



CHAPTER 1

概要

この章では、Cisco NX-OS でのレイヤ 3 ユニキャスト ルーティング プロトコルの基盤となる概念を紹介しします。

この章では、次の内容について説明します。

- 「レイヤ 3 ユニキャスト ルーティングについて」 (P.1-1)
- 「ルーティング アルゴリズム」 (P.1-8)
- 「レイヤ 3 仮想化」 (P.1-10)
- 「Cisco NX-OS 転送アーキテクチャ」 (P.1-11)
- 「N7K-F132-15 モジュールとのレイヤ 3 相互運用」 (P.1-13)
- 「レイヤ 3 ユニキャスト ルーティング機能のまとめ」 (P.1-14)
- 「関連資料」 (P.1-17)

レイヤ 3 ユニキャスト ルーティングについて

レイヤ 3 ユニキャスト ルーティングには、最適なルーティングパスの決定とパケットの交換という、2つの基本的動作があります。ルーティング アルゴリズムを使用すると、ルータから宛先までの最適パス（経路）を計算できます。この計算方法は、選択したアルゴリズム、ルート メトリック、そしてロード バランシングや代替パスの探索などの考慮事項により異なります。

ここでは、次の内容について説明します。

- 「ルーティングの基本」 (P.1-2)
- 「パケット交換」 (P.1-2)
- 「ルーティング メトリック」 (P.1-3)
- 「ルータ ID」 (P.1-5)
- 「自律システム」 (P.1-5)
- 「コンバージェンス」 (P.1-6)
- 「ロード バランシングおよび等コスト マルチパス」 (P.1-6)
- 「ルートの再配布」 (P.1-6)
- 「アドミニストレーティブ ディスタンス」 (P.1-7)
- 「スタブ ルーティング」 (P.1-7)

ルーティングの基本

ルーティング プロトコルは、**メトリック**を使用して、宛先までの最適パスを調べます。メトリックとは、パス帯域幅などの、ルーティング アルゴリズムが宛先までの最適パスを決定するために使用する測定基準です。パスを決定しやすいように、ルーティング アルゴリズムは、ルート情報 (IP 宛先アドレスや、次のルータまたは**ネクスト ホップ**のアドレスなど) を含むルーティング テーブルを初期化して維持します。宛先とネクスト ホップの関連付けにより、ルータは、宛先までの途中にあるネクスト ホップとなる特定のルータにパケットを送信すると、最適パスで IP 宛先まで届けられることを判定できます。ルータは、着信パケットを受信すると、宛先アドレスをチェックし、このアドレスをネクスト ホップと関連付けようとします。ルート テーブルの詳細については、「**ユニキャスト RIB**」(P.1-11) を参照してください。

ルーティング テーブルには、パスの優先度に関するデータなどのその他の情報も含まれる場合があります。ルータはメトリックを比較して、最適ルートを決定します。また、これらのメトリックは、使用されるルーティング アルゴリズムの設計により異なります。「**ルーティング メトリック**」(P.1-3) を参照してください。

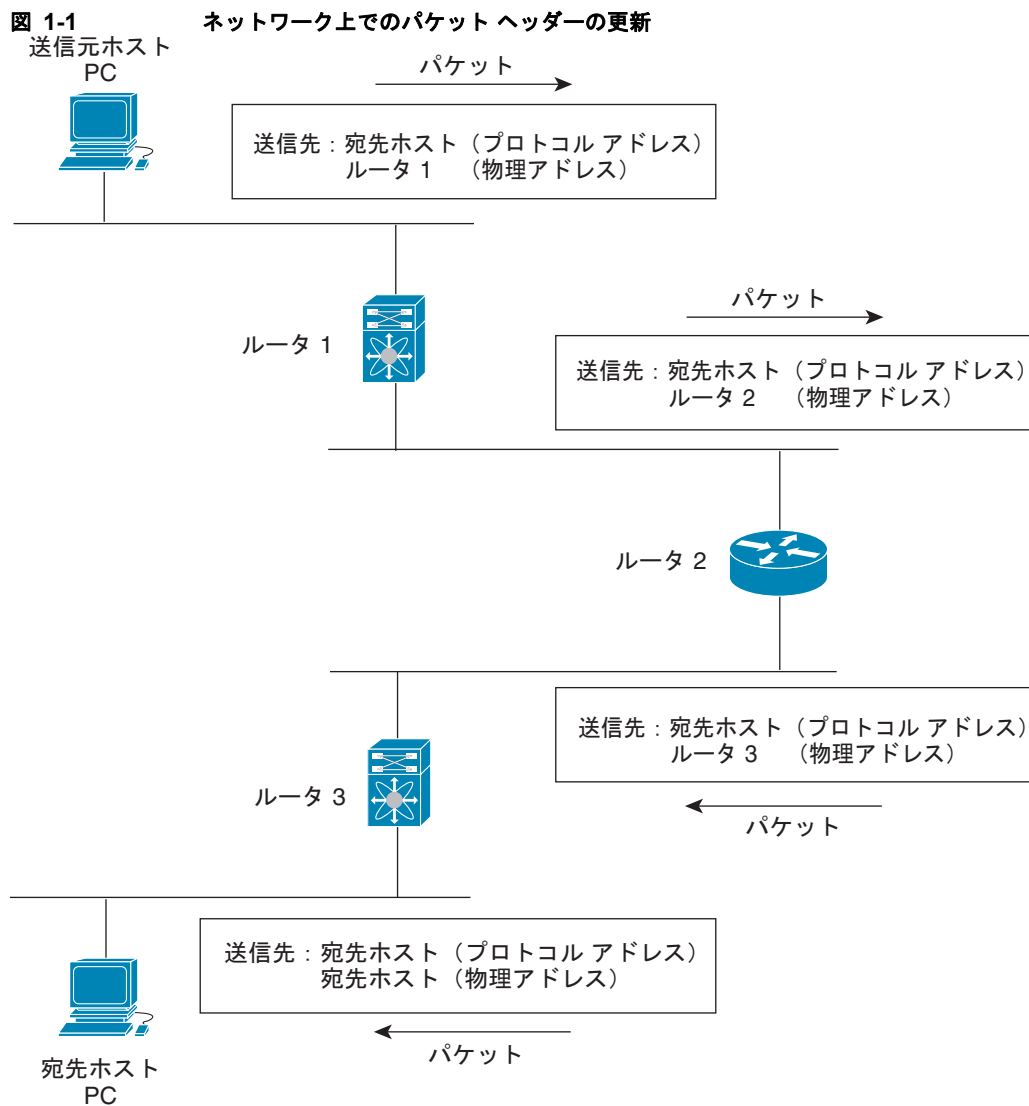
各ルータは互いに通信し、さまざまなメッセージを送信して、そのルーティング テーブルを維持します。ルーティング更新メッセージは、ルーティング テーブルの全部または一部で構成されるメッセージです。ルータは、他のすべてのルータからのルーティング更新情報を分析して、ネットワーク トポロジの詳細な図を構築できます。ルータ間で送信されるメッセージの別の例であるリンクステートアドバタイズメントは、他のルータに送信側ルータのリンク状態を通知します。リンク情報を使用して、ルータが、ネットワーク宛先までの最適ルートを決定できるようにすることもできます。詳細については、「**ルーティング アルゴリズム**」(P.1-8) を参照してください。

