



概要

この章では、Cisco NX-OS のマルチキャスト機能について説明します。

- [ライセンス要件](#) (1 ページ)
- [マルチキャストについて](#) (1 ページ)
- [マルチキャストに関する注意事項と制限事項](#) (12 ページ)
- [マルチキャストのハイ アベイラビリティ要件](#) (12 ページ)
- [仮想デバイス コンテキスト](#) (12 ページ)
- [SW と HW マルチキャスト ルート間の不一致のトラブルシューティング](#) (13 ページ)

ライセンス要件

Cisco NX-OS ライセンス方式の推奨の詳細と、ライセンスの取得および適用の方法については、『[Cisco NX-OS Licensing Guide](#)』を参照してください。

マルチキャストについて

IP マルチキャストは、同一セットの IP パケットをネットワーク上の複数のホストに転送する手法です。IPv4 ネットワークで、マルチキャストを使用して、複数の受信者に効率的にデータを送信できます。

マルチキャストには、グループと呼ばれる IP マルチキャストアドレスに送信されたマルチキャストデータの送信側と受信側の配信と検出の両方の手法が含まれます。グループと送信元 IP アドレスが入ったマルチキャストアドレスは、しばしばチャンネルと呼ばれます。Internet Assigned Number Authority (IANA) では、IPv4 マルチキャストアドレスとして、224.0.0.0 ~ 239.255.255.255 を割り当てています。詳細については、次の URL を参照してください。
<http://www.iana.org/assignments/multicast-addresses>

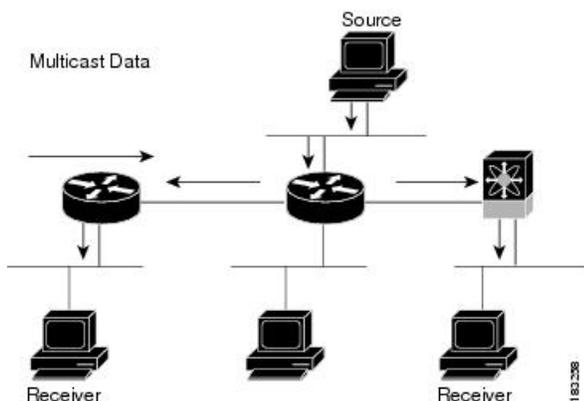


(注) マルチキャストに関連する RFC の完全なリストについては、「[IP マルチキャストに関する IETF RFC](#)」の章を参照してください。

ネットワーク上のルータは、受信者からのアドバタイズメントを検出して、マルチキャストデータの要求対象となるグループを特定します。その後、ルータは送信元からのデータを複製して、対象の受信者へと転送します。グループ宛のマルチキャストデータが送信されるのは、そのデータを要求する受信者を含んだ LAN セグメントだけです。

次の図に、1つの送信元から2つの受信者へと、マルチキャストデータを送信する場合の例を示します。この図で、中央のホストが属する LAN セグメントにはマルチキャストデータを要求する受信者が存在しないため、このホストは受信者にデータを転送しません。

図 1: 1つの送信元から2つの受信者へのマルチキャストトラフィック



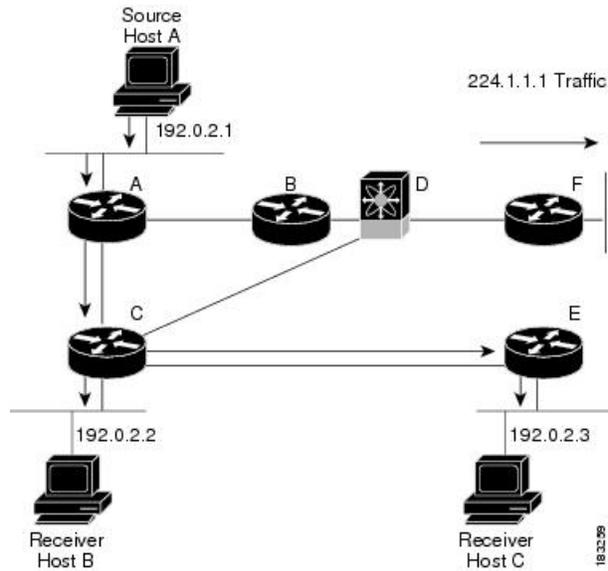
マルチキャスト配信ツリー

マルチキャスト配信ツリーとは、送信元と受信者の中継するルータ間の、マルチキャストデータの伝送パスを表します。マルチキャストソフトウェアはサポートするマルチキャスト方式に応じて、タイプの異なるツリーを構築します。

