この1冊でスッキリ解る!

vSphereの基礎の基礎

新卒SEの視点で学ぶ、vSphereの「はじめの一歩」

vmware®

さあ、vSphereの扉を開こう!

今や仮想化ソリューションの定番として
欠かせない存在となっているVMware vSphereですが、
いざその中身や機能を理解しようとすると戸惑うことも多いのではないでしょうか?
本書は、VMwareのスタッフブログで
多くのアクセスを獲得した新卒SEによる入門シリーズです。
vSphereの基本的な構成から機能までをわかりやすく解説して、
仮想環境の構築に挑戦・活用していくための足掛かりを得られるようになっています。
これから仮想化をはじめたいという皆さまはもちろん、
すでに仮想化を導入している皆さまも、ぜひともご一読ください。

Index

1 vSphereの基本構成を理解しよう!

VMware vSphere ～仮想化基盤の中心的存在～

VMware vSphere（以下vSphere）とはVMware vSphere ESXi（以下ESXi）とVMware vCenter Server（以下vCenter）を含む仮想化ソフトウェアのスイートの総称です。このvSphereにより、仮想化プラットフォームを実現することができます。vSphereの基本コンポーネントは、以下のようになっています。

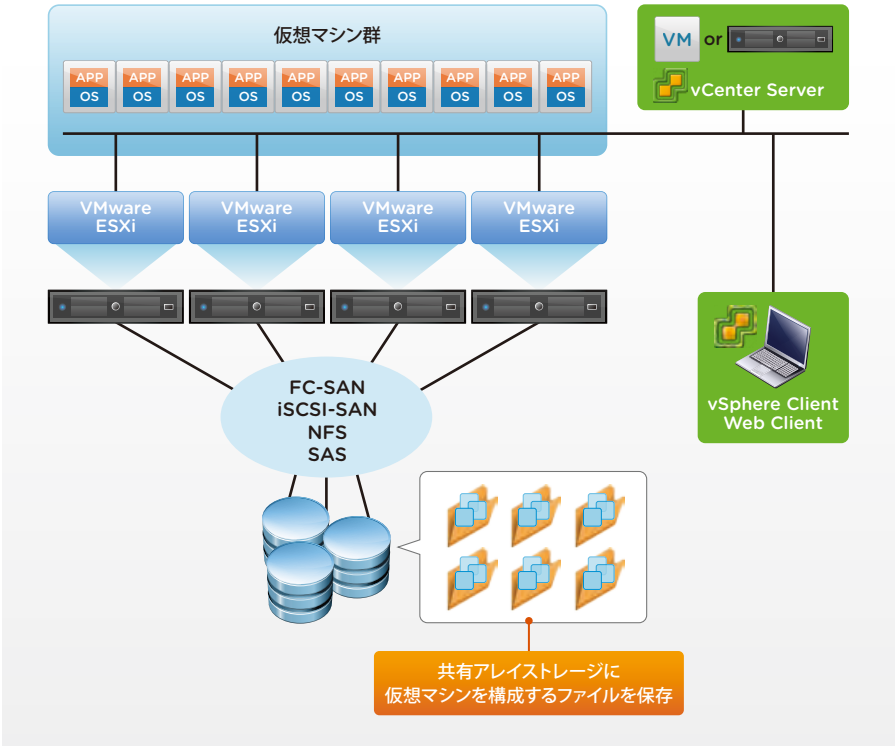


図1：vSphereの基本コンポーネント

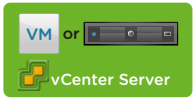
この図は、vSphereの全体像を表しています。各物理ホストの上にハイパーバイザーであるESXiが敷かれており、その上でゲストOS、アプリケーションが動いています。そしてESXiの管理を束ねているのがvCenterです。vSphere環境の管理者はvSphere Client、またはvSphere Web Client（以下Web Client）を用いてvSphereという仮想環境の管理を行うことができます。それでは登場する各用語を押さえていきましょう!

VMware ESXi ～vSphereの根幹をなす仮想化ソフトウェア～

ESXiは、vSphereの中核となるハイパーバイザー型仮想化ソフトウェアでホストOSの代わりにハードウェア上で直接動作する仮想化のためのOSようなものです。ハイパーバイザー上では、複数の仮想マシンを実行することができます。各物理サーバの上にWindowsやLinuxといったOSを直接インストールするのではなく、ESXiをそれぞれの物理サーバにインストールしておくことによって、1つの物理サーバ上で複数のOSを動かすことが可能となります。ESXiのインストールも簡単で、慣れていれば10分程で終えることができます。

図2：ESXiインストーラーブート画面

VMware vCenter Server ～仮想基盤の司令塔～



vCenter Serverとは、仮想基盤を管理するために必須のコアサービスで、vSphere環境の管理一元化を行います。1台の物理サーバにvCenter Serverをインストールしてvcenter Serverとして使うこともできますし、仮想マシンにvCenter Serverをインストールして使うこともできます。またvCenterがインストールされたアプライアンスも用意されておりますので、簡単にvCenterを展開することも可能です。vCenter Serverの役割として大きく2つあります。

- **統合管理（複数のESXiを束ねて管理）**
- **vSphereにあるさまざまな機能を有効化**

vSphereにある「さまざまな機能」に関しては、後ほど解説していきます。

vSphere Client/vSphere Web Client ～仮想基盤の入り口～

vSphere Clientとは仮想環境にアクセスするための、言わばvSphereへの入り口となるインタフェースを提供します。vSphere ClientはWindowsマシンにインストールして使用しますが、Webベースのvsphere Web Client（以下 Web Client）を用いることによって、ブラウザベースのvSphere環境の管理ツールによるvSphere基盤の運用・監視が可能になります。従来のvSphere管理機能はvSphere Clientのみでしたが、Web Clientが登場し、今後はWeb Clientに統一されます。

vSphere 5.5以降の追加機能はWeb Clientで対応しますので、Web Clientの操作に慣れておくことをお勧めします。一部アドオン製品のプラグインに関しては、vSphere Clientでしか使えない機能もあります。

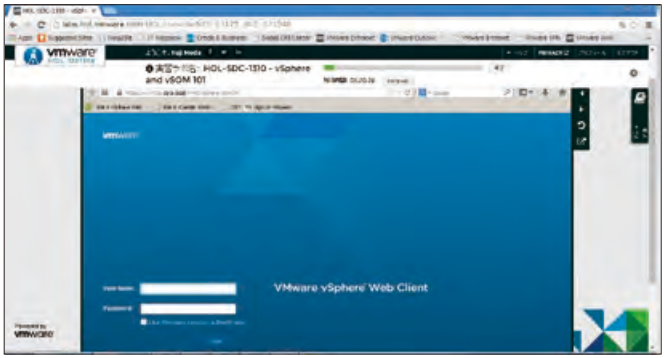


図3：vSphere Web Clientのインタフェース

仮想マシン ～仮想マシンの実体はファイル～

本章の冒頭で各物理ホストの上にESXiが敷かれており、その上でOSやアプリケーションが動いているとお話しました。ここで言う物理ホストとは物理サーバのことを指し、この物理サーバ上に直接ESXiがインストールされています。このESXiがインストールされたサーバのことを、通称「ESXiサーバ」と呼んでいます。そしてこのESXiが、WindowsやLinux等のOS＝ゲストOSやアプリケーションを入れる器を作り出します。この器が「仮想マシン」なのです。

この仮想マシンでは、物理環境でいうCPUやメモリ、HDDといった装置もESXiによって仮想化されたソフトウェアとして定義され、仮想CPU、仮想メモリ、仮想HDDとして存在しています。vSphere 5.5では、仮想マシンに対して最大64 vCPU（仮想CPU）、1TBのメモリを割り当てることができます。

また、仮想マシンの実体は「ファイル」です。全ての仮想マシンの情報はファイル（.vmdkや.vmx等）としてストレージに保存されています。ファイルなので簡単に複製することができ、ネットワークを通じて遠隔地に同じ構成の仮想マシンを作成することも、簡単にバックアップをとることも可能になります。物理環境では多大な時間とコストがかかってしまい、敷居が高くなってしまう災害対策ですが、vSphere環境なら、そういった敷居を大幅に下げることができます。

下図は仮想マシンのイメージになります。この例では、ESXiサーバ上に3台の仮想マシンが載っています。それぞれ仮想マシンにはCPU、メモリ、NIC、Diskが載っていますが、実際にはハイパーバイザーが各仮想マシンに物理リソースを割り当てています。仮想マシンに入るゲストOSは「仮想環境で動いている」ということを意識することなく、割り当てられたリソースを使って動いているのです。

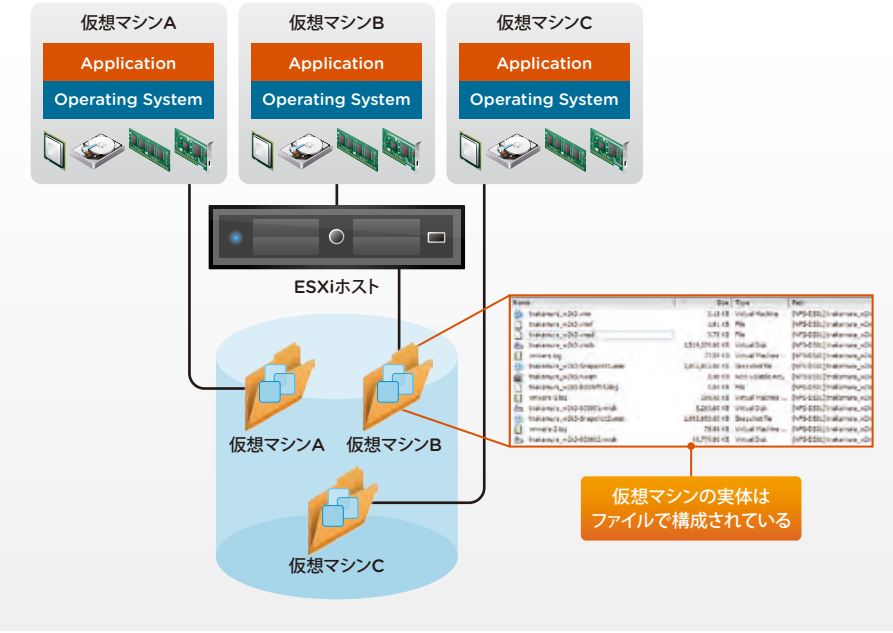


図4：仮想マシンの実体

共有ストレージ ～仮想マシンの家～

vSphere環境では共有ストレージがほぼ必要となってきます。共有ストレージへの接続方法としては、FC、iSCSI、NFS等、選択可能です。この共有ストレージに仮想ディスクファイルが保存され、共有ストレージに保存された仮想ディスクファイルを読み込むことで、仮想マシンを動かしています。

2 仮想マシンはなぜネットワークに接続できるのか

仮想マシンのネットワーク概要 ～仮想化されたネットワーク機器～

仮想マシンは、1台の物理サーバ上で複数台動作させることができる一方で、物理サーバに搭載できるNICの枚数は限られています。では、どうやって仮想マシンはネットワーク接続を行っているのでしょうか？

仮想マシンにはCPU、メモリ、ストレージ等が割り当てられており、仮想マシンに入っているOS（＝ゲストOS）は、あたかも物理サーバ上で動作していると思い込んでいます。ネットワークの接続を行うため、NICも他のハードウェアと同様に仮想的なハードウェアとして仮想マシンに搭載できます。この仮想NICを「vNIC」と呼びます。ゲストOSは本物のNICだと思い、このvNICにIPアドレスを割り当て、通信を行います。

vNICは仮想的なNICなので、仮想マシンがvNICから送信しようとする信号を物理ネットワークへ送るためには、ESXiサーバに搭載された物理NICとの紐付けが必要になります。しかし前述の通り、1つのvNICにつき1つの物理NICを割り当てるとなると、ESXiサーバには膨大なNICが必要になってしまいます。そこでvSphereは、ESXiサーバの内部で仮想的なスイッチ「vSwitch＝仮想スイッチ」を作り、ネットワークコントロールを行っています。下の図1をご覧ください。まさに物理サーバにケーブルを接続してスイッチに接続するという行為を、ESXiサーバ内部でも実施しているイメージです。

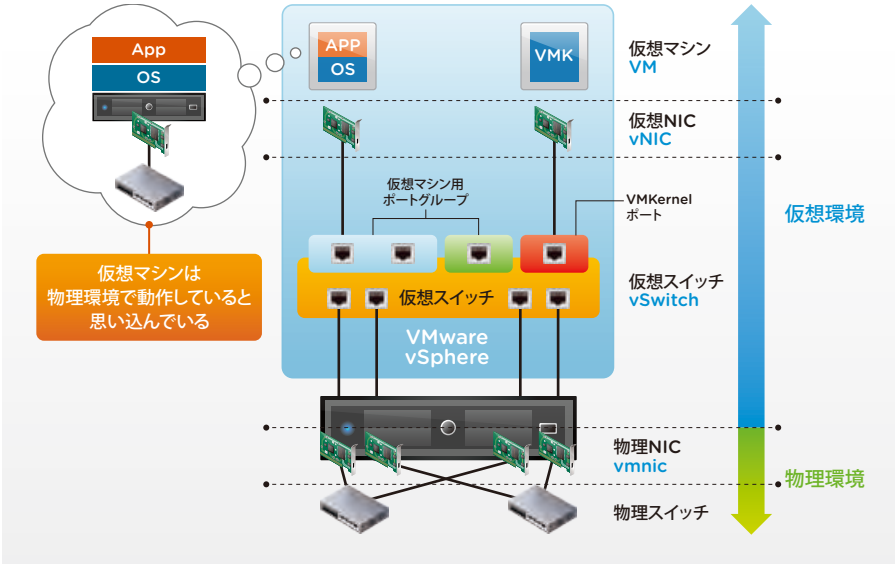


図1：仮想マシンの物理ネットワークへの接続

図1のうち、「物理」と書かれた2つの機器（物理NICと物理スイッチ）が実際に「目に見える」機器で、「仮想」と書かれた機器は全てESXiサーバ内部で実現される「目に見えない」機器です。つまりESXiサーバ内部では物理ネットワークと同じように仮想的にネットワークの構築に必要な機器を作り上げ、仮想マシンがネットワークへ接続できるようにしています。

vSwitch ～物理と仮想を繋ぐ装置～

vNICとvSwitch、２つの機能についてご紹介させていただきましたが、vSwitchには、vSphereならではの考え方がありますので補足しましょう。図２をご覧ください。

