行政院國家科學委員會專題研究計畫　期中進度報告

子計畫一：多媒體內容傳遞網路封包分類排程技術之研究

(1/2)

計畫類別：整合型計畫
計畫編號：NSC92-2213-E-002-086-
執行期間：92 年 08 月 01 日至 93 年 07 月 31 日
執行單位：國立臺灣大學電機工程學系暨研究所

計畫主持人：蔡志宏

報告類型：精簡報告
報告附件：出席國際會議研究心得報告及發表論文
處理方式：本計畫可公開查詢

中　華　民　國 93 年 6 月 1 日

行政院國家科學委員會專題研究計畫成果報告
多媒體內容傳遞網路前瞻技術之研究-子計畫一
多媒體內容傳遞網路封包分類排程技術之研究(1/2)
計畫編號：NSC 93-2213-E-002-086
執行期限： 92 年 8 月 1 日至 93 年 7 月 31 日
主持人：蔡志宏　國立台灣大學電機工程學系暨研究所教授

前言

本研究至目前為止，在多媒體內容傳遞網路封包分類技術架構的有以下三個領域，已有較明顯之工作之進度，分別涵蓋從頭端、傳遞網路、以及到接收端的完整實驗。

一、頭端視訊伺服器群之封包分流架構

由於網際網路的普及，越來越多的人上網收看網路所提供多媒體影片所播放之球賽及娛樂節目。對串流影片伺服器提供者而言，單一伺服器已經無法滿足使用者的需求。如何有效增加串流伺服器端的處理能力，以滿足使用者端的需求，便成為一個嚴肅且亟待解決的問題。

在本研究中，我們研究如何增加串流伺服器的數目以服務龐大的客戶需求，本研究中的實作方式是在串流伺服器群的前端安裝 Linux Virtual Server，運用負載平衡的機制將收到的影片需求往後端的串流伺服器群送，以有效增加伺服器端的處理能力，來達到使串流影片伺服器提供者，能夠提供影片給更多的客戶，並確保其收看的品質。而其中之關鍵便是將封包流適當加以分類再進行路由。

1.1 封包分流架構

我們在伺服器群中，挑選其中一台來作為前端伺服器，作適當的封包分流連結以及軟體安裝設定修改，來達到平衡負載的效果。

我們應用 Linux Virtual Server(LVS)這套軟體來達成在 Linux 系統上架構出負載平衡的效果。

另外，我們在後端伺服器上安裝 Darwin Streaming Server-5.0(DSS)，來作為提供串流媒體撥放之用。本封包分流架構圖如圖 1。

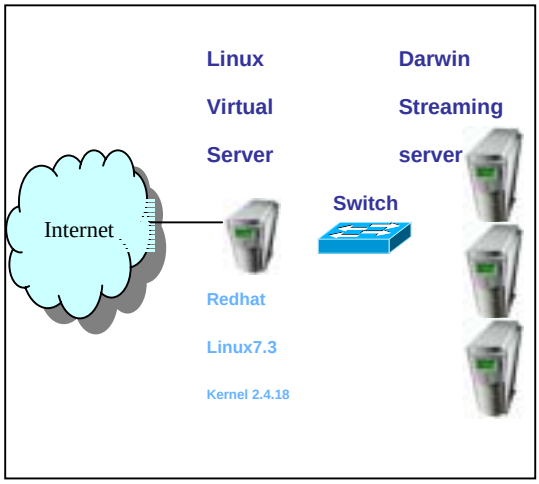

圖 1.視訊分流架構

使用者所發出的需求會先經由前端的 LVS 伺服器，LVS 再將使用者端的需求轉傳給後端 DSS 伺服器。LVS 會監測後端 DSS 伺服器的負載情形，以保障所有負載能平均的分布在後端伺服器群中。

