

## Windows Server® 2008 によるリモート拠点間クラスタリングのサポート

SteelEye DataKeeper for Windows によるフェールオーバー・クラスタリング

### 概要

SteelEye DataKeeper は、Windows Server 2008 フェールオーバー・クラスタリングとシームレスに連携することで、WAN 環境に有効なレプリケーション・ソリューションを提供します。Windows Server 2008 フェールオーバー・クラスタリングでは、異なるネットワーク・セグメント間のクラスタリングやハートビートの設定を調整するパラメーターなどの新機能が追加され、地理的に分散したクラスターを簡単に管理できるようになりました。SteelEye DataKeeper は、データ・レプリケーション機能により Windows Server 2008 フェールオーバー・クラスタリングのディザスタリカバリー構成を拡張します。

### リモート拠点間クラスター技術を振り返る

Microsoft®は 1997 年にWindows NT向けにMicrosoft Cluster Services (MSCS、コードネーム「Wolfpack」)をリリースしました。MSCSは当初、共有SCSIを使用していましたが、ストレージ製品が充実すると、共有ファイバー・チャネルやiSCSI SANを使用するようになりました。Windows Server 2008 がリリースされる以前は、MSCSで地理的に分散したクラスターを構成するオプションには限界がありました。MSCSでクラスターを構成するには、事前に認定された共有ストレージ・アレイでデータを複製しなければならず、MSCSの遅延要件を満たした上で、ノードを同じサブネット上に配置しなければなりませんでした。そこで、地理的に分散した遠隔拠点間クラスターが求められるようになりました。

こうした課題を解消するために、一部の顧客企業は、サードパーティーが開発した技術を導入して、データ・レプリケーション、アプリケーションとシステムレベルの監視、フェールオーバーを実行していました。サードパーティー製ソリューションは、ミッション・クリティカルなアプリケーションとしてポピュラーであった Microsoft Exchange、SQL Server、BlackBerry Enterprise Server、VMware Virtual Center などに対応していました。SteelEye LifeKeeper もそうしたアプリケーションの一つでした。

Microsoft Exchange Server 2007 では、クラスター連続レプリケーション (Continuous Cluster Replication: CCR) という技術が追加されて、共有ストレージ・アレイを使用しないクラスタリングが初めて可能になりました。CCR はアプリケーションベースの複製機能であり、アプリケーション自身が複製機能を提供しています。Exchange 2007 CCR は、現在に至るまで、マイクロソフトのクラスタリング技術と協調的に動作する唯一のネイティブな複製技術です。

### Windows Server 2008 フェールオーバー・クラスタリングの新機能

Windows Server 2008 フェールオーバー・クラスタリング (WSFC) は、マイクロソフト社が開発したフェールオーバー・クラスタリング・ソリューションの新しい名称です。この名称は、マイクロソフトが提供する新しい「高性能」な Windows HPC Server 2008 と区別するために命名されました。

WSFC で追加された新しい機能は、マイクロソフトが提供する技術の中で最も堅牢なフェールオーバー・クラスターを簡単に構築できるソリューションとなっています。次節で説明する新機能は、地理的に分散したクラスターを構成する際の問題を巧みに解消しています。

## クラスター検証ツール

マイクロソフトが WSFC で実現した最も重要なポイントは、クラスター・ハードウェア互換性リスト (HCL) を廃止したことです。WSFC がリリースされる以前は、MSCS クラスターに使用するハードウェアは、マイクロソフトによるテストを受けて、クラスター・ハードウェア互換性リスト (HCL) に登録しなければなりません。WSFC で追加されたクラスター検証ツールは、HCL を参照しなくても、クラスターを自由に構成することを可能にしました。クラスター検証ツールは、ハードウェア互換性リストに代わり、「システム、ストレージ、ネットワーク構成がクラスターに適していることを検証します。」

クラスター検証ツールは、クラスターに最適なハードウェアとソフトウェアを自動的に検証するため、マイクロソフトによる認定を事前に受けなくても、それらを自由に使うことができます。

