在廣域網路上多工傳送壓縮視訊之技術

A Multiplexing Scheme for Transmission Compressed Video Streams over Wide Area Networks

童曉儒(Sheau-Ru Tong)、劉維真(We-Jeam Liu)
資訊管理研究所
國立屏東科技大學
Taiwan
srtong@mail.npust.edu.tw

摘要

經過壓縮後的視訊資料量變化極大,因而引發諸多關於如何有效利用網路資源來傳送的研究。就資料大小而言,事實上有規律可循,如果善加利用此規率性,可控制視訊的傳送時間點,以某種特定的時序,彼此交錯,我們稱之為「視訊多工(video multiplexing)」。如此一來,視訊彼此間互補資源的需求,進而提升資源(頻寬與暫存區)的使用率,同時能亦能無妥協地滿足視訊所要求的服務品質。因此在本文中,我們將此觀念運用在多節點的廣域網路傳輸,提出一套資源預留的演算法,來達成如何善用上層應用程式所允許的延遲上限,給予各網路節點適度的延遲彈性,以使他們能達到最佳的「視訊多工」效果。為了驗證所提方法的正確性與優越性,我們以 Simscript 模擬軟體模擬系統運作。從實驗結果顯示,當視訊流的數量大時,我們明顯的觀察到「視訊多工」所帶來的好處,一般頻寬使用率能夠達到60~70%之譜,而額外暫存區的需求亦十分有限。

關鍵字:廣域網路,VBR 視訊多工,Deterministic guarantee service

ABSTRACT

It is well known that compressed videos, exhibiting significant rate variability, are hazardous to the network resource utilization. In fact, traffic also exhibits a regular pattern. If we multiplex video streams with a specific temporal relationship, they will compensate with each other on their resource requirements. As a result, the resource utilization is improved without scarifying any video quality. We term such a technique *video multiplexing*. In this paper, we demonstrate how to apply video multiplexing over multi-hop wide-area networks. A resource reservation mechanism is proposed to determine the transmission schedule in each intermediate hop for achieving the best multiplexing effect subject to the end-to-end delay bound tolerable to the upper-layer applications.

We simulate the system in Simscript simulation package and observe that when the number of video streams are many, the performance gain brought by video multiplexing is evident. In general, a 60~70% utilization on bandwidth can be achieved with moderate buffer space deployed in intermediate hops.

1.Introduction

目前的壓縮技術可將視訊壓縮成固定速率(Constant-Bit-Rate；CBR)資料流，雖然網路資源管理得以簡化，但造成了每張畫面品質不一的效果。如果欲達到固定的畫面品質，經過壓縮後，將會使每張畫面的資料量不同，因而產生了可變速率(Variable-Bit-Rate；VBR)資料流，如何有效的運用網路資源來傳送 VBR 資料流，是一項熱門且重要的研究課題。現今新一代的寬頻網路，(如 ATM (Asynchronous Transfer Mode[4]))就網路傳輸服務方面，可根據使用者不同的需求，區分為 deterministic guarantee 及 statistic guarantee 二大類。其中 Statistic guarantee 傳輸服務強調只要整體的服務品質能達到某一可容忍的限度即可。而 Deterministic guarantee 則是對於傳送服務品質要求百分之百的正確性。至於該選擇何種傳送服務品質做為傳輸策略，則須視在不同的環境下，對視訊傳送服務品質的要求程度，一般相信 statistic guarantee 可藉由統計式多工(statistic multiplexing)的效果,而達到較 deterministic guarantee 為高的頻寬使用率,但相對的卻付出資料可能遺失的代價,而本文重點則突破此種認知,提出一套創新的視訊多工的方法來提升 Deterministic guarantee 的頻寬使用率。

在 Deterministic guarantee 傳輸服務品質下，我們可以在視訊流當中出現最大的資料流量(peak bandwidth)做為預留頻寬的大小,以達到無遺失的傳輸，但針對壓縮後的視訊流而言，其某一瞬間的資料變化量往往有相當大的起伏，但若以 peak bandwidth 做為預留的依據，會造成頻寬嚴重的浪費問題。因此有許多的研究在探討如何對視訊流模型化，企圖能以簡單的參數更精準的描述視訊流的變化,以便配合流量控制的機制做更有效率的資源預留[5][6][7][17][18]。另一類做法則是追縱實際資料量的變化,以求得在來源端與接收端所需的暫存區大小、啟始延遲與網路頻寬四者之間的關係[1] [2] [9] [10] [11] [12] [13] [14] [15] [16] 。然而這些研究大部份都只針對單獨視訊流討論，視訊流之間無法共享資源,同時討論的重點放在兩端點的設計,而對在廣域網路環境下路徑必須經過多點(multiple hops)的狀況鮮有著墨。

