



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.1-10)
- 「レイヤ 3 ユニキャスト ルーティング機能のまとめ」 (P.1-13)
- 「関連資料」 (P.1-16)

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、宛先までの最適なパスを調べます。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、および次のルータまたは**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もう 1 つのメッセージ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。

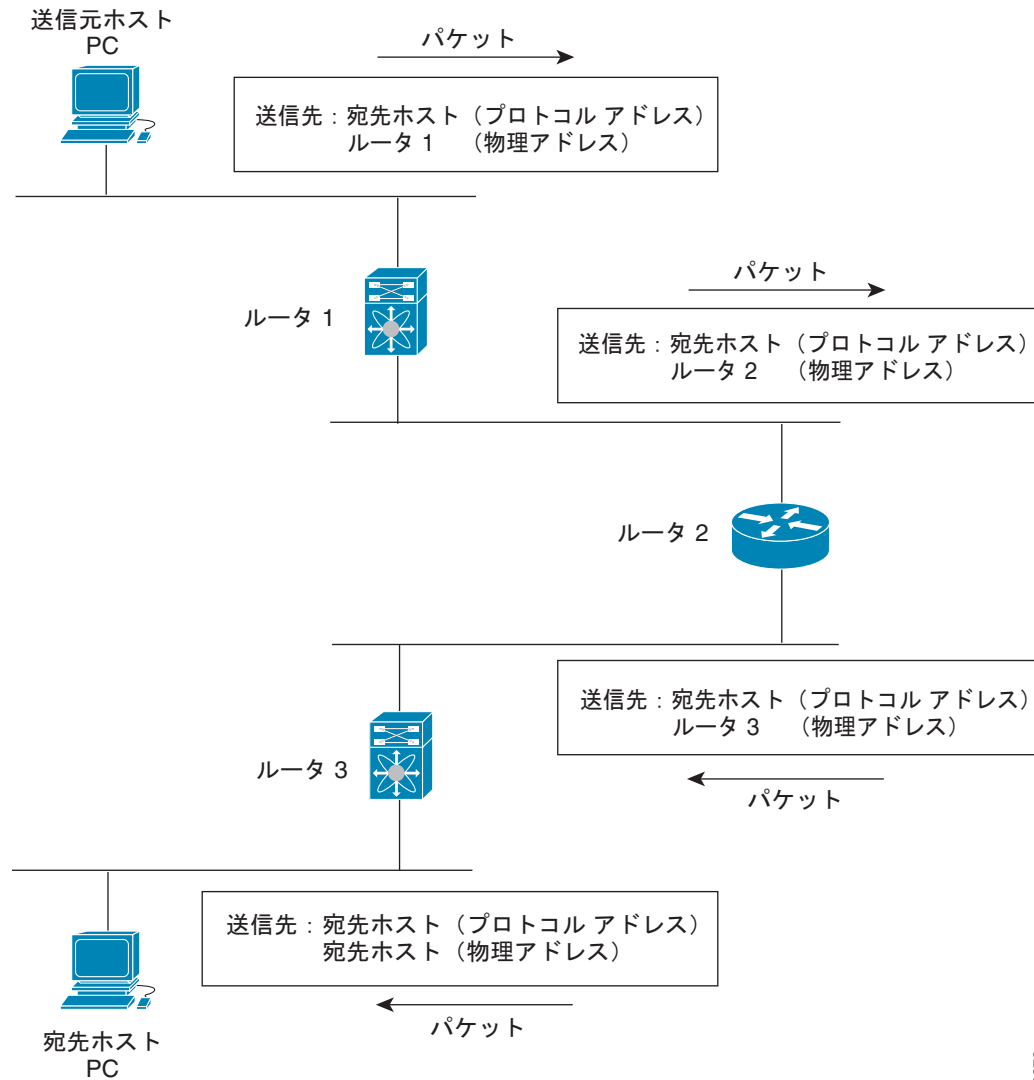
パケット交換

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。なんらかの方法でルータのアドレスを入手したら、送信元ホストはパケットを明確に、宛先ホストの IP (ネットワーク層) アドレスを含むルータの物理 (**Media Access Control (MAC; メディア アクセス コントロール)** レイヤ) アドレス宛に送信します。

