

フォグコンピューティング環境を考慮したIaaSクラウド基盤ソフトウェアにおけるVMマイグレーションに関する研究

M2015SE009 坂本光司

指導教員：野呂昌満

1 はじめに

現在、コンピュータネットワーク経由で様々なサービスを提供するクラウドコンピューティング(クラウド)が活発に使用されている。さらに、低いレイテンシが要求されるコンピューティングサービス(サービス)に対応するためにクラウドよりもユーザに近い場所でサービスを提供するフォグコンピューティング(フォグ)といったコンピュータの利用方法が提案されている[1]。フォグとはクラスタ化した物理ホスト群を指し、クラウドとユーザの間を中継することで、クラウドに比べてネットワーク距離的によりユーザに近い場所でサービスを提供できる。フォグは、Internet of Things(IoT)などの、低いレイテンシが要求されるようなサービスで有効であると考えられている。クラウドまたはフォグはそれぞれ独立して仮想マシン(VM)を管理することもできるが、IaaS基盤ソフトウェアによってクラウドとフォグのVMを一元管理すれば、ユーザの要求するサービスに応じて柔軟にVMをホストしてユーザに提供することができると考えられる。このようなシステムを考えたとき、ネットワークレイテンシに起因するVMライブマイグレーションのQoS(Quality of Service)の低下が懸念される。VM上で動作しているアプリケーションの性能やネットワークレイテンシなどの情報、ユーザに指定させる必要のある自動で取得できない情報に従って適切なVMライブマイグレーション方法を選択する機能がなければVMライブマイグレーションのQoSを保証することはできない。本稿ではVMライブマイグレーションのQoSとは、VMライブマイグレーション実行時の合計時間の短さとサービスのレイテンシの低さ、ネットワーク転送による他サービスへの影響の少なさを指す。このことからクラウド、フォグのVMを一元管理するIaaS基盤ソフトウェアにおいて、VMライブマイグレーションのQoSの保証は不可欠である。

本研究ではクラウドとフォグのVMを一元管理するコンピューティングサービスで利用されるVMライブマイグレーションのQoSを保証するための予備的な情報収集を行う。具体的には、アプリケーションの動作や、ネットワークレイテンシの違いが、VMライブマイグレーションの性能にどのような影響を及ぼすかを、複数のVMライブマイグレーション方式を比較して調査する。

2 背景

本節ではクラウドやフォグ、IaaSクラウド基盤の基本的な構成と、新たな着想であるクラウド及びフォグ複合型基盤ソフトウェア、そしてこれらのシステムに使用されるVMライブマイグレーションについて述べる。

2.1 クラウドコンピューティング

クラウドコンピューティングとはIaaSクラウド基盤ソフトウェアによって集約された物理ホスト上にVMを作成し、サービスをホストすることでユーザへサービスを提供するコンピュータの利用体系である。ユーザはクラウドへアクセスするだけでどこでも同じサービスを利用することができる。しかし、ネットワーク距離が遠い物理ホスト上のVMと通信を行うので、低いレイテンシが必要なサービスには向かない。

2.2 フォグコンピューティング

クラウドでは想定されていなかったサービスに対応するため、クラウドとユーザの間のよりネットワーク的にユーザに近い場所へ物理ホスト群を用意し、ユーザへサービスを提供する技術をフォグコンピューティング[1]と呼ぶ。フォグはユーザから送られる情報をフォグで処理してから、クラウドへ送信することでWANの通信量の減少させることができる。また、フォグとユーザがネットワーク距離的に近いことから、コンピュータゲームのような低レイテンシな通信が求められるアプリケーションを、外出先でも自宅と同じ環境で利用できるサービスの構築が期待される。

2.3 IaaSクラウド基盤ソフトウェア

IaaSクラウド基盤ソフトウェアとはクラウド処理の基盤を提供するソフトウェアである。多数の物理ホスト群を統括しており、物理ホスト上でVMを作成、管理し、必要に応じて物理ホスト間をライブマイグレートさせる。実際に利用されているIaaSクラウド基盤ソフトウェアとして、商用ではAmazon EC2[3]、オープンソースソフトウェアではCloudStack[4]等がある。

