



(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 I612786 B

(45) 公告日：中華民國 107 (2018) 年 01 月 21 日

(21) 申請案號：105118391

(22) 申請日：中華民國 105 (2016) 年 06 月 13 日

(51) Int. Cl. : H04L12/931 (2013.01)

H04L12/24 (2006.01)

(30) 優先權：2016/04/11 美國

15/096,032

(71) 申請人：廣達電腦股份有限公司 (中華民國) QUANTA COMPUTER INC. (TW)

桃園市龜山區文化二路 188 號

(72) 發明人：董彥屏 TUNG, YEN PIN (TW)

(74) 代理人：祁明輝；林素華；涂綺玲

(56) 參考文獻：

TW 201519607A

US 2012/0110164A1

US 2013/0238802A1

US 2016/0014039A1

審查人員：黃偉倫

申請專利範圍項數：9 項 圖式數：6 共 37 頁

(54) 名稱

節點管理系統、節點管理方法與電腦可讀取儲存裝置

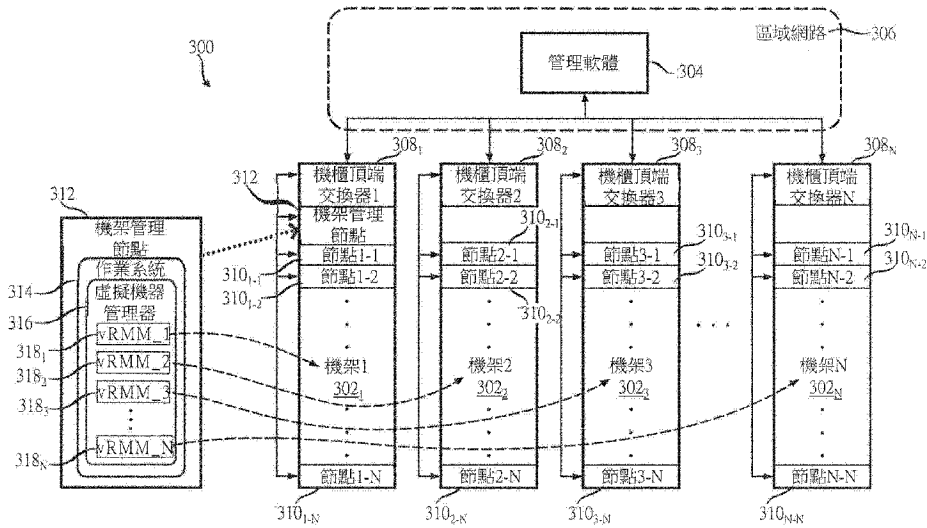
NODES MANAGING SYSTEM, NODES MANAGING METHOD AND COMPUTER-READABLE STORAGE DEVICE

(57) 摘要

一種節點管理系統、節點管理方法與電腦可讀取儲存裝置。節點管理系統可包含一第一機架。第一機架包括一第一機櫃頂端交換器及一第一群組之數個節點。此系統亦可以包含一第二機架。第二機架包括一第二機櫃頂端交換器及一第二群組之數個節點。此節點管理系統可以包括執行一虛擬機器管理器之一機櫃管理節點。虛擬機器管理器可以執行一第一虛擬機架管理模組及一第二虛擬機架管理模組。第一虛擬機架管理模組及第二虛擬機架管理模組可以分別管理第一群組之此些節點及第二群組之此些節點。

A nodes managing system, a nodes managing method and a computer-readable storing device are provided. The nodes managing system can have a first rack that includes a first top-of-rack (ToR) switch and a first group of nodes. The system can also have a second rack that includes a second ToR switch and a second group of nodes. The system can include a rack management node that executes a hypervisor. The hypervisor can run a first virtual rack management module (vRMM) and a second vRMM. The first vRMM and the second vRMM can manage the first group of nodes and the second group of nodes, respectively.

指定代表圖：



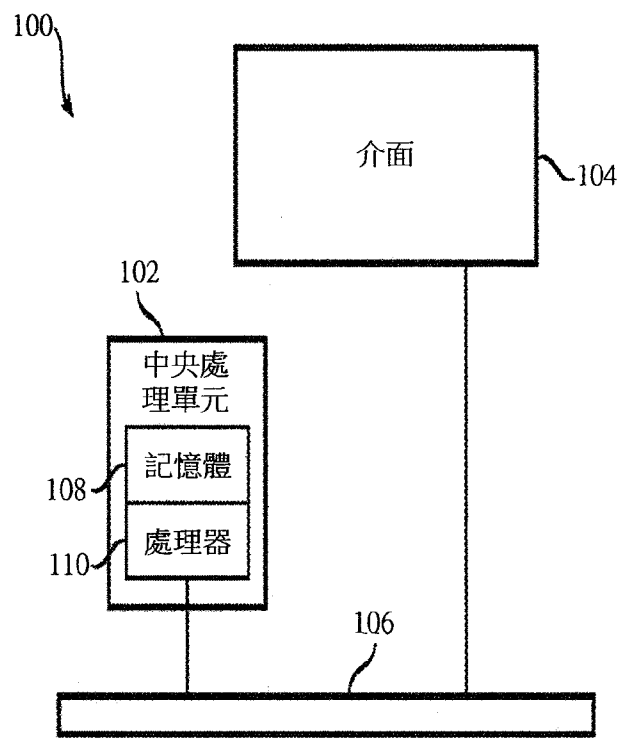
第 3 圖

符號簡單說明：

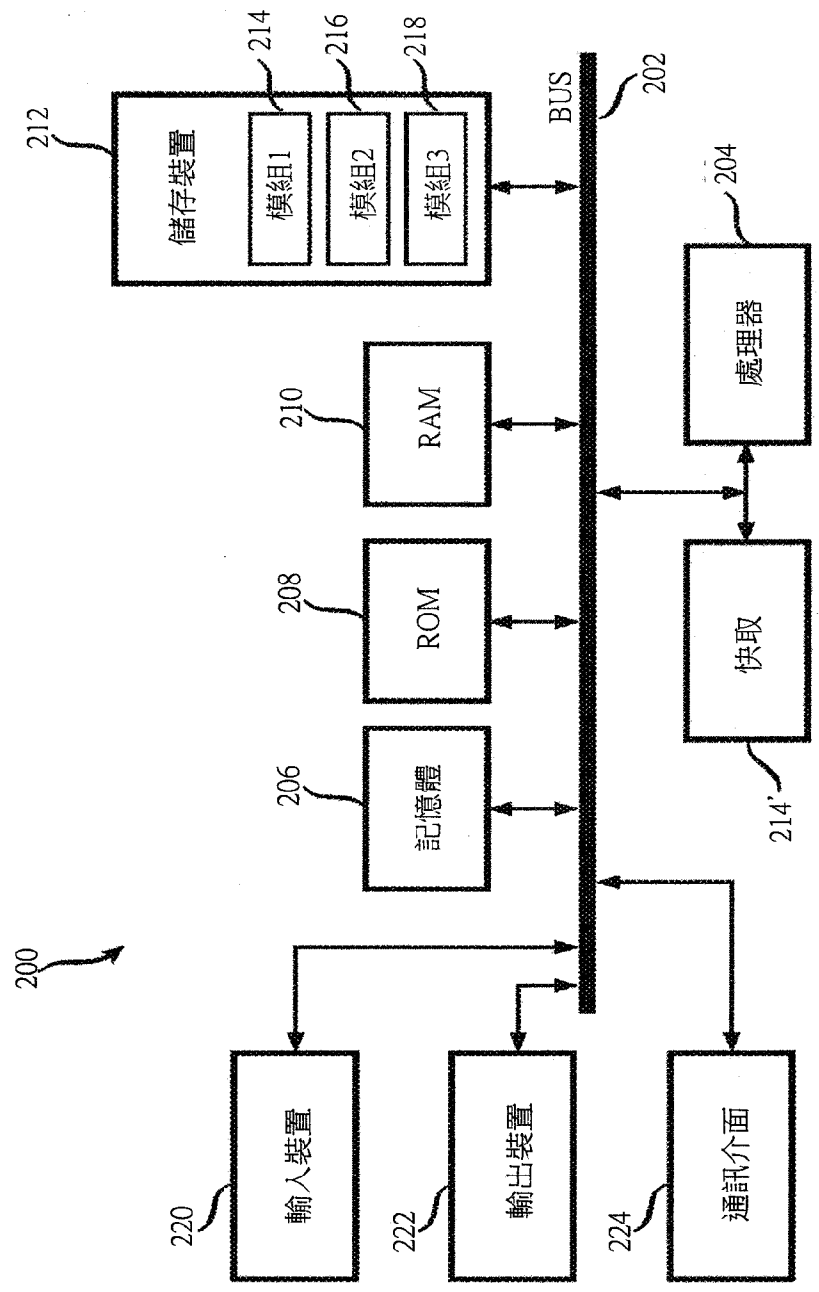
- 300 . . . 資料中心
- 302₁、302₂、302₃、
302_N . . . 機架
- 304 . . . 管理軟體
- 306 . . . 區域網路
- 308₁ . . . 機櫃頂端
交換器 1
- 308₂ . . . 機櫃頂端
交換器 2
- 308₃ . . . 機櫃頂端
交換器 3
- 308_N . . . 機櫃頂端
交換器 N
- 310₁₋₁ . . . 節點 1-1
- 310₁₋₂ . . . 節點 1-2
- 310_{1-N} . . . 節點 1-
N
- 310₂₋₁ . . . 節點 2-1
- 310₂₋₂ . . . 節點 2-2
- 310_{2-N} . . . 節點 2-
N
- 310₃₋₁ . . . 節點 3-1
- 310₃₋₂ . . . 節點 3-2
- 310_{3-N} . . . 節點 3-
N
- 310_{N-1} . . . 節點
N-1
- 310_{N-2} . . . 節點
N-2
- 310_{N-N} . . . 節點 N-
N
- 312 . . . 機架管理節
點
- 314 . . . 作業系統
- 316 . . . 虛擬機器管
理器

