

支援可適性視訊編碼之內容傳遞網路的設計

周韶芬、吳和庭*、柯開維、楊士萱、唐偉倫
國立臺北科技大學資訊工程學系

摘要—內容傳遞網路 (Content Delivery Network) 將檔案放在不同的邊緣伺服器中，以減輕原始伺服器的負擔，並有效的服務使用者請求，增進服務品質。本論文以支援可適性視訊編碼 (Scalable Video Coding) 的內容為訴求，設計相對應的內容配置機制。同時，內容管理伺服器也會持續監控所有的邊緣伺服器，移除使用率過低的內容。此外，考慮到系統負載過重時使用者必須等待的狀況，本論文設計了頻寬分配機制，讓持有較多頻寬的使用者釋放部份頻寬，以減少使用者的平均等待時間。我們利用模擬結果來評估此設計的系統效益。¹

一、緒論

網際網路技術近年來的進步，使用網際網路的人口數和網路資訊量也急遽增加，早期的網路架構已不敷使用，因此內容傳遞網路 (Content Delivery Network, CDN) 被提出。CDN 主要是利用一台原始伺服器及數台邊緣伺服器來服務使用者[1,2]。依照管理人員的決策，將邊緣伺服器放在不同的區域提供當地使用者服務，再根據使用者喜好決定邊緣伺服器中的內容。使用者請求也可以被分散到不同的邊緣伺服器上，邊緣伺服器便可就近服務使用者。當邊緣伺服器發生滿載或當機的狀況時，使用者也能在其他邊緣伺服器取得需要的檔案[3]。

由於現在多媒體服務的興起，多媒體業者除了使用 CDN 來提供服務外，還利用可適性視訊編碼技術 (Scalable Video Coding, SVC) 為影片做分割，省去系統重複做影片編碼的動作，同時也能夠給予使用者相對應的服務[5]。然而，系統管理者無法預知這個地區的使用者偏好的檔案，並將檔案放在邊緣伺服器中[4]。當使用者不能在近的邊緣伺服器取得需要的檔案時，他必須向較遠的原始伺服器下載檔案，除下載時間增加外，還有可能發生封包遺失的情形。當系統管理者放置了使用者不需要的檔案時，這些檔案將佔用一部分的儲存空間。

本論文以 SVC 技術所分割出來的影片為背景，設計一個在邊緣伺服器中新增和刪除內容的機制。同時，考慮到邊緣伺服器頻寬的限制，可能發生系統沒有足夠頻寬服務使用者的狀況，本論文提出了分配使用者頻寬的機制，期望能夠降低使用者的平均等待時間。

二、內容傳遞網路系統之設計

2.1 內容傳遞網路系統架構

本論文使用的影片解析度格式為常用視頻標準化格

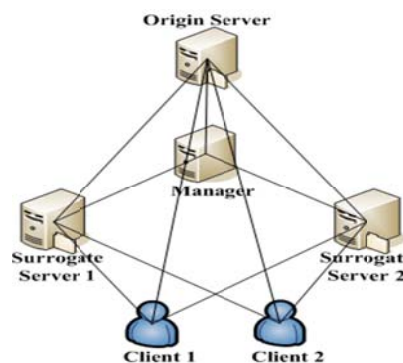
式 (Common Intermediate Format)。我們不考慮影片解析度大小的改變，首先利用時間可適性編碼技術將影像切割成五個階層，再利用品質可適性編碼技術將影像分成六個分層。因此最後我們可以得到三十個分層，也就是三十個內容，如圖一。我們以兩位數字來表示每一個內容，十位數字代表的是時間層，個位數字代表的是品質層。其中，00 為基礎層，其餘皆為增強層[6]。

	Low fps					High fps				
High quality	05	15	25	35	45					
	04	14	24	34	44					
	03	13	23	33	43					
	02	12	22	32	42					
	01	11	21	31	41					
Low quality	00	10	20	30	40					

