I821373

(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 I821373 B**

(45)公告日： 中華民國 112 (2023) 年 11 月 11 日

(21)申請案號：108130070 (22)申請日： 中華民國 108 (2019) 年 08 月 22 日

(51)Int. Cl.： ***H04L45/00 (2022.01)*** ***H04L45/28 (2022.01)***
H04L12/46 (2006.01)

(30)優先權：2018/08/23 美國 62/722,003

(71)申請人：美商阿爾克斯股份有限公司 (美國) ARRCUS INC. (US)
美國

(72)發明人：馬赫卓 尼拉傑 MALHOTRA, NEERAJ (US)；派德 凱優 PATEL, KEYUR (US)；楊 民傑德瑞克 YEUNG, MAN-KIT DEREK (US)；克雷格 洛夫勞倫斯 KREEGER, ROLFE LAWRENCE (US)；沙阿 歇坦舒 SHAH, SHITANSHU (US)；庫馬 拉利特 KUMAR, LALIT (IN)；拉賈拉曼 卡揚尼 RAJARAMAN, KALYANI (US)；瑞谷庫瑪 維克朗 RAGUKUMAR, VIKRAM (IN)；派 納林納許 PAI, NALINAKSH (IN)

(74)代理人：林鼎鈞

(56)參考文獻：

US 2010/0189117A1 US 2014/0112122A1

US 2016/0014025A1 US 2018/0034665A1

網路文獻 Sajassi, BGP MPLS-Based Ethernet VPN ISSN: 2070-1721 20150201 https://www.rfc-editor.org/rfc/rfc7432.txt

網路文獻 Juniper Understanding EVPN Pure Type 5 Routes Juniper 20160922 https://www.juniper.net/documentation/us/en/software/junos/evpn-vxlan/topics/concept/evpn-route-type5-understanding.html

審查人員：林宥辰

申請專利範圍項數：20 項 圖式數：11 共 54 頁

(54)名稱

網路運算環境中的第一跳轉閘道的冗餘機制系統

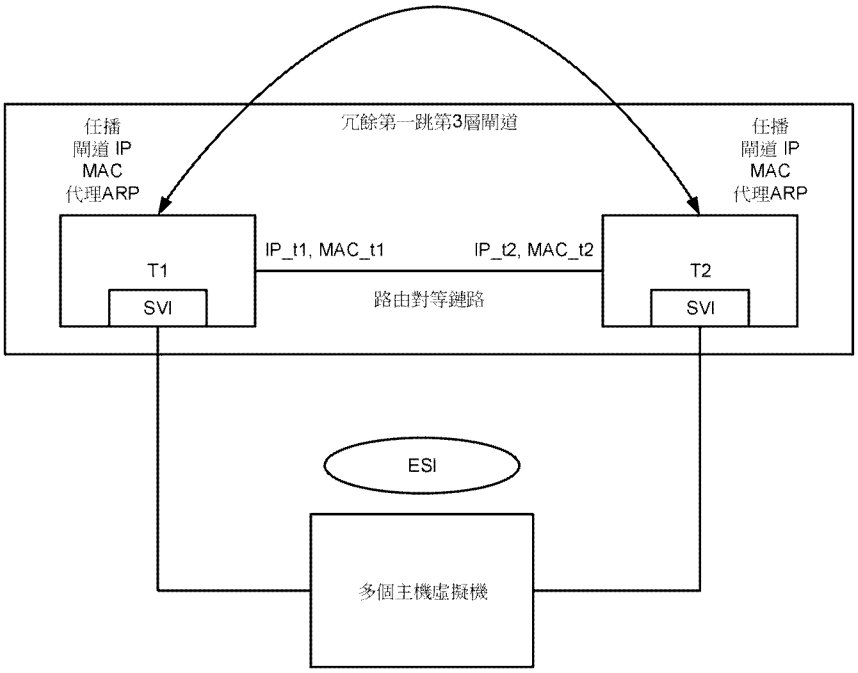
(57)摘要

用於在網路計算環境中改善路由操作的系統、方法與設備。系統包括網路拓樸中的一個第一交換器與一個第二交換器。所述系統包括與所述第一交換器與所述第二交換器中的至少一個通信的一個主機虛擬機。所述系統包括將所述第一交換器連接到所述第二交換器的一條路由對等鏈路（routed peer link）。所述系統使得所述第一交換器與所述第二交換器具有相同的網際網路協定（Internet protocol，IP）位址與媒體存取控制（media access control，MAC）位址。

Systems, methods, and devices for improved routing operations in a network computing environment. A system includes a first switch and a second switch in a network topology. The system includes a host virtual machine in communication with at least one of the first switch and the second switch. The system includes a routed peer link connecting the first switch to the second switch. The system is such that the first

switch and the second switch have the same Internet protocol (IP) address and media access control (MAC) address.

指定代表圖：



【第 2 圖】

符號簡單說明：

ESI:乙太網段標識符

IP_t1、IP_t2:本地對等鏈路網際網路協定

MAC_t1、MAC_t2:媒體存取控制位址

SVI:交換式虛擬介面

T1、T2:交換器

專利案號：108130070

公告本

I821373

【發明摘要】

【中文發明名稱】網路運算環境中的第一跳轉閘道的冗餘機制系統

【英文發明名稱】SYSTEM FOR FIRST HOP GATEWAY REDUNDANCY IN A NETWORK COMPUTING ENVIRONMENT

【中文】

用於在網路計算環境中改善路由操作的系統、方法與設備。系統包括網路拓樸中的一個第一交換器與一個第二交換器。所述系統包括與所述第一交換器與所述第二交換器中的至少一個通信的一個主機虛擬機。所述系統包括將所述第一交換器連接到所述第二交換器的一條路由對等鏈路（routed peer link）。所述系統使得所述第一交換器與所述第二交換器具有相同的網際網路協定（Internet protocol，IP）位址與媒體存取控制（media access control，MAC）位址。

【英文】

Systems, methods, and devices for improved routing operations in a network computing environment. A system includes a first switch and a second switch in a network topology. The system includes a host virtual machine in communication with at least one of the first switch and the second switch. The system includes a routed peer link connecting the first switch to the second switch. The system is such that the first switch and the second switch have the same Internet protocol (IP) address and media access control (MAC) address.

【指定代表圖】第 2 圖

【代表圖之符號簡單說明】

ESI…乙太網段標識符

IP_t1、IP_t2…本地對等鏈路網際網路協定

MAC_t1、MAC_t2…媒體存取控制位址

SVI…交換式虛擬介面

T1、T2…交換器

【發明說明書】

【中文發明名稱】網路運算環境中的第一跳轉閘道的冗餘機制系統

【英文發明名稱】SYSTEM FOR FIRST HOP GATEWAY REDUNDANCY IN A NETWORK COMPUTING ENVIRONMENT

【技術領域】

【0001】本發明涉及一種多個計算網路，特別是涉及在一個計算機網路環境中的網絡拓樸與路由協定。

【先前技術】

【0002】網路計算是使多個電腦或節點一起工作並通過網路相互通信的一種方式。存在有廣域網路（wide area network，WAN）與區域網路（local area network，LAN）。廣域網路與區域網路皆允許電腦之間的互連。區域網路通常用於較小的、更加本地化的網路，這些網路可用於家庭、企業、學校等。廣域網路覆蓋較大的區域，例如；城市，甚至可以允許不同國家的電腦進行連接。區域網路通常比廣域網路更快、更安全，但是廣域網路可以實現廣泛的連接。區域網路通常由部署它們的組織在內部擁有、控制與管理，而廣域網路通常需要兩個或多個區域網路通過公用網際網路或通過由電信提供商所建立的專用連接進行連接所組成。

【0003】區域網路與廣域網路使電腦可以相互連接並傳輸資料與其他資訊。對於區域網路與廣域網路而言，皆必須有一種方法來確定將資料從一個計算實例傳遞到另一個計算實例的路徑。這稱為路由。路由是為一個網路中、多

個網路之間或跨多個網路的流量選擇路徑的過程。路由過程通常基於路由表來指導轉發，該些路由表維護著到各個網路目的地的路由記錄。路由表可以由管理員指定，可以通過觀察網路流量來學習，也可以藉助路由協定來構建。

【0004】小型網路可以使用手動配置的路由表來確定資訊應如何從一台電腦傳播到另一台電腦。一個路由表可以包括“最佳路徑”的列表，該列表指示一個起始電腦與一個最終目標電腦之間的最有效或最理想的路徑。較大的網路（包括連接到公用網際網路的網路）可能依賴於複雜的拓樸結構，該些拓樸結構可能會快速變化，因此所述手動構建路由表是不可行的。動態路由試圖通過基於路由協定承載的資訊自動建構路由表來解決此問題。動態路由使網路能夠幾乎自主地採取行動，以避免網路故障與阻塞。存在多種路由協定，其提供用於確定網路設備之間的最佳路徑的規則或指令。動態路由協定與演算法的示例包括路由資訊協定（Routing Information Protocol，RIP）、開放式最短路徑優先（Open Shortest Path First，OSPF）、加強型閘道間選徑協定（Enhanced Interior Gateway Routing Protocol，EIGRP）與邊界閘道協定（Border Gateway Protocol，BGP）。

【0005】在某些情況下，路徑選擇涉及將路由度量（routing metric）應用於多個路由以選擇或預測最佳路由。大多數路由演算法一次僅使用一個網路路徑。多路徑路由技術允許使用多個替代路徑。在計算機網路中，一個路由演算法可用於預測兩個計算實例之間的最佳路徑。所述路由演算法可以基於多個因素，例如：頻寬、網路延遲、中繼段個數（hop count）、路徑成本、負載、最大傳輸單位、可靠性與通信成本。所述路由表儲存最佳路徑的列表。拓樸資料庫可以儲存最佳路徑的列表，並且可以進一步儲存其他資訊。

【0006】在某些網路中，沒有一個實體負責選擇最佳路徑的事實，使路由變得複雜。相反，在選擇最佳路徑或單一路徑的事件部分時會涉及多個實體。在通過網際網路進行計算機網路的環境中，網際網路被劃分為多個自主系統（autonomous system，AS），例如：多個網際網路服務提供商（Internet Service Providers，ISPs）。每一自主系統控制涉及其網路的路由。基於所述邊界閘道協定（BGP）選擇多個自主系統級路徑。每一自主系統級路徑包括一系列自主系統，多個資訊流封包通過該些自主系統從一個計算實例流到另一個計算實例。每一自主系統可以具有由相鄰自主系統提供的多個路徑來選擇。

【0007】有許多網路拓樸其對於不同的計算應用程序具有不同的優點與缺點。一種網路拓樸為葉-脊（leaf-spine）網路拓樸，其包括與多個葉節點進行通信的多個脊柱節點。用於葉-脊網路拓樸的傳統路由協定存在許多缺陷，當一個葉節點變為非活動狀態時，可能導致無效的資料環路（data loops）。存在對用於葉-脊網路拓樸的改善的標籤協定與路由協定之需求。

【0008】鑑於前述，本文公開了用於網路計算環境中的改進的路由操作的系統、方法與設備。

【發明內容】

【0009】本發明揭露一種網路運算環境中的第一跳轉閘道的冗餘機制。

【0010】首先，本發明揭露一種網路運算環境中的第一跳轉閘道的冗餘機制系統，其包括：在網路拓樸中的第一交換器、在網路拓樸中的第二交換器、與第一交換器與第二交換器中的至少一個通信的主機虛擬機以及將第一交換器

連接到第二交換器的路由對等鏈路。其中，第一交換器與第二交換器具有相同的網際網路協定（IP）位址與媒體存取控制（MAC）位址。

【0011】在一實施例中，第一交換器與二交換器配置為主機虛擬機的冗餘任播集中式閘道。

【0012】在一實施例中，第一交換器與第二交換器配置有代表主機虛擬機上主埠之公共乙太網段標識符（ESI）。

【0013】在一實施例中，第一交換器與第二交換器中的每一個配置有每個VLAN的EVPN實例（an EVPN instance per-VLAN），其具有以下一項或多項：自動導出的媒體存取控制-虛擬路由與轉發路由目標（MAC-VRF RT）；或手動配置的MAC-VRF RG。

【0014】在一實施例中，路由對等鏈路是啟用了第3層的對等鏈路。

【0015】在一實施例中，第一交換器或第二交換器中的一個或多個被配置為通告路由對等鏈路IP位址作為主機虛擬機的下一跳。

【0016】在一實施例中，系統是使用RT-1保護信號的一個乙太網虛擬私有網路（EVPN）。

【0017】在一實施例中，第一交換器與第二交換器通過路由對等鏈路交換每個乙太網段標識符（per-ESI）路由，以信號通知通過第一交換器與第二交換器形成的冗餘組上的本地ESI連接。

【0018】在一實施例中，per-ESI路由是在主機虛擬機與第一交換器或第二交換器其中之一之間的鏈路發生故障時使用的修復路徑。

【0019】在一實施例中，per-ESI路由作為一個邊界閘道協定（BGP）信息在第一交換器與第二交換器之間傳輸。

【0020】在一實施例中，第一交換器與第二交換器被配置為通過路由對等鏈路來同步位址解析協定（ARP）表。

【0021】在一實施例中，第一交換器包括一個或多個處理器，所述一個或多個處理器可配置為執行儲存在非暫態電腦可讀取媒體中的多個指令，該些指令包括：從主機虛擬機接收指示對ARP表進行更新之信息；以及通過BGP信息將更新通知第二交換器。

【0022】在一實施例中，第一交換器與第二交換器組成冗餘組，並且第一交換器與第二交換器中的一個或多個配置為通告乙太網虛擬私有網路（ EVPN）MAC位址，以同步冗餘組。

【0023】在一實施例中，第一交換器與第二交換器組成冗餘組，以使流向或來自主機虛擬機的流量在第一交換器與第二交換器之間實現負載平衡。

【0024】在一實施例中，通過跨路由對等鏈路重新路由流量來負載均衡流量。

【0025】在一實施例中，進一步包括第一交換器與主機虛擬機之間的鏈路，其中，鏈路終止於第一交換器上的一個虛擬區域網路（VLAN）。

【0026】在一實施例中，第一交換器與第二交換器配置為充當主機虛擬機的虛擬第一跳閘道。

【0027】在一實施例中，系統還包括主機虛擬機上的乙太網段標識符（ESI），其中，第一交換器與第二交換器通過路由對等鏈路具有對ESI的可達性。

【0028】在一實施例中，第一交換器被配置為向第二交換器發送邊界閘道協定（BGP）信息，該信息指示第二交換器透過第一交換器通過下一跳具有對ESI的可達性。

【0029】在一實施例中，第一交換器被配置為響應於第一交換機學習對主機虛擬機上的位址解析協定（ARP）表的更新而經由路由對等鏈路自動向第二交換器發送邊界閘道協定（BGP）信息。

