

HPC 環境における動的変精度演算に対する精度回復手法の検討

Investigation on Accuracy Recovery for Dynamic Transprecision Computing on HPC Systems

和田康孝
明星大学情報学部

1. 研究目的

近年では、従来の限界を超えてシステムの実行性能を向上させるために、演算結果にある程度の誤りが含まれることを許容し、その誤差と引き換えに性能を向上させる Approximate Computing (以下 AC)、アプリケーションの特性に応じて演算精度を調整する Transprecision Computing (以下 TC) といった手法が注目を集めている。この AC あるいは TC の考え方は有効である一方、HPC アプリケーションにおいては、収束計算の反復回数に悪影響を与えるなど、適用には細心の注意が必要となる。

そのため、演算精度と実行性能トレードオフを事前に調整する、実行時に動的に演算精度を変更してアプリケーションの構造に応じた対応をとる、といった最適化が必要となる。一方、事前の静的な調整では、入力データ等アプリケーション実行時の動的な要素を考慮することは難しく、事前に想定した動的な演算精度変更を行う場合であっても、一度下げた演算精度を再び上げた場合に、演算結果の精度がそれに伴って向上するかどうかは未知数である。

以上のような背景から、本プロジェクトでは、TC により動的に演算精度を変更する際、一度下げた演算精度を再び上げた場合に、いち早く演算結果の精度を取り戻す方法について、評価・検討を行う。そのための最初のステップとして、単純に演算精度を上下した際に、演算結果がどのようにそれに追従するかを評価する。

2. 研究成果の内容

本研究課題では、上述の通り、HPC アプリケーションにおいて演算精度が動的に変更された際に、どのように演算結果への影響が変化していくのかを評価することを目的としている。HPC アプリケーションにおける処理の多くはノード間の通信と各ノードにおける演算に大別することができる。

本研究課題では、まず、通信時にデータ型変更を行った際の、データ量やデータ型変換のタイミングに応じた通信性能の変化を評価し、その変化をモデル化した[1-2]。動的に演算精度を変化させる際には、そのために必要となる変換オーバーヘッドを考慮した上で、演算結果への影響を考慮する必要があるが、特に通信に関して、その基礎的な知見を得ることができた。

さらに、演算処理における特性を理解するため、HPCG ベンチマークなどのベンチマ

ークアプリケーションを用いた性能評価を実施した。動的な演算精度調整・最適化は様々なパターンで適用することが想定されるため、今後さらに多様な特性を持ったアプリケーションを対象として、ある程度網羅的に評価データを収集・分析する必要がある。引き続き、演算精度の変更に対する性能と演算結果の変化について評価を進め、モデル化を行うなど発展的な内容につなげていきたい。

3. 学際共同利用プログラムが果たした役割と意義

本学際共同利用プログラムにおける取り組みを通じ、Cygnus および Wisteria という異なる特性・構成を持った複数のシステム上で評価作業を行うことができた。また、本プログラムにおける成果・連携をもとに、さらに内容を発展させた研究提案が科研費研究課題（科研費基盤 C「多要素協調型 Approximate Computing 実現に向けた HPC アプリケーション解析手法」）として採択され、より高度な研究開発・共同研究へとつなげることができた。計算機環境の利活用のみならず、研究者間の連携、およびより発展的な研究課題への基盤形成、といった様々な面で本学際共同利用プログラムの果たした役割は大きい。

4. 今後の展望

HPC システムに求められる演算性能とそれを制約する消費電力その他の物理的制約はトレードオフの関係にある。将来にわたり引き続き HPC システムの性能・効率を向上させるためには、このトレードオフを最適化する AC は有効な手法の一つと考えられる。本プログラムで取り組んだ知見は、新たに採択となった科研費研究課題を中心とした研究開発において活用・発展させることが可能である。特に、コンパイラによるアプリケーション解析技術、FPGA や GPU によるアクセラレーションなどと連携させることで、HPC システム・HPC アプリケーションにおいて、より効果的に AC を適用する手法、およびその自動化に向けた検討を進めたい。

5. 成果発表

(1) 学術論文

- [1] Yoshiyuki Morie, Yasutaka Wada, Ryohei Kobayashi and Ryuichi Sakamoto, "Data Transfer API and its Performance Model for Rank-Level Approximate Computing on HPC Systems", International Journal of Networking and Computing, Vol. 13, No. 1, pp. 48-61, Jan., 2023.

(2) 学会発表

- [2] Yoshiyuki Morie, Yasutaka Wada, Ryohei Kobayashi, and Ryuichi Sakamoto, "Performance Evaluation of Data Transfer API for Rank Level Approximate Computing on HPC Systems", 2022 IEEE International

Parallel and Distributed Processing Symposium Workshops (IPDPSW),
pp. 445-448, May., 2022

- [3] Yasutaka Wada, Yoshiyuki Morie, Ryohei Kobayashi, Ryuichi Sakamoto,
"Toward Dynamic Approximate Computing for HPC Applications", The
International Conference for High Performance Computing, Networking,
Storage, and Analysis (SC22), Exhibition, Nov., 2022.

(3) その他

なし

| 使用計算機 | 使用計算機に ○ | 配分リソース※ | |
|------------------|-------------|---------|------|
| | | 当初配分 | 追加配分 |
| Cygnus | ○ | 3,150 | 0 |
| Wisteria/BDEC-01 | ○ | 31,500 | 0 |

※配分リソースについてはノード時間積をご記入ください。