第五章 IP 通訊協定

5-1 網際層簡介

網際層（Internet Layer）相當於 OSI 參考模式的網路層（如圖 1-6 所示），負責處理資料在工作站間傳送時的路徑選擇問題，其中包含建立、維護、以及結束兩部工作站之間的連線。我們用圖 5-1 來說明網際層在 Internet 網路之中所扮演的功能，其中，網路存取層提供兩網路端點之間的媒介存取，以及資料傳輸流動的管理，基本上還是在同一邏輯網路內。但在網際層所提供的通訊連線，也許會跨越連結許多不同型態的網路，也就是說，在浩瀚無涯的網路之中，如何達成兩部工作站之間的連線，這就是網際層所提供的功能，因此，網際層又稱之為**『工作站對工作站的連線』（Station-to-Station Connection）**。如圖 5-1 中，工作站 A 和工作站 B 也許位於地球的兩端，如何來達到它們之間的連線，這是網際層所必須提供的服務，其中包含如何尋找到達的路徑，以及它們之間的資料傳送，在這連線之間也許會經由許多不同型態的實體網路，譬如 Ethernet 網路、Token-Ring 網路、廣域網路之傳輸網路等等，這些實體網路之間的連線問題，是由該地區的網路存取層負責。也就是說，一條網際層連線，也許是由許多不同型態的網路存取層所構成，亦或許是由多條網路存取層連線銜接而成。

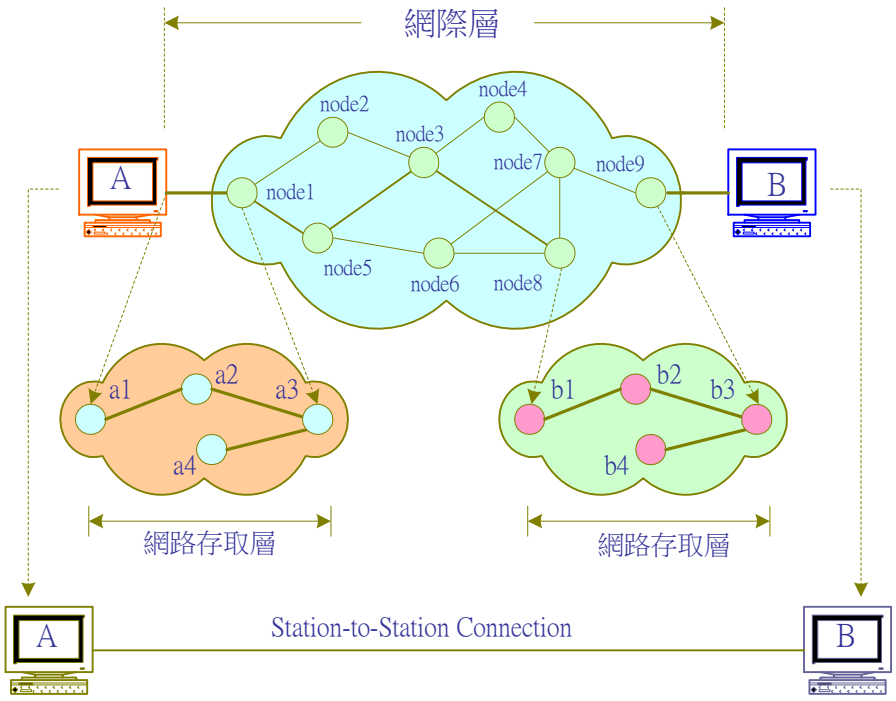

圖 5-1 網際層之工作站對工作站連線

5-1-1 網際層通訊協定

Internet 網路為了能連結遠端地區（全球網路）電腦之間的通訊，因此網際層採用變異性較高的**電報傳輸（Datagram）**方式，亦是**『網際協定』（Internet Protocol, IP）**，其中包含下列幾項重要通訊軟體：

- **IP（Internet Protocol）：**網際通訊協定。提供複雜網路之間路徑選擇的功能，遠端之間的電腦可透過 IP 協定尋找出對方位址，並將其連接在一起。
- **ARP（Address Resolution Protocol）：**位址解析協定。工作站可透過 ARP 協定以對方（使用者）的網路位址（IP 位址），來查問它的網路實體位址（Ethernet 位址）。
- **RARP（Reverse Address Resolution Protocol）：**反向位址解析協定。使用者透過 RARP 協定以本身的網路實體位址（Ethernet 位址），向網路上其它工作站（如名稱伺服器）詢問（或要求）本身的網路位址（IP 位址）。

5-1-2 網際層訊框包裝

在 Internet 網路上，上層通訊協定（TCP 或 UDP）、網際層的 ICMP 與 IGMP 封包都包裝在 IP 封包之內，以 IP 方式傳輸，但 ARP 和 RARP 因牽涉到實體網路位址，而以獨立封包格式傳送，因此，在網路存取層（如 Ethernet 網路）上，所接受的上層訊框格式有如圖 5-2 與 5-3 所示之三種格式。

早期 Ethernet 委員會是以圖 5-2 定義 Ethernet 與 IP 封包格式（RFC 894），在此封包內並沒有 802.2 LLC 的標頭。當時的想法是 IP 封包為非連接方式，針對每一封包都是獨立的傳送，並不需要再闢一個連接導向的 LLC 層。後來 IEEE 委員會希望 802.3 協定也可能適用於連接導向的區域網路上，譬如，NetBIOS（Windows 2000） 或 AppleTalk 上，因此再設計圖 5-3 的訊框格式（RFC 1024），其中就包含了 IEEE 802.2 的封包標頭。

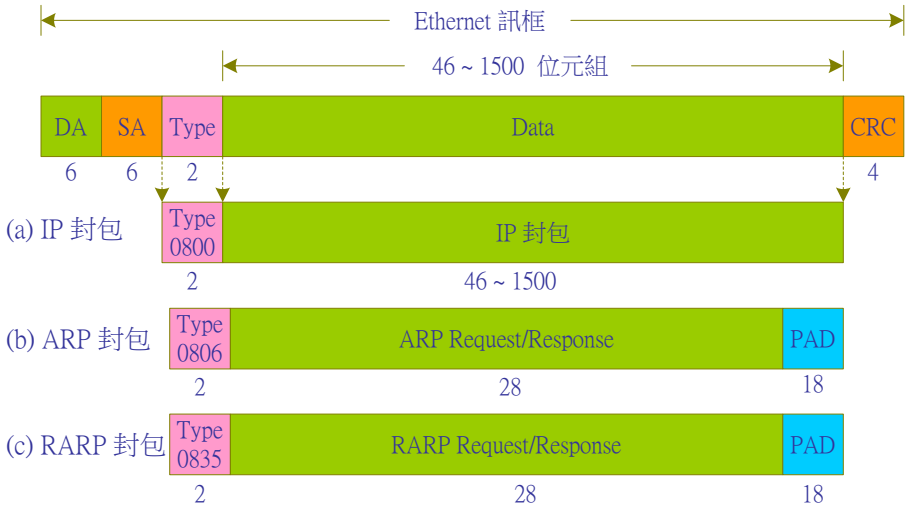

圖 5-2 網際層封包與 Ethernet 訊框包裝

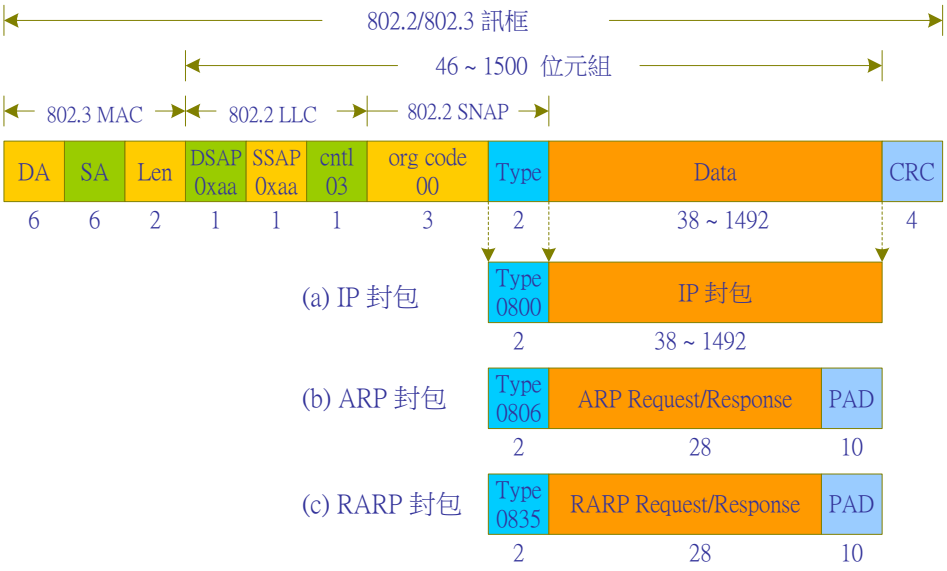

圖 5-3 網際層封包與 802.2/802.3 訊框包裝

一般網路卡都必須能接收與傳送這兩種訊框格式的功能，兩者最大的區分在於 Ethernet 封包並沒有長度（Length）欄位，而直接以型態（Type）欄位取代。區分這兩種訊框有兩個方法：

(1) 802.2/802.3 訊框的長度大小不可能會和型態欄位內容相同（0x0800、0x0806、0x0835），因為訊框長度最高為 1500 位元值，不可能會超過 0x0800（2048）。因此我們只要由該欄位的內容就可判斷出是屬於 802.2/802.3 或 Ethernet 訊框。

(2) 在訊框的前置訊號（Preamble）加入特殊碼代表。一般會在 Ethernet 訊框的前置訊號最後一個位元（Start Delimiter, SD）以 0xD5 表示，可用來區分兩種不同的訊框格式。

在圖 5-2 與 5-3 中，各欄位大小皆以位元組表示，其中 802.3 MAC 為 802.3 訊框的標頭；802.2 LLC 為邏輯鏈路控制的標頭包裝。802.2 SNAP（Sub-Network Access Protocol）為次網路存取

標示包裝，其中 org code 都是 0，再以 Type 欄位分辨封包格式，以 0x0800 表示所承載之協定單元為 IP 封包；0x0806 為 ARP 封包；0x0835 為 RARP 封包。如果是 Ethernet 封裝的話，IP 最大承載的資料為 1500 位元組；而 802.3 訊框包裝，則 IP 最大傳輸量為 1492 位元組。基本上，ARP 和 RARP 所傳送的訊息只有 28 位元組，但為了補足 802.3 的最小訊框長度，因此以 10（802.3）或 18（Ethernet）位元組的 PAD 欄位來填補。

圖 5-4 是以 Windows 2000 上的網路監視器所擷取的 Ethernet 訊框格式（有關網路監視器的使用方法請參閱附錄 A），該訊框的產生是由 Windows 2000 執行下列命令所得之訊框。

> telnet 163.15.2.62 <Enter>

其中訊框長度（Frame Length）是由計算所得（66 Bytes），並非在訊框欄位上，此訊框之格式如圖 5-5 所示。

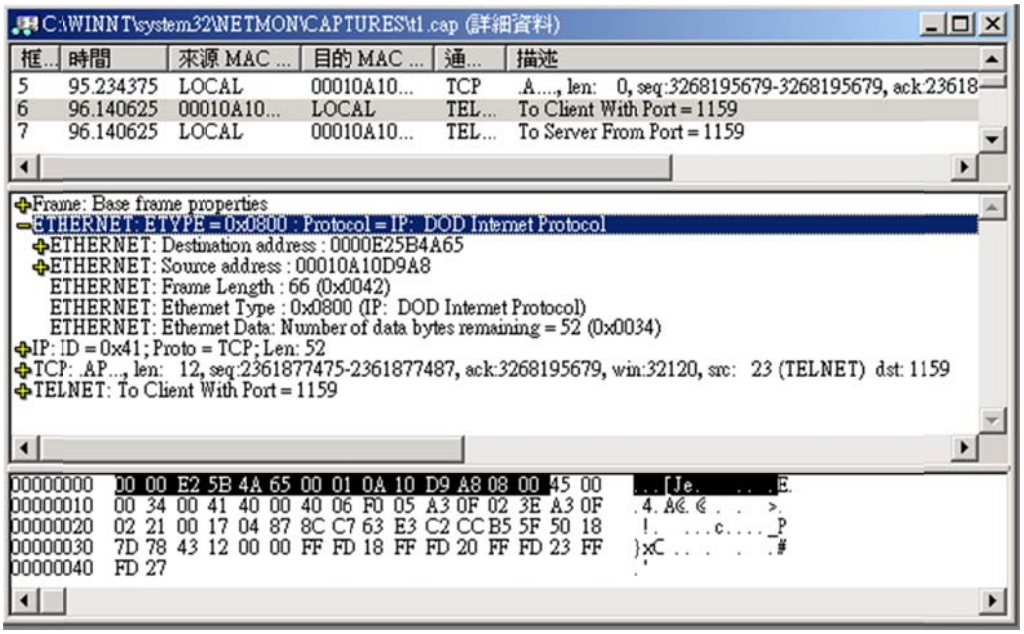

圖 5-4 Telnet 命令所擷取之 Ethernet 封包格式

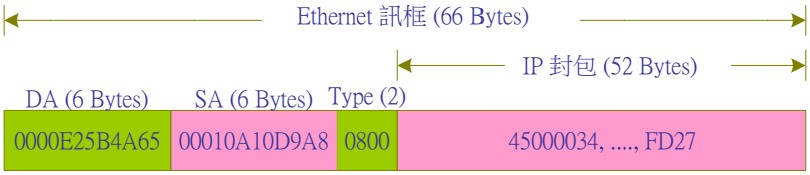

圖 5-5 Ethernet 封裝擷取範例

5-2 IP 通訊協定

IP 通訊協定是 Internet 網路之中最主要的協定，規定在廣泛複雜的網路上，如何尋找到所欲連接之工作站，並負責雙方的連線。IP 通訊協定所採用的非連接方式的『**電報傳輸**』**(Datagram)**，主要的工作有二：(1) 每一部工作站如何去定名？使成為網路唯一的識別名稱，有了這個呼叫名稱，才可以連結其它工作站，或被其它工作站所連結，宛如電話號碼一般；(2) 如何尋找連結路徑？在廣泛複雜網路之中，如何尋找出可以到達目的地的最佳路徑。因此，IP 通訊協定是 Internet 網路中既重要、又複雜的通訊軟體。目前 Internet 網路已連結全世界上億的電腦，任何地區的網路只要透過簡單的路由器就可銜接上 Internet，每一部路由器上主要功能就是實現 IP 通訊協定的功能，所以 IP 必須有足夠的共通性和連結性，本節就依照這個方向介紹 IP 通訊協定。

5-2-1 IP 協定特性

IP 是整個 Internet 協定中的靈魂，任何通訊軟體（ICMP 或 TCP）都是利用 IP 來傳送，亦即，無論資料的功能為何，所有資料都必須經由 IP 協定包裝後再轉送，其功能如下：

- 在網路層和傳輸層之間資料的轉送。
- 進行封包的拆解與重組。
- 將封包傳送到目的主機，包含跨越多個網路找到目的主機所在的網路位址。

基本上，IP 的連接技術是非連接方式的電報傳輸（Datagram），因此在網路通訊的技術上，具有下列特性：

- IP 是非連接(Connectionless)服務。表示 IP 在傳送封包之前，雙方並未事先建立連線。發送端直接將封包發送到網路上，由網路上各個路由器（或網路閘門）負責轉送到達目的地。
- IP 無錯誤偵測功能。當封包進入路由器時，路由器只針對封包標頭上的目的位址轉送，並不保證該封包是否可安全無恙傳送到目的地，因此沒有針對封包內的資料作錯誤偵測。封包在傳輸當中，是否受到干擾或其他原因影響，而發生錯誤，這必須由委託傳送（或上層）的通訊軟體自行負責偵測。
- 封包彼此之間的次序，經傳送到對方後，可能和原來的不同，因為每個封包都自行依當時網路情況尋找路徑（電報傳輸方式），所經過的路徑不一定相同，到達的次序當然也不一定相同。

- 封包可能被重複傳送。重複傳送的原因可能是發送端的上一層協定在溢時後，未收到接收端的回應確認（中途遺失？），再發送另一個同樣的封包。而實際上，接收端在收到第一個封包時已作回應（對方沒收到？），結果又再收到同樣的封包。IP 本身只負責傳送封包，封包是否重複傳送，這必須由上層通訊軟體負責偵測。

- IP 並不驗證目的主機是否有確實收到正確的資料。

由以上看出，IP 協定是一個非常不可靠的通訊協定。因此，有關可靠性的處理動作（例如，錯誤偵測、封包次序等）都必須由另一個可靠的上一層協定來處理，所以 IP 協定必須配置可靠的 TCP 協定。正因如此，TCP/IP 必須整合在一起，對整個網路的通訊而言仍維持可靠的。IP 採取不可靠傳輸的原因是希望在廣泛的網路上，能以既快速且簡單的傳送方式來傳送，如此較容易結合各地區的網路。如果採用可靠的傳輸（如 X.25）要連結全世界的電腦可能要花費不少成本，自然也限制網際網路的發展（如電話系統）。

