

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 23 日現在

機関番号：82636

研究種目：若手研究(A)

研究期間：2011～2014

課題番号：23680002

研究課題名(和文)クラウド基盤におけるメモリ管理の高度化による消費電力削減手法の研究開発

研究課題名(英文)Research on power-efficient cloud infrastructure with advanced memory management

## 研究代表者

河合 栄治(Kawai, Eiji)

独立行政法人情報通信研究機構・テストベッド研究開発推進センターテストベッド研究開発室・室長

研究者番号：40362842

交付決定額(研究期間全体)：(直接経費) 19,900,000円

研究成果の概要(和文)：本研究では、大規模化し、仮想化され、非常に複雑なものになってきた情報システムを対象に、メモリやネットワークなどプロセッサ以外のリソースに着目、物理リソースから独立した論理システムの集合体として管理される仮想化に対応した最適化手法を開発した。具体的には、仮想化されたホストを対象としたメモリ管理、広域分散システムを対象としたSDN技術の応用による資源管理における最適化手法を開発した。これらの手法は、広域分散システムとしてクラスタシステム向けのジョブ管理システムや広域ネットワーク環境に展開された可視化システムをユースケースとして検証し、その有効性を検証した。

研究成果の概要(英文)：The goal of this research is to develop optimization mechanisms in resource management for large-scale, virtualized, and highly-complicated information systems. Especially, we focused on resource management mechanisms for memory and networks, in contrast to the conventional ones for processors. We developed a memory management mechanism for virtual host environments and application-oriented network management mechanisms for widely distributed systems which use SDN (Software-Defined Networking) technologies. Those developed mechanisms were proved to achieve performance improvement in use-cases: an SDN-enhanced job management system for cluster computing and an SDN-enhanced TDW (Tiled Display Wall) system deployed on wide area networks.

研究分野：ICT基盤技術

キーワード：省エネルギー 情報システム オペレーティングシステム メモリ管理 クラウドコンピューティング

### 1. 研究開始当初の背景

情報システムの最適化は非常に重要になっている。従来の最適化技術は、計算機システムのプロセッサが有する処理能力に着目し、それを最大限生かすことを目的としたものが多い。消費電力削減の観点からも、プロセッサが消費する電力が占める割合が高いことから、その処理能力をできる限り犠牲にせず、最適化を行うものに注目が集まっている。

一方で、情報システムは、大規模化し、仮想化され、非常に複雑化なものになってきており、最適化手法もそうした変化に対応することが求められている。近年では、多数のコンピュータがネットワークで接続され、分散処理を前提にしたものが主流になってきている。また、構成する個々のシステムは仮想化され、物理リソースから独立した論理システムの集合体として管理される。仮想ホスト環境では、プロセッサも仮想化されており、オペレーティングシステムでさえ、物理プロセッサを専有して利用することを前提にできなくなっている。その結果、情報通信インフラの管理技術は、非常に複雑化してきている。

### 2. 研究の目的

本研究の目的は、大規模化し、仮想化された情報通信インフラにおける消費電力の削減などを含む多様な最適化技術を検討し、開発することである。大規模化については、特に広域ネットワークによって相互接続された分散システムをターゲットに、処理能力だけではなく、メモリや、通信などもターゲットにする。さらに、仮想化技術にも配慮し、大規模分散システムにおける仮想リソースと物理リソースの関係をアプリケーション性能とリソース消費のバランスの観点から最適化する手法を検討する。

### 3. 研究の方法

#### (1) メモリ管理技術の最適化

従来、プロセッサを中心にリソース管理の最適化技術が開発されてきたが、消費電力の観点からはメモリもプロセッサに匹敵する電力を消費しており、その最適化が求められる。メモリ管理の最適化では、ウィークポイントを活用したメモリ管理技術を開発した(図1)。これは、従来のメモリ管理のように一度割り当てたメモリは明示的にあるいは暗黙的に開放しない限りは割り当て続けられるというモデルではなく、キャッシュのようにその利用者とは独立して管理され、開放されるモデルである。

仮想化されたホスト環境では、物理ホストに収容される他の仮想ホストの状況が急激に変化する状況が発生する。その際、従来のようなメモリ管理モデルでは、その変化に追従することができず、大幅な性能低下が発生する。また、消費電力の観点から、利用する

