

大阪大学における仮想化基盤の設計とその増強計画

柏崎 礼生^{1,a)} 森原 一郎^{1,b)}

概要: 大阪大学では 2009 年度に VMware 社による vSphere をハイパーバイザとして利用する仮想化基盤を構築し、その上で動作する仮想マシンにより様々なサービスを提供している。2014 年度に構築される学内データセンタ「IT コア棟」(仮称)への移設に伴い、また仮想化環境技術の急速な変化に追従するため、大阪大学では今後 5 年に一度のリプレイスではなく 2 年単位での段階的な増強とリプレイスを繰り返すことを検討している。本発表では現環境と新環境の設計について評価する。

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

Abstract: A virtualized infrastructure on Osaka University started its service in 2009. Hypervisor software of the infrastructure is vSphere that is products of VMware. Virtual machines (VM) on the infrastructure serve various services. On the other hand, data center in the university called IT core building will be completed construction in 2014. The infrastructure will have to be migrated to the building. Now, the authors are designing plans of migration and a future generation virtualization infrastructure. They consider to enhancement the infrastructure every 3 or less years but conservative every 5 years renewal to keep up with evolution of technique on virtualization infrastructure. In this paper, they evaluate the present infrastructure and new design.

Keywords:

1. はじめに

大阪大学は学部学生約 1 万 5 千人、大学院生約 8 千人、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。大阪大学は中長期的な目標として大学の ICT リソースを集約し、大学全体としての ICT 投資を効率化することを掲げている。それに合わせて大阪大学全体としては業務フロー全体の最適化を行い、業務の効率化も目指している。大阪大学程度の規模の総合大学では、部局ごとに独自に 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場合がある。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が少なくない。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、大阪大学情報推進機構は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でいる。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、現在は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る [1, 2].

大阪大学では本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。本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。現在、キャンパスクラウド上で事務系基幹システム、Web システム、そしてキャンパスメールサービスが動作している。本稿では本学のキャンパスクラウドへの取り組みの現状について述べる。

2. 現システム構成

キャンパスクラウドの構成図を図 1 に示す。本システムでは、仮想マシン (VM) を動作させるハードウェアとしてブレードサーバシステム複数台導入し、さらにファイアウォールやコアスイッチ、ロードバランサなど全てのネットワーク機器を冗長構成にすることでシステムの高可用性を実現している。ブレードサーバシステムは VM が利用す

¹ 大阪大学
Osaka University

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

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

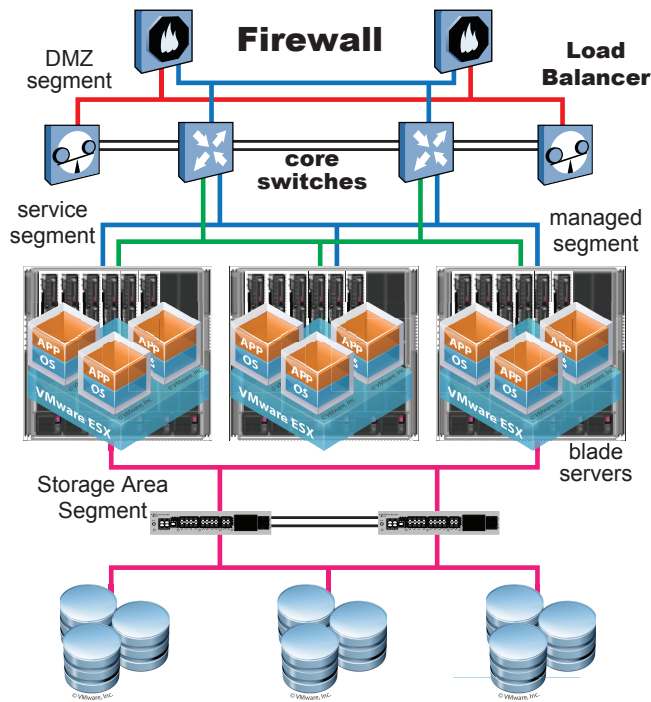


図 1 大阪大学共通基盤プラットフォームシステム「大阪大学キャンパスクラウド」概念図

Fig. 1 A Diagram of Osaka University Common Infrastructure System for Information Systems a.k.a “Osaka University Campus Cloud”

