

申請日期	90.6.26
案 號	90115366
類 別	H04L 12/54

A4
C4

(以上各欄由本局填註)

576048

發 明 專 利 說 明 書

一、發明 名稱	中 文	在一網路交換埠上以線率來識別資料封包之裝置及方法
	英 文	APPARATUS AND METHOD FOR IDENTIFYING DATA PACKET AT WIRE RATE ON A NETWORK SWITCH PORT
二、發明 創作人	姓 名	1. 高 帕 S. 可瑞席那 GOPAL S. KRISHNA 2. 周家輝 PETER KA-FAI CHOW 3. 曾世杰 SHR-JIE TZENG 4. 索姆那 維斯萬那史 SOMNATH VISWANATH
	國 籍	1. 印度 2. 4. 美國 3. 中華民國
三、申請人	住、居所	1. 美國·加州 95148·聖喬西市·亞當斯伍得道 3260 號 3260 Adamswood Drive San Jose, California 95148 U.S.A. 2. 美國·加州 95131·聖喬西市·馬克斯威爾徑 1416 號 1416 Maxwell Way San Jose, California 95131 U.S.A. 3. 美國·加州 94555·法蒙特市·米摩沙泰瑞·34263 號 34263 Mimosa Terrace Fremont, California 94555 U.S.A. 4. 美國·加州 95124·聖喬西市·馬達道 3538 號 3538 Manda Drive San Jose, California 95124 U.S.A.
	姓 名 (名稱)	高級微裝置公司 ADVANCED MICRO DEVICES, INC.
三、申請人	國 籍	美國
	住、居所 (事務所)	美國·加州 94088-3453·桑尼威·第 1AMD 區·M/S 68· 郵政信箱 3453 號 One AMD Place, M/S 68, P. O. Box 3453, Sunnyvale, CA 94088-3453 U.S.A.
三、申請人	代 表 人 姓 名	理查·J·諾迪 Richard J. Roddy

經濟部智慧財產局員工消費合作社印製

裝

訂

線

(由本局填寫)

承辦人代碼：
大類：
I P C 分類：

A6
B6

本案已向：

美

國(地區)申請專利，申請日期：

案號：

，☒有 ☐無主張優先權

2000 年 8 月 14 日 09/637,015 (主張優先權)

有關微生物已寄存於：

，寄存日期：

，寄存號碼：

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

經濟部智慧財產局員工消費合作社印製

五、發明說明(1)

[發明背景]

[發明範疇]

本發明係關於：在一經過配組而用來供資料封包於子網路間做交換的無阻塞(nonblocking)網路交換器中，層2及層3資料封包之交換。

[背景技藝]

區域網路(LAN)使用網路纜線或其他媒介來鏈接網路上的網站。每一區域網路架構，在其每一網路節點上皆使用一媒體存取控制(MAC)賦能網路介面裝置，以存取網路媒體。

乙太網路協定 IEEE 802.3(美國電機電子工程師學會 802.3 標準)已經演進為指定半雙工(half-duplex)媒體存取機制及全雙工(full-duplex)媒體存取機制來傳輸資料封包。全雙工媒體存取機制係在兩個網路元件間，例如在一網路節點與一交換集線器間，提供雙向的點對點通訊鏈接。

交換的區域網路現正愈益要求更為高速的連接性、更具彈性的交換性能，以及容納更複雜網路架構的能力。例如，共有讓與的 U.S. Patent No. 5,953,335 (美國專利 5,953,335 號)揭露一種網路交換器，係經配組而在不同網路節點間用以交換層2型態的乙太網路(IEEE 802.3 標準)資料封包；一個接收的資料封包可依 IEEE 802.1q 協定而包含一 VLAN(虛擬區域網路 [virtual LAN])標籤框，該協定係指定另一子網路(經由一路由器)或既定的網站群。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(2)

既然交換是發生在層 2 等級，則典型地需要一路由器，在子網路間傳送資料封包。

為做層 3(如，網際網路協定)處理，加強網路交換器的交換性能之做法有嚴重的缺點，此乃因現行的層 2 交換器宜經配組而以無阻塞模式供作業之用，其中資料封包能以同於接收時的速率自交換器輸出。為確保較高速的交換能提供層 2 交換及層 3 交換能力，供較快速的網路(諸如每秒 100 百萬位元 [100 Mbps])之用，則需有較新的設計。

然而，諸如此等設計要求，冒有網路交換器無阻塞特徵之損失風險，此係因網路交換器之交換加工漸難能以線率(即，網路資料速率)來施行層 3 處理。

習知的路由器(router)及高級的交換器(high-end switches)係使用以 CPU 密集方案為主的集中式軟體來檢查資料封包，以供分類之用；並決定其類別或優先性，供封包交換/選路之用。第 1 圖顯示，一資料封包抵達一習知的網路交換埠(T1)直到其在一輸出埠上離開網路交換器(T10)，該資料封包的時間序列。既然封包必須發送至一處理器(於 T3)，再由該處理器做分類(於 T5-T6)，則其潛時(latency)(第一個進入位元至第一個輸出位元)高，且可以式表為 $S1+S2+S3+S4$ 。會有高潛時，係因封包甚至在完全被接收後(T2 之後)，猶未開始被處理。

[發明概述]

邇來有一種配備之需要，該配備其使一網路交換器能提供層 2 交換及層 3 交換能力，供 100 Mbps 及十億位元

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(3)

鏈接(gigabit links)之用，而不阻塞資料封包。

也需要一種配備，其使一網路交換器能在內有最小緩衝下(否則會影響交換的資料封包之潛時)，提供層2交換及層3交換能力。

也需要一種配備，其使一網路交換器可容易程式化，以辨別不同型態之層3資料封包，而獲致服務品質(QoS)。

也需要一種配備，其使一網路交換埠能立即對一進入的資料封包估值，並決定層3或更高的協定，以有充份時間提供關聯的交換構件(switch fabric)，而依所偵測的協定來處理該進入的資料封包。

本發明達成此等及其他需求，係在一網路交換埠中，於資料封包之最後一個位元一經接收即評估整個資料封包，基於此而生成一個框標籤。

本發明一方面係提供一種評估網路交換埠上之一進入資料封包的方法。該方法包含：儲存複數個經配組而用以識別各別資料框之模板(template)，而每一模板有至少一個最小項(min term)，該最小項係經配組而用以比較一對應的既定值與該進入資料封包之一對應的選定位元組。在該網路交換埠接收該選定的位元組之時，立即將該等對應於選定位元組之最小項同時與該選定位元組比較。基於該等最小項與該網路交換埠所接收的整個資料封包之資料位元組所做之比較，生成一識別該進入的資料封包之比較結果。當資料封包之最後一個位元在該網路交換埠一經接收，即基於比較結果而生成一個框標籤。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(4)

本發明另外的優點及新穎特徵一部份陳述於以下說明中；一部份則為熟習此項技藝者查酌下文即可明白，或者可藉實作本發明而習得。本發明之優點可藉助於工具或工具組合，特別是附錄之申請專利範圍所指出者，來實現及達成。

[圖式簡單說明]

參考附圖，其中指定為相同參考號數的元件在全文中代表相似元件，又其中：

第1圖顯示一資料封包之潛時，其係由封包抵達一習知的網路交換埠直到封包離開該埠的時間所定義。

第2圖為一含多個網路交換器之封包交換網路的方塊圖，係用來依據本發明之一實施例而供資料封包於各別子網路間做交換。

第3圖闡示一習知的層2乙太網路型態之資料封包，其係運載一個層3網際網路協定(IP)封包。

第4圖為一流程圖，闡示一種評估一IP封包的習知(先前技藝)方法。

第5圖為一方塊圖，依本發明之一實施例，闡示圖2網路的網路交換埠，其包含網路交換埠過濾器(filter)。

第6A圖及第6B圖闡示：藉第5圖之最小項生成器，而同時處理一方程式的四個模板。

第7圖更詳細地闡示：藉第5圖之最小項生成器，而同時處理最小項。

第8圖係依本發明之一實施例，闡示第5圖之最小項

(請先閱讀背面之注意事項再填寫本頁)

