

TWAREN IPv6 群組廣播之規劃與建置

陳俊傑 許瑞明 楊哲男

財團法人國家實驗研究院

國家高速網路與計算中心

jjchen@nchc.org.tw、rmhsu@nchc.org.tw、cnyang@nchc.org.tw

摘要

本篇主要介紹 TWAREN 在「IPv6 群組廣播 (Multicast)」之架構規劃與建置狀況，並且以 TWAREN 目前網路設備上的 IPv6 Multicast 設定為例，說明 IPv6 Multicast 的建置與路由協定的相關設定，以期推廣 IPv6 新一代網際網路協定的群組廣播網路服務。

關鍵詞：IPv6，群組廣播、Multicast、PIM Sparse-Mode、Rendezvous Point、MP-BGP。

Abstract

The design and deployment of TWAREN IPv6 Multicast are mainly introduced in this paper. We take the IPv6 multicast configurations of network facilities currently on TWAREN for example. They describe the deployment of IPv6 Multicast and relative routing protocol configuration. Wish to popularize the IPv6 Multicast network service.

1. 前言

由於網際網路的日漸普及，現有的 IPv4 位址已經不敷使用，新一代的網際網路協定-IPv6 於是產生，其目的主要是要解決 IPv4 位址的缺乏與增進網路使用的效能。

在 IPv4 網際網路協定中，群組廣播 (multicast) 的功能的在於可以減少發送端 (sender) 與接收端 (receiver) 間網路通道 (network channel) 和封包 (packet) 的數量，因此也有助於提升多媒體資訊的傳輸效能。這些特性對某些多媒體應用來說是很重要的，例如視訊會議 (video conference) 和隨選視訊 (Media on Demand, MOD)。因為這些應用都有需要支援多個使用者並自同時自相同伺服器接收多媒體資訊，而有些多媒體資料因具有即時性而對網路傳輸的效能相對要求也較高。

對於解決 IPv4 位址不足的困境，IPv6 網路的建置與應用提供一個解決之道，IPv6 unicast 網路已經普遍存在各個學術研究網路，但是 IPv6

Multicast 群組廣播協定仍是一個比較先進的技術，有建置 IPv6 Multicast 網路服務的學術研究網路仍是少數。

所以本篇主要是介紹 TWAREN 在 IPv6 Multicast 網路架構的規劃與建置現況，但由於受限於網路設備 IPv6 Multicast 功能的完整性，會以 TWAREN 四個主節點的設備設定做例子解說 IPv6 Multicast 的運作但考量到必須讓末端使用者能夠使用 IPv6 Multicast 功能，所以利用設備廠商提供的臨時變通的設定方法讓 TWAREN 各個 GigaPOP 設備可支援 IPv6 Multicast 功能達到可以提供 IPv6 Multicast 服務給連線單位的目的。

底下本篇文章會分四個章節：在第 2 章節會介紹網路上群組廣播的運作方式與 IPv6 群組廣播 (Multicast) 路由協定的種類。TWAREN IPv6 Multicast 網路架構規劃與建置會在第 3 章節中介紹並列舉網路設備的設定與解釋 IPv6 Multicast 路由協定的運作。因為 IPv6 Multicast 為一較先進的網路協定，因應網路設備的功能支援不完全，為使各個 GigaPOP 與連線單位可以使用 IPv6 Multicast 服務，在第 4 章節中會介紹與 IPv6 Unicast 搭配變通的設定方式，供 GigaPOP 與連線單位使用 IPv6 Multicast 網路。最後，第 5 章節對本篇文章作一個簡要的結論與未來 TWAREN IPv6 Multicast 建置待改善的地方。

2. IPv6 群組廣播 (Multicast)

IPv4 網路上封包傳遞的方式大致可分為 Unicast、Multicast 及 Broadcast 等型式，其個別之使用時機及目的亦各有所不同。Unicast 封包主要用於兩個網路節點互相交換資料，此類封包資料只對此兩個節點有實質之作用，網路上的其他節點可以不必理會些類封包。Broadcast 封包主要是要將封包傳送給連接在網路上的所有網路節點，無論該節點需要，所有連接在該網路上的節點均要接收並處理 Broadcast 封包。Multicast 則是用於單一設備傳送封包給多個網路節點的一種方式，雖然 Multicast 是將封包傳送給網路上的多個節點，但與 Broadcast 不一樣的是並非網路上所有的節點均需要接收並處理 Multicast 封包。