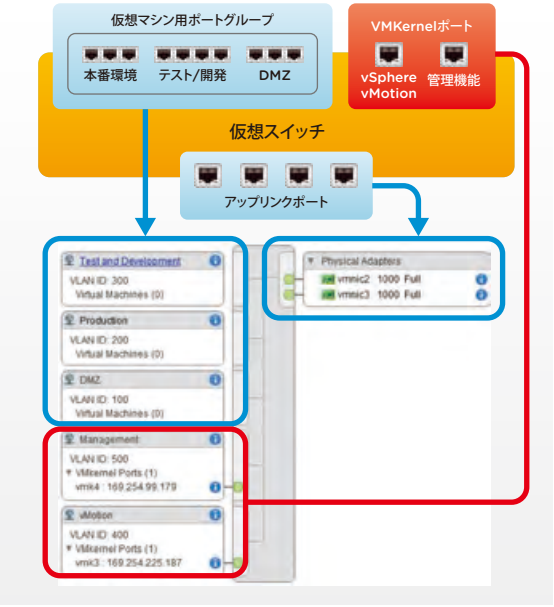


図２：vSwitchのポートの種類（上：概念図、下：実際の管理画面）

vSwitchには３種類のポートが存在します。まず、物理NICと対応づけられる「アップリンクポート」、ESXiサーバの管理やvMotion、vSphere HA、vSphere FTなどvSphereの機能を使用するための「VMkernelポート」、そして最後に仮想マシンのvNICを接続するための「仮想マシンポートグループ」です。

ここで、仮想マシンの接続用ポートだけ「ポートグループ」とされていることに注目してください。vSwitchは、各ポートのポート番号で接続を管理するのではなく、ひとまとまりのポート群＝「ポートグループ」でポートを管理しています。このポートグループはL2レイヤのネットワークを形成しており、VLAN IDもポートグループ毎に割り当てることができます。

図２の下の画像はvSphere Web Clientからみたvswitchの管理画面のキャプチャです。この例ではESXiサーバ内のネットワーク（左側）はポートグループ毎にわかれてvSwitchに接続され、アップリンクポート（右側）には物理NICが割り当てられていることが解ります。物理NICは「vmnic0、vmnic1、vmnic2、vmnic3、…」など、「vmnic」というラベルを付けてESXiサーバが管理していることが確認できます。ここでアップリンクポートにvmnicが２つある理由は、物理NICの冗長性を確保するためです。vSwitchに複数の物理NICを割り当てることによって、冗長性を確保する設定を簡単に行えます。

今回ご説明したvSwitchは「標準vSwitch」と呼ばれ、ESXi内部に複数作成することが可能です。ESXiサーバの台数が増えてしまうと仮想スイッチの管理も複雑になってしまうため、複数のESXiサーバにまたがって仮想スイッチを一元管理できる「vSphere Distributed Switch」という機能も存在します。これについては、後ほど解説することにしましょう。

3 vSphereにおける ストレージの考え方

共有ストレージの概要 ～実際の作業からキーワードを知る～

vSphereには、仮想マシンが動的に他のESXiサーバに移行する「vMotion」や、ESXiサーバが停止した際に他のESXiサーバから仮想マシンを再起動させる「vSphere HA」、2つのESXiサーバで同一の仮想マシンを動作させてダウンタイムなしの可用性を実現する「vSphere FT」などの機能がありますが、この機能の実現にはストレージが大きく関わっています。ここではまず、vSphere環境でストレージを使用するまでの手順を追いながら用語と概念を理解していきましょう。

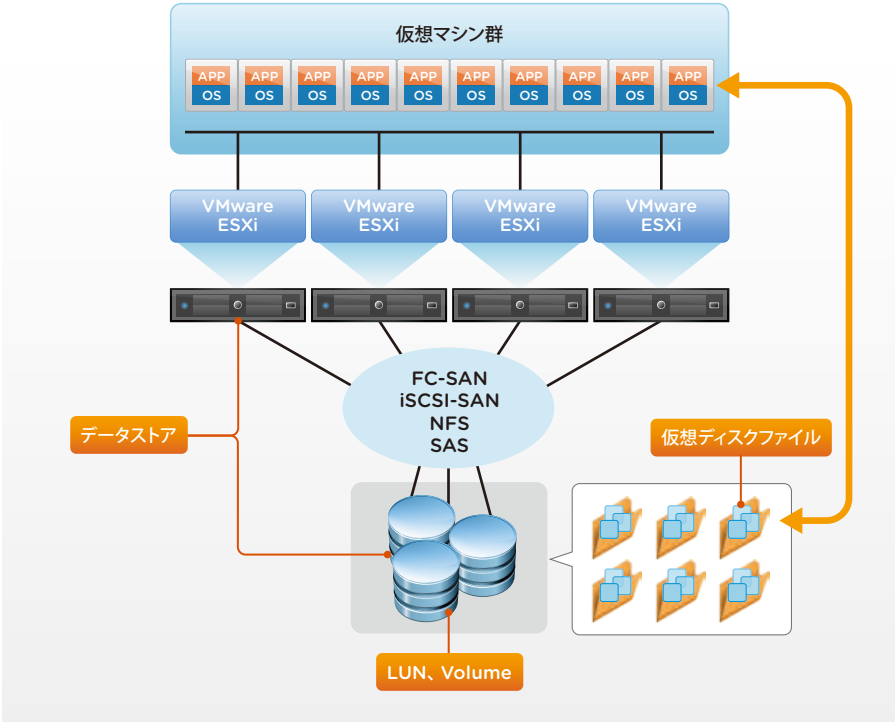


図1：vSphereの構成とストレージの関係

LUN、ボリューム ～ESXiとストレージの接続～

ESXiサーバが新しいストレージを使用するためにはまず、ESXiサーバがストレージを認識する必要があります。設定を行うと、ESXiサーバはLUNやボリュームを検出し、次に検出したLUNを仮想マシンファイル等を収容するための「データストア」として登録します。

図２はiSCSIソフトウェアアダプタを確認した際にESXiサーバに接続されたiSCSIストレージを確認している画面です。画面中央下段の「デバイス」タブで、検出したLUNを確認できます。ここでは45GBのLUNが見えています。

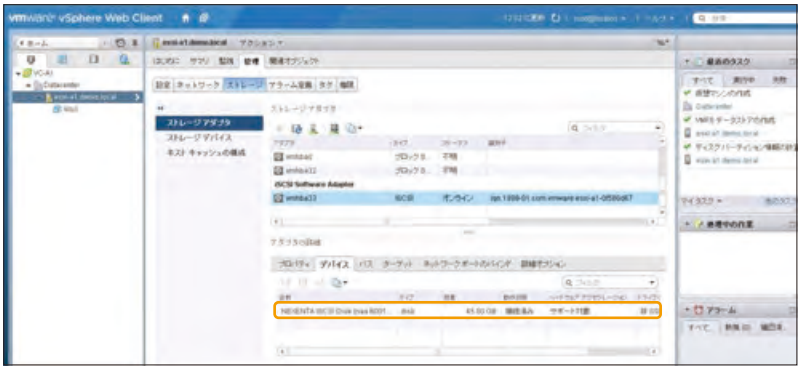


図２：ストレージの検出

次に、検出したLUNをvSphereが管理するためにデータストアとして登録します。登録画面では、図３のようにESXiサーバに接続されているストレージを参照することができ、この中からストレージを選択してデータストアとして登録します。

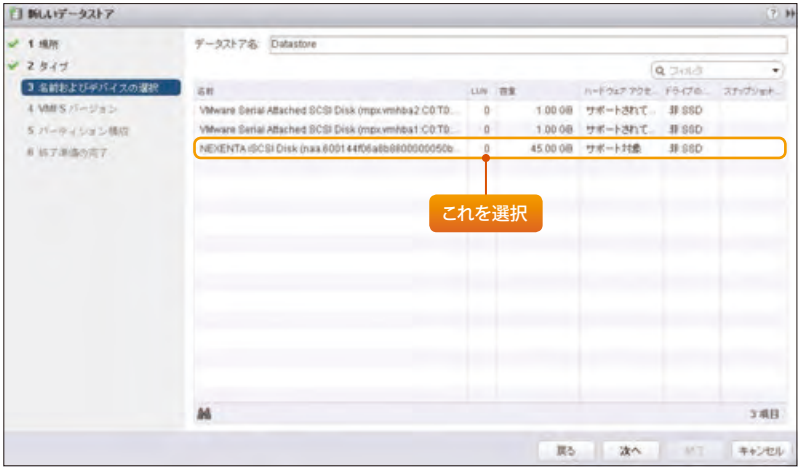


図３：データストアの追加（ストレージの選択）

また、データストアに追加する際、FCやiSCSIのブロックアクセスストレージであれば「VMFS」というファイルシステムでフォーマットします（図４）。

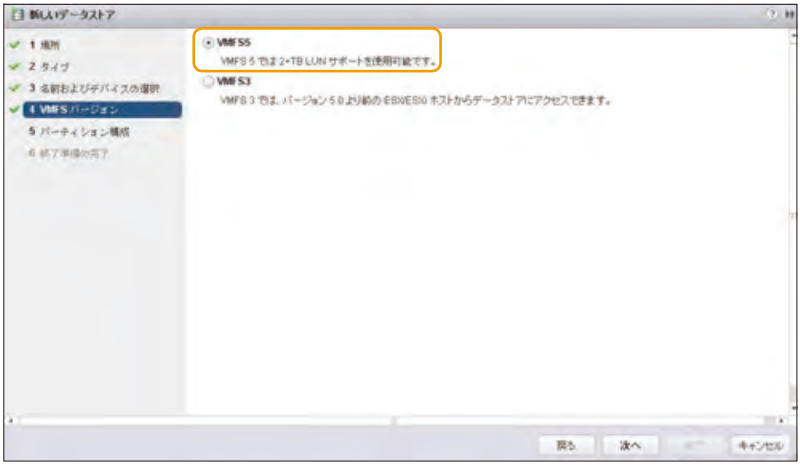


図４：データストアの追加（VMFS）

VMFSは「Virtual Machine File System」の略で、仮想マシンを収容するためにVMwareが開発した仮想環境に最適なファイルシステムです。LUN、ボリュームを設定されたストレージは、VMFSにフォーマットされることによって、データストアとしてESXiサーバ内で使用できるようになります。登録を終えるとデータストアが図５のように追加され、vSphereからデータストア＝仮想マシン等がおかれる倉庫として使用されます。

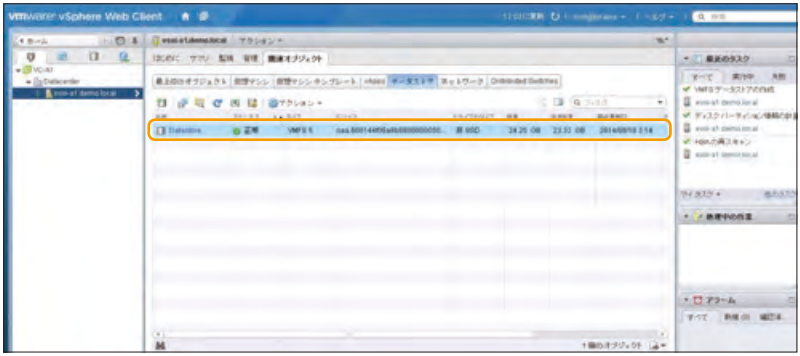


図５：データストア一覧

仮想ディスク（vmdk）～仮想マシンのディスクもファイル～

データストアとして登録が終了すると、いよいよ仮想マシンのファイルをESXiサーバから収容できるようになります。第1章で、仮想マシンの実体は複数のファイルであると解説しましたが、ここではそのファイルの1つ、「仮想ディスク=vmdk」について説明します。

仮想ディスクは「.vmdk」という拡張子のファイルとして存在し、仮想マシンにおいてローカルディスクの役割を果たします。この仮想ディスクは、ゲストOSがWindowsであればCドライブやDドライブなどとして認識され、このローカルディスクにファイルを書き込もうとしますが、実際にはESXiサーバが書き込み命令を検知して仮想ディスクファイルに内容を書き込んでいます。

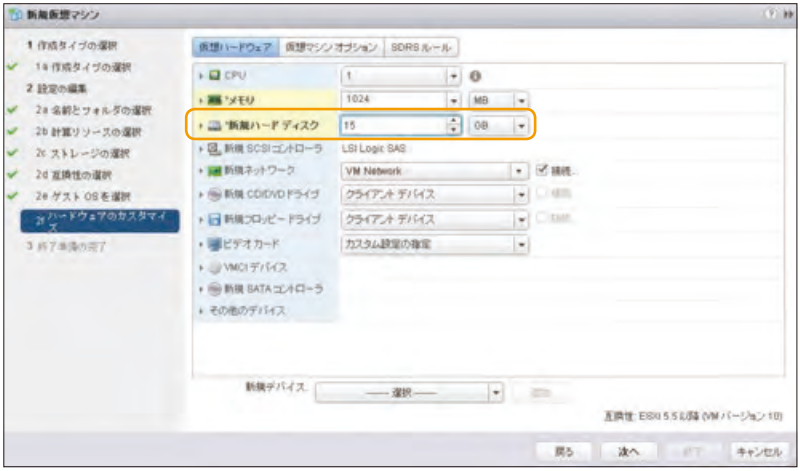


図6：仮想マシンのディスク容量の設定

図6の様に、新しい仮想マシン「kiso」を、15GBのディスクを搭載する設定で作成すると、データストアには約15GBの仮想ディスクファイル「kiso.vmdk」が作成されていることがわかります（図7）。

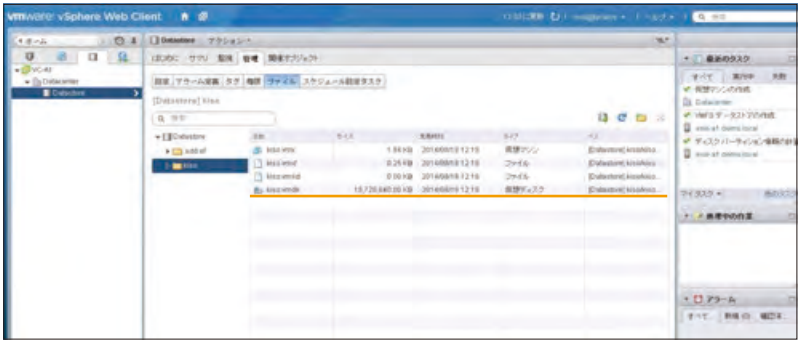


図7：データストアから見た仮想ディスクファイル

4 vMotion、HA/FTの違いとは?