圖一：使用時間可適性以及品質可適性編碼技術將影像做分層

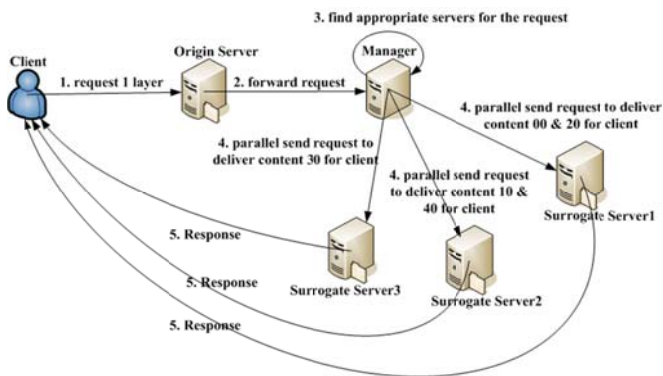
我們主要討論使用者請求不同品質的影片的狀況，因此使用者必須以品質層做為單位來請求，取完一層才能繼續取次一層的內容。舉例來說：當使用者請求兩個品質層的影像時，他必須先取得第一層的 00、10、20、30、40 後再請求第二層的 01、11、21、31、41。必須整層的畫面能夠傳送到使用者端時，使用者才能夠解碼出這一層品質的影片。由於在 SVC 技術中，即使只取得基礎層也能夠觀看最低品質的影片。因此我們將基礎層定義為最低階的品質層。

本論文使用的 CDN 系統主要由使用者、原始伺服器、內容管理伺服器 (Content Manager)、和邊緣伺服器構成。原始伺服器存放所有的內容，每台邊緣伺服器固定放置基礎層和不一定數量的增強層內容。圖二為本論文所使用之 CDN 系統的網路拓樸架構示意圖，由於連線數較多，我們以兩台邊緣伺服器 and 兩位使用者為例。



圖二：CDN 系統網路架構示意圖

¹ 本研究由國科會贊助，計畫編號 NSC102-2219-E-027-003。



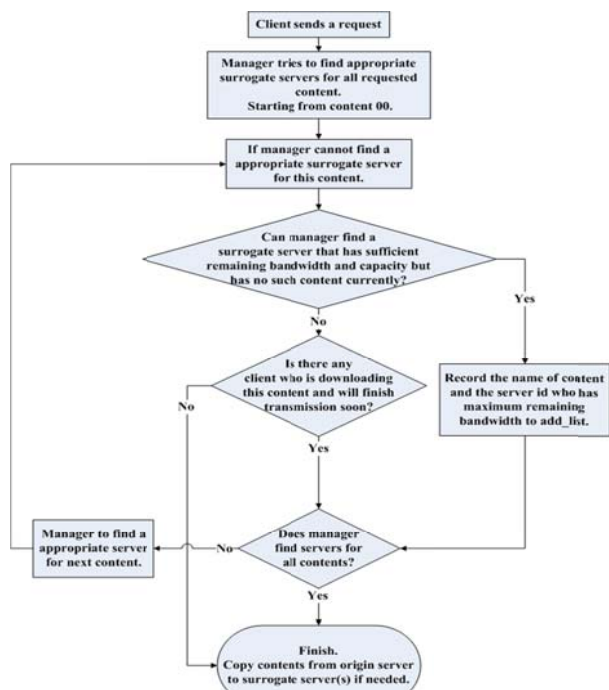
圖三：CDN 系統運作圖

圖三為 CDN 系統的運作圖。當 Manager 接收到使用者請求後，Manager 會按順序對使用者需要的內容找出合適的邊緣伺服器。若無法為某個內容找到能夠提供服務的邊緣伺服器時，便終止配對的動作，並就目前配對的結果平行傳送請求給邊緣伺服器。舉例來說：當使用者請求兩個品質層時，Manager 便從內容 00 開始做配對的動作。假設 Manager 無法為內容 11 找到可以提供服務的邊緣伺服器，便終止動作，並將先前配對的結果以階層為單位平行傳送請求給邊緣存伺服器。也就是說，Manager 將捨棄內容 01，並利用平行的方式傳送內容 00、10、20、30、和 40 的請求給邊緣伺服器。

