



(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 I531908 B

(45) 公告日：中華民國 105 (2016) 年 05 月 01 日

(21) 申請案號：103114860

(22) 申請日：中華民國 103 (2014) 年 04 月 24 日

(51) Int. Cl. : **G06F15/167 (2006.01)****G06F9/455 (2006.01)****H04L12/947 (2013.01)**

(71) 申請人：財團法人國家實驗研究院(中華民國) (TW)

新竹市研發六路 7 號

(72) 發明人：李慧蘭(TW)；劉德隆(TW)；胡仁維(TW)；陳敏(TW)

(74) 代理人：蔡秀玫

(56) 參考文獻：

TW 201329398A

CN 103685250A

US 2009/0070760A1

WO 2012/119390A1

審查人員：梁中明

申請專利範圍項數：9 項 圖式數：8 共 24 頁

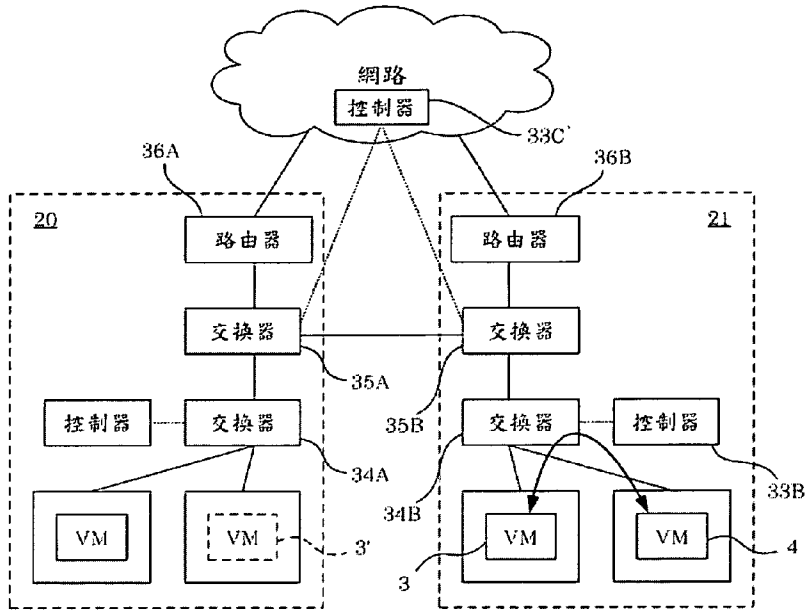
(54) 名稱

以軟體定義網路(SDN)支援虛擬機器遷移之方法

(57) 摘要

本發明係關於一種虛擬機器遷移之方法，其利用軟體定義網路技術之協定，讓虛擬機器在經過跨網域遷移時，可快速通知本地端的控制器，預先派送前述虛擬機器的資訊至交換器，以在無需更改網路組態設定的優點之下，讓遷移後的虛擬機器能夠持續提供服務，達到最佳路由，並有效改善三角路由問題。

指定代表圖：



第5圖

符號簡單說明：

20 . . . 第一網域

21 . . . 第二網域

3 . . . 虛擬機器

3' . . . 虛擬機器

33B . . . 第二控制器

33C . . . 中央控制器

34A . . . 第一存取層交換器

34B . . . 第二存取層交換器

35A . . . 第一匯聚層交換器

35B . . . 第二匯聚層交換器

36A . . . 第一路由器

36B . . . 第二路由器

4 . . . 虛擬機器



105年1月25日修正  
p1~p2

105年 01月 25日 修正替換頁

申請日: 103.4.24

IPC分類: G06F 15/167 (2006.01)

G06F 9/455 (2006.01)

H04L 12/947 (2013.01)

### 【發明摘要】

【中文發明名稱】 以軟體定義網路(SDN)支援虛擬機器遷移之方法

【中文】

公告本

本發明係關於一種虛擬機器遷移之方法，其利用軟體定義網路技術之協定，讓虛擬機器在經過跨網域遷移時，可快速通知本地端的控制器，預先派送前述虛擬機器的資訊至交換器，以在無需更改網路組態設定的優點之下，讓遷移後的虛擬機器能夠持續提供服務，達到最佳路由，並有效改善三角路由問題。

**【指定代表圖】 第5圖****【代表圖之符號簡單說明】**

20 第一網域

21 第二網域

3 虛擬機器

3' 虛擬機器

33B 第二控制器

33C 中央控制器

34A 第一存取層交換器

34B 第二存取層交換器

35A 第一匯聚層交換器

35B 第二匯聚層交換器

36A 第一路由器

36B 第二路由器

4 虛擬機器

**【特徵化學式】**

無

## 【發明說明書】

【中文發明名稱】 以軟體定義網路(SDN)支援虛擬機器遷移之方法

【技術領域】

【0001】 本發明係關於一種虛擬機器遷移之方法，尤指一種在不中斷對外服務的條件下完成異地跨網域遷移之一種虛擬機器遷移之方法。

【先前技術】

【0002】 基於技術的演進及人們對於網路需求的增加，諸如雲端服務、虛擬化技術、巨量資料等網路應用隨之而生。這些網路應用需要大量的網路資源、良好的頻寬管理及安全政策。這些需求在現有的分散式網路架構下不容易實現。而隨著網路拓撲的增大，網路也變得難以有效管理。為解決這些重要問題，軟體定義網路（Software Defined Networking, SDN）的觀念及技術因此在近幾年被提出且受到廣泛的重視及研究。

【0003】 隨著雲端運算的快速發展，雲端應用服務的使用需求激增，為提高雲端服務的可靠度，因此對虛擬機器遷移的需求也因應而生。尤其是為確保雲端服務維持不中斷，首當其衝要面對的是虛擬機器跨不同網域遷移時的網路問題。

