

TWAREN 第二代架構無中斷移轉之設計與實現

梁明章

國家實驗研究院國家高速網路與計算中心

liangmc@nhc.org.tw

摘要

當 TANet[1]與 TWAREN[2]共用骨幹之後，TWAREN 第一代架構喪失備援能力導致維運風險升高，為解決問題，TWAREN 第二代架構(後文簡稱 TWAREN2)因應而生，然因兩代架構無法並行，在沿用原有設備的情況下，如何移轉方能使施工期間的網路服務中斷時間減少，本文將說明整個移轉設計所面臨的問題與解決方法，最終達成網路服務無中斷的移轉，令 TWAREN2 順利上線。

關鍵詞： TWAREN，TWAREN2，無中斷移轉。

1. 前言

當 TANet 與 TWAREN 共用骨幹之後，TWAREN 第一代架構面臨了原始設計未預見的應用與變化，原本兩套純路由網路可互為備援，然因設備無法提供 Layer2 多點 VPN (Multi-Point to Multi-Point Virtual Private Network 多點虛擬私有網路) 給 TANet 使用，為徹底區隔路由，不得不由 TANet 與 TWAREN 研究網各自使用一套，導致備援能力消失，形成許多單點失能風險，故維運團隊設計 TWAREN2 架構以解決問題，並達成可提供複數多點 Layer2 VPN 給使用者的目標，然因經費有限且設備使用年限未到等因素，TWAREN2 無法全新建置，在沿用原有設備的情況下，兩代架構不能並行，如何替換方能使施工期網路服務中斷縮短，使最大用戶 TANet 願意接受 TWAREN2 上線，頓成維運團隊亟需克服的難題。

歸納移轉替換所面臨的困難，可分四個大議題，詳述如后：

1.1 介面不足使新舊實體線路無法並存運行

參考圖 1與圖 2，顯示兩代架構在實體線路上的異同之處，許多線路從 10GE (Ten Gigabit Ethernet) 改成 SDH (Synchronous Digital Hierarchy)，而 TWAREN 本身的設備並無新增，無足夠的介面可供新舊骨幹同時上線，僅能以替換方式進行，在拔插之間必然斷線，而且新線路無法保證與 TWAREN 設備介面的匹配度可達到一接即通，根據維運團隊過往的經驗，固網與我方異種設備之間的接線有相當機率無法在第一次順利接通，故單純的拔插替換方法造成的中斷時間無法精

準預估，難以說服用戶接受 TWAREN2 施工，若能設計出不受拔插中斷時間影響的移轉方式，如此一來可完全閃避此議題所產生的風險。

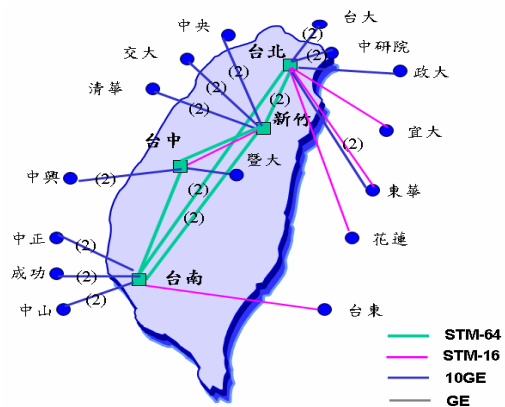


圖 1 TWAREN 第一代架構實體線路圖

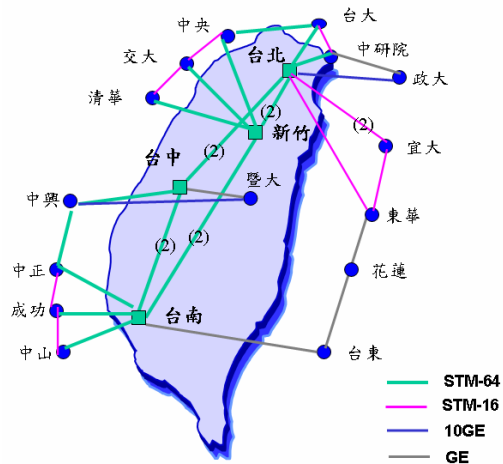


圖 2 TWAREN 第二代架構實體線路圖

1.2 兩代網路架構原理迥異

TWAREN2 架構核心乃 VLAN[3]搭配 EoMPLS (Ethernet over Multi Protocol Label Switching) 的方式(如圖 3)，骨幹已非純路由網路，乃是 Hybrid