## 新しいクォーラム・モデル

マイクロソフトは、WSFC をリリースする前から、マジョリティー・ノード・セット (MNS) と File Share Witness (FSW) 機能による新しいクォーラム (クラスター情報) オプションを導入しました。この斬新なオプションは、WSFC で新たに導入された次の 4 つのクォーラム・モデルの前身となりました。

- ・ マジョリティーなし (ディスクのみ)
- ・ ノード・マジョリティー
- ・ ノードとディスク・マジョリティー
- ・ ノードとファイル共有マジョリティー

[この文書に記載されている用語と製品名は、それぞれの会社と所有者の商標であり、著作権で保護されています。]

この中で「ノードとファイル共有マジョリティー」は、遠隔拠点間クラスタリングに最適なモデルです。このモデルでは、単純なファイル共有をクラスターの「証人 (witness)」として識別することで 2 つのノードを WAN 上でクラスタリングします。ファイル共有と 2 つのクラスター・ノードは、それぞれ「票 (vote)」を獲得します。大半の票が有効な限り、クラスターはオンラインで生き続けられます。

## サブネットを超えたクラスタリング

WSFC が登場するまで、クラスター・ノードは同じサブネット上に配置しなければなりません。地理的に分散したクラスターを展開するには、ネットワークを変更しなければならないため、VLAN を使用して、WAN 回線でサブネットを接続する必要がありました。しかし、Windows Server 2008 では、論理和 (OR) の使用が認められ、2 つの IP アドレスが共存できるようになりました。2 つの IP アドレスが別のサブネットに常駐できるようになり、VLAN は不要になりました。

## 設定可能なハートビート・タイムアウト

WSFC では、ラウンド・トリップ遅延が 500 ミリ秒未満に緩和されました。ハートビート値を設定できるようになったため、遠隔拠点間クラスターで狭帯域ネットワークをサポートできるようになりました。

## SteelEye DataKeeper とは

SteelEye DataKeeper は、WAN 環境に適合したレプリケーション・エンジンであり、データを効率的かつ高速に複製する高度な機能を備えています。その特長は次のとおりです。

- ・ ブロック単位でボリュームを複製 - ホスト上で同期または非同期オプションを選択可能
- ・ 継続的データ保護 (CDP) により任意の時点のデータを復旧可能
- ・ WAN 環境への適合 - WAN アクセラレータ (ハードウェア) を使用せずに、SteelEye DataKeeper だけで高速ネットワークと狭帯域ネットワークを集約可能
- ・ 圧縮アルゴリズムで帯域を有効活用
- ・ ビットマップによるロギング - 完全な再同期が不要
- ・ 帯域幅制限でレプリケーション・トラフィックを完全に制御
- ・ 直感的に使用できるウィザードベースの MMC 3.0 ユーザー・インターフェース
- ・ フェールオーバー構成に LAN/WAN 上の複数のターゲットをサポート
- ・ 従来の 2 ノード共有ストレージ・クラスターを 3 ノードに拡張 - 障害回復機能が強化

以下の節では、前記の特長を個別に取り上げて、各機能がデータ・レプリケーション・ソリューションで重要な役割を担っていることを紹介します。

## ブロック単位のレプリケーション

SteelEye DataKeeper はデータをブロック単位で複製するため、レプリケーションを高速かつ効果的に作成できます。SteelEye DataKeeper は Windows のフィルター・ドライバとして実装されており、ファイル・システムとデータ・ボリュームの間に位置しています。そのため、ファイルのオープンとロック、NTFS のアクセス権、暗号化ファイルがあっても、問題なく動作します。こうしたファイルは、保護対象のデータブロックと同様に、単なるデータブロックとして処理されます。そのため、SteelEye DataKeeper はホストシステムを過負荷にすることなく、レプリケーションをリアルタイムで作成できます。