る CPU およびメモリ資源を省スペースな領域に高密度に配備できることから、キャンパスクラウドのように様々なサーバを集約するようなシステムにとっては非常に効率が良い。また、システム全体の要求リソースに応じてブレードを追加導入するといった拡張が可能であり、効率のよい ICT 投資を計画することが可能である。しかし一方で、ブレードエンクロージャのファームウェアアップデート時にはそのエンクロージャに含まれる全てのブレードサーバが電源断状態になるため、高可用性を実現するためには2つ以上のエンクロージャが必須となる。エンクロージャは6Uに8台、あるいは10Uに16台のサーバを搭載することができるモデルが一般的であるため、6Uエンクロージャであれば1エンクロージャあたり7台以上、10Uエンクロージャであれば11台以上のサーバを搭載し、かつ2エンクロージャ以上用意することでその省スペース性と高可用性を実現できる。そのためあらかじめ十分な CPU およびメモリ資源が必要であることが見込めているか、あるいは1エンクロージャで稼働させ、ダウンタイムを運用体制で吸収することが可能であることを見込む必要がある。本キャンパスクラウドでは導入当初はブレードエンクロージャ2台の構成から運用を開始したが、現在は学内からのサービスに対する要求の増加に対応して、1台を追加導入し、3台構成で運用している。

2.1 仮想プラットフォーム環境

本システムでは VM のプラットフォームとして VMware vSphere ESX を導入した。VMware vSphere は VM の実行環境(ハイパーバイザ)である VMware ESX をインストールしたブレードサーバ(ESX サーバ)群をクラスタ構成し、クラスタ内で実行される VM の集中管理を可能としている。VM の起動・終了・移動やインストール・設定などが全て集約管理可能なため、運用の統一による効率化が容易である。またクラスタ内の ESX サーバに障害が発生した際に、VM 単位でフェイルオーバーを行う HA(High Availability) 構成を設定することが可能であり、上述のスイッチ群の冗長構成と組み合わせることによって、システムの高可用性をより高めている。このような構成は、サービスごとに機器を導入する従来手法では低コストでの実現が困難であったが、現在のシステムではサービスを共用のプラットフォーム上に集約することで互いの障害リスクをカバーし合う運用を可能としている。

2.2 仮想ネットワーク環境

本システムでは仮想プラットフォームにおけるネットワーク環境を、Cisco 社の分散仮想スイッチソフトウェアである Nexus 1000V 用いて一元管理・設定を行なっている。本システムのように複数の ESX サーバ機をクラスタ構成し、仮想プラットフォームを構築する際は、ネットワークの管理・設定作業が非常に煩雑になる。VMware ESX は、個々の ESX サーバ内に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ことを標準でサポートしている。したがって、仮想ネットワークに関する設定は各 ESX サーバに分散し、一元管理を行うことが困難である。また、VMware が提供する仮想スイッチの設定は設定・運用方法は従来の一般的なネットワークスイッチと異なり、仮想プラットフォームの管理者とネットワークの管理者の作業の切り分けを困難にしている。

Nexus 1000V は上記の問題を解決する手段であり、複数の ESX サーバにまたがる、単一の巨大なネットワークスイッチを仮想的に実現することが可能である。これによって、仮想プラットフォームにおけるネットワーク設定が一元的に管理可能となる。また、ネットワークスイッチ設定のインターフェースとしてネットワーク管理者が従来から使用してきた Cisco IOS の CLI が利用可能であり、仮想プラットフォームの管理者の作業と、ネットワーク管理者の作業を容易に切り分けることを可能としている。本システムでは、この Nexus 1000V を用いて、VM に対しては基本的には以下の3つのネットワークセグメントを提供している。

- (1) サービスセグメント
- (2) 管理セグメント(含むバックアップセグメント)
- (3) ストレージセグメント

サービスセグメントは実際にエンドユーザからファイアウォールを経由してアクセスされるセグメントである。管理セグメントは VM 自体の管理に必要な通信や、バックアップデータが通信される。ストレージセグメントメールサービスのスプール領域や文書管理サービスのファイルストレージなど、ストレージ上にデータを確保する必要があるサービスに関して提供されている。これらのサービスセグメントは予め Nexus 1000V で設定されているため、VM を作成する際には VM のネットワークカードをサービスセグメントに接続するかを選択すればよい。この操作は、物理的なサーバのネットワークカードからケーブルをネットワークスイッチのポートに接続するかを選択することに似ており、直感的な理解を助ける。

3. 運用サービス

現在、キャンパスクラウドでは事務系基幹システム、キャンパスメールシステム、他部局の情報システムのホスティングを主に行っている。本章では特にキャンパスメールシステムについて説明する。

