台灣高品質學術研究網路整合式監控平台之設計與實作

曾金山　劉德隆
國家高速網路與計算中心
{tsengjs, tlliu}@nchc.org.tw

摘要

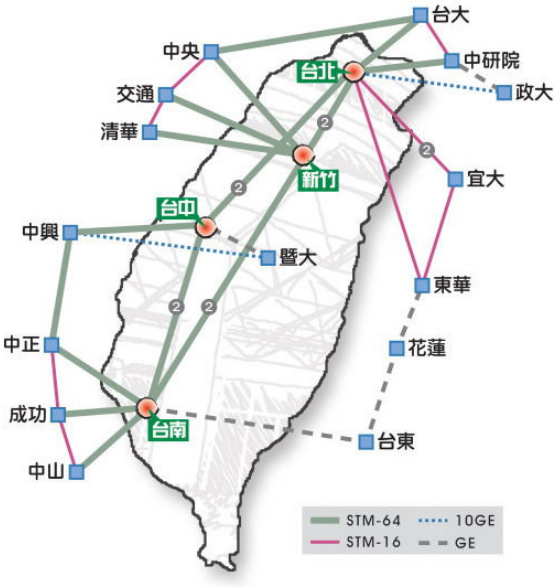
台灣高品質學術研究網路[8]為提供台灣產官學研界之骨幹網路，面對此一大型骨幹網路，單一類型之網路監控、管理軟體已不足以滿足網路維運人員之需求，為應付多樣化之網路設備及降低維運成本，本文以一網路維運人員角度設計並實作一套適用於大型骨幹網路之整合式監控平台並實際部署於網路監控管理上。

關鍵詞：網路監控、網路管理

1. 前言

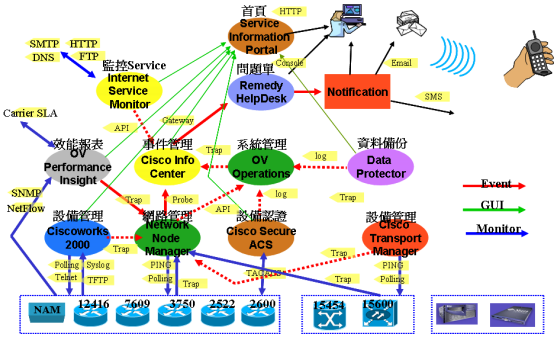
台灣高品質學術研究網路(TWAREN，TaiWan Advanced Research and Education Network)為提供台灣產官學研界之骨幹網路，提供高可靠度及高可用性骨幹網路為維運團隊之重要任務，如何能快速且有效偵測出網路問題所在是當前重要的課題。

圖表 1 為目前 TWAREN 國內實體網路架構圖，包含 4 個骨幹核心點(core node)及 15 個區網中心(GigaPOPs)，其中網路設備包含：光交換器、路由器、乙太交換器及伺服器等，面對此一多元化網路設備，傳統使用單一類型之網路監控、管理軟體組合而成之網管系統已不足以滿足網路維運人員之需求，本文以一網路維運人員角度設計並實作一套適用於大型骨幹網路之整合式監控平台。



圖表 1 TWAREN 國內實體網路架構圖

TWAREN 網路維運中心(NOC，Network Operation Center)於網路建置初期購置數套商用網管軟體以利監控管理骨幹網路，經過幾年來維運諸多問題逐漸浮現檯面，數套網管軟體各自擁有不同樣式之呈現介面，對於維運人員無疑會增加維運上負擔，此外，為提供使用者更完整之服務，客製化報表需求與日劇增，目前數套網管軟體各自採用不同類型儲存資料庫，且針對其資料庫存取亦有諸多限制，造成報表開發上之困擾。為解決上述問題維運團隊著手開發符合此目標之整合式監控平台並實際部署於現有網路監控管理上。



圖表 2 TWAREN 舊網管系統架構圖

圖表 2 為 TWAREN 現有運行中之舊網管系統架構圖，由數套商用軟體所組合而成，而現有網管系統所面臨問題：1.管理軟體非完整版，僅有局部功能、2.軟體授權、升級及維護費用昂貴、3.各軟體有其特定之資料儲存方式及呈現介面且無原始碼參考，難以進一步整合、4.軟體與硬體及作業系統相依程度高、5.國內廠商技術支援度不足。

