

# 定量的な評価に基づく組織内仮想化基盤の増強設計

柏崎 礼生<sup>1,a)</sup> 藤本 祥人<sup>1,b)</sup> 宮永 勢次<sup>1,c)</sup> 森原 一郎<sup>1,d)</sup>

**概要:** 組織内に散在する計算機を集約する目的で構築された仮想化基盤は、必要とされる計算機資源の増大やベンダーによるサポート期限を契機として、その更改が必要となる場合がある。その更改における定量的な設計の根拠となるパラメータを示し、増強の導入とサービス運用のコストについて、100VMs程度が動作する仮想化基盤における例示を用いて評価を行う。

## Reinforcement designs of a virtualization infrastructure in a organization based on quantitative evaluations

HIROKI KASHIWAZAKI<sup>1,a)</sup> FUJIMOTO YOSHITO<sup>1,b)</sup> SEIJI MIYANAGA<sup>1,c)</sup> ICHIROU MORIHARA<sup>1,d)</sup>

**Abstract:** Sometime it may be necessary for a virtualization infrastructure that is designed to server virtualization and consolidation to reinforcement its computing resources triggered by a growth of its requirement and limitation of its support time. This paper shows parameters that are required to design quantitative reinforcement and evaluate its cost of implementation and operation referring to an example of a virtualization infrastructure supporting about 100VMs.

**Keywords:** virtualization infrastructure, quantitative design, quantitative evaluation

### 1. 背景と目的

少子高齢化とそれに伴う人口減少が労働投入の減少や国内需要の縮小を招き、中長期的な経済成長を阻害すると懸念される先進国において、情報通信技術 (Information and Communication Technology) が社会活動の解決に貢献することが期待されている\*1。ICTを活用したグリーン化は2010年以前から注目されており\*2、組織における電力使用量の削減や施設空間の有効活用を目的として組織が所有する情報システム (以下、簡潔な表現のため「組織が所有す

る情報システム」を「情報システム」と表す) を集約するための組織内仮想化基盤 (以下、簡潔な表現のため「組織内仮想化基盤」を「仮想化基盤」と表す) の構築もグリーン化を目的とする具体例の一つである。

情報システムの集約は段階的に行われるため、仮想化基盤の導入初期では情報システムが要求する計算機資源の総量は大きくない場合があり、要求総量は次第に増大する。そのためサーバ集約を目的とする仮想化基盤の初期導入ではスモールスタートによる設計にコストメリットがある [1]。これまで拡張性を考慮した仮想化基盤の構築に関する事例や、仮想化基盤をネットワーク処理に利用する事例が報告されているが、その基盤の撤退戦略について言及しているものは乏しい [2,3]。集約された情報システムは永遠に仮想化基盤上で動作し続けるわけではなく、以下のような理由から仮想化基盤の利用を終了する場合がある。

- パブリッククラウドへの移行
- オンプレミスへの回帰
- 情報システムのサービス終了  
終了への過渡期においてはデータセンターの利用による

<sup>1</sup> 大阪大学  
Osaka University

a) reo@cmc.osaka-u.ac.jp

b) fujimoto-yo@office.osaka-u.ac.jp

c) miyanaga-s@office.osaka-u.ac.jp

d) morihara@cmc.osaka-u.ac.jp

\*1 総務省: 平成 28 年度版情報通信白書  
<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/>

\*2 総務省: 平成 22 年度版情報通信白書  
<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h22/>

オンプレミスとパブリッククラウドの混合利用という選択肢もある [4]。設置場所が自組織内であろうとデータセンターであろうとオンプレミスである限り、仮想化基盤を構成するハードウェアおよびソフトウェアにはサポート期限があり、このサポート期限までに全ての情報システムがこの仮想化基盤の利用を終了しないのであれば、仮想化基盤はサービスを継続して提供するため、増強設計が必要となる。本稿では既にサービスの提供が始まっている仮想化基盤の増強更新に必要なパラメータを示す。このパラメータに基づいて設計された仮想化基盤の実例を示し、増強設計の妥当性の検証とそのコスト評価を行う。

## 2. 提案する設計手法

仮想化基盤は計算機資源を提供する仮想化ハイパーバイザ、ストレージ機器、ネットワーク回線、および電源供給の4要素により構成され得る。2016年現在、ネットワーク機能仮想化技術 (Network Function Virtualization: NFV) やソフトウェアにより定義されるストレージ技術 (Software Defined Storage: SDS) が製品として登場しているが、本提案では前述の4要素からなるものとして設計を行う。

### 2.1 仮想化ハイパーバイザの設計

仮想化ハイパーバイザは、中央演算装置 (Central Processing Unit: CPU) と主記憶 (メモリ) を提供するハードウェアと、仮想化ハイパーバイザ機能を提供するソフトウェアから構成される。CPU は物理コア数が多いほど高額であり、同じ物理コア数で比較すると動作周波数が高いほど高額であり、かつ低消費電力であるほど高額である。仮想化ハイパーバイザソフトウェアは仮想計算機 (virtual machine: VM) の仮想 CPU (virtual CPU: vCPU) が要求する命令を物理コアに対して割り当てる。コンテキストスイッチにより複数の vCPU が要求する命令を1つの物理コアに割り当てることができる。コンテキストスイッチのオーバーヘッドは無視できないが、1物理コアに対する vCPU 数の比は1以上とすることにより物理コアの利用率を高めることができる。これを CPU のオーバーサブスクリプションと呼ぶ。1物理コアに割り当てられた複数の vCPU の命令とそのオーバーヘッドがその物理コアの演算能力を超えた時、VM から観測すると vCPU が提供する演算能力は低下したように見える。

