



(19) 中華民國智慧財產局

(12) 發明說明書公告本

(11) 證書號數：TW I482460 B

(45) 公告日：中華民國 104 (2015) 年 04 月 21 日

(21) 申請案號：098117841

(22) 申請日：中華民國 98 (2009) 年 05 月 27 日

(51) Int. Cl. : **H04L12/951 (2013.01)**

(30) 優先權：2008/05/30 瑞典 0801280-9
2008/05/30 美國 61/057,614

(71) 申請人：馬維爾國際股份有限公司 (百慕達) MARVELL INTERNATIONAL LTD. (BM)
百慕達

(72) 發明人：諾德馬克 甘那 NORDMARK, GUNNAR (SE)；柏頓 湯瑪斯 BODEN, THOMAS (SE)；卡斯崔姆 賈克柏 CARLSTROM, JAKOB (SE)；蘇科尼克 維塔立 SUKONIK, VITALY (IL)；皮森 馬帝思 PERSSON, MATTIAS (SE)

(74) 代理人：李貞儀

(56) 參考文獻：

TW	200922221A	US	2002/0080789A1
US	2006/0248376A1	US	2007/0230475A1

審查人員：蔡鴻璟

申請專利範圍項數：25 項 圖式數：5 共 38 頁

(54) 名稱

網路處理器單元及其相關方法

A NETWORK PROCESSOR UNIT AND A METHOD FOR A NETWORK PROCESSOR UNIT

(57) 摘要

一種在一網路中處理封包的方法及網路處理器單元(10)，該網路處理器(10)包括：經配置以接收及發送封包的通信介面(14)；至少一個處理封包或其部分的處理構件(16)；一經配置以在該通信介面(14)及該處理構件(16)之間交換封包的嵌入式交換器(12)；且其中該嵌入式交換器(12)經配置以分析一被接收之封包並判定該封包是否應丟棄；如果該封包不應被丟棄，該交換器經配置以儲存該被接收封包、發送該封包之一第一部分至該處理構件(16)以進行處理、從該處理構件(16)接收該封包之經處理的第一部分以及發送該封包之經處理的第一部分。

A method of and a network processor unit (10) for processing of packets in a network, the network processor (10) comprising: communication interface (14) configured to receive and transmit packets; at least one processing means (16) for processing packets or parts thereof; an embedded switch (12) configured to switch packets between the communication interface (14) and the processing means (16); and wherein the embedded switch (12) is configured to analyze a received packet and to determine whether the packet should be dropped or not; if the packet should not be dropped, the switch is configured to store the received packet, to send a first part of the packet to the processing means (16) for processing thereof, to receive the processed first part of the packet from the processing means (16), and to transmit the processed first part of the packet.

- 10 . . . 網路處理器
- 12 . . . 嵌入式交換器
- 14 . . . 通信介面
- 16 . . . 處理構件

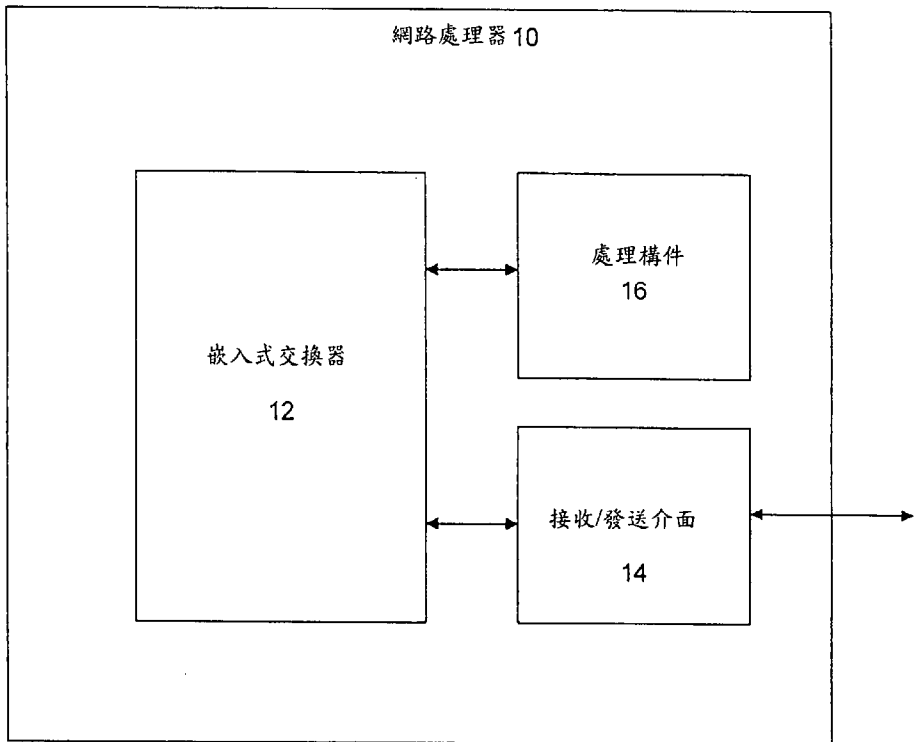
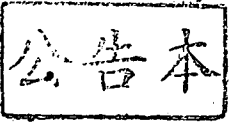


圖 2

發明專利說明書



(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※申請案號：98117841

※申請日：98.5.27

※IPC分類：H04L 12/95 | 2013.01

一、發明名稱：(中文/英文)

網路處理器單元及其相關方法

A NETWORK PROCESSOR UNIT AND A METHOD FOR A
NETWORK PROCESSOR UNIT

二、中文發明摘要：

一種在一網路中處理封包的方法及網路處理器單元(10)，該網路處理器(10)包括：經配置以接收及發送封包的通信介面(14)；至少一個處理封包或其部分的處理構件(16)；一經配置以在該通信介面(14)及該處理構件(16)之間交換封包的嵌入式交換器(12)；且其中該嵌入式交換器(12)經配置以分析一被接收之封包並判定該封包是否應丟棄；如果該封包不應被丟棄，該交換器經配置以儲存該被接收封包、發送該封包之一第一部分至該處理構件(16)以進行處理、從該處理構件(16)接收該封包之經處理的第一部分以及發送該封包之經處理的第一部分。

三、英文發明摘要：

A method of and a network processor unit (10) for processing of packets in a network, the network processor (10) comprising: communication interface (14) configured to receive and transmit packets; at least one processing means (16) for processing packets or parts thereof; an embedded switch (12) configured to switch packets between the communication interface (14) and the processing means (16); and wherein the embedded switch (12) is configured to analyze a received packet and to determine whether the packet should be dropped or not; if the packet should not be dropped, the switch is configured to store the received packet, to send a first part of the packet to the processing means (16) for processing thereof, to receive the processed first part of the packet from the processing means (16), and to transmit the processed first part of the packet.

四、指定代表圖：

(一)本案指定代表圖為：第 (2) 圖。

(二)本代表圖之元件符號簡單說明：

10	網路處理器
12	嵌入式交換器
14	通信介面
16	處理構件

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

(無)

六、發明說明：

【發明所屬之技術領域】

本發明關於一種在一網路中處理封包的網路處理器單元及其一種方法，具體而言關於一種具有嵌入式封包交換器的網路處理器單元及其一種方法。

【先前技術】

一種網路處理器可被用作一單晶片交換器/路由器或用於一模組交換器/路由器中之一進線卡或出線卡。許多先前技術之網路處理器包括多個中間緩衝器，該等緩衝器被使用於該網路處理器之不同模組之間，促使該網路處理器具有超過需要的更多緩衝空間。此外，該等多個中間緩衝器可在處理時引起延遲。

圖1概要顯示一種具有多個中間緩衝器的先前技術網路處理器。如圖所示，該網路處理器1包括一流量管理器2、介面3、具有佇列及排程器的緩衝器4a、4b、4c、處理構件5及可選外部記憶體6a、6b，參見圖1。從該等介面3接收之封包資料在被寫入該流量管理器2之緩衝器之前被緩衝於一輸入緩衝器4a中、被該處理構件5處理、緩衝於一中間緩衝器4b中。來自該流量管理器2的封包資料經由一輸出緩衝器4c在該等介面3上排程及傳送。由於不同系統間不同的要求，封包處理之理想順序亦不同。舉例來說，離開該流量管理器之流量的全部或部分可在如一環回路徑7所示的傳輸之前需要額外通過該處理構件一次。在其他情況下，可能需要通過該流量管理器兩次。

先前技術之該網路處理器1之一缺點在於其需要較大的緩衝器尺寸，比所需更大。舉例來說，如果封包資料之一叢發通過圖1之該網路處理器1，其首先儲存於該輸入緩衝器4a中，之後儲存於該中間緩衝器4b中，最後儲存於該輸出緩衝器4c中。這意味著該等緩衝器之各者：輸入緩衝器、中間緩衝器及輸出緩衝器，必須為一普通數量之封包資料及該叢發大小提供緩衝空間。

先前技術之該網路處理器的另一個缺點在於經由該網路處理器之封包的循序流動。如圖1所示，第一，該等封包於該介面上被接收；第二，其於該輸入緩衝器中被緩衝；第三，其被該處理構件處理；第四，其於該中間緩衝器中被緩衝；第五，其被寫入該流量管理器之緩衝器；第六，其於一輸出緩衝器中被緩衝；及第七，其於該等介面上被傳送。

