

行政院所屬各機關因公出國人員出國報告書  
(出國類別：實習)

## 赴美國實習「網際網路交換設備技術」

### 報 告 書

服務機關：中華電信股份有限公司  
國際電信分公司

出國人職稱：助理工程師 專員  
姓名：張士瑞 李家康

出國地點：美國  
出國期間：92.9.7.~92.9.20  
報告日期：92.12.16

行政院研考會／省(市)研考會 編號欄

H6/C09203885

公務出國報告提要

頁數: 14 含附件: 否

報告名稱:

實習網際網路交換設備技術

主辦機關:

中華電信國際電信分公司

聯絡人/電話:

/23445280

出國人員:

張士瑞 中華電信國際電信分公司 網路處 助理工程師  
李家康 中華電信國際電信分公司 網路處 專員

出國類別: 實習

出國地區: 美國

出國期間: 民國 92 年 09 月 07 日 - 民國 92 年 09 月 21 日

報告日期: 民國 92 年 12 月 18 日

分類號/目: H6/電信 /

關鍵詞: 網際網路交換設備技術

內容摘要: 職等二人奉派赴美實習「網際網路交換設備技術」內容為參加網路設備製造商Juniper公司在維吉尼亞州Herndon訓練中心所開「Introduction to Juniper Networks Routers」、「Advanced Juniper Networks Routing」各五日的課程。前後兩週的課程以一貫性與系統性的做網際網路交換技術介紹。本報告就課程中研討MPLS有關訊務工程(Traffic Engineering)及Juniper所採用的RSVP(Resource reSerVation Protocol)方式做進一步討論,選擇此一課題乃基於職等所負責維運的網際網路服務網路TWGate已是MPLS enabled的IP網路,而RSVP為提供訊務工程的重要技術之一,經實作在TWGate骨幹網路達到訊務流量管控目的,得以進一步提供客戶更優質的網路服務。本報告分為五個單元:一、前言二、MPLS概述三、訊務工程 Traffic Engineering 四、資源預訂協議RSVP 五、結論與心得

本文電子檔已上傳至出國報告資訊網

# 摘要

職等二人奉派赴美實習「網際網路交換設備技術」內容為參加網路設備製造商 Juniper 公司在維吉尼亞州 Herndon 訓練中心所開「Introduction to Juniper Networks Routers」、「Advanced Juniper Networks Routing」各五日的課程。前後兩週的課程以一貫性與系統性的做網際網路交換技術介紹。本報告就課程中研討 MPLS 有關訊務工程(Traffic Engineering)及 Juniper 所採用的 RSVP (Resource reSerVation Protocol)方式做進一步討論，選擇此一課題乃基於職等所負責維運的網際網路服務網路 TWGate 已是 MPLS enabled 的 IP 網路，而 RSVP 為提供訊務工程的重要技術之一，經實作在 TWGate 骨幹網路達到訊務流量管控目的，得以進一步提供客戶更優質的網路服務。

本報告分為五個單元：

一、 前言

二、 MPLS 概述

三、 訊務工程 Traffic Engineering

四、 資源預訂協議 RSVP

五、 結論與心得

# 目 錄

一、 前言.....	1
二、 MPLS 概述.....	2
三、 訊務工程 Traffic Engineerig .....	4
四、 資源預訂協議 RSVP .....	11
五、 結論與心得.....	13

# 一、前言

Juniper 路由器以其獨家訂製晶片組 ASIC 及特有硬體架構在高階路由器產品市場以其優異的性能表現而取得領先的地位，在作業軟體 Juniper Operation System 方面亦以單一但完整性高又功能齊備得到使用者給予極高評價。由於本公司陸續增購是類設備以建置更堅實穩定的 IP 服務網路，因而奉派參加原廠訓練中心所開課程，期能習得有關技術。

實習日程如下：

(一)、九十二年九月七日

去程：台北 - 美國

(二)、九十二年九月八日至九十二年九月十二日

「Introduction to Juniper Networks Routers」訓練課程

(三)、九十二年九月十三日至九十二年九月十四日

例假日

(四)、九十二年九月十五日至九十二年九月十九日

「Advanced Juniper Networks Routing」訓練課程

(五)、九十二年九月十九日至九十二年九月二十日

回程：美國 - 台北

## 二、MPLS 概述

MPLS Multi-Protocol Label Switching--多重協定標籤交換。MPLS 屬於第三代網路架構，是新一代的 IP 高速骨幹網路交換標準，由 IETF (Internet Engineering Task Force, 網際網路工程專案小組) 所提出，由 Cisco、ASCEND、3Com 等網路設備大廠所主導。