送信元ツリー

送信元ツリーは、送信元からネットワーク経由でマルチキャストトラフィックを伝送する場合の最短パスです。特定のマルチキャストグループへと送信されたマルチキャストトラフィックが、同じグループのトラフィックを要求する受信者へと転送されます。送信元ツリーは、最短パスとしての特性から、最短パスツリー (SPT) と呼ばれることがあります。この図は、ホスト A を起点とし、ホスト B および C に接続されているグループ 224.1.1.1 の送信元ツリーを示しています。

図 2: 送信元ツリー

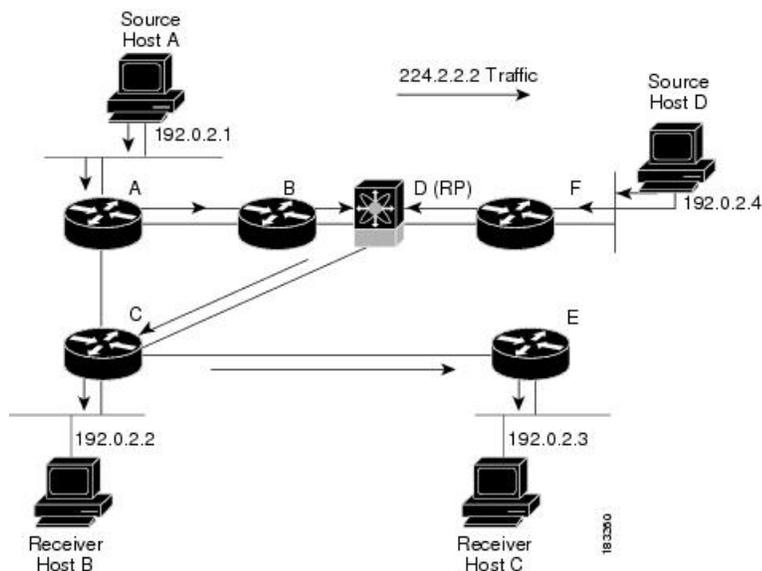


表記 (S,G) は、グループ G の任意の送信元からのマルチキャストトラフィックを表します。この図の SPT は、(192.0.2.1, 224.1.1.1) と記述されます。同じグループの複数の送信元からトラフィックを送信できます。

共有ツリー

共有ツリーとは、共有ルート、つまりランデブーポイント (RP) から各受信者に、ネットワーク経由でマルチキャストトラフィックを伝送する共有配信パスを表します (RP は各ソースへの SPT を作成します。) 共有ツリーは、RP ツリー (RPT) とも呼ばれます。この図は、ルータ D に RP を持つ、グループ 224.2.2.2 の共有ツリーを示しています。データは送信元ホスト A およびホスト D からルータ D (RP) に送信され、そこから受信者ホスト B およびホスト C にトラフィックが転送されます。

図 3: 共有ツリー

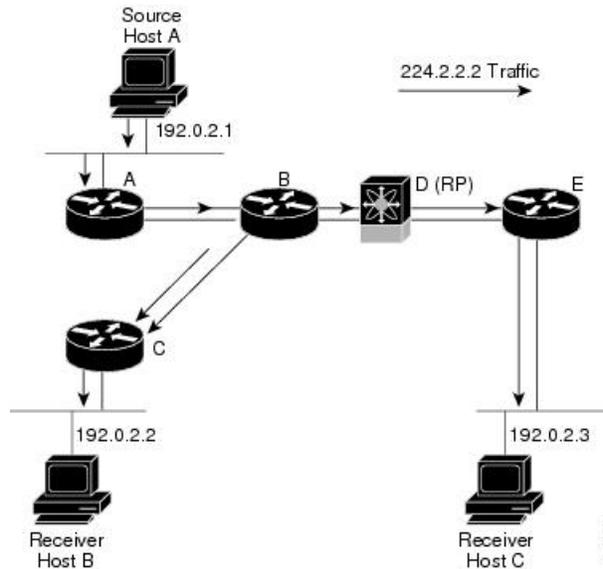


表記 (*,G) は、グループ G の任意の送信元からのマルチキャストトラフィックを表します。図の共有ツリーは、(*, 224.2.2.2) と記述されます。

双方向共有ツリー

双方向共有ツリーとは、共有ルート、つまりランデブーポイント (RP) から各受信者に、ネットワーク経由でマルチキャストトラフィックを伝送する共有配信パスを表します。マルチキャストデータは、RP への経路上にある受信者に転送されます。次の表に、双方向共有ツリーの利点を示します。マルチキャストトラフィックは、ルータ B および C を通して、ホスト A からホスト B に直接送られます。共有ツリーの場合、送信元ホスト A から送信されたデータは、まず RP (ルータ D) に送信され、ルータ B に転送されてからホスト B に伝送されます。

