

## 並列気候モデルNJR-SAGCMのプロセス競合を考慮した動的負荷分散手法

浅野 俊幸 † 堀口 進 †

本論文では、並列気候モデル向きのプロセス競合を考慮した動的負荷分散手法を提案する。近年、従来から使用されていた逐次処理型の気候モデルからの並列化への移行、特に大規模計算を念頭に置いた並列化への移行が試みられている。並列化によって新たに数値計算手法の並列化効率やマルチユーザ環境でのプロセス競合による実行時間の遅延等が問題視されている。一般的に数値計算手法はスペクトル法が広く用いられている。スペクトル法では格子空間と波数空間それぞれで計算が行われるが、各空間の使用ノード数が異なる事による並列効率の低下が指摘されている。本手法では時計算量が小さいノードへ、他のプロセスと競合状態にある負荷集中状態のノードの計算を移行させることにより効率よく負荷分散を実現した。

### Load-Balancing Algorithms for a Parallel Climate Model NJR-SAGCM

TOSHIYUKI ASANO † and SUSUMU HORIGUCHI †

We propose a Dynamic load balancing Scheme in consideration of the process rivalry which is effective in mounting of the parallel climate model. Recently, it shifts from sequential operation model which had been used from before to the weather model model, climate of parallel operation model. The problem of the delay of the effective time by the problem of the parallel efficiency of the numerical value calculation technique and the process rivalry in the multi-user environment is pointed out with an arrangement in a row-ization model. Generally spectral transforms is being used for the numerical value calculation technique widely. As for the spectral transforms, a calculation is done in the lattice space and the wave number space. To be a problem here is that parallel efficiency decreases because the number of nodes to use in each space is different. It aims at dispersing a load efficiently by shifting the calculation of a node to compete with each other with other processes to the node which isn't being used for the calculation with this technique.

#### 1. はじめに

大規模並列計算機上での性能を高めるために通信負荷の軽減や均等な計算負荷分散について研究の重要性が増している。気候シミュレーションを大規模並列計算機にて実行すると、領域分割による最大並列化可能プロセッサ数の約半分まで速度向上比が飽和することが報告されている<sup>1)</sup>。また、各PE毎の計算負荷は気象状態(力学過程)の基礎方程式の計算に用いられているスペクトル法やモデルの分解能より細かい現象(物理過程)の計算の1つである放射の計算、さらにPE間データ通信に顕著な不均衡が確認されている。また、並列計算機でシミュレーションを行う場合、一般的には専有することは難しくマルチユーザ環境での使用となる。並列計算機をマルチユーザ環境で並列プロセス

を実行すると、使用ノードのうち1つでも他のプロセスと競合すると、並列プロセス全体の終了時間が遅延するという影響を受ける。このような競合状態の並列計算機上で実行される並列プログラムでは、計算領域を均等に計算させるようなモデルであっても並列効率は減少する。我々は、こうした競合による負荷の変化を指標として利用することで、競合状態にある並列計算機上でも効率的に負荷分散制御を行うことができると考えた。従来の並列気候モデルでは、負荷(競合)による遅延への対処を考慮されていなかった。そこで本稿では、並列大気大循環モデル(NJR-SAGCM:Spectral Atmospheric General Circulation Model)を並列計算機(IBM RS/6000 SP)に実装し、ノードの負荷量を考慮したスペクトル法の負荷分散手法を提案した。

#### 2. 気候モデル

気候モデルとは気候現象の予測・研究のために用いられる数値シミュレーションモデルである。気候モデ

† 北陸先端科学技術大学院大学  
Japan Advance Institute of Science and Technology

ルのなかでも大気大循環モデルは、地球上の大気の流れを計算機上で再現するものである。モデルの支配方程式は、運動方程式、熱力学第一法則、水物質の保存則等である。これらを連立偏微分方程式の時間発展問題として数値的に解を求めることが力学過程と呼ばれる計算フレームである。一方、雲の物理過程と水の相変化に伴う熱交換（雲粒の生成、凝結等）や放射伝達過程、光化学反応や植生等目的によって計算される物理過程と呼ばれる計算フレームがある。これら力学過程と物理過程の計算フレームにより数値シミュレーションを行う。

