



VERITAS

now from



VERITAS Global Cluster Manager™

**Heterogeneous Cluster Management and
Disaster Recovery Protection**

V
E
R
I
T
A
S

W
H
I
T
E

P
A
P
E
R

ペリタスソフトウェア株式会社および VERITAS Software Corporation は、本資料に含まれている誤り、本書によるソフトウェアの動作、および本資料の使用に関して、いかなる損害（直接損害、間接損害、付随的損害、結果的損害、逸失利益、懲罰的損害を含むがこれらに限定されない）に関して責任を一切負わないものとします。これはペリタスソフトウェア株式会社および VERITAS Software Corporation がかかる損害の発生する可能性について知らされていた場合でも同様です。

本書の内容は、2001 年 8 月現在のものです。

発行元：ペリタスソフトウェア株式会社

目次

はじめに.....	1
VERITAS Global Cluster Manager	1
VERITAS Global Cluster Manager のアーキテクチャ.....	4
災害対応の取組みと管理	7
レプリケーションの機能	8
レプリケーションエージェントの処理例	9
Global Cluster Manager の運用.....	10
まとめ	12

はじめに

企業のシステム環境で用いられるクラスタ技術は、これまで基幹になるデータベースの連続稼働を維持することが主眼でした。厳しいデータ主導のデジタル情報社会で生き残るためには、社内データの心臓部にあたるこの部分の可用性を保証する方策は欠かせないものとなってきました。最近では社内システムが成長してくるにつれて、こうしたデータ保護に対する要求が広がりを見せ、クラスタはデータベースの可用性維持に限らずいろいろなアプリケーションサービスにも使われるようになってきました。システムに異常が発生しようと保守作業を行おうと、できる限りの連続稼働が、データベースにとどまらずトランザクションやアプリケーションの段階でも求められ、様々な業務に使われるシステムすべてにその要求が広がっています。企業では、事業単位の区分に沿ってシステムの連続稼働の問題に対処する動きもあります。アプリケーションやリソースのクラスタをこれまでよりも大きな規模で定義して、大きな単位で管理するこの傾向は、フェイルオーバーの対象や方法、そしてリソースグループのフェイルオーバーなどに、厳密なガイドラインの設定を必要とします。このようなより高い可用性へのニーズに応えるために、これまでよりも高度なクラスタを構成する要求が徐々に高まっています。

一方、災害対応の面からも新しい要求が浮かび上がってきました。クラスタ内のノードからノードにアプリケーショングループをフェイルオーバーするのではなく、クラスタ全体をそのまま別のクラスタにフェイルオーバーする機能です。自然災害、あるいは計画停電で社内の情報中枢が止まってしまうというのでは、情報経済の中で望ましい姿とはいえません。こうした時に、地理的に離れた別の拠点にリソースグループやアプリケーションを速やかに移行できれば、情報へのアクセスを守ることができます。地理的分散によるフェイルオーバーの必要性は、今日の企業可用性・インフラストラクチャに新しい次元の要求をもたらしています。

時として企業の競争力にも関わるこうした複雑な要求の中で、システムに含まれるクラスタの数は数十または数百あるいはそれ以上に膨れ上がり、それぞれが20あまりのノードをもち、さらにその各ノードが同程度の数のリソースグループを抱えています。それだけの数のクラスタが、地理的には数千キロの範囲にまで広がり、しかも様々なオペレーティングシステムが入り交じっています。連続稼働を維持し整備された社内システムを実現して円滑に運用していくためには、こうした環境を効率よく管理できることが課題となっています。

本書は、VERITAS Global Cluster Managerの概要を解説したものです。Global Cluster Managerはアプリケーションの連続稼働を広域レベルにおいても保護するソフトウェアで、WANに広がるクラスタやアプリケーショングループのビジュアルな把握、レプリケーション管理、ディザスタリカバリ保護などの機能によって、総合的なハイ可用性計画の一翼を担います。

VERITAS Global Cluster Manager

VERITAS Global Cluster Managerは、WebベースのインタフェースからVERITAS Cluster ServerTM環境をビジュアルに集中管理できる総合的なクラスタ管理ソフトウェアです。各地に分散したいくつものCluster ServerクラスタをGUIでまとめて表示し、管理することができます。クラスタの境界を超えてイベント/アクションが設定できるので、実質的にはグローバルなクラスタを構成することになり、広

