



CHAPTER 26

シスコ エクスプレス フォワーディングの設定

この章では、Catalyst 4000 ファミリースイッチ上の Cisco Express Forwarding (CEF; シスコ エクスプレス フォワーディング) について説明します。この機能の注意事項、設定手順、および設定例についても紹介します。

この章の主な内容は、次のとおりです。

- 「CEF の概要」 (P.26-1)
- 「Catalyst 4500 シリーズ スイッチでの CEF の実装」 (P.26-3)
- 「CEF 設定の制限事項」 (P.26-6)
- 「CEF の設定」 (P.26-6)
- 「CEF の監視およびメンテナンス」 (P.26-8)



(注) この章で使用するスイッチ コマンドの構文および使用方法の詳細については、『*Catalyst 4500 Series Switch Cisco IOS Command Reference*』および次の URL の関連マニュアルを参照してください。

<http://www.cisco.com/univercd/cc/td/doc/product/software/ios124/124cr/index.htm>

CEF の概要

ここでは、CEF の動作を構成する 2 つの主要コンポーネントについて説明します。

- 「CEF の利点」 (P.26-1)
- 「FIB」 (P.26-2)
- 「隣接関係テーブル」 (P.26-2)

CEF の利点

CEF は、先進的なレイヤ 3 IP スイッチングテクノロジーです。ダイナミック トラフィック パターンを特長とする大規模ネットワーク、または Web ベースのアプリケーションおよび対話型セッションを多用するネットワークで、パフォーマンスおよびスケーラビリティを最適化します。

CEF には次の利点があります。

- ルーティング テーブル情報が変更されると、キャッシュ全体がフラッシュされる場合がありますが、CEF はマルチレイヤ スイッチのキャッシュ方式のパフォーマンスを向上させます。
- レイヤ 3 ルーティング情報に基づいて、複数のリンクにパケットを分散させる形式でロード バランシングを行います。ネットワーク デバイスが宛先に対して複数のパスを検出した場合、その宛先に対応する複数のエントリでルーティング テーブルが更新されます。その宛先へのトラフィックは、それらのパスの間で分散されます。

CEF は、マルチレイヤ スイッチのルート キャッシュではなく、複数のデータ構造で情報を保存します。このデータ構造によって検索が最適化され、効率的なパケット転送が実現されます。

FIB

Forwarding Information Base (FIB; 転送情報ベース) は、IP ルーティング テーブルの転送情報のミラー イメージを格納したテーブルです。ネットワークでルーティングまたはトポロジが変更されると、ルート プロセッサが IP ルーティング テーブルを更新し、CEF が FIB を更新します。FIB エントリとルーティング テーブル エントリは 1 対 1 の関係なので、FIB には既知のすべてのルートが含まれます。したがって、高速スイッチング、最適スイッチングなど、スイッチング パスに関連するルート キャッシュのメンテナンスが不要になります。CEF は FIB を使用し、IP 宛先プレフィクスに基づいてスイッチングを決定し、IP ルーティング テーブルの情報に基づいてネクストホップアドレス情報を維持します。

Catalyst 4500 シリーズ スイッチでは、CEF は Integrated Switching Engine ハードウェアに FIB をロードして、転送のパフォーマンスを向上させます。Integrated Switching Engine には、ルーティング情報を保存できる、決まった数のフォワーディング スロットがあります。この限度を超えると、CEF は自動的にディセーブルになり、すべてのパケットがソフトウェアで転送されます。この場合、スイッチ上のルート数を減らしてから、`ip cef` コマンドでハードウェア スwitchングを再度イネーブルにする必要があります。

隣接関係テーブル

CEF は、FIB の他に、隣接関係テーブルを使用してレイヤ 2 アドレス情報を前に付け加えます。ネットワーク内のノードは、相互間の距離が 1 ホップ以内の場合に、隣接しているといいます。隣接関係テーブルで、すべての FIB エントリのレイヤ 2 ネクストホップアドレスが維持されます。

