



CHAPTER 1

概要

この章では、ハイ アベイラビリティ (HA) の概念と Cisco NX-OS ソフトウェアの機能について説明します。この章の構成は次のとおりです。

- 「ハイ アベイラビリティについて」 (P.1-1)
- 「サービスレベル ハイ アベイラビリティ」 (P.1-2)
- 「システムレベルのハイ アベイラビリティ」 (P.1-3)
- 「ネットワークレベル ハイ アベイラビリティ」 (P.1-3)
- 「VSM から VSM へのハートビート」 (P.1-4)
- 「参考資料」 (P.1-6)

ハイ アベイラビリティについて

ハイ アベイラビリティ (HA) の目的は、システム内で発生したハードウェアの障害およびソフトウェアのエラーの両方による影響を一定範囲内に抑えることです。Cisco NX-OS オペレーティング システムは、ネットワークレベル、システムレベル、およびサービスレベルでのハイ アベイラビリティを実現するよう設計されています。

次の Cisco NX-OS の機能により、障害発生時にトラフィックの中断が最小化または防止されます。

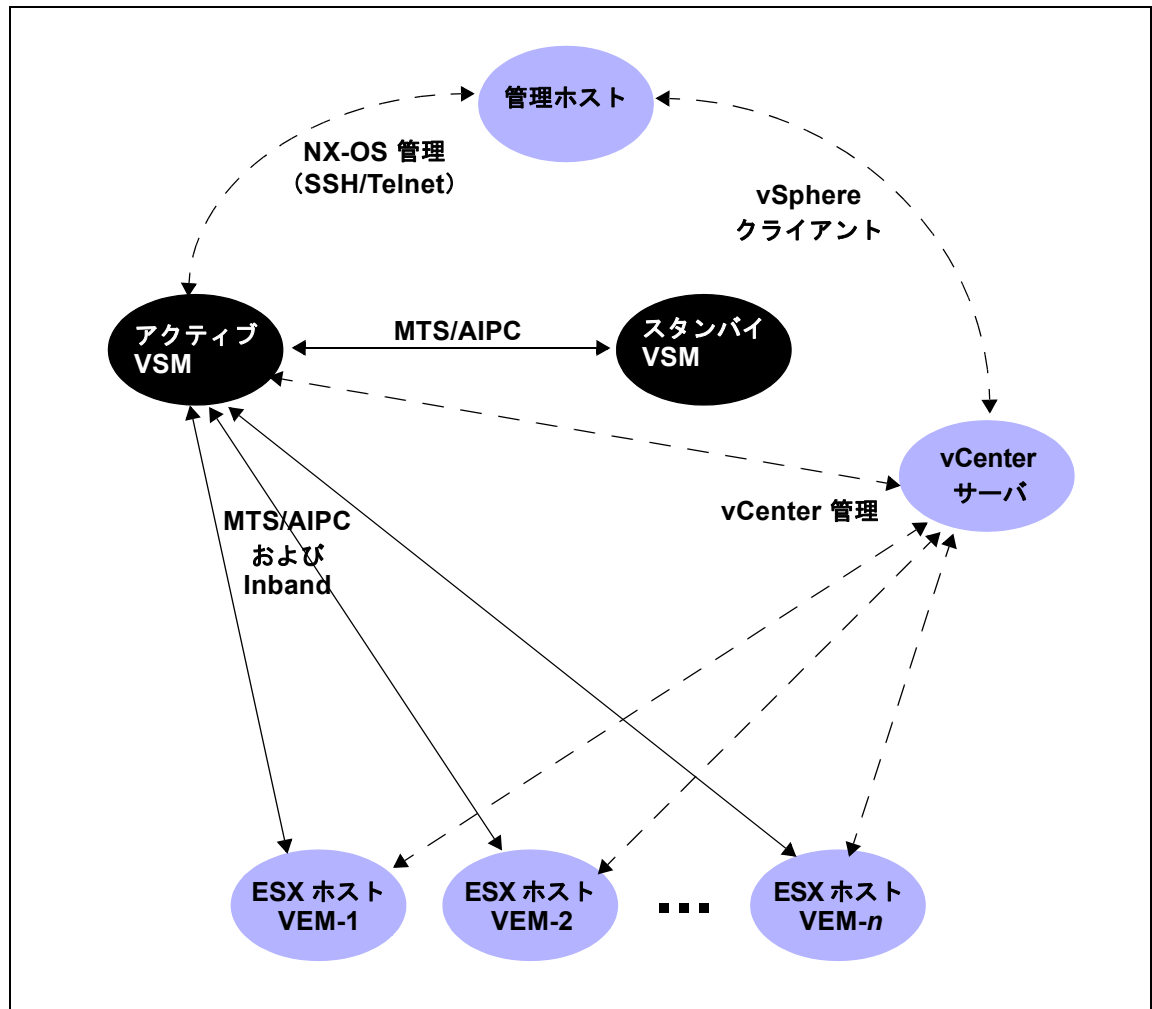
- 冗長性：ソフトウェア アーキテクチャのあらゆる面での冗長性。
- プロセスの分離：1 つのプロセス内で発生したエラーで他のプロセスが中断されるのを防ぐためのソフトウェア コンポーネント間の分離。
- 再起動性：ほとんどのシステム機能およびサービスが分離されているため、エラーが発生しても、他のサービスは実行され続けている中で独立して再起動が可能。さらに、ほとんどのシステム サービスはステートフルな再起動を実行するため、その他のサービスに対して透過的に稼働を再開できます。
- スーパーバイザ ステートフル スイッチオーバー：アクティブ/スタンバイ デュアル スーパーバイザ設定。状態および設定は、障害の発生時にシームレスかつステートフルなスイッチオーバーを提供するために 2 つの Virtual Supervisor Module (VSM) 間で常に同期されます。

