

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 30 日現在

機関番号：14301

研究種目：若手研究（B）

研究期間：2010～2012

課題番号：22700030

研究課題名（和文） 後戻りに基づく動的負荷分散による並列化技法の実用化

研究課題名（英文） Practical Application of Parallelization with Dynamic Load Balancing Based on Backtracking

研究代表者

平石 拓 (HIRAISHI TASUKU)

京都大学・学術情報メディアセンター・助教

研究者番号：60528222

研究成果の概要（和文）：提案する動的負荷分散計算フレームワーク Tascell を用いた実アプリケーションの開発，およびフレームワークそのものの改良を行った．アプリケーション開発では，多体問題シミュレーションやスパニング木の構築，グラフマイニング等の実装を行い，既存のフレームワークより高速な並列化が可能であることを確認できた．Tascell 自体の改良としては，ワーカローカルストレージ，広域分散環境への対応や確率的ガードという機能の提案・実装を行い，プログラムの書き易さや性能を向上させた．

研究成果の概要（英文）：We developed practical applications using our dynamic load balancing framework Tascell. We also developed some additional features to Tascell itself. As the application development, we parallelized a many-body simulation program, a spanning tree construction algorithm, and a graph mining algorithm using Tascell. We ensured that we can parallelize these programs more efficiently than existing frameworks. As the improvements to Tascell, we added support for worker local storages and widely distributed environments. We also proposed and developed a novel feature called “probabilistic guards.” These features contribute to productivity and performance of Tascell.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	1,800,000	540,000	2,340,000
2011年度	700,000	210,000	910,000
2012年度	600,000	180,000	780,000
総計	3,100,000	930,000	4,030,000

研究分野：総合領域

科研費の分科・細目：情報学・ソフトウェア

キーワード：並列処理・分散処理

1. 研究開始当初の背景

近年の計算システムは，単一のプロセッサの性能向上ではなく，マルチコア／プロセッサ，クラスタ，グリッドによる大規模並列化による高性能化が進んでおり，この傾向は今

後も続くものと予想される．そのため，計算機による応用計算を行っている多くの研究者が従来型の逐次プログラミングではなく，分散環境にも対応した並列プログラミングへの移行を余儀なくされている．しかし，そ

のような複雑化する並列計算環境に対応した、効率良いプログラムを書くのは容易ではない。特に、充足可能性問題 (SAT) やグラフアルゴリズムに代表される、事前に平等な仕事の配分を見積ることができない「不規則アプリケーション」の並列化は困難であり、可能であったとしても既存の逐次版から根本的なプログラムの書き換えが必要となることも多い。

共有メモリ環境においては、たとえば言語が (例えば Intel TBB などと比較して) そのようなアプリケーションの並列化を比較的容易に実現できる言語の一つである。すなわち、多数の論理スレッドを生成して最古優先のワークスティールを採用することで全ワーカを有効活用し、良好な負荷分散を実現する。これに対し我々は、論理スレッドフリーなプログラミング/実行フレームワーク Tascell を提案している。Tascell は生産性を損なうことなく、クラスタを含む多様な並列計算環境に対応する。Tascell ワーカは本物のタスクを生成 (spawn) するが、それは他の遊休 (idle) ワーカから要求されたときであり、一時的な後戻り (バックトラック) による最古のタスク生成可能状態 (task-spawnable state) の復元に基づく。この手法は、論理スレッド生成・管理のコストを削減する。また並行に計算を進める可能性がある論理スレッドそれぞれに対して作業空間を準備する必要がないため、作業空間の再利用を促進するとともに参照局所性を改善する。さらに、本フレームワークが提供する Tascell 言語では、作業空間の遅延コピーを伴う、すっきりとした効率良いバックトラック探索アルゴリズムが実現できる。また Tascell は、Cilk は未対応である分散メモリ環境にも対応している。

2. 研究の目的

本研究では、前述の後戻りに基づく負荷分散フレームワーク Tascell をより現実的な応用問題に適用して、並列計算環境における高いパフォーマンスを得ることを目的とする。さらに、その開発を通してフィードバックを得ることで Tascell に必要な機能追加等を行い、実用に耐え得る言語として完成度を高めることも目指す。

3. 研究の方法

平成 22 年度は、具体的な既存の逐次実装のグラフマイニングの例題について、主に共有メモリ環境での並列化を、Tascell を用いて施すことで性能向上を実現する。開発中、必要に応じて Tascell システムの機能拡張も並行して行う。平成 23 年度以降は、主に分散メモリ環境 (クラスタ) にも対応した並列化を施し、同様に必要な機能拡張も行う。

研究の進捗状況によっては、対象とする例題を SAT ソルバ等他の既存逐次プログラムに広げたり変更したりすることも有り得る。また、グリッド等の広域分散並列環境における評価も行う。