Su等人之美國專利US 2005/0169259號揭示一種包括若干埠、一封包交換控制單元、一微處理器資料傳送介面、一微處理器及一封包緩衝器的封包交換單元。該封包緩衝器包括一封包區域、一索引緩衝器及一傳送佇列電路系統。該封包交換控制單元及該微處理器資料傳送介面之兩者都可分別從該等埠及該微處理器傳送封包/傳送封包至該等埠及該微處理器。

Su等人的該封包交換單元之一缺點在於其亦需要較大的緩衝空間，因為該等埠具有一儲存及轉發功能，意即該等封包在被移動至該封包緩衝器之前被該等埠儲存。

Langner的美國專利US 2007/0230475號揭示一種基於交換器的網路處理器，其包括若干輸入介面、若干輸出介面、若干處理元件及一具有雙向埠的交換器。該網路處理器將資訊與一經由一輸入介面接收的封包聯繫在一起，且此資訊基於該交換器之交換操作以連續之順序經由該等處理元件之多個循序處理。該網路處理器可被視作依據應用而在該等多個處理元件之輸入及輸出之間提供可選擇之互聯。對於某一種應用，封包將因此以適當順序只流向所需的特定處理元件。處理元件之次序藉由該交換器及該等處理元件之配置而獲得。

Langner之網路處理器的一個缺點在於封包交換對於一種給定應用的預定連續順序，例如一種安全交易伺服器應用、一種安全封包層(SSL)VPN防火牆應用或一種安全網際網路協定(IPSEC)VPN防火牆應用。在配置一種具體應用之後，所有封包將具有相同的連續處理順序且由此將採用相同的路徑通過該網路處理器。

【發明內容】

本發明之一目的為克服先前技術之缺點及不足。更具體而言，本發明之一目的為最小化一網路處理器中緩衝器之尺寸並為封包提供通過該網路處理器的彈性路徑以便適用於不同的系統要求。

其中上述該目的利用一種根據該等獨立請求項的方法及網路處理器單元實現。本發明之進一步實施例將指定於該等附屬請求項中。

【實施方式】

本發明將參考附圖而被更詳細地描述，在該等圖式中相同的參考數字指示相同或相應的特徵、元件或構件。

根據本發明，圖 1 所顯示之先前技術網路處理器的輸入緩衝器、中間緩衝器及輸出緩衝器由一種嵌入式交換器替代。根據本發明，一功能塊，例如，諸如埠、處理構件、流量管理器、封包格式器或另一個與該嵌入式交換器通信並寫入封包資料至該交換器之元件的流量介面，經配置以選擇該封包資料之輸出位址。在實施例中，該嵌入式交換器包括一緩衝記憶體，該緩衝記憶體在朝向該網路處理器之不同模組(例如功能塊)的佇列之間共用，例如一處理構件、一個或更多個介面及一可能的流量管理器之佇列，藉此減小所需的總緩衝尺寸。

由此，根據本發明，相同的封包叢發在該嵌入式交換器中一段時間內只發生一次。進一步，寫入封包資料至該嵌入式交換器的功能塊所藉以選擇輸出位址，例如輸出埠的本發明之定址機構啟用經由該網路處理器之模組，例如經由該處理構件及一流量管理器的資料路徑之彈性組織。此外，藉由本發明基於設定的屬性將來自該等接收介面的封包直接交換至該等傳送介面係可行的，該等屬性可基於利用該等封包接收的資訊，例如接收於該等封包標頭中的資訊。進一步，本發明之統一通信介面及彈性定址使不同的子系統，例如處理構件，能連接至該嵌入式交換器之不同埠。

應理解使用於本申請案文件中的術語「封包」可涉及多種類型的資料單元，包含但不限於網際網路協定第4版(IPv4)封包、網際網路協定第6版(IPv6)封包、IEEE 802.3訊框、點對點協定(PPP)訊框、ATM單元或其他協定資料單元(PDU)。進一步，一封包可為如此一PDU，例如單元、資料段或記憶體頁；一控制訊息；或在一網路處理系統中互通的另一類型之資料單元之一部分。

圖2概要顯示一種根據本發明之一實施例的網路處理器10，有時其亦被稱作一種網路處理單元。該網路處理器10包括一交換器12、一通信介面14及一處理構件16。該交換器12經配置以便與該介面14及該處理構件16通信。

圖3概要顯示一進一步包括一流量管理器18的網路處理器10之一實施例，該流量管理器18經配置以與該交換器12及該等通信介面14通信。選擇性的外部記憶體20、22可經配置以分別與該處理構件16及該流量管理器18通信。

該網路處理器可為多種熟知類型中的任一種；包含描述於全文以引用方式併入本文的國際專利申請案第PCT/EP2007/055777號中之處理器，但其亦可為一種加密處理器、一種字串搜索處理器及/或一種通用CPU。該處理器可包括多種熟知類型的處理構件；包含一同步處理管線，如全文以引用方式併入本文的國際專利申請案第PCT/SE2005/001969號所描述。進一步，該處理器可包括諸如一封包格式器的處理構件，及/或一流量管理器。

該交換器12經配置以提供流量管理功能，包含服務分

級、階層式排程、流量調整以及空間及邏輯多播。

進一步，交換器12，較佳的係一嵌入式交換器，此處亦被稱為一種共用記憶體交換器(SMS)，經配置以在該等通信介面之間交換封包，例如該等接收/傳送介面14之一者及該處理機構16之間。在諸實施例中，較佳的係該交換基於資訊，例如利用該等封包接收的設定屬性。此外，該交換器12經調適以按一種與該資料路徑上之系統要求匹配(即與該線速匹配)的資料率交換封包資料。

圖4a、圖4b概要顯示一種網路處理器10的實施例，其中該交換器12包括一封包緩衝器24、輸入埠(有時亦被稱為寫入埠)26a、輸出埠(有時亦被稱為讀取埠)26b、一輸入仲裁器28、一佇列映射器30、一丟棄單元34、一佇列管理器36及一封包排程器38。如圖4所示，該網路處理器10之實施例亦可包括一經配置以與該交換器12通信的封包格式器32。

在實施例中，該交換器12經配置以便使用該交換器12之一嵌入式記憶體中的一個緩衝器來完成封包交換，參見圖4a及圖4b。該緩衝器可例如為一種諸如一封包緩衝器24的嵌入式緩衝器。

如圖4a及圖4b所例示，有時亦被稱為主封包緩衝器的該封包緩衝器24包括一用於儲存封包資料的資料儲存區24-1及一控制單元24-2，該控制單元包括實施用於多播及/或單播之佇列及構件的邏輯及控制記憶體。該控制記憶體經配置以儲存鏈結列表元素，包括例如指標到封包資料、指標

到其他鏈結列表元素及諸如封包類型及封包長度的封包屬性。該共用記憶體交換器 12 之封包緩衝器 24 可使用一多重埠 26 (26a、26b)；一處理構件埠，例如一可編程管線埠 (PP 埠)、一流量介面埠 (TI 埠)、一可選封包格式器埠 (PF 埠) 及/或一可選流量管理器埠 (TM 埠) 中的一個或更多個讀取或寫入。

該等 PP、TI、PF、TM 埠 26 可包括一個或更多個子埠，即該埠中的一個或更多個邏輯位址。該等子埠可由來自該接收子系統，例如來自該處理構件 16、該流量管理器 18 及/或該封包格式器 32 的背壓節流。舉例來說，一相應於一 TX 乙太網路 MAC 的 TI 子埠可在該排程器中為朝向該等流量介面 14 之佇列節流一映射至此子埠的節點。該 PP 讀取埠可具有 16 個相應於流量介面的子埠。該 TI 讀取埠可具有 64 個相應於流量介面的子埠。該 TM 讀取埠可具有一單個子埠。該 PF 讀取埠可具有 2 個子埠。具有不同子埠的封包可在其經由一埠 26 讀取或寫入時交錯；即，一屬於一第一子埠的第一封包之一第一部分可在該第一封包之一第二部分經由一埠讀取或寫入之前經由該埠讀取或寫入，之後為一第二封包之一第一部分。

該封包緩衝器 24 可經配置以便以一種儲存-轉發方式操作，即該封包緩衝器 24 經配置以便在發送出一完全進入之封包之前儲存該封包，因此該記憶體交換器 12 可確認被接收之封包的完整性並丟棄誤標封包。進一步，在實施例中，該封包緩衝器 24 經配置以儲存 4 MB 之封包資料並吸

收上至32 K之封包。其可支援一400 MHz的封包寫入率及封包讀取率。在另一實施例中，其可支援一800 MHz的封包寫入率及一400 MHz的封包讀取率。然而，應理解該封包讀取率及該封包寫入率可改變且分別可為另一個適當的寫入率及讀取率。該封包緩衝器24可進一步包括64個多播佇列，其各者可持有64個封包。