在本次實驗中，我們利用圖 1 之架構，並改變後端 DSS 的數目來作測量，並借由三個參數(Streaming

Response Time，Packet loss，Jitter)，來判斷整個系統可提供最多的使用者人數。而這次實驗的結果如圖 2 所示。

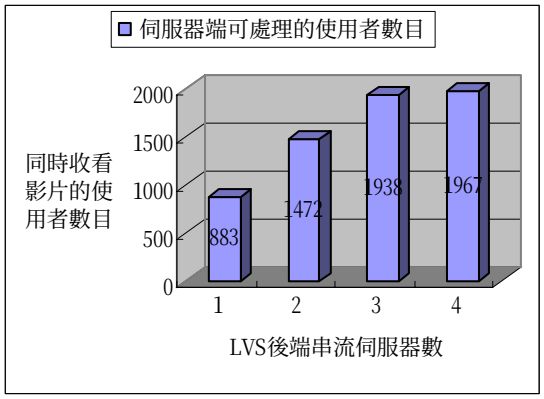

圖 2 視訊半流分流實驗結果

圖 2 的結果是在 ***Streaming Response Time*** < 1 second，***Packet loss*** < 1%，及 0.3 second < ***Jitter*** < 1 second 所作出的判斷。

由圖 2 結果我們可以看到，在 linux virtual server 封包分流之後所接的串流伺服器數目達到三台時，可同時觀看影片的使用者數目可達到約 1900 人，比起只有一台串流伺服器時來的佳，所以由此負載平衡運用在串流伺服器電腦群上的技巧，確實可以增加串流伺服器的能力。

二、傳送網路中防火牆/匣道器之規則簡化

2.1 防火牆種類及功能

防火牆大致上可分為兩類：封包過濾(packet filtering)、應用層閘道(application gateways)，但這兩類防火牆並不衝突。封包過濾是防火牆與路由器中最主要，也是最有效的功能，內部對外溝通是透過路由器(router) 傳遞封包，利用管理員所指定的規則來判斷是否允許封包傳遞通過防火牆。這些規則是針對每個封包的協定標頭 (head) 中所提供的資訊加以查核，查核的資訊包括：IP 來源與目的地位址、封裝的協定、來源與目的地連接埠、ICMP 訊息類型、內傳與外送介面，透過這些資訊，就可以指定哪些封包允許通過防火牆。

應用層閘道也稱為 Proxy 伺服器，它並不使用原本通用的傳遞資料方式而是特別設計過。看起來似乎是多此一舉，但比其它方法都有用。一點也不用擔心外面環境如何(使用系統是否有千瘡百孔)，因為它是獨立存在。正由於它的複雜，它所提供的安全性比封包過濾更高但只能控制特定應用程式的存取。 應用層閘道本上是用戶端和伺服器之間特定服務的中介者。封包過濾可以阻斷用戶端與伺服器之間關於該服務的直接通訊；流量全部改為送到應用層閘道伺服器。目前最常用的是 Web 上的 Proxy 伺服器。 應用層閘道另一個優點是縱使在十分危險的環境，它依然可以記錄所有進、出系統的資料。

本研究所提出的封包分類架構及改善方法皆是採用封包過濾型的防火牆/匣道器而設計。

2.2 分類規則建立

防火牆的運作方法主要是根據使用者所建立的規則去判斷是否允許封包進出的，因此防火牆的效能與使用者建立的規則有很大的關係，在防火牆的規則中包括了比對的內容和執行的動作。一般來講比對的內容有：來源 IP 位址(Source IP address)、目的 IP 位址(Destination IP address)、來源埠(Source port)、目

的埠(Destination port)，執行的內容包括了允許通過(Allow)或是拒絕通過(Deny)，防火牆會先看封包標頭(Packet header)的內容來跟已建立的規則一條條的依序做比較，如果符合其中的規則時，再執行規則中的動作，值得注意的是，當符合其中一條規則時，防火牆就不會繼續比對下去，所以在建立規則時，要特別注意先後順序，順序一旦弄錯，它就無法正確過濾封包。以下是以 iptables 為例，介紹一個防火牆的流程：