裝
訂

五、發明說明(5)

記憶體中之一最小項表項(table entry)。

第9圖顯示本發明一實施例之一資料封包的潛時，係由封包抵達一網路交換埠直到封包離開該埠的時間所定義。

[元件符號說明]

10	分封交換網路	12、12a、12b、12c	交換器
14	網站	16	乙太網路鏈接
18a、18b、18c	子網路	20	交換埠
22	媒體存取控制(MAC)	24	封包分類器模組
25	交換(器)構件	26	中央處理單元(CPU)
28	緩衝記憶體	29	標頭修正器
30、32	封包	34	框起始定界符
36	乙太網路頭標		
38	循環冗餘核對(CRC)或框核對序列(FCS)欄位		
62、62a、62b、62c、62d	模板	64	起始位置
70	最小項記憶體	72	框識別器
74	最小項控制器	76	最小項生成器
78	方程式核心	80	埠
81	標籤優先性解析裝置	82	位置轉換器
90	表項	94	遮罩欄位
96	期望值欄位	98	算符欄位
100	分支部份	102、106	回應部份
110	方程式識別符		

[實行本發明之最佳模式]

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(6)

第 2 圖為一方塊圖，闡示一封包交換網路 10，諸如乙太網路(IEEE 802.3)之類的網路。該封包交換網路包含積體(即，單一晶片)多埠交換器 12，而使網站 14 間能有資料封包通訊。每一網站 14(例如一客戶工作站)皆典型地經配組而以 10 Mbps 或 100 Mbps(依 IEEE 802.3 協定)發送及接收資料封包。每一積體多埠交換器 12 皆藉由十億位元乙太網路鏈接 16 而互連，使子網路 18a、18b 與 18c 間能傳送資料封包。由此，每一子網路皆包含一交換器 12，及一關聯的網站 14 群。

每一交換器 12 皆含一交換埠 20，而交換埠 20 含一媒體存取控制(MAC)模組 22 及一封包分類器模組 24。MAC 模組 22 依 IEEE 802.3u 協定，跨越 10/100 Mbps 實體層(PHY [physical layer])資料收發器(未示出)，而發射並接收資料封包至關聯的網站 14。每一交換器 12 也含一交換構件 25，而構件 25 係經配組，成為用於接收的資料封包之框轉接決策(frame forwarding decision)。特別是，交換構件 25 係經配組，成依據來源位址、目的位址及乙太網路(IEEE 802.3)標頭(header)內之 VLAN 資訊，供層 2 交換決策之用；交換構件 25 也經配組成為依據乙太網路封包內之一 IP 資料封包之評估，供選擇性的層 3 交換決策之用。

如第 2 圖所示，每一交換器 12 皆有一關聯的主機 CPU 26 及一緩衝記憶體 28(例如一 SSRAM(同步靜態隨機存取記憶體 [synchronous SRAM]))。主機 CPU 26 控制對應的交換器 12 之全部作業，包括交換構件 25 之程式化。緩衝

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(7)

記憶體 28 為對應的交換器 12 所使用，以儲存資料框；而交換構件 25 則為接收的資料封包處理轉接決策。

如上述，交換構件 25 係經配組成為用以施行層 2 交換決策及層 3 交換決策。子網路 18a 內的終端網站 14 若欲發送一電子郵件訊息至子網路 18b、18c 或兩者內的選定網站，則層 3 交換決策之可用性是格外有效的；若僅有層 2 交換決策可用，則交換器 12a 之交換構件 25 之發送電子郵件訊息至交換器 12b 及 12c，將無特定的目的位址資訊，致使交換器 12b 及 12c 在所有的埠間四處選路。若非如此，交換器 12a 之交換構件 25 則須發送電子郵件至一路由器（未示出）；路由器會引入額外的延遲。以交換構件 25 做層 3 交換決策，則交換構件 25 能聰明地決定如何處置一個封包（包括先進轉接決策），以及應否視一封包為一供潛時敏感應用之高優先性封包。以交換構件 25 做層 3 交換決策，也使交換器 12a 之主機 CPU 26 能遠端程式化另一交換器（例如交換器 12b），其係藉發送含 IP 位址的訊息，而該 IP 位址對應於交換器 12b 之 IP 位址；當偵測到一定址於交換器 12b 的訊息，交換器 12b 在回應下能轉接該訊息至對應的主機 CPU 26，供交換器 12b 程式化之用。

第 3 圖闡示一乙太網路(IEEE 802.3)封包 30，而封包 30 運載一 IP 封包 32 做為運荷資料。特定地說，乙太網路封包 30 包含一框起始定界符(SFD[start frame delimiter]) 34、一乙太網路標頭 36、IP 封包 32，及一循環冗餘核對(CRC[cyclic redundancy check])或框核對序列(FCS[frame

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(8)

check sequence])欄位 38。由此，經配組而供層 3 交換決策用之交換器構件 25 必須能在接收的乙太網路框 30 之內快速處理 IP 封包 32，以免該框在交換器內阻塞。

第 4 圖為一流程圖，闡示可對一進入資料封包施行之層 3 處理型態範例。第 4 圖之流程圖係習知地實作於一軟體中，而將核對進入資料封包是否為一超文字傳輸協定(HTTP)封包(於步驟 50)、一 SNMP(簡單網路管理協定)封包(於步驟 52)，或一高優先性封包(於步驟 54)。然後指定妥當的標籤，來識別該封包(於步驟 56、57、58 或 60)。

然而，從實務立場來說，第 4 圖之配置並不能在硬體中實作以提供無阻塞交換而供 100 Mbps 或十億位元網路之用。特別是，第 4 圖中決策過程的順序本質將對進入的資料封包造成過份的潛時。

依據所揭露的實施例，第 2 圖之封包分類器模組係經配組，而在進入的資料流與識別該進入的資料流資料格式的模板之間，用來做多重而同時之比較。特定地說，主機處理器 26 之使用者將規定策略，以定義含某 IP 協定之資料封包應如何為交換構件 25 所處置。此等策略係將一組框轉接決策(供每一對應的 IP 協定型態所用)載入交換構件 25，藉而得以實作。由此，交換構件 25 可包含一組框轉接指令供一 HTTP 封包之用、另一組框轉接指令供一 SNMP 封包之用，以及另一組框轉接指令供一高優先性封包之用(如，視頻或語音等)。

第 5 圖為一方塊圖，闡示依據本發明一實施例之封包

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(9)

分類器模組 24。如第 5 圖所示，網路交換埠 20 包含一 MAC 22、一接收 FIFO(先入先出)緩衝器 27、一標頭修飾器(modifier)29，及封包分類器模組 24。封包分類器模組 24 也稱做網路交換埠過濾器，而係經配組，用以識別(即，估值)在網路交換埠 20 之進入的資料封包，並供應一標籤給交換構件 25，而該標籤係基於正被接收的資料封包之型態來規定對該資料封包的動作。特定地說(且如第 9 圖所示)，封包分類器模組 24 係將進入的資料封包與複數個模板同時比較，經配組而當資料於交換埠 20 被接收時立即識別各別的資料格式。封包分類器模組 24 比較進入的資料封包與該複數個模板，基於此而識別一將執行的方程式，而該方程式規定將要供應給交換構件 25 的標籤。如第 9 圖所示，既然封包在交換埠 20 一經接收即被分類，則不論封包是否驗效，其潛時 $S1+S2+S3$ 比起參考圖 1 之上述潛時，有所縮減。