在實施例中，該交換器12經配置以便不修改封包校直，即不修改該封包或該封包中所包括的資料。進一步，為了向該網路處理器10所接收之封包的附加標頭保留空間，較佳的係該等流量介面14應被配置以移動封包偏移，例如藉由添加一封包偏移值PBase。該流量介面14亦可在傳輸至例如一連接單元之前從該交換器12除去該等封包之過量標頭空間。其他經配置以接收及傳送完整封包的處理構件亦可移動該封包偏移。這些處理構件包括，例如，該流量管理器、該封包格式器、加密處理單元或字串搜索處理單元。

在實施例中，當一封包從一佇列讀取以插入該處理構件16中時，該第一個256位元組被發送至該處理構件16中。如果該封包長度PLen及該封包偏移值PBase之總和為256位元組或更少，該整個封包進入該處理構件16。如果該封包更長，即如果PLen及PBase之總和大於256位元組，該封包之標尾保留於該封包緩衝器24之一專用區域中，例如該資料儲存區24-1中，並在該封包標頭從該處理構件16離開時與該封包標頭重連。然而，在另一實施例中，該整個封包

被發送至該處理構件16而不管該封包長度。進一步，該處理構件16可經配置以處理一封包之若干部分、整個封包及其一組合。亦應理解被交換器接收之該等封包或其若干部分可被交錯。

在實施例中，如果被接收之封包具有不同長度，該封包率在封包較短(例如在64位元組與256位元組之間並包括64位元組與256位元組)時變得比封包較長(例如大於256位元組)時更高。在僅長封包之一第一部分被發送至該處理構件16的實施例中，該處理構件16之容量可超額；即低於以短封包之最大封包率處理封包所需之容量。如果該處理構件16超額，該嵌入式交換器12可儲存短封包之臨時叢發。如果這些叢發之範圍有限且該處理構件16之容量大於平均封包率，則所有被接收之封包可仍被該處理構件16處理，因為該處理構件16可趕上並處理先前在具有長封包之期間儲存的叢發。

在長封包之標頭被發送至該處理構件16且處理構件16充當一種適於以一受保證的150 MHz之封包率或另一適用封包率處理短封包的同步管線之實施例中，若封包較長則可能存在未使用的處理容量，因為長封包之封包率可小於受保證之封包率，促使封包可插入該同步管線的時槽不被使用。在如此之實施例中從該處理構件16離開的封包標頭可儲存於該嵌入式交換器12中並重新進入該處理構件16進行下一步處理、使用原本未使用之處理容量。在處理完成後，該等封包標頭從該處理構件16離開並與該嵌入式交換

器12中的封包標尾重連。如此之擴展處理係有用的，例如如果該封包組之一子組比其他封包需要更多處理，例如因為其需添加附加之協定標頭。

該交換器12經配置以支援單播、邏輯多播及/或空間多播。封包類型及附加輸入參數作為封包屬性用該封包傳至該寫入埠中，如下所述。

在實施例中，該嵌入式交換器12包括用於邏輯多播的構件，藉此至少部分基於用該封包接收之資訊且視情況亦部分基於一個或更多個查詢表中之資訊的該嵌入式交換器12藉由用一寫至一佇列的封包儲存一屬性NumCopies而儲存副本數量。當該封包從該佇列讀取時，該封包之NumCopies副本被傳送。

由此，藉由邏輯多播，一封包被多次讀取。當如此一封包出隊時，副本數量之一參考計數器可被減量。然而，應理解一參考計數器可為各個副本增量直到達到最大副本數量。如果該副本數量達到一最小數量例如零或最大數量，該封包從該佇列釋放。否則該封包保留於該佇列中並因被該相應排程器選擇而重複出隊。邏輯多播可支援上至每封包64 K副本。其適用於一較大範圍的應用，包含第3層多播，即ISO OSI模型第3層多播；例如IP封包多播及VPLS(虛擬私人LAN服務)泛播。

在實施例中，該嵌入式交換器12包括用於空間多播之構件，藉此至少部分基於用該封包接收之資訊且視情況亦部分基於一個或更多個查詢表中之資訊的該嵌入式交換器12

選擇至少兩個佇列並寫入一被接收之封包至該等兩個被選擇之佇列。

由此，藉由空間多播，一封包描述符平行寫入多個空間多播佇列而非普通佇列，一參考計數器被初始化。對於各個空間多播佇列，一多播副本被建立且該參考計數器被增量或減量。如果該副本數量達到一最小數量例如零或最大數量，該封包從該佇列釋放。這些空間多播佇列可經配置以遮蔽該交換器 12 中的佇列。這意味著一空間多播佇列被讀取而非一被該相應排程器選擇之佇列。空間多播適用於至實體埠的低延遲多播，其中多媒體傳播或第 2 層多播或廣播為代表性應用。各個空間多播佇列可持有上至 64 個任意尺寸的封包。

應理解邏輯及空間多播可被組合。對於寫入至該空間多播佇列的封包，副本數量可上至 64 K。此組合可用於在一實體埠內產生邏輯子埠之副本。

在實施例中，該交換器 12 之輸入仲裁器 28 經配置以服務於來自該交換器 12 之該等輸入埠 26a 的交錯封包並將其傳遞至朝向該封包緩衝器 24 的佇列映射器 30，參見圖 4a 及圖 4b。由此，該輸入仲裁器 28 可經配置以便對從一個或更多個輸入埠 26 到該交換器 12 的封包資料進行緩衝及排程。來自一個或更多個輸入埠 26 的封包資料 43a 被該輸入仲裁器 28 所包括並服務的佇列接收。該輸入仲裁器 28 經配置以交錯來自不同輸入埠 26a 的封包資料 43a、並一次寫入一個記憶體頁至該交換器 12 中。如果佇列填滿，諸如背壓 44a、

44b、標尾丟棄或標頭丟棄的溢出保護機制可被啟動。

在實施例中，該佇列映射器30經配置以從該輸入仲裁器28接收封包屬性42a。封包屬性被該流量介面14、該可選封包格式器32、該可選流量管理器18及/或該處理構件16中的一個或更多個用封包發送。在其他實施例中，封包屬性亦可獨立發送而不用發送封包。該佇列映射器30經配置以便對該等封包屬性解碼並產生一從該等封包屬性映射的佇列。該佇列映射包括一佇列號碼及視情況一個或更多個該封包所寫入之多個多播佇列，以及一旦該封包出隊產生何等數量之副本。該佇列映射亦可被一丟棄單元34使用，其經配置以決定是否丟棄該封包或使該封包入隊。

在實施例中，該流量介面14、該處理構件16、該流量管理器18及該可選封包格式器32中的至少一個基於從該交換器12接收之封包資料及/或封包屬性的處理而設定封包屬性。此處理可藉由電腦程式之執行、藉由有限狀態機或藉由使用諸如內容可定址記憶體或RAM中之查詢表的記憶體查詢表而分類或藉由其他可被技術熟練者理解之適當方式而實現。在其他實施例中，該等封包屬性逐埠、子埠、流或一埠中的其他邏輯單元配置。

在一個實施例中，封包屬性42a為一種在該輸入埠26a上的帶外信號，即該等屬性不經由與該等封包相同的通道發送。在另一實施例中，封包屬性42a被帶內發送且該交換器12，例如該佇列映射器30，經配置以從該封包資料43a獲取該等封包屬性43a。舉例來說，該佇列映射器30或該

輸入仲裁器 28 可經配置以獲取/取回儲存於該封包之一標頭及/或一標尾中的屬性。該標尾可包括置於該封包末端的補充資料。

該佇列映射器 30 包括一經配置以儲存關於該佇列映射之資訊的映射資料庫 31。

該等封包屬性可含有一直接或間接佇列映射。在直接映射之情況下，該等封包屬性含有一佇列映射，該佇列映射被直接轉達至一丟棄單元以判定該封包資料是否應被該佇列管理器寫至相應於該直接佇列映射的佇列。在間接佇列映射之情況下，該佇列映射器 30 使用封包屬性 42a 以便從該映射資料庫 31 查詢該佇列映射，然後該等封包屬性被轉達至該丟棄單元以判定該封包資料是否應被該佇列管理器寫至被查詢佇列。然而，應理解該等屬性可被轉達至該丟棄單元以判定該封包是否應在直接或間接映射該封包至一佇列之前被丟棄。

進一步，該等封包屬性可含有關於封包丟棄、封包開始、封包結束、錯誤、封包長度、至封包欄位之偏移、來源節點、來源埠、來源子埠、來源佇列、終點節點、終點埠、終點子埠、終點佇列、服務級別、顏色、反射、儲存位址、記憶體佔用、多播狀態、資料保護的資訊及其它可與一封包關聯的資訊。

在實施例中，該交換器 12 包括拋棄一被接收之封包而非將其寫入一佇列的構件。拋棄該封包之決定至少部分基於當前佇列長度及/或平均佇列長度及/或儲存於表中之資訊

及/或用該封包接收之資訊。

拋棄一封包之構件可藉由前述之一丟棄單元34實現。

在實施例中，該交換器12進一步包括在一佇列頭丟棄一封包而非將其傳送至一輸出埠的構件。丟棄該封包之決定可至少部分基於當前佇列長度及/或平均佇列長度及/或儲存於表中之資訊及/或用該封包接收之資訊及/或該佇列中封包所消耗之時間。

