

スーパー SINET を利用した大規模遠隔可視化処理の評価

滝沢寛之 小林広明

東北大学情報シナジーセンター スーパーコンピューティング研究部

概要

広域に分散した各種資源を共有するグリッド環境では、遠隔地にある資源とネットワークを介して効率的に協調作業を行う必要がある。大規模科学技術計算の途中結果を実時間可視化し、さらにその可視化画像を対話的に操作可能なシステムを遠隔協調作業の実例として考察する。本稿では、そのようなシステムのプロトタイプを実装し、その性能を評価する。スーパー SINET で接続された北海道大学と東北大学との間の帯域幅やフレームレートを実測したところ、画像を圧縮転送することにより十分実用的な性能が得られることが分かった。

1. はじめに

近年、物理的に広域に分散した計算資源を高速ネットワークで接続して有効活用するための技術として、グリッド (GRID) が盛んに研究・開発されている。グリッドに期待される応用先として、単一計算機では実現できないほど大規模な数値計算 (計算グリッド) やデータ処理 (データグリッド) への応用 [1, 2]、遊休資源の有効活用 [3, 4]、および遠隔地の各種資源との協調作業 [5, 6] などが挙げられる。

スーパー SINET[7] は、日本全国の主要な大学や研究機関の間を 10Gbps の広帯域で接続している研究用広域ネットワークである。グリッドの今後の発展および普及のためには高速かつ広帯域なネットワークの整備が必要であり、そのようなネットワークが整備された近未来の環境下での広域グリッドの性能評価を先進的に行うために、スーパー SINET は必要不可欠な情報基盤である。現在、スーパー SINET を活用して全国共同利用大型計算機センター群のスーパーコンピュータ間を接続し、計算グリッドとして利用する計画 [8] が進行している。

本稿では、今後グリッド技術の応用が期待されている全国共同利用施設間の相互利用を実証実験するため、北海道大学と東北大学との計算資源間の遠隔協調作業環境を実装し、その性能評価を行う。東北大学情報シナジーセンターからスーパー SINET を介して北海道大学情報基盤センターの超高速可視化サーバ [9] を対話的に遠隔利用し、実効帯域幅や描画性能を評価する。

2. 可視化サーバ遠隔利用システム

広域に分散した計算資源を共有するグリッド環境では、スーパーコンピュータなどのデータを産出する計算資源 (以下、演算サーバとする) とそのデータを可視化・解析するための計算資源 (可視化サーバ) が物理的に離れている可能性が高い。さらには、これらの高価な演算サーバおよび可視化サーバを効率よく運用するためには、遠隔地からの共同利用といった利用形態が考えられる。特に協調作業によりシミュレーション結果を解析するような場合には、可視化情報を遠隔地にいる共同研究者が共有することが必要となる。本稿で想定する演算サーバ、可視化サーバ、および利用者の端末の関係を図 1 に示す。

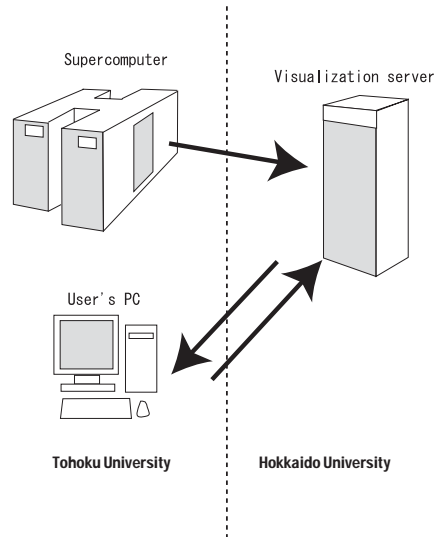


図 1. 可視化サーバ遠隔利用システム

本稿では、演算サーバによるシミュレーションの途中結果として得られた大容量データを、遠隔地の可視化サーバに順次転送し、実時間可視化するシステムを構築する。可視化の結果得られた画像は、可視化サーバから遠隔地にある利用者の端末へと転送される。利用者は可視化結果を表示し、それに対して回転や拡大縮小などの対話的操作を行うことが可能である。つまり、このシステムを構築することで、物理的に離れた場所にいる複数の利用者が高価な可視化サーバを共有し、限られた計算機資源を有効に活用できるようになる。