特定地說，封包分類器模組 24 生成一比較結果，係從複數個模板偵側出至少一個匹配的模板，以識別進入的資料封包。然後封包分類器模組 24 識別哪一個方程式包含該匹配的模板，並生成該方程式所規定的標籤。

第 6A 圖及第 6B 圖闡示，藉封包分類器模組 24 而對一方程式的兩個模板同時所做之處理。第 6A 圖闡示，藉封包分類器模組 24 所做之方程式：

$$Eq1=M1*M2*M3*M4*(M5+M6+M7+M8)$$

的邏輯評估。

(請先閱讀背面之注意事項再填寫本頁)

五、發明說明 (10)

第 6B 圖闡示，方程式 Eq1 將如何實際地儲存在最小項記憶體 70 中。方程式 Eq1 含四個模板 62a、62b、62c 及 62d：模板 62a 含最小項 M1、M2、M3、M4 及 M5；模板 62b 含最小項 M1、M2、M3、M4 及 M7；模板 62c 含最小項 M1、M2、M3、M4 及 M6；而模板 62d 含最小項 M1、M2、M3、M4 及 M8。每一模板 62 皆對應一特定的 IP 資料格式，而該格式可基於 IP 資料封包 32 之標頭來辨識。例如，模板 62a 及 62c 可經配組而用以識別一 HTTP 封包；模板 62b 及 62d 可經配組而用以識別一 SNMP 封包。特定地說，一封包會被識別為 HTTP 封包，如果：其為 IPv4 格式、IP 中的存活時間(time to live)欄位大於一、IP 標頭中的協定欄位為 TCP、標頭核對和(checksum)為正確、來源 TCP 埠為 80 或目的 TCP 埠為 80。一封包會被識別為 SNMP 封包，如果：其為 IPv4 格式、IP 中的存活時間欄位大於一、IP 標頭中的協定欄位為 TCP、標頭核對和為正確、來源 TCP 埠為 25 或目的 TCP 埠為 25。

由此，可建立以下最小項來代表所有的上述判據：

M1=封包為 IPv4 格式

M2= IP 中的存活時間欄位大於一

M3= IP 標頭中的協定欄位為 TCP

M4=標頭核對和為正確

M5=來源 TCP 埠為 80

M6=目的 TCP 埠為 80

M7=來源 TCP 埠為 25

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明 (11)

M8=目的 TCP 埠為 25

由此，模板 62a 及 62c 識別 HTTP 封包，而模板 62b 及 62d 識別 SNMP 封包。如此，方程式一(Eq1)規定，如果模板 62a、62b、62c 及 62d 任一為真，則應輸出一特定結果(如，有一規定值之標籤)至交換構件 25。

甚且，最小項 M1...M8 係以一既定次序安排在關聯的模板 62a 及/或 62b 內，而該次序係對應進入資料流中一資料位元組之相對位置。如第 7 圖所示，最小項 M1 係經配組，而用與 IP 封包 32 之第一位元組(B1)比較；最小項 M2 係經配組，而用與 IP 封包 32 之繼 B1 後的位元組(B2)比較；最小項 M3 係經配組，而用與 IP 封包 32 之繼 B2 後的位元組(B3)比較，等。由此，基於進入的資料流中一資料位元組之相對位置，以此為使用的模板 62 之最小項次序，則能在該進入資料流與該等最小項間做比較。由此，進入資料流能與多個模板比較，不但決定了進入資料封包的資料格式，也決定交換構件 25 須施行何種動作。

如第 5 圖所示，封包分類器模組 24(也稱做網路交換埠過濾器)包含一最小項記憶體 70，用以儲存最小項之值(如，M1、M2，等)，如第 8 圖所示。以下說明之：封包分類器模組 24 也含一框識別器(identifier)72，其係經配組而用以識別正被接收的層 2 框之型態；尤其，識別接收中的層 2 框之型態(如，乙太網路、IEEE802 至 3，等)，即能識別層 2 封包 30 內的 IP 封包 32 起始位置 64。封包分類器模組 24 也含一最小項控制器 74、一最小項生成器 76、一

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明 (12)

方程式核心(core)78，及一估值結果記憶體 80。使用一處理器介面模組(pi_mod)，從主機 CPU 26 傳送生成的最小項至最小項記憶體 70。

最小項控制器 74 係經配組，對應於 IP 框 32 之一選定的位元組而從最小項記憶體 70 提取最小項。最小項控制器 74 也含一位置轉換器 82，位置轉換器 82 係經配組而當接收一來自框識別器 72 之框型態(frm_type)訊號(規定層 2 框之型態)時，在回應下規定起始點 64 之實際位元組位置(byte_location)。由此，最小項控制器 74 偵測 IP 封包的開端，在回應下提取將與 IP 封包之第一位元組(B1)比較的所有最小項，例如第 7 圖中用於方程式 Eq1、Eq2 及 Eq3 之最小項 M1、M9 及 M14。然後，最小項控制器 74 轉接最小項之值(M_STRU INFO)給最小項生成器 76 及方程式核心 78。

最小項生成器 76 在最小項控制器所提取之最小項與進入的資料流之選定的位元組之間，施行實際的最小項比較。例如在第 7 圖中，最小項生成器 76 將進入的資料位元組 B1 同時與最小項 M1、M9 及 M14 比較，並提供最小項比較結果(mt_result)給方程式核心 78。在次一比較循環期間，最小項生成器 76 將進入的資料位元組 B2 同時與最小項 M2、M10 及 M15 比較。依據所揭露的實施例，最小項生成器係經配組而用來將進入的資料流同時與多達八個最小項比較。

方程式核心 78 係經配組，基於接收自最小項生成器

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明 (13)

76 之最小項比較結果而生成一個框標籤，相對於有關的模板 62。舉例來說，方程式核心 78 在最小項生成器供應出結果時，對此最小結果順序地估值，藉而估算方程式 1 之值(闡示於第 6 圖)。舉例來說，若對於最小項 M1、M2、M3、M4、M5 及 M6 每一項所做的比較導致條件為真(對下述第 8 圖而言)，則終結條件在方程式 1 中為匹配，而致使方程式核心 78 生成一對應於方程式 1 規定條件的標籤。該框標籤在必須施行的動作之外，也識別進入的資料封包之本質。

由此，封包抵達交換埠 20 時，封包分類器模組 24 立即以一可程式化的量子(位元組字 [byte word]、雙字 [double word]，等)邊界(quantum boundary)，來檢查進入的資料是否匹配模板中的適當欄位。如果發現是匹配的，則封包分類器模組 24 持續追蹤哪些模板已匹配，並依抵達次序來繼續檢查其餘進入的封包欄位。循此方式，不再匹配該封包的模板即遭廢除，不做進一步的比較。在封包終端，或在一完全匹配的模板終端，封包分類器模組 24 之方程式核心 78 以關聯於該等匹配模板的標籤，來標記該封包。或許某些情況中會有多於一個模板匹配該封包；在諸如此種情形中，標籤優先性解析裝置 81 施行一基於優先性的決定(由使用者所定義)，而在框終端(或附近)選取單一個框標籤值並發送至交換構件 25。

第 8 圖闡示最小項記憶體 70 中的最小項結構之資料格式。依據所揭露的實施例，最小項係對應於將做比較的資

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明 (14)

料位元組之相對位置，以此次序而儲存在最小項記憶體 70 中。由此，所有將與第一資料位元組比較的最小項係一起儲存在最小項記憶體的第一部份中，繼之則是所有將與第二資料位元組比較的最小項，等等。

