

Cisco Express Forwarding (CEF) について

目次

[概要](#)

[前提条件](#)

[要件](#)

[使用するコンポーネント](#)

[表記法](#)

[概要](#)

[CEF の処理](#)

[GRP のルーティング テーブルの更新](#)

[OC48 および QOC12 を除くすべてのラインカードの packets 転送](#)

[OC48 および QOC12 ラインカードの packets 転送](#)

[関連情報](#)

概要

この文書では、Cisco Express Forwarding の概要と、Cisco 12000 シリーズ インターネット ルータでの実装方法について説明します。

前提条件

要件

このドキュメントに関する固有の要件はありません。

使用するコンポーネント

このドキュメントは、特定のソフトウェアやハードウェアのバージョンに限定されるものではありません。

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されたものです。このドキュメントで使用するすべてのデバイスは、クリアな (デフォルト) 設定で作業を開始しています。ネットワークが稼働中の場合は、コマンドが及ぼす潜在的な影響を十分に理解しておく必要があります。

表記法

ドキュメント表記の詳細は、『[シスコ テクニカル ティップスの表記法](#)』を参照してください。

概要

Cisco Express Forwarding (CEF) スイッチングは、キャッシング要求に関連した問題に対処するための、Cisco 独自のスケーラブルなスイッチングの形態です。CEF スイッチングでは、従来の方法でルート キャッシュに保管された情報が、複数のデータ構造をまたがって分割されます。CEF コードは、これらのデータ構造を Gigabit Route Processor (GRP; ギガビット ルート プロセッサ) だけでなく、12000 ルータのラインカードなどのスレーブ プロセッサでも維持できます。効率的なパケット フォワーディングのために最適化されたルックアップを実現するデータ構造には、次のものがあります。

- Forwarding Information Base (FIB; 転送情報ベース) テーブル : CEF は FIB を使用して、IP の宛先プレフィックススペースのスイッチング決定を行います。FIB は概念的には、ルーティング テーブルや情報ベースに類似します。FIB は、IP ルーティング テーブルに含まれるフォワーディング情報のミラー イメージを維持します。ネットワーク内でルーティングまたはトポロジの変更が発生した場合、IP ルーティング テーブルは更新され、これらの変更が FIB に反映されます。FIB は 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。FIB エントリとルーティング テーブルのエントリには 1 対 1 の相関関係があるため、FIB には既知のルートのすべてが含まれ、高速スイッチングや最適スイッチングなどのスイッチング パスに関連するルート キャッシュのメンテナンスは不要になります。
- 隣接テーブル : ネットワーク内のノードは、リンク層の全域で 1 ホップで相互に到達できれば、隣接とみなされます。FIB 以外にも、CEF は隣接関係テーブルを使用してレイヤ 2 アドレッシング情報を先頭に付加します。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。

CEF は次の 2 つのうちどちらかのモードでイネーブルされます。

- セントラル CEF モード : CEF モードが有効のとき、CEF FIB および隣接関係テーブルはルート プロセッサに常駐し、ルート プロセッサが高速フォワーディングを実行します。CEF スイッチングに使用可能なラインカードがない場合、または分散 CEF スイッチングと互換性のない機能を使用する必要がある場合、CEF モードが使用できます。
- 分散 CEF (dCEF) モード : dCEF が有効のとき、ラインカードは FIB テーブルおよび隣接関係テーブルとまったく同じコピーを保持します。ラインカードはそれ自体が高速フォワーディングを実行できるので、スイッチング操作にはメイン プロセッサ (Gigabit Route Processor (GRP)) の関与は不要です。これは、Cisco 12000 シリーズ ルータで使用可能な唯一のスイッチング方式です。dCEF では、ルート プロセッサとラインカード上の FIB および隣接テーブルの同期を保証するために、Inter-Process Communication (IPC; プロセス間通信) メカニズムが使用されています。

CEF スイッチングについての詳細は、『[Cisco Express Forwarding \(CEF \) White Paper](#)』を参照してください。

