

分散式平行放置檔案傳輸協定

吳炯達

朝陽科技大學

網路與通訊研究所

s9530607@cyut.edu.tw

廖俊鑑

朝陽科技大學

網路與通訊研究所

jjliaw@cyut.edu.tw

摘要

本論文提出一個分散式平行放置檔案傳輸協定(Distributed Parallelized File Transfer Protocol, 簡稱 DP-FTP), 這個方法在固定檔案伺服器數量的環境下, 提供分散且平行的傳輸方式, 使得使用者可以快速地下載大型檔案, 並且達到各個檔案伺服器之間的負載平衡。當使用者提出一個檔案傳輸請求時, 最快回應的檔案伺服器會提供這次檔案傳輸的服務, 並找尋附近可以支援的檔案伺服器來支援此次的檔案傳輸服務。我們的方法會依照各檔案伺服器之間的傳輸速度, 決定各個檔案伺服器所要傳輸的檔案大小。也就是說, 每個檔案伺服器依其傳輸能力提供服務, 使網路提供最有效率的傳輸服務。

關鍵詞：平行放置檔案傳輸協定、分散式傳輸、平行處理。

Abstract

We propose a Distributed Parallelized File Transfer Protocol (DP-FTP) to provide a distributed and parallelized environment of a fixed number of file servers. DP-FTP makes users to download large size file quickly and it also balances the loading of file servers. When a client requests a file, the server which responds the request fastest is called the main server. The main server not only transfers the request file to the client, but also finds adjacent servers to support the transferring work. The request file is segmented in to several parts. According to the transmission speed of each support server, the main server decides the transfer loading of each support server. In the other words, support servers transfer different rate of the request file based on their transmission ability to improve the efficiency of the network.

Keywords: Parallelized File Transfer Protocol, Distributed Transmission, Parallel Transmission

1. 前言

在網際網路的時代, 使用者對資訊檔案的下載日漸頻繁, 使得檔案下載服務跟網際網路密不可分。網路上的使用者可以經由檔案下載服務向某個檔案伺服器進行下載, 進而取得所需檔案。藉由網路來傳輸檔案, 必須依靠一種可靠且有效率的傳輸協定。目前在檔案下載服務協定中, 檔案傳輸協定(File Transfer Protocol, 簡稱 FTP[5])是目前網路上最可靠且廣泛使用的檔案傳輸協定。當使用者要利用 FTP 進行檔案下載時, 一般是透過地理位置上最近的檔案伺服器來提供服務, 因為選擇地理位置最近的檔案伺服器理論上可以提供最快的檔案傳輸服務, 進而取得最少的傳輸時間。以這種選擇檔案伺服器方式的技術稱為單一下載技術(Single Download Technique, 簡稱 SDT)[7], 所謂的 SDT 是從單一檔案伺服器透過 TCP/IP 協定來下載檔案。這個技術中, 又可以區分為兩種方法來選擇檔案伺服器: SDT-R 是以隨機的方式選擇可利用的檔案伺服器; SDT-L 是以地理位置上選擇距離客戶端最近的檔案伺服器。

然而隨著傳輸的資料越來越大, 單一檔案伺服器的服務已經無法滿足使用者的需求, 因此除了選擇最短距離之外, 也開始加入其它因素來探討多個檔案伺服器同時服務。Guyton 和 Schwartz 學者提出以跳躍計數(hop counts)和來回延遲(round trip)的方法[4]來增加檔案傳輸效率。Carter 和 Crovella 提出以伺服器和客戶端的可利用頻寬和擁塞路徑來選擇對映的檔案伺服器[1]。Fu 和 Venkatasubramanian 利用 CPU 週期、I/O 頻寬和 Memory 的使用量來選擇檔案伺服器[3]。以上這些方法都是用來選擇多個檔案伺服器或將檔案傳輸適當的分配到各個檔案伺服器中, 來進行同時間的下載, 因此可將檔案下載工作分散到各個檔案伺服器上做負載平衡, 來加速檔案傳輸。

