A109

大規模流体解析における遠隔並列可視化システムの検討

伊藤浩之[○](菱友システムズ(株)), 元木伸治(JAXA), 堤誠司(JAXA)

高木亮治(JAXA), 嶋英志(JAXA), 松尾裕一(JAXA)

Study on a Cluster System for Remote Visualization of Large-Scale CFD Results

Hiroyuki ITO, Shinji MOTOKI, Seiji TSUTSUMI,
Ryoji TAKAKI, Eiji SHIMA, Yuichi MATSUO

ABSTRACT

Testbed PC cluster is installed to evaluate the parallel and remote visualization of large-scale computational results produced by the supercomputer in the remote site. Three commercial and open-source software are employed in this study. It is observed that the partitioned file format such as VTK multi-block is desirable for reading large-scale data. High scalability more than 80 % is obtained by using the parallel rendering and image compositing. In spite of the insufficient network speed, interactive performance for visualization is satisfactory when the PC-over-IP technology is implemented. This testbed cluster with 80GB system memory can handle the numerical result up to 0.7 billion nodes. The knowledge obtained here will be utilized for the post-processing system of the large-scale computations done by remotely connected supercomputers.

Keywords: Large-Scale Data, Remote Visualization, Parallel Visualization

1. 結 論

JAXA/情報・計算工学(JEDI)センターでは予てより H-IIA, M-V, 及び次期固体ロケットの打上げ時音響解析を行ってきた。¹⁾²⁾ これらの解析では, より広範囲の周波数帯域を捉えるため, 密な格子を必要とし, かつ長いタイムスケールを取得する必要があることから, その計算規模は必然的に大規模となる。

現行計算ではベクトル型マシンである SX-6(理論演算性能約 1.2TFLOPS)を使用しており, 出力されるデータ量は 1 ケース当たり 25GB 程度であるが, 今後 JAXA の新統合スーパーコンピュータシステム(JSS)(理論演算性能 120TFLOPS)の利用により, さらに大規模な計算が実施可能になる。JSS がスカラー型マシンであることから, その実効性能を理論性能の 10~20%程度とした場合でも, 現行の 10 倍以上の計算性能を持つこととなり, 同様に 1 ケースあたりのデータ量も 250~500GB 程度に達すると予想される。

JSS 本体が設置されている調布航空宇宙センターと作業場所の相模原キャンパスは現在 1Gbps の回線で接続されており, bcbp を利用すると 20MB/sec 以下の転送速度しか得られない。従って, 500GB のファイルの転送にはおよそ 7 時間を要すると予想され, ファイルを手元の WS に転送するという従来の方法は, 転送時間, ローカルディスク容量の両面から考えて現実的ではない。

また流体解析という側面から考えると, 画像をバッチ

Fig.1 Visualization Cluster System

Fig.2 Remote desktop by the PC-over-IP technology

Table 1 Machine Configuration (Frontend and Nodes)

CPU	AMD Opteron 2380 2.5GHz(quad-core) x 2
Main Memory	16GB
GPU	NVIDIA Quadro FX 4800, 4600 (x2 SLD)
OS	RedHat Linux 5

処理で取得するのみでは不十分であり, 得られた結果をインタラクティブに表示・操作しながら, 流れ場の観察を行えることが望ましい。しかし, 大規模解析では,

Iso-surface 等のポリゴン表示、I/O 処理に多大な時間を要し、ファイル操作や描画の待ち時間が作業の大半を占めることになるため、結果としてインタラクティブ性が損なわれると予想される。

従って、今後の大規模流体解析においては、データファイルを計算機上から移動せずに遠隔可視化を行い、かつ可視化プロセスの並列化等の手段を用いて、インタラクティブ性を確保することが重要な課題である。

今回、当センターでは大規模解析における可視化システムのテストベッドとして遠隔並列可視化クラスタを構築し、将来の可視化システムにとって最適なハード/ソフトウェア構成の検討を行った。

2. 大規模 CFD 結果の可視化に要求される機能

大規模データを遠隔から可視化し、加えて GUI 操作を用いて、任意の領域の指定した物理量を素早く表示できるようなインタラクティブ性の確保を目指した。そのためには以下に代表される機能が必要であると考えられる。

