

# 運用鏈路負載平衡機制 提昇單位網路傳輸效能之研究

作者/吳世達 上尉

## 提要

針對多條聯外路徑環境(Multihoming)於單位網路環境的運用，並將鏈路負載平衡演算法對於不同的網路運用作分析及比較。Multihoming 是一種提升單位網路運作可靠性的機制，同時能達到節省成本的目的；目前有關 Multihoming 網路環境的研究與實務運用包含有 BGP 與 RON 等，由於 BGP 和 RON 需要藉由所連接的 ISP 額外提供大量路由訊息的交換；而使用線路負載平衡 (Link Load Balance) 則是一種不需要 ISP 額外提供路由資訊的運作機制，它採用類似「伺服器叢集負載平衡」的概念來分配每條線路的負載流量。亦將此想法運用至「國軍網路軍團級」以上之單位 (含防衛部單位)，提供不間斷的服務並具有容錯的功能，亦能提昇國軍單位網路之服務效能。

本研究以二個網路上常見的應用實例：「壅塞反應」及「斷線反應」進行探討；壅塞反應為避開其中某些鏈路相對較為壅塞的鏈路，將流量導向相對較為暢通的鏈路，增加使用者資料傳輸的滿意度。斷線反應為偵測網路健康狀況，適時避開中斷鏈路，提高資料傳輸的可靠度。

## 壹、前言

國軍各單位對於網路的依賴性日趨增加，亦是國軍推動資訊化的腳步愈來愈快之結果所致，如不能提供一個穩定且持續性的網路服務，相信對國軍各單位勢必造成非常大之困擾；基於此想法，本研究企圖以鏈路負載平衡的觀點去建構一穩定性、可靠性之網路服務。

## 貳、緒論

### 一、研究動機

現今，單位為了能夠支援單位網路運作所急劇增加的網路流量，提升單位運作效能並減少單位網路聯外鏈路的連線成本，都會降低對網際網路連接專線的使用，以減低專線線路的租用成本，並配置多條低成本的聯外鏈路(Multipaths)取而代之，亦可獲得較大的連線頻寬。此外，單位網路透過多條聯外鏈路可能透過相同或不相同的 ISP 業者，避免單一對外連結中斷造成連線失敗，提高了

單位網路聯外鏈路運作的可靠性與穩定性<sup>1</sup>。

當單位網路透過多條聯外鏈路連接網際網路，單位如何有效利用其所配置的多條聯外鏈路的網路環境，避免發生部份聯外鏈路閒置，或是造成部份聯外鏈路過度壅塞，形成聯外鏈路負載不均的現象。為解決上述單位網路在多條聯外鏈路環境可能遭遇聯外鏈路負載不均衡的現象，單位可透過鏈路負載平衡機制之運作，整合單位所有聯外鏈路的頻寬，期使單位聯外鏈路能夠發揮最大的連線運作效能。

## 二、研究目的

本研究將針對鏈路負載平衡運作機制進行瞭解，探討構成鏈路負載平衡機制的運作項目，然後以仿真模擬（emulation）的方式，瞭解不同的鏈路負載平衡演算法在單位網路面臨不同的網路連線狀況時所產生的運作效能，以供單位網路使用鏈路負載平衡機制之參考。

本研究以「壅塞反應」和「斷線反應」做為探究鏈路負載平衡機制對於單位網路聯外鏈路流量傳輸效能的影響。壅塞反應為避開其中某些鏈路相對較為壅塞的鏈路，將流量導向相對較為暢通的鏈路，增加使用者資料傳輸的滿意度。斷線反應為偵測網路健康狀況，適時避開中斷鏈路，提高資料傳輸的可靠度。

## 三、研究方法與架構

為不影響單位組織正常網路運作之前提下，本研究嘗試以仿真模擬的方式進行單位多條鏈路負載平衡機制網路環境之建構，藉由不同的設定鏈路負載平衡機制仿真模擬想定之執行，分析不同的負載平衡演算法對不同設定項目所造成的影響，並將所仿真模擬之數據加以整理分析，並將分析結果做為進一步探究單位網路應用鏈路負載平衡機制之參考。

## 四、研究範圍與限制

通常鏈路負載平衡依連線方向可以區分為二種狀況，一種是單位網路內部對外（outbound）的連線請求，另一種則是由外部鏈路經由單位路由對內（inbound）的連線請求。對外的連線請求會使用 NAT 負載平衡技術的方式來定址，使連線請求和回覆的封包會在同一條線路上。對內的連線請求是由外部 DNS 做解析<sup>2</sup>，決定連線請求該走那一條鏈路，負載平衡伺服器會監看所有廣域網路的連線狀況，若某條網路連線中斷，便自動調整 DNS 紀錄，將流量導引至正常的線路上。由於這需要 DNS 階層式架構的支援，而控制權也不全然在負載平衡

---

<sup>1</sup> Goldenberg, D.K., Qiuy, L., Xie, H., Yang, Y.R. and Zhang, Y., "Optimizing Cost and Performance for Multihoming", Proceedings of the 2004 ACM SIGCOMM Conference, USA, New York, pp.79-92, 2004.

<sup>2</sup> Brisco, T. [1995], "DNS Support for Load Balancing", RFC 1794.

