

平成 22 年 5 月 28 日現在

研究種目：若手研究 (B)
 研究期間：2008～2009
 課題番号：20700021
 研究課題名 (和文) メニーコア超並列クラスタ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 Many-Core Massively Parallel Clusters

研究代表者
 高橋 大介 (TAKAHASHI DAISUKE)
 筑波大学・大学院システム情報工学研究科・准教授
 研究者番号：00292714

研究成果の概要 (和文)：並列一次元 FFT アルゴリズムの並列多次元 FFT アルゴリズムへの拡張を行った。さらに、並列多次元 FFT をマルチコア並列クラスタで実行する際に、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った。その結果、メニーコア超並列クラスタにおける並列三次元 FFT では、二次元方向に分割する方が通信時間の削減という観点からは優れていることが分かった。

研究成果の概要 (英文)：In this research, we extended a parallel one-dimensional FFT algorithm to a parallel multi-dimensional FFT algorithm. Furthermore, we investigated an appropriate data distribution and a communication method. For parallel three-dimensional FFTs, a two-dimensional decomposition improves performance effectively by reducing the communication time for many-core massively parallel clusters.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008 年度	1,100,000	330,000	1,430,000
2009 年度	700,000	210,000	910,000
年度			
年度			
年度			
総計	1,800,000	540,000	2,340,000

研究分野：ハイパフォーマンスコンピューティング

科研費の分科・細目：情報学・ソフトウェア

キーワード：並列処理・分散処理

1. 研究開始当初の背景

(1) 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。並列計算機の普及に伴い、並列 FFT アルゴリズムが様々な研究者によって低減されており、ライブラリとなっているものも多い。

近年、CPU のアーキテクチャは消費電力当

たりの性能を向上させるために、マルチコア／メニーコア化が進んできている。

今後、ペタスケールコンピューティング環境としては、メニーコア CPU を数個～数十個搭載した計算ノードを数千～数万台接続するというメニーコア超並列クラスタが主流になることが予想された。

(2) このようなメニーコア超並列クラスタ

においては、CPU 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、メニーコア超並列クラスタで高い性能を発揮させるためには、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必要がある。

しかし、これまで PC クラスタ向けに提案されてきた並列 FFT アルゴリズムを、そのままメニーコア超並列クラスタに用いたのでは、高い性能が発揮できないことが予想された。

(3) 並列 FFT アルゴリズムの研究については、この 10 年程研究が活発に行われている。その代表的なものの一つとして、MIT のグループが開発した FFTW (The Fastest Fourier Transform in the West) 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。

しかし、この自動チューニングはプロセッサ内についてだけ行われており、MPI を用いた大規模な並列実行については考慮されていない。

また、デュアルコア CPU から成る PC クラスタにおける並列三次元 FFT の例としては、デュアルコア CPU で構成された超並列マシンである IBM Blue Gene/L を用いて評価した既存研究がある。

本研究のターゲットは、メニーコア超並列クラスタにおける 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ほとんどないのが現状であった。

(4) これまでに研究代表者の高橋は FFT ライブラリとして FFTE を開発してきており、現在 <http://www.ffte.jp/> でソースコードを公開している。

FFTE ライブラリに関しては、FFTW が多くのプロセッサで高速なライブラリとして知られているが、FFTE は複数のアーキテクチャの計算機において FFTW よりも高速であることを性能評価により客観的に示しており、現在世界で最も高速である FFT ライブラリの一つとして認知されている。

(5) また、Tennessee 大学を中心として Lawrence Berkeley National Laboratory, MIT などのグループが 2004 年から公開している、並列スーパーコンピュータ向けの次世代のベンチマークテストである、HPC Challenge Benchmark (HPCC) の 7 つのベンチマークのうち、FFTE がその一つとして取り入れられている。2007 年現在世界最高速であった並列スーパーコンピュータである IBM Blue Gene/L (131,072 プロセッサ, ピーク性能 360TFlops) において HPCC ベンチマークが実