【圖式簡單說明】

【0030】參考以下附圖描述了本公開的非限制性與非詳盡性的實施方式，其中，除非另外指明，否則在各個視圖中相同的附圖標記表示相同的元件。關於以下描述與附圖，將更好地理解本公開的優點，其中：

【0031】

第1圖是通過網際網路進行通信的多個網路設備的系統之示意圖。

第2圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖。

第3圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其實現修復路徑信號。

第4圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其實現多個位址解析協定（address resolution protocol，ARP）表的同步。

第5圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其說明了穩態的東西向流量。

第6圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其說明了穩態南北向流量。

第7圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其說明了東西向流量中的鏈路故障。

第8圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其說明了南北向流量中的鏈路故障。

第9圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其說明了向孤立的乙太網段標識符（Ethernet segment identifier，ESI）主機的位址解析協定（ARP）請求。

第10圖是在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路之示意圖，其說明了來自具有ARP的孤立ESI主機的回覆。

第11圖是說明一個示例計算設備的多個元件之示意圖。

【實施方式】

【0032】本申請案主張2018年8月23日申請之美國臨時專利申請案第62/722,003號發明名稱為“資料庫系統的方法與設備（DATABASE SYSTEMS METHODS AND DEVICES）”的優先權，該專利申請案係據此以引用方式併入本文中，包括但不限於下文中具體出現的那些部分，除以下情況外，以引用方式併入：在以上引用的申請案的任何部分與本申請案不一致的情況下，本申請案取代上述引用的申請案。

【0033】本文公開了用於在網路計算環境中改善網路拓樸、路由標籤與路由協定的系統、方法與設備。本公開的實施例是一種在第一交換器與第二交換器之間具有第一跳閘道冗餘的網路。在所述網路中，一個或多個主機虛擬機通過一個虛擬介面連接到所述第一交換器與所述第二交換器。所述第一跳閘道冗

餘提供了使用一個多機箱綁定介面（multi chassis bond interface）實現最佳路徑冗餘的解決方案。

【0034】在一實施例中，一個系統包括具有一個第一跳閘道冗餘的網路。所述系統包括在一個網路拓樸中的一個第一交換器與一個第二交換器。所述系統包括與所述第一交換器與所述第二交換器中的至少一個通信的一個主機虛擬機。所述系統包括將所述第一交換器連接到所述第二交換器的一條路由對等鏈路。所述系統使得所述第一交換器與所述第二交換器具有相同的網際網路協定（IP）位址與媒體存取控制（MAC）位址。

【0035】在計算機網路環境中，諸如交換器或路由器之類的網路設備可用於將資訊從一個目的地傳輸到一個最終目的地。在一實施例中，一個資料封包與一個信息可以在諸如個人家中的電腦的第一位置處生成。該資料封包與該信息可以從該個人與 Web 瀏覽器互動且向可通過網際網路存取的遠程伺服器請求資訊或向其提供資訊時生成。在一示例中，該資料封包與該信息可以是該個人輸入到連接至網際網路的網頁上存取的表格中的資訊。該資料封包與該信息可能需要傳輸到該遠程伺服器，該伺服器可能在地理位置上離該個人的電腦很遠。該個人家裡的路由器與該遠程伺服器之間很可能沒有直接通信。因此，該資料封包與該信息必須通過“跳（hopping）”到不同的網路設備，直到到達該遠程伺服器的最終目的地。在該個人家中的該路由器必須確定通過連接到網際網路的多個不同設備傳輸該資料封包與該信息直到該資料封包與該信息到達該遠程伺服器的最終目的地為止的路徑。

【0036】確定從第一位置到最終目的地的最佳路徑以及將多個資料封包與多個信息轉發到下一目的地的過程是由諸如交換器或路由器之類的網路設備

所執行的重要功能。網路中多個網路設備之間的連接稱為網路拓樸。網路拓樸是通信網路中諸如鏈路與節點之類的元素的配置。一個網路拓樸可以包括有線鏈路、無線鏈路或在網路中節點之間的有線與無線鏈路的組合。所述有線鏈路的一些示例包括同軸電纜、電話線、電源線、帶狀電纜（ribbon cables）、光纖等。所述無線鏈路的一些示例包括衛星、蜂巢信號（cellular signals）、無線電信號、自由空間光通信（free-space optical communication）等。所述網路拓樸包括在網路中所有節點（例如：電腦、路由器、交換器與其他設備）的指示以及對節點之間的鏈路的指示。本文公開了用於改善網路拓樸與網路路由的系統、方法與設備。

【0037】 為了進一步理解本公開，將為眾多的網路計算設備與協定提供一些解釋。

【0038】 一個BGP實例是在網路中用於路由資訊的一個設備。一個BGP實例可以採用一個路由反射器設備的形式。所述BGP實例可以在交換器、路由器或交換器上的BGP發言者（BGP speaker）上運行。在較高的級別上，所述BGP實例將已學習的所有路徑作為前綴發送給最佳路徑控制器。所述最佳路徑控制器以這些路徑中的一組最佳路徑作為響應。允許所述最佳路徑控制器修改任何路徑的下一跳與屬性。一旦收到最佳路徑後，所述BGP實例將更新本地選路資訊庫（Routing Information Base，RIB）並將所述最佳路徑通知給其鄰居。

【0039】 一個交換器（也可以稱為交換集線器、橋接集線器或MAC橋接器）創建一個網路。大多數內部網路都使用多個交換器來連接一個建築物或園區內的電腦、印表機、電話、攝影機、燈與伺服器。一個交換器用作一個控制器，使網路連接的多個設備可以有效地相互通訊。交換器通過使用資料封包交換來

接收、處理並轉發資料至目標設備，從而連接計算機網路上的多個設備。一個網路交換器是一個多埠（multiport）網路橋接器，其使用硬體位址以在開放式系統互聯（Open Systems Interconnection，OSI）模型的資料鏈路層（第2層）處理與轉發資料。一些交換器還可以通過附加合併的路由功能來處理網路層（第3層）上的資料。這種交換器通常被稱為第3層交換器或多層交換器。

【0040】 一個路由器連接多個網路。交換器與路由器執行相似的功能，但是每一個都有自己的獨特功能，以在網路上執行。路由器是一種在電腦網路之間轉發資料封包的網路設備。路由器在網際網路上執行流量定向功能。通過網際網路發送的資料（例如：網頁、電子郵件或其他形式的資訊），以資料封包的形式發送。資料封包通常透過構成內部連接網路（例如，網際網路）的網路從一個路由器轉發到另一路由器，直到所述資料封包到達其目的地節點為止。路由器連接到來自不同網路的兩條或多條數據線。當資料封包進入其中一條數據線時，路由器讀取該資料封包中的網路位址資訊，以確定最終目的地。然後，路由器使用其路由表或路由策略（routing policy）中的資訊，將資料封包定向到其行程中的下一個網路。BGP發言者（BGP speaker）是啟用邊界閘道協定（Border Gateway Protocol，BGP）的路由器。

【0041】 客戶邊緣路由器（CE路由器）是位於客戶房屋內的路由器，它在客戶的區域網路與提供者的核心網路之間提供介面。CE路由器、提供者路由器與提供者邊緣路由器是多協定標號交換（multiprotocol label switching）結構中的元件。提供者路由器位於提供商或運營商網路的核心。提供者邊緣路由器位於網路的邊緣。客戶邊緣路由器連接到提供者邊緣路由器，提供者邊緣路由器通過提供者路由器連接到其他提供者邊緣路由器。

【0042】路由表或路由資訊庫（routing information base，RIB）是儲存在路由器或網路電腦中的資料表，列出了到特定網路目的地的路由。在某些情況下，路由表包括多個路線度量，例如：距離、重量等。所述路由表包括有關網路拓樸的資訊，該網路拓樸緊鄰其儲存的路由器。路由表的構建是路由協定的主要目標。靜態路由是通過非自動方式在路由表中創建的條目，這些條目是固定的，而不是某些網路拓樸發現過程（network topology discovery procedure）的結果。路由表可以包括至少三個資訊欄位，包括用於網路ID、度量與下一跳的欄位。所述網路ID為目標子網。所述度量為通過發送資料封包的路徑的路由度量。路由將朝度量最低的閘道的方向行進。下一跳是資料封包在到達其最終目的地的途中要發送到的下一站的位址。路由表可以進一步包括與路由相關的服務品質、與到與該路由相關的過濾標準列表的鏈路以及用於乙太網卡的介面等等。

【0043】為了說明路由表的概念，可以將所述路由表模擬為使用地圖來傳遞包裹。路由表類似於使用地圖，用以將包裹遞送到其最終目的地。當一個節點需要將資料發送到網路上的另一個節點時，所述節點必須首先知道將資料發送到哪裡。如果該節點無法直接連接到該目標節點，則該節點必須沿著到該目標節點的正確路由將資料發送到其他節點。大多數節點不會嘗試找出哪些路由可能有用。相反，一個節點發送一個IP封包至區域網路中的一個閘道，然後由所述閘道決定如何將資料路由到正確的目的地。每一閘道都需要追踪傳遞各種資料封包的路線，為此，它使用路由表。路由表是一個資料庫，它追踪路徑（類似一張地圖），並使用這些路徑來確定轉發流量的路線。閘道還可以與其他請求資訊的節點共享其路由表的內容。

【0044】對於逐跳路由（hop-by-hop routing），每一路由表針對所有可到達的目的地列出沿著通往目的地的路徑的下一設備的位址（即下一跳）。假設路由表是一致的，那麼將資料封包中繼到其目的地的下一跳之演算法就足以在網路中的任何地方傳遞資料。逐跳是IP網際網路層（IP Internetwork Layer）與OSI模型的特徵。

【0045】OSI模型是一個概念模型，其可描述與標準化計算系統的通信功能，而無需考慮其底層內部結構與技術。所述OSI模型的目標為具有標準通信協定的各種通信系統之互通性。所述 OSI模型將通信系統劃分為多個抽象層。一層服務於其上方的層，而被其下方的層服務。例如，在網路上提供無錯誤通信的層提供了其上層應用程序所需的路徑，同時它要求下一個較低的層來發送與接收構成該路徑內容的資料封包。將同一層的兩個實例可視化為通過該層中的水平連接進行連接。通信協定使一個主機中的一個實體能夠與同一層的另一主機中相應的一個實體進行相互作用。像OSI模型一樣，服務定義抽象地描述了由（N-1）層提供給（N）層的功能，其中，N是在本地主機中操作的協定層之一。

【0046】路由控制是一種網路管理，旨在改善網際網路連接並降低頻寬成本與整體網際網路操作。一些路由控制服務包括一套基於硬體與基於軟體的產品與服務，它們可以一起工作以提高整體網際網路性能並以最小的成本微調可用網際網路頻寬的使用。在網路或自主系統正在從多個提供商獲取網際網路頻寬的情況下，路由控制可能會成功。路由控制可以幫助選擇最佳的資料傳輸路徑。

【0047】一些網路通訊系統是大型的企業級網路，具有數千個處理節點。所述數千個處理節點共享來自多個網際網路服務提供商（Internet Service

Provider，ISPs）的頻寬，並且可以處理大量網際網路流量。這樣的系統可能非常複雜，且必須正確配置才能獲得可接受的網際網路性能。如果未正確配置系統以實現最佳資料傳輸，則網際網路的存取速度可能會降低，並且系統可能會消耗大量頻寬與流量。為了解決此問題，可以實施一組服務來消除或減少這些問題。這組服務可以稱為路由控制。

【0048】路由控制機制的一個實施例係由硬體與軟體組成。所述路由控制機制通過其與一個ISP的連接來監視所有發出的流量。所述路由控制機制有助於選擇最佳路徑以進行有效的資料傳輸。所述路由控制機制可以計算所有ISP的性能與效率，並僅選擇在適用區域中表現最佳的ISP。可以根據與成本、性能與頻寬有關的已定義參數來配置路由控制設備。

【0049】用於確定資料傳輸的最佳路徑之已知演算法稱為邊界閘道協定（BGP）。BGP是一種路徑向量協定，其為網際網路上的自主系統提供路由資訊。如果BGP配置不正確，可能會導致嚴重的可用性與安全性問題。此外，修改後的BGP路由資訊可允許攻擊者重定向大流量塊，從而使所述流量在到達其預期目的地之前先行到達某些路由器。可以實施所述BGP最佳路徑演算法，以確定要安裝在網際網路協定（IP）路由表中以進行流量轉發的最佳路徑。BGP路由器可以配置為接收到同一目的地的多個路徑。

【0050】所述BGP最佳路徑演算法將第一個有效路徑分配為當前最佳路徑。所述BGP最佳路徑演算法將最佳路徑與列表中的下一個路徑進行比較，直到所述BGP到達有效路徑列表的末尾。所述列表提供了用於確定最佳路徑的規則。例如，所述列表可以包括以下指示：首選具有最高權重的路徑、首選沒有本地偏好（local preference）的路徑、首選通過網路或聚集BGP在本地發起的路

徑、首選最短路徑、首選具有最低多出口鑑別器（the lowest multi-exit discriminator）的路徑，依此類推。可以定制BGP最佳路徑選擇過程。

【0051】在BGP路由的背景下，每一選路領域（routing domain）都稱為一個自主系統（autonomous system，AS）。BGP協助選擇通過網際網路連接兩個選路領域的路徑。BGP通常選擇經過最少自主系統的路由，稱為最短AS路徑。在一實施例中，一旦啟用了BGP，一個路由器將從可能是ISPs的BGP鄰居中提取網際網路路由的列表。隨後BGP將仔細檢查所述列表，以查找具有最短AS路徑的路由。這些路由可以被輸入到路由器的路由表中。通常，一個路由器會選擇到一個AS的最短路徑。BGP使用路徑屬性來確定如何將流量路由到特定網路。

【0052】等價多重路徑（Equal cost multipath，ECMP）路由是一種路由策略，其中，轉發到單一目的地的下一跳資料封包可以通過多個“最佳路徑”進行。基於路由度量計算，所述多個最佳路徑是等效的。多路徑路由（Multiple path routing）可以與許多路由協定結合使用，因為路由是僅限於單一路由器的逐跳決策。多路徑路由可以通過負載平衡多條路徑上的流量來顯著增加頻寬。但是，在實際部署策略時，ECMP路由存在許多已知問題。本文公開了用於改善ECMP路由的系統、方法與設備。

【0053】一個Clos網路可以部署在電信中。Clos網路是多級電路交換網路，其代表了多級交換系統的理想化。一個Clos網路包括三個階段，包括入口階段、中間階段與出口階段。每一階段都由許多縱橫交換機（crossbar switch）組成。每一單元（cell）都進入一個入口縱橫交換機，該入口縱橫交換機可以通過任何可用的中間級縱橫交換機路由至相關的出口縱橫交換機。如果將入口交換