## 同期モードと非同期モード

SteelEye DataKeeper は、最適なデータ保護構成を選択できるように、同期モードと非同期モードをサポートしています。同期モードを選択すると、ターゲット・システムは常にソース・システムと同期されます。ただし、ソース・システムは、ターゲット・ボリュームへの書き込みが完了するまで、次のデータを送信できないため、このオプションを使用するときは注意が必要です。特に、狭帯域ネットワークでは、レプリケーションを作成する際の遅延が大きな問題となります。そのため、同期モードは、高速ネットワークに使用するか、アプリケーションのパフォーマンスよりも、データ保護を優先する環境で使用することをお勧めします。

一方、非同期モードは、データをターゲット・システムに送信する間にソース・ボリュームへの書き込みを完了します。ディスクへの過剰な書き込みが発生すると、SteelEye DataKeeper はソース・システムのキューを使用して、データを書き込む順番を保証します。非同期モードには、狭帯域ネットワークでもソース・サーバーのパフォーマンスに影響を与えることなく、レプリケーションを作成できる利点があります。ただし、ソース・システムに予期せぬ障害が発生すると、キュー内のデータがターゲット・システムで使用できなくなる欠点もあります。

## WAN 環境への適合

狭帯域ネットワーク上でデータを送信する場合は、遅延による影響を考慮しなければなりません。同じ WAN 回線を使用して、高速ネットワークと狭帯域ネットワークに接続する場合は、遅延が大きな問題となります。SteelEye DataKeeper は、最初からこうした問題を考慮して開発されているため、遅延による影響は出ません。SteelEye DataKeeper を使用すると、たとえ遅延が発生した場合でも、ハードウェア・ベースの WAN アクセラレータを使用せずに、利用可能な帯域を最大 90% まで活用できます。SteelEye DataKeeper は 9 種類の圧縮レベルをサポートしています。最大で 5:1 の圧縮率を適用できるため、帯域を効率的に使用できます。

非同期モードを選択して、帯域幅が異なる WAN 回線上で SteelEye DataKeeper を使用したときのスループットを図 1 に示します。この図では、圧縮モードをオフ (0) に設定したときと、中程度の「3」に設定した場合のスループットを比較してあります。データ・タイプによっては、圧縮が非常に効果的であることが分かります。また WAN 回線の遅延はスループットに影響していません。遅延をそれぞれ 1 ミリ秒、32 ミリ秒、55 ミリ秒、128 ミリ秒以下に設定しても、同じ結果になっています。

