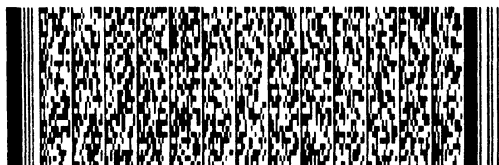


申請日期： 911074	IPC分類	588250
申請案號： 91122965	G06F13/36	

(以上各欄由本局填註)

發明專利說明書

一、 發明名稱	中 文	電腦系統之輸入/輸出節點
	英 文	COMPUTER SYSTEM I/O NODE
二、 發明人 (共2人)	姓 名 (中文)	1. 史帝芬C·艾尼斯 2. 雷利D·修伊特
	姓 名 (英文)	1. STEPHEN C. ENNIS 2. LARRY D. HEWITT
	國 籍 (中英文)	1. 美國 US 2. 美國 US
	住居所 (中文)	1. 美國·德州78704·奧斯汀·艾羅路2100-B號 2. 美國·德州78746·奧斯汀·班歐河路6103號
	住居所 (英文)	1. 2100-B Airrole Way, Austin, TX 78704, U.S.A. 2. 6103 Bend O' River Drive, Austin, TX 78746, U.S.A.
三、 申請人 (共1人)	名稱或 姓 名 (中文)	1. 高級微裝置公司
	名稱或 姓 名 (英文)	1. ADVANCED MICRO DEVICES, INC.
	國 籍 (中英文)	1. 美國 US
	住居所 (營業所) (中文)	1. 美國·加州94088-3453·桑尼威·第1AMD區·M/S 68·郵政信箱3453號 (本地址與前向貴局申請者相同)
	住居所 (營業所) (英文)	1. One AMD Place, P.O. Box 3453, Sunnyvale, Mail Stop 68, CA, 94088-3453, U.S.A.
	代表人 (中文)	1. 理查·J·諾迪
	代表人 (英文)	1. RICHARD J. RODDY



一、本案已向

國家(地區)申請專利	申請日期	案號	主張專利法第二十四條第一項優先權
美國 US	2001/10/15	09/978,349	有

二、☐主張專利法第二十五條之一第一項優先權：

申請案號：

日期：

三、主張本案係符合專利法第二十條第一項☐第一款但書或☐第二款但書規定之期間

日期：

四、☐有關微生物已寄存於國外：

寄存國家：

寄存機構：

寄存日期：

寄存號碼：

☐有關微生物已寄存於國內(本局所指定之寄存機構)：

寄存機構：

寄存日期：

寄存號碼：

☐熟習該項技術者易於獲得，不須寄存。

五、發明說明 (1)

[發明所屬之技術領域]

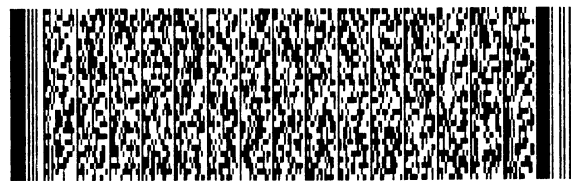
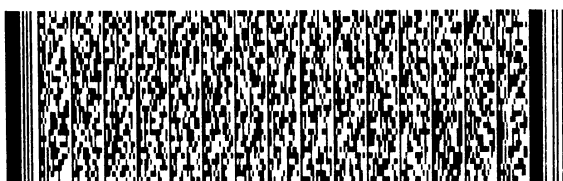
本發明係關於電腦系統輸入/輸出 (Input/Output, I/O), 並且尤其係關於在輸入/輸出節點中之資訊交易處理。

[先前技術]

在典型的電腦系統中, 一個或一個以上之處理器可以經由一個或一個以上之匯流排與輸入/輸出裝置做通訊。該輸入/輸出裝置可以透過輸入/輸出橋接器連接至該處理器, 該橋接器處理在連接至該輸入/輸出裝置之周邊匯流排與連接至該處理器之共用匯流排之間之資料的轉移。此外, 該輸入/輸出橋接器可處理在系統記憶體與該輸入/輸出裝置之間或在該系統記憶體與該處理器之間之資料的轉移。

遺憾的是很多匯流排系統仍然具有許多缺點。例如, 許多連結至匯流排之裝置對於在該匯流排上之裝置驅動訊號可能存在相對大的電容量。此外, 在共用匯流排上之該多重連結點在高訊號頻率下產生減少訊號完整性之訊號反射。因此, 在匯流排上之訊號頻率通常保持相對較低藉以維持訊號完整性於可接受的位準。該相對低的訊號頻率降低訊號頻寬, 而限制了連結至該匯流排之裝置的效能。

對於較大數量的裝置之擴充性的缺乏是共用匯流排系統之另外的缺點。共用匯流排之可獲得的頻寬實質上是固定的 (並且若加入額外的裝置造成在訊號頻率於匯流排上之降低, 則可能減少)。一旦連結至該匯流排 (不論是直接

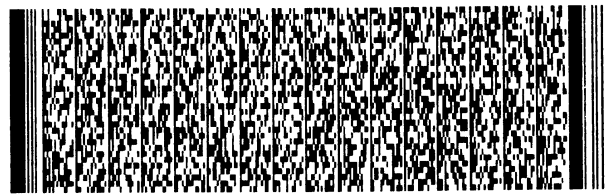
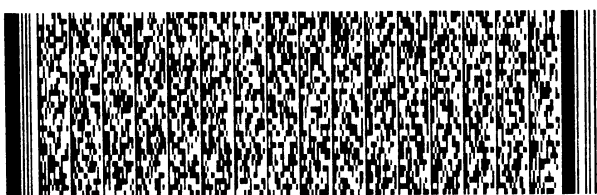


五、發明說明 (2)

地或非直接地)之裝置的頻寬需求超過該匯流排之可獲得的頻寬時，當嘗試存取該匯流排時，裝置將會常常延遲，並且包含該共用匯流排之電腦系統之整體的效能將非常有可能降低。一個藉由輸入/輸出裝置所使用之共用匯流排之例子為周邊組件互連接(peripheral component interconnect, PCI)匯流排。

很多輸入/輸出橋接裝置使用緩衝機制以緩衝許多來自該PCI匯流排至最終的目的匯流排之待解決的資訊交易。然而緩衝可能導致在該PCI匯流排上之延遲。當一連串的資訊交易(後文中簡稱交易)緩衝於佇列中以及等待傳輸至目的匯流排時，延遲將可能造成，並且延遲發生於該目的匯流排上，該目的匯流排停止傳遞前進。接著將容許那些等待的交易去完成之交易到達於該佇列處並且儲存於其它交易後面。為了打破該延遲，在佇列中之該交易必須設法重新排序以允許該新到達交易傳送於該待解決的交易之前。因此，為了避免諸如此種的情況，該PCI匯流排規範訂定一組支配PCI匯流排交易之處理及排序之重新排序規則。

為了克服共用匯流排之某些缺點，某些電腦系統可能在裝置或節點之間使用封包模式通訊。在此類的系統中，節點可以藉由交換資訊之封包而彼此通訊。通常，"節點"為能夠在內連接上參與交易之裝置。例如，該內連接可以是封包型模式的，並且該節點可以經由配置以接收及發送封包。一般來說，"封包"為在兩個節點之間之通訊，該兩