器連接到中間級交換機的鏈路以及將中間級交換機連接到出口交換機的鏈路都是空閒的時候，則中間級交叉開關（middle stage crossbar）可用於特定的新呼叫。

【0054】 可以部署一個葉-脊網路拓樸結構，以連接計算機網路中的多個節點。所述葉-脊拓樸具有兩層，包括葉層與脊柱層。所述葉層由多個存取交換器組成，這些存取交換器連接到伺服器、防火牆、負載平衡器與邊緣路由器等設備。所述脊柱層由執行路由並形成網路主幹的多個交換器所組成，每一葉交換器與每一脊柱交換器互連。在一個葉-脊拓樸中，所有設備彼此之間的鏈路數量相同，並且包括可預測且一致的延遲或等待時間，以傳輸資訊。

【0055】 虛擬區域網路（VLAN）是一個廣播域，在計算機網路中的資料鏈路層被分區與隔離。一個VLAN可以將多個標籤應用於多個網路框架，並在多個網路系統中處理這些標籤，以創建實際上位於單一網路上但看起來好像在各個網路之間拆分之網路流量的外觀與功能。多個VLAN即使連接到同一實體網路，也可以使網路應用程序分離，並且不需要部署多套電纜與網路設備。

【0056】 交換式虛擬介面（switched virtual interface，SVI）是一個虛擬介面與連接埠，其可為一個受管理的交換器傳輸多個未標記的VLAN封包（untagged-VLAN packets）。傳統上，交換器僅將流量發送到同一廣播域（單一VLAN）內的多個主機，而路由器則處理不同廣播域（不同VLANs）之間的流量。在這樣的實現中，如果沒有路由器，則不同廣播域中的網路設備將無法通信。當實施SVI後，一個交換器可以使用一個虛擬第3層介面將流量路由到其他第3層介面。這消除了對實體路由器的需求。多個VLAN通過將一個LAN分成多個較小的區段（segments）並將本地流量保留在一個VLAN內，從而減輕了網路負載。但是，由於每一VLAN都有其自己的域，因此需要一種機制，使該些

VLAN將資料傳遞到其他VLAN而不將資料通過路由器傳遞。所述SVI就是這樣一種機制。通常在多個交換器（例如，第3層與第2層交換器）上找到SVI。當實施SVI後，一個交換器可以識別發送VLAN的多個本地封包目的地，並可以交換那些發往不同VLAN的封包。在一實施例中，一個VLAN與一個SVI之間存在一對一的映射。在這樣的實施例中，單一個SVI僅可以被映射到一個VLAN。

【0057】 為了促進對本公開的原理的理解，將參考附圖中示出的實施例，並且將使用特定語言來描述它們。然而，將理解的是，由此並不意圖限制本公開的範圍。相關領域的技術人員通常會想到並擁有本公開，對本文所示的發明特徵進行任何更改和進一步修改，以及本文所示的本公開原理的任何其他應用，應當被認為在所要求保護的公開的範圍內。

【0058】 在公開與描述用於追踪網路計算環境中的目標的生命週期之結構、系統與方法之前，應當理解，本公開不限於本文所公開的特定結構、配置、處理步驟與材料，如：這樣的結構、配置、處理步驟與材料可能會有所不同。還應理解，本文所採用的術語僅用於描述特定實施例的目的，而無意於限制本公開的範圍，因為本公開的範圍僅由所附屬的申請專利範圍及其等同物限制。

【0059】 在描述與要求保護本公開的主題時，將根據以下闡述的定義使用以下術語。

【0060】 必須注意的是，在本說明書與所附屬的申請專利範圍中使用時，除非上下文另外明確指出，否則單數形式「一」、「一個」與「該」也預期包含複數形式。

【0061】如本文所使用的術語「包括」、「具有」、「含有」、「由...表徵」及其語法等同物係為包括性或開放性的術語，且不排除額外未敘述的要素或方法步驟。

【0062】如本文所使用的詞組「由...組成（consisting of）」及其語法等同物排除了申請專利範圍中未指定的任何要素或步驟。

【0063】如本文中所使用的詞組「基本上由...組成（consisting essentially of）」及其語法等同物將限制申請專利範圍為指定的材料或步驟，以及那些實質上不影響基本且新穎的特徵或者要求保護的公開的特徵。

【0064】參考「第1圖」，「第1圖」示出了用於將設備連接到網際網路的系統100的示意圖。系統100包括通過交換器106連接的多個區域網路110。每一區域網路110可以透過路由器112通過公共網際網路彼此連接。在「第1圖」的系統100中，有兩個區域網路110。但是，應該理解，可能有許多區域網路110通過公共網際網路相互連接。每一區域網路110包括通過交換器106彼此連接的多個計算設備108。多個計算設備108可以包括例如桌上型電腦、筆記型電腦、印表機、伺服器等。區域網路110可以通過路由器112透過公共網際網路與其他網路通訊。路由器112將多個網路彼此連接。路由器112連接到一個網際網路服務提供商102。網際網路服務提供商102連接到一個或多個網路服務提供商104。如「第1圖」所示，網路服務提供商104與其他本地網路服務提供商104通訊。

【0065】交換器106通過使用封包交換連接區域網路160中的多個設備，以接收、處理並轉發資料到一個目的地設備。舉例而言，交換器106可以被配置為從一台電腦接收發往印表機的資料。交換器106可以接收所述資料、處理所述資料，並將所述資料發送到印表機。交換器106可以是第1層交換器、第2層交換器、

第3層交換器、第4層交換器、第7層交換器等。一台第1層網路設備傳輸資料，但不管理通過它的任何流量。第1層網路設備的一個示例是乙太網集線器。第2層網路設備是使用硬體位址在資料鏈路層（第2層）處理與轉發資料的多連接埠設備。第3層交換器可以執行路由器通常執行的部分或全部功能。但是，某些網路交換器僅限於支持單一類型的實體網路，通常是乙太網，而路由器可能在不同的連接埠上支持不同種類的實體網路。

【0066】 路由器112是在計算機網路之間轉發資料封包的網路設備。在「第1圖」所示的示例性系統100中，路由器112在區域網路160之間轉發資料封包。然而，路由器112不一定必須用於區域網路110之間的轉發資料封包，並且可以用於廣域網路之間轉發資料封包等等。路由器112在網際網路上執行流量導向功能。路由器112可以具有用於不同類型的實體層連接的多個介面，例如：銅電纜、光纖或無線傳輸。路由器112可以支持不同的網路層傳輸標準。每一網路介面用於使資料封包從一個傳輸系統轉發到另一個傳輸系統。路由器112也可以用於連接多個計算機設備的兩個或更多個邏輯組，稱為子網，每一子網具有不同的網路前綴（prefix）。路由器112可以提供企業內部、企業與網際網路之間或網際網路服務提供商的網路之間之連接，如「第1圖」所示。一些路由器112被配置為互連各種網際網路服務提供商，或者可以在大型企業網路中使用。較小的路由器112通常為家庭與辦公室網路提供到網際網路的連接。「第1圖」中所示的路由器112以表示用於網路傳輸的任何合適的路由器，例如：邊緣路由器、訂戶邊緣路由器、提供商間邊界路由器、核心路由器、網際網路骨幹、埠轉發、語音/資料/傳真/影像處理路由器等等。

【0067】 網際網路服務提供商（ISP）102為提供用於存取、使用或參與網際網路的服務之組織。ISP 102可以以各種形式來組織，例如：商業、社區所有、非營利性或私有。ISP 102通常提供的網際網路服務包括網際網路存取、網際網路傳輸、域名註冊、Web寄存、Usenet服務與寄存。「第1圖」所示的ISP 102可以代表任何合適的ISP，例如：代管ISPs（hosting ISPs）、中轉ISPs（transit ISPs）、虛擬ISPs、免費ISPs（free ISPs）、無線ISPs等等。

【0068】 網路服務提供商（network service provider，NSP）104係為透過向網際網路服務提供商提供直接網際網路主幹（Internet backbone）存取，提供頻寬或網路存取的組織。網路服務提供商可以提供對網路進出點（network access points，NAPs）的存取。網路服務提供商104有時被稱為骨幹提供商或網際網路提供商。網路服務提供商104可以包括提供高速網際網路接取（Internet access）的電信公司、資料載體、無線通信提供商、網際網路服務提供商與有線電視運營商。 網路服務提供商104也可以包括信息資訊技術公司。

【0069】 應當理解，「第1圖」所示的系統100僅是示例性的，其可以創建用於在網路與計算設備之間傳輸資料的系統與許多不同的配置。因為在網路形成中存在大量可定制性，所以需要在確定用於在電腦之間或在網路之間傳輸資料的最佳路徑時創建更大的可定制性。鑑於前述內容，本文公開了用於將最佳路徑計算卸載到外部設備以使得能夠在確定非常適合於特定類別的電腦或特定企業之最佳路徑演算法時實現更大可定制性的系統、方法和設備。

【0070】 「第2圖」至「第10圖」說明了用於實現第一跳閘道冗餘的網路的多個實施例。在一實施例中，多個主機虛擬機連接到多個交換器。在「第2圖」

至「第10圖」中將多個交換器描繪為T1與T2。「第2圖」至「第10圖」中的實施例說明了使用一個多機箱綁定介面來建立最佳路徑冗餘的方法。

【0071】「第2圖」至「第10圖」的實施例的拓樸包括跨越多個主機虛擬機到交換器T1與交換器T2的介面。所述介面是在所述主機虛擬機上相同綁定介面的一部分。該些鏈路終止於交換器T1與T2上的一個虛擬區域網路（VLAN）。該些實施例可以部署在用作第一跳閘道的一個第3層路由介面中。在這樣的實施例中，如果一個主機虛擬機需要到達另一個主機虛擬機，則可以通過交換器T1與T2中的一個或多個來促進通信。交換器T1與T2一起充當該些主機虛擬機的一個虛擬第一跳閘道。從一個主機虛擬機的角度來看，交換器T1與T2配置有相同的閘道IP位址與相同的閘道MAC位址。因此，從一個主機虛擬機的角度來看，該主機虛擬機正在與單一個閘道IP通信，而不是與位於兩個不同的交換器T1與T2上的兩個閘道IP通信。

【0072】通過在交換器T1與T2上配置相同的IP與MAC位址來實現冗餘。此外，在交換器T1與T2之間配置了包括所述IP與MAC位址的一條路由對等鏈路。

【0073】交換器T1與T2可以通過BGP信號互相發信號。在一實施例中，交換器T1與T2各自發送信號通知它們的修復路徑的末端以處理鏈路故障。

【0074】在一實施例中，存在用於處理鏈路故障的一條修復路徑。例如，交換器T1與一個主機虛擬機之間的鏈路故障。需要使交換器T2利用一條修復路徑來重定向來自所述主機虛擬機的流量。這樣可以通過所述路由對等鏈路達到乙太網段標識符（ESI）。交換器T2可以從交換器T1接收BGP信息，所述信息指示通過交換器T1到下一跳之方式使交換器T1具有到達ESI的能力。交換器T1通過綁定學習的任何主機虛擬機皆將安裝為了直接連接的主機虛擬機的一條固定路

徑（anchor path）。在一條鏈路故障的情況下，一條自動修復路徑被啟動以通過交換器T2發送流量。

【0075】在一實施例中，為交換器T1執行ARP SYNC以將封包路由到所述主機虛擬機。所述ARP SYNC包括在交換器T1與T2中同步多個ARP表。如果交換器T1從所述主機虛擬機獲悉更改，則交換器T1可以使用BGP EVPN信號與交換器T2同步。當發生更改時，BGP信號可能會自動發送。當交換器T1在本地SVI介面上的主機H1上學習ARP綁定時，交換器T1可能會向交換器T2生成一BGP VPN路由類型2的信息，所述信息包括所述IP與學習IP的站點。

【0076】在一實施例中，從一個孤立的ESI主機生成ARP回覆。可以使用交換器T1與T2之間的信息傳遞來執行所述ARP回覆，以發送一個ARP請求。可以使用交換器T1與T2之間的BGP EVPN路由類型2的信息發送回所述ARP回覆。

【0077】在一實施例中，交換器T1、交換器T2與一個主機虛擬機或虛擬消費電子設備之間的鏈路可能會故障。如果鏈路故障，則可以從轉發中刪除該路徑。鏈路發生故障的交換器可能會自動刪除故障路徑。所述交換器可能已經了解了所有路由並將這些路由聚合到一個路由覆蓋協定中，然後可以從這些路由中撤銷故障路徑，使得目的地的所有流量將不再通過故障的鏈路發送。

【0078】「第2圖」是具有第一跳閘道冗餘的網路之示意圖。所述網路包括代表網路設備（例如交換器或路由器）的T1與T2。T1與T2中的每一個都包含一個交換式虛擬介面（SVI）。T1與T2之間存在一條路由對等鏈路。T1與T2分別包括一個任播閘道IP（anycast gateway IP）、一個任播MAC（anycast MAC）與一個任播代理ARP（anycast proxy-ARP）。T1通告稱為IP_t1的一個本地對等鏈路IP。T2通告稱為IP_t2的一個本地對等鏈路IP。T1通告稱為MAC_t1的一個媒

體存取控制（MAC）位址。T2通告稱為MAC_t2的一個MAC位址。T1與T2通過所述SVI連接與多個主機虛擬機通信。

【0079】配置所述網路使得T1與T2充當冗餘任播集中式閘道（anycast centralized gateways）。T1與T2是通過第2層LAG捆綁進行多宿主的多個主機的閘道。T1與T2配置有一個SVI與任何任播閘道MAC以及用於南北路由的任播閘道IP。T1與T2配置有代表鏈路聚合（link aggregation，LAG）主埠的一個公共乙太網虛擬私有網路（ethernet virtual private network，EVPN）乙太網段標識符（ESI）。T1與T2配置有每個VLAN的EVPN實例（an EVPN instance per-VLAN），其具有媒體存取控制-虛擬路由與轉發路由目標（media access control-virtual routing and forwarding route targets，MAC-VRF RTs）。所述 MAC-VRF RT可以自動導出或手動配置。T1與T2配置有用於保護的一個第3層啟用對等鏈路。在一實施例中，在T1與T2之間建立的一個BGP-EVPN通信期（session）以通告本地對等鏈路IP（可以稱為IP_t1與IP_t2）作為下一跳。

【0080】在所述網路中，存在一個BGP-EVPN控制平面，用於通過RT-1通知一條修復路徑。所述BGP-EVPN控制平面還通過RT-2發出一個ARP SYNC信號，並通過RT-2發出一個ARP請求信號。