パケット交換

パケット交換では、ホストが、パケットを別のホストに送信する必要があることを決定します。何らかの手段でルータ アドレスを取得したら、送信元ホストは、明確にルータの物理 (メディア アクセス コントロール (MAC) レイヤ) アドレスにアドレス指定されているが、宛先ホストの IP (ネットワーク層) アドレスを含むパケットを送信します。

ルータは宛先の IP アドレスを調べ、ルーティング テーブルでその IP アドレスを探します。ルータでパケットの転送方法がわからない場合は、パケットは通常、廃棄されます。パケットの転送方法がわかった場合、ルータは、宛先の MAC アドレスをネクストホップルータの MAC アドレスに変更し、パケットを送信します。

ネクスト ホップが宛先のホストである場合や、同じ交換決定処理を行う別のルータである場合があります。パケットがネットワーク間を移動するにつれ、その物理アドレスは変更されますが、そのプロトコルアドレスは変わりません (図 1-1 を参照)。



182978

ルーティング メトリック

ルーティング アルゴリズムは、多くの異なるメトリックを使用して最適ルートを決定します。高度なルーティング アルゴリズムは、複数のメトリックに基づいてルートを選択している場合があります。

ここでは、次のメトリックについて説明します。

- 「パス長」 (P.1-4)
- 「信頼性」 (P.1-4)
- 「ルーティング遅延」 (P.1-4)
- 「帯域幅」 (P.1-4)
- 「負荷」 (P.1-4)
- 「通信コスト」 (P.1-4)

パス長

パス長 は、最も一般的なルーティング メトリックです。一部のルーティング プロトコルでは、各ネットワーク リンクに恣意的なコストの割り当てが可能です。この場合、パスの長さは、経由した各リンクに関連付けられたコストの合計となります。その他のルーティング プロトコルでは、パケットが送信元から宛先までに経由する必要がある、ルータなどのネットワーク間製品の通過の回数を指定するメトリックであるホップ数が定義されます。

信頼性

ルーティング アルゴリズムとの関連における**信頼性**は、各ネットワーク リンクの信頼性（ビット誤り率で示される）です。一部のネットワーク リンクは、他のネットワーク リンクよりダウンする頻度が高い場合があります。ネットワークがダウンしたあと、特定のネットワーク リンクが他のリンクより容易に、または短時間に修復される場合もあります。信頼性のランクを割り当てるときに考慮できる信頼性係数は、一般的にネットワーク リンクに割り当てる任意の数値です。

ルーティング遅延

ルーティングは、**遅延** 送信元から宛先に、インターネットワークを通過してパケットを移動するために必要な時間の長さです。遅延は、中間ネットワーク リンクの帯域幅、経由する各ルータでのポートキュー、中間の全ネットワーク リンクでのネットワークの輻輳状況、パケットが移動する必要がある物理的な距離など、多くの要素に応じて異なります。ルーティング遅延はいくつかの重要な変数の組み合わせであるため、一般的で便利なメトリックです。

帯域幅

帯域幅 は、リンクで使用可能なトラフィック容量です。たとえば、10 ギガビット イーサネット リンクは 1 ギガビット イーサネット リンクより優れています。帯域幅は、リンクで達成可能な最大スループットですが、帯域幅のより大きいリンクを経由するルートが、帯域幅のより小さいリンクを経由するルートより優れているとは限りません。たとえば、帯域幅の大きいリンクの方が混雑していると、実際には、パケットを宛先に送信するためにさらに長い時間がかかる場合があります。

負荷

負荷 は、ルータなどのネットワーク リソースが使用状況の程度です。負荷は、CPU 使用率や 1 秒あたりに処理されるパケット数などのさまざまな方法で計算できます。これらのパラメータを継続的にモニターすると、リソースに負担がかかる場合があります。

通信コスト

通信コスト は、リンク上でルーティングするための稼働コストの測定単位です。通信コストは重要なメトリックの 1 つで、特にパフォーマンスより稼働コストの削減が優先される場合に使用されます。たとえば、専用回線での回線遅延が公衆回線より大きくても、使用時間に応じて課金される公衆回線上でなく、自身の専用回線上でパケットを送信できます。

ルータ ID

各ルーティング処理には、**ルータ ID** が関連付けられています。ルータ ID は、システムのあらゆるインターフェイスに設定できます。ルータ ID を設定しないと、Cisco NX-OS が次の基準に基づいて、ルータ ID を選択します。

- Cisco NX-OS は、他のあらゆるインターフェイス上で **loopback0** を優先します。loopback0 が存在しない場合、Cisco NX-OS は、他のあらゆるインターフェイス タイプ上で最初のループバックを優先します。
- ループバック インターフェイスを設定していない場合、Cisco NX-OS はコンフィギュレーション ファイル内の最初のインターフェイスをルータ ID として使用します。Cisco NX-OS がルータ ID を選択したあとにいずれかのループバック インターフェイスを設定した場合は、ループバック インターフェイスがルータ ID となります。ループバック インターフェイスが loopback0 でない場合に、loopback0 を IP アドレスで設定した場合は、ルータ ID が loopback0 の IP アドレスに変更されます。
- ルータ ID の元であるインターフェイスが変更されると、新しい IP アドレスがルータ ID となります。他のどのインターフェイスの IP アドレスが変更されても、ルータ ID はまったく変更されません。

自律システム

自律システム (AS) とは、単一の技術的管理エンティティにより制御されるネットワークです。AS により、グローバルな外部ネットワークが個々のルーティング ドメインに分割され、これらのドメインでは、ローカルのルーティング ポリシーが適用されます。この構成により、ルーティング ドメインの管理と一貫したポリシー設定が簡素化されます。

