

サーバ仮想化技術を利用した社内サーバの集約事例



テクニカルコンピテンシー部
基盤技術チーム
シニアITスペシャリスト

高橋 一博

Kazuhiro Takahashi
kazuhiro-takahashi@exa-corp.co.jp

サーバ仮想化技術の利用により、経営上および運用上のメリットがもたらされる。本稿では、オープン系のサーバ仮想化技術を概観した上で、Windows・Linux社内サーバを対象に、物理サーバを集約し仮想マシンを構築した事例を報告する。物理サーバとして、サーバ集約効果を高めるためにブレード・サーバを採用しており、その特徴についても紹介する。

1. はじめに

従来、仮想化といえばネットワークにおけるVLANやRAIDにおける論理ディスク・ボリュームといった技術を指していたが、近年はサーバ分野における仮想化が脚光を浴びている。これは、オープン系の仮想化を実現する製品が、パフォーマンス、機能、安定性、管理機能、さらにサポート体制などの向上が図られ、実用的になってきたことを示している。本稿では、そのサーバ仮想化技術(以後、サーバ仮想化技術を単に仮想化と略す)の紹介とともに、その技術を利用した当社内サーバの集約事例について報告する。

仮想化とは、単一の物理サーバをあたかも複数のサーバであるかのように使える技術のことである。仮想化では、単一の物理サーバ上においてハードウェア・リソースを複数に分割して提供し、おのおののリソース上でコンピュータの基本ソフトおよびアプリケーションを動作させることができる。仮想サーバとは、仮想化により物理的なリソースを割り当てられ、その上で稼働するオペレーティング・システム(以下、ゲストOSと略す)とアプリケーションを含めたものである。リソースが論理的に分割されているため、ゲストOS上で動くアプリケーションのトラブルにより、そのゲストOS自身に支障が生じても、ほかの仮想サーバが影響を受けてダウンしてしまうことはない。また仮想サーバの管理も物理サーバ同様に行うことができる。仮想化はこのように、複数の物理サーバで行っていたことをほとんどそのまま、単一のサーバで実現する技術である。

複数のOSを単一マシンに統合できる有利さは、社内に多数存在しているサーバの集約による効果をもたらす。弊社内には多数のサーバが存在しており、ファイル・サーバや開発・検証用サーバに使用しているx86系サーバは、本社地区だけでも200~250台と推定されている。それぞれが、物理サーバを占有することによる執務スペースの圧迫、サーバ台数分におよぶ管理の複雑さ、そのために個人が管理に費やす負担、マシンの老朽化などの問題を抱えている。これらの問題を解決するため、仮想化を利用して社内サーバの集約に取り組むこととした。ただし仮想化を実施するにあたり、物理サーバ上に仮想化を実現するためのサーバ・ヘッドや、既存サーバの仮想マシンへの移行方法確立など解決すべき課題がある。そのためパイロット・プロジェクトとして部内運用のファイル・サーバやWEBサーバ(以後、部内サーバと略す)の仮想化による集約に取り組み、併せて仮想化を実施する上での課題洗い出しと解決方法の

検討を行った。

以降の第2章では、仮想化技術とその特徴に検討を加え、第3章から第5章では当部でのパイロット・プロジェクト事例を紹介する。最後に、今回積み残した課題を対象とした、今後の展開について述べる。

