

科学研究費助成事業 研究成果報告書

平成 28 年 9 月 13 日現在

機関番号：82626

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330097

研究課題名(和文) ディザスタリカバリを可能にする高速退避型遠隔ライブマイグレーションの研究

研究課題名(英文) A study on fast, remote live migraton enabling disaster recovery

研究代表者

広淵 崇宏 (Hirofuchi, Takahiro)

国立研究開発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

交付決定額(研究期間全体)：(直接経費) 2,300,000円

研究成果の概要(和文)：本研究では、地震発生直後の猶予時間を利用したサーバ退避手法に一定の実現可能性を付与することを目指し、仮想計算機群の遠隔移動の高速化手法について研究を進めてきた。(1) 平時から実行状態を退避先に逐次ベストエフォートで転送し遠隔移動時には未転送の実行状態のみを転送する手法、(2) 転送データ中の重複を排除する手法、(3) 多くの仮想計算機を同時に移動する際に全体の移動時間を最小化するスケジューリング手法、(4) 仮想計算機移動のシミュレーション手法を開発した。

研究成果の概要(英文)：This project studied mechanisms to speed up the live migration of a group of virtual machines to a distant location. It aims at realizing safe server evacuation upon a disaster by exploiting a short period of time of backup power supply. We developed 1) a mechanism to transfer only updated states of a VM to a remote location upon a live migration, which is enabled by periodically sending the states of a VM during normal operation in a best-effort manner; 2) a mechanism to eliminate duplicated data in the data transferred during a live migration; 3) a mechanism that optimizes the scheduling of multiple live migrations to minimize an elapsed time of relocation of whole virtual machines; 4) a simulation mechanism of live migrations of virtual machines.

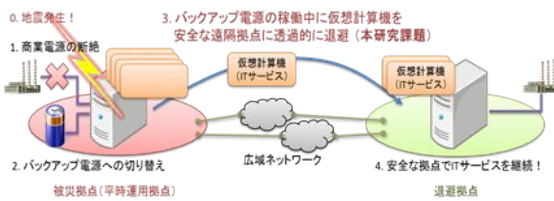
研究分野：システムソフトウェア

キーワード：仮想計算機 ハイパーバイザ ライブマイグレーション ディザスタリカバリ 仮想マシン 広域ネットワーク シミュレーション 分散システム

1. 研究開始当初の背景

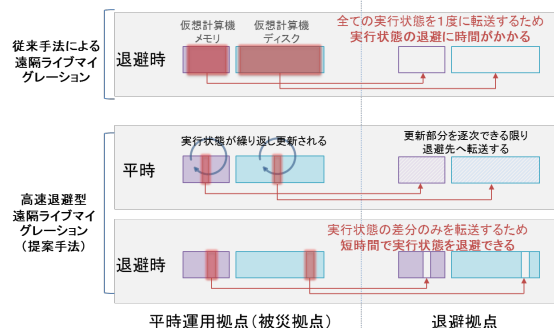
今日多くのデータセンタでは、仮想計算機とよばれる特殊なソフトウェア上にオペレーティングシステムやアプリケーションを導入し IT サービスを構築している。仮想マシンモニタとよばれるプログラムが、計算機を構成する CPU・メモリ・ストレージ等のハードウェアをエミュレーションすることで、ソフトウェア上に仮想的な計算機を作り出している。仮想計算機の実行状態をネットワーク越しにコピーすれば、仮想計算機を一切停止することなく他の計算機ハードウェア上に実行状態をそのまま移動(ライブマイグレーション)できる。

我々は、もしこの移動技術が遠隔拠点間で利用できるのであれば、広域災害発生時においてサーバを一切停止することなく安全な拠点へ退避して、IT サービスの継続が可能になるはずであると考えた。アプリケーションに依存しない退避技術として幅広い IT サービスに災害対策を提供できると期待される。しかし、我々の知る限り過去に提案されたいずれのライブマイグレーション手法も、災害発生直後のごく短い猶予時間中に多数の仮想計算機群を遠隔拠点に移動することが難しいという問題があった。



2. 研究の目的

本研究では、我々が過去に開発した遠隔移動技術を発展させて、短時間の間に完了できる遠隔移動技術を開発する。平時から仮想計算機の実行状態をベストエフォートで退避先へ転送し、退避時には残りの実行状態のみを転送する。評価実験を通して多様なネットワーク状況やワークロードに対する提案手法の有効性を明らかにする。



3. 研究の方法

初年度に提案機構のプロトタイプを実装し、翌年度以降にネットワークやワークロードのエミュレータを用いて様々な状況下における提案機構の有効性を評価する。

プロトタイプの実装においては、広く一般に普及している仮想計算機プログラムの一つである Qemu/KVM を拡張して、提案機構を実装する。評価実験において必要となるシミュレーション技術についても新たに開発する。フロリダ大学 ACIS 研究室やフランス研究機関 INRIA と適宜連携しながら進める。

4. 研究成果

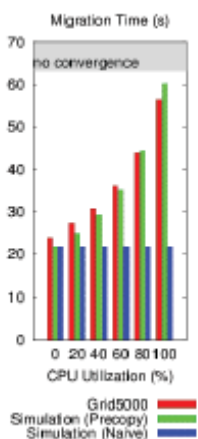
第一に、仮想計算機の主要な実行状態であるメモリおよびストレージに対して、そのデータをあらかじめ宛先に転送する手法について研究を行った。あらかじめ仮想計算機の実行状態を宛先へ転送(キャッシュ)しておけば、移動時には残りの実行状態のみを転送することで移動時間を短縮できる。メモリに関しては、我々が以前発表した「過去に存在した仮想計算機のメモリデータを再利用することで移動時間を短縮する手法」を発展させて、「現在実行中のメモリデータをあらかじめ宛先に転送することで移動時間を短縮する手法」を開発した。またストレージについては、既存のストレージサーバプログラムを拡張して、ディスクブロックをあらかじめ宛先に転送する機構を開発した。

