

# 行政院國家科學委員會專題研究計畫成果報告

## 以 TCP TRUNKING 達成階層式頻寬管理

### Using TCP Trunking to Hierarchically Manage Network Bandwidth

計畫編號：NSC 90-2213-R-009-157

執行期限：90 年 8 月 1 日至 91 年 7 月 31 日

主持人：王協源助理教授 國立交通大學資訊工程系

計畫參與人員：周智良、黃致銓、楊宗明 學生

#### 一、中文摘要

本計畫使用階層式 TCP Trunking 來達成階層式線路頻寬分配的目標。TCP Trunking 是個新穎及威力強大的頻寬分配的方法。本計畫將基本單層式的 TCP Trunk 延伸為階層式 TCP Trunk 且在由多台 FreeBSD UNIX 個人電腦所組成的測試網路上設計及實作階層式 TCP Trunk 的一個雛形系統。實驗結果顯示 TCP Trunking 能有效地達成階層式線路頻寬分配的目標。

關鍵詞：線路分享、頻寬分配、流量管理

#### Abstract

In this paper, we use TCP trunking to hierarchically manage a link's bandwidth. TCP trunking is a novel and powerful bandwidth allocation method. This project implements the single-level TCP trunking scheme and extends it to the multi-level TCP trunking scheme. We have implemented and tested the multi-level TCP trunking scheme on a network testbed composed of many FreeBSD machines. Experimental results show that TCP trunking can effectively achieve the goal of hierarchically managing a link's bandwidth.

Keywords: Link sharing, bandwidth allocation, traffic management

#### 二、緣由與目的

隨著網際網路 (Internet) 漸漸轉變為一

個有價值的資訊基礎建設，愈來愈多的使用者將不再滿足於現有的盡力而為 (best effort) 模式；相反的，當他們願意付較多的錢來取得更好的網路服務時，他們將會希望網路能提供一種品質保證或差別待遇 (Differentiated) 的服務以取得較多的頻寬或較低的傳輸延遲。

在現實中，一個機構通常是屬於階層式的架構，而此一機構之所有部門的頻寬分配亦是屬於這種階層式架構。因此，在將總頻寬分配給機構中的所有部門時，首先是將總頻寬依每部門內的人口數、願意付擔的費用百分比及工作型態等，分配給屬於等級 1 的部門，然後各等級 1 的部門再將其所分配到的頻寬依某種所需條件分配給其下較次級的部門。上述這種頻寬分配策略可應用至機構內各等級的部門，以達到階層式頻寬分配的目的。

階層式的網路頻寬分配可以以控制的方法來達到管理網路的流量。若在一个擁塞的網路環境下，沒有一個部門可以使用超過其所被允許的最大頻寬；然而，在一个不擁塞的網路環境下，所有部門可以在某種控制的機制下分享此一網路所提供的頻寬以達到頻寬的充份利用，在此一情況下，每一部門可以使用的頻寬是可以超過在擁塞下所被允許的頻寬。

