

# 土星磁気圏シミュレーション結果の大規模可視化に伴う AVS/Express PCE の評価

深沢 圭一郎

九州大学情報基盤研究開発センター

黒木 勇

サイバネットシステム株式会社

## 1. はじめに

惑星磁気圏とは惑星の固有磁場と太陽から吹いてくる磁場を伴ったプラズマの風（太陽風）が相互作用して出来る領域のことを指す。この領域は惑星磁場に依るが、惑星半径の 100 倍以上の大きさを持ち様々な電磁気現象が起こる。これらは宇宙空間で起きる現象であるため探査機を打ち上げて観測を行うが、探査機では基本的に”その場”の観測しか行えない。そのため、惑星磁気圏計算機シミュレーションがこの分野の理論の発展、また観測結果の理解の促進に非常に重要な役割を果たしてきている。

このように巨大な磁気圏をシミュレーションするには、プラズマを流体として扱う電磁流体力学（MHD）方程式が利用される。MHD 方程式は式(1)から成る。

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\nabla \cdot (\vec{v}\rho) + D\nabla^2 \rho \\ \frac{\partial \vec{v}}{\partial t} &= -(\vec{v} \cdot \nabla)\vec{v} - \frac{1}{\rho} \nabla P + \frac{1}{\rho} \vec{J} \times \vec{B} + g + \frac{\Phi}{\rho} \\ \frac{\partial P}{\partial t} &= -(\vec{v} \cdot \nabla)P - \gamma P \nabla \vec{v} + D_p \nabla^2 P \\ \frac{\partial \vec{B}}{\partial t} &= \nabla \times (\vec{v} \times \vec{B}) + \eta \nabla^2 \vec{B} \\ * \vec{J} &= \nabla \times (\vec{B} - \vec{B}_d)\end{aligned}\tag{1}$$

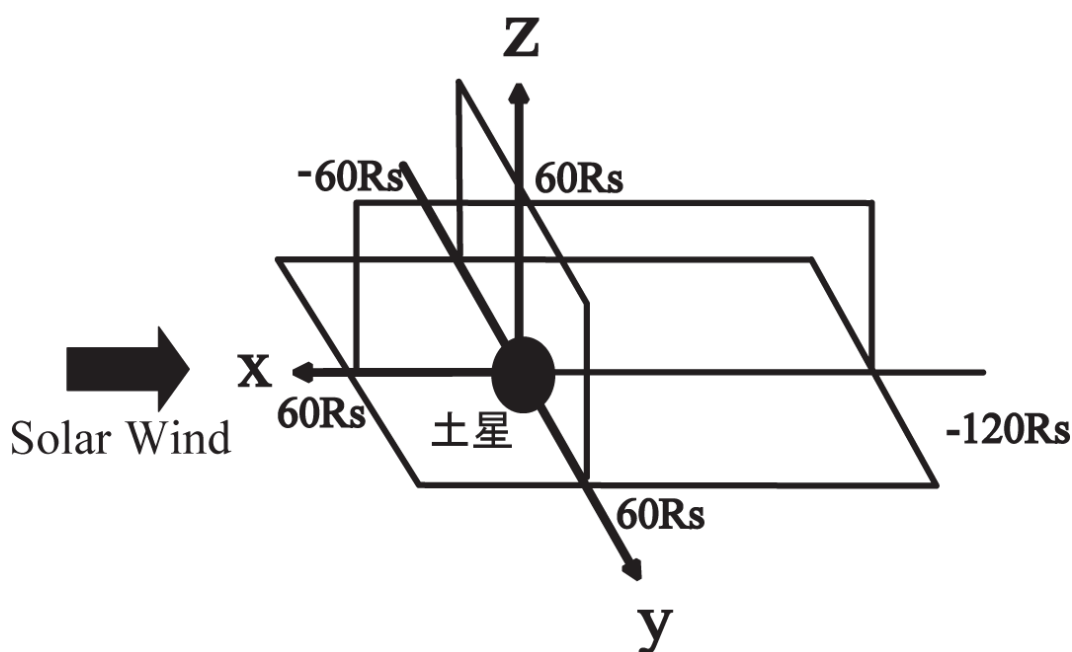
ここで  $\rho$  はプラズマの密度、 $\vec{v}$  は速度、 $P$  はプラズマ圧力、 $\vec{B}$  は磁場、 $\vec{J}$  は電流密度、 $D = D_p$  は拡散係数、 $g$  は重力加速度、 $\Phi \equiv \mu \nabla^2 \vec{v}$  は粘性、 $\gamma = 5/3$  は 3 次元の比熱定数、 $\eta$  は電気抵抗である。 $\vec{B}_d$  は惑星の固有磁場を示す。

