

広域分散仮想化環境における計算機資源利用状況に基づいた課金・フィードバック方式の評価

柏崎 礼生^{1,a)}

概要: 地理的に広域に分散した拠点間をネットワークで接続した仮想化計算機環境を実現した場合、仮想マシン (Virtual Machine: VM) はこの拠点間をマイグレーションするため、VM が使用した計算機リソースに応じた金額をそのリソースを提供した拠点に支払う課金モデルでは利用者の支払い処理が煩雑になり、利用のハードルを高くする欠点がある。そこで本研究では拠点で計測されたリソース使用情報を集約して課金額を決定し、集約された計算機利用料を各拠点に還元するモデルについて考察を行う。

キーワード: 広域分散, クラウドコンピューティング, 仮想化基盤, 課金, エコシステム

An Evaluation of an algorithm of accounting and feedback based on information of usage of computing resource on wide area distributed virtualized infrastructure

HIROKI KASHIWAZAKI^{1,a)}

Abstract: When a virtualization infrastructure which consist of widely distributed institutions connected each other by the Internet, it is so annoying for users to pay charge of VMs because VMs will be migration from one instution to another. The annoyance may make users leave the infrastructure. In this paper, I propose an algorithm of accounting and feedback based on information of usage of computing resource in every institutions.

Keywords: widely distributed, cloud computing, virtualization infrastructure, account charging, ecosystem

1. はじめに

仮想化ハイパーバイザの高性能化と CPU のマルチコア化が進む中、組織が持つ情報システムを仮想マシン (Virtual Machine: VM) 上で構築し動作させる仮想化基盤の構築とその利用が広範に普及している。代表的なパブリッククラウドコンピューティングサービスでは、多くの仮想化ハイパーバイザの VM イメージ形式を変換しインポートする機能を提供しているため、情報システムをパブリッククラウドコンピューティングサービスで動作させるためのステップとしても仮想化基盤の利用は有効である。教育研

究機関においては、静岡大学や北陸先端科学技術大学院大学 (JAIST)、北海道大学がクラウドコンピューティング環境を整備し IaaS (Infrastructure As A Service) や DaaS (Desktop As A Service) と呼ばれるクラウドコンピューティングサービスを提供している [1-4, 6]。教育研究組織におけるクラウドコンピューティングサービスの利用という側面においては、メールサービスをプライベートクラウドコンピューティング環境で構築した佐賀大学の事例の他 [5]、パブリッククラウドコンピューティングサービスのメール SaaS (Software As A Service) を利用している組織も多い。ここで用いるクラウドコンピューティングという用語については、Gartner, UC Berkeley, そして NIST による定義に基づき [7-9]、本稿では「仮想化技術を用いて

¹ 大阪大学サイバーメディアセンター
Cybermedia Center, Osaka University

^{a)} reo@cmc.osaka-u.ac.jp

実現されるスケールアウト可能な基盤の上に構築された、規模を取縮可能なサービス」の意味で用いることとする。

しかし SaaS のメールサービスについては訴訟が発生した場合の管轄裁判所が海外となる場合がある。米国企業がサービスしている SaaS であっても、サービスの利用組織が存在する国の法を適用するという約款となっているものもあるが、サービス提供企業が米国企業である場合は、米国愛国者法によりたとえそのサービスが扱うデータが米国外のデータセンターに存在していたとしても、米国の捜査機関が裁判所の令状なしにリモート操作でクラウド上のデータを複製することが許されている。結果として、法的に安心できる SaaS のメールサービスとは国内事業者が国内データセンターで稼働させる SaaS という事になるが、国内で複数のデータセンターにまたがりサービスを提供する SaaS のメールサービスというものは現在存在しないと思われる。地震をはじめとする自然災害に事欠かない日本において、高い可用性とサービス継続性と実現した SaaS のメールサービス IaaS, VPS サービスは、海外のそれらに対して価格競争力で劣ってしまう。