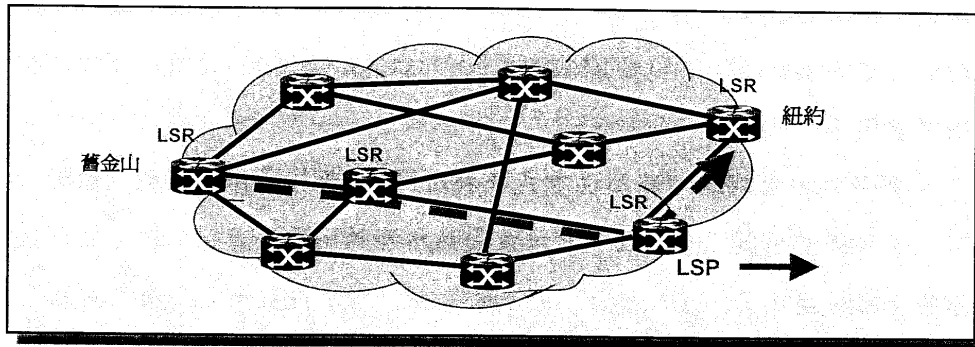
MPLS 最初是取代傳統 IP 封包傳送--以複雜的位址配置演算法，根據標籤，用快速、簡單、非同步傳輸模式(ATM)，改善了 IP 網路的效能和可擴充性，這項功能具備了基本的訊務工程能力，以及為 IP 網路帶來企業虛擬網路 (VPN) 服務。

MPLS 是集成式的 IP Over ATM 技術，在 Frame Relay 及 ATM Switch 上結合路由功能，封包透過虛擬電路來傳送，只須在 OSI 第二層 (資料鏈結層) 執行硬體式交換 (取代第三層 [網路層] 軟體式選徑)，整合了 IP 選徑作業與第二層標籤交換作業為單一的系統，因此可以解決 Internet 路由的問題，使網路封包傳送的延遲時間減短，更適合多媒體訊息的傳送，增加網路傳輸的速度。因此，MPLS 最大技術特色為可以指定封包傳送的先後順序，使用標籤交換式 (Label Switching)，網路路由器只需要判別標籤，進行轉送處理。

MPLS 的運作原理是提供每個 IP 封包一個標籤，由此決定封包的路徑以及優先順序，與 MPLS 相容的路由器 (Router)，會將封包轉送到其路徑前，僅讀取封包標籤，無須讀取每個封包的 IP 位址以及標頭 (因此網路速度便會加快)，將所傳送的封包置於 Frame Relay 或 ATM 的虛擬電路上，迅速將封包傳送至終點的路由器，進而減少封包的延遲，同時藉由 Frame Relay 及 ATM 交換器所提供的 QoS (Quality of Service) 對所傳送的封包加以分級，因而大幅

提升網路服務品質提供更多樣化的服務。

下列網路圖介紹 MPLS 相關名詞及意含：

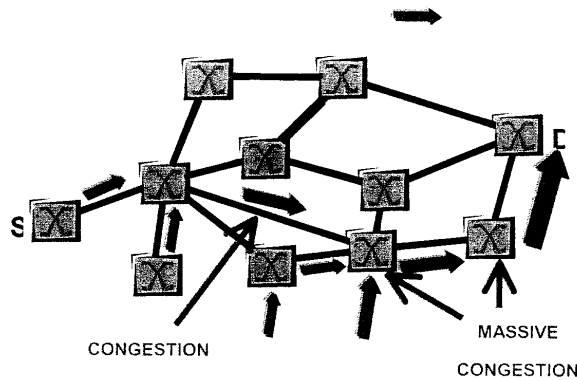


- Label Switch Routers (LSRs): 以封包標籤(label)來決定如何傳送封包的路由器或交換器。
- Label Edge Routers (LERs): 位於 MPLS 網路與一般 IP 網路邊界上的路由器。
- 入口 LERs: 負責將進入 MPLS 網路的 IP 封包分類與貼上適當的標籤。
- 出口 LERs: 負責移除標籤並將這些封包傳送到正確的 IP 網路下一節點。
- Forwarding Equivalence Class (FEC): 當所有的封包在 MPLS 網路內部遵循相同路徑並每一節點都有相同的處理方式時稱為 FEC。
- MPLS Routing Protocol
  - 1、既有 IGP  
傳送 Topology 網路架構的訊息, 例如 OSPF, IS-IS, BGP-4
  - 2、Routing based constraint: 常作為 TE 流量調整之方法  
例如線路容量, 使用率, 優先權, 線路恢復後使用移轉。  
如: OSPF-TE、IS-IS-TE

