

在 NBEN 以政策為基礎的 QoS 管理 — Differentiated Services Provisioning

Policy-based QoS Management in NBEN – Differentiated Services Provisioning

孫雅麗 李正帆

台灣大學資訊管理學系
sunny@im.ntu.edu.tw

1. 簡介

隨著科技進步，通訊與電腦軟硬體的技术、速度都呈現快速的進步；這使得網路上多媒體應用與日遽增，視訊會議(video conference)及網路電話(Voice over IP)就是最明顯的例子。在網際網路變得越來越普及的同時，使用者已經越來越難忍受現有網際網路通訊協定及相關技術在資料傳輸效能的不可預測性與無品質保證的限制。我們知道許多未來的應用(applications)必須仰賴網路提供傳輸服務品質保證(QoS guarantees)的能力以及多種服務類別，來滿足各個不同使用者、應用的傳輸需求。這些服務品質的保證通常以下列的型態出現：有上界的端點間傳輸延遲(bounded end-to-end delay)、頻寬(bandwidth)、傳輸延遲的基差(delay jitter)、封包遺失率(packet loss)等，或是上述幾項的組合。

Internet IETF 針對未來網際網路就定義了兩個新的服務架構：整合服務(Integrated Services) [1][2][3]以及差別性服務(Differentiated Services) [4][5][6][7]。這個新架構的目標，就是要能提供多個不同層次等級的傳輸服務以滿足各種應用的傳輸要求，例如保證型服務(Guaranteed

Service)、負載控制服務(Controlled Load Service) 或是 Expedited Forwarding service 與 Assured Forwarding service (in addition to 現有的 best-effort service)。

現有的網路尤其是對一個企業體而言，連外網路的頻寬通常相對於 enterprise intranet 的頻寬相對小很多、因此當資料匯集於對外出口時就很容易發生擁塞，例如 TANet 的出國通道。因此，如何支援網路傳輸線頻寬共享(Link Sharing)的理想就日益重要。也就是需要網路能提供機制允許資源的共享可藉由行政考量，通訊協定及其資料型態或其他重要因素分類[8][9][10][11][12][13]。它是一個以應用(application-)、企業(enterprise) 以及業務(business-oriented) 為導向的管理。過去的文獻資料顯示 Internet 的資料傳送效能(例如回應時間、throughput, etc.) 一直是不可預測，完全要看當時的網路使用留量。如何提供使用者良好的上網傳輸品質與效能對網路的管理者是一大挑戰；這也是目前網路服務者最頭痛的問題之一。相關議題包含網路傳

輸品質的保證，特別是即時資訊的傳遞例如股市即時報價、下單或是緊急電子訊息的傳遞；網路管理；網路容量規劃。最重要的是一個網路服務的提供者如何能保證他的客戶的 *business-critical applications and traffic* 能夠正確的、即時的傳遞到目的地。

我們相信階層式的頻寬共享 (*hierarchical link sharing*) 的架構是絕對必要的，因為組織的階層架構已不是短暫的現象，組織必須將網路資源供多個單位使用，而每個單位希望當網路發生擁塞時能得到保證的頻寬；相對的，當頻寬沒有被一單位使用而剩餘時，其剩餘頻寬應如何分配或是希望以何種方式供其他單位使用，這些都是我們所謂的網路資源使用政策制訂的議題。傳統的存取控制列 (*access list*) 已不能滿足現今頻寬管理的需求，提供階層化的頻寬管理架構將使組織能更自然，更有彈性的管理網路頻寬。

圖一即是一個企業對外連線的可能頻寬管裡的政策架構的例子：第一層代表總的 link；第二層有 3 個節點，各代表一個單位 (*agency*)，例如業務部、研發部以及公司所有其它的部門。當這些所有的單位的資料要出公司，共享對外的網路頻寬時，網路管理者設定不管網路流量如何，代理節點 A1 業務部至少得到此對外網路 40% 的頻寬；代理節點 A2 研發部至少得到此對外網路 50% 的頻寬；而公司所有其它的部門（屬於盡力傳輸類）為避免完全得不到頻寬 (*Starvation*)，管理者設定在所有的頻寬中，盡力傳送類別可以得到至少 10% 的頻寬。

