



(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 I690789 B

(45)公告日：中華民國 109 (2020) 年 04 月 11 日

(21)申請案號：107142431

(22)申請日：中華民國 107 (2018) 年 11 月 28 日

(51)Int. Cl. : G06F1/04 (2006.01)

G06F13/38 (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：許繼委 XU, JI-WEI (CN)；呂宏根 LV, HONG-GEN (CN)

(74)代理人：許世正

(56)參考文獻：

CN 104579772A

CN 105094242A

CN 106649014A

CN 107239346A

CN 206147604U

US 2009/0167771A1

US 2014/0237156A1

審查人員：游象甫

申請專利範圍項數：11 項 圖式數：8 共 28 頁

(54)名稱

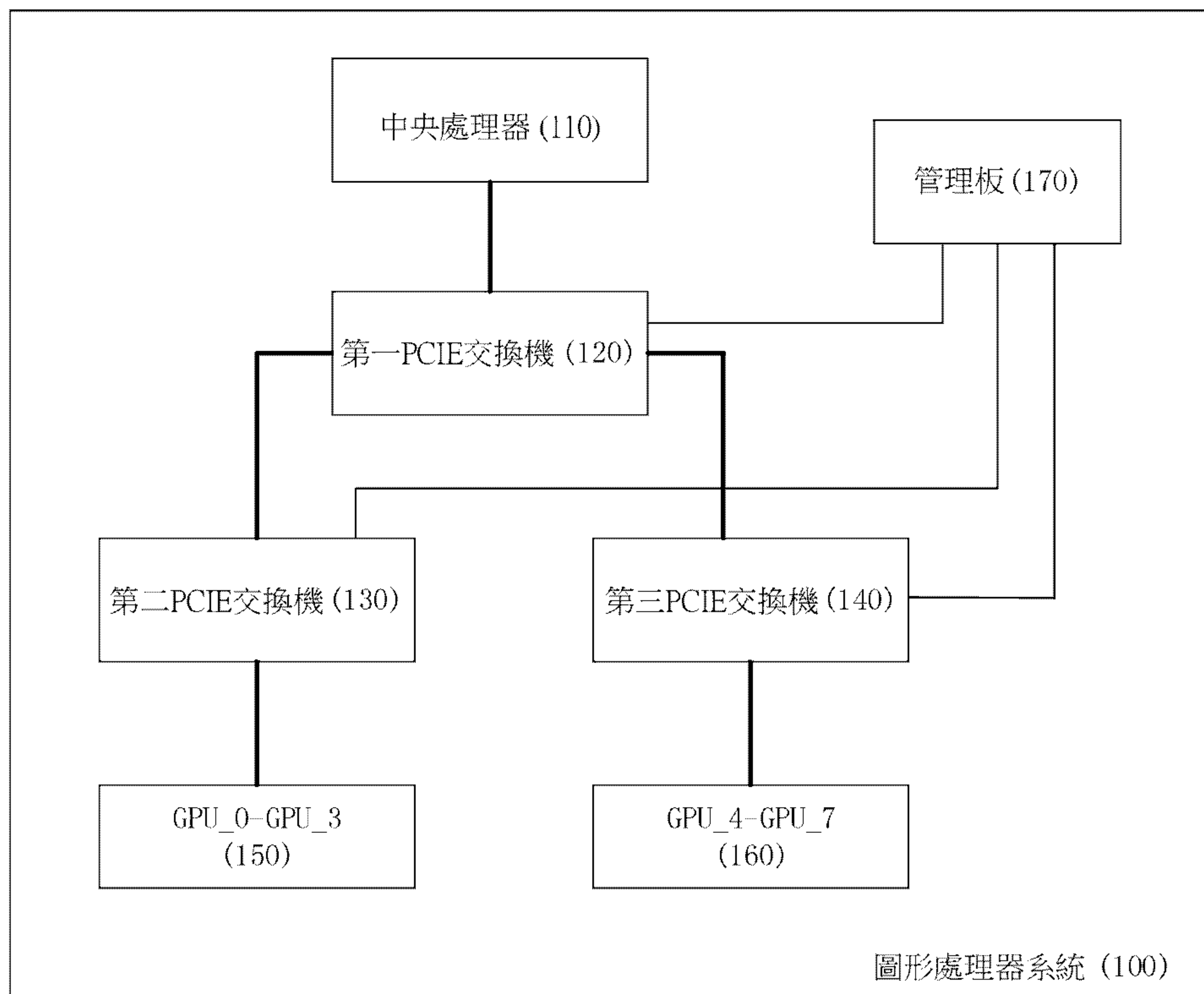
圖形處理器系統

(57)摘要

本發明提供一種圖形處理器系統，包括：一中央處理器，一第一 PCIE 交換機與中央處理器相連，一第二 PCIE 交換機與第一 PCIE 交換機相連，下行連接複數個圖形處理器，一第三 PCIE 交換機與第一 PCIE 交換機相連，下行連接複數個圖形處理器，一管理板與第一 PCIE 交換機、第二 PCIE 交換機和第三 PCIE 交換機分別相連，進行管理設置。本發明實現了多主機和多終端之間的靈活通信，解決了 GPU 相互之間的 P2P 通信頻寬偏低的問題。

This disclosure discloses a graphic processor system comprising: a central processor, a first PCIE switch connected to the central processor, a second PCIE switch connected to the first PCIE switch, and the second PCIE switch downlink connected to the plurality of graphic processors; also, a third PCIE switch connected to the first PCIE switch, and the third PCIE switch downlink connected to the plurality of graphic processors; additionally, a management board connected to the first PCIE switch, the second PCIE switch, and the third PCIE switch respectively for managing the configuration; this disclosure discloses the flexible communication between a plurality of hosts and a plurality of terminals, and the problem of the low bandwidth of the P2P communication between the GPUs is improved.

指定代表圖：



符號簡單說明：

100:圖形處理器系統

110:中央處理器

120:第一 PCIE 交換機

130:第二 PCIE 交換機

140:第三 PCIE 交換機

150, 160:圖形處理器

170:管理板

圖 1



I690789

【發明摘要】

【中文發明名稱】 圖形處理器系統

【英文發明名稱】 GRAPHIC PROCESSOR SYSTEM

【中文】

本發明提供一種圖形處理器系統，包括：一中央處理器，一第一 PCIE 交換機與中央處理器相連，一第二 PCIE 交換機與第一 PCIE 交換機相連，下行連接複數個圖形處理器，一第三 PCIE 交換機與第一 PCIE 交換機相連，下行連接複數個圖形處理器，一管理板與第一 PCIE 交換機、第二 PCIE 交換機和第三 PCIE 交換機分別相連，進行管理設置。本發明實現了多主機和多終端之間的靈活通信，解決了 GPU 相互之間的 P2P 通信頻寬偏低的問題。

【英文】

This disclosure discloses a graphic processor system comprising: a central processor, a first PCIE switch connected to the central processor, a second PCIE switch connected to the first PCIE switch, and the second PCIE switch downlink connected to the plurality of graphic processors; also, a third PCIE switch connected to the first PCIE switch, and the third PCIE switch downlink connected to the plurality of graphic processors; additionally, a management board connected to the first PCIE switch, the second PCIE switch, and the third PCIE switch respectively for managing the configuration; this disclosure discloses the flexible communication between a plurality of hosts and a plurality of terminals, and the problem

of the low bandwidth of the P2P communication between the GPUs is improved.

