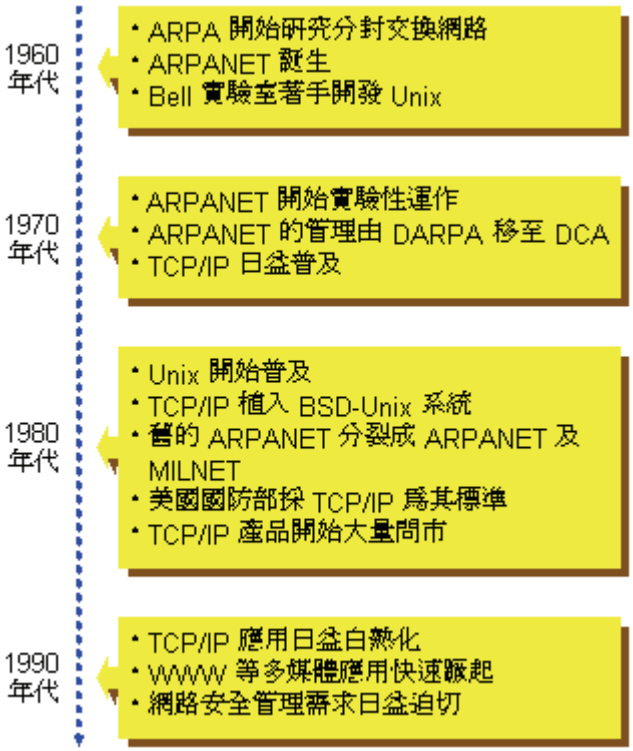
TCP/IP 課程講義

13.0 、 Internet 史觀

官商之間的關係一直是敏感的，政府單位為避免瓜田李下之嫌，通常不會一直採購同一家公司的產品，這對一些文具、食品可能沒什麼大礙，但對於電腦之類的設備可能就會有嚴重的影響，最明顯的即是購自不同廠商的電腦或通訊設備之間彼此互不相容的問題。

圖 13.0 、 Internet 大事紀要



1960 年代，美國國防部 (DOD) 及各處的軍事基地已充斥著各種廠牌的電腦及通訊設備，為令資料能在不同規格的機器間互通，在 1968 年，美國國防部接受若干公司與大學的建議，成立尖端研究企劃署 (ARPA)，初期，APRA 嘗試架設一個實驗性的封包交換網路－ ARPANET，並以它連結一些受政府補助的研究單位，最初，網路的傳輸率僅 56K BPS，僅連接四個節點，所使用的通訊程式名為網路控制程式 (NCP)，當時正處於冷戰時期，ARPANET 主要在研發一不受戰事 (包括核戰) 破壞、堅固、可信賴、與廠商無關、且獨立於電話線路之外的分散式全國性數位通訊網路，未料此網路漸行漸遠，許多當代的通訊技術的基礎即是在那段時期建立的。

1969 年，ARPANET 開始運作，雖然它是個實驗性網路、但卻很成功，不同地點的人員可用它互傳電子信件與檔案，也可透過它使用其它地點的電腦，它的成功吸引許多機構先後連上該網路，用它進行日常的資料交換，最初網路上僅 4 個節點，十年之後，與之連接的主機已超過 100 部。

1970 年初期，ARPA 進一步研究將傳輸介質延伸至移動式無線電、甚至衛星連線等技術，由於 ARPANET 最初的通訊協定彈性不大，造成網路的擴充困難，因此，在 1970 年中期，研究人員開始研發一種新的通訊技術，他們試

圖透過新規格將不同的通訊設備併入同一個網路內，該規格即是後來的 TCP/IP 原型。

1975 年，ARPANET 由實驗性網路改制為操作性網路，整個網路則轉由國防部通訊署 (DCA) 管理，而 ARPA 則改名為 DARPA，1979 年，DARPA 成立一個正式的委員會 ICCB，即目前 IAB 的前身，ICCB 主要在協調並指導新網際協定的開發，那時，TCP/IP 已有基本的雛型，而原本的 ARPANET 亦逐漸將舊有的 NCP 改換為 TCP/IP。

1980 年，TCP/IP 正式問世，DARPA 為推廣 TCP/IP，便以極低的價格提供各界試用，那時，許多大學亦存在不同廠牌電腦設備間難以互通的問題，且也找不到合適的通訊協定，TCP/IP 的出現適時提供他們最佳解答，同時期也正是 BSD Unix 的流行期，為將 TCP/IP 推廣至 Unix 系統，DARPA 提供資金贊助 BBN 公司將 TCP/IP 植入 Unix 環境，並和柏克萊 (Berkeley) 大學合作將之整合於 BSD Unix 環境內，TCP/IP 就在此時與 Unix 相結合。

隨著 TCP/IP 普及於各大學，ARPANET 的規模亦快速擴張，但由於 ARPANET 最初主要的用途是在軍事上，為顧及國防安全，在 1983 年，原先的 ARPANET 分裂成兩個網路，其一是 MILNET，此為 DDN 的非機密部份，僅供美國國防部使用，另一個是新的、且較小的 ARPANET，僅供有和政府簽約合作的研究單位使用，就在同一時期，Internet 這名詞也開始被廣泛引用，當時它代表由 MILNET 與 ARPANET 所構成的整個網路。

由於 ARPANET 隸屬國防部，所以，未和政府簽約的機構即無法使用它，為此，美國國家科學基金會 (NSF) 即輔助電腦技術、工程界的教學研究機構建立採 TCP/IP 規格的網路－ CSNET。

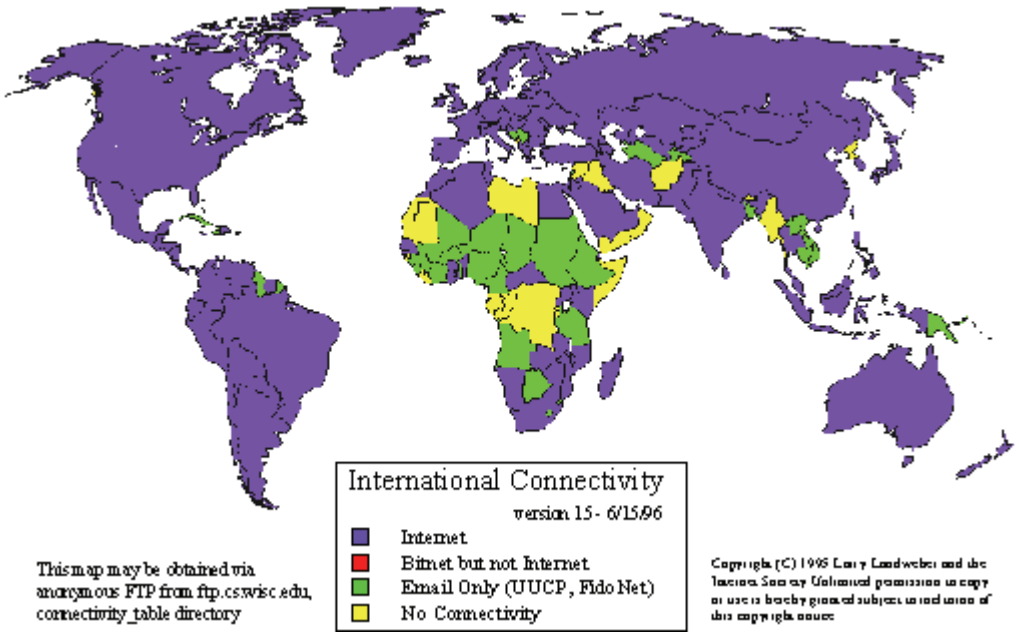
1984 年，NSF 開始規劃超級電腦中心與高速網路，在 1987 年獲得聯邦政府撥款補助，並在全美各地架設七個超級電腦中心，1980 年代末期，美國國防部開始刪減 ARPANET 的預算，當時，NSF 已開始建立一採用 TCP/IP 協定的網路 — NSFNET，並使用較 ARPANET 快三倍的 T1 (1.544M BPS) 傳輸線路，NSFNET 屬於一般性的研究網路，該網除了提供學術界免費的服務外，也服務業界，並酌收些許使用費，由於 NSFNET 的廣泛使用，它逐漸成為許多網路的主幹 (backbone)，之後又由美國國家研究教育網路 (NREN) 取以代之成為骨幹。

1980 年代末期，隨著 Internet 的普及，網路商業化應用一時如雨後春筍般冒出，TCP/IP 的觸角也伸入了辦公室，Internet 不再是學術界教授或學生的特權，所以，網路由軍方發展起，並歷經了電腦工程界、教育界，乃至商業界。

不同性質的網路都得採 TCP/IP 協定與 Internet 連接，這使得各種機構開始熟悉 TCP/IP，並瞭解它的威力，TCP/IP 除了用在 Internet，也適用於其它類型的網路，例如，在 Unix 環境下，僅管實際上不須與 Internet 連接，但 TCP/IP 也常用在區域網路，並且，在區域網路內採 Ethernet 搭配 TCP/IP、而和遠端電腦則利用 UUCP 連線的情形也很常見。

當初，internet 特指採用 TCP/IP 相連結的網路群，現在，小寫 "internet" 泛指採用某種共同協定相結合的一群網路，例如 IPX 網際網路，大寫 "Internet" 特指由原先 ARPANET 發展起、並於全球各地透過 TCP/IP 相連結的整個實體網路。

圖 13.1、1996 年 6 月全球 Internet 連線概況 (引自 ISOC，copyright 1995 A.M.Rutkowski and Internet Society)



當代正處於一技術高速發展、資訊爆炸的反曲點，網路的發展亦同，根據臺灣網路資訊中心 (TWNIC) 公佈的流量統計數據顯示，1996 年間 TANET 與國際 Internet 間平均每月資料流量已為 1995 年的 1.5 倍強，此不包括 SEEDNET 及 HINET 的流量。

根據 ISOC 於 1996 年 6 月 15 日公佈的資料顯示，目前全球幾乎僅剩非洲及東南亞少部份國家未連上 Internet (圖 13.1)。

根據 1995 年 10 月 O'Reilly & Associates 的統計顯示，美國當地約有 580 萬名成人直接存取 Internet，此數字不包含其它商業網路，如 America Online 或 CompuServe 的用戶。