2. 仮想化技術

本章では、仮想化技術の一般的特徴について解説する。

2.1. 仮想化技術の概要

(1) 仮想化の分類

オープン系の仮想化を大別すると、図1に示すようにホストOSベースと専用カーネル・ベースに分けられる。

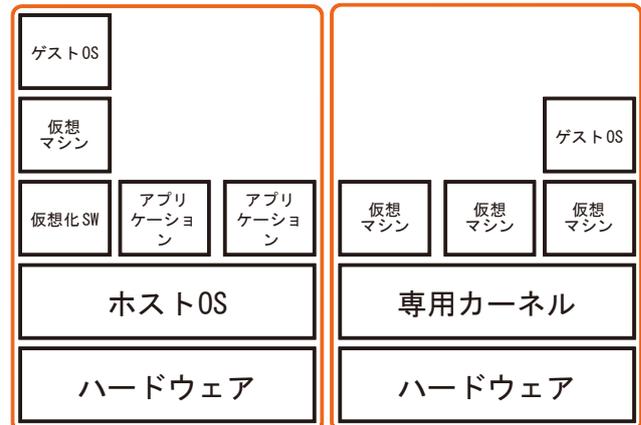


図1 オープン系仮想化の分類

ホストOSベースは、ホストOS(Microsoft Windows や Red Hat Enterprise Linuxなど)上に、仮想化ソフトウェアをミドルウェアとしてインストールし、その上に仮想サーバを提供するタイプである。専用カーネル・ベースは仮想化ソフトウェア自身が専用OSとなっており、その上に仮想サーバを提供するタイプである。

ホストOSベースは専用カーネル・ベースに比べ、ハードウェアとの間にホストOSが一階層余分に介在するため、仮想サーバとしての実効性能は劣る。そのため、高いパフォーマンスが要求されるIT基盤としての利用には限界が

ある。以降では、より汎用的なIT基盤としての利用を対象とすることとし、専用カーネル・ベースを中心に述べる。

なお専用カーネル・ベースの製品には、x86系ではVMware ESXサーバ、オープンソース Xen、Microsoft Hyper-Vなどがあり、XenをカスタマイズしたものとしてSun xVM Hypervisor、Oracle VMなどが発表されている。ハードウェアに依存した製品としてはIBM LPAR (Logical Partitioning)や日立製作所 MLPF (Multiple Logical Processor Feature)がある。

(2) 専用カーネル・ベースの仮想化

仮想化では、ソフトウェア(専用カーネル)がコンピュータ・ハードウェア(コア、メモリ、ハードディスクなど)を論理的に分割し、仮想マシンとして提供する。仮想マシンにインストールされるゲストOSは、仮想化されていない場合と同様、一台のコンピュータを占有しているものとして稼働する。それらの仮想マシンが単一の物理サーバ上で稼働する。この仮想マシン稼働環境を提供する専用カーネル(仮想マシン・モニタ、Virtual Machine Monitor、ハイパーバイザなどと呼ばれる。以後、ハイパーバイザと略す。またハイパーバイザはIBM製品の仮想化機構で使用される用語でもある)がハードウェア・リソースを制御し、スケジュールに従い個々のゲストOSに制御を移し、CPU特権が必要な処理についてはハイパーバイザが代行する。

当然であるが、ハイパーバイザは仮想マシンのオーバ・ヘッドとなる。

x86アーキテクチャでは、リング0からリング3までの4つのレベルで動作が可能となっており、デバイス・ドライバ、メモリ管理、プロセス管理など、OSの低レベル処理はリング0で実行される。しかし複数のゲストOSがリング0で実行され、あるゲストOSが特権命令でハードウェア・リソースを制御すると、ほかのゲストOSが正常に動作できなくなる。

そこでVMware ESX Serverなどでは、図2に示すようにハイパーバイザを専用カーネルとしてリング0で稼働、ゲストOSをリング1で実行させ、ハイパーバイザがリング0でハードウェア・リソースの制御を管理している。ところが、ゲストOS自身はリング1上で実行されているにもかかわらず、リング0での実行を前提に設計されているため、リング0レベルの特権命令を発行する。しかし特権命令はリング0でのみ実行可能な命令であるため、リング1で実行され

ているゲストOSからの発行はエラー(例外処理)となる。ハイパーバイザはそのエラーを監視し、特権命令の処理をエミュレーションしてゲストOSに戻すという処理を行っている。また一部の特権命令(例えば、セグメント・レジスタの値を読み出すMOV命令を実行すると、現在の特権レベル情報を知ることができるが、この命令をリング0以外で実行してもトラップされないなど)は例外処理で対応できないため、ゲストOSの命令発行を常時監視するという処理を行っている。これにより、元のOSの改変無しにゲストOSは通常の物理マシン環境の場合と同様に動作するが、このエミュレーションや監視という処理が、仮想環境でOSを実行する際の大きなオーバ・ヘッドとなっている。