既にサービスの提供が始まっている仮想化基盤において、現在必要とされる vCPU 数は既知である。スモールスタートで仮想化基盤のサービスを開始した場合、開始直後から増強設計を開始するまでの間で、仮想化基盤の利用を開始した情報システム、およびその情報システムが要求する vCPU の数の時間的変化は把握することができる。この時間的変化を線形補間して、今後の一定期間に要求されることが想定できる vCPU の数を見積もることも可能である。

しかし組織内に存在する情報サービスは無限ではなく有限である。情報サービスには、仮想化基盤が提供するサービスと同様、情報サービスを構成するハードウェア、ソフトウェア、および開発と運用のサポートの期限がある。そのため情報サービスが仮想化基盤の利用を開始するのは、そのサポートが有効な間ではなく、サポート期限を迎えるタイミングであるケースが想定できる。

これら既に稼働している実際値から得られた一定期間後に要求されるであろう vCPU 数と、この一定期間内に仮想化基盤の利用を開始する予定のある情報システムが要求する vCPU 数から得られる想定要求総 vCPU 数の間には乖離がある。この乖離は、想定していなかった仮想化基盤の利用に対応できるための緩衝となる。緩衝は大きく見積もりすぎれば無駄な計算機資源となる (図1)。CPU はこのようにオーバーサブスクリプションを前提とした設計を行うが、メモリのオーバーサブスクリプションがもたらす VM のパフォーマンスの劣化は無視できないため、物理メモリの容量は各情報システムを構成する VM が要求するメモリ量と一致させる。

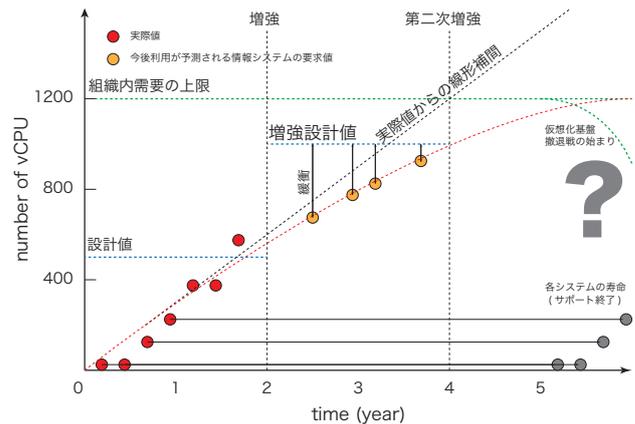


図1 vCPU 数の見積もり

Fig. 1 Estimation of a number of vCPUs

これらの定量的な判断基準から一定期間後に要求される vCPU 数の推定値を決定することができれば、1物理コアに対して割り当てる vCPU 数を定めれば増強設計における総物理コア数を決定することができる。前述の通り1CPUあたりの物理コア数が大きいほど高額ではあるが、仮想化ハイパーバイザソフトウェアは、そのソフトウェアが稼働する CPU のソケット数に応じてライセンス費用を要求される場合がある。ライセンス費用は仮想化ハイパーバイザが動作する間、定常的に要求される金額であるため、この金額の総額とプロセッサの費用を比較する。また、組織においてはハウジング費用を要求される場合があり、専有するラック数を少すことで定常費用を低減することが可能である。ハウジング費用には電力利用料が含まれているものとし、本稿では CPU や計算機が消費する電力の利用

料金は無視できるものとする。本稿執筆時点においては 1 ラックユニット (Rack Units: RU) あたり 8 ソケットの CPU を動作させる計算機が販売されている\*3。

ここまでの議論をまとめる。要求される総物理コア数を  $C$ 、1CPU でまかなうコア数を  $p$ 、コア数  $p$  の CPU の単価を  $u_p$  とすると、要求される総物理コア数を満足するための CPU に要する値段  $V_c$  は

$$V_c = \frac{C}{p} \times u_p \quad (1)$$

で表すことができる。一方、仮想化ハイパーバイザソフトウェアの 1 ソケット・1 年あたりのライセンス費を  $l$  とし、仮想化基盤の稼働予定期間を  $y$  年とすると、 $y$  年間で要する仮想化ハイパーバイザソフトウェアのライセンス費用  $V_l$  は、

$$V_l = \frac{C}{p} \times l \times y \quad (2)$$

で表すことができる。1RU に集約する CPU の数を  $r$  とし、その組織における 1RU・1 年あたりのハウジング費用を  $u_h$  とし、仮想化基盤の稼働予定期間を前述の通り  $y$  年とすると、 $y$  年間で要するハウジング費用  $V_h$  は、

$$V_h = \frac{C}{p} \div r \times u_h \times y \quad (3)$$

で表すことができる。ゆえに増強設計時における市場調査を行い、複数のコア数  $p$  に対応する CPU の単価  $u_p$  の組み合わせを得れば、仮想化ハイパーバイザの増強設計に要する総額  $V$  は、

$$V = V_c + V_l + V_h = \frac{C}{p} \left( u_p + \left( l + \frac{u_h}{r} \right) \times y \right) \quad (4)$$

で表すことができる。変数は  $p$  および  $u_p$  なので  $V$  は 2 変数からなる関数  $V(p, u_p)$  である。この関数に従い  $u_p$  の離散点をプロットすることにより最小の  $V$  を導く 1CPU でまかなうコア数  $p_{min}$  を得ることができる。