大規模な科学技術演算の場合、演算の結果得られるデータのサイズもまた巨大なものになる場合が多いため、演算サーバと可視化サーバ間には広帯域な通信回線が必要である。現在、演算サーバと可視化サーバ間のデータ転送には FTP や NFS といった手段が一般的であるが、演算結果の対話的な可視化の実現には適していない。また、可視化サーバを対話的に利用するためには、可視化サーバと利用者の端末との間にも広帯域で、かつ通信遅延の小さい高速な通信回線が必要である。例えば、 512×512 ピクセルの 24 ビットカラー無圧縮画像はおよそ 6.3Mbits であり、これを例えば無圧縮のまま毎秒 30 フレーム転送するためには、およそ 189Mbps もの広帯域が必要である。

3. システムの設計および構築

可視化サーバで生成した画像を利用者の端末に表示するために、SGI 社の OpenGL VizServer(以下、VizServer とする)を用いる [10]。VizServer は、可視化処理の結果得られた画面をネットワークを介して転送するシステムのため、利用者の端末に可視化用の高性能グラフィックスハードウェアを用意する必要が無い。画像に対するクリックやドラッグ、キー押下などのイベントが可視化サーバへ通知されるため、対話的操作の実現も可能である。また、VizServer は送信(可視化サーバ)側で画面データを実時間で圧縮しながら送信し、受信(クライアント)側ではその圧縮されたデータを実時間で伸張して表示する機能を有している。データ圧縮法として可逆圧縮と不可逆圧縮の双方を備えており、特に不可逆圧縮の場合には画面の画質の劣化と引き換えにサーバ・クライアント間の画面転送による占有帯域幅を大幅に削減できる。さらに、同じ画面を同時に複数の端末へと転送することができるため、遠隔地にいる

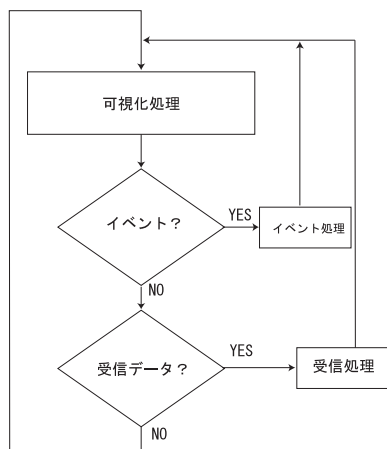


図 2. 可視化プログラムの流れ

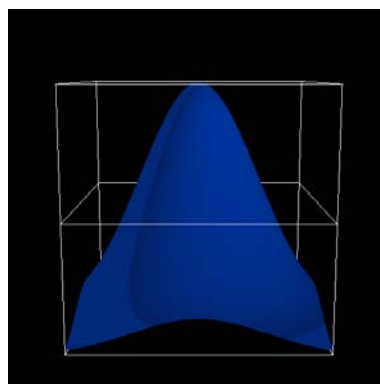


図 3. 三次元グラフデータ

人との共同作業や可視化の結果に対する認識の共有を図る際にも有用である。

可視化サーバと利用者端末との間の通信部分に関しては VizServer で実現可能であるため、本研究ではそれ以外の実装を行う。演算サーバ側のプログラムは単純に演算の途中結果をデータ転送するだけで良いため、実装は容易である。しかし、可視化サーバ上のプログラムは演算サーバからのデータ受信に加えて可視化処理、および対話的操作への対応を兼ね備える必要がある。本稿では、OpenGL のユーティリティライブラリである GLUT[11] を用いて、図 2 に示すプログラムを作成する。GLUT では、マウスやキーボードなどの操作のイベントに対してコールバック関数を設定し、可視化処理とイベント処理を両立している。さらに、イベントが無い、つまりアイドル状態の時に呼ばれるコールバック関数も定義できる。このため、UNIX のシステムコール `select` で受信データの有無を判定する関数を作成し、アイドル時のコールバック関数として登録する。この構成により、可視化処理、イベント処理、およびデータ受信処理の 3 つの処理を行う。本実装ではアイドル状態の時にのみデータ受信処理を行うため、対話的操作が行われている期間にはデータ受信速度が遅くなる。

