

行政院及所屬各機關出國報告

(出國類別：實習)

多協定標籤交換技術
及IP網際網路電話
出國報告

服務機關：中華電信行動通信分公司

出國人 職 稱：副工程師(三)

姓 名：駱襄陽

出國地區：美 國

出國期間：91.11.10~91.11.23

報告日期：92.4.25

系統識別號:C09105206

公 務 出 國 報 告 提 要

頁數: 40 含附件: 否

報告名稱:

實習「多協定標籤交換技術及IP網際網路電話」

主辦機關:

中華電信行動通信分公司

聯絡人/電話:

陳月雪/23442808

出國人員:

駱襄陽 中華電信行動通信分公司 網路維運處 副工程師

出國類別: 實習

出國地區: 美國

出國期間: 民國 91 年 11 月 10 日 - 民國 91 年 11 月 23 日

報告日期: 民國 92 年 04 月 25 日

分類號/目: H6/電信 H6/電信

關鍵詞: 多協定標籤交換

內容摘要: 現今在電信界之規約何者可以存活，這可能是許多電信資深人員之問題，目前Everything(業務) over IP或者IP over everything已為各界所認知，但IP網路與現今使用之技術相比尚有不足，如時間延遲之掌握，流量控管之機制，Qos之能力等再再都點出IP網路之弱點，因此如何結合其它之技術解決上述問題即成為熱門話題。如何使 IP網路擁有如ATM網路之保證端對端之Qos、流量控制之機制、擁塞控制及頻寬分配與管理等特性，但又沒有ATM較複雜、缺少端點之介面及高價實網路卡之缺點，則成為各界思考之方向。

本文電子檔已上傳至出國報告資訊網

摘 要

現今在電信界之規約何者可以存活，這可能是許多電信資深人員之問題，目前 Everything(業務) over IP 或者 IP over everything 已為各界所認知，但 IP 網路與現今使用之技術相比尚有不足，如時間延遲之掌握，流量控管之機制，Qos 之能力等再再都點出 IP 網路之弱點，因此如何結合其它之技術解決上述問題即成為熱門話題。如何使 IP 網路擁有如 ATM 網路之保證端對端之 Qos、流量控制之機制、擁塞控制及頻寬分配與管理等特性，但又沒有 ATM 較複雜、缺少端點之介面及高價網路卡之缺點，則成為各界思考之方向。

目 錄

第一章、 前言.....	1
第二章、MPLS 概要.....	2
2.1 MPLS.....	2
2.2 MPLS 網路架構.....	2
2.3 MPLS 網內資料流.....	7
第三章、MPLS 標籤分配及信號.....	10
3.1 分配及信號概要.....	10
3.2 標籤分配之控制.....	11
3.3 標籤分配之啟動.....	13
3.4 信號規約.....	17
3.5 MPLS 訊務控制.....	19
3.5.1 RSVP-TE.....	21
3.5.2 CR-LDP.....	23
3.6 CR-LDP 與 RSVP-TE 之比較.....	26
第四章、MPLS 網路信賴度及復原.....	26
4.1 網路保護之需求.....	27
4.2 障礙偵測.....	27
4.3 網路保護機制.....	30
4.4 RSVP-TE 之保護機制.....	33
第五章、MPLS 訊務規劃.....	34
5.1 Over-Provisioning 及 Under-Provisioning.....	34
5.2 需求頻寬之計算.....	36
第六章、心得及建議.....	40

第一章、前言

現今在電信界之規約何者可以存活，這可能是許多電信資深人員之問題，目前 Everything(業務) over IP 或者 IP over everything 已為各界所認知，但 IP 網路與現今使用之技術相比尚有不足，如時間延遲之掌握，流量控管之機制，Qos 之能力等再再都點出 IP 網路之弱點，因此如何結合其它之技術解決上述問題即成為熱門話題。如何使 IP 網路擁有如 ATM 網路之保證端對端之 Qos、流量控制之機制、擁塞控制及頻寬分配與管理等特性，但沒有 ATM 較複雜、缺少端點之介面及高價質網路卡之缺點，則成為各界思考之方向。

MPLS 則在各方看好下產生，其主要目地為讓屬於第三層之 IP 網路可 over 多數之第二層技術如 Frame Relay、ATM 等，MPLS 全稱為 MultiProtocol Label Switch-多協定標籤交換。以下層之技術支援 IP 網路之 Qos 等等特性。MPLS 之前身為 CISCO 之 Tag 交換技術，而 MPLS 則為 IETF 國際標準組織所制定之規約架構。

而為了配合本公司 ATM 設備之引進，了解本公司未來提供新機能及新服務之準備方向，並吸收新知，擴大視野；故本次奉派赴美學習多協定標籤交換技術以及 IP 網路電話之相關知識，為期含行程共計 14 天。

第二章、MPLS 概要

2.1 MPLS

MPLS 為 Internet Engineering Task Force (IETF) 國際組織所制定之網路架構，其全稱為 Multiprotocol Label Switching (MPLS)。此架構支援有效的訊務設計、路由設定、傳送及交換技術使訊務有效率的流通於網路間。

MPLS 主要執行下列功能：

1. 規範並管理使多樣之資料訊務流通於不同之硬體、機種甚至應用層面的管理技術。
2. MPLS 為獨立於第二層及第三層規約以外之部份。
3. 提供一種技術使 IP 位址能轉換成簡單及長度固定之標籤以協助封包之傳送及交換。
4. 做為與現用路由協定之介面例如 Resource reservation protocol (RSVP)、Open Shortest Path First (OSPF) 等。
5. 支援 IP、ATM、Frame Relay 等第二層規約。

2.2 MPLS 網路架構

在介紹 MPLS 架構前先介紹在規約中用到的數個名稱，以利後續說明：

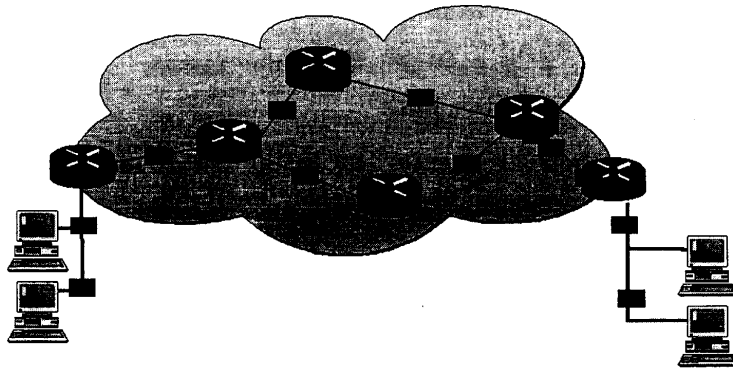
LER	Label Edge Router
LSR	Label Switch Router

LIB	Label Information Base
LSP	Label Switch Path
FEC	Forward Equivalence Class; also, Functional Equivalent Class

MPLS 網路運作概觀如下：

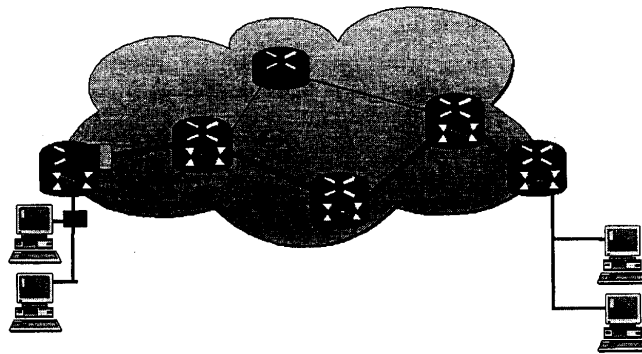
1. MPLS 以覆加於 IP 之外的規約使得 IP 可與其他第二層之技術結合。
2. 加裝 MPLS 機能之路由器則處理附加於 Shim Header 之 MPLS 標籤。
3. 純 IP 之訊務由 LER 進入，並由 LER 賦予標籤後，即可於 LSR 與 LSR 間流通。
4. 在 MPLS 網路之出口(egress)，由 LER 消除標籤，並以 IP 選路方式送出訊務。
5. 若訊務流經許多網路時，可以堆疊標籤方式建立 tunneled 機制。

在 IP 網路中可將路由器想像為一個郵局，在郵局若無一套標示、分類、監視檢查郵件之機制，則根本無從去辨識郵件之等級高低，亦無法決定何者須優先傳遞。IP 網路中亦有著類似的狀況，圖一為單純之一般 IP 網路示意圖