透過 IPv4 Multicast 可以將一份相同的資料，送給多個接收者，而使用 Unicast 方式則需要複製

多筆相同的資料。如圖 1 所示，使用 Unicast 方式，則來源端需要複製多筆資料出去，而 Multicast 方式則只需要一筆即可。也正是如此，與 Unicast 相較起來，Multicast 有降低不必要之重複流量、增進頻寬的使用率及減低伺服器的負載等優點。[1]

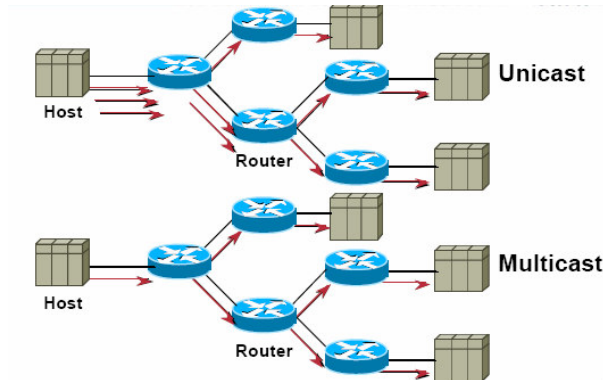


圖 1. Unicast 與 Multicast

正因為 IPv4 Multicast 有這樣的優勢，因此目前許多影音應用皆可支援 multicast，如：Access Grid、多點視訊以及 MOD (Media on Demand) 或是網路電視 (NetTV) 等，與數年前相較起來，IPv4 Multicast 技術日益受到重視。

然而，Multicast 也並非全然沒有缺點，正因為他的某些特性，也可能是他的致命傷：群組廣播主要是 UDP Based。也因此，Multicast 是採取盡力傳輸 (Best Effort Delivery)、沒有碰撞避免機制、以及可能會有封包失序的現象發生。

IPv6 Multicast 運作的方式也是如上面所描述的 IPv4 Multicast 運作的方式一致，其優點在於 IP 位址與群組廣播的群組位址增加且增加了避免 IPv4 Multicast PIM Dense-Mode flooding 的機制。

IPv6 Multicast 相關的協定大致上可分成三部分來處理：

(1). Host 對 Router：MLDv1 (RFC 2710) [2]、MLDv2 (RFC 3810) [3]。MLD 是存在於區域網路內的一種 IPv6 Multicast 路由協定，其主要的功能在於協助 Router 找出位於其區域網路中之參與群組廣播的主機 (Host)，其功能運作方式與 IPv4 中的 IGMP 相同。MLDv1 相當於 IGMPv2，MLDv2 相當於 IGMPv3。MLDv2 完全往後相容於 MLDv1

(2). Intra-Domain Router 間：在 IPv6 Intra-Domain multicast 交換協定為 PIM 路由協定，其主要是用於 Multicast Router 與其他 Multicast Router 進行 multicast 訊息交換，如 (S, G) 以瞭解 multicast 整體架構。在 IPv4 Multicast 中有區分為 Dense-Mode(PIM-DM) 與 Sparse-Mode(PIM-SM)，但由於 IPv6 Multicast 去除 router flooding 的交換方式，所以在 IPv6 Multicast PIM 路由協定中只有 Sparse-Mode。Sparse-Mode 會先透過相關機制，詢問哪裡有接收者，先將派送樹建立起來，再將

multicast 訊務送下去，避免設備需要定時做 flooding 的動作，減少設備資源的浪費。目前 IPv6 PIM Sparse-Mode，IETF 有一最新版本為 draft-ietf-pim-sm-v2-12[4]，但是尚未完整成為 RFC 標準版本，有些廠商使用新的 PIM Sparse-Mode 規格在 IPv6 Multicast 協定上，而使用舊的 PIM Sparse-Mode(RFC 2362)在 IPv4 Multicast 上。如圖 2 所示，MLD 與 PIM Sparse-Mode 協定在 IPv6 Multicast 網路環境中的關係。