2.4 クラウド及びフォグ複合型基盤ソフトウェア

IaaSクラウド基盤ソフトウェアは元々クラウド処理の基盤を提供するものであるが、フォグにもこの技術を利用しようという研究がある[2, 8]。これによりフォグはクラウドと同じようにIaaS基盤ソフトウェアを活用して物理ホストを集約し、多数のVMを稼働することができるようになる。これによりクラウドとフォグのVMを一元管理することが可能となり、例えばフォグの物理ホスト群を離れるユーザのVMをクラウドへライブマイグレートすることで、サービスを継続して提供したり、逆にクラウドの物理ホスト群のVMを利用してユーザがフォグの物理ホスト群に近いなら、VMをフォグへライブマイグレートすることで、パフォーマンスの向上を図るなど、クラウドとフォグの両方のサービスに対応できる基盤ができると考えられる。

2.5 システムを実現するために必要な要素

このシステムを実現するにあたって IaaS クラウド処理基盤における VM ライブマイグレーションの QoS を向上させる必要があると考えられる。それは現状の IaaS クラウド処理基盤では、VM 上のアプリケーションの動作やそれを利用するユーザの要求、及びネットワークレイテンシによって適切に VM ライブマイグレーションを行う機能が無く、ユーザとサービス間の QoS を保証できないからである。この問題に対して、我々は動作しているアプリケーションの性質やネットワークレイテンシなどの情報を利用したり、自動的に取れない情報をユーザに指定させたりすることによって、VM ライブマイグレーションを最適化するアプローチを考えている。

2.6 VM ライブマイグレーション

物理ホスト上の VM を稼働させたまま別の物理ホストへ移動する技術である。単なる VM マイグレーションと違い、VM を稼働させたまま VM マイグレーションを行うことで、ユーザが利用するサービスを停止させずに VM を移動させられる利点がある。VM ライブマイグレーションには VM のメモリを全て移動してから実行ホストを切り替えるプレコピー型 [5] と、実行ホストを切り替えてから、メモリのコピーを始めるポストコピー型 [6] がある。

2.6.1 プレコピー型の手順

1. 移動先へメモリページをコピーする。移動元でページが変更（ダーティ）した場合、ページを再コピーする。これを再コピーした割合がダーティ率以上になるまで繰り返す。
2. 移動元にある VM を停止し、残りのダーティページをコピーし、VM の実行ホストを移動先へ切り替える。（移動元の VM を停止し、移動先で再開するまでの時間を「ダウンタイム」と呼ぶ）

プレコピー型 VM ライブマイグレーションはすべてを移動し終えないと実行ホストの切り替えを行わないので、仮想マシンのサイズが大きければ、VM ライブマイグレーションに要する時間も長くなり、実行ホストを切り替えが開始するまでの時間も増加する。また、頻繁にメモリの更新を行う VM に対してプレコピー型 VM ライブマイグレーションを行うと、その都度メモリをコピーするため、実行ホストを切り替えが開始するまでの時間が増加する。

2.6.2 ポストコピー型の手順

1. VM の実行状態の最小サブセット（CPU 状態、レジスタ、及びオプションでページング不可メモリ）を移動先へコピーし、すぐに VM の実行ホストを移動先へ切り替える。
2. 移動元の VM の残りのメモリページを移動先へプッシュする。

ポストコピー型 VM ライブマイグレーションは即座に実行ホストを切り替えるので、VM の状態によらず、ダウンタイムが一定になり、メモリページは再コピーする必要

がない。しかし、実行ホストの切り替え後、VM がまだ転送されていないページにアクセスしようとする、ページフォルトが生成される（ネットワーク障害）。ネットワーク障害は移動先で検知され、障害が発生したページに対応する移動元のメモリページへリダイレクトされる。ネットワーク障害が多いと、VM で稼働するアプリケーションのパフォーマンスが低下する。また、プレコピーは VM ライブマイグレーション中の VM が最新の状態に保持されるのに対し、ポストコピーは移動元と移動先の両方に分断される。この違いから VM ライブマイグレーション中に移動先で障害が発生すると、プレコピーでは VM をリカバリできるが、ポストコピーではできない。