在這個計劃中，我們提出一套適合各種組織，非常有彈性的網路頻寬共享階層

架構，使組織或企業能對其昂貴的網路頻寬做更有效及更彈性的管理；另外，為了同時滿足即時應用 (*Real-time application*) 服務品質，*Best-Effort Traffic* 及階層式頻寬共享，我們實作一 *PC-based QoS router prototype*，能做到資料分流 (*classification*)，*Policing*，及提供服務品質保證的排程器 (*scheduler*)，圖二為我們的整合測試環境，利於我們進行各種測試及實驗。

2. NBEN - Nortel Open Policy Service

根據我們目前對 NBEN 上 Nortel Networks 的 IP switch router 以及 Open Policy Service (OPS) – a software package running on a remote workstation [14] – 的研讀與實驗，它們並沒有提供 IETF RFC2475[4]、RFC2474[5] 所制定的 *Differentiated Services* 的功能。雖然它可以根據使用者設定之 *Policy*，對不同的網路流量去 Mark 其 *codepoint values*，即其 IPV4 之 TOS 欄位，但在目前 NBEN 的 IP router BH5000 上我們並不能指定 *Differentiated Service* 不同 *codepoint values* (代表不同 *classes of service*) 可受到如何的對待 (例如 *bandwidth assignment*)。故我們打算在這個 router 的前端，裝置我們所設計、開發的 *PC-based QoS router prototype*。實驗 *policy-based differentiated service provisioning* 以研究探討所可能遭遇到的議題；也就是說我們會做一些實驗，實驗中所有要進入 NBEN backbone 的 *traffic*，將會在這個 *QoS router* 作資料分類 (*classification*) 與分流 (*segregation*)。不同類別的將給予不同的頻寬，提供不同等級的傳輸服務品質。另外，為了提供 *Policy-based QoS management*，我們也制

定一套網路資源分享架構，使網路管理者可更有效且彈性的管理其珍貴的網路資源。

3. PC-based QoS Router Prototype

如圖三，我們實作了一 PC-based QoS router prototypes，其作業平台為 FreeBSD 3.4 stable，ALTQ 1.2[15]。其模組包含 Porter、Call Admission Control、Classifier、Policers、Flow Manager and Packet Scheduler（目前支援 WFQ[16]及 WF2Q[17]）。在 QoS router 的架構中，Porter 負責接收來自 Policy Manager 之服務需求，將其 Parse 後，轉送給 Call Admission Control(CAC)，當 CAC 做出決策後，再把結果及相關參數分別送給 Policy Manager 及 Flow Manager；Call Admission Control 用來決定 Router 內現行網路資源是否能接受一項新的資源管理政策，或一個將進入路由器的連線，以確保現行網路流量及欲進來流量之服務品質；Policer 保證每個 flow 都符合其所描述的 Traffic Specification，在 packet 被放入 queue 之前先作檢查的動作；Flow Manager 接收 Porter 經過計算之後傳來的控制命令，將控制命令中的參數取出，根據所收到不同命令，回應所提供的不同功能，例如新增或刪除 queue：對 QoS 相關機制資料作初始化；另一方面對應不同功能，將不同的資料型態寫入 Kernel，如此一來，Scheduler 就可以根據建立出來的資訊表，完成保留頻寬的任務。目前我們實作了 WFQ 及 WF2Q 兩種 Scheduling algorithms，使用這兩種 Scheduler，我們對每個不同的服務類別建立不同的 queue，計算每個 queue 中每個封包的 virtual finish time，以此決定封包傳送的前後順序，以達到公平的保留頻寬和頻寬分配的效果。在此實作中，

我們達到了 per class queueing 和動態改變頻寬分配的要求。

4. Link Sharing Hierarchy

爲了同時滿足不同應用(Real-time application)服務品質需求與 Best-Effort Traffic 傳輸，以及彈性的(flexible)階層式頻寬共享，我們提出一套階層性網路頻寬共享架構：Integrated Services 的精神是針對每個 session 或 connection 保留特定頻寬，而 Differentiated Services 則對於不同的服務類別給予不同服務保證，但在我們的網路資源共享階層架構下，可以同時支援 Integrated Services 及 Differentiated Services，即網路管理者可針對單一連線(connection)做頻寬的保留，或對特定服務類別提出服務品質需求。圖四是此網路資源共享階層示意圖，每個封包會被分流到一個最底層的節點(leaf node)，接受其應有的服務品質。在此架構下，網路管理者可針對其內部所有路由器之介面做資源管理之政策設定，管理者可利用 Bpipe 服務類型(service class)替單一連線保留頻寬；也可透過 Class 服務類型爲一服務類別制定最小保證頻寬及指定其服務品質需求；Agency 服務類型則是一個比較邏輯性的單位，它可以表示一個部門，單位或一個別企業的網路流量；另外，爲了避免盡力傳輸類別完全得不到頻寬，可讓管理者設定盡力傳輸類別的最小保證頻寬。在下面的章節中，會對此架構做更詳盡的介紹。在管理者設定之後，於 Web 介面將會詳細的表現出整個網路資源使用情形，及其階層架構，可讓管理者清楚的明白目前已做的設定，如下圖五。

