

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年 4月 30日現在

機関番号：82626

研究種目：若手研究（B）

研究期間：2011～2012

課題番号：23700048

研究課題名（和文） 大規模仮想計算機群の遠隔移動に関する研究

研究課題名（英文） A Study on Wide-area Migration of Large-Scale Virtual Machine Clusters

研究代表者

広瀬 崇宏（HIROFUCHI TAKAHIRO）

独立行政法人産業技術総合研究所・情報技術研究部門・主任研究員

研究者番号：20462864

研究成果の概要（和文）：大規模な仮想計算機群を効率的に遠隔移動する手法を開発した。発電量が不安定な再生可能エネルギーでデータセンタを駆動するためには、仮想計算機のライブマイグレーション機能を利用して、発電量に応じて一部の計算処理を中断することなく他のデータセンタに移動する技術が必要である。しかし、現状では遠隔拠点へ仮想計算機を移動する技術が確立されているとは言い難い。本研究では、我々が先に開発した単一仮想計算機の遠隔移動技術を発展させて、複数の仮想計算機を効率的に遠隔移動できる技術を多角的に開発した。メモリページの重複性やメモリアクセスの局所性を利用して、データ転送量と移動時間を削減し性能低下を抑制する手法を開発した。仮想計算機群の遠隔移動を透過的に可能にするネットワーク技術を開発した。

研究成果の概要（英文）：This project has studied wide-area live migration of large-scale virtual machine clusters. In order to drive datacenters with renewable energy, it is necessary to perform load balancing among datacenters in response to fluctuating production of electricity. However, wide-area load balancing technology using virtual machines has not yet been established. In this project, we have developed virtualization mechanisms to efficiently migrate multiple virtual machines over wide-area network, extending our preliminary work (i.e., a migration mechanism designed for a single virtual machine). It successfully reduced live migration times and performance degradation by exploiting redundancy of migration data and locality of memory access. We have also developed a networking mechanism providing transparent network connectivity for migrated virtual machines.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
交付決定額	1,500,000	450,000	1,950,000

研究分野：総合領域

科研費の分科・細目：情報学・ソフトウェア

キーワード：並列処理・分散処理、仮想化、仮想計算機、マイグレーション

1. 研究開始当初の背景

仮想計算機ライブマイグレーションを利用したデータセンタ間での計算処理移動は概

念的に提唱されてはいるものの、未だに学術界および産業界において十分な研究が進んでいなかった。特に大量の仮想計算機を遠隔データセンタ間で移動する技術が未確立で

あった。いずれの研究でも、将来の大容量ネットワークに着目し、仮想計算機のライブマイグレーションをデータセンタ間での計算処理移動手法として位置付けている。しかし、既存のライブマイグレーション機能は LAN 環境を前提として設計されており、大量の仮想計算機を遠隔拠点に移動するためには、解決すべき課題が存在した。

2. 研究の目的

本研究では、複数の仮想計算機を遠隔拠点へ効率的に移動するために、ライブマイグレーションの機能向上手法について多角的に研究を行う。メモリページの重複性やメモリアクセスの局所性を利用した、仮想計算機群の実行状態移送手法を開発する。

3. 研究の方法

最初に既存のライブマイグレーション機構を用いて予備的な評価実験を行い、WAN 環境における課題を明らかにした。その結果、多数の仮想計算機を同時に遠隔移動する際には、データ転送量の増大、仮想計算機の性能低下という問題が存在することがわかった。また遠隔移動の前後において仮想計算機のネットワーク到達性を維持することも難しいことがわかった。そこで、仮想計算機の実行状態データの重複性に着目したデータ転送量削減手法、ゲスト OS の実行とデータの転送を非同期化することで性能低下を抑制する手法、複数の仮想計算機の遠隔移動を支えるネットワーク構築手法について研究を進めた。それぞれ提案機構のプロトタイプを実装し評価実験によりその有効性を明らかにした。