5-2-2 IP 位址

(A) IP 位址結構

在 Internet 網路中，任何一部連線的電腦或工作站設備都稱為主機（host）。早期 TCP/IP 被設計成適合在不同類型、位置之全球各地的電腦系統之間連接，為了方便標定每部主機，TCP/IP 定義了一套通用的定址方法。當時理想的定址格式必須能提供足夠的路徑選擇（routing）資訊，而且不要佔用太多記憶體空間，因此，將 IP 位址（IP Address）的長度設定為 32 位元。為方便表達，我們將此 32 位元分割成四段，連續 8 位元為一組，每組並以十進位值（0 ~ 255）表示，每組之間以點（dot）分隔。整個 IP 位主表示法就如下所示：

dec3.dec2.dec1.dec0（如 163.15.2.1）

雖然 IP 位址長為 32 位元，但其中包含兩種號碼：網路號碼（Network number）及主機號碼（Host number），因此 IP 位址也可以表示成：（如圖 5-6 所示）

【network number, host number】

圖 5-6 IP 位址結構

在實務連接上，並非每一部主機上都只有一個 IP 位址，一般 IP 位址都依照網路介面卡（Ethernet 網路卡）設定。如主機有特殊需要安裝多個網路介面卡（如當路由器使用），每只網路介面卡上都必須設定一個 IP 位址，因此，一部主機上就擁有多個 IP 位址，同時也容許類似虛擬主機的設定，表示一只網路介面卡上可設定多個 IP 位址。

(B) IP 位址分級

在 32 位元長度的位址之中，應該多少位址長度來表示網路位址或主機位址。TCP/IP 網路依照所能容納的主機和網路的數量多寡分成 A、B、C 、D 和 E 五種類級（class），如圖 5-7 所示，其中 Class D 目前為實驗性多點廣播（Multicast）位址，Class E 則保留未來發展之用。分級技巧是配置不同數目的網路位址，網路位址的位元數愈多，所能指定的網路數量就愈多，但相對應的主機位址就愈少。Class C 所能容納的網路位址最多，所以在 Internet 網路上定址方式皆採用 Class C 模式。Class A 所能容納的主機位址最多，但相對應的所能容納的網路位址最少，一般使用在區域網路的定址模式。各類級的特性如下：

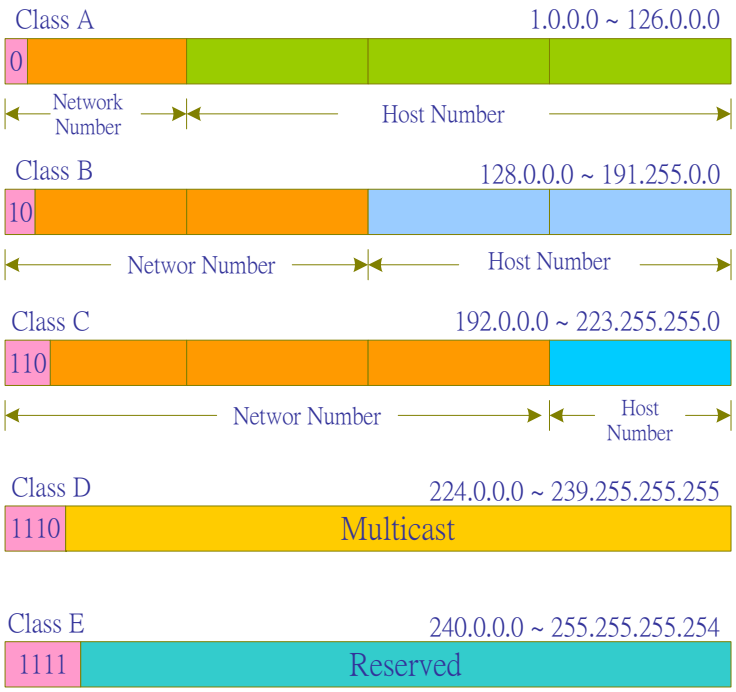

圖 5-7 各類級的 IP 位址結構

(1) Class A：以最高位元（第 31 位元）為 0 表示 Class A 模式。前一位元組（8 位元）表示網路位址；而後 24 位元表示主機位址。網路位址由 1.0.0.0 ~ 126.0.0.0，所能表示的主機位址是 1.0.0.0 ~ 126.255.255.255 的範圍之內。Netmask = 255.0.0.0（下小節說明）。

(2) Class B：以二個最高位元為 10 表示 Class B 模式。前 16 位元表示網路位址；而後 16 位元表示主機位址。網路位址由 128.0.0.0 ~ 191.255.0.0，所能表示的主機位址為 128.0.0.0 ~ 191.255.255.255 的範圍之內。Netmask = 255.255.0.0。

(3) Class C：以前三個最高位元為 110 表示 Class C 模式。前 24 位元是網路位址；而後 8 位元為主機位址。網路位址由 192.0.0.0 ~ 223.255.255.0，所能表示的主機位址為 192.0.0.0 ~ 223.255.255.255 的範圍之內。Netmask = 255.255.255.0。

(4) Class D：以前四個位元為 1110 表示 Class D 模式。其主要應用於多點廣播 (Multicast)，一些特殊應用軟體皆用此模式，來對某些定點 (工作站) 廣播，如隨選視訊 (VOD) 就用此定址模式，對若干個定點工作站廣播視訊。

(5) Class E：以前四個最高位元為 1111 表示 Class E 模式。目前保留尚未使用。

經過 IP 分級後，IP 位址可分為兩大部份：網路位址和主機位址。如以 Class B 中的 163.15.3.42 為例，此 IP 位址可區分為下列兩種位址：

- 網路位址：163.15.0.0
- 主機位址：0.0.3.42

如果僅由 IP 位址來觀察，如此分類好像不是很重要，但是在作路徑選擇時就非常重要，因為一般路徑選擇只觀察網路位址，而不用理會主機位址是多少。

除了上述 IP 分級外，還有一些網路位址保留為特殊使用，其它應用必需避免使用下列位址：

- **預定閘門 (Default Router):** 以網路 0 保留特殊使用，並以『**0.0.0.0**』作為預定閘門位址。
- **回繞位址 (Loopback Address):** 網路 127 也保留於特殊用途，並以『**127.0.0.0**』作為回繞位址，『**127.0.0.1**』表示主機位址，主要作為測試網路使用。
- **廣播位址 (Broadcast Address):** IP 位址全部為 1 時，表示是對所有主機的廣播位址『**255.255.255.255**』。一般應用上只對本網路廣播，因此，將所有主機位址設定為 1 時，表示對本網路所有主機管播，譬如，163.15.0.0 網路的廣播位址為『**163.15.255.255**』。

剛開始設計 TCP/IP 網路時，電腦還未普及，網路也非常少，TCP/IP 協定開始應用時也只連結大型主機，因此 32 位元容量的 IP 位址對當時來講已足足有餘。但沒有想到 Internet 網路大風

行，理論上，任何一部電腦連結上 Internet 網路都需要一個獨一無二的 IP 位址，因此 IP 位址將會在短期內被耗盡。雖然目前已提出 IPv6 的解決方案，但要網路上使用中的路由器和主機都更新為 IPv6 的通訊協定，也並非易事。目前網路上大多透過**『網路位址轉換器』(Network Address Translator, NAT)**來增加私人網路位址，以解決 IP 位址不足的問題。

(C) 網路遮罩（Network Mask, Netmask）

IP 位址是由網路號碼和主機號碼組成的 32 位元，為方便起見，一般都用 [network#, host#] 表示。網路號碼決定主機所屬的網路位址，因此主機在傳遞封包之前，會先過濾出網路號碼，再決定封包應該往哪一個網路傳送。為使能由 IP 位址中過濾出網路號碼，我們使用**『網路遮罩』(Netmask)**來過濾。網路遮罩亦為 32 位元，在位元中 "1" 表示網路位址；而 "0" 表示主機位址，其遮罩方式如圖 5-8 所示。如 IP 分級方式，各等級之網路遮罩為：

- **Class A**：網路遮罩為 255.0.0.0
- **Class B**：網路遮罩為 255.255.0.0
- **Class C**：網路遮罩為 255.255.255.0

在一般網路遮罩，皆是 IP 位址的前面若干位元設定為 "1"，因此，我們以網路號碼的長度 (也是網路遮罩的長度) 來代表網路位址的範圍，以**『主機號碼/網路號碼長度』**來表示一個網路位址，如：

IP = 163.15.2.3/16 表示網路遮罩長度為 16 位元，則：

網路位址 = 163.15.0.0

主機位址 = 0.0.2.3

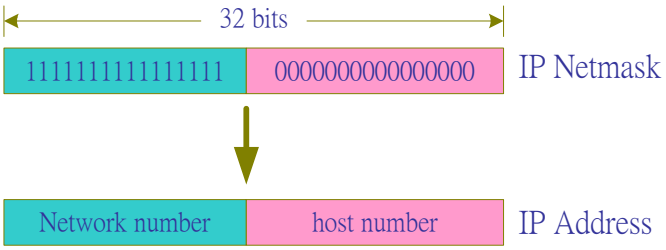

圖 5-8 IP Netmask

(D) 次網路遮罩（Subnet Mask）

如果硬是將網路位址區分為三個類級（Class A ~ C），恐怕很難滿足各種環境需求。例如，目前 TANet 使用 Class B 的定址模式，分配到每一個學校 2 ~3 個網路號碼。每一個學校裡的網路，是由多個系所的區域網路所構成，在技術上，每一個區域網路都要有一個網路號碼，因此網路號碼一定不符所需，這也是 IP 分級所衍生的問題。解決的方法就是再劃分次網路（Subnet），產生次網路的基本原理，是將原有主機號碼的幾個位元拿來當網路號碼，並沒有改變原來 32 位元的 IP 位址格式。

如欲劃分次網路，就必須有次網路的編號，原來 IP 位址所表達的是 [network#, host#] 方式，就必須更換為 [network#, subnet#, host#] 形式。而增加次網路號碼，就必須犧牲原來主機號碼的數量，次網路號碼增加愈多，主機號號碼就減少愈多。對整個位址格式並未改變，因為原來網路號碼並未改變，對外部網路而言，次網路位址也被視為主機位址，因此連結到外部網路並不影響。我們以一個例子說明，假使希望在 163.15.0.0 的網路上增加 8 個次網路，基本上，163.15.0.0/16 網路是屬於 Class B 格式，它原來的 IP Netmask 為 255.255.0.0。我們為了增加 8 個次網路，必須將 3 個位元的主機號碼拿來當次網路號碼，因此 Netmask 為 255.255.224.0，如圖 5-9 所示。

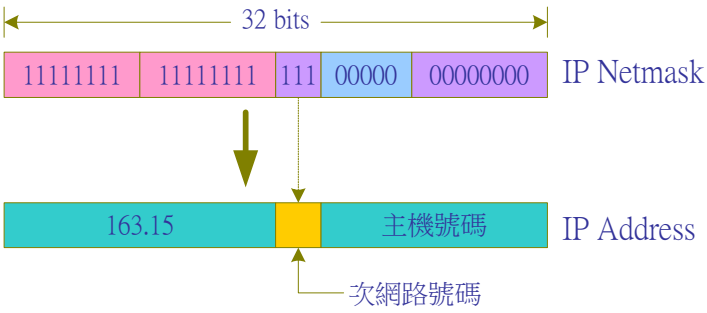

圖 5-9 次網路號碼

分割後所產生的網路位址為：（Network number = 163.15.0.0、Netmask = 255.255.224.0）

163.15.0.0/19、163.15.32.0/19、163.15.64.0/19、163.15.96.0/19、163.15.128.0/19、163.15.160.0/19、163.15.192.0/19、162.15.224.0/19。

主機位址和網路位址範圍為：

第 1 個次網路範圍：163.15.0.0 ~ 163.15.31.255。

第 2 個次網路範圍：163.15.32.0 ~ 163.15.63.255。

第 3 個次網路範圍：163.15.64.0 ~ 163.15.95.255。

第 4 個次網路範圍：163.15.96.0 ~ 163.15.127.255。

第 5 個次網路範圍：163.128.0.0 ~ 163.15.159.255。

第 6 個次網路範圍：163.160.0.0 ~ 163.15.191.255。

第 7 個次網路範圍：163.15.192.0 ~ 163.15.223.255。

第 8 個次網路範圍：163.15.224.0 ~ 163.15.255.255。

如果依照 IP 分級中，網路位址全部為 0 和 1，保留給預定路由器和廣播位址，則上述第 1 和 8 不可以作為次網路位址(RFC 950)，因此，次網路位址剩下 6 個。但最新規範(RFC 1518、1519) 允許使用這兩個網路位址，但一般路由器必須有支援此功能（CIDR）才可以。

5-2-3 IP 封包結構

圖 5-10 為 IP 封包之資料結構，其中包含 IP 標頭和 IP 承載（IP Payload）(圖中 Data) 兩大部份。IP 標頭的長度可以由 20 Bytes 到 60 Bytes 不等（由 IHL 欄位登錄），對整個 IP 封包長度可以是 46 ~ 1500 Bytes 之間（如圖 5-2 所示）。IP 標頭的各欄位功能如下：

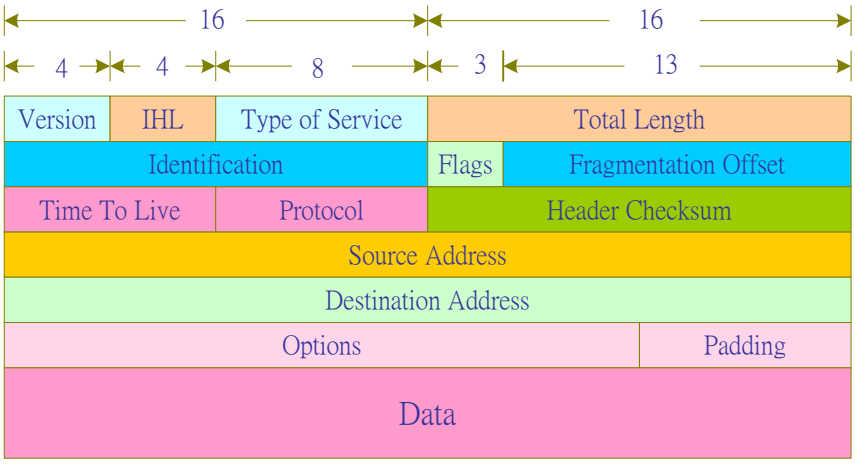

圖 5-10 IP 封包結構

- **版本（Version）：**4 位元。表示 IP 封包的 IP 版本。
- **網際標頭長度（Internet Header Length, IHL）：**4 位元。表示本 IP 封包標頭的長度（位元組表示）。範圍為 5 ~ 15，預設值為 5。
- **服務型態（Type of Service, TOS）：**8 位元。其內容為 "PPPDTRUU"，PPP 表示本 IP 封包的優先權（Precedence）；D = 0 表示一般延遲（Delay），D = 1 表示低延遲；T = 0 為

一般傳送量（Throughput），T = 1 為高傳送量；R = 0 為一般可靠度，R = 1 為高可靠度（Reliability）；UU 保留未用。

- **總長度（Total Length）：**16 位元。是指該封包的總長度，包括封包標頭及資料的長度，範圍由 576 ~ 65535 位元組。

- **辨識碼（Identification）：**8 位元。表示資料分割（fragmentation）的識別編號。同一筆資料如被分割成若干個區段，每段給予相同的辨識碼，接收端再依辨識碼組合回原來資料封包。

- **旗標(Flags)：**3 位元。其格式為 "0DM"。D = 0 表示可以分段，D = 1 為不可分段(Don't fragment)；M = 0 表示最後分段，M = 1 為不是最後分段（More fragment）。

- **分段偏移（Fragment Offset）：**13 位元。表示目前的分段在原始的資料中所在的位址。原始分段允許有 8192 個分段，以每 8 個位元為一個基本偏移量，所以最大容許 65536 位元資料，此值和總長度一樣。因此範圍為 0 ~ 8191，預設值為 0。

- **存活時間（Time to Live, TTL）：**8 位元。表示該封包在網路上的存活時間，封包每經過一個路由器（或網路閘門），該值就被減一。如路由器發現某封包的 TTL = 0，便將該資料片丟棄而不轉送。範圍 0 ~ 255。

- **協定號碼（Protocol Number）：**8 位元。表示該 IP 封包所承載通訊協定的協定號碼。在 TCP/IP 協定中，任何通訊協定（如 IP、ICMP、TCP、UDP 等等）的資料在傳送中都被包裝成 IP 封包，而以 IP 通訊協定發送。所以，在 IP 封包裡必須有協定號碼欄位，來表示目前所承載之資料是屬於哪一個通訊協定（IP、ICMP、TCP 等等）。一些較常用著名（well-known）的協定號碼如表 5-1 所示。

- **標頭檢查集（Header Checksum）：**16 位元。做標頭錯誤檢查用。

- **來源位址（Source Address）：**32 位元。發送此 IP 封包的來源位址。

- **目的位址（Destination Address）：**32 位元。目的主機之位址。

- **選項欄位（Options）：**可變長度。提供多種選擇性服務。目前已定義使用有下列：

  (1) 安全處理機制：有關資料加密與認證。

(2) 路由紀錄：當 IP 封包經過路由器時，讓該路由器登錄其 IP 位址。當封包到達目的地時，可追蹤它所經過的路徑。

(3) 時間戳記：當 IP 封包經過路由器時，讓路由器登錄其 IP 位址和時間。

(4) 寬鬆來源路由(Loose Source Routing): 記錄該封包所必須經由之路徑，為一 IP 位址的序列表。

(5) 嚴格來源路由 (Strict Source Routing): 如同寬鬆來源路由，但嚴格規定祇能依照 IP 序列表傳送該封包。

- **填補欄位 (Padding):** IP 資料片的標頭一定是 32 位元的整數倍，當 Options 欄位不足 32 位元整數倍時由 Padding 欄位補足。