由於我們將基礎層放在每一台邊緣伺服器中，因此若系統無法提供基礎層時，代表當前系統過於繁忙，必須等待一段時間再讓 Manager 重新做配對的動作。

2.2 內容新增機制

內容新增機制主要是在 Manager 無法找到邊緣伺服器來提供某個內容時啟動，圖四為流程圖。



圖四：內容新增機制流程圖

Manager 根據以下條件篩選掉不適合的邊緣伺服器：

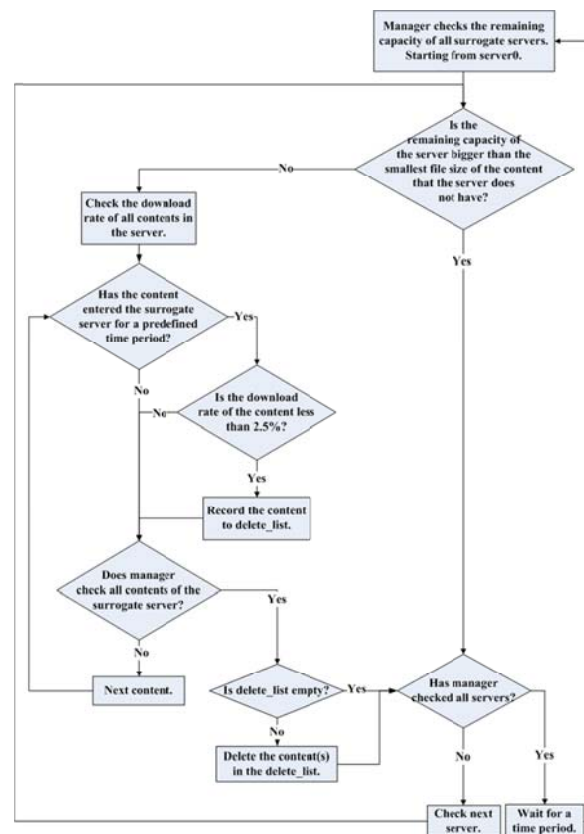
- (1) 已經有此內容的邊緣伺服器
- (2) 容量不夠存放此內容的邊緣伺服器
- (3) 頻寬不足以傳送此內容的邊緣伺服器

接著從剩下的邊緣伺服器中，選出一台剩餘頻寬量最多的做為內容新增的邊緣伺服器，並在 add_list 中。當 Manager 無法找到一台能夠做內容新增的邊緣伺服器時，若我們能夠找到一個正在取得該內容並即將結束傳輸的使用者，則我們就可以假設在很短的時間之內，原使用者能夠取得該內容。但是，如果 CDN 系統內不存在一個即將結束傳輸的使用者，則判斷原使用者無法取得該內容，內容配對的動作就必須在此結束。

待整個流程結束後，若 add_list 中有記錄存在，則 Manager 會以階層為單位，發送請求複製內容的封包給原始伺服器，讓原始伺服器複製內容到邊緣伺服器中。

2.3 內容刪除機制

在 CDN 系統中，Manager 會定期的對所有邊緣伺服器做監控，當邊緣伺服器的容量達到上限時，便啟動內容刪除機制以刪除使用者不需要的內容，或是因為系統資源限制，導致使用者無法取得的內容。我們使用下載率作為刪除檔案的指標，當內容的下載率低於刪除門檻時，我們認定該內容不被使用者所需要。我們根據測試的結果，將刪除門檻設定為 2.5%。圖五為內容刪除機制流程圖。



圖五：內容刪除機制流程圖

當邊緣伺服器剩餘的容量無法再放入他所沒有的檔案中檔案大小最小的檔案時，便判斷他的容量達到上限。

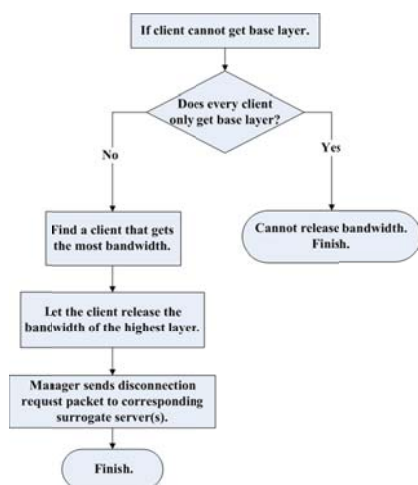
Manager 會逐一檢查該伺服器中每一個內容的進入時間以及內容在邊緣伺服器中實際被下載的次數，將下載次數過低的內容刪除。我們利用下載次數來定義下載率 (download rate) 如下：

$$\text{下載率} = \frac{\text{內容進入邊緣伺服器後，內容被下載的次數}}{\text{內容進入邊緣伺服器後，邊緣伺服器被請求的次數}} \quad (1)$$