図 4: 双方向共有ツリー



表記 (*,G) は、グループ G の任意のソースからのマルチキャストトラフィックを表します。図の双方向ツリーは、(*, 224.2.2.2) と記述されます。

マルチキャスト転送

マルチキャストトラフィックは任意のホストを含むグループ宛に送信されるため、ルータはリバースパスフォワーディング (RPF) を使用して、グループのアクティブな受信者にデータをルーティングします。受信者がグループに加入すると、送信元方向へ向かうパス (SSM モード)、または RP 方向へ向かうパス (ASM または Bidir モード) が形成されます。送信元から受信者へのパスは、受信者がグループに加入したときに作成されたパスと逆方向になります。

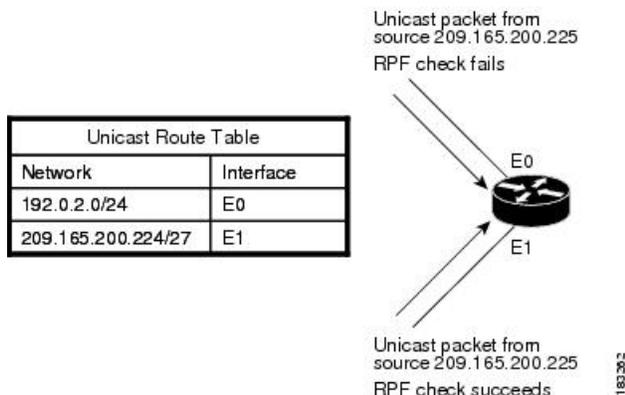
マルチキャストパケットが着信するたびに、ルータは RPF チェックを実行します。送信元に接続されたインターフェイスにパケットが着信した場合は、グループの発信インターフェイス (OIF) リスト内の各インターフェイスにパケットが転送されます。それ以外の場合、パケットはドロップされます。



- (注) Bidir モードでは、パケットが非 RPF インターフェイスに着信した際に、インターフェイスが指定フォワーダ (DF) として選択されていれば、パケットは RP に向かうアップストリーム方向にも転送されます。

次の図に、異なるインターフェイスから着信したパケットについて、RPF チェックを行う場合の例を示します。E0 に着信したパケットは、RPF チェックに失敗します。これは、ユニキャストテーブルで、対象の送信元ネットワークがインターフェイス E1 に関連付けられているためです。E1 に着信したパケットは、RPF チェックに合格します。これは、ユニキャストルートテーブルで、対象の送信元ネットワークがインターフェイス E1 に関連付けられているためです。

図 5: RPF チェックの例



Cisco NX-OS の PIM

Cisco NX-OS は、Protocol Independent Multicast (PIM) スパース モードを使用したマルチキャストをサポートします。PIM は IP ルーティング プロトコルに依存せず、使用されているすべてのユニキャストルーティングプロトコルが提供するユニキャストルーティングテーブルを利用できます。PIM スパース モードでは、ネットワーク上の要求元だけにマルチキャストトラフィックが伝送されます。Cisco NX-OS では、PIM デンス モードはサポートされません。



(注) このマニュアルで、「PIM」という用語は PIM スパース モードバージョン 2 を表します。

マルチキャストコマンドにアクセスするには、PIM 機能をイネーブルにする必要があります。ドメイン内の各ルータのインターフェイス上で、PIM をイネーブルにしないかぎり、マルチキャスト機能はイネーブルになりません。PIM は IPv4 ネットワーク用に設定できます。デフォルトでは、IGMP がシステムで稼働しています。

マルチキャスト対応ルータ間で使用される PIM は、マルチキャスト配信ツリーを構築して、ルーティング ドメイン内にグループ メンバーシップをアドバタイズします。PIM は、複数の送信元からのパケットが転送される共有配信ツリーと、単一の送信元からのパケットが転送される送信元配信ツリーを構築します。