【0004】 請參考第一圖，其係為先前技術當中的封包路徑圖；如圖所示，第一網域10與第二網域11分別屬於不同之網域，第二虛擬機器（VM2）12原先運行於第一網域10，而於其遷移至第二網域11後，維持原來IP網路組態，由於第二虛擬機器12與第三虛擬機器13分別屬於不同網域之IP位址，因此第三虛擬機器13與遷移後之第二

虛擬機器12通訊時，封包會經由第三虛擬機器13之閘道路由器而進入網際網路，經由網際網路上路由表查找，然後傳送至第一網域10的入口路由器，再往介接其下之Layer2交換器轉送。此架構中的匯聚層交換器(aggregation layer switch)14A、14B之間有以隧道15(tunnel)連接，使第一網域10的存取層交換器(access layer switch)16A、匯聚層交換器14A，以及第二網域11的存取層交換器16B、匯聚層交換器14B均屬於同一個Layer2廣播網域，因此封包一路由第一網域10之匯聚層交換器14A經過隧道15，流經第二網域11之匯聚層交換器14B和存取層交換器16B，最後傳送到目的端。

【0005】 這樣的繞路就是三角路由(Triangle Routing)，這種非對稱性的路徑對於通訊延遲與網路資源浪費帶來了顯著與額外的負擔。尤其資料中心的流量特徵以東西向流量居多，更突顯其封包傳送延遲造成的效能不彰。

【0006】 中國專利申請號CN 201210567450.X曾揭露了一種「虛擬機器網路控制策略的遷移處理方法及系統」，利用虛擬機器遷移前後介接於交換器埠號的異動，來獲知虛擬機器已由一方遷移至另一方；其通過對虛擬機器遷移消息的監控以及對虛擬交換機資料轉發規則的修改，使虛擬機器在不同的物理主機上遷移時，能夠保持其對應的網路策略不變，從而為虛擬機器提供連續一致的網路控制服務。

【0007】 美國專利公開號US 20130151661也揭露了一種虛擬機器遷移之技術，其利用網路位址轉譯(Network Address Translation, NAT)之技術來處理虛擬機器之遷移。其係在虛擬機器自第一主機遷移

至第二主機時，自動發送網路位址轉譯的訊息到第二主機，以透過在第二主機上運行的轉發管理單元得以基於該網路位址轉譯的訊息來處理所遷移入的虛擬機器。

#### 【發明內容】

【0008】 本發明之主要目的，係提供一種以軟體定義網路支援虛擬機器遷移之方法，其在虛擬機器跨網域遷移時，透過軟體定義網路之技術，能快速地通知網路控制器，可預先派送該虛擬機器的轉送路由表到交換器，此方法可於虛擬機器遷移後，無須更改網路組態之設定而仍持續提供服務，達到最佳路由，並有效改善三角路由問題。

【0009】 本發明之另一目的，係提供一種以軟體定義網路支援虛擬機器遷移之方法，其所運用之技術係為軟體定義網路，以其架構中分離之控制器和交換器之溝通，達到最佳路由。任何可透過開放式或封閉式介面而修改交換器路由表之方式，皆屬於本案所指之軟體定義網路之範疇。

【0010】 本發明之再一目的，係提供一種以軟體定義網路支援虛擬機器遷移之方法，可利用於跨多網域之雲端資料中心，以及進一步作為虛擬機器遷移通知代理程式、支援跨網域虛擬機器遷移之網路控制器應用程式，或支援跨網域虛擬機器遷移之SDN網路交換器等產品之技術基礎，存在商業價值。

【0011】 為了達到上述之目的，本發明揭示了一種以軟體定義網路支援虛擬機器遷移之方法，其於將原運行於一第一網域之一第一主機之一虛擬機器遷移至一第二網域時，進行步驟：該虛擬機器或該第

一主機提供一第一遷移通知至該第一網域之一第一控制器，且該虛擬機器或該第二主機亦提供一第二遷移通知至該第二網域之一第二控制器，該第一遷移通知以及該第二遷移通知分別至少包含該虛擬機器之IP位址、MAC位址以及閘道位址；以及該第二控制器寫入一路由規則(Flow Entry)至該第二網域之一交換器，使目的為該虛擬機器且需經該第二網域之一路由器之封包被轉發至該虛擬機器所遷移至該第二網域之一第二主機，以及使該虛擬機器發向該第二網域之封包被轉發至該第二控制器處理。依此方法之步驟，即可在不中斷對外服務的條件下，有效完成虛擬機器之異地跨網域遷移。