Network，架構與設定皆不同於第一代，卻需運行在原有相同設備上，與線路同樣面臨無法並存運行的難題，當設備設定變更時會形成服務中斷，因此流量不僅需事先導向不同線路，也需經過不同的設備，才能徹底閃避設定變更造成的風險。

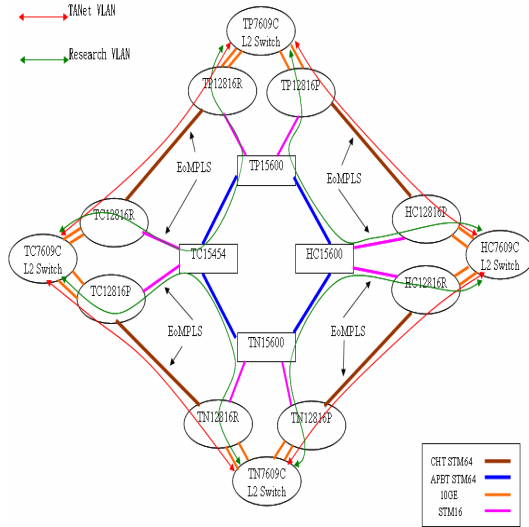


圖 3 TWAREN 第二代架構核心原理圖

1.3 替代路徑需足以容納 TANet 龐大流量

TANet 由於出國口位於台北，流量有越北越龐大的累積特性[4]，進台北端流量超越 5Gbps，由於採購時程的關係，移轉時間無法趕上流量低的暑假內進行，故替代路徑需面對學期中龐大的流量與使用者，因此中斷時間或頻寬壅塞若超過半天將令 TANet 無法接受，此為第三個難題。

1.4 牽涉規模太大

全國性架構的改變，若全體設定沒有一次變更，將導致網路分裂孤立，然而此次變更工作範圍遍及所有區網中心，In-Band 與 Out-Band 合計近百個設備，數十條電路與上百條短跳線，規模之大遠超過維運團隊之人數，不可能全面同時變更，因此第四個難題就是如何切割工作，使能分批進行又不至於網路分裂孤立。

總結上述難題，最大的關鍵莫過於服務中斷時間的長短難以估測且可能超過用戶的承受能力，導致用戶不敢同意新架構施工，因此我們集中思考無中斷移轉的可能性，且因人力不足，移轉工作需能切割分批進行，除避免上述風險之外，移轉方法需能明確地讓用戶瞭解施工不會影響網路服務，爭取用戶的信任與准許，方能使 TWAREN2 獲得實現機

會。

2. 移轉方法之設計內容

基於前述議題與規劃原則，設計出長達十三個工作天的移轉步驟，分成四個階段、十幾個大步驟，數十個小步驟，簡述如后。

2.1 建立新核心之雛形

第二代核心乃以四台 7609[5]（後續稱為 7609C）組成一四邊形 VLAN 環境為主體，因骨幹電路乃 STM64（Synchronous Transport Module level-64, 10Gbps），無法直接上 7609C，故中間透過 GSR[6]以 EoMPLS 方式令 Ethernet 得以跨過 SDH 相接，然因 GSR 與 ONS[7]的 Port 皆不足，故新核心 7609C 無法先接新骨幹電路。

此外，新核心四個 Core 之間的接法也不同於第一代，故先利用 OSPF[8]路由網路自動尋徑的特性，逐次調動原光網路骨幹電路，如圖 4，不會造成服務中斷，且 Port 可逐步挪用。

