



(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 I513260 B

(45) 公告日：中華民國 104 (2015) 年 12 月 11 日

(21) 申請案號：102146229

(22) 申請日：中華民國 102 (2013) 年 12 月 13 日

(51) Int. Cl. : H04L29/06 (2006.01)

H04L12/729 (2013.01)

H04L12/28 (2006.01)

(71) 申請人：英業達股份有限公司 (中華民國) INVENTEC CORPORATION (TW)

臺北市士林區後港街 66 號

(72) 發明人：劉得彥 LIU, TE YEN (TW)

(74) 代理人：許世正

(56) 參考文獻：

TW 201246843A

CN 102595504A

審查人員：蔡鴻璟

申請專利範圍項數：6 項 圖式數：3 共 16 頁

(54) 名稱

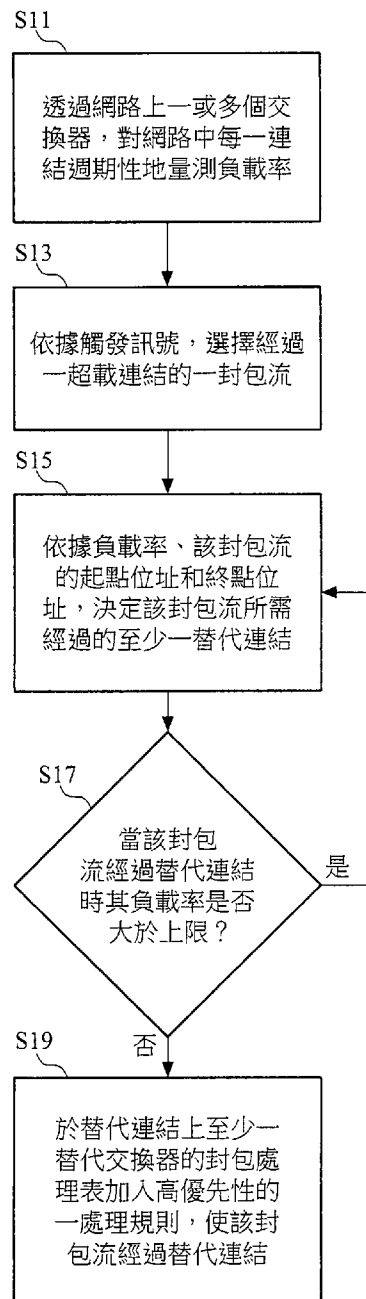
路由控制方法與裝置

ROUTING CONTROL METHOD AND APPARATUS

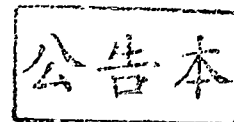
(57) 摘要

本發明提供一種用以平衡網路中多個連結的負載的路由控制方法。透過網路上一或多個交換器，對每一連結週期性地量測負載率。依據觸發訊號，選擇經過該些連結中一超載連結的一封包流，並依據該些連結的負載率、該封包流的起點位址和終點位址，決定該封包流所需經過的該些連結中至少一替代連結。選擇性地設定替代連結上至少一替代交換器的封包處理表，使該封包流經過替代連結。替代交換器係該或該些交換器其中之一。

Disclosed herein is a routing control method for balancing the load of links in a network. For each link, a load rate is measured through one or more switches on the network. Based on a trigger signal, a packet flow passing through an overloaded link is selected. At least one substitute link through which the packet flow is to pass is determined based on the load rates, a source address, and a destination address of the packet flow. A packet handling table of at least one substitute switch on the substitute link is configured so as to make the packet flow pass through the substitute link.



第2圖



發明摘要

※ 申請案號： 102146229

※ 申請日： 102. 12. 13

※IPC 分類： H04L 29/06 (2006.01)

【發明名稱】 路由控制方法與裝置

H04L 12/29 (2013.01)

H04L 12/28 (2006.01)

ROUTING CONTROL METHOD AND
APPARATUS

【中文】

本發明提供一種用以平衡網路中多個連結的負載的路由控制方法。透過網路上一或多個交換器，對每一連結週期性地量測負載率。依據觸發訊號，選擇經過該些連結中一超載連結的一封包流，並依據該些連結的負載率、該封包流的起點位址和終點位址，決定該封包流所需經過的該些連結中至少一替代連結。選擇性地設定替代連結上至少一替代交換器的封包處理表，使該封包流經過替代連結。替代交換器係該或該些交換器其中之一。

【英文】

Disclosed herein is a routing control method for balancing the load of links in a network. For each link, a load rate is measured through one or more switches on the network. Based on a trigger signal, a packet flow passing through an overloaded link is selected. At least one substitute link through which the packet flow is to pass is determined based on the load rates, a source address, and a destination address of the packet flow. A packet handling table of at least one substitute switch on the substitute link is configured so as to make the packet flow pass through the substitute link.

