

jh150050-IS04

## 大規模計算結果の共同活用に向けた高精細可視化イメージ 遠隔配信システムの実証

阿部洋丈（筑波大学）

**概要** 遠隔地の高性能計算システムで出力された大規模計算の結果を、近隣の大型タイルディスプレイや、手元のディスプレイで簡便に表示させるためのシステムの実用化に向け、可視化イメージ遠隔配信システムの研究開発を以前から実施してきた。今回、JHPCN による支援を受け、大阪大学・東北大学の二拠点を中心とした実環境に配備して当システムの実用化に向けた実証を行った。また、今年度は、利便性をより向上させるため、SAGE2 対応や AutoVFlow 対応といった、システム内部の機能改善も実施した。

### 1. 共同研究に関する情報

#### (1) 共同研究を実施した拠点名

大阪大学、東北大学

#### (2) 共同研究分野

超大規模数値計算系応用分野

超大規模データ処理系応用分野

超大容量ネットワーク技術分野

超大規模情報システム関連研究分野

#### (3) 参加研究者の役割分担

阿部 洋丈（筑波大学）： 代表者

木戸 善之（大阪大学）： 副代表・ソフトウェア整備

江川 隆輔（東北大学）： 東北大拠点の整備・運用

小松 一彦（東北大学）： 東北大拠点内の整備・運用

高野 了成（産業技術総合研究所）：

アプリケーションサーバの環境整備

伊達 進（大阪大学）： 阪大拠点の整備・運用

安福 健祐（大阪大学）： 阪大拠点の整備・運用

市川 晃平（奈良先端科学技術大学院大学）：

阪大拠点の整備・運用

山中 広明（情報通信研究機構）：

ネットワークの整備・運用

河合 栄治（情報通信研究機構）：

ネットワークの整備・運用

### 2. 研究の目的と意義

高性能計算システムを利用した大規模計算は、多くの研究分野において今や無くてはならないも

のになっている。しかし、そのような高性能計算システムの数に限られているため、今後はネットワークを介した遠隔利用がより一般的になることが予想される。我々は、遠隔地で実行された大規模計算結果の迅速かつ直接的な理解を可能とするための可視化イメージ遠隔配信システム（以降、当システム）の実現を目指している。当システムは、計算機利用者側に設置された大型表示装置、とりわけ TDW (Tiled Display Wall) と呼ばれる種類の表示装置と、遠隔地の高性能計算システムを広域ネットワーク経由で相互に接続し、遠隔地で行われた大規模計算結果の可視化イメージを直接的に TDW に表示することを可能とする（図 1）。

研究代表者らは、情報通信研究機構(NICT)との共同研究プロジェクトにおいて当システム実現のための要素技術開発 (Software Defined Network (SDN) に基づいたフロー制御技術等)、およびプロトタイピングを行ってきた。今回申請した共同研究においては、それらの成果を活用する形で、実際の高性能計算機システムと TDW を用いた実証実験を実施した。

### 3. 当拠点公募型共同研究として実施した意義

当拠点公募型共同研究の枠組みを利用することで、NICT や大阪大学が中心となって従来実施されてきた、前述の共同研究の成果が当初想定していたのとほぼ同じ環境 (=広域の OpenFlow ネット



図 1 広域環境における高精細可視化イメージ遠隔配信の概念図

数の高性能計算拠点)での実証実験を実施することが可能となった。特に、東北大学拠点である東北大学サイバーサイエンスセンターと大阪大学拠点である大阪大学サイバーメディアセンターの間には、大地震発生時における津波被害予測のために双方のスーパーコンピュータを連携させて計算を行なうことを可能とするシステムが開発されており、その開発や運用を経て得られた経験を活かしたフィードバックを得られることが大いに期待できる。

#### 4. 前年度までに得られた研究成果の概要

(新規課題のため省略)

#### 5. 今年度の研究成果の詳細

今年度は、当初の申請内容である遠隔配信の実証実験に加え、これまでに開発してきたソフトウェアの改良も合わせて実施している。以下、それぞれについて述べる

##### (1) 遠隔配信の実証実験

##### (a) 東北大学-大阪大学での高精細映像ストリーミング実験

当システムの実用化に向けた第一歩として、当システムの基盤に採用されている SAGE (Scalable Adaptive Graphics Environment) による広域での高精細映像ストリーミング実験を実施した。実験には、大阪大学が実施している「組込み適塾」という社会人向け講義の場を利用し、SAGE を用いた講師映像の共有を利用して大阪大学と当報告大学で同時に実施するという形で行われた。

