

# 最先端並列計算機における次世代気候モデル開発に 係わるワークショップ報告\*

山 岸 米二郎\*<sup>1</sup> 室 井 ちあし\*<sup>2</sup> 保 坂 征 宏\*<sup>3</sup>

## 1. はじめに

標記ワークショップが、東京大学気候システム研究センター (CCSR) 所長住教授と NCAR の全球大気力学と気候研究部長 Blackmon 両氏をコンビーナーとして、1999年3月1日から3日間ハワイ大学東西センターで実施された。この会議は科学技術振興調整費総合研究「高精度の地球変動予測のための並列ソフトウェア開発に関する研究」の大気・海洋分野分科会の活動として、開催された(財団法人高度情報科学技術研究機構(以後 RIST と略称)主催、科学技術庁後援)。参加者は延べ83名(日本からは38名、その他は米国)で、最先端並列計算機、次世代気候モデル、計算機メーカーの現状と将来、討論の4つのセッションで実施され、29の発表(日本から12)があった。

我が国の気象界でも分散主記憶型の並列計算機がかなり利用可能になっているので、会員の参考に供するために、並列計算を重点に並列計算機、気候研究、モデル開発、討論の項目に分けてワークショップの内容を報告する。最後に3人の感想も付加した。なおこれまで並列計算機の記事が少なかったので、付録に簡単な用語解説をつけた。

## 2. 並列計算機

現在我が国で開発されている地球シミュレータ(ES)の全体計画が、谷(地球シミュレータ研究開発センター)によって紹介された。ESは全部で640N(ノー

ド)あり、1Nあたり8PE(プロセッサエレメント)、1PEのピークスピードが8Gflops(1ギガフロップス=每秒 $10^9$ 浮動小数点演算)、記憶容量は1N当たり16G(ギガ)バイトの共有メモリーという構成である。全体で記憶容量10T(テラ= $10^{12}$ )バイト、ピークスピード40Tflopsという文字通りのベクトル型超並列計算機である。この振興調整費研究はESの活用を目指した並列ソフト開発研究である。

2002年春の運用開始を目指して着々進行中の計画の具体的紹介なので、米国側参加者から強い関心が示された。関心の1つは勿論計算機ハード関連で、使用電力量など細部についてまで質問があった。関心のもう1つは膨大な計算結果の処理である。データ処理は基本OSとか画像表示技術等に加え、シミュレータからユーザまでのデータ転送、ユーザが個別に使用するハード等、シミュレータ計画だけでは解決できない事項にも関係するので直ちに解決案が示され得るわけではない。ユーザにとっては計算機の性能向上により高分解能モデルの運用が可能になったとき、データの洪水に溺れずにそこからどのように効率的に有効な情報を得るかという挑戦となる。

Buzbee (Buzbee Enterprises) は今後の HPC (High Performance Computer) のアーキテクチャが、DSMタイプになるとの予想のもとに、計算機の性能と価格の見通しを示した(第1表)。

つまり2004-2007年の間に実効速度(sustained speed)1Tflopsの計算機が1000~2000万USドルの価格で入手可能という予想である。なおここでは実効速度はピークスピードの5-10%と仮定されている。

## 3. 気候研究

住(CCSR)はセッション開始前の基調的講演で気候の自然変動の理解向上の必要性、高分解能モデルの必要性を述べ、物理過程に関する知識の向上の果たす役

\* International Workshop on Next Generation Climate Models in the Advanced Computing Facilities.

\*<sup>1</sup> Yonejiro Yamagishi, 財団法人高度情報科学技術研究機構。

\*<sup>2</sup> Chiashi Muroi, 気象庁予報部数値予報課。

\*<sup>3</sup> Masahiro Hosaka, 気象研究所気候研究部。

© 1999 日本気象学会

第1表 計算機の性能と価格の見通し (Buzbee).

