

## 第二章 多重協定標籤交換式網路相關研究背景

### 2.1 MPLS[15] 概要

#### 2.1.1 MPLS 的由來

在傳統的 IP 網路[16]，封包從起始點到其目的地的傳送過程中，網路上每個路由器都會個別獨立地分析封包網路層標頭中所含的目的地 IP 位址，而根據靜態或動態路由協定（static or dynamic routing protocol）所建立的路由表（routing table），選擇適當的輸出介面，將封包傳送至下一個路由器或是目的地。這種路由方式稱為 hop-by-hop destination-based unicast routing。

此傳統 IP 網路有許多的限制及缺失：

1. 封包在傳送時是以最短路徑選路(shortest path routing)的問題：根據 IGP 演算法所做的路徑選擇，可能造成某些鏈接壅塞及資料的流失，但某些鏈接卻仍有許多可用頻寬閒置，因此導致不良的網路資源利用率和網路效能，傳輸品質也受到影響。
2. 當路由器收到 IP 封包時會先儲存 IP 封包，分析目的地 IP 位址，並索引路由表，選擇適當的輸出介面，將封包送出，當路由器收到下一個 IP 封包時，又要重覆的動作，因此造成傳送速度慢，且效率差。
3. 傳統 IP 網路傳送封包時，對每一個封包都相同看待，對於如語音、視訊要求低的延遲變化(delay variation)的封包，無法優先處理。

因此傳統 IP 網路的封包轉送的方法雖然很成功，而且已經很普及，但因為它的限制及缺失，所以需要新的技術來解決。此新的技術即是 MPLS。傳統的 IP 網路有很好的連接性(connectivity)，透過 IP 網路可以連至全世界各地；ATM 及 Frame Relay 具有很好的私密性(Privacy)和品質保證(QoS)。而新的技術 MPLS 則結合以上二者的優點。

MPLS(Multi-protocol label switching)技術之 multi-protocol 乃是指其可支援資料鏈接層（data link layer）如 ATM、Frame Relay、Ethernet 或 PPP 等，以及網路層如 IPV4、IPV6、IPX 等。而在 MPLS 網路裡，路由器會建立一個 LFIB(Label Forwarding Information Base)：包括 input I/F，input label 和 output I/F，output label。當路由器收到含夾層標籤之封包的標籤時，根據標籤內標籤值和進入的

介面索引 LFIB，然後快速地將整個封包往 output I/F 送出，且將標籤內的標籤值換成 output label 值，因此稱為 label switching。因為標籤是固定長度且整個遞送過程全部用硬體處理，所以此種交換技術可以很快將封包遞送（forwarding），因此可以解決傳統存轉式 IP 網路存轉速度慢的缺失。MPLS 應用在 ATM 或 Frame Relay 的網路中也是利用類似的方式來作交換。

### 2.1.2 標籤包裝 (Label Encapsulation)

## Label Encapsulation

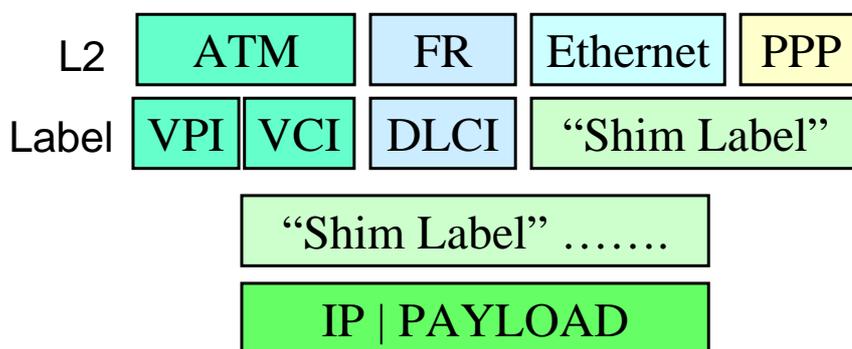


圖 1 MPLS 之標籤包裝

參考圖 1，在不同協定裡，MPLS 標籤包裝情形各有不同：

1. 在 ATM 網路裡，使用 VPI/VCI 當作標籤，而隨著細胞 (cell) 被傳送。
2. 在 Frame Relay 網路裡，使用 DLCI 當作標籤，而隨著碼框 (frame) 被傳送。
3. 當資料鏈接層 (data link layer) 為 Ethernet 或 PPP 時，則是在資料鏈接層和網路層之間加入一個夾層標籤 (Shim Label)，而隨著碼框或封包 (packet) 被傳送。

