



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 201415236 A

(43)公開日：中華民國 103 (2014) 年 04 月 16 日

(21)申請案號：102134179

(22)申請日：中華民國 102 (2013) 年 09 月 23 日

(51)Int. Cl. : **G06F13/38 (2006.01)**

G06F13/42 (2006.01)

(30)優先權：2012/09/21 美國

13/624,781

(71)申請人：P L X 科技公司 (美國) PLX TECHNOLOGY, INC. (US)

美國

(72)發明人：雷古拉 傑克 REGULA, JACK (US)；沙伯拉馬尼耶 那加拉杰 SUBRAMANIYAN, NAGARAJAN (IN)；多德森 傑弗瑞 麥克 DODSON, JEFFREY MICHAEL (US)

(74)代理人：陳展俊；林聖富

申請實體審查：無 申請專利範圍項數：14 項 圖式數：3 共 25 頁

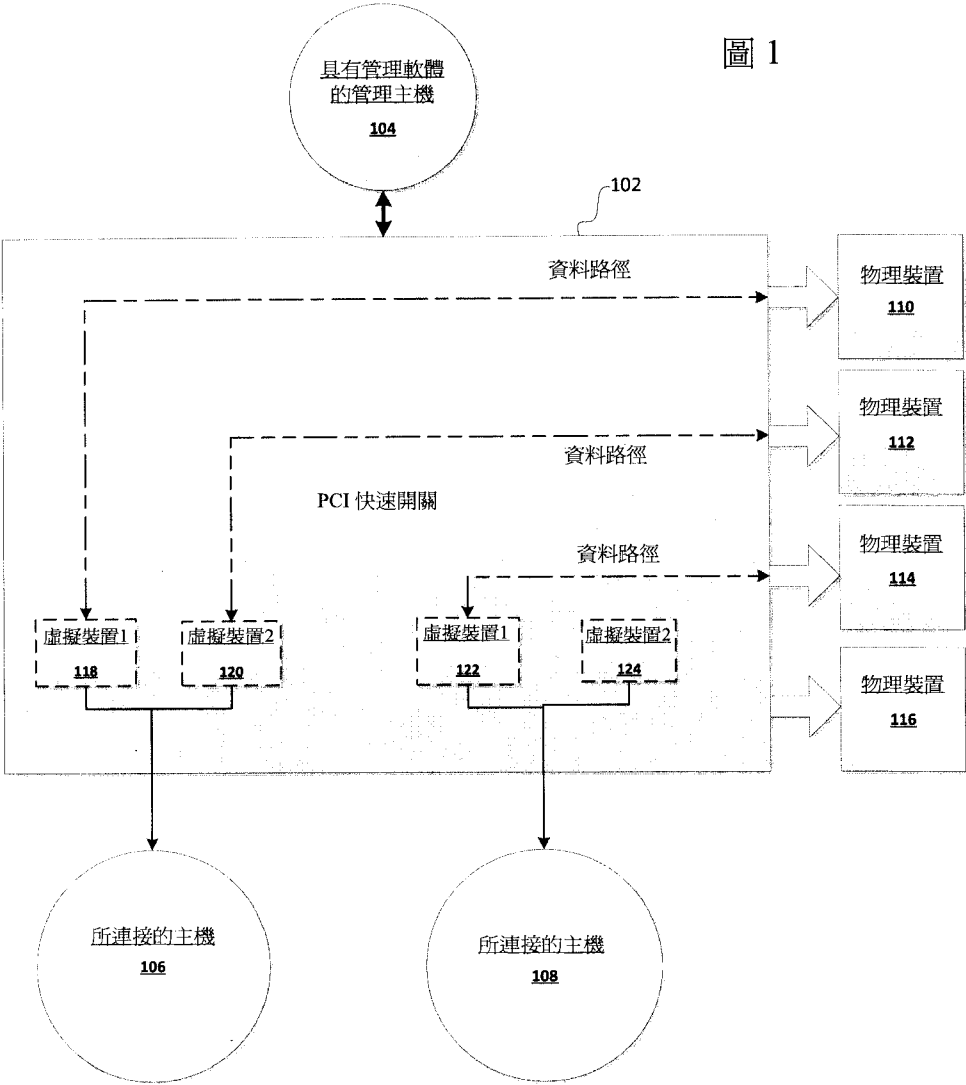
(54)名稱

具有邏輯裝置能力的 P C I 快速開關

PCI EXPRESS SWITCH WITH LOGICAL DEVICE CAPABILITY

(57)摘要

本發明涉及具有邏輯裝置能力的 PCI 快速開關，其實現供所連接的主機系統使用的邏輯裝置。該邏輯裝置由主機管理系統上運行的邏輯裝置啟用軟體創建。該邏輯裝置能夠合併一個或多個物理裝置或者可以是完全基於軟體的。來自所連接的主機的命令被在該主機中的命令和回應佇列中進行處理並且也在該管理系統中所儲存的影子佇列中得到反映。與所連接的主機相關聯的 DMA 引擎被設置為對所連接的(本地)主機中的佇列進行自動觸發。命令被發送至物理裝置以完成工作，並且完成信號被發送至管理軟體，而針對該工作的回應被直接發送至所連接的主機，其並不知曉該邏輯裝置是不存在的並且由管理系統中的軟體所實現。



- 102：開關
- 104：管理主機
- 106：主機
- 108：主機
- 110-116：物理裝置
- 118-124：虛擬裝置



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 201415236 A

(43)公開日：中華民國 103 (2014) 年 04 月 16 日

(21)申請案號：102134179

(22)申請日：中華民國 102 (2013) 年 09 月 23 日

(51)Int. Cl. : **G06F13/38 (2006.01)**

G06F13/42 (2006.01)