ずばりvMotionとvSphere HAの違いとは!?

この章で扱うのは、vSphereの持つ機能のうち、vMotionとvSphere HA（以下HA）/vSphere FT（以下FT）です。これらの機能は少し紛らわしい部分がありますので、その違いをクリアにしていきましょう。はじめに、vMotionとHAの違いは何か、どのような時に使う機能なのか、ということから触れていきたいと思います。まずは、それぞれがどのような場合に有用な機能なのか見てみましょう。

● vMotionの使い時

例えば…

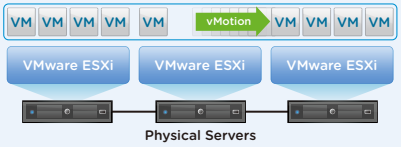
物理サーバにCPU予防交換の必要があるため一度停止したいが、そこで稼動しているサービスは平日には止めることができない。土日に出勤してメンテナンス作業を実行する必要がある。

例えば…

負荷分散の最適化のためにシステムの構成を変更したいが、日中は仮想マシンを停止できない。夜間に一度仮想マシンを停止して、別の物理サーバに移行することで適正な負荷バランスにしよう。

vMotionを用いると…

稼働中の仮想マシンを別の物理サーバに移行でき、仮想マシンで動いているシステムを止めずに、物理サーバのメンテナンスや負荷分散が可能!



● HAの使い時

例えば…

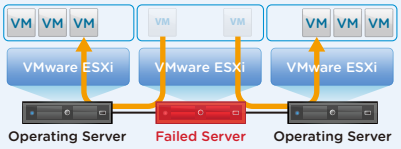
物理サーバが障害で停止してしまったため、その上で動いていたサービスも停止してしまった。早急に復旧が必要だが、データセンターまで出向いての対応には多くの時間を要する。

例えば…

仮想化はしたものの、突発的な障害に対処するため土日昼夜を問わず監視をしている。

HAを用いると…

1台の物理ホストが障害で停止したが、HAの機能によってすべての仮想マシンは別ホストで問題なく稼動しており、IT管理者は余裕を持って対応できた。



これらのケースから読み取れるように、vMotionは計画的な物理サーバの停止に対応する機能である一方、HAは非計画的な物理サーバの障害に対応して可用性を確保する機能です。したがって、vMotionは物理サーバのメンテナンスなど計画的に物理サーバを停止する必要がある場合に使用する移行機能であるのに対し、HAは機能として常に有効にしておき、いざ物理サーバに障害が起きた際に自動で保護してくれる復旧の仕組みとなります。

vMotion ～仮想マシンのホット移行～

vMotionは、図1のように、起動している仮想マシンをシャットダウンすることなく、動かしたまま別の物理サーバに移動する機能です。起動したまま移行するため「ホット移行」とも表現されます。

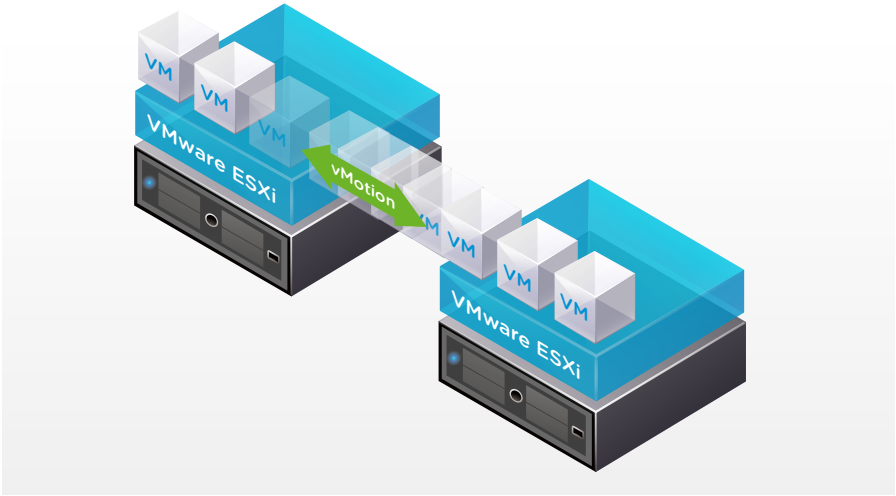


図1：vMotionによる仮想マシンのホット移行

vMotionによる仮想マシンの移行は、管理画面から仮想マシンを指定し、図２のようなウィザードに従って進めることにより、数クリックの簡単な操作で完結できます。

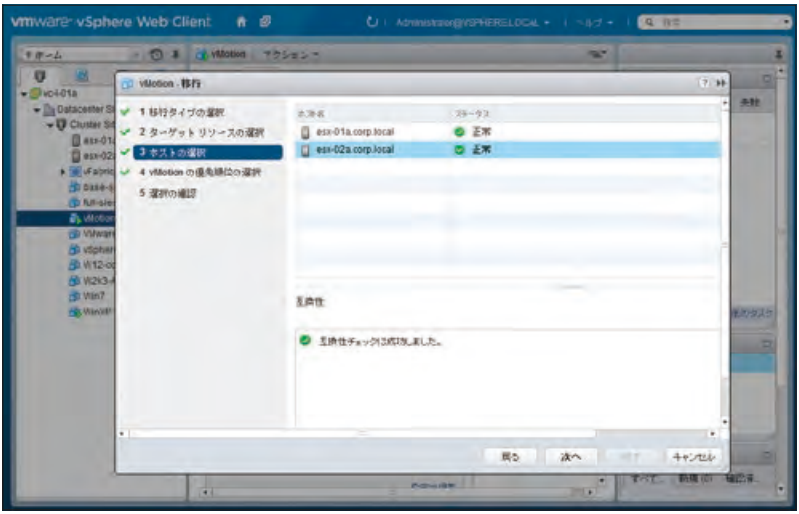


図２：vMotionによる移行は数クリックで完了

vMotionの機能は、ホストの定期メンテナンスや一部パーツの交換等で、物理サーバを計画的に停止しなければならない際に有効です。vMotionによって停止する物理サーバから別の物理サーバへ仮想マシンを退避しておくことにより、仮想マシンとして、あるいはその仮想マシンの提供しているITサービスとしてはダウンタイムがなくなります。

なお、vMotionを行うためには、対象物理サーバ（=ESXiサーバ）がvCenterに登録されていること、移行元、移行先の物理サーバのCPUに互換性があること、共有ストレージが構成されていることが必要です。CPUの互換性に関しては、同じメーカーかつ同一の互換性グループに属するファミリのもの同士でなければなりません。詳細は次のURLでご確認ください。

http://kb.vmware.com/kb/1991、http://kb.vmware.com/kb/1992

FAQ ～vMotion～

Q. 移行の前後ではMACアドレスやIPアドレスは変わりますか？

A. vMotionによる移行ではMACアドレスとIPアドレスは保持されます。仮想マシンの場合IPアドレスはvNICごとに割り当てられるため、これがvMotionによる移行前後でそれぞれ保持されることになります。

Q. 後日物理サーバを追加していくとCPUの互換性確保ができなくなりそうですが…？

A. Enhanced vMotion Compatibility（EVC）により異なるCPU世代間のvMotionが可能です。クラスタ内でEVCのベースラインを定義することにより、クラスタ内の全ての物理サーバを同一のCPU機能に統一します。詳細は次のURLをご覧ください。
http://kb.vmware.com/kb/2011037

Q. 移行先の物理サーバとの間に共有ストレージがありません。

A. vMotionとvSphere Storage vMotionという機能を同時にご利用いただくことで、共有ストレージがない物理サーバ間でも移行することが可能です。（クロスホストvMotionとも呼ばれます）

Q. 移行中に加えられた変更について整合性は保たれますか？

A. vMotionは実行中のメモリおよび全てのシステム状態を移行先の物理サーバにコピーし、移行元の仮想マシンをサスペンドして切り替えます。実行中のメモリトランザクションをコピーした後に移行先で仮想マシンを再開するため、トランザクションの整合性も保たれます。

Q. 一般的にvMotionに要する時間はどの程度ですか？

A. ネットワークの状況に依存しますが、数秒から数分程度で完了する場合が一般的です。

「クラスタ」の構成について

ここで、HA/FTの紹介を行う前に、クラスタという概念について説明いたします。なぜならHA/FTを利用するためには、クラスタの構成が必須だからです。クラスタは複数の物理サーバを論理的にグループ化したもので、まとめられたサーバはあたかも1つの大きなリソースであるかのように扱うことができます（図3）。

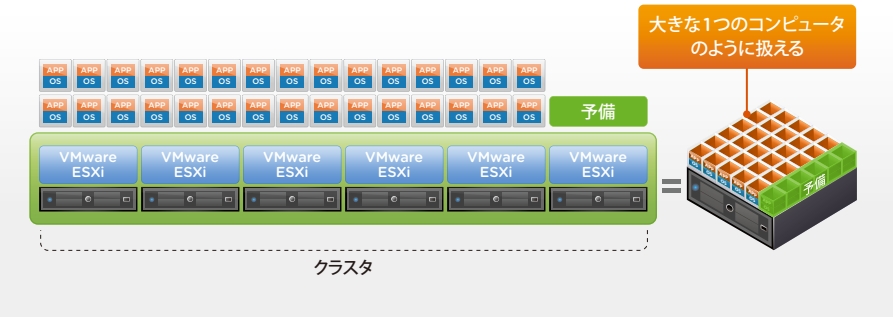


図3：クラスタ構成図

このように物理ホストをグルーピングするメリットは、それらをひと括りにして1つの大きなコンピュータのように扱うことにより、個別に稼動するよりも優れたサービス品質を提供できることです。これら複数の物理ホストは、クラスタ内で各自の持つリソースを互いに共有するため、余剰のリソース能力（CPU、Memory）を最適に配分することによって処理能力を上げたり、計画的／非計画的なホストの停止に対応する可用性の確保を実現したりします。

そのクラスタに対してHA機能を有効にすれば、クラスタ内に含まれる仮想マシンはすべてHAによって保護されることになります。また、FTの保護を施したい場合には、仮想マシンを選択してFTを有効化することによって、自動的にクラスタ内の別ホストにセカンダリが作成されます。なお、vMotionの利用にはクラスタの構成は不要です。

HA/FT ～物理サーバ障害における可用性を向上～

計画外停止（＝物理ホスト障害）に対して可用性を向上する機能がHAとFTです。HAは「High Availability」（＝高可用性）を意味し、アクティブ ― スタンバイの可用性を提供する機能です。HAを使用しない場合、ある物理サーバが障害等で機能を停止するとその上で起動している仮想マシンも停止してしまいます。それに対し、図４のようにあらかじめクラスタを構成してHAを有効にしておけば、同じクラスタ内の別の物理サーバで自動的に再起動できます。HAの場合、仮想マシンが再起動するまで数分の停止が発生しますが、仮想マシンが自動的に再起動するだけでも管理者としては助かります。

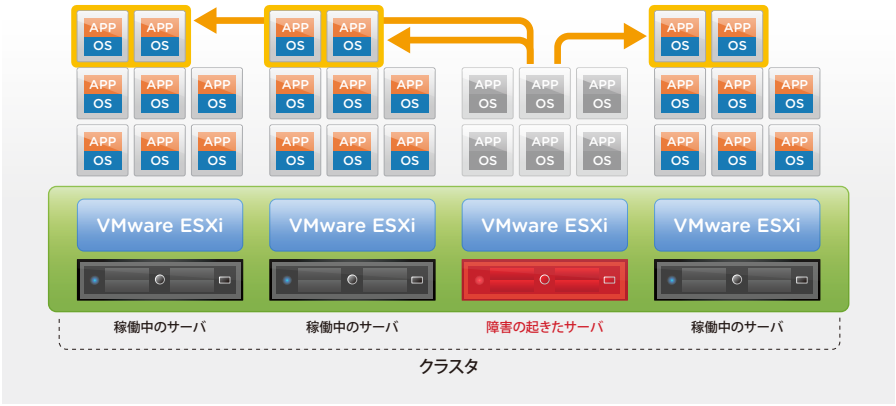


図４：HAにより仮想マシンを別ホストで再起動

FT（Fault Tolerance）は、物理サーバ障害が発生しても無停止でサービスを継続する機能です。図５のように、保護対象となる仮想マシン（プライマリ）に対し、別の物理ホスト上にセカンダリというコピーマシンを作成します。これらは常に同期しているため、仮にプライマリ仮想マシンが起動している物理サーバが停止しても、すぐに切り替わってセカンダリで動作し続けることが可能です。これにより物理サーバ障害によるダウンタイムを０にできますから、特にダウンタイムが許容されないシステムがある場合は使用をご検討ください。vSphere 5.5では、FT機能が対象にできる仮想マシンはvCPUが１つの仮想マシンに限られています。

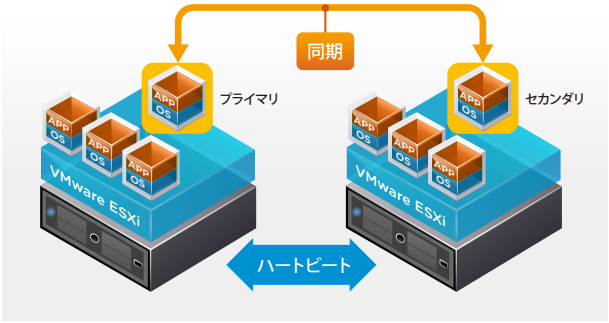


図５：FTのしくみ

FAQ ～HA/FT～

Q. HA で仮想マシンが再起動した場合、実行中だったアプリケーションはどうなりますか？

A. 仮想マシンが再起動されるため、アプリケーションは一度終了されます。Crash Consistent（＝OS が起動している状態で電源が落ちる状態）ではありますが、仮想マシンの起動とともに特定のアプリケーションが起動するよう設定しておくことで、アプリケーションやサービスの再開までを自動化することも可能です。

Q. クラスタ内に HA に必要なリソースの余裕があるか確認できますか？

A. クラスタで「許容するホスト障害数」を設定したり一定割合を予約したりすることができます。これにより常に物理サーバ障害時に必要なリソースを確保した計画的なリソース使用が可能です。

Q. HA で再起動される先の物理サーバは指定できますか？

A. アドミッションコントロールポリシーにより特定のホストをフェイルオーバーホストとして再起動する物理サーバに指定可能です。（ただし、リソースの空き具合により他のホストで再起動する可能性もあります）

Q. FT で保護されている仮想マシンのセカンダリに対して操作を行うとどうなりますか？

A. セカンダリに対する操作は行えず、プライマリに対する操作のみが反映されます。

Q. 一度物理サーバの障害に対応すると FT の保護はなくなりますか？

A. プライマリ、またはセカンダリのホストに障害が発生した場合、クラスタ内にある別の物理サーバに新たなセカンダリが生成されて保護状態が継続されます。

vMotion と HA の使い分け

これまで見てきたように、vMotion と HA は、仮想マシンを移行して別のホスト上で動かすという点では共通していますが、移行の際に起動したままか再起動するか、利用シーンが計画的な移行か非計画的な障害対応か、クラスタの構成は不要か必要か、といった違いがあります。このような違いを FT も含めて整理したのが表1です。

機能	使用目的	設定対象	仮想マシン停止	ダウンタイム	設定
vMotion	計画停止削減	仮想マシン単位	なし	ゼロ	手動で移行先を設定
HA	物理サーバ障害対策	クラスタ単位	あり	数分	自動でフェイルオーバー
FT	物理サーバ障害対策	仮想マシン単位	なし	ゼロ	自動でフェイルオーバー

表1：vMotion と HA/FT の比較

表1にあるような特徴を把握すれば、vMotion と HA/FT の違いを明確に整理できます。特に、それぞれの機能を使用するシーンや目的は全く異なるため、機能をよく理解することによって vSphere をこれまで以上に使いこなしていただけたらと思います。

5 仮想マシンの配置管理は DRS にお任せ!

