

特別推進課題（平成 28 年度第 1 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 1 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 1 期特別推進課題は平成 28 年 4 月から平成 28 年 7 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いただきたい。

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 「近未来及び 21 世紀末気候の北上する台風の雲解像モデルを用いたダウンスケーリング実験」への技術支援

本課題は昨年度の 3 月～5 月期および 6 月～9 月期に実施した課題の継続課題である。今回は大気モデル（雲解像度モデル：CReSS）を用いた多数の台風のダウンスケーリング実験の技術支援を行った。

過去の課題では、非静力学大気波浪海洋結合モデル（CReSS-NHOES）を用いてシミュレーション実験を行うため、以下のモデルに対して地球シミュレータへの移植、動作検証からシステムの特徴にあわせたモデルのバージョン選定および高速化作業の技術支援を行った。

- ・海洋モデル（波浪モデル：NHOES）
- ・大気モデル（雲解像度モデル：CReSS）
- ・大気海洋結合モデル（CReSS-NHOES）

本課題では大気モデル（雲解像度モデル：CReSS）を使用して、近未来気候と今世紀末の気候の台風 60（事例）、合計約 120（事例）のシミュレーションを実施するための技術支援を行った。

2.2.1 ジョブ実行、計算規模の検討

大規模かつ多数の計算を行うため、課題の実施期間内に効率良く計算が進むジョブ実行規模について課題研究者と検討を行った。

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	坪木 和久 名古屋大学	近未来及び 21 世紀末気候の北上する台風の雲解像モデルを用いたダウンスケーリング実験	齋藤友一
#2	小玉 知央 海洋研究開発機構	近未来台風予測は実現できるか？ – 全球非静力学大気モデルを用いた高解像度・大アンサンブル実験	池田美紀子

ここで地球シミュレータのシステム構成、運用・利用方法を紹介した後、検討方針および検討結果について説明する。

地球シミュレータは5,120ノードで構成され、2,048ノードの拡張クラスタと512ノードの基本クラスタ(6クラスタ)に分かれている。システムの利用には計算規模にあわせて各クラスタを使い分けている。更に、1つの基本クラスタではデバッグ等の用途にインタラクティブノードを16ノード、プリ・ポスト処理等の小規模実行用途にS系バッチノードを48ノード用意している。その他の5,056ノードは中規模、大規模計算を行うL系バッチノードとなっている。

1) 計算ノードの使い分け

- ・インタラクティブノード(16ノード)
(主にセルフ開発、デバッグ等で利用する。)
- ・S系バッチノード(48ノード)
(主に小規模ジョブの計算、プリ・ポスト処理等で利用する。ファイルのステージングは不要。)
- L系バッチノード(5,056ノード)
(通常の計算、中規模から大規模ジョブの計算で利用する。ファイルのステージングが必要。)

2) 各クラスタの計算規模による利用方法

- ・基本クラスタ(計算ノード512ノード):
1～512ノードを使用するジョブ
- ・拡張クラスタ(計算ノード2,048ノード):
256～2,048ノードを使用するジョブ

512ノード以上を使用するジョブは拡張クラスタで実行することになるが、256ノードから512ノードを使用するジョブについては基本クラスタ、拡張クラスタの両クラスタを利用者が選択して実行できる仕様となっている。

上記システムの利用形態とシステム全体のジョブスケジュールを考慮し、1台風あたりの計算で使用するノード数を検討した。

CRESSは計算領域の分割数から512ノード以上の実行も可能だったが、スケジューリング効率面から判断すると512ノード程度の計算規模になると1つの計算は短時間で終了するが、ジョブ実行のために必要なノード数を確保する時間がかかるため、複数の台風計算を同時に効率良く進めていくには困難と予想した。計算にかかる時間を抑えるため、ある程度のノード数を使用し且つ複数の計算を並行して効率的に進められる計算規模は256ノードと判断し、この計算規模で進める案を研究者に提案した。また、ジョブの投入時には、スケジューリング状況を見ながら基本クラスタと拡張クラスタへのジョブ投入を使い分けて実行することを推奨した。

- ・計算規模(1台風計算)
L系ジョブ(256ノード、約20時間)