隣接関係の検出

隣接関係テーブルは、新しい隣接ノードが検出されたときに更新されます。(Address Resolution Protocol [ARP; アドレス解決プロトコル] などを使用して) 隣接関係エントリが作成されるたびに、その隣接ノードのリンクレイヤ ヘッダーが隣接関係テーブルに保存されます。ルートが確定すると、リンクレイヤ ヘッダーはネクストホップおよび対応する隣接関係エントリを指し示します。リンクレイヤ ヘッダーはそのあと、CEF パケット スwitchング時のカプセル化に使用されます。

隣接関係の解決

ロード バランシングと冗長性に同時に対応するようにルータが設定されている場合など、宛先プレフィクスへのルートに複数のパスが含まれることがあります。解決されたパスごとに、そのパスのネクストホップ インターフェイスに対応する隣接関係にポインタが追加されます。このメカニズムは、複数パス間でのロード バランシングに使用されます。

特殊な処理が必要な隣接関係タイプ

特定の例外条件が存在する場合、スイッチングを迅速に処理する目的で、ネクストホップ インターフェイスの隣接関係（ホスト ルートの隣接関係）に加え、他のタイプの隣接関係が使用されます。プレフィクスが定義されると、例外処理が必要なプレフィクスは、表 26-1 に示した特殊な隣接関係の 1 つを指定して、キャッシュに格納されます。

表 26-1 例外処理を伴う隣接タイプ

隣接タイプ	処理
Null	インターフェイス Null0 を宛先とするパケットはドロップされます。インターフェイス Null0 は、効果的なアクセス フィルタリング形式として使用できます。
Glean	ルータが複数のホストに直接接続されている場合、ルータ上の FIB テーブルは、個々のホストではなく、サブネットに対応するプレフィクスを維持します。サブネットプレフィクスは、Glean 隣接関係を指し示します。特定のホストにパケットを転送する必要がある場合、隣接データベースから特定のプレフィクスが集められます。
Punt	特殊な処理を必要とする機能や CEF スイッチングでまだサポートされていない機能は、次に上位のスイッチング レベルに送られます（パントされます）。
Discard	パケットが廃棄されます。
Drop	パケットがドロップされます。

未解決の隣接関係

パケットの前にリンクレイヤ ヘッダーが追加された場合、FIB はネクストホップに対応する隣接関係を指し示すプリペンドを要求します。隣接関係が FIB によって作成され、ARP などのメカニズムでは検出されなかった場合、レイヤ 2 アドレッシング情報が不明であるため、その隣接関係は不完全と見なされます。レイヤ 2 情報が認識されると、パケットはルート プロセッサに転送され、ARP によって隣接関係が判別されます。

Catalyst 4500 シリーズ スイッチでの CEF の実装

ここでは、次の内容について説明します。

- 「ハードウェアおよびソフトウェアのスイッチング」(P.26-4)
- 「ロード バランシング」(P.26-6)
- 「ソフトウェア インターフェイス」(P.26-6)

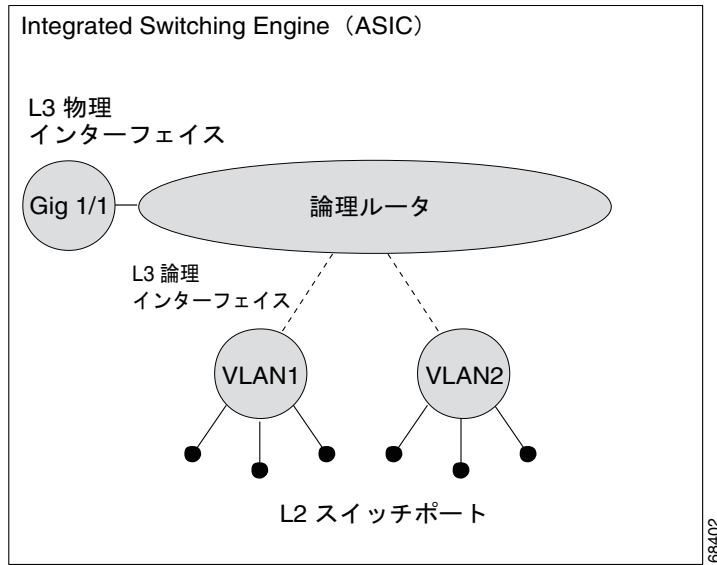
Catalyst 4000 ファミリー スイッチは、ASIC ベースの Integrated Switching Engine をサポートすることにより、次の処理を行います。

