

機関番号：12612

研究種目：若手研究（B）

研究期間：2009年度～2010年度

課題番号：21700055

研究課題名（和文）マルチパスメモリ構成によるマルチコアプロセッサの高性能化技術の研究

研究課題名（英文）A study on a Multi-path Memory for High-Performance Multi-core Processors

研究代表者

近藤 正章 (MASAAKI KONDO)

電気通信大学・大学院情報システム学研究科・准教授

研究者番号：30376660

研究成果の概要（和文）：複数メモリアクセスパスを構成したプロセッサにおいて、プログラムの特徴に応じて動的にメモリアクセスを各パスに割り当てる手法を考案した。最新規格の主記憶メモリでは、バンド幅でなくレイテンシに性能が制約され、複数メモリパスの恩恵が十分に得られないことがわかった。今後のコア数増加により、メモリアクセスの多いアプリケーションを複数実行させた場合は、本手法のバンド幅拡大の効果により性能が大きく改善すると期待される。

研究成果の概要（英文）：We study a multi-path main memory technique and a dynamic path-allocation mechanism to overcome the memory-wall problem for high-performance multi-core processors. Assuming high-end DRAM devices, we could not obtain enough performance improvement with our method since the performance limiter is not memory access bandwidth but latency. However, we expect a benefit of the method for future many-core processors due to its extended memory bandwidth.

交付決定額

（金額単位：円）

| | 直接経費 | 間接経費 | 合計 |
|--------|-----------|-----------|-----------|
| 2009年度 | 2,100,000 | 630,000 | 2,730,000 |
| 2010年度 | 1,300,000 | 390,000 | 1,690,000 |
| 年度 | | | |
| 年度 | | | |
| 年度 | | | |
| 総計 | 3,400,000 | 1,020,000 | 4,420,000 |

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、情報システム、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、メモリシステム

1. 研究開始当初の背景

近年のマイクロプロセッサは、消費電力や設計コスト増大などの問題から、クロック周波数の向上に頼らずに、並列・並行処理を行うことで高い性能が得られるマルチコアプロセッサが主流となっている。マルチコア化により、チップあたりの演算性能が大きく向上する反面、チップ外にある主記憶とのデータ転送性能が相対的に低下し、メモリアクセ

スがシステム全体の性能向上を阻害するという「メモリウォール問題」が深刻化する。半導体デバイスの将来動向を予測した ITRS ロードマップによると、今後はチップ内に搭載されるコア数の増加が見込まれる一方、物理的な限界から、チップの I/O ピン数はほとんど増加しないとされている。そのため、このままでは演算性能に対してデータ供給バンド幅が絶対的に不足し、将来的なマイクロ

プロセッサの性能向上は望めない。これまでも、メモリウォール問題の緩和を目的とした研究が多く行われてきた。たとえば、DDR (Double-Data-Rate) デバイスを用いた DRAM バンド幅の増強、キャッシュメモリの大型化やアクセススケジューリングによるメモリバンド幅の有効利用等である。しかし、既存技術にも限界があり、今後の演算処理能力の向上にはとうてい追いつかないと考えられる。

2. 研究の目的

これまでも、キャッシュメモリの有効利用技術や、コア間の主記憶バンド幅割り当てを最適化するためのプロセススケジューリング技術など、メモリアクセスに着目してプロセッサの性能や消費電力効率を改善する研究が行われている。その中で、現状の主記憶バンド幅のもとでコア数を増加させると、主記憶バス競合の影響などとも相まって、最悪の場合、コア数を増加させると逆に性能が悪化してしまう可能性もあることが指摘されている。このように、メモリウォール問題は深刻であり、半導体技術の進歩の恩恵を享受しつつ、将来に渡ってプロセッサの性能向上を維持するためには、本質的な主記憶バンド幅の向上と、それを最大限に活用できる新しいアーキテクチャ技術が必要である。そこで、ボードを経由した I/O ピンからのデータ供給だけでなく、種々のデバイス技術、特にチップ間無線転送技術を応用し、異なる経路からデータを供給することで本質的にバンド幅を向上させ、メモリウォール問題の解決を目指す「マルチパスメモリ構成によるマルチコアプロセッサの高性能化技術」を研究開発することが目的である。

