

大阪大学における仮想化基盤の増強とクラウド戦略

柏崎 礼生^{1,a)} 宮永 勢次^{1,b)} 森原 一郎^{1,c)}

概要: 大阪大学では 2010 年にサーバ集約を目的として、VMware 社の VMware ESX 4.0 をハイパーバイザとして利用する仮想化基盤を構築した。この基盤上で 2014 年 3 月末時点で 32 システムによる 105 VMs が稼働し、227 仮想 CPU を利用している。また本基盤の上で稼働するメールシステムは学内 38 部局、8,450 アカウントを提供する。この基盤は当初 2014 年度末にリプレイスされる予定であったが、2014 年 9 月に竣工したの学内データセンタ「IT コア棟」への移設に伴い、半年繰り上げて「拡張」が行われる事となった。本稿ではこの仮想化基盤の増強の設計と構築、および大阪大学における今後のクラウドコンピューティング戦略について紹介する。

Extension of a virtualization infrastructure and ICT strategies of cloud computing environment in Osaka University

HIROKI KASHIWAZAKI^{1,a)} SEIJI MIYANAGA^{1,b)} ICHIROU MORIHARA^{1,c)}

Abstract: A virtualized infrastructure on Osaka University started its service to aggregate server machines inside the university in March 2010. A hypervisor software of the infrastructure have been VMware ESX 4.0 even now and 105 VMs of 32 systems that using 227 virtual CPU is working on the infrastructure. An university mail systems is one of the systems and this mail system provide 8,450 mail accounts with 38 departments. Although the infrastructure was supposed to be replaced until March 2015, replacement was moved up in September 2014 because of datacenter of the university a.k.a. IT core building was completed of construction in September 2014 and the current infrastructure have to be moved to the building. In this paper, authors show the design and construction of new virtualized infrastructure. They also show ICT strategies of cloud computing environment in Osaka University.

Keywords: virtualization infrastructure, capacity planning, cloud computing, design

1. はじめに

「人間の仕事をやってくれる、新しくかつ最も有能な機械的奴隷の集団を人類が持つことになるのである。(中略) 機械のおかげで不快な卑しい仕事をやる必要がなくなるのは、人類にとってひじょうな福祉かもしれないが、あるいはそうでないかもしれない。私にはわからないことである。」とノーバート・ウィーナーは語った [1,2]。人間は単純な計算を機械 (計算機) に委託し、やがて計算の組み

合わせからなる複雑な処理をも委託するに至ったその処理はやがて人間の植生に適合するよう地理的に拡大され、分散システムと融合しながら情報通信技術 (Information and Communication Technology: ICT) を開拓した。人間がより多くの時間を思考に費やすため、より多くの処理を計算機に委託した結果、ICT 投資は肥大の一途を辿り、組織は ICT 投資の効率化に着目せざるを得なくなった。この潮流と、計算機科学において古くから研究開発されていた仮想化技術の親和性は高く、2000 年代からエミュレーションによるレガシー OS 動作環境の仮想化や、コスト削減のためのサーバマシンの集約を目的として用いられるようになった。米 VMware 社^{*1} は代表的な仮想化ソフトウェア

¹ 大阪大学

Osaka University

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

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

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

^{*1} <http://www.vmware.com>

ア企業であり、VMware 社が VMware Infrastructure 3 をリリースした 2006 年からサーバ集約のための仮想化技術の活用が様々な組織で進められるようになった。奇しくも当時 Google の CEO であった Eric Emerson Schmidt 氏が Search Engine Strategies Conference において Danny Sullivan 氏との対談で「クラウドコンピューティング」という言葉を紹介したのも 2006 年のことである *2。

日本国内の大学におけるサーバ資源集約のための仮想化技術の利用例としては、2007 年から「クラウド化」を推進している静岡大学の事例が有名である [3]。また北陸先端科学技術大学院大学では仮想デスクトップサービスを提供するためにプライベートクラウドを構築しており [4,5]、佐賀大学は専用線で接続された外注先にプライベートクラウドを構築しメールサービスの提供を行うなど [6]、仮想化技術を用いた様々なアプローチで ICT 投資の効率化を図っている。一方、情報技術に関する研究とコンサルティングをもっぱら行う米ガートナー社の research director である Michael Warrilow 氏は、2014 年はプライベートクラウドからハイブリッドクラウドやパブリッククラウドへと移行するようになるだろうという趣旨の発言をしている *3。2000 年代後半に仮想化基盤を構築した組織の中には、数年間の運用とリソースの計測を通して、その基盤構築への投資が期待した効果を提供しないことに気付いた組織もある。2014 年 8 月に米ガートナー社が公表した “Hype Cycle for Emerging Technologies, 2014” *4 によると、クラウドコンピューティングは過剰な期待のピーク (Peak of inflated expectations) を越えて幻滅・目覚めの谷 (Trough of disillusionment) に入っている (図 1)。一方、米ガートナー社は 2013 年 10 月のプレスリリースで「約半数の大企業は 2017 年末までにハイブリッドクラウドを展開するようになる」と述べている *5。また同社の research director である Michael Warrilow 氏は、2014 年にはクラウドコンピューティング環境で処理される仕事総量のうち 70～80% がパブリッククラウドサービスで処理されるようになる」と指摘している *6。2014 年からの「プライベートクラウド」あるいは仮想化基盤の構築には、これまでの 8 年の過程を考慮に入れ、かつハイブリッドクラウドの利用を経たパブリッククラウドへの移行を視野に入れた設計が求められる。