## 2.2 ストレージ機器の設計

仮想化ハイパーバイザはストレージ機器をデータストアとして認識し、VM が要求する補助記憶装置 (ストレージ) はこのデータストア上に仮想マシンディスクファイルとして確保される。ストレージ機器の設計においては、VM が要求する総ストレージ容量と、VM が要求する単位時間あたりの総 I/O 性能 (IOPS) の双方を満たす必要がある。

ストレージ機器は耐障害性の向上を目的として、機器を

構成する複数のドライブで冗長構成を採用する。RAID5 はその RAID グループを構成するドライブの中の同時に 1 台のドライブの故障に対してパリティからの復元を可能とし、RAID6 は同時に 2 台のドライブの故障に対して復元を可能とする。RAID グループを構成するドライブ数構成について RAID5 は 4+1 (5 台で構成され、1 台がパリティ用) か 8+1、RAID6 は 6+2 か 8+2 が推奨されている\*4。1 ドライブあたりの IOPS はハードディスクドライブ (HDD) の場合、7,200 rpm で 75-100 IOPS、10,000 rpm で 125-150 IOPS、15,000 rpm で 175-210 IOPS という参考値が用いられる\*5。RAID では Write が行われる際に RAID のレベルに応じた処理の複雑さがあるため、RAID レベルごとにペナルティ係数 (Penalty Coefficient:  $C_p$ ) が定められている。

例えば RAID5 では、あるチャンクが書き換えられる際に、もともとのチャンクの読み出し、その箇所のパリティの読み出しを行い、更新後のパリティを計算してパリティの書き込みとチャンクの書き込みを行うため、合計 4 回の入出力処理を必要とする。そのため RAID5 のペナルティ係数  $C_{p5}$  は 4 であり、RAID6 のペナルティ係数  $C_{p6}$  は 6 である。 $R$  rpm ( $i$  IOPS) の HDD を用いて  $N_{np} + N_p$  ( $N_{np}$  はパリティが保存されないドライブ数、 $N_p$  はパリティが保存されるドライブ数) の RAID $n$  グループを作成した場合、 $N_{np} + N_p$  台のドライブからなる生 IOPS (Raw IOPS) は  $i \times (N_{np} + N_p)$  IOPS だが、ペナルティ係数が  $C_{pn}$  であるため、Write は  $i \times (N_{np} + N_p) \div C_{pn}$  IOPS となる。このとき Read と Write の比率を Read:Write= $r:w$  とした場合、この RAID $n$  グループが提供する IOPS,  $I_n$  は、

$$I_n = \frac{i(N_{np} + N_p)}{r + w} \left( r + \frac{w}{C_{pn}} \right) \quad (5)$$

となる。RAID グループは複数個をまとめて連結することができる。この場合、ストレージ機器が保有する全ての RAID グループの IOPS の総和が、ストレージ機器が提供することができる IOPS として推定できる。

仮想化ハイパーバイザの設計の CPU 単価と同様、ディスクドライブは回転数に応じて高額になり、同じ回転数のディスクドライブにおいては容量に応じて高額になる。容量が高く回転数が高いディスクドライブを用いれば少ない RU で要求される総ストレージ容量と総 IOPS を満足することができる。容量が低く回転数も低いディスクを用いても要求される総ストレージ容量と総 IOPS を満足することができるが、より多くの RU を専有するでハウジング費用が加算される。またストレージ製品のシャーシの制限があり、その数のディスクドライブを格納できない制約もある。ストレージ機器の設計においては、前述の IOPS の

\*3 Cisco UCS M-Series Modular Servers

<http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-m-series-modular-servers/index.html>

しかしこの製品で 1RU あたり 8 ソケットを実現すると 1 ソケットで 4 物理コア、32GB メモリであるため仮想化ハイパーバイザソフトウェアの価格が占める割合が高くなるため仮想化基盤を構築する用途には適切とは言い難い。10 物理コア、256GB メモリのノードで構成した場合は 1RU あたり 4 ソケットとなる。

\*4 EMC VNX2 Unified Best Practices for Performance  
<https://www.emc.com/collateral/software/white-papers/h10938-vnx-best-practices-wp.pdf>

\*5 wikipedia IOPS  
<https://en.wikipedia.org/wiki/IOPS>

算出式に従い、 $R$  rpm のディスクドライブを用いて要求される IOPS を満たすために必要な総ドライブ数  $N_d$  を算出し、要求される総ストレージ容量  $S$  を満たすために必要な 1 台あたりのドライブ容量を算出する。ストレージ製品ではドライブを格納するシャーシとは別にコントローラがある構成をとる場合もある。このコントローラが  $r_c$  RU のラックを専有するとする。2.5 インチのドライブを用いた時には 2RU で 24 台のドライブを格納できる製品があるが、1RU あたりに格納できるドライブ数を  $u_d$  とすると、 $y$  年間で要するストレージのハウジング費用  $S_h$  は、

$$S_h = \left( \frac{N_d}{u_d} + r_c \right) \times u_h \times y \quad (6)$$

で表すことができる。HDD でストレージ装置を構成する場合、要求される総 IOPS を満たせない場合もある。たとえば要求される総 IOPS を 10,000 IOPS とし、15,000 rpm (200 IOPS) の HDD を 4+1 の RAID5 で構成しようとする。R:W=1:1 と仮定した場合、この RAID5 グループが提供する IOPS は 625 なので 16 グループ、80 台の HDD が必要となり、シャーシは 4 基、8RU となる。これはまだ実現可能な範疇と言えるが、要求される総 IOPS が 10 倍になればストレージのみでラック 2 基が埋まる構成となり、ハウジング費用が膨大になる。この場合、フルフラッシュストレージの導入が現実的な選択肢として存在する (図 2)。