現在、広島大学、金沢大学、国立情報学研究所、大阪大学を中心として複数拠点からなる地理的に広域に分散した環境で分散ストレージを実現し、これを仮想化基盤で利用することにより広域のライブマイグレーションを実現し、高い可用性とサービス継続性を実現する仮想化基盤を構築する研究開発を行っている [13-15]。将来的には、十分にサービス断時間の短いライブマイグレーションを実現し、高可用性を実現するための仕組みが整うことにより、仮想化サーバの障害や拠点の停電、回線の切断といった唐突な障害が発生してもサービス不通時間の短い広域分散仮想化基盤を実現することができる。広域ライブマイグレーションに伴うサービス不通時間が十分に短くなれば、ユーザや運用担当者が VM がどの拠点、どの仮想化サーバ上で稼働しているかを意識することのない仮想化基盤を実現することができる。その前提の上では、ユーザあるいは VM 上に存在するサービスが、自分が保有する VM あるいは VM 自身をどの拠点、どの仮想化サーバ上で稼働するかを明示的に指定することができる基盤を実現することも技術的に困難ではない。これにより、VM 上で稼働するサービスにアクセスする複数のクライアントから情報を収集し、複数のクライアントとの通信応答時間を最小化し得る拠点に VM をマイグレーションする仕組みを実現することも可能となる。大規模な計算処理を実行したいユーザは、より計算機の空き資源が大きい仮想化サーバに VM をマイグレーションすることも可能となる。またその観点からは、頻繁に仮想化サーバがダウンする拠点は信頼性を失い、ユーザが利用しなくなるという事が想定される。

この広域分散仮想化基盤は研究開発段階ということもあり、研究協力組織であれば現在は無料で使うことができる

が、無料で計算機資源を提供し続けていては持続可能性に乏しい。研究資金を獲得し続け常に新しい研究開発課題を発見することで、永遠に研究開発途上のテストベッドとするモデルも考えられるが、健全な運用とは言い難い。国内の研究組織が構築するクラウドコンピューティングサービスや大型計算機サービスでは利用者に課金をしており、得られた収入をサービスのランニングコストや運用経費の一部として利用していると考えられる。しかし前述のように夢想する広域分散仮想化環境においては VM がどの拠点のどの仮想化サーバで稼働しているか、ユーザは意識しているにせよしていないにせよ、計算機資源を提供する拠点に対して利用料金を支払うモデルでは、支払いが煩雑になり過ぎて積極的な利活用を阻害する要因となり得る。複数データセンターをまたぐ大規模なクラウドコンピューティング環境においてマイグレーションを用いた消費電力削減手法が提案されているが [16]、これらの研究では一事業者による均一な運用レベルで構成された環境での検討であり、またマイグレーションをするのがユーザの意思ではなくクラウドコンピューティングサービス事業者の判断によるものである点において広域分散環境への適用が困難と言える。本研究では、参加組織を抽象化した機関が課金と集金を行い、実態に沿った計算機利用状況に応じて各拠点に利用料金への支払いを代行するモデルを提案する。シミュレーションを用いた評価により、提案モデルの優位性と問題点を論じる。

2. 広域分散仮想化基盤のモデル化

仮想化基盤における収益性について、図1のように考える。仮想化基盤の事業者はユーザへの課金による収益を得て仮想化サーバの増強を行う。増強された仮想化サーバ群は VM を提供し、ユーザへの課金とこの VM のユーザの所有が結びつく。事業者は VM のコア数は開示するか提供周波数については開示しない。事業者が VM に提供できる周波数とコア数の乗算結果は他のユーザの利用率などにより変動する。ユーザは VM 上にサービスを作成しエクスペリエンス (経験による対価) を得るが、そのサービスが実際に仮想マシンに要求しているリソース量をつぶさに観測しているユーザもいれば、していないユーザもいる。観測していないユーザや、観察していてもその観測結果を適切に解釈できなければ課金される額を減らす (VM のプランのグレードを下げる) という判断が出来ずに余剰リソースに金額を支払い続ける。ユーザが自分が構築したサービスに対する妄想のリソース要求と現実のサービスが仮想マシンに対して要求するリソース要求の差が事業者の利益となり、これにより得られた収益の一部が設備投資に回される。