キャンパスメールシステムはメールサーバと、それに接続されたメールスプール、バックアップ装置、アプリケーションサーバと DB サーバ、さらにサービス監視システムからなる。メールサーバは負荷分散と冗長構成を目的に 3 重化している。アプリケーションサーバと DB サーバはそれぞれ 2 重化している。これらを構成するサーバは全てキャンパスクラウド上において VM として動作している。アプリケーションサーバは、アカウント作成・管理用の Web ユーザインタフェースを利用者に提供するものであり、キャンパスメールシステム用に今回独自に開発したものである。ここで設定された内容に従い、メールサーバの設定が生成される仕組みになっている。

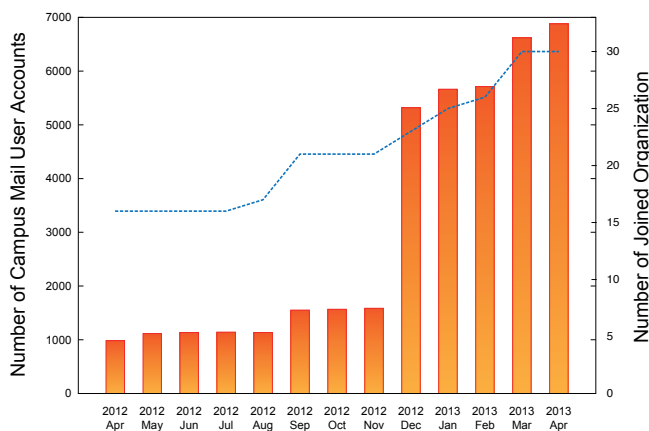


図 2 大阪大学キャンパスメールシステムのアカウント数と参加組織数の変化

Fig. 2 Changes of Number of User Accounts and Joined Organizations on Osaka University Campus Mail System

本学では、部局、専攻、講座の 3 階層の組織構造を持ち、部局によっては専攻ごとや講座ごとに管理者をおいて、それぞれ独自のドメインでアカウントの管理をしている場合もある。一方、部局全体で一人の管理者で運用している場合もある。キャンパスメールシステムにおいては、このような実際の管理運用の実態に柔軟に対応するように、大学内部の組織構造をドメインおよびそのドメインの管理者と関連させ、階層的に管理する仕組みを実装している。上位階層の組織に関連付けられた管理者は、下位の組織とドメイン、アカウントの管理を可能としている。本設計にあたっては、学内で既に大規模な統合サーバを構築していた工学部の統合メールサーバの設計を参考にした。本システムは、運用開始した平成 23 年 4 月から半年間で 7 組織、約 700 アカウントの移行を完了し、その後平成 25 年 4 月までに 30 組織、6,880 アカウントの移行を完了した。非常勤職員等を含めると大阪大学の教職員数は約 9,300 人であり、おおまかな計算では 74% 程度のアカウントがキャンパスメールシステムに収容されていると考えられる (図 2)。

4. 次期システムの設計

現行のシステムは物理コア 96 コア、432GB の主記憶、3.6TB の補助記憶装置からなるシステムであり、この上で 164 コアの仮想コアを利用して各 VM が動作している。現行システムのリプレイスは、導入から 5 年後の 2015 年に行われる予定であったが、2014 年にスーパーコンピュータやネットワークノードなどを格納するデータセンターである IT コア棟 (仮称) が竣工されることとなり、それに伴い現行システムの移転が必要となった。現行のシステムはブレードシャーシ 2 台から構成されているものの、VM が提供するサービスへの通信が不通となる時間をなるべく短くするために、通常より一年繰り上げて次期システムを導入することを検討している。現行のシステムと次期システムの間でストレージマイグレーションを行うことによりサービス断時間を短くするのが目的の一つとしている。またお互いに数 km 離れた複数の校地間で仮想化基盤を分散させることを検討しており、来年度に予定されている増強はその礎となるものである。

4.1 仮想化サーバの規模

現行のシステムでは物理コアと仮想コアの比率が 1:1.7 であるが、各仮想化サーバの CPU 利用率は低い状態で安定している。この比率を 1:4 に向上させることを設計の一つの指標としている。2 年後には比較的大規模な財務会計システムの完全な仮想化基盤への移行が検討されており、このシステムから現時点で要求されているスペックは仮想コアで 100 コア以上となっている。今後、学内において物理マシンで稼働するシステムをより積極的に集約することを鑑み、2 年後に要求される仮想コア数を 400 コアと設定

