



## 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。

この章は、次の項で構成されています。

- 「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.1-12)
- 「関連項目」 (P.1-14)

## 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なパス（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。

この項では、次のトピックについて取り上げます。

- 「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な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、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、最適なパスで 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.1-10) 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なルートを決めます。これらのメトリックは、使用しているルーティング アルゴリズムの設計によって異なります。「[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うち 1 つの例であるリンクステート アドバタイズメントは、送信ルータのリンク状態を他のルータに通知します。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なルートを決め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。

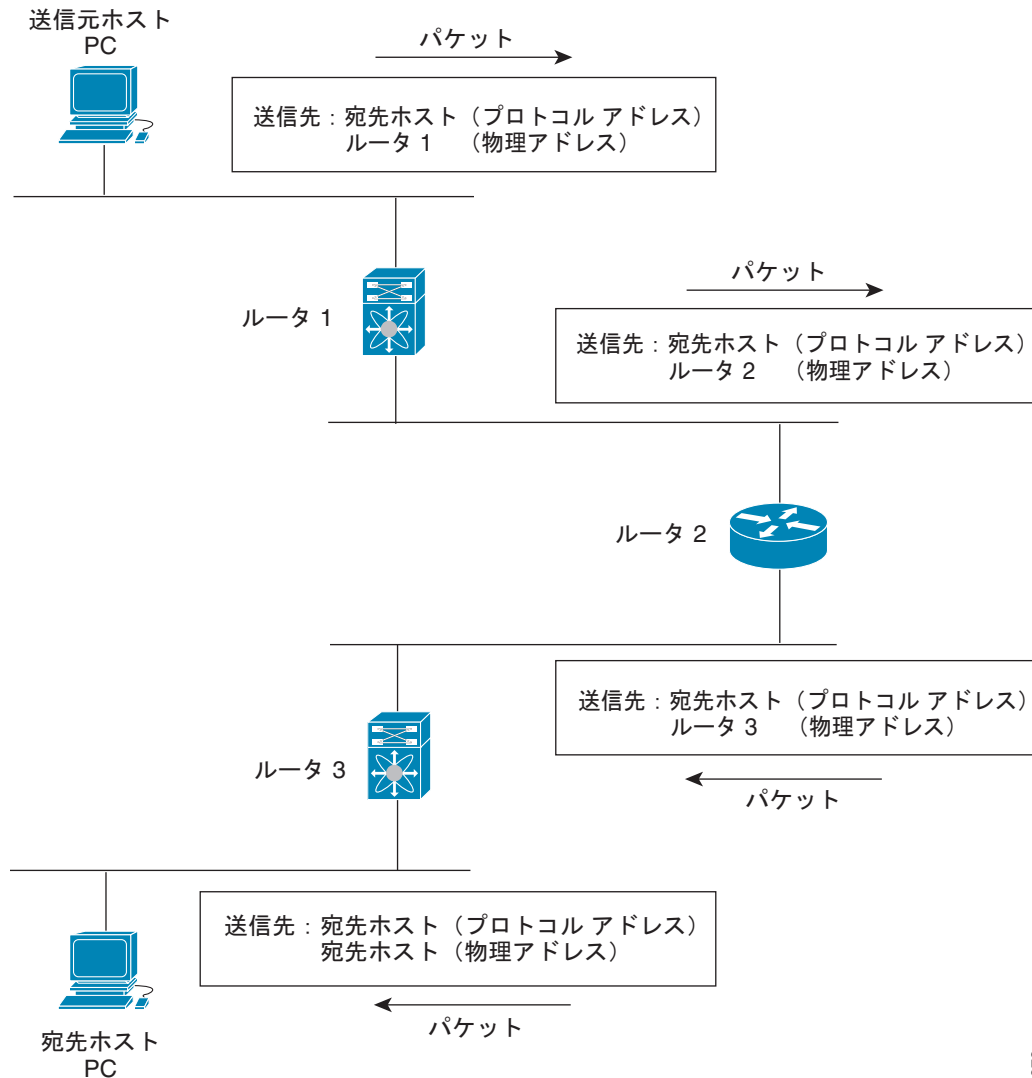
## パケット交換

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。ルータがパケットの転送方法を認識していない場合は、通常はパケットをドロップします。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なルートを決めます。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

パスの長さは、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後、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

ルーティング遅延は、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 プロセスに関連付けられているルータ ID があります。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なかった場合、Cisco NX-OS はルータ ID としてコンフィギュレーション ファイルの最初のインターフェイスを使用します。Cisco NX-OS がルータ ID を選択した後、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、`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せん。
- 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。自律システムにより、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。

