



Dell OpenManageと VMwareの仮想化環境で実現する プロアクティブなメンテナンスと 電源管理

文=Balasubramanian Chandrasekaran
Puneet Dhawan

Dell OpenManage™ システム管理スイートとVMware® Infrastructure 3のクラスタリソースを組み合わせると、プロアクティブなメンテナンス体制が整い、継続的なサービス提供が可能になる。また、状況に応じて電源利用がコントロールできるため、データセンタの電気料金や冷却費の節約にも効果的だ。



クラスタを仮想化するメリットは多い。既存ワークロードのコンソリデーションはもちろんのこと、データセンタの電源/冷却コストの削減にも役立つうえ、ビジネスニーズの変化にも即応できるようになる。Dell™ PowerEdge™ サーバのクラスタ環境を仮想化し、これを運用管理するのであれば、是非、Dell OpenManageスイートとVMware Infrastructure 3をご活用いただきたい。Dell OpenManage™ とVMware® Infrastructure 3から提供される豊富なツール群を組み合わせることで、プロアクティブなメンテナンス体制が整い、その結果、サービスの連続提供能力が高まる。また、状況に応じた電源利用も可能になるため、データセンタの電気料金や冷却費の削減にも貢献するのだ。

関連分野:

Dell OpenManage

Dell PowerEdge
サーバ

電力 & 冷却対策

仮想化

VMware

全カテゴリ:

www.dell.com/power/solutions

Dell OpenManageとVMware Infrastructure 3のツール群

Dell OpenManageとVMware Infrastructure 3ソフトウェアは、「Dell PowerEdgeサーバで構築した仮想化環境をシンプルに管理する」という設計理念のもとに開発された。たとえば、Dell OpenManage Server Administrator、Dell OpenManage IT Assistant、VMware分散型リソーススケジューラ(DRS)、VMware Infrastructureソフトウェア開発キット(SDK)などは、仮想化クラスタの管理に役立つ代表的なツールだ。

Dell OpenManage

Dell OpenManageスイートには、Dell PowerEdgeサーバ用のシステム管理アプリケーションが各種含まれており、サーバのディプロイメント(導入/展開)から、監視、変更管理まで、標準ベースの相互運用ツールが総合的に取り揃う。今回、仮想化の管理に最も活躍するツールとして取り上げるのは、2つのDell OpenManageアプリケーション「OMSA」と「IT Assistant」だ。

Dell OpenManage Server Administrator (OMSA): OMSAを使うと、個々のサーバと内蔵ストレージレイが容易に管理できる。たとえば、サーバのステータスやインベントリの確認、BIOSやRAIDの構成、イベントに応じた処理(アクション)の設定、サーバの電源オン/オフなどが実行可能だ。OMSAは、VMware Infrastructure 3環境での完全な動作確認も済んでいる。

Dell OpenManage IT Assistant (ITA): 総合的な中央コンソールとして活躍する標準ベースのIT Assistantは、デルサーバおよびストレージはもちろん、テープライブラリ、ネットワークスイッチ、果てはプリンタ、クライアントシステムに至るまで一元管理が可能だ。さらに、OMSA搭載のデルサーバ上で発生したイベントやアラートを調べたり、それらのイベント/アラート内容に応じたアクションを構成したり、サーバのパフォーマンス統計情報を監視することができる。

VMware Infrastructure 3

VMware Infrastructure 3スイートからは、ITデータセンタのコンソリデーション、スムーズな管理、リソースの最適化、高可用性を達成する、エンタープライズクラスの仮想化ソフトウェアが各種提供される。数々の最先端機能もご利用可能で、iSCSI (Internet SCSI)、ネットワーク接続ストレージ(NAS)、64ビット仮想マシン(VM)、VMware High Availability (VMware HA) およびDRSをサポートするほか、今や仮想マシンでは、4-Wayの対称型マルチプロセッシングと最大16 GBのVMメモリにも対応できるようになった。VMware Infrastructure 3は、Dell PowerEdgeサーバでの完全な検証が済んでおり、これには、クアッドコアプロセッサ、インテル® Virtualization Technology、AMD™ Virtualizationなどの革新的なプロセッサ強化技術も含まれる。