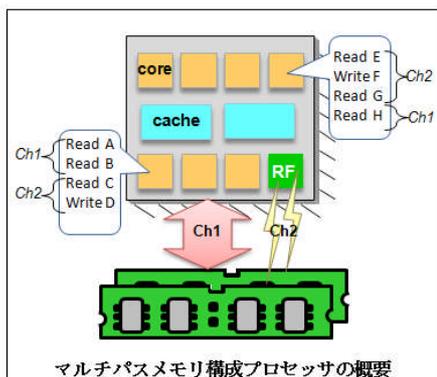


図 1 : マルチパスメモリ構成プロセッサ

3. 研究の方法

図 1 に示すようなマルチパスメモリ構成により、マルチコアプロセッサの性能向上の可能性を明らかにするため、主記憶データ供給に関するデバイス技術の調査、マルチパスメモリ構成を評価可能なシミュレーション

環境の構築、マルチコアプロセッサにおけるアクセスの特徴解析、データの特徴に応じたアクセスパスへの割り当てアルゴリズムの開発、などの要素技術を開発した上で、最終的に本手法の有効性を示すために各要素技術を統合し動的にマッピングを最適化した場合の性能評価が必要となる。以下、それぞれについて簡単に述べる。

(1) 主記憶データ供給に関する半導体デバイス技術調査

複数のメモリアクセスパスを構成するための各種デバイス技術に関して調査を行い、レイテンシやバンド幅、1 チップに対し接続可能なチャンネル数といった、主記憶アクセスパスに適用した場合の性能パラメータを明らかにする。

(2) マルチパス主記憶構成を評価可能なシミュレーション環境の構築

マルチパスメモリ構成を適用した際のマルチコアプロセッサの性能について、メモリアクセス特性データの取得や割り当てアルゴリズムの開発、性能の評価などに用いるためのシミュレーション環境を構築する。この際、米国ウィスコンシン大学で開発された単一プロセッサ評価用のシミュレータである SimpleScalar をベースにして開発する。

(3) アプリケーションデータ毎の主記憶アクセス特徴解析

種々のプログラムの各メモリアクセスについて、レイテンシセンシティブリティや、連続アクセスの度合いといった主記憶アクセスの特徴を、構築したシミュレーション環境を用いて解析し、マルチコア上でアクセス特性が性能に及ぼす影響を定量的に評価・解析する。マルチコア環境では、キャッシュやメモリバスなどシステム中の資源が密に結合されていることが多く、メモリアクセスの特徴は、各プログラム自身を持つものだけでなく、同時に実行される他のプログラムとの相互作用によっても変化する。したがって、単純に特徴を解析することは難しい。そこで、本研究では、種々のプログラムの組み合わせについて、メモリバンド幅やレイテンシのパラメータを変化させつつシミュレーションを行い、網羅的なデータを取得した上で、統計的にそのデータを解析する手法を用いる。

(4) 特徴に応じて動的にデータを複数パスに割り当てるアルゴリズムの開発

解析した種々のプログラムにおけるメモリアクセスの特性を参考に、各データを動的に複数のメモリパスに対して割り当てるアルゴリズムを開発する。

(5) 統合評価

開発した複数メモリパスへの動的割り当てアルゴリズムを、種々アプリケーション、および種々のマルチコアプロセッサ構成に対して評価し、本手法の有効性を評価する。また、評価結果をアルゴリズムのパラメータ最適化などにフィードバックすることで、アルゴリズムを改良する。

4. 研究成果

(1) 主記憶データ供給に関する半導体デバイス技術

まず、複数のメモリアクセスパスを構成するための技術として、チップ間無線伝送技術、および光伝送技術に関しての調査を行った。特に、主記憶アクセスパスに適用した場合の性能パラメータを明らかにする必要があるため、レイテンシやバンド幅、1チップに対するチャンネル数といった点を中心に調査を行った。特に極近接のチップ間通信技術を用いると、1Tbps 程度の速度以上で通信できるとされており、現状のメモリの大域に比べても遜色ない転送速度である。しかし、チップ間の距離に応じて転送速度が急激に低下することから、実装方法にも大きく依存することがわかった。そこで、まずは現状のDDR3-SDRAM メモリと同程度の転送速度を仮定して本技術を検討することにした。

