

瞭解Cisco Express Forwarding

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[慣例](#)

[概觀](#)

[CEF操作](#)

[更新GRP路由表](#)

[除OC48和QOC12外的所有線卡的封包轉送](#)

[OC48和QOC12線卡的資料包轉發](#)

[相關資訊](#)

簡介

本檔案將說明思科快速轉送(CEF) switching 以及如何在Cisco 12000系列網際網路路由器中實施。

必要條件

需求

本文件沒有特定需求。

採用元件

本文件所述內容不限於特定軟體和硬體版本。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除 (預設) 的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

慣例

如需文件慣例的詳細資訊，請參閱思科技術提示慣例。

概觀

Cisco Express Forwarding(CEF)交換是一種專有的可擴展交換形式，旨在解決與需求快取相關的問題。通過CEF交換，通常儲存在路由快取中的資訊在幾個資料結構上被拆分。CEF代碼能夠在Gigabit路由處理器(GRP)中以及在輔助處理器 (例如Gigabit路由器中的線卡) 中維護這些數12000結構。為高效資料包轉發提供最佳化查詢的資料結構包括：

- 轉發資訊庫(FIB)表 — CEF使用FIB做出基於IP目標字首的交換決策。FIB在概念上類似於路由表或資訊庫。它維護IP路由表中包含的轉發資訊的映象。當網路中發生路由或拓撲更改時，IP路由表將更新，這些更改將反映在FIB中。FIB根據IP路由表中的資訊維護下一跳地址資訊。由於FIB條目和路由表條目之間存在一對一的關聯，因此FIB包含所有已知路由，並且無需進行與交換路徑（例如快速交換和最佳交換）相關聯的路由快取維護。
- 鄰接表 — 如果網路中的節點能夠通過鏈路層的一跳到達彼此，則稱其為相鄰節點。除FIB外，CEF還使用鄰接表新增第2層地址資訊。鄰接表為所有FIB條目維護第2層下一跳地址。

可以在以下兩種模式之一中啟用CEF：

- 中央CEF模式 — 啟用CEF模式時，CEF FIB和鄰接表駐留在路由處理器上，而路由處理器執行快速轉發。當線卡不可用於CEF交換時，或者需要使用與分散式CEF交換不相容的功能時，可以使用CEF模式。
- 分散式CEF(dCEF)模式 — 啟用dCEF時，線卡會維護FIB和鄰接表的相同副本。線卡可以自己執行快速轉送，這樣就可以避免主處理器 — 千兆路由處理器(GRP)參與交換操作。這是Cisco 12000系列路由器上可用的唯一交換方法。

dCEF使用進程間通訊(IPC)機制來確保FIB和路由處理器和線卡上的鄰接表的同步。

有關CEF交換的詳細資訊，請參閱[Cisco Express Forwarding\(CEF\)白皮書](#)。

CEF操作

更新GRP路由表

圖1說明了將路由更新資料包傳送到Gigabit路由處理器(GRP)的過程，以及將生成的轉發更新消息傳送到線卡上的FIB表。

為了清楚起見，後面段落的編號與圖1中的編號相對應。

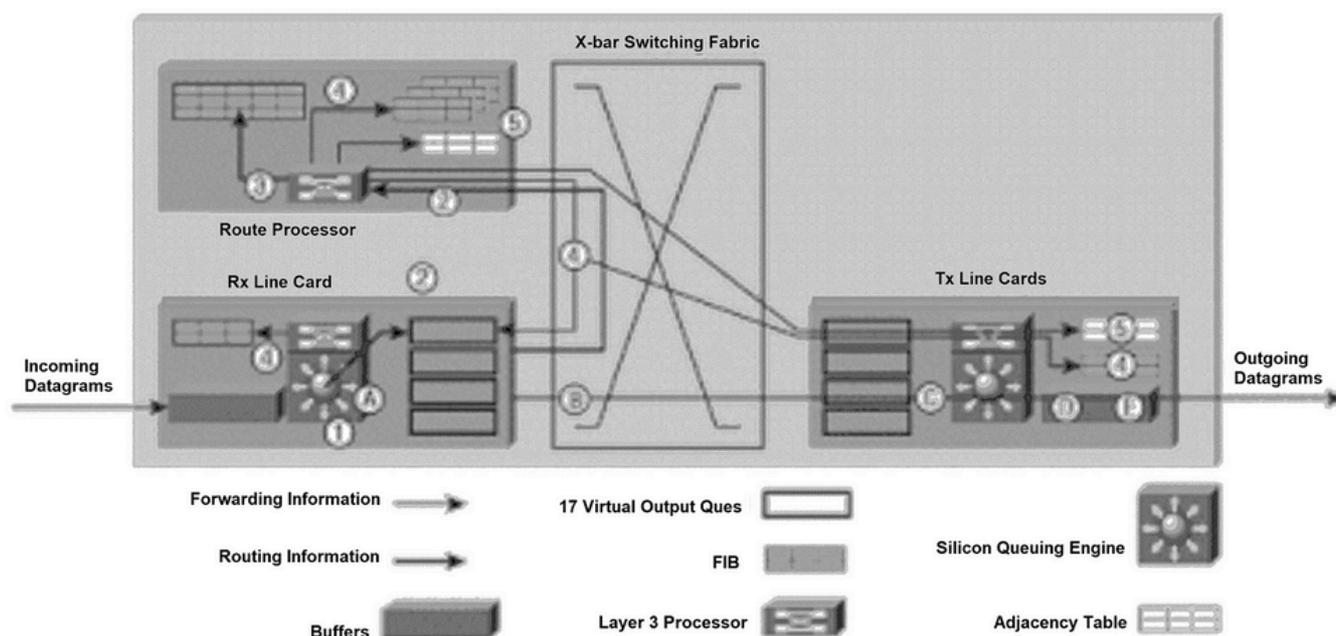
下一個過程發生在路由表初始化期間，或者網路拓撲發生變化的任何時間（新增、刪除或更改路由時）。圖1中所示的過程包含五個主要步驟：

