHT アーキテクチャ」について以下のように研究を進めた。まず、着想したアーキテクチャの具体化として、パイプラインロジックの概要を設計し、性能および電力シミュレーションにより性能向上の余地を明らかにする。次に命令セット初版を策定し、対応するアセンブラを作成する。さらに、この初期命令セットを実行可能なエミュレータを作成すると共にパイプライン詳細ロジックを明らかにして詳細な性能シミュレータを作成する。これらの手順を経て、仕様を HDL で記述し、ハードウェア

として実現可能なことを示す。

次に、性能最適化技術について、ソフトウェアおよびハードウェア両面から研究を進める。ソフトウェア面では、まずレジスタのバージョン管理の必要がない命令セットについて、そのような制約を満たすコードをどのようなプログラムに対しても生成可能とするための仕様を策定する。さらに充分高クオリティなコードを自動生成するコンパイラのアルゴリズムを明らかにし、STRAIGHT アーキテクチャによって初めて実現可能な最適化技術を開発する。ハードウェア面では主にレジスタおよびメモリアクセスについて、静的なアプローチでは対処しきれない動的な最適化技術を研究する。これらの技術の評価は、前記シミュレータにより定量的に行う。

また、本方式のマルチコア構成のための技術として、半導体3次元積層技術およびキャッシュ制御に着目して研究を行う。3次元積層技術では、高性能プロセッサを3次元実装する場合の傾向について明らかにし、STRAIGHT が前提とするような、トランジスタ容量と熱の交換が現実的な設計であることを示す。キャッシュ制御の研究では、従来の LRU を越えた適応的なデータ取捨手法を開発し、高性能なプロセッサが複数実装された場合でも外部からのデータ供給遅延の影響を最小化する技術を明らかにする。

### 4. 研究成果

#### アーキテクチャ検討

アーキテクチャ検討では、i) 一度値が生成されたレジスタに上書きがされないこと ii) レジスタの寿命は一定時間に設定され、それ以降は参照されないことの二点をコード側で保証することが可能であれば、高性能プロセッサのクリティカルパスの一つであり、主要な電力消費源であるレジスタ・リネーム機構を排除できることを明らかにした。この影響を見積もるため、まず従来のスーパスカラ・モデル・シミュレータを利用してパラメタ変更により新アーキテクチャの性能を模し、シングルスレッド実行能力を30%以上向上させつつ電力効率を向上させることを確認した。このことにより提案アーキテクチャの有効性が見積もられ、当初の予定どおりにアーキテクチャの詳細化を進めることとなった。

上記二点の制約を自然に満たす命令形式として、値を生成する命令との距離でオペランドを指定する手法を採用し、この方針に基づいて初期命令セットアーキテクチャを決定した。

#### 評価環境作成

策定した命令セット仕様を元に、これに対応するエミュレータ、アセンブラ、サイクルレベルシミュレータを作成した。ベンチマークとして、Livermore ループ・プログラムの実行バイナリをハン

ド・アセンブリングにより作成し、性能評価を行った。

性能評価の結果、並列性抽出能力向上とメモリアクセス数削減の効果が確認され、同じ計算を従来アーキテクチャの数倍の性能で実行可能なことが明らかになった。また、よく知られたループ・アンローリング最適化を施すとさらに数倍の性能に反映されることが分かり、提案命令セットアーキテクチャがレジスタを大量に指示可能なことが、当初の予想を超えた効果を持つことが明らかになった。

#### コンパイラ技術

独特な命令形式を持つ STRAIGHT アーキテクチャのためのコードを自動生成するコンパイラ技術の開発を行った。Single Static Assignment 形式の中間元後を持つ LLVM を用い、このポスト・コンパイラの形で実装を進めた。ソース・オペランドをその生成命令との距離で指定する方式のため、制御フローのバリエーションがあっても値が正しく指定できるアルゴリズムを開発し、正しく実行可能な Livermore ループを生成できることを確認した。