為避免內容進入系統後馬上被刪除的狀況，我們會給予每一個新進入邊緣伺服器的內容一段適應時間。適應時間定義為一個影片長度的時間。若某內容已經進入系統一段時間且該下載率低於門檻，則把他加到 delete_list 中。當檢查完畢且 delete_list 中存有記錄，Manager 會發送請求刪除內容的封包給各別的邊緣伺服器。邊緣伺服器收到請求刪除內容的封包後，便會將該內容刪除。

2.4 頻寬分配機制

由於基礎層被固定在每一台邊緣伺服器上，因此當系統負載過重，連基礎層都無法提供時，使用者就必須等待其他使用者結束傳輸，才可能得以進入系統。當邊緣伺服器的頻寬愈少，部份的使用者佔用了太多頻寬時，這種等待的情形也發生的愈頻繁。本論文設計了頻寬分配機制，讓系統中持有最多頻寬的使用者釋放最高層的頻寬。希望藉此減少上述使用者等待的情形。圖六為頻寬分配機制流程圖。



圖六：頻寬分配機制流程圖

在本論文中，為了盡量提供使用者最高品質的服務，我們設定當有使用者無法進入系統時，頻寬分配機制才會被啟動。Manager 首先會檢查所有使用者持有頻寬的情況。由於基礎層為最低品質影片的層數，若所有使用者皆只取得基礎層，則此機制便終止。否則，Manager 會找出一個持有最多頻寬的使用者，強制讓他的影片品質下降，以釋放部份頻寬。

三、 模擬結果及分析

3.1 模擬環境及效能指標

本論文使用的模擬工具為 OMNeT++[7]。我們的 CDN 環境由一台原始伺服器、十五台邊緣伺服器、一個

Manager 和使用者所組成。原始伺服器中存放所有內容，邊緣伺服器放置基礎層和不一定數量的增強層。當邊緣伺服器接收到使用者請求時，會根據該內容的 bit-rate 大小決定該內容需要的傳輸頻寬。我們只模擬單一影片檔案下載的狀況，並使用以下指標評估 CDN 系統效能的改變，分別為使用者平均等待時間 (avg. waiting time)、平均滿意度 (avg. satisfaction)、平均 PSNR (avg. PSNR) [8]。定義如下：

等待時間 = 使用者收到第一張畫面的時間 - 送出請求的時間

$$\text{滿意度} = \frac{\sum_{i=1}^n (\text{使用者第 } i \text{ 段時間取得的檔案大小} * \text{第 } i \text{ 段持續時間})}{\text{使用者請求的檔案大小} * \text{影片總時間長度}}$$

$$\text{PSNR} = \frac{\sum_{i=1}^n (\text{使用者第 } i \text{ 段時間取得內容的 PSNR 值} * \text{第 } i \text{ 段持續時間})}{\text{影片時間長度}}$$

n = 在檔案下載的總時間長度中，檔案根據影片下載品質的差異，被切割成的時間區段之個數

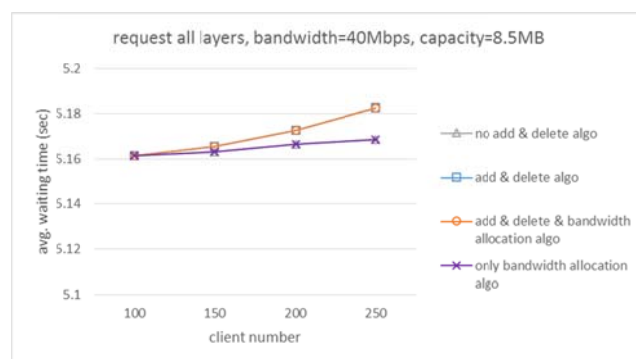
在計算上述三個指標時，在模擬過程中，我們將所有使用者每次下載影片所得之數值平均，即可得到使用者平均等待時間、平均滿意度、與平均 PSNR 值。