域規模のフェイルオーバが可能になるほか、それを利用した災害からの復旧（ディザスタリカバリ）にあたることができます。

Global Cluster Manager は VERITAS Cluster Server を使っています。VERITAS Cluster Server は高い拡張性を備えたクラスタソフトウェアで、クラスタ内のアプリケーションの稼働状態を監視し、ハードウェアやソフトウェアに異常が発生すると別のノードでアプリケーションを再起動します。最大 32 個のノードに対応でき、現在のところオペレーティングシステムは Solaris、HP-UX をサポートしています（将来、Windows と IBM AIX のサポートを予定）。

VERITAS Global Cluster Manager は、Cluster Server アプリケーションとクラスタ情報を管理し、イベントアクションやレプリケーションのコマンド、状態情報を遠隔拠点間で TCP/IP 通信を介してやり取りすることで、Cluster Server 環境の機能を広域レベルにまで拡張します。操作インターフェースには Web ブラウザを利用し、システム管理者がどこからでも全クラスタの監視と管理ができるコンソール機能を提供しています。

開発の背景

近頃では複数のクラスタを集約管理する要求が、いろいろな場面からもち上がっています。

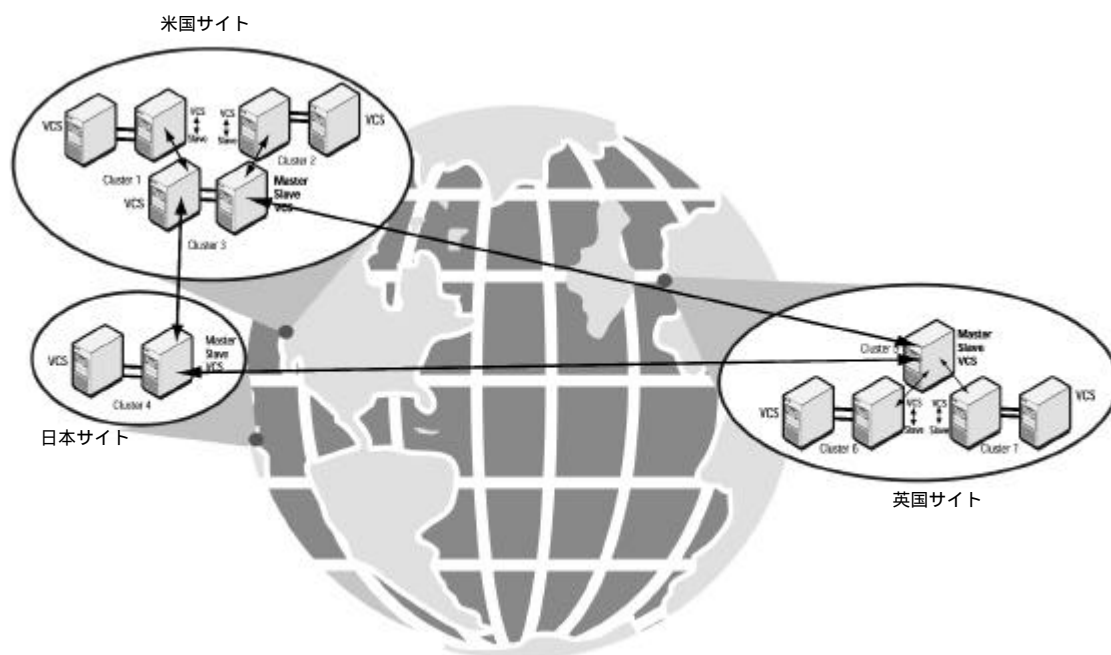
- **クラスタの集約管理** - すでにたくさんのクラスタを抱えているシステムでは、全部のクラスタとアプリケーションを、たとえば一つのツリー構造でまとめて表示できればという要望があります。
- **機能による区分** - 互いにわずかに関連するアプリケーションは、クラスタを分割してそれぞれで独立して処理させることを考える方が得策です。複数の異なるサービスを一つのクラスタにまとめてしまうのは、欠点こそあれ利点はあまりありません。
- **事業単位に沿った分割** - システムのリソースやアプリケーション運用を事業単位ごとに分割する要求もしばしば見られます。クラスタを分ければ、フェイルオーバや災害対策の方針を事業単位で独自に立てることができます。
- **多様なハードウェアへの対応** - ハードウェアが混在した環境では、無理に一つのクラスタにするよりもクラスタを分けてしまった方が良い場合があります。
- **SAN のスケーラビリティ** - SAN の利用が広がるにつれて、複数クラスタへのニーズが高まるのは必然です。
- **地理的分散** - Cluster Server のノード間専用通信の GAB (Global Atomic Broadcast) が使えなくなる水準までサーバの配置が広がると、クラスタを分けざるをえません。
- **災害対応** - クラスタ環境がすべて停止した場合に、他のクラスタ(サイト)でアプリケーションを引継ぐ機能への要求が高まっています。