また、コンパイラ技術研究を進める過程で、本アーキテクチャのレジスタ指定方式は、データ構造のレジスタ・プロモーションや複数の値を返す関数など、充らにない最適化を可能とすることが新たに分かり、コンパイラ最適化技術の研究を行った。まずループコードに着目し、 $n$  イテレーションのループにまたがるデータ依存をレジスタ経由とする手法を提案・実装し、Livermore ループによる評価を行った。

#### マイクロアーキテクチャ技術

提案アーキテクチャが保持するレジスタ・ファイルのアクティブ電力削減のため、レジスタ・キャッシュを用いる手法を提案し、その置き換えアルゴリズムについて FIFO および LRU の比較検討を行い、いずれの方法も効果的であることを確認した。

また、投機メモリ・フォワーディング技術の適用が従来よりも効率的に行えることを示し、その実装方法を明らかにした。

#### プロセッサ 3 次元積層

高性能プロセッサを 3 次元化した時の効果を見積もるため、アーキテクチャレベルの 3D フロアプランナを新規開発し、傾向を調べた。作成したフロアプランナはシミュレーテッド・アニーリング手法を用い、フットプリントと通信電力とモジュール熱拡散の 3 要素を織り込んだ評価関数を最小化する。フロアプランナによる評価の結果、積層技術が進むほど、データバスや他ポート RAM の実装の制約が

緩和され、高性能プロセッサが構成しやすくなる見積もりが得られた。また、熱拡散の面からは、マルチコア構成のフロアプランは、2 次元のプロセッサを積み重ねるよりも、ホットスポットの拡散を優先した 3 次元プロセッサを並べる方が有利である予測が得られた。

#### キャッシュ制御技術

プロセッサにメモリ・データを供給するキャッシュ・メモリの制御技術について、アクセス予測を組み合わせてキャッシュ内の置き換えを理想に近づける pre-promotion、大容量キャッシュを活用するために超長期再参照データを保護する Stubborn cache を提案し、キャッシュ性能の改善を行った。またこれらの適応制御を的確に行うために、メモリアクセス傾向に着目してプログラムフェーズの移り変わりを判断する手法を提案し、従来手法よりも精度が向上することを確認した。

#### HDL 実装

詳細化した STRAIGHT アーキテクチャのパイプラインを HDL により記述し、ハードウェア実装可能であることを明らかにした。これらの研究成果により、本研究は次の段階に進められることが分かり、企業との共同研究を視野に入れて、システムレベルの評価をするためのフレームワーク作りや、ディープ・ラーニングなどデータ並列タスク実行の効率化検討などを進めているところである。

#### 5. 主な発表論文等

(研究代表者、研究分担者および連携研究者には下線)

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 20 件)

福田 晃史, 中江 哲史, 入江 英嗣, 坂井 修一: 「STRAIGHT アーキテクチャにおけるループ内ロード命令の削減手法」, 情報処理学会研究, 2017 年 3 月 9 日 ~ 2017 年 3 月 10 日 具志川農村環境改善センター (沖縄県島尻郡久米島町) 2017-ARC-225, No.3, pp. 1-6, Mar., 2017.

<http://id.nii.ac.jp/1001/00177854/>

洪江 陽人, 野村 隼人, 入江 英嗣, 坂井 修一: 「ストライドアクセスの階層構造に着目したフェーズ検出」, 情報処理学会研究報告, 2017 年 3 月 9 日 ~ 2017 年 3 月 10 日 具志川農村環境改善センター (沖縄県島尻郡久米島町) 2017-ARC-225, No.2, pp. 1-6, Mar., 2017.

<http://id.nii.ac.jp/1001/00177853/>  
酒井 一憲, 中江 哲史, 入江 英嗣, 坂井 修一: 「STRAIGHT における投機メ

モリフォワードイングの実装の検討」, 情報処理学会研究報告, 2017年1月23日~2017年1月25日 慶応大学来往舎(神奈川県横浜市) 2017-ARC-224, pp. 1-6, Jan., 2017.