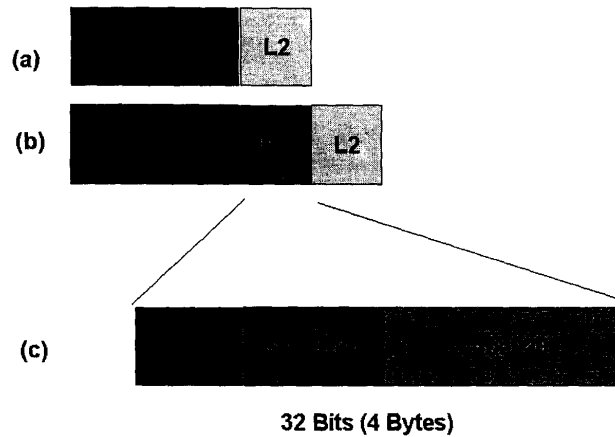
圖一、IP 網路示意圖

為了區別不同等級或者是不同優先序之服務業務，訊務必須在進入服務網路時先貼上特別之標籤標示其等級，以告知後續之處理者如何對待。在 MPLS 網路中貼標籤之工作由 LER (Label Edge Routers) 擔任。LER 作為 IP 封包及 MPLS 封包之轉換橋樑。在 MPLS 網路之入端 (ingress)，LER 檢視進入之訊務並決定是否加標籤，且在 LER 中會有一資料庫以作為目的位址與標籤之轉換用，圖二即為 IP 封包加上標籤之示意圖，



圖二、封包加標籤示意圖

而在 MPLS 之標籤即為 MPLS shim header，其內容如圖三

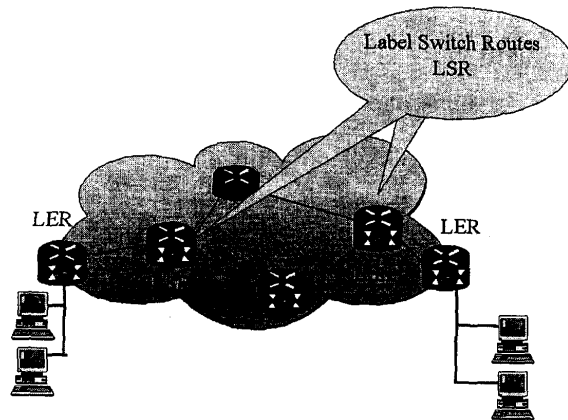


圖三、 The MPLS Shim Header 及格式

Shim Header 介於 OSI 模型的第二層與第三層間，分為 4 個部分共占 32 個 Bits。其中有 20 個 Bits 作為標籤用，3 個 Bits 歸實驗用，1 個 Bit 給堆疊功能，8 個 Bits 決定封包的生存期限(time to live)。

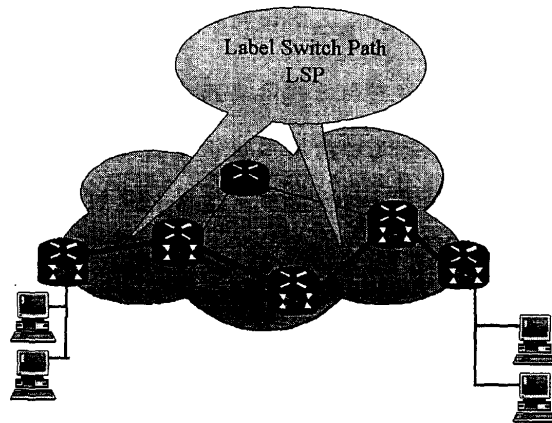
一旦標籤分貼好後則由非邊界之路由器執行 LSR(Label Switch Routers) 之角色，LSR 之功能為檢視進來封包之標籤，解析其內容並依該內容之指示選擇下一傳送路徑，通常 LSR 並會置換標籤內容，圖四顯示網路內之 LSR

封包傳送路徑由 LER 及 LSR 建立，這些路徑稱為 LSP (Label Switch Paths)，為依其訊務特性設計，非常類似建立 ATM 之路徑。而每一 LSP 所能承載之訊務量皆經過精算，其計算之依據有忙時訊務量、訊務量變動率、封包遺失率等等



圖四、LSR

圖五顯示 MPLS 設備建立之 LSP，由於 MPLS 不屬於 IP 之網內，因此可共存而不互相干擾

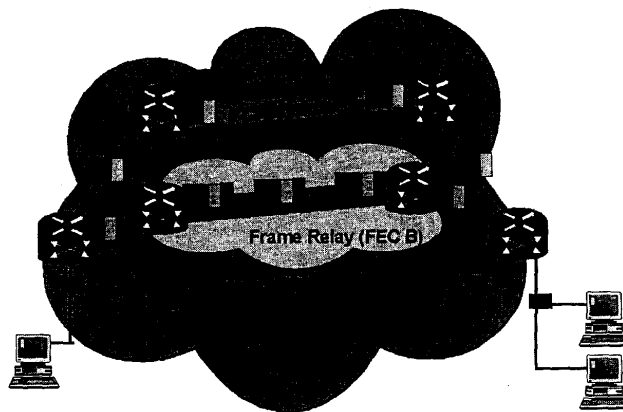


圖五、LSP (Label Switch Paths)

如此說來，MPLS 到底有何好處？

若一間公司其資料庫使用上正受資料遺失或不穩定之苦，為了確

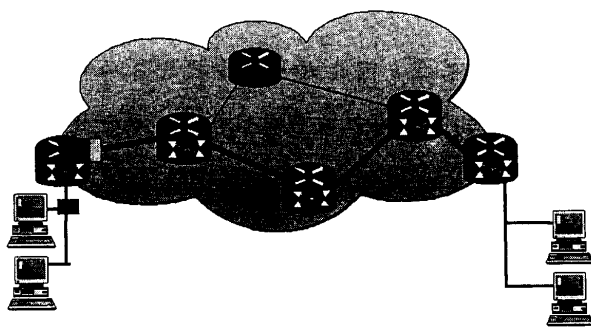
保重要資料之安全，公司可能會承租高單價高頻寬之電路因應，而該電路必須以最高標準估算，因此大部分時間其有大量之頻寬為閒置。若使用 MPLS 則可由 LER 將封包以重要性排序，最重要之部分走高單價穩定之電路，而一般例行訊務則走其他較便宜路徑，亦即在 MPLS 中可對訊務作分類，LER 將訊務分成不同的 FECs (Forward Equivalence Classes)，再分別傳送，如圖六所示



圖六、MPLS 網內之 FECs

2.3 MPLS 網內資料流

當單純之 IP 訊務出現在 LER 時，LER 先將訊務以例行性、優先等級等方式分類，並貼上不同標籤，此分類方式稱為 forward equivalence class 簡稱 FEC。LER 使用多種不同的模式來區別訊務，下列表一為簡單例子，LER 之 FEC 使用預先規劃好之表格為撥接之 IP 訊務分類，如圖七及表一所示。



圖七、Ingress LER 貼上 Shim Header

Destination / IP	Port Number	FEC	Next Hop	Label	Instruction
191.150.3.1	81	B	x.x.x.x.	20	Push
191.150.3.1	243	A	y.y.y.y	34	Push
191.150.3.1	25	IP	z.z.z.z		(Do nothing; native IP)

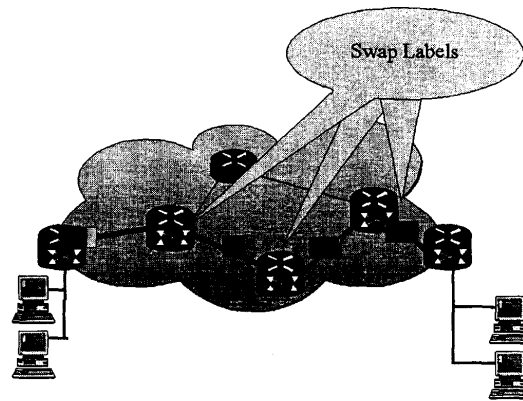
表一、LER Instruction Set

當 MPLS 封包離開 LER 並進入 LSR 後，LSR 僅檢視該封包之標籤，並比對 LSR 本身送出表內容以決定如何送出，該送出表稱為 Label Information Base [LIB] 或連結表(connectivity table)。LSR 會根據 LIB 之指令置換標籤如表二所示

Label/In	Port In	Label/Out	Port/Out	FEC	Instruction Next Hop
20	B	60	D	B	Swap
34	A	72	C	A	Swap

表二、Label Information Base (LIB)

圖 8 為 LSR 標籤置換功能之展示。



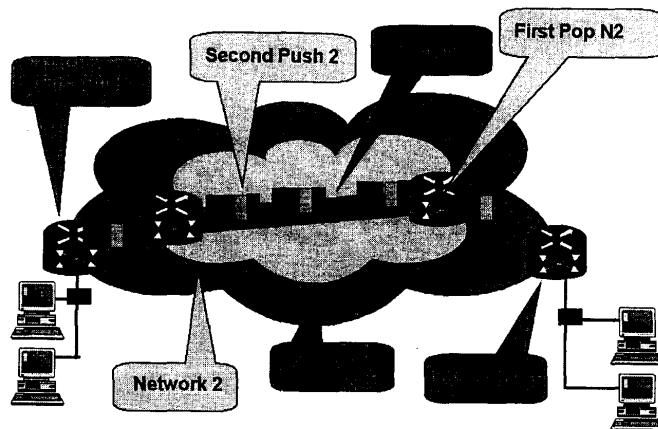
圖八、label-swapping functions.