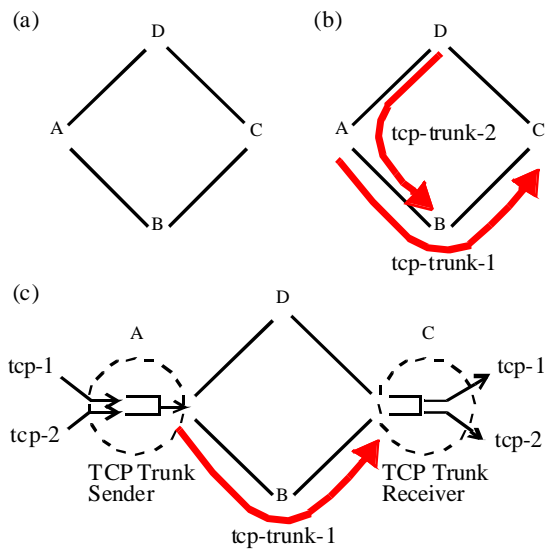
### 三、相關研究

在參考文獻 [1] 中，階層式網路共享之所以被提出是為了解決在一群機構相互競爭頻寬的擁塞網路中以階層式方式來分配頻寬的方法。參考文獻 [2] 所使用的這個方法叫作 class-based queueing (CBQ)，因為某些特定資料封包 (packet) 的集合就叫作一個 class，例如，某些從某特殊部間所發出的資料封包 (packet)。在各 class 中，CBQ 是用一種加權式的 round-robin 排班 (schedule) 方式來達到頻寬的分配。在一個 class 中，對於每一個到達的資料封包 (packet) 都給予一個服務的時間戳記 (service timestamp)，這個服務時間戳記的計算是基於該 class 所被分配的頻寬所計算出來的。所有 class 的資料封包 (packet) 都將全部一起被排班繼而依各資料封包 (packet) 的時間戳記 (timestamp) 將它們依序的送上網路。當一個 class 的緩衝區 (buffer) 已滿，則屬於該 class 的資料封包 (packet) 若再進來則會被丟棄。至於 CBQ 的運作和效能問題在參考文獻 [2, 3] 中有被深入的討論。

所有其它有關階層式網路共享的方法中所使用到的機制大致和上述所使用的機制是一樣的，它們間唯一不同的地方是在對於分配頻寬上公平性的定義有所不同。然而它們都是以時間戳記為基礎的方法而且都需要一個資料封包 (packet) 排班者 (scheduler) 來明確地排班資料封包 (packet)。

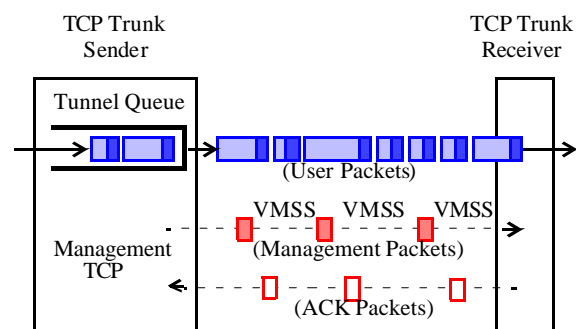
### 四、設計及實作

TCP Trunking 如下圖所示是一個 end-to-end 且使用 TCP congestion control 的頻寬管理機制。



(a) An IP network; (b) two TCP trunks over the network; and (c) two user flows (tcp-1 and tcp-2) over tcp-trunk-1.

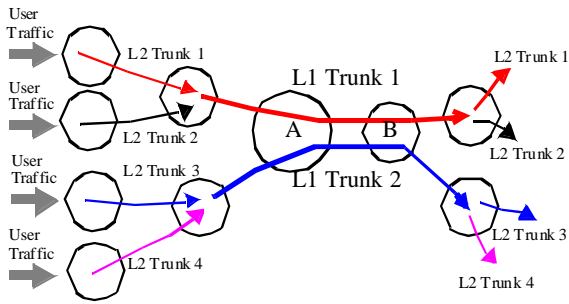
我們已經在 FreeBSD 平台上將它設計及實作出來。它在傳送端及接收端的設計原理如下所示。



Overview of TCP trunking implementation. At the trunk sender, arriving packets are redirected to the tunnel queue. The tunnel queue plays the same role as a TCP socket send buffer for the management TCP, except that once a user packet has been forwarded by the management TCP, its occupied buffer space is immediately reclaimed. (Since there will be no retransmission of lost user packets, there is no need to hold transmitted user packets in the buffer.) Each time after VMSS bytes of user packets have been forwarded out, the management TCP can send out a management packet that contains only a TCP/IP header but no TCP payload. At the trunk receiver, arriving user packets are immediately forwarded out based on their headers. All operations are done efficiently in the kernel; no user level operations or overhead are involved.