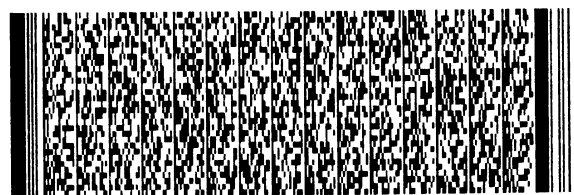
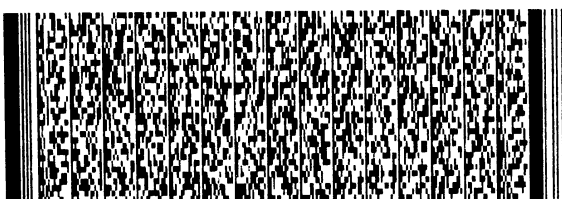
五、發明說明 (3)

個節點為：發送該封包之初始或"來源"節點及接收該封包之目的或"目標"節點。當封包到達該目標節點時，該目標節點接受由該封包所傳遞之資訊並且於內部處理該資料。位在該來源及目標節點之間之在通訊路徑上之節點將可轉送或傳遞來自該來源節點之該封包給該目標節點。

此外，也有系統使用封包模式通訊及匯流排模式通訊之組合。例如，系統可連接至PCI匯流排及諸如AGP之繪圖匯流排。該PCI匯流排可以連接至封包匯流排介面，該封包匯流排介面接著可以轉譯PCI匯流排交易成為用於在封包匯流排上傳輸之封包交易。同樣地該繪圖匯流排可以連接至可轉譯AGP交易成為封包交易之AGP介面。每個介面可以與結合其中一個處理器之主橋接器或在某些例子中與另一個周邊裝置做通訊。

當PCI裝置初始化交易時，如同於PCI區域匯流排規範中所提及的，封包型交易可能受到由相同的排序規則所限制。該相同對於指定於該PCI匯流排之封包交易而言可能是合法的。由於可能發生在封包匯流排介面處之交易延遲可能在該封包匯流排介面處造成停滯，這些排序規則在該封包模式交易中仍然可以觀察到。此停滯可能造成進一步的延遲回到該封包匯流排線路上。此外，AGP交易可以依循一組交易順序規則以確保資料之適當的傳送。

視輸入/輸出節點之配置而定，不論是在主橋接器之方向上或遠離該主橋接器之方向上，交易可以透過節點傳遞至另一個節點。另外，交易可以在特定的節點處引入封



五、發明說明 (4)

包傳輸中。在任一個情況下，當該交易沿著該通訊路徑發送時，可以控制該交易之輸入/輸出節點架構可能是必須具備的。

[發明內容]

在此將揭露電腦系統輸入/輸出節點之各種實施例。在其中一項實施例中，用於電腦系統之輸入/輸出節點包含經由配置以在第一通訊路徑上接收第一指令之第一接收器單元及經由連接以在第二通訊路徑上發送對應該第一指令之第一對應指令之第一發送器單元。該輸入/輸出節點亦包含經由配置以在第三通訊路徑上接收第二指令之第二接收器單元及經由連接以在第四通訊路徑上發送對應該第二指令之第二對應指令之第二發送器單元。在一項特定的實現中，該通訊路徑可以是諸如例如 HyperTransport™ 連結之點對點通訊連結。再者，該輸入/輸出節點可包含經由連接以接收來自於該第一接收器及該第二接收器之選擇的指令之橋接器單元，並且經由配置以在周邊匯流排上發送對應於該選擇的指令之指令。

在一項特定的實施例中，輸入/輸出節點包含經由連接以控制來自該第一通訊路徑之指令至該第二通訊路徑及至該周邊匯流排之傳送的控制單元。此外，該控制單元經由連接以控制來自該第三通訊路徑之指令至該第四通訊路徑及至該周邊匯流排之傳送。該控制單元亦經由配置以控制來自該周邊匯流排之指令至該第二通訊路徑及該第四通訊路徑之傳送。基於來自該第一接收器、該第二接收器及



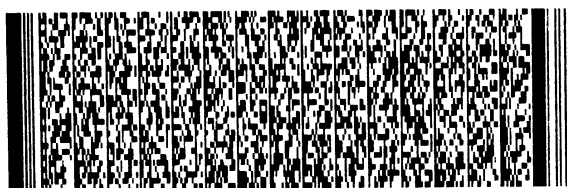
五、發明說明 (5)

該橋接器單元所接收之複數個控制指令，該控制單元更經由配置以選擇性地控制該指令之傳送。

[實施方式]

茲參照第 1 圖，該圖顯示電腦系統之一項實施例之方塊圖。該電腦系統包含每個藉由相干封包匯流排 15 做內連接之處理器 10A 至 10D。每個區段的相干封包匯流排 15 在每個處理器 10A 至 D 之間可形成點對點連結。雖然該圖顯示使用點對點連結之四個處理器，需要注意的是其它數量的處理器亦可以使用並且其它形式的匯流排亦可在該處理器之間做內連接。該電腦系統亦包含三個標示為 20、30 及 40 之輸入/輸出節點，每個節點分別藉由輸入/輸出封包匯流排 50B 及 50C 以串連方式連接在一起。輸入/輸出封包匯流排 50A 連接於主節點/處理器 10A 及輸入/輸出節點 20 之間。處理器 10A 描述為主節點，該節點可包含用於與輸入/輸出封包匯流排 50A 做通訊之主橋接器。處理器 10B 至 D 亦可包含用於与其它輸入/輸出封包匯流排 (未顯示) 做通訊之主橋接器。由輸入/輸出封包匯流排 50A 至 C 所形成之該通訊連結亦可以稱為點對點連結。輸入/輸出節點 20 連接至一對周邊匯流排 25A 至 B。輸入/輸出節點 30 連接至繪圖匯流排 35，而輸入/輸出節點 40 連接至額外的周邊匯流排 45。

處理器 10A 至 10D 例如每個可以用諸如 Athlon™ 微處理器之 x86 微處理器做為例證說明。此外，諸如輸入/輸出封包匯流排 50A 至 50C 之封包匯流排的其中一個例子可以是非同步 HyperTransport™。周邊匯流排 25A 至 B 及周邊匯流排



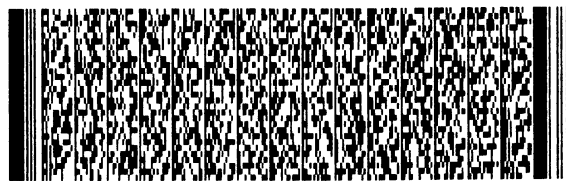
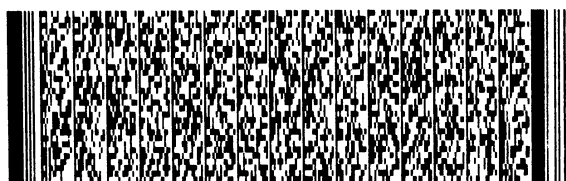
五、發明說明 (6)

45可以用諸如周邊組件互連接(PCI)匯流排之共用周邊匯流排來做為例證說明。匯流排35例如可以用加速繪圖埠(accelerated graphics port, AGP)做為例證說明。然而，需要瞭解的是其它形式的微處理器及其它形式的周邊匯流排亦可以使用。