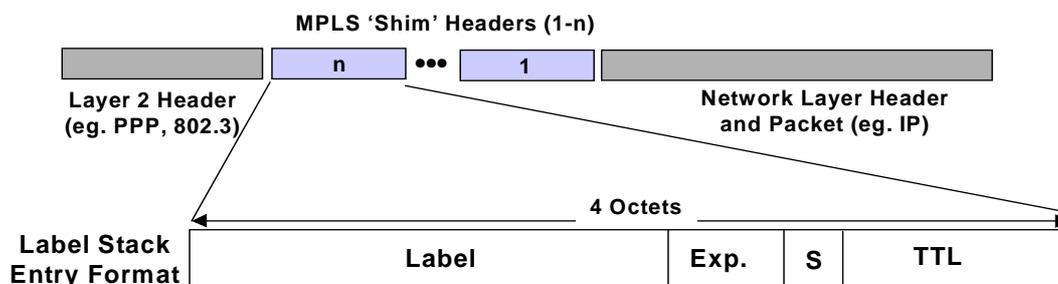


圖 2 夾層標籤之標籤格式

圖 2 為夾層標籤之標籤格式，我們可以看到它被插入在 Layer 2 和 Layer 3 之間，並且可插入二個標籤以上。一個標籤長度為 32 位元，其中 label 佔用 20 位元，放置標籤值。Exp. 佔用 3 個位元，當作實驗用途，在 DiffServ 中，可當作服務類別（Class of Service）使用。S 佔用 1 個位元，用來指出此標籤是否在所有 MPLS 標頭最底層，如果是，其值為 1；否則其值為 0。TTL 欄位與 IP 標頭的 TTL 相同，當 TTL 減至 0 時，則此封包將被丟棄。

