

Cisco Express Forwardingについて

内容

[はじめに](#)

[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パケット転送](#)

[OC48 および QOC12 ラインカードのパケット転送](#)

[関連情報](#)

はじめに

このドキュメントでは、Cisco Express Forwarding(CEF)について説明します switching Cisco
12000シリーズインターネットルータでの実装方法を説明します

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

このドキュメントの内容は、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ました。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。本稼働中のネットワークでは、各コマンドによって起こる可能性がある影響を十分確認してください。

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セカンダリプロセッサでも維持できます。効率的なパケット転送のために最適化されたルックアップを提供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に加えて、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追加します。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で高速転送を実行できるため、スイッチング操作にメインプロセッサ(Gigabit Route Processor (GRP ; ギガビットルートプロセッサ))が関与する必要がなくなります。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 ; ギガビットルートプロセッサ) 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FIBテーブルに送信されるプロセスを示しています。

分かりやすくするために、次の段落の番号付けは、図1の番号付けに対応しています。

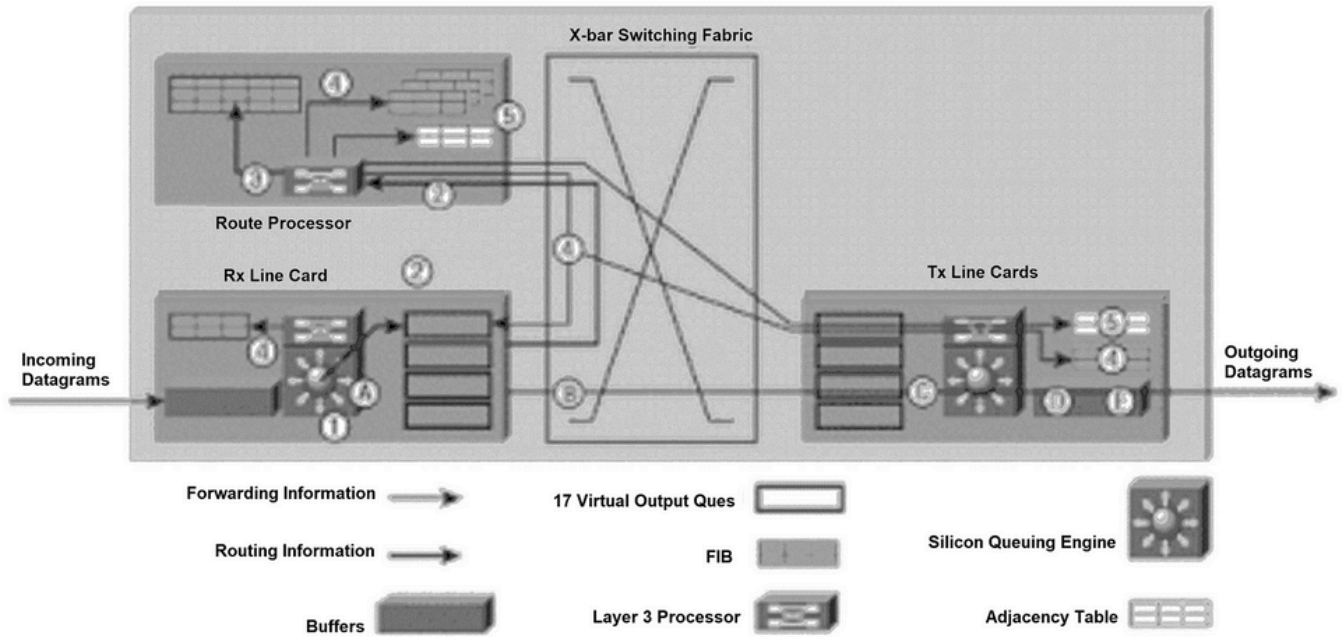
次のプロセスは、ルートテーブルの初期化中、またはネットワークポロジが変更されたとき (ルートが追加、削除、または変更されたとき) に発生します。図 1 に示すプロセスには、5 つの主なステップが含まれます。

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仮想出力キュー(VOQ)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。パケット内のルーティング情報に依存して、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]の場合) リンクステート情報を隣接ルータに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。さらに、GRPは、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]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計算します。

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。これにより、ラインカード上のレイヤ3プロセッサがパケットの転送に集中でき、再帰ルートを計算しないので、転送プロセスが大幅に高速化されます。

4. GRPは、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。これには、GRP上のアップデートも含まれます。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には各ラインカードのFIBテーブルのコピーがあるため、新しいラインカードがシャーシに挿入されると、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(通常はポイントツーポイントプロトコル(PPP)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



