

# 仮想マシン移送によるデータセンタ内通信の効率化手法と通信負荷推定

野田 利一<sup>†1</sup> 堀江 光<sup>†1</sup> 河野 健二<sup>†1</sup>

## 1. はじめに

近年クラウドコンピューティング基盤(クラウド)が広く用いられており、クラウド上で稼働するサービスも普及している。クラウドを実現するための施設であるデータセンタ(DC)では仮想化技術が用いられており、1台の物理マシン(PM)上で複数の仮想マシン(VM)が稼働している。PMのメモリやCPU使用率などに応じてVMを他のPM上に移送することも可能であり、計算資源や電力の効率化を考えたVM配置手法が多く提案されている。また、クラウド上で稼働するサービスには他のVMと頻りに通信を行うものも多い。通信の負荷が特定のパスに集中すると、そのパスがボトルネックとなりパケットの損失が発生し、サービスのパフォーマンスが低下する。このような事態を防ぐためにDC内のトポロジは冗長化されており、冗長性の高さを効率良く利用するためのルーティング手法の研究が盛んに行われてきた。

しかし現状のルーティングシステムは、VMの配置によって通信の効率化が制限されてしまうという問題点がある。具体例を図1に示す。src1-dst1とsrc2-dst2がそれぞれ図1の左に示すように通信を行っており、スイッチとPMの間のパスがボトルネックとなっている。スイッチとPMの間のパスは通信を行う際に必ず通らなければいけないパスであるため、このような状況においてルーティングシステムは通信を効率化することができない。この問題を解決するためには図1の右に示すように、VM移送を用いてVMの配置を動的に変える必要がある。

VM配置手法の既存研究の中には、帯域の負荷を考慮したもの<sup>1)</sup>があるが、VM移送の判断を行う管理サーバが、VMの移送先を決定するためにパスの負荷を動的に調査する必要がある。そのため、DC内全てのスイッチに調査パケットを送る必要があり、オーバーヘッドが大きくなってしまいう問題点がある。

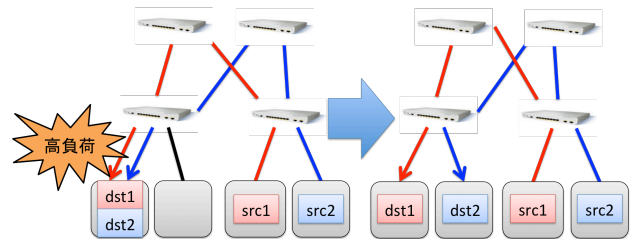


図1 VM配置によるボトルネックの発生

## 2. 提案

### 2.1 概要

本研究ではVM移送を用いてDC内の通信を効率化する手法を提案する。

パスが高負荷な状態の時にVMを適切に移送することでパスの負荷を分散し、各VMの通信スループットを向上させる。また、パスの負荷を調べるのではなく、各VMの通信情報からパスの負荷を推定することによってどのVMをどこに移送すれば良いかを判断する。これによりパスの負荷を調べるための調査パケットをスイッチへ送る必要がなくなり、帯域にかかるオーバーヘッドを削減することが可能となる。

### 2.2 アプローチ

管理サーバはトポロジ情報、各VMが使用しているパス、通信スループットが低下したVMからの通知によって移送の判断を行う。管理サーバは通信スループットが低下したという通知を受けた時に各VMが使用しているパス情報を基に、最も多く共有されているパスがボトルネックとなっていると推定し移送処理を行う。この手法は管理サーバとVMMの間だけで完結できるため、スイッチへ調査パケットを送る必要が無く、帯域にかかるオーバーヘッドを削減することが見込まれる。

