



## CHAPTER 3

# Subscriber Manager フェールオーバー

## 概要

この章では、Subscriber Manager (SM) をクラスタと冗長性ととも使用する方法について説明します。

1 次サービス プロバイダー環境に配置された Cisco Service Control Application for Broadband (SCA BB) ソリューションで重要な役割を担うことに加え、SM バージョン 2.2 以降はフェールオーバー動作モードもサポートします。この機能により、SM 障害 (詳細については「[Subscriber Manager フェールオーバーに関する情報](#)」(P.3-1) を参照してください) によって引き起こされたシステム ダウンタイムが最小化されます。

この項では、2 つの SM ノードから構成されるクラスタをフェールオーバー動作モードで使用する場合はさまざまなコンセプトについて説明します。



(注) ここでは、Veritas Cluster テクノロジーに関する知識があることを前提としています。

- 「[Subscriber Manager フェールオーバーに関する情報](#)」(P.3-1)
- 「[フェールオーバーからの回復方法](#)」(P.3-5)

## Subscriber Manager フェールオーバーに関する情報

- 「[概要](#)」(P.3-1)
- 「[通常の動作](#)」(P.3-2)
- 「[フェールオーバー トポロジ](#)」(P.3-3)
- 「[フェールオーバーの動作](#)」(P.3-4)

## 概要

SM に実装されたフェールオーバー スキームは、Veritas Cluster テクノロジーに基づいています。クラスタには 2 つのマシンが含まれ、各マシンでは SM TimesTen と Veritas ソフトウェアが実行されます。Veritas Cluster Server (VCS) ソフトウェアは、クラスタ全体に単一の仮想 IP アドレスを提供することにより複数の SM を統合し、単一エンティティを提供します。

このクラスタ ソフトウェアはアクティブ マシンとスタンバイ マシンを区別します。アクティブ マシンは仮想 IP アドレスとすべてのネットワーク接続を「所有」し、スタンバイ マシンはフェールオーバーが発生するまでパッシブの状態になります。フェールオーバーが発生すると、障害が発生したサーバからバックアップ サーバに IP アドレスが渡され、バックアップ サーバがアクティブになり、すべてのネットワーク接続が再び確立されます。

フェールオーバーの発生時に、Login Event Generator (LEG) は障害が発生した SM との接続を失い、アクティブになった (バックアップ) SM に再接続し、コミットされていないメッセージを再送信します。アクティブになった SM は Service Control Engine (SCE) プラットフォームに接続し、SCE と再同期します。

TimesTen データベース レプリケーション エージェントは、アクティブ ノードからスタンバイ ノードに SM データベースを継続的に複製します。アクティブになったマシンの加入者データは常に有効であるため、ある SM から別の SM へのフェールオーバーは高速になります。この 2 つの SM ノードは加入者データを受け渡す以外に通信しません。

VCS は「クラスタ エージェント」と呼ばれるソフトウェア コンポーネントを使用して Network Interface Card (NIC; ネットワーク インターフェイス カード)、ディスク、IP アドレス、プロセスなどのリソースのステータスをモニタおよび制御します。シスコは、SM および TimesTen データベース デーモンをモニタするクラスタ エージェントとレプリケーション エージェントを提供します。

クラスタ動作の一部として、TimesTen データベース デーモンとレプリケーション エージェントはフェールオーバーの状態に関係なく稼動したままになります。SM Veritas エージェントは、TimesTen データベース デーモンとレプリケーション エージェント プロセスをモニタします。これらのいずれかで障害が発生すると、フェールオーバーが行われます。



(注)

アクティブ マシンとスタンバイ マシンの SM ソフトウェアの設定は同一である必要があります。両方のマシンに同じコンフィギュレーション ファイルを適用してください。