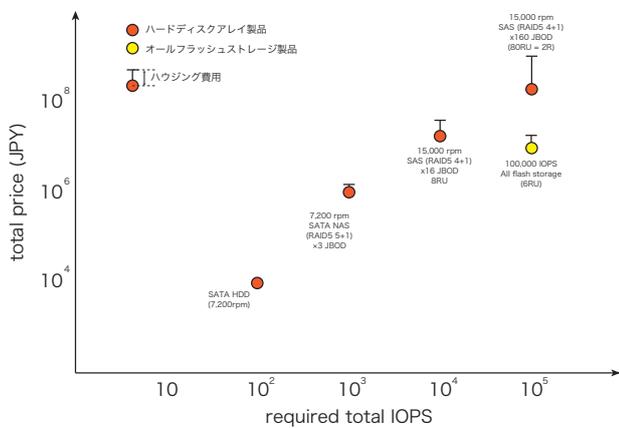


図 2 ストレージ機器の IOPS と価格の関係

Fig. 2 Scatter plot of IOPS versus its prices of storage products

### 2.3 ネットワーク機器と電源供給

ネットワーク機器と電源供給については機能に特化した製品である。ネットワーク機器については VM が要求する総トラフィック量、単位時間あたりの接続数、要求スループットをもとにファイアウォール、ロードバランサの仕様を決定する。仮想化ハイパーバイザとストレージ機器の間で発生するトラフィック要求量によりこれらの機器のインターコネクットの性能を決定する。ネットワーク機器を含めた全ての想定される製品の電力容量を算出し、

電力会社からの電力供給が途絶えた際に稼働させ続けることを想定する時間を定め、その時間、要求される電力容量を提供することができる無停電電源装置を調達する。この時間には仮想化ハイパーバイザ、ストレージ機器、およびネットワーク機器を停止させるのに必要な時間を含める。

## 3. 設計の実証実験

100VMs 程度が動作する大阪大学の仮想化基盤「大阪大学キャンパスクラウド」の増強設計に本手法を導入し、提案する設計手法の有効性を実証する。

### 3.1 大阪大学の仮想化基盤

大阪大学は学部学生約 1.5 万人、大学院生約 7.9 千人、教職員数約 6.4 千人 (非常勤職員等を含めると約 1 万人) からなる国立大学である\*6。大阪大学もまた ICT 投資を効率化するという目的のもと、大阪大学全体として業務フロー全体の最適化を行い、業務の効率化を目指すという目標が掲げられた。大阪大学程度の規模の総合大学では、部局ごとに独自の ICT 投資が行われ、事務業務フローも部局独自で構築されるケースがある。大学全体を俯瞰すると ICT 投資が分散し、業務改革も局所的な最適化に留まり、非効率な状態にあることもある。業務の全体最適化と ICT 投資の集約を実現する手段として、大阪大学は情報推進機構を設置し、仮想化技術を中心に据えたクラウド技術の活用に取り組んだ。本機構は、将来的には ICT リソースを外部にアウトソースする可能性も選択肢の一つとして考えながら、プライベートクラウド方式のプラットフォームシステムの構築を目指していた [5,6]。

大阪大学では上述の背景のもと、事務業務の効率化・改革の第一歩として事務系基幹システムを 2010 年に刷新した。この刷新時には、今後の大学の様々なサービスを集約し実行可能なシステムの構築を目指し、仮想化技術を採用した共通基盤プラットフォームシステム「大阪大学キャンパスクラウド」(以下、キャンパスクラウド)を設計、構築した。このキャンパスクラウドは事務系基幹システムの計算機資源だけでなく、財務会計システムからも計算機資源の提供を受けて 2010 年からサービスを開始した。このキャンパスクラウドは hp のブレードサーバを 12 台利用し、物理コア数は合計 96、主記憶容量は合計 432GB で動作するものである。ここにさらに 2010 年度に教員基礎データベースシステムと呼ばれる計算機資源が追加され、2つの業者が導入した3つのシステムによる3ブレードシャーシからなるキャンパスクラウドが出来上がった。

一方、これらの機器が設置されている大阪大学吹田キャンパスのサイバーメディアセンター本館では耐震・改修工事が 2014 年度に行われる事が決定し、スーパーコンピュー

