

MPLS網路中的跟蹤路由

目錄

[簡介](#)

[背景資訊](#)

[MPLS網路中的ICMP路徑追蹤](#)

[從PE到遠端PE觸發的ICMP跟蹤](#)

[從CE觸發到遠端CE的ICMP跟蹤](#)

[MPLS網路中的MPLS LSP路徑追蹤](#)

[從PE到遠端PE觸發LSP跟蹤](#)

[從CE觸發到遠端CE的LSP跟蹤](#)

[相關資訊](#)

簡介

本檔案介紹多重通訊協定標籤交換(MPLS)網路中的網際網路控制訊息通訊協定(ICMP)追蹤行為，以及與LSP追蹤的快速比較。

背景資訊

在IP環境中，任何節點在收到封包且存留時間(TTL)到期時，預期會產生「TTL Exceeded」ICMP錯誤訊息(型別=11，代碼=0)並將它傳送到封包來源位址。利用此概念，通過從1開始按順序傳送TTL的UDP資料包，跟蹤從源到目的地的IP路徑。請注意，此功能的基本要求如下：

- 可從傳輸節點到達封包的來源位址
- 沿路徑未過濾ICMP

在MPLS環境中，傳輸提供商LSR可能並不總是能夠到達源地址，因此需要對MPLS域中的ICMP處理進行一些增強。

MPLS網路中的ICMP路徑追蹤

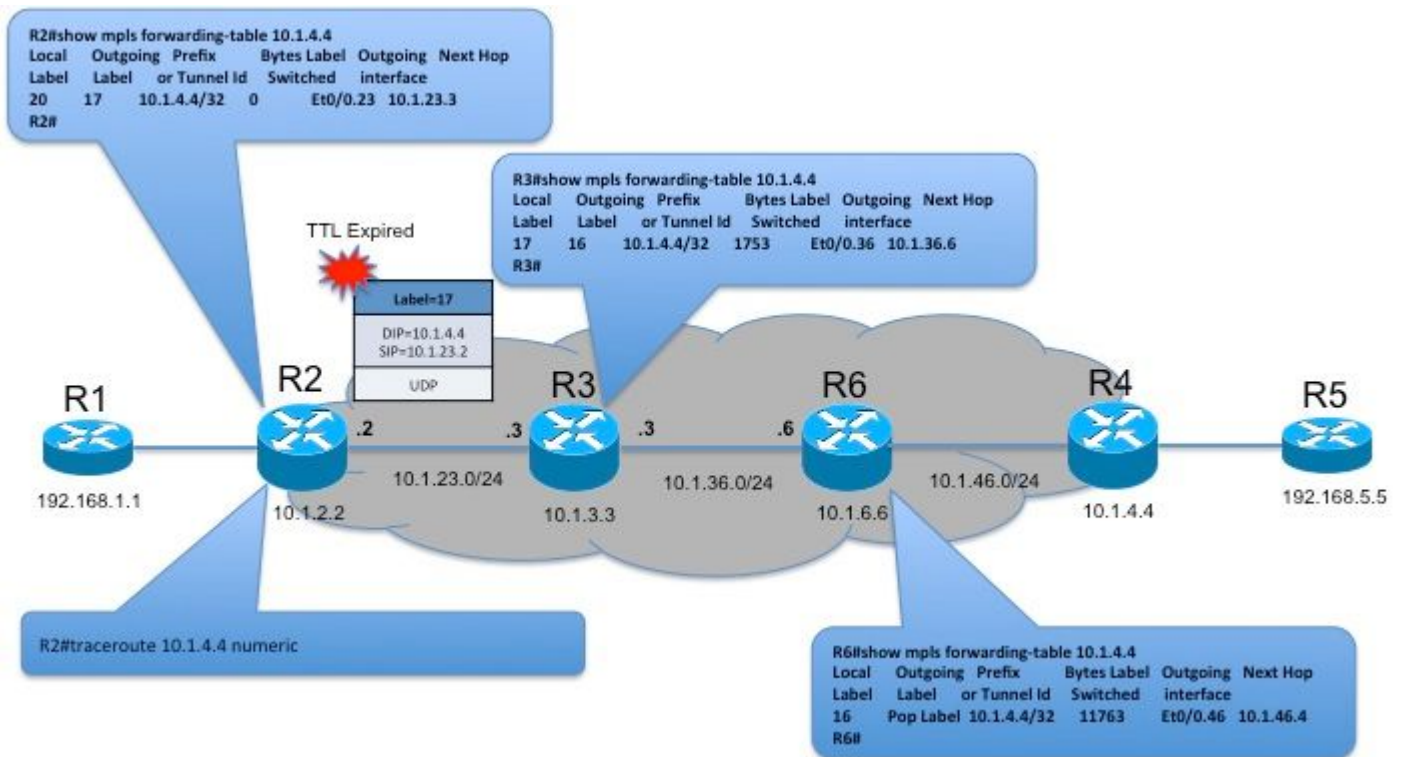
在接收頂部標籤上帶有TTL = 1的資料包時，任何LSR的預設行為都遵循傳統的IP行為，即丟棄資料包並觸發ICMP錯誤消息。若要將ICMP訊息路由到來源，LSR將執行以下操作：

