

演算精度の動的変更によるノード間ロードバランシング

Load Balancing among Compute Nodes with Dynamic Precision Adjustment

和田 康孝
明星大学情報学部

1. 研究目的

今日の HPC システムはより多くのハードウェア資源を搭載し、より並列度の高いシステムとなっている。更には GPU や FPGA などのアクセラレータを搭載することで、ヘテロジニアスな構成を持つことも多い。多種のハードウェアを多数搭載することで理論性能は飛躍的に向上するものの、並列ソフトウェア開発の視点からは、負荷バランスを考慮した高効率なソフトウェア開発がより難しくなるという問題がある。

また、半導体プロセス技術の進歩に伴い CPU やメモリの製造時のばらつきがより顕著となっており、これがプロセス間・ノード間の負荷バランス最適化をより難しくしている。特に、システムの負荷状況によって CPU の動作周波数を動的に調整するシステムが一般的になっているため、事前に負荷バランスを調整することも困難である。

他方、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という取り組みも行われている。この Approximate Computing (以下、AC) 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。

以上のような背景から、本プロジェクトでは、ノード間の負荷バランスを最適化することを目的として、ノードあるいはランクごとに演算精度を個別かつ動的に調整する仕組みの実現を目指す。具体的には、通信や同期のタイミングごとに待ち時間を観測し、その待ち時間の長短に応じて以降の演算における演算精度を選択する手法を提案・実装する。特に、HPC アプリケーションの主要部分・カーネルをマルチバージョン化し、異なるデータ型による実装の中から、状況に応じてノード・ランクごとに適切なものを選択・実行する手法の実現を目指す。

2. 研究成果の内容

本研究課題は、上述の通り、HPC アプリケーションに対してノード・ランクごとに AC を適用することを目指すものである。これを実現するためには、まず、アプリケーションの構造・特性によって、演算精度と実行性能がどのような関係にあるかを把握する必要がある。また、異なる演算精度で処理を行なっている各ノード・ランク間で適切にデータ転送を実施すること、つまりデータ転送時に送信元・送信先データ型が

異なる場合に、齟齬がないよう変換を行いつつデータを授受することが必要となる。

そのため、本研究課題では、まず、さまざまな特性のアプリケーションを含む Rodinia ベンチマークセットを用い、演算精度、具体的にはプログラム中で用いるデータ型を変更してどのような性能の変化があるかを評価した。定性的には、演算精度を下げることで、キャッシュメモリや演算機の利用効率が上がり、最終的に得られる結果の精度が下がることを許容すれば、アプリケーションの実行時間を短縮することが期待できる。しかし、今回調査を行った結果、アプリケーションの特性によっては、演算精度を下げることで、逆にアプリケーションの実行時間が延びる場合があることが確認された[1]。

次に、MPI アプリケーションにおいて、ランクごとに異なる演算精度にて AC を適用するランクレベル AC を実現するために、データ型を変更しつつデータ転送を行う API の提案と実装をおこなった。提案した API を用い、データ型の変換を含むデータ転送オーバーヘッドの評価を行った結果、低オーバーヘッドでの転送が可能であることをしめした。また、転送データサイズ等に基づいた通信コストのモデル化も行うことができ、今後、演算とデータ転送のオーバーラップ等に向けた検討も可能となった[2]。

以上の成果をベースに、引き続きランクレベル AC を効果的に適用する手法について検討を実施している。

3. 学際共同利用が果たした役割と意義

本研究課題の実施においては、プロセッサ構成、インタコネクト、コンパイラとうさまざまな要素において、可能な限り幅広く評価実験を行う必要があった。今回の学際共同利用を通じて、Cygnus と Oakforest-PACS という複数の異なる特性を持つシステムを活用することができたことから、重要な基礎データを取得するとともに、通信ライブラリの実装・評価までを信頼性高く行うことが可能となった。

4. 今後の展望

本研究課題の実施を通じ、HPC アプリケーションに対してランクレベル AC を適用することで、負荷不均衡やプロセッサの製造時ばらつきなどに起因する実行効率の低下を軽減するための基盤を整えることができた。今後は、コンパイラによるアプリケーション解析技術などと連携させることで、より効果的にランクレベル 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
なし
- (2) 学会発表

- [1] Yoshiyuki Morie, Yasutaka Wada, Ryohei Kobayashi, and Ryuichi Sakamoto, "Performance Evaluation of Data Transfer API for Rank Level Approximate Computing on HPC Systems", 24th Workshop on Advances in Parallel and Distributed Computational Models (in conjunction with 36th IEEE International Parallel and Distributed Processing Symposium), May., 2022 (To Appear)
- [2] 和田 康孝, 小林 諒平, 坂本 龍一, 森江 義之, "演算精度の動的制御による Approximate Computing の実現に向けた予備評価", 情報処理学会研究報告 ハイパフォーマンスコンピューティング (HPC) , Vol. 2021-HPC-181, No. 2, pp.1-6, Sept., 2021.

(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	○	3,645	0
Oakforest-PACS	○	45,000	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。			