広域クラスタ管理の機能

Web ベースのコンソールからは、様々なハードウェアやオペレーティングシステムが混在したたくさんのクラスタを一括して監視し管理することができます。各種のイベントが広域レベルで処理でき、アクションの自動実行も設定できます。一つのコンソールからの統合管理方式によって、大きく 4 つ

の機能を満たすことができます。

- **クラスタの拡張** - Global Cluster Manager の構成をオンラインにしたままでシステムへのクラスタの追加を行えます。新しいクラスタを設定すると、そのクラスタは Global Cluster Manager のフレームワークに自動的に組み込まれます。
- **監視と管理** - クラスタの集合単位である「サイト」には、様々なオペレーティングシステムが混在したクラスタが共存でき、システム全体を一つのツリー構造で表示して、共通のコマンドセットを介した管理が行えます。
- **サイトの移行** - アップグレードや保守作業の場合に、サイトの稼働リソースを代替サイトに移行することで、システム停止をほぼゼロにすることも可能です。世界中に拠点をもつ企業の場合は、時間帯に応じて稼働サイトを順次移行していくことで、日中のアクセスが集中する地域に近接したサイトから常時サービスを提供していくことができます（この方式を「フォロー・ザ・サン」と呼んでいます）。
- **災害対応** - 地理的に分散したシステムでは、万一の災害などによってある拠点のサーバがすべて停止した場合、Global Cluster Manager によってアプリケーションを他の地域のクラスタへフェイルオーバーするポリシーを設定しておくことができます。本来一つにすべきクラスタが、地理的に分散しているために複数のクラスタに分けざるをえなくなったケースでも、Global Cluster Manager がいわばクラスタのクラスタとなってアプリケーションのフェイルオーバーを可能にします。



VERITAS Global Cluster Manager のアーキテクチャ

Global Cluster Manager は、非常に拡張性の高いアーキテクチャ上に管理機能を構成しています。ノードの最大数は 16,000 個、リソースグループの数は事実上無制限で、そのすべてを一つのコンソールから管理することができます。

オブジェクトの階層

- **サイト** - VERITAS Cluster Server クラスタの集まり。全体で最大 8 個のサイトが作成できます。
- **クラスタ** - GAB でつながれたシステムの集まり。1 サイトには最大 64 個のクラスタが入ります。
- **システム** - クラスタに入ったコンピュータ。1 クラスタは最大 32 台のシステムで構成できます。
- **グループ** - 一つのサービスに必要なリソースの集まり。

プロセス構成

Global Cluster Manager Master (サイトマスタ) - サイトマスタは、サイトあたり一つのクラスタ内の一つのシステム上で実行されます。サイトマスタ自身も VCS (VERITAS Cluster Server) サービスグループの一つなので、サーバやサービスに異常が起こるとクラスタ内で自動的に別のノードにフェイルオーバーされます。ここでいうサイトとは、上にあげた Cluster Server クラスタの集まりのことです。サイトマスタは、一つ一つのサイトスレーブから接続・状態情報を受け取って、サイト内のすべてのクラスタ情報をまとめて表示することができます。

