

Cisco Catalyst 6500 の NSF/SSO (ノンストップ フォワーディング / ステートフル スイッチオーバー)

Cisco® Catalyst® 6500 は、企業ネットワークとサービス プロバイダー ネットワークの要所に数多く展開されています。Cisco Catalyst 6500 はネットワーク内できわめて重要な位置に設置されているため、100% に近いアベイラビリティを実現する必要があります。Cisco Catalyst 6500 のプラットフォームは何年にもわたって進化し続け、高度な復元メカニズムを提供することにより、より高レベルのアベイラビリティを実現することを可能にしました。ハイ アベイラビリティ メカニズムの例としては、冗長スーパーバイザ、冗長スイッチと冗長リンク、Cisco EtherChannel® テクノロジー、Spanning-Tree Protocol (STP; スパニングツリー プロトコル)、Unidirectional Link Detection (UDLD; 単一方向リンク検出) プロトコル、Hot Standby Router Protocol (HSRP)、Gateway Load Balancing Protocol (GLBP)、およびルーティング プロトコルの等コスト パスなどがあります。このハイ アベイラビリティ機能に追加された最新機能が、Nonstop Forwarding (NSF; ノンストップ フォワーディング) / Stateful Switchover (SSO; ステートフル スイッチオーバー) です。NSF/SSO は、Cisco IOS® ソフトウェア リリース 12.2(18)SXD で Supervisor Engine 2 と Supervisor Engine 720 に導入されたスーパーバイザの冗長メカニズムであり、レイヤ 2 ~ 4 のシャーシ内でのステートフル スイッチオーバーを提供します。NSF/SSO では、スーパーバイザのスイッチオーバーがすばやく実行されるため、パケット損失は 0 ~ 3 秒程度で済み、Mean Time To Repair (MTTR; 平均修復時間) が短縮されます。NSF/SSO は、企業ネットワークまたはサービス プロバイダーのネットワークの最重要部分に展開可能です。これはネットワーク内の終端ポイントに不可欠の機能であり、Voice over IP (VoIP)、映像、およびパケット損失の影響を受けやすいその他のアプリケーションが関与するネットワークのダウンタイムを最小限に抑えることができます。

本書では、Cisco IOS ソフトウェアの Cisco Catalyst 6500 における NSF/SSO スーパーバイザの冗長動作について紹介します。NSF/SSO に関するプラットフォーム固有の詳細事項、NSF/SSO でサポートされている機能 (Multicast Multilayer Switching [MMLS] NSF/SSO など)、および NSF/SSO のパフォーマンス実績を紹介します。本書では扱いませんが、NSF/SSO を備えたアベイラビリティの高いネットワークの設計方法を理解することは非常に重要です。ハイ アベイラビリティ キャンパス ネットワークの設計に関する情報、NSF/SSO の動作全般に関する詳細情報、NSF/SSO のコンフィギュレーションガイド、Cisco Catalyst 6500 のハイ アベイラビリティ メカニズムを網羅した完全なリスト、および Cisco Catalyst 6500 の Cisco Catalyst OS (オペレーティング システム) に関するスーパーバイザの冗長性に関する情報は、この文書の「参考資料」を参照してください。

スーパーバイザの冗長性

スイッチの冗長コンポーネント

Cisco Catalyst 6500 シリーズ スイッチは、アベイラビリティの高いシステムを実現するための基盤として、冗長ハードウェア システム アーキテクチャが採用されています。Cisco Catalyst 6500 スイッチでは、次のコンポーネントによってスイッチの冗長性を実現しています。

- スーパーバイザ エンジン — すべての Cisco Catalyst 6500 シャーシは、システムのハイ アベイラビリティを提供するため、冗長スーパーバイザ エンジンをサポートしています。スーパーバイザはアクティブおよびスタンバイ モードで動作し、フェールオーバーのためのさまざまな冗長メカニズムをサポートしています。
- スイッチ ファブリック — スイッチ ファブリックは、ファブリック対応のライン カードにデータ パスを提供し、使用可能なシステム帯域幅を共有バス キャパシティの 32 Gbps から、SFM2 (スイッチ ファブリック モジュール 2) 搭載の Supervisor Engine 2 では 256 Gbps、Supervisor Engine 720 では 720 Gbps に増やします。スイッチ ファブリックに障害が発生した場合は、冗長スイッチ ファブリックに切り替わります (冗長ファブリックが存在する場合)。
- 電源装置 — すべての Cisco Catalyst 6500 シャーシでは、冗長電源装置がサポートされているので、1 台の電源装置に障害が発生しても動作に影響はありません。

- ファントレイ — 各ファントレイには複数のファンがあります。また、Cisco Catalyst WS-C6509-NEB-A シャーシには、オプションの冗長ファントレイが用意されています。
- ラインカードの Online Insertion and Removal (OIR; ホットスワップ) — システムに影響を与えることなく新しいモジュールを追加でき、ラインカードを交換しても設定が失われません。ローカルフォワーディングエンジンを備えたモジュール (Distributed Forwarding Card [DFC]) を取り付けると、ローカルフォワーディングエンジンのハードウェアテーブルには、最新のフォワーディング情報が読み込まれます。

スーパーバイザの冗長性の定義

スーパーバイザの冗長性を Cisco Catalyst 6500 で実現するには、以下が必要です。

- シャーシあたり 2 台のスーパーバイザ
- 2 台のスーパーバイザ間で情報を同期させる冗長プロトコル

最初に起動したスーパーバイザエンジンが、アクティブスーパーバイザエンジンになります。アクティブスーパーバイザでは、コントロールプレーンとフォワーディングの決定が実行されます。もう 1 台のスーパーバイザはスタンバイスーパーバイザとなり、コントロールプレーンまたはデータプレーンの決定には参加しません。アクティブスーパーバイザは、設定とプロトコルのステート情報をスタンバイスーパーバイザと同期させます。そのため、アクティブスーパーバイザに障害が発生した場合は、スタンバイスーパーバイザでただちにアクティブスーパーバイザの役割を引き継ぐことができます。アクティブスーパーバイザからスタンバイスーパーバイザへの「交代」のプロセスは、スイッチオーバーと呼ばれます。

一度にアクティブにできるスーパーバイザは 1 台だけであり、冗長化した場合でも、スーパーバイザエンジン間の負荷分散は行われません。ただし、スタンバイスーパーバイザエンジンが動作しているときは、インターフェイスがアクティブになっているため、トラフィックのフォワーディングに使用することができます。

スーパーバイザの冗長動作

スーパーバイザの冗長動作は、Route Processor Redundancy (RPR) と RPR plus (RPR+) から、Single Router Mode (SRM) /SSO、そして NSF/SSO へと発展してきました。それぞれの冗長動作モードは、以前よりも機能が向上しています。