我們從 MPEG 壓縮過後的資料流中觀察，每一條視訊流其實是以一個固定的週期(group of picture；GOP)不斷的重覆本身的資料流型態(frame pattern)，根據此一特性，我們假設能夠控制視訊流進入網路傳輸的時間點,視訊流和視訊流間則可以某種預期的時序關係相互多工交織，藉以錯開突發(burst)的流量來避開資料量擁塞，以提昇頻寬使用率。我們稱此種方法為「視訊多工(video multiplexing)」。在[8]中提出了一套尋找多道視訊流之間最佳多工組合的法則。但未提及如何將此方法延伸至多節點的廣域網路環境。

因此我們在此研究中探討當視訊多工應用在多節的廣域網路環境下所遭遇到的難題與對策。首先由於節點間互為獨立，為了讓視訊流在特定的時間傳輸，我們在節點上必須加上緩衝區(buffer)，以對視訊流採行必要的滯留，以等待適當的時機再將此視訊流送出。明顯的,就一道資料流而言，在路徑上各節點的滯留時間總和不得超過上層應用程式所允許的延遲上限。如果節點上滯留的時間太短,我們可能必須遷就較不理想的多功組合,因而無法達到預期的多工效益;反之, 如果滯留的時間太長,我們可能超過延遲上限。因此如何能夠在諸多的因素中間找到一個平衡點,來決定各節點上應分配多少的滯留的時間,是一個值得探討的問題。再者,在真實的網路情境中,

資料流隨時可能自任一節點產生,也隨時可能自任一節點結束,換句話說,多工組合必須不斷的更新,此更新必須要在不影響目前正在進行傳送的視訊流為原則,因此如何在此動態的環境中,尋找到最佳或適當的多工組合,是一個必須解決的問題。我們將針對這些問題分別提出演算法。為了驗證所提方法的正確性與優越性,我們以 Simscript 模擬軟體模擬一個四個節點的網路系統運作,四種不同型態的 MPEG 視訊流，以 Poisson process 的方式到達。從實驗結果顯示,當視訊流的數量大時,我們明顯的觀察到「視訊多工」所帶來的好處,一般頻寬使用率能夠達到 60~70%之譜,而額外暫存區的需求亦十分穩定與有限。

以下章節的組織如下:第二節將描述 MPEG 視訊流的特性;第三節則在探討視訊多工的原理和當中所引發的問題與我們的演算法；第四節是描述整個的系統的細部運作；第五節是系統模擬的結果,而最後的第六節是我們的結論。

2.MPEG 視訊壓縮標準

在這節當中，我們來探討組成 MPEG 壓縮視訊流的三種 frame 分別所代表的意義和特性[3]。

1. I frame (Intracoded frame) : 為其他 frame 所須參考的依據，也就是本身不須參考其他 frame 就可自行解壓縮的，又稱為 reference frame。
2. P frame (Predicated frame):必須參考之前的 I frame 或 P frame (以最接近本身的為準)。
3. B frame (Bidirectional frame):必須參考之前的 I frame 或 P frame 及後面最先預測到的 I frame 或 P frame。其中各 frame 間相互參考的方向如圖 1。

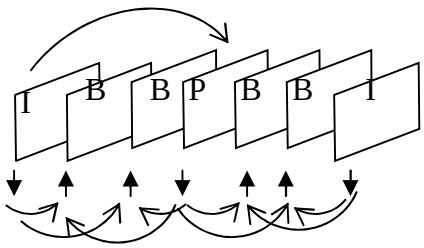

圖 1:不同畫面間的(解)壓縮參考關係圖

根據 I、B、P frame 間的間隔，我們可定義下面的參數:

L：表兩個相鄰 I frame 間 frame 數目+1。從一個 I frame 到下一個 I frame 之前的 frame 所組成的 frame pattern 會不斷的返覆出現，也是我們所稱的 Group of Picture；GOP，而 L 值也正代表著 GOP 的長度，。

Q：表兩個相鄰 I 或 P frame 間 frame 數目+1。

例如當視訊流以 I B B P B B I B B P B B I ….的 frame pattern 出現時，其 GOP = I B B P B B、L=6、Q=3。一般根據 L 的值，我們可以獲得影片中場景變化的程度，L 值愈大表示場景變化有限，故可以以較大比例壓縮，而產生的資料的變化量也愈大。雖然在 MPEG 的標準中未要求在一部壓縮的影片中 L 與 Q 值必須保持定值，也就是說 L 與 Q 理論上應能隨場景的複雜度做動態調整，以達最佳的視覺效果，但是一般而言，一部影片的 L 與 Q 值多採定值，以簡化在用戶端解壓縮時之複雜度。而就 frame 的大小而言，I frame 最大，其次為 P frame， 最小的則為 B frame。其間的差異可達數十到數百倍。

3. 多節點視訊多工(Multi-hop Video Multiplexing)