MPLS 封包進入終端 LER 後即由其將 MPLS 之表頭去除將 IP 封包送入 IP 網路中。

LER 執行許多封包分析之功能，例如：將 Layer 2 對應到 MPLS、將 MPLS 對應到第三層、訊務分類等等。LER 同時還需要判別何者為 MPLS 訊務，通常 LER 使用 Triggered mode 來判定，所謂 Triggered mode 即對特定目的點之訊務在某一個時間範圍內流入者即視為 MPLS 訊務。

為了使 MPLS 能更彈性運用因而有所謂的 the label-stacking method，如圖九所示。

在圖九中若你屬於網路一的成員，但網路一與網路二分屬不同公司，但網路一之訊務又須經由網路二傳送，此時網路一必須確認網路二為根據雙方所簽署之 service-level agreements (SLA) 來傳遞網路一之訊務。但此兩網路之標籤又不盡相同，在此情況下則需使用堆疊式(stack labels)之標籤以建立隧道(tunnel)之方式使網路一之訊務通過網路二，這種方式不但保存了本身之標籤，又能使其他之標籤獨立運作。



圖九、Tunneled 網路中的標籤堆疊

第三章、MPLS 標籤分配及信號

3.1 分配及信號概要

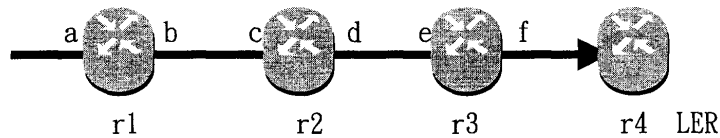
早期之電路交換接線生須熟習許多之規則，到了 MPLS 仍然如此，只是接線生之角色由自動化之設備取代，如路由器、交換器等。而這些設備仍須經過學習之階段，學習之目的為如何建立一可靠之交換路徑以使發話端能接續至被叫端。此種過程約可分為兩類，其一為固接式使用程式化硬體達成，如路由器之 Static routing 即是；另一種則為近代較風行之動態式交換路徑學習。

動態式交換路徑學習，即動態的信號及標籤分配。目前有許多之規約可提供該項功能，但也各有優缺點，而目前也尚未有一統天下之規約共識，因此如何選擇則屬各自之判斷。唯不論如何，標籤分配及信號都屬各項規約不可缺少之功能。

在 MPLS 之動態規約中，各轉送點至少需學習到如何處理攜帶標籤進來之封包，此項處理在 MPLS 中通常由 cross-connect table 達成。

例如由 Port A 進來攜帶標籤為 241 之封包，將由 Port B 轉送出，同時將標籤置換為 304 即為 cross-connect table 之功能。而 cross-connect table 優於 Routing table 之處在於前者可於 Data link 層處理，一般而言其速度將較後者為快。

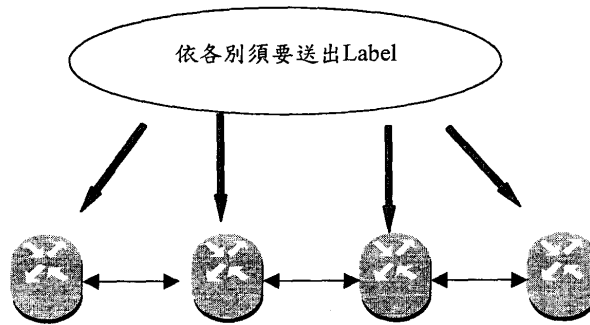
以下舉一網路例說明。在本例之網路中擁有 4 個路由器分別以 r1、r2、r3 及 r4 表示，各路由器中各有適當數量之 Port，各以 a、b、c、d、e 及 f 表示本例使用之 port，資料流向則由 router 1 之 a port 進入，最後由 router 3 之 f port 流出並送入 router 4，如圖十所示，而該網路如何升級至以 MPLS 信號處理？



圖十、 Basic MPLS Network with 4 Routers

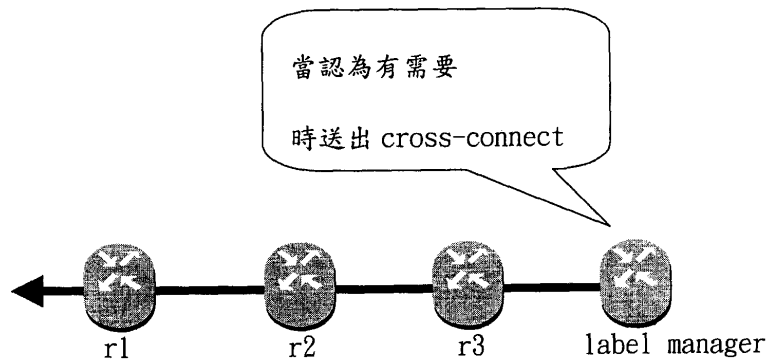
3.2 標籤分配之控制

標籤分配有兩種不同之模式，第一種為獨立控制者，每一顆 router 都可以收取 routing table 訊息將之置入本身之 cross-connecting table，並通知其他之 router 有關本身之訊息。而這些 router 都各自獨立運作，亦即所有 router 都可收取網路上傳送之路由之規約訊息建立自身之 cross-connect tables 並將之再轉送出去，如圖十一。



圖十一、獨立控制式標籤分配

另一種模式稱為 ordered control 如圖十二所示。在 ordered control 模式下，只有一個 router(通常為 egress LER)具有分配標籤之責任或稱標籤管理者



圖十二、Ordered Control (pushed)

前述兩種模式各有值得取捨處。獨立控制之網路當任何 router 收聽到 routing 之變化時可轉送給其他所有 router，因此網路收斂速度較快。但因沒有訊務主控點，使得控制設計較為困難。

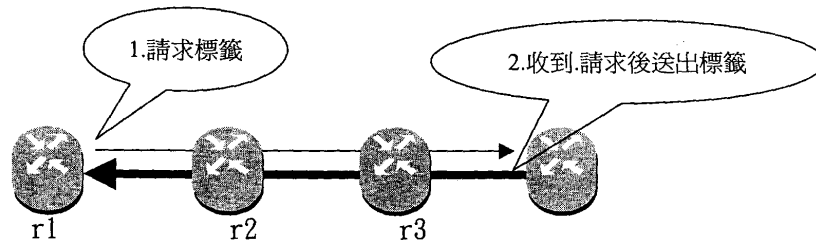
orderd control 模式則對訊務控制設計較為容易，並使網路控制較為緊密，但其缺點則是收斂速度較慢，且若該點障礙時也會提昇控制上的風險。

3.3 標籤分配之啟動

在 ordered control 模式中有兩種主要的方法啟動標籤的分配，第一種稱為 down-stream unsolicited，第二種稱為 down-stream on demand。