【指定代表圖】：圖 1。

【代表圖之符號簡單說明】

- 100 圖形處理器系統
- 110 中央處理器
- 120 第一 PCIE 交換機
- 130 第二 PCIE 交換機
- 140 第三 PCIE 交換機
- 150，160 圖形處理器
- 170 管理板

【特徵化學式】

無

【發明說明書】

【中文發明名稱】 圖形處理器系統

【英文發明名稱】 GRAPHIC PROCESSOR SYSTEM

【技術領域】

【0001】 本發明屬於資料處理技術領域，涉及一種資料處理系統，特別是涉及一種圖形處理器系統。

【先前技術】

【0002】 隨著虛擬實境和人工智慧技術的快速發展，對具有海量資料計算和處理能力的系統的需求逐漸攀升。圖形處理器（Graphics Processing Unit，GPU）正是以高計算性能著稱，近年來顯得格外受歡迎。

【0003】 GPU又稱顯示核心、視覺處理器、顯示晶片，其是一種專門在個人電腦、工作站、遊戲機和一些移動設備（如平板電腦、智慧手機等）上執行圖像運算工作的微處理器。GPU的用途是將電腦系統所需要的顯示資訊進行轉換驅動，並向顯示器提供行掃描信號，控制顯示器的正確顯示，是連接顯示器和個人電腦主機板的重要元件，也是“人機對話”的重要設備之一。

【0004】 在實際應用中，採用多GPU系統來增強圖形處理能力是常用手段，但由於GPU與CPU（中央處理器）之間通信介面的限制，以及

GPU與GPU之間通信頻寬的限制，多GPU系統的圖形處理能力也是受限的，並不能隨著GPU數量的增多而呈現指數增強的效果。

【發明內容】

【0005】 本鑒於以上所述現有技術的缺點，本發明的目的在於提供一種圖形處理器系統，用於解決現有多GPU系統的通信速率低的問題。

【0006】 為實現上述目的及其他相關目的，本發明提供一種圖形處理器系統，所述圖形處理器系統包括：一中央處理器；一第一快捷外設互聯標準交換機，與所述中央處理器相連；一第二快捷外設互聯標準交換機，與所述第一快捷外設互聯標準交換機相連；所述第二快捷外設互聯標準交換機下行連接複數個圖形處理器；一第三快捷外設互聯標準交換機，與所述第一快捷外設互聯標準交換機相連；所述第三快捷外設互聯標準交換機下行連接複數個圖形處理器；一管理板，與所述第一快捷外設互聯標準交換機、第二快捷外設互聯標準交換機和第三快捷外設互聯標準交換機分別相連，對所述第一快捷外設互聯標準交換機、第二快捷外設互聯標準交換機和第三快捷外設互聯標準交換機進行管理設置。

【0007】 於本發明的一實施例中，所述第一快捷外設互聯標準交換機包括：至少一第一主機埠，用於與所述中央處理器通信相連；至少二個第一光纖通信埠，用於分別與所述第二快捷外設互聯標準交換機和所述第

三快捷外設互聯標準交換機通信相連；至少一第一上行埠，用於與所述管理板通信相連。

【0008】 於本發明的一實施例中，所述第二快捷外設互聯標準交換機包括：至少一第二光纖通信埠，用於與所述第一快捷外設互聯標準交換機的一第一光纖通信埠通信相連；至少一第二上行埠，用於與所述管理板通信相連；至少一第二下行埠，用於與圖形處理器通信相連。

【0009】 於本發明的一實施例中，所述第三快捷外設互聯標準交換機包括：至少一第三光纖通信埠，用於與所述第一快捷外設互聯標準交換機的另一第一光纖通信埠通信相連；至少一第三上行埠，用於與所述管理板通信相連；至少一第三下行埠，用於與圖形處理器通信相連。

【0010】 於本發明的一實施例中，所述第二快捷外設互聯標準交換機下行連接的一圖形處理器發出一傳輸線脈衝時，所述第二快捷外設互聯標準交換機的對應第二下行埠接收所述傳輸線脈衝，並向上發出第二匯流排彙報；所述第二快捷外設互聯標準交換機的其他第二下行埠監聽所述第二匯流排彙報，若所述第二匯流排彙報包含的位址屬於所述第二快捷外設互聯標準交換機的其他第二下行埠之一，則所述第二快捷外設互聯標準交換機對應的第二下行埠接收所述傳輸線脈衝；若所述第二匯流排彙報包含的位址不屬於所述第二快捷外設互聯標準交換機的其他第二下行埠之一，

則所述第二快捷外設互聯標準交換機的第二光纖通信埠向上轉發所述傳輸線脈衝至所述第一快捷外設互聯標準交換機。

【0011】 於本發明的一實施例中，所述第一快捷外設互聯標準交換機對應的一第一光纖通信埠接收所述傳輸線脈衝，並向上發出第一匯流排彙報；所述第一快捷外設互聯標準交換機的另一第一光纖通信埠監聽所述第一匯流排彙報，若所述第一匯流排彙報包含的位址屬於所述第一快捷外設互聯標準交換機的另一第一光纖通信埠，則所述第一快捷外設互聯標準交換機的另一第一光纖通信埠接收所述傳輸線脈衝，並向下轉發所述傳輸線脈衝。

【0012】 於本發明的一實施例中，所述第三快捷外設互聯標準交換機的第三光纖通信埠接收所述傳輸線脈衝，並向下轉發第三匯流排彙報；所述第三快捷外設互聯標準交換機的第三下行埠監聽所述第三匯流排彙報，若所述第三匯流排彙報包含的位址屬於所述第三快捷外設互聯標準交換機的第三下行埠之一，則所述第三快捷外設互聯標準交換機對應的第三下行埠接收所述傳輸線脈衝，並向下轉發所述傳輸線脈衝；與所述第三快捷外設互聯標準交換機對應的第三下行埠相連的圖形處理器接收所述傳輸線脈衝。

【0013】 於本發明的一實施例中，所述第二快捷外設互聯標準交換機下行連接的複數個圖形處理器彼此間通過NVLink埠通信或通過所述第二快捷外設互聯標準交換機通信。

【0014】 於本發明的一實施例中，所述第三快捷外設互聯標準交換機下行連接複數個圖形處理器彼此間通過NVLink埠通信或通過所述第三快捷外設互聯標準交換機通信。

【0015】 於本發明的一實施例中，所述管理板對所述第一快捷外設互聯標準交換機、第二快捷外設互聯標準交換機和第三快捷外設互聯標準交換機進行動態管理。

【0016】 於本發明的一實施例中，所述第二快捷外設互聯標準交換機的一第二下行埠與網卡相連；或/和所述第三快捷外設互聯標準交換機的一第三下行埠與網卡相連。

【0017】 如上所述，本發明所述的圖形處理器系統，具有以下有益效果。

【0018】 本發明所述的快捷外設互聯標準交換機的光纖通信模式主要用於多交換機組網，利用管理埠對交換機的動態管理，可以實現多主機（CPU中央處理器）和多終端（GPU圖形處理器）之間的靈活通信。

【0019】 本發明可以解決GPU相互之間的P2P通信頻寬偏低的問題，可以給GPU和CPU之間提供高頻寬，最多可以提供六組X16 PCIE（快捷外設互聯標準）通信頻寬。

【0020】 本發明將二級快捷外設互聯標準交換機通過網卡接入網路中後，可實現同一網路中的不同GPU系統間的資料直接交換，而不需要經過CPU及其記憶體，極大地提高了GPU系統間的資料交換能力。

【0021】 以上之關於本揭露內容之說明及以下之實施方式之說明係用以示範與解釋本發明之精神與原理，並且提供本發明之專利申請範圍更進一步之解釋。

【圖式簡單說明】

【0022】

圖1顯示為本發明實施例所述的圖形處理器系統的一種示例性實現結構示意圖。

圖2顯示為本發明實施例所述的第一快捷外設互聯標準交換機的一種示例性埠結構示意圖。

圖3顯示為本發明實施例所述的第二快捷外設互聯標準交換機的一種示例性埠結構示意圖。

圖4顯示為本發明實施例所述的第三快捷外設互聯標準交換機的一種示例性埠結構示意圖。

圖5顯示為本發明實施例所述的第三快捷外設互聯標準交換機的一種示例性圖形處理器之間的通信流程示意圖。

圖6顯示為本發明實施例所述的GPU與GPU之間的一種示例性通信結構示意圖。

圖7顯示為本發明實施例所述的GPU與CPU之間的一種示例性通信結構示意圖。

圖8顯示為本發明實施例所述的GPU與CPU之間的另一種示例性通信結構示意圖。

【實施方式】

【0023】 以下通過特定的具體實例說明本發明的實施方式，本領域技術人員可由本說明書所揭露的內容輕易地瞭解本發明的其他優點與功效。本發明還可以通過另外不同的具體實施方式加以實施或應用，本說明書中的各項細節也可以基於不同觀點與應用，在沒有背離本發明的精神下進行各種修飾或改變。需說明的是，在不衝突的情況下，以下實施例及實施例中的特徵可以相互組合。