配信ツリーは、リンク障害またはルータ障害のためにトポロジが変更されると、トポロジを反映して自動的に変更されます。PIM はマルチキャスト対応の送信元および受信者を動的に追跡します。ただし、Bidir モードの場合、送信元ステートは生成されません。

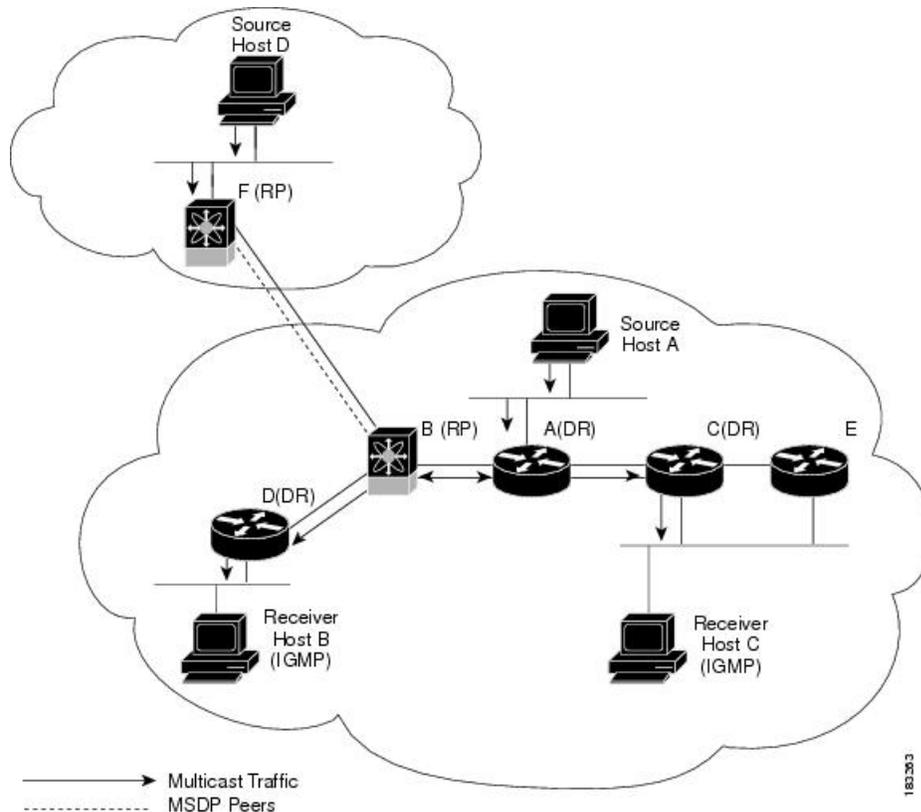
ルータはユニキャストルーティングテーブルおよび RPF ルートを使用して、マルチキャストルーティング情報を生成します。Bidir モードの場合は、追加のルーティング情報が生成されます。



- (注) このマニュアルでは、「IPv4用のPIM」という表現は、Cisco NX-OSにおけるPIM スペースモードの実装を表します。

次の図に、IPv4 ネットワーク内の2つのPIMドメインを示します。

図6: IPv4 ネットワーク内の PIM ドメイン



- 矢印の付いた直線は、ネットワークで伝送されるマルチキャストデータのパスを表します。マルチキャストデータは送信元ホストのAおよびDから発信されます。
- 点線でつながれているルータBおよびFは、Multicast Source Discovery Protocol (MSDP) ピアです。MSDPを使用すると、他のPIMドメイン内にあるマルチキャスト送信元を検出できます。
- ホストBおよびホストCではマルチキャストデータを受信するため、インターネットグループ管理プロトコル (IGMP) プロトコルを使用して、マルチキャストグループへの加入要求をアドバタイズします。
- ルータA、C、およびDは指定ルータ (DR) です。LANセグメントに複数のルータが接続されている場合は (CやEなど)、PIMソフトウェアによってDRとなるルータが1つ選択されます。これにより、マルチキャストデータの窓口として、1つのルータだけが使用されます。

ルータ B とルータ F は、それぞれ異なる PIM ドメインのランデブーポイント (RP) です。RP は、複数の送信元と受信者を接続するため、PIM ドメイン内の共通ポイントとして機能します。

PIM は送信元と受信者間の接続に関して、これらのマルチキャストモードをサポートしています。

- Any Source Multicast (ASM)
- Source Specific Multicast (SSM)
- 双方向共有ツリー (Bidir)

