

6C-06

# 分散ストレージシステム向け SDN 型ネットワーク制御手法の提案

和泉 諭<sup>†1</sup> Guillen Luis<sup>†1</sup> 阿部 亨<sup>†1,†2</sup> 菅沼 拓夫<sup>†1,†2</sup> 村岡 裕明<sup>†3</sup>

<sup>†1</sup> 東北大学大学院情報科学研究科 <sup>†2</sup> 東北大学サイバーサイエンスセンター

<sup>†3</sup> 東北大学電気通信研究所

## 1. はじめに

データの大容量化やクラウド技術の普及に伴い、複数のストレージに分散してデータを管理・複製する分散ストレージシステムに注目が集まっている [1]. 分散ストレージシステムでは複数のストレージにネットワーク経由でデータを伝送し、その管理や複製を行う。しかし、大量のデータを伝送する場合にはストレージの読込・書込速度やネットワーク帯域がボトルネックとなり、データ伝送に多くの時間を有する問題が発生する。

そこで本研究ではネットワーク帯域のボトルネックの解消に着目し、Software Defined Networking (SDN) 技術を用いたネットワーク制御手法を提案する。本手法はネットワーク状況やデータサイズに応じて転送する経路を SDN を用いて動的に制御することで、分散ストレージにおけるデータ伝送の効率化を実現する。本稿では、提案手法の概要と初期実験による提案手法の効果について述べる。

## 2. 関連研究

隣接する分散ストレージ間のネットワーク帯域を向上させる技術の 1 つとしてリンクアグリゲーション (Link Aggregation: LAG) がある [2]. これは複数の物理リンクを束ねて 1 つの論理リンクとして扱う技術である。複数のリンクを用いることで、あるリンクに障害が生じた場合も残りのリンクを活用することで通信を継続することが可能となり耐障害性の向上も期待できる。また、比較的安価なケーブルやスイッチなどでも高速通信が可能となる。

リンクアグリゲーションでは、MAC アドレスや IP アドレスなどを基にしたハッシュテーブルによりスイッチが利用するポートを分散するよう制御している。しかし、複数のストレージ間で通信が行われる場合、使用されるポートが偏り、スループットが 1 つのリンクの上限値となる場合があり、十分なスループットの確保が困難である課題が存在する。

## 3. 分散ストレージシステム向け SDN 型ネットワーク制御手法の提案

前節で述べた課題を解決するため、本研究では分散ストレージシステム向け SDN 型ネットワーク制御手法を提案す

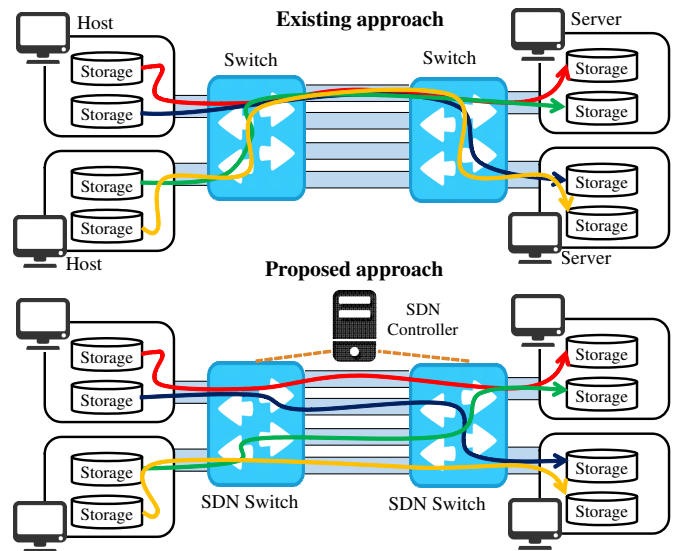


図 1: SDN 型ネットワーク制御手法の概要

る。その概要を図 1 に示す。リンクアグリゲーションなどの従来手法ではストレージ間通信において、特定のリンクにトラフィックが集中して、スイッチ間のリンクを十分に活用できず、スループットが低下する課題が存在した。本提案手法では SDN を利用してリンクの混雑状況に応じてトラフィックを制御し、各トラフィックが利用するリンクを動的に選択する。これによりスイッチ間のリンクを活用し、スループットの向上を実現できる。