3.1 視訊流的起始延遲

所謂視訊多工(video multiplexing)，即表示經由同一窗口傳輸的視訊流間，保持一定的時序關係，假設兩道視訊流 S_1 與 S_2,

其 GOP 分別為 $I_1B_1B_1P_1B_1B_1$ 與 $I_2B_2B_2P_2B_2B_2$，如果沒有強制時序關係，則兩道視訊流 S_1 與 S_2 有可能同時傳送 I frame，故網路必須以兩道視訊流中各別最大 I frame 大小的和為預留頻寬,也就是 $|I_1^{max}|+|I_2^{max}|$ (其中$|K_i^{max}|$表示 S_i 中 K 類型 frame 的最大 size)。如果強制 S_2 的視訊流較 S_1 的視訊流晚一個 frame time (如 NTSC 系統中的 33ms)送出，則所須預約之頻寬可望降至 S_1 與 S_2 相對映 frame 大小的最大組合，也就是 $\max\{(|I_1^{max}|+|B_2^{max}|),(|B_1^{max}|+|I_2^{max}|),(|B_1^{max}|+|B_2^{max}|),(|P_1^{max}|+|B_2^{max}|),(|B_1^{max}|+|P_2^{max}|),(|B_1^{max}|+|B_2^{max}|)\}$,其值明顯地比前者小。而對 S_2 所強制的延遲，也就是從它抵達節點至開始傳送所經歷的時間,我們稱之為起始延遲(initial delay)。如果所在節點是視訊流的起源點（如視訊伺服器)，則我們可以控制 S_2 的啟始時間，來達到強制的起始延遲的效果，但是如果節點為路徑上的某一點，則我們將無法控制 S_2 的抵達時間，故必須藉助緩衝區將資料暫存於內，俟起始延遲到時，方允許 S_2 傳送，所有的 S_2 資料都必須經過緩衝區的延遲。

3.2 多工週期與頻寬預留

實際上我們在多工後的整合視訊流中可觀察到一個不斷返覆的 frame pattern 週期，其中每個 frame time 內出現來自每個視訊流的一個 frame,若以多工 S_1(其 GOP=$I_1B_1B_1$) 與 S_2(其 GOP=$I_2B_2P_2B_2$) 為例，我們假設 S_2 晚 S_1 一個 frame time，則經由多工後所得的 frame pattern 週期為 $\{I_1,B_2\}$ $\{B_1,I_2\}$ $\{B_1,B_2\}$ $\{I_1,P_2\}$ $\{B_1,B_2\}$ $\{B_1,I_2\}$ $\{I_1,B_2\}$ $\{B_1,P_2\}$ $\{B_1,B_2\}$ $\{I_1,I_2\}$ $\{B_1,B_2\}$ $\{B_1,P_2\}$ (其中{*}表示在同一 frame time 網路所傳送的 frame pattern)。我們稱此週期為多工週期(Multiplexing Cycle；MC)，而多工週期的長度即為所有參與多工者的 GOP 長度的最小公倍數(LCM 或 Least Common Multiplier)。就前例而言，LCM(3,4)=12。正式的說,假設既有 k 道視訊流正在多工,其 GOP 長分別為 m_1, m_2, ..m_k,故其多工週期長為 n=LCM(m_1, m_2, ..m_k),有一新視訊流欲加入,其 GOP 長為 m_{k+1},伺其加入後,多工週期長為調整為 LCM(n,m_{k+1});同樣的,如果視訊流 i 欲離去,則多工週期長必須調整為 LCM(m_1,m_2,….,m_{i-1}, m_{i+1}, …,m_k,m_{k+1})。

我們所需預留的頻寬為,在多工週期中任何一個 frame time 內產生的最大可能資料量,理論上我們欲尋求一個可預留最少的頻寬的多工組合,然而多工的組合究竟有多少種?就多工 k 道視訊流而言,我們可以推論如下,視訊流 1 與 2 多工時有 min(m_1, m_2)個組合,多工後的視訊流 (多工週期長為 LCM(m_1,m_2)), 再與視訊流 3 多工,其組合又有 min(LCM(m_1,m_2),m_3)種,故至此總共有 min(m_1, m_2)* min(LCM(m_1,m_2),m_3)種,以此類推,當 k 大時,可能的組合數將以指數的形式成長,每加入或離開一道視訊流就必須重新找一次最佳組合,且可能會改變正在進行中的視訊流的啟始延遲,故此種窮舉法在多節點網路中不可行。有鑒於此我們採行以下簡化的方法:

- 新加入的視訊流,尋找與目前多工整合視訊流的最佳組合,如此一來只有 min(n,m)種組合必須比較(其中 n、m 分別為新視訊流的 GOP 長與整合視訊流的多工週期長),
- 當視訊流離開時,僅將所屬的 frame pattern 自多工週期中移除,同時調整多工週期長度。