2. 大阪大学の仮想化基盤

大阪大学は学部学生約 1 万 5 千人、大学院生約 7 千 5 百

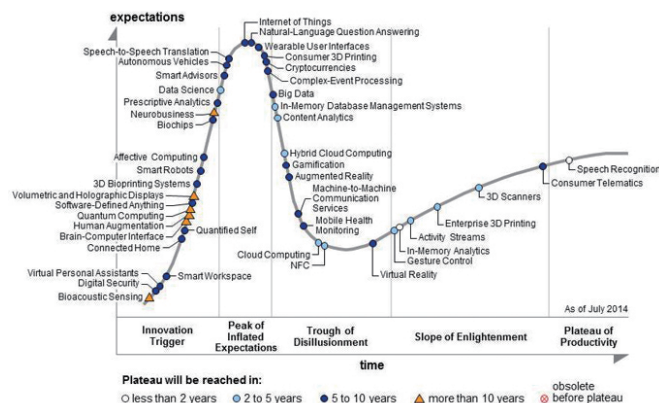


図 1 米ガートナー社による先進テクノロジーのハイブサイクル 2014 年度版

Fig. 1 Hype Cycle for Emerging Technologies, 2014

人、教職員数約 6 千人 (非常勤職員等を含めると約 9 千 3 百人) からなる国立大学である。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 1 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 [7,8]。

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。事務系基幹システムの刷新の際には、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。この基盤上で 2014 年 3 月末時点で 32 システムによる 105VMs が稼働し、227 仮想 CPU を利用している。また本基盤の上で稼働するメールシステムは学内 38 部局、

*2 <http://www.google.com/press/podium/ses2006.html>

*3 http://www.cio.com.au/article/534196/where_cloud_computing_heading_2014_/

*4 <http://www.gartner.com/newsroom/id/2575515>

*5 <http://www.gartner.com/newsroom/id/2599315>

*6 <http://news.idg.no/cw/art.cfm?id=D50FA809-A35F-4B39-866F2B7569F7F2B0>

8,450 アカウントを提供している。

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タ等を収納する地上2階建て、建物延べ面積2,040m²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月末を予定しており、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ている。サイバーメディアセンターの改修にあたって、全ての計算機資源はITコア棟に移設される必要があるが、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日間で完了することが可能と見積もられたが、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。それ以外にも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である。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で稼働し続けている。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日で終了する。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。

2.1 仮想化基盤の拡張設計

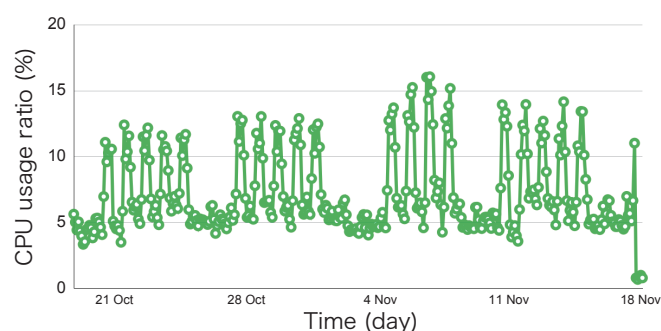


図2 仮想化ホストのCPU利用率の平均値

Fig. 2 Average of CPU usage ratio of 12 virtualization hosts

拡張したキャンパスクラウドの性能的な要求要件を策

定するため、まず現キャンパスクラウドの計測を行った。キャンパスクラウドを構成する12台の仮想化ホストの中のCPU利用率を4週間計測し、その平均値をグラフ化した(図2)。仮想化ホスト1台は2つのCPUから構成されており、Intel Xeon 5500番台、2.53GHzのCPUが仕様で要求されている。この数値は2時間ごとの平均値から算出されているため、より短い時間でのピークは図よりも大きくなるが、参考値として利用するものとする。より短い時間でのCPU利用率を計測し、偏差を算出することが望ましい。この図では平日の9時から17時にピークがあり、ピーク時の平均CPU利用率は20%未満、夜間および休日の平均CPU利用率は10%未満であることが分かる[9]。

