

仮想マシン配置の最適化のためのライブマイグレーション手法に関する検討

櫻井 邦彦* 佐野 浩士 石井 啓之 (東海大学)

A Proposal on Live Migration Method for Optimised Virtual Machine Deployment

Kunihiko Sakurai*, Hiroshi Sano, Hiroshi Ishii, (Tokai University)

Virtualisation technology which can reduce the number of physical machines for effective use of resources has become a research trend. Several methods of Live Migration (LM) that allows us to migrate a Virtual Machine (VM) from one physical server to another without disrupting its normal operation for optimised VM deployment have been proposed. However, they do not consider some significant factors, for example, Round-Trip Delay Time (RTT) between each base and Random-Access Memory (RAM) usage of each VM that have impacts on the actual transfer time when LM is implemented. Considering these backgrounds, we have already proposed a method of LM in case of the “On-premise” system structure to balance the load of each physical server by assuming the Internet Virtual Private Network (VPN). Additionally, this paper assumes RTT proportional to the distance between each base where the server is placed as an additional condition and shows its effectiveness.

キーワード：仮想マシン，ライブマイグレーション，負荷分散，往復遅延時間

(Keywords: Virtual Machine, Live Migration, Load Balancing, Round-Trip Delay Time)

1. はじめに

仮想マシンとは、オペレーティング・システム (OS) 上で実際のコンピュータである物理マシンの様に動作するコンピュータ・ファイルのことであり⁽¹⁾、仮想化技術によりエンドユーザは物理マシンを使用する際と同様に仮想マシンを使用することができる。物理マシンと比較した際の安全性、可用性、管理性といった利点および、資源の有効活用を目的としたシンクライアントなどを代表とする仮想化技術の活用が近年注目されている。特に、仮想化技術の 1 つであるライブマイグレーションは物理サーバ上で起動中の仮想マシンの動作を止めず、他のサーバに移動させることを可能にする。そのため、物理サーバのメンテナンスや負荷分散に活用されている。ライブマイグレーションを実行し、仮想マシンが現在稼働している物理サーバから他の物理サーバに転送させる際、転送が完了するまで仮想マシンのメモリ内容のコピーが行われており、転送中もユーザは仮想マシンを使用することができる。この利点から、ライブマイグレーションはメンテナンスなどに活用されている。ライブマイグレーションの実行の可否については、いくつかの要素を検討する必要がある。既存研究において仮想マシンが動作する物理サーバからクライアント端末への画面転送に使用されるリモートデスクトッププロトコル、ネットワーク品質とライブマイグレーションについての提案や評価が行われている^{(2)~(6)}。しかし、これらの既存研究

においては、実際のライブマイグレーションの実行の可否を左右する仮想マシンの転送時間に関わるネットワーク品質、メモリ使用量の関係が十分検討されていない。

これらを踏まえ、我々は各拠点間がインターネット VPN によって接続された主に小規模な組織によって運用されるオンプレミス型のシステム構成を想定したライブマイグレーション手法を提案している。十分な帯域の確保が可能な専用線を使用しない環境を想定しているため、仮想マシンのメモリ使用量に応じてその転送時間も長時間化する。よって、各拠点間の往復遅延時間および仮想マシンのメモリ使用量がライブマイグレーションの実行の可否の判断のために重要となる。加えて、ライブマイグレーションにより各物理サーバの負荷分散を行う際、ユーザの仮想マシンの使用時間や転送時間も加味される必要がある。これらを反映させ、我々は物理サーバが設置された各拠点間の往復遅延時間と仮想マシンのメモリ使用量を加味し、仮想マシンを最適配置する手法を提案している⁽⁶⁾。引用文献⁽⁶⁾においては、その一次評価を行っており、有用性を示している。

さらに、本稿においては、一次評価の結果に加え、追加条件として各拠点間の距離に比例した往復遅延時間の設定を行うことにより再検討を行い、その有用性を示している。

本稿は以下の構成である。2章では関連研究およびそれらの課題点を述べ、3章では我々の提案手法およびその一次評価について述べ、4章では引用文献⁽⁶⁾をもとに各拠点間の距離を勘案した再検討を行い、5章で本稿をまとめている。

2. 関連研究

関連研究においては、データセンタ間においてライブマイグレーションを実行することでネットワーク全体の遅延を削減する方式が提案されている⁽⁷⁾。データセンタ間の距離および TCPセッション数をコストとして計算し、仮想マシンを低コストのデータセンタに転送することでネットワークの遅延が削減される。しかし、実際に仮想マシンを転送する際に要する時間が勘案されておらず、これはデータセンタ間の往復遅延時間および仮想マシンのメモリ使用量に依存する。