3.2 頻寬限制的改變

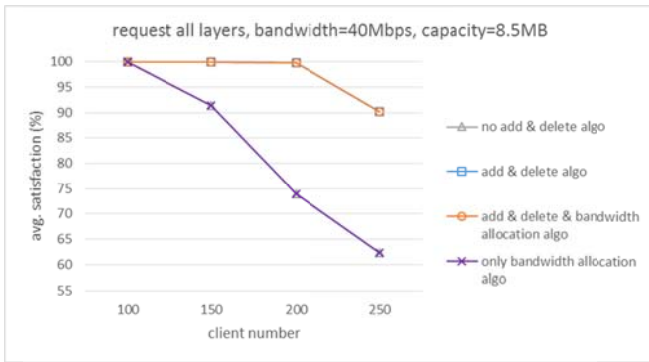
我們觀察在不同的使用者數量和邊緣伺服器的頻寬下，四種機制組合在效能上的變化。四種演算法組合為：無演算法、頻寬分配機制、內容新增機制 & 內容刪除機制、內容新增機制 & 內容刪除機制 & 頻寬分配機制。

表 I
頻寬改變環境下之模擬參數設定

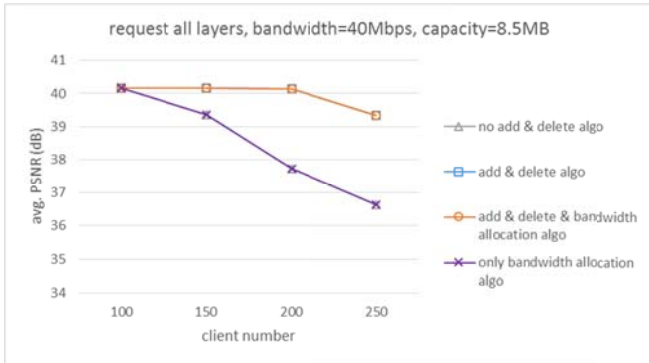
Surrogate server capacity	8.5MB
Simulation time	7200s
Video length	60s
Request arrival pattern	Exponential distribution, mean inter-arrival time = 30s
Bit-rate of content (kbps) (starting from content 00、10、20、...、45)	183.1、75.8、75、88、101.3、170.6、47.2、50.5、69.4、63.8、116.1、38.1、47、57.8、73.1、119.8、54.1、62.9、98、92.4、227.8、116.8、177.3、252.7、313.4、170.5、93.1、129.3、220.1、241
Client request pattern	Request all layers
Initial content placement	Put each enhancement layer content in only 1 surrogate server, and put base layer content in every surrogate server.