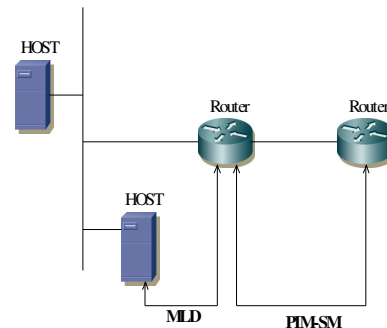


圖 2. MLD 與 PIM-SM 路由協定

(3). Inter-Domain Router 間：IPv6 Inter-Domain Multicast 沒有如 IPv4 Multicast 的 MSDP (Multicast Source Discovery Protocol) 協定可以用來交換群播網路間的 RP 訊息，所以只能用一個 Global Static RP 位址來統籌跨領域的 Multicast 網路中的群播群組或是使用內嵌式的 RP 位址表示方式[8]，將 RP 位址內藏在群播位址中，經由 IPv6 的路由找到 RP 的所在地，如此不需有固定 RP 位址，經由群組位址解譯出 RP 位址，在經由 IPv6 路由找到 RP 所在路由設備，即可發現群播群組的資訊。

3. TWAREN IPv6 群組廣播

3.1 Intra-Domain multicast protocol

為因應 IP 位址的不足與影音多媒體應用的發展，提供這些 IPv6 網路應用的群組廣播平台，TWAREN 骨幹網路也支援 IPv6 Multicast 服務。在 TWAREN IPv6 Multicast Intra-domain 架構中內，TWAREN IPv6 Multicast 使用與 IPv4 Multicast 相同的 PIM Sparse-Mode 協定。

PIM (Protocol Independent Multicasting)，顧名

思義，該協定跟其下層的 Unicast 路由協定（如 RIP、IS-IS、OSPF、BGP 等）是無關的，不像 MOSPF 其一定需要與 OSPF 搭配，這對於我們網路路由協定的選擇上會有相當大的彈性。而在 IPv6 PIM 協定中只有 Sparse-Mode，由於 Dense-Mode 需要定期性將 multicast 訊務廣泛地送到整個網路內，且 PIM Dense-Mode 是使用 source-tree (shortest-path tree)，會消耗較多的 memory ($O(S \times G)$)，而 PIM Sparse-Mode 則部分採用 shared-tree 使用較少的 memory ($O(S)$)。所以在 IPv6 Multicast PIM 協定中將 Dense-Mode 捨去，只採用 Sparse-Mode 而 PIM Sparse-Mode 相對也比較穩定。因此 TWAREN 在 Intra-domain 內，是選擇 PIM Sparse-Mode (Protocol Independent Multicast Sparse-Mode) 為其 multicast protocol。

RP (Rendezvous Point) 的選擇

在 PIM Sparse-Mode 中，都會選擇一個 RP (Rendezvous Point) 位址，因為 RP 位址所在的設備中存在所有在此 Domain 中的 source sender 資訊，因此 RP 位址的選擇也相對重要。

發送端 (Sender) 透過”註冊 (register)”的動作，驅使 RP 加入與該發送端間的 shortest-path tree，從而建立發送端與 RP 之間的 source-tree；而接收端 (Receiver) 則是透過”加入 (Join)”的動作，建立 shared-tree。換言之，兩種分別由發送端與收端建立的派送樹，皆在 RP 中相會，以建立發送端到接收端的完整的路徑。圖 3 所示即是 PIM Sparse-Mode 之 multicast tree 建立情形。

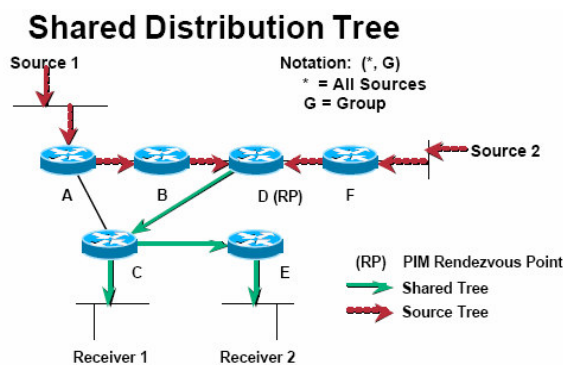


圖 3. PIM Sparse-Mode multicast tree

TWAREN 國內 IPv6 Multicast 網路架構與 IPv4 Multicast 網路架構在 RP 的規劃與設定略有所不同。TWAREN IPv6 Multicast 網路架構如圖 4。在 IPv4 架構中 [10]，因為有 MSDP 的協定，採取 Anycast RP 的架構，以四個主節點 GSR 為 RP，各個 RP 間跑 IMSDP 協定，達到各個 RP 間有互相備援與負載平衡的功能。但是因為 IPv6 Multicast 沒有 MSDP 這項協定，所以 TWAREN 採用 Static RP 位

