

SMP クラスタにおける高速多重極展開法の並列化

PARALLELIZATION OF FAST MULTIPOLE METHOD FOR SMP CLUSTERS

鈴木 正昭¹⁾, 奥田 洋司²⁾

Masaaki SUZUKI and Hiroshi OKUDA

- 1) 東京大学大学院工学系研究科 (〒153-8656 東京都目黒区駒場, E-mail: suzuki@nihonbashi.race.u-tokyo.ac.jp)
 2) 東京大学人工物工学研究センター (〒153-8656 東京都目黒区駒場, E-mail: okuda@race.u-tokyo.ac.jp)

The authors have applied MPI/OpenMP hybrid parallel programming model to the fast multipole method in molecular dynamics simulations on a symmetric multiprocessor (SMP) cluster architecture. The parallel performance of the MPI/OpenMP hybrid programming style was compared with that of a conventional flat programming style, which uses only MPI both for Inter- and Intra-SMP nodes. The computer used here is Hitachi SR8000/MPP placed at the University of Tokyo. When using 16 SMP nodes (128 PEs), it was observed that the performance with the hybrid style was 25% higher than that with the flat-MPI for the simulation of a system of charged particles.

Key Words: SMP cluster, MPI/OpenMP hybrid, Fast multipole method, Molecular dynamics

1. 序論

境界要素法および多体問題を扱う分子動力学法などに共通の特徴として各要素(粒子)が全ての要素(粒子)と相互作用をもつという点があり, この特徴は大規模問題の解析を困難なものにしている. 境界要素法においては離散化により生成される行列が密となり, それを解く際に必要な計算量は境界要素数を N として, 直接法を用いる場合 $O(N^3)$, 反復法を用いる場合 $O(MN^2)$ (但し M は反復法の反復回数) である. 同様に, 例えば分子動力学法においても遠距離相互作用であるクーロン項を有限距離で打ち切ることなく計算するには, 粒子数を N として $O(N^3)$ の計算量を要する. ゆえに, いずれの場合においても N が数万以上のとき, その計算の実施は現実的ではない. この問題を克服する手段として高速多重極展開法⁽¹⁾⁽²⁾が広く研究され, 利用されている. 高速多重極展開法はツリー構造による階層的な空間分割を行い, 遠方の要素からの作用を多重極展開を用いて集団で評価することで計算量を $O(N)$ へ減らすものである.

一方で, 大規模数値解析の短時間・高精度の実施に並列計算技術は必要不可欠であり, 特に近年の超並列計算機においては地球シミュレータに代表されるように SMP クラスタ型アーキテクチャの採用が一般的となっている. そのような並列計算機上で高い性能を得るには, SMP ノード間通信に対し MPI などのメッセージパッシングを, SMP ノード内並列化に OpenMP などのループディレクティブを用いるハイブリッド並列プログラミングモデル (Fig. 1) が適していると考えられる. 理由として, SMP ノー

ド内のデータ転送に関してメッセージパッシングによる通信よりメモリ共有による直接参照のほうが高速であること, また MPI における MPI_ALLREDUCE のような全対全通信に関してそこに参加するプロセッサ数が少ないほど通信時間が短くてすむことなどが挙げられる. Dual SMP-PC クラスタや 4 way SMP-PC クラスタを用いた場合にはメモリアクセスの競合などによる性能低下が大きく, 従来の MPI のみによる並列化と比較してむしろ性能が下がることが多く報告されているが⁽³⁾⁽⁴⁾, 筆者らは大型 SMP クラスタ上の並列分子動力学計算において MPI/OpenMP ハイブリッド並列プログラミングの有効性を確認している⁽⁵⁾.

本研究では大型 SMP クラスタ型並列計算機における高速多重極展開法の MPI/OpenMP ハイブリッド並列化について検討する. 荷電粒子系の分子動力学計算に高速多重極展開法の一つである Anderson の方法⁽²⁾を適用し, MPI/OpenMP ハイブリッド並列化を行い従来の MPI のみを用いた並列化 (以下, フラット MPI と呼ぶ) との並列性能比較を実施・検討する.