<http://id.nii.ac.jp/1001/00176972/>  
Hayato Nomura, Hiroyuki Katchi, Hidetsugu Irie, Shuichi Sakai: "Stubborn Strategy to Mitigate Remaining Cache Misses", Int. Conf. on Computer Design, 2016年10月3日~2016年10月5日 Boston Marriott Newton(Phoenix,USA)pp.388--391, Oct., 2016.(査読有)

<http://ieeexplore.ieee.org/document/7753310/>

甲地 弘幸, 入江 英嗣, 坂井 修一: 「Pre-Promotion の動的切り替え手法の検討」, 情報処理学会研究報告 2016年8月8日~2016年8月10日 長野県松本文化会館(長野県松本市) Vol. 2016-ARC-221, No. 17, pp. 1-5, Aug., 2016.

<http://id.nii.ac.jp/1001/00172879/>  
中江 哲史, 入江 英嗣, 坂井 修一: 「STRAIGHT コンパイラにおける不要コードの削減手法の検討」, 情報処理学会研究報告, 2016年8月8日~2016年8月10日 長野県松本文化会館(長野県松本市) Vol. 2016-ARC-221, No. 5, pp. 1-6, Aug., 2016.

<http://id.nii.ac.jp/1001/00172867/>  
中江 哲史, 入江 英嗣, 坂井 修一: 「STRAIGHT アーキテクチャのためのコンパイラ技術」, 電子情報通信学会総合大会, 2016年3月15日~2016年3月18日 九州大学伊都キャンパス(福岡県福岡市) D-6-16, pp. 70, Mar., 2016.

[https://i-scover.ieice.org/iscover/page/ARTICLE\\_CONF\\_2016G\\_D-6-16](https://i-scover.ieice.org/iscover/page/ARTICLE_CONF_2016G_D-6-16)

赤木 晟也, 入江 英嗣, 坂井 修一: 「STRAIGHT アーキテクチャの HDL 実装と評価」, 電子情報通信学会総合大会, 2016年3月15日~2016年3月18日 九州大学伊都キャンパス(福岡県福岡市) D-6-15, pp. 69, Mar., 2016.

[https://i-scover.ieice.org/iscover/page/ARTICLE\\_CONF\\_2016G\\_D-6-15](https://i-scover.ieice.org/iscover/page/ARTICLE_CONF_2016G_D-6-15)

甲地 弘幸, 入江 英嗣, 坂井 修一: 「置き換えアルゴリズムとプリフェッチが協調動作するキャッシュマネジメント・プリプロモーションの評価」, 電子情報通信学会総合大会, 2016年3月15日~2016年3月18日 九州大学伊都キャンパス(福岡県福岡市) D-6-14, pp. 68, Mar., 2016.

[https://i-scover.ieice.org/iscover/page/ARTICLE\\_CONF\\_2016G\\_D-6-14](https://i-scover.ieice.org/iscover/page/ARTICLE_CONF_2016G_D-6-14)

入江 英嗣: 「ユーザに寄り添うシステムとそのコンピュータアーキテクチャ」,

電子情報通信学会技術研究報告(招待講演), 2015年11月17日~2015年12月18日 京都工芸繊維大学(京都府京都市) Vol. 115, no. 373, ICD2015-67, pp. 25, Dec., 2015.

[https://i-scover.ieice.org/iscover/page/ARTICLE\\_TECH\\_77839](https://i-scover.ieice.org/iscover/page/ARTICLE_TECH_77839)

Midoriko Chikara, Hidetsugu Irie, Makoto Sahoda, Masato Yoshimi and Tsutomu Yoshinaga: "Pre-Promotion: Synergizing Prefetching and Anti-thrashing Replacement Policy", Int. Conf. on Computer Design WIP session, 2015年10月18日~2015年10月21日 New York University(New York City, USA)Oct., 2015. (査読有)  
村田 篤志, 野村 隼人, 吉見 真聡, 入江 英嗣, 吉永 努, 坂井 修一: 「3次元積層プロセッサ向けフロアプランナの可視化」, 2015年10月8日 幕張メッセ(千葉県千葉市)電子情報通信学会技術研究報告, Vol.115, No. 243, pp. 63-65, Oct., 2015.