2.1 遠隔可視化機能

ローカルへのファイルの転送を省き、解析結果を遠隔地のクライアント上で可視化するために、Server-Client 機能に対応したソフトウェアを導入し、ファイルの転送時間を省く。また、リモートデスクトップによる可視化システムの直接操作も可能とすることで、より高いインタラクティブ性を目指す。

2.2 I/O, ポリゴン生成, レンダリング処理の並列化

大規模データを高速で処理するためには可視化の各段階における並列化が要求される。I/O 処理では解析後の結果ファイルをマージする時間を省き、かつ高速な読み込み速度を実現するため、スーパーコンピュータのディスク装置にマウントを行い、そこから計算直後の分散ファイルをそのまま読み込めることが望ましい。レンダリングの並列化に関しては、インタラクティブ性確保のために、GPU を用いた並列ハードウェアレンダリングが必須と考えられる。米ローレンスリバモア国立研究所、米サンディア国立研究所など^{3,4)}では、既に多数の GPU を用いた並列レンダリングが実施されており大規模可視化において成果を上げている。

2.3 より大規模なデータへの柔軟な対応

さらに大規模なデータを可視化するためには、処理時間及びメモリ消費量の削減のため、I/O の階層にまで及ぶ ROI (Region of Interest) や LOD (Level of Detail) 等の活用が必要であり、それら幅広いスキームに対応したソフトウェアの導入、及びシステムの構築を行う必要がある。

3. ハード・ソフトウェア環境

3.1 可視化クラスタシステム構成

今回構築した遠隔並列可視化クラスタシステム(以下可視化システム)の概要を Fig.1 に示した。各ノード、及

びフロントエンドのマシン構成はほぼ同一であり、その性能諸元を Table.1 に示した。

運用上の制約から JSS のディスク装置には直接マウントを行っておらず、データは RAID 内に置かれている。ファイルサーバとクラスタはそれぞれ 10G の Ethernet で接続され、RAID からフロントエンドへの転送速度は Bonnie++ を用いた実測で約 132.2 [MB/sec] である。

3.2 可視化ソフトウェア概要・特徴

本稿における評価で用いた遠隔・並列化対応の可視化ソフトウェアは以下のとおりである。

- ・ FIELDVIEW-Parallel
- ・ EnSight DR
- ・ ParaView

2. の要求実現に近付けるため、遠隔可視化が可能な Server-Client 機能を有し、大規模な並列化に対応していると思われるソフトウェアを選択した。

いずれのソフトウェアも、可視化クラスタの各ノード上で並列処理にてポリゴン生成を行い、クライアント側にポリゴンを転送して表示を行う機能を有している。これを本稿では『ポリゴンモード』と呼称する。また EnSight と ParaView に関しては、各ノードの GPU を使用した並列レンダリングによる画像転送、重畳の機能を有しており、こちらを『ビットマップモード』と呼称する。後者は各ノードでレンダリングした画像をクライアントにて重畳するため、ネットワークの負荷が軽く、クライアントのメモリ消費が非常に少ないことが特徴である。

なお、FIELDVIEW に関してはビットマップモードこそ実装していないものの、並列化に要するセットアップが容易であり、Server-Client を使用する際にも、通常版の FIELDVIEW を使用する場合とほとんど変わらない感覚で操作でき、他のソフトに比べユーザーフレンドリーである点が特徴である。

ParaView⁵⁾はオープンソースの並列対応可視化ソフトウェアであり、FIELDVIEW と EnSight が PLOT3D フォーマットを使用しているのに対し、ParaView では VTK Multi Block と呼ばれる独自のフォーマットを利用している。この VTK Multi Block は各格子ブロックを別々のファイルとして取扱っている分散ファイルフォーマットであり、元より並列可視化を意図したフォーマットであるといえる。

3.3 遠隔可視化方式

可視化システム本体は調布に設置されており、実際の操作はネットワークを介して相模原キャンパスにて行う。その際に利用できる遠隔可視化の方法として以下の2つがあげられる。

一つは相模原の WS 上でクライアントプロセスを立ち上げ、Server-Client 機能により、可視化システム上のサーバープロセスからポリゴン/画像を WS へ送る方法であり、もう一つは可視化システム内のフロントエンド上