仮想マシンの配置に絶大な効果を発揮する DRS

DRSというのは「Distributed Resource Scheduler」という単語の頭文字を繋げた略称です。翻訳すると「分散リソーススケジューラ」となります。「ESXi サーバの物理リソース CPU/メモリを効率的に使いましょう!」と、そんなニュアンスで解釈された方もおられるのではないでしょうか。果たしてDRSとはどんな機能なのでしょう?

解説に入る前に、前章でお話しした「クラスタ」を思い出してください。クラスタとして1つにまとめられたサーバ群は、あたかも1つの大きなコンピュータのように扱えます。この章でご紹介するDRS でも、クラスタの構成が必須となります。

では、本題に入りましょう。ここから少しの間、会社の IT 管理者になったつもりで考えてみてください。

● 状況

あなたは IT 管理者として自社の仮想基盤の整理を任されています。今、自社の仮想基盤では、図1のように10台のESXi サーバ上で100 台の仮想マシンが動いています。

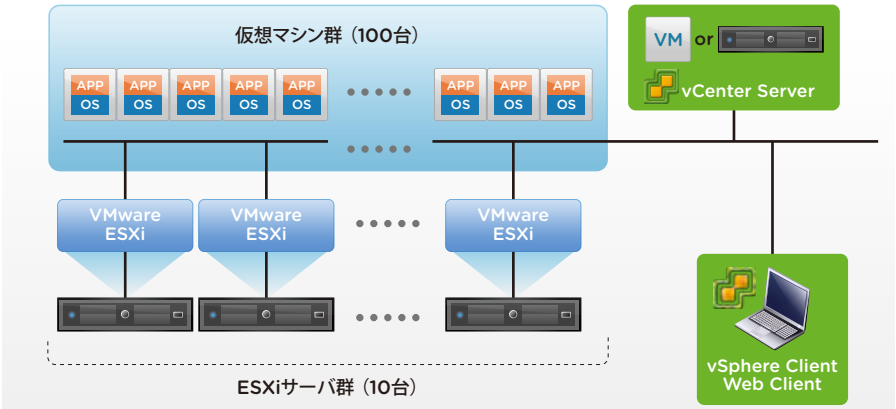


図1：自社の vSphere の環境構成図

あなたの会社がある新規サービスを立ち上げるために、仮想マシンを展開することになりましたが、自社のESXi サーバは、リソースが飽和状態であったり、時間帯によって大きく変化したり、さまざまな状況を展開しています。仮想環境は、まるで生き物のようです。

せっかく仮想基盤にしたにもかかわらず、悩ましい課題が発生してしまいそうです。こういった状況で存在感を示すのが「DRS」という機能です。先ほど、クラスタは複数のESXi サーバを1つの大きなコンピュータ（リソース）として扱える、と説明しました。管理者は「クラスタ上に仮想マシンが存在する!」と意識していますが、実際どこのESXi サーバ上に仮想マシンが配置されるかは、このDRS にお任せできてしまいます。具体的な課題と照らし合わせながら、その便利さを見ていきましょう。

● 課題1：どこのESXiサーバ上で新規の仮想マシンをパワーオンすべき?

おそらくESXiサーバ1台1台のリソースの消費具合を確認し、展開先のESXiサーバを探そうと考えたのではないでしょうか。ESXiサーバの台数が多くなればなるほど、各ESXiサーバのリソースを調べるには大変な労力と時間を消費してしまいます。見つかったとしても、すぐに負荷状況が変わる可能性もあります。困りました…。

DRSを利用すれば、仮想マシンはクラスタ内で最適なESXiサーバ上に
自動で（もしくは管理者が承認後に）展開されます。

● 課題2：ESXiサーバ間に負荷の偏りが出てきた場合

図2のように、手動で仮想マシンを他のESXiサーバに移行してESXiサーバ間の負荷の均衡をとります。移行先のESXiサーバのリソースに余裕があればよいですが、どのESXiサーバにどの仮想マシンを移行すればよいのか? 判断が難しく、困りました…。

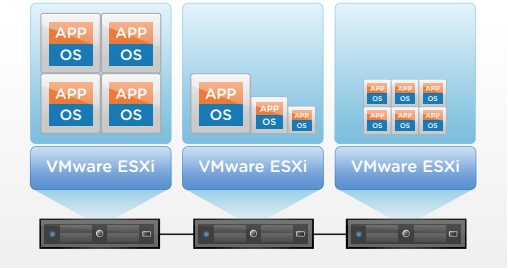


図2：ESXiサーバ間の負荷の偏り

DRSを利用すれば、
負荷の偏りが発生した時点で、
図3のように自動で
（もしくは管理者が承認後に）
適切なESXiサーバ上に移行されます。

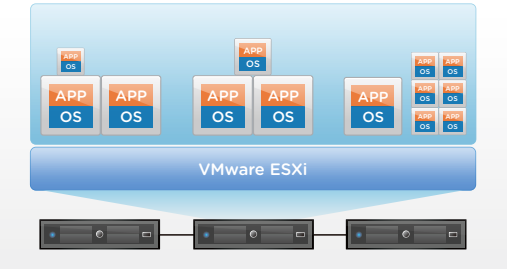


図3：DRS発動後、VMが最適に配置される

● 課題3：物理サーバのメンテナンス、ハードウェア交換、パッチ更新、メンテナンスの時期

各ESXiサーバのリソースを調べながら、手動で仮想マシンをリソースに余裕のあるESXiサーバへ移行していくのも根気のいる作業。こちらも課題2と同様、移行先にある仮想マシンに影響がでないようにしながら、どのESXiサーバにどの仮想マシンを退避したらいいのかを考えなくてはなりません。

DRSを利用すれば、物理サーバメンテナンス時も、ESXiサーバをメンテナンスモードに
することによって、仮想マシンの再配置を自動的に行ってくれます。

このようにDRSを利用すれば、仮想マシンをどのESXiサーバ上へ展開するか?といったことを考える必要はなく、単にクラスタに仮想マシンを展開するという感覚で仮想マシンを展開できます。課題1～3について考慮する必要はなくなりますね。クラスタ単位で考えれば、今まで以上に仮想基盤を有効利用できるかもしれません。

DRSの設定

ではDRSの設定を行ってみましょう。DRSとして仮想マシンの再配置が行われるタイミングは、以下の２つです。

①：仮想マシンのパワーオン時
②：クラスタ内のリソースに偏りが生じたとき

この２つを意識しながら、DRSの設定を行います。

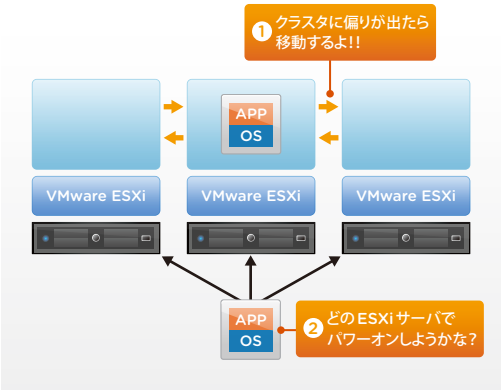


図４：DRSによって再配置が行われるタイミング

DRSの設定で特徴的なのが「自動化レベル」と「移行のしきい値」です。DRSを有効にしても仮想マシンを移行するタイミングは自分で確認したい! という方には、自動化レベルの設定が役に立ちます。

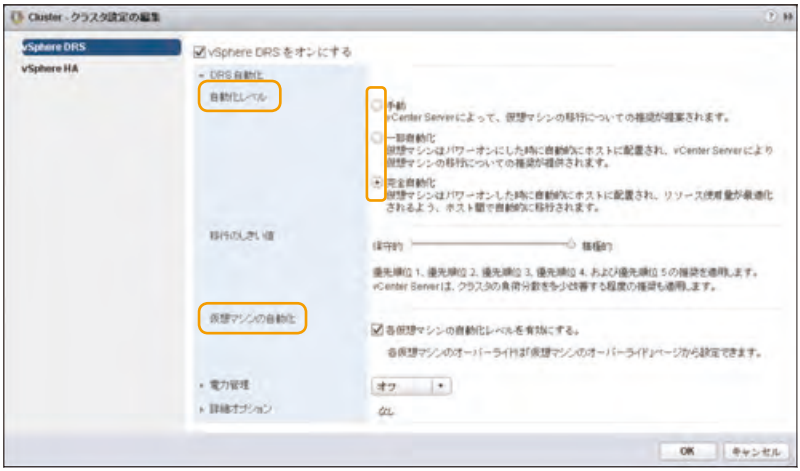


図５：DRS設定画面

● 自動化レベル

DRSには３種類の自動化レベルが提供されています。

■ 1. 完全自動化

仮想マシンをパワーオンすると、仮想マシンが最適なESXiサーバに自動で移行されます。また、DRSがクラスタ内の負荷の偏りを検出し、自動で仮想マシンの移行を行ないます。IT管理者は仮想マシンがどのESXiサーバで動いているかあまり意識しません。自動化レベルの設定ではこの完全自動化がデフォルト値となっています。

■ 2. 一部自動化

仮想マシンをパワーオンした段階は、完全自動化と同じくDRSにより仮想マシンが最適なホストに配置されます。しかし、クラスタ内のリソースに偏りが出てくると、仮想マシンの移行推奨が表示され、IT管理者が承認後、仮想マシンの再配置が行われます。

■ 3. 手動

この場合、自動的な仮想マシンの移行は行われません。つまり、仮想マシンをパワーオンすると、推奨のESXiサーバのリスト表示、またクラスタのリソースに偏りが出た場合、仮想マシンの移行を推奨する表示がされ、いずれもIT管理者の承認後仮想マシンの配置、再配置が行われます。
ではDRSが発動するタイミング「②のクラスタのリソースに偏りが出た場合」ですが、少しの偏りでも再配置をするのか、大きく偏りが出た場合に再配置をするのか、を定義するのが「移行しきい値」です。

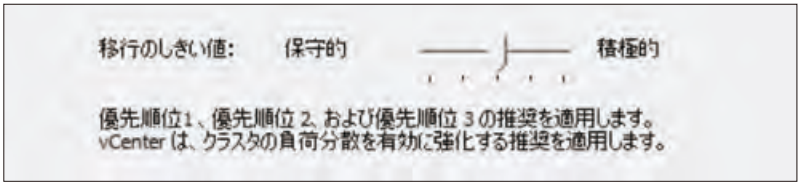


図6：移行しきい値設定画面

● 移行しきい値

クラスタ内のESXiサーバ間のリソースの偏り具合によって、移行するかしないかを決定します。決定の基準となる値のことを「移行しきい値」と呼びます。図6に示す通り、しきい値は1（保守的）～5（積極的）までの5段階があり、デフォルトは3に設定されています。しきい値1はメンテナンスモードと呼ばれ、仮想マシンの再配置はメンテナンスモードが実行された場合にのみ行われます。移行しきい値が大きくなるにつれ、少しの偏りでも仮想マシンの再配置（積極的な再配置）が行われるようになります。

● 再配置先を限定する ～ホストアフィニティ～

DRSを使用すると、仮想マシンの再配置先はクラスタ上のすべてのESXiサーバとなります。しかし、「ゲストOSで使用しているソフトウェアライセンスなどの関係上、再配置先のESXiサーバを限定したい!」というご要望もあるでしょう。

このような状況で役に立つのが、DRSのホストアフィニティという機能です。前もって仮想マシンをグルーピングしておき、その仮想マシンが動くESXiサーバを限定することにより、ソフトウェアライセンスを節約したり、仮想マシンの所在をはっきりさせたりすることが可能になります。グルーピングはDRSのみならずHAでも有効に働きます。

6 さまざまな仮想マシンが混在＆混雑しても大丈夫!?

