Oracle VM - ゲストVMの高可用性環境の構築とメンテナンス

オラクル・テクニカル・ホワイトペーパー
2008年9月

Oracle Corporation 発行「Oracle VM Creating & Maintainnig a Healthy AvailableMDM as a Foundation for SOA」の翻訳版です。

ORACLE

Oracle VM - ゲストVMの高可用性環境の構築とメンテナンス

Oracle Corporation 発行「Oracle VM - Creating & Maintaining a Highly Available Environment for Guest VMs」の翻訳版です。

Oracle VM - ゲストVMの高可用性環境の構築とメンテナンス

オラクルの認定済み仮想化ソリューション: ORACLE VM
- ライセンス・コストがかからない、総合的なサーバー仮想化および管理
- Oracle VMテンプレートによるアプリケーション配置の迅速化
- オーバーヘッドが低い最新のアーキテクチャで、業界トップのコスト・パフォーマンスを実現
- セキュア・ライブ・マイグレーション、VM高可用性、その他の高度な機能の組込み

概要

Oracle VMは、基礎となるハードウェアの物理的制約からワークロードを切り離すことで、サーバー・ハードウェアのダウンタイムが計画的なものであれ、予期しないものであれ、それに伴うサービス停止の大幅な削減を可能にします。

このホワイトペーパーでは、サーバー・プールのロード・バランシングやセキュア・ライブ・マイグレーション、ゲストVM（Virtual Machine上で稼働するゲストOS）の高可用性など、ゲストVMとワークロードのアップタイムを劇的に改善するのに役立つOracle VM付属の機能について概説します。

高度なエンタープライズ仮想化プラットフォーム

Oracle VMは、エンタープライズ・アプリケーションの配置、管理、サポートを容易にする、無償で利用可能な最先端のサーバー仮想化および管理ソリューションです。Oracle環境、Oracle以外の環境の両方を対象に手頃な価格で世界的に提供されるエンタープライズ品質のサポートに支えられ、Oracle VMは完全に認定されたプラットフォームでエンタープライズ・アプリケーションの配置と運用を促進し、運用とサポートのコストを削減すると同時に、ITの効率と俊敏性を高めます。

Oracle VMは、企業のサーバー・ワークロード、つまり最も重要なアプリケーション、ミドルウェア、データベースをホスティングする目的で開発されました。このような環境では、サービスの可用性を最大化することが不可欠ですが、予算の制約もあり、複雑さを対処可能な範囲にとどめておかなければならないため、コンポーネントの1つ1つに従来の高可用性クラスタを導入することはできません。

Oracle VMは、高度なVM可用性管理を実現することで、企業を支援します。

- **ゲスト VM 高可用性機能により、計画外停止を最小化 -**
  - 組込みの VM 高可用性機能により、障害が発生したサーバーで実行中だった個々の VM または一連の VM 全体を自動的に再起動します。
  - 柔軟な配置: NFS（NAS）または OCFS2（SAN/iSCSI）とともに使用します。
- **セキュア・ライブ・マイグレーションにより、計画的なダウンタイムが原因で発生する停止を排除 -**
  - 実行中の VM を物理サーバー間で安全に移行します。
  - Oracle VM は、デフォルトで移行トラフィックの SSL 暗号化を行うことにより、機密データを不正利用から保護する、初めての主要仮想化ソリューションです。ほとんどの移行製品はネイティブ暗号化機能を備えていないため、脆弱性が生じます。また、余分な複雑さと費用をかけて SSL ハードウェアを購入しないかぎり、専用の移行ネットワークが必要になります。

Oracle Corporation 発行「Oracle VM - Creating & Maintaining a Highly Available Environment for Guest VMs」の翻訳版です。

セキュア・ライブ・マイグレーション: Oracle VMは、デフォルトで移行トラフィックのSSL暗号化を行うことにより、機密データを不正利用から保護し、専用の移行ネットワークの必要性を排除する、初めての主要仮想化ソリューションです。

- **サーバー・プールのロード・バランシングにより、VM 起動時のブロックを防止 -**
  - ゲスト VM の物理ホストは、ゲスト VM の起動時に、プールのロード・バランシングおよび可用性アルゴリズムに基づいて、正常かつ使用可能なサーバーのプールから自動的に選択されます。
  - Oracle VM のサーバー・リソース・プールおよび共有ストレージ・アーキテクチャにより、停止中のサーバーがゲストの起動をブロックしないことが保証されます。これは、予測可能なサービス・レベルを維持するのに役立ちます。
  - オプションとして、個々のゲスト VM ごとに、そのゲスト VM のホスティングに使用される名前付きサーバー（優先サーバー）の固有のリストを指定できます。これにより、パフォーマンスや可用性に関する独自のニーズへのきめ細かな対応が可能になります。
- **Oracle VM Manager の分散アーキテクチャとクラスタリング（オプション）により、管理アップタイムを最大化 -**
  - Oracle Unbreakable Linux サポートを契約している Oracle Enterprise Linux 管理サーバーは、追加のライセンス・コストをかけずに、Oracle Clusterware を使用してクラスタリングすることができます。これにより、管理サービスのフェイルオーバーとリカバリを自動化して、手動操作なしでダウンタイムを最小化することが可能になります。
  - 分散管理アーキテクチャにより、管理サーバーが一時的に使用不能になっても、セキュア・ライブ・マイグレーションやゲスト HA 自動再起動といったほとんどの VM 操作を正常に実行することができます。

