

TWAREN 光網路上之光通道控制技術

林書呈 曾惠敏 張宜正 劉德隆

國家高速網路與計算中心

{daniellin, n00hmt00, andy, tlliu}@nchc.org.tw

摘要

隨著各種科學研究軟體 (e-Science Applications) 運算與資料量之增加, 帶來了高容量與頻寬保證之新型態網路傳輸需求, 亦帶動了光網路在學研界之發展。光網路之特性在於可動態建立使用者專屬之光通道, 使得無擁塞之獨享網路頻寬得以實現, 而光通道的配置是經由控制層 (control plane) 協定設定路徑中途經之光交換器, 本論文即針對現在國際間研發中之光控制技術進行探討, 並介紹 TWAREN 目前所試用之 UCLP 軟體; 同時簡述光交換器之程式語言介面, 以對於未來光通道之管理規劃出具體之方案。

關鍵詞: 臺灣高品質學術研究網路 (TWAREN)、光網路 (Optical Networks)、光通道 (Lightpath)。

1. 前言

近年來網路應用除了個人與企業之日常使用外, 科學研究軟體如高能物理與天文觀測等應用亦蓬勃發展, 這些為了特殊科學研究所開發之新一代網路應用程式其共同特徵為需要大量的網路頻寬傳遞即時性的資料。然而在傳統以 IP 路由交換為主的網路上, 所有的應用程式共享有限的網路頻寬, 除了傳輸速率的問題外尚有網路擁塞導致延遲的可能, 並無法滿足科學研究軟體所需, 而電信業者營運多年之光網路正可符合此要求; 光網路之運作原理自電話網路演進而來, 在通話前先建立起連結兩端之虛擬電路, 在電話中斷之前, 此電路上之通話頻寬為通話者與受話者獨享, 無須擔心延遲與擁塞之問題。在網路頻寬方面, 傳統的 SONET/SDH 光網路可以在 10Gbps 的頻寬中切割出 51.84Mbps/155Mbps 以上之光通道, 而在最近的 DWDM 光網路架構中, 更發展出自單一光纖切割出 32 個 10Gbps 頻寬供 SONET/SDH 使用之技術, 足以應付新一代網路應用程式之頻寬所需。

國際研網業已注意到此頻寬發展之趨勢, 於 1998 年, 加拿大的 CA*net 成為世界上第一個採用光網路架構之研究網路[1], 之後由於光網路產業景氣衰退導致光纖之佈建供過於求, 因此歐美各國研網皆有能租得光纖並進行光網路之規劃, 如美國的 NLR (National Lambda Rail) 網路[2], 底層以 DWDM 光網路架構為主, 提供了 Layer 1 光通道服務、Layer 2 交換服務與 Layer 3 的路由服務; 荷蘭

的 SURFnet 亦自 2000 年起採用光網路技術[3], 並於今年年初更新為全光網路之架構。國家高速網路與計算中心於 2003 年設計臺灣高品質學術研究網路時, 首度將光網路納入規劃[4][5], 以 SDH 技術為主提供光通道服務, 今年年底預計將更新架構, 將光網路與 IP 分封網路結合成單一骨幹, 以提供更完善的服務品質。

在光網路架構下, 光通道之配置需經由控制層協定設定路徑中途經之光交換器, 猶如電話網路中建立連線 (signaling) 之運作方式, 目前有 GMPLS (Generalized MPLS) 與 ASON (Automatically Switched Optical Networks) 等協定尚在制定中, 廠商亦成立 OIF (Optical Internetworking Forum) 進行各項實作驗證等工作, 因此現今各國皆有研發團隊開發適合研究網路使用之光網路控制技術[6][7]。

為了於 TWAREN 的光網路上有效地控制光通道之配置, 我們綜觀目前各國所開發之光網路控制技術, 加拿大所開發的 UCLP (User Controlled LightPath) 由於與 TWAREN 使用相同之設備且架構類似, 因此我們試用 UCLP 並協助將 SONET 程式碼移植為適合 SDH 使用之版本, 未來將進行光通道管理之研究。第二章簡介 TWAREN 整合光網路平台之架構, 第三章為各國光網路控制技術之研究心得, 第四章為 TWAREN 試用 UCLP 之經驗, 在第五章中, 我們研究了光交換器的程式介面, 並規劃光通道管理之方式。