根據 1995 年 MIDS 的統計，Internet 用戶約有 2 至 3 千萬左右。

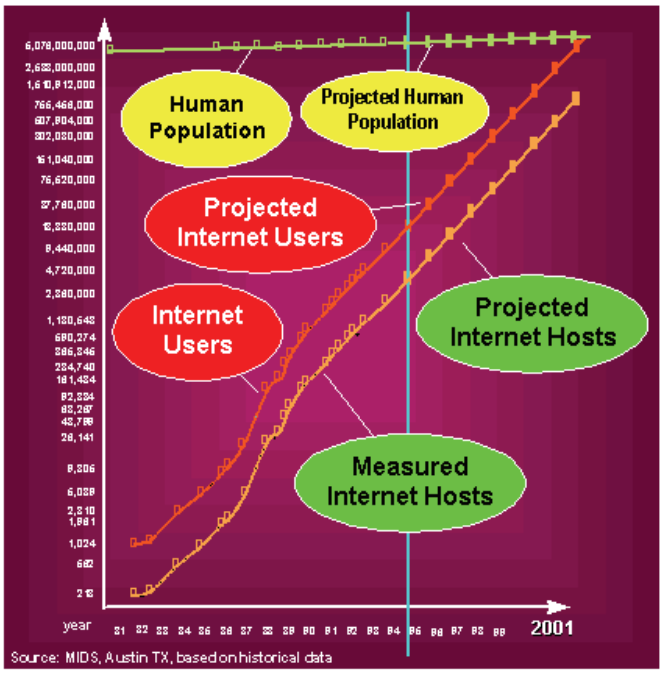
根據一資訊技術調查公司 Zona Research 於 1996 年 8 月的預估，企業 Intranet 將有 280 億美元的市場。

根據專職國際間網域名稱註冊的 Network Solutions 公司的統計，自 1995 年間短短 16 個月內，已註冊的網域由 63,000 增加至 500,000，其中以 .com 網域的為大宗。

專家預估，在公元 2000 年，Internet 的用戶將達 10 億人次，約地球人口總數的 1/15 云云。

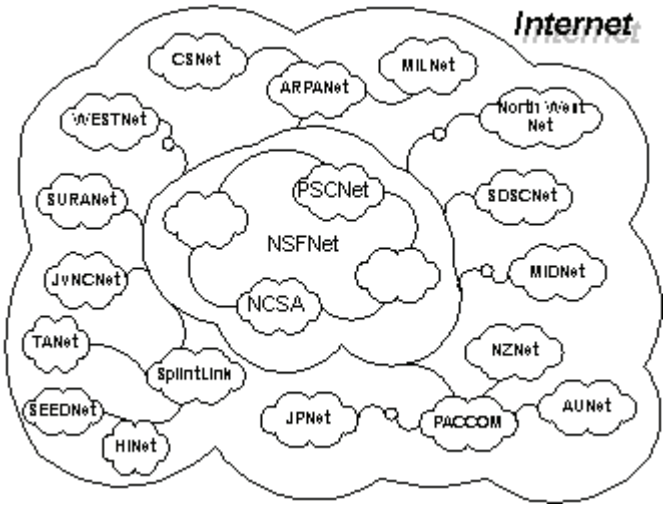
這類數據還有很多，在目前詭譎多變Internet，其各項數據並不容易推算精確，許多研究報告皆顯示，Internet目前的成長率正以每年2倍呈指數方式遞增，可以肯定的是，Internet的用戶及各式應用在短期內必將有增無減。

圖 13.2、2001 年 Internet 主機、用戶成長預估(引自 ISOC，original source MIDS，Austin TX，based on historical data)



Internet的蓬勃發展在積極方面的確引來許多商機、大幅提昇生產力、更便利的通訊服務、多國文化會萃，然而，在消極方面，它也引來許多新問題、新風險，如非法入侵、色情泛濫、網路律法不夠完備、用戶電子信件遭受網路廣告疲勞轟炸等等，Internet的網路頻寬仍未普及到足以應付資料量大的服務(如VOD)，其介面尚未簡化到人人皆可上手，目前的網路安全性亦未成熟到能令用戶安心地透過它進行電子交易，而網路駭客仍有恃無恐，凡此種種皆顯示這國際共同參與的網路遊戲才剛開始。

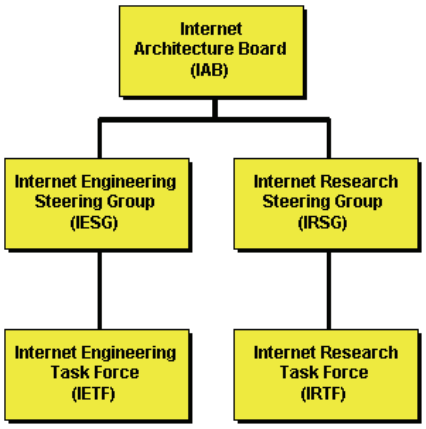
圖 13.3、Internet 網路成員一覽



13.1 、 Internet 組織與標準

Internet 是個開放性網路，和私人企業相較下，其組織架構並不算嚴密，早期負責監督整個 Internet 網路技術之發展的組織是 1983 年成立的 Internet 活動部會 (IAB)，當時 Internet 仍屬研發性網路，由 IAB 負責監督整個 Internet 網路技術的發展與規劃，其下設有兩個指導團體－ IESG 與 IRSG，這兩個指導團體各領導自己附屬的特別小組，由 IETF 負責短期工程，IRTF 負責長期研發。

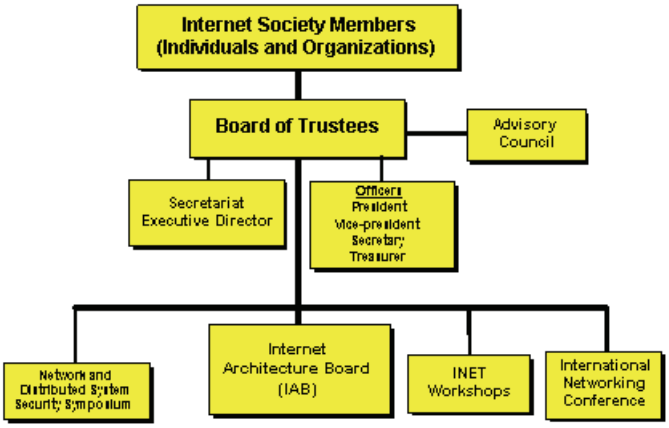
圖 13.4 、 IAB 組織架構



而隨著 Internet 的國際化與網路社會的日益複雜，人們逐漸意識到須要一更具代表性、非官方的組織以整合當時現有的網路資源，包括相關機構及各式標準，以利 Internet 加速進化，並推廣其應用層面，為此，在 1991 年 6 月於哥本哈根 (Copenhagen) 舉行的一國際性網路會議中即公佈籌組一 Internet 的國際性組織－網際網路協會 (ISOC)，並於 1992 年 1 月正式成立，其成員包括來自全球各地與 Internet 有關的企業、非營利機構、官方部門、乃至個人。

ISOC 的功能是多方面的，包括發展並推廣 Internet 的技術及應用，搜集並傳播與 Internet 相關資訊，強化整個 Internet 的結構，推廣和 Internet 有關的教育及研發，扮演整個 Internet 社會的各式活動、集會的仲裁、協調者，輔助開發中國家、地域發展 Internet 基礎建設，最後，ISOC 也與其它相關國際機構、官方部門接觸，以利上述目的的推展。

圖 13.5 、 ISOC 組織架構 (1995 年 9 月)



Internet 大部份的文件、程式、或測試都是由感興趣的個人或團體完成的，IAB 只負責監督、並要求這些文件依照 Internet 的建議文件 RFC 系列編號，RFC 自 1969 年 ARPANET 成立已來即不斷增加，每份 RFC 皆有個唯一的號碼，截至 1996 年底已有超過 2060 份 RFC。

IAB 的 RFC Editor 負責每份 RFC 的最後編校及發行，若某份 RFC 的內容被校定過，則為了避免混淆，須將之編為新的號碼，若有新版的 RFC 取代某些舊的 RFC，則在新的 RFC 的封面會標示此事。

RFC 的風格及內容並不若一般的規格文件來的嚴謹，它所涵蓋的資訊範圍極廣，並不僅限制在通訊協定的制定上，不過與 Internet 基礎協定有關的規格則多由 IETF 及 IESG 所定義，所有 RFC 皆可於全球各大網路免費取得，對 RFC 的產生方式感興趣的讀者可參考 RFC-1310，其分類方式可參考 RFC-1000。

Internet 的通訊協定存在各種號碼，如 IP 位址、協定號碼、網域名稱等，為統一管理，Internet 亦設有一專責的號碼配置機構－IANA，IANA 由 ISOC 及美國聯邦網路議會 (FNC) 授權，專職負責 Internet 號碼的配置、登錄，並定期以 RFC 型式公佈。

從 RFC-1150 開始，又分支出另一名為 FYI (For Your Information) 的文件系列，FYI 實際上是整個 RFC 的一個子集，每份 FYI 即是另一份 RFC，例如 FYI 1 對等於 RFC-1150，FYI 不涉及標準或規格的制定，其內容多是各地各階層的 Internet 參與者的資訊報導，或是一些與 Internet 有關的使用經驗、常問問題 (FAQ) 的解答等文獻，FYI 也有自己的編號，並且自成一系列。

除了 RFC 之外，還有一群由 IETF 所發表的 Internet 草案 (Internet Draft) 系列文件，這些文件僅報告工作進度，並未明確定義任何標準，不過，若某份 Internet 草案的發展在經多次改版已趨於成熟，它即可能被考慮成為標準，但，Internet 草案並不像一般標準文件被永久保存著，它通常僅存在某段短暫的時間，之後經常就被刪除了，因此，廠商或著述者最好不要引用 Internet 草案的資料，而改採 RFC 的。

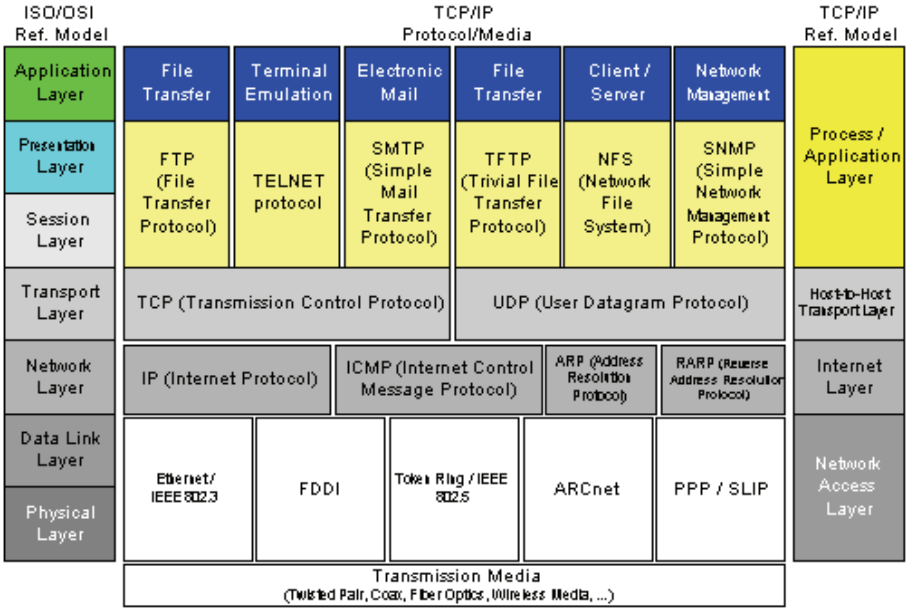
目前已發表的 RFC 數量已超過兩千份，但這些文件皆被 IAB 標示為資訊的 (informational)，意即，它們僅供參考、討論，不是官方的標準，為了避免混淆，RFC Editor 引進另一系列的文件－STD，凡由 IAB 標示為已標準化狀態 (standardization state) 的 RFC 即被歸類為 STD 系列。