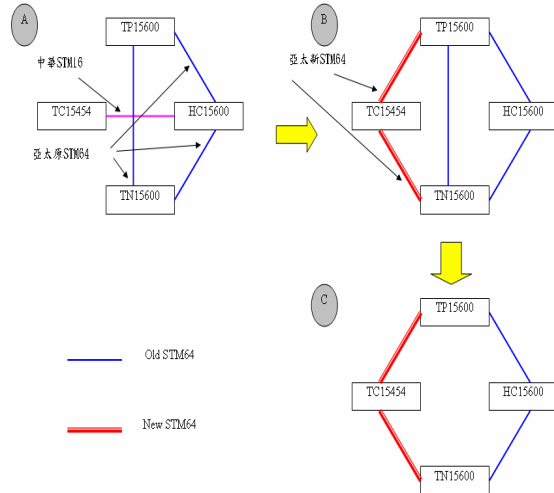


圖 4 光骨幹移轉

光骨幹移轉成功後，可從 ONS15454 切出 STM8（1.2Gbps）光路徑，轉成 1GE 將四個 Core 7609C 銜接成 VLAN 核心之雛形以為臨時新骨幹，如圖 5。

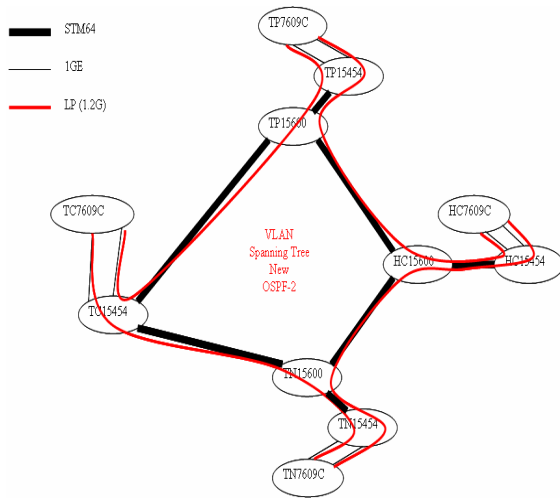


圖 5 建立 VLAN 核心雛形為臨時骨幹

2.2 移轉研究網

當臨時新核心完成時，表示 7609C 已可運作，此時 7609C 先切好各個 VLAN 以等待即將移入的用戶。然因此臨時核心僅有 1GE 之骨幹頻寬，即使先將流量較小的研究網移入也會造成壅塞，依照流量往北累積的特性，我們由南而北逐次進行研究網移轉，每完成一個 Core (TWAREN 骨幹核心機房) 及其下游 GigaPOP (TWAREN 骨幹區域網路中心)，就調度 10GE 卡版至 Core 7609C 上，以擴張骨幹頻寬。

由於所有 GigaPOP 無法同時施工，為避免網路分割，研究網所用的 GSR 先以 GE 連上 7609C 進入 VLAN 運行 OSPF，各 GigaPOP 之 7609 (後續稱為 7609P) 以 GE 經光網路至 7609C 接上 VLAN 亦運行 OSPF，此時 7609P 原 10GE 骨幹尚未截斷，故同時存在兩骨幹，舊的 10GE-GSR 骨幹與新的 GE-VLAN 骨幹，OSPF 在兩骨幹間可互通，再來於 7609P 上利用 OSPF 之優先權設定將流量切入 VLAN 骨幹，如圖 6，觀察無異常後，就截斷 10GE，並將 10GE 卡版帶回 Core 7609C 安裝。

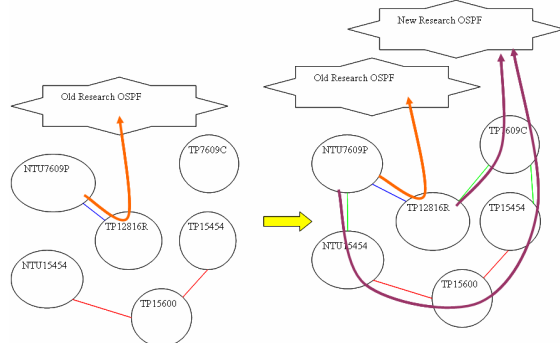


圖 6 移轉 GigaPOP 7609P 之流量

當 Core 7609C 獲得 10GE 卡版之後，與 GSR 之連線即可升級為 10GE，且 GSR 因下游 GigaPOP 已經移轉上臨時骨幹，可功成身退，變更設定為 MPLS 環境，建構 EoMPLS 路徑使 7609C 能以超越 1GE 的頻寬跨越 SDH，如圖 7，當 GSR-EoMPLS 骨幹設定好，此時調整 1GE Light-Path 與 EoMPLS 兩邊的 Spanning Tree[9]優先權，以 EoMPLS 優先，之後將 EoMPLS 介面啟動，就能將 7609C 的流量轉入 EoMPLS，至此臨時骨幹的頻寬可擴張至 STM16 (2.5Gbps) 等級，足以應付全研究網。