【0081】在一個資料中心網路中，可以使用一個基於乙太網虛擬私有網路與RT-1的保護信號來提供第一跳閘道冗餘。在此配置中，假設在T1/T2以北的第3層路由網路。在這種情況下，存取與IP單播流量上只有第2層連接。在一個實現中，僅第2層連接被允許存取。

【0082】「第3圖」說明了通過EVPN RT-1提供修復路徑信號的網路。在所述網路中，T1與T2對等體（peer）與IP_t1和IP_t2下一跳交換per-ESI RT-1（乙

太網AD路由）。這表示冗餘組對等體（redundancy group peers）之間的本地ESI連接。此外，針對在一個ESI主埠上配置的多個VLAN，EVI-RTs與per-ESI RT-1一起通告，以用於導入到MAC-VRF中。這個per-ESI RT-1可以通過所述冗餘組對等體向一個給定ESI上所有直接連接的多個主機發出一個第3層修復路徑的信號。

【0083】在一實施例中，修復路徑信號需要所述RT-1，因為取決於學習ARP的位置，RT-2可能並不總是由T1與T2生成。在這樣的實施例中，將不具有從任何對等體發信號通知的修復路徑之ESI視為並處理為一個孤立的ESI。

【0084】「第4圖」是提供一個主機鄰接同步與修復路徑編程的網路之示意圖。所述網路可以從一個本地ARP快取記憶體中學習本地（例如，MAC+IP與/或SVI）。給定從所述SVI到所述VLAN的EVPN背景（context）。所述網路可以通過源自所述MAC源之給定EVPN背景中的HW MAC學習更新來學習本地（MAC到AC）。所述網路可以進一步通過一個本地MAC執行本地MAC+IP解析，以導出用於ARP學習的MAC+IP的連接埠與/或ESI。一旦解決，所述網路便可以通告EVPN MAC+IP RT-2，以實現跨冗餘組對等體（例如，從T1到T2）的MAC+IP同步。

【0085】所述網路可以參考T2以通過EVI-RT映射將所述MAC+IP RT-2導入MAC-VRF。所述網路可以通過來自T1的一個per-ESI RT-1與一個查找本地ESI DB的方式解析來自T1的MAC+IP RT-2，以檢查接收到的MAC+IP與ESI的本地連接。如果接收到的ESI是本地的，則如果網路不是動態學習的，網路可以在本地VLAN SVI介面上為接收到的IP安裝靜態ARP條目。可以通過相應的ESI的一個RT-1學習下一跳來存取FIB，以帶有保護地安裝ARP學習鄰接路由。

【0086】在一個本地ESI（動態或已同步）上學習到的所有主機鄰接都通過該ESI的RT-1學習修復路徑安裝保護。

【0087】所述網路可以提供T2 ESI故障處理。在出現故障的情況下，所述網路可以通過一個冗餘對等體啟動一條修復路徑。

【0088】「第5圖」是具有在本地路由的穩態東西向流量（steady state East-West flows）的網路之示意圖。所述網路包括具有所述乙太網段標識符（ESI）的多個主機虛擬機。一組主機虛擬機與ESI-2相關，另一組與ESI-1相關。所述穩態東西向流量包括子網內與子網間流量的本地路由。如「第5圖」所示，存在從儲存ESI-2的多個主機虛擬機到儲存ESI-1的多個主機虛擬機的穩態流量。所述主機虛擬機與T1與T2中的每一個之間皆有通信。

【0089】所述網路可以提供東西向子網內流量，以避免任何第2層泛洪（flooding）或東西向流量的橋接。這可以通過代理ARP機制來完成。所述ARP機制可以傳輸從橋接到所述本地SVI介面的多個存取面向主機（access-facing hosts）所接收到的多個廣播ARP請求。在一實施例中，所述ARP請求不泛洪在所述VLAN中的其他第2層埠上。在這樣的實施例中，在所述SVI介面上所接收到的多個ARP請求被代理回覆予所述任播閘道MAC。同樣地，源自所述閘道的多個ARP請求在所述本地ESI與多個本地孤立ESI埠上泛洪，而不會泛洪到所述對等閘道。因此，可以為所述SVI介面配置代理ARP與代理ND，以用於已通過本地ARP/ND條目或通過遠程MAC+IP RT-2建立可達性的完整主機。在這樣的實施例中，包括子網內流量的任何東西向流量可以在所述閘道上被第2層終止並且被路由到所述目的地鄰接。這可以類似於「第6圖」所示的南北向流量。

【0090】「第6圖」是具有在本地路由的穩態南北向流量的網路之示意圖。所述穩態南北向流量包括由北向南子網間流量之本地路由。如「第6圖」所示，存在從T1到儲存ESI-2的多個主機虛擬機以及從T2到儲存ESI-1的多個主機虛擬機的穩態流量。「第6圖」說明了到一個多宿主主機的穩態南北向流量。在TOR處接收到的發往所述主機IP的流量直接被路由到所述主機。

【0091】「第7圖」是在東西向流量中具有鏈路故障的網路之示意圖。從T1到所述主機H2的鏈路處於非活動狀態並且故障。可以通過所述對等路由的對等鏈路重新路由通過子網路由路由到發生故障的ESI上的主機的所有流量。此外，所述網路可以撤回本地的per-ESI RT-1（大規模撤回），而路由的流量以負載平衡的方式繼續通過所述子網路由跨T1與T2進行路由。以這種方式，通過所述路由對等鏈路將到達T2的流量重新路由到T1。將流量直接路由到現在孤立的ESI上連接的一個主機（如「第8圖」所示）。

【0092】「第8圖」說明了在南北向流量中具有鏈路故障的網路之示意圖。所述網路通過到T2的路由對等鏈路啟動到H2上的孤立ESI主機的一條修復路徑。從T1到所述主機H2的鏈路處於非活動狀態並且故障。如「第6圖」所示，在穩態南北向流量中，通常將流量從T1與/或T2直接路由到相應的主機虛擬機群。當從T1到主機H2的鏈路發生故障時，可以通過所述路由的對等鏈路將流量從T1路由到T2。然後可以將所述流量路由到適當的主機H2。

【0093】此外，到達T2的流量可以通過所述路由對等鏈路重新路由到T1，T1將把所述流量直接路由到現在孤立的ESI上的連接主機，如「第8圖」所示。所述網路可以提供T1“孤立ESI”處理。對於每一個ESI RT-1，所述網路可以承諾

從T2大規模撤回。這可能導致本地ESI移至所述“孤立”狀態。所述網路可以重新編程轉發，以在從所述對等體大規模撤回時刪除修復路徑編程。

【0094】所述網路可以從T1大量撤回RT-1。這可能導致無法解析來自T1的MAC+IP RT-2路徑。作為響應，如果由於T1的MAC+IP RT-2而存在，則所述網路可能會刪除靜態來源的SYNC-ARP，並將多個主機路由注入到在一個孤立ESI上學習到的所有主機鄰接關係（即ARP條目）之所述預設路由控制平面中。一旦注入，更特定的路由將使發往一個孤立埠（orphan port）上的多個主機之流量匯聚到通向T1的直接路徑。

【0095】「第9圖」是向孤立ESI主機執行ARP請求的網路之示意圖。所述網路可以執行一個孤立ESI主機的“ARPing”，其中ARP是一個位址解析協定。例如，要通過T2保持T1孤立ESI上的多個主機之東西向與南北向可達性，T2必須能夠ARP T1孤立ESI上的一個主機。在T1與T2之間沒有第2層擴展的情況下，所述對等閘道上的多個ARP孤立主機需要一個備用機制。

【0096】「第9圖」所示的網路解決所述孤立的ESI。所述網路可能會使BGP RT-2超載，以將ARP請求發送到所述對等閘道，如「第9圖」所示。以這種方式，T2接收到對SVI上的主機IP1的一個ARP請求，或者由於拾取（glean）而需要ARP主機IP1。T2可以通過MAC+IP RT-2向T1發送ARP請求。作為響應，T1在所述本地ESI與本地孤立ESI埠上生成ARP請求。T1學習所述本地主機IP1的ARP條目並生成一個MAC+IP1 RT-2。T2可以通過所述路由對等鏈路安裝IP1的可達性，如「第9圖」所示，以回覆所述孤立ESI主機（如「第10圖」所示）。

【0097】在「第2圖」至「第10圖」中所示的任何網路的實施例中，可以提供虛擬路由與轉發（VRF）支持。為了促進VRF，使用[VRF，ESI] RT-1與ESI

RT-1的網路學習了一條修復路徑。這是通過L3-VPN標籤屬性執行的。在沒有一個覆蓋的情況下，可以通過具有一個per-VRF MPLS VPN封裝之直接連接的對等鏈路發送一修復路徑，如下所示：[VRF，IP/32]到鄰接的[IP/32，SVI]到[MAC，ESI-port]（主要路徑）；或[VRF，IP/32]到鄰接的[IP_t1，P]到MAC_t1 + VPN標籤（備用路徑）。

【0098】 或者，可以將第3層VLAN標記的子介面用作一個VPN標籤的多個對等鏈路，以在多租戶環境中實現修復路徑轉發。

【0099】 在一實施例中，提供了覆蓋VPN支持與對等鏈路的一種替代方式。在一示例中，啟用一個VXLAN覆蓋後，可能不再需要一個直接連接的對等鏈路。在沒有對等鏈路的情況下，一個VPN覆蓋遍及所述冗餘組。為此，通過pre-[VRF，ESI] EAD RT-1通告的一個L3-VNI /VSLAN封裝的修復路徑可以代替所述直接連接的對等鏈路修復路徑。

【0100】 在一示例中，在T2處於穩定狀態的情況下，可以按以下方式實現封裝（encapsulation）：[VRF，IP/32]到鄰接的[IP/32，SVI]到[MAC，ESI-port]（主要路徑）；或[VRF，IP/32]到L3-VNI + VXLAN到VTEP-T1的隧道路徑（備用路徑）。

【0101】 在T2上發生ESI後故障的情況下，流量將按照以下方式在所述覆蓋修復路徑上路由：[VRF，IP/32]到L3-VNI + VXLAN到VTEP-T1的隧道路徑。

【0102】 但是，如果所述子網超出了所述冗餘組，則可以通過所述隧道路徑建立多個遠程葉節點的可達性，類似於上面的內容：[VRF，IP/32]到L3-VNI + VXLAN到VTEP-T1的隧道路徑 。

【0103】拾取處理（Glean handling）可能類似於上述對孤立主機的處理。但是，所述網路可能會用MAC將MAC+IP RT-2通告為所述擴展EVI中所有的MAC，以觸發來自參與所述EVI的所有ToR之本地ARP。

【0104】現在參考圖。參照「第11圖」，說明了示例性計算設備1100的方塊圖。計算設備1100可以用於執行各種過程，諸如本文所討論的那些過程。在一實施例中，所述計算設備1100可以執行異步物件管理器（asynchronous object manager）的功能，並且可以執行一個或多個應用程序。計算設備1100可以是多種計算設備中的任何一種，例如：桌上型電腦、汽車/航空/海事內嵌式電腦（in-dash computer）、車輛控制系統、筆記型電腦、伺服器電腦、掌上型電腦、平板電腦等。

【0105】計算設備1100包括一個或多個處理器1102、一個或多個儲存設備1104、一個或多個介面1106、一個或多個大容量儲存設備1108、一個或多個輸入/輸出（I/O）設備1110與一個顯示設備1130，其全部耦合到匯流排1112。處理器1102包括一個或多個執行儲存在儲存設備1104與/或大容量儲存設備1108中的指令之處理器或控制器。處理器1102還可以包括各種類型的電腦可讀取媒體，例如：快取記憶體。

【0106】儲存設備1104包括各種電腦可讀取媒體，例如：揮發性記憶體（如隨機存取記憶體（RAM）1114）與/或非揮發性記憶體（如唯讀記憶體（ROM）1116）。儲存設備1104還可以包括可重複錄寫ROM，如快閃記憶體。

【0107】大容量儲存設備（Mass storage device）1108包括各種電腦可讀取媒體，例如：磁帶、磁碟、光碟、固態記憶體（如快閃記憶體）等。 如「第11圖」所示，特定的大容量儲存設備為硬碟驅動器1124。各種驅動器也可以包括

在大容量儲存設備1108中，以實現從各種電腦可讀取媒體讀取與/或寫入各種電腦可讀取媒體。大容量儲存設備1108包括可移除媒體（removable media）1126與/或非可移除媒體（non-removable media）。

【0108】輸入/輸出（I/O）設備1110包括允許將資料與/或其他資訊輸入到計算設備1100或從計算設備1100取得資料與/或其他資訊的各種設備。例如，I/O設備1110包括游標控制設備、鍵盤、數字小鍵盤、麥克風、監視器或其他顯示設備、揚聲器、印表機、網路卡、數據機等。

【0109】顯示設備1130包括能夠向計算設備1100的一個或多個用戶顯示資訊之任何類型的設備。顯示設備1130的示例包括監視器、顯示終端、視頻投影設備等。

【0110】介面1106包括允許計算設備1100與其他系統、設備或計算環境互動的各種介面。舉例而言，介面1106可以包括任何數量的不同網路介面1120，諸如到區域網路（LAN）、廣域網路（WAN）、無線網路與網際網路的介面。其他介面包括用戶介面1118與周邊設備介面1122。介面1106還可以包括一個或多個用戶介面元件1118。介面1106還可以包括一個或多個周邊介面，例如：用於 印表機、指示設備（滑鼠、觸控板或本領域普通技術人員現在已知的或以後發現的任何合適的用戶介面）、鍵盤等。

【0111】匯流排1112允許處理器1102、儲存設備1104、介面1106、大容量儲存設備1108、I/O設備1110以及耦合到匯流排1112的其他設備或元件彼此通信。匯流排1112代表幾種類型的匯流排結構中的一種或多種，例如：系統匯流排、PCI匯流排、IEEE匯流排、USB匯流排等等。

【0112】為了說明的目的，程序與其他可執行程序元件在此處顯示為分離的方塊，儘管應當理解，這樣的程序和元件可以在計算設備1100的不同儲存元件中於不同時間駐留並且由處理器1102執行。或者，本文所描述的系統與過程可以以硬體或硬體、軟體與/或韌體的組合來實現。舉例而言，可以對一個或多個特殊應用積體電路（application specific integrated circuit，ASIC）進行編程以執行本文所述的一個或多個系統或過程。

【0113】出於說明與描述的目的已經給出了前面的描述。它不旨在全面性徹底或將本公開限制為所公開的精確形式。根據以上教導，許多修改與變化是可能的。 此外，應注意，可以以要求的任何組合使用任何或所有上述替代實施方式，以形成本公開的其他混合實施方式。

【0114】此外，儘管已經描述與圖示了本公開的特定實施方式，但是本公開不限於如此描述與圖示的部件之特定形式或佈置布置。本公開的範圍將由所附屬的申請專利範圍、此處與在不同申請中提交的任何將來的申請專利範圍（如果有的話）及其等效形式來定義。

