

jh170052-MDJ

スケジューラと連動した広域データステージングに関する 検証・評価

阿部洋丈（筑波大学）

概要 様々な科学研究分野において、大規模なデータを処理・解析するための計算要求はますます増加している。しかし、セキュリティ上の制約のために、たとえ基盤センターに大規模な計算機システムが導入されていたとしてもそれらを十分に活用することができていないケースがある。我々の研究グループでは、大規模計算機システムに配備されたスケジューラシステムを基盤技術とし、ベアメタル構成技術（OpenStack Ironic）、およびシステム仮想化技術 Exp-Ether を連動させることにより、ユーザのジョブ要求に基づき GPU や SSD などの周辺機器を計算ノードに動的に割り当てた計算環境を提供するフレキシブル計算環境構成技術のプロトタイプ開発、検証を推進してきた。本研究課題では、上述の問題点に着眼し、これまでに開発してきたフレキシブル計算環境構成技術を基に、ユーザ占有型 SSD を所有する計算環境を形成し、当該 SSD に対してジョブ実行前後に計画的にユーザ所有データをステージイン/ステージアウトする機構の研究開発を行った。

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 サイバーメディアセンター

野崎一徳（大阪大学）

ニーズ分析

高橋雅彦（日本電気）

スケジューラと SDN 連携の設計・実装

(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

三澤明寛（大阪大学）

設計・実装補助

2. 研究の目的と意義

近年の計測技術、プロセッサ技術、ネットワーク技術の発展は、科学研究で扱われるデータを大規模化・多様化しつつある。そのため、様々な科学研究分野において、それらデータを処理・解析するための計算要求はますます増加している。しかし、それらの科学データにはそのデータに要求されるセキュリティ要件のために、たとえ基盤センターに大規模な計算機システムが導入されていたとしても、それらセキュリティ要件のためにそれら大規模計算機をそのデータ処理・解析の