する。物理コア:仮想コア比を 1:4 とするため、必要とされる物理コアは 100 コアである。現在 Ivy Bridge の Xeon で 1 プロセッサ 12 コアの製品が販売されている。1 プロセッサ 8 コアの製品は安価であるが、プロセッサ数は VMware vSphere のライセンス数に直結する。120 物理コアを実現するために 10 プロセッサが必要な場合と 15 プロセッサが必要な場合では VMware vSphere Enterprise Plus エディション (アカデミック版) の場合、初期費用で 135 万円、年間保守料で 30 万円の差が現れる。Xeon E5-2697V2(2.7GHz, 12 コア) と E5-2650V2 (2.6GHz, 8 コア) の実勢価格を比較すると、1 プロセッサあたり 15 万円の差があるが、120 コアを構成した場合の差額 75 万円は VMware の初期費用差あるいは年間保守料金 2.5 年分で回収することが可能となる。

4.2 ソフトウェア化

機器費用および年間保守料金において大きな割合を占めるのがファイアウォールとロードバランサーである。これらの機器は十分な性能が要求されており、静的ファイアウォール 1Gbps 以上の性能が謳われた製品が導入されている。ロードバランサーは SSL アクセラレーションの機能を有し、4,400TPS(新規 SSL セッション) の性能を有する。実際にこれらのファイアウォールやロードバランサーでどの程度の通信が行われているかについて、運用開始後 3 年に渡り計測と保存が行われていなかったため不明であるが、半年分の計測結果からこの要求仕様がオーバースペックであることが分かっている。そのためこれらの機能についてはソフトウェアファイアウォールおよびソフトウェアロードバランサーで構成し、現在のハードウェアからの設定移行を構築費用に含めることを検討している。当初は Vyatta^{*1} の利用を検討したが、VMware NSX が提供するソフトウェアファイアウォールおよびソフトウェアロードバランサーの利用を検討している。現行環境のハードウェアで投入されている設定が移行可能であるかどうかを調査中であり、また Cisco NEXUS 1000V と機能的に重複がないかを確認している。

4.3 計測体制の確立

また今回の設計では、現行の運用において行われていない各種リソースの計測を行うことも次期システムにおける検討事項の一つである。仮想化基盤における CPU リソースや主記憶使用量、補助記憶使用量の推移はもちろんだが、各コンテナごとの消費電力を計測することで、各仮想化サーバにおける CPU 利用率の変化と消費電力の変化の関係を調査する。これまで大学の情報センターにおける仮想化基盤への集約には電力消費の節約による TCO の削減と

いう文言が並ぶことがあったが、しかしその実態において実測値でどの程度の金額の削減を実現できたかについて言及された事例は多くない。次期の設計の要は、計測を根拠とした身の丈に合った仮想化基盤の設計を 2~3 年周期で行う体制を構築することにもある。

また、研究開発段階にある広域分散ストレージについて導入することを検討しているほか、利用状況や負荷情報をリアルタイムで監視することができるシステムを構築することを検討している。研究開発段階にある広域分散ストレージについても導入を検討しており、3 校地にストレージを分散させて停電時の対応が可能となるような運用を実現することを目標としている。

5. まとめ

本稿では、大阪大学が取り組んでいるキャンパスクラウドについて述べた。キャンパスクラウドは事務系基幹システムの刷新の際に今後の学内システムの集約と効率化を目指して構築されたもので、来年度に向けた段階的移行のための設計と仕様書を作成中である。将来的な学外のパブリッククラウドサービスの利用も考慮に入れつつ、学内で個別運用されているシステムに関して、このキャンパスクラウドの集約を進めていく予定である。キャンパスクラウドの運用により各種リソースの利用率が把握することが可能となったため、今後の増強はリソース情報を根拠とした設計を行うことができる。今後の課題として、集約するシステムの増加に伴って、保守・運用手法の整理が問題となっている。現状の学内システムでは様々な運用がなされており、それぞれのシステムを開発・導入したベンダも異なる状況にある。そのような状況の中で、キャンパスクラウドとして一貫性のとれた保守・運用手法を定義するのは困難な状況にある。現在、本学でもこの問題を整理しきれず、解決方法を模索している段階にある。

参考文献

- [1] 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.
- [2] 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.
- [3] 秋山豊和, 寺西裕一, 岡村真吾, 坂根栄作, 長谷川剛, 馬場健一, 中野博隆, 下條真司, 長岡亨: 大阪大学における全学 IT 認証基盤の構築, 情報処理学会論文誌, Vol.49, No.3, pp.1249-1264, 2008.

^{*1} <http://www.vyatta.org>