アップタイムの最大化: ORACLE VMのゲストHA機能

仮想環境を持つことの最大の利点の1つは、単一サーバーが提供する限られたレベルの可用性の制約を受ける必要がないことです。従来のHAクラスタリング・ソリューションは、複数のサーバーにまたがるワークロードに対して優れた可用性を提供できますが、ライセンス契約と複雑さの点で高いコストがかかることが少なくありません。きわめて重要かつ極端な状況で継続的な可用性が必要とされる場合、こうしたコストは正当化されるのが一般的です。しかし、非常に多くのサーバーやサービスが存在し、継続的な可用性が絶対に必要であるとはおそらく言えないものの、ダウンタイムを極限まで抑えることが求められるケースは十分に考えられます。さらに、サーバーの数だけを考えても、運用、トレーニング、サポートのコストを低く抑えるためには、構成やメンテナンスが複雑でない可用性ソリューションが不可欠です。

Oracle VMのゲストVM HA機能は、Linuxカーネルに採用された初めてのクラスタ・ファイル・システムや最初のクラスタ・データベースを生み出した専門知識を活用し、ほぼすべてのゲストVMのワークロード・アップタイムを最大化する強力で管理の容易なソリューションを提供します。このソリューションは、VM内部のカスタマイズを必要とせず、VMの設定、使用、メンテナンスを簡素化します。Oracle VMのゲストHA機能には、次のような利点があります。

- 予期しない障害が発生した個々の VM をサーバー・プール内の別のサーバーで自動的に再起動します。
- 予期しない物理サーバー障害が発生したときに、すべてのゲスト VM をサーバー・プール内の別のサーバーで自動的に再起動します。
- 強力なクラスタベースのネットワーク・ハートビート・アルゴリズムとストレージ・ハートビート・アルゴリズムにより、障害が発生

したサーバーや隔離されたサーバーをサーバー・プール内ですばやく確定的に識別して、迅速かつ正確なリカバリを保証します。

- NFS、SAN、iSCSI ストレージの高度な分散ロック管理機能により、データ破損のリスクを冒すことなく、VM またはサーバー全体をすばやく確実に再起動できます。

ゲストHA機能の仕組み

Oracle VMのゲストHA機能が業界で最も高度な機能に数えられる理由を理解するために、このホワイトペーパーではソリューションの主な側面のいくつかを詳しく見ていきます。

- Oracle VM の HA 関連アーキテクチャと概念
- VM またはサーバーの障害を確実に検出する方法
- データの一貫性を確保しながら VM の再起動を行う方法
- 運用中のアップタイムの最大化を保証するために、HA 機能が製品全般にわたってどのように統合されているか

アーキテクチャの概要: Oracle VMのHA関連の概念

Oracle VMには2つの主要コンポーネントが含まれています。

- **Oracle VM Server:** Intel および AMD の x86/x86_64 を搭載したベアメタル・サーバーにインストールされ、ゲスト VM をホスティングするための環境を提供します。Oracle VM Server には、より規模の大きなオラクル開発の仮想化サーバー（Oracle VM Server）に統合された、xen.org 製のオープン・ソース・ハイパーバイザ・コンポーネントが組み込まれています。
- **Oracle VM Manager:** 多数の Oracle VM Server およびゲスト VM を一元管理するための Web ベースの管理ソリューションです。Oracle VM Manager は次のコンポーネントで構成されています。
  - Java ベースの管理サーバー。
  - Oracle Database 管理リポジトリ（Database Express Edition、Standard Edition、Enterprise Edition、Real Application Clusters のいずれか）。このデータベースは、スケーラビリティと可用性の要件に応じて、管理サーバー上に配置することも、別のサーバー上に配置することもできます。
  - Web ブラウザベースの GUI。
  - 各 Oracle VM Server 上の Oracle VM Server 管理エージェント。管理サーバーとの通信、および Oracle VM Server への管理コマンドの発行に使用されます。