- 318₁ . . . vRMM_1
- 318₂ . . . vRMM_2
- 318₃ . . . vRMM_3
- 318_N . . . vRMM_N

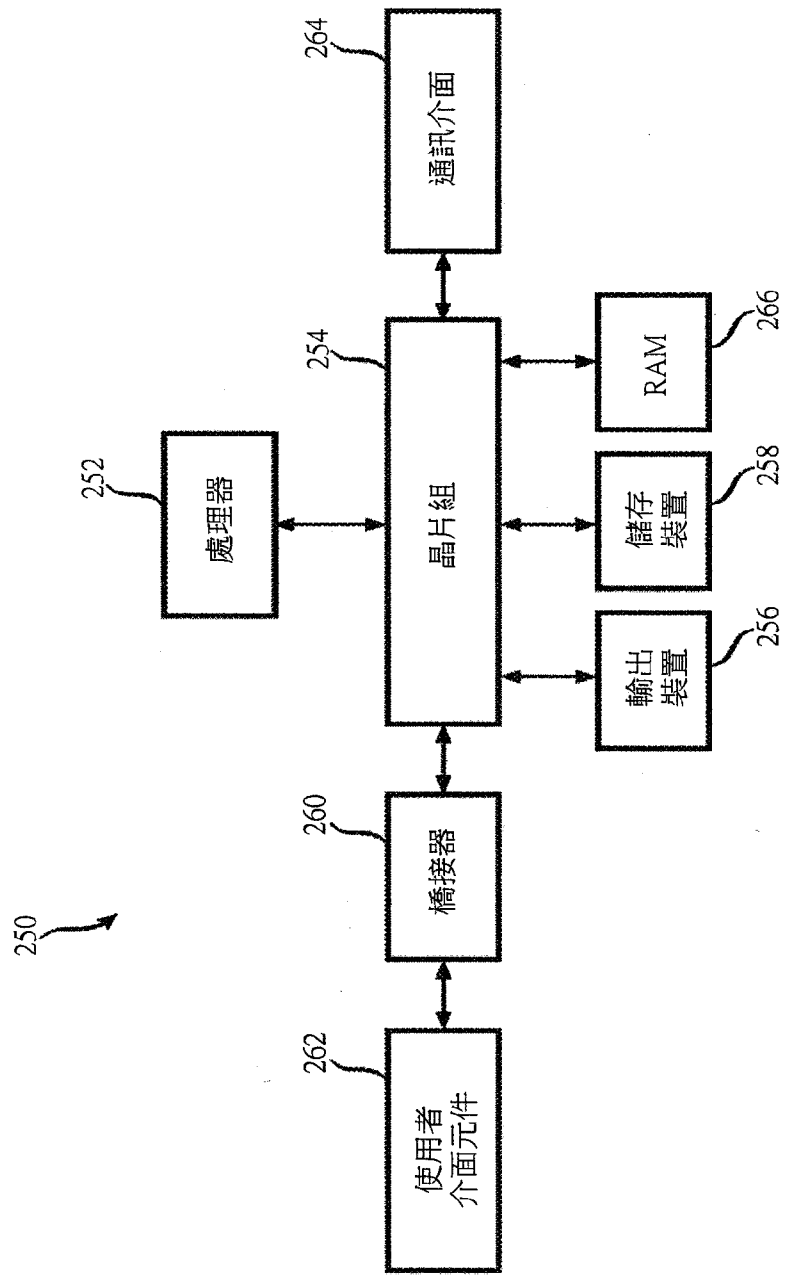
圖式



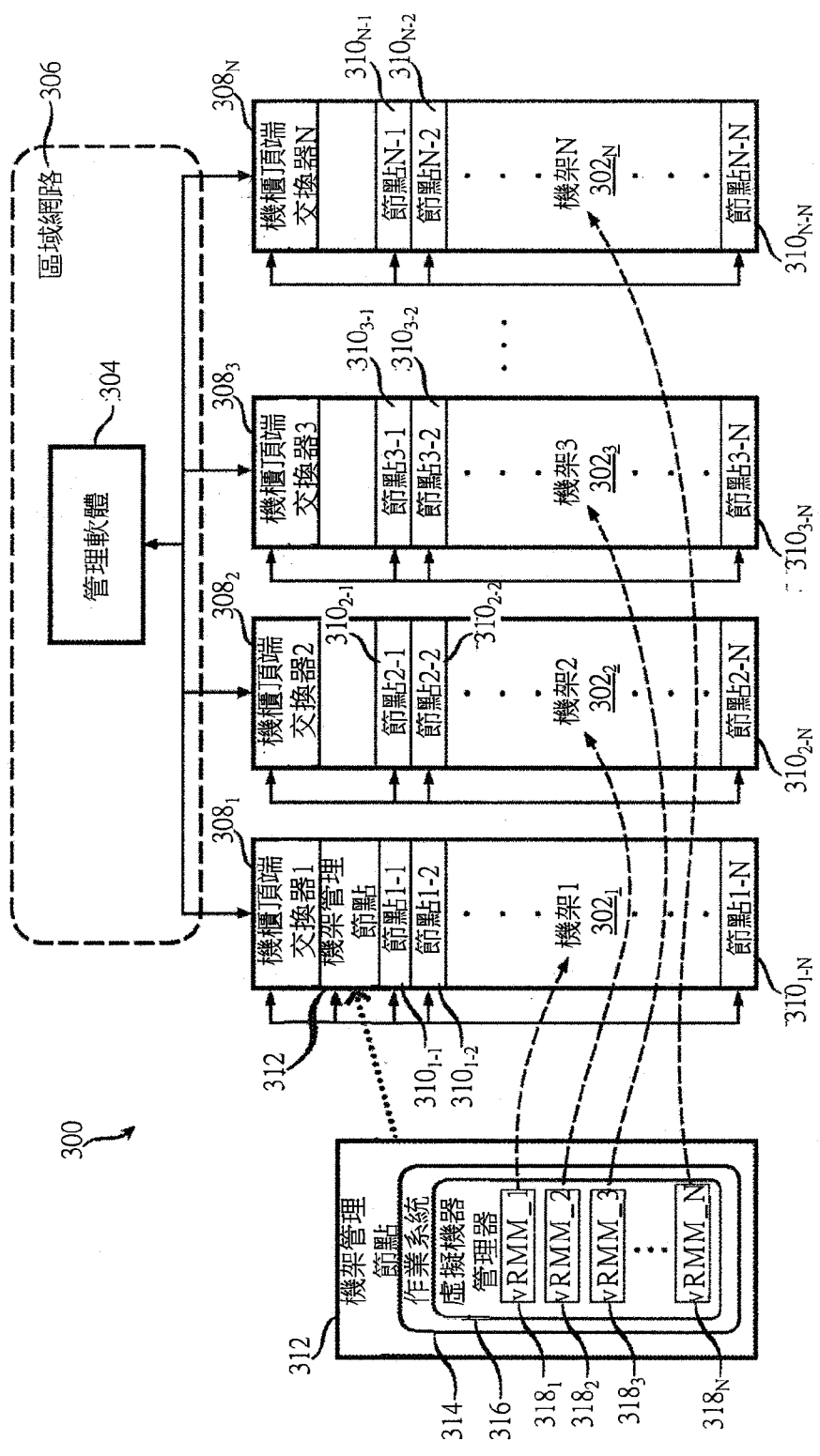
第 1 圖



第 2A 圖



第 2B 圖



第3圖

400

402 404 406

機櫃頂端交換器	位址	虛擬機架管理模組
1	192.168.1.1	1
2	192.168.1.2	8
3	192.168.1.15	未註冊
4	192.168.1.22	4
5	192.168.1.5	未註冊

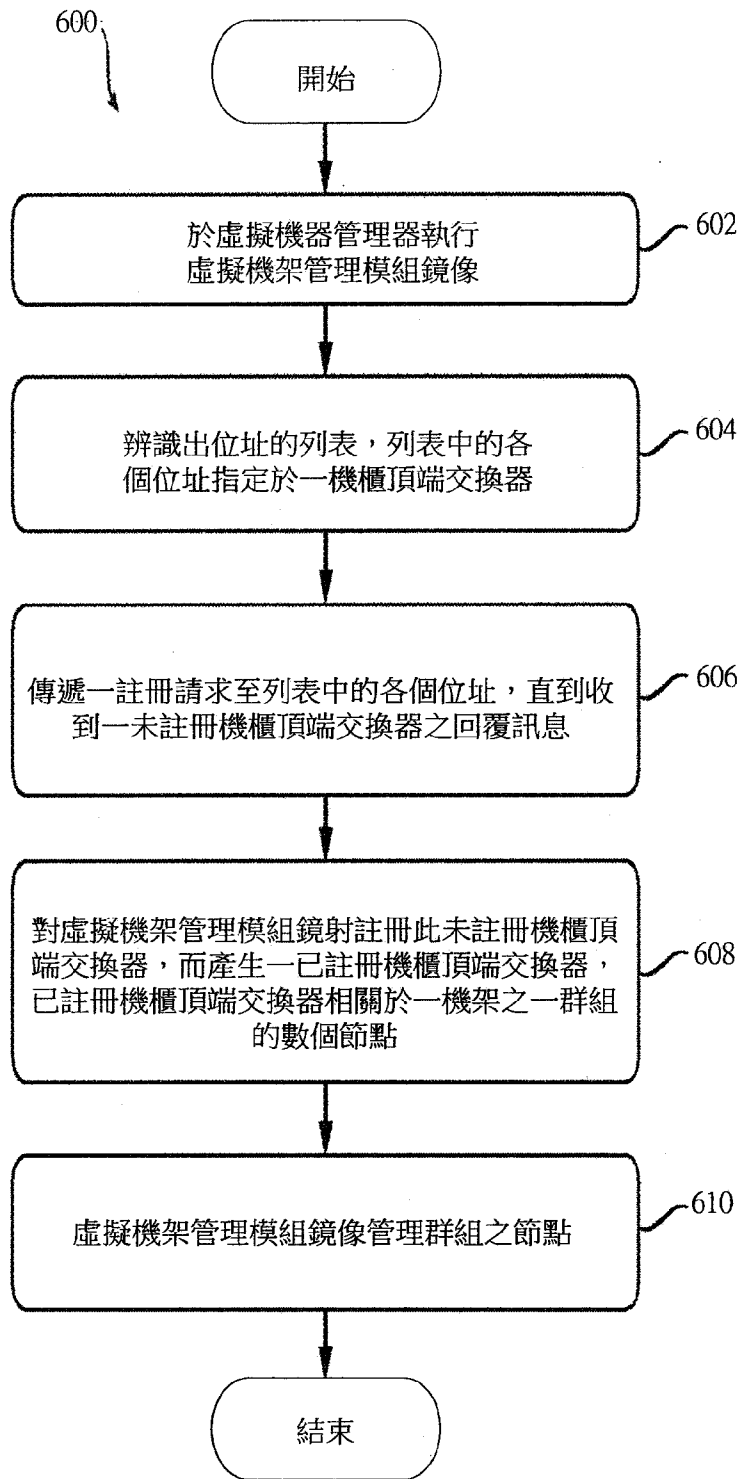
第 4 圖

500

502 504 506 508

節點	主IP	BMC IP位址	位置
1	192.168.1.100	10.2.2.1	1
2	192.168.1.122	10.2.2.2	2
3	192.168.1.114	10.2.2.16	5
⋮	⋮	⋮	⋮
12	192.168.1.183	10.2.2.9	11

第 5 圖



第 6 圖

發明專利說明書

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

節點管理系統、節點管理方法與電腦可讀取儲存裝置
/NODES MANAGING SYSTEM, NODES MANAGING METHOD
AND COMPUTER-READABLE STORAGE DEVICE

【技術領域】

【0001】 本發明是有關於一種管理系統、管理方法與儲存裝置，且特別是有關於一種節點管理系統、節點管理方法與電腦可讀取儲存裝置。

【先前技術】