- RPR — RPR は、Cisco IOS ソフトウェアに初めて導入された冗長動作モードです。RPR モードでは、スタートアップコンフィギュレーションとブートレジスタがアクティブスーパーバイザとスタンバイスーパーバイザ間で同期化されます。ただし、スタンバイスーパーバイザは完全には初期化されず、両方のスーパーバイザ間のイメージは同一である必要がありません。スイッチオーバーが発生すると、スタンバイスーパーバイザは自動的にアクティブになりますが、ブートプロセスを完了させなければなりません。さらに、すべてのラインカードがリロードされ、ハードウェアが再プログラミングされます。RPR でのスイッチオーバーには、2 分以上の時間がかかります。
- RPR+ — RPR+ は RPR の強化版です。スタンバイスーパーバイザは完全に起動されていて、スイッチオーバーが発生してもラインカードのリロードは行われません。ランニングコンフィギュレーションはアクティブスーパーバイザとスタンバイスーパーバイザ間で同期化されます。また、RPR で行われる同期化動作のすべてが実行されます。同期化はスイッチオーバーの前に完了し、スタンバイスーパーバイザがアクティブになったときには同期化された情報が使用されるので、ダウンタイムが最小限に抑えられます。ただし、リンク層またはコントロールプレーンの情報の同期化は行われません。インターフェイスはスイッチオーバー後にただちに復旧し、ハードウェア情報は再プログラミングされます。RPR+ によるスイッチオーバーには 30 秒以上の時間がかかります。
- SRM/SSO — SSO は RPR+ の機能をさらに拡張し、スーパーバイザに障害が発生した場合にレイヤ 2 プロトコルのトランスペアレントなフェールオーバーを提供します。SSO はレイヤ 2 プロトコルについてステートフルとなっています。Policy Feature Card (PFC; ポリシーフィーチャカード) と DFC のハードウェアテーブルは、スイッチオーバーが発生しても維持されます。これにより、レイヤ 2 およびレイヤ 4 でトランスペアレントなフェールオーバーが可能になります。SSO は、SRM/SSO および NSF/SSO の要件です。SRM および NSF は追加のレイヤ 3 ルーティング機能を提供するもので、SSO は SRM および NSF とは独立して使用できます。SRM/SSO を使用しているときにスイッチオーバーが発生すると、ルーティングプロトコルは再起動します。ただし、SRM/SSO は、設定可能なルートコンバージェンス間隔の間、既存の PFC および

DFC のレイヤ 3 スwitチング情報を使用してトラフィックを転送します。この間に、新しくアクティブになった Multilayer Switch Feature Card (MSFC; マルチレイヤ スwitチ フィーチャ カード) は、独自のルーティング テーブルを構築します。これにより、ダウンタイムは最小限に抑えられますが、スーパーバイザに障害が発生するとピアでは再コンバージェンスが必要です。SRM/SSO のスwitチオーバーにかかる時間は、レイヤ 2 のユニキャスト トラフィックで 0 ~ 3 秒です。

- NSF/SSO — NSF は SSO と連携して動作し、スwitチオーバー後のレイヤ 3 の完全性を保証します。これによって、ルータはアクティブ スーパーバイザに障害が発生しても既知のルートでデータ パケットの転送を継続することができ、その間にルーティング プロトコル情報が回復され、検証が行われます。再起動するルータで近接ルータとのピアリング アレンジメントが失われた場合でも、転送を続行できます。NSF が機能するためには、スーパーバイザのスwitチオーバー中にコントロール プレーンとデータ プレーンが隔離されていることが必要です。データ プレーンは、スwitチオーバー以前の Cisco Express Forwarding (CEF) の情報に基づいてパケット フォワーディングを継続します。コントロール プレーンにはグレースフル リスタート ルーティング プロトコル拡張機能が実装されており、スwitチオーバー後にスーパーバイザが再起動したことを近接する NSF 認識ルータに通知して、ネイバ隣接関係を再編し、さらにルーティング プロトコル データベースを再構築します。NSF 対応ルータには NSF 機能が実装されており、スーパーバイザに障害が発生したあともデータ パケットの転送を続行できます。NSF 認識ルータは、NSF のグレースフル リスタート メカニズムに対応しています。NSF 認識ルータは、再起動する NSF 対応ルータとのネイバ関係を破棄せず、近接する NSF 対応ルータの再起動を支援できるため、不要なルート フラップやネットワークの不安定が発生しにくくなります。NSF 対応ルータも NSF を認識することができます。

MMLS NSF/SSO によって、システムはスーパーバイザ エンジンのスwitチオーバーの間、PFC3 および DFC3 ハードウェア内のマルチキャスト フォワーディングの状態を維持できるので、マルチキャスト サービスの中断を最小限に抑えることができます。MMLS NSF/SSO 以前は、マルチキャスト フォワーディング エントリはスタンバイ スーパーバイザ エンジンに同期化されていませんでした。NSF/SSO のスwitチオーバー時間は、レイヤ 2 ~ 4 のユニキャストまたはマルチキャスト トラフィックで 0 ~ 3 秒です。

表 1 に、Cisco Catalyst 6500 でサポートされている各冗長モードに必要なソフトウェア バージョンの最低要件を示します。

表 1 スーパーバイザの冗長モードのサポート

スーパーバイザ エンジン	RPR および RPR+	SRM/SSO	NSF/SSO
Supervisor Engine 1A	12.1(13)E	-	-
Supervisor Engine 2	12.1(13)E または 12.1(17d)SXB	-	12.2(18)SXD
Supervisor Engine 720	12.2(14)SX	12.2(17b)SXA および 12.2(17d)SXB のみ	12.2(18)SXD

2 台の Supervisor Engine 720 を使用する場合、デフォルトの冗長動作モードは SSO で、Cisco IOS ソフトウェア リリース 12.2(17b)SXA 以降のリリースでサポートされています。2 台の Supervisor Engine 2 を使用する場合、デフォルトの冗長動作モードは SSO で、Cisco IOS ソフトウェア リリース 12.2(18)SXD 以降のリリースでサポートされています。Cisco IOS ソフトウェア 12.2SX より以前のリリースでは、デフォルトの冗長動作モードは RPR+ となります。

SRM/SSO は、Cisco IOS ソフトウェア リリース 12.2(17b)SXA および 12.2(17d)SXB では、デフォルトで有効になっています。SRM/SSO は、Cisco IOS ソフトウェア リリース 12.2(18)SXD 以降では NSF/SSO に置き換えられています。NSF 認識と NSF 機能を有効にする方法は、ルーティング プロトコルごとに異なります。NSF/SSO の設定の詳細については、次の URL を参照してください。

<http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/122sx/swcg/nsfss0.htm>

ソフトウェアのアップグレード

SSO を機能させるには、各スーパーバイザ エンジンが同じリリースの Cisco IOS ソフトウェアを稼働している必要があります。Fast Software Upgrade (FSU) を使用すれば、ソフトウェアのアップグレード計画に伴うダウンタイムを最小限に抑えることができます。このプロセスが行われている間、冗長モードはアップグレード中に RPR に戻ります。手順の詳細は次の URL に記載されています。

http://www.cisco.com/jp/service/manual_j/sw/cat60/iosscg/chapter07/07_redund.shtml#62923

RPR+ または SSO 冗長モードで稼働するには、イメージのバージョンがアクティブ スーパーバイザとスタンバイ スーパーバイザとで同一である必要があります。これらの冗長モードでは、スタンバイ スーパーバイザ エンジンがオンラインになると、スタンバイ スーパーバイザ エンジンに搭載されているイメージのバージョンがアクティブ スーパーバイザ エンジンによってチェックされます。スタンバイ スーパーバイザ エンジンのイメージがアクティブ スーパーバイザ エンジンのイメージと一致しない場合は、ソフトウェアは冗長モードを RPR に切り替えてアップグレードを実行し、アップグレードが完了してから SSO に再度切り替えます。

将来の Cisco In Service Software Upgrade (ISSU) では、SSO 冗長モードでソフトウェアのアップグレードが可能になります。NSF/SSO は Cisco ISSU の基本的な構成要素です。

スーパーバイザの障害検出

スーパーバイザの冗長性は、システムのハイアベイラビリティの一部に過ぎません。ハードウェアおよびソフトウェアの障害を検出することが、復元力を備えたスーパーバイザのスイッチオーバー メカニズムを実現するための主要な要件となります。この障害検出のフレームワークを提供するのが、Generic Online Diagnostics (GOLD) とプラットフォーム依存の診断です。