重要な概念: サーバー・プールとサーバー・プール・マスター

配置の視点から見ると、複数のOracle VM Serverがグループ化されてサーバー・プールを構成しており、1つのプール内のサーバーはそれぞれ同じ共有ストレージ（NFSまたはSAN/iSCSIストレージ）にアクセスします。そのため、プールに関連付けられたVMを、プール内の使用可能かつ最も空きリソースが多い物理サーバーで起動、実行できます。また、共有ストレージへの均一なアクセスにより、VMのセキュア・ライブ・マイグレーションや自動（再）起動をプール内の任意のサーバーで行うことも可能です。

動的ロード・バランシング: VMは起動時に、ロード・バランシング・アルゴリズムに基づいて、あるいはそのVMのホストとして使用される名前付きサーバーのユーザー定義リストに基づいて、プール内の特定のサーバーに動的に関連付けられます。

VMはその起動時に、ロード・バランシング・アルゴリズムに基づいて、あるいはPreferred Serverと呼ばれる、任意のVMを起動するときに選択肢となるあらかじめ指定されたサーバリストに基づいて実際に稼動するサーバが決定されます。停止中の実行されていないVMは、共有プール・ストレージに停止状態で存在しているだけであるため、どの物理サーバーにも関連付けられていません。

このアーキテクチャにより、個々のサーバーの障害、あるいはプール内の複数のサーバーの障害によってブロックされることなく、VMの起動、停止、移行、再起動を容易に行うことができます。ただし、リソースの合計量が、プール内で同時に実行されているすべてのVMの合計必要量をサポートするのに十分であることが前提となります。

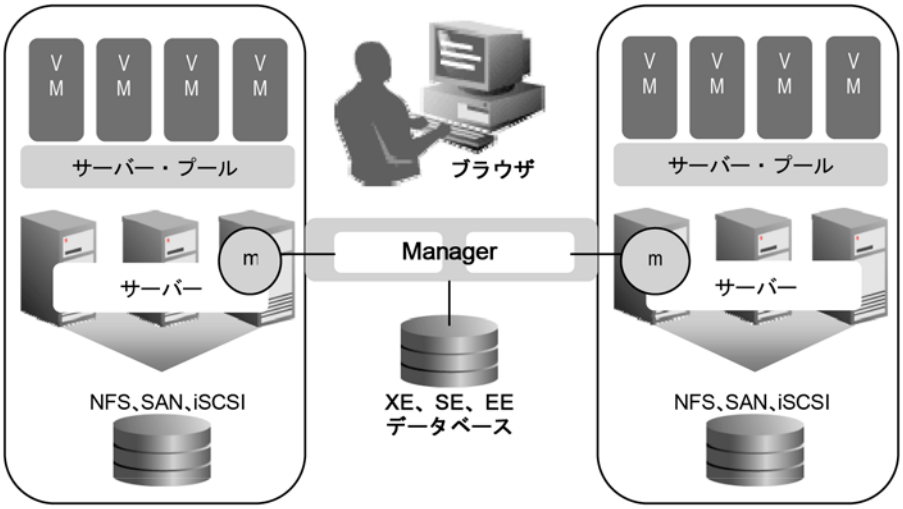

図 1. 配置アーキテクチャの概略

各サーバー・プールには、プールのアクティビティのいくつかを調整するサーバー・プール・マスター・エージェントが存在します。マスターが調整を行うのは、特に処理が複数のサーバー間での調整を必要とする場合であり、セキュア・ライブ・マイグレーション処理やゲストVM HA自動再起動処理といった重要なアクティビティが調整の対象となります。

プール内のすべての管理エージェントがマスターと直接通信し、マスターはOracle VM管理サーバーと通信します。このアーキテクチャには数多くの利点があります。その1つが大規模環境における高度な管理スケーラビリティですが、それだけではなく、機能を分散および隔離して単一障害の影響を最小化することで、Oracle VM Managerのインスタンス・レベルでの可用性を高めています。たとえば、管理サーバーが停止した場合でも、プール・マスターはセキュア・ライブ・マイグレーション処理や障害が発生したVMの自動フェイルオーバー/再起動を実行できます。同様に、1つのプールでプール・マスターが停止しても、他のプールの動作には影響を与えません。

プール・マスター・サーバーは、スケーラビリティと可用性の要件に応じて、専用のサーバーでも、ゲストVMをホスティングしているサーバーでもかまいません。最大の可用性を得るためには、プール・マスターを専用のサーバーとして配置することをお薦めします。

ゲストVM HA管理インフラストラクチャ

障害が発生したVMの予測可能で信頼性が高く、正確な再起動を保証するためには、きわめて緊密に統合されたHA管理システムによって、VMの障害検出からゲストVMの正常な再起動までのあらゆる処理を調整すること

が不可欠です。処理の途中でデータが破損する可能性がないことを保証することも同じく重要です。

OCFS（Oracle Cluster File System）、Oracle Real Application Clusters（RAC）、および関連のOracle Clusterwareに関する深い専門知識を活用し、オラクルはゲストの可用性を管理するための高度なアーキテクチャを開発しました。これは、単純なネットワークpingを利用してゲストが実行中かどうかを判別するだけのものではありません。その結果生まれたのは、より信頼性が高く、VMの障害の有無を判別する際の誤検出の可能性を大幅に低減するソリューションです。このソリューションは、共有データ破損のリスクを回避し、VMが正しく再起動されることも保証します。