#### 【圖式簡單說明】

#### 【0012】

第1圖：其係為先前技術中，虛擬機器在遷移後存在三角路由之問題，其並不利於封包傳輸；

第2圖：其係為本發明一較佳實施例之步驟流程圖；

第3圖：其係為本發明一較佳實施例中，虛擬機器所遷移至之第二網域之元件架構示意圖；

第4圖：其係為本發明一較佳實施例中，第二控制器獲得遷移通知而寫入路由規則於第二存取層交換器之路由表之示意圖；

第5圖：其係為本發明一較佳實施例中，位於第二網域之另一個虛擬機器與遷移至第二網域之虛擬機器之間有通訊產生時，封包傳輸路徑與元件架構示意圖；



第6圖：其係為本發明一較佳實施例中，依據路由規則1而傳輸封包之路徑示意圖；

第7圖：其係為本發明一較佳實施例中，依據路由規則3、4而傳輸封包之路徑示意圖；以及

第8圖：其係為本發明一較佳實施例中，遷移至第二網域之虛擬機器要與第一網域之另一虛擬機器之間有通訊產生時，封包傳輸路徑與元件架構示意圖。

### 【實施方式】

【0013】 為使本發明之特徵及所達成之功效有更進一步之瞭解與認識，謹佐以較佳之實施例及配合詳細之說明，說明如後：

【0014】 首先，請參考第2圖，其揭示了本發明一較佳實施例之方法流程，其在將原本運行於一第一網域之一第一主機上之一虛擬機器遷移至一第二網域時，係包含步驟：

步驟S1：由第一主機代理程式或該虛擬機器發送一第一遷移通知至該第一網域之一第一控制器，該第一遷移通知包含該虛擬機器之IP位址、MAC位址以及閘道位址；

步驟S2：該虛擬機器遷移至該第二網域之該第二主機後，第二主機代理程式或該虛擬機器發送一第二遷移通知至該第二網域之一第二控制器，該第二遷移通知包含該虛擬機器之IP位址、MAC位址以及閘道位址；

步驟S3：該第二控制器於獲知資訊表更新時，寫入一路由規則至該第二網域之一交換器，使目的為該虛擬機器且需經該第二網域

之一路由器之封包被轉發至該虛擬機器所遷移至該第二網域之所介接埠號，以及使該虛擬機器為來源端發送之封包被轉發至該控制器處理。

其中，上述之第一主機代理程式以及上述之第二主機代理程式是分別裝載於該第一主機以及該第二主機。

【0015】請配合參考第3圖之元件架構示意圖；如圖所示，前述較佳實施例在運作時，元件架構係分散於第一網域20以及第二網域21，元件則包含了第一主機22A、第二主機22B、第一控制器23A、第二控制器23B、中央控制器23C、第一存取層交換器24A、第一匯聚層交換器25A、第二存取層交換器24B、第二匯聚層交換器25B、第一路由器26A以及第二路由器26B；其中第一主機22A、第一控制器23A、第一存取層交換器24A、第一匯聚層交換器25A、第一路由器26A係位於第一網域20，而第二主機22B、第二控制器23B、第二存取層交換器24B、第二匯聚層交換器25B以及第二路由器26B則是位於第二網域21。中央控制器23C位於網路可存取處，亦可與區域控制器安裝於同一台主機。為縮短傳輸延遲第一控制器23A和第二控制器23B為區域控制器。

【0016】為了有效改善三角路由的問題，此較佳實施例中，原本運行於第一主機22A之虛擬機器3' 在遷移至第二主機22B而為虛擬機器3時，基於兩者主機網域不相同，其具有不同的IP位址網域，但虛擬機器3的IP位址於遷移後未異動。

【0017】虛擬機器遷移時，第一主機22A的代理程式會提供第一遷移通知至第一控制器23A，以告知其「虛擬機器3' 將遷移出去」，也可

由虛擬機器提供該第一遷移通知至第一控制器23A；而遷移之後第二主機22B的代理程式則會提供第二遷移通知至第二控制器23B，以告知其「虛擬機器3將進駐」，也可由虛擬機器提供該第二遷移通知至第二控制器23B。上述之第一遷移通知以及第二遷移通知包含了該虛擬機器3之IP位址、MAC位址以及閘道位址等資訊；接著，第二控制器23B自動寫入一路由規則(flow entry)至第二存取層交換器24B，此第二存取層交換器24B係為一種軟體定義網路(SDN)交換器。前述之路由規則係於虛擬機器在遷移時自動寫入於第二存取層交換器24B之路由表(flow table)，使第二網域21得以具備縮短封包傳輸的條件，上述路由規則至少包含了一路由規則1(flow entry1)、一路由規則2(flow entry2)以及一路由規則3(flow entry3)，可參考第4圖。本發明當中所傳輸之封包的格式包含但不限於來源端MAC位址(S-MAC)、目的端MAC位址(D-MAC)、來源端IP位址(SIP)、目的端IP位址(DIP)、TCP來源端埠號(TCP sport)以及TCP目的端埠號(TCP dport)。

【0018】 依據上述步驟而將虛擬機器完成遷移後，關於封包傳輸的細節實施例，可請參考第5圖；當位於第二網域21之另一個虛擬機器4與前述虛擬機器3之間有通訊產生時，其中一種情況為第二虛擬機器4要發送封包至虛擬機器3，對於這種情況，可一併參考第6圖，由於第二存取層交換器34B路由表(flow table entry)之路由規則1(flow entry1)中已有目的端位址，也就是虛擬機器3的IP位址，前述之路由規則1(flow entry1)係指定虛擬機器3為封包發送目的，因此其會指示原來必須送給第二網域21的第二路由器36B的封包直接轉送給虛擬機器3的埠號(port)，使得虛擬機器3

能以較短路徑收到第二虛擬機器4之封包而不需要再繞經第一網域20。

【0019】 虛擬機器3為屬於第一網域之IP位址，因此遷移至第二網域後，當其對外通訊時，根據TCP/IP標準，該虛擬機需要獲知所屬網域(即第一網域)之閘道位置。此時虛擬機器3對外通訊時，會先送出ARP請求第一網域之閘道MAC位址，並產生以太網路類別為0806(ether-type:0806)之廣播封包，第二控制器33B會觸發其轉送模組，使用路由規則2(flow entry2)，以快速幫助虛擬機器3取得第一網域20的第一路由器36A之閘道MAC位址。虛擬機器3取得閘道位址後，方始與其不同網域之其他機器開始通訊。