次に今後キャンパスクラウドへの移転を検討する物理マシンの計測を行った。業務系システムの中でも負荷が高いとされる財務会計システムのデータベースサーバを対象とし、CPU負荷だけでなくストレージのIOPSも評価対象とした。ここではネットワンシステムズ株式会社の協力のもと、VMware社が提供するCapacity Plannerを利用して計測を行った。財務会計システムのデータベースサーバは2つのCPUから構成されており、動作周波数は2.93GHz、メモリ容量は24GBである。2週間の計測を行い、毎日0時から23時までの平均値を計算した(図3)。図の左はCPUリソース利用状況であり、右はディスクリソース利用状況である。夜間バッチが実行される2時から5時が最も高いプロセス待ち行列長(茶色)を示している他、業務開始後の午前中もまた比較的高い値を示している。一方CPU利用率(青色)は業務時間内である日中は比較的高い値を示すが、24時間の平均CPU使用率は8.3%、日中に限定しても14.9%、ピークは20.5%である。ディスクリソースも、夜間バッチが実行される2時から5時の値が高くなっているが、日中は比較的安定しており、IOPSの平均値は216、ピーク値は490であった。またReadとWriteの割合は平均時で77:23、書き込みのピーク時においても65:35であった。

この結果を前提とし、今後3年間のシステム数の増大を推測した上で、拡張したキャンパスクラウドは物理コア数は現行と同程度の120コアとし、この上で動作させるVM数を約4倍の400にすることとした。また潤沢な主記憶容量の要求に応じることができるよう、拡張当初は合計640GBとし、400VMsが平均200IOPSを要求するものとして80kIOPSの要求に応じることのできるストレージを具備するものとした。その上で費用の増大を抑制することを考える必要がある。

2.2 ライセンス費の抑制

現構成において導入されたハイパーバイザソフトウェア群であるVMware vSphere 4.0のエディションはEnterprise Plusである。このソフトウェア群は、動作する物理CPUの数に比例したライセンス費用体系となっている。

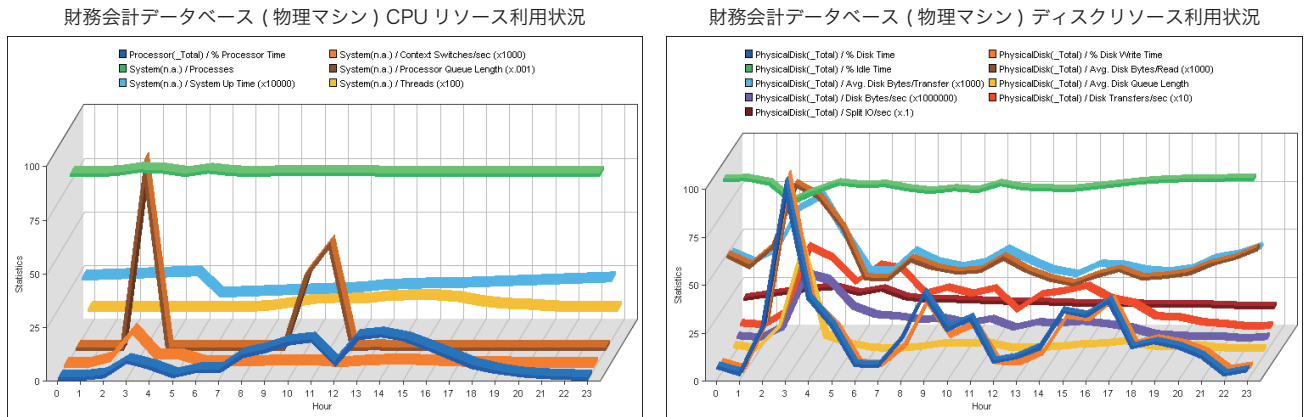


図 3 物理ホストの CPU・ディスクリソース利用状況
Fig. 3 Status of CPU and storage resource usages on a physical database server

Enterprise Plus は vSphere における最高級のエディションであるが、その理由は Storage vMotion や Distributed Resource Scheduler を利用するためではなく、Cisco Nexus 1000V を利用することが目的であった。

Cisco Nexus 1000V は Cisco 社の分散仮想スイッチソフトウェアである。複数の仮想化ホストをクラスタ構成する場合はネットワークの管理・設定作業が非常に煩雑になるためモダンなハイパーバイザソフトウェアにおいてはその内部で仮想スイッチを作成し、VM をその仮想スイッチに接続する機能を標準で提供する。各仮想化ホストに分散した仮想スイッチの設定を一元管理するのが Nexus 1000V である。キャンパスクラウドにおいては主に VM のライブマイグレーション後の VLAN の設定を主な目的として Cisco Nexus 1000V が導入されたが、全ての VM が 1 つの VLAN の配下にあるキャンパスクラウドにおいて利活用される場面は極めて限定的であった。そして Cisco Nexus 1000V もまた物理 CPU の数に比例したライセンス費用体系である。キャンパスクラウドは 24 物理 CPU からなる計算機資源により構築されているため、これらのライセンス費用が高額であることもまた問題点として挙げられていた。