- 從傳入資料包(TTL=1接收的資料包)緩衝標籤堆疊
- 從接收的資料包生成ICMP錯誤消息，其中源地址為自己的地址，目標地址為源地址。
- 附加標籤堆疊底部的所有標籤(在步驟1之前已緩衝)，TTL = 255(頂部標籤除外)。
- 從緩衝的標籤堆疊中獲取頂標籤並執行本地LFIB查詢，以獲得要交換的標籤和關聯的下一跳。
- 以TTL = 255將新標籤追加到堆疊頂部，並進行傳送。

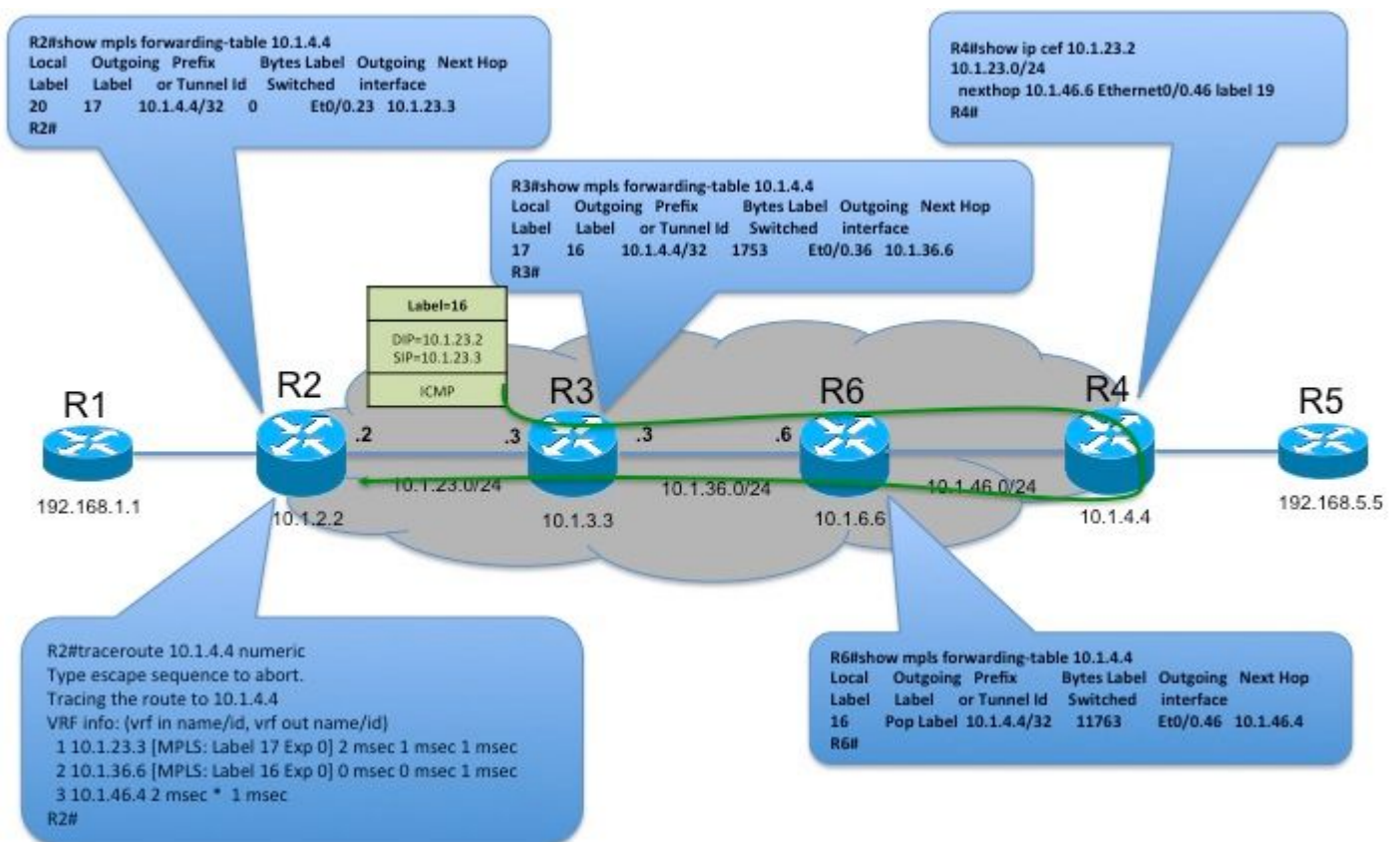
透過此方法，ICMP錯誤訊息會從傳輸LSR遍歷到輸出LER，然後返回到輸入LER以到達實際來源。

