

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 5 日現在

機関番号：37111

研究種目：挑戦的萌芽研究

研究期間：2011～2013

課題番号：23650026

研究課題名(和文) 先端記憶デバイスを利用する記憶階層の再構築に関する研究

研究課題名(英文) A Study on Reconstructing Memory Hierarchy that Utilizes Emerging Devices

研究代表者

佐藤 寿倫 (SATO, Toshinori)

福岡大学・工学部・教授

研究者番号：00322298

交付決定額(研究期間全体)：(直接経費) 2,700,000円、(間接経費) 810,000円

研究成果の概要(和文)：依然として上昇している要求性能に答えることが可能な、プロセッサ性能の向上が達成されている。しかしシステム全体を眺めると、十分な性能を提供できているとは言い難い。メモリが足枷となっている。加えて、メモリはエネルギー消費量が大きい点でも問題である。このような問題意識から、本課題を実施した。

近い将来に磁気抵抗メモリ等の次世代メモリが利用可能になるという仮定の下で、アーキテクチャ上の工夫により高性能・低電力・高信頼なメモリシステムを実現するために、従来の記憶階層を解体して新たに構築することを目標とした。

提案する記憶階層を用いると最良の場合で、エネルギー遅延積を49%改善できることが確認できた。

研究成果の概要(英文)：Processor performance is still improved in order to achieve increasing performance demand. However, computer system can not provide us with enough performance to meet the demand. This is due to the poor performance of memory system. In addition, memory consumes huge power. Considering the above, we started our project.

We assume that we can use next-generation memory devices such as STT-MRAM. Under this assumption, we try to reconstruct memory hierarchy in order to achieve high-performance, low-power, and dependable memory system.

The proposed memory system can improve energy-delay product by up to 49%.

研究分野：総合領域

科研費の分科・細目：情報学，計算機システム・ネットワーク

キーワード：計算機システム システムオンチップ ハイパフォーマンス・コンピューティング フラッシュメモリ
相変化メモリ

1. 研究開始当初の背景

温暖化ガス排出量削減の野心的な目標設定により、エネルギー消費量削減は喫緊の技術課題となった。中でも IT 分野のエネルギー消費量は年を追って増加しており、ガートナーによるとデータセンターが排出する温暖化ガスの量は航空業界に匹敵する。クラウドコンピューティングの急速な浸透を鑑みると、データセンターの規模と数はともに増大を続け、IT 機器のエネルギー利用効率改善が強く求められるであろうことは容易に想像できる。要求性能の提供はもはやコンピュータシステム構築の第一の目標とは成り得ず、エネルギー消費量あるいは温暖化ガス排出量の削減が支配的な目標と成りつつある。つまり、実装技術よりも運用技術の方が重要性を増している。

コンピュータシステムに要求される性能は依然として上昇しているが、幸いなことに、それに応えることが可能なプロセッサ性能向上も達成されている。しかし、コンピュータシステム全体を眺めると、十分な性能を提供できているとは言い難い。記憶装置が足枷となっている。加えて、記憶装置はエネルギー消費量が多い点でも問題である。すなわち、記憶装置こそがコンピュータシステムの運用上解決すべき課題を提示している。

デバイス技術に目を向けると、既に市民権を得ているフラッシュに加えて、実用化が現実味を増してきた相変化メモリ (PCM)、磁気抵抗メモリ (STT-MRAM) や強誘電体メモリ (FeRAM) 等、様々な先端記憶デバイスが登場している。これらは、従来の SRAM や DRAM、そして磁気ディスク等が備えていない機能を提供する。不揮発性・高速性・高密度・低電力等、これらのデバイス特性を有効活用できれば、省エネルギーかつ高速で大容量といった夢の記憶装置を構築できる道が開けるに違いない。

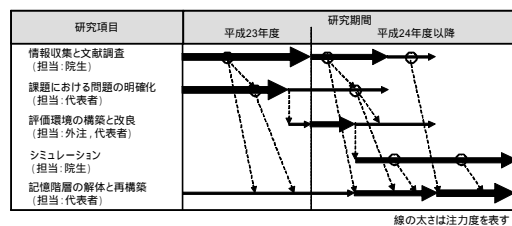
まとめると、記憶装置の構築という最適化問題において、目的関数・制約条件・解空間の全てが、古典的なコンピュータアーキテクチャの想定から変化したわけである。既存の記憶階層を解体し、新たに再構築することが喫緊の課題であることが理解出来る。

2. 研究の目的

近い将来に PCM や STT-MRAM 等の次世代メモリデバイスが利用可能になるという仮定の下で、アーキテクチャ上の工夫により高性能・低電力・高信頼なメモリシステムを実現するために、従来の記憶階層を解体して新たに構築することを目標とする。その意義は記憶階層の解体と再構築を検討する価値があるかどうかを明らかにすることにある。短期的にはフラッシュや PCM 等に共通する問題を解決し、高速・低電力・高信頼な記憶階層の構築を目標とする。

