

perfSONAR 在網路交換中心的佈署和應用

陳敏¹、李慧蘭²、古立其³

¹²³財團法人國家實驗研究院國家高速網路與計算中心

{¹minchen、²gracelee、³kuli}@narlabs.org.tw

論文摘要

為維護網路品質與可靠性，持續監控與即時警示至關重要。perfSONAR 作為成熟、可擴展且開源的量測工具，被廣泛應用於國際科學研究網路，本研究以國內 Formosa Open eXchange 交換中心為例，探討其在網路交換中心的應用。透過將 perfSONAR 量測節點設置在各成員網路邊際路由器，我們成功實現對經交換中心的網路品質監測。同時，新一代 perfSONAR 使量測資料的儲存與展示更具彈性。這為網路交換中心的營運管理提供了技術支援。

關鍵詞：perfSONAR、網路效能量測、網路交換中心、FOX

Abstract

To maintain the quality and reliability of the network, continuous monitoring of the network and early warning should any problem arise is the most important foundation. As an advanced, extensible and open source measurement tool, perfSONAR is widely adopted by international research and education networks. Its architecture and capability are introduced in this paper, as well as its current deployment in an open network exchange point, the Formosa Open Exchange is elaborated.

Keywords: perfSONAR, Measurement, Formosa Open Exchange.

1 前言

perfSONAR[1](performance Service -Oriented Network monitoring Architecture)是以服務為導向的網路效能量測平台，廣泛被佈建在科學研究網路設施上，用來監控和測量網路效能品質，可協助識別網路可能影響網路效能問題，以確保網路高可靠度。2000 年代初期，學術研究網路發展漸趨成熟後，對於網路效能監控的標準方法需求變得明顯，初始由美國能源科學網路(ESNet)[2]、美國學術研究網路 Internet2[3]和泛歐學術研究網路(GEANT)[4]等多個組織合作發起。經過數十年的發展，perfSONAR 可以呈現跨域網路路徑上點對點的網路效能，例如：傳輸速率、網路延遲、封包遺失和網路抖動等。目前已成為全球網路工程師、管理員和研究人員必備的工具。且被佈署在超過 2000 各註冊主機和超過 300 各網域，分散在世界上 1000 個以上的組織。

國研院國網中心所維運的 TWAREN[5]研究網路和 FOX 網路交換中心，長期以 perfSONAR 平台

監控網路效能品質，以作為高可度網路服務的保證。perfSONAR 在研網社群多年的研發下，已演進至 perfSONAR 5.0。加強了異質網路環境中的實用性和通用性，也發展容器版本，易於雲端環境佈署。本文就新一代 perfSONAR 平台於 FOX 網路交換中心佈署的經驗進行探討。

2 網路效能量測

2.1 網路效能量測背景

在網路發展初期，由於技術尚未成熟，為了廣泛普及及長距離廣域網路傳輸並以相對低成本實現，TCP/IP 協定被設計為七層結構[6]。網路層(Layer3)和傳輸層(Layer4)，與網路傳輸的軟體控制密切相關，由作業系統核心程式統一負責，以降低硬體設計的複雜度和成本。實體層(Layer1)和資料鏈結層(L2)則與硬體設計相關，保持較簡單，而應用層(Layer7)提供給網路應用的介面得以統一，使得應用開發更容易，更容易普及。這一設計為網路帶來了巨大成功，塑造了今日網路的蓬勃發展。然而，這種分層設計導致應用程式及使用者難以了解網路的實際運作狀態，增加了故障診斷的複雜性。

在這樣的背景下，主動式網路效能量測技術應運而生。以應對上述挑戰，透過主動發送量測封包並回收特定資訊，這種技術使量測發起端能夠在不依賴底層網路狀態資訊的情況下，判定網路品質和問題。進行網路品質量測時，需要傳輸的兩端共同配合。例如，ping 和 traceroute 通常需要量測彼端在收到量測封包後，回傳特定的 ICMP 訊息。這些功能通常已內建在現今流行的各種作業系統中，由作業系統核心負責，因此量測雙方一般不需要額外交裝程式。然而，進行更複雜的網路量測，如單向 ping 或頻寬量測，則需要量測的兩端都安裝相同的量測程式並搭配特定參數。針對不特定的目標主機通常只能進行基本量測，如 ping。當基本量測無法滿足查找網路障礙的需求時，進行較複雜的量測則需要在測試網路的兩端安裝配套的量測軟體，以進行進一步的評估。