從PE到遠端PE觸發的ICMP跟蹤

以下是一個簡單的示例，說明在同一MPLS域中從PE向遠端PE觸發ICMP跟蹤時的行為：



在此拓撲中，當ICMP traceroute從R2觸發到10.1.4.4時，第一個資料包的TTL為1。R3收到資料包時會將TTL遞減到0並觸發ICMP生成機制。



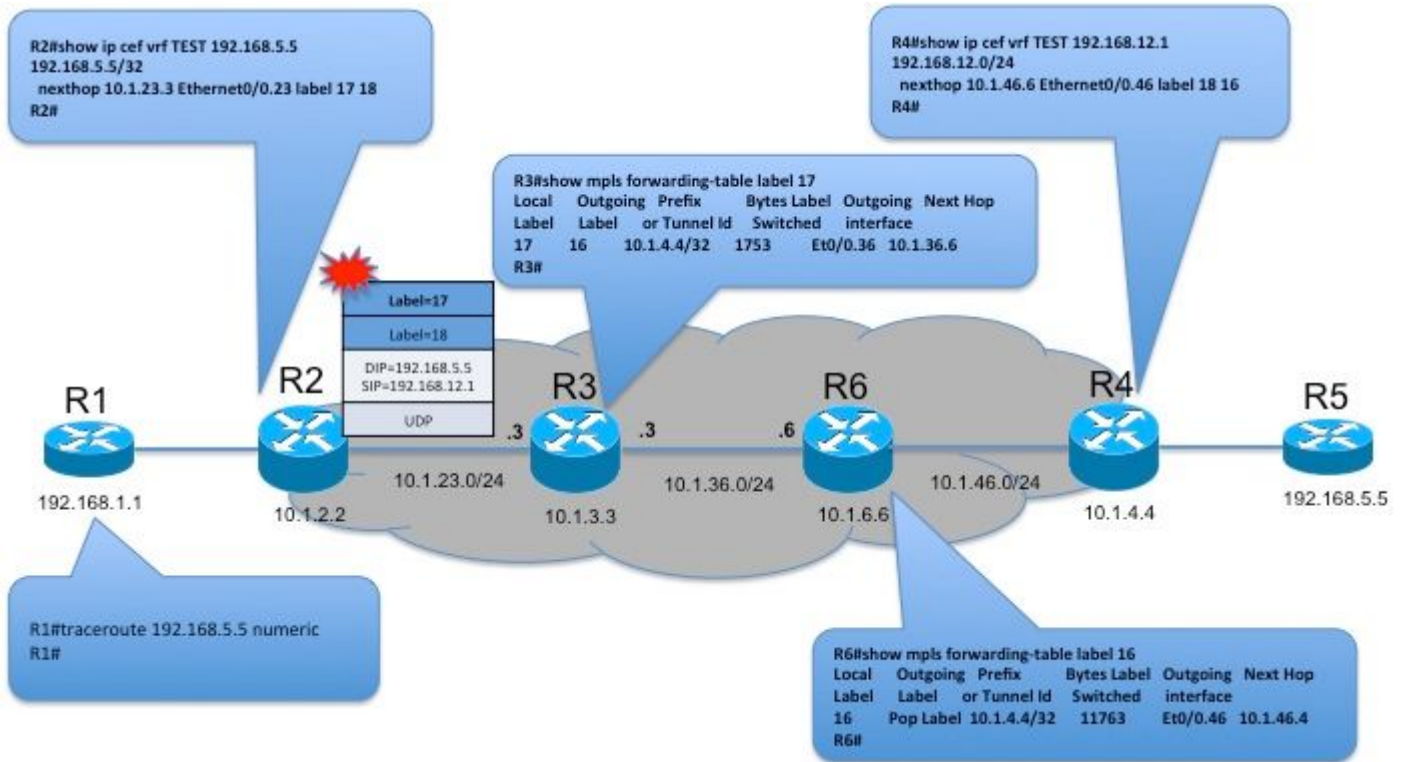
R3將緩衝標籤堆疊並生成ICMP錯誤消息，並在ICMP負載中包括來自緩衝區的傳入標籤堆疊。它進一步使用來自標籤資料包的傳入介面的源地址填充IP報頭，目標地址作為標籤資料包的源。TTL設定為255。它現在從緩衝區推送標籤堆疊，並查詢LFIB表以轉發頂標籤上的操作。在此拓撲中，收到的標籤堆疊為17。在LFIB表中執行查詢時，標籤17與標籤16交換並向下一跳R6轉發。R6將彈出