Cisco NX-OS では上記モードを組み合わせて、さまざまな範囲のマルチキャストグループに対応することができます。マルチキャスト用の RPF ルートを定義することもできます。

アーキテクチャ セールス マネージャ (ASM)

Any Source Multicast (ASM) は PIM ツリー構築モードの 1 つです。新しい送信元および受信者を検出する場合には共有ツリーを、受信者から送信元への最短パスを形成する場合は送信元ツリーを使用します。共有ツリーでは、ランデブーポイント (RP) と呼ばれるネットワークノードをルートとして使用します。送信元ツリーは第 1 ホップルータをルートとし、アクティブな発信元である各送信元に直接接続されています。ASM モードでは、グループ範囲に対応する RP が必要です。RP は静的に設定することもできれば、Auto-RP プロトコルまたはブートストラップルータ (BSR) プロトコルを使用して、グループと RP 間の関連付けを動的に検出することもできます。RP が学習されている場合、かつ Bidir-RP であるかどうか不明な場合、グループは ASM モードで動作します。

RP を設定する場合、デフォルトモードは ASM モードです。

Bidir

双方向共有ツリー (Bidir) は ASM モードと同様、受信者と RP の間の共有ツリーを構築する PIM モードです。ただし、グループに新しい受信者が追加された場合、送信元ツリーに切り替えることはできません。Bidir モードの場合、受信者に接続されたルータは代表フォワード (DF) と呼ばれます。これは、RP を経由することなく、代表ルータ (DR) から受信者に直接マルチキャストデータを転送できるためです。Bidir モードを利用するには、RP を設定する必要があります。

Bidir モードを使用すると、マルチキャスト送信元が多数存在する場合に、ルータに必要なリソース量を削減するとともに、RP の動作ステータスや接続ステータスに関係なく、運用を継続できます。

SSM

送信元固有マルチキャスト (SSM) は、マルチキャスト送信元への加入要求を受信する LAN セグメント上の代表ルータを起点として、送信元ツリーを構築する PIM モードです。送信元ツリーは、PIM 加入メッセージを送信元方向に送信することで構築されます。SSM モードでは、RP を設定する必要がありません。

SSM モードの場合、PIM ドメインの外部にある送信元と受信者を接続できます。

マルチキャスト用 RPF ルート

静的マルチキャスト RPF ルートを設定すると、ユニキャストルーティングテーブルの定義内容を無効にすることができます。この機能は、マルチキャストトポロジとユニキャストトポロジが異なる場合に使用されます。

IGMP

デフォルトでは、PIM のインターネット グループ管理プロトコル (IGMP) が、システムで実行されています。

IGMP は、マルチキャストグループのメンバーシップを要求するため、マルチキャストデータを受信する必要があるホストで使用されます。グループメンバーシップが確立されると、対象のグループのマルチキャストデータが要求元ホストの LAN セグメントに転送されます。

インターフェイスには IGMPv2 または IGMPv3 を設定できます。SSM モードをサポートするには、(S, G) を使用して IGMPv3 を設定する必要があります。デフォルトでは IGMPv2 がイネーブルになっています。

IGMP スヌーピング

IGMP スヌーピングは、VLAN で既知の受信者に接続された一部のポートだけにマルチキャストトラフィックを転送する機能です。対象ホストからの IGMP メンバーシップ レポートメッセージを調べる (スヌーピングする) ことにより、マルチキャストトラフィックは対象ホストが接続された VLAN ポートだけに送信されます。システムでは、IGMP スヌーピングがデフォルトで稼働しています。

ドメイン内マルチキャスト

Cisco NX-OS では、PIM ドメイン間でマルチキャストトラフィック送信を実行するための方法が提供されます。

SSM

PIM ソフトウェアは SSM を使用して、受信者の指定ルータから既知の送信元 IP アドレスへの最短パス ツリーを構築します。この場合、送信元は別の PIM ドメイン内にあってもかまいません。ASM および Bidir モードの場合、別の PIM ドメインから送信元にアクセスするには、別のプロトコルを使用する必要があります。

ネットワークで PIM をイネーブルにすると、SSM を使用し、受信者の指定ルータが IP アドレスを把握している任意のマルチキャスト送信元への接続パスを確立できます。

MSDP

Multicast Source Discovery Protocol (MSDP) は、PIM と組み合わせて使用することで、異なる PIM ドメイン内にあるマルチキャスト送信元を検出できるようにするマルチキャストルーティングプロトコルです。