2. 網管軟體

網管軟體為網路管理之重要工具，目前常見商用網管軟體如：Cisco CiscoWorks 2000[1]、HP OpenView[4]、IBM Tivoli 等，除有安裝硬體限制外，其維護成本亦不低，以 CiscoWorks 2000 為例，其針對每一模組功能採獨立販售方式，對於一大型網路維運單位而需完整功能者，成本是一大考量，此外，商業軟體針對監控設備、監控標的亦有限制，對於一擁有多樣化設備之網路，如 TWAREN，並無法滿足需求。

目前國際教育研究網路 NOC，如 GRNOC(Global Research NOC)[5]等，所採用的監控軟體大多為開放式原始碼，如 Nagios[3]、OpenNMS[9]、MRTG、OWAMP[6]等，TWAREN

維運團隊除網路維運致力於與國際研網接軌外，亦著重於網管系統開發，期望能降低維運成本並將相關成果提供於開發社群。

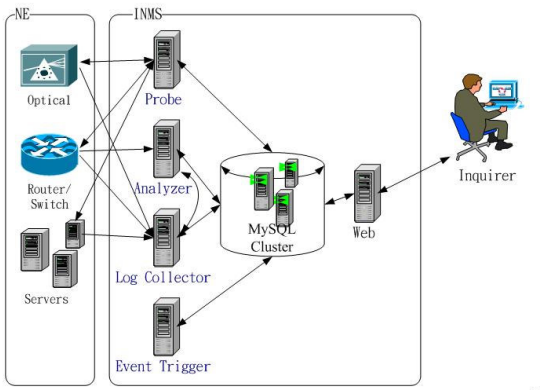
本文將於第三章節說明系統之設計概念，於第四章節說明系統實作方式，第五章節為系統展示。

3. 系統設計

本章節將介紹我們如何設計一套適合於大型網路之整合監控平台。

3.1. 系統架構

圖表 3 NE(Network Element)區塊代表網路設備，包含：光交換器、路由器、乙太交換器及伺服器等；INMS(Integrated Network Management System)區塊為本系統之架構圖及相關元件，本系統區分成四大模組，分別為：探測(Probe)、分析(Analyze)、紀錄收集(Log Collect)及事件處理(Event Trigger)，針對效能上考量，此四模組分別安裝於各自獨立之伺服器，除此外，針對資料完整性及可用性，本系統資料庫採用開放式原始碼 MySQL，架構上採用 Cluster 設計以達到高容錯(fault tolerant)及高可用(availability)性，監控人員則透過前端 Web 伺服器介面以掌握目前網路及設備狀態並可進行相關處理。

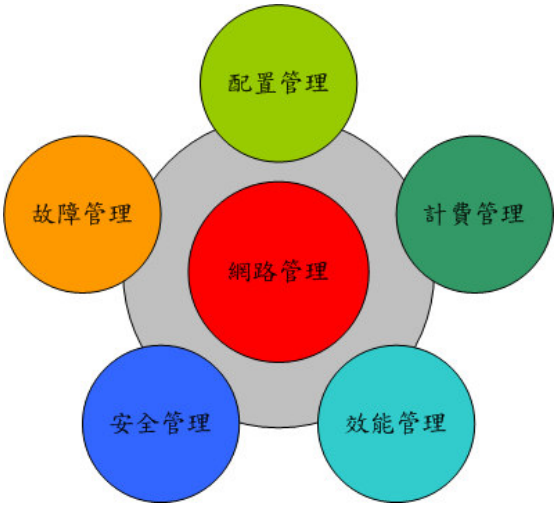


圖表 3 新網管系統架構圖

本系統目前所支援之監控設備包含：Cisco 光交換器 ONS 15600、ONS 15454、Cisco 路由器 GSR、7609、Cisco 乙太交換器 3750 及網際網路服務伺服器(e.g. DNS、FTP、Mail、Web 伺服器)等，監控範圍包含 TWAREN Layer1 至 Layer7 設備及服務。