表 5-1 著名協定號碼 (節錄)

協定名稱	協定號碼	協定全名 (協定包裝在 IP 資料片內)
ICMP	1	Internet Control Message Protocol
IGMP	2	Internet Group Management Protocol
GGP	3	Gateway-to –Gateway Protocol
IP	4	IP in IP encapsulation
TCP	6	Transmission Control Protocol
EGP	8	Exterior Gateway Protocol
IGP	9	Interior Gateway Protocol
UDP	17	User Datagram Protocol

圖 5-11 為 IP 封包所擷取的結果 (以 Telnet 命令所得)，而圖 5-12 為該封包的封裝格式。由 Protocol 欄位 (6) 可以得知 Telnet 是使用 TCP 連接。

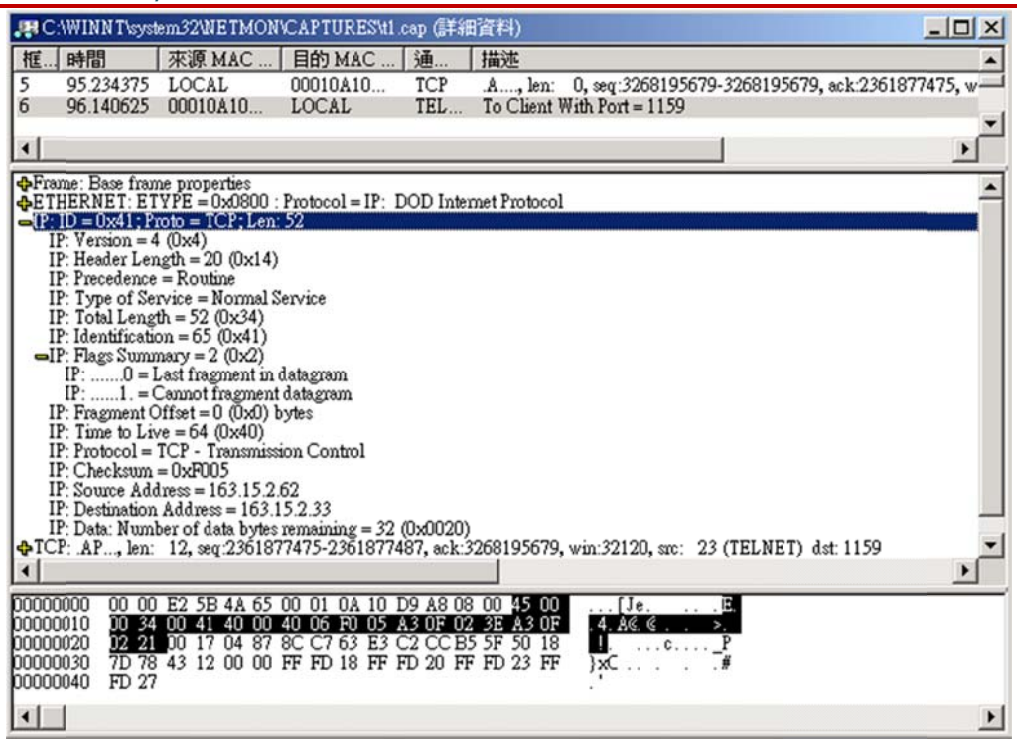

圖 5-11 IP 封包擷取結果

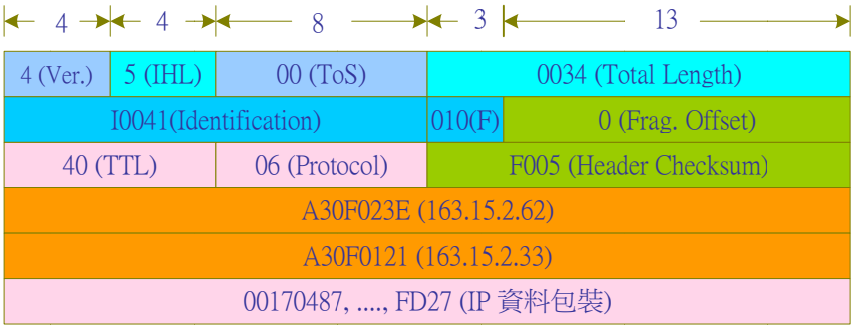

圖 5-12 IP 封裝範例（依圖 5-11 擷取結果）

5-2-4 IP 路徑選擇

IP 路徑選擇（Routing）是 Internet 網路上最主要的中心樞紐。上層的應用程式如何在浩瀚的網路之中找到欲通訊的對象，這就是利用 IP 路徑選擇的功能。基本上 IP 路徑選擇是屬於分散式處理方式，網路上任一路由器（Router）或網路閘門（Gateway）並不瞭解整個網路的狀況，它的路徑選擇功能只知道下一個路徑可能到達哪些地方。當 IP 封包由傳送端發出後，它將不知道，也不可能知道，這筆資料是否可安全到達目的地，和可能經過哪些路徑。所經過的路徑都是先到達某一路由器，再由該路由器決定應該往哪一下個路徑（Next hop）傳送，IP 封包就在這層層轉送之中到達目的地。因此，一筆資料可安全到達目的地，是由網路上所經過的路由器共同來完成，並非發送端事先可以控制的，也不是某一特殊設備事先建立好通訊連線（如虛擬電路）。

為了實現 IP 路徑選擇的機制，在每部主機或路由器上都必須維護一只路由表（Routing Table），在路由表內每一筆記錄至少必須包含下列訊息：

- **目的位址（Destination Address）：**此欄位可以是 IP 的完整主機位址或只有網路位址，如是完整主機位址必須包含主機號碼（Host ID）與網路號碼（Network ID）。
- **下一站位址（Next Hop Address）：**此欄位內表示封包經由轉送之位址。下一站位址也許是該封包的目的位址或在經由下一站在轉送。
- **旗標（Flags）：**旗標用來標示目的 IP 位址是否為一網路位址或是主機位址。一般路由器有下列五種旗標：
  - ★ **U：**該路徑可使用。
  - ★ **G：**該路徑為網路閘門（Gateway）。如此旗標未設定表示該路徑與目的端直接連接。
  - ★ **H：**該路徑通往主機位址。亦是目的位址為一完整的主機位址（網路號碼 + 主機號碼）。
  - ★ **D：**該路徑因轉向所產生。（ICMP Redirect）
  - ★ **M：**該路徑被轉向所改變。（ICMP Redirect）
- 下一路徑之網路介面：該路徑所必須經由之網路介面。

通常區域網路的範圍較小，在同一網路內，主機和主機之間無須借助路徑選擇即可通訊，一般都使用廣播方式（Broadcast）通訊。傳送主機將封包廣播到網路上，同一網路上所有主機收到該封包後，會先檢查封包標頭的目的位址是否和自己的位址相符，如果是就收下該封包；否則忽略它。但在廣域網路上就無法利用廣播方式來傳送封包，如果每部主機都廣播，則整個網路將被廣播封包所佔滿。因此網路之間需設置一個以上的對外網路閘門或路由器，網路閘門就是負責封包對外的所有通訊，隔離內外網路，只讓必要的封包通過，阻擋不必要的封包經過。

路徑選擇的處理方式是當某一部主機或路由器要發送封包時，首先檢查封包之目的 IP 位址，其處理程序如下：

(1) 若目的位址的網路號碼與本主機（或路由器）的網路號碼相同時，則表示目的主機在本網路內，不需選擇路徑，直接將封包送出，例如採用廣播方式（首先查問 ARP）。

(2) 若目的位址的網路號碼與本主機（或路由器）的網路號碼不同時，則表示目的主機不在本網路內，必須跨越路由器到其他網路。就將該封包轉送給另一路由器，由它去負責轉送。

(3) 一般路由器（或網路閘門）都安裝有兩個以上的網路卡，每個網路卡連接一個網路也代表一個網路號碼。當它由任一個網路（或網路卡）上收到一個封包後，依照步驟 (1)、(2) 處理。

當主機（或路由器）發現所要發送（或轉送）的封包不在本網路內，而必須將該封包轉送給其它路由器，由它再負責傳送時，如果網路上有多個路由器，應該轉送給哪一個路由器？這就牽涉到路徑選擇的問題。因此，主機（或路由器）查詢路由表的動作如下：

(1) 在路由表中，尋找完全符合的目的 IP 位址（網路號碼與主機號碼）。如果找到，便將封包轉送給下一路徑路由器或網路連接介面；否則，接下一步驟。

(2) 在路由表中，尋找吻合之網路號碼之紀錄，如果找到，便轉送給該路徑之下一站；否則，接下一步驟。

(3) 在路由表中，尋找**『預設』**路徑位址，便將該封包轉送到預設位址；否則，便回送目的主機無法到達(ICMP Destination Unreachable)或網路無法到達(ICMP Network Unreachable) 之訊號給原發送該封包者。

在整個路徑選擇過程，首先尋找是否相吻合之主機位址，在尋找吻合之網路位址，如果兩者皆找不到適合路徑，再將該封包傳送給**『預設』(Default)**路徑。每一路徑是否有轉送功能，必須視其旗標的設定，如該路徑的旗標設定為 G （Gateway）便有轉送的功能。如設定為 H (Host)，表示最終位址，並不負責轉送，如目的位址不在該網路位址（或網路埠口）所屬網路上，它便會回送 ICMP 給原發送端，表示無法到達目的地。

(A) 運作範例

我們以圖 5-13 來說明 IP 路徑選擇的運作程序。圖中有三個網路：163.15.2.0/24、163.15.3.0/24 和 163.15.4.0/24，利用兩個路由器（Router_A 與 Router_B）所連結而成。每一路由器上都有兩個連接埠口，並個別設定 IP 位址。圖 5-14 為兩個路由器的路由表，其可能是靜態路由(Static Routing) 或動態路由（Dynamic Routing）。我們以下列三種情況來說明路徑選擇的運作程序：

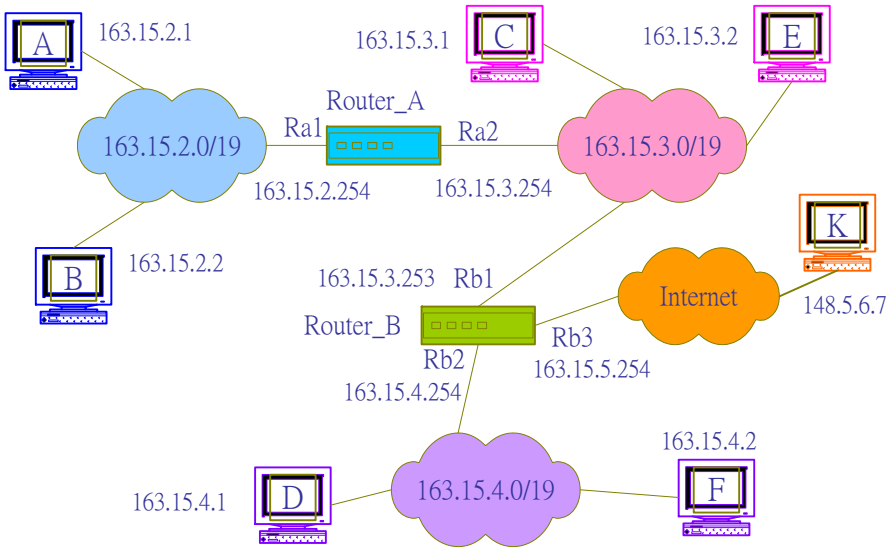

圖 5-13 IP 路徑選擇範例

Router_A 路由表

目的位址	下一站位址	旗標	介面
163.15.2.0/24	163.15.2.254	U	Ra1
163.15.3.0/24	163.15.3.254	U	Ra2
163.15.4.0/24	163.15.3.254	UG	Ra2
Default	163.15.3.254	UG	Ra2

Router_B 路由表

目的位址	下一站位址	旗標	介面
163.15.2.0/24	163.15.3.253	UG	Rb1
163.15.3.0/24	163.15.3.253	U	Rb1
163.15.4.0/24	163.15.4.254	U	Rb2
Default	163.15.5.254	UG	Rb3

圖 5-14 Router A 和 B 的路由表

(1) 主機 A (163.15.2.1) 欲傳送封包給主機 B (163.15.2.2): 主機 A 發現目的網路位址和自己網路號碼相同 (163.15.2.0/24)，便將封包直接廣播到網路上 (由 ARP 查詢對方實體位址後，再傳送給主機 B)，主機 B 由網路上收到封包。

(2) 主機 A (163.15.2.1) 欲傳送封包給主機 C (163.15.3.1): 目的網路號碼不同，必須透過路由器轉送。主機 A 便將封包傳送給 Router_A 的 R1a(預定路由器)。Router_A 由 Ra1 埠口收到封包 (目的位址為 163.15.3.1) 後，便由路由表上得知網路位址是 163.15.3.0/24 的封包，必須轉送到 163.15.3.254(Ra2)，便將該封包由 Ra2 埠口廣播出去(163.15.3.0/24 網路)。主機 C 由網路上讀取到該封包。

(3) 主機 A(163.15.2.1)欲傳送封包給主機 D(163.15.4.1): Router_A 是網路(163.15.2.0/24) 的唯一出口，主機 A 勢必將封包傳送給 Router_A (與網路號碼無關)。Router_A 由路由表上查詢後，便由 Ra2 (163.15.3.254) 廣播出去 (也可直接傳給 163.15.3.253)。Router_B 由 Rb1 收到封包，再由路由表上查出，應轉送到 Rb2 (163.15.4.254)，便由 Rb2 埠口將封包廣播出去。主機 D 由網路上 (163.15.4.0/24) 收到該封包。

(4) 主機 A 欲傳送封包給外部網路之主機 K（148.5.6.7）：該封包會經由 Router_A（Ra1→Ra2），再經 Router_B（Rb1→Rb3），被傳送到外部網路上。至於該封包在 Internet 上如何被傳送到主機 K，是由外部網路之路由器負責，本地網路是幫不上忙的。

另外，我們可以在主機上執行 route 命令，來觀察該主機所建立的路由表：

（必須以 root 使用者登入 Linux 主機）

```
# route
Kernel IP routing table
Destination     Gateway         Genmask         Flags   Metric  Ref    Use Iface
linux-1.cu.edu. *               255.255.255.255 UH      0       0        0 eth0
163.15.2.0      *               255.255.255.0   U       0       0        0 eth0
192.168.0.0     *               255.255.255.0   U       0       0        0 eth1
127.0.0.0       *               255.0.0.0       U       0       0        0 lo
default         192.168.0.1     0.0.0.0         UG      0       0        0 eth1
default         linux-1.cu.edu. 0.0.0.0         UG      1       0        0 eth0
```

(B) 路由表建構方式

路由表的建構牽涉到整個 IP 網路的封包轉送效益，當網路的狀況隨時改變，IP 所經過的路由也隨時變化，因此，建構（或維護）這路由表有兩種方法：

- **靜態路徑選擇（Static routing）：**路由表是由人工建立，完成後除非再由人工修改，否則不會變更。系統維護人員會依照網路狀況或規劃網路拓樸圖架構來建立靜態路由表。（一般使用 route 或 ifconfig 命令來建立）

- **動態路徑選擇（Dynamic routing）：**路由表是由網路上路由器（或網路閘門）隨時交換訊息所建立而成。因此，路由器會隨時依照網路狀況自動維護路由表。至於路由表如何建立的技術，於第七章我們再詳細介紹其演算法。

關於靜態路由表的設定方法，在本書的第九章有針對 Linux 系統上的實作說明，並在第十八章有詳述 Microsoft 網路上的設定方法。有關動態路徑選擇的理論將再第六章說明。

5-2-5 IP 廣播與多點傳送

基本上 IP 通訊協定有：單點傳送（Unicast）、廣播（Broadcast）及多點傳送（Multicasting）等三種定址方式。單點傳送採用 Class A、Class B 和 Class C 三種定址方法，針對單一目的主機

傳送封包，也就是點對點（Point-to-Point）的傳送。針對某一網路（或子網路）的廣播位址的定址方法是將該網路位址之下的主機位址設定為 "1"，譬如在 163.15.0.0/16 的網路下，廣播封包之目的位址設為 163.15.255.255，該網路上任一工作站接收到廣播封包後，再判斷是否傳送給自己，並決定是否保留或拋棄。如果是針對某一子網路廣播，所定址的方式也如同網路一樣，譬如，欲將某封包廣播到 163.15.2.0/24 之子網路上，而它的目的位址為 163.15.2.255。又如果針對網路上所有工作站廣播，定址為 255.255.255.255。

在許多情況下廣播訊息是由傳輸層（TCP 或 UDP）發送到某一特定的傳輸埠口（Port）上，以作為特殊用途（如 VOD），如果還是採用廣播訊息，則網路上每一工作站將必須接收該封包，再從事判斷的工作。例如，一部視訊伺服器（Video Server）發送視訊給網路上收視的工作站，雖然當時網路上有 100 工作站，但只有 30 部工作站收視該視訊，如果採用廣播定址方式，100 部工作站都必須接收該封包，並透過網路存取層、網際層和傳輸層（TCP 或 UDP）判斷，才可決定是否接受該訊息，如此會造成其它未收視工作站不必要的負荷，如採用另一種**『多點傳送』**定址，祇針對已收視之工作站發送訊息，其它工作站在網際層（IP）就可決定是否拋棄該封包，也祇有已收視的 30 部工作站會將該封包拆解到傳輸層。