【0024】 需要說明的是，以下實施例中所提供的圖示僅以示意方式說明本發明的基本構想，遂圖式中僅顯示與本發明中有關的組件而非按照實際實施時的元件數目、形狀及尺寸繪製，其實際實施時各元件的型態、數量及比例可為一種隨意的改變，且其元件佈局型態也可能更為複雜。

【0025】 多圖形處理器（Graphics Processing Unit，GPU）系統除了對 GPU 與中央處理器（Central Processing Unit Processor，CPU）之間有高頻寬性能有需求之外，還要求 GPU 之間有很高的頻寬能力，即 P2P（Peer to Peer，點對點）頻寬。

【0026】 為便於描述，後續說明書和附圖中將採用業內英文縮寫簡稱來描述，中央處理器為 CPU，圖形處理器為 GPU，快捷外設互聯標準為 PCIE。

【0027】 為了提高多 GPU 系統的頻寬，本發明提供一種圖形處理器系統，如圖 1 所示，所述圖形處理器系統 100 包括：一中央處理器（CPU）110，一第一快捷外設互聯標準（Peripheral Component Interconnect Express，PCIE）交換機 120，一第二快捷外設互聯標準（Peripheral Component Interconnect Express，PCIE）交換機 130，一第三快捷外設互聯標準（Peripheral Component Interconnect Express，PCIE）交換機 140，或/和一管理板 170。

【0028】 所述第一 PCIE 交換機 120 與所述 CPU110 相連。

【0029】 參見圖 2 所示，於本發明一實施例中，所述第一 PCIE 交換機 120 包括：至少一第一主機埠 121，至少二個第一光纖通信埠 122，或/和至少一第一上行埠 123。所述至少一第一主機埠 121 用於與所述 CPU 通信相連；所述至少二個第一光纖通信埠 122 用於分別與所述第二 PCIE 交

換機和所述第三 PCIE 交換機通信相連；所述至少一第一上行埠 123 用於與所述管理板通信相連。

【0030】 所述第二 PCIE 交換機 130 與所述第一 PCIE 交換機 120 相連；所述第二 PCIE 交換機 130 下行連接複數個 GPU150。所述第二 PCIE 交換機 130 下行連接的複數個 GPU150 彼此間通過 NVLink 埠通信或通過所述第二 PCIE 交換機 130 通信。

【0031】 參見圖 3 所示，于本發明一實施例中，所述第二 PCIE 交換機 130 包括：至少一第二光纖通信埠 131，至少一第二上行埠 132，或/和至少一第二下行埠 133。所述至少一第二光纖通信埠 131 用於與所述第一 PCIE 交換機的一第一光纖通信埠 122 通信相連；所述至少一第二上行埠 132 用於與所述管理板 170 通信相連；所述至少一第二下行埠 133 用於與 GPU150 通信相連。所述第二 PCIE 交換機的一第二下行埠 133 可與網卡相連。

【0032】 所述第三 PCIE 交換機 140 與所述第一 PCIE 交換機 120 相連；所述第三 PCIE 交換機 140 下行連接複數個 GPU160。所述第三 PCIE 交換機 140 下行連接複數個 GPU160 彼此間通過 NVLink 埠通信或通過所述第三 PCIE 交換機 140 通信。

【0033】 參見圖 4 所示，于本發明一實施例中，所述第三 PCIE 交換機 140 包括：至少一第三光纖通信埠 141，至少一第三上行埠 142，或/和

至少一第三下行埠 143。所述至少一第三光纖通信埠 141 用於與所述第一 PCIE 交換機的另一第一光纖通信埠 122 通信相連；所述至少一第三上行埠 142 用於與所述管理板 170 通信相連；所述至少一第三下行埠 143 用於與 GPU160 通信相連。所述第三 PCIE 交換機的一第三下行埠 143 可與網卡相連。

【0034】 所述管理板 170 與所述第一 PCIE 交換機 120、第二 PCIE 交換機 130 和第三 PCIE 交換機 140 分別相連，對所述第一 PCIE 交換機、第二 PCIE 交換機和第三 PCIE 交換機進行管理設置。于本發明一實施例中，所述管理板可對所述第一 PCIE 交換機、第二 PCIE 交換機和第三 PCIE 交換機進行動態管理。

【0035】 于本發明一實施例中，若所述第二 PCIE 交換機下行連接的一 GPU 與所述第三 PCIE 交換機下行連接的一 GPU 之間需要進行通信，則通信過程可參見圖 5 所示，包括：

【0036】 S501，所述第二 PCIE 交換機下行連接的一 GPU 發出一傳輸線脈衝；

【0037】 S502，所述第二 PCIE 交換機的對應第二下行埠接收所述傳輸線脈衝，並向上發出第二匯流排彙報；

【0038】 S503，所述第二 PCIE 交換機的其他第二下行埠監聽所述第二匯流排彙報；

【0039】 S504，若所述第二匯流排彙報包含的位址屬於所述第二 PCIE 交換機的其他第二下行埠之一，則所述第二 PCIE 交換機對應的第二下行埠接收所述傳輸線脈衝；

【0040】 S505，若所述第二匯流排彙報包含的位址不屬於所述第二 PCIE 交換機的其他第二下行埠之一，則所述第二 PCIE 交換機的第二光纖通信埠向上轉發所述傳輸線脈衝至所述第一 PCIE 交換機；

【0041】 S506，所述第一 PCIE 交換機對應的一第一光纖通信埠接收所述傳輸線脈衝，並向上發出第一匯流排彙報；

【0042】 S507，所述第一 PCIE 交換機的另一第一光纖通信埠監聽所述第一匯流排彙報；

【0043】 S508，若所述第一匯流排彙報包含的位址屬於所述第一 PCIE 交換機的另一第一光纖通信埠，則所述第一 PCIE 交換機的另一第一光纖通信（fabric）埠接收所述傳輸線脈衝，並向下轉發所述傳輸線脈衝；

【0044】 S509，所述第三 PCIE 交換機的第三光纖通信埠接收所述傳輸線脈衝，並向下轉發第三匯流排彙報；

【0045】 S510，所述第三 PCIE 交換機的第三下行埠監聽所述第三匯流排彙報；

【0046】 S511，若所述第三匯流排彙報包含的位址屬於所述第三 PCIE 交換機的第三下行埠之一，則所述第三 PCIE 交換機對應的第三下行

埠接收所述傳輸線脈衝，並向下轉發所述傳輸線脈衝；

【0047】 S512，與所述第三 PCIE 交換機對應的第三下行埠相連的 GPU 接收所述傳輸線脈衝。

【0048】 本發明中，NVLink 是英偉達（NVIDIA）開發並推出的一種匯流排及其通信協定。NVLink 採用點對點結構、串列傳輸，用於 GPU 與 GPU 之間的連接，也可用於 CPU 與 GPU 之間的連接。NVLink 埠是 GPU 與 GPU 之間或者 CPU 與 GPU 之間的點對點通訊連接埠。

【0049】 如圖 6 所示，NVLink 埠實現了 GPU 與 GPU 之間的直接通信，但每個 GPU 只有六組 NVLink，當系統中的 GPU 數量達到八個時，NVLink 就無法涵蓋所有 GPU 的通信。

【0050】 如圖 7 所示，在八 GPU 系統中，CPU 端至少要提供兩組 X16 的 PCIE 埠，然後利用 PCIE 交換機擴展出八組 X16 的 PCIE 埠；當只有一級 PCIE 交換機時，左右兩半部分的 GPU 只能透過 CPU 做 P2P 通信。當有兩級 PCIE 交換機時，可以利用 PCIE 交換機同一橋下兩個不同下行埠之間可以 P2P 通信的特性，但僅限於只有一個上行埠的基本模式。這種拓撲結構下，GPU 系統與 CPU 之間只有一組 X16 PCIE 埠，頻寬遠遠不夠。因此，將三個 PCIE 交換機的工作模式調整為光纖通信（fabric）模式，向下八組 PCIE 埠和八個 GPU 相連，向上兩組 PCIE 埠和 CPU 相連，一級 PCIE 交換機分別和兩個二級 PCIE 交換機相連，三個 PCIE 交換機各有一