需要注意的是雖然該圖顯示三個輸入/輸出節點連接至主處理器10A，但是其它實施例可具有其它數目的節點並且那些節點可以以其它形態做連接。說明於第1圖中之該串連形態是為了本身的容易瞭解而顯示。

在說明的實施例中，處理器10A之主橋接器可接收諸如輸入/輸出節點20、30或40之來自下行節點之上行封包交易。另外，處理器10A之主橋接器可發送封包下行至諸如周邊裝置(未顯示)之裝置上，該周邊裝置例如可以連接至周邊匯流排25A。

在運作期間，輸入/輸出節點20及40可轉譯PCI匯流排交易而成為在輸入/輸出資料流中傳輸之上行封包交易，並且另外可轉譯下行封包交易而成為PCI匯流排交易。所有源自於節點處而非處理器10A之主橋接器的封包在受到傳遞至任何其它的節點之前可流通上行至處理器10A之主橋接器。所有源自於處理器10A之主橋接器的封包可流通下行至諸如輸入/輸出節點20、30及40之其它節點。如同在此所使用的，"上行"意指封包傳輸在處理器10A之主橋接器的方向上流通，而"下行"意指封包傳輸在遠離處理器10A之主橋接器的方向上流通。每個輸入/輸出資料流可以

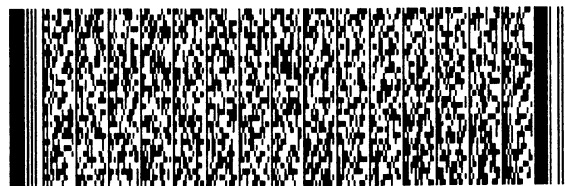
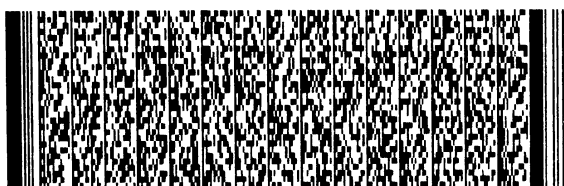


五、發明說明 (7)

藉由稱為單元 ID(Unit ID)之識別器所識別。需要考量的是該單元 ID可以是部分封包標頭或者該單元 ID可以是在單一封包或複數個封包中之某些其它標示數目之位元。如同在此所使用的，"輸入/輸出資料流"意指含有該相同的單元 ID並且因此源自於該相同的節點之所有封包交易。

舉例說明，在周邊匯流排 45 上之周邊裝置初始化導向在周邊匯流排 25 上之周邊裝置之交易。該交易首先可以轉譯成為一個或一個以上之具有獨特的單元 ID 之封包並且接著可以發送上行。需要注意的是每個封包可以編碼成具有能識別該封包之特定的資訊。例如，該單元 ID 可以編碼成為該封包標頭。此外，該交易之形式亦可以編碼成為該封包標頭。每個封包可以指定識別該初始節點之單元 ID。由於輸入/輸出節點 20 可能無法從下行而傳遞封包給在周邊匯流排 25 上之周邊裝置，故該封包發送上行至處理器 10A 之主橋接器。處理器 10A 之主橋接器接著可發送具有處理器 10A 之主橋接器之單元 ID 的該封包而反向下行，直到輸入/輸出節點 20 識得並且取走 (claims) 給予在周邊匯流排 25 上之該周邊裝置的該封包。輸入/輸出節點 20 接著可轉譯該封包成為周邊匯流排交易並且發送該交易給在周邊匯流排 25 上之該周邊裝置。

當封包交易進行上行或下行時，該封包可穿越一個或一個以上之輸入/輸出節點。該穿越有時稱為通道並且該輸入/輸出節點有時稱為通道裝置。從上行至下行或從下行至上行所發送之封包稱為"前遞"傳輸。此外，源自於特



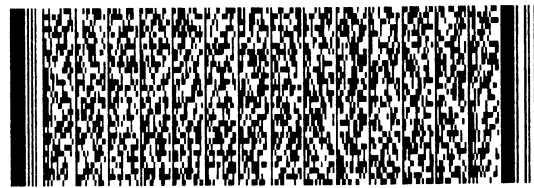
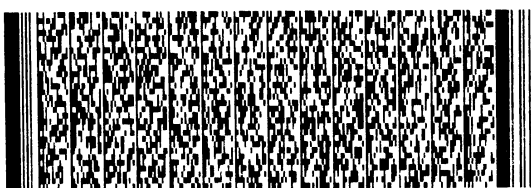
五、發明說明 (8)

定的輸入/輸出節點並且插入至該上行傳輸之封包傳輸稱為"注入"傳輸。

如同將於下文中更加詳細說明的是，為了保留可能連接至輸入/輸出節點之各種匯流排之順序規則，該輸入/輸出節點可提供交易重新排序以及封包緩衝。該輸入/輸出節點藉由前遞及注入傳輸兩者亦可包含控制封包之流通進入該通道及由該通道出來之控制邏輯。

參考第2圖，該圖顯示輸入/輸出節點之一項實施例之方塊圖。該輸入/輸出節點以第1圖之輸入/輸出節點20、30或40表示並且為了簡化起見將於下文中稱為輸入/輸出節點20。第2圖之輸入/輸出節點20包含透過指令匯流排111連接至發送器140及透過指令匯流排112連接至周邊介面150之交易接收器110。輸入/輸出節點20亦包含透過指令匯流排121連接至發送器130及透過指令匯流排122連接至周邊介面150之交易接收器120。周邊介面150亦透過指令匯流排151連接至發送器130及140並且連接至周邊匯流排152。此外，輸入/輸出節點20包含透過控制指令匯流排101連接至每個接收器、每個發送器及周邊介面之交易控制單元100。如同於此所使用的，指令匯流排意指包含用於指令、控制及資料之訊號。因此，當交易或指令宣稱經由個別的指令匯流排而發送時，該交易或指令意指包含指令及資料位元。

在說明的實施例中，接收器110及發送器140形成輸入/輸出通道之其中一個通訊路徑並且接收器120及發送器



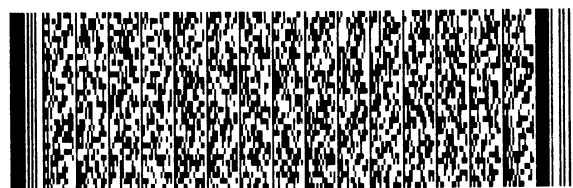
五、發明說明 (9)

130形成該輸入/輸出通道之第二個通訊路徑。由於每個該通訊路徑為單一方向的，故任何一個路徑可以連接成為該上行或下行路徑。因此，來自周邊介面150之該注入傳輸將提供給任何一個傳輸器130及140。

每一個接收器110及120接收封包交易進入接收緩衝區(未顯示)內。當接收到每個交易時，將產生含有內含於該接收的指令中之子集合資料之控制指令。該控制指令可例如包含該來源節點之單元ID、目的資訊、資料計數及交易形式。需要注意的是該控制指令可包含其它資訊或可不包含某些於此所列舉的資訊。該控制指令將發送給交易控制單元100。

回應於周邊介面150接收來自於周邊匯流排152之交易，周邊介面150亦可產生含有類似於上文所描述之該控制指令之資訊的控制指令。周邊介面150亦可儲存該交易於一個或一個以上之緩衝區中並且發送該控制指令給交易控制單元100。