【0115】例子

【0116】以下示例涉及其他實施例。

【0117】示例1是一個系統。該系統包括在一個網路拓樸中的一個第一交換器。該系統包括在該網路拓樸中的一個第二交換器。該系統包括與該第一交換器與該第二交換器中的至少一個通信的一個主機虛擬機。該系統包括將該第一交換器連接到該第二交換器的一條路由對等鏈路。該系統使得該第一交換器與該第二交換器具有相同的網際網路協定（IP）位址與媒體存取控制（MAC）位址。

【0118】 示例2是如示例1所述的系統，其中，該第一交換器與該第二交換器配置為該主機虛擬機的一個冗餘任播集中式閘道。

【0119】 示例3是如示例1-2中任一示例所述的系統，其中，該第一交換器與該第二交換器配置有代表該主機虛擬機上一個主埠（main port）之一個公共乙太網段標識符（ESI）。

【0120】 示例4是如示例1-3中任一示例所述的系統，其中，該第一交換器與該第二交換器中的每一個配置有每個VLAN的EVPN實例（an EVPN instance per-VLAN），其具有以下一項或多項：一個自動導出的媒體存取控制-虛擬路由與轉發路由目標（MAC-VRF RT）；或一個手動配置的MAC-VRF RG。

【0121】 示例5是如示例1-4中任一示例所述的系統，其中，該路由對等鏈路是一個啟用了第3層的對等鏈路。

【0122】 示例6是如示例1-5中任一示例所述的系統，其中，該第一交換器或該第二交換器中的一個或多個被配置為通告該路由對等鏈路IP位址作為該主機虛擬機的一個下一跳。

【0123】 示例7是如示例1-6中任一示例所述的系統，其中，該系統是使用RT-1保護信號的一個乙太網虛擬私有網路（EVPN）。

【0124】 示例8是如示例1-7中任一示例所述的系統，其中，該第一交換器與該第二交換器通過該路由對等鏈路交換一每個乙太網段標識符（per-ESI）路由，以信號通知通過該第一交換器與該第二交換器形成的一個冗餘組上的本地ESI連接。

【0125】 示例9是如示例1-8中任一示例所述的系統，其中，該per-ESI路由是在該主機虛擬機與該第一交換器或該第二交換器其中之一之間的鏈路發生故障時使用的一修復路徑。

【0126】 示例10是如示例1-9中任一示例所述的系統，其中，該per-ESI路由作為一個邊界閘道協定（BGP）信息在該第一交換器與該第二交換器之間傳輸。

【0127】 示例11是如示例1-10中任一示例所述的系統，其中，該第一交換器與該第二交換器被配置為通過該路由對等鏈路來同步一個位址解析協定（ARP）表。

【0128】 示例12是如示例1-11中任一示例所述的系統，其中，該第一交換器包括一個或多個處理器，該一個或多個處理器可配置為執行儲存在非暫態電腦可讀取媒體中的多個指令，該些指令包括：從該主機虛擬機接收一個指示對該ARP表進行更新之信息；以及通過一個BGP信息將該更新通知該第二交換器。

【0129】 示例13是如示例1-12中任一示例所述的系統，其中，該第一交換器與該第二交換器組成一個冗餘組，並且該第一交換器與該第二交換器中的一個或多個配置為通告一個乙太網虛擬私有網路（ EVPN）MAC位址，以同步該冗餘組。

【0130】 示例14是如示例1-13中任一示例所述的系統，其中，該第一交換器與該第二交換器組成一個冗餘組，以使流向或來自該主機虛擬機的流量在該第一交換器與該第二交換器之間實現負載平衡。

【0131】 示例15是如示例1-14中任一示例所述的系統，其中，通過跨該路由對等鏈路重新路由流量來負載均衡該流量。

【0132】示例16是如示例1-15中任一示例所述的系統，其中，進一步包括該第一交換器與該主機虛擬機之間的一鏈路，其中該鏈路終止於該第一交換器上的一個虛擬區域網路（VLAN）。

【0133】示例17是如示例1-16中任一示例所述的系統，其中，該第一交換器與該第二交換器配置為充當該主機虛擬機的一個虛擬第一跳閘道。

【0134】示例18是如示例1-17中任一示例所述的系統，其中，該系統還包括該主機虛擬機上的乙太網段標識符（ESI），其中，該第一交換器與該第二交換器通過該路由對等鏈路具有對該ESI的可達性。

【0135】示例19是如示例1-18中任一示例所述的系統，其中，該第一交換器被配置為向該第二交換器發送一個邊界閘道協定（BGP）信息，該信息指示該第二交換器透過該第一交換器通過一個下一跳具有對該ESI的可達性。

【0136】示例20是如示例1-19中任一示例所述的系統，其中，該第一交換器被配置為響應於該第一交換機學習對該主機虛擬機上的位址解析協定（ARP）表的更新而經由該路由對等鏈路自動向該第二交換器發送一個邊界閘道協定（BGP）信息。

【0137】應當理解，上述布置、示例與實施例的任何特徵可以在單一個實施例中進行組合，該單一個實施例包括取自所公開的布置、示例和實施例中的任何一個特徵的組合。

【0138】應當理解，本文所公開的各種特徵提供了本領域的顯著優點與進步。以下請求項是這些特徵中的一些特徵之示例。

【0139】在前述的本公開的詳細描述中，出於簡化本公開的目的，在單一個實施例中將本公開的各種特徵組合在一起。本公開的方法不應被解釋為反映

了這樣一種意圖，即所要求保護的公開需要比每一請求項中明確敘述的特徵更多的特徵。而是，發明方面在於少於單一個前述公開的實施例的所有特徵。

【0140】 應當理解，上述布置僅是對本公開原理的應用說明。在不脫離本公開的精神和範圍的情況下，本領域技術人員可以設計出多種修改與替代布置，並且所附屬的申請專利範圍旨在涵蓋這種修改與布置。

【0141】 因此，儘管已經在附圖中示出了本公開並且在上面詳細地描述了本公開，但是顯而易見的是，對於本領域普通技術人員而言，在不脫離本文闡述的原理和概念的情況下，可以進行許多修改，包括但不限於尺寸、材料、形狀、形式、操作的功能與方式、組裝與使用的變化。

【0142】 此外，在適當的情況下，本文所描述的功能可以在以下一項或多項中執行：硬體、軟體、韌體、數位元件或類比元件。例如，可以對一個或多個特殊應用積體電路（ASICs）或可程式邏輯陣列（field programmable gate arrays，FPGA）進行編程，以執行本文所述的一個或多個系統或過程。貫穿以下描述與申請專利範圍，使用某些術語來提及特定系統元件。如本領域技術人員將理解的，可以用不同的名稱來提及元件。本文件無意區分名稱不同但功能相同的元件。

【0143】 出於說明與描述的目的已經給出了前面的描述。它不旨在徹底或將本公開限制為所公開的精確形式。根據以上教導，許多修改與變化是可能的。此外，應注意，可以以要求的任何組合使用任何或所有上述替代實施方式，以形成本公開的其他混合實施方式。

【0144】 此外，儘管已經描述與圖示了本公開的特定實施方式，但是本公開不限於如此描述與圖示的部件之特定形式或布置。本公開的範圍將由所附屬

的申請專利範圍、此處與在不同申請中提交的任何將來的申請專利範圍及其等效形式來定義。

【符號說明】

【0145】

100…系統

102…網際網路服務提供商

104…網路服務提供商

106…交換器

108、1100…計算設備

110…區域網路

112…路由器

1102…處理器

1104…儲存設備

1106…介面

1108…大容量儲存設備

1110…輸入/輸出設備

1112…匯流排

1114…隨機存取記憶體

1116…唯讀記憶體

1118…用戶介面

1120…網路介面

1122…周邊設備介面

1124…硬碟驅動器

1126…可移除媒體

1130…顯示設備

EVPN…乙太網虛擬私有網路

ESI…乙太網段標識符

H1、H2…主機

IP_t1、IP_t2…本地對等鏈路網際網路協定

MAC_t1、MAC_t2…媒體存取控制位址

NH…下一標頭（Next Header）

SVI…交換式虛擬介面

T1、T2…交換器

RD…路由資料庫

RT…路由

VLAN RT List…虛擬區域網路的路由表

VLAN RT…虛擬區域網路路由

VM_IP…主機虛擬機的網際網路協定位址

VM_MAC…主機虛擬機的媒體存取控制位址

【發明申請專利範圍】

【第1項】一種網路運算環境中的第一跳轉閘道的冗餘機制系統，其包括：

在一網路拓樸中的一第一交換器；

在該網路拓樸中的一第二個交換器，其中，該第一交換器以及該第二交換器中的每一個包含與一虛擬第3層介面通信的一交換式虛擬介面（switched virtual interface，SVI），用以將流量路由到其他虛擬第3層介面；

一主機虛擬機，包含與該第一交換器與該第二交換器中的每一個連接的一通信鏈路；以及

連接該第一交換器與該第二交換器的一路由對等鏈路；

其中，該第一交換器與該第二交換器具有相同的網際網路協定（IP）位址與媒體存取控制（MAC）位址；

其中，該路由對等鏈路是在該主機虛擬機與該第一交換器或該第二交換器其中之一之間的該通信鏈路發生故障時用以重定向流量的一修復路徑；以及

其中，該路由對等鏈路使用一個多機箱綁定介面以提供最佳路徑冗餘解決方案。

【第2項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器被配置為該主機虛擬機的一冗餘任播集中式閘道。

【第3項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器配置有代表該主機虛擬機上的一個主埠之一個公共乙太網段標識符（ESI）。

【第4項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器分別配置有每個VLAN的EVPN實例（an EVPN instance per-VLAN），其具有以下一項或多項：

一個自動導出的媒體存取控制-虛擬路由與轉發路由目標（MAC-VRF RT）；或

一個手動配置的MAC-VRF RG。

【第5項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該路由對等鏈路是一個啟用了第3層的對等鏈路。

【第6項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器或該第二交換器中的一個或多個被配置為將該路由對等鏈路IP位址通告為該主機虛擬機的一個下一跳。

【第7項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該系統是使用RT-1保護信號的一個乙太網虛擬私有網路（EVPN）。

【第8項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器通過該路

由對等鏈路交換一每個乙太網段標識符（per-ESI）路由，以信號通知通過該第一交換器與該第二交換器形成的一個冗餘組上的本地ESI連接。

【第9項】根據申請專利範圍第8項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該per-ESI路由是在該主機虛擬機與該第一交換器或該第二交換器其中之一之間的鏈路發生故障時使用的一修復路徑。

【第10項】根據申請專利範圍第8項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該per-ESI路由作為一個邊界閘道協定（BGP）信息在該第一交換器與該第二交換器之間傳輸。

【第11項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器被配置為通過該路由對等鏈路來同步一位址解析協定（ARP）表。

【第12項】根據申請專利範圍第11項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器包括一個或多個處理器，該一個或多個處理器可配置為執行儲存在非暫態電腦可讀取媒體中的多個指令，該些指令包括：

從該主機虛擬機接收一個指示對該ARP表進行更新之信息；以及

通過一個BGP信息將該更新通知該第二交換器。

【第13項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器組成一個

冗餘組，並且該第一交換器與該第二交換器中的一個或多個配置為通告一個乙太網虛擬私有網路（ EVPN）MAC位址，以同步該冗餘組。

【第14項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器組成一個冗餘組，以使流向或來自該主機虛擬機的流量在該第一交換器與該第二交換器之間實現負載平衡。

【第15項】根據申請專利範圍第14項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，通過跨該路由對等鏈路重新路由流量來負載均衡該流量。

【第16項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，還包括該第一交換器與該主機虛擬機之間的一鏈路，其中該鏈路終止於該第一交換器上的一個虛擬區域網路（VLAN）。

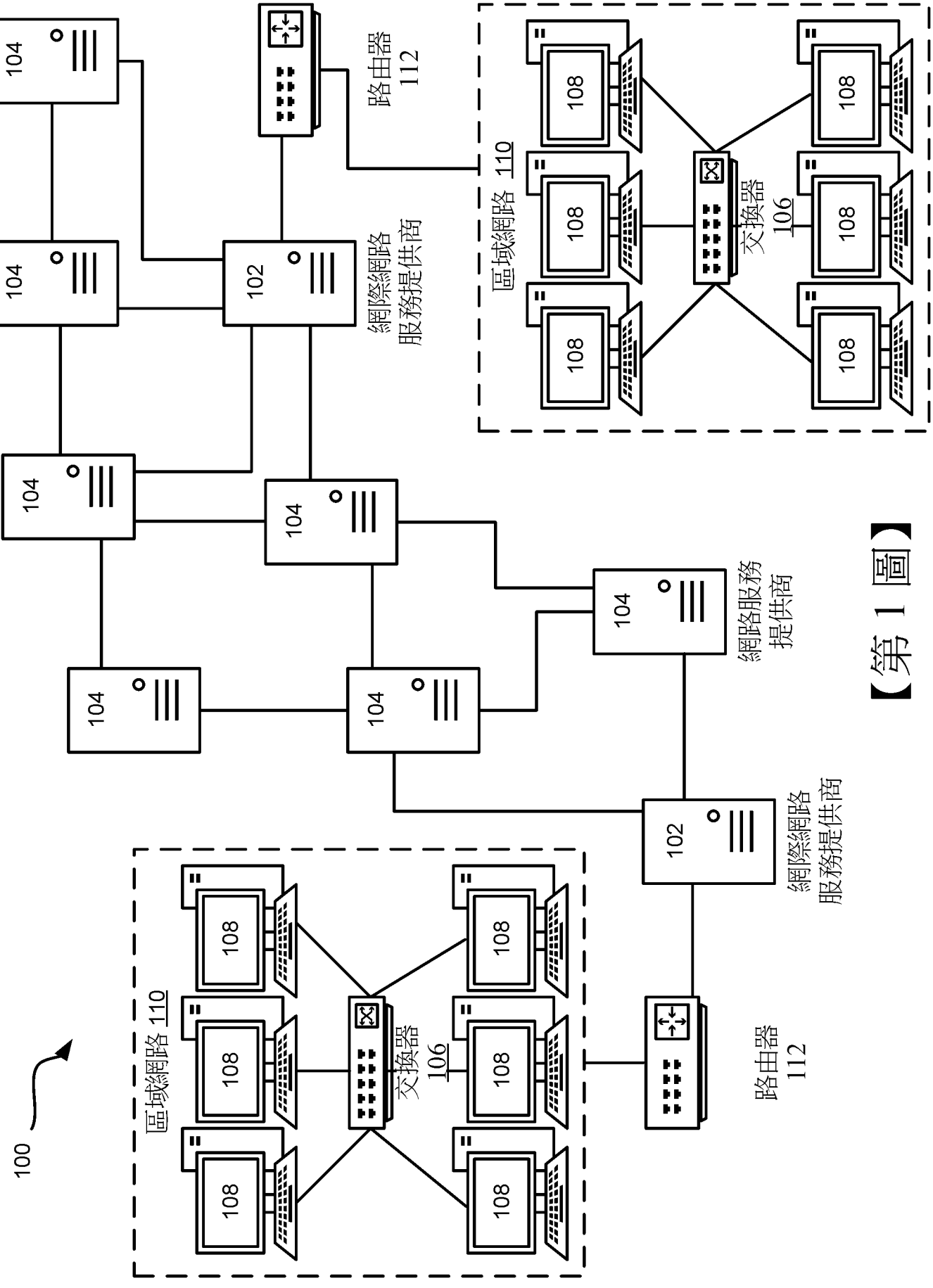
【第17項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器與該第二交換器配置為充當該主機虛擬機的一個虛擬第一跳閘道。

