

## 科学研究費助成事業 研究成果報告書

平成 28 年 5 月 30 日現在

機関番号：13901

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330057

研究課題名(和文)高性能・低電力コンピュータの方式に関する研究

研究課題名(英文)Computer architecture for high performance and low power consumption

## 研究代表者

安藤 秀樹 (Ando, Hideki)

名古屋大学・工学(系)研究科(研究院)・教授

研究者番号：40293667

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：本研究では、プロセッサの性能向上を阻む壁の1つであるメモリの壁を打破する方式として、動的命令ウィンドウ・リサイジングと呼ぶ手法を提案した。本方式により、プロセッサの性能を21%向上させることに成功した。また、プロセッサの電力を削減する研究として、発行キュー2段階タグ比較手法を提案した。わずか1%の性能低下と引き換えに、発行キューで最大の電力を消費するタグ比較回路の電力を50%減少させることに成功した。

研究成果の概要(英文)：In this study, I have proposed a scheme called dynamic instruction window resizing to overcome the memory wall, which limits the performance improvement of processors. My evaluation results using SPEC2006 benchmark programs show that this scheme successfully improves the performance of a processor by 21% on average. We have also proposed a scheme called issue queue double-stage tag comparison to reduce the power consumption of a processor. My evaluation results using SPEC2006 benchmark programs show that this scheme successfully reduces the power consumption of the tag comparison circuit, which is a maximum consumer in the issue queue, by 50% with only 1% performance degradation on average.

研究分野：コンピュータ・アーキテクチャ

キーワード：コンピュータ・アーキテクチャ

## 1. 研究開始当初の背景

現在のコンピュータの性能向上は、「メモリの壁と電力の壁」という2つの壁により大きく阻まれている。実際、1つのプログラムを実行したときの性能(単一スレッド実行性能)は、この数年、ほとんど改善していない。

メモリの壁とは、プロセッサと主記憶の間の非常に大きな速度ギャップのことをいう。現在では、数百倍もの速度ギャップがある。この影響を避けるため、現在のプロセッサは、極力主記憶をアクセスせずにプログラムの実行を行えるよう大容量のキャッシュを備えている。現在では、その面積は、プロセッサ・チップの半分近くを占めているが、それでもなお十分ではなく、非常に大きな性能低下を被っている。

一方、電力の壁とは、次のようなものである。現在のコンピュータを構成するプロセッサは、非常に多くの電力を消費し、空冷の限界近くまで熱くなる。このため、クロック速度を向上させる単純な方法はいうまでもなく、これまで提案された多くのアーキテクチャ技術も電力を多く消費し、性能をあげることが困難となっている。

## 2. 研究の目的

(1) メモリの壁の問題に対し、メモリ・アクセスを並列化する手法について研究を行った。このとき、これにより、他の性能向上要素、たとえば、命令実行の並列化に悪影響を与えないように考慮した。また、限られた電力予算においても、高い効果を発揮するよう考慮した。

(2) 電力の壁の問題に対し、プロセッサの中でも多くの電力を消費する発行キューの電力を削減する手法について研究を行った。このとき、これにより、性能がほとんど低下しないよう考慮した。

## 3. 研究の方法

考案した方式の性能改善量を測定するため SimpleScalar ツール・セットと呼ばれるシミュレータを修正して用いた。SimpleScalar ツール・セットは、現在主流のスーパースカラ方式を採るプロセッ

サのマイクロアーキテクチャ上での動作を C 言語で記述したシミュレータである。世界の多くの研究者の間で用いられている。

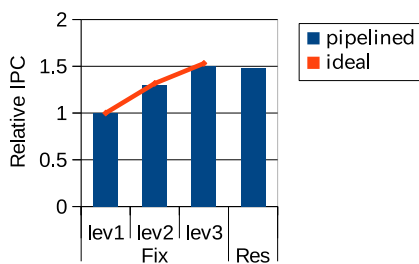
電力評価については、解析的に、あるいは、CACTI または McPAT と呼ばれるツールを修正して行った。これらのツールは、回路の種々の容量を計算し、種々の構成での電力を計算するツールである。

## 4. 研究成果

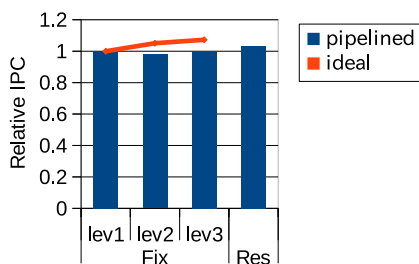
(1) メモリ・インテンシブなプログラムの単一スレッド性能を向上させることは難しい。その原因の1つとして、メモリの壁がある。メモリの壁とは、プロセッサと主記憶との間の大きな性能差のことをいう。この問題を克服するために、メモリ・レベル並列 (MLP: memory-level parallelism) の利用は有効である。MLP を利用する1つの方法は、積極的なアウト・オブ・オーダー実行である。これを実現するには、大きな命令ウィンドウ資源(リオーダー・バッファ、発行キュー、ロード/ストア・キュー)が必要となる。しかし、単純にこれらの資源を拡大すると、クロック速度を低下させてしまう。この問題は、これらの資源をパイプライン化すれば解決可能であるが、今度は、命令の発行に遅延が生じ、命令レベル並列 (ILP: instruction-level parallelism) の有効利用を妨げてしまう。その結果、計算インテンシブなプログラムの性能を著しく低下させてしまう。