他の関連研究においても、ライブマイグレーションを用いた仮想マシンの配置手法が提案されている⁽⁸⁾。ネットワーク負荷が存在する場合、仮想マシンをユーザのシンクライアント端末に近い物理サーバに転送することにより、迅速な画面遷移など、ユーザエクスペリエンスを向上させることができる⁽⁹⁾と述べられている。一方、ライブマイグレーションには欠点もあり、本稿がすでに述べた様に仮想マシンがライブマイグレーションにより他の物理サーバに転送される際、そのメモリ内容のコピーが行われ、ネットワーク帯域を占有してしまう。さらに、各拠点間の往復遅延時間の値が大きい場合、実際の転送時間も長時間化する。よって、ライブマイグレーションは場合によりユーザエクスペリエンスに悪影響を及ぼす可能性がある。

限られた資源の有効活用のため、仮想マシンごとに割り当てる資源量および物理サーバへの資源の割り当て量は慎重に判断される必要がある⁽⁹⁾。CPU使用率、ディスク Input / Output (I/O) 遅延を指標とし、資源の充てん率を分析することにより、限られた物理マシン群にできるだけ多くの仮想マシンを配置可能な方式が提案されている。しかし、CPU使用率やディスク I/O 遅延の遷移は実際のシステム環境に依存する。

加えて、仮想マシンを運用するネットワークを仮想マシンが使用するテナント用、管理用およびストレージ用に分け、再配置を行う方式も提案されている⁽¹⁰⁾。しかし、我々は各拠点間がインターネット VPNによって接続された主に小規模な組織によって運用されるオンプレミス型のシステム構成を想定しているため、ネットワークの帯域が限られている。また、ユーザのサービス利用時間帯を考慮した仮想マシン・ディスク領域の割当て手法も検討されている⁽¹¹⁾。ユーザからサービスの利用時間帯に関する情報を取得することで資源を有効活用することができると述べられている。一方、実際に資源を消費するユーザのサービス利用時間帯と仮想マシン・ディスク領域の割当ては密接に関係しており、必ずしも資源を有効活用できるわけではない。場合によりこの手法は資源の有効活用に向く作用しない可能性がある。

以上の様に、既存の仮想マシン配置の最適化のライブマイグレーション手法は実際の状態を十分事前に検討していない。ゆえに、より現実的なライブマイグレーション手法

を我々は提案している⁽⁶⁾。引用文献⁽⁶⁾において、我々は仮想マシン最適配置法の提案を行い、一次評価を行っている。

3. 提案手法

〈3-1〉 提案手法の概要

(1) メモリ使用量と往復遅延時間、転送時間の関係

手法の提案を行う前の段階として、実際にライブマイグレーションを行い、仮想マシンのメモリ使用量と各拠点間の往復遅延時間、ライブマイグレーションに要する転送時間の関係を検証した。その際の実験環境を図 1、表 1 に示す。

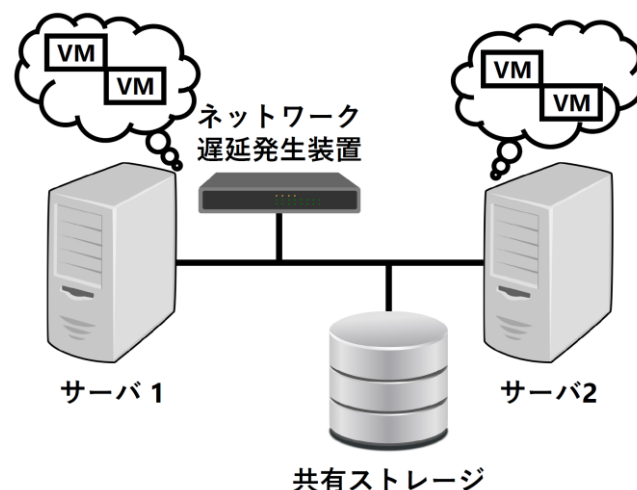


図 1 実験環境

Fig. 1. Experiment Environment

表 1 実験環境

Table 1. Experiment Environment

項目名	内容
ハイパーバイザ	Proxmox VE 5.1-35
物理サーバ	ML30 Gen9 NHP 4LFF
物理サーバのメモリ搭載量	32.00 [GiB]
仮想マシンの OS	Xubuntu 16.04 LTS
仮想マシン 1 台のメモリ搭載量	2.00 [GiB]
共有ストレージの種類	NFS Storage
ネットワーク帯域	1.00 [Gbps]
ネットワーク遅延発生装置	Ethdelay Pro NA-820