(3) 参加研究者の役割分担

阿部洋丈（筑波大学）

代表。統括・設計・実装。

吉川隆士（大阪大学）

副代表。統括補助・設計・実装。

伊達進（大阪大学）

副代表。統括補助・設計・実装。

渡場康弘（大阪大学（申請時は奈良先端大））

スケジューラと SDN 連携の設計・実装

Chonho Lee（大阪大学）

Exp-ether との連動設計

木戸善之（大阪大学）

SDN コントローラ設計

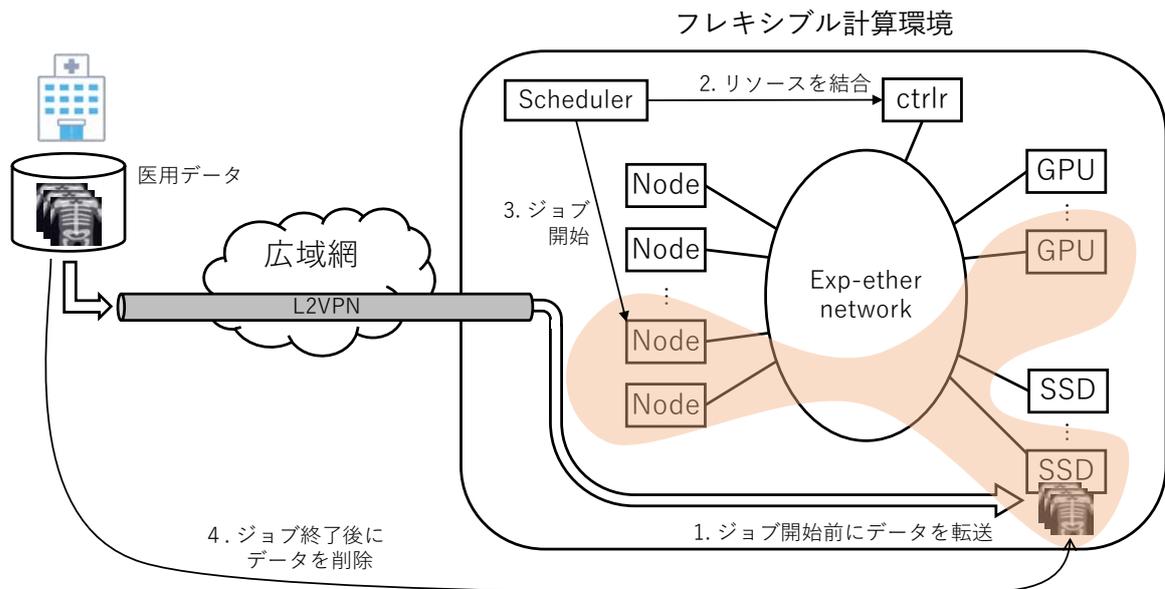


図 1. 提案システムの概要図

ために利用することができないという問題がある。

一方、われわれ研究グループでは、大規模計算機システムに配備されたスケジューラシステムを基盤技術とし、ベアメタル構成技術 OpenStack Ironic およびシステム仮想化技術 Exp-Ether を連動させることにより、ユーザのジョブ要求に基づき GPU や SSD などの周辺機器を計算ノードに動的に割り当てた計算環境を提供するフレキシブル計算環境構成技術のプロトタイプ開発、検証を推進してきた。当研究では、上述の問題点に着眼し、これまでに開発してきたフレキシブル計算環境構成技術を基に、ユーザ占有型 SSD を所有する計算環境を形成し、当該 SSD に対してジョブ実行前後に計画的にユーザ所有データをステージイン／ステージアウトする仕組みを設計・実装する(図 1)。具体的には、センター所有計算基盤とユーザの所有するデータ源との間のネットワーク接続 (network reachability) を SDN (Software Defined Networking) 技術により制御するとともに、ジョブ実行時のみにアクセス可能となるユーザ占有型 SSD に対するデータステージングにより、第三者へのデータ露出可能性

を低減した仕組みを提案する。特に、本研究では、キャンパス内での実証実験計画を拡大し、広域に分散する基盤センター間での応用可能性について検証することを目的とする。

3. 当拠点公募型共同研究として実施した意義

本研究を学際大規模情報基盤共同利用・共同研究拠点として実施する意義は、以下の 2 点より説明できると考えている

- (1) 本研究を推進する上で必須となる Exp-Ether が導入された大規模計算機システムは、大阪大学サイバーメディアセンターの大規模可視化対応 PC クラスタ以外に存在しない。そのため、当該センターの計算機システムを中核とした研究推進が必須となること。
- (2) 本研究で検討する広域分散環境化での