器本身，因此，本研究將以單位網路內部對外的連線請求為研究範疇。

## 參、文獻探討

### 一、負載平衡技術<sup>3</sup>

通常以網際網路為運作環境的鏈路負載平衡機制，其常加以運用的 IP 負載平衡技術包括有：NAT、IP Tunneling 封裝技術與 Direct routing 等三種。

#### (一) NAT

網路位址轉換技術主要應用於單位私有網路環境之中，網路位址轉換技術係在私有 IP 網址與合法 IP 網址進行轉換工作，藉由網路位址轉換技術之運作將可使單位私有網路配置私有 IP 網址的電腦連線至網際網路；網路位址轉換技術亦具備疏緩 IPv4 網路位址不足等問題。

鏈路負載平衡機制應用網路位址轉換技術，將可將一個外部 IP 網址導引至多個內部伺服器 IP 位址，當使用者的連線請求時，動態轉換為一個內部伺服器的 IP 位址，其封包的來源位址與目的地址都會被更改，藉此以單一 IP 的方式而達到負載平衡的目的。

#### (二) IP Tunneling 封裝技術

IP Tunneling 負載平衡技術將 IP 資料包(datagram) 封裝另一個 IP 資料包，使資料包中指定一個 IP 網址被封裝並導引至另一個 IP 網址；當用戶端提出連線請求時，將請求封包經封裝後導引至不同的伺服器，並將處理結果直接回應給用戶端。由於這種作法須要每台伺服器能支援 IP Tunneling 協定，且建置成本相對的高。鏈路負載平衡機制可以應用 IP Tunneling 封裝技將網路應用封包導引至特定網路伺服器或特定網址，以達到聯外鏈路負載平衡的效果。

#### (三) Direct routing 技術

Direct routing 的方法是利用終端伺服器與負載平衡器共用同一個虛擬 IP 網址，鏈路負載平衡器改變終端伺服器資料訊框中的 MAC address，使得請求連線由伺服器接收並直接轉送回應到使用者。IBM 的 Net Dispatcher 的運作方式與此方法非常相似。表一 說明這三種技術運作的優缺點並進行相關之比較。

表一 IP 負載平衡三種方法優缺點之比較

| 負載平衡方法 | 優點                 | 缺點           |
|--------|--------------------|--------------|
| NAT    | 1. 支援 TCP/IP 協定的任意 | 1. 性能擴展能力有限。 |

<sup>3</sup> Jin, S., Wu, Q. and Zhang, W., "Scaling Internet services by LinuxDirector", The Fourth International Conference/Exhibition on High Performance Computing in the Asia-Pacific Region, vol.1, China, Beijing, 2000, pp.176-183.

|                |                                 |                                                                |
|----------------|---------------------------------|----------------------------------------------------------------|
|                | 作業系統。<br>2.使用私有位址。.             | 2.伺服器的節點數量達到 20 或更多時，負載平衡伺服器可能成為整個系統的瓶頸。                       |
| IP Tunneling   | 負載平衡伺服器和實際伺服器可以在不同的區域網路，甚至廣域網路。 | 1. 所有伺服器必須使用 IP Tunneling 協定。<br>2. 直接路由的效率比較低，而且並不是所有OS都支援。   |
| Direct routing | 1.效能最高。<br>2.網路延遲（Delay）最小。     | 1.負載平衡伺服器和所有應用伺服器必須在同一個 Physical segment。<br>2.終端伺服器不能做ARP的回應。 |

(資料來源：Jin, S., Wu, Q. and Zhang, W. “Scaling Internet services by Linux Director”, The Fourth International Conference/Exhibition on High Performance Computing in the Asia-Pacific Region, vol.1,China, Beijing,2000, pp.176-183.)

此外，還有一個常見的鏈路負載平衡技術，係利用網域名稱服務機制的運作，此一技術將利用 DNS 網域名稱伺服器的運作<sup>4</sup>，將網際網路中特定一個網路領域名稱配置多個 IP 網址，同時採用簡單的 Round Robin 演算法，將來自於網路用戶端的連線請求，依序導引至網域名稱伺服器內所對應不同的 IP 網址，如此，將使得不同用戶端的連線請求連結至不同的伺服器，以避免連限制特定伺服器的鏈路發生壅塞線，藉此達到負載平衡的目的，提昇網路整體效能。

## 二、Multihoming簡介<sup>5,6</sup>

Multihoming 主要的運作網路環境在於一個電腦主機同時擁有兩個或是兩個以上的網路連結介面；Multihoming 網路連結可以區分成「一個網路介面、多個網路位址」以及「多個網路介面」兩種。本研究主要針對應用於「多個網路介面」之 Multihoming 網路環境之探討。

### (一) Multihoming 之特性

#### 1. 冗餘備援

當電腦主機主要的網路連線鏈路突然中斷時，該電腦主機可以透過 Multihoming 機制之運作，從其他網路介面建立連線，以避免各種因網路斷線而造成單位營運上的損失。

#### 2. 負載平衡

在網路連線建立前，網路節點可以先行偵測網路封包流量狀態，依據相

<sup>4</sup> 同註 2。

<sup>5</sup> 楊憲其，「在 IPv6 環境下 Multihoming 設計之研究」，國立中正大學資訊工程研究所碩士論文，民國 94 年。

<sup>6</sup> 蘇柏州、朱振練，「IPv6 Multihoming 與家庭網路技術探討」，電腦與通訊，第 114 期，民國 94 年，頁 55-59。

關設定，自動選擇符合頻寬要求的網路介面建立連線，避免造成單一網路鏈路傳輸的沈重負擔。

### 3.路由政策

網路應用的多樣化，其對於網路頻寬的需求意有所不同；因此，許多不同的網路服務，可依其運作原理與所需服務品質需求之不同，而有不同的網路頻寬配置。網路管理者可以依據相關網路應用的特性，針對單位網路運作之需求、不同的線路性質、網路使用者的特性之不同與相關因素，訂定不同的網路路由策略，使得單位網路應用路由導引能夠符合單位網路運作之需求。