にクライアントプロセスを立ち上げ、Fig.2に示したハードウェア圧縮技術（以下 PC-over-IP）を利用してシンクライアント上から遠隔操作を行う方法である。

インタラクティブ性の確保という観点から考えると、後者の方法では可視化システム上のデスクトップ画像を直接ハードウェア上で圧縮して遠隔地まで転送するため途中のネットワークの負荷を最小限に抑えることができ、その点で WS へポリゴンや画像を単純に送るよりも有利である。ただし、転送される画像の画質、及びフレームレートは回線速度の制約を受けるため、低速の回線では、ブロックノイズやマウス操作の遅延が発生する可能性もある。

3.4 性能評価に使用した格子・データファイル

Fig.3 に本評価にて使用した格子、及びデータファイルを示した。格子サイズは $512 \times 512 \times 512$ (格子点数 0.13billion), 64Block のマルチブロック格子である。なお、Iso-Surface の生成ではデータファイルとして、sin 関数からなるダミーデータを 1 変数のみ与え、関数値 $f(x)=0$ の等値面を描画している。

本評価ではロケット打上げ時の音響計算に使用している構造格子のみを対象としており、非構造格子での検証は今回行っていない。

4. 結果及び考察

4.1 可視化処理時間

本評価では、可視化システムのフロントエンド上でクライアントを立ち上げ、1-32 並列時におけるファイルの読み込みに要する時間と、Iso-Surface の描画に要する並列化効率を得た。並列時に使用される CPU 数及びコア数の内訳は Table.2 に示した通りである。これらの評価に当たっては毎回メモリキャッシュをクリアし、計 3 回の計測を実施してその平均を使用した。

ファイルの読み込みに要する時間を並列度別に Fig.4 に示した。また、Iso-Surface 描画時における各ソフトウェアの並列化効率(PE)を Fig.5 に示している。

Fig.4 より、非並列時の読み込み時間はいずれのソフトウェアも 80(Sec)~110(Sec) となっている。並列時ではファイル形式に VTK Multiblock フォーマットを用いている ParaView が最も高速の結果を出しているが、非並列時の速度は FIELDVIEW とほぼ同等であり、ファイルの読み込み時間の並列化は各ブロックをあらかじめ分散配置したフォーマットがより有効であることがうかがえる。EnSight は他のソフトウェアよりも若干読み込みに時間を要しているが、これは EnSight が PLOT3D の自動分割処理を行った上で読み込みを行う仕様になっているためと考えられる。

いずれのソフトウェアも並列度に応じて読み込み時間は高速化するものの概ね 8~16 並列で高速化は限界に達している。特に ParaView に関しては 16 並列と 32 並列で処理時間がほぼ一致しており、ノード間の通信速度が

Table.2 Comprising item of parallel level

並列度	Frontend	Node1	Node2	Node3	Node4
1	1CPU × 1Core				
4	1CPU × 1Core	1CPU × 1Core	1CPU × 1Core	1CPU × 1Core	1CPU × 1Core
8	1CPU × 1Core	2CPU × 1Core	2CPU × 1Core	2CPU × 1Core	2CPU × 1Core
16	1CPU × 1Core	2CPU × 2Core	2CPU × 2Core	2CPU × 2Core	2CPU × 2Core
32	1CPU × 1Core	2CPU × 4Core	2CPU × 4Core	2CPU × 4Core	2CPU × 4Core

Fig.3 Preliminary test data

Fig.4 File reading time

Fig.5 Parallel Efficiency (PE)
(Iso-surface drawing)

ボトルネックになっていると推測される。

Iso-Surface の描画については、Fig.5 より、ビットマップモードの 8 並列までの PE が EnSight, ParaView 共におよそ 80% 程度となっており、比較的高い値を維持している。これに対し、ポリゴンモード時の並列化効率はビットマップモードと比較して低く、また、並列度の増大に伴う PE の低下も激しい事が分かる。特に ParaView の場合、ポリゴンモードのみの状態では 8 並列時でも 10% 程度の効率しか得られていない。今回、評価に用いた Iso-Surface のポリゴン数では、ポリゴン転