3.2. 功能設計

本系統監控標的設計遵照國際標準組織(ISO)所提出之網路管理[10]架構，功能如**圖表 4** 所示分為五大部分 Fault Management、Configuration Management、Accounting/Administration Management、Performance Management、Security Management (FCAPS)，內容涵蓋實體網路至應用服務層。各監控標的之分類詳列如下：



圖表 4 FCAPS

A. Fault Management：

機房溫度監控，針對各 GigaPoPs 機房環境溫度狀態監控。

設備溫度狀態監控，針對各設備卡版溫度狀態監控。

介面狀態監控，針對各連線介面狀態監控。

服務主機服務狀態監控，針對各服務主機服務埠(port)監控。

B. Configuration Management：

設備設定異動監控，針對現有網路設備設定監控。

C. Accounting /Administration Management：

路由監控，針對連線單位路由內容監控、國內外 Peer BGP 狀態監控及重點路由(e.g. Google、Microsoft、Yahoo 路由)監控。

連線單位使用狀況監控，針對連線單位使用量及 IP 數監控。

D. Performance Management：

光網路品質監控，監控光路徑 SES(Severely Errored Second)、ES(Errored Second)、UAS(Unavailability Second)等。

RTT(Round-Trip Time)監控，監控路由器與路由器間的 RTT 及各 GigaPoPs 效能量測主機間的 RTT 監控。

設備效能監控，針對設備之 CPU、記憶體等使用率監控。

封包遺失率監控，針對骨幹介面及各連線單位介面封包遺失率監控。

E. Security Management：

阻斷服務攻擊(DDoS)[7]偵測，針對骨幹之大量特定應用埠連線進行監控。

針對監控標的訂定相關門檻(threshold)值，超過設定門檻值觸發對應處理動作。目前已完成處理

動作包含：郵件通知、簡訊通知、事件單紀錄，以連線單位路由內容監控為例，當比對出連線單位路由表異常，系統發信通知 NOC 維運人員及所屬單位負責人，並於問題追蹤系統(bug tracking system)開立事件單以利維運人員進行事件追蹤，事件發生後每兩小時再次通知。

4. 系統實作

本系統功能實作主要概念為利用 SNMP 介面及 CLI(Command Line Interface)進行相關資料取得。資料取得方式分成兩類：由探測(Probe)模組主動進行資料收集及紀錄收集(Log Collect)模組被動接收 Trap、Syslog 及 Netflow[2]等資料，探測模組將收集資料存於後端資料庫。分析(Analyze)模組透過後端資料庫及紀錄收集器將收到的資料進行分析並進行門檻值計算及異常判斷，計算出之結果則回存於資料庫。事件觸發(Event Trigger)模組則取出資料庫結果針對不同的策略(policy)做出回應。

探測模組根據設定收集相關 MIB OID(object identifier)值，如機房環境溫度感應器、設備卡版溫度感應器、設備介面狀態、設備效能狀態、設備 BGP 狀態、線路品質狀態等皆可透過 SNMP 方式取得資料；此外，針對無法透過 SNMP 方式取得的資料探測模組則透過 CLI 方式，如設備設定異動監控、連線單位路由內容監控、重點單位路由監控、路由器與路由器間 RTT 監控等資料，探測模組模擬 telnet 連線進入設備進行資料查詢，針對各服務主機服務狀態監控部分，探測模組模擬應用服務客戶端行為與服務主機進行資料交換以確認服務狀態，**圖表 5** 為探測模組 pseudo-code；探測模組每五分鐘進行監控標的資料收集。

```
foreach(targetHosts){
  foreach(monitorObjects){
  result = doProbe(targetHost,monitorObject);
  storeResult(result);
   }
 }
doProbe( ){
  if( monitorObject.type = { mib } )
    snmpget( targetHost,monitorObject );
  else
    cliProxy(targetHost,monitorObject);
 }
```

圖表 5 探測模組 pseudo-code

紀錄收集模組收集由網路設備所送至之 Trap、Syslog 及 Netflow 資料，並以網路檔案系統(NFS，Network File System)方式提供分析模組進行分析處理。