基本的には一般流体方程式に電磁場を考慮した式系になっており、電磁場を解く分、計算量が多い特徴がある。このように計算量が多い方程式系に加えて、宇宙空間という広大な領域を自己無撞着に解くためには、膨大な計算機資源が必要となる。そのため、東京大学 HA8000 を 1024 コア利用し、シミュレーションを行っている。シミュレーションサイズ等は後述するが、現在この大規模計算結果の可視化、解析に苦勞している。そこで、今回新しく東京大学 HA8000 に導入された大規模可視化用 AVS/Express PCE を用いて可視化を行い、その評価を行った。

## 2. シミュレーションモデル

MHD 方程式を解く数値計算法としては、差分法の一つである Ogino et al. [1]によって開発された Modified Leap Frog 法を使用した。これに Flat MPI 並列を行い、シミュレーションボックスを領域分割することで並列計算した。

本シミュレーションでは、GSM (geocentric solar magnetospheric coordinates) と呼ばれる直交座標系を用いており、太陽方向を  $x$  軸の正、夕方方向を  $y$  軸の正、北向きを  $z$  軸の正にとっている。この座標系を第 1 図に表す。そして、 $x$  軸の正の方向から太陽風を流し続ける。また今回の計算では  $xy$  平面に対して対称であるので  $z < 0$  の領域は解かなくて良いこととなり、計算領域は  $60RS \leq x \leq -120RS$ 、 $-60RS \leq y \leq 60RS$ 、 $0 \leq z \leq 60RS$  の直方体となる。格子間隔を  $\Delta x = \Delta y = \Delta z = 0.1RS$  の均一格子としたので、シミュレーションボックスの格子数は  $(n_x, n_y, n_z) = (1800, 1200, 600)$  となる。この領域で MHD 方程式を解いていく。初期条件など詳細なシミュレーションモデルは参考文献[2]を参照されたい。



第 1 図:シミュレーション座標系。

左側から太陽風が吹いてくる。このモデルでは計算時間節約のため、南北対称モデルを用いている。

MHD 方程式では未知数が 8 (プラズマ密度、速度 3 成分、圧力、磁場 3 成分) があるために、シミュレーション結果として約 42GB のデータがサンプリング時間毎に書き出される。サンプリング時間は 0.069 時間であり、磁気圏構造、ダイナミクスを調べるためには 20 時間以上の結果が必要となり、約 300 個のデータを利用した (合計 12TB)。計算結果の書き出し自体は 1024 コア毎に行い、各時間に 1024 ファイルずつを分散に書きだした。今回はこの中から 1 データ (42GB) を利用し、可視化を行った。

## 3. AVS/Express PCE とその利用法

サイバネットシステムの Web[3]によると、AVS/Express PCE(Parallel Cluster Edition)と

は、PC クラスタ対応版の AVS/Express であり、クラスタ化された複数の Linux マシンで、各計算ノードが持つ部分領域のみを可視化し、最終的な可視化結果のみ制御ノード上で表示するという構成になっている。そのため、並列計算の結果、出力される大規模データを可視化する場合でも、高い精度を保ったまま、可視化処理を実現することが可能なアプリケーションである[3]。また並列可視化機能の特徴として、1)データの読み込みからレンダリング処理まで、可視化処理のほとんどを計算ノードで実行することにより、大規模データの可視化を実現、2)計算ノード複数台で処理することにより、1台では取り扱えないような大規模データを可視化することが可能、3)分散して格納したデータを、制御ノードに集めることなく可視化することが可能、がある。このような特徴により分散書き出ししてある大規模計算結果の可視化に適しているということで、シミュレーション結果の可視化を行った。