交易控制單元100可儲存每個控制指令，該交易控制單元接收該控制指令以該單元所接收之順序進入一個或一個以上之緩衝區構造內。基於該交易控制單元100儲存該控制指令於本身的緩衝區內，交易控制單元100可決定該對應的指令等待於來源緩衝區(意即接收器及/或周邊介面)中可以發送之順序。如同將於下文中結合第3圖至第6圖以更詳細之描述，交易控制單元100基於諸如交易之形式之因素可以在本身的緩衝區中的交易之間做仲裁，是否

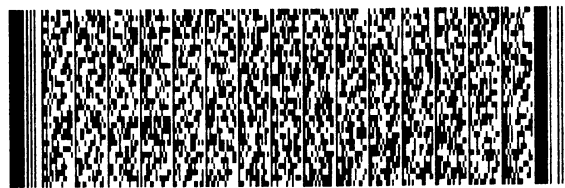


五、發明說明 (10)

在該發送器及目的處具有閒置的緩衝空間，是否該交易為前遞傳輸或注入傳輸。因此該交易控制單元 100 透過輸入 / 輸出節點之通道可以負責於交易之溢位。

一旦交易控制單元 100 仲裁那一個交易將要處理時，交易控制單元 100 可針對該個別的來源裝置以發送待解決的交易給該目的裝置。例如，該交易控制單元 100 選擇來自本身的緩衝區之控制指令，該緩衝區代表由接收器 110 至發送器 140 之待傳遞之交易。交易控制單元 100 通知接收器 110 以透過指令匯流排 111 發送該交易給發送器 140。發送器 140 接著可發送該交易給在該串連中之下一個節點。該下一個節點可以是另一個不管是上行或下行之輸入 / 輸出節點，或者該下一個節點可以是諸如第 1 圖之主處理器 10A 之主節點。此外，交易控制單元 100 及發送器 140 可包含指示另一個節點是否於該接收緩衝區中未具有閒置空間之邏輯（未顯示）。

茲參照第 3 圖，該圖顯示交易控制單元之一項實施例之方塊圖。為了簡化及簡潔，顯示於第 2 圖中之對應於那些組件之電路組件將以相同圖式標號做標示。交易控制單元 100 包含三個分別標示為 160、170 及 180 之排程器。每一個排程器 160、170 及 180 包含一對虛擬通道指令緩衝區及仲裁器與緩衝區處理單元。排程器 160 之虛擬通道指令緩衝區標示為 V.C. FIFO 166 及 167 並且仲裁器與緩衝區處理單元標號為 168。同樣地，排程器 170 之虛擬通道指令緩衝區標示為 V.C. FIFO 176 及 177 並且仲裁器與緩衝區處理單

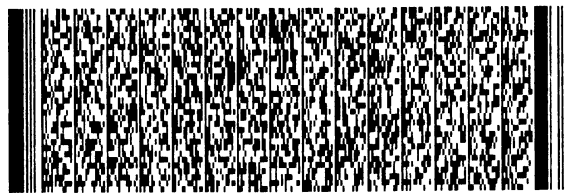
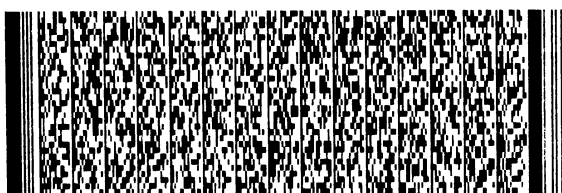


五、發明說明 (11)

元標號為 178，以及排程器 180 之虛擬通道指令緩衝區標示為 V.C. FIFO 186 及 187 並且仲裁器與緩衝區處理單元標號為 188。

一般來說，"虛擬通道"為用於運送封包於各種處理節點之間之通訊路徑。每個虛擬通道為資源獨立於該其它的虛擬通道（亦即，藉由實體的傳輸，在其中一個虛擬通道流通之封包通常並不會受到封包於另一個虛擬通道中之存在或缺乏而影響）。封包基於封包形式而指定給虛擬通道。在該相同的虛擬通道中之封包實體上可能與彼此的傳輸衝突（亦即，在該相同的虛擬通道中之封包可能經歷資源衝突），但是實體上可能不會與在不同的虛擬通道中的封包之傳輸衝突。

某些封包於邏輯上可能與其它封包衝突（亦即，對於通訊協定之原因、相干性之原因或其它此類的原因，其中一個封包可能於邏輯上與另一個封包衝突）。對於邏輯/通訊協定之原因，若第一封包必須在第二封包抵達該封包的目的地節點之前抵達本身的目的節點，而若該第二封包實體上阻滯該第一封包的傳送（藉由佔據衝突資源），則電腦系統可能停滯。藉由指定該第一及第二封包以分離虛擬通道，以及藉由在該電腦系統內實現該傳送媒介使得在個別的虛擬通道中之封包無法阻滯每個其它的傳送，停滯免除之操作則可以達成。需要注意的是來自不同的虛擬通道之封包將透過該相同的實體連結而傳送。然而，由於接收緩衝區於發送之前是可以獲得的，即使當使用此種共用資源



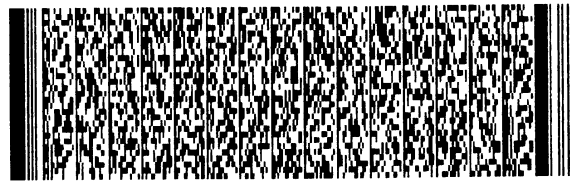
五、發明說明 (12)

時，該虛擬通道並未彼此阻滯。

從其中一個觀點，每個不同的封包形式（例如每個不同的指令編碼）可以指定給屬於本身的虛擬通道並且因此在一項實施例中，個別的緩衝器分派給每個虛擬通道。由於個別的緩衝器可以使用於每個虛擬通道，來自其中一個虛擬通道之封包可能不會與來自另一個虛擬通道之封包形成實體衝突（由於此類封包將放置於另一個緩衝區中）。

每個排程器對應於特定的目的及對應於兩個來源。在該說明的實施例中，排程器 160 控制具有第 2 圖之發送器 130 作為目的與接收器 120 及周邊介面 / 橋接器 150 作為來源之交易。同樣地，第 3 圖之排程器 170 控制具有第 2 圖之發送器 140 作為目的與接收器 110 及橋接器 150 作為來源之交易。最後，第 3 圖之排程器 180 控制具有第 2 圖之橋接器 150 作為目的與接收器 110 及接收器 120 作為來源之交易。在第 3 圖中，每個虛擬通道指令緩衝區接收來自個別的接收器或對應於由該個別的接收器或橋接器所接收之交易的橋接器控制指令。控制指令可包含指定該控制指令將會發送至那一個排程器之目的位元。通常控制指令將僅具有一個目的位元設定。然而，在交易為廣播訊息之事件中，超過一個的目的位元將會設定以允許超過一個的排程器以接收該控制指令。

為了簡化討論，將只詳細說明排程器 160。當控制指令接收進入 V.C. FIFO 166 或 167 時，該指令視交易之形式而放置於個別的 FIFO 區段內。由於 V.C. FIFO 166 及 167 是