或者，最小項之儲存次序除基於將做比較的資料位元組之相對位置外，也可基於 IP 標頭內的相關資訊。由此，最小項之儲存次序可供資料位元組序列比較之用，而提供來源 IP 位址、目的 IP 位址及來源和來源/目的埠等條件；在此情形中，在 IP 框開端之無關的資料位元組將不會在最小項記憶體 70 開端儲存其關聯的最小項，最小項記憶體 70 的效率進而獲得改良。

每一表項 90 包含一最小項部份及一估值部份。最小項部份含一遮罩欄位(遮罩[MASK])94、一期望資料欄位(期望資料[EXP_DATA])96，及一算符欄位(算符[OPERATOR])98。基於表項 90 在最小項記憶體 70 中的位置，最小項控制器 74 能決定 IP 封包 32 的哪一個位元組須與對應的最小項比較(相對於 IP 封包的開端 64)。遮罩欄位 94 為最小項生成器 76 施行比較所用的遮罩，若該遮罩有一個位元設定為 1，則比較其值；而若遮罩值在欄位中有零值，則毋庸比較。期望資料欄位 96 規定將與 IP 封包 32 之有關資料位元組比較的期望資料。算符欄位 98 規定最小項生成器所將施行的比較之型態，例如：小於、小或等於、大於、大於或等於，及不等於。

估值部份則含一支部份 100、一回應部份(RINP 1)

五、發明說明 (15)

102 用於最小項部份比較為真之情形、一第二回應部份 (RINP 0)106 用於最小項部份比較為偽之情形，及一方程式識別符 110。分支部份 100 規定方程式中 OR 項之次序；例如第 6 圖及第 7 圖所示之最小項 M1，其分支部份將設定為 0000 1111，標示著：方程式識別符 110 中所規定的方程式最先四個分支應包含對應的最小項。以八個位元用於分支部份，係假設任何給定方程式中最多有八個分支。

回應部份 102 在最小項部份相對於比較的位元組之估值為真時，規定所將施行的作業。特別是，若方程式結果在最小項結果為真時被決定，則完結位元 (FIN) 設定為一；若估值過程在最小項結果為真時應返回初始狀態 (init)，則「回到初始 (BINIT)」設定為一。例如在最小項 M1 之情形中，最小項結果為真時既須做其他比較，則 RINP 1 之 FIN 位元及 BINIT 位元係設定為零。在最小項 M5、M6、M7 及 M8 之情形中，為「真」之比較結果既會結束估值，則 RINP 1 之 FIN 位元係設定為一，如第 6 圖所示。

回應部份 106 在最小項部份相對於比較的位元組之估值為偽時，規定將施行的作業。特別是，若方程式結果在最小項結果為偽時被決定，則完結位元 (FIN) 設定為一；若估值過程在最小項結果為偽時應返回初始狀態 (init)，則「回到初始 (BINIT)」設定為一。例如在最小項 M1 之情形中，FIN 位元係設定為零而 RINP1 之 BINIT 位元係設定為一，以致方程式在最小項 M1 結果為偽時將返回 init 狀態，如第 6 圖所示。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

五、發明說明(16)

方程式識別符 110 識別最小項所對應的方程式(或者，若方和式中僅有一個模板，則對應該模板)。

由此，方程式核心 78 決定任何規定的方程式是否有一模板 62 匹配進入的資料流。對帶有多個模板 62 之進入的資料流做多重同時比較，基於此則方程式核心 78 能識別一匹配方程式；並對於交換構件 25 上的框轉接決策，生成對應於匹配的方程式之適當標籤。必要時，核心 78 在傳送資料給交換器之前，也輸出一指令至標頭修正器 29，以修正層 2 標頭、層 3 標頭，或兩者。

依據所揭露的實施例，一網路交換埠包含一過濾器，而該過濾器能在資料封包之進入的資料流與經配組而供識別一對應協定用之多個模板之間，施行多重同時比較。既然封包分類器模組 24 能處理 IP 框 32 的任何位元組，則封包分類器模組 24 能直譯 IP 封包 32 中所有的標頭資訊，從層 3 達層 7 協定。甚且，該等多重同時比較使網路交換器 12 能施行層 3 交換，供 100Mbps 及十億位元網路之用，而在網路交換器中無阻塞。最後，該等多重同時比較以資料接收之次序，而能施行即時(real time)比較，此與可程式化邏輯陣列(PLAs)之類的替代方案相反；此等替代方案要求在開始處理之前接收整個標頭。

由此，依據所揭露的實施例，封包分類解析度及關聯的交換控制屬性在封包之最後一個位元一抵達交換埠時，可立即發生，從而縮減全部的交換潛時，如此，封包分類器可經配組，以線速(wire speed)，基於任何協定來轉接封

(請先閱讀背面之注意事項再填寫本頁)

裝
訂

五、發明說明 (17)

包，而常駐在 ISO 模型之任一層。

本發明雖經目前所知最為實用的較佳實施例而有所說明，但應了解，本發明並不限於所揭露的實施例；反之，則係希望涵蓋附錄之申請專利範圍之精神及範圍內所含的各種不同的修正及同等的配備。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

四、中文發明摘要（發明之名稱： 在一網路交換埠上以線率來識別資料封包之裝置及方法 ）

一種網路交換器，經配組而用以施行在一乙太網路(IEEE 802.3)中的層 2 及層 3 交換，而不阻塞進入的資料封包，包含具有一過濾器(即，一封包分類器模組)之一網路交換埠，而該過濾器係經組構而在網路交換埠接收進入資料封包之時，立即對該資料封包評估。該過濾器在進入的資料封包資料流與經組構以識別各別資料協定的多個模板之間，施行同時之比較。每一模板由複數個最小項所組成，其中每一最小項規定在進入資料封包之一選定的位元組內的一個既定比較作業。該等模板可由使用者程式化，且儲存於一內部的最小項記憶體。甚且，該多重同時比較使該網路交換器能施行層 3 交換，供 100 Mbps 及十億位元網路之用，而在網路交換器中無阻塞。

英文發明摘要（發明之名稱：

APPARATUS AND METHOD FOR IDENTIFYING DATA PACKET AT WIRE RATE ON A NETWORK SWITCH PORT

A network switch, configured for performing layer 2 and layer 3 switching in an Ethernet (IEEE 802.3) network without blocking of incoming data packets, includes a network switch port having a filter (i.e., a packet classifier module) configured for evaluating an incoming data packet on an instantaneous basis, immediately upon receipt at the network switch port. The filter performs simultaneous comparisons between the incoming data stream of the data packet and multiple templates configured for identifying respective data protocols. Each template is composed of a plurality of min terms, wherein each min term specifies a prescribed comparison operation within a selected data byte of the incoming data packet. The templates may be programmed by a user and stored in an internal min term memory. Moreover, the multiple simultaneous comparisons enable the network switch to perform layer 3 switching for 100 Mbps and gigabit networks without blocking in the network switch.

