

令和 6 年 5 月 15 日現在

機関番号：12102

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K11989

研究課題名（和文）エクサスケールシステムにおける高速フーリエ変換のアルゴリズムに関する研究

研究課題名（英文）Research on algorithm of fast Fourier transform in exascale system

研究代表者

高橋 大介（Takahashi, Daisuke）

筑波大学・計算科学研究センター・教授

研究者番号：00292714

交付決定額（研究期間全体）：（直接経費） 2,700,000円

研究成果の概要（和文）：単体のメニーコアプロセッサにおいてSIMD命令を有効活用する高速フーリエ変換（Fast Fourier Transform、以下FFT）アルゴリズムの実装を行った。また、二次元分割を用いた並列三次元FFTにおいて演算と通信をオーバーラップした実装と評価をメニーコア超並列クラスタ上で行った。さらに、複素数上の離散フーリエ変換（Discrete Fourier Transform）を環や体上に一般化した数論変換（Number-Theoretic Transform）の並列化および性能評価を行った。

研究成果の学術的意義や社会的意義

エクサスケールシステムにおけるFFTアルゴリズムを実現するとともに、メニーコアプロセッサやGPUを搭載した超並列クラスタにおいて性能評価を行うことにより、エクサスケールシステムに適したアルゴリズム及び最適化手法を見出すことができた。今後エクサスケール計算環境でFFTを用いた科学技術計算が行われる際に、計算時間を短縮することができるものと期待される。

研究成果の概要（英文）：We implemented the Fast Fourier Transform (FFT) algorithm, which effectively uses SIMD instructions on a single many-core processor. We also implemented and evaluated a parallel three-dimensional FFT with overlapping computations and communication using two-dimensional decomposition on a massively parallel cluster of many cores. Furthermore, we parallelized and evaluated the performance of the Number-Theoretic Transform, which generalizes the Discrete Fourier Transform over complex numbers to a ring and a finite field.

研究分野：高性能計算

キーワード：エクサスケールシステム 高速フーリエ変換 数論変換 メニーコアプロセッサ SIMD化 並列化 GPU 最適化

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

高速フーリエ変換 (fast Fourier transform、以下 FFT) は、科学技術計算において今日広く用いられているアルゴリズムである。FFT において大量のデータを高速に処理するために、並列 FFT アルゴリズムが様々な研究者によって提案されており、ライブラリとなっているものも多い。研究開始当初の 2019 年 4 月の時点では、100 ペタフロップスを超える理論ピーク演算性能を持つスーパーコンピュータは世界で 4 システムが存在していたが、当時の技術動向から 2021 ~ 2022 年頃にはエクサフロップスを超える性能を持つ次世代のスーパーコンピュータが出現すると予想されていた。エクサスケールシステムを実現するにあたって最大の壁となるのは消費電力であると言われていた。この制約を克服するために、メニーコアプロセッサや GPU を搭載した計算ノードを数万台から数十万台接続した超並列クラスタがエクサスケールシステムの一つとして検討されていた。このようなエクサスケールシステムにおいては、プロセッサ内の並列性、計算ノード内の並列性、そして計算ノード間の並列性という、複数階層にまたがった並列性を引き出すようにアルゴリズムを設計する必要がある。さらに、エクサスケールシステムで高い性能を発揮させるためには、それぞれの階層におけるスケラビリティを確保するとともに、メモリアクセスやノード間通信に要する時間をできるだけ少なくする必要がある。

しかし、研究開始当初までに提案されてきた並列 FFT アルゴリズムは、マルチコアプロセッサからなる計算ノードを数万ノード程度接続したシステムまでしか対応できておらず、そのままエクサスケールシステムに用いたのでは、高い性能が発揮できないことが予想された。並列 FFT アルゴリズムにおいては、行列の転置を全対全通信により行っているが、この全対全通信が実行時間に占める割合は非常に大きく、場合によっては実行時間の大部分を占めることもある。全対全通信のメッセージサイズは MPI プロセス数に反比例するため、各 MPI プロセス当たり数 GB のデータがあったとしても、MPI プロセス数が数万のオーダーになると、メッセージサイズは数十 KB 程度になるため、レイテンシが無視できなくなるという問題点があった。研究開始当初の 2019 年 4 月の時点で世界最高速のスーパーコンピュータであった米国オークリッジ国立研究所 (ORNL) の Summit や、2 番目に高速であったスーパーコンピュータである中国の国立スーパーコンピュータセンターの Sunway TaihuLight では、ノード当たりの演算性能に対して通信バンド幅が相対的に低くなっていた。この傾向はエクサスケールシステムにおいても続くことが予想されたため、ノード内の演算性能だけでなく全対全通信の性能を改善し、計算と通信をオーバーラップさせることが必要であった。