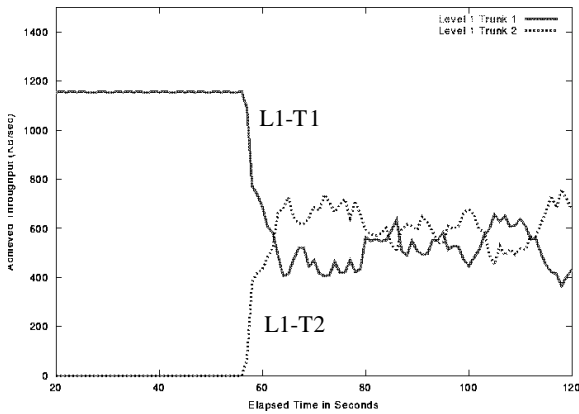
### 五、實驗結果

為了測試 TCP trunking 及階層式 TCP Trunking 的效能，我們使用以下的實驗平台。



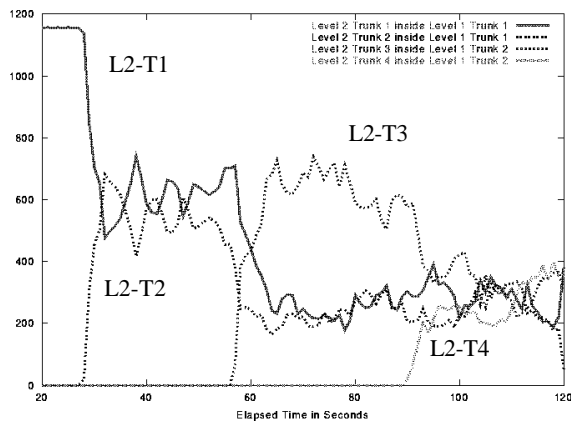
The experimental testbed network to demonstrate TCP trunk's hierarchical bandwidth allocation capability. In total, 16 nodes are used. Two of them are used as the shared traffic source and sinker node respectively, for the four level-2 trunks. (For presentation clarity, not all used nodes are shown in this figure.) All links are 10 Mbps except for those that connect to the shared traffic source and sinker nodes, which are 100 Mbps. There are two level-1 trunks contending for the bandwidth of the bottleneck link from node A to node B. In each level-1 trunk, there are two level-2 trunks contending for the achieved bandwidth of the level 1 trunk.

以下顯示第一層 TCP trunks 互相競爭頻寬的實驗結果。



Achieved bandwidth of the two level-1 trunks.

以下則顯示第二層 TCP trunks 在第一層 TCP trunks 內競爭頻寬的實驗結果。



Achieved bandwidth of the four level-2 TCP trunks.

## 六、討論

本研究的內容與原計畫非常相符。研究計畫的成果也達成原預期的目標。本研究成果現已經寫成論文正送到某期刊尋求發表中。

## 五、參考文獻

- [1] Floyd, S., and Jacobson, V., "Link-sharing and Resource Management Models for Packet Networks," *IEEE/ACM Transactions on Networking*, Vol. 3 No. 4, pp. 365-386, August 1995.
- [2] Fulvio Risso and Ranos Gevros, "Operational and Performance Issues of a CBQ Router," *ACM Communication Review*, Vol.29 No.5 October 1999.
- [3] J.C. R. Bennett and H. Zhang, "Hierarchical Packet Fair Queuing Algorithms," *ACM SIGCOMM'96* Stanford, CA, September, 1996, pp. 143-156

附件：封面格式

# 行政院國家科學委員會補助專題研究計畫成果報告

以 TCP TRUNKING 達成階層式頻寬管理

計畫類別：個別型計畫      整合型計畫

計畫編號： NSC 90-2213-R-009-157

執行期間： 90年8月1日至91年7月31日

計畫主持人：王協源助理教授      國立交通大學資訊工程系

共同主持人：

計畫參與人員：

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：國立交通大學資訊工程系

中 華 民 國      91 年      10 月 2      日