

TWAREN IPv6 網路效能量測平台建置

¹ 楊哲男, ² 古立其

¹ 財團法人國家實驗研究院國家高速網路與計算中心

E-mail: yangcn@nchc.narl.org.tw

² 財團法人國家實驗研究院國家高速網路與計算中心

E-mail: lku@nchc.narl.org.tw

摘要

網路維運者對網路定時進行效能量測，可偵知網路品質改變，幫助釐清網路障礙發生位置，使故障得以迅速排除的重要機制，然而大部分僅限於 IPv4 之監控。IPv6 網路目前已建置於學術研究網路及部分之商業網路，隨著物聯網的蓬勃發展，IPv6 之需求也越來越大，針對 IPv6 骨幹進行之網路效能量測已刻不容緩，本文將敘述 TWAREN 利用 perfSONAR 建置骨幹 IPv6 效能量測之方法。

關鍵字: IPv6; TWAREN; 效能量測; 網路效能。

連線，連接全球研究網路，提供 IPv4、IPv6、VPLS VPN、SSLVPN 及 SDN 連線應用服務等。

由於 TWAREN 網路管理之設備眾多，當設備故障、接頭老化、管溝施工等眾多物理性因素，甚至是人為設定失誤、網路攻擊等人為因素影響，一個網路在建置完成之後，品質隨時都在變化。為了在第一時間發現網路問題，進行修護，以確保最佳的網路品質，網路效能量測是網路維運中重要且不可或缺的一環。

1. 前言

在物聯網即將爆發的時代，未來的資訊服務將透過各式有線或無線網路進行訊息傳遞，再透過遠端大資料數據分析進行智慧處理來創造加值服務，若透過 IP 管理實現網狀拓模架構，必然伴隨而來龐大的 IP 位置。隨著 IPv4 已經不敷使用，IPv6 將會成為往後物聯網的重要標準，在台灣之學術及研究網路皆已提供 IPv6 之骨幹服務。國家高品質學術研究網路 (Taiwan Advanced Research and Education Network; TWAREN) [1] 已於 105 年 10 月升級骨幹頻寬至 100G，提供一國際級實驗研究網路平台。目前國內骨幹網路有五個核心主節點及十二個區域網路中心(GigaPOP)，主節點分別位於中央研究院、台北國家教育研究院、國網中心新竹本部、國網中心台南分部，以及國網中心台中分部。GigaPOP 部分包含臺灣大學、中央大學、清華大學、交通大學、中興大學、暨南國際大學、中正大學、成功大學、中山大學、東華大學、宜蘭大學及政治大學，以 20Gbps 至 50Gbps 線路連上主節點，實際之骨幹線路如圖 1 所示。各連線單位(大學院校/研究單位/縣市網路中心) 以 1Gbps 以上連結區域網路中心，連線單位亦可透過都會型網路連上 TWAREN GigaPOP，東部宜蘭及花蓮地區分別與台北及台南核心點連接，藉東西部串連構成環島網路並以 20Gbps 國際



圖 1: TWAREN 100G 網路架構圖

測量網路效能可提供管理者或使用者清晰的網路資訊以協助排除故障，當遇到網路障礙時能更快釐清問題的成因[2]，更有效率的解決網路面臨的瓶頸。一個網路效能觀測網之結果的精確度就取決於它所用來測量網路效能的工具以及維護管理人員之校調方式及經驗。在本文中，我們將介紹我們如何利用放置於 TWAREN GigaPOP 之伺服器建置 Server based 的 IPv6 網路效能量測架構及如何利用 Internet2 所開發之開放原始碼(Open Source)建置 IPv6 網路效能量測平台，方便 TWAREN 維運工程

師能透過此工具發覺可能的網路問題。在第二章節中我們將介紹網路效能量測的一些基本概念及幾種影響網路效能的重要因素。在第三章節中我們將介紹我們所使用的網路效能量測工具，並於第四章節中說明我們的設計架構及方法。

2.網路效能量測

2.1.檔案傳輸之效能影響參數

傳統的科学應用，例如從資料擷取點傳輸海量資料至資料分析之設施會利用網路檔案工具來傳遞資料，這些傳輸工具有一個共通的特性，就是預設會選用傳輸控制協定（Transmission Control Protocol, TCP）來做為網路傳輸層之協定。TCP 是一種連接導向的、可靠的、基於位元組流的運輸層通信協議。有許多的原因皆會影響 TCP 的效能，其中最主要的因素有三個[3]：