1. 將IP資料包放置在接收線卡（入口線卡）上的輸入緩衝區中，並且L2/L3轉發引擎訪問資料包中的第2層和第3層資訊，並將其傳送到轉發處理器。轉發處理器確定資料包包含路由資訊。轉送處理器將指標傳送到GRP虛擬輸出隊列(VOQ)，並指示必須將緩衝區儲存器中的資料包傳送到GRP。
2. 線路卡向時鐘和排程器卡(CSC)發出請求。排程器卡發出授權，資料包通過交換網狀架構傳送到GRP。
3. GRP處理路由資訊。GRP上的R5000（處理器）更新網路路由表。依賴資料包中的路由資訊，第3層處理器可能必須將鏈路狀態資訊泛洪到相鄰路由器（如果內部路由協定是開放最短路徑優先[OSPF]）。處理器產生攜帶鏈路狀態資訊和FIB表的內部更新的IP資料包。此外，GRP還會計算當同時支援內部協定和外部網關協定（例如，邊界網關協定[BGP]）時產生的所有遞迴路由。

計算的遞迴路由資訊被傳送到每個線卡上的FIB。這顯著加快了轉發過程，因為線卡上的第3層處理器可以專注於轉發資料包，而不計算遞迴路由。

4. GRP向所有線卡上的FIB表傳送內部更新，並包括GRP上的更新。線路卡的FIB更新會受到監控並需要限制。GRP具有每個線卡FIB表的副本，因此，如果將新的線卡插入機箱中，GRP會在新卡變為活動狀態時將最新的轉發資訊下載到新卡。
5. 每當新鄰居路由器連線到GRP路由器時，都會通過線卡通12000GRP。線卡上的處理器會將包含新第2層資訊(通常是點對點通訊協定(PPP)標頭資訊)的封包傳送到GRP。GRP使用此第2層資訊更新位於GRP和線卡上的鄰接表。每個線卡在資料包從IP路由器傳送時，會將此第2層資訊新增到每12000資料包。鄰接表的副本會保留在GRP上，以便進行初始化。

圖1：路徑確定和第3層交換圖



路徑決定和第3層交換圖

除OC48和QOC12外的所有線卡的封包轉送

一旦線卡有足夠的轉發資訊來確定通過交換結構的路徑（例如，下一跳的目標），12000路由器就可以轉發資料包了。接下來的步驟將概述路由器使用的簡單快速12000發技術(請參見圖1)。為清楚起見，各段落的字母與圖1中的字母相對應。

- A. 將IP資料包置於接收線卡（Rx線卡）上的輸入緩衝區中，L2/L3轉發引擎訪問資料包中的第2層和第3層資訊，並將其傳送到轉發處理器。轉發處理器確定資料包包含資料，而不是路由更新。根據FIB表中的第2層和第3層資訊，轉發處理器將指標傳送到適當的線卡VOQ，指示緩衝區儲存器中的資料包將傳送到該線卡。
- B. 線卡排程式向排程式發出請求。排程式發出授權，資料包通過交換網狀架構從緩衝儲存器傳送到線卡（Tx線卡）。
- C. Tx線卡緩衝傳入資料包。

- D. Tx線卡上的第3層處理器和相關聯的特定應用積體電路(ASIC)將第2層資訊 (PPP地址) 附加到每個傳輸的資料包。線卡上的每個連線埠都會複製封包 (如果需要)。
- E. Tx線路卡發射器透過光纖介面傳送封包。