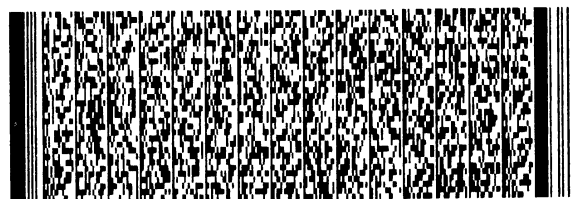
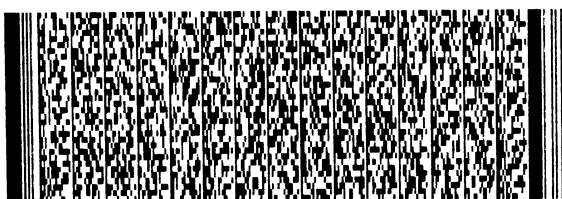
五、發明說明 (13)

相同的，只有 V.C. FIFO166將詳細說明。V.C. FIFO166含有三個個別的 FIFO區段，該區段對應於三個交易形式：宣告 (posted)、非宣告 (non-posted)、及回應 (response)。該控制指令依照所接收之順序放置於該個別的 FIFO內。然而，為了維持結合可能已經產生該初始的指令之各種匯流排或元件形式之排序規則，該交易可能必須在不同的交易形式之間脫序處理。

將於下文結合第 4圖至第 6圖以更詳細之描述，仲裁器及緩衝器處理邏輯 168可以經由配置以仲裁那一個交易可以在 V.C. FIFO166或 V.C. FIFO167之間依照第一、第二順序而處理。例如，在回應指令之前到達於 V.C. FIFO166之宣告的指令，由於該順序規則而可能必須在該回應指令之後處理。此外，仲裁器及緩衝器處理邏輯 168亦可以仲裁那一個 V.C. FIFO，交易可以基於一組公正性規則及是否在該下一個輸入/輸出節點或該主橋接器之接收緩衝區內之空間是可以獲得的而做處理。若該目的為第 2圖之橋接器 150，則該特定的仲裁規則可能不同於上文所描述之仲裁規則。

需要注意的是，上文所描述之說明的實施例為交易控制單元之特定的實現。經由考量之另一種實施例可包含可執行額外的功能之不同的實現。

參考第 4圖，該圖顯示排程器之一項實施例之方塊圖。為了簡化及明確起見，對應於第 3圖中所顯示之電路組件將標示相同的圖式標號。交易排程器 400包含連接至

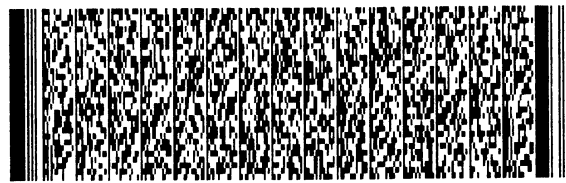
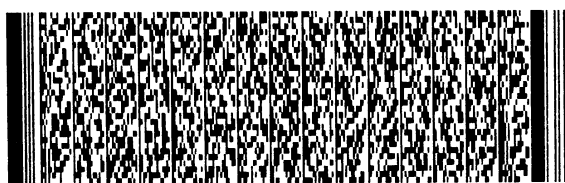


五、發明說明 (14)

仲裁及公平性邏輯 450之虛擬通道 FIFO緩衝區 410。交易排程器 400包含亦連接至仲裁及公平性邏輯 450之虛擬通道 FIFO緩衝器 420。仲裁及公平性邏輯 450連接至 FIFO緩衝區 460，該 FIFO緩衝區連接至緩衝處理邏輯 470。緩衝處理邏輯之輸出由輸出暫存器 480所門鎖。

如同上文結合第 3圖之描述，每個虛擬通道 FIFO緩衝區 410及 420可接收來自個別的來源輸入之控制指令，該來源輸入諸如例如第 2圖之接收器 110或橋接器 150。藉由該控制指令所表示之交易形式，該控制指令可以放置於虛擬通道 FIFO緩衝區 410及 420內。尤其，該控制指令可表示不論是宣告、非宣告或回應指令並且因此表示該個別的宣告、非宣告或回應佇列。

在說明的實施例中，仲裁及公平性邏輯 450包含仲裁單元 430與 440以及公平性單元 445。仲裁單元 430可以經由配置以選擇其中一個儲存於虛擬通道 FIFO緩衝區 410內之控制指令。將於下文中更進一步說明的是，該選擇程序可包含藉由預定的仲裁演算法選擇其中一個入選者。同樣地，藉由如同仲裁單元 430所使用類似的演算法，仲裁單元 440可以經由配置以選擇其中一個儲存於虛擬通道 FIFO緩衝區 420內之控制指令。公平性單元 445接著可以選擇其中一個由仲裁單元 430及 440所選擇之該入選交易。公平性單元 445基於是否該交易為前遞或注入，而可使用公平性演算法。視該排程器之目的，仲裁單元 430及 440亦可含有在該對應之交易的目的緩衝區中保持追蹤緩衝區空間之邏



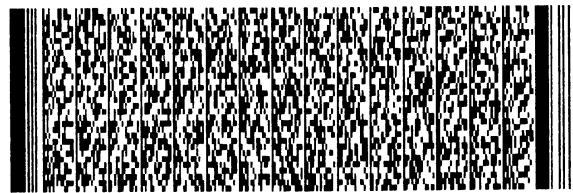
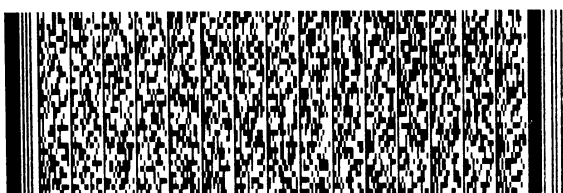
五、發明說明 (15)

輯(未顯示)，該目的緩衝區諸如該下一個輸入/輸出節點之接收緩衝區或於第1圖之處理器10A之主橋接器中。

在說明的實例中，可能具有三個時脈周期之等待時間以透過排程器400傳送控制指令。旁路415及旁路425顯示以分別繞過虛擬通道FIFO緩衝區410及420。當排程器400接收來自來源之控制指令並且在給定的虛擬通道FIFO緩衝區內之每個佇列為空的時，則可能有機會藉由繞過該虛擬通道FIFO緩衝區以節省時脈周期。例如，非宣告控制指令在虛擬通道FIFO緩衝區410處所接收，該緩衝區目前為空的。若仲裁單元430顯示在該交易的目的緩衝區具有可獲得的緩衝空間，則在仲裁單元430內之邏輯可允許該非宣告控制指令以繞過虛擬通道FIFO緩衝區410並且即時放置於FIFO緩衝區460內。此外，公平性單元445視如同上所描述之公平性演算法而定，可允該繞道。因此在該上文的例子中，其中一個時脈周期之減少等待時間將可以實現。需要注意的是雖然該說明的實施例顯示具有三個時脈周期的等待時間，所考量的是其它實施例可具有較少的或較多的時脈周期等待時間。此外，由旁路415及425所實現之該實際的等待時間減少可能是更多或更少。

第4圖之FIFO緩衝區460可接收該入選的控制指令。在該說明的實施例中，FIFO緩衝區460為雙延伸(two-deep)緩衝區，然而，所考量的是在其它實施例中，FIFO緩衝區460可以具有更多或更少的緩衝區位置。

