

科学研究費助成事業 研究成果報告書

平成 28 年 6 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：2012～2015

課題番号：24500029

研究課題名(和文) エクサスケール計算環境に向けた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 FFT Algorithms for Exa-Scale Computing Environment

研究代表者

高橋 大介 (TAKAHASHI, Daisuke)

筑波大学・システム情報系・教授

研究者番号：00292714

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：エクサスケール計算環境に向けた並列高速フーリエ変換(FFT)アルゴリズムの検討を行った。並列FFTにおいては全対全通信が性能に大きく影響する。そこで、通信隠蔽や基底、キャッシュブロックサイズなどのパラメータを自動的にチューニングする機構を実装した。性能評価の結果、提案する並列FFTの実装は性能を向上させる上で効果が高いことが示された。

研究成果の概要(英文)：In this research, we investigated an implementation of parallel fast Fourier transform (FFT) algorithm for exa-scale computing environment. Parallel FFTs require intensive all-to-all communication, which affects the performance of parallel FFTs. An automatic tuning facility for selecting the optimal parameters of the communication hiding, the radices, and the block size is implemented. The performance results demonstrate that the implementation of parallel FFTs with automatic tuning is efficient for improving the performance.

研究分野：高性能計算

キーワード：高速フーリエ変換 エクサスケール計算環境

1. 研究開始当初の背景

2011年10月の時点で、ペタフロップスを超える性能を持つスーパーコンピュータは世界で10システムが存在した。その後の技術動向からは2018~2019年頃にはエクサフロップスを超える性能を持つ次々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。エクサスケール計算環境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、GPUやMIC(Many Integrated Core)等のメニーコアアクセラレータと従来のマルチコアCPUが密結合されたヘテロジニアスアーキテクチャの計算ノードを数十万台接続したシステムがエクサスケール計算環境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計算環境においては、CPU内の並列性、アクセラレータ内の並列性、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と共に、主記憶アクセスやノード間通信の時間を極力少なくするようにアルゴリズムを設計する必要がある。しかし、これまでに提案されてきた並列高速フーリエ変換(FFT)アルゴリズムは、ホモジニアスアーキテクチャの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計算環境に用いたのでは、高い性能が発揮できないことが予想されていた。

並列FFTアルゴリズムについては、この10年程研究が活発に行われている。その代表的なものとしては、MITのグループが開発したFFTWが挙げられる。FFTWでは、自動チューニング技術によりFFTの性能を向上させているのが特徴であるが、アクセラレータを搭載した超並列クラスタには対応していない。さらに、マルチコア超並列クラスタにおいて、二次元分割を用いることで通信時間を削減した並列三次元FFTについての研究が挙げられる。本研究のターゲットは、ヘテロジニアスアーキテクチャの計算ノードを数十万台接続したエクサスケール計算環境に向けたFFTアルゴリズムであるが、このような例についてはまだ研究事例が存在していなかった。

2. 研究の目的

本研究の目的は、エクサフロップスを超える性能を持つ次々世代のスーパーコンピュータに向けたFFTアルゴリズムを実現すると共に、アクセラレータ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、エクサスケール計算環境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。

これまでの研究成果を十分に活用し、以下の点について明らかにする。(1)エクサスケール計算環境に向けたFFTアルゴリズムを実現する。(2)アクセラレータを搭載した超並

列クラスタにおいて性能評価を行い、提案する並列FFTアルゴリズムの高速性を実証する。(3)これらの結果より、エクサスケール計算環境に適したアルゴリズム及び最適化手法について明らかにする。

3. 研究の方法

研究の方法としては、まずGPUなどのアクセラレータ単体における並列FFTアルゴリズムの検討を行い、次にアクセラレータとマルチコアCPUのヘテロジニアスアーキテクチャにおける並列FFTアルゴリズムの検討を行う。その後、複数ノードにおける並列FFTアルゴリズムについて検討し、最後にアクセラレータを搭載した超並列クラスタにおける性能評価を行うというように、ボトムアップで研究を進める。

平成24年度は、アクセラレータ単体における並列一次元FFTの検討を行うと共に、ヘテロジニアスアーキテクチャにおける並列FFTの検討を行う。アクセラレータ単体における並列一次元FFTの検討においては、GPUにおいて高い並列性を確保すると共に、メモリアccessが少なくなるようにアルゴリズムを構築する。ヘテロジニアスアーキテクチャにおける並列FFTの検討においては、アクセラレータの性能に加えてマルチコアCPUの性能も発揮できるように負荷を適切に分散するアルゴリズムを構築する。

平成25年度は、ヘテロジニアスアーキテクチャにおける並列FFTの性能評価を行うと共に、複数ノードにおける並列FFTアルゴリズムの検討を行う。ヘテロジニアスアーキテクチャにおける並列FFTの性能評価の結果を、並列FFTアルゴリズムの改良に反映させる。複数ノードにおける並列FFTアルゴリズムの検討においては、ノード数に対するスケラビリティを良くすることが重要になる。したがって、アクセラレータを搭載した超並列クラスタ上においてどのようなデータ分散および通信方法が望ましいかについて検討を行う。

平成26年度はアクセラレータを搭載した超並列クラスタにおける性能評価を行う。性能評価にあたっては、これまでに提案されているFFTWライブラリや既存のベンダー製の並列FFTライブラリと、今回の研究で実現する並列FFTの性能を比較し、その優位性を検証する。

4. 研究成果

(1)エクサスケール計算環境に向けたFFTアルゴリズムとして、並列一次元FFTをマルチコア超並列クラスタで実行する際に、どのようなアルゴリズムが望ましいかについて検討を行った。その結果、メニーコア超並列クラスタにおける並列一次元FFTでは、これまでに提案されてきた三次元表現よりも、二次