Dell OpenManage のシステム監視機能は、VMware VirtualCenter に統合することができる。これは、クラスタのフォルトトレランス強化策としても、また、プロアクティブなメンテナンス対策としても有効だ。

VMware Infrastructure 3の主要コンポーネントのうち、プロアクティブなメンテナンスや、状況に応じた電源管理に役立つのがDRSとVMware Infrastructure SDKだ。DRSは、動的なリソーススケジューリングと、物理リソースの最適化を提供する。また、SDKを活用すれば、サードパーティ製のアプリケーションからESX ServerホストやVMが管理できるようになる。

VMware DRS(分散型リソーススケジューラ): VMware Infrastructure 3では、複数のESX Serverホストを疎結合し、1システムとして扱う「ESX Serverクラスタ」の概念を取り入れている。DRSはその名が示すとおり、複数の物理サーバ間に分散しているコンピューティングリソースを1つのプールにまとめ、各サーバ上のVMをスケジュールすることで、これらのVMのリソース要件を効率的に満たしていく技術だ。VMware VMotion™ テクノロジーに基づいて開発されたDRSは、次のような機能を提供する。

- それぞれのVMに応じて「最適な」クラスタホストを選び、そこにVMを配置
- ホストの追加や削除などによるクラスタコンピューティングリソースの変化に応じて、リソースの最適化と配置換えを自動的に実施
- VMのリソース要件に応じて、VMを自動的に配置換え

VMware Infrastructure 3からは、ホストを「メンテナンスモード」という状態に移行できるようになった。このモードに入ったESX Serverホスト上のVMは、すべて、他のホストに移される。これらのVMは、通常移動している残りの全ホスト間に分配されるが、DRSは、その移動状況を自動チェックし、リソースが最適に割り当てられるようにする。

VMware Infrastructureソフトウェア開発キット(SDK): 開発者は、このVMware Infrastructure SDKを活用することで、ESX ServerホストとVMを管理するSOAP(Simple Object Access Protocol)ベースのカスタムアプリケーションを作成することができる。また、システム管理者は、既存の管理アプリケーションをVMware Infrastructure 3に統合することで、VMクローンの自動作成、VMの構成、性能レポートの表示など、様々な操作が行える。

Dell OpenManageとVMware Infrastructure 3の統合

システム管理者は、Dell OpenManageスイートとVMware Infrastructure 3の多種多様なコンポーネントを統合することで、物理および仮想インフラストラクチャを総合的に管理し、タスクを自動化することができる。このプロセスを説明するため、本書では2つの例を挙げた。1つは、ハードウェア障害にプロアクティブに対応する例、もう1つは、状況に応じて電源利用を調節する例だ。

この2例を実行するスクリプトおよびプログラムファイルは、www.dell.com/downloads/global/solutions/prctv.zip から入手可能である¹。これらのスクリプトやプログラムは、VMware Infrastructure SDKを通してシステム管理とVMware VirtualCenterを統合するための枠組みとなる。システム管理者は、お使いの環境に合わせてコードを変更しても構わない。

ハードウェア障害へのプロアクティブな対応

Dell OpenManageのシステム監視機能は、VMware VirtualCenterに統合することができる。これは、クラスタのフォルトトレランス強化策としても、また、プロアクティブなメンテナンス対策としても有効だ。たとえば、あるサーバ上で一部のハードウェアが故障したとしよう。この場合、劣化したサーバ上のVMを、DRSクラスタ内で移動している他のサーバへ早急に移動させることができる。このようにプロアクティブに対処すれば、その後、この劣化サーバにもっと深刻なハードウェア障害が発生しても、VMのダウンが回避されたことになる。