丟棄一封包之構件可藉由一丟棄單元34實現，該丟棄單元34經配置以便為各個接收封包檢查對緩衝器保留及主動佇列管理之規則的符合性並產生一通過或丟棄決定。該丟棄單元34可經配置以包含用於加權隨機早期檢測(WRED)之構件以及用於視情況共用記憶體池之佇列之標尾丟棄的構件。該丟棄單元34亦可包括經配置以檢測封包錯誤的構件。該丟棄單元34亦可維持該等佇列長度並包含用於生成背壓的配置暫存器。

藉由該丟棄單元34中的配置，該封包緩衝器24可分成數個不同記憶體使用類型的區域。該丟棄單元34可經配置以監督這些區域之使用並為封包丟棄或背壓應用一組機制以保護緩衝記憶體區域並防止致命的緩衝狀況。

在實施例中，該交換器12亦可包括一佇列管理器36，該管理器36包括一經配置以保持佇列狀態並管理佇列及自由控制記憶體之鏈結列表的入隊鎖。該入隊鎖可進一步藉由因該丟棄單元之要求將封包及記憶體頁附加至鏈結列表之標尾而實行封包入隊。該佇列管理器26可進一步包括一出

隊鎖。該出隊鎖經配置以藉由因一封包排程器 38 之要求從鏈結列表彈出封包及記憶體頁而實行封包出隊。

在實施例中，該交換器 12 包括一個或更多個內部佇列組，組數量相應於埠數量。

如圖 5a 所示，在實施例中，該交換器 12 包括一佇列組；朝向該處理構件 16 的 PP 佇列，一佇列組；朝向該流量介面 14 或該可選流量管理器 18 的 TX 佇列，一佇列組；至該可選流量格式器 32 的 PF 佇列。各個佇列組具有一封包排程器 38。

如圖 5b 所示，在實施例中，該交換器 12 包括一個佇列組；朝向該處理構件 16 的 PP 佇列，一個佇列組；朝向該流量介面 14 的 TX 佇列，一個佇列組；朝向該可選流量管理器的 TM 佇列，一個佇列組；至該可選封包格式器 32 的 PF 佇列，以及一個佇列組；用於待丟棄之封包資料之排程讀取操作的丟棄佇列。各個佇列組具有一封包排程器 38，其中一些顯示於圖 4a 及圖 4b 中。

進一步，如圖 4a、4b 及圖 5a、5b 所示，該交換器 12 包括一封包排程器 38，該排程器 38 經配置以便將來自該交換器 12 之該等佇列的封包排程至一個或更多個與該交換器 12 通信的功能塊，該排程利用一種排程演算法執行。該交換器 12 可在每個輸出埠具有至少一個排程器，其判定該等佇列之服務順序。

在實施例中，該封包排程器 38 經配置以利用欠額加權輪詢法 (DWRR)、嚴格優先權佇列 (SPQ) 及輪詢演算法 (RR) 中

的一個或一組合對該等封包排程。然而，應理解亦可使用其他適用的排程演算法。舉例來說，該等輸出埠排程器可根據該排程演算法輪詢、加權輪詢、欠額輪詢、欠額加權輪詢、嚴格優先權佇列、先到先服務、最早時限優先、多時分工或另一種適用排程演算法中的至少一個服務佇列。

在實施例中，該封包排程器38可包括一個或更多個包括至少兩個排程器階層的階層式排程器。舉例來說，該封包排程器38可包括一處理構件排程器；PP排程器，其經配置以將封包資料排程至該處理構件16；一流量管理器排程器；TM排程器，其經配置以將封包資料排程至該流量管理器；一介面排程器；TX排程器，其經配置以將封包資料排程至該等流量介面；TX介面14b；一封包格式排程器；PF排程器，其經配置以將封包資料排程至該封包格式器32，及/或一丟棄排程器，其經配置以排程封包丟棄。

交換器之該等輸出埠排程器，例如TX排程器，包括諸如符記桶、漏桶的流量調整器或其他適於實現流量調整的方式或演算法，其限制從各個排程器輸出的資料率。

該PP排程器及該TX排程器之各者可經配置以服務於512個佇列，該PF排程器可經配置以服務於16個佇列。

在實施例中，該交換器可包括一輸出仲裁器41，參見圖4b及圖5b。該輸出仲裁器41可相應於該輸入仲裁器28。該輸出仲裁器41經配置以從該封包排程器38接收對於封包讀取操作的請求，並將這些操作轉換為來自該主封包緩衝器24的資料讀取操作。該輸出仲裁器41經配置以控制輸出埠

26a之間的讀取帶寬共用。在實施例中，該輸出仲裁器41經配置以利用欠額加權輪詢法(DWRR)、嚴格優先權佇列(SPQ)及輪詢演算法(RR)中的一個或一組合以對該等封包排程。然而，應理解亦可使用其他適用排程演算法。舉例來說，該等輸出埠排程器可根據該排程演算法輪詢、加權輪詢、欠額輪詢、欠額加權輪詢、嚴格優先權佇列、先到先服務、最早時限優先、多時分工或另一種適用排程演算法中的至少一個服務佇列。在實施例中，至少一個輸出埠具有子埠。在這種情況下，該輸出仲裁器41進一步對這些子埠排程並控制該輸出埠之該等子埠之間的帶寬共用。

該交換器12之實施例包括一狀態格式器40，該狀態格式器經安排以與該封包緩衝器24及該封包排程器38通信並經配置以便在佇列之階層超過經配置之浮水印時產生至該封包排程器的背壓。藉由可配置之連接矩陣的優點，該等背壓信號可彈性映射至該封包排程器中的佇列或節點。

如圖5a及圖5b所示，排程可被來自該處理構件16、該封包格式器32、該等流量介面14或該流量管理器18的背壓開啟/關閉調節。換言之，封包資料至該處理構件16、該封包格式器32、該等封包介面14或該流量管理器18的排程可被中斷，如果它們的緩衝器全滿且不能接收更多的封包資料。該封包排程器38停止排程及發送封包資料直到該等緩衝器被清空並能夠再次接收封包資料。

如下之實例描述該網路處理器對四個例示封包的處理。這些及其它封包可同時用該網路處理器之相同配置處理。

由此，相同網路處理器配置可導致經由該網路處理器的不同流動路徑。

所有四個封包43a首先採取相同路徑。它們被一流量介面14接收。藉由該流量介面14之配置，該等封包屬性42a被設定為一至一朝向該處理構件16之PP佇列的直接映射。除非該封包被該丟棄單元34丟棄，否則該封包將進入被選擇之PP佇列且其封包資料儲存於該主封包緩衝器24之資料儲存區24-1中。從這個PP佇列該封包被該封包排程器38排程，該封包資料之部分或所有被該輸出仲裁器41從該主封包緩衝器24讀取並經由該PP埠26b發送至該處理構件16。該等封包屬性基於佇列號碼被該交換器12設定。該處理構件16藉由一執行程式編碼序列的電腦程式處理該封包，該電腦程式可執行封包資料及/或封包屬性之分類及編輯以及技術熟練者所熟知的其他封包資訊及資料庫操作。該處理構件16檢驗諸如封包標頭的封包內容並執行表查詢。基於這些操作該等四個封包以四種不同方式處理，其將以如下四點1)-4)描述。

1)該處理構件16將該第一例示封包識別為一應通過該交換器12、該流量管理器18並在其被發送至該等流量介面14之前再次通過該交換器12的用戶封包。為實現此，該處理構件16將一持有至該流量管理器18之輸入參數的標頭添加至該封包並將封包屬性設定為一種至一朝向該流量管理器18之佇列的直接映射。在該封包經由該交換器12傳遞後，如上述，該流量管理器18使用該輸入標頭中的參數以設定

以至一朝向該等流量介面14之佇列的直接映射。在另一次通過該交換器12至該等流量介面14後，該封包被發送至一外部接收器(未顯示)。

2)該處理構件16將該第二例示封包識別為一種應直接通過該交換器12至該等流量介面14的低延遲控制封包。為實現此，該處理構件16將封包屬性設定為一種至一朝向該等流量介面14之佇列的直接映射。在另一次通過該交換器12至該等流量介面14後，該封包被發送至一外部接收器(未顯示)。

3)該處理構件16將該第三例示封包識別為一多播封包，例如一多播視頻封包。此封包應複製至多個例項並發送至該等流量介面14之不同埠。為實現此，該可編程管線16將封包屬性設定為一識別一多播群的直接映射。然而，應理解一種至一多播佇列的直接映射亦係可行的，但因為存在許多多播佇列，其要求資源以提供一直接映射，因此，為節省帶寬，較佳的係提供一種間接映射。該佇列映射器30將該多播群用作用於在一映射資料庫31中查詢的密鑰。此查詢返還該封包入隊的多播佇列之一列表。除非一副本被該丟棄單元34丟棄，各個副本被寫至一多播佇列。在通過該交換器12至該等流量介面14之後，各個封包被發送至一外部接收器(未顯示)。

4)該處理構件16將該第四例示封包識別為一種應被例如被IPv4分片操作分片的用戶封包。此封包應通過該交換器12、該封包格式器32、再次通過該交換器12、該流量管理