広域分散仮想化基盤を構成する利害関係を考察し、広域分散仮想化基盤の構成要素を以下の6要素に抽象化する。

- 仮想化基盤管理センター

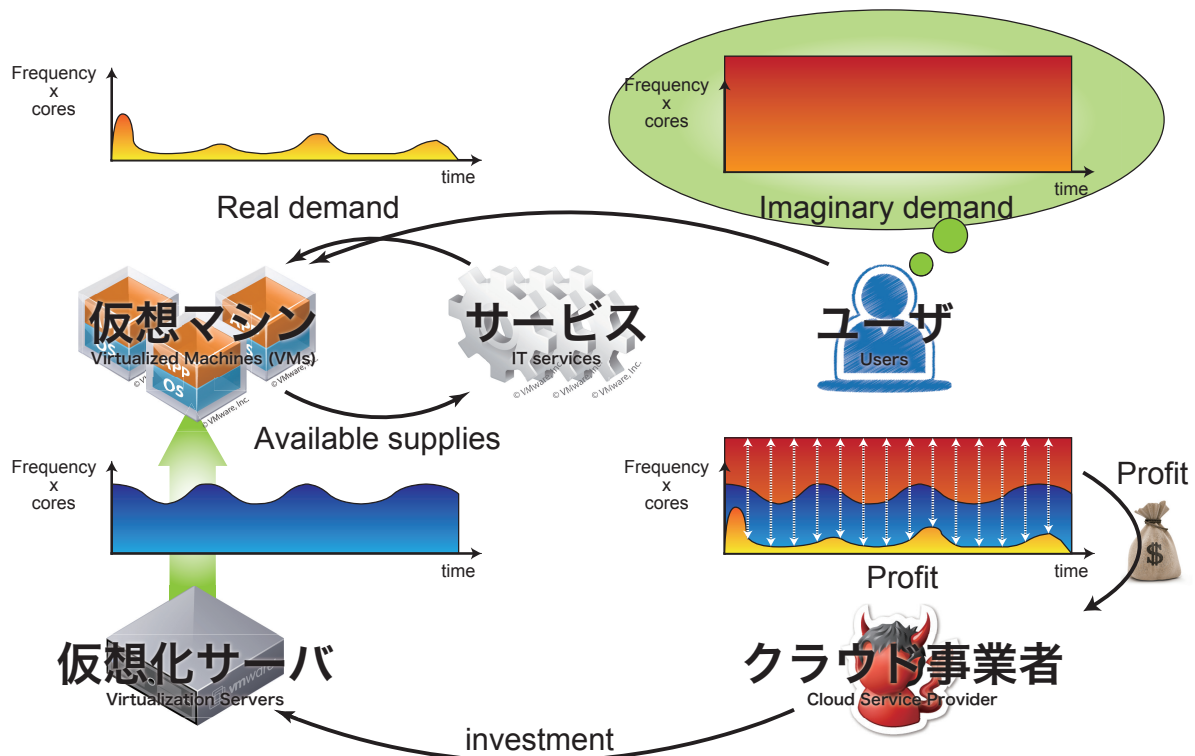


図 1 仮想化基盤の経済学の一側面
Fig. 1 One economical aspects of virtual infrastructure

- 拠点
- 仮想化サーバ
- VM
- サービス
- ユーザ

図 2 は 6 要素が相互にどのように作用するかを表した模式図である。以下の節で各要素について説明する。

2.1 仮想化管理センター

広域分散仮想化基盤に参加する組織を横断的に監視し、ユーザにサービスを提供するインターフェイスとなるのが仮想化管理センターである。仮想化管理センターの役割と特徴は以下の通りである。

- 仮想化管理センターは、各拠点が持つ仮想化サーバを拠点横断的に管理する。
- 仮想化管理センターは、ユーザからの VM 作成・変更・削除要求に対応し、要求に応じた VM の作成を適切な仮想化サーバに要求する。
- 仮想化管理センターは、ユーザからの VM 移動要求に対応し、要求に応じた VM のマイグレーションを行う。
- 任意の仮想化サーバが電源 ON 状態から電源 OFF 状態に変わった時、その仮想化サーバ上にある全ての VM を他の適切な仮想化サーバ上にマイグレーションさせる。
- マイグレーションは 0 秒より大きいサービス途絶時間

を VM に与える。

- 各仮想化サーバが使用している計算機リソース情報を集約する。
- 集約された計算機リソース情報に従い、各拠点到バックする金額を決定する。

2.2 拠点

広域分散仮想化基盤に参加する組織が持つデータセンターが拠点である。拠点は回線使用料や電力使用量を支払い、設備投資をして、仮想化基盤を運用する。拠点の役割と特徴は以下の通りである。

- 拠点は仮想化サーバに電力を提供する。
- 拠点は回線使用料、電力使用量、設備費用により構成される設備投資を行う。
- 拠点は複数の仮想化サーバを持つことができる。
- 拠点は仮想化サーバの電源を ON にする、OFF にする権限を有する。