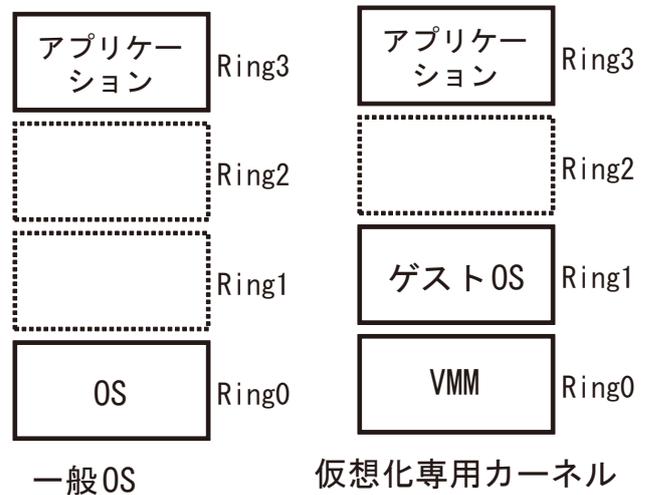


図2 仮想化専用カーネルの特権レベル

2.2. 仮想化のもたらす効果

仮想化によりさまざまなメリット、デメリットが生じるが、当部内のサーバ集約では、次のようなメリットを期待した。

(1) ITリソース利用率の向上

仮想化が脚光を浴びる前は、システムごとにITリソースが確保されてきた。あるシステムでの余剰リソースを他のシステムが利用するリソース共有は、RAID利用時に一部で行われてきた程度である。企業の中には目的に応じて多くのシステムが導入されている。例えば経理システム、人事システム、営業支援システムや在庫管理システムなどだ

が、これらは、処理のピークが特定の日に偏っており、ピーク時期そのものも異なる場合が多い。これらが別々に稼働している環境では、それぞれのピーク性能に合わせたリソースを確保しなければならず、個々のシステムが定常的に性能を維持するために要求されるピーク性能に合わせた構成をとっている。

また、システムが高可用性を要求するならば、システムごとに待機系のリソースを確保する必要も生じる。そのためシステムごとに余剰なリソースがさらに生まれ、全体として投資効率の悪いものとなっている。

しかし、仮想化を用いて複数のシステムを集約し、おのおのシステムが利用する資源を仮想的なリソース・プールとして共有すれば、ピーク時のリソースを必要な期間だけ占有すればよいことになる。負荷変動や可用性に備えていた余剰リソースは、複数のシステムで共有することで低減が可能となるなど、キャパシティ管理が効率化できる。

このように仮想化を用いれば、ITリソースの総量を最適化することで、リソースの利用効率を高められるメリットがある。

(2) 運用性の向上

運用面でも一部効率化が期待できる。例えば単一の物理サーバ内に複数の仮想サーバを集約できるため、CPU、メモリ、ディスクなどのパフォーマンス監視を行うことによって、一時的なリソース不足に対しては迅速に増強するなど、サービスレベルの向上が図れる。ただし、複数の物理サーバで仮想環境を構成した場合は、物理サーバの負荷を均一化させるために、仮想サーバの計画的配置が必要になる。また、個々の仮想サーバが提供しているサービスの監視は行えないので、運用性の向上は一部の局面に限られる。

また、開発用サーバを例にとれば、必要となった時点で必要なリソースを提供すれば、新たな購入手続きのための手間と時間を削減できる。本番機で発生したアプリケーション・エラーの再現は、同一構成の検証環境を仮想サーバとして構築することによって可能となる。また不要となった時点でそのリソースを開放すれば、有効利用の促進が図れる。

3. 仮想化製品の選定

社内x86系サーバの集約を行うにあたり、VMware ESXサーバまたはXen(およびXen系列の製品)が候補となる。

今回は以下の理由によりVMware ESXサーバを選択した。

- 対応できるゲストOSの種類が多い。
- オンライン・トランザクション処理を対象としたTPC-Cベンチマーク結果が良好であった(2007年度に実施)。
- VMware社が提供するハードウェア・リソースのパフォーマンス監視や仮想マシンのバックアップ作成などの管理ツールが充実している。
- サポートが充実している。

