

令和 3 年 5 月 19 日現在

機関番号：14303

研究種目：基盤研究(C) (一般)

研究期間：2018～2020

課題番号：18K11214

研究課題名(和文) 投機メモリ機構の構成方式に関する研究

研究課題名(英文) Study on mechanisms to implement speculative memory

研究代表者

平田 博章 (Hirata, Hiroaki)

京都工芸繊維大学・情報工学・人間科学系・教授

研究者番号：90273549

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：マルチプロセッサ構成のマイクロプロセッサが一般的となった現在でも、並列化できないプログラムもまだ多く存在し、最新のマイクロプロセッサの利点を活かすことができない状況にある。そこで、並列化不可能(静的に解析すること自体が不可能)と諦められていたプログラムに対しても、すでに提案していた「投機メモリ」という概念の実現方式を探求することを通して並列化を可能とし、プログラムの実行時間の短縮を達成した。

研究成果の学術的意義や社会的意義

本研究は、単なる従来技術の改良ではなく、これまで並列化不可能として扱われていたプログラムを並列化可能にする点で革新的であり、学術的意義も大きい。本研究の成果は広く一般のプログラムに対して利用可能であるが、その中には、例えば、人工知能やビッグデータの分野で多用されるプログラムも含まれる。このため、本研究分野内だけに留まらず、他の重要な研究分野を含む広い範囲の研究にも貢献でき、有用性が高い。

研究成果の概要(英文)：Today, microprocessors organized as multiprocessors are commonplace, but there still remain many programs that are not parallelizable. Those programs can not get the benefit of modern microprocessors. In this study, we have explored implementation methods to realize "speculative memory", which is our original concept we had already presented to enlarge the opportunity of program parallelization. Consequently, we could achieve the speedups of program execution through our speculative memory system.

研究分野：情報工学

キーワード：計算機システム ハイパフォーマンスコンピューティング スレッドレベル並列処理 投機実行

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

(1) 並列プログラミングやプログラムの並列化の研究は古くから行われ、特に命令レベルの並列性の抽出については多大な成果が得られていた。しかし、スレッドレベルの並列性の抽出については、実際のプログラムにおいて、そもそも並列性を抽出できる機会が少ないという見解が一般化していた。

(2) 一方、コンピュータに搭載するマイクロプロセッサは、複数のプロセッサコアを備えてスレッドレベル並列処理を行うマルチプロセッサ(マルチコア)構成(マルチスレッドアーキテクチャ)とすることが一般的となってきた。つまり、マイクロプロセッサの高性能化の拠り所となる技術は、命令レベル並列処理からスレッドレベル並列処理へと移り変わってきていた。

(3) したがって、スレッドレベルの並列プログラムを記述しなければ実行時間の短縮が望めない状況へと変化してきていた。それにも関わらず、スレッドレベル並列性が容易に抽出できるプログラムのみを対象に並列実行することに力が注がれ、スレッドレベル並列性の抽出機会を増やす新たな方策の検討については停滞していると言わざるを得ない状況であった。つまり、プログラムを静的に解析して並列化可能と判定できないものについては、高速化を諦めざるを得ない状況であった。

(4) このような状況にあって、本研究代表者らは、データ依存性があるかどうかを実行前にわからない場合でも、並列実行可能なものと仮定して並列に実行を開始し、実行中にデータ依存性を検出した場合のみ、誤った計算部分の結果のみをキャンセルしてやり直す、というスレッドレベル並列投機実行方式の研究に着手し、その枠組みとして「投機メモリ」と呼ぶ概念を考案・提唱するに至っていた。

### 2. 研究の目的

(1) 本研究の目的の1つは、スレッドレベル並列性抽出の機会を拡大できる可能性を学術的な見地に立って明らかにすることである。これは、投機メモリという1つのアプローチの可能性と限界を明らかにすることにより、「スレッドレベル並列性の抽出機会をどれくらい増やすことが可能か」という普遍的な課題に対する挑戦と言える。