AVS/Express PCE の詳細な利用方法については本号に掲載されている記事「AVS/Express PCE の使用方法」を参照されたい。ここでは少し注意が必要な起動方法について簡単に説明する。

## 1) X サーバー環境

AVS/Express PCE を起動するために、ログインノードに X のフォワーディングを行うようにログインする。例えば、ssh コマンドでログインする場合、

```
ssh -X (ha8000-1 ~ ha8000-4 のいずれか)
```

のようにログインする。

## 2) PCE の起動（ジョブの投入と操作画面の起動）

以下のコマンドを実行する（以下の例では 4 並列）。

```
/opt/itc/avs/pce_scripts/pce_start.sh 4
```

起動のシェルに、並列数を引数に与えて起動する。

上記のコマンドを実行後、ターミナル画面には以下のような表示があり、AVS/Express PCE が起動し、並列ジョブ（各並列ノード）と通信が確立するまで、待ち状態となる。

```
Parallel num = 4
```

```
Request 749226.batch1 submitted to queue: avs.
```

```
$
```

```
AVS/Express Developer Edition
```

```
Version: 7.3 fcs linux_64_el4
```

```
Project: /opt/itc/avs/xp_pce73/express_pce73/avs_disp
```

続けて作業していると、前のソケットがクローズできておらず、次の通信ができない場合がある。

```
can not create server socket.
```

上記メッセージが出て、すぐに終了する場合、少し時間を置いてから、起動すると起動できる。また、後述のジョブが実行中の場合、次の起動ができないため、すぐに終了した場合には、ジョブを qdel コマンドで削除する。

また、起動したターミナルで、リターンキーを入力すると、プロンプトが戻るので、qstat コマンドでジョブを確認すると下記のような状態が表示される。

2011/06/01 (Wed) 14:27:34: REQUESTS on HA8000 cluster

NQS schedule stop time : 2011/06/24 (Fri) 8:55:00 (Remain: 546h 27m 26s)

REQUEST	NAME	OWNER	QUEUE	PRI	NODE	E-TIME	MEM	STATE
749227.batch1	pce_job	z39912	avs	63	4	7200s	112GB	RUNNING

ジョブの STATE が「RUNNING」になり、操作画面上に [Default user\_area] と [Select File] ボタンを持つウィンドウが表示されると、AVS/Express PCE が操作できる環境となっている。ジョブの状態が「QUEUED」のときは、他に誰かが使用しているか、何らかの都合で avs キューが停止しているため、使用可能となるまで待機する。

環境によって、このウィンドウが表示されない場合があるが、ジョブのエラーメッセージに Host key verification failed.

とある場合は、ノード間で通信が行えていない。その場合は以下の手順でパスフレーズ無し  
のノード間ログインが出来るようにする。

i) ログインノードでパスフレーズなしの SSH 鍵を作成

ssh-keygen コマンドを使ってパスフレーズなしの SSH 鍵を作成

ii) home において鍵を共有

```
cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
```

iii) 計算ノードのホスト公開鍵を取得。

下記のようなジョブスクリプトを作成し、qsub でジョブを実行する

```
#!/bin/csh
```

```
#@$-q avs
```

```
ssh-keyscan -t rsa y045
```

```
ssh-keyscan -t rsa y046
```

```
ssh-keyscan -t rsa y047
```

```
ssh-keyscan -t rsa y048
```

iv) ホスト公開鍵を登録

iii) で取得したホスト公開鍵を \$HOME/.ssh/knownhosts に追加する

※一回に同時に利用できるユーザーは 1 名で、avs キューの制限時間は 120 分である。

起動シェルの引数には、並列数 (2 ~ 64) を指定する必要があり、起動可能なプロセスは 1 ノードあたり 16、使用可能ノード数は 4 ノードまでとなっている。したがって並列数は最大 64 となる。