13.2 、 TCP/IP 協定堆疊

TCP/IP 是組通訊協定的名稱，它源於該組協定中最重要的兩個協定 TCP 與 IP，通訊協定是套定義完善的溝通規則，不同種類的機器只要遵循相同的協定即可互通，而 TCP/IP 正是 Internet 網路社會的共通語言，主機間須利用 TCP/IP 互通訊息。

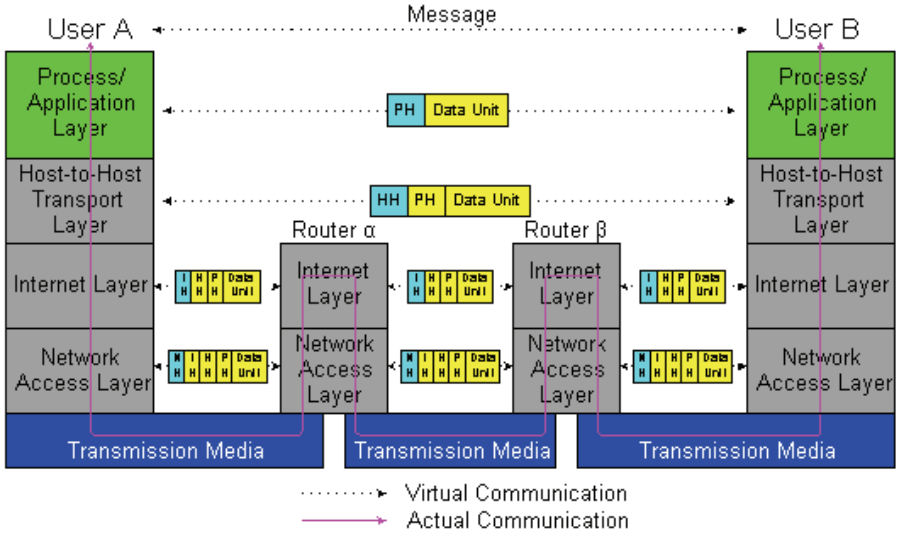
TCP/IP 也是多層的協定堆疊架構，不過經常被引用的 ISO/OSI 參考模型 (2.0 節) 並不很適合直接描述 TCP/IP 協定堆疊，較恰當的模型是如圖 13.6 所示的四層模型，該圖也展示了它與 OSI 模型的關聯性，TCP/IP 未定義網路存取層，換言之，它可架構於多種網路存取介面之上，如 Ethernet、FDDI、Token Ring、或串列線路，只須提供這些介面的驅動器即可。

圖 13.6、TCP/IP 協定堆疊



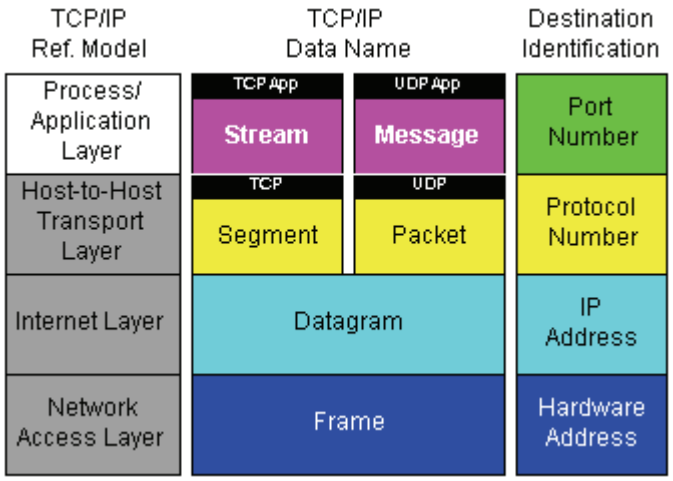
在 TCP/IP 協定堆疊中，協定間的對談只發生在同一層的相同協定之間，這稱作虛擬連線，實際的資料流動則是由發源層依序傳至最底層，之後透過傳輸介質送抵對方的最底層，再依序傳至目標層，每一層將資料傳至下一層之前會先於其資料區塊的前端附加一稱作表頭 (header) 的控制資訊，此表頭記錄了該資料塊相對於該層的特性及資訊，每一層會將上一層傳來的資料連同其表頭一同視為上層的資料，並附加該層的表頭之後再送至下一層，這種資料封裝 (encapsulation) 過程大抵上與 OSI 描述的相同，當資料送抵對方時也會發生解封裝 (decapsulation) 動作，意即，每一層由下一層收到資料之後，會先剝去該層的表頭，之後再將剩餘的部份送至上一層。

圖 13.7、TCP/IP 通訊及資料封裝 (encapsulation)



TCP/IP 協定堆疊的每一層資料皆有固定的結構與名稱，理論上，每一層皆可忽略其它層的資料結構，但實際上，由於考量了傳輸效率等因素，每一層的資料結構皆被設計成相容於該層相鄰的上、下兩層的資料結構，僅管如此，每一層仍保有描述該層資料結構的專門術語，TCP/IP 每一層的資料名稱皆不相同，例如，使用 TCP 的應用程式稱呼它的資料為串流 (stream)，TCP 本身稱它的資料為資料段 (segment)，IP 稱它的資料為資料片 (datagram)，最底層則稱之為框架 (frame)。

圖 13.8、TCP/IP 協定堆疊各層的資料名稱



13.3.0、IP 位址 (IP Address)

在 Internet，任何使用 TCP/IP 協定提供或接受服務的電腦或設備都算是主機 (host)，TCP/IP 被設計成適用在不同類型、位於全球各地的傳輸介質及電腦系統，為方便標定每部主機，Internet 定義了一套通用的定址方法，當時理想的定址格式須提供足夠的跨網選徑 (routing) 資訊、且不佔太多儲存空間，TCP/IP 的定址方式即是給定每部主機一組在整個 Internet 為唯一的號碼，稱作 IP 位址 (IP Address)。

IP 位址長 32 位元，為方便表達，人們將此 32 位元數值切成四段，每連續 8 個位元一組，改以如下的四個整數值表達一 IP 位址

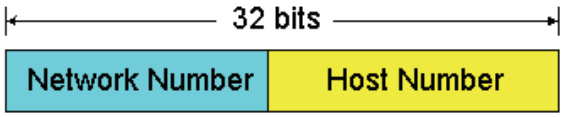
dec3.dec2.dec1.dec0 或 [dec3.dec2.dec1.dec0]

其中，dec0 至 dec3 為 10 進位數值，並取左高右低的階層性，例如 192.168.5.10，此種 IP 位址表示法以句點隔開數字，故又稱作點標記法 (dotted notation)。

雖然 IP 位址長 32 位元，但實際上僅內含兩項資訊，即網路號碼 (network number) 與主機號碼 (host number)，故 IP 位址也可表達成

[network number,host number]

圖 13.9 、 IP 位 址 結 構



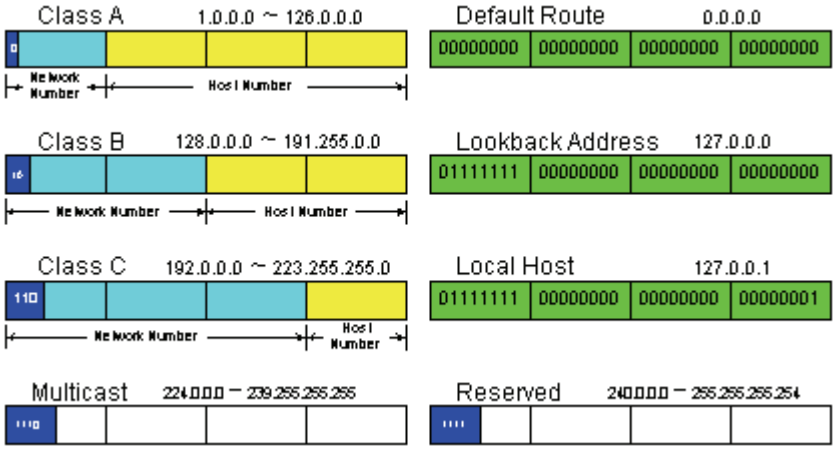
TCP/IP 是 網 路 導 向 的 ， 意 即 ， 屬 於 相 同 網 路 的 主 機 其 IP 位址 的 網 路 號 碼 也 相 同 ， 而 主 機 號 碼 則 是 該 主 機 在 該 網 內 唯 一 的 編 號 ， 同 網 路 的 主 機 間 可 直 接 互 通 ， 不 須 藉 助 第 三 者 ， 不 同 網 路 主 機 則 須 藉 助 選 徑 器 (router) 於 網 路 間 遞 送 封 包 ， 而 選 徑 器 賴 以 判 斷 的 選 徑 資 訊 即 是 IP 位 址 中 的 網 路 號 碼 。

IP 位 址 又 名 主 機 位 址 ， 但 此 名 稱 有 著 誤 導 作 用 ， 它 會 令 初 學 者 誤 以 為 一 部 主 機 僅 有 一 個 IP 位 址 ， 但 實 際 上 ， 主 機 的 每 個 網 路 介 面 皆 可 擁 有 數 個 IP 位 址 ， 實 際 上 ， 主 機 的 每 個 網 路 介 面 皆 可 配 置 數 個 IP 位 址 ， 當 然 ， 若 無 特 殊 需 求 ， 一 般 還 是 一 個 介 面 、 一 個 位 址 。

13.3.0.0 、 IP 位址分級

TCP/IP 網 路 依 其 中 所 能 容 納 的 主 機 數 量 多 寡 分 成 A 、 B 、 C 三 級 ， D 級 目 前 為 實 驗 性 多 點 投 射 (multicast) 位 址 ， E 級 則保 留 作 為 未 來 發 展 之 用 ， 分 級 的 技 巧 是 配 置 不 同 的 位 元 數 目 予 網 路 號 碼 部 份 ， 網 路 號 碼 的 位 元 數 多 ， 該 級 的 網 路 數 目 就 多 ， 但 相 對 的 ， 其 主 機 號 碼 的 位 元 數 就 變 少 ， 而 該 網 能 容 納 的 主 機 數 目 也 少 。