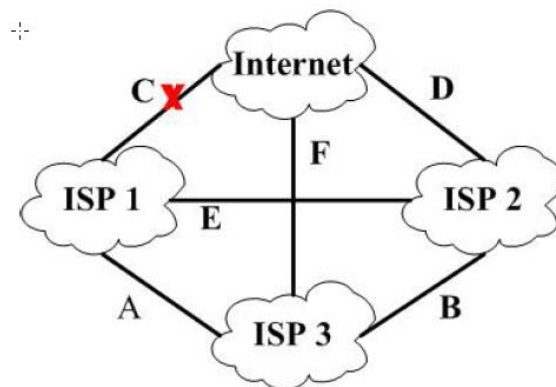
### 4.頻寬聚合

頻寬聚合係指可將多條頻寬較小的鏈路聚合在一起，使得這些多條鏈路如同一條具有較大頻寬的鏈路一般，供各類網路應用進行流量之傳送。一般而言，ISP 對於頻寬較大的鏈路租用，其租用成本相對高許多，單位往往可藉由租用多條頻寬較小的鏈路滿足其所需之頻寬需求，且其所需之租用費用相對少許多。由於，多條頻寬較小的鏈路可以聚合，使其運作效能相當於同樣頻寬大小的單一鏈路，因此，單位為了符合經濟效益，並且可使單位網路聯外鏈路可以同時擁有多重聯外鏈路的高可靠性，頻寬聚合概念對於單位網路是一種很好的應用。

## (二) Multihoming 分類

### 1.ISP level

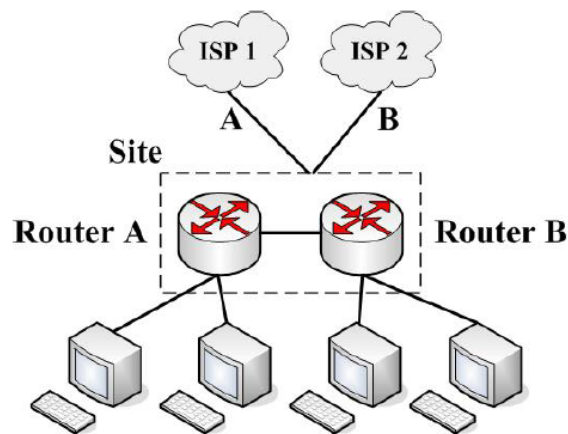
此一層級主要係針對網際網路上 ISP 與 ISP 間的 Multihoming 進行相關的探討，ISP-level 的 Multihoming 架構圖，如圖一所示：透過 Multihoming 機制之運作，ISP1 可以把自己所屬的位址網域經由 ISP2 或 ISP3 的路由，傳輸網路應用封包，因此，當 ISP1 主要對外線路 C 發生中斷時，其所屬網路應用封包可以經由鏈路路由設定的運作，繞送到其它的 ISP 連到 Internet。



圖一 ISP level 示意圖(資料來源：楊憲其，「在 IPv6 環境下 Multihoming 設計之研究」，國立中正大學資訊工程研究所碩士論文，民國 91 年。)

## 2.Site level

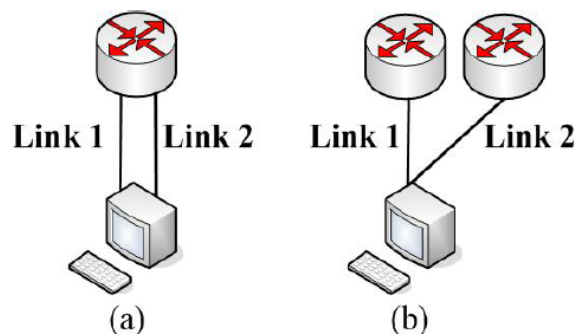
Site level 的 Multihoming 技術在 IETF RFC3178 中有深入的探討，其主要是指當一 Site 同時有兩個以上的路由器，連接至兩個以上不同的 ISPs；如圖二所示：當網際網路某一特定網路位置分別由路由器 A 與路由器 B 所建構而成，分別協調該特定網路位置內部電腦的連線運作情形；在正常狀況之下，網路應用封包流量將可透過路由器 A 經由線路 A 連往 ISP1，並透過路由器 B 經由線路 B 連往 ISP2。當線路 A 中斷時，使用者可以透過路由器 B 經由線路 B 連接至網路上。



圖二 Site level 示意圖(資料來源：楊憲其，「在 IPv6 環境下 Multihoming 設計之研究」，國立中正大學資訊工程研究所碩士論文，民國 91 年。)

## 3.Subnet/Host level

Subnet/Host level 的 Multihoming 技術係指一部電腦主機上有兩個或兩個以上的網路連結分別連往相同或是不同的路由器，如圖 3 所示。圖 3(a)為兩個連結接上相同的路由器，而圖 3(b)分別接往不同的路由器；當其中此一電腦主機連線至路由器的鏈路斷線時，透過 Multihoming 機制的運作，該部電腦網路應用封包將可藉由另一鏈路進行傳送。本研究所探討之重點將以圖 3(a)之作法做為研究之想定基礎。



圖三 Subnet/Host level 示意圖(資料來源：楊憲其，「在 IPv6 環境下 Multihoming 設計之研究」，國立中正大學資訊工程研究所碩士論文，民國 91 年。)

### (三) Multihoming 網路環境相關封包路由技術之探究

目前可運用於 Multihoming 網路環境中封包路由技術可以概略區分為三大類型：BGP、RON 及 Load Balance 等三種技術，這三種技術分別應用於不同的網路環境，具有不同特點，將於下列小節中加以說明與探討。

#### 1. BGP<sup>7,8</sup>

BGP 是標準的外部閘道器路徑協定，其主要功能是在多個網域自治系統之間進行路由資訊的交換，透過 BGP 路由協定之運作，可以強化封包路由政策之執行與避免封包傳送所可能產生路由迴路的問題，同時 BGP 路由協定可以提供負載平衡功能。

BGP 路由協定運作涉及許多的相關問題，尚需進一步探討，如不同定址方式帶來不同的路由成本及 BGP 路由協定在選擇路徑上不夠即時也不能反映實際流量的狀態的問題等。在 Multihoming 的環境下，BGP 路由協定要交換整個路由表，如果用戶端都與其所接連的多家 ISP 業者進行 BGP 路由訊息的交換，那麼這些 ISP 業者進行路由訊息交換的複雜度將會變成  $O(N^2)$ ，相對的，路由器運作成本亦隨之增加；因此，在實務上 BGP 路由協定只會運用於較大型的企業、機構或是 ISP 業者。