3.1. VMware ESXサーバが実現する主な応用機能

前章では仮想化の効果および仮想化を実現する基本技術について述べた。さらに、仮想化製品は仮想化の特長を活かした応用機能を充実させ、効果の向上を図っている。当部ではVMware ESX Server上に仮想マシンを構築した。ここでは、VMwareが提供している主な応用機能を紹介する。

(1) 仮想マシンの移行

VMware VMotionは稼働中のゲストOSをESXサーバ間でアプリケーションの中断無しに移行するサービスであり、仮想化された共有ストレージの利用が前提条件となっている。仮想マシン全体の状態は、共有ストレージに保存されたファイル・セットとしてカプセル化されており、移動元および移動先のESXサーバからのアクセスが可能となっている。そして、仮想マシンのメモリおよび稼働状態が、ネットワークを介して移動先ESXサーバに準備された仮想マシンに伝達される。情報の伝達が終了した時点でゲストOS稼働先の切り替えが行われるが、経験上反応が途切れるのは1秒以下である。

ダウンタイム無しでライブ・マイグレーションが可能なので、ハードウェア監視機構がウォーニングを発している場合や、ゲストOSの過負荷によってパフォーマンスが低下しているESXサーバから、仮想マシンを他のESXサーバに移行することにより、ESXサーバの故障によるダウンや、リソース不足を原因としたパフォーマンスの低下を回避することに利用できる。

また、VMotionは次に述べる高可用性やリソースの動的割り当てを実現する基本機能となっている。

(2) 高可用性

VMware HA (High Availability) をESXサーバに適用した場合、リソース・プール内の対象ESXサーバの稼働が監視され、ESXサーバが障害によってダウンした場合、稼働中だった仮想マシンは他のESXサーバ上で再起動する。

ただしサーバに障害が発生してもアプリケーションが稼働し続けることに主眼を置いたMicrosoft Cluster Serverでは、サーバは数秒で切り替わるのに対し、VMware HAではESXサーバの障害により仮想マシンがダウンし、ダウンした仮想マシンをクラスタ内の残りのESXサーバ上で再稼働させるフェイル・オーバー方式になっているため、サービス再開に数分を要する。なお、いずれにしてもデータ処理中のメモリやキャッシュの内容は失われるため、その対策は必要である。

(3) リソースの動的割り当て

VMware DRS (Distributed Resource Scheduler) は、メモリリソースやCPUリソースなどのクラスタ内にある全ESXサーバと全仮想マシンのリソース利用率データを収集・監視し、DRS設定パラメータに従い、仮想マシンに対する推奨配置先を含めたリソース割り当てを自動的に決定する。この機能により特定のESXサーバに負荷が集中することを回避し、ゲストOSの実行時性能を適切に維持するとともに、クラスタ全体の利用効率を高めることができる。また管理者が手動でリソースを割り当てることもできる。

3.2. サーバ管理ツール

サーバ管理には物理環境管理および論理環境管理がある。この両管理を統一ツールで行えることが、効率的な運用および管理者への負担軽減に欠かせない。

VMware VirtualCenterサーバは、物理環境である複数のESX Serverと論理環境である仮想マシン群を一つのコンソールで管理するための集中管理アプリケーションであり、GUIであるVMware Infrastructure Client (以後、VI Clientと略す)から利用する。

VI Clientの機能は以下のとおりである。

- VMware VMotion、VMware HA、VMware DRS管理機能。

- ウィザード形式の手順による新規仮想マシンのプロビジョニングの簡素化、テンプレートによる新規仮想マシン導入の簡素化。
- ESXサーバおよび仮想マシンのパフォーマンスと使用率を監視し、CPU、メモリおよびI/Oパフォーマンスの詳細を表示。
- タスク・スケジュール設定およびアラート設定による定期的な管理タスクの自動化。
- 承認された担当者にカスタマイズしたロールおよび権限を付与し、ESXサーバに関する権限を委譲。

4. サーバ集約の実施

3章で述べた、VMware ESXサーバを用いて、社内サーバを集約した事例を紹介する。

4.1. 概要

(1) 対象

当部で使用しているx86系サーバは約40台である。その内、老朽化している20台を集約対象とした。対象としたサーバのOSの内訳は、Windowsサーバ (Windows 2000 Server, Windows XP, Windows Server 2003) などが10台、Linuxサーバが10台となっている。