ネットワークとストレージの帯域を維持する仕組み

この章では、vSphereなら「さまざまな仮想マシンが混在し、かつネットワークやストレージI/Oが混雑している状況であっても、各仮想マシンのサービスレベルを維持できる」ということについてお話しします。

仮想環境を最大限に生かすには、図1のようにサーバリソースをプール化し、システムごとに切り分けるというアプローチが大切です。サーバリソースをプール化することによって、特定のESXiサーバの負荷を他のサーバで補うことが可能になり、サーバ統合率を向上させることができます。管理者の皆さんは、どのサーバ上でどの仮想マシンが動いているかを気にする必要がなくなります。(詳しくは前章をご参照ください)

図1：システムごとにリソースを切り分ける

しかし、このような環境では、1つのESXiサーバ上にさまざまなシステムの仮想マシンが混在するため、各仮想マシンのサービスレベルを維持できるのかという不安を持たれる管理者も少なくないと思います。この不安はもっともなことです。

実際にDRSを適用した場合、CPUやメモリなどのサーバリソースは最適化できますが、ネットワークやストレージの利用帯域については考慮されていません。仮想マシンがどこに移動しても安心なように、CPU、メモリの他に、ネットワークやストレージの利用帯域を含めたサービスレベルを担保したいものです。

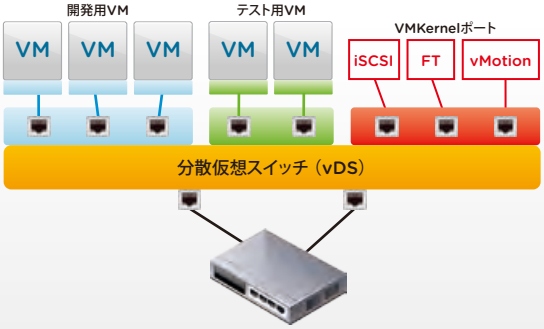
そこでこの章では、このような問題を一気に解決できる「Network I/O Control」と「Storage I/O Control」という機能についてご紹介します。これらの機能を有効にすることにより、同一のESXiサーバ上にさまざまな仮想マシンが混在していても、各仮想マシンのサービスレベルを簡単に維持することができます。また、ネットワークやストレージの帯域を効率よく利用するための機能である「LBT (Load Based Teaming)」や「Storage DRS」についてもあわせてご紹介します。

ネットワーク編 ～混在＆混雑時でも仮想マシンのトラフィックを維持する仕組み～

● Network I/O Control（NIOC）とは?

まずは、ネットワークリソースを各VMに適切に分配する仕組みである「Network I/O Control」からご紹介しましょう。Network I/O Control（以下NIOC）とは、物理NICのトラフィックが輻輳している時に、優先的に送出するトラフィックの種類を設定できる機能です。

最初に、VMware vSphereにおけるトラフィックの種類についてご説明します。vSphere環境では、ネットワーク帯域もリソースの1つとして捉え、各種トラフィックリソースがESXiサーバの帯域をみんなで仲良く使います。図2のように、ネットワークのトラフィックリソースは、FTトラフィックやvMotionトラフィックなど、事前に定義されたものがいくつかありますが、ユーザ側で特定のポートやポートグループを1つのネットワークリソースとして定義することも可能です。



事前に定義されたネットワークリソース

- FTトラフィック
- vMotionトラフィック
- マネージメントトラフィック
- NFSトラフィック
- iSCSIトラフィック
- 仮想マシントラフィック
- vSphere Replicationトラフィック

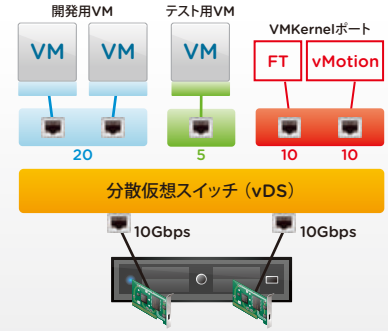
ユーザ定義のネットワークリソース（例）

- 開発用VM
- テスト用VM

図2：ネットワークリソースの定義

NIOCでは、定義されたネットワークリソースにサービスレベルを設定をすることにより、優先して帯域を利用できるトラフィックや仮想マシンを指定できます。

具体的には、図3のように各ネットワークリソースにシェア値というものを設定し、ネットワークに輻輳が起きた場合、このシェア値の割合に基づいて、ESXiサーバの帯域を割り当てるという仕組みです。



事前に定義されたネットワークリソースのシェア値

- FTトラフィック……10
- vMotionトラフィック……10
- マネージメントトラフィック……10
- NFSトラフィック……10
- iSCSIトラフィック……10
- 仮想マシントラフィック……10
- vSphere Replicationトラフィック……10

ユーザ定義のネットワークリソース（例）のシェア値

- 開発用VM……20
- テスト用VM……5

図3：ネットワークリソースのシェア値を設定

では、実際に輻輳が起きた場合、開発用 VM トラフィックにどの程度の帯域幅を割り当てられるか計算してみましょう。

図３をベースとした場合、開発用 VM のシェア値の割合は全体値（20 ＋ 5 ＋ 10 ＋ 10）分の 20、すなわち、20 ÷（20 ＋ 5 ＋ 10 ＋ 10）＝ 0.444 となります。NIC 一枚あたり 10Gbps となりますので、10Gbps × 0.444 ＝ 4.44Gbps の帯域を割り当てられることになります。図３では、ESXi サーバに NIC が２枚搭載されていますから、開発用 VM のネットワーク用に担保されている帯域は、合計で 8.88Gbps となります。

このように NIOC を利用することによって、ネットワークのサービスレベルが異なる仮想マシンが混在していても、それぞれの仮想マシンのサービスレベルを制御できます。言い換えれば、重要な仮想マシンのトラフィック（シェア値：大）が、重要ではない仮想マシンのトラフィック（シェア値：小）に影響されないよう設定できるのです。

シェア値は、ネットワークに輻輳が起きたときのみ発動されるので、輻輳が起きていない状態であれば、どのような仮想マシンも上限なく自由にネットワーク帯域を利用できます!

● LBT（Load Based Teaming：物理 NIC に基づいた負荷分散）とは?

次に、ESXi サーバ上の物理 NIC を最大限活用する機能である LBT（Load Based Teaming）についてご説明します。

同一の ESXi サーバ上で稼働する仮想マシンは、限られた物理 NIC をみんなで仲良く使わなければならないので、すべての物理 NIC をできるだけ有効活用することが重要です。

vSphere には、どの仮想マシンがどの物理 NIC を利用するかを紐付ける方式がいくつかありますが、デフォルトの設定では、仮想マシンがつながっているポートと物理 NIC が 1 対 1 で結びつきます（ポート ID ベース）。しかし、これではある仮想マシンが多くのネットワーク帯域を利用しようとした場合、同じ物理 NIC に紐付いている仮想マシンが影響を受けてしまう可能性があります。仮想マシンが利用する物理 NIC が通信相手の IP によって変わる方式（ターゲット IP ハッシュベース）もありますが、この方式でも、ある仮想マシンが同一の宛先に大量のデータを送信する場合、同じ物理 NIC を利用している仮想マシンへの影響を無視できません。

前置きが長くなりましたが、vDS という仮想スイッチ（後述）を利用している場合に限り、仮想マシンと物理 NIC に特別な紐付けを行うことができます。これこそが、今回ご紹介する LBT です。

LBT では、物理 NIC の負荷に基づいて、各仮想マシンがどの物理 NIC を利用するかを決定します。具体的には、図４のように 30 秒ごとに物理 NIC の使用率をチェックし、ある物理 NIC の使用率が 75% 以上になった場合、負荷が均等になるように仮想マシンと物理 NIC の紐付けを更新します。LBT を利用していれば、特定の仮想マシンのトラフィックが幅を利かせていても、他の仮想マシンのトラフィックが逃げ場を失うことはありません。

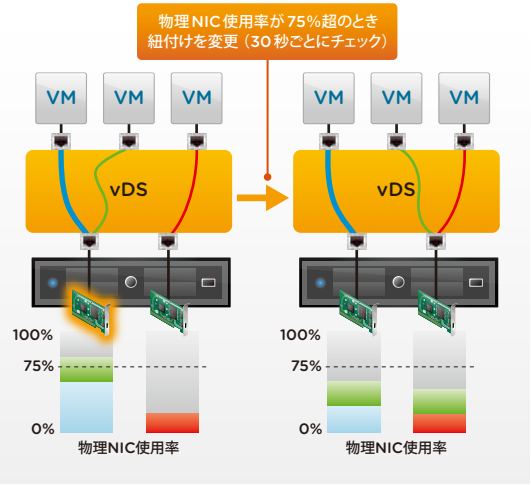


図４：LBT（物理 NIC に基づいた負荷分散）

● 分散仮想スイッチ（vDS：vSphere Distributed Switch）

最後に、NIOCやLBTを利用するために必須となる分散仮想スイッチ（vDS）について簡単にご説明します。

これまで、仮想マシンをESXiサーバ間で移行することによって、さまざまなメリット（DRS、HAなど）が得られることをご紹介してきましたが、実は、仮想マシンを他のサーバ上に移動させる際には、あらかじめ両サーバに同一の仮想スイッチを設定しておく必要があります。ESXiサーバが２台や３台ならまだ大丈夫ですが、それ以上になってくると、すべてのサーバに全く同じ仮想スイッチを設定する作業は面倒で、設定ミスのリスクも増大してしまいます。

しかし、分散仮想スイッチを利用すれば、図５のように、複数のESXiサーバに同じ仮想スイッチを一気に展開することが可能になります。もちろん設定の変更も一発でOKです!

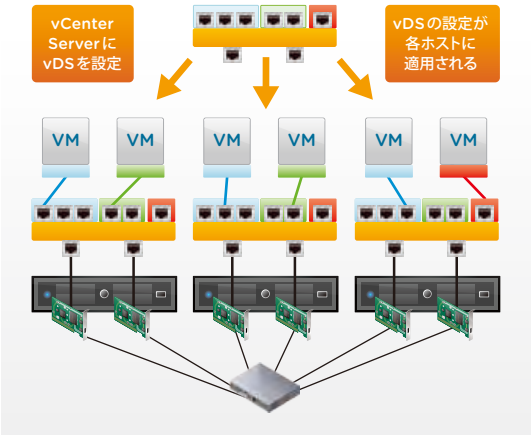


図５：分散仮想スイッチ（複数のESXiサーバに同じ仮想スイッチを一気に展開）

図６のように、この分散仮想スイッチは、論理的には「複数のESXiサーバにまたがった１つの仮想スイッチ」と捉えることができます。

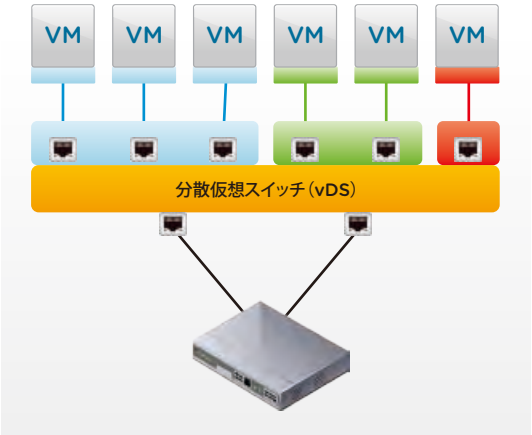


図６：分散仮想スイッチ（論理的には１つの仮想スイッチとなる）

分散仮想スイッチを利用することにより、複数のESXiサーバへのネットワーク設定が楽になるほか、さまざまな機能を利用できるようになります。今回ご紹介したNIOCやLBTは、それらの機能のほんの一部です。

分散仮想スイッチについて詳しく知りたいという方は、下記の記事をご覧ください。

- **「押さえておきたいvSphereの基本～ネットワーク編 第２回～」**
  http://blogs.vmware.com/jp-cim/2014/02/vsphere-basic-network02.html

ストレージ編 ～混在＆混雑時でも仮想マシンのストレージI/Oを維持する仕組み～

● Storage I/O Control（SIOC）とは?

それでは次に、仮想マシンがストレージを快適に利用するための仕組みについてご説明します。

Storage I/O Control（以下SIOC）とは、特定のストレージへのI/Oが集中してレイテンシが大きくなった場合、優先的にI/Oを行う仮想マシンを設定できる機能です。先ほど出てきたNIOCのストレージ版と言っても過言ではありません。ストレージI/Oを「シェア値に基づいて各仮想マシンに割り当てる」という考え方も同じです。

ただし、ネットワークとは異なり、ストレージには複数のESXiサーバからアクセスがあるため、SIOCではストレージを利用しているサーバ間でシェア値を共有する必要があります。図7をご覧ください。

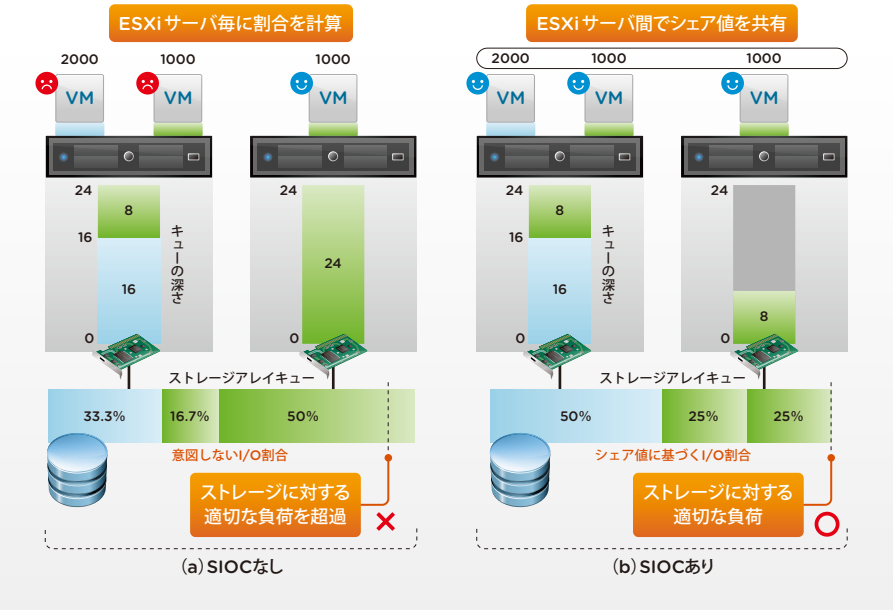


図7：SIOC（Storage I/O Control）

実は図7（a）のように、SIOCを使わなくても、単体のESXiサーバの中だけであればI/Oを優先する仮想マシンを指定できます。しかし、この仕組みは他のESXiサーバからのストレージI/Oを意識していないので、他のESXiサーバに存在する優先度の低い仮想マシンにストレージ帯域を奪われてしまう可能性があります。ストレージ側から見れば、管理者が意図しないI/O割合になるのは明らかです。

そこでSIOCは、図7（b）のように、特定のストレージを利用している仮想マシンのシェア値をESXiサーバ間で共有してから、各VMのシェア値割合を計算します。これによって、重要な仮想マシンのI/Oが重要でない仮想マシンに影響されないように、サービスレベルを担保することができます。

ただし、SIOCを利用して仮想マシンのストレージサービスレベルが維持できていたとしても、特定のストレージの高負荷状況が長く続くのは良くありません。この場合には、次に説明するStorage DRSが有効に働きます!

● Storage DRSとは?