圖 13.10 、 各 級 IP 位 址 結 構



網 路 分 級 的 原 因 是 考 慮 到 不 同 規 模 的 網 路 ， 因 為 Internet 是 網 路 導 向 的 ， 當 申 請 一 個 網 路 時 ， 申 請 者 須 考 慮 其 網 路 內 可 能 有 的 主 機 數 目 ， 並 申 請 適 當 等 級 的 網 路 ， 申 請 超 過 實 際 所 須 的 網 路 將 使 得 大 部 份 IP 位 址 被 閒 置 ， 這 在 目 前 IP 位 址 短 缺 的 情 況 下 是 不 被 允 許 的 。

網 路 等 級 的 區 別 方 式 是 由 其 IP 位 址 的 最 高 特 定 個 位 元 的 值 判 定 ， 圖 13.10 中 ， A 級 網 路 的 最 高 位 元 值 為 0 ， B 級 網 路 的 為 10 ， C 級 網 路 為 110 ， 最 高 三 個 位 元 值 為 111 的 位 址 則 保 留 做 為 其 它 特 殊 用 途 。

另外，網路 0 與 127 已保留於特別用途，網路 0 代表預設選徑 (default route) 位址，供應用程式將之做為預設選徑的位址表示方式，網路 127 代表回繞 (loopback) 位址，此方便主機自己定位自己。

在各等級的 IP 位址中，有兩種特別的位址已保留它用，主機號碼的位元值皆 0 的位址代表該網路，主機號碼的位元值皆 1 的位址為該網路的廣播位址 (broadcast address)，此位址可同時定址當地網路的所有主機，例如 26.0.0.0 代表網路 26，其廣播位址為 26.255.255.255，192.168.5.0 代表網路 192.168.5，其廣播位址為 192.168.5.255。

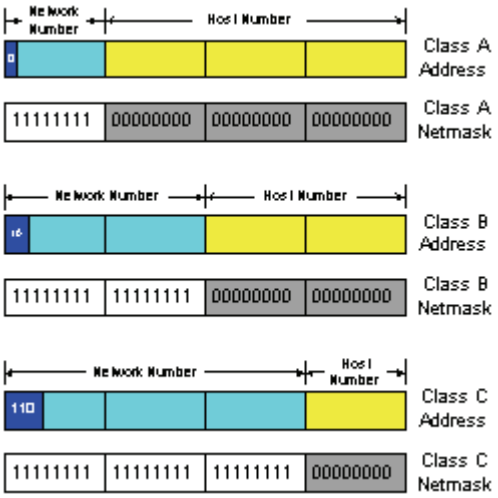
雖然 IP 位址的容量是 32 位元，但 Internet 目前遭遇的瓶頸卻是 IP 位址將在短期內耗盡，解決之道尚在研擬中，造成此問題的主因即是網路分級的設計，它使得實際可用的 IP 位址較理論上的 47 億個還少，以一個 B 級網路為例，此種網路最多可有約 64,000 部主機，但申請此級網路的機構可能沒有那麼多部主機，因此該網中未分配到的 IP 位址就被閒置了，無論如何，此問題並不能完全怪罪當初的設計者無先見之明，在 TCP/IP 誕生的年代，網路作業僅限制在大型機構，當時尚無 Unix 系統，一個 32 位元的定址方式在當時的環境而言實在夠大了，天知道網路會發展成今天的規模。

13.3.0.1 、 網路遮罩 (Netmask)

IP 位址是組由網路號碼及主機號碼構成的 32 位元定址數值，為方便討論，一般也以 [network#, host#] 表示，IP 位址的網路號碼決定了主機所屬網路，因此主機在傳遞封包之前，會先由其中過濾出網路號碼，以決定封包的歸宿。

為由 IP 位址濾出網路號碼，人們引進網路遮罩 (netmask) 概念，該遮罩同樣是 32 位元數值，根據 IP 位址的分級，網路遮罩中與網路號碼對應的位元保留為 1，主機號碼的位元皆為 0，將這樣的遮罩與 IP 位址進行 AND 運算的結果即是網路號碼 (圖 13.11)。

圖 13.11 、 網路遮罩 (Netmask)



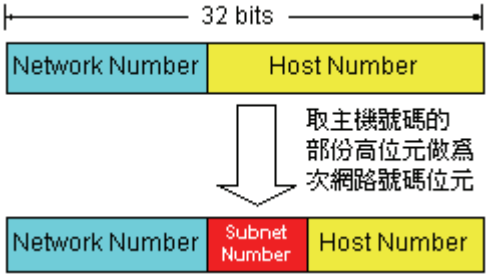
例如，A級網路的網路號碼為其IP位址中的最高8位元，故其網路遮罩為255.0.0.0，依此類推，B級網路的遮罩為255.255.0.0，C級為255.255.255.0。

13.3.0.2 、 次網路遮罩 (Subnet Mask)

將IP位址規劃成網路號碼及主機號碼兩部份已可應付大部份需求，但網路分級所衍生的問題是，有些大型機構、組織、或企業內部實際上是由數個區域網路所組成，在技術上，每個區域網路皆得擁有自己的網路號碼，否則網路之間無法遞送封包，解決方式可以是為每區域申請一組網路號碼，但這通常會造成IP位址的浪費，而實際上也沒有一個區域網路可在不使用選徑器或閘道器裝置的情況下含滿A級或B級網路所能容納的主機數目，所以，網路有必要再劃分成次網路(subnet)。

欲設置次網路，就必須有次網路的編號，換言之，原本IP位址所表達的[network#, host#]資訊即須變換成[network#, subnet#, host#]形式，而為了在原本的網路號碼之下再獲得次網路號碼，就必須犧牲主機號碼的些許位元作為次網路號碼的位元，至於犧牲的位元數量須由次網路數量決定，挪做次網路號碼的位元越多，剩餘主機號碼的位元就越少，則每個次網路所能容納的主機數量即相對減少。

圖 13.12 、 次網路號碼



至於如何由主機號碼位元挪出次網路號碼位元?實際的作法是將網路遮罩的1的範圍由網路號碼位元延伸至主機號碼的高位元，例如，若欲將一個B級網路劃分成8個次網路，則須由主機號碼位元挪出3個位元作為次網路號碼位元，由IP的運算習慣，我們可將預設的B級網路遮罩255.255.0.0延伸成255.255.224.0，以二進位表示即為

11111111.11111111.11100000.00000000

其中，第三個位元組左側的3個位元即對應至次網路號碼位元，此網路遮罩可將任一B級網路位址劃分成八個次網路位址，即

x.y.0.0 x.y.32.0 x.y.64.0 x.y.96.0
x.y.128.0 x.y.160.0 x.y.192.0 x.y.224.0

其中的x.y為原本的網路號碼部份，每個次網路可提供2^13個主機號碼，其中，0已保留，31為該次網路的廣播位址，所以實際上有2^13 - 2個主機號碼可用。

變動網路遮罩對於IP的影響只在於由IP位址計算得的網路號碼，當我們將網路遮罩延伸至主機號碼位元時，對於IP只是網路號碼的位元數目增

加而已，只要同一個網路內的所有主機皆使用相同的網路遮罩，即不會影響IP對於資料片的遞送，換言之，網路遮罩的引入令IP位址的使用變得更有彈性。

次網路的設置尚可解決不同網路類型之並存以及長距離或主機數量過多的問題，例如，透過次網路的設置，不同種類的網路即可擁有自己的次網路位址，再透過IP選徑器即可連接這些次網路，而主機數量過多的區域網路也可使用次網路技巧將之打散成數個小網路，以減少主機間封包碰撞頻率過高、導致網路傳輸率下降的問題。

13.3.1 、 網際協定 (IP)

IP (Internet Protocol) 是網際網路層最主要的協定，它也是整個TCP/IP協定堆疊的靈魂，其它協定都得靠IP傳輸資料，無論資料的最終目的為何，所有流進流出的資料皆會經過IP，其功能包括

- 於網路存取層及端對端傳輸層之間搬移資料
- 進行資料片 (datagram) 的拆解與重組
- 將資料片傳送至目標主機，含位在它網的主機

資料片 (datagram) 是TCP/IP最小的資料傳輸單位。

在連線技術上，IP具有下列特性

• IP是非連線導向 (connectionless) 的，意即，IP是直接將封包傳至目標，並無事前的握手 (handshake) 程序。

• IP無錯誤偵測動作，透過IP傳送資料的協定可視需要自行偵測傳輸時可能發生的錯誤。

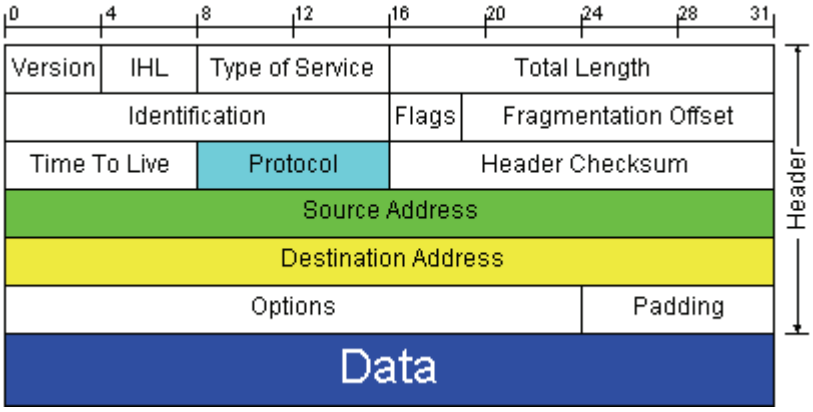
• 資料片彼此間的次序經傳送至對方時有可能和原來的不同，例如若資料片通過動態選徑區段時，不同的資料片可能流經不同的路徑，此時即可能發生先來後到的情形。

• 相同資料片可能被重複發送，而IP不做此類偵測，重複發送的情形可能發生在發送端的上層協定在逾時未收到接收端的認可回應 (中途遺失?)，再次透過IP發送同一個資料片，而實際上，接收端在收到第一個資料片時已回應過 (對方沒收到?)，結果又再次收到同樣的資料片。

• IP不驗證目標主機是否確實收到正確的資料。

以上幾點似乎意味著IP不可被信賴，實際上也並非如此，特別是在目前通訊品質普遍良好的網路環境，並且，若真正需要做到零錯誤的資料傳輸，也可在其上層協定進行，TCP即是個例子。