注意我們對既有視訊流不做任何延遲調整。如此做法雖然我們無法在節點上獲至最佳的多工組合,但是每有新視訊加入時,節點會盡其所能的朝向此目標修正,如此一

來視訊多工的優點得以在節點上實踐。

3.3 節點中的起始延遲「容許度」

但是這裡又演生了另一個議題,因為應用環境不同,對應的應用程式所能容忍的端點至端點的延遲範圍亦不同,如視訊會議(video conference)需要不超過 500ms 的延遲,否則使用者將無法忍受,又如隨選視訊(video-on-demand),使用者可允許數秒的延遲。因此如何在上層應用程式的所能容許延遲的限制下,在所經過的節點上進行多工是必須解決的問題。我們的做法是將端點至端點的延遲分配於所經的節點之上,在此我們稱每一個節點的此時間為「容許度」。當在節點上尋找最佳多工組合時,從資料抵達到傳送出去,所經歷的延遲時間 (即起始延遲) 必須小於此值。也就是說,我們所尋找的最佳多工組合局限於此範圍內,其它會導至大於此延遲值的組合將不予考率。(當此容許度值大於前述的多工組合數 min(n,m)時,則以 min(n,m)為準。) 接下來的問題是依據何種法則來分配容許度？一般而言,容許度愈大，就有愈大的彈性找出較滿意的多工組合,如果節點上目前多工週期的資料量變化很大,則表示多工效果不彰,需要加強多工組合的選取;反之,如果節點上目前多工週期的資料量十分穩定,則表示多工效果顯著,故可多遷就較差的多工組合。依據此觀察,我們採用各節點多工週期的頻寬的變異度大小,來做為分配容許度的依據。也就是變異度較大的節點將分配到較大部份的端點至端點延遲,來做為它的容許度,如此一來它就有較大的機率找到理想的多功組合。

3.4 延遲機制的硬體設計

為了要對視訊流執行必要的起始延遲,我們設計了如圖 2 的機制來完成任務，frame 到達時如果需要被延遲，則將被存入 shift register 中，(shift register 所存取的單元為 frame)，frame 將在 shift register 中累積，直到其 frame 個數為起始延遲之值時，(此值以 frame time 為單位),爾後每一個 frame time，新到的 frame 加到尾部，而既有的 frame 向右位移一格，被擠出的 frame 就交付網路模組傳送。此項控制由 delay regulator 來達成，於每個 frame time 網路必須傳送完畢來自所有視訊流所委託傳送的 frame。由於網路已預留多工後最大可能的頻寬，故此傳送工作確定可在一個 frame time 之內完成。

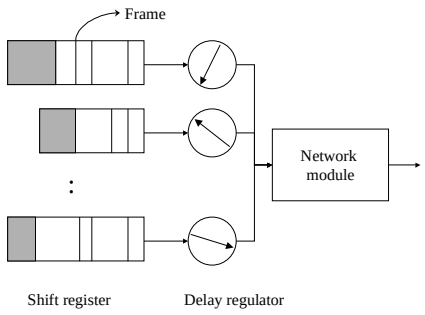

圖 2: 視訊多工延遲機制模型

4. 系統細部說明

4.1 系統定義與假設

本節將更系部介紹我們先前所提的系統架構,同時以舉例的方式描述系統運作的過程,以利讀者瞭解。假設網路骨幹為一串 n 個節點所組成，節點以 1 至 n 編上號碼，我們首先定義有關變數如下：S_i(start,end,D)：表第 i 條進來系統的視訊流，start 是起始節點的編號, end 是終結節點的編號,D 為視訊流所允許的端點至端點的總延遲上限(Delay Bound) 。GOP(i)：表第 i 條視訊流的週期長度； HOP(j)表示第 j 個節點。MC(j)表示第 j 個節點上的多工週期長度；T(j) 表示第 j 個節點上的容許度,I_i^{max} 、P_i^{max}、B_i^{max} 分別為 S_i 中 I frame、P frame、B frame 的最大 frame size。我們

將以此值做為頻寬預留的依據。在之前，我們已說明每一個節點上的多工週期長度的求法，接著我們將舉例說明在我們架構中如何尋找多工的組合及運算容許度時所須要用到的數學式子和其原理。

4.2 最佳多工組合及預留頻寬法則

假設一道新進的視訊流 S_k 欲與 HOP(j)上的多工視訊流以 ID 的起始延遲整合，理論上 ID 值的變化可以從 0 至 min{GOP(k), MC(j)}-1,每一 ID 值的變化,都對應一組多工組合,但是我們必須同時考慮容許度 T(j)的限制,因此 ID 的範圍應介於 0 至 min{T(j)-1, GOP(k)-1, MC(j)-1}之間。換句話說我們必須在這些多工組合中找到一個需求頻寬最小者,在此，我們列舉一例(如圖 3)：