行されているが、LINPACK ベンチマークの性能がピーク性能の約 72% であるのに対し、FFTE の性能は約 2.3TFlops と、ピーク性能の 10% 以下の性能にとどまっている。

したがって、今後メニーコア超並列クラスタをはじめとするペタスケールコンピューティング環境において FFT を実行する上で、さらなる性能の向上が求められていると言える。

2. 研究の目的

(1) 本研究の目的は、メニーコア超並列クラスタに向けた FFT アルゴリズムを実現すると共に、現在利用可能なマルチコア超並列クラスタにおいて性能評価を行うことにより、メニーコア超並列クラスタ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。

性能評価の結果より、メニーコア超並列クラスタに適した FFT アルゴリズムおよび最適化手法について明らかにする。

(2) これまでに PC クラスタ等における並列 FFT アルゴリズムはほぼ確立されてきたと言えるが、メニーコア超並列クラスタは数万~数十万コアから構成されており、PC クラスタ等におけるアルゴリズムや最適化手法が必ずしも有効とは言えない可能性がある。

また、メニーコア超並列クラスタの性能を十分発揮できるような並列 FFT アルゴリズムについては、まだ確立されているとは言いがたい。さらに、メニーコア超並列クラスタにおけるアルゴリズムや最適化手法は、今後ペタスケールコンピューティングも視野に入れたものである。

(3) FFT は、科学技術計算において非常に多く出現し、FFT が全実行時間のほとんどを占めるような計算も多い。したがって、メニーコア超並列クラスタにおける並列 FFT アルゴリズムは、今後ペタスケールコンピューティング環境が使われる際に、科学技術計算の計算時間を短縮することができるものと期待される。

また本研究から得られた知見は、メニーコア超並列クラスタにおける他の並列数値計算アルゴリズムの最適化手法についても役立てることができると考えられる。

3. 研究の方法

(1) まず、単体コアにおける一次元 FFT の検討を行う。最近のプロセッサでは Intel Pentium4 の SSE2/SSE3 や、Motorola PowerPC の AltiVec など、Short Vector SIMD 命令を搭載しているものが多いことから、これらの命令を活用することについても考慮する。

(2) 次に、単体 CPU (複数コア) における並列一次元 FFT の検討を行う。複数コアで共有されているキャッシュメモリを活用し、主記憶へのアクセス回数が少なくなるようにアルゴリズムを構築することが重要となる。さらに、複数 CPU の場合に拡張することにより、単体ノード (複数 CPU) における並列一次元 FFT の検討を行う。

(3) 複数ノードにおける並列一次元 FFT を実現するには、ノード数に対するスケーラビリティを良くすることが重要となる。さらに、マルチコア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の検討を行う必要がある。

(4) 並列一次元 FFT アルゴリズムの並列多次元 FFT アルゴリズムへの拡張を行う。この並列多次元 FFT への拡張に関しては、本研究で実現した並列一次元 FFT を元に行う。さらに、並列多次元 FFT をマルチコア超並列クラスタ上で実行する際に、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。

(5) 本研究で実現した並列一次元 FFT および並列多次元 FFT をマルチコア超並列クラスタ上に実現し、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、その優位性を検討する。

4. 研究成果

(1) 単体コアにおける一次元 FFT アルゴリズムを検討した。具体的には、キャッシュメモリを活用し、主記憶へのアクセス回数が少なくなるようにアルゴリズムを構築した。

(2) 近年の超並列クラスタにおいては、性能を向上させるためにコア数やプロセッサ数が増える傾向にある。例えば、PFlops を超えるピーク性能を有する超並列システムである Jaguar や Roadrunner においては、コア数が 10 万個を超えている。

このようなシステムにおいては、MPI と OpenMP のハイブリッド実行を行い、MPI プロセス数を少なくすることが通信時間を削減する上で有効であるが、それでも MPI プロセス数は 1 万個以上になる。

並列三次元 FFT において、従来の一次元分割による手法では、x, y, z 各方向のデータ数が 1 万点以上でなければならないことになり、三次元 FFT の問題サイズに制約を受けることになる。