各 AS は、ルート **再配布** により動的にルーティング情報を交換する、複数の内部ルーティング プロトコルをサポートできます。地域インターネット レジストリにより、インターネットに直接接続する各公共 AS に一意の番号が割り当てられます。この AS 番号で、ルーティング処理と AS の両方が識別されます。

Cisco NX-OS は 4 バイト AS 番号をサポートしています。表 1-1 は、AS 番号の範囲を示します。

表 1-1 AS 番号

2 バイト番号	AS ドット表記での 4 バイト番号	プレーンテキスト表記での 4 バイト番号	目的
1 ~ 64511	0.1 ~ 0.64511	1 ~ 64511	公共 AS (RIR により割り当てられる) ¹
64512 ~ 65534	0.64512 ~ 0.65534	64512 ~ 65534	専用 AS (ローカルの管理者により割り当てられる)
65535	0.65535	65535	予約済み
N/A	1.0 ~ 65535.65535	65536 ~ 4294967295	公共 AS (RIR により割り当てられる)

1. RIR = 地域インターネット レジストリ (Regional Internet Registries)

専用 AS 番号は内部ルーティング ドメインに使用されますが、インターネット上にルーティングされたトラフィック向けに、ルータにより変換される必要があります。ルーティング プロトコルを、専用 AS 番号が外部ネットワークにアダプタイズされるように設定しないでください。デフォルトでは、Cisco NX-OS は専用 AS 番号をルーティング更新情報から削除しません。



(注)

公共ネットワークおよび専用ネットワークの AS 番号は、Internet Assigned Number Authority (IANA; インターネット割り当て番号局) により管理されています。自律システム番号 (予約された番号の割り当てを含む) については、または自律システム番号の登録を申請するには、次の URL を参照してください。

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

コンバージェンス

ルーティング アルゴリズム測定の際となる要素の 1 つは、ルータがネットワーク トポロジの変化に対応するために要する時間です。リンク障害など、なんらかの理由でネットワークの一部が変化すると、さまざまなルータのルーティング情報が一致しなくなる場合があります。変化したトポロジに関する情報が更新されているルータと、一方では古い情報が残っているルータがあるためです。**コンバージェンス**は、ネットワーク内のすべてのルータが更新され、ルーティング情報が一致するまでにかかる時間の長さです。コンバージェンス時間は、ルーティング アルゴリズムによって異なります。コンバージェンスが速い場合は、不正確なルーティング情報によるパケット損失の可能性が小さくなります。

ロード バランシングおよび等コスト マルチパス

ルーティング プロトコルでは、**ロード バランシング**または等コスト マルチパス (ECMP) を使用して、複数のパス上のトラフィックを共有できます。ルータは、特定のネットワークへのルートを複数検出すると、最もアドミニストレーティブ ディスタンスの低いルートをルーティング テーブルにインストールします。ルータが、同じアドミニストレーティブ ディスタンスと宛先までのコストを持つ複数のパスを受信し、インストールすると、ロード バランシングが発生する場合があります。ロード バランシングでは、すべてのパス上にトラフィックが配布され、負荷が共有されます。使用されるパスの数は、ルーティング プロトコルによりルーティング テーブルに配置されるエントリの数に制限されます。Cisco NX-OS は、宛先までの 16 のパスをサポートします。

Enhanced Interior Gateway Routing Protocol (EIGRP) は、等コストでないロード バランシングもサポートしています。詳細については、**第 8 章「EIGRP の設定」**を参照してください。

ルートの再配布

ネットワークに複数のルーティング プロトコルが設定されている場合は、各プロトコルでルート再配布を設定して、ルーティング情報を共有するように設定できます。たとえば、**Open Shortest Path First (OSPF)** プロトコルを、**ボーダー ゲートウェイ プロトコル (BGP)** から学習されたルートをアドバタイズするように設定できます。また、**スタティック ルート**を、どの**ダイナミック ルーティング プロトコル**にも再配布できます。別のプロトコルからのルートを再配布しているルータは、これらの再配布されるルートに固定のルート メトリックを設定するため、異なるルーティング プロトコル間で互換性のないルート メトリックが防止されます。たとえば、**EIGRP** から **OSPF** に再配布されたルートには、**OSPF** が認識できる固定リンク コスト メトリックが割り当てられます。



(注)

ルーティング情報の再配布を設定する場合は、ルート マップを使用する必要があります。

ルート再配布では、アドミニストレーティブディスタンス（「アドミニストレーティブディスタンス」(P.1-7)を参照）の使用によっても、2つの異なるルーティングプロトコルで検出されたルートが区別されます。優先ルーティングプロトコルには、より低いアドミニストレーティブディスタンスが与えられており、そのルートが、より高いアドミニストレーティブディスタンスが割り当てられた他のプロトコルからのルートに優先して選択されます。

アドミニストレーティブディスタンス

アドミニストレーティブディスタンスは、ルーティング情報の送信元の信頼性のランクです。値が大きいほど、信頼性のランク付けは低くなります。一般的にルートは、複数のプロトコルを通じて検出されます。アドミニストレーティブディスタンスは、複数のプロトコルで検出されたルートを区別するために使用されます。最もアドミニストレーティブディスタンスが低いルートがIPルーティングテーブルにインストールされます。

スタブルーティング

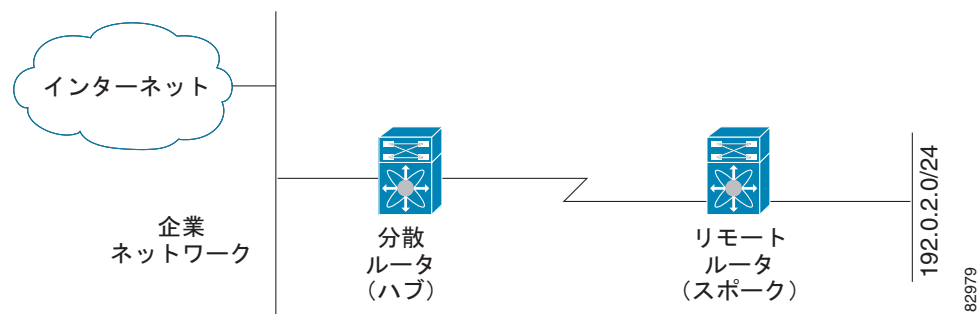
スタブルーティングはハブアンドスポーク型ネットワークトポロジで使用できます。このトポロジでは、1つ以上の終端（スタブ）ネットワークが、1つ以上の分散ルータ（ハブ）に接続されたりリモートルータ（スポーク）に接続されています。リモートルータは、1つ以上の分散ルータにのみ隣接しています。リモートルータへと流れるIPトラフィックのルートは、分散ルータ経由のルートのみです。このタイプの設定は、分散ルータが直接WANに接続されているWANトポロジで使用されるのが一般的です。分散ルータは、さらに多くのリモートルータに接続できます。分散ルータが100台以上のリモートルータに接続されていることも、よくあります。ハブアンドスポーク型トポロジでは、リモートルータがすべての非ローカルトラフィックを分散ルータに転送する必要があります。これにより、リモートルータが完全なルーティングテーブルを保持する必要はなくなります。通常、分散ルータは、デフォルトのルートのみをリモートルータに送信します。

