



RRPB88030716

(4.P)

PB 8803-0716

行政院國家科學委員會專題研究計畫成果報告

高速多層交換器之研製(I)

子計畫二：高速多層交換器系統平台之研製 (I)

計劃編號：NSC 88-2213-E-032-007

執行期限：87年8月1日至88年7月31日

主持人：許獻聰 執行機構及單位名稱：淡江大學電機工程學系

一、摘要

近年來網際網路盛行全世界，同時也帶來的各種網路問題和困擾，其中包括網路頻寬及效能不足等問題。本計畫我們將研究一種能有效解決上述問題的技術 Multi-layer Switch。它主要是整合了layer 3 (IP) 的通訊協定和layer 2 (Switching) 的交換技術，使得網際網路的頻寬與效能大大的提昇，成為可供下一世紀使用的網路技術和架構。本計劃「高速多層交換器系統平台之研製(I)」為群體研究計劃「高速多層交換器之研製(I)」之子計劃二。計劃進行的第一年裡，我們先研讀和分析目前現有的技術，作為規劃了高速多層交換器系統平台的完整架構之參考。除了選定符合系統需要的現有晶片模組和元件外，須配合自行開發的網路處理晶片以支援一個高速多層交換器之硬體平台，目前已完成其評估和規劃，將進入實作的階段。其目標便是利用多層交換的技術實作一高效率、低成本、並具有傳輸品質保證與虛擬區域網路等功能之交換器，以取代目前低速、高價、功能欠缺的路徑器，以符合目前與未來網際網路與多媒體應用的需求。

Recently, the Internet comes into fashion in the world. Simultaneously, it causes us many various problems and trouble on the network. One of the problems includes the problem of the insufficient bandwidth and efficiency.

This project will research into an efficient technique of multi-layer switch to solve above problem. It mainly integrates the communication protocol of Layer 3 (IP) and the switching technique of Layer 2. It makes the bandwidth and efficiency of the Internet a great deal increase and it also provides the network technique and architecture using in the next century. The project "The design and implementation of System Platform for High-Speed Multi-layer Switches (I)" is the subproject two of "The design and implementation of High-Speed Multi-layer Switches (I)" in the group-research project. As the project is proceeded in the first year, we have read and analyzed the available new technique to be a reference of planning the complete architecture of high-speed multi-layer switches. Besides the present chips selected to conform to the system requirements, we must coordinate the self-developing network processor to support a hardware platform of high-speed multi-layer switches. At present, we have finished the evaluation and plan, and will enter the implementation step. Our goal is to use the multi-layer switching technique to implement a high-efficiency, low-cost switches which have the QoS and VLAN functions to replace the present low-speed, high-cost and imperfection-function router. It will conform the present and future requirements in the Internet and multi-media applications.

二、計劃緣由與目的

近年來網際網路用戶急速成長，再加上傳送的資料型態已不限於純文字資料，而是轉變為需要大量頻寬的多媒體資料（即：聲音、影像、圖畫、超文字），造成目前網際網路（Internet）的使用量成指數成長，大約每三年成長十倍；而 Internet 上位址的傳送正是 IP 位址。但反觀負責處理和轉送這些資料的路徑器，其處理能力和效能卻以逐漸不敷使用，僅能以每兩年提升一倍。時至今日，傳統架構的路徑器已經無法有效率的處理日益增加的網路資料，而為了滿足需求殷切的即時（Real-time）和多媒體應用程式（Multimedia Application），IP 也必須能夠提供傳輸服務品質（QoS）的選項。於是新的路徑器架構便應運而起。目前學術界與工業界一致認為利用硬體的交換技術（Switching）比起軟體的路徑選擇（routing）要來得更有效率，再加上 Layer 2 switching 的技術已相當成熟，使得更高層（Layer 3 和 Layer 4）的交換技術相繼被提出，不論效率和價錢上皆比傳統路徑器要來得有競爭力。由此可見多層交換器勢必成為未來市場之主流。

傳統的 IP 路徑器架構，它有以下幾點的缺失：

1. 它無法支援多媒體服務
2. 頻寬有限
3. 無法作到流量的監管
4. 太多潛伏不利即時服務的因素
5. 這是以路由為基礎，所以無法評估 RSVP
6. Datagram 只能單一向前傳送
7. 是分享的服務品質保證(QoS)而不是真正的 QoS