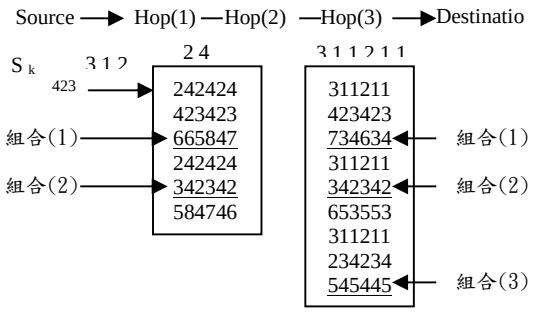

圖 3:視訊多工範例

假設 S_k 從 Source 到 Destination 共有 3 個節點，目前三個節點上的多工週期內資料大小，分別為 <3,1,2>, <2,4> 與 <3,1,1,2,1,1>,有一新的視訊流從 HOP(2)進來，HOP(3)出去，其 T(2)=3 與 T(3)=4,此視訊流必須要跟 HOP(2)及 HOP(3)上現有的多工週期進行多工處理，我們算出新的 MC(2)=LCM(3,2) 與 MC(3)=LCM(3,6) 為 6。以 HOP(2)而言，GOP(k)為 3，而 MC(2)原本為 2，min { GOP(k) , MC(2) , T(2) } = min {3 , 2 , 3} = 2，所以新進來的視訊流在 HOP(2)上最多只能嘗試 2 種多工組合,假設 S_k 在 HOP(2)的多工週期的第一個 frame time 到達,則 2 種多工組合如圖所示,我們選擇組合(2),也就是延遲一個 frame time,因為組合後的頻寬為 4,較組合(1)的 8 為小；同理，在 HOP(3)上，最多能嘗試 3 種多工組合,也假設 S_k 也是於 HOP(3)的多工週期的第一個 frame time 到達, 則 3 種多工組合如圖所示,我們選擇組合(2),它的頻寬需求為 4。如果有一個以上的組合有相同的頻寬需求,我們則選擇,延遲最小的。

4.3 容許度法則

首先我們先算出 HOP(i)上多工週期的平均資料量 μ_i，其公式如下：

$$\mu_i = \frac{\sum_{j=1}^{\mathrm{MC}(i)} X_j}{\mathrm{MC}(i)},$$

其中 X_j 為多工週期中第 j 個 frame times 內的總資料量。接著再算出 HOP(i)上多工週期的資料量變異度 σ^2，其公式如下：

$$\sigma^2 = \sum_{j=1}^{\mathrm{MC}(i)} \left(X_j - \mu_i \right)^2 \text{ 。}$$

最後，我們從起始點至終點，一站一站的依序求出各節點上的容許度 T(j)，其公式如下：

$$T(j) = \left\lfloor \left(D - \sum_{p=start}^{j-1} ID_p \right) \times \frac{\sigma_j^{\,2}}{\sum_{i=j}^{end} \sigma_i^{\,2}} \right\rfloor,$$

ID_j 為在 HOP(j)上，所強制的起始延遲。在此，為了更容易了解，我們舉出一個例子來說明如何運算各站間的頻寬變異度和容許度。如圖 4，共有 3 個節點，其多工週期的平均資料量分別為

$$\mu_1 = \frac{3+1+2}{3} = 2 \quad ,$$

$$\mu_2 = \frac{3+1+1+2+1+1}{6} = 1.5 \quad ,$$

$$\mu_3 = \frac{3+1+2}{3} = 2$$

所以多工週期的資料量變異度分別為

$\sigma_1^2 = (3-2)^2+(1-2)^2+(2-2)^2 = 2$

$\sigma_2^2 = (3-1.5)^2+4*(1-1.5)^2 + (1-1.5)^2 + (2-1.5)^2 = 3.5$

$\sigma_3^2 = (3-2)^2+(1-2)^2+(2-2)^2 = 2$

我們假設目前有一條新的視訊流欲從 HOP(2)進來，HOP(3)離開，且 Total Delay Bound D 為 8 個 frame times。所以：

$$T(2) = \left\lfloor (8-0) * \frac{3.5}{3.5+2} \right\rfloor = \lfloor 5.09 \rfloor = 5 \text{。}$$

(我們在計算容許度的時候，若容許度不恰為整數的時候，則取其下限值，原因是因為所有節點上的容許度相加的值，應該要不大於當初給定的 D，所以便取其下限來避免到最後容許度相加後的值大於 D 的問題產生。) 圖 4 列出可能的多工組合，根據上節頻寬預留的法則得知較佳的組合為新進來的視訊流不須要有任何的 initial delay (也就是 ID=0)就可和 HOP(2)上的視訊流合併，並達到最佳的頻寬使用率，接著根據上述容許度的公式，我們可推算出 HOP(3)上的容許度為

$$T(3) = \left\lfloor (8-0) * \frac{2}{2} \right\rfloor = \lfloor 8 \rfloor = 8 \text{。}$$