對於高容量檔案的傳輸, 在 2003 年 Sohail、Jha 和 Elgindy 提出了一個嶄新的方

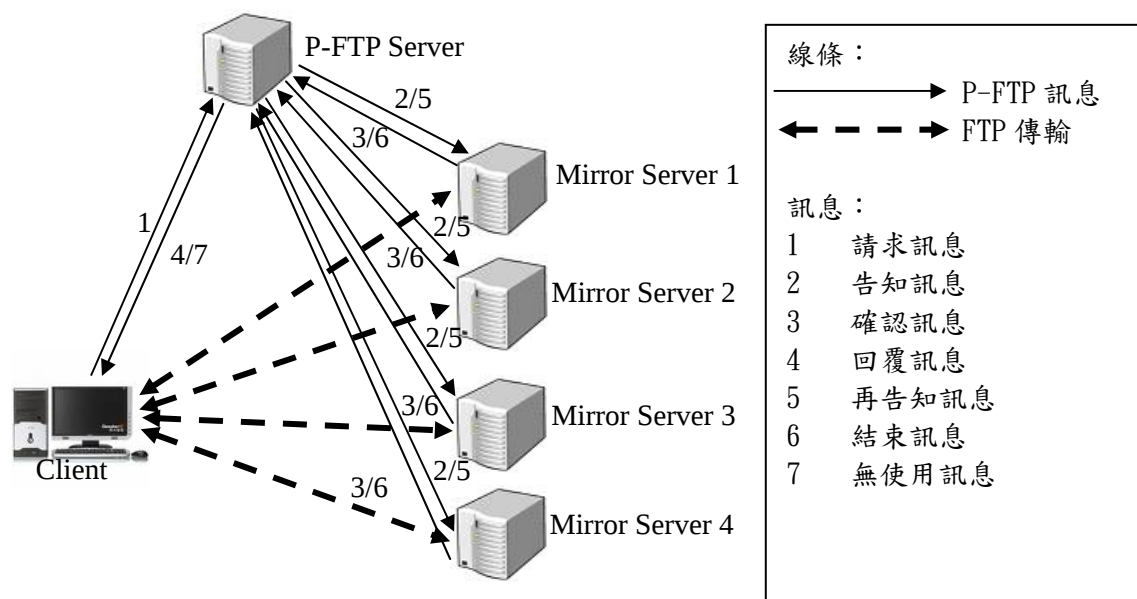


圖 1、P-FTP 實體訊息傳遞架構圖[6]

法，我們稱為平行放置檔案傳輸協定 (Parallelized File Transfer Protocol，簡稱 P-FTP[6,7])。所謂的 P-FTP 是透過設立一個自主系統 (Autonomous System，簡稱 AS)[2] 區域中心點的 P-FTP 伺服器，來進行會話式訊息傳遞，經由訊息傳遞來達到客戶端對多個檔案伺服器進行平行下載。

平行放置檔案傳輸協定 (P-FTP) 在運作時，可分為三個實體來說明，分別為請求的客戶端、P-FTP 伺服器和對映伺服器 (Mirror Server) 來進行互動流程。P-FTP 在檔案傳輸的流程如圖 1 所示，並說明如後：

- 1) 當一個客戶端發出請求時，會將請求訊息發送到 P-FTP 伺服器，
- 2) P-FTP 伺服器接收到請求後會開始收集對映伺服器的相關資訊，這些相關資訊包含 P-FTP 伺服器與對映伺服器間的延遲與對映伺服器本身的使用率，並且以這些參數去執行適當演算法 (Suitability algorithm)[6] 計算，評估這些對映伺服器的適當性和每個對映伺服器應分配多少比例的檔案大小提供下載，當 P-FTP 伺服器計算完成後會送出告知訊息給所有的對映伺服器，這些告知訊息包含對映伺服器本身的適當性、請求檔案和請求的客戶端位置，
- 3) 接收到告知訊息的對映伺服器會保存此資訊並回送確認訊息給 P-FTP 伺服器，
- 4) 在 P-FTP 伺服器接收到確認訊息後，以回覆訊息告知請求客戶端，此回覆訊息