### 三、訊務工程 Traffic Engineering

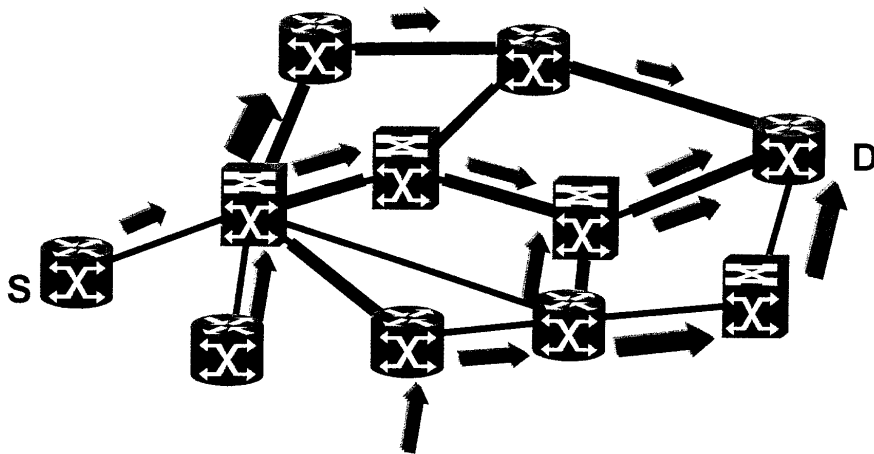
Traffic Engineering(TE) 訊務工程是以網路管理的觀點來看資料封包在目前甚至預期其在網路中繞送所可獲得的可用頻寬的處理機制，同時也是在網路服務提供 COS(Class of Service)及 QOS(Quality of Service)的重要因素。

傳統的控制網路訊務流的方式是利用 IGP 運算選擇壅塞較少的鏈路。但 IGP 仍有其限制，因 IGP 依據最短路徑的演算法選徑，但會造成嚴重的壅塞問題，SPF 演算法使用簡單的路徑值並未使用頻寬使用度與訊務特性作為選徑考量。



TE 可以是人工手動方式，經由監視網路狀況伺機增減訊務所需頻寬，當然透過路由協定或其他的方式經由回饋自動調整頻寬是維護大型網路所極為需要的。TE 能幫助網路提供者最有效的頻寬利用及透過 L2 分攤訊務負載甚至保留部分鏈路提供給服務等級較高的客戶使用。MPLS 網路訊務工程是從資源管理與目前及未來期待之流量角度來處理調整訊務流量，正因為 MPLS 網路能提供更有效的 TE 機制，因此 TWGate 骨幹網路採用 MPLS 網路架構來提供更高等級的網路服務。



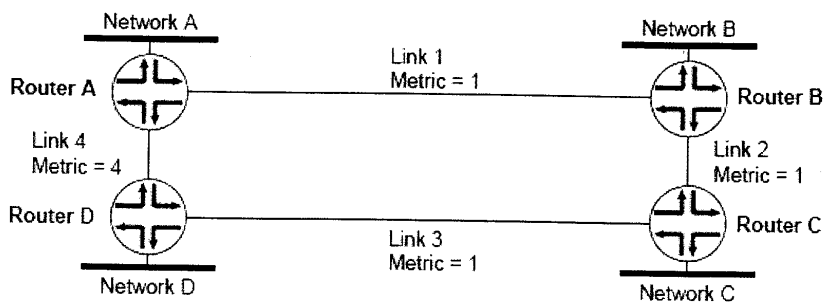


Traffic distributed over network resources by (MPLS) traffic engineering

以下就 MPLS 網路在 TE 技術上的演變做簡短介紹：

1、過往的 Traffic Engineering 方法：

在路由器平台的 Traffic Engineering 是以簡單的路由路徑值為運算標準。這種方式在較小的網路架構以鏈路數、路由器數目、流量的角度來看是適用的。下圖解釋路由路徑值計算的原理，假設網段 A 送出大量的流量往網段 C、D。Link 1 與 Link2 會很壅塞，因為網段 A 到網段 D 與網段 A 到網段 C 的流量都經過 Link 1 與 Link2。如果 link 4 的 metric 改為 2，網段 A 到網段 D 的流量將改道 link 4，網段 A 到網段 C 的流量還是經過 Link 1 與 Link2。