以降に示す手順は、Dell OpenManageと旧バージョンのESX Server 2およびVirtualCenter 1の活用法を述べたデルホワイトペーパー²を基にした。このホワイトペーパーは、移行プール内にある各候補サーバのプロセッサ負荷に応じて、VMの移行先サーバを選択する、というアルゴリズムを解説したものだ。しかし、このような手法は複雑で、開発者は、対象サーバを選ぶための正しいアルゴリズムを組み立て、移行する個々のVMに対して最適な決定を下さなければならない。

そこで、本書では、このような概念をDRSクラスタに発展させ、完全に自動化されたVMotion機能をフル活用することにした。これなら容易に負荷分散が行え、VMのリソース要件も満たすことができる。たとえば、冗長電源装置の片方が故障し、その旨をOMSAエージェントがレポートしたとしよう。このようなハードウェア障害をIT Assistantサーバレイヤで対処する場合は、ESX Serverホストを自動的にメンテナンスモードに置き、VMを他の移動ホストに移行させるといった方法が取れる。

VMware Infrastructure 3 スイートからは、IT データセンタのコンソリデーション、スムーズな管理、リソースの最適化、高可用性を達成する、エンタープライズクラスの仮想化ソフトウェアが各種提供される。

¹ これらのスクリプトやプログラムは現状での提供となり、サポートや保証は一切提供されない。

² 2005年9月、デルエンタープライズ製品グループ発行『Implementing Fault-Tolerance Through Dell OpenManage and the VMware Software Development Kit』(Dave Jaffe, Todd Muirhead著)を参照: <http://www.dell.com/downloads/global/solutions/OM-VMware-Integration.pdf>

図1は、ハードウェア障害が起きたときに発動されるアクション手順を示したものだ。その流れは次のとおり。

1. OMSAがSNMP(簡易ネットワーク管理プロトコル)トラップをIT Assistantサーバに送信する。このサーバには、問題のサーバとアラートに関する情報が格納されている。
2. IT Assistantサーバは、システム管理者の設定どおりにトラップをフィルターにかけ、Javaプログラムを呼び出す。このとき引数として、当該サーバの名前と緊急度がプログラムに渡される。
3. Javaプログラムは、システム管理者が選択したアラートを対象に、VMware Infrastructure SDKを使ってVirtualCenterサーバに接続し、障害ESX Serverホストをメンテナンスモードに移すコマンドを実行する。
4. 障害ホスト上の全VMは、クラスタ内の他ホストへと分配される。このとき移行先を決めるのがDRSのアルゴリズムだ。

こうしてシステム管理者は、障害サーバのトラブルシューティングと保守(メンテナンス)を進めることができ、稼働VMのダウンが回避された。サーバが障害から復旧すると、今度は次のようなアクションが実行される。