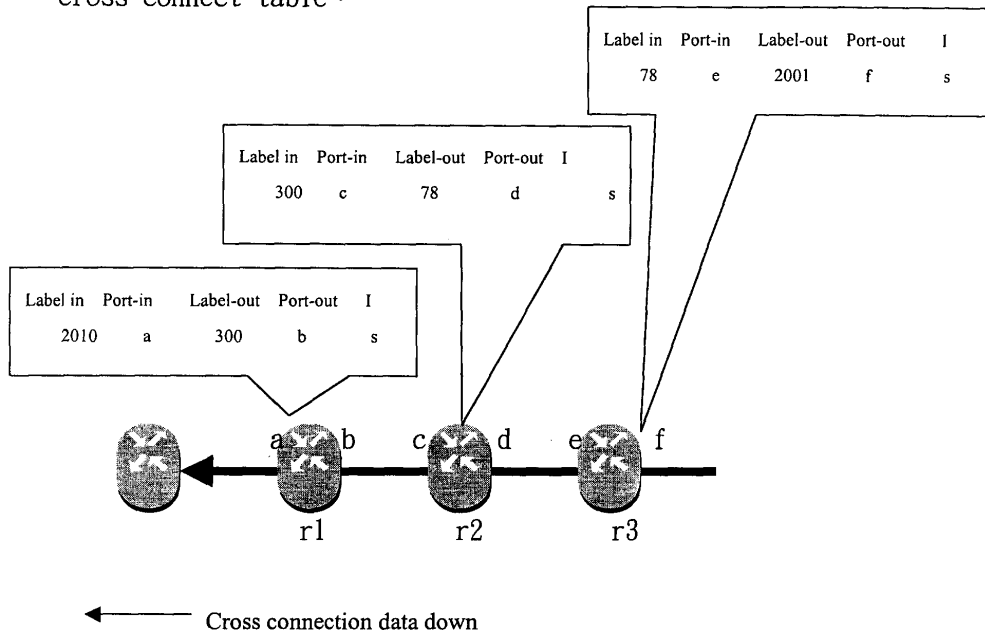
down-stream unsolicited 簡稱(DOU)，在圖十二中可見到標籤送至 down-stream routers，而此送出之動作是基於標籤管理路由器之決定，當標籤非因其他之要求而由標籤管理者主動送出，此型式稱為 down-stream unsolicited (DOU)。例如標籤管理者可訂定啟動點(如一段時間)送出標籤或每 45 秒送出更新訊息，或標籤管理者可於每次標準之路由表格異動時送出標籤更新訊息給所有相關之路由器。

down-stream on demand 則為當需要標籤時才執行，如圖十三中在第一步驟使用者提出標籤需求後，第二步驟才是管理者送出標籤。



圖十三： Down-stream on Demand (DOD)

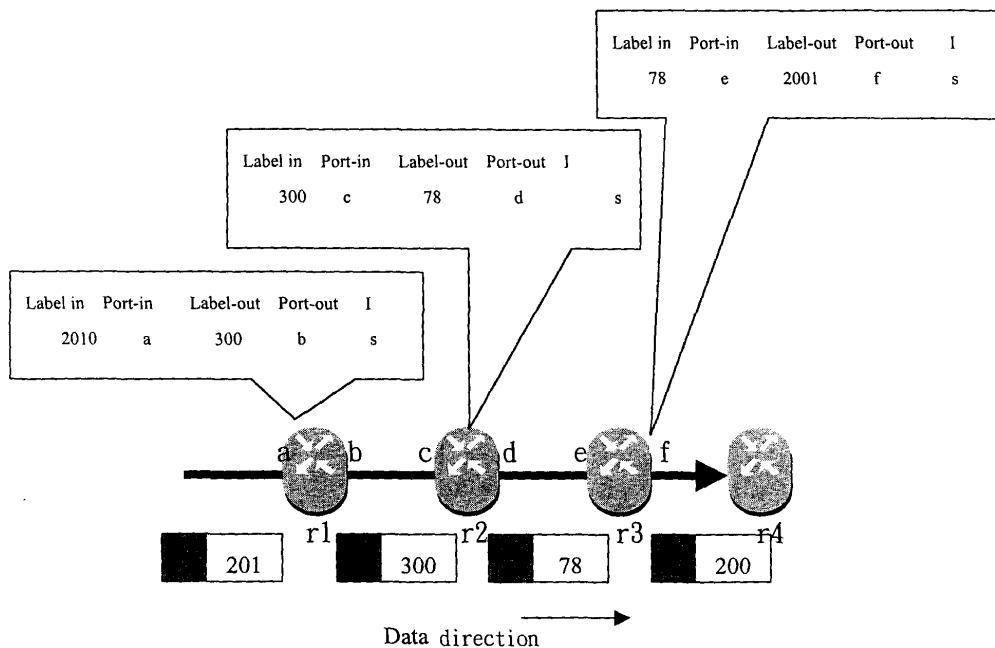
不論標籤是經由 DOD 或 DOU 方式取得，LSR 都會建立一個類似圖十四之 cross-connect table。



圖十四：LSR with Cross-connect Tables Populated

連結表格由 r3 送至 r1，而其欄位包含 label-in, port-in, label-out, port-out, and instruction (I)，在此例中之指令為置換 (swap)，而重點是標籤及 cross-connect table 即在說明路由。一旦 cross-connect table 建立後，訊務資料才能由各路由器置換標籤並沿著 r1 送至 r4。

圖十五則顯示在 cross-connect table 建立後訊務資料沿著設定好的 LSP(label switch path)由 r1 送至 r4



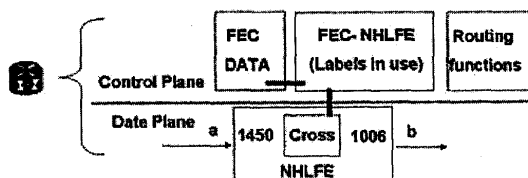
圖十五： Data Flow on LSP

重述前面之講法，路由器需要 cross-connect table 以作為交換路徑之依據，而 cross-connect table 可由鄰近之路由器依據獨立控制或標籤管理等控制方式獲得。而標籤管理者可依要求(down-stream on demand)或其本身認為有需要時(down-stream unsolicited)向前位局路由器送出標籤，此外尚需考慮更多之觀點如標籤如何送給路由器、由什麼媒介傳送而服務等級(quality of service)資訊又如何傳送給路由器等等。

在前述中有提及 MPLS 封包攜帶標籤但封包中並無欄位告知路由器該封包之 QOS 資訊，然前述亦提及訊務可分類成不同之 Forward Equivalence Classes (FECs)，而 FEC 可於 LSP 中指配，吾人可於訊

務設計時強迫高優先序之 FECs 走高品質之 LSPs，而低優先序之 FECs 走較低品質之 LSPs，而不同訊務對 QOS 標準不同也使得標籤分配之對應更顯複雜。

圖十六顯示 LSR 內部之處理圖示，在 LSR 內分為兩面，一面為資料面，一面為控制面。若一個由 port a 攜帶標籤為 1450 之封包進入 LSR 經由標籤置換為 1006 且由 port b 送出封包，上述過程為 cross-connect table 之功能，而該表格亦可稱為 *next hop label forwarding entry table* (NHLFE)。



圖十六: A Closer Look at the Router

上述資料庫並非一獨立之資料庫，而是由控制面的 FEC 及 FEC-to-NHLFE 兩個資料庫連結而成。FEC 資料庫包含目的點之 IP 位址、訊務特性、封包處理要求。而 FEC 又需與標籤相連結，FEC 與標籤相連結之處理程序稱為 binding。

表三為 FEC 資料庫與 FEC-to-NHLFE 資料庫之一例

FEC Database

FEC	Protocol	Port	
192.168.10.1	06	443	guaranteed no packet loss
192.168.10.2	11	69	best efforts

192.168.10.3 06 80 controlled load

FEC to NHLFE Table

FEC	Label in	Label out
192.168.10.1	1400	100
192.168.10.2	500	101
192.168.10.3	107	103

NHLFE Table

Label in	Label out
1400	100
500	101
107	103

表三、FEC 資料庫與 FEC-to-NHLFE 資料庫

因此當攜帶標籤之封包進入資料面，而該標籤於 FEC 存在時即能快速的處理。由此可見在資料封包能夠被處理前必須先有前置程序以建立 cross-connect table。

3.4 信號規約

如何找到一種傳送方式建立上述之表格為設計者最關心的問題，而該規約必須符合能即時、快速、具有自我回復及高信賴度的承載所有需要之訊息。制定 MPLS 規約的工作群組則以定義一新的規約因應，稱為標籤分配規約(Label Distribution Protocol)，簡稱 LDP。

在 LDP 中，當一個標籤配出後即會一直存在，除非有另一指令將其取消，此種特性使其不需要經常更新。因此 LDP 規約屬於隱含式之路由方式。

但也有其他之群體認為使用新的又未經測試之規約不如使用現成已有之規約加以修改來的實際，故一些現存之路由規約已被修改成可攜帶標籤之規約，而如 Border Gateway Protocol (BGP) 及 IS-IS 已被認可攜帶標籤且不影響其原有路由規約之功能。

不論是 LDP、BGP 或是 IS-IS 都可建立 LSP(Label Switch Path)，但對訊務設計則很少著墨，原因是當出現訊務壅塞時可經由更高等級之 LSP 迂迴訊務。