次に、これらのコンセプトについて詳しく説明します。

## 通常の動作

2 つの SM ノードはホットスタンバイ モードで動作します。この場合は常に、一方のノード (アクティブ ノード) がすべての SM イベントを受信および処理し、もう一方のノード (スタンバイ ノード) は待機し、フェールオーバー発生時に動作できます。シームレスなフェールオーバーを実現し、フェールオーバー時間を最小化するために、2 つの SM ノードは外部ストレージ デバイスなしで動作します。

クラスタの通常の動作時に、アクティブ ノード (クラスタで選択されたノード) は次のことを行います。

- 非クラスタ環境のすべての SM 機能を実行します。
- クラスタ エージェントの「状態」に関する情報を提供します。
- 定期的に加算者データベースをスタンバイ ノードに複製します。

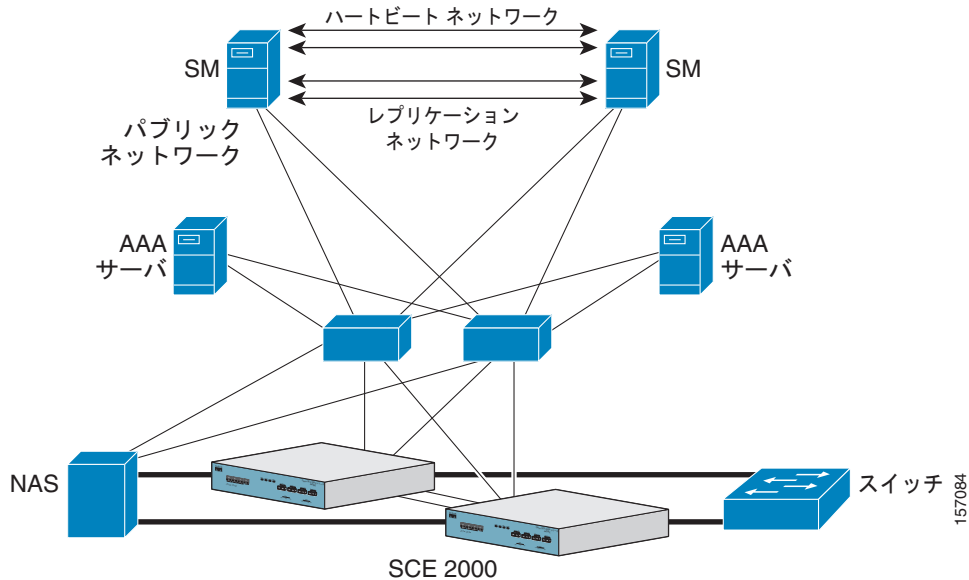
スタンバイ ノードでは、SM と TimesTen ソフトウェアの両方が実行されています。

- SM が完全に設定されます (アクティブ ノードと同じコンフィギュレーション ファイルが適用されますが、アクティブ ノードの作業は干渉されません)。
- SM は TimesTen データベースに接続しますが、LEG デバイスと SCE デバイスには接続しません。
- TimesTen ソフトウェアは加入者データベースのレプリケーション クライアントとして動作し、アクティブ ノードの TimesTen ソフトウェアからアップデートを受信して適用します。

## フェールオーバー トポロジ

図 3-1 は、1つの冗長 AAA サーバと、冗長性を実現するためにカスケード接続された2つの SCE 2000 プラットフォームのトポロジの SM クラスタ構成を示しています。

図 3-1 フェールオーバー トポロジの SM クラスタ構成



すでに説明したように SM フェールオーバー トポロジには、クラスタ スキームで接続された2つの SM ノードが含まれます。

2つのノードは、次の2つの専用（プライベート）の冗長ネットワークによって相互接続されます。

- ハートビート ネットワーク：クラスタのモニタと制御を実行するために Veritas Cluster Server によって使用されます。
- レプリケーション ネットワーク：加入者レコードを渡すためにレプリケーション プロセスによって使用されます。