GOLD は、シスコシステムズのプラットフォーム上の診断動作を行うための、共通のアーキテクチャを定義しています。GOLD はプラットフォーム固有のオンライン診断と連携して動作し、起動中および稼働中のシステムの健全性を保証するのに役立ちます。Cisco Catalyst 6500 シリーズでは、情報の大部分がハードウェア ベースで処理されるため、ハードウェアの機能を定期的にテストすることが非常に重要です。Cisco Catalyst 6500 システムでは、アクティブおよびスタンバイ スーパーバイザを含むほとんどのモジュールで障害検出診断メカニズムが有効になっています。診断テストの結果は、スイッチオーバーの決定に利用することができます。Cisco Catalyst 6500 にはオンライン診断が組み込まれているため、スイッチオーバーのトリガーとなるものがソフトウェアのクラッシュやキープアライブ メカニズムのみに限定されません。スイッチオーバーは、スーパーバイザの制御パスとデータパスに不整合や障害が生じた場合、またはランタイム診断でハードウェアの誤作動が検出された場合にも起動されることがあります。GOLD はスイッチオーバーの決定と起動に役立つだけでなく、スタンバイ スーパーバイザを定期的に監視して、スイッチオーバーのときにいつでも交代できることを保証します。GOLD には、スイッチオーバーのスケジューリング機能も組み込まれています。管理者は、オンライン診断 CLI (コマンドライン インターフェイス) によって、スイッチオーバーが特定の時間に行われるようにスケジューリングすることができます。

GOLD では、次の問題を検出してスーパーバイザのスイッチオーバーを決定することができます。

- ハードウェア コンポーネントの障害
- コネクタの障害
- インターフェイスの障害
- メモリ エラー
- データ プレーンとコントロールプレーン間の不整合

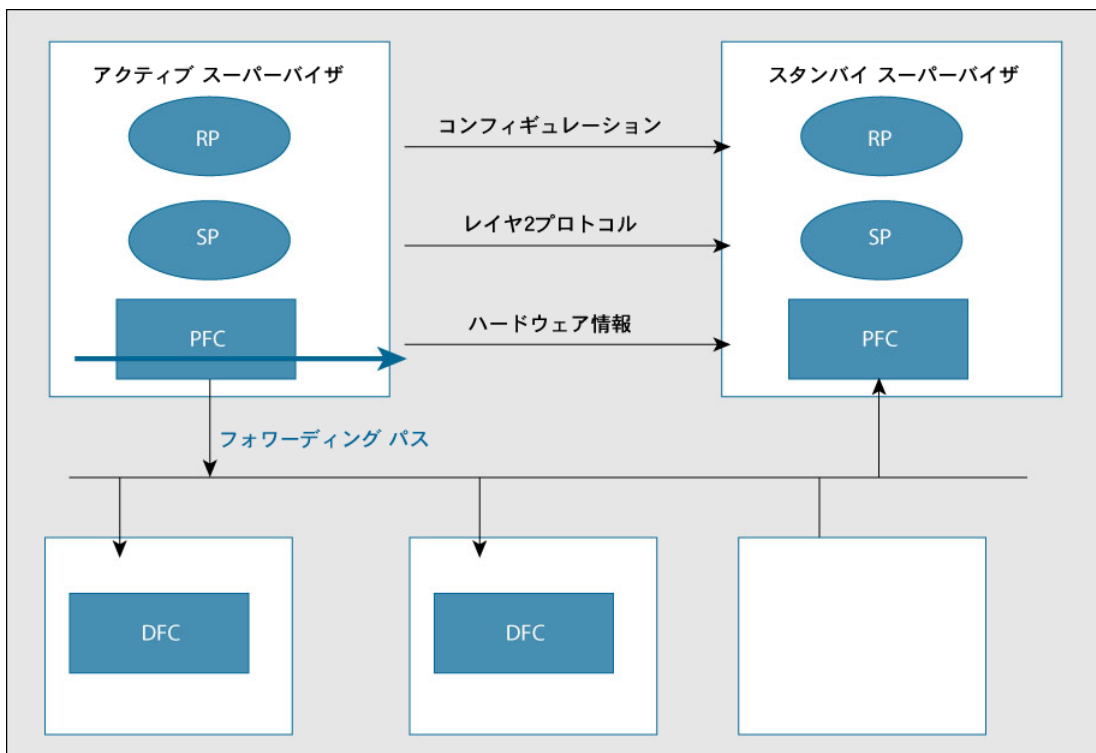
SSO

SSO の動作

SSO の同期化動作

図 1 は、通常の動作時の SSO の同期化を図示したものです。SSO モードでは、レイヤ 2 プロトコルと PFC のハードウェア情報がアクティブ スーパーバイザからスタンバイ スーパーバイザへと同期化されます。この図では、RP は Route Processor (RP; ルート プロセッサ)、SP は Switch Processor (SP; スイッチ プロセッサ)、PFC は Policy Feature Card (PFC; ポリシー フィーチャ カード)、DFC は Distributed Forwarding Card (DFC; 分散フォワーディング カード) を表します。

図 1 SSO の同期化動作



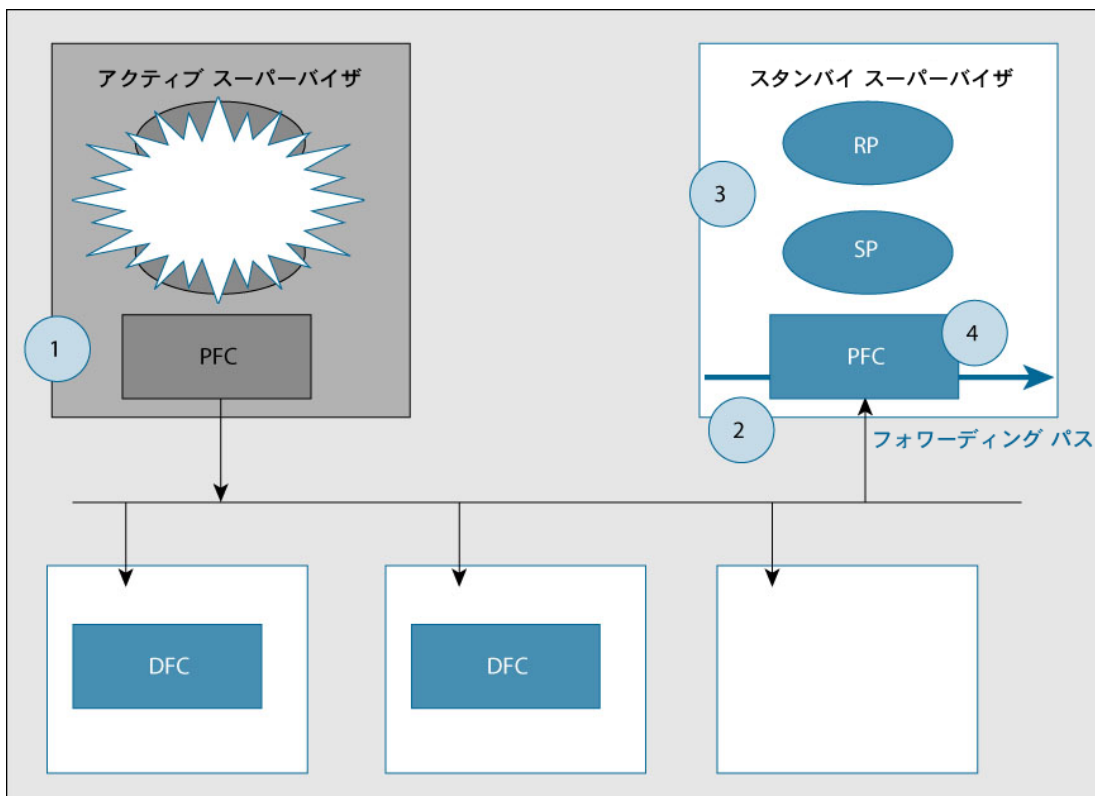
SSO では RPR+ の同期化機能が拡張され、レイヤ 2 およびレイヤ 4 でトランスペアレントなフェールオーバーが可能となっています。アクティブ スーパーバイザからスタンバイ スーパーバイザへの同期は、スタートアップ コンフィギュレーション、スタートアップ変数、およびランニング コンフィギュレーションだけでなく、ランタイム データにも適用されます。フェールオーバーの際にアクティブ スーパーバイザおよびスタンバイ スーパーバイザ上のピア プロトコル プロセス間で規則性と信頼性の高い通信を実現するため、このダイナミックなデータ同期化 (チェック ポインティング) では、Cisco IOS Redundancy Facility と Checkpoint Facility を使用しています。起動時には、SSO のバルク同期化が実行されます。システム稼働時にシステム内で変更が生じると、設定の同期化と状態のチェック ポインティングがさまざまなプロトコルで実行されます。