3 本研究の目的

本研究の目的は VM ライブマイグレーションの QoS を保証するという観点から、プレコピー、ポストコピーの比較実験を行うことである。具体的には以下の環境の違いにおける各 VM ライブマイグレーションの比較を行う。

3.1 アプリケーションの動作

アプリケーションによるメモリの書き換えの頻度の違いによって、適切なライブマイグレーション方式が変わると考えられる。例えばプレコピーはメモリページの再コピーが必要なために、メモリ書き換えが頻繁な VM のライブマイグレートには向かないが、移動先の物理ホストが故障した際のリカバリの点でポストコピーに勝る。このことからメモリ書き換えが少ないアプリケーションを稼働させた VM をライブマイグレートさせる際はプレコピーが有効な可能性があり、調査する必要がある。

3.2 ネットワークのレイテンシ

ネットワークレイテンシが増加すると VM ライブマイグレーション全体の時間も増加してしまうことが考えられる。ユーザが迅速な VM の追尾を求める場合、どちらの VM ライブマイグレーション方式が有効であるか調査する必要がある。

4 実験

この節では実験の内容や環境、実験によって得られた結果とそれをふまえた考察を述べる。

4.1 実験内容

ネットワークで繋がった物理ホストを 2 つ用意し、一方で VM を作成し、もう一方へライブマイグレートさせる。計測するデータ、及び変化させる条件は以下の通りである。

- 計測するデータ
 - ダウンタイム、合計時間、セットアップ時間、メモリ転送量
- 変化させる条件
 - VM ライブマイグレーション方式、アプリケーションの動作、ネットワークレイテンシ

表 1 物理ホストの実験環境

ホスト OS	Ubuntu Server 16.04
RAM	8GB
CPU コア数	8
ホスト間のネットワーク	1Gb/s イーサネット

表 2 VM の実験環境

ゲスト OS	CentOS7
RAM	1GB
CPU コア数	1
ストレージ	8GB

4.2 実験環境

物理ホストの仕様を表 1 に、VM の仕様を表 2 に示す。ネットワーク遅延は物理ホストで tc コマンド (Traffic Control) を実行し、擬似的なネットワーク遅延を発生させる。VM 内で稼働させる負荷プログラムはおよそ 614MB のメモリを確保し、その内容を書き換え続ける。

4.3 実験結果

実験によって得られた各 VM ライブマイグレーション方式毎の平均ダウンタイムを図 1、平均合計時間を図 2、平均セットアップ時間を図 3、平均メモリ転送量を図 4 に示す。ダウンタイムはポストコピーのほうがプレコピーよりも短い。これはメモリの書き換えが起こらない環境だとしても、VM ライブマイグレーションが開始されると同時に実行ホストが切り替わるからである。平均合計時間はポストコピーがより短かった。しかし、ポストコピーの合計時間は、実行ホストが切り替わった後の更新されたメモリページを、移動元の物理ホストから取得する時間を含めるとより長くなるはずである。平均メモリ転送量はポストコピーのほうがプレコピーより少なかった。メモリ転送量は「ポストコピーのマイグレーションにともなうネットワーク転送量は、更新ページの再送が不要なためプレコピー型よりも短い。ほぼ VM のメモリサイズと同等である。マイグレーションにともなうネットワーク資源消費量が少ない [7]」とあり、VM ライブマイグレーション方式によらず、少なくとも RAM1GB のメモリ転送が起こる。しかし、実験で測定した平均メモリ転送量は最大約 207MB しかない。つまり、図 4 の値はライブマイグレーション全体のメモリ転送量でなく、RAM1GB 分を差し引いたメモリ転送量と考えられる。そして、参考文献にあるようにメモリページを再帰的に移動先の物理ホストへ送る必要のあるプレコピーの方がメモリ転送量が多かった。平均セットアップ時間はほぼ均一であった。

4.4 考察

実験の結果、メモリ消費状態やネットワーク遅延によらずポストコピーの方が、ダウンタイムを短縮できると分かった。それはポストコピーは VM ライブマイグレーションが開始されたと同時に実行ホストを切り替えられ、VM でアプリケーションが稼働していなくてもプレコピーが実行ホスト切り替え開始までの時間の短さでポストコピーに勝ることがないからである。しかし、合計時間や