4. 評価実験

可視化サーバ遠隔利用システムの性能を実験により評価する。本実験では、図 1 における演算サーバ兼利用者端末として Linux PC (Intel Pentium4 2.8GHz processor, 512MB main memory) を用い、北海道大学の超高速可視化サーバ SGI Onyx 300[9] を遠隔利用する。まず、図 3 のガウス関数を 64×64 、 128×128 、あるいは 256×256 のメッシュに分割し、それらの頂点の座標および法線ベクトルを上記の PC で算出して、北海道大学の可視化サーバに転送する。ガウス関数の大きさを少しずつ変化させ、そのデータを可視化サーバに送信するという処理を繰り返す。可視化サーバでは受信したデータを順次可視化し、その画像を東北大学の PC に配信する。その結果、PC では大きさの異なるガウス関数がアニメーション表示される。また、利用者による対話的操作も可視化サーバに伝えられ、その後の可視化処理に反映される。描画ウィンドウサイズを 128×128 ピクセルから 1024×1024 ピクセルまで変化させて実験を行う。また、VizServer の機能である無圧縮、可逆圧縮、不可逆圧縮による画像転送の結果も評価する。

可視化サーバおよび PC にはギガビットイーサ (1000base-SX) のネットワークインタフェースが備わっており、スーパー SINET に光ケーブルで直結されている。ネットワーク性能を事

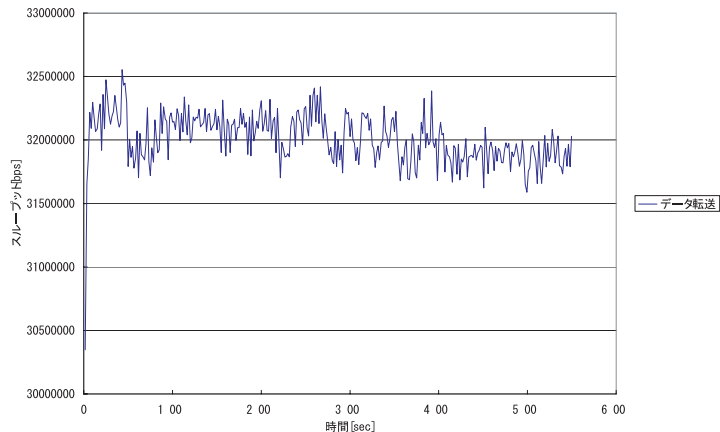


図 4. データ転送の帯域幅

前に調査した結果を表 1 に示す。実効帯域幅の測定には `ttcp`[12] を用いた。実測の結果、全体的に北海道大学から東北大学への方向の通信の帯域幅が逆方向の半分程度しか無いことが分かった。東北大学側で Linux カーネルの TCP に関連するいくつかのパラメータを変化させて計測したが、この非対称性を解消することはできなかった¹。

次に、可視化サーバ遠隔利用システムの評価実験結果について説明する。まず、本実装では対話的操作時にデータ受信速度が低下するため、操作されていない時の帯域幅を測定した結果を図 4 に示す。スーパー SINET を利用した場合、常に安定して 30Mbps 以上の帯域幅でデータを転送し続けることが可能であった。100Mbytes もの大容量データを途中結果として出力するシミュレータの場合でも、その経過をおよそ 26 秒に 1 回の間隔で転送可能である。また、以下の実験で VizServer による画像圧縮率や描画ウィンドウサイズを変化させているが、同じ経路を通る画像配信のトラフィックによる影響は見られなかった。表 1 の結果からも、帯域幅にはまだ十分な余裕があると考えられる。これらの結果から、本システムの用途に関して言えば、十分な帯域幅で安定してデータ転送できていると言える。