為解決此問題另一信號規約出現，用建立訊務隧道(Tunnels)方式(此為顯性路由)以更簡便的包含訊務規劃，此規約如 Constraint Route Label Distribution Protocol (CR-LDP) 及 Resource Reservation Setup Protocol (RSVP-TE)，而 OSPF 亦曾經修改為 OSPF-TE，只是現今鮮少人使用。

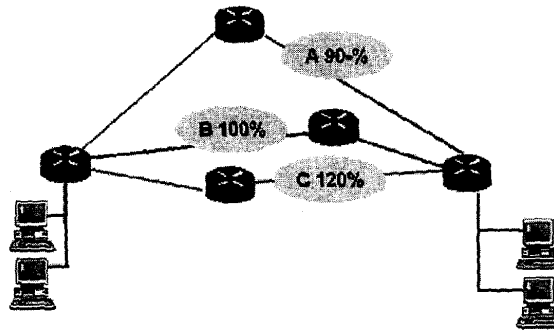
表四歸納上述之說明

Protocol	Routing	Traffic engineering
LDP	Implicit	NO
BGP	Implicit	NO
IS-IS	Implicit	NO
CR-LDP	Explicit	YES
RSVP-TE	Explicit	YES
OSPF-TE	Explicit	YES

表四、各信號規約之路由型式及訊務設計

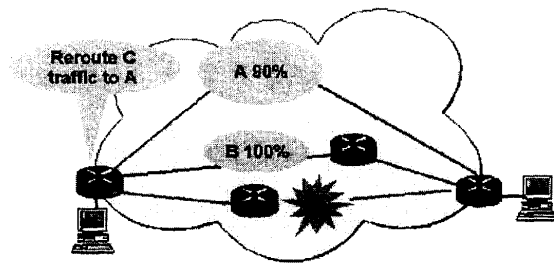
3.5 MPLS 訊務控制

在 MPLS 網路中可能會因為一些例行之訊務佔用高效通道而產生壅塞及速率下降等情事，由於 MPLS 是架設在一般 IP 網路上，因此在意外障礙無法預期下 IP 網路的自我修復機能針對資源調整部分，可能使得原先之一般訊務因路由調整而導入高效路由。而如何避免已被預留之路徑及頻寬被一般訊務佔用即成為一重要課題。

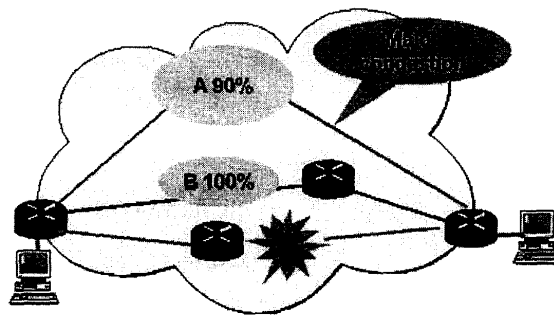


圖十七

圖十七為一擁有三條經過 WAN 不同路徑之 MPLS 網路，其中路徑 A 為規劃在忙時訊務佔用 90%頻寬，路徑 B 為忙時訊務佔用 100%頻寬，路徑 C 為忙時訊務佔用 125%頻寬。理論上言，路徑 A 不會有壅塞之情況，因為其設計量高於忙時之使用量；路徑 C 在未規劃重荷處理之情形下在忙時一定會壅塞。因此路徑 C 之 QOS 即可能變得不可預期，而路徑 A 的 QOS 則可獲得很好之控制。



圖十八、 MPLS with a Failed Path C



圖十九、 MPLS with Congestion Caused by a Reroute

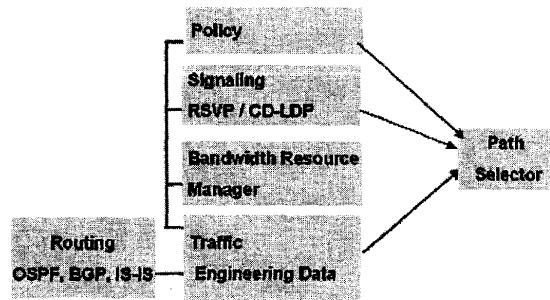
想像在圖十八中若路徑 C 發生障礙時，路經路徑 C 之訊務其路由將重新調整，假設為重新導入路徑 A(如圖十九)，則原有路徑 A 之 Qos 將很難確保，因此為保障 Qos，必須有方法來控制路徑上的訊務量，及規劃每一路徑上之承載百分比。

為了解決路由重新調整造成之訊務壅塞情形，IETF 之相關工作群組試圖尋找可能之解決方案，而其重點則在規約及架構在路由器上的軟體系統。

為了有良好之 Qos，系統必須能夠對訊務作註記、分類、篩選等動作。由前述之敘述範圍中 MPLS 可經由標籤對訊務作註記及分類，但是

卻獨缺篩選的功能。選路及標籤分配建立 LSP 時卻無法限制訊務及對鏈路上之訊務作控制。

有鑑於此，需在路由管理軟體的路徑選擇功能中增加篩選及頻寬管理之機制，此機制附加於兩種規約即可達成訊務限制功能。



圖二十、MPLS Routing State Machines

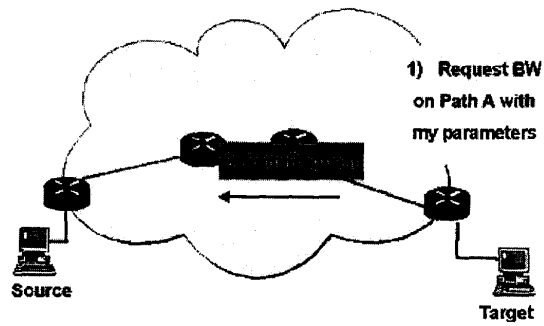
此兩種可以限制及控制 MPLS 訊務的規約為 RSVP-TE 及 CR-LDP。

3.5.1 RSVP-TE

資源須於呼叫前預定之概念與原有之電話呼叫概念雷同，當 Qos 成為資料網路之議題時此概念也適用之。

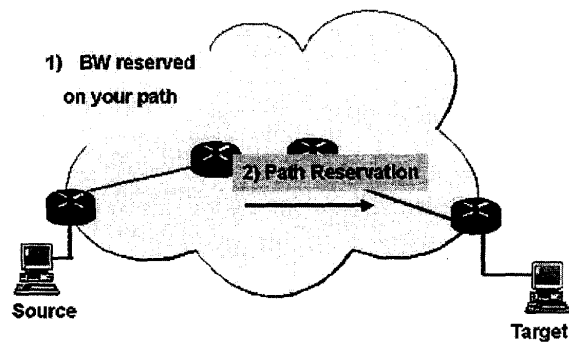
IETF 早於 1997 即設計 Resource ReSerVation Protocol (RSVP) 執行此功能，此規約於定義呼叫路徑時即要求適當的頻寬，若當時頻寬合乎所需即可建立鏈路。此鏈路將訊務分類為最高級 (respectively)、保證級 (guaranteed load) 及盡力級 (best-effort load)。

而 RSVP 加入 MPLS 訊務規劃機能後稱為 RSVP-TE。



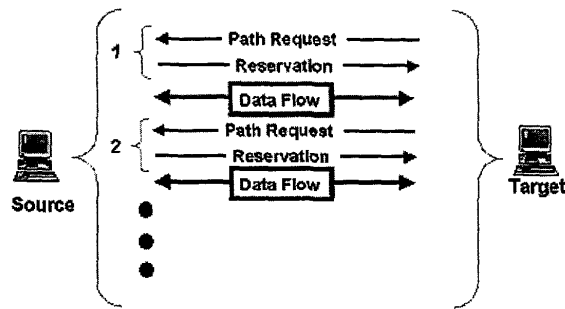
圖二十一、RSVP-TE 路徑建立要求

於圖二十一及圖二十二中可看到如何在兩個端點間建立路徑。被叫端點將路徑所需之訊務條件、對待參數等包含於路徑建立要求之訊息中，網路端收到此訊息後在網路中保留適當頻寬，並告知被叫端，一旦被叫端收到該訊息，訊務即可開始於兩端流通。



圖二十二、RSVP-TE 路徑保留

呼叫建立及信號處理稱為 soft state，因為在更新時限內未更新其狀態則呼叫會被終止。因此在圖二十三之路徑要求及保留之訊息會在資料流通之過程中持續送收，以維持資料傳送路徑。



