

非同期入出力機構を用いた大規模乱流直接数値シミュレーションコードの開発

横川三津夫（神戸大学）

概要

スペクトル法による非圧縮一様等方性乱流直接数値シミュレーション (DNS) は、その精度の高さから乱流の特性を明らかにするために広く用いられている。しかし、これまで行われてきたシミュレーションよりも格子点数の大きい DNS は、膨大な計算時間とメモリ容量が必要となるため、スーパーコンピュータのメモリ容量を超える DNS は実現できない。本課題では、コード内の変数を細分化し、複数の一時ファイルに格納することにより、使用メモリ容量を削減する省メモリ DNS コードを開発した。時間積分において、これらのファイルを順次処理することにより DNS を行う。この DNS コードを用いて、速度など 16 個の変数の分割数を変えながら、それらの値を一時ファイルに保存し、同期型及び非同期型の入出力 (I/O) 機能を用いて初期性能評価を行った。この結果、8 分割時の同期 I/O による実行方法が最も計算時間が短かった。また、非同期型 I/O による計算時間は、同期型 I/O の計算時間よりも大きい結果となった。

1 共同研究に関する情報

1.1 共同研究を実施した拠点名

- 東北大学 サイバーサイエンスセンター

1.2 課題分野

- 大規模計算科学課題分野

1.3 共同研究分野 (HPCI 資源利用課題のみ)

- 超大規模数値計算系応用分野
- 超大規模情報システム関連研究分野

1.4 参加研究者の役割分担

- 横川三津夫（神戸大学大学院システム情報学研究科）：総括及びコード性能評価
- 撫佐昭裕（東北大学サイバーサイエンスセンター）：コード高速化検討
- 石原卓（岡山大学大学院環境生命科学研究

科）：スキーム及び計算結果評価

- 小松一彦（東北大学サイバーサイエンスセンター）：コード性能評価
- 曾我隆（NEC ソリューションイノベーション関西支社）：コード高速化検討の総括
- 山口健太（NEC ソリューションイノベーション関西支社）：コード高速化検討、評価
- 山根悠輝（NEC ソリューションイノベーション関西支社）：コード高速化検討、実装、評価
- 松本泰生（神戸大学大学院システム情報学研究科）：コード高速化性能評価
- 畑中裕翔（神戸大学大学院システム情報学研究科）：コード高速化性能評価

2 研究の目的と意義

乱流は、自然や科学技術のさまざまな問題に現れる現象であり、簡単な流れを除いては未だに解析的な解が得られない未解決な問題として知られている。したがって数値シミュレーションが唯一の解決手段と考えられている。

非圧縮性乱流 DNS コードは、基礎方程式であるナビエ・ストークス方程式をフーリエ・スペクトル法により離散化し、フーリエ係数に関する常微分方程式を数値積分するものである。計算時間のほとんどが、非線形項を求めるための3次元高速フーリエ変換 (Fast Fourier Transform, 3D-FFT) に費やされており、この部分の高速化が必須である。また、問題規模を大きくすると必要なメモリ量が増加するため、使用するスーパーコンピュータのメモリ量を超えた大規模な計算格子を扱えない問題がある。近年のスーパーコンピュータは、各計算ノードにファイルアクセス速度の大きい Solid State Drive (SSD) を使用したローカルファイルシステムを持つものが多く見られる。DNS コードの中の変数の一時退避場所として、これを利用することにより実行時に必要なメモリ容量を削減することが考えられる。ファイル入出力時間が追加され全体の計算時間が増加するが、大規模計算を可能にする方法である。

本課題では、一様等方性乱流 DNS コードに対し、計算ノード直結のファイルシステムに変数の退避場所としての一時ファイルを作成し、実行時の使用メモリ容量を削減する実装を行い、その計算性能及びファイル入出力性能を評価することを目的とした。非圧縮性乱流 DNS においては、大きな格子点数による DNS データを得ることにより乱流研究に貴重なデータを示すことができる。

3 当拠点公募型研究として実施した意義

東北大学サイバーサイエンスセンターは、現在ベクトル型プロセッサのスーパーコンピュータ SX-Aurora TSUBASA (AOBA-A) を運用している。SX-Aurora TSUBASA は、x86 アーキテクチャのプロセッサをホスト計算機 (VH) とし、PCI-e カードとして実装された複数のベクトルエンジン (VE) から構成されている。VE 部にはベクトル計算ユニットなどの他にダイレクトメモリアクセス (DMA) エンジンが実装されており、この DMA エンジン部を活用した高速 I/O 機能を提供している。また、計算と I/O を同時に実行できる非同期 I/O 機能を持ち、本課題でファイルへの一時退避した変数の I/O に非同期型 I/O を用いることができた。