2. TWAREN 整合光網路平台

今年度所規劃之 TWAREN 整合光網路平台架構如圖 1 所示, 結合了中華電信公司與亞太固網公司之線路, 以台北、新竹、台南與台中四主節點間之 STM-64 線路配合光設備構成骨幹網路; 各 GigaPOP 與主節點間之光設備再以 STM-64 線路構成星狀之都會網路。除此之外, 為了考慮故障時之備援, 各 GigaPOP 間另以 STM-16 或 STM-64 線路連接構成一環狀網路, 並加強東部宜蘭大學與東華大學間之連線。在光設備方面, 以思科公司之光網路系統 ONS (Optical Networking System) 光交換器為主, 由於台北、新竹與台南三主節點所需交換頻寬較高, 故採用背版頻寬高達 320Gbps 之 ONS 15600 設備, 而其餘節點均使用背版頻寬為 40Gbps 之 ONS 15454 設備。各節點間均能切割 STM-1 至 STM-8 不等之光通道, 然而 ONS 15600 並無乙太網路介面, 故在台北、新竹與台南三主節點另配置一

部 ONS 15454 供 GE 介面之終端設備接取。

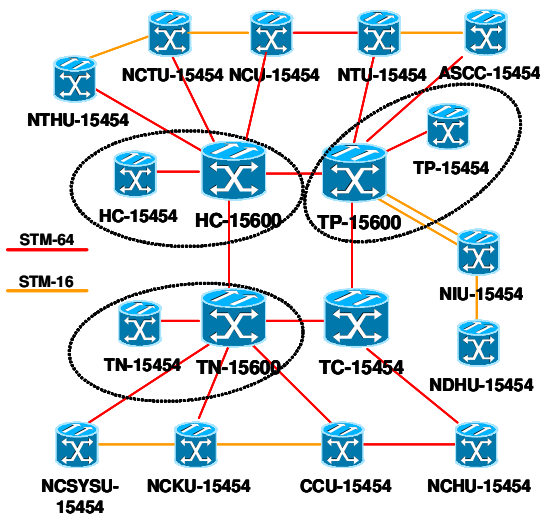


圖 1 TWAREN 整合光網路平台

3. 光通道控制技術

為了控制光通道的配置，我們需對光通道的控制技術有更深入的瞭解，光網路在各國際學研網路發展已有一段時間，各學研網針對各式各樣不同的網路需求與應用，皆發展出一些有用的管理技術，本章將介紹各網路研究計畫所開發各種不同的控制技術。

3.1 Bandwidth Reservation for User Work (BRUW)

隨著網路發展，經常有大量資料需要透過網路來傳送，為避免大量傳輸造成網路癱瘓，提供專屬網路服務便是各學術研究網路的重要課題。但每條專屬網路的建立，都需要瞭解網路架構的網管人員對連線上的各路由器進行相對的設定，相當麻煩。為了減少人力的浪費，美國 Internet2 希望能建立一套頻寬預留系統，根據使用者的連線需求，由系統自動建立一條專屬使用的線路，供使用者運用。

BRUW[8]是 Internet2 Hybrid Optical and Packet Infrastructure(HOPI)計畫的一部份，允許授權使用者針對分封交換網路(Packet Switched Network)的骨幹頻寬進行預留，以建立專屬的連線頻寬提供大量資料傳輸應用來使用。系統分為前端系統與後端系統，使用者需先向 Internet2 提出申請，獲得認證之後才可登入前端系統進行連線頻寬預留。

前端系統以網頁介面為基礎，使用者在認證後可登入前端系統進行頻寬預留，系統使用 Shibboleth 中介軟體技術，對使用者的權限做控管。使用者僅需輸入起始端 IP、目的端 IP、連線所需的頻寬及連線建立的期間，不需瞭解複雜的骨幹架構，即可自動建立起始端至目的端之間的連線。系統會檢查當時骨幹頻寬使用之情況，並回覆使用者頻寬預留是

否成功。

後端系統以排程的方式，每小時檢查是否有連線建立或刪除的需求發生。當有連線異動需求時，系統使用 traceroute 的指令，找出距離起始端與目的端最近的兩個路由器，接著利用 JUNOScript[9]，找出預約時間內骨幹可用頻寬符合需求的路徑，最後由後端系統對路徑上的各路由器下指令，以建立專屬的 MPLS(Multi-Protocol Label Switching)通道。

