

## 為隨選視訊而設計之分散式層次輪置檔案伺服器\* A Distributed Hierarchical Striping File System for Video-On-Demand

童曉儒、沈紋正、黃一鋒

國立屏東科技大學資訊管理技術系(所)  
email: {srtong,m8556007}@mail.npust.edu.tw

### 摘要

在目前的隨選視訊(Video-On-Demand)系統發展中,不斷的開發儲存容量與傳輸頻寬,以提供大量的遠距連續媒體資料(如:視訊/音訊)的讀取,乃是主要的設計目標之一。針對此項設計目標,本文將描述一套適用於分散式環境的視訊檔案系統架構,稱之為分散式層次輪置檔案系統(Distributed Hierarchical Striping file system)或簡稱為 DHS-FS。DHS-FS 能夠有效率的集結大量的網路資源以提供視訊讀取使用,其主要藉助於我們所提出的不同系統層次的資料輪置技術,其中包括:伺服器層、磁碟機層與磁區層。

我們也想像未來的系統會不斷的擴大,同時極有可能並存著不同能力的伺服器,而系統必須要能夠提供不同的服務品質來滿足不同的用戶群。為處理如此多樣且複雜的環境,我們在伺服器群之上提出了一套“邏輯伺服器(Logical Video Server)”的觀念,將伺服器的資源重新分配,以滿足不同的效能需求。

DHS-FS 目前尚在開發階段,但從先導的研究分析與實驗結果顯示,如果網路頻寬足夠,它應能彈性且經濟地支援不同規模(數十至上千用戶)的隨選視訊系統。

### Abstract

Continuous exploitation on storage and bandwidth for supporting a large-scale distant retrieval of continuous media (such as video/audio) has been one of the major design goals in contemporary Video-On-Demand systems. This paper describes the system architecture of a distributed video file system, termed *Distributed Hierarchical Striping file system* or *DHS-FS*, for this purpose. DHS-FS effectively aggregates a large amount of network resources for video accesses by employing proposed data striping techniques on various system levels, including *video server*, *disk* and *disk zone* levels.

\*本研究之部份經費由國家科學委員會專題計畫 NSC-86-2213-E-020-001 補助。

We also vision that the system of the future will continuously grow with size and comprise of video servers with different capability. The system should also be able to serve different user groups with different quality-of-service needs. Facing such a complicate situation, we propose a *logical video server* concept upon video servers to re-allocation video server resource to meet different needs.

DHS-FS is currently under development. Some preliminary analyses and experiment results reveal that with the sufficient network bandwidth, DHS-FS is expected to be able to flexibly and economically supporting a VOD system in various scales (ranging from several tens to thousands users).

關鍵字: 隨選視訊 Video-On-Demand、視訊伺服器 video server、分散式視訊檔案系統 distributed video file system、資料輪置 data striping、儲存管理 storage management.

### 1 簡介

隨著電腦與通訊科技的快速進步,網路所傳遞的資訊也由過去傳統的文字型態衍生成今天的多媒體資料,透過網路的連結,人們可以享受到各種不同的多媒體服務,如即時視訊會議、隨選視訊及全球資訊網等。而這些多媒體服務依其應用環境迥異,對系統的架構、資料的即時性與服務品質的需求亦有所不同。以隨選視訊(Video-On-Demand)的應用環境為例,系統大多呈現主從的架構,影音資料事先錄製成檔案存放於伺服器中,再讓使用者再隨時讀取。在資料傳遞的過程,講究的是播映畫面品質的穩定、影音的同步與連續性,而對系統或網路所造成的傳送延遲,要求較不嚴格,只要其變異度能保持在某個固定的範圍之內,適度的暫存區管理即可確保資料流的順暢。由於資料的型態以影片為主,其儲存與讀取的管理策略亦有別於傳統的文字資料,以一部 100 分鐘 MPEG 的影片而言,需佔用 1.1G Bytes 的儲存空間,當播放時需佔用約 1.5Mbps 的頻寬 100 分鐘之久。以目前的技術而言,一部磁碟機最多能儲存的影片數不超過十部;如資料未經特別的放置,由於磁碟機的讀寫頭移動的額外負擔,導致僅能同時提供不到十個使用者。