元表現を再帰的に適用して計算することで、キャッシュヒット率をさらに向上させることができることが分かった。並列一次元 FFT においては、全体に関わる性能パラメータとして、全対全通信方式と基底、そしてブロックサイズが存在するが、これらについて最適なパラメータを自動的にチューニングする機構をソフトウェアで実現した。その結果、並列一次元 FFT において実行時間の多くを占めている全対全通信に要する時間を大幅に削減することに成功した。

(2) GPU クラスタにおける並列一次元 FFT の実装について検討を行った。GPU クラスタにおいて並列一次元 FFT を行う際には、全対全通信が3回行われることから、計算時間の大部分が全対全通信によって占められることになる。さらに CPU と GPU 間を接続するインターフェースである PCI Express バスの理論ピークバンド幅は PCI Express Gen 2 x 16 レーンの場合には一方向あたり 8GB/sec となっていることから、CPU と GPU 間のデータ転送量を削減することも重要になる。GPU 上のメモリを MPI により転送する場合、基本的には 1. GPU 上のデバイスメモリから CPU 上のホストメモリへデータをコピーする。2. MPI の通信関数を用いて転送する。3. CPU 上のホストメモリから GPU 上のデバイスメモリにコピーする、という手順で行う必要がある。この場合、CPU と GPU のデータ転送を行っている間は MPI の通信が行われないという問題がある。そこで、CPU と GPU 間のデータ転送とノード間の MPI 通信をパイプライン化してオーバーラップすることができる MPI ライブラリである MVAPICH2 を用いることで、この問題を解決した。さらに、FFT の処理において出現する行列の転置の処理を GPU 上で行うなどの工夫も行った。その結果、GPU クラスタにおける並列一次元 FFT が PC クラスタにおける並列一次元 FFT よりも高速に実行できることを確認した。

(3) エクサスケール計算環境に向けた FFT として、Xeon Phi クラスタおよび Fujitsu PRIMEHPC FX100 における並列一次元 FFT の実装について検討を行った。演算と通信を分割しパイプライン方式でオーバーラップさせることで計算時間の短縮を図った。演算と通信を分割しパイプライン方式でオーバーラップさせる場合、通信メッセージサイズの分割数(パイプラインの段数)を大きくすれば、オーバーラップの割合が高くなる。その一方で、通信メッセージサイズの分割数を大きくすれば、通信1回あたりの通信メッセージサイズが小さくなるため、通信バンド幅も小さくなる。したがって通信メッセージサイズの分割数には最適な値が存在すると考えられる。そこで、最適な通信メッセージサイズの分割数を自動チューニングする機構を実装した。Xeon Phi クラスタおよび Fujitsu PRIMEHPC FX100 における性能評価の結果、従来の並列一次元 FFT の実装よりも高速である

ことを確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表](計10件)

Daisuke Takahashi, Automatic Tuning for Parallel FFTs on Intel Xeon Phi Clusters, SIAM Conference on Parallel Processing for Scientific Computing (PP16), 2016年4月14日, Universite Pierre et Marie Curie, Cordeliers Campus (Paris, France).

Daisuke Takahashi, Automatic Tuning for Parallel FFTs on Intel Xeon Phi Clusters, 2016 Conference on Advanced Topics and Auto Tuning in High-Performance and Scientific Computing (2016 ATAT in HPSC), 2016年2月19日, National Taiwan University (Taipei, Taiwan).

高橋 大介, Xeon Phi における並列 FFT の実現と評価, 日本応用数理学会 2015 年度年会, 2015 年 9 月 11 日, 金沢大学角間キャンパス(石川県金沢市).

高橋 大介, Xeon Phi クラスタにおける並列 FFT の自動チューニング, 第 20 回計算工学講演会, 2015 年 6 月 8 日, つくば国際会議場(茨城県つくば市).

Daisuke Takahashi, Automatic Tuning for Parallel FFTs on GPU Clusters, 2015 SIAM Conference on Computational Science and Engineering (CSE15), 2015 年 3 月 18 日, Salt Palace Convention Center (Salt Lake City, Utah, USA).

高橋 大介, GPU クラスタにおける並列 FFT の自動チューニング, 日本応用数理学会 2014 年度年会, 2014 年 9 月 3 日, 政策研究大学院大学(東京都港区).

Daisuke Takahashi, Implementation of Parallel FFTs on GPU Clusters, 2014 Conference on Advanced Topics and Auto Tuning in High Performance Scientific Computing (2014 ATAT in HPSC), 2014 年 3 月 14 日, National Taiwan University (Taipei, Taiwan).

Daisuke Takahashi, Implementation of Parallel 1-D FFT on GPU Clusters, Proc. 2013 IEEE 16th International Conference on Computational Science and Engineering (CSE 2013), pp. 174-180 (2013).

Daisuke Takahashi, Experience of Implementing Parallel FFTs on GPU Clusters, Special Session: Legacy HPC Application Migration 2013 (LHAM) (held in conjunction with IEEE MCSoc-13), 2013 年 9 月 27 日, National

Institute of Informatics(Chiyoda-ku ,
Tokyo , Japan).

Daisuke Takahashi, Atsuya Uno and
Mitsuo Yokokawa, An Implementation of
Parallel 1-D FFT on the K computer,
Prof. 2012 IEEE 14th International
Conference on High Performance
Computing and Communications
(HPCC-2012), pp. 344-350 (2012).

6 . 研究組織

(1)研究代表者

高橋 大介 (TAKAHASHI DAISUKE)
筑波大学・システム情報系・教授
研究者番号 : 00292714