インストール時にユーザーが気付くことはありませんが、Oracle VMではOracle OCFS2のクラスタスタックがOracle VMインフラストラクチャの一部としてコア製品に組み込まれています。そのため、高可用性の観点から見ると、サーバー・プールは実質的にクラスタに変換されます。これにより、VMまたはサーバーの障害の有無を判別する手段として、ごく単純なネットワークpingとタイムアウトを組み合わせて使用する競合他社のHA手法に比べ、より堅牢な信頼性の高いレベルでゲストVMのHAを管理することができます。他社のソリューションはあまりに基本的であるため、VMまたはサーバーに障害が発生したと誤って判断したり、さらに悪いことに、正常なVMを実際に停止する、あるいは必要なときにVMを再起動しないといった結果を引き起こす可能性が大きくなります。Oracle VMのアーキテクチャは、頭を悩ませるこうしたシナリオを事実上排除します。

注意: この後の項では、特に指定のないかぎり、「プール」と「クラスタ」の2つの用語は同じ意味であると考えてください。

重要な概念: クォーラムとフェンシング

先に進む前に、クラスタリングに関するいくつかの概念を理解しておくことは、Oracle VM HA/自動再起動がどのように行われるかを理解する上で有益です。

HAクラスタリングでは、すべてのアクティビティがクラスタにとって適切であることを保証する上で、クラスタ内のすべてのノードが自身のステータスとクラスタ内の他のノードのステータスを把握することが重要になります。

そのため、プール内の各Oracle VM Serverは次の機能をサポートしています。

- クラスタ・ノード管理 - プール内のすべてのサーバーの構成情報を保持します。
- ハートビート管理（ネットワーク接続とストレージ接続の両方が対象） - 後述のように、VM/サーバーの障害検出と管理を目的としています。
- 分散ロック管理（DLM） - 後述のように、共有データを同時に変更しようとした場合に発生する破損のリスクを回避し、VM の安全な再起動を保証します。

各ノードが上記の機能を備えているため、どのノードでいつ障害が発生しても、残存しているノードが調整を行い、クラスタ内のすべての仮想マシンおよびサーバーにとって安全な方法でリカバリを迅速に実行することができます。

クォーラム

クラスタが正常に稼働し、各ノードが他のすべてのノードに問題なく接続できるとき、そのクラスタにはクォーラムが存在すると言います。

これは、そのクラスタ（プール）で稼働しているすべてのサーバーが、セキュア・ライブ・マイグレーションやHA再起動などのアクティビティを許可するために必要なすべてのサーバーおよび共有ストレージに等しくアクセスできることを意味します。したがって、クラスタ内の残りのノードへのネットワーク接続やストレージ接続を失ったためにクォーラムを外れたサーバーは、プールのステータスに完全にアクセスできないので、できるだけ速やかに処理する（たとえば、適切な管理または停止ができるように、クォーラムが存在するクラスタに戻す）必要があります。クラスタ内のすべてのノードは絶えず相互にチェックインしてクォーラムが存在することを確認し、障害のためにクォーラムが失われたと考えられる場合にはその状況に対処します。

正常に稼働している有効なサーバー・プールは、完全な機能とデータの一貫性を保証するために、クォーラムを一度に1つだけ持つことができます。このことを実現する方法については、後の項で様々な障害シナリオに沿って説明します。

フェンシング

クラスタ内のノードとそのノードでホスティングされているVMがネットワーク障害などの理由でメイン・クラスタから隔離されると、リソースの最新の状態を一貫して把握することができなくなるため、それらがクラスタのリソース、特にストレージへのアクセスを試みるのをただちに防止することが重要です。したがって、隔離された状態のノードは、属していないクラスタのリソースにアクセスできないよう切り離す（フェンシングする）必要があります。Oracle Real Application Clustersをはじめとする他のHAクラスタリング・ソリューションのほとんどは、セルフ・フェンシングを実行することにより、クラスタまたはネットワーク内の他のどのノードとも通信できない場合など事前に定義された特定の状況下で、自身を自動的に停止するか、クラスタから分離します（あるいはその両方を行います）。多くの場合、これが最も実際的で信頼できるフェンシング手法です。なぜなら、そもそも前述のような状況ではノードに外部からアクセスできないため、外部的な力で確実に停止することは不可能だからです。

ゲストVMまたはホスト・サーバーの信頼性の高い障害検出

確定的なHA管理: Oracle VMは、常にVMのステータスをかぎりなく正確に把握しておくために、ネットワーク・ハートビート管理だけでなく、ストレージ・ハートビート（クォーラム・ディスクとも呼びます）も使用します。