因此受制於有限的硬體資源下，如何設計一套視訊檔案系統，而能夠應付上百甚至上千人的大型 VOD 系統，是一項迫切且極富挑戰性的工作。

本文將描述一套適用於分散式環境的視訊檔案系統架構，稱之為“分散式層次輪置檔案系統 (Distributed Hierarchical Striping File System or DHS-FS)”。該系統運作在網路環境，將多重伺服器的資源依影片服務品質的需求，重新分配成不同的邏輯伺服器，以供使用者端讀取使用，這個架構的特色在於將如何分配資源以滿足影片服務品質與實際的伺服器的資源分開考量，例如某群影片對應於某個邏輯伺服器，它需要某些頻寬與容量來滿足特定的服務品質需求 (如:起始延遲上限、blocking 機率、解析度等)，而這些頻寬與容量則可取自於任何伺服器中；相對的，一部伺服器的資源亦可與數個邏輯伺服器共享。在此架構下伺服器資源可以更有效的被整合利用，資料更新時對系統的影響度減小，不同能力的伺服器可以隨意地整合，同時提供優越的系統擴充性。

同時該系統為了提升頻寬與容量的使用率，同時確保整個系統負載均衡，在不同的系統層次運用了資料輪置 (data striping) 的技術，首先一部影片切割成數段 video clip，長度約數分鐘，輪置於邏輯伺服器所屬的伺服器間；video clip 再進一步切割成 video segment，長度約數十秒，輪置於伺服器內的磁碟機間；最後 video segment 切割成 video block，長度約數秒，輪置於磁碟機內的磁區間。在我們先導的研究 [Tong97a] 中，以 SCSI-II 為架構的 PC 級伺服器為例，提出了一套稱之為“Skewed Zone-major Round Robin (SZRR)”的資料輪置演算法，運用該演算法後，單一伺服器 (聯結 5 台磁碟機) 最多能提供約 70 道 1.5Mbps 的 MPEG-1 視訊流。我們預測在 DHS-FS 的架構下，結合數十台類似功能的伺服器，只要網路頻寬足夠，提供上千道 MPEG-1 視訊流實際上應屬可行。

在以下的文章中，我們討論的重點在於系統如何在分散式的環境中運作與相關的資料輪置演算法。我們將於第二節中首先對 DHS-FS 的工作原理、系統架構做一介紹。在第三節中將討論各系統層次資料輪置的實際設計考量，同時描述資料輪置演算法。第四節中我們將與其它相關系統做一比較，第五節中以報告目前系統發展現況與未來的發展做為最後的結論。

## 2. 層次輪置檔案系統 (DHS-FS)

### 2.1 工作原理

整個系統可能的架構如圖一所示，數個伺服器 (video server or VS)、管理站 (manager) 與大容量儲存設備 (archive server or AS) (如:磁帶櫃、光碟櫃) 同時連接上高速網路，伺服器將視訊流透過網路送抵用戶端，用戶端可以是 PC，配有 SET-TOP BOX 的 TV，或者經由分配器 (dispatcher) 將信號送至幅員更廣的區

域。較常點選的影片則常駐於伺服器中，如果用戶點選的影片在伺服器中無備份，則系統必需至高容量儲存設備下載全部資料，然後才能供用戶端點選。

我們將伺服器的資源重新配置成數個邏輯視訊伺服器 (logical video server or LVS)，邏輯視訊伺服器是一個抽象觀念，係由許多伺服器所組成，就用戶端而言，它是一個儲存設備實體，邏輯視訊伺服器配有“通道 (stream channel)”，通道的總數即代表該邏輯視訊伺服器最多可同時支援視訊流的總數，通道可能實際散置於數伺服器間，所有儲存於同一邏輯視訊伺服器內的影片稱之為影片叢集 (video cluster)，它們具有相同的服務品質要求 (如:起始延遲上限、blocking 機率、解析度等)，而儲存的方式是將每部影片切割成固定大小的 video clip，然後輪置於通道間。舉例而言，如圖一的系統中，有兩個邏輯視訊伺服器 LVS X、LVS Y，分別儲存了影片叢集 Cluster X、Cluster Y，Cluster X 的影片較熱門但數量少，共需 3Gbytes 的空間與 5 個通道，系統將 VS A 的 4 個通道與 VS B 的一個通道配置給 LVS X，因此佔用了 VS A 2.4GB 與 VS B 0.6GB 的空間；而 Cluster Y 的影片較冷門但數量多，共需 5Gbytes 的空間與兩個通道，系統自 VS B 與 C 中各保留一個通道配置給 LVS Y，因此佔用了 VS B 與 C 各 2.5GB 的空間。