2.2 perfSONAR 架構

perfSONAR 是一套整合開放原始碼工具平台，是由美國能源科學網路(ESNet)、美國學術研究網路 Internet2、印第安納大學(Indiana University)、密西根大學(University of Michigan)和泛歐學術研究網路(GEANT)等單位所共同參與執行，集合了知名的網路效能量測工具。我們以最新的 perfSONAR 5.0 架構圖，六大模組來說明其元件

功能(如圖 1 所示)：

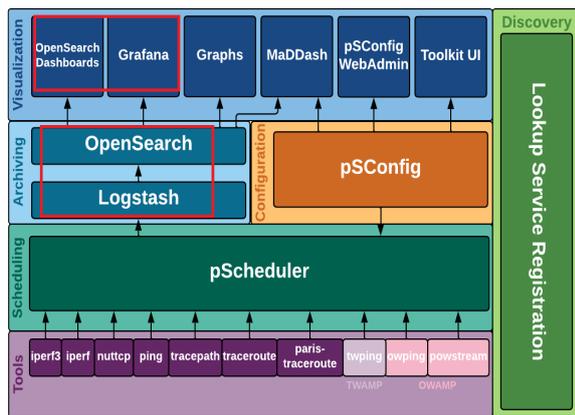


圖 1. perfSONAR5.0 架構圖

- (1) 量測工具(Tools)：提供的測項相當多元，包含有以下項目
 - iper3、iperf 和 nuttcp 網路有效頻寬量測，用於量測網路傳輸速率。
 - ping 用於檢測網路可達性。
 - tracepath、traceroute 和 paris-traceroute 用於網路路徑追蹤。
 - TWAMP (Two-Way Active Measurement Protocol)[7] 用於量測封包遺失和雙向延遲，套件中的 twping 是用於客戶端工具，相較 ping 而言，提高了準確性。
 - OWAMP (One-Way Active Management Protocol)[8] 用於量測封包遺失和單向延遲，套件中的 owping 是單一短期測試命令，powstream 是長期後台執行的命令。
- (2) 排程器(Scheduling)：可依量測的需求讓各量測節點之間定時排程，進行必要的網路量測，當一端主機要進行量測排程需執行 pScheduler Server 服務，並透過 API 呼叫客戶端的量測工具執行。
- (3) 量測資料歸檔(Archiving)：作為量測資料長期儲存庫，提供視覺化模組使用。在 perfSONAR 4.0 版本由 Esmond 套件實現，搭配底層 cassandra 和 postgresQL 資料庫。Esmond(ESnet Monitoring Daemon)是由 ESnet 所開發的資料搜尋引擎，用來收集、儲存和分析大量時序型的資料，主要提供給 MaDDash 視覺化使用。perfSONAR 5.0 捨棄 Esmond，改以 OpenSearch 和 Logstash 取而代之。
- (4) 設定檔配置(Configuration)：pSConfig 模組利用 JSON 文件建立自動化的大規模量測節點管理。也提供 pSConfig WebAdmin，搭配 web 介面定義和發佈設定檔配置。
- (5) 視覺化呈現(Visualization)：支援之模組有 OpenSearch 儀表板、Grafana、Graphs、MaDDash(Monitoring and Debugging Dashboard)和 pSConfig WebAdmin。透過執行 pSConfig 的設定檔配置，以搜尋引擎

OpenSearch 獲取其他量測節點的拓模資訊和量測數據。

- (6) 探索設定 Discovery：垂直分布與上述各層相關的是 Lookup Service Registration Daemon。量測節點的 lookup service 會註冊和維護網卡硬體是否執行、各式 perfSONAR 的元件是否正常運行。