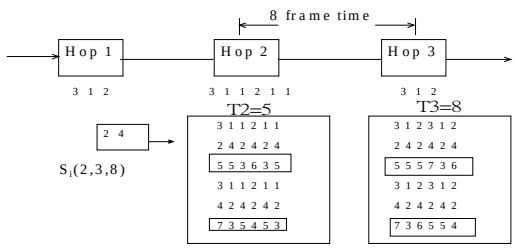

圖 4:求得各節點延遲容許度之範例

5. 系統模擬實驗結果與分析

5.1 模擬架構

我們選用了四部以 MPEG-1 壓縮的視訊影片資料來測試本系統的效能[19]。四部影片抱括 CNN 新聞(CNN)、廣告片(ADVERT)、演講(LECTURE)、與一部電影 (PBRIDE)。其中 CNN 與 PBRIDE 的 GOP 的 frame pattern 為 IBBPBBPBBPBBPBB；而 ADVERT 與 LEC 的 GOP 的 frame pattern 為 IBBPBB。以壓縮比例來說，CNN 與 PBRIDE 兩部影片比 ADVERT 與 LECTURE 兩部影片壓縮比例來得高。圖 5 與圖 6 表示這四部影片前 1000 個畫面的資料量；CNN 與 PBRIDE 兩部視訊影片的 I frame 分佈得比較稀鬆，而 ADVERT 與 LECTURE 的 I frame 則分佈的比較緊密，也就是說前者的交通流量的變異度比較明顯。

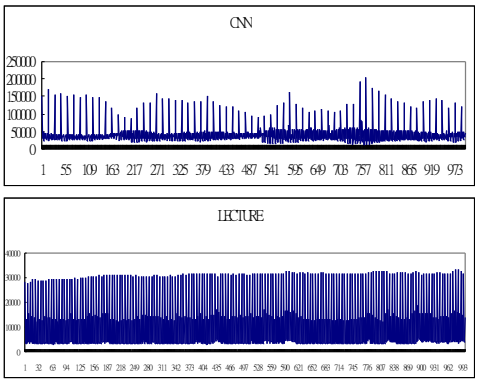

圖 5：CNN、PBRIDE 影片前 1000 個畫面資料量表示圖

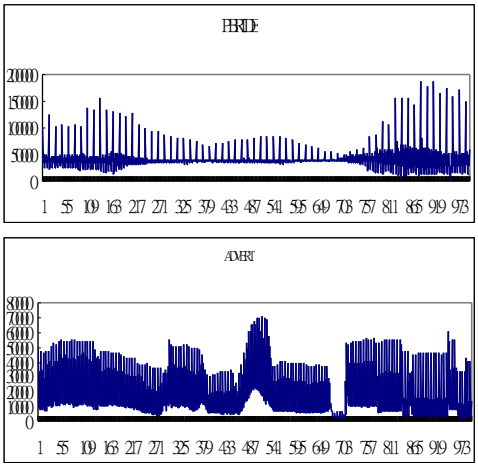

圖 6 : ADVERT、 LECTURE 影片前 1000 個畫面資料量表示圖

在我們所模擬的網路中，共設置四個串接的節點，影片和影片間的播放時間間隔為 exponential 的時間分佈，而新產生的每一條視訊流皆可由四個節點中的任一點進入，然後由下游當中的任一節點離開(所有視訊流皆同向)。每一部影片的播放時間皆為 300 個畫面，由於每一條視訊流都將以 30 frame/sec 的速率產生每一張畫面，所以就相當於我們讓每一部影片播放 10 秒鐘，我們將系統整個的模擬時間定在 3000 個 frame time，也就是 100 秒，在這 100 秒鐘內所產生的每條視訊流在其所經過的節點當中，必須和其他的視訊流進行本論文所提供的視訊多工機制[1]。我們將針對同質視訊流和異質視訊流兩種環境分別進行模擬實驗，以分析頻寬和緩衝區的利用率。同質視訊流環境，係指只選用 ADVERT 一部影片；而異質視訊流環境，則表示任意選取四部影片。

5.2 頻寬使用率

我們定義「頻寬使用率」為「實際傳送的資料量/最大可傳送的資料量」。圖 7 與 8 則分別是在同質環境與異質環境下頻寬使用率對 arrival rate 的變化。由圖 7 中我們可看出，在同質環境下，當我們的 arrival rate 愈高的時候，頻寬使用率愈高，這和我們當初的假設是吻合的，因為 arrival rate 愈高表示視訊流彼此產生的時間間隔愈密集，也正意味著，同一時間點上，節點裏欲進行視訊多工處理的視訊流數目也就愈多，透過彼此某種時序交錯的關係，將會大大的提高頻寬使用率；反之，當 arrival rate 愈低的時候，表示視訊流彼此產生的時間間隔愈長，這正意味著，同一時間點上，節點裏欲進行視訊多工處理視訊流數目就愈少，此時視訊多工的機制能發揮的效用就相對的降低。