年	100万 USドル (\$M) 当たりの実効速度	
2001年	6-12	Gflops/\$M
2004年	24-128	Gflops/\$M
2007年	96-192	Gflops/\$M

割等について論じた。なかでも分解能を向上させただけでは改善が保証されない水循環過程の理解と扱いの重要性を強調し、それが高分解能モデル成功の鍵であると主張した。また今は観測資料が不十分なのではなく、衛星などによる膨大な観測結果を統一的に理解する知的作業が必要とされている時代であるという現状認識を述べた。Gates (LLNL) は ACPI (American Climate Program Initiative) の包括的紹介を行った。ACPI 計画の柱は、気候予測に Tflops 級計算機を確保すること、モデル開発の支援、モデル開発のインフラストラクチャの整備、気候変化に関する情報、特に地域スケールの気候情報の流通の促進の4つである。インフラとしては計算機と気候モデルに加えて、通信ネットワークの整備が重視されている。包括的な計画だが、予算も決まっていない段階なので計算機仕様等の具体的な「もの」は何も見えない。

大気・海洋結合モデル (Blackmon, NCAR ; Mechoso, UCLA 等) や高分解能 (1/10度×1/10度) 海洋モデル (Semtner, UCAR 等) によるシミュレーション結果の議論、大気の全球モデルと地域モデルを two-way にした大気・海洋結合モデルの報告 (Suarez, NASA/GSFC) 等があったが、個別の紹介はやめ、気候予測の信頼度を如何に高めるかの議論をまとめる。Blackmon は過去100年の気候の再現実験に続き、複数の二酸化炭素濃度増加率のもとでの21世紀の気候予測結果を示して議論した。モデルが1940年代を中心とした全球平均地上気温の上昇をシミュレートしていないこと、温暖化予測で二酸化炭素濃度増加率の小さいケースが短期間ではあるが、増加率の大きいケースよりも高い全球平均地上気温を示したことなどを示し、気候の自然変動の大きさの程度、原因を明らかにすることの重要性を指摘した。また多くのモデルによる多くの初期条件でのアンサンブル実験で、予測のばらつきの程度を明らかにして、気候予測の信頼度を高める必要性を強調した。気候予測におけるアンサンブル実験の必要性には誰も異存はなかろう。計算機の性能向上によりこれを実現できる時期が近づいてきたとい

第2表 気候変動と気候変化の理解及びその影響を予測するために必要な結合気候モデルの実験事例数及び必要な計算時間の見積もり (Semtner).

実験目的	積分年数 /1実験	事例数	全積分年
自然変動	300年	50アンサンブル	15000年
観測時代の気候	150	20初期状態	3000
二酸化炭素変化による気候変化	600	4×10 アンサンブル	24000
硫酸塩/生物体存在量の効果	200	10例/10例	20000
エネルギー使用戦略	500	5×10 アンサンブル	25000

うことである。

Semtner は気候変動と気候変化を理解するために必要なアンサンブル実験の計算量の具体的見積もりを提示した(第2表)。彼は湾流や黒潮など幅の狭い強い流れによる熱輸送を正しく評価するために、海洋モデルは1/10度の水平分解能が必要と主張している。また結合モデルで大気と海洋を同じ計算精度で扱うのに海洋と大気モデルの水平分解能の比は2/9とする必要があるとし、第2表の実行に要する計算時間を推定した。

彼は2015年には実効速度1ペタ(=10<sup>15</sup>)フロップス級の計算機が実現し、分解能1/10度の海洋モデルと相当する分解能の大気モデルの結合モデルで、第2表、第1欄の目的のためのアンサンブル気候予測実験が可能になるとしている。

#### 4. モデル開発

モデル開発の議論は2つあった。1つは並列効率の観点であり、もう1つは限られた資源(マンパワー、計算機)で効率的にモデルの高度化を図る方策の議論である。

先に述べたように HPC のアーキテクチャーは並列マシン(DSM または MPP)になるだろうから、並列効率向上の努力は避けて通れない。Mechoso は512 PE (T3E900, 1PE/1N) を用いて大気・海洋結合モデル(大気に327PE, 海洋に185PE)を36.6Gflopsの実効速度で計算できたことを報告した。Semtner は NCAR を中心として実施された並列気候モデル(PCM)の計算効率評価実験(T3E900使用)について報告した。海洋部分の結果を第3表に示す。カップラーとは大気、海洋、海水モデルを別々のモジュールにしたとき、相

第3表 モデルの1年間積分に必要な所要時間（ウォールクロックタイム）(Semtner).