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

六、申請專利範圍

1. 一種評估網路交換埠上之進入資料封包之方法，此方法包含：

儲存複數個經配組而用以識別各別資料框之模板，而每一模板有至少一個最小項，該最小項係經配組而用以比較一對應的既定值與該進入的資料封包之一對應的選定位元組；

在該網路交換埠接收該選定的位元組之時，立即將該等對應於選定位元組之最小項同時與該選定位元組比較；

基於該等最小項與該網路交換埠所接收的整個資料封包之資料位元組所做之比較，生成一識別該進入的資料封包之比較結果；以及

當資料封包之最後一個位元在該網路交換埠一經接收，即基於比較結果而生成一個框標籤。

2. 如申請專利範圍第 1 項之方法，其中該同時比較之步驟包含：

將對應於一第一資料位元組之最小項載入一最小項生成器；

平行比較該載於最小項生成器之最小項與該第一資料位元組；以及

輸出該載於最小項生成器之最小項的比較結果，至一評估核心。

3. 如申請專利範圍第 2 項之方法，其中該同時比較之步驟進一步包含：將對應於一第二資料位元組之最小項，接

六、申請專利範圍

續於該第一資料位元組之後而載入該最小項生成器。

4. 如申請專利範圍第 1 項之方法，進一步包含：輸出該框標籤至一交換構件，該交換構件係經配組而基於該對應的框標籤，用以交換該進入的資料封包。
5. 如申請專利範圍第 1 項之方法，其中該儲存步驟包含：將每一最小項在一記憶體中儲存為一表項，而每一表項基於對應的選定位元組之位置而在該記憶體中有一位置，而該表項含一最小項表式部份以規定該對應的既定值和一比較運算元，並含一模板識別符欄位以規定該等使用對應的最小項之模板。
6. 如申請專利範圍第 5 項之方法，其中該生成步驟包含：

暫時儲存該等最小項與該進入的資料封包之選定位元組的比較結果；

基於該等最小項之比較結果，而從複數個模板偵側出至少一個匹配的模板；以及

基於該偵側出的至少一個匹配的模板，而生成比較結果。
7. 如申請專利範圍第 6 項之方法，其進一步包含：當有多於一個模板匹配該進入的資料封包，將模板優先性解析為一個最終框標籤。
8. 如申請專利範圍第 5 項之方法，其中該第一資料位元組係對應於一既定格式之封包的資料位元組，而該同時比較之步驟包含：對相對於該既定格式之封包開端的選定資料位元組，做評估。

六、申請專利範圍

9. 如申請專利範圍第 8 項之方法，其中該既定格式為網際網路協定(IP)格式。

10. 如申請專利範圍第 6 項之方法，其中該基於偵側出的至少一個匹配模板而生成比較結果之步驟包含：

對於與該進入的資料封包比較的最小項之每一個，識別一對應的方程式，而每一方程式為一選定的模板群規定唯一的結果；以及

藉該含有偵側出的至少一個匹配模板之方程式，而生成該比較結果。

11. 一種經配組而用來評估進入資料封包評估之網路交換埠過濾器，其包含：

一最小項記憶體，經配組而用以儲存最小項之值，每一最小項之值係基於該進入資料封包之一對應的選定位元組位置而儲存，且有一表式部份以規定該對應的比較作業，以及一模板識別符欄位以規定該等使用對應的最小項之模板；

一最小項生成器，經配組而當該進入資料位元組一經接收，用來將進入的資料封包同時與對應於該接收的位元組之最小項比較，並生成各別的最小項比較結果；以及

一方程式核心，經配組而基於相對於該等模板之最小項比較結果，用來生成一識別該進入的資料封包之框標籤。

12. 如申請專利範圍第 11 項之過濾器，進一步包含一框識

(請先閱讀背面之注意事項再填寫本頁)

訂
線

六、申請專利範圍

別器而係經配組以識別層 2 封包之型態，而該進入資料封包之選定位元組係基於該經識別的層 2 封包型態而決定。

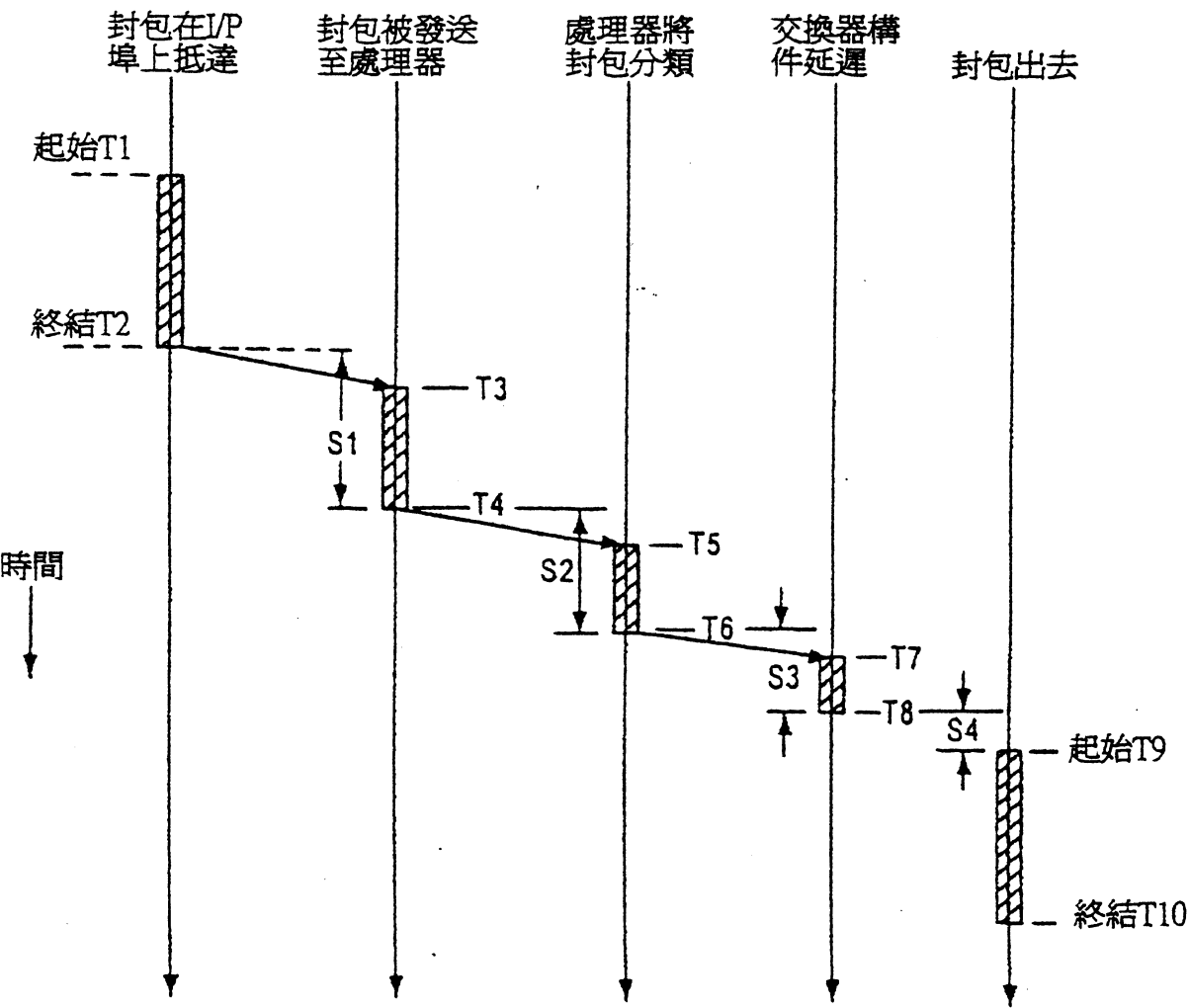
13. 如申請專利範圍第 12 項之過濾器，其中每一儲存的最小項之值的位置係相對於該層 2 封包內之一 IP 框的開端。
14. 如申請專利範圍第 13 項之過濾器，進一步包含一最小項控制器而係經配組，對應於該進入的資料封包內的 IP 框之一選定位元組而從最小項記憶體提取該等最小項。
15. 如申請專利範圍第 11 項之過濾器，其中該方程式核心以該進入的資料封包之線率，而在該進入資料封包終端之前生成該框標籤。
16. 如申請專利範圍第 11 項之過濾器，進一步包含一標籤優先性裝置而係經配組，當有多於一個模板匹配該進入資料封包，將模板優先性解析為一個最終框標籤值。