(2) また、投機メモリの実現方式を探求することを通して、スレッドレベルの並列投機実行技術を確立する。

### 3. 研究の方法

(1) 投機メモリを実現するための全体的なアーキテクチャ設計を行う。機能メモリの外部仕様はそれまでの研究でほぼ決まっていたが、内部の実装方式については、特に、データの依存解析方法に多くの選択肢が考えられ、投機実行の成功率と依存解析オーバーヘッドとの関係を十分に把握する必要がある。そこで、ソフトウェアのみで実装した投機メモリのプロトタイプを用いて予備評価を行いながら、投機メモリの実現方式について、そのデザインスペース(設計において考えられる選択肢の範囲)を明らかにし、その中で最適な設計を発見する(最適設計を行う)。そして、ハードウェア・ソフトウェアの詳細な機能分担を決定する。

(2) 上記(1)の結果に従って、ハードウェア機能の詳細設計を行う。実際にハードウェアを設計するのではなく、ここでの設計結果は下記(3)のシミュレータ内の機能モジュールとして実装する。

(3) 下記(5)の性能評価のために、マルチプロセッサ上でのマルチスレッドプログラムの実行時間をマシンスイクル単位で計測することが可能なシミュレータを開発する。

(4) 上記(1)の結果を基に、上記(2)のハードウェア機構を利用するソフトウェアライブラリを開発する。

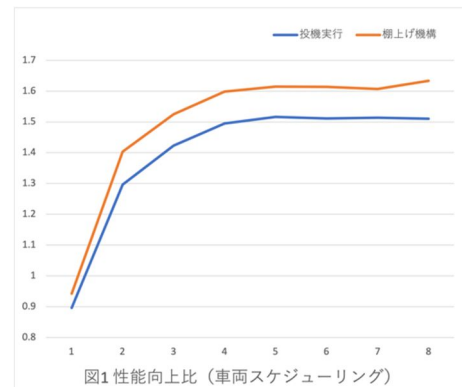
(5) 上記(3)で開発したシミュレータ上で、上記(4)のライブラリを用いたアプリケーションプログラムを実行し、投機メモリシステムの性能評価を行う。これにより、投機メモリの有効性を明らかにするとともに、さらなる並列投機実行の可能性の有無や指針を探る。

#### 4. 研究成果

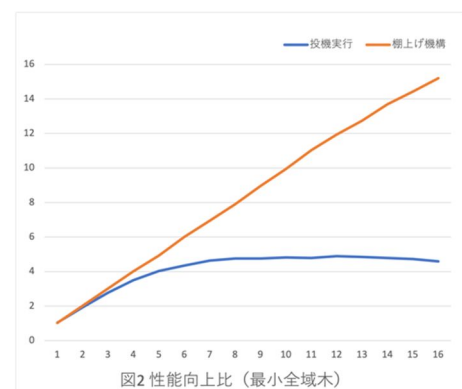
(1) 従来のようにハードウェアのみで投機実行を制御するのではなく、また、ソフトウェアのみでスレッドレベル並列投機実行機能を実装するのでもなく、両者の利点を活かした「投機メモリ」の実装方式を検討した。使いやすさ・適用範囲の広さの点からソフトウェアでの明示的な投機実行制御を可能とし、また、速度性能の点からその投機実行制御機能をサポートするハードウェア機構を設計した。これにより、スレッドレベル並列投機実行システムのハードウェア・ソフトウェアトレードオフを明らかにすることができた。

(2) 従来は並列化不可能として高速化をあきらめざるを得なかったプログラムに対しても並列化を可能とし、並列化によるプログラム実行の高速化の対象範囲を拡大した。以下のように、高速化の度合いは対象プログラムにもよるが、概ね優れた性能向上が達成できることを確認した。