BRUW 目前已成功的使用在美國 Abilene 的分封交換網路骨幹上，可在分封交換網路中有效建立專屬的通道，未來希望進一步結合其他非分封交換網路，例如光網路系統，提供異質網路環境中專屬頻寬的建立。

3.2 JRA3-BoD (Joint Research Activity 3 - Bandwidth on Demand)

近年來各學術研究計畫使用點對點的傳輸的頻率越來越高，尤其歐洲一些大型的學術研究計畫，往往需要建立跨國、跨網路環境的點對點路徑來進行資料傳輸，歐洲各學術研究網路單位為解決這樣的問題，於是發起了 JRA3-BOD 的計畫。

JRA3-BOD[10]系統專注於建立跨異質網路的點對點路徑，包括整合 SDH 網路的 GFP 協定、Layer2 的 MPLS 協定與 Native Ethernet 等跨領域的技術，希望由此系統的協助，減少建立連線的時間與人力的浪費。

JAR3-BOD 可整合各網路環境的格式，建立連線的過程由 IDM(Inter-Domain Manager)負責溝通與協調。IDM 是整個 JRA3-BOD 系統的核心，主要功能有：

- JRA3-BOD 系統的唯一入口。
- 接收所有來自使用者與其他 IDM 系統提出的頻寬需求。
- 與 AAI(Authentication and Authorization Infrastructure)服務進行互動，對使用者進行身份認證與權限控管的任務。
- 根據提出的頻寬需求，對所管控的網路設備進行頻寬預留。
- 決定建立連線所需的其他 IDM，並發出頻寬需求以通知下一個 IDM 進行頻寬預留的動作。
- 在不同研究網路的 IDM 間進行資料同步與資料交換的工作。

除了在各異質網路環境間建立點對點路徑的功能外，JRA3-BOD 系統也針對 Layer 2 網路，開發相關的監控機制。目前 JRA3-BOD 系統尚處於開發階段，待系統開發完成之後，將可提供一個泛歐洲學術網路的研究平台，提供跨網路平台的環境讓各種研究與應用在此平台進行使用。

3.3 OSCARS (On-Demand Secure Circuits and Advance Reservation System)

由於隨著一些對頻寬高度敏感性需求的應用

出現(如視訊會議、遠端實驗操作、有時間限制的大量資料傳輸...等),對於美國能源部網路 ESnet 的網路架構來說,似乎更須要增加一些網路頻寬保證的服務。OSCARs 的目的在於動態地建立 QoS(Quality of Service)路徑給使用者來使用。使用者並不需要自行選擇路徑,只要提出其所需頻寬要求即可。所有的保證頻寬路徑是由一個頻寬預留管理者(RM, Reservation Manager)以及網路上的路由器所控制協調來達成。以下將對於實作的架構以及機制逐一說明。

- (1) QoS、MPLS and RSVP: 要達成頻寬保證的實作是由兩個要素所組成:
 - 點對點的路徑。
 - 路徑的頻寬保證。而 MPLS 將會達成點對點的繞徑產生, QoS 會實現路徑的頻寬保證。
- (2) Reservation Manager(RM): RM 是由三項子系統所組成(AAAS), 且具有與 WS-RF(Web service Resource Framework)服務模組一致性的架構, 其利用 WSDL(Web Service Description Language)定義了一個語言獨立的標準介面。整個系統狀態是由 RM 來管理, 而且是由標準的通訊服務介面-SOAP(Simple Object Access Protocol)所呼叫實作。經由這些機制, 允許更多以 WS-RF 所實作的應用程式作一個整合。
- (3) Authentication、Authorization and Auditing System(AAAS): AAAS 負責使用者的認證以及頻寬要求使用的認可。藉由一個發佈的標準 WSDL 介面, AAAS 會接收一個 SOAP 所簽章的訊息, 其中內容包括頻寬預留的資訊與被認證者名稱。另外, 所產生的歷史記錄將於未來配置及管理頻寬所使用。
- (4) Bandwidth Scheduler Subsystem(BSS): BSS 負責紀錄各點對點網路鏈結上所預留頻寬的資訊。BSS 必須能通知回報目前的網路狀態中, 是否能有保證 QoS 的能力。所以 BSS 必須有能力了解所有網路上目前最新的資源使用狀況。
- (5) Path Setup Subsystem(PSS): PSS 負責建立以及釋放根據需求所建立的路徑。這可以藉由改變網路上的路由器設定而達成 LSP(Label Switched Path)的建立與釋放的動作。
- (6) Deployment and User Access: OSCARS 正在發展能夠藉由 WBUI(Web-Based User Interface)來允許使用者預留其所需的保證頻寬, 在如同 ESnet 的網路上。期間藉由 WBUI 和 RM 的溝通協調, 並經由 HTTPS 的安全協定認證。一旦對於使用者來說, 頻寬的建立、維護以及釋放都是通透性且不需要了解的。