プロセス数	海洋部分 (時間)	海水部分 (時間)	カップラー (時間)	合計 (時間)
16	4.50	5.75	0.09	10.34
32	2.33	1.82	0.07	4.22
64	1.15	0.89	0.05	2.09
128	0.59	0.46	0.04	1.09
256	0.31	0.23	0.04	0.58
512	0.20	0.12	0.04	0.36

互の運動量, エネルギー, 物質(水分等)のフラックス交換量の内挿計算等を実行するモジュールをいう。なお第3表の場合は海洋と海水のみである。また陸面水収支過程を大気とは別のモジュールにする構成の報告もあった。

第3表によれば海洋モデル全体ではPE数を16から512と32倍に増やしたとき, 所要時間(ウォールクロックタイム)が10.34時間から0.36時間と1/28.7になり, よい scalability を示している。第3表によるとカップラーの計算は並列効率が悪く, 16PEでは全計算時間の0.87%を占めるに過ぎないが, 512PEでは全体の11.1%となっている。カップラーはモジュール間の分解能の違いを吸収する内挿計算等を実行するが, 計算量が少ない割に通信が多いためであろう。結合モデルの構成を検討するとき留意すべきことと思われる。大気部分(CCM3)は粗水平分解能(T42)であるが, 多くのPEを利用可能とするため2次元領域分割とした結果が示された。16PEから256PE(PE数16倍)で所要時間が11.78時間から1.30時間(1/9.1)に減少していて scalability はかなり良い。Semtnerは2次元領域分割を採用すれば, 将来大気モデルを高水平分解能にした場合, PEが数千の並列計算機でもかなりよい scalability が確保されるとの見通しを述べている。勿論ここに示した結果(数字)は使用した計算機に依存する。特にベクトル計算機とスカラー計算機では領域分割の方式に差があり得る。

これからもソフトレベルで効率を高めるために利用者側での経験蓄積が必要だが, 一方では今後PEが1000個を越えるような大規模なものになるとき, OSなど計算機設計側での並列効率向上の一層の努力が望まれる。

放射や降水, 乱流等の物理過程, 大気と海洋の結合, 効率的で安定な計算方式, 並列効率の向上等々, 複雑化する気候モデルの開発・維持には多様な人材の協力作業が必要である。そのためにはモデルの構成のわか

りやすさ, 変更しやすさというようなことも考慮する必要がある。Duffy(LLNL)は多数のモデルを並列化した経験から portability の確保を強調した。現在多くの並列計算機はMPIを備えているから, ノード間通信をMPIで記述することが portability 確保の前提であろう。Anderson(GFDL)はモデルソフトの読みやすさと同時に変更の容易さを強調し, 気候モデルの大気・陸面水収支・海水・海洋の各部分をモジュール化してカップラーで結合する構造を示していた。このような構造はソフトウェア間での可搬性という意味での portability 確保の1つの方向であろう。モデルのわかりやすさ(読みやすさ)を破壊するような並列最適化はなすべきでないというのが彼の主張であった。

Kanamitsu(NCEP)とAndersonはモデル開発で多くの機関・グループ(大学, 研究所, 官庁の研究者等)の協力体制を推進するために, モデルの構造を統一(Common Modeling Infrastructure)あるいは標準化(American Standard for Climate Model)するという考えを議論した。portability を高める方策ともいえる。豊富な計算機資源と多数の研究者を有するところがセンターとなり, いろいろな機関, 研究者により提案される改良(例えばパラメタリゼーションのサブルーチン)を, 標準構造のモデルでテストして採否を決め, モデル高度化を効率的に行うという考えが根底にあるようである。モデルの統一といっても, 気象予測か気候予測かという目的によりモデル構造は当然異なるから, 複数のセンターが想定されているようである。魅力的な構想ではあるが, このためには多くの研究者がモデルに容易にアクセスできることが前提となる。多くのモデルがインターネット等で入手可能な米国と現在の日本の事情はかなり異なる。

