

2パス限定投機システムにおける効率的なメモリアクセスによる省電力化

菅野 智之[†] 十鳥 弘泰[†] 大津 金光[†] 横田 隆史[†] 大川 猛[†]
[†]宇都宮大学大学院工学研究科

1 はじめに

プロセッサ技術の進歩により複数のプロセッサを持つ計算機が登場し、複数のプロセッサを有効活用する並列実行手法が提案されている。我々は、ループの実行経路(パス)の実行頻度が高い2本のパスを選定し、並列実行時にそのどちらかを予測し、投機的にマルチスレッド実行する2パス限定投機システム PALS を開発している [1].

また、PALS 独自機能の追加により逐次実行よりも消費電力量が増加し、電力効率が悪くなる可能性がある。したがって、実行性能だけでなく消費電力を評価する必要がある。我々はこれまで、逐次実行時と並列実行時の PALS 全体の消費電力量と実行サイクル数、電力効率を示すエネルギー遅延積を比較して、逐次実行よりも電力効率が悪い事を明らかにした [2].

本研究では PALS の電力効率を改善することが目的である。そのため、PALS 中の消費電力量が高いユニットに着目し、実行サイクル数を増加させずに省電力化する手法を提案する。

2 2パス限定投機システム PALS

PALS はプログラム内のループの実行頻度が高い2つのパスの実行に特化した投機コードを予め用意しておき、プログラム実行時にどちらかのパスを予測し、予測したパスを投機的にマルチスレッド実行するシステムである。

図1は PALS の構成を示している。TMU (Thread Management Unit) は、投機実行時に次に実行するパスの予測を行うユニットである。TU (Thread Unit) は TMU によって予測されたパスの実行を行うユニットである。MB (Memory Buffer) は各 TU が持つ投機的なメモリアクセス機構である。また、MB の補助的な記憶装置 LS (Load Shelter) がある。

我々は、PALS の各ユニットの消費電力量についての分析を行った結果、PALS 全体の消費電力量の中で MB の消費電力量が高いことを明らかにした [2].

MB の機能の1つにメモリ依存違反への対応がある。メモリ依存違反とはあるアドレスのデータを書き込む前に投機しているスレッドが同じアドレスのデータを読み込むことで古いデータをロードすることである。MB はこれに対応するためにデータのロードを行う場合に経由した MB にデータのコピーを保存する。ストアによってデータのコピーを更新する場合はメモリ依

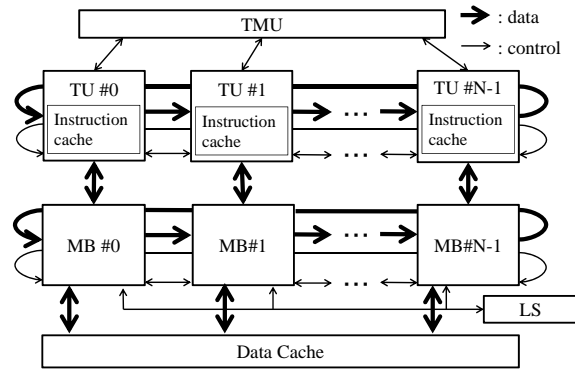


図 1: PALS の構成図

存違反であると判定され、それ以降のスレッドを再実行させる。

図2にロード命令が実行時の処理を示す。最後尾スレッドを実行している TU#3 がアドレス X のデータをロードする命令を実行した場合、メモリアクセス要求を MB#3 に出す。MB は要求するデータがなければ、先行するスレッドの MB にメモリアクセス要求を出す。先頭スレッドの MB にも要求するデータがなければデータキャッシュにメモリアクセス要求を出す。そしてデータキャッシュからロード要求する TU までデータを送信する。このとき、経由したすべての MB にデータのコピーを保存する。もし先行スレッドのストアによって、データのコピーを更新する場合は、メモリ依存違反と判定し、後続のスレッドの実行を破棄し再実行させる。しかし、ロード時に経由するデータのコピーを保存する処理はすべてのロード要求に対して行うので、電力を無駄に消費する問題がある。

3 提案手法

経由する MB にデータのコピーを保存する処理は、メモリ依存違反の可能性があるアドレスのロードには必要な処理である。しかし、メモリ依存違反がない場合は先行スレッドにメモリアクセス要求を出し、経由す

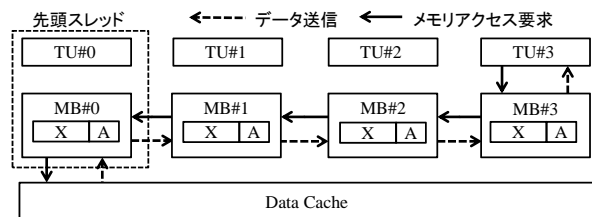


図 2: ロード命令実行時の処理

Reduction of Power Consumption for Two-Path Limited Speculation System by Efficient Memory Access

[†]Tomoyuki Kanno, Hiroyoshi Jutori, Kanemitsu Ootsu, Takashi Yokota, Takeshi Ohkawa
 Graduate School of Engineering, Utsunomiya University (†)