【0020】 待虛擬機器3取得第一路由器36A之閘道MAC位址之後，此時虛擬機器3開始送出的封包目的IP位址為同樣位於第二網域21之第二虛擬機器4的位址，符合了路由規則3(flow entry3)。此flow entry3係指定虛擬機器3所發送之封包轉向送至第二控制器33B，也就是當封包的來源IP位址是屬於虛擬機器3時，此封包就會轉向給第二網域21的第二控制器33B處理，如第7圖所示之虛線段。

【0021】 第二控制器33B的資訊表記錄介接其下所有的虛擬機器資訊，包含IP位址、MAC位址和閘道位址和所介接之埠號。請參考第7圖所示之鏈線段，因此當第二控制器33B收到封包後，便擷取封包之目的IP位址(DIP)當成關鍵搜尋字串，進到第二控制器33B的的資訊表進行搜尋，於是獲知了目的端的MAC位址和與第二存取層交換器34B介接之埠號。控制器將所獲知之資訊產生一筆路由規則4(flow entry4)到第二存取層交換器34B的路由表。

- 【0022】 最後，請參考第7圖所示之實線段，傳輸路徑即依據前述路由規則4(flow entry4)之定義，將原應送到第一網域20閘道的封包，直接修改為第二虛擬機器4之目的MAC位址並轉向至其所介接之埠號。
- 【0023】 經過上述流程，虛擬機器3所發送之第二虛擬機器4之第一個封包被轉送至第二控制器33B處理後，之後的封包在到達第二存取層交換器34B時，由於可成功比對路由規則4(flow entry4)，因此就不需要再將封包轉送至第二控制器33B處理；換言之，優先順序較低的路由規則3(flow entry3)被忽略。
- 【0024】 進一步而言，第二存取層交換器34B的路由表(flow table entry)當中會有四筆路由規則，分別為前述的路由規則1(flow entry1)、路由規則2(flow entry2)、路由規則3(flow entry3)和路由規則4(flow entry4)，其中的路由規則4(flow entry4)之優先順序係高於該路由規則3(flow entry3)。虛擬機器3派送封包給第二網域之虛擬機器4時，最後只會比對路由規則4(flow entry4)；反向由第二網域之虛擬機器4派送封包給虛擬機器3的路徑最後只會比對路由規則1(flow entry1)，因此有效改善了三角路由的問題。
- 【0025】 若遷移後之虛擬機器3要與原第一網域20的另一虛擬機器5通訊，則請參考第8圖，來自虛擬機器5之封包於其傳輸路徑會依序經由Layer2中的第一存取層交換器34A、第一匯聚層交換器35A、隧道52、第二匯聚層交換器35B以及第二存取層交換器34B。在此情境下，中央控制器33C、區域控制器33A和33B載入轉送模組之後，自動產生路由規則至交換器便可自動通訊。另外，當數量龐大的

虛擬機器需要遷移時，則路由規則可事先自動安裝於SDN交換器與前一實施例係為相同，避免Layer2廣播網域被拖垮效能。中央控制器負責維持隧道建置資訊與管理各網域之匯聚層交換器。

【0026】 若前述第一網域、第二網域以外之網域的機器要與前述遷移後的虛擬機器通訊，則基於該虛擬機器的IP 位址未異動，可經網際網路上之傳統路由表(routing table)上查找出最佳路徑是經第一網域的邊界路由器走，再經過如前一實施例當中的第一匯聚層交換器、隧道、第二匯聚層交換器以及第二存取層交換器而到達目的端，其路由規則的設計與前一實施例係為相同。

【0027】 綜上所述，本發明詳細揭示了一種以軟體定義網路支援虛擬機器遷移之方法，其利用軟體定義網路技術之協定，讓虛擬機器在經過跨網域遷移時，可快速通知本地端的控制器，預先派送前述虛擬機器的資訊至交換器，以在無需更改網路組態設定的優點之下，讓遷移後的虛擬機器能夠持續提供服務，達到最佳路由，並有效改善三角路由問題。利用本發明，可解決雲端資料中心在大量使用虛擬化技術時，同時要面臨實體伺服器超載、異地備援、資源分配和負載平衡等問題而必須進行虛擬機器遷移。故總結而言，本發明確實提供了一種充分展現實用與經濟價值之虛擬機器遷移之方法。

【0028】 惟以上所述者，僅為本發明之較佳實施例而已，並非用來限定本發明實施之範圍，舉凡依本發明申請專利範圍所述之形狀、構造、特徵及精神所為之均等變化與修飾，均應包括於本發明之申請專利範圍內。

## 【符號說明】

- 10 第一網域
- 11 第二網域
- 12 第二虛擬機器
- 13 第三虛擬機器
- 14A、14B 匯聚層交換器
- 15 隧道
- 16A、16B 存取層交換器
- 20 第一網域
- 21 第二網域
- 22A 第一主機
- 22B 第二主機
- 23A 第一控制器
- 23B 第二控制器
- 23C 中央控制器
- 24A 第一存取層交換器
- 24B 第二存取層交換器
- 25A 第一匯聚層交換器
- 25B 第二匯聚層交換器

26A 第一路由器

26B 第二路由器

3 虛擬機器

3' 虛擬機器

33B 第二控制器

33C 中央控制器

34A 第一存取層交換器

34B 第二存取層交換器

35A 第一匯聚層交換器

35B 第一匯聚層交換器

36A 第一路由器

36B 第二路由器

4 虛擬機器

5 虛擬機器

52 隧道

S1~S3 步驟



## 【發明申請專利範圍】

【第1項】 一種以軟體定義網路支援虛擬機器遷移之方法，其於將原運行於一第一網域之一第一主機之一虛擬機器遷移至一第二網域時，進行步驟：

該虛擬機器或該第一主機提供一第一遷移通知至該第一網域之一第一控制器，且該虛擬機器或該第二主機亦提供一第二遷移通知至該第二網域之一第二控制器，該第一遷移通知以及該第二遷移通知分別至少包含該虛擬機器之IP位址、MAC位址以及閘道位址；以及