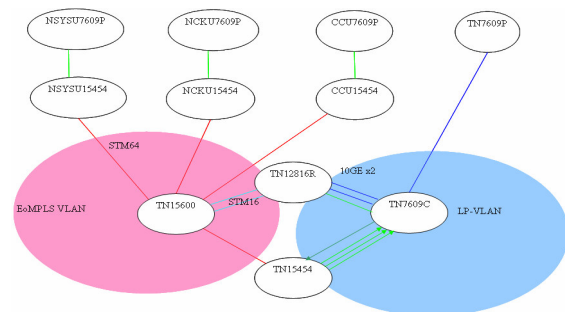


圖 7 移轉 Light-Path 流量至 EoMPLS

2.3 移轉 TANet

移轉進行至此，確保服務皆未曾中斷，後續進行最大用戶 TANet 之移轉。儘管新骨幹已經具有 2.5G 之頻寬，然 TANet 由南至北的流量卻超過 2.5G，因此在移轉 TANet 之前，必須給予臨時骨幹足夠的頻寬。

觀察新舊中華電信骨幹的差異，可發現應變的契機，如圖 8 所示，僅需一條新建電路及遷移一條舊電路，使台南可經由台中 GSR-TC12816R 到達台北，即可提供足夠頻寬 STM64 予 TANet 移轉使用。

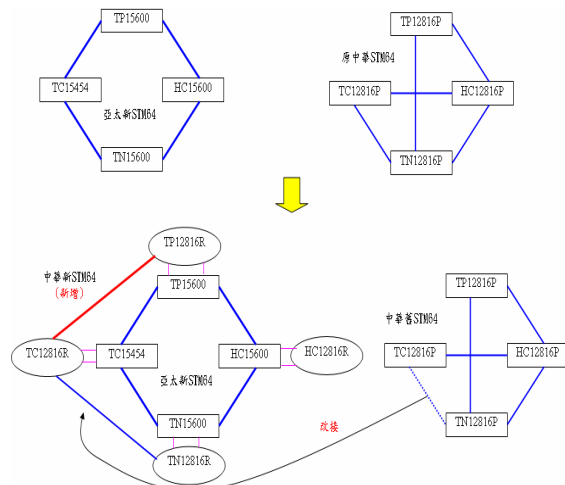


圖 8 調度中華電信骨幹以應付 TANet 頻寬

骨幹電路變更之後，調整核心骨幹 EoMPLS 設定如圖 9，令 TANet 移入後，可暫時與研究網共用新竹-台北之 2.5G 頻寬，台南-台中-台北則由研究網使用亞太骨幹 2.5G 頻寬，TANet 使用中華電信 STM64 頻寬，則此臨時骨幹足可應付 TANet 龐大流量，並可與舊的 TANet 骨幹並行一段時間。