SSO は、レイヤ 2 のダイナミック プロトコルのランタイム データを同期化します。レイヤ 2 のコントロールプレーン、設定、その他のネットワーク関連の変更が生じると、アクティブ スーパーバイザおよびスタンバイ スーパーバイザ上のピア プロセス間で動作する Cisco IOS Checkpoint Facility が変更に関する情報を交換します。表 2 は、SSO でサポートされているレイヤ 2 プロトコルの一覧を示します。たとえば、スタンバイ スーパーバイザ上の STP データベースは、アクティブ スーパーバイザからのプロトコル情報とポート ステータスの両方のチェック ポインティングを実行することによって最新の状態に維持されます。

SSO は、アクティブ スーパーバイザとスタンバイ スーパーバイザ間のハードウェア フォワーディング テーブルも同期します。PFC はスーパーバイザのドータカードであり、ハードウェア スイッチングを実行する Application Specific Integrated Circuit (ASIC; 特定用途向け IC) を搭載しています。ハードウェア テーブルの新しいエントリを PFC にダウンロードするときには、システム内の他のすべてのフォワーディング エンジンにも同時にダウンロードされます。これにより、スタンバイ スーパーバイザの PFC は、アクティブ スーパーバイザの PFC および DFC と同じフォワーディング情報を保持できます。MAC アドレス テーブル、Forwarding Information Base (FIB; フォワーディング情報ベース)、隣接テーブル、Access Control List (ACL; アクセス制御リスト)、および Quality of Service (QoS; サービス品質) のハードウェア テーブルの内容は、スイッチオーバー後のスイッチングの決定にも使用できます。

図 2 は、スーパーバイザのスイッチオーバー動作を図示したものです。スイッチオーバーが発生しても、トラフィックは混乱なく転送できます。1、2、3、4 という番号は、スイッチオーバーの順番を表します。以下、これらのステップについての説明をします。

図 2 スーパーバイザのスイッチオーバー動作



通常の動作時には、ハードウェア テーブルとレイヤ 2 プロトコルの状態が同期されています。図 1 には、SSO 冗長モードでのスーパーバイザのスイッチオーバー動作が図示されています。スイッチオーバーのステップ 1 ~ 4 についての説明は、次のとおりです。

1. システムによってアクティブ スーパーバイザでソフトウェアまたはハードウェアの障害が検出されると、スイッチオーバーが起動されます。障害は、ソフトウェアの例外ハンドラ、GOLD のバックグラウンド チェック、RP と SP 間のキーブアライブの失敗、Supervisor Engine 720 でのファブリック スイッチング モジュールの状態変化などによって検出されます。また、スイッチオーバーはユーザによって起動される場合もあります。

2. ライン カードの同期によって、システム内のすべてのモジュールでスイッチオーバーの発生が認識されます。スタンバイスーパーバイザはアクティブスーパーバイザの役割を引き受け、データは新たにアクティブになったスーパーバイザ上のPFCによって転送されます。
3. 新たにアクティブになったスーパーバイザ上のSPとRPは、プロトコルとデータパケットの処理を開始します。SSO認識プロトコルはスイッチオーバーの影響を受けず、ネットワークからのアップデート処理を開始します。
4. 非SSO認識プロトコルとルーティングプロトコルが起動されます。SRM/SSOは、設定可能なルートコンバージェンス間隔が経過したあとにスイッチオーバー以前のFIB情報を除去するため、ルーティングプロトコルが収束されている間に、レイヤ3の転送をハードウェア内で継続できます。ピアでは障害の前後で再コンバージェンスが必要です。スタティックルートは、動的ではなく静的な設定に基づいているため、スイッチオーバーの前後で維持されたままになります。QoSまたはACLポリシーから派生するサポート対象のレイヤ2制御プロトコルとレイヤ4ポリシーは、スイッチオーバーの影響を受けません。

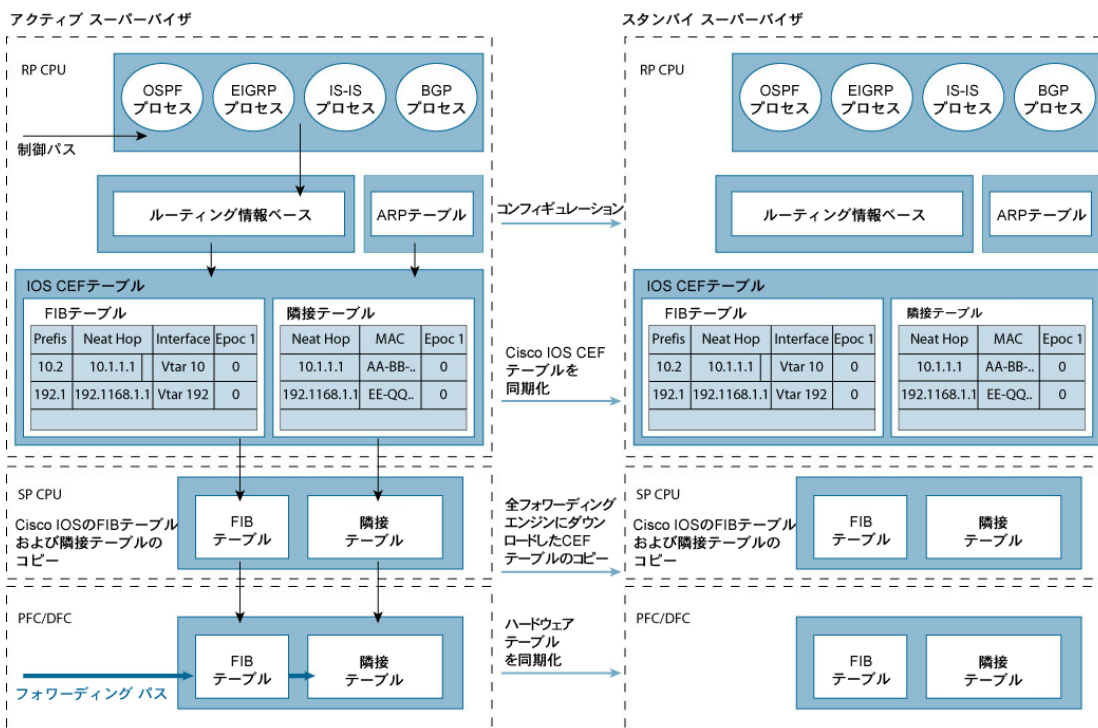
NSF/SSO

NSF/SSOの動作

NSF/SSOの同期化動作

図3は、NSF/SSOによるスーパーバイザの同期化を図示したものです。

図3 NSF/SSOの同期化動作



シスコ製ルータでのパケットフォワーディングは、CEFによって実行されます。CEFは、FIBテーブルと隣接テーブルの2つテーブルを保持しています。FIBテーブルはルーティングテーブルから抽出して作成されたテーブルで、フォワーディングプロセスに関連する情報のみが含まれ、特定のルーティングプロトコルに関する情報は含まれません。