4 前年度までに得られた研究成果の概要

該当なし

5 今年度の研究成果の詳細

オリジナルの非圧縮一様等方性乱流 DNS コードは、倍精度変数を用いた Fortran コードであり、また y 軸、z 軸の2軸で領域分割された並列化コードである。このコードに対し、ファイル I/O を用いて使用メモリ量を節約するコードを実装した。

メモリ節約型 DNS コードで扱う変数の分割イメージを図1に示す。図1の左の立方体は、非圧縮性乱流 DNS コードで使用する変数がプロセス分割によって領域分割されていることを表している。また、右上の立方体は1つのプロセス中の変数がさらに分割されることを表す。ここで分割された変数を一時ファイルに退

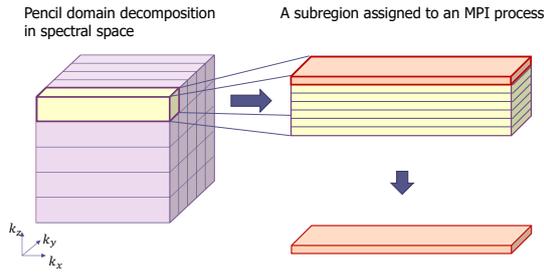


図1 各プロセスにおける変数の分割イメージ

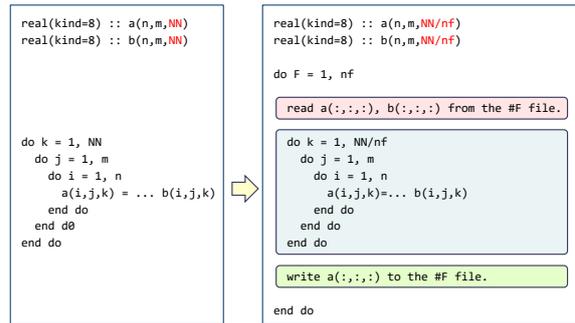


図2 メモリ削減コードの実装イメージ

避させ、必要な時に順に入出力処理を行い、計算に利用する。したがって、1つの変数に対してメモリ上で確保する記憶領域は、ファイル数を N_f とすると、既存コードの使用メモリ容量の $\frac{1}{N_f}$ となる。ただし、3D-FFT に用いる変数については、その実行時に all_to_all 通信が頻繁に現れ、ファイルアクセス回数が多くなるため、ファイル退避しないこととした。

この方法の実装イメージを図2に示す。3次元配列の3次元目を $\frac{1}{N_f}$ し、分割された各部分を別々のファイルに書き込む。実行時には、各ファイルを順番に読み込み、その部分の計算を行い、結果を同じファイルに上書きする。この方法では、入出力と計算を順番に行っており、同期型 I/O により実行する。

非同期 I/O については、入出力時間より計算時間が大きい場合に、入出力時間が計算時間に隠蔽されることを確認しており、本コードについても非同期 I/O 機能を用いたコードも実装した。実装のイメージを図3に示す。配列変数を、4次元目にサイズ3を持つ4次元配列に拡張し、それぞれ読み込む領域、計算する領域、書き込む領域とし、ラウンドロビンにより計算を進める方法である。ループの最初と最後に工夫が必要であるが、これにより、ピンク色の部分と青色の部分がオーバーラップして実行できると考えた。

オリジナルコード、同期 I/O コード、及び

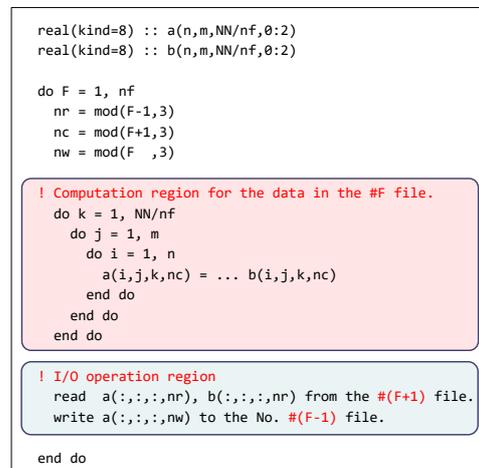


図3 非同期 I/O の実装イメージ

非同期 I/O コードに対し、東北大学サイバーサイエンスセンターのスーパーコンピュータ AOBA-A を用いて、DNS の1時間ステップの計算時間を計測した。計算格子数は 512^3 、並列数を 8×8 とし、各変数の一時ファイル数を 2, 4, 8, 16, 32 と変えながら計算性能を比較した。64 並列プロセスは、64 個の VE に割り当て、各 VE ではスレッド並列計算を行った。また、VH 直結の SSD で構成されたファイルシステム、HDD で構成された共用の並列ファイルシステムとの比較も行った。ファイルに退避した変数は、速度など 23 変数である。図4に、各ケースの実行時間を示す。