Cisco Nexus 1000V システムは、次のコンポーネントで構成されます。

- 仮想マシン (VM) 内で実行される 1 つまたは 2 つの VSM。
- 仮想サーバ内で実行される Virtual Ethernet Modules (VEM)。VEM は VSM 内でモジュールとして表されます。
- リモート管理コンポーネント。たとえば、VMware vCenter Server はリモート管理コンポーネントです。

図 1-1 は HA コンポーネントと、HA コンポーネント間の通信リンクを示しています。

図 1-1 HA コンポーネントと通信リンク



サービスレベル ハイ アベイラビリティ

Cisco NX-OS ソフトウェアは、プロセスをコンパートメント化して障害の分離、冗長性、および効率性を実現します。

ここでは、次の内容について説明します。

- 「プロセスの分離」(P.1-3)
- 「プロセスの再起動性」(P.1-3)

サービスレベル HA の詳細については、第 2 章「サービスレベル ハイ アベイラビリティの理解」を参照してください。

プロセスの分離

Cisco NX-OS ソフトウェアでは、サービスと呼ばれる独立したプロセスが、サブシステムまたはフィーチャセットの機能またはフィーチャセットを実行します。各サービスおよびサービス インスタンスは、独立した保護プロセスとして実行されます。このアプローチにより、高いフォールトトレラントを備えたソフトウェア インフラストラクチャとサービス間での障害の分離を実現できます。サービス インスタンスの障害は、障害時に実行されている他のサービスに影響しません。また、サービスの各インスタンスは独立したプロセスとして実行できます。つまり、ルーティング プロトコルの 2 つのインスタンスは別々のプロセスとして実行できます。

プロセスの再起動性

Cisco NX-OS のプロセスは、保護メモリ領域内で互いに独立に、またカーネルとも独立に動作します。このようにプロセスが分離されているため、障害が閉じこめられ、迅速な再起動が可能になります。プロセスの再起動性により、プロセスレベルの障害によってシステム全体に障害が及ぶのを防ぐことができます。また、大半のサービスはステートフルな再起動を実行できます。これにより、プラットフォーム内の他のサービス、およびネットワーク内の隣接デバイスへ透過的に、障害の発生したサービスを再起動し、動作を再開できます。

システムレベルのハイ アベイラビリティ

Cisco Nexus 1000V は冗長な VSM 仮想マシン（プライマリおよびセカンダリ）を HA ペアとして実行することをサポートします。デュアル VSM は、同時にはどちらか片方の VSM しかアクティブになれず、他方はスタンバイ バックアップとして機能する、アクティブ/スタンバイ キャパシティで稼働します。VSM は、Cisco Nexus 1000V インストールの一部としてプライマリまたはセカンダリとして設定されます。アクティブな VSM で障害が発生した場合にステートフル スイッチオーバーが提供されるよう 2 つの VSM 間で状態と設定は常に同期されます。