### 3) AVS/Express PCE の終了

作業終了後は、上記制限のため、AVS/Express PCE を正常終了させる。第 2 図にあるように通常の AVS/Express と異なり Exit ボタンで終了する。終了後、qstat コマンドで avs キューが空いていることを確認する。



第 2 図: AVS/Express PCE の終了ボタン

今回は 1024 ファイルを 64 並列で読み込むため、並列数とファイル数は一対一対応では無く、64 個のプロセスが 16 回繰り返し読み込みを行い、データ処理も 64 個のプロセスが 16 回処理を行うことで可視化処理を行った。Field ファイルに工夫があるが、詳細については後日公開予定のサイバネットシステムによる説明書を参照されたい。このように並列数よりデータが多い場合でも対応は可能である。

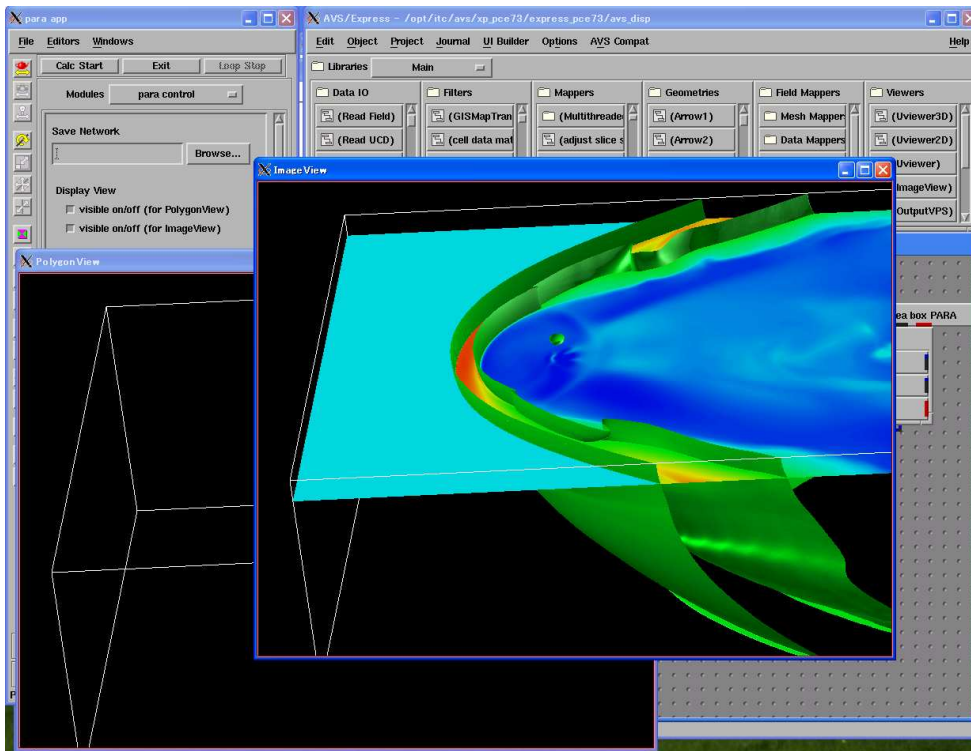
#### 4. AVS/Express PCE の評価

可視化を行うデータは前述のように  $1800 \times 1200 \times 600 \times 8$  の 3 次元データ + MHD8 変数であり、約 42GB のデータ量である。このデータは 1024 個に分割されており、それぞれ約 42MB となっている。これら 1024 個のデータを 64 プロセス利用し、読み込みを 16 回繰り返すことでデータを読み込むことが出来た。また分散データを読み込む場合、1 メッシュ分ののりしろ（隣り合うデータと重なり合う部分）が必要となるので注意が必要である。

今回 AVS/Express PCE は九州大学情報基盤研究開発センターから東京大学情報基盤センターに ssh で接続し、X フォワーディングを利用し、実行した。クライアントマシンは Widows PC である。