【第18項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，進一步包括該主機虛擬機上的乙太網段標識符（ESI），其中，該第一交換器與該第二交換器通過該路由對等鏈路具有對該ESI的可達性。

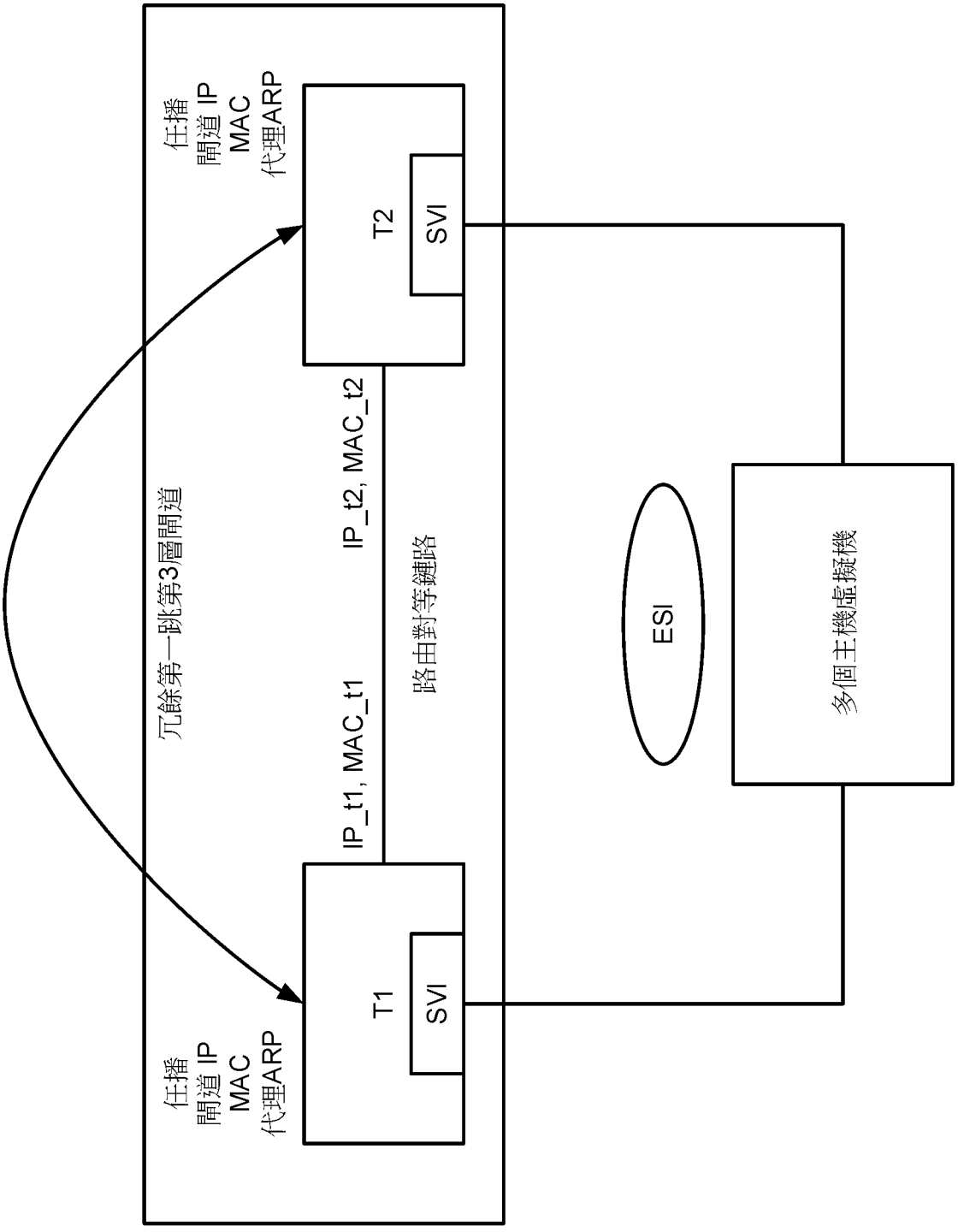
【第19項】根據申請專利範圍第18項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器被配置為向該第二交換器發送一個邊界閘道協定（BGP）信息，該信息指示該第二交換器透過該第一交換器通過一個下一跳具有對該ESI的可達性。

【第20項】根據申請專利範圍第1項所述之網路運算環境中的第一跳轉閘道的冗餘機制系統，其中，該第一交換器被配置為響應於該第一交換機學習對該主機虛擬機上的位址解析協定（ARP）表的更新而經由該路由對等鏈路自動向該第二交換器發送一個邊界閘道協定（BGP）信息。

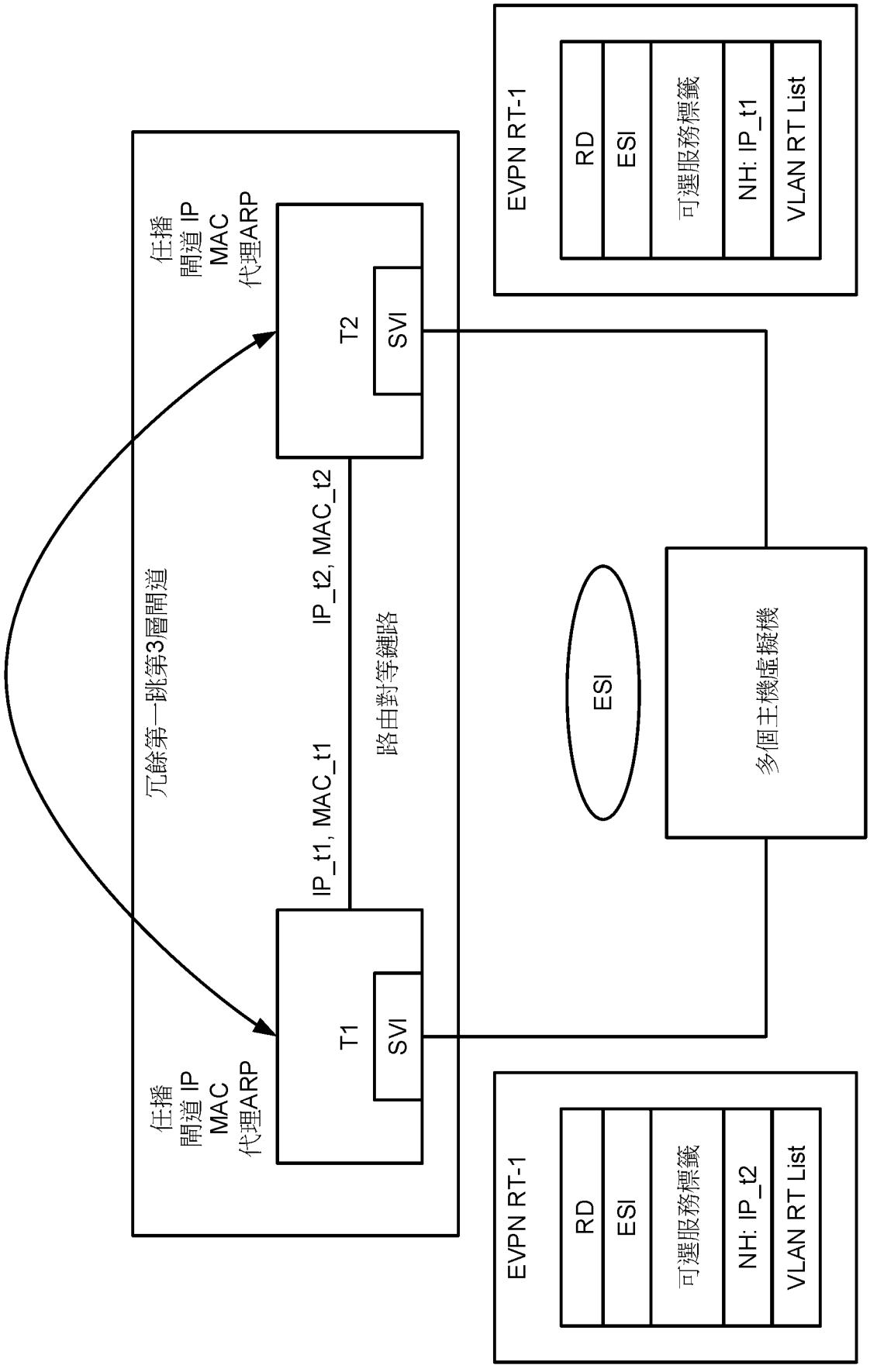
【發明圖式】



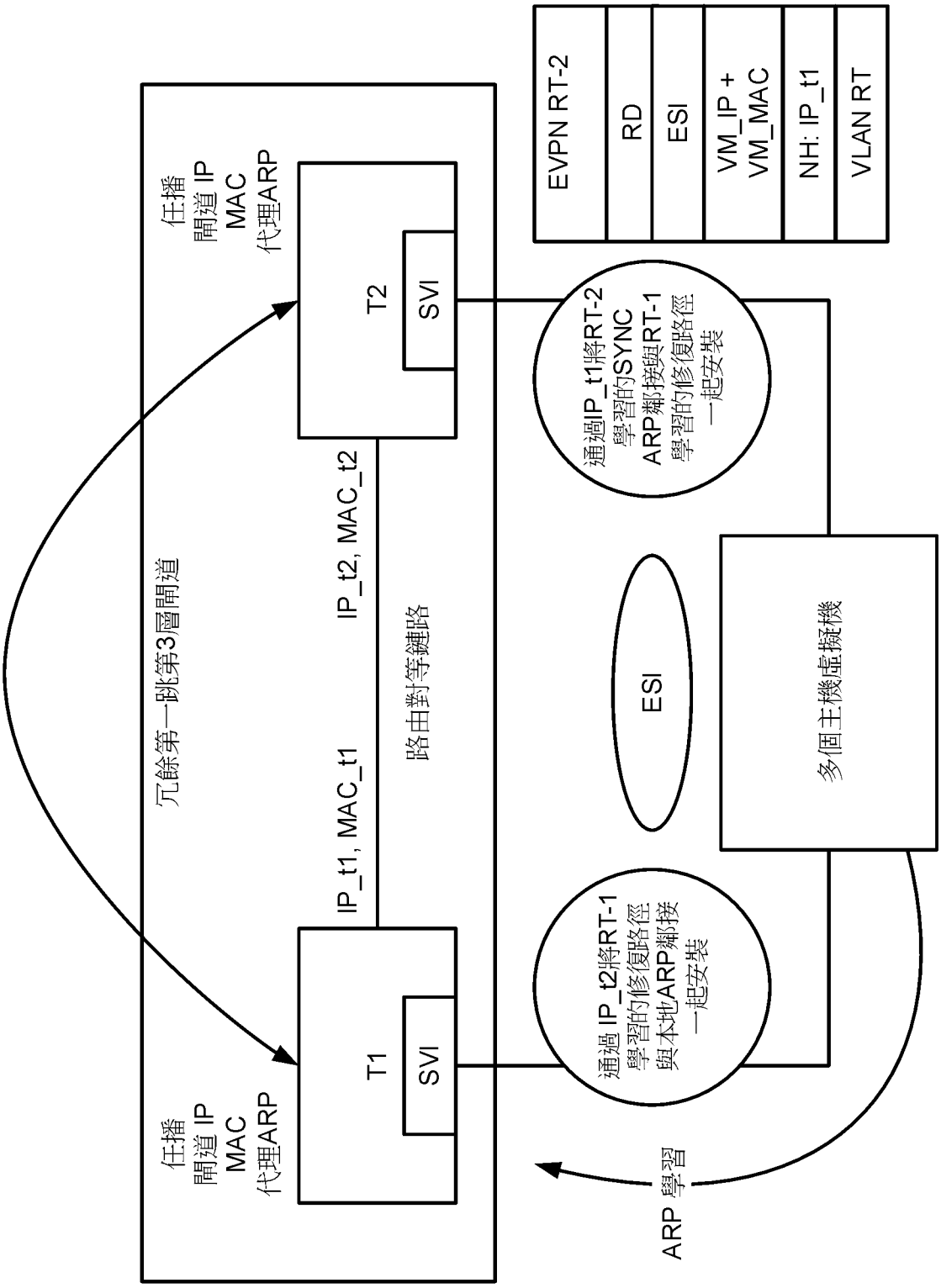
【第 1 圖】


任播
閘道 IP
MAC
代理ARP
T2
SVI
冗餘第一跳第3層閘道
IP_t2, MAC_t2
路由對等鏈路
IP_t1, MAC_t1
ESI
多個主機虛擬機
任播
閘道 IP
MAC
代理ARP
T1
SVI