- Buffer/window size：

TCP 使用 CWND 來決定同一時間內有多少封包數可以同時被傳送，也就是說月大的 CWND 數，可以得到越高的性能。TCP 利用一些演算法，透過 Slow Start 及避免封包擁塞之 congestion avoidance 來決定 CWND 之大小。TCP socket 程式利用 window size 及 buffer size 來決定 CWND 之大小，因此最佳的 buffer size 為 BDP 之兩倍。也就是說，若作業系統的 window size 設定太小或是應用程式的 buffer size，將嚴重影響整個 TCP 的傳輸效能。

- 網路封包的延遲(Latency)

根據 BDP 的算法[4]，它與網路封包的延遲時間息息相關，也就是說當兩點之間的 RTT 越長，若想要得到叫好的網路傳輸效能，它所需要的 BDP 將越大，若是作業系統的 window size 設定以及應用程式 buffer size 的設定沒做調整，將會得到非預期的網路傳輸效能。當一個頻寬很大、RTT 很長的網路，我們稱為 Long Fat Network，例如台灣連到美國的跨洋網路。在 TCP 表頭(header)裡，由於只有 16 個 bits 大小的位元數來規定 window size 的值，因此最大的值為 2 的 16 次方，也就是 65536Bytes。根據 BDP 的算法，這一個值完全無法滿足 Long Fat Network 的需求。所幸在 RFC1323 裡有規範，透過啟動 TCP window scale option，可以利用其他的位元數來擴大 window size 的大小，避免在 Long Fat Network 下使用小的 BDP 來限制網路的傳輸效能。

- 網路封包遺失(Packet loss)

當傳輸過程發生封包遺失的現象時，根據 TCP 的機制，它將會減慢傳輸的速度，若是封包一直的遺失，則網路傳輸速度會越來越慢，直到封包遺失的現象消失。若是在 RTT 比較長的狀況下，又發生了封包遺失的現象，情況就更加的惡化了，因為 RTT

較長的網路，其所需的 BDP 較大，但當發生封包遺失時，它會慢慢的增加 window size，但因為封包遺失的關係，整個 window size 值無法提昇，且封包必須重傳，並且降速，造成效能之惡性循環，RTT 也會因此變得更長。

2.2.網路效能量測之目的

為了能準確的了解骨幹網路即時之狀態，我們利用一般伺服器來建置量測平台，我們稱為 Server based 之網路效能量測。將安裝 perfSONAR 工具程式的伺服器安裝至 TWAREN 各核心節點及各 GigaPOP 節點，直接接於骨幹路由器之下。由各台伺服器通過 TWAREN 骨幹網路彼此 Fully mesh 互相進行量測。量測項目包含 ping (來回延遲時間)、One Way ping (取得單向封包遺失率)，必要時亦可進行頻寬量測。為了使 One Way ping 能得到正確的單向封包延遲時間值，各台伺服器需經由 NTP 網路校時的方式，維持各台伺服器時間一致。Server based 網路量測的結果預設會自動儲存在各個 perfSONAR 的量測資料庫中。之後以自行開發的程式每 5 分鐘向各台伺服器索取最新量測結果的方式將量測結果集中，並修正為 TWAREN NOC 監控及告警所需的格式後統一寫入 TWAREN NOC 網路監控資料庫中。

3.效能量測相關工具介紹

本節將說明我們所使用到的量測工具及技術：

3.1.perfSONAR 介紹

歐洲高速研究暨教育網路 GÉANT2，提供最新的網路基礎設施，並致力於在各方面提高教育科技研究網路技術。為了加強國際合作與交流，GÉANT2、Internet2 和 ESnet 合作開發 perfSONAR (Performance Focused Service Oriented Network Monitoring Architecture) 網路效能監控基礎設施。perfSONAR 是介於效能監控與視覺化應用程式之間，運作於中間層 (Intermediate Layer)，這一層為通訊協定主要在處理數個網路間效能測量的交換。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：

- 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。

3.2.perfSONAR Toolkit

perfSONAR Toolkit 集合了 perfSONAR-PS 所開發的各種軟體 [5]，原生於 CentOS 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 Apache、MySQL、NTP 及 esmond 等服務，目前的穩定版本為 3.5.1。