### 3. プロセス競合を考慮した動的負荷分散手法

#### 3.1 概要

これまでの気候モデルの負荷分散手法は、あらかじめ計算領域を均等にして負荷を均一化するという静的負荷分散が用いられてきた。このようなデータ並列だと物理過程の計算によるロードインバランスに対しては並列化効果を高めることは難しい。そこで本論分は、データ並列手法によって負荷均一化を試みてきた力学過程で、新たに動的負荷分散の仕組みの導入を検討した。これは気候モデルの水平解像度に影響されない力学過程計算で、計算量の不均衡を調節する仕組みとなる。従って、全球気候モデルのあらゆる解像度にロードバランスの効果が期待される。また、気候モデルのシミュレーションの実行時には、計算機がマルチユーザー環境のために専有することは難しく、並列プログラムを構成するプロセスのうち1つでも他のプロセスと実行ノードが重なると、並列プログラム全体の終了時間の遅延は避けられない。

今後さらに大規模並列化された気候モデルでは、このようなプロセス競合に対する手段を検討する必要性がある。しかしこの問題は、気候モデルの大規模並列化がここ数年のことであり、モデル自体の並列化の効率化を試みている段階であるため、このような実環境での問題はほとんど議論されていない。そこで本論文は、プロセス競合状態での動的負荷分散手法を検討する。

#### 3.2 プロセス競合を考慮した動的負荷分散手法

本方式では、マルチユーザー環境での並列プロセスの実行を対象とする。動的負荷分散の対象となる力学過程の計算は球面調和関数展開による格子空間から波数空間への変換によって行っている。計算手法ではフーリエ変換を行った後にルジャンドル変換を施す。この時、注目すべきはルジャンドル変換に用いる領域数である。本計算手法は、ルジャンドル変換に用いる領域数

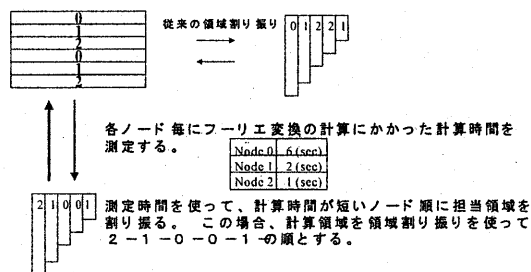


図1 計算領域の決定

はフーリエ変換に用いる領域数よりも少なく、また各PEの計算量も異なっている点が特徴である。例えば解像度がT106L20の計算では、フーリエ変換の計算時には最大160の領域で計算できるが、ルジャンドル変換の計算時には106の領域となる。従って、160ノードで計算した場合にはルジャンドル変換の計算には106ノードしか使わずに54ノード余ってしまうことになる。そこで我々は、計算途中のプロセス競合状態にあるノードを観測し、競合状態のノードの計算をルジャンドル変換の計算に使われていない（もしくは計算負荷の小さい）ノードへ動的に移行することにより競合を避ける手法を提案する。本論文で提案するプロセス競合を考慮した動的負荷分散手法は、負荷分散制御に用いる情報として、力学過程計算のフーリエ変換のルーチン時間を測定し指標としている。

プロセス競合を考慮した動的負荷分散手法の処理手順を示す図1。

- (1) 指定されたノード入れ替えのタイムステップであるか確認。
- (2) ノード入れ替えのステップであれば、力学過程計算のうちフーリエ変換にかかった時間を各ノード毎に測定する。
- (3) さらに測定した時間を使ってノードを昇順にソートする。
- (4) ソートされた順序に従って、次計算ステップで担当する計算領域を決定する。
- (5) 次計算領域の計算データを取得するため、各領域を担っていたノードから計算データをMPI通信により取得しノード入れ替えを行う。
- (6) 次計算ステップで担当領域を計算し、次回ノード入れ替えまで計算担当領域を変更せずに計算する。

計算領域の振り分けは図2のような領域振り分けを行い、計算量の均等化がされた。Barros等の並列化手法は、本論文で用いたNJR-SAGCMで実装されているものである。また、一般的には並列化手法は我々に

