

スーパーコンピュータとクラウドシステムの連携による高性能計算環境の提案 Proposition of High Performance Computing with Cooperation between Supercomputer and Cloud System

後藤 佑介[†] 岩下 武史[‡]
Yusuke Gotoh Takeshi Iwashita

1. はじめに

多種多様なデータを多くのユーザがやり取りする近年の計算機環境では、データの規模や種類に応じて効率的に処理できる仕組みが望まれている [1].

代表的な計算機環境として、スーパーコンピュータ（以下、スパコン）およびクラウドシステムの2種類が挙げられる。スパコンは、高性能なプロセッサや内部データネットワークを備え、多くの場合バッチ処理形式での運用がなされ、大規模シミュレーション処理に活用されている。例えば、気象予測や経済予測といった現状のデータから未来を予測する技術 [2]、自然現象から導かれる法則を用いた自動車の安全技術 [3] 等がその活用事例として挙げられる。一方、クラウドシステムは、仮想化技術等を活用し、ユーザがより簡単にインターネットを介したソフトウェアの利用環境を設計・運用することができる。クラウドシステムは保守運用が容易である利点もあり、多くの企業で利用されている。

クラウドシステムで大規模データを処理する場合、計算量が膨大となり計算時間が長大化する恐れがある。このような場合、スパコンを併用することでその処理時間を短縮することが考えられる。しかし、従来のスパコンとクラウドシステム間の連携には実際上の問題点があり、その円滑な連携事例は必ずしも多くない。

本研究では、北海道大学情報基盤センターに新たに導入されるスパコンとクラウドシステムの特徴を活用し、両システムの連携による高性能計算環境の構築を提案する。具体的には、スパコンとクラウドシステムが地理的・ネットワーク的な意味で近接して設置される点やシステムで使用されるプロセッサアーキテクチャが統一化されることを利用した連携の仕組みを構築する。

2. 計算機環境

2.1 スーパーコンピュータ

スパコンを用いた処理では、多数の高性能プロセッサを用いた並列処理を行うことで、ユーザは大規模なシミュレーション評価にかかる処理時間を短縮でき、評価結果に基づいた技術開発の進捗を加速できる利点がある。一方で、小～中規模のデータ処理や並列効果が低い処理は処理時間の短縮効果が低くなる。また、当該のスパコンがサポートしていないプログラミング言語で構築されたプログラムを実行できない点や、並列台数効果を向上させるために専門的なノウハウが必要になる点が課題として挙げられる。

2.2 クラウドシステム

クラウドシステムは、外部に設置したクラウドサーバでデータを管理・運用することで、ユーザがデータの保存場

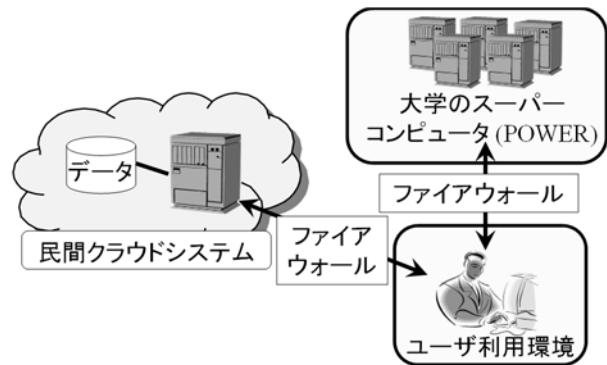


図1 大学のスパコンと民間クラウドシステムの連携

所を意識せずにアプリケーションを利用してデータを処理できるシステムであり、小～中規模のデータ処理に適している。また、企業や研究施設において複数のユーザでデータを共有したい場合、クラウドシステムを用いてデータを管理することで、情報システムの構築や運用の手間を削減できる。さらに、root 権限を用いてアプリケーションのインストールやシステムの設定変更が可能であり、拡張性が高い計算機環境を構築できる。

一方で、クラウドシステムはスパコンに比べて、並列台数効果が得にくい構成となっている場合が多い。このため、大規模なデータ処理を行う場合はデータサイズに応じて処理時間が長大化し、サービス全体の処理性能の低下、および一時的な利用資源の拡大が必要となる。

3. スーパーコンピュータとクラウドシステムの連携における問題点

クラウドシステムで大規模データを高速に処理する場合、スパコンを利用することで、計算時間を短縮できる可能性がある。本研究では、スパコンとクラウドシステムが連携することで互いの利点を生かした高性能計算機環境の構築を目指す。

現在考えられるスパコン・クラウド連携システムとして、北海道大学情報基盤センターの SR16000 [4] と民間事業者によるクラウドシステムの連携を考える。本連携システムの概要を図1に示す。本連携システムの問題点として以下の3つを挙げることができる。

(1) システムの設置場所

現在、スパコンは国内では大学や研究機関を中心に運用されており、このうちの多くは大学に設置されている。一方

[†] 岡山大学大学院自然科学研究科, Graduate School of Natural Science and Technology, Okayama University