3.4 DRAGON (Dynamic Resource Allocation via GMPLS Optical Networks)

DRAGON 計畫的發展目的, 在於提供動態的網路服務資源配置及跨異質性網路的資源分配研

究技術。而發起的源由, 包括以下所需克服的技術: 1) 對於 Label Switched Paths (LSPs) 而言, 尚無標準化的跨網域路由架構, 而導致在跨多重管理網域與異質網路的環境下建立 LSP 仍是問題; 2) 沒有簡易的使用者介面; 3) 沒有具體的點對點實作案例(包含認證、紀錄系統); 4) 沒有標示非 General Multi-Protocol Label Switching(GMPLS)網段的能力。因此, DRAGON 計畫希望能從中研發出相應的對策。

而 DRAGON 計畫發展出了三個元素: Network Aware Resource Broker (NARB), Virtual LSR (VLSR), and an Application Specific Topology Definition Language (ASTDL), 用來解決上述的問題。

- 每一個 Autonomous Domain (AD)都會有一個 NARB, 用來收集自己網域內的路由訊息, 並與其他網域的 NARB 交換必要的路由訊息。NARB 還具備資源管理與計算出最佳 LSP 的功能, 並負責身份驗證與紀錄。因此當使用者的應用程式要與另一端的應用程式作溝通時, 所屬網域的 NARB 就會被觸動而去建立一條(跨)網域的 LSP, 以供此應用程式傳訊息用。
- 當 LSP 需要跨過不支援 GMPLS 的網段時, 就需要 VLSR 的幫忙。VLSR 有著 OSPF-TE 與 RSVP-TE 的能力, 並透過 SNMP 與不支援 GMPLS 的路由設備做溝通, 讓此設備就像是支援 GMPLS 的設備一樣, 也能讓 LSP 順利建立通道。
- ASTDL 則是用來方便使用者與他的 NARB 做溝通的介面, 只要輸入來源端與目的端的 IP 還有其他的參數, 就能建立所需要的最佳化的 LSP。

此計畫的架構已經在美國的 Washington 特區被具體實現出來, 並用兩個應用程式 electronic Very Long Baseline Interferometry (eVLBI) 與 High Definition Collaboration and Visual Area Networking (HD-CVAN)來做測試。此兩個應用程式對於網路的 QoS 有著極大的需求, 因此非常適合運作於擁有快速規劃並建立 QoS LSP 能力的 DRAGON 架構。

4. UCLP (User Controlled LightPath) 與 UCLPv2 測試

UCLP 是由加拿大 CANARIE 開發於 CA*net4 上管理光通道的光網路管理系統。由於 CANARIE 所使用的設備與 TWAREN 相同且架構類似, 本章我們將描述試用 UCLP 系統之經驗, 並從中了解光通道控制的概念。

隨著 UCLP 版本的演進, CANARIE 已在今年五月推出 UCLP version 2(UCLPv2)[11], UCLPv2 強調更多設備的支援、更強大的網管功能、更容易操作與更美觀的使用者介面, 因此我們選擇 UCLPv2 做為我們測試使用的版本。UCLPv2 主要的組成元件

如下所述：

- NE-WS(Network Equipment Web Service)：負責與光設備連通，存放與光設備溝通的資訊，包含使用協定、IP、通訊埠、登入光設備的帳號密碼等。
- Lightpath WS(Lightpath Web Service)：負責存放兩台光設備實體連線間，欲配置的光通道連線資訊，包含兩邊介接之 slot、port、channel 等。
- Interface WS(Interface Web Service)：負責存放光設備與網路設備連接介面的資訊，包含介接的 slot、port、頻寬等。