指定されたルートのみが、リモート（スタブ）ルータから伝播されます。スタブルータは、要約、接続したルート、再配布されたスタティックルート、外部ルート、内部ルートに対する照会のすべてに、「アクセスできない」メッセージで対応します。スタブとして設定されたルータは、すべての隣接ルータに特別なピア情報パケットを送信して、自身のスタブルータとしての状態を報告します。

スタブステータスを通知するパケットを受信したネイバーは、スタブルータにルートについて照会しません。また、スタブピアを持つルータはそのピアに照会しません。スタブルータは、配布ルータに依存して適切なアップデートをすべてのピアに送信します。

図 1-2 は、単純なハブアンドスポーク型設定を示します。

図 1-2 単純なハブアンドスポークネットワーク



スタブ ルーティングを使用する場合でも、リモート ルータにルータをアドバタイズできます。図 1-2 は、リモート ルータが、分散ルータのみを使用して企業ネットワークとインターネットにアクセスできることを示しています。この例では、企業ネットワークやインターネットへのパスが常に分散ルータを経由するため、リモート ルータ上の完全なルート テーブルは機能的に意味がありません。より大規模なルート テーブルを使用しても、リモート ルータに必要なメモリの量が削減されるだけです。使用される帯域幅とメモリは、分散ルータでルートを要約し、フィルタリングすると、削減できます。このネットワーク トポロジでは、リモート ルータはすべての非ローカル トラフィックを（その宛先には関係なく）分散ルータに送信する必要があるため、リモート ルータが他のネットワークから学習されたルートを受信する必要はありません。真のスタブ ネットワークを設定するには、リモート ルータへのデフォルト ルートのみを送信するよう、分散ルータを設定する必要があります。

OSPF はスタブ エリアをサポートしており、EIGRP はスタブ ルータをサポートしています。

ルーティング アルゴリズム

ルーティング アルゴリズムは、ルータが到達可能性の情報を収集し、報告する方法、トポロジの変化に対応する方法、および宛先までの最適ルートを決定する方法を決定します。ルーティング アルゴリズムにはさまざまなタイプがあり、各アルゴリズムがネットワークやルータ リソースに与える影響もさまざまです。ルーティング アルゴリズムは、最適ルートの計算に影響するさまざまなメトリックを使用します。ルーティング アルゴリズムは、スタティックまたはダイナミック、内部または外部など、タイプで分類できます。

ここでは、次の内容について説明します。

- 「スタティック ルートおよびダイナミック ルーティング プロトコル」 (P.1-8)
- 「内部および外部ゲートウェイ プロトコル」 (P.1-9)
- 「ディスタンス ベクトル プロトコル」 (P.1-9)
- 「リンクステート プロトコル」 (P.1-9)

スタティック ルートおよびダイナミック ルーティング プロトコル

スタティック ルートは、手動で設定するルート テーブル エントリです。スタティック ルートは、手動で再設定しない限り、変更されません。スタティック ルートは設計が簡単で、ネットワーク トラフィックが比較的予想しやすい環境や、ネットワーク設計が比較的単純な環境での使用に適しています。

スタティック ルーティング システムはネットワークの変化に対応できないため、絶えず変化する大規模なネットワークには使用しないでください。今日のほとんどのルーティング プロトコルは、着信ルーティング更新メッセージを分析することにより変化するネットワーク状況に合わせて調整する、ダイナミック ルーティング アルゴリズムを使用しています。メッセージがネットワークの変化を示している場合は、ルーティング ソフトウェアがルートを計算し直して、新しいルーティング更新メッセージを送信します。これらのメッセージがネットワークを通過すると、ルータがそのアルゴリズムを再実行し、それに従ってルーティング テーブルを変更します。

適切であれば、ダイナミック ルーティング アルゴリズムをスタティック ルートで補完することができます。たとえば、各サブネットワークに、IP デフォルト ゲートウェイまたはラスト リゾート ルータ（ルーティングできないすべてのパケットが送信されるルータ）へのスタティック ルートを設定する必要があります。

内部および外部ゲートウェイ プロトコル

ネットワークを、一意のルーティング ドメインまたは AS に分割できます。AS は、管理ガイドラインの特定のセットで規制された共通の管理機関の下の内部ネットワークの一部です。AS 間でのルートを設定するルーティング プロトコルは、外部ゲートウェイ プロトコルまたはドメイン間プロトコルと呼ばれます。ボーダー ゲートウェイ プロトコル (BGP) は、外部ゲートウェイ プロトコルの例です。1 つの AS 内で使用されるルーティング プロトコルは、内部ゲートウェイ プロトコルまたはドメイン内プロトコルと呼ばれます。EIGRP および OSPF は、内部ゲートウェイ プロトコルの例です。

ディスタンス ベクトル プロトコル

ディスタンス ベクトル プロトコルは **ディスタンス ベクトル** アルゴリズム (Bellman-Ford アルゴリズムとも呼ばれます) を使用します。このアルゴリズムにより、各ルータは、そのルーティング テーブルの一部または全部を隣接ルータに送信します。ディスタンス ベクトル アルゴリズムでは、ルートが、ディスタンス (宛先までのホップ数など) および方向 (ネクスト ホップ ルータなど) により定義されます。その後、これらのルートは、直接接続された隣接ルータにブロードキャストされます。各ルータは、これらの更新情報を使用して、ルーティング テーブルを確認し、更新します。