Oracle VMは、クラスタベースの高度なハートビート管理手法を使用して、クォーラムが存在することを常に確認します。クォーラムが失われ、相互に通信できないノードがある場合は、ハートビートを使用すれば、どのノードで障害が発生しているか、あるいはどのノードがクラスタから隔離されたかを判別し、適切な処理を開始することができます。

Oracle VMは、常にVMの状態をかぎりなく正確に把握しておくために、ネットワーク・ハートビート管理を使用しますが、競合他社のソリューションとは異なり、ストレージ・ハートビート（クォーラム・ディスクとも呼びます）も使用します。クォーラム・ディスクを使用する場合、クォーラム内のすべてのノードがスケジュールに従って定期的に、共有ストレージの同じセクションに対してごく小さなステータス・データの読み書きを行います。各ノードは自身のステータスを書き込み、他のすべてのノードのステータスを読み取るので、他のノードがステータスの更新に失敗すればただちに気付きます。他のノードで発生したハートビートの問題（ネットワークまたはストレージ）を最初に検出したノードは、隔離された、あるいは障害の発生したVMをクォーラムに戻す処理を開始します。

ここで、次の2つのシナリオについて考える必要があります。

1. 正常な物理サーバーにおける個々の VM の障害
2. 複数のゲスト VM を含む物理サーバーの障害

正常な物理サーバーにおける単一ゲストVMの障害の検出とリカバリ

前述のように、Oracle VMの分散アーキテクチャでは、サーバー・プール内のアクティビティの一部を調整するプール・マスター・エージェントが各サーバー・プールに必ず含まれています。このマスターは、サーバー・プール内のすべてのサーバーと絶えず通信し、管理サーバーやサーバー・プールに代わってステータスの監視とコマンドの発行を担当します。

プール・マスターによる定期チェック時に、個々のVMが予期しない状態にある、たとえばVMのステータスがまだ「実行中」だが、そのサーバーで実行中のVMのリストに該当するVMが含まれていないと判別された場合、プール・マスターはそのVMの再起動を試みます。Oracle VMはロード・バランシング・アルゴリズムを使用してプール内の最善の（使用可能なリソースが最も多い）サーバーにVMを配置するため、VMはプール内の同じノードまたは異なるノード（あるいは、そのVMの優先サーバー・リストに指定された特定のサーバーのいずれか）で再起動されます。

ロック管理機能は、処理中に共有ストレージでデータ破損が発生しないことを保証します。VMがまだ実行中でネットワークやストレージにアクセスしている、またその他の障害により報告されたステータス自体が誤っていたなどのきわめてまれな状況でも、重複したVMインスタンスが作成されることはないため、データ破損のリスクも発生しないことを、分散ロック管理機能が保証します。

正常なサーバーにおける単一VMの障害の場合、それは特定のVMに限った単発的な障害であり、ホスト・ノード自体には影響を与えないことから、フェンシングは行われません。

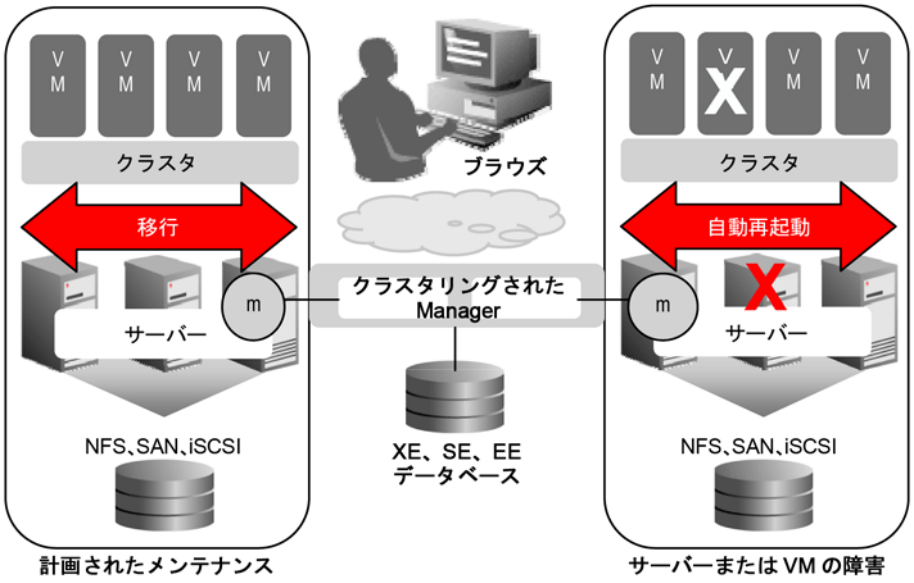

図 2. 計画的または計画外のイベントによるダウンタイムの最小化または排除

物理サーバー障害の検出とリカバリ

複数のゲストVMを持つサーバーに障害が発生した場合の検出とリカバリは、概念上は非常に似通っています。大部分のケースでは、単一のサーバーまたはネットワーク接続が停止するだけですが、最悪の場合、あるいはネットワーク構成が不適切な場合には、単一のネットワーク・デバイスの障害によって複数のサーバーがクラスタから隔離されることがあります。