まず、AVS/Express PCE を立ち上げ、V ファイルを読み込むと 70 秒程度かかった。その後、実際にイメージビューに可視化結果が表示されるまでに約 2 分かかった。この処理は 1 回目は時間がかかり、2 回目以降は少し早くなると聞いていたが、実測では 2 回目も同様の時間がかかった。このときに可視化した図は第 3 図の通りである。この図のように AVS/Express PCE では二つのビュー画面があり、一つは枠線を表示し、位置変換、拡大縮小を指示するポリゴンビュー（第 3 図中では左下部）、実際に絵が表示されるのがイメージビュー（第 3 図中の中央部）と呼んでいる。

イメージビューへの表示が出来た後、画面の視点を変えてみると、表示が DirectX に対応していないためか、動きが遅く、画面がフリーズする。実行環境を Windows では無く、Mac 上でもテストしたところ、非常に処理は遅いが動作はできた。また、各ウインドウのサイズを変更すると、動作が重くなりスムーズにサイズ変更が行えないことがある。スライス面の変更や isosurface の閾値変更などは 30 秒程度であり、42GB を読み込んで処理をしていると考えれば、比較的快適であった。



第 3 図: AVS/Express PCE での土星磁気圏可視例

処理系の問題として、通常の AVS/Express では可能だが、AVS/Express PCE では実行できない処理がいくつか存在する。AVS/Express PCE では複数プロセスでデータを読み込み、可視化作業を行っているために、自プロセスと異なるデータを参照するような可視化は行えない。例えば通常の AVS/Express であれば、ストリームラインは簡単に描けるが、AVS/Express PCE では他プロセスのデータまでは追いかけて積分しないために、自プロセス内のデータ範囲でしかストリームラインを描くことは出来ない。この問題への対処はベクトル値だけを別ファイルに書き出し、1 プロセスで読み込み、結果を重ね合わせるということが提案されている。

最後にストレージの違いによる読込速度を調べた。今までの作業は HSFS (Hitachi Striping File System) 上にあるデータを利用していたが、東京大学情報基盤センターには並列アクセス対応の Lustre が導入されているので、AVS/Express PCE 上でデータ読み込み時間のテストを行った。テストでは一つにまとめたデータを HSFS と Lustre 上に置き、読込速度を測った。その結果、3 回の読込平均で、HSFS では 1820 秒、Lustre では約 10 秒かかった。このように圧倒的に Lustre を利用した方が速かった。このことから AVS/Express PCE 利用時は Lustre の利用を強く薦める。

#### 4. まとめと問題点

今回東京大学情報基盤センターに導入された AVS/Express PCE を利用し、性能評価を行った。可視化データは電磁流体力学シミュレーションに依って計算された土星磁気圏であり、サイズは 42GB である。AVS ではすべてのデータを読み込みデータ処理を行うために、並列処理の AVS/Express PCE の良さを生かすためにもデータファイルは NFS では無く Lustre に置く

方が良い。東京大学では 64 プロセスまで PCE で利用できるが、Field ファイルを工夫することで、64 個以上の分散ファイルでも読み込みに対応している。

データを読み込んだ後は 42GB にもかかわらず、比較的レスポンス良く動作した。ハードウェアの設定もあるが、OpenGL や DirectX に対応させておかないと描写に時間がかかり、ウィンドウサイズの変更でさえレスポンスは悪くなる。

まだ、いくつか問題はあるが現実には 42GB 程度のデータを読み込んでインタラクティブに可視化できる環境はほぼ無い事を考えると、現状では利用する価値のあるポスト処理アプリケーションである。

### 参 考 文 献

- [1] T. Ogino, R. J. Walker, M. Ashour-Abdalla, A global magnetohydrodynamic simulation of the magnetopause when the interplanetary magnetic field is northward, IEEE Trans. Plasma Sci.20, 817-828, 1992
- [2] 深沢圭一郎、梅田隆之、荻野瀧樹、高効率並列電磁流体コードによる HA8000 クラスシステムの性能評価、2010 年ハイパフォーマンスコンピューティングと計算科学シンポジウム論文集、133-140、2010.
- [3] <http://kgt.cybernet.co.jp/feature/pce/>