第二に、遠隔移動時のデータ転送中に含まれる重複データを排除する機構を開発した。過去に発表した手法では仮想計算機を移動する前に重複排除を完了する必要があるため時間短縮効果が限定的であったが、今回新たに仮想計算機の移動と並行して重複排除を行う機構を開発したことで大きな時間短縮効果を得られた。詳細な評価実験を行った結果、部分的に重複排除を行った場合に最も仮想計算機の移動時間を短縮できることがわかった。そこで時間短縮効果が最大となるように自動的に重複排除率を調整する機構を新たに開発した。時間短縮効果に関する理論的な分析も行った。

第三に、多くの仮想計算機を遠隔拠点に一度に移動する際に、仮想計算機群全体の移動時間を最小化する機構について研究を行った。過去に開発したプロトタイプに対して詳細な実験および理論的な考察を行い、その効率性を詳しく評価した。

第四に、仮想計算機の移動技術に関するシミュレーション手法について研究を進めた。フランスの研究者らと共同で、過去に開発したシミュレーションフレームワークを強化し、より正確に移動時間を見積もることを可

能にした。本フレームワークにより数万台以上の仮想計算機からなる大規模環境のシミュレーションが高い精度で可能になる。クラウドコンピューティング分野におけるトップレベルの論文誌 IEEE Transactions on Cloud Computing に投稿し、最終的に採録を得ることができた。フレームワークのソースコードを整理し、使用方法を記載した文書を準備するなど、開発技術の普及活動を行った。



仮想計算機の移動時間 (Migration Time、単位は秒) をシミュレーションおよび実機で比較。開発したシミュレーション手法 (緑) は実機 (赤) とほぼ同様の移動時間を見積もることができる。詳細は雑誌論文を参照のこと。

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 7 件)

Postcopy Live Migration with Guest-cooperative Page Faults, Takahiro Hirofuchi, Isaku Yamahata, Satoshi Itoh, IEICE Transactions on Information and Systems, 査読有, pp.2159-2167, Vol.E98-D, No.12, IEICE, Dec 2015

10.1587/transinf.2015PAP0011

SimGrid VM: Virtual Machine Support for a Simulation Framework of Distributed Systems, Takahiro Hirofuchi, Adrien Lebre, Laurent Pouilloux, IEEE Transactions on Cloud Computing, 査読有, pp.1-14, IEEE, Sep 2015

10.1109/TCC.2015.2481422

Design and Implementation of Middleware for Cloud Disaster Recovery via Virtual Machine Migration Management, Tae Seung Kang, Mauricio Tsugawa, Andrea Matsunaga, Takahiro Hirofuchi, Jose Fortes, Proceedings of the 7th IEEE/ACM International Conference on Utility and Cloud Computing, 査読有, pp.166-175, Dec 2014

10.1109/UCC.2014.25

Evaluating Impact of Live Migration on

Data Center Energy Saving, Soramichi Akiyama, Takahiro Hirofuchi, Shinichi Honiden, Proceedings of the IEEE 6th International Conference on Cloud Computing Technology and Science, 査読有, pp.759-762, Dec 2014

10.1109/CloudCom.2014.42

Fast Live Migration with Small IO Performance Penalty by Exploiting SAN in Parallel, Soramichi Akiyama, Takahiro Hirofuchi, Ryousei Takano, Shinichi Honiden, Proceedings of the 7th International Conference on Cloud Computing, 査読有, pp.40-47, Jun 2014

10.1109/CLOUD.2014.16

A WAN-optimized Live Storage Migration Mechanism Toward Virtual Machine Evacuation Upon Severe Disasters, Takahiro Hirofuchi, Mauricio Tsugawa, Hidemoto Nakada, Tomohiro Kudoh, Satoshi Itoh, IEICE Transactions on Information and Systems, 査読有, pp.2663-2674, Vol.E96-D, No.12, IEICE, Dec 2013

10.1587/transinf.E96.D.2663

Fast Wide Area Live Migration with a Low Overhead through Page Cache Teleportation, Soramichi Akiyama, Takahiro Hirofuchi, Ryousei Takano, Shinichi Honiden, Proceedings of the 13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 査読有, pp.78-82, IEEE, May 2013

10.1109/CCGrid.2013.57

[学会発表](計 5 件)

Tae Seung Kang, Design and Implementation of Middleware for Cloud Disaster Recovery via Virtual Machine Migration Management, The 7th IEEE/ACM International Conference on Utility and Cloud Computing, 2014年12月9日、ロンドン(イギリス)

Soramichi Akiyama, Evaluating Impact of Live Migration on Data Center Energy Saving, The IEEE 6th International Conference on Cloud Computing Technology and Science, シンガポール(シンガポール) 2014年12月17日

Soramichi Akiyama, Fast Live Migration with Small IO Performance Penalty by Exploiting SAN in Parallel, The 7th International Conference on Cloud Computing, アンカレッジ(アメリカ) 2014年6月28日

Soramichi Akiyama, Fast Wide Area Live Migration with a Low Overhead through Page Cache Teleportation, The 13th IEEE/ACM International Symposium on

Cluster, Cloud and Grid Computing, デ
ルフト(オランダ), 2013年5月14日

6. 研究組織

(1) 研究代表者

広淵 崇宏 (HIROFUCHI TAKAHIRO)

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