ルーティング ループを防ぐために、ほとんどのディスタンス ベクトル アルゴリズムは **ポイズン リバース** を指定した **スプリット ホライズン** を使用します。これは、インターフェイスで検出されたルートを到達不能として設定し、それをそのインターフェイスで、次の定期更新中にアドバタイズするという意味です。このプロセスにより、ルータによるルート更新が、そのルータ自体に返信されなくなります。

ディスタンス ベクトル アルゴリズムは、一定の間隔で更新を送信しますが、ルート メトリックの値の変更に応じて、更新を送信することもできます。このように送信された更新により、ルート コンバージェンス時間の短縮が可能です。Routing Information Protocol (RIP) はディスタンス ベクトル プロトコルの 1 つです。

リンクステート プロトコル

リンクステート プロトコルは、最短パス優先 (SPF) とも呼ばれ、情報をネイバー ルータと共有します。各ルータは、各リンクおよび直接接続されたネイバー ルータに関する情報を含むリンクステートアドバタイズメント (LSA) を構築します。

各 LSA にはシーケンス番号があります。ルータが LSA を受信し、そのリンクステート データベースを更新すると、その LSA はすべての隣接ネイバーにフラッディングされます。ルータが (同じルータから) 同じシーケンス番号を持つ 2 つの LSA を受信した場合、そのルータは LSA 更新ループを防止しようとするため、最後に受信した LSA をそのネイバーにフラッディングしません。ルータは LSA を受信するとすぐにフラッディングするため、リンクステート プロトコルのコンバージェンス時間は最小限になります。

隣接ルータの探索と隣接関係の確立は、リンクステート プロトコルの重要な部分です。隣接ルータは、特別な hello パケットを使用して探索されます。このパケットは、各隣接ルータのキープアライブ通知としても機能します。隣接関係は、隣接ルータ間のリンクステート プロトコルの一般的な動作パラメータ セットで確立されます。

ルータが受信した LSA は、そのルータのリンクステート データベースに追加されます。各エントリは、次のパラメータで構成されます。

- ルータ ID (LSA を構築したルータの)
- ネイバー ID
- リンク コスト

- LSA のシーケンス番号
- LSA エントリの作成時からの経過時間

ルータは、リンクステート データベース上で SPF アルゴリズムを実行し、そのルータの最短パス ツリーを構築します。この SPF ツリーを使用して、ルーティング テーブルにデータが入力されます。

リンクステート アルゴリズムでは、各ルータがそのルーティング テーブル内に、ネットワーク全体の図を構築します。リンクステート アルゴリズムが小さな更新を全体的に送信するのに対し、ディスタンス ベクトル アルゴリズムは、より大きな更新を隣接ルータのみに送信します。

リンクステート アルゴリズムは、より短時間でコンバージェンスするため、ルーティング ループが発生する可能性はディスタンス ベクトル アルゴリズムより低くなります。ただし、リンクステート アルゴリズムにはディスタンス ベクトル アルゴリズムより多くの CPU パワーとメモリが必要なため、実装とサポートがより高価になる場合があります。リンクステート プロトコルは通常、ディスタンス ベクトル プロトコルよりスケーラブルです。

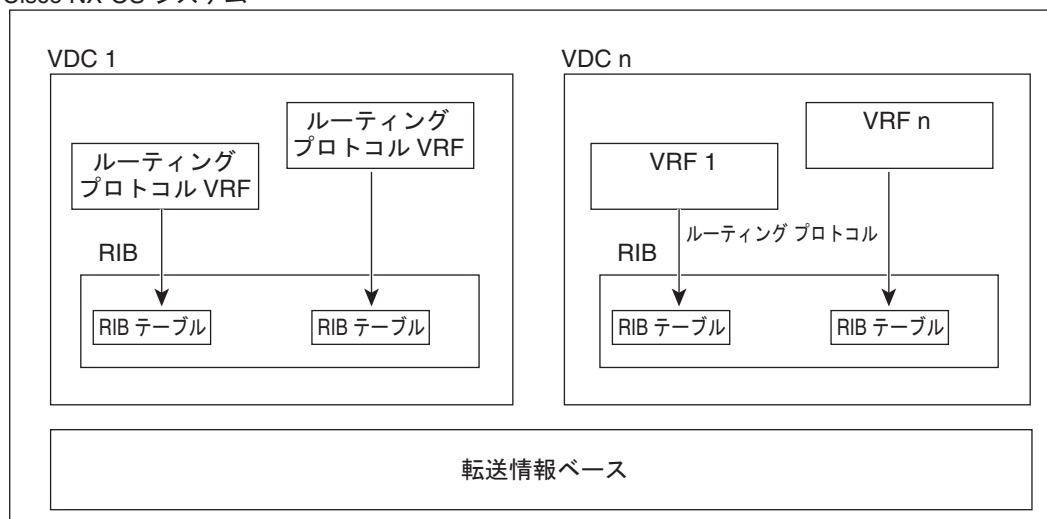
OSPF は、リンクステート プロトコルの一例です。

レイヤ 3 仮想化

Cisco NX-OS は、仮想デバイス コンテキスト (VDC) を使用して、VDC ごとに分離された管理ドメインや、ソフトウェア障害の分離機能を提供します。各 VDC は、複数の仮想ルーティング/転送インスタンスおよび複数のルーティング情報ベース (RIB) をサポートすることにより、複数のアドレスドメインをサポートします。各 VRF は RIB と関連付けられ、この情報が転送情報ベース (FIB) によって収集されます。図 1-3 は、VDC、VRF、および Cisco NX-OS デバイスの間の関係を示します。

図 1-3 レイヤ 3 仮想化の例

Cisco NX-OS システム



VRF は、レイヤ 3 アドレス指定ドメインを表します。各レイヤ 3 インターフェイス (論理または物理) は、1 つの VRF に属します。VRF は、1 つの VDC に属します。各 VDC は複数の VRF をサポートできます。詳細については、第 14 章「レイヤ 3 仮想化の設定」を参照してください。