日本側からは, 丸山(電力中央研究所)がCCM3を用いて, scalabilityなどを詳しく調べた結果を報告した。隈と室井(気象庁)は気象庁の数値予報モデル開発の将来計画, 並列化の方針等を紹介した。気象庁は全球モデルからメソスケールモデルまで同一のモデル(Unified Model)で運用するという方針を報告した。杉(気象研究所)は気候が温暖化したときの台風の発生数の変化をT106のモデルで調べた結果を示した。対流のパラメタリゼーションや海水温の分布を変えても温暖化したときの方が温暖化前よりも, 台風の発生数が減るという結果であった。山中(気象研究所)は気象研究所で開発された海洋モデルによるシミュレーション結果を紹介した。斉藤(気象研究所)は気象研

究所の非静力学モデルで、冬季の季節風の吹き出しの雲列、海陸風による対流の生成、梅雨期の集中豪雨時の持続する降雨帯等のシミュレーション結果を示した。山岸 (RIST) は NJR モデル (科学技術庁の方針のもとで、宇宙開発事業団 (NASDA) と海洋科学技術センター (JAMSTEC) の支援で、RIST が開発している並列気候モデル) の構成と並列の方針等を説明した。

## 5. 討論

セッション 4 の討論では、‘より多くの計算機資源’から派生する問題として、莫大なデータ量、可視化、計算機ネットワークの3つが議論された。シミュレーション結果の解析に今後も主にワークステーションが利用されるとするとデータ転送が大きな問題点になるというのが共通の認識であった。杉は高解像度モデルの実現には、並列計算機など計算機ハードの改善のみならず、セミラグランジュ法など、計算方式の工夫による計算効率化も重要であることを強調した。

今後の日米協力のトピックスとコンタクトパーソンが以下のように決められた。数値計算法と力学スキーム (Kanamitsu, Suarez, Anderson, 杉, 隈, 木本, CCSR), 解像度と気候予測 (Blackmon, 住), データ関連 (Mechoso, 中村, RIST), 静力対非静力 (Adcroft, MIT, 齊藤), 季節変動と季節内変動 (Kanamitsu, 木本), 標準化 (Zebiak, IRI, 杉, 隈)。次のワークショップを開く場合は、データと可視化のセッションを追加することになった。

## 6. 感想

ベクトル計算機という‘高級マシーン’を便利に利用できた日本が、並列計算機対応の応用ソフト開発で、米国に比して大きく遅れたのは当然である。しかしこれからは並列を避けては通れないので取り組み強化が必要である。応用ソフト技術の開発に限定すれば、後発の有利さを生かして追いつくという例のパターンで差を縮める見通しはたつ。

今は研究者が1種類のマシンを使用するだけで済む時代ではない。並列ソフトはマシンアーキテクチャーに大きく依存するし、並列化の作業はモデラーには少々煩わしい。計算機アーキテクチャーに関する知識を有して並列化を行うグループとモデラーとの協力が望ましい。協力の範囲は、モデラーが並列にできるだけ煩わされずにモデルの開発あるいは修正ができるようにモデル構造や変数構造を設計するというような面

から、並列効率と精度を勘案した計算方式の選択といった面まで幅が広い。ノード間通信の send, receive の message の挿入という作業だけでは無いことは明らかである。並列化を行う人が、大気・海洋の基本的知識を有していれば非常に助かるが、それは必ずしも望めない。ドキュメントが整備されたモデルの自由流通とか並列化に携わる職種の方の専門家としての評価の確立等により人材を確保して行く努力が前提となるう。