3.3.OWAMP

OWAMP(One Way Active Measurement Project)是用來測量兩點之間的單方向延遲時間(one way delay)的一個軟體工具；如圖 2 所示，它是採用 Client-Server 的架構來作測量：當作 Server 的那一端先啟動 owampd 以等待另外一端的 Client 連過來做測試，然後當作 Client 的那一點以指令“owping Server_IP”來啟動作測試的動作，數秒後，one-way delay 的平均值的結果就會被顯示在 Client 端的螢幕上了，相當的簡單易用；若想要一個指令就測出兩點之間的“往”和“返”兩個分別的 one-way delay 的值，則兩點都需要同時先啟動 owampd，然後其中一點發起“owping Server_IP”的指令，則數秒後，“往”和“返”兩個分別的 one-way delay 的平均值都會被顯示在這點的螢幕上了，結果中會包含著“往”和“返”兩個 one-way 的測量值，而每個測量值中包含來源點 IP、目的地點 IP、封包遺失率、one-way delay 平均值、精確度等資料，這就是我們最常採用的方式。需要注意的一點的是，因為兩點間測量 delay 的準確度跟此兩點的系統時間是否有作同步有密切的關係，因此在將此兩點的 owampd 跑起來之前必須要先執行 Network Time Protocol (NTP) 的 daemon，即 ntpd。還有另外一點，因為 OWAMP 所出來的數據只是一堆數字而已，因此還需藉由統計繪圖軟體的幫忙，例如 RRDTool，將長期的 OWAMP 所得到的數據，封包遺失率及 one-way delay 平均值等統計並以圖表的方式顯示出來，以利相關人員閱讀方便。ping 只能測來回全路徑的 Delay、Jitter 值，但 OWAMP 則能測量出到底是出去還是回來的方向發生擁塞。

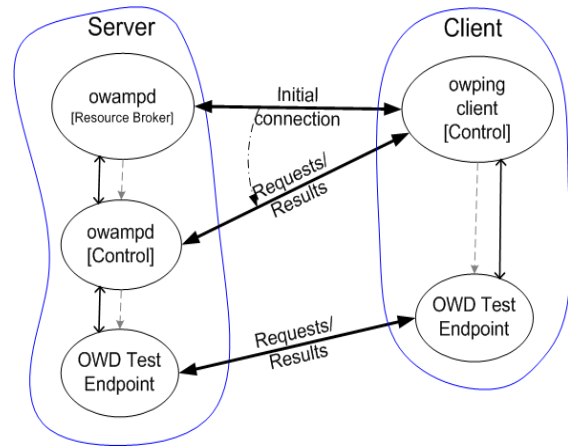


圖 2 OWAMP Control Flow

3.4.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，其中亦包含國網中心。後續 perfSONAR Toolkit 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。

4.TWAREN perfSONAR 架構及監控程式

在本章節中我們將介紹 TWAREN perfSONAR 之設計架構及所使用之工具，並介紹如何使用擷取程式獲取我們所需要之資訊，最後並介紹 TWAREN 網路效能之告警系統。

4.1.選擇使用 perfSONAR 之原因

利用各地的路由器(Router)兩兩相互執行 ping 指令並收集 RTT(Round Trip Time)及計算可用率，此監控模式我們可以稱為是 Router based 的網路效能測量方式。由於 TWAREN 網路完成建置的同時，各地的 Router 皆已備妥，因此使用 Router 進行網路量測不需要額外的設備。此外由於 Router 一般皆全時運作，需要的維護較一般伺服器少且可事先預期，因此量測的結果產生較為連續且穩定。然而因 Router 的設計會在設備忙碌時專注於封包 Forwarding，優先忽略回應 ICMP。因此 Router 間互 ping 易受 Router 負載影響，在網路仍然順暢，客戶端體驗正常時，即有可能在量測結果中出現封包遺失。易高估網路封包遺失率，也可能誤導網路品質的判斷結果。Server 間的量測封包對於 Router 來說屬於使用者封包，優先權與一般訊務相同。因此量

測封包是否遺失或被主動丟棄，可以等量反應正常訊務流經相同路段時的狀況。量測的結果一般較貼近使用者的感受，更能反映客戶端實際的網路品質。因此我們選擇了 Server based 之方式，並且搭配使用 perfSONAR。perfSONAR 除了一般的 ping 之外，還支援單向 (One Way) ping，能夠反映網路單向的品質，更有利於網路單邊障礙，例如接頭鬆脫、GBIC/SFP 模組故障的位置確認，有必要時亦可進行定時頻寬量測。

4.2.TWAREN 骨幹主節點

透過 perfSONAR 來安裝相關軟體，除了台北及中研院外，其餘三個節點均位於科學園區內，計有：中研院、台北(三峽國教院)、新竹：國網中心本部、台中國網中心分部及台南網中心分部。另外於台北科技大樓有一節點(舊網路之主節點)。每個節點建置兩台伺服器互為備援，兩台伺服器為高可用(HA)架構，每台主機之 IPv6 Address 皆為固定 IP 並註冊主機名稱以方便管理。整體之架構如圖 3，以中研院為例，兩台伺服器皆執行 IPv4 及 IPv6 網路封包遺失、RTT 即可用率之計算，收集到之資料會回傳至 TWAREN 維運中心資料庫並產生報表，格式如表 1。各點之詳細之架構如圖 4。