2つのノードは、2つのノード間のバックツーバック接続または冗長スイッチを使用してハートビートが実装された同じサイトに存在する必要があります。クラスタ内の各ノードには、SM が通信するすべての外部エンティティ（AAA、LEG、SCE）に接続する冗長なネットワーク パス（NIC）が存在します。

クラスタ内の各ノードには、最低6つのイーサネット NIC が存在します。6つのイーサネット NIC は次のように使用されます。

- 2つの NIC が（プライベートの）ハートビート ネットワーク用に使用されます。
- 2つの NIC が（プライベートの）レプリケーション ネットワーク用に使用されます。
- 2つの NIC がパブリック ネットワーク用（SCE および LEG への接続と SM の管理）に使用されます。

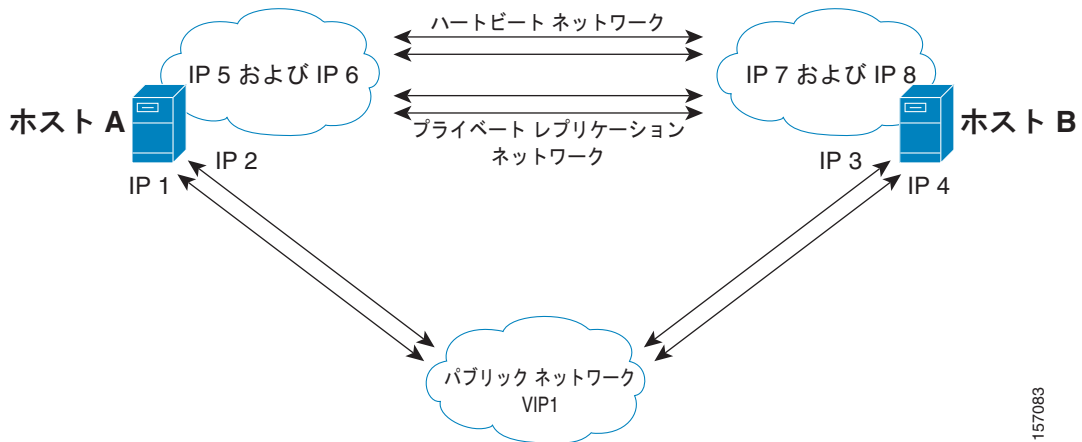
クラスタは、外部エンティティとの接続に使用される Virtual IP (VIP; 仮想 IP) を持ちます。クラスタ内の各ノードもノード/クラスタの管理用の IP アドレスとレプリケーション用の IP アドレスを持ちます。

パブリック ネットワークのプライマリ NIC で障害が発生すると、同じノードのセカンダリ NIC にフェールオーバーされます。同じ IP アドレス (VIP1) が保持され、クラスタのフェールオーバーは行われません。レプリケーション ネットワークまたはハートビート ネットワークのプライマリ NIC で障害が発生すると、同じノードのセカンダリ NIC にフェールオーバーされます。同じ IP アドレス (VIP2 および VIP3) が保持され、クラスタのフェールオーバーは行われません。

図 3-2 は、クラスタ構成で 사용되는通常の IP アドレスと仮想 IP アドレスの使用例を示しています。

- ノードの管理では IP1/IP2 と IP3/IP4 がそれぞれ使用されます。
- パブリック ネットワークを介して接続された外部クライアント用クラスタ IP アドレスは VIP1 を使用します。

図 3-2 クラスタ構成における通常の IP と仮想 IP



レプリケーション IP 構成の詳細については、「Veritas Cluster Server」(P.E-1) を参照してください。

## フェールオーバーの動作

通常の動作時に Veritas Cluster Server メカニズムによってアクティブになる SM サーバとスタンバイ状態になる他の SM サーバが自動的に選択されます。

アクティブ SM サーバは、通常のすべての SM 機能を実行します。これら 2 つのサーバは相互間でハートビートメカニズムを保持し、アクティブサーバは継続的に加入者データベースをスタンバイサーバのデータベースに複製します。