『多點傳送』是採用 Class D 的群組定址方法，如圖 5-15 所示。它不同於單點傳送 IP 位址（Calss A、B 和 C）模式，多點傳送群組 ID 只需另外配置 28 位元即可，不需另外的結構，而位址範圍由 224.0.0.0 到 239.255.255.255。任一工作站可以選取加入一個或多個多點傳送群組的位址內，而工作站的所在位置並非限於某一固定網路內，亦可跨越不同網路。工作站和群組關係是動態的，工作站可以隨時選擇加入或退出。當工作站接收到多點廣播封包，便依照該封包上的群組位址判斷是否給予接收或拋棄，因此，針對多點傳送封包只拆解到網際層。

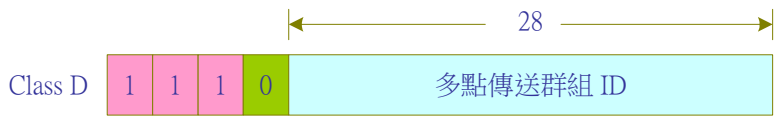

圖 5-15 Class D 多點傳送位址格式

有些多點傳送群組被 IANA（Internet Assigned Number Authority）設定為公認位址，又稱為**『永久主機群組』**。例如，224.0.0.1 為**『子網路上的所有系統』**、224.0.0.2 為**『子網路上所有路由器』**，另外 RIP-2 的 NTP 為 224.0.0.9，其它有關網路閘門之通訊協定都有指定某一特殊多點傳送位址。

5-2-6 IP 分段

在 Internet 網路上，各種訊息（TCP、UDP、ICMP、IGMP）都被包裝在 IP 封包內傳送。雖然整個 IP 封包的最大長度為 65535 位元組，但所經過的實體網路大多有限制一定的長度，譬如 Ethernet 網路最大訊框的長度為 1518 位元組，其所能承載之訊息長度為 1500 位元組，如扣除 IP 封包本身的標頭，最高僅能傳送 1440 ~ 1480 位元組（如圖 5-10 所示）。又當一個 IP 傳送路徑之中，也許會經由許多不同型態的網路，各種網路都有其限制最大框訊框的大小，因此就牽涉到該 IP 封包應該多大才適合，該封包大小就稱之為**『最大傳輸單位』（Maximum Transission Unit, MTU）**。例如，Ethernet 網路之 MTU 為 1492 位元組；IBM Token-Ring 之 MTU = 17914 位元組；IEEE Token-Ring 之 MTU = 4464 位元組；X.25 之 MTU = 576 位元組。

當兩部工作站互相通訊時，它們之間的 MTU 就非常重要，因為通訊之中的封包不一定經由同一路徑傳送，每一路徑所經過之實體網路都有不同的 MTU 限制。因此，MTU 大小並非由兩工作站之間的網路所限制，而是通訊之間所有可能經由之網路的最小 MTU，也稱之為**『路徑 MTU』（Path MTU）**。如何來協議傳送之 IP 封包的 MTU 是利用 ICMP 協定之**『目的無法到達』（Destination Unreachable）**（Type 3）的**『Dragmentation Needed and DF set』**（Code = 4）訊息，來通知發送端必須修改 MTU 大小。或是利用 traceroute 命令來探索所欲到達目的之路徑 MTU，一般國際性網路所經過的路徑較為複雜，路徑 MTU 大多限制在 576 位元組以內，但也都以 512 位元組傳送較多。

我們以圖 5-16 來說明 IP 分段的情形，如上層所傳輸為 TCP 訊息，基本上，TCP 訊息最大長度也可以達到 65535 個位元組，但 IP 封包大小必須依照所經過之網路的 MTU 而定，假設所經過為 Ethernet 網路，Ethernet 網路最大訊框為 1518 位元組，扣除本身的訊框表頭與 CRC 欄位（如圖 5-2 所示），所能包裝的最大訊息為 1500 個位元組，分段的情形，就是將 TCP 訊息以最大 1480 個位元組分割填入 IP 封包內，亦可以發現只有分割後的第一個 IP 封包上有 TCP 的協定表頭，這對於防火牆的製作，將會產生很大的困擾。

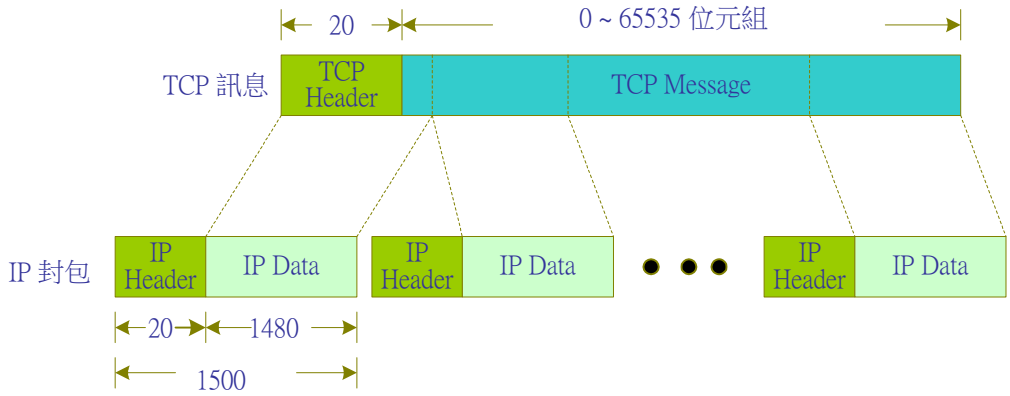

圖 5-16 IP 分段情形（TCP 訊息）

原傳送訊息之工作站並不能預知 IP 可能經過的路徑，因此當分段後的 IP 可能到達另一型態之網路，會被要求再分段。如圖 5-17 中，工作站 A 欲傳送訊息給工作站 B，起先工作站 A 所連接為 Ethernet 網路，而以 1500 位元組型態來分段，但當 IP 封包經由路由器 A 進入 X.25 網路時，該網路的 MTU 為 512 位元組，因此 IP 封包勢必再分割成若干個封包，再傳送給工作站 B。

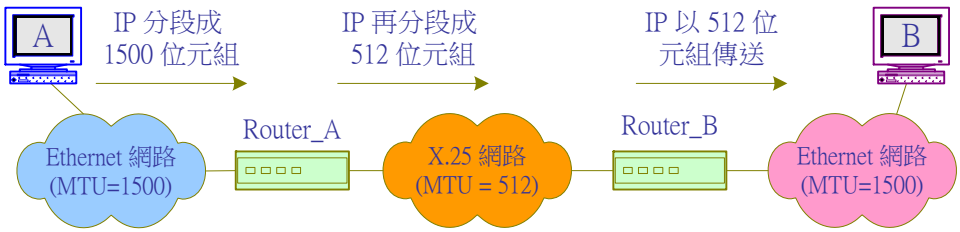

圖 5-17 IP 封包分段之情形

IP 分段並不同於一般通訊協定的情形（如 1-3-2 節說明），IP 分段的機會可能會發生在原傳送端，或路徑中的路由器，但分段後的封包必須到達目的地，才會被組裝回來，其主要的原因是只有原發送端分段的第一個封包有上層通訊協定（如 TCP）的表頭（如圖 5-16 所示），如在中途組裝會產生許多不必要的困擾。因此，IP 封包如果經過分段後，我們將儘可能不要讓它在傳送途中再被分段，因此會在 IP 封包的表頭的旗標（Flags）上設定 DF（Don’t Fragment），並以**『分段偏移』（Fragment Offset）**標示從原來資料開頭到此分段資料的偏移量；旗標中 M = 1 表示後面還有分段之封包；M = 0 表示這是分段封包的最後一個封包。

圖 5-18 為網路之間 MTU 協調情況，被分段的封包進入路由器 A，而準備轉送到 X.25 網路上，路由器 A 發現該封包必須再分段，而 IP 表頭上已設定 DF 旗號，表示不可再分段封包，

因此，路由器 A 會發送 ICMP Fragmentation Needed and DF set 回原發送端（工作站 A）要求重新設定 MTU，如此，便可協議出封包應該分段的大小了。

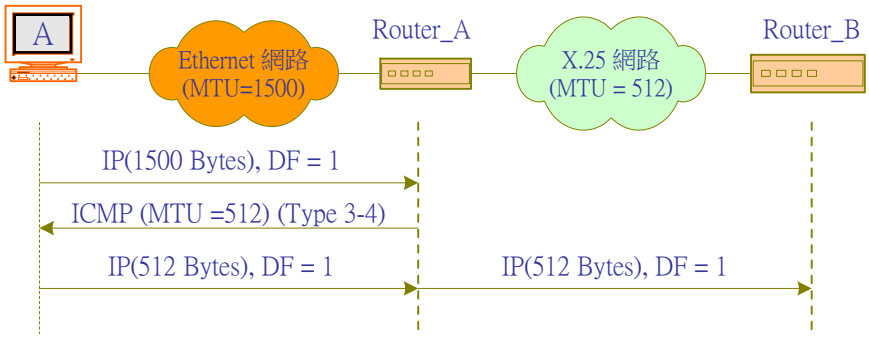

圖 5-18 網路之間 MTU 的協調情形

5-2-7 IP 來源路徑選擇

基本上，當某一 IP 封包被發送出去之前，傳送端並不能預估該封包可能經過之路徑，而是由網路上之路由器轉送而到達目的地。但如果傳送端能預估 IP 所經由路徑，或許可由網路架構圖上觀察出來，或經由『**traceroute**』程式去探索出可到達的路徑，就可以在傳送之前指定該封包到達目的地的路徑，其設定方法是在 IP 封包表頭上的『**選項**』（**Option**）欄位上，註明所欲經過路徑的路由器序列。選項欄位的結構如圖 5-19 所示（請配合圖 5-10 IP 封包結構），一般最高可紀錄 9 個路由器的 IP 位址，因此，選項欄位最長為 39 個位元組，不及 4 位元組的整數倍以 Padding 欄位補足。選項欄位下的各子欄位功能如下：

- **編碼（Code）**：0x83 表示寬鬆來源路徑選擇；0x89 表示嚴格來源路徑選擇。
- **長度（Length）**：表示該來源路徑的總位元數。
- **指標（Pointer）**：標明目前已經過到第幾個路徑中的路由器（以位元組位址表示）。
- **IP 位址**：路徑中所經過路由器的 IP 位址，由前至後的 IP 位址排列表示整個路徑。

圖 5-19 IP 來源路徑選擇之選項

如果使用嚴格路徑選擇，表示該封包所傳送的路徑，一定必須符合 IP 位址序列傳送，倘若有某一路由器無法到達，便回送 ICMP Source Route Faile（Type = 3，Code = 5）回原發送端。寬鬆路徑選擇採用較寬鬆的方法，只要依照該序列路徑，而允許跨越其它位址的路由器。又當封包每經過一個路由器，該路由器便將指標（Pointer）往後一個 IP 位址的空間，再傳送給下一個路由器。

5-3 ARP 與 RARP 通訊協定

在 Internet 網路下，每一部主機都有一個獨一無二的 IP 位址。理論上，欲傳送封包給任何一部主機只要知道它的 IP 位址，便可依照此 IP 位址傳送給該主機。但事實上，IP 位址只是一個網路位址，它是讓較高層的網路程式設計者（如應用層）在編寫程式時不用去考慮網路實際連線的問題（通訊協定的基本功能），但 IP 封包在網路上傳遞時，還是必須透過實體網路位址，才能送達目的位址。例如，IP 網路架設在 Ethernet 上，則網路上所有工作站依照 Ethernet 位址傳送資料，接收端收到訊框後也依照 Ethernet 位址判斷是否傳送給自己，由此可見，在實體網路上，並沒有使用到 IP 位址。因此，某一部主機欲依照 IP 位址傳送給另一部主機，首先必須知道該主機的實體網路位址（一般稱為 MAC 位址），再依 MAC 位址傳送，亦即，必須擁有該主機的 IP/MAC 對照表，同樣的，任何一部主機也必須知道自己的 IP/MAC 對照關係。但問題是當某部主機欲依照 IP 位址發送資料給另一主機時，它怎麼知道對方的 MAC 位址（如 Ethernet 位址）？一般主機當然知道自己的 MAC 位址，但如何得知自己的 IP 位址呢？

在 Internet 通訊協定裡有兩個協定來解決上述的問題，一為**『位址解析協定』（Address Resolution Protocol, ARP）**；另一為**『反向位址解析協定』（Reverse Address Resolution Protocol, RARP）**。ARP 是用來查問欲傳送之目的主機的 MAC 位址，也就是說，由已知的 IP 位址查問其相對應的網路實體位址；而 RARP 是由已知的網路實體位址查詢其相對應的 IP 位址。

5-3-1 ARP 協定

位址解析協定（ARP）是被用來以 IP 位址查詢其相對應的網路實體位址（MAC 位址），其運作方式如圖 5-20 所示。首先主機 A（163.15.2.1）的網際層（Internet Layer）發送出 ARP Request（查問 163.15.2.4）訊息給網路存取層（Network Access Layer），網路存取層之 MAC 層（Ethernet 層）再將 ARP Request 訊息包裝在 Ethernet 訊框內，並廣播在網路上。網路上（區域網路內）所有主機接收到廣播訊框再拆解訊框。主機 C（163.15.2.4）的網際層收到 ARP Request 後，並由其中瞭解是詢問自己，便回應 ARP Reply（包含 Ethernet 位址） 給網路存取層，網路存取層再發送

訊框給主機 A（163.15.2.1）。當然，其他主機也會收到 ARP Request 訊號，但皆判斷不是詢問自己而不予理會。

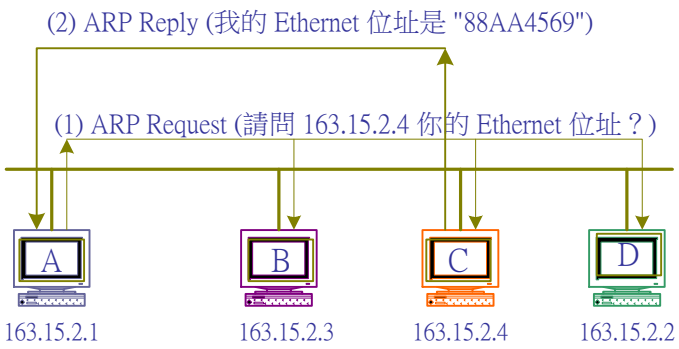

圖 5-20 ARP 運作方式

不論 ARP 的 Request 和 Reply，或 RARP 之 Request 和 Reply 都使用 ARP 封包格式。ARP 封包格式如圖 5-21 所示，各欄位功能如下：

- **Hardware Type**：表示發送主機使用之網路實體介面種類，如 1 表示 Ethernet 網路介面。
- **Protocol Type**：表示所使用的通訊協定，如 0x0800 表示 IP 協定，其它通訊協定模式如表 5-1 所示。
- **Operation Type**：表示此封包的工作模式：
  - **1** → ARP 要求（ARP Request）
  - **2** → ARP 回應（ARP Reply）
  - **3** → RARP 要求（RARP Request）
  - **4** → RARP 回應（RARP Reply）
- **HLEN**：網路介面卡硬體位址長度。若 Ethernet 位址的長度為 6。
- **PLEN**：網路協定位址長度。因為 IP 位址長 4 個位元組，此值為 4。
- **Sender HW**：發送端的硬體位址。如果是 Ethernet 網路的話，此為 6 個位元組長的地址，如 0x8823AA112233。
- **Target HW**：目的地的硬體位址。
- **Sender IP**：發送端的 IP 位址，如 163.15.2.1。

- **Target IP**：目的地主機的 IP 位址，如 163.15.2.4。

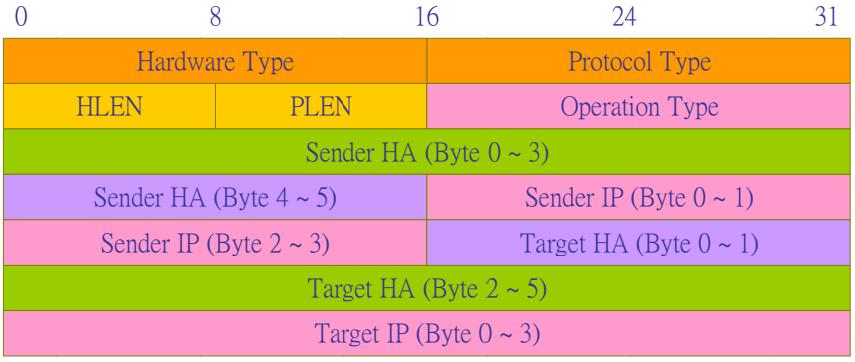

圖 5-21 ARP 封包格式

圖 5-22 為 RAP 封包的擷取結果，此為主機要求連線時 (執行 telnet 163.15.2.62)，並不知道對方的 Ethernet 位址，所廣播的 ARP Request 訊息的封包格式。由圖中可觀察 Ethernet 訊框標頭的 Type 欄位為 0x0806，表示為 ARP 訊框包裝。