発表内容 (シミュレーション結果等) の密度の濃さ、計算機有効利用の試みの多様さ、モデル開発を有効に進めるための資源 (人材, ハード) 有効活用への意気込み等、総体的に見れば、日米にかなりの差を感じた。これはまねて追いつける類のものでないだけにいささか深刻である。

(山岸米二郎)

筆者は同じ数値予報課の隈 健一とともに本会議に参加し、気象庁の数値予報モデリングの現状と将来計画、およびその並列処理計算機への対応についてプレゼンテーションを行った。本会議で議論されたことからのうち、特に印象に残ったものについてその感想を記したい。

現在の並列計算機のアーキテクチャについては、ベンダー間でその戦略に若干の差があり、日本の地球シミュレータ、米国のアスキー計画といった最先端並列処理計算機の成否がその将来を握っていると言っても過言ではない。そしてアプリケーションプログラムが高い演算効率を上げるためには、そのアーキテクチャを強く意識したソフトウェア開発をする必要がある。これは、我々のような毎日の天気予報を目的とする場合は勿論であるが、能率よく研究を進めるためにも重要であるし、それによってより大規模な計算も可能になる。今後はよい意味でお互い刺激を与え、より高速な計算機が開発されることを期待したい。

米国のモデル統一化あるいは標準化計画については、その意気込みに少々驚いた。現在はプログラミングルールの統一に向けてほぼ定期的に作業部会を開催し、まず物理過程の可搬性を高めることを目標にしているという。マンパワーの面で不利な日本においても、今後は組織の垣根を越えた共同開発へ向けた取り組みが必要となるかもしれない。この計画に刺激され、数値予報課と気象研究所の有志との間でモデルの共同開発およびプログラミングルールの統一に関する意見交換を行っているところである。

3月とはいえ気候のよいハワイにおいて、スーツ姿でワイキキビーチを散策するのには多少違和感があった。また筆者は外国出張といえども何らかのトラブルに会うのであるが、今回は大勢の会議出席者が飛行機に乗りあわせ、バスでホテルと会場の間を移動するというきわめて快適なものであった。(室井ちあし)

絶好の好天続きのもと、会議に出ていなければならぬのを少々惜しく感じながら、気楽な立場で会議に参加させていただいた。

今回の会議では、米国がモデル構造の統一・標準化の方針を強力に押し進めている、という印象を強く受けた。一人が開発したプログラムを皆が容易に使えるようにする事が主たる目的のようである。欧州の一人勝ちの現状に刺激されてのことだとは思いますが、日本も何とかしないとイケないという気にさせられた。ただし、単に真似してもプログラマーの少なさを考えると勝ち目はないだろう。さまざまな意味で開発効率を上げることができるような足場固めからはじめる必要を感じた。

並列化への取り組みは、想像通り米国の方が進んでいた。日本の気象コミュニティーでは気象庁グループと(財)高度情報科学技術研究機構とが多少の経験を持っているにすぎないと思う。米国と日本との計算機事情の違い、マンパワーの不足等理由はあるが、早晚並列化が必要になることは明らかである。せめて並列化に向けたプログラムを作るよう、目を向けていくべきであろう。

もちろん、高解像度モデルに必要な物理過程・力学フレームの開発が必要であることは言うまでもない。

筆者にとって意外なことに、気象研究者および計算機ベンダの他に、計算工学関係者や可視化の専門家等も日本から参加していた。気象研究者と彼らとの間であまり交流はかれなかったようで、残念である。高速アルゴリズムの開発、大量データの取り扱いとその可視化など、我々にとって未知でかつ必要とする事柄は多い。今後も懲りずに加わってほしいと思う。

(保坂征宏)

## 謝辞

報告作成にあたり、気象庁気象研究所杉室長に大変お世話になりました。感謝致します。

## 略語一覧

CCM3: Community Climate Model (Version 3)

共用気候モデル (バージョン3)

DSM: Distributed Shared Memory

(分散共有メモリ型: 付録参照)

ES: Earth Simulator

(地球シミュレーター; 本文参照)

GSFC: Goddard Space Flight Center

(ゴダード宇宙飛行センター)

