IP 通訊協定解析

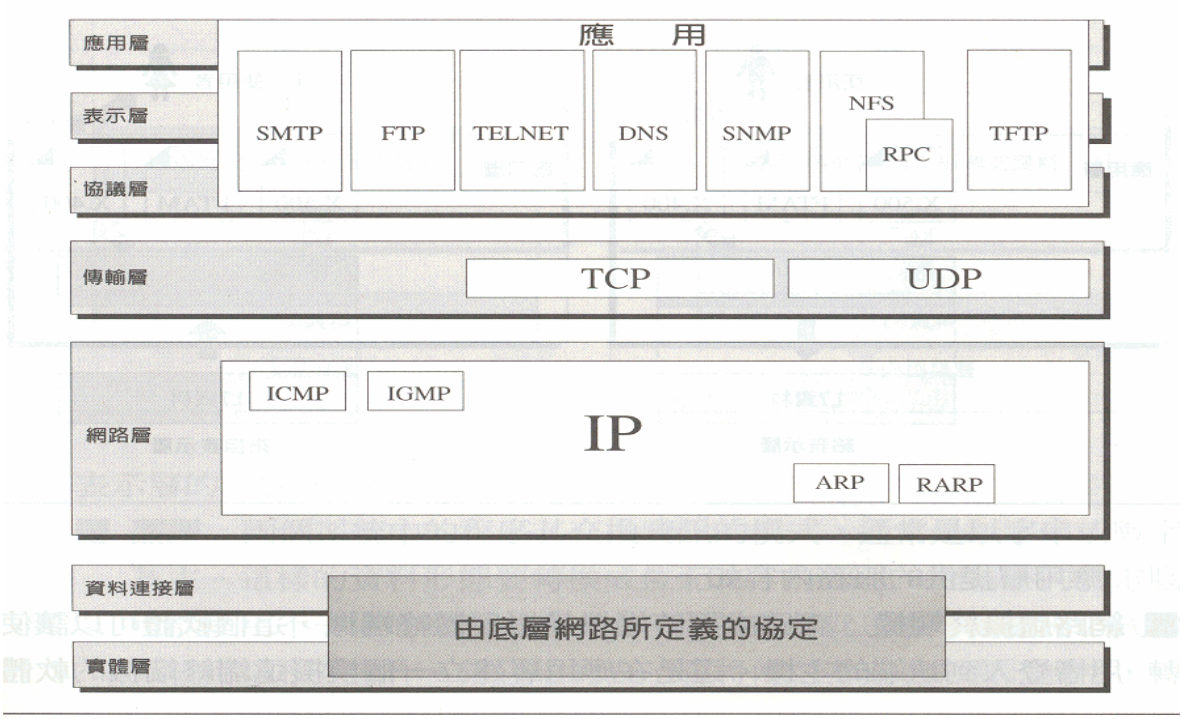
孫善傑

eMail: petersun@mail.ur-solution.com
www : http://www.ur-solution.com

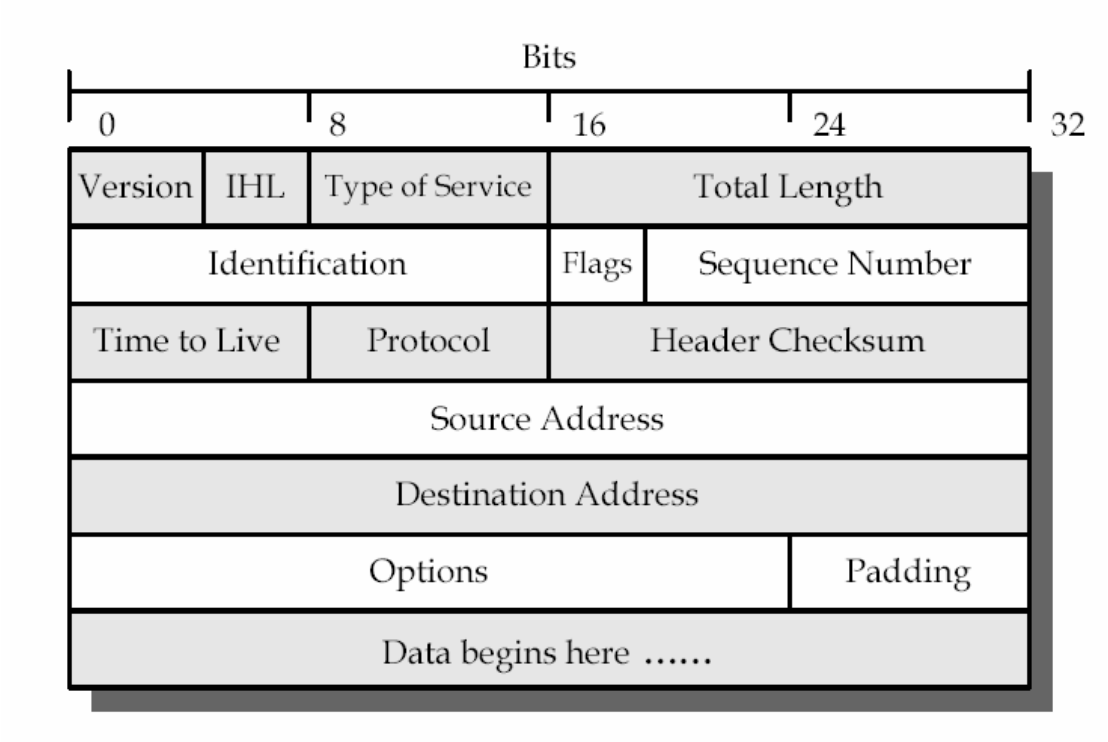
一， **導讀** **3**
二， **IP封包** **4**
1. VERSION(版本) 4
2. IHL(標頭長度) 4
3. TYPE OF SERVICE(服務種類) 4
4. TOTAL LENGTH(總長度) 5
5. IDENTIFICATION(識別代碼) 6
6. FLAG(旗標) 6
7. SEQUENCE NUMBER(分段差量) 6
8. TIME TO LIVE(TTL/存活時間) 6
9. PROTOCOL(協定) 7
10. HEADER CHECKSUM(檢查碼) 7
11. SOURCE ADDRESS(來源端地址) 7
12. DESTINATION ADDRESS(目的端地址) 7
三， **分段(FRAGMENTATION)** **7**
1. 最大傳輸單位(MTU) 8
2. 與分段有關的欄位 9
四， **OPTIONS(選項)** **10**
1. 格式(FORMAT) 11
2. 選項種類 12
五， **檢查碼(CHECKSUM)** **15**
1. 傳送端 CHECKSUM 之計算 15
2. 接收端 CHECKSUM 之計算 16
3. IP封包使用的 CHECKSUM 16
六， **IP設計** **17**

