



CHAPTER 2

サービスレベルハイ アベイラビリティの理解

この章では、サービスレベルのハイ アベイラビリティを実現する Cisco NX-OS サービスの再起動性について説明します。

この章の構成は次のとおりです。

- 「Cisco NX-OS サービスの再起動について」(P.2-1)
- 「再起動性インフラストラクチャ」(P.2-1)
- 「プロセスの再起動性」(P.2-3)
- 「スタンバイ スーパーバイザ サービスの再起動」(P.2-5)
- 「スイッチング モジュール サービスの再起動」(P.2-5)
- 「再起動のトラブルシューティング」(P.2-5)
- 「参考文献」(P.2-6)

Cisco NX-OS サービスの再起動について

Cisco NX-OS のサービス再起動機能では、スーパーバイザを再起動せずに障害の発生したサービスを再起動することによって、プロセスレベルの障害がシステムレベルの障害に拡大するのを防ぐことができます。サービスは、現在のエラー、障害状況、サービスのハイ アベイラビリティ ポリシーに基づいて再起動できます。サービスの再起動には、ステートフルな再起動とステートレスな再起動があります。Cisco NX-OS では、サービスが実行時の状態情報とメッセージを保存することで、ステートフルな再起動を実現しています。ステートフルな再起動では、サービスが保存されていた状態情報を取り出して、直前のチェックポイント サービス状態から動作を再開します。ステートレスな再起動では、サービスは、初めて起動するときのように、初期化および実行されます。

再起動性インフラストラクチャ

Cisco NX-OS では、大部分のプロセスおよびサービスのステートフルな再起動が可能です。プラットフォーム内のプロセス、サービス、アプリケーションのバックエンドでの管理および調整は、このセクションで説明する一連の高レベルのシステム コントロール サービスによって実行されます。

ここでは、次の内容について説明します。

- 「システム マネージャ」(P.2-2)
- 「永続ストレージ サービス」(P.2-2)

- 「メッセージおよびトランザクション サービス」 (P.2-2)
- 「HA ポリシー」 (P.2-2)

システム マネージャ

システム マネージャは、あらゆるシステム機能、システム管理、システム ヘルス モニタリングの実行を制御し、ハイ アベイラビリティ ポリシーを実施します。システム マネージャは、サービスの起動、停止、モニタリング、再起動を担当し、サービス状態とスーパーバイザ状態の同期を開始および管理してステートフル スイッチオーバーを実現します。

永続ストレージ サービス

Cisco NX-OS サービスは、Persistent Storage Service (PSS; 永続ストレージ サービス) を使用して、運用の実行時情報とプラットフォーム サービスの設定を保存および管理します。PSS コンポーネントは、システム サービスを使用して、サービス再起動時に状態を回復します。PSS は状態および実行時情報のデータベースとして機能します。これにより、各サービスは、必要なときにいつでも、サービス自体の状態情報のチェックポイントを作成できます。サービスを再起動すると、障害が発生する直前の既知の動作状態を回復できるので、ステートフルな再起動が可能になります。

PSS を使用する各サービスは、保存された情報をプライベート情報（当サービスだけ読み取り可能）または共有情報（他のサービスも読み取り可能）として定義できます。情報を共有する場合は、ローカル（同一スーパーバイザ上のサービスだけ読み取り可能）またはグローバル（スーパーバイザまたはモジュール上のサービスが読み取り可能）のどちらかを指定できます。

メッセージおよびトランザクション サービス

Message and Transaction Service (MTS; メッセージおよびトランザクション サービス) は、ハイ アベイラビリティ セマンティクスに特化した高パフォーマンス Interprocess Communication (IPC; プロセス間通信) メッセージ ブロカーです。MTS は、モジュール内とモジュール間、およびスーパーバイザ間でメッセージのルーティングとキューイングを行います。また、イベント通知や同期などのメッセージ交換を容易にし、システム サービス間およびシステム コンポーネント間のメッセージ永続性を促進します。MTS では、永続メッセージおよびログ メッセージをキュー内に保管できるため、サービスの再起動後もそれらのメッセージにアクセスできます。

HA ポリシー

Cisco NX-OS では、各サービスに、障害の発生したサービスの再起動方法を定義する一連の内部 HA ポリシーのセットを作成できます。サービスごとに 4 つの定義済みポリシーを用意できます。つまり、スーパーバイザが 2 つの場合のプライマリ ポリシーとセカンダリ ポリシー、スーパーバイザが 1 つだけの場合のプライマリ ポリシーとセカンダリ ポリシーです。HA ポリシーが定義されていないサービスでは、サービスの障害発生時に実行されるデフォルトの HA ポリシーは、スーパーバイザが 2 つの場合はスイッチオーバー、スーパーバイザが 1 つの場合はスーパーバイザのリセットとなります。