(30)優先權：2012/09/21 美國

13/624,781

(71)申請人：P L X 科技公司 (美國) PLX TECHNOLOGY, INC. (US)

美國

(72)發明人：雷古拉 傑克 REGULA, JACK (US)；沙伯拉馬尼耶 那加拉杰 SUBRAMANIYAN, NAGARAJAN (IN)；多德森 傑弗瑞 麥克 DODSON, JEFFREY MICHAEL (US)

(74)代理人：陳展俊；林聖富

申請實體審查：無 申請專利範圍項數：14 項 圖式數：3 共 25 頁

(54)名稱

具有邏輯裝置能力的 P C I 快速開關

PCI EXPRESS SWITCH WITH LOGICAL DEVICE CAPABILITY

(57)摘要

本發明涉及具有邏輯裝置能力的 PCI 快速開關，其實現供所連接的主機系統使用的邏輯裝置。該邏輯裝置由主機管理系統上運行的邏輯裝置啟用軟體創建。該邏輯裝置能夠合併一個或多個物理裝置或者可以是完全基於軟體的。來自所連接的主機的命令被在該主機中的命令和回應佇列中進行處理並且也在該管理系統中所儲存的影子佇列中得到反映。與所連接的主機相關聯的 DMA 引擎被設置為對所連接的(本地)主機中的佇列進行自動觸發。命令被發送至物理裝置以完成工作，並且完成信號被發送至管理軟體，而針對該工作的回應被直接發送至所連接的主機，其並不知曉該邏輯裝置是不存在的並且由管理系統中的軟體所實現。

發明摘要

※ 申請案號：102134179

※ 申請日：

102.9.23

※IPC 分類：

G06F13/38 2006.01

G06F13/42 2006.01

【發明名稱】(中文/英文)

具有邏輯裝置能力的 PCI 快速開關

PCI EXPRESS SWITCH WITH LOGICAL DEVICE CAPABILITY

【中文】

本發明涉及具有邏輯裝置能力的 PCI 快速開關，其實現供所連接的主機系統使用的邏輯裝置。該邏輯裝置由主機管理系統上運行的邏輯裝置啓用軟體創建。該邏輯裝置能夠合併一個或多個物理裝置或者可以是完全基於軟體的。來自所連接的主機的命令被在該主機中的命令和回應佇列中進行處理並且也在該管理系統中所儲存的影子佇列中得到反映。與所連接的主機相關聯的 DMA 引擎被設置為對所連接的（本地）主機中的佇列進行自動觸發。命令被發送至物理裝置以完成工作，並且完成信號被發送至管理軟體，而針對該工作的回應被直接發送至所連接的主機，其並不知曉該邏輯裝置是不存在的並且由管理系統中的軟體所實現。

【英文】

A PCIe switch implements a logical device for use by connected host systems. The logical device is created by logical device enabling software running on a host management system. The logical device is able to consolidate one or more physical devices or may be entirely software-based. Commands from the connected host are processed in the command and response queues in the host and are also reflected in shadow queues stored in the management system. A DMA engine associated with the connected host is set up to automatically trigger on queues in the connected (local) host. Commands are sent to the physical devices to complete the work and a completion signal is sent to the management software and a response to the work is sent directly to the connected host, which is not aware that the logical device is non-existent and is implemented by software in the management system.

【代表圖】

【本案指定代表圖】：第（ 1 ）圖。

【本代表圖之符號簡單說明】：

102 開 關

104 管 理 主 機

106、108 主 機

110-116 物 理 裝 置

118-124 虛 擬 裝 置

【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

(本說明書格式、順序，請勿任意更動)

【發明名稱】(中文/英文)

具有邏輯裝置能力的 PCI 快速開關

PCI EXPRESS SWITCH WITH LOGICAL DEVICE
CAPABILITY

【技術領域】

【0001】 本發明總體上涉及開關 (switch) 和電子通信。更特別地，本發明涉及在智慧 PCIe 開關中啓用合成或邏輯裝置以便合併多個裝置的功能。

【先前技術】

【0002】 電腦架構多年來已經有了大幅發展。近來，對於晶片設計師而言，將諸如通用串列匯流排 (USB) 介面控制器之類的外部資料介面包括在其母板之中變得越來越常見。這些介面被稱之為主機控制器。處理器則通常經由輸入/輸出 (I/O) 互連系統連接至電腦系統的其他元件。

【0003】 存在著許多不同的可用電腦 I/O 互連標準。多年來最為流行的之一是週邊裝置互連 (PCI) 標準。PCI 允許匯流排像橋那樣進行工作，其將本地處理器匯流排與外設隔離開來，允許電腦的中央處理單元 (CPU) 通過該互連而連接至 IO 裝置的主機。

【0004】 近來，PCI 的後繼已經普及，其被稱作 PCI Express 5

（或者簡單地被稱作 PCIe）的。PCIe 提供了更高的性能、增加的靈活性以及針對下一代系統的可擴縮性，同時保留了與現有 PCI 應用的軟體相容性。與傳統的 PCI 相比，PCI Express 協議明顯更為複雜，其具有三層—事務層、資料連結層和物理層。

【0005】 在 PCI Express(快速 PCI)系統中，根聯合體(root complex)裝置將處理器和記憶體子系統連接至包含一個或多個開關裝置（然而，實施例也可能沒有開關）的 PCI 快速開關構造。在 PCI Express 中，使用點對點架構。與 PCI 系統中的主機橋相類似，根聯合體代表通過本地 I/O 互連而進行互連的處理器生成事務請求。根聯合體功能可以被實施為離散的裝置，或者可以與處理器進行整合。根聯合體可以包含多於一個的 PCI Express 埠，並且多個開關裝置可以連接至根聯合體上的埠或者被級聯。