2.3 仮想化サーバ

拠点が持つ仮想化サーバ上には VM が存在し、仮想化サーバは電力と引き替えに VM に計算機資源を提供することができる。仮想化サーバの役割と特徴は以下の通りである。

- 仮想化サーバは状態として電源 ON と電源 OFF を持つ。

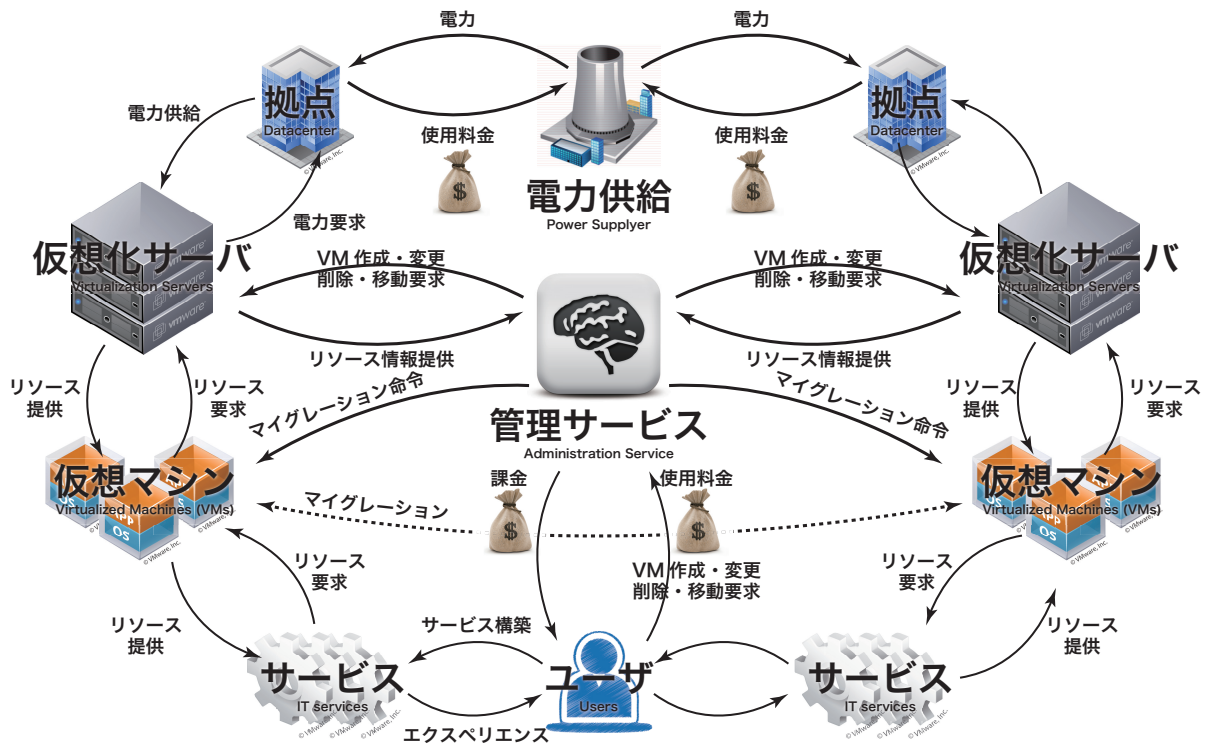


図 2 広域分散仮想化環境を構成する要素の関係模式図

Fig. 2 A correlation diagram of elements which widely distributed virtualization infrastructure consists of.

- 仮想化サーバは電力を拠点に対して要求し、拠点は電力を仮想化サーバに提供する。
- 仮想化サーバと拠点の間の電力の需給は遅延を発生させずに行われる。
- 仮想化サーバはその能力に応じて VM にリソースを提供することができる。
- 仮想化サーバの能力は、メモリ容量、コア数、周波数、ストレージ容量によって決定される。
- 仮想化サーバは複数の VM を持つことができるが、その能力を超えたリソースの提供はできない。
- センターからの VM 作成要求には有限時間を要する。
- 仮想化サーバは与えられた確率に従い故障 (電源断) する。

2.4 仮想マシン (VM)

仮想化サーバ上の VM はユーザが仮想化管理センターに作成を要求した時に作られ、削除を要求した時に消失する。そのスペックはユーザが変更を要求した時に変更される。VM の役割と特徴は以下の通りである。