このため物理コア数の大きな CPU を調達するメリットを算出する必要がある。現環境では 96 コアのために 24 個の CPU を利用している。2013 年に発売された Ivy Bridge の Xeon では 1CPU あたり 12 コアの製品がリリースされている。増強設計においては 120 コアを仕様として定めたが、12 コアの CPU を用いると 120 コアは 10 個の CPU で実現することが可能である。比較的安価な Xeon は 8 コア、6 コアのものがリリースされているが、こちらを使うとそれぞれ 15 個、20 個の CPU が必要である。ハイパーバイザの初期費用および保守費用は CPU 単位で課金され、また CPU が消費する電力の料金も考慮に入れる必要がある。12 コア CPU は Xeon E5-2697v2、8 コア CPU は Xeon E5-2640、6 コア CPU は Xeon E5-2630 とし、ハイパーバ

イザの初期費用を 1CPU あたり 32 万円、ライセンス費用を年額 1CPU あたり 10.8 万円、電力使用量を 1KWh あたり 20 円とする。12 コア CPU の消費電力を 130W、8 コア、6 コア CPU の消費電力を 95W として、5 年間運用することを前提として比較すると、12 コア CPU で構成した場合は 1633 万円、6 コア CPU で構成した場合は 2206 万円と算出され、570 万円以上の差がある (図 5)。この比較から増強設計においては 12 コア CPU である Xeon E5-2697v2 による構成を採用した。

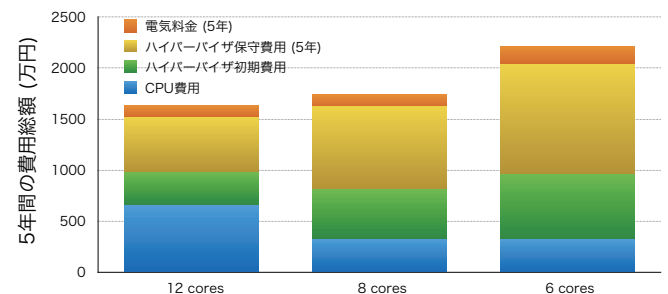


図 5 120 コアを実現するための 5 年間の総額比較
Fig. 5 A compare of total cost of 120 cores CPUs for 5 years

VMware vSphere 5.5 update1 の Enterprise Plus では VM の I/O 特製を考慮した格納先の最適配置を自動化できる Storage DRS, VM の IOPS を動的に制限することでストレージアクセスの優先度の決定を自動化できる Storage I/O Control が提供されている他、Nexus 1000V が提供していた機能と同等の分散スイッチを実現できる Distributed Switch が具備されている。そこで今回の調達では vSphere のエディションを変更しないかわり、Nexus 1000V の導入を削除することとした。