(請先閱讀背面之注意事項再填寫本頁)

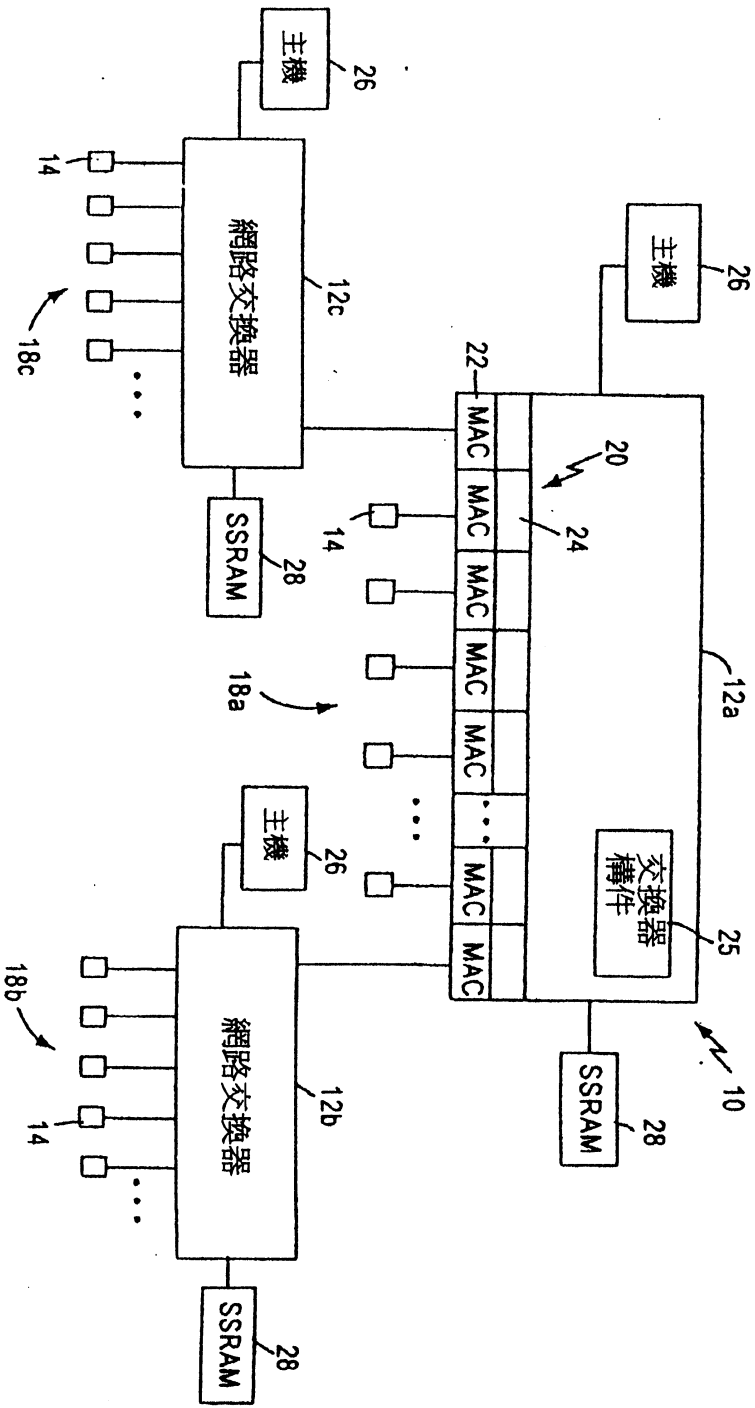
裝

訂

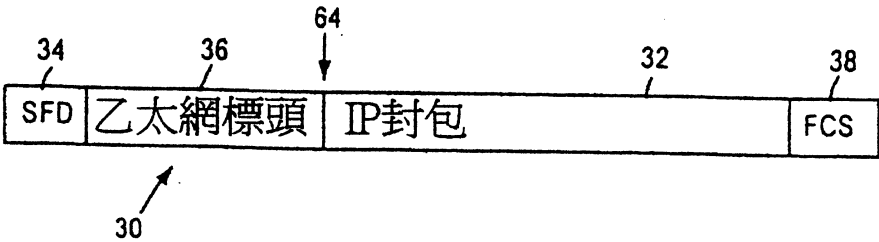
線



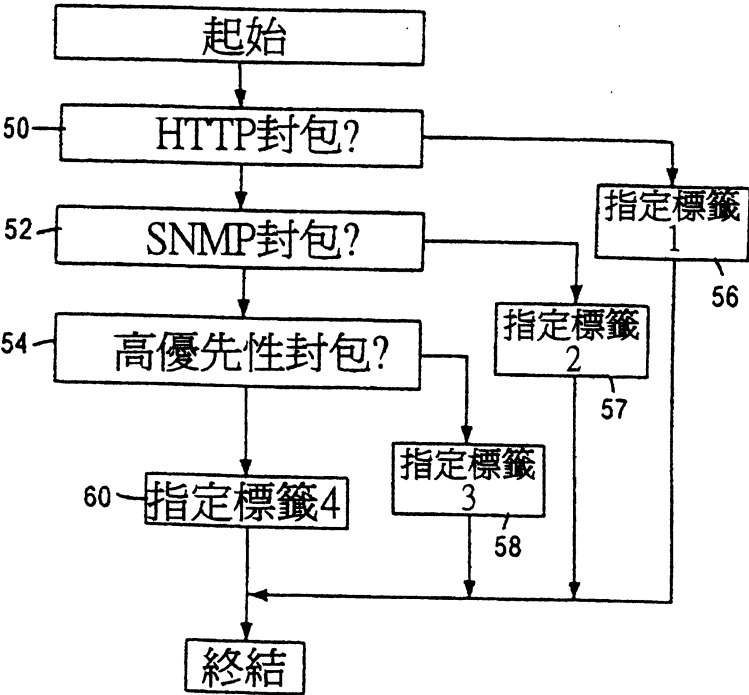
第 1 圖
(先前技藝)