頂部標籤並轉發到R4，後者將IP將資料包轉發回R2。

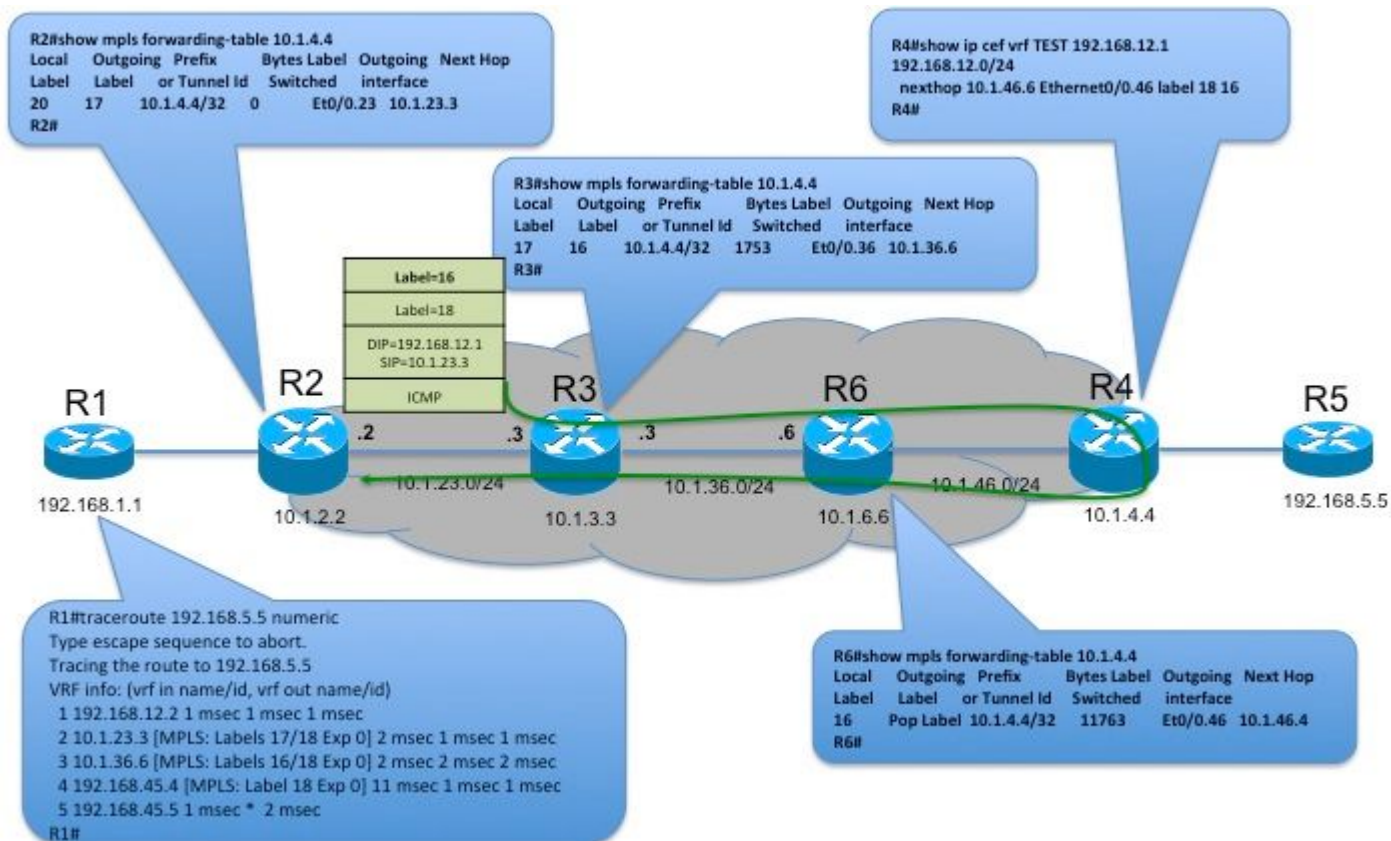
在R2的traceroute輸出中可以看到，傳入標籤將按路徑上的每一跳列出。

從CE觸發到遠端CE的ICMP跟蹤

以下是一個簡單的示例，說明通過MPLS域從CE觸發到遠端CE的ICMP跟蹤的行為：



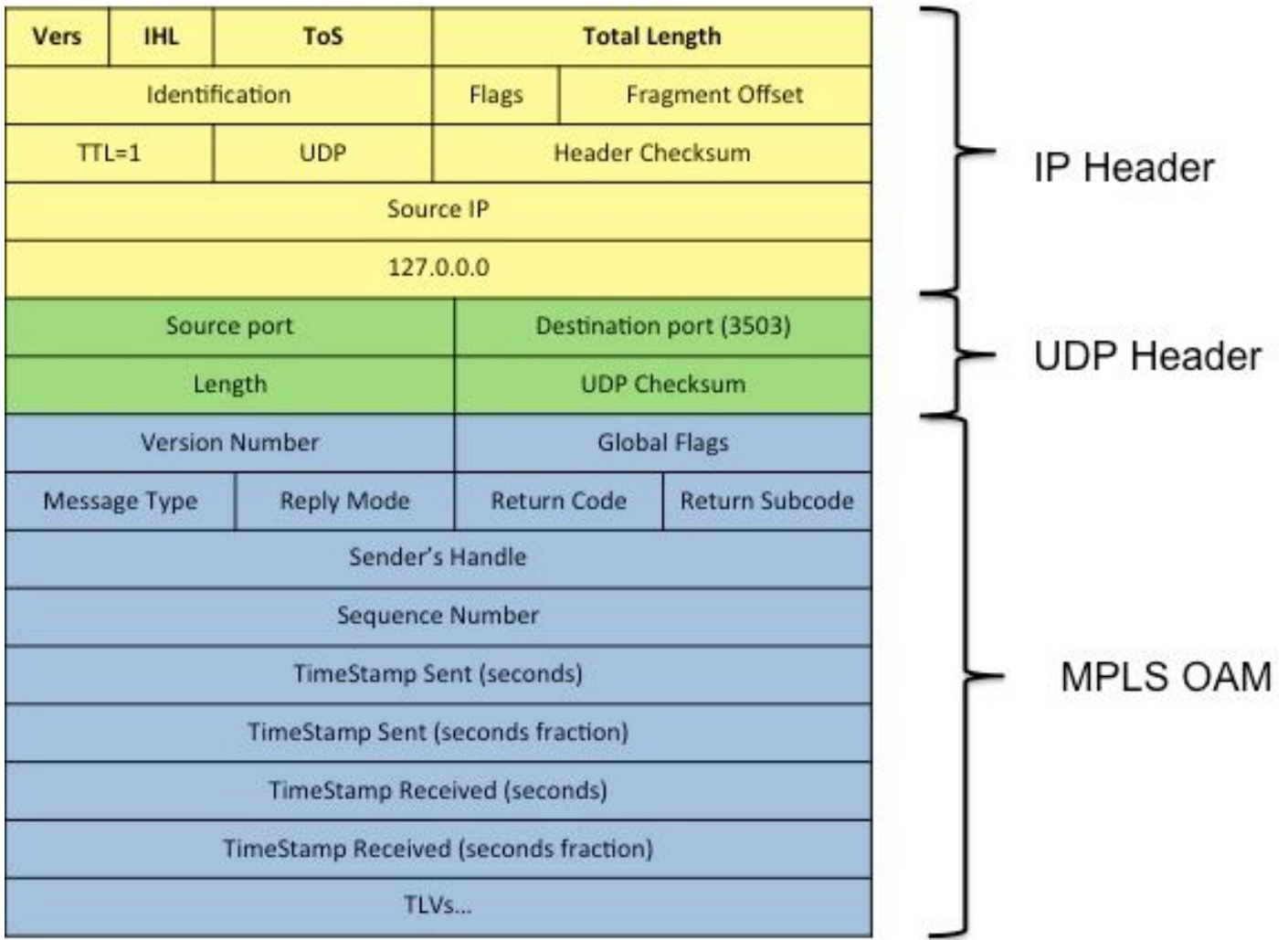
在此拓撲中，當從R1(CE)觸發到192.168.5.5 (遠端CE)的ICMP traceroute時，第一個資料包的TTL為1。這是正常的IP資料包，因此R2遵循生成ICMP並直接傳送到R1的傳統行為。使用TTL=2傳送的第二個資料包將在R3過期。