ネットワーク遅延発生装置を用いることでサーバ 1 およびサーバ 2 間に人工的に往復遅延時間を発生させた。また、実験環境において、各サーバおよび共有ストレージはカテゴリ 5e の LAN ケーブルによって接続されている。

図 2 は仮想マシンのメモリ使用量が 0.60 [GiB] の場合における往復遅延時間と仮想マシンの転送時間との関係を示している。

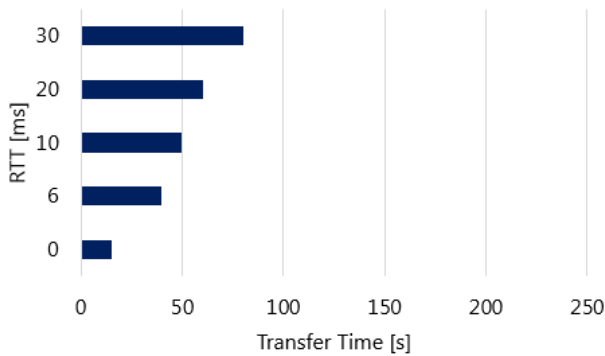


図 2 往復遅延時間と仮想マシンの転送時間との関係
(RAM usage: 0.60 [GiB])

Fig. 2. Preset RTT vs Actual Transfer Time

図 3 は仮想マシンのメモリ使用量が 1.00 [GiB] の場合における往復遅延時間と仮想マシンの転送時間との関係を示している。

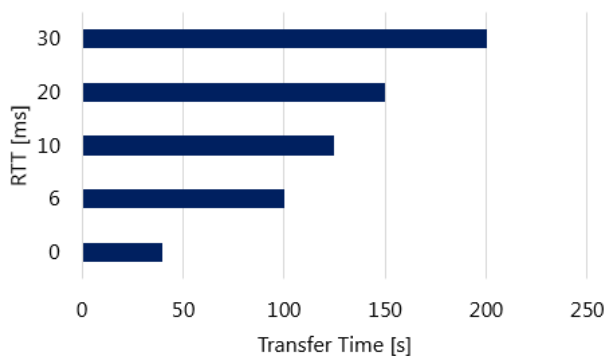


図 3 往復遅延時間と仮想マシンの転送時間との関係
(RAM usage: 1.00 [GiB])

Fig. 3. Preset RTT vs Actual Transfer Time

これらの結果から、6 ミリ秒の往復遅延時間の場合であっても、往復遅延時間が 0 ミリ秒の場合と比較して実際の仮想マシンの転送時間が 2 倍以上に長時間化することが伺える。さらに、仮想マシンのメモリ使用量が 0.60 [GiB] の際と 1.00 [GiB] の際を比較すると、メモリ使用量の増加に伴い、転送時間が同様に長時間化することが分かった。

(2) アルゴリズムの概要

予備実験の結果を踏まえ、我々は図 4 に示すライブマイグレーション手法を提案している。それぞれのステップの詳細は下記の通りである。

- Step 1 : 開始。
- Step 2 : 各拠点の物理サーバで仮想マシンが起動。
- Step 3 : 各物理サーバの仮想マシンによって消費されるメモリ使用量の合計を調べ、物理サーバの負荷を分析。
- Step 4 : 各物理サーバのメモリ使用量が閾値を超える場合、負荷が高い状態と定義し次ステップへ

進む。

- Step 5 : 本提案手法が想定するシステム構成はインターネット VPN であるため専用線と異なり、ネットワーク帯域に限られる。仮想マシンの転送時間を短縮化させ帯域への負担も考慮し、メモリ使用量の少ない仮想マシンを選出。
- Step 6 : 各拠点間の長い往復遅延時間は仮想マシンの転送時間を長時間化させる。よって、仮想マシンが稼働している現在の物理サーバと仮想マシンの転送先である新物理サーバ間の往復遅延時間が閾値を超えないことを確認。閾値を超えない場合、次ステップへ進む。
- Step 7 : ライブマイグレーション実行。仮想マシンを転送。
- Step 8 : 終了。

