**The TWAREN 100G Backbone Path Fluctuation and Its Detection by Using Measurement Tools**

**古立其 楊哲男 陳俊傑**

**國家高速網路與計算中心**

**[lku, yangcn, jjchen]@narlabs.org.tw**

# 摘要

TWAREN在100G骨幹設計所採用的多層容錯架構具備高度的容錯能力，使得單線段障礙、甚至單接取點障礙皆不致影響服務。但複雜的多層容錯架構本身卻大幅增加了故障偵測及查修的困難。本研究發揮TWAREN骨幹量測伺服器所構成的主動式量測網所具備的靈敏偵測能力，長時間追蹤各網路線段細微的線路延遲時間變化，以便能在底層網路故障發生、骨幹線路自動切換時主動發現，以提供網路管理團隊更即時、更精確的資訊參考，使網路障礙能更快確認位置，提高整體網路品質。

**關鍵詞**：網路效能量測、100G骨幹、路徑更動偵測

# Abstract

A multi-layer fault tolerant architecture has been incorporated while TWAREN 100G new backbone was designed. While extremely effective on defending single fiber or single point of failure, the architecture itself also shields many symptoms of failure from being properly detected by the operation center. This study proposed a way to detect such backbone path fluctuation automatically by utilizing the perfSONAR performance monitoring mesh network thus the network failure can be easier identified.

**Keywords**: Network measurement, TWAREN 100G, perfSONAR, Backbone path fluctuation.

## 前言

網路的頻寬及其品質為影響各種網路應用使用經驗的關鍵。網路頻寬的提高主要透過所採用的傳輸技術的升級及可用線路的增加來獲得。而網路的可靠度則與所使用的網路設備是否可靠、線路是否穩定、供電是否恆定、維運是否積極等諸多因素相關。這些因素甚至會受天災的發生及大環境的整體因素影響，因此網路的障礙因素不可能完全排除。要獲得高度可靠的網路品質，勢必需要搭配良好的網路架構設計，使不可避免的網路障礙發生時，能藉由架構所提供的額外備援能力而保持訊務的相對完整。

然而提高架構的容錯能力，勢必增加網路設計的複雜度。當網路複雜度提高到某種程度時，必然為網路的維運及故障的偵測及排除帶來挑戰。本文即以網路效能量測的觀點切入，提供高複雜度的多層容錯式網路一個路徑變動偵測的新觀點。

1. **TWAREN 100G骨幹架構**

國家高品質學術研究網路 (Taiwan Advanced Research and Education Network; TWAREN)[1] 甫於2016年完成100G骨幹網路升級，其架構示意如圖1。完成升級之後，TWAREN除了具備高達100G的骨幹網路頻寬之外，其多層容錯架構更確保網路品質不會受到單點中斷、單線段故障等障礙的影響，而能持續提供最高品質的網路傳輸服務。