傳統路由骨幹的 Traffic Engineering 限制

- (1). software based 路由器在大流量的狀況之下將造成傳輸的瓶頸。
- (2). 以路徑值 metric 做 Traffic Engineering 運算基礎的是不具擴充與延展性的。因為當 ISP 的網路越複雜，在調整 metric 時，很難保證一些調整過程的失誤與副作用 side effect 不會發生。Traffic Engineering 在 metric based 的方法下，面對複雜問題時是採用嘗試錯誤的方法，而非以科學方式解決。
- (3). IGP 的路由計算是拓樸導向與簡單的路徑值運算。但 IGP 並不會傳遞頻寬使用度與流量特性的訊息。所以流量的負載並不會在 IGP 的運算之內，如此流量不會均勻的分佈於網路的鏈路裡，而使得珍貴的網路資源無法有效利用。

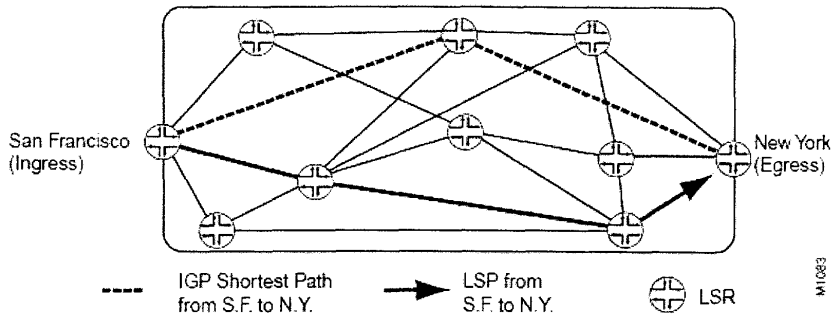
2、現今的 Traffic engineering 方式：

本文僅就 Juniper 所提解決方案：

- Packet forwarding
- Information distribution
- Path selection
- Signaling

A、封包轉送元件(Packet - forwarding component)

Juniper 在 Traffic engineering 的 Packet forwarding 架構在 MPLS 方式。MPLS 主要負責 IP 封包傳送在一條是先定義好的路徑之上，稱之為標籤轉送路徑(L S P) label switch path，其類似 ATM PVCs 的特性，一條 L S P 是以單方向方式傳送流量，所以雙向流量必須使用二個 L S P。L S P 是以多個標籤轉送的交換器構成，允許封包從一個標籤轉送的交換器到另一個標籤轉送的交換器。



當一個入口 LSR 接收到封包，它會在封包的前面加上標籤然後送到下一個 LSR，並且沿著 LSP 送到出口端的 LSR。出口端的 LSR 會將標籤移除然後根據 IP 標頭內的目的地地址，送到目的地。下圖是 MPLS forwarding table 每一個 entry 都包含了 input/output interface 與 in/out label 對。當封包抵達一個 LSR 時，LSR 會根據 Forwarding table 的資料對送到出口的介面上。

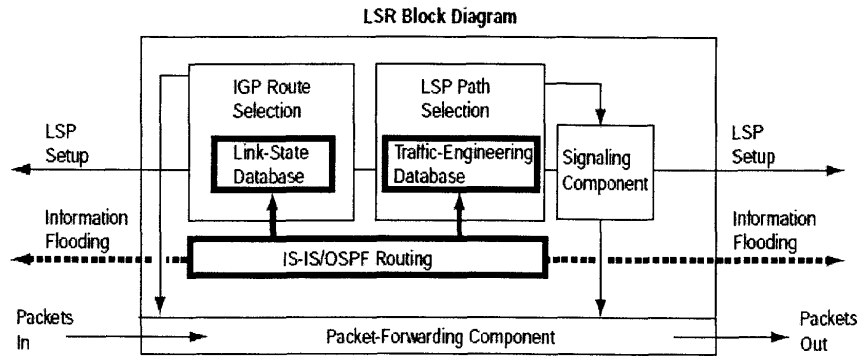
In Interface	In Label	Out Interface	Out Label
⋮	⋮	⋮	⋮
3	21	4	18
3	56	6	135
⋮	⋮	⋮	⋮