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)。

図 1-1 ネットワーク上でのパケット ヘッダーの更新



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めます。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。それ以外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、中間の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容量が大きく、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 使用状況や処理される 1 秒あたりのパケット数など、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。
- 設定済みのループバック インターフェイスがない場合、Cisco NX-OS は、コンフィギュレーション ファイル中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 ではなく、後で 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 バイトの AS 番号をサポートします。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。予約済み番号の割り当てを含む AS 番号の詳細について、または、AS 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、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ッケージ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 6 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、OSPF (Open Shortest Path First) を設定して、ボーダー ゲートウェイ プロトコル (BGP) で検出したルートをアドバタイズできます。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、その再配布ルートに対する固定ルート メトリックを設定します。これにより、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 メトリックの問題が回避されます。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。

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。値が高いほど、信頼性のランクは低くなります。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

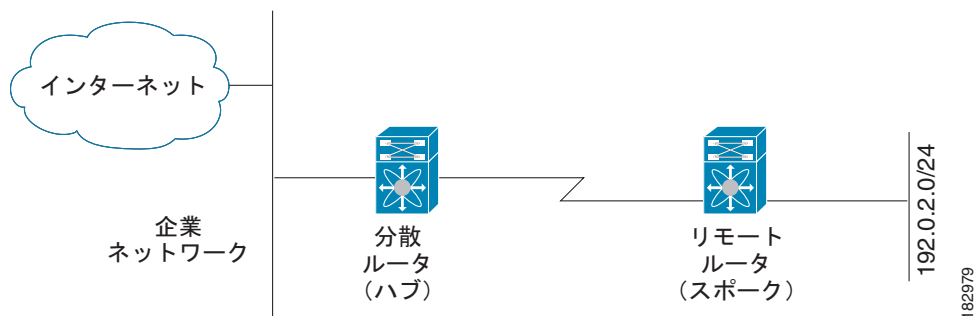
スタブ ルーティングはハブ アンド スポーク型ネットワーク トポロジで使用できます。このトポロジでは、1つ以上の終端（スタブ）ネットワークが、1つ以上の分散ルータ（ハブ）に接続されたリモートルータ（スポーク）に接続されています。リモートルータは、1つ以上の分散ルータにのみ隣接しています。リモートルータへと流れる IP トラフィックのルートは、分散ルータ経由のルートのみです。このタイプの設定は、分散ルータが直接 WAN に接続されている WAN トポロジで使用されるのが一般的です。分散ルータは、さらに多くのリモートルータに接続できます。分散ルータが 100 台以上のリモートルータに接続されていることも、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。

スタブルータの状態を通知するパケットを受信した隣接ルータは、ルートについてはスタブルータに照会しません。また、スタブピアを持つルータは、そのピアについては照会しません。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、企業ネットワークと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 テーブルの機能は無意味です。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 トポロジでリモートルータは、他のネットワークから検出されたルートを受信する必要はあり

ません。これは、宛先がどこであっても、リモート ルータは、すべての非ローカル トラフィックを分散ルータに送信する必要があるためです。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.1-8)
- 「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、今日の大規模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、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。このアルゴリズムでは、着信ルーティング更新メッセージを分析して、ネットワーク状況の変化に合わせて調整します。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。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。この機能により、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、Shortest Path First (SPF; 最短パス優先) と呼ばれ、情報を隣接ルータと共有します。各ルータは Link State Advertisement (LSA; リンクステート アドバタイズメント) を構築し、ここに、各リンクおよび直接接続された隣接ルータに関する情報が含まれます。

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、LSA がすべての隣接ネイバーにフラッドされます。ルータが同じシーケンス番号の 2 つの LSA (同じルータからの) を受信した場合は、LSA 更新ループを防ぐために、ルータは最後に受信した LSA を隣接ルータにフラッドしません。ルータは、受信直後に LSA をフラッドするため、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となります。