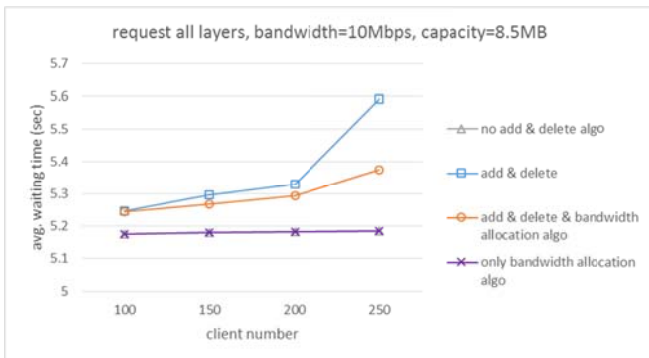
圖七：使用者請求最高品質且頻寬 40Mbps 時使用者平均等待時間



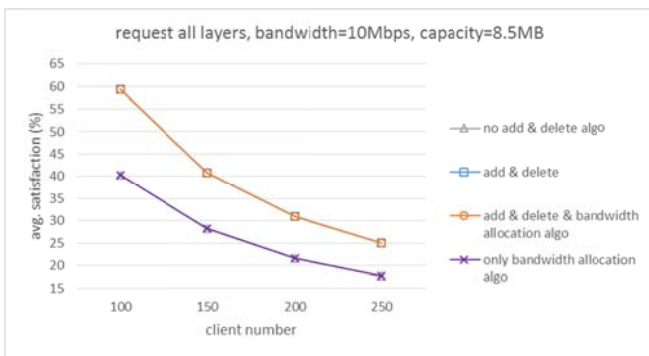
圖八：使用者請求最高品質且頻寬 40Mbps 時使用者平均滿意度



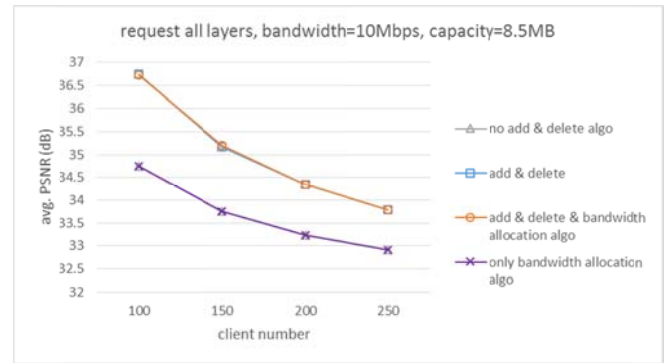
圖九：使用者請求最高品質且頻寬 40Mbps 時使用者平均 PSNR



圖十：使用者請求高品質且頻寬 10Mbps 時使用者平均等待時間



圖十一：使用者請求高品質且頻寬 10Mbps 時使用者平均滿意度



圖十二：使用者請求高品質且頻寬 10Mbps 時使用者平均 PSNR

當邊緣伺服器的頻寬充足時，如 40Mbps，內容配置（新增及刪除）機制能夠大幅提高使用者的滿意度，如圖八，使用者收看的影片品質也相對的提高，如圖九，而使用者所要付出的只是少許的等待時間，如圖七。然而，當邊緣伺服器的頻寬太小時，如 10Mbps，即使 CDN 系統替每位使用者進行內容新增，使用者仍受限於頻寬因素而無法取得較高品質的內容，如圖十二，整體使用者的平均滿意度無法如預期的提升，如圖十一。甚至還會發生因為系統新增內容，讓部分使用者佔用過多頻寬，造成使用者的平均等待時間增加的情形，如圖十。另一方面，當邊緣伺服器的頻寬充足時，系統有足夠的頻寬應付所有使用者，因此並沒有發生使用者無法進入系統的情形，頻寬分配機制也未被觸發。但是當邊緣伺服器的頻寬縮小至 10Mbps 時，無法進入系統的使用者數量也開始增加。若系統中只有內容配置演算法，整體的使用者平均等待時間比較高一些。當系統加入頻寬分配演算法後，使用者無法進入系統的情形獲得改善，使用者平均等待時間相對的也減少了。

參考文獻

- [1] M. Day, B. Cain, G. Tomlinson, and P. Rzewski, "A Model for Content Internetworking (CDI)," Internet Engineering Task Force, RFC 3466, Feb. 2003.
- [2] A. Vakali and G. Pallis, "Content Delivery Networks: Status and Trends," IEEE Internet Computing, IEEE Computer Society, pp. 68-74, Nov.-Dec. 2003.
- [3] K. Pathan, and R. Buyya, "A Taxonomy and Survey of Content Delivery Networks," Tech Report, University of Melbourne, 2007.
- [4] G. Pallis and A. Vakali, "Insight and Perspectives for Content Delivery Networks," Communications of the ACM, vol. 49, no. 1, ACM Press, NY, USA, pp. 101-106, Jan. 2006.
- [5] H. Schwarz, D. Marpe and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, pp. 1103-1120, Sep. 2007.
- [6] A. Segall and G. J. Sullivan, "Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension," IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, pp. 1121-1135, Sep. 2007.
- [7] OMNeT++, URL: <http://www.omnetpp.org/>.
- [8] Peak signal-to-noise ratio, URL: http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio/.