新一代的 perfSONAR 5.0 版本最重要的改革是後端的量測資料儲存檢索系統 Measurement Archive (MA) 的搜尋引擎從 Elastic Search 改為 OpenSearch 及 Logstash。優勢是 OpenSearch 是更開放的生態系統，具更好的擴展性和彈性，Logstash 可以用於數據收集、轉換和發送到 OpenSearch 中進行搜索和分析。OpenSearch 是開放原始碼的搜尋引擎，也包含 OpenSearch Dashboards 資料可視化儀表板。此軟體始於 2021 年，作為 Elasticsearch 與 Kibana 的分支，主要由 AWS(亞馬遜雲端運算服務)開發。

perfSONAR 是分散式架構的量測系統，依量測範圍進行大規模安裝，這些量測節點(testpoint)會依據排程定期進行測試，目前也提供 docker 容器映象檔，讓佈建 perfSONAR 量測節點增加了一個極為便利且易於管理的途徑。perfSONAR 量測節點(testpoint)容器化的量測工具擁有以下優點：

- 相比於採用虛擬機器 (VM) 的量測節點，容器化的量測軟體啟動快速許多，且方便動態在啟動時進行參數設定，利於搭配網路自動化軟體，納為自動化流程的一環。
- perfSONAR 需使用特定版本的 Java、資料庫等元件。所在主機如同時擔任其他任務 (例如 www 伺服器或資料庫) 往往會造成上述系統元件的版本衝突。容器化後量測軟體可自帶適用的元件版本，與所在主機軟體環境隔離，可同時安裝互不衝突。利於網路自動化的佈建及主機資源共用。

perfSONAR 5.0 採用的架構對於量測計畫的派送及量測資料的儲存具備高度的彈性。因此使用在大量佈署的場合時，可以在各量測節點安裝套件 (或 docker image)，共同訂閱同一量測計畫發佈節點的 JSON 文件量測設定，即能依同一份量測計畫的指示分別進行量測。量測的結果可以各自寫入本地端的儲存資料庫，亦可以依量測計畫的設定，共同集中寫入同一資料儲存節點 (Measurement Archive; MA)。此特點對於擁有大量量測節點的資料需要收集的場合非常方便。同時 perfSONAR 支援將量測結果寫入多個目的資料庫，產生備援。可以避免在單一儲存節點故障時，量測結果遺失的不便。

2.3 perfSONAR 案例

perfSONAR 量測平台廣為國際學術研究網路使用。台灣知名的研究網路 TWAREN 於中研

院、台北、新竹、台中、台南設置 5 個骨幹核心主節點，於全台 12 所大學設置區域網路中心 (GigaPOP)，基於高度備援性考量，我們於每各節點均建置兩套 perfSONAR 量測平台。如圖 2 的 MaDDash 儀表板所示，為 TWAREN 節點上點對點的量測紀錄，點選圖中任一點即可檢視目前量測結果，以圖 3 為例說明，自台中主節點 (TC_TWAREN) 至新竹主節點 (HC_TWAREN)，紀錄網路節點長期的效能監控數據，可檢視 1 天至 1 年的效能數據，例如圖中封包遺失率折線於 4/18 至 4/24 日凸起高於 0% 數次 (粉色線)，且於 4/23 下午 3 點是一周來的高峰期，顯示數據為 5.46e-5，是網路效能衰減的跡象，需進行潛在問題識別，以維持高度網路可用率。