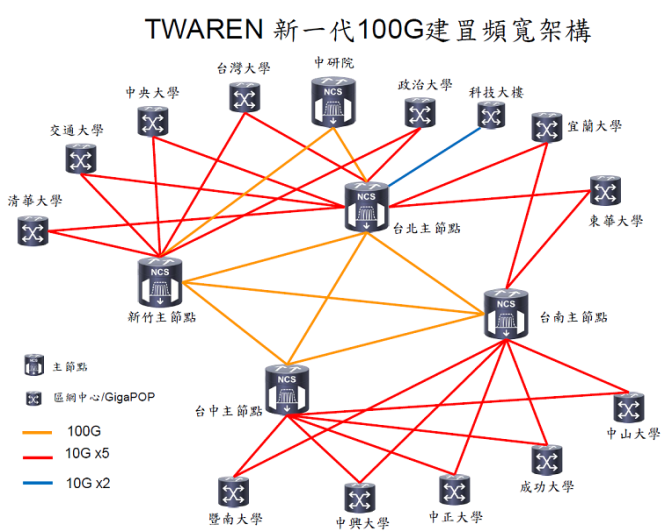
**圖1 TWAREN 100G骨幹架構示意圖**

TWAREN骨幹由核心節點 (Core Node)、網路接取節點 (GigaPOPs) 及連接其間的高頻寬網路所構成。5個核心節點分別位於台北、新竹、台中、台南及中央研究院，核心節點間由100G線路彼此互連。12個網路接取節點則分別為台灣大學 (NTU)、東華大學 (NDHU)、中央大學 (NCU)、清華大學 (NTHU)、交通大學 (NCTU)、中興大學 (NCHU)、暨南大學 (NCNU)、中正大學 (CCU)、成功大學 (NCKU)、中山大學 (NSYSU)、政治大學 (NCCU) 以及宜蘭大學 (NIU)，各節點擁有10條10G線路，分別連接兩個上游核心節點。各節點間則使用租自電信業者的暗光纖搭配100G ROADM及OTN設備提供最高彈性之高品質連線頻寬。

TWAREN所採用的多層容錯架構分別於 Layer 3、Layer 2及Layer 1提供對應的容錯機制，以提供骨幹網路面對來自各種可能障礙的能力。各層的容錯設計分述如下。

* 1. **Layer 3 容錯設計**

TWAREN 100G新骨幹與前代架構相較，最顯著的不同就是位於網路第三層的「異地雙中心」傳輸架構。在此架構下，任一網路接取節點都同時與兩個不同的核心節點相連接。其路由器採用Dual Homing到兩個核心節點的傳輸架構，使任一主節點故障時，網路接取節點的網路路由均能自動繞至尚能正常連線的另一核心節點提供訊務服務，不致於影響網路接取節點與核心節點層的連線。



**圖2 TWAREN新骨幹Layer3架構圖**

以上圖中的台灣大學為例，其以50G頻寬連接台北核心節點，同時以另外50G頻寬連接新竹核心節點。因此若台北或新竹兩個核心節點的其中之一發生障礙時，台灣大學仍保有50G的連網頻寬與整個骨幹相連，因此台灣大學及連接於台灣大學之下的其他連線單位均不會發生斷線。

實務上這個異地雙中心的架構中，各網路接取節點的兩條上行連線均同時使用，並非以其中之一保持閒置等待另一條線路故障才接手的純備援架構，以確保網路的最佳使用效益。目前的設計是以IPv4訊務優先走於其中一路、而IPv6訊務則優先走於另一路的方式來切分。

* 1. **Layer 2 容錯設計**

各網路接取節點連向核心節點的50G線路是由5條10G線路建立Port Channel的方式合併頻寬而成。

**圖3 Layer 2虛擬線路示意圖**

當其中任何線路發生障礙時，網路交換器會自動將訊務分配至其餘正常的線路之中。因此單一線路的故障僅造成總可用頻寬減少10G，並不影響實際訊務的傳輸。

* 1. **Layer 1 容錯設計**

為了減少單點故障造成的影響，由各網路接取節點連向核心節點的各條10G線路實際上均繞行不同路徑，或是途經其他網路接取節點才到達核心節點，實際訊號路徑長度各異，因此從Layer 3的觀點來看，各段邏輯線路所量測得到的封包延遲時間與節點間的直線地圖距離並不一定成比例。

## 網路量測方式

在眾多網路效能量測工具中，本研究採用高度模組化、容易客製化、同時安裝容易的perfSONAR工具包[2]進行骨幹網路的相關量測。

* 1. **perfSONAR**

由於網路效能及品質的監測是眾多跨單位網路合作的必備基石，特別是跨國學術合作更容易受到網路品質的影響。為了加強合作國家間的國際合作與交流，同時為彼此共同推動的學術研究計畫提供一致的網路基礎工具，歐洲的GÉANT2與美國的Internet2及ESnet合作開發 perfSONAR (Performance Focused Service Oriented Network Monitoring Architecture) 網路效能監控基礎設施。perfSONAR是同時具備效能監控與視覺化結果呈現的應用程式。這套工具設計上運作於中間層 (Intermediate Layer，主要在處理數個網路間效能測量結果的交換)。目的在於幫助網路管理者找出影響網路效能的原因並解決問題，即使跨越不同網域 (Multi- Domain) 時，仍能輕易的解決點對點的網路效能問題，保證網路使用者能夠擁有高品質、高效能的網路。perfSONAR以服務導向架構 (Service-Oriented Architecture; SOA) 為基礎，主要提供的服務為:

* Measurement Point (MP) Service:網路量測服務。
* Measurement Archive (MA) Service: 儲存和提供量測結果。
* Lookup Service (LS):搜尋其它 perfSONAR服務並提供。
* 服務註冊的功能以分享其所擁有的服務。
* Authentication Service (AS):授權和認證服務 。
* Topology Service:提供拓撲資訊服務。
* Transformation Service:轉換資料服務 。
* Resource Protector Service:控管有限的量測資源 。

因為perfSONAR 是一個可跨網域之網路效能監控平台，使得能更簡單的解決跨不同網域間之點對點網路問題，在一個聯盟化的環境下，利用多種不同的Services來做為效能量測之用，根據共通的標準(OGF NM-WG)，可自行開發所需程式大家可互相分享資料。

* 1. **perfSONAR Toolkit**

perfSONAR Toolkit集合了perfSONAR-PS專案所開發的各種軟體，以CentOS為運作環境所客製化的一個LiveCD。包含了包含了BWCTL, OWAMP, NDT, NPAD, Pinger, Traceroute, Reverse CGIs等工具，並且還啟動了Apache2、MySQL、NTP及esmond等服務，目前的穩定版本為4.0.1。

* 1. **PingER**

PingER (Ping End-to-end Reporting)原生於Internet End-to-end Performance Measurement (IEPM) 計畫，是用來量測點對點間之RTT資訊。目前有超過150個國家之節點正執行PingER程式中，其中亦包含TWAREN。在後續版本的perfSONAR Toolkit中亦將PingER收納到網路效能監控之服務中，並使用了perfSONAR Toolkit中的另一個服務BWCTL來做排程控制，透過ping程式來兩兩監控點對點間的效能。使用ping做測試的好處，例如很多節點可接受ping測試，不需修改防火牆設定或是在目標主機上執行特別的程式。

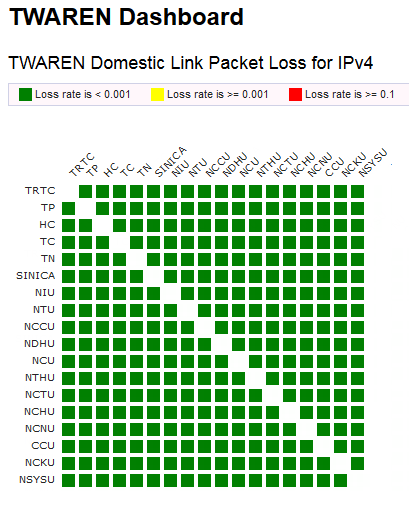
* 1. **TWAREN量測佈建現況**

TWAREN目前在5個核心節點、12個網路接取節點以及3個國際交換節點皆佈建了perfSONAR量測主機，直接連接於各節點的TWAREN路由器下。

其中位於國內各節點的perfSONAR量測主機設計為彼此定時兩兩對測。以各點間的封包遺失狀況為例，統整後的骨幹量測結果如下圖所示[5]。由於TWAREN國際線路在邏輯上由台北及新竹核心節點出發至美國三地交換中心，因此位於國際交換中心端的3台perfSONAR主機除了彼此對測之外，僅與國內核心節點間進行對測。

於各台量測主機安裝完成perfSONAR Toolkit之後，可透過中央統一管理的方式，讓每台主機讀取遠端的同一個json設定檔案，方便管理者只要設定一次，即可完成所有量測主機的設定。亦可避免每一台主機之管理設定策略不一樣，減少執行錯誤的發生。每台主機開始執行各項量測工具，包括每分鐘定期量測one-way之延遲時間及封包遺失率，每5分鐘定期量測RTT之延遲時間。收集到的資料會存在本機之資料庫裏。

由於perfSONAR為獨立的量測工具，每台perfSONAR僅具備與其進行對測的各點至自己間的量測數據。為了得到全骨幹的綜合數據，TWAREN另行開發資料收集程式，自所有perfSONAR量測主機取得各點間的量測資料，由資料收集程式匯整之後統一存入網管中心的數據資料庫中。以供維運人員參考各線路現況及判斷障礙發生位置之用。其中網路延遲時間的匯整數值即用於本研究之骨幹路徑變動偵測之中。



