

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 4 日現在

機関番号：14603

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24700067

研究課題名(和文)クラウド環境向け資源管理機構の研究開発

研究課題名(英文)A Resource Management Mechanism for Cloud Computing Environment

研究代表者

奥田 剛 (Okuda, Takeshi)

奈良先端科学技術大学院大学・情報科学研究科・特任准教授

研究者番号：30304067

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：クラウドコンピューティングの普及により、計算機資源を共有することで資源の効率的な利用が可能になった。しかし依然としてクラウドコンピューティングで用いられる計算機クラスター内部では断片化により有効活用できない資源が残っている。

本研究課題では、複数の計算機をまたがって資源の利用状況を把握し、断片化により利用されることがない資源を高速なネットワークを介して集約し、資源が不足している計算プロセスや仮想計算機に提供する事で資源の有効活用を可能にした。この研究課題で得られた知見は論文誌等で発表し、試作したシステムはオープンソースソフトウェアとして公開することで資源管理機構の更なる発展に貢献している。

研究成果の概要(英文)：The spread of cloud computing has enabled high utilization of computational resources by sharing resources among multiple users. Even under such a condition, there still remains unused resources within a compute cluster due to fragmentation of resources.

In this research project, we proposed a resource management mechanism which enables resource accommodation to the virtual machines and processes in need of more resources by monitoring resource usage and aggregating fragmented resources. To expedite and promote researches in this area, we publicized several papers in an academic journal, international conference, and domestic workshop and a prototype as an open source software.

研究分野：情報科学

キーワード：資源管理機構 クラウドコンピューティング 高速ネットワーク 資源共有 データセンター

### 1. 研究開始当初の背景

クラウドコンピューティングと呼ばれる複数の仮想計算機環境をクラスタ化して利用する形態が一般化するに従って、これまでの計算機内部の資源が未使用のままになる断片化だけではなく、仮想計算機環境クラスタ全体の資源の断片化が顕在化してきている。従来は実計算機の保持する資源は単一のOSの資源管理機構によって管理することで断片化を解消していた。しかし、仮想計算機環境では、実計算機上で複数の仮想計算機上のOSがそれぞれに割り当てられた資源を管理するため、仮想計算機上と実計算機上の双方で資源の断片化が発生する。また、複数の仮想計算機環境をクラスタ化する場合、各仮想計算機環境での資源の断片化が積み重なり、クラウドコンピューティング環境での資源の有効利用の妨げになっている(図1)。

私の研究チームでは、従来実計算機で用いられてきた資源共有手法などを仮想計算機環境に拡張し、特にメモリをネットワークを介して共有する手法を研究してきた。ここでは、仮想計算機が透過的にリモートメモリを利用することを可能としている。さらにクラウドコンピューティング環境でのメモリ共有の有効性を確かめている。これらの研究をとおして、今後複数の仮想計算機環境をまたいで、クラスタ全体として資源断片化を解消し、資源利用効率を改善するには、クラスタ全体を統轄する資源管理機構が必要であると考えた。

そこで、本研究では、仮想計算機環境内の資源の利用状況を把握、未使用領域を集約、資源が不足している仮想計算機環境に対して資源を貸与する資源管理機構を提案する。仮想計算機環境内では、自律的に資源の利用状況を監視し、資源管理機構に対して報告する。資源が不足している仮想計算機環境は、資源管理機構に対して利用可能な資源の要求を自律的に行い、未使用領域の共有を行うことで断片化の解消と資源利用率の向上を自律的に行う。

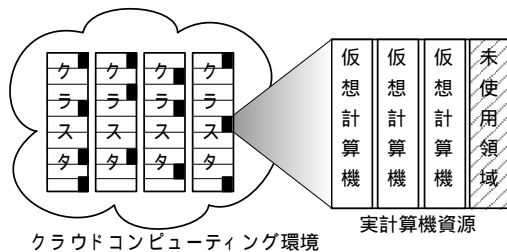


図1 クラウドコンピューティング環境内の資源断片化

### 2. 研究の目的

本研究では、仮想計算機クラスタ内の資源の利用状況を把握、断片化して利用されないことのない資源を、高速ネットワークを介して集約し、資源を必要とする仮想計算機やプロセスに提供する資源管理機構を提案する。この資源管理機構を導入することにより、計算機クラスタ内の計算資源をより有効に活用することを目的とする。