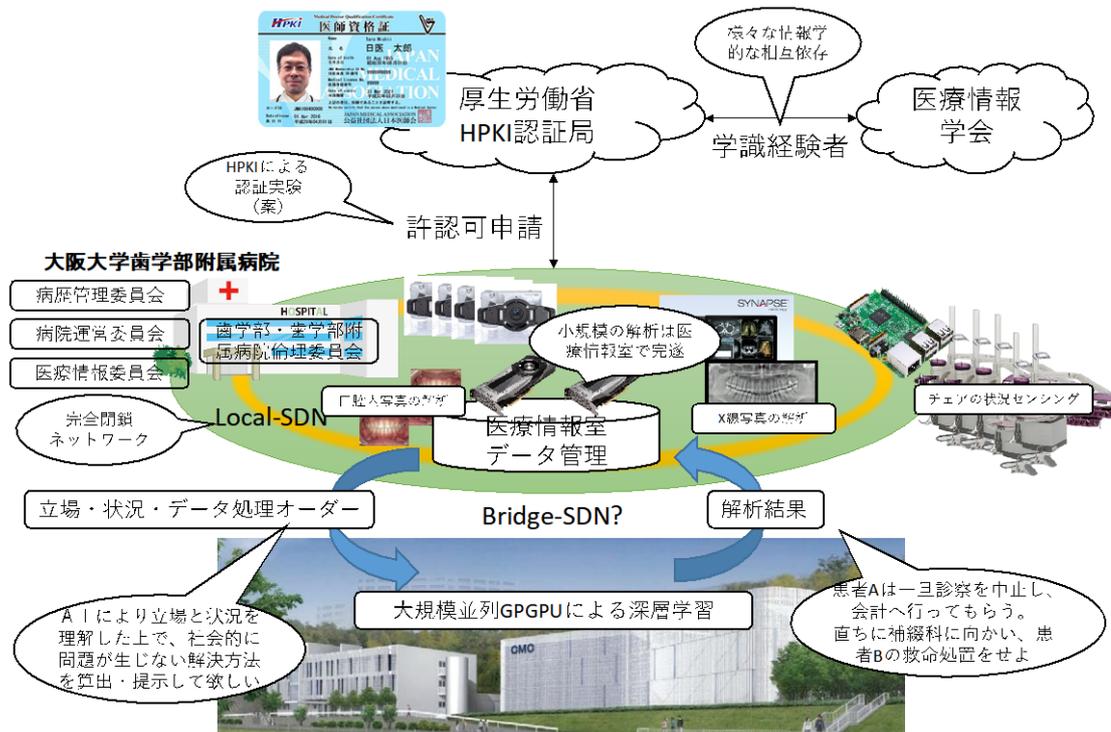


図 2. 医療情報の高性能計算環境に対するニーズの概要図

データステージングの応用可能性を検証するためには、学際大規模情報基盤共同利用・共同研究拠点で募集をしているネットワーク課題の趣旨に合致するものであるとともに、その提供する資源を利用しなければ本研究の目的達成がなされていないこと。

4. 前年度までに得られた研究成果の概要

われわれの研究グループでは、これまで SDN (Software Defined Networking) の有するネットワークプログラミング性を活用する高性能計算・可視化の実現を目的とした研究開発を推進してきた。その結果、今日までに、ユーザジョブ要求に対して大規模計算機資源のネットワーク資源、すなわちインターコネクト、および、プロセッサ資源の最適な組み合わせを提供可能なジョブ管理システム (SDN-JMS)、ネットワークフロー制御を動的に行うことで集合通信の実行時間の短縮をはかるメッセージ通信ライブラリ (SDN-MPI)、ユーザ操作により変動するネットワークト

ラフィックにシームレスに経路制御を行う遠隔可視化技術 (VisCloud)、複数の研究拠点の計算資源を集約した仮想クラスタ構成技術 (On-demand Cloud) 等のプロトタイプ化に成功している。同時に、査読付き国際会議論文、学術論文での成果発表実績も有している。

本研究は、これらの研究開発で得られた知見と経験を活用し、基盤センターの計算機システム運用における問題に着眼し、推進する研究である。本研究での基盤技術となる、ユーザのジョブ要求に基づき GPU や SSD などの周辺機器を計算ノードに動的に割り当てた計算環境を提供するフレキシブル計算環境構成技術については、すでに 2016 年 11 月に開催された国際会議・展示会 SC において大阪大学サイバーメディアセンターが出展を行う研究展示ブースにて展示実績もある。また、本研究の着想やアイデアについては、国際ワークショップ WSSP において発表実績もある。

5. 今年度の研究成果の詳細

i. ニーズ分析

当年度の研究における第一の成果は、大阪大学歯学部附属病院医療情報室の野崎一徳助教を中核として実施された、歯学部附属病院内に存在する計算ニーズ・要求の分析である。その上で、本研究の推進に際して利用可能な医用データの洗い出しを行うとともに、歯学部附属病院の歯科医師や研究者らが当該データを処理・解析する上で想定するシナリオ及び想定環境を策定した。

図 2 は、医療従事者および病院運営者側の視点から見た、我々が実現を目指しているシステムへの要望と要求を一枚の絵にまとめたものである。当システムの開発を始めた当初の想定では、大量に病院内に蓄積されている大量のデータをバッチ的に処理して研究上の新しい知見を得ることが主な用途であると想定していた。実際に医療に従事している方からは、それだけでなく、むしろそれよりも個々の患者のケースについての診断や診療方針の決定に利用したいという意見が得られた。そうすると、当システムの利用に習熟することが可能な一部の研究者だけでなく、日々の診療に忙殺される多数の医療従事者が利用し得ることを想定したシステム設計が課題となることが明らかになった。