個管理埠通過 PEX8608 掛在一個管理板（mCPU）上，如圖 8 所示。

【0051】 圖 8 所示的三個 PCIE 交換機（PEX9797）均設置為光纖通信（fabric）模式，第一個 PCIE 交換機的 port0 和 port4 設置為主機埠（host port），與 CPU 的兩個根埠（root port）相連，如果 CPU 的 PCIE 資源充足，另外可以將第一個 PCIE 交換機的 port16 和 port20 也設置為主機埠（host port）和 CPU 相連，這樣可以給 GPU 和 CPU 的通信提供最大的頻寬性能。第一個 PCIE 交換機的 port8 和 port12 分別與第二、三個 PCIE 交換機的 port0 相連，均設置為光纖通信埠（fabric port）；三個 PCIE 交換機的 port24 均設置為管理埠（management port），作為上行埠，通過 PEX8608 與管理板（mCPU）相連；其餘均為下行埠和 PCIE 設備相連，包括八個 GPU。如果接網卡的話，插在第二、三個 PCIE 交換機的下行埠上即可。

【0052】 圖 8 所示的系統啟動順序如下：

【0053】 1) 啟動 GPU 板子；

【0054】 2) 啟動 mCPU 系統，並啟動 mCPU fabric driver（光纖驅動器），分配 GPU0-3 給 Host port 0，分配 GPU4-7 給 Host port 4；

【0055】 3) 啟動 Host 系統。

【0056】 以 GPU0 與 GPU4 通信為例，TLP（Transmission Line Pulse，傳輸線脈衝）傳輸過程為：

【0057】 1) GPU0 發出 TLP，switch 2 的 port4 接收該 TLP，並向上發出匯流排事物；

【0058】 2) Switch2 (即第二 PCIE 交換機) 上的其它 port 監聽，發現沒有 port 的 Address Trap (位址陷阱) 包含這個位址，於是由該 switch2 的 upstream port (上行埠 port 0) 向上轉發；

【0059】 3) Switch1 的 port 8 收到 TLP，並向上發出匯流排事物，switch1 的 port 12 包含這個位址，所以 port 12 接收這個 TLP，並向下轉發該 TLP；

【0060】 4) Switch 3 的 upstream port (上行埠) 接收該 TLP，並向下發出匯流排事物，switch 3 的 port4 包含該 TLP 的位址，將接收該 TLP，並向下發 TLP；

【0061】 5) GPU4 接收該 TLP。

【0062】 本發明所述的 PCIE 交換機的光纖通信模式主要用於多交換機組網，利用管理埠對交換機的動態管理，可以實現多主機 (CPU) 和多終端 (GPU) 之間的靈活通信。

【0063】 本發明可以解決 GPU 相互之間的 P2P 通信頻寬偏低的問題，可以給 GPU 和 CPU 之間提供高頻寬，最多可以提供六組 X16 PCIE 通信頻寬。

【0064】 本發明將二級 PCIE 交換機通過網卡接入網路中後，可實現

同一網路中的不同 GPU 系統間的資料直接交換，而不需要經過 CPU 及其記憶體，極大地提高了 GPU 系統間的資料交換能力。

【0065】 綜上所述，本發明有效克服了現有技術中的種種缺點而具高度產業利用價值。

【0066】 上述實施例僅例示性說明本發明的原理及其功效，而非用於限制本發明。任何熟悉此技術的人士皆可在不違背本發明的精神及範疇下，對上述實施例進行修飾或改變。因此，舉凡所屬技術領域中具有通常知識者在未脫離本發明所揭示的精神與技術思想下所完成的一切等效修飾或改變，仍應由本發明的權利要求所涵蓋。

【符號說明】

【0067】

- 100 圖形處理器系統
- 110 中央處理器
- 120 第一快捷外設互聯標準交換機
- 121 第一主機埠
- 122 第一光纖通信埠
- 123 第一上行埠
- 130 第二快捷外設互聯標準交換機
- 131 第二光纖通信埠

- 132 第二上行埠
- 133 第二下行埠
- 140 第三快捷外設互聯標準交換機
- 141 第三光纖通信埠
- 142 第三上行埠
- 143 第三下行埠
- 150 , 160 圖形處理器
- 170 管理板
- S501~S512 步驟

【發明申請專利範圍】

【第1項】 一種圖形處理器系統，其特徵在於，所述圖形處理器系統包括：

一中央處理器；

一第一快捷外設互聯標準交換機，與所述中央處理器相連；

一第二快捷外設互聯標準交換機，與所述第一快捷外設互聯標準交換機相連；所述第二快捷外設互聯標準交換機下行連接複數個圖形處理器；

一第三快捷外設互聯標準交換機，與所述第一快捷外設互聯標準交換機相連；所述第三快捷外設互聯標準交換機下行連接複數個圖形處理器；以及

一管理板，與所述第一快捷外設互聯標準交換機、第二快捷外設互聯標準交換機和第三快捷外設互聯標準交換機分別相連，對所述第一快捷外設互聯標準交換機、第二快捷外設互聯標準交換機和第三快捷外設互聯標準交換機進行管理設置。

【第2項】 根據請求項 1 所述的圖形處理器系統，其特徵在於，所述第一快捷外設互聯標準交換機包括：

至少一第一主機埠，用於與所述中央處理器通信相連；

至少二個第一光纖通信埠，用於分別與所述第二快捷外設互聯標準交換機和所述第三快捷外設互聯標準交換機通信相連；

以及

至少一第一上行埠，用於與所述管理板通信相連。

【第3項】 根據請求項 2 所述的圖形處理器系統，其特徵在於，所述第二快捷外設互聯標準交換機包括：

至少一第二光纖通信埠，用於與所述第一快捷外設互聯標準交換機的一第一光纖通信埠通信相連；

至少一第二上行埠，用於與所述管理板通信相連；以及

至少一第二下行埠，用於與圖形處理器通信相連。

【第4項】 根據請求項 3 所述的圖形處理器系統，其特徵在於，所述第三快捷外設互聯標準交換機包括：

至少一第三光纖通信埠，用於與所述第一快捷外設互聯標準交換機的另一第一光纖通信埠通信相連；

至少一第三上行埠，用於與所述管理板通信相連；以及

至少一第三下行埠，用於與圖形處理器通信相連。

【第5項】 根據請求項 4 所述的圖形處理器系統，其特徵在於：所述第二快捷外設互聯標準交換機下行連接的一圖形處理器發出一傳輸線脈衝時，所述第二快捷外設互聯標準交換機的對應

第二下行埠接收所述傳輸線脈衝，並向上發出第二匯流排彙報；所述第二快捷外設互聯標準交換機的其他第二下行埠監聽所述第二匯流排彙報，若所述第二匯流排彙報包含的位址屬於所述第二快捷外設互聯標準交換機的其他第二下行埠之一，則所述第二快捷外設互聯標準交換機對應的第二下行埠接收所述傳輸線脈衝；若所述第二匯流排彙報包含的位址不屬於所述第二快捷外設互聯標準交換機的其他第二下行埠之一，則所述第二快捷外設互聯標準交換機的第二光纖通信埠向上轉發所述傳輸線脈衝至所述第一快捷外設互聯標準交換機。

【第6項】 根據請求項 5 所述的圖形處理器系統，其特徵在於：所述第一快捷外設互聯標準交換機對應的一第一光纖通信埠接收所述傳輸線脈衝，並向上發出第一匯流排彙報；所述第一快捷外設互聯標準交換機的另一第一光纖通信埠監聽所述第一匯流排彙報，若所述第一匯流排彙報包含的位址屬於所述第一快捷外設互聯標準交換機的另一第一光纖通信埠，則所述第一快捷外設互聯標準交換機的另一第一光纖通信埠接收所述傳輸線脈衝，並向下轉發所述傳輸線脈衝。