【代表圖】

【本案指定代表圖】：第 2 圖。

【本代表圖之符號簡單說明】：

無

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無

發明專利說明書

【發明名稱】 路由控制方法與裝置

ROUTING CONTROL METHOD AND
APPARATUS

【技術領域】

【0001】 本發明係關於網路路由（routing），特別係關於軟體定義網路（software-defined network，簡稱 SDN）上的路由控制方法與裝置。

【先前技術】

【0002】 在一網路中如何平衡各連結（link）的負載而極大化封包總通量一向是電腦網路技術的重要議題，尤其在網路位相（topology）益形複雜、資料量與日俱增的趨勢下，傳統等成本多路徑（equal-cost multi-path，簡稱 ECMP）盲目均分流量的做法已不合時宜。

【發明內容】

【0003】 鑒於上述問題，本發明旨在利用軟體定義網路中切分交換器（switch）的資料面（data plane）與控制面（control plane）而得以集中管理路由的特性，提供一種獨立運作不影響其他應用程式的路由控制方法與裝置。

【0004】 本發明提供一種用以平衡網路中多個連結的負載的路由控制方法，包含：透過網路上一或多個交換器，對每一連結週期性地量測負載率；依據觸發訊號，選擇經過該

些連結中一超載連結的一封包流 (flow)；依據該些連結的負載率、該封包流的起點位址和終點位址，決定該封包流所需經過的該些連結中至少一替代連結；以及選擇性地設定替代連結上至少一替代交換器的封包處理表，使該封包流經過替代連結。替代交換器係該或該些交換器其中之一。

【0005】 本發明提供一種路由控制裝置，用以平衡網路中多個連結的負載，包含通訊模組、處理模組與路由計算模組。通訊模組用以透過網路上一或多個交換器，對每一連結週期性地量測負載率，且用以產生觸發訊號。處理模組耦接通訊模組，用以依據觸發訊號，選擇經過該些連結中一超載連結的一封包流，且用以選擇性地設定該些連結中至少一替代連結上至少一替代交換器的封包處理表，使該封包流經過替代連結。替代交換器係該或該些交換器其中之一。路由計算模組耦接處理模組，用以依據該些連結的負載率、該封包流的起點位址和終點位址，決定該封包流所需經過的替代連結。

【0006】 綜上所述，本發明提供之路由控制方法與裝置可以軟體定義網路的控制器 (controller) 實作，當觸發訊號指示某連結超載時，設定替代交換器，重新導向超載連結上的某個封包流；此流程可被重複執行，直到該連結的負載不再那麼吃緊。

【0007】 以上關於本發明內容及以下關於實施方式之說明係用以示範與闡明本發明之精神與原理，並提供對本發明

之申請專利範圍更進一步之解釋。

【圖式簡單說明】

【0008】

第 1 圖係依據本發明一實施例路由控制裝置的高階方塊圖。

第 2 圖係依據本發明一實施例路由控制方法的流程圖。

第 3 圖係一網路的高階方塊圖。

【實施方式】

【0009】 以下在實施方式中敘述本發明之詳細特徵，其內容足以使任何熟習相關技藝者了解本發明之技術內容並據以實施，且依據本說明書所揭露之內容、申請專利範圍及圖式，任何熟習相關技藝者可輕易地理解本發明相關之目的及優點。以下實施例係進一步說明本發明之諸面向，但非以任何面向限制本發明之範疇。

【0010】 請參見第 1 圖。第 1 圖係依據本發明一實施例路由控制裝置的高階方塊圖。如第 1 圖所示，路由控制裝置 1 可以是軟體定義網路的控制器，包含處理模組 10、通訊模組 12 與路由計算模組 16。處理模組 10 分別耦接通訊模組 12 和路由計算模組 16。軟體定義網路可以 OpenFlow 協定實作，例如交換器開放控制面而能與控制器以訊息溝通。路由控制裝置 1 以通訊模組 12 所包含的網路介面卡 (network interface card，簡稱 NIC) 透過連結耦接網路上的一或多臺交換器，如