3. 移行プラン

現行のキャンパスクラウドにおける問題点の 1 つとして、構成する機器が使用するリソース情報のほとんどを長期間

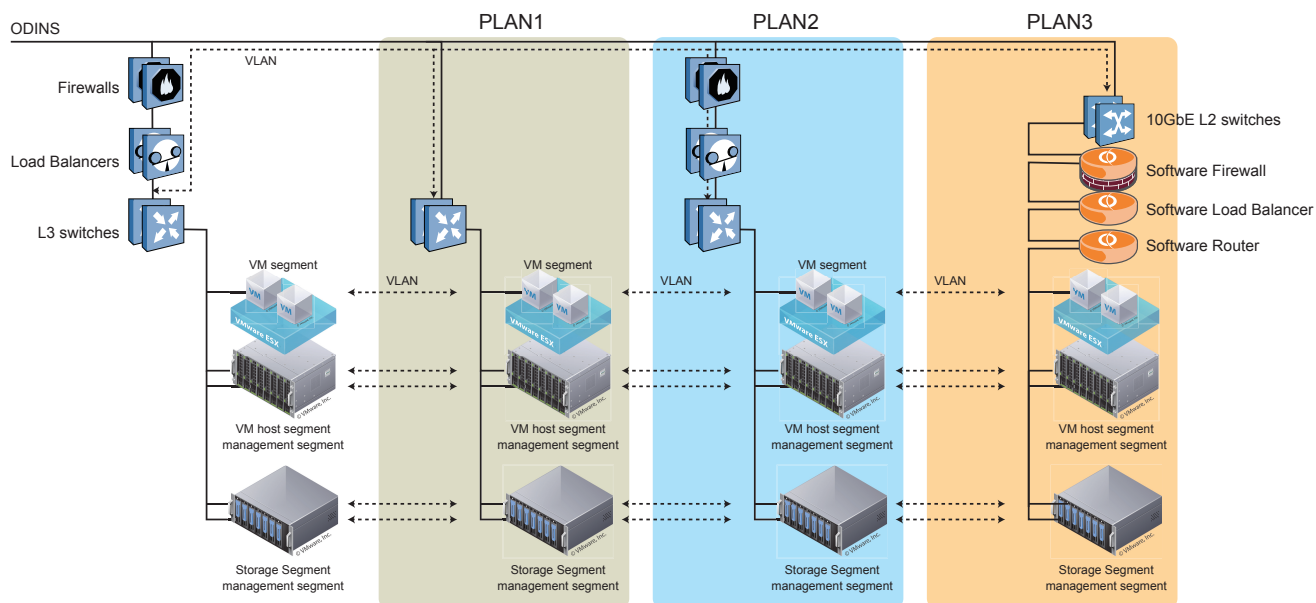


図 4 移行計画の議論の過程模式図
Fig. 4 A diagram of discussion process of the migration

に渡って収集、蓄積する環境になっていないことであった。VMware vSphere および vCenter の利用により仮想ホストのリソース利用率は計測できたが、その他のネットワーク機器、ファイアウォール、ロードバランサについてはログも極めて限定的にしか保存されていなかった。そのため拡張したキャンパスクラウドの要件として、全ての機器の情報を収集可能であることを要求している。

ファイアウォールやロードバランサにどの程度の負荷がかかっているかを把握することが出来ないため、根拠となるのは現行機種と同程度以上の機能を有することとしか要求することができない。移行に関する試行実験としては、

- (1) 現行のファイアウォール、ロードバランサを継続して利用する方法
- (2) 新たにファイアウォール、ロードバランサを購入し、現環境とは VLAN で仮想化ホスト、ストレージが通信可能な状態とする方法
- (3) ファイアウォール、ロードバランサはソフトウェアで実現する方法

が提案された (図 4)。最終的にプラン 2. が採択され、その後の移行のステージについても同様に図を用いて設計部門内で共有され議論が進められた。言葉による議事録ではなく、このような図を用いて共有をすることで、設計や仕様策定に携わる担当者が異動したり、あるいは新たに加わった時でも迅速に理解を共有することができ、また異動にまつわるドキュメントの執筆で作業が遅延することも未然に防ぐことが可能となる。

また、当初はファイアウォールとロードバランサーが個別の機器として調達することを前提としていたが、昨今のファイアウォール機器には十分なロードバランサー機能を

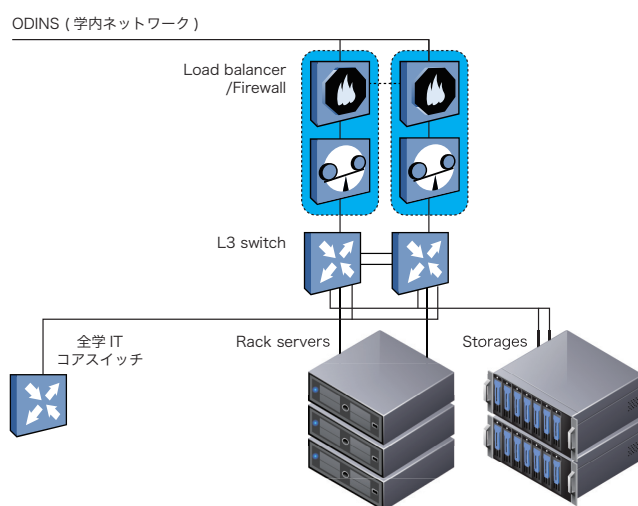


図 6 キャンパスクラウド拡張版の模式図
Fig. 6 A diagram of enforcement of the campus cloud

具備するものが少なくないため、最終的に図 6 に示すシンプルな構成で構築が行われることとなった。

4. VMware Virtual SAN (VSAN)

VMware 社^{*7} が 2014 年 3 月に General Availability をアナウンスしたソフトウェアである Virtual SAN (VSAN) は、VMware 社曰く Software-defined storage であるとされている^{*8}。VMware 社 CTO の Richard McDougall は VSAN について計算機資源とストレージを converge したモデルであると述べている^{*9} (図 7)。

ストレージベンダーの業界団体である The Storage Net-

^{*7} <http://www.vmware.com>

^{*8} <http://www.vmware.com/products/virtual-san/features>.

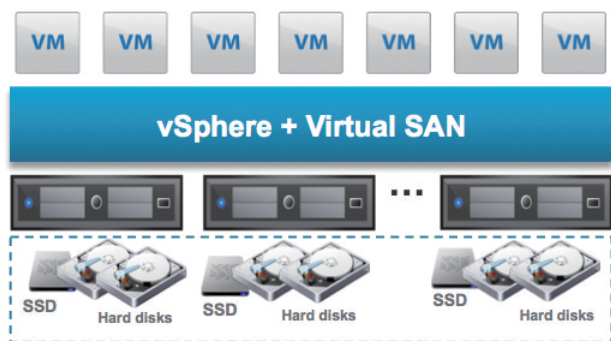


図 7 VSAN の模式図 (Office of the CTO の記事より)