### 2.1.3 MPLS 基本運作程序

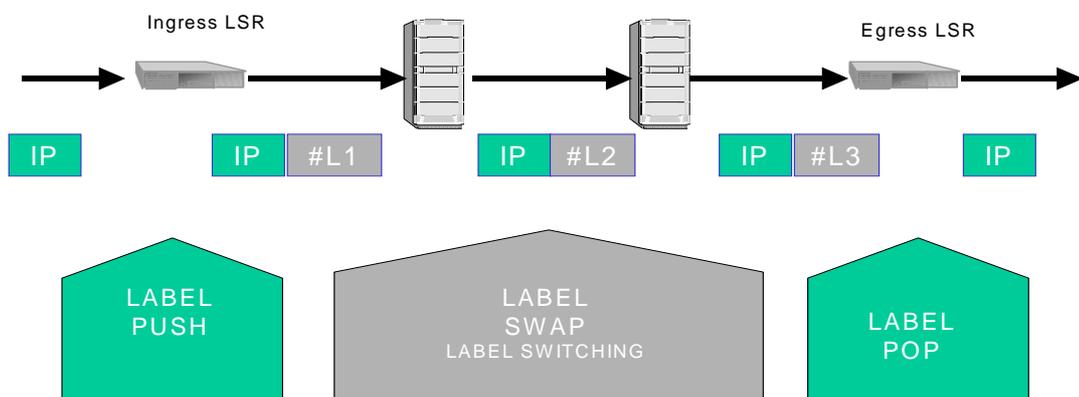


圖 3 MPLS 基本運作程序

我們對照圖 3 說明一些 MPLS 術語，以及其基本運作程序。

**LSR**：Label Switching Router（標籤交換路由器），是具有 MPLS 技術功能的路由器

**LER（或 ELSR）**：Label Edge Router（標籤邊緣路由器），LER 位於 MPLS 網路的邊緣，負責將未標籤的封包加上標籤或是將已標籤的封包去掉標籤。

**Ingress LSR**：當封包進入 MPLS 網域時之 LER，其負責將未標籤的封包加上標籤。

**Egress LSR**：當封包離開 MPLS 網域時之 LER，其負責將已標籤的封包去掉標籤。

**Label Push**：在 Ingress LSR 將未標籤的封包加上標籤的動作。

**Label Pop**：在 Egress LSR 將已標籤的封包去掉標籤的動作。

**Label Swap**：在 MPLS 網路內部之 LSR，當其收到一個封包時，根據此封包的標籤值到 LFIB 中做查詢，決定輸出標籤值為何，然後將新的標籤值取代原標籤值，此更換標籤值的動作稱為 **Label Swap**。

**FEC**：Forwarding Equivalence Class（同一運送等級），此指資料流的等級，此等級依據很多因素來分類，例如來源端位址、目的地位址、資料形態（語音、視訊、資料等）、IP Protocol 類別、服務的種類。同一運送等級的資料流在 Ingress LSR 和 Egress LSR 之間使用相同的 LSP。

**LDP**：Label Distribution Protocol（標籤分散協定），在 MPLS 網域作為標籤分散的協定，此協定使相鄰的路由器有一個共同的 FEC <-> label binding。目前標籤分散協定有三種：LDP、RSVP-TE 和 CR-LDP，其中 RSVP-TE 和 CR-LDP 為具有流量工程（Traffic Engineering）能力的標籤分散協定。

**LSP**：Label Switched Path（標籤交換路徑），指一資料流通過 MPLS 網域的某一用 label 建立起來的特定路徑。此路徑稱為 LSP。

## 2.2 MPLS 技術的基本運作



MPLS 技術將控制面（control plane）和傳遞面（forwarding plane）分開。在控制面方面使用二種通信協定：IP 尋徑協定（IP routing protocol，例如 OSPF、IS-IS），以及標籤分散協定（e.g.LDP、CR-LDP、RSVP-TE）。而傳遞面主要作標籤的交換。其基本運作為先以傳統第三層的 IP 尋徑協定建立起整個網路的基本路由資料，再根據這些路由資料以標籤分散協定建立 LSP，當封包進入此 MPLS 網域後，根據此封包所屬的 FEC 對應到它所屬的 LSP，而將此封包載送出去。

目前標籤分散協定有三種：LDP、CR-LDP 和 RSVP-TE，其中 CR-LDP 是 LDP 的延伸，允許網路管理者去建立一個符合需求的 explicitly routed LSP，而 RSVP-TE 是使用原 RSVP 的 PATH 和 RESV 功能延伸，使其可用於標籤分散，因此 CR-LSP 和 RSVP-TE 都具有流量工程能力的標籤分散協定。

我們以圖 4 至圖 7 進一步說明 MPLS 的運作，從圖 4 可知 MPLS 的運作為當網路剛開始時，所有 LSR 利用 IP 的尋徑協定如 OSPF、IS-IS 等去交換路由資訊，建立路由表以及傳遞信息資料庫（FIB，Forwarding Information Base）。