【第7項】 根據請求項 6 所述的圖形處理器系統，其特徵在於：所述第三快捷外設互聯標準交換機的第三光纖通信埠接收所

述傳輸線脈衝，並向下轉發第三匯流排彙報；所述第三快捷外設互聯標準交換機的第三下行埠監聽所述第三匯流排彙報，若所述第三匯流排彙報包含的位址屬於所述第三快捷外設互聯標準交換機的第三下行埠之一，則所述第三快捷外設互聯標準交換機對應的第三下行埠接收所述傳輸線脈衝，並向下轉發所述傳輸線脈衝；與所述第三快捷外設互聯標準交換機對應的第三下行埠相連的圖形處理器接收所述傳輸線脈衝。

【第8項】 根據請求項 1 所述的圖形處理器系統，其特徵在於：所述第二快捷外設互聯標準交換機下行連接的複數個圖形處理器彼此間通過 NVLink 埠通信或通過所述第二快捷外設互聯標準交換機通信。

【第9項】 根據請求項 1 所述的圖形處理器系統，其特徵在於：所述第三快捷外設互聯標準交換機下行連接複數個圖形處理器彼此間通過 NVLink 埠通信或通過所述第三快捷外設互聯標準交換機通信。

【第10項】 根據請求項 9 所述的圖形處理器系統，其特徵在於：所述管理板對所述第一快捷外設互聯標準交換機、第二快捷外設互聯標準交換機和第三快捷外設互聯標準交換機進行動態管理。

【第11項】 根據請求項 4 所述的圖形處理器系統，其特徵在於：所述第二快捷外設互聯標準交換機的一第二下行埠與網卡相連；或/和所述第三快捷外設互聯標準交換機的一第三下行埠與網卡相連。