圖 2. TWAREN 的 PerfSONA 儀表板

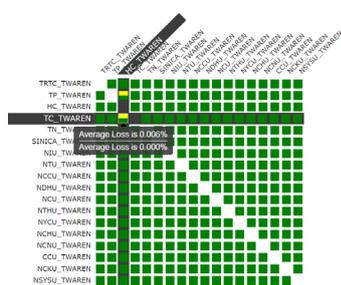


圖 2. TWAREN 的 PerfSONA 儀表板

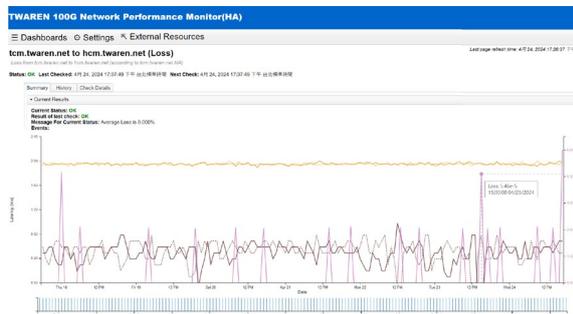


圖 3. TWAREN 的長期監控效能品質圖表

3 佈署於網路交換中心

3.1 FOX 網路交換中心

我國政府於 110 年起推動數位國家創新經濟發展方案，為加速疫後的數位轉型，擴大國家數位基礎建設的需求，持續投入強化數位韌性相關建設，其中建置公共服務網路交換中心至關重要，透過公共服務網路交換中心整合現有公共網路、提升各公部門跨域傳輸效率及備援能力。國研院國網中心配合我國數位發展政策於 111 年完成建置公共服務網路交換中心，並命名為福爾摩沙開放網際網路交換中心 (Formosa Open eXchange; FOX)[9] 簡稱為 FOX 交換中心。如圖 4 所示，國內四大公共服務網路：GSN 政府服務網路[10]、TANet 台灣學術網路[11]、TWAREN 研究網路和 ASNet 中研院研究網路[12] 分別以 20G 或 100G 頻寬接入 FOX 交換中心，進行網路傳輸高速直連交換，提升公共服

務網路跨網傳輸效率。

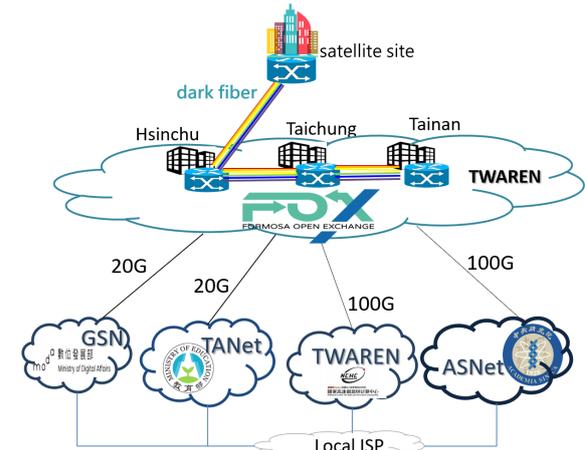


圖 4. FOX 架構圖

3.2 佈署 perfSONAR 於網路交換中心

對於網路交換中心而言，它本身在架構上相當於一個網路交換器，並不配置 IP 位址於交換介面上。因此，無法在交換中心內部設置主動式網路量測節點。為了測量各成員網路經過交換中心後的真實網路品質，必須將量測節點設立在參與交換中心的各成員網路的邊緣路由器之後。因此，我們將新一代 perfSONAR 佈署於 FOX 四大公網環境中，量測節點分別位於 國網中心 TWAREN、教育部 TANET[7]、中央研究院 ASNet[8]，以及政府服務網路 GSN[9] 上。在安裝 perfSONAR 量測節點前，需要使用 traceroute 指令驗證各量測節點 IP 之間的路徑是否經過 FOX 交換中心。這是因為在 FOX 交換中心設立之前，各公共服務網路已經存在其他交換路由的機制。確認各量測節點之間的路徑經過 FOX 交換中心後，其量測結果才能作為 FOX 交換中心的網路品質參考。

perfSONAR 提供 Maddash 圖形化套件。Maddash 套件發展已久，其功能完整，CPU 資源耗用節省。Maddash 使用方格圖呈現大量量測節點彼此兩兩互量的全貌總覽，如圖 5 所示。方格圖中，位於左邊縱向排列的是量測的來源端，位於上方橫向排列的是目的端。方格圖以顏色做為量測值狀況嚴重程度的區別。綠色為封包遺失值小於 0.001%，黃色為 0.001% 至 0.1% 之間，粉色則為大於 0.1%。從該視覺化面板立即可得知整體網路品質狀況。

圖 5 是 perfSONAR MaDDash 儀表板，上半部分顯示使用 perfSONAR 內的 TWAMP 工具量測的封包遺失率和雙向延遲。當檢視 GSN 網路時，發現以 GSN 為起點到其他網域的網路品質，包括橫向和縱向的遺失率均超過 0.1%。因此，需要關注 GSN 網域量測節點所在的網路效能問題。