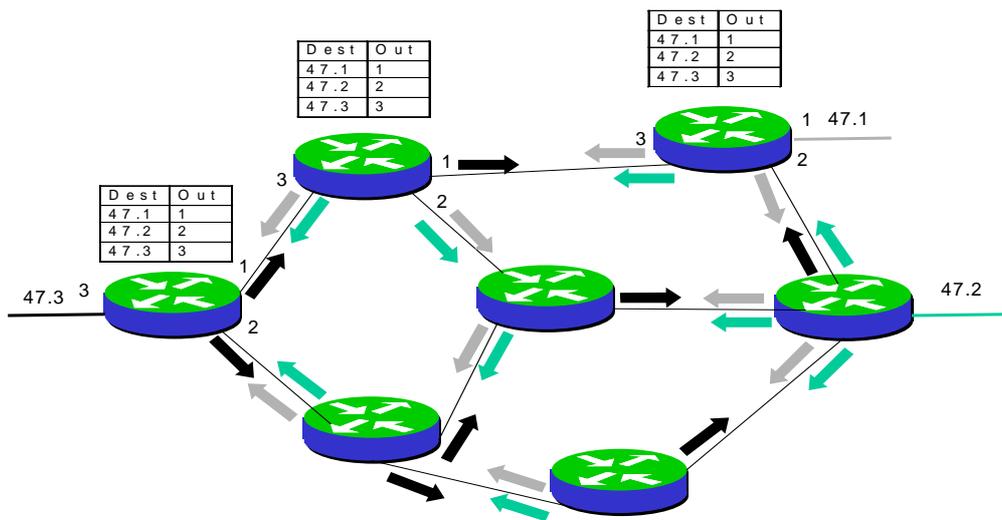


圖 4 傳統 IP 之路由資訊

然後如圖 5 根據 IP 尋徑協定建立的 FIB，啟動 LDP 散佈 label-binding 資訊，從 Ingress LSR 送一 label request message 至 Egress LSR，當 Egress LSR 收到 label request message 後，送出 label mapping message 至 Ingress LSR，替 Ingress LSR 和 Egress LSR 建立 LSP，LSP 建立後各 LSR 會產生一個標籤傳遞信息資料庫 (LFIB, Label forwarding Information Base)。

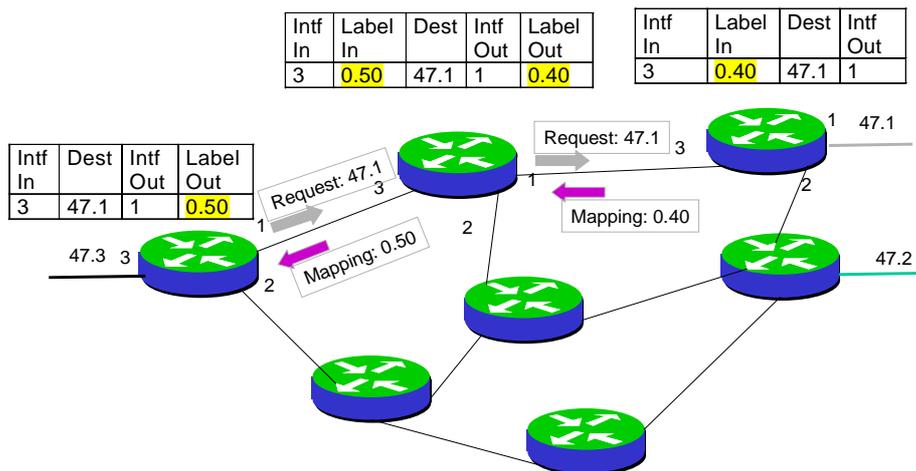


圖 5 MPLS 標籤交換路徑之建立程序

圖 6 為控制面所建立的 LSP，此 LSP 之建立以最短路徑為原則。當 Ingress

LSR 收到封包，根據此封包所屬的 FEC 索引 LFIB，加上標籤。在 core LSR 只負責 label swapping，而 Egress LSR 移除標籤，還原回原 IP 封包，並繼續將封包往目的端傳送。