#### 2. RON<sup>9</sup>

RON 的方式主要是針對 BGP 無法即時反應網路狀況的變動，而且也無法及時準確地找到最佳路徑的問題提出改善方案。RON 架構係以一種點對點的量測方式做為尋找最佳路徑的方式，其量測的方式是每隔一段時間（約十幾秒鐘）在兩個節點相互交換 UDP 封包，RON 藉由這種量測提供三種路徑選擇評量值：Latency、Packet loss rate 與 Available throughput。每個網路節點將會對其所連接的網路節點進行量測，並將所獲得的三個評量值，存入一個屬於這個網路節點的傳輸效能資料庫之中，而不同網路節點會透過線路狀態路由交換協定來交換彼此的傳輸效能資料庫的資料，以便知曉其他網路節點路由狀況，進而進行網路拓樸的路徑選擇計算。

經由網路節點路由訊息的交換，RON 可以選擇以 Latency、Loss rate 或 Throughput 找到點對點之間的最佳封包傳送路徑；由於不同節點之間是以非同

---

<sup>7</sup> Ahuja, A., Bose, A., Labovitz, C. and Jahanian, F., "Delayed Internet routing convergence", Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, USA, New York, 2000, pp. 175-187.

<sup>8</sup> Li, W., "Inter-domain Routing: Problems and Solutions", Technical Report TR-128, Department of Computer Science, USA, New York, 2003.

<sup>9</sup> Andersen, D., Balakrishnan, H., Kaashoek, F. and Morris, R. "Resilient Overlay Networks," Proceedings of the eighteenth ACM symposium on Operating systems principles, USA, New York, 2001, pp.131-145.



步的方式持續進行量測，因此對於問題路徑的反應時間比較快，相較於 BGP 路由協定需要數分鐘的時間反應路徑狀態資訊，RON 可以在幾十秒內避開問題路徑。此外，一些研究顯示 RON 的運作可以降低封包傳輸的延遲。

RON 的運作成本往往會隨著 RON 所在網域節點數量的多寡而有指數性成長的變化；假設在 RON 的網域中有  $N$  個節點，那麼每個節點會對其他  $N-1$  個節點進行量測，其彼此量測的複雜度也將是  $O(N^2)$ 。而 RON 的量測成本，在 50 個節點相互量測的時候會有 33kbps 頻寬需求，而當有 500 個節點可能會造成有數個 mbps 的量測頻寬需求，其量測成本呈現指數性的成長；因此，RON 的應用雖然可以改善問題路徑的反應時間，但其運作於大型網域時，其運作成本對於網路是一個很大的負擔，如此一來，限制 RON 架構運作的規模。

### 3. Load Balance

Load-Balance<sup>10</sup>是別於 BGP 和 RON 的運作，它不需要藉由所連接的 ISP 額外提供大量路由訊息的交換，藉以獲得網路各個節點間的最佳路徑封包傳輸路徑。Load-Balance 機制利用網路位址轉換技術的方式來進行定址的工作，它可以利用不同的網路連結單位（如：connection、session 與 packet）來分配線路，同一個網路連結單位會被導引至同一條鏈路，進行封包傳輸。Load-balance 機制採用類似伺服器叢集負載平衡的概念，它會依據每條鏈路當時的流量負載來分配流量，透過不同的演算法則，進行不同的鏈路選擇，傳送網路應用封包。

在 RFC2391<sup>11</sup>中提到伺服器負載平衡機制的運作方式，其中 Load Sharing using IP Network Address Translation (LSNAT) 路由器扮演連線請求分配者的角色，如圖 4 所示，它依據網路或伺服器的負載量及存取成本來分配連線，並透過 NAT 平衡技術將該連線的位址轉換為被指定的伺服器位址。



圖四 LSNAT 運作架構圖(資料來源:Gan, D. and Srisuresh, P., “Load Sharing using IP Network Address Translation(LSNAT)”, RFC2391,1998.)

<sup>10</sup> Chen, J., Chiueh, T., Guo, F. and Li, W., “Experiences in building a multihoming load balancing system”, Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, vol.2, USA, New York, 2004, pp.1241- 125.

<sup>11</sup> Gan, D. and Srisuresh, P., “Load Sharing using IP Network Address Translation (LSNAT)”, RFC2391,1998.



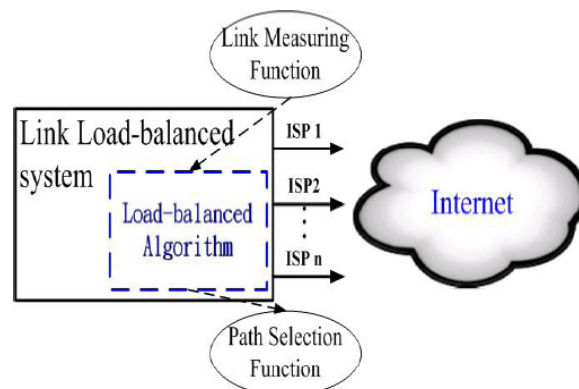
當負載平衡的對象不是伺服器而是網路鏈路的時候，負載平衡機制需要做一些調整；首先對伺服器負載和流量的統計得轉換成對網路鏈路的統計，其次，不再需要對目的地地址的轉換。

相對於伺服器負載平衡機制，鏈線路負載平衡機制主要功能在於平衡單位網路每一條聯外鏈路負載的負荷，避免其中一條聯外鏈路因過多或過少的封包傳送服務需求而產生鏈路封包傳輸超載或是發生鏈路閒置；所以，透過該機制之運作，將可有效提高單位網路的運作效能，增強系統的可靠度及穩定度，加快對使用者的回應時間。