- レイヤ 2 でのイーサネットブリッジング
- レイヤ 3 での IP ルーティング

ASIC はパケット転送専用で設計されているので、Integrated Switching Engine ハードウェアは、CPU サブシステム ソフトウェアより、はるかに高速でこの処理を実行できます。

図 26-1 に、Integrated Switching Engine 上の ASIC ベース レイヤ 2 およびレイヤ 3 スイッチング プロセスの概念を示します。

図 26-1 論理 L2/L3 スイッチング コンポーネント



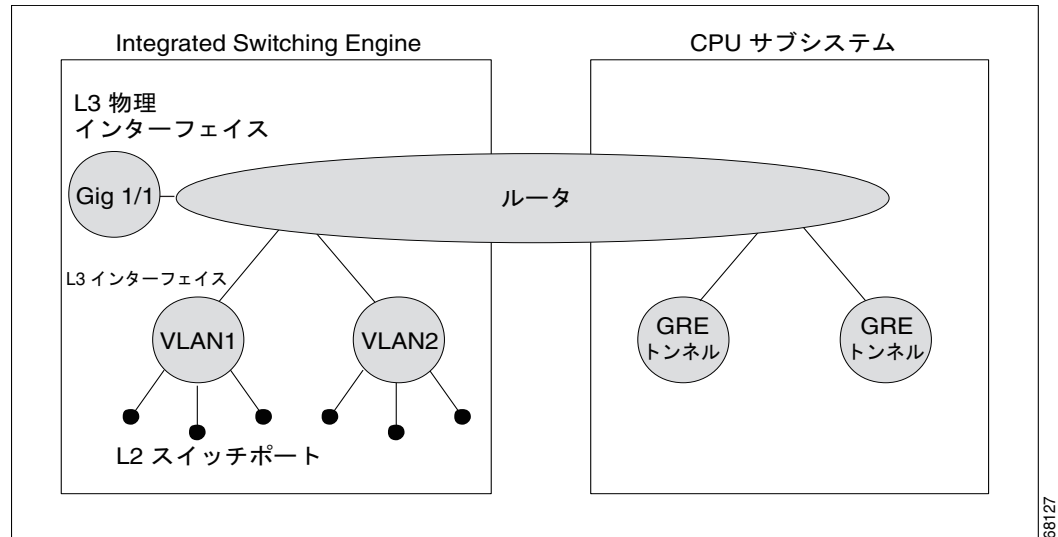
Integrated Switching Engine は、ASIC ハードウェアを使用して、論理レイヤ 3 インターフェイス上で VLAN 間ルーティングを実行します。ASIC ハードウェアは、ホスト、スイッチ、またはルータに接続するように設定できる物理レイヤ 3 インターフェイスもサポートしています。

ハードウェアおよびソフトウェアのスイッチング

通常の packets については、Integrated Switching Engine はハードウェアで packet 転送機能を実行します。これらの packets は非常に高速でハードウェアスイッチングされます。例外の packets は、CPU サブシステム ソフトウェアによって転送されます。Integrated Switching Engine が大部分の packets をハードウェアで転送していることは、統計レポートからわかります。ソフトウェア転送の速度はハードウェア転送に比べて大幅に下がりますが、CPU サブシステムによって packets が転送されても、ハードウェア転送の速度が下がることはありません。