交換器 40。

【0011】 請參見第 2 圖。第 2 圖係依據本發明一實施例路由控制方法的流程圖。如第 2 圖所示，於步驟 S11 中，通訊模組 12 透過各交換器，對每一連結週期性地量測負載率。實務上，交換器需被開啓 sFlow 或 NetFlow 等網路量測功能，並將包括可見連結（如交換器的實體接線）的負載率的所得資訊回報給具監視功能的通訊模組 12。交換器可以即時地（real-time）也可以較低的頻率量測或回報。

【0012】 通訊模組 12 可設定有觸發條件，並於其發生時產生觸發訊號。舉例而言，觸發條件可以是「一連結的負載大於其頻寬的 90%」，則觸發訊號指示此連結超載。在其他實施例中，觸發條件中連結負載率可以有不同的上限。在一實施例中，觸發條件更可以考量超載是否為持續性的，如通訊模組 12 在「一連結的負載率大於 90% 超過 4 秒」才產生觸發訊號。

【0013】 於步驟 S13 中，處理模組 10 依據觸發訊號，選擇經過超載連結的一個封包流。將任何封包流移出超載連結顯然有助於減輕此連結的負載。實務上，處理模組 10 可向關聯於超載連結的交換器取得所有經過此連結的封包流的列表，並選擇流量最可觀的封包流以利步驟 S15 的執行。於步驟 S15 中，路由計算模組 16 依據選擇出的封包流的起點位址和終點位址，重新規劃遞送此封包流的最短路徑，著眼於避

開前述超載連結。規劃時可以以步驟 S11 中週期性量測的各連結的負載率為連結成本；重新規劃出的路徑會使此封包流經過相異於超載連結的至少一替代連結和其上至少一替代交換器。

【0014】 在一實施例中，處理模組 10 於步驟 S17 中判斷若將選擇出的封包流導向至替代連結，是否反而造成替代連結超載。顯然若此封包流會造成替代連結超載，則通訊模組 12 會於步驟 S13 中對替代連結產生觸發訊號，路由計算模組 16 又需執行步驟 S15 搬動封包流，形成無限循環。在一實施例中，若處理模組 10 於步驟 S17 中判斷替代連結不可行，則路由控制裝置 1 回到步驟 S15 重新對此封包流規劃路徑。在另一實施例中，路由控制裝置 1 可回到步驟 S13，選擇原超載連結上其他的（如流量次大的）的封包流並執行步驟 S15，或放棄並等待網路滿載的情況稍解再行路由控制。

【0015】 若處理模組 10 於步驟 S17 中判斷承擔選擇出的封包流不會使替代連結超載時，則可於步驟 S19 中設定替代交換器的封包處理表，使替代交換器接收此封包流的封包後能將之導向至替代連結。封包處理表可以是但不限於 OpenFlow 所定義的流表（flow table）。具體而言，於步驟 S19 中，處理模組 10 於替代交換器的封包處理表中加入針對此封包流（如以其起點位址和終點位址辨識）的一條處理規則。這條處理規則用於平衡負載，因此具緊迫性，應用上應比其

他來源的處理規則優先。又這條處理規則具臨時性，當原超載連結恢復正常或此封包流結束後即不再適用，因此處理模組 10 亦可在其中設定一適用時限。

【0016】 請配合第 1 圖和第 2 圖參見第 3 圖。第 3 圖係一網路的高階方塊圖。如第 3 圖所示，在網路 2 上有終端主機 (host) 30 和 32，以及交換器 40、42、44 與 46。終端主機 30 以連結 21 耦接交換器 40，並透過交換器 40 和網路 2 上其他可能存在的終端主機溝通。終端主機 32 以連結 27 耦接交換器 46。交換器 40 和 46 因耦接終端主機而常被稱為緣 (edge) 交換器，位於網路 2 內部的交換器 42 和 44 則是核 (core) 交換器。交換器 40 以連結 22 耦接交換器 42。交換器 40 以連結 23 耦接交換器 44。交換器 42 以連結 24 耦接交換器 46。交換器 42 以連結 25 耦接交換器 44。交換器 44 以連結 26 耦接交換器 46。概念上未繪於第 3 圖中的路由控制裝置 1 係分別耦接交換器 40、42、44 與 46，但實際上路由控制裝置 1 可以與網路 2 中某交換器整合，或是作為終端主機呈現的伺服器。