器18及在其被發送至該等流量介面14以便傳輸之前最後一次通過該交換器12。為實現此，該處理構件16將一輸入參數之標頭添加至該封包格式器32並將封包屬性設定為一種至一朝向該流量管理器18之佇列的直接映射。在該封包經由該交換器12傳遞後，如上述，該封包格式器32使用配置參數或該輸入標頭中之參數以控制待執行之功能及/或操作。進一步，該封包格式器32設定一直接映射至一返回該處理構件16的佇列。此後將採用從該處理構件16開始處理該第一例示封包的相同步驟。

雖然本發明已根據所顯示的實施例而被描述，但一般技術者將瞭解改變可被做出而不脫離本發明之範圍。因此，如上描述中所包含及該等圖式中所顯示的所有內容僅具說明性，而不應理解為限制性。

【圖式簡單說明】

本發明之實施例將參考如下圖示而被更詳細地描述，其中：

圖1概要顯示一種先前技術之網路處理器的方塊圖；

圖2概要顯示一種根據本發明之一實施例的網路處理器之一方塊圖；

圖3概要顯示一種根據本發明之一實施例的網路處理器之一方塊圖；

圖4a概要顯示一種根據本發明之網路處理器之一實施例的方塊圖；

圖4b概要顯示一種根據本發明之網路處理器之一實施例

的方塊圖；

圖 5a 概要顯示一種根據本發明之網路處理器之一實施例的方塊圖，其中顯示一種嵌入式交換器之一排程圖，及

圖 5b 概要顯示一種根據本發明之網路處理器之一實施例的方塊圖，其中顯示一種嵌入式交換器之一排程圖。

【主要元件符號說明】

1	網路處理器
2	流量管理器
3	介面
4a	輸入緩衝器
4b	中間緩衝器
4c	輸出緩衝器
5	處理構件
6a	可選外部記憶體
6b	可選外部記憶體
7	環回路徑
10	網路處理器
12	嵌入式交換器
14	通信介面
14a	流量介面
14b	TX 介面
16	處理構件
18	流量管理器
20	可選外部記憶體

22	可選外部記憶體
24	主封包緩衝器
24-1	資料儲存區
24-2	控制單元
26a	輸入埠
26b	輸出埠
28	輸入仲裁器
30	佇列映射器
31	映射資料庫
32	封包格式器
34	丟棄單元
36	佇列管理器
38	封包排程器
40	狀態格式器
41	輸出仲裁器
42a	封包屬性
42b	封包屬性
43a	封包資料
43b	封包資料
44a	背壓
44b	背壓

七、申請專利範圍：

1. 一種在一網路中處理諸多封包的網路處理器單元(10)，該網路處理器單元(10)包括：
 - 經配置以接收並傳送諸多封包的通信介面(14)；
 - 至少一個用於處理諸多封包或其部分的處理構件(16)；
 - 一經配置以在該通信介面(14)及該處理構件(16)之間交換諸多封包的嵌入式交換器(12)；及
 - 其中該嵌入式交換器(12)經配置以分析一被接收之封包並判定該封包是否應丟棄；如果該封包不應被丟棄，該交換器經配置以儲存該被接收封包；發送該封包之第一部分至該處理構件(16)以進行處理；從該處理構件(16)接收該封包之該經處理的第一部分以及發送該封包之該經處理的第一部分。
2. 如請求項1之網路處理器單元(10)，其中該通信介面(14)、該處理構件(16)或諸如經配置以與該嵌入式交換器(12)通信之一流量管理器(12)或一封包格式器(32)的另一個功能塊(12、14、16、32)經配置以基於該封包資料之處理而設定各封包屬性。
3. 如請求項1的網路處理器單元(10)，其中該通信介面(14)、該處理構件(16)或諸如經配置以與該嵌入式交換器(12)通信之一流量管理器(12)或一封包格式器(32)的另一個功能塊(12、14、16、32)經配置以基於從該嵌入式交換器(12)接收的資訊而設定各封包屬性。
4. 如請求項2至3中任一項的網路處理器單元(10)，其中該

- 交換器(12)依據該等被設定的封包屬性而判定一功能塊(12、14、16、32)並發送該封包至該功能塊(12、14、16、32)。
5. 如請求項1至3中任一項的網路處理器單元(10)，其中該封包之該第一部分包括該整個封包。
 6. 如請求項1至3中任一項的網路處理器單元(10)，其中該嵌入式交換器(12)經配置以儲存該封包之一第二部分、從該處理構件(16)接收該封包之該經處理的第一部分及在發送經重組之封包之前將該封包之該經處理的第一部分及該封包之該第二部分重組。
 7. 如請求項6之網路處理器單元(10)，其中該封包之該第一部分為該封包之標頭，該封包之該第二部分為該封包之主體，且其中該嵌入式交換器(12)包括一用於在該第一部分被處理時儲存該第二部分的封包緩衝器(24)。
 8. 如請求項1至3中任一項的網路處理器單元(10)，其中該嵌入式交換器(12)經配置具有來自該通信介面(14)及該處理構件(16)的輸入埠(26a)；並具有至該通信介面(14)及該處理構件(16)的若干輸出埠(26b)，藉此一依據該等被設定之封包屬性接收於任一輸入埠(26a)上的封包可被寫入至任一輸出埠(26b)之一輸出佇列並經由該輸出埠(26b)發送。
 9. 如請求項8的網路處理器單元(10)，其進一步包括一經配置與該嵌入式交換器(12)及該通信介面(14)通信的流量管理器(18)，其中該嵌入式交換器(12)經配置具有一來

自該流量管理器(18)的輸入埠(26a)及一至該流量管理器(18)的輸出埠(26b)，藉此一接收於任一輸入埠(26a)上的封包可被寫入至任一輸出埠(26b)之一輸出佇列並經由該輸出埠(26b)發送。

10. 如請求項8的網路處理器單元(10)，其中該嵌入式交換器(12)包括一經配置以對來自一個或更多個輸入埠(26a)之封包資料進行緩衝及排程的輸入仲裁器(28)、及一經配置以從該輸入仲裁器(28)接收若干封包屬性的佇列映射器(30)。
11. 如請求項10的網路處理器單元(10)，其中該嵌入式交換器(12)包括一經配置以利用一諸如輪詢演算法、加權輪詢、欠額輪詢、欠額加權輪詢、嚴格優先權佇列、先到先服務、或最早時限優先的排程演算法對來自該交換器(12)之佇列的諸封包進行排程的封包排程器(38)。
12. 如請求項8的網路處理器單元(10)，其中各個輸入埠(26a)及/或各個輸出埠(26b)可包括一個或更多個子埠。
13. 如請求項1至3中任一項的網路處理器單元(10)，其中該嵌入式交換器(12)經配置以支援單播、邏輯多播及/或空間多播。
14. 一種網路處理器單元(10)在一網路中處理諸多封包的方法，該方法包括：
 - 藉由一通信介面(14)接收並發送各封包；
 - 藉由至少一個處理構件(16)處理各封包或封包之部分；

藉由一嵌入式交換器(12)，

分析一被接收之封包，

判定該封包是否應丟棄；如果該封包不應被丟棄，

藉由該交換器(12)：

儲存該被接收之封包；

發送該封包之一第一一部分至該處理構件(16)以對其進行處理，

從該處理構件(16)接收該封包之該被處理的第一一部分，及

發送該封包之該被處理的第一一部分。

15. 如請求項14的方法，其中，藉由該通信介面(14)、該處理構件(16)或諸如一經配置以與該嵌入式交換器(12)通信之一流量管理器(18)或一封包格式器(32)的另一個功能塊(12、14、16、32)，各封包屬性基於該封包資料之處理而被設定。

16. 如請求項14之方法，其中，藉由該通信介面(14)、該處理構件(16)或諸如一經配置以與該嵌入式交換器(12)通信之一流量管理器(18)或一封包格式器(32)的另一個功能塊(12、14、16、32)，各封包屬性基於從該嵌入式交換器(12)接收之資訊而被設定。

17. 如請求項15至16中任一項的方法，其進一步包括如下步驟：藉由該交換器(12)並依據該等被設定之封包屬性判定一功能塊(12、14、16、32)並發送該封包至該功能塊(12、14、16、32)。

18. 如請求項14至16中任一項的方法，其中該封包之該第一部分包括該整個封包。
19. 如請求項14至16中任一項的方法，其進一步包括：
儲存該封包之一第二部分，
從該處理構件(16)接收該封包之該被處理的第一部分，及
在發送被重組之封包之前將該封包之該被處理之第一部分及該封包之該第二部分重組。
20. 如請求項14至16中任一項的方法，其中該封包之該第一部分為該封包之標頭，該封包之該第二部分為該封包之主體，且其中該第二部分在該第一部分被處理時儲存於該嵌入式交換器(12)之一封包緩衝器(24)中。
21. 如請求項14至16中任一項的方法，其中一接收於該交換器(12)之任一輸入埠(26a)上的封包依據該等被設定的封包屬性可被寫入至該交換器(12)之任一輸出埠(26b)並經由該輸出埠(26b)發送。
22. 如請求項21的方法，其中，藉由一包括於該嵌入式交換器(12)中的輸入仲裁器(28)，接收於一個或更多個該等輸入埠(26a)上的封包資料被緩衝及排程，且其中，藉由一包括於該嵌入式交換器(12)中的佇列映射器(30)，來自該輸入仲裁器(28)的各封包屬性被接收。
23. 如請求項22的方法，其中，藉由一包括於該嵌入式交換器(12)中的封包排程器(38)，封包資料使用一諸如輪詢演算法、加權輪詢、欠額輪詢、欠額加權輪詢、嚴格優