図 26-2 は、Integrated Switching Engine および CPU サブシステムのスイッチング コンポーネントの概念図です。

図 26-2 ハードウェアおよびソフトウェア スイッチングのコンポーネント



Integrated Switching Engine は、ハードウェアで VLAN 間ルーティングを実行します。CPU サブシステム ソフトウェアは、Subnetwork Access Protocol (SNAP) カプセル化を使用する、VLAN へのレイヤ 3 インターフェイスをサポートします。CPU サブシステム ソフトウェアはさらに、Generic Routing Encapsulation (GRE; 総称ルーティング カプセル化) トンネルもサポートしています。

ハードウェア スイッチング

ハードウェア スイッチングは、Supervisor Engine III および Supervisor Engine IV の標準動作です。

ソフトウェア スイッチング

ソフトウェア スイッチングは、ハードウェアでトラフィックを処理できない場合に行われます。次のタイプの例外パケットがソフトウェアで処理されますが、速度は大幅に低下します。

- IP ヘッダー オプションを使用しているパケット



(注) TCP ヘッダー オプションを使用しているパケットは、転送の決定に影響しないので、ハードウェアでスイッチングされます。

- IP Time To Live (TTL; 存続可能時間) カウンタが満了しているパケット
- トンネル インターフェイスに転送されるパケット
- サポート対象外のカプセル化タイプを指定して送られてきたパケット
- サポート対象外のカプセル化タイプが設定されたインターフェイスにルーティングされるパケット
- 出力インターフェイスの MTU を超えているため分割が必要なパケット
- IGMP リダイレクトのルーティングが必要なパケット
- 802.3 イーサネット パケット

ロード バランシング

Catalyst 4000 ファミリー スイッチは、Integrated Switching Engine ハードウェアで、パケット ルーティングのロード バランシングをサポートします。ロード バランシング機能は常にイネーブルです。この機能は、同一ネットワークに複数の異なるネクストホップ アドレスで複数のルートが設定されている場合に動作します。このようなルートは、静的に設定することも、OSPF、EIGRP などのルーティング プロトコルによって設定することもできます。

ハードウェアは、ハードウェア ハッシュ機能を使用し、送信元および宛先 IP アドレス、送信元および宛先 TCP ポート番号（該当する場合）に基づいて値を計算することによって、転送を決定します。さらに、この負荷分散型ハッシュ値を使用して、パケット転送に使用するルートが選択されます。特定のフロー（TCP 接続など）内のすべてのハードウェア スイッチングが同じネクストホップにルーティングされるので、パケットの順序変更が発生する可能性が小さくなります。1つのネットワークに対して最大 8 種類のルートがサポートされます。

ソフトウェア インターフェイス

Catalyst 4000 ファミリー スイッチ対応の Cisco IOS は、ハードウェア転送エンジンには組み込まれていない GRE および IP トンネル インターフェイスをサポートしています。これらのインターフェイスとの間を流れるパケットはすべて、ソフトウェアで処理する必要があります。転送速度はハードウェア スイッチング インターフェイスに比べて大幅に低下します。また、これらのインターフェイスでは、レイヤ 2 機能はサポートされません。

CEF 設定の制限事項

Integrated Switching Engine がハードウェアでのレイヤ 3 スイッチングでサポートするカプセル化タイプは、ARPA および ISL/802.1Q だけです。CPU サブシステムは、レイヤ 2 対応の SNAP など、ソフトウェアでのレイヤ 3 スイッチングに使用できるさまざまなカプセル化をサポートします。

CEF の設定

ここでは、CEF を設定する手順について説明します。

- 「CEF のイネーブル化」(P.26-6)
- 「CEF のロード バランシングの設定」(P.26-7)

