

特別推進課題（平成 28 年第 2 期）に対する技術サポート報告

課題責任者

上原 均 海洋研究開発機構 地球情報基盤センター 情報システム部計算技術グループ

著者

齋藤 友一^{*1}, 池田 美紀子^{*1}, 上原 均^{*1}

*1 海洋研究開発機構 地球情報基盤センター 情報システム部計算技術グループ

地球シミュレータの平成 28 年度第 2 期特別推進課題に対して地球情報基盤センターが実施した技術サポートについて報告する。

キーワード：特別推進課題, 地球シミュレータ, 技術支援

1. はじめに

地球シミュレータの平成 28 年度第 2 期特別推進課題は平成 28 年 8 月から平成 28 年 11 月にかけて実施された。特別推進課題の推進と地球シミュレータの運用管理を行う地球情報基盤センターでは、各課題に対する利用高度化と運用管理面からの技術支援を全面的かつきめ細やかにを行い、研究促進と成果創出に貢献した。以下では、その技術サポートの概要と種々の技術的なトピックスを述べる。

2. 技術支援の詳細

2.1 各課題と技術サポート担当

表 1 に示すように各特別推進課題に 1 名の担当者を設定して、技術サポートにあたった。

各課題の研究目標や学術的詳細については、各課題からの報告をご覧ください。

技術支援として、利用プログラムの地球シミュレータへの移植、計算結果確認、性能評価、高速化、効率的なプログラムの実行条件調査、ジョブスクリプト作成支援などを行った。運用上のジョブ数制限や実行時間制限などの各種利用制限の緩和のほか、利用上の疑問点や不具合にも随時対応して、効率的な研究実施環境の維持・向上にも努めた。このように、プログラム実行前からプロダクトラン、後処理に至るまで、全面的かつきめ細かな技術支援を行った。

2.2 課題 #1 「全球非静力学大気モデルを用いた台風特異年を対象とした高解像度・大アンサンブル実験」への技術支援

2.2.1 課題概要

本課題は第 1 期からの継続課題であり、第 1 期では全球雲解像モデル NICAM (Non-hydrostatic Icosahedral Atmospheric Model) を用いて、地球温暖化の停滞 (ハイエイタス) 期間およびその終焉後を想定したアンサンブル実験を多数実施したが、今期は 2016 年をはじめとする、台風発生が極端に遅れた特異年における台風の再現実験を行った。NICAM は地球シミュレータにおいて十分動作実績があり、且つ第 1 期において多数アンサンブル実験もすでに実施されているため、移植や計算性能最適化、パラメータ最適化などの準備期間を設けることなく実行を開始できた。今期も第 1 期に引き続き、多数アンサンブル計算を効率よく実施するための技術支援を行った。

2.2.2 実施環境 (計算資源)

本課題への計算資源割り当ては以下の通り。CPU 資源やディスク資源の利用内訳については、次で述べる。

CPU 資源：600,000 ノード時間

ディスク資源：HOME 領域 128GB、DATA 領域 300TB

表 1 各特別推進課題と技術支援担当者の一覧

No.	課題代表者・所属	課題タイトル	担当
#1	小玉 知央 海洋研究開発機構	全球非静力学大気モデルを用いた台風特異年を対象とした高解像度・大アンサンブル実験	池田美紀子
#2	青野 光 東京理科大学	高精度流体シミュレーションによる小型ファンから発生する音響予測レベルの確認と向上による「騒音」という環境課題の改善にむけて	齋藤友一

2.2.3 計算モデル及び計算条件

計算モデル及び計算条件を以下に示す。

計算モデル：全球雲解像モデル NICAM
(水平解像度 約 14km)

アンサンブル数：50 (2 ケース)

1 ジョブあたりの使用ノード数：40

1 アンサンブルの計算には 5,000 ノード時間 (40 ノード × 125 時間, 5 ヶ月積分) を要する。台風発生が遅れた 2016 年及び 1998 年の 2 ケースを計算対象とし、50 アンサンブル実験を実施した。1 アンサンブル計算のデータ出力量は約 10.5TB で、総容量は初期値ファイルをあわせて約 535TB になる。ポスト処理においてデータ圧縮を行い (圧縮前 10.4TB、圧縮後 2.4TB)、随時アーカイブにデータを移動しながら計算を実施した。

2.2.4 技術支援

第 1 期同様、12 アンサンブルを同時に実施するため、地球シミュレータの利用制限値を以下の表 2 のように緩和した。設定変更は基盤システムグループの協力による。

短期間で計画的に大量のジョブを実行する必要があるため、密に計算実施者とジョブ実行状況や ES 運用状況の情報をやり取りし、計算ノードを占有するか否かなど、実行の仕方について随時検討しながら実験を実施した。