**圖4 TWAREN perfSONAR 偵測封包遺失**

## 骨幹路徑變動偵測

* 1. **骨幹路徑變動原因探究**

TWAREN骨幹擁有多層的容錯設計，因此當單一實體線路發生斷路時，原有訊務會改走同一Layer 2 虛擬線路中的其他實體路徑，Layer 2 的連線狀態尚不會受到影響。唯有當某網路接取節點連向上游某核心節點的　Layer 2　虛擬線路中的每一條實體 Layer 1 光路徑全數斷線時，該　Layer 2 虛擬線路才會斷線，繼而觸發路由器選擇另一條可用之 Layer 2 虛擬線路連向另一核心節點，訊務仍然不受影響。直到此條 Layer 2 虛擬線路中的所有實體線路亦全數故障，才會導致該網路接取節點的網路完全中斷。

以交大的網路連線為例。交大有一路 Layer 2 虛擬線路連向新竹核心節點、另有一路 Layer 2 虛擬線路連向台北核心節點。當交大直接連向新竹核心節點的某條實體 Layer 1 線路發生斷線時，訊務即可能改繞經另一條邏輯上由交大連向新竹核心節點，但實際 Layer 1 線路先走到清大，再由清大連達新竹核心節點的路徑。如下圖所示。



**圖5 Layer 1 線路斷線時訊務繞徑示意圖**

若交大連向新竹核心節點的這條 Layer 2 虛擬線路內的所有實體線路均中斷，此時位於交大的 TWAREN路由器會改走台北主節點的路由，原有訊務不受影響，如下圖所示。