【發明圖式】

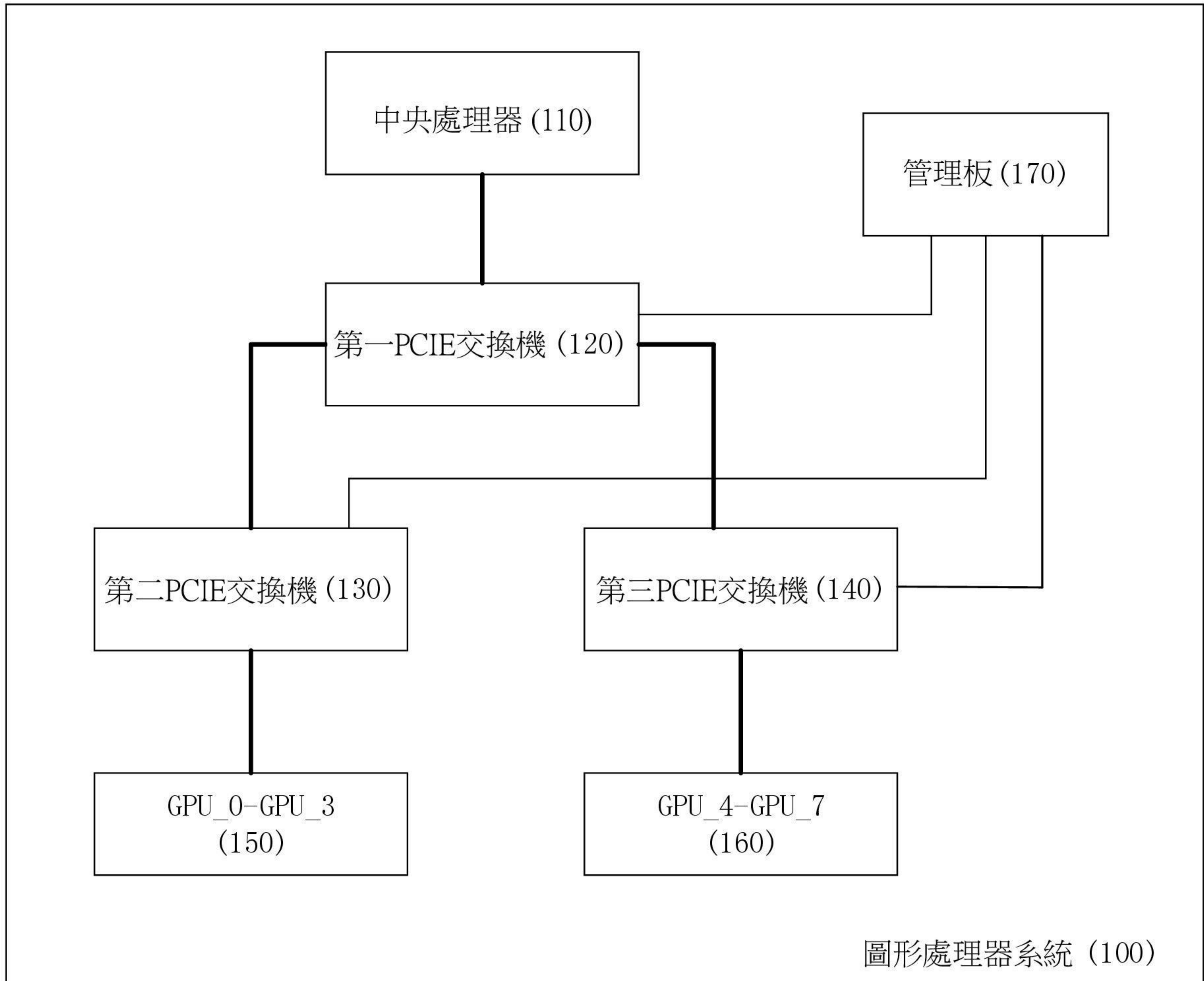


圖 1

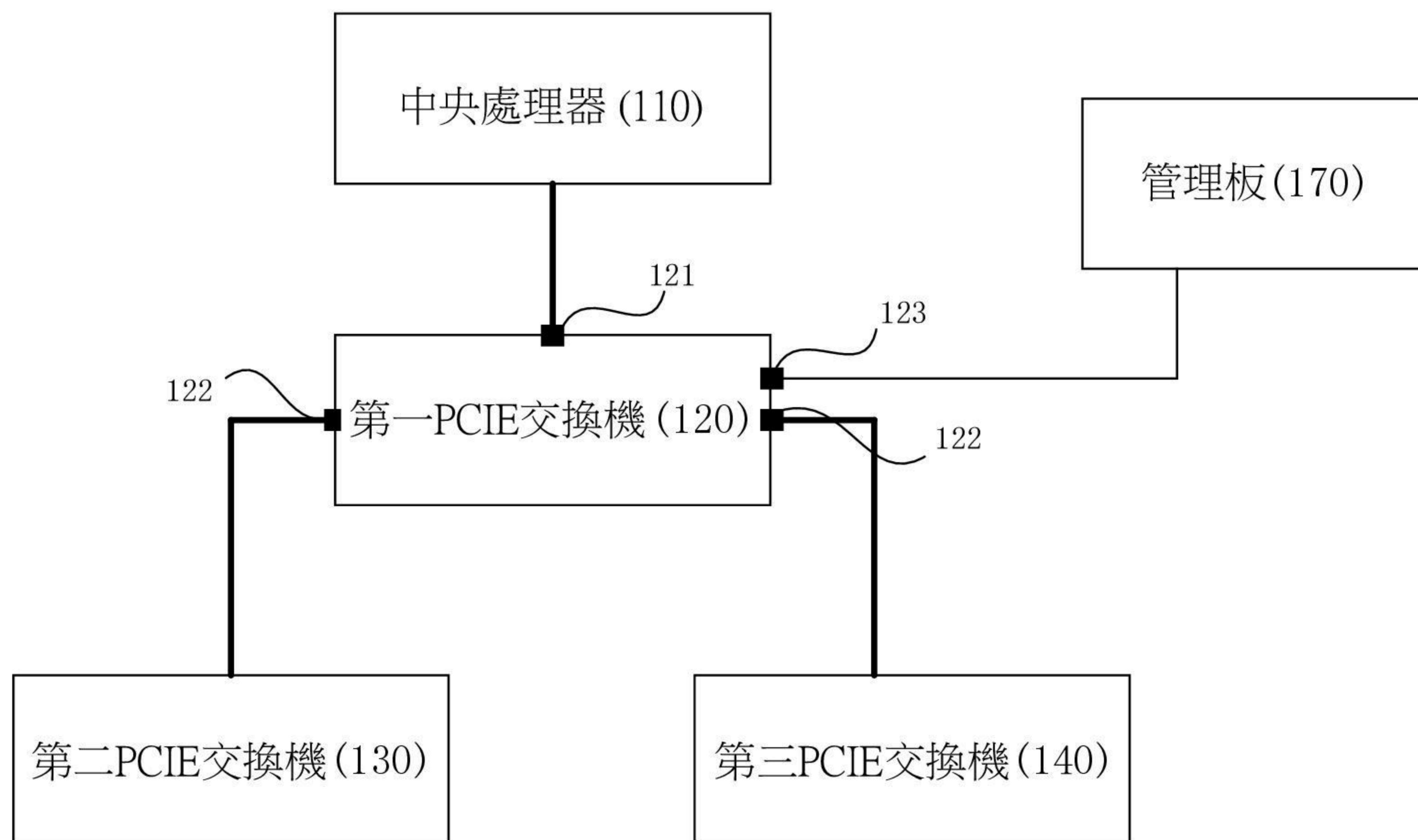


圖 2

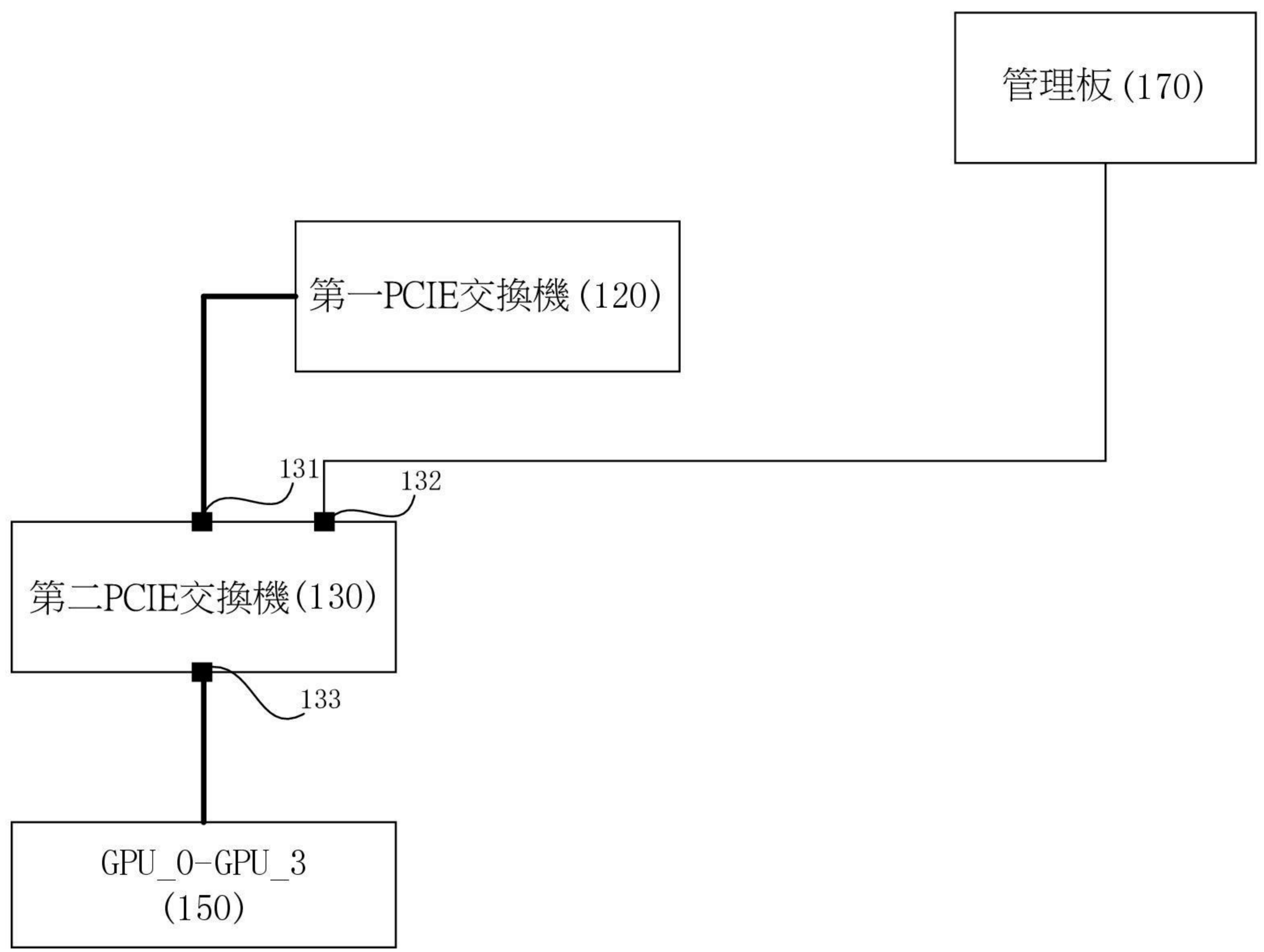