このとき使用されたのは、SAGE2 と呼ばれる、新しいバージョンの SAGE である。我々がこれまで実施してきたソフトウェア開発は、2004 年から開発が続いていた古いバージョンの SAGE (以降、便宜的に SAGE1 と呼ぶ) をベースに開発が行われていた。その後、2013 年に、従来の SAGE1 の設計を大幅に改め、ウェブ技術を積極的に取り込んだ SAGE2 の開発が、SAGE の開発元であるシカゴ大学イリノイ校の研究グループから発表された。SAGE2 では、映像データの転送に WebSocket を使用するように変更がされており、SAGE1 における独自プロトコルに比べ汎用性が向上している半面、広域



図 2 SAGE2 を用いた映像配信実験の様子。東北大から送られてきた映像が、大阪大の TDW 上に表示され、それを見ながらの講義が円滑に実施されている。

かつ高精細の映像配信に耐えられるのかという性能面の懸念があった。そのため、まずは SAGE2 が広域での高精細イメージの転送に耐えうるかを評価することが目的であった。

組込み適塾当日の様子を図 2 に示す。当日は、性能上の問題や安定性上の問題も特に発生せず、円滑に遠隔講義を実施することができた。なお、当日の実験の様子については 7-(4) に示した論文により詳しい情報が記載されている。

**(b) 国際会議 SC15 における実証実験**

前述の東北大学-大阪大学間での実証に加え、日米間での実証実験についても実施した。こちらの実験の目的は、昨年度までに開発した SAGE1 ベースの遠隔配信システムが、太平洋を跨いだ高遅延の環境でも動作可能であることを実証することであった。

実験の計画は以下の通りであった： 国際会議 SC15 の開催地である米国テキサス州オースティン市のオースティンコンベンションセンターにタイトルディスプレイを設置し (図 3)、大阪大学のアプリケーションサーバ、ならびに産業技術総合研究所 (茨城県つくば市) のアプリケーションサ

ーバでそれぞれ実行されるアプリケーションを同時に表示する。各拠点間の接続には、情報通信研究機構の JGN-X テストベッドネットワークをはじめとした実験用ネットワークを利用する。JGN-X は RISE という仮想 OpenFlow テストベッドサービスを提供しており、その機能を活用することで今回の SDN 実験を実施することができる。JGN-X 以外のネットワーク (たとえば、日米間の海底ケーブルによるリンク) については OpenFlow サービスが提供されていないが、その部分は L2 VLAN を利用し、OpenFlow からは通常のリンクであるように扱う。

SC15 の会場の TDW に日本からの映像を表示するという実験は、残念ながら完全には成功しなかった。我々の開発している遠隔配信システムには、複数の表示先 TDW に対し、それぞれのサイズやレイアウトに合わせて映像のサイズを変換するためのトランスコーダが実装されており、必要に応じてネットワーク内の仮想マシン上に動的に配置される仕組みになっている。その機能が、開発時の環境では動作していたが、SC15 の現地の環境では正しく動作しなかった。

トランスコーダが正しく動作しなかった原因の



図 3 国際会議 SC15 の NICT ブースに設置された TDW

1 つは、実際の複雑なネットワーク環境に起因する問題であると考えられる。ソフトウェアの開発段階では、環境セットアップの省力化のため、ソフトウェアベースの仮想スイッチである OpenVSwitch のみを用いた均質なネットワークを用いていた。それに対し、SC15 では複数のネットワークを経由した上での実験であったことに加え、ハードウェアスイッチとソフトウェアスイッチが混在する環境であった為、その途中経路上で想定外の事が起き、ソフトウェアが正しく動作しなかったものと考えられる。SC15 で利用したネットワークは時限的に構築されたものであるために、SC 終了後において、原因追求は難しいが、このような事が起こり得るということを十分念頭に置いて今後の開発を継続したい。