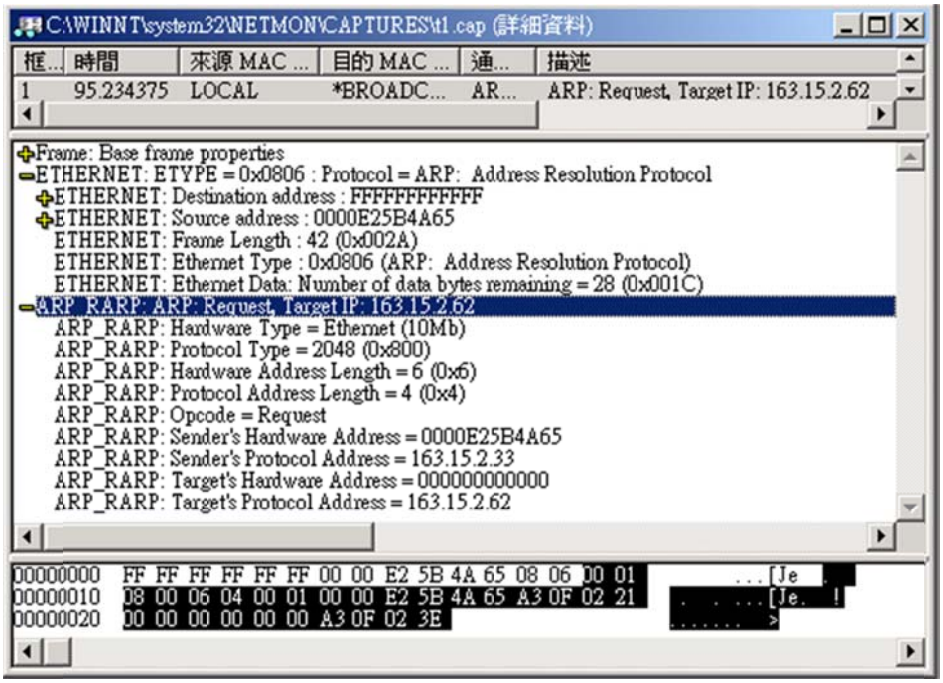

圖 5-22 擷取 ARP-Request 封包

並非主機每次欲發送封包之前，都要詢問對方的網路硬體位址，一般主機上的快取記憶體 (Cache memory) 上，都有維護一只**『ARP 快取表』(ARP cache，由 ARP 協定建立)**。當主機電腦欲依照 IP 位址發送封包時，首先到快取記憶體上的 ARP 索引表查詢是否有相對應之硬體位址，如沒有再發送 ARP Request 到網路上詢問。詢問後將相關的 IP 和硬體位址登錄在索引表上，下次可以再使用。相對應的，網路上任何一部電腦接收到 ARP Request 封包時，不管該封包是否詢問自己，在 ARP 封包內都有紀錄傳送者的 IP 位址和硬體位址 (MAC)，便將其 IP 位址和硬

體位址紀錄在本身的索引表內。又在 Internet 網路上的任何主機啟動時，都會公佈（announce）ARP 訊息告訴網路上所有工作站自己的 IP 和硬體位址。快取記憶體上的索引表必須隨時更新，否則所紀錄的資料也許會失去時效性（也許主機更換 IP 位址），這必須由系統維護人員去設定自動更替時間（通常為 20 分鐘）。主機電腦（或路由器）可以利用 arp 命令來查詢 ARP 快取表的內容，查詢結果如下：

```
[root@linux-1 /sbin]# arp
Address         HWtype      HWaddress          Flags Mask          Iface
163.15.2.30     ether       00:00:E2:5B:4A:65     C                 eth0
192.168.0.1     ether       00:C0:02:25:46:18     C                 eth1
163.15.2.34     ether       00:00:E2:2E:A3:76     C                 eth0
```

一般在跨接網路上，網路之間所連接的路由器都具有『**代理 ARP**』（**Proxy ARP**）的功能，讓路由器代理回應其它網路上詢問他所管轄之工作站的 MAC 位址。但這會欺騙 ARP 發送端，以為路由器就是目的主機，而事實上目的主機位於路由器的另一端。路由器扮演了目的主機的代理身份，回應 ARP Response 給詢問端。如圖 5-23 中，工作站 A 欲詢問工作站 B 的 MAC 位址，於是廣播 ARP Request (163.15.3.50) 到網路上，Router_A 收到該訊息，便回應 ARP Response 給工作站 A，當然，Router_A 必須事先知道工作站 B 的 MAC 位址。如果 Router_A 不知道工作站 B 的 MAC 位址，則它必須將該詢問訊號收取，再另外發送 ARP Request (163.15.3.50) 到另一網路上詢問，工作站 B 回應訊號給 Router_A，Router_A 取得它的 MAC 位址後，再回應給工作站 A（如圖 5-23 中的訊號傳遞順序）。

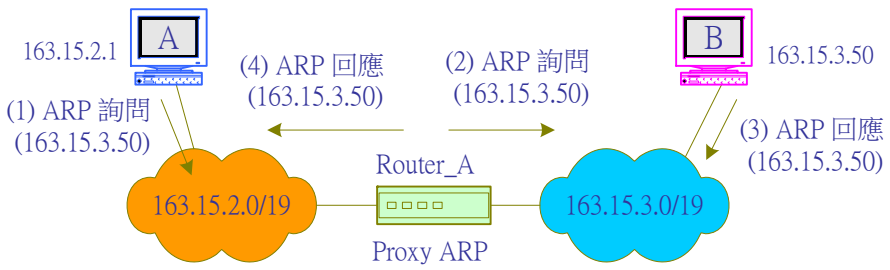

圖 5-23 代理 ARP 之功能

5-3-2 RARP 協定

一般主機電腦上的 IP 位址設定有靜態設定和動態設定兩種方法。靜態設定是於主機電腦上直接用人工設定 IP 位址，設定後如沒有再用人工修改，則 IP 位址永遠不變；動態設定並未設定 IP 位址，每當主機啟動時，再由網路上某部伺服器給予 IP 位址，因此，每次啟動時得到的 IP 位址

不一定相同。如果主機電腦採用動態指定 IP 位址模式，啟動就必須利用『反向位址解析協定』（Reverse Address Resolution Protocol, RARP），向網路上的伺服器（如 DHCP Server）要求給予一個 IP 位址。

RARP 的動作是，主機電腦用自己硬體位址（MAC 位址）向伺服器詢問自己的 IP 位址，如圖 5-24 所示。圖中假設各主機皆以動態設定 IP 位址，主機 A 為『動態主機組態伺服器』（Dynamic Host Configuration Protocol, DHCP），負責分配 IP 位址給網路上主機。電腦主機啟動時便要求 DHCP 伺服器分配一個 IP 位址，電腦主機關機時便釋放該 IP 位址，下次開機時再要求重新分配。例如主機 C 啟動時立即在網路廣播（或傳送給 DHCP 伺服器） RARP Request（要求 IP 位址），主機 A 收到 RARP Request 並驗證其 Ethernet 位址是否可以給予 IP 位址，再回應 RARP Reply 給主機 C，其中並攜帶著對方的 IP 和 MAC 位址。RARP 封包格式和 ARP 一樣，不再另述。

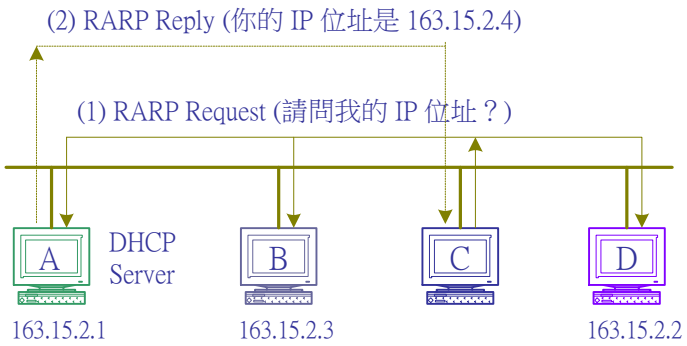

圖 5-24 RARP 運作方式

5-4 ICMP 通訊協定

根據我們的瞭解 IP 網路是一種不可靠的傳輸方式，傳送中的封包必須經過多層路由器的轉送才能到達目的地，因此，在發送封包之前，我們很難預測該封包是否可以安全到達目的地。我們也很迫切地想知道目前網路的狀況，尤其在傳送失敗時，更想瞭解問題出在什麼地方。Internet 網路中提供一種稱之為『網際控制訊息協定』(Internet Control Message Protocol, ICMP)的通訊軟體，用來偵測網路的狀況。在 IP 網路上，任何一部主機或路由器皆設置有 ICMP 協定，它們之間就可以利用 ICMP 來互相交換網路目前的狀況訊息，例如，主機不存在、網路斷線等等狀況。ICMP 訊息的產生有下列兩種情況：

(1) 障礙通知：當 IP 封包傳送當中，在某一網路上發生問題而無法繼續傳送，則會回應 ICMP 訊息給原封包傳送端。如圖 5-25 所示，訊號_1 是由 Router_A 回應；或是由 Router_B 回應訊號_2；也有可能是由主機 B 回應訊號_3。

(2) 狀況查詢：可以發送 ICMP 來查詢目前網路的情況。如圖 5-26 中，主機 A 發送 ICMP 查詢訊息，有可能由路由器回應（訊號_1 和 訊號_2），或由主機 B 回應訊號_3。

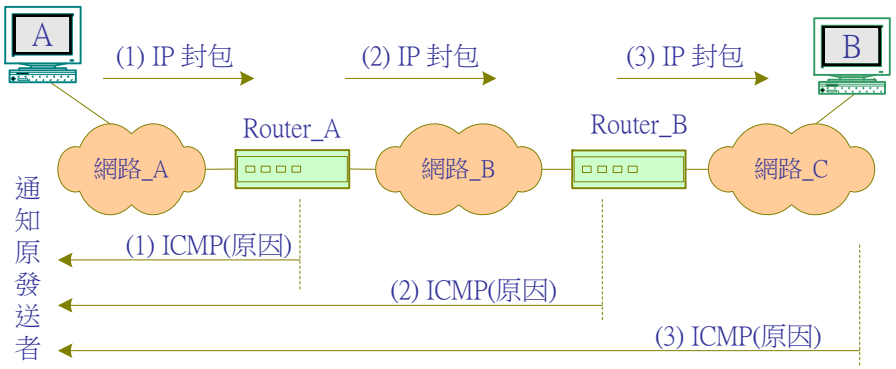

圖 5-25 ICMP 障礙通知

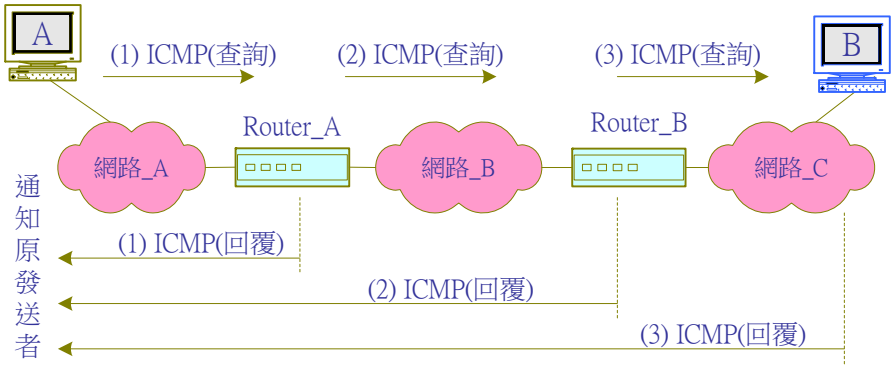

圖 5-26 ICMP 網路狀況查詢

ICMP 封包無法直接傳送到網路上，必須如同 TCP 封包一樣被嵌入 IP 封包內（如表 5-1），以 IP 方式傳送，包裝在 IP 內的封包格式，如圖 5-27 所示。

圖 5-27 ICMP 封包嵌入 IP 封包內傳送

ICMP 封包的長度並不固定，隨著各種訊息型態而有不同的長度，圖 5-28 為 ICMP 封包格式，其各欄位功能如下：

- **訊息型態（Message Type）：**表示該 ICMP 所欲控制之訊息型態，共有 13 種型態，訊息型態之型態代表值如表 5-2 所示。
- **編碼（Code）：**對各種訊息型態進一步說明工作內容。
- **檢查集檢查碼（Checksum）：**對該封包檢查集錯誤偵測。
- **訊息說明（Message description）：**依照不同的控制訊息，而有不同的說明方式。
- **訊息資料（Message Data）：**依照不同的控制訊息，而有不同的資料表示。

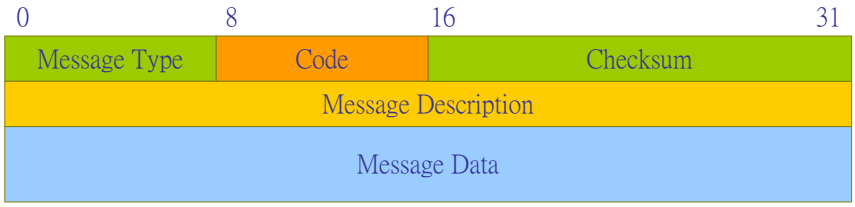

圖 5-28 ICMP 封包格式

表 5-2 ICMP 訊息型態

Message Type	ICMP 訊息功能
0	Echo Reply（回應答覆）
3	Destination Unreachable（目的地無法到達）
4	Source Quench（來源抑制）
5	Redirect（改變傳輸路徑）
8	Echo Request（回應要求）
9	Router Advertisement（路由器宣傳）
10	Router Solicitation（路由器懇請）
11	Time Exceeded for a Datagram（溢時傳輸）
12	Parameter Problem on a Datagram（參數問題）
13	Timestamp Request（時間標籤要求）
14	Timestamp Reply（時間標籤回覆）
15	Information Request（資訊要求）（停用）
16	Information Reply（資訊回覆）（停用）
17	Address Mask Request（位址遮罩要求）
18	Address Mask Reply（位址遮罩回覆）

以下分別說明各種不同型態的控制訊息：

5-4-1 回聲要求/回聲回應

『回聲要求』(Echo Request)(Type 8) 是用來要求對方回聲，如有『回聲回應』(Echo Reply) (Type 0) 表示對方主機或路由器工作正常。也可以用來測試網路路徑是否確實可以到達，被測試端用 Echo Reply 之 ICMP 封包答覆對方的回聲要求。TCP/IP 的主機電腦 (或路由器等設備) 上皆有提供 "ping" 指令，用來實現 Echo Request 命令。例如，在主機電腦上執行 ping 163.15.2.1 指令，當 163.15.2.1 之網路設備接收到該訊號時，必須立即回應 Echo Reply，而在主機電腦上會顯示經歷時間，或主機電腦在溢時之後，未收到 Echo Reply 訊號，則會顯示 Request timed out 訊息。被測試端和測試端之間也許經過多個路由器 (或網路閘門)，但如果收到 Echo Reply 訊號，則表示該路徑是可到達的。以下為執行 ping 命令的結果範例：

```
[root@linux-1 /sbin]# ping 192.168.0.1
PING 192.168.0.1 (192.168.0.1) from 192.168.0.50 : 56(84) bytes of data.
64 bytes from 192.168.0.1: icmp_seq=0 ttl=64 time=1.8 ms
64 bytes from 192.168.0.1: icmp_seq=1 ttl=64 time=1.6 ms
64 bytes from 192.168.0.1: icmp_seq=2 ttl=64 time=1.6 ms
64 bytes from 192.168.0.1: icmp_seq=3 ttl=64 time=1.4 ms
--- 192.168.0.1 ping statistics ---
4 packets transmitted, 4 packets received, 0% packet loss
round-trip min/avg/max = 1.4/1.6/1.8 ms
```

圖 5-29 為 Echo Request 及 Echo Reply 之封包格式，其中 Identifier (識別碼) 和 Sequence Number (順序號碼) 是用來檢查回聲回應封包是針對哪一個回聲要求所產生的。一般我們可以連續發送多個 Echo Request 給被測試端，每個回聲要求都給予一個順序編號放置在 Identifier 內，被測試端發送 Echo Reply 時將 Echo Request 內之 Identifier 的值放置於 Sequence Number 欄位內，告訴發送端是針對哪一個 Echo Request 的回聲回覆。另外，封包內之 Option Data 可有可無，如 Echo Request 有放置資料，回覆 Echo Replay 就依照該資料送回。

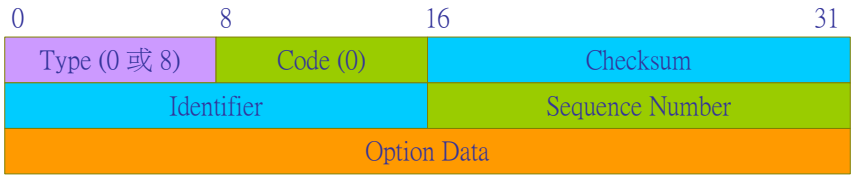

圖 5-29 Echo Request/Reply 封包格式

圖 5-30 為執行 "ping 163.15.2.62" 所擷取封包的結果，由圖中可以看出 ICMP 封包是包裝在 IP 封包內（Protocol = ICMP 或 = 01），而 IP 封包也包裝在 Ethernet 訊框內（Type = 0x0800）。ICMP 之 Packet Type = Echo（8）、Code = 0。