先權佇列、先到先服務或最早時限優先的排程演算法從該交換器(12)之各佇列排程。

24. 如請求項14至16中任一項的方法，其進一步包括單播、邏輯多播及/或空間多播。
25. 一種電腦程式產品，其包括在被一電腦化單元上執行時用於執行如請求項14至16項中任一項的方法之電腦程式編碼序列。

八、圖式：

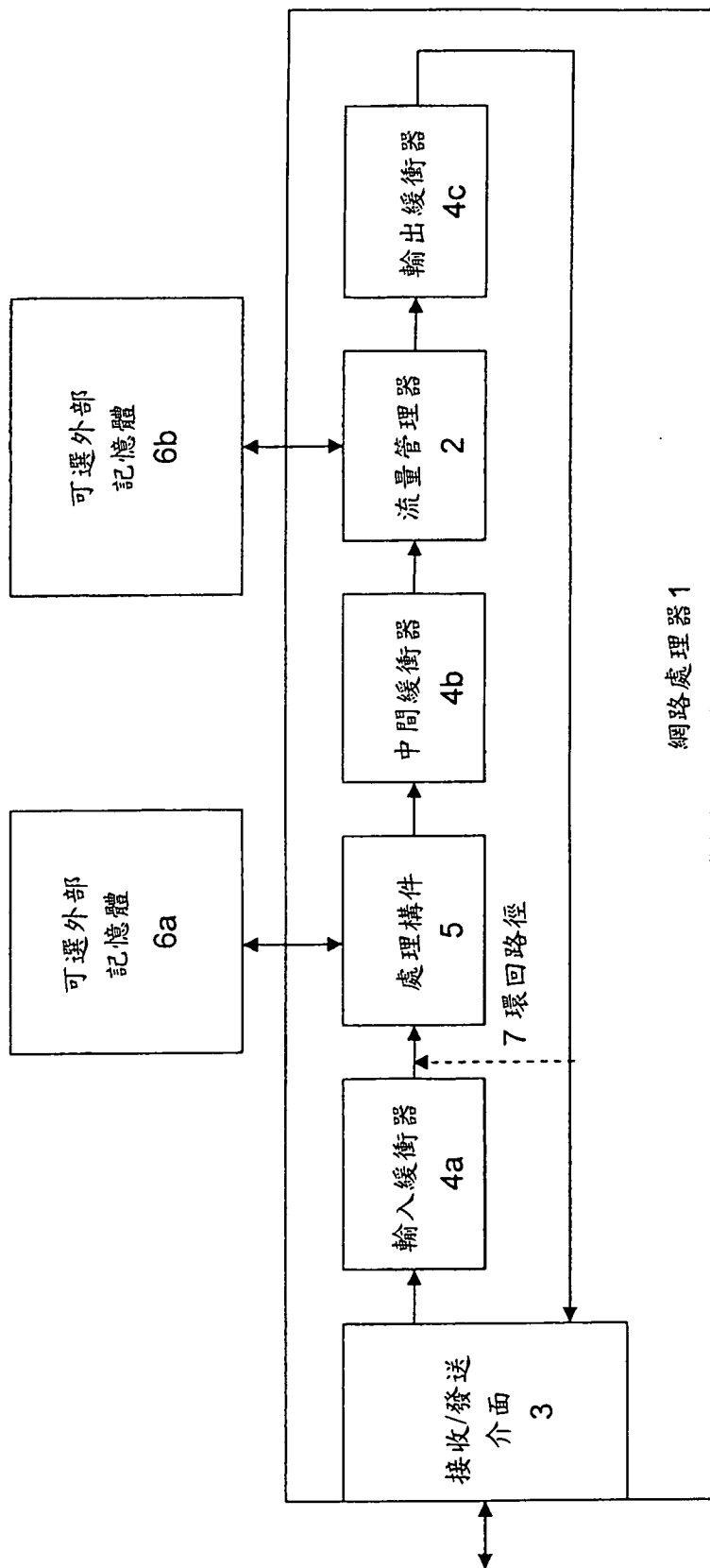


圖 1

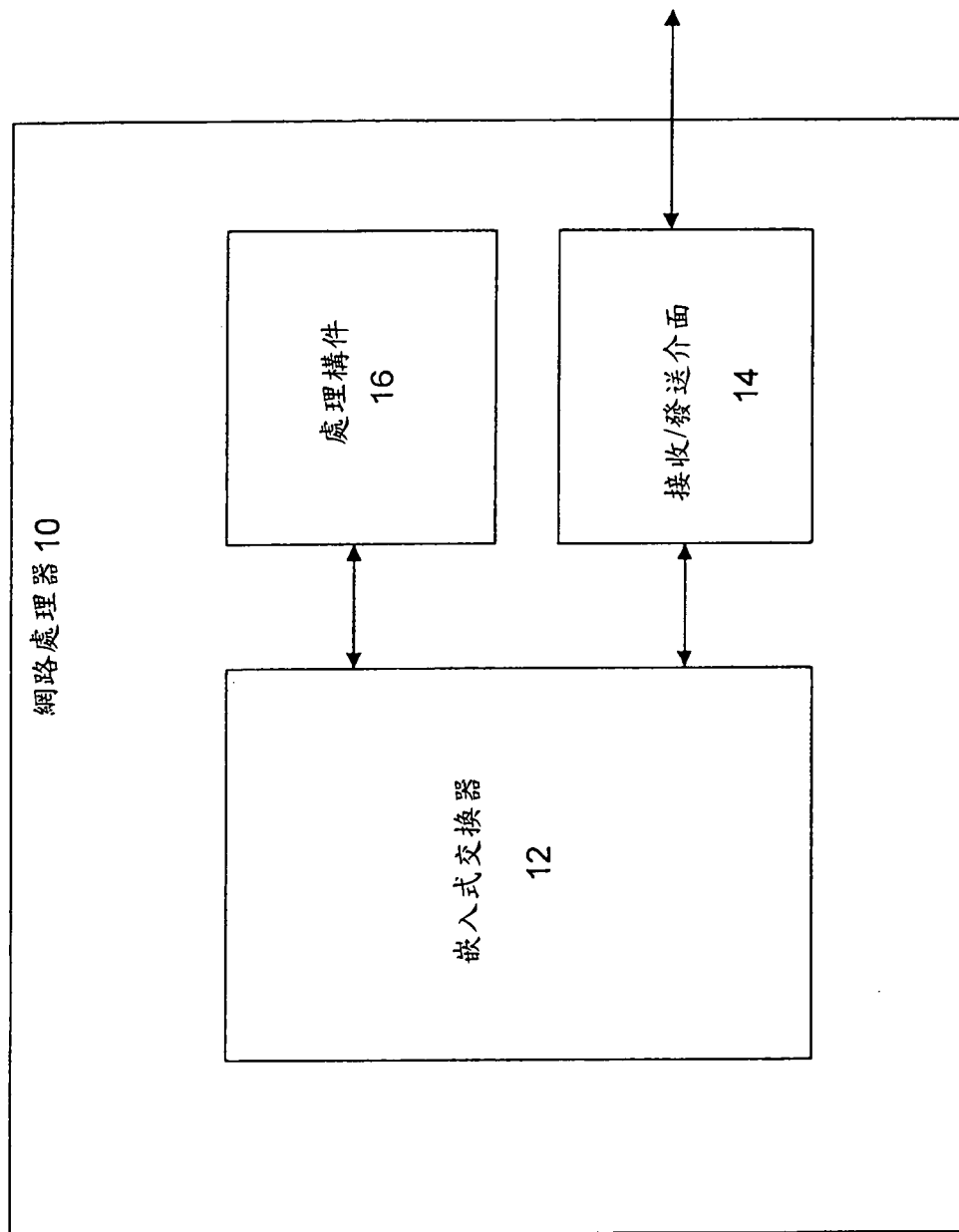


圖 2

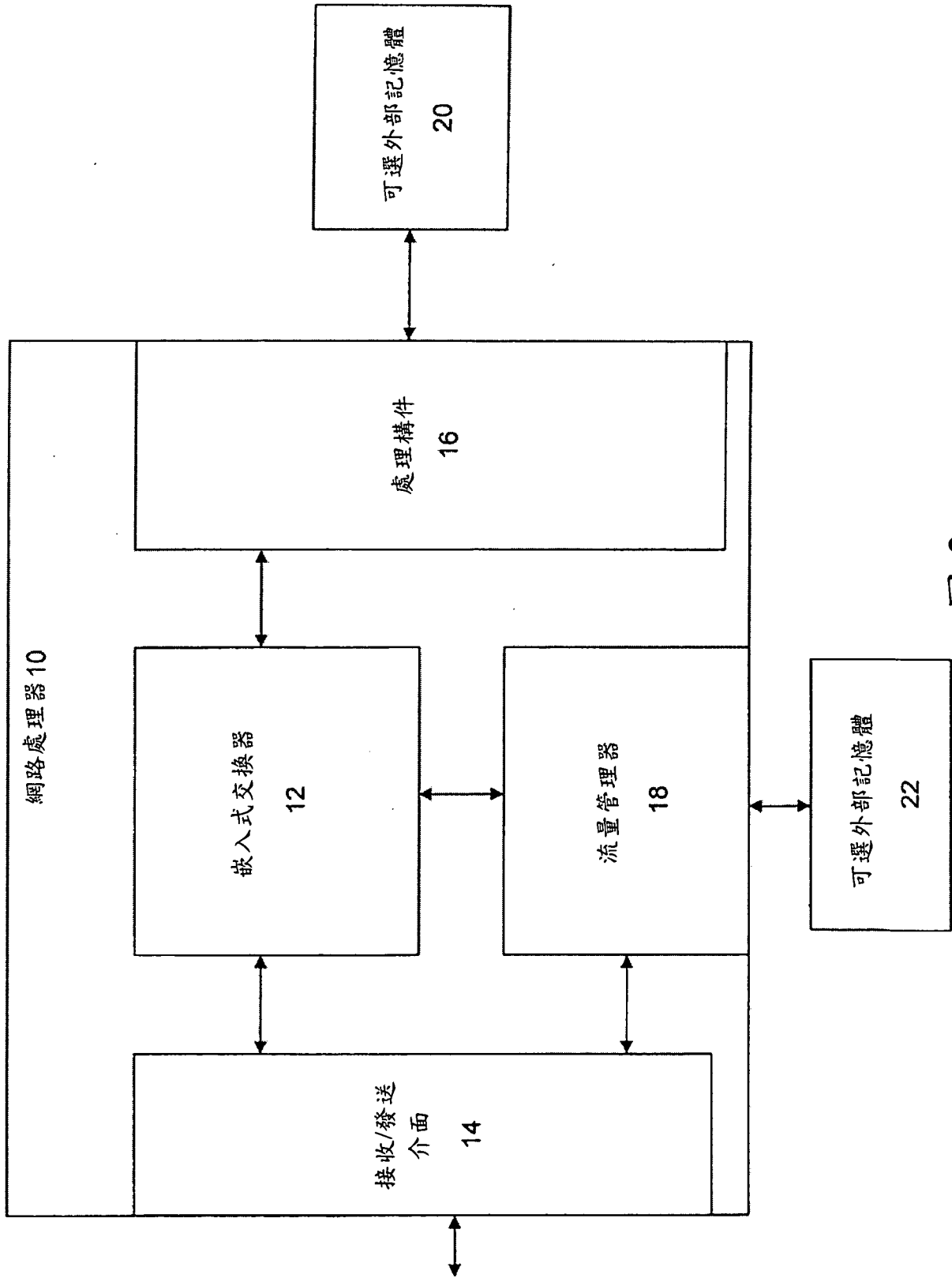


圖 3

10

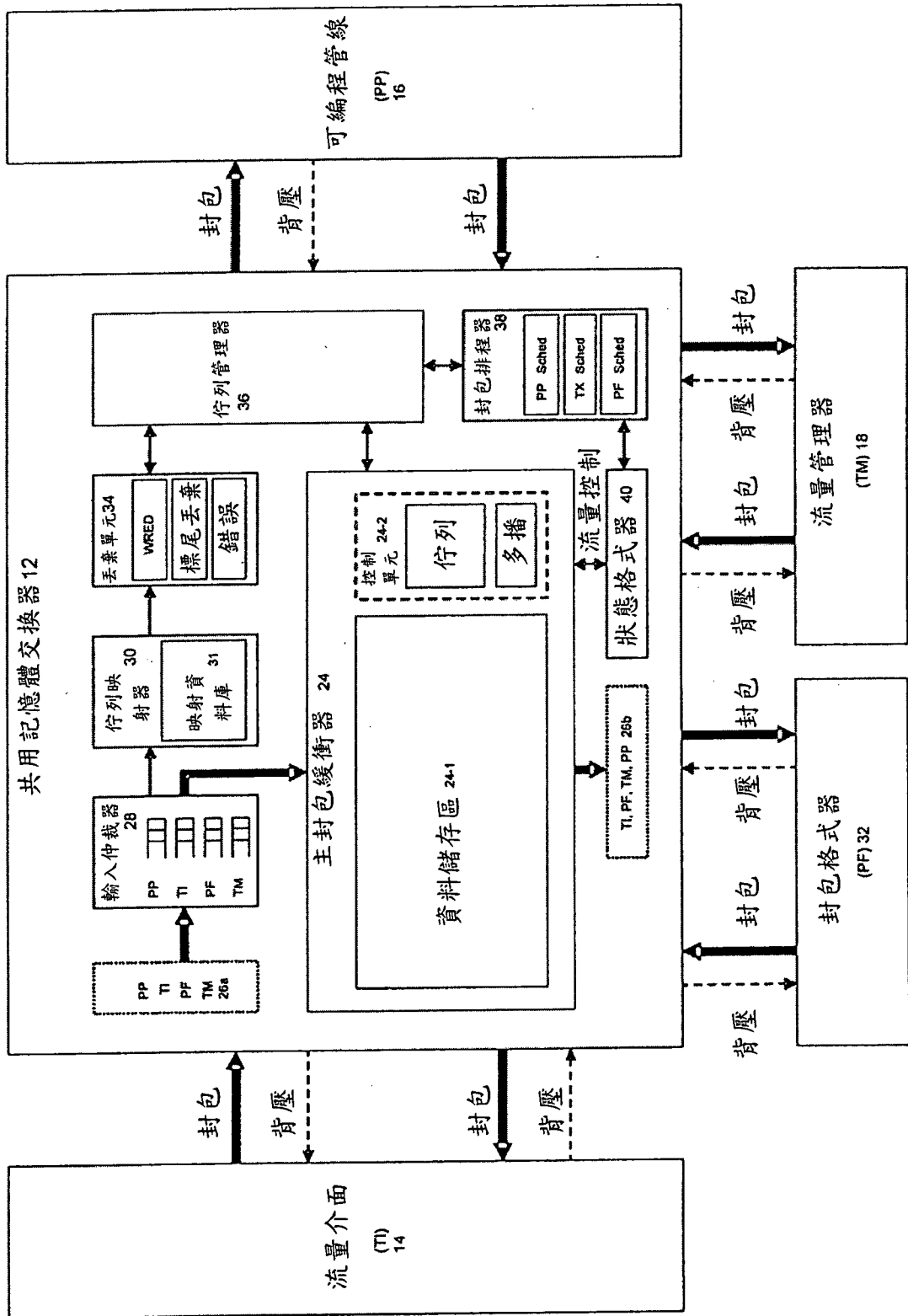