【0002】 於資料中心，各種伺服器及網路裝置設置於一或多個伺服器機架 (server rack) 上，以利於架設、調整與管理的方便性。設置於一機架 (rack) 上之各個節點 (node) 可以掛接至一機櫃頂端交換器 (top-of-rack switch, ToR switch)，其與各節點設置同一機架上 (通常位於機架之頂端)。各個節點依序由一機架管理模組 (rack management module) 管理。機架管理模組一般也設置於機架內，並負責管理機架內的所有節點。

【0003】 然而，由於機架管理模組需要額外的空間、功耗與佈線，於每個機架設置一個機架管理模組將會花費相當多的成

本。隨著資料中心的規模成長與機架數量的增加，這將成為一項相當困難的任務。更多的機架加入資料中心時，相同數量之機架管理模組也需要被安裝與維護。因此，業界急需一個改良架構來管理機架及機櫃頂端交換器。

【發明內容】

【0004】 本揭露之特徵和優點將在下面的描述中闡述並瞭解，或者可以透過實踐本文公開的原理而獲知。本揭露之特徵與優點可以由申請專利範圍之元件及其組合來實現與獲得。本揭露之其他特徵可以由以下的敘述及請求項來得知，或者可以由下述原理的實踐來得知。

【0005】 所揭露的內容包含利用虛擬機架管理模組（**virtual rack management module**）之系統、方法與電腦可讀取儲存裝置。在一些實施例中，系統可以包含一第一機架（**first rack**）。第一機架包括一第一機櫃頂端交換器（**first top-of-rack switch**）及一第一群組之數個節點（**node**）。第一機櫃頂端交換器可以透過一網路連接於第一群組之此些節點。此系統亦可包含一第二機架（**second rack**）。第二機架包括一第二機櫃頂端交換器（**second top-of-rack switch**）及一第二群組之數個節點。第二機櫃頂端交換器可以透過網路連接於該第二群組之此些節點。第二機櫃頂端交換器也可以透過網路連接於第一機櫃頂端交換器。此系統可以包括一機架管理節點（**rack management node**）。機架管理節點執行一虛擬機