4.1 Organization

整個網路資源共享階層架構的頂

端，依使用者而定，它可以是一家企業，一個組織如學校或政府部門，或一個 Service Provider，如 ISP，ASP 或 Data Center 等。代表以下的網路資源是屬於這個組織或企業所管理。

4.2 Router and Interface

Router 即上述單位可以管理，或網路管理者欲對其做資源管理之路由器；Interface 則代表路由器上的介面卡或模組。管理者可針對從介面卡出去(Upstream)的網路流量，或從介面卡進來(Downstream)的網路流量做個別的網路資源管理政策。在 Interface 之下，管理者可根據資源管理政策需求，定義 Bpipe，Class 或是 Agency 服務類型。以下就各種類型作介紹。

4.3 Agency

Agency 是一個因行政上或管理上方便而衍生出來的邏輯單位，可以是一個部門，一棟大樓，或一家公司(例如一 ISP 租一條專線給一家公司)等。網路管理者指定一段 IP Address 範圍來定義一個 Agency，即 Source 或 Destination IP Address。在這個 IP 範圍內的網路流量，都屬於這個 Agency 的網路流量，管理者必須設定此 Agency 的最小保證頻寬以供其網路流量使用；此外，為了避免盡力傳送類別封包一直得不到網路頻寬，在建立 Agency 時，系統會要求管理者設定給盡力傳送類別的最小保證頻寬，以達到整個系統的完整性。在 Agency 下，管理者也可依需求建立 Bpipe 及 Class。如圖四：在臺灣大學的某個介面模組上，管理者建立一個 IM 的 Agency，其 IP 範圍為 140.112.106.1 到 140.112.106.254，指定其最小保證頻寬為 50Mbps。

為了避免盡力傳送類別封包一直得不到網路頻寬，在建立 Agency 時，會要求管理者設定給盡力傳送類別的最小保證頻寬，以達到整個系統的完整性。事實上，在建立一個 Agency 同時，系統會馬上建立一個 Best-Effort 的 queue，此時這個 queue 的最小頻寬是管理者給此 Agency 的最小保證頻寬，當管理者在此 Agency 下建立其他 Bpipe 或 Class 時，此 Best-Effort queue 的最大頻寬會慢慢減少，但最終不能低於管理者設定盡力傳送類別的最小保證頻寬，若管理者欲建立一個 Bpipe 或一 Class，使得 Best-Effort queue 的頻寬低於盡力傳送類別最小保證頻寬的話，於路由器內的 Call Admission Control 將會拒絕此設定。如圖七：當管理者建立一 agency，其最小保證頻寬為 4Mbps，而其 Best-Effort 的保證頻寬為 1Mbps 時，系統會建立一個頻寬為 4Mbps 的 Default queue 供 Best-Effort 流量使用；當管理者在此 Agency 下建立一 Ftp 服務類別，其保證頻寬為 1Mbps 時，系統在建立 Ftp，1Mbps queue 的同時，也會將 Default queue 之頻寬降為 3Mbps；同樣的，管理者再建立一個 Web 服務類別時，系統會降低 Default queue 之保證頻寬；但這時，若管理者欲新增一個 Telnet 類別，其保證頻寬為 1Mbps 時，CAC 將會拒絕這項設定，因為此設定將使 Default queue 的頻寬低於之前設定的 Best-Effort queue 之最小頻寬。

4.4 Class

我們定義 Class 為一服務類別，例如網路管理者可將所有的 Ftp 網路流量歸納為一個 Ftp 服務類別，在此服務類別內的每一連線(connection)，都將享有一致的服務品質保證。管理者可彈性的定義一服務