各自律システムは、ルートの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 (RIR) 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。この自律システム番号で、ルーティング処理と自律システムの両方が識別されます。

ボーダー ゲートウェイ プロトコル (BGP) は、`asplain` と `asdot` 表記で表示できる 4 バイトの AS 番号をサポートします。

- `asplain` : 10 進表記方式。2 バイトおよび 4 バイト AS 番号をその 10 進数値で表します。たとえば、65526 は 2 バイト AS 番号、234567 は 4 バイト AS 番号になります。
- `asdot` : AS ドット付き表記方式。2 バイト AS 番号をその 10 進数値で表し、4 バイトの AS 番号をドット付き表記で表します。たとえば、2 バイト AS 番号 65526 は 65526 として表され、4 バイトの AS 番号 65546 は 1.10 として表されます。

BGP の 4 バイト AS 番号機能は、4 バイト AS 番号をサポートしていない BGP スピーカーをまたがって、4 バイトをベースとする AS パス情報を伝播するために使用されます。



(注)

RFC 5396 は部分的にサポートされます。`asplain` と `asdot` 表記はサポートされますが、`asdot+` 表記はサポートされません。

専用自律システム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、専用自律システム番号が外部ネットワークにアドバ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 は専用自律システム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自律システム番号は、インターネット割り当て番号局 (IANA) 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自律システム番号の詳細について、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、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 7 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) プロトコルを設定して、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。他のプロトコルからのルートを再配布するルータは、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 メトリックを防ぐ再配布されたルータの固定ルートを設定します。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にルート 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、ルーティング情報源の信頼性を示す評価基準です。値が高いほど信頼性の評価は低くなります。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、複数のプロトコルから学習し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 テーブルに組み込まれます。

## スタブ ルーティング

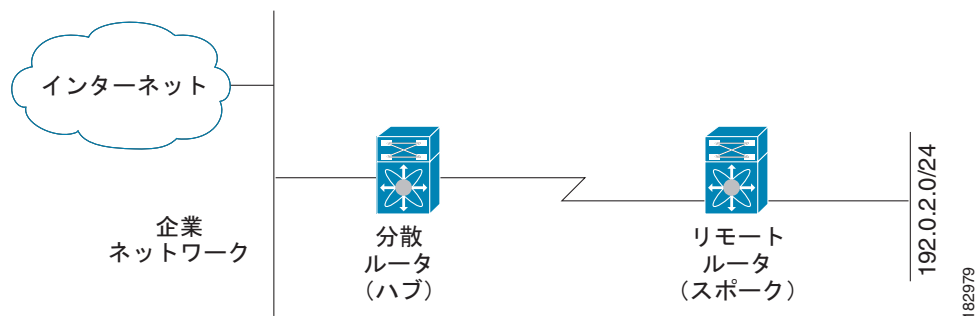
スタブ ルーティングはハブ アンド スポーク型ネットワーク トポロジで使用できます。このトポロジでは、1 つ以上の終端 (スタブ) ネットワークが、1 つ以上の分散ルータ (ハブ) に接続されたリモートルータ (スポーク) に接続されています。リモートルータは、1 つ以上のディストリビューションルータにのみ隣接しています。リモートルータへ流れる IP トラフィックのルートは、ディストリビューションルータ経由のルートのみです。このタイプの設定は、ディストリビューションルータが直接 WAN に接続されている WAN トポロジで使用されるのが一般的です。ディストリビューションルータは、さらに多くのリモートルータに接続できます。ディストリビューションルータが 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ディストリビューション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、ディストリビューションルータを使用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、Enhanced Interior Gateway Routing Protocol (EIGRP) はスタブ ルータをサポートします。

## 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な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、最適な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。

この項では、次のトピックについて取り上げます。

- 「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が変化したことを示している場合は、ルーティング ソフトウェアはルートを再計算し、新しいルーティング アップデート 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 ドメインまたは自律システムに分割できます。自律システムは、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。自律システム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



イ プロトコルの例です。1 つの自律システム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、直接接続されたネイバー 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) と呼ばれ、情報を隣接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ドされます。ルータが (同じルータから) 同じシーケンス番号の 2 つの LSA を受信した場合、ルータは LSA アップデートのループを回避するため、ネイバーによって受信された最後の LSA をフラッドしません。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。

ネイバー 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。ネイバー 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、各ネイバー ルータのキーアライブ通知としても機能します。隣接関係は、ネイバー 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、より大きな更新をネイバー 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、より多くの CPU パワーとメモリを必要とし、実行とサポートをするにはよりコストが高くなります。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、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 に関連付けられており、この情報が転送情報ベース (FIB) によって収集されます。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。詳細については、第 13 章「レイヤ 3 仮想化の設定」を参照してください。