メモリ量をコンパクト化し、仮想ホストを集約することで、稼働する物理ホスト数を削減するといったことも可能である。

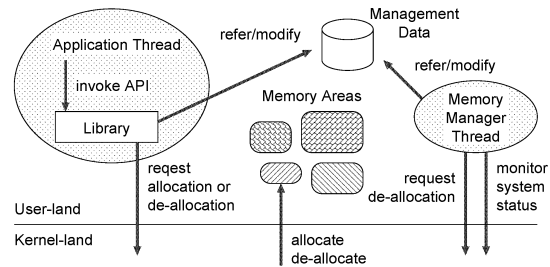


図1: 開発したメモリ管理システムの概観

#### (2) SDNによる広域ネットワーク管理技術

大規模分散システムにおいてはネットワークが重要な要素となるが、従来はL2VLANなどの静的な仮想回線を用いて制御するか、アプリケーション層でオーバーレイネットワークを構築して制御するぐらいしか現実的な方法はなかった。近年、SDN(Software-Defined Networking)技術に代表されるような柔軟なネットワーク制御技術が登場し、状況が変化してきている。そこで、通信も含めた分散システムの最適化の観点から、広域ネットワークにおいてSDNを活用する手法の開発を行った。これは、従来データセンター等のLAN環境での利用が想定されていたSDNを既存の広域ネットワーク上に展開し、L2VLANのようなネットワーク機器での仮想回線制御をプログラムから動的に行うことを可能にすると共に、SDNをオーバーレイ的に構築することで、アプリケーション層オーバーレイネットワークで生じていたパケット処理に関わるオーバーヘッドの削減も実現するものである。

また、広域ネットワーク環境でSDNを使う場合、その集中制御であるが故の規模拡張性の問題の解決、特に仮想化によるマルチユーザ(マルチテナント)環境の実現が必要となる。そこで、広域ネットワークにおけるSDNの仮想化管理技術を開発した。これは、従来のSDNの仮想化手法が、ネットワーク制御の内容にユーザ間で衝突がないように別途調停が必要なため規模拡張性に欠けていた問題を、ネットワーク制御の内容の自動的に変換することで事前調停を不要にして解決するものである。

#### (3) 大規模分散システムと連動したインフラ制御

大規模化し、仮想化された情報通信インフラにおいて最適化を考える上で、ユースケースとして2つのアプリケーションを対象として検証を行った。

1つは、クラスタコンピューティングでのジョブ割り当てにおけるリソース管理の最適化である。従来プロセッサ負荷の状況が重視されていたのに対して、ネットワークの構

成を含め他の状況を考慮するものである(図2)。

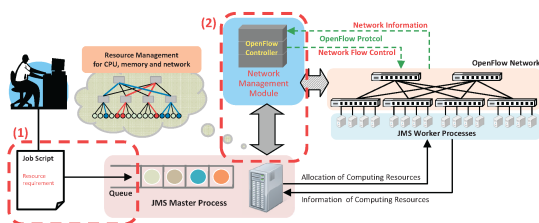


図2: 開発したジョブ管理システムの概観

もう1つは、広域ネットワーク上での TDW (Tiled Display Wall) を共有利用である。TDW とは、多数の映像表示ディスプレイを並べて一体化して大量の可視化情報を提示することを実現したシステムであり、データソースサーバ、映像管理サーバ、映像表示サーバなどが連携する分散システムである。これを広域ネットワーク上に展開し、その上でのユーザの振る舞い(ユーザによるシステム操作)に連動した最適化(図3)について検証した。

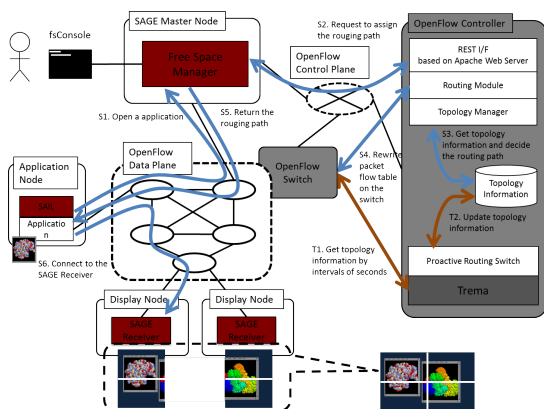


図3: 開発した TDW ネットワーク連携システムの概観

#### 4. 研究成果

##### (1) メモリ管理技術の最適化

メモリ管理の最適化では、仮想ホストを想定した環境において、急激にメモリリソースが欠乏する状況を模擬し、その中での提案手法を実装したシステム(プロセス)のメモリアクセス性能について評価実験をおこなった。具体的には、メモリアクセス中心のベンチマークプログラムに提案手法を組み込み、一定量のメモリを利用する形で実行を開始した後、徐々にシステム全体の空きメモリが欠乏するように別プロセスを実行し、その経過におけるベンチマークプログラムの性能を計測した。

その結果、本研究で開発したメモリ管理機構により、利用可能なメモリ量が減少していく中でメモリの非自発的開放により、過度のページアウトによりシステムが機能不全に陥るような急激な性能低下の状況を生じさ

せず、非開放のメモリ領域についてはアクセス性能を十分に維持することができた。現行のサーバシステムでは、システム上のメモリモジュール毎の消費電力の制御まではできないことから、提案手法により、サービス品質をある程度維持しつつもメモリ使用量を削減することで消費電力の削減が可能であることを示している。