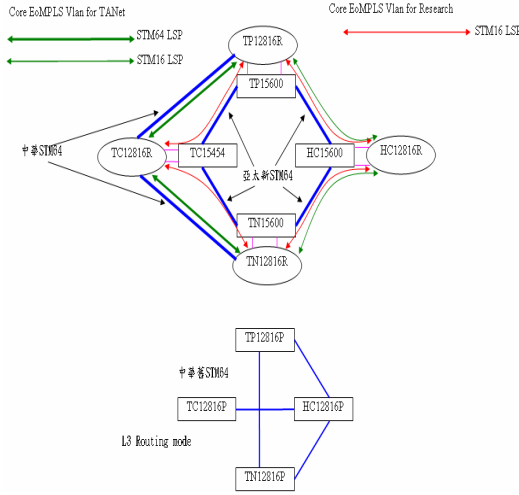


圖 9 調整骨幹 EoMPLS 路徑

當臨時骨幹就緒，開始於各區網中心進行流量移轉，原理如同研究網，區網 6509 以四條 GE 運行 802.3ad (Cisco 稱為 Ether-Channel[10]) 透過 ONS15454 連至 Core 7609C 接上配置給 TANet 的 VLAN，而 TANet 所使用的 GSR-P 亦透過 10GE 連上 7609C 加入此 TANet VLAN，兩套骨幹由此相接，通行 OSPF，此時區網 6509[11]同時連接新舊骨幹，透過 OSPF 優先權設定將流量切入新骨幹，確認無異常之後就截斷舊的 10GE 骨幹，隨即將 10GE 卡版調度至 Core 作後續步驟使用。

如圖 10 所示，由於移轉仍是利用 OSPF 優先權變更切換流量路徑，因此服務不會中斷，更因兩套骨幹有 OSPF 相連，各區網無須同時施工。

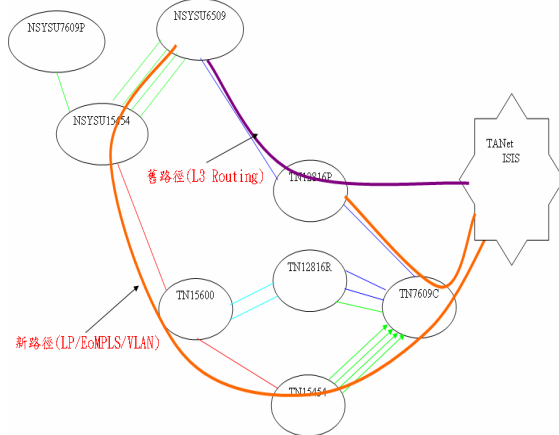


圖 10 移轉 TANet 區網流量上新骨幹

當 TANet 所有區網中心皆移轉完成時，整個新骨幹架構如圖 11 所示，原本 TANet 使用的 GSR-P 與中華電信骨幹因而閒置，此時將進行移轉步驟的最後階段。

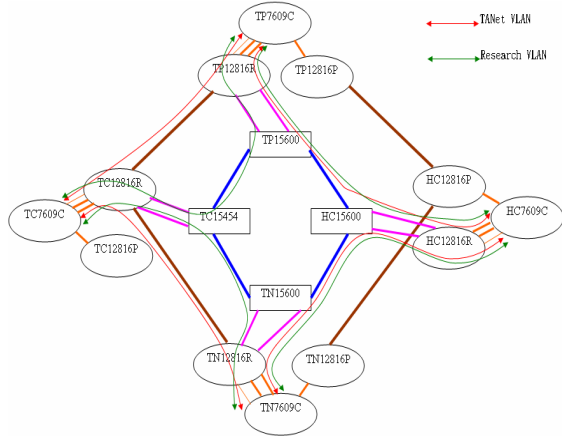


圖 11 移轉 TANet 區網流量上新骨幹

2.4 調整為 Dual GSR 架構

由 TWAREN 的維運事件紀錄分析，在設備失能事件中佔比例最高的是 GSR 失能，為了降低 GSR 失能造成的影響，且每個 Core 正好有兩個 GSR，因此將 GSR 調整為備援架構，未來單一 GSR 失能僅會造成 Spanning Tree 重新收斂而不會導致服務中斷。由於此時新骨幹已經運行，四邊形 VLAN 環形架構實際上僅能運行三邊，必有一邊需以 Spanning Tree Block 以避免迴圈產生，利用這個特性，以每次施工一邊的方式，逐步將閒置的 GSR-P 與中華電信電路整合入新骨幹，連續動作如圖 12、圖 13、圖 14：