VDC については、『Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide』を参照してください。

Cisco NX-OS 転送アーキテクチャ

Cisco NX-OS 転送アーキテクチャにより、すべてのルーティングの更新処理と、シャーシ内のすべてのモジュールへの転送情報の入力が行われます。

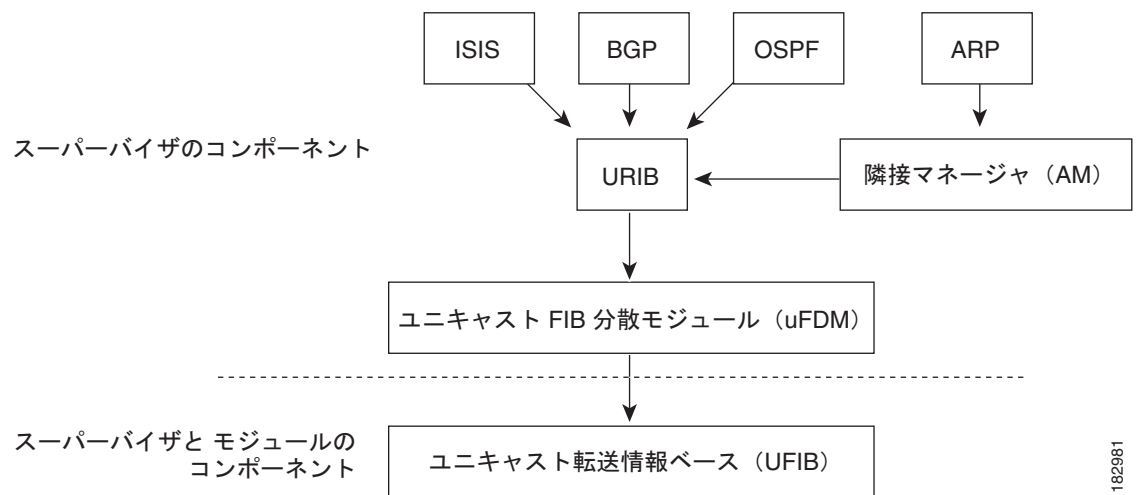
ここでは、次の内容について説明します。

- 「ユニキャスト RIB」 (P.1-11)
- 「隣接マネージャ」 (P.1-12)
- 「ユニキャスト転送分散モジュール」 (P.1-12)
- 「FIB」 (P.1-12)
- 「ハードウェア転送」 (P.1-12)
- 「ソフトウェア転送」 (P.1-13)

ユニキャスト RIB

Cisco NX-OS 転送アーキテクチャは、図 1-4 に示すように、複数のコンポーネントで構成されます。

図 1-4 Cisco NX-OS 転送アーキテクチャ



ユニキャスト RIB は、アクティブなスーパーバイザ上にあります。ユニキャスト RIB は、直接接続のルート、スタティックルート、ダイナミックユニキャストルーティングプロトコルで検出されたルートを含むルーティングテーブルを維持しています。また、Address Resolution Protocol (ARP; アドレス解決プロトコル) などの送信元から、隣接情報を収集します。ユニキャスト RIB は、特定のルートのための最適なネクストホップを決定し、ユニキャスト FIB 分散モジュール (FDM) のサービスを使用して、スーパーバイザおよびモジュール上のユニキャスト FIB にデータを入力します。

各ダイナミックルーティングプロトコルは、タイムアウトしたあらゆるルートについて、ユニキャスト RIB を更新する必要があります。その後、ユニキャスト RIB はそのルートを削除し、そのルートに最適なネクストホップを再計算します (代わりに使用できるパスがある場合)。

隣接マネージャ

隣接マネージャは、アクティブなスーパーバイザ上にあり、ARP、ネイバー探索プロトコル (NDP)、スタティック設定などのさまざまなプロトコルの隣接情報を維持します。最も基本的な隣接情報は、これらのプロトコルで探索されたレイヤ 3 からレイヤ 2 へのアドレス マッピングです。発信レイヤ 2 パケットは、隣接情報を使用して、レイヤ 2 ヘッダーの作成を終了します。

隣接マネージャは、ARP 要求による、レイヤ 3 からレイヤ 2 への特定のマッピングの探索をトリガーできます。新しいマッピングは、対応する ARP 返信を受信し、処理すると、使用できるようになります。IPv6 の場合は、隣接マネージャが NDP からの、レイヤ 3 からレイヤ 2 へのマッピング情報を検索します。詳細については、第 3 章「IPv6 の設定」を参照してください。

ユニキャスト転送分散モジュール

ユニキャスト転送分散モジュール (FDM) はアクティブなスーパーバイザ上に存在し、ユニキャスト RIB やその他の送信元からの転送パス情報を配布します。ユニキャスト RIB は、ユニキャスト FIB によってスタンプスーパーバイザおよびモジュール上のハードウェア転送テーブルにプログラミングされる転送情報を生成します。ユニキャスト FDM はまた、新規挿入されたモジュールへの FIB 情報のダウンロードも行います。

ユニキャスト FDM は隣接関係情報を収集し、ユニキャスト FIB でのルート更新時に、この情報およびその他のプラットフォーム依存の情報を書き直し (リライト) します。隣接情報およびリライト情報は、インターフェイス、ネクスト ホップ、およびレイヤ 3 からレイヤ 2 へのマッピング情報で構成されます。インターフェイスとネクスト ホップの情報は、ユニキャスト RIB からのルート更新情報で受信します。レイヤ 3 からレイヤ 2 へのマッピングは、隣接マネージャから受信します。

FIB

ユニキャスト FIB は、スーパーバイザ モジュールとスイッチング モジュール上にあり、ハードウェア転送エンジンが使用する情報を構築します。ユニキャスト FIB は、ユニキャスト FDM からルート更新を受信し、この情報を送信してハードウェア転送エンジンでプログラミングされるようにします。ユニキャスト FIB は、ルート、パス、隣接関係の追加、削除、変更を管理します。

ユニキャスト FIB は、VRF ごと、および address-family ごとに維持されます。つまり、設定された各 VRF について、IPv4 用に 1 つ、IPv6 用に 1 つ維持されます。ルート更新メッセージに基づいて、ユニキャスト FIB は、VRF ごとのプレフィクスとネクスト ホップ隣接情報データベースを維持します。ネクスト ホップ隣接データ構造には、ネクスト ホップの IP アドレスとレイヤ 2 リライト情報が含まれます。同じネクスト ホップ隣接情報構造を複数のプレフィクスで使用できます。

ハードウェア転送

Cisco NX-OS は、分散パケット転送をサポートしています。入力ポートは、パケット ヘッダーから該当する情報を取得し、その情報をローカル スwitchング エンジンに渡します。ローカル スwitchング エンジンはレイヤ 3 ルックアップを行い、この情報を使って、パケット ヘッダーをリライトします。入力モジュールは、パケットを出力ポートに転送します。出力ポートが別のモジュール上にある場合は、スイッチ ファブリックを使って、パケットが出力モジュールに転送されます。出力モジュールは、レイヤ 3 転送決定には関与しません。