\*6 平成 28 年 5 月 1 日現在の数  
<http://www.osaka-u.ac.jp/ja/guide/about/outline>

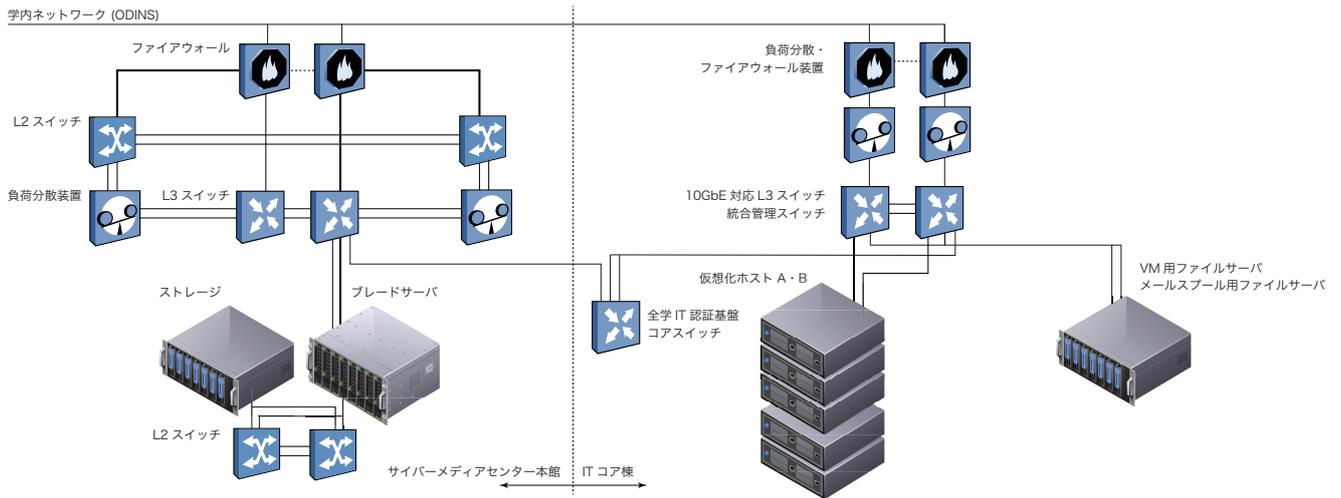


図 3 仮想化基盤の増強 2014 の模式図

Fig. 3 Diagram of virtualization infrastructure reinforcement in 2014

ター等を収納する地上 2 階建て、建物延べ面積 2,040m<sup>2</sup> の IT コア棟が新設されることとなった。IT コア棟の竣工は 2014 年 9 月末、その直後にサイバーメディアセンター本館の耐震・改修工事が着工することとなった。サイバーメディアセンターの改修にあたり、全ての計算機資源は IT コア棟に移設する必要があるが、移設作業中、キャンパスクラウドのサービスが全て停止することが懸念された。移転作業は週末の休日と祝日を含めた 3 日間で完了するものと見積もられたが、メールシステムを 3 日間ダウンさせることにはユーザからの反発が想像することが難しい。ソフトウェアのサポート上の問題点も指摘された。

2010 年のサービス構築時において仕様として要求されたハイパーバイザソフトウェアは VMware 社の VMware Infrastructure 3 であった。このソフトウェアはバージョンアップを繰り返し、現在では VMware ESXi™5.5 となっているが、キャンパスクラウドは VMware ESX 4.0 で稼働し続けていた。VMware 社ではハイパーバイザソフトウェアのメンテナンスアップデート、アップグレード、不具合とセキュリティの修正、および技術的な支援が提供される“General Support”の期限を製品ごとに設定しており、VMware ESX 4.0 の General Support は 2014 年 5 月 21 日で終了した。その後は 2016 年 5 月 21 日まで技術的な支援は提供されるが、その他のサポートは提供されなくなるため、ハイパーバイザソフトウェアに深刻な脆弱性が発見された場合は対処を行うことが極めて困難となる。これらの問題を解決するために、基幹系プラットフォームの機器を拡張することとなった [7]。

IT コア棟への移行によるサービス断時間の抑制、および新しいハイパーバイザソフトウェアへの全ての VM の移行を実現するため、サイバーメディアセンター本館と IT コア棟をまたぐ構成で仮想化基盤の増強設計を行った (図 3)。

2014 年 10 月から稼働を開始したこの仮想化基盤は 120 物理コア、640GB メモリ、VM 用ストレージ 20TB、メール用ストレージ 14TB の計算機資源からなり、物理コア対 vCPU 比は 4.0 として設計が行われた。2016 年 3 月末時点でこの仮想化基盤は 29 システムの 60VMs を稼働させていた。このうち最も高い可用性を要求されるキャンパスメールシステムは、同時期において 52 ドメイン、約 11,000 アカウントの利用があった。しかし 2016 年 4 月以降にはグループウェア、勤務管理、旅費、学務情報といった基幹システムがキャンパスクラウドでの動作させることを 2015 年度中に決定しており、CPU、メモリ、ストレージの逼迫が懸念された。そのため 2015 年度に、2016 年度から稼働させるための第 2 次増強設計が行われた。

### 3.2 状況の定量的計測

2015 年 8 月の時点でキャンパスクラウド上で動作する全ての VM が要求する vCPU 数は検証環境を含めて 340 であり、物理コア対 vCPU 比は 2.83 であった。この比率を 4.0 として設計されたキャンパスクラウドの状況を定量的に計測し、その設計指針値の妥当性を評価する。

図 4 は 2014 年度に導入された 2.7GHz、12 コアの CPU を 2 ソケット搭載した仮想化ハイパーバイザ 5 台の CPU 使用率の平均値と標準偏差を表している。2015 年 9 月末から 2016 年 7 月末までの値を示す。年間を通して平均 CPU 使用率は 10% から 15% を推移しており、標準偏差も最大で 30% である。図 5 は同じ仮想化ハイパーバイザ 5 台の CPU が提供する動作周波数の平均値と標準偏差を表している。計測期間は CPU 使用率の計測期間と同一である。動作周波数は全ての物理コアの積算値となっており、年間を通して平均動作周波数は 4000MHz・cores から 10000MHz・cores を推移している。全ての仮想化ハイパーバイザは 24

物理コアを搭載しているので、1物理コアあたりの平均値では166.7MHzから416.7MHzを推移している。最大値で12887.2MHz・cores(1物理コア平均535MHz)であり、2016年4月7日に発生している。この時点で動作周波数はこれは新年度にまつわるデータ投入とその処理に関連するバッチ処理が発生させたものと考えられる。これらの値は2.7GHzの動作周波数を提供するCPUからなる仮想化ハイパーバイザにおいて、全てのVMの処理要求は適切に処理されていることを示している。