(2) 期待する効果

社内サーバにまつわる問題とサーバ集約による効果を表1にまとめた。この表に示したように、管理、運用面で効果が大きく、コスト、スペースの大幅な削減が期待できる。

4.2. 要求事項

今回のサーバ集約にあたっての要求事項を次に示す。

(1) 集約効果と拡張性

サーバ設置場所の集中・集約化により物理的統合を果たし、フロア・スペースの無駄な消費を抑える。なお、ハードウェア・リソースの追加に柔軟に対応できる構造とする。仮想化統合によりリソース有効利用を図り、開発・検証用

サーバの提供要求に対して柔軟に対応する。また、仮想サーバのダウンから復旧に要する時間を60分以内とする。

(2) 性能の保証

既存の部内サーバ・開発用サーバが提供していたCPU・ディスクI/O・ネットワークに関する実行時性能を保証する。

(3) 既存サーバの移行

既存サーバの仮想化は、移行ツールの利用によって、移行に伴う作業時間の低減を図る。

移行ツールが適用不可能の場合は、仮想マシンを新規作成しOSなどの再インストールを行った後、データの移行により対応する。

(4) 管理ツール用サーバの仮想化

管理対象のESXサーバが増大するにつれて、VirtualCenter

サーバの能力は増強しなければならない。柔軟な対応を可能とするため、VirtualCenterサーバを仮想化しESXサーバ上で運用し、性能の増強はリソース・プールからの追加で対応する。

(5) 運用

基本的には全体管理者がすべてのESXサーバおよびディスク装置の管理を行うが、ユーザの要求に応じて、ESXサーバ単位での管理権限を委譲可能とする。

仮想サーバはその構築が終了した時点でバックアップを取得する。仮想サーバが生成するデータについては、ESXサーバ管理者およびユーザがバックアップ可能とする。バックアップ経路については、FibreChannelおよびEthernetの2経路を提供する。

5. ハードウェア構成の設計とESXサーバの設定

本章では、4章で述べられた要求事項について、その実

表1 サーバ集約効果

社内・部内サーバの現状	サーバ仮想化による効果
デスクサイド型のサーバが多く執務スペースを侵略している。またスペース維持費用の負担も大きい。	一台のサーバ上で複数の仮想マシンを実行することにより、省スペース化が図れる。ブレードサーバを採用すれば、より集約効果が高められる。ただし初期導入費用が発生する。
各サーバの管理者は基本的にユーザが兼ねており、各管理者に負担が生じている。	物理的な管理および保守をサーバ管理者に委ねることにより、管理負荷が軽減される。
既存サーバの多くが数年以上前に導入されたため、不具合の発生も見受けられるようになったが、障害復旧が短期間で行える保証がない。	新規サーバに移行することによってMTBFの向上が見込まれるだけでなく、HA機能の利用も可能となる。
サーバの管理が個々に委ねられているため、データのバックアップが確実とは言えない。	FC-RAIDの利用によりデータの冗長性が確保できる。またサーバ管理者により、統一的・効果的管理が可能となる。
急な開発用(および検証用)サーバの調達には時間がかかる。	仮想リソース・プールに開発用(および検証用)サーバの利用分を見込んでおくことにより、柔軟に対応できる。
開発が終了した時点で、次の転用までの非稼働期間が存在する。	仮想リソース・プールの活用で、ある程度の回避が可能となる。
現在、サーバの設置は執務室内である。入室時社員認証が行われているが、セキュリティが万全とは言えない。	入室制限を課した部屋にサーバを集約すれば、セキュリティを確保できる。
プロジェクト毎に開発環境構築コストが必要。	標準化された環境であれば、開発環境構築コストが低減できる。
既存サーバを更新するとなると、新しいOSにアプリケーションを対応させねばならない。	ほとんどが既存OSに対応可能なので、アプリケーションの更新が不要となる。
開発用サーバに既存のサーバを手当てした場合、仕様に完全に合致できない場合がある。	CPU、メモリ、ハードディスクについては、仕様に対応可能なケースが多くなる。