類別，例如根據 IP 範圍，通訊協定，連接埠或特定的應用程式等。為了因應現行網路應用程式趨勢，我們將 Class 細分為以下三類：1) Non-CAC-Class：這種服務類別沒有特定的 Delay bound 及 Loss rate 需求，如 Telnet，Ftp 或 WWW 類型之網路流量。管理者只需指定此服務類別的最小保證頻寬。因此當使用者在使用這個服務類別的網路頻寬之前，並不用經過 Flow setup 過程及 Call Admission Control 的同意，屬於這個服務類別的網路流量會一起競爭此類別的最小保證頻寬。2) VoIP Class：為 Voice Over IP 特定之服務類別，管理者必須設定此類別的最大可容忍 Delay bound 與 Loss rate，及其整個類別之最小保證頻寬。為了保證其 Delay bound 及 Loss rate，使用者在使用此服務類別之網路頻寬前，必先透過 Flow setup 過程通過 Call Admission Control 認證，若此類別之頻寬已不足夠新的連線加入，CAC 將會拒絕此連線加入；3) Video Class：Video Class 和 VoIP Class 的概念相同，為了保證 Video 網路應用流量的即時性，及最大封包遺失率的容忍度，使用者必須先通過 CAC 之後，才能使用此類別的網路頻寬。

在實作上，當管理者建立一個 VoIP 或 Video Class 時，系統會先在路由器上建立一個頻寬很小的 queue，因為這時並沒有實際的網路流量。當網路使用者欲開始應用程式，通過 CAC 後，系統會根據此應用程式的壓縮演算法，或不同產品之服務品質需求，動態增加這個 queue 的保證頻寬；當然，當此使用者結束此應用程式時，系統會動態減少此 queue 之保證頻寬，如此一來，不但可確保此應用程式之服務品質，也可更有效率的使用網路資源。

4.5 Bpipe

Bpipe 即 bandwidth pipe，管理者可利用指定 Source、Destination IP pair，protocol number and /or source 和 destination ports 建立 Bpipe 為特定連線保留頻寬。例如管理者可能為 CEO，CIO 或其他較重要的管理者及客戶連線保留頻寬。

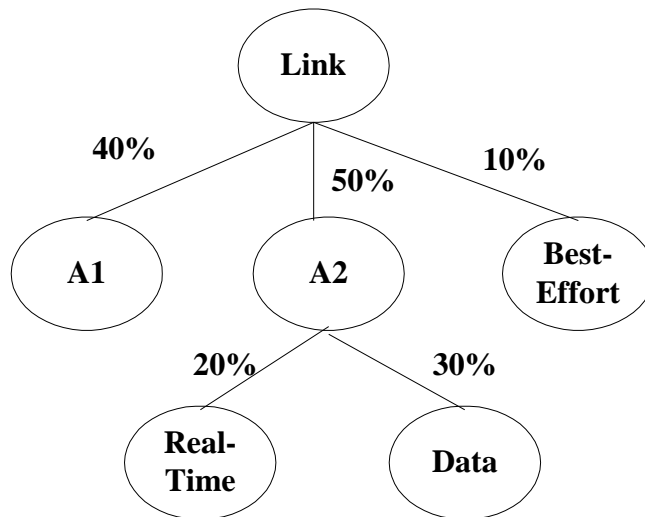
5. 結論

本論文實作一個 PC-based QoS router prototype，並提出一套適用於各企業或組織的網路資源管理共享架構，在此架構下，網路管理者不但可自行定義服務類別，也可針對特定的連線作保留頻寬。有了此架構後，管理者將可更彈性且更有效率的使用，及管理珍貴的網路資源。此 prototype 配合此架構已能因應目前整合性網路環境需求，經過初期測試，對於及時性的視訊流量及網路電話，都可正常的運作且達到公平的保留頻寬和頻寬分配的效果及動態改變頻寬分配的要求。