**DataKeeper のスループット**  
MS Exchange とファイル・サーバーによるデータ転送

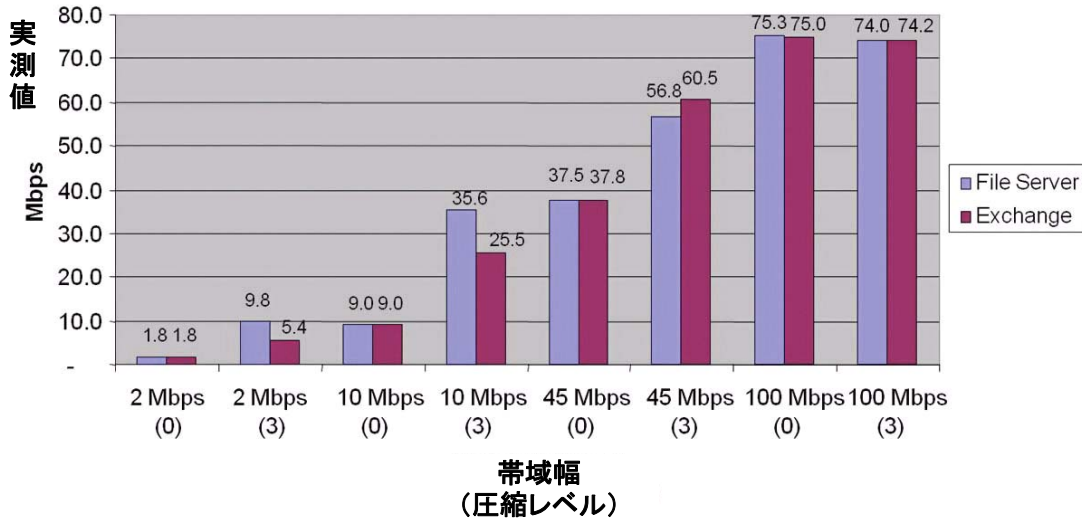


図 1

### ビットマップ・ファイル

WAN 経由でデータを複製する場合、すでに同期したデータは、データ・セット全体を再同期する際に使用しないことが重要です。低速な WAN 回線で大量のデータ・セットを同期すると時間がかかります。SteelEye DataKeeper は、永続的なビットマップを保持しており、どのような障害が発生しても、データを再同期する必要がありません。たとえば、ミラーリングのターゲットが停止しても、ソースはブロックをトラッキングして、ビットマップを更新します。

ビットマップはミラーの作成時に生成されて、固定サイズのファイルとして、メモリおよびミラー化されていないソース・システムのボリュームに格納されます。ミラーリングが停止した場合でも、ソース上で変更されたブロックは、ビットマップに「ダーティ」としてマークされます。ターゲット・システムと接続が再開されると、ダーティ・ブロックだけがターゲットに転送されます。

### 継続的データ保護 (CDP)

継続的データ保護 (Continuous data protection : CDP) とは、ストレージ関連の業界団体である SNIA (Storage Networking Industry Association) の定義によると、変更されたデータを継続的に記録・トラッキングし、それをプライマリ・データから独立して格納する方法です。CDP を使用すると、過去の任意の時点 (Any Point in Time: APIT) のデータを復旧できます。CDP システムは、トラッキングするレイヤーによって、ブロック、ファイル、アプリケーションの 3 タイプに分類されており、非常に細かい単位 (グラニュラリティー) でデータの復旧ポイントを設定できます。前記の定義によると、CDP ソリューションは次の 3 つの特長を備えています。

1. 保護対象のデータ変更を継続的に記録・トラッキングする。

2. データの変更は、プライマリ・ストレージと異なるロケーションに格納する。
3. リカバリー・ポイントを任意に設定できるため、リカバリー開始前にポイントを定義する必要がない。

ブロック・ファイル、アプリケーション・ベースで CDP を実装した新しい技術が次々に登場しています。現在では、多数のベンダーが特定のアプリケーションやデータ環境に対応した製品をリリースしています。しかし、どのようなアプローチを採用しても、CDP は、安価なコストで高速なデータ取得、データ保護の拡張、ビジネス・コンティニューイティ(事業継続性)を実現するものです。

SteelEye DataKeeper は、SNIA の定義に準拠した CDP ソリューションであり、ターゲット・システムで発生した書き込みを循環ログ・ファイルに記録してトラッキングしています。ネットワーク管理者は、ログ・ファイルを参照することでターゲット・ボリュームのロックを解除し、最適なリカバリー・ポイントが見つかるまで、複製したデータ・ストリーム内を移動できます。理想的な RPO (Recovery Point Objective : リカバリー・ポイント目標)を探して、データを任意の時点のデータまでロールバックできます。

## DataKeeper と WSFC の統合

WSFC では、地理的に分散した堅牢なクラスターを簡単に構成することが可能になりましたが、この機能をフル活用するには、WAN 上でデータを複製する機能が大切になります。マイクロソフトは、サードパーティのハードウェアとソフトウェア・ベンダに対して、データ複製機能を WSFC に追加することをサポートしています。ただし、Exchange 2007 CCR は、サポート対象から除外されています。SteelEye DataKeeper は、利便性、高機能、価格性能比を見事にバランスさせたソリューションとして、遠隔拠点間クラスターの構築を支援します。

## 構成例

最も基本的な構成(図 2)では、2つのノードで Microsoft フェールオーバー・クラスターを構成して、一方をプライマリ・データ・センター、他方を WAN に接続した遠隔地のデータ・センターに配置します。SteelEye DataKeeper はプライマリ・サーバー上のボリューム・リソースをセカンダリ・サーバーに複製します。ボリューム・リソースは、ローカル・ディスクに格納するか、iSCSI やファイバ・チャネルで SAN(ストレージ・エリア・ネットワーク)に接続したボリュームに格納します。