1. OMSAは、通常稼働を知らせるSNMPアラート(緊急度「通常」=異常無しを示すSNMPトラップ)をITAサーバに送信する。

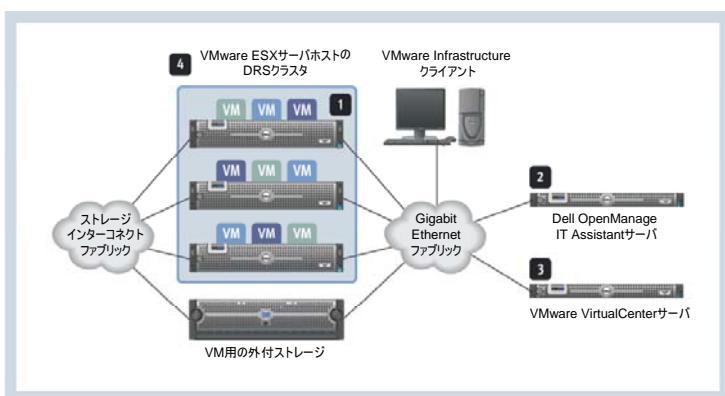


図1. ハードウェア障害へのプロアクティブな対応

電力を効率良く使い、インフラストラクチャ経費を抑えることは、効果的なデータセンタ管理の必須条件だ。

2. IT Assistantサーバは、システム管理者の設定どおりにトラップをフィルターにかけ、Javaプログラムを呼び出す。このとき引数として、当該サーバの名前と緊急度がプログラムに渡される。
3. アラートが「通常」だった場合、Javaプログラムは、サーバにSNMPクエリーを送信してグローバルヘルス情報を調べ、他のサーバサブシステムも順調に稼働していることを確認する。サーバのグローバルステータスに問題がなければ、Javaプログラムは、VMware Infrastructure SDKを使ってVirtualCenterサーバに接続し、メンテナンスモードに入っていたサーバを通常稼働に戻す。
4. するとDRSサービスは、クラスタに新しいサーバが加わったことを検出し、VMを分配し直してクラスタの負荷分散を実施する。

状況に応じた電源管理

電力を効率良く使い、インフラストラクチャ経費を抑えることは、効果的なデータセンタ管理の必須条件だ。Gartner社の調査によると、ほとんどの大企業のIT部門では、IT総予算の約5%が光熱費に割かれており、しかも、向こう5年間でこれが2倍、3倍に膨れ上がる可能性もあるという。さ

らに同調査から、ほとんどのエンタープライズデータセンタでは、機材の冷却に60%以上のエネルギーを無駄遣いしているという概算結果も出ている³。

増え続ける消費電力と冷却要件に対応せねばならない大規模データセンタにとって、これらの支出は、常に悩みの種となる。そんな現状を打破するのがサーバの仮想化だ。利用率の低い多数のサーバをコンソリデーションすれば、出費を抑えることができるからだ。

典型的なデータセンタでは、ワークロードの増減に一定のパターンがある。したがって、クラスタ内の全サーバを始終、点けたままにするのは効率的でない。そこで、処理量が減る時間帯に、少数のサーバ上へワークロードを自動コンソリデーションすれば、必要以上のサーバ数を稼働させずに済み、それにまつわる経費も削減できる。一般に、社内のデータ共有や電子メール、印刷、Webサーバといったワークロードのリソース利用率は、営業時間中にピークを迎えるが、夜間や週末は減る傾向にある。

システム管理者は、Dell OpenManageシステム管理スイートとVirtual Infrastructure 3 DRSクラスタを組み合わせることで、状況に応じた電源管理を実施することが可能だ。一例として次のような運用法が考えられる。VMware Infrastructure SDKを使ってDRSクラスタのリソース利用率を継続的に監視し、クラスタの平均利用率が事前設定したスレッシュホールド(しきい値)を下回った場合、VMを自動的に少数のサーバ上へコンソリデーションするのだ。そして、不要となったサーバは、DRAC(Dell Remote Access Controller)を通じて電源を自動的に落とせば、節電に役立つ。その後、VMのリソース利用率が増えた場合は、落としていたサーバを自動的に立ち上げ、処理量の増加に対応すれば良い(図2)。

³ 出典: 2006年10月10日、Gartner, Inc.発行『Why 'Going Green' Will Become Essential for Data Centers.』(Rakesh Kumar 著)

状況に応じた電源管理アルゴリズム: 状況に応じた電源管理の目的は、仮想インフラのワークロードをサポートする物理サーバ稼働数を最小限に抑え、クラスタ全体の平均利用率が、指定された下限を割らないようにすることである。状況に応じた電源管理アルゴリズムは、次のようなステップを踏む。

1. Javaプログラムが VirtualCenterサーバを照会して、クラスタ全体のプロセスサーバおよびメモリ利用率を測定する。
2. プロセッサやメモリの平均利用率がシステム管理者の定めた最小スレッシュドを下回る場合、サーバを1台ずつメンテナンスモードに移す。すると、DRSサービスによってVMが自動的に他サーバへ移行される。移行が完了し次第、そのサーバの電源を落とす。クラスタの利用率が最小スレッシュドを上回るまで、この処理を繰り返す。
3. 逆に、プロセッサやメモリの平均利用率がシステム管理者の定めた最大スレッシュドを超えてしまった場合は、1台ずつサーバの電源を点けていく。するとDRSサービスは、クラスタに新しいサーバが加わったことを検出するため、VMを分配し直してクラスタの負荷分散を実施する。クラスタの利用率が最大スレッシュドを下回るまで、この処理を繰り返す。