サイトマスタは、サイト内の全クラスタのシステムとアプリケーションについての情報をまとめると、ほかのサイトのサイトマスタと通信してその情報を共有します。ですから、どのサイトマスタもシステム全体についての情報を共有しています。クラスタ内の情報に変化があれば、Cluster Server がローカルのサイトスレーブに情報を送り、そこからサイトマスタ、そしてほかのサイトマスタへと情報が伝達されます。

サイトマスタは、なにかイベントが発生した時に、たとえサイトスレーブが止まっている場合でも問題の発生や変化を間接的に知ることができる冗長性を備えていて、WAN やノード間リンクの問題や切断に対して強健性があります。

Global Cluster Manager Slave (サイトスレーブ) - クラスタ毎に一つ実行されるスレーブのプロセスで、クラスタ内のどのシステムで実行していてもかまいません。サイトスレーブは Cluster Server の管理下にあるアプリケーショングループの構成や状態の情報を受け取って通知します。このサイトスレーブも Cluster Server アプリケーションの一つで、Cluster Server の管理機能によって、それぞれのクラスタに必ず一つのサイトスレーブだけが稼動するようになっています。

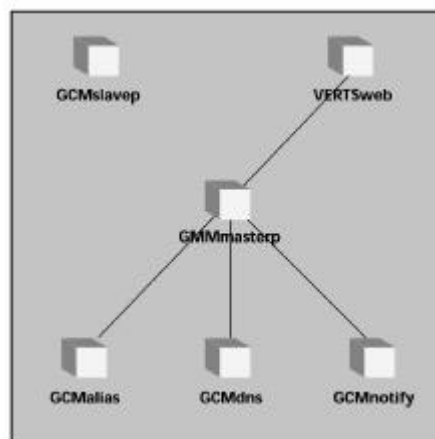


コンポーネント構成

Global Cluster Manager の各コンポーネントは、独立した Cluster Server アプリケーションとして実行されるので、ほかのアプリケーションと同じように管理とフェイルオーバーの機能の恩恵を受けます。コンポーネントがフェイルオーバーする時には、サイトスレーブや他のサイトマスタとの接続がすみやかに再確立されます。ソフトウェアをアップグレードする時の作業も、ほかの Cluster Server アプリケーションと変わりません（停止、パッケージの削除、再インストール、再起動）。

Cluster Service Group のコンポーネント

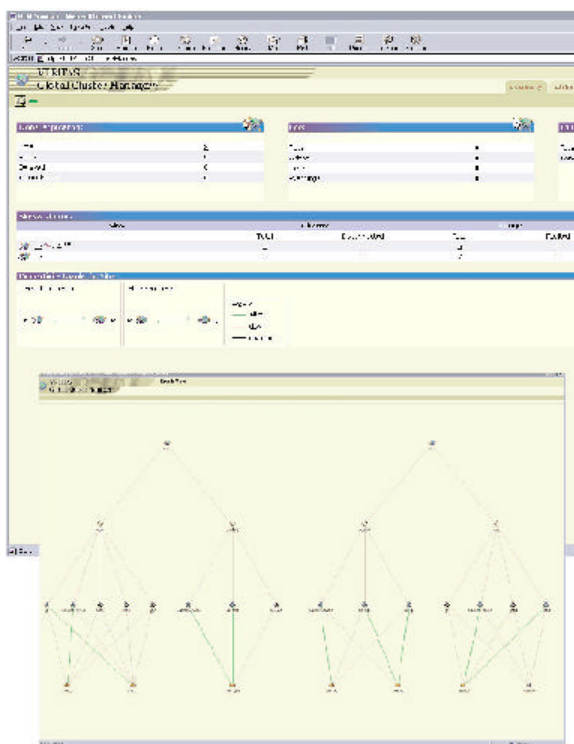
- サイトマスタ：GCMmaster
- サイトスレーブ：GCMslave
- DNS の更新：GCMdns
- SNMP/SMTP 通知：GCMnotify
- IP エイリアス：GCMalias
- グラフィカル・インタフェース用の Web サーバ：VRTSweb