HA ポリシーには、次の 3 つのパラメータを指定します。

- システム マネージャによって実行されるアクション：
 - ステートフルな再起動
 - ステートレスな再起動

- スーパーバイザのスイッチオーバー（または再起動）
- 最大再試行回数：システム マネージャによって実行される再起動試行回数を指定します。再試行をその回数行ってもサービスが正常に再起動しない場合、その HA ポリシーは失敗したものと見なされ、定義されている次の HA ポリシーが使用されます。他の HA ポリシーが定義されていない場合はデフォルトのポリシーが適用されます。つまり、スーパーバイザのスイッチオーバーまたは再起動が実行されます。
- 最小ライフタイム：再起動の試行のあとにサービスを実行する時間を指定します（再起動が正常に行われたと見なします）。最小ライフタイムは最低でも 4 分です。

プロセスの再起動性

プロセスの再起動性により、データ プレーンやその他のサービスを中断せずに、障害の発生したサービスを回復し動作を再開することができます。システム マネージャは、サービスの HA ポリシー、前回の再起動の失敗、同じスーパーバイザ上で実行されているその他のサービスのヘルス状態に応じて、サービスの障害発生時に実行するアクションを決定します。

表 2-1 は、さまざまな障害発生時にシステム マネージャが実行するアクションを示しています。

表 2-1 障害発生時のシステム マネージャのアクション

障害	アクション
サービス/プロセスの例外	サービスの再起動
サービス/プロセスのクラッシュ	サービスの再起動
サービス/プロセスの応答がない	サービスの再起動
サービスの障害が繰り返される	スーパーバイザのリセット（シングル スーパーバイザの場合）またはスイッチオーバー（デュアル スーパーバイザの場合）
システム マネージャからの応答がない	スーパーバイザのリセット（シングル スーパーバイザの場合）またはスイッチオーバー（デュアル スーパーバイザの場合）
カーネル障害	スーパーバイザのリセット（シングル スーパーバイザの場合）またはスイッチオーバー（デュアル スーパーバイザの場合）
ウォッチドッグ タイムアウト	スーパーバイザのリセット（シングル スーパーバイザの場合）またはスイッチオーバー（デュアル スーパーバイザの場合）

障害の発生したサービスは、HA の実装および HA ポリシーに応じて、ここで説明するいずれかの方法で再起動されます。

ここでは、次の内容について説明します。

- 「ステートフルな再起動」(P.2-4)
- 「ステートレスな再起動」(P.2-4)
- 「スイッチオーバー」(P.2-5)

ステートフルな再起動

再起動可能なサービスで障害が発生すると、サービスは同じスーパーバイザ上で再起動されます。サービスの新しいインスタンスは、以前のインスタンスがオペレーティング システムによって異常終了させられたと判断した場合、永続コンテキストがあるかどうかを確認します。新しいインスタンスは初期化時に永続コンテキストを読み込んで、実行時コンテキストを構築します。この結果、新しいインスタンスは障害発生前のインスタンスと同じ状態になります。初期化が完了すると、サービスは、停止したときに実行していたタスクを再開します。新しいインスタンスが再起動および初期化されている間、他のサービスは、そのような障害が発生していることを認識していません。他のサービスから障害が発生したサービスに送信されたメッセージは、サービスが再開された時点で MTS から取得できます。

新しいインスタンスでステートフルな初期化を完了できるかどうかは、前のインスタンスの障害の原因に依存します。サービスで再起動を数回実行できない場合、そのサービスの再起動は失敗したと見なされます。その場合、システム マネージャは、再起動に失敗したサービスの HA ポリシーに指定されたアクション（ステートレスな再起動、再起動しない、スーパーバイザのスイッチオーバーまたはリセットのいずれか）を実行します。

ステートフルな再起動に成功した場合、システムが矛盾のない状態に到達するまでに遅延が発生することはありません。ステートフルな再起動により、障害発生後の回復に要する時間が短縮されます。

ステートフルな再起動の前後および最中に発生するイベントは次のとおりです。

1. 実行中のサービスが、実行時状態情報のチェックポイントを PSS に作成します。
2. システム マネージャがハートビートを使用している実行中サービスのヘルス状態を監視します。
3. システム マネージャが、クラッシュまたはハングしたサービスを即座に再起動します。
4. 再起動のあとに、サービスは、PSS から状態情報を回復し、保留中のすべてのトランザクションを再開します。
5. 何度か再起動してもサービスの動作が安定しない場合、システム マネージャはスーパーバイザのリセットまたはスイッチオーバーを開始します。
6. Cisco NX-OS はプロセス スタックとコアをデバッグ用に収集します。また、オプションでコア ファイルをリモートに転送します。

ステートフルな再起動が行われると、Cisco NX-OS がレベル LOG_ERR の Syslog メッセージを送信します。SNMP トラップがイネーブルになっている場合は、SNMP エージェントがトラップを送信します。

ステートレスな再起動

Cisco NX-OS インフラストラクチャ コンポーネントは、ステートレスな再起動を管理します。ステートレスな再起動中、システム マネージャは、障害の発生したプロセスを特定し、新しいプロセスに置き換えます。障害の発生したサービスは再起動時に実行時状態を保持していないため、実行コンフィギュレーションから実行時状態を構築するか、必要な場合は、他のサービスと情報を交換して実行時状態を構築します。