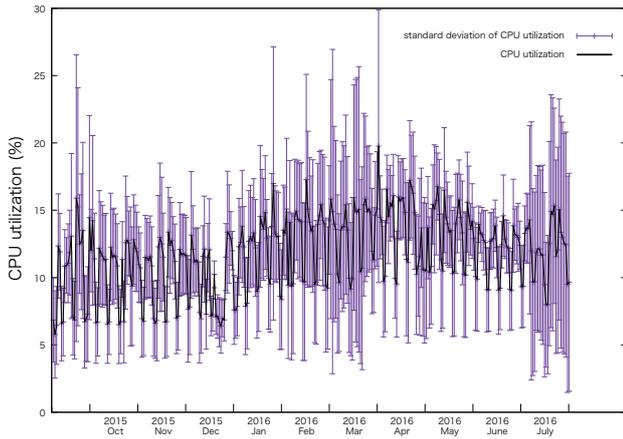


図4 仮想化ハイパーバイザのCPU使用率の平均値

Fig. 4 Averaged changes of CPU utilization of virtualization hypervisors

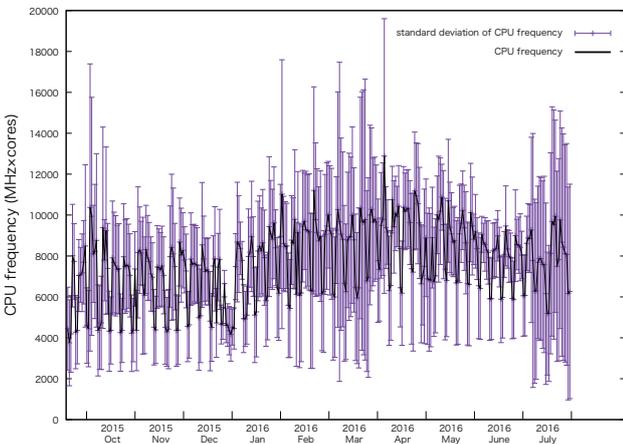


図5 仮想化ハイパーバイザのCPU動作周波数の平均値

Fig. 5 Averaged changes of CPU frequency of virtualization hypervisors

### 3.3 増強の設計

2015年8月において2016年度にキャンパスクラウド上での稼働を開始予定の情報システムが要求するvCPU数は検証環境を含めて合計で382であり、稼働中のVMが要求するvCPU数を加えると2016年度に722vCPUを要求

されることになる。これを2014年度に構築した120物理コアのキャンパスクラウドで賄う場合、物理コア対vCPU比は6.0となり、当初設計が目標としていた比率である4.0を大きく逸脱し、CPUオーバーサブスクリプションによるVMの性能低下が懸念された。

2017年度以降にキャンパスクラウドでの動作が予定している情報システムで100vCPU以上を要求することが想定されるものは図書館情報システムのみであり、2017年度以降、要求されるvCPU数がこのまま2014年度から2015年度の利用VM数の変化を線形補間した推定値に沿って増加しないことも想定された。そこで2018年度までの2年間で要求されるvCPU数を200vCPUと定め、この上限の要求が発生した場合においても物理コア対vCPU比4.0以下を保持するために必要な物理コア数を算出すると、 $922 \div 4.0 \div 230$ 物理コアとなる。これにより2016年度の増強での追加物理コア数は110と定められた。

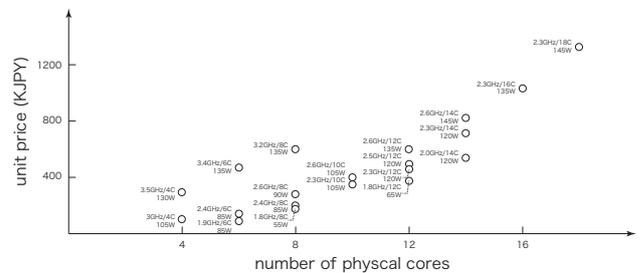


図6 CPUの物理コア数に対する価格の散布図

Fig. 6 Scatter plot of a number of physical cores versus unit price of CPUs

110物理コアを調達するため、その時点でのCPU提供価格(参考見積)を物理コア数別でプロットしたものが図6である。ハイパーバイザソフトウェアのライセンスは新規購入ではなく、2009年度に調達したプラットフォームに導入したVMware ESXi 4.0のライセンスからのアップグレードを利用し初期費用の低減を実現した。アップグレード先であるVMwareのvClud Suiteの提供価格とハウジング費用から2.1節で示した式(4)から増強費用を求めると、110物理コアを実現するための仮想化ハイパーバイザの価格比較は図7となる。1つの物理コア数の価格に幅がある場合は最高価格と最低価格の2本の棒グラフを記した。この図から14物理コアのCPUで構成した112物理コアの構成が最安、18物理コアのCPUで構成した108物理コアの構成が次点(価格差4%)であることが分かる。

稼働を予定するVMの中に、そのVM上でORACLEのデータベースを動作させる可能性のあるVMがあることが分かった。この場合、純粋なコスト比較のみでなくVM上で動作させるORACLEのライセンスに関する問題も制約条件として与えられる\*7。最終的にこの制約条件を満たす