スレーブ側には GCMslave と VERITAS Web コンポーネントだけが入ります。

コマンドの処理

- まずクライアントがサイトマスタに接続してコマンドを発行します（コマンドラインまたは GUI を使用）。
- サイトマスタは該当する Cluster Server クラスタにコマンドを転送します（必要ならば別のサイトマスタを経由します）。
- 処理結果がクライアントに戻されます。
- Cluster Server はスレーブをほかのクライアントと同じように扱うので、データ破壊などの異常の発生を未然に防ぐことができます（競合するオンラインコマンドを排除するなど）。
- Global Cluster Manager には VERITAS Cluster Server が必要です。その他に特に必要な製品はありません。



操作インターフェース

Web ブラウザを利用した GUI (グラフィカル・ユーザ・インターフェース) を備え、オンラインでどこにいてもわかりやすい操作が可能です。画面からは、集計情報や、階層構造、リソースの状態、その他のグローバル情報表示、ポリシー設定、ログ操作、リアルタイムの状態表示などの機能が利用でき、オンラインヘルプも備わっています。

コマンドラインからも、GUIと同じ機能セットのコマンドが利用できます。Global Cluster Manager では、グループのオンライン/オフライン・コマンドが、どの場所のグループにも利用できます。操作履歴を記録するログは、レプリケーションされながらディスクに保存されます。ログの検索や並べ換えは GUI から実行できます。

Cluster Server との間では、Cluster Server の Web GUI へのリンクを通して操作インタフェースを統合しています。クラスタや、グループ、システムのレベルで情報を監視する機能が強化され、管理機能が拡張されています。

将来はクラスタ管理の枠組を SAN に広げ、各種の機器が混在した SAN 環境から一つのストレージリソースまで、詳細なビジュアル表示と管理ができるようになる予定です。

災害対応への取組みと管理

災害に強いインフラストラクチャ（ローカルシステム）の構築

広域にわたる災害対応も、ローカルレベルで堅固な基盤を構築することが第一歩です。例えば、

- VERITAS Cluster Server によるアプリケーションフェイルオーバー
- VERITAS Volume Manager™ によるディスクミラーリング
- VERITAS File System™ によるファイルの高速リカバリ
- ノード間やディスク間の経路の二重化
- 冷却ファン、筐体、電源、そしてバックアップ媒体管理などの適切な計画

ローカルの概念の拡大による強健性の強化

- ローカルの機器を離れた建物に分散すると、地域的な停電に耐えることができます。
- Volume Manager のミラーリング機能は、ファイバチャネルの限界の 10km まで距離を広げることができます。
- クラスタ（VCS クラスタ）のノード間距離は、ノード間専用リンクの許す限り延ばせます。
- これらによってローカル内の機器を地理的に分散しておくこと、計画停電などの影響も受けなくなります

広域レベルでの管理とフェイルオーバー

ローカルレベルで強靭なシステムが組み上がれば、その次の段階が広域レベルでの対策です。広域レベルのフェイルオーバーを導入すると以下のようなことが可能になります。

- アプリケーションとクラスタのグローバルな配置
- サイト全体に対する災害対応
- サポートシステムやコールセンターシステム（国際規模での 24 時間運用）
- ファイバチャネルや SCSI によるミラーリングの距離制限を超えるフェイルオーバー

- これらを総合して業務システムの安定した連続稼働が実現できます。
- ノード間高速通信の LLT (Low Latency Transport) リンクの多重化は広域レベルでは困難です。

レプリケーションの機能

Global Cluster Manager は外部のレプリケーション機能とやり取りするレプリケーションのフレームワークとエージェント、そしてレプリケーションの状況を監視するグローバルアプリケーションを備えています。システム管理者は、未完了の書込みの程度や通信リンクの状態などのレプリケーションに関する各種の情報を手もとで知ることができ、拠点全体が停止してフェイルオーバーする場合のほか、計画停電の時などにも、状況を正しく把握しながらフェイルオーバー実施の要否やタイミングを判断することができます。

レプリケーション機能には、VERITAS Volume Replicator™ などのソフトウェアベースのレプリケーション製品のほかに、ハードウェアによるレプリケーション機能を備えた機器も利用することができ、それらの製品によるレプリケーションの設定、制御、管理、状態情報の収集などが、Global Cluster Manager の Web 画面から行えます。