(注) Cisco NX-OS では、MSDP 設定が不要な PIM Anycast-RP をサポートしています。

MBGP

Multiprotocol BGP (MBGP) は BGP4 の拡張機能であり、ルータによるマルチキャストルーティング情報の伝送を可能にします。このマルチキャスト情報を使用すると、PIM を介して、外部の BGP 自律システム (AS) 内の送信元と通信できます。

MRIB

Cisco NX-OS IPv4 マルチキャストルーティング情報ベース (MRIB) は、PIM や IGMP などのマルチキャストプロトコルで生成されるルート情報を格納するためのリポジトリです。MRIB はルート情報自体には影響を及ぼしません。MRIB はの仮想ルーティングおよびフォワーディング (VRF) インスタンスごとに、独立したルート情報を保持します。

Cisco NX-OS リリース 10.2(1)以降、グローバル境界マルチキャスト設定がサポートされています。

グローバル マルチキャスト境界で許可または拒否される IP マルチキャストグループおよびチャンネルのグローバル範囲を定義するには、VRF コンフィギュレーションモードで **{ip | ipv6} multicast group-range prefix-list <prefix-list-name>** コマンドを設定する必要があります。このコマンドは、ルータのすべてのインターフェイスで、未認可グループまたはチャンネルのマルチキャストプロトコルアクションおよびトラフィック転送をディセーブルにするために使用されます。prefix-list は境界を構成します。次に設定例を示します。

```
vrf context enterprise
ip multicast group-range prefix-list test
```

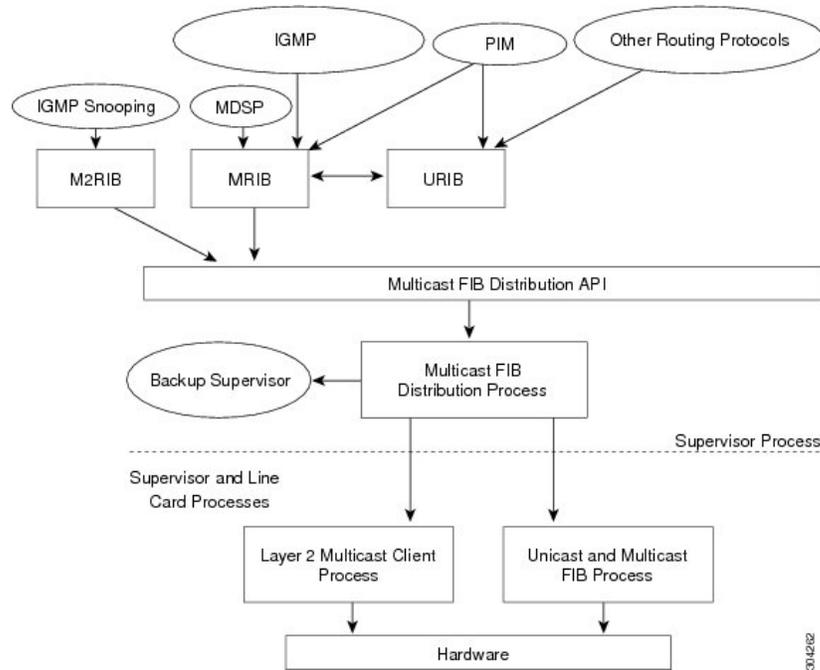
Cisco NX-OS マルチキャスト ソフトウェア アーキテクチャの主要コンポーネントは次のとおりです。

- マルチキャスト FIB (MFIB) 分散 (MFDM) API は、MRIB を含むマルチキャストレイヤ 2 およびレイヤ 3 コントロールプレーンモジュールと、プラットフォーム転送プレーン間のインターフェイスを定義します。コントロールプレーンモジュールは、MFDM API を使用してレイヤ 3 ルートアップデートを送信します。
- マルチキャスト FIB 配信プロセス：すべての関連モジュールおよびスタンバイ スーパーバイザに、マルチキャストアップデートメッセージを配布します。このプロセスはスーパーバイザだけで実行されます。

- レイヤ2 マルチキャストクライアントプロセス：レイヤ2 マルチキャストハードウェア転送パスを構築します。このプロセスは、スーパーバイザとモジュールの両方で実行されます。
- ユニキャストおよびマルチキャストFIBプロセス：レイヤ3ハードウェア転送パスを管理します。このプロセスは、スーパーバイザとモジュールの両方で実行されます。