表 2 ジョブ実行制限の緩和

	変更前	変更後
グループあたりの最大ジョブ実行数	50	100
ユーザあたりの最大ジョブ実行数	10	50

2.2.5 課題 #1 への技術サポートのまとめ

第 1 期と同様に、短期間で大量のジョブを実行する必要があったが、予定通りジョブ実行を完了することができた。状況によってはノードを占有してジョブを実行する予定であったものの、最大ジョブ投入数の制限値を緩和したこと等により、実施期間中、効率よくジョブが実行された。実行により第 1 期とあわせて 1PB を超えるデータセットが作成されたが、データ保存やデータ配信について、引き続き支援を行いたい。

2.3 課題 #2 「高精度流体シミュレーションによる小型ファンから発生する音響予測レベルの確認と向上による「騒音」という環境課題の改善に向けて」への技術支援

本課題の研究は、地球シミュレータ利用することで実現可能になる小型ファンの大規模数値流体解析を行うことが目的であった。技術支援では、地球シミュレータで計算するモデル「LANS3D」の移植、性能評価、高速化およびジョブの最適な実行方法の検討を中心に行った。

2.3.1 モデルの移植と動作検証

LANS3D はこれまで「SX-9」、「京」での動作実績があったが地球シミュレータでの動作実績がなかったため、は

じめにモデルの移植と動作検証を行った。

1) 移植

初期化漏れの確認を行うためコンパイルオプション `-init heap=nan stack=nan` を指定し実行エラーのチェックを行った。計算結果に影響のない箇所であったが 1 箇所の初期化漏れについて修正を行った。

2) 動作検証と計算時間の測定

動作検証のため計算ステップ数を 1000 程度に設定し、格子点数を変えた複数のケースで結果を検証した。また、計算性能を把握するため、以下のケースについて 1 ノード 1 コア、2 コア、4 コアで計算時間を測定した。測定結果からコア数を増やした計算の加速率が得られ、ハイブリッド並列の有効性を確認した。(図 1)

- ・ 格子点数 50 万 (1node_0p5MPts)
- ・ 格子点数 100 万 (1node_1MPts)
- ・ 格子点数 500 万 (1node_5MPts)
- ・ 格子点数 1000 万 (1node_10MPts)
- ・ 格子点数 2000 万 (1node_20MPts)
- ・ 格子点数 3000 万 (1node_30MPts)

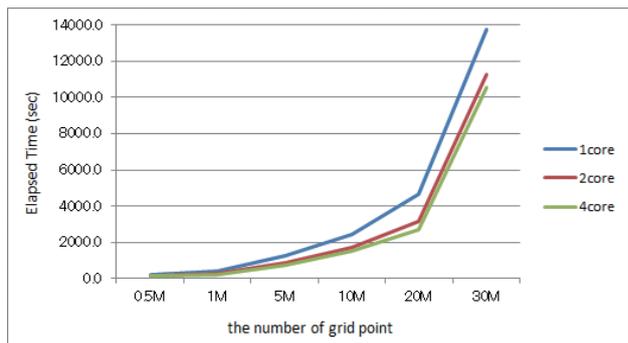


図 1 格子点数、使用コア数別の計算時間

2.3.2 性能測定と高速化

研究課題で計算する最大総格子点数約 5 億 (500 プロセス、プロセスあたり 100 万格子点) を想定し、高速化の評価は 1node_1MPts で性能を測定した。取得した ftrace 情報から時間コストの高いサブルーチンについて高速化作業を行った。

1) サブルーチン A、B の高速化

時間コストが高くバンクコンフリクトが発生していると考えられる DO ループについてループ分割によるベクトル化を行った。

2) サブルーチン C の高速化

インライン展開のオプションを外した性能測定を行い、時間コストの高いサブルーチンの調査を行った。依存関係のないループに対し アウターアンロールを指定し性能測定を行ったところ効果が見られた。その後、インライン展開のオプションを戻し性能測定を実施し性能向上を確認した。

3) その他

オリジナルのコード中に 並列化指示行 (`!CDIR PARALLEL DO`) の指定誤りにより有効となっていな

かった箇所があったためこれを修正（有効化）した。

4) 高速化のまとめ

上記作業の実施により、1 ノード、4 コアの性能はオリジナルの 213.4 秒から 124.4 秒（約 1.72 倍）に高速化された。

2.3.3 マルチノードの性能測定

性能測定には格子点数 $840 \times 840 \times 840 = 5.9$ 億点を 64 プロセスから 1000 プロセスで実行を行った。測定結果を以下の図 2 に示す。

本計算で想定する計算規模において測定の結果オリジナルに比べ高速化版は約 1.3 倍高速化された。なお、図中の 16 ノード実行 (16,64,1) はメモリ不足となった。