一， 導讀

TCP/IP 爲網際網路所使用的傳輸機制。IP 是一種非可靠性、非預接式的資料封包協定。IP 僅提供一種盡力而爲的服務，也就是說 IP 沒有提供錯誤檢查或追蹤的機制。IP 會盡全力將封包送達目的地，但是不保證一定能收到。如過可靠性對你是很重要的，那 IP 就必須與俱備可靠度傳輸的協定，像是 TCP 一起搭配使用。IP 是一種非預接式協定，這表示每一個封包都是獨立的個體，將會個別被處理，封包可以走不同的路徑到達目的地。這表示是送出的個別封包，會走不同的路徑到達目的地，這些資料到達的順序可能會不一樣，也許其中一些封包會遺失或者在傳送中受到損壞。因此 IP 必須依靠更高一層的協地來處理問題。



二， IP 封包



1. Version(版本)

這 4 個位元定義 IP 協定的版本。目前是第四版，如果某一台電腦用的是別的版本，則所收到的封包要被丟棄掉，而非以不正確解釋來處理。

2. IHL(標頭長度)

定義標頭的總長度，單位是 word，而每個 word 是 4 Bytes。因爲標頭的長度在 20~60 Bytes 之間，所以必須有此欄位。在沒有選項(Options)時，標頭的長度爲 20 Bytes，此時本欄位是 5 (5 X 4 =20)，當選項全部都使用時，標頭長度爲 15 (15 X 4 = 60)。

3. Type of Service(服務種類)

定義路由器要如何處理這個封包。本欄位分爲兩個子欄位即優先權（3 個 Bite）及服務種類（4 Bits），剩下一個位元不用。

D：最小延遲　　R：最大可靠度

R：最大單位流量　C：最小成本

			D	T	R	C	

優先權

A. 優先權爲一個 3 位元數字，由 0~7，定義封包在路壅塞時送出的優先權，如果一台路由器發生壅塞，需要丟棄封包時，優先權最低的先被丟掉。在 IPv4 中，此欄位沒有被使用。

B. 服務種類（TOS）有 4 個位元，每個位元都有其特殊意義。雖然每個位元不是 0 就是 1，但是這四個位元，一次只能有一個位元爲 1，，總共有五個不同的服務。應用程式可以要求某一特定服務。

TOS 位元	說明
0000	正常（預設）
0001	最小成本
0010	最大可靠度
0100	最大單位流量
1000	最小延遲

4. Total Length(總長度)

16 個位元。定義包括 IP 資料封包的標頭及資料總長度。單位是位元組，要找出上一層送來的資料長度，可將總長度減去標頭長度，標頭長度以 IHL 欄位的數值乘以 4 而獲得。

因爲總長度是 16 位元，所以 IP 封包資料總長度最大爲 65535 個位元組，其中 20 到 60 位元組爲標頭，剩下的即是來自上一層的資料。

預設服務

協定	TOS 位元	說明
ICMP	0000	正常
BOOTP	0000	正常
NNTP	0001	最小成本
IGP	0010	最大可靠度
SNMP	0010	最大可靠度
TELNET	1000	最小延遲
FTP（資料）	0100	最大單位流量
FTP（控制）	1000	最小延遲

TFTP	1000	最小延遲
SMTP（命令）	1000	最小延遲
SMTP（資料）	0100	最大單位流量
DNS（UDP 詢問）	1000	最小延遲
DNS（TCP 詢問）	0000	正常
DNS(zone)	0100	最大單位流量

5. Identification(識別代碼)

使用於分段（Fragmentation），稍後說明。

6. Flag(旗標)

使用於分段（Fragmentation），稍後說明。

7. Sequence Number(分段差量)

使用於分段（Fragmentation），稍後說明。

8. Time to Live(TTL/存活時間)

每一個封包都有一定的存活時間。由每個經過的路由器扣減，當數值爲 0 時，此封包即被丟棄。此欄位現在用來控制一個封包所能經過的路由器個數。當來源端送出封包時，它會設定一個 TTL 數值，這個數值大約是任一兩台電腦可能通過所有路由數目的兩倍。路由器在收到一個封包後，會將此封包內的 TTL 數值減一，如過減了以後的數值爲零，該路由就將此封包丟棄。倘若路由器故障了，封包可能在兩台或多台路由器間繞來繞去，形成資源浪費，因此 TTL 限制了一個封包的生命。

TTL 另一個用途是讓來源端限制封包旅行的距離。例如，來源端要將封包限制在區域網路內，TTL 就設定成 1，當這個封包到達第一個路由器時，TTL－1 會等於 0，即被丟棄。

9. Protocol(協定)

8 個位元的欄位，定義 IP 的上層服務。IP 封包可以包裝來自 TCP、UDP、ICMP 及 IGMP 等較高層的資料。這個欄位指定了 IP 封包最終目的要跑得協定。

數值	協定
1	ICMP
2	IGMP
6	TCP
8	EGP
17	UDP
41	Ipv6
89	OSPF

10. Header Checksum(檢查碼)

有關檢查的觀念及計算方式，稍後說明。

11. Source Address(來源端地址)

定義來源端的 IP。

12. Destination Address(目的端地址)

定義目的端的 IP。

三， 分段(Fragmentation)