なお、トランスコーダの互換性の問題については、後述するソフトウェア改良に際しては十分な配慮を行い、現在はハードウェアスイッチの環境でも動作している。

## (2) ソフトウェアの改良

### (a) SAGE2 対応

昨年度までに情報通信研究機構と大阪大学が共同研究で開発した遠隔配信システムは、前述した SAGE1 に基づいて設計・構築されている。SAGE1 は、アプリケーションとディスプレイノード間の通信には TCP/IP や UDP/IP といった一般的な通信プロトコルを利用していた。それに対し、SAGE2 では、ウェブ技術の大胆な取り込みが行われ、ディスプレイノード側のソフトウェアは Google Chrome ウェブブラウザを利用することとなり、アプリケーションと Chrome の通信には WebSocket が採用された。SAGE1 と SAGE2 には互換性は考慮されておらず、異なるバージョン間で可視化イメージを送受信することはそのままでは不可能である。

当研究課題を開始した当初は、今後も引き続き SAGE1 ベースのシステムの開発を継続していく予定であった。それは、前述の共同研究が SAGE2 公開前の 2012 年に開始されており、SAGE1 ベースのソフトウェアが多く蓄積されていること、また、SAGE1 のアーキテクチャはネットワーク側からの制御と親和性が高いという理由によるものであった。また、それに加え、SAGE2 では WebSocket が性

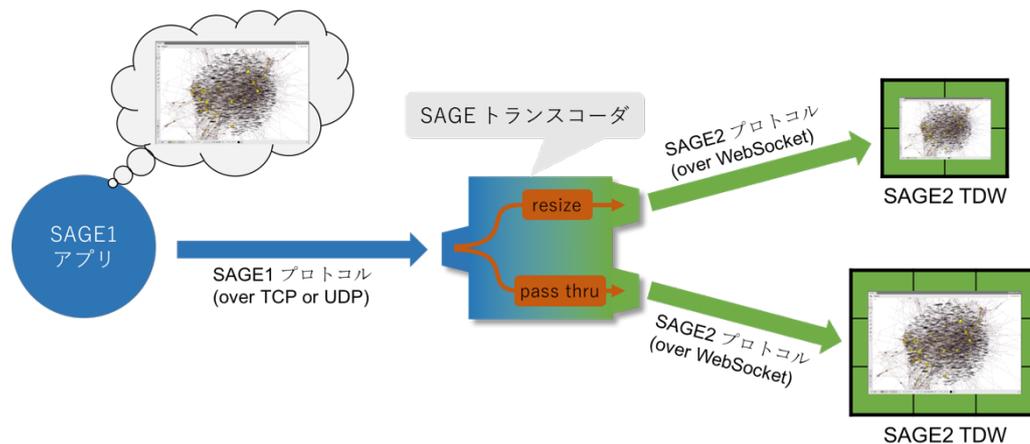


図 4 SAGE 用トランスコーダの概念図。フローのミラーリングと画像のリサイズを行う。今回、さらに SAGE1 から SAGE2 へのプロトコル変換機能が追加実装された。

能上のボトルネックとなり、同じ環境化であって SAGE1 に比べて SAGE2 のフレームレートが低下することも懸念されていた。

しかしながら、今回我々は、これまでに開発されてきたシステムの SAGE2 対応を今年度中に進めることを決めた。それは、第一に、本申請の審査時に審査員から SAGE2 に対応すべきであるという付帯コメントが付いたことを真摯に受け止めた結果である。また、第二に、前述の東北大-大阪大の SAGE2 による映像配信実験を通じて SAGE2 でも遠隔配信に耐えうるパフォーマンスを達成可能である見通しが経ったことも理由の 1 つである。

前述のように、SAGE1 と SAGE2 では通信方式が異なるため、完全な統合は難しい。そこで我々は、SAGE1 用に作成されたアプリケーションからの出力を SAGE2 のタイルとディスプレイ上に表示できるようにするためのソフトウェア開発を実施した (図 4)。具体的には、前述の SAGE1 ベースのシステム用のトランスコーダに、SAGE1 と SAGE2 のプロトコル変換をする機能を実装した。SAGE1 用のアプリケーションは、トランスコーダを SAGE1 のディスプレイノードとして認識し、TCP/IP や UDP/IP を利用して可視化イメージをそこに送信する。そして、トランスコーダで受信された可視化イメージは、内部でデコードされ、必要に応じてサイズ変換を実施した上で、さらに SAGE2 のディスプレイである Chrome へ WebSocket を介し