また、今回のニーズ分析を通じて、我々がシステム設計の際に主に想定する、性質の異なる 2 種類のアプリケーションシナリオを策定した。図 2 の中にはそれらのアプリケーションが表現されている。一つは、市中の一般利用者が自身の口腔内をスマートフォンで撮影し、クラウドにアップロードすることで、人工知能による歯科医療のアドバイスを獲得することができるというものである。アップロードされた結果は医療ビッグデータとして蓄積され、アドバイス内容の更なる改善のために役立てられる。もう一つは、歯科医師が治療行為を行っている際の診察台（チェア）

の周辺動画（ただし患者は含めない）を撮影し、その動画をクラウド上で人工知能により分析することで、実施に行われた治療内容を推定し、治療記録の自動生成や歯科医師に対するアドバイスの作成などを行うというものである。これらの 2 つのアプリケーションは、歯科医療の充実という共通の目的の元ではあるものの、データの性質や結果の取り扱いに関しては異なる配慮が必要となる。我々は、この 2 つを包括的にカバーできる仕組みを目指すことで当システムに対するニーズの分析が可能であると考えている。

ii. プロトタイプ実装

当年度における第二の成果は、当研究グループで以前から研究開発を進めてきた、SDN (Software Defined Networking) 技術を活用したデータステージング制御管理モジュールの設計とプロトタイプ実装である。このモジュールは、これまでに大阪大学で開発されてきたフレキシブル計算環境構成技術の中核技術であるスケジューラシステムと連動する形で、ジョブの投入に合わせて利用者のサイトと計算機センターとの間に専用の仮想計算環境を構成し、利用者サイト側にあるデータを使って計算機センター側で大規模計算を実施する。

図 1（前掲）で示されているのは、当研究の開始当初に想定していた、医療機関外に出すことが容易でない大規模データの処理を計算機センターでバッチ的に処理する場合を想定した設計である。この設計では、計算の進行に応じてデータの転送がオンデマンドで発生してしまうと計算時間の増大につながってしまうために、計算機センター側に用意された専用のストレージに予めすべてのデータを転送し、それに対して処理を行う方法を想定していた。

その後、前述のニーズ分析の結果でも述べたように、そのような計算の他にも、データ量は多くないがオンデマンドで迅速に計算



図 3. 動態デモ展示の様子（医療情報学連合大会 大阪大学ブース）

結果が欲しいというニーズも存在することが明らかになった。そのため、当初想定していた図 1 のような使い方の他に、利用者サイトに設置されたストレージに対して Exp-Ether でアクセスし、データのみならず計算環境に対してもオンデマンドでアクセスすることで計算機センター側での計算をする方式も新たに実装した。これによる利点は二つある。一つは、個々のデータへのアクセス自体は遅くなるものの、アクセスされるデータが比較的少量であるならばトータルの処理時間の短縮が望めるということである。もう一つは、計算環境ごと利用者サイトから持って来るために、環境の違いを意識せずに利用することができ、計算機センターの環境に習熟するコストが軽減できるという点である。

今回開発したプロトタイプシステムは、大阪大学サイバーメディアセンターの大規模可視化対応 PC クラスタの一部を専有利用させて頂き、その上で動作実験を実施、問題なく動作することを確認した。動作実験には、歯学部附属病院矯正科で利用されることを想定して作成されたアプリケーションを使用した。そのアプリケーションは、深層学習の技法を用いて、患者の顔画像から所見を構成する特徴記述文を自動で生成するものである。モデルの学習には、患者の正面顔画像