該第二控制器寫入一路由規則至該第二網域之一交換器，使目的為該虛擬機器且需經該第二網域之一路由器之封包被轉發至該虛擬機器所遷移至該第二網域之所介接埠號，以及使該虛擬機器為來源端發送之封包被轉發至該第二控制器處理。

【第2項】 如申請專利範圍第1項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該第二網域之一路由器之封包及該虛擬機器為來源端發送之封包之格式包含來源端MAC位址(S-MAC)、目的端MAC位址(D-MAC)、來源端IP位址(SIP)、目的端IP位址(DIP)、TCP來源埠號(TCP sport)以及TCP目的埠號(TCP dport)。

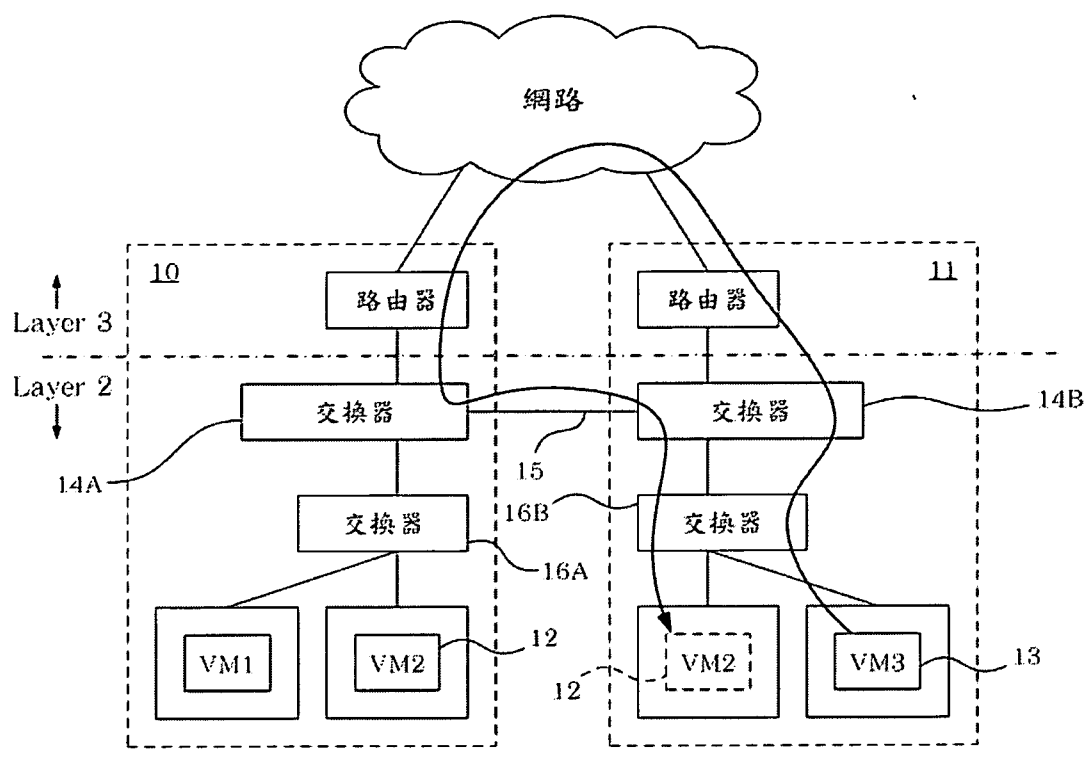
【第3項】 如申請專利範圍第1項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該交換器係為軟體定義網路(SDN)交換器。

【第4項】 如申請專利範圍第1項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該路由規則(flow entry)係寫入於該交換器之路由