表 1: TWAREN 報表

TWAREN POP Packet Loss Rate (%) - 2017-03-14 00:00:00 - 2017-03-14 23:59:59 --																
Src/Dst	TP	HC	TC	TN	NTU	SINICA	NIU	NDHU	NSYSU	NCKU	CCU	NCNU	NCHU	NTHU	NCTU	TRTC
TP	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HC	0	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TC	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0	0
TN	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0
NTU	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0
SINICA	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0
NIU	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0
NDHU	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0
NSYSU	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0
NCKU	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0
CCU	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0
NCNU	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0
NCHU	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0	0
NTHU	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0	0
NCTU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	0
NCU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-
TRTC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NCCU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

4.3.TWAREN GigaPOP

TWAREN 13 個 GigaPOP 主要分佈於各縣市主要國立大學及研究機關，負責該地區網路管理等任務，包含台灣大學 NTU、宜蘭大學 NIU、政治大學、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 及中山大學 NSYSU。每個節點亦建置兩台高可靠度架構之伺服器互為備援，整體之架構如圖 3，以中研院為例，詳細之架構如圖 4。

4.4.TWAREN 國外節點

除了國內 18 個節點之外，TWAREN 在美國也建置了 3 個節點以方便與國外研究網路建立網路互連，因此我們為了監控跨洋線路之品質，亦在國外 3 個節點建置了 perfSONAR 監控點，包含了：洛杉磯 LA、芝加哥 CHI 及紐約 NY，目前國外節點僅監控