R3將緩衝標籤堆疊並生成ICMP錯誤消息，並在ICMP負載中包括來自緩衝區的傳入標籤堆疊。它進一步使用來自標籤資料包的傳入介面的源地址填充IP報頭，目標地址作為標籤資料包的源。TTL設定為255。它現在從緩衝區推送標籤堆疊，並查詢LFIB表以轉發頂標籤上的操作。在上述拓撲中，收到的標籤堆疊為{17, 18}。在LFIB表中查詢頂部標籤時，17將與標籤16交換，並轉發到下一跳R6。R6將彈出頂部標籤並轉發到R4。R4將使用VRF標籤標識VRF，並將資料包轉發到R1。

在R1的traceroute輸出中可以看到，沿途的每一跳都會列出傳入的標籤堆疊。

MPLS網路中的MPLS LSP路徑追蹤

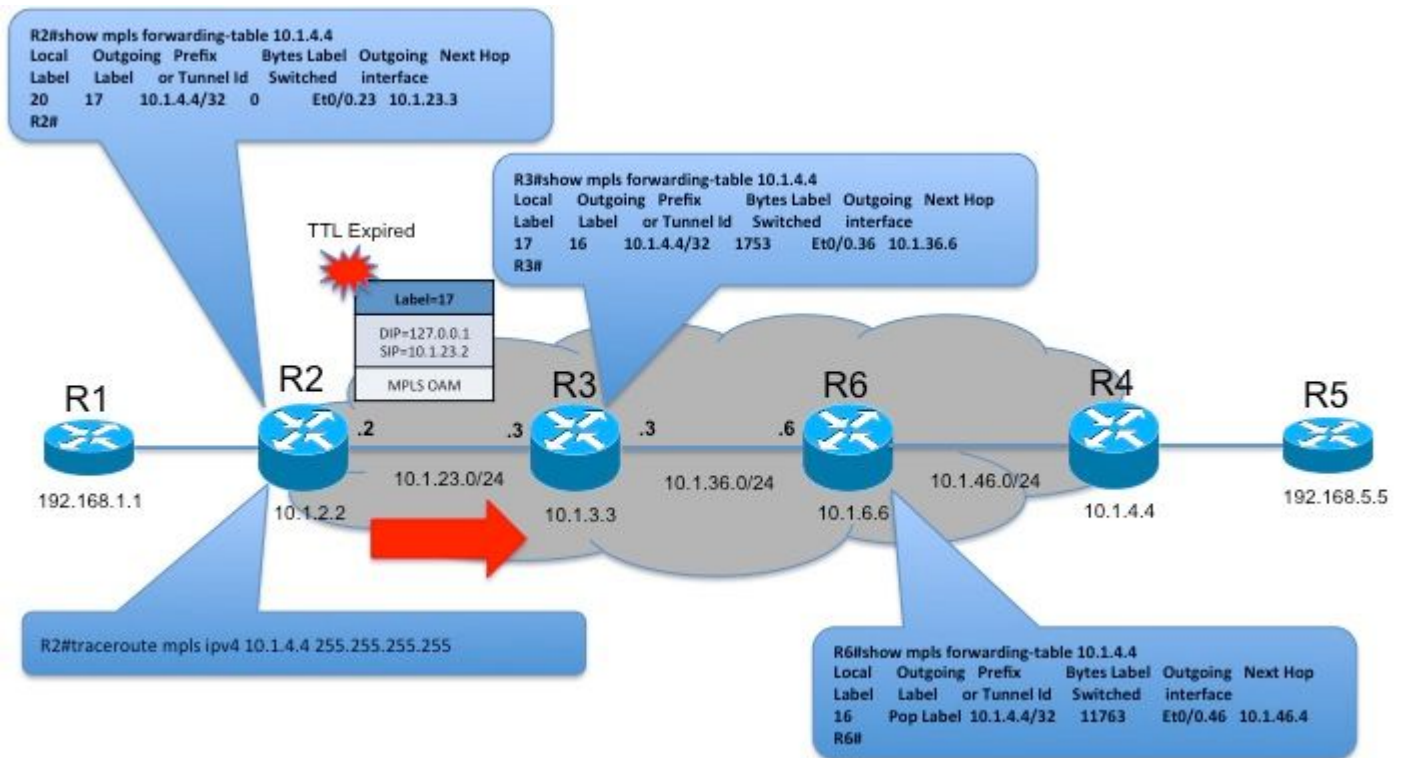