スーパーバイザ上とすべてのモジュール上の転送テーブルは同じです。

また、**show platform fib** または **show platform forwarding** コマンドを使用すると、ハードウェア転送の詳細が表示されます。

ソフトウェア転送

Cisco NX-OS 内のソフトウェア転送パスは、主に、ハードウェアでサポートされていない機能进行处理したり、ハードウェア処理中に発生したエラー进行处理したりするために使用されます。通常、IP オプション付きのパケットまたはフラグメンテーションの必要なパケットは、アクティブなスーパーバイザ上の CPU に渡されます。ソフトウェアでの切り替えが必要なパケットや終端される必要のあるパケットはすべて、スーパーバイザに渡されます。スーパーバイザは、ユニキャスト RIB および隣接マネージャから提供された情報を使用して、転送の決定を下します。モジュールは、ソフトウェア転送パスには関与しません。

ソフトウェア転送は、コントロールプレーン ポリシーおよびレート リミッタによって管理されます。詳細については、『Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 6.x』を参照してください。

N7K-F132-15 モジュールとのレイヤ 3 相互運用



(注) N7K-F132-15 モジュールでレイヤ 3 ルーティングを実行するには、Cisco Nexus 7000 シリーズ シャーシ内にいずれかの N7K-M シリーズ モジュールをインストールする必要があります。同じ VDC で M シリーズと N7K-F132-15 の両方のモジュールからのインターフェイスが必要です。(VDC の詳細については、『Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide』参照してください)



(注) Cisco Nexus 7000 シリーズ シャーシでは、N7K-F132-15 モジュールでレイヤ 3 ルーティングを実行するために F2 シリーズ モジュールを使用することはできません。

N7K-F132-15 モジュールを含むシャーシ内にいずれかの N7K-M シリーズ モジュールがインストールされている場合は、レイヤ 3 ルーティング機能が自動的にアップになります。N7K-F132-15 モジュールと M シリーズ モジュールの両方を含むシャーシ、または混在シャーシは通常、レイヤ 2 ネットワークとレイヤ 3 ネットワークの間の境界に配置します。

混在シャーシ内の、プロキシルーティング機能を使用する N7K-F132-15 モジュール上で、VLAN ごとに VLAN インターフェイスを設定する必要があります。(VLAN インターフェイスの設定については、『Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide, Release 6.x』を参照してください)

デフォルトでは、VDC 内の N7K-M シリーズ モジュール上の物理インターフェイスはすべて、同じ VDC 内のレイヤ 2 専用 N7K-F132-15 モジュール上の VLAN インターフェイスが設定された VLAN に対し、プロキシルーティング ポートとして機能するようになります。M シリーズ モジュール上の物理インターフェイスは、管理上ダウンにしなが、引き続きプロキシ転送としてトラフィックを通過させることができます。

N7K-F132-15 モジュール上のインターフェイスに着信したパケットは、ルーティングされる同じ VDC 内の M シリーズ モジュール上のいずれかのインターフェイスに自動的に転送されます。また、M シリーズ モジュール上のインターフェイスでは、同じ VDC 内の N7K-F132-15 モジュール上のインターフェイスに着信したレイヤ 3 マルチキャスト パケットに対する出力レプリケーションも実行されます。

M シリーズ モジュールがすでに処理しているトラフィックに N7K-F132-15 モジュールからのレイヤ 3 (プロキシルーティング) トラフィックが追加されるため、デバイスは、VDC 内の使用可能な M シリーズ モジュールの前面パネル ポート間の合計のトラフィック負荷のためのロード バランシングを自動的に提供します。VDC 内の M シリーズ モジュールにインターフェイスを追加または削除した場合、

デバイスはトラフィックを自動的に再分散します。プロキシルーティングは M シリーズ モジュールの転送容量を共有していることに注意してください。インターフェイスを削除すると、使用可能な容量が削減されます。

M シリーズ モジュール上の自動的に設定されるプロキシルーティング インターフェイスを使用する代わりに、必要に応じて、VDC 内の M シリーズ モジュール上のどのインターフェイスがプロキシルーティングを実行するかを設定できます。

レイヤ 3 ユニキャスト ルーティング機能のまとめ

ここでは、Cisco NX-OS でサポートされるレイヤ 3 ユニキャスト機能およびプロトコルを簡単に説明します。

ここでは、次の内容について説明します。

- 「IPv4 および IPv6」 (P.1-14)
- 「IP サービス」 (P.1-14)
- 「OSPF」 (P.1-15)
- 「EIGRP」 (P.1-15)
- 「IS-IS」 (P.1-15)
- 「BGP」 (P.1-15)
- 「RIP」 (P.1-15)
- 「スタティック ルーティング」 (P.1-16)
- 「レイヤ 3 仮想化」 (P.1-16)
- 「Route Policy Manager」 (P.1-16)
- 「ポリシーベース ルーティング」 (P.1-16)
- 「ファースト ホップ冗長プロトコル」 (P.1-16)
- 「オブジェクト トラッキング」 (P.1-17)

IPv4 および IPv6

レイヤ 3 は、IPv4 プロトコルまたは IPv6 プロトコルを使用します。IPv6 は新しい IP プロトコルで、世界中で広く展開され、使用されているインターネット プロトコルである IPv4 に代わるものとして設計されました。IPv6 では、ネットワーク アドレス ビット数が 32 ビット (IPv4 の場合) から 128 ビットに増やされています。詳細については、第 2 章「IPv4 の設定」または第 3 章「IPv6 の設定」を参照してください。

IP サービス

IP サービスには、DHCP クライアントおよび Domain Name System (DNS; ドメイン ネーム システム) クライアントがあります。詳細については、第 4 章「DNS の設定」を参照してください。

OSPF

Open Shortest Path First (OSPF) プロトコルは、自律システム内のネットワーク到達可能性情報を交換するために使用されるリンクステート ルーティング プロトコルです。各 OSPF ルータは、そのアクティブなリンクに関する情報を隣接ルータにアドバタイズします。リンク情報は、リンク タイプ、リンク メトリック、およびそのリンクに接続されたネイバー ルータで構成されます。このリンク情報を含むアドバタイズメントは、リンクステート アドバタイズメントと呼ばれます。詳細については、[第 6 章「OSPFv2 の設定」](#)を参照してください。