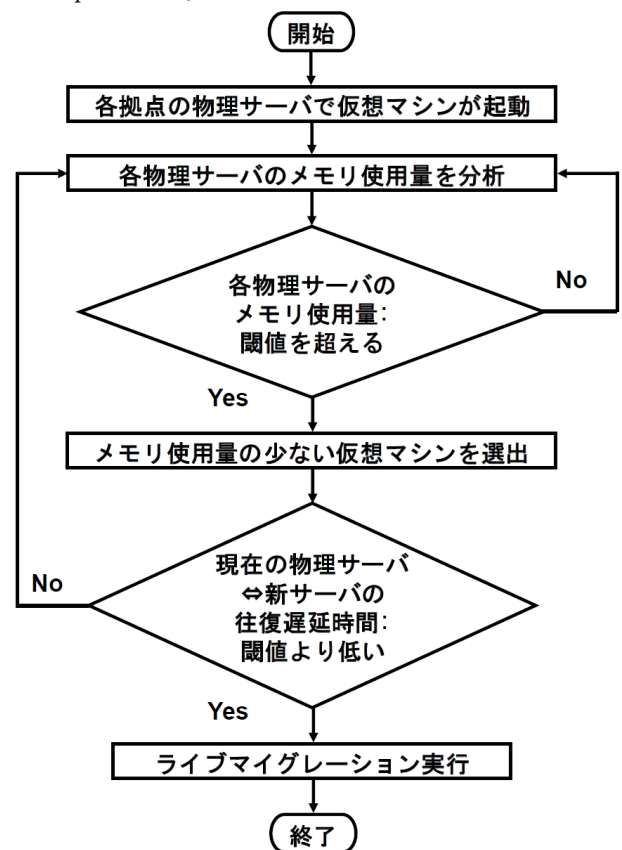


図 4 本提案手法のフローチャート

Fig. 4. A Flow Chart of Our Method

〈3・2〉 一次評価

引用文献⁶⁾において、我々は一次評価を行い、提案手法の有用性を示している。Python 3.6 を用いてシミュレータを作成し、仮想マシンの処理時間、待ち時間、ライブマイグレーションによる仮想マシンの転送回数との関係を明らかにした。表 2 はその際のシミュレーション環境を示す。各拠点間の往復遅延時間についてはランダムに設定し、30 [ms] の場合、仮想マシンの転送時間が長時間化するためライブマイグレーションを行わないものとした。

表 2 シミュレーション環境
Table 2. Simulation Environment

項目名	設定値
拠点数	3 か所
仮想マシン数	20 台
物理サーバのメモリ搭載量	12.00 [GiB]
仮想マシン 1 台のメモリ搭載量	2.00 [GiB]
仮想マシンが行う処理の種類	3 種類 (負荷小・中・大)
各拠点間の往復遅延時間	6・10・20・30 [ms]

表 3 はシミュレーション結果を示す。本稿において、「平均処理時間」とは、物理サーバで 1 台目の仮想マシンが起動、定められた処理を開始してから 20 台目の仮想マシンの処理が終了するまでの時間についてのシミュレーションによる計算を 5 回行った際の平均を指す。一方、「平均待ち時間」とは、物理サーバの空き容量が不足している状態により仮想マシンが起動を待つ時間についてのシミュレーションによる計算を 5 回行った際の平均を指す。ライブマイグレーションの実行により、平均処理時間はライブマイグレーションを行わない場合と比べ、約 6%短縮化された。これは、本提案手法により、物理サーバの負荷分散が行われたためであり、平均待ち時間は約 1/6 に短縮化された。

表 3 シミュレーション結果
Table 3. Simulation Results

	平均処理時間 (分)	平均待ち時間 (分)	平均転送回数 (回)
本提案手法	708.4	9.4	2.7
ライブマイグレーションなし	755.6	56.6	

4. 各拠点間の距離を勘案した再検討

〈4.1〉 再検討の概要

3 章において、我々の提案手法およびその一次評価について説明し、実際に手法の適用により仮想マシンの処理時間や待ち時間の短縮化が可能であることを証明した。一方、提案手法の一般化が必要であるため、想定する拠点数を増加させ、追加条件として各拠点間の距離に比例した往復遅延時間の設定を行うことにより再検討を行った。

〈4.2〉 アルゴリズムの概要

一次評価を踏まえ、我々は図 5 に示すライブマイグレーション手法を再検討した。それぞれのステップの詳細は下記の通りである。

- Step 1: 開始。
- Step 2: 各拠点の物理サーバで仮想マシンが起動。
- Step 3: 各物理サーバの仮想マシンによって消費されるメモリ使用量の合計を調べ、物理サーバの負荷を分析。