【0006】 作為示例，圖 1 是描繪正常共用的 IO 架構的方塊圖，該 IO 架構具有由運行開關管理軟體的管理主機 104 所控制的標準 PCIe 開關 102。開關 102 對一個或多個主機進行服務，所述主機被示為所連接的主機 106 和所連接的主機 108（也被稱作“本地主機”），例如同伺服器、PC 和其他計算裝置。還連接至開關的是一個或多個裝置 110-116，所述裝置通常為所連接的主機提供一些類型的功能或服務。虛擬裝置 118-124 處於開關 102 之內。虛擬裝置 118 和 120 連接至所連接的主機 106 且虛擬裝置 122 和 124 連接至所連接的主機 108。這些虛擬裝置中的某些具有去往物理裝置 110-114 的資

料路徑。虛擬裝置 118-124 的功能和角色在題為

“MULTI-ROOT SHARING OF SINGLE-ROOT INPUT/OUTPUT VIRTUALIZATION” 的美國專利申請號 12/979,904 的共同未決申請中有所描述，其中描述了一種在多個主機使用支援共用 I/O 機制的 PCIe 開關的非透明埠進行連接時使用資源重定向方法的解決方案。

【0007】 將期望的是：通過使用 DMA 引擎、位址映射和記憶體空間重定向而對連接至開關的物理裝置的功能進行合併來進一步增強 PCIe 開關的功能。此外，將期望的是：針對具有或沒有諸如 RAID 的附加特徵的一組 SSD 驅動器（或其他儲存裝置）啓用（enable）用於所連接的主機的合成 NVMe 端點，或者與若干所連接的主機共用一組 SSD/儲存驅動器。最後，即使在資源原本並不允許共用的情況下，如果通過軟體和硬體幫助而使得能夠在多個所連接的主機之間共用稀缺或昂貴的資源，則對於開關的用戶而言也將是有利的。

【發明內容】

【0008】 本發明的一個方面是一種通過啓用開關中的非存在的、邏輯（或合成）裝置的操作而增強 PCI 快速開關的功能的方法。該邏輯裝置由管理系統主機中對開關的操作進行控制的邏輯裝置軟體來加以實現。該合成裝置被呈現給連接至該開關的本地主機。該本地主機內的寫入操作被擷取，從而使得管理系統能夠創建本地主機元件佇列的影子副本（shadow copy）。本地主機載入用於該合成裝置的驅動器。在

該本地主機中發生的寫入被反映在管理系統中。在該管理系統上創建影子佇列，該管理系統反映本地主機中的命令和回應佇列。與本地主機埠相關聯的 DMA 引擎被設置為對本地主機中的佇列進行自動觸發。

【圖式簡單說明】

【0009】

圖 1 是描繪正常共用的 IO 架構的方塊圖，該 IO 架構具有由運行開關管理軟體的系統管理主機所控制的標準 PCIe 開關；

圖 2 是描繪依據一個實施例的具有邏輯裝置和管理主機系統的 PCIe 開關的方塊圖，該管理主機系統具有邏輯裝置啓用軟體；和

圖 3 是依據一個實施例的在 PCIe 開關上安裝合成或邏輯裝置的處理的流程圖。

【實施方式】

【0010】 現在將對本發明的具體實施例進行詳細參考，其包括發明人為了執行本發明所預期的最佳模式。這些具體實施例的示例在附圖中進行圖示。雖然結合這些具體實施例對本發明進行描述，但是將要理解的是，並不旨在將本發明侷限於所描述的實施例。與之相反，旨在覆蓋如可以包括在由所附申請專利範圍限定的本發明的精神和範圍內的替換物、修改和同等物。在以下描述中，闡述具體細節以便提供對本

發明的全面理解。本發明可以在沒有這些具體細節中的一些或全部的情況下實施。此外，可能沒有對公知特徵進行詳細描述以避免對本發明造成不必要的混淆。

【0011】 依據本發明，可以使用各種類型的作業系統、編程語言、計算平臺、電腦程式和/或通用機器來實現所述元件、處理步驟和/或資料結構。此外，本領域普通技術人員將會認識到，在不背離本文所公開的發明原理的範圍和精神的情況下，也可以使用諸如硬線裝置、現場可編程閘陣列（FPGA）、專用積體電路（ASIC）等的具有較不通用屬性的裝置。本發明還可以被有形地體現為儲存在諸如儲存裝置的電腦可讀媒體上的電腦指令的集合。

【0012】 在各個附圖中描述了用於將物理裝置合併為 PCIe 開關中的邏輯裝置的方法和系統，其中該邏輯裝置使用切換管理系統中的軟體來加以實現。通常，使用命令佇列和採用門鈴和中斷的完成或回應佇列 FIFO 模型以用於那些 FIFO 的任何裝置驅動器模型能夠被使用以下所描述的方法和系統而呈現給所連接的主機。在繼續進行之前，對基本元件進行描述是有用的。如所提到的，一個是用於使用所連接的主機記憶體而從所連接的主機驅動器向 DMA 引擎進行輸入和輸出的 FIFO 模型。另一個是使用用於所連接的主機驅動器的門鈴以向硬體指示新的工作已經被添加或者正在等待。為了負荷平衡並且指示工作完成，針對每個完成佇列的每個完成或回應佇列使用中斷。這些特徵存在於一些 PCIe 開關之中，諸如來自加利福尼亞 Sunnyvale 的 PLX Technologies 的