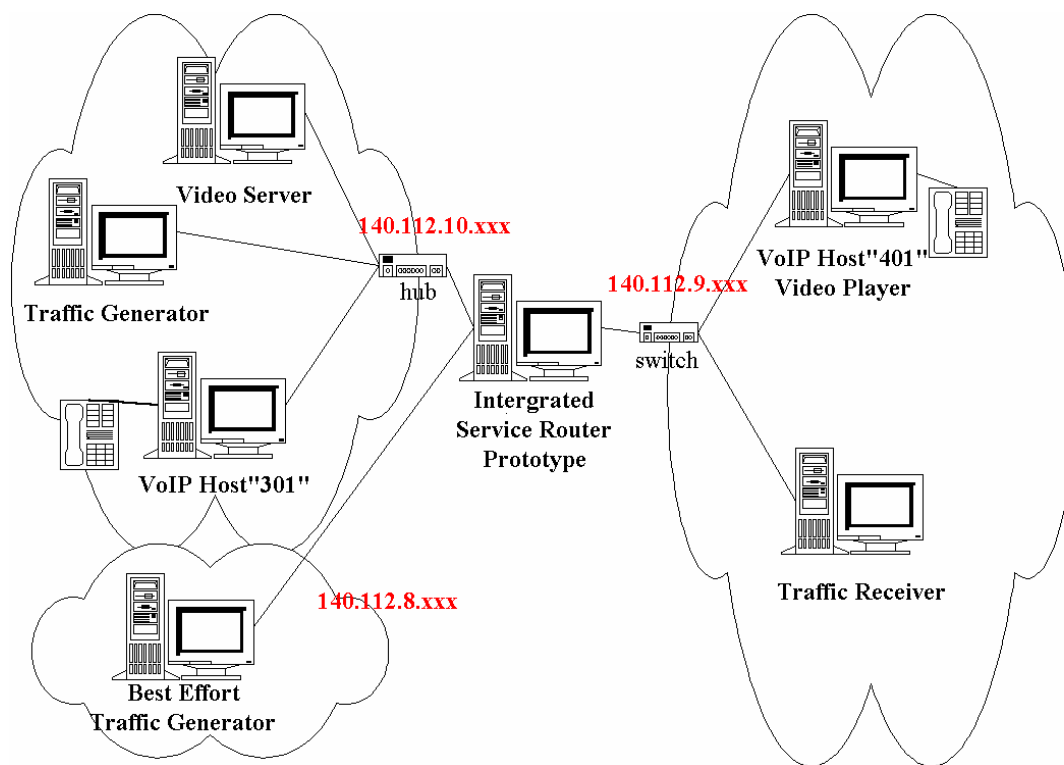
參考文獻

- [1] J. Wroclawski, "The Use of RSVP with IETF Integrated Services", IETF RFC 2210, September 1997.
- [2] J. Wroclawski, "Specification of the Controlled-Load Network Element Service", IETF RFC 2211, September 1997.
- [3] S. Shenker, C. Partridge and R. Guerin, "Specification of Guaranteed Quality of Service", IETF RFC 2212, September 1997.
- [4] K. Nichols et. al., "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers",