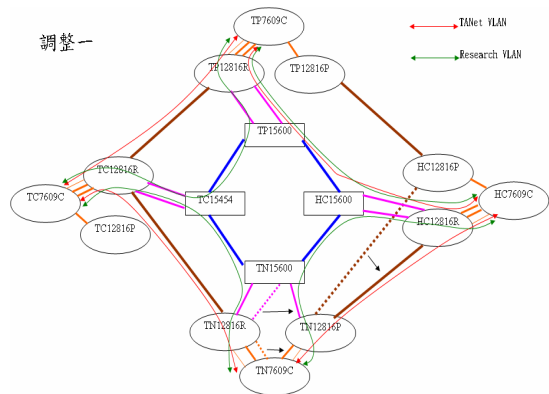


圖 12 調整 Dual GSR 架構步驟一

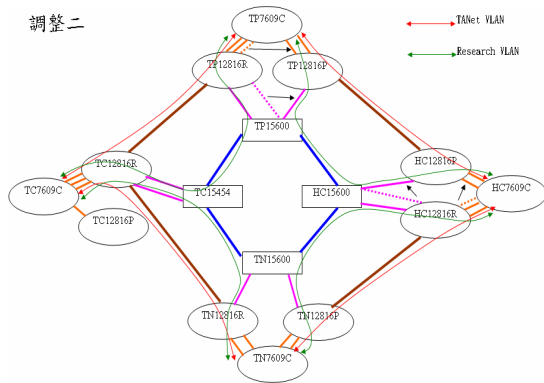


圖 13 調整 Dual GSR 架構步驟二

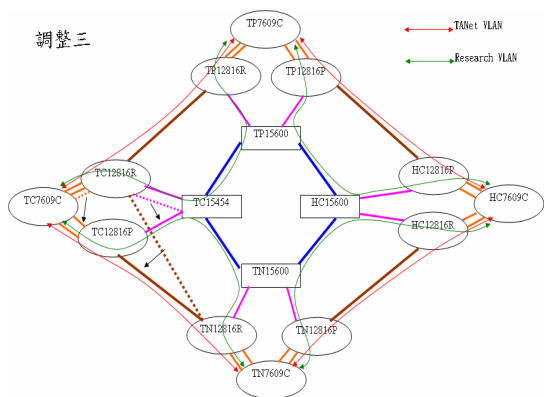


圖 14 調整 Dual GSR 架構步驟三

最終完成的核心架構如前述圖 3，細部的設備與連線架構圖如下圖 15 所示。

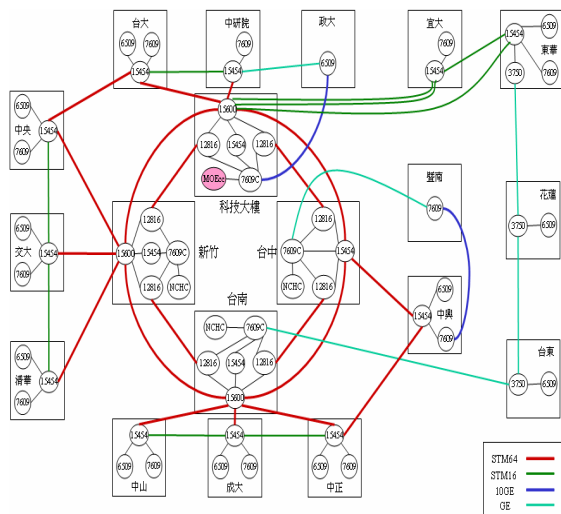


圖 15 TWAREN2 設備與連線架構圖

3. 移轉之實現

移轉設計雖完成，但欲使 TWAREN2 實現，仍

須一些事前準備工作，例如獲得用戶的施工同意是最為關鍵，其他尚有維運團隊、區網機房管理人、設備廠商、電路廠商等各方人員的工期協調與人力調度等等，說明如后。

3.1 爭取用戶之認同

為使 TWAREN2 獲得實現，維運團隊必須說服既有用戶同意 TWAREN2 施工實作，既然服務無中斷的移轉設計已經完成，接續的工作乃如何說服大用戶們，亦即教育部電算中心與各區縣網中心，明瞭 TWAREN2 之優點與移轉設計無風險，為此維運團隊歷經教育部電算中心會議、TANet 技術小組會議兩次當場簡報，終於獲得同意施工。簡報主旨說明如后。