次に、VizServer の描画性能をフレームレートにより評価する。フレームレートの計測には、可視化サーバにインストールされている SGI Performance Co-Pilot Tools の `pmval` コマンドを利用した。VizServer は描画処理時にのみ画像を転送するため、図 3 のグラフをアイドル時に呼ばれるコールバック関数内で回転させ、その時のフレームレートを測定した。複数の実験結果を公平に比較するために、本実験では利用者による操作に依らずに回転処理を行っている。なお、利用者によるマウスやキーボードからの操作時には、本実験結果よりも高いフレームレートを達成できる条件もあり、本実験結果は必ずしもその条件下での最大フレームレートを

表 1. 各計算資源間の実効帯域幅

送信	受信	帯域幅 [Mbps]
PC(東北大)	Onyx(北大)	64.3
Onyx(北大)	PC(東北大)	23.5

¹現在、北海道大学側に設定の確認を依頼中である。

示す結果ではない。結果を図5に示す。図5において、異なるメッシュ数で行った結果が色分け表示されている。

図5から、いずれの描画ウィンドウサイズにおいても画像を無圧縮で転送している場合(128(NC)、256(NC)、512(NC)、1024(NC))にはフレームレートが低くなっていることが分かる。描画ウィンドウサイズ(総ピクセル数)すなわち無圧縮転送における画像データサイズを4倍にすると、メッシュ数に関わらずフレームレートがおおよそ1/4になっていることから、無圧縮の場合には通信の帯域幅がボトルネックになっていることが分かる。128×128ピクセルの描画ウィンドウの場合の結果を見ると、メッシュ数が128×128あるいは256×256の場合には圧縮機能を用いた場合と同等のフレームレートが出ているにも関わらず、メッシュ数が64×64の場合だけ圧縮した場合よりも低いフレームレートになっている。これは、画像転送の帯域幅がボトルネックになってフレームレートが低く抑えられたものと考えられる。達成されたフレームレートがおおよそ39fpsであることから、(総ピクセル数)×(24ビットカラー)×(フレームレート)=128×128×24×39≒15[Mbps]程度の帯域幅で画像が転送されていることが分かる。

画像を圧縮して転送する場合、メッシュ数がフレームレートを定める大きな要因になっている。描画ウィンドウが比較的小さい128×128ピクセルや256×256ピクセルの場合には、画像の圧縮率を高くしてもフレームレートが改善されないことから、画像無圧縮時とは異なって通信はボトルネックになっていないことが分かる。描画ウィンドウが大きい場合には、可逆圧縮から不可逆圧縮に変更したときにフレームレートの改善が見られる(例えば512(LL)から512(1/4))が、それ以上圧縮率を高めてもフレームレートは改善されない。このことから、描画ウィンドウが大きい場合でも、不可逆圧縮を適用している場合には通信がボトルネックになっていない。メッシュ数は可視化処理におけるピクセル数に比例しており、全処理時間の中で可視化サーバ上での描画処理時間が支配的になっているため、その描画時間の増加に伴ってフレームレートが低下しているものと考えられる。

本システムはシミュレーションの途中結果を可視化することを目的としている。利用者が対話的に見る方向を調整したとしても、毎秒数フレーム程度の描画性能で実用上十分であると考えられる。したがって、図5の結果から、本システムによる実時間可視化処理の有効性が示されたと言える。

また、本実装では描画性能が可視化サーバ上での描画速度に大きく依存していることが分かった。可視化ソフトウェアの実装上の問題、描画ハードウェアの性能、VizServerを経由することによるオーバーヘッドなどいくつかの要因が推測されるため、今後より詳細な原因調査を行って性能を改善していく予定である。

5. まとめ

本稿では、スーパー SINET を介した全国共同利用施設間の遠隔協調作業の実証実験を行った。東北大学青葉山キャンパスから、物理的には500km以上離れた北海道大学の超高速可視化サーバSGI Onyx 300に接続し、3次元グラフを十分な描画性能で遠隔操作することができた。また、演算サーバと可視化サーバとの間の通信においても安定してデータを転送し続けることができることを確認し、スーパー SINET を介した全国共同利用施設間の相互利用の可能性を示した。