現方法と注意点について述べる。

5.1. ハードウェア構成の設計

サーバの物理的統合のため標準ラックにブレード・シャーシ、ブレード・サーバ、FC(FibreChannel) RAID、VirtualCenterサーバ、バックアップ・サーバおよびテープ・ライブラリを搭載する。ブレード・サーバにはFC HBA(Host Bus Adapter)を装着し、FC RAIDによって記憶装置の集中と仮想化も実現する。

実際のブレードの構成を表2 ブレード・サーバのスペックおよび構成に示す。ブレード・サーバ式は物理サーバ式に相当し、1~2CPU(8コア)、2~16GBメモリ、システム・ディスク、NIC(Ethernet 1000Base-T x 2)およびオプション・カードを搭載可能であり、サーバとしての基本機能を満たしている。また1ベイを占有する拡張ボックスを用いることにより2~16GBメモリやディスクの追加が可能となっている。ブレード・シャーシに採用したIBM/BladeCenter Hは4Uの筐体にブレード・ベイを14持つ。そのためブレード・サーバを最大14枚収容可能と省スペースを実現できる。

ESXサーバは2CPUで1ライセンスなので、1ブレード・サーバあたり1ライセンスとなる。今回、ブレード・サーバはType-1を5式、Type-2を3式とし、おのにおにESXサーバをインストールする。4式は当部内サーバの集約を目的とし、残りの4式は他部署向けの開発・検証用サーバとして試験的に提供し、全社展開に備えた基礎データ取得を目的とする。

5.2. ESXサーバの設定

4章の要求事項を満足し、既存利用環境と同等以上のパフォーマンスを保証するために以下の構成とした。

(1) ブレード・サーバ

ブレードを構成するにあたり、主な決定すべき仕様は、CPU(クロック、コア数)、メモリ量およびオプション・カードの種類である。なお、現在のブレードはデュアル全二重Ethernet 1000Base-T NICの標準装備が前提となっている。

「専用カーネル・ベースの仮想化技術」で述べたハイ

パーバイザに相当する機能を、VMwareではVMkernelが遂行する。ゲストOSの実行用には仮想マシンを構成する仮想CPU(クロック、コア数)、仮想メモリ、仮想ディスクおよび仮想ネットワークが提供されるが、実際のリソースの制御はVMkernelが行っており、主な役割は以下のとおりである。

- ・ ゲストOSでのCPU例外処理バイナリ・トランスレーションによるコード変換。
- ・ ゲストOSと物理サーバ間のメモリアドレス変換。
- ・ ESXサーバ内仮想ネットワーク・スイッチの生成とそのパケット処理。
- ・ ゲストOSによるディスクI/Oをエミュレーション。

現在のところVMkernelは一つのコアを使用するシングルタスクとして働き、上記のサービスを処理しているのでVMkernelそのものの過負荷に留意せねばならない。また、メモリを十分に確保することによってゲストOSでのスワップ発生を防ぐ必要がある。スワップの発生はゲストOSのパフォーマンス低下だけでなく、それを処理するVMkernelが行っているメモリ管理に大きな負荷を与え、ESXサーバとしてのパフォーマンス低下を引き起こす。そのため他のゲストOSのパフォーマンスも低下させることになる。

ESXサーバのライセンス費用だけを考慮すると、一台のESXサーバ(ブレード・サーバ)に実装可能なCPU、メモリを最大限搭載すべきであるが、VMkernelの処理能力を超えるゲストOS数を稼働させても性能が低下するだけであり、ハードウェアに対する投資が無駄に終わる。最適なゲストOS数を稼働させた場合よりも、ESXサーバとしてのアウトプットが下回ることに留意すべきである。今回、既存サーバの集約では、5~7ゲストOS/ESXサーバを目安とし、運用実績を見た上でさらにゲストOSの追加を検討することとした。

(2) ファイル・システム

ESXサーバが仮想マシンに提供するディスクの多くはiSCSIやFibreChannel RAID(以後、FC RAIDと略す)である。今回VMFS(ESXサーバが使用するファイル・システム)用として、以下の考えのもとにFC RAID(今回はIBM/DS4700)を選択、構成した。