仮想マシンの実体が共有ストレージ上のファイルであることを、第３章でご説明しました。仮想マシンの台数が増えてくると、当然ストレージへのI/O要求が増加するため、ストレージ間でのI/O負荷の分散が重要になります。そのためインフラ管理者は、仮想マシンを展開する際に、各データストアの空き容量や予想されるI/O量などを確認し、適切な配置先を選択する必要がありました。

Storage DRSは、図８のように、この煩わしい仮想マシンの初期配置を自動で行ってくれます。さらに、特定のデータストアへのI/O負荷が慢性的に高くなっている場合には、そのデータストア上に配置されている仮想マシンを他のストレージへ自動的に移すことでI/O負荷を分散してくれます。仮想マシンのデータストアを移行する際には、Storage vMotionが使われるので、仮想マシンが停止する心配はありません。

仮想マシンのデータストア初期配置やストレージI/O負荷分散は、管理者が「データストアクラスタ」として定義したプール化されているストレージに対して行われます。(実際には、データストアクラスタに対してStorage DRSを有効にするという形になります)

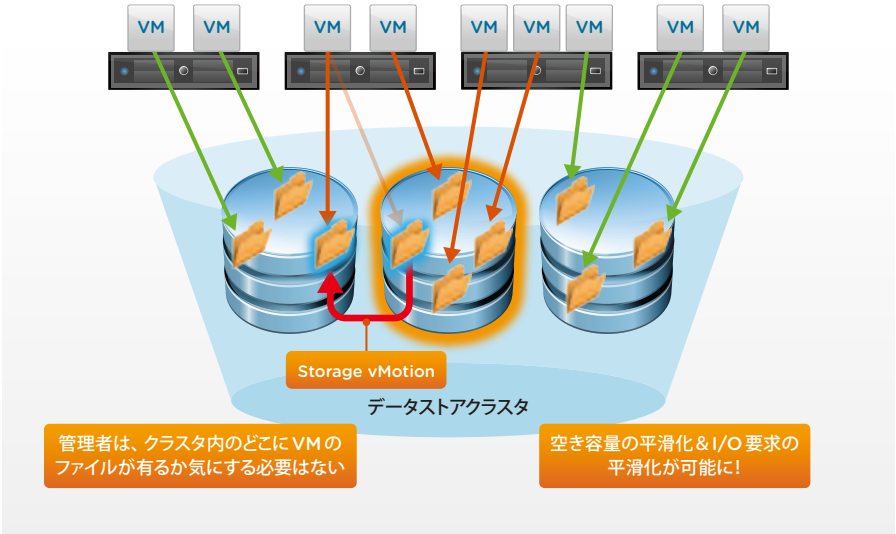


図８：Storage DRSによるストレージI/O負荷分散

ESXiサーバをクラスタ化した場合にはDRSという便利な機能を利用できましたが、データストアもクラスタ化することによってStorage DRSという便利な機能を利用できるようになるのですね。

7 vSphereでココまでできる! データ保護

vSphere ReplicationとvSphere Data Protection

VMware vSphereのデータ保護機能には「vSphere Replication（VR）」と「vSphere Data Protection（VDP）」の２つがあります。どちらも仮想マシンおよび仮想マシン内のデータの保護に利用できる機能なのですが、この章では「用途に応じて使用していただく」ための注意点についてご説明します。

「どちらもデータ保護のためのソリューションであれば、２つも必要無いのでは?」と思われるかも知れませんが、vSphere Replicationには「サイト切り替え時における仮想マシンのすばやい復旧」、vSphere Data Protectionには「仮想マシンのバックアップデータの保存」という主要目的があり、その目的に応じた違いがあります。図1は、その違いをわかりやすく図案化したものです。

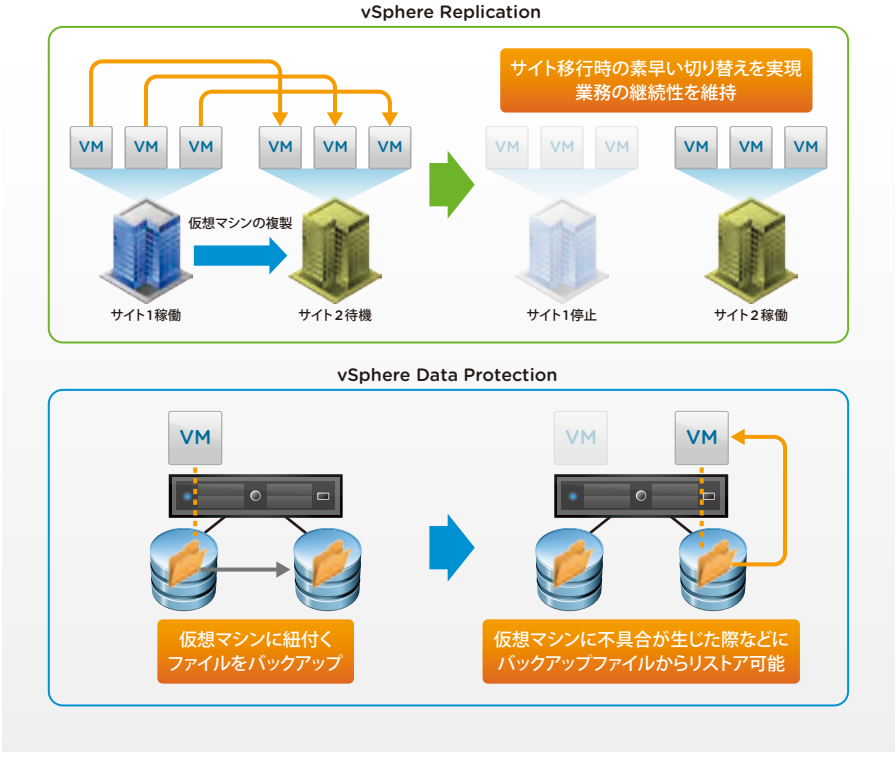


図1：VRとVDP

vSphere Replicationとは?

● vSphere Replicationによる仮想マシンの保護

vSphere Replicationは、vSphereに組み込まれたレプリケーション（複製）の仕組みで、仮想マシンにサイトレベルの障害に対する可用性を提供します。全体的な構成イメージは、図2のようになります。

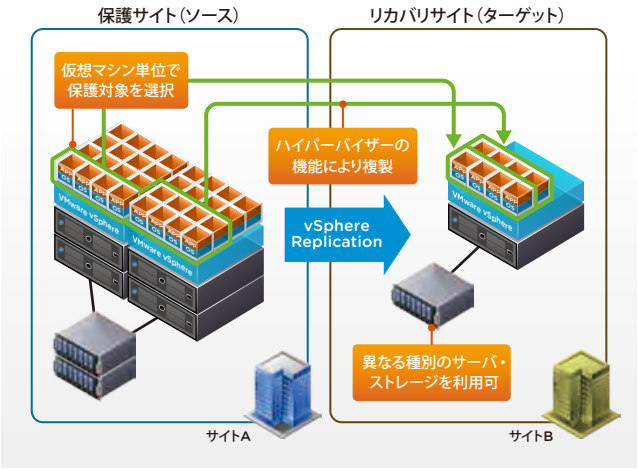


図2：vSphere Replicationによる仮想マシンの保護

図2で示した環境では、保護サイトにある仮想マシンのうち、緑色にハイライトされたものがビジネス継続性の観点から特に重要で、非常時には早期の復旧が必要とされる仮想マシン群です。vSphere Replicationの保護対象に指定されており、リカバリサイトに複製されてサイト単位の障害に備えています。

vSphere Replicationの特長は、仮想マシン単位でレプリケーションが行えることと、vCenterから一元管理できることです。仮想マシンを1つの単位としてレプリケーションを行うことにより、システムの中でも早期の復旧が必要とされる重要な仮想マシンを自由に保護対象として選択できます。

また、仮想マシン単位のレプリケーションを行うことにより、「ハードウェア非依存」という仮想マシンの特長をリカバリにおいても活用できます。このため、ストレージアレイベースのレプリケーションのように、保護サイトとリカバリサイトで同等のストレージを持つ必要がなくなり、容易にディザスタリカバリ（災害対策）を行う環境を構築できるようになります。

さらに、vSphere Replicationによる仮想マシンの複製、リカバリといった管理は、図3のようにvSphere Web Clientから一元的に行うことができます。レプリケーションの管理が仮想マシンの一般的な管理と統合されているため、システム管理者はツールを使い分ける必要がなく、レプリケーションの計画、作成からリカバリまでを1つの画面で簡単に行えます。

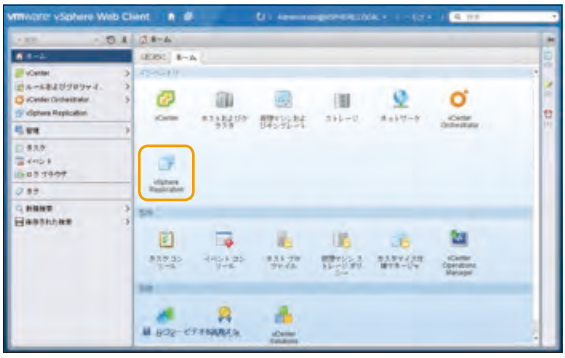


図3：vSphere Replicationはvcenterから一元的に管理

● vSphere Replicationの構成要素

まずは、vSphere Replicationの全体的な構成と、そこに登場する要素を把握していきましょう。例として、レプリケーション先が別のサイトであり、別のvCenterによって管理されている図４のような構成を取り上げます。

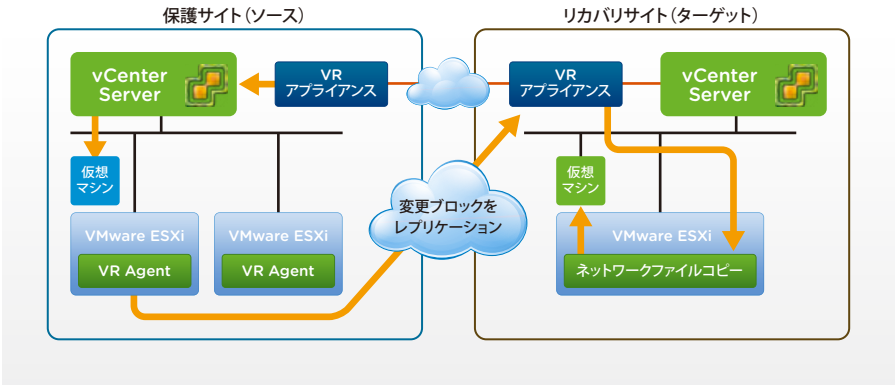


図４：vSphere Replicationの全体構成と仮想マシン複製の流れ

図４では、左の保護サイトにある仮想マシンを、右のリカバリサイトにレプリケーションしています。鍵となる以下の構成要素を覚えておきましょう。

■ vSphere Replicationアプライアンス（VRアプライアンス）

vSphere Replicationを司る仮想アプライアンスです。仮想アプライアンス管理インターフェース（VAMI）が用意されており、vSphere Replicationデータベース、ネットワーク設定、公開鍵証明書、アプライアンスのパスワード再構成といった設定は、このインターフェースから行えます。このアプライアンスはovaファイルとして提供されており、vSphere ESXiサーバ上に簡単に展開できます。

■ vSphere Replication Agent（VR Agent）

各ESXiサーバ内にインストールされ、仮想マシンの変更データをリカバリサイトのVRアプライアンスに送信します。あらかじめESXiにインストールされているため、ユーザは意識せずに使用できます。

■ ネットワークファイルコピー（NFC）

リカバリサイトのVRアプライアンスは、仮想マシンの変更データを受け取ると、問題が無いかを確認した上で対象となるESXiサーバを通じて書き込みます。この際、ネットワークファイルコピーを通じて書き込みが行われます。NFCもVR Agentと同様、あらかじめESXiにインストールされています。

●「導入→構成→リカバリ」の流れ

では実際に導入から、レプリケーションの構成、そしてリカバリまでの流れを見てみましょう。全体の流れとしては、下に示されるようにレプリケーションの構成までは３ステップ、リカバリとフェイルバックもそれぞれ簡単な操作で行えるようになっています。

● 1. VRアプライアンスの展開

VRアプライアンスを展開するとホーム画面にvSphere Replicationというアイコンが出現し、クリックするとvCenterが登録されていることがわかります。レプリケーション先が別のvCenterとなる場合は、vCenterごとにVRアプライアンスを展開します。

● 2. ターゲットサイトの登録

ターゲットサイト（リカバリサイト）のvCenterを登録します。これで、レプリケーションを行うための構成は完了です。ただし、ターゲットサイトとして同一vCenter管理下のリソースを使用したい場合は、改めて登録する必要はありません。繰り返しになりますが、別のvCenterを登録する際には、ターゲットサイト側にも事前にVRアプライアンスを展開しておく必要があります。

● 3. レプリケーションの構成

対象とする仮想マシンを選択し、レプリケーションの構成を行います。レプリケーションの構成時にはいくつかのオプション機能を設定することが可能です。オプションとしてカスタマイズできる設定には、「ゲストOSの静止（VSS対応）」「RPO」「複数時点のスナップショット」の3つがあります。

■ ゲストOSの静止

vSphere Replicationによる移行時にアプリケーションの整合性を保ち、データ損失を防ぐ仕組みで、ゲストOSが対応している場合に有効にすることができます。
対応OSは「vSphere Replication 5.5互換性マトリックス」をご覧ください。
https://www.vmware.com/support/vsphere5/doc/vsphere-replication-compat-matrix-5-5.html

■ RPO

RPOは復旧ポイントオブジェクティブを指し、リカバリ時に何時間前（あるいは何分前）の状態に戻せるようにレプリケーションを作成するか、というレプリケーションの頻度を定める指標です。図5のように、最短15分～24時間の範囲で設定可能です。

■ 複数時点のスナップショット

仮想マシンのレプリケーションは、スナップショットのように複数時点の履歴を同時に保持することができます。1日あたりの数と日数を決めて「1日3ポイント×1週間」「1日1ポイント×20日」といった設定を行うことにより、直近だけではなく一定期間前の状態にもリカバリ可能になります。