我々はこれまでに SDN を用いたネットワーク経路の動的制御手法を提案してきた [3]. この動的制御手法を分散ストレージ向けに拡張することで、提案するネットワーク制御手法の設計並びに実装を行う。具体的にはスイッチにおけるポート毎の単位時間当たりのトラフィック量並びにフロー毎の単位時間当たりのトラフィック量を SDN コントローラが計測する。そして、各ポートで送信するトラフィック量の差が小さくなるようにフローが利用するポートを選択し、そのフローエントリの生成・設定を行う。これをリアルタイムで行うことで、使用されるポートの偏りをなくし、効率的なデータ伝送を実現する。

## 4. 実験

提案手法によるスループット向上の検証のために初期実験を行った。具体的には図 2 に示すように実機の SDN スイッチを 2 台用意して、各スイッチにホスト (H1, H3) とサーバ (H2, H4) を接続した。SDN スイッチ間は 100Mbps のリンクを 9 本接続し、SDN スイッチとホスト・サーバ間は 1Gbps のリンクを接続した。そして、H1 から H2 に

A Proposal of SDN-Based Network Control for Distributed Storage Systems

Satoru IZUMI<sup>†1</sup>, Guillen LUIS<sup>†1</sup>, Toru ABE<sup>†1,†2</sup>, Takuo SUGANUMA<sup>†1,†2</sup>, and Hiroaki MURAOKA<sup>†3</sup>

<sup>†1</sup> Graduate School of Information Sciences, Tohoku University

<sup>†2</sup> Cyberscience Center, Tohoku University

<sup>†3</sup> Research Institute of Electrical Communication, Tohoku University

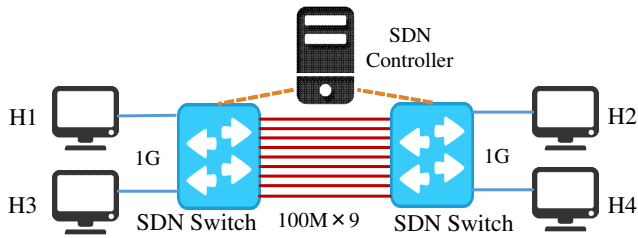


図 2: 実験環境

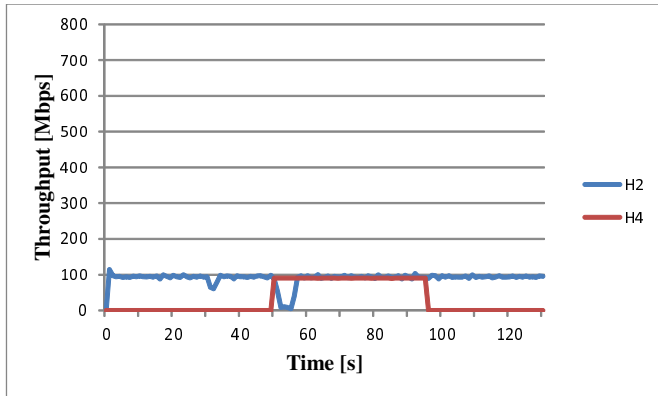


図 3: 各ホストのスループット (LAG の場合)

1Gbyte のデータを伝送するトラフィックを 8 本送信し、その 50 秒後に H3 から H4 へ 90Mbps のストリーミングデータを送信した。

この時、リンクアグリゲーションにより 8 本の物理リンクを 1 つの論理リンクと設定した場合 (LAG), SDN により各トラフィックが利用するリンクを静的に選択した場合、動的に選択した場合について、それぞれ実験を行い、各ホストのスループットを計測した。