\*7 <https://faq.oracle.co.jp/app/>

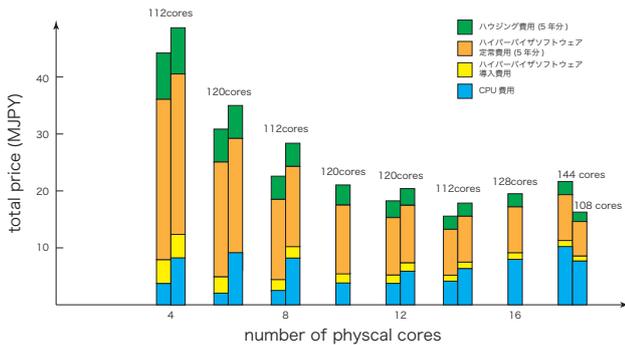


図 7 物理コア別のハイパーバイザの総費用

Fig. 7 Total cost of hypervisor versus a number of physical cores

ために、12 物理コア CPU を 4 基、18 物理コアを 4 基で 120 物理コアからなる構成に決定した。

ストレージの設計においては 2014 年度の設計で 10,000rpm の SAS HDD からなる RAID5 構成の領域を複数確保し、全体としてアクセスパターンを Read:Write=1:1 と仮定した場合、4500IOPS を提供可能なストレージを導入した。ストレージベンダから提供されたストレージ監視アプリケーションおよびハイパーバイザの管理ソフトウェアの双方で 2015 年 2 月から 8 月までの一日あたり、一週間あたり、一ヶ月あたりの IOPS の平均を求めたところ、ピーク時で 2600IOPS、平均値で 1200IOPS の I/O 性能が要求されていることを確認した。2016 年度にキャンパスクラウド上での稼働を開始予定の情報システムが要求する IOPS は、要求仕様から積算した結果、4800IOPS となった。既存のストレージが提供する I/O 性能では不足するため I/O がネックとなりアプリケーションの動作に影響を及ぼす可能性が懸念された。2017 年度以降に利用を開始する情報システムも想定に入れて現行のストレージと併せて 10,000IOPS を提供可能なストレージとすることとした。

I/O 性能だけでなくストレージ容量も逼迫していた。2014 年度の設計では VM 用のファイルサーバとして 20TB の領域を要求したが、2015 年 8 月時点で 14TB の領域が要求されていた。これに対して 2016 年度にキャンパスクラウド上での稼働を開始予定の情報システムが要求するストレージ容量は合計 13TB であった。VM に接続するディスクをシンプロビジョニングで確保しているため各ディスクの空き容量領域は利用可能であったが、今後の各情報システムの利用拡大により空き容量が逼迫した際に本来確保されていたストレージ容量を提供できないという事態に陥ることを回避したいという運用の要望もあった。2016 年度にキャンパスクラウド上での稼働を開始予定の情報システムが要求するストレージ容量の合計は 18TB である。

要求される IOPS を満足し、かつストレージ容量を満

たす構成を得るために、2.2 節で示した総ドライブ数  $N_d$  とその単価、および式 (6) を用いて算出されたハウジング費用から算出する。安価に構築できる構成は最安価が 10,000rpm、900GB の HDD を用いた構成 (5625IOPS, 43.2TB) となり、次いで 10,000rpm、600GB の HDD を用いた構成 (5625IOPS, 28.8TB, 価格比 0.01%)、10,000RPM、1.2TB の HDD を用いた構成 (5625IOPS, 57.6TB, 価格比 4.5%) であった。高い IOPS を要求する VM のみを配置するオールフラッシュストレージの導入も検討されたが、ストレージ全体で合計 10,000IOPS を満たすという最低限度の要件に対してオールフラッシュストレージが提供は過剰投資であると判断されたため、2014 年度に導入したストレージと同一スペックの 20TB のディスクアレイを導入し、VM 用ストレージ領域として増設することとした。

この他の増強要件としてはファイアウォール・負分散装置のスループットを向上するために 2014 年度に導入した製品のライセンス購入による性能向上を図った。2009 年度に導入したプラットフォーム基盤の機器が撤去されることに伴い、事務職員向けネットワークとの接続を実現するための L3 スイッチ、および保守ネットワークを引き継ぐための L2 スイッチを導入することになった。これらを全て合わせ、2016 年度の増強設計は図 8 に示す構成となった。赤い矩形で括った部分が 2016 年度の増強部分であり、保守ネットワーク用 L2 スイッチおよび事務情報ネットワーク用 L3 スイッチより左側の領域が 2016 年度で撤去される 2009 年度導入のプラットフォーム基盤機器である。

#### 4. まとめと展望

組織内仮想化基盤において定量的な指標を導入した設計手法を提案した。本提案手法に従った増強の設計と、大阪大学の仮想化基盤である「大阪大学キャンパスクラウド」の増強設計に適用し、調達価格の抑制に成功することができた。CPU のオーバーサブスクリプションによる VM 上で動作するアプリケーションのパフォーマンス劣化を抑制可能な範囲でサーバの集約度を上げることに成功している。2016 年度の拡張は 2016 年 7 月末に納品を終えて、8 月から増強された環境での動作が始まっている。2016 年度に大阪大学キャンパスクラウドでのサービス提供の開始を予定していたシステムは 10 月以降に本番稼働が始まるため現時点ではこれらの新システムの要求に応じることができているかどうかを定量的に判断することができないが、2016 年 12 月には稼働開始後 1~2 ヶ月における状況を報告することができる予定である。

大阪大学キャンパスクラウドという名前のシステムではあるが、大阪大学に所属するユーザが自由にこの仮想化基盤上で VM を作成することはできない。クラウドコンピューティングというモデルにおいて重要な点は“pay as you go”，すなわち従量課金にある。大阪大学キャンパス