### 3. 研究の方法

(1)これまで研究開発を行ってきた仮想計算機環境間でのメモリ共有機構を、仮想計算機クラスタ全体で利用可能とするための資源管理機構を開発し、基礎的な評価を行う。まず、仮想計算機が利用していないメモリ領域を仮想計算機環境に解放し、他の仮想計算機環境が利用できるようにする。さらに、メモリ不足の仮想計算機に対して共有を仲介する資源管理機構を開発する。これまでの我々の方式ではメモリ共有プロトコルとしてiSCSIのみを用いていたが、NBDなど他の共有プロトコルへの対応を検討し、仮想計算機環境のアーキテクチャに依存しない管理機構とする。次に、開発した資源管理機構を用いてテスト環境において評価実験を行う。テスト環境構築のため、実験・開発用PCを2台導入し、それらと現有する学内ネットワークを接続する。最後に、基本機能の試作をオープンソースソフトウェアとして公開し、国内外の開発者に広く利用してもらうことにより、研究へのフィードバックを得ることを計画している。

(2)前項の基本機能に対して機能拡張を行うとともに、資源共有の更なる効率化を行う。研究開発した資源管理機構で、仮想計算機クラスタ全体の資源利用率、計算機の信頼性、性能向上などの最適化指標を指定できるようにする。また、資源の利用率、信頼性、性能向上を行うため、空き資源の集約により、利用率と性能向上を測り、ミラーリングやストライピングなどの高信頼化手法を用いることで信頼性と性能の両方の向上を図る。次に、異種アーキテクチャ混在環境での資源共有割り当て最適化について検討する。クラスタ内の資源とクラスタ間の資源では、利用可能なネットワーク帯域や遅延に違いがあるため、共有後の性能に差が生じる。この性能差を最小化あるいは平滑化するため、共有開始時に計測した帯域や遅延等の情報を資源共有ポリシーと照らし合わせ共有先を決定する機能を実現する。

(3)前項に対して、さらに仮想計算機支援技術のSR-IOVや、オフロード技術のiSCSI off-load等の技術の適用可能性とこれらを適用した際の効果も計測する。これまでに得られた成果を国際会議並びに学術雑誌に投稿し、公表する。

#### 4. 研究成果

(1) ある仮想計算機環境で未使用のメモリ領域を、他の仮想計算機環境が利用できるようにメモリ共有の仲介を行う資源管理機構を開発した。各仮想計算機環境内に資源の利用状況をモニタする資源モニタを設置し、空き状況を常に資源管理機構に伝える。資源が不足する仮想計算機環境は、必要な資源量を資源管理機構に要求し、資源の貸し出しが可能な資源モニタに対して貸し出し準備要求を送信し、借り手には URL を伝えることでメモリ共有を行う(図 2)。このメモリ共有では iSCSI プロトコルを利用して、メモリ領域の共有を行っている。メモリ共有のプロトコルとして NBD への対応を検討し、基礎性能の評価をおこなったが、NBD ではメモリ共有を実現するのに十分なパフォーマンスが得られない事が分かり、NBD への対応は取りやめた。今回開発した資源管理機構はオープンソースソフトウェアとして公開している。

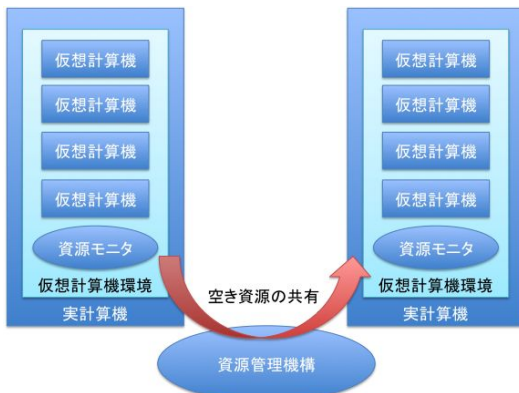


図 2 資源管理機構の概要図

(2) 仮想計算機クラスタ全体の資源利用率と性能向上は資源利用の多重化という観点からはトレードオフの関係にある。例えばネットワークや仮想計算機の多重度を増すと資源利用率は向上するが、テナント(仮想計算機インスタンス)あたりの性能は低下する。そこで、開発した資源管理機構に対して、資源利用率、性能向上のうちどちらを優先するかを選択できるようにした。性能に影響する仕様としてはネットワークの帯域(1Gbps or 10Gbps)を選択できるようにした。また、資源の利用率、信頼性、性能向上の3者間のトレードオフも検討し、従来の高信頼手法では、資源利用率と性能の2つが犠牲になる事を確認した。

(3) 異種アーキテクチャ混在環境での資源共有割当の最適化について検討した。クラスタ内とクラスタ間の資源共有では、利用可能なネットワーク帯域や遅延に差があるため、共有後の性能に影響を及ぼす。例えば別のデータセンターに存在するクラスタとの資源共有では、性能が著しく低下するにもかかわらず、