緩衝區處理邏輯470經由配置以保持在第2圖之發送器



五、發明說明 (16)

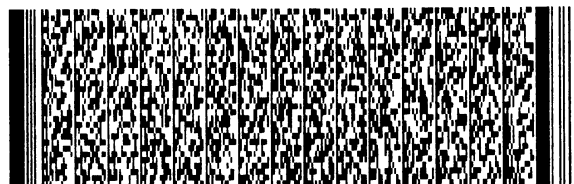
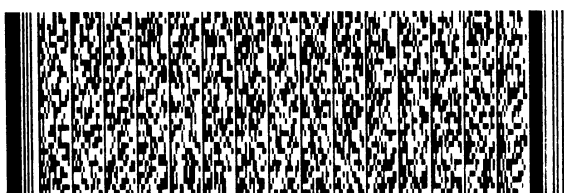
130或140或者在橋接器150中之緩衝區空間的追蹤。當交易儲存於FIFO緩衝區460中時，緩衝處理邏輯470檢查該下一個緩衝區之可取得性並且也將保留該控制指令直到緩衝區空間可以獲得或允許該控制指令進行至輸出暫存器

480。一旦該控制指令已經由輸出暫存器480門鎖，該個別的交易來源受到通知，對應於該控制指令之交易可以分別發送到第2圖之發送器130或140或者發送至橋接器150。

茲參照第5圖，該圖顯示包含標記邏輯排程器之一項實施例之方塊圖。排程器500包含連接至虛擬通道FIFO緩衝區505之標記邏輯510。虛擬通道FIFO緩衝區505包含對應於該三種形式的交易之三個個別的佇列，該三種形式的交易為：宣告、非宣告及回應。標記比較/仲裁邏輯單元520連接至虛擬通道FIFO緩衝區505。此外，該圖亦顯示虛擬通道FIFO緩衝區505之分解圖。該分解圖顯示每個該非宣告及回應佇列具有對應的標記。然而，該宣告佇列具有兩個標記：每一個分別對應於該非宣告及回應佇列。

標記邏輯510在控制指令儲存於虛擬通道FIFO緩衝區505之前，可指定標記給每個控制指令。控制指令可以以來自諸如第2圖之接收器110或橋接器150之來源單元所接收之順序而指定標記。該標記可以附加於該控制指令。

當控制指令到達在虛擬通道FIFO緩衝區505內之上方位置時，標記比較/仲裁邏輯單元520可以經由配置以在該三個虛擬通道之間做仲裁以選取入選的控制指令。入選者使用基於一組順序規則之演算法來挑選，該順序規則可符

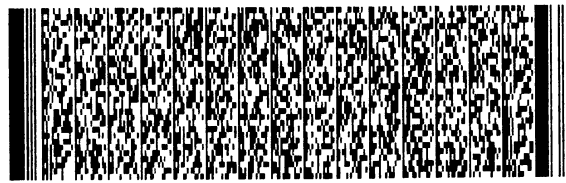
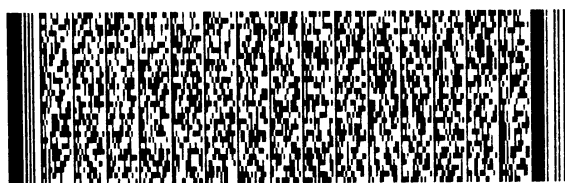


五、發明說明 (17)

合由連接至該輸入/輸出節點之周邊匯流排所維持的順序規則。在一項實施例中該順序規則可符合該PCI順序規則。在另一個實施例中，該順序規則可符合AGP序順序規則。

在說明的實施例中，由於虛擬通道FIFO緩衝區505可以上達16個位置延伸，故該標記可以是四個位元。然而，需要注意的是在其它實施例中，虛擬通道FIFO緩衝區505可含有其它數目的位置並且因此該標記可含有對應數目的位元。標記邏輯510可包含計數器邏輯(未顯示)，該計數器邏輯保持已經儲存於虛擬通道FIFO緩衝區505之用於該非宣告及回應控制指令之該目前的標記之追蹤。若至少其中一個可阻滯的非宣告或回應控制指令是於該目前的宣告控制指令之前並在上一個宣告控制指令之後接收，則當接收到宣告控制指令時，該標記可以依據增加該個別的非宣告或回應計數器之演算法而指定。如同在此所使用的，若特殊的位元設定於該相對的控制指令而確認該非宣告或回應控制指令在虛擬通道FIFO緩衝區505內將通過宣告控制指令，則非宣告或回應控制指令為可阻滯的。在一項實施例中，該特殊的位元稱為PassPW位元。

為了說明標記演算法之使用，第1表說明控制指令進入該三個宣告、非宣告及回應佇列之接收順序及標記。該第一欄顯示該九個控制指令之接收順序。該第二欄顯示指令所接收之形式。該第三欄顯示指定給該非宣告及回應指令之標記，並且該第四及第五指令顯示在該等計數器值由



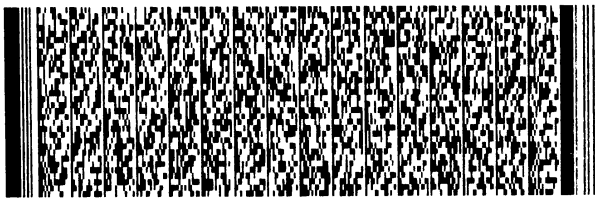
五、發明說明 (18)

宣告控制指令所遞增之後之該非宣告及回應指令計數器數值。由於該宣告控制指令接收兩個標記，對於該個別的非宣告及回應計數器，該宣告控制指令接收顯示於該目前的計數器數值內之標記兩者。虛擬通道 FIFO 緩衝器 505 之分解圖說明第 1 表之該指令如何儲存。

第 1 表

接收順序	控制指令	Pass PW 位元	標記值	非宣告計數器	回應計數器
1	宣告 1	0		0	0
2	回應 1	0	0	0	0
3	宣告 2	0		0	1
4	非宣告 1	0	0	0	1
5	回應 2	0	1	0	1
6	非宣告 2	0	0	0	1
7	宣告 3	0		1	2
8	回應 3	0	2	1	2
9	非宣告 3	0	1	1	2

第 2 表說明 PassPW 位元如何可影響顯示於第 1 表中之該接收的控制指令之標記。若該 PassPW 位元設定為非宣告或回應控制指令，則下一個宣告控制指令將不會增加該相對的計數器。例如，在第 2 表中，該 PassPW 位元設定為回應 1，因此當接收到宣告 2 控制指令時，該回應計數器並未遞增。然而，該 PassPW 位元設定為非宣告 2，並且該宣告 3 控制指令造成該非宣告及回應計數器兩者遞增。這是因為該非宣告 1 控制指令在該上一個宣告控制指令之後及在該目前的宣告控制指令之前接收到具有該 PassPW 位元清除，因



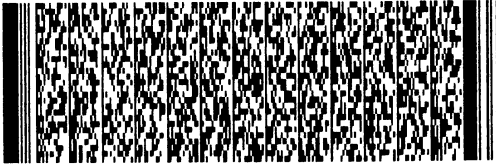
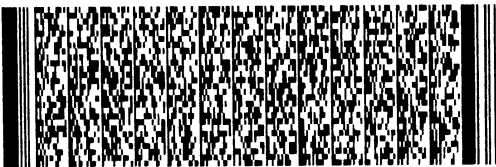
五、發明說明 (19)