- VM は状態として電源 ON と電源 OFF を持つ。
- VM は仮想化サーバに対してリソースを要求し、仮想化サーバはリソースを VM に提供する。
- VM と仮想化サーバの間のリソースの需給は遅延を発生させずに行われるが、仮想化サーバへ要求したリソ

スが必ず提供されるとは限らない。

- VM はその能力に応じてサービスにリソースを提供することができる。
- VM の能力は、メモリ容量、コア数、周波数、ストレージ容量によって決定される。
- VM は複数のサービスを持つことができるが、その能力を超えたリソースの提供はできない。
- VM は仮想化管理センターによりマイグレーションをされる。マイグレーション中に 0 秒より大きい途絶時間が発生する。

2.5 サービス

VM 上のサービスはユーザが作成し、サービスはユーザの要求を代弁し、自動化することができる。VM の役割と特徴は以下の通りである。

- サービスは VM にリソースを要求し、VM はサービスにリソースを提供する。
- サービスと VM の間のリソースの需給は遅延を発生させずに行われるが、VM へ要求したリソースが必ず提供されるとは限らない。
- サービスが VM に要求するリソースは、メモリ容量、スレッド数 (コア数)、ストレージ容量によって決定される。
- サービスが VM に要求するリソースは周期的に変化

する。

- サービスはユーザに対してエクスペリエンスを提供する。
- エクスペリエンスは実際に使用されたメモリ容量，コア数×周波数，ストレージ容量の累計で表現される。
- サービスはユーザの要求を代弁することができるので，仮想化管理センターに対してマイグレーション要求をすることができる。

2.6 ユーザ

ユーザは使用料金を支払い，VM を所有し，VM の上でサービスを構築して，サービスが実現するエクスペリエンスを期待する。ユーザの役割と特徴は以下の通りである。

- ユーザは仮想化管理センターに使用料金を支払い，VM を所有することができる。
- ユーザは仮想化管理センターに VM のスペックの変更を要求することができる。
- ユーザは仮想化管理センターに VM が稼働する拠点もしくは仮想化サーバを指定することができる。
- ユーザは VM 上で複数のサービスを構築することができる。
- ユーザは VM 上のサービスからエクスペリエンスを得ることを期待する（メモリ容量，スレッド数×コア数，ストレージ容量の潜在的期待値が存在する）。
- ユーザは満足度閾値があり，サービスから得られるエクスペリエンスと潜在的期待値の割合がこの閾値を下回るとマイグレーションを行う。

これらの要素の要求と提供の関係により，広域分散仮想化環境を模擬することができる。

3. シミュレータの構成

要素の要求と提供を 1 sec 単位で模擬することにより，この仮想化環境の振る舞いを模擬する。多様な構成要素により成立する環境を模擬するため，拠点およびサービスには以下のような複数のポリシーのうちいずれかを反映させる。拠点は「年に連続する n 日の電力供給断を m 回行う」という年周期性のある運用ポリシーに準じる。 n は 1~3， m は 1~2 で与えられる整数値とする。サービスは以下に示す日周期性，週周期性，年周期性の乗算によって与えられるものとする。日周期性については，

- l 秒間負荷が F_h となり，他の $86400 - l$ 秒間は負荷が F_l となる矩形波
- 周期 T 秒で振幅が F_h の正弦波

の 2 パターンであり，週周期性については上記の日周期性を D_w 日間繰り返した後で $7 - D_w$ 日間は R_w を日周期性に乗した値とする。 $D_w = 7$ の時は一週間変わらない周期性を持ち， $D_w = 5$ などとすると週末は負荷の低いサービスとなる。年周期性についても同様で， D_y 日間は 1 を乗