Fig. 7 A diagram of VSAN

working Industry Association (SNIA)^{*10} は Software Defined Storage に関する White Paper をリリースしている^{*11}。この文書では Software defined storage が備える 10 の属性が述べられている。VMware 社が VSAN を Software defined storage であると位置づける理由については語られていないが、VSAN が提供するソフトウェアにより、一般的な x86 サーバ上でスケールアウト可能なストレージシステムを構築することができることから Software defined storage と位置づけられていることが推測できる。

VSAN では仮想化ホストによりストレージプールが構築されるため、仮想化ホストのハイパーバイザソフトウェアのメンテナンスを考慮にいれる必要がある。VMware ESXi は再起動を必要とするアップデートの数は (例えば Microsoft Windows Server と比較すると) 少ないと主張されるが、EMC 社や NetApp 社などのストレージの製品において再起動を必要とするアップデートがどの程度の頻度で発生するかと比較することが合理的な比較であると考えられる。こういった性能上・運用上の問題点を解決するために、現在増強を進めている仮想化基盤を用いて VSAN の検証を行うこととした。

4.1 検証環境の論理構成

図 8 に VSAN 検証環境のネットワーク構成を示す。仮想化ホストである UCS C240 は統合管理スイッチ、Fabric Interconnect である UCS 6248UP を使わず、10GbE 対応スイッチである Nexus 5548UP と Twinax 接続している。GbE を使い Cisco Integrated Management Controller への接続用ネットワークと、ESXi サービスへのリモート接続用ネットワーク、および VM 用セグメント用のネットワークを提供している。検証環境の機材はセキュリティエリア

に設置されているため、これら GbE 接続のネットワークは学内ネットワーク ODINS を経由してセキュリティエリア外の操作端末が設置されている区画まで VLAN により接続性を確保している。

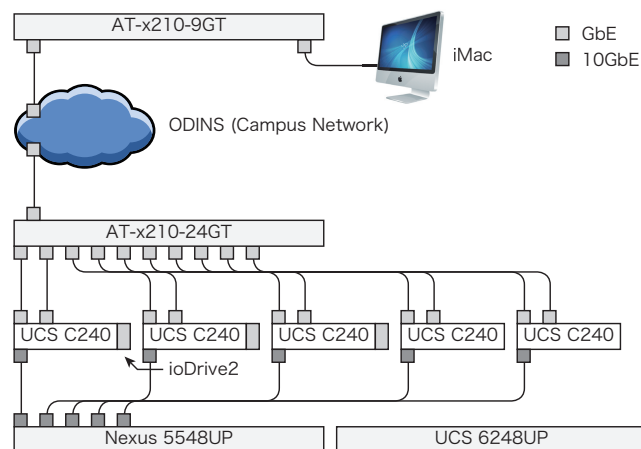


図 8 VSAN 検証環境のネットワーク構成

Fig. 8 A network diagram of verification environment for VMware VSAN

仮想化ホストである UCS C240 には ESXi 用の 4GB の USB フラッシュメモリが搭載されており、ハイパーバイザはこのメモリ上にインストールされる。VSAN を構築するためには VMware vCenter サーバが必要であるため、全ての仮想化ホストが備える 300GB のローカルストレージ上に VM を作成し、Windows Server をインストールして VMware vCenter サーバの環境を構築する。ESXi の仮想 NIC がこの VM と接続され、仮想スイッチにより UCS C240 の物理 NIC と接続されて各仮想化ホストを管理できる状態としている。VSAN クラスタを作成することで各仮想化ホストが VSAN をストレージプールとして利用することができる。このストレージプール上に VM を作成し、各種ベンチマークを実行することができる。

4.2 実運用への投入

実運用では UCS C240 を UCS 6248UP に接続して統合管理を行う構成としているため追加機器の購入なく検証環境から実運用環境への移行が可能となった。実運用で VSAN を利用するためには VSAN 専用の 10GbE ネットワークを用意することが推奨されているため、10GbE NIC の追加が必要となる。また検証環境で用意した RAID カードは 2014 年 9 月現在では VSAN 用に利用できる確認が取れていないため、実運用のためには RAID カードの追加が必要であることが判明した。機器を追加する際には PCI スロットの残り数の都合上、現状用いている RAID カードを取り外して、VMware から VSAN での利用が保証された RAID カードを追加し、ハイパーバイザを再インス

html

*9 <http://cto.vmware.com/the-dawn-of-vsan/>

*10 <http://snia.org>

*11 <http://snia.org/sites/default/files/SNIA\%20Software\%20Defined\%20Storage\%20White\%20Paper-\%20v1.0k-DRAFT.pdf>