## 肆、系統架構

### 一、運作架構

鏈路負載平衡機制主要運作於具有多條聯外鏈路的單位網路環境之中，圖 5 其運作架構示意圖，亦為本研究進行鏈路負載平衡機制運作探究之單位網路環境運作示意圖。由圖 5 所示，其中  $ISP_1$  至  $ISP_n$  為企業所租賃的 ISP 連線數目，企業可以租賃相同或不相同的 ISP 業者連接網際網路。



圖五 系統架構圖(作者繪製)

由鏈路負載平衡系統架構圖中，可以發現鏈路負載平衡演算法運作將包含聯外鏈路狀況量測及路徑選擇與分配；聯外鏈路量測功能可依鏈路流量的多寡及反應時間的快慢來決定鏈路的好壞；而鏈路選擇與分配的功能可以使用最佳（Best）或是權重（Weighted）的方式來選擇網路應用流量的傳輸鏈路，鏈路分配的單位通常可以為 connection、TCP/IP session 或是 packet，每個網路應用連線會依據鏈路量測的結果，為鏈路分配單位選擇某一條 ISP 鏈路做為其傳輸的線路。

### 二、運作要素分析

#### （一）鏈外鏈路流量量測構面

##### 1. 線路流量量測要素

線路流量量測構面可區分為三個必要元素：量測的時間點、量測的方式和量測的距離。

### (1)量測時間

線路流量探測的頻率越高，越能反映網路的流量變化情形，但相對量測花費的成本亦越多，亦造成更多的網路冗餘。量測時間點的選擇可以為被動或主動。被動的方式係指當每個封包或每個連線(TCP/IP session)到來的時候，鏈路負載平衡機制才進行量測；而主動的方式則是鏈路負載平衡機制每隔一個時段主動對網路聯外鏈路進行量測工作，以瞭解當時網路聯外鏈路的流量狀況。

### (2)量測方式

一般而言，鏈路量測有三種形式：連線數、流量統計及來回時間。

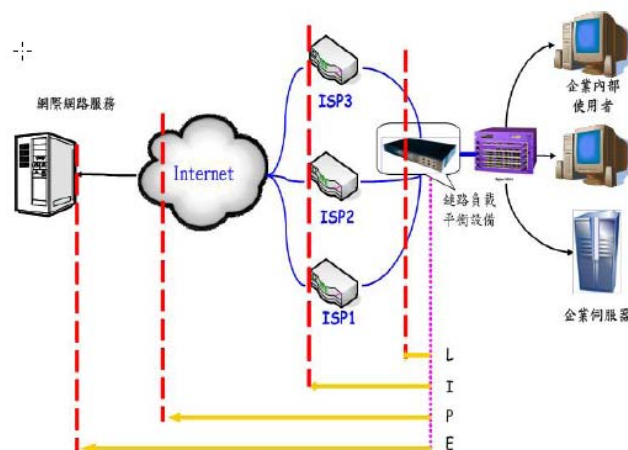
A.連線數是指目前還在進行連線的網路應用數目，連線數可以代表在固定的頻寬下，每個連線平均可以得到的服務。

B.流量統計主要是觀察過去的時間所使用頻寬的多寡，通常可以封包長度的累計來獲得，但是須考慮「已使用」頻寬是有時間性的，舉例來說，以 A、B 兩線路為例，假設 A 線路在五秒前已使用 R 個位元的頻寬，B 是在這一秒用掉了 R/2 個位元的頻寬，那麼 B 的線路可能還比較壅塞。

C.來回時間主要依發出某個封包到有回應的來回時間，如 ICMP 的 echo request 到 echo reply 或 TCP connection 的 SYN sent 到 SYN ACK；一般而言，來回時間和量測的距離有一定的關係。

### (3)量測距離

係指從來源網路位置到目的網路位置間所行經路由中某一特定節點與鏈路負載平衡設備之間的距離；圖六顯示從企業網路鏈路負載平衡設備到目的端之間數個不同的量測距離。



圖六 距離量測圖(作者繪製)

A.量測至 L 點：在鏈路負載平衡設備上觀察流量資訊，如每條線路的連線數。

B.量測至 I 點：量測至 ISP 並觀察 Last mile 的流量，如用 Last mile 的頻寬和使用的頻寬來計算可用的頻寬。

C.量測至 P 點：量測至網際網路特定網址，在路由路徑中，對所有不同的目的端選擇一個會經過的定點來計算回應時間，這個定點可能是 ISP 一個出口路由器。

D.量測至 E 點：量測至接收端，點對點進行回應時間的量測。

## 2.斷線偵測

線路的斷線偵測可和線路流量量測分開或合併進行。線路斷線偵測可藉由反應時間量測的方式進行，利用反應時間的量測結果判斷線路斷線與否；因為得不到反應時間的線路，通常代表有異常情況的發生；但是這和量測距離亦有關聯，如果使用圖 6「量測至 I」的量測距離，那麼這個量測點以外的異常都就無法量測。

此外，使用其他的量測方式，（如：連線數、流量統計）需要額外的斷線偵測機制，因為線路斷線時會造成流量減少的現象，這會使以連線數和流量統計量測方式誤以為線路傳輸狀況很好；額外的線路斷線偵測機制通常會針對幾個穩定的網路節點進行反應時間的量測，這些機制包括：ICMP echo request、UDP traceroute 或是直接使用 TCP 來連線到該網路節點。線路斷線的判定可藉由多個節點共同失敗的連線來判定線路斷線，它將會延長判定線路斷線的時間；亦可利用單一固定節點的連線失敗來判定，但有可能發生誤判的現象。如何取捨線路斷線判斷的速度與判斷的精準度是一個值得考量的議題。