圖 13.13 、 資 料 片 (datagram) 結 構



在 資 料 片 (圖 13.13) 的 目 標 辨 識 上 ， IP 使 用 IP 位 址 及 協 定 號 碼 ， 資 料 片 是 Internet 最 基 本 的 傳 輸 單 位 ， 它 包 含 一 組 IP 附 加 的 表 頭 及 資 料 本 身 ， 表 頭 中 內 含 該 資 料 片 的 來 源 位 址 、 目 標 位 址 、 及 目 標 協 定 號 碼 ， 若 該 資 料 片 是 往 外 的 ， 則 IP 會 根 據 目 標 主 機 的 IP 位 址 、 配 合 其 選 徑 機 制 (13.3.1.0 節) 將 該 資 料 片 送 至 選 徑 器 (router) 或 目 標 主 機 ， 若 該 資 料 片 由 外 往 內 的 ， IP 會 根 據 其 表 頭 登 記 的 協 定 號 碼 將 已 剝 去 IP 表 頭 的 資 料 區 塊 轉 交 指 定 協 定 (如 TCP 或 UDP) 處 理 。

端 對 端 傳 輸 層 的 每 個 協 定 皆 有 個 協 定 號 碼 (protocol number) ， 方 便 IP 傳 遞 資 料 時 辨 識 ， 有 許 多 號 碼 已 配 置 予 著 名 (well-known) 協 定 ， 也 就 是 被 廣 泛 使 用 、 已 成 為 標 準 的 協 定 ， 例 如 TCP 的 協 定 號 碼 是 6 、 UDP 的 是 17 ， 其 它 著 名 協 定 的號 碼 可 參 考 表 13.14 。

表 13.14 、 著 名 協 定 號 碼 (節 錄)

協 定 名 稱	協 定 號 碼	協 定 全 名
IP	0	Internet Protocol
ICMP	1	Internet Control Message Protocol
IGMP	2	Internet Group Multicast Protocol
GGP	3	Gateway-Gateway Protocol
TCP	6	Transmission Control Protocol
PUP	12	PARC Universal Packet Protocol
UDP	17	User Datagram Protocol

由 於 IP 底 下 的 網 路 存 取 層 協 定 可 批 次 承 載 的 資 料 塊 長 度 視 下 層 協 定 而 定 (如 Ethernet 、 SLIP 、 PPP) ， IP 須 將 上 層 協 定 遞 交 給 它 的 過 大 的 資 料 塊 切 割 成 其 下 層 協 定 所 能 接 受 的 尺 吋 ， 之 後 再 一 塊 塊 傳 送 出 ， 而 在 接 收 端 的 IP 則 負 責 將 被 切 割 的 資 料 片 收 集 成 原 先 的 完 整 一 塊 之 後 再 遞 交 其 上 層 協 定 處 理 。

13.3.1.0 、 IP 選徑機制 (Routing)

在TCP/IP協定堆疊模型中，由最上層應用程式發出的資料經過其間許多協定的處理，最後一定會由IP經手、傳送至目標主機，當資料抵達IP層時已成為資料片(datagram)形式，其表頭含有來源及目標的IP位址，目標主機可能就在同一段網路上，也可能遠在天邊，IP如何將資料片送抵目標主機?

通常，區域網路本身是自給自足的，在同一段網路中，主機與主機間無須借助選徑(routing)即可通訊，一般以廣播(broadcast)方式通訊，傳送主機將它的封包丟上網路，同一段網路上的所有主機在收到該封包時會先檢查封包表頭的目標位址是否與自己的位址相符，若是，就收下該封包，否則忽略它。

廣域網路間則無法靠廣播之類的方式傳送封包，若每部主機都廣播，則整個網路將被廣播的封包所佔滿，因此，網路間須設置一至數部對外的閘道器(gateway)或選徑器(router)，這種機器負責區域網路對外的所有通訊，它們卡在網路之間，只讓必要的封包通過它們，遮住其它不須跨網的封包。

對IP而言，一旦目標主機的IP位址確定了，則選徑的工作幾乎是件極容易的事情，由於Internet是網路導向的，因此，當IP發現

・若目標位址的網路號碼與來源的網路號碼相同，此表示通訊在當地網路進行，不須選徑，直接將封包送出，例如採廣播方式。

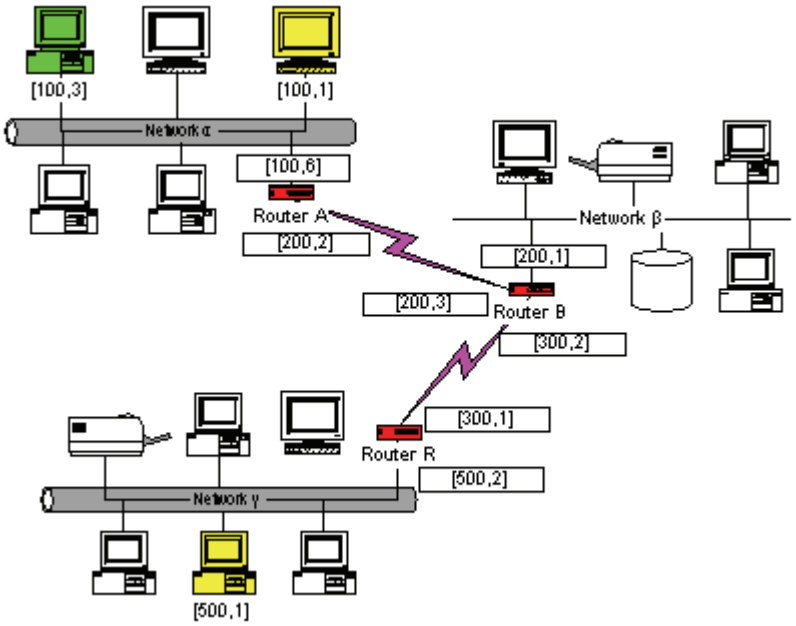
・若目標位址的網路號碼與來源的網路號碼不同，此表示通訊須跨網進行，這時須將封包交由當地選徑器代送。

當來源主機的IP將資料片託予當地選徑器時，該選徑器將根據目標位址的網路號碼選擇與該網連接、且最恰當的另一網的選徑器，將資料片交由該選徑器再代為遞送，如此這般，資料將在由不同網路間的選徑器連成的路線上旅行，直到抵達目標網路時，目標網的選徑器即直接將資料片送至該網的目標主機。

另外，選徑器所連接的各個網路可能採不同的傳輸介質，因此，當資料片由一網遞送至另一網之前，若其長度超過傳輸介質能負荷的最大量時，IP尚得負責將資料片拆解成較小的資料片，至於接收主機的IP則負責將這些小的資料片重組成原來的尺吋。

IP看似無遠弗屆、可即時通達Internet的每個角落，但對於不採TCP/IP的網路(如BITNET、Fidonet)，IP的觸角就到不了了，這時，網際間的資訊即須透過閘道器(gateway)遞送，IP可將資料交給適當的閘道器，再由閘道器將資料轉換為另一網可接受的格式之後、送進目標網路。

圖 13.15 、 IP 選 徑 機 制 展 示



接著我們舉個IP選徑實例，圖 13.15 中有三個網路，分別是 α 、 β 、及 γ ，我們以 [network#, host#] 表達主機的位址，在網路 α ，假設其中的 [100, 3] 送出封包給 [100, 1]，來源與目標的網路號碼都是 100 ，所以 [100, 3] 的 IP 在網路 α 廣播封包，[100, 1] 立即收到該封包，這是不須選徑器的情形。

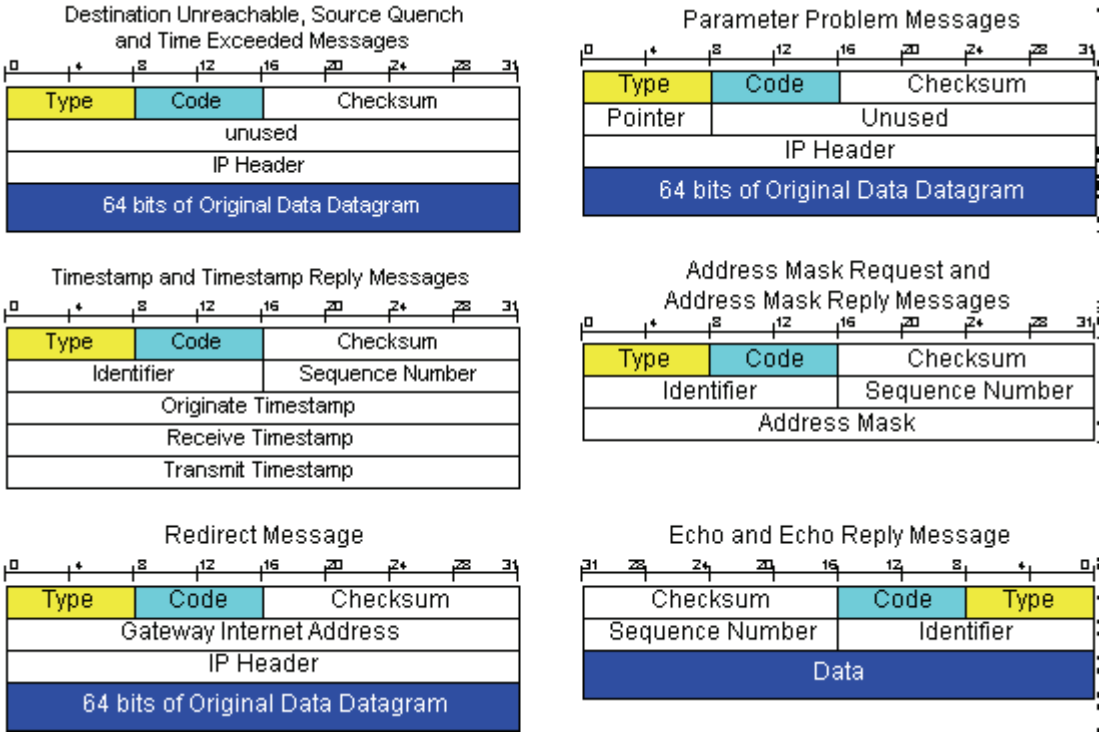
若 [100, 3] 送出封包給 [500, 1]，來源與目標的網路號碼不同，所以 [100, 3] 的 IP 在網路上廣播、將封包交給當地的選徑器 A 代為遞送，A 的唯一出口的 [200, 2] 介面，所以它直接將封包送至 β 網的選徑器 B，B 發現，封包的目標網路號碼非當地 β 網的，它還發現，由 [300, 2] 介面可將封包送至目標網路，所以它由此介面將封包送至 γ 網的選徑器 R，R 發現收到的封包的目標網路是自己的，R 即以廣播方式由其 [500, 2] 介面將封包交給當地的目標主機。

13.3.2 、 網際控制訊息協定 (ICMP)

ICMP 是與 IP 模組整合在一起的控制訊息協定，它透過 IP 收發 ICMP 訊息，ICMP 被用於報告在傳輸資料片 (datagram) 的過程中發生的各種狀況，包括資料片的目標不存在、遞送路徑不正確等訊息，也可透過 ICMP 測試主機之間的連接是否中斷，甚至利用 ICMP 控制特定主機的資料片流出量。