1.初始化：開始防火牆程式，並清空規則。

2.阻擋惡意 IP 位址：讀入 deny 的 script 檔。

3.允許 LAN 下的 IP 位址：讀入 allow 的 script 檔。

4.其他的 IP 則根據開放應用規則決定是否通過防火牆。

5.已建立之連線直接放行。

對於一些安全上的顧慮較高的網路環境，會建立相當複雜的規則，這樣的確可以過濾絕大部分來路不明的封包，但相對的會造成防火牆處理封包的負擔；另一方面，規則設的簡單可使防火牆做快速的判斷，但會使防火牆過濾的準度下降。

2.3 防火牆架構

單一防火牆主要用在一般小型的企業網路(圖 2-1)，且區域網路(LAN)內的成員是完全可信任的，防火牆只要抵擋網際網路上的不良封包，以及管理 LAN 成員向外連結的部份。

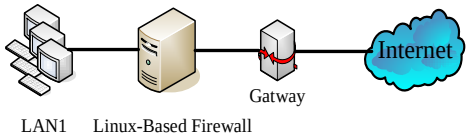

圖 3 單一防火牆架構

在一些大型網路中，公司內部有一些比較重要的電腦，並不希望所有的員工都能存取內部資料，所以會在防火牆內再設置防火牆(圖 4)，以達到層級更高的安全性，在這種架構下，內部的防火牆所要建立的規則就是以只允許某台電腦連至另一台電腦，在這類的架構中，內部防火牆的規則數目很明顯的會大很多，此時就需要有一個化簡規則的演算法來降低防火牆的負擔。

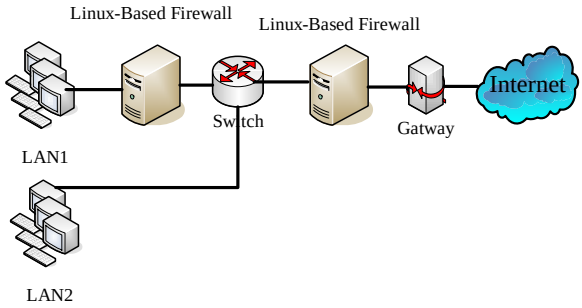

圖 4 分層式防火牆

在分層式的防火牆架構中，其主要得功能是能防止 LAN1 與 LAN2 之間不當的資料存取，但是對 LAN1 的使用者而言，如果需要對外做連結的時候會經過兩個防火牆的規則比對，這樣會增加內部防火牆的複雜度，可能會造成管理方面的問題，為了要能解決 LAN 與 LAN 之間的安全性以及對外聯結的問題，可以使用多重介面防火牆(圖 5)，防火牆只要根據管理者所制定的規則將封包轉送到指定的介面上，就可以與 LAN2 或是外部網路做連結。

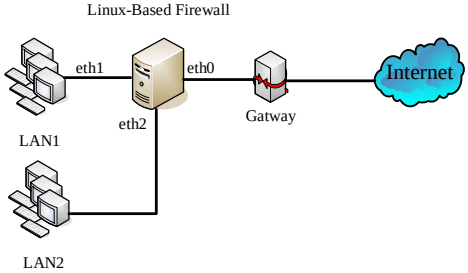

圖 5 多重介面防火牆

2.4 封包分類規則化簡的範例

前一章提到，規則數量與安全性的問題，以下舉一個例子來用數學式來描述這個現象。表 2-1 為原始防火牆的規則表，此防火牆是只允許這些子網路連上網際網路，每個子網路分別有 64 台主機。總共設了 16 條規則，架構如下圖 2-4。如果將防火牆的規則的網路遮罩(netmask)改為 24 個位元組的話，則每個規則就可處理 255 台主機，因此規則表會簡化成表 2-2，便可將規則從原先的 16 個化簡成只剩下 5 個，根據前述，規則化簡之後安全性會降低，接下來便解釋此現象的發生。根據表 1，在這個網路中的 140.112.21.193~140.112.21.254、140.112.23.193~140.112.23.254、140.112.24.193~140.112.24.254、140.112.25.193~140.112.25.254 均不存在，假設每個子網路的正常封包量為 ρ，病毒封包為 γ，一共有 16 個子網路，所以規則還沒化簡前病毒封包與正常封包的比值為 $16\gamma/16\rho = \gamma/\rho$，一旦規則化簡之後，中毒的電腦就可能會任意偽裝成其他的主機發送封包，這些原本不存在的 IP 位址就可通過防火牆到外部網路去了，所以病毒封包的量總共多了 4 個子網路的量，變成了 20γ，病毒封包與正常封包的比值就變為 $20\gamma/16\rho = 5\gamma/4\rho$，在這個例子裡發現當防火牆的規則從 16 個規則簡化成 5 個規則，內部網路的病毒封包增加了 $\gamma/4$，安全上的威脅自然就增加了。