### (二)傳輸線路選擇構面

依連線線路的選擇而言，通常鏈路負載平衡機制會選擇將流量分配至當時量測最好的線路，此一分配方式稱之為最佳模式；但最佳模式可能造成「自我引發壅塞」的狀況，亦就是在同一時段內的連線都會被分配到量測最好的線路，進而造成該線路的壅塞。為了避免這種情形，可以選擇另外一種權重模式的分配方式。權重模式的線路分配方式會依每條線路的量測結果賦予一個權重值，然後依權重比例將當時所有連線需求依公式(1)分配到不同的線路。

$$\text{Min } (S_i / W_i) \text{ for } i = 1..n \quad (1)$$

其中  $S_i$  是每條線路的累計連線數， $W_i$  是那條線路的權重值，權重可以依量測的結果來定其大小。

## 三、負載平衡演算法

本研究將依據 IETF RFC2391 所提的負載平衡演算法和前面述及鏈路負載平衡的相關構面進行相關之分析並提出修訂方法。IETF RFC 2391 的負載平衡方法包括：Round robin (RR)、Least loading first (LLF)、Least traffic first (LTF) 與 Ping to find the most responsive host (PMR)。

### (一) RR 與 LLF

RR 與 LLF 此二種演算法均可直接運用於鏈路負載平衡機制中，RR 是將請求連線依序分配到各個聯外鏈路上，而 LLF 是將連線分到目前擁有連線數最少的鏈路，可以用公式(2)來代表。與 RR 的不同之處在於 LLF 使用還沒有結束的連線數目做為選擇線路分配的依據，而 RR 則是採用輪流分配的方式選擇線路並不會考慮每條線路上面的流量，然而只有還沒有結束傳輸的連線會對流量有影響。

RR 和 LLF 還有另外一項特點，那就是它們不會產生「自我引發壅塞」的情形，因為 RR 和 LLF 不會把相鄰的連線分配到同一個線路。RR 會循序的分配線路，而 LLF 觀察公式(2)可以發現是把公式(1)的  $W_i$  設為 1 的特例，它也會用平均的方式分配。因此 RR 和 LLF 有和權重分配般平均分配線路的特性。

$$\text{Min}(SI_i) \text{ for } i=1..n \quad (2)$$

其中  $SI_i$  係指尚未完成傳輸作業連線的總數。

### (二) LTF

LTF 可以延伸成兩類的線路演算法:Maximum Inbound/Outbound Remaining Bandwidth First (MIRBF/MORBF) 和 Weighted Maximum (Inbound/Outbound) Remaining Bandwidth First (WMIRBF / WMORBF)。這樣的延伸改變主要有兩個考量因素：方向與可用頻寬；方向係因為 LTF 主要根據線路的流量做統計，但 ISP 線路在 inbound 和 outbound 的頻寬可能是非對稱的，而可用頻寬則是需考慮所統計的流量對當下頻寬的影響。

MIRBF 和 MORBF 是使用前面所提到流量統計的方法，在最後一哩線路使用不同的方向來把線路分配到有最大可用頻寬的線路。而 WMORBF 及 WMIRBF 則是考量了前面所提到的權重模式的分配分式，讓 MIRBF 和 MORBF 根據量測的結果與賦予的權重來分配每個連線。

### (三) PMR

PMR 可以延伸成三種演算法: Fastest Round Trip Time to each Destination First (FRTTDF)和 Fastest Round Trip Time to a Fixed Node First (FRTTFNF) 及 Weighted Fastest Round Trip Time to a Fixed Node First (WFRTTFNF)。FRTTDF 和 FRTTFNF 的差異主要在量測的距離。FRTTDF 使用「量測 E 點」即點對點

的量測；FRTTFNF 使用「量測至 P 點」即選擇一個固定的點進行量測。FRTTDF 和 FRTTFNF 都會使用反應時間最好的線路來分配連線。WFRTTFNF 則是依據 FRTTFNF 量測每條線路反應時間的結果賦予權重，然後使用權重模式方式來分配連線。

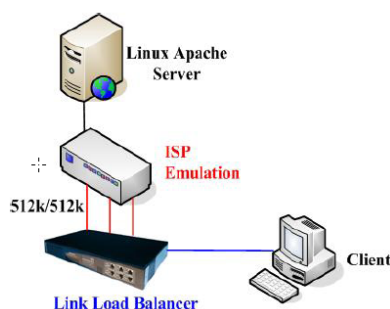
FRTTDF 演算法並沒有賦予一個對應的權重轉換方式，主要是因為它對每個目的地同時進行量測，如果要以權重模式進行線路分配則需得到每條線路的量測結果，其成本相對比較高，因此只它使用最快反應回來的那一條線路進行線路分配的指定。此外，此一類別演算法本身的量測結果即可做為判斷線路是否斷線的依據，可以用於線路斷線的偵測。表二 將上述三種類型的演算法加以整理與比較。

表二 線路負載平衡演算法比較表

| 演算法名稱         | 量測距離 | 分配方式   | 量測方式 | 可否反應斷線 |
|---------------|------|--------|------|--------|
| LTF           | L    | weight | 連線數目 | X      |
| MIRBF/MORBF   | I    | best   | 統量統計 | X      |
| WMIRBF/WMORBF | I    | weight | 統量統計 | X      |
| FRTTDF        | P    | best   | 反應時間 | O      |
| FRTTFNF       | E    | best   | 反應時間 | O      |
| WFRTTFNF      | P    | weight | 反應時間 | O      |
| RR            | L    | weight | 連線數目 | X      |

#### 四、模擬結果與探討

圖七顯示本研究仿真模擬環境，其中 ISP emulator 具有數條 512k/512k 的 ISP 線路，並可以產生近端和遠端的塞車行為和斷線行為。這個仿真模擬環境會由用戶端產生 http 的下載流量到遠端的伺服器端，藉由不同的資料流量（Traffic Load）和不同的網路情況，本研究將從壅塞反應及斷線反應來瞭解不同線路負載平衡演算法的運用及特性。