この研究では、ウィンドウ資源を MLP が利用可能なときのみ拡大、パイプライン化し、ILP が利用可能なときは、それらを縮小し非パイプライン化する適応的な動的命令ウィンドウ・リサイジング手法 (DIWR: dynamic instruction window resizing) を提案した。本手法では、MLP が利用可能かどうかを最終レベル・キャッシュ・ミスの発生に基づいて予測し、ウィンドウ資源のサイズを変更する。本手法は非常に単純であり、また、ハードウェアの変更は、既存のプロセッサの構成によく適合するため極めて実用的である。

図1に評価結果を示す。評価プログラムとして



(a) メモリ・インテンシブ・プログラム



(b) 計算インテンシブ・プログラム

図 1: 基準プロセッサの性能に対する DIWR の性能向上率

SPEC2006 を用いた。同図 (a) は、メモリ・インテンシブなプログラムの平均であり、(b) は、計算インテンシブなプログラムの平均である。両図において、Fix の 3 本の棒グラフは、ウィンドウ・サイズを固定 (3 種類: 小さいものから lev1, 2, 3) した場合の性能向上率である。この場合、lev2, 3 では、資源をパイプライン動作させている。これに対して、赤色の折れ線は、パイプライン化していない、すなわち、ILP の損失がない理想の状態での性能を示す。Res は、提案の DIWR により適応的にウィンドウ・サイズを変更した場合の性能向上である。2つの図よりわかるように、メモリ・インテンシブなプログラムにも、計算インテンシブなプログラムにも、うまく適応し、ほぼ理想の性能を達成している。既存のプロセッサと比較した場合、ベンチマークの平均で 21% の性能向上を達成している。このとき、必要な追加コストは、既存のプロセッサ・コアのわずか 6%、あるいは、プロセッサ全体の 3% である。このコスト性能比は、

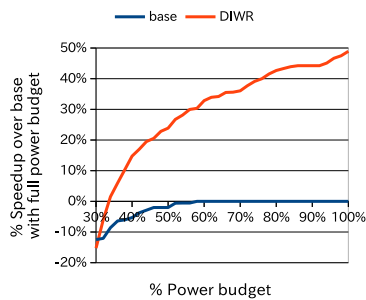
従来のトレンドであるポラックの法則に基づく性能向上率 (6%) をはるかに凌ぐものである。また、エネルギー効率においても、既存のプロセッサを 8% 改善することができた。

(2) 先に示したように、DIWR は、命令ウィンドウ・サイズを利用可能な並列性、すなわち、MLP か ILP に対応して変更し、それぞれを有効に利用することができる。(1) で DIWR は大きな性能向上を得ることができることを示したが、残念ながら、プログラムのメモリ・インテンシブ・フェーズでは消費電力は増加する。もしも、消費電力が何らかの要求によって定まる電力予算を超えたなら、DIWR プロセッサは、何らかの方法で電力を抑制する必要がある。これにより、達成できる性能は低下する。

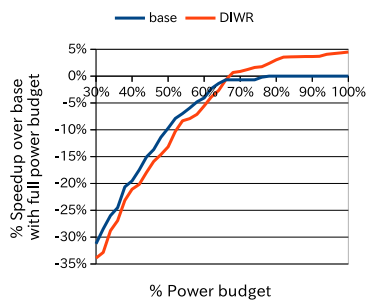
この研究では、電力削減の手法として、DVFS(dynamic voltage and frequency scaling) を仮定し、与えられた電力予算でどれほどの性能が達成できるかを調査した。

図 2 に、評価結果を示す。評価プログラムとして SPEC2006 を用いた。横軸は電力予算であり、DIWR が最も多く電力を消費するプログラムでの電力を 100% としている。縦軸は、電力予算が無制限の既存のプロセッサに対する性能向上率である。青色の折れ線グラフは、既存のプロセッサの、赤色の折れ線グラフは、DIWR の性能向上率を示している。図からわかるように、電力予算を低下させれば、性能も低下していく。しかし、メモリ・インテンシブ・プログラムにおいては、非常に電力予算が小さくない限り、既存プロセッサに対し、DIWR は依然として大きな性能向上を達成している。最も重要な測定点は、既存プロセッサの最大消費電力を電力予算とした場合の DIWR の性能である。この点において、DIWR は既存のプロセッサに対して、メモリ・インテンシブ・プログラムで 49%(全プログラムで 16%) の性能向上を達成することがわかった。