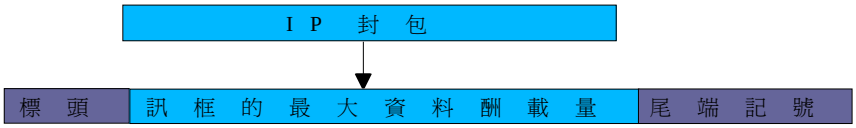
封包在網際網路上要經過不同的網路，每個路由器從收到的訊框分解出 IP 資料，經過處理將之將之再封裝成另一個訊框。路由器所收到的訊框格式與長短取決於該訊框所使用的的實體網路協定。相同的，路由器送出的訊框其格式與大小同樣取決於該訊框所使用的實體網路協定。

例如，某路由器將一乙太網路接到到一個記號環網路，那麼，它接收的訊框是乙太網路，

而送出的是記號環網路（Token ring）。

1. 最大傳輸單位(MTU)

每種資料連接層協定都有自己的訊框格式，而訊框中有一個欄位定義最大的資料酬載量。也就是說當一個資料封包將成訊框時，該資料封包的大小必須受限於訊框的最大資料酬載量。



最大訊匡資料酬載量因實體網路協定不同而不同，下表中展示不同協定的最大值。

協定	*MTU*
超通道	65535
記號環(16 Mbps)	17914
記號環(4 Mbps)	4464
FDDI	4352
乙太網路	1500
X.25	576
PPP	296

爲了讓 IP 協定與實體網路不相依，所以 IP 的封包最大長度等於表 3.1 中的最大值，即 65535 位元組。但是在表 3.1 中，每個網路的訊框長度都不一樣，因次我們必須依據實體網路的規格，將封包分成小段，才能在這些網路上傳輸，這個過程稱爲分段。

當封包被分段之後，每一個小段都會有自己得標頭，但是如果這些小段要經過 MTU 更小的網路，那這些小段還會在被分段，由此可知，一個封包在送達目的地之前可能被分段好幾次。

先前提過，IP 是一種非預接式的協定，所以被分段的封包會由不同的路由送出，我們無法控制一的被分段的封包會經過那一個路由。所以最後重組的工作必須由目的地端的電腦來完成。

當封包被分段之後，標頭會複製到每一個小段，但是選項欄位可能不會被複製，這些下一節我們將介紹。將資料封包分段的主機或則路由必須更改三個欄位的數值，即旗標，分段差量及總長度。其他的個欄位複製即可。當然 Checksum 的數值必須重新在計算過。

2. 與分段有關的欄位

- 識別代碼：

用來辨識來自來源電腦的封包，爲了確保唯一性，IP 協定使用一個計數器來作爲封包編號，所以只要計數器在主記憶體中，那封包的唯一性就可以確定，當一格封包被分段之後，這個編號也被複製到所屬的分段中，那目的端的電腦在重組封包時，就知道那幾個分段該組成同一個封包。

- 旗標

3 個位元的旗標，第一個位元保留不用。第二個位元稱爲不要分段欄位。如果爲 1，表示不要將此封包分段，但是因此而無法將封包送到實體網路時，該電腦或路由會將此封包丟棄，然後送一個 ICMP 的錯誤訊息給來源端電腦。假如數值爲 0，表示如果有需要，封包可以被分段。

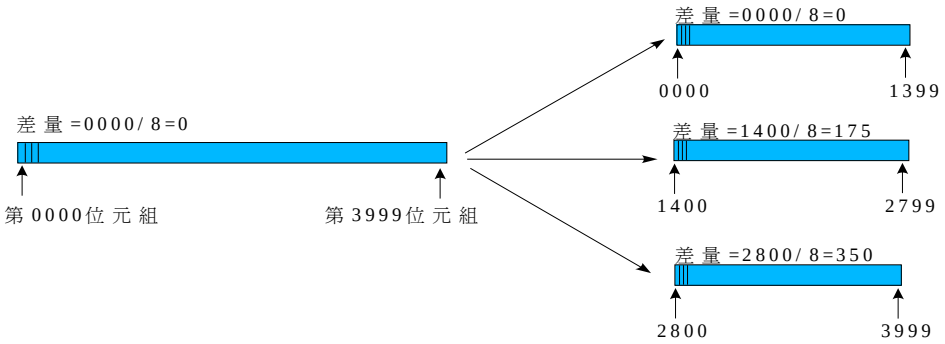
第三個位元稱爲尚有分段位元，如過數值爲 1，表示該封包不是最後的分段，後面還有更多的分段。如過數值爲 0，表示此封包是最後或者是唯一的分段。