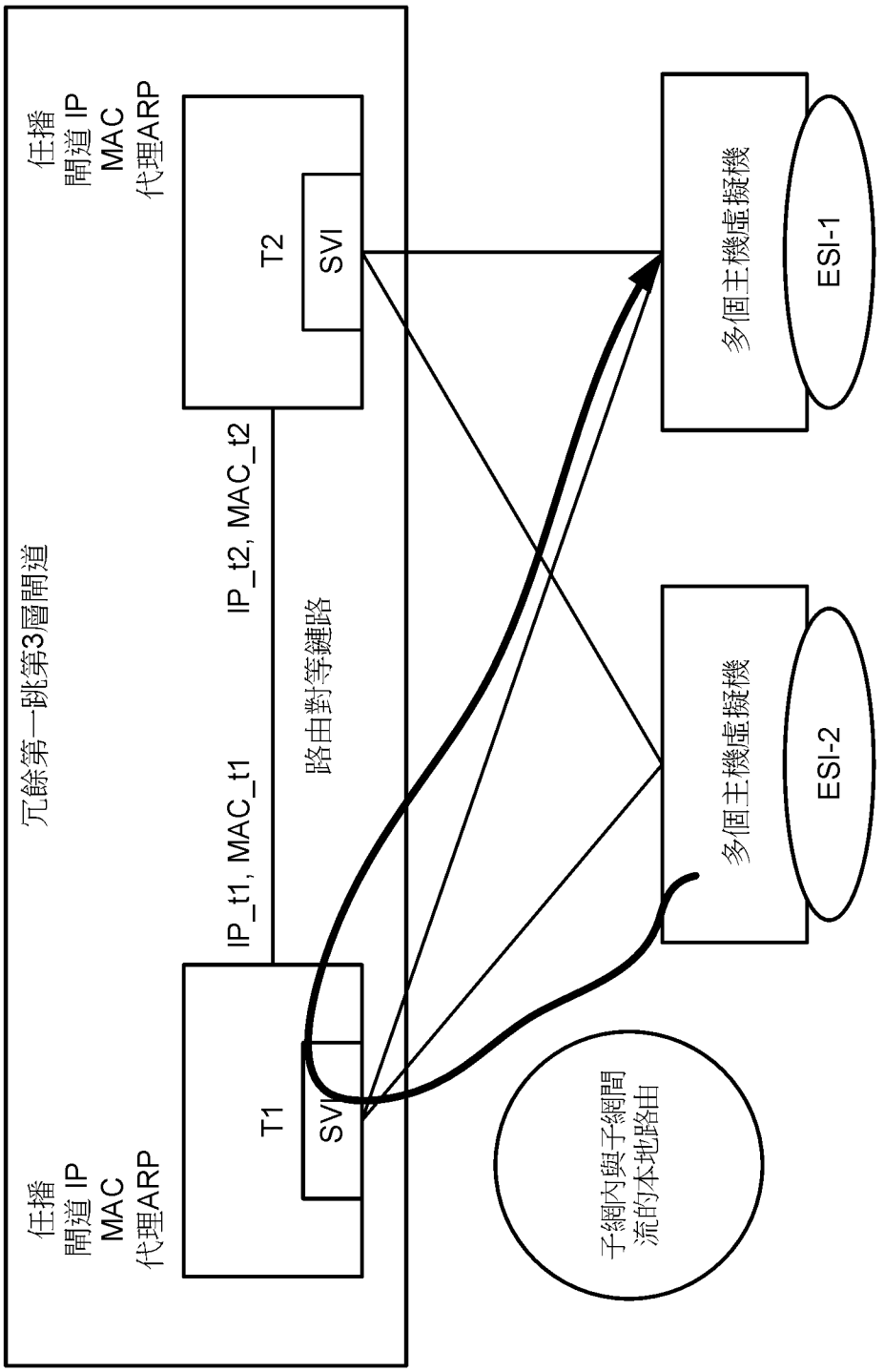
【第 2 圖】


EVPN RT-1
RD
ESI
可選服務標籤
NH: IP_t1
VLAN RT List
任播
閘道 IP
MAC
代理ARP
T2
SVI
冗餘第一跳第3層閘道
IP_t2, MAC_t2
路由對等鏈路
IP_t1, MAC_t1
T1
SVI
任播
閘道 IP
MAC
代理ARP
ESI
多個主機虛擬機
EVPN RT-1
RD
ESI
可選服務標籤
NH: IP_t2
VLAN RT List

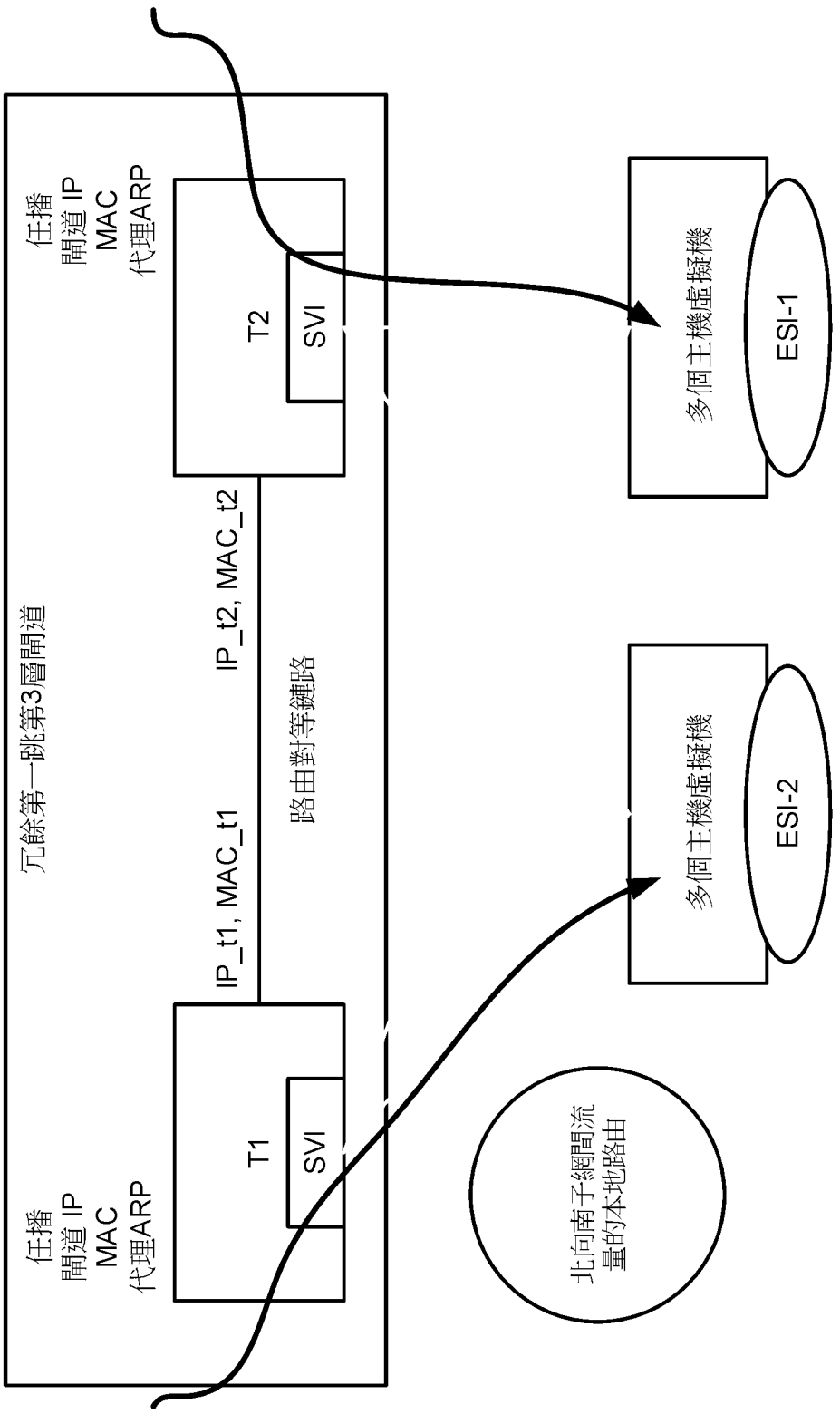
【第 3 圖】


任播
閘道 IP
MAC
代理ARP
T2
SVI
EVPN RT-2
RD
ESI
VM_IP +
VM_MAC
NH: IP_t1
VLAN RT
通過IP_t1將RT-2
學習的SYNC
ARP鄰接與RT-1
學習的修復路徑
一起安裝
冗餘第一跳第3層閘道
IP_t2, MAC_t2
路由對等鏈路
IP_t1, MAC_t1
ESI
多個主機虛擬機
任播
閘道 IP
MAC
代理ARP
T1
SVI
通過 IP_t2將RT-1
學習的修復路徑
與本地ARP鄰接
一起安裝
ARP 學習

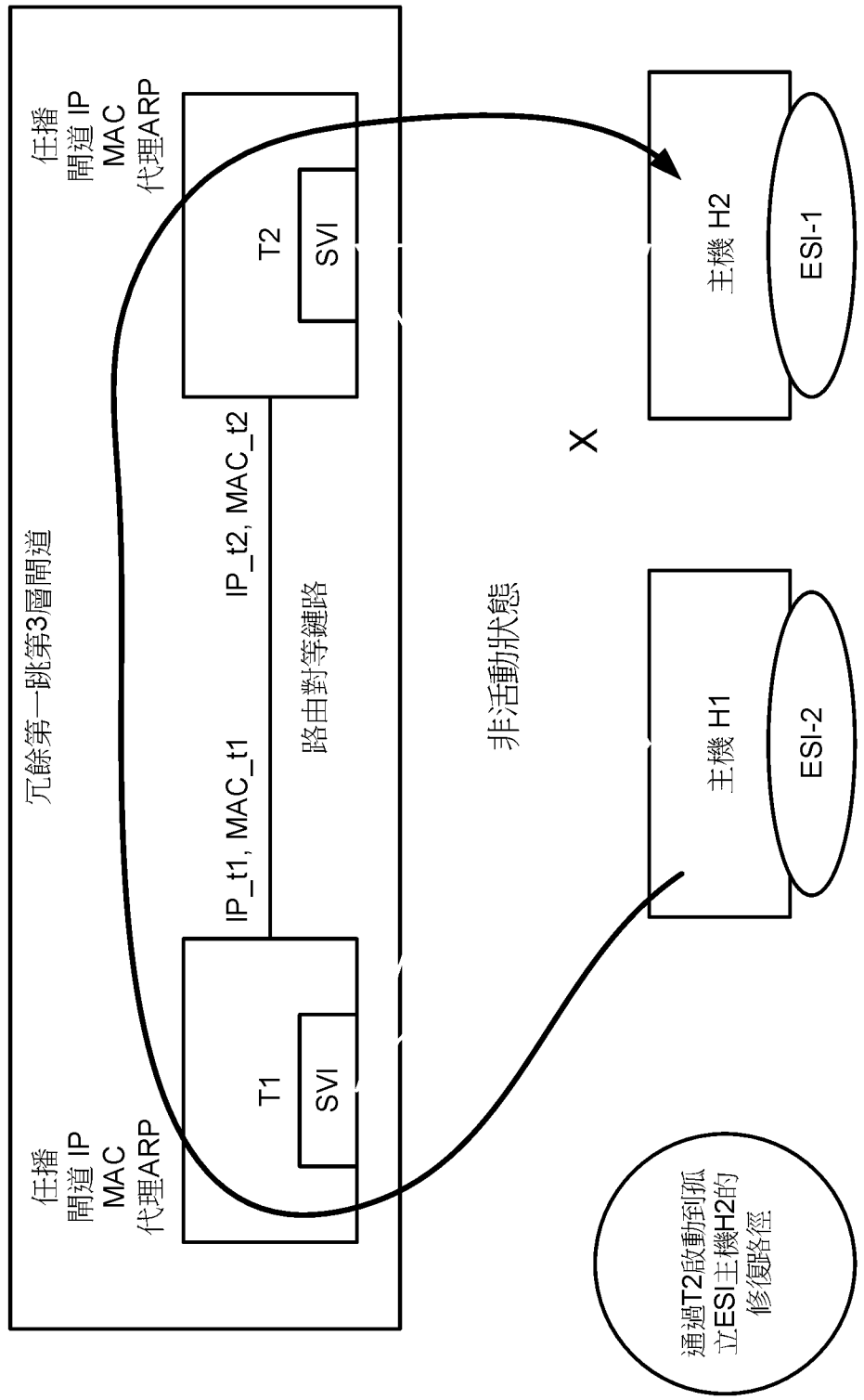
【第 4 圖】


冗餘第一跳第3層閘道
任播
閘道 IP
MAC
代理ARP
T1
SVI
IP_t1, MAC_t1
路由對等鏈路
IP_t2, MAC_t2
T2
SVI
任播
閘道 IP
MAC
代理ARP
多個主機虛擬機
ESI-1
多個主機虛擬機
ESI-2
子網內與子網間
流的本地路由

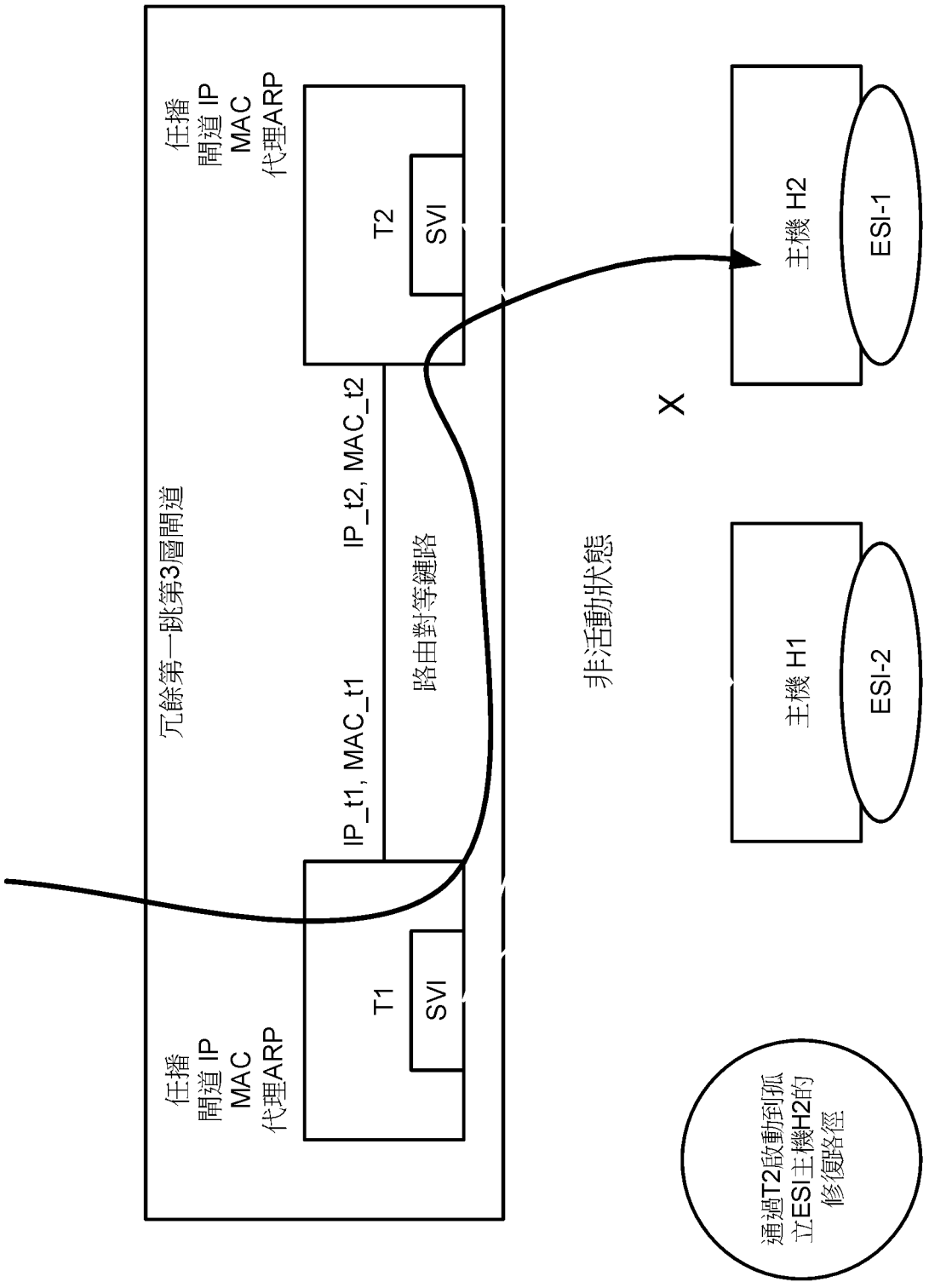
【第 5 圖】


冗餘第一跳第3層閘道
任播
閘道 IP
MAC
代理ARP
T1
SVI
IP_t1, MAC_t1
路由對等鏈路
IP_t2, MAC_t2
T2
SVI
任播
閘道 IP
MAC
代理ARP
多個主機虛擬機
ESI-2
多個主機虛擬機
ESI-1
北向南子網間流量的本地路由