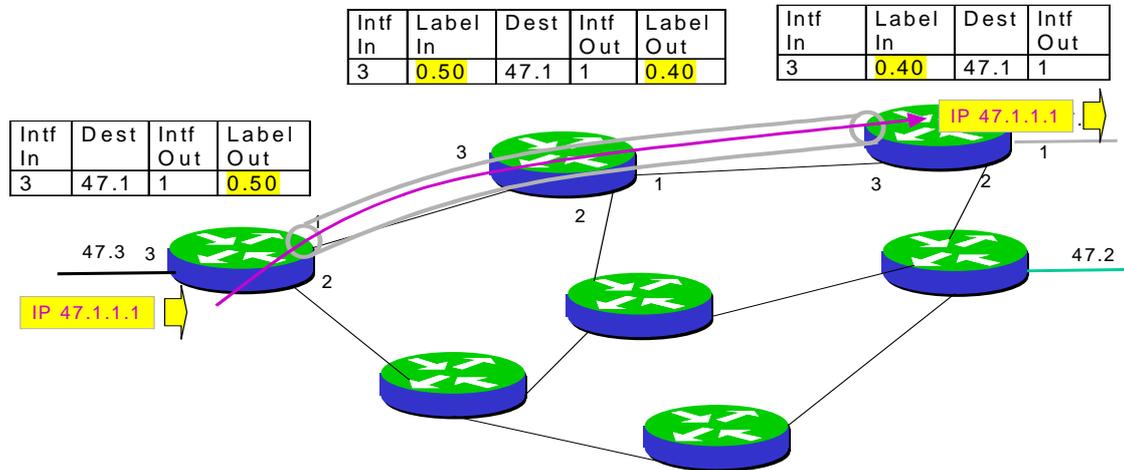


圖 6 MPLS 標籤交換路徑

圖 7 為用來說明 CR-LSP，CR-LSP 是藉由 CR-LDP 通訊協定在 MPLS 網路上所建立具有傳輸服務品質保證的 LSP。它的建立以滿足 QoS 需求為原則，未必為最短路徑。我們將於下一小節詳加介紹 CR-LSP。

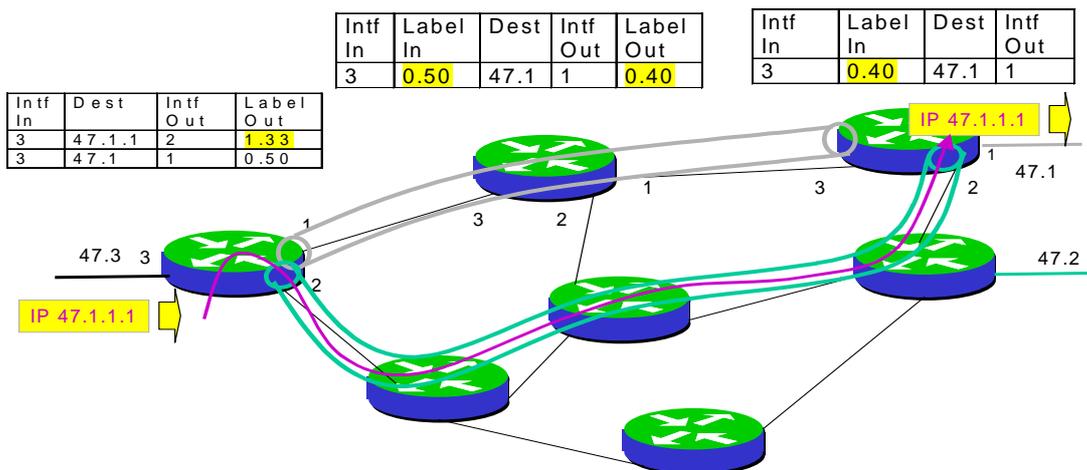


圖 7 Explicitly Routed LSP

## 2.3 標籤分散協定 (Label Distribution Protocol)

我們在 2.2 節中提到標籤分散協定有三種：LDP、CR-LDP 和 RSVP-TE，我們將在本節中詳加說明。

### 2.3.1 LDP 運作流程

LDP 有四種信息分別是發現信息(Discovery messages)、會議信息(Session messages)、散播信息(Advertisement messages)、通知信息(Notification messages)。我們分別說明如下：

- 發現信息：主要用於宣布和維護 LSR 的存在，週期經由 UDP 送出 Hello 信息給其鄰近所有路由器。
- 會議信息：LDP peers 間建立、維護或終止會議，這些操作必須經由 TCP 連線傳送初始信息。
- 散播信息：用來建立、改變或刪除標籤和 FEC 的對映。
- 通知信息：傳送有關網路狀態、診斷或錯誤信息