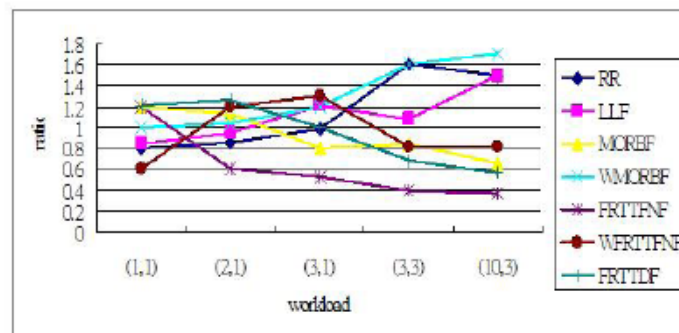
圖七 模擬環境架構圖(作者繪製)



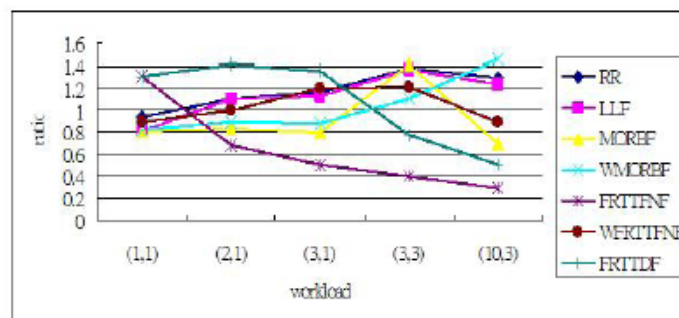
### (一)線路壅塞反應

模擬想定將在一條線路上(如：ISP2)產生網路壅塞的狀況，產生網路壅塞的方式係利用將上傳的頻寬減小，使上傳頻寬只有 5kbits，同時反應時間會從 50ms 變成 2000ms；在使用 HTTP 下載的情況下，上傳頻寬將影響 TCP 通訊協定 acknowledgement 的運作，進而影響下載的速度。網路壅塞的地點會先放在最後一哩線路，讓大部分的演算法都能偵測出流量的變化，接著讓網路壅塞的地點發生在最後一哩線路以外的地方，再來觀察演算法的反應。

此外，對於使用者產生資料流量的模擬將以  $L(u,s)$  的形式表示，但模擬想定資料流量設定將會從  $L(1,1)$  (即同一時段內一個使用者，一個 session 數) 到  $L(10,3)$  (即同一時段內十個使用者，三個 session 數)；研究嘗試收集從最小資料流量到較多資料流量的模擬數據，藉以觀察企業網路不同壅塞狀況下聯外線路被選擇的情形。為了能夠清晰觀察鏈路負載平衡機制中相關演算法對於不同資料流量的資料傳送量，研究將以總平均資料傳送量做為相關演算法模擬數據分析比較之基準；因此圖九是將總平均資料傳送量比例設定為 1，而總平均資料傳送量則是取所有演算法資料傳送量的平均數，透過此一方式的表達比較能夠明顯顯示出不同演算法在不同資料流量狀況下的資料傳送效能，亦於進行相關的分析與比較。



(a) 近端壅塞



(b) 遠端壅塞

圖九 Load Balance 演算法對網路壅塞之反應效能表示圖(作者繪製)



圖九(a)顯示在最後一哩線路中及不同資料流量產生的狀況下，各個演算法的平均資料傳送量；觀察模擬所獲得的數據，可以發現下列現象：

1.當資料流量最小為  $L(1,1)$  時，此時 FRTTDF，FRTTFNF 及 MORBF 的平均資料傳送量效能最佳，此三種演算法在資料流量最小  $L(1,1)$  時為最佳之運作演算法；前述之討論，因為此三種演算法會找到量測最佳線路來傳送資料，其中 FRTTDR 及 FRTTFNF 是利用反應時間的方式來找尋最佳線路，而 MORBF 是利用上傳剩餘頻寬最佳模式來找尋最佳線路，所以此三種演算法均能避開壅塞的線路。

2.當資料流量最小為  $L(1,1)$  時，其他採用權重模式分配線路的演算法如：WMORBF、WFRTTFNF、RR 與 LLF，它們的平均資料傳送量效能則相對比較差。此一結果顯示：當企業網路資料流量少同時聯外鏈路又發生壅塞現象時，將流量以比較平均地分配在不同聯外鏈路傳送，鏈路負載平衡機制運作所獲得的平均資料傳送量效能相對較差。

3.當資料流量逐漸增加，使用權重模式分配線路的演算法，其所獲得的平均資料傳送量效能亦逐漸提升；因此，在企業網路資料流量為  $L(3,1)$  的時候，WFRTTFNF 及 WMORBF 演算法的平均資料傳送量效能最好，因為此二種演算法具量測線路壅塞的能力，亦使用權重模式的線路分配方式。此外，在企業網路資料流量為  $L(10,3)$  的時候，RR 與 LLF 演算法也呈現較好的平均資料傳送量效能，因為它們使用權重模式的線路分配方式，雖然它們並不具有量測線路壅塞的能力。

圖九(b)顯示聯外線路壅塞發生在最後一哩線路之外的地方。使用最後一哩作為量測距離的演算法(如 MORBF,WMORBF)其平均資料傳送量效能與圖九(a)相互比較，呈現較差的表現；使用最佳模式分配線路方式的 MORBF 在低交通量的時候 ( $L(1,1)$ )，其平均資料傳送量效能明顯的下降；此外，使用權重模式分配線路的演算法，如：WMORBF，在交通量增加的時候 ( $L(3,1)$ )，其平均資料傳送量亦未能顯現較佳之效能。

比較圖九(b)及圖九(a)在交通量最大的時候( $L(10,3)$ ),這時候使用權重分配方式的演算法(如 RR,LLF 和 WMORBF)則繼續保持較高的資料傳輸量。由上述的模擬結果，可以發現下列現象：

