行政院所屬各機關因公出國報告書
(出國類別：實習)

「實習企業 VPN 新技術及其應用」報告

服務機關： 中華電信股份有限公司
數據通信分公司
出 國 人： 職稱 姓名
助理工程師 馮仁傑
出國地點：美國
出國期間：93 年 10 月 3 日至 93 年 10 月 16 日
報告日期：93 年 12 月 14 日

目　　錄

第一章 前 言 4

第二章 研習行程及課程 5

第三章 影響 VPN 服務品質(QoS)的因素.....6

3.1 頻寬不足: 有限頻寬多種服務同時佔用 6

3.2 傳送延遲: 封包經各種網路設備處理與轉送結果 6

3.3 延遲變化: 多個資料流相互影響所產生的變化 7

3.4 封包丟棄: 網路壅塞的結果 7

第四章 服務品質(QoS)的網路應用 11

4.1 服務品質的類型.....11

4.2 各種服務品質(QoS)的機制.....12

4.3 整合型服務(Integrated Service；IntServ).....14

4.4 差異型服務(Differentiated Service；DiffServ) 19

4.5 流量調整器.....22

4.6 緩衝管理機制.....24

4.7 整合型服務(IntServ)與差異型服務(DiffServ)之比較 28

第五章 多重協定標籤交換技術之服務品質技術 ...31

5.1 概說.....31
5.2 多重協定標籤交換技術.....32
5.2.1 MPLS 基本運作.....35
5.2.2 標籤指定(Assignment)與傳遞(Distribution).....37
5.2.3 Packet 在 MPLS 網路中傳送的過程.....41
5.2.4 標籤交換路徑(Label Switched Path；LSP)建立的方式.....42
5.2.5 其他有關應用在 VPN 的解決方案.....45
5.2.6 Layer 3 MPLS VPN.....48
5.2.7 Layer 3 MPLS VPN 服務供應商之間的介接.....58
5.3 MPLS VPN 的服務品質(QoS).....60
5.3.1 MPLS VPN 邊緣路由器與核心網路之 QoS 管理.....61
5.3.2 MPLS VPN DiffServ 與 IntServ 的整合.....64
5.3.3 流量工程.....65
5.3.4 MPLS 動態路由及資源分配.....66
5.3.5 MPLS VPN 自動化服務品質管理.....70
5.4 VPN 解決方案之相互比較.....72
5.4.1 技術面.....72

5.4.2 架構建置面.....73

5.5 結 論.....74

第六章 研習心得.....76

第一章 前 言

近幾年來企業 VPN 的技術不斷的推陳出新，網路加值應用也呈現多樣化，其目的在於創造更便利、更安全、更快速的網路傳輸技術。而多重協定標籤交換技術(MPLS) 是實現寬頻 VPN 最佳的方式，它整合了目前各種交換式路由器技術的優點，其結合了 ATM 快速化、簡單傳輸的優點，以及傳統 IP 的普遍性(ubiquity)、延展性(scalability)及彈性度(flexibility)優點。在以寬頻為主的今天，不同的形態的資料串流(語音、影音、視訊)與數據資料整合在同一個網路，不但可以節省網路設備投資的成本，更可以輕易的維運與創造豐富的營收。

回顧寬頻網路發展的歷史，第一代寬頻網路是以路由器為主幹的架構，雖然語音、影像等多媒體資訊均可在 IP 網路上傳送，但是當時路由器設計屬於非連接導向(Connectionless Oriented)，無法區分封包的優先權，只能以「先進先出」的方式運作，因此一旦封包過大，路由器 CPU 的乘載能力就成了挑戰，造成無法有效利用頻寬的問題，使得企業客戶在備援及頻寬分享上有很大的問題。

第二代寬頻網路則是以訊框傳送(Frame Relay)和非同步傳送模式(Asynchronous Transfer Mode，ATM) ，建置於網路的第二層(Level 2) ，路由器跑在第三層(Level 3) ，由於 FR/ATM 的封包交換速度遠比路由器快上千倍，所以整個網路傳送速度都會變快。形成具有交換功能的網路，可讓企業客戶直接做到點對點傳輸、備援及頻寬分享。

第三代寬頻網路則是著眼於未來網路會結合語音、影像、數據加值的廣泛應用，使得高速、低延遲(low latency) 、即時應用服務的品質保證服務品質、頻寬使用效率與網路安全更加重要，因此多重協定標籤交換技術(MPLS) 成了第三代網路的重要需求，而 MPLS 服務品質(QOS)技術可將網路資源與各種服務資料串流的利用達到最佳化，來提高寬頻網路整體效能。

第二章 研習行程及課程

研習行程及課程

93 年 10 月 3 日(星期日)：行程，搭機赴美國舊金山。

93 年 10 月 4 日~10 月 8 日：實習課程

1. Juniper Nerworks Security Solutions
2. NetScreen Advanced VPN Implementation

93 年 10 月 11 日~10 月 14 日：實習課程

Juniper Networks QoS Solution with Integrating Voice ,Video ,Data M/T/J Series

93 年 10 月 15 日~10 月 16 日(星期五~星期六)：

返程，搭機回台北。

第三章 影響 VPN 服務品質(QoS)的因素

3.1 頻寬不足: 有限頻寬多種服務同時佔用

頻寬不足將是往後 VPN 所面臨到的一個主要問題，頻寬的競爭和衝突嚴重危害了整體網路的傳輸效能，突發性資料量往往導致網路上壅塞等問題，造成網路交通量變得緩慢而遲鈍，進而直接衝擊到網路整體的效能。 傳統解決網路壅塞問題的方法是唯有持續加大網路可提供的頻寬，即是超載供給(Over-provisioning)網路架構，然而，網路人口及頻寬需求的迅速發展遠遠超過目前網路技術所能供給頻寬的能力，所以超載供給網路架構一時之間亦無法有效解決網路不斷突發的壅塞量，唯有依靠有效率的管理整體網路流量，改善壅塞與衝突的發生，妥善地使用頻寬，以解決頻寬不足的當務之急。

3.2 傳送延遲: 封包經各種網路設備處理與轉送結果

以傳統的路由器構成的網路為例，路由器接收到任何一封包均將其拆開後分析封包性質、判斷是否符合安全機制等後，交由路由處理器決定其所要經過之路徑，由於每一封包所經之路徑，因當時的網路狀況不同而有所不同，因此每一封包傳送(Forward)的速度便受到相當大的限制。

如果在網路上又加上語音應用，則延遲的來源有無形中增加了累

積延遲，壓縮處理延遲、網路延遲。

1. 累積延遲(accumulation delay)：這是語音編碼器(voice coder)在收集音框(frame)時所造成的延遲，每種音框導向編碼/解碼器(Frame- Oriented CODEC)都有其 frame time，G.723.1 一個 frame 為 30ms。
2. 壓縮處理延遲(processing delay)：這是語音編碼器(voice coder)在壓縮及編製封包 (packet) 時，所造成的延遲。
3. 網路延遲(network delay)：這是封包在網路傳遞及接收端緩充器在移除 packet jitter 所造成的延遲。

3.3 延遲變化: 多個資料流相互影響所產生的變化

有些數據資料的傳送由於首重資料的完整性(Data Integrity)會賦予加權設定，資料的正確與否是最重要的，時間的延遲上反而不那麼重要，雖然可使高優先權的資料流可以得到較好的服務品質，但是對於即時多媒體(Multimdeia)的資料傳輸而言，可能會因為網路的流量隨時間改變而導致多媒體資料的延遲變化(Delay Jitter)無法有效控制，最後造成多媒體的播放品質低落。

3.4 封包丟棄: 網路壅塞的結果

以 IP 網路而言，通常數據資料的應用程式使用的是 TCP 協定相互通訊，TCP 的機制還可以簡單控制網路的擁塞情況，但是即時多媒體(Multimdeia)的資料傳輸大多使用 UDP 協定來傳輸。而這種型式的流量最大的問題是，它不僅產生很大的流量而且它

並不會像使用 TCP 協定的應用程式，可以回應網路的擁塞情況。因此，使用 UDP 的串流媒體會引發兩個問題：擁塞崩潰(congestion collapse)及不公平的頻寬分配。

然而隨著網路流量大量的成長，但 TCP 這個擁塞控制協定卻只有微小的改變。在所有使用者使用相同流量控制協定的情況下，TCP 的終點對終點適應性流量控制(end to end adaptive flow control)可以讓使用者公平的分享網路頻寬。 然而這樣的情況卻面臨一個重大的威脅：串流媒體流量的大量成長。 串流媒體使用 UDP 協定傳輸，但 UDP 沒有提供流量及擁塞(congestion)控制。而且 UDP 所使用的演算法會儘量去使用可用的頻寬，因此它會產生兩個問題：擁塞崩潰(congestion collapse)及造成競爭的流量間不公平的頻寬分配。

第一個問題：Congestion collapse，造成的原因是網路過度擁擠，以致於有些封包在到達目的地之前就會被丟棄。 許多網路應用程式都使用 UDP，但 UDP 並不回應網路的擁擠，或是當有封包被丟棄時，減少封包的傳輸量。 事實上，在網路擁擠時，許多應用程式會增加封包的傳輸量，甚至傳輸多個相同的封包，以確保封包能到達目的地，這只會使擁塞崩潰的情況更加嚴重。第二個問題，不公平的頻寬分配，有許多的因素引發這樣的問題，其中之一的原因是目前的網路應用程式不會去適應(adapt)擁塞的情況。當網路擁擠時 TCP 的應用程式會減少它的傳輸量，以致於頻寬被使用 UDP 的串流媒體搶去，造成 TCP 的頻寬縮小的不公平情況。 而 IP 協定本身也會造成頻寬分配不公的情形，

例如：TCP 演算法會使有較短的折返時間(round trip time)的 TCP flow 擁有較大的頻寬。

除了根本改變串流應用程式之外，解決這些問題的其它方法可以分成兩大類： 一是管理網路邊緣(the edges of the network)的交通量；二是依靠網路本身的服務。 在第一類的方法中，網路邊界巡邏(network border patrol)會比較網路的兩端，進入和出去的流量，確保進入網路的封包量不會大於出去的封包量。 這樣可以預防 congestion collapse 的情況。

第二類的方法是使用 ATM 與 MPLS 網路。

ATM 支援不同的應用及服務，在多媒體服務的傳輸架構上，它可以確保在 TCP 及 UDP 流量間公平的分配網路資源。ATM 網路提供四種服務：constant bit rate(CBR)、variable bit rate(VBR)、unspecified bit rate(UBR)、available bit rate(ABR)。 CBR 及 VBR 提供保證服務品質(Quality of Service)給使用者。 CBR 及 VBR 的優先權高於 ABR 及 UBR，CBR 及 VBR 未使用到的資源才會分配給 ABR 及 UBR。

MPLS 技術可提供流量工程 (Traffic Engineering)的方法, 針對不同的資料流容忍度及客戶的需求，給予適當的網路決策。 達到妥善管理網路上的資源、控制網路流量與特定路由的流動，減少擁塞情形並改進流量效率，使得網路效能達到最高。

綜合所有可能影響 VPN 服務品質的因素可概略分為四大類，如下表 4.1 所示。

	流通率/吞吐量 Throughput	延遲 Delay	封包遺失 Loss	延遲變化 Jitter
互動式(interactive) ex: Telnet	低	低	低	不重要
背景式 (background) ex: FTP	高	不重要	低	不重要
交談式 (conversational) ex: voice	低	低/可預測	低	低
串流式(streaming) ex: video	高	低/可預測	低	低

表 4.1 影響 VPN 服務品質(QoS)的因素

第四章 服務品質(QoS)的網路應用

4.1 服務品質的類型

●盡力而為(Best effort)

盡力而為(Best-Effort)方式可以用尖峰時段的高速公路上的交通狀況來說明，如果每一個人都使用同樣的道路，在此同時每個人都竭盡所能地想要儘快地達到自己的目的地。 道路上交通的流量是採取盡力而為(Best Effort)的運輸方式，因為沒有一個駕駛能保證自己到達時間。 壅塞與碰撞偶爾會阻礙網路效能造成交通變得緩慢與遲頓，而嚴重的壅塞與碰撞更會造成架駛者永遠無法到達目的地。 同樣地在大部份的網路上，封包是採取盡力而為(Best Effort)的方式傳輸，但是網路上沒有提供任何保證頻寬的工具可保障傳輸無障礙，而有些簡單的壅塞控制與優先權設定，不但無法解決網路壅塞的問題，相反的會限制網路整體的效率。 因為這些技術無法辨識 Layer1-Layer4 不同應用資料流並保護它們傳輸的品質。

●整合型服務(Integrated Service)

整合型服務(IntServ)是以每一個應用程式為單位，透過網路信號方式進行，所以可以得到精確的服務品質，而相對的，網路路由器必須處理的資料量會大幅增加，因為它必須為每一個應用程式建立一組相關的服務品質參數，並且不斷地去計算、儲存相關參數。 雖然以

目前的交換式路由器而言，這些工作皆由硬體完成，然而其容量終究有所限制，所以整合型服務一般是置於網路邊緣端，以 Per-flow 方式，在資料流傳送之前使用資源保留協定(Resource Reservation Protocol； RSVP)技術針對資料流所需要的品質事先找出一條路徑來保留頻寬，之後才開始傳送整個資料流。 然而針對資料流種類的增多，其網路容量將受到限制，且難以管理各個不同資料流的需求。

●差異型服務(Differentiated Service)

差異型服務(DiffServ)為 Per-class 方式，依各種規則或方法將資料簡化與分類，並決定各類別的特性及傳送優先次序。 而不需對每個資料流個別提供特定服務，只需針對所分類的等級來提供不同服務。換言之，差異型服務是將這些相關應用程式依規則加以分類，針對每一個分類給予一組服務品質之參數，依參數的定義來傳輸資料，如此一來將可以避免整合型服務資料量太多的缺點，所以分類型服務適用於核心的網路架構，但是這樣的差異型服務的服務品質會較不精確，所以並不能夠全程適用於通訊協定中。

不論何種服務型態，路由器均須具備監控及排程功能，以作為網路監管之用。

4.2 各種服務品質(QoS)的機制

● 排隊(佇列)方式 Queuing

同一網路節點的介面會同時載送多種服務，因此直覺式的先進先出(First In First Out)之處理機制並不適用，必須有新的機制來管理緩衝器和處理資料，如：優先權佇列(Priority Queuing；PQ)、客制化

佇列(Custom Queuing；CQ)、比重公平佇列(Weighted Fair Queuing；WFQ)、改良式逆差循環排程(Modified Deficit Round Robin；MDRR)、分級加權公平佇列(Class-based Weighted Fair Queuing；CB-WFQ)、分級低延遲佇列(Class-based Low-latency Queuing；CB-LLQ)。

● 流量塑型 Traffic Shaping

在封包上印記標籤，以區分其服務的等級，並且網路的節點根據標籤的不同而做不同的處理。當使用者將資料送入網路的入口，入口節點根據服務的等級將封包的標頭做對應的標籤映射，而網路的節點將根據標籤做出相對應的處理，以確保其服務品質如：一般流量塑型(Generic Traffic Shaping；GTS)、訊框傳送流量塑型(Frame Relay Traffic Shaping；FRTS)、分級流量塑型(Class-based Shaping)。

● 流量管理 Traffic Policing

在網路上建立一條端點對端點的路徑，並在此路徑上預留頻寬，以保障此流量的服務品質，而因為此法以點對點的資料流為基礎來預留並保證頻寬，故網路上的各節點都必須紀錄此流量的狀態如優先頻寬保證(Committed Access Rate；CAR) 、分級管理(Class-based Policing)。

● 丟棄方式 Dropping/Discard

當使用者要傳送資料前會透過信號通信規範與網路提供者定義其可使用的資源，網路提供者根據協定監視使用者傳送資料的行為，若使用者未依規定傳送資料、為了保護網路上其他的使用者，網路提供者會將未依規定的資料捨棄如：加權隨機預先偵測(Weighted

Random Early Detection；WRED) 。

● 信號方式 Signaling

當使用者要傳送資料時，必須先通知網路提供者其所需的服務品質為何，網路提供者根據使用者的需求，判斷網路是否有足夠資源來提供該項服務，而這些機制統稱為信號通信規範，如：資源保留協定(Resource Reservation Protocol；RSVP)。

4.3 整合型服務(Integrated Service；IntServ)

IntServ 基本概念是“所有封包流 (packet flow，即一串具有相同來源與目的地之 IP 位址與 port numbers 的封包)相關狀態訊息，應該會出現在端點或邊緣系統上”並參考封包標頭所攜帶的服務型態(Type of Service；ToS)(圖 4.1)值，使得每個封包流（單個的或是匯聚的）提供點到點的保證或是受控負載的服務（controlled-load service）。 Int-Serv 框架使 IP 網能夠提供具有 QoS 的傳輸，可用於對 QoS 要求較為嚴格的即時服務（聲音/訊頻）。

Int-Serv 使用一種類似 ATM 的 SVC 的方法，它在發送端和接收端之間用 RSVP 作為每個封包流的信號。 RSVP 訊息橫跨整個網路，假設從接收端到發送端之間沿途的每個路由器都要為每一個要求 QoS 的資料流預留資源。 路徑沿途的各路由器(包括核心路由器)必須為 RSVP 資料流維護狀態。

在 Int-Serv 流中，定義了三種類型的服務：

1. 保證服務(Guaranteed-Service；GS)： 對於 GS 服務流的最大延遲是受到控制的，路由上的任何的延遲都會影響最大排隊延遲。

2. 受控負載服務(Controlled-load-Service；CLS)： CLS 沒有固定的延遲的保證，但服務流要與在網路輕載情況下的流量相等，實際上 CLS 要求有長期的寬頻保證。

3. 盡力而為的服務(Best-Effort)： 類似目前 Internet 在多種負載環境(由輕到重)下提供的盡力而為的服務。

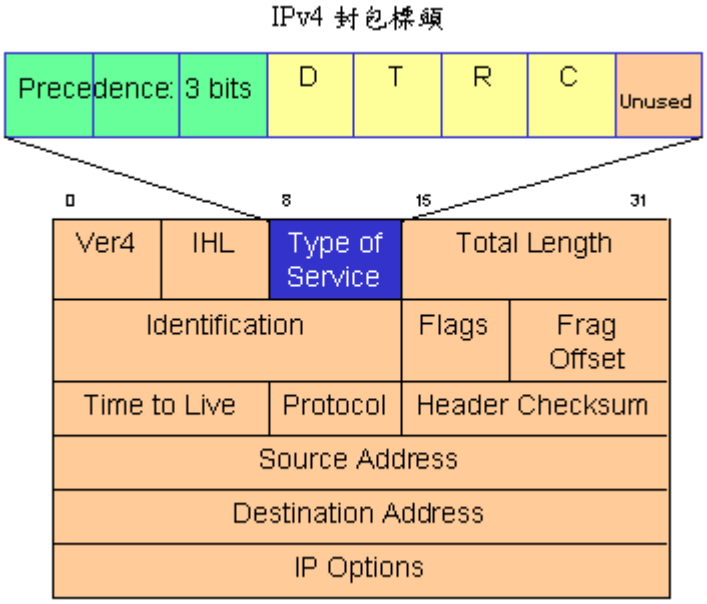

圖 4.1 Type of Service

在 IntServ 網路中，使用者必須透過 RSVP 通訊協定如下表 4.2 所示。 來宣告封包流 (packet flow) 的特性 (透過 PATH 訊息之 Tspec 參數)以及作頻寬保留 (透過 RESV 訊息之 Rspec 參數)。路徑上的所有路由器都必須處理這些 RSVP 訊息，包括了記錄 PATH 訊息的路徑、參數，以及依據 RESV 訊息之參數和路由器資源與 link 頻寬使用的狀況對使用者要求之頻寬或服務執行登錄管制

(admission control)。 除了 RSVP 協定訊息之外，路由器還必須對每一封包執行 Multi-Field (MF) 篩選分類以辨識封包流，並對每一封包流執行管理(policing)與排程(scheduling)。 換言之，在 IntServ 的架構裡，無論是核心或邊緣路由器都必須要能夠辨識、記錄、並且管控每一封包流的狀態。 如此一來，隨著網路的逐漸擴張，封包流大量增加，無論是在儲存設備或是處理速度方面對於路由器而言都是一大挑戰，這就是 IntServ 一直為人所詬病的延展性(scalability)問題。