Video clip 對 VS 而言是一個個獨立的視訊檔案 (相繼的 video clip 若位於同一台 VS 上，則事實上可連結成一個檔案，因此在同一 LVS 中，各所屬 VS 中的檔案大小不一定相等)，VS 會將 video clip 再次切割成大小相等的 video segment，輪置於磁碟機間，video segment 再依磁區的容量與頻寬配置切割成大小不盡相等的 video block，然後輪置於磁碟機的磁區間。

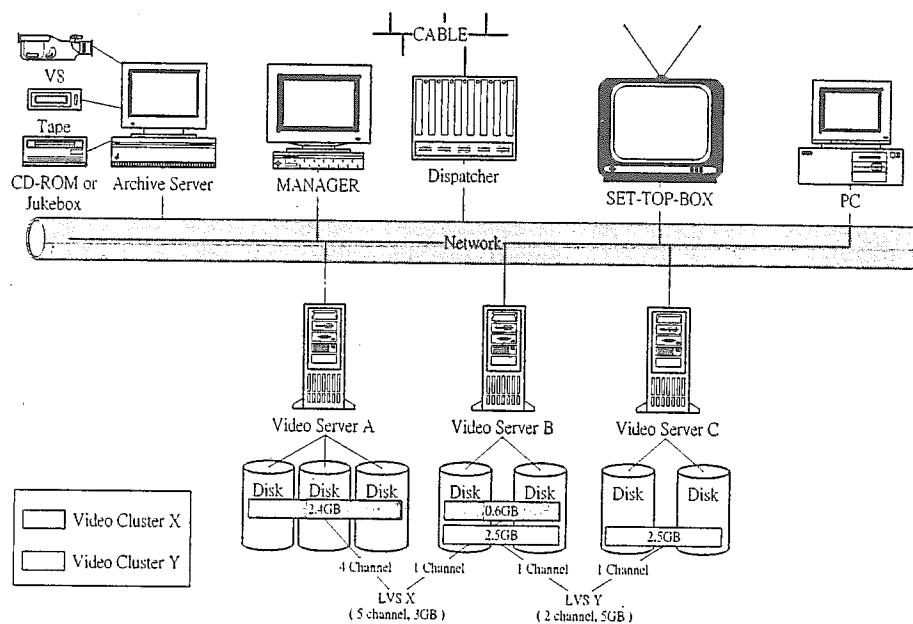
當影片實際播放時，LVS 會負責時間排程的管理，依特定的時間啟動有關的 VS，將資料以資料流的方式送至用戶端，也就是說用戶所看到無間斷的影片，事實上可能來自數個 VS。這個架構有以下幾個優點：

- 邏輯視訊伺服器有效的將伺服器資源整合運用，彈性的提供了各種不同規模的系統的建構方式，同時能夠滿足各種不同的服務品質需求。
- 分散式的架構有效的將系統負載平均於機器間，提升了系統效能，同時提供較佳的容錯性。
- 不同層次的資料輪置，充分的發揮平行輸出的優點，滿足了 VOD 環境中所期盼的高輸出頻寬。

### 2.2 系統設計

本節我們將討論本系統的設計，概略介紹各 process 間的功能與交互關係，整個系統組成包含了五種 process，分別為用 client、manager、logical video server、video server 與 archive processes，其主要功能分述如后：

- Client process：為視訊流的接收端，一般執行在用戶端的機器上，允許使用者可以透過 SET-



圖一：DHS 系統架構圖

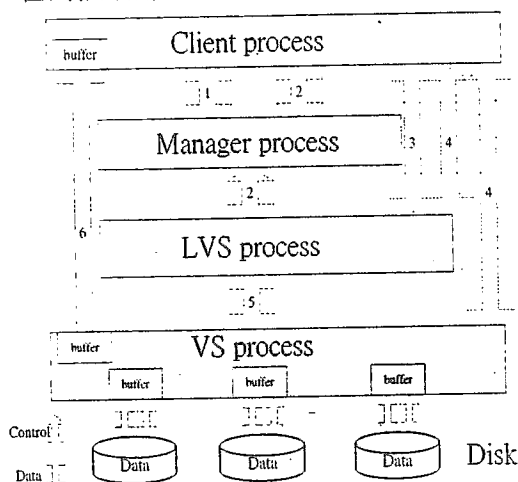
TOP-BOX 或瀏覽器 (Browser) 連上網路觀看影片, client process 在讀取某部影片之前必須先向某個 manager process 註冊。