スタンバイ SM サーバはホットスタンバイマシンとして動作するため、最小のフェールオーバー時間で引き継ぐ (アクティブになる) ことができます。

フェールオーバーメカニズムは次のような障害によって作動します。

- SM アプリケーションの障害 (TimesTen データベースの障害など)
- TimesTen レプリケーションプロセスの TimesTen デーモンの障害
- SUN サーバの障害 (両方のパブリック ネットワーク NIC の障害など、サーバのいずれかのリソースの障害が原因)
- 手動によるフェールオーバーのアクティブ化



(注)

冗長な NIC が存在する場合は、通信障害によってフェールオーバーは発生しません。したがって、各 SUN マシンには外部デバイスに接続する 2 つの NIC が存在するため、いずれかの NIC で障害が発生した場合、冗長な NIC に切り替わるだけでフェールオーバー メカニズムは作動しません。

障害が検出されたら、スタンバイ SM はアクティブになり、次のことが起こります。

- アクティブになった SM が仮想 IP メカニズムの IP リソースを引き継ぎます。
- LEG がアクティブになった SM に再接続します。
- アクティブになった SM が SCE との IP 接続を確立し、SCE と再び同期します。
- アクティブになった SM が、さまざまな LEG から送信された情報の処理を開始し、情報を SCE に転送します。

## フェールオーバーからの回復方法

障害の種類によって、回復作業も異なります。ノード間ポートリンクの障害など、一部の障害は自動的に回復しますが（リンクが回復したときに自動的に回復する）、他の障害は手動で回復する必要があります。

障害が発生した SM が自己回復した場合、または SM が置き換えられた場合（必要な場合）、回復が行われます。回復作業の目的は、クラスタを完全に機能するモードに戻すことです。回復作業が終了すると、動作がインストール後と同じになります。

障害が発生した SM サーバは、発生した障害の種類に応じて手動で回復するか、または自動的に回復されます。回復作業と、それぞれの回復作業をいつ行うかについては、次の項で説明します。

- 「マシンのリブート」(P.3-5)
- 「サーバの置き換え」(P.3-6)
- 「データベース複製の回復」(P.3-6)
- 「データベース複製回復の管理」(P.3-7)

## マシンのリブート

マシン リブートからの回復は完全に自動的な回復プロセスです。障害が発生した SM サーバはリブートされ、他のサーバとの接続を確立し、データベースを同期した後に 2 つの SM サーバのクラスタで再びフェールオーバーを行えるようになります。



(注)

次の作業の手順は自動的に行われます。

- ステップ 1** ノードでリブート プロセスが実行されます。
- ステップ 2** VCS によりノードがスタンバイ状態になります。
- ステップ 3** ノードがブートします。
- ステップ 4** VCS がノード間通信を確立し、新しいノードがクラスタに参加します。
- ステップ 5** TimesTen データベースのレプリケーション プロセスが、リブート前の時点から開始されます。

回復されたサーバの SM は、データベース回復プロセスを実行した後の使用可能な状態になり、SM が初期状態からスタンバイ状態に移行します。

## サーバの置き換え

サーバの置き換えは、回復不可能な物理的な障害がマシンで発生した場合に必要です。サーバは、新しい SM、TimesTen、および VCS がインストールされた新しいマシンによって置き換えられます。

サーバの置き換えは手動による回復であり、障害が発生した SM サーバは物理的に置き換えられます。新しい SM サーバをネットワークに接続し、サーバを設定し、2 つのデータベースを同期した後に、2 つの SM サーバのクラスターで再びフェールオーバーを行えるようになります。

**ステップ 1** 新しいサーバをノード間ポートとノード内ポートに接続します（ただし、ネットワーク ポートは未接続のままにします）。

**ステップ 2** ネットワークとクラスターの基本的な設定を手動で行います（初回）。

**ステップ 3** アクティブなノードからコンフィギュレーション ファイルをコピーします。

**p3sm.cfg** ファイルだけをコピーする必要がある場合は、次の Command Line Utility (CLU; コマンドラインユーティリティ) コマンドを使用します。

```
p3sm --load-config --remote=NEW-SM_IP
```