與基於ICMP的traceroute不同，LSP traceroute使用RFC4379中定義的機制。它使用IP/UDP封裝，請求的目標地址設定為環回地址(127.0.0.0/8範圍)。預計在同一個MPLS域內觸發LSP Ping，因此回覆將直接傳送到發起方。

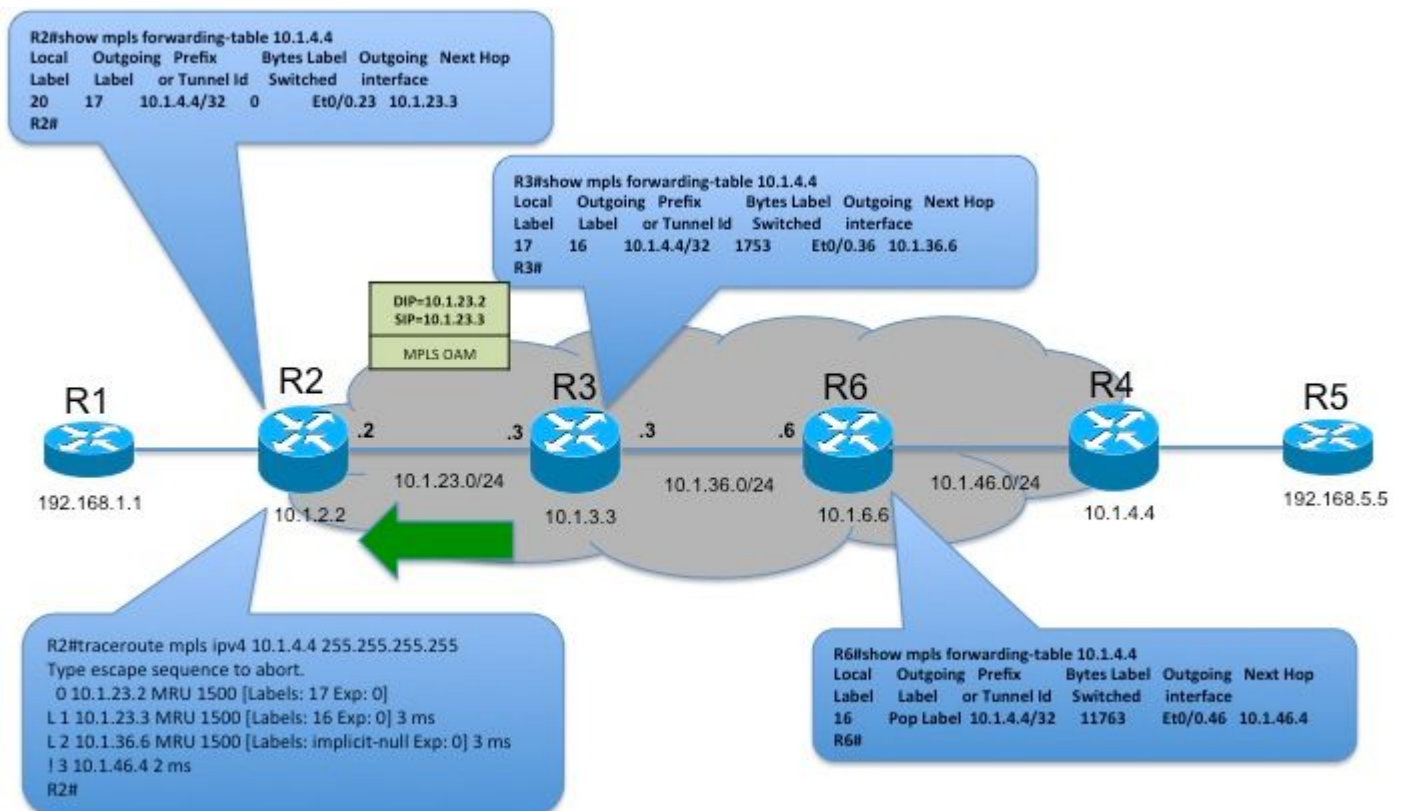
從任何LSR觸發LSP路徑追蹤(「traceroute mpls ipv4 <FEC>」)時，有關要驗證的FEC的詳細資訊將作為「目標FEC堆疊」包含在MPLS回應請求中的TLV中。此消息將與Label堆疊上的TTL一起從1開始按順序傳送。接收資料包時的任何傳輸LSR以及TTL到期都將處理IP資料包，因為目的地址是環回地址。並轉換為CPU以處理MPLS OAM。