すべてのノードがネットワークおよびストレージのハートビートを使用して、わずか1～2個のIPアドレスではなく、すべてのクラスタ・メンバーの

ステータスを定期的にチェックしているため、ノードの障害はクラスタ内のいずれかのノードですばやく正確に検出されます。検出したノードが処理を開始すると、それに伴って他のすべてのノードは相互に通信するよう要求されます。これにより、どのサーバーまたはネットワーク・リンクに障害が発生しているか、あるいはクラスタから隔離されているかを確認するための接続性マップが作成されます。

相互に問題なく接続できるすべてのノードがクォーラムを形成します。ネットワーク障害/隔離またはサーバー障害によってプールとの通信が完全に行えないと判断されたノードは、クラスタ全体としての正常な稼働を保護するためにクラスタから切り離されます。

このセルフ・フェンシング処理により、サーバーとそのVMは自動的に再起動しますが、起動時のロード・バランシング・アルゴリズムとロック管理に基づいて、クラスタのクォーラムに属するノードでのみ再起動が行われます。その結果、VMは、必要に応じてセキュア・ライブ・マイグレーションやさらなるHA処理を通常どおり実行できる正常なプールに戻されます。

特殊なケース: スプリットブレイン・クラスタとノード隔離

エンタープライズクラスのHAソリューションの一部として必ず考慮しなければならない興味深いケースが2つあります。スプリットブレイン・クラスタとノード隔離です。

スプリットブレイン・クラスタは、ネットワーク障害などが原因でクラスタが2つ（以上）の小規模な同一クラスタに分割された場合に発生します。残存している各クラスタは、共有ストレージなどへのアクセス権を持ち、かつアクセスを管理する制御クォーラムとして自身を宣言しようとします。Oracle VMを使用すると、新規クォーラムの形成時にこのような状況が発生していないかをチェックするアルゴリズムが個々の各ノードに組み込まれます。この場合、ストレージ・ハートビートとその他の情報に基づいて、クォーラムを形成すべきか、それともノードを再起動すべきかをノード自身が判別します。再起動を行う場合は、ノードが切り離され、VMが自動的に再起動して、ストレージのロックが解放されます。その結果、残存している他のクラスタは自身をクォーラムとして宣言して、再起動したVMのホスティングを開始できるようになります。

ノード隔離とは、他の点では正常なサーバーが稼働を継続し、ストレージI/Oを実行しているときに、おそらくネットワーク障害が原因で、あるノードがクラスタ内の他のノードから完全に隔離され、そのノードと他のノードが相互にまったく通信できない状況のことを言います。問題のノードは現在正常に稼働しているかもしれませんが、いかなる期間であれ、この状況が続くことは一般に望ましくありません。その間、ノードはステータスを提供できず、積極的な管理も行えないのに加えて、最終的に予期しない障害が発生した場合にVMのセキュア・ライブ・マイグレーションや自動再起動に参加できないためです。

また、このケースでは、前述のような状況を検出してサーバーとそのサーバーでホスティングされているVMの再起動を開始するアルゴリズムが各ノードに組み込まれています。停止処理が安全に終了すると、ストレージのロックが解放され、クォーラム内のVMが自動的に再起動されます。VMの再起動が自動的に行われるのは、Oracle VM Managerのプール・マスターが、そのノードのVMが実行中であると予想し、クォーラム内のVMが実行中でない場合は起動処理を開始するためです。隔離されたノードでディスクの停止処理が行われているためにディスクにロックが存在し、起動処理を実行できない場合、Oracle VMはしばらく待ってから再試行します。ロックが存在しない場合、VMはロード・バランシング・アルゴリズムに基づいてプール内で正常に起動します。

この処理は確定的です。また、他の製品とは異なり、Oracle VMでは、隔離されたノードのステータスに関する推測や仮定に頼ることはありません。推測や仮定を根拠にした場合、隔離されたノードの停止後にクォーラム内のVMが再起動されないといった問題が発生する可能性があります。

セキュア・ライブ・マイグレーションによるサービス停止の排除

Oracle VMのゲストVM HA機能を使用すると、予期しない障害が発生した場合でも、無理のないコストで複雑さを低く抑えながらサービスのアップタイムを最大化できます。では、ダウンタイムの発生が予測可能な場合の停止の完全な排除についてはどうでしょうか。