與 IP 的上層協定相似，ICMP 既然透過 IP 收發控制訊息，其訊息在經 IP 傳送前，自然也被裹上一層 IP 表頭，ICMP 不做錯誤偵測，因此它與 IP 同樣不可完全被信賴，ICMP 訊息內含錯誤報告或回應，其訊息種類有許多，每種訊息的結構不盡相同，主要結構有 6 種，共同的部份為其前導的三個欄位－訊息型別 (Type)、訊息代碼 (Code)、及核對合 (checksum)，後續部份則視訊息型別而有不同。

圖 13.16 、 ICMP 主 要 6 種 控 制 訊 息 之 格 式



訊 息 型 別 (Type) 記 錄 了 該 訊 息 的 種 類 (圖 13.16) ， 例 如 目 標 不 可 觸 及 (Type 3) ， 訊 息 代 碼 (Code) 則 記 錄 了 更 進 一 步 的 細 節 ， 以 目 標 不 可 觸 及 為 例 ， 訊 息 代 碼 則 進 一 步 指 出 是 網 路 、 主 機 、 協 定 、 或 是 目 標 埠 不 可 觸 及 ， 有 數 個 訊 息 型 別 是 成 對 出 現 的 ， 一 個 訊 息 由 一 方 主 動 發 出 ， 屬 於 要 求 (request) 訊 息 ， 另 一 個 訊 息 由 被 要 求 端 被 動 回 覆 ， 屬 回 應 (reply) 訊 息 ， Type 8 和 0 、 13 和 14 、 15 和 16 、 17 和 18 即 屬 於 這 樣 的 配 對 ， 底 下 即 針 對 幾 個 主 要 的 訊 息 進 一 步 說 明 :

表 13.17 、 ICMP 訊 息 型 別 (Type) 一 覽

型 別	描 述
0	Echo Reply
3	Destination Unreachable
4	Source Quench
5	Redirect
8	Echo
11	Time Exceeded
12	Parameter Problem
13	Timestamp
14	Timestamp Reply
15	Information Request (已 停 用)
16	Information Reply (已 停 用)
17	Address Mask Request

18	Address Mask Reply

・目標不可觸及訊息 (Type 3 、 Destination Unreachable)

傳送 IP 資料片 (datagram) 的過程中，若任一主機發現無法繼續將資料片遞送至下個目標，該主機會利用 ICMP 向資料片的發源主機發出這樣的訊息，此訊息的型別 (Type) 為 3 ，其訊息代碼 (Code) 欄則內含不可觸及的目標種類 (表 13.18)。

選徑器 (router) 在遞送 IP 資料片時會取資料片的目標位址與它的選徑表 (routing table) 做一查照，若發現目標位址不在它的遞送範圍、或因其它理由無法遞送時，則較可能發出代碼為 0 、 1 、 4 、或 5 的訊息。

若資料片已抵達目標主機的 IP 、甚至更上層協定，但無法送至對應的協定 (經由 Protocol)、或服務 (經由 Port)，則該主機即可能發出代碼為 2 或 3 的訊息。

表 13.18 、 ICMP 目標不可觸及之訊息代碼 (Code)

代碼	描述
0	Net Unreachable
1	Host Unreachable
2	Protocol Unreachable
3	Port Unreachable
4	Fragmentaion Needed and DF Set
5	Source Route Failed

・重導遞送路徑 (Type 5 、 Redirect)

一般用戶對於選徑器的概念及設置可能不若網路管理員來的敏感，若同一段區域網路上存在兩部或更多部選徑器，則可能發生與網外的某部主機連線時，經由選徑器 A 遞送資料片的效率會較由用戶預先指定的選徑器 B 的效率為優，此時，當主機的 IP 將資料片交由選徑器 B 遞送時，此選徑器即可能發出重導遞送路徑訊息告知來源主機將資料片轉交由另一部選徑器遞送，意即，將資料片的遞送路徑重導至較佳的路徑。

・主機輸出量控制 (Type 4 、 Source Quench)

此訊息用在抑制特定主機的資料片輸出流量，當任一主機 (可能是選徑器或普通主機) 感覺來源主機送出資料片的速度過快時，會利用 ICMP 送出此訊息要求來源主機降緩送出資料片的速率。

・探索遠端主機

在 Internet，與遠方主機的連接隨時都有可能中止，一般是利用 ping 或 traceroute 之類的工具來檢查點與點之間的導通狀況，以 ping 為例，它即利用了 ICMP 的回音訊息 (Type 8) 向目標主機發出回音要求，若目標主機確有收到，即會發出回音回應訊息 (Type 0)，要求主方若到回應，即知道雙方之間的連接是正常的，順便也可計算這些訊息一來一返之際所花的時間，多進行數次這樣的偵測、並將所得的時間值加以平均之後的數值也可做為雙方之間傳輸效率的評估。

以上是 ICMP 主要訊息的說明，其它訊息在網管上可能不若前者來得重要，有興趣的讀者可查閱進階書籍及 RFC 文件。

13.3.3 、 位址解析協定 ARP 與 RARP

TCP/IP 利用 IP 位址定址，其下層的通訊介面也有自己的定址方式，其位址稱作物理或硬體位址，例如，Ethernet 的 hh.hh.hh.hh.hh.hh，通訊介面並不認得 IP 位址，須將它映射成當地網路的硬體位址，所以，在通訊介面及 IP 之間須存在一機制進行位址的映射。

負責將 IP 位址映射為硬體位址的是位址解析協定 (ARP)，ARP 會在系統內動態維護一份 IP 位址與硬體位址的對照表，當 ARP 被要求進行映射時，它先檢查其對照表，若於其中發現要求的 IP 位址，則傳回對應的硬體位址，否則，ARP 會在當地區域網路廣播內含欲映射的 IP 位址封包，當地網路的所有主機皆會收到該封包，若其中的一部發現該封包上的 IP 位址和自己吻合，則會回應另一個內含它的硬體位址的封包，發問的 ARP 收到後，除了將對應的硬體位址傳回要求者外，也將它擺在自己的動態對照表。

系統有時也須將硬體位址映射成 IP 位址，這時，位址反解析協定 (RARP) 即派上用場，常見的情形是在未裝設磁碟機的工作站，這類系統僅安裝網路介面，在初開機時，它們得透過網路介面發出 RARP 廣播封包，向伺服主機詢問自己的 IP 位址。

13.4.0 、 傳輸控制協定 (TCP)

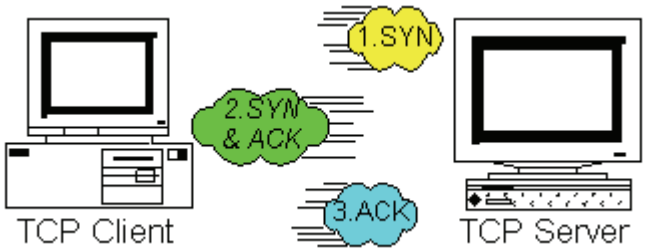
TCP (Transmission Control Protocol) 是端對端傳輸層內最重要的協定之一，另一個同級協定是 UDP (13.4.1 節)，這兩個協定負責在程序應用層及網際網路層之間搬移資料，TCP 的功能包括

- 提供連線導向式及可信賴的端對端資料傳輸服務。
- 滑動窗式流量控制。

所謂連線導向 (connection-oriented) 係指 TCP 首先利用控制資訊和對方建立連線，也就是連線前的握手 (handshake) 動作，之後再傳送資料，最後還有終止連線的動作。

TCP 採三向式握手(three-way handshake)建立連線，首先，由客戶端向伺服端發出SYN訊息，表示要求建立TCP連線，若伺服端接受連線，則回應SYN/ACK訊息，客戶端收到之後再回應ACK訊息，然後即可開始傳送資料。

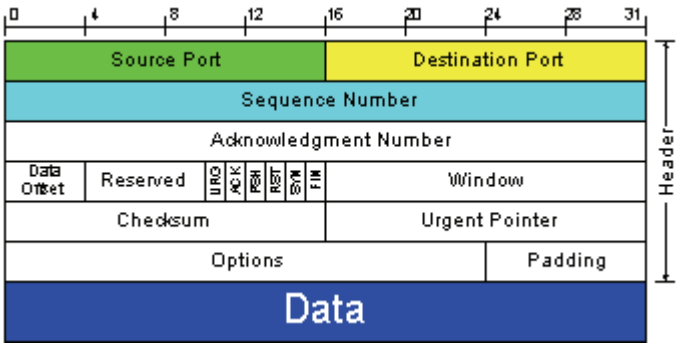
圖 13.19 、 TCP 之三向式握手(Three-Way Handshake)



TCP稱呼它的資料塊為資料段(圖 13.20)，每個資料段由一TCP添加的表頭與實際資料所組成，為令傳送的資料可信賴，TCP運用數種技巧，首先，為令資料段之間在傳送時不至失去原有的次序，TCP在其資料段表頭內設置了序號(sequence number)，接收端TCP可依照序號將之重組回原始資料的順序。

TCP資料段可承載的最大資料段長度(MSS)一般預設為1460位元組，當然並不一定得採用此值，此值的理由在於目前最普遍的區域網路介質是Ethernet，其最大傳輸單元(MTU)為1500位元組，將此值減去標準的IP表頭(20位元組)及TCP表頭(20位元組)的長度即得到1460，較正確的計算方式是將當地的傳輸介質的MTU - 40方可獲得正確的MSS值，MSS的值若過大，則在IP時段即得拆卸成數個更小的封包，增加更多額外的IP表頭，MSS的值若過小，同樣長度的資料即得分更多批傳送，同樣也會增加TCP的表頭數量。

圖 13.20 、 TCP 資料段(Segment)結構



當接收端收到一段資料即會回應一個認可(ACK)訊息，其中包含一所收到資料段之序號+1的認可號碼(acknowledgement number)，TCP在回應認可訊息時還使用一挾帶(piggy backing)技巧將認可訊息附帶在可能有的資料段中一併傳給對方，如此可降低單獨的認可訊息對網路頻寬的耗費。

TCP另一個確保資料正確傳輸的技巧是正向認可與重傳(PAR)機制，發送端若於指定時段內未收到另一端對於已送出資料段的認可訊息時，會重新送出相同的資料段，TCP將重複嘗試數次，直到對方回應認可訊息之後