D：不要分段

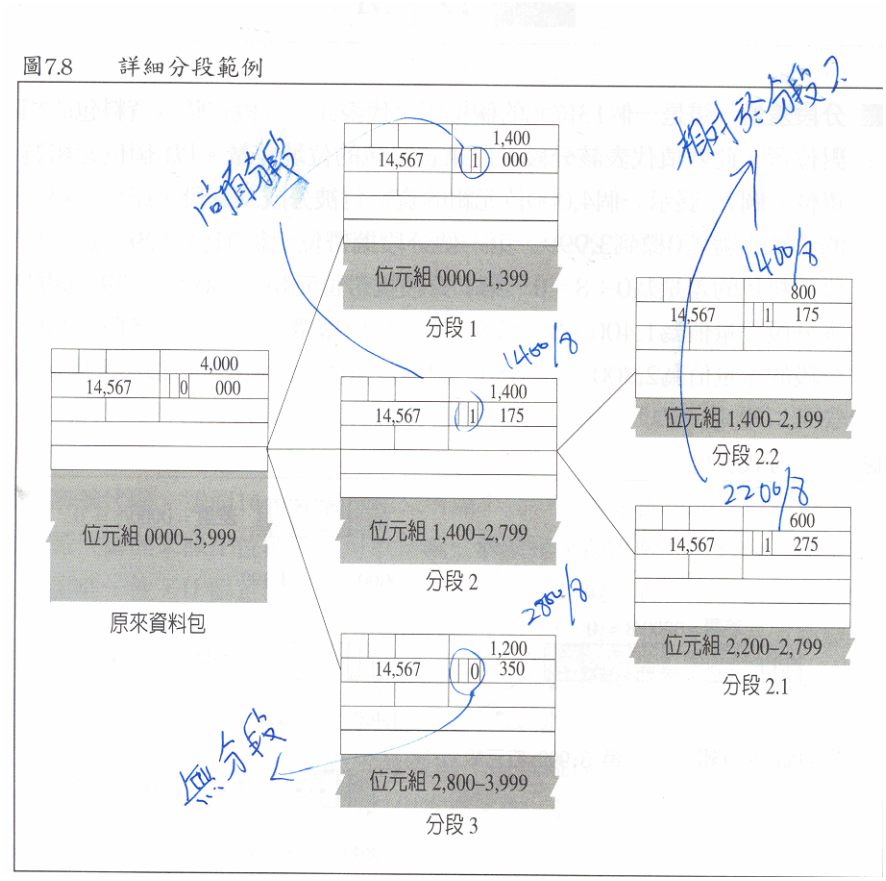
M：尚有分段

- 分段差量

這是一個 13 位元的欄位，用來表示此一分段在封包中的相對位置。此欄位的數值代表該分段在原來封包中的位置差量。以 8 個位元組爲單位。



上圖中展示一個 4000 位元組的資料封包被分成三段，第一個分段由 0 到 1399，這個分段的差量爲 0/8=0。第二個分段由 1400 到 2799，差量爲 1400/8=175。第三個分段由 2800 到 3999，差量爲 2800/8=350。請注意，差量的數值是以 8 個位元爲一單位。這樣做是因爲差量欄位只有 13 個位元，無法表示一串位元組超過 8191 的資料。同時以 8 位元組爲單位，電腦或路由在分段時必須選用的第一個位元組號碼要可被 8 除盡。



在上圖中，請注意所有的分段的辨識欄位數值都是一樣的。而旗標中的尚有分段位元除了最後一個分段外都是 1，同時每個分段的差量也表示在圖中。同時，我們也表現出一個分段再被分段的情形。對於這種情況，其差量是以鄉對於原來的封包爲主。例如分段 2 被分成兩個分段後，分別是 800 及 600 位元組。而這兩的分段差量是相對於原來的資料。所以說，假設沒有分段遺失的話，即使每一個分段走不同的路，不按照順序到達目的地，最終目的電腦可以依照下列的方法重組回原來的封包。

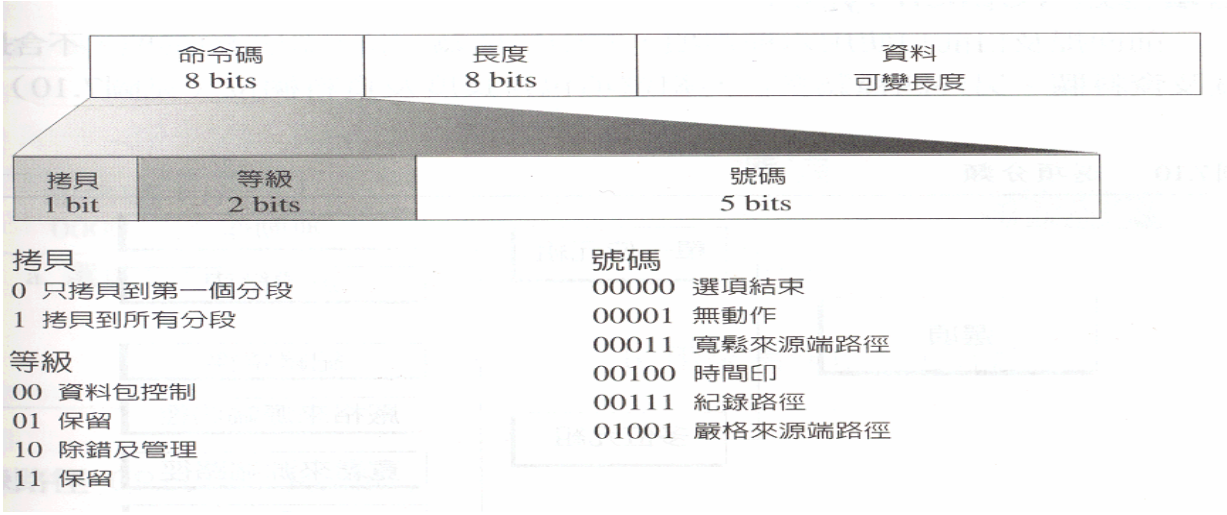
四， Options(選項)

IP 封包分爲兩部份，分別是固定部份及可變部份。固定部份爲 20 位元組，在前面已經介紹過了。而可變部份包含的選項可有 40 位元組。選項顧名思義不是每個封包都需要，選項是

給網路測試及除錯用，雖然選項不是 IP 標頭必要的一部分，但是對於選項的處理卻是 IP 軟體必備的。也就是說，如果選項出現在標頭，所有以 IP 爲標準的軟體必須能夠處理。

1. 格式(Format)

下圖是選項的格式，包括一個位元組稱爲命令碼(Code)，一個位元組的長度欄位及可變長度的資料欄位。



- 命令碼

8 個位元，包括 3 個欄位：拷貝(Copy)，等級(Class)及號碼(Number)。

- ✓ 拷貝(Copy)：這個位元控制選項如何出現在分段裡，當值爲 0 時表示選項只拷貝到第一分段，如果值爲 1 表示選項要拷貝到所有分段。
- ✓ 等級(Class)：這兩個位元定義選項的一般用途，當值爲 00 時表示選項是給封包控制用的，值爲 10 時表示選項是給除錯與管理用的，另外 10，11 目前尚未被定義。
- ✓ 號碼(Number)：這五個位元用來定義選項的種類。雖然 5 個位元可以定義 32 種不同的種類，但是只有 6 種在使用，稍後我們會來討論。