Capella 2 開關。附圖在以下描述了本發明的創新性 PCIe 開關所需要的附加特徵。首先提供這些附加特徵的概述。

【0013】 在一個實施例中，存在著當所連接的主機通過添加至其命令佇列而生成工作時在管理系統中對佇列/FIFO 內容的自動更新。DMA 引擎（開關埠的一部分）將命令佇列中的排隊的工作請求自動傳輸至管理系統，該管理系統也被稱作消費者/合併系統。另一個特徵是從物理裝置到所連接的主機的直接資料路徑。在命令（控制路徑）通過管理系統（MCPU）時，該資料路徑被允許直接從物理裝置去往所連接的主機。這提供了顯著的性能優勢並且通過防止管理系統處的瓶頸而使得延遲有所改進。例如，典型的裝置（包括 SSD 儲存器）將從該裝置向所連接的主機記憶體進行 DMA 請求（讀或寫入），並且這些將直接從物理裝置去往所連接的主機。所連接的主機的請求最初去往管理系統，並且該管理系統隨後向個體裝置（或驅動器）重新定向或發出等同的命令。另一個特徵是位址隔離以及從任意物理裝置端點存取不同的所連接的主機的能力。這可以通過由管理系統軟體啓用的硬體的位址映射/俘獲特徵而被啓用。爲了向所連接的主機呈現邏輯裝置，要求開關俘獲（trap）所連接的主機 I/O 枚舉週期（enumeration cycle）並且還使用管理系統中的邏輯裝置軟體對它們進行正確應答。

【0014】 在本發明的一個實施例中，連接至 PCIe 開關的邏輯或非存在裝置被呈現或顯示給所連接的主機（連接到 PCIe 開關的主機）。該邏輯裝置在物理上並不存在（其是非存

在的)，從而使得虛構裝置能夠在任意時間連接至所連接的主機。邏輯裝置是合成的 I/O 合併 PCIe 端點。該邏輯裝置提供了一種用於物理裝置的合併的機制。該邏輯裝置也可以是單純的軟體創建的裝置而其並不對應於任何物理裝置（即，在其之後沒有物理裝置）。該邏輯裝置表現為 PCIe 端點或者表現為具體實施方式或部署，諸如 RAID 裝置。其使用 PCIe 開關的具體軟體和硬體加速特徵。

【0015】 圖 2 是描繪依據一個實施例的具有邏輯裝置和管理主機系統的 PCIe 開關的方塊圖，該管理主機系統具有邏輯裝置啟用軟體。示出的是一種連接至主機 204 和三個物理裝置 206-210 的創新性 PCIe 開關 202。具有邏輯裝置軟體以及用於操作管理系統的其他必要軟體的管理系統主機 212 連接至開關 202。邏輯裝置啟用軟體實施了由虛線框所示的邏輯裝置 214（也被稱之為合併的虛擬裝置）。邏輯裝置 214 虛擬地與所連接的主機 204 一起操作。其具有去往物理裝置 206-210 的資料路徑。管理系統主機 212 具有由虛線所示的到實現邏輯裝置 214 的物理裝置 206-210 的控制路徑。這些控制路徑以虛線表示以指示它們實質上源自於管理系統 212 中的邏輯裝置啟用軟體並且負責在物理上實現邏輯（虛擬）裝置 214。

【0016】 在以下所描述的各個實施例中，一種對 PCIe 開關實現邏輯裝置的方法以新的能力、位址映射和變換及配置，以及儲存空間和配置空間重定向特徵來利用 DMA 引擎。例如，可能需要位址映射以使得邏輯裝置中的基礎位址暫存器（BAR）被主機管理系統中的軟體映射至管理記憶體。以

這種方式，各個物理裝置的特徵被合併在一個邏輯裝置之下。另一種優勢在於，如果一個物理裝置具有錯誤或者變得禁用，則邏輯裝置仍然能夠進行操作（例如，主機系統將不會遇到“藍屏”的情形）。中斷合併和生成也是可能的，從而緩解了主機系統上的負荷。在沒有這種合併的情況下，如果存在“n”個物理裝置，則可能有“n”個那麼多的裝置發送中斷，這在一些情況下對於所連接的主機而言可能是超載的。來自一個邏輯裝置的中斷將會緩解主機系統的工作負荷。在另一種部署情形中，邏輯裝置可以是 NVMe 端點，其可以對物理 NVMe 或非 NVMe 儲存裝置（例如，SATA 驅動、SSC 驅動）進行合併。

【0017】 相關申請中所描述的方法描述了物理裝置如何被呈現給主機。如以上所指出的，本發明通過將邏輯裝置的位址空間映射至主機管理系統 CPU 中而解決了向主機呈現邏輯、非存在裝置。邏輯裝置因此還在裝置被呈現給主機時綁定至 DMA 引擎，該主機僅能夠看到邏輯裝置而不是 DMA 功能。通常，如果僅有邏輯裝置被連接至 PCIe 開關，則所連接的主機將不會看到或瞭解到潛在的 DMA 或非透明（NT），儘管該 DMA 和 NT 功能正在進行操作而只是沒有向主機進行顯示而已。

【0018】 如以下所解釋的，DMA 引擎被附著至佇列和門鈴（doorbell）。佇列被分配在主機記憶體中並且被寫入到邏輯裝置。佇列被映射至主機管理系統記憶體（FIFO）中。DMA 隨後被編程以將邏輯裝置中的佇列綁定至主機系統記憶體佇

