



Title	分散メモリ型並列計算機におけるタスクスケジューリングに関する研究
Author(s)	藤本, 典幸
Citation	
Issue Date	
Text Version	ETD
URL	<a href="https://doi.org/10.11501/3169503">https://doi.org/10.11501/3169503</a>
DOI	10.11501/3169503
rights	

***Osaka University Knowledge Archive : OUKA***

<https://ir.library.osaka-u.ac.jp/repo/ouka/all/>

氏名	藤本典幸
博士の専攻分野の名称	博士(工学)
学位記番号	第 15540 号
学位授与年月日	平成12年3月24日
学位授与の要件	学位規則第4条第1項該当 基礎工学研究科物理系専攻
学位論文名	分散メモリ型並列計算機におけるタスクスケジューリングに関する研究
論文審査委員	(主査) 教授 萩原 兼一  (副査) 教授 谷口 健一 教授 柏原 敏伸

### 論文内容の要旨

並列計算を表す重み付きの DAG (Directed Acyclic Graph)  $G$  と利用可能なプロセッサ数  $p$  が与えられたときに、 $p$  台のプロセッサを用いて  $G$  が表す並列計算を計算するためには、すべてのタスクについて、そのタスクをいつ、どのプロセッサで実行すればよいかを表すスケジュールを求める問題をタスクスケジューリング問題 (以降 TSP) という。ここで  $G$  の節点はタスクを表し、有向辺はタスク間の実行順序の依存関係を表す。節点の重みと有向辺の重みはそれぞれタスクの処理時間とタスクの計算結果がそれを必要とするタスクで利用可能になるまでの通信遅延時間を表す。 $G$  の有向辺の重みがすべて定数  $\tau$  である場合の TSP を  $TSP_\tau$  と呼ぶ。スケジュールですべてのタスクの計算が完了するまでの時間をそのスケジュールのメイクスパンという。 $TSP_\tau$  でメイクスパン最小のスケジュールを求める問題は NP 困難である。本論文では、 $TSP_\tau$  に関する 2 つの問題を扱っている。

本論文では、まず、単にメイクスパンが小さいだけのスケジュールを並列プログラム化してもメイクスパン値から予測される性能は出ないことを評価実験により指摘し、その原因の 1 つが通信のソフトウェアオーバーヘッドであることを明らかにしている。さらに本論文では通信のソフトウェアオーバーヘッドが少ない並列プログラム化が可能なスケジュールのクラスの一つとしてバルク同期スケジュールを提案し、任意の  $G$ 、 $p$ 、 $\tau$  に対して高速な並列プログラムとなるバルク同期スケジュールを生成するアルゴリズム BCSH を提案している。さらに本論文では、分散メモリ型並列計算機である富士通 AP1000 上で、BCSH が生成するスケジュール通りに動作する並列プログラムの性能評価を行い、BCSH を用いれば、32 台のプロセッサを用いたとき、問題サイズ 128 の LU 分解で速度向上率 8 倍、問題サイズ 8192 の高速フーリエ変換で速度向上率 12 倍、問題サイズ 512 の連立一次方程式の反復解法のヤコビ法で速度向上率 18 倍の性能をもつ並列プログラムを生成できることを示している。

次に本論文では、プロセッサ数が任意に設定でき、各タスクの処理時間がすべて等しく、DAG が完全 2 分木である場合の  $TSP_\tau$  に対するアルゴリズム ALG 1、ALG 2 を提案している。アルゴリズムが生成するスケジュールのメイクスパンの最適スケジュールのメイクスパンに対する比率をそのアルゴリズムの近似精度という。本論文では 2 から 10000 の範囲の  $\tau$  と、ノード数 100 万以内の完全 2 分木に対して、ALG 1 と ALG 2 の近似精度が平均 1.1 から 1.3 以内であること、プロセッサ割当てが非常に単純であるにもかかわらず、ALG 2 の近似精度が平均 1.1 から 1.4 以内であることを実験により示している。さらに本論文では、利用できるプロセッサ数が限られる場合のアルゴリズム ALG 3 を提案し、ALG 3 の近似精度が高々  $10/3$  であることを解析的に示している。さらに本論文では、任意の DAG と  $\tau$

に対してメイクスパンの既知の下界の改良方法を提案し、完全二分木、バタフライ、ダイヤモンドについて既知の下界が改良できることを示している。

## 論文審査の結果の要旨

本論文は、並列計算のための主流アーキテクチャである分散メモリ型並列計算機を対象として、並列計算の研究分野における最も重要な基本問題のひとつであるタスクスケジューリング問題に関する研究成果をまとめている。

タスクスケジューリング問題の既存研究ではスケジュールの評価をメイクスパンによってのみ行っている。これに対して本研究ではスケジュールの評価を対応する並列プログラムの実行時間によって行っている。本研究は、まず、高い並列性をもつという利点がある細粒度のタスクグラフをタスクスケジューリングの対象とする場合、単にメイクスパンが小さいだけのスケジュールを並列プログラム化してもメイクスパン値から予測される性能は出ないことを並列計算機上でのプログラム評価実験により指摘している。そして、分散メモリ型並列計算機のハードウェアと通信ライブラリの実装の分析をもとに、スケジュールのメイクスパンと並列プログラムの実行時間の乖離の主要な原因が通信のソフトウェアオーバーヘッドの無視であることを明らかにしている。また、並列プログラムの実行時間が遅い原因が細粒度実行と通信の一括化が並列プログラムに適用できないためであることを明らかにしている。すなわち、既存のタスクスケジューリングアルゴリズムではスケジュールの並列計算機上での実行時間を正しく予測できず、また高速な並列プログラムに対応するスケジュールを生成できない。これに対して本研究は、並列プログラムの実行時間をメイクスパンで高精度に近似でき、かつ、通信の一括化が適用できる粗粒度プログラムに対応するスケジュールのクラスの1つとしてバルク同期スケジュールを提案している。さらに高速な並列プログラムとなるバルク同期スケジュールを生成するアルゴリズムを提案している。並列計算機上でプログラム評価実験を行い、提案アルゴリズムでは予測が正しく、かつ、高速な並列プログラムとなるスケジュールを生成できることを示している。

次に、並列プログラム中でサブルーチンとしてよく用いられるリダクション計算を表すタスクグラフを対象としたタスクスケジューリング問題に対するアルゴリズムを提案している。考案したアルゴリズムは既知のものに比べ効率がよい。さらに、メイクスパンの既知の下界の改良方法を提案し、タスクグラフが完全二分木状、バタフライ状、ダイヤモンド状の場合について、下界を改良できることを示している。

以上の研究成果は、分散メモリ型並列計算機を対象としたタスクスケジューリング問題の基礎研究の発展に貢献しており、博士（工学）の学位論文として価値あるものと認める。