以致於無法滿足現今多媒體的傳輸服務、即時或是服務品質的保證，所以才會有 Multi-layer Switch 的

架構提出。Multi-layer Switch 的架構是結合了 Layer 3 (IP)的通訊協定和 Layer 2 (Switching)的交換技術，不但具有傳輸品質保證與虛擬區域網路等功能之交換器；同時為了增進處理效率，我們儘可能利用硬體的架構來執行封包交換的功能，將硬體交換技術（Switching Technology）應用在 IP 路徑器的設計上，以儘可能達到 Wire speed 的要求。如此研製出的 Multi-layer Switch，不但可以達成傳輸品質保證與虛擬區域網路等功能，還兼具即時高速處理的能力，是適合未來需要的交換器。

三、研究方法與成果

目前網路晶片廠商已開發許多網路晶片組供系統廠商搭配使用，我們可以直接選用已經開發出來的晶片組，這樣不但可以縮短研發和設計所耗費的時間，而且也是符合經濟效益的作法。因此在研究的方法上，我們在規劃系統硬體平台時的基本原是先找市面上已有的現成而且符合所需的晶片和元件為主，分析各個晶片和元件的特性和功能，然後找出符合我們所需要的元件。如果找不到適合的晶片則必須由自己以 FPGA 開發出新的網路晶片。對於此次的架構裡，其中的三顆晶片(Lookup chip、Qos Chip 以及 Switching fabric chip)是非常重要的，可以說是這片 Switching board 的關鍵所在。它們分別負責網路 IP 的 Lookup、QoS 和封包交換的動作。然而這三顆晶片在目前市面上是尚未開發出來的，這樣就無法符合我們所需的規格，對於這樣的情況下，唯有自行設計開發這三顆晶片，才能順利達成預定的目標。

關於本計畫的完成結果有以下幾項：

1. 我們已審查了目前市面上的交換器、路徑器和 PCI 網路卡製作技術。而且檢視並評估現今業

- 界所採用的軟硬體平台架構。
2. 已完成研讀及分析現有的 IPv4/IPv6 Routing Table Lookup 的方法與技術。
 3. 已分析目前現有的 Fast Ethernet Switching Engine 系統，並加以評估其效能，以作為系統模組設計之參考。
 4. 完成以目前現有的 Port-based 第二層交換器和撰寫軟體驅動程式來模擬高效能交換器的運作行為。
 5. 熟悉硬體設計所需工具與相關資料，預計將來將繼續接觸到如 Ethernet module 等多層交換器所需設備。

四、結論與討論

有關本計畫所建立的多媒體路徑器架構還需要考慮底下幾個較重要的部分：

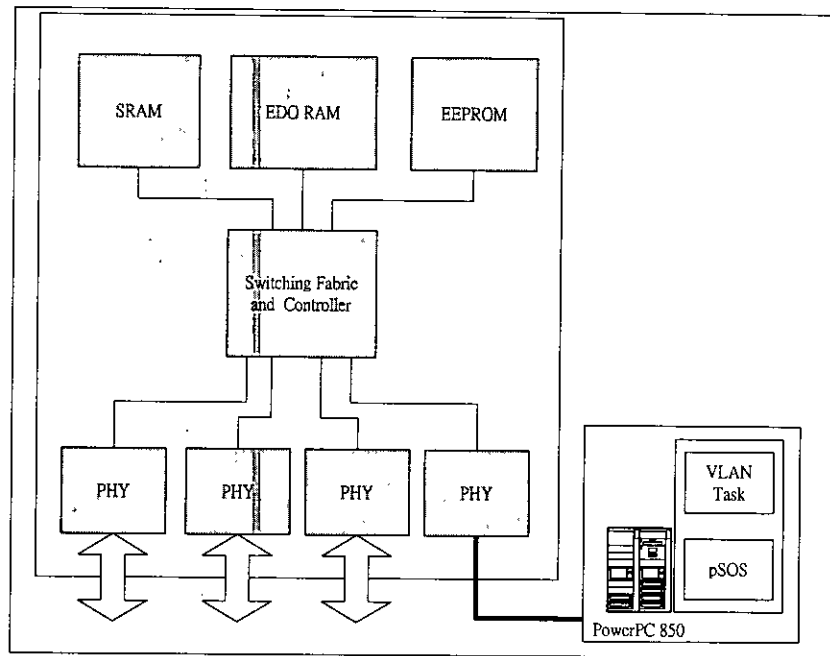
1. 選擇較適當的網路晶片組可以簡化系統的硬體平台架構之複雜度。
2. 除了建立硬體平台外，另一方面也必須修改 Switching board 內的韌體程式，讓 Switching board 能依據我們所規劃的架構模式順利驅動。