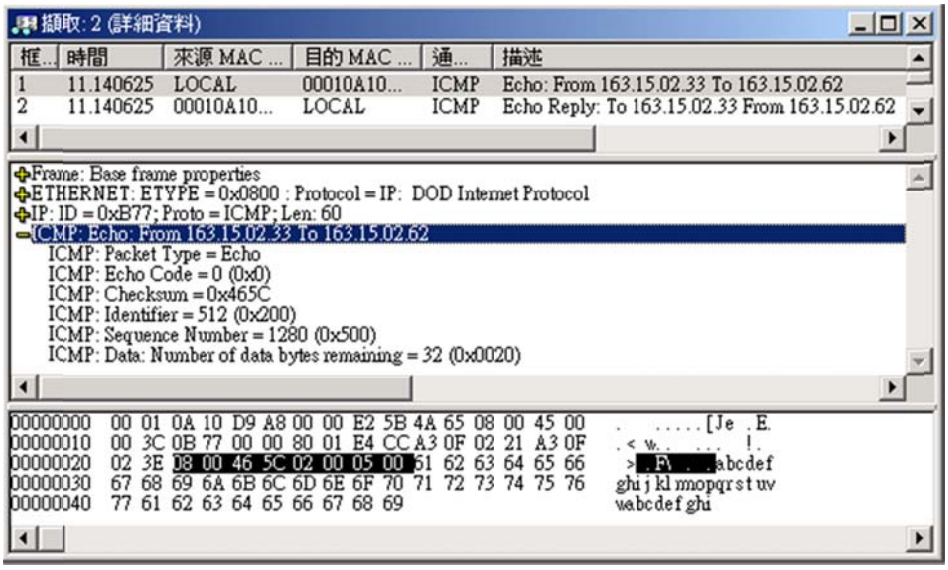

圖 5-30 擷取 ICMP Echo 之封包內容

5-4-2 目的地無法到達

當路由器（或網路閘門）發現，某一個 IP 封包無法往下一個路徑傳送時，便發送 Destination Unreachable 之 ICMP 封包（Type 3）給原始發送該封包者，並將封包丟棄。因此，在回應封包之 Message Data 欄位內必須註明是哪一個 IP 封包（原 IP 封包表頭和其資料前 8 個位元組）被丟棄，如圖 5-31 所示。又另封包內之 Code 欄位註明無法到達目的地的原因，其原因分類如下：

- **0:** Network Unreachable（無法到達目的網路）
- **1:** Host Unreachable（無法到達目的主機）
- **2:** Protocol Unreachable（通訊協定不存在）
- **3:** Port Unreachable（無法到達連接埠）
- **4:** Fragmentation Needed and DF set（資料需分割並設定不可分割位元）
- **5:** Source Route Failed（來源路徑選擇失敗）
- **6:** Destination Network Unknown（無法識別目的地網路）
- **7:** Destination Host Unknown（無法識別目的地主機）

- **8:** Source Host Isolated（來源主機被隔離）
- **9:** Communication with Destination Network Administratively Prohibited（管理上禁止和目的地網路通訊）
- **10:** Communication with Destination Host Administratively Prohibited(管理上禁止和目的地主機通訊）
- **11:** Network Unreachable for Type of Service（無法到達此型態的網路服務）
- **12:** Host Unreachable for Type of Service（無法到達此型態的主機服務）

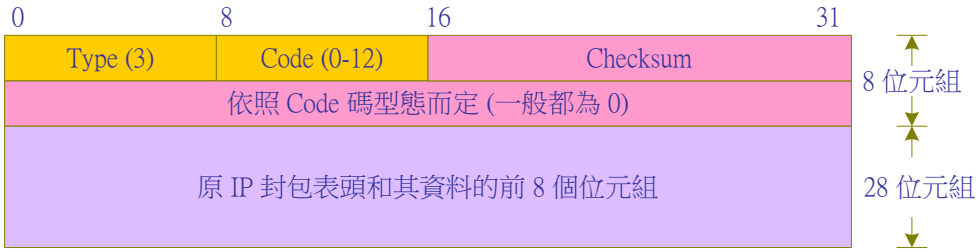

圖 5-31 Destination Unreachable 封包格式

ICMP 目的無法到達訊息在 Internet 網路上扮演非常重要的角色，一般網路偵測各種異常狀況都會使用到它，我們以下列幾種較常用的訊息來介紹，其它未介紹之訊息的作業方式也大同小異。

(1) 無法到達目的網路或主機（Code 0/1）

當 IP 封包被轉送到一部路由器（或主機），而該路由器（或主機）在路由表無法找到適當的路徑傳送時，如果是路由器，便傳送『**無法到達網路**』之訊息；如果是主機，則發送『**無法到達主機**』之訊息給該封包的原發送端。

(2) 通訊協定不存在（Code 2）

當路由器（或主機）接收到某一封包，當拆封該封包時，發現其格式並不符合 IP 封包格式，便發送『**通訊協定不存在**』的訊息回原發送端。所發生的原因可能是封包在傳送過程中已被損壞，或不同網路之間封包格式轉換錯誤。

(3) 無法到達連接埠（Code 3）

當主機接收到某一封包，該封包所承載之傳輸層（TCP 或 UDP）所指定的埠口（Port）不存在（該埠口並未啟動執行），則回應『無法到達連接埠』訊息給原發送端。如果是 TCP 訊息，則回應如圖 5-32 之封包結構，其中原錯誤封包的 IP 表頭（如圖 5-10 所示）佔用 20 個位元組，另 8 個位元組為 TCP 表頭的前 8 位元組資料，原發送端收到該 ICMP 封包，便可由訊息欄位中得知哪一個埠口無法連接上。圖 5-33 是回應 UDP 訊息的格式，其中前 8 個位元組資料剛好是 UDP 的表頭。

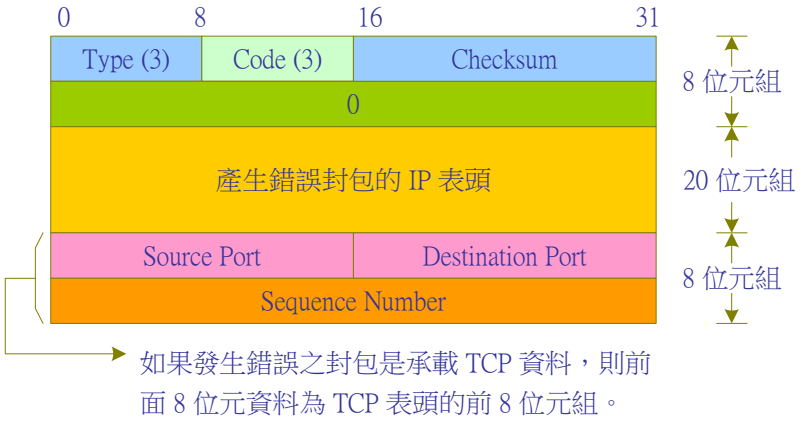

圖 5-32 無法到達連接埠（承載 TCP 表頭）

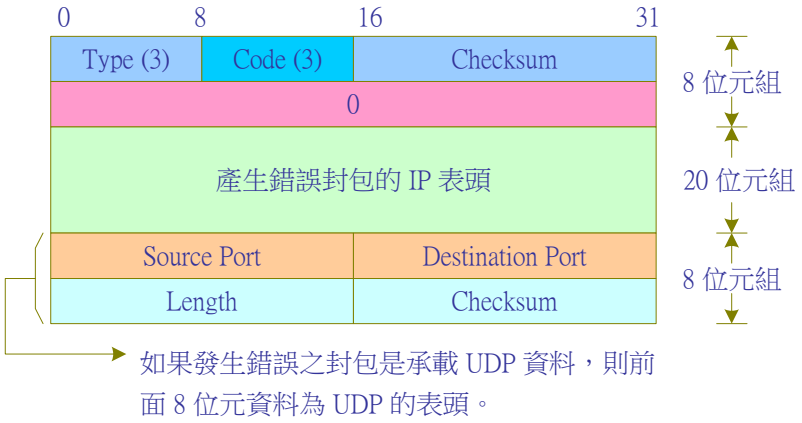

圖 5-33 無法到達連接埠（承載 UDP 表頭）

(4) 資料需分割但已設定不可分割位元（Code 4）

如路由器接收到一個需要分段的封包（MTU 太大），而該封包上已設定 DF（不可再分段），則會傳送一個 ICMP Fragmentation Needed and DF set 之訊息給原發送端，而 ICMP 封包上會標明下一個路徑的 MTU，請原發送端依照該 MTU 來分段封包。ICMP 封包格式如圖 5-34 所示。

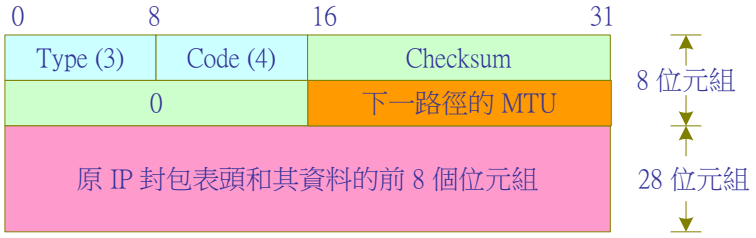

圖 5-34 ICMP Fragmentation Needed and DF set 封包格式

(5) 來源路徑選擇失敗（Code 5）

路由器接收到 IP 封包上有指定嚴格來源路徑選項（如 5-2-7 節），但所指定之下一路由器無法到達時，則發送 ICMP Source Route Failed 給原發送端。

5-4-3 來源抑制

當網路上路由器（或網路閘門）對於進入的封包速度過快，而來不及處理時（本身的緩衝器已溢滿），這時候會將再進入的封包丟棄，並發送一個來源抑制（Source Quench）之 ICMP 封包（Type 4）給該 IP 封包的來源主機。原發送主機接收到 Source Quench 封包後會暫停或降低發出封包的速度，一直到沒有接收到路由器所發送的 Source Quench 後，再恢復原來的發送速度。Source Quench 之封包格式如圖 5-35 所示。

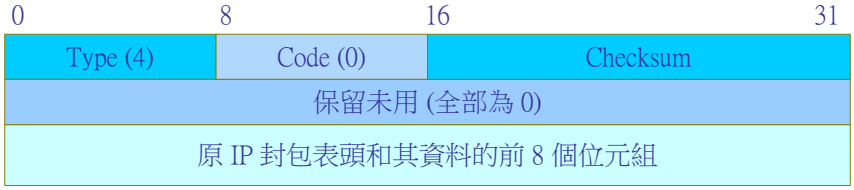

圖 5-35 Source Quench 之封包格式

5-4-4 傳輸路徑改變

在某些情況下（如網路拓樸圖改變或網路斷線），路由器發現原來由它轉送的目的位址，經由別的路由器轉送會比較快速，則該路由器便利用 ICMP Redirect 訊息（Type 5）通知原發送端，請它改變傳輸路徑，傳送到另一個路由器上。圖 5-36 為 Redirect 的封包格式，其中 "Gateway IP Address" 表示新轉向傳輸路徑的路由器之位址，Code 欄位表示轉向的原因：

- **0:** Redirect Datagram for the Net（網路變更而轉向）
- **1:** Redirect Datagram for the Host（主機變更而轉向）

- **2:** Redirect Datagram for the Type of Service and Net（網路和服務型態變更而轉向）
- **3:** Redirect Datagram for the Type of Service and Host（主機和服務型態變更而轉向）

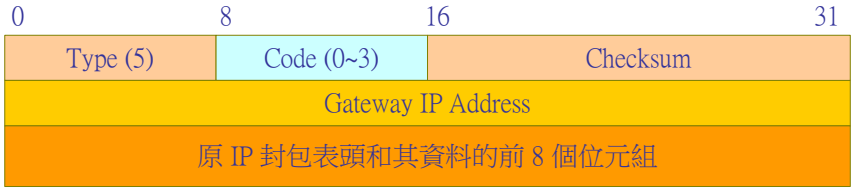

圖 5-36 Redirect 之封包格式

5-4-5 逾時傳輸

當路由器（或網路閘門）發現某一個 IP 封包內之 TTL（Time to Live）為 0 時，會發送一個溢時傳輸（Time Exceeded）封包（Type 11）給原發送封包者，並將該封包丟棄。為了避免封包在網路上不停的環繞，封包發送之前會在 TTL 欄位設定一個值，表示該封包的存活時間。封包在傳送後，每經過一個路由器，就將 TTL 內的值減一。如果路由器發現某個封包內的 TTL 值為 0，便判斷該封包已在網路上環繞，找不到適當路徑到達目的地，而將其丟棄。另外一種情況，接收端也會發送 ICMP time exceeded 訊息給來源端，此情況是資料片重組（Fragment reassembly）發生問題。當 IP 封包被分為若干個資料片段傳送時，接收端收到一筆資料片段時會設定一個計時時間，如在溢時(timeout)後未收到下一筆資料片段，則將發送 ICMP time exceeded 給原傳送端。圖 5-37 為 Time Exceeded 之封包格式，其中 Code = 0 表示 TTL = 0 時發送該封包；而 Code = 1 表示資料片段重組失敗而發送封包。

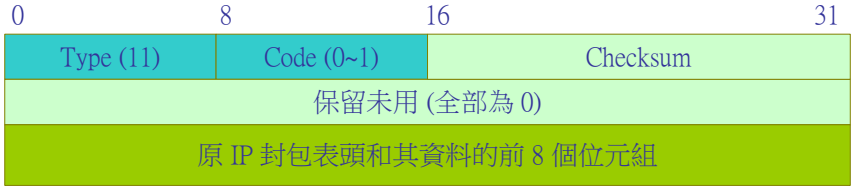

圖 5-37 Time Exceeded 之封包格式

5-4-6 參數問題

當路由器（或網路閘門）發現 IP 封包內的某些欄位的值（參數）不正確，而無法處理該封包時，便發送參數問題（ICMP parameter problem）封包（Type 12）給原發送者。封包格式如圖 5-38 所示，其中 Pointer 欄位表示錯誤參數之位址。

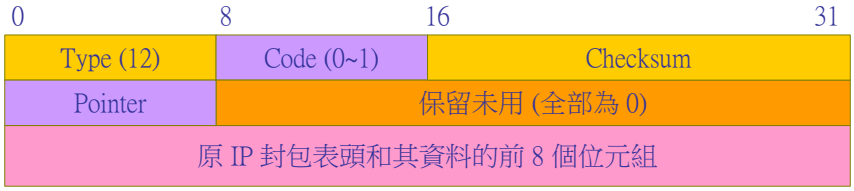

圖 5-38 Parameter Problem 之封包格式

5-4-7 時間訊息要求及回覆

時間要求訊息（ICMP timestamp request）封包（Type 13）是用來詢問某部主機的系統時間；而時間訊息回應（ICMP timestamp reply）封包（Type 14）則用來回應系統時間，這兩個封包是用來使網路上各設備的時間達到同步。圖 5-39 為 Timestamp 的封包格式，其中 Originate Timestamp 表示詢問者自己的時間、Receive Timestamp 表示被詢問者接收到 ICMP timestamp request 封包的時間、Transmit Timestamp 為被詢問者發送回應的時間。前一項為詢問者的時間基準；後二項為被詢問者的時間基準。時間單位都是 millisecond，並以格林威治時間為基準。

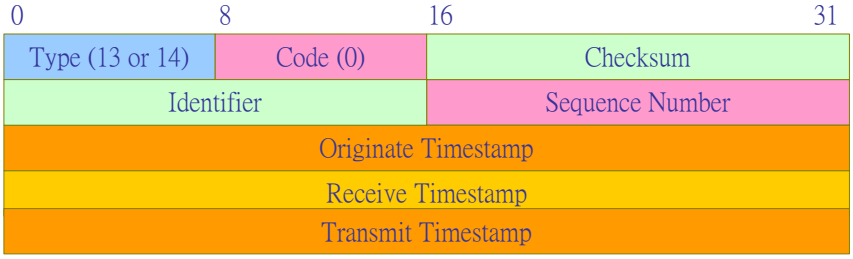

圖 5-39 Timestamp Request/Reply 之封包格式

5-4-8 遮罩位址要求及回覆

遮罩位址要求及回覆（ICMP address mask request/reply）（Type 17/18）是被用來要求獲得或回應某一個次網路（Subnet）的位址遮罩（mask）時所使用。圖 5-40 為 Address Mask Request/Reply 的封包格式。

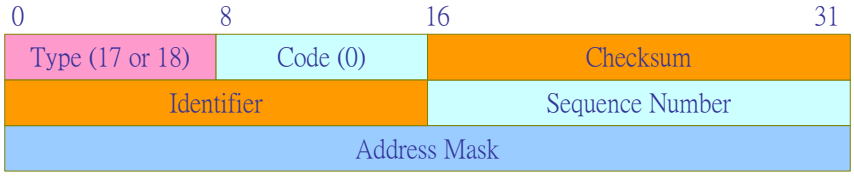

圖 5-40 Address Mask Request/Reply 之封包格式

5-5 IGMP 通訊協定

『網際群組管理協定』(Internet Group Management Protocol, IGMP) (RFC 1112) 是作為管理 Internet 網路群組使用，譬如，主機或路由器如何加入某一群組或退出群組。當主機或路由器屬於某一群組後，便可依照群組位址向群組中的成員作多點傳送，如不屬於群組成員就無法收到該訊息。我們在 5-2-5 節有介紹過 IP 多點傳送是利用 Class D 位址格式，而 IGMP 是管理網路群組中，有關 Class D 位址的設定及分配問題。簡單的說，我們將一些主機或路由器組成一個群組，並給予一個 Class D 的位址，如果某一主機向該群組傳送資料時，只有該群組的成員可以接收到訊息，也稱之為『多點傳送』。基本上，每一個主機都有一個獨一無二的 IP 位址，如何再增加一個網路群組位址 (Class D)，則需要利用 IGMP 協定來管理群組之中的成員。

5-5-1 多點傳送位址

依照 IP 分級的歸類，Class D 等級位址為多點傳送位址，它的位址格式如圖 5-41 所示。

圖 5-41 Class D IP 多點傳送位址

如以一般 IP 位址表示法，Class D 的位址範圍是：

224.0.0.0 ~ 239.255.255.255

一般應用上會將 224.0.0.0 保留而不分配給任何群組。另外 224.0.0.1 為特殊群組，表示網路上任何主機或路由器都屬於該群組的成員，雖然如此定義這個群組的成員，已非常接近廣播位址 (所有成員)，其間仍有差別，如果主機或路由器沒具有接收群組位址的功能，它仍不會收到目的地為 224.0.0.1 的訊息。