[http://i-scover.ieice.org/iscover/page/ARTICLE\\_TECH\\_75893](http://i-scover.ieice.org/iscover/page/ARTICLE_TECH_75893)

Hayato Nomura, Takuma Nakajima, Masato Yoshimi, Tsutomu Yoshinaga, Hidetsugu Irie: "Stubborn Cache: A Novel Strategy for Repeating Thrashing Access Patterns", COOL Chips XVIII poster, 2015年4月13日~2015年4月15日 Yokohama Joho Bunka Center at Yokohama, Japan pp. 19, Apr. 2015. (Featured Poster Award)

村田 篤志, 稲場 朋大, 吉見真聡, 入江 英嗣, 吉永 努: 「TSV モジュールの配置最適化アルゴリズムの提案」, 電子情報通信学会技術研究報告, 2015年3月6日~2015年3月7日 奄美市社会福祉協議会会議室(鹿児島県奄美市)Vol. 114, No. 506, pp. 43-48, Mar. 2015.

[http://i-scover.ieice.org/iscover/page/ARTICLE\\_TECH\\_72580](http://i-scover.ieice.org/iscover/page/ARTICLE_TECH_72580)

力翠湖, 吉見真聡, 吉永努, 入江英嗣: 「プリフェッチ精度に基づくキャッシュライン保持手法」, 情報処理学会研究報告, 2014年10月6日~2014年10月7日 ホテルニューツルタ(大分県別府市)Vol. 2014-ARC-212, No.2, Oct. 2014.

<http://id.nii.ac.jp/1001/00103168/>  
佐保田誠, 山中崇弘, 吉見真聡, 吉永努, 入江英嗣: 「STRAIGHT シミュレータによるループ実行の評価」, 情報処理学会研究報告, 2014年10月6日~2014年10月7日 ホテルニューツルタ(大分県別府市) Vol. 2014-ARC-212, No. 10, Sep. 2014.

<http://id.nii.ac.jp/1001/00103176/>  
野村隼人, 力翠湖, 吉見真聡, 吉永努, 入江英嗣: 「動的推定によるキャッシュ

パーティショニング最適化」, 情報処理学会研究報告, 2014年7月28日 朱鷺メッセ 新潟コンベンションセンター( ) 新潟市中央区) Vol. 2014-ARC-211, No. 2, Jul. 2014.

<http://id.nii.ac.jp/1001/00102238/>  
力翠湖, 眞島一貴, 藤原大輔, 吉見真聡, 吉永努, 入江英嗣: 「プリフェッチ情報から再参照予測を行うキャッシュライン置き換えアルゴリズム」, 情報処理学会研究報告 2013年7月31日~2013年8月1日 北九州国際会議場(福岡県北九州市) Vol. 2013-ARC-206, No. 20, Aug. 2013.

<http://id.nii.ac.jp/1001/00094555/>  
入江英嗣, 山中崇弘, 佐保田誠, 吉見真聡, 吉永努: 「もしILPプロセッサのレジスタファイルが分散キーバリューストアになったら」, 情報処理学会研究報告, 2013年7月31日~2013年8月1日 北九州国際会議場(福岡県北九州市) Vol.2013-ARC-206, No. 5, Aug. 2013.

<http://id.nii.ac.jp/1001/00094540/>  
稲場朋大, 放地宏佳, 藤原大輔, 眞島一貴, 吉見真聡, 入江英嗣, 吉永努: 「配線アクティビティを考慮した3次元積層プロセッサ向けフロアプランナのための熱評価手法」, 情報処理学会研究報告, 2013年1月31日~2013年2月1日 早稲田大学グリーン・コンピューティング・システム研究開発センター(東京都新宿区早稲田) Vol.2013-ARC-203, No. 12, Jan. 2013.

<http://id.nii.ac.jp/1001/00095278/>

〔図書〕(計 0件)

〔産業財産権〕

出願状況(計 0件)

取得状況(計 0件)

〔その他〕

ホームページ等

<http://www.mtl.t.u-tokyo.ac.jp/researches/>

6. 研究組織

(1)研究代表者

入江 英嗣 (Irie Hidetsugu)

東京大学・大学院情報理工学系研究科・准教授

研究者番号: 50422407

(2)研究分担者

分担者なし

(3)連携研究者

連携研究者なし