サイエンティフィック・システム研究会　2010 年 10 月 20 日
SS 研科学技術計算分科会

3次元MHDコードによる FX1の性能評価

荻野竜樹
名古屋大学太陽地球環境研究所

［アブストラクト］

並列型スパコン FX1 や HX600 の利用により、地球磁気圏のグローバル構造とメソスケールの境界層渦乱流の自己無撞着な大規模高精度シミュレーションができる時代が訪れてきたが、その実現にはさらに高効率の並列計算コードの開発が不可欠である。ここに、キャッシュヒット率向上と多次元領域分割／一括転送による通信量最小化を可能にした高効率の 3 次元 MHD コード開発に成功し、FX1 の 3072 コア数まで 15-20%以上の実行効率を得ることができた。Flat MPI と Hybrid (Impact)の比較では、Flat MPI が最速だったが、コア数によらず Hybrid もよい実行効率を得た。通信速度と計算速度の比は 3072 コアでも 4%以下であり、通信速度は 10 万コア程度までボトルネックにならずにスケーラビリティが伸びることが期待できる。もちろん、実際にテストして実証する必要がある。

［キーワード］

地球磁気圏の MHD シミュレーション、高効率の並列コード開発、多次元領域分割と通信量最小化、Flat MPI と Hybrid の比較、スーパーコンピューティング

１．はじめに

スペースプラズマや核融合プラズマの最も特異な性質はその異常輸送現象にあるといえる。特に顕著な例として温度の輸送や磁場の拡散は粒子間衝突から得られる古典的な輸送係数の千倍から 1 万倍に達することもあることが観測や実験から知られている。しかし、その基本的な物理過程の説明は今日でも一致した考えに達していないといえる。プラズマの異常輸送の有力な説明には二つある。一つは、マクローメソーミクロスケールが必ず結合していてメソスケール現象が決定的な役割を果たしているとする説である。もう一つは、マクロとミクロスケールの直接結合が本質的でメソスケールは重要な役割を果たしていないとする説である。ここに、最近の並列型スパコンの急速な進歩により MHD（電磁流体力学的）モデルからマクローメソスケール現象を自己無撞着に解くことができるような時代が実際に訪れてきた。即ち、メソスケール現象がプラズマの異常輸送現象に本質的な役割を果たしているかどうかに決着をつけるような大規模な 3 次元 MHD シミュレーションを実行できる段階に達してきたといえる。もちろん、ミクロスケール現象は MHD モデルで扱うことができないので、超粒子モデルやブラソフモデルなどの運動論的モデルによるシミュレーションが不可欠である。

そのマクローメソスケール結合を調べるには、太陽風と地球磁気圏相互作用のグローバル３次元 MHD シミュレーションが一つの典型的な対象となる。なぜなら、磁気圏のマクロ構造の中に物理量が急峻に空間変化する境界層があり、そこにメソスケールのプラズマ乱流が発生してプラズマの異常輸送を引き起こしている可能性が高く、グローバルモデルから境界層のプラズマ不安定とその非線形発展と結果としての渦乱流発生を自己矛盾なく解ける段階に達したからである。しかし、その大規模 MHD シミュレーションを実行するためには大規模並列型スパコンの効率を最大化するプログラムの開発が不可欠である。

２．MHD コードの並列化と超並列スパコンの有効利用

MHD モデルでは流体の方程式と Maxwell 方程式を高精度の数値計算法の一つである Modified Leap-Frog 法で初期値境界値問題として解く。その 3 次元 MHD コードを MPI Fortran を用いて並列化する。その際、CPU 数が 1 千個を越えて 1 万個、10 万個程度まで計算速度が劣化しないようなよいスケーラビリティが得られる並列計算コードの開発を目指す。