RSVP 訊息內容說明	
訊息種類	功　　能
PATH	從傳送端的電腦中傳送資料流資訊給接收端電腦。
RESV	從接收端電腦中傳送保留項目的請求，其中的內容包括了頻寬大小、服務層級以及來源的 IP 位址。
PATH-ERROR	回應 PATH 訊息所產生的錯誤。
RESV-ERROR	回應 RESV 訊息所產生的錯誤。
PATH-TEAR	沿著運作的路徑移除 PATH 的狀態。
RESV-TEAR	沿著運作的路徑移除保留項目 。
RESV-CONF	選項設定。假如接收端需要一個確認訊息，那麼傳送端就會發出這個訊息給接收端。

表 4.2 RSVP 訊息內容說明

為了執行整合型服務(Int-Serv)的服務，Int-Serv 定義了 4 個功能元件，網路中每個路由器皆需要執行這 4 個元件。

1.RSVP(RFC2205)： RSVP 即資源保留協定，它是 Internet 上的信號協定。透過 RSVP 如圖 4.2 所示，用戶可以給每個服務流(或連接)申

請資源預先保留，要預留的資源可能包括緩衝區及頻寬的大小。這種預留需要在路由器上的每一 hop 都要進行，這樣才能提供點到點的 QoS 保證。 RSVP 是單向的預留，適用於點到點以及點到多點的通信環境。

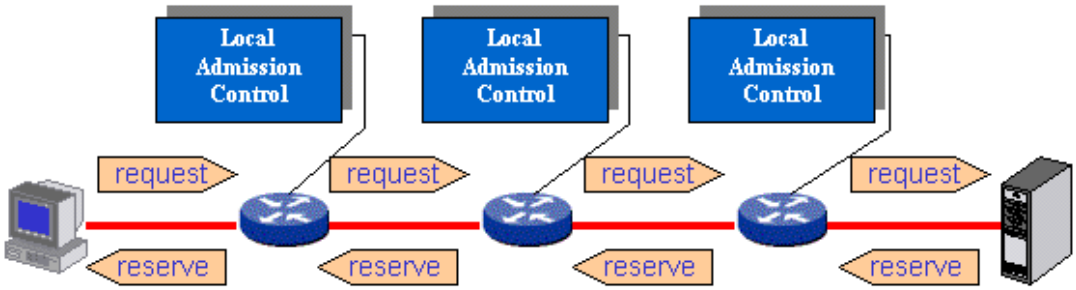

圖 4.2 點對點 RSVP

2.登錄管制(Admission Control)： 根據用戶和網路業者達成的服務協議，對用戶的登錄進行一定的監視和控制，有利於保證雙方的共同利益。

3.分類器(Classifier)： 根據預設的一些規則，它對進入路由器的每一個分組進行分類。 這可能需要查看 IP 分組裡的某些區域：IP 來源位址、IP 目的位址、上層協定類型、來源 Port、目的 Port，分組經過分類以後被放到不同的排隊列中等待接收服務。 這方面的技術還不很成熟，是一個有待研究的領域。

4.隊伍排程/調度器(Scheduler)： 根據一定的隊伍調度演算法對分類後的分組隊列進行調度服務。這方面的技術目前已比較成熟，常見的調度算法有比重公平佇列(Weighted Fair Queuing；WFQ)、改良式逆差循環排程(Modified Deficit Round Robin；MDRR)、分級加權公平佇列(Class-based Weighted Fair Queuing；CB-WFQ)、分級低延遲佇列

(Class-based Low-latency Queuing；CB-LLQ)等等。

另一種實現整合型服務(Int-Serv)的方式為共同開放策略服務(Common Open Policy Service；COPS) 如圖 4.3 所示，是一種以策略為基礎的網路管理 (Policy-based Network Management)它可以用來解決彈性與協調性不足的問題。 該協定運作是把策略的決定權委託給策略伺服器 (Policy Servers)，策略伺服器又可以稱為策略決定點(Policy Decision Points；PDP)。 不同的 PDP 可以透過 SNMP 或 LDAP 來存取中央策略資訊庫裡的資訊內容，而系統管理者再來管理中央策略資訊庫裡的資訊內容，策略服務者可以依照需求來提供恰當的組態資訊給策略執行的主機或網路設備。

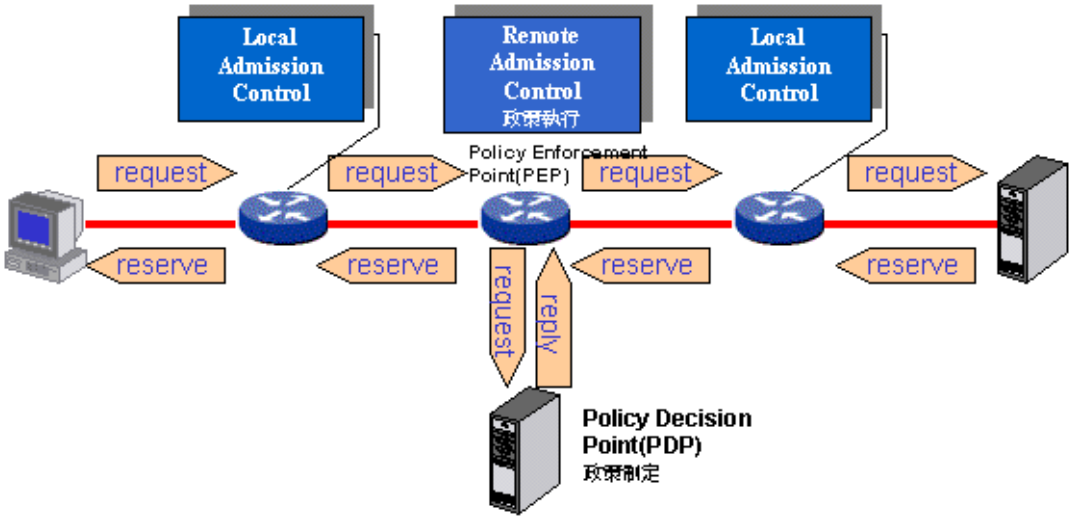

圖 4.3 共同開放策略服務(Common Open Policy Service)

4.4 差異型服務(Differentiated Service；DiffServ)

為了依照每一資料流的需求特性給予不同的服務品質保證，IEFT 組織提出了差別性服務（Differentiated Services；DiffServ）QoS 機制。當封包進入邊緣路由器時，邊緣路由器便將封包分類，並標上類別碼(DiffServ Code Point；DSCP)，當封包進入核心網路時，核心路由器依照其類別進行相對應服務等級的排程，以提供承諾的服務品質。

針對各種不同的流量等級所做的轉送機制的方法，稱為封包排程行為(Per Hop Behavior)。 就其涵意，簡單而言，是對可觀察的各點其轉送行為的描述，舉例而言，某個 PHB 的內容就是在保證某等級的聚集可以使用某百分比的頻寬。而 PHB 的實作則可能是以某種排隊/佇列規則(Queuing Discipline)及某種排程機制(Scheduling Mechanism)來完成的。

下圖 4.4 為一個典型的 DiffServ 架構，當發送端（Source）的 DiffServ 網域要發送封包到目的端（Destination）時，途中到達另一個 DiffServ 網域，當要進入口節點（Ingress）時，封包便被進行分類，進入核心網路後，便會依照類別碼給予對應之服務，再從出口節點（Egress）送往目的端。

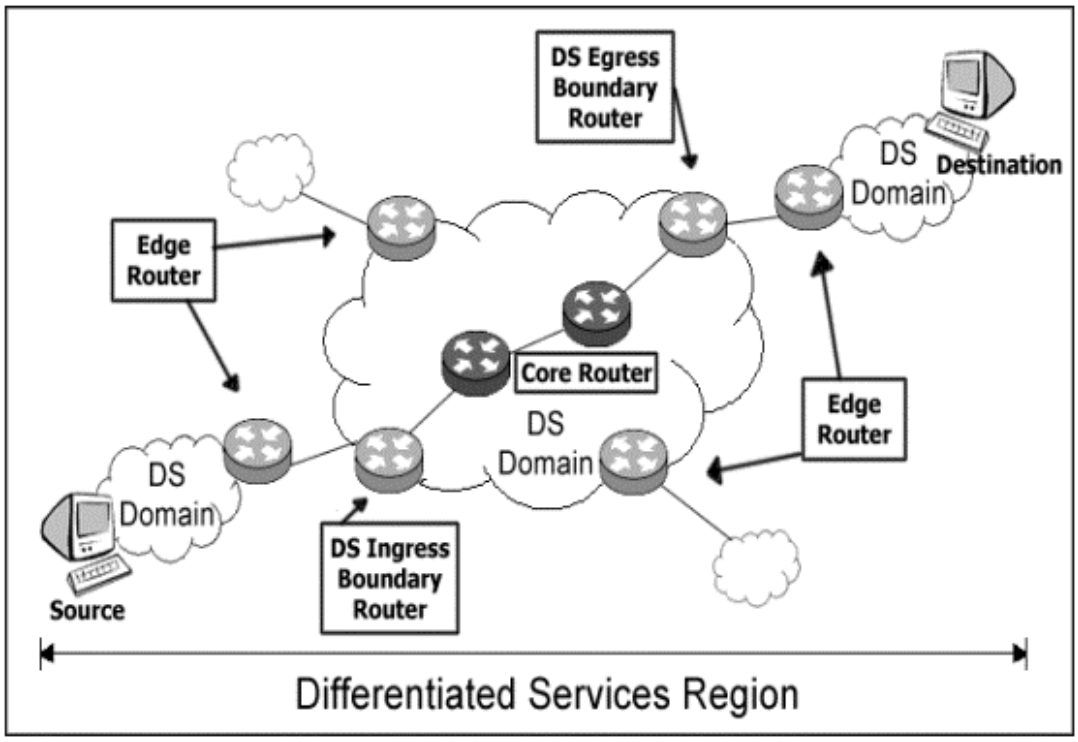

圖 4.4 差異型服務(DiffServ)

IETF 目前建議了三種 PHB 轉送封包的服務等級(如表 4.3)：

- **快速轉送（EF, Expedited Forwarding）**：此類的封包擁有最高的轉送優先權，保證延遲時間(delay time)，通常用來傳遞重要訊息，例如路由器上的信號等。 此種服務等級只有一個類別碼，就是所謂的保證服務(Guaranteed Service)。

- **保證轉送（AF, Assured Forwarding）**：這個服務等級比 EF 低一點，但比 BE 高。 AF 又細分為四個等級，每個等級又分成三個不同的封包丟棄優先順序，若遇到網路擁塞現象，則依其優先順序丟棄封包。

● **盡力轉送（BE, Best-Effort Forwarding）**：即為傳統 IP 網路上所使用的 BE 封包轉送機制，封包一律排隊，頻寬夠時就繼續送，不夠時就隨機丟棄封包。 如果沒有設轉送規則時，封包之等級預設為 BE。 原有的 IP Precedence 仍被保留，將封包分為八個等級。

封包排程行為 PHB(Per Hop Behavior)				
PHB 等級	相對優先權	特 性	對於設定頻寬	應用環境
Expedited Forwarding(EF)	最高	低延遲、低遺失率、保證頻寬、不受其它傳輸之影響，有如私人專線	無法高於該設定頻寬，有如頻寬限制一般	網路電話、視訊會議、高品質傳輸、網路信號傳輸等
Assured Forwarding(AF)	較高	高傳輸量、保證最小使用頻寬	網路未擁塞時，可使用較設定頻寬為高之頻寬	網路資料交換、網站瀏覽、檔案傳輸等
Best Effort(BE)	最低	服務品質受到網路使用率影響，無法預測	無設定頻寬，一切只能視網路負載而定	電子郵件、或較不急需回應之應用

表 4.3 封包排程行為 PHB(Per Hop Behavior)

4.5 流量調整器

流量調整器是差異型服務中一個相當重要的組成部份。 它的目標是將控制的功能應用在之前分類過的封包上。 一個流量調整器(如圖 4.5 所示)通常包含了以下各個子元件：

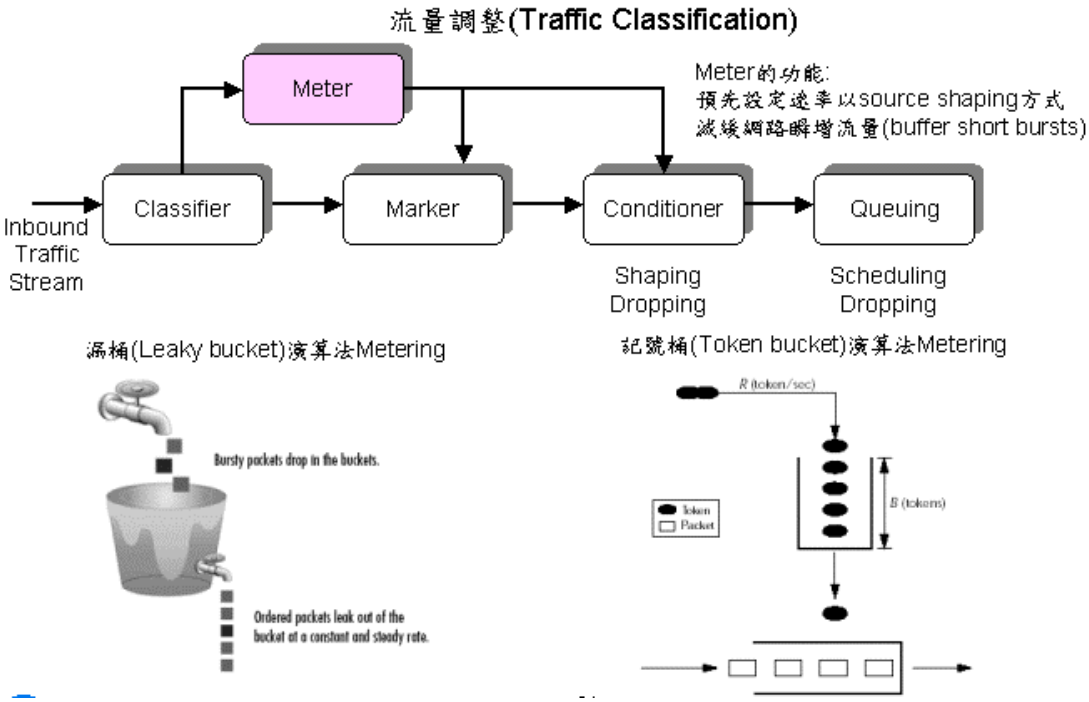

圖 4.5 流量調整(Traffic Classification)

● Meter： 用來測量由分類器所選擇的資料流之特性。 Meter 是用來監視封包的到達時間，藉此來計算速率或是延遲，以便和先前預定 Profile 內之速率或延遲要求比較其符合程度。 Meter 也是一個 1：N device，它輸入了由 classifier 給它的單一 flow，卻輸出該 flow 內符合程度有差異性的封包(例如依據未達到協議速率與超過協議速率的封包粗淺分為兩類)，而可分別交給不同的後端元件處理。一般實作 meter 時，常見的

有 average rate meter 與 token bucket meter 兩種。 average rate meter 十分簡單,只要計算在上一封包離開的時間至下一封包到達時間之時間差除以下一封包的封包長度，便可粗略的估計出該封包的速率。 而 token bucket meter 則是由系統依時間產生固定數量的 token，封包通過 meter 需取得相對數量的 token，而 token 若未被使用殆盡可留至下一個封包使用，好處是能夠應付大量爆發性的封包來到。

● Classifier： 封包的分類是執行差異型服務時一個重要的功能，它通常在差異型服務網路領域的端點進行，它是要辨別封包的服務種類，並且提供不同的服務類型。目前有兩種分類器已經被定義；「類行為聚集」(Behavior Aggregate；BA)分類器與「多欄位」(Multi-Field；MF)分類器。 BA Classifier 只依據 DSCP 來做分類，而 MF Classifier 則可能根據了一個以上的欄位來做分類，如來源位址(埠)、目的位址(埠)以及傳輸層所使用的協定(TCP 或 UDP)等資訊，以 DSCP 來做 QoS 較為方便簡單。

● Marker： 根據已決定的規則來設定 DSCP。 Marker 是一個 1：1 的 device 用來對 IP header 內的 DSCP 做標記的動作。DSCP 的寫入，取決於 Classifier 以及 meter 對 IP flow 內封包的分類，而 DSCP 決定了該封包在核心網路裡面將會如何被處理的方式。

● Shaping/Dropping： 將某一資料流的流速予以減緩以符合之前已定義的規則，Shaper 是用來將爆發性(burst)大資料流予以平滑化的元件，而這跟前端所使用之 meter 有關，若是使用 token bucket meter，由於該演算法本身就有將資料流平滑化的特性，故 shaper 於其中便不適用了。

● Scheduling/Dropping： 將某些封包丟棄以符合之前已定義之規則。 Dropper 簡單而言就是丟棄封包的元件，對整個 data-path 來講是個 terminal node。 Meter 判別過的資料若符合約定速率的程度太低或是依 QoS 特性的不同若需要被丟棄便會被送至這裡來。

4.6 緩衝管理機制

緩衝區管理機制必需滿足三個要件：能儲存封包、能調解及排列封包送出的順序、有丟棄封包的機制。

而三個要件分別可成為三個獨立元件的： Queue、Scheduler、Discarder。 其中 Queue 代表暫存封包的資料結構，Scheduler 代表了排程的機制，Discarder 則是選擇丟棄封包的機制。 一般暫存網路封包的緩衝區資料結構是以 FIFO(First In First Out)的 Queue 為主，因為上層的通訊協定通常會將到達次序顛倒之封包丟棄，因此路由器或是網際網路協定都必需能夠保證所提供的架構不會發生 misorder 的情形。 而 Queue 符合網路先到先處理的特性，封包進入佇列後不能馬上被傳送出去的話，就暫存在佇列內，而能保證先

到的封包會比後到的封包先傳送。 FIFO Queue 最少需要 en-queue 和 de-queue 兩種功能，以及保留一個表現佇列長度的參數深度值 (depth)。而如何選擇 Queue 深度值的最大值也是佇列排隊理論中一門重要的課題，因為如果深度值選擇太大，會造成緩衝區停留太多的封包，造成許多已經逾時的封包仍存留在佇列中；但若選擇的太小，又會造成頻寬無法充份被利用的問題。 而 scheduler 的變化就較為多樣了。視該路由器是針對什麼樣的特性而設計的，如果是針對 IntServ/RSVP 設計的路由器，則需要能夠處理 Per-flow 等級的機制。一般常見的有 FCFS(First Come First Serve)或是以 Priority 為主或是以 Fair Sharing 為原則的 scheduler。 FCFS 類的排程，優點僅有易於實作，但是只能夠達到 Best Effort 的品質。 Priority 則是對某些特定的串流(Stream)做某些資源上的保留，讓高優先權的串流比其它低優先權的串流擁有更高使用資源的機會。而 Fair Sharing 的機制則是期望讓各類串流能夠公平的分享頻寬，或是透過給予權重(Weight)的方式，讓頻寬得以重新分配與利用。 scheduler 是從 Queue 的尾端去減少緩衝區內封包的數量， Discarder 相對來講則是在 Queue 的前端利用丟棄封包的方式來減少緩衝區滿溢的機會。它通常有個計算平均佇列長度的方式，並依該數據來評估是否丟棄一個正要被 en-queue 的封包，而丟棄封包決定於計算出來的機率，例如若是佇列此時長度超過某個標準，可能丟棄的機會就會比低於該標準時大的多。

● Discarder

隨機預先偵測(Random Early Detection；RED)：

它是一種主動丟棄封包的機制。 當各一個封包到達後，它便計算平均的佇列長度(Average Queue Size；avg-qsize)，如果這個值超過了某個上限(Threshold)，便計算出機率並以該機率來決定是否丟棄該封包，而計算機率的方式是由一個avg_qsize 當自變數之函數計算求得的。而一個RED 閘道，有以下三個參數應該被設定： min_th，max_th 及Pmax。在RED 中存在三種狀態：由 avg_qsize 所依序組成的(0, min_th)，(min_th, max_th)與(max_th, ∞)三種狀態，這三種狀態分別為「正常運作狀態」(Normal Operation)、「擁塞避免狀態」(Congestion Avoidance)與「擁塞控制狀態」(Congestion Control)。在正常狀態下，RED 閘道會計算出丟棄封包的機率為0，意即不丟棄該封包。在擁塞避免狀態下，RED 閘道會得到一個與avg_qsize 增加成正比的丟棄機率來決定是否丟棄該封包。而在avg_qsize 達到max_th 後，進入擁塞控制狀態，進來的封包一律丟棄，藉此期能降低avg_qsize 回到穩定狀態。

● Scheduler

優先權佇列(Priority Queue；PQ) ：

是由一個或一個以上的 queue 與一個具有優先權的scheduler 所組成的。 scheduler 可藉由設定等方式，使得某一佇列內的封包可以較其它佇列內的封包具有優先傳送的機會，這就是優先權的概念。例如佇列A 若較佇列B 擁有較高之優先權，則scheduler 會優先服

務停留在佇列A 中之封包，直到佇列A 中不再有封包時才會服務佇列B。但隨之而來的問題是，一但佇列A 中封包到達的速率大於scheduler所能服務的速率時， scheduler 就沒有足夠剩餘的時間分配給佇列B，便會造成了飢餓狀態(starvation)，這樣子會使得網路的效能更加低落。解決的方法是給予佇列權重，意即優先權非絕對而是相對，例如佇列A在連續處理了m 個封包之後，就必須處理佇列B 之n 個封包，藉此來取得減少飢餓的機會。因此有了WRR演算法。

佇列循環(Round Robin；RR)： 是一種很簡單的公平排程機制(Generalized Processor Sharing)。 假設一佇列循環(RR)的系統中存在q0 … qn 個佇列，則具有RR 處理能力的scheduler 會從q0 開始處理固定量的資料(單位可以是packet 也可以是byte)，而超過該量之後，就接著服務q1…依次下去，直到服務 qn 後、再跳回服務q0，完成一個回合(round)。而具有權重的RR，則代表了每次服務各佇列時可以有不同的量，藉此來提高優先權。

一般而言，加權佇列循環(WRR) 雖然能夠簡單的做到某一程度的公平性，但是因為服務的量不同，還是會影響理想分配頻寬的期望，理想的WRR 應該每次服務以一最小不可分割的單位為基準，如bit或是byte，但是實作上卻難以實現，於是退而使用網路的最小單位packet 來處理，意味著我們若限定某一WRR 機制在每一回合中，跳躍至下一佇列的條件是處理現前佇列內的一個封包，但是由於封包的長度不固定，如果封包長度較長的，相對來講就比封包長度較短的佔有較多的頻寬。

改良式逆差循環 (Modified Deficit Round Robin；MDRR)： 改良自逆差循環(DRR)的方法，就是為每個佇列設定一計數器Cn，然後在scheduler 上設定處理各個佇列的基本量是以packet 為單位，假設為一常數d (credits)。 而在scheduler 初始化時，便將C0…Cn 皆設為d，然後開始處理q0。 處理以packet 為單位，而處理完後將C0 的值減去該封包的長度，如果尚未小於等於0，則繼續處理該佇列之下一個封包，但是倘若該下一個被處理的封包長度大於C0，則不再予以處理，而跳至q1 服務…依此類推，而到處理完qn 時結束一個回合時，再將C0... Cn 的值加上常數d，接著繼續處理q0。 藉由計數器的運作，間接的使得從長程(long-term)看來，賦予了WRR 某種程度的公平性。 在DiffServ 中，由於需要提供低延遲低時間的非同步轉送(Expedited Forwarding)的服務，所以具有優先權的排程服務是極為適當的方式，而又由於DiffServ 在領域邊界有控制進入流量速度的機制，所以只要是一個控制良好的網路，搭配適當的排程機制，不要讓高優先權的流量大於可處理的流量，應該可以達到提供EF 服務的目的。

4.7 整合型服務(IntServ)與差異型服務(DiffServ)之比較

盡其所能服務(Best-effort service；BES)是現今 IP 網路最基本的資料連結傳送方式，沒有所謂的頻寬保証，所以對於特殊型態服務則無法提供其完整需求。 為了解決此問題，IETF 於是最先提出了整合型服務架構(Integrated Services Architecture)，並定義了兩種新

的服務等級:一、保證服務(Guaranteed service；GS)，二、受控負載服務(Controlled-Load service；CLS) 。 保證服務(GS)以保障特別服務所需的頻寬或延遲時間限制，來滿足應用服務需求，但對於頻寬最大值以及最小延遲限制，較不具彈性。 受控負載服務(CLS)其主要保證即使在網路過載(overload)時，其服務可以提供如類似沒有過載的服務，而在網路無過載則可提供高速率、高頻寬、低遺失率等服務，以變通且動態服務為觀念來提供資源，並且較 BES 有較高優先權。

由於 IntServ 可以透過 RSVP 協定來宣告封包特性以及保留其所需頻寬，達到服務品質的保障，然而針對資料流種類的增多，其網路容量將受到限制，且難以管理各個不同資料流的需求。 所以有差異型服務架構(Differentiated Services Architecture)提出，DiffServ 則將 IntServ 的訊務資料盡量簡化分類，不需對每個資料流個別提供特定服務，只需針對所分類的等級來提供不同服務。 DiffServ 架構也定義了兩種新服務等級： 一、優質服務(Premium Service；PS) ，二、確保服務(Assured Service；AS)。 優質服務(PS)主要為使用者提供低延遲、低劇跳(jitter)、低遺失率且頻寬保證的傳輸服務。

目前定義的服務等級最高者，亦可稱為“虛擬專線”服務。 確保服務著重於頻寬及遺失率，不涉及延遲、劇跳。 AS 服務原則為無論是否網路壅塞，盡量保證使用者能獲得其所預約的最低限量頻寬。 由表 4.4 所示，可清楚比較 IntServ、DiffServ 以及 Best-effort 三種架構的一些特性差異，由於 IntServ 是採 Per-flow 來提供 QoS，所以其執行複雜度為最高，DiffServ 採 Per-class 次之，而 Best-effort

無服務品質考量故為最低。 在資源分配方面，IntServ 採針對每一資料流給於動態資源,而 DiffServ 則可以靜態或動態方式指定資源。對於控制管理方面，IntServ 在於主機和路由器皆能配合，DiffServ 則在邊緣節點執行標示分類，而使核心節點能專心處理資料封包流程，而 Best-effort 只有 First-In-First-Out(FIFO)。

IntServ、DiffServ、Best-Effort 之比較			
	IntServ	DiffServ	Best-Effort
QoS 解析度	Per-flow	Per-class	None
服務	保證服務 GS、 負載控制服務 CLS	優先服務 PS、 保證服務 AS	Best-Effort
資源分配	動態	靜態或動態	無
控制管理	主機和路由器皆有控制管理	Edge 端 Marking 標示分類，核心端則專責處理資料封包佇列排程 Queue mngt	僅有 FIFO
複雜度	高	中等	低

表 4.4　IntServ、DiffServ、Best-Effort 之比較

第五章 多重協定標籤交換技術之服務品質技術

5.1 概說

傳統 IP 網路是由眾多的路由器將不同區段所構成的網路,儲存和傳送(Store and Forward)程序則為傳統路由器運作模式，當路由器收到 IP 封包會先作儲存 IP 封包、分析路由，然後轉送封包至下一個路由器，如此完成一個 IP 封包的處理。 該動作會反覆進行直到目的地,使的網路運作顯得沒有效率。 後來為了提升網路的效率而有了 ATM 網路的虛擬私有網路(VPN)建置方式，但是 ATM 技術屬第二層的協定，因此傳送 IP 封包前，必須先將 IP 封包轉換成 ATM 細胞(cells)。ATM 網路技術屬連接式(Connection-Oriented)通訊協定，傳送資料需要事先建立好一條條虛擬通道，由於虛擬通道數量有限，因此有效率地運用虛擬通道極為重要。 ATM 技術因為屬於鏈結層(Layer 2)協定，沒有能力來分析第三層各種不同資料的流量特性。 然而現今的 MPLS 卻能結合了 IP 及 ATM 技術的優點，將目前各種交換式路由器技術的優點整合，並引用與 ATM 交換技術類似的標籤（Label）觀念，來提升整體網路的運作效率，可說是一種極具彈性及擴充性的 IP 交換技術。

5.2 多重協定標籤交換技術

多重協定標籤交換技術(MPLS)是最早由IETF所提出的一個先進的傳送技術，著重於有關封包傳送(Packet forwarding)以及路徑控制(Path controlling) 方面的路由。MPLS 全名為 Multi-protocol Label Switching，而其中從”multi-protocol ”字面定義可知此技術適用任何一種網路層的通訊協定的意思。 MPLS 將目前各種交換式路由器技術的優點整合，基本概念是引用與 ATM 交換技術類似的標籤（Label）的觀念，來提升整體網路的運作效率，可說是一種極具彈性及擴充性的 IP 交換技術。 在傳統 IP 網路中，IP 封包的 Layer 3 網路層包含著封包的路由訊息，路由器將根據 Layer 3 的訊息來獨立完成封包的路由(hop by hop routing)。 傳統的 IP Forwarding 中，一般的路由器如果接收兩個封包具有相同的 address prefix，則路由器將認為它們是同屬於一個相同傳送等級(Forwarding Equivalence Class；FEC)。 一 FEC 指出同一類封包將會以相同的方式對待傳送。並且在傳統網路中，當封包經過網路每一路由器時，都將有被重新檢查及對應到特定的 FEC 的情況。 然而在 MPLS 網路中，封包分配 FEC 只進行一次，映射到 FEC 的封包都被分配有一個固定長度數值的 Label。 當數據封包傳送到下一節點或路由器時，這固定長度的 Label 將跟隨著封包一起傳送。因而在頻繁的資料傳送過程中, MPLS 網路中的路由器將不再分析各封包的 Layer 信息，而藉以根據封包上的 Label 來決定下一個傳送的節點。 當封包傳送到達下一個節點時，Label 將被新的 Label 所取代，此過程稱為標籤轉換(Label

Swapping)。每一個 MPLS 封包都有一個 32-bit 的標頭(Header)，我們將此標頭稱作 Label。 Label 是一個簡短的、固定長度且只具有本地性意義的識別符(Identifier)。 因此 Label 沒有紀錄任何在網路層標頭的資訊，換句話說，Label 不紀錄網路層的 Source/Destination addresses，它只用來識別相對傳送等級(Forwarding Equivalence Class；FEC)，然後各路由器之間根據此 FEC 建立其所對應的轉送路徑，如此即可讓 MPLS 網路每一節點簡易地藉由查詢 Label，得知封包其下一傳送節點(Next hop)的資訊。 Label 資訊將告知路徑上每個路由器與交換器，不用再耗費多餘時間在進行查詢路徑的動作，節省網路運算資源。

一個 Label (如圖 5.1、表 5.1 所示)包含四個欄位： 一個 20-bit Label 欄位，一個 3-bit 實驗用的欄位或稱 Class of Service 欄位(CoS field)，一個 1-bit 的 Label Stack 標示欄位，以及一個 8-bit 的 Time-to-Live(TTL)欄位。

在鏈結層的技術下，如 ATM 技術，Label (如圖 5.2 所示)可以被載入在 ATM 標頭中的 VCI/VPI 欄位中，被視為鏈結層標頭的一部分，或者是被當作一個”shim” label 標頭穿插放置於鏈結層和網路層之間。

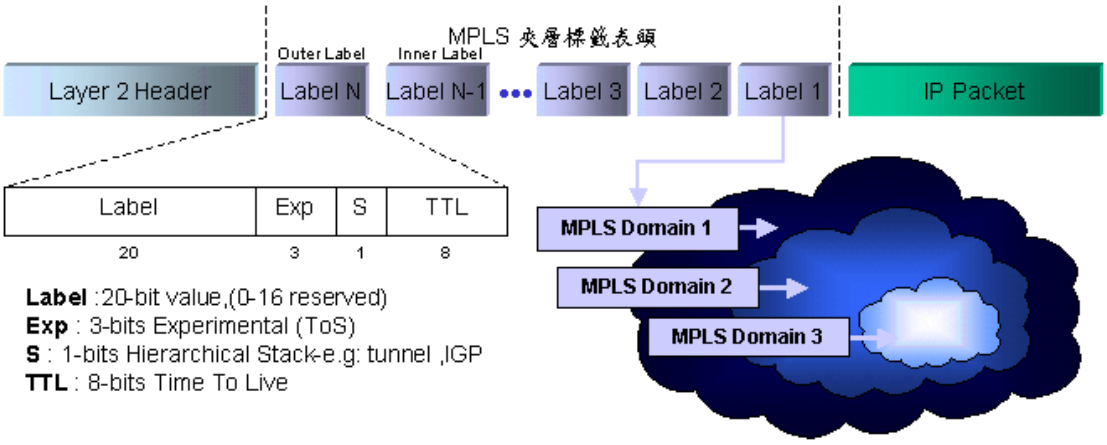

圖 5.1 MPLS 夾層標籤表頭

欄位名稱	長度	功　　用
Label	20	記錄 MPLS 堆疊實際的值
CoS	3	在網路傳遞時做為 Queuing 或是 discard 的依據
Stack(S)	1	支援階層性的堆疊，允許再封裝一個以上的標籤,當封包傳送時,永遠以最上層的標籤作為資料傳送依據,該被封裝多個標籤的空間稱為標籤堆疊
TTL	8	提供方便的 Time-to-live 的功能

表 5.1 MPLS 夾層標籤表頭

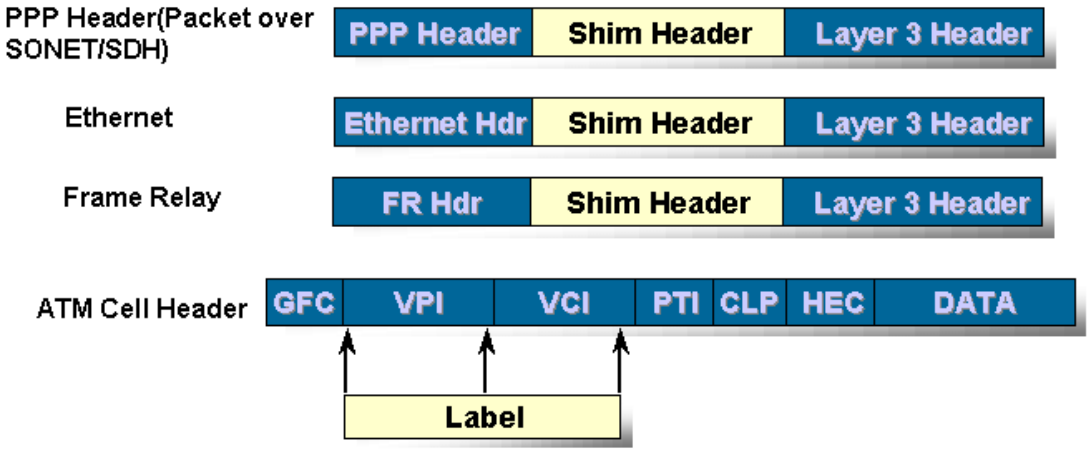

圖 5.2 其他夾層標籤表頭

5.2.1 MPLS 基本運作

MPLS 網路是由多個具有標籤交換能力的路由器 LSR(Label Switch Router)互相連結所組成，LSR 主要負責檢測標籤(Label)並且查詢其 LSR 本身的轉送表(Forwarding table)，然後再進行標籤傳送，以轉換到下一個節點。

而在本公司的 MPLS VPN 應用裡頭 LSR 能同時支援 MPLS 與 ATM 交換機的功能，我們稱之為 ATM-LSR。而 LSR 的運作的架構可分為兩塊：

1. 控制面(Control Plane)

   控制面也就是我們一般所說的路由引擎模組。該模組的功能是用來和其他 LSR 交換三層路由訊息，以此建立路由表和交換標籤對路由的綁定訊息，來產生標籤訊息資料庫(Label Information Base；LIB)。同時再根據路由表和 LIB 建立轉送訊息資料庫(Forwarding Information Base；FIB)和標籤轉送訊息資料庫(Label Forwarding Information Base；LFIB)。

2. 數據面(Data Plane)

   數據面的功能主要是根據控制面生成的 FIB 表和 LFIB 表轉發 IP 封包和標籤封包。

對於控制面中所使用的路由協定，可以使用任何一種，如 OSPF、RIP、BGP 等等，這些協定的主要功能是和其他設備交換路由訊息，產生路由表。這是實現標籤交換的基礎。在控制面中導入了一種標籤傳遞協定(LDP)，該協定的功能是用來針對本地路由表

中的每個路由條列生成一個本地的標籤，由此生成 LIB 表，再把路由表和本地貼上的綁定標籤封包通告給鄰居 LSR，同時把鄰居 LSR 告知的路由表和貼上的綁定標籤封包接收下來放到 LIB 表裡，最後在網路路由收斂的情況下，參照路由表和 LIB 表的訊息生成 FIB 表和 LFIB 表。實際的標籤指定(Assignment)與傳遞(Distribution)方式下一節再詳述。

根據在 MPLS 網路內扮演角色的不同 LSR 可以分為三種類型：
(1) Ingress Edge LSR(LER)： 負責將進入 MPLS 網路的 IP Packet 進行分類且設定其 Label 作為往後路由的資訊，此時的每個 IP 封包都將被貼上 MPLS 標籤(Push Label)，然後繼續轉送到下一個 LSR。
(2) Core LSR： 則位於 MPLS 網路的核心，負責做標籤轉換(Label Swap)。而運作方式則是將檢查進來封包的 Label 以當作 Incoming Label 的資訊，緊接著查詢其 Table，並決定 Outgoing label 來置換之，進而封包繼續轉換傳送至下一個 LSR。基於這種傳送操作原則，封包的 Label 不停地在於各個 LSR 之間作轉換，此轉換過程類似於 ATM 的 VCI/VPI 欄位轉換的情形。
(3) Egress Edge LSR(LER)： 當封包要離開 MPLS 網路到一般 IP 網路時，負責移除標籤(Pop Label)。

這些由入口的 LSR 到出口的 LSR 之間封包所行進的路線，統稱作標籤交換路徑(Label Switching Path；LSP)。 在 MPLS 網路中，通常使用一些訊號通訊協定，如資源保留協定(Resource Reservation Protocol；RSVP)、以及標籤傳遞協定(Label Distribution Protocol；LDP)等，來安排設定標籤交換路徑。 LDP 為 MPLS 技術用來分配

標籤而定義的通訊協定，主要目的是作各標籤的指定(Assignment)、對應(Mapping)以及傳遞/轉送(Distribution/Forwarding)的工作。換句話說,標籤交換路徑的安排是透過各個 LSR 執行標籤傳遞協定(LDP)協定，沿著路由路徑來分佈路由與作標籤(Label)的對映。LSR 會將所得到的標籤(Label)相關資訊儲存於標籤訊息資料庫(Label Information Base；LIB)中，LSR 之間會以 LDP 協定彼此交換資訊所建的 LIB，MPLS 封包則根據 LIB 來轉換標籤，同時以 LDP 協定的管理維護此 LIB 資料的正確性。由於標籤具有長度較短且固定的特性，故 LSR 可以快速地搜尋 LIB，使得 MPLS 網路運作將比傳統網路而更快、更有效率。

5.2.2 標籤指定(Assignment)與傳遞(Distribution)

● LSR 路由表的建立

在 MPLS 網路中所有的 LSR 利用繞徑協定(Routing protocol)來交換路由資訊，建立自己的 IP 路由表(Routing Table)，並根據路由表建立自己的轉送訊息資料庫(Forwarding Information Base；FIB)，此時的 FIB 中並沒有標籤的資訊。

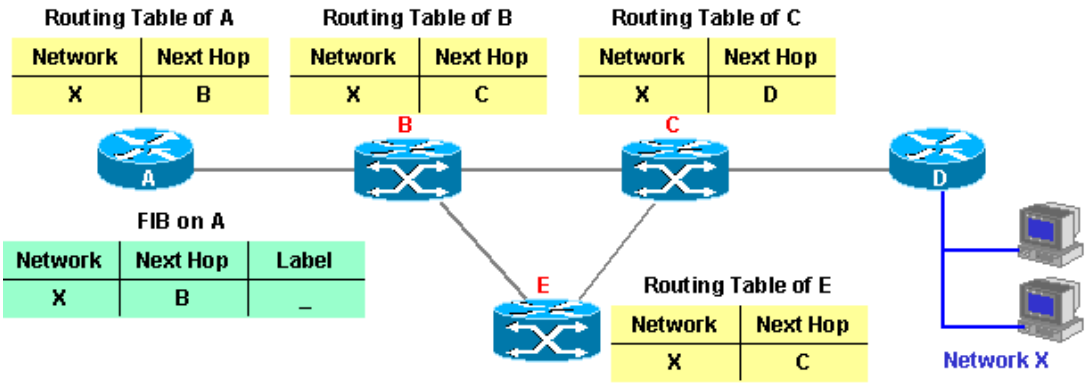

圖 5.3 LSR 路由表的建立

● LSR 分派路由表的過程

當 LSR 路由器開始啟動 MPLS 功能時,會根據由 IGP(如 RIP、OSPF)學來的路由表(Routing Table)內容，對於使用相同處理方式、相同路徑、到達相同目的地 IP 子網路的 Routing Entry 做彙整(aggregation)及分類後指定標籤(Label)。

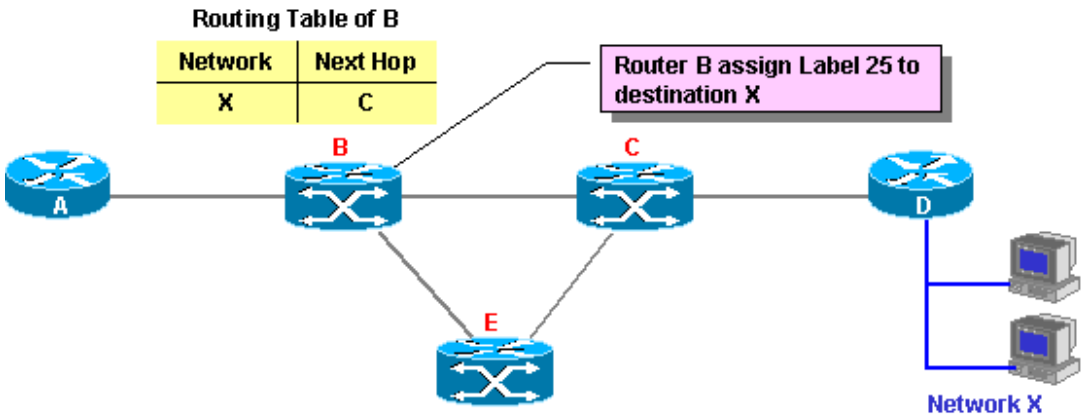

圖 5.4 LSR 分派路由表的過程

● LSR 初步建立自己的 LIB 及 LFIB

將前面步驟分派路由表後的本地標籤(Local Label)資訊儲存於 LIB 和 LFIB 中，此時的 LFIB 中只有本地標籤的資訊並沒有 Outgoing Label 的資訊。

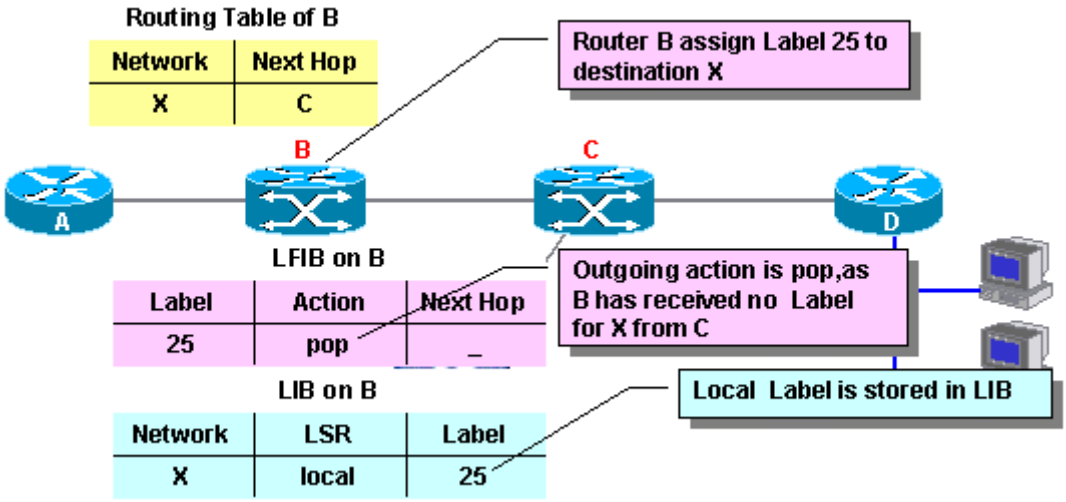

圖 5.5 LSR 初步建立自己的 LIB 及 LFIB

● LSR Label 傳遞的過程

LSR 將他本地端指定的標籤資訊傳遞(Distribution)給相鄰的 LSR，不論這相鄰的 LSR 是 Local LSR 的 downstream 或 upstream 都會傳送，而標籤傳遞靠的是相鄰的 LSR 間要執行標籤傳遞協定(Label Distribution Protocol；LDP)，來互相交換彼此的標籤資訊。另外談到 LDP 的特性， MPLS 設備會發送/接收 LDP 封包，LDP 封包透過發掘(Discovery)去和相鄰路由器(Neighbor)溝通對方是否有啟動 MPLS 及交換標籤資訊，而 LDP 是用 UDP 協定去發掘相鄰的路由器，並利用 TCP 協定去交換彼此標籤資訊。

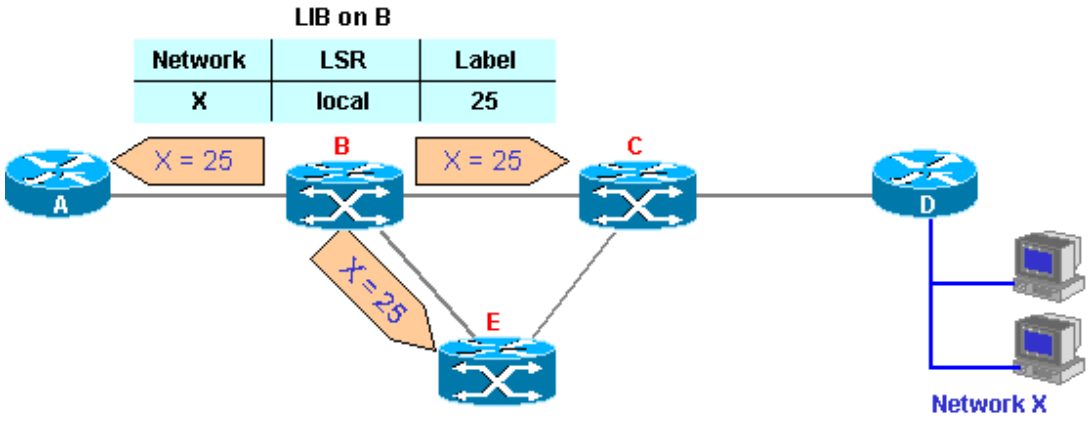

圖 5.6 LSR Label 傳遞的過程

● LSR 收到相鄰 LSR 送來的 Label 資訊作資訊的彙整過程

最後每個 LSR 根據接收到相鄰 LSR 送來的標籤資訊後，新增這些標籤資訊於自己的 LIB 中，並根據路由表(Routing Table)得到的最佳路徑，獲知到某網段的 Next-hop LSR 所送來的標籤資訊，插入到 LFIB 的 Outgoing Label 資料結構中。

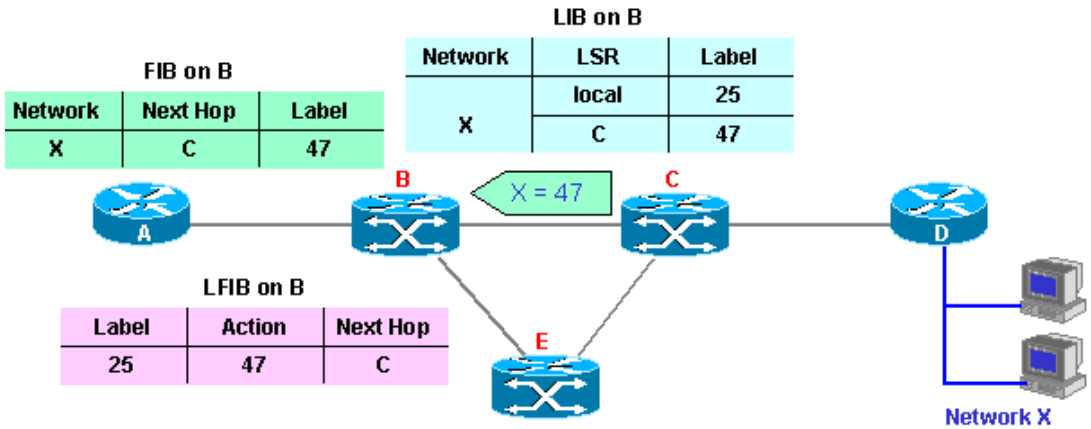

圖 5.7 LSR 收到相鄰 LSR 送來的 Label 資訊作資訊的彙整過程

5.2.3 Packet 在 MPLS 網路中傳送的過程

1. Ingress LSR(Router A)： IP Packet 進入 MPLS 網路的第一顆 LSR 路由器稱為 Ingress LSR，當 IP Packet 進入 Ingress LSR 首先會查看 Packet 中的 Destination IP address，並且在 FIB 中 lookup 是否有符合的 IP network，如果有則進一步查看 FIB 中相對應的 Label 欄位其值為何？(例如：IP =X ，Label=25)，當 Packet 從 Ingress LSR 送出時，會在此 Packet 中打上 Label=25 的標示，再傳送出去。

2. Core LSR (Router B)： 當帶有 Label=25 的 Packet 傳到 Router B 時，Router B 會查看(lookup)他的 LFIB 的資料，看看是否有 Inbound Label=25 的 entry，如果有則再查看此 entry 中 Outgoing Label 的欄位值為何？(例如 Outgoing Label=47)，所以 Packet 中的 Label 快速的被置換(Label=25 aLabel=47)並往下一個節點傳送出去。

3. Egress LSR(Router C)： 當帶有 Label=47 的 Packet 傳到 Router C 時，Router C 會查看(lookup)他的 LFIB 的資料，看看是否有 Inbound Label=47 的 entry，如果有則再查看此 entry 中 Outgoing Label 的欄位值為何？(例如 Outgoing Label=Pop)，所以 Packet 中的 Label 被移除，此時已離開 MPLS 網路再進入到 IP 的網路中，因此重新查看 Packet 中的 Destination IP address 為何？並查看其 FIB 以決定 Packet 要傳送的下一個節點。

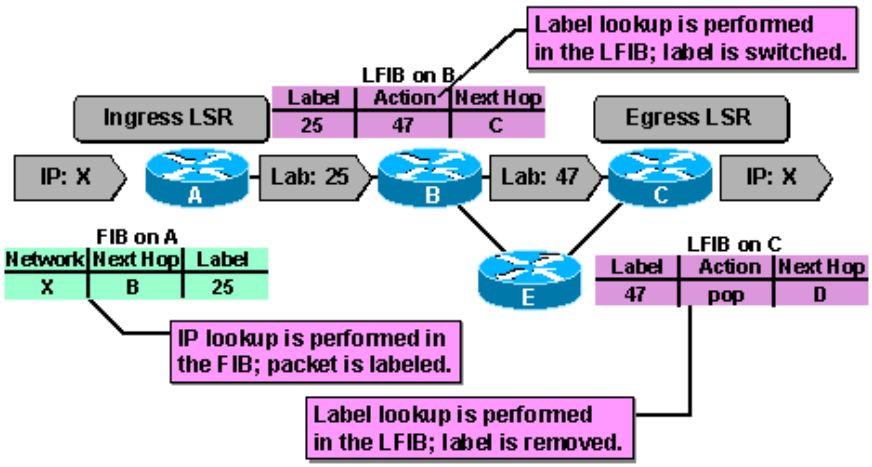

圖 5.8 Packet 在 MPLS 網路中傳送的過程

● 倒數第二個 Hop 移除(Penultimate Hop Popping)

由於 Egress LSR 不但要查看 LFIB 中的資料以便移除 Packet 中的 Label，而且還要查看 FIB 中的資料以決定將 Packet 往 IP 網路的下一個節點傳送，這樣的作法會使 Egress LSR 的負擔太重，而且對傳送有 Label 的封包也不是最有效的方式。所以解決的方式就是在原來 Egress LSR 前一個節點就把 Label 移除，最後一顆 Router 只要做 IP lookup 就好了，此種運作方式稱為 Penultimate Hop Popping。

5.2.4 標籤交換路徑(Label Switched Path；LSP)建立的方式

在 MPLS 網路裡頭，封包藉由標籤交換方式抵達目的地，期間所經過的路徑稱之為 LSP，其建立方式乃透過一標籤傳遞(Label Distribution)過程來完成，以下列出二種協定的組合來建立 LSP：

● 使用 LDP(Label Distribution Protocol)，支援 hop-by-hop routed MPLS。路徑建立方式將 LSR 使用傳統的繞徑協定(Routing Protocol)

建立其路由資料時，便同時進行 MPLS 的標籤傳遞工作。

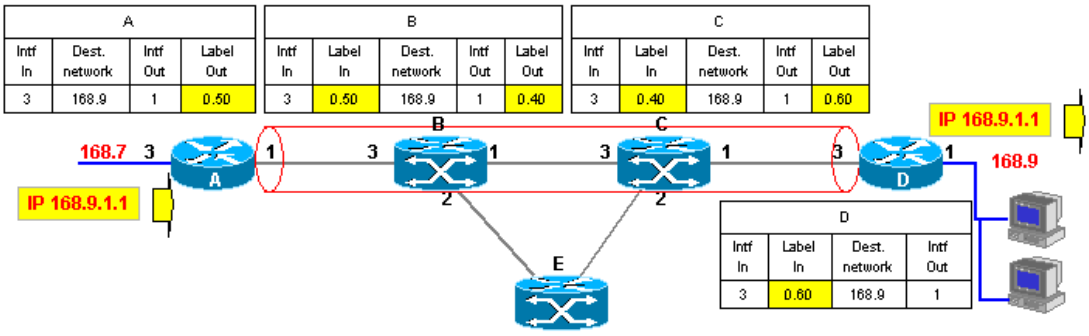

圖 5.9 標籤交換路徑 LSP(Label Switched Path)

● 使用 CR-LDP(Constraint-based Route；LDP)

CR-LDP 為使用 LDP 與其它延伸(some extensions)方式來建立 LSP，能夠建立一條符合所需頻寬的路徑，並能確保該路徑不會被其他 LSP 所侵占。 而 LSR 建立完成其路由資料之後，才進行 MPLS 的標頭指定與傳播工作。CR-LDP 實現方式包含：

1. 明確標籤繞徑 Explicit Routing：

在 CR-LDP 內 Explicit Route 也稱為 Constraint-based Route 或 CR-LSP，目的在建立一條符合所需頻寬的路徑。

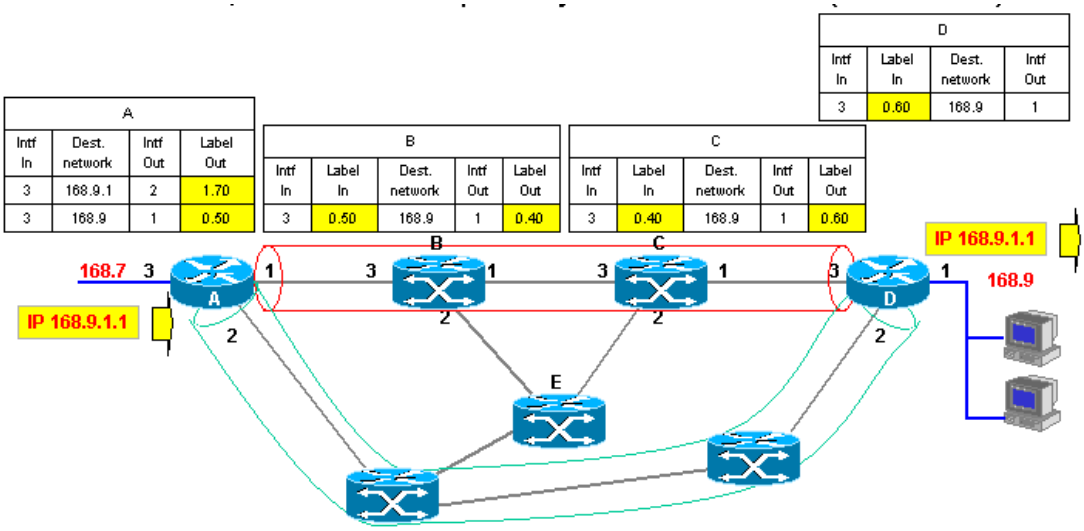

圖 5.10 明確標籤繞徑 Explicit Routing

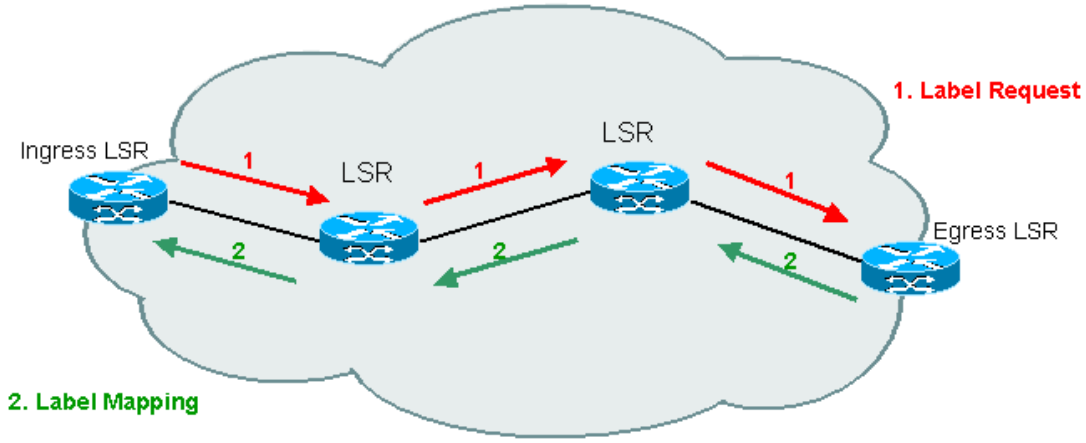

圖 5.11 CR-LDP(Constraint-based Route LDP)

2. 與 RSVP(Resource Reservation Protocol)組合來建立 LSP：

   CR-LDP 允許為明確標籤路由(Explicit Route)保留資源，所以在設定 CR-LDP 時，假設使用的頻寬不足而無法保留時，便會拒絕建立連線。

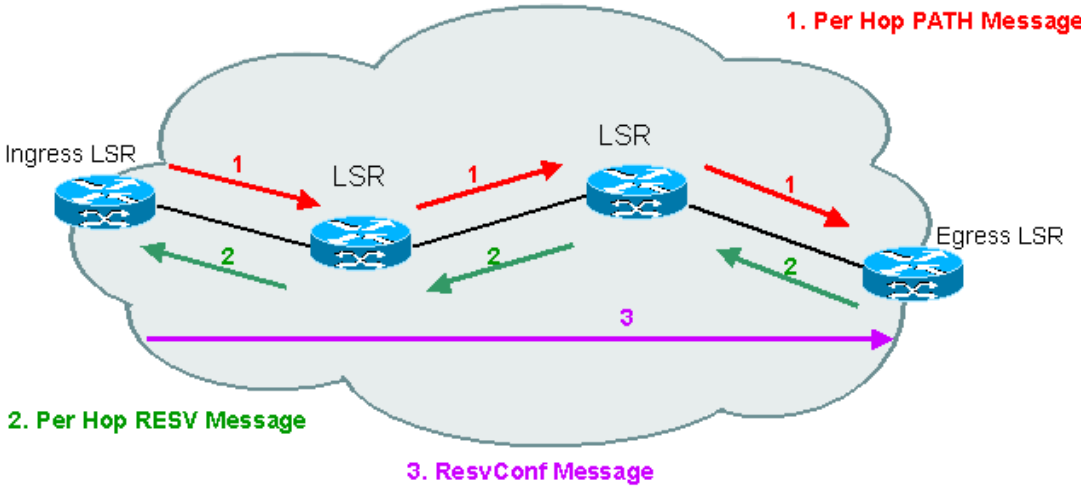

圖 5.12 RSVP(Resource Reservation Protocol)

3. 路徑搶奪與優先權(Path Preemption and Priority)：

當所剩的頻寬不足於要求時，除了拒絕建立該連線外，還可以搶奪其他連線的資源以達需求的頻寬，能夠搶奪的連線是依照設定的 set priority 與 holding priority 兩項參數而定，當 set priority 數值小於某條連線的 holding priority 時，表示能夠搶奪該連線的資源。

5.2.5 其他有關應用在 VPN 的解決方案

●客戶自主管理 VPN 解決方案(CPE-VPNs)(參考圖 5.13)

●Layer 3: IPSec

IPSec 為第三層的穿隧技術，專門為 IP 所設計加解密技術，不但符合現有 IPv4 的環境，同時也是 IPv6 與 IETF 所制定的業界標準。企業用戶透過網際網路來建構虛擬私有網路(IP-VPN)，使

企業透過網際網路來進行交易，以確保資訊傳送安全並節省通訊成本。而在 IP-VPN 之網路安全核心技術中之加解密技術，能將所有傳送的封包予以加密，直到封包抵達目的地再予以解密來確保資料傳輸的隱密性。整個 IPSec 協定包括了 Encryption、Authentication、封包分類、查詢及處理，是目前網際網路之網路安全中最完整、最成熟及最普遍的網路安全通訊協定。

●Layer 2: L2TP and PPTP

L2TP 與 PPTP 均為第二層的穿隧技術，適合具有 IP/IPX/Apple Talk 等多種協定的環境。

IPSec、L2TP、PPTP 三者，最大的不同點在於運用 IPsec 的技術，使用者可以同時使用 Internet 與 VPN 的功能。 而 L2TP 與 PPTP 建構在點對點協定(PPP)上，所以只能在 Internet 或 VPN 兩種功能中，選擇一種使用，也就是說，利用 L2TP 或 PPTP 執行穿隧傳輸時，就無法同時連結網際網路其他節點，使用起來比較不方便，所以較完整的虛擬私有網路整合性設備（VSU）均採用 IPSec 協定。

●網路供應商代管 VPN 解決方案(PP-VPNs) (參考圖 5.13)

●Layer 3: MPLS-Based VPNs (RFC 2547bis)

下一節再詳細介紹

●Layer 3: Non-MPLS-Based VPNs (Virtual Routers)

服務供應商及最終用戶可享有在其 VPN 內最全面的透明度就是讓客戶將 CPE 管理的路由器功能整合至邊緣平台上虛擬路由

器，以減少客戶維護網路的成本。

●Layer 2: VPLS VPNs

當企業客戶考慮同一個 Site 有多點要同時接取 VPN 時會有大量的 Multicast/broadcast 的問題，就可以採用 VPLS VPN 技術, 以 MPLS 建立 LSP Tunnel 方式傳送 Layer 2 封包，以 Muti-point Ethernet 路由，輔以 MPLS 標籤堆疊做為訊務區隔，核心走 MP-BGP 使企業能輕易管理 Multicast/broadcast 的資訊流。

●Layer 2: ATM and Frame Relay

企業客戶採用 ATM 技術建置 VPN，可將所欲傳送之大型資料，分割成 53Byte 為一細包(Cell)的基本單位，可與既有訊框傳送(Frame Relay)網路之客戶整合，達到快速交換傳送之目的。又可整合語音、數據及視訊等資訊在同一網路，同時針對各種資訊型態提供最佳的傳輸環境，提供最有效率的多媒體服務。而且利用 ADSL 供裝的中華電信 ADSL 業務，配合一般數據專線供裝方式架設 HiLink VPN 網路，並配合動態的頻寬管理，提高網路頻寬的利用率及降低網路成本。

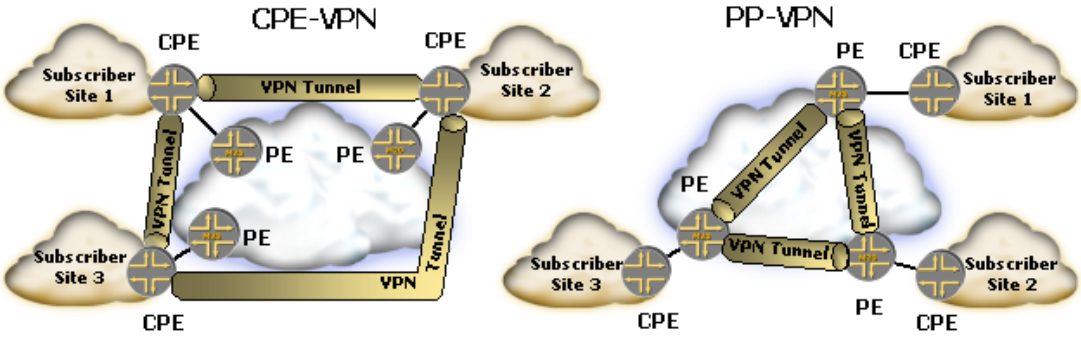

圖 5.13 客戶自主管理 CPE-VPN、網路供應商代管 PP-VPN 之比較

5.2.6 Layer 3 MPLS VPN

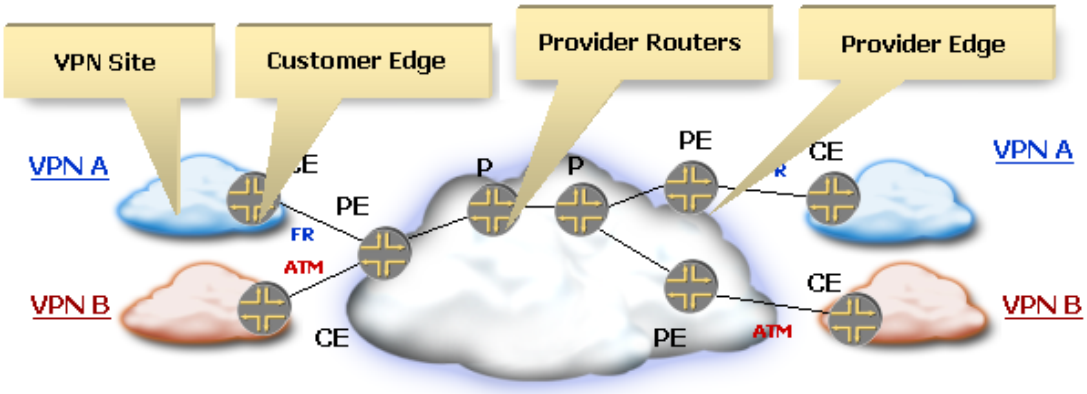

圖 5.14 Layer 3 MPLS VPN

●VPN 是由一個區域共享相同的路由資訊所組成的集合體。
●VPN 可視為封閉使用者的公共領域獨享共同的利益。
●為了讓 PE 路由器能區分出是哪一個 VPN 用戶的路由，在 PE 路由器上建立大量的虛擬路由器,每個虛擬路由器都有各自的路由表和轉送表，這些 IP 路由表和 IP 轉送表(又稱 CEF 表)統稱為 (VPN Routing and Forwarding instances；VRF) 。
●由於 VRF 定義了連到 PE 路由器上的 VPN 成員,使用 VRF 的 IP 路由表和 IP 轉送表處理 PE 介面的路由協定與路由導入(import)導出(outport)規則，使得虛擬路由器能隔離不同的 VPN 用戶之間的路由,因此能解決不同 VPN 之間 IP 位址重疊的問題。
●如果兩個以上的 VPN 共享一個領域，IP 位址配置在 VPN 裡頭必須是唯一的。

● 路由轉送實例(VRF)、路由區分器(RD)、路由目標(RT)

●路由轉送實例(VPN Routing and Forwarding instances; VRF)

連接到 PE 路由器上的每個點(Site)都分配一個 VRF 來存放 VPN 路由，但連接在同一 PE 路由器上的點，如果滿足以下三個條件，則可以共享一個 VRF：(1)各點屬於同一個 VPN(2)路由資訊相同(3)每個點之間允許相互直接通信。通常 PE 路由器上每個用戶端與一個特定的 VRF 相關聯，從該用戶端輸入的 VPN 分組將根據各自對應的 VRF 找出其 VPN 標籤和下一個 Hop 的 PE 路由器位址。因此，VRF 隔離了不同的 VPN。

●路由區分器(Route Distinguisher ;RD)

允許不同的 VPN 客戶使用相同的 IPv4 位址，對於不同 VPN 中相同的位址，採用 RD 可以實現 IPv4 位址的重覆使用。PE 路由器透過 MP-iBGP 協定通報各點的路由時，將同時攜帶各路由的 RD，即將 IPv4 位址轉化為 VPN-IPv4 位址，PE 路由器在收到 MP-iBGP 通報的路由後，將找出該路由的 RD，然後將 VPN-IPv4 位址轉化成 IPv4 位址，即將位址中的 RD 去掉，將其導入到相應的 VRF 中。由於不同 VPN 被 VRF 隔離了，因此不同 VPN 中相同的 IPv4 位址，將被導入到不同的 VRF 中。在同一個 VPN 中，位址必須是唯一的，當多個 VPN 之間需要相互通訊時，則會要求位址必須在多個 VPN 中是唯一的，此時多個 VPN 只能使用同一個 RD。

●路由目標(RouteTarget ;RT)

當每個 VPN 路由要從 VRF 匯入或匯出給其他的 VRF 時，必須貼上一個或多個路由目標的標籤。因此，RT 是用來控制 VRF 的匯入和匯出的策略，以構成各種複雜的 VPN 拓撲。一個 VPN 有可能不止使用一個 RT，RT 的實際的運作與 VPN 的拓撲架構有密切關係，對於全網狀相連的 VPN 可以使用一個 RT，對於非全網狀相連的 VPN，一個 VPN 往往需要多個 RT 配合。當從 PE 匯出 VPN 路由時，會用 RT 對 VPN 路由進行標記，而在 VRF 中匯入路由時，會使用到多個 RT，只要有一個 VPN 路由中攜帶的 RT 與匯入路由中的任意 RT 相同，都將被匯入到該 VRF 中。

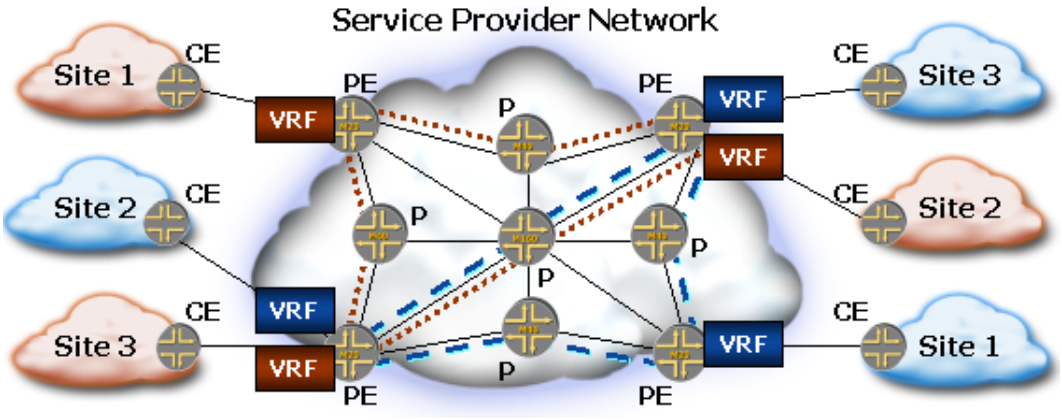

圖 5.15 Layer 3 MPLS VPN Service Provider Network

● **邊界閘道協定(Border Gateway Protocol；BGP)**

選擇 BGP 做為 VPN 路由的繞送協定(如圖 5.16 所示)，主要是由於：

1. MPLS VPN 網路的 VPN 路由數量可能會變得越來越龐大，BGP 是唯一能支援大量路由的繞送協定。

2. BGP 功能可讓 VPN 路由資訊可以在埠相連接的路由器之間交換資訊，不必進入 ISP 的核心網路。

3. BGP 可以傳送依附在路由上的任何資訊，使得 PE 路由器之間傳播路由目標變得非常簡單，所以在 PE 路由器能夠選擇路由之前，他必須知道有哪些 VPN 路由存在，才能由 BGP 來對它們進行比較。

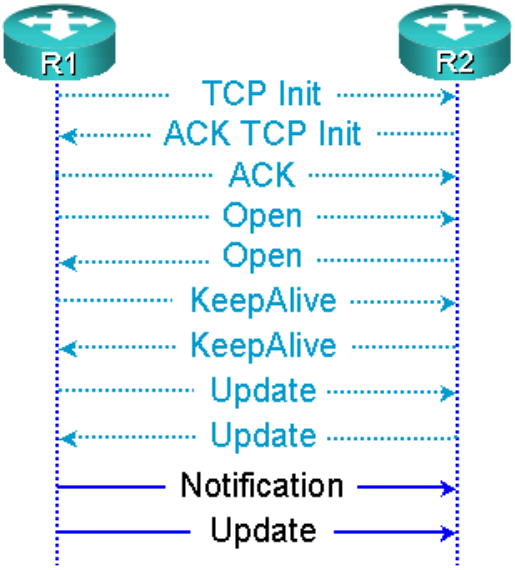

圖 5.16 BGP 協定信號方式

● MPLS VPN 路由信號交換方式

藉由個別 VPN 的 OSPF 程序，PE 路由器從各 CE 端收集路由資訊後，重新分發給多重協定 MP-BGP。 這些路徑在重新分送傳遞時會在 VPN 位址中加上路由識別碼，也會附加 VRF 中所設定的匯出(outport)路由目標，所產生的資訊會由 MP-BGP 傳播給目標路由器(RD) 。

該步驟如下：

1. 執行各個 VRF 的繞送協定，以便從 VPN CE 端收集路由資訊。
2. 重新分送傳遞 VRF 中的路由資訊到 MP-BGP 中，並透過骨幹網路傳播。
3. 由 MP-BGP 中所接收的路由資訊將被插入 VRF 中，並選擇性地轉送給 CE 路由器。

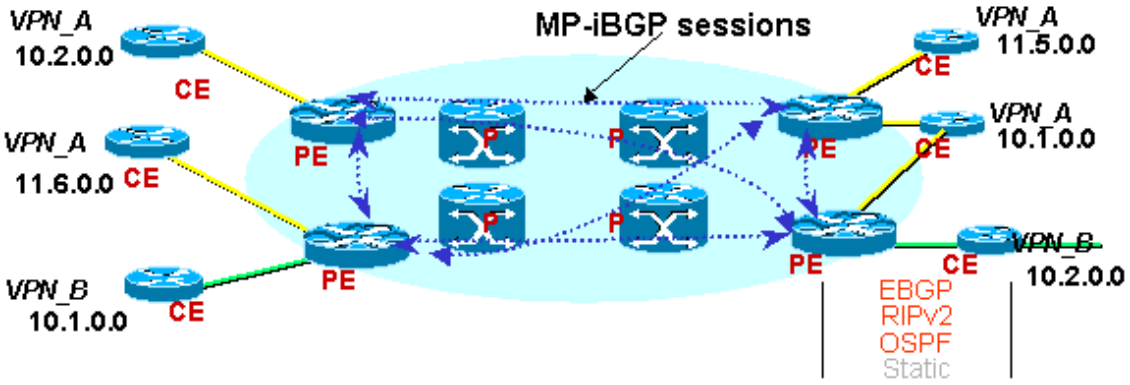

圖 5.16 Layer 3 MPLS VPN 路由信號交換方式

● P 與 PE 路由器之間必須使用共同的 IGP 交換路由訊息。

●對 MP-BGP 來說 PE 路由器的彼此界接會是全網狀(fully mesh)。

●PE 路由器對於所面對的 CE 路由器必須負責透過 MP-BGP 傳遞 VPN 資訊(VPN-IPv4 addresses、Extended Community、Label)給其他 PE 路由器。

● P 路由器不執行 BGP 當然也不處理任何 VPN 有關的訊息。

● PE 與 CE 路由器之間是透過 EBGP、 RIPv2、 OSPF、 Static

routing 來彼此交換路由資訊。
●對 CE 路由器而言，他只執行一般標準的路由功能。

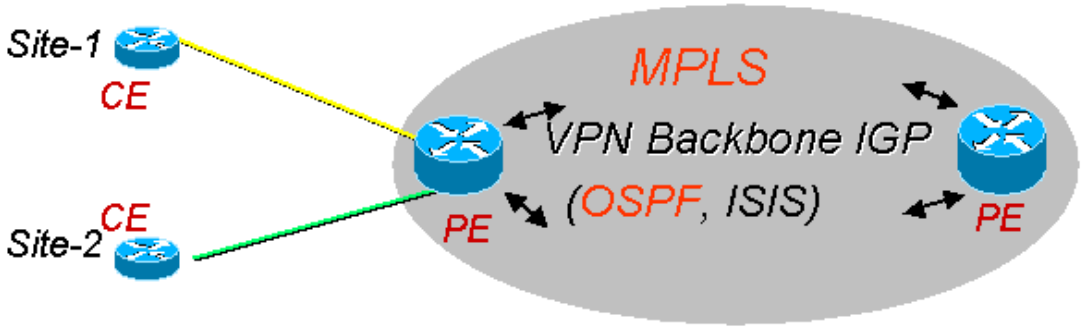

圖 5.17 Layer 3 MPLS VPN 路由信號交換方式

● PE 路由器彼此維護不同的路由表(routing tables)。
● 所有具備 PE 及 P 路由的總體路由表(global routing table)是透過 VPN 骨幹的 IGP (ISIS or OSPF) 來存放的。
● VRF (VPN Routing and Forwarding) 路由與轉送表直接關係著許多個 CE 點(site)。
● VRF 跟 (sub/virtual/tunnel)是直接界接。
● 如果相連的點(site)共享相同路由資訊,那麼與介面之間也會共同享有相同 VRF。
●同屬一個 VPN 的 Sites 會享有相同的 VRF。
●路由若是 PE 透過 CE 路由器接收到的，會被放在 VRF 適當的位置。
●路由若是 PE 透過 backbone IGP 接收到的，會被放在總體路由表 (global routing table) 。
●VPN 之間因為分開使用不同的 VRF， 所以 IP 位址的配置不一定要唯一。

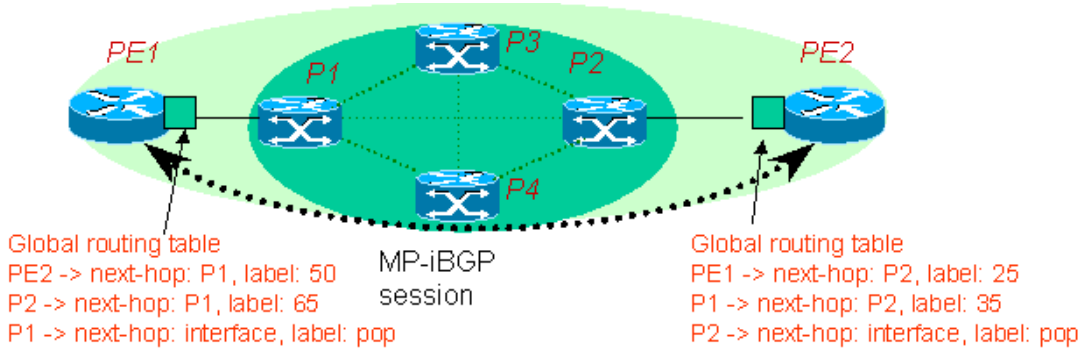

圖 5.18 Layer 3 MPLS VPN 路由信號交換方式

●所有 P 與 PE 路由器之間皆會採內部閘道協定(Interior Gateway Protocol；IGP)與標籤傳遞協定(Label Distribution Protocol；LDP)相互溝通。

●每一個 P 與 PE 路由器都保有所有骨幹節點的路由而且每個標籤會對應到每個路由。

●PE 可以同時處理多個 VRF 路由表。

●每個 VRF 都會保存企業客戶的路由。

●企業客戶 IP 位址可以重疊 (overlap)使用，不會產生衝突。

●VPN 之間是被隔離的。

●MP-BGP 是用來傳播 PE 路由器間的 IP 位址。

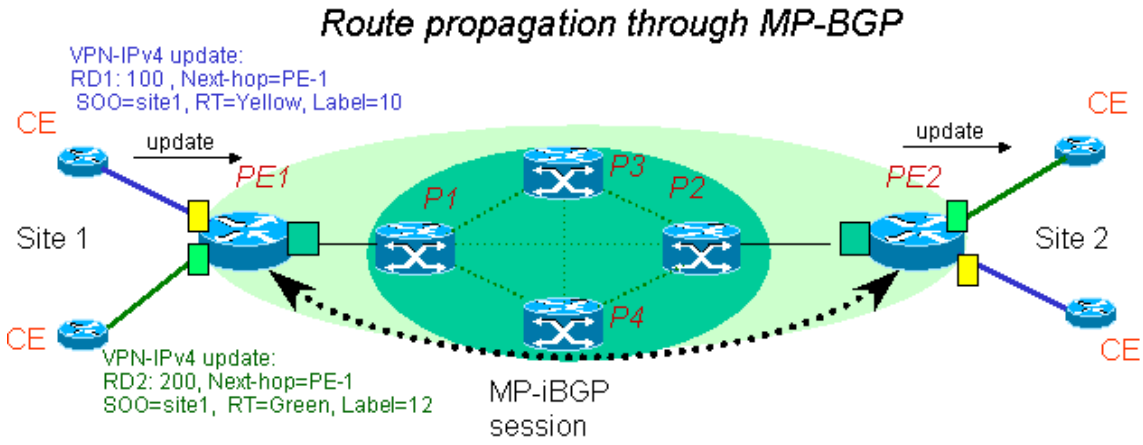

圖 5.19 Layer 3 MPLS VPN 路由信號交換方式

●VPN-IPv4 更新封包可以解讀為 IPv4 位址與對應 RT 值後要加入 VRF 的位址。
●MP-BGP 指定 RD 到每個路由使 RD 看起來像是唯一的，以便整個傳播作業。
●MP-BGP 指定路由目標(Route-Target) 讓遠端的 PE 路由器加入相關的路由到相對應的 VRF 路由表。
●VPN-IPV4 位址是由 Route Distinguisher 64 bits ＋ IPv4 address 32bits 組成。
●可包含 Extended Community attribute (64 bits) 。

● MPLS VPN 資料流傳遞方式

詳細情形如下說明：
●在總體路由表(global tables)，PE 路由器會儲存 IGP 路由與相關的標籤 Label，而標籤的傳遞是透過 LDP 來進行。
●在 VRF 裡頭，PE 路由器儲存 VPN 路由與相關的標籤，而標籤的傳遞是透過 MP-BGP。
●Ingress PE 只單純的從 CE 路由器接收一般的 IP 封包。
●PE2 路由器從 VRF 得到 IP 最遠的配對資訊(IP Longest Match)後，利用 iBGP 找到下一個 hop PE1，然後再另外又加上一層標籤後變成： 外部標籤 IGP (Ti) ＋ 內部標籤 VPN (Tv) 之標籤堆疊封包，所以 P 路由器之間與 PE 路由器之間 IGP 的收斂與標籤的傳送是完全獨立的。
●接下來的 P 路由器只單純地將帶有 IGP 標籤的封包作交換。
●Egress PE 路由器會移除 VPN 標籤。

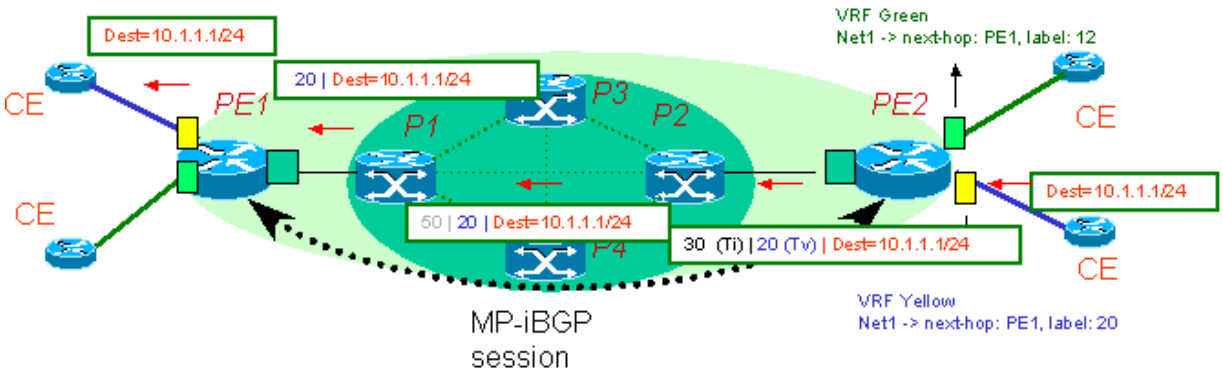

圖 5.20 Layer 3 MPLS VPN 資料流傳遞方式

●當不同的 VPN 封包流向同一個 PE 路由器時，PE 讀取內部標籤值(Inner Label)- VPN (Tv)以決定往哪個 CE 走。也就是說，封包從 PE 路由器轉送到 CE 路由器只根據內部標籤(Inner Label) - VPN (Tv)而非 VPN 的 IP 位址，而這些 IP 位址都是可允許重複。

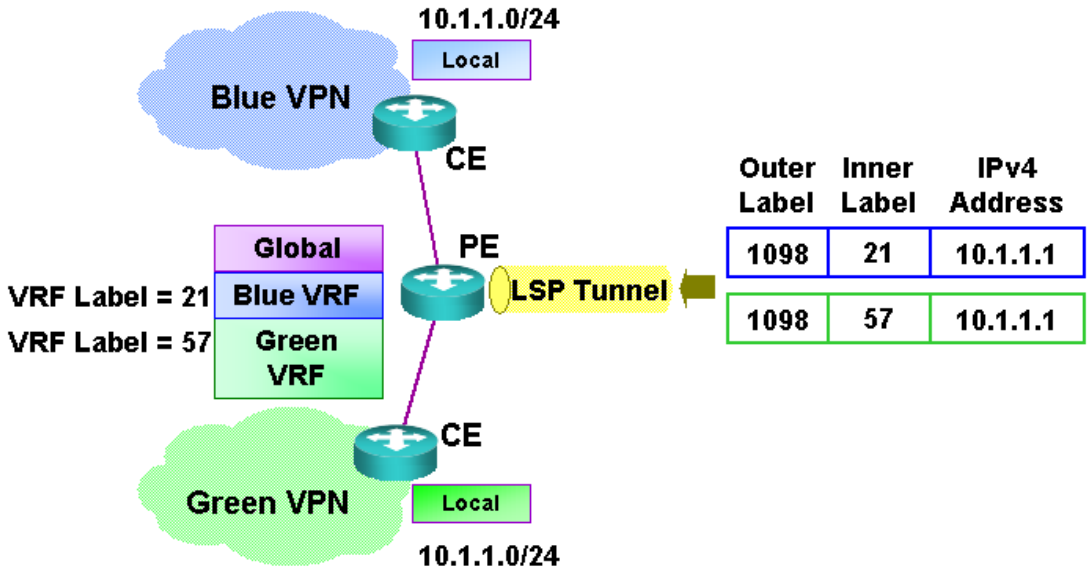

圖 5.21 Layer 3 MPLS VPN 資料流傳遞方式

● MPLS VPN 封包轉送方式-倒數第二個 Hop 的移除

步驟 1 : PE2 收到被當地 VRF 查詢註記的 IP 封包，找到標籤與攜帶 Next-Hop 的 BGP 路由，藉由標籤所對應的 IGP 路由便可抵達 BGP

下一個 hop (PE1) 。

步驟 2： P 路由器再根據外部標籤 IGP (Ti)來交換封包。

步驟 3： 倒數第二個 Hop 的標籤移除(Penultimate Hop Popping) 。

P1 對於 BGP 的下一個 hop 而言，為所謂倒數第二個 Hop 。

P1 移除最上層的標籤。

由 PE1 透過 LDP 向 P1 主動提出要求。

步驟 4： PE1 收到的封包含有對應 outgoing interface (VRF) 的標籤時，執行封包標籤解析(lookup)與標籤移除動作，然後最後把產生的 IP 封包傳送到相鄰的 CE 路由器。

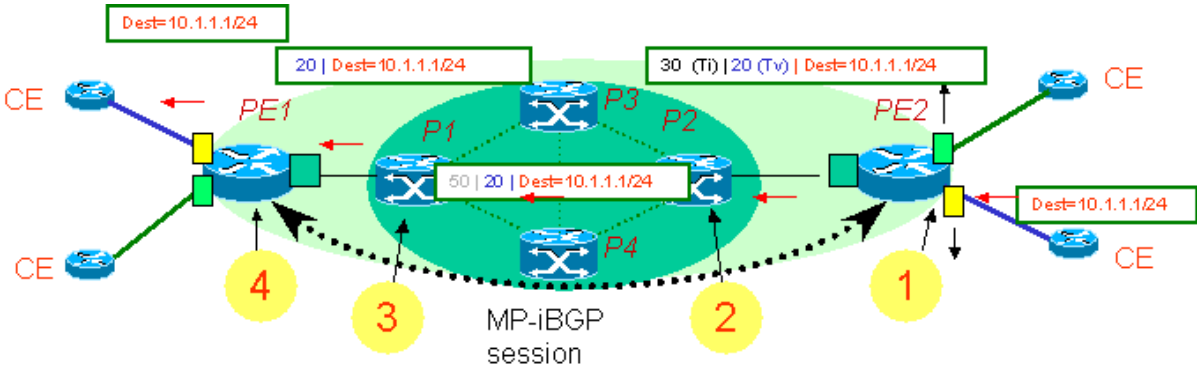

圖 5.22 Layer 3 MPLS VPN 倒數第二個 Hop 的移除

● MPLS VPN 與 IP VPN 之比較

以 MPLS VPN 與傳統的 IP VPN 相比較，最大的差異是 VPN 通道建立的不同，傳統的 IP VPN 是由用戶的 CPE 設備來產生 VPN 通道，與電信業者無關，而 MPLS VPN 則是由電信業者的設備來產生 VPN 通道。在骨幹網路部分，由於 IP VPN 並不能給電信廠商產生利潤，傳統 IP VPN 通常是利用與一般用戶相同的網際網路來進行，缺乏 QoS 機制。但 MPLS VPN 大多以專屬網路為主，相對地，也要進行 MPLS TE 或是 MPLS QoS 的機制。傳統 IP VPN 成本低廉但缺乏服務品質保證的缺點，因此在 VPN 技術的應用上，MPLS

VPN 逐漸受到青睞。MPLS VPN 與傳統 IP VPN 的比較如下表：

MPLS VPN 與傳統 IP VPN 的比較			
項目		MPLS VPN	IP VPN
VPN 數	可提供	2^{64}	250
	已提供	442	84
目前用戶數		7825	886
使用設備		Router	ATM Switch
使用技術		MPLS	VR(Virtual Router)
用戶介面		ADSL、FR T1/E1、ATM T1/E1/T3/STM-1、Ethernet 50M~1G	無 FR 及 Ethernet 介面
服務品質		可提供 SLA (Service Level Agreements)，並可依客戶需求（如影像、語音、資料等）提供不同等級服務	缺乏 QoS 機制
骨幹網路		通常為網路服務供應商所建置的核心網路	利用公共網路建立虛擬通道，進行資料傳遞
網路功能	可提供語音、視訊及 Data 不同等級之 QoS 服務	有	無
	同時可提供 Internet、Intranet、Extranet	有	無
	可提供兩岸 MPLS IP VPN	有	無
加值服務	網路電話 (VoIP)	有	無
	視訊會議	有	無
	遠距教學	有	無
	遠端管理	有	無
	異地備援	有	無
	電路備援	有	無

表 5.2 MPLS MPLS VPN 與傳統 IP VPN 的比較

5.2.7 Layer 3 MPLS VPN 服務供應商之間的介接

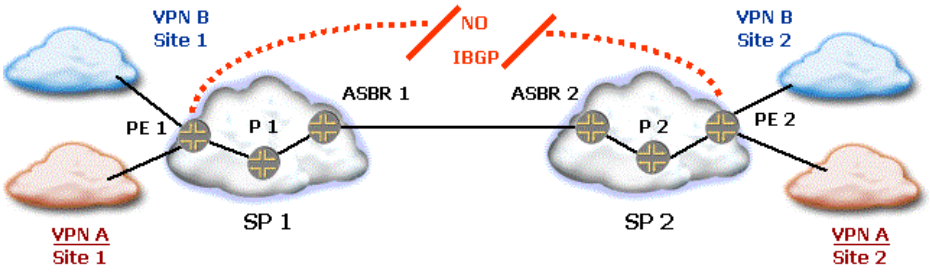

圖 5.23 Layer 3 MPLS VPN 服務供應商之間的介接

困難點:

如上圖 5.23 所示，MPLS VPN 服務供應商視為一個自治系統(AS)，而自治系統之間的 PE 採 IBGP 不能形成網狀(mesh)，又必須透過 EBGP 傳送 VPN-IPv4 路由。

解決方案:

自治系統邊緣路由器 ASBR1 與 ASBR2 採 VRF-to-VRF 直接連線方式，讓 EBGP 重新分佈於自治系統(AS)間，收斂後形成 Multi-hop EBGP，使兩個自治系統看起來像一個自治系統。

● **範例**

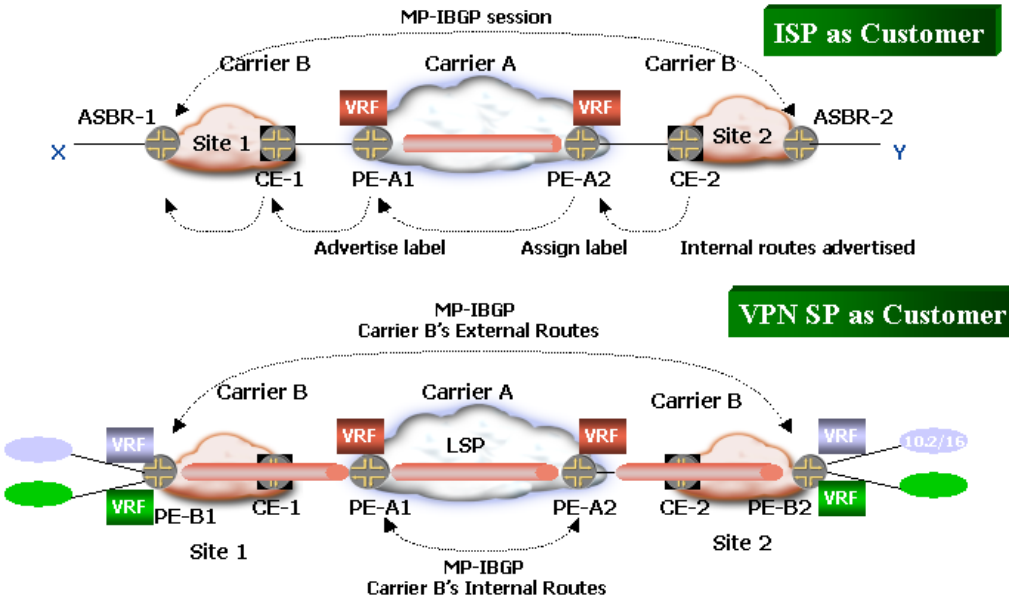

圖 5.24 Layer 3 MPLS VPN 服務供應商之間的介接

- 服務供應商提供服務的對象為服務供應商時(如圖 5.24 所示)

  服務提供商能夠允許其他服務提供商以透通的方式使用它的骨幹網路。

- 網際網路供應商(ISP)為企業客戶時

  當服務供應商 B 的客戶走的是 IPv4 服務時, 其外部路由必然是 IPv4，服務供應商 A 透過服務供應商 B 所管轄的用戶邊緣路由器 ASBR1 與 ASBR2，指派標籤建立 LSP Tunnel 方式，提供 VRF-to-VRF 直接連接 MPLS 骨幹服務，使得服務供應商 B 所管轄的不同據點(site)能透過內部路由的 MP-IBGP 協定交換路由，當內部路由數量遠小於外部路由數量時，能有效管制 ISP 客戶與服務供應商 B 之間路由交換的頻率，就服務供應商 B 而言,可以減少對 PE 路由器維護的數量。因此可以提升 ISP 效能、降低骨幹網路維護的成本、使網路的擴充性與延展性大大的提高。

- MPLS VPN 網路供應商為企業客戶時

  資料傳送方式有別於 ISP,當服務供應商 B 的客戶走的是 MPLS VPN 服務時，其外部路由必然是 VPN-IPv4 與服務供應商 B 的 MP-IBGP 協定，使得內部路由也走服務供應商 B 的 MP-IBGP 協定，這樣對服務供應商 B 而言，無須由服務供應商 A 指派標籤建立 LSP Tunnel，因為 MP-IBGP Session 可以透過 LDP 動態指派標籤，大幅度提高服務供應商 B 網路的擴充能力。

5.3 MPLS VPN 的服務品質(QoS)

MPLS VPN QoS 機制啟動後 IP 封包在透過 CE 路由器進入 MPLS VPN 網路時(如圖 5.25 所示)，會將 IP Precedence 複製到 MPLS 標籤中的 EXP，並根據規則進行對照，做為 Queuing 或是 Discard 的依據。

MPLS VPN QoS 的管理要針對網路的品質狀態及可保留的資源隨時調節 QoS 策略與相關的管理措施。但是，QoS 管理必須能隨時獲知網路的品質狀態，並能精確預測網路的品質，才有能力適時調整管理措施。網路規模越大，該項工作越不容易做好，所以 QoS 管理必須對靜態與動態 QoS 管理做個取捨，一般來說，上層的管理者需要用到較多的靜態管理。下層管理者因管理範圍較小，也較容易預測，可以採用較為動態的管理方式。

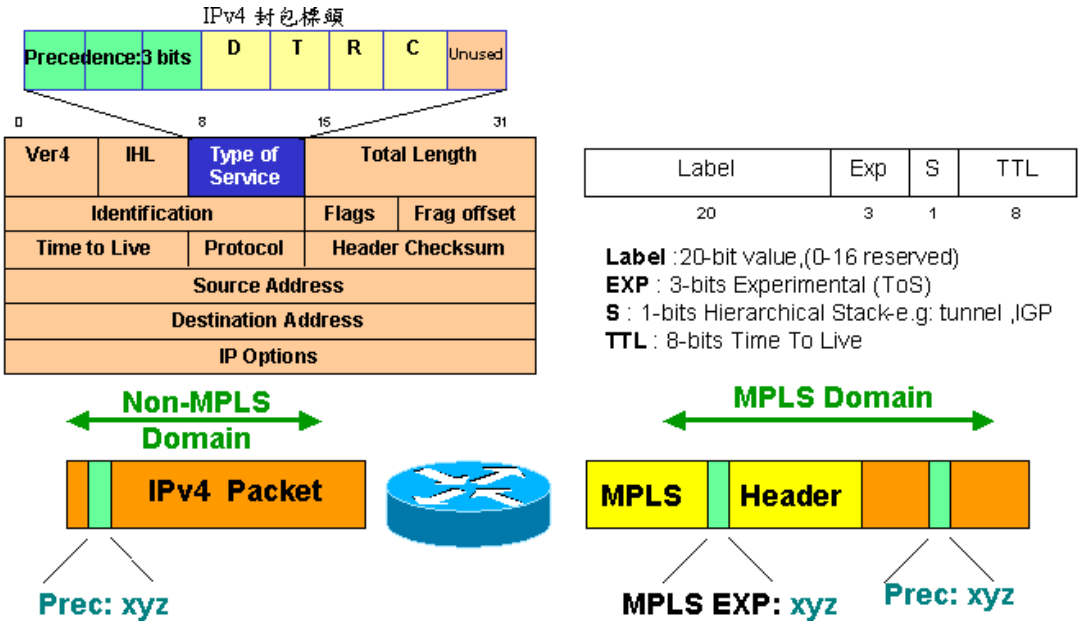

圖 5.25 Layer 3 MPLS VPN 的服務品質(QoS)

