

大規模分子動力学計算の並列化効率

渡辺宙志¹, 鈴木将², 伊藤伸泰³

¹ 東大物性研, ² 九大工, ³ 東大工

hwatanabe@issp.u-tokyo.ac.jp

概要 空間分割により並列化された分子動力学法コードの開発、及び並列化効率について調べた。ウィークスケールリングにて最大 8192 プロセス、41 億粒子までの計算を行ったところ、通信時間はほぼ無視できると思われるにもかかわらず、並列化効率の顕著な低下が見られた。これはオペレーティングシステムによる干渉 (OS Jitter) の影響であると思われる。

我々は分子動力学法を用いた沸騰現象の全粒子計算を行っている。マルチスケールな現象の直接計算には大規模計算が必須となる。そこで、大規模並列計算向けの分子動力学法コードを開発し、その並列化効率について調べた。コードは単純空間分割により並列化され、通信には MPI を用いた。なお、開発言語は C++ である。相互作用ポテンシャルには Lennard-Jones 型を用い、カットオフ距離は相互作用半径を σ として 3.0σ とし、一つのプロセスには一辺 100σ の立方体領域を割り当てた。プロセスあたりの粒子数を 50 万粒子に固定したままプロセス数を増やす、ウィークスケールリングにより核融合科学研究所のプラズマシミュレータ (HITACHI SR16000)、及び東京大学物性研究所のシステム B (SGI Altix ICE 8400EX) にてベンチマークを取った。その結果、並列化効率が 1024 プロセスで 80% から 90% 程度、8192 プロセスで 60% から 70% 程度にまで低下した。分子動力学法における計算のホットスポットは粒子間に働く力の計算であり、その計算時間に比べて通信時間は無視できると思われる。この並列化効率の低下要因について調べるため、通信を含まない、力の計算時間のみをプロセスごとに調べた。その結果、計算時間の平均値はプロセス数にほとんど依存しないが、計算時間の揺らぎがプロセス数が増えるにつれて増大することがわかった。

図 1 (a) は力の計算にかかった時間を全プロセスについて平均したものをステップごとに表示したものである。ノード数 (プロセス数) が増えても平均値はほとんど変わらないことがわかる。図 1 (b) はステップごとに最も遅かったプロセスの計算時間から最も速かったプロセスの計算時間を引き、平均時間で規格化したものを表示した。2 ノード (128 プロセス) ではほとんど計算時間に揺らぎが見られないのに対して、128 ノード (8192 プロセス) では大きく揺らいでいることがわかる。今回開発したコードでは高速化のために Bookkeeping 法を用いており、これはステップごとに全てのプロセスの同期を必要とする。プロセス数が増えるにつれて増大する計算時間の揺らぎが実質的なロードバランスの悪化となり、並列

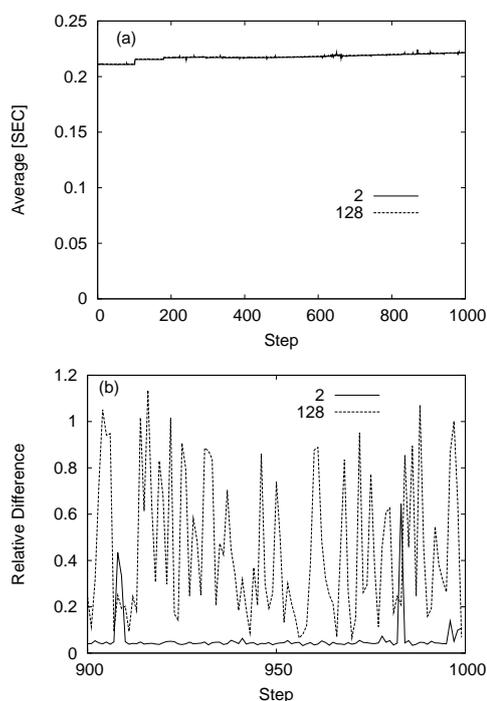


図 1: (a) ステップごとの力の計算にかかった平均時間。2 ノードの計算でも 128 ノードの計算でも力の計算の平均時間はほとんど変わらない。(b) プロセスごとの計算時間の揺らぎ。2 ノードの計算に比べ、128 ノードでの計算では揺らぎが増大している。

化効率を下げていることがわかった。この揺らぎの原因は、通信時間を含まない計算時間が揺らぎ、かつ最も遅いプロセスが試行ごとに異なることからオペレーティングシステムの干渉によるもの、すなわち OS Jitter による可能性が高い。今後、1 万個以上のコアを持つ MPP 型の計算機にて高い並列化効率を達成するためには、ユーザプログラムだけでなく、OS 側でもなんらかの工夫が必要になると思われる。

本研究は KAUST GRP(KUK-I1-005-04)、名古屋大学 COE 計算科学フロンティア、及び科学研究費補助金 若手研究 (B) 課題番号 19740235 の助成を受けて行われた。