圖 4a

10

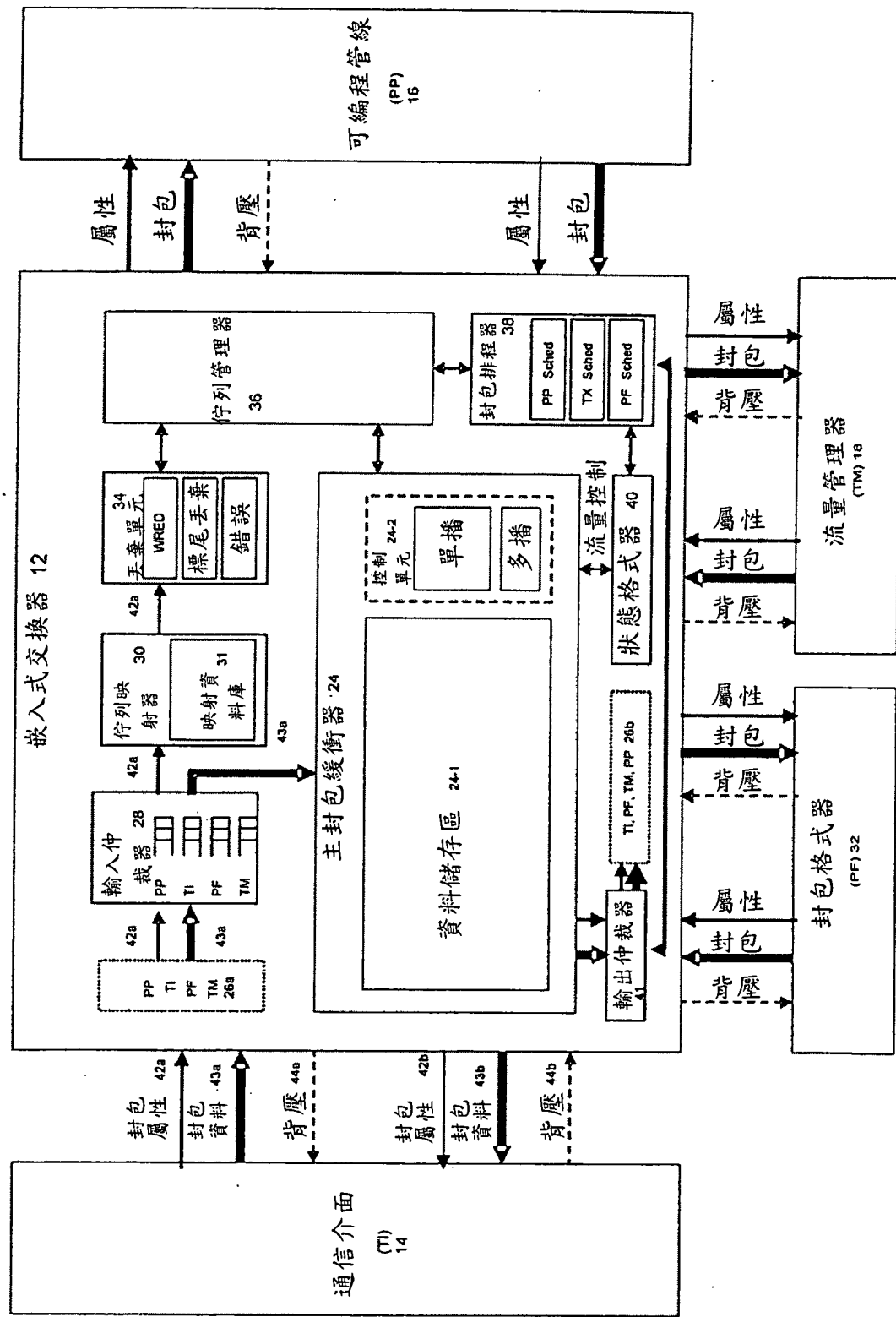


圖 4b

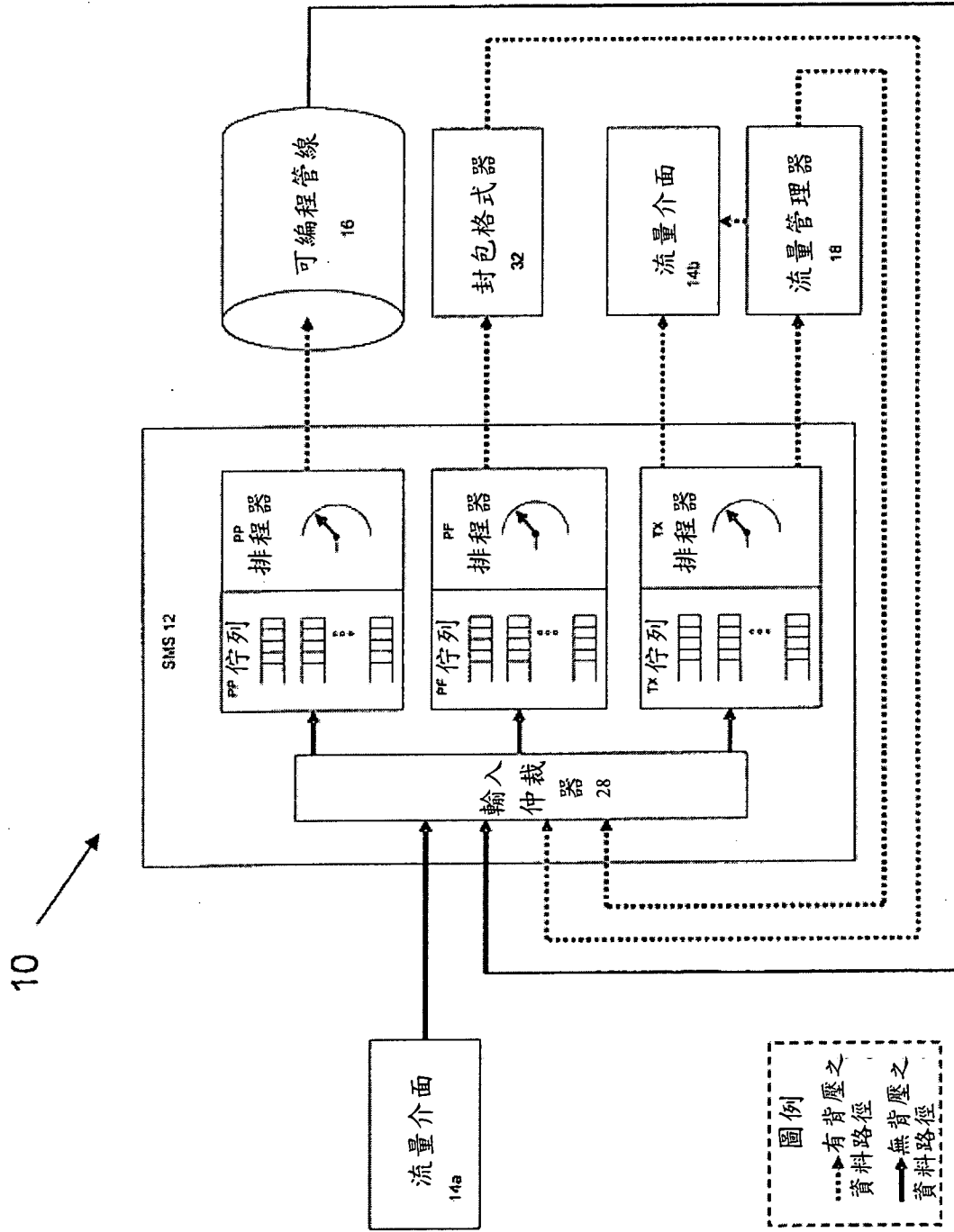


圖 5a

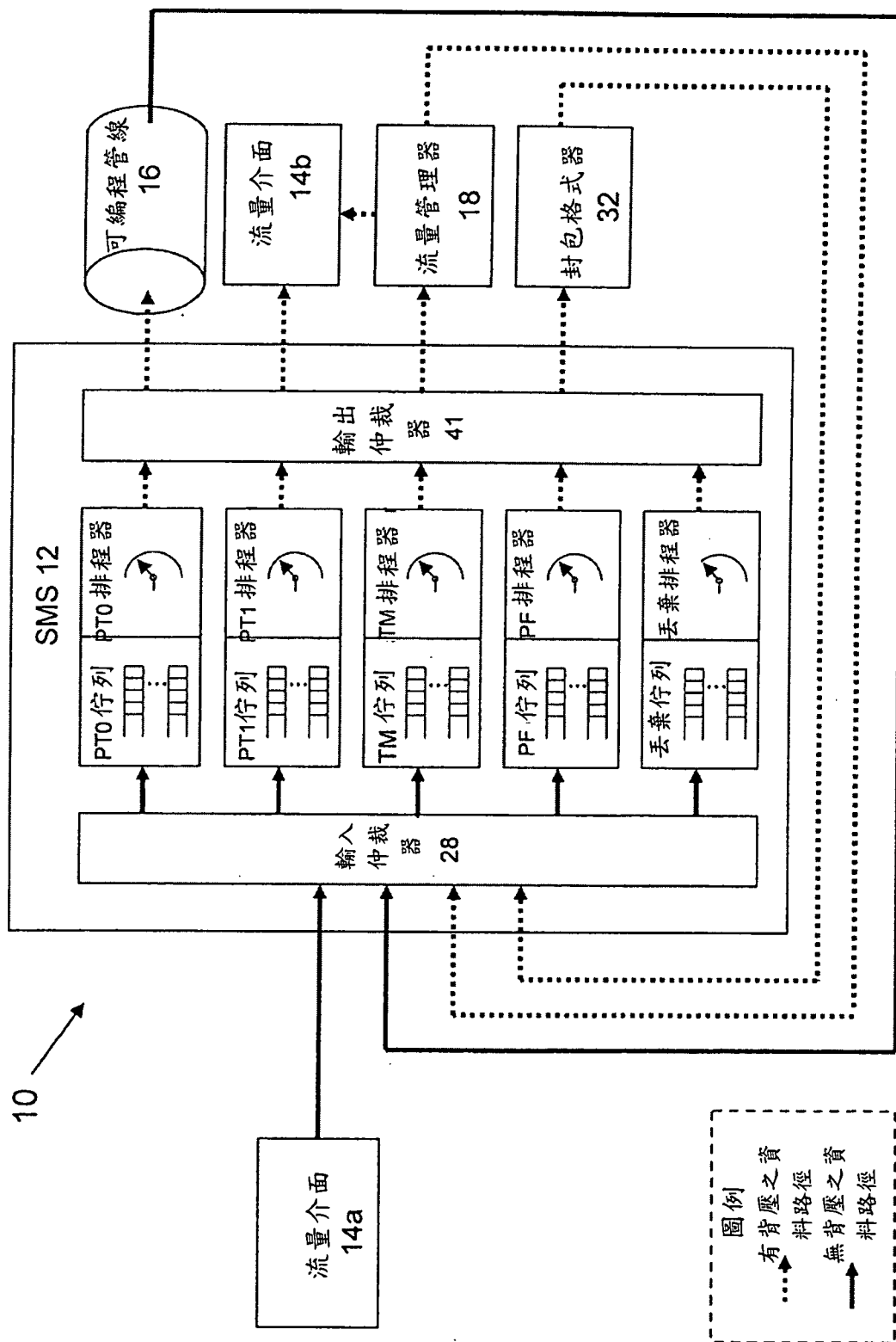


圖 5b