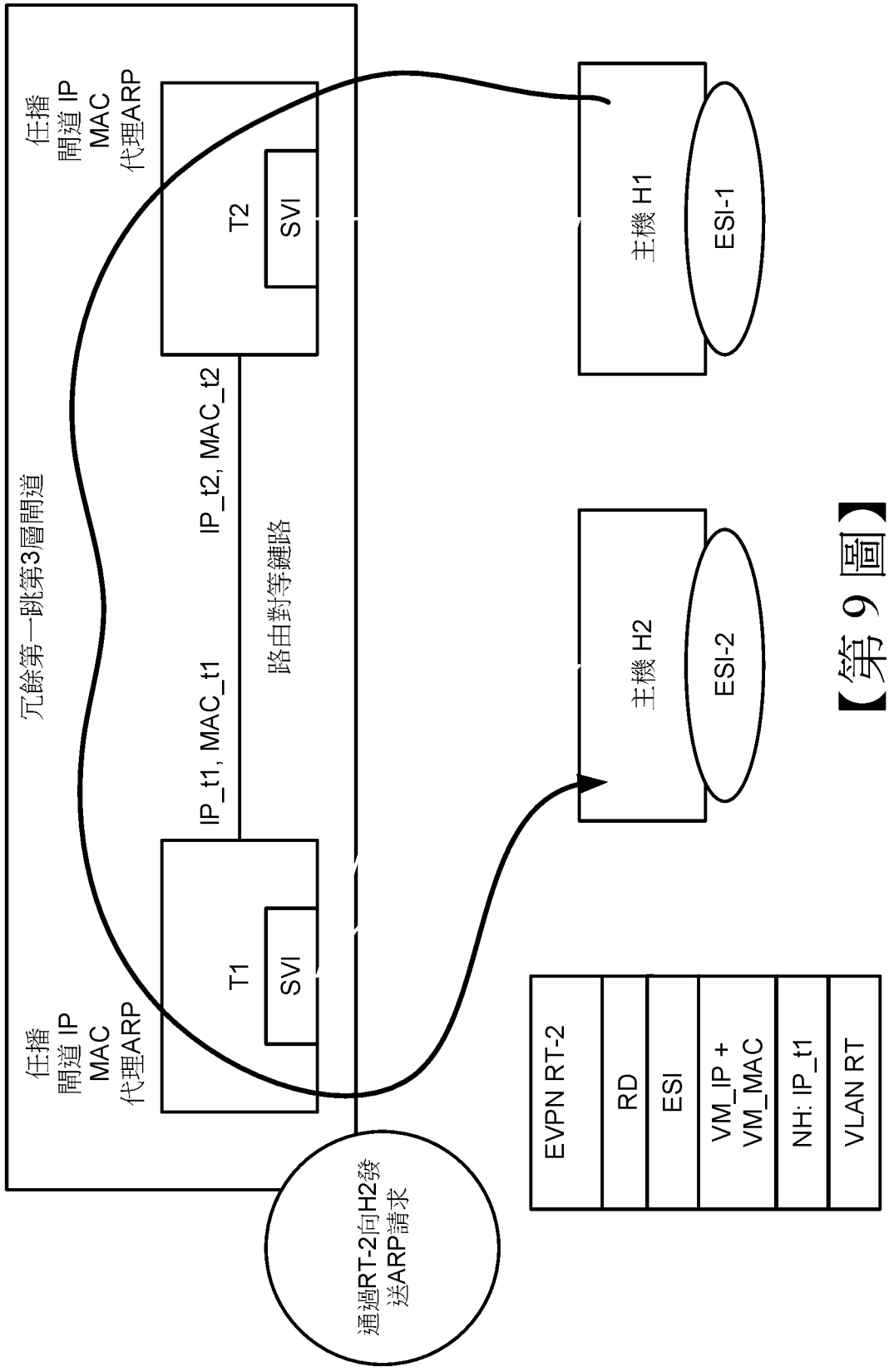
【第 6 圖】


冗餘第一跳第3層閘道
任播
閘道 IP
MAC
代理ARP
T1
SVI
IP_t1, MAC_t1
路由對等鏈路
IP_t2, MAC_t2
T2
SVI
任播
閘道 IP
MAC
代理ARP
非活動狀態
X
主機 H1
ESI-2
主機 H2
ESI-1
通過T2啟動到孤立ESI主機H2的修復路徑

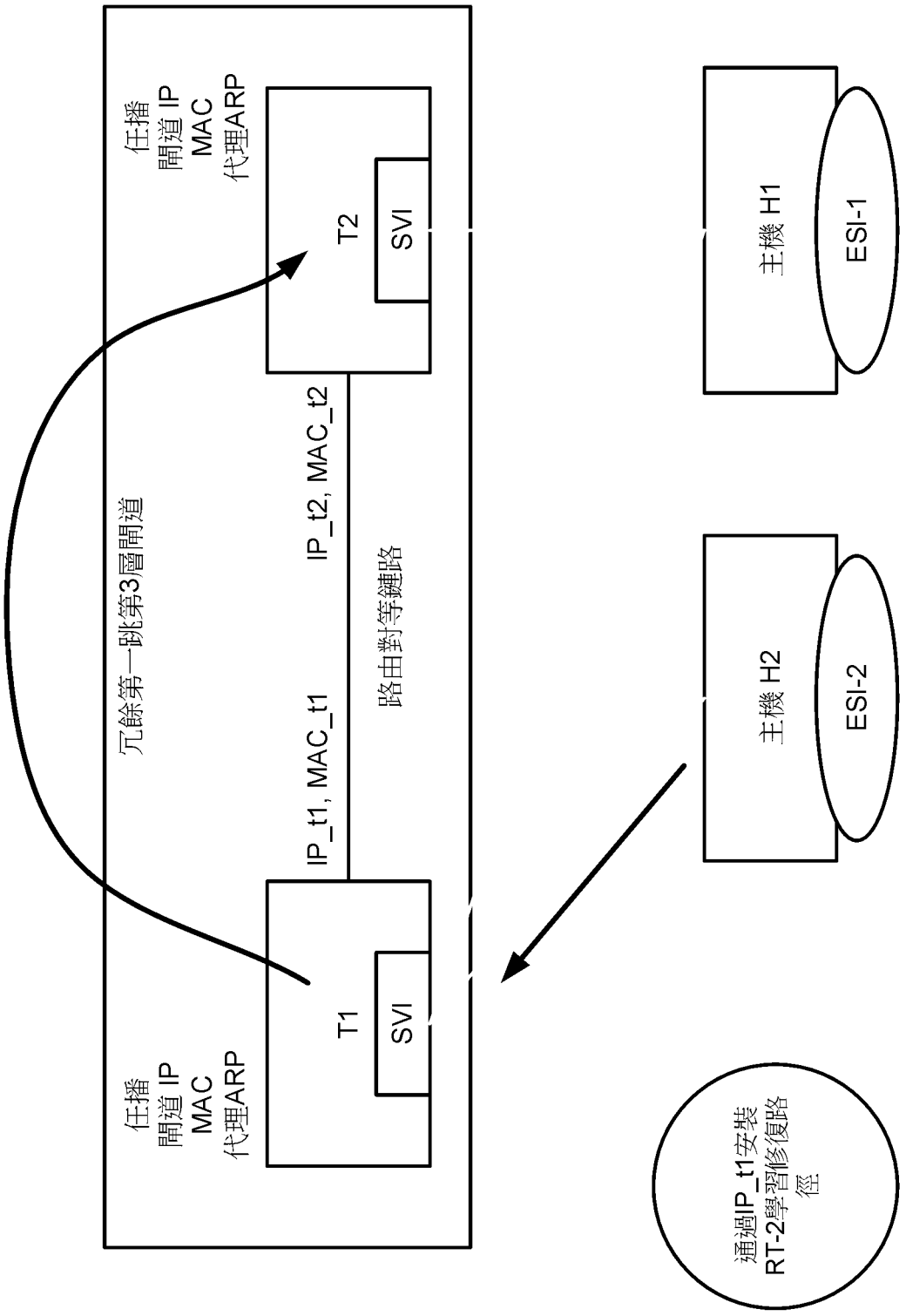
【第 7 圖】


冗餘第一跳第3層閘道
任播
閘道 IP
MAC
代理ARP
T1
SVI
IP_t1, MAC_t1
路由對等鏈路
IP_t2, MAC_t2
T2
SVI
非活動狀態
X
主機 H1
ESI-2
主機 H2
ESI-1
通過T2啟動到孤立ESI主機H2的修復路徑

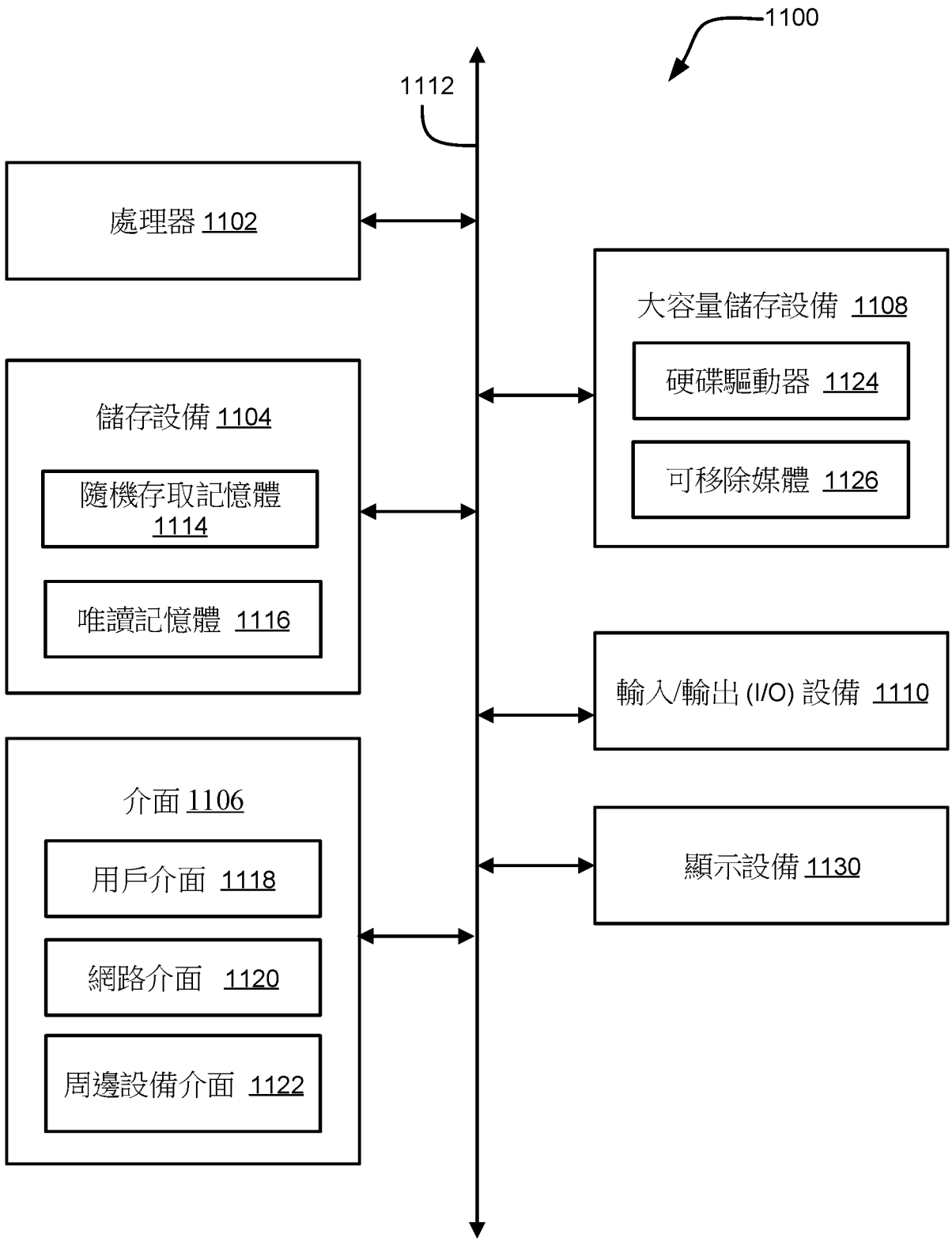
【第 8 圖】


冗餘第一跳第3層閘道
任播
閘道 IP
MAC
代理ARP
T1
SVI
IP_t1, MAC_t1
路由對等鏈路
IP_t2, MAC_t2
T2
SVI
任播
閘道 IP
MAC
代理ARP
主機 H1
ESI-1
主機 H2
ESI-2
通過RT-2向H2發送ARP請求
EVPN RT-2
RD
ESI
VM_IP + VM_MAC
NH: IP_t1
VLAN RT

【第 9 圖】


冗餘第一跳第3層閘道
任播
閘道 IP
MAC
代理ARP
T1
SVI
T2
SVI
IP_t1, MAC_t1
IP_t2, MAC_t2
路由對等鏈路
主機 H1
ESI-1
主機 H2
ESI-2
通過IP_t1安裝
RT-2學習修復路
徑

【第 10 圖】


1100
1112
處理器 1102
儲存設備 1104
隨機存取記憶體 1114
唯讀記憶體 1116
介面 1106
用戶介面 1118
網路介面 1120
周邊設備介面 1122
大容量儲存設備 1108
硬碟驅動器 1124
可移除媒體 1126
輸入/輸出 (I/O) 設備 1110
顯示設備 1130

【第 11 圖】