我們在網路實驗室已建置一台 UCLPv2 主機，並以 CA*net4 實驗室提供的 West 測試區作為測試環境，設備包含四台 Cisco 的 ONS-15454/SONET 光設備與 2 台路由器，West 測試區網路架構如圖 2 所示。

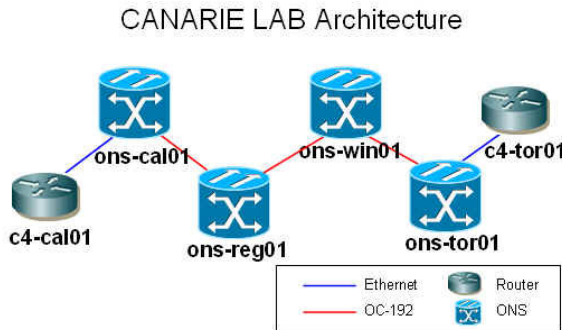


圖 2 CA*net4 實驗室測試架構圖

建立連線之相關資訊，茲分述如下：

(1) 建立實體網路環境

利用 UCLP v2 中實體網路之圖形化編輯介面，針對實體網路架構配置出適合的地圖及新增、修改實體網路環境。

(2) 新增光網路設備

利用拖曳方式方便配置光網路設備於實體網路環境中適當的位置，畫出實體網路設備之拓撲圖。將相關光網路設備連結設定資訊，新增於 NE-WS 中。

(3) 建立實體網路拓撲圖

根據實際網路架構，建立各光設備間實體連線。圖 3 為完成後實體網路架構圖。

(4) 切割光通道

光網路的特性在於可動態在一條實體線路上切割多個光通道，此步驟將實體線路中欲提供給 UCLPv2 系統使用之光通道頻寬切割出來，並配置給 UCLPv2。

(5) 新增 Lightpath Web Service

將各條光通道與設備間界面的資訊，新增於 Lightpath WS 中。

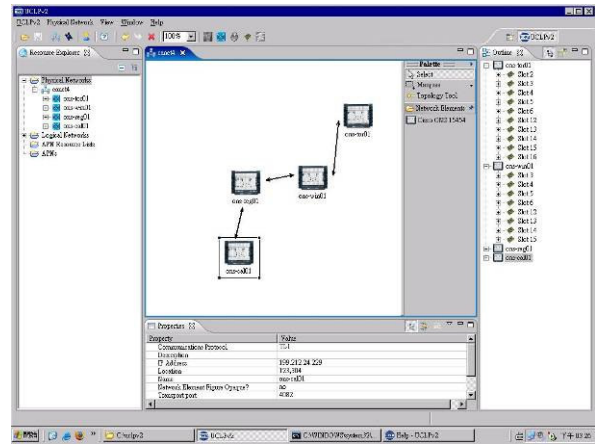


圖 3 CA*net4 實驗室實體網路架構圖

(6) 新增 Interface Web Service

將光設備與網路設備間界面介面的資訊，新增於 Interface WS 中。完成後之邏輯網路架構如圖 4。

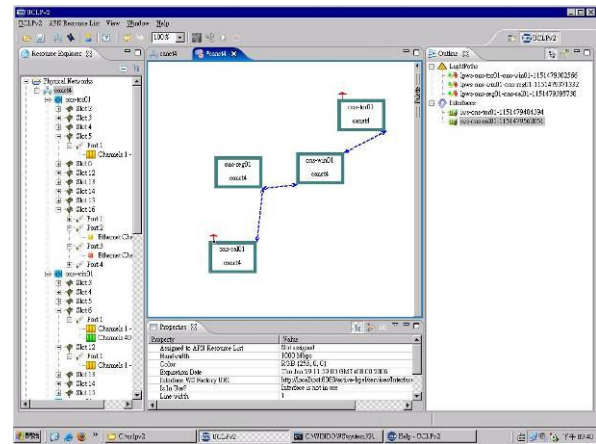


圖 4 CA*net4 實驗室邏輯網路架構圖

(7) 建立連線

點選起始端至目的端所有相連的 Interface WS 與 Lightpath WS，建立點對點連線，完成光網路連線。

(8) 測試連線成功