此滿足用於遞增該計數器之標記規則。需要注意的是雖然該 PassPW位元描述設定為顯示非宣告或回應控制指令可傳遞宣告控制指令，需考量的是在其它實施例中該邏輯可以反置。

第 2表

接收順序	控制指令	Pass PW 位元	標記值	非宣告計數器	回應計數器
1	宣告 1	0		0	0
2	回應 1	1	0	0	0
3	宣告 2	0		0	0
4	非宣告 1	0	0	0	0
5	回應 2	0	0	0	0
6	非宣告 2	1	0	0	0
7	宣告 3	0		1	1
8	回應 3	0	1	1	1
9	非宣告 3	0	1	1	1

回頭參考第 5圖，在仲裁期間，標記比較及仲裁邏輯單元 520經由配置而藉由比較附屬於每個控制指令之標記以選取來自虛擬通道 FIFO緩衝區 505之入選控制指令。此外，在該標記比較之前，標記比較及仲裁邏輯單元 520可確認是否在下一個輸入/輸出節點處對於每個該虛擬通道具有可獲得的緩衝區空間。若任何虛擬的通道為阻滯的，則該通道對於該周期並未參與該仲裁。若所有三個通道對於仲裁為開啟時，則該仲裁在該非宣告及回應通道之間以循環 (round robin)方式發生並且該宣告通道將個別與該非宣告及回應通道做比較。最後，標記比較及仲裁邏輯單



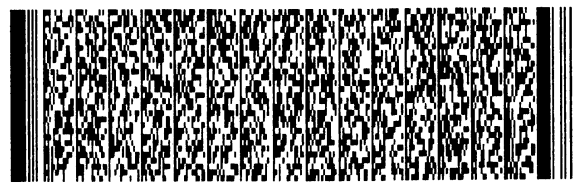
五、發明說明 (20)

元 520 可使用公平性演算法來決定入選者。該公平性演算法結合第 6 圖之描述將於下文中做更詳細的描述。

茲參考第 6 圖，該圖顯示包含匱乏避免邏輯之交易排程器之一項實施例之方塊圖。交易排程器 600 包含連接至仲裁電路 650 之虛擬通道 FIFO 緩衝區 610。交易排程器 600 更包含亦連接至仲裁電路 650 之虛擬通道 FIFO 緩衝區 620。仲裁電路 650 連接至 FIFO 緩衝區 670，該 FIFO 緩衝區連接至緩衝處理邏輯 680。緩衝處理邏輯之輸出由輸出暫存器 690 所門鎖。

類似於上文之第 3 圖及第 4 圖之說明，每個虛擬通道 FIFO 緩衝區 610 及 620 可接收來自個別的來源輸入之控制指令，諸如例如第 2 圖之接收器 110 或橋接器 150。該控制指令依據該控制指令所表示之交易之形式可以儲存於虛擬通道 FIFO 緩衝區 610 及 620 內。例如，該控制指令可表示宣告、非宣告或回應指令並且因此可以儲存於該個別的宣告、非宣告或回應佇列內。

在說明的實施例中，仲裁電路 650 包含仲裁單元 630 及 640，以及公平性電路 645。在仲裁周期期間，仲裁單元 630 及 640 可以經由配置以選擇分別儲存在虛擬通道 FIFO 緩衝區 610 及 620 內之控制指令。此外，公平性電路 645 將提供可決定那一個仲裁單元 630 或 640 將選擇該入選者交易之選擇條件。如同結合第 7 圖之說明而將於下文中更詳細之描述，公平性電路 645 可使用公平性演算法以建立仲裁優先權以分配頻寬。該公平性演算法將視匱乏避免邏輯而定



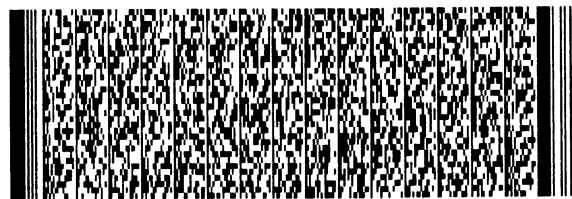
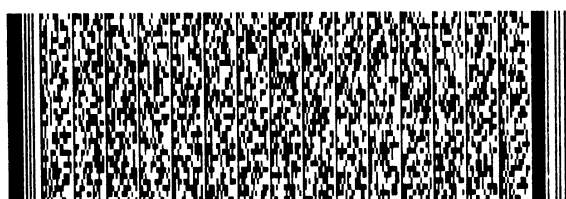
五、發明說明 (21)

以決定何時任何交易已經阻滯持續預定數目的仲裁周期。

茲參照第 7 圖，該圖顯示第 6 圖之公平性電路之一項實施例之方塊圖。公平性電路 645 包含含有連接至 8 位元計數器 705 之三十二個三位元計數器 0 至 31 之公平性單元 700。門鎖 710 連接計數器 705。插入速率邏輯 715 連接至門鎖 710。公平性電路 645 亦包含連接至匱乏門限暫存器 760 之三個虛擬通道計數器 755 至 757 的匱乏單元 750 及連接至公平性單元 700 及匱乏單元 750 之交易選擇單元 775。

在一項實施例中，每個交易透過該通道向前傳遞的時刻，發送該交易之對應於該輸入/輸出節點之該三位元計數器 0 至 31 可以遞增。此外，計數器 705 在每個任何的該三位元計數器 0 至 31 遞增的時刻將會增加。當其中一個該三位元計數器 0 至 31 溢位時，在計數器 705 中之該數值可以由門鎖 710 所擷取。該擷取的值可表示於時間內在該特定點處該下行節點交易請求速率。插入速率邏輯 715 接著使用該擷取的值可計算對於該節點之允許的插入速率。

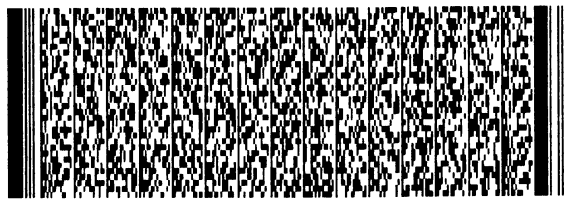
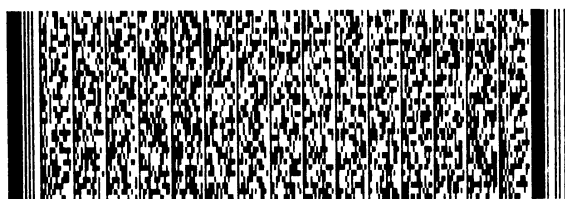
交易選擇單元 775 在第 6 圖之仲裁單元 630 或 640 內可造成指標以指向在仲裁循環期間所考量的虛擬通道緩衝區。當具有交易於前遞的虛擬通道緩衝區之虛擬通道中並且具有交易於插入緩衝區之相同的虛擬通道中可能需要時間。在第 7 圖中，交易選擇單元 775 依據由公平性單元 700 所建立之優先權可以另外選擇該兩個虛擬通道緩衝區。然而在虛擬通道在交易的目的處由於缺乏可獲的緩衝區空間之此類因素而變成阻滯之事件中，仲裁邏輯可略過該阻滯通道