圖 3

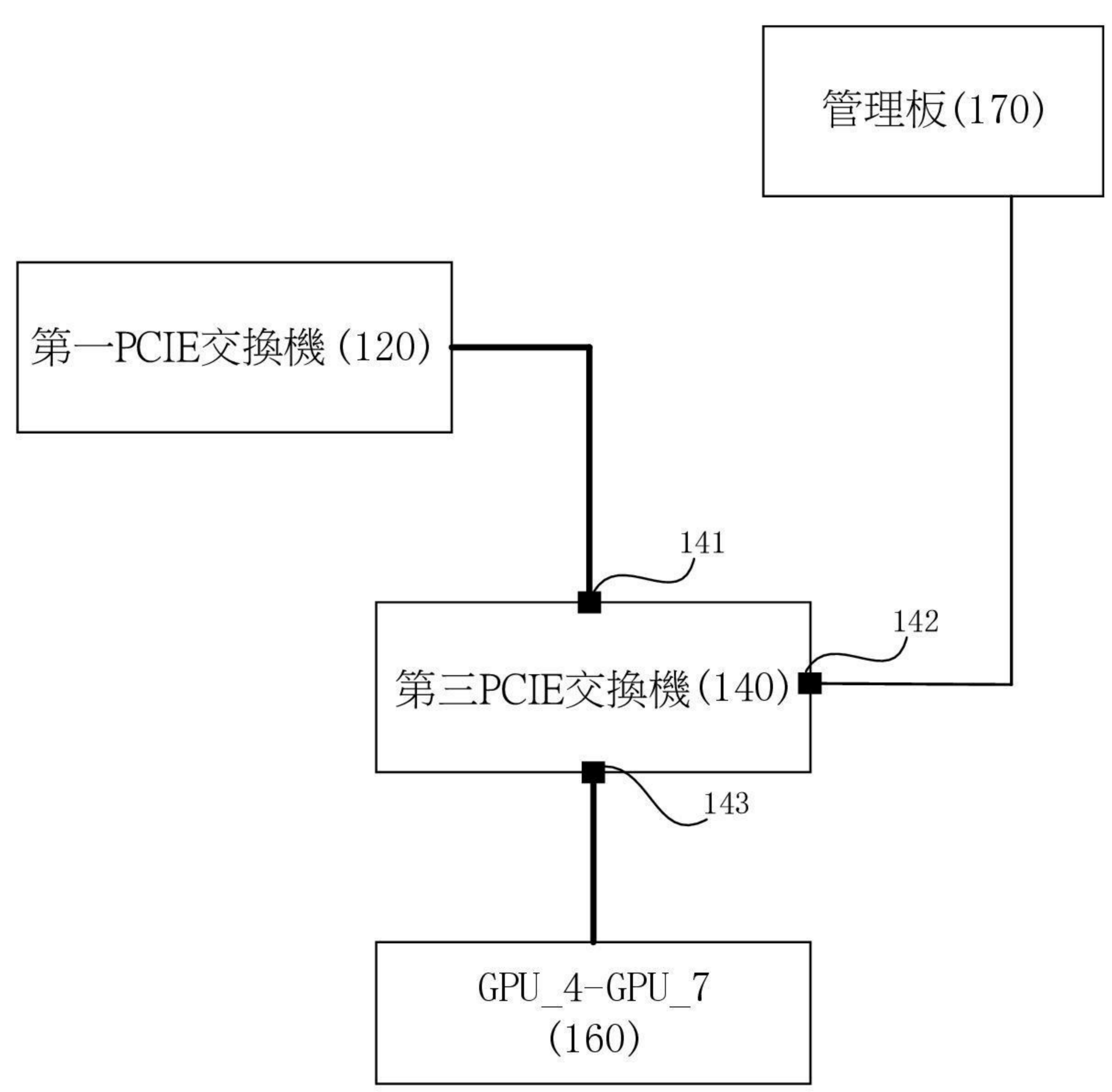


圖 4

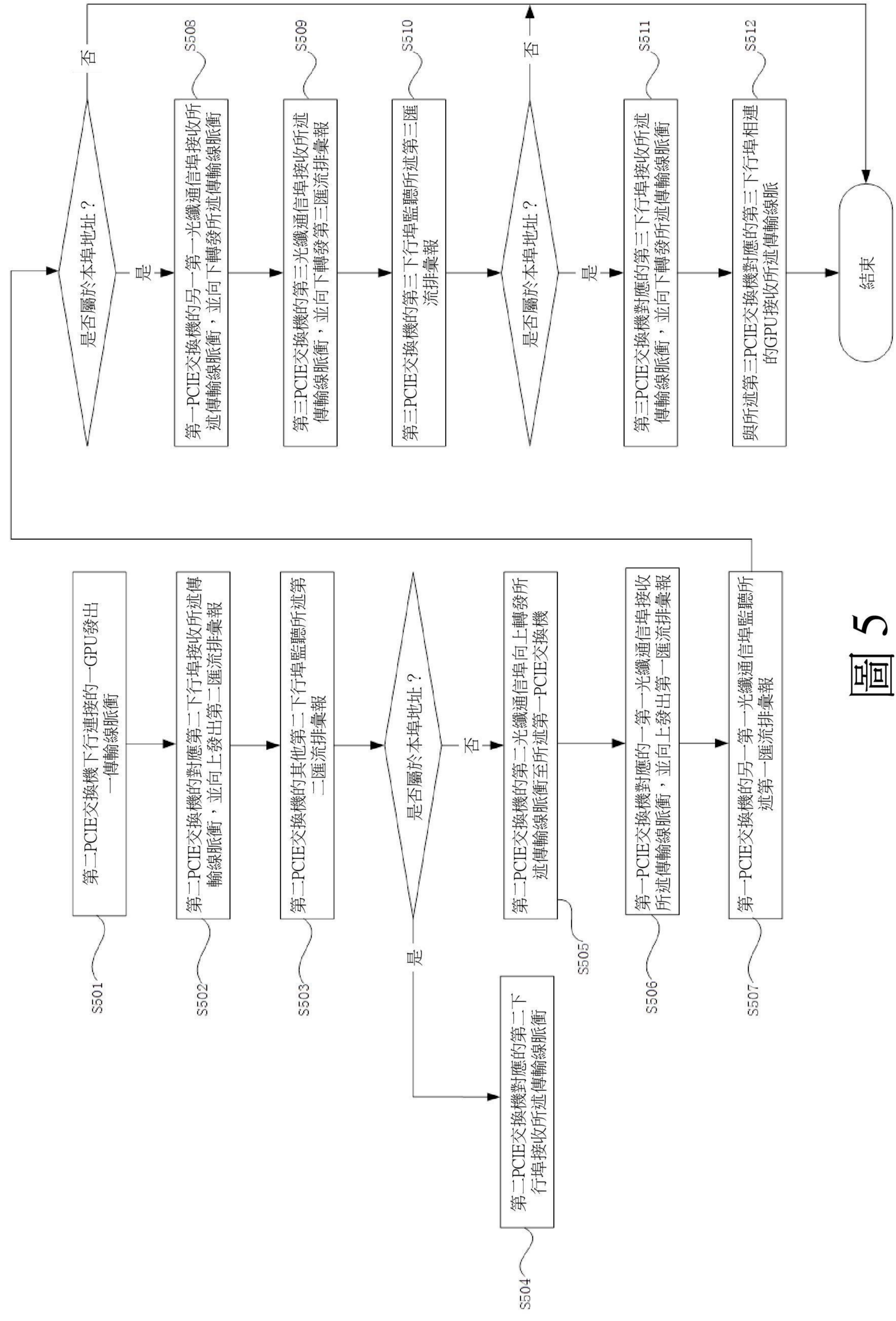


圖 5