2.2.2 ポスト処理の検討

計算結果のポスト処理として、小規模実行用のS系ジョブで計算結果のデータを結合する処理を行った。CRESSは大規模なデータ処理を行うため30時間程度かかる見積

もりとなった。通常、S系ジョブは最大8時間の実行制限で運用を行っているが、課題推進のため48時間まで実行可能な設定に変更することで対応した。

- ・ポスト処理(本計算結果の解析用処理)
S系ジョブ(1ノード、約30時間)

2.2.3 ディスク割り当ての検討

1台風あたりの計算出力量は1TB程度となる見込みであった。ディスク領域については本計算や検証計算が保存できる容量として、DATA領域(通常の利用者割り当て領域)を110TB用意した。また、計算が終了したデータの保存領域(MSS領域、ファイルステージング不可)200TBを別途本課題用に用意した。DATA領域100TB以上の割り当ておよびMSS領域の利用については期間内に計算を終了することを優先した特別な処置である。

- ・ディスク割り当て(DATA領域、MSS領域)
DATA領域 割り当て容量:110TB
MSS領域 割り当て容量:200TB

2.2.4 課題#1への技術サポートのまとめ

本課題の研究目標であった、CRESSを用いた120事例の台風シミュレーションは課題期間内に全ての計算を終了することができた。これは地球シミュレータのジョブスケジュールを考慮した最適な計算規模の選択と研究課題実施者の適切なジョブ投入により実現した。また、ポスト処理やディスク割り当てについて柔軟な運用を取ったことも重要であった。大規模計算によるディスク領域の確保や課題終了後の実施機関へのデータ回収方法については今後も検討が必要である。

2.3 課題#2「近未来台風予測は実現できるか？」

ー全球非静力学大気モデルを用いた高解像度・大アンサンブル実験」への技術支援

本課題では、数年から十数年先の近未来の台風予測の実現可能性を示すため、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。

NICAMは初代地球シミュレータにおいて開発が開始された計算モデルであるが、今では京コンピュータやその他各大学の大型計算機など、多種多様なシステムにおいて動作実績がある。地球シミュレータは2015年3月に3代目となる現システムに更新されたが、NICAMは地球シミュレータ現行機でもすでに計算性能最適化が実施されている。今回、地球シミュレータにおいて、多数のアンサンブル計算を短期間で実施するための技術支援を行った。

2.3.1 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：510,000 ノード時間

ディスク資源：HOME 領域 128GB、DATA 領域 300TB

2.3.2 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 アンサンブル実験には 5,000 ノード時間 (40 ノード×125 時間, 5 ヶ月積分) を要し、ハイエタス前後の台風変化を調査するため、1997 年と 2015 年の 2 ケースについて、50 アンサンブルの計算を実施した。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。よって、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。

2.3.3 技術支援

1) 地球シミュレータの利用制限値緩和

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 に示すように緩和した。設定変更は基盤システムグループの協力による。

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) NetCDF に関する技術支援

NICAM (二十面体格子) を用いた計算で得られるデータを緯度経度座標系に変換する際、ポスト処理ツール fio_ico2ll を用いる。データ出力は NetCDF 形式で行うが、zlib のコンパイルで FASTEST マクロを適用した場合と適用していない場合についてそれぞれ NetCDF の環境を作成し、経過時間や圧縮率を比較した。

FASTEST マクロを適用した場合、テストケース (圧縮前 200MB のデータを使用) で経過時間が 14.0 秒から 5.9 秒に時間短縮した。どちらのケースも圧縮後ファイル互換性が損なわれていないことが確認できたが、圧縮率は適用していない場合と比べ、適用した場合はわずかに (約 2%) 低下した。また、課題側で測定した結果によると、FASTEST マクロを適用していない場合に比べて、経過時間の削減は約 3 割程度であった。尚、この測定とテストケースでは、変数の種類が異なる。

このように、データの種類やデータ容量によって、圧縮率が低下する可能性がある。FASTEST マクロを適用するか、場合によって検討が必要である。

2.3.4 課題 #2 への技術サポートのまとめ

NICAM は地球シミュレータにおいて十分動作実績があるため、移植や計算性能最適化といった作業を行う準備期間を設けることなく、すぐにプロダクトランを実行することができた。ジョブの実行状況によってはノードを占有する予定であったが、最大ジョブ投入数の制限値を緩和したことにもより、実施期間中、効率よくジョブが実行されたため、ノード占有は行っていない。NetCDF に関する調査については、NetCDF 形式は広く用いられており、本課題のみならず他課題にも今後参考になると考えられる。

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 (Apr. 2016 – Jul. 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 Apr. 2016 to Jul. 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 Apr. 2016 to Jul. 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 aimed to compute the cloud-resolving downscaling of northward migrating typhoons in warming climates of the near future and late twenty-first century using the program CReSS.