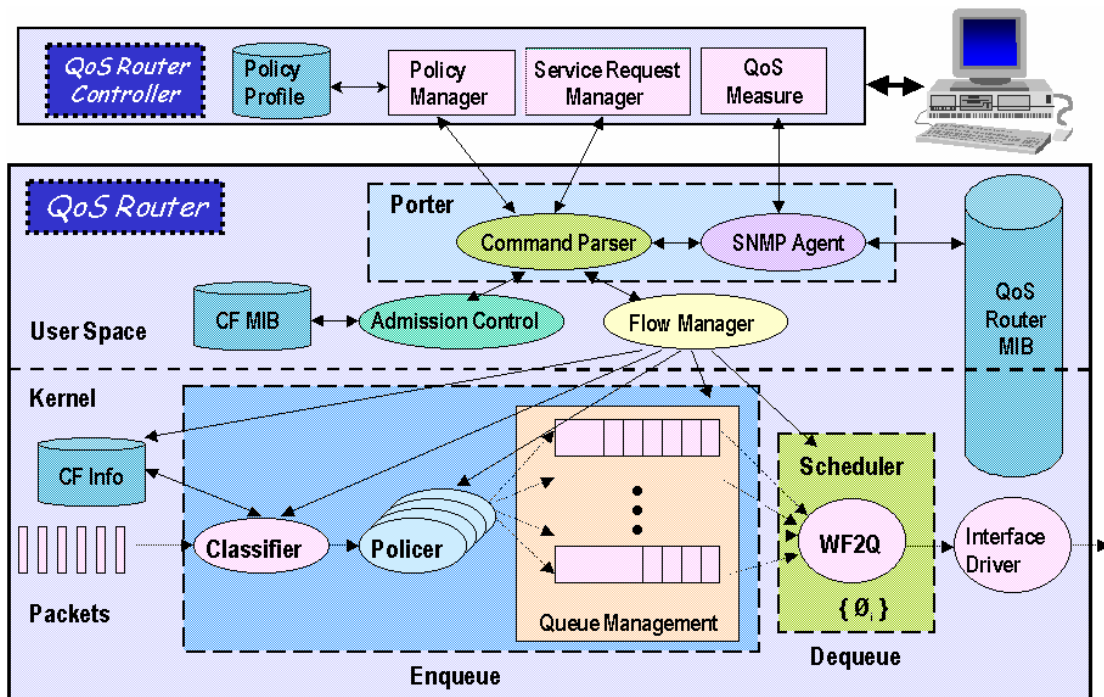
- IETF RFC2474, December 1998.
- [5] S. Blake, et. al., "An Architecture for Differentiated Services," IETF RFC2475, December 1998.
- [6] J. Heinanen, et. al., "Assured Forwarding PHB Group," IETF RFC2597, June 1999.
- [7] V. Jacobson, et. al., "An Expedited Forwarding PHB," IETF RFC2598, June 1999.
- [8] S. Floyd and V. Jacobson, "Link-sharing and resource management models for packet networks," *IEEE/ACM Trans. Networking*, vol. 3 pp. 365-386, Aug. 1995.
- [9] J. C. R. Bennett and H. Zhang, "Hierarchical packet fair queueing algorithms," *IEEE/ACM Trans. Networking*, vol. 5, pp. 675-689, Oct. 1997.
- [10] I. Stoica and H. Zhang, "A hierarchical fair service curve algorithm for link-sharing, real-time, and priority services," *IEEE/ACM Trans. Networking*, vol. 8 pp. 185 -199 , Apr. 2000.
- [11] R. Yavatkar, D. Pendarakis and R. Guerin, "A Framework for Policy-based Admission Control," IETF RFC 2753, January 2000.
- [12] M. Wahl, T. Howes and S. Kille, "Lightweight Directory Access Protocol (v3)," IETF RFC 2251, December 1997.
- [13] J. Boyle, R. Cohen, D. Durham, S. Herzog, R. Rajan and A. Sastry, "The COPS (Common Open Policy Service) Protocol," IETF RFC2748, January 2000.
- [14] http://www.nortelnetworks.com/corporate/news/newsreleases/1999b/6_16_9999420_Optivity_PS.html.
- [15] <http://www.csl.sony.co.jp/person/kjc/programs.html>.
- [16] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 3, pp.344-357, June 1993.
- [17] J. C. R. Bennett and H. Zhang, "WF2Q: worst-case fair weighted fair queueing," in Proc. *IEEE INFOCOM'96*, San Francisco, CA, Mar. 1996.