送時にネットワークがボトルネックとなり性能が伸びていないと考えられる。従って、大規模解析の可視化において、大量のポリゴン転送を必要とする場合、ビットマップモードを使用の方がより高いスケーラビリティを確保できるといえる。

また、ビットマップモードを用いた場合でも、高いPEを得られているのは8並列時までで、それ以上プロセス数を増やした場合、効率が急減する。その傾向はEnSightの場合で特に顕著であり、8並列時80%程度だったPEが、16並列では50%程度まで低下している。本可視化システムが計8CPUであり、1CPU内で複数コアを使用した場合に並列化性能が低下していることから、CPUのメモリバンド幅の影響を受けていると考えられる。

4.2 遠隔可視化結果

相模原からの遠隔可視化について、PC-over-IPを用いて可視化システムのフロントエンドをシンククライアント上で操作した場合とWSをフロントエンドとしてServer-Client機能を用いた場合に関して、FIELDVIEWの結果を用いて比較を行った。

その際のIso-Surfaceの描画に要する時間をFig.6に示した。WS上でServer-Clientを使用した場合、シンククライアントと比較しておよそ2倍～3倍程度の時間を要しており、調布・相模原間の転送速度がボトルネックになっていることが分かる。ただし、FIELDVIEWはポリゴンモードのみの検証になるため、転送負荷は他のソフトのビットマップモード時より高いと考えられる。

一方シンククライアントでの操作ではフロントエンド操作時において、画面の見かけ上のフレームレートが回線速度に応じて低下し、マウスの操作にストレスを感じた。

4.3 大規模可視化に向けて

大規模格子への対応においては、速度はもとよりデータ読み込み時のメモリ消費量が重要となる。

Fig.7に32並列の下、896³(格子点数0.72Billion)の格子を用いてFig.3と同様のIso-Surfaceを描画した場合のメモリ消費量を示した。ここでFIELDVIEWはポリゴンモードを用いているため、フロントエンドのメモリを3GB程度消費している。メモリ消費の観点では、いずれのソフトウェアも計80GBのメモリを搭載した本可視化システム上で格子数7億点程度の可視化処理を十分行えることが分かる。

5. 結論

新たに構築した可視化クラスタシステムを用いて、その性能評価と、より有効な遠隔並列可視化の手段についての検討を行った。得られた結果は以下の通りである。

- ①I/O 時間ではVTK-MultiBlockを使用するParaViewが最速の結果を出しており、I/O時間の短縮のためには分散ファイルフォーマットの使用が有効である。
- ②Iso-Surfaceの表示では、ビットマップモードがポリゴンモードと比較して高いスケーラビリティを示す。

Fig.6 Remote visualization performance

Fig.7 Memory consumption of 0.72 billion nodes (Iso-surface drawing)

③1CPU内において複数コアを使用した場合、PEが低下する。これはメモリバンド幅の制約によるものと考えられる。

④描画速度の点ではPC-over-IPに代表されるリモートデスクトップ技術を用いたシンククライアントの使用が効果的である。ただし回線速度によっては、操作性にストレスを感じる程度にフレームレートが低下するので注意が必要。

今後、実際の解析データを用いたより大規模で実用的な使用状況を想定した評価を行い、将来の本格的な可視化システム導入に向けての基礎的なデータとしたい。

参考文献

- 1) Tsutsumi, S., Fukuda, K., Takaki, Shima, E., Fujii, K., and Arita, M., "Generation and Propagation of Pressure waves from H-IIA Launch Vehicle at Lift-off", AIAA Paper 2008-390, Jan.2008.
- 2) Tsutsumi, S., Fukuda, K., Takaki, Shima, E., Fujii, K., and Ui, K., "Numerical Study on Acoustic Radiation for Designing Launch-Pad of Advanced Solid Rocket", AIAA Paper 2008-5148, July, 2008.
- 3) http://www.nvidia.com/object/IO_27539.html
- 4) https://publicaffairs.llnl.gov/news/news_releases/2005/NR-05-11-04p.html
- 5) Andy Ceilnik, Berk Geveci, Kenneth Moreland James Ahrens, and Jean Favre, "Remote Large Data Visualization in the ParaView Framework", EGPGV2006