rule	Network-layer Source (addr/mask)	Action
R1	140.112.21.0/26	Permit
R2	140.112.21.64/26	Permit
R3	140.112.21.128/26	Permit
R4	140.112.22.0/26	Permit
R5	140.112.22.64/26	Permit
R6	140.112.22.128/26	Permit
R7	140.112.22.192/26	Permit
R8	140.112.23.0/26	Permit
R9	140.112.23.64/26	Permit
R10	140.112.23.128/26	Permit
R11	140.112.24.0/26	Permit
R12	140.112.24.64/26	Permit
R13	140.112.24.128/26	Permit
R14	140.112.25.0/26	Permit
R15	140.112.25.64/26	Permit
R16	140.112.25.128/26	Permit

表 1　原始防火牆的規則表

rule	Network-layer Source (addr/mask)	Action
R1	140.112.21.0/24	Permit
R2	140.112.22.0/24	Permit
R3	140.112.23.0/24	Permit
R4	140.112.24.0/24	Permit
R5	140.112.25.0/24	Permit

表 2　化簡後的規則表

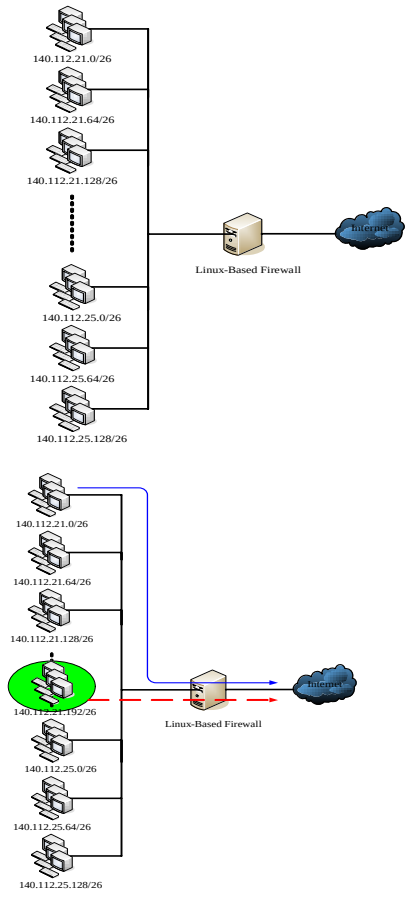

圖 6　防火牆/匣道器化簡後被穿透的情形

對一個內部防火牆而言也會有相同的情形發生，表 3 為規則化簡前後

的結果。根據之前的假設，可算出未化簡前的病毒封包與正常封包的比值為 $3\gamma/4\rho$，化簡後的比值為 γ/ρ。

rule	Network-layer Source (addr/mask)	Network-layer Destination (addr/mask)	Action
R1	140.112.21.0/28	140.112.22.0/28	Permit
R2	140.112.21.0/27	140.112.22.16/28	Permit
R3	140.112.21.32/27	140.112.22.0/27	Permit
R4	140.112.21.0/28	140.112.22.0/27	Permit
R5	140.112.21.0/27	140.112.22.32/27	Permit
R6	140.112.21.0/26	140.112.22.32/27	Permit

表 3 內部防火牆化簡前的規則表

rule	Network-layer Source (addr/mask)	Network-layer Destination (addr/mask)	Action
R1	140.112.21.0/27	140.112.22.0/27	Permit
R2	140.112.21.32/27	140.112.22.0/27	Permit
R3	140.112.21.0/28	140.112.22.0/27	Permit
R4	140.112.21.0/26	140.112.22.32/27	Permit