構成可能なパラメータ: システム管理者は、構成ファイル「PowerSave.xml」(www.dell.com/downloads/global/solutions/prctv.zip から入手可)を編集することで、各種の構成パラメータを調整することができる。

- **最小利用率:** cpuMin パラメータとmemMin パラメータは、それぞれクラスタのプロセッサとメモリ利用率の下限を、利用可能な合計リソースにおける割合(%)で設定することができる。クラスタの平均利用率がこれらの設定値を下回った場合、サーバの電源を自動的に切ることが可能だ。

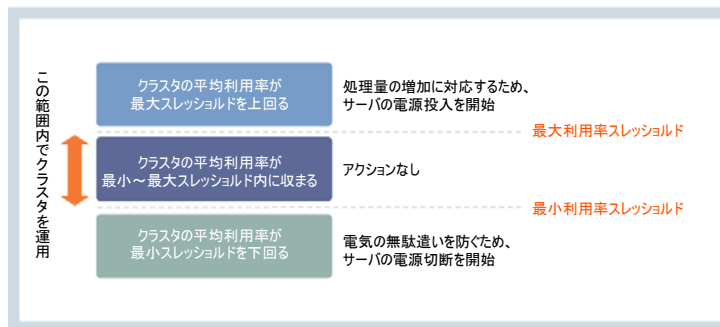


図2. 状況に応じた電源管理モデル

- **最大利用率:** cpuMax パラメータとmemMax パラメータは、それぞれクラスタのプロセッサおよびメモリ利用率の上限を、利用可能な合計リソースにおける割合(%)で設定することができる。クラスタの平均利用率がこれらの設定値を上回った場合、サーバの電源を自動的に入れることが可能だ。
- **サーバあたりアクティブにするVM数:** VMMinとVMMMaxは、物理サーバ1台につき同時実行するVM数を制限するパラメータだ。たとえば、サーバあたりの平均アクティブVM数が最小値(VMMin)を下回ったら、これ以上、サーバの電源を入れないようにし、逆に、サーバあたりの平均アクティブVM数が最大値(VMMMax)を上回ったら、これ以上、サーバの電源を落とさないようにする、といった運用法が考えられる。また、VMMinに1を設定し、VMMMaxに桁違いの数値を指定すれば、制限を実質取り除くことも可能だ。
- **タイミング:** 利用率情報のチェック頻度を指定するには、opTimeoutパラメータが利用できる。また、パラメータスレッシュドを送信してから対応アクションを発動するまでの待機時間は、steadyStateパラメータで指定可能だ。これらのパラメータを適切に設定すれば、「ワークロードが突発的に増減するたびに、サーバの電源が投入されたり切断されたりしてしまう」といった事態が避けられる。

仮想化クラスタの統合管理

本書で示したように、Dell OpenManageとVMware Infrastructure 3スイートの強力なシステム管理能力を結集すれば、ハードウェア障害へのプロアクティブな対応や、状況に応じた電源管理などの重要なクラスタ機能がスムーズに制御できる。本書に示した方法を、物理リソースのシンプルな管理と有効活用には是非お役立ていただき、データセンタの経費削減につながれば幸いです。

Balasubramanian

Chandrasekaranは、デル仮想化ソリューションエンジニアリンググループに所属するシステムエンジニア。データセンタの仮想化、高速インターコネクト、ハイパフォーマンスコンピューティング(HPC)分野に携わる。オハイオ州立大学コンピュータ科学修士課程卒。

Puneet Dhawan

は、デル仮想化ソリューションエンジニアリンググループに所属するシステムエンジニア。パンジャブ工科大学(PEC)電気工学科卒、および、テキサス農工大学コンピュータ工学科修士過程卒。

more
Online
www.dell.com/powersolutions

クイックリンク

デルとVMware社:
www.dell.com/vmware

スクリプトとプログラム:
www.dell.com/downloads/global/solutions/prctv.zip