CEF のイネーブル化

Catalyst 4000 ファミリー スイッチでは、CEF はデフォルトの設定でグローバルにイネーブルになっています。したがって設定作業は不要です。

CEF を再度イネーブルにするには、次の作業を行います。

コマンド	目的
Switch(config)# <code>ip cef</code>	標準の CEF 動作をイネーブルにします。

CEF のロード バランシングの設定

CEF のロード バランシングは、送信元および宛先パケット情報の組み合わせに基づいて行われます。宛先ヘッダを転送するときに、複数のパスにトラフィックを分散させることによって、リソースの利用を最適化できます。ロード バランシング機能は、宛先単位で設定できます。ロード バランシングの決定は、発信インターフェイスで行われます。CEF のロード バランシングを設定するときには、宛先単位で発信インターフェイスに設定してください。

ここでは、次の内容について説明します。

- 「宛先別ロード バランシングの設定」(P.26-7)
- 「負荷分散型ハッシュ機能の設定」(P.26-7)
- 「CEF 情報の表示」(P.26-8)

宛先別ロード バランシングの設定

CEF がイネーブルの場合、宛先別ロード バランシングはデフォルトでイネーブルです。宛先別ロード バランシングを使用する場合は、CEF をイネーブルにすれば、追加の作業は必要ありません。

宛先別ロード バランシングにより、ルータは複数のパスを使用して負荷を分散させます。特定の送信元または宛先ホストのペアに対応するパケットは、複数のパスが使用できる場合でも、必ず同じパスを通ることになります。異なるペア宛てのトラフィックは、それぞれ異なるパスを通る傾向があります。CEF をイネーブルにすると、宛先別ロード バランシングはデフォルトでイネーブルになります。このロード バランシング方式はさまざまな環境に適しています。

宛先別ロード バランシングは、統計上のトラフィック分散に依存するので、送信元ペアまたは宛先ペアの数が増えるほど、負荷分散の効果が大きくなります。

宛先別ロード バランシングを使用すると、特定のホスト ペア宛てのパケットが順番に届くようになります。特定のホスト ペア宛てのパケットはすべて、同じリンク（1 つまたは複数のリンク）を使用してルーティングされます。

負荷分散型ハッシュ機能の設定

宛先 IP プレフィクスに対して複数のユニキャスト ルートが存在する場合、ハードウェアは、プレフィクスに一致するパケットを使用可能なすべてのルートで送信するので、すべてのネクストホップ ルータで負荷が分散されます。デフォルトでは、送信元および宛先 IP アドレスのハッシュを計算し、ルート選択用に導き出された値を使用して使用されるルートが選択されます。この結果、1 つの IP 送信元フローまたは宛先フロー内のすべてのパケットが必ず同じルート上で送信されるため、複数のパケットについてパケット順序が保持されますが、ルートに対してフローをほぼランダムに分散します。

負荷分散型ハッシュ機能は変更できるので、送信元および宛先 IP アドレス以外に、送信元 TCP/UDP ポート、宛先 TCP/UDP ポート、またはその両方をハッシュに組み込むことができます。

送信元または宛先ポート、あるいはその両方を使用するように負荷分散型ハッシュ機能を設定するには、次の作業を行います。

コマンド	目的
Switch (config)# [no] ip cef load-sharing algorithm include-ports source destination]	送信元および宛先ポートを使用する負荷分散型ハッシュ機能をイネーブルにします。 デフォルトの Cisco IOS 負荷分散アルゴリズムを使用するようにスイッチを設定するには、 no キーワードを指定します。

負荷分散の詳細については、次の URL で Cisco IOS マニュアルの「*Configuring Cisco Express Forwarding*」モジュールを参照してください。
http://www.cisco.com/en/US/docs/ios/12_2/switch/configuration/guide/xcfcefc.html



(注) **include-ports** オプションは、Catalyst 4500 シリーズ スイッチ上のソフトウェアでスイッチングされるトラフィックには適用されません。