- 長度

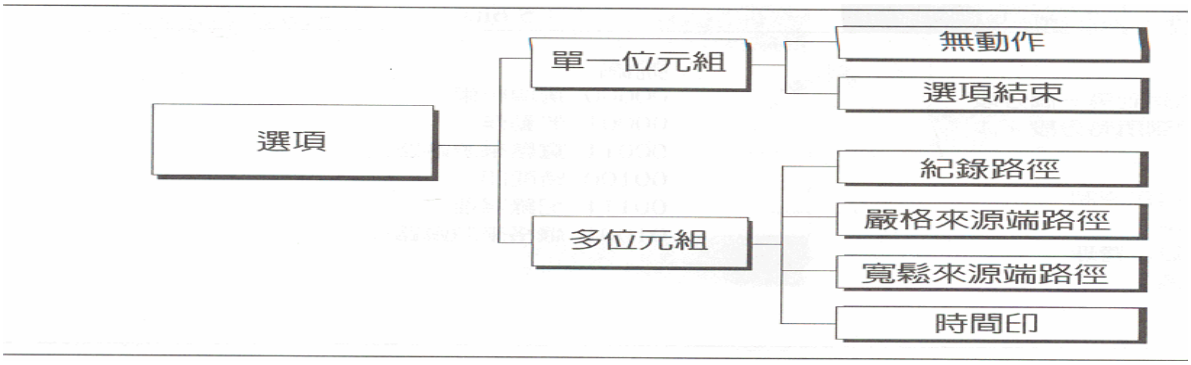
定義選項的總長度。包括命令碼及長度欄位本身，這個欄位不是在所有的選項中都會出現。

- 資料

資料欄位包含了某一選項所需的資料，與長度欄位一樣，這個欄位不是在所有的選項中都會出現。

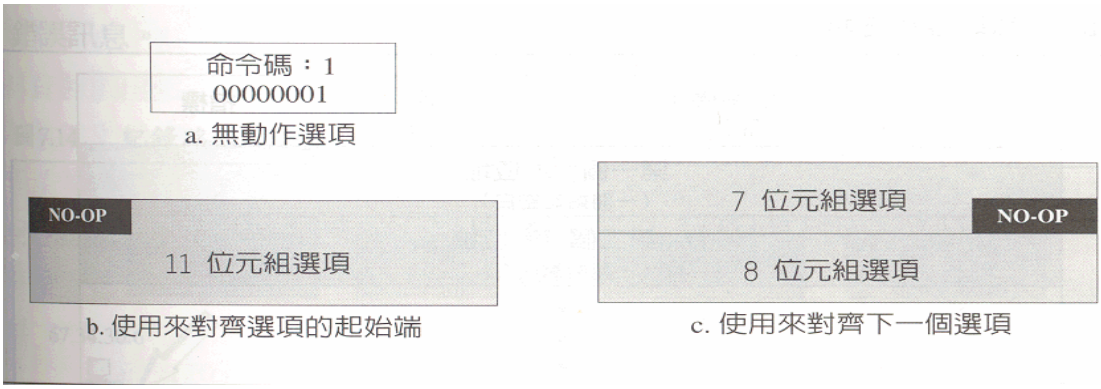
2. 選項種類

前面提到選項種類只使用 6 種，其中兩種爲一位元組長的選項，不含長度及資料欄。另外四種爲多位元組選項包含長度及資料欄。



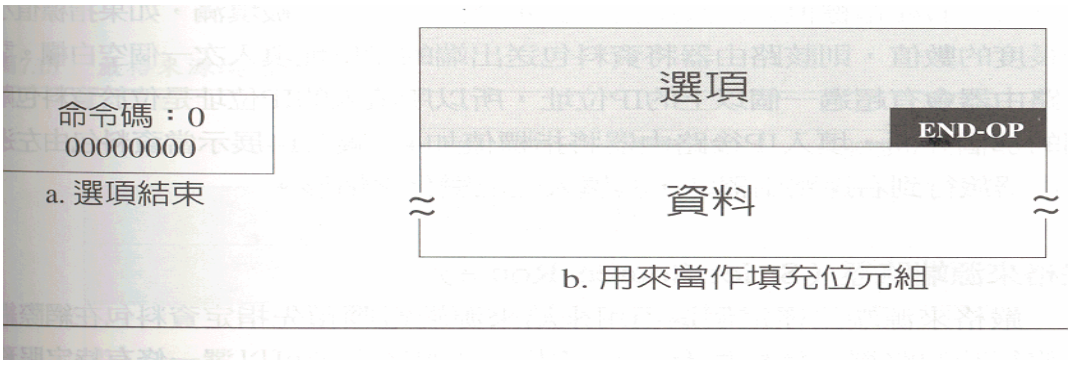
➢ 無動作(NO-OP)

這個選項的長度是一個位元組，它作爲兩個選項間的填補位元組之用。



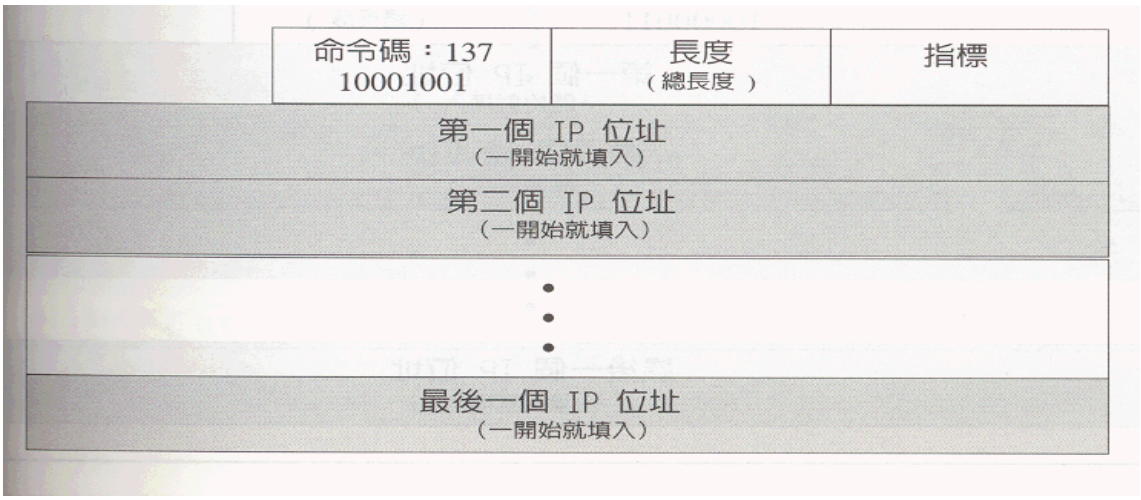
➢ 選項結束(END-OP)