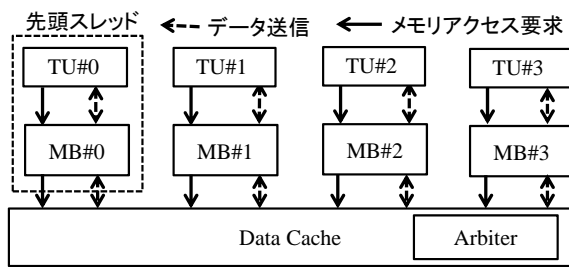


図 3: メモリ依存がないロード命令の処理

る MB にデータのコピーを保存する必要がない。このためメモリ依存違反がないロード命令を実行することにより、経由する MB においてデータのコピーに無駄な電力を消費してしまう。これに対処するために、メモリ依存違反がないロード命令は直接データキャッシュにアクセスすることにより、経由する MB とデータのコピーを削減するメモリアクセス最適化を提案する。

図 3 に TU4 台で本手法を適用した場合の概略図を示す。TU がメモリ依存違反のないロード命令を実行した場合、先行スレッドの MB にアクセスを送ることなく直接データキャッシュにアクセスする。これにより先行する MB にアクセスすることなくデータのコピーを保持することもない。本手法では同時に複数の MB からデータキャッシュにアクセスすることを許容するが、従来方式であればデータキャッシュにアクセスするのは先頭スレッドの MB のみである。したがって複数の MB から同時にアクセスがある場合は先頭スレッドに近い MB からのアクセスを優先するアービタの機能を追加した。

4 評価

本手法による電力効率の改善の効果を評価するために、Wattch を追加した PALS シミュレータ [2] を用いて並列実行部分の消費電力量、実行サイクル数、電力効率を示すエネルギー遅延積を求めた。エネルギー遅延積は消費電力量と実行サイクル数を乗算したものであり、値が小さいほど電力効率が良いものである。また、評価対象プログラムとして SPEC CINT2000 の 3 つのベンチマークプログラム 181.mcf, 175.vpr, 164.bzip を用いた。

図 4 の縦軸が TU, MB, データキャッシュ, PALS 全体を示し、横軸は適用前と比較した消費電力量比を示す。実験の結果、全てのベンチマークプログラムで MB の消費電力量が削減された結果となった。これはメモリ依存違反のないロード命令を実行し、データのコピーを保存しなかったために MB の消費電力量が削減できたからである。特に 181.mcf の MB の消費電力量が大幅に削減できたのは対象ループのパスがメモリ依存違反のないロード命令として扱うことができ、ストア命令が存在しなかったためである。したがって全体の消費電力量を平均して 50% 程度削減することができた。

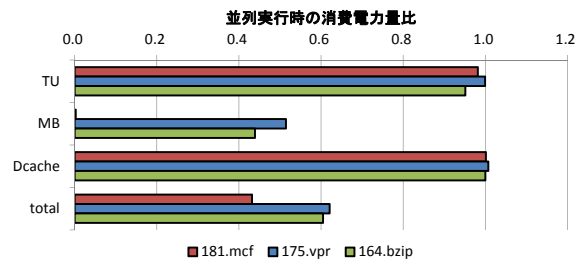


図 4: 並列実行部分の消費電力量比

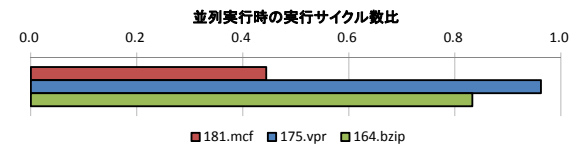


図 5: 並列実行部分の実行サイクル数比

図 5 は適用前と比較した実行サイクル数比を示す。適用前と比較して、全てのベンチマークプログラムにおいて実行サイクル数を削減することができた。本手法では先行スレッドの MB にアクセスしないので、後続スレッドのロード命令であるほど、先行スレッドの MB にアクセスする回数を削減することができる。したがって、ロードにかかる時間を短縮し、実行サイクル数を削減することができる。特に 181.mcf は同期待ちといった先行スレッドの実行を待つ時間が少なく、後続スレッドのロードが頻繁に実行されるために実行サイクル数が大きく削減できた。結果として、実行サイクル数は平均して 30% 程度削減することができた。さらに、消費電力量と実行サイクル数の積であるエネルギー遅延積を平均して 65% 削減することができ、電力効率の良いシステムとなった。

5 おわりに

本稿では、2パス限定投機システム PALS の電力効率を改善するために、消費電力量の MB のアクセス数を削減する手法を検討した。メモリ依存違反がないロード命令に対して、直接データキャッシュにアクセスする手法を提案した。この提案手法を 3 つのベンチマークプログラムに適用し、平均して 65% 削減することができた。

謝辞

本研究は、一部日本学術振興会科学研究費補助金 (基盤研究 (C)24500055, 同 (C)24500054, 同 (C)25330055, 若手研究 (B)25730026) の援助による。

参考文献

- [1] 十島弘泰ほか: “2パス限定投機方式を実現するマルチコアプロセッサ PALS の提案”, 信学技報, Vol.109, No.319(CPSY2009-46), pp.19-24, 2009.
- [2] 菅野智之ほか: “2パス限定投機システムにおけるプロセッサコアの省電力化”, 情報処理学会 第 75 回全国大会論文集, pp.1-173-1-174, 2013.