- **Manager process**：扮演仲介者的角色，負責用戶帳號與所屬的 LVS process 群的建立與使用收費管理，提供影片資料庫 (video meta-database) 以供查詢，建立 client process 與 LVS process 之間的初始聯繫，同時追蹤 LVS process 與 client process 間的動態資料。整個系統可能同時存在一個以上的 manager process，且彼此影片資料庫之間串連成一個影片蒐尋網路，而每個 manager process 所管轄的 LVS process 互不重疊。
  - **Logical Video Server (LVS) process**：負責將影片自 archive process 所管轄的高容量儲存設備中，下載至 LVS process 所專屬的暫存磁碟機中，並切割成 video clip，然後輪置於所屬的 VS 中，當播放影片時，扮演資料傳輸監督者的角色，計算資料傳輸的排程，然後依照決定的排程，命令所屬 VS process 依序將資料流送抵 client process，以確保影片播放的流暢。
  - **Video Server (VS) process**：一個 video server process 控管一台伺服器，負責將 video clip 依 Skewed Zone-major Round-Robin (SZRR) 演算法 (將於第 3.2 節中描述)，切成 video segment，video segment 再細分成 video block 輪置於各個磁碟機的磁區間。當播放影片時，將 video clip 以固定速率資料流的形式送至指定的 client process。
  - **Archive process**：負責將存放在高容量儲存設備的影片資料下載至 LVS process 所控制的暫存磁碟機中。
- 在實際的運作中，除了 VS process 因牽涉到磁碟

機的低階動作，且必須保證達到預期的效能，因而建議用專屬的機器外，其餘的 process 並不侷限於非使用不同機器不可。以下我們以一用戶要求觀看一部影片為例，描述 process 之間的互動關係 (如圖二)：

1. 用戶端啟動一 client process，client process 連線至其所註冊的 manager process，要求觀看某一部影片，當 manager process 接收到要求，且確認用戶的身份無誤後，manager process 便會先從自己的 video meta-database，搜尋影片是屬於哪一個影片叢集中，是由哪一個 LVS process 所控制的。
2. 當找到 LVS process 後，manager process 先檢查影片所屬之 LVS process 是否能接受新的使用者 (manager process 本身有保留其所管轄的 LVS 使用情形的資料)，如果 LVS 仍可服務新加入的使用者，manager process 便會將 LVS process 的 ID 傳回給 client process，一方面 manager process 也會通知 LVS process 有新的使用者加入，並將所欲觀看的影片名稱傳遞給 LVS process，這時接下來的工作便轉移至 client process 與 LVS process 去進行溝通。如果不幸的，此 LVS 已滿載，並不能服務新的使用者，那麼系統便會繼續找是否有其他的 LVS 有相同的影片備份，且有空閒的通道，能提供使用者加入。如果沒有，manager process 會詢問其他的 manager process 是否有同樣一部影片可供使用者加入，如果仍沒有，manager process 會告知 client process，是否等待至最近一個系統有空時再觀賞或離開。
3. 當 client process 接收到 manager process 所回報的 LVS process 位置後，便會和該 LVS process 建立聯繫，等待 LVS process 對 client process 的請求做排程。
4. LVS process 計算好排程後，回覆 client process，何時那一個 VS 的資料會開始到達，client process 便依所得之 VS 的 ID 與他們建立連線，接著等候資料流