レプリケーションのフレームワーク

Global Cluster Manager は、他社のレプリケーション製品を組み合わせることができるフレームワークを備えています。おもな特徴を列挙すると、

- ・ 計画的なシステム停止による移行と、異常が発生した時のフェイルオーバーの両方に対応する柔軟な管理ができます。
- アクションの送出や結果のログ記録はレプリケーションが行われる場所で実行され、リモートコマンド (rsh や rcp など) を経由してではなく、各所でローカルにレプリケーションコマンドが実行されます。
- レプリケーションのフレームワークは、Cluster Server のエージェントフレームワークになっています。
- Cluster Server や、Volume Replicator、あるいは他社製品とのやり取りをエージェントが担当し、統一されたポリシー設定をベースとしてフェイルオーバーの柔軟な操作を可能にしています。
- エージェントはレプリケーションが実行されるすべてのノードにインストールされます。
- レプリケーションは同期/非同期のふたつのモードが利用でき、フェイルオーバーもどちらにも対応しています。
- Solaris、HP-UX などオペレーティングシステムの違いを超えた、共通のコマンドセットを提供しています。
- イベントによるアクションのトリガは、リソースグループ、クラスタ、サイトなどいろいろな段階をベースとして実施できます。

- エージェントは一定周期で `getinfo` エントリポイントを呼び出してレプリケーションに関する情報を集め、それを Global Cluster Manager に渡します。これによって、システム管理者はレプリケーションに関する問題と、アプリケーションの状態を一つのコンソールから知ることができ、システム異常が発生した時に同期の遅れの程度によって移行を実施するかどうかを判断するなど、様々な役立てることができます。アクションはモジュール化に適し、処理をそれぞれの場所だけで実行させることが保証できます。エージェントはレプリケーションが実行される場所でローカルにコマンドを発行し、その結果がサイトマスタに戻ってきます。

グローバルアプリケーション

グローバルアプリケーションはクラスタやサイトをまたいだアプリケーショングループの情報を保持することができ、Cluster Server やグループオブジェクトより上のレベルで監視を実行できます。

レプリケーションは Global Cluster Manager のグローバルアプリケーションモジュールを経由して構成します。クラスタ内の複数のアプリケーションをグローバルアプリケーションが仮想的にリンクします。レプリケーション関係を指定すると、Global Cluster Manager がレプリケーションエージェントを自動的に設定してアプリケーショングループに加え、処理の開始、停止、フェイルオーバーを制御します。

グローバルアプリケーションにエージェントパラメータを送ると、そこでインテリジェントに処理されて、該当するアプリケーション処理部 (Cluster Server コマンド、Volume Replicator コマンドなど) に適切な順序で渡されます。その結果は、GUI やコマンドラインで統合されて表示されます。

レプリケーションエージェントの処理例

VERITAS Volume Replicator エージェント

通常は、`getinfo` から `vxrlink` ステータスコマンドを一定周期で実行して、同期遅れの程度といった情報を表示します。

アクションは下のものが実行できます。

- 昇格 (Promote)
- 降格 (Demote)
- 変換 (Convert)
- 再同期 (Resynch)

グローバルアプリケーションコマンドの `gagrp` を実行すると、該当グループのエージェント (どこで実行されているかにかかわらず) にパラメータが送られます。