器管理器 (hypervisor)。虛擬機器管理器執行一第一虛擬機架管理模組 (first virtual rack management module) 及一第二虛擬機架管理模組 (second virtual rack management module)。第一虛擬機架管理模組可以透過網路管理第一機架之第一群組的節點。同樣地，第二虛擬機架管理模組可以透過網路管理第二機架之第二群組的節點。

【0006】 虛擬機器管理器可以執行更多的虛擬機架管理模組。這些虛擬機架管理模組管理更多的機架及其節點。機架管理節點可以是架設於這些機架之其中之一之機架掛載伺服器 (rack mount server)，並減少其他機架對於各自的機架管理模組的需求。虛擬機架管理模組可以攔截 (retrieve) 來自對應之機櫃頂端交換器的數個 IP (Internet Protocol) 位址之列表。列表內之各個 IP 位址對應於一節點。IP 位址可以是基板管理控制器 (baseboard management controller, BMC) IP 位址。虛擬機架管理模組可以用以攔截來自機櫃頂端交換器之數個位置的列表。列表內之各個位置對應於一節點。節點可以是伺服器、電腦、或網路裝置等。系統也可以包括執行於虛擬機器管理器之數個虛擬機架管理模組與數個群組之節點的映射關係之列表。虛擬機架管理模組可以攔截及/或更新映射關係的列表。

【0007】 為了對本發明之上述及其他方面有更佳的瞭解，下文特舉較佳實施例，並配合所附圖式，作詳細說明如下：

【圖式簡單說明】**【0008】**

本揭露之各種實施例搭配以下圖式進行說明：

第 1 圖繪示根據本技術一實施例之計算裝置的示意圖。

第 2A 圖及第 2B 圖繪示根據本技術一實施例之系統的示意圖。

第 3 圖繪示根據一實施例具有虛擬機架管理模組之資料中心架構。

第 4 圖繪示根據一實施例之機架管理模組的映射資料表 (mapping table)。

第 5 圖繪示根據一實施例之機架管理資訊 (rack management information) 的表格。

第 6 圖繪示根據一實施例之攔截虛擬機架管理模組之方法。

【實施方式】

【0009】 所揭露的技術解決了上述資料中心的需求。所揭露的內容係為採用虛擬機架管理模組 (virtual rack management module) 來管理伺服器機架 (server rack) 之系統、方法與電腦可讀取儲存媒體。示例性的系統與網路簡介於第 1 圖及第 2A~2B 圖。虛擬機架管理模組及其相關概念與示例性的變化將詳細說明於後。這些變化以各種實施例說明如下。

【0010】 電腦網路是數個節點的集合。個人電腦或工作站等終端透過這些節點的連結傳輸資料。在此，各種型式的網路均可

採用，例如區域網路（local area networks, LANs）、廣域網路（wide area networks, WANs）、軟體自定義網路（software-defined networks）（如虛擬可擴展局域網路（virtual extensible local area networks, VXLANs））。

【0011】 區域網路通常透過於建築物或校園等同一實體區域內之專用的私人通訊連結來連接節點。在另一方面，廣域網路通常透過長距離的通訊連結來連接離散的節點。長距離的通訊連結例如是電路線、光路（optical lightpaths）、同步光纖網（synchronous optical networks, SONET）或同步數位階層連結（synchronous digital hierarchy links, SDH links）。區域網路及廣域網路可以包括第二層（layer 2, L2）及/或第三層（layer 3, L3）網路及裝置。

【0012】 網際網路（Internet）係為廣域網路之一例，其可連接世界各地的網路，而提供各個節點於各種網路進行全球通訊。節點通常根據一預定的協定（protocols）於網路交換訊框（frames）或封包（packets），例如是傳輸控制協定/網際網路協定（Transmission Control Protocol/Internet Protocol, TCP/IP）。在本文中，協定可以是指節點之間如何交流的規則。電腦網路可以是透過一中介網路節點（例如是路由器）來連接，以延伸各個網路的有效範圍。本揭露由第 1 圖開始說起。

【0013】 第 1 圖繪示用以實施本揭露之計算裝置 100。計算裝置 100 包括一主中央處理單元（central processing unit, CPU）102、介面（interfaces）104、及匯流排（bus）106（例如是 PCI

匯流排)。透過適當的軟體或韌體的控制，中央處理單元 102 可用來執行封包管理 (packet management)、錯誤偵測 (error detection) 及/或路由功能 (routing functions) (例如是佈線錯誤偵測功能 (miscabling detection functions))。較佳地，中央處理單元 102 可以透過含有作業系統之軟體及任何適當的應用軟體的正來達成所有的這些功能。中央處理單元 102 可以包括一或多個處理器 (processors) 110，例如是 Motorola 家族微處理器、MIPS 家族微處理器、或 ARM 家族微處理器。在另一實施例中，處理器 110 可以用來控制計算裝置 100 之運作的特製硬體。在一實施例中，一記憶體 (memory) 108 (例如是一非揮發性隨機存取記憶體 (non-volatile RAM)) 及 / 或非揮發性唯讀記憶體 (non-volatile ROM) 也可以是中央處理單元 102 的一部份。然而，記憶體可採用各種方式來附加於系統內。

【0014】 介面 104 通常是介面卡 (interface cards (或稱為線路卡 (line cards)))。一般而言，他們可以控制資料封包於網路的傳遞與接收，並提供計算裝置 100 之周邊設備的支援。介面例如是乙太網路介面 (Ethernet interfaces)、訊框中繼介面 (frame relay interfaces)、纜線介面 (cable interfaces)、數位用戶線路介面 (DSL interfaces)、標記環介面 (token ring interfaces) 及其類似。此外，也可以是各種超高速介面 (very high-speed interfaces)，例如是高速標記環介面 (fast token ring interfaces)、無線網路介面 (wireless interfaces)、乙太網路介面、十億位元乙太網路介面 (Gigabit

Ethernet interfaces)、ATM 介面、高速串列介面(HSSI interfaces)、POS 介面(POS interfaces)、光纖分散式數據介面(FDDI interfaces)及其類似。一般而言，這些介面可以包含對應於各種媒介的適當連接埠。在一些情況下，這些介面也可以含有獨立的處理器及隨機存取記憶體(volatile RAM)。這些獨立的處理器可以控制通訊密集的任务，例如封包交換(packet switching)、媒體控制與管理(media control and management)。藉由提供各自的處理器來進行通訊密集的任务，使得主微處理器 102 能夠有效率地執行路由計算(routing computations)、網路診斷(network diagnostics)或安全功能(security functions)。

【0015】 雖然第 1 圖之系統係為本揭露之一特定網路裝置，但這並不代表這是本揭露唯一能夠實現的架構。舉例來說，採用單一處理器來處理通訊與路由計算也是常用的架構。或者，其他型式的介面與媒介也能夠應用於計算裝置 100。

【0016】 不論網路裝置的架構為何，一或多個記憶體或記憶體模組(包括記憶體 108)可以用來儲存程式指令，來進行路由功能的網路操作與機制。此些程式指令可以共至作業系統及/或多個應用程式。記憶體可以用以儲存各種表格，例如是移動連結表(mobility binding table)、註冊表(registration table)及關聯表(association table)。

【0017】 計算裝置 100 可以是能夠透過網路處理資料的任何裝置。舉例來說，計算裝置 100 可以是一電腦、一桌上型電腦、

一移動計算裝置、一伺服器、一交換器、一路由器、一閘接器 (gateway)、一橋接器 (bridge)、一媒體存取控制器 (media access controller)、一虛擬機器 (virtual machine)、一虛擬機器管理器 (hypervisor)、一基板管理控制器系統 (Baseboard Management Controller system, BMC system)。網路可以是一區域網路 (LAN) 或一廣域網路 (WAN)。