し，その他の日は R_y を乗することにより季節的に負荷の高いサービスを表現する（図 3）。

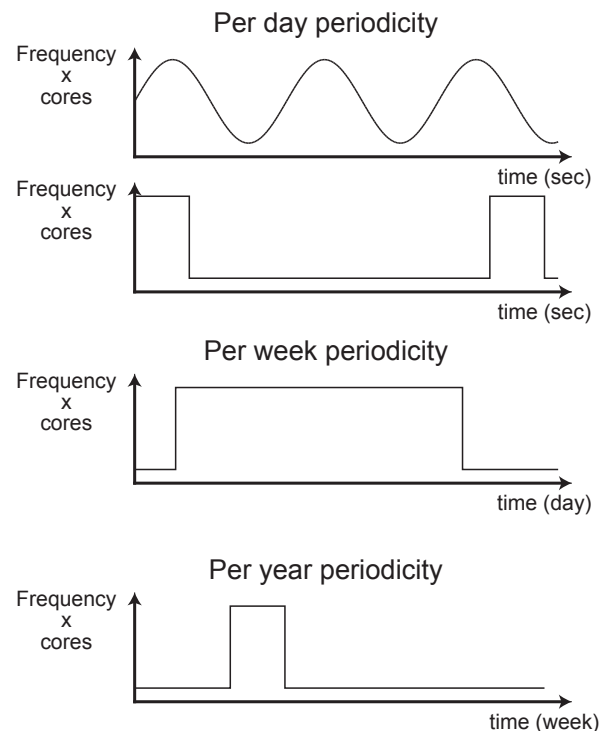


図 3 多様なサービスの周期性の表現
Fig. 3 Expressions of various periodicity of services

ユーザは $[1, n]$ の間の整数値の VM を購入し， $[1, m]$ の間の整数値のサービスを 1 つの VM の上に構築する。整数値は一様分布に従って決定する。周期性を決定するパラメータについてもそれぞれ与えられた区間内での一様分布に従った値を与えて各ユーザが各 VM に構築するサービスが決定される。各 VM は周期に従ってリソースを要求し，仮想化サーバは仮想マシンにリソースを提供する。仮想化サーバは仮想マシンから要求されたりリソースの総和 R_r が仮想化サーバが提供できるリソース R_s よりも大きかった場合は全ての仮想マシンには要求リソース量の $\frac{R_s}{R_r}$ 倍のリソースを提供する。リソースを提供された仮想マシンは各サービスに対して同様に減じたリソースを提供する。このリソースと，各ユーザごとに設定された潜在的期待値の差が満足値閾値を下回るとユーザは他拠点あるいは他仮想化サーバへのマイグレーションを管理サービスに要求する。管理サービスが収集した各仮想化サーバのリソース情報はユーザからも可視であり，あらかじめ定められた優先度に応じて移動先を決定する。ユーザが希望したマイグレーションおよび拠点の運用による仮想化サーバの電源断によるマイグレーションの双方においてサービスは与えられた秒数の間，サービス断状態となる。この間，ユーザが受け取るエクスペリエンスは 0 となる。

このようにユーザが受け取るエクスペリエンスに従って

マイグレーションをするアルゴリズムを適用すると、仮想マシンに偏りが見られるようになる。そこで拠点において粒度の細かい運用をする拠点と、荒い運用をする拠点という特性を与える。粒度の細かい運用をする拠点は一週間単位での全仮想化サーバの利用率を観測し、それに応じて翌週に稼働させる仮想化サーバの数を決定する。荒い運用をする拠点においては年単位で決定をする。細かい運用をする拠点においては、仮想化サーバの月単位の平均利用率が閾値を超えた場合、仮想化サーバを月単位で増強するものとする。荒い運用をする拠点においては増強は平均利用率の算出と増強を年単位で行う。

管理サービスはユーザから得られた利用料金からマージンを引き、残った金額を各拠点が持つ仮想化サーバの稼働率に応じてキャッシュバックする。年単位で各拠点が増強に要した金額および電力使用料の総計とキャッシュバックされた利用料金の値を比較することで、どのような拠点戦略がこの仮想化基盤において有効であるかについて示すとともに、より長い年月を経過させた場合どのような変化が生じるかについて評価を行うことができる。

4. まとめ

本研究では現在構築中の広域分散仮想化基盤において適用可能な、ユーザが自由にマイグレーションすることができるとともにユーザが意識せずともマイグレーションが行われる仮想化基盤における課金方式とキャッシュバックモデルについて解説を行った。このモデルにおける各拠点の戦略の有効性を検証するためのシミュレーションモデルについて示した。現在シミュレータは実装中であり簡単な評価実験とその結果を出力できる段階にある。今後提案手法に沿ってシミュレータの実装を進めるとともに、例えば各拠点が増強の際に購入できる仮想化サーバのスペックに多様性を持たせるといった各抽象化要素の行動様式の多様化、またより長時間をかけたシミュレーション評価を行うため、現実に存在しない未来技術をムーアの法則 [17] になぞらえて CPU のコア数、メモリ容量、ネットワークバンド幅の価格帯を推定しシミュレータに反映させるといった手法を導入することが今後の課題である。