址的方式，以 Taipei 主節點設備為 TWAREN Intra-domain PIM Sparse-Mode 的 RP 位址所在，TWAREN 所屬國內 IPv6 Multicast 網路群組：FF3E:40:2001:E10::/64 皆以 Taipei 主節點為 RP 位址。

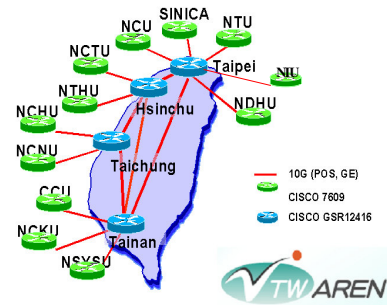


圖 4. TWAREN 國內 IPv6 multicast 網路架構

因此，TWAREN 的 intra-domain 選擇使用 PIM Sparse-Mode 的 multicast protocol，並且指定了 TP-GSR 的 Loopback0 為 RP 位址。以下是 CISCO GSR 與 7609 的相關指令 [5][10]：

Step 1：在 global mode 下啟動 router ipv6 multicast 路由功能

```
!
ipv6 unicast-routing
ipv6 multicast-routing
!
```

Step2：因 Cisco router 在啟動 IPv6 multicast 功能後，預設會啟動 MLDv2 與 PIM-SM，所以此步驟只需在需要啟動 ipv6 multicast 的介面上啟動 ipv6 功能：

```
!
interface Loopback0
  ipv6 address 2001:E10:FFFF:1000::2/128
  ipv6 enable
  ipv6 cef
!
```

Step 3：指定網路 domain 內 RP 的位址，其中 ipv6 access-list 為 TWAREN 國內 IPv6 Multicast Domain 的群組範圍。

```
!
ipv6 pim rp-address 2001:E10:FFFF:1000::2
  TWAREN-RP
ipv6 access-list TWAREN-RP
  \permit ipv6 any FF3E:40:2001:E10::/64
!
```

各 GigaPOP 點 7609 對主節點間的 IPv6 multicast

啟動方式與 GSR 相同，步驟如下（以台大 7609P 對台北主節點 GSR 為例）：

```
Step 1：啟動 router multicast 功能
!
ipv6 unicast-routing
ipv6 multicast-routing
!
Step 2：在需要啟動 ipv6 multicast 的介面上啟動 IPv6 功能：
!
interface TenGigabitEthernet4/1
description "10GbE to TP-12416R"
ipv6 enable
ipv6 cef
interface GigabitEthernet9/3
description "台灣大學 GE-1"
ipv6 address 2001:E10:FFFF:1401::1/64
ipv6 enable
ipv6 cef
!
Step 3：指定 RP 位址：
!
ipv6 pim rp-address 2001:E10:FFFF:1000::2
\TWAREN-RP
ipv6 access-list TWAREN-RP
\permit ipv6 any FF3E:40:2001:E10::/64
!
```

3.2 Inter-Domain multicast protocol

3.2.1 MP-BGP (Multi-Protocol BGP)

MP-BGP 定義在 RFC2283 (Multiprotocol Extensions for BGP-4) [6]。顧名思義，MP-BGP 不同於傳統 BGP，主要是其可以傳遞多種路由資訊，包含了：

- IPv4 unicast 路由所需要的 IPv4 prefix
- IPv4 multicast RPF (Reverse Path Forwarding) 檢查所需的 IPv4 prefix
- IPv6 unicast 路由所需的 IPv6 prefix
- IPv6 multicast RPF (Reverse Path Forwarding) 檢查所需的 IPv6 prefix

此外，MP-BGP 也負責維護 Unicast 及 Multicast 路由資訊表 (Routing Information Base, RIB)。Unicast RIB (U-RIB) 主要是存放了路由器在做 unicast forwarding 時所需要的 unicast prefix；M-RIB 則是存放了當進行群組廣播時必須先進行的 RPF 檢查時所需要的 unicast prefix (此為 unicast prefix，只是當有 multicast 流量經過時，它會被路由器用來