【0017】 假設於步驟 S11 和 S13 中連結 22 超載，而處理模組 10 選擇重新導向自終端主機 30 至終端主機 32 的一個封包流；此封包流現在的路徑是連結 21、22、24 和 27，亦即依序經過交換器 40、42 和 46。復假設於步驟 S15 和 S17 中路由控制裝置 1 決定此封包流的一條可用的替代路徑是連結 21、23、26 和 27 (連結 23 與 26 是替代連結)，因此處理模組 10

於步驟 S19 中在交換器 40 的封包處理表中加入「將此自終端主機 30 至終端主機 32 的封包流送往連結 23」的規則，並設定交換器 44「將此自終端主機 30 至終端主機 32 的封包流送往連結 26」，以此類推。

【0018】 一般而言，步驟 S15 中引用的起點和終點位址，如第 3 圖中用以辨識終端主機 30 和 32 者，指的是媒體存取控制（media access control，簡稱 MAC）位址。值得一提的是，本發明之路由控制方法亦適用於網路基礎建設的一般性維護。具體而言，在一實施例的步驟 S11 中，交換器可以即時量測網路中所有連結的負載率，但同時以較低頻率並行對網路內部連結（如連結 22 至 26）的另一種量測。對內部連結而言，通訊模組 12 仍在觸發條件符合時產生觸發訊號，唯此觸發條件可能略為寬鬆，如「負載率大於 80%」，相較於即時量測所有連結時的 90% 上限。路由計算模組 16 於步驟 S15 中可捨封包流兩終端主機而取兩緣交換器（如交換器 40 和 46）對網路內部的媒體存取控制位址來規劃替代路徑。舉例而言，當連結 24 超載時，任何封包流仍能沿連結 25 和 26 從交換器 42 被遞送到交換器 46（以及任何後續的目的地）。處理模組 10 於步驟 S19 中設定的針對內部連結的處理規則將不會與終端主機的位址有關。在一實施例中，這種規則的優先性低於即時量測、觸發造成者（但仍高於非出自本發明的處理規則），但適用時限較長。

【0019】 綜上所述，本發明提供之路由控制方法與裝置可以軟體定義網路的控制器實作，藉修改替代交換器的封包處理表，重新導向超載連結上的某個封包流；此流程可被重複執行，直到該連結的負載不再那麼吃緊。在一實施例中，只有重新導向的封包流不會使替代連結超載時處理模組才設定替代交換器，且新增的處理規則具高優先性和適用時限。在一實施例中，觸發訊號指示超載連結的負載率係大於某上限超過某預設時間。在一實施例中，網路基礎建設的一般性維護可與即時量測並行。

【0020】 本發明以實施例揭露如上，然前述實施例並非用以限定本發明。在本發明之範圍內不脫離本發明之精神而為之更動與潤飾均屬本發明之保護範圍。關於本發明所界定之保護範圍請參考所附之申請專利範圍。

【符號說明】

【0021】

1	路由控制裝置	2	網路
10	處理模組	21 - 27	連結
12	通訊模組	30、32	終端主機
16	路由計算模組	40 - 46	交換器
S11 - S19	步驟		

申請專利範圍

1. 一種路由控制方法，用以平衡一網路中多個連結的負載，該路由控制方法包含：

透過該網路上一或多個交換器，週期性地量測該些連結的多個負載率，每一該連結係分別對應該些負載率其中之一；

依據量測時產生的一觸發訊號，選擇一封包流，該封包流經過該些連結中一超載連結；

依據該些連結的該些負載率、該封包流的一起點位址和該封包流的一終點位址，決定該封包流所需經過的該些連結中至少一替代連結；以及

選擇性地設定該替代連結上至少一替代交換器的一封包處理表，使該封包流經過該替代連結，該替代交換器係該或該些交換器其中之一；

其中該觸發訊號指示該超載連結對應的該負載率大於一上限；

其中選擇性地設定該替代交換器的該封包處理表的步驟包含：

判斷當該封包流經過該替代連結時，該替代連結對應的該負載率是否大於該上限；

其中當該替代連結對應的該負載率不大於該上限時，於該封包處理表加入高優先性的一處理規則，且該處理規則具有一適用時限。

2. 如請求項 1 所述的路由控制方法，其中該觸發訊號更指示該超載連結對應的該負載率大於該上限超過一預設時間。
3. 如請求項 2 所述的路由控制方法，其中該起點位址係該或該些交換器其中之一的一媒體存取控制位址，該終點位址係該或該些交換器其中之一的一另一媒體存取控制位址。
4. 一種路由控制裝置，用以平衡一網路中多個連結的負載，該路由控制裝置包含：