現在の可視化プログラムの実装では、図2に示すとおりアイドル状態時にデータを受信するようになっているため、対話的操作時にはデータ受信速度が低下する問題がある。このため、可視化プログラムのマルチスレッド化が今後の課題として挙げられる。また、本実装では利用

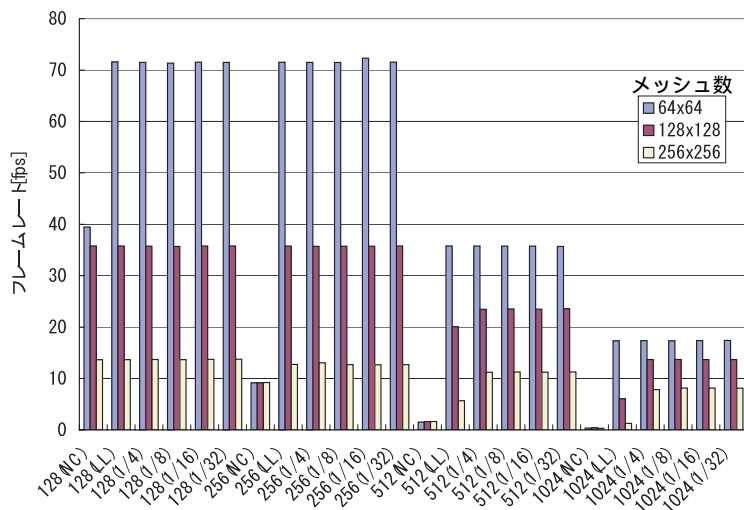


図 5. フレームレート

者端末から演算サーバ上のシミュレータを操作することが考慮されていないが、可視化結果を見ながらシミュレータのパラメータ調整やモデルの変更を行えるようにシステムを改善していく予定である。

謝辞

本研究を遂行するに当たって、数多くの方々にご協力・ご支援いただきました。なかでも、北海道大学情報基盤センターの高井昌彰教授、長尾光悦博士、および相樂恭宏氏には多大なるご協力をいただきました。また、東北大学情報シナジーセンターの首根秀昭教授、後藤英昭助教授、水木敬明助教授、伊藤英一システム管理掛長、および千葉実ネットワーク掛長にも貴重なご助言をいただきました。日本電気株式会社、日本 SGI 株式会社、および東日本電信電話株式会社からは技術提供をいただきました。あらためてここに感謝の意を表します。

本研究の一部は、総務省戦略的情報通信研究開発推進制度の研究助成金および科研費(基盤研究(B)14380132、若手研究(B)15700124)によって行われました。

参考文献

- [1] NAREGI: National Research Grid Initiative, <http://www.naregi.org/>
- [2] The DataGrid Project, <http://www.eu-datagrid.org/>
- [3] SETI@home: Search for Extraterrestrial Intelligence at home, <http://setiathome.ssl.berkeley.edu/>
- [4] Folding@home distributed computing project, <http://folding.stanford.edu/index.html>
- [5] The Access Grid Project, <http://www.accessgrid.org/>
- [6] 福田正太, “ITBL とは何か、何を指すのか,” 第 1 回 ITBL シンポジウム, 2001 年 11 月 30 日
- [7] 国立情報学研究所ニュース, 第 9 号, 2002 年 3 月
- [8] “平成 14 年度グリッド研究報告書,” 全国共同利用大型計算機センターグリッド研究推進委員会, 2003
- [9] 大宮学, “グリッド・コンピューティング・システム,” 北海道大学大型計算機センターニュース vol.34, no.4, 2002
- [10] SGI 社 OpenGL VizServer 製品ページ, <http://www.sgi.co.jp/visualization/vizserver/>
- [11] “OpenGL プログラミングガイド 第 2 版 新装版,” ピアソン・エデュケーション, 2003
- [12] Netcordia tools download page, <http://www.netcordia.com/tools/tools.html>