包含了對映伺服器各別的適當性和位置等資訊，

- 5) 當 P-FTP 伺服器偵測到網路環境改變的資訊後，會對所有運作中的對映伺服器重新評估適當性並且以再告知訊息告知對映伺服器，對映伺服器會根據此資訊來調整檔案大小的比例，並且回送調整後的確認訊息給 P-FTP 伺服器，
- 6) 最後客戶端就以一般的 FTP 來對所有對映伺服器進行檔案傳輸，在完成傳輸之後，會由各別的對映伺服器來送出結束訊息告知 P-FTP 整體流程結束。
- 7) 如果偵測到所有的對映伺服器都故障或是無法使用時，會由 P-FTP 伺服器向客戶端發出無使用訊息，並且中止檔案傳輸。

在整個傳輸過程中，都是由 P-FTP 伺服器去對所有的對映伺服器進行評估，因此，實際上會產生對 P-FTP 伺服器所評估出來最好的對映伺服器，不見得對客戶端是最好的，因為客戶端可能與自主系統的中心點距離很遠，再加上所有的計算都集中在 P-FTP 伺服器上，當使用者的數量變多時會對中央 P-FTP 伺服器產生很大的負擔。

經由上述，本論文將以無中央控制的伺服器且結合 SDT-L 的基礎，讓客戶端以最近的檔案伺服器為計算伺服器，且計算伺服器可自主尋找附近能力較佳的檔案伺服器來支援此次的檔案下載。本論文中將提出一個具有自我運算能力的檔案傳輸協定，以分別套用在各別