這個選項也是一個位元組，用來補在選項欄的最後面。它只可作爲最後的選項，而且只能使用一次。在這個選項之後，接收者開始尋找封包的酬載資料，以就是說，如果需要用超過一個位元組來對齊選項欄，那麼就必須先用一些無動作選項然後再接一個選項結束。

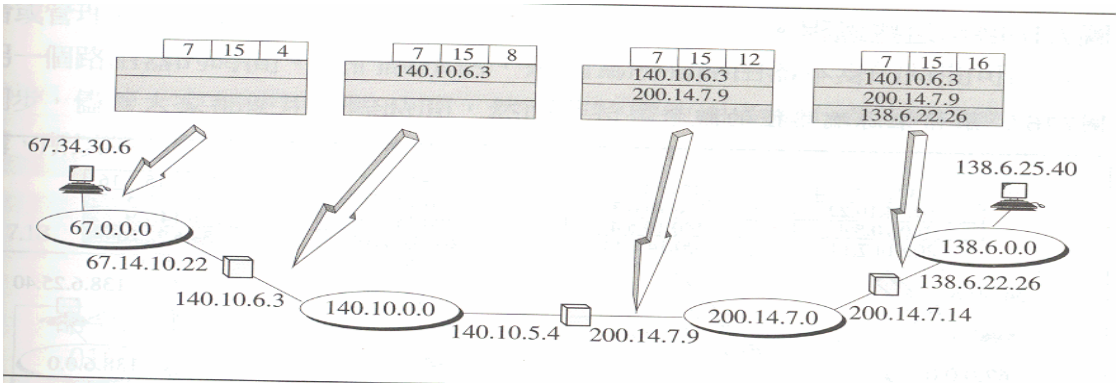


➢ 記錄路徑(Record Route)

記錄路徑選項是用來記錄封包所經過的路由器。它可以記錄到 9 個路由器的 IP 位置。因爲標頭最大爲 60 位元組，包括 20 位元組的固定部分，另外有 40 個位元組可以給選項用。來源端主機在選項裡空出放 IP 位置的欄位，這些空白欄位由所經的路由器分別填入其 IP。

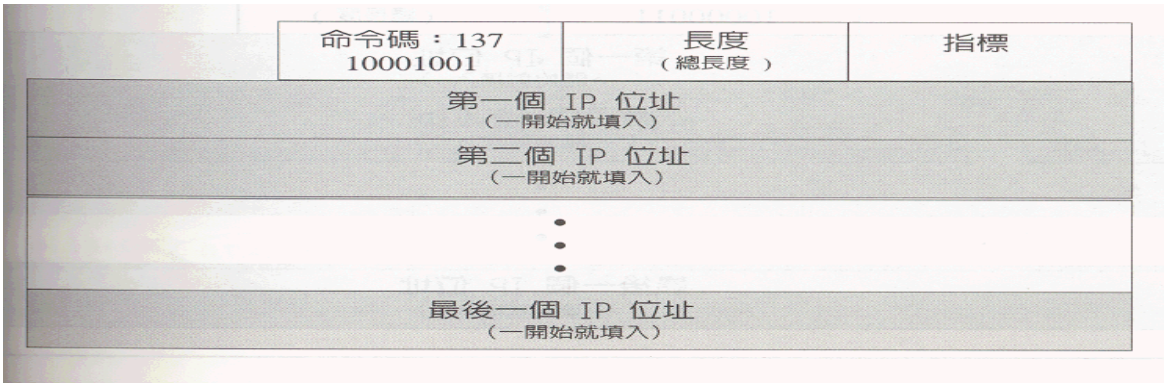


上圖中的命令碼及總長度前面已經提過了，而這裡的指標欄存放一個整數差量，代表第一個空白欄位的位元組位置，也就是這個指標（數值）指到第一個要存入的地方。當封包離開來源端主機，所建立的存放 IP 欄位都是空白，此時指標的值爲 4，指向第一個欄位，當封包經過一個路由器時，路由器會將其 IP 填入，此時指標的值會加 4，只向下一個空白的欄位。當指標的值大於長度值，表示空白欄位都已經被填滿。一個路由器會有超過一個以上的 IP，所填入的 IP 是封包離開的那個介面之 IP。



➢ 嚴格來源端路徑(Strict Source Route)

此指令是用來給來源端電腦指定封包在網際網路旅行時的路徑。這樣做有一個好處，送出者可以選一條有特定服務，例如最低延遲或者最大傳輸服務。或者送出者可以選擇安全一點或可靠一點的路徑，譬如送出者可以選擇一條路徑讓他們的封包不會經過其他競爭者的網路。如果封包由來源嚴格指定其路徑，那麼封包選項中所定義的所有路由器都必須經過，而沒有被指定的路由器該資料封包是不會通過的，如果資料封包路過一台不在指定路徑中的路由器，該資料封包會被丟掉，而將資料封包丟掉的路由器會發出錯誤訊息。如果封包到達目的地而部分指定的路由沒有經過，這個封包一樣會被丟棄，然後發出錯誤訊息。