列。此外，所連接的主機中的門鈴也可以被映射至主機管理系統門鈴堆疊。如本領域中已知的，裝置驅動器向裝置中的命令佇列發送命令。可以存在一個或多個命令佇列。硬體對命令佇列中的資料執行操作並且結果被儲存在回應佇列中。如所提到的，邏輯裝置具有用於實現這些佇列的 FIFO，並且 DMA 引擎被附著到它們或者被編程以綁定至邏輯裝置中的佇列。

【0019】 當所連接的主機驅動器啟動或“敲響”邏輯裝置中的門鈴時（即，當所連接的主機想要使用邏輯裝置所提供的功能時），邏輯裝置 DMA 引擎由於其被綁定至該邏輯裝置並且更具體地由於其瞭解邏輯裝置命令佇列而喚醒。DMA 引擎將此傳輸至管理系統中斷。所連接的主機所請求的工作被插入佇列中並且處理遵循其正常操作過程。

【0020】 主機上的邏輯軟體能夠在主機上沒有外來或附加軟體的情況下執行工作。這能夠僅借助於硬體來完成。

【0021】 所連接的主機驅動器向裝置配置空間或 BAR 寫入操作。特別地，其將寫入到主機管理系統記憶體中的影子空間中所反映的 BAR 中的裝置的暫存器。照此，BAR 的影子副本在主機管理系統上。因此，當在邏輯裝置中存在著對 BAR 的寫入時，其在管理系統中有所反映。

【0022】 主機管理系統中的影子記憶體擷取原本在所連接的主機中的寫入。如本領域中已知的，當所連接的主機對驅動器進行載入時，其開始使用 BAR 中的暫存器，所述 BAR 專用於所呈現的裝置。驅動器開始進行操作、配置命令並且

初啓（bring up）裝置。PCIe 開關將所有寫入和命令重新定向至管理系統中。如所提到的，所連接的主機中的所有寫入將經由 PCIe 開關而出現在主機管理系統影子記憶體中。

【0023】 一個步驟是在本地主機記憶體上設置資料命令和回應佇列。用於設置或配置空間（諸如，需要支援多少佇列，邏輯盤的數量，等等）的命令以及其他配置問題可以由主機管理系統上的邏輯軟體進行應答或照料（tend）。驅動器通過為 FIFO 分配本地系統記憶體並且向硬體寫入基礎位址和大小而對佇列進行配置。

【0024】 圖 3 是依據一個實施例的在 PCIe 開關上安裝合成或邏輯裝置的過程的流程圖。在步驟 302，該合成裝置被呈現在連接至該開關的本地主機上或顯示給連接至該開關的本地主機。該本地主機（本文為“主機”）通過執行裝置所擷取的匯流排上的配置週期而對 PCIe 上的裝置進行枚舉。這在母專利申請中有所描述。在步驟 304，主機系統軟體在配置空間上寫入 BAR 並且啓用邏輯裝置。在步驟 306，該開關擷取步驟 304 處的寫入，從而使得管理系統能夠構建與本地主機上相同的配置。其能夠分配相同的儲存空間，因此在 BAR 和儲存空間記憶體有著一對一映射。因此，現在在管理系統中存在著本地主機 BAR 的影子副本。

【0025】 在步驟 308，本地主機對用於邏輯裝置的驅動器進行載入。管理系統經由開關而被通知以相同的操作。也就是說，當本地主機對主機記憶體中的 BAR 寫入某些東西時，這些寫入將在管理系統中的影子記憶體中得到反映。在步驟

309，本地主機驅動器執行驅動器初始化和裝置配置查詢。經由開關給管理系統通知這些。在步驟 310，本地主機中的所有寫入出現在管理系統上的影子記憶體中。該驅動器執行設置命令。

【0026】 在步驟 312，驅動器就像其正常情況下那樣設置用於本地主機中的命令和回應的佇列。這些用於儲存命令和回應的 FIFO 佇列在本領域是已知的。在步驟 314，由管理系統上的軟體創建影子佇列並且將影子佇列儲存在記憶體中。本地主機中的佇列實質上被映射至管理系統記憶體。在步驟 316，用於本地主機的埠的 DMA 引擎被設置為對本地主機上的佇列進行自動觸發。也就是說，該 DMA 引擎被編程為將本地主機中的命令佇列綁定至管理系統記憶體中的命令佇列。該映射還在步驟 218 處利用本地主機中的門鈴來完成。因此，當主機驅動器敲響門鈴時，該 DMA 喚醒（即，被自動觸發）。

【0027】 在步驟 320，本地主機驅動器設置本地主機上的中斷向量和處理機（handler）。這些中斷被發送至管理系統，在此時可能發生兩種不同動作之一。在步驟 322，管理系統儲存中斷向量以便在運行時間期間直接存取。在步驟 324，中斷向量與將管理本地主機回應佇列的 DMA 引擎佇列相關聯。在這一階段，邏輯裝置在 PCIe 開關上的安裝完成。

【0028】 如以上所描述的，由管理系統中的軟體實現邏輯裝置。正是該軟體對資料進行工作並且將其指向物理裝置。當物理裝置完成功能或工作時，完成信號被發送至管理系統，該管理系統經由物理裝置的特徵被合併在其中的邏輯裝