システムレベルのハイ アベイラビリティの詳細については、「[システムレベル ハイ アベイラビリティの設定](#)」(P.3-1) を参照してください。

ネットワークレベル ハイ アベイラビリティ

ネットワークレベルでの Cisco Nexus 1000V のハイ アベイラビリティには、ポート チャネルと Link Aggregation Control Protocol (LACP) が含まれます。ポート チャネルは、物理リンクをまとめて 1 つのチャネル グループに入れ、最大 8 つの物理リンクの帯域幅を集約した単一の論理リンクを作ります。ポート チャネル内のメンバー ポートに障害が発生すると、障害が発生したリンクで伝送されていたトラフィックはポート チャネル内のその他のメンバー ポートに切り替わります。

さらに、LACP では最大 16 個のインターフェイスが 1 つのポート チャネルに入るように設定できます。最大 8 つのインターフェイスをアクティブにすることができ、最大 8 つのインターフェイスをスタンバイ状態に入れることができます。

ポート チャネルと LACP の詳細については、『*Cisco Nexus 1000V Layer 2 Switching Configuration Guide, Release 4.2(1)SV1(5.1)*』を参照してください。

VSM から VSM へのハートビート

ここでは、次の内容について説明します。

- 「説明」 (P.1-4)
- 「コントロール インターフェイスおよび管理インターフェイスの冗長性」 (P.1-4)
- 「部分的な通信」 (P.1-5)
- 「通信損失」 (P.1-5)

説明

プライマリ VSM とセカンダリ VSM は VSM から VSM へのハートビートを使用して、ドメイン内で次の操作を実行します。

- 自分自身の存在を通知するブロードキャスト。
- 別の VSM の存在の検出。
- アクティブ/スタンバイ冗長性状態のネゴシエート。

VSM は、最初に起動したときに、ドメインにディスカバリ フレームをブロードキャストして、別の VSM の存在を検出します。他の VSM がない場合、起動 VSM がアクティブになります。別の VSM がアクティブであることが判明した場合、起動 VSM はスタンバイになります。別の VSM が（たとえば、システム リロード中に）初期化中であることが判明した場合、プライマリ VSM がセカンダリ VSM よりも優先してアクティブになります。

初期接続とロール ネゴシエーションの後、アクティブ VSM とスタンバイ VSM は 1 秒間隔のハートビート メッセージで次の情報をユニキャストします。

- 冗長性状態
- 他方の VSM によるアクションを要求するコントロール フラグ

この項の後半では、VSM から VSM へのハートビートのメカニズムについて、次の間隔の順守を含め説明します。

間隔	説明
1 秒	ハートビート要求が送信される間隔。
3 秒	失われたハートビートによってコントロール インターフェイスでの通信劣化が示され、管理インターフェイスにもハートビートが送信されるまでの間隔。
6 秒	アクティブ VSM がスタンバイ VSM にリロードを指示するまでの通信損失の間隔。

コントロール インターフェイスおよび管理インターフェイスの冗長性

アクティブ VSM がコントロール インターフェイスを介してハートビート応答を 3 秒間受信しない場合、通信は劣化していると見なされ、VSM も管理インターフェイス経由での要求の送信を開始します。この場合、管理インターフェイスは冗長性を提供し、両方の VSM がアクティブ（アクティブ-アクティブ状況やスプリットブレイン状況とも呼ばれる）にならないように機能します。ただし、管理インターフェイスのみでハートビート要求およびハートビート応答が処理されるため、通信は完全冗長ではありません。VSM 間の AIPC およびデータの同期は、コントロール インターフェイスだけを使用して行われます。

部分的な通信

コントロール インターフェイス上の通信がある程度劣化すると、2 つの VSM は同期を失います。同期は、スタンバイ VSM のリロードによってのみ復元されます。VSM 間のこの通信が 6 秒を超えて中断されても、VSM が管理インターフェイスでまだ通信できる場合、アクティブ VSM はスタンバイ VSM にリロードを指示します。通信が劣化したままの場合は、コントロール通信が復元されるまで、スタンバイ VSM が複数回リロードすることがあります。