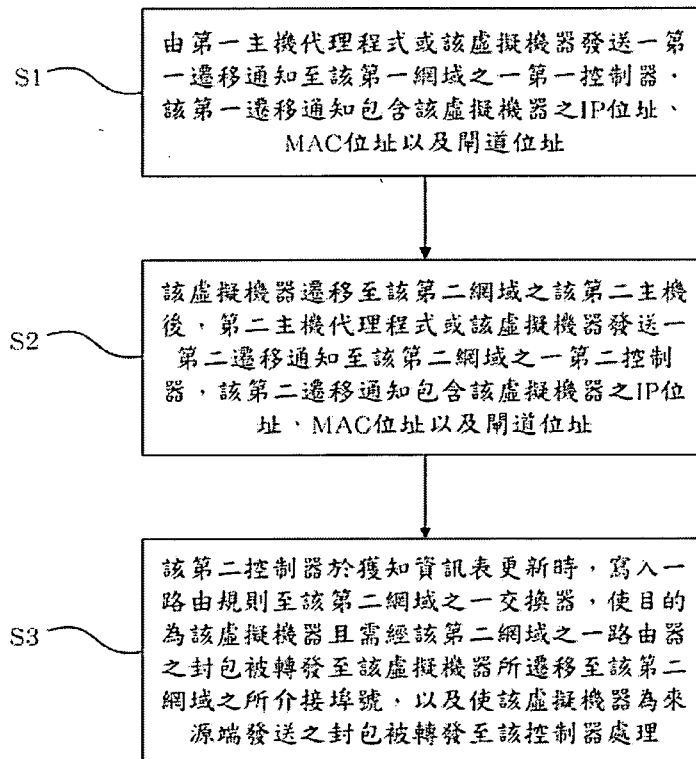
表(flow table)。

- 【第5項】 如申請專利範圍第1項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該路由規則係包含一路由規則1(flow entry1)以及一路由規則3(flow entry3)，該路由規則1(flow entry1)係指定該虛擬機器為封包發送目的，該路由規則3(flow entry3)係指定該虛擬機器所發送之封包轉向送至該控制器。
- 【第6項】 如申請專利範圍第5項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該路由規則更包含一路由規則2(flow entry2)，該路由規則2(flow entry2)係用以幫助該虛擬機器取得該第一網域之路由器之閘道MAC位址。
- 【第7項】 如申請專利範圍第5項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該路由規則更包含一路由規則4(flow entry4)，該路由規則4(flow entry4)係用以將原應送到該第一網域之閘道的封包，直接修改為目的端之一第二虛擬機器之目的MAC位址且轉送至所介接之埠號。
- 【第8項】 如申請專利範圍第7項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該路由規則4(flow entry4)之優先順序係高於該路由規則3(flow entry3)。
- 【第9項】 如申請專利範圍第1項所述之以軟體定義網路支援虛擬機器遷移之方法，其中該交換器係為存取層交換器。

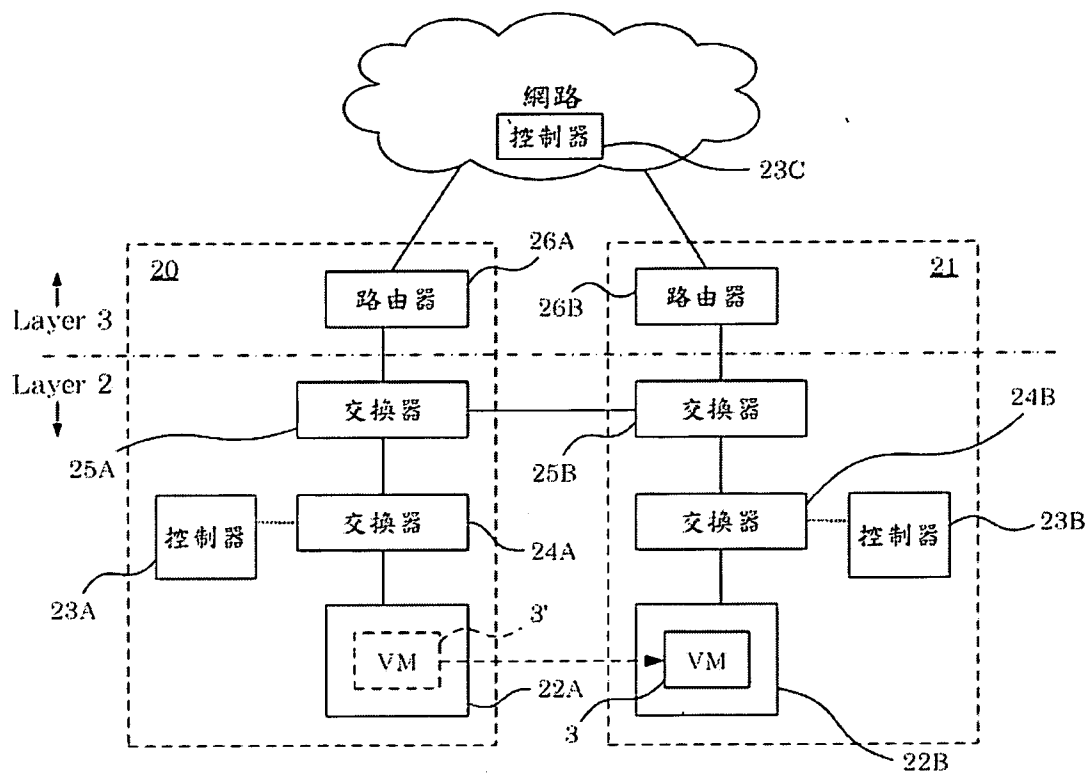
【發明圖式】



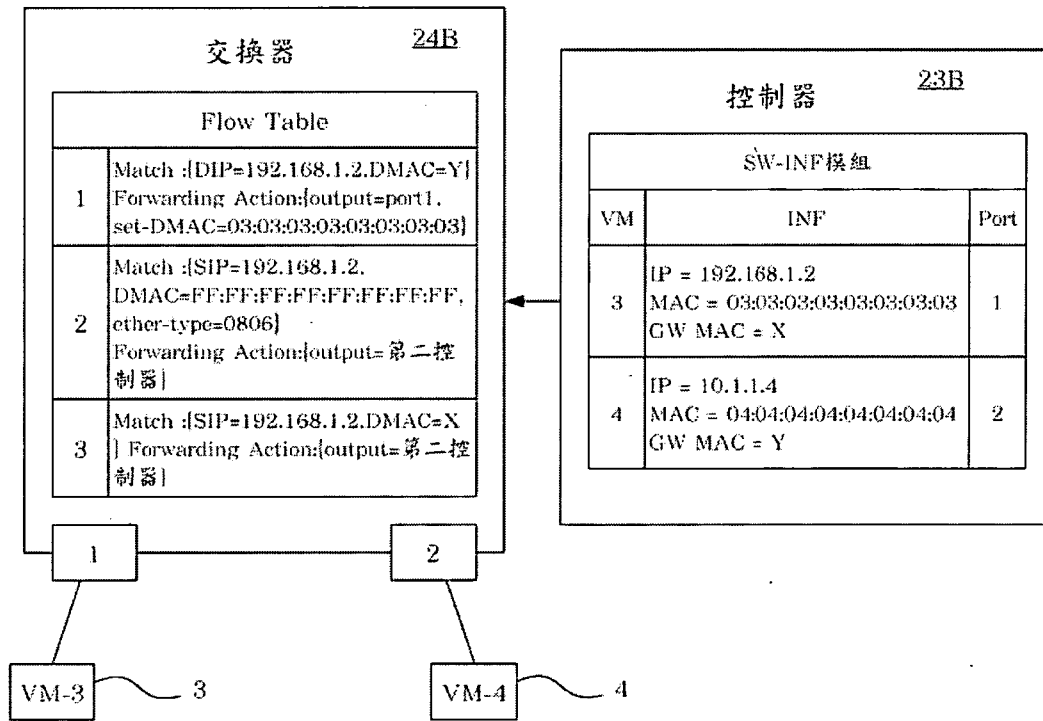
第1圖  
(先前技術)