## 3. 実験

提案手法がスループットを向上させることを示すためにJavaで作成したシミュレータを用いて実験を行った。

<sup>†1</sup> 慶應義塾大学

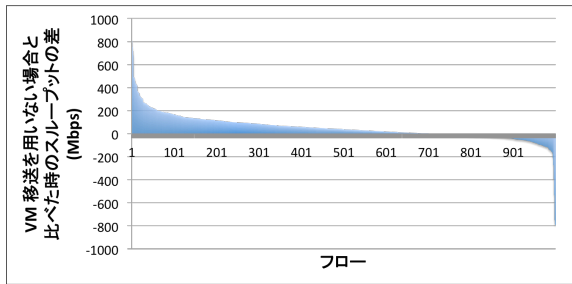


図 2 提案手法によるスループットの変化

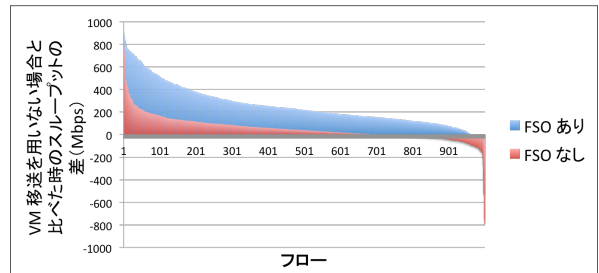


図 4 FSO を用いた場合と用いなかった場合のスループットの変化の比較

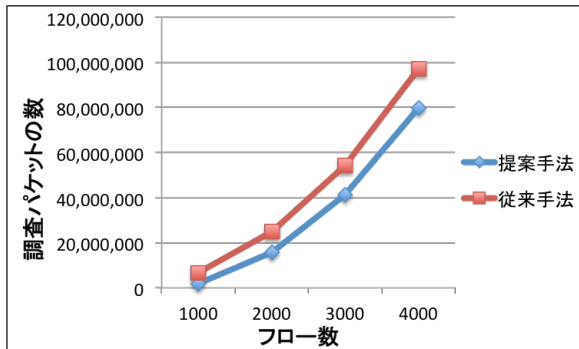


図 3 調査パケットの数

実験ではトポロジサイズ 8 の Fat-tree<sup>2)</sup> を用い、パスの帯域幅は 1 Gbps、VM のメモリ量は 100 MB とし、各フローのデータ量は 100 MB から 1 GB の間でランダムに発生させることとした。

スループットに関する評価の結果を図 2 に示す。提案手法は通信のスループットを平均 30 % 向上させた。しかし一方で、一部のフローのスループットが大幅に低下していることが分かる。これは移送によって帯域が使われることで、パスが一時的に高負荷になることが原因であると考えられる。この問題の解決策は次のセクションで述べる。次に調査パケットの数に関する評価を行った。結果を図 3 に示す。提案手法は調査パケットの数を大幅に削減できていることが分かった。

#### 4. 移送による通信スループット低下の解決策

現在、自由空間光通信 (FSO) を用いて DC 内のトポロジを動的に変える手法が提案されている<sup>3)4)</sup>。これらの手法を応用し VM 移送専用のバイパスを張ることで移送による通信スループットの低下を防ぐ。この機能をシミュレータに実装し、再度実験を行った。結果を図 4 に示す。FSO を用いて VM 移送専用のバイパスを張ることで、スループットを平均 140 % 向上させた。

## 5. まとめ

パスの負荷を推定することによって VM の配置を決定し、DC 内通信を効率化する手法を提案した。提案手法をシミュレーションによって評価した結果、提案手法は調査パケットを削減しつつスループットを平均 30 % 向上させた。さらに、FSO を用いて移送専用のバイパスを張ることでスループットを平均 140 % 向上させた。

## 参考文献

- 1) Vijay Mann, Akanksha Gupta, Partha Dutta, Anilkumar Vishnoi, Parantapa Bhattacharya, Rishabh Poddar, and Aakash Iyer. Remedy: Network-aware steady state vm management for data centers. In *Proceedings of the 11th International IFIP TC 6 Conference on Networking*, pages 190–204, 2012.
- 2) Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. A scalable, commodity data center network architecture. In *Proceedings of the ACM SIGCOMM 2008 Conference on Data Communication*, pages 63–74, 2008.
- 3) Navid Hamedazimi, Zafar Qazi, Himanshu Gupta, Vyas Sekar, Samir R. Das, Jon P. Longtin, Himanshu Shah, and Ashish Tanwer. Firefly: A reconfigurable wireless data center fabric using free-space optics. In *Proceedings of the 2014 ACM Conference on SIGCOMM, SIGCOMM '14*, pages 319–330, 2014.
- 4) I. Fujiwara, M. Koibuchi, T. Ozaki, H. Matsutani, and H. Casanova. Augmenting low-latency hpc network with free-space optical links. In *IEEE 21st International Symposium on High Performance Computer Architecture (HPCA)*, pages 390–401, 2015.