一通訊模組，用以透過該網路上一或多個交換器，週期性地量測該些連結的多個負載率，每一該連結對應該些負載率其中之一，且量測時用以產生一觸發訊號；

一處理模組，耦接該通訊模組，用以依據該觸發訊號，選擇一封包流，該封包流經過該些連結中一超載連結，且用以選擇性地設定該些連結中至少一替代連結上至少一替代交換器的一封包處理表，使該封包流經過該替代連結，該替代交換器係該或該些交換器其中之一；以及

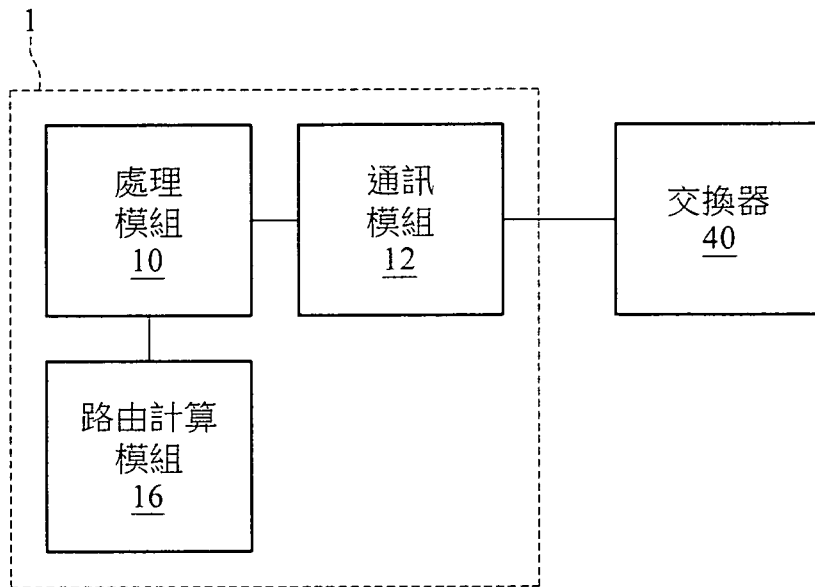
一路由計算模組，耦接該處理模組，用以依據該些負載率、該封包流的一起點位址和該封包流的一終點位址，決定該封包流所需經過的該替代連結；

其中該觸發訊號指示該超載連結對應的該負載率大於一上限；

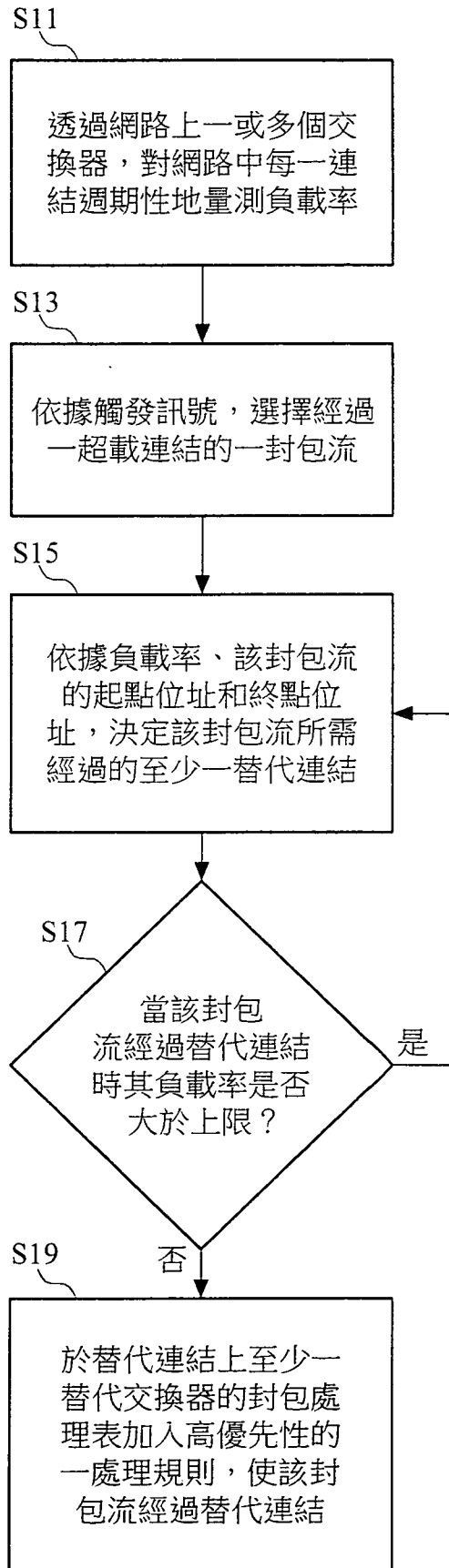
其中該處理模組選擇性地設定該替代交換器的該封包處理表包含判斷當該封包流經過該替代連結時，該替代連結對應的該負載率是否大於該上限，且當該替代連結對應的該負載率不大於該上限時，該處理模組於該封包處理表加入高優先性的一處理規則，且該處理規則具有一適用時限。