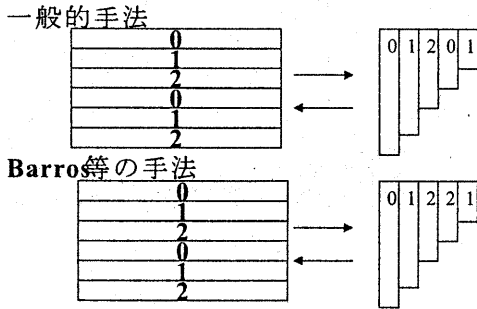


図2 力学過程の並列化手法

マシン	IBM RS/6000 SP(PPC)
プロセッサ	PPC604e
動作クロック	332MHz
トポロジー	SP スイッチ
PE 間通信性能	150MB/s

表1 並列計算機の仕様

Truncation	Physical Grid	Vertical Levels
T106	320 x 160	20

表2 気候シミュレーションサイズの分解能

よって NJR-SAGCM に実装し、両手法を選択できるようにした。一般的な手法では南北方向の領域をノード 0,1,2,0,1,2,... と割り振り、Barros 等<sup>2)</sup>の手法はノード 0,1,2,2,1,0,... の順に割り振る。両並列化手法は、Barros 等の手法のほうが各ノードに割り当てられる計算量が平均化されており、並列化効率が良い。

### 3.3 評価実験

本論文で使用した並列計算機 IBM RS/6000 SP(SP2)の仕様を表1に示す。IBM RS/6000 SPは1ノード4PE(PowerPCのシステム)によるシステム構成で、プロセッサ結合ネットワークにはSPスイッチと呼ばれるクロスバースイッチを用いている。

NJR用の入力初期データは、地球表面には陸地がなく全て海であるとする全球海面としている。その上にある大気は全球平均気温で風や雲もない気候的に安定で平均的な物理量の分布をしており、基本となる気候現象を表現するには十分なものである。

気候シミュレーションの空間分解能については、表2に示すような T106L20を用いてシミュレーションを行った。これら入力初期データを用いた場合、領域分割で最大の並列化可能数は T106L20では160ノードとなる。

評価実験では上記入力データを用い、IBM SP2にてシミュレーションを行った。本研究で用いた気候モデ

ルは、並列化手法として4種類のシミュレーションを行うことができる。それぞれは、各ノードへの割り振り方として、

- 一般的な手法
- Barros等の手法
- 一般的な手法にプロセスを考慮した動的負荷分散手法の組み合わせ
- Barros等の手法にプロセスを考慮した動的負荷分散手法の組み合わせ

である。これら4種のシミュレーションを、シミュレーション時間で160ノードで1日の積分を行った。さらにシミュレーション実行時には競合用プロセスとして Node 4~7に行行列計算のプロセスを継続して実行させ、プロセス競合時の動的負荷分散の効果を確認した。また、領域割り振りのタイミングは、3タイムステップ毎とした。

## 4. 考察

同じ物理条件で4種のシミュレーションを行った結果を図3, 図4, 図5, 図6, 図7, 図8に示す。これらはあるタイムステップで、ルジャンドル変換にかかる計算時間を total time に対して%で示したものである。図3, 図4に関しては、プロセス競合がない状態の結果である。ノード0~105までは与えられた計算領域のルジャンドル変換を行っており、ノード106~159までは計算領域が与えられていない。またプロセス競合状態では、図5, 図6の結果となった。それぞれプロセス競合があるノードで、約2倍の計算時間となっている。これは、プロセス競合により、並列プロセスが待たされていることを示している。この計算時間の遅延は、並列プロセスの同期合わせのタイミングで、全てのノードが最も計算時間がかかったノードの終了を待つことを意味する。従って、この結果の場合、プロセス競合がない場合と比べて約10%程度の実行時間が延長することになる。一方、同じくプロセス競合状態の図7, 図8では、遅延時間と思われるノードが確認されない。その理由は、提案した動的負荷分散手法の効果によって、プロセス競合が発生しているノードの領域計算を、本来計算をせずに Wait しているはずのノード(106~159のうちどれか)に計算領域を割り振った為である。そのためにノード4~7の計算時間の遅延は読み取れない。この結果、先の図5, 図6の結果に比べて、約2倍の速度向上を示している。これは提案手法を用いることにより、ルジャンドル変換の計算ルーチンに関してはプロセス競合によって計算時間が影響を受けにくいと理解できると考えられる。