做反向路徑檢查時所用)。

儘管 MP-BGP 可以支援多種的路由資訊，但是除了上述這些主要的分別外，其餘如路徑選擇等參數，皆與傳統 BGP 無異。

在 TWAREN CISCO GSR IOS 中有完全支援 IPv6 MP-BGP，所以在 TWAREN IPv6 Multicast 架構中，四顆主節點設備利用 MP-BGP 功能，達到 full-mesh IPv6 MP-BGP neighbor。GSR 利用 MP-BGP IPv6 multicast address-family 做 IPv6 Multicast RPF 的檢查。底下為 Taipei 主節點 GSR 的 IPv6 MP-BGP full-mesh iBGP 的設定：

```
router bgp 7539
neighbor BetweenRV6 peer-group
neighbor BetweenRV6 remote-as 7539
neighbor BetweenRV6 description [V6 iBGP
\ between R]
neighbor BetweenRV6 update-source Loopback0
neighbor 2001:E10:FFFF:1100::2 peer-group
\ BetweenRV6
...
!
address-family ipv4
neighbor 211.79.59.130 activate
...
!
address-family ipv4 multicast
neighbor 211.79.59.130 activate
...
!
address-family ipv6
neighbor BetweenRV6 activate
neighbor BetweenRV6 next-hop-self
neighbor 2001:E10:FFFF:1100::2 peer-group
...
!
address-family ipv6 multicast
neighbor BetweenRV6 activate
neighbor BetweenRV6 next-hop-self
neighbor 2001:E10:FFFF:1100::2 peer-group
\ BetweenRV6
...
exit-address-family
!
```

從上面的設定例子，發現 MP-BGP 協定可以區分為四大 address-family 區塊，而在每個區塊內，管理者可以針對每個互連的網路決定是否啟動與該互連網路的該路由資訊交換 (neighbor <ip-address> active) 與訂定各自獨立的路由政策。

所以，在 Inter-Domain Multicast 路由選擇上會

以 MP-BGP IPv6 Multicast address-family 做選擇並以 MP-BGP peering 的網路交換 IPv6 address family 路由資訊來作為 RPF check 的主要選擇，底下為 TWAREN 與工研院、HiNet 做 MP-BGP IPv6 Multicast peering 設定範例：

```
!  
router bgp 7539  
  neighbor 2001:E10:FFFF:1001::2 remote-as 18422  
  neighbor 2001:E10:FFFF:1001::2 description \  
  [ITRI_IPv6_peering]  
  ...  
  !  
  address-family ipv6 multicast  
  neighbor 2001:E10:FFFF:1001::2 activate  
  ...  
  exit-address-family  
  !
```

3.2.2 Global Static RP 位址設定

因為 IPv6 Multicast 沒有支援類似 IPv4 Multicast 架構中的 MSDP 協定，所以在 TWAREN IPv6 Multicast Inter-Domain 架構採用 Satic RP 位址架構，針對 Global 群播群組的範圍採用 M6bone[7] 建置的一個 Global RP 位址，互連網路的群播群組範圍採用互連網路提供的 RP 位址，底下為台北主節點的 Static RP 設定範例：

```
!  
...  
!  
ipv6 pim rp-address  
  \ 2001:238:0:1:20:3066:770:1200 Hinet-RP  
ipv6 pim rp-address 2001:660:3007:300:1::  
  \ Global-RP  
!  
ipv6 access-list Hinet-RP  
  permit ipv6 any FF3E:40:2001:238::/64  
!  
ipv6 access-list Global-RP  
  permit ipv6 any FF0E::/16  
  permit ipv6 any FF1E::/16  
  permit ipv6 any FF3E::/16  
!
```

其中 FF0E::/16、FF1E::/16、FF3E::/16 為 Global 群播群組範圍，RP 位址為 2001:660:3007:300:1::，該 RP 位址為 M6bone 提供給 Global 群組對應的 RP 位址[7]。FF3E:40:2001:238::/64 為 HiNet 網路的群播群組範圍，HiNet 提供的 RP 位址為 2001:238:0:1:20:3066:770:1200。

利用群播群組範圍與相對應網路的 RP 位址設定，可以提供 TWAREN 使用者 Inter-Domain IPv6 Multicast 的網路服務。

3.2.3 Embedded RP 位址

在 IPv6 Multicast Inter-Domain 架構中，所有的 Global 群播群組都對應到唯一的 Global RP 位址，對使用者與網路服務提供者使用上不方便與 RP 缺乏備援性。Embedded RP 位址的機制[8]就是為了解決此一問題，將 RP 位址內嵌在他們的群播群組中，當 Embedded RP[8] 啟動的路由器碰到這個群組，會解析出 IPv6 unicast RP 位址，再利用 IPv6 unicast routing 即可知道該群組的 RP 位址，所以每個群組都可以擁有自己的 RP 位址，達到一個很彈性的 RP 設定機制。要啟用 Embedded RP 位址機制需要在網路內的所有路由器皆須支援 Embedded RP 功能且都已經啟動運作正常，目前 TWAREN 骨幹網路正在測試 Embedded RP，當測試告一階段，未來就可以提供 TWAREN 使用者更便利的 IPv6 Multicast 網路環境。