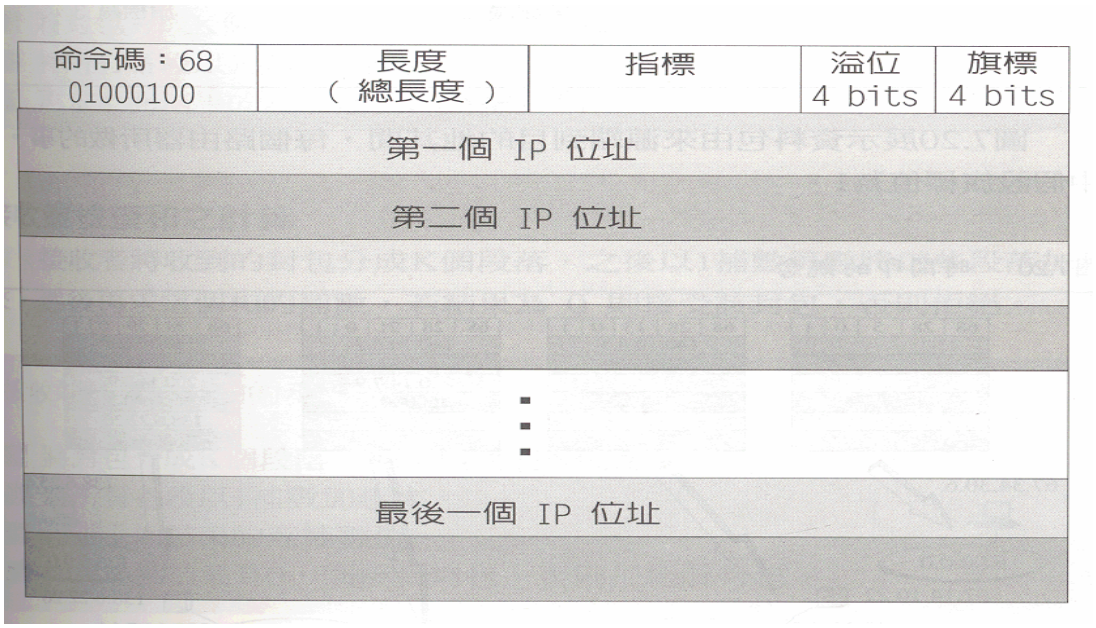


➢ 寬鬆來源端路徑(Loose Source Route)

與嚴格來源路徑相似，不過它放寬了一些限制。寬鬆來源路徑中所有的路由器都必須經過，但是封包也可以經過其它路由器。

➢ 時間印(Timestamp)

用來記錄路由器處理封包的時間。用國際時間，從午夜計算，千分之一秒爲單位。知道封包被處理的時間可以讓使用者追蹤網際網路上路由器的行爲，可以預估封包從一個路由器到另一個路由器的時間。我們說預估，是因爲所有路由器本身的時間可能沒有同步，儘管大家都使用國際時間，然而分系統管理員身份者都不清楚其網路架構，所以 Timestamp 這個選項也不是大多數人能用的。



上圖中有一個溢位欄位，用來記錄因爲 IP 空位不夠不能將時間記下來的路由器數目。旗標欄位規範路由器的任務，如果旗標爲 0，路由器填入的時間到所提供的欄位表。如果旗標爲 1，路由器填入送出端的 IP 位址與時間。如果旗標爲 3，表示有提供 IP 位址，每個路由器必須檢查提供的 IP 位址與封包進入自己這端的 IP，如果一樣，路由器將原提供的 IP 位址換成自己送出端的 IP 位址並填入時間。

五， 檢查碼(Checksum)

TCP/IP 協定組中大部分的協定所使用的錯誤偵測方式稱爲 Checksum。Checksum 是爲了封包在傳輸的過程中可能遭受破壞所使用的一種保護措施。Checksum 是根據封包所加入的一些訊息。送出者計算 Checksum，然後將之與封包一起送出。接收者針對整個封包、包括 Checksum 本身，以相同的計算程序，如過其結果正確，則接受封包，否則便拒絕。

1. 傳送端 Checksum 之計算

傳送者將封包分成若干個 N 位元的段落(N 通常是 16)，之後以 1 補數算數，將這些段落加起來，而其和依然是 N 位元長，然後再求這個和的補數(所有 0 變成 1，1 變成 0)，其結果就是 Checksum。步驟：

A. 將封包分成 K 個段落，每個段落有 N 個位元。
B. 將所有個段落以 1 補數相加。
C. 再求步驟(B)的結果的互補數及爲 Checksum。

2. 接收端 Checksum 之計算

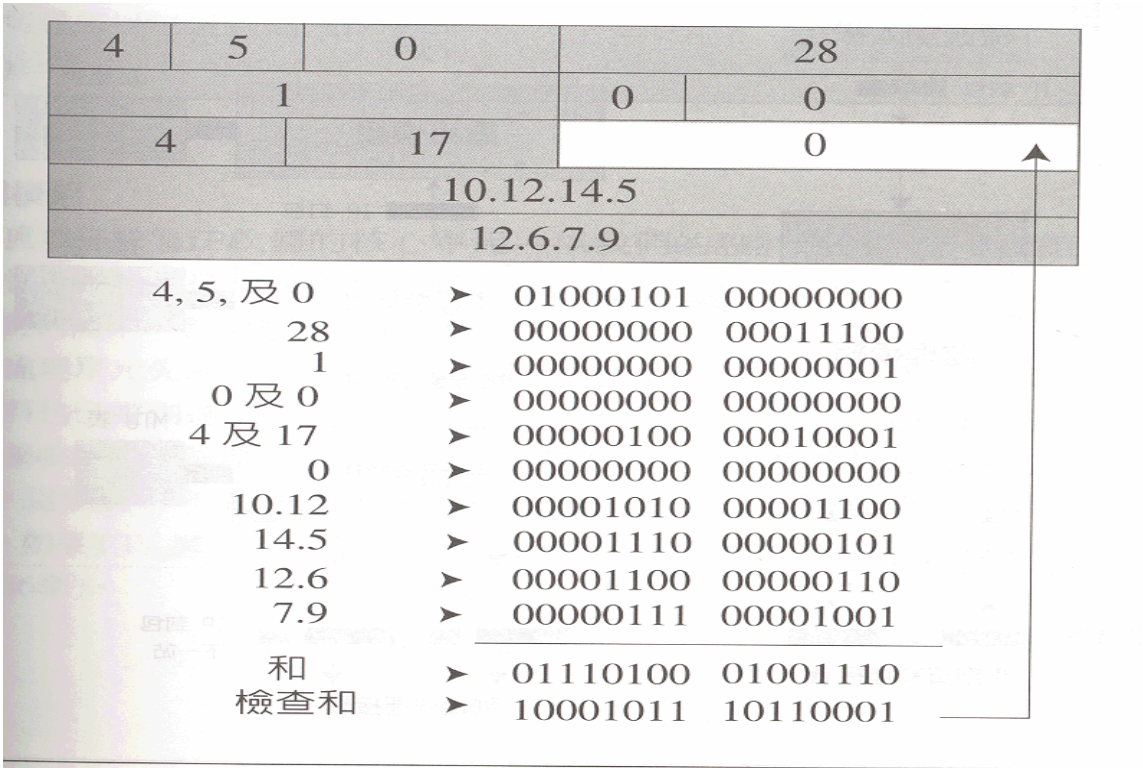
接收者收到封包分成 K 個段落，之後以 1 補數算數接這些值加起來，然後再求這個和的補數，若結果爲 0，則接受該封包，否則拒絕。步驟：