2. 研究の目的

本研究の目的は、エクサスケールシステムにおける FFT アルゴリズムを実現するとともに、メニーコアプロセッサや GPU を搭載した超並列クラスタにおいて性能評価を行うことにより、エクサスケールシステムに適したアルゴリズム及び最適化手法を見出すことである。

具体的な研究項目は、(1)メニーコアプロセッサや GPU の性能を最大限に引き出す並列 FFT の実装、(2)超並列クラスタにおいて並列 FFT を実行する上でボトルネックとなる全対全通信の性能の改善、(3)計算と通信のオーバーラップの最適化、の 3 つである。

3. 研究の方法

近年のメニーコアプロセッサにおいては、Intel Advanced Vector Extensions 512 (AVX-512) 命令や Arm Scalable Vector Extension (SVE) 命令などの SIMD 命令を用いて演算性能を向上させていることから、これらの SIMD 命令の有効活用を図る FFT アルゴリズムの実装を行う。また、次世代の GPU も視野に入れ、性能を最大限に引き出せるような FFT アルゴリズムの実装も行う。2 段階全対全通信アルゴリズムを用いた場合でも、ノード数やデータサイズによって 1 プロセッサ当たりの最適な MPI プロセス数が異なると考えられるが、まだ十分な知見が得られていないためこれを明らかにする。また、並列 FFT においては計算と通信をオーバーラップさせることが有効であるが、二次元分割を用いた並列 FFT ライブラリにおいて計算と通信をオーバーラップした例はまだ知られていないことから、実装および評価を行う。1 プロセッサ当たりの最適な MPI プロセス数や、計算と通信の最適なオーバーラップの割合を静的に求めることは困難であるため、自動チューニング手法を活用する。さらに、エクサスケールシステムの性能パラメータを想定して、開発した並列 FFT の性能予測も行う。

研究スケジュールとしては、まずは単体のメニーコアプロセッサや GPU における並列 FFT アルゴリズムの検討ならびに性能評価を行い、その後複数ノードにおける並列 FFT アルゴリズムの検討ならびに性能評価を行う。

4. 研究成果

単体のメニーコアプロセッサにおいて SIMD 命令を有効活用する FFT アルゴリズムの実装を行った。さらに、Intel Xeon Phi クラスタにおける二次元分割を用いた並列三次元実数 FFT の実現と評価を行った。最近のスーパーコンピュータでは、MPI プロセス数が 1 万個を超える場

合もあるが、z 軸で次元分割を行った場合、z 軸におけるデータ点数も 1 万点を超えることになり、三次元 FFT の問題サイズに制約が生じることになる。この問題に対処する方法として、二次元分割を用いた並列三次元 FFT が提案されている。しかし、Intel Xeon Phi クラスタにおける二次元分割を用いた並列三次元実数 FFT の実装はまだ報告されていなかった。そこで Intel Xeon Phi クラスタにおいて二次元分割を用いた並列三次元実数 FFT を実装して評価を行った。二次元分割を用いた並列三次元 FFT の提案した実装は、離散フーリエ変換 (Discrete Fourier Transform、以下 DFT) の共役対称性および row-column FFT アルゴリズムに基づいている。FFT カーネルは Intel AVX-512 命令を用いてベクトル化を行った。性能評価の結果、二次元分割が多数の MPI プロセスに対して通信時間を短縮することによって性能を効果的に改善することを示した。さらに、Intel Xeon Phi クラスタにおいて既存の並列 FFT ライブラリ (FFTW、P3DFFT) よりも高速に並列三次元実数 FFT を計算できることを確認した。

二次元分割を用いた並列三次元 FFT において計算と通信をオーバーラップした実装と評価を Intel Xeon Phi クラスタ上で行った。計算と通信のオーバーラップには、MPI 非同期通信が広く使われているが、OpenMP による通信スレッドを導入した計算と通信のオーバーラップ手法が提案されているため、これを用いた。1 プロセッサ当たりの最適な MPI プロセス数や、計算と通信の最適なオーバーラップの割合を静的に求めることは困難であるため、自動チューニング手法を活用した。MPI プロセスグリッド、計算と通信のオーバーラップ、キャッシュブロッキングサイズの最適なパラメータを選択する自動チューニング機能を実装した。性能評価の結果、実装した二次元分割を用いた並列三次元 FFT の自動チューニングは性能向上に有効であることを示した。さらに、Intel Xeon Phi クラスタにおいて既存の並列 FFT ライブラリである FFTW よりも高速に並列三次元 FFT を計算できることを確認した。