【0018】 第 2A 圖及第 2B 圖繪示系統的各种實施例。熟知此技藝者再實現本技術時可以得知更多適當的實施例。第 2A 圖之系統 200 及/或第 2B 圖之系統 250 可以相同於第 1 圖之系統 100。

【0019】 第 2A 圖說明傳統系統匯流排計算系統架構 (system bus computing system architecture) 200，系統的元件透過匯流排 202 進行溝通。系統 200 包括一處理單元 (processing unit) 204 (CPU 或處理器) 及系統匯流排 202。系統匯流排 202 耦接各種系統元件 (含系統記憶體 206、唯讀記憶體 (ROM) 208 及隨機存取記憶體 (RAM) 210) 至處理器 204。系統 200 可以包括設置於處理器 204 內的快取。系統 200 可以從記憶體 206 及/或儲存裝置 212 複製資料至快取 214'，使處理器 204 能夠以快速存取。透過此方式，快取能夠避免處理器 204 等待資料而發生延遲，以促進效能。各種模組可以控制處理器 204 來執行各種動作。記憶體 206 可以包括具備不同特性的多種型態記憶體。處理器 204 可以包括任何一般目的處理器及硬體模組或軟體模組，例如儲存於儲存裝置 212 的模組 1 (214)、模組 2 (216) 及模組 3 (218) 可

以用來控制處理器 204 及結合軟體指令之特殊目的的處理器。處理器 204 實質上可以是含有多核心或處理器、匯流排、記憶體控制器、快取的完全獨立計算系統。多核心處理器可以是對稱型或不對稱型。

【0020】 為了讓使用者能夠操作計算裝置 200，可採用一輸入裝置 (input device) 220。輸入裝置可以是任何的輸入機構，例如是用以輸入語音的麥克風、用以輸入手勢或圖案的觸控螢幕、鍵盤、滑鼠、動作輸入元件、語音輸入元件等。輸出裝置 (output device) 222 可以是一或多個熟知此技藝者瞭解的輸出機構。在某些實施例中，可採用多型態系統讓使用者提供多種型態的輸入至計算裝置 200。通訊介面 (communications interface) 224 可以管理使用者的輸入及系統輸出。這裡並未特別限制任何特定的硬體設置，可以透過簡單的變換來改善硬體或軟體。

【0021】 儲存裝置 212 係為非揮發記憶體且可以是硬碟或者是其他型式的電腦可讀取記錄媒體例如是 (磁帶盒 (magnetic cassettes)、快閃記憶卡 (flash memory cards)、固態記憶裝置 (solid state memory devices)、數位光碟 (digital versatile disks)、盒式磁帶 (cartridges)、隨機存取記憶 (random access memories, RAMs) 210、唯讀記憶體 (read only memory, ROM) 208 及其組合。

【0022】 儲存裝置 212 可以包括控制處理器 204 之軟體模組 (software modules) 214、216、218。亦可以採用其他的硬體或軟體模組。儲存裝置 212 可以連接於系統匯流排 202。在一方面

中，執行特定功能之硬體模組可以包括儲存於電腦可讀取媒體之軟體元件或連接於必要硬體元件（如處理器 204、匯流排 202、顯示器 222 等）之裝置。

【0023】 第 2B 圖說明具有晶片組架構之電腦系統 250，其可用來執行所述的方法並產生與顯示圖形化使用者介面（**graphical user interface, GUI**）。電腦系統 250 例如是本揭露技術能夠實現的電腦硬體、軟體、及韌體。系統 250 可以包括處理器 252，其可以是實體上或邏輯上分離的各種資源，以執行軟體、韌體及硬體。處理器 252 可以與晶片組 254 進行通訊，晶片組 254 可以控制來自處理器 252 的輸入與輸出。在此例中，晶片組 254 輸出資訊至輸出裝置 256（例如是顯示器），並可以讀取與寫入資訊至儲存裝置 258（可包含磁性媒體及固態媒體）。晶片組 254 也可以從隨機存取記憶體 266 讀取資料並寫入資料至隨機存取記憶體 266。橋接器 260 用以橋接各種使用者介面元件（**user interface components**）262 及晶片組 254。使用者介面元件 262 可以包括鍵盤、麥克風、接觸偵測與處理電路、指點裝置（如滑鼠）。一般來說，系統 250 的輸入可以來自於任何的來源或機器。

【0024】 晶片組 254 可以連接於具有不同實體介面之一或多個通訊介面（**communication interfaces**）264。此些通訊介面可以包括有線或無線區域網路、寬帶無線網路及個人區域網路的介面。產生、顯示及使用圖案化使用者介面之方法的一些應用可以包括透過實體介面接收資料或由處理器 252 分析儲存於儲存裝置

258、266 而產生資料。再者，可以透過使用者介面元件 262 接收使用者之輸入並執行適當的功能。例如可由處理器 252 解讀這些輸入，以瀏覽功能。

【0025】 值得一提的是，系統 200 及 250 可以具有多於一個的處理器 204 或者可以是計算裝置之群組或叢集（cluster）的一部份，此些群組或叢集透過網路連接，以提供更強大的處理能力。

【0026】 第 3 圖繪示具有虛擬機架管理模組之資料中心。資料中心 300 包括一或多個機架 302₁、302₂、302₃、…、302_N、（統稱「302」），其透過網路連接於管理軟體（management software）304 及/或其他網路節點。網路可以是區域網路（LAN）306 或其他型式的網路，例如是廣域網路（WAN）、虛擬可擴展局域網（VXLAN）、同步光纖網（SONET）、同步數位階層（SDH）、儲存區域網路（storage area network, SAN）、或光路（optical lightpaths）。管理軟體 304 可以是資料中心基礎架構管理（data center infrastructure management, DCIM）軟體，其可以管理資料中心之實體上及邏輯上的資產（physical and logical assets）。機架 302 可以是實體的框架（frame）、機殼（enclosures）或機櫃（cabinets），其可以架設各種裝置與設備模組，例如是伺服器、交換器、路由器、閘接器、儲存裝置。此些裝置與模組亦可稱為機架式裝置（rack-mounts）、機架式工具（rack-mount instruments）或機架式系統（rack mounted systems），且可透過螺絲或固定元件實質上設置於機架 302 上。

【0027】 在此實施例中，機架 1 (302₁) 具有機櫃頂端交換器 1 (top-of-rack switch 1, ToR switch 1) (308₁) 及節點 1-1 (310₁₋₁) 至節點 1-N (310_{1-N})。同樣地，機櫃頂端交換器 2 (308₂) 及節點 2-1 至節點 2-N (310₂₋₁、310₂₋₂、…、310_{2-N}) 設置於機架 2 (302₂)；機櫃頂端交換器 3 (308₃) 及節點 3-1 至節點 3-N (310₃₋₁、310₃₋₂、…、310_{3-N}) 設置於機架 3 (302₃)；機櫃頂端交換器 N (308_N) 及節點 N-1 至節點 N-N (310_{N-1}、310_{N-2}、…、310_{N-N}) 設置於節點 N (302_N)。熟知此技藝者可以瞭解機架可以設置任何數量（包括零個）的節點。機架 302 並不需要設置相同數量的節點。機櫃頂端交換器（例如是交換器 308₁、308₂、308₃、…、308_N（統稱「308」））係為一交換器，例如是乙太網路交換器，其設置於機架內並透過網路纜線連接至機架上的所有節點。然而，機櫃頂端交換器 308 不需要實質上設置於機架的頂端，只要連接於機架之多數或全部節點即可。因此，節點 1-1 至節點 1-N 連接於機櫃頂端交換器 1 (308₁) 並可透過機櫃頂端交換器 1 (308₁) 與其他節點工溝通。同樣地，節點 2-1 至節點 2-N 可以透過機櫃頂端交換器 2 (308₂) 與區域網路及其他網路溝通。此外，機櫃頂端交換器 308 可以具有自己的 IP 位址且可以管理所有獨立位址（例如是 IP 位置、BMC IP 位址(基板管理控制器系統 IP 位址)）及機架之所有節點的位置。舉例來說，機櫃頂端交換器 3 (308₃) 可以追蹤節點 3-1 至節點 3-N (302₃₋₁、302₃₋₂、…、302_{3-N}) 的所有 IP 位址、BMC IP 位址及位置（例如是機架 3 (302₃) 之插槽數）。