EIGRP

Enhanced Interior Gateway Routing Protocol (EIGRP) は、ディスタンス ベクトルとリンクステートの両方のルーティング プロトコルの特長を備えたユニキャスト ルーティング プロトコルです。これは、シスコ専用ルーティング プロトコルである IGRP の改良バージョンです。EIGRP は、ルートの提供をそのネイバーに依存しています。また、リンクステート プロトコルのように、隣接ルータからアドバタイズされたルートからネットワーク トポロジを構築し、この情報を使用して、ループの発生しない、宛先までのパスを選択します。詳細については、[第 8 章「EIGRP の設定」](#)を参照してください。

IS-IS

Intermediate System-to-Intermediate System (IS-IS) プロトコルは、国際標準化機構 (ISO) 10589 で指定されたドメイン内オープン システム インターコネクション (OSI) ダイナミック ルーティング プロトコルです。IS-IS ルーティング プロトコルはリンクステート プロトコルです。IS-IS 機能は次のとおりです。

- 階層型ルーティング
- クラスレス動作
- 新情報の高速フラッディング
- 短時間でのコンバージェンス
- 高いスケールビリティ

詳細については、[第 9 章「IS-IS の設定」](#)を参照してください。

BGP

BGP は AS 間ルーティング プロトコルです。BGP ルータは、信頼性の高い転送メカニズムとして Transmission Control Protocol (TCP) を使用し、他の BGP ルータにネットワーク到達可能性情報をアドバタイズします。ネットワーク到達可能性情報には、宛先ネットワーク プレフィクス、宛先に到達するまでに通過する必要のある AS のリスト、およびネクスト ホップ ルータが含まれます。到達可能性情報には、ルートの優先度、ルートの始点、コミュニティなどの詳細なパス属性が含まれます。詳細については、[第 10 章「ベーシック BGP の設定」](#) および [第 11 章「拡張 BGP の設定」](#)を参照してください。

RIP

RIP は、ホップ数をメトリックとして使用するディスタンス ベクトル プロトコルです。RIP は、世界中のインターネットでトラフィックのルーティングに広く使用されています。また、IGP であるため、単一の AS 内でルーティングを行います。詳細については、[第 12 章「RIP の設定」](#)を参照してください。

スタティック ルーティング

スタティック ルーティングを使用して、宛先までの一定のルートを入力できます。この機能は、単純なトポロジの小規模ネットワークでは便利です。また、スタティック ルーティングは、他のルーティング プロトコルとともに、デフォルト ルートおよびルート配布の管理に使用されます。詳細については、[第 13 章「スタティック ルーティングの設定」](#)を参照してください。

レイヤ 3 仮想化

仮想化を使用すると、複数の管理ドメインにわたる物理リソースを共有できます。Cisco NX-OS は、Cisco NX-OS システム内に個別の仮想システムを作成できる仮想デバイス コンテキスト (VDC) をサポートしています。各 VDC は互いに孤立しているため、1 つの VDC 内に問題が発生しても、他のどの VDC にも影響しません。VDC はまた、相互にセキュリティが確保されています。各 VDC にそれぞれのネットワーク オペレータを割り当てることができます。これらのネットワーク オペレータは、別の VDC の設定を表示することも、管理することもできません。

また、Cisco NX-OS は、仮想ルーティング/転送 (VRF) を含むレイヤ 3 仮想化もサポートしています。VRF では、レイヤ 3 ルーティング プロトコルを設定するための個別のアドレス ドメインが提供されます。詳細については、[第 14 章「レイヤ 3 仮想化の設定」](#)を参照してください。

Route Policy Manager

Route Policy Manager は、Cisco NX-OS でルート フィルタリング機能を提供します。Route Policy Manager はルート マップを使用して、さまざまなルーティング プロトコルや、特定のルーティング プロトコル内のさまざまなエンティティ間で配布されたルートをフィルタリングします。フィルタリングは、特定の一致基準に基づいて行われます。これは、アクセス コントロール リストによるパケット フィルタリングに似ています。詳細については、[第 16 章「Route Policy Manager の設定」](#)を参照してください。

ポリシーベース ルーティング

ポリシーベース ルーティングは、Route Policy Manager を使用してポリシー ルート フィルタを作成します。これらのポリシー ルート フィルタでは、パケットの送信元またはパケット ヘッダーのその他フィールドに基づいて、指定されたネクスト ホップにパケットを転送できます。プロトコル タイプやポート番号に基づいてルーティングできるように、ポリシー ルートを拡張 IP アクセス リストにリンクすることができます。詳細については、[第 17 章「ポリシーベース ルーティングの設定」](#)を参照してください。

ファースト ホップ冗長プロトコル

ゲートウェイ ロード バランシング プロトコル (GLBP)、ホットスタンバイ ルータ プロトコル (HSRP)、仮想ルータ冗長プロトコル (VRRP) などのファースト ホップ冗長プロトコル (FHRP) を使用すると、ホストで接続の冗長性を実現できます。アクティブなファーストホップ ルータで障害が発生した場合は、処理を引き継ぐスタンバイ ルータが FHRP によって自動的に選択されます。アドレスは仮想のものであり、FHRP グループ内の各ルータ間で共有されているため、ホストを新しい IP アドレスで更新する必要はありません。GLBP の詳細については、[第 18 章「GLBP の設定」](#)を参照してください。HSRP の詳細については、[第 19 章「HSRP の設定」](#)を参照してください。VRRP の詳細については、[第 20 章「VRRP の設定」](#)を参照してください。

オブジェクト トラッキング

オブジェクト トラッキングを使用すると、インターフェイス回線プロトコル状態、IP ルーティング、ルート到達可能性などの、ネットワーク上の特定のオブジェクトをトラッキングし、トラッキングしたオブジェクトの状態が変化したときに対処することができます。この機能を使用すると、ネットワークの可用性を向上させ、オブジェクトがダウン状態となった場合の回復時間を短縮することができます。詳細については、第 21 章「オブジェクト トラッキングの設定」を参照してください。

関連資料

次のシスコ マニュアルは、レイヤ 3 機能に関連するものです。

- 『Cisco Nexus 7000 Series NX-OS Multicast Routing Configuration Guide』
- 『Cisco Nexus 7000 Series NX-OS High Availability and Redundancy Guide』
- 『Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide』
- AS 番号の詳細については、次のページを参照してください。
http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_9-1/autonomous_system_numbers.html