一般 IP (Class A ~ Class C) 欲被發送到網路上之前，在資料連結層 (Ethernet 層)，必須將 IP 位址轉換成 Ethernet 位址，也許會透過 ARP 協定來詢問該 IP 位址的相對應 Ethernet 位址。但如果 IP 位址為 Class D 群組位址時，便無法透過 ARP 協定來詢問相對應的 Ethernet 位址，這應如何來達成多點傳送之目的呢？如果主機或路由器的網路卡沒有特殊的處理能力，當它準備發送多點傳送的 IP 位址時，它的 Ethernet 層便將目的位址 (Ethernet 位址) 設定為廣播位址 (全部都為 1)，讓網路上所有主機都接收到該訊框，再由上一層的通訊軟體來決定是否接受該訊息。雖然這種做法可以達到多點廣播的目的，但是整個網路的負荷會增加許多。譬如，一個視訊伺服器

（VoD）欲廣播視訊給網路上 20 個收視的主機，但當時連線的主機有 100 部，則 100 部主機都必須接收該訊息，再由上層通訊軟體決定是否接收。

為了達到 Ethernet 層有多點傳送的功能，我們可以利用 Ethernet 網路的群組位址，其為 0x01.00.5E.00.00.00（48 位元）。而將 IP 多點傳送位址植入 Ethernet 群組位址上，植入的方法是將 IP 多點傳送位址的最後 23 位元，放置到 Ethernet 位址的後 23 位元內，如以 224.0.0.1 的位址對應到 Ethernet 位址是 0x01.00.5E.00.00.00.01，其對應方法如圖 5-42 所示。

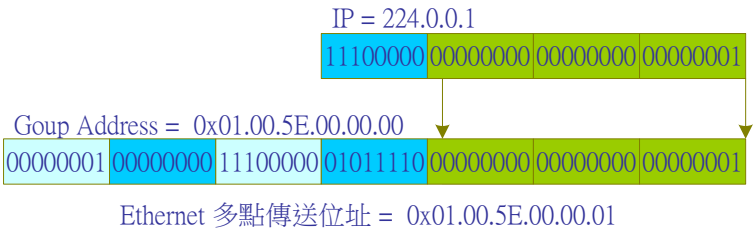

圖 5-42 IP Class D 位址植入 Ethernet 群組位址

這種植入法也有先天的缺點，多點傳送位址有 28 位元（32 - 4）的變化，而 Ethernet 位址只能植入 23 位元，如果超過 23 位元的地址便無法被植入 Ethrnet 位址上，這可能會有不同的多點傳送位址而產生相同的 Ethernet 群組傳送位址。為了解決這個問題，我們儘量將 IP 多點位址設定在低位元組，在一般網路上的群組位只要超過 23 位元的表示量，這種機會可以說非常渺小，因此在使用上應該不會發生問題。

5-5-2 主機分類

在多點傳送的網路環境裡，我們可以將參與主機（或路由器）區分為下列三個等級：

- **等級 0（Level 0）:** 此等級的主機沒有能力傳送或接收 IP 多點傳送封包。也有可能是該主機不參與網路任何群組的成員，當它決定參與群組時，可能會升級到等級 1 或 2。
- **等級 1（Level 1）:** 此等級的主機可以傳送但無法接收 IP 多點傳送封包。一般此等級的主機或路由器都是週期性的廣播訊息給同一群組的成員，而它的功能也比較簡單，只要具有能將 IP 群組位址轉換成 Ethernet 的群組位址即可。
- **等級 2（Level 2）:** 此等級主機具有傳送和接收 IP 多點傳送封包的功能。如具有接收 IP 多點傳送的功能則較為複雜，因為主機上的應用程式隨時都有可能加入或退出某一群組，主機必須能隨時決定是否接收某一群組的訊息。

一般網路上有參與群組工作的主機都具有等級 2 的能力。我們從另一觀點來看，主機是否參與某一群組是由它的應用程式（或稱為處理程序）來決定。也就是說，某一主機上也許會有許多應用程式各自參與不同群組，或者某一應用程式同時成為若干個群組的成員，因此，一個主機也許會屬於多個群組的成員，也可能隨時退出某一群組，這完全決定於該主機的應用程式。所以一般主機都必須維護一個多點傳送的表格，以紀錄當時每一個多點傳送位址有哪些應用程式銜接，或退出哪一群組。當主機接收到 IP 多點傳送封包時，再查閱該群組位址是否有應用程式使用。至於網路上主機如何要求加某一群組，或退出某一群組，這必須由參與群組的成員共同來協調，它們之間就必須利用 IGMP 協定來互相傳送訊息來從事協調的工作。

5-5-3 IGMP 封包結構

IGMP 也是屬於網際層的通訊協定之一，傳輸方式如同 ICMP 一樣，被包裝在 IP 封包內傳送。但 IGMP 有固定大小的封包，並沒有選擇性的訊息資料。IGMP 的封包為 8 位元組，再加上 IP 標頭 20 位元組，合計有 28 位元組，如圖 5-43 所示。

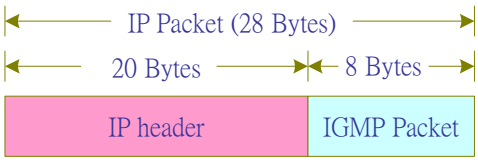

圖 5-43 IGMP 嵌入 IP 封包中

圖 5-44 為 IGMP 封包格式，其中各欄位功能如下：

- **版本（Version）：**目前版本都為 1。
- **型態（Type）：**IGMP 的封包型態區分為兩種：
  - ★ **查詢（Inquiry）：**（Type = 1）由多點傳送路由器（或主機）發送的查詢訊息。
  - ★ **回應（Response）：**（Type = 2）由主機傳送的回應訊息。
- **標頭檢查集（Header Checksum）：**封包標頭的檢查集。
- **群組位址（Group Address）：**是一個 Class D 的 IP 位址。當 IGMP 封包為查詢時（Type = 1），該欄位內容為 0。

圖 5-44 IGMP 封包格式

5-5-4 IGMP 協定運作

IGMP 協定的主要功能是作為網路上路由器 (或主機) 之間協調如何加入某一群組，或由某一群組中退出，當路由器成功加入群組之後，便可依照群組位址和同一群組的成員相互通訊，而達到多點傳送的目的，以下介紹 IGMP 協定的運作情形。

(A) 加入一個多點傳送群組

加入多點傳送群組的基本概念，是讓處理程序 (Process) 加入主機給予的特定介面。一個群組位址給予一個特定介面，給予處理程序隨時加入或退出，在一部主機上也許維護多個群組界介面，而一個處理程序也可能加入多個群組。此群組介面是讓處理程序和群組中的成員連結在一起，當程序加入群組介面後，或許會由介面上接收訊息或發送多點傳送訊息，這些訊息都必須經過 IP 層包裝或拆裝成 IP 多點傳送封包。因此，在主機系統裡必須提供群組傳送的 API 介面，以提供多點傳送連接，一般在 Unix 系統裡都有提供多點傳送的 Socket 介面，其連結情形如圖 5-45 所示。

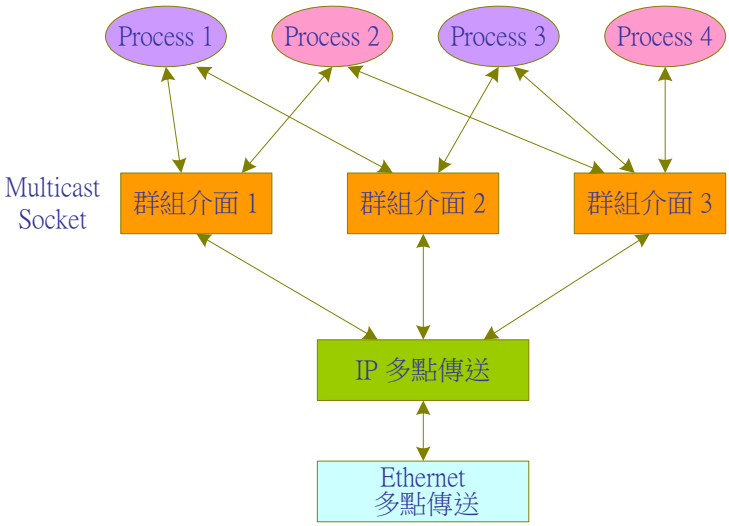

圖 5-45 處理程序加入群組之連結情形

由圖中可以看出，主機是藉由群組位址和 Socket 介面來辨識群組，因此，主機必需維護一個包含所有群組(至少有一個處理程序屬於該群組成員)的表格，以及屬於該群組之處理程序的參考數量。

(B) IGMP 報告與查詢

IGMP 訊息被多點傳送路由器用來追蹤網路上各路由器連接群組的情形，它的運作規則如下：

(1) 當第一個處理程序加入群組時，主機傳送一個 IGMP Response 報告。假如有多個處理程序加入同一群組，只需送出一份報告即可。這個報告被送到的介面與處理程序所加入的群組介面相同（如圖 5-46 所示）。

(2) 當處理程序離開群組時，主機不會送出報告（IGMP Response），即使是最後一個成員也是一樣。如果主機知道給定的群組中沒有成員，而收到 IGMP Inquiry 查詢時，也不會向該群組報告（不會送出 IGMP Response）。

(3) 一個多點傳送路由器（或主機），會在固定時間週期內送出 IGMP Inquiry 查詢，以查看是否有任何主機仍然有屬於任何群組的處理程序。路由器必需針對每一個介面傳送查詢訊息，查詢中的群組位址為 0（圖 5-44 中的 Group Address 欄位），因為路由器希望每個群組（在主機上有一個或數個成員）的主機都有一個回應。

(4) 主機收到 IGMP Inquiry 訊息後，以 IGMP Response 回應給群組，也表示所回應之群組至少還有一個以上的處理程序連接。

多點傳送路由器必需維護一個記錄著多點傳送群組中擁有一個或數個主機介面的表格，當路由器收到一個必須往前送（Forward）的多點傳送封包，它只將該封包往前送到下一個介面，該介面仍有主機和屬於該群組的處理程序。圖 5-46 為 IGMP Inquiry 和 IGMP Response 訊息發送情形，IGMP Inquiry 是路由器用來查詢主機上群組登錄的情形，而 IGMP Response 是主機用來回應路由器的查詢。

圖 5-46 IGMP 協定運作情形

(C) 存活時間（TTL）

多點傳送封包上的存活時間（Time-to-Live, TTL）顯得特別重要，因該欄位是在 IP 封包標頭上（請參考圖 5-10）。TTL 上的值乃決定多點傳送封包可跨越多少網路傳送，TTL =1 表示此封包只在同一子網路上傳送。路由器收到封包時，由 TTL 上的值決定是否往其它子網路傳送，如果（TTL –1）的值大於 0，則往下一個子網路作多點傳送。一般為了限制封包的傳輸量，都不會將 TTL 的值設定太大，應用實際環境下，同一群組成員也不會位於跨越太多的網路上。當路由器判斷 TTL = 0 時，就不會回應 ICMP 封包給原發送端。

(D) IGMP 協定實作

為了提高 IGMP 協定的執行效率，在實作上必須增加下列細節：

(1) 當主機傳送一個初始的 IGMP 報告時，並不保證該報告一定會被送達（因 IP 是電報傳輸），通常主機都會等待 0 ~ 10 秒的隨機時間，再傳送另一個報告。

(2) 當主機收到來自路由器的查詢，並沒有立即回應，而把回應排在最後（或等待一個隨機延遲時間）。當它發現有相同群組的成員回應給路由器，便可省略而不予回應，由圖 5-46 中可看出主機回應報告的目的位址為群組位址，因此，同一群組的成員也可以收到其它主機的回應訊息。

如圖 5-46，路由器所發送查詢訊號的目的位址為網路上所有主機和路由器（IP = 224.0.0.1），這也表示任何有群組介面的主機都可以給予回應，如果大家都即時回應恐造成封包風暴。由另一方面，對於多點傳送路由而言，乃希望知道所管轄的群組是否還有成員存在，至於多少成員並不重要，因此無需每一個主機都回應。

5-6 IPv6 通訊協定

自從 1990 年網際網路風行之後，網路專家們漸漸感覺到所使用的 IP（IPv4，Version 4）通訊協定已不敷使用。尤其隨著應用層次的提高，IP 網路不再只須提供檔案傳送及遠端登入等簡單的應用，更必須進一步處理有關資料庫系統的查詢與更新，也進入電子商務的應用。因此，有必要再發展出功能更完整的通訊系統，**『IPv6』（IP Version 6）**就在這迫切需求之下被發展出來。首先，我們用下列幾點來介紹 IPv6 如何彌補 IPv4 的不足：

(A) IPv6 提供更寬廣的 IP 定址空間

IPv4 用 32 位元（4 個位元組）來表示 IP 位址空間，已漸不能滿足目前網際網路上主機電腦連結的成長。IPv6 提供 128 位元（16 個位元組） 的空間來表示 IP 位址。

(B) IPv6 提供認證服務

一般 IP 封包在傳送過程中，每經過一個中途路由器就必須被拆裝和重新包裝，又要經過許多不可預測的路由器（無法事先預估），而且為了要使封包能順利到達目的地，在封包上又無法做太多的保護措施。正因如此，任何有心人士，皆可輕而易舉在網路上窺視他人資料，甚至竄改它，使傳送的訊息失去正確性。尤其在企業內的區域網路連接更為困擾，早期，跨區域的區域網路連接都透過專線，但為了節省費用及符合出差人員方便連線，目前都希望能直接透過網際網路連線，如**『虛擬私人網路』（Virtual Private Network, VPN）**。雖然目前有許多技術可以克服，譬如，使用**『通道技術』（Tunneling Technique）**來保護資料在 IP 網路上不被偷竊，但這類技術大多必須利用上一層通訊協（TCP）定來完成，而且並非每一部路由器都提供到上一層的服務（一般路由器只提供到 IP 層服務）。IPv6 提供認證的功能，未經認證通過的連線， IPv6 路由器將不給予轉送或拆封，以保證資料的隱密性，也不需要透過上層協定的處理。

(C) IPv6 提供流量標籤（Flow Label）

一般 IPv4 網路上，封包經過中途路由器，路由器只負責將封包轉送到適當的路徑上，並未做任何的紀錄。在 IPv6 網路上，每一個封包提供一個流量標籤，同一筆資料串列給予相同的標籤號碼，因此可以做流量控制及統計。

(D) IPv6 提供傳輸流量等級（Traffic Class）的分類

雖然在 IPv4 封包上有提供 ToS（Type of Service），但有關 QoS（Quality of Service）的服務都由上層通訊軟體所提供，ToS 幾乎沒有發揮功能。IPv6 提供傳輸流量等級的分類，再配合流量標籤使用，就可以依照每一個封包的服務性質，給於路徑選擇的優先次序及適合路徑，便可達到 QoS 的需求。

(E) IPv6 減少封包分段（Fragmentation）的機率

在 IPv4 的協定裡，每個封包大小並未嚴格限制，當封包經由不同網路存取層（Network Access Layer, NAL）時，會將封包分割成不同區段，並給予相同的分段號碼（Fragment Number），好讓對方接收到後，依照相同的分段號碼，再組合回原來封包（向上/向下多功能）。但在一序列的封包可

能經由不同的 NAL 傳輸，其中有任一封包沒有到達目的，則將使整串列資料失效，導致必須全部重傳。因此，在 IPv6 協定裡只允許傳送中的起始和終點可以做分割和組合，減少中途路由器的分割動作，不但可以減少中途路由器的負擔，對傳輸效率也較高。

(F) IPv6 簡化封包標頭

不像 IPv4 的標頭裡存放太多欄位，IPv6 捨棄不必要的欄位，作較有效率的處理。

5-6-1 IPv6 封包格式

圖 5-47 為 IPv6 的封包格式，各欄位功能如下：

- **版本（Version）：**表示本封包的 IP 版本，如 IPv6 的值為 6。
- **交通流量等級(Traffic Class):** 標示該封包的流量等級，等級愈高者優先轉送。可區分為 16 個等級，0 ~ 7 可以提供回應壅塞（如 TCP），8 ~ 15 則不提供回應壅塞（如壅塞時便拋棄該封包）。等級 0 ~ 7 所提供之服務如下：

  0：非結構化交通流量（Uncharacterized traffic）

  1：填滿型交通流量（Filler traffic）

  2：未處理資料傳送（Unattended data transfer）

  3：保留（Reserved）

  4：經處理區塊傳送（Attended bulk transfer）

  5：保留（Reserved）

  6：交談式交通流量（Interactive traffic）

  7：網路間控制交通流量（Internet control traffic）

- **流量標籤（Flow Label）：**同一筆資料給予相同的標籤，可作為流量控制。
- **承載長度（Payload Length）：**本封包所承載資料的長度。表示承載其他通訊軟體（TCP 或 UDP 等封包）的長度。