第一年的規劃和準備工作已經就緒，並將於第二年展開實作。卻由於此計畫後二年的研究經費被完全刪除，以致往後無法繼續執行實作的部分，甚為遺憾。面對業界網路技術不斷地創新，而學術界卻無經費發展高速多層交換器系統平台實作，實在令人感到無限可惜。

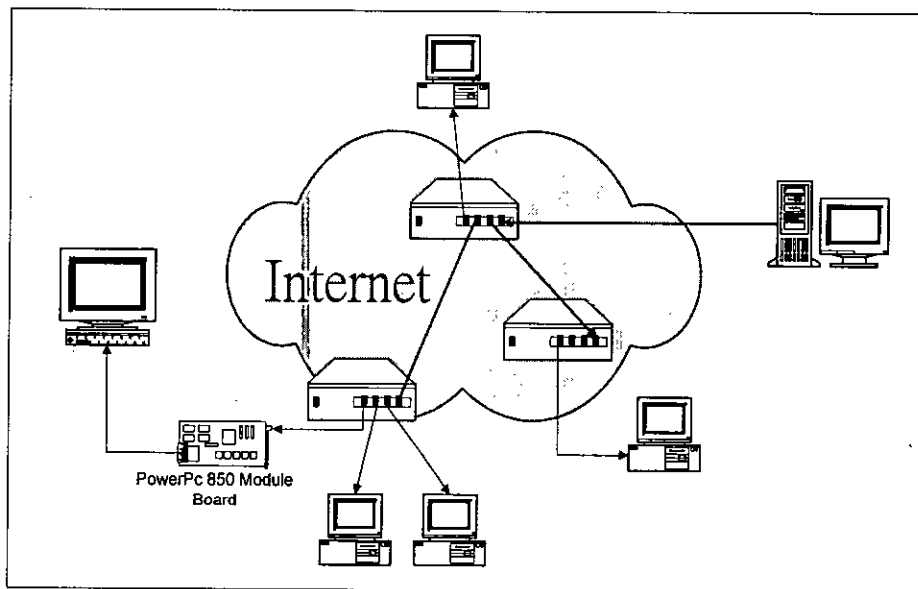
五、參考文獻

- [1]P. Newman et al., "IP Flow Management Protocol Specification for IPv4," IETF RFC 1953, May 1996.
- [2]P. Newman, T. Lyon, and G. Minshall, "IP Switching and Gigabit Routers," *IEEE Commun. Mag.*, January 1997.
- [3]D.D. Clark, S. Shenker, and L. Zhang, "Supporting Real-time Application in An Integrated Services Packet Network," Proc. ACM SIGCOMM, Comp. Commun. Review 22(4), Sep. 1992, 14-26.
- [4]R. Braden, D. Clark, and S. Shenker, "Integrated Services in The Internet Architecture: An Overview," IETF RFC 1633, Jul, 1994.
- [5]A. S. Thyagarajan, S. L. Casner, and S. E. Deering, "Making The Mbone Real," Proc. INET, Honolulu, Jun. 1995, 465-473.
- [6]Internet Domain Survey, July 1996, Network Wizards, <<http://www.nw.com>>.
- [7]RFC 1953 I P. Newman, W. Edwards, R. Hinden, E. Hoffman, F. Liaw, T. Lyon, G. Minshall, "Ipsilon Flow Management Protocol Specification for IPv4 Version 1.0," 05/23/1996
- [8]RFC 1987 I P. Newman, W. Edwards, R. Hinden, E. Hoffman, F. Liaw, T. Lyon, G. Minshall, " Ipsilon's General Switch Management Protocol Specification Version 1.1," 08/16/1996.
- [9]B. Braden, J. Postel, and Y. Rekhter, "Internet Extensions for Shared Media," IETF RFC 1620, May. 1994.

六、圖表



圖一. Multi-layer switch 內部架構圖



圖二. Multi-layer switch 外接架構圖