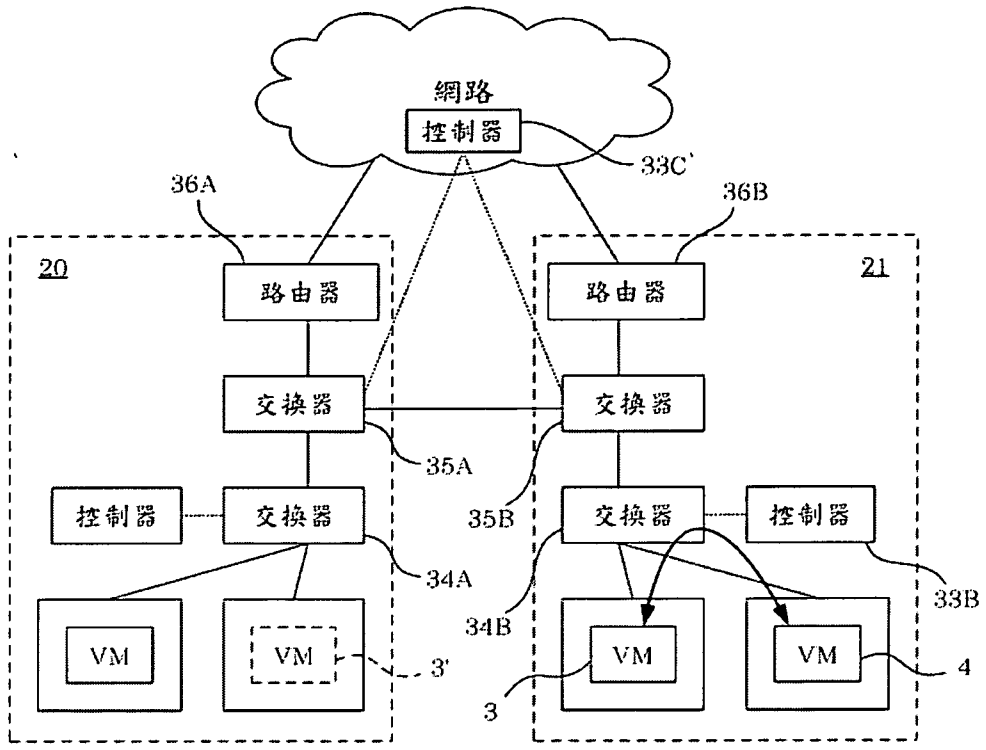
第2圖



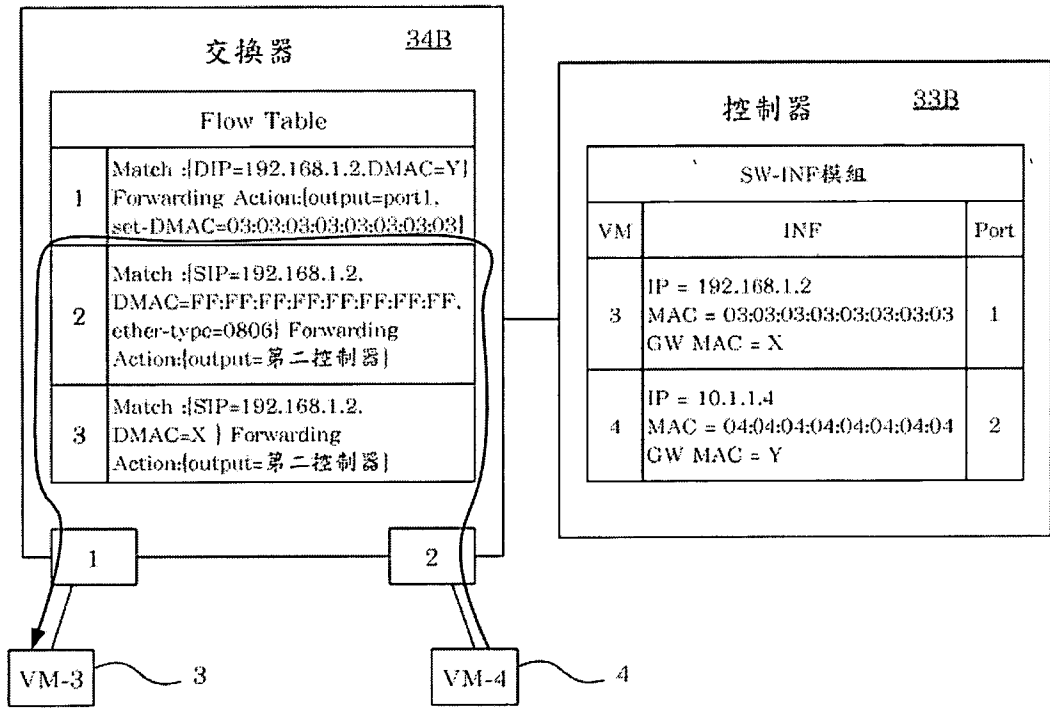
第3圖



第4圖