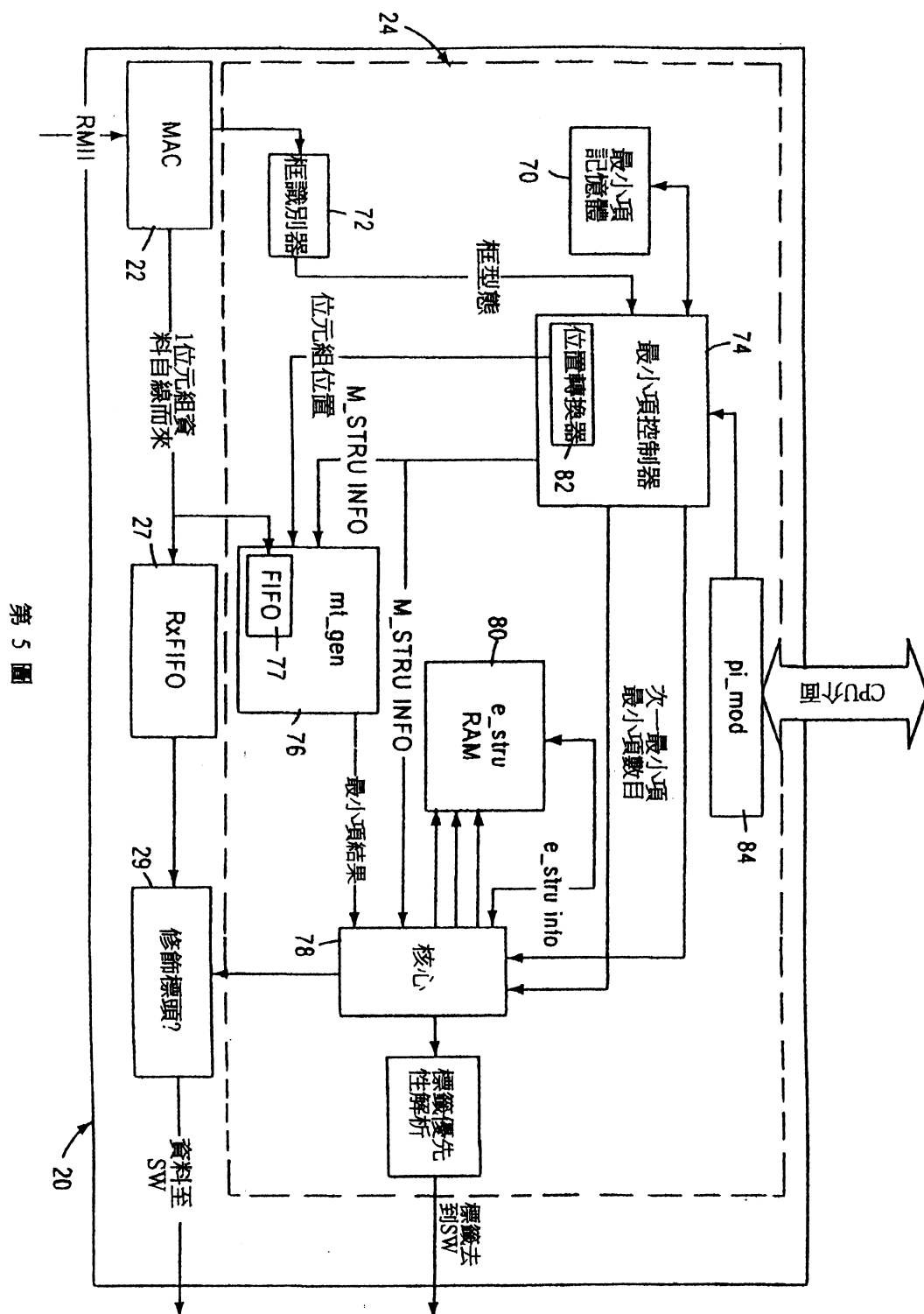
第 2 圖

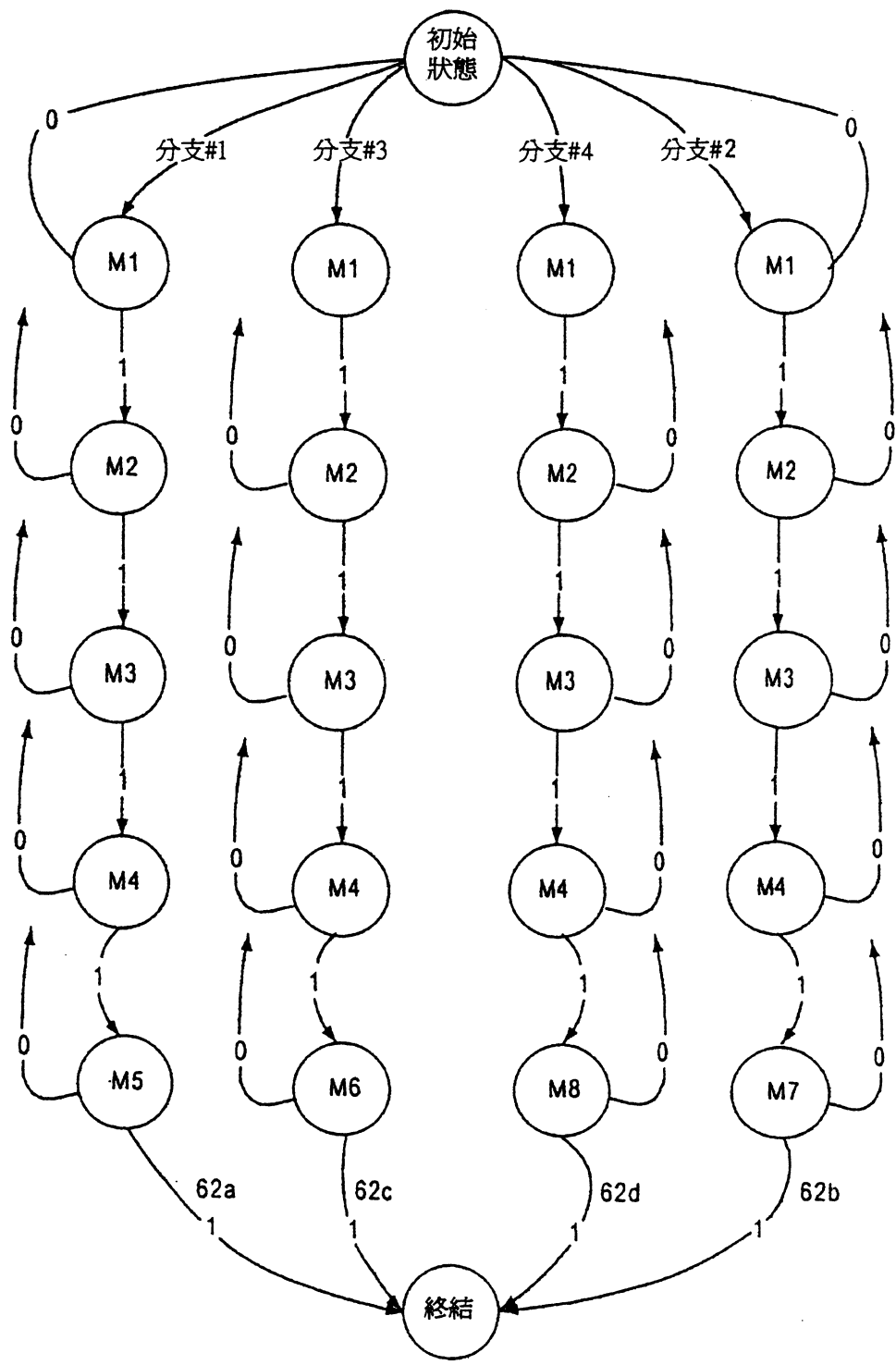


第 3 圖

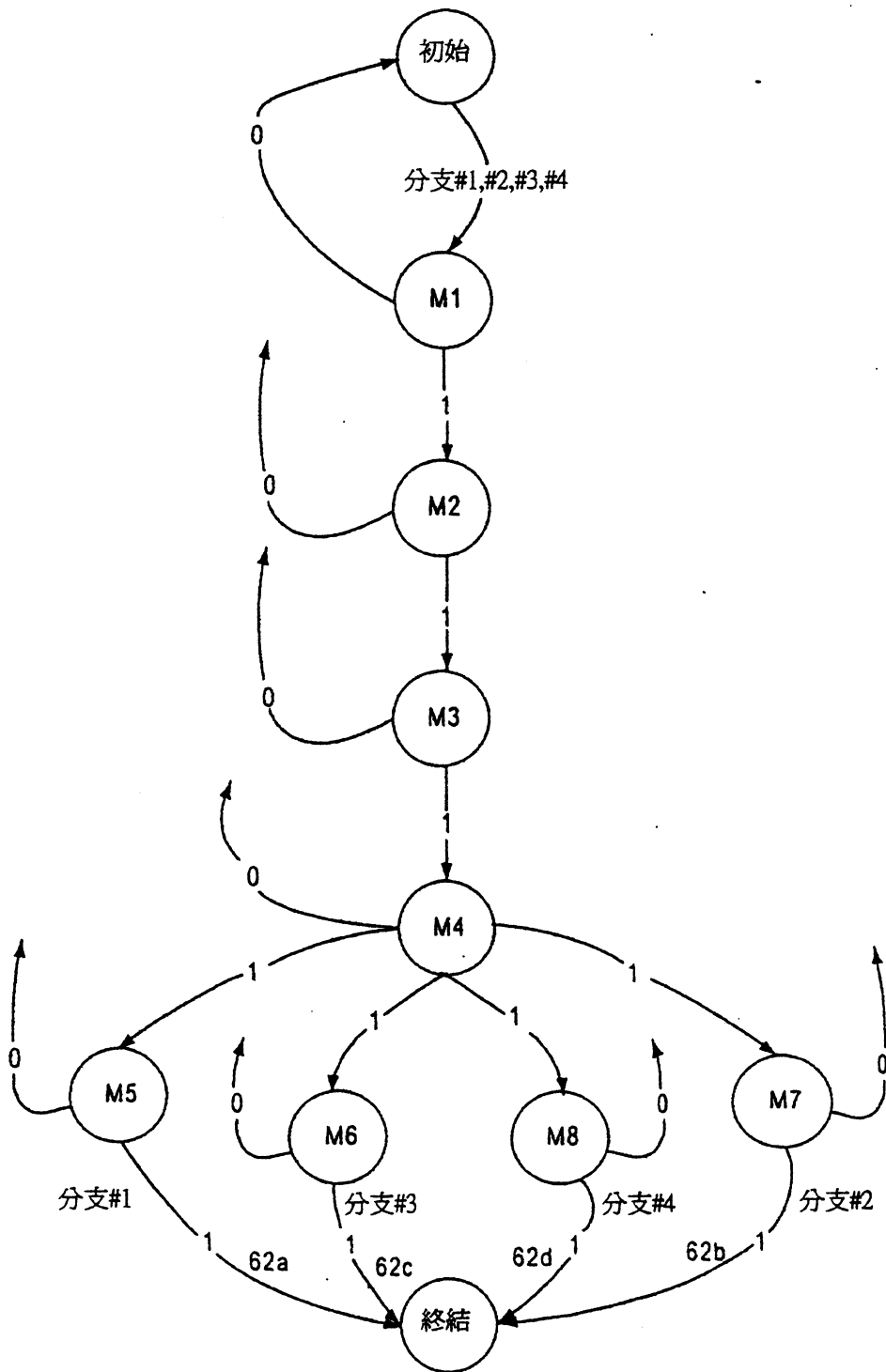


第 4 圖
(先前技藝)