5. 如請求項 4 所述的路由控制裝置，其中該觸發訊號更指示該超載連結對應的該負載率大於該上限超過一預設時間。
6. 如請求項 4 所述的路由控制裝置，其中該起點位址係該或該些交換器其中之一的一媒體存取控制位址，該終點位址係該或該些交換器其中之一的一另一媒體存取控制位址。

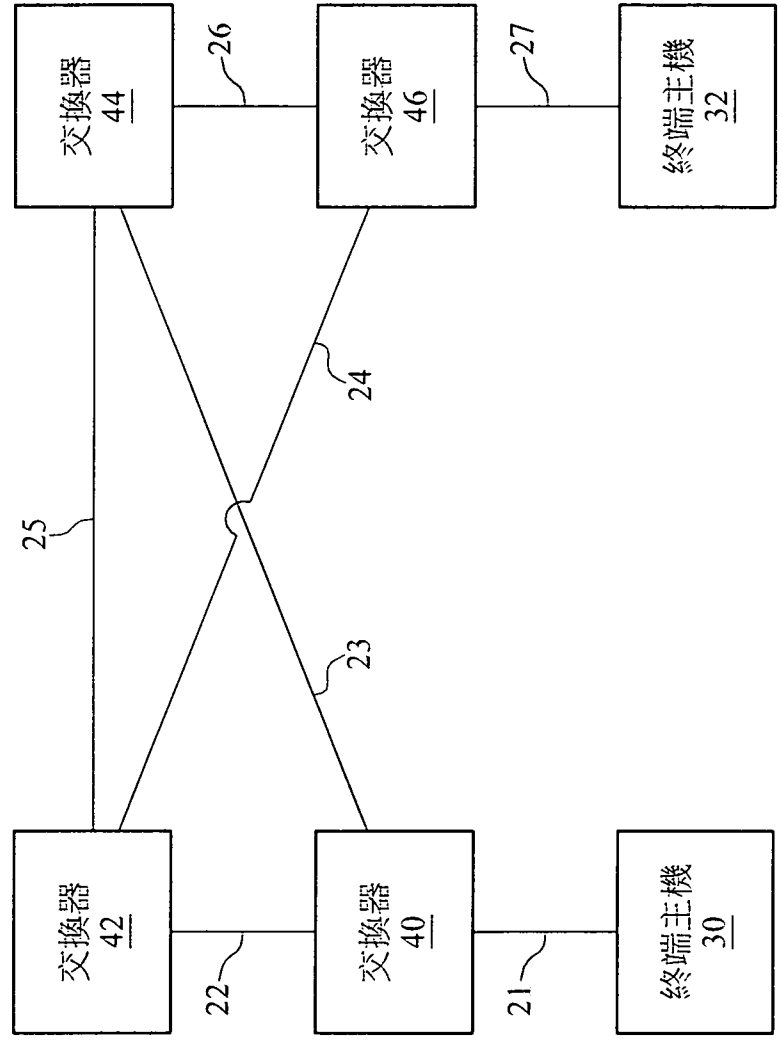
圖式



第1圖



第2圖



2

第3圖