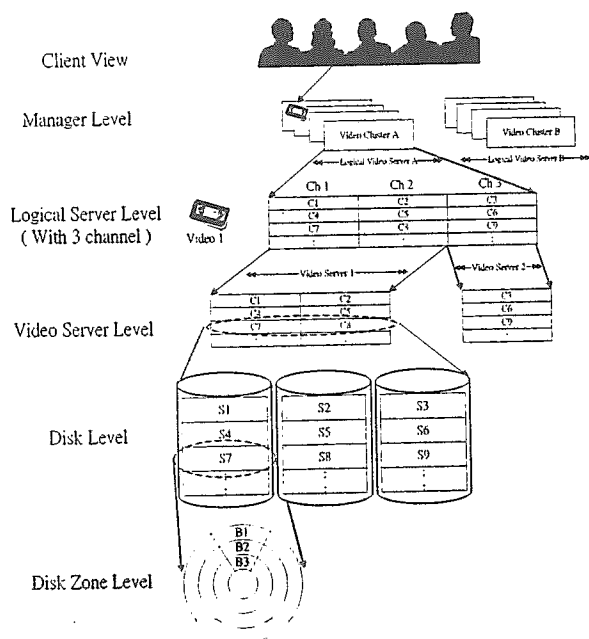
- 的到達。每當某一 VS process 資料流快結束時，client process 會順帶通知 LVS process。
5. LVS process 依據 client process 所傳回的“Video Clip 結束”信號，命令下一個 VS process 在一定的時間範圍內啟動資料的傳送。
  6. VS process 群依序將視訊資料直接送至使用者端，直到影片播完為止。



圖二：系統流程

### 3. 層次輪置原理

資料在各層所呈現的處理單元與各層上下相依的關係，可以用圖三表示。其中資料輪置分別應用在伺服器層、磁碟機層、磁區層。因各層的實際環境不同，在設計上的考量亦迥異，以下將對他們各別討論。

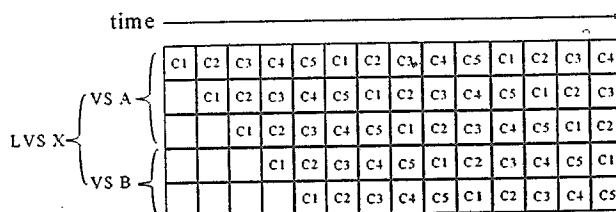


圖三：資料在各層相依的關係

#### 3.1 伺服器層

影片切割成大小相同的 video clip，然後以 round robin 的方式輪流放置於某個 LVS 所屬的通道群中，

通道群則散佈伺服器之間。時間軸分割成許多對應一段 video clip 播映長度的“時區”，某個時區內每道通道，最多僅能允許一個 client 佔用，而 client 會輪流使用所有通道，且不中斷，如果 client 能彼此錯開通道的使用時間，則最多可以讓等於通道總數的 client 同時上線讀取影片（如圖四所示）。要能有效達到上述的多人上線的功能，則有賴由 VS 與 VS、VS 與 client 之間的同步協調；然而在分散式的環境中，網路或系統所造成的延遲變異度較大，要達到不同機器間精確同步實屬不易，因此一種變通的做法就是考慮最大的可能誤差，然後於 VS 或 Client 中安置 buffer 來互補誤差，由於處理機器間的同步是較為繁瑣也消耗資源的工作，因此同步的次數應盡量減少，這也隱喻 video clip 的大小不宜過小，數分鐘到數十分鐘應屬合理的範圍。我們也稱此種資料輪置為“coarse-grained data striping”。在我們先前的研究[Tong96]中已實際驗證，在 VS 與 client 端適度的 buffer 控制，的確能解決不定性延遲的問題，並提供前述的多人同時上線功能。



圖四：一個配有 5 道通道 (3 道來自 VS A; 2 道來自 VS B) 的 LVS，最多 5 個 client 同時讀取 video clip 的時間關係圖。

另一個重要的問題即：如何決定 LVS 所需的資源？我們的方法是將影片依其需求的服務品質加以分類，目前所考慮的服務品質有影片的畫質（頻寬）、影片的起始延遲、影片的讀取率、影片讀取成功率等（在實際的 VOD 系統中這些服務品質應反應在計價公式上），需求相近的影片聚集在一起，稱之為影片叢集（video cluster），一個 video cluster 即對應一個 LVS。LVS 所預留的資源（空間、通道數、通道頻寬）必須要能滿足 video cluster 的服務品質。