這種簡單轉發過程的優勢在於，大多數資料傳輸任務可在ASIC中完成，並允許12000務以千兆速率運行。此外，資料包永遠不會傳送到GRP。

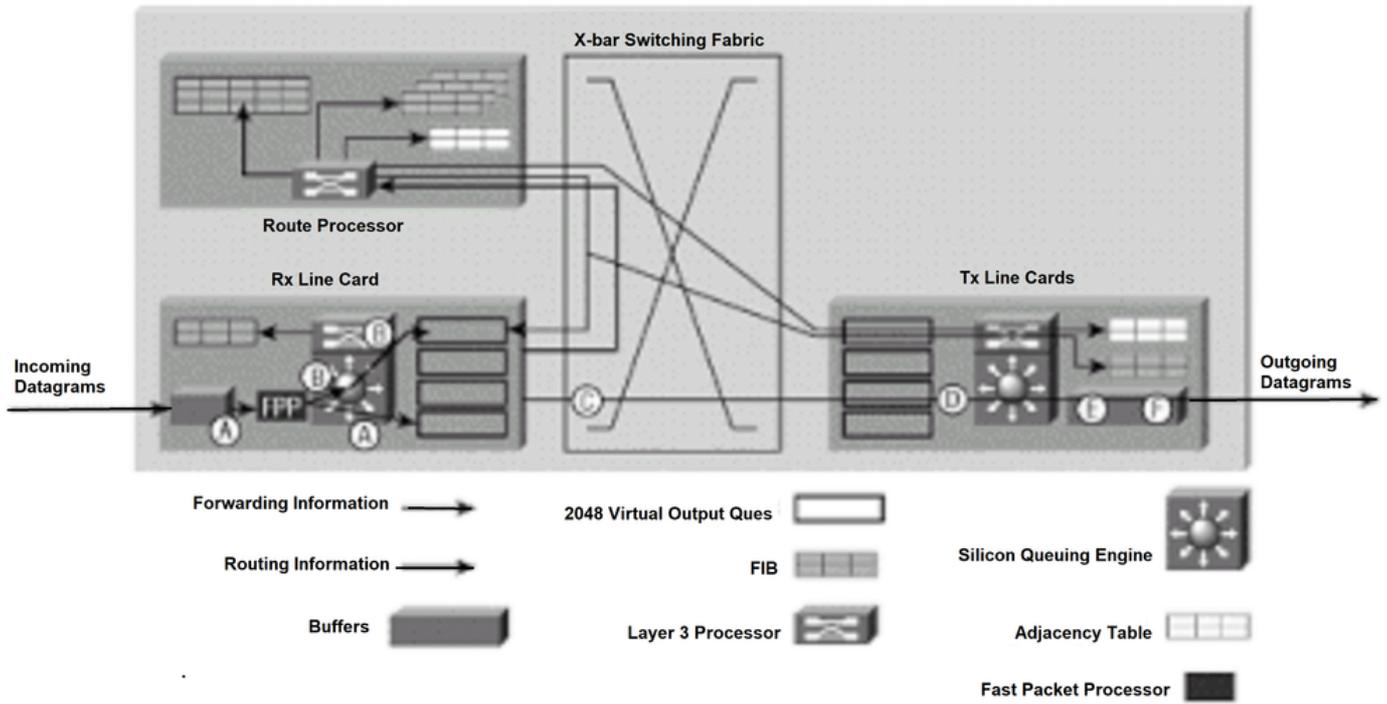
OC48和QOC12線卡的資料包轉發

當線卡有足夠的轉發資訊來確定通過交換結構的路徑 (例如，下一跳的目標) 時，12000路由器就可以轉發資料包了。接下來的步驟彌補了路由器使用的簡單和超快速轉發12000術(請參見圖2)。為清楚起見，各段落的字母與圖2中的字母相對應。

- A. 將IP資料包(不是路由更新、網際網路控制訊息通訊協定(ICMP)和具有選項的IP封包)接收到線卡並經過第2層處理。根據本地FIB表中的第2層和第3層資訊，快速資料包處理器確定資料包的目的地，並修改資料包報頭。然後，根據目的地，將封包放在適當的線路卡VOQ中。
- B. 在快速封包處理器無法正確轉送封包的極少數情況下，轉送處理器會處理封包。轉發處理器基於其本地FIB表的第2層和第3層資訊，將指標傳送到適當的線卡VOQ，指示緩衝區儲存器中的資料包將傳送到該線卡。
- C. 一旦封包位於適當的VOQ中，線路卡排程器就會向排程器發出要求。排程程式發出授權，資料包通過交換網狀架構從緩衝儲存器傳送到線卡 (Tx線卡)。
- D. Tx線卡緩衝傳入資料包。
- E. 第3層處理器和Tx線卡上的相關ASIC將第2層資訊 (PPP地址) 附加到每個傳輸的資料包。線卡上的每個連線埠都會複製封包 (如果需要)。
- F. Tx線路卡發射器通過光纖介面傳送資料包。

新轉發過程的優點是它專門為更快的速度最佳化了卡，如OC48/STM16。

圖2：用於速度更快的線卡的資料包交換



適用於更快線路卡的分組交換

相關資訊

- [技術支援 - Cisco Systems](#)

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。