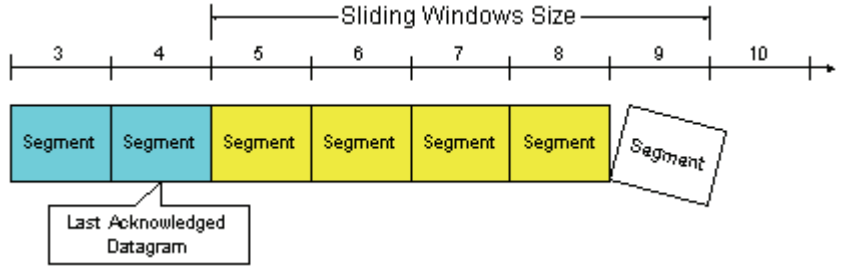
再送出下個資料段，若重複嘗試失敗，則TCP將通知應用層"失去連線"這類的訊息。

TCP也採核對合(checksum)計算資料段的正確性，該核對合位於資料段的表頭，當TCP收到資料段時，會將它所計算的合與表頭的核對合相比對，若相同，則送出認可訊息，表示接收無誤，否則忽略該資料段，在一小段等待之後，對方會再次送來同一個資料段。

在滑動窗式(sliding window)流量控制方面，TCP的考量是，由於每個資料段在被送出之後不會立即抵達目標，而是會有一段時間的旅行，為避免因等待接收端的認可訊息所造成的閒置，發送端可在未收到前一個資料段的認可訊息的情況下持續發送數個資料段，此處的數個即是所謂的滑動窗長度。

以圖 13.21 為例，其中的號碼為資料段的序號，由圖可看出，發送端已收到資料段3、4的認可訊息，換言之，接收端已"確實"收到這些資料段，發送端雖尚未收到5、6、7等資料段的認可訊息，但由於尚未超出滑動窗的範圍(5～9)，故可繼續送出資料段，直到送出第9個資料段之後，即得停下來等待接收端的認可訊息，由於在傳送過程中，發送端隨時可收到接收端的認可，故滑動窗的起始位置將不斷地往前移動，該名稱的由來即緣於此。

圖 13.21 、 TCP之滑動窗(Sliding Window)



滑動窗的長度視網路連線狀態而定，通常可由參數設定之，但參數設定值僅作為參考之用，TCP尚會依據實際的連線品質自動調整滑動窗長度。

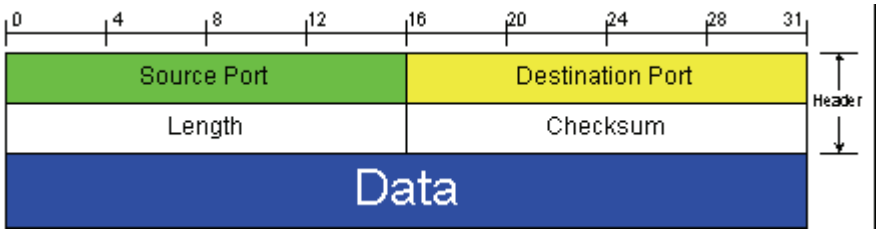
易言之，TCP的功能在於提供上層各種應用程式一組無錯誤的連線管道，UDP雖與TCP同等級，但各有各的使用時機，對於資料量較小、且資料的正確性不是很重要的情況可考慮使用UDP(如RIP、DNS)，而在資料的正確性要求較嚴格的狀況下(如TELNET、FTP)，TCP即是較佳的選擇。

13.4.1 、 用戶資料片協定(UDP)

UDP(User Datagram Protocol)和TCP屬於同等級的協定，它們的功能也相仿，只是和TCP相較下，UDP提供較高速、但不可信賴、非連線導向的資料傳輸服務，意即，它既不事先與對方建立連線(握手)，也不偵測傳輸過程可能引入的錯誤。

UDP 稱呼它的資料為訊息(message)，訊息表頭較 TCP 的小，且 UDP 無煩瑣的握手、認可、重送等動作，故在同一網路環境下，其效率較 TCP 為高。

圖 13.22 、 UDP 之訊息(Message)結構



由於 UDP 不可信賴的特性，對於資料量較小、且資料的正確性不是很重要、或是其上層應用程序可自行驗證資料正確性的情況下可考慮使用 UDP (例如 RIP 、 DNS)，其它情況則須使用 TCP 。

13.5.0 、 應用層服務

應用層的程序通常扮演著人機介面的角色，也是整個 TCP/IP 協定堆疊中、最接近人類用戶的一層，常見的 Telnet 、 FTP 、電子信件等服務皆屬於此層。

TCP/IP 的大部份應用皆是主從模式的服膺者，意即，一種分散式服務架構，位在網路的一端扮演客人，專門取得服務，另一端扮演主人，專門提供服務，且客戶扮演主動角色，主動提出連線要求、主動發出命令、主動結束連線，這也是伺服器之所以取名伺服的主因，"伺"也者，"等待"之意。

在技術上，應用層的伺服程序隨時監聽 TCP 或 UPD 特定埠是否有來自客戶端的要求送抵，例如，SMTP 伺服器程式將一直監聽埠 25 是否有連線要求，若有，即建立連線，然後根據客戶端的要求提供對應的服務。

客戶程序可透過 TCP 或 UDP 與伺服程序建立連線，在佔多數的連線導向式應用上，一旦 TCP 的三向式連線建立完成，服務端即會回應一個初始化訊息，然後等待客戶端送來命令，當服務結束時，客戶端會送出一個終止命令，並等待服務端的回應，之後關閉 TCP 連線，通常在雙方之間傳達的命令與回應皆是文字列型態(以 CR/LF 結束)，其中的字元則採 NVT ASCII 字元集。

NVT ASCII 是 ASCII 的美國版變體，它是種每個字元長 7 個位元的字元集，當傳送時，發送端會先清除字元的最高位元，接收端則忽略該位元，這也是為何 8 位元的中文信件無法順利於 Internet 傳送的主因， NVT ASCII 字元集將所有字元分成三類

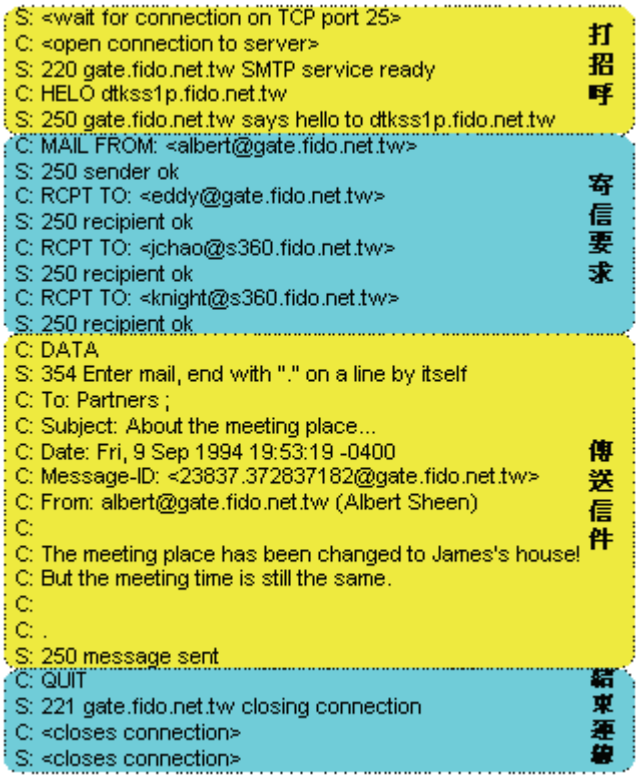
- 控制碼　　其值介於 0 至 31
- 圖形碼　　其值介於 32 至 126
- 未定義的碼　其值介於 127 至 255

NVT ASCII 的規格中，每一列皆以一組 CR/LF 結束，若僅欲表達單一的 CR 或 LF，則在 CR 或 LF 之後須伴隨一個 NUL (ASCII 0)。

應用程序使用的命令內含一個關鍵字與伴隨其後的零或數個引數，引數之間以空白字元隔開，回應則包含一組結果指示符號(或稱回應碼)及一些文字資訊，這些文字在人類看來非常直覺的，我們可由這樣的訊息清楚觀察到整個對談的流程。

圖 13.23 是個 SMTP 交談樣本，其中，S: 代表服務端，C: 代表客戶端，其中可看出，一開始，伺服端監聽著埠 25，接著，某客戶端向該 SMTP 伺服端發出連線要求，然後雙方建立連線，連線初建立時，雙方先向對方打聲招呼以確定對方身份，接著，由客戶端發出寄信要求，在服務過後，即由客戶端主動發出結束連線要求，此交談過程已經過筆者簡化，實際的交談內容會較此處的樣本還複雜些。

圖 13.23 、 SMTP 主從對話樣本



13.5.1 、 埠號碼 (Port Number)

在應用層的每種服務皆有個唯一的埠號碼，例如 TELNET 的埠號碼是 23 、FTP 的是 21 ，當 TCP 或 UDP 由 IP 收到資料後，會根據表頭的埠號碼將資料轉交對應的程序處理，須注意的是，TCP 及 UDP 的應用程序可分配到相同的埠號碼，必須配合埠號碼及傳輸協定種類才可決定資料所對應的程序。

許多埠號碼已保留予一些著名(well-known)的網路服務之用，如 TELNET 、FTP 、 SMTP 、 DNS 此類服務皆是網路上經常用到、且已成為標準的服務，

所有著名服務的埠號碼皆記錄在一份名為已配置號碼(Assigned Numbers)的RFC，目前是RFC-1700，此份文件由IANA負責維護(13.1節)。

埠號碼0至255已保留給著名服務，256至1023則分配給Unix特有的服務，僅管當初的分配方式是這樣的，但其中大部份的服務已不再是Unix特有的，後來，IANA也已將著名服務的埠號碼擴展為0至1023，介於1024至65535的號碼則未定義，IANA並未限定這些埠的用途，主機可自行決定如何配置，一般是用在動態配置埠，稍侯說明。

由於著名埠(well-known port)的號碼已標準化，故任兩部主機在建立某個著名服務的連線之前即已知道須使用的埠號碼，如此可簡化雙方的連線程序，例如，所有Internet的主機皆統一透過埠23提供TELNET服務，當使用者欲以客戶端TELNET程式登入遠端主機時，即不須特別指定埠號碼。

表13.24、著名服務的埠號碼(節錄)

服務名稱	埠號碼	傳輸協定	服務名稱全名／服務描述
tcpmux	1	TCP	
echo	7	UDP	
echo	7	TCP	
systat	11	TCP	
netstat	15	TCP	
ftp-data	20	TCP	File Transfer Protocol (data)
ftp	21	TCP	File Transfer Protocol
smtp	25	TCP	Simple Mail Transfer Protocol
time	37	TCP	Time Server
time	37	UDP	Time Server
name	42	UDP	Name Server
whois	43	TCP	nicname
domain	53	UDP	
domain	53	TCP	
tftp	69	UDP	
rje	77	TCP	
finger	79	TCP	
link	87	TCP	ttylink
supdup	95	TCP	
hostname	101	TCP	hostname
pop-2	109	TCP	Post Office Protocol
uucp-path	117	TCP	
nntp	119	TCP	Network News Transfer Protocol

ntp	123	TCP	Network Time Protocol

TCP 資料段 (segment) 或 UDP 訊息 (message) 的表頭內含兩個埠號碼，其一是來源埠 (source port)，其二是目標埠 (destination port)，理論上，當 TCP 或 UDP 收到資料後，只須根據表頭的埠號碼將資料轉交對應的應用程序即可，並且，來源程序的埠號碼應與目標程序的相同，所以雙方應只須共用同一個埠號碼，但實際上可能遇到的狀況是，同一個應用程序執行數份的情形，例如，當用戶同時執行兩份 TELNET 程式登入兩部主機，此時，光憑一個埠號碼無法辨識資料是屬於哪一份 TELNET 程序的，解決方式是巧妙利用來源埠、目標埠、及動態配置埠。