図1は公共の交通機関において車両のスケジューリングを行うプログラムを実行した結果である。(単一スレッド)プロセッサの性能評価に用いられる標準的なベンチマークプログラムの中の1つであり、従来はまったく並列化が不可能であったが、本研究で並列化を可能とした。縦軸が単一スレッド実行に対するマルチスレッド実行の性能向上比であり、横軸がスレッド数である。青線が並列投機実行による性能向上比を表し、並列化対象とした部分以外の実行時間も含まれているために性能向上比は少なく見積もられているが、全体として約1.5倍の高速化を達成している。なお、本研究の過程において並列投機実行のための最適化技術に関する新たな着想を得た。投機実行の失敗を回避する機構(棚上げ機構)で、これを併用した場合の性能向上比を赤線で示している。棚上げ機構を用いることで、最大1.6倍以上の高速化を達成した。



また、図2は最小全域木を求めるプログラムを実行した結果である。全域木の概念はコンピュータネットワークやグラフ理論において重要であり、また、ビッグデータにおける階層化クラスタリングにも関連する。部分的には並列化が可能ではあるが、粒度が細かく、従来は性能改善は望めなかった。それを並列投機実行によって並列化し、5倍程度の性能向上を実現した。さらに、並列投機実行に加えて棚上げ機構を用いることで、ほぼスレッド数に比例する性能向上を達成することができた。



(3) 本研究では、スレッドレベル並列投機実行に関する基礎的な実現方式を探求することを研究の主眼としていたが、研究過程において、上記の「棚上げ機構」を含む新たな最適化・高速化技術の着想を得た。ただし、現時点ではそれらの適用範囲は限定的で、十分に検討し尽くしたとは言えない段階にある。これらについては、今後の研究課題として取り組む予定である。

(4) スレッドレベル並列投機実行方式の最適な実装方式を明らかにした(上記(1)(2)に記載)とともに、上記(2)の 以外にも、 $k$ - $d$ 木などのデータ構造を扱うアルゴリズムなど、ビッグデータ解析やAIで利用されながらも並列化が困難であったアルゴリズムに対して、投機メモリを用いて並列化を行う手法を提案した。十分な高速化が達成できることを示し、重要な分野のアプリケーションに対するスレッドレベル並列投機実行手法の適用可能性・有効性を確認した。

(5) 研究計画時には本研究の研究目的に含めていなかったが、従来技術で並列化可能なプログラムに対しても、本研究で開発した投機メモリを利用することで並列プログラムのデバッグが格段に容易になることを見出した。これにより、並列プログラム開発手法としても投機メモリが有効であることを確認した。

## 5. 主な発表論文等

〔雑誌論文〕 計10件（うち査読付論文 10件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 D.Matsunaga, A.Nunome and H.Hirata	4. 巻 1
2. 論文標題 Shelving a Code Block for Thread-Level Speculation	5. 発行年 2019年
3. 雑誌名 Proceedings of the 20th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019)	6. 最初と最後の頁 427-434
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SNPD.2019.8935751	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 H.Kitano, A.Nunome and H.Hirata	4. 巻 1
2. 論文標題 Performance Evaluation of Parallel Heapsort Programs	5. 発行年 2019年
3. 雑誌名 Proceedings of the 20th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019)	6. 最初と最後の頁 435-442
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SNPD.2019.8935767	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 A.Nunome and H.Hirata	4. 巻 1
2. 論文標題 An Improvement of Migration Efficiency in a Distributed Storage System with Dynamic Tiering	5. 発行年 2019年
3. 雑誌名 Proceedings of the 20th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019)	6. 最初と最後の頁 455-460
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SNPD.2019.8935656	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 A.Nunome and H.Hirata	4. 巻 8
2. 論文標題 Performance Evaluation of Data Migration Policies for a Distributed Storage System with Dynamic Tiering	5. 発行年 2019年
3. 雑誌名 International Journal of Networked and Distributed Computing (IJNDC)	6. 最初と最後の頁 1-8
掲載論文のDOI (デジタルオブジェクト識別子) 10.2991/ijn/dc.k.191115.001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 H.Yamasaki, A.Nunome and H.Hirata	4. 巻 1
2. 論文標題 Parallelizing the Construction of a k-Dimensional Tree	5. 発行年 2018年
3. 雑誌名 Proceedings of the 3rd International Conference on Big Data, Cloud Computing, and Data Science & Engineering (BCD 2018)	6. 最初と最後の頁 33-40
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/BCD2018.2018.00012	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 A.Nunome and H.Hirata	4. 巻 1
2. 論文標題 A Data Migration Scheme Considering Node Reliability for an Autonomous Distributed Storage System	5. 発行年 2018年
3. 雑誌名 Proceedings of the 5th International Conference on Computational Science/Intelligence and Applied Informatics (CSII 2018)	6. 最初と最後の頁 160-165
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CSII.2018.00035	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hirata Hiroaki, Nunome Atsushi	4. 巻 8
2. 論文標題 A Modified Parallel Heapsort Algorithm	5. 発行年 2020年
3. 雑誌名 International Journal of Software Innovation	6. 最初と最後の頁 1~18
掲載論文のDOI (デジタルオブジェクト識別子) 10.4018/IJSI.2020070101	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hirata Hiroaki, Nunome Atsushi	4. 巻 8
2. 論文標題 Decoupling Computation and Result Write-Back for Thread-Level Parallelization	5. 発行年 2020年
3. 雑誌名 International Journal of Software Innovation	6. 最初と最後の頁 19~34
掲載論文のDOI (デジタルオブジェクト識別子) 10.4018/IJSI.2020070102	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hirata Hiroaki, Nunome Atsushi	4. 巻 1
2. 論文標題 Reducing the Repairing Penalty on Misspeculation in Thread-Level Speculation	5. 発行年 2021年
3. 雑誌名 Proceedings of the 8th International Virtual Conference on Applied Computing & Information Technology	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Nunome Atsushi, Hirata Hiroaki	4. 巻 1
2. 論文標題 An Adaptive Tiering Scheme for an Autonomous Distributed Storage System	5. 発行年 2021年
3. 雑誌名 Proceedings of the 8th International Virtual Conference on Applied Computing & Information Technology	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