ネットワーク帯域を無駄に消費することになり、資源を効率よく利用できていないとは言えない。このため、クラスタ間での資源共有は現実的ではない。例えば並列分散計算で広く利用されている Hadoop では、あらかじめラックトポロジというネットワーク構成を知識として与え、計算をなるべく同じラックに集約するように制御している。同様に本資源共有もクラスタ内の異種アーキテクチャ混在環境での資源共有に限定する事が可能である。本提案方式自体はアーキテクチャに依存せず、実装もアーキテクチャに依存しないオープンスタンダードを選択しているため、仮想計算機モニタの種別に依存せず利用可能である。

資源管理機構の信頼性向上と性能向上を目的として iSCSI off-load 機能を有する NIC を用いた実験を行った。仮想計算機サーバが高負荷の状況においては、iSCSI off-load 機能を有効にする事で応答性が改善する事を確認した。ただし、iSCSI off-load を利用するには、この機能に対応した NIC とデバイスドライバ、専用の iSCSI プロトコルスタック(イニシエータとターゲット)の整合性が必要となるという制約がある。

(4) 太陽光発電や外気冷却等を取り入れた環境指向のクラウドデータセンター実現に関する提案も行った。これは計算資源の利用率向上と電力消費の関係に対する知見を活かした提案で、仮想計算機とそのストレージを太陽あるいは月を追跡するように地球を周回させるように移動するというものである。データセンターで消費される電力の多くは計算機等の冷却に使われているが、この提案によりその冷却に使われる電力を削減する、あるいは再生可能エネルギーでまかなう事ができる事を示した。

また、クラウド環境内で共有される資源に対して、セキュリティリスクに基づいてアクセス制御を施す機構を提案した。仮想計算機環境で用いられるソフトウェアと各テナントで用いられる OS やソフトウェアの脆弱評価と、セキュリティリスクの伝搬をモデル化し、定式化する事により動的にアクセス制御を行う方式である。

本研究課題で得られた知見や提案内容は、国内研究会、国際会議、学術雑誌、オープンソースソフトウェア等で発表しており、今後のこの分野の発展に寄与すると期待している。

<引用文献>

山本和典,石川裕.テラスケールコンピューティングのための遠隔スワップシステム Teramem, 情報処理学会論文誌 (ACS), Vol.2, No.3, pp.142-152, September 2009.

Tia Newhall and Douglas Woos. Incorporating Network RAM & Flash into Fast Backing Store for Clusters, in Proceedings of IEEE Cluster, September 2011.

T. Okuda et al. A Remote Swap Management Framework in a Virtual Machine Cluster. In IEEE 3rd International Conference on Cloud Computing, pp.546-547, July 2010.

T. Okuda et al. A Mechanism of Flexible Memory Exchange in Cloud Computing Environments. In 2nd IEEE International Conference on Cloud Computing Technology and Science, pp.76-80, December 2010.

Dan Williams, Hakim Weatherspoon, Hani Jamjoom and Yew-Huey Liu. Overdriver: Handling Memory Overload in an Oversubscribed Cloud, in Proceedings of ACM VEE '11, pp. 205-216, March 2011.

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 1件)

1. Doudou Fall, Takeshi Okuda, Youki Kadobayashi, and Suguru Yamaguchi. "Security Risk Quantification Mechanism for Infrastructure as a Service Cloud Computing Platforms", Journal of Information Processing, 査読あり, Vol. 23, 2015.

[学会発表](計 2件)

1. 澤田直也,岡田和也,奥田剛,山口英.「差分ディスクを利用した環状マイグレーション方式の提案とその評価」,電子情報通信学会 IA 研究会,2015年3月5日,小名浜オーションホテル(福島県いわき市).

2. Doudou Fall, Takeshi Okuda, Youki Kadobayashi, and Suguru Yamaguchi. "Towards a Vulnerability Tree Security Evaluation of OpenStack Logical Architecture", The 7th International Conference on Trust & Trustworthy Computing (TRUST2014), July 30, 2014, Hersonissos (Greece).

[図書](計 0 件)

[産業財産権]

出願状況(計 0 件)

名称:  
発明者:  
権利者:  
種類:  
番号:  
出願年月日:  
国内外の別:

取得状況(計 0 件)

名称:  
発明者:  
権利者:  
種類:  
番号:  
出願年月日:  
取得年月日:  
国内外の別:

[その他]  
ホームページ等  
<http://iplab.naist.jp/>

6. 研究組織

(1)研究代表者

奥田 剛 (OKUDA, Takeshi)  
奈良先端科学技術大学院大学・情報科学研究科・特任准教授  
研究者番号: 30304067

(2)研究分担者

( )

研究者番号:

(3)連携研究者

( )

研究者番号: