第四章 MPLS VPN 系統規劃與功能設計

本節將針對論文第三章提出的企業網路需求分析，參考【7】為 ISP 規劃一個具擴充性的 MPLS VPN 骨幹網路架構，並針對企業全球化趨勢，需要將企業網路延伸至中國大陸與國際各地的情況發生，因而與國際 ISP 業者合作提供跨 ISP 的企業網路延伸架構，可為企業客戶提供單一窗口、價格優惠、保證網路品質、穩定性高的 VPN 服務，另外針對這樣的網路架構提供網路管理系統，在網管系統上使用以政策為基礎的技術概念規劃，透過 Policy Server 直接管理區域的網路設備，並將其資訊傳送至 Policy Manager 集中儲存，並呈現有價值的整合資訊以網頁的方式傳送給網管人員，此外網管人員也可透過 Policy Manager 做整體政策設定，再分別交由各區域的 Policy Server 對當地的網路設備做即時組態設定，另外此系統區分為兩部分，第一部分是提供給 ISP MPLS VPN 骨幹網路使用的，包括有組態自動設定系統及障礙與品質監控系統，第二部分是提供給企業客戶的加值服務，讓其可以掌控自己 VPN 網路的線路障礙與品質監控系統，這種設計主要是讓企業客戶不但享有網路委外的經濟性與便利性，並且透過提供企業可以監看自己 VPN 網路狀態的網管加值服務，來增加客戶網路自主性與安全性的感覺，也是此篇論文的重點。

4.1 MPLS VPN 骨幹網路規劃

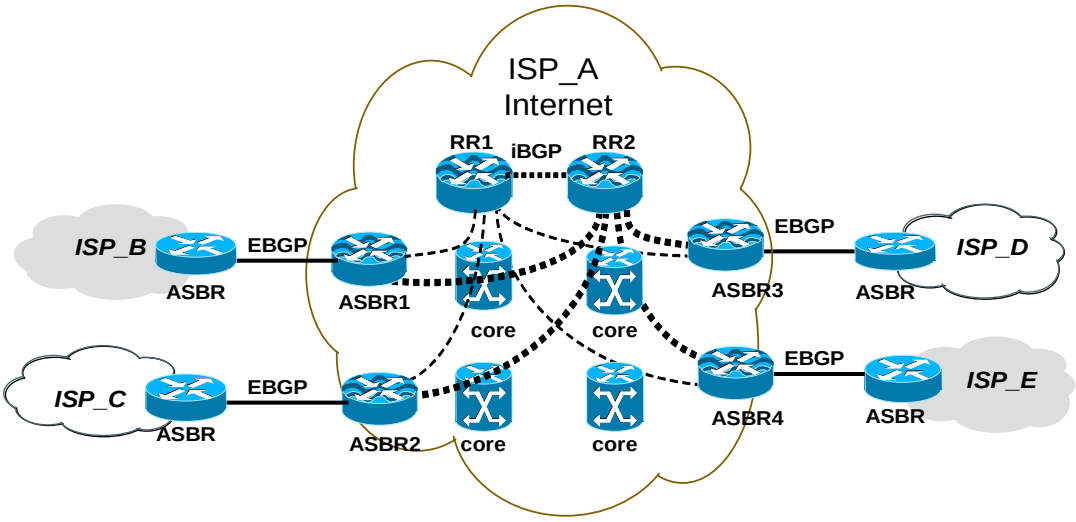
4.1.1 MPLS VPN 骨幹網路擴充性架構設計

ISP 在網際網路上為了防止網路在 Autonomous System(AS)內形成路徑迴圈，所以 BGP 有所謂水平分割(Split Horizon)的原則，也就是說本地路由器經由 BGP 得知的路由資訊，不可再轉手傳送給其他 BGP 對等節點之路由器，因此造成 AS 內部的所有 BGP 路由器必須構成邏輯上完全網狀(Full-Mesh)連結才能彼此交換路由資訊，如果在 AS 中的每個 BGP 路由器都要與其他對等 BGP 路由器建立 BGP TCP session 形成所謂完全網狀(Full-Mesh)連結，則需要建立的 BGP TCP session number 是 n(n-1)/2，其中 n 是 BGP 路由器的數目，這樣在大型網路中會有擴充性受限的問題，這是因為路由器越多需要建立及維護的 BGP TCP session 越多，同時會浪費路由器大量的記憶資源，使每個 BGP 路由器重覆儲存相同的路由資訊，並且增加不必要的網路流量。

為了解決這個問題而提出了 BGP 路徑反映器(Route Reflector,RR)的可擴充性解決方案，請參考【圖 4-1】，概念是我們在 AS 網路中選擇一個路由器當成

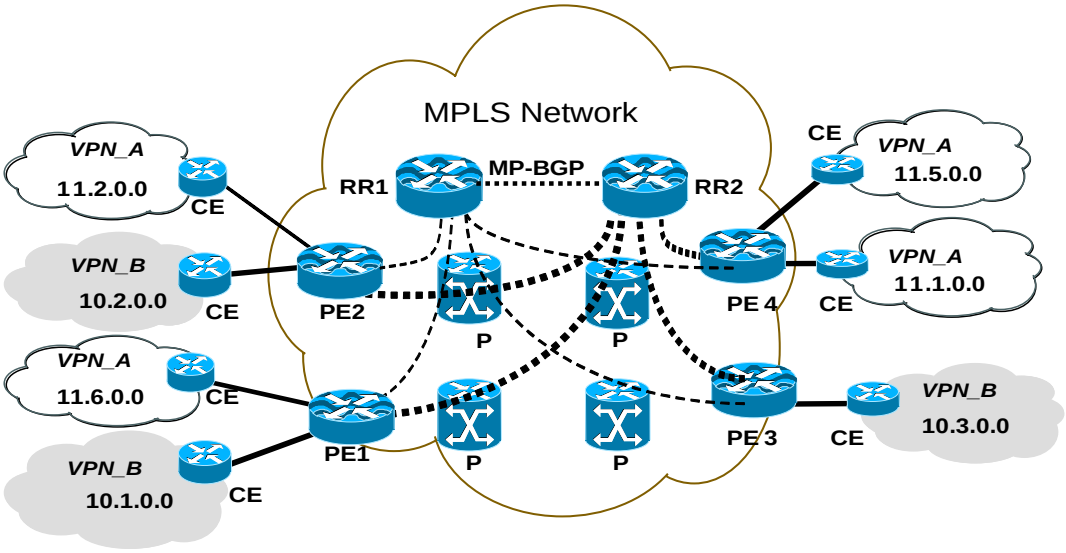
BGP 路徑反映器(Route Reflector,RR)，允許它可以將它從一個 BGP 對等節點路由器取得的路由資訊轉送給另一個鄰接 BGP 對等節點的路由器，因此縮減了必須建立的 BGP TCP session 數目，換言之與 BGP RR(路徑反映器)建立對等關係的路由器稱為客戶端 BGP 路由器，它會傳送自己的路由資訊給 RR，而 RR 再將路由資訊轉送給鄰接節點的其他 BGP 路由器，但客戶端 BGP 路由器還是必須遵守 BGP 有所謂水平分割原則。

為了防止在 AS 網路中唯一的 BGP 路徑反映器發生故障，必須規劃備援方案，所以在 AS 網路中建立多個 RR，為了確保 RR 會轉送客戶端更新的路由資訊給其他非客戶端的 RR 所以在 AS 網路中多個 RR 間必須建立邏輯上完全網狀連結的 BGP TCP session。



【圖 4-1】ISP 網際網路 BGP 擴充性網路架構

同樣的我們在規劃 MPLS VPN 骨幹網路時，使用 RR 的架構，不但可使 MPLS VPN 客戶路資訊的傳遞具有擴充性(因為不需在 PE Router 之間，建立完全網狀連結的 BGP TCP session 請參考 2.2.2 的【圖 2-17】)，而且每一個 PE Router 只需儲存跟他有直接連線的 VPN 客戶節點之路由資訊，請參考【圖 4-2】，同時此 VPN 客戶其他節點之路由資訊都集中儲存於 RR，因此可大量節省路由器的記憶體資源，所以網路上所有的 PE Router 都要與 RR 建立一條 BGP TCP session，以將 VPN 客戶之各節點的路由資訊集中儲存於 RR 路由器上，這種作法可以說是 VPN 路由資訊集中儲存於 RR，而且分散執行於各節點 PE Router，簡而言之 RR 是 VPN 路由資訊的 Server 端，分散在各區的 PE Router 是 VPN 路由資訊的 Client 端，Client 端的 PE Router 如有任何新增客戶的路由資訊，都會對 Server 端的 RR 做 Update。



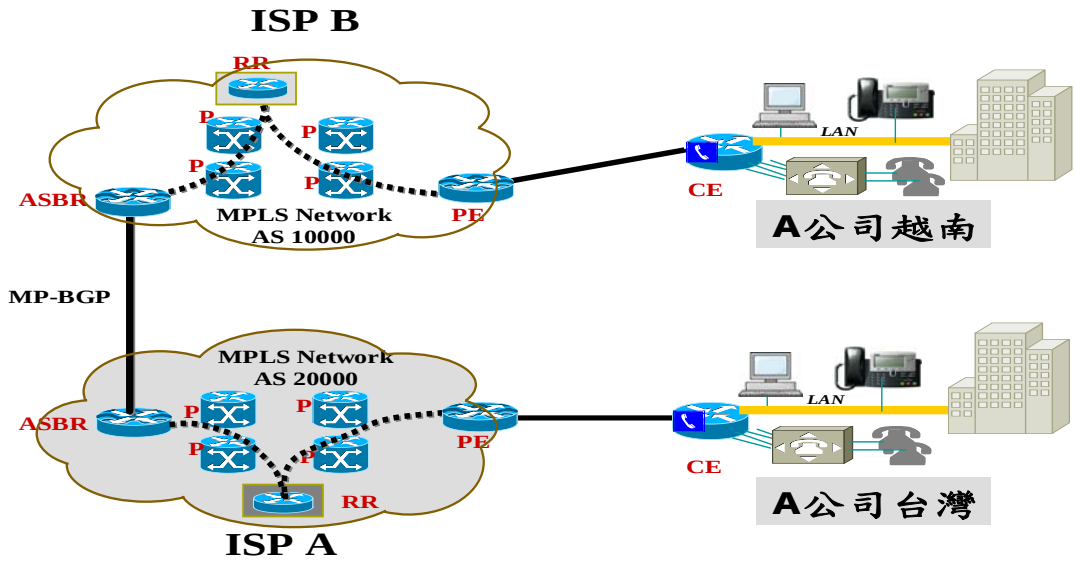
【圖 4-2】MPLS VPN 擴充性網路架構圖設計

4.1.2 MPLS VPN 骨幹網路國際延伸規劃

ISP 在幫企業虛擬私有網路找尋網路延伸國際端分公司解決方案時，早期大多採用國際海覽專線或衛星連線的方式，前者線路月租費用非常昂貴，並且網路頻寬的使用缺乏彈性，沒有較好的備援方案與品質保證，海纜斷線時需修復的時間太久等問題，然而若採用衛星連接的方式則有 delay time 過長的問題，若要傳送 voice 等 delay sensitve traffic 時無法使用。所以本論文建議 ISP 因應國際上 MPLS 技術已普及的情況下，與有國際佈點的 ISP 簽署合作契約，與其在台灣的 Pop 點界接，透過 MP-BGP 協定將企業國內節點與國外節點的網路連接起來，在 ISP 之間的網路合作關係，很類似在網際網路上的不同 ISP 間 peering 及 transit 流量交換的概念，差異在交換的資料流量種類是公開性的資訊還是企業私密性的資訊。

關於兩個 ISP 之間的合作界接模式有兩種，一種是 MPLS trunk 的界接方式，請參考【圖 4-3】，ISP 間界接的路由器稱為 ASBR(Autonomus System Border Router)透過 MP-BGP 的協定來交換企業網路國內端與國外端的路由資訊，這種合作模式通常是兩個 ISP 間已有大量的企業私有網路資訊在交換，同時要提供企業網路的國內端總公司到國際端分公司間 end-to-end 的網路品質保證 Qos 時，另一方面也是當在企業客戶國際端界接在此 ISP 的節點數達到一定程度時，所採用的合作模式，另一種合作模式是 PE-to-PE 的界接方式，是當企業客戶國際端界接在此合作 ISP 的節點數與客戶數很少時，可將企業購買國際端企

業網路服務的需求，代為向國際 ISP 購買此服務，所以本地 ISP 與國際 ISP 的合作模式較不密切，只是類似為單一客戶單一節點連網服務的轉賣行為，但是在本地 ISP 與國際 ISP 初期合作階段，客戶數不多且企業客戶國際端節點數不多的情況下，可以暫時採用這樣的合作方式提供企業客戶網路服務。



【圖 4-3】MPLS VPN 骨幹網路國際延伸架構

4.2 MPLS VPN 網管系統架構規劃

MPLS VPN 網管系統架構設計

本論文將以 Policy-based framework 技術的概念，應用於管理 MPLS VPN 網路之組態設定管理(Configuration Management)、障礙管理(Fault Management)、品質管理(Performance Management)，運作方式為管理者制定某些有意義的網路政策(Policy)，交由 Policy-based 網路管理系統執行，系統將依據其所訂政策內的規則與條件(Rules)，自動執行網路資源管理服務，請參考【圖 4-4】例如對所管轄的網路設備進行相關組態設定、頻寬等級管理或安全管理等服務，提供網路及通訊業者簡單而完整的網路管理技術解決方案，最後希望能達到網路資源集中管理，網路政策分散執行的目標。同時建議當此系統要控管的服務越來越複雜時，使用者越來越多時，必須利用角色為基礎的存取控制技術，以組織中員工所扮演的角色作為授權控制之標的，使身分鑑別與存取控制做適當的整合，可以減少資訊系統連結之負荷，提昇整體之安全性以及可用性。

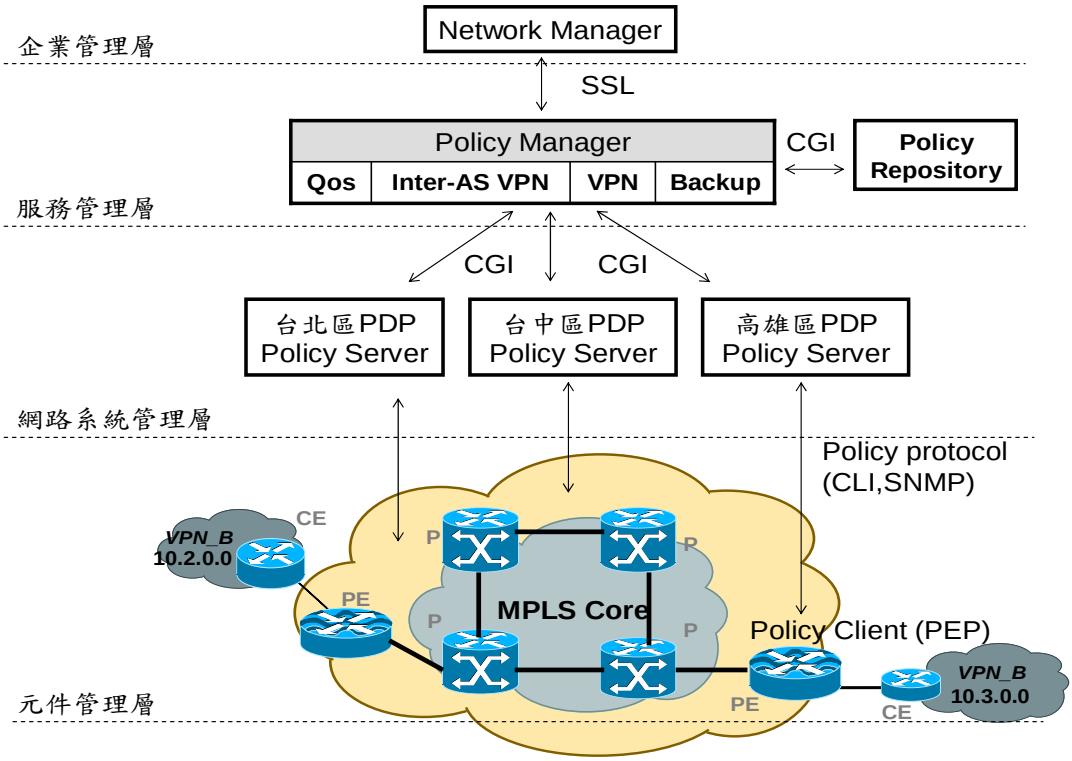
根據 DMTF 與 IETF 這兩組織提出的 Policy-based Network 的階層式系統架構說明(請參考 2.5.2)，管理系統中四大階層中的每一個階層都是獨立的子系

統，各個階層均可透過一定的管理協定來互相溝通，Network Manager 與 Policy Manager 間使用通用的網路安全協定 Secure Socket Layer(SSL)，在上層應用協定間提供資料加密與認證的功能，它利用了非對稱式加密演算法來交換交談金鑰，再利用對稱式加密演算法的交談金鑰來加密應用程式資料。本研究根據 ISP 產業的實際需求做調整如 Policy Manager 和 Policy Repository 間是直接採用 CGI 程式互相傳遞網管資訊，而不使用 LDAP 的目錄方式來傳遞資訊，以簡化系統複雜性增加執行效能，PM 和 PDP 同樣也直接採用 CGI 程式互相傳遞網管資訊，PDP 根據網路服務功能的不同而區分不同的模組，例如對於專門控管 MPLS 骨幹網路的流量調整政策的 PM for TE 模組及專門用於企業客戶 VPN 網路組態設定參數用的 PM for VPN 模組。

針對系統安全性設計，在 PDP 與 PEP 模組間加入 Secure Shell (SSH)及 ACL(Acess Control List)的安全模式，SSH 是一種在不安全網路上用於遠端登入其他安全網路服務的協定，它提供了安全的遠端登入可以自動加密、認證並壓縮所傳輸的資料，以此加強在做自動組態設定時的安全防護，同時在 PDP 與 PEP 模組間原先採用 Common Open Policy Service (COPS)作業協定來傳輸資料，但為因應 ISP 產業實際狀況中有各種不同廠牌設備及為了要相容於老舊網路設備因此採用 telnet script 程式模擬 CLI 的方式進行設備的組態設定行為。

MPLS VPN網路管理系統架構分為四層，每一層的定義及服務內容是參考【8】並根據ISP產業實務經驗做修改，最上層的「企業管理」讓網路管理人員可以透過底下幾層來獲得全面的管理資訊，「服務管理」是涉及ISP提供給企業客戶訂購的網路服務商品之功能模組。例如VPN服務、高級Qos服務及線路備援服務(像企業向ISP訂購台灣與美國之間通訊，是透過IPLC國際海底電纜來傳遞訊息的，如果發生纜線斷裂的情況，造成對外通訊中斷，則ISP如何利用備援線路來維持客戶連線的暢通，這類應變的模式就是「服務管理」層應該要想到的)。在「網路管理作業系統」層是執行骨幹網路區域節點內設備與線路實際運作的相關警告、資料過濾、與參數設定等工作，「元件管理作業系統」層則是以代理程式（agent）的方式，在網路的各個元件中執行收集網路狀況資料的工作（例如集線器、交換器、路由器或是伺服器上），並且可以把設備狀態以被動或主動的方式傳遞給「網路管理作業系統」層。

網管系統分層架構是為了達到較佳的管理效率，讓網路服務的資訊集中管理，並分散同步控管各區域節點的網路設備與線路，降低網管資訊在骨幹內傳遞的流量，另外也可突顯網路服務的商業價值，隱藏背後複雜的管理規則，特別是針對企業VPN網路有上百個節點分散在各區時，獲得的管理效率最高。



【圖 4-4】MPLS VPN 網路管理系統架構

4.3 MPLS VPN 組態管理系統規劃

4.3.1 MPLS VPN 組態參數規劃

VRF 名稱的規劃

PE Router 會根據設定不同 VRF 名稱的 VPN 客戶，切割一個獨立的空間，存放此 VPN 客戶的 Routing Table，這個各自獨立的 VPN 客戶 Routing Table 和 PE Router 本身網際網路的 Grobal IP Routing Table 是分開的，彼此不會互相影響。因此我們以[客戶編號]例如 F8208102 做為 VRF 的名稱，方便我們查詢辨識該客戶的路由資訊。

RD 值的規劃

RD 是一個長度 64bits 的數字，附加於 IPv4 位址的前面使其達到絕對唯一(global unique)，這樣組合長度為 96bits 的位址表示法稱為 VPNv4，此乃為了防止 VPN 企業客戶規劃 IP 位址時發生重覆的問題，此外每一個 VRF 只能設定一個 RD 值，同時因為企業 VPN 有可能跨 ISP MPLS 網路建立的情況，所以將其分成兩部分命名(以冒號分開)，前面用 ISP 的 BGP AS no.命名，後半部分用客戶之客戶編號數字部分例如 4782:8208102。

RT 值的規劃

Route Targets 是延伸利用原本 BGP communities 的屬性，我們稱這個屬性為 Extended communities。為了更彈性的規劃客戶的某個節點可以加入一個以上的 VPN 群組，而新增了 RT 這個屬性，由 RT 值來決定某個節點是否要和另一個節點網路互通，這個概念類似於區域網路的 VLAN，所以為了區別網路各節點路由資訊或服務種類，同時考慮到企業 VPN 有可能跨 ISP MPLS 網路建立的情況發生，所以延用 RD 的命名原則，另外增加 2 bits 來區別各種網路服務。例如：

- Full Mesh 服務架構的 RT 值→

  General site 是 4782:820810200。

- Hub & Spoke 服務架構的 RT 值→

  Hub site 是 4782:820810201(01~09 可用)。

  Spoke site 是 4782:820810210 (10~29 可用)。

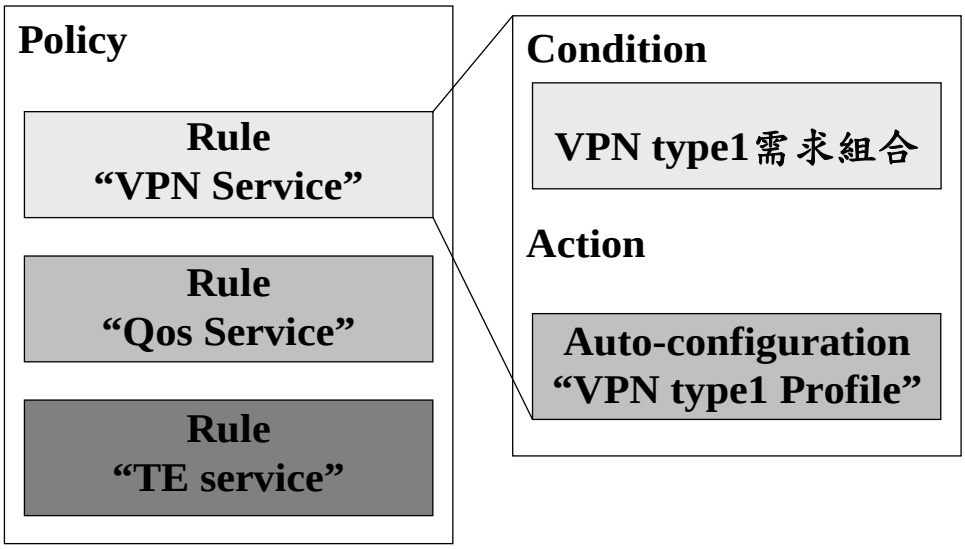
- Internet 服務架構的 RT 值→4782:890810230 (30~49 可用)。

- Extranet 服務架構的 RT 值→4782:890810250 (50~69 可用)。

- Voice 服務架構的 RT 值→4782:890810270 (70~79 可用)。

4.3.2 MPLS VPN 組態政策規則設計

網路管理者可透過 Policy Manager 的政策管理工具產生政策規則(policy rule)，Policy server 再根據 policy rule 來設定網路設備的組態資訊，用來啟動整個網路服務或影響網路服務的行為，一個 Policy 包含多個 rule，每一個 rule 又包含兩個元素 Condition 與 Action，當 Condition 成立時就執行 Action 的程序，以企業虛擬私人網路服務自動設定組態機制為例，ISP 業務人員根據客戶的需求，從 ISP 的企業網路服務需求訂購選單如【表 3-1】中挑選所需的服務項目(如 A1B1C1G2)，此需求組合對應到 VPN type1 的 Condition，根據這個 Condition 找到對應的 VPN type1 profile 的設定組態程序，請參考【圖 4-5】，最後根據此設定內容及程序，經由區域節點的 Policy server 啟動相關程式，自動登入相關網路設備進行設定。



【圖 4-5】Policy Rule 範例簡圖

在 VPN type1 Profile 簡易範例中，請參考【圖 4-6】，包括三個部份：(1)一般設定參數包括 VRF 名稱、RD、RT 的值(2)連線界面參數包括 WAN 界面的 IP address 與將 VRF 轉入該界面(3)VPN 路由參數設定。

PE對客戶端的設定參數

Global command (設定vrf name , RD , RT 的值)

```
ip vrf F8908102    #客戶編號#
rd 4780:8908102   # AS :客戶編號#
route-target export 4780:890810200 #客戶編號+00 ,00表示full mesh #
route-target import 4780:890810200
```

Interface command

```
interface Serial3/0
 description ***connect to 聯華電子 223D89119 *****
 ip vrf forwarding F8908102
 ip address 172.31.1.1  255.255.255.252
 encapsulation ppp
```

routing command

```
ip route vrf F8908102 192.168.1.0 255.255.255.0 Serial3/0 172.31.1.2
```

【圖 4-6】VPN type1 Profile 範例

4.3.3 MPLS VPN 組態自動設定

使用 Telnet script 模式，進行 MPLS VPN 組態自動設定功能，運作的流程是網管人員以瀏覽器經過 Policy Administration User Interface 模組的 SSL 認證後，登入企業網路服務工程設定系統，根據客戶相關網路資料填上系統規劃好的政策選項參數,按下確認後會啟動 VPN Policy Manager 的 Command Dispacher 程式，將政策參數傳送到各節點的 Policy Server，由 Command Daemon 接收到政策參數後，就會去資料庫讀取相關設備組態參數，再由 PDP Rule Locator 將組態設定參數組成 profile，由 telnet script 程式執行 profile，請參考【圖 4-7】。

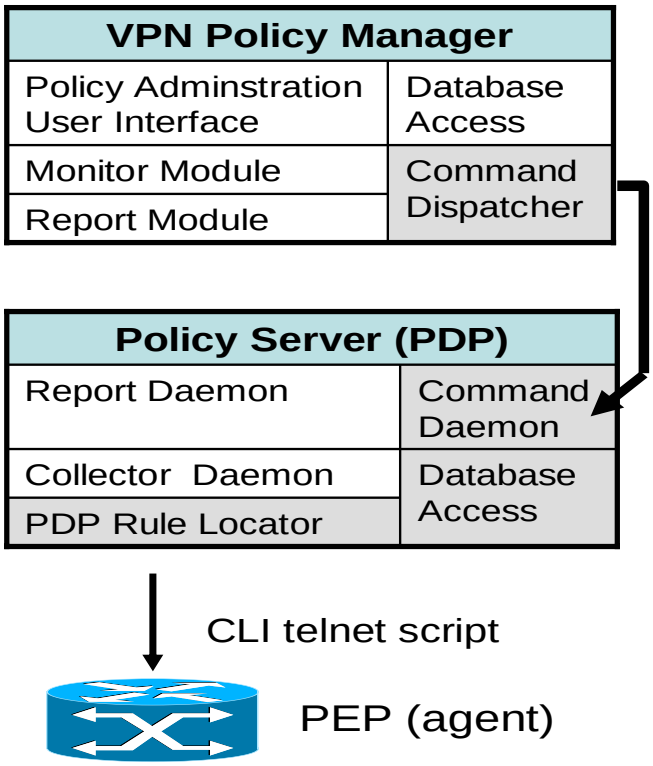
本論文使用 telnet script 的方式做網路組態設定，有兩大因素說明如下：

(1)加強網管系統跨異質網路的管理能力

根據 Policy-based Network 的階層式系統架構中建議，PDP 和 PEP 採用 Common Open Policy Service (COPS)作業協定來傳輸資料，但此協定還不普遍並且一般 ISP 企業內部網路有各種不同廠牌或型號及老舊設備存在的問題，為了能加強此一網管系統跨異質網路的能力所以改用 telnet script CLI 的方式做遠端設備組態設定。

(2)增加網管系統的彈性與可擴充性

因為所有的網路設備皆支援標準 Command Line Interface (CLI)模式，可直接使用 telnet 來將所有的組態指令，以剪貼的方式，將資料傳送到設備上。所以必須發展 telnet script 程式，以外掛（plug-in）方式，提供系統使用，系統只需將必要的組態指令寫成檔案就可以了，即使是未來設備更新，只要支援標準的 CLI，即可使用 telnet script，管理者只要稍加修改組態指令以符合新設備，大大的增加網管系統的彈性與可擴充性。



【圖 4-7】組態自動設定架構圖

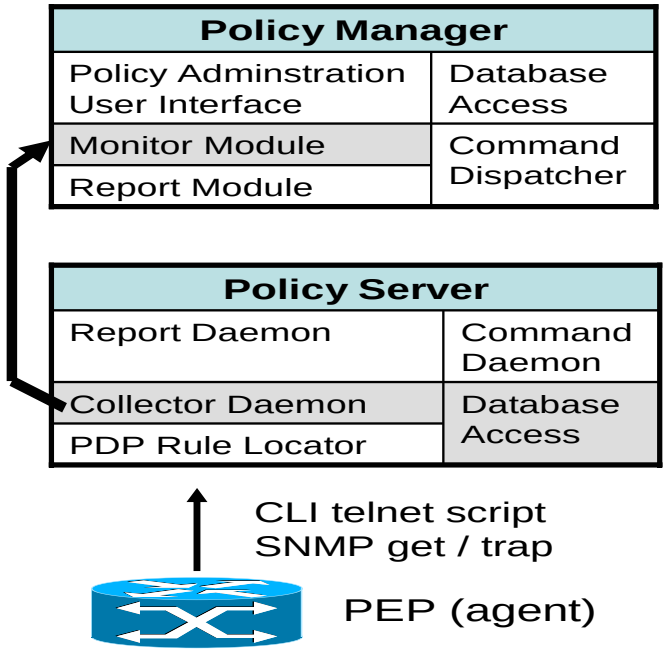
4.4 障礙與品質監控系統規劃

本論文將透過Polocy server內的collector daemon收集從監控設備PEP送出的SNMP trap，能即時發現網路設備與線路的問題，但為了避免線路不穩造成的大量trap混亂網管系統對網路狀況的誤判，所以會在收的trap的同時由collector daemon再發出SNMP get的要求，再次確認網路目前的狀態，因此我們會將SNMP trap與SNMP get功能搭配使用，請參考【圖4-8】，務必掌握網路狀況的即時性與正確性。

MRTG 並不受限於使用SNMP 協定所提供的資料，它也提供彈性的方式，可以接受其它的任意資料，同樣地透過其圖形介面，將結果顯示出。目前系統有兩個有效的方法可以處理當警報爆發時通知系統操作員：視覺方法 --在螢幕上出現一些閃動的圖案聲音方法--可以產生一些聲音引起系統操員的注意另外我們將考慮另一個方法，告警管理例如：發送預設的消息至相關人員的Email及手機簡訊(SMS)即時通知。