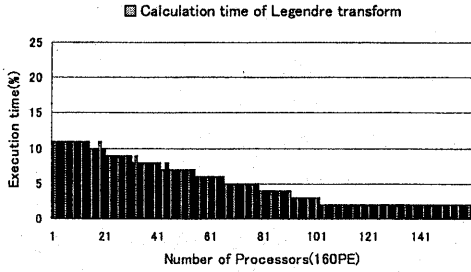


図3 プロセス競合のない一般的な手法

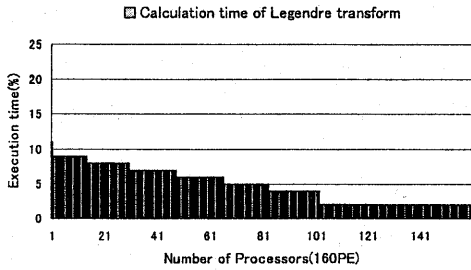


図4 プロセス競合のないBarros等の手法

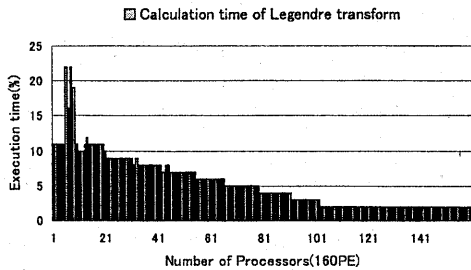


図5 プロセス競合のある一般的な手法

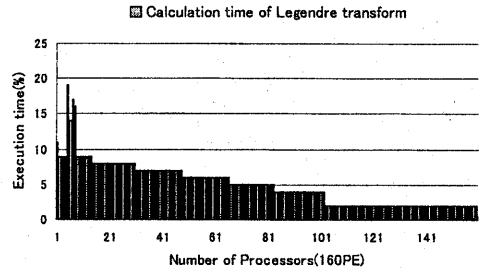


図6 プロセス競合のあるBarros等の手法

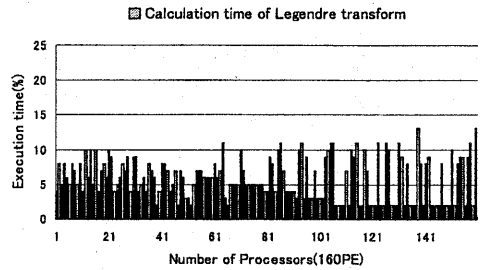


図7 一般的な手法を用いたプロセス競合の動的負荷分散手法

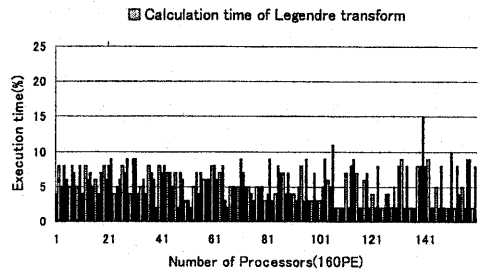


図8 Barros等の手法を用いたプロセス競合の動的負荷分散手法

## 5. おわりに

本論文では、力学過程の計算特徴を考慮することにより、プロセス競合が起きやすいマルチユーザ環境でもプロセスを効率良く分散させる動的負荷分散手法を提案した。本提案手法は、ルジャンドル変換で計算できる波数よりも多くのノードを使ったシミュレーションでは、プロセス競合時には並列化効果を上げることがわかった。これは一般の並列シミュレーションに見られる、大規模並列化による並列化効率の減少という特長に対して、本提案手法を用いることによりマルチユーザ環境では、大規模並列時でも並列化効率の減少を抑えられることを意味する。今後は本手法を用いて、

各種分解能の並列気候モデルにて評価実験を行う予定である。

## 参考文献

- 1) Toshiyuki Asano, Takahiro Inoue, Susumu Horiguchi, *Parallel Performance of Global Climate Models NJR on Parallel Computers*, The 3rd International Workshop on Next Generation Climate Models for Advanced & High performance Computing Facilities, pp. 28-30(2001).
- 2) S.R.M.Barros, T.Kauranne, *On the parallelization of global spectral weather models* Parallel Computing, 20, 1335-1356(1994)