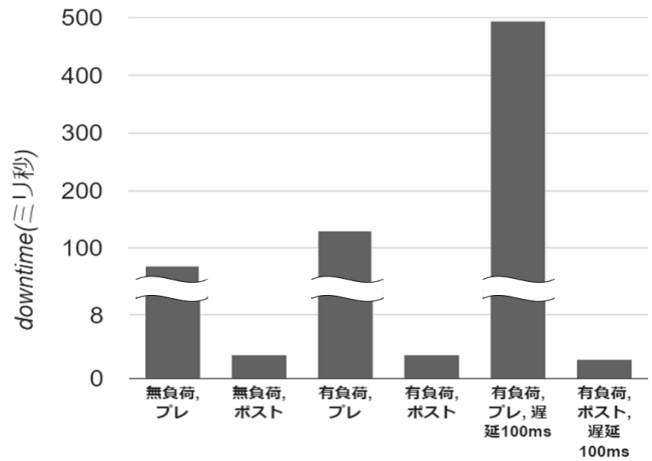


図 1 各 VM ライブマイグレーション方式毎の平均ダウンタイム

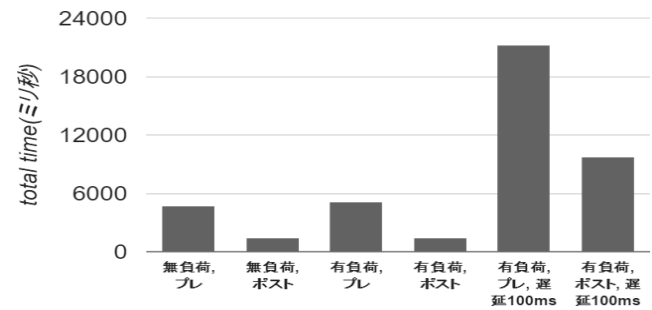


図 2 各 VM ライブマイグレーション方式毎の平均合計時間

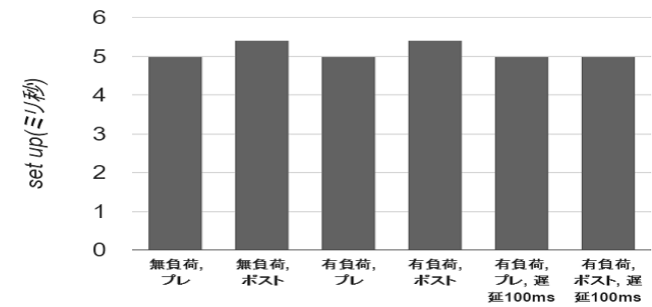


図 3 各 VM ライブマイグレーション方式毎の平均セットアップ時間

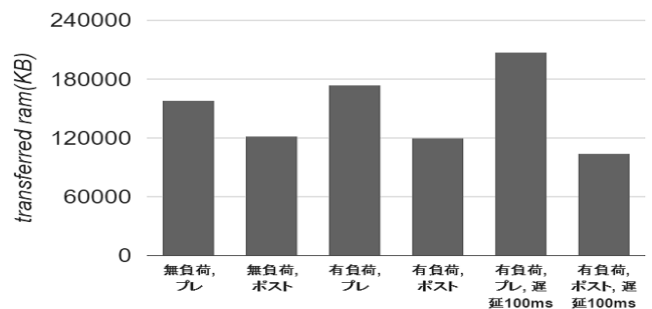


図 4 各 VM ライブマイグレーション方式毎の平均メモリ転送量

メモリ転送量の値に疑問が残るので、今後追加の実験による確認が必要である。

5 関連研究

Zhang らはモバイルユーザがクラウドコンピューティングサービスを使用する際に問題となる高レイテンシや帯域幅の低さを解決する「CloudAP」の導入を提案する過程で、3ステップVMライブマイグレーションを開発している [8]。これはVMライブマイグレーションをプレコピー、フリーズアンドコピー、ポストコピーの3ステップに分けて行うことで、QoSの低下を軽減することに成功している。本研究に適用すれば純粋なポストコピーの欠点であるメモリコピー中のパフォーマンスの低下を防ぐことができると考えられるが、この研究はモバイルアプリケーションのQoSを考慮するものであり、クラウドとフォグのVMがIaaS基盤ソフトウェアによって一元管理されたサービスにおけるQoSを考慮する本研究とは方向性が異なる。