圖二十三、 RSVP-TE 路徑保持

早期曾有對 RSVP 在規模、路徑建立、會產生更多更新路徑之訊息、更新路徑之訊息造成網路重荷等方面提出質疑，而解決之方式則為不允許訊務鏈路及路徑切割成太小及聚集相似之路徑。

3.5.2 CR-LDP(Constraint-based Routing over Label Distribution Protocol)

將 LDP 規約修改加入訊務之說明即成為 CR-LDP，亦即在現存之 LDP 規約中加入訊務規劃之能力即成。CR-LDP 在 LDP 規約中加入 peak、committed、excess-data rates 等欄位，非常類似 ATM 使用之項目。

以下概略說明依 CR-LDP 建立 LSP 之過程

IP 封包在進入 MPLS 網域前會根據該封包所攜帶之信息將之分類成許多不同之類別，而這些類別即為 MPLS 網路為該封包所提供的服務等級之分類，而 MPLS 網路之入端 LER 依據下列之內容將其分類：

- (1). 承載 IP 封包之 DLCI 及 VCC 等之相關資料
- (2). TOS 或 DS 所攜帶之內容

- (3). 訊息來源或目的點所使用之 port 編號
- (4). 訊息來源或目的點之 IP 位址
- (5). 該訊息使用之 IP 上層規約，如 UDP 或 TCP 等

LER 對 IP 封包分類後，當欲建立 LSP 時即將 Qos 的參數塞入所使用之 Label Request 訊息的 Traffic parameter 中，而在 MPLS 骨幹網路中的 LSR 在接收到 Label Request 訊息後即根據其內之 Traffic Parameter TLV 參數之內涵作連結許可控制(Connection Admission Control)，為該特定之業務預留需求之頻寬，在 CR-LDP 之 Label Request 訊息中有 FEC 提供疏轉訊務等級分類、ER-HOP 提供路由資訊、Qos 提供 Traffic Parameter 等 TLV，圖二十四 提供 CR-LDP 之 Label Request 訊息之格式。

U	F	Label Request	Message Length
Message ID TLV			
Return Message IDTLV			
FEC TLV			
LSPID TLV			
ER-TLV			
Traffic Parameters TLV			
Pining TLV			
“Ressource Class” TLV			
Pre-Emption TLV			

圖二十四、CR-LDP Label Request 訊息格式

Label Request 訊息順著 ER-HOP 標示的路徑向其後之路由轉送點流過，而所有經過的 LSR 都會根據 Label Request 中的 Traffic Parameters TLV 的指示為該 LSP 執行連結許可控制，以預留頻寬及建立轉送之出口 port。該 Label Request 訊息一路送至出端 LER(Ergess LER，即 Label

管理者)，Label 管理者則回送一個標籤對映(Mapping)訊息通知提出 Label Request 之 LER(Ingress LER)，告知其所申請之 LSP 之網路資源已保留。

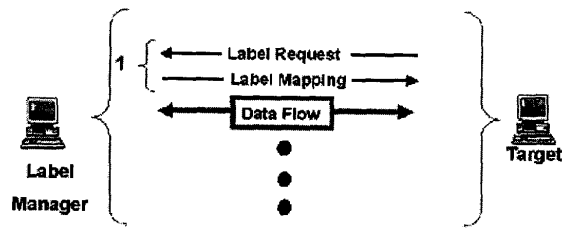
在上述程序中 Qos 的維持即是透過 Traffic Parameters TLV 來達成，而此 TLV 中即涵蓋了建立 LSP 所要求的 Qos 的參數，其內容如圖二十五所示。

U	F	Traffic TLV	Message Length	
Flags		Frequency	Reserved	Weight
Peak Data Rate(PDR)				
Peak Burst Size(PBS)				
Committed Data Rate(CDR)				
Committed Burst Size(CBS)				
Excess Burst Size(EBS)				

圖二十五、 CR-LDP 格式

上圖中，具體表現了該使用 IP 之業務對服務品質之要求，如 Frequency 為時限參數，Peak Data Rate(PDR)為該業務最高需求之頻寬，Committed Data Rate 為該業務之保證頻寬，這些參數提供了 MPLS 之 Qos 之機制。

在上述 CR-LDP 之 LSP 建立過程中控制訊息總計分兩階段處理，標籤要求及標籤對應，如圖二十六所示。而 CR-LDP 之所以簡單是因為其屬於 hard-state protocol，呼叫路徑一旦建立後除非再次要求拆線，否則永遠有效



圖二十六、CR-LDP 呼叫建立

3.6 CR-LDP 與 RSVP-TE 之比較

以下表五以條列方式比較前述兩者

項目	CR-LDP	RSVP-TE
Vendors	Nortel	Cisco, Juniper, Foundry
State	Hard State	Soft State
QoS Type	ATM	IntServ
Recovery Time	A little slower	Faster
Chat Overhead	Low	High
Transported on	LDP over TCP	RSVP on IP
Path Modifications	Make before break	Make before break

表五、 CR-LDP 與 RSVP-TE 之比較

雖然在現今之製造商 Cisco, Juniper 偏好 RSVP-TE，而 Nortel 支持 CR-LDP，但大部分之設備製造商都能提供此兩規約。

第四章、MPLS 網路信賴度及復原

在電信網路中當一條電路被終止前必須找出一條相同品質之替代電路以疏轉原電路所承載之訊務，也就是服務不能中斷且品質不能打

折。在傳統的 IP 網路中 IP 封包會自行尋找路徑以繞過有問題之節點，通常不會有影響。但當語音在 IP 網路中傳遞時就必須標示出障礙時可使用且現存之替代路徑，以避免影響有即時性需求之訊務。

4.1 網路保護之需求

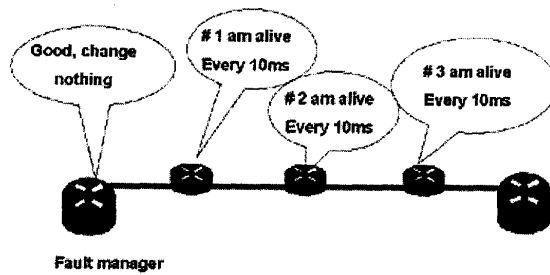
MPLS 已在可在 IP 的核心網路運行，通常其對障礙之處理，一般會將之與 ATM 網路之障礙處理逐項比對，期望其能達到 ATM 對障礙處理之能力。

在一般 IP 網路通常期待一路徑之障礙可於數秒或至少於數分鐘內完成切換，但 MPLS 則要求切換必須於 60ms 內完成，且不可影響原有服務之品質。因此對 MPLS 網路之復原程序之挑戰為如何於 60ms 時間內發現問題並完成切換之動作。

4.2 障礙偵測

前面提到 MPLS 網路復原程序必須在很短時間內偵測到障礙，目前有兩種方式可行：heartbeat detection (or polling) and error messaging。heartbeat detection 可以很快速地找出網路問題，但卻需使用大量的網路資源；而 error messaging 方式對資源之需求較少但相對的其速度亦較慢，如何取捨則取決於成本與速度間之權重。

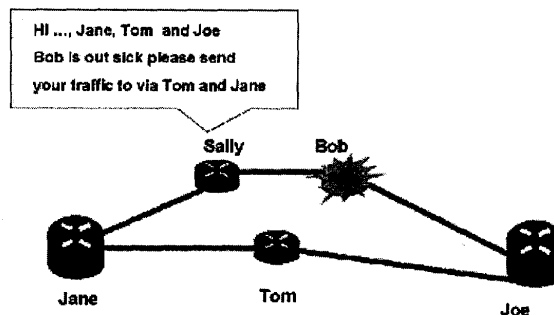
heartbeat detection 使用簡單之方法偵測障礙，如圖二十七所示。



圖二十七、Heartbeat Method

網路上之設備在一固定之時間內向網路管理者報告本身仍存活著，若逾時未報告者則由網路管理者宣告其路徑、鏈路或節點為障礙，則切換動作開始執行，此法需要佔用相當大之訊息訊務量，當切換需要 40ms 的工作時間時，則必須於每 20ms 發送一次訊息，如此才能達到 60ms 之要求。