13.5.2 、 動態配置埠

動態配置埠 (dynamically allocated port) 是當應用程序須要它們時才由系統動態地配置出，其埠號碼範圍是 1024 至 65535，其目的在令同一個 TCP/IP 應用程序能同時存在許多份，其運作邏輯是，客戶端採動態方式配置埠號碼，伺服端則直接使用著名埠號碼。

以 TELNET 為例，假設一名位於 192.72.5.2 的用戶先執行第一份 TELNET 程式登入主機 192.72.5.1，此時，192.72.5.2 為此份 TELNET 程序動態配置一個埠編號 3001，接著，該用戶又執行另一份 TELNET 登入同一部主機，這時，該 TELNET 被分配到的埠號碼是 3002。

在第一份連線建立之初，雙方進行三向式握手程序 (圖 13.19)，192.72.5.1 的 TCP 收到來自 192.72.5.2 的 SYN 訊息，其中，資料段表頭的來源埠為 3001、目標埠為 23，由目標埠，192. 72.5.1 知道對方要求 TELNET 服務。

在第二份連線建立之初，同樣的，192.72.5.1 的 TCP 由目標埠知道對方要求 TELNET 服務，但該 TCP 也發現此次連線有別於上次的，因為它們的來源埠不同，於此例，該用戶建立的兩組連線可分別表示成

192.72.5.23001 → 192.72.5.123
192.72.5.23002 → 192.72.5.123

由此我們也可觀察出 IP 位址與埠號碼相結合之後的唯一性，意即，一個 IP 位址與一個埠號碼可構成一個插槽 (socket)，一個插槽即可表達在整個 Internet 中唯一的一個網路程序，兩個插槽即構成在整個 Internet 中唯一的一組連線，有人也將著名服務一詞稱作著名插槽。

13.6 、 TCP/IP 協定堆疊總覽

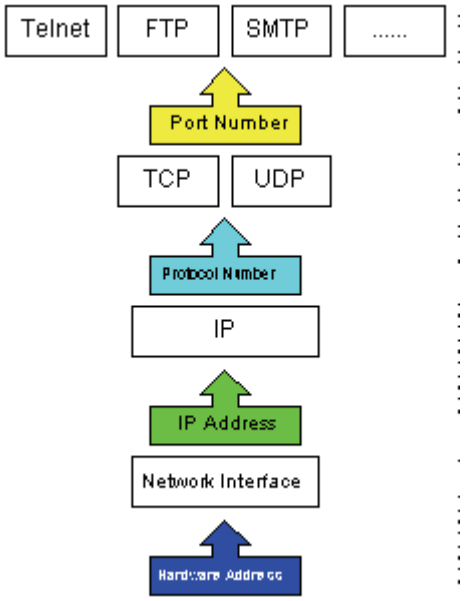
前幾節，我們介紹了 TCP/IP 協定堆疊的幾個主要協定，讀者可能有見樹不見林之感，本節，筆者即舉一簡單的應用實例以便對整個協定堆疊來個總瀏覽。

雙方在通訊時，資料的來源通常是用戶透過應用程式發出的，例如收發電子信件、傳檔等，目標則是另一端的應用層，以 SMTP 的應用為例，假設主機 α 的用戶正利用程式發信至主機 β。

首先，α 的 SMTP 客戶端透過 TCP 與 β 的 SMTP 伺服器連線，TCP 將應用層送來的資料串流 (stream) 封裝成小包資料，也就是資料段 (segment)，並於這些資料段表頭登錄其上層協定的埠號碼 25 (代表 SMTP)，之後，TCP 再將這些資料段交由 IP 傳送，IP 將這些資料段封裝成資料片 (datagram)，並將其上層協定的協定號碼 6 (代表 TCP) 及來源與目標主機的 IP 位址登錄於資料片表頭，接著，IP 根據目標的 IP 位址透過網路存取層的協定將資料片送抵目標主機。

以上是發送部份，當目標主機的網路存取層協定收到資料包時，它先判別該封包是否為 IP 資料片，若是，則交由 IP 協定處理，IP 先確定該資料片上登錄的目標位址的確是該主機的，接著，根據資料片表頭的協定號碼 6、將剝去 IP 表頭的資料段交由 TCP 處理，同樣的，TCP 由該資料段的表頭發現目標埠號碼為 25，所以將資料交由 SMTP 伺服程序處理。

圖 13.25 、 TCP/IP 目標識別方式



主機間的資料大抵是以此種形式流傳著，每當資料流經某協定之際，該協定即須藉由表頭中的資訊判別資料的下個目標協定為何，在接收資料 (由下而上) 的過程中，IP 利用協定號碼識別傳輸層的協定，傳輸層協定利用埠號碼識別應用層的程序，在送出資料 (由下而上) 的過程中，應用層的程序透過傳輸層協定送出資料，傳輸層協定將資料轉交 IP 送出，IP 根據目標位址決定遞送路線，若在目標位址位在同一網路，則直接將資料送交對方，否則將資料轉交選徑器繼續遞送。

13.7.0 、 選徑資訊協定 (RIP)

RIP 源於全錄 (Xerox) 公司的 XNS 網路系統，它的正式定義文件發表於 1981 年，並在 1982 年與 BSD Unix 及 TCP/IP 相結合，RIP 在 Unix 環境是以選徑伺服器 (routed) 的形式提供服務，目前也是 TCP/IP 網路普遍採用的動態選徑機制。

許多廠商也將 RIP 移植至他們的網路產品中，例如 Mac 的 AppleTalk 協定的 RTMP 即修改自 RIP，Novell、3Com 的產品則直接引用 Xerox 標準的 RIP，Banyan、Ungermann-Bass 等廠商則對 RIP 小幅修正以滿足各自產品的需求。

有別於傳統以手工設置主機的靜態選徑表 (routing table)，動態選徑允許多部選徑器 (router) 彼此交換選徑資訊，自動進行動態的選徑工作，主機可由這些選徑器詢問到即時的選徑資訊，以獲得離目標主機最近的封包遞送路線。

RIP 主要用於企業內部大型區域網路的動態選徑，RIP 服務基本上維護了一份動態選徑表，其中登錄了每一目標主機的相對距離、及下個最近的選徑器，距離代表其間的選徑器數目，RIP 以傳統的最短距離演算法提供詢問者一最佳的遞送路徑，當網路拓撲發生變動時，例如選徑器偵測到另一選徑器當機、或網路上出現新的選徑器時，RIP 會將新的變動訊息廣播至所有支援 RIP 的選徑器。

圖 13.26、典型的 RIP 選徑表

目標 (destination)	下一站 (next hop)	距離 (distance)	計時器 (timers)	旗號 (flags)
網路 A	選徑器 2	5	t1,t2,t3	x,y
網路 B	選徑器 1	2	t1,t2,t3	x,y
. . .	. . .	. . .	. . .	. . .
網路 X	選徑器 2	4	t1,t2,t3	x,y

由一選徑器至另一選徑器的這段路徑也暱稱一跳 (hop)，除了相隔的選徑器數目之外，網路間的傳輸線路流量也是選徑距離的因子，一般也以尺度 (metric) 或成本 (cost) 來描述主機間的"距離"。

RFC-1058 描述的 RIP 規格有些地方不太適合大型區域網路，例如，一遞送路徑可經過的最大距離 (選徑器數目) 為 15，又，由於選徑資訊的更新有段時差，這有可能導致選徑迴圈，另外，每段路徑的尺度須是固定值，這樣的設置即無法以尺度隨時反應動態變化的傳輸線路流量。

參與動態選徑的主機間以 RIP 封包 (圖 13.27) 交換資訊，RIP 利用 UDP 的埠 520 收發封包，此封包的固定表頭長 32 位元，其後可接至多 25 筆相同格式的訊息，其中的訊息內含該網路的選徑資訊，每個封包皆有固定作用，此於其命令 (Command) 欄描述，參考 表 13.28 。

圖 13.27 、 RIP 封 包 結 構

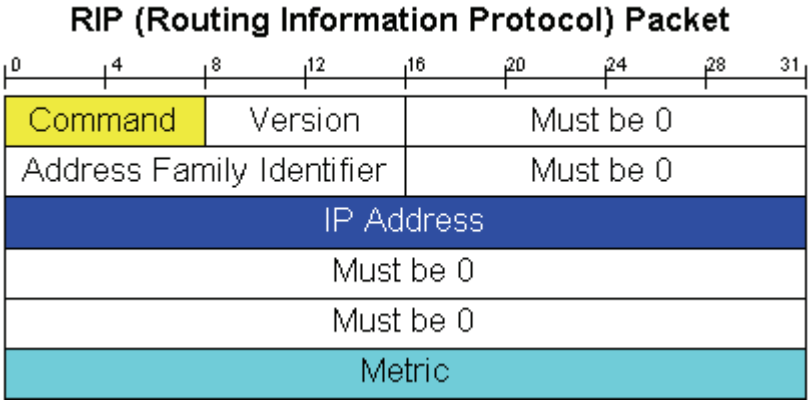


表 13.28 、 RIP 封 包 命 令

命 令 碼	描 述
1	要 求 (Request) 選 徑 表 資 訊
2	回 應 (Response) 選 徑 表 資 訊
3	追 蹤 開 啟 (Traceon) (此 命 令 已 宣 布 停 用)
4	追 蹤 關 閉 (Traceoff) (此 命 令 已 宣 布 停 用)
5	Sun Microsystems 保 留 命 令

僅 管 RFC-1058 定 義 了 RIP ， 但 目 前 有 許 多 區 域 網 路 作 業 系 統 並 不 完 全 遵 照 該 文 件 描 述 的 規 格 ， 這 會 造 成 不 同 產 品 間 的 RIP 無 法 協 調 的 問 題 ， 導 致 網 路 上 存 在 各 種 產 品 各 自 的 RIP 服 務 ， 彼 此 不 相 關 聯 。

為 因 應 日 益 複 雜 的 網 路 社 會 ， 新 版 的 RIP-2 新 支 援 了 選 徑 網 域 (routing domain)、 外 部 選 徑 標 籤 、 次 網 路 遮 罩 、 下 一 跳 (next hop) 位 址 、 及 認 證 等 。