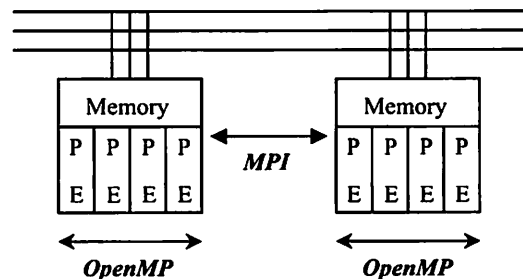


Fig. 1 MPI/OpenMP hybrid programming on SMP cluster

2. 高速多重極展開法

高速多重極展開法では、まず八分木構造を得るために全粒子を含む立方体領域を再帰的に八分割していき、最下層のセルに含まれる粒子の個数がある数以下になるまで分割を繰り返す。

実際の相互作用計算では、はじめに最下層のセルに対して、そのセル内の粒子による多重極展開を計算する。そして一段上の階層のセルについて、子セルの多重極展開の展開中心をシフトし集めることで展開を構成する。これを順次ルートセルに到達するまで繰り返す。(多重極展開のシフト)

次に、各レベルのセルについて、遠方のセルの多重極展開を局所展開に変換する。そのとき、遠いセルの相互作用ほど上の階層を通じて計算を行う。つまり、Fig. 2に示すような二次元での階層分割において黒く塗られたセルに含まれる粒子への作用を求める場合、黒く塗られたセル自身とその周りの8個のセルからの寄与はより下の階層で計算し、さらにその周りの27個のセルに対して多重極展開を変換する。それより遠いセルからの寄与は親セルのレベルで同様の評価を行う。(多重極展開から局所展開への変換)

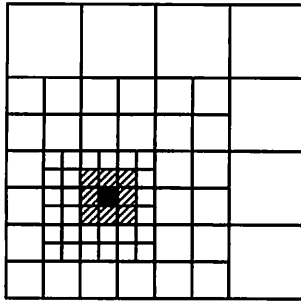


Fig. 2 Hierarchical clustering of particles

さらに、ルートセルから最下層セルに向かって局所展開をシフトしながら一段下のセルの局所展開を順次求めていく。子セルでは親セルからシフトした局所展開と同一階層の遠方セルの局所展開を足し合わせる。最下層セルにおいて、求めた局所展開からそのセル内の粒子への作用を求め、最後に近傍粒子との相互作用を直接計算したものをそこに加える。(局所展開のシフト)

以上の手順により多体相互作用の計算量を $O(N^2)$ から $O(N)$ へ下げることができる。しかしながら、プログラム実装に際しては、球面調和関数のシフトや多重極展開から局所展開への変換式が複雑であり、高速化あるいは開発自体の困難さが実用上の問題として挙げられる。

高速多重極展開法の実装を容易にする Anderson により提案された手法では、球面調和関数の展開係数をデータとする代わりにポテンシャルの値そのものを用いる。半径 a の球に含まれる粒子が球面の外側の点 \vec{x} に及ぼすポテンシャル $\Psi(\vec{x})$ は、球面上での

ポテンシャルの値が分かればポアソンの公式を用いて次式のように書ける。

$$\Psi(\vec{x}) \approx \sum_{i=1}^K \left[\sum_{n=0}^M (2n+1) \left(\frac{a}{r} \right)^{n+1} P_n(\vec{s}_i \cdot \vec{x}_p) \right] g(a\vec{s}_i) w_i \quad (1)$$

ここで、半径 a の球の中心を原点として、 P_n は n 次ルジャンドル関数、 \vec{s}_i は単位球面上の積分点 i の位置ベクトル、 w_i は積分点の重み、 r は \vec{x} の大きさ、 \vec{x}_p は \vec{x} を向いた単位ベクトル、 g は球面上におけるポテンシャルの値をそれぞれ表す。Anderson の方法では、各階層の全てのセルに対してそれを内包するような球面を定義する。まず最下層のセルでは、セル内の粒子が球面上に及ぼすポテンシャルを直接計算する。子セルの多重極展開を親セルでの多重極展開に変換するには、式(1)に従って親セルを内包する球面上でのポテンシャルを求めればよい。また、多重極展開を局所展開に変換する際にも式(1)を用いて局所展開する球面上でポテンシャルを評価する。さらに親セルから子セルへの局所展開の変換も次式で同様に求めることができる。

$$\Psi(\vec{x}) \approx \sum_{i=1}^K \left[\sum_{n=0}^M (2n+1) \left(\frac{r}{a} \right)^n P_n(\vec{s}_i \cdot \vec{x}_p) \right] g(a\vec{s}_i) w_i \quad (2)$$