## B、資訊散佈元件(Information Distribution component)

Traffic engineering 需要整個網路架構的詳細的資料以及動態的網路負載訊息，所以需要完整而動態的資訊散佈。資訊散佈元件可由簡單的 IGP extensions 來完成，如 ISIS extensions 就可用 TLV 欄位完成，OSPF extensions 使用 LSA 完成。每一個 LSR 都會維護一個 TED (Traffic engineering database) 紀錄網路的屬性與架構，並且用來計算出一條特定的標籤轉送路徑(LSP)，TED 與 IGP 的資料庫是分開的，使得 Traffic engineering 的計算與 IGP 最短路徑運算可以分開。

因此 I G P 運算將不會受到 Traffic engineering 的影響。

Figure 10: Information Distribution Component



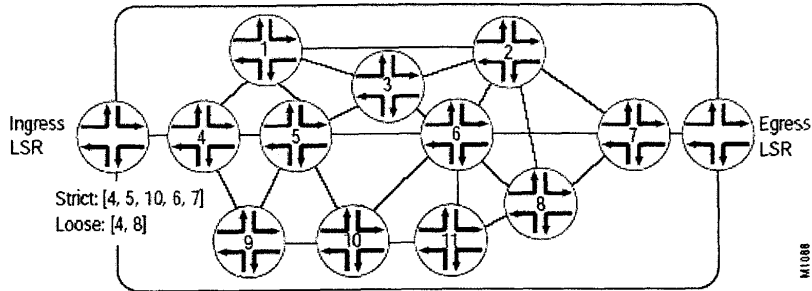
Traffic Engineering extensions 被加入 OSPF 的 L S A 有：

- 鏈路最大頻寬
- 最大可保留頻寬
- 目前保留頻寬
- 頻寬使用率
- 鏈路的分類

C、路徑選擇元件(Path selection component)：

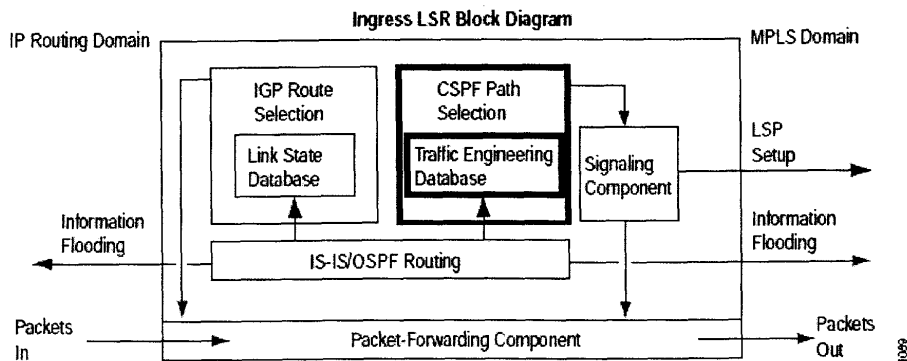
當網路 LINK 屬性與架構資訊藉由 IGP 散佈，並且放入 TED，每一個入口端的 L S R 運用 TED 計算在整個路由網域中自己的 L S P 路徑組合。參考下圖，入口端的 L S R 的 L S P 路徑組合是一個預定路徑與一個鬆寬路徑。一個預定路徑是先定義好順序的 L S R 組合，如果只有部分的

L S R 被決定的話稱為鬆寬路徑。



入口端的 L S R 將限制的最短路徑優先演算法 (Constraint Shortest Path First) 應用於 T E D 內。CSPF 已被修改以適合一些特定限制，包含了下列屬性：

1. 從 IGP 學習到的拓樸與鏈路狀態訊息 (存於 T E D 內)
2. 由 IGP extensions 載入的網路資源狀態，例如：link 總頻寬、使用率、可用頻寬、保留頻寬，並保存於 T E D 內。
3. 管理用的屬性需要在預定的標籤轉送路徑 (L S P) 上被傳遞 (如：頻寬需求、最大轉送值)，其值由使用者自訂。