3. 研究の方法

代表者と 1~2 名の院生で研究体制を整える。研究課題における問題の整理とそれらの解決策の考案は、代表者の分担とする。それらのための材料を揃えることを院生の分担とする。すなわち、研究会などでの情報収集や文献調査に従事することが院生の役割である。これらの結果を簡潔かつ必要十分なレポートとして院生がまとめ、それらから得られる知見をもとに、代表者が問題を整理し解決策を考案する。また、考案した解決策を評価するために、計算機シミュレーションを実施する。新規な記憶階層を提案するため、評価環境も新たに構築する必要がある。シミュレーションの実施は院生の役割である。



4. 研究成果

平成 23 年度は情報収集と文献調査の実施に注力し、国際会議の会議録を中心に調査を実施した。調査結果に基づき、本課題の研究推進方針を決定した。次世代メモリとして PCM、STT-MRAM、FeRAM を比較検討し、動作の高速性、書き換え耐性、消費電力の項目で他を凌いでいる STT-RAM を選択した。

STT-RAM は次世代メモリの中では高性能であるが、現在キャッシュとして使用されている SRAM と比較すると、書き込みに非常に大きな処理時間を要する。加えて、STT-RAM は不揮発性であるため待機時の消費電力は無視できるほどに小さいが、書き込み処理に要する消費電力が極めて大きい。つまり、書き込みにおける処理時間と消費電力に課題がある。

上記の課題を勘案すると、高速性が最重要視される一次キャッシュに STT-MRAM を採用することは検討に値しない。そこで本研究では、二次キャッシュおよび三次キャッシュに STT-MRAM を採用することを検討した。

平成 24 年度には、前年度の実施状況報告書で掲げた最優先課題であった評価環境の構築を行った。オンチップメモリの記憶階層を評価するための環境をほぼ立ち上げることが出来た。

基盤となるシミュレータには、ウイスコンシン大らで開発されたプロセッサシミュレータ gem5 と、ヒューレット・パッカード社で開発された電力見積りシミュレータ McPAT とを選択した。いずれのシミュレータにも、STT-MRAM の書き込み特性を反映できるように、我々の手で改良を加えている。

平成 25 年度は、前年度に構築した環境を用いてシミュレーション実験を行った。

評価モデルは 1 コア・1 スレッド実行の A

ウトオブオーダ・プロセッサであり、2GHz 動作で同時に 8 命令の実行が可能であると仮定している。キャッシュは一次、二次、三次でそれぞれ 32KB、256KB、8MB の容量としている。評価した記憶階層は、全てのキャッシュが SRAM、三次のみ STT-MRAM で残りが SRAM、一次のみ SRAM で残りが STT-MRAM の三通りである。ベンチマークには SPEC CPU2006 を選択した。

以下に、429.mcf のシミュレーション結果を示す。図 1 は消費電力である。三本のグラフは左から順に階層の結果で、階層の値で正規化されている。STT-SRAM を採用すると、静的電力が大いに削減できることが期待される。現在のプロセッサでは大容量のキャッシュが消費する静的電力が深刻であるため、このことによる効果が大きく、階層とでそれぞれ 76%と 81%の電力削減が達成できた。

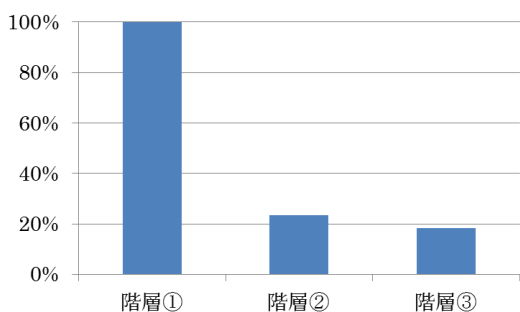


図 1 消費電力

図 2 は実行時間である。ここでも三本のグラフは左から順に階層の結果で、階層の値で正規化されている。一次キャッシュと比較すると、二次キャッシュや三次キャッシュにアクセスする頻度は小さいが、SRAM と比べて STT-MRAM 書き込み時間は 10 倍大きい。そのため、その影響が無視できないことが分かる。階層で約 1.8 倍に、階層では約 2.8 倍もの実行時間増大となった。この問題を解決することが喫緊の課題である。