- ① ESXサーバからディスクへの経路がボトルネックに陥らないこと。そのために、4GbのFibreChannel Host Bus Adapter(以後、FC HBAと略す)を採用した。
- ② FCスイッチの採用により、FC HBAやFC RAID側コントローラの障害に対するFCの経路切り替えは、VMwareが提供する機能をそのまま使用した。
- ③ 共有ディスクには常にディスクI/Oが集中することを念頭に置いた。仮想化で実現される個々の機能、特にゲストOSをESXサーバ間でライブ・マイグレーションする機能を使用するならば、FC RAIDの利用は欠かせない。この場合、ゲストOSから見たローカル・ディスクはFC RAID内に生成される。これは複数のゲストOSを載せたESXサーバがクラスタとなり、一台のFC RAIDを共有している形となる。表現を変えれば多数のサーバが一台のFC RAIDに対してランダムなRead/Write処理を行うことを意味する。つまり、VMkernelの処理能力ではなく、FC RAIDの処理能力が稼働するゲストOS数を制限する要因となる。今回は、10個のFCディスクを4+1、4+1のRAID 5アレイとし、パリティ・ディスクの負荷を軽減する構成とした。

仮想サーバがVMotionやHAを必要としないケースもある。そのような場合は、FC RAID上にゲストOSのローカル・ディスクを持たせる必要はなくなる。今回は、表2に示すように、ブレードに2種類のタイプを用意した。Type-2は基本ブレード+ストレージ拡張モジュールとなっている。ストレージ拡張モジュールはESXサーバから見るとVMFS用ローカル・ストレージであり、VMware VMotionやHAなどのサービス対象外となるが、共有ディスクへのI/O集中回避が可能となっている。これは開発用サーバなどの、必ずしもサーバとしての可用性要件が厳しくない用途を想定している。

なお、FC RAIDの導入時にWindows上のアプリケーションである管理マネージャと、装置内ファームウェアのメジャー・バージョンの一致作業などが発生したが、手順については別途問い合わせをいただきたい。

(3) NICチーミング

ESXサーバには仮想ネットワーク・スイッチへの接続タイプとして、VMkernel、VMconsole、仮想マシンの3種類がある。それぞれ、機能および使用されるケースが違い、

表2 ブレード・サーバのスペックおよび構成

		構成可能範囲	Type-1	Type-2
CPU	CPU数	1 or 2	2	2
	クロック	2GHz ~ 3.33GHz	2.5GHz	2.5GHz
	コア数	2 or 4 / CPU	4 / CPU	4 / CPU
メモリ		1GB ~ 32GB	16GB	16GB
ディスク	内蔵	73GB or 146GB x 2 (SAS), RAID0, 1, 10可	146GB x 2 (SAS), RAID 1	146GB x 2 (SAS), RAID 1
	増設	73GB or 146GB x 3 (SAS), RAID0, 1, 5, 6, 10可	—	146GB x 3 (SAS), RAID 5
ネットワーク(標準で装備)		デュアル全二重 Ethernet 1000Base-T	デュアル全二重 Ethernet 1000Base-T	デュアル全二重 Ethernet 1000Base-T
オプション	ネットワーク	デュアル全二重 Ethernet 1000Base-T	—	—
	Fibre Channel	4Gb SFF FC HBA	4Gb SFF FC HBA	4Gb SFF FC HBA
	ネットワーク&FC	デュアル全二重 Ethernet 1000Base-T & 4Gb SFF FC HBA	—	—
拡張ボックス	メモリ拡張モジュール	メモリ拡張 I/O拡張カード x2	—	—
	ストレージ拡張モジュール	ディスクド x3 I/O拡張カード x3	—	—

理想的には個別の仮想ネットワーク・スイッチを用意し、おののちに物理NICをアップリンクすべきである。しかし今回のブレード・サーバのNIC数は2であり、要求数を満足することはできない。VMkernelとVMconsoleを一つの仮想ネットワーク・スイッチに接続し、仮想マシン群をもう一つの仮想ネットワーク・スイッチに接続する方法が考えられるが、この場合、NICの障害に対しての冗長性が確保できない。