[学会発表] 計7件 (うち招待講演 0件 / うち国際学会 7件)

1. 発表者名 D.Matsunaga, A.Nunome and H.Hirata
2. 発表標題 Shelving a Code Block for Thread-Level Speculation
3. 学会等名 The 20th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 H.Kitano, A.Nunome and H.Hirata
2. 発表標題 Performance Evaluation of Parallel Heapsort Programs
3. 学会等名 The 20th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019) (国際学会)
4. 発表年 2019年

1 . 発表者名 A.Nunome and H.Hirata
2 . 発表標題 An Improvement of Migration Efficiency in a Distributed Storage System with Dynamic Tiering
3 . 学会等名 The 20th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2019) (国際学会)
4 . 発表年 2019年

1 . 発表者名 H.Yamasaki, A.Nunome and H.Hirata
2 . 発表標題 Parallelizing the Construction of a k-Dimensional Tree
3 . 学会等名 The 3rd International Conference on Big Data, Cloud Computing, and Data Science & Engineering (BCD 2018) (国際学会)
4 . 発表年 2018年

1 . 発表者名 A.Nunome and H.Hirata
2 . 発表標題 A Data Migration Scheme Considering Node Reliability for an Autonomous Distributed Storage System
3 . 学会等名 The 5th International Conference on Computational Science/Intelligence and Applied Informatics (CSII 2018) (国際学会)
4 . 発表年 2018年

1 . 発表者名 H.Hirata and A.Nunome
2 . 発表標題 Reducing the Repairing Penalty on Misspeculation in Thread-Level Speculation
3 . 学会等名 The 8th International Virtual Conference on Applied Computing & Information Technology (ACIT2021) (国際学会)
4 . 発表年 2021年

1. 発表者名 A.Nunome and H.Hirata
2. 発表標題 An Adaptive Tiering Scheme for an Autonomous Distributed Storage System
3. 学会等名 The 8th International Virtual Conference on Applied Computing & Information Technology (ACIT2021) (国際学会)
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	布目 淳 (Nunome Atsushi) (60335320)	京都工芸繊維大学・情報工学・人間科学系・准教授  (14303)	
研究分担者	柴山 潔 (Shibayama Kiyoshi) (70127091)	京都情報大学院大学・その他の研究科・教授  (34323)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------