表 4 內部防火牆化簡後的規則表

2.5 分類規則叢集

防火牆是比對封包的來源 IP 位址、目的 IP 位址、來源埠、目的埠，所以可以將比對的內容映射到一個四維的座標系中，為了方便說明，以下都只考慮來源 IP 位址和目的 IP 位址。根據網路層的協定得知 IP 位址有 4 個位元組，如果要將完整的 IP 位址映射到平面上需要 $2^{32}\times 2^{32}$ 的平面，但從本章所提的防火牆架構得知，在分層式的防火牆與多重介面的防火牆中會有比較大量的規則數量，所以本研究所設計的演算法主要就適用在這兩種架構，而在這兩種架構下所架設的防火牆通常只在兩個特定的子網域之間，例如 140.112.20.0 和 140.112.21.0 這兩個子網域之間。在這種情形下，不需要知道整個完整的 IP 位址，只需要取得最後一個 byte 的值就可以將比對內容映射成(src_ip，des_ip)，此二維的座標點稱為 IP 對(IP pair)。如果防火牆收到從 140.112.20.61 到 140.112.21.110 的封包的話，經過映射的結果就變成(61，110)，這個二維座標就表示當防火牆收到從 140.112.20.61 轉送到 140.112.21.110 的封包的話，就執行動作(接受、丟棄或傳回)。

當所有的規則都映射在平面上時，就可以根據每個座標點在平面上分佈的情形來加以分群，在這裡將分群之後的結果稱為一個個的叢集(cluster)，有關分群的方法根據不同的準則會有不同的結果，包括叢集的數量、準則函數的定義，這些將在下一節做介紹。這些叢集是由座標點所構成的，在防火牆中設定了多少條的規則就表示平面上有多少個座標點，如果有辦法讓這些座標點用一些區塊來涵蓋住的話，就可以達到規則數量化簡的目的了。在一個網路環境裡，為了管理的方便常常會將網路分成好幾個子網域，所用的方法就是用網路遮罩(netmask)，讓路由器只根據 IP 封包中沒有被遮住的部分去做判斷路由，同樣的在防火牆比對規則的時候，也可以根據網路遮罩的方式將一些規則包含在一個子網域中，而這些子網域在經過映射在平面上的結果就是矩形，換句話說，就是要用一些矩形區塊來取代這些散佈在平面上的座標點。

2.6 規則化簡演算法

本研究所提出的演算法是根據 k-means 演算法改良而成的，k-means 演算法是一種常見的資料分群演算法，它是根據一個準則函數(Criterion function)來判斷資料所屬的叢集，定義如下

n_k：第 k 個叢集中所包含的資料數量。

N_c：叢集的個數。

$x_i^{(k)}$：第 k 個叢集中的第 i 個座標點。

$m^{(k)}$：第 k 個叢集中的質心座標。

Criterion function

$$e_k = \sum_{i=1}^{n_k} \left\| x_i^{(k)} - m^{(k)} \right\|^2$$

$$E = \sum_{k=1}^{Nc} e_k$$

e_k 的物理意義就是在這個叢集中所有的座標點與質心距離的平方，另外由於此演算法最後所要做的是將座標點用矩形的區塊涵蓋住，所以距離的公式並不適用於歐幾里德距離公式，另外定義如下：

$$Let$$

$$x = (x_1, x_2), x_1, x_2 \in R$$

$$y = (y_1, y_2), y_1, y_2 \in R$$

$$\|x - y\|^2 \equiv \max\{(y_1 - x_1), (y_2 - x_2)\}$$

根據這個定義，距離相等的點就只單純跟橫軸的距離或縱軸的距離相等，在平面上所表示的就是一個矩形。當有一個新的座標點要做資料分群時，就先將資料分別分配的第 1~N_c 個叢集，再根據以上的公式找出 E 的最小值，如此一來就可以判斷出此座標點是歸類於哪一個叢集的。