5.3.1 MPLS VPN 邊緣路由器與核心網路之 QoS 管理

● **邊緣路由器到核心網路**

方法一： 以優先頻寬保證(Committed Access Rate；CAR)方式在進入 MPLS 網路前的邊緣路由器上對所有進入的 IP 訊務進行管制，並針對訊務型態與策略設定 IP 標頭中優先權欄位值,同時將此 IP 封包標頭內的 IP Precedence 複製到 MPLS 標籤內的 EXP(CoS)欄位上。

方法二： 以優先頻寬保證(CAR)方式在進入 MPLS 網路前的邊緣路由器上對所有進入的 IP 訊務進行管制，並針對訊務型態與合約設定 MPLS 標籤內的 EXP 欄位，與方法一不同的是 IP 封包標頭內 IP Precedence 值並不會被改變。

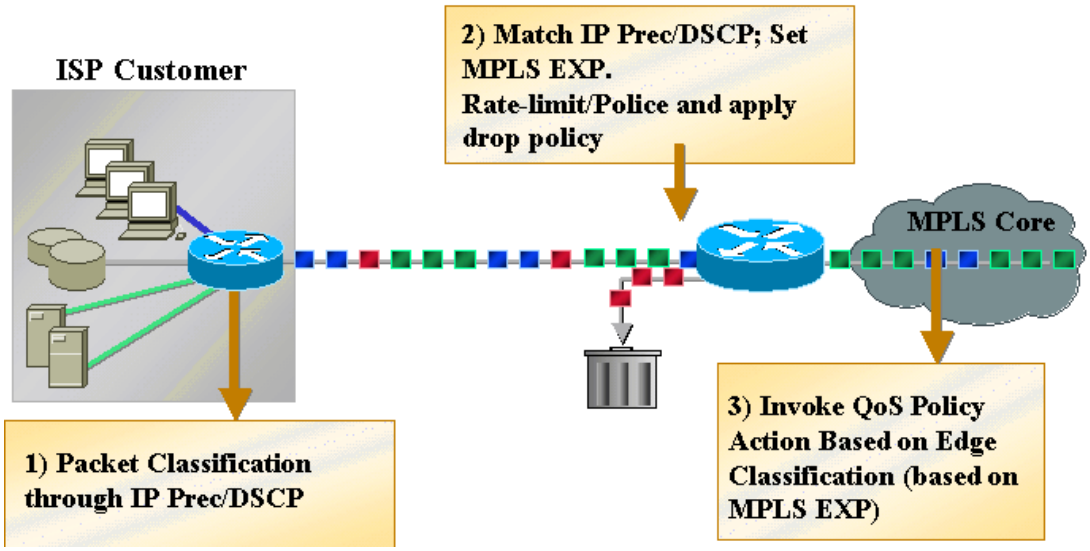

圖 5.26 Layer 3 MPLS VPN 的服務品質(QoS)

● **核心網路**

在 MPLS 骨幹網路上，訊務的區別與分類是依據 MPLS EXP 欄位值，其使用的機制為 IP QOS 中優先權排序(Queuing)與拋棄(Discard)程序(如 Weighted Fair Queuing；WFQ 、Weight Random Error Detector； WRED、Weight Round Robin；WRR 等等) 。

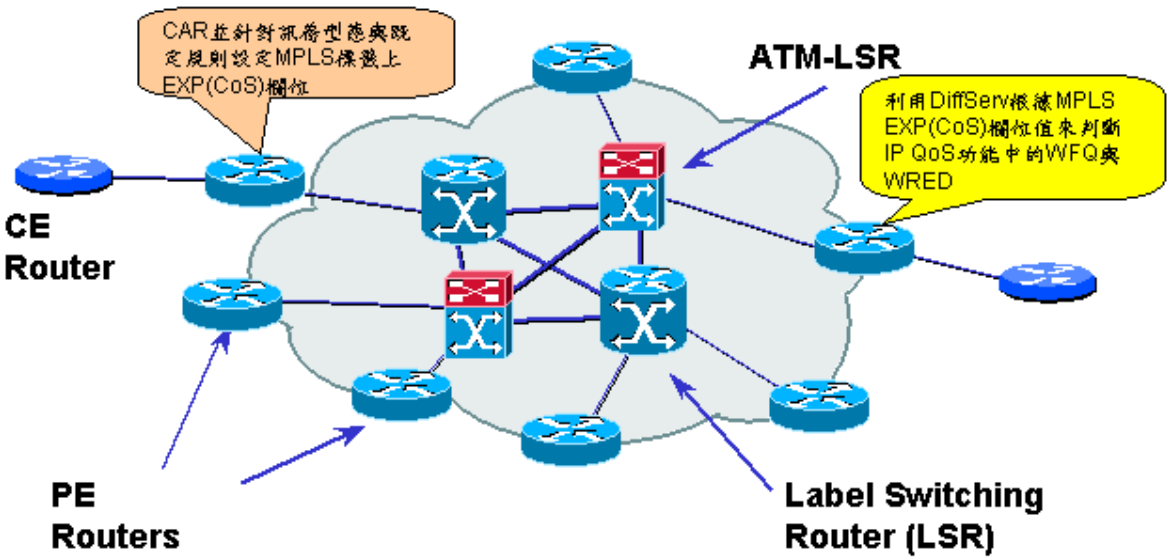

圖 5.27 Layer 3 MPLS VPN 的服務品質(QoS)

● MPLS VPN DiffServ 服務分級 Class of Service 範例

目的是允許服務供應商提供各種等級的網路服務，對企業客戶而言，如果是事先偵測出所謂敏感資料並即時分類後，再依既定優先權進行傳送作業，便可以將企業客戶數據、語音、多媒體資料流收斂整合起來，以達到頻寬有效的利用，可增加服務供應商額外服務的營收與提高客戶信賴度，接下分別舉兩個範例來作說明。

範例一： 分級方式是根據各種服務所事先定義的緩衝區大小與加權循環(Weighted Round Robin；WRR)空間分佈之規則，來對網路封包進行優先權排序與拋棄程序。

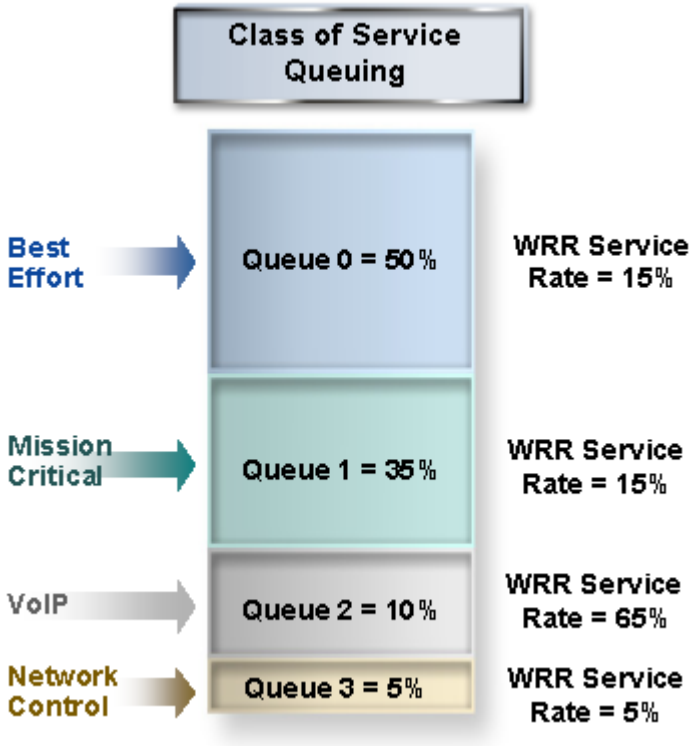

圖 5.28 範例一

範例二： 分級方式是根據各種服務所事先定義的緩衝區大小與頻寬使用率規則，來對網路封包進行優先權排序與拋棄程序。

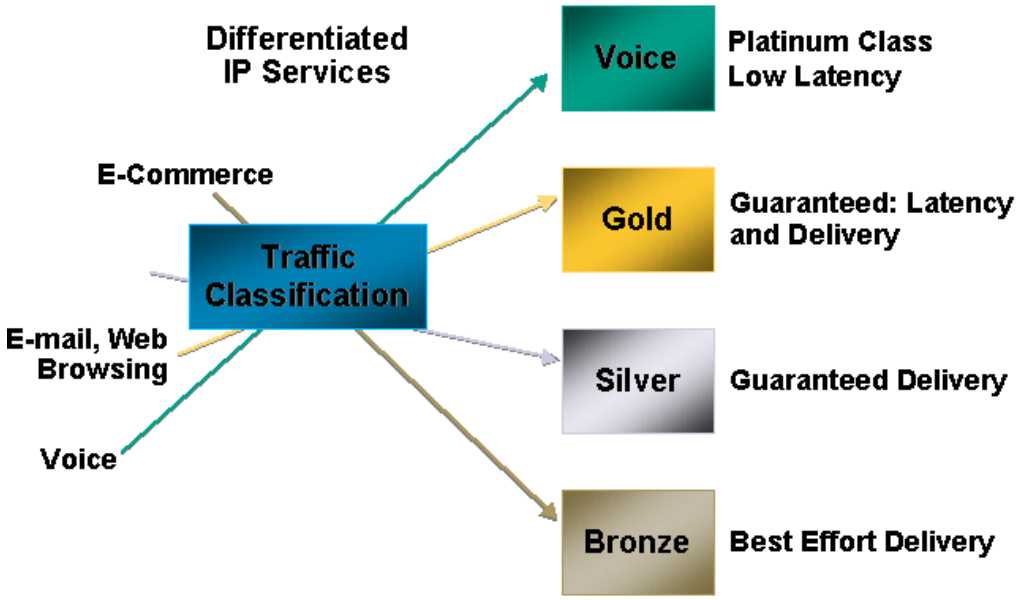

圖 5.29 範例二

5.3.2 MPLS VPN DiffServ 與 IntServ 的整合

- **E-LSP**

建立 E-LSP 必須透過標籤綁定協定(LDP or RSVP)。
假設 E-LSP 支援 EF 與 AF1 服務等級，在單獨一條 LSP (單一標籤)內移動的 EF 與 AF1 封包，會根據不同 EXP 值選擇佇列規則來進行隊伍排列。

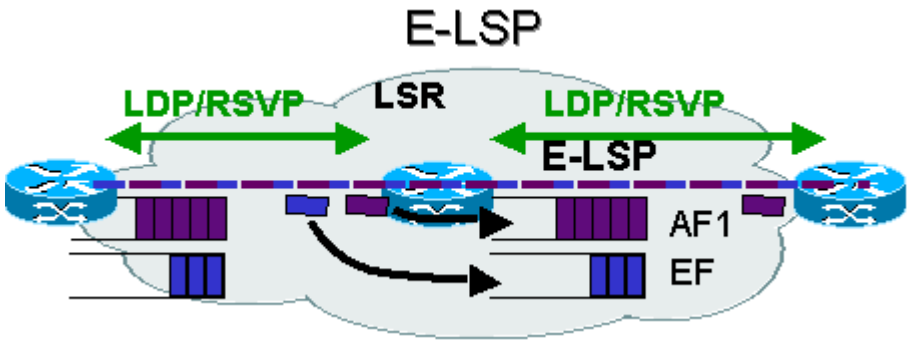

圖 5.30 E-LSP

- **L-LSP**

建立 L-LSP 必須透過標籤綁定協定(LDP or RSVP)。
假設分開的兩條的 L-LSP 都支援 EF 與 AF1 服務等級，
在分開不同的兩條 L-LSP 內移動的 EF 與 AF1 封包，會根據不同標籤值選擇佇列規則來進行隊伍排列，並根據 EXP 值來進行丟棄動作。

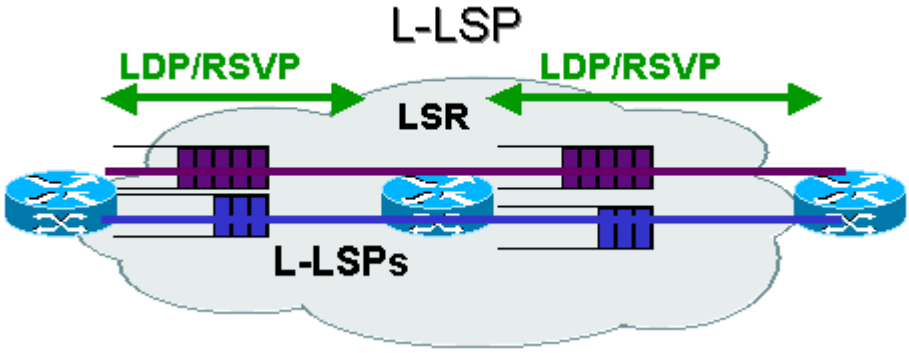

圖 5.31 E-LSP

● **MPLS VPN 整合作業與平行作業**

整合作業以 DiffServ 網路為骨幹架構，而 IntServ 網路為接取網路，兩網路銜接處的路由器將根據網路負荷來決定是否接受 RSVP 的請求。

平行作業則為在同一網路上同時提供整合式服務以及差異式服務，如果此網路沒有支援訊號方式的策略，就一律必須採用 DiffServ 網路來提供服務品質。

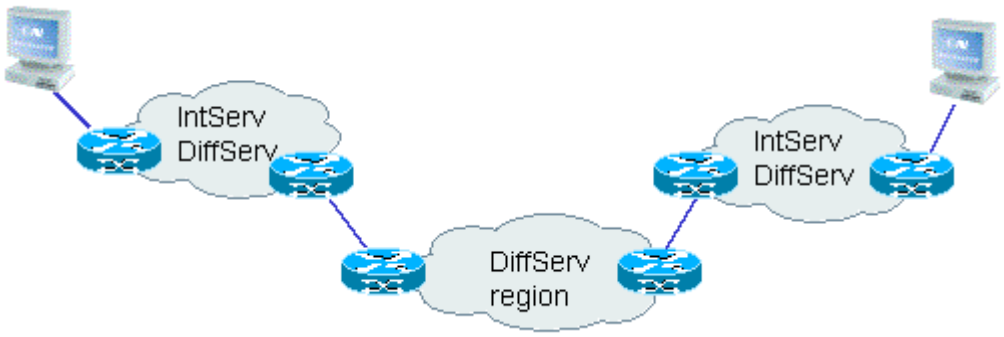

圖 5.32 整合作業與平行作業

5.3.3 流量工程

在 VPN 的 QoS 中有一項重要的議題，那便是流量工程(Traffic Engineering)。流量工程便是在處理我們要如何善用網路上的資源、控制網路流量的流動，使得網路績效達到最高。而且流量工程能控制網路上的特定路由，減少擁塞情形並改進流量效率。

傳統的路由協定通常在尋找至目的位址的最短路徑，然而這不一定是最佳的路徑；透過流量工程，ISP 可以建立不同於最短路徑

之符合特定需求的路徑，例如符合 QoS 的頻寬與延遲需求、或是建立備援標籤交換路徑(Label Switching Path；LSP) 以提高網路的可靠度。 MPLS 網路中，只要在入口邊緣標籤交換路由器(Label Edge Router；LER) 處給予封包不同的標記，封包即會依循不同的 LSPs 到達目的地，這使得流量工程中的 data plane 部分，在 MPLS 網路中變得相當容易實現，而剩下的問題在於 control plane 如何建立 Explicit-routed LSPs。這涉及兩部分，一是 Explicit-routed LSP signaling 目前有兩個主要相關協定：CR-LDP 與 RSVP-TE；至於哪一種協定會成為主流，尚待市場決定。另一部分是執行 CSPF (Constrained Shortest Path First) 演算法，以決定路由，這部分則不會有標準，端視電信業者自己的作法。

在流量工程上我們所採用的是重疊模式 (overlay model)，而且運用基礎的連接導向網路技術 (connection-oriented network technology)，例如：非同步傳輸模式 (ATM, asynchronous transfer mode) 或 frame relay。但是，重疊模式有好幾個問題存在。最重要的問題便是我們在操作 IP 和基礎的連接導向網路技術時非常難以管理。因此 mesh-like 網路無法有彈性地運用於大型的 VPN 之上。

5.3.4 MPLS 動態路由及資源分配

不同服務頻寬的競爭與衝突是目前 VPN 所面臨到的一個重要問題。網路資源的不足或不可預知的突發性資料量，都是導致網路壅塞和促使網路交通量變得緩慢遲鈍的主要原因，也直接衝擊到網

路整體的效能。

訊務資料量不平均的分佈在 VPN 網路的原因，大部分是由於現今的動態路由通訊協定所造成的，如: RIP、OSPF 等。因為這些路由通訊協定大多採最短路徑所計算的路由來傳遞資料，所以往往造成某些鏈結將重複被最短路徑所計算而使用，進而造成這些鏈結發生頻寬的爭搶、壅塞的情形。如何有效率的管理整體網路資源及路由，改善壅塞和頻寬衝突的發生，則需要透過流量工程技術。

以下範例說明各種流量工程計算所造成的結果。

考慮以下的網路架構(如圖 5.33 所示)：

所有的 link 的 cost 均為 10(OSPF=100M/Bandwidth) 。

從 RtrA 到 RtrE 路徑經過 A⇨B⇨E，cost 20。

所有從 A 到 Rtr {E，F，G}的訊務流都會經過 A⇨B⇨E。

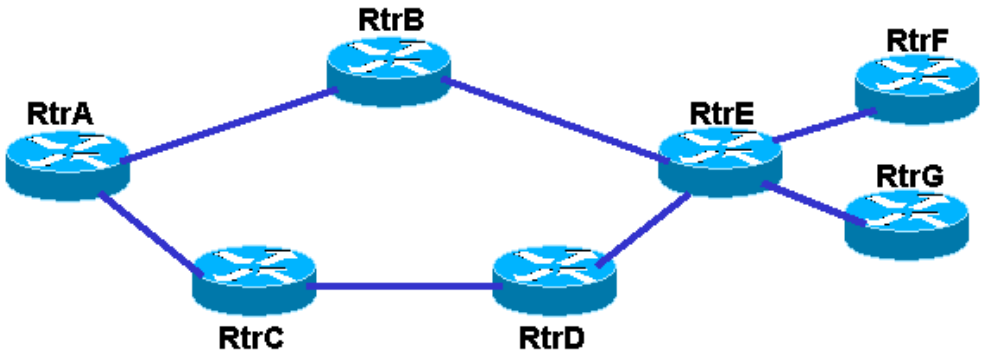

圖 5.33

● 距離向量協定(Distance Vector Protocol)

RtrA 看不見所有 link。

RtrA 只知道最短路徑(shortest path)。

所有 link 要事先經過規劃。

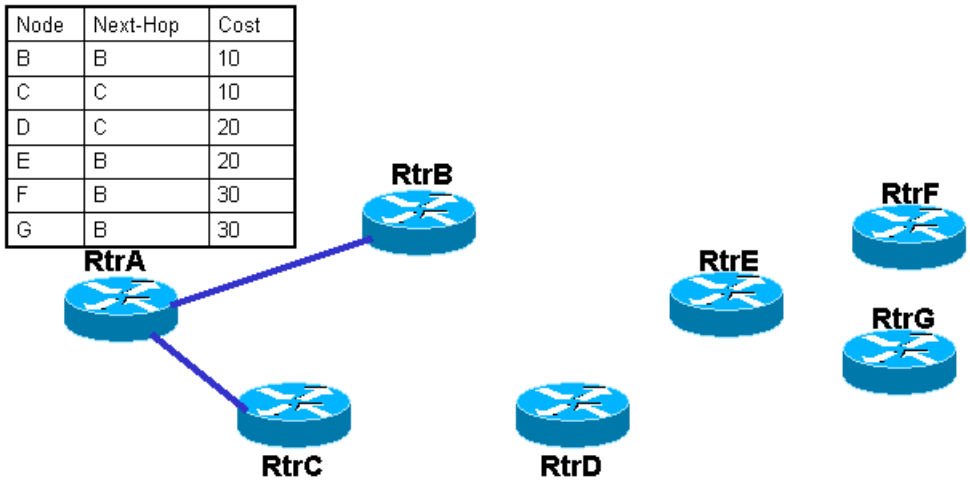

圖 5.34 距離向量協定

● 鏈結狀態協定(Link State Protocol)