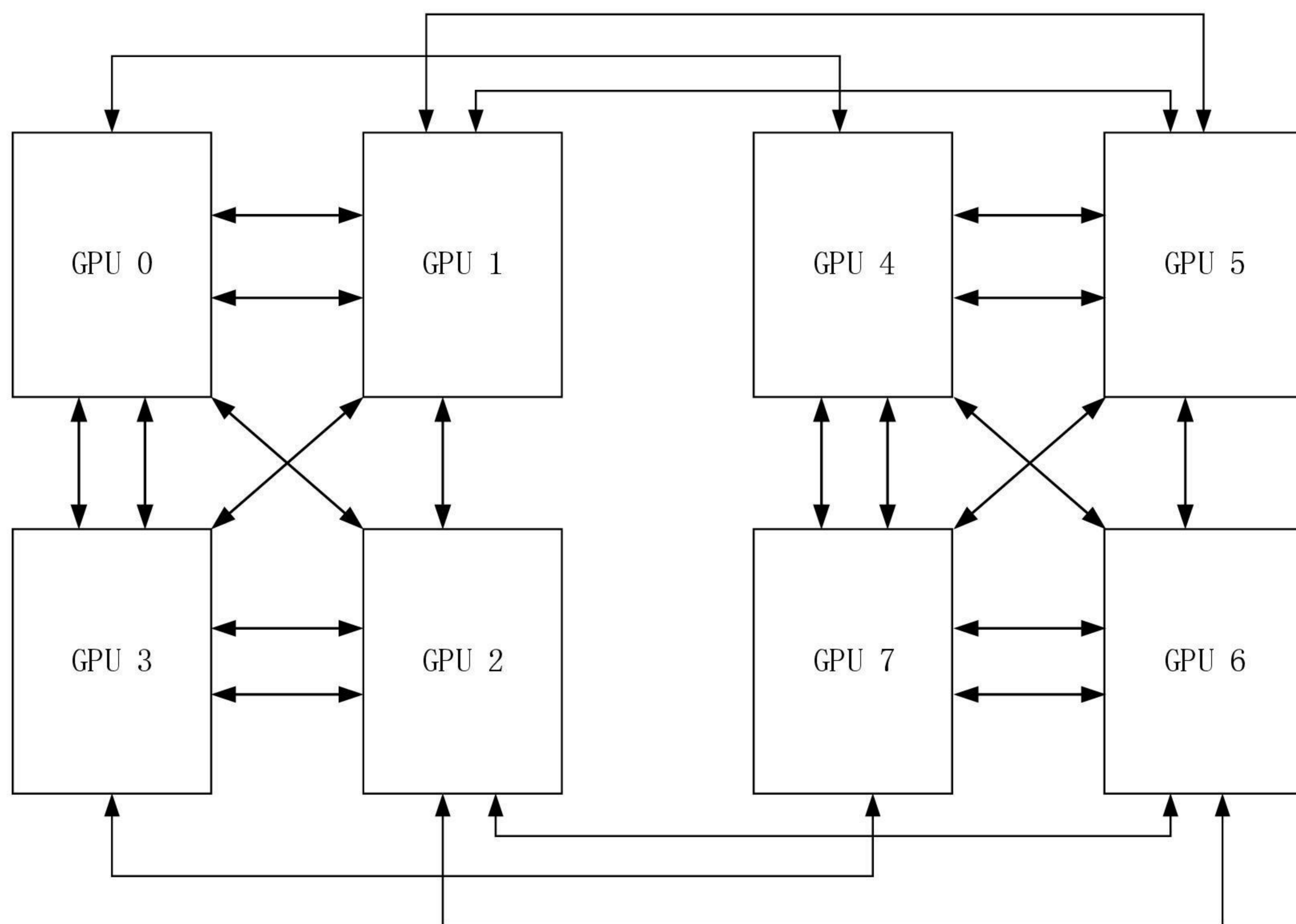


圖 6

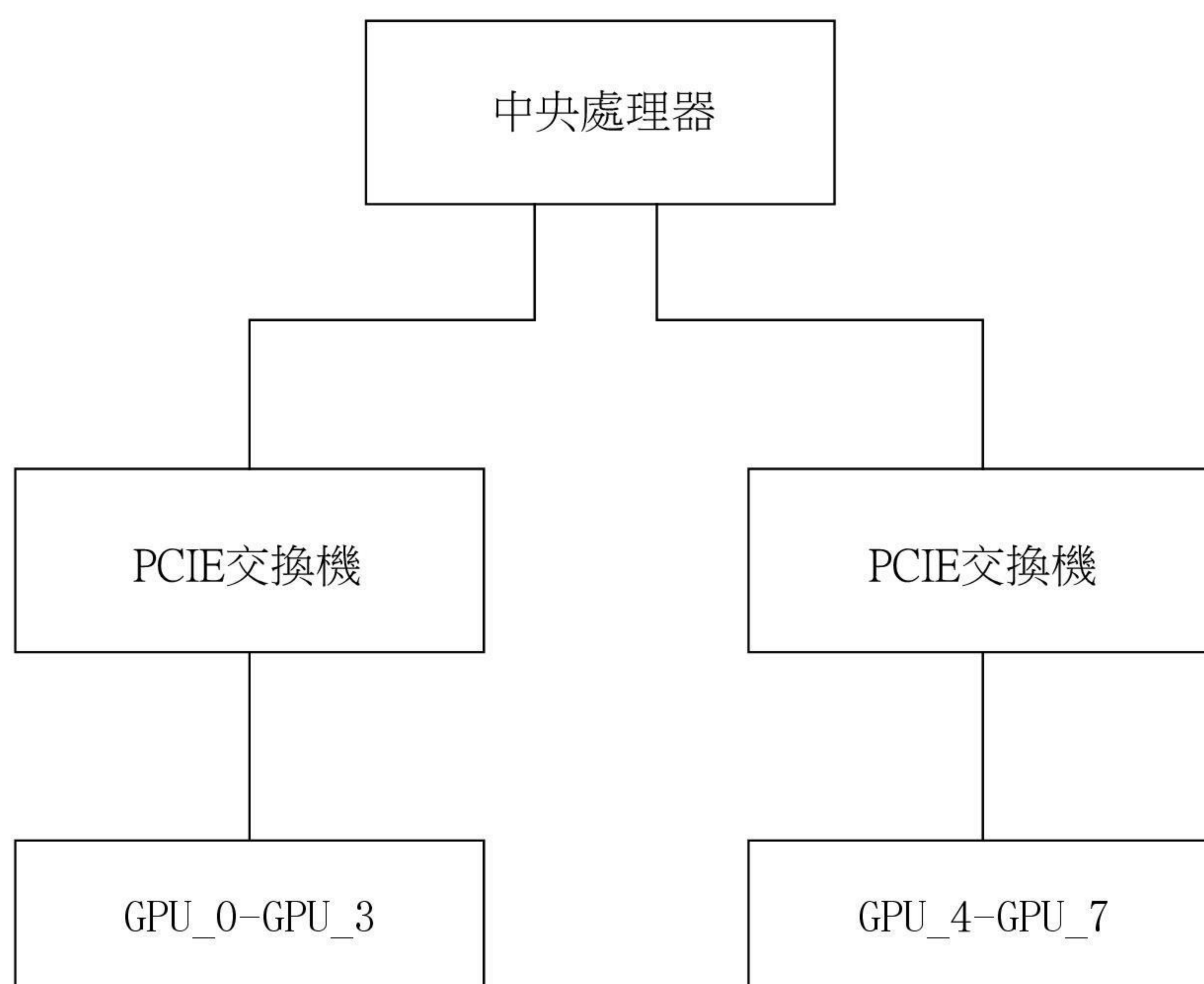


圖 7

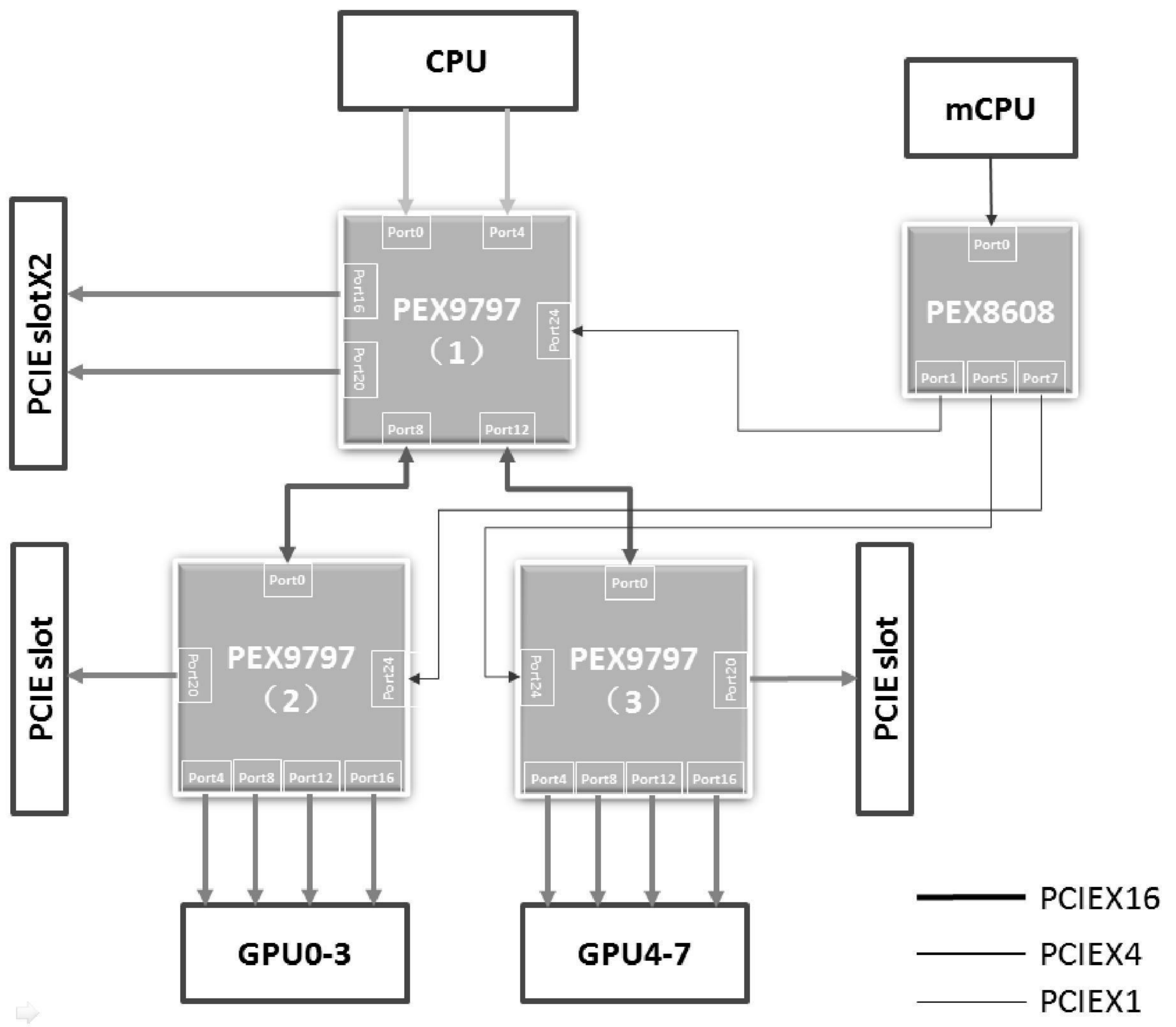


圖 8