【0028】 機架管理模組 (rack management module, RMM) 可以是一電腦、系統晶片 (system-on-chip, SoC)、或嵌入式電腦模組 (computer-on-module, CoM)，其設置於或鄰近於機櫃頂端交換器且設置於機架內。舉例來說，機架管理模組可以執行 Linux® 作業系統使用於 x86 處理器電腦之內建的管理服務。機架管理模組可以自機櫃頂端交換器攔截機架上之節點的 BMC IP 位址，以監控各個節點。一般而言，各個機架具有自己的機架管理模組，且管理軟體 304 可以監控與管理機架管理模組。然而，由於各個機架管理模組需要外加的電源、空間及佈線，於各個機架設置機架管理模組將會增加資料中心的成本與複雜度。

【0029】 在資料中心 300 中，不於每一機架 302 設置獨立的機架管理模組，而僅有單一機架管理節點 (rack management node) 312 設置於機架之其中之一 (例如是機架 1 (302₁))。熟知此技藝者可以瞭解機架管理節點 312 可以設置於機架 302 之外。機架管理節點 312 可以是依標準機架掛載伺服器 (standard rack mount server)。再者，機架管理節點 312 可以執行作業系統 314，例如是 Linux® 或 Unix® 作業系統。機架管理節點 312 可以具有執行一或多個虛擬機器 (virtual machines, VMs) 之虛擬機器管理器 (hypervisor) 316。各個虛擬機器可以是基於機架管理模組之虛擬化版本 (virtualized version)。因此，虛擬機器管理器 316 可以執行一或多個虛擬機架管理模組 (virtual rack management modules, vRMMs) 之鏡像 (images)。虛擬機架管理模組例如是

vRMM_1(318₁)、vRMM_2(318₂)、vRMM_3(318₃)、…、vRMM_N(318_N) (統稱「318」)。虛擬機架管理模組 318 可以是實體機架管理模組之虛擬化版本 (virtualized version)，且分別可指定或註冊於機架 302 之其中之一。在此實施例中，vRMM_1(318₁) 註冊於機架 1(302₁)，vRMM_2(318₂) 註冊於機架 2(302₂)，vRMM_3(318₃) 註冊於機架 3(302₃)，…，vRMM_N(318_N) 註冊於機架 N(302_N)。虛擬機架管理模組 318 可以是基於機架管理模組之同一鏡像或者可以是基於兩個或多個不同的鏡像。換句話說，虛擬機架管理模組 318 可以關聯於不同的用戶作業系統、不同的應用程式、不同的驅動程式、不同的服務。

【0030】 當一或多個虛擬機架管理模組 318 被實例化 (instantiated) 且聯機(come online)時，管理軟體 304、機架管理節點 312 或獨立的虛擬機架管理模組可以確認是否有任何的機架 302 未被註冊。若發現有一未註冊機架時，則對其註冊，並管理新註冊之機架。因此，vRMM_2(318₂) 可以發現未註冊機架 (例如是機架 2(302₂))，並與其機櫃頂端交換器 308₂ 進行溝通，以攔截關於節點 310₂₋₁、310₂₋₂、…、310_{2-N} 的資訊 (例如是 IP 位址、BMC IP 位址、位置)。一旦 vRMM_2(318₂) 從機櫃頂端交換器 2(308₂) 攔截到所有必要的資訊，vRMM_2(318₂) 可以開始對機架 2(302₂) 的節點進行管理。因此，當資料中心取得新的機架之節點時，機架管理節點 312 可以虛擬機架管理模組 318 之其中之一至新的機架，或對新機架使用虛擬機架管理模組之鏡像

的新虛擬機器，而不需要修改管理軟體 304 或已存在之虛擬機架管理模組之鏡像。

【0031】 第 4 圖繪示機架管理模組之映射資料表。表 400 可以是示例性的列表或資料結構，其儲存於節點、機架、機櫃頂端交換器、機架管理節點、虛擬機器管理器、虛擬機架管理模組、及/或管理軟體。表 400 可以具有機櫃頂端交換器 402、位址 404、虛擬機架管理模組 406 等欄位。舉例來說，表 400 顯示出機櫃頂端交換器 1 具有的 IP 位址為 192.168.1.1，且註冊於 vRMM 1。表 400 亦顯示出機櫃頂端交換器 3 具有的 IP 位址為 192.168.1.15，且尚未註冊於任何虛擬機架管理模組。當有新的虛擬機架管理模組，或者存某一虛擬機架管理模組移除與先前之機櫃頂端交換器之註冊時，虛擬機架管理模組可以查找表 400，以發現到機櫃頂端交換器 3 或機櫃頂端交換器 5 尚未被註冊。若虛擬機架管理模組之註冊狀態 406 的資訊無法取得，則新的虛擬機架管理模組可以傳送一註冊請求至各個位址 404（每次傳送一個位址），直到機櫃頂端交換器之其中之一回覆一訊息來指出未註冊之交換器。當發現了未註冊機櫃頂端交換器（unregistered ToR switch）及其未註冊機架（unregistered rack），虛擬機架模組可以將其註冊於未註冊機櫃頂端交換器及其未註冊機架，並標註此機櫃頂端交換器已被註冊，以更新表 400。

【0032】 第 5 圖繪示機櫃管理資訊之表格。表 500 係為虛擬機架模組從機櫃頂端交換器攔截到之列表或資料結構。舉例來

說，一旦某一虛擬機架管理模組註冊於一機櫃頂端交換器及其對應之機架的節點，虛擬機架管理模組可以攔截此機架之節點及機櫃頂端交換器的資訊。用來進行機架管理的資訊可以包括機架識別碼（node identifiers）502、IP 位址 504、BMC IP 位址 506、或位置等。節點識別碼 502 可以是數字或字母串，其可以識別出伺服器機架上之節點。節點可以是電腦、伺服器、虛擬機器、交換器、路由器、閘接器或網路裝置等。主/節點 IP 位址 504 可以對應於節點，且虛擬機架管理模組可以使用 IP 位址 504 與對應之節點進行溝通。BMC IP 位址 506 可以邊帶（side-band）或帶外（out-of-band）通道以與對應之節點進行溝通。這些通訊通道可以根據智慧平台管理介面（Intelligent Platform Management Interface, IPMI）個別地專用於管理區域網路連接點。BMC IP 位址 506 可以指定於節點內部之個別的 BMC 模組，並當節點離線或故障時，可以允許虛擬機架管理模組及/或管理軟體去監控節點的狀態。位置 508 可以表示個別節點於機架中的位置。舉例來說，位置 508 可以表示機架上之掛載處或插槽號碼。

【0033】 除了上述的系統元件與概念，以下更進一步以第 6 圖說明方法步驟。為了清楚說明，下述方法將以第 1 圖之系統 100 為例作說明。然而，所述之方法可以由任何軟體、硬體元件、裝置來實現，例如是第 2A 圖之系統 200、第 2B 圖之系統 250、第 3 圖之資料中心 300、第 3 圖之管理軟體 304 第 3 圖之機架管理節點 312、第 3 圖之虛擬機器管理器 316、第 3 圖之虛擬機架管