CEF の処理

GRP のルーティング テーブルの更新

図 1 は、ルーティング アップデート パケットが Gigabit Route Processor (GRP) に送信され、結果のアップデート転送メッセージがラインカードの上の FIB テーブルに送信されるプロセスを示しています。

説明をわかりやすくするため、次の文中の番号は、図 1 の番号に対応しています。次のプロセスは、ルート テーブルの初期化時、またはネットワークトポロジが変更されるあらゆる時点 (ルー

トの追加、削除、または変更時)で発生します。図 1 に示すプロセスには、5 つの主なステップが含まれます。

1. IP データグラムが受信ラインカード (入力ラインカード) の入力バッファに配置され、L2/L3 フォワーディング エンジンがパケット内のレイヤ 2 およびレイヤ 3 の情報にアクセスし、その情報を転送プロセッサに送信します。転送プロセッサは、パケットにルーティング情報が含まれていることを確認します。転送プロセッサは、ポインタを GRP Virtual Output Queue (VOQ; 仮想出力キュー) に送信し、バッファ メモリ内のパケットを GRP に送信する必要があることを示します。
2. ラインカードは、Clock and Scheduler Card (CSC; クロック スケジューラ カード) に要求を発行します。スケジューラ カードが許可を発行し、パケットはスイッチング ファブリックを経由して GRP に送信されます。
3. GRP はルーティング情報を処理します。GRP 上の R5000 (プロセッサ) が、ネットワーク ルーティング テーブルを更新します。パケットのルーティング情報に応じて、レイヤ 3 プロセッサは、(内部ルーティング プロトコルが Open Shortest Path First (OSPF) である場合) 隣接ルータにリンクステート情報をフラッディングする必要がある場合があります。プロセッサは、FIB テーブルのリンクステート情報と内部アップデートを運ぶ IP パケットを生成します。また、内部プロトコルと外部ゲートウェイ プロトコル (Border Gateway Protocol (BGP; ボーダーゲートウェイ プロトコル) など) の両方がサポートされている場合に発生する、すべての再帰ルートを、GRP が計算します。事前に計算された再帰ルート情報は、各ラインカードの FIB に送信されます。ラインカード上のレイヤ 3 プロセッサはパケットの転送に重点を置き、再帰ルートの計算には重点を置いていないため、これにより転送プロセスが大幅に高速化されます。
4. GRP は、GRP 上のものを含めて、すべてのラインカード上の FIB テーブルに内部アップデートを送信します。ラインカードへの FIB のアップデートは監視され、必要に応じて制限されます。GRP は各ラインカードの FIB テーブルのコピーを持っているため、新しいラインカードがシャーシに装着された場合、そのカードがアクティブになると、GRP は最新の転送情報をその新しいカードにダウンロードします。
5. 新しい近隣ルータが 12000 ルータに接続されるたびに、GRP はラインカードから通知を受けます。ラインカード上のプロセッサは、新しいレイヤ 2 情報 (通常は Point-to-Point Protocol (PPP; ポイントツーポイント プロトコル) ヘッダー情報) を含むパケットを GRP に送信します。GRP はこのレイヤ 2 情報を使用して、GRP 上およびラインカード上に存在する隣接関係テーブルを更新します。パケットが 12000 ルータから送信された時点で、各ラインカードはこのレイヤ 2 情報を各パケットに追加します。隣接関係テーブルのコピーは、初期化のために GRP で維持されます。

図 1: [パスの決定とレイヤ 3 スイッチングの図](#)

[OC48 および QOC12 を除くすべてのラインカードのパケット転送](#)

ラインカードに、スイッチング ファブリックを通じてパスを決定するための十分な転送情報があれば (ネクストホップの宛先など)、12000 ルータはパケットの転送が可能です。次のステップに、12000 ルータにより使用されるシンプルかつ高速な転送テクニックの概要を示します ([図 1](#) を参照)。説明をわかりやすくするため、文中のアルファベットは、図 1 のアルファベットに対応しています。

- A. IP データグラムが受信ラインカード (Rx ラインカード) の入力バッファに配置され、L2/L3 フォワーディング エンジンがパケット内のレイヤ 2 およびレイヤ 3 の情報にアクセスし、その情報を転送プロセッサに送信します。転送プロセッサは、パケットにデータが含ま