Jin らはプレコピー型のVMライブマイグレーションをベースに、メモリ圧縮を利用したVMライブマイグレーション機構「MECOM」の設計と実装を行う過程で、メモリ圧縮を使用することでVMライブマイグレーションを高速かつ安全に実現することに成功している [9]。本研究に適用すれば純粋なプレコピーに比べ、高速なVMライブマイグレーションを実現できると考えられるが、この研究はデータセンタでの運用が主な関心事であり、クラウドやフォグにおけるQoSを想定しているとは言えない。

6 おわりに

本稿ではクラウドとフォグのVMを一元管理するIaaS基盤ソフトウェアで利用されるVMライブマイグレーションのQoSを保証するための予備的な情報収集を行った。具体的には、プレコピー、ポストコピーにおける、動作しているアプリケーションの種類とマイグレーション時間の関係について調べるための実験を行った。今回の比較実験によって、VMの状態に関わらず実行ホストの切り替えを始めに行うポストコピーの方が、クラウドとフォグの物理ホスト間、及びフォグの物理ホスト間で素早くユーザの移動を追尾することに適していることが分かった。しかし、実行ホストが素早くユーザを追尾しても、移動元の物理ホストに残されたメモリページへアクセスした際の遅延によって、サービスのQoSがどれ程低下するか、今回の実験だけでは把握しきれない。つまり、今回の実験では、ポストコピー方式でVMライブマイグレーションをすれば、確実にVMライブマイグレーションのQoS保証に繋がるとは断言できないが、VMの実行状態の最小サブセットを素早くユーザに追尾させ、サービスへアクセスする時間を短縮できることが分かった。

参考文献

- [1] Flavio Bonomi, Rodolfo Milito, Jiang Zhu, and Sateesh Addepalli, “Fog computing and its role in the internet of things”, in the first edition of the MCC workshop on Mobile cloud computing pp.13-16, 2012.
- [2] Mahadev Satyanarayanan, Paramvir Bahl, Ramon Caceres, and Nigel Davies, “The Case for VM-Based Cloudlets in Mobile Computing”, in IEEE Pervasive Computing, Volume: 8, Issue: 4, 2009.
- [3] Amazon Web Services: “Amazon EC2”, <http://aws.amazon.com/jp/ec2/>, 2017/1/17 にアクセス。
- [4] The Apache Software Foundation: “Apache CloudStack: Open Source Cloud Computing”, <http://cloudstack.apache.org/>, 2017/1/17 にアクセス。
- [5] Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen, Eric Jul, Christian Limpach, Ian Pratt, Andrew Warfield, “Live migration of virtual machines”, in NSDI'05 Proceedings of the 2nd Conference on Symposium on Networked Systems Design & Implementation - Volume 2, Pages 273-286, May 02 - 04, 2005.
- [6] Michael R. Hines, Kartik Gopalan, “Post-copy based live virtual machine migration using adaptive pre-paging and dynamic self-ballooning”, in VEE '09 Proceedings of the 2009 ACM SIGPLAN/SIGOPS international conference on Virtual execution environments Pages 51-60, March 11 - 13, 2009.
- [7] 広瀬 崇宏, 中田 秀基, 伊藤 智, 関口 智嗣, “高速マイグレーションを利用した仮想マシン配置最適化システムの検討”, 研究報告システムソフトウェアとオペレーティング・システム (OS), 2010-OS-115 巻, 14号, pp1-13, 2010.
- [8] Yunkai Zhang, Renyu Yang, Tianyu Wo, Chunming Hu, Junbin Kang, and Lei Cui, “CloudAP: Improving the QoS of Mobile Applications with Efficient VM Migration”, in 2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing (HPCC_EUC), 2013.
- [9] Hai Jin, Li Deng, Song Wu, Xuanhua Shi, Xiaodong Pan, “Live virtual machine migration with adaptive, memory compression”, in 2009 IEEE International Conference on Cluster Computing and Workshops, 2009. CLUSTER '09, 2009.