- **下一標頭（Next Header）：**此欄位類似 IPv4 的 Protocol 欄位，表示本封包所承載的標頭型態，如 TCP 或 UDP 等等。
- **跳躍次數（Hop Limit）：**類似 IPv4 上的 TTL 欄位，每經過一個路由器其值就被減一，如路由器發現該值為一時，便將封包丟棄。在 IPv6 的封包標頭上沒有 Checksum 檢查，因此，路由器計算跳躍值後不用再重新計算 Checksum。
- **來源位址（Source Address）：**以 128 個位元（16 個位元組）表示。
- **目的位址（Destination Address）：**以 128 個位元（16 個位元組）表示。

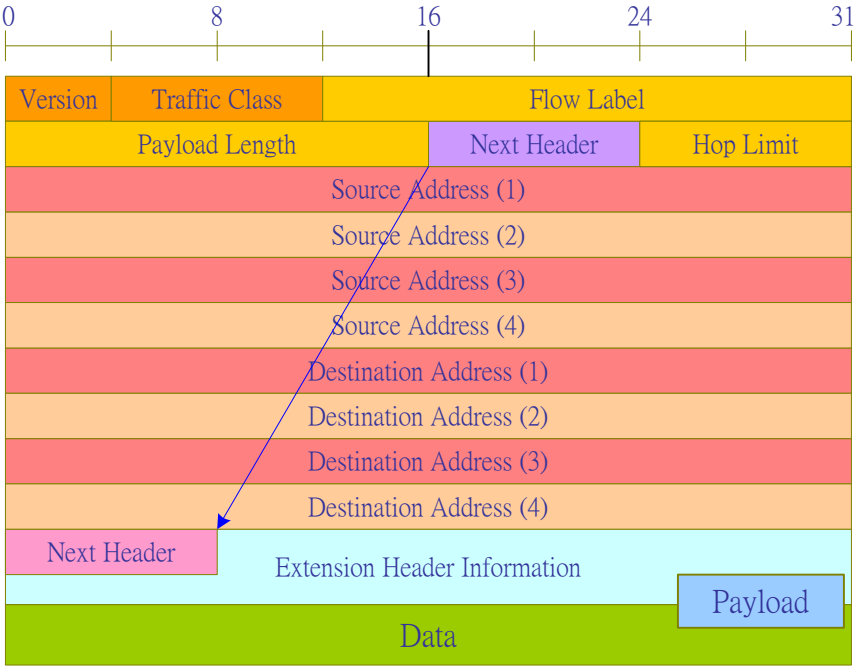

圖 5-47 IPv6 封包格式

5-6-2 IPv6 位址格式

IPv6 用 128 位元（16 個位元組）來表示位址，比 IPv4 增加許多，對於它的表示方法也較複雜。IPv4 只將位址區分為兩個部份：網路號碼和主機號碼，但 IPv6 就有更多型態的區分。首先我們來看如何表示這 128 位元，將其分成 8 組位置，每組 16 個位元又區分為 4 個字元，每個字元（4 個位元）以 16 進位法表示。如下面格式：

X:X:X:X:X:X:X:X

例如：A2E5:56EF:906B:4590:12EC:D532:7812:0001

在 IPv6 位址表示中常有許多組為 0 的情形，為了簡化組內數值都為 0 時，以「：：」來表示，例如：

0:0:0:0:0:0:0:1→ ::1 為 loop back 位址

0:0:0:0:0:0:0:0 → :: 為未指定位址

緊接著，我們來看 IPv6 的位址型態。在 IPv4 協定之下，將主機位址都設定為 1 時，表示針對這個網路號碼之下的所有主機廣播。但廣播封包容易造成風暴，嚴重影響網路效能。IPv6 不再使用這任意廣播的方式，而是使用『**所有節點**』的多點廣播位址。多重節點廣播並非所有節點都會處理廣播訊息，而是有關聯的節點才會處理，如此就可以減低所有節點對廣播訊息的處理。IPv6 將位址型態區分為下列三種：

- 單一廣播位址（Unicast Address）
- 任一廣播位址（Anycast Address）
- 多點廣播位址（Multicast Address）

以下分別介紹各種型態，對於 IPv6 前置位址以『prefix/prefix-length』表示之，例如： 2000::/3 表示該 IPv6 位址的前置碼長度為 3，第一組字元為（0010），其餘皆為 0。

(A) 單一廣播位址（Unicast Address）

單一廣播型態如同 IPv4 一樣，對每一個主機而言，都有一個獨一無二的位址，但 IPv6 採用 64 位元的介面位址，取代 IPv4 的主機位址。IPv6 期望如同 48 位元之 Ethernet 位址一樣可嵌入硬體介面上。單一廣播可區分為四種位址型態：

(a) 可集合式整體位址（Aggregatable Global Address）

(b) 地區本地位址（Site-Local Address）

(c) 鏈路本地位址（Link-Local Address）

(d) IPv4 相容（IPv4-Compatible）之位址格式

以下分別介紹各位址格式之特性：

(1) 可集合式整體位址（Aggregatable Global Address）

一般主機或網路設備都可被指定一個或一個以上以上的整體位址（Global Address），以便與其他電腦設備區分。它的位址格式如圖 5-48 所示，我們將 128 位址區分為三個部份，第一部份（Provider）表示網路連接的供應者，指一般的 ISP 位址；第二部分為區域，讓使用者自行區分網路，也就是一般所言的次網路；第三部份為主機，計有 64 位元，也希望類似 Ethernet 位址使用方法。各欄位功能如下：

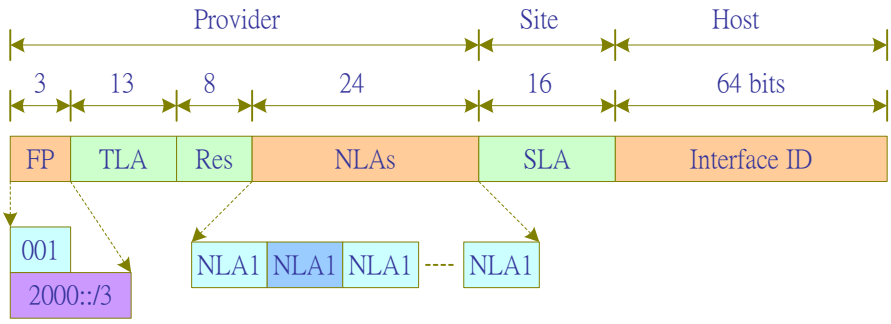

圖 5-48 可集合式整體位址格式

- **固定前置碼（Fixed Prefix, FP）：**以 001 表示 IPv6 的可集合式整體位址格式，對整個位址表示為 2000::/3。
- **最上層集合碼（Top-Level Aggregator, TLA）：**讓網路提供者做最上層的區分，類似電話號碼的國家碼區分法。
- **保留（Reserve, Res）：**保留未使用。
- **下一層集合碼（Next-Level Aggregator, NLA）：**第二層次的集合碼，類似電話號碼的區域碼。它可以再區分若干個區域碼組合而成。
- **地區層次集合碼（Site-Level Aggregator, SLA）：**讓使用者環境區分集合碼，如 IPv4 的次網路（Subnet）位址碼。
- **介面識別碼（Interface ID）：**64 位元的主機位址。

(2) 地區本地位址（Site-Local Address）

在某個區域裡自行構成網路就可以使用 IPv6 的地區本地定址格式，對於該地區內所用的位址格式不同於整體位址（Global Address）（類似 IPX 定址），IP 路由器不會將該類封包轉送出去，只在地區內廣播。至於需要轉送到地區外，則必須使用整體位址，它的位址格式如圖 5-49 所示。

其固定前置碼為 FEC0::/10，並允許地區內再劃分多個小地區，因此有 Subnet-ID 欄位來表示次地區的編碼。Interface-ID 也是 64 位元組來表示（但 IPv6 對 IPX 及 NASP 有另外定義位址）。

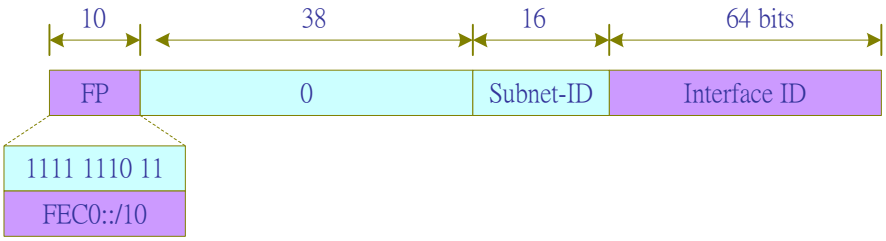

圖 5-49 地區本地位址格式

(3) 鏈路本地位址（Link-Local Address）

如同地區本地位址的特性一樣，但鏈路本地位址是指在每一鏈路上而非地區，因此無法再區分次區域。鏈路本地位址可被自動規劃於任何網路介面上，並用來發現鄰近端點的協定或無狀態轉換的自動規劃使用。任何端點在本地網路上可用該位址互相通訊。路由器也不會將該類封包轉送出去，它的封包格式如圖 5-50 所示，其前置碼為 FE80::/10。

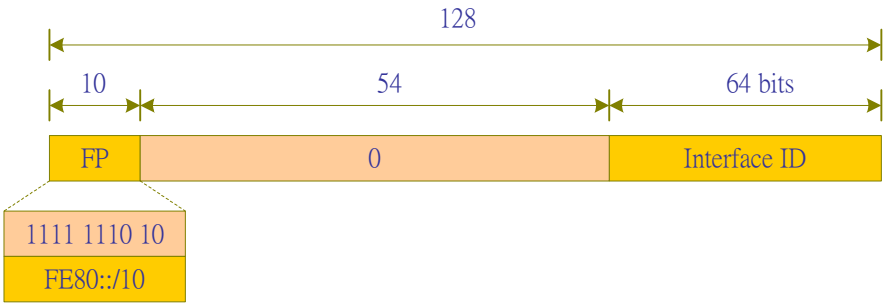

圖 5-50 鏈路本地位址格式

(4) IPv4 相容（IPv4-Compatible）之位址格式

依目前 IPv4 的網路要轉換到 IPv6，可能需要一段長的時間，不僅封包格式需要修改，通訊協定也必須更新。因此，首先讓兩種協定能夠相容，再漸進式慢慢更新。IPv4 有 32 位元位置格式，IPv6 有 128 位元，其中介面位址為 64 位元，我們可將 IPv4 的 32 位元置入 64 位元的介面位址，所剩之位元皆放 0，如圖 5-51 所示。

圖 5-51 IPv4 相容之 IPv6 格式

(5) 任一廣播位址（Anycast Address）

在任一廣播位址型態下，一個位址可屬於多個介面共同使用，這些介面也允許在不同的端點上，當廣播任一廣播位址時，同一位址的介面都可收到。一般最常用是在路由器之間的廣播，不屬於該位址的網路設備就不用去拆裝封包，因此較 IPv4 的廣播方式更節省頻寬。圖 5-52 為次網路之路由器（Subnet router）的任一廣播位址格式，它是由可伸縮的前置碼（Prefix）欄位來代表次網路的位址。

圖 5-52 次網路路由器任一廣播位址格式

(6) 多點廣播位址（Multicast Address）

多點廣播可以指定較多的位址，可以針對若干個端點廣播，不像任一廣播只針對一個端點廣播。如採用多點廣播位址型式，IPv6 可向一組介面位址廣播，而每一個介面位址可以在不同的端點上。位址格式如圖 5-53 所示，它的前置碼為 FF00::/8；Lifetime 欄位表示該包產生的時機，其中 "0" 表示該位址為永久式（Permanent Unicast Address）；"1" 表示該位址為暫時性的（Temporary Unicast Address），使用後便不存在。另 Scope 欄位表示位址型態："1" 表示端點位址（Node Address）；"2" 表示鏈路（Link）位址；"5" 表示地區（Site）；"8" 表示組織（Organization）；"E" 表示整體（Global）位址。但在 IPv6 規格裡有一些特殊的多點廣播位址如下：

- FF02:0:0:0:0:0:0:1 為所有端點的多點廣播群組（All-node multicast group）。
- FF02:0:0:0:0:1:FF00:0000/104 為徵求端點的多點廣播群組（Solicited-node multicast group）。
- FF02:0:0:0:0:0:0:2 為所有路由器多點廣播群組（All-node multicast address），所有路由器都必須加入。

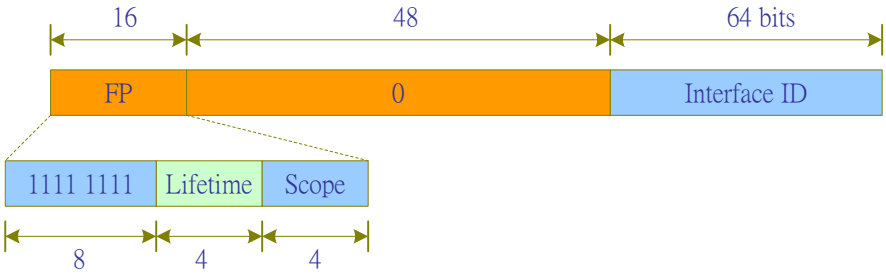

圖 5-53 多點廣播位址格式

5-6-3 ICMPv6 協定

如同 IPv4 上的 ICMP 一樣，ICMPv6 （Internet Control Message Protocol version 6）是用來測試或回報 IPv6 網路訊息，也是包裝在 IPv6 封包內（Next header = 58）。例如，某一路由器由於某種原因無法處理 IPv6 封包時，便發送該原因類別的 ICMPv6 給原始發送封包者，再由它做適當的處理。而其封包型態也如同 IPv4 包含終點無法到達、封包太大、時間超過、參數問題、回聲要求、以及回聲回應等。其封包格式也如同 ICMPv4（如圖 5-28）。

習題

1. 請列出一般網際層有哪些較重要的通訊協定？並敘述其功能。
2. 請敘述 Ethernet 訊框對網際層封包的組裝格式。
3. 請利用網路監視器擷取 ARP 的封包格式，並說明各欄位的組裝。
4. 請利用網路監視器擷取 802.2/802.3 的封包格式，並說明各欄位的組裝。
5. 請敘述 IP 通訊協定的特性。
6. 何謂 TCP/IP 網路？為何這兩個通訊協定要結合在一起？
7. 何謂**『IP 分級』（IP Classic）**？並請敘述每一分級（Class）的網路範圍。
8. 何謂**『網路遮罩』（Network Mask）**？
9. 如果有一網路位址為 138.45.0.0/16，請設計增加 16 個次網路位址，並請列出次網路遮罩及各網路範圍之主機的 IP 位址。
10. 某一 Class B 等級的網路，其網路遮罩為 255.255.240.0，請問該網路有多少個子網路？並且每一子網路最多可分配多少部主機位址。
11. 當某一路由器收到封包時，請敘述其路徑選擇機制的運作程序。
12. 請說明主機（或路由器）的路由表至少應具有哪些訊息？
13. 何謂 IP 廣播和多點傳送？請分別說明其特性。
14. 何謂 IP 分段？每一分段在 Ethernet 網路上最高可以承載多少位元組？
15. 請敘述 IP 分段後可能會發生哪些問題？
16. 何謂 MTU（Maximum Transission Unit）？網路之間如何來協調最高 MTU？
17. 何謂 IP 來源路徑選擇？應如何實現它？
18. 如下圖（圖 5-54）之 IP 網路架構，請規劃出各主機電腦的 IP 位址，並設計出路由器（主機 A）的路由表。

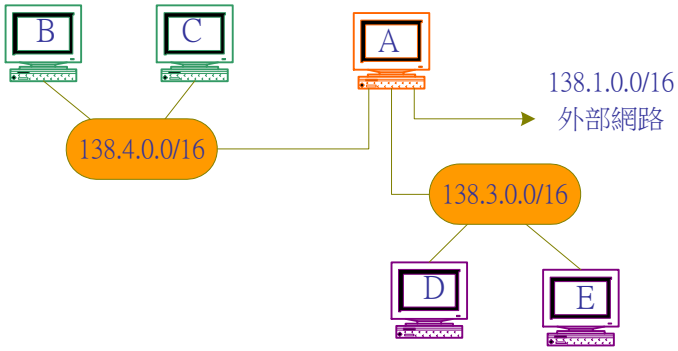

圖 5-54 IP 網路範例

19. 同上題，請利用一部 Linux 作業系統之電腦，安裝成主機 A 之路由器，並測試其路徑選擇之結果（請參考第九章說明）。

20. 同上題，請利用 Windows Server。

21. 何謂 ARP 通訊協定？並說明其運作程序。

22. 何謂『代理 ARP』（Proxy ARP）？請說明其功能。

23. 請擷取 RARP 封包，並說明各欄位功能。

24. 何謂 RARP 通訊協定？並說明其運作程序。

25. 何謂 ICMP 通訊協定？並說明其運作程序。

26. 請執行 "ping" 命令，並擷取 ICMP Echo Request 之封包，並說明各欄位功能。

27. 請執行 "tracert www.nsysu.edu.tw"命令（Win 2k 主機），並擷取 ICMP Echo Request、ICMP Echo Replay 與 ICMP Time Exceede 封包，並解釋各欄位的功能。

28. 請說明下列 IP 網路之命令的功能，並列出在電腦上執行的結果（Linux 主機）。

(1) ping

(2) ifconfig

(3) netstat

(4) route

(5) arp

(6) traceroute

(7) finger

29. 何謂 IGMP 通訊協定？其運作程序為何？

30. 請敘述 Ethernet 層如何達成 IP 多點傳送的功能？

31. 請說明存活時間（TTL）欄位的數值對於 IP 多點傳送有何關聯？

32. 請敘述 IPv6 比 IPv4 增加了哪些功能。

33. 請說明 IPv6 的定址模式。

34. IPv6 位址模式如何相容於 IPv4 的定址方式？