- Step 4: 各物理サーバのメモリ使用量が閾値を超える場合、負荷が高い状態と定義し次ステップへ進む。
- Step 5: 本提案手法が想定するシステム構成はインターネット VPN であるため専用線と異なり、ネットワーク帯域が限られる。仮想マシンの転送時間を短縮化させ帯域への負担も考慮し、メモリ使用量の少ない仮想マシンを選出。
- Step 6: 仮想マシンによって消費されるメモリ使用量の合計が少ない物理サーバを負荷が低い状態と定義。仮想マシンの転送先の物理サーバとして選出。
- Step 7: 各拠点間の長い往復遅延時間は仮想マシンの転送時間を長時間化させる。よって、仮想マシンが稼働している現在の物理サーバと仮想マシンの転送先である新物理サーバ間の往復遅延時間が閾値を超えないことを確認。閾値を超えない場合、次ステップへ進む。
- Step 8: ライブマイグレーション実行。仮想マシンを転送。
- Step 9: 終了。

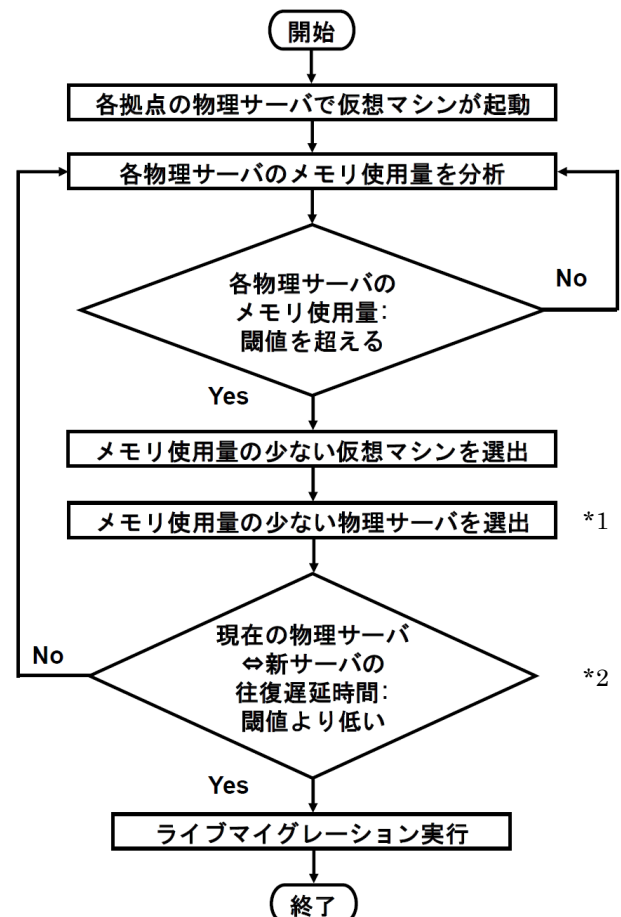


図 5 本提案手法を再検討したフローチャート
Fig. 5. A Revised Flow Chart of Our Method

- *1. 一次評価と異なり、3拠点以上の拠点数を想定しているため、適切な物理サーバを選択する処理を追加し、フローチャートの改良を行っている。
- *2. このステップにおいて、シミュレーション実行時には各拠点間の距離に比例した往復遅延時間の設定が行われる。

〈4・3〉 再評価

一次評価と同様、Python 3.6を用いてシミュレータを作成し、仮想マシンの処理時間、待ち時間、ライブマイグレーションによる仮想マシンの転送回数との関係を明らかにした。表4はその際のシミュレーション環境を示す。5か所の各拠点が一直線上に存在しているものとし、各拠点間の距離に比例した往復遅延時間の設定が表4に記載の往復遅延時間の候補からシミュレータの重み付きランダム計算により行われる。また、各拠点間の距離が遠い場合など、30 [ms]が設定された場合、一次評価を行った際と同様に仮想マシンの転送時間が長時間化するため、ライブマイグレーションを行わないものとした。

表4 シミュレーション環境

Table 4. Simulation Environment

項目名	設定値
拠点数	5か所
仮想マシン数	20台
物理サーバのメモリ搭載量	7.20 [GiB]
仮想マシン1台のメモリ搭載量	2.00 [GiB]
仮想マシンが行う処理の種類	3種類 (負荷小・中・大)
*各拠点間の往復遅延時間	6・10・20・30 [ms]