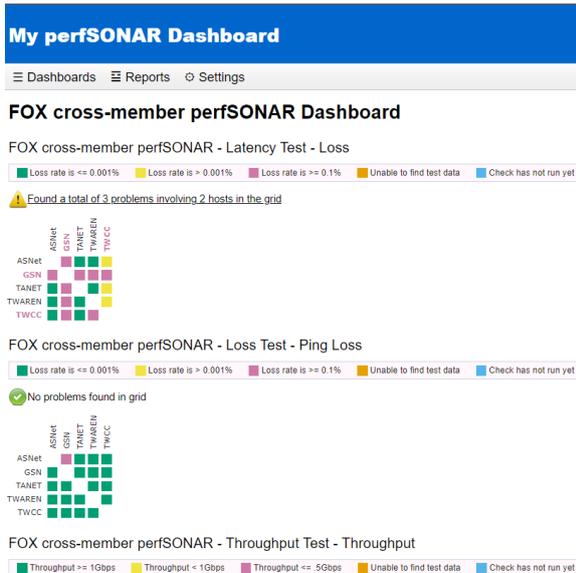


圖 5. perfSONAR 佈署於 FOX

新一代 perfSONAR 5.0 引入了 Measurement Archive (MA) 系統，可以解決過去量測節點需要儲存量測資料而導致資源不足的問題。MA 將量測資料集中儲存在中央位置，而量測節點則透過透過 psconfig remote 指令指定遠端 MA 進行資料的儲存。這樣一來，資料庫及網頁 Java 元件所佔用的記憶體不會隨時間增加，也不會導致資源不足而影響量測節點的回應速度。將儲存的數據透過 OpenSearch 搜尋引擎進行檢索，再利用 Grafana 進行圖形化展示，可以輕鬆地將 perfSONAR 整合至現有的監控方案中。管理者就可以在現有的 Grafana 環境中新增 perfSONAR 的 Panel，將其圖形化展示，進行監控作業。

在圖 6 中展示該段時間內的有效頻寬量測。網路傳輸速率介於 12.5G 和 15G 之間，平均封包遺失率在 0.03 和 0.04 之間。當傳輸速率下降或封包遺失率曲線飆升時，這表明網路可能存在問題，需要進一步調查和處理。(圖 6 未呈現異常)



圖 6. perfSONAR 量測結果呈現於 Grafana 平台

結論與未來工作

本研究介紹了 perfSONAR 在網路交換中心的佈署及其應用情況。通過將 perfSONAR 量測節點佈署在四大公共服務網路的邊際路由器之後，我們成功地實現了對於經過交換中心的真實網路品質的監測和量測。透過量測節點定期進行測試，我們可以及時地發現網路問題，提高了網路效能的管理和維運水準。同時，我們利用新一代 perfSONAR 的特性，如 Measurement Archive 和 Grafana 的整合，使得量測資料的儲存和展示更加方便和可靠。這為網路交換中心的營運管理提供了重要的技術支援。

未來，我們將繼續深入研究 perfSONAR 在網路交換中心的應用，進一步優化量測節點的佈署方式和量測策略，提高對網路品質的監測精準度和效率。同時，我們還將探索如何進一步整合其他監測工具和技術，以建構更加完善和強大的網路監測系統，為網路營運管理提供更多的支援和保障。此外，我們還將關注新興技術和應用場景，不斷拓展 perfSONAR 的應用範圍，以滿足不斷變化的網路管理需求。

參考文獻

- [1] The perfSONAR Project, "perfSONAR", <https://www.perfsonar.net/>
- [2] ESnet, <https://www.es.net/>
- [3] Internet2, <https://internet2.edu/>
- [4] GEANT, <https://geant.org/>
- [5] 台灣高品質學術研究網路 (TWAREN), <https://www.twaren.net/>
- [6] 古立其, 李慧蘭, 陳敏, "交換中心網路品質量測系統建置實例", TANET 2023 論文集, 台北, 2023 年 11 月.
- [7] Two-Way Active Measurement Protocol (TWAMP), RFC6308, <https://datatracker.ietf.org/doc/html/rfc6308>
- [8] Internet2, One-Way Ping (OWAMP), <https://software.internet2.edu/owamp/owping.man.html>
- [9] 福爾摩沙開放網際網路交換中心 Formosa Open eXchange, <https://www.fox.net.tw/>
- [10] 臺灣學術網路 (Taiwan Academic Network; TANet), <https://noc.tanet.edu.tw/>
- [11] Academia Sinica Network (ASNet), <https://www.peeringdb.com/net/1356>
- [12] 政府網際服務網 GSN, <https://gsn.nat.gov.tw/>