##### (2) SDN による広域ネットワーク管理技術

広域ネットワークにおける SDN の展開技術については、開発した技術を独立行政法人情報通信研究機構が運用する SDN テストベッドである RISE (Research Infrastructure for large-Scale network Experiments) 上で検証し、複数の SDN コントローラから衝突のある制御情報を同時に投入しても、それぞれ論理的に隔離された形で制御が実施されることが実証された。こうしたマルチテナント SDN 技術は、次で述べる大規模分散システム制御における SDN の適用においても応用した。

##### (3) 大規模分散システムと連動したインフラ制御

ネットワーキングを含めた大規模システムでの最適化では、開発したクラスタシステム用のジョブ管理システムにおいて、ベンチマークシステムを用いた性能評価を行った。その結果、ネットワークを考慮しない場合と比較して、SDN を応用した制御により最大で 15% 程度の性能向上が見られることが分かった(図4)。



図4: 開発したジョブ管理システムの効果

また、SDN によるネットワーキングの最適化機構の応用例として、Tiled Display Wall (TDW) に組み込んだユースケースでは、ユーザの振る舞い(ユーザによるシステム操作)とネットワーキングの最適化を連携させる仕組みを実現し、ネットワーク上の輻輳を自動的に回避するなどの効果が得られることが分かった。

#### 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 2件)

渡場康弘、木戸義之、伊達進、阿部洋文、市川昊平、山中広明、河合栄治、竹村治雄、計算資源とネットワーク資源を考慮した割り当てポリシーを配備可能とするジョブ管理フレームワーク、電子情報通信学会和文論文誌 D, Vol. J97-D, No. 6, 2014、pp.1082-1093、DOI: [http://search.ieice.org/bin/summary.php?id=j97-d\\_6\\_1082](http://search.ieice.org/bin/summary.php?id=j97-d_6_1082)  
Yasuhiro Watashiba, Susumu Date, Hirotake Abe, Yoshiyuki Kido, Kohei Ichikawa, Hiroaki Yamanaka, Eiji Kawai, Shinji Shimojo, and Haruo Takemura, Efficacy Analysis of an SDN-enanced Resource Management System through NAS Parallel Benchmarks, The Review of Scionetwork Strategies, Vol. 8, Issue 2, 2014, pp.69-84, DOI: 10.1007/s12626-014-0045-9

〔学会発表〕(計 7件)

Yasuhiro Watashiba, Susumu Date, Hirotake Abe, Yoshiyuki Kido, Kohei Ichikawa, Hiroaki Yamanaka, Eiji Kawai, Shinji Shimojo, and Haruo Takemura, Performance Characteristics of an SDN-enanced Job Management System for Cluster Systems with Fat-tree Interconnect, Emerging Issues in Cloud (EIC) Workshop, 2014年12月15日, Singapore (Singapore)  
Hiroaki Yamanaka, Eiji Kawai, Shuji Ishii, and Shinji Shimojo, AutoVFlow: Autonomous Virtualization for Wide-area OpenFlow Networks, European Workshop on Software Defined Networks 2014, 2014年9月1日, Budapest (Hungary)  
Wasuhiro Watashiba, Yushiyuki Kido, Susumu Date, Hirotake Abe, Kohei Ichikawa, Horoaki Yamanaka, Eiji Kawai, and Haruo Takemura, Prototyping and Evaluation of a Network-aware Job Management System on a Cluster System Leveraging OpenFlow, The 19th IEEE International Conference on Networks (ICON 2013), 2013年12月11日, Singapore (Singapore)  
Yasuhiro Watashiba, Susumu Date, Hirotake Abe, Kohei Ichikawa, Hiroaki Yamanaka, Eiji Kawai, and Haruo Takemura, An Architectural Design of a Job Management System Leveraging Software Defined Network, IEEE International Workshop on High-Speed Network and Computing Environment (HSNCE 2013), 2013年6月26日, 京都テルサ(京都府京都市)  
山中広明、河合栄治、石井秀治、下條真

司、フロースペース仮想化における任意のフロー定義が可能なフロー変換手法の検討、電子情報通信学会情報ネットワーク研究会、2012年12月13日、広島市立大学(広島県広島市)

Eiji Kawai, Can SDN help HPC?, Computer Software and Applications Conference Workshop, 2012年7月17日, Izmir (Turkey)

Eiji Kawai, VMMF: Volatile Memory Management Framework, The 2012 IEEE International Systems Conference, 2012年3月22日, Vancouver (Canada)

6. 研究組織

(1) 研究代表者

河合 栄治 (KAWAI, Eiji)

独立行政法人情報通信研究機構・テストベッド研究開発推進センターテストベッド研究開発室・室長

研究者番号：40362842