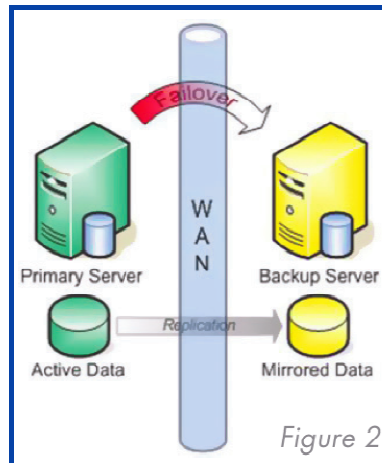


図 2

定期メンテナンスや計画外のネットワーク停止によりバックアップ・サーバーが起動すると、複製化されたボリュームのロックが解除され、逆方向のミラーリングが自動的に実行されます。その結果、プライマリー・サーバーが復旧すると、バックアップ・サーバーのボリュームに書き込まれた変更データは、プライマリー・サーバーへ自動的に複製されます。SteelEye DataKeeper は、固定サイズのビットマップ・ファイルを保持しているため、定期点検や計画外のネットワーク停止が長期化しても、データは再同期することなく残ります。

共有ストレージまたは複製化したハイブリッド・モデルの構成例を図 3 に示します。ハイブリッド構成では、プライマリー・サーバーとバックアップ・サーバーをプライマリー・データ・センターに置き、共有ストレージ・デバイスに接続します。SteelEye DataKeeper は、保護対象のボリュームを遠隔拠点の障害回復サーバーに複製します。

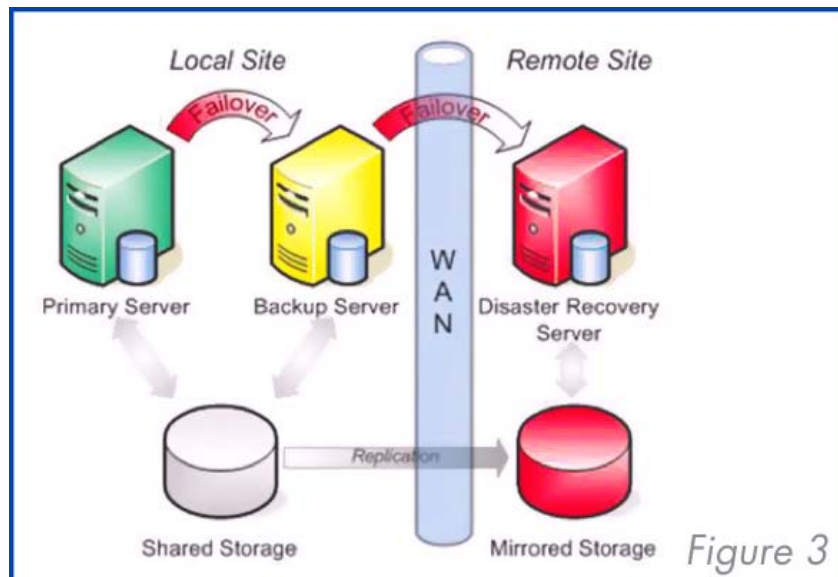


図 3

プライマリー・サーバーに障害が発生すると、セカンダリ・サーバーが瞬時にオンラインとなり、遠隔拠点へのレプリケーションを引き継ぎます。ローカル・サイト全体が停止した場合は、障害回復

サーバーがオンラインとなり、逆方向のミラーリングを再開します。障害回復サーバー上で変更されたボリュームは、ビットマップでトラッキングされます。そして、プライマリ・サーバーが復帰すると、レプリケーションがプライマリ・サーバーにミラーリングされます。

SteelEye DataKeeper は複数ターゲットへの複製機能もサポートしています。このモデルでは、複数のノードで Microsoft フェールオーバー・クラスターを構成し、ノードを遠隔拠点間に分散するオプションを設定します。この機能では、複数の構成が可能となり、障害がローカル、地域、国全体で発生した場合でも、ネットワークを確実に保護できます。ハイブリッド構成の例を図 4 に示します。ここでは、高可用性(HA)用のローカル・ノードを持つサーバー1 台をニューヨークに置き、地域ネットワークの停止に備えて、シカゴに別のノードを置き、さらに、国全体のネットワークが停止した場合に備えて、4 つ目のノードをロンドンに置いています。