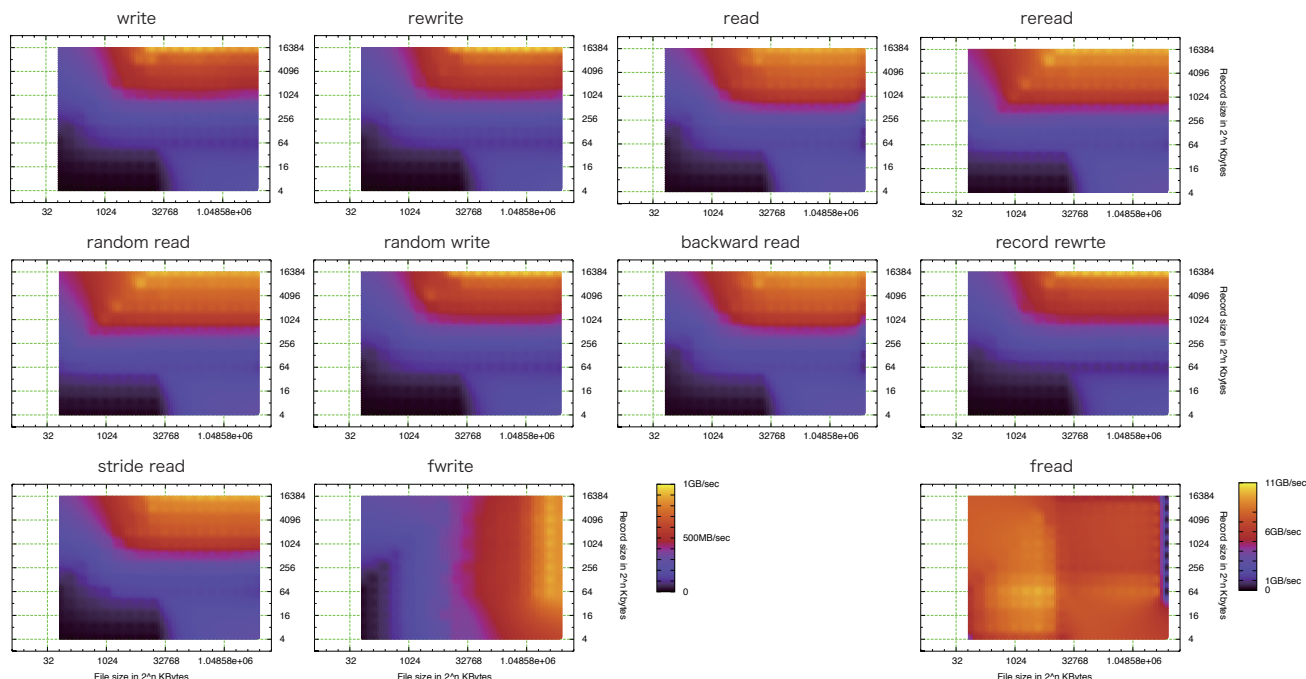


図 9 vCloud Air のストレージ性能評価
Fig. 9 Evaluations of vCloud Air's storage

ツールする必要があるため、仮想化ホストを一基ずつメンテナンスモードにする作業が発生する。このコストを削減するためには増強構築中に機器追加をする必要があった。

2014 年 7 月に学内組織から共同研究が持ちかけられ、この機器追加の資金の目処が立った。複数財源からなる資源の集合体として一つのシステムを構築することであっても責任境界を明確にすることで運用は可能であると考えられたが、その増強を行う事の明確な根拠に乏しいことと、責任境界の調停に要するコストが未知数である（あるいは経験的にはコストが高い）ことから、この共同研究による増強は見送られる事となった。仮想マシンモニタを用いてストレージへの負荷を定期的に計測し、共用ストレージの性能の限界が示されれば、今後 vSAN の実運用環境への投入も有り得る。それまでの間、PCI スロットに接続されたフラッシュストレージ機器は読み込みキャッシュ用の機器として利用することも可能である。

5. パブリッククラウド戦略

本増強の設計においては Peter F. Drucker の “The Deadly Sins in Public Administration” を参考に行っている [11]。この文献で Drucker は以下の 6 つの罪のうち 2 つ犯せばプロジェクトは失敗すると述べている。

- (1) To have a lofty objective.
- (2) To try to do several things at once.
- (3) To believe that “fat is beautiful”.
- (4) Don't experiment, be dogmatic.
- (5) Make sure that you cannot learn from experience.

(6) The inability to abandon.

そこで我々はこの仮想化基盤の増強をプライベートクラウドを目指すものではなく、パブリッククラウドへと展開するための踏み台と位置づけ、またこの位置づけすらも変更不可とならぬよう 2.5 年ごとに見直しを行う事とした。この基盤上で動作する VM は将来的にはパブリッククラウド上の IaaS へと移行することとなる。

組織が保有するプライベートクラウド、あるいは仮想化基盤の上で動作する仮想マシン (Virtual Machine: VM) をパブリッククラウドへと移行する過渡期においてハイブリッドクラウド状態が現れるが、ハイブリッドクラウドの利用においてはハイパーバイザーの差異が問題となる。ハイパーバイザーは VM を制御するプログラムだが、ある組織が保有するプライベートクラウドや仮想化基盤で用いられるハイパーバイザーと、利用予定のパブリッククラウドで用いられるハイパーバイザーは異なる場合がある。VM のイメージフォーマット形式はハイパーバイザーにより異なるため異なるハイパーバイザー間で VM を移行するためには、仕様の公開されたフォーマットである Open Virtual Machine Format (OVF) を経てフォーマットの変換を行う Virtual to Virtual (V2V) 処理が必要となる。V2V 処理の間、変換元となる VM のイメージに変更が加えられてはいけなため、VM は停止状態となる。そのため変換中、その VM が提供するサービスは停止することとなり、V2V 処理には相応のコストを必要とする。