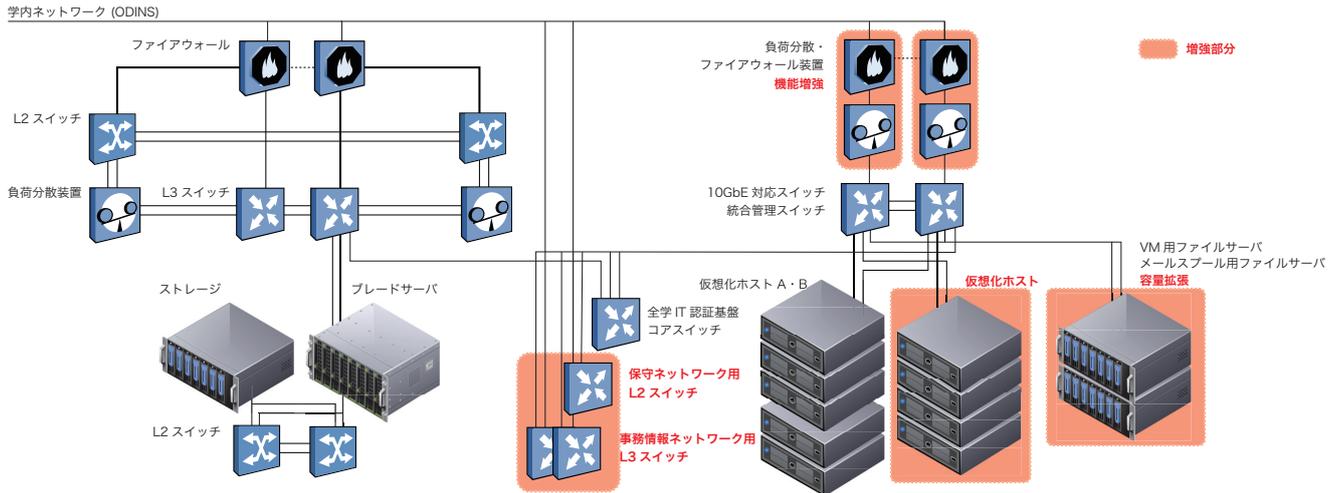


図 8 仮想化基盤の増強 2016 の模式図

Fig. 8 Diagram of virtualization infrastructure reinforcement in 2016

クラウドでは vCPU 数、メモリ容量、ストレージ容量に応じて月単位で料金請求額が定められており、年末までに年度分の使用料を支払うモデルとなっている。VM が要求する資源を途中で柔軟に変更するインターフェイスは提供されておらず、また月単位の課金という粒度では、サービスの利用変化に柔軟に対応することができるというクラウドコンピューティングの利点を発揮することができていない。組織内の環境だからこそ理想的な環境を目指すことが望ましいのかもしれない。

また、パブリッククラウドプロバイダでも実現できていないことではあるが、vCPU 数やストレージ容量という粒度の大きな単位での課金ではなく、vCPU に割り当てられた動作周波数、ストレージ容量だけでなく IOPS に対する粒度の細かい課金を設定することも、組織内仮想化基盤であればこそ実現可能なことと言える。今回の増強設計では詳細まで調査をしなかったファイアウォール・ロードバランサー機能ではあるが、これも VM 毎の利用度を計測し、機能ごとの課金、例えば SSL アクセラレーションを秒間にどれだけ使ったか、という指標で課金を行うことができれば、仮想化基盤への投資に対して妥当性のある課金額の設定を行うことができ、理想的なクラウドコンピューティング環境を構築することができると期待している。

プライベートクラウドにおける問題点として、組織内において使用料金の節約をしようとする動機付けがないことが挙げられる。例えば組織内においては仮想化基盤上で動作する情報システムの担当部署が、その情報システムを稼働させるための使用料金の単年度の概算金額を要求し、翌年度にその金額が割り当てられるという仕組みである場合、情報システムの構成や設計に対して適応的な仕組みを導入するなどして利用料金を低く抑えることができたとし

ても、その担当部署は余剰金を自由に利用できるわけではない。より有効に計算機資源を使おうとさせる誘引要素が存在しない。これは単に課金制度の問題だけでなく、組織内における予算執行プロセスそのものまで手を入れなければ改善できない問題点と言える。

大阪大学では 2016 年度のキャンパスクラウドの増強設計を計算機資源の規模拡大を行う最後の増強と考えており、今後はゆるやかに計算機資源の規模を縮小させ、稼働中の VM はパブリッククラウドへと移行させていく撤退戦に入ることを検討中である。撤退戦略をどのように定量化して策定していくかがこれからの課題となる。

#### 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] 松浦知史, 森健人, 金勇, 友石正彦: 拡張性を考慮した小規模仮想化基盤の構築, 研究報告インターネットと運用技術 (IOT), Vol. 2016-IOT-32, pp.1-8 (2016)
- [3] 鍛冶秀紀, 安東孝二, 小野成志: 武蔵大学における仮想化基盤を活用したキャンパスネットワークの構築と運用, 研究報告インターネットと運用技術 (IOT), Vol. 2013-IOT-22, No. 6, pp. 1-5 (2013)
- [4] 前田香織, 末松伸朗, 北村俊明: 広島市立大学における情報ネットワークシステムのクラウド環境移行, 研究報告インターネットと運用技術 (IOT), Vol. 2015-IOT-28, No. 19, pp. 1-6 (2015)
- [5] 市川昊平, 江原康生, 長岡亨, 森原一郎: 大阪大学のキャンパスクラウドへの取り組み, 大学 ICT 推進協議会 2011 年度年次大会論文集, pp. 312-325 (2011).
- [6] 宮永勢次, 市川昊平, 小林兼: 大阪大学のキャンパスクラウドシステムについて, 全国共同利用情報基盤センター研究開発論文集, No. 34, pp. 77-82 (2012).
- [7] 柏崎礼生, 宮永勢次, 森原一郎: 大阪大学における仮想化基盤の増強とクラウド戦略, インターネットと運用技術シンポジウム 2014 論文集, Vol.2014, pp. 93-100 (2014).