たとえば、ルーティングプロトコルのアドミニストレーティブディスタンスは、フォワーディングプロセスとは関連性がありません。隣接テーブルは、隣接ノードのネクストホップリライト情報を集約したものです。

通常の動作時には、各ルーティングプロトコルで計算されたルートがシステムによって Routing Information Base (RIB; ルーティング情報ベース) という共通データベースに集約されます。すべてのルーティングプロトコルの情報が RIB 内に存在するときは、RIB の情報に基づいて、各ネットワークおよびサブネットに対応する最低コストのネクストホップ宛先が決定されます。その時点で、最低コストパスのルーティングプレフィクスと隣接情報が CEF テーブルに入力されます。ルーティングプロトコルに変更が生じると、ソフトウェアの CEF データベースではアクティブスーパーバイザからスタンバイスーパーバイザへのチェックポインティングが行われ、CEF テーブルがシステム内に存在するすべての PFC および DFC のハードウェアにダウンロードされます (スタンバイ PFC を含む)。これにより、フォワーディングテーブルの同期化がソフトウェアおよびハードウェアレベルで保証され、スイッチオーバー後も最も正確な最新のフォワーディングテーブル情報に基づいてフォワーディングが実行されることが保証されます。

新旧の CEF エントリを区別できるように、CEF エントリごとのエポック番号が割り当てられています。これは、FIB および隣接データベースのバージョンングと呼ばれています。エポック番号はソフトウェアの CEF テーブルのみに管理され、フォワーディングパスには影響しません。スイッチオーバーが発生すると、「グローバルエポック番号」の数値が 1 つ増えます。スイッチオーバー後、新たにアクティブになったスーパーバイザで新しいルーティング情報が入力されると、CEF エントリのバージョン番号はグローバルエポック番号によって更新されます。ルーティングプロトコルによってコンバージェンスの完了が通知されると、現在のエポックよりも前のバージョン番号を持つ FIB および隣接エントリはクリアされます。

スーパーバイザのスイッチオーバー動作

スイッチオーバーの発生時に、NSF を正常に機能させるには、データプレーンとレイヤ 3 のコントロールプレーンが隔離されていることが必要です。コントロールプレーンが新しいルーティングプロトコルデータベースを構築し、ピアリングアグリメントを再起動させるのに対して、データプレーンはスイッチオーバー以前のフォワーディングテーブルの同期化に基づいてトラフィックフォワーディングを継続します。次の項では、NSF 認識ネイバが存在することを前提としています。NSF 認識ネイバによる支援がないと、NSF 対応システムではスイッチオーバー中にデータベースを再構築することも、ネイバ隣接関係を維持することもできません (ただし、Cisco Intermediate System-to-Intermediate System [IS-IS] の NSF 実装では、NSF 認識ネイバは不要)。

スイッチオーバー動作は、図 2 で説明したとおりに発生します。ただし、NSF 対応ルーティングプロトコルの再初期化によってルートフラップが発生することはありません。図 4 は、実行される汎用ルーティングプロトコルの NSF/SSO の動作について説明しています。図 4 は、近接する NSF 認識ルータと NSF 対応の Cisco Catalyst 6500 について図示しています。ここには、新たにアクティブになった Cisco Catalyst 6500 のスーパーバイザと NSF/SSO 動作の手順が示されています。ただし、この図には障害を起こしたアクティブスーパーバイザは含まれていません。スーパーバイザの SP と PFC に適用される手順は、ラインカードプロセッサと DFC にも適用されることに注意してください。

図4 NSF/SSOの動作

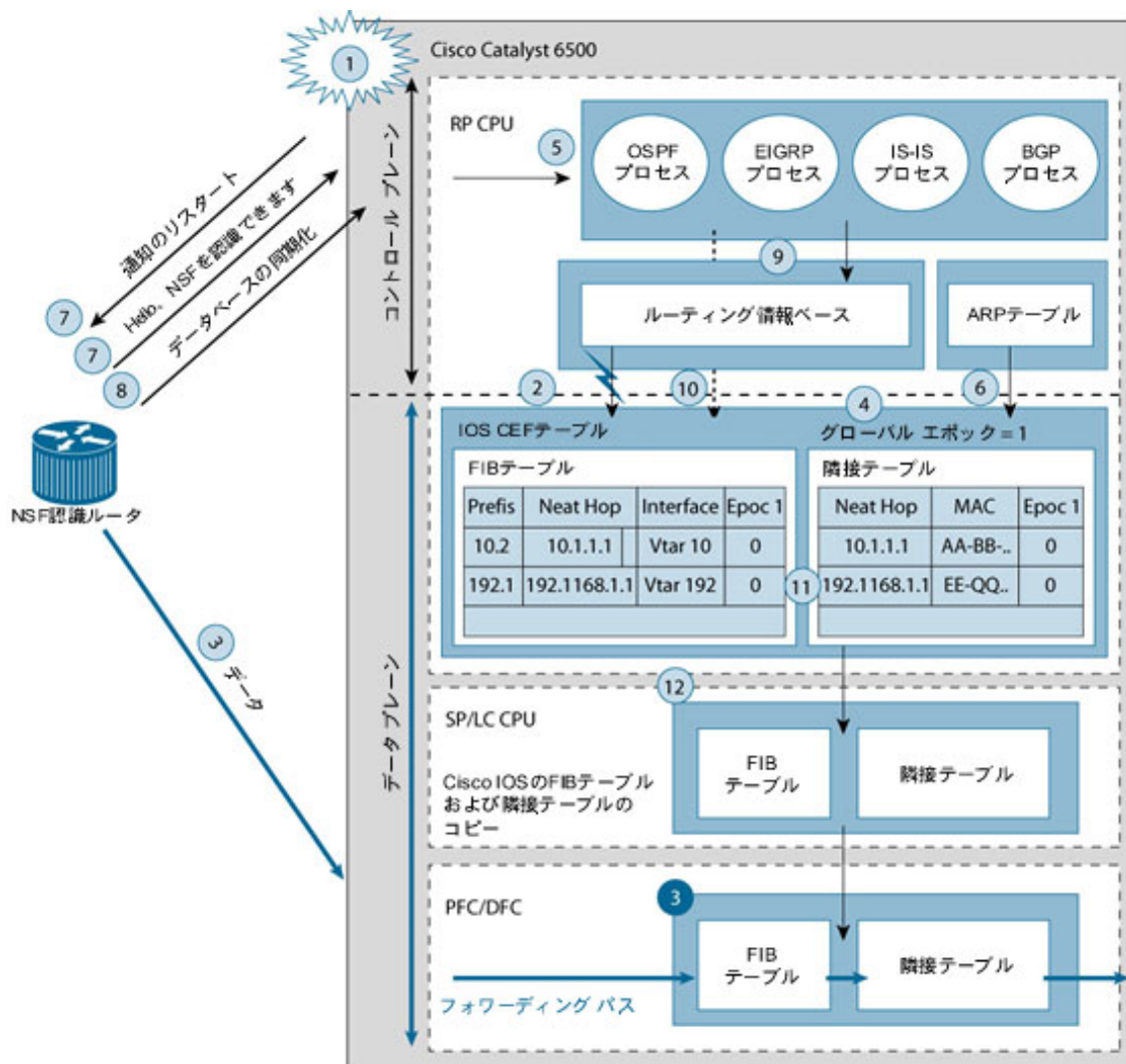


図4のステップ1～12についての説明は、次のとおりです。これらのステップは、すべて「新たにアクティブになった」スーパーバイザで実行されます。

1. スイッチオーバーが起動されます。
2. ルーティング プロトコルのプロセスによって、スーパーバイザのフェールオーバーが発生したことが通知されます。コントロールプレーンとデータプレーンを隔離するため、ルーティングプロトコルの再コンバージェンスが実行されるまでFIBがRIBから隔離されます。
3. パケット転送は、スタンバイスーパーバイザへの切り替えが行われている間、既知の最新FIBおよび隣接エントリに基づいて継続されます。
4. グローバル エポック番号の数値が1つ増えます。たとえば、スイッチオーバー以前にグローバル エポック番号が0であった場合は、1になります。
5. スーパーバイザがコントロールプレーンのトラフィックの処理を開始します。