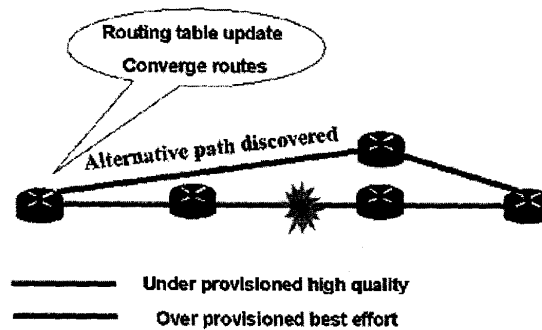
另外一種障礙偵測系統稱為 error-detection 如圖二十八所示，當網路上之設備偵測到障礙時即對其鄰近節點發送訊息告知對方重新指定路由或重新調整話務路徑方向，許多之路由規約使用方本方式。而本方式之缺點為送出 error-and-redirect 訊息給其他網路設備時需花許多時間，另一不利處則為錯誤訊息有可能永遠到不了前位局。



圖二十八、Error Message

在不要求時限之情況下 error-detection 之方式可工作的很好，但若時限是一個很重要的因素時，則選用 heartbeat 方式為較佳之選擇。

在使用 IP 之網路中(見圖二十八)資料為 connectionless 在沒有 Qos 之前提下，封包在不同網路間流通主要是透過路由器及路由表。若一條鏈路或路由器障礙時，可找到另一替代路由並送出訊務，若過程中封包不幸遺失，則由第四層之規約如 TCP 將遺失資料重送。

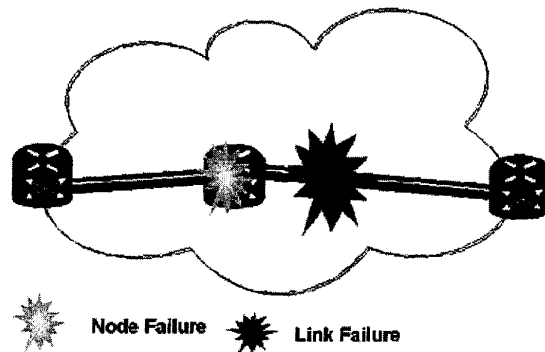


圖二十八、 Standard Routing

此種方式可適用於非即時系統，但當其傳送之資料為即時資料時如語音或影像，則延遲及封包遺失是不可忍受的，對於上述問題 OSPF 及 IGP 制定之工作群組已發展出 IGP 快速收斂之技術，將路由之障礙造成網路之影響限制於一秒中內。對於 IGP 之快速收斂技術只解決了 MPLS 一半的問題，至於 Qos 問題則尚待處理。

4.3 網路保護機制

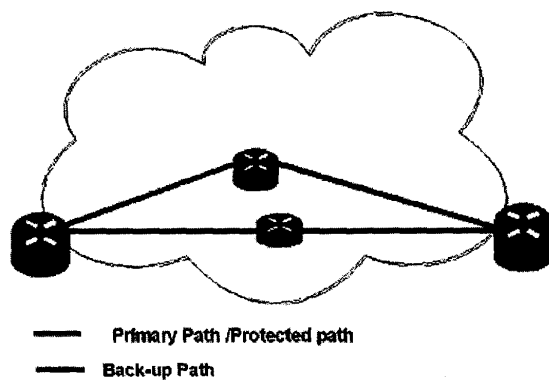
在網路中可能造成障礙的點有許多，其中二個主要的地方為鏈路及節點(參考圖二十九)，而主要的障礙有可能是交換硬體設備、交換軟體程式、資料庫及鏈路能力下降等。



圖二十九、網路障礙

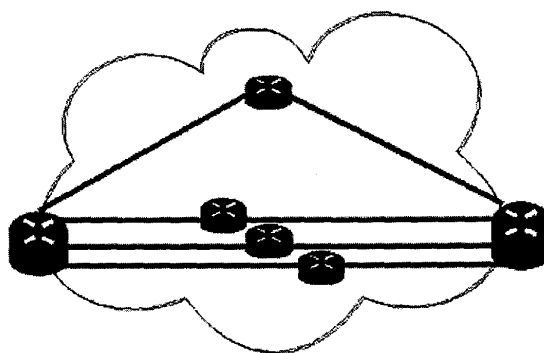
電信界傳統對於網路鏈路障礙之解決方案設計通常採用一對一之備援或一對多之備援方式處理，而另一種常用之觀念則為硬體之備援。

在 MPLS 網路之一對一備援方式則以預先預留同等級之 Qos 及訊務處理能力之路徑，而此路徑通常為閒置並可執行相關之測試程序，惟在被其保護之主路徑正常時該備援路徑通常處於離線狀態如圖三十所示，此種方式有很好之信賴度，但相對的成本則較高。



圖三十、一對一備援

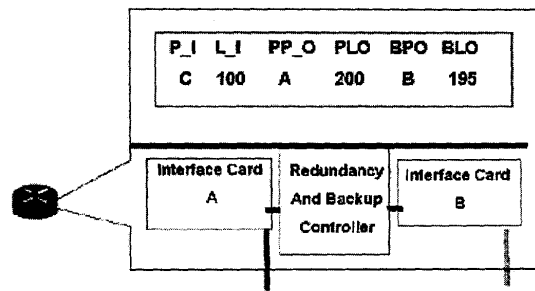
另一種一對多之備援方式如圖三十一 所示



圖三十一、一對多之備援

此種方式為當一條路徑障礙時由備援路徑取代，如上圖。但此法之風險為當網路上產生第二個障礙時則無備援設備可使用，因此只能備援單一障礙。

第三種備援方法則為用硬體備援模式，如圖三十二所示。



圖三十二、硬體備援模式

每一元件皆有另一備援硬體設備，從電源卡、控制卡、介面卡皆如是，如上圖為顯示控制卡之備援圖示，此法可應付硬體設備之障礙，但當 cross-connection table 出問題時，則此種方式無解決之能力。

從前述已了解備援之模式及偵測網路障礙之方法後，再來由哪一層及使用何種規約來實現上述之程序則為另一重點。

從 OSI 之架構而言，執行偵測障礙及回復網路能力之功能架得越上層，則其處理時間將托得越長，因此對於偵測及回復之功能架在越底層越好，最號是在實體層即能處理。

MPLS 於底層之錯誤偵測及更正可依據其下層承載之實體系統原有之功能，如 SONET ring、ATM、Frame Relay 等等，但對於 MPLS 中之鏈路之 Qos、tunnel、MPLS 呼叫路徑等之保護無法由底層之錯誤偵測及更正等機制來達成，MPLS 需保證其備援路徑之 Qos 與原路徑擁有相同之品質，而若訊務承載出現問題，MPLS 亦須有能力將低優先續之訊務移走。

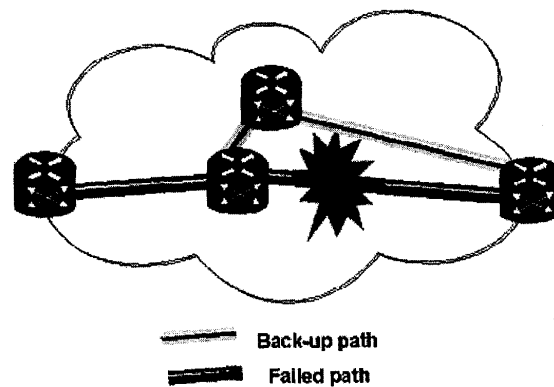
針對 MPLS 上述之功能，回過頭來檢視目前 MPLS 使用的兩種錯誤偵測方法。

在 LDP 及 CR-LDP 規約中包含一錯誤訊息欄 error message TLV，以報告鏈路或節點之錯誤，此法有兩項相當不利之因素：第一，傳送 error message 需耗費許多時間；第二，LDP 屬於連結導向 (connection-oriented) 之訊息，當鏈路障礙時告知之訊息根本無法送達。

另一方訪則是在 RSVP-TE 規約中使用的 heartbeat 方式，RSVP-TE 屬於顯示狀態之規約需定時更新，亦即當更新資料無法送達時即表示鏈路中斷，無須 error message 只要未收到更新資料即將路徑切換至預先設定之備援路由，許多的設備商都採用 RSVP-TE 執行呼叫及 tunnel 之備援。

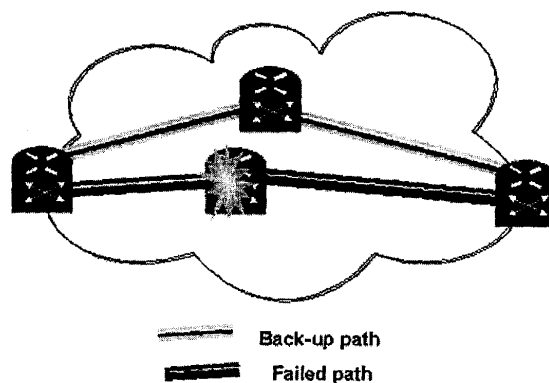
4.4 RSVP-TE 之保護機制

在 RSVP-TE 之保護機制中主要是保護鏈路及節點。在鏈路保護之機制為採用預先設定之備援鏈路，只要鏈路之障礙發生則馬上切換到備援鏈路上(如下圖所示)