4. GigaPOP 與使用者 IPv6 Multicast 連線設定建議

目前 TWAREN GigaPOP 7609 還沒支援 IPv6 Multicast MP-BGP 協定，採取一個暫時變通的方法[10]，將 MP-BGP address-family ipv6 unicast 的路由在主節點的 GSR 設備中轉換到 MP-BGP address-family ipv6 multicast 路由表中，供 RPF check 使用，在 GigaPOP 7609 的設定不用做特殊的改變，只需在主節點 GSR 做如下的設定：

```
!  
address-family ipv6  
  neighbor 2001:E10:FFFF:1400::1 activate  
  neighbor 2001:E10:FFFF:1400::1 translate-update  
  \ ipv6 multicast unicast  
exit-address-family  
!  
address-family ipv6 multicast  
  neighbor 2001:E10:FFFF:1400::1 activate  
exit-address-family  
!
```

在連線單位與 GigaPOP 7609 交換 IPv6 路由，GigaPOP 7609 取的連線單位的 IPv6 路由，將此路由倒入 MP-BGP IPv6 unicast 路由表中，在主節點的 GSR 會再將連線單位的 IPv6 路由轉換成 MP-BGP IPv6 multicast 路由表中，並在骨幹中傳送且與互連網路交換使用者的 MP-BGP IPv6 multicast 路由。

5. 結論

IPv4 群組廣播 (multicast) 在現階段 TWAREN 骨幹[9]，已經是個成熟且穩定的網路服務，也已經

有做過許多的測試與展示，例如 Access Grid 超級視迅格網，研討會議網路即時轉播等。都是影音串流利用群組廣播網路傳送至使用者端，且降低了頻寬的消耗與伺服器的負載。

在 IPv4 位址日漸不足的今天，發展 IPv6 協定是研究網路的一項使命也是未來網路解決 IPv4 位址的一個解決方案，在 IPv6 協定設計中就有把 Multicast 的精神導入 IPv6 位址的欄位規劃，所以當我們將 IPv4 Multicast 網路運作的很穩定成為一項常態的網路服務，需考量到未來 IPv4 位址不足時的解決方案，TWAREN 目前在主節點間已經完全提供 IPv6 Multicast 網路服務，未來將繼續與廠商協調，加快 GigaPOP 設備 IPv6 Multicast 功能的完整性，期能使整個 TWAREN 的 IPv6 Multicast 功能完整，TWAREN 將繼續測試 Embedded RP 位址的機制，提供使用者更方便、更有效率的 IPv6 Multicast 群組群播服務並設法在 TWAREN 骨幹上實際架設 IPv6 Multicast 視迅播放主機，由末端使用者接收資料來量測 IPv6 Multicast 的效能數據與檢驗 TWAREN 骨幹傳輸 IPv6 Multicast 資料的穩定度。

參考文獻

- [1] Developing IP Multicast Networks, Cisco Press.
- [2] Multicast Listener Discovery Version 1(MLDv1), RFC 2710.
- [3] Multicast Listener Discovery Version 1(MLDv1), RFC 3810.
- [4] Draft-ietf-pim-sm-v2-new-12, IETF INTERNET-DRAFT.
- [5] Implementing IPv6 Multicast, Cisco Press
(http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123cgcr/ipv6_c/sa_mcast.pdf)
- [6] Multiprotocol Extensions for BGP-4, IETF RFC2283, T. Bates, R. Chandra, D. Katz, Y. Rekhter, February 1998
- [7] M6Bone : an experimental IPv6 Multicast network (<http://sem2.renater.fr/m6bone>)
- [8] Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address, RFC 3596.
- [9] 許瑞明, 楊哲男, 陳俊傑, “TWAREN 群組廣播之規劃與建置”, TANet 2004, 11 月 2004, 台東.
- [10] IPv6 Multicast Deployment and Configuration Guide, Cisco Press
(http://www.cisco.com/application/pdf/en/us/guest/products/ps6594/c1650/cdcont_0900aecd80320fb8.pdf)