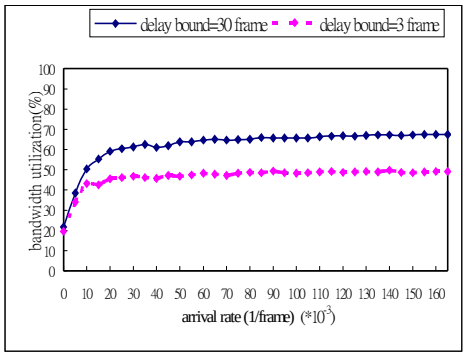

圖 7: 同質環境下，頻寬的使用率

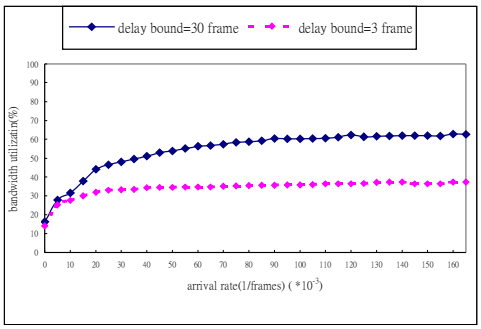

圖 8: 異質環境下，頻寬的使用率

另外，同樣由圖 7 我們也可看出不同的 Delay Bound，也會影響頻寬使用率的高低，Delay Bound 給定的愈高，表示每一個節點所能分配到的容許度就愈高，也表示視訊流間能有更多的時間進行視訊多工處理來找出視訊流彼此間時序交錯的組合，所以相反的，當 Delay Bound 的值給定的愈低，表示每一個節點所能分配到的容許度就愈少，先前的章節我們也提到容許度減少將會造成視訊流間某些時候無法找出視訊流間彼此時序交錯的所有組合，因此便降低了頻寬的使用率。類似的現象也出現在圖 8 的異質環境。

再者，在同質環境下，由於參與視訊多工的影片皆為同一部影片，所以視訊流彼此間資料量的互補性也就愈高，所以和異

[1] 100 秒的模擬範圍雖然短,但影片長度也相對短,而多工調整發生於有視訊流進出活動,與影片長短無關,在此我們觀查到了上百次的視訊流進出活動,故已合理的模擬出網路實際的狀況。

質環境下有不同性質的四部影片參與視訊多工比較起來，某一瞬間的資料量變化量較接近平均值，所以我們可由圖 7 和圖 8 的比較中看出同質環境下的頻寬使用率較異質環境下的頻寬使用率為高。

5.3 緩衝區使用率

我們定義「緩衝區使用率」為「實際經歷緩衝的資料量/最大可緩衝的資料量」，如何求得最大緩衝區的大小，與求得預留的頻寬類似，而主要的不同點是，考慮被延遲的資料量變化。圖9與10則分別是在同質環境與異質環境下緩衝區使用率對arrival rate的變化。由於提升頻寬使用率是我們首要的目標，而非緩衝區使用率，故緩衝區使用率在兩種環境下皆並不穩定，不過一般來說，當arrival rate愈高的時候，大數法則的效應也使得緩衝區的使用率漸漸改善，反之，當arrival rate愈低的時候，緩衝區的使用率也就愈低愈不穩定，這一點和頻寬使用率方面是一樣的。

圖9：同質環境下，緩衝區的使用率

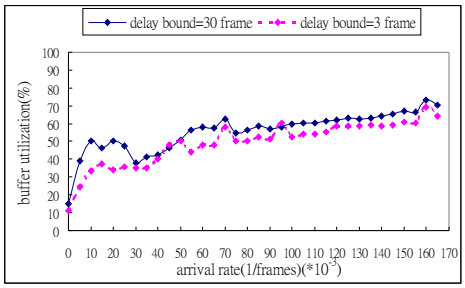

圖10: 異質環境下，緩衝區的使用率

6. 結論

在網路的傳送服務方面，主要區分為兩大類：即 deterministic guarantee 及 statistic guarantee，而各自有不同的需求。Deterministic guarantee 要求沒有資料丟失及些微的延遲，因而使得整體網路的利率降低。相對地，statistic guarantee 則充份利用網路頻寬，使得視訊的品質維持在一可容忍的限度。 就提供 deterministic guarantee 傳送服務而言，通常的做法是依據實際交通的特性推導出一個最佳狀況的上限，然後依據此上限執行允許控制。由於資源的預留是依據最差狀況，而實際的交通量僅偶爾達到最差狀況，因此通常資源的利用率偏低。因此我們透由視訊多工的機制將視訊流間彼此以某一種程度的時序將突發資料的流量錯開，如此一來，便可提降低頻寬的浪費，提高其使用率。