れること、およびそれがルーティング アップデートではないことを確認します。FIB テーブルのレイヤ 2 およびレイヤ 3 情報に基づいて、転送プロセッサは、適切なラインカードの VOQ にポインタを送信して、バッファ メモリ内のパケットがそのラインカードに送信されることを示します。

- **B.** ラインカードのスケジューラが、スケジューラに要求を出します。スケジューラが許可を発行し、パケットはバッファ メモリからスイッチング ファブリックを経由してラインカード (Tx ラインカード) に送信されます。
- **C.** Tx ラインカードは着信パケットをバッファに入れます。
- **D.** Tx ラインカード上のレイヤ 3 プロセッサと、関連付けられた Application-Specific Integrated Circuits (ASIC; 特定用途集積回路) が、レイヤ 2 情報 (PPP アドレス) を、送信される各パケットに添付します。(必要に応じて) ラインカード上の各ポートに対してパケットが複製されます。
- **E.** Tx ラインカードのトランスミッタにより、ファイバ インターフェイス経由でパケットが送信されます。

このシンプルな転送プロセスの利点は、大部分のデータ転送タスクが ASIC で行えるため、12000 がギガビット レートで動作できることです。また、データ パケットは GRP には送信されません。

OC48 および QOC12 ラインカードのパケット転送

ラインカードに、スイッチング ファブリックを通じてパスを決定するための十分な転送情報があれば (ネクストホップの宛先など)、12000 ルータはパケットの転送が可能です。以降のステップでは、12000 ルータで使用されるシンプルかつ超高速の転送テクニックを示しています ([図 2](#) を参照)。説明をわかりやすくするため、文中のアルファベットは、図 2 のアルファベットに対応しています。

- **A.** (ルーティング アップデート、Internet Control Message Protocol (ICMP; インターネット制御メッセージ プロトコル)、オプション付きの IP パケットではなく) IP データグラムがラインカード内に受信され、レイヤ 2 処理を通過します。ローカル FIB テーブルのレイヤ 2 およびレイヤ 3 情報に基づいて、Fast Packet Processor はパケットの宛先を決定し、パケットのヘッダーを修正します。宛先に基づいて、パケットは適切なラインカードの VOQ に配置されます。
- **B.** Fast Packet Processor がパケットを適切に転送できない、まれな場合では、パケットは転送プロセッサにより処理されます。ローカル FIB テーブルのレイヤ 2 およびレイヤ 3 情報に基づいて、転送プロセッサは、適切なラインカードの VOQ にポインタを送信し、バッファ メモリ内のパケットがそのラインカードに送信されることを示します。
- **C.** パケットが適切な VOQ に入れられれば、ラインカードのスケジューラがスケジューラに要求を発行します。スケジューラが許可を発行し、パケットはバッファ メモリからスイッチング ファブリックを経由してラインカード (Tx ラインカード) に送信されます。
- **D.** Tx ラインカードは着信パケットをバッファに入れます。
- **E.** Tx ラインカード上のレイヤ 3 プロセッサと、関連付けられた ASIC が、レイヤ 2 情報 (PPP アドレス) を、送信される各パケットに添付します。(必要に応じて) ラインカード上の各ポートに対してパケットが複製されます。
- **F.** Tx ラインカードのトランスミッタにより、ファイバ インターフェイス経由でパケットが送信されます。

この新しい転送プロセスの利点は、OC48/STM16 などの、特により高速な転送を目的とするカードを最適化することにあります。

図 2 : [高速ラインカードの packets 交換](#)

[関連情報](#)

- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ - シャーシ](#)
- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ - スイッチ ファブリック](#)
- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ : ルート プロセッサ](#)
- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ : ライン カードの設計](#)
- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ : メモリの詳細](#)
- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ : メンテナンス バス、電源と
ブローワー、およびアラーム カード](#)
- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ : ソフトウェアの概要](#)
- [Cisco 12000 シリーズ インターネット ルータのアーキテクチャ - パケット スイッチング](#)
- [テクニカルサポート - Cisco Systems](#)