4. 研究成果

前述の研究方針に従い、Tascell によるアプリケーションの開発およびその過程で必要となった Tascell への機能拡張を実施した。これにより、これまで並列化が難しかった種類のアプリケーションを小さいプログラミングコストで効率良く並列化できるようになった。

以下、研究成果をアプリケーションと機能拡張に分けて説明する。

(1) アプリケーションの開発

① Spanning tree 構築アルゴリズム

Spanning tree とは、任意のグラフから枝を取り除き閉路を持たないようにしたグラフ (ツリー) である。Spanning tree を構築するためのアルゴリズムは、従来効率的に並列化することが困難であったが、深さ制限付きの深さ優先探索を行いつつ溢れた呼び出しをワーカごとの記憶領域 (ワーカローカルストレージ) に蓄え、それを次回の並列ステージで呼び出すというプログラミングの工夫を用いて、従来より高い並列性能を得ることに成功した。なお、ここでワーカローカルストレージの機能は、後で述べる通り、アプリケーション開発時判明した Tascell に不足していた機能として新たに追加したものである。

② Barnes-Hut アルゴリズムによる多体問題シミュレーション

Barnes-Hut アルゴリズムは、各質点を leaf とする木を構築してそれを一回走査することで全質点間の力を近似的に計算するものである。生成される木は一般にアンバランス

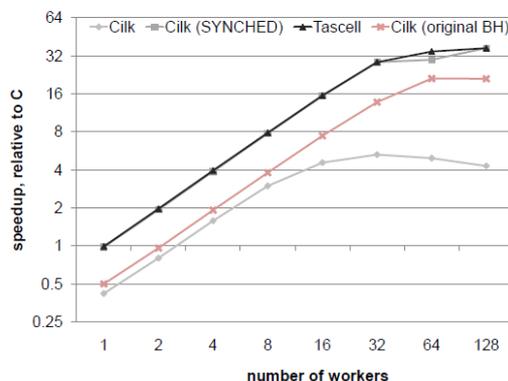


図 1 Barnes-Hut アルゴリズムの並列化実装の性能評価 (共有メモリ環境)。

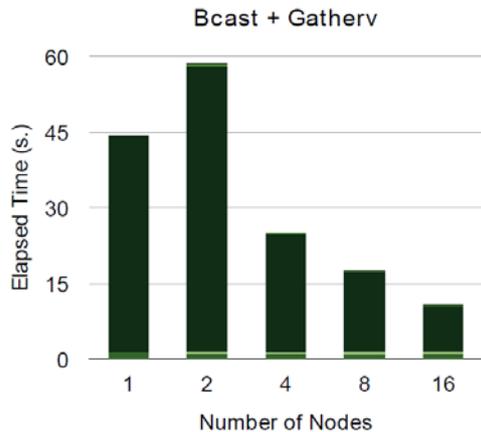


図 2 Barnes-Hut アルゴリズムの並列化実装の性能評価 (分散メモリ環境)。

であるため、Tascell による並列化が効果的に適用できる。木の構築と木の走査それぞれについての並列化が考えられるが、最初は計算時間の多くを占める木の走査の並列化を行った。実装および評価の結果、128 ハードウェアスレッドを備える共有メモリ並列計算サーバにおいて、素朴に並列化した Cilk プログラムより約 7 倍高速に計算させることに成功した (図 3)。

また、分散環境における効率良い並列化を実現するために、MPI ライブラリの集団通信機能を利用した実装を行った。京都大学学術情報メディアセンターのスーパーコンピュータ 16 ノードを用いた評価を行い、高い並列性能を得られることを確認した (図 2)。

並列数が増加すると、Barnes-Hut アルゴリズムのうち、木の構築にかかる時間が無視できなくなるため、共有メモリ環境において木の構築の並列化を行った。複数のワーカが同時に木の同一箇所の書き換えを行うことを防ぐためにロックを利用すると、頻繁に起こるロック獲得のオーバーヘッドが大きくなるため十分な性能を得ることができない。そこで、書き込むアドレスが重複しないように担当する粒子の集合を分割する空間スティーリングと呼ぶタスク分割手法を提案・実装した。また、タスクの分割数の増大に伴い増加する余剰計算の量を抑えるため、後述する確率的ガードの機能を Tascell に追加し、木の構築の並列実装に適用した。性能評価を行い、図 3 に示すように高い並列性能を得られることを確認した。

③ グラフマイニング問題

本研究で対象としたグラフマイニング問題は、各頂点がアイテムの集合を持つようなグラフから、構成する頂点が共通して持つアイテムの数が指定された閾値以上であるような連結部分グラフを全て列挙するもので