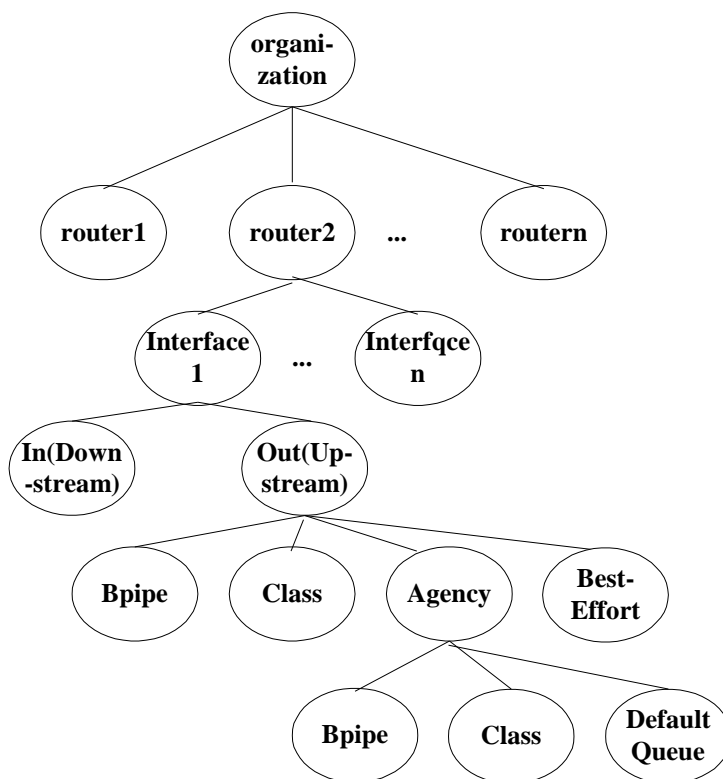
圖一、階層式網路資源分配範例



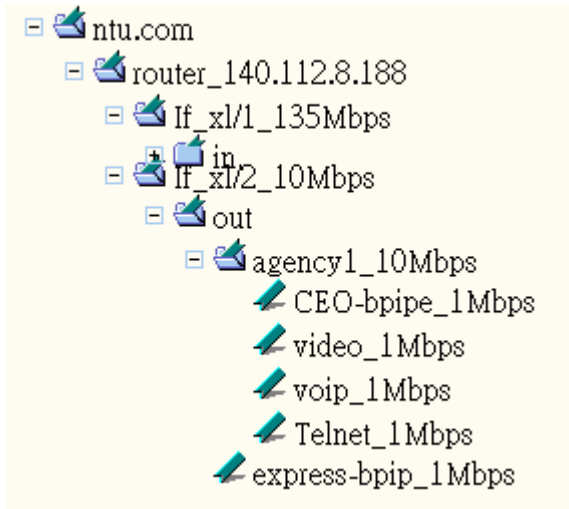
圖二、整合測試環境



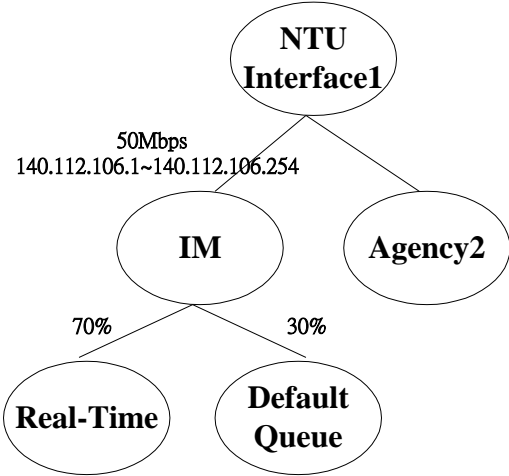
圖三、PC-based QoS router prototypes Implementation Architecture



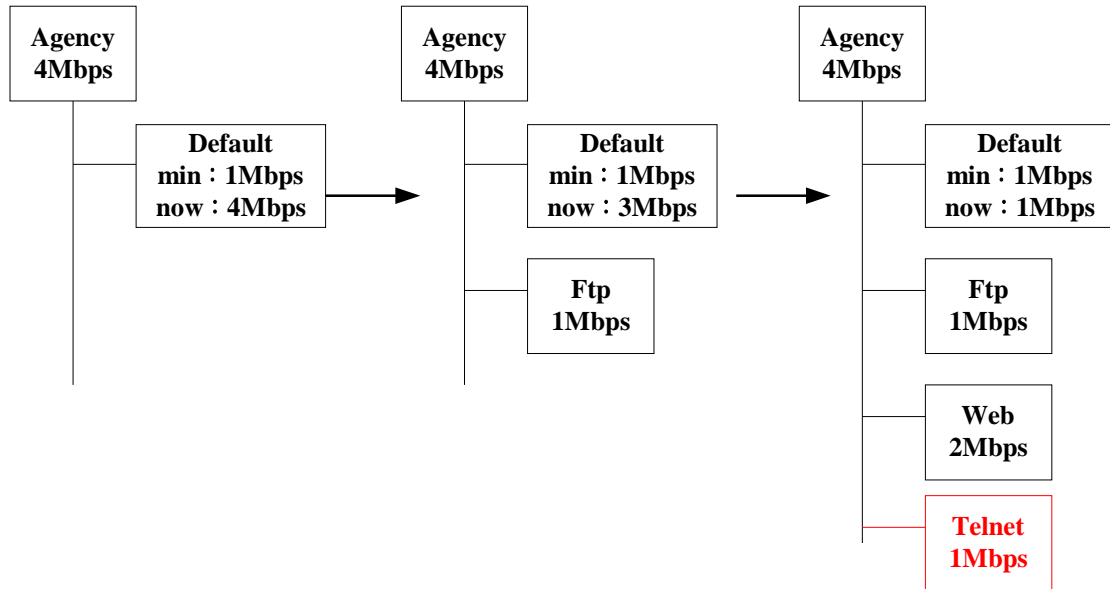
圖四、網路資源共享階層架構圖



圖五、實作之管理介面



圖六、Agency 範例



圖七、Agency queue allocation example