VMkernelを使用するケースとして想定されるのがVMotionである。今回、VMotionを行うケースはESXサーバの障害発生時がほとんどであると判断し、仮想ネットワーク・スイッチは一つとし、すべてのタイプを接続することとした。NIC数は2であるので、vnic0,1をアップリンクさせNICチーミングを有効とした。この方法によりネットワーク・アダプタは2つとも有効になるので、仮想マシン群が両NICを利用でき、スループット向上にもつながる。

5.3. 既存サーバ移行時の注意事項

今回行った既存サーバの移行作業から、一般的に考慮すべき点について記す。

(1) 移行先環境

- ・ 仮想マシンは時刻が狂う。移行先には時刻同期の仕組みが必要。
- ・ 利用中のファイル・サーバの移行は作業時間も長く、ユーザに対する周知が必須。
- ・ 移行元が生きてると同じホスト名が使用できないので、切り替え時期の調整が必要。
- ・ (Windowsの)ドメインを利用している場合、移行先でのドメイン・サーバ参照を解決しておく。

(2) 移行ツール利用

- ・ 移行の成功は仮想マシンのブート成功までわからない。
- ・ 移行ツールによる移行はパーティション単位で行われる。システム・ディスクにデータを含む場合は多くの時間が必要で、移行が失敗した場合、成功までに費やす作業時間とネットワークに関する負担が大きくなる。
- ・ 特別なデバイスを有するサーバは、ツールが対応できない。

- ・ ウィルスチェック、OSやデバイスの監視などを行っているサーバは、移行に失敗する可能性が高くなる。

5.4. 管理ツール用サーバの仮想化

VMware ESXサーバの管理はWindows Server 2003上のアプリケーションであるVirtualCenterサーバによって行われる。今回、VirtualCenterサーバが稼働するWindows Server 2003を、VMwareが提供する移行ツールによって仮想化し、ゲストOSとしてESXサーバ上に移植した。

VirtualCenterサーバをゲストOSとして移植するメリットであるが、仮想化によってハードウェア障害による影響を低減できる、CPUやメモリ追加に柔軟に対応できることが挙げられる。デメリットは管理すべきESXサーバの影響下に監視用サーバが入ってしまうことである。

実際、すべてのゲストOSがブートできなくなるというESXサーバのクリティカル・バグが発生し、パッチを適用するためにVirtualCenterサーバを含めたゲストOSを停止しなければならない事態が発生した。パッチ適用後、ゲストOSのブートに成功し、通常管理体制に復帰できたが、このケースなどはデメリットがそのまま表面化したものである。

6. おわりに

今回は仮想化製品を用いて社内サーバの集約化を実施し、設計、構築、移行上の課題とその解決策について事例を紹介した。まだ稼働を開始したばかりであり、運用面・性能面での検証はこれからである。運用面の検証の手始めとして、複数のバックアップ手法を対象にそれぞれの有効性の比較・検証に着手したところである。また、バックアップとともに重要なのが、ESXサーバの性能管理である。これについては、VirtualCenterサーバをとおした監視を中心にを行いデータを収集し評価を行っていく予定である。

現在は、当部の既存サーバおよび一部の開発用サーバの利用となっているが、運用実績を蓄積し、ユーザにとって快適な運用環境・運用規則を整えた段階で、全社展開する予定である。

参考文献

- 1) 平谷靖志 独自の仕組みでx86アーキテクチャ上に仮想化を実現 ITpro 2006/10

Microsoft、Windows は、米国 Microsoft Corporation の米国 およびその他の国における登録商標です。

Linux は、Linus Torvalds の米国およびその他の国における登録商標 あるいは商標です。

Red Hat は、米国Red Hat, Inc.ならびにその子会社の米国およびその 他国における登録商標若しくは商標です。

VMware, VMotion は、VMware, Inc.の米国および各国での登録商標または商標です。

Sun に関連するすべての商標は、米国 Sun Microsystems, Inc. の米国およびその他の国における商標または登録商標です。

Oracle は、米国Oracle Corporationおよびその子会社、関連会社の登録商標です。その他の名称は、Oracle Corporationまたは各社が所有する商標または登録商標です。

IBM は、IBM Corporation の登録商標です。

その他の会社名、製品名およびサービスは、それぞれ各社の商標または登録商標です。