6. ソフトウェアの隣接テーブルには、スイッチオーバー前の Address Resolution Protocol (ARP) テーブルの内容が入力されています。アップデートされた CEF エントリに新しいグローバル エポック番号が付与されます。このエポック番号は、RP ソフトウェアの CEF エントリでのみ使用可能で、ハードウェア テーブルには存在しません。ハードウェアには、新しい隣接エントリがダウンロードされます。
7. ルーティング プロトコル固有のネイバおよび隣接関係の再取得が実行されます。再起動する NSF 対応ルータは、そのネイバに対して、隣接関係が再取得されていること、および NSF 認識ネイバはネイバ関係を再初期化しないことを通知します。再起動の通知が受信されると、プロトコル固有の手順が実行されて隣接関係を維持できるようになります。ほとんどの場合、再起動の通知は、Hello パケットに再起動フラグを設定して、復旧プロセス中に短い間隔で Hello パケットを送信することによって行われます。また、NSF 認識ネイバは NSF 認識機能を持つことを、再起動するルータに通知することがあります。非 NSF 認識ネイバでは再起動の通知は無視され、隣接関係が停止されます。現在の NSF の実装では、複数の NSF 可能ネイバの同時再起動はサポートされていません。
8. ルーティング プロトコル固有のデータベース同期化が実行されます。ルーティング プロトコルのプロセスによって、NSF 認識ネイバからのデータベース情報を利用してデータベースが再構築されます。
9. ルーティング データベースが同期化されると、ディスタンス ベクタ、パス ベクタ、または Shortest Path First (SPF) アルゴリズムの計算によって、特定のプレフィクス宛先への最適ルートが決定されます。RIB には、新しいルーティング エントリが入力されます。対応する CEF エントリが更新されます。
10. 更新された情報がソフトウェアの CEF データベースに入力されると、更新されたエントリには、リフレッシュされたことを示すグローバル エポック番号が付与されます。対応する FIB エントリとハードウェア エントリが更新されます。
11. 各ルーティング プロトコルは、CEF にコンバージェンスの完了を通知します。すべてのプロトコルでコンバージェンスが完了すると、最後のルーティング プロトコルは古くなったルートとの隣接情報を消去します。現在のグローバル エポック番号に該当しない番号を持つソフトウェアの CEF エントリは消去されます。対応する FIB および隣接ハードウェア エントリも消去されます。
12. これで、RP 上の Cisco IOS ソフトウェアの CEF テーブル、および SP、PFC、DFC 上のフォワーディング テーブルが同期化されました。

NSF のグレースフル リスタート ルーティング プロトコル拡張機能は、IETF ドラフトと RFC に準拠しています。NSF プロトコル固有の詳細については、「参考資料」を参照してください。

MMLS NSF/SSO

Cisco IOS ソフトウェア リリース 12.2(18)SXD 以前は、IPv4 マルチキャストのスイッチオーバー モデルは RPR+ でした。SRM/SSO が IPv4 ユニキャストで設定されていても、PFC3 ハードウェア内のマルチキャスト フォワーディング エントリはスタンバイ スーパーバイザ エンジンに同期化されません。

スイッチオーバーが発生すると、PFC3 および DFC3 (存在する場合) 上のマルチキャスト フォワーディング エントリは消去され、マルチキャスト トラフィックのサービスが中断されます。新しいアクティブ RP がオンラインになったあとで、その RP で PFC3 および DFC3 のフォワーディング エンジン内のハードウェア フォワーディング エントリへの再入力を行うには、Protocol Independent Multicast (PIM) のネイバ関係を確立し、Internet Group Management Protocol (IGMP) パケットを処理するか、またはマルチキャスト状態の再コンバージェンスを行う必要があります。

MMLS NSF/SSO によって、システムはスーパーバイザ エンジンのスイッチオーバー中に PFC3 および DFC3 ハードウェア内のマルチキャスト フォワーディングの状態を維持できるので、マルチキャスト サービスの中断を最小限に抑えることができます。

準備完了の状態にあるアクティブ スーパーバイザ エンジンは、スタンバイ スーパーバイザ エンジンをハードウェアのマルチキャスト フォワーディング エントリで同期化します。スーパーバイザ エンジンのスイッチオーバーが発生しても、PFC3 および DFC3 のハードウェア フォワーディング テーブルは保持され、システムはマルチキャスト フォワーディング テーブルの有効な最新コピーを使用して、マルチキャスト トラフィックの転送を継続します。

新しいアクティブ RP がオンラインになり、ネットワークでコンバージェンスが行われ、マルチキャスト転送の状態が再取得されると、その RP によって PFC3 および DFC3 上のハードウェア フォワーディング テーブルに新しい情報が再入力されます。

SSO と NSF/SSO の機能

表 2～5 は、SSO と NSF/SSO 用に同期化されるレイヤ 2、レイヤ 3、WAN、およびハードウェアの機能を説明しています。「共存」として指定される機能は SSO および NSF/SSO と併用することが可能ですが、そのプロトコルの状態はアクティブ スーパーバイザからスタンバイ スーパーバイザには同期化されず、スイッチオーバーが発生した時点で再初期化されます。リリースの詳細は、Cisco Catalyst 6500 のリリース ノートに記載されています。

レイヤ 2 機能の同期化

表 2 は、レイヤ 2 SSO 用に同期化されるレイヤ 2 プロトコルについて説明しています。記載されている機能は、Cisco IOS ソフトウェア リリース 12.2(17b)SXA を搭載した Supervisor Engine 720、および Cisco IOS ソフトウェア リリース 12.2(18)SXD を搭載した Supervisor Engine 2 で使用できます。

表 2 レイヤ 2 SSO でサポートされている機能

レイヤ 2 SSO でサポートされている機能	
Cisco Discovery Protocol (CDP) (電力関連の情報についてのみステータス)	ポート セキュリティ
診断	Switched Port Analyzer (SPAN; スイッチド ポート アナライザ) と Remote SPAN (RSPAN)
802.1q	STP
802.1X	トラフィック ストーム
Dynamic Trunking Protocol (DTP; ダイナミック トランキング プロトコル)	UDLD
IGMP スヌーピング	VLAN トランク
インターフェイスとポート ステート	VLAN Trunking Protocol (VTP; VLAN トランキング プロトコル)
レイヤ 2 プロトコル トンネリング	音声 VLAN とインライン パワー
ポート チャネリング : Port Aggregation Protocol (PAgP; ポート集約プロトコル) と Link Aggregation Control Protocol (LACP)	