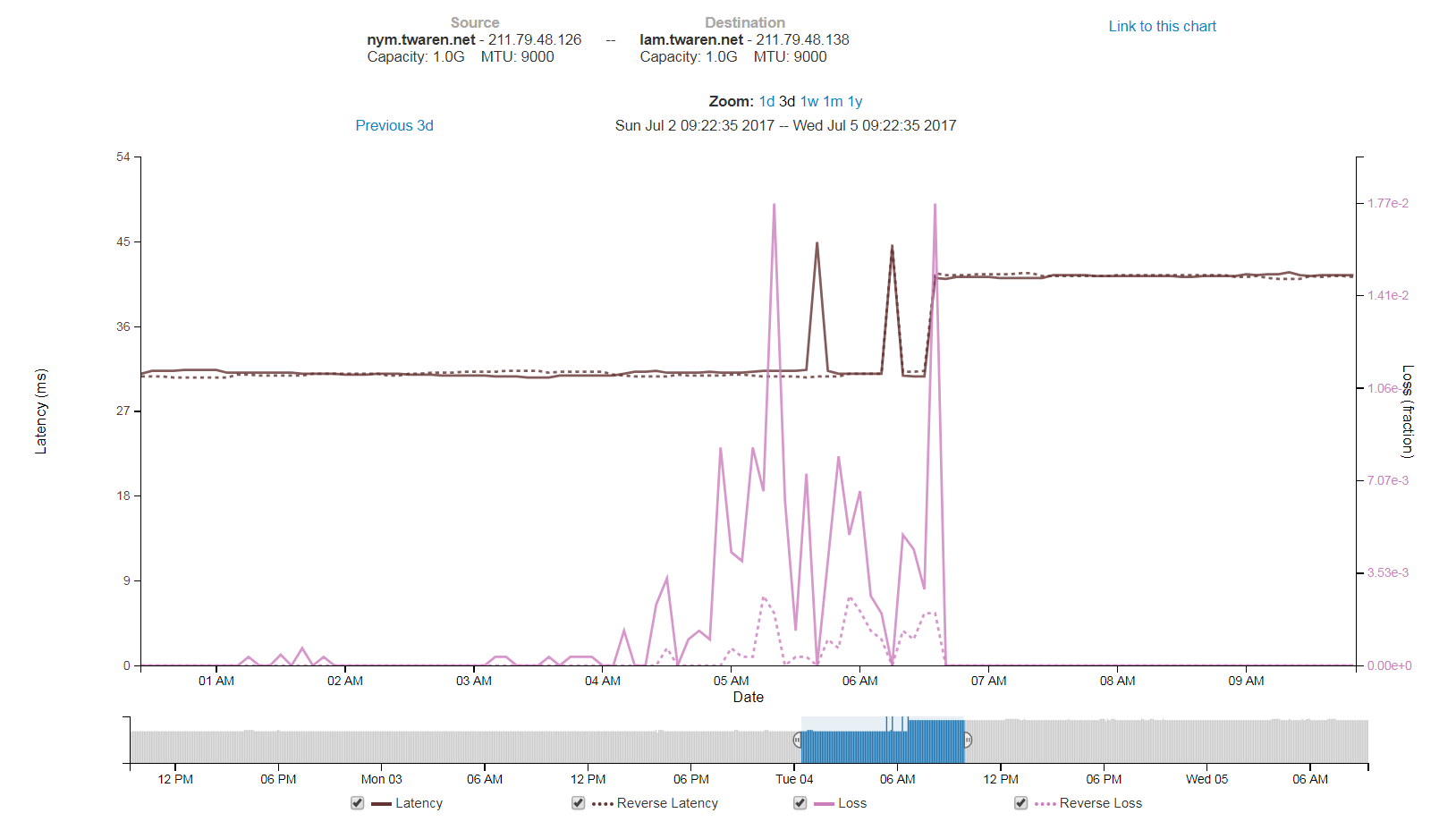


**圖6 Layer 2 線路斷線時訊務繞徑示意圖**

必須全數 Layer 2 虛擬線路均斷線，交大的對外網路才會中斷。

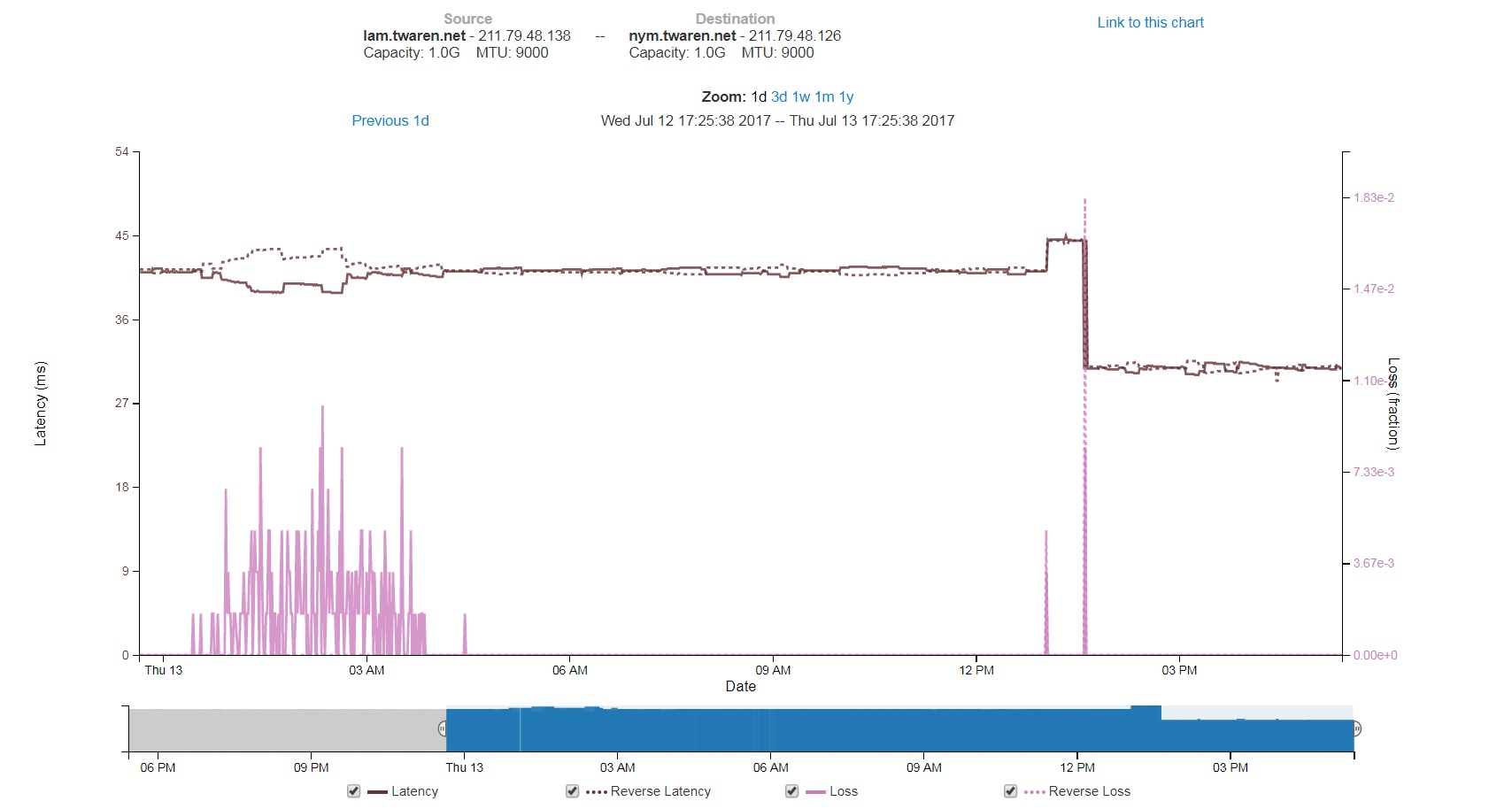
然而多層容錯架構雖然帶來對於線路及節點障礙的高度容錯能力，卻也大幅增加了網路架構的複雜度。以上例來看，當清大至新竹核心節點間的某條暗光纖發生斷線時，有可能在邏輯上卻是交大至新竹核心節點的某條 Layer 2 虛擬線路內的實體線路障礙。而由於IPv4訊務及IPv6訊務走在不同 Layer 2 虛擬線路的設計，當交大至新竹核心節點的 Layer 2 虛擬線路中發生部份實體線路中斷時，理論上 Layer 2 虛擬線路仍然存在，但實際可用頻寬減少，就有可能在網路流量較大時發生 IPv4 訊務壅塞，封包遺失過高，而IPv6訊務卻仍然順暢的怪異現象。