ステートレスな再起動が行われると、Cisco NX-OS がレベル LOG_ERR の Syslog メッセージを送信します。SNMP トラップがイネーブルになっている場合は、SNMP エージェントがトラップを送信します。

スイッチオーバー

スタンバイ スーパーバイザが使用可能な場合で、複数の障害が同時に発生したとき、Cisco NX-OS は常にスーパーバイザの再起動ではなくスーパーバイザのスイッチオーバーを実行します。こうしたケースは、同一スーパーバイザ上では回復不可能と見なされるからです。たとえば、複数の HA アプリケーションで障害が発生すると、回復不可能と見なされます。

デュアル VSM を持つシステムでは、スイッチオーバー後にアクティブ スーパーバイザがリセットされ、スタンバイ スーパーバイザとして復帰します。

スーパーバイザのスイッチオーバーおよび再起動の詳細については、第3章「システムレベルハイ アベイラビリティの設定」を参照してください。

スタンバイ スーパーバイザ サービスの再起動

スタンバイ状態のスーパーバイザ上のサービスで障害が発生した場合、システム マネージャは HA ポリシーを適用せず、30 秒待ってからサービスを再起動します。30 秒待つことで、スタンバイ サービスの障害と同期化が繰り返されたときにアクティブ スーパーバイザが対応しきれなくなるのを避けることができます。再起動されるサービスをアクティブなスーパーバイザ上のサービスと同期させる必要がある場合、スタンバイ スーパーバイザは、当該サービスの再起動と同期化が完了するまでホットスタンバイ モードではなくなります。サービスが再起動不可能な場合は、スタンバイ スーパーバイザがリセットされます。

スタンバイ サービスの再起動が行われると、Cisco NX-OS はレベル LOG_ERR の Syslog メッセージを送信します。SNMP トラップがイネーブルになっている場合は、SNMP エージェントがトラップを送信します。

スイッチング モジュール サービスの再起動

非スーパーバイザ モジュール サービスでサービス障害が発生した場合、スーパーバイザ スwitchオーバーは必要ありません。

VEM で DPA が再起動されます (クラッシュした場合)。この状況では、モジュールが削除され VSM に再び追加されます。

再起動のトラブルシューティング

サービスで障害が発生すると、システムは障害の原因を判定するために使用できる情報を生成します。次の情報ソースが使用可能です。

- サービスの再起動によって、LOG_ERR レベルの Syslog メッセージが生成されます。
- SNMP トラップがイネーブルになっている場合は、サービスが再起動されると、SNMP エージェントがトラップを送信します。
- VSM でサービス障害が発生すると、イベントがログに記録されます。ログを参照するには、そのモジュールで **show processes log** コマンドを使用します。プロセスのログは、スーパーバイザのスイッチオーバーまたはリセット後も保持されます。
- サービスの障害が発生すると、システムのコア イメージ ファイルが生成されます。最新のコア イメージを表示するには、アクティブなスーパーバイザ上で **show cores** コマンドを入力します。コア ファイルはスーパーバイザのスイッチオーバーまたはリセット後に保持されませんが、Trivial

File Transfer Protocol (TFTP; 簡易ファイル転送プロトコル) などのファイル転送ユーティリティを使用してコア ファイルを外部サーバにエクスポートするようにシステムを設定することができます。

サービスの障害に関する生成情報の収集および使用については、『Cisco Nexus 1000V Troubleshooting Guide, Release 4.0(4)SV1(3)』を参照してください。

参考文献

サービスレベルの HA 機能の実装に関する詳細は、次の各セクションを参照してください。

- 「関連資料」(P.2-6)
- 「標準」(P.2-6)
- 「MIB」(P.2-6)
- 「RFC」(P.2-7)
- 「テクニカル サポート」(P.2-7)

関連資料

関連トピック	マニュアル名
スーパーバイザ スイッチオーバー	第 3 章「システムレベル ハイ アベイラビリティの設定」
トラブルシューティング	『Cisco Nexus 1000V Troubleshooting Guide, Release 4.0(4)SV1(3)』
Cisco NX-OS の基礎	『Cisco Nexus 1000V Getting Started Guide, Release 4.0(4)SV1(3)』

標準

標準	タイトル
この機能によってサポートされる新規の標準または変更された標準はありません。また、サポートされている既存の標準は、この機能でもサポートされます。	—

MIB

MIB	MIB 関連のリンク
CISCO-PROCESS-MIB	MIB を検索およびダウンロードするには、次の URL にアクセスしてください。 http://www.cisco.com/public/sw-center/netmgmt/cmtk/mibs.shtml

RFC

RFC	タイトル
この機能によってサポートされている RFC はありません。	—

テクニカル サポート

説明	リンク
Technical Assistance Center (TAC) のホームページには、製品、テクノロジー、ソリューション、技術的なヒント、ツールへのリンクを含め、30,000 ページに及ぶ検索可能な技術コンテンツが含まれています。 Cisco.com の登録済みユーザは、このページからログインして、さらに広範なコンテンツにアクセスできます。	http://www.cisco.com/public/support/tac/home.shtml