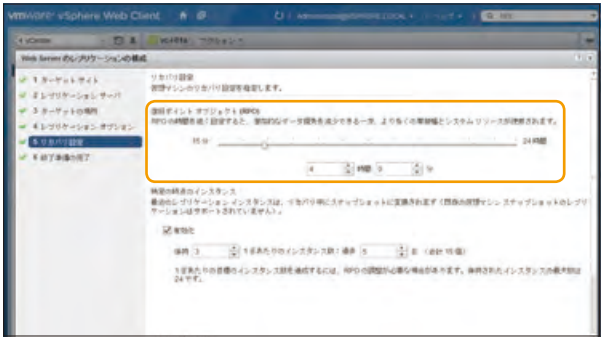


図5：RPOを15分～24時間で設定可能。複数時点の履歴を保持可能

● 4. リカバリ

いざ障害が生じて復旧が必要になった際には、リカバリを行います。「ようやくvSphere Replicationの本領発揮か!?」と意気込みたくなるところですが、操作としてはごく簡単で、数クリックで完了してしまいます。

図6に示すように、まずレプリケーション元の仮想マシンが生きているかどうかに応じて、リカバリ前に改めて同期するかを選択します。次いでリカバリ先の所属データセンターとフォルダ（選択は任意）、リカバリ先で使用するリソース（ESXiサーバ）を選択すれば完了です。

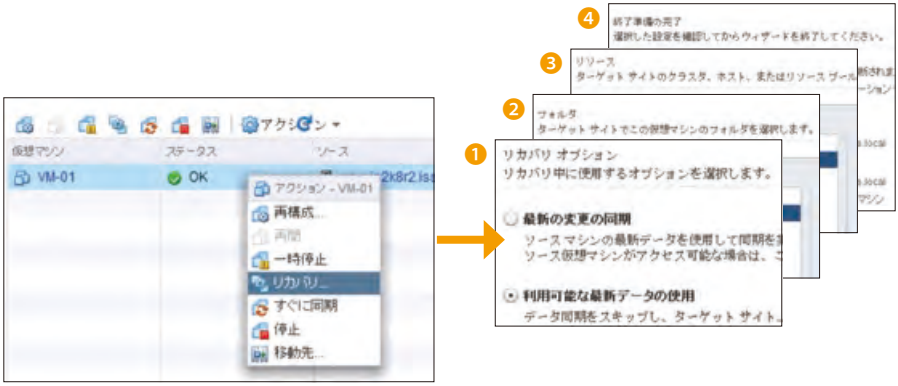


図6：リカバリは数ステップの選択で完了

● 5. フェイルバック

フェイルバックは、「リカバリサイトで一時的に稼動させていたが、元のサイトが復旧したため戻したい」という状況で必要になる作業です。リカバリ先のサイト（ターゲットサイト）から元のサイト（ソースサイト）に向けて、手動で逆方向のレプリケーション（リバースレプリケーション）を構成することにより、フェイルバックが可能です。ただし、リバースレプリケーションの構成前に、ソースサイトの該当仮想マシンはインベントリから登録解除しておく必要があります。これらはすべて手動操作となります。

ちなみに、VMware vCenter Site Recovery Managerという別の製品を使えば、操作を自動化できます!

FAQ ～vSphere Replication～

Q. ストレージアレイベースでのレプリケーションとの違いを教えてください。

A. 一言で言えばストレージの機能を用いるか、ホスト（ESXiサーバ）を用いるかの違いとなります。ハイパーバイザーベースのレプリケーションのメリットとしては、低コストでの各仮想マシンのデータ保護、ストレージベンダー選択の柔軟性、リカバリ用リソースを平常時に有効活用可能、といった点が挙げられます。

Q. 単一のvCenterで管理された環境内でもレプリケーションは可能ですか?

A. vSphere Replicationは同一のサイト内、または同一のvCenter内でも利用可能です。登録されているvCenterが1つでも、レプリケーション先のストレージに別のものを指定して耐障害性を高めるといった使用が考えられます。

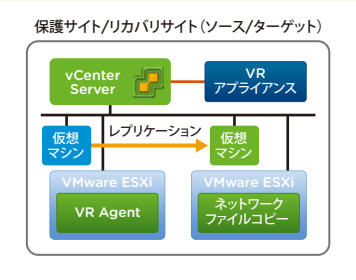


図7：単一のvCenter内でのVRの利用

Q. VDPのバックアップでは不十分なのでしょうか?

A. まず、バックアップでは同サイト内でデータのコピーが行われる構成も一般的に考えられますが、サイト単位の障害への対策という意味では別サイトへのレプリケーション（複製）が必要です。また、遠隔サイトへのバックアップとの違いとしては、保存されているデータの形式が異なります。vSphere Replicationでは、立ち上げまでの時間が短くなるよう仮想マシンごとに.vmdk形式で保存されていますが、VDPを用いたバックアップでは仮想マシンのデータ形式にリストアするまでに余計に時間がかかることが予想されますので、用途に応じて使い分けることが重要です。

Q. 定期的なデータ更新となるとネットワーク帯域をかなり消費するのでは?

A. 初回の同期時には全てのデータを転送するためそれなりに時間を要しますが、その後は変更された差分のみ（ブロック単位）を送信するため、ネットワーク帯域の消費を抑えることができます。ネットワーク帯域の要件に関してはRPOの設定にも依存するため、マニュアルを参考に加味してご検討ください。

Q. レプリケーション対象が多い場合、負荷が集中するのでは?

A. VRアプライアンスを追加で展開することにより、負荷を分散したり、レプリケーション可能な仮想マシン数を増やしたりすることが可能です。詳細は下記のURLをご参照ください。
http://kb.vmware.com/kb/2034768、http://kb.vmware.com/kb/2087771

Q. 9時～18時などと時間指定して、更新がある時間のみレプリケーションしたいのですが?

A. 残念ながらvSphere Replicationでは、指定された時間帯のみのレプリケーションには対応しておりません。しかしながら、vSphere Replicationは変更されたブロックのみを送信するため、変更が加えられていない場合のレプリケーションデータはほぼ0となり、ネットワーク等への負荷はありません。また、固定のスケジュールで縛らずRPOでデータの新しさを担保することで、例外的な操作に対しても一定したサービスレベルを維持しております。

vSphere Data Protectionとは?

● 仮想アプライアンスによるバックアップ

本書の第1章で、仮想マシンはファイルで構成されていることをご説明しました。vSphere Data Protectionもこの特徴を利用して、仮想マシンを構成するファイルをコピーすることによってバックアップを行っています。

vSphere Data Protectionは仮想アプライアンスとして、ESXiサーバ上で動作します。管理対象の仮想マシンを構成するファイルをデータストアから取得し、VDPアプライアンスの仮想ディスク、「デデュープストア」にバックアップを保管します。

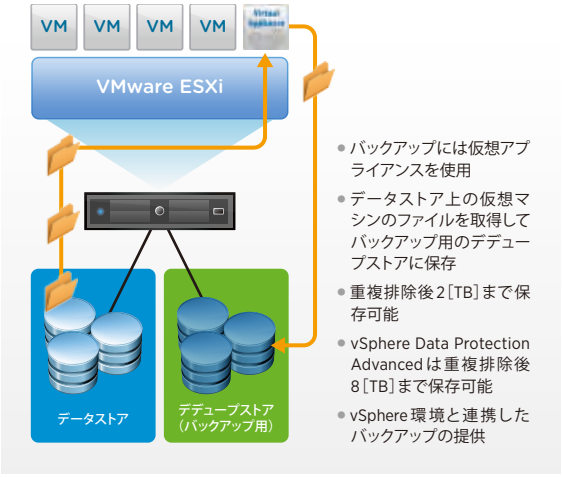


図8：VDPの仕組み

● 「導入→構成→リストア」の流れ

バックアップおよびリストアをvSphere Web Clientから行うことができるのも、vSphere Data Protectionの大きな特長です。vSphere環境にvSphere Data Protectionを導入し、バックアップ、リストアを行うまでの流れをご紹介しましょう。

● 1. VDPの導入 ～仮想アプライアンスによる簡単な展開～

vSphere Data Protectionは仮想アプライアンスとしてESXiサーバ上で稼動させます。Open Virtualization Aechive(.ova)ファイルとしてvSphere Data Protectionをダウンロードし、vSphere Web Client上で展開します。

展開が終了して設定が終わると、vSphere Web ClientにVDPプラグインが追加されます。

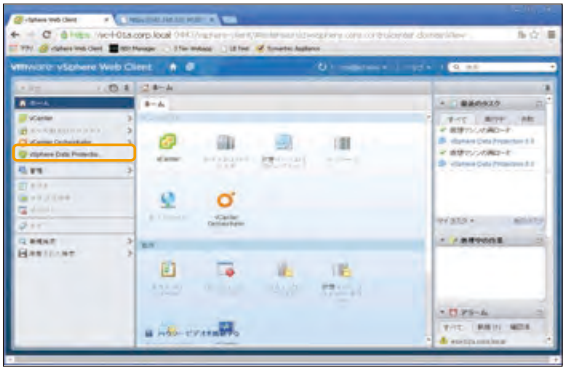


図9：VDPプラグイン

● 2. バックアップジョブの作成 ～5ステップで作成～

バックアップジョブは、図10のようにVDPプラグインから簡単に作成できます。

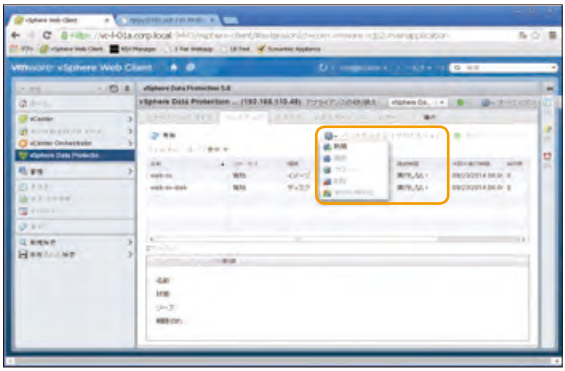


図10：バックアップジョブの作成

また、バックアップジョブの作成も図11のように、

1．バックアップするデータは仮想マシンのフルイメージか、仮想マシン内のディスクのデータか

2．どの仮想マシンに対してバックアップを実行するか

3．どのくらいの頻度でバックアップを行うか（スケジューリング）
　・毎日／週に1回／月に1回

4．バックアップしたデータの保存期間の設定
　・無期限／日、月、年単位

5．バックアップジョブの名前の決定

という5つのステップで簡単に作成できます。vSphere Data ProtectionとvSphere Replicationの大きな違いの1つは、バックアップデータを長期保存できることです。vSphere Replicationは最長で24日前のデータまでしか保存できませんが、vSphere Data Protectionは（データストアの容量が許せば）毎日取るバックアップデータを無期限に保存でき、いつでも昔のシステムに戻せます。

一方でvSphere ReplicationのRPOは最短15分前に設定できますが、vSphere Data ProtectionのRPOは最短1日となっており、「災害時になるべく最近のデータを保持したシステムを復旧させたい」といったニーズに対してはvSphere Replicationの方が適した機能を提供できます。

図11：5ステップでジョブが作成終了

● 3. データのリストア ～仮想マシンからファイルまで～

データのリストアも、vSphere Web Clientから行います。リストアするデータは仮想マシンごとに選択でき、各仮想マシンのデータはバックアップを行った時間別に並んでいるため、好きな世代のデータをリストアすることができます。また、仮想マシン内のディスク単位（vmdk単位）でリストアを行うことも可能です。

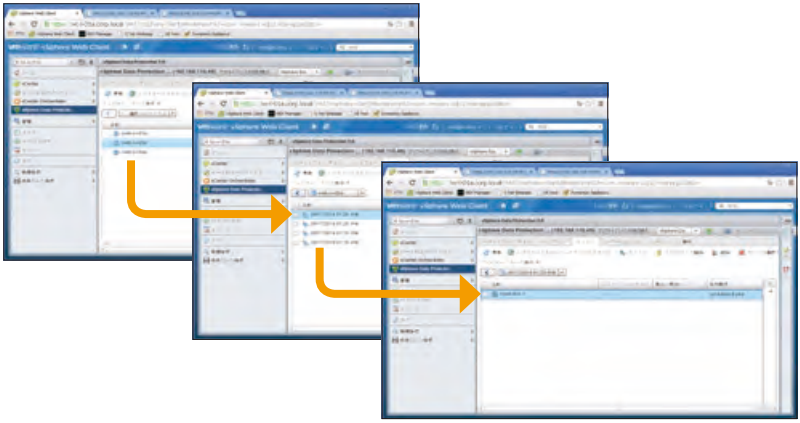


図12：仮想マシン単位のリストア

仮想マシンをリストアする際、既存の稼働中の仮想マシンにリストアするデータを上書きすることもできますが、別の仮想マシンとしてリストアすることもでき、これによって世代の異なる仮想マシンの状態を同時に確認することが可能になります。

例えば、仮想マシンに不具合が生じた場合に、どのくらい前まで仮想マシンの状態を戻せばよいかを検証できます。

また、各仮想マシンを利用しているシステム管理者は、仮想マシンのゲストOSレベルのファイル（Windowsであればレジストリやプログラムファイルなど）をリストアするファイルとして選択できます。これを、ファイルレベルのリストアと呼んでいます。ユーザは、自分の使用している仮想マシンのWebブラウザから専用のリストアクライアントにログインすることにより、必要なファイルをリストアできます。

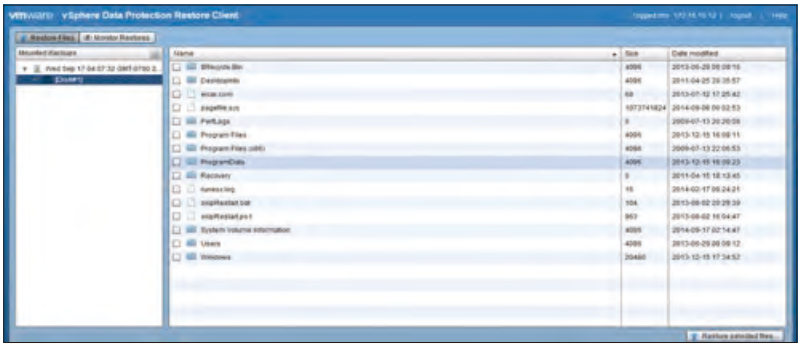


図13：ファイルレベルのリストア

● VDPの高機能バージョン vSphere Data Protection Advanced

vSphere Data Protectionは、仮想マシン同士で同じデータがあれば1つにまとめてバックアップを行う重複排除機能や、仮想マシンのデータに変更があった部分のみをバックアップする変更ブロックトラッキング機能など、仮想基盤のバックアップに必要な機能が充実しています。

さらに、VDPのアップグレード版として、遠隔地へのデータ保護、バックアップに用いるデータストアのサイズ増加、自動でバックアップ検証を行う機能などを利用できるようになるvSphere Data Protection Advanced（VDPA）もありますので、用途に応じて選択してください。