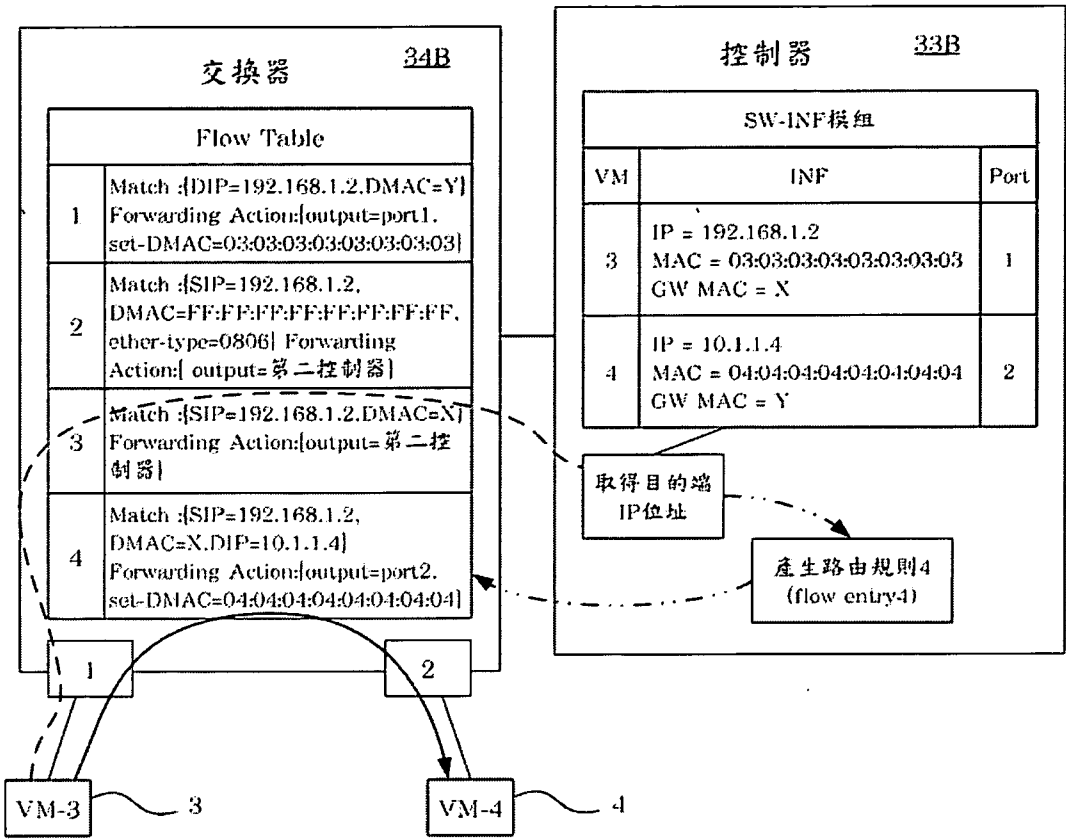


第5圖

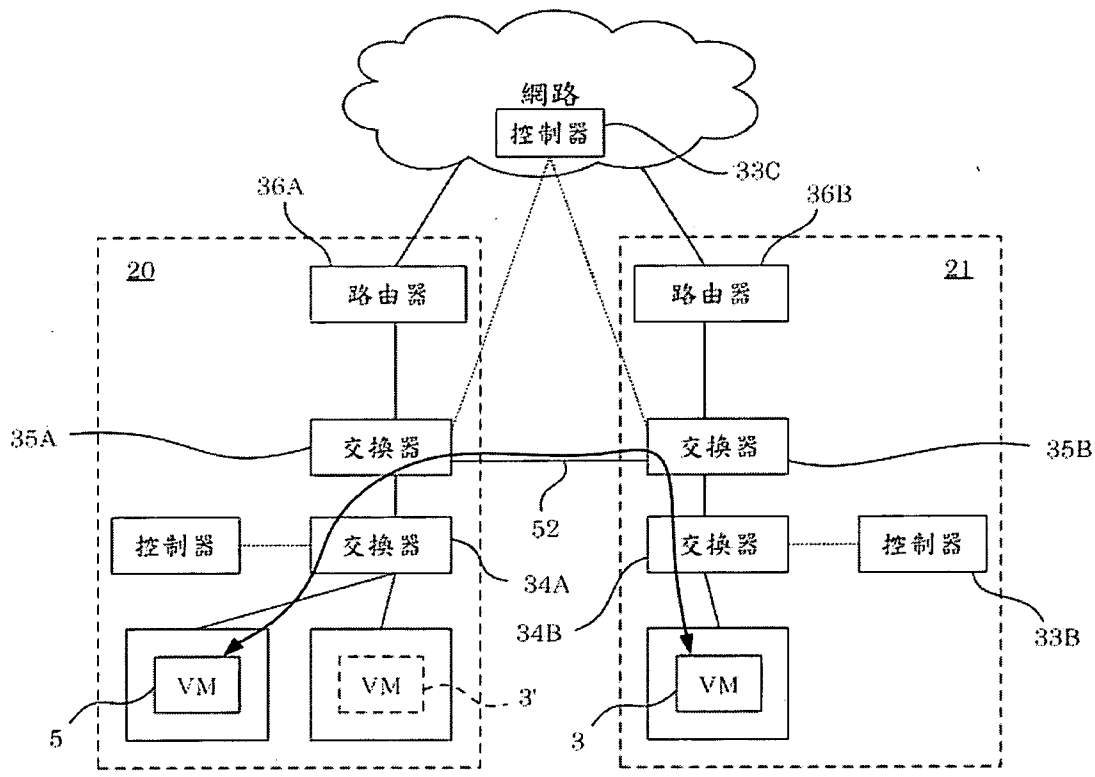


第6圖





第7圖



第8圖