希望讓系統操作員可以更容易透過網管系統來一覽所有產生之警報/事件，在螢幕上有固定的位置可以顯示即時的警告統計 - 目前有幾個緊急警告、

重要警告、次要警告、事件警告及解除警告。當警報/事件未被告知接收時，警報/事件將會在螢幕上閃動。另外，拓樸映圖上的管理節點將會擁有自己的狀態列，狀態列會以為色指示出其節點是否包括警報/事件。



【圖 4-8】障礙與品質監控架構圖

4.4.1 障礙與品質監控項目規劃

(1)網路設備品質與錯誤監控

(a)CPU Loading：可判定網路設備本身運作狀況與是否有網路攻擊。

(b)Memory Utilization：可判定網路設備本身運作狀況與是否路由發生異常。

(c)Uplink Port Status：可判定網路擁塞的情況與是否發生設備介面異常。

(d)Reachability：可判定網路設備是否還活著。

(e)設備介面 Packet Drop：可判定是否發生設備介面異常與線路品質異常。

(f)設備介面 Packets/sec：當流量不大但封包數異常增加可能有網路攻擊。

(2)網路線路品質與障礙監控

(a)Line Down(透過 Trap)：可透過設備發 trap 主動通知網管系統。

(b)Line Error (reliability)：定期監測設備介面的 reliability 可得知線路品質。

(c)Delay time/Packet Loss：透過系統定期的 ping 可得知網路品質的狀況。

(d)Traffic Load：系統對線路流量的監控可做為未來線路頻寬擴充的依據。

(3)告警通知方式

(a)Web display status：提供網管中心的網管人員圖像式的顯示方式。

(b)SMS / Email：提供在外的 Mobile user 即時的告警。

4.4.2 網路管理監控政策規劃

客戶端線路(PE2CE)監控規劃

- 斷線(透過 Trap)：監控網路介面狀態，發現斷線時，透過 E-mail 通知客戶技術聯絡人一次，E-mail 通知該地工程及 Hub 端工程每三十分鐘一次。
- Error (reliability)：監控網路介面狀態，發現 Reliability < 250 ，透過 E-mail 通知該地工程及 Hub 端工程每三十分鐘一次。
- Reachable：定期 Ping 客戶設備(透過建 Extranet，從固定監控機器 Ping)，透過 E-mail 通知客戶技術聯絡人一次，E-mail 通知該地工程及 Hub 端工程每三十分鐘一次。
- Delay time/Packet Loss(網管主機 to CE)：Delay time 通知標準為>180ms，透過 E-mail 通知該地工程每三十分鐘一次。Packet Loss 通知標準為>20%，透過 E-mail 通知該地工程每三十分鐘一次。Traffic(所購買頻寬)通知標準為>80%，透過 E-mail 通知工程及業務每天一次。

客戶端VPN網路拓樸(CE2CE)監控規劃

- 監控 Routing Prefix：Hub Sites、Spoke Sites、General Sites、Extranet Sites，透過 E-mail 通知該地工程及 Hub 端工程每三十分鐘一次。

- QoS 監控：PE2PE Delay Time/packet loss(參考 Core 監控)，Traffic Pattern 監控：針對 Lease-line 線路 Sites，各個 CE 間之 Traffic Pattern(By Netflow)，提供 Topology 圖形顯示，點數過多時另外處理。

MPLS VPN骨幹網路端(PE2PE)監控規劃

- PE 路由設備 CPU Loading 監控：通知標準為> 50%時，透過 E-mail 通知該地工程每三十分鐘一次。

- PE 路由設備 Memory 監控：通知標準為< Free 5Mbytes，透過 E-mail 通知該地工程每三十分鐘一次。

- PE Uplink Port 監控：知標準為: Packet loss > 10%，Traffic > 80%，Reliability < 252，Unreachable – 簡訊(SMS)通知該地工程及監控中心。

- PE2PE 監控(高雄到 PE)：通知標準為: Packet loss > 10% 、Delay Time > 40ms，Unreachable – 簡訊(SMS)通知該地工程及監控中心。

- Inter-AS/ ASBR 監控：比對 RR 與 ASBR 的 VPNv4 路由資訊是否異常。

- Route Reflector 網路設備監控：RR 路由器是否存活(Reachable) – 簡訊通知骨幹工程及監控中心。RR2PE 的 IBGP Session 是否正常– 簡訊通知骨幹工程及監控中心。比對 RR 路由器上所有 VPN 客戶的 VPNv4 Routing 資訊是否異常。