表5はシミュレーション結果を示す。ライブマイグレーションの実行により、平均処理時間はライブマイグレーションを行わない場合と比べ、約14%短縮化された。一次評価と同様、物理サーバの負荷分散が行われ、平均待ち時間は約1/2に短縮化された。

表5 シミュレーション結果

Table 5. Simulation Results

	平均処理時間 (分)	平均待ち時間 (分)	平均転送回数 (回)
本提案手法	906.4	123.2	3.6
ライブマイグレーションなし	1053.4	264.4	

5. おわりに

〈5・1〉 まとめ

本稿では、我々がすでに提案している各拠点間の往復遅延時間と仮想マシンのメモリ使用量を加味し、仮想マシンを最適配置するライブマイグレーション手法を踏まえ、追加条件として各拠点間の距離に比例した往復遅延時間の設定を行い、再検討を行った。結果として、平均処理時間は

提案手法の適用により、ライブマイグレーションを行わない場合と比較し、約14%短縮化された。平均待ち時間については、約1/2に短縮化された。シミュレーションによるすべての計算結果において、処理時間および待ち時間の短縮化が確認されたが、シミュレーション環境に依存する可能性があるために各パラメータの調整の検討が必要である。

〈5・2〉 今後の課題

今後の課題を以下に挙げる。1点目は、各物理サーバで仮想マシンが起動する間隔に対し、指数分布など確率分布の概念を取り入れることである。本稿では、仮想マシンが一定間隔で起動するという条件のもとでシミュレーションを行ったが、提案手法をより現実的に実現可能なものとするために必要である。2点目は、1点目に関連し、現実的な手法の提案のため、シミュレーションにおける各設定値など、さらなる要件の定義および整理が必要である。システムへの提案手法の実装には様々な課題が存在するため、今後はこれらを解決する必要がある。

文 献

- (1) <https://pubs.vmware.com>
- (2) K. Kushiyama, T. Hata, T. Ota, T. Hakata, M. Uruga, and A. Hiura: "Demonstration of live migration between distributed data centers", IEICE Technical Report, IN2011-39, pp61-66 (2011) (in Japanese)
- (3) Y. Nagafuchi, T. Kishi, T. Inoue, T. Koyama, and H. Kitazume: "Evaluation of virtual machine migration across remote datacenters", IEICE Technical Report, IN2012-40, pp43-48 (2012) (in Japanese)
- (4) R. Komatsu and S. Kuribayashi: "Impact of network quality and live-migration in virtual desktop environments", IEICE Technical Report, ICM2013-52, pp13-17 (2013) (in Japanese)
- (5) T. Nishijima, Y. Nakai, H. Ohsaki, N. Yokoi, K. Nakagawa, C. Lethanhman, Y. Nakamoto, Y. Takatani, and M. Imase: "On the Impact of Network Environment on Remote Desktop Protocols", IEICE Technical Report, CQ2012-21, pp23-28 (2012) (in Japanese)
- (6) K. Sakurai, H. Sano, H. Ishii: "A Proposal on Optimised Virtual Machine Placement by using Live Migration", the International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2018), pp294-297 (2018-7)
- (7) Y. Nagafuchi, Y. Teramoto, T. Kishi, T. Koyama, and H. Kitazume: "Proposal of the usage of the VM migration to reduce whole network delay in the live migration environment between datacenters", IEICE Technical Report, IN2013-88, pp7-12 (2013) (in Japanese)
- (8) S. Fukahori, A. Suzuki, and H. Shigeno: "Optimal placement method of Virtual Machine in accordance with Communication situation", The 76th National Convention of IPSJ, 3A-2, pp1-29-1-30 (2014-3) (in Japanese)
- (9) T. Ogawa, S. Tokunaga, S. Aiba, Y. Iida, H. Kitayama, S. Matsumoto, and R. Horisaka: "A Study of an Efficient Virtual Machine Deployment Strategy", IEICE Technical Report, ICM2009-30, pp23-28 (2009) (in Japanese)
- (10) S. Enomoto, T. Kimura, and T. Fujiura: "A Proposal of Virtual Machine Relocation Method Using Hierarchical Reinforcement Learning", IEICE Technical Report, ICM2018-2, pp27-32 (2018) (in Japanese)
- (11) H. Nakazato and T. Fujimoto: "Allocation Method of Virtual Machine and Disk Domain based on Service Use Time of Users", IEICE Technical Report, ICM2014-37, pp27-32 (2014) (in Japanese)