私達は、これまで富士通の VPP5000/64、PRIMEPOWER HPC2500 1536core などで MPI を利用してそれぞれ 60%、20%を超える実行効率を出す 3 次元 MHD コードを開発してきた。その知識と経験を活かし、高効率の並列計算 3 次元 MHD コードを開発して、平成 21 年度に名古屋大学情報基盤センターに導入された富士通の FX1 3072core と HX600 2048core を用いてテストした。作成した MHD コードは基本的に 4 種類に分類できる。それらは、1 次元領域分割法（1Da）、2 次元領域分割法（2Da）、3 次元領域分割法（3Da）およびキャッシュヒットを上げるために MHD 方程式の配列成分を第 1 変数に移動した 3 次元領域分割法（3Db）である。2、3 次元領域分割法を用いると分散した CPU 間の通信量が小さくなり、これまでのテスト計算からは、スカラー並列機 HPC2500 では（3Db）が最速、ベクトル並列機 Earth Simulator では（2Da）が最速であった。

３．FX1 と HX600 における３次元 MHD コードの計算速度

FX1 は計算速度とメモリバンド幅が同程度との特徴があるスパコンで、Flat MPI と Hybrid（Impact）両方の利用ができて、Hybrid は 1 ノード 4 コアを共有メモリとしてスレッド並列化し、ノード間をプロセス並列（MPI）で利用する。そのどちらが高速計算できるかが最初の注目点であったが、ほとんどの場合 Flat MPI が最速になっている。ただし、Hybrid の場合もそれほど劣化しない。また、かなりバラつきはあるが、3072 コア利用までスケーラビリティはよく保たれていた。前章の 4 種類の MHD コードを比較した場合、計算速度の速い順に並べると（3Db）＞（3Da）＞（2Da）＞（1Da）となり、（3Db）で 15-20%以上の実行効率を、（1Da）で 5-15%程度の実行効率を得ることができた。特に、（3Db）ではコア数が 3072 まで多くなってもほとんど計算速度の劣化は起らず、バランスが良い場合は若干ではあるがむしろ計算効率が向上する場合もあった。JAXA の FX1 でテストした場合は、多数コア（4096、5760 コア）で Hybrid（Impact）が Flat MPI よりも高速になる場合があった。これは大変興味ある結果であるが、MHD 変数の大きさや 3 次元方向の並列分割数がより適切に取られていたかの疑問もあり、継続的なテストが必要である。また、計算時間と通信時間を分離して計測したところ、通信時間と計算時間の比は 256-3072 コアの範囲で 3-4%であり、しかもその比は 3072 コアまであまり変らないので、スケーラビリティは 1 万コアを超えてもよく成り立つことが期待される。しかし、実際はより多くのコア数で実証確認する必要がある。

HX600 も Flat MPI が最速であり、スレッド並列数を増やして Hybrid にすると計算速度がかなり劣化する傾向が明確であった。私のテストでは、最適値を求めた場合 HX600 は FX1 より常に少し計算速度が低めであった。一方、共同研究者のほとんどは逆の結果を得ていて、HX600 が FX1 よりも並列効率がより良いと結論付けている。その結果の違いがどこからきているかはまだ明確ではないが、一つの可能性として 1 コア当りのメモリ使用量、即ち粒度が関係していると予想される。

４．おわりに

学際大規模情報基盤共同研究・共同拠点のプロジェクトで他の大学の情報基盤センターのスパコンを利用する機会を得た。それらは SX-9、HA8000、SR16000、TSUBAME である。4 種類の MHD コードでテストしたが、どのスパコンも期待したとおりのよい結果が得られ、スケーラビリティもテストしたコア数（3072 コア）の範囲でよく成り立っていた。気がついた点として、T2K のスパコン（HX600，HA8000）は（2Da）が最速で（3Da）がそれに続き、（3Db）は逆に遅かった。一方、SR16000 は FX1 と同様に（3Db）が最速で両者は似た傾向を示した。これらの結果は FX1 と SR16000 ではキャッシュのヒット率向上が重要であることを意味している。また、SR16000 は SMT が非常によく効き約 2 倍の速度がでた。一方、FX1 では SMT による計算速度向上の効果は僅かであった。T2K のスパコンで（2Da）と（3Da）が速いのは「擬似ベクトル機能が効くためと推測しているが、理由究明にはさらなる系統的なテストが必要である。