(2) 主記憶アクセスの特徴解析

複数データ転送経路の効率的利用を図る上では、実行するプログラムの特徴を考慮したアクセスパスの割り当てが重要となる。まず、既存のシミュレーション環境上で、SPEC CPU2006 ベンチマークを用い、種々のデバイス特性を仮定し、プログラムの性能との関係性について解析を行った。この際、SimpleScalar プロセッサシミュレータをベースに開発され、DDR3-SDRAM などのメモリシステムを詳細に評価可能なプロセッサシミュレータである Zesto を用いた。まず、マルチパス構成による性能向上の可能性に関し評価を行った。ここでは、マルチパス構成によるバンド幅の向上効果をFSBクロック周波数の変更により仮想的に評価を行った。図2に評価を示す。図2より、メモリアクセスのバンド幅がある程度以上になっても、性能に大きな変化がないことがわかる。特に近年のDRAM 技術ではベースのクロック周波数が400MHz 超程度であり、そこから2倍の転送クロック(すなわち2つのパスを設けた場合と同様の効果)にしたとしても、ほとんど性能に変化がない。これは近年のDRAM の性能を仮定した場合、バンド幅が大きなボトルネックにならないことを示している。

次に、無線技術等を用いたメモリアクセス

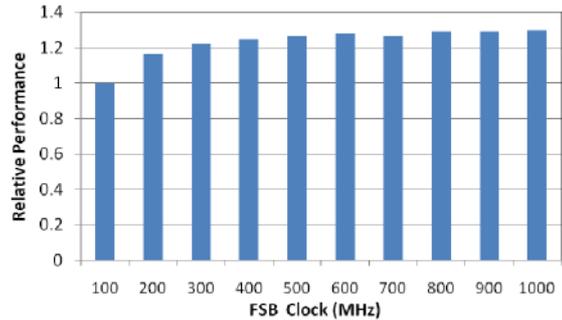


図2：FSB周波数の違いによる性能変化

パスを構成した際に想定されるレイテンシ増大の影響を評価するため、同様の評価環境を用いてメモリアクセスのレイテンシを変えて評価を行った。結果を図3に示す。

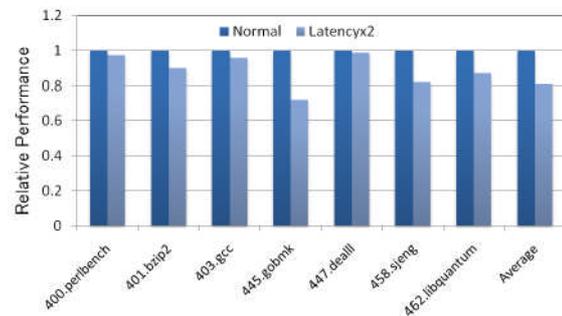


図3：レイテンシ増大の影響

図3は、通常のDRAMアクセスレイテンシの場合(Normal)に対し、レイテンシが2倍に増大したと仮定した場合(Latencyx2)の結果を示している。図より、レイテンシが増大すると、性能が大きく低下するアプリケーションがあることがわかる。そのため、あるメモリアクセスをレイテンシが大きなアクセスパスに割り当てる際にはこの性能低下に注意しなければならないことがわかる。

(3) 複数パスへの割り当て手法

複数のメモリアクセスパスを構成したプロセッサにおいて、各コアで実行するプログラムの特徴に応じて動的にメモリアクセスを複数パスに割り当てる必要がある。そこで、最終レベルキャッシュのミス率情報を基に、基本的にはミス率の高いプロセスに対してバンド幅の高いチャンネルを割り当てる手法を考案した。本手法の有効性を評価するために、アプリケーションとしてSPEC CPU2006 ベンチマークプログラムを、また評価環境としてZestoシミュレータをベースに開発したシミュレーション環境を用い、いくつかのマルチコアプロセッサ構成に対して評価を行った。その結果、DDR3-SDRAM規格の主記憶メモリを仮定すると、数コア程度のマルチコアプロセッサの場合、図2や図3の図からも想像できるように、バンド幅ではなく