オリジナルコードの計算時間は、約 1.8 秒

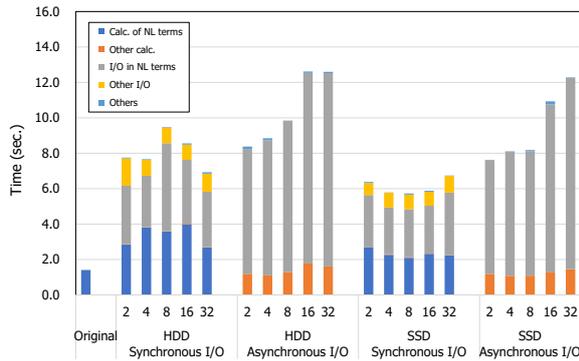


図 4 計算時間

程度であるのに対し、メモリ節約型コードは、I/O 時間が必要なため、約 4 倍から約 7 倍の計算時間が必要だった。I/O 時間が計算時間よりもかなり大きかったためである。非同期 I/O については、ファイル分割数の増加とともに計算時間が増加した。非同期 I/O では、VE 側の入出力命令に対し、VH 側に対応するスレッドを生成すると聞いており、このスレッド数は回避変数の数とファイル分割数の積となるため、スレッド数増加によるオーバーヘッドが原因と考えているが、今後さらに調査したい。HDD と SSD の比較では、やはり計算ノードにローカルに接続された SSD の方が計算時間が短くなる結果となった。

一方、並列化圧縮性乱流 DNS コードについては、新規にスクラッチから開発しており、各部分の動作確認を行っている。コード実装では、移流項に対し 8 次精度コンパクト差分、拡散項に対し 8 次精度中心差分を適用するため、並列領域の袖領域のデータ交換の実装を行った。また、低波数成分への外力挿入部分について、離散フーリエ変換 (Discrete Fourier Transform, DFT) への変更を行い、全体の計算を確認した。非圧縮性 DNS コードへのメモリ削減実装により非同期 I/O による計算時間短縮は現状では見込めないことから、当面は

コード開発と動作検証を行うこととした。

6 今年度の進捗状況と今後の展望

計算ノードに直結した第一階層のファイルシステム上のファイルに変数退避によりメモリ削減する DNS コードの計算時間の評価を行った。同期型 I/O と非同期型 I/O の計算時間を比較すると、非同期型 I/O の性能が十分に発揮されておらず、今後さらに実装の検討が必要と考えている。本課題では、計算時間が増加することを許容して、スーパーコンピュータが持つメモリ容量を超える大規模 DNS を実行することが目標である。今後、メモリ節約型 DNS コードをスーパーコンピュータ富岳への実装を行い、本コードの性能評価と大規模な DNS を目標としたい。

7 研究業績一覧 (発表予定も含む)

学術論文 (査読あり)

国際会議プロシーディングス (査読あり)

- M. Yokokawa, Y. Yamane, K. Yamaguchi, T. Soga, T. Matsumoto, A. Musa, K. Komatsu, T. Ishihara, and H. Kobayashi, "I/O Performance Evaluation of a Memory-Saving DNS Code on SX-Aurora TSUB-ASA," Proc. of 2023 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), pp. 692-696, (2023). <https://doi.org/10.1109/IPDPSW59300.2023.00117>

国際会議発表 (査読なし)

国内会議発表 (査読なし)

- 畑中裕翔, 松本泰生, 山根悠輝, 山口健太, 曾我隆, 撫佐昭裕, 小松一彦, 今村俊幸, 石

原卓，横川三津夫，“メモリ節約型非圧縮
性乱流直接数値シミュレーションコードの
挙動解析，” Vol. 2023-HPC-188, No. 5,
pp. 1-8 (2023)

公開したライブラリ等

その他（特許，プレス発表，著書等）