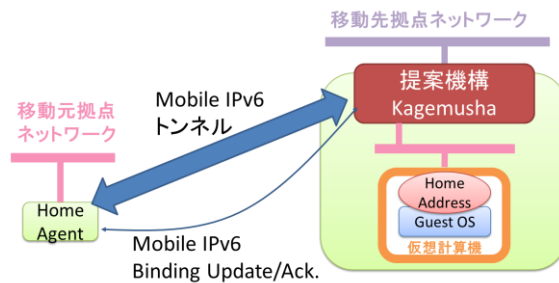
4. 研究成果

第一に、複数の仮想マシンを同時に遠隔移動する際の課題を洗い出すべく、日米間のネットワークを用いて既存移動技術の基礎的な性能評価を行った。高遅延のネットワークにおいては、同時に実行する遠隔移動セッション数を動的に最適化することで、仮想計算機群全体の移動時間を短縮できることがわかった。また移動機構自体の抜本的な改良が必要であることもわかった。

第二に、仮想計算機を遠隔移動する際に、データ転送量を削減する手法について研究を進めた。仮想計算機の動作を解析した結果、ワークロードを実行中であっても、更新されないメモリページが相当数存在することが

わかった。そこで仮想計算機のメモリページをあらかじめキャッシュしておくことで、遠隔移動の際にもなうデータ転送量を削減する機構を開発した。同様の手法を仮想ディスクに対しても開発し、その有効性を評価実験において確認した。また、仮想計算機のメモリにおいてページキャッシュとして使用されている領域を分析し、仮想計算機の移動時に重複排除を行う機構を開発した。

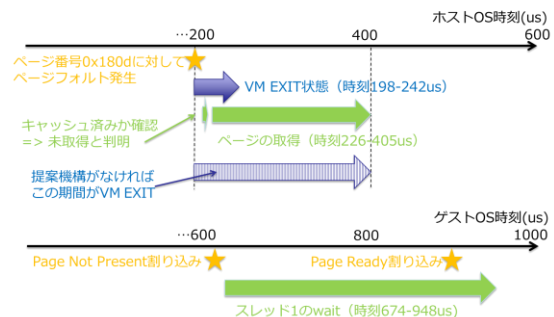
第三に、仮想計算機群が遠隔拠点に移動した際にもネットワーク到達性を透過的に維持できる手法について研究を進めた。Client Mobile IPv6 プロトコルのシグナリングメッセージやトンネリングメッセージをゲスト OS の代わりにホスト OS 上で処理する機構を開発した。仮想計算機内部を改変することなく透過的にトンネリングを可能とし、また複数の仮想計算機を遠隔移動する際にも、一つずつ柔軟に遠隔移動することを可能にした。



仮想計算機群が遠隔拠点に移動した際にもネットワーク到達性を透過的に維持する機構 (Kagemusha) の概要

第四に、遠隔移動の際に発生する性能低下を緩和する手法について研究を行った。ポストコピー型の移動機構においては、未だキャッシュされていないメモリページに対してゲスト OS がアクセスすると、メモリページが移動元から転送されるまでの間ゲスト OS 全体の実行が停止してしまう。特に同時並行的に複数の仮想計算機が移動する場合にその影響が大きい。そこで、仮想計算機のページフォルトハンドラを準仮想化することで、対象メモリページにアクセスしたゲスト OS 上のプロセスのみを停止し、残りのプロセスの実行を継続する手法を開発した。

提案機構 (性能低下抑制機構) の動作事例
スレッド1がページ番号0x180dに対してページフォルトを発生した場合



以上の研究成果を内外の学会で発表し、論文誌等にまとめた。特に、仮想計算機メモリをキャッシュし再利用する機構を発表した論文、および、ポストコピー型移動機構の性能向上手法を発表した論文は、それぞれシンポジウムにて論文賞を獲得した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 10 件)