メモリアクセスレイテンシに性能が制約され、複数のメモリアクセスパスを利用することの恩恵が十分に得られないことがわかった。

(4) まとめ

今後は、1チップに搭載されるコア数が数十から数百コアに増加すると予想されている。そのようなメニーコアプロセッサでは、バンド幅の側面においてもメモリアクセスのボトルネックが深刻化すると考えられる。そのような環境において、メモリアクセスが多いアプリケーションを複数実行させた場合は、複数メモリアクセスパスによるバンド幅拡大の効果により、性能が改善することが予備評価でわかっている。今後、そのようなメニーコアプロセッサが主流になると、提案手法の有効性が高まっていくと考えられる。

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)

[雑誌論文] (計4件)

- ① 佐々木広, 高木紀子, 近藤正章, 中村宏, “共有資源の競合を考慮したチップマルチプロセッサ向け低消費電力化手法”, 情報処理学会論文誌 コンピューティングシステム, Vol. 4, No. 2, pp. 40-58, 2011, (査読有り).
- ② 関直臣, レイジャオ, 小島悠, 池淵大輔, 長谷川揚平, 大久保直昭, 武田晴大, 香嶋俊裕, 白井利明, 宇佐美公良, 砂田徹也, 金井遵, 並木美太郎, 近藤正章, 中村宏, 天野英晴, “MIPS R3000 プロセッサにおける細粒度動的スリープ制御の実装と評価”, 電子情報通信学会論文誌 D, Vol. J-93D, No. 6, pp. 20-930, 2010, (査読有り)
- ③ 近藤正章, 高木紀子, 中村宏, “Pipeline Blocking: 走行時パワーゲーティングのための命令実行制御手法”, 情報処理学会論文誌 コンピューティングシステム, Vol. 2, No. 3, pp. 40-58, 2009, (査読有り)
- ④ H. Sasaki, M. Kondo, H. Nakamura, “Energy-Efficient Dynamic Instruction Scheduling Logic Through Instruction Grouping”, IEEE Transactions on VLSI, Vol. 17, No. 6, pp. 848-852, 2010, (査読有り)

[学会発表] (計10件)

- ① 穂園智哉, 近藤正章, 平澤将一, 本多弘樹, “機械学習により抽出されたアプリケーションの特徴を利用したタスク配置の検討”, 情報処理学会計算機アーキテク

チャ研究会, 2011年3月11日, 高知工科大 (高知県)

- ② 山藤友紀, 近藤正章, 平澤将一, 本多弘樹, “走行時パワーゲーティングにおけるスラック解析を用いた動的命令スケジューリングの検討”, 情報処理学会計算機アーキテクチャ研究会, 2011年3月11日, 高知工科大 (高知県)
- ③ L. Zhao, D. Ikebuchi, Y. Saito, M. Kamata, N. Seki, Y. Kojima, H. Amano, S. Koyama, T. Hashida, Y. Umahashi, D. Masuda, K. Usami, T. Sunata, K. Kimura, M. Namiki, S. Takeda, H. Nakamura, M. Kondo, “Geysers-1 and Geysers-2: MIPS R3000 CPU Chips with Fine-grain Runtime Power Gating”, IEEE Symposium on Low-Power and High-Speed Chips, 2010年4月16日, 横浜情報文化センター (神奈川県)
- ④ N. Takagi, H. Sasaki, M. Kondo, H. Nakamura, “Cooperative Shared Resource Access Control for Low-Power Chip Multiprocessors”, 14th ACM/IEEE International Symposium on Low Power Electronics and Design, 2009年8月20日, 米国・San Francisco
- ⑤ 近藤正章, 薦田登志矢, 佐々木広, 中村宏, “コンパイラによる細粒度スリープ制御のためのアーキテクチャ支援技術の検討”, 情報処理学会計算機アーキテクチャ研究会, 2009年8月4日, フォレスト仙台 (宮城県)

[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.hpc.is.uec.ac.jp/kondo>

6. 研究組織

(1) 研究代表者

近藤 正章 (MASAAKI KONDO)

電気通信大学・大学院情報システム学研究科・准教授

研究者番号: 30376660

(2) 研究分担者

なし

(3) 連携研究者

なし