下一個問題是 LVS 的資源應如何配置於 VS 間？LVS 資源應配置於那些 VS 間，從放映的角度而言 VS 的配置並無任何差別，但從整體資源使用情形上來看，則須做全盤考量。消極的來說，一些使用不平衡的狀況應予以避免，例如某些 VS 頻寬幾乎用盡，而尚餘大量空間；亦或空間所剩無幾，但頻寬仍閒置許多。因此在每次決定 LVS 資源預留時，我們盡可能的遵循以下簡單的法則，即所有的 VS 盡可能的保持相同的頻寬與空間使用率，如此一來，更多的影片可以納入系統中，也減少影片蒐尋失敗，而必須自 archive server 中下載的機率。詳細的分析可參閱[Tong97b]。另外一個重要的策略會影響系統效能的是如果有影片自 archive server 中載入，而系統資源已滿，則必需自原有 VS 中刪除掉那些影片，以存新進影片？我們希望這種取代對未來系統效能的負面影響減至最低的程度。這方面的研究尚在進行中。

### 3.2 磁碟機層與磁區層

由於此二層的資料輪置息息相關，我們將一併討論。Video clip 在 VS 中會進一步分割成等大小的 video segment，然後以 round-robin 的方式輪置於磁碟機中。爲了節省 VS 中記憶體耗用量，真正自磁碟機中讀取的資料單元並非 video segment，而是將 video segment 進一步的切割成更小的 video block，使得磁碟機在讀出目前的 video block 的同時，系統可以傳送上一個已完整讀出的 video block。

每一次 video block 的讀取，磁碟機讀寫頭必須移動位置到 video block 的起始位址，然後才開始讀的動作，這個過程十分耗時，因此如何安排 video block 所擺放的位置，以致於能減少讀寫頭來回地移動，便成了提升磁碟機效能的重要設計。最常採用的方法就是“constrained allocation”，基本上將磁碟機的磁碟上區分成幾個環形區域，稱之爲“磁區”，然後將 video block 依次輪置於其間，讀寫頭朝一個方向移動，每移動到一個磁區，一組 video block 就被取出，每一個 video block 對應一道資料流，讀寫頭反覆此種掃描動作，數道資料流因而形成。

除此之外，目前市面上較先進的磁碟機，在構造上爲了要提升磁碟容量，往往事先將磁碟規劃成幾個資料密度相等的環形磁區，在固定的磁碟轉速下，導致外圍磁區有著較內圍磁區高的傳輸速率，且高低比率差異相超過 20%，這個設計增加了資料輪置與讀取的複雜度。爲了解決這個問題，我們採用 fixed- $\pi$  演算法，磁碟重新規畫成數個等“磁區容量/傳輸速率”比率的磁區，而 video block 也依磁區容量的大小有所調整，如此一來，磁區容量與傳輸速率皆能接近 100% 的被利用。

另一個實際的問題是有關於連結磁碟機的 SCSI bus，由於 SCSI bus 所採用的 access protocol 爲 contention-based，也就是說有可能同時有一個以上的磁碟機同時企圖傳送資料，而只有優先權較高的磁碟機能夠成功，這樣的設計當交通量過於頻繁時，會導致 SCSI

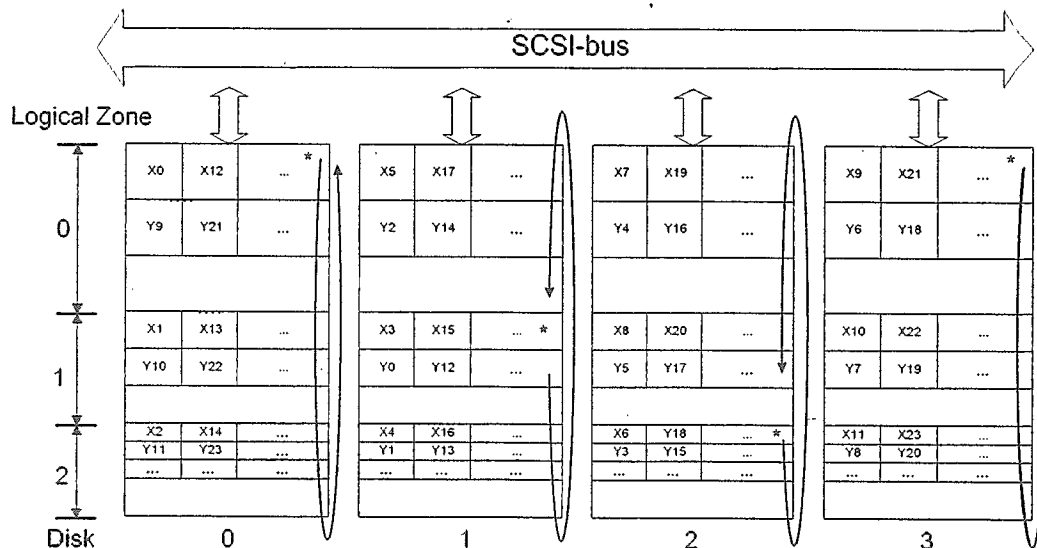
bus 效能降低至 70% 左右。回過頭來看，不同的磁區產生不同的資料傳輸速率，如果磁碟機彼此不加以約束，則很可能產生瞬間的高資料量，因此限制了 SCSI bus 的效能，爲解決這個問題，我們採用所謂的 Skewed Zone-major Round Robin (SZRR) 輪置演算法(如圖五)。基本的原理爲規律地錯開相鄰磁碟機資料擺放的磁區位址，使得在任何時間所有磁碟機輸出頻寬總合趨近一個常數值。經過實驗的驗證，5 部磁碟機使用 SCSI-II bus (20Mbps) 串聯，可提供到 70 道 1.5Mbps 的 MPEG-1 視訊資料流。詳細的討論請參閱[Tong97a]。