置軟體而對工作進行排序。來自工作的回應被直接發送至本地主機，其在那裏被儲存在回應佇列中。針對本地主機回應佇列的這一更新仍然在管理系統的影子記憶體中有所反映。

【0029】 本文的描述提供了多個新的特徵以及針對現有 PCIe 切換技術的改進。在一個實施例中，對用於啓用 I/O 物理裝置合併的硬體輔助進行描述，同時使用該合併增加或啓用增值特徵。描述了通過使用所連接的主機中的邏輯或合成端點而對從所連接的物理裝置到所連接的埠的熱插拔事件進行的分離和隔離。還存在著由於所連接的物理裝置生成的到所連接的主機的中斷的合併和隔離而導致的性能提升。最後，應當注意的是，以上所描述的實施例可利用管理系統中的軟體進行擴展。可以存在多個主機和物理裝置（例如，可以存在 15 個物理裝置，其被通過全部連接至一個 PCIe 開關的 5 個所連接的主機進行投影）。

【0030】 所描述的實施例的各個方面、實施例、實施方式或者特徵能夠被分開使用或者以任意組合來使用。所描述的實施例的各個方面可以由軟體、硬體或者硬體和軟體的組合來加以實現。所描述的實施例也可以被體現為電腦可讀媒體上的電腦可讀代碼。電腦可讀媒體被定義為能夠儲存隨後能夠由電腦系統進行讀取的資料的任意資料儲存裝置。電腦可讀媒體的示例包括唯讀記憶體、隨機存取記憶體、CD-ROM、DVD、磁帶和光學資料儲存裝置。電腦可讀媒體也可以在網路耦合的電腦系統上進行分佈以使得電腦可讀代碼以分散式方式進行儲存和運行。

【0031】 雖然已經參考其具體實施例特別示出並描述了本發明，但是本領域技術人員將會理解的是，在不背離本發明的精神或範圍的情況下可以對所公開的實施例的形式和細節方面進行改變。此外，雖然已經參考各個實施例在本文中對本發明的各種優勢、方面和目標進行了討論，但是將要理解的是，本發明的範圍並不應當通過參考而限定於這樣的優勢、方面和目標。相反，本發明的範圍應當參考所附申請專利範圍來加以確定。

【符號說明】

【0032】

- 102、202 開關
- 104、204 管理主機
- 106、108 主機
- 110-116、206-210 物理裝置
- 118-124 虛擬裝置
- 212 管理系統主機
- 214 邏輯裝置

申請專利範圍

1. 一種在 PCIe 開關中啓用合成裝置的操作的方法，該方法包括：

將該合成裝置呈現給連接至該開關的本地主機元件；

擷取該本地主機內的讀取和寫入操作，從而使得該開關的管理系統能夠創建本地主機元件佇列的影子副本；

向該管理系統通知本地主機驅動器操作，其中該本地主機載入驅動器；

使得在該本地主機中發生的寫入能夠被反映在該管理系統上的記憶體中；

使得能夠在該管理系統上創建影子佇列，所述影子佇列反映了該本地主機中的命令和回應佇列；

設置與本地主機埠相關聯的 DMA 引擎以對該本地主機中的佇列進行自動觸發；並且

使得合成裝置門鈴能夠映射至 DMA 引擎門鈴。

2. 如申請專利範圍第 1 項的方法，進一步包括：

將該本地主機上的中斷向量發送到該管理系統。

3. 如申請專利範圍第 1 項的方法，進一步包括：

針對適當主機裝置回應將中斷向量與 DMA 引擎佇列相關聯。

4. 如申請專利範圍第 1 項的方法，其中多個本地主機和

多個物理裝置使用該合成裝置。

5. 如申請專利範圍第 1 項的方法，其中該管理系統具有合成裝置啓用模組，其提供用於在該開關中實現該合成裝置的可執行指令。

6. 如申請專利範圍第 1 項的方法，其中該管理系統具有在該本地主機上所儲存的佇列的影子副本。

7. 如申請專利範圍第 1 項的方法，其中當該物理裝置完成該本地主機所請求的功能時，將完成信號發送至該管理系統並且將回應直接發送至該本地主機。

8. 如申請專利範圍第 5 項的方法，其中該啓用模組對來自該本地主機的資料執行操作並且將該資料定向至該物理裝置。

9. 如申請專利範圍第 1 項的方法，其中在該管理系統和物理裝置之間存在用於傳送與該合成裝置相關的命令的控制路徑以及在該物理裝置和該本地主機之間存在資料路徑。

10. 如申請專利範圍第 1 項的方法，其中存在針對本地主機中的反映在該管理系統中的影子佇列中的佇列的更新的一對一映射。

11. 如申請專利範圍第 1 項的方法，其中該合成裝置是沒有相關聯物理裝置的、由軟體所創建的裝置。

12. 如申請專利範圍第 1 項的方法，進一步包括：在該本地主機和該管理系統之間映射位址，從而啓用在該物理裝置和該本地主機之間的資料傳輸。

13. 如申請專利範圍第 1 項的方法，進一步包括在該本地主機中分配佇列並且將該佇列寫入至該合成裝置。

14. 一種與管理系統相結合的 PCI 快速開關，該管理系統具有用於在該開關中實現物理裝置的邏輯合併的邏輯裝置啓用器模組，其中該邏輯合併被所連接的主機使用，並且其中該管理系統中的記憶體儲存如在該邏輯裝置中所儲存的命令佇列、響應佇列和門鈴的影子副本。