Oracle VMのセキュア・ライブ・マイグレーションを使用すると、実行中のVMを物理サーバー間ですばやく容易に移行できるため、計画的なダウンタイムに伴う停止が排除されます。サービスを停止する必要はなく、大規模なハードウェアへの移行後に様々な処理がどの程度高速化したかに気付く場合を除けば、サービスのユーザーは変化に気付きさえしません。

現在、ほとんどの移行製品にはセキュリティの欠如という問題があります。サーバー間の移行トラフィックが暗号化されないため、移行時には、口座番号やパスワードといった機密データのすべてが、メモリーに存在する場合と同様に暗号化されない状態で通信回線を介して転送されます。

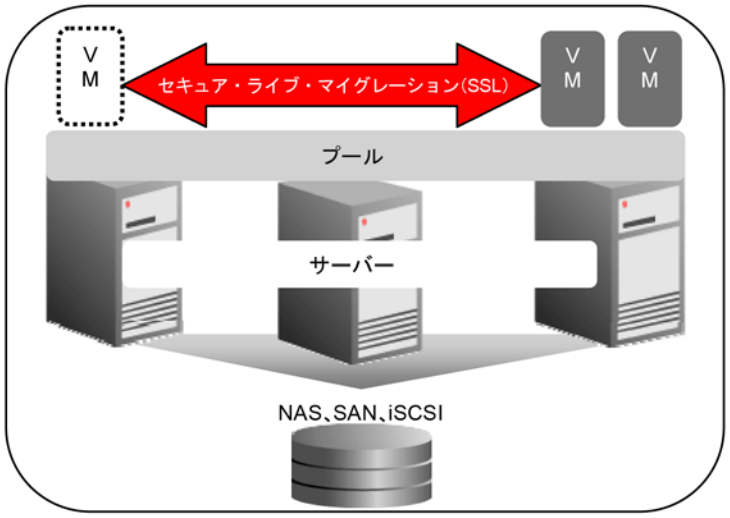

図 3. 移行トラフィックの SSL 暗号化による脆弱性の解消

Oracle VMは、重要なワークロードを実行している本稼働中の企業を対象に設計されています。セキュア・ライブ・マイグレーションでは、ハードウェアを追加しなくてもデフォルトでネイティブSSL暗号化が使用されるため、現在販売されている主要他社製品のほぼすべてに存在する脆弱性が解消します。このトラフィック暗号化機能により、機密データの不正利用を心配せずに、共有ネットワークを移行トラフィックに使用することも可能になります。

予測可能な起動と可用性: プール・ロード・バランシング

前述のように、Oracle VMの配置アーキテクチャはサーバー・プールを利用しており、プール内のすべてのサーバーはストレージへの共有アクセスが可能です。ゲストVMは共有ストレージに格納され、次のいずれかの方法でサーバーの1つに配置されます。

- Oracle VM のプール・ロード・バランシング・アルゴリズムを自動的に使用して、最も多くのリソースが利用可能なホスト・サーバーを自動的に選択する。
- ユーザー指定の優先サーバー・リストから選択する。このリストには、VM の起動と実行が可能な名前付きサーバーのサブ・プールをユーザーが指定できます。

VMは、優先サーバー・リストによって明示的に指定しないかぎり、プール内のどの物理サーバーにもバインドされないため、個々のサーバーがメンテナンスの目的で偶然停止している、あるいは他の理由でそのとき使用できないというだけで、VMの起動が妨げられることはありません。さらにロード・バランシング・アルゴリズムにより、VMは使用可能なリソースが最も多いサーバーに必ず配置されるので、プールの合計パフォーマンスを確実に最大化することができます。