IRI: International Research Institute for Climate

Prediction (気候予測国際研究所)

LLNL: Lawrence Livermore National Laboratory

(ローレンスリバモア国立研究所)

MIT: Massachusetts Institute of Technology

(マサチューセッツ工科大学)

MPI: Message Passing Interface (付録参照)

NCAR: National Center for Atmospheric Research

(米国立大気研究センター)

NCEP: National Centers for Environmental Pre-

diction (米国立環境予測センター)

NASA: National Aeronautics and Space Adminis-

tration (米国航空宇宙局)

UCLA: University of California Los Angeles

(カリフォルニア大学ロスアンジェルス分校)

## 付録. 並列計算機に関する用語解説.

報告を読んでいただくための初歩的解説を心がけた。計算機の専門用語解説の観点では不備がありうるがご容赦乞う。

○ハードシステム・OSに関する用語

PE: 1つのCPUをプロセッサあるいはプロセッサエレメント(PE)とよぶ。

shared memory: 複数個のPEが1つの主メモリ空間を共有しているとき、共有メモリ(shared memory)とよぶ。各PEから同時に読み書きできる。

node(N): 1つのPEまたは主メモリ空間を共有する複数PEが1つの単位を構成しているときノード(Node)と呼ぶ。

distributed: 1つの計算機内で複数個のPEまたはノードが通信で接続されているとき分散(主記憶)型(distributed)と呼ぶ。計算機の仕様によりPEがベクトルプロセッサの場合もあるし、RISCプロセッサの場合もある。PEが他の主記憶の読み書き

可能な場合と不可能な場合の2つの形態がある。以後の説明は（暗に）後者を想定している。

並列計算機のアーキテクチャの呼び名は必ずしも統一されていないが、ここでは Semtner が使用した呼び名を示す。説明は著者が付した。

MPP (massively parallel processor)：非常に多数の PE が分散型で1つの計算機を構成している場合。RISC プロセッサはベクトルプロセッサに比較してスピードが遅いので、数百以上の PE からなる、このタイプの並列計算機である場合が多い。

DSM (distributed shared memory)：主記憶共有の複数 PE からなるノードがいくつか集まって分散型で1つの計算機を構成している場合。

○並列化手法に関する用語

OpenMP：主記憶を共有している PE (shared memory) 間の計算作業の分割では通信を必要としない。PE 間の作業分担の均等化や同期の回数の削減等の工夫により、並列効率が向上する。コンパイラが自動で作業の分担をする自動並列化とかコンパイラ指示行を用いる方式などがある。OpenMP は shared memory 内での作業分割でコンパイラ指示行を付加する標準的方式である。

MPI：分散メモリ型の並列計算機で、ノード間の情報交換を指示する方式で最もよく用いられるのが

MPI (Message Passing Interface) である。単純な具体例でノード間通信を説明する。1000×1000個の格子点からなる数値予報モデルがあるときに100×100の100個の領域に分け、100個のノードで並行してそれぞれの領域の計算を行う。各ノードは100×100個の格子と境界値の計算に必要な範囲（いわゆる袖領域、Shadow area：Halo region）の値を持つのが普通である。袖領域の値のアップデートが必要ときにノード間の通信を行う。1つ1つの作業当たりでは、CPU 内の計算に比べてノード間の情報交換は圧倒的に多くの時間を要する。並列ソフト作成には作業分担の均等化の他、情報交換が少ない時間で効率的に実行できる工夫が必要である。

○その他の用語

Scalable, Scalability：ある計算機で使用する PE の数を増やしても計算効率（1PE あたりの実効スピード）が変わらなければ、この並列ソフトは scalable であるという。プログラムサイズを変えずに PE の数を増やすと、通信に要する時間が CPU 内での計算時間に対して相対的に増える等の理由で、一般に計算効率は低下する。

Portable, Portability：ある並列ソフトが機種が異なる並列計算機でも実行可能なこと、あるいはあるプログラムがソフト間で可搬的なこと。