### 4 與其它相關系統之比較

以往的研究文獻提出不少的視訊檔案系統，以下僅對其中幾個具有類似於本系統所提的分散式與層次資料輪置設計的系統做一比較，首先 Microsoft Tiger 系統[Bolosky96]提出將資料輪置於磁碟機與伺服器的架構，伺服器間藉由高速 ATM 網路聯結，伺服器(磁碟機)必需具備相同硬體架構，輪置的資料單元大小必需一致(64K~1Mbytes)，因此僅能提供單一服務品質視訊，由於資料單元相對的較小，在資料讀取時，伺服器間要求較精確的同步，爲此伺服器間另配置一個專屬控制網路，負責控制信號的交換。

University of Minnesota [Hsieh95] 研究由許多 RAID-3 磁碟機陣列所組成的高效能伺服器，提出了 application level、logical volume、single byte-striped RAID-3 等不同的層次資料輪置，其中 application level 與 DHS-FS 中的 LVS 機制類似，可以彈性的針對應用程式組成所需的資源，但其動機與 DHS-FS 稍有不同，且對 application level 資源分配的問題並無深入討論。其它兩層則著重在如何在多重 RAID-3 的環境中輪置資料，主機則爲高效能工作站。

UC Berkeley[Brubeck96, Rowe95]提出了 Distributed VOD System (DVS)，該系統架構上是採分散式的策略，系統中同時存在數個 video server 與 archive server，但彼



圖五：以 4 部磁碟機爲例的 SZRR 資料擺放法，\* 表示起始磁區

此獨立，video server 內部構造並未詳述，video server 之 server 與 archive server 之間的互動關係，希望減低必須到 archive server 下載資料的機率。

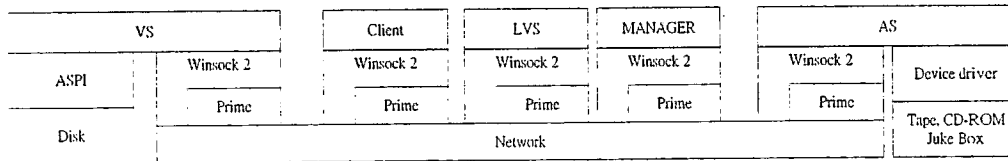
### 5. 結論

一個適合 Video-On-Demand 的檔案系統，為了因應能即時點播影片的需求，需要大量的儲存空間與頻寬，為達此目的 data striping 不失為一個值得考慮的技術，主要的原理為結合一群設備，實施輪流短暫服務的方式，分散系統負載並達到平行處理的效果。在以往的 data striping 研究中，多針對以磁碟機為主的區域環境。而本文中提出一個分散式的檔案系統，其中將 data striping 的觀念落實到不同的系統層次，其中包括：video server、disk、disk zone，因此系統效能更能提升。