一次サイトが停止した時：

データの喪失が許容できるかどうかを判断

許容できればフェイルオーバーを実行

その後、元の一次サイトが再立上げた時：

二次サイトに変換

一次サイトへ再同期

同期がとれれば元の一次サイトに戻します：

代替の一次サイトを降格

二次サイトを昇格

Global Cluster Manager の運用

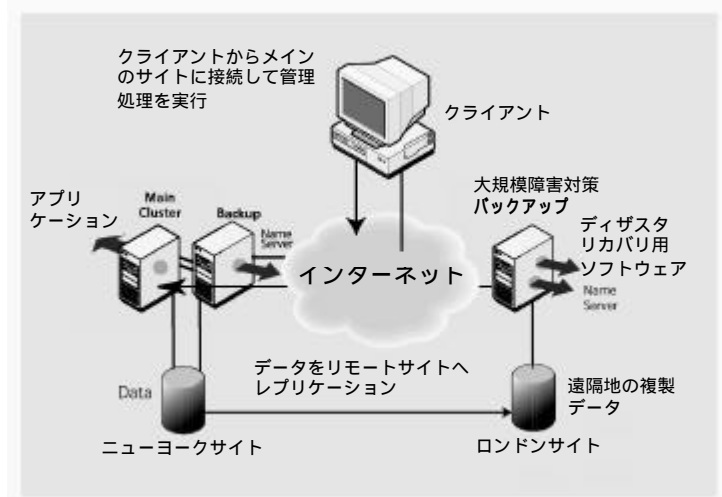
ローカルレベルで堅固なシステムを構築し、適切なレプリケーション製品を導入したら、Global Cluster Manager がコマンドセンターとして機能し、WAN をまたぐポリシーベースのフェイルオーバーが可能なハイアベイラビリティ構成を実現します。

Global Cluster Manager の動作

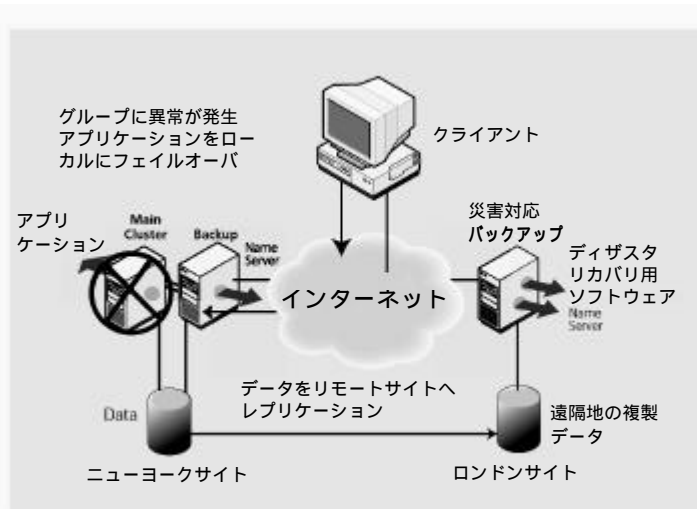
- サイト稼働状態の監視 - サイトマスタは、WAN を介した他のサイトマスタのクラスタへ ping コマンドを送ってその稼働状態を知ります（この処理をハートビートと呼びます）。デフォルトではサイトマスタのエイリアス宛に ICMP ping が送られます。そのほかにも無線や衛星通信などの手段を加えて、状態検出の精度を高めることができます。
- サイト間には複数のリンクを設定でき、遠隔サイトの稼働状態の検出を補強することができます。あるサイトに対して他のすべてのサイトがすべてのハートビートで停止を検出した時に、そのサイトが停止したと判断されます。またサイト内部のリソース（アプリケーション）グループは、クラスタの全ノードで異常を起こした時、またはクラスタやサイト全体が停止すると停止したと判断されます。
- アプリケーションのフェイルオーバーを実施する時には、システム管理者の確認操作を経るのがデフォルトです。設定によって自動的にフェイルオーバーさせることも可能です。
- クラスタ間のアプリケーション移行制御 - リモートのサイトマスタから該当するクラスタの Cluster Server にコマンドを発行すると、アプリケーションをそのクラスタで起動することができます。
- レプリケーション管理 - 他社のレプリケーション製品ともやり取りができ、クラスタやサイトの移行の時にはパラメータが自動的に変更されます。

通常の稼働状態

Global Cluster Manager は全サイトにわたってアプリケーションとクラスタの状態を監視しています。データはリモートサイトに複製され、Global Cluster Manager がレプリケーションの状態を追跡して表示し、遅れや異常を報告します。ソフトウェア/ハードウェアを問わないレプリケーション製品との統合によって、万一の災害への備えを普段から整えています。