圖三十三、RSVP-TE 鏈路保護機制

針對節點之障礙，RSVP-TE 採用預先建立之 tunnel 來避過障礙之節點如下圖所示



圖三十三、RSVP-TE 節點保護機制

第五章、MPLS 訊務規劃

5.1 Over-Provisioning 及 Under-Provisioning

在資料及語音網路中，訊務規劃之主要目的為將訊務導向可用的資源處。若能簡易的移動訊務而使得網路能平穩的運轉，則速率下降及忙時將成為空響。

在網際網路上對訊務之規劃分為四個階段：查測、特性分析、模型建立、移轉訊務至適當地點。

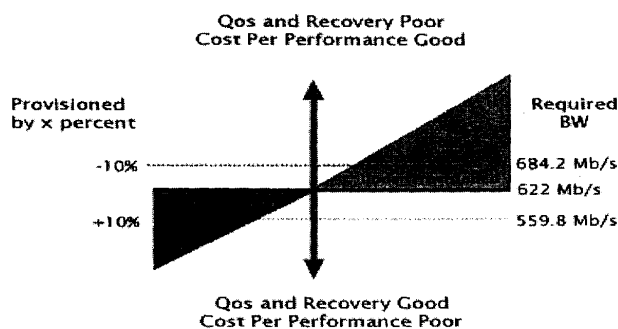
在找出 MPLS tunnel 話務設計之簡易公式前必須先討論頻寬之提供及需求。在此先下一定義，所謂 Over-provision 表示網路資源所擁有之頻寬大於網路之需求頻寬，under-provision 表示需求頻寬大於系統所

能提供之頻寬。而相對 provision 之另一字眼為 subscribe，所謂 Over-subscribing 表示需求頻寬大於所能提供者，under-subscribing 表示網路所擁有之頻寬大於需求頻寬。

一般在強調頻寬時指的是流經實體線路上所有之有效及無效之 bits 之總和，如 OC-12 之頻寬為 622Mb/s 表示其能承載之總流量。而要了解訊務在某一階段之流量時可以用一些量測工具或軟體達成，因此也可計算出每秒之實際需求量，再加上一些變動因素之 overhead 即為訊務估算之基礎，而 overhead 之估算為最困難之處。

若設計於 OC-12 上之訊務為 622Mb/S 稱為 fully provisioned，若實際放入之訊務少於 622Mb/S 稱為 over-provisioned。在 over-provisioned 狀況下 Qos 比較有機會達成，但相對的成本則較高。

若設計於 OC-12 上之訊務為高於 622Mb/S 稱為 under-provisioned，如航空定位若其估算會有若干百分比之定位不會搭乘，則會以 under-provisioned 方式設計，以保證能夠滿載，但若所有定位旅客皆要搭乘時則問題就產生了。而於電信網路上之訊務設計亦然，若以 under-provisioned 方式設計，則可能會有訊務無法傳遞之問題，此法之優點為節省成本但對於 Qos 及信譽則有影響。



圖三十四、Over-Provisioning v. Under-Provisioning

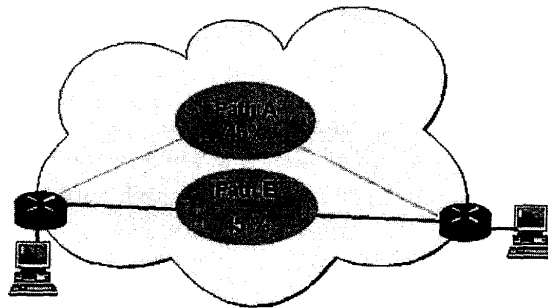
上圖顯示 over- or under-provision 之百分比關聯性

5.2 需求頻寬之計算

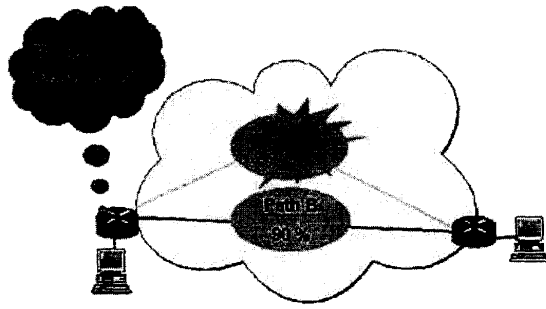
為簡化設計假設網路特性為已知，唯一需注意的為需集合所有之訊務。

CASE 1：某兩點間忙時以兩條負載平衡之 tunnels 承載訊務，其實體線路採用 OC-12 之設計。

若設計需有快速回復機能，則可設計每一條 tunnel 之訊務量占其使用之實體線路之 45%，此為 under-provisioned 之設計，如此一旦有一條 OC-12 障礙時，快速回復機能可將訊務導入另一路 tunnels，而此時該路之承載量為 90%仍為 under-provisioned 之設計，如下兩圖所示



圖三十五、 Sample Network Diagram Example One



圖三十六、 Sample Network Failure

而每一 tunnel 所承載之訊務內容如下

100 voice calls at 200Kb/s each
3 video calls at 500Kb/s each
250 Mb/s committed information rate

表六、 承載訊務內容

我們可用下表計算出上述訊務之總需求量，若結果少於設計值，則在安全範圍內，若結果高於設計值，則需要想辦法減少需求或移出部分之業務。

Traffic Demands		Totals and subtotals	
Number of voice calls	100		
b/s/call	200.000		
Total voice streams in b/s	20,000,000	20,000,000	
Number of video calls	3		
b/s/call	500.000		
Total video streams in b/s	1,500,000	1,500,000	
Committed information	250.000.000	250.000.000	

rate			
Other traffic	0	0	
Total traffic demand	271,500,000	271,500,000	BW required
Bandwidth Available			
Circuit bandwidth for OC-12	622,000,000		
Percentage used	45%		Over-provisioned
Total BW for over-provisioned	279,900,000	279,900,000	BW on-hand
		271,500,000	BW required
Remaining Bandwidth		8,400,000	BW remaining

表七、CASE 1 之話務計算

在上述之計算中若再考慮加入其它之機制以減少某種業務之使用頻寬時，則能容納多之業務或更多之呼叫

CASE 2：承前一 CASE，若在語音通道中加入靜音抑制機能，在靜音時不傳遞資料，使得每一通語音通道佔用之頻寬降至 100Kb/s，則網路上可使用之資源將相對較多。

Traffic Demands		Totals and subtotals	
Number of voice calls	100		
b/s/call	100,000		

Total voice streams in b/s	10,000,000	10,000,000	
Number of video calls	3		
b/s/call	500,000		
Total video streams in b/s	1,500,000	1,500,000	
Committed information rate	250,000,000	250,000,000	
Other traffic	0	0	
Total traffic demand	261,500,000	261,500,000	BW required
Bandwidth Available			
Circuit bandwidth for	622,000,000		
Percentage used	45%		Over-provisioned
Total BW for over-provisioned	279,900,000	279,900,000	BW on-hand
		261,500,000	BW required
Remaining Bandwidth		18,400,000	BW remaining

表七、CASE 2 之話務計算

MPLS 可透過 CR-LDP 或 RSVP-TE 規約來調整訊務，而目前來看應為 RSVP-TE 較佔優勢，雖然其頻寬成本較高，但至少較有保障，至

於是否可用 over-provisioned 之方式來設計並期待不會有高峰訊務出現，則是各自網路訊務設計者需根據其個別業務對訊務遺失之容忍度自行斟酌之課題。

第六章、心得與建議

在電信界目前最流行的路線行動通信及網際網路，以及現在正當紅之VoIP等等技術，再再的宣示IP網路的未來的主導性。對於此潮流我們會被其淹沒或站在其頂峰，端賴對新技術的掌握，以及對市場的敏感嗅覺。此行主要學習之MPLS為電信核心網路所使用之技術，若能使用於目前之IP網路，結合ATM之設備，建立電信網路層級之VPN，則可利用MPLS建構之VPN提供各類增值業務之服務，如日本電信使用MPLS backbone連結各ISP等。而本公司現有之設備某些已具有MPLS之功能，善加利用將能更使網路靈活的選擇路由、加強流量控制功能並達成用戶要求之QOS，使本公司於電信激烈競爭的年代中，永遠居於領導地位。