Fig. 3 に Anderson の方法の概略図を示す。

分子動力学法では粒子に働く力としてポテンシャルの空間微分が必要となるが、式(2)より導くことが可能である。Anderson の方法はオリジナルの高速多重極展開法と比較して計算量のオーダーが下がるわけではないが、プログラムの最適化やチューニングが容易になる点は実用上重要である。

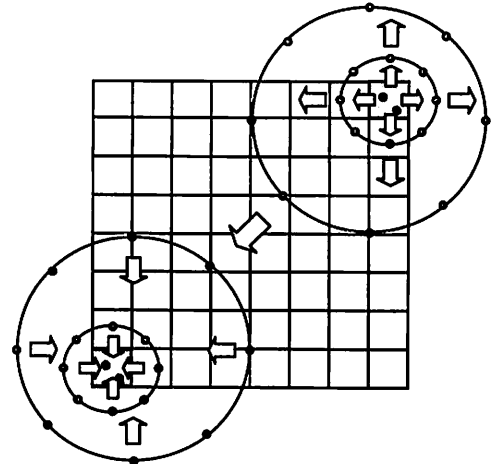


Fig. 3 Schematic drawing of Anderson's method

3. MPI/OpenMP ハイブリッド並列化

分子動力学法では粒子の初期座標・初速度を初期データとして入力する。その後はタイムステップごとに座標値を用いてエネルギー

ギ関数から各粒子に働く力を計算、その力に基づき座標を更新することを繰り返して系の時間発展を追う。一般に分子動力学法において最も計算コストを要するのは高速多重極展開を含む力の計算部であり、前述したようにクーロン力の計算量はカットオフなしの直接計算では粒子数の二乗に比例して増加する。ゆえに力の計算部を効率よく並列処理することが計算時間の短縮に有効である。

Fig. 4 は 2 SMP ノード（ノード内 2 プロセッサ）使用時の並列分子動力学計算の流れをフラット MPI, MPI/OpenMP ハイブリッドの両並列化手法について示したものである。Fig. 4 において、プロセッサ P0, P1 および P2, P3 がそれぞれ同一 SMP ノードに属し、また、 F は粒子に働く力、 X は粒子の座標、 V は粒子の速度をそれぞれ表す。フラット MPI 並列化の場合、同一 SMP ノードに属するプロセッサ間においてもメッセージパッシングによる通信を行う。つまり、実際に行われる処理は分散メモリ型並列計算機を用いた場合とまったく変わらない。一方、ハイブリッド並列化の場合、同一 SMP ノードに属するプロセッサ間のデータ転送については OpenMP によりメモリを共有し直接参照することが可能であり、通信は各 SMP ノードに一つずつ存在するマスタースレッド (Fig. 4 における P0, P2) 間でのみ行う。

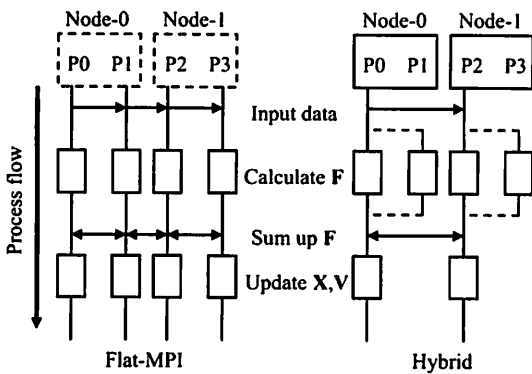


Fig.4 Process of parallel molecular dynamics method

OpenMP のようなスレッド型並列処理ではマスタースレッド以外のスレッドはプログラム中で並列実行の指示行があった時点で生成され、並列実行部分が終了した時点で消滅し、そこにはそれぞれオーバーヘッドが生じる。ハイブリッド並列化においてはマスタースレッドのみが MPI 通信を行うことを保証しなければならないが、通信のたびにその前後でスレッドを生成・消滅させては計算時間に占めるオーバーヘッドの割合が大きくなるため、極力スレッドの生成・消滅の回数を減らすことが高い並列性能を得るために重要である。そのため本研究では MASTER ディレクティブと呼ばれる OpenMP ディレクティブを用いている。プログラム中で MASTER / END MASTER ディレクティブに囲まれた箇