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 packets 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。次の手順では、12000ルータで使用されるシンプルで高速な転送技術の概要を示します(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）上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情報にアクセスし、それをフォワーディングプロセッサに送信します。転送プロセッサは、パケットに含まれること、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 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。
- B.ラインカードスケジューラは、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。
- 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プロセッサおよび関連する特定用途向け集積回路(ASIC)は、送信される各パケットにレイヤ2情報（PPPアドレス）を添付します。（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は、ファイバインターフェイス経由でパケットを送信し

ます。

このシンプルな転送プロセスの利点は、ほとんどのデータ転送タスクがASICで実行でき、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。

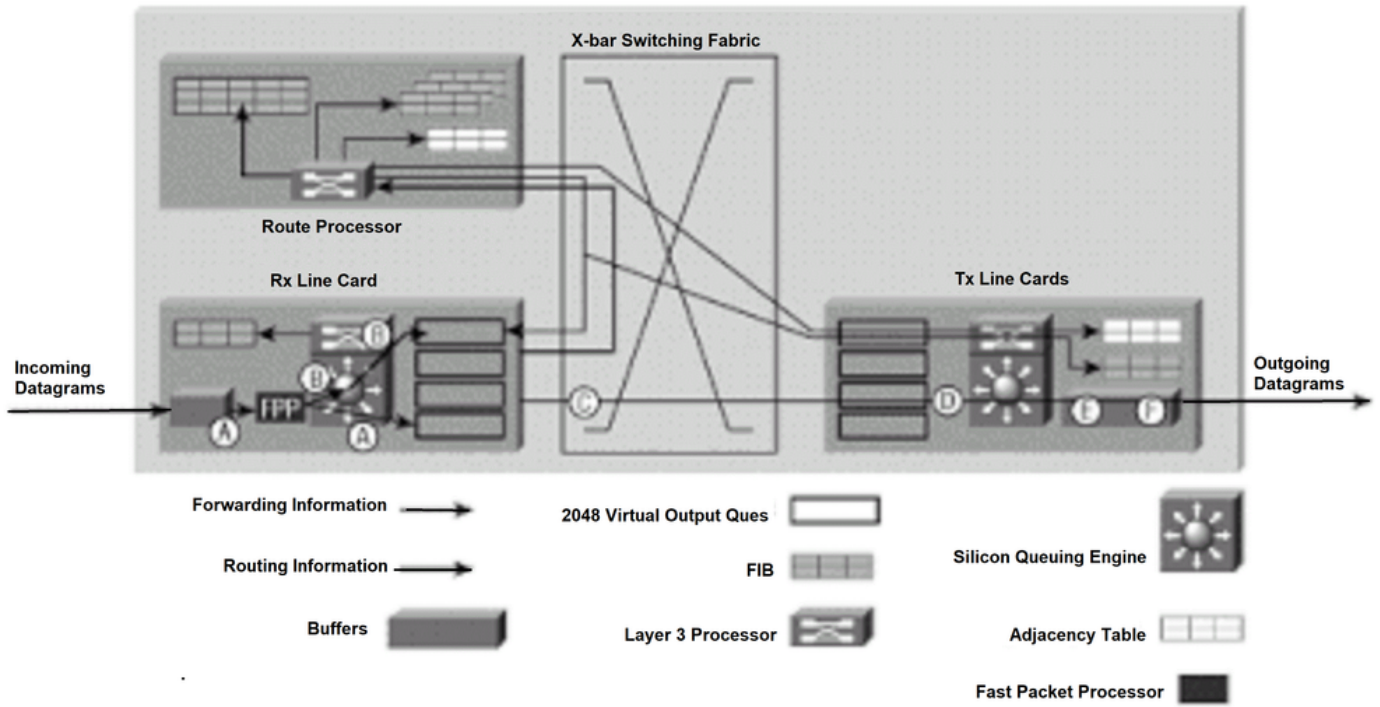
OC48 および QOC12 ラインカードのパケット転送

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。次のステップは、12000で使用されるシンプルで超高速な転送テクニックを構成します(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. IPデータグラム(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)、およびオプション付きのIPパケットではありません)がラインカードで受信され、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VOQに配置されます。
- B. まれな場合ですが、Fast Packet Processor がパケットを適切に転送できない場合、パケットは転送プロセッサにより処理されます。転送プロセッサは、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、適切なラインカードVOQにポインタを送信します。これは、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に入ると、ラインカードスケジューラは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. 着信パケットは、Tx ラインカードによりバッファに入れられます。
- E. Tx ラインカード上のレイヤ 3 プロセッサと関連 ASIC により、送信される各パケットにレイヤ 2 情報 (PPP アドレス) が添付されます。(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パケット交換



高速ラインカードの packets 交換

関連情報

- [テクニカルサポート - Cisco Systems](#)

翻訳について

シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。