五、發明說明 (22)

並且當仲裁繼續時進行至下一個虛擬通道。當該阻滯通道變成可獲得時，可能僅對於該前遞通道具有可獲得的空間並且由於該通道依據公平性演算法而具有高於該插入通道之優先權，來自該前遞通道之交易將可以發送。接著當下一個該先前阻滯通道變成可獲得時，該通道在該目前的仲裁周期中可能再次不具有優先權並且因此可再次略過。此種情況可能持續並且因此"缺乏(starve)"該插入虛擬通道。視特定的節點之配置而定，所考量的是在其它實施例中，該前遞通道亦可能藉由注入通道以類似的事件而變成缺乏。

為了避免給定通道之缺乏，匱乏單元 750 可保持交易為阻滯之次數的數目之追蹤。每次交易選擇單元 775 決定那一個交易有資格仲裁但是為阻滯，交易選擇單元 775 造成該對應的虛擬通道計數器 755 至 757 將遞增。匱乏臨界暫存器 760 保持對應於通道在仲裁期間可能略過之次數的最大數目之值。儲存於匱乏臨界暫存器 760 之值在任何時刻視該計算的請求速率而可以動態改變。當任何該虛擬通道計數器 755 至 757 符合在匱乏臨界暫存器中之數值時，該對應的虛擬通道之優先權將可以改變以便交易選擇單元 775 可以造成該阻滯的交易在後續的仲裁周期期間受到選擇。因此，藉由動態改變該優先權，特定的通道之匱乏可以避免。需要注意的是所描述之該邏輯處理僅為其中一個特定的實現。所考量的是另外的實施例可以使用包含不同數目計數器及在那些計數器中之較多或較少的位元之另一個邏



五、發明說明 (23)

輯處理以達到上文所描述的功能。

對於熟習此項技藝之人士一旦完全瞭解上文之揭露後，各種改變及修正將變得顯而易見。下列之申請專利範圍應解釋為意在含括所有此類的改變及修正。



圖式簡單說明

[圖式簡單說明]

第 1圖為電腦系統之一項實施例之方塊圖。

第 2圖為輸入/輸出節點之一項實施例之方塊圖。

第 3圖為交易控制單元之一項實施例之方塊圖。

第 4圖為排程器之一項實施例之方塊圖。

第 5圖為包含標記邏輯之排程器之一項實施例之方塊圖。

第 6圖為包含匱乏避免邏輯之交易排程器之一項實施例之方塊圖。

第 7圖為公平性電路之一項實施例之方塊圖。

雖然本發明容許於各種的修正及替代形式，但本發明之特定的實施例藉由在圖式中之例子而呈現並且將於此詳細描述。然而，應該要瞭解的是對於該例子之圖式及詳細描述並非意在限定本發明於所揭露之該特定的形式，而相反地，本發明是在於含括落在由附加的申請專利範圍所定義之本發明之精神及範疇內的所有修正、等同及替代。

10A至 10D 處理器

15 相干封包匯流排

20、30、40 輸入/輸出節點

25A至 25B、45、50A至 50C、152 周邊匯流排

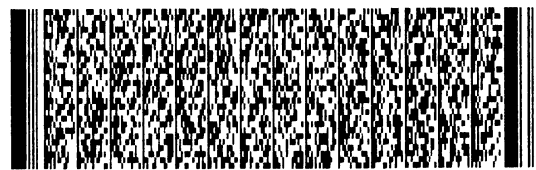
35 繪圖匯流排

100 資訊交易控制單元

101 控制指令匯流排

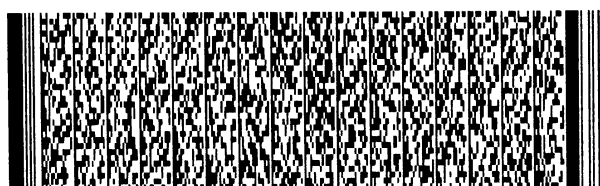
110 接收器

111、112、121、122、151 指令匯流排



圖式簡單說明

120 交易接收器	130、140 發送器
150 周邊介面	160、170、180 排程器
166、167、176、177、186、187 V.C. FIFO	
168 仲裁器及緩衝區處理邏輯	
178、188 仲裁器與緩衝區處理單元	
400、600 交易排程器	
410、420、505、610、620 虛擬通道 FIFO緩衝區	
415、425 旁路	
430、440、630、640 仲裁單元	
445、700 公平性單元	450 仲裁及公平性邏輯
460、670 FIFO緩衝區	470、680 緩衝處理邏輯
480、690 輸出暫存器	510 標記邏輯
520 標記比較 /仲裁邏輯單元	
645 公平性電路	650 仲裁電路
705 計數器	710 門鎖
715 插入速率邏輯	750 匱乏單元
755至 757 虛擬通道計數器	760 匱乏門限暫存器
775 交易選擇單元	



四、中文發明摘要 (發明名稱：電腦系統之輸入/輸出節點)

本發明提供一種電腦系統輸入/輸出節點。用於電腦系統之輸入/輸出節點包含經由配置以接收在第一通訊路徑上之第一指令的第一接收器單元及經由連接以發送第一對應的指令之第一發送器單元，該第一對應的指令對應於在第二通訊路徑上之第一指令。該輸入/輸出節點亦包含經由配置以接收在第三通訊路徑上之第二指令的第二接收器單元及經由連接以發送第二對應的指令之第二發送器單元，該第二對應的指令符合在第四通訊路徑上之該第二指令。再者，該輸入/輸出節點包含經由連接以接收來自該第一接收器及該第二接收器之選擇的指令及經由配置以發送對應於該選擇的指令之指令於周邊匯流排上的橋接器單元。

本案代表圖：第 1圖

10A至 10D 處理器

15 相干封包匯流排

陸、英文發明摘要 (發明名稱：COMPUTER SYSTEM I/O NODE)

A computer system I/O node. An input/output node for a computer system includes a first receiver unit configured to receive a first command on a first communication path and a first transmitter unit coupled to transmit a first corresponding command that corresponds to the first command on a second communication path. The input/output node also includes a second receiver



四、中文發明摘要 (發明名稱：電腦系統之輸入/輸出節點)

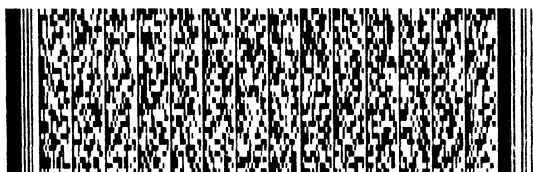
20、30、40 輸入/輸出節點

25A至25B、45、50A至50C 周邊匯流排

35 繪圖匯流排

陸、英文發明摘要 (發明名稱：COMPUTER SYSTEM I/O NODE)

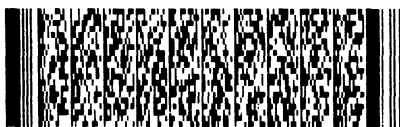
unit configured to receive a second command on a third communication path and a second transmitter unit coupled to transmit a second corresponding command that corresponds to the second command on a fourth communication path. Further, the input/output node includes a bridge unit coupled to receive selected commands from the first receiver and the second receiver and configured to