Cisco NX-OS では、仮想デバイスをエミュレートする Virtual Device Context (VDC) に、OS およびハードウェア リソースを分割できます。Cisco Nexus 9000 シリーズ スイッチは、現在複数の VDC をサポートしていません。すべてのスイッチ リソースはデフォルト VDC で管理されます。

## 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。

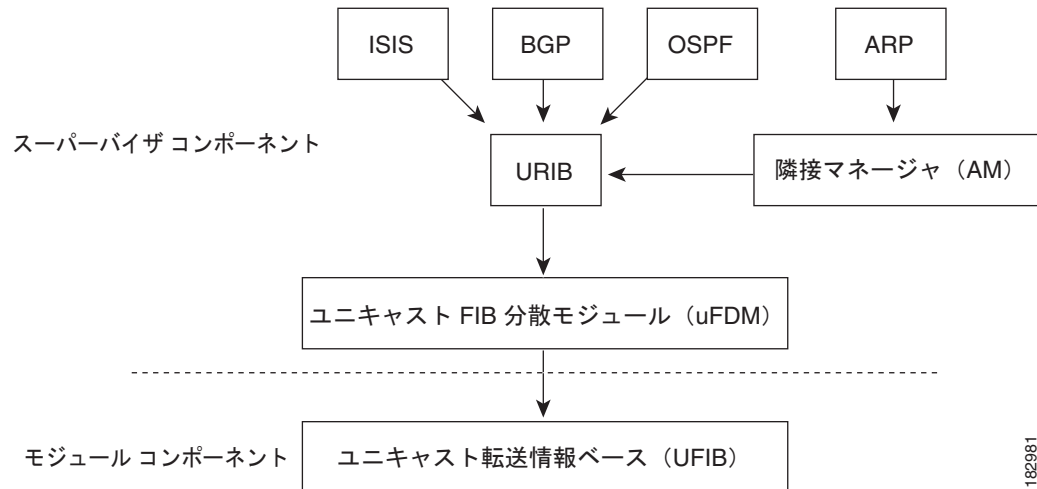
この項では、次のトピックについて取り上げます。

- 「ユニキャスト RIB」 (P.1-10)
- 「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.1-11)
- 「FIB」 (P.1-12)
- 「ハードウェア転送」 (P.1-12)
- 「ソフトウェア転送」 (P.1-12)

## ユニキャスト RIB

Cisco NX-OS 転送アーキテクチャは、図 1-3 に示すように、複数のコンポーネントで構成されます。

図 1-3 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、アドレス解決プロトコル (ARP) 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、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。

## 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 によってスタンバイ 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。また、ユニキャスト FDM 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 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。隣接およびリライト情報は、インターフェイスとネクスト ホップ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。

## 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 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 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 IP アドレスが含まれます。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。

## 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。

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。

ソフトウェア転送はコントロール プレーン ポリシーによって制御されます。詳細については、『Cisco Nexus 9000 Series NX-OS Security Configuration Guide』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。

この項では、次のトピックについて取り上げます。

- 「IPv4 および IPv6」(P.1-13)
- 「IP サービス」(P.1-13)
- 「OSPF」(P.1-13)
- 「EIGRP」(P.1-13)
- 「IS-IS」(P.1-13)

- 「BGP」 (P.1-14)
- 「RIP」 (P.1-14)
- 「スタティック ルーティング」 (P.1-14)
- 「レイヤ 3 仮想化」 (P.1-14)
- 「Route Policy Manager」 (P.1-14)

## 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 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ドメイン ネーム システム (DNS) 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、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ネイバー 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 5 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 はネイバーに依存し、ルートを提供します。また、リンクステート プロトコルのように、ネイバー 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 7 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開放型システム間相互接続 (Open System Interconnection) 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
- 高いスケラビリティ

詳細については、[第 8 章「IS-IS の設定」](#)を参照してください。

## BGP

BGP は自律システム間ルーティング 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伝送制御プロトコル (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がある自律システムのリスト、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せん。詳細については、[第 9 章「ベーシック BGP の設定」](#)および[第 10 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、単一の自律システム内でルーティングを行います。詳細については、[第 11 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 12 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、仮想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 13 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 15 章「Route Policy Manager の設定」](#)を参照してください。

## 関連項目

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 9000 Series NX-OS Multicast Routing Configuration Guide](#)』
- 『[Cisco Nexus 9000 Series NX-OS High Availability and Redundancy Guide](#)』

- 自律システム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](http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_9-1/autonomous_system_numbers.html)