図 3 にリンクアグリゲーション時の各ホストのスループットのグラフを示す。グラフから、特定のリンクにトラフィックが偏っているため、十分なネットワーク性能を確保できていないことが確認できる。図 4 に SDN により静的にリンクを選択した場合のスループットのグラフを示す。グラフから、トラフィックの偏りを減らし、十分なスループットを確保できていることが確認できる。しかし、他のトラフィックが発生した場合、その影響によりネットワーク性能が低下した。

図 5 に SDN により動的にリンクを選択した場合のスループットのグラフを示す。グラフから、トラフィックの偏りを減らし、十分なネットワーク性能を確保でき、他のトラフィックが発生した場合もその影響を考慮して、他の空いているリンクを利用することでスループットの低下を防ぐことができた。

以上より、提案手法によりトラフィック状況に応じて利用するリンクを動的に選択することで、スループットの向上が確認できた。従って、提案手法により分散ストレージにおけるデータ伝送の効率化が期待できる。

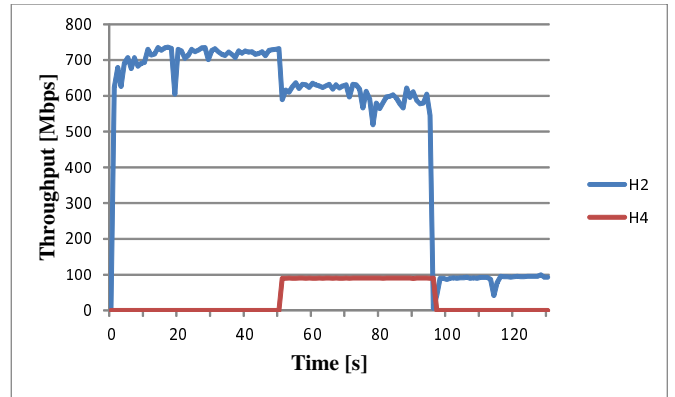


図 4: 各ホストのスループット (SDN 静的リンク選択)

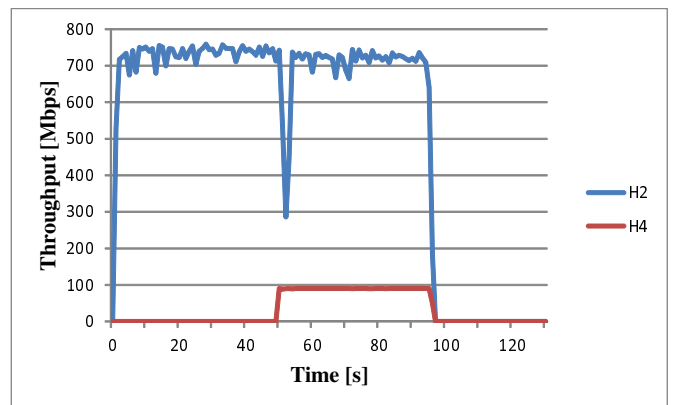


図 5: 各ホストのスループット (SDN 動的リンク選択)

## 5. おわりに

本稿では、分散ストレージシステム向け SDN 型ネットワーク制御手法を提案し、初期実験によりその効果を示した。今後は GlusterFS など分散ストレージシステムに提案手法を適用し、実機の OpenFlow スイッチを用いた実験により提案手法の有効性や実用性について検証する。

謝辞 本研究の一部は、文科省委託研究「高機能高可用性情報ストレージ基盤技術の開発」の支援を受けて実施している。

## 参考文献

- [1] Liu, Y., et al.: ProRenaTa: Proactive and Reactive Tuning to Scale a Distributed Storage System, *Proc. of the 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid2015)*, pp.453-464, 2015.
- [2] IEEE Standard for Local and metropolitan area networks—Link Aggregation, IEEE Std 802.1AX-2008, pp.1-163, 2008.
- [3] Izumi, S., et al.: Disaster-Aware Smart Routing Scheme based on Symbiotic Computing for Highly-available Information Storage Systems, *Proc. of the 15th IEEE International Conference on Cognitive Informatics & Cognitive Computing (ICCI\*CC2016)*, pp.137-142, 2016.