理模組 318 等等。此處所述之步驟可以採各種順序組合來實現，包括刪減、增加或修改部分步驟。

【0034】 第 6 圖繪示註冊虛擬機架管理模組之方法 600 的示意圖。系統 100 可以於虛擬機器管理器執行虛擬機架管理模組鏡像 (602)。虛擬機器管理器可以執行於一機架上之機架管理節點。機架管理節點可以是一機架掛載伺服器。系統 100 可以辨識出位址的列表。列表中的各個位址指定於一機櫃頂端交換器 (604)。系統 100 可以傳遞一註冊請求至列表中的各個位址，直到收到一未註冊機櫃頂端交換器之回覆訊息 (acknowledgement) (606)。未註冊機櫃頂端交換器回應註冊請求而發出回覆訊息。舉例來說，系統 100 可以對列表中的各個位址送出註冊請求。每一時間點傳送一個請求。當未註冊機櫃頂端交換器和器收到其中之一請求時，將會傳回回覆訊息至系統 100。回覆訊息能夠指出未註冊機櫃頂端交換器之身份。或者，系統 100 可以藉由查找映射資料表來發現未註冊機櫃頂端交換器。映射資料表映射機櫃頂端交換器與虛擬機架管理模組鏡像。

【0035】 系統 100 可以對虛擬機架管理模組鏡射註冊此未註冊機櫃頂端交換器，而產生一已註冊機櫃頂端交換器。已註冊機櫃頂端交換器相對於一機架之一群組的數個節點 (608)。機架可以是機架管理節點所設置之同一個機架，或者是不同的機架。此群組之節點可以包括一伺服器、一電腦、及/或一網路裝置。已註冊機櫃頂端交換器可以設置於同一機架而作為群組之節點。註冊

虛擬機架管理模組鏡像可以包括從未註冊機櫃頂端交換器接收節點位址。各個節點位址可以對應於群組中各自的節點。節點位址可以是基板管理控制器 IP 位址。註冊虛擬機架管理模組鏡像也可以包括從未註冊機櫃頂端交換器接收位置。各個位置對應於機架上之群組的各個節點的各自位置。已註冊機櫃頂端交換器可以於位址的列表中註記已註冊。虛擬機架管理模組鏡像接著可以管理群組之節點（610）。

【0036】 系統 100 可以於虛擬機器管理器執行第二虛擬機架管理模組鏡像。系統 100 可以傳遞一第二註冊請求至列表內之各個位址，直到從一第二未註冊機櫃頂端交換器接收到一第二回覆訊息。系統 100 可以對第二虛擬機架管理模組鏡像註冊此第二未註冊機櫃頂端交換器，以產生一第二已註冊機櫃頂端交換器。第二已註冊機櫃頂端交換器可以是相關於設置於第二機架之第二群組的數個節點。第二虛擬機架管理模組鏡像接著可以管理第二群組之節點。

【0037】 更清楚的說，在一些實施例中，本技術可以採用獨立功能模塊，包括於軟體、或軟硬體中之元件與方法步驟。電腦可讀取儲存裝置及媒體可以排除能量、負載訊號、電磁波及訊號等暫態媒體。

【0038】 上述各實施例之方法可以透過儲存於或取得於電腦可讀取媒體或裝置的電腦可執行指令。這些指令例如是包括能夠讓一般電腦、特殊電腦或特殊處理裝置去執行一些功能的指令及

資料。一些電腦資源可以透過網路取得。舉例來說，電腦可執行指令可以是二位元、中間格式指令（*intermediate format instructions*）（如組合語言（*assembly language*）、韌體、或原始碼（*source code*）。電腦可讀取媒體可以用來儲存指令、使用的資訊、及/或方法執行過程所產生之資訊，例如是磁片或光碟片、快閃記憶體、採用非揮發記憶體之 **USB** 裝置、網路儲存裝置等。

【0039】 實現上述方法的裝置可以包括硬體、韌體及/或軟體，且可以採用各種變化。例如可採用筆記型電腦、智慧型電話、小型個人電腦、個人數位助理（*personal digital assistants*）、機架掛載裝置、獨立裝置（*standalone devices*）等。各種功能可以透過周邊裝置或外接卡（*add-in cards*）來實現。各種功能也可以透過設置不同晶片之電路板。

【0040】 指令、攜帶指令之媒體、執行的計算資源及其他結構用來提供本揭露之功能。

【0041】 雖然各種實施例之變化及其他資訊係用來解釋申請專利範圍，但這些申請專利範圍不應該被侷限於這些實施例的安排。任何熟知此技術者接能夠透過這些實施例來提供更寬廣的應用。並且，雖然此處已敘述一些結構特徵及/或方法步驟，但可以瞭解的是申請專利範圍並不侷限於這些特徵或動作。舉例來說，可以採用不同的功能，或以不同元件來執行。再者，實施例之系統與方法所述之特徵與步驟僅僅是作為示例。此外，申請專利範圍所述之「至少一」包括一個或多個。

【0042】 實施例所述之這些特徵或配置可以實現或結合其他實施例。「實施例」、「變化」、「方面」、「例子」、「配置」、「實施」、「示例」及其他用詞僅僅是用來說明各種實施例，並未侷限為一個特定實施例或獨立的個別實施例，也沒有限定不能組合於其他實施例、變化、與配置。換句話說，一特定實施例之例子可以與其他實施例結合。更精確的說，本發明所屬技術領域中具有通常知識者均可瞭解這裡所述的各種實施例及其相關特徵可以相互結合。

【0043】 「方面」一詞並非指這是本技術所必要的內容。一某方面的敘述可以應用於所有的配置、或一或多種配置。「配置」「一方面」可以是指一或多個方面。「配置」並非是指這是本技術所必要的內容。一配置可以應用於所有配置、或一或多個配置。「示例」一詞是指用來說明或解釋的範例。任何於示例中所述的内容均並非用以限制本發明。再者，申請專利範圍所述之「至少一」係指一或多個。

【0044】 綜上所述，雖然本發明已以較佳實施例揭露如上，然其並非用以限定本發明。本發明所屬技術領域中具有通常知識者，在不脫離本發明之精神和範圍內，當可作各種之更動與潤飾。因此，本發明之保護範圍當視後附之申請專利範圍所界定者為準。

【符號說明】

【0045】

- 100：計算裝置
- 102：中央處理單元
- 104：介面
- 106：匯流排
- 108：記憶體
- 110：處理器
- 200：系統
- 202：匯流排
- 204：處理器
- 206：記憶體
- 208：唯讀記憶體
- 210：隨機存取記憶體
- 212：儲存裝置
- 214：模組 1
- 214'：快取
- 216：模組 2
- 218：模組 3
- 220：輸入裝置
- 222：輸出裝置
- 224：通訊介面
- 250：系統
- 252：處理器
- 254：晶片組

- 256：輸出裝置
- 258：儲存裝置
- 260：橋接器
- 262：使用者介面元件
- 264：通訊介面
- 266：隨機存取記憶
- 300：資料中心
- 302₁：機架 1
- 302₂：機架 2
- 302₃：機架 3
- 302_N：機架 N
- 304：管理軟體
- 306：區域網路
- 308₁：機櫃頂端交換器 1
- 308₂：機櫃頂端交換器 2
- 308₃：機櫃頂端交換器 3
- 308_N：機櫃頂端交換器 N
- 310₁₋₁：節點 1-1
- 310₁₋₂：節點 1-2
- 310_{1-N}：節點 1-N
- 310₂₋₁：節點 2-1
- 310₂₋₂：節點 2-2
- 310_{2-N}：節點 2-N
- 310₃₋₁：節點 3-1

- 310₃₋₂：節點 3-2
- 310_{3-N}：節點 3-N
- 310_{N-1}：節點 N-1
- 310_{N-2}：節點 N-2
- 310_{N-N}：節點 N-N
- 312：機架管理節點
- 314：作業系統
- 316：虛擬機器管理器
- 318₁：vRMM_1
- 318₂：vRMM_2
- 318₃：vRMM_3
- 318_N：vRMM_N
- 400：表
- 402：機櫃頂端交換器
- 404：位址
- 406：虛擬機架管理模組
- 500：表
- 502：機架識別碼
- 504：IP 位址
- 506：BMC IP 位址
- 508：位置
- 600：方法
- 602、604、606、608、610：流程步驟

發明摘要

※ 申請案號：105118391

※ 申請日：105/06/13

※IPC 分類：H04L 12/931 (2013.01)
H04L 12/24 (2006.01)

【發明名稱】(中文/英文)