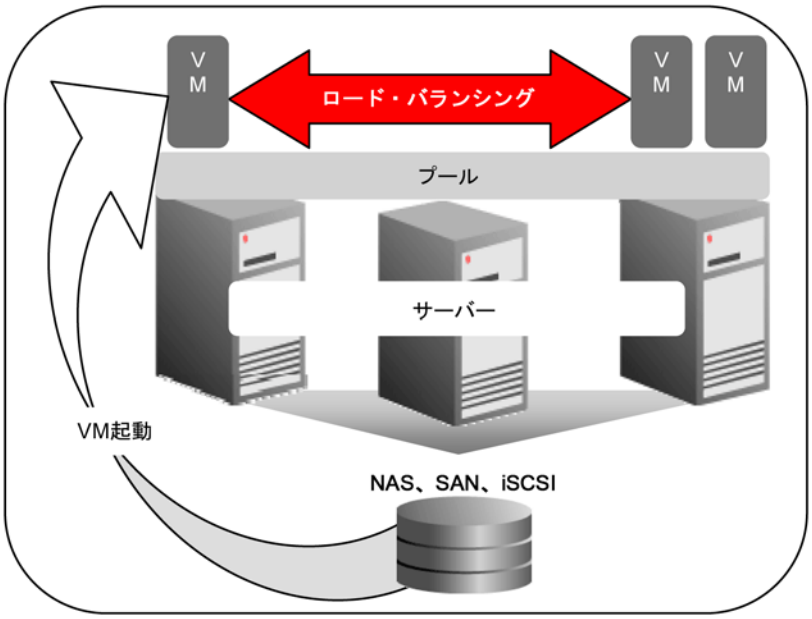

図 4. 停止中のサーバーの回避 -
VM 起動時の自動プール・ロード・バランシング

パフォーマンスやスケーラビリティをきめ細かく調整するために、あるいは単にライセンス契約や組織内の問題に基づいて、プール内のどのサーバーをホストとして使用するかを指定したい場合は、優先サーバー・リストを定義することができます。優先サーバー・リストは可用性を微調整する目的でも利用可能できます。たとえば、優先サーバーのサブ・プールを使用して、同じアプリケーション・スタック内の複数のコンポーネントが同じ物理サーバーに同時に存在しないようにすることができます。

優先サーバー・リストはVM単位で作成されるので、事実上、個々のVMごとにサブ・プールを作成することが可能です。このリストは、VMをどこで起動、実行、再起動（HA）できるか、どのサーバーをセキュア・ライブ・マイグレーションに使用できるかを決定します。そのため、環境を微調整が可能であることは明らかですが、最大の可用性が得られるのは一般に、VMがプール内の最大数のノードにアクセスできる場合です。

セキュア・マイグレーションおよびサーバー再起動とのHA統合

Oracle VMのゲストVM HA機能は、計画的なメンテナンスや予定されているその他のアクティビティの間も高可用性を容易に維持できるように、製品全体でも統合されています。

- 物理サーバーが Oracle VM Manager から停止するよう指示された場合、ユーザーはホスティングされている VM のいずれかまたはすべてのセキュア・ライブ・マイグレーションを停止前に実行するよう選択できます。これにより、VM の可用性を容易に維持できます。
- サーバー停止時に、移行の対象ではないが HA オプションがまだ有効になっている VM は、物理サーバーが停止されるとプール内の他のサーバー上で自動的に再起動します。そのため、サーバー停止時に管理者が戻って各 VM を手動で再起動する必要はありません。
- HA 機能は、メンテナンス期間中の意図しない再起動を防止するために、プール・レベル（プール内のすべての VM が対象）または個々の VM レベルで簡単に無効化したり、再度有効化したりすることができます。

Oracle VMテンプレートを使用して、高度なエンタープライズ・ソフトウェアを含む事前構築済みのVMを迅速に配置する方法については、Oracle VMテンプレートのWebサイトを参照してください。

高可用性管理インフラストラクチャ

管理操作は1日24時間途絶えることがないため、インフラストラクチャのアップタイムの最大化が求められます。Oracle Unbreakable Linuxサポートを Basic レベルおよびPremierレベルで契約しているOracle Enterprise Linuxベースの管理サーバーは、追加のライセンス・コストをかけずに、Oracle Clusterwareを使用してクラスタリングすることができます。これにより、管理サービスのフェイルオーバーとリカバリを自動化してダウンタイムを最小限に抑えることが可能になります。サーバーの再起動とリストアを手動で行う必要はありません。

管理サーバーの一時的な停止でさえ、VMの実行を妨げることはありません。Oracle VMの分散管理アーキテクチャにより、管理サーバーが一時的に使用不能になっても、セキュア・ライブ・マイグレーションやゲストHA自動再起動といったほとんどのVM操作を正常に実行することができます。

可用性の高いエンタープライズ仮想化

仮想化は、新たな問題を生み出すのではなく、問題の解決に役立つものでなければなりません。Oracle VMは本稼働中の企業のデータ・センターにおける使用を想定して開発されており、アプリケーションの配置、管理、サポートを容易にするだけでなく、セキュアな環境での可用性を高めることを目的としています。

基礎となるハードウェアの物理的制約からワークロードを切り離すことで、Oracle VMはサーバー・プールのロード・バランシングやセキュア・ライブ・マイグレーション、ゲストVMの高可用性などの機能を活用し、ゲストVMとワークロード・アップタイムの劇的な改善を支援します。

詳細情報

Oracle VMの詳細、およびフル機能を備えた完全なリリースの無償ダウンロードについては、Oracle.comのOracle VM Webページを参照してください。

Oracle Corporation 発行「Oracle VM - Creating & Maintaining a Highly Available Environment for Guest VMs」の翻訳版です。

Oracle VM - ゲストVMの高可用性環境の構築と管理
2008年9月
作成者: A. Hawley

Oracle Corporation
World Headquarters
500 Oracle Parkway Redwood Shores, CA 94065 U.S.A.
海外からのお問合せ窓口:
電話: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Oracle Corporation 発行「MDM as a Foundation for SOA」の翻訳版です。