次の図に、Cisco NX-OS マルチキャストソフトウェアのアーキテクチャを示します。

図 7: Cisco NX-OS マルチキャストソフトウェアのアーキテクチャ



仮想ポートチャネルおよびマルチキャスト

仮想ポートチャネル (vPC)：1 台のデバイスで 2 台のアップストリームスイッチのポートチャネルを使用できるようにします。vPC を設定すると、次のマルチキャスト機能に影響が及ぶ可能性があります。

- PIM：Cisco Nexus 9000 シリーズスイッチ対応の Cisco NX-OS ソフトウェアは、vPC での PIM Bidir をサポートしません。
- IGMP スヌーピング：vPC ピアの設定を同一にする必要があります。

より低い IP アドレスを持つ L2 デバイスでスヌーピングクエリアを設定して、L2 デバイスをクエリアとして強制することをお勧めします。これは、マルチシャーシ EtherChannel トランク (MCT) がダウンしているシナリオの処理に役立ちます。

マルチキャストに関する注意事項と制限事項

- Cisco NX-OS リリース 10.2(1q)F 以降、Cisco Nexus N9K-C9332D-GX2B プラットフォームスイッチではレイヤ 2 およびレイヤ 3 マルチキャストがサポートされます。
- Cisco NX-OS リリース 10.1(2) 以降、N9K-X9624D-R2 ラインカードではレイヤ 3 マルチキャストがサポートされます。
- レイヤ 3 イーサネットポートチャネルサブインターフェイスは、マルチキャストルーティングではサポートされていません。
- レイヤ 2 IPv6 マルチキャストパケットは、着信 VLAN でフラッディングされます。
- 不明なマルチキャストトラフィックによるトラフィックストーム制御はサポートされていません。
- FEX ポートでのレイヤ 3 マルチキャストルーティングおよび FEX ポートチャネルでのレイヤ 3 マルチキャストルーティングは、Cisco Nexus 9300-FX および -EX プラットフォームスイッチでサポートされています。
- 双方向モードは、-R ラインカードを備えた Cisco Nexus 9500 プラットフォームスイッチではサポートされていません。
- IPv6 マルチキャストは、Cisco Nexus 9500 R シリーズラインカードではサポートされていません。

マルチキャストのハイアベイラビリティ要件

マルチキャストルーティングプロトコルを再起動すると、MRIB プロセスによってステータスが回復されます。スーパーバイザのスイッチオーバーが発生した場合、MRIB はハードウェアからステータスを回復し、マルチキャストプロトコルは定期的なメッセージアクティビティからステータスを回復します。ハイアベイラビリティの詳細については、『Cisco Nexus 9000 シリーズ NX-OS ハイアベイラビリティおよび冗長性ガイド』を参照してください。

仮想デバイスコンテキスト

Cisco NX-OS では、仮想デバイスをエミュレートする Virtual Device Context (VDCs) 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、現在のところ、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。

SW と HW マルチキャスト ルート間の不一致のトラブルシューティング

症状

このセクションでは、アクティブなフローで MRIB に表示されるが、MFIB でプログラムされていない *、G、または S,G エントリに関連した症状、考えられる原因、および推奨されるアクションについて説明します。

考えられる原因

この問題は、ハードウェアの容量を超えて多数のアクティブフローを受信した場合に発生します。これにより、空きハードウェアインデックスがなくなって、一部のエントリがハードウェアでプログラムされなくなります。

ハードウェアリソースを解放するためにアクティブなフローの数が大幅に削減された場合、ハードウェアテーブルがいっぱいであったときに以前影響されていたフローについては、エントリ、タイムアウト、再入力が生じ、プログラミングがトリガーされるまで、MRIB と MFIB の間で不整合が見られることがあります。

現在、ハードウェアリソースが解放された後に、MRIB テーブルを調べて、ハードウェアの欠落しているエントリを再プログラムするメカニズムはありません。

改善処置

エントリを確実に再プログラミングするには、**clear ip mroute *** コマンドを使用します。

翻訳について

このドキュメントは、米国シスコ発行ドキュメントの参考和訳です。リンク情報につきましては、日本語版掲載時点で、英語版にアップデートがあり、リンク先のページが移動/変更されている場合がありますことをご了承ください。あくまでも参考和訳となりますので、正式な内容については米国サイトのドキュメントを参照ください。