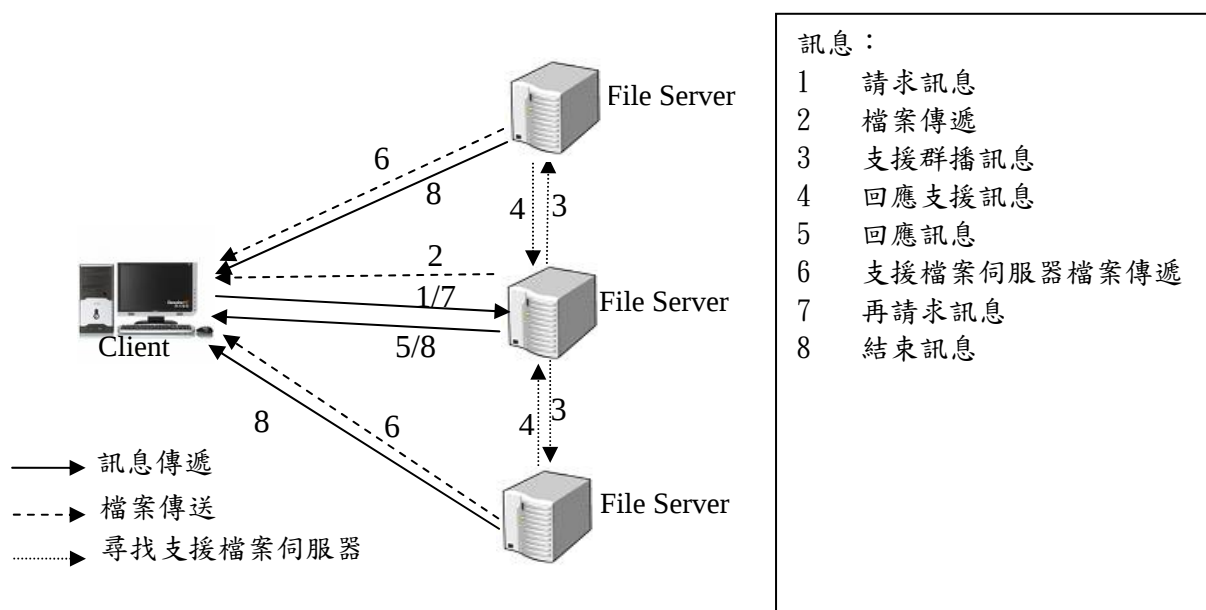


圖 2、DP-FTP 架構流程圖

的檔案伺服器，以解決固定數量的檔案伺服器情況下加速檔案的下載，並且尋找對客戶端最佳的檔案伺服器。

2. 分散式平行放置檔案傳輸協定

本論文所提出的分散式平行放置檔案傳輸協定(Distributed Parallelized File Transfer Protocol, 簡稱 DP-FTP), 是以傳統的 SDT-L 和 P-FTP 兩種技術為基礎, 使客戶端尋找最近的檔案伺服器為計算伺服器, 並且以計算伺服器自主的找尋支援檔案伺服器, 來進行網路上分散且平行的檔案傳輸, 並且針對 P-FTP 上的問題加以改進。

2.1 DP-FTP 之架構

DP-FTP 的架構中並沒有特定中央伺服器, 藉由網路各個檔案伺服器之間相互的溝通, 進行檔案傳輸時的適當分配, 以距離客戶端最近的檔案伺服器為中心去尋找對客戶端最好的支援檔案伺服器, 來建構 DP-FTP 的分散且平行的檔案傳輸處理方式。

DP-FTP 在傳輸檔案的時候, 主要分成兩個主要的部份, 分別為請求的客戶端和檔案伺服器。在請求的客戶端要求一個檔案時, 利用 SDT-L 的觀念, 會以距離客戶端最近的檔案伺服器先接收到客戶端所群播出去的請求訊息, 而由這個伺服器開始提供檔案傳輸的服務。DP-FTP 的流程如圖 2 所示, 並說明如後:

- 1) 首先當客戶端請求檔案下載時, 會以群播的方式發送一個請求訊息給最先回應的檔案伺服器,
- 2) 最先回應的檔案伺服器在接收到請求訊息後建立連線, 以及開始透過 FTP 進行檔案傳送給客戶端,
- 3) 並且以它發送一個支援群播訊息, 尋找支援的檔案伺服器, 此群播訊息包含客戶端所請求的檔案資訊和位置,
- 4) 附近的檔案伺服器在接收到支援群播訊息後, 進行回應支援訊息, 此訊息包含自身的位置和資訊, 回應給群播的檔案伺服器,
- 5) 當支援的檔案伺服器回應後, 立刻詢問客戶端目前剩下多少的檔案部份還未傳送, 再利用回應的資訊(連同本身檔案伺服器)去執行適當演算法[6], 計算各個檔案伺服器的適當性, 並且以適當性去分配各個檔案伺服器所負責傳送的檔案大小比例, 之後發送回應訊息告知客戶端, 在此回應訊息中包含支援的檔案伺服器位置、負責的檔案大小百分比比例,
- 6) 客戶端藉由此訊息進一步向各支援檔案伺服器進行平行的 FTP 檔案下載, 因為客戶端從回應訊息中知道各個檔案伺服器下載的部份, 因此可利用此訊息重組檔案,
- 7) 當網路環境改變時, 由於客戶端會週期

性的向計算伺服器發出再請求訊息，再請求訊息包含了目前未完成傳輸的檔案部份，提供計算伺服器對剩下未完成的檔案部份重新計算適當性，使所有的檔案伺服器可以動態調整傳輸檔案比例，