A. 經封包分成 K 個段落，每個段落有 N 個位元。
B. 將所有各段以 1 補數相加。
C. 再求步驟(B)結果的互補數值。
D. 如果最後結果爲 0，則接受該封包，否則拒絕該封包。

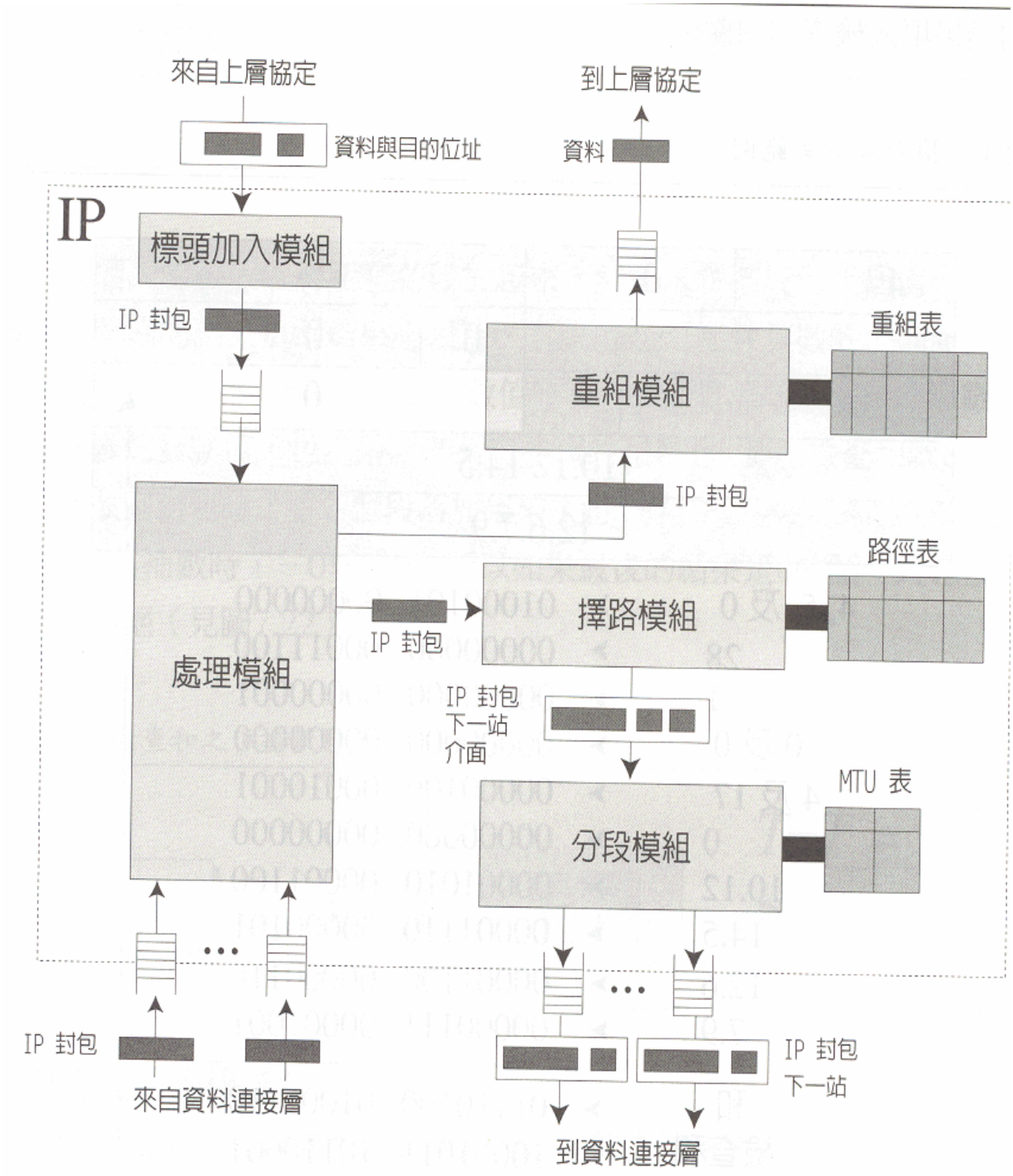
3. IP 封包使用的 Checksum

IP 封包使用的 Checksum 依照上述的方法來實現，一開始 Checksum 的數值訂爲 0，然後將整個封包的標頭分成數個 16 位元的段落，然後加在一起，求其結果之互補值，將結果放入 Checksum 欄位。

IP 封包中的 Checksum 只包含標頭部分，並不包含資料。這有兩個原因，第一，所有將資料放到 IP 封包的較高層協定都有自己的 Checksum 包含到資料部分，所以 IP Checksum 不用去檢查封包的資料。第二，因爲 IP 的標頭在經過每個路由器時都會改變，但是資料部分並不會改變，所以 Checksum 只會包含會改變的部分。如果要把資料部分也算進來，那麼每個路由器必須以整個封包來計算 Checksum，那意味者路由器要花更多的處理時間。下圖中說明 IP Checksum 的計算過程。



六， IP 設計


來自上層協定
資料與目的位址
到上層協定
資料
IP
標頭加入模組
IP 封包
重組表
重組模組
IP 封包
路徑表
處理模組
IP 封包
擇路模組
IP 封包
下一站
介面
MTU 表
分段模組
IP 封包
來自資料連接層
IP 封包
下一站
到資料連接層