經由本論文中模擬分析的結果，我們得到以下結論：

1. 以視訊多工方式，建構在Deterministic guarantee傳送策略下，在一定可容許的延遲(Delay Bound)，頻寬和緩衝區使用率最高皆可達到60%~70%
2. 在頻寬使用率方面
   - arrival rate愈高，頻寬使用率也就愈高。
   - 同質環境較異質環境好
   - Delay Bound愈高，頻寬使用率愈高
3. 在緩衝區方面
   - 由於緩衝區預留的大小，和每一部視訊流進來系統的時間點有關，所以較難推測某一瞬間緩衝區所須預留的大小。
   - 雖然難以較精準的預測其預留大小，但是一般來說，當arrival rate愈高的時候，緩衝區的使用率也愈高，反之，當arrival rate愈低的時候，緩衝區的使用率也就愈低，這一點和頻寬使用率方面是一樣的。

總而言之,將視訊多工以提升頻寬使用率,而又能同時達到deterministic guarantee是實際可行的傳送策略。

參考文獻

[1] Du D, Hsieh J, Lee H, and Chang T (1996), PCR-Assist CBR for Delivering Pre-Recorded MPEG-2 Transport Streams, INFOCOM96.

[2] Feng W, Jahanian F, and Sechrest S (1997), An Optimal Bandwidth Allocation Strategy for the Delivery of Compressed Prerecorded Video, ACM Multimedia Systems , 1997(5) pp.297-309.

[3] Gall D. (1991), MPEG: A Video Compression Standard For Multimedia Applications, Communications of the ACM, vol.34, no.4, pp.46-58, April 1991.

[4] Handel R, Huber MN. , and Schroder S (1994), ATM Networks : Concepts , Protocols , Applications, Addison-Wesley, ISBN 0-201-42274-3.

[5] Knightly E and Zhang H (1995), Traffic Characterization and Switch Utilization using a Deterministic Bounding Interval Dependent Traffic Model, IEEE INFOCOM95.

[6] Knightly E, Wrege D, Liebeherr J and Zhang H (1995), Fundamental Limits and Tradeoffs of Providing Deterministic Guarantees to VBR Video Traffic, in Proceeding of ACM SIGMETRICS95.

[7] Knightly E (1996), H-BIND : A New Approach to Providing Statistical Performance Guarantees to VBR Traffic, In Proceeding of IEEE INFOCOM ’96

[8] Krunz M and Tripathi SK (1997), Impacet of Video Scheduling on Bandwidth Allocation for Multiplexed MPEG streams, ACM Multimedia Systems , Number (6), Volume (5), pp.347-357.

[9] Lam S, Chow S and Yau D (1994), An algorithm for lossless smoothing of MPEG video, Proc., ACM SIGCOMM, Lodon, England, August 1994.

[10] McManus J and Ross K (1996), Video-on-Demand Over ATM: Constant-Rate Transmission and Transport, IEEE Journal on Selection Areas in Communications, Vol.14, No.6, August 1996, pp.1087~1098.

[11] Ng J and S Song (1997), A video smoothing algorithm for transmitting MPEG video over limited bandwidth, IEEE Fourth International workshop on Real-Time Computing Systems and Applications, Taipei, Taiwan, Oct. 1997, pp229-236.

[12] Ott T, Lakshman T and Tabatabai A (1992), A scheme for smoothing delay-sensitive traffic offered to ATM networks, IEEE INFOCOM92, pp. 776-785, 1992..

[13] Reininger D and Raychaudhuri D (1996), Bandwidth Renegotiations for VBR Video Over ATM Networks, IEEE Journal on Selected Areas in Communications , Vol.14 No.6 , August 1996.

[14] Salehi J, Zhang Z, Kurose J and Towsley D(1996), Supporting stored video: reducing rate variability and end-to-end resource requirements through optimal smoothing, SIGMETRICS96, PA, pp222-231.

[15] Sheau-Ru Tong and Sho-Chi Lee (1999), An Efficient Flow Control Plan for End-To-End Delivery of Pre-stored Compressed Videos, IEEE Multimedia, IEEE Multimedia System99, Florence, Italy.

[16] Vogt C (1995), Quality-of-service management for multimedia streams with fixed arrival periods and variable frame sizes, ACM Multimedia Systems Volume 3, pp.66-75.

[17] White P (1997), RSVP and Integrated Services in the Internet: A Tutorial, IEEE Communications Magazine, pp.100-106, May 1997.

[18] Zhang H and Knightly E (1997), RED-VBR: A Renegotiations-based Approach to Support Delay-sensitive VBR Video, Multimedia Systems , 1997 (5) pp164-176.

[19] Movie tracks are available via anonymous FTP from ftp://tenet.berkeley.edu /pub/dbind.