在傳統的 k-mean 演算法中，N_c 的大小為一定值，但是這並沒辦法適用於所有的資料，因此必須要想辦法根據資料的特性去增加後減少 N_c 的值。在這裡的作法是當我新的資料加到叢集之後，如果它與質心之間的距離比質心之間的最大距離還有大的話，就表示需要將它從其他的叢集中分出來，此時的 N_c 就需要比之前的叢集個數加 1；另一方面，如果某一個叢集的質心與屬於該質心資料最大的距離比質心之間最短的距離還要小的話，就必須要將這兩個叢集合併成一個，N_c 就比之前的叢集個數少 1。

2.7 重要成果

為了要找到可以表示成網路遮罩的矩形來包含這些座標點，必須要將 k-means 演算法做進一步的修改，原本的 k-means 演算法只能做到資料分群，但根據本論文所提出方法還必須要找到適當的區塊將一些座標點涵蓋住，首先紀錄每一個叢集中與質心相距最遠的點，根據它與質心的橫軸與縱軸的距離作一個矩形，為了要能做到網路遮罩，此矩形的邊長要有一些限制，也就是說，矩形的邊長要能用 2 的冪次方來表示，但是如果就以這個矩形來涵蓋整個叢集會造成很大的誤差，所以第二步就是在這個矩形的範圍內用一個特定大小的矩形來掃描，當這個矩形所包含的座標點大過某一個數量時，就將這些座標點用這個網路遮罩來表示，之後再將掃描的矩形縮小重複做上述之動作直到無法在化簡為止，當所有的叢集都做完時，演算法也就結束了。以下為此演算法的流程：

1. 將防火牆原始規則的來源 IP 位址與目的 IP 位址映射到平面上。
2. 在平面上隨機選取 N_c 個座標點作為叢集的質心。
3. 任選剩下的一個座標點加入第 1 個叢集，並重新計算質心與 E。
4. 重複 3.，將此點加入第 2~N_c 個叢集求出其他的 E。
5. 比較將此點加到哪一個叢集

後此時的 E 為最小，並將此點納入這個叢集。

6. 判斷叢集間可否合併或叢集數是否要增加。
7. 重複 3.~6.。
8. 決定每個叢集的掃描範圍。
9. 用特定大小的矩形掃描各個叢集。
10. 當矩形可包含超過某個數量的座標點時，將此網路遮罩取代這些點。
11. 將掃描的矩形縮小重複 10.

當防火牆的規則數目減少之後，會使得封包過濾的情形跟原來所設定的結果有一些差別，這些差別的結果稱之為誤判。誤判率定義為

$$誤判率 = \frac{誤放行交通量}{總交通量}$$

三、無線區域網路接收電視之實測

3.1 系統架構

為實際測試多媒體內容傳遞效果，我們利用 WLAN 來接收電視視訊封包，在無線網路電視伺服端方面，主要由一台作業系統為 Windows 2000 Advanced Server 的單 CPU 桌上型電腦充當閘道器，區隔無線電視網路與校園網路。其中尚需啟動其網路位址轉譯、DHCP 及 IIS（網際網路服務）管理功能，以做為 NAT、DHCP 及 Web 伺服器。

在視訊編碼方面，則分別由多台 Pentium III 雙 CPU 桌上型電腦，搭配 Windows 2000 Professional 及 Windows Media Encoder 來完成。轉換成數位格式後的封包經由一台搭配 Windows 2000 Server 及 Windows Media Administrator 之 Pentium III 雙 CPU 桌上型電腦來多點廣播。

多台 802.11a+b、一台 802.11a 及一台 802.11b 之 Access Point 目前遍佈於台大電機新館一樓及五樓的數間會議室、教室、視聽教室及實驗室。閘道器與各 Access Point 之間，有一台 throughput 為 45 Mbps 的頻寬管理器做頻寬控管，給予各頻道一定的最低保證頻寬。防火牆則阻擋了所有多點廣播的封包，以避免影響到台大電機系的網路。使用者端方面，經測試最低的軟體要求為 Windows 作業系統、Internet Explorer 5.0 或以上以及 Windows Media Player 7.1 或以上；硬體方面，必須配有 IEEE 802.11a 或 802.11b 之無線網路卡。只要符合上述的軟硬體要求，即可連上本系統的網頁觀看無線網路電視。