分析模組將存於資料庫之資料與基準資料進行比對，**圖表 6**為分析模組 pseudo-code，針對設備資源使用率、溫度、RTT 等其歷史資料以前七日平均方式計算出對應門檻值；針對流量等使用率其門檻值計算分成上下界，**圖表 7**為門檻值計算 pseudo-code，根據觀察，骨幹網路流量一天內約凌晨 2 點至早上 9 點較其他時段所呈現之特性不同，且星期一至星期五與週末六日所呈現之特性亦不同，因此本系統針對當天不同時段，當週不同星期，採用不同組之門檻值；針對路由部分則檢查相關路由器路由表以確認路由是否存在；針對阻斷服務攻擊部分，藉由路由器所輸出之 NetFlow 資料進行判斷，此一類型攻擊反應於 flow 資料中除有短時間多連線特性外，尚有連線大多為單方向，此類攻擊如 SYN flooding 等。

```
foreach(targetHosts){
  foreach(monitorObjects){
   if(type = route | APservice | int erfaceStatus | packetLoss){
      if(current.Value ≠ baseValue)
        Alarmbit = 1;
   }elseif(type = cpu | memory | temperature | RTT){
      if(currentValue ≥ baseValue * 110%)
        Alarmbit = 1;
   }elseif(type = traffic){
      if(currentValue ≤ lowBound * 90% || currentValue ≥ upBound * 110%)
        Alarmbit = 1;
   }elseif(type = config){
      if(currentConfig ≠ preConfig)
        Alarmbit = 1;
   }
  }
}
```

圖表 6 分析模組 pseudo-code

```
if(type = cpu | memory | temperature | RTT){
  (sum_{i=1}^{i=288} X_i) / 288;
 }elseif(type = traffic){
  lowBound = (sum_{i=24}^{i=108} X_i) / 84;
  upBound = (sum_{i=1}^{i=288} X_i - sum_{j=24}^{j=108} X_j) / 204;
}
```

圖表 7 門檻值計算 pseudo-code

事件處理模組則依據分析結果將異常事件依據事件處理原則進行通知、紀錄、封鎖。通知部分，由監控系統主動發出異常警告聲以提醒監控人員，郵件通知各權責單位，重要事件則簡訊通知主管單位；紀錄部分，結合問題追蹤系統開立事件單，以利維運人員追蹤事件；封鎖部分，於骨幹路由器以 ACL(Access Control List)方式進行封鎖，以保護受害設備及網路。

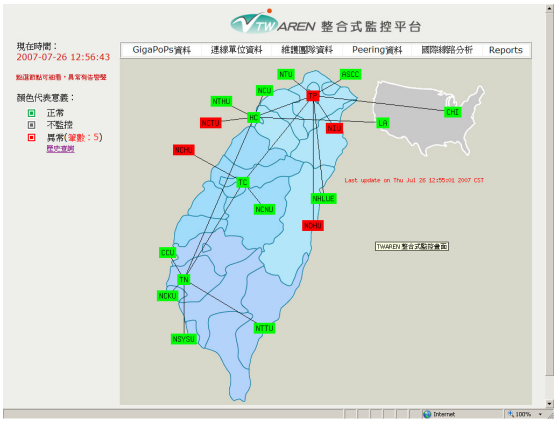
針對資料庫使用，本系統採用開放式原始碼 MySQL，架構上採用 Cluster 設計以達到高容錯(fault tolerant)及高可用(availability)性。

TWAREN 目前已全面啟用 IPv6 功能，因此對於 IPv6 網路之監控亦是重要，本系統目前已支援 IPv4 及 IPv6 網路監控。

5. 系統展示

本章節擷取部分目前已完成之成果。

爲使維運人員清楚了解網路狀況，本系統呈現方式採以階層式概念設計，以每一 GigaPoPs 為主體，第二層為 GigaPoPs 詳細狀況，包含所有相關之設備、連線狀態等。本系統呈現介面採用 PHP 撰寫，**圖表 8** 為呈現介面，當設備或連線狀況出現異常時，其結果會反應於 GigaPoPs，將所有異常事件皆歸屬於其所屬 GigaPoPs，於 GigaPoPs 呈現顏色狀態變化並發出告警聲響，將能使維運人員更容易了解網路狀況。