所の演算はマスタースレッドのみが実行し、その他のスレッドは何もせず通過する。よって、MPI 通信の前後に MASTER / END MASTER ディレクティブを挿入することでマスタースレッドのみによる通信が保証され、かつスレッドの生成・消滅を最小限に抑えることが出来る。他に、OMP_GET_THREAD_NUM 関数などの OpenMP 関数についてもその呼び出しに時間を要するため、出来る限り呼び出し回数を削減することが必要である。

本研究では各 MPI プロセスに全粒子の座標・速度のデータを持たせ、相互作用の計算部のみ領域分割型並列化をしている。Fig. 5 にフラット MPI およびハイブリッドのそれぞれの場合の、分割領域に対するプロセッサの割り当て例を示す (2 SMP ノード (ノード内 8 プロセッサ) 使用時)。ここで、セルの番号付けには Morton オーダリングを用いている。フラット MPI では 1 PE が 1 領域に対応するのにに対してハイブリッドでは 1 SMP ノードが 1 領域に対応する。MPI 通信に関して、相互作用計算部では多重極展開を局所展開に変換する際に他領域の情報が必要となるが、ここでは各プロセスが担当領域について各階層の各セルの多重極展開を全て求めた後に、それらをまとめて全対全通信を行っている。また、全プロセスが座標・粒子の更新を行うために、相互作用の計算後に各プロセスの力の足し合わせを要する。

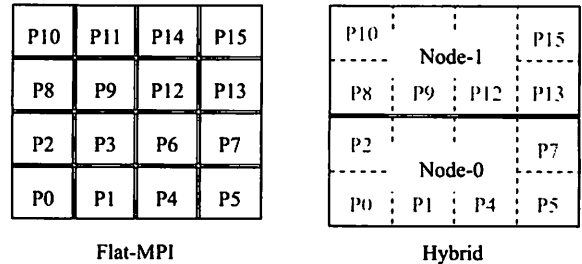


Fig. 5 Assignment of PEs

4. 数値実験

立方体領域に一様分布した荷電粒子系に対して、SMP クラスタ型並列計算機上で高速多重極展開法を用いた分子動力学計算を行い、MPI/OpenMP ハイブリッド並列化手法とフラット MPI 並列化手法の計算性能比較を行った。問題規模の異なる二つの系を考え、以下それぞれ Small モデル (59,319 粒子)、Large モデル (117,649 粒子) と呼ぶ。性能評価に際しては分子動力学法の 1 タイムステップに要する時間を計測し、式(3)に示すスピードアップ S_n および並列化効率 P_n をみた。

$$S_n = \frac{T_1}{T_n} \quad P_n = \frac{S_n}{n} \times 100 \quad (3)$$

ここで、 n は計算に用いるプロセッサ数、 T_1 は 1 プロセッサ用いたときの実行時間、 T_n は n プロセッサ用いたときの実行時間をそ

それぞれ表す。また、SMP クラスタとして東京大学情報基盤センターHitachi SR8000/MPP を 16 ノード使用した。SR8000/MPP は 8 プロセッサの SMP 型アーキテクチャを 1 ノードとして 144 ノードを多次元クロスネットワークでつないだ SMP クラスタである。演算プロセッサは擬似ベクトル処理機構をもち、1 ノードあたり 8 GFLOPS のピーク性能を有する。

5. 結果および考察

Fig. 6 にハイブリッドおよびフラット MPI 両並列化手法の並列化効率を、Fig. 7 に両手法の性能比をそれぞれ示す。この場合、Small, Large どちらのモデルにおいても使用 SMP ノード数によらずハイブリッドの性能がフラット MPI を上回っていることが確認できる。Large モデルにおける 16 SMP ノード (128 PE) 使用時にフラット MPI と比較して 25% の性能向上を示している。本計算では相互作用計算およびその足し合わせに全対全通信を要する。ここでは SMP クラスタとして 1 SMP ノードが 8 プロセッサからなる SR8000/MPP を使用しているため、ハイブリッドでは

