

データセンタ間ライブマイグレーションにおける ネットワーク遅延削減方式の比較検討

寺本 泰大† 岸 寿春† 永瀨 幸雄† 小山 高明† 北爪 秀雄†

†日本電信電話株式会社 NTT セキュアプラットフォーム研究所

1 はじめに

近年、災害時の事業継続やネットワーク負荷分散などのために、データセンタ(DC)間を跨って仮想マシン(VM)を移動を行うライブマイグレーションが利用されている。テナント間の接続が論理L2NWによって接続されている場合、DC間マイグレーション後に、VMとデフォルトゲートウェイが異なるDCに存在するために外部への通信がDC間を跨った冗長経路になる“トロンボーン現象(Traffic Trombone)”が発生してしまう。例えば、図1のように、距離の離れた拠点Xから拠点YにVMをライブマイグレーションした際、当該VMのデフォルトゲートウェイはマイグレーション後もRxであるため、近隣に存在するUyとの通信が経路2を経由し冗長な経路となり、遅延が発生してしまう。

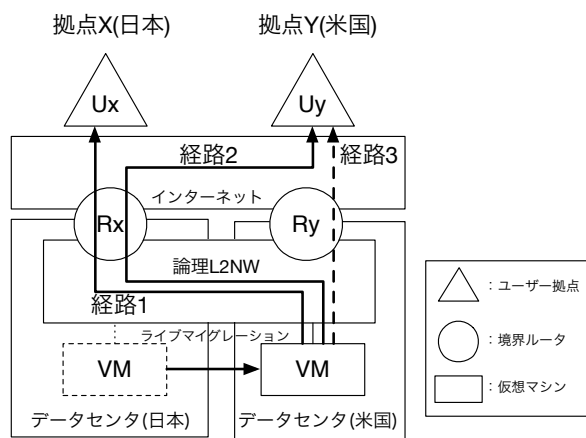


図1: トロンボーン現象の例

我々は、トロンボーン現象を解決するために、各拠点にゲートウェイ(拠点YにおけるRy)を設置し、マイグレーションと同時にVMのゲートウェイを最寄りものに切り替える事によって、経路を動的に変化させる(経路2→経路3)という方式の提案を行ってきた[1][2]。本稿では、ゲートウェイ切り替え技術として新たにDHCP方式を提案し、各方式の有効性および問題点について比較評価を行う。

なお、外部からVMへの通信時に発生するトロンボ-

ーン現象の解消については、グローバルIP網でのルーティングの経路切り替え等の技術が必要となり、外部への通信のための技術と実現方法が大きく異なるため、本稿では扱わない。また、OpenFlow等の特定の経路制御技術に依存した方式についても本稿では用いないものとする。

2 方式の比較

本章では、それぞれのデフォルトゲートウェイ切り替え技術の方式について有効性および問題点について述べる。

2.1 VM制御方式

VMのマイグレーション時に、クラウドコントローラーが当該VMに対して、拠点移動のイベントを通知する事によって、VMがデフォルトゲートウェイ情報を書き換える。

この方式では、VMの実装次第では、既存のセッションを維持したままデフォルトGWを切り替える事が可能である。しかしながら、VMが通知を受け取りデフォルトGW情報を書き換える必要があり、利用者がVMの構築を行うようなサービスでは提供が困難であるという問題点が存在する。

2.2 ARP方式

ARP方式では、ゲートウェイが同一拠点からのVMのARPメッセージのみに反応する事で、複数ゲートウェイに同一IPを持たせる。また仮想マシンのマイグレーション時にはGARPメッセージを送信し、仮想マシンのMACアドレステーブルをクリアさせる事で、宛先ゲートウェイの情報の書き換えを行う。

しかしながら当方式は、ルータープログラムの改変が必要になるため、一般的なハードウェアルーターでの実装が困難であるという問題点が存在する。ソフトウェアルーターについてはvArmour社仮想L3FW製品[4]などの実装が存在する。また、本方式はVMごとに保持しているルーターのIPアドレスに対応するMACアドレスが異なるため、ネットワークの問題発生時に

†Ysuhiro Teramoto †Yukio Nagafuchi †Toshiharu Kishi †Takaaki Koyama †Hideo Kitazume
†NTT Secure Platform Laboratories, NTT

原因の究明を困難にする可能性があるという問題点が存在する。

2.3 DHCP 方式

上記2方式以外の実現方法として、DHCPを利用してゲートウェイIPを切り替える方法が考えられる。

当方式では、DHCPサーバーがクラウドコントローラーと連携し、マイグレーションと同時にVMに対してDHCP Reconfigure[3]メッセージを送信、もしくはlease timeを短かく設定する事で、VMのゲートウェイ情報を更新させる。

DHCP方式は、前述のVM制御方式と比較すると、VMにDHCPクライアントを除く特別なソフトウェアをインストールする必要がなく、ARP方式と比較すると、DHCPサーバーを実装もしくは変更するのみで対応する事が出来るため、実装の容易が容易であるという点で優れている。

しかしながら、現在広く利用されているオープンソースのDHCPクライアントでは、Reconfigureメッセージに対応しているものが少なく、またlease timeを短かくして再割当する方法は、DHCP情報の更改時に一時的にIP割り当て情報が削除されてしまうため、現在の通信セッションが維持出来ないという問題点が存在する。

3 考察

第2章の各方式の特徴をまとめると、表1のようになる。L2レイヤ(MACアドレス)を用いて切り替え制御を行うARP方式が、セッションを維持しつつゲートウェイIP等の構成情報も変化しないため機能面では最も優れているが、ハードウェアルータを用いる場合はルータOSを改変出来ないなど、実装の困難性がある事が分かる。一方、VM制御方式およびDHCP方式ではゲートウェイIP等の構成情報が変化するものの、制御箇所がARP方式と比較し、実装が比較的容易であるという事が分かる。各方式の切り替え速度などの性能については、今後実装し評価を行う。

表1: 各方式の比較

方式	制御箇所	レイヤ	GWIP	セッション
VM制御	VM	L3	変化	実装依存
ARP	ルータ	L2	不変	維持
DHCP	DHCPサーバ	L3	変化	切断

4 まとめ

本稿では、トロンボーン現象の解決手段として提案されている、VM制御方式とARP方式について調査し、また同様の仕組みを実現するDHCP方式の提案を行なった。

さらにそれらの方式について、制御箇所・構成情報の変化・セッション維持の可否等の機能面の観点で比較を行った。

今後は、VM制御方式およびDHCP方式の実装を行い、ARP方式の既製品も含め、ゲートウェイ切り替えまでの時間や通信断時間などの、性能面について調査および比較を行う予定である。また、本稿では対象に含めなかった、ユーザ拠点からVMへの経路のトロンボーン現象解決技術についても今後評価を行っていく必要がある。

参考文献

- [1] 永淵 幸雄, 寺本 泰大, 小山 高明, 北爪 秀雄, “データセンタ間ライブマイグレーションにおける冗長経路回避に向けた経路制御方式の提案”, 信学技報, IN2013-48, pp.71-76, Jul. 2013.
- [2] 小山 高明, 岸 寿春, 井上 朋子, 永淵 幸雄, 北爪 秀雄, “複数アクティブ型仮想ルータ, 仮想L3FWの効果に関する一報告”, 信学技報, IN2013-89, pp.13-18, Nov. 2013.
- [3] Y. T’Joens, “Rfc3202: DHCP reconfigure extension”
- [4] vArmour Networks website: <http://varmour.com>