- 8) 當檔案傳輸結束時，由各檔案伺服器發出結束訊息，告知客戶端它所負責的檔案部份已經傳送完成，客戶端如果接收到所有檔案伺服器的結束訊息，就可以開始重組檔案。

在這整個傳輸過程中，都是由網路上各檔案伺服器之間相互支援，來完成檔案傳輸，並不需要透過任何中央伺服器，而所有支援的檔案伺服器，都是由距離客戶端最近的檔案伺服器所尋找到，相對的距離也比較近，並且每個客戶端與檔案伺服器的距離不一樣，因此所選擇的計算伺服器也不相同，所以當使用者的數量增加時，計算伺服器也不會集中落在同一台檔案伺服器上造成負擔，因此 DP-FTP 可達到負載平衡。

2.2 DP-FTP 之評估

整體而言，DP-FTP 比 P-FTP 少了建置中央控制的伺服器並且實現了將計算工作分散在網路上各個檔案伺服器之中，來完成分散且平行的檔案傳輸處理，避免中央控制的伺服器故障後整個網路無法進行傳輸。由於是以客戶端最近的檔案伺服器去尋找的支援檔案伺服器，所以對客戶端來說相對的距離比較近，傳輸擁塞比較小，傳輸時間也比較短。

3. 實驗

實驗中我們採取 DP-FTP 和 P-FTP 相互比較的方式，實驗是在相同大小的自主系統 (Autonomous System, 簡稱 AS) 中，每次由一個客戶端發出 1Giga Bytes 的資料請求。實驗一開始是以 10 個檔案伺服器和 1 個客戶端來執行 100 次結果，將每一次結果所產生的下載時間加總取得平均值，每次執行時必須重新產生 10 個檔案伺服器和 1 個客戶端的位置，而 P-FTP 的中央伺服器則固定在中心點，其後再每次增加 10 個檔案伺服器執行 100 次，直到檔案伺服器的數量到達 50 個。實驗的下載時間，是以計算伺服器或 P-FTP 的中央伺服器對所找尋到的檔案伺服進行適當性演算法後，進行適當的分配傳輸大小，再以檔案伺服器與客

戶端實際的延遲進行檔案傳輸所獲得的平均時間。實驗結果如圖 3 所示。

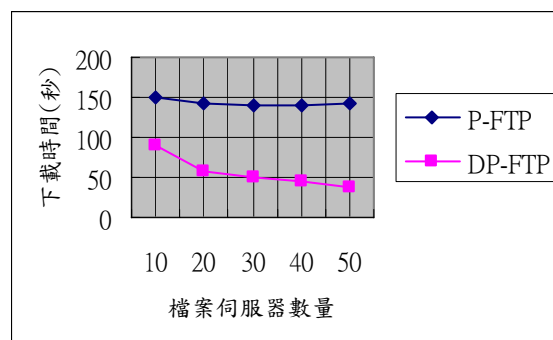


圖 3、固定檔案伺服器數量之下載時間

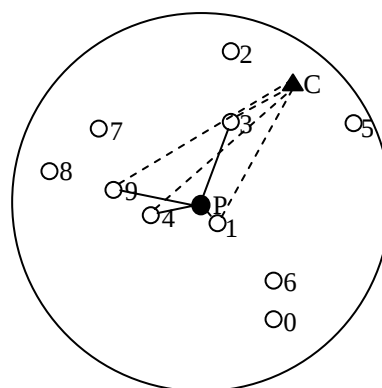


圖 4、模擬 P-FTP 實際傳輸示意圖

表 1、P-FTP 延遲和下載時間

No.	Delay 1	Percent	Delay 2	Time
1	0.036	0.60	0.439	273.077
4	0.111	0.20	0.509	102.149
9	0.202	0.11	0.544	60.323
3	0.247	0.09	0.21	19.12