當 CSPF 為了新的 L S P 會考慮每一個候選節點與鏈路，會根據資源的可使用度或者是否被選擇者是否超過使用者定義的一些政策。CSPF 運算的輸出屬於 explicit route，包含了一序列的 L S R IP 位址，提供最

短路徑以符合限制。接著 explicit route 被轉送到 signaling 元件，並使 LSP 上所有的 LSR 處於 forwarding 狀態。

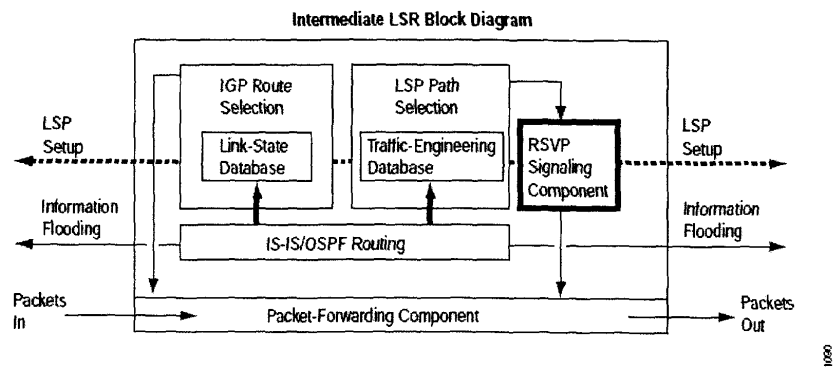
不管 on line path calculation 所帶來降低管理工作的好處，一個 offline 的規劃與分析工具仍然是必須的。” on line path calculation” 必須考慮網路資源限制與替代路由。

一個 offline 的規劃與分析工具同時檢查每一鏈路的資源限制與 LSP (入口到出口端的 LSR) 的需求。Offline 可能花上數個小時才會完成整個網路的運算，並比較每次計算的結果，最後選出最佳的路徑。

#### D、信號元件(signaling component):

因為網路狀態訊息存在入口端的 LSR 內的 TED 是較舊的資料，CSPF 計算出的路徑將被認為可接受的 LSP，但此路徑卻可能是還未運作的，必須等到信號元件(signaling component)將其建立起來。

以下是 Signaling component 方塊圖：

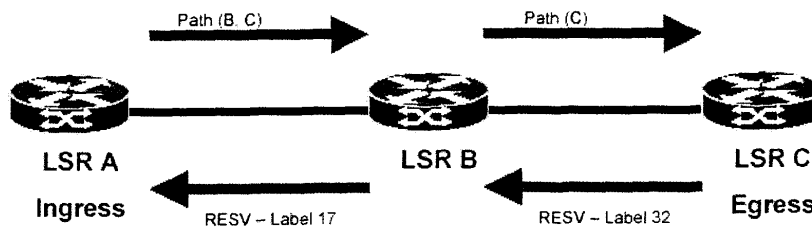


有關 Juniper 所採用的信號元件主要為 RSVP 將在下一單元做進一步討論。

## 四、資源預訂協議 RSVP

RSVP為IETF所規範的網路資源預訂協議，基本它使用信息交換(message exchange)在IP網路節點間預訂訊務流所需的網路資源，延伸應用到MPLS L S P的建立時必須加強其所帶信息才能傳佈MPLS標籤。

RSVP不是一般的IP封包，它使用IP datagrams在L S R的互連節點間通訊而不需要維持TCP的連線，但是因此它需要有遺失「控制信息」的處理機制。下圖顯示使用RSVP建立L S P的流程：



入口L S R A 因需要而決定建立一條新的L S P到L S R C，依據網路管理者考量這條路徑須是經由L S R B 再到L S R C，而不是由所經各節點根據其路由表決定到達L S R C 所要經過的下一站。L S R A 將此路徑所需的指定轉送點(B, C)及訊務所需的一些相關參數以 IP datagram 方式傳送給L S R B，稱為「建徑信息」(Path message)。