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。

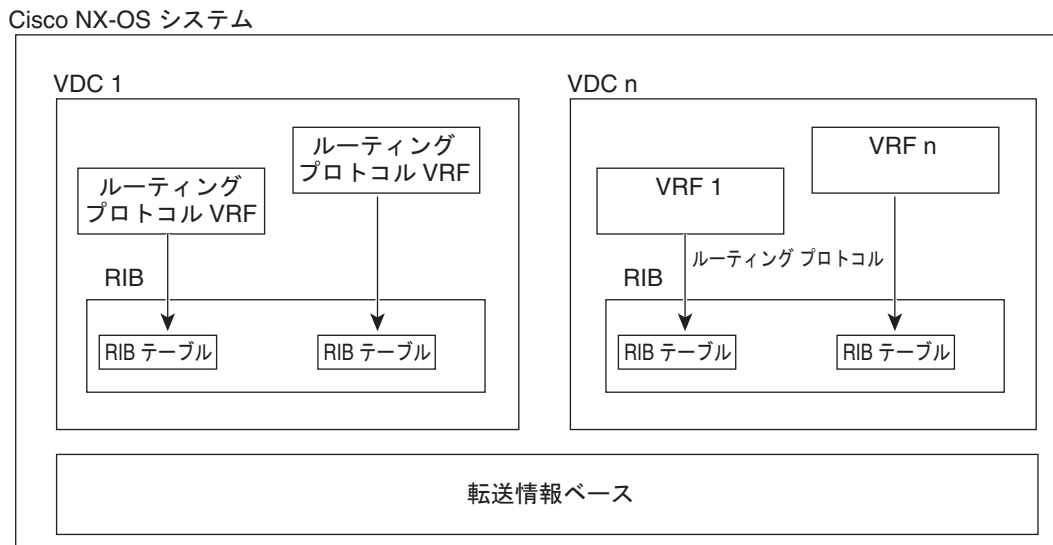
リンクステート アルゴリズムは、より短時間でコンバージェンスするため、ディスタンス ベクトル アルゴリズムより、ルーティング ループがやや発生しにくくなっています。ただし、リンクステート アルゴリズムはディスタンス ベクトル アルゴリズムより、大きな CPU パワーとメモリを必要とします。リンクステート アルゴリズムは、実装とサポートにより多くの費用がかかる場合があります。リンクステート プロトコルは通常、ディスタンス ベクトル プロトコルよりスケーラブルです。

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 には、仮想デバイス コンテキスト (VDC) が採用されており、切り離された VDC およびソフトウェア障害ごとに別の管理ドメインを提供します。各 VDC は、複数の仮想ルーティングおよび転送 (VRF) インスタンスおよび複数のルーティング情報ベース (RIB) をサポートしているため、複数のアドレス ドメインがサポートされます。各 VRF は RIB と関連付けられ、この情報が Forwarding Information Base (FIB; 転送情報ベース) により収集されます。図 1-3 は、VDC、VRF、および Cisco NX-OS システムの間の関係を示します。

図 1-3 レイヤ 3 仮想化の例



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 の詳細については、『Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide, Release 5.x』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 アップデートの処理およびシャーシ内のすべてのモジュールへのスイッチの転送情報の入力が行われます。

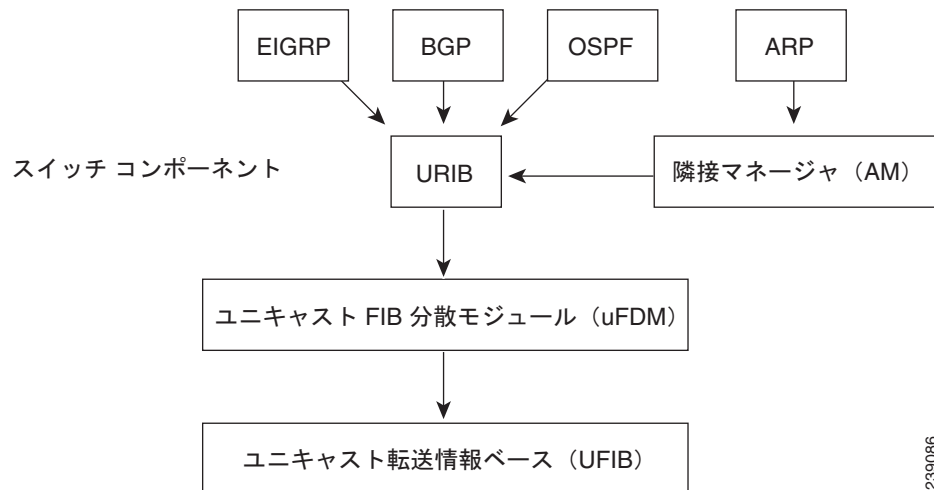
ここでは、次の内容について説明します。

- 「ユニキャスト RIB」 (P.1-11)
- 「隣接マネージャ」 (P.1-11)
- 「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



239086

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。これは、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、ルートに最適なネクストホップを決定し、さらにユニキャスト 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転送情報ベース (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、Open Shortest Path First version 2 (OSPFv2)、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探索します。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は、アクティブなスーパーバイザ上にあり、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 は転送情報を生成し、この情報がユニキャスト FIB により、スタンバイ スーパーバイザとスタンバイ 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ます。また、ユニキャスト転送分散モジュールは、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。