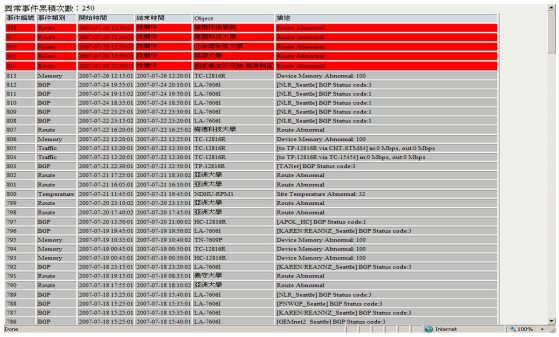


圖表 8 整合式監控平台

針對每一 GigaPoPs 狀況，可點擊該 GigaPoP 進一步了解詳細狀況，**圖表 9** 為某一 GigaPoP 之詳細狀態，包含：機房溫度、連線單位路由、設備狀態、線路流量、BGP 狀態等，當異常發生亦會同時呈現正常狀態基準值及相關所需處理資訊，讓維運人員能快速處理事件。

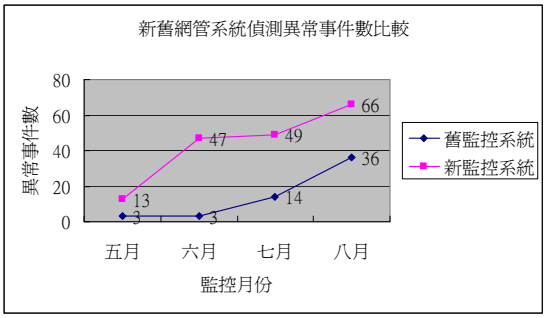
圖表 9 GigaPoP 狀態監控

針對持續中或已結束之事件，系統亦支援事件追查功能，如**圖表 10** 所示，方便維運人員隨時追蹤掌握事件狀況。



圖表 10 歷史資料查詢

圖表 11 為新網管系統與舊網管系統於 5 月至 8 月間偵測到的異常事件數比較表，兩套監控系統目前仍持續並行中，新監控系統較舊監控系統能監控更多的事件類別。



圖表 11 新舊網管系統成效比較

6. 結論與未來工作

本文初步已完成一套實際運用於 TWAREN 網路維運之系統，系統強調即時偵測及告警，除系統所使用之門檻值計算方法持續改進外，本系統亦持續開發管理功能，以方便維運人員能就單一介面解決問題，目前尚在開發之功能如：1.物件化圖形呈

現介面，目前系統呈現介面採用 PHP 撰寫，並無法達到物件化任意操控，未來期望能透過管理介面操控相關設備、2.自動化網路拓樸探測，由於網路設備日亦複雜，如何快速了解網路拓樸可有效降低維運人員除錯時間、3.身分認證權限區分管理介面，根據不同身分給予不同使用權限，管控及觀看所授權之設備、4.使用者自訂客製化報表，依據使用者需求自訂所需報表等，未來期望能將此套監控、管理平台提供給各 GigaPoPs 及連線單位使用讓網路管理能更加容易及透明化。

參考文獻

[1]Cisco, CiscoWorks2000, http://www.cisco.com/en/US/products/sw/netmgtsw/index.html
[2]Cisco, NetFlow, http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html
[3]Ethan Galstad, Nagios, http://www.nagios.org/
[4]HP, OpenView, http://managementsoftware.hp.com/
[5]Indiana University, Global Research NOC, http://globalnoc.iu.edu/
[6]Internet2, One-way Ping (OWAMP), http://e2epi.internet2.edu/owamp/
[7]Kevin J. Houle, George M. Weaver, ”Trends in denial of service attack technology”, Technical report, CERT/CC, Carnegie Mellon University, 2001.
[8]NCHC, TWAREN, http://www.twaren.net/
[9]OpenNMS, http://www.opennms.org/
[10]R. Boutaba and J. Xiao, “Network Management: State of the Art”, in Proceedings of IFIP World Computer Congress, pp. 127-146, 2002