本系統之建立過程主要分為以下四個部份：

1. IEEE 802.11 無線區域網路之架設
2. 視訊編碼及群播環境之建立
3. 建立首頁
4. 頻寬調節及防火牆管理

系統建立的過程中遭遇到許多的問題，像是並非所有品牌的 Access Point 都支援多點廣播（multicast），即使支援，又往往無法有效設定其功能，我們必須多方研究軟硬體的相容性及勤於更新韌體版本，才逐步解決問題；頻率干擾也是一大問題，正因 802.11b 已相當普及，且其頻道較窄使得干擾更為明顯且嚴重，以台大電機新館為例，區區五層樓的建築物中就佈有超過 80 台的 802.11b 的 Access Point，外加各實驗室自行架設的 Access Point，干擾情況頗為嚴重，

初期我們致力在頻率規劃問題上，也設法與各單位協調頻率分配問題，但改善有限，當 802.11a 的產品推出後，我們選擇全面昇級無線區域網路，避開訊號干擾，以增加無線網路頻寬。

在克服了許多的困難之後，我們成功實作出無線電視網路試播系統，使用者欲觀看無線網路電視的使用者，必須具備有：

1.配有 IEEE 802.11a 或 b 無線網卡之桌上型電腦、筆記型電腦或 PDA 等終端設備

2.Internet Explorer 5.0 或以上

3.Windows Media 7.1 或以上（PDA 則利用 Windows Media 7.1 for Pocket PC 或 Windows Media Player for Pocket PC 2002）

圖 7 是 Tablet PC 及 PDA 播放無線網路電視的畫面。

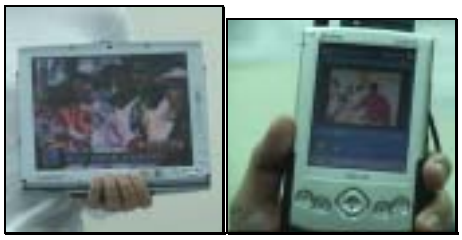

圖 7 電視串流接收實驗

結語

目前本計畫尚未完成深入之整合，而實驗之資料也在陸續整理中，但預計未來下一年度計畫中，將有更完整之研究成果及論文產出，並對產業界及學術界做出貢獻。

參考資料

1. W. Wang, C.Edward Chow, “Design and Implementation of a Linux-based Content switch, ” The Second International Conference on Parallel and Distributed Computing, Applications and Technologies(PDCAT), 2001
2. W. Zhang, S. Jin, Q Wu, “Scaling Internet service by LinuxDirector,” Proc. The Fourth International Conference/Exhibition on High Performance Computing in the Asia-Pacific Region, 2000., Vol. 1 , pp. 176-183, 14-17 May 2000,
3. V. Srinivasan, S. Suri, and G. Varghese. “Packet Classification Using Tuple Space Search.” In *ACM SIGCOMM,* September *1999*, pp.135-146.
4. Gupta, McKeown. “Algorithm for Packet Classification.” *Network, IEEE , Volume; 15, Issue: 2*, March-April 2001.
5. S. Hazelhurst, A. Attar. “Algorithm for Improving the Dependability of Firewall and Filter Rule Lists.” In *Proceedings International Conference on* , 25-28 June 2000, pp.576 – 585
6. Michael R. Lyu and Lorrien K. Y. Lau. “Firewall Security: Policies, Testing and Performance Evaluation.” In *The 24th Annual International* , 25-27 Oct. 2000, pp.116-121.
7. William R. Cheswick, Steven M. Bellovin, Aviel D, Rubin, *Firewalls And Internet Security: Repelling The Wily Hacker*, second edition, Addison-Wesley Professional.