ローカルレベルのアプリケーション異常によるグループ・フェイルオーバー

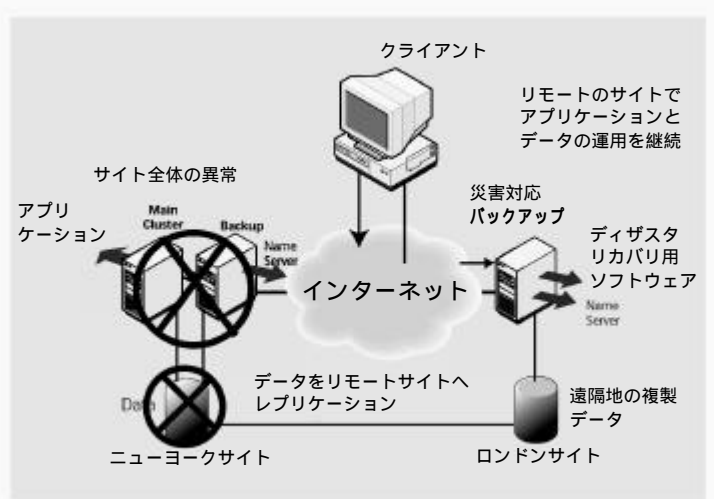


Global Cluster Manager は VERITAS Cluster Server を介してクラスタ内のアプリケーションの異常を検出し、VERITAS Cluster Server によるフェイルオーバーを監視して報告します。必要があれば、VERITAS Cluster Server のイベントとトリガを Global Cluster Manager の操作インタフェースから変更することもできます。

異常が発生してもデータのレプリケーションは継続して行われ、アプリケーションへのアクセスも途絶えることはありません。

サイト規模の異常によるフェイルオーバー

災害によるフェイルオーバーの場合、あるいは保守作業などによるサイト移行の場合、Global Cluster Manager がポリシー設定に基づいて効率的に処理を行います。ネームサーバの更新がトリガされて、リモートの移行先サイトへクライアントアクセスが途切れることなく振り替えられます。



通常の継続的なレプリケーションによってデータの整合性ができる限り保たれているので、クライアントはリソース移行による差異をほとんど感じないはずで、リモートサイトの昇格とともにレプリケーションの役割も変更されます。元のサイトの問題が解消して再び起動し同期が取れば、すみやかに再移行ができるように準備をしています。

- ネームサービスの更新 (Bind v.8 をサポート) - IP アドレスの引継ぎも確実にいきます。
- レプリケーション製品はソフトウェア/ハードウェアのどちらのものも利用できます。
- 同期/非同期のレプリケーションが設定できます。
- 同期モード - レプリケーション先の書込みが完了するまでアプリケーションへ I/O 完了が通知されません。データの整合性が確実に保てる反面、通信速度の影響を受けます。
- 非同期モード - レプリケーション先への書込みはキューに入り、先にアプリケーションに I/O 完了が通知されます。その後ある時点でデータが送信されるので、同期が取れていない時に移行が起こるとデータが失われます。しかしミラーリングなどと同程度のパフォーマンスを提供します。

まとめ

システムの連続稼働への要求が高まるにつれて、広い地域に分散したクラスタ構成を効率的に管理する重要性が増してきました。Global Cluster Manager はそのようなシステムの基盤として活躍するものです。広域システムのビジュアルな管理から、サイト規模のフェイルオーバー、さらにはカスケード方式の多重フェイルオーバーによって、Global Cluster Manager は、数千キロの範囲に広がる何百ものクラスタグループを能率よく処理しながら、企業の重要データを守るフレームワークとなるはずで



V
E
R
I
T
A
S

W
H
I
T
E

P
A
P
E
R

Copyright ©2001 VERITAS Software Corporation. All rights reserved. VERITAS, VERITAS SOFTWARE, VERITAS ロゴ、VERITAS Global Cluster Manager, VERITAS Cluster Server, VERITAS Volume Manager, VERITAS File System, VERITAS Volume Replicator は、米国および各国のベリタスソフトウェア・コーポレーションの商標または登録商標です。また、本書で言及されているその他の名称は、それぞれ各社の商標または登録商標です。

製品の仕様・性能等は予告なく変更する場合がありますので、ご了承ください。 本資料の無断転載を禁じます。

ベリタスソフトウェア株式会社

〒100-0011 東京都千代田区内幸町2丁目2番2号 富国生命ビル
TEL 03-5532-8242 FAX 03-5532-0889
<http://www.veritas.com/jp>

お問い合わせ先