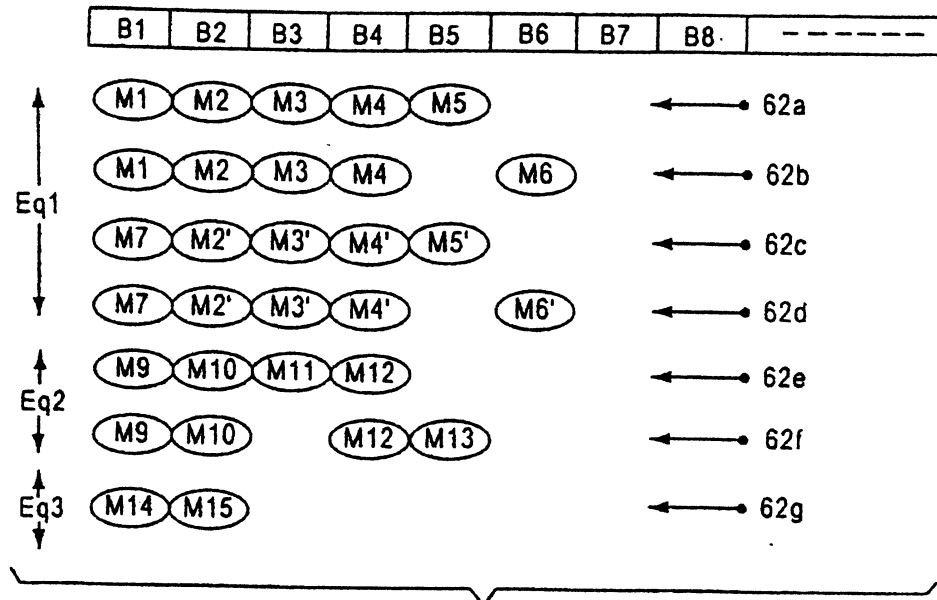




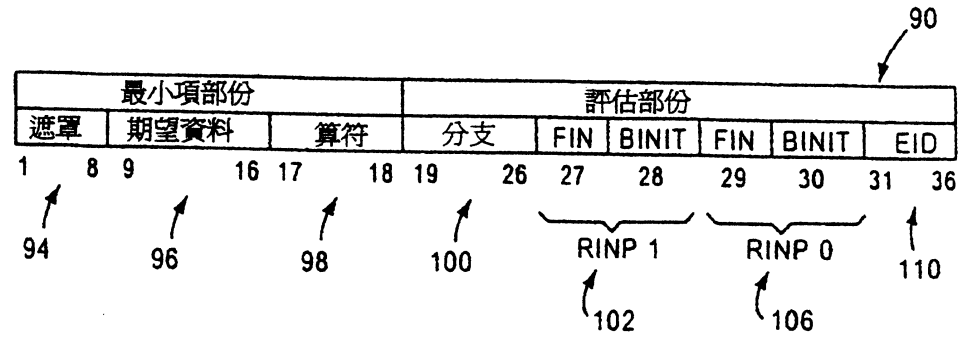
第 6A 圖



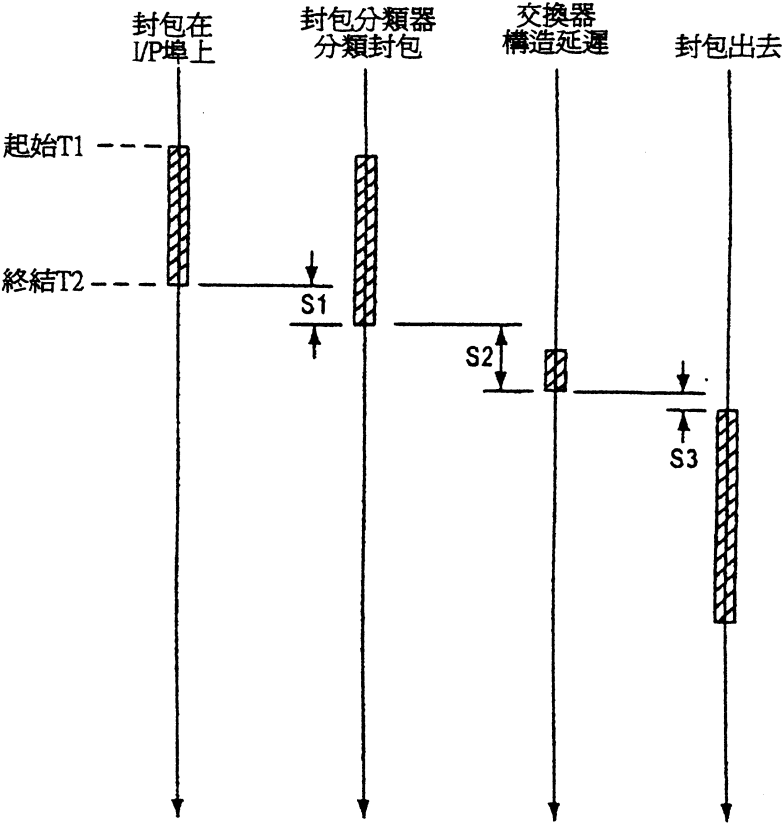
第 6B 圖



第 7 圖



第 8 圖



第 9 圖