通信損失

冗長 VSM 間または Cisco Nexus 1010 間に通信がない場合、いずれも他方の存在を検出できません。アクティブがスタンバイをドロップします。スタンバイは、応答の欠如をアクティブに障害が発生したと解釈し、スタンバイもアクティブになります。これは、両方が Virtual Center に接続し、VEM と通信してシステムを制御しようとする、アクティブ-アクティブまたはスプリットブレインと呼ばれる状況です。

冗長 VSM または Cisco Nexus 1010 が管理インターフェイスの同じ IP アドレスを使用するため、この IP アドレスへのパスがネットワークで変更され、リモート SSH/Telnet 接続に失敗する可能性があります。したがって、スプリットブレイン競合の間はコンソールを使用することを推奨します。

スプリットブレイン競合中に、設定を変更しないでください。コントロール通信および管理通信が冗長デバイス間で復元されると、プライマリのリロードによってスプリットブレイン競合が解決され、プライマリがスタンバイ モードに戻り、セカンダリと同期されます。スプリットブレイン競合中に行われたプライマリ設定への変更は、リロード後に失われます。

スプリットブレイン競合中にプライマリに変更を加える必要がある場合は、セカンダリをシャットダウンします。これにより、2 台のデバイスが接続を再開して、リロードされないようにすることができます。この場合、スプリットブレイン競合中にプライマリに加えた変更は保持されます。



(注)

スプリットブレイン状態が解決され、プライマリおよびセカンダリの通信が再開された後は、セカンダリの設定が優先されます。両方がアクティブなときに変更した場合は注意してください。



(注)

アクティブからスタンバイへの遷移では、常に、Cisco Nexus 1000V と Cisco Nexus 1010 の両方をリロードする必要があります。

VSM-VEM の通信損失

原因となった特定のネットワーク障害によって、各 VSM は、場合によっては重なり合う VEM の別のサブセットに到達します。スタンバイ状態だった VSM が新たにアクティブになると、現在のアクティブ デバイスとしてこの VSM に切り替えるという要求をすべての VEM にブロードキャストします。VEM が新しいアクティブ VSM に切り替わるかどうかは、次によって決定されます。

- 各 VEM と 2 つの VSM の間の接続。
- VEM がこの切り替え要求を受信するかどうか。

VEM は、新しいアクティブ VSM からハートビートを受信しても、元のアクティブ VSM に接続されたままです。ただし、新しいアクティブ VSM からの切り替え要求を受信すると、元のアクティブ VSM から切断され、新しいアクティブ VSM に接続されます。

VEM は、元のアクティブ デバイスから切断され、新しいアクティブ デバイスからのみハートビートを受信する場合、ヘッドレス モードになるまでこれらのハートビートを無視します。これは、元のアクティブ デバイスからハートビートを受信しなくなってから約 15 秒後に発生します。この時点で、VEM に新しいアクティブ デバイスへの接続がある場合、VEM は新しいアクティブ デバイスに接続されます。

単方向通信

スタンバイ VSM がハートビート要求を受信してもアクティブ VSM が応答を受信しないというネットワーク通信の障害がある場合、次のようになります。

- アクティブ VSM がスタンバイ VSM がいないことを宣言する。
- スタンバイ VSM がスタンバイ状態のままとなり、アクティブ VSM からハートビートを受信し続ける。

冗長性状態が矛盾しており (**show system redundancy state**)、2 つの VSM の同期が失われます。

双方向通信が再開されると、スタンバイ VSM が不一致とリロードを検出します。



(注)

アクティブからスタンバイの方向の単方向通信の障害が発生した場合、通信ができないに等しくなります。なぜなら、スタンバイ VSM がアクティブ VSM 要求への応答としてハートビートを送信するだけになるからです。

参考資料

Cisco Nexus 1000V を設定する前に、次のマニュアルをお読みになることを推奨します。

- 『Cisco Nexus 1000V Getting Started Guide, Release 4.2(1)SV1(5.1)』
- 『Cisco Nexus 1000V Port Profile Configuration Guide, Release 4.2(1)SV1(5.1)』
- 『Cisco VN-Link: Virtualization-Aware Networking』 (ホワイト ペーパー)