節點管理系統、節點管理方法與電腦可讀取儲存裝置
/NODES MANAGING SYSTEM, NODES MANAGING METHOD
AND COMPUTER-READABLE STORAGE DEVICE

【中文】

一種節點管理系統、節點管理方法與電腦可讀取儲存裝置。節點管理系統可包含一第一機架。第一機架包括一第一機櫃頂端交換器及一第一群組之數個節點。此系統亦可以包含一第二機架。第二機架包括一第二機櫃頂端交換器及一第二群組之數個節點。此節點管理系統可以包括執行一虛擬機器管理器之一機櫃管理節點。虛擬機器管理器可以執行一第一虛擬機架管理模組及一第二虛擬機架管理模組。第一虛擬機架管理模組及第二虛擬機架管理模組可以分別管理第一群組之此些節點及第二群組之此些節點。

【英文】

A nodes managing system, a nodes managing method and a computer-readable storing device are provided. The nodes

managing system can have a first rack that includes a first top-of-rack (ToR) switch and a first group of nodes. The system can also have a second rack that includes a second ToR switch and a second group of nodes. The system can include a rack management node that executes a hypervisor. The hypervisor can run a first virtual rack management module (vRMM) and a second vRMM. The first vRMM and the second vRMM can manage the first group of nodes and the second group of nodes, respectively.

【代表圖】

【本案指定代表圖】：第 3 圖。

【本代表圖之符號簡單說明】：

300：資料中心

302₁、302₂、302₃、302_N：機架

304：管理軟體

306：區域網路

308₁：機櫃頂端交換器 1

308₂：機櫃頂端交換器 2

308₃：機櫃頂端交換器 3

308_N：機櫃頂端交換器 N

310₁₋₁：節點 1-1

310₁₋₂：節點 1-2

310_{1-N}：節點 1-N

申請專利範圍

1. 一種節點管理系統，包括：

一第一機架（**first rack**），包括一第一機櫃頂端交換器（**first top-of-rack switch**）及一第一群組之複數個節點（**node**），該第一機櫃頂端交換器透過一網路連接於該第一群組之該些節點；

一第二機架（**second rack**），包括一第二機櫃頂端交換器（**second top-of-rack switch**）及一第二群組之複數個節點，該第二機櫃頂端交換器透過該網路連接於該第二群組之該些節點，且該第二機櫃頂端交換器透過該網路連接於該第一機櫃頂端交換器；以及

一機架管理節點（**rack management node**）執行一虛擬機器管理器（**hypervisor**），其中該虛擬機器管理器執行一第一虛擬機架管理模組（**first virtual rack management module**）及一第二虛擬機架管理模組（**second virtual rack management module**），該第一虛擬機架管理模組管理該第一機架之該第一群組的該些節點，且該第二虛擬機架管理模組管理該第二機架之該第二群組的該些節點。

2. 如申請專利範圍第 1 項所述之節點管理系統，其中該機架管理節點僅設置於該第一機架上，且該機架管理節點包括一機架掛載伺服器（**rack mount server**）。

3. 如申請專利範圍第 1 項所述之節點管理系統，其中該第一虛擬機架管理模組用以攔截 (retrieve) 該第一機櫃頂端交換器之複數個 IP (Internet protocol) 位址之一列表，該列表之各該 IP 位址對應於該第一群組之該些節點之其中之一。

4. 如申請專利範圍第 3 項所述之節點管理系統，其中各該 IP 位址包括一基板管理控制器 IP 位址 (baseboard management controller IP address)。

5. 如申請專利範圍第 1 項所述之節點管理系統，其中該第一虛擬機架管理模組用以攔截 (retrieve) 該第一機櫃頂端交換器之複數個位置 (location) 之一列表，該列表之各該位置對應於該第一群組之該些節點之其中之一。

6. 一種節點管理方法，包括：

於一虛擬機器管理器 (hypervisor) 執行一虛擬機架管理模組鏡像 (virtual rack management module image)；

辨識出複數個位置之一列表，該列表之各該位址相關於對應之一機櫃頂端交換器 (top-of-rack switch)；

傳送註冊請求 (registration request) 至該列表之各該位址，直到收到來自一未註冊機櫃頂端交換器 (unregistered top-of-rack switch) 之一確認訊息 (acknowledgement)；

於該未註冊機櫃頂端交換器(**unregistered top-of-rack switch**) 註冊該虛擬機架管理模組鏡像，以產生一已註冊機櫃頂端交換器 (**registered top-of-rack switch**)，其中該已註冊機櫃頂端交換器相關於設置於一機架 (**rack**) 之一群組的複數個節點 (**node**)；以及藉由該虛擬機架管理模組鏡像管理該群組之該些節點；其中註冊該虛擬機架管理模組鏡像之步驟包括：自該未註冊機櫃頂端交換器接收複數個節點位址 (**node address**)，各該節點位址對應於該群組之該些節點之其中之一。

7. 如申請專利範圍第 6 項所述之節點管理方法，其中註冊該虛擬機架管理模組鏡像之步驟包括：

自該未註冊機櫃頂端交換器接收複數個位置 (**location**)，各該位置對應於該機架之該群組之該些節點之其中之一。

8. 如申請專利範圍第 6 項所述之節點管理方法，更包括：於該列表之該些位址，標記該已註冊機櫃頂端交換器為已註冊。

9. 一種電腦可讀取記錄裝置，用以儲存複數組指令，該些指令由一處理器執行，以執行以下步驟：

於一虛擬機器管理器 (**hypervisor**) 執行一虛擬機架管理模組鏡像 (**virtual rack management module image**)；

由該虛擬機架管理模組鏡像辨識出複數個位置之一列表，該列表之各該位址相對於對應之一機櫃頂端交換器（**top-of-rack switch**）；

由該虛擬機架管理模組鏡像傳送註冊請求（**registration request**）至該列表之各該位址，直到收到來自一未註冊機櫃頂端交換器（**unregistered top-of-rack switch**）之一確認訊息（**acknowledgement**）；

於該未註冊機櫃頂端交換器（**unregistered top-of-rack switch**）註冊該虛擬機架管理模組鏡像，以產生一已註冊機櫃頂端交換器（**registered top-of-rack switch**），其中該已註冊機櫃頂端交換器相對於設置於一機架（**rack**）之一群組的複數個節點（**node**）；以及

藉由該虛擬機架管理模組鏡像管理該群組之該些節點；

其中註冊該虛擬機架管理模組鏡像之步驟包括：

自該未註冊機櫃頂端交換器接收複數個節點位址（**node address**），各該節點位址對應於該群組之該些節點之其中之一。

managing system can have a first rack that includes a first top-of-rack (ToR) switch and a first group of nodes. The system can also have a second rack that includes a second ToR switch and a second group of nodes. The system can include a rack management node that executes a hypervisor. The hypervisor can run a first virtual rack management module (vRMM) and a second vRMM. The first vRMM and the second vRMM can manage the first group of nodes and the second group of nodes, respectively.

【代表圖】

【本案指定代表圖】：第 3 圖。

【本代表圖之符號簡單說明】：

300：資料中心

302₁、302₂、302₃、302_N：機架

304：管理軟體

306：區域網路

308₁：機櫃頂端交換器 1

308₂：機櫃頂端交換器 2

308₃：機櫃頂端交換器 3

308_N：機櫃頂端交換器 N

310₁₋₁：節點 1-1

310₁₋₂：節點 1-2

310_{1-N}：節點 1-N

- 310₂₋₁：節點 2-1
- 310₂₋₂：節點 2-2
- 310_{2-N}：節點 2-N
- 310₃₋₁：節點 3-1
- 310₃₋₂：節點 3-2
- 310_{3-N}：節點 3-N
- 310_{N-1}：節點 N-1
- 310_{N-2}：節點 N-2
- 310_{N-N}：節點 N-N
- 312：機架管理節點
- 314：作業系統
- 316：虛擬機器管理器
- 318₁：vRMM_1
- 318₂：vRMM_2
- 318₃：vRMM_3
- 318_N：vRMM_N

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：無