CEF 情報の表示

収集された CEF 情報を表示できます。CEF を表示するには、次の作業を行います。

コマンド	目的
Switch# show ip cef	収集された CEF 情報を表示します。

CEF の監視およびメンテナンス

IP トラフィックに関する情報を表示するには、次の作業を行います。

コマンド	目的
Switch# show interface type slot/interface begin L3	IP ユニキャスト トラフィックのサマリーを表示します。

次に、インターフェイス FastEthernet 3/3 上の IP ユニキャスト トラフィック情報を表示する例を示します。

```
Switch# show interface fastethernet 3/3 | begin L3
  L3 in Switched: ucast: 0 pkt, 0 bytes - mcast: 12 pkt, 778 bytes mcast
  L3 out Switched: ucast: 0 pkt, 0 bytes - mcast: 0 pkt, 0 bytes
  4046399 packets input, 349370039 bytes, 0 no buffer
  Received 3795255 broadcasts, 2 runts, 0 giants, 0 throttles
<...output truncated...>
Switch#
```



(注) IP ユニキャスト パケット カウントは、約 5 秒間隔で更新されます。

IP 統計情報の表示

IP ユニキャスト統計情報は、インターフェイス単位で収集されます。IP 統計情報を表示するには、次の作業を行います。

コマンド	目的
Switch# show interface type number counters detail	IP 統計情報を表示します。

次に、インターフェイス FastEthernet 3/1 の IP ユニキャスト統計情報を表示する例を示します。

```
Switch# show interface fastethernet 3/1 counters detail

Port          InBytes      InUcastPkts  InMcastPkts  InBcastPkts
Fa3/1         7263539133  5998222     6412307      156

Port          OutBytes      OutUcastPkts  OutMcastPkts  OutBcastPkts
Fa3/1         7560137031  5079852     12140475     38

Port          InPkts 64      OutPkts 64      InPkts 65-127  OutPkts 65-127
Fa3/1         11274    168536    7650482     12395769

Port          InPkts 128-255  OutPkts 128-255  InPkts 256-511  OutPkts 256-511
Fa3/1         31191    55269     26923       65017

Port          InPkts 512-1023  OutPkts 512-1023
Fa3/1         133807    151582

Port          InPkts 1024-1518  OutPkts 1024-1518  InPkts 1519-1548  OutPkts 1519-1548
Fa3/1         N/A        N/A                N/A                N/A

Port          InPkts 1024-1522  OutPkts 1024-1522  InPkts 1523-1548  OutPkts 1523-1548
Fa3/1         4557008    4384192           0                 0

Port          Tx-Bytes-Queue-1  Tx-Bytes-Queue-2  Tx-Bytes-Queue-3  Tx-Bytes-Queue-4
Fa3/1         64                0                  91007              7666686162

Port          Tx-Drops-Queue-1  Tx-Drops-Queue-2  Tx-Drops-Queue-3  Tx-Drops-Queue-4
Fa3/1         0                  0                  0                  0

Port          Rx-No-Pkt-Buff    RxPauseFrames     TxPauseFrames     PauseFramesDrop
Fa3/1         0                  0                  0                  N/A

Port          UnsupOpcodePause
Fa3/1         0
Switch#
```

CEF (ソフトウェア スイッチング) およびハードウェア IP ユニキャスト隣接テーブル情報を表示するには、次の作業を行います。

コマンド	目的
Switch# show adjacency [<i>interface</i>] [detail internal summary]	オプションの detail キーワードを指定すると、レイヤ 2 情報を含めた詳細な隣接情報が表示されます。

次に、隣接統計情報を表示する例を示します。

```
Switch# show adjacency gigabitethernet 3/5 detail
Protocol Interface          Address
IP          GigabitEthernet9/5  172.20.53.206(11)
                                     504 packets, 6110 bytes
                                     00605C865B82
                                     000164F83FA50800
ARP          03:49:31
```



(注) 隣接統計情報は、約 10 秒間隔で更新されます。