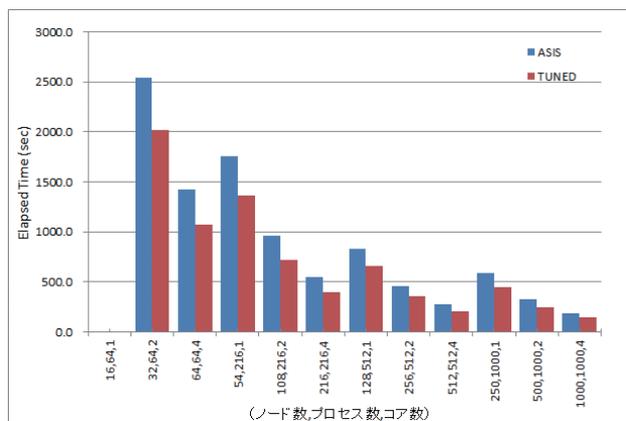


図 2 使用ノード数、プロセス数、コア数別の計算時間

2.3.4 ジョブ実行の検討

2.3.3 での測定結果により MPI のみを使用して並列化するフラット MPI より MPI とマイクロタスクを使用して並列化するハイブリッド並列の並列化率は高い結果となった。

本計算においてはこれらの測定結果および以下の点を考慮しジョブ実行の計算規模を検討した。

- ・ システム全体のジョブスケジュール混雑状況
- ・ 計算規模によるジョブ実行の待ち時間
- ・ 多数のジョブ実行

測定結果は 1000 ノードでの実行が最速であったがこれらの点を考慮し、256 ノードで本計算を進めることを研究課題実施者に提案した。

2.3.5 占有ノード利用の検討

本計算の実行は連続して 256 ノードを使用して計算を進め、その後約 3 週間ノードを占有（占有ノード、512 ノード）し計算を進めた。

2.3.6 課題 #2 への技術サポートのまとめ

技術支援では、モデルの移植から性能評価、高速化を行い、その後マルチノードでの性能測定を行った。更に本計算において最適なジョブ実行規模を検討し、高速化を反映したプログラムで計算を進めることができた。運用面では実施期間の後半には占有ノードを用意し研究の推進を図ることができた。

3. おわりに

特別推進課題は、限られた期間内、限られた人的リソースの中で、新 ES で画期的なサイエンスとしての成果を出さなければならない、という非常に強いプレッシャーのもとで実施された。その中でサポート要員も従来以上に業務で創意工夫し、技術的な知見を深めた。これらは今後の特別推進課題や他の課題の技術サポートに活かしていきたい。

謝辞

各課題の先生方や関係各位には技術支援担当者との頻繁な打ち合わせを寛容かつ前向きに受け入れていただきました。また地球シミュレータ運用面では地球情報基盤センター情報システム部 塚越部長、基盤システムグループ各位に大変なご尽力をいただきました。さらに運用面での調整やプログラムの移植・最適化等で日本電気株式会社の関係各位に大変なご助力をいただきました。ここに記して感謝の意を表します。

Report of Technical Support for Strategic Project with Special Support (Aug. 2016 – Nov. 2016)

Project Representative

Hitoshi Uehara Center for Earth Information Science and Technology, Japan Agency for Marine-Earth Science and Technology

Authors

Yuichi Saito^{*1}, Mikiko Ikeda^{*1} and Hitoshi Uehara^{*1}

^{*1} Center for Earth Information Science and Technology, Japan Agency for Marine-Earth Science and Technology

Using the Earth Simulator opened in March 2015, two research projects were completed in the Strategic Project with Special Support from Aug. 2016 to Nov. 2016. This report covers the details of the technical support for these research projects.

Keywords: Strategic Project with Special Support, Earth Simulator, Technical Support

1. Report

Strategic project with special support is a short-term project spanning four months. Significant results were achieved using the state-of-the-art supercomputer, the Earth Simulator. To deliver significant results within a limited period, technical support in all aspects of high performance computing played an important role.

In the Strategic project with special support from Aug. 2016 to Nov. 2016, Technical staff took charge of the support for each project (as shown in the Table 1). The technical staff proactively provided support for all aspects of the research project implementation.

Project 1 was continued from the previous phase. In the previous phase, a lot of ensemble experiments for the period during and after a global warming stagnation period “hiatus,” were carried out using the global cloud-resolving model “NICAM”(Non-hydrostatic Icosahedral Atmospheric Model). In this phase, reproduction experiments of typhoons were performed in the years, when typhoon outbreaks were extremely late.

Fifty ensemble jobs were performed with a horizontal

resolution of 14 km using 40 nodes per job and resources allocated for calculation in this project were 600,000 node-hours.

We relaxed the usage limits on the Earth Simulator, as shown in Table 2 below, so as to be able to simultaneously execute 12 ensemble jobs.