圖 8 為 LDP 運作流程，LSR 送出 Hello 封包去尋找周圍有無支援 LDP 協定的路由器，如果收到從其它路由器送來的 Hello 封包，這時候表示對方也支援 LDP 協定，所以我們就可以進一步建立 TCP 連線當作 LDP session，往後的 LDP 封包就利用此 TCP 連線來做傳送，如路經的建立或釋放等。當 LSR 之間要建立 LSP 時，就會傳送 label request 信息和 label mapping 信息，以便完成 LSP 的建立。

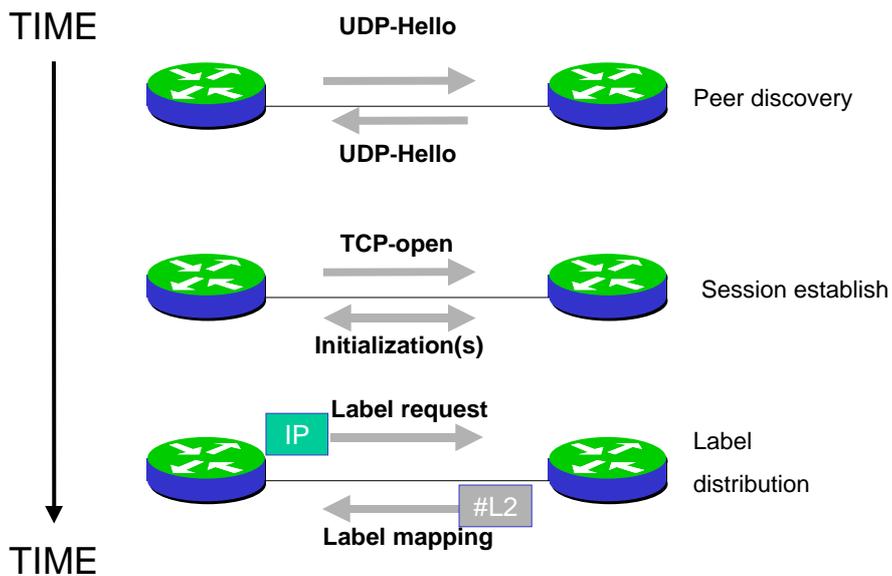


圖 8 LDP 運作程序

### 2.3.2 CR-LDP 運作流程

CR-LDP 是 LDP 的延伸，LDP 是 best-effort 的服務，它只跑一般的尋徑協定，而 CR-LDP 具有處理流量工程和 QoS 的能力，因此它除了跑一般的尋徑協定之外，亦加入一些限制(constraints)，使建立的 LSP 能符合需求。利用 CR-LDP 所建立的 LSP 稱為 CR-LSP，CR-LSP 建立的選徑是在 Ingress LSR 處理的，也就是 Ingress LSR 根據資料流需求和網路流量狀況，替資料流選擇一條最適合的路徑。因此 CR-LSP：

- 1.未必一定找尋最短路徑。
- 2.符合需求及流量工程的路徑
- 3.在 connection-less 的網路提供 connection-oriented 的服務
- 4.做 constraint-based 的路徑選擇。