這就是之前所指的信號元件使用了RSVP的擴充格式，來建立L S P與標籤傳遞，包含有：

- Explicit Route Object 允許RSVP Path 訊息傳送於整個特定的路徑，其與傳統最短路徑路由方式互為獨立。特定路徑路由可以是指定或鬆寬兩種。
- Label Request Object 允許RSVP Path 訊息要求中間的L S R s在L S P建立時將標籤的連結完成。
- Label Object 允許RSVP支援標籤的傳送從下游節點送到上游節點，並不需要改變現有運作架構。

當 L S R B 收到這項建立路徑的要求後，發現自己不是這條 L S P 的出口，因此經過更改參數後轉送給 L S R C。L S R C 收到後發現自己是這條 L S P 的出口後再由「建徑信息」中取得此路徑所須保留的資源並在此節點予以保留後，再選配一個標籤給此新建路徑，將以上相關參數反向發出「保留信息」(Resv message)給其上游節點—即 L S R B。

L S R B 收到「保留信息」後根據 L S P ID, 找出相同的「建徑信息」中的資訊，同樣經選配標籤及保留所需資源後再發出「保留信息」給 L S R A。L S R A 也做同樣的處理，但因為它是此新 L S P 的入口 L S R，因此不需選配標籤給它的上游節點。

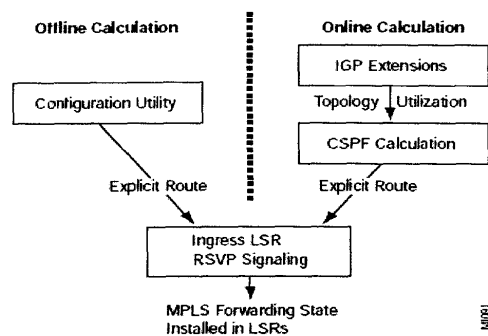
基於下列因素，RSVP 是建立 L S P 很理想的 signaling protocol：

- RSVP 是網路標準的資源保留協定。
- RSVP 可在 MPLS 的環境中建立 L S P。
- RSVP 允許在建立 L S P 時指定與保留網路資源。
- RSVP 允許以指定的方式建立提供如 A T M、Frame Relay 的負載平衡與訊務工程之 L S P。
- MPLS domain 中，端點對端點的 RSVP signaling 是易於擴充的，因為 L S P 的數目是關連於端點 L S R 的數目而非資料流的數目。



## 五、結論與心得

綜言之，Juniper 在 MPLS 網路實作上對於 L S P 的計算與設定的原則是：決定 L S P 的實體路徑是實現訊務控制的核心，可以用離線方式的設定工具或線上的限制型路由方式 (constraint-based routing)。至於傳送機制使用 R S V P 的 signaling 能力與實體路徑計算方式獨立。



Juniper 在網路訊務工程提供一些不同的方式設定 L S P

- ISP 可以計算 L S P 全段路徑在離線模式下並且個別設定每個 L S R 靜態的傳送模式，類似 IP over ATM 的方法。
- ISP 可在入口端的 L S R 計算全段路徑離線模式和靜態設定全段路徑，然後入口端的 L S R 用動態訊息協定如 RSVP 來設定 L S P 上的每一個 L S R 內之傳送狀態。
- ISPs 可以靠 constraint-based routing 來實行動態線上 L S P 的運算。在限制模式下，網路管理者為 L S P 設定一些限制，然後各路由器會依據這些限制尋找符合的 L S R。Juniper 的策略是讓入口的 L S R 根據特定的限制，計算整個 L S P，然後在網域內啟動 signaling。
- I S P 可以用離線方式計算部分路徑，其在入口的 L S R 內先設定部份特定路由器，然後啟動線上計算方式來決定全段路徑。

I S P可在入口的L S R上設定無限制的動態線上運算,然後使用 IGP 的最短路徑原則算出最佳L S P ,這是最簡單的模式。

此次奉派赴美參加 Juniper 原廠訓練中心所開的訓練課程是極難得的機會,因為 Juniper 的訓練地點及課程不多,在數量上相較其競爭者—Cisco 來說可喻為天壤之別。此次訓練時間雖不長,但上課環境設施完善且授課老師皆為 Juniper 網路技術的重要核心人物,配合即時實機操作使得上課效果極佳。職二人非常珍惜上課與討論的機會,並將維運網路所遇到的一些問題請教老師並得到解答。因此前後十天的訓練收獲頗豐,也同時開啟了我們對網際網路交換技術更深的探討空間,願將此篇心得報告做為一個好的開始。