**ステップ 4** TimesTen データベースの複製作業を実行します。

「データベース複製の回復」(P.3-6) を参照してください。

**ステップ 5** 回復されたノードで VCS を起動します。

**ステップ 6** ネットワーク ポートに接続します。

回復されたサーバの SM は、データベース回復プロセスの完了後に使用可能な状態になり、SM が初期状態からスタンバイ状態に移行します。

## データベース複製の回復

データベース複製の回復は手動による回復であり、スタンバイ ノードのデータベースがアクティブ ノードのデータベースとの同期を失った場合に必要です。同期は、SM マシンのいずれかが置き換えられた場合や、アクティブ ノードのレプリケーション プロセスが、データベースに挿入されたすべてのデータの複製に失敗した場合（レプリケーション NIC が未接続だった場合）に失われることがあります。

**ステップ 1** Cluster Server (VCS) によるリソースのモニタを停止します。

VCS CLU **hastop -local** コマンドを使用して VCS を停止します。

**ステップ 2** データベースの消去によって影響を受けないよう SM を停止します。

CLU コマンド **p3sm --stop** を使用します。

**ステップ 3** レプリケーション エージェントを停止します。

CLU コマンド **p3db --rep-stop** を使用します。

- ステップ 4** データベースを破棄します。  
CLU コマンド `p3db --destroy-rep-db` を使用します。
- ステップ 5** リモート データベースをローカル マシンに複製します。  
CLU コマンド `p3db --duplicate` を使用します。
- ステップ 6** Cluster Server によるリソースのモニタを開始します。  
VCS CLU `hastart` コマンドを使用します。レプリケーション プロセスと SM が自動的に開始されま  
す。

## データベース複製回復の管理

2 つの SM サーバは、コマンドラインユーティリティ (CLU) とコンフィギュレーション ファイルを使用して設定されます (「設定と管理」および「加入者管理ソリューションの設定方法」を参照)。実際の設定はアクティブ SM に対して実行され、スタンバイ SM に対しては手動で複製します。

- ステップ 1** アクティブ マシンとスタンバイ マシンとの間で FTP 接続を確立します。
- ステップ 2** コンフィギュレーション ファイルをコピーします。  
~`pcube/sm/server/root/config/` フォルダのすべてのコンフィギュレーション ファイルをアクティブ ノードからスタンバイ ノードにコピーし、CLU コマンド `p3sm --load-config` を使用して SM コンフィギュレーション ファイルを適用します。
- ステップ 3** データベース関連のコンフィギュレーション ファイルを必要な場所に手動でコピーします。  
データベース関連のコンフィギュレーション ファイルで変更を行った場合は、コンフィギュレーション ファイルを `/etc/system` (Solaris の場合) または `/etc/sysctl.conf` (Linux の場合) にコピーし、`/var/TimesTen/sys.odbc.ini` をアクティブ ノードからスタンバイ ノードにコピーします。



(注) この手順を実行した場合は、スタンバイ ノードのリポートが必要です。



(注) データベースが 2 つのノードの異なるディレクトリに存在する場合は、両方のノードのファイル `sys.odbc.ini` が同一ではないため、このファイルで変更された実際のパラメータをコピーする必要があります。

- ステップ 4** Veritas ツールを使用して Veritas Cluster Server を設定および管理します。  
Veritas Cluster Server が提供する SNMP トラップによって通知が有効になります。Veritas Cluster Server は次のように SNMP トラップをサポートします。
- 致命的な障害を検出 (ローカルまたはリモート)
  - セカンダリ ノードがフェールオーバー作業を開始
  - セカンダリ ノードが稼動 (フェールオーバーの終了)