在我們所研究的環境中，假想當系統因用戶的增加而不斷擴張時，勢必同時存在許多不同效能的 video server，同時系統必需提供不同價位的影片服務來滿足用戶不同的消費觀，因此我們提出了 logical video server 的觀念，彈性的重新分配 video server 的資源，然後影片在 video server 間輪置。而就 video server 本身的設計而言，我們重點放在最為簡單、經濟且普遍的 SCSI-based, multiple disk 架構上，我們假設所有相連的 disk 有相同的型號，且具有

zoning 的特性，我們首先將 disk 依 fixed- $\pi$  演算法重新切成數個磁區，再採用 SZRR 輪置演算法輪置資料。先前的研究中顯示一部 Pentium 133 的 PC 伺服器可連結 5 部 SCSI-II 的 disk，能同時支援到 70 道 1.5Mbps 的 MPEG-1 視訊資料流。若結合數十部 video server，只要網路頻寬夠寬，上千人同時上線應是可行的。

目前本系統仍在開發階段，disk 與 disk zone 層的資料輪置已開發完成，目前正積極開發 video server 層的有關 process。發展平台為 WINDOWS NT 4.0，磁碟系統使用 fast and wide SCSI 磁碟系統，由於 WINDOW NT 所提供的檔案系統並不能充分支援連續媒體的讀取，因此我們透過 ASPI (Advanced SCSI Programming Interface) 介面[ANSI93]，對磁碟機做低階的資料存取；另在網路部份，我們遵照 WINSOCK 2.0 協定介面，提供上層設定服務品質的機制，而實質上資料流的傳輸乃透過我們在 WINSOCK 2.0 內部所加入自行發展的即時排程協定，稱之為 Prioritized-channel Real-time Scheduling (PRIME) (如圖六所示)。PRIME 能提供具有服務品質的固定速率資料通道，目前服務品質是以“資料週期”與“每週期資料量”為主要表示參數。因空間有限，網路傳輸的細節將予以省略，有興趣的讀者請參閱相關論文[Tong94, Tong95]。



圖六：系統平台

### 參考文獻

[ANSI93] American National Standard of Accredited Standards Committee X3. Small Computer System Interface-2, ANSI X3T9.2.Draft Revision 10L, 1993.

[Bolosky96] William J. Bolosky, Joseph S. Barrera, III, Richard P. Draves, Robert P. Fitzgerald, Garth A. Gibson, Michael B. Jones, Steven P. Levi, Nathan P. Myhrvold, Richard F. Rashid. The Tiger Video File server. In Proceedings of the Sixth International Workshop on Network and Operating System Support for Digital Audio and Video. IEEE Computer Society, Zushi, Japan, April, 1996. <http://www.research.microsoft.com/research/os/>

[Brubeck96] David W. Brubeck and Lawrence A. Rose, University of California at Berkeley, "Hierarchical Storage Management in a Distributed VOD System", IEEE MultiMedia, Fall 1996

[Hsieh95] J. Hsieh, M. Lin, C. L. Liu, and D. H. C. Du, "Performance of a Mass Storage System for Video-On-Demand", Journal of Parallel and Distributed Computing: Special Issue on Multimedia Processing and Technology, Vol.30, No.2, November, 1995, pp. 147-167.

[Rowe95] Lawrence A. Rose, John S. Boreczky, David A. Berger, David W. Brubeck and J Eric Baldeschwieler, University of California at Berkeley, "A Distributed

Hierarchical Video-On-Demand System", 1995 International Conference on Image Processing, Washington DC, October 1995.

[Tong94] S. R. Tong, *etal*, "PRIME: A Prioritized-Channel Real-Time Service Architecture for Multimedia Communications", HD-Media Technology and Applications Workshop, Oct. 1994.

[Tong95] S. R. Tong, L. Y. Lu and S. G. Jiang, "Design of Integrated Virtual Video Servers in a Distributed Heterogenous Environment", Proceedings of National Computer Symposium, Dec. 1995, pp.275-282. <http://140.127.22.195/lab/mnlab/chinese/FTP.HTM>

[Tong96] S. R. Tong and L. Y. Lu, "A Network-based Cooperative Server Cluster: An Adaptive Coarse-Grain Striping Approach for Multimedia Retrieval", The Pacific Workshop on Distributed Multimedia Systems, Hong Kong, June 1996, pp171-180. <http://140.127.22.195/lab/mnlab/chinese/FTP.HTM>

[Tong97a] Sheau-Ru Tong and Yee-Foon Huang, "Study on Disk Zoning for Video Servers", submitted to Infocom97. <http://140.127.22.195/lab/mnlab/chinese/FTP.HTM>

[Tong97b] Sheau-Ru Tong, Bandwidth and Storage Allocation Strategies for Adaptive Coarse-grained Data Striping, working paper.