謝辞 本研究は平成 25 年度北海道大学情報基盤センター共同研究「インタークラウド環境での広域分散ストレージ実験と実証」による支援を受けました。

参考文献

[1] 柏崎礼生: スモールスタートで始める大学の仮想化基盤の構築と運用の実情, インターネットと運用技術シンポジウム 2012 論文集, pp.94-101 (2012).
[2] 坂田智之, 長谷川孝博, 水野信也, 永田正樹, 井上春樹: 情報セキュリティの観点からみた静岡大学の全面クラウド化, 情報処理学会研究報告, 2011-IOT-14, Vol.7, pp.1

(2011).
[3] 松原義継, 大谷誠, 江藤博文, 渡辺健次, 只木進一: プライベートクラウドによる電子メール管理コストの低減とサービスレベルの改善 —佐賀大学の事例—, 情報処理学会研究報告, 2011-IOT-14, Vol.8, pp.1-6 (2011).
[4] Shikida Mikifumi, Miyashita Kanae, Ueno Mototsugu, Uda Satoshi: An evaluation of private cloud system for desktop environments, Proceedings of the ACM SIGUCCS 40th annual conference on Special interest group on university and college computing services (SIGUCCS '12), pp.131-134 (2012).
[5] 宮下夏苗, 上埜元嗣, 宇多仁, 敷田幹文: 大学におけるプライベートクラウド環境の構築と利用, 第 3 回インターネットと運用技術シンポジウム, pp.17-24 (2010).
[6] 棟朝雅晴, 高井昌彰: 北海道大学アカデミッククラウドにおけるコンテンツマネジメントシステムの展開, 第 10 回情報科学技術フォーラム 情報科学技術レターズ pp.15-18 (2011).
[7] Daryl C. Plummer, Thomas J. Bittman, Tom Austin, David W. Cearley and David Mitchell Smith: Cloud Computing: Defining and Describing an Emerging Phenomenon, Gartner Research, G00156220 (2008).
[8] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica and Matei Zaharia: Above the Clouds: A Berkeley View of Cloud Computing, UCB/EECS-2009-28 (2009).
[9] Lee Badger, Tim Grance, Robert Patt-Corner, Jeff Voas: DRAFT Cloud Computing Synopsis and Recommendation, NIST Special Publication 800-146 (2012).
[10] Jeffrey Shafer: I/O virtualization bottlenecks in cloud computing today, Proceedings of the 2nd conference on I/O virtualization (WIOV'10), pp.5-5 (2010).
[11] S. Mikami, K. Ohta, O. Tatebe: Using the Gfarm File System as a POSIX Compatible Storage Platform for Hadoop MapReduce Applications, Grid Computing (GRID), 2011 12th IEEE/ACM International Conference on, pp.181-189 (2011).
[12] Sanjay Ghemawat, Howard Gobioff, Shun-Tak Leung: The Google file system, Proceedings of the nineteenth ACM symposium on Operating systems principles (SOSP '03), pp.29-43 (2003).
[13] 中川郁夫, 橋本好史, 楠田友彦, 北口善明, 近堂徹, 柏崎礼生, 市川昊平, 下條真司: 大学間連携による頑強な広域分散データ基盤アーキテクチャの提案, 情報処理学会研究報告インターネットと運用技術, Vol. 2013-IOT-20, No. 20, pp.1-6 (2013).
[14] 柏崎礼生, 近堂徹, 北口善明, 楠田友彦, 大沼善朗, 中川郁夫, 市川昊平, 棟朝雅晴, 高井昌彰, 阿部俊二, 横山重俊, 下條真司: 広域分散ストレージ検証環境における I/O 性能評価, 情報処理学会研究報告インターネットと運用技術, Vol. 2013-IOT-20, No. 19, pp.1-6 (2013).
[15] 北口善明, 近堂徹, 柏崎礼生, 中川郁夫, 下條真司: 広域分散ストレージを用いた長距離ライブマイグレーションの評価実験, 信学技報, vol. 113, no. 94, IA2013-7, pp. 37-42 (2013).
[16] Toshiaki Suzuki, Shinichi Kuwahara: A Large-scale Power Saving Cloud System Composed of Multiple Data Centers, Proc of The Third International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies (Energy2013), pp.127-133 (2013).
[17] Gordon E. Moore: Cramming more components onto integrated circuits, Electronics, Volume 38, Number 8 (1965).