- ① 広渕崇宏、山幡為佐久、伊藤智、準仮想化ページフォルトを用いたポストコピー型ライブマイグレーションの性能向上手法、情報処理学会コンピュータシステム・シンポジウム (ComSys2012) 論文集、pp. 2-12、査読有、Dec 2012
<http://id.nii.ac.jp/1001/00087509/>
- ② 広渕崇宏、中田秀基、工藤知宏、伊藤智、関口智嗣、仮想マシンに対して透過的な Client Mobile IPv6 トンネリング機構、電子情報通信学会論文誌 B、pp. 1239-1252、Vol. J95-B、No. 10、査読有、Oct 2012
<http://ci.nii.ac.jp/naid/110009517394>
- ③ Soramichi Akiyama、Takahiro Hirofuchi、Ryousei Takano、Shinichi Honiden、MiyakoDori: A Memory Reusing Mechanism for Dynamic VM Consolidation、IEEE Fifth International Conference on Cloud Computing、pp. 606--613、査読有、Jun 2012
10.1109/CLOUD.2012.56
- ④ Takahiro Hirofuchi、Hidemoto Nakada、Satoshi Itoh、Satoshi Sekiguchi、Kagemusha: A Guest-Transparent Mobile IPv6 Mechanism for Wide-Area Live VM Migration、Proceedings of the IEEE/IFIP Network Operations and Management Symposium 2012 (IFIP/IEEE International Workshop on Cloud Management)、pp. 1319-1326、査読有、Apr 2012
10.1109/NOMS.2012.6212069
- ⑤ 穂山空道、広渕崇宏、高野了成、本位田真一、都鳥：メモリ再利用による連続するライブマイグレーションの最適化、情報処理学会論文誌：コンピューティングシステム、pp. 74-85、査読有、Mar 2012
<http://id.nii.ac.jp/1001/00081515/>

- ⑥ Takahiro Hirofuchi、Hidemoto Nakada、Satoshi Itoh、Satoshi Sekiguchi、Reactive Cloud: Consolidating Virtual Machines with Postcopy Live Migration、IPSJ Transactions on Advanced Computing Systems、pp. 86-98、査読有、Mar 2012
<http://id.nii.ac.jp/1001/00081516/>

[学会発表] (計 12 件)

- ① Akiyama Soramichi、Fast Wide Area Live Migration with a Low Overhead through Page Cache Teleportation、The 13th IEEE/ACM International Symposium on Cluster and Grid Computing、2013年5月14日、Delft University of Technology (オランダ)
- ② 広渕崇宏、準仮想化ページフォルトを用いたポストコピー型ライブマイグレーションの性能向上手法、コンピュータシステム・シンポジウム (ComSys2012)、情報処理学会、2012年12月07日、筑波大学東京キャンパス (東京都)
- ③ Takahiro Hirofuchi、Kagemusha: A Guest-Transparent Mobile IPv6 Mechanism for Wide-Area Live VM Migration、The IFIP/IEEE International Workshop on Cloud Management、IEEE、2012年4月20日、マウイ (アメリカ)
- ④ 広渕崇宏、仮想マシンの広域ライブマイグレーションを実現するゲスト透過な Mobile IPv6 トンネリング機構、インターネットカンファレンス 2011、日本ソフトウェア科学会、2011年10月28日、九州産業大学 (福岡県)
- ⑤ Takahiro Hirofuchi、Reactive consolidation of virtual machines enabled by postcopy live migration、The 5th international workshop on Virtualization technologies in distributed computing、ACM、2011年6月8日、サンノゼコンベンションセンター (アメリカ)
- ⑥ 穂山空道、メモリの再利用により移動後の性能低下を抑えたライブマイグレーション、先進的計算基盤システムシンポジウム SACSIS2011、情報処理学会、2011年5月26日、秋葉原コンベンションホール (東京都)

[その他]

ホームページ等

<http://grivon.apgrid.org/>

6. 研究組織

(1) 研究代表者

広渕 崇宏 (HIROFUCHI TAKAHIRO)

独立行政法人産業技術総合研究所・情報技術研究部門・主任研究員

研究者番号：20462864