ユニキャスト転送分散モジュールは、隣接情報を収集し、ユニキャスト FIB でのルートの更新時に、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には、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が含まれています。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、ユニキャスト転送分散モジュールからルート更新情報を受信し、ハードウェア転送エンジンにプログラミングされるよう、この情報を送信します。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、およびアドレス ファミリ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 ごとのプレフィッ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ックスで使用できます。

またユニキャスト FIB は、インターフェイスごとのユニキャスト Reverse Path Forwarding (RPF; リバースパス転送) チェックをイネーブルまたはディセーブルにします。Cisco Nexus 3000 シリーズは、各入力側インターフェイスに設定される、次の 2 つの RPF モードをサポートします。

- RPF Strict チェック：ルータ転送テーブルで検証可能な送信元アドレスを持たないパケット、または送信元へのリターンパスに到着しないパケットはドロップされます。
- RPF Loose チェック：パケットはルータ転送テーブルで検証可能な送信元アドレスを持ち、送信元は物理インターフェイスを通じて到達可能です。パケットを受信する入力側インターフェイスは、FIB 内のインターフェイスに一致する必要はありません。

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イッチングエンジンに渡します。ローカルスイッチ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 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ない機能、またはハードウェア処理中に発生したエラーへの対処に使用されます。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。ユニキャスト RIB および隣接マネージャは、ソフトウェアでスイッチされるかまたは終了されるパケットに基づいて転送を決定します。

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 (『Cisco Nexus 7000 Series NX-OS Security Configuration Guide, Release 5.x』を参照してください)。

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.1-13)
- 「OSPF」 (P.1-14)
- 「EIGRP」 (P.1-14)
- 「IS-IS」 (P.1-14)
- 「BGP」 (P.1-14)
- 「RIP」 (P.1-14)
- 「スタティック ルーティング」 (P.1-15)
- 「レイヤ 3 仮想化」 (P.1-15)
- 「Route Policy Manager」 (P.1-15)
- 「ポリシーベース ルーティング」 (P.1-15)
- 「ファーストホップ冗長プロトコル」 (P.1-15)
- 「オブジェクト トラッキング」 (P.1-16)

IPv4 および IPv6

レイヤ 3 は IPv4 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。

OSPF

OSPF プロトコルは、AS 内のネットワーク到達可能性情報の交換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。リンク情報には、リンクタイプ、リンクメトリック、およびリンクに接続された隣接ルータが含まれます。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 4 章「OSPFv2 の設定」](#)を参照してください。

EIGRP

EIGRP プロトコルは、ディスタンスベクトルとリンクステートの両ルーティングプロトコルの特徴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、典型的なディスタンスベクトルルーティングプロトコルのように、ルートを提供するために隣接ルータを必要とします。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 6 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、International Organization for Standardization (ISO; 国際標準化機構) 10589 で指定されたドメイン内 OSI (Open System Interconnection; 開放型システム間相互接続) 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 には、次の機能があります。

- 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
- 高いスケーラビリティ

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ックス、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 7 章「ベーシック BGP の設定」](#) および [第 8 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 VDC をサポートしているため、Cisco NX-OS システム内に複数の仮想システムを作成できます。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、相互にセキュリティが確保されています。各 VDC にそれぞれ別のネットワーク 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。

Cisco NX-OS は、また VPN ルーティングおよび転送 (VRF) を含むレイヤ 3 仮想化を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 ヘッダーの他の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。ポリシー ルートは、拡張 IP アクセス リストにリンクできるため、プロトコル タイプやポート番号などに基づいてルーティングが行われる場合もあります。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

ファーストホップ冗長プロトコル

First-hop Redundancy Protocol (FHRP; ファーストホップ冗長プロトコル) は、ホストへの冗長接続を可能にします。アクティブなファーストホップ ルータがダウンした場合は、その機能を引き継ぐスタンバイ ルータが、FHRP により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。ゲートウェイ ロード バランシング プロトコル (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。Hot Standby Router Protocol (HSRP; ホットスタンバイ ルータ プロトコル)

の詳細については、第 17 章「HSRP の設定」を参照してください。Virtual Router Redundancy Protocol (VRRP; 仮想ルータ冗長プロトコル) の詳細については、第 18 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

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。この機能により、ネットワークの可用性が向上し、オブジェクトがダウン状態となった場合の回復時間が短縮されます。詳細については、第 19 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, Release 5.x』
- 『Cisco Nexus 7000 Series NX-OS High Availability and Redundancy Guide, Release 5.x』
- 『Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide, Release 5.x』
- 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