This project was a continuation project from the previous period; therefore, The CReSS had been ported to the Earth Simulator and optimized. The project was carried out in three stages: job execution, post-processing, and data-storage.

Before we describe the three stages on which the project was based, we first outline the construction of the Earth Simulator. The Earth Simulator consists of 5,120 nodes that are divided into one extended cluster and six standard clusters (One extended cluster has 2,048 nodes, and One standard cluster

has 512 nodes). In addition, it has two-type queues: The S batch queue and the L batch queue. In one standard cluster, 16 interactive nodes are prepared for debugging and 48 nodes of the S batch nodes are prepared for small-sized execution such as pre- or post-processing. The other 5,056 nodes are L batch nodes that perform medium-sized and large-sized calculations. Furthermore, standard clusters can execute jobs using 1 to 512 nodes the extended cluster can execute jobs using 256 to 2,048 nodes. However, jobs requiring the use of maximum available nodes often need long job-allocation times before execution. The long job-allocation time affects the progress of a research project, especially in the case where the project requires many job executions. Thus, a project using ensemble execution in a short period needs to consider the job-allocation time for effective project implementation.

We now describe the job-execution stage. Based on the usage of the system and the job schedule, we considered the optimal number of nodes to use in one typhoon calculation and proposed the use of a 256-node batch job that can efficiently execute multiple calculations in parallel.

Next, with regard to the post-processing stage, CReSS outputs large amounts of data and therefore, it took about 30 h to calculate them. Basically, the S batch jobs are operated with a maximum limit of 8 h, and hence, we shifted to a 48-h executable setting to perform this project.

Table 1 Each Project and Support Staff

No.	Project Name	Name of Support Staff
#1	Cloud-resolving Downscaling Simulations of Northward Migrating Typhoons in Warming Climates of The Near Future And Late Twenty-first Century	Yuichi Saito
#2	Is Near-future Projection of Tropical Cyclone Feasible? High-resolution Large-ensemble Experiments Using Global Non-hydrostatic Atmospheric Model	Mikiko Ikeda

Finally, in the data-storage stage, we provided a massive storage space for this project to store large amounts of data. The following disk space was prepared for this project.

1. DATA area: 110 TB (for main calculation and verification calculation)
2. MSS area: 200 TB (to save the calculated data)

The disk space prepared for the DATA area was larger than usual and the space for the MSS area was specially provided. These were special measures that were adopted to give priority to calculating the data within the period.

In Project 2, to indicate the feasibility of typhoon prediction in the near future, the researchers have carried out a lot of ensemble experiments for the period during and after “hiatus” (a global warming stagnation period) using the global cloud-resolving model “NICAM” (Non-hydrostatic Icosahedral Atmospheric Model). In that project, we provided technical support to run a lot of ensemble jobs efficiently.

Fifty ensemble jobs were performed with a horizontal resolution of 14-km using 40 nodes per job and the resources allocated for calculation were 510,000 node-hours.

We relaxed the limits on the number of concurrent job executions on the Earth Simulator, as shown in the following table (Table 2).

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

By relaxing the limit on the number of concurrent job executions, it was possible to efficiently execute the jobs.

In addition to relax the limits for the ensemble job execution, we also investigated the performance of Network Common Data Form (NetCDF), which is the output data format of these experiments. We compared the elapsed time and the compression rate in the cases where FASTEST macro was applied or not applied while compiling zlib. In fact, we were able to reduce the elapsed time by approximately 30 % when the FASTEST macro was applied. In both cases, we were able to confirm that data was not destroyed by compressing, and the compression rate was reduced only by a few percent. Based on the results of the investigation, we concluded that not only was the NetCDF format helpful for this job but it could also be helpful for others as well.

By providing above supports 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.