響應器可選擇地執行FEC驗證，方法是從收到的MPLS回應請求的標籤堆疊提取標籤，並從目標FEC堆疊TLV提取FEC詳細資訊，以根據本地控制平面資訊驗證該資訊。在跟蹤情況下，響應方會將下游資訊(如傳出標籤和下游鄰居地址等)包括在作為下游對映(DSMAP)TLV的TLV中。(DSMAP將被DDMAP否決，因為它比DSMAP更靈活)。

從PE到遠端PE觸發LSP跟蹤



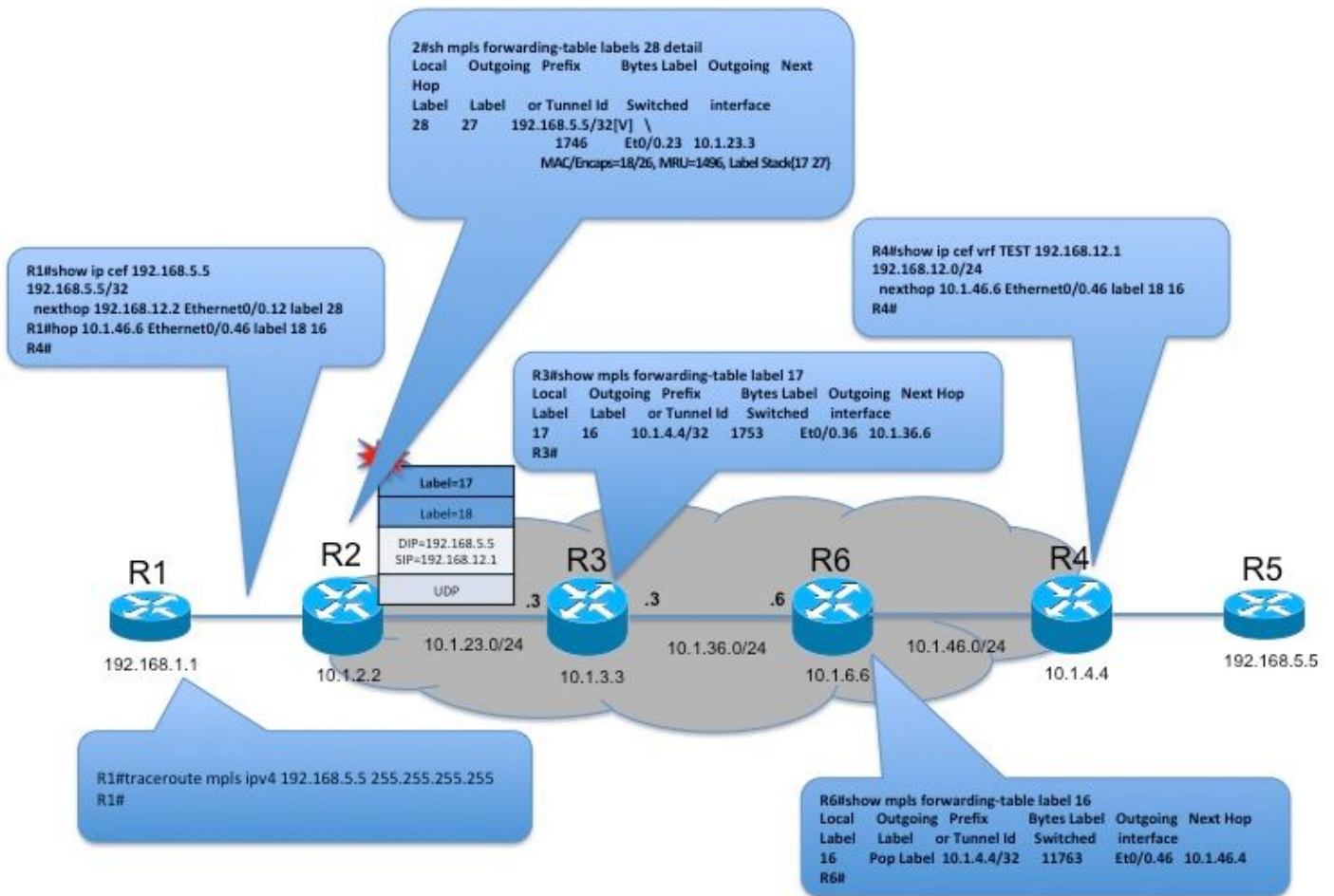
在此拓撲中，從R2觸發LSP跟蹤，以將LSP驗證為字首10.1.4.4/32。標籤上的TTL將從1設定。R3收到標籤時將其推送到CPU進行OAM處理。



R3將使用MPLS Echo Reply with DSMTP TLV回覆R2,DSMTP TLV攜帶傳出標籤16和諸如下游鄰居詳細資訊等附加資訊。與ICMP消息不同，MPLS回應回覆將從響應方R3直接轉發到發起方R2。

在R2的LSP traceroute輸出中可以看到，傳出標籤堆疊將按路徑上的每一跳列出。這與基於ICMP的traceroute不同，其中，輸出中列出的標籤是傳入標籤堆疊。

