

課題名 (タイトル) :

大規模並列化に適した MD プログラム開発

利用者氏名 : ○大野洋介, 小山洋, 長谷川亜樹, 舛本現
森本元太郎, 似鳥啓吾, 古石貴裕所属 : 社会知創成事業 次世代計算科学研究開発プログラム
生命体基盤ソフトウェア開発・高度化チーム

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

次世代生命体統合シミュレーション研究推進グループでは「京」を活用するグランドチャレンジとして生命科学分野でのソフトウェア開発を行っている。その中で生命体基盤ソフトウェア開発・高度化チームは他のチームのソフトウェアの開発支援や共通で利用できるソフトウェア基盤の開発を担当している。本課題の目的は開発中の基盤ソフトウェアの一つである分子動力学 (MD) コアプログラムの性能評価、改良点の分析、並列性能改良である。

2. 具体的な利用内容、計算方法

本課題では分子動力学計算で最も計算負荷のかかる二体相互作用計算の性能評価を行なった。クーロン相互作用のような二体相互作用は直接計算では演算量が粒子数の二乗に比例して増加する。そのため、原子間相互作用が遠距離では弱いことに着目し、直接計算はカットオフ距離内のみで行うカットオフ法と、原子の物理量をメッシュ点に分配して空間メッシュで遠距離相互作用を計算する Particle Mesh (PM) 法を組み合わせるのが主流であり、開発中の MD コアプログラムでも採用した。カットオフ法では近距離粒子の探索の効率化のため空間座標で粒子を分割・グループ化するセルインデックス法を採用した。計算も空間分割の単位であるセルを単位として並列化する。セルインデックス法による並列効率を向上させるとともにカットオフ計算の効率化を図る。計算対象は水分子の系とした。水分子内の結合力による運動のタイムスケールは二体相互作用によるものより短いため、結合力として計算せず結合距離を固定する拘束条件として処理する SETTLE 法を使用した。並列化は MPI のみでスレッド並列

は使用していない。

3. 結果

スケーラビリティの評価のためプロセスあたりの粒子数をほぼ一定で並列度を変えた性能測定を行なった (図 1)。ほぼ並列数に比例した性能

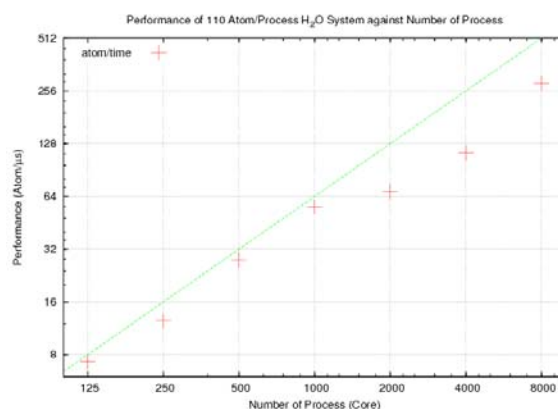


図 1

向上がみられたが、1000 並列までと 2000 並列以上では倍程度の性能差がある。これはおもに通信時間の差に起因する。通信のほとんどは近接通信であるのでプロセスあたりの通信量がほぼ一定の今回の評価では本来通信時間は変化しないが、RICC はファットツリーで上流帯域が少ないため、論理的に近接でも上流を超えた通信が 2000 並列から発生し性能低下したと考えられる。特にフラット MPI であるため帯域不足の影響がしやすいと考えられる。

実装が完了した PME 法の評価を行なった。計算サイズとカットオフ計算部分のプロセス数 (N_c) を固定 (128) し PME 法に割り当てるプロセス数 (NPME) を変化させた結果が表 1 である。

表 1

N_c	NPM E	Ntot al	Ttotal	T_c	T_{comm} +TPME	Tot her
128	1	129	92859	3516	89191	152
128	2	130	51142	3522	47470	151

平成 22 年度 RICC 利用報告書

128	4	132	29991	3521	26322	148
128	8	136	20564	3522	16894	148
128	16	144	16479	3518	12812	149
128	32	160	14815	3525	11136	154
128	64	192	17943	3527	14256	160
128	128	256	37958	3486	34308	163

きいと考えられ、単体性能評価としても RICC で確認する必要性が低下した。

以上のような計画上の必要性の変化により RICC 上での性能評価、特に週末特別利用での大規模実行を当初の予定より大幅に減らしたため、演算時間を使いきらなかった。

Ttotal, Tc, Tcomm, TPME, Tother は合計、カットオフ、通信、PME、その他の処理時間（1 ステップあたり、マイクロ秒）である。測定の都合で通信と PME の処理時間は分離されていない。PME 並列数 3 2 までは処理時間が短縮されているが 6 4 からは増加に転じている。PME プロセスは FFT 計算に fftw2 を使用したため次元分割されているが、カットオフ計算部は三次元分割されており、トポロジーの違いにより両者の間の通信がグローバル通信に近いとため、並列度を上げると通信性能が低下しているのだと考えられる。

4. まとめ

ウィークスケール性能評価でも高並列時の帯域不足が確認された。

遠距離相互作用に PME を使う場合、FFT も三次元並列しないとカットオフ計算部との通信が問題になることが確認できた。

5. 今後の計画・展望

大規模並列時の近接通信性能の変化はトラスネットワークの「京」では RICC とは大きく異なると考えられるため、これ以上の通信性能の評価・改善は実機（に近い環境）で行なうことが望ましい。

FFT に関してはカットオフ部との通信の影響が予想以上に大きく、FFT 自体の性能向上も問題があると考えられるため、FFT に依存しない手法の評価・移行を行なう必要がある。

6. 一般利用で演算時間を使い切れなかった理由

大規模並列時には RICC のファットツリーの上流帯域の影響が大きいと考えられたため、「京」を想定した開発では、RICC での大規模並列時の性能評価の必要性が低下した。また、別途行なっていた「京」での予測単体性能がかなり低く、通信性能よりも単体性能向上が優先された。

予測単体性能自体も RICC での評価よりかなり低く、CPU アーキテクチャやコンパイラの違いが大

平成 22 年度 RICC 利用報告書

平成 22 年度 RICC 利用研究成果リスト

【その他】

ポスター発表

小山洋、舛本現、大野洋介、長谷川亜樹、森本元太郎、泰地真弘人 “cppmd” current status and the future
第 2 回 バイオスーパーコンピューティングシンポジウム 平成 22 年 3 月 18 日、19 日、東京

大野洋介、小山洋、長谷川亜樹、舛本現、森本元太郎 Towards High-Performance Computing with “K
computer” 計算生命科学研究センターキックオフミーティング 平成 22 年 10 月 26 日 大阪大学 OLABB

展示会出展（ポスター）

大野洋介、小山洋、長谷川亜樹、舛本現、森本元太郎 Towards High-Performance Computing with “K
computer” SC10 平成 22 年 11 月 15 日～18 日 New Orleans, USA