RtrA 看見所有 link。

RtrA 能計算出最短路徑(shortest path)。

路由表(Routing table)不會改變。

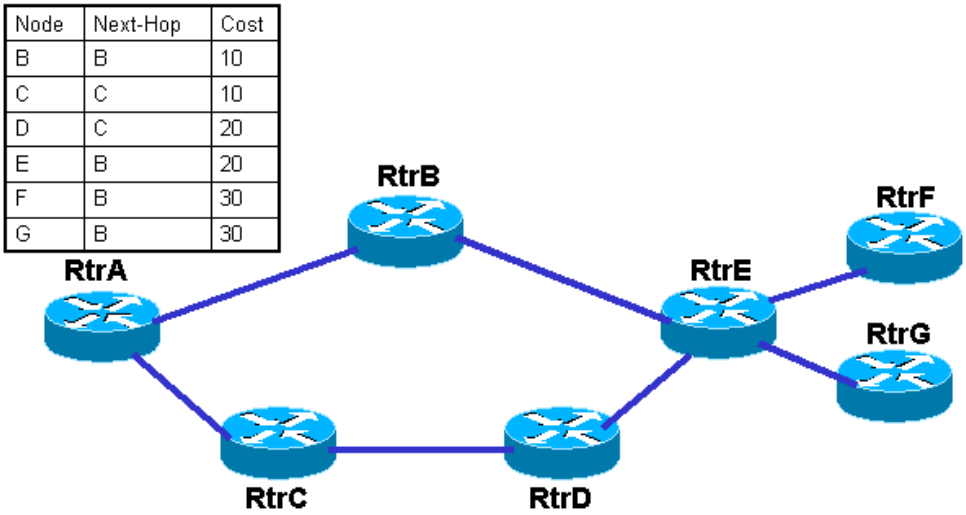

圖 5.35 鏈結狀態協定

● 最短路徑(Shortest-Path)問題

RtrA 看見所有 link。

一部分 link 為 DS3(45Mbps),一部分為 OC3(150Mbps)。

RtrA 有 40Mbps 的訊務預定傳送給 RtrF, 40Mbps 的訊務給 RtrG。

在 RtrB⇨RtrE 間傳送的封包會有 44%大量遺失。

改走 A⇨C⇨D⇨E 同樣情形仍會產生。

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	B	30
G	B	30

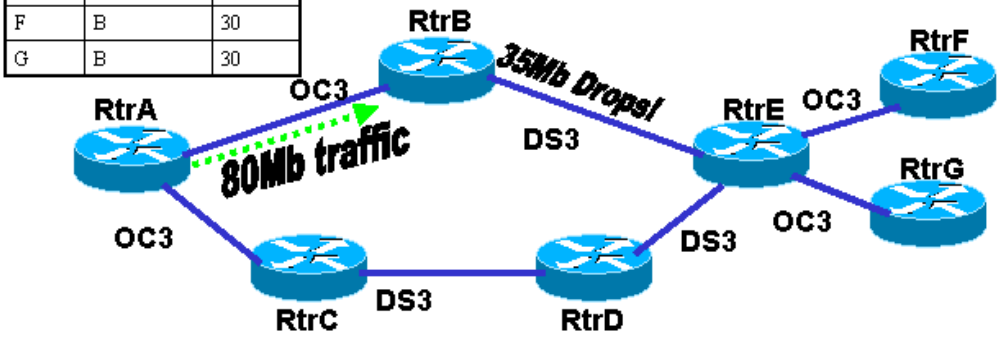

圖 5.36 最短路徑

如圖 5.36 所示，根據最短路徑(shortest path)路由協定，訊務流會被導向最短的路徑，因而造成某些路徑發生壅塞，而其他路徑的頻寬使用率極低。例如 RtrA 有 40Mbps 的訊務流要到 RtrF，有 40Mbps 的訊務流要到 RtrG，若依照 shortest path 的路由協定，將會有 80Mbps 訊務流由 RtrA 流向 RtrB，然後再由 RtrB 流向 RtrE，因此會有 35Mbps 的訊務流會因為頻寬不足而被丟棄。

● MPLS 流量工程(Traffic Engineering)

RtrA 可以看到所有 link。

RtrA 會適當地計算出路徑而非僅僅是採最短路徑的最小 cost。

不會產生壅塞。

Node	Next-Hop	Cost
B	B	10
C	C	10
D	C	20
E	B	20
F	Tunnel0	30
G	Tunnel1	30

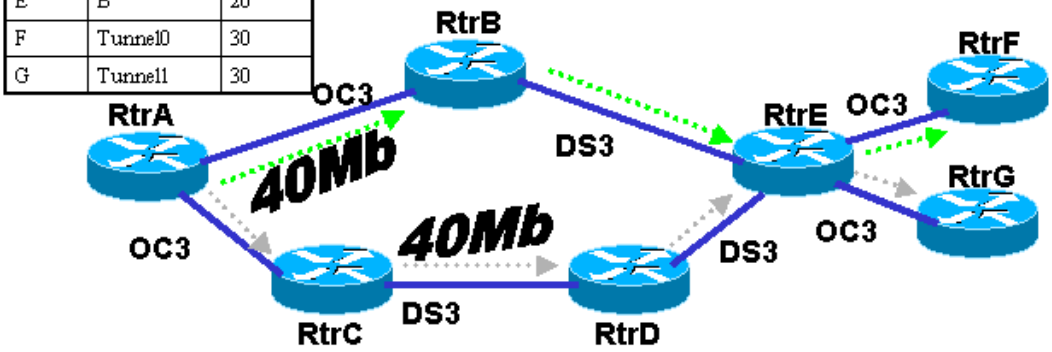

圖 5.37 流量工程

如圖 5.37 所示，採取流量工程(Traffic Engineering)技術，RtrA 會考慮整體網路的狀況，然後得到與最短路徑(shortest path)不同的路由結果，所以 RtrA 會將 40Mbps 的訊務流導向 RtrB，然後傳送至 RtrF，另外會將 40Mbps 的訊務流導向 RtrC，然後傳送至 RtrG。如此整個網路不會發生壅塞現象，並且整體網路頻寬的使用效能可以提升。

5.3.5 MPLS VPN 自動化服務品質管理

VPN 網路設備能提供服務品質的功能，若是沒有網路代理系統進行自動化的管理，網路服務品質的提供將會服務供應商的惡夢。

網路代理系統必須遵循標準化的電信網路管理 (TMN, telecommunications management network)模型來設計，這個模型提供了一個管理「服務提供者 (IPS)」 的方法。它包含五個階層，分別是網路元件層 (network element layer)、元件管理層 (element management layer)、網路管理層 (network management layer)、服務管理層 (service management layer)、商業管理層 (business management layer)。

低階層必需提供支援給高階層，而高階層則可以使用低階層的所有能力、資源。也就是說高階層會具備有低階層的能力。低階層的元件較為分散也較偏技術。高階層的資訊量較為集中，屬於較高層次的概念。高階層也因此比較容易做到集中控管與維護。為了自動化的 VPN 服務品質管理運作，必須整合 TMN 模型、SNMP、DiffServ 管理架構，透過集權式的智慧型軟體代理主機(service broker)可以做到網路組態的集中管理。這個代理主機必需根據服務策略來做相關的處理動作，這也就是典型的政策伺服器 (PDP) 功能。它也會提供一些額外的功能，例如：DiffServ 的服務等級分類。在服務層所擁有的 IPSec 安全策略或 DiffServ 策略。

在一自動化服務品質管理架構中，結合 SLA 是能確保客戶所付出的成本是否有得到應得的服務。服務等級協議(Service Level Agreement; SLA)定義了服務的類型、服務提供的資料傳輸率，以及依照接受服務的使用者所付出的代價，能夠獲得什麼樣的效能水準，有些 SLA 還會指定當服務業者無法提供客戶所預期的效能時，客戶將可獲得賠償。

一般 SLA 常定義的一些衡量指標包含：

(1) 效能：效能衡量指標通常以回應時間與封包遺失率來表示。
(2) 使用率：此項目定義在保證的回應時間與封包遺失率下，最大的服務使用量。
(3) 可靠性：包括在一段時間內的各項可用性保證。
(4) 服務的提供：保證會以某種方式來提供服務。
(5) 客戶支援：包括安裝時程、付費條款、典型的問題報告與問題解決的保證。
(6) 業務的持續運作及故障回覆：包括發生事故時的回覆計畫、備

援設備、分散式備份與儲存計畫。
(7) 終止條件與相關法律議題：包括保證、賠償和責任限制等。

5.4 VPN 解決方案之相互比較

一個電信業者能在競爭激烈的網路環境生存，不單是靠頻寬，服務的多樣化與市場區隔更是致勝的關鍵，以下就技術面與架構建置面 Layer 3 MPLS 、Frame Relay 與 ATM 分別做一個客觀的比較。

5.4.1 技術面

技術面			
	L3 MPLS	Frame Relay	ATM
VPN 建構方式	Private Label	Virtual Circuit	Virtual Circuit
VPN 分類	Network based Peer to Peer	Network based Overlay	Network based Overlay
OSI 模型處理層級	3	2	2
QoS	第 3 層的 QoS	無	第 2 層的 QoS
CoS	Yes	可, 但因架構屬 Overlay,實際從事點對點的 Cos 相當困難	可, 但因架構屬 Overlay,實際從事點對點的 Cos 相當困難
網路管理	容易	複雜	複雜
網路私密性	高	高	高
Traffic Engineering	有	無	有
VC Number	n	Σ(n-1)	Σ(n-1)
Scalability	高	低	低

表 5.3 技術面

5.4.2 架構建置面

架構建置面(一)			
	L3 MPLS	Frame Relay	ATM
網路架構	單純	隨著據點增加而複雜	隨著據點增加而複雜
建置成本	低	高	高
接續方式	多	多	少(專線或 ADSL)
架構擴充性	高/成本低	低/成本高	低/成本高
Extranet 整合	容易/成本低	困難/成本高	困難/成本高
客戶設備	Data:一般 Router Voice: Voice gateway/Router with voice function	Data:一般 Router Voice: Voice gateway/Router with voice function	•現有 FR 設備(T3 以下速率) •需加裝 ATM 模組(T3 以上速率)

表 5.4 架構建置面

架構建置面(二)			
	L3 MPLS	Frame Relay	ATM
服務提供者	第一或第二類電信業者	第一或第二類電信業者	第一類電信業者
服務內容	1.高度安全性的 VPN 服務 2.遠端撥接連結企業 Intranet 3.多重選擇的備援電路方案 4.網際網路連接 5.Intranet/Extranet 的整合	1.支援多種通信協定 2.可執行各種 OSI 模組 Layer 3 通訊協定,如: TCP/IP、SPX/IPX 、Apple Talk 、NetBIOS…等 3.提供多點接駁速	1.建置 ATM VPN 網路 2.接取政府電子採購領標投標系統 3.接取農勞保系統 4.主機資料更新 5.接取 HiB2B 6.接取 HiVideo

	6.不同等級服務(Class of Service)	率,範圍:9.6k~E1 4.彈性的超頻寬政策 5.CIR 最小達 4K,最大達 2048K 6.滿足 ITU-T Frame Relay 與 ANSI 標準	
IP Application 整合	IP based,彈性大且具第二層的 privacy	屬第二層,高 privacy 但整合不易	屬第二層,高 privacy 但整合不易
Voice/Video/Data Integration	有, 且具訊務分級能力	可, 但無訊務分級能力	可, 但無訊務分級能力
適用對象	特別適合拓點迅速或需全網狀(fully mesh)之 VPN	適用對網路傳輸安全性有相當高要求的企業	超高速數據交換網路適用於頻寬需求高於 T1 或以上的企業

表 5.5 架構建置面

5.5 結 論

提升到以 MPLS 技術為基礎的核心網路，可提供企業 VPN 網路服務供應商整合與收容第二層(Layer 2)客戶群及第三層(Layer 3)客戶群，能夠讓各種不同網路技術收斂到單純的核心網路(如下圖 5.38 所示)。

以第二層(Layer 2)客戶群而言，目前已經可支援乙太網路規格的多協定標籤交換技術(Ethernet over MPLS;EoMPLS) 、AAL5 層之上的多重協定封裝技術(Multiprotocol Encapsulation) 、ATOM(Any Transport over MPLS) 、ATM、Frame Relay。MPLS 技術可支援非常大型且分散的第二層網路環境。第二層(Layer 2)客戶除了彼此可以透過 VPLS/Martini/Kompella 建 LSP tunnel 方式互通，也可利用第二層的 VLAN 與虛擬電路(VC)的透通性，使得第二層(Layer 2)客戶可以跨越 MPLS 核心網路與第三層(Layer 3)客戶互通。

以第三層(Layer 3)客戶群而言，可支援 MPLS VPN、IPSec 與 IP VPN，但由於 MPLS VPN 可支援第三層的服務(如服務品質(QOS)、訊務工程(TE) 、快速回復(Resiliency)等) ，在第三層控制面(Control Plane)和資料面(Data Plane)上，能提供虛擬私有網路服務且擴充性非常高，所以較能被服務供應商所青睞。在互通性方面，所有類型的第二層服務如 Gigabit 乙太網路(GigE)、ATM、Frame Relay、都會型乙太網路、ADSL、Cable Access、Wireless LAN 及 Dialup，都可同時接取 MPLS VPN 網路，也可和其他電信業者的 MPLS 服務互通和合作共存。

MPLS 核心網路的部分，可以採用的各種傳輸技術如：高密度分波多工(DWDM)、同步光纖網路(SONET)、同步數位階層(SDH)、Gigabit 乙太網路(GigE)。除了使整個核心網路架構能滿足今天網路用戶不斷成長的頻寬需求外，更可以達到異質網路集中維運與降低核心網路投資成本之目標。

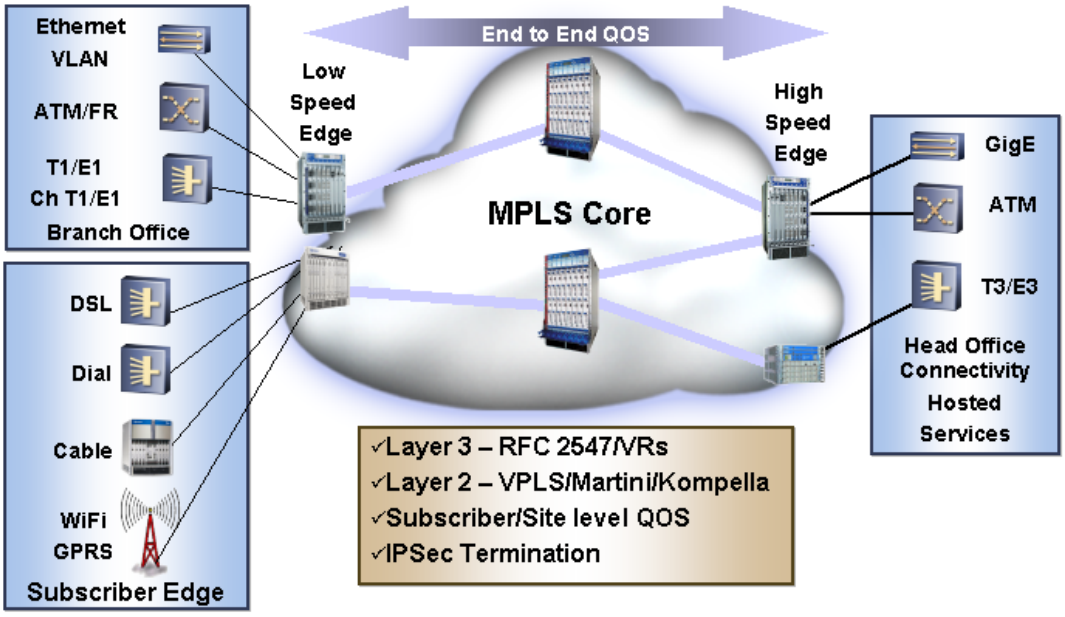

圖 5.38 高效能 MPLS 核心網路

第六章 研習心得

隨著資訊化的高度發展，網路的資料流量隨著電子商務、多媒體資訊傳輸、大量檔案下載等應用呈現大幅的成長。瞬間大量的資料傳輸更影響了企業網路的使用效能。加大頻寬不但非常昂貴，並且不能保證能夠解決網路效能不足的問題。然而沒有設定 QOS 的網路傳輸服務是以所謂盡力而為(Best Effort)的方式來提供頻寬服務，並無法保護傳輸的頻寬。這樣的方式對於以往如 E-mail、FTP 等服務尚可以被接受。但面對 ERP、電子商務、VoIP 及 VoD 等多媒體資訊傳輸服務，則無法滿足這些應用的頻寬，以及低延遲的要求。

雖然 QoS 並不能產生新的頻寬，但是它可依據網路服務的需求以及網路管理的設定來有效的管理網路頻寬，使得 VPN 網路能提供穩定、可預測的資料傳送服務，來滿足網路服務的需求。

然而以 QoS 架構實現的難易度來說，整合式服務因為有擴充性的問題，較適合企業網路的建置，對 ISP 而言並不恰當。差異式服務則適合其型態，差異式服務網路是一種屬於服務級基礎型(Class-based)之保證服務品質之機制。與資料流基礎型(Flow-based)機制不同之處，便是它減去了許多在路由器上儲存資料流資訊的多餘負擔。它在 Edge Router 將流量分類後，將具有同一服務品質特質的流量形成聚集的形式，再經核心路由器處理。因為路由器只需對其有相同服務品質特性之數種服務等級處理而不需對每個資料流做處理，便可使得負載降低，並又簡化路由器之設計。差異式服務架構最大的優點是其擴充性，因此非常適用於大型網路，而能具有擴

充性的好處主要是其將複雜的處理程序由網路的中心節點移至 Edge 節點，因 Edge 節點處理的資料量和流量數通常較核心路由器為少，又其提供的服務是「集合性質」之資料流型態 (Aggregated Traffic)，而不是「單一資料流」之型態 (Per -flow)。 目前新一代的路由器皆有提供差異式服務與 MPLS 的功能，所以就 ISP 而言，現行要實現 QoS 還是以差異式服務架構較為可行。

多協定標籤交換(MultiProtocol Label Switching；MPLS)是新一代確保網路通訊品質的通訊協定，這個通訊協定萃取了 IP(Internet Ptotocol)、ATM(Asychronous Transfer Mode)與 Frame Relay 等網路架構的優點，可以運行於現有的 IP、ATM 與 Frame Relay 網路架構之上，不過目前應用於 IP 網路之上居多，如 IP VPN(Virtual Private Network)網路上加上 MPLS 的功能為目前當紅之應用。

MPLS 的網路運作方式為 Label Switching，有別於 IP 網路的 Packet Switching 與 ATM 網路的 Cell Switching，MPLS 在資料封包上加上一個標籤(Label)，以設定不同等級的服務品質，同時增加網路路由(Routing)選擇路徑的效率。IP、ATM 與 Frame Relay 網路均為技術已相當成熟且普遍的網路架構，而彼此均有各自的優缺點：IP 網路由於設備價格便宜，且應用廣泛，未來發展潛力深受肯定，但是 IP 網路最為人詬病的為其網路品質(QoS)問題一直無法有較完善的解決方案；而 ATM、Frame Relay 網路雖可針對不同的應用提供不同等級的網路品質服務，但是其設備十分昂貴且架構複雜、管理不易，因此漸漸不受到青睞。在 IP 網路架構逐漸成為網路架構的主流之際，MPLS 可以補強 IP 網路在 QoS 的不足之處，因此在 IP 網路上搭配 MPLS 通訊協定，將是未來廣域網路上通訊協定的主流。