レイヤ 3 機能の同期化

表 3 は、SRM/SSO および NSF/SSO 用に同期化されるレイヤ 3 プロトコルについて説明しています。Enhanced Interior Gateway Routing Protocol (EIGRP)、Open Shortest Path First (OSPF)、IS-IS、および Border Gateway Protocol (BGP) 用の NSF 機能は、Cisco IOS ソフトウェア リリース 12.2(18)SXD を搭載した Supervisor Engine 720 と Supervisor Engine 2 で使用できます。Cisco IOS ソフトウェア リリース 12.2(18)SXD で導入されている MMLS NSF/SSO は、Supervisor Engine 720 でのみ使用できます。

表 3 レイヤ 3 機能の同期化

レイヤ 3 情報の同期化 NSF/SSO でサポートされている機能	SRM/SSO でサポートされている機能	共存機能
ARP	ARP	HSRP、GLBP、および Virtual Router Redundancy Protocol (VRRP)
BGP	ハードウェアの CEF テーブル	Multiprotocol Label Switching (MPLS; マルチプロトコル ラベル スイッチング)、 Internetwork Packet Exchange (IPX)、および IPv6
CEF テーブル (ソフトウェアおよびハードウェア)		PIM スヌーピング
EIGRP		Routing Information Protocol (RIP)
IS-IS		
MMLS NSF/SSO		
OSPFv2		
VRF Lite		

WAN 機能の同期化

表 4 は、SSO 用に同期化される WAN プロトコルについて説明しています。記載されている機能は、Cisco IOS ソフトウェア リリース 12.2(18)SXD で使用できます。

表 4 WAN 機能の同期化

レイヤ 2 SSO でサポートされている機能 SSO でサポートされている機能	共存機能
Asynchronous Transfer Mode (ATM; 非同期転送モード)	Distributed Link Fragmentation and Interleaving over ATM (dLFloATM) dLFI over Frame Relay (dLFloFR)
Automatic Protection Switching (APS; 自動保護スイッチング)	Multilink PPP (MLPPP) と Multilink Frame Relay (MFR)
フレーム リレー	IP ヘッダー圧縮
High-Level Data Link Control (HDLC; ハイレベル データリンク制御)	MPLS と Any Transport over MPLS (AToM)
PPP (ポイントツーポイント プロトコル)	WAN カード用 QoS
Spatial Reuse Protocol (SRP)	
ATM	dLFloATMdLFloFR
APS	MLPPP と MFR

ハードウェアのレイヤ 2 ~ 4 機能の同期化

表 5 は、SSO スイッチオーバーの発生時にスタンバイ スーパーバイザで利用できる PFC 情報について説明しています。記載されている機能の大部分は、Cisco IOS ソフトウェア リリース 12.2(17b)SXA を搭載した Supervisor Engine 720、および Cisco IOS ソフトウェア リリース 12.2(18)SXD を搭載した Supervisor Engine 2 で使用できます。

表 5 ハードウェアのレイヤ 2 ~ 4 機能の同期化

ハードウェアの L2 ~ 4 機能の同期化	
ハードウェアの ACL ベースの機能	ハードウェアの FIB
ハードウェアの隣接テーブル	ハードウェアの MAC アドレス テーブル
ハードウェアの IP マルチキャスト情報	ハードウェアの QoS ベースの機能

スイッチオーバーのパフォーマンス

NSF/SSO のフェールオーバー時間

図 5 は、NSF/SSO のパフォーマンス テスト用のセットアップを図示したものです。

図 5 NSF/SSO のパフォーマンス テスト

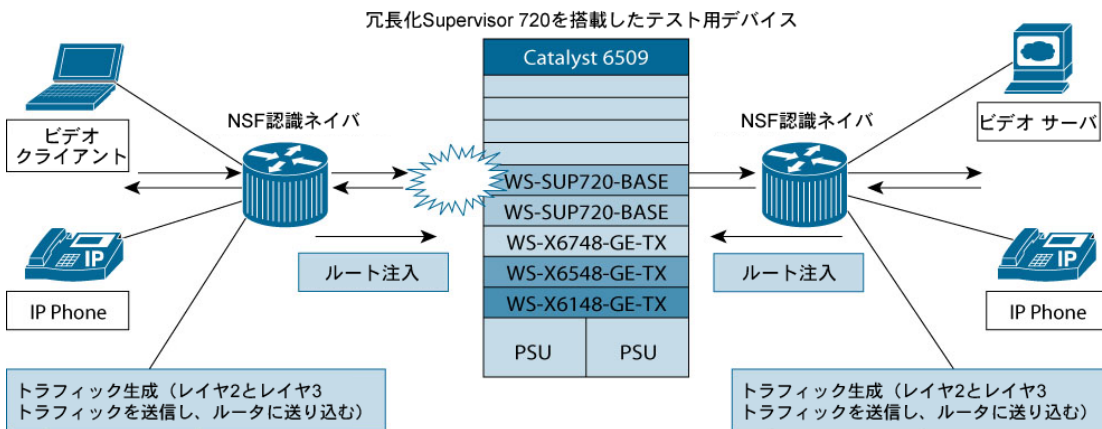


図 5 のようなネットワーク構成で、Cisco Catalyst 6500 での NSF/SSO のパフォーマンスを測定しました。このセットアップには、各種動作モードに対応するフェールオーバー時間を記録するためのシミュレーション デバイスが組み込まれています。また、実際に使用されるアプリケーションが NSF/SSO スイッチオーバーによって影響を受けないかどうかを確認するため、アプリケーションのトラフィックも流れています。テストしたアプリケーションは映像および VoIP アプリケーションなどです。

テスト用セットアップは、3 台の Cisco Catalyst 6500 スイッチで構成されます。このセットアップの全スーパーバイザには、Cisco IOS 12.2(18)SXD がロードされています。テスト用デバイスの Cisco Catalyst 6500 には、近接ルータからのレイヤ 2 およびレイヤ 3 トラフィックをスイッチングする冗長構成の Supervisor 720 が搭載されています。この双方向トラフィックには、トラフィック シミュレータによって 100,000 pps で生成されるレイヤ 2 およびレイヤ 3 トラフィック、IP Phone および映像クライアント / サーバから送信される音声および映像トラフィックが含まれます。テスト手順は、RPR+、SSO (NSF 機能は無効)、および NSF/SSO の各フェールオーバー メカニズムについて実行されるレイヤ 2 およびテスト 3 テストで構成されます。レイヤ 3 テスト ランは、OSPF、EIGRP、IS-IS、および BGP 用に挿入した 1000 のルートで実施されました。ネイバはすべて NSF 認識機能を持っています。

フェールオーバー時間は、スイッチオーバー中に送信されたパケット数と受信されたパケット数を比較することによって算出されます。パケット レートが指定されている場合 (テスト ランでは 100,000 pps)、フェールオーバー時間は (送信パケット数 - 受信パケット数) / パケット レートとなります。

表 6 は、トラフィックが WS-X6748-GE-TX モジュールの 2 つのポート間を伝送されているときのフェールオーバー時間を、各種のシナリオに基づいて示したものです。全体として、NSF/SSO ではフェールオーバー時間はテスト条件に応じて 0 ~ 3 秒の範囲になります。

表 6 NSF/SSO のフェールオーバー時間

フェールオーバー時間	レイヤ 2 トラフィック	EIGRP でルーティングされるレイヤ 3 トラフィック	OSPF でルーティングされるレイヤ 3 トラフィック	IS-IS でルーティングされるレイヤ 3 トラフィック (シスコ方式)	IS-IS でルーティングされるレイヤ 3 トラフィック (IETF 方式)	BGP でルーティングされるレイヤ 3 トラフィック
RPR+	62.00 秒	70.00 秒	140.00 秒	82.00 秒	82.00 秒	130.00 秒
SSO (NSF 機能は無効)	0.50 秒	6.00 秒	11.00 秒	0.55 秒	0.55 秒	54.00 秒
NSF/SSO	0.50 秒	0.55 秒	0.55 秒	0.55 秒	0.55 秒	0.55 秒