(3) この問題点を改善する方法として、これまでは一次元方向に分割されていた並列三次元 FFT アルゴリズムを、二次元方向に分割するアルゴリズムの検討を行った。この方法を用いることで、比較的少ないデータ数でも高いスケーラビリティを得ることが可能になった。

(4) この二次元分割による並列三次元 FFT アルゴリズムを実現し、マルチコア超並列クラスタである T2K 筑波システムの 256 ノードにおいて性能評価を行った。性能評価にあたっては、提案した二次元分割による並列三次元 FFT と、一次元分割を行った並列三次元 FFT の性能比較を行った。

測定に際しては、 32^3 , 64^3 , 128^3 , 256^3 点の順方向 FFT を連続 10 回実行し、その平均の経過時間を測定した。通信ライブラリとしては、MVAPICH 1.2.0 を用いた。なお、今回の性能評価では 1 コアあたり 1 個の MPI プロセスを用い、flat MPI 実行としている。これは、今回の評価は MPI プロセス数が大きい場合のスケーラビリティに重点を置いているためである。

(5) 性能評価の結果であるが、 32^3 点 FFT では良好なスケーラビリティが得られていない。これは問題サイズが小さい (1MB) ことから、全対全通信が全実行時間のほとんどを占めているためであると考えられる。それに対して、 256^3 点 FFT では 4,096 コアまで性能が向上している。

また、4,096 コアにおける性能は約 401.3GFlops であった。64 コア以下の場合には、通信量の少ない一次元分割が二次元分割よりも性能が高くなっているが、128 コア以上では通信時間を少なくできる二次元分割が一次元分割よりも性能が高くなっている。

(6) なお、一次元分割を用いた場合、 256^3 点 FFT は 256MPI プロセス以下でのみ実行が可能であるが、二次元分割を用いた場合、 256^3 点 FFT は 65,536MPI プロセスまで実行が可能である。

二次元分割を行った 256^3 点並列三次元 FFT における通信時間についてであるが、4,096 コアにおいては 96%以上が通信時間に費やされている。この理由としては、全対全通信において各プロセッサが一度に送る通信量がわずかに 1KB となるため、通信時間においてレイテンシが支配的になるからであると考えられる。

(7) 通信時間をさらに削減するためには、全対全通信に MPI_Alltoall 関数を使わずに、より低レベルな通信関数を用いて、レイテン

シを削減するなどの工夫が必要になると考えられる。これは今後の課題である。

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)

[雑誌論文] (計 1 件)

- ① Daisuke Takahashi, An Implementation of Parallel 3-D FFT with 2-D Decomposition on a Massively Parallel Cluster of Multi-Core Processors, Proc. 8th International Conference on Parallel Processing and Applied Mathematics (PPAM 2009), Lecture Notes in Computer Science, Springer-Verlag, 査読有, 2010, 印刷中.

[学会発表] (計 4 件)

- ① Daisuke Takahashi, Automatic Tuning for Parallel 3-D FFT with 2-D Decomposition, 2010 SIAM Conference on Parallel Processing for Scientific Computing, 2010年2月25日, Grand Hyatt Seattle (Seattle, Washington, USA).
- ② 高橋大介, ペタスケール計算環境に向けた FFT ライブラリ, 日本応用数理学会 2009 年度年会, 2009 年 9 月 28 日, 大阪大学 (豊中市).
- ③ 高橋大介, 次世代スーパーコンピュータに向けた高速フーリエ変換アルゴリズム, 第 58 回理論応用力学講演会, 2009 年 6 月 9 日, 日本学術会議 (東京都港区).
- ④ Daisuke Takahashi, A Volumetric 3-D FFT on Clusters of Multi-Core Processors, 2009 SIAM Conference on Computational Science and Engineering, 2009 年 3 月 5 日, Miami Hilton Downtown (Miami, Florida, USA).

6. 研究組織

(1) 研究代表者

高橋 大介 (TAKAHASHI DAISUKE)

筑波大学・大学院システム情報工学研究科・准教授

研究者番号: 00292714