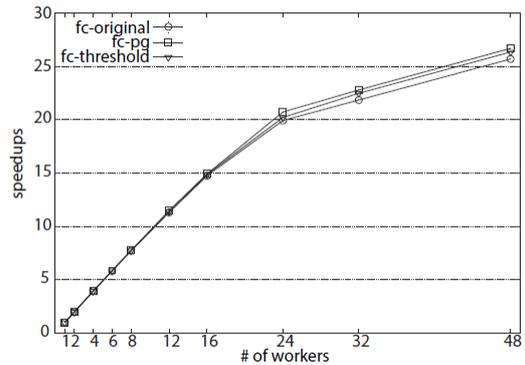


図 3 Barnes-Hut アルゴリズムにおける木の構築の並列化実装の性能評価。

ある。このマイニング問題は、タンパク質からある試薬に反応する部分や、SNS において共通の興味を持つ「友達」関係を抽出するような応用に適用できる。

本研究期間では、枝刈りを用いる逐次版のアルゴリズムが単純には並列化できないことが判明したため、枝刈りを行いつつ並列に正しく探索を実施できるアルゴリズムおよびその実装方針を考案し、ノード内並列・ノード間並列の双方で Tascell を用いて効率良く実装可能であることを確認した。

(2) Tascell へ機能拡張

① ワーカーローカルストレージ

従来の Tascell では、計算に使用するデータは、グローバル領域を除くと、タスクオブジェクト (タスクをスティーリングした際に渡されるオブジェクト) に全て蓄えられており、そのデータはタスク完了時に破棄されていた。しかし、前述の spanning tree 構築アルゴリズムの並列実装のように、ワーカごとに永続的にデータを保持しておく領域を利用したいこともある。そこで、Tascell 言語を拡張し、そのようなデータ構造を C の構造体定義と同様の記述により定義し、利用できるようにした。

② 広域分散環境への対応

Tascell は複数の拠点に設置されたクラスターを WAN で接続した広域分散環境にも適用可能なように設計、実装しているが、本研究では、そのような環境における Tascell の有効性を検証するために、複数の大学等の計算クラスターを WAN で接続した計算環境 InTrigger のうち最大 5 クラスター 288 コアを用いた環境における評価を行った。その結果、メモリバンド幅や通信遅延、ワークスティーリングの偏り等の影響によるスケラビリティの制約は観測されたものの、本フレームワークが広域分散環境を含む高並列環境においてもおおむね効率良い計算が行えることが確認

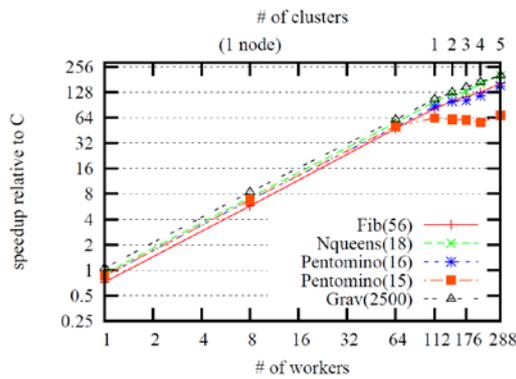


図 4 広域分散環境における Tascell の並列性能の評価。

できた (図 4)。また、さらなる高速化に向けた詳細な解析のために Tascell のシミュレータを実装し、その上で、広域分散環境における計算をシミュレートして分析を行った。その結果、局所性を考慮したスティーリングと完全にランダムにワーカを選択するスティーリングを混合させることで性能を改善できることがわかった。

③ 確率的ガード

計算の分割数の増大に伴うオーバーヘッドの増加を抑えるため、ユーザが与えた確率式に基づいてワークスティーリングを阻止する「確率的ガード」の言語機能を開発した。本機能は、通常用いられる「閾値」によるワークスティーリングの抑制（自らの残りの仕事がある閾値より小さいと判断すればワークスティーリングを受け付けない）とは異なり、潜在的な並列性を失わずに確率的に仕事の分割単位を大きくすることができる。実際、Barnes-Hut アルゴリズムやいくつかのマイクロベンチマークにおける評価によって、特に分割のコストが大きいアプリケーションにおいて「閾値」による手法より良好な性能が得られることが確認できた。なお、この機能は広域分散環境等、スティーリングのコストが大きくなるような実行環境においても有効にはたらくことが期待できる。

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 6 件)

- ① 田附 正充, 八杉 昌宏, 平石 拓, 馬谷 誠二, L-Closure の呼び出しコストの削減, 情報処理学会 プログラミング, 査読有, 2013 (採録決定)
- ② Masahiro YASUGI, Tasuku Hiraishi, Parallel Graph Traversals using Work-Stealing Frameworks for