圖 1

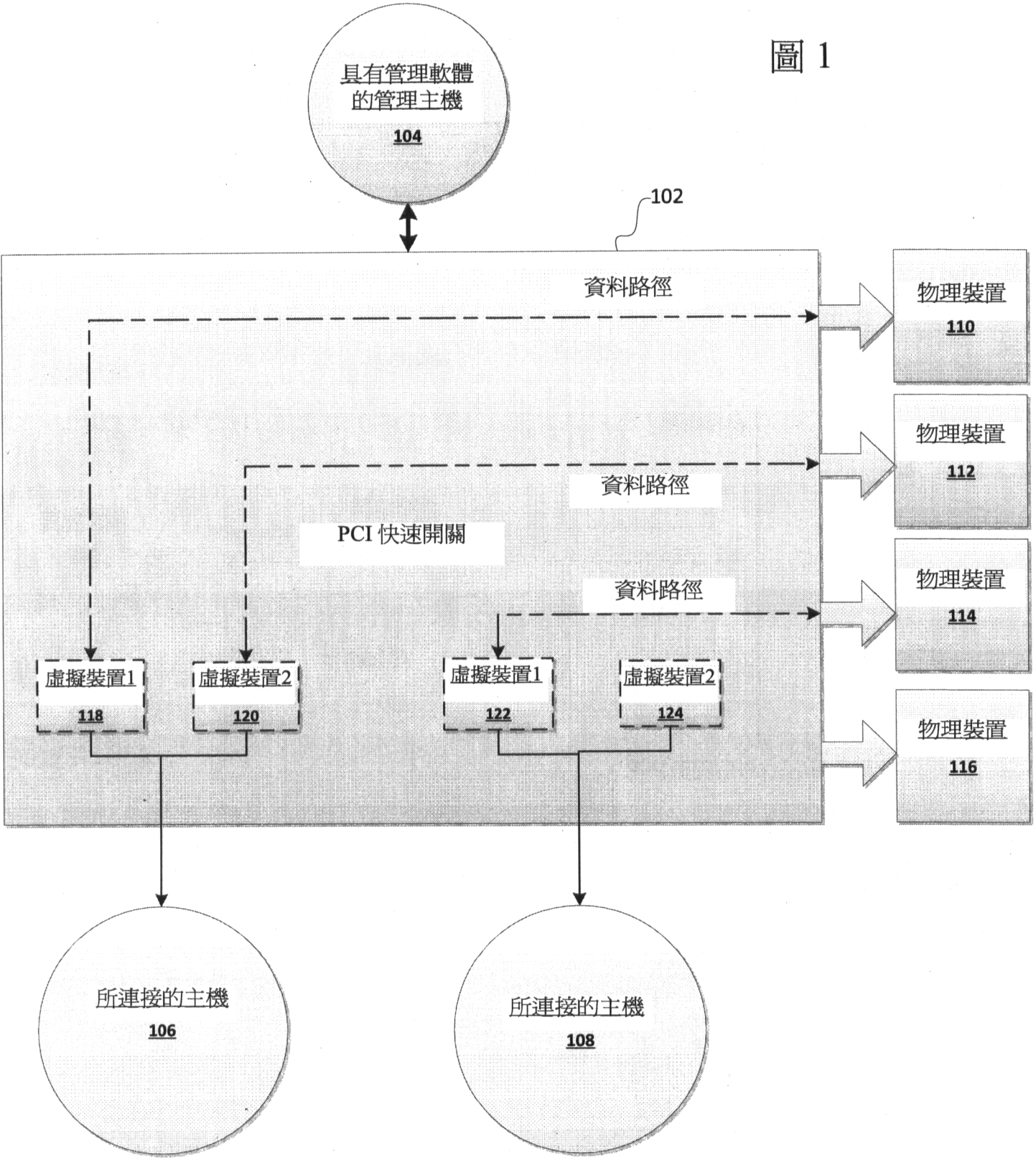
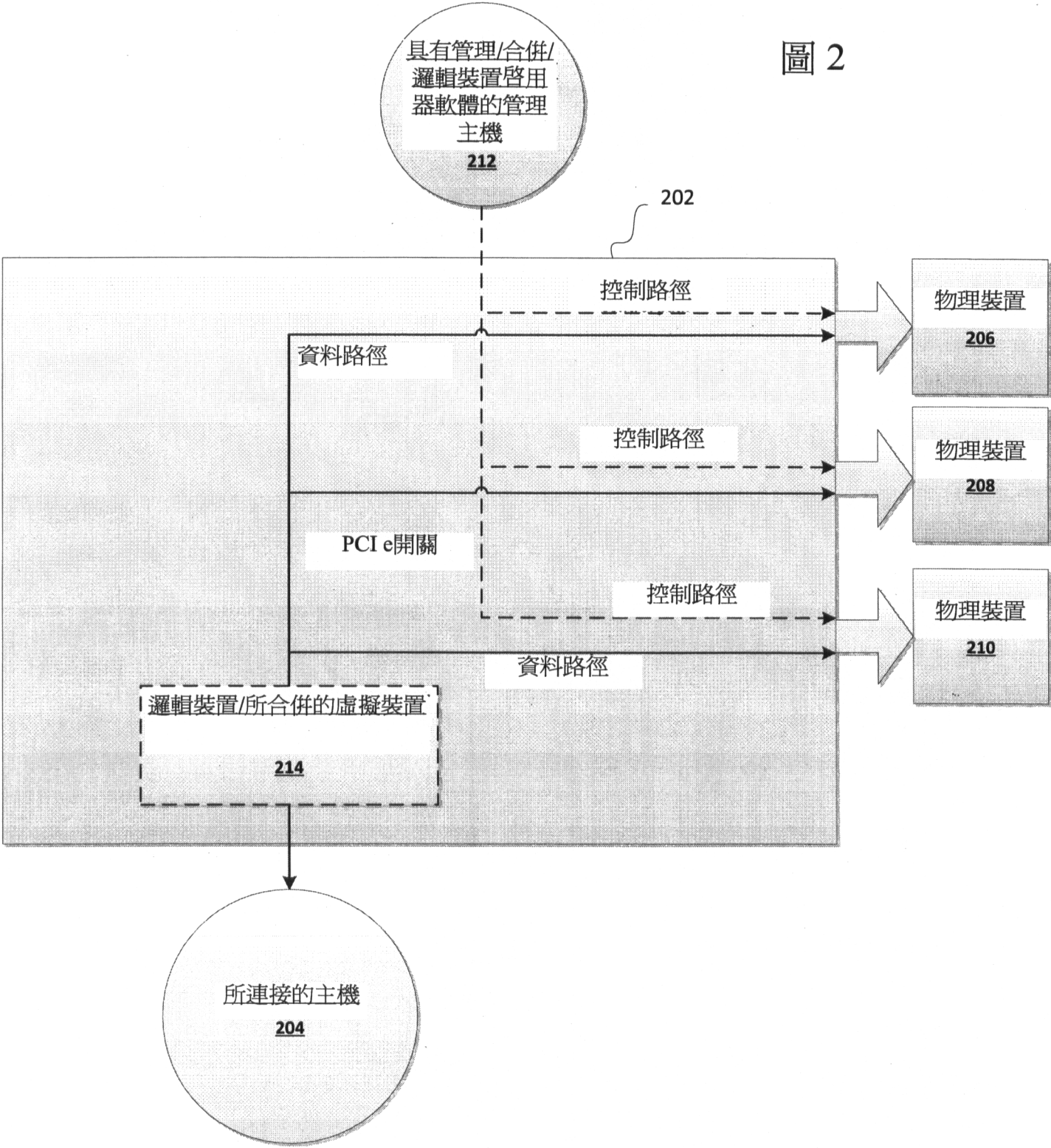