とその患者に対する医師の所見をラベルとした訓練データを用いた。使用したデータは、倫理審査を通して提供された患者の顔画像と所見のみで、名前や年齢などの情報は使っていない。技術的には、Convolutional Neural Network (CNN) と Recurrent Neural Network (RNN) を組み合わせた Neural Image Caption (NIC) モデルを基盤技術として採用した。サーバ側に深層学習フレームワーク (Tensorflow, Keras など) が実行可能な Docker イメージを事前に用意し、Job スクリプトにはユーザの実行ファイル (モデルの学習など) とデータセットのパスを記載する。Job 実行時に Docker コンテナが立ち上がり、実行ファイルが起動する。このアプリケーションの設計と開発については IEEE e-Science 2017 で発表されている。

iii. 国際会議等でのデモ展示

第三の成果は、昨年 11 月に米国デンバーで開催された国際会議・展示会 SC17、および、同 11 月に大阪で開催された第 37 回医療情報学連合大会において、大阪大学サイバーメディアセンターが出展する研究展示ブースでの動態デモ展示 (図 3) の実施である。両会場に、利用者サイトと計算機センターを模した計算環境、および、SDN 環境を実現するための OpenFlow スイッチとコントローラを搬入し、来場者に対して研究内容の説明とデモ

を実施した。特に SC17 では多数のブース来場者を迎え（3 日間を通じてのブース来場者は 539 名）、その中には同様の課題に直面して困っているという研究者もおり、詳しい内容に関する質問やポジティブなフィードバックを得ることができた。

6. 今年度の進捗状況と今後の展望

前節での述べた 3 つの課題（ニーズ分析、プロトタイプ実装、およびデモ展示）については、当初予定していた以上の成果が得られたと考えている。特に、ニーズ分析によって得られた新たな知見を、プロトタイプ実装に早速盛り込むことに成功し、さらにそれを SC17 等でデモ展示できたことは想定を大きく超える成果であった。

その一方で、課題申請時に予定していたキャンパス内実験（大阪大学歯学部病院と大阪大学サイバーメディアセンター間）、および、広域環境実験（SINET L2VPN を介して大阪大学と筑波大学との間を結ぶ）は残念ながら年度内に実施することはできなかった。その原因は、ネットワーク環境の整備が間に合わなかったこと（キャンパス内、および、広域環境）、そして、Exp-Ether 関係の機器の準備が間に合わなかったこと（筑波大学側）である。これについては大いに反省すべき点であると認識している。

今回の申請における期間は既に終了しているものの、この研究自体は今後も継続して実施する予定である。大阪大学でのキャンパス内実験、および、大阪大学-筑波大学間での広域環境実験のいずれも、2018 年度中には当初の計画通り実施できるように今後も研究を継続して行きたい。また、それらの実験で得られる知見を踏まえて、当システムの更なる発展と利用推進を今後進めて行きたい。

7. 研究成果リスト

(1) 学術論文

なし

(2) 国際会議プロシーディングス

- Seiya Murata, Chonho Lee, Chihiro Tanikawa, Susumu Date, “Towards a Fully Automated Diagnostic System for Orthodontic Treatment in Dentistry”, The thirteenth IEEE eScience Conference (e-science2017), pp. 1-8, Auckland, New Zealand, Oct. 2017.

(3) 国際会議発表

- 国際会議 SC17 におけるブース展示、2017 年 11 月、デンバー、アメリカ。
- 第 37 回医療情報学連合大会におけるブース展示、2017 年 11 月、大阪。

(4) 国内会議発表

- 伊達進, 吉川隆士, 野崎一徳, 渡場康弘, 木戸善之, Lee CHONHO, 下條真司. 医療データを高性能計算機システムで利用するためのダイナミック・セキュア・ステージングシステム. 第 37 回医療情報学連合大会. 2017 年 11 月。
- 渡場康弘, 伊達進, 吉川隆士, 阿部洋丈, 野崎一徳, 木戸善之, Lee CHONHO, 下條真司. 高性能計算環境における秘匿性データ解析に向けたダイナミックセキュアステージングシステムの構築. 日本ソフトウェア科学会 第 15 回ディペンダブルシステムワークショップ (DSW 2017), 2017 年 12 月. (ポスター発表)

(5) その他（特許, プレス発表, 著書等）

なし