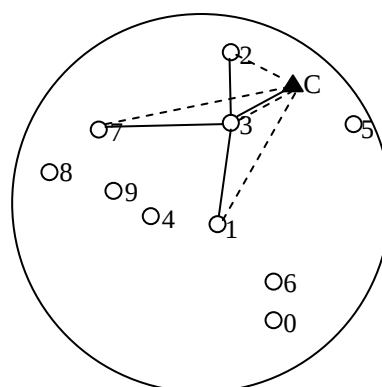


圖 5、模擬 DP-FTP 實際傳輸示意圖

表 2、DP-FTP 延遲和下載時間

No.	Delay 1	Percent	Delay 2	Time
3	0.021	0.30	0.21	64.242
2	0.34	0.25	0.228	58.597
1	0.25	0.23	0.439	103.346
7	0.27	0.22	0.483	109.894

從圖 3 可以觀察到，當檔案伺服器增加時，對 DP-FTP 來說，會逐漸減少傳輸時間，而 P-FTP 並不會有太大的影響，因為 P-FTP 的中央伺服器是固定的，影響 P-FTP 最大的因素是客戶端發出請求時的位置，距離中央 P-FTP 伺服器越近，傳輸速度越快，而 DP-FTP 則是可由檔案伺服器的密度越高，傳輸速度越快。所以整體來說，DP-FTP 執行下載的速度比 P-FTP 來的快速。

而圖 4 和圖 5 為檔案伺服器數量為 10 時，其中一次的實驗所產生的伺服器分佈圖，其中圖 4 表示 P-FTP 的實驗，而圖 5 表示 DP-FTP 的實驗。檔案傳輸相關數據則分別如表 1 和表 2 所示。示意圖中○表示檔案伺服器、●表示 P-FTP 伺服器和▲表示客戶端。表格中 No.代表此次所分別挑選出來的檔案伺服器編號、P-FTP 的 Delay 1 是 P-FTP 伺服器到各檔案伺服器的延遲；DP-FTP 的 Delay 1 是各檔案伺服器到計算伺服器的延遲(這裡的計算伺服器是編號 3 的檔案伺服器)、Percent 是分別計算出來的適當分配率、Delay 2 是分別挑出來的檔案伺服器與客戶端的延遲和 Time 是各別的下載時間。由這兩組圖和表格可以看出來 P-FTP 所挑出來的檔案伺服器與客戶端的相對位置存在著相當大的差距，因此在相同大小的範圍下，DP-FTP 會比 P-FTP 的傳輸時間短。

4. DP-FTP 與 P-FTP 的比較

在這個部份，我們將分別針對執行時的負載平衡和故障回覆兩方面，對 DP-FTP 和 P-FTP 兩個做比較，

4.1 分散式負載計算

在 P-FTP 中主要是以一個中央控制的 P-FTP 伺服器來處理所有對映伺服器的適當性計算，再以平行處理的方式來提供檔案下載服務。雖然將同一個檔案分散在各個對映伺服器進行平行下載，不過不能完全算是分散式處理，因為所有對映伺服器需要分別提供多少百分比的檔案下載，全部都要透過 P-FTP 伺服器計算才會知道，而所有演算法的計算負載，全部集中於同一個自主系統區域的 P-FTP 伺服器上，因此在使用者大量增加時，中央的 P-FTP 可能會面對超過負載情況。

而 DP-FTP 是沒有特定計算的中央伺服

器，將計算分散到網路各個檔案伺服器上處理，每個請求所產生的適當演算法，只須要在客戶端所對映到最近的檔案伺服器上運行，並不會對檔案傳輸造成額外的負載，也不需要建立一個只有計算能力的中央伺服器，因此可達到負載平衡。此外，由於 DP-FTP 沒有特定計算的中央伺服器，所以不需要建立自主系統區域來服務特定範圍的使用者，並且範圍的大小不會影響到 DP-FTP 的演算法能力，因此可達到分散且平行處理的下載服務。