3.1.1 可提供多點 Layer2 VPN 服務

TWAREN2 以光路徑、EoMPLS 與 VLAN 構築交換骨幹，對用戶而言，骨幹內不再有 Layer3 路由轉送點存在，因此 TWAREN2 可以打造多個多點對多點的 Layer2 VLAN 提供給多個用戶而不會形成路由混淆，比第一代 Layer3 骨幹更具有 Routing Hop 少、用戶網路可邏輯區隔的優點。

3.1.2 可提供多重保護備援

TWAREN2 所租用的固網電路，皆為雙路由電路，若工作路由斷線，需在 50ms 之間切換至保護路由，此為第一層保護。

TWAREN2 架構由許多環狀構成，參考圖 15，所有網路節點皆具備至少兩組對外線路，任一路徑斷線，皆可切換至備援路徑而不至於完全中斷，此為第二層保護。

設備部分的備援保護，在 Core 機房內的設備如有失能，可視為 Core Failure，整個 Core 下游的流量將轉向其他 Core 進入骨幹 VLAN，服務仍可勉強持續，此為第三層保護。

至於 GigaPOP 機房的設備，受限於區網內集中式架構與單一設備經過點的問題，無法以流量轉向處理，但因 TWAREN2 所使用的設備少於第一代架構，維運團隊有多餘設備可預為備品縮短設備替換的時間。

3.1.3 移轉施工期間服務保證不中斷

詳如前章節所述之設計，於說明之後，各用戶之代表皆能認同此移轉設計在理論上確實不會導致服務中斷，可達成無痛移轉。

3.2 人力與工期之協商與調度

為降低成本，本移轉設計以時間換取人手，以縝密的設備暨卡版調度步驟避免大量購買卡版，於

設備與設定部分參與的人未超過七人，以免人多手雜、溝通費時，且每次施工均無顧慮服務中斷時間長短的壓力（因服務並未中斷），參與者可慢工出細活，精確且高品質地完成任務，相對降低了意外狀況發生的風險。

4. 結論

為實現 TWAREN 第二代架構以改善第一代面臨的問題，維運團隊設計了無中斷移轉的步驟，克服了兩代設備相同、介面不足、流量龐大等困難，並以 TWAREN2 的優點及移轉規劃爭取到用戶們的認同，令 TWAREN2 得以實現，一般的使用者甚至未察覺到每天使用的網路已經歷了巨大的架構變化，事實證明了移轉設計的實現確實成功。

至本文撰畢，TWAREN2 已歷經半年以上的運行，維運報告亦顯示，儘管設備失能與線路中斷的事件仍不斷發生，但會影響用戶的服務中斷事件數卻已大大降低，足可證明 TWAREN2 之實現確實對 TWAREN 之維運有極大助益。

參考文獻

- [1] TANet 台灣學術網路：
http://www.edu.tw/EDU_WEB/EDU_MGT/MO_ECC/EDU0688001/tanet/environ/introduction.html。
- [2] TWAREN 台灣高品質學術研究網路：
<http://www.twaren.net/>。
- [3] <http://www.ieee802.org/1/pages/802.1Q.html>。
- [4] TANet 骨幹流量圖：
<http://mrtg.twaren.net/mrtg/wmap/tanet.htm>。
- [5] Cisco 7609 Router：
<http://www.cisco.com/en/US/products/hw/routers/ps368/ps367/index.html>。
- [6] Cisco 12816 Router：
<http://www.cisco.com/en/US/products/hw/routers/ps167/ps5632/index.html>。
- [7] Cisco ONS 15454：
<http://www.cisco.com/en/US/products/hw/optical/ps2006/ps2008/index.html>。
- [8] OSPF, Open Shortest Path First IGP
<http://www.ietf.org/html.charters/ospf-charter.html>。
- [9] <http://www.ieee802.org/1/pages/802.1D.html>。
- [10] Cisco Ether-Channel
<http://www.cisco.com/warp/public/473/4.html>。
- [11] Cisco Catalyst 6509 Switch
<http://www.cisco.com/en/US/products/hw/switches/ps708/ps711/index.html>。