Table 2 Relaxation of the execution limits

	Before	After
Number of concurrent job executions per group	50	100
Number of concurrent job executions per user	10	50

Because it was necessary to execute a lot of jobs systematically in a short period of time, we closely exchanged information on the jobs execution status and the Earth Simulator operation status with the person in charge of the jobs execution. The experiments were then carried out while considering how to best execute them, for instance whether or not to occupy some calculation nodes from time to time.

While it was planned to occupy some calculation nodes to execute jobs in time, the jobs were efficiently executed during the implementation period due to a relaxation of the maximum value on the number of concurrent job executions. The data set exceeding 1 PB was created in this issue and the previous phase. We would like to continue to provide support for data storage, data distribution and so on.

Project 2 was designed to perform high-accuracy fluid dynamic numerical simulation of aerodynamic noise radiated from a propeller fan by the program LANS3D using the Earth Simulator. As technical support, we examined the program optimization and tuned the program as well as the job execution

Table 1 Each Project and Support Staff

No.	Project Name	Name of Support Staff
#1	High-resolution Large-ensemble Experiments for Anomalous Tropical Cyclone Year Using Global Non-hydrostatic Atmospheric Model	Mikiko Ikeda
#2	High-accuracy Fluid Dynamic Numerical Simulations of Aerodynamic Noise Radiated From a Propeller Fan Towards the Improvement in Noise-related Environmental Problems	Yuichi Saito

size and optimal execution method.

We first checked that the LANS3D is clean (i.e., non-initialization and an access of array outside). Next, in order to grasp the rate of performance of the program, we measured the calculation time using a single node (1 core/2 cores/4 cores).

The number of grid points and calculation time by the number of used cores are shown below in Fig. 1.

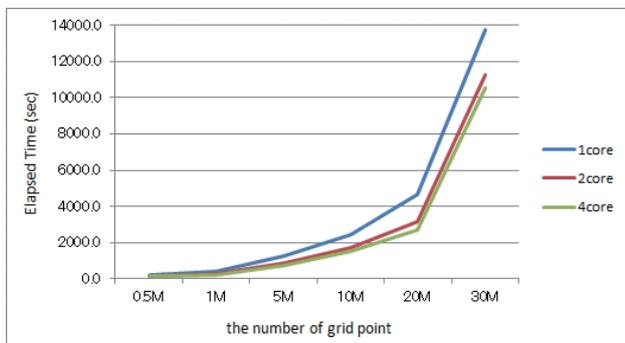


Fig. 1 Measurement results under various conditions

We measured the performance at the assumed total number of grid points of about 500 million (500 processes, 1 million grid points per process). For the evaluation, we used a case of 1 node_1MPts (the number of grid point here is 1 million) and we tuned the high cost subroutine.

As a result of tuning, the performance of using 1 node and 4 cores was increased from the original 213.4 seconds to 124.4 seconds (about 1.72 times).

For the performance measurement, a total of $840 \times 840 \times 840 = 590$ million grid points were executed from 64 processes to 1000 processes. The measurement results are shown below in Fig. 2.

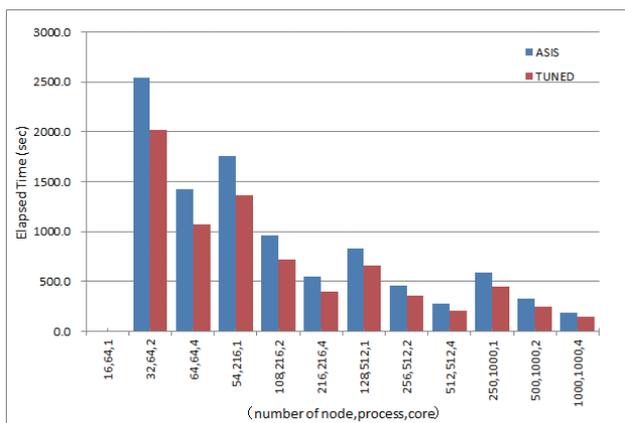


Fig. 2 Measurement results of program scalability

The measurement results of the tuning version, compared with the ASIS version, was sped up by about 1.3 times.

In addition to the result of the performance measurement, taking into account the usage of the system and the job schedule, we considered the optimal number of nodes and, as a result, we proposed to use 256 nodes.

For the purpose of promoting job execution, this project implemented occupation use of nodes for about 3 weeks.

By providing above support for each project, the support staff gained high-level technical knowledge. We will make full use of this knowledge to support other research projects.

Acknowledgement

All project members accepted meetings with us, even though the meetings were frequent. The Director, Mr. Tsukakoshi, and the Earth Simulator operation staff in CEIST strongly supported projects from the aspect of system administration. The NEC staff were invaluable to the in successfully providing support. We are deeply grateful to them.