圖 2



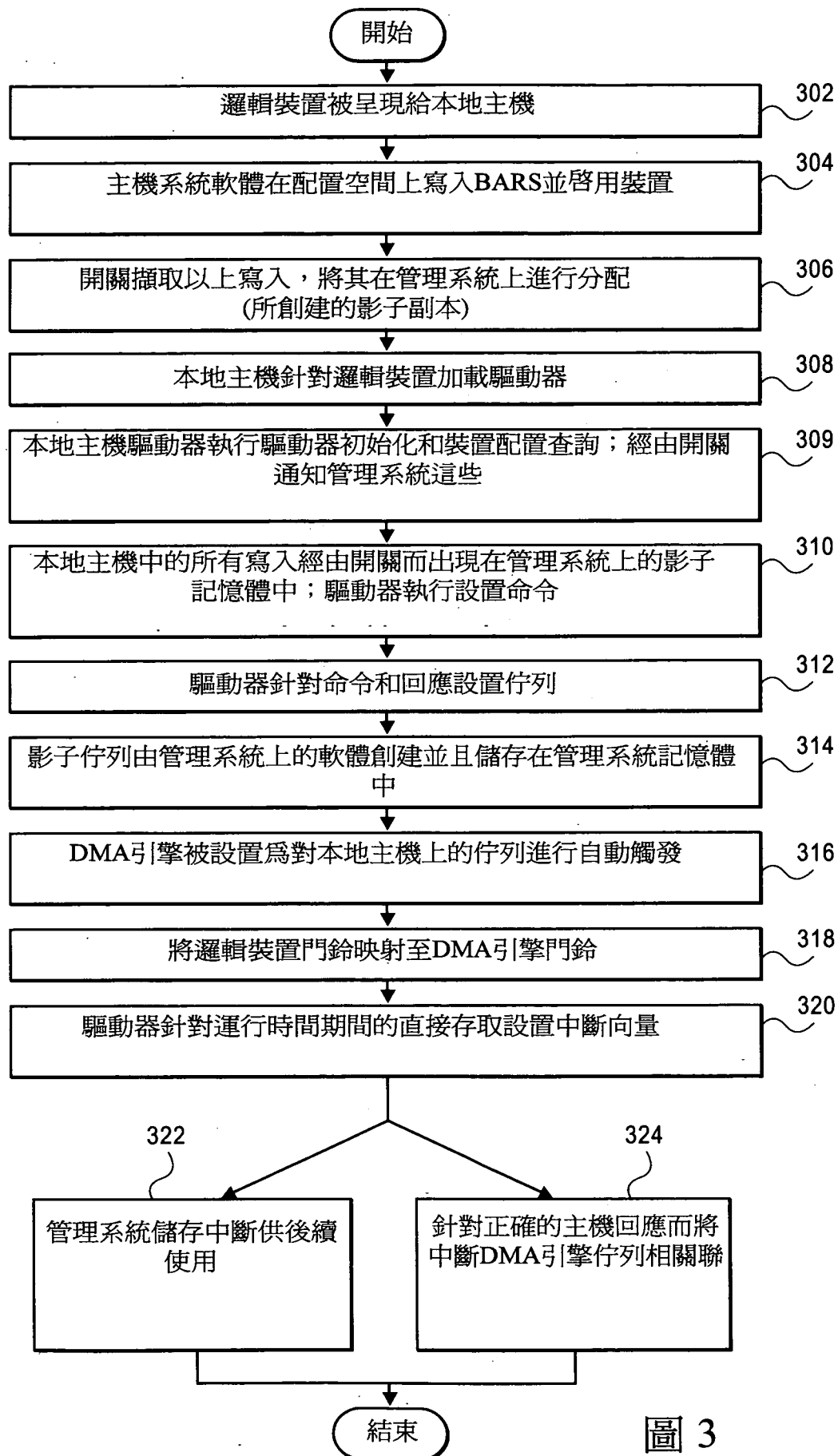


圖 3