從CE觸發到遠端CE的LSP跟蹤

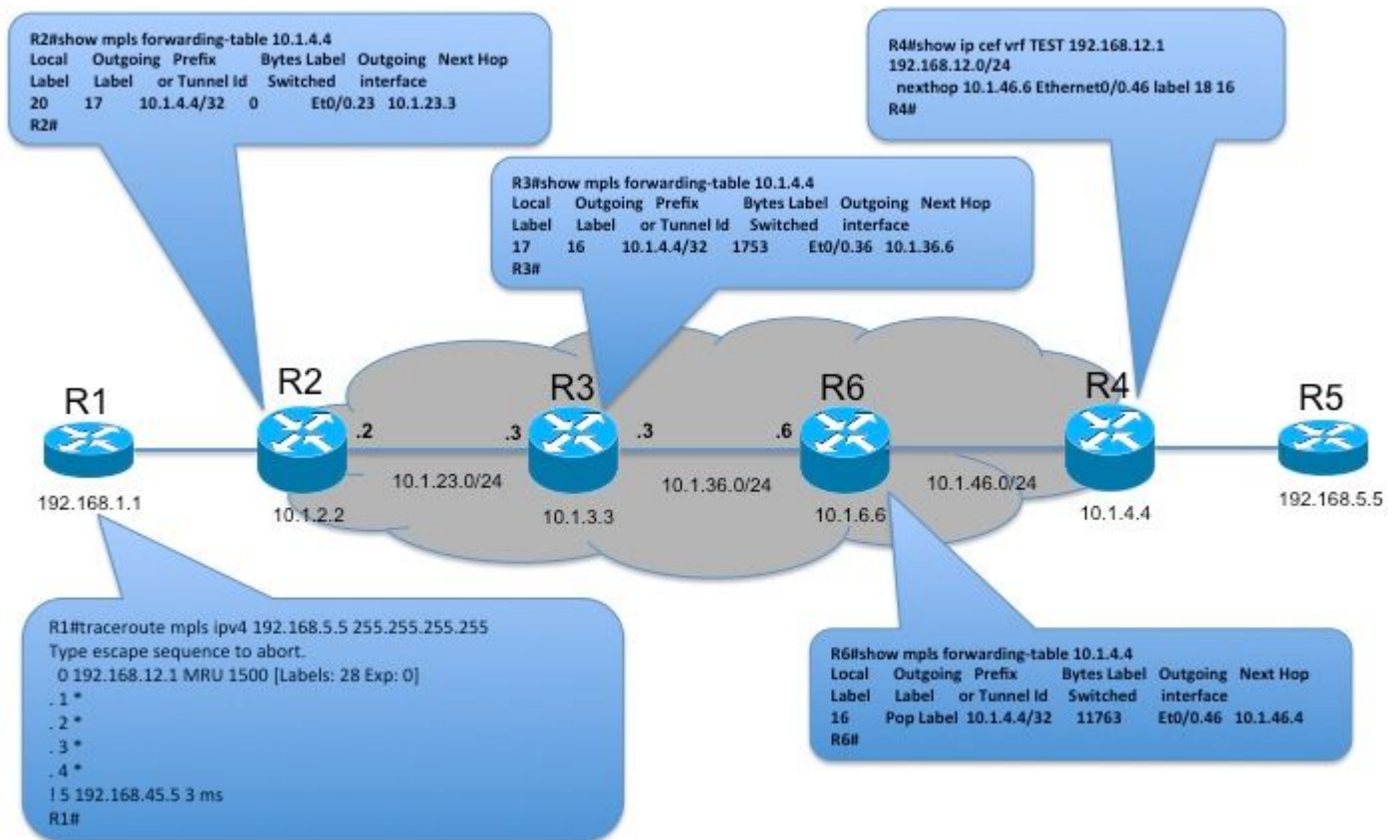


這適用於在PE-CE之間啟用MPLS的類似CSC的情況。通過運營商MPLS域執行從CE到遠端CE的LSP跟蹤存在2個挑戰，如下所示：

- LSP回應回覆將直接傳送到啟動器。因此，響應方必須能夠到達發起方。在以上拓撲中，R3可能無法與R1通訊，因為它在VRF中。
- 對於標籤堆疊中的每個標籤，目標FEC堆疊中應包含相關的FEC詳細資訊以進行驗證。發起方包含的FEC將為1，而PE將推送2個標籤。在上述拓撲中，R1傳送了FEC={192.168.5.5/32}的MPLS回應請求，並在堆疊中包括標籤28。由於R2使用{17, 27}交換標籤28，因此R3將接收堆疊中帶有2個標籤的請求，而TLV中的1個FEC會混淆FEC驗證。

RFC6424定義了「FEC堆疊更改TLV」的概念，以解決問題2。此TLV將包括在回覆中，並帶有相關FEC，作為發起方可以在後續回應請求中包括的PUSH/POP。

draft-ietf-mpls-lsp-ping-relay-reply定義了TLV中承載中繼節點地址堆疊的概念，響應方可以使用該堆疊來中繼響應，即使它無法到達發起方。



Cisco IOS®目前不支援這2個問題，因此從CE到遠端CE的LSP追蹤將只列出輸入PE和遠端CE。僅為了完整起見才包括此項。

相關資訊

- [RFC 3032](#)
- [RFC 4379](#)
- [RFC 6424](#)