Cisco Catalyst OS と Cisco IOS ソフトウェアのハイアベイラビリティの比較

表 7 は、Cisco Catalyst OS、ハイブリッド、および Cisco IOS ソフトウェアのスイッチオーバーのパフォーマンス値を、同等の機能について比較したものです。Cisco Catalyst OS のスーパーバイザの冗長メカニズムの詳細については、次の URL を参照してください。

http://www.cisco.com/warp/public/cc/pd/si/casi/ca6000/tech/hafc6_wp.pdf

Cisco Catalyst OS の高速スイッチオーバー機能は、Cisco IOS ソフトウェアの RPR+ 機能に相当します。Cisco Catalyst OS のハイアベイラビリティ機能は、Cisco IOS ソフトウェアの SSO 機能に相当します。両機能ともレイヤ 2 プロトコルの同期化とレイヤ 2 ~ 4 のハードウェア同期化を提供しています。

SRM 冗長方式によるハイブリッドのハイアベイラビリティは、SSO のハイブリッドに相当します。ハイブリッド モデルでは、NSF 機能はサポートされていません。

表 7 は、Cisco Catalyst OS、ハイブリッド、および Cisco IOS ソフトウェアの冗長機能のパフォーマンス値を比較したものです。

表 7 Cisco Catalyst OS と Cisco IOS ソフトウェアのスーパーバイザの冗長機能

Cisco Catalyst OS	ハイブリッド	Cisco IOS ソフトウェア
-	-	RPR : > 120.00 秒
高速スイッチオーバー : > 30.00 秒	高速スイッチオーバー : > 30.00 秒	RPR+ : 30.00 秒
ハイアベイラビリティ : 0.50 ~ 5.00 秒	SRM でのハイアベイラビリティ : 0.50 ~ 5.00 秒	SSO : 0.00 ~ 3.00 秒
ハイアベイラビリティ : 0.50 ~ 5.00 秒	SRM でのハイアベイラビリティ : 0.50 ~ 5.00 秒	NSF/SSO : 0.00 ~ 3.00 秒
ハイアベイラビリティバージョンング	ハイアベイラビリティバージョンング	FSU

統計情報と SNMP

統計情報

アクティブ スーパーバイザで保持されている各種の統計情報は、スタンバイ スーパーバイザには同期化されません。統計情報は頻繁に変化するため、相当な規模の同期が必要となるからです。影響を受ける統計情報を定期的にポーリングして正確さを維持するには、ネットワーク管理システムを使用する必要があります。

SNMP

SNMP（簡易ネットワーク管理プロトコル）のデータは、スーパーバイザが SSO モードで動作しているときは冗長化したスーパーバイザ間で同期化されます。これは、スタンバイスーパーバイザとアクティブスーパーバイザがネットワーク管理システムから見て区別できないことを保証するためです。同期化される SNMP オブジェクトには、ifindex などのインターフェイス関連の機能や SNMP 設定などがあります。

Cisco High-Availability MIB である CISCO-RF_MIB は、冗長情報を管理者にレポートします。この情報には、プライマリスーパーバイザとセカンダリスーパーバイザの ID、現在の冗長性の状態、および最新のスイッチオーバーの発生原因と発生時間が含まれます。スイッチオーバーが発生すると、ciscoRFSwactNotif 通知によってスイッチオーバーの信号が送られます。

Cisco-RF MIB が使用されるほか、Syslog メッセージと SNMP トラップが送信され、管理者にコンポーネントの障害が通知されます。

SNMP データの同期化は、RPR および RPR+ 動作モードでは利用できません。

SSO の SNMP サポートの詳細については、次の URL を参照してください。（英語）

<http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120limit/120s/120s26/ssomibs3.htm>

サービス モジュールでのスーパーバイザの SSO 機能のサポート

NSF/SSO によるスーパーバイザのフェールオーバー イベントの間、サービス モジュールの動作が継続されるのは重要なことです。現在のサービス モジュールの多くは特定のハイアベイラビリティメカニズムを備えており、シャーシ内またはシャーシ間でモジュール間のスイッチオーバーを実行できます。サービス モジュールでのスーパーバイザの NSF/SSO のサポートは、スーパーバイザのフェールオーバーの影響を最小限に抑えることによって、各サービス モジュールのハイアベイラビリティメカニズムを補完します。

標準のスイッチング モジュールの SSO に関連する各特性は、サービス モジュールにも当てはまります。サービス モジュールは再起動せず、そのインターフェイスは動作を継続します。また、サービス モジュールはラインカードが同期化される短い時間を除いて、スーパーバイザのスイッチオーバーの影響を受けません。

サービス モジュールの SSO モード互換性については、次のリリース ノートで確認してください。（英語）

http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/122sx/ol_4164.htm

WAN モジュールでのスーパーバイザの SSO のサポート

OSM（オプティカル サービス モジュール）と FlexWAN モジュールはスタンバイスーパーバイザエンジンでサポートされており、NSF/SSO によるスーパーバイザのフェールオーバー イベントの間も動作を継続します。

参考資料

ハイアベイラビリティ：

ハイアベイラビリティに関する技術文書：

<http://www.cisco.com/warp/public/732/Tech/grip/splash/>

Cisco Catalyst 6500 のハイアベイラビリティ：

Cisco Catalyst 6500 シリーズ スイッチのハイアベイラビリティ（Catalyst OS のハイアベイラビリティ）：

http://www.cisco.com/warp/public/cc/pd/si/casi/ca6000/tech/hafc6_wp.pdf

スーパーバイザエンジンの冗長性の設定：

http://www.cisco.com/jp/service/manual_j/sw/cat60/iosscg/chapter07/07_redund.shtml

NSF/SSO を使用したスーパーバイザ エンジンの冗長性の設定 :

<http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/122sx/swcg/nsfsso.htm>

Cisco IOS ソフトウェア リリース 12.2SX のリリース ノート (英語) :

http://www.cisco.com/univercd/cc/td/doc/product/lan/cat6000/122sx/ol_4164.htm

Cisco Catalyst 6500 シリーズ スイッチのマニュアル :

http://www.cisco.com/jp/service/manual_j/index_sw_cat6500.shtml

ステートフル スイッチオーバー (英語) :

SNMP SSO :

<http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120limit/120s/120s26/ssomibs3.htm>

NSF/SSO :

NSF :

<http://www.cisco.com/univercd/cc/td/doc/product/software/ios122s/122snwft/release/122s20/fsnsf20s.htm>

IETF のルーティング プロトコル拡張機能 (英語) :

<http://www.ietf.org/>

©2005 Cisco Systems, Inc. All rights reserved.

Cisco、Cisco Systems、および Cisco ロゴは米国およびその他の国における Cisco Systems, Inc. の商標または登録商標です。
この文書で説明した商品、サービスはすべて、それぞれの所有者の商標、サービスマーク、登録商標、登録サービスマークです。
この資料に記載された仕様は予告なく変更する場合があります。



シスコシステムズ株式会社

URL: <http://www.cisco.com/jp/>

問合せ URL: <http://www.cisco.com/jp/go/contactcenter/>

〒 107-0052 東京都港区赤坂 2-14-27 国際新赤坂ビル東館

TEL: 03-6670-2992

電話でのお問合せは、以下の時間帯で受付けております。

平日 10:00 ~ 12:00 および 13:00 ~ 17:00

お問合せ先