利用 CANARIE NOC 網站上提供之 Looking Glass[12]為測試工具，在兩路由器上分別以 ICMP 協定測試所建立之連線是否正常運作。

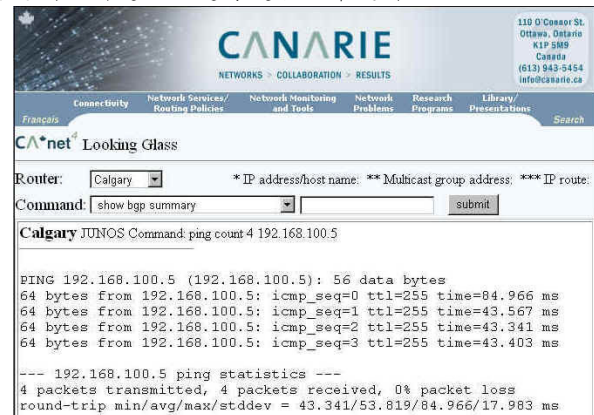


圖 5 測試連線成功圖

在本測試中，實際使用 CA*net4 實驗室所提供真實網路環境之光網路設備，成功的完成建立光網路連線及連通，證明 UCLPv2 對於光通道的操控是可行且容易使用的。

UCLP v2 目前雖支援 CANARIE 的 SONET 系統，但不適用於 TWAREN 光網路之 SDH 系統，無法實際利用 UCLP v2 建立光網路連線配置，因此目前只先行利用 UCLP v2 所提供之測試模式規劃 TWAREN GigaPOP 之光網路拓撲圖，如下圖 6。未來將協助把 SONET 版本的程式碼移植為適合 SDH 使用的版本，並規劃成為 TWAREN 的光通道管理系統。

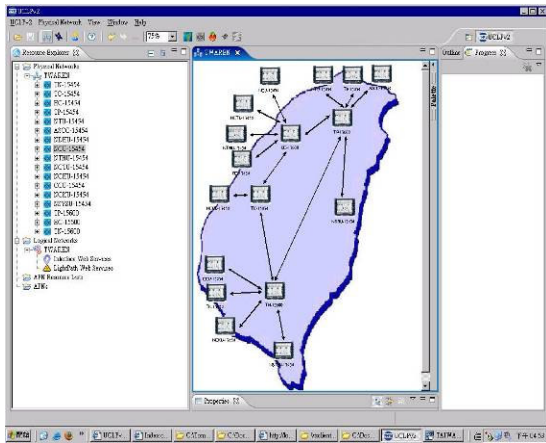


圖 6 TWAREN 光網路拓撲圖

5. 光通道管理

光網路的維運工作除了光通道的建立與移除外，對於目前系統中所佈建之光通道，如何監控其狀態亦為管理者的一大挑戰。針對此一目標，Cisco 提供了 CTC (Cisco Transport Controller) 與 CTM (Cisco Transport Manager) 兩套管理軟體，CTC 為內嵌於 ONS 內之 Java Applet，管理者可藉由 Web 瀏覽器下載執行，於其內可設定並觀察光交換器內之各種狀態；CTM 則為安裝於管理者電腦內之軟體，搭配 Oracle 資料庫，可將光交換器之狀態儲存供日後查詢與分析。

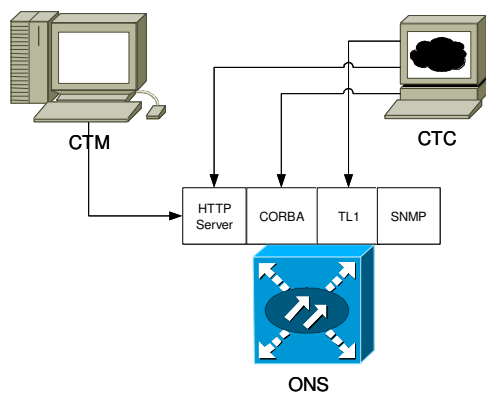


圖 7 CTM/CTC 與 ONS 間之界面

Cisco ONS 亦開放了不同的程式語言介面供 CTC、CTM 等網管程式存取光交換器內之相關資訊，圖 7 為 CTC 與 CTM 和光設備之介面溝通之示意圖，茲分述如下：