* 1. **使用perfSONAR偵測骨幹路徑變動**



**圖7 路徑變動時之RTT變化圖(1)**

由於perfSONAR可以長時間收集骨幹各點間的量測數據，其中 PingER 工具所收集的RTT數據便可做為骨幹路徑異動的重要參考。以上圖為例，在特定時間點，這段骨幹線路原本恆定的RTT值突然發生變化，之後恆定於另一較高的值。實際上在該時間點發生線路的中斷，導致訊務改繞至 Layer 2 虛擬線路內的另一實體線路傳輸。邏輯上在同一虛擬線路內的各條實體線路均位於同樣的兩個端點間，但實際上的實體路徑可以有極大的差距。由於 perfSONAR可以精確測得骨幹RTT的變化，因此可以發現線路切換至另一實體路徑所產生的改變。



**圖8 路徑變動時之RTT變化圖(2)**

上圖為紐約至台灣的RTT變化圖。由於TWAREN連至美國的兩條國際線路長度差異甚大，因此發生線路中斷時，產生的RTT影響也較大。圖中右側可以看出線路原本恆定的RTT先微幅變大，而後較大幅度變小之後保持新的恆定。這是由於TWAREN的國際線路發生單纜斷線，封包改繞至另一路徑，導致RTT上升，而後電信業者改以另一距離較短的海纜備援，訊務RTT因而下降。

由上例可以看出，網路效能量測系統運用於多層容錯架構下偵測骨幹路徑異動的價值。未來若能加以進一步發展，相信可以對於網路障礙的偵測發揮更重要的角色。

**參考文獻**

1. Taiwan Advanced Research and Education Network (TWAREN), http://www.twaren.net/
2. perfSONAR toolkit, http://psps.perfsonar.net/toolkit/
3. TWAREN Network Operation Center (NOC), http://noc.twaren.net
4. 100G Circuit switch announcement, <http://211.79.61.102/noc_2008/NOCBulletin/>
5. TWAREN Dashboard, http://ps-dashboard.twaren.net/