(3) プロセッサの全ての資源の中で、発行キューは、最も電力を消費する資源の 1 つである。その多



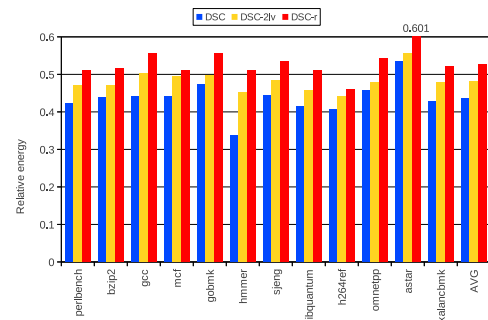
(a) メモリ・インテンシブ・プログラム



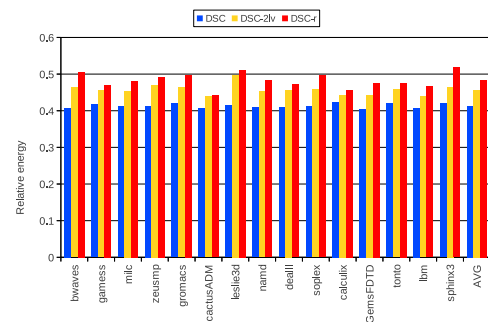
(b) 計算インテンシブ・プログラム

図 2: 与えられた電力予算に対する基準プロセッサの性能に対する性能向上率

くの電力は、ウェイクアップ論理で消費する。ウェイクアップ論理の中では、ソース・オペランドが用意できたかどうかをチェックするタグ比較器が最も多くの電力を消費する。この研究では、このタグ比較に要する電力を削減する2段階タグ比較方式(DSC: double-stage tag comparison)を提案した。この方式では、最初に、全エントリについてタグの下位ビットを比較する。そして、それが一致したエントリに対してのみ、上位ビットを比較する。タグ比較で消費する電力は、比較するタグのビット数におよそ比例するので、この数を削減すれば、消費電力も削減される。しかし、この直列比較は発行キューの遅延を増加させてしまい、その結果、クロック・サイクル時間を増加させる。クロック・サイクル時間の問題は、発行動作に余



(a) 整数プログラム



(b) 浮動小数点プログラム

図 3: タグ比較に要するエネルギーの既存回路に対する相対値

分に1サイクルを割り当てれば解決するが、今度は、IPCが悪化するという問題が生じる。この問題に対し、本研究では、2つの拡張方式を提案した。1つは、2階層発行キュー方式であり、もう1つは後方1段化方式である。これらの方式は、性能に悪影響を与える可能性のある少数の命令が保持されているエントリのタグ比較回路を動的に変更し、タグ比較を1段化する方式である。

図3に、タグ比較に要するエネルギーの既存回路に対する相対値を示す。ベンチマークは、SPEC2006であり、同図(a)、(b)は、それぞれ、整数プログラム、浮動小数点プログラムにおける相対エネルギーである。DSC、DSC-2lv、DSC-rの3つの棒グラフは、すべてDSCを適用した場合であり、それぞれ、拡張方式適用なし、2階層方式の拡張、後方1段化方式の拡張を適用した場合

である。タグの2段階比較に2階層方式を適用した場合、0.91%の性能低下でタグ比較の消費エネルギーを従来の46.8%に、後方1段化方式を適用した場合、0.28%の性能低下で消費エネルギーを50.3%に削減できることを確認した。

## 5. 主な発表論文等

[雑誌論文](計3本)

①H. Ando and R. Shioya, “Performance of Dynamic Instruction Window Resizing for a Given Power Budget under DVFS Control,” IEICE Transactions on Information and Systems, Vol.E99-D, No.2, pp.341-350, February 2016. (査読あり、<https://search.ieice.org/>)

②Y. Kora, K. Yamaguchi, and H. Ando, “MLP-Aware Dynamic Instruction Window Resizing in Superscalar Processors for Adaptively Exploiting Available Parallelism,” IEICE Transactions on Information and Systems, Vol.E97-D, No.12, pp.3110-3123, December 2014. (査読あり、<https://search.ieice.org/>)

[学会発表](計15本)

①Y. Kora, K. Yamaguchi, and H. Ando, “MLP-Aware Dynamic Instruction Window Resizing for Adaptively Exploiting Both ILP and MLP,” In Proceedings of the 46th Annual International Symposium on Microarchitecture, pp.37-48, December 2013. (12/09/2013, Davis, CA, USA)

②小林誠弥, 塩谷亮太, 安藤秀樹, “タグの2段階比較による発行キューの消費エネルギー削減,” 2013年先進的計算基盤システムシンポジウム SACSIS 2013, pp.2-9, 2013年5月. (5/22/2013, 仙台国際センター, 仙台)

## 6. 研究組織

### (1) 研究代表者

安藤秀樹 (Hideki Ando)、名古屋大学大学院工学研究科 電子情報システム専攻、教授、研究者番号:40293667