Many-core Platforms, Journal of Information Processing, 査読有, Vol. 20, No. 1, 2012, pp. 128~139

DOI: 10.2197/ipsjjip.vol.20.128

- ③ 平石 拓, 八杉 昌宏, バックトラックに基づく動的負荷分散フレームワーク Tascell, 京都大学学術情報メディアセンター 全国共同利用版広報, 査読無, Vol. 10, No. 1, 2011, pp. 32~42

- ④ 平石 拓, 八杉 昌宏, 湯浅 太一, SC 言語処理系における変形規則の再利用機構, コンピュータソフトウェア, 査読有, Vol. 28, No. 1, 2011, pp. 258~271

- ⑤ 八杉 昌宏, 小島 啓史, 小宮 常康, 平石 拓, 馬谷 誠二, 湯浅 太一, L-Closure を用いた真に末尾再帰的な Scheme インタプリタ, 情報処理学会論文誌 プログラミング, 査読有, Vol. 3, No. 5, 2010, pp. 1~17

https://ipsj.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=71435&item_no=1&page_id=13&block_id=8

[学会発表] (計 21 件)

- ① 八杉 昌宏, 平石 拓, 安全な計算状態操作機構の実用化, 自動チューニング技術の現状と応用に関するシンポジウム, 2012 年 12 月 25 日, 東京大学 山上会館
- ② 松井 健, 平石 拓, 八杉 昌宏, 馬谷 誠二, ワークスティーリングフレームワークにおける集団通信機能, 並列/分散/協調処理に関するサマー・ワークショップ (SWoPP2012), 2012 年 8 月 1 日, 鳥取市 とりぎん文化会館
- ③ 松井 健, 平石 拓, 八杉 昌宏, 馬谷 誠二, 高速版 Barnes-Hut 多体シミュレーションの並列実装, 先進的計算基盤システムシンポジウム (SACSYS2012), 2012 年 5 月 18 日, 神戸国際会議場
- ④ 平石 拓, 八杉 昌宏, 安全な計算状態操作機構の実現と応用, 自動チューニング研究会 第 3 回自動チューニング技術の現状と応用に関するシンポジウム, 2011 年 12 月 5 日, 東京大学
- ⑤ 松井 健, 平石 拓, 八杉 昌宏, 馬谷 誠二, 湯浅 太一, ワークスティーリングフレームワークにおけるブロードキャスト機能, 並列/分散/協調処理に関するサマー・ワークショップ (SWoPP2011), 2011 年 7 月 29 日, かがしま県民交流センター
- ⑥ 平石 拓, 八杉 昌宏, 馬谷 誠二, 動的負荷分散フレームワーク Tascell の広域分散およびメニーコア環境における評

価, 先端的計算基盤システムシンポジウム SACSIS 2011, 2011年5月25日, 東京都秋葉原コンベンションホール

- ⑦ Masahiro Yasugi, Tasuku Hiraishi, Seiji Umatani, Taiichi Yuasa, Dynamic Graph Traversals for Concurrent Rewriting using Work-Stealing Frameworks for Multicore Platforms, 16th International Conference on Parallel and Distributed Systems (ICPADS 2010), 2010年12月10日, 中国上海市 衡山賓館
- ⑧ 平石 拓, バックトラックに基づく負荷分散フレームワーク Tascell, 京都大学 学術情報メディアセンターセミナー 「Models and frameworks for programming systems and languages」招待講演, 2010年11月17日, 京都大学
- ⑨ Masahiro Yasugi, Tsuneyasu Komiya, Tasuku Hiraishi, Seiji Umatani, Managing Continuations for Proper Tail Recursion, 2010 International Lisp Conference (ILC 2010), 2010年10月21日, 米国ネバダ州リノ市 John Ascuaga's Nugget
- ⑩ 平石 拓, 河野卓矢, 八杉昌宏, 馬谷誠二, 湯淺太一, バックトラックに基づく負荷分散の高並列環境における評価, 並列/分散/協調処理に関するサマー・ワークショップ(SWoPP2010): 第126回ハイパフォーマンスコンピューティング研究発表会, 2010年8月4日, 石川県金沢市文化ホール
- ⑪ Tasuku Hiraishi, Masahiro Yasugi, Takuya Kouno, Seiji Umatani, Taiichi Yuasa, Tascell: a Backtracking-based Load Balancing Framework, 24th International Conference on Supercomputing (ICS10), poster presentation, 2010年6月2日, 茨城県つくば市 エポカルつくば

[その他]

ホームページ等

<http://super.para.media.kyoto-u.ac.jp/tascell/>

6. 研究組織

(1) 研究代表者

平石 拓 (HIRAISHI TASUKU)

京都大学・学術情報メディアセンター・助教

研究者番号: 60528222