VMware 社は 2013 年にパブリッククラウドサービスであ

る vCloud Hybrid Service (vCHS) の提供を開始した^{*12}。また 2014 年にはこのサービスの日本でのサービス提供が開始された。ハイブリッドクラウドにおいては組織内のシステムとパブリッククラウドとのデータの送受信が頻繁に、かつ大量に行われることが予想されるため、サービスが提供されるデータセンターとの遅延が低い方が有利であると考えられる。大阪大学における仮想化基盤は今後さらに増強するのではなく、パブリッククラウドへの移行を積極的に推進することを検討しており、vCloud Air は V2V のコストにおいて優位点があるため有力な移行先の候補となっている。そこでベータテスト中の vCloud Air の日本リージョンを利用させて頂き、vCloud Air の計算機資源のパフォーマンス計測を行った。

仮想化環境においてその規模が大きくなるにつれボトルネックとなるのはストレージであることが指摘されている [10]。そこで vCloud Air のストレージの評価を行うこととした。ストレージのベンチマークソフトとして iozone^{*13} を用い、vCloud Air 上の VM に iozone をインストールしてベンチマークを行った。iozone は指定したファイルサイズとブロックサイズにおけるスループットを評価することができ、13 のアクセスパターン (read, write, reread, rewrite, fread, fwrite, freread, frewrite, randread, randwrite, bkwrdread, strideread, recrewrite) が用意されている。本評価において、ブロックサイズは 4KB からサイズを倍にして最大 16MB までとし、ブロックサイズはファイルサイズを超えない事とした。ファイルサイズは 64KB からサイズを倍にして 8GB までとした。スループットは close コールを含めた時間を計測した。また Direct IO を利用し、open システムコールがカーネル空間のページキャッシュを利用しないよう指定する。試行は 100 回行い、全てのアクセスパターンに対する全てのファイルサイズとブロックサイズの組み合わせに対し欠損のない 98 の結果を対象として平均スループットを算出した (図 9)。

vCloud Air がベータテスト中ということもあってか、ストレージのスループットは全てのパターンで最大 1GB/sec のスループットを計測した。ストレージがフラッシュメモリを用いたキャッシュを用いているか、あるいはフラッシュメモリのみからなるストレージを用いている可能性がある。fread, fwrite の結果においては 11GB/sec を記録しており、iozone がカーネルキャッシュの影響を除外できているとするならば、VM とストレージ間のネットワーク帯域が極めて広いことが類推できる。このストレージを何台の仮想化ホストで共有しているかが問題となるため、あくまでベータテストの結果に過ぎないが、V2V のコストが低

い点も合わせて、パブリッククラウドへの移行における候補として十分な性能を保持することが分かった。

6. まとめと今後の課題

大阪大学における仮想化基盤の増強とその構築、および今後のクラウドコンピューティングの利用に関する取り組みを紹介した。2014 年 9 月現在において旧環境から新環境への VM の移行はこれからの課題であり、全ての VM の移行と新たな VM の誘致は今後の課題である。仮想化基盤の設計はハードウェアのみならずハイパーバイザソフトウェアまで更新されるため、設計が基盤に対してどのような問題解決を提供したかを明示することが難しい。今後は設計や運用技術の優位性を明示する手法について模索することとしたい。

参考文献

- [1] Norbert Wiener: Cybernetics, Second Edition: Or the Control and Communication in the Animal and the Machine, The MIT Press (1965).
- [2] ノーバート・ウィーナー (著), 池原止戈夫, 彌永昌吉, 室賀三郎, 戸田巖 (訳): サイバネティックス 動物と機械における制御と通信, 青 948-1, 岩波文庫 (2011).
- [3] 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).
- [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] 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).
- [6] 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 (IOTS2010), Vol. 2010, pp.17-24 (2010).
- [7] 市川晃平, 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).
- [8] 宮永勢次, 市川晃平, 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).
- [9] 柏崎礼生, 宮永勢次, 森原一郎: 大阪大学の仮想化基盤の増強と設計, 情報処理学会研究報告, Vol.2014-IOT-25, No.26, pp.1-6 (2014).
- [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] Peter F. Drucker: The Deadly Sins in Public Administration, Public Administration Review, Vol. 40, No. 2, pp. 103-106 (1980)

^{*12} vCHS は 2014 年 8 月に vCloud Air へと名称を変更したため、本稿では vCHS ではなく vCloud Air の名称を用いることとする。

^{*13} <http://www.iozone.org>