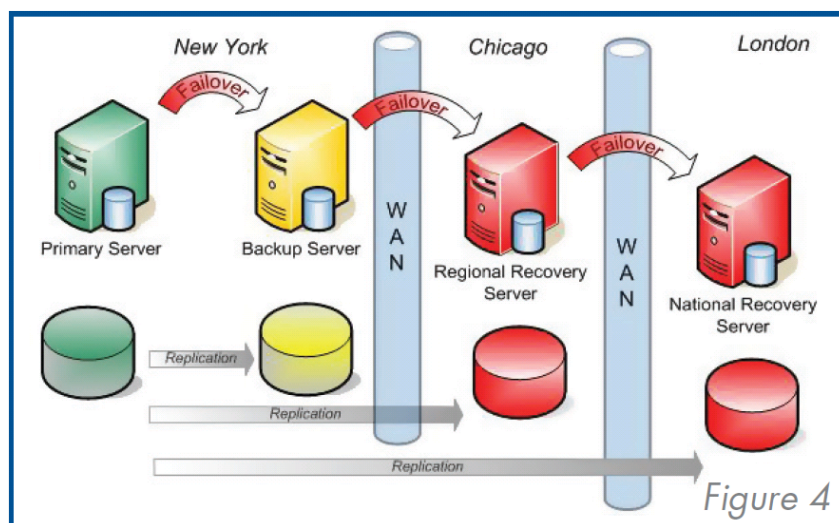


図 4

ハイブリッド・モデルでは、アクティブなノードに障害が発生すると、次に利用可能なノードにサービスを引き継ぎます。リカバリーは、アプリケーションが完全にリストアされるまで、バックアップ・サーバー、リージョナル・リカバリー・サーバー、ナショナル・リカバリー・サーバーへと、段階的に実行されます。サービスを提供するノードに関係なく、サービス開始状態となったノードがミラー・ソースとなり、障害から復旧したノードに対して、変更済みデータを複製します。

## 結論

マイクロソフトは、Windows Server 2008 により Windows Server フェールオーバー・クラスタリングを柔軟に構築できる強力なツールを提供しました。遠隔拠点間クラスタリングのフレームワークを提供して、アプリケーションの可用性と事業継続性を大幅に向上しました。SteelEye DataKeeper は、使用性と効率に優れた高速データ・レプリケーション・エンジンとして、レプリケーションの同期/非同期オプション、継続的データ保護(CDP)、複数ターゲットのサポート、データ圧縮技術など、多彩な機能を提供し、マイクロソフトのクラスター・モデルを拡張しています。

## SteelEye Technology 社について

SteelEye はデータとアプリケーションの可用性管理ソリューションを提供するリーディング企業であり、Linux、Windows、仮想環境を対象として、事業継続性とディザスタリカバリーをサポートしています。

データ・レプリケーション、HA クラスティング、障害回復ソリューションから成る SteelEye 製品ファミリーは、あらゆる規模の企業を対象として、ミッション・クリティカルなアプリケーション、サーバー、データの継続的な可用性を保証しています。

SteelEye は、ソフトウェア・ソリューションを補完するために、専門家によるコンサルティングなど、多彩なサービスを提供して、皆様を支援しています。高可用性を保証するエンタープライズ・ネットワークの評価、設計、実装につきましては、ぜひ、当社にご相談ください。

SteelEye は、SIOS Technology, Incの子会社です。当社へのお問い合わせは、[www.steeleye.com](http://www.steeleye.com) にアクセスするか、次のアドレスにメールをお送りください。  
米国/カナダ 866.318.0108 欧州 +44 (0)1223 208701 その他の地域 +1.650.843.0655

SteelEye Technology, Inc. 4400 Bohannon Drive, Suite 110, Menlo Park, CA 94025