a) HTTP server

ONS 內建 HTTP 伺服器，供管理者下載 CTC 的 Java Applet 與 CTM 所需之各項資料，一般自行開發之軟體並不會使用此介面。

b) CORBA

ONS 提供 CORBA 架構，藉由介面定義語言 (interface definition language, IDL) 可以讓程式開發者撰寫管理程式，CTC 與 CTM 即為使用此 CORBA 介面所開發而成；但是一般之開發者需額外購置 CTC CORBA Gateway 作為 Object Request Broker (ORB) 方能使用 IDL 定義檔以撰寫網管程式。

c) TL1

Transaction Language 1 (TL1) 為傳統電信設備之標準控制語言介面，透過 telnet session 傳遞 ASCII 命令至光設備內，並亦可讀取光設備以 ASCII 形式回傳之資訊，此介面較為易學且使用時無需購置額外軟體，UCLP 軟體即為透過此介面撰寫而成。

d) SNMP

Simple Network Management Protocol (SNMP) 為網路設備之管理標準，除了一般狀態之查詢外，於異常事件發生時亦會主動送出告警訊息 (SNMP Trap)。

表 1 ONS Trap 格式

變數號碼	變數名稱
1	sysUpTime
此 SNMP Trap 發生時之系統開機時間	
2	snmpTrapOID
此 SNMP Trap 所對應之 Object ID，包含描述此事件之字串	
3	cerentNodeTime
Trap 發生時光設備之系統時間	
4	cerent454AlarmState
告警訊息之嚴重程度與影響範圍，嚴重程度分為 Critical、Major 和 Minor 三種等級；影響範圍則分為 NonServiceAffecting 及 ServiceAffecting	
5	cerent454AlarmObjectType
事件起因物件所屬之 SNMP Table，如 STM-n、backplane、DWDM Trunk 等，網管系統利用此值進行更進一步的分析	
6	cerent454AlarmObjectIndex

此事件起因物件於 SNMP Table 內之索引值	
7	cerent454AlarmSlotNumber
事件起因物件所屬之卡版編號，若此物件與卡版無關則此值為 0	
8	cerent454AlarmPortNumber
事件起因物件所屬之埠號，若此物件與網路埠無關則此值為 0	
9	cerent454AlarmLineNumber
事件起因物件所屬之網路線號碼，若此物件與網路線無關則此值為 0	
10	cerent454AlarmObjectName
事件起因物件以 TL1 格式表達之名稱	

表 1 為 ONS Trap 之格式，針對光設備之維運所需，我們將針對 SNMP 之主動告警訊息進行搜集並分析其內容，未來將可望由告警訊息找出問題之根源 (root cause)，減少維運除錯之時間。

6. 結論

在可預見的未來將會有更多的新一代網路應用程式需要光網路所提供的專屬頻寬，因此光網路之佈建將是我們需要持續努力的方向，以確保頻寬供應無虞；同時為使光通道配置更為便利，光通道控制技術亦為國際研網發展的重點，UCLPv2 已有完善的界面與設定模式，但是尚未支援 SDH 系統，我們將會協助 UCLPv2 移植到適用 TWAREN 的 SDH 系統上；為了管理已配置的光通道，未來也將分析 SNMP 告警訊息並建立偵測系統平臺，以提供完整的光網路服務。

參考文獻

- [1] CA*net 4 website,
<http://www.canarie.ca/canet4/index.html>
- [2] National Lambda Rail, <http://www.nlr.net>
- [3] SURFnet website, <http://www.surfnet.nl>
- [4] TWAREN website, <http://www.twaren.net>
- [5] 劉德隆、王冠超，「TWAREN 光網路與 DWDM 系統整合測試」，TANET2005 論文集，台中，2005 年 10 月。
- [6] B. St-Arnaud, et al., “Web services architecture for user control and management of optical Internet networks,” vol. 12, issue 9, Proceedings of the IEEE, September 2004.
- [7] Cees de Laat, “Optical/Photonic Networking and Grid Integration,”
<http://staff.science.uva.nl/~delaat/talks/cdl-2005-06-21.pdf>
- [8] BRUW website ,
<http://vbvod.internet2.edu/internship/bruw/index.html>
- [9] JUNOScript website ,

- <http://www.juniper.net/support/junoscript/>
- [10] JRA3-BOD website ,
<http://www.geant2.net/server/show/nav.756>
- [11] UCLPv2 website ,<http://www.uclp.ca/>
- [12] CA*net NOC Looking Glass
<http://dooka.canet4.net/lg/lg.php>