て送信される。

#### (b) AutoVFlow 対応によるマルチテナント化

AutoVFlow は、NICT が開発を行っている、自律分散型の OpenFlow ネットワーク仮想化のためのシステムである。Internet Protocol (IP) の成功と普及は、利用者に利便性をもたらすと共に、方式の硬直化を引き起こすという負の側面もある。そのような状況で、IP のような既存の方式にとらわれず、いわゆる clean-slate なアプローチによる研究開発の基盤として、OpenFlow ネットワークの仮想化技術は多いに期待を集めている。AutoVFlow は、従来の OpenFlow 仮想化が持っていた、集中型制御ゆえの問題を解決するために開発された新しいシステムである。AutoVFlow では、仮想化ネットワーク構築の土台となるサブネットワークがそれぞれ自律的な制御を行い、適切なパケット書き換えを行うことで仮想化が実現される。それにより、より頑健なネットワーク仮想化が実現可能である。

我々は、今年度、我々の遠隔配信システムを AutoVFlow に対応するようにソフトウェアの開発を実施した。これまでは、我々の遠隔配信システムは、すべての利用者が単一の OpenFlow ネットワークを共有していることを前提としていた。そのため、特定のリンクを一部のユーザに対してのみ

利用可能にするなど、いわゆるマルチテナント的な機能を実現するためには、その都度ごとに別の OpenFlow ネットワークを設定する必要があった。今回、遠隔配信システムが AutoVFlow に対応したことにより、利用者毎に異なる仮想 OpenFlow ネットワークを見せることが可能となり、単一のネットワーク、単一のシステムインスタンスでマルチテナント機能を実現することが可能となった。

## 6. 今年度の進捗状況と今後の展望

今年度の研究開始当初は、SC15 の会期に合わせた日米間実験の後に、図 1 に示してあるような国内の実験を実施することを予定していた。しかし前述の通り、日米間実験においてソフトウェアの不備が見つかった為、その後で予定していた国内実験を予定通り実施し完了させることができなかった。この点については大変残念に思っている。

その一方で、当初はあまり予定していなかった、ソフトウェア面の機能強化という大きな成果が得られた。特に、SAGE2 でのパフォーマンスの懸念が払拭されたことと、SAGE2 への対応が大きく進んだことは、当システムの利便性の大幅な向上に寄与するものと期待している。また、SAGE2 対応や AutoVFlow 対応と同時に、日米間実験の際に問題を生じたトランスコーダの制御プログラムの部分の改善も進んだ。それにより、同じ失敗を重ねることは無いと考えている。

当研究課題は、今年度のみで発展的解消を行い、次年度以降はより広いテーマを扱う別の研究課題（課題名：Toward a resilient software defined infrastructure to support disaster management applications. 代表：渡場康弘（奈良先端大））へ合流することを予定している。来年度は、その枠組の中で、今年度やり残した国内実験の完遂することを目指したい。

## 7. 研究成果リスト

### (1) 学術論文

- [Yoshiyuki Kido](#), [Kohei Ichikawa](#), [Susumu Date](#), Yasuhiro Watashiba, [Hirotake Abe](#), [Hiroaki Yamanaka](#), [Eiji Kawai](#), Haruo Takemura, Shinji Shimojo. SAGE-based Tiled Display Wall Enhanced with OpenFlow-based Dynamic Routing Functionality Triggered by User Interaction, Future Generation Computer Systems, vol. 56, pp. 303-314, Elsevier, Mar. 2016. [[DOI:10.1016/j.future.2015.09.033](#)]

### (2) 国際会議プロシーディングス

- [Susumu Date](#), [Hirotake Abe](#), Dashdavaa Khureltulga, Keichi Takahashi, [Yoshiyuki Kido](#), Yasuhiro Watashiba, Pongsakorn U-Chupala, [Kohei Ichikawa](#), [Hiroaki Yamanaka](#), [Eiji Kawai](#) and Shinji Shimojo. "An Empirical Study of SDN-accelerated HPC Infrastructure for Scientific Research", 2015 International Conference on Cloud Computing Research and Innovations (ICCCRI 2015), October 2015. [[DOI:10.1109/ICCCRI.2015.13](#)]

### (3) 国際会議発表

- なし

### (4) 国内会議発表

- [木戸善之](#), [瀬野恭彦](#), [西原秀明](#), “「組込み適塾」の遠隔講座拡大に向けた超臨場感メディア技術による実証実験”, ソフトウェアエンジニアリングシンポジウム 2015, pp. 216-217, Tokyo, Japan, 7-9, Sep. 2015. (poster)  
[<http://id.nii.ac.jp/1001/00144891/>]

### (5) その他（特許、プレス発表、著書等）

- なし（プレス発表を検討中）