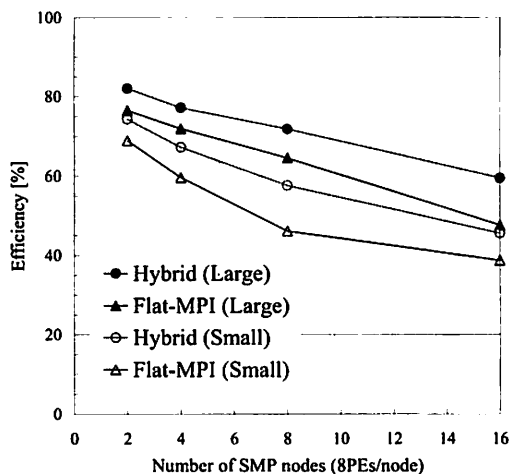


Fig. 6 Parallel efficiency (Small: N=59,319, Large: N=117,649)

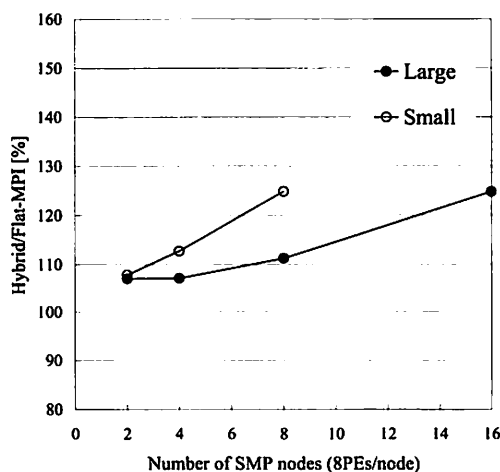


Fig. 7 Performance ratio (Small: N=59,319, Large: N=117,649)

そのような通信に参加するプロセッサ数がフラット MPI の 8 分の 1 となり通信時間の短縮が見込まれる。その効果は使用 SMP ノード数が増加するにつれて大きくなると予想できるが、実際 Large モデルにおいて最大ノード使用時に両手法の差が最も大きい。

一方で、並列化効率の数値自体には改善の余地が残る。本計算では高速多重極展開法の並列化に領域分割を用いているが、使用 SMP ノード数が増すに連れてルートセルに近い階層ではプロセッサ数が領域数を上回り、遊休プロセッサが生まれる事態が生じる。高速多重極展開法を使用した分子動力学計算においてもなお近傍粒子との相互作用の直接計算部が最も計算コストを要するため、最下層において十分な並列化がなされていれば一定の並列化効率を出すことが可能だと考えられるが、プロセッサ数が千台を超えるような場合には問題が残ることは明らかである。

6. 結論

大型 SMP クラスタ型並列計算機上の分子動力学計算において、高速多重極展開法に MPI/OpenMP ハイブリッド並列プログラミングを適用し、従来の MPI のみを用いた並列化と比較して性能が向上することを確認した。通信量が多いアプリケーションに対しては、十分なメモリバンド幅をもつ大型 SMP クラスタの使用で MPI/OpenMP ハイブリッド並列化が最適なプログラミングモデルになりうる。

本研究では数値実験に一様分布した粒子系を用い、規則的な空間分割を行った。非均一な計算対象に対し適応的な空間分割を行う場合にはさらに負荷分散の考慮が必要である。また、高速多重極展開法の効率的なベクトル化はより重要な課題である。

参考文献

- (1) L. Greengard and V. Rokhlin: A fast algorithm for particle simulations, *J. Comput. Phys.*, **73**(1987), pp. 325-348.
- (2) C. R. Anderson: An implementation of the fast multipole method without multipoles, *SIAM J. Sci. Stat. Comput.*, **13**(1992), pp. 923-947.
- (3) 立石絢也, 樫山和男: SMP クラスタでの MPI と OpenMP による並列計算効率の比較検討, 第 51 回理論応用力学講演会講演論文集, (2002), pp. 495-496.
- (4) 中林靖: SMP クラスタにおける効率的な並列化手法の検討, 日本機械学会第 15 回計算力学講演会講演論文集, **2**(2002), pp. 65-66.
- (5) 鈴木正昭, 奥田洋司, 矢川元基: SMP クラスタ上の MPI/OpenMP ハイブリッド並列分子動力学計算, 日本機械学会論文集 (A 編), **70**(2004), pp. 519-524.