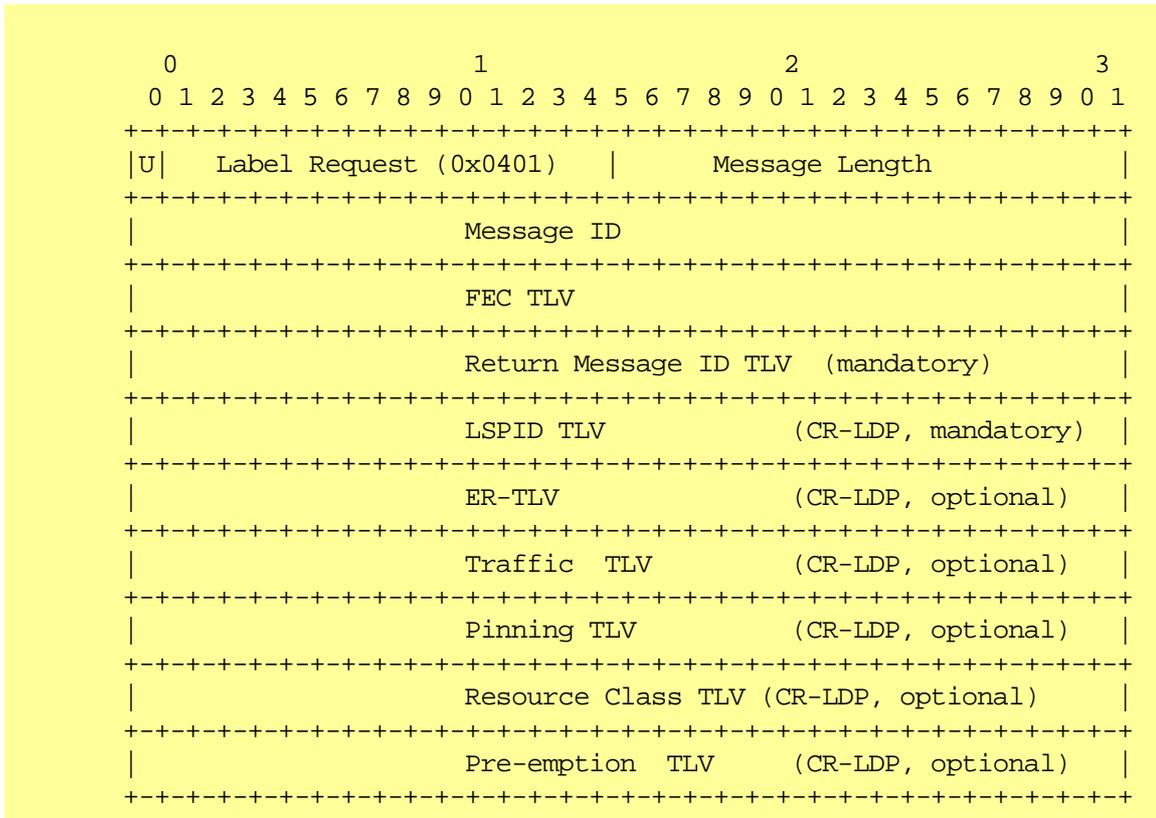


圖 9 CR-LDP Label Request Message 的格式

圖 9 為 CR-LDP Label Request Message 的格式，它比 LDP 多出一些做為流量工程的欄位，我們針對這些欄位加以說明如下：

- LSPID TLV：用來代表 LSP 的編號，用以區分不同的 LSP。
  - ER-TLV：用來指明正要被建的此 LSP 所欲通過的是那些節點與其順序。
- CR-LDP 建立的路徑為顯路徑(Explicit Route)，Explicit Route 可分為 Strict Route 和 Loose Route。
- ✧ Strict Route：下一個節點必須直接連到上一個節點。
  - ✧ Loose Route：中間經過的路徑根據各節點的路由表來決定。
  - ✧ Strict Route 和 Loose Route 可以混合使用，圖 10 為此狀況的舉例，從 Ingress LSR 到節點 C 為 strict，而節點 D 為 loose，所以節點 C 和節點 D 之間所經路徑由路由表來決定。節點 F 為 strict，所以節點 D 和節點 F 之間直接相連。