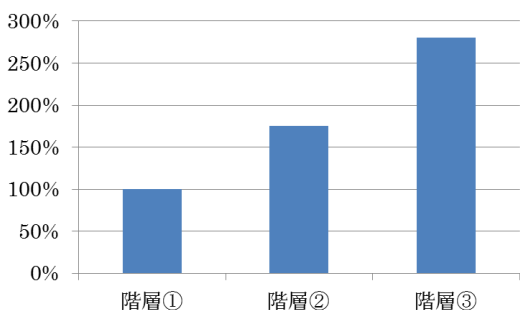


図 2 実行時間

図 3 はエネルギー遅延積 (EDP: Energy-Delay Product) である。一般に消費電力とコンピューティング性能はトレード・オフの関係にあるため、EDP を指標として用いる。

今回も同様に三本のグラフは左から順に階層の結果で、階層の値で正規化されている。階層では消費電力削減の効果が大きく EDP を 49%改善出来ているが、階層では実行時間のペナルティを無視できず EDP は 20%改悪された。

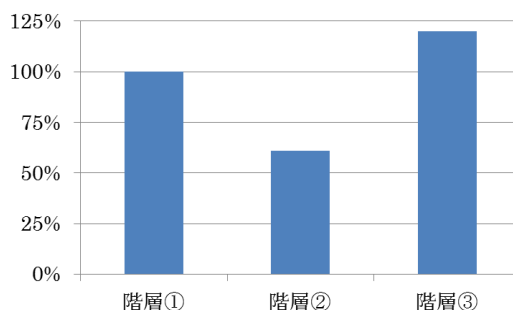


図 3 エネルギー遅延積

残念ながら好ましい結果が得られなかったため机上で考察した。その結果、以下の問題が明らかになった。すなわち、既存の SRAM 等の次世代メモリデバイスへの置き換えによる階層や容量の最適化だけでは大幅な性能改善は不可能、低速かつ大電力なデータ書き込みが次世代デバイスへの置き換えを阻害、メモリデバイスの特性改善だけでは不十分でメモリリクエストのスケジューリング改善なども必要、である。つまり、記憶階層の再構築に価値は認められるが、次世代メモリの特性を活かすことの出来ない利用法では期待される大きな性能改善を達成できない、ということである。

次世代デバイスを採用する新しい記憶階層構造には、同時に新しい利用法が必要であると考えられる、というのが結論である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表](計 9 件)

枝元正寛, 林田隆則, 佐藤寿倫, 2 命令インオーダ発行仕様のスーパースカラプロセッサの設計, 情報処理学会研究報告 ARC, Vol. 2014-ARC-208, No. 2, 1 page, 2014 年 1 月 23 日.

吉田康洋, 林田隆則, 佐藤寿倫, 高性能・省電力プロセッサ向けメモリスケジューリング手法の調査, 2013 年度電子情報通信学会九州支部学生講演会, D-6, 1 page, 2013 年 9 月 23 日.

村上隆俊, 林田隆則, 佐藤寿倫, 巡回セールスマン問題を題材とした OpenCL による並列化, 情報処理学会九州支部火の国情報シンポジウム, 5 pages, 2013 年 3 月 14 日.

Yuji Kunitake, Toshinori Sato, Hiroto Yasuura, and Takanori Hayashida, Guidelines for Mitigating NBTI Degradation in On-chip Memories, 12th

International Symposium on Communications and Information Technologies, pp.827-832, October 5 2012.

Yoshimi Shibata, Takanori Hayashida, Toshinori Sato, and Shinya Takahashi, Simultaneous Dynamic and Static Power Reduction Utilizing Power Heterogeneous Functional Units, 27th International Technical Conference on Circuits/Systems, Computers and Communications, 4 pages, July 17 2012.

Toshinori Sato, Takanori Hayashida, and Ken Yano, Dynamically Reducing Overestimated Design Margin of MultiCores, 10th International Conference on High Performance Computing & Simulation, pp.403-409, July 4 2012.

田端隼人, 林田隆則, 佐藤寿倫, 高橋伸弥, 分岐予測器の状態数とエントリ数のトレードオフに関する考察, 2012年電子情報通信学会総合大会 ISS 特別企画「学生ポスターセッション」, ISS-SP-042, pp.197, 2012年3月22日.
柴田善水, 林田隆則, 佐藤寿倫, 高橋伸弥, スーパスカラ型CPUコアの低消費電力化手法の提案, 2012年電子情報通信学会総合大会 ISS 特別企画「学生ポスターセッション」, ISS-SP-043, pp.198, 2012年3月22日.

小川哲平, 林田隆則, 佐藤寿倫, GPUを用いた巡回セールスマン問題の並列化解法, 2012年電子情報通信学会総合大会 ISS 特別企画「学生ポスターセッション」, ISS-SP-044, pp.199, 2012年3月22日.

〔その他〕

ホームページ等

<http://uarch.jimdo.com/>

6. 研究組織

(1) 研究代表者

佐藤寿倫 (SATO, Toshinori)

福岡大学・工学部・教授

研究者番号：00322298