また、倍精度浮動小数点 FFT だけでなく複素数上の DFT を環や体上に一般化した数論変換 (Number-Theoretic Transform、以下 NTT) についても実装の検討を行った。NTT は準同型暗号、多項式乗算、多倍長精度数乗算などに用いられている。NTT のバタフライ演算には、剰余加算、剰余減算、および剰余乗算が含まれている。剰余加算は加算と条件付き減算に置き換えることができる。剰余乗算は Montgomery 乗算や Shoup 乗算を用いることで時間の掛かる除算を実質的に行うことなく、乗算、加減算およびシフト演算のみで剰余乗算を行えることが知られている。単体プロセッサにおいて、NTT で用いられる複数の Montgomery 乗算を SIMD 命令である Intel AVX-512IFMA (Integer Fused Multiply-Add) 命令を用いて高速化し、性能評価を行った。52 ビット以下のオペランドに対して、Intel AVX-512IFMA 命令を用いた実装は、Intel 64 および Intel AVX-512F (Foundation) 命令を用いた実装に比べて高速に複数の剰余乗算が行えることを示した。

さらに、FFT の並列化に適している six-step FFT アルゴリズムを NTT に適用し、six-step NTT アルゴリズムを構築するとともに、OpenMP を用いて並列化を行った。メニーコアプロセッサにおける性能評価の結果、提案する NTT の実装は並列化が行われていない Intel Homomorphic Encryption (HE) Acceleration Library に含まれている NTT の実装よりも高速であることを示した。

また、GPU クラスタにおける NTT の並列化および性能評価を行った。FFT の並列化に適している four-step FFT アルゴリズムを NTT に適用し、four-step NTT アルゴリズムを構築するとともに、GPU クラスタにおいて MPI と OpenACC を用いて並列化を行った。NVIDIA H100 GPU を搭載した GPU クラスタの 32 ノードにおける性能評価の結果、提案する NTT の実装は 462 Gops (giga-operations per second) の性能を示した。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Daisuke Takahashi	4. 巻 13366
2. 論文標題 An Implementation of Parallel Number-Theoretic Transform Using Intel AVX-512 Instructions	5. 発行年 2022年
3. 雑誌名 Proc. 24th International Workshop on Computer Algebra in Scientific Computing (CASC 2022), Lecture Notes in Computer Science	6. 最初と最後の頁 318-332
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-031-14788-3_18	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Daisuke Takahashi	4. 巻 12253
2. 論文標題 Fast Multiple Montgomery Multiplications Using Intel AVX-512IFMA Instructions	5. 発行年 2020年
3. 雑誌名 Proc. 20th International Conference on Computational Science and Its Applications (ICCSA 2020), Part V, Lecture Notes in Computer Science	6. 最初と最後の頁 655-663
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-030-58814-4_52	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Daisuke Takahashi	4. 巻 12043
2. 論文標題 Implementation of Parallel 3-D Real FFT with 2-D Decomposition on Intel Xeon Phi Clusters	5. 発行年 2020年
3. 雑誌名 Proc. 13th International Conference on Parallel Processing and Applied Mathematics (PPAM 2019), Part I, Lecture Notes in Computer Science	6. 最初と最後の頁 151-161
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-030-43229-4_14	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計7件（うち招待講演 0件／うち国際学会 3件）

1. 発表者名 高橋大介
2. 発表標題 Intel AVX-512IFMA命令を用いた並列数論変換の実現と評価
3. 学会等名 日本応用数理学会2022年度年会
4. 発表年 2022年

1. 発表者名 Daisuke Takahashi
2. 発表標題 Parallel Implementation of FFT in a Finite Field
3. 学会等名 SIAM Conference on Parallel Processing for Scientific Computing (PP22) (国際学会)
4. 発表年 2022年

1. 発表者名 高橋大介
2. 発表標題 二次元分割を用いた並列三次元FFTにおける計算と通信のオーバーラップの自動チューニング
3. 学会等名 日本応用数理学会2021年度年会
4. 発表年 2021年

1. 発表者名 Daisuke Takahashi
2. 発表標題 Automatic Tuning of Computation-Communication Overlap for Parallel 3-D FFT with 2-D Decomposition
3. 学会等名 SIAM Conference on Computational Science and Engineering (CSE21) (国際学会)
4. 発表年 2021年

1. 発表者名 高橋大介
2. 発表標題 Intel AVX-512IFMA命令を用いた複数のMontgomery乗算の高速化
3. 学会等名 日本応用数理学会2020年度年会
4. 発表年 2020年

1. 発表者名 Daisuke Takahashi
2. 発表標題 Implementation of Parallel 3-D Real FFT with 2-D Decomposition on Intel Xeon Phi Clusters
3. 学会等名 SIAM Conference on Parallel Processing for Scientific Computing (PP20) (国際学会)
4. 発表年 2020年

1. 発表者名 高橋大介
2. 発表標題 Xeon Phiクラスタにおける二次元分割を用いた並列三次元実数FFTの実現と評価
3. 学会等名 日本応用数理学会2019年度年会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関