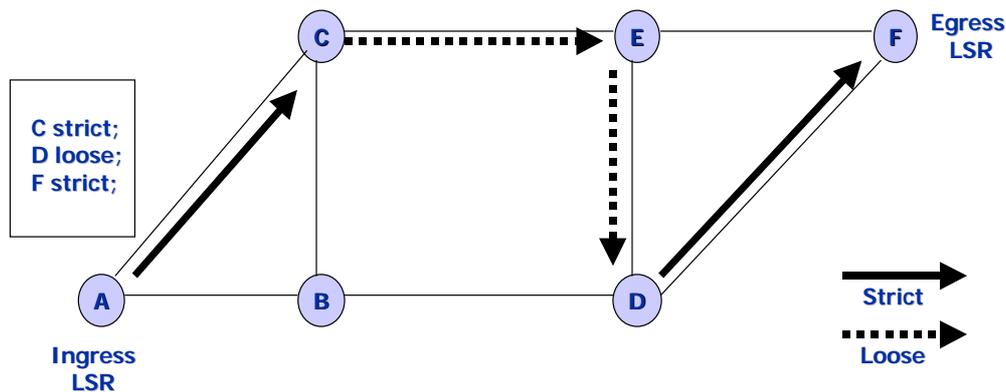


圖 10 Explicit Route：Strict/Loose Route

- Traffic TLV：用來描述要建的 CR-LSP 的流量需求，包括 peak data rate，peak burst size，committed data rate，committed burst size，excess burst size。
- Route Pinning TLV：在 Loose Route 時，因為 request message 走的路徑是根據路由表來決定，但有時候因為路由表改變的關係，導致 request message 和 mapping message 會走不同路徑，這時候可利用 Route Pinning TLV 讓 mapping message 依照 request message 所走的路徑傳送到來源端。
- Resource Class TLV：網路管理者可以將網路資源作各種分類，如用“colors”或“administrative groups”，Resource Class TLV 用來描述要建的 CR-LSP 欲使用或避免的資源屬性。
- Preemption TLV：在建 CR-LSP 時的 Label Request Message 中之 Preemption TLV 裡面有二個參數：建立優先權（Setup-Priority）及保持優先權（Holding-Priority）。建立優先權及保持優先權值的範圍從 0 到 7，0 的優先權最高，7 的優先權最低。當正要建立之 CR-LSP，發生網路頻寬不足時，若此資料流的建立優先權高於現已存在之 CR-LSP 之保持優先權，則可以侵佔其保留的頻寬，此侵佔行為稱為頻寬資源的侵佔（Resource Preemption）。

### 2.3.3 RSVP –TE

RSVP-TE 是將 RSVP（Resource-Reservation Protocol）加以延伸，使其可以成爲一種標籤分散協定並具有流量工程的能力，其功能與 CR-LSP 相近。它將 5 個物件（Objects）加至 RSVP Path 和 Resv message 內。

RSVP-TE 之 Path Message 是由 Ingress LSR 送出，它是在原 Path Message 內增加了以下物件：

1. Label\_Request Object: 要求中間的 LSR 和出口 LSR 提供 label binding。
2. Explicit\_Route Object：用來指定此 LSP 要經過的節點。
3. Session\_Attribute Object：有關於設定優先權、保持優先權、本地重新路由資訊。
4. Record\_Route Object：記載路徑上往下游各節點的 IP 位址、路徑上的錯誤信息、迴路偵測等。

RSVP-TE 之 Resv Message 由 Egress LSR 送出，它是在原 Resv Message 內增加了以下物件：

1. Label Object：執行 downstream on demand 標籤分散程序，提供標籤給上游 LSR。
2. Record\_Route Object：記載路徑上往上游各節點的 IP 位址、路徑上的錯誤信息、迴路偵測等。

## 2.4 MPLS 技術的用途



1. 快速的遞送。
2. 流量工程 (Traffic Engineering)。
3. 虛擬私人網路 (Virtual Private Networks): 使用二個標籤，外層的標籤用來通過 MPLS 網路作為標籤交換使用，內層的標籤用來辨識 VPN 客戶。
4. 支援差異式服務 (DiffSev-capable MPLS network): 可利用不同的標籤值來分辨此封包所屬的服務類別，而用標籤格式中的 EXP 來記錄封包被丟棄的優先順序；或只利用 EXP 來分辨此封包所屬的服務類別。