比較資料：vSphere ReplicationとvSphere Data Protectionの違い

	vSphere Replication	vSphere Data Protection
レプリケーション対象	仮想マシン	バックアップデータ
どのくらい前の状態に復旧できるか（RPO）	15分～24時間	24時間～
復旧までにかかる時間（RTO）	VMあたり3～5分間	VMあたり数分～数時間
保存期間	DR用に短期間保持 最大24世代分のレプリカ保持	長期間保持 一般的に30～180日間 （永久に保持する設定も存在）
主な利用例	個別VMの災害対策 短時間でなるべく最近の状態に 復旧したい場合	バックアップデータの長期保持
データの復旧	仮想マシン単位のリカバリ	仮想マシン単位または ファイル単位のリストア
世代管理の有無	24のレプリカを保持 複数世代のレプリカを用いた リカバリは不可	仮想マシンごとに世代管理可能 複数の世代の仮想マシンを 同時にリストア可能

8 仮想環境とながく〜
お付き合いしていくために

VMware vRealize Operations Manager（vR Ops）
（旧称：VMware vCenter Operations Manager）

環境の運用管理とは、何ができていれば上手く行えていると言えるでしょうか。状況を正確に把握し、効果的なアクションを行い、より少ないコストで効率的にパフォーマンスを高く維持できれば、良い運用が行えていると言えるのではないでしょうか。このうち「状況を正確に把握する」という部分が、物理環境から仮想環境へ移行した際に少し難しくなるポイントです。

図1：物理環境と仮想環境でのリソース使用の比較

本書でもたびたび触れておりますが、仮想環境ではリソースが仮想化され、複数のOSで同一の物理ホストのリソースを共有します。また、物理ホスト間でもリソースをプール化し、全体として最適な運用が行えるようになっています。

仮想環境では、ゲストOSから見てリソースの使用率が高くないかだけではなく、他のOSとのリソースの取り合いによるパフォーマンス低下はどれくらいか、物理サーバ間でのリソース融通による最適化の機会はないかといったことも考える必要があります。実際、すでに仮想環境をご利用いただいているお客様にも、管理に課題を感じておられる方は多く、以下のようなお声をいただいております。

公共系A機関 ご担当者様

どの物理サーバでどの仮想マシンが稼働しているかをExcelで管理していましたが、何かが起こった際の影響範囲の特定が非常に困難でした。また、一部のシステムでパフォーマンスが落ちてきた場合に、リソースが不足しているのか、アプリケーションに問題があるのかなどの切り分けが困難で、原因究明までにかなりの時間を要していました。

メディア系B社 ご担当者様

仮想マシンの配置やキャパシティプランニングはすべてExcelを使って手作業で行っており、非常に手間がかかっていました。さらに仮想マシンで必要となるリソースの割り当てについても、妥当性の基準が明確化されておらず、システムの健全性やリソース配分の効率性を評価できていませんでした。

このようなお客様には、vRealize Operations Manager（vR Ops）の導入をおすすめし、課題の解決に役立ていただきました。まずは、vCenterとvR Opsでの管理方法を比較することによって、vR Opsはどのように役に立つのか見ていきましょう。

vCenterのみでOK？

vCenterのみの場合とvR Opsも利用した場合、何がどのように変わるのか、参考ケースを見ながら比較してみましょう。

参考ケース

IT管理者のAさんは、vSphereをベースとした仮想環境の管理を任されています。担当している物理サーバの数は20台ほどですが、仮想マシンのパフォーマンスが悪い場合には、まずAさんが問題を切り分け、必要に応じてネットワークやストレージ、アプリケーションの各担当者に連絡して対処をお願いしています。
物理から仮想に移行したことによって、追加のハードウェアなしにサーバ数を増加させることができ重宝していますが、最近は仮想マシン数の増加とともに環境が複雑になり、障害原因の特定に時間がかかるようになってきました。また、来年の予算を考えるにあたって追加リソースの申請が必要ですが、仮想マシンごとに負荷も違うためどう考えれば良いかわかりません。

● 障害原因の特定

まずは障害原因の特定を行う場合を例にとり、管理方法の変化を見てみましょう。障害発生を確認してから対処するまでの流れについて、vCenter Serverのみを用いた場合と、vR Opsを用いた場合を比較します。全体の流れの一例を示したのが、図2です。

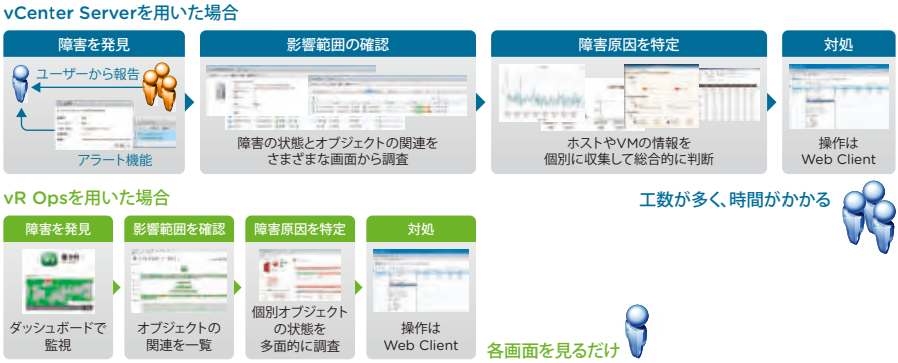


図2：障害対応方法の違い：vSphere Web ClientとvR Ops

vCenter Serverの管理画面であるvSphere Web ClientやvSphere Clientからも、ホストや仮想マシンに関するさまざまな情報を収集することは可能です。しかしながら、その情報は多岐にわたるため、広範囲な参照先から得た情報を管理者が統合して判断に結び付ける必要があります。

一方で、vR Opsを用いた場合には、障害の監視、関連オブジェクトの参照、各オブジェクトの詳細情報が一括して得られる設計となっているため、管理者の判断を助け、対処にとりかかるまでの時間も短縮できます。

● キャパシティ管理

次に、リソースを追加する際の容量計算の流れを例に、キャパシティ管理方法の違いを見てみましょう。障害原因特定のケースと同様に、vCenter Server のみを用いた場合と vR Ops を用いた場合を比較します。全体の流れの一例を示したのが、図３です。

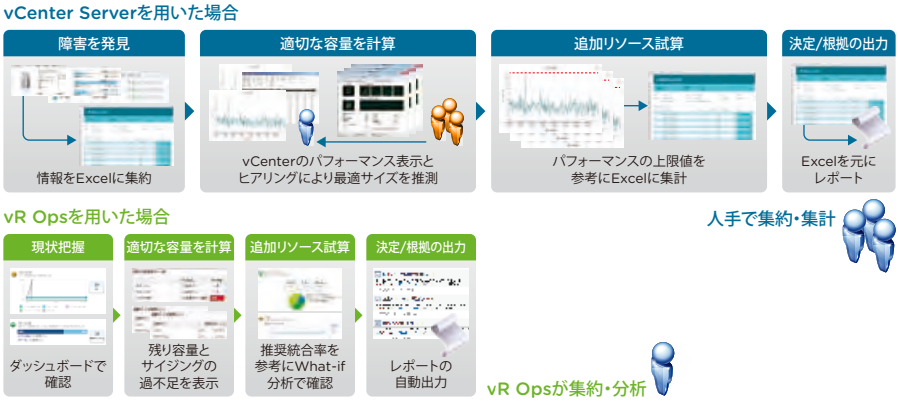


図３：キャパシティ管理方法の違い：vSphere Web Client と vR Ops

vCenter Server を用いた場合、多くの状況ではそのデータを Excel 等で集計しなおすことが多いのではないでしょうか。ホストや仮想マシンの割り当て容量を Excel に集計し、vCenter Server で得られるパフォーマンス情報とユーザーからのヒアリングを元に、適切な容量を考えます。そして、それを元にキャパシティ不足の仮想マシンや追加の仮想マシン用のリソースを計算するのです。

このような流れの問題点は、Excel での管理に時間がかかる点、また、本当に必要な容量を知ることはかなり困難で、結局のところ安全策として必要以上のリソース割り当てとなりがちな点です。

一方で vR Ops を用いた場合、現状の把握はダッシュボードから一目瞭然で、適切な容量の計算も、ホストとしては残り容量の表示が、仮想マシンとしてはサイジングの過不足に関する表示があり、クリックするだけで一覧できます。

追加リソースに関しては、推奨されるオーバーコミット率に従って考え、実際に任意のサイズの仮想マシンとリソースを追加して試算することにより、必要な容量であることを簡単に知ることができます。これらの情報はファイルとして出力し、説明の根拠にすることができます。

ポイントをまとめておきましょう。

vCenter Server を用いた管理の課題

① vCenter Server だけでは長年の経験と専門スキルを要する（工数と時間がかかる理由）

② 運用メンバーの管理手法が属人化している（人手で集約・集計している結果）

vR Ops による解決

① vR Ops 上の表示の確認で済む（工数と時間を短縮）

② vR Ops が自動で集計・分析する（工数の削減と根拠の明確化）

VMware vRealize Operations Manager（vR Ops）の概要

vRealize Operations Manager（vR Ops）を導入すると、システム構成は図４のようになります。vR Opsは仮想アプライアンスとして展開され、vCenter Serverから収集したデータを解析し、表示します。1つまたは複数のvCenter Serverを対象にすることができます。

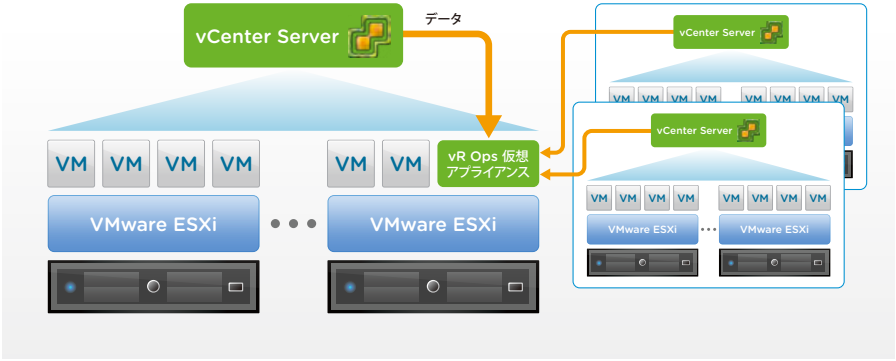


図４：vRealize Operations Managerの構成

集めたデータはvR Opsで解析されます。この解析により、各リソースの数値データは単にグラフ化されるのではなく、仮想環境は良いパフォーマンスを発揮しているのか、何かとるべきアクションはあるのか、といった管理者にとって役立つ情報として表示されます。ダッシュボードと呼ばれるvR Opsの基本画面は図５のような見た目となっています。

図５：vRealize Operations Managerダッシュボード画面

このダッシュボード画面は、左から「健全性」、「リスク」、「効率」の縦３列に分かれており、それぞれ以下のような内容を表しています。

- **健全性：「現在の状況」＝ 障害やパフォーマンス低下について**
- **リスク：「将来の予測」＝ 今後のパフォーマンス低下要因はないか、リソースは十分か**
- **効率：「構成の最適化」＝ 構成を変更することで、リソース節約の可能性はあるか**

vR Opsのダッシュボードに表示される「健全性」「リスク」「効率」の指標を「バッジ」と呼びます。このバッジによって、管理者はひと目で環境の特徴を把握することができ、必要であれば、詳細な情報を掘り下げて見ていくことも可能です。例えば、ある仮想マシンについて性能劣化の要因を調べる場合には、関連するオブジェクトを一括表示することができ、どこに原因があるか突き止める大きな助けとなります。

図６：関連するオブジェクトを一括表示

また、個別のオブジェクトに関して、それ自体の情報を詳細に見るという場合には、図７のような画面からも確認できます。

図７：個別オブジェクトの詳細情報表示

他にもvR Opsは、俯瞰的な見方から詳細に特化した見方まで、多彩な情報の表示方法を用意しており、実際の運用管理の場面でも、その時々の目的に応じた情報を得ることができるようになっています。このようなvR Opsの機能は、評価版を実際にご使用いただくことによって、さらによく確認していただけます。

評価版のインストールガイドは操作ガイドとともに、下記の記事にございますので、ぜひご活用ください。

http://blogs.vmware.com/jp-cim/2014/07/vcops_operations_guide.html

● vRealize OperationsとvSOM

「VMware vSphere with Operations Management」という製品をご存知でしょうか？vSOMという略称でも呼ばれるこの製品は、vSphereとvR Opsのスタンダードエディションが合わさったものになっています。vR Opsのライセンスは、通常「25仮想マシン数単位のライセンス」ですが、vSOMのライセンスはvSphere同様、CPU単位のライセンスとなっています。したがって、vR Opsを利用される際に「仮想マシン数が将来増加する可能性もあるなぁ…」という場合は、vSOMを購入することによって仮想マシン数を意識する必要がなくなります。

詳細は弊社ウェブページをご覧ください。

http://www.vmware.com/jp/products/vsphere-operations-management/

図８：vSOMはvSphereとvR Opsのセット製品

本書をお読みくださった皆さまへ

本書は、VMwareの４名の新卒SEが全７回で連載したブログ記事「新卒SE社員が贈るvSphereのキソ!」を印刷物として再構成したものです。「新入社員の目線」だからこそ、vSphereをわかりやすく解説できた部分もあると思いますが、いかがでしたでしょうか？ 何はともあれ、本書を通じて少しでもVMware製品への理解を深めていただけたなら、とても嬉しいです。

私どもは今年の４月に新卒１期生としてVMwareに入社し、ほぼ知識がないところからのスタートでした。VMwareに関する勉強には少なからず苦労しておりますが、わかってくると楽しく、ついつい時間が過ぎてしまうこともしばしばです。今後、初めてVMware製品を導入されるユーザ様や提案されるパートナー様も、新しい概念や用語で苦労されるかもしれません。その際は本書を読み返していただければ幸いです。

お読みいただき誠にありがとうございました。

著者：ヴイエムウェア（株）2014年新卒社員SE
氏田 裕次 / 川崎 一青 / 楠木 正博 / 野田 裕二

監修：ヴイエムウェア（株）シニアシステムズエンジニア
中村 朝之

ヴイエムウェア株式会社 〒105-0013 東京都港区浜松町1-30-5 浜松町スクエア13F www.vmware.com/jp

2014.10