4.2 伺服器故障回覆

在計算伺服器故障回覆方面，P-FTP 的中央控制伺服器如果產生故障的情況時，客戶端必須去尋找另一個新的 P-FTP 伺服器，這個伺服器可能是合併另一個 AS 區域，或是在這個 AS 區域中重新建立一個 P-FTP 的伺服器，在重新建立 P-FTP 伺服器會產生一些問題，例如原本的 P-FTP 伺服器無法正常服務，所以備用伺服器接手運作，但是新的 P-FTP 伺服器必須重新對整個網路環境進行即時的重新調整，而所有的檔案傳輸都是透過 P-FTP 伺服器，因此，當更換為備用伺服器時，所有的傳輸連線必須重新建立，並且所有客戶端的資料也要重新下載，而同一時間中央伺服器必須對所有請求重新計算，會造成伺服器擁塞或是負擔太大的情況。

在 DP-FTP 方面，由於計算伺服器不是集中在同一個地方，所以當某一個計算伺服器發生故障時，只要客戶端重新對下一個距離最近的檔案伺服器發出請求，重新建立連線，並不會影響太多的客戶端，以及整個網路的檔案傳輸。由於 DP-FTP 不像 P-FTP 以中央伺服器為中心去找尋對映伺服器，所以每個重新請求的客戶端不見得會選擇同一個檔案伺服器做為計算伺服器，因此不會讓重新建立的計算伺服器，在建立時產生太大的負擔。

5. 結論

我們提出一個分散式平行放置檔案傳輸協定 (Distributed Parallelized File Transfer Protocol, 簡稱 DP-FTP)。我們將 AS 網路上的每個檔案伺服器皆視為一個 P-FTP 伺服器，並利用 P-FTP 平行處理的優點，再加上分散式計算的特性，讓每個客戶端的服務不會集中在同一個伺服器上，以完成分散且平行的檔案傳輸處理。DP-FTP 不但將檔案進行平行下載，還

利用每個伺服器的傳輸能力，將傳輸負擔適當地分散到各個支援檔案伺服器上，使得檔案傳輸更加快速完成。

本方法是基於 P-FTP 為雛型，加以改進缺點，使得 DP-FTP 可以更加的快速以及穩定的傳輸。未來我們將針對 DP-FTP 做更進一步的改進。我們會考慮有關服務品質(Quality of Service, 簡稱 QoS)的參數，以及如何在傳輸過程中，不中止檔案傳輸，就可以更換計算伺服器，達到計算伺服器故障時，更換計算伺服器後，不需要重新傳輸檔案的方式。

參考文獻

- [1] Carter, R. and Crovella, M., "Dynamic Server Selection Using Bandwidth Probing in Wide Area Networks," **Proc. INFOCOM. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies**, 1997.
- [2] Doyle, J., *Routing TCP/IP Volume I (CCIE Professional Development)*. Cisco Press, 1998.
- [3] Fu, Z. and Venkatasubramanian, N., "Directory Based Composite Routing and Scheduling Policies for Dynamic Multimedia Environment," **Proc. IEEE International Parallel and Distributed Processing Symposium 2001**, vol. 9, no. 5, pp. 398-410, 2004
- [4] Guyton, J. and Schwartz, M., "Locating Nearby Copies of Replicated Internet Server," **Proc. ACM Special Interest Group on Data Communications**, pp. 288-298, 1955.
- [5] Linington, P.F., "File Transfer Protocols," **IEEE Journal on Selected Areas in Communications**, vol. 7, no. 7, pp. 1052-1059, 1989.
- [6] Sohail S., Jha S. and Elgindy H., "Parallelized File Transfer Protocol (P-FTP)," **Proc. IEEE Local Computer Network Workshop High-Speed Local Network**, pp. 624-631, 2003
- [7] Sohail S., Jha S., Kanhere S.S. and Chun T.C., "QoS Driven Parallelization of Resources to Reduce File Download Delay, " **IEEE Trans. on Parallel and Distributed Systems**, vol. 17, no. 10, pp. 1204-1215, 2006