IPv4 之網路效能，未來亦將納入 IPv6 之監控。

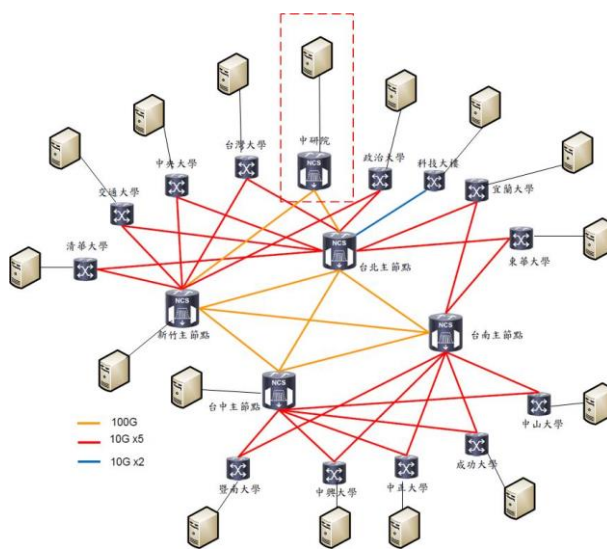


圖 3: TWAREN 效能量測網路架構圖

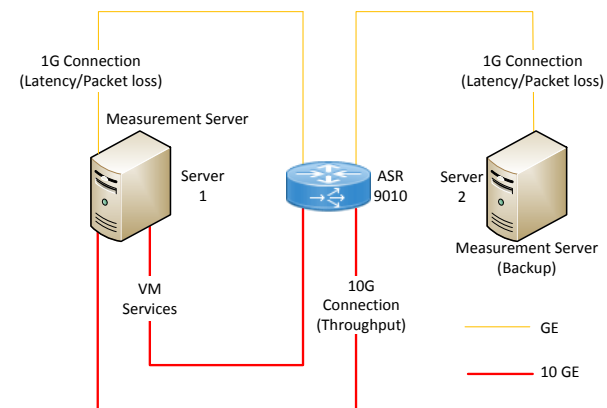


圖 4: TWAREN 100G 網路架構圖

4.4.量測及告警方式

於 18 台主機安裝完成 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，其測試參數片段資料如圖 5 這樣方便管理者只要設定一次，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 OWAMP 及 PingER 等 daemons，開始做 fully-mesh 之互相監控的。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裡，如圖 6 所示。使用者或管理者可從每台主機之各別網頁上選取有興趣之時間軸放大圖形做進一步之觀察，亦可透過統一管理之查詢方式做管理，以方便管理，因此需要製作一個統一管理之入口網站，我們稱為 TWAREN Dashboard，如圖 7 所示。

```
{
  "members": {
    "members": [
      "trtc-ow6.twaren.net",
      "tp-ow6.twaren.net",
      "hc-ow6.twaren.net",
      "tc-ow6.twaren.net",
      "tn-ow6.twaren.net",
      "sinica-ow6.twaren.net",
      "niu-ow6.twaren.net",
      "ntu-ow6.twaren.net",
      "ndhu-ow6.twaren.net",
      "ncu-ow6.twaren.net",
      "nthu-ow6.twaren.net",
      "nctu-ow6.twaren.net",
      "nchu-ow6.twaren.net",
      "ncnu-ow6.twaren.net",
      "ccu-ow6.twaren.net",
      "ncku-ow6.twaren.net",
      "nsysu-ow6.twaren.net",
      "nccu-ow6.twaren.net"
    ],
    "type": "mesh"
  },
  "parameters": {
    "force_bidirectional": "1",
    "bucket_width": "0.001",
    "ipv6_only": "1",
    "packet_padding": "0",
    "sample_count": "600",
    "packet_interval": "0.1",
    "type": "perfsonarbuoy/owamp"
  },
  "description": "Loss Test Between TWAREN Domestic Latency New Hosts for IPv6"
},
```

圖 5: 相關監控統一 json 檔案設定參數

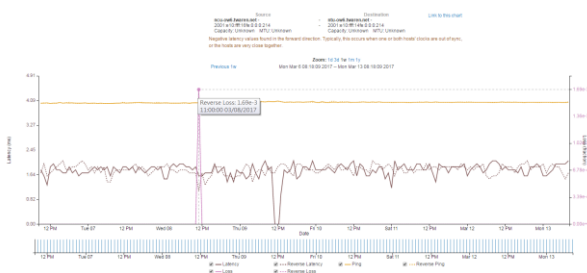


圖 6: OWAMP 查詢之曲線圖

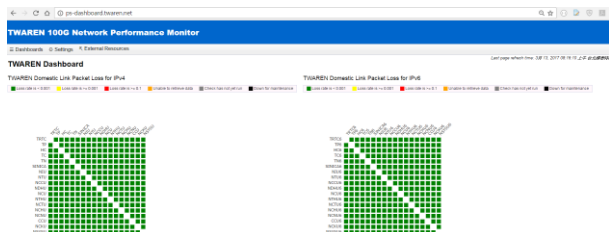


圖 7: TWAREN Dashboard 網站

目前我們透過國內 18 台主機之兩兩互相監測，共會收集到 306 筆資料，18 台主機會先統計每 30 分鐘內之 one-way 延遲時間之平均值及封包遺失率之數值再回傳給統一管理之主機，透過矩陣式的顯示方法，讓使用者及管理者透過網路觀察全部的量測資訊。我們設定了幾個臨界值，並利用顏色來表現封包遺失率之影響情形，其中綠色代表封包遺失率低於百分之 0.001，黃色代表封包遺失率介於百分之 0.001 至百分之 0.1 之間，紅色代表封包遺失率大於百分之 0.1，橘色代表管理主機無法收到各個節點的回傳值，灰色代表尚未啟動資料收集之節點，黑色代表伺服器正在進行維護狀態。

結論

IPv6 的網路效能監控常常都被網路管理者忽略，再越來越多 IPv6 應用的同時，一個品質良好的網路才可提供真正有效率的網路頻寬，而如何管理網路品質成為各個網路管理者的重要目標，透過健全的量測規劃以及測試工具來確保網路上的設備及線路能運作正常亦是個重要課題。使用 perfSONAR 監控程式透過長期的網路效能量測並透過我們所設計的自動告警系統，它可發現比較難發現的軟故障(soft failures)。當骨幹發生了軟故障時，它並不會出現無法傳輸的問題，但卻會出現傳輸效能降低的問題。透過更多的資料收集，透過大資料之技術，來找出真正發生問題之原因。也因此若有更多的區網或是單位能夠將資料收集在一起，將可利用更多的資料分析來找出更精準之問題所在點。

參考文獻

- [1] Taiwan Advanced Research and Education Network (TWAREN), <http://www.twaren.net/>
- [2] Jason Zurawski, Sowmya Balasubramanian, Aaron Brown, Ezra Kissel, Andrew Lake, Martin Swamy, Brian Tierney, and Matt Zekauskas, 2013 IEEE International Conference on Big data, perfSONAR: On-board Diagnostics for Big Data
- [3] 楊哲男，“海量資料傳輸基於負載平衡協定之比較”，Oct. 2012.，2012 台灣網際網路研討會, Oct. 2012.
- [4] speedguide, <http://www.speedguide.net/bdp.php>
- [5] perfSONAR toolkit, <http://www.perfsonar.net/>