1.在交通量小的時候，演算法具備壅塞量測能力是提高資料傳送量的重要因素。

2.在交通量大的時候，演算法使用權重模式線路分配方式則是提高資料傳送量的重要因素；因為在交通量大的時候，藉由使用權重模式線路分配方式可充

份利用較壅塞線路的可用頻寬，進而增加所有可供使用線路的頻寬，但這將帶來資料傳輸延遲(delay)的問題。

## (二)線路斷線反應

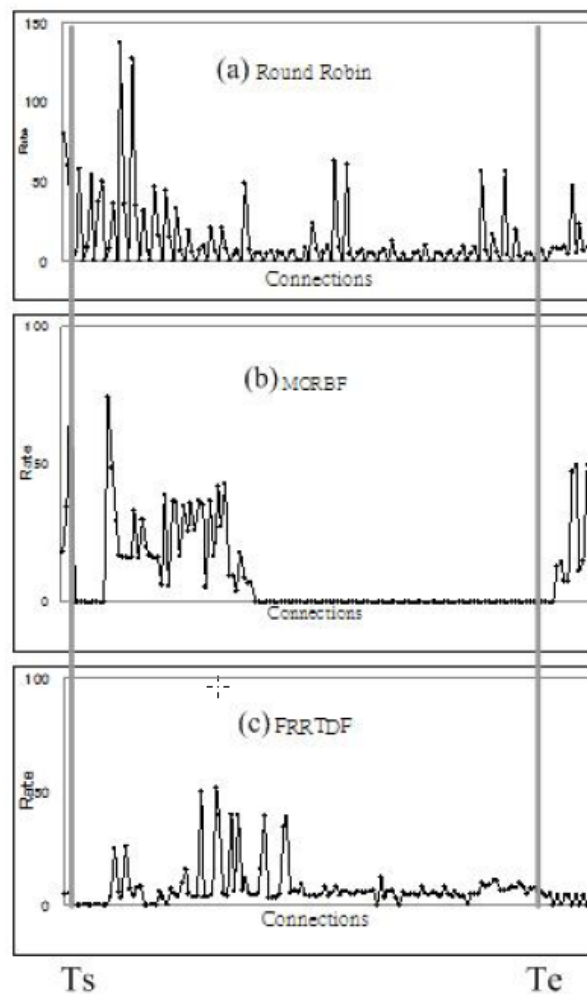
這小節的模擬想定將由用戶端連續產生 300 個 http 下載連線，並在一段時間之後在第二條聯外鏈路發生斷線的情況，並在一段時間之後，讓聯外鏈路線路斷線狀況能夠被偵測察覺。本研究將比較需要額外使用斷線偵測機制的演算法(如：RR 與 MORBF)和本身就可達成線路斷線偵測的演算法(如：FRTTDF)，如表 2 的分類；此外，亦將觀察使用權重模式分配線路的演算法(如：RR)和使用最佳模式分配線路的演算法(如：FRTTDF)的運作效能。圖 10(a)、(b)與(c)分別為 RR、MORBF 和 FRTTDF 不同演算法在企業網路聯外鏈路發生斷線狀況下所量測每個 session 的資料傳送量，其中  $T_s$  是網路斷線的時間， $T_e$  是偵測到網路斷線的時間。

由圖十的模擬結果，可以發現下列現象：

1.圖十(a)顯示：在  $T_s \sim T_e$  這段時間，許多的 session 的資料傳送量是 0，這代表這些 session 被分配到斷線的線路。由於 RR 演算法分配到不同線路的機率是均等的，因此會有 1/3 的 session 會被分配到斷線的線路。

2.圖十(b)顯示：MORBF 演算法在  $T_s$  時間點的附近的 session 都呈現傳輸失敗的情形，這代表這段時間正在進行傳輸的 session 都使用第二條已斷線的聯外鏈路，接下來有段時間 session 都成功地傳輸，然後接下來又所有 session 全數無法成功傳輸直到  $T_e$  時間點；由於 MORBF 演算法是使用最佳模式來分配線路，所以會有一段時間內把所有 session 分配至當時量測流量最好的線路；因此在斷線一段時間之後，由於這條已斷線的鏈路幾乎沒有流量，會被最佳模式認為是最好的線路，所以把所有的連線都分配到此一斷線的線路。

3.圖十(c) 顯示：FRTTDR演算法的量測時間是採被動式的方式，是以每個 session 來到的時間做為其量測的時間點；模擬結果顯示在  $T_s$  時間點的附近的 session 都呈現傳輸失敗的現象，這是因為雖然 FRTTDR 可以為所有 session 找到最佳路徑，但是正在進行傳輸的 session 則因無法再次進行線路的選擇，所以會發生 session 傳輸失敗的現象；接下來幾乎所有的 session 都呈現成功傳輸的現象，並不因有鏈路斷線而受到影響，這代表 FRTTDR 演算法本身即有能力避開斷線鏈路的功能。



圖十 Load Balance 演算法對斷線影響的資料傳送變化圖(作者繪製)

研究發現大部分的 session 資料傳送量均偏低，這是由於所有 session 的頻寬負擔需求都落在剩餘的兩條線上，因此甚至會有少數連線因傳輸超時(timeout)而連線失敗。但總體而言，FRR TDF 成功率還是最高，在 300 個連線中，RR、MORBF 與 FRR TDF 的失敗連線數分別為 46、88 及 30。

## 伍、結論

為了追求穩定的網路服務品質以及保持網路暢通不斷線，許多的民間企業選擇配置多條對外的網路線路，因為可以使內部網路對外時可以擁有較高的可靠性及較好的效能，亦能達到節費的成效。藉由鏈路負載平衡機制之運作可以避開壅塞和斷線的路徑，提升整個單位的網路品質。

對於我們現有國軍網路的服務品質一樣地重要，本研究利用許多負載平衡演算法去模擬測試在鏈路壅塞反應及斷線反應所呈現的不同網路效能與分析數據，期使日後在國軍網路應用上所面臨的線路壅塞問題和斷線問題能夠利用所分析的數據，調整負載平衡演算法以能避開有問題之鏈路，提昇網路的穩定性和可靠性，並能提供一個不間斷的網路服務。