[‡] 北海道大学情報基盤センター, Information Initiative Center, Hokkaido University

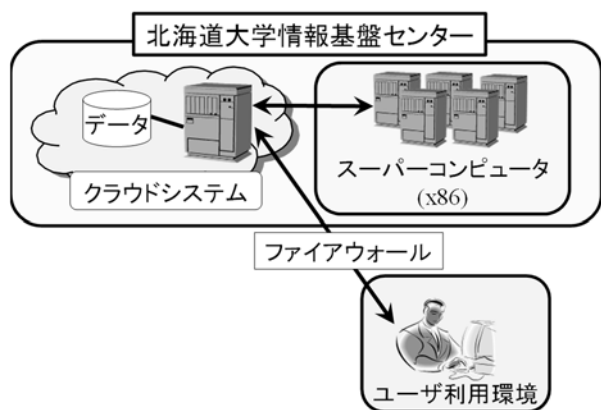


図 2 提案方式における計算機利用環境

で、クラウドシステムは、システムを運用する環境を考慮して日本のみならず世界中で設置されており、ユーザは自身にとって最適な利用環境を選択して利用する。このため、スパコンとクラウドシステムが連携する場合、両者の設置場所が物理的に大きく離れることになり、データ通信に関する処理時間が長大化する問題がある。

(2) セキュリティへの配慮

図 1 に示すように、多くの場合スパコンとクラウドシステムは異なるネットワークシステムに属しており、通信時は外部のネットワークを介する必要がある。そのため、ユーザは、それぞれのネットワークが設置しているファイアウォールを超えて、データ処理を行う必要があり、データ漏洩の危険性等のリスクが 0 であるとはいえない。

(3) プログラムの動作環境

スパコンとクラウドシステムのプロセッサアーキテクチャや OS の種類が異なっていた場合、クラウドシステムで運用しているサービスプログラムがそのままスパコンで実行できないという問題が生じる。例えば、図 1 のシステムではスパコンは POWER アーキテクチャによるプロセッサを利用しており、現在の多くの民間クラウド業者のシステムとの整合性は高くない。

4. 提案方式

4.1 概要

北海道大学情報基盤センターの次期システムを利用し、スパコンとクラウドシステムが連携した高性能計算環境の構築を提案する。提案方式における計算機利用環境を図 2 に示す。

4.2 特徴

提案方式の特徴を以下で述べる。

(1) 物理的距離が近接した計算機利用環境

提案方式では、スパコンとクラウドシステムの物理的距離が近接していることを利用する。これにより、ユーザはデータ通信時に発生する処理時間を短縮することができる。

(2) 同一ネットワーク環境による管理運用

スパコンとクラウドシステムが同じネットワーク環境で運用されることで、ユーザはそれぞれのファイアウォールを経由せずにデータをシステム間で送受信できる。これにより、データの漏洩を防ぎ、機密性を確保したデータ処理が可能となる。

(3) x86 プロセッサを搭載したスパコンによるプログラム動作環境の拡張

スパコンシステムとクラウドシステムの両者が x86 アーキテクチャによるプロセッサにより構成され、OS としても Linux が用いられるため、クラウドシステムにおける利用プログラムをそのままスパコン上で実行できることが期待される。

5. 提案する高性能計算環境の利用例

本研究で提案する高性能計算環境の利用例として、タクシーの配車システムを説明する。従来のクラウドシステムを導入したタクシー配車システム [5] では、タクシー会社とクラウドシステムがインターネットを介して接続することで、リアルタイムの配車サービスを実現している。提案方式では、スパコンとクラウドシステムが連携することで、従来のクラウドシステムを用いたリアルタイムの配車サービスに加えて、今後想定される繁忙期の渋滞予測および配車分布の見積もりをスパコンで実行する。スパコンによる実行結果に基づいてタクシーの配車分布をクラウドシステムで管理することで、タクシーの配車システムを効率的に運用できることが期待される。

6. おわりに

本研究では、スパコンとクラウドシステムが連携した高性能計算環境の構築について検討を行った。それぞれのシステムの特徴を述べた上で、両方のシステムが連携する場合の問題点を具体的な事例をもとに示した。また、これらの問題点を解決する提案方式についての特徴を述べ、具体的な利用例を挙げて提案方式の有用性を検討した。

今後は、北海道大学情報基盤センターのクラウドシステム上で提案方式のプロトタイプを作成し、連携可能な基盤の構築に向けて検討を行う [6]。また、これまでに提案してきた位置情報システムの探索手法 [7] を本プロトタイプに適用し、提案方式を想定した高性能計算環境で性能評価を行う。

謝辞

本研究は、北海道大学情報基盤センター萌芽型共同研究採択課題、ならびに学際大規模情報基盤共同利用・共同研究拠 (JHPCN) 萌芽型共同研究課題による成果である。

参考文献

- [1] 総務省, 平成 28 年度情報通信白書, <http://www.soumu.go.jp/johotsusintokei/whitepaper/h28.html>.
- [2] 気象庁気象研究所, スーパーコンピュータ, <http://www.mri-jma.go.jp/Facility/supercomputer.html>.
- [3] Onishi K. and Tsubokura M., "Optimized preprocessing of tens of billions of grids in a full-vehicle aerodynamic simulation on the K-computer", Proc. 2014 International Vehicle Aerodynamics Conference, pp.149-158 (2014).
- [4] 北海道大学情報基盤センター, スーパーコンピュータ HITACHI SR16000/M1, http://www.hpci-office.jp/pages/iic_hokudai_2014-1.
- [5] 富士通テン (株), クラウド型タクシー配車システム, <http://www.fujitsu-ten.co.jp/c-system/cloud/tokucho.html>.
- [6] 北海道大学情報基盤センター, 萌芽型共同研究, http://www.iic.hokudai.ac.jp/kyodo_kenkyu/kyodo_kenkyu.html.
- [7] Gotoh Y. and Okubo C., "A Proposition of Querying Scheme with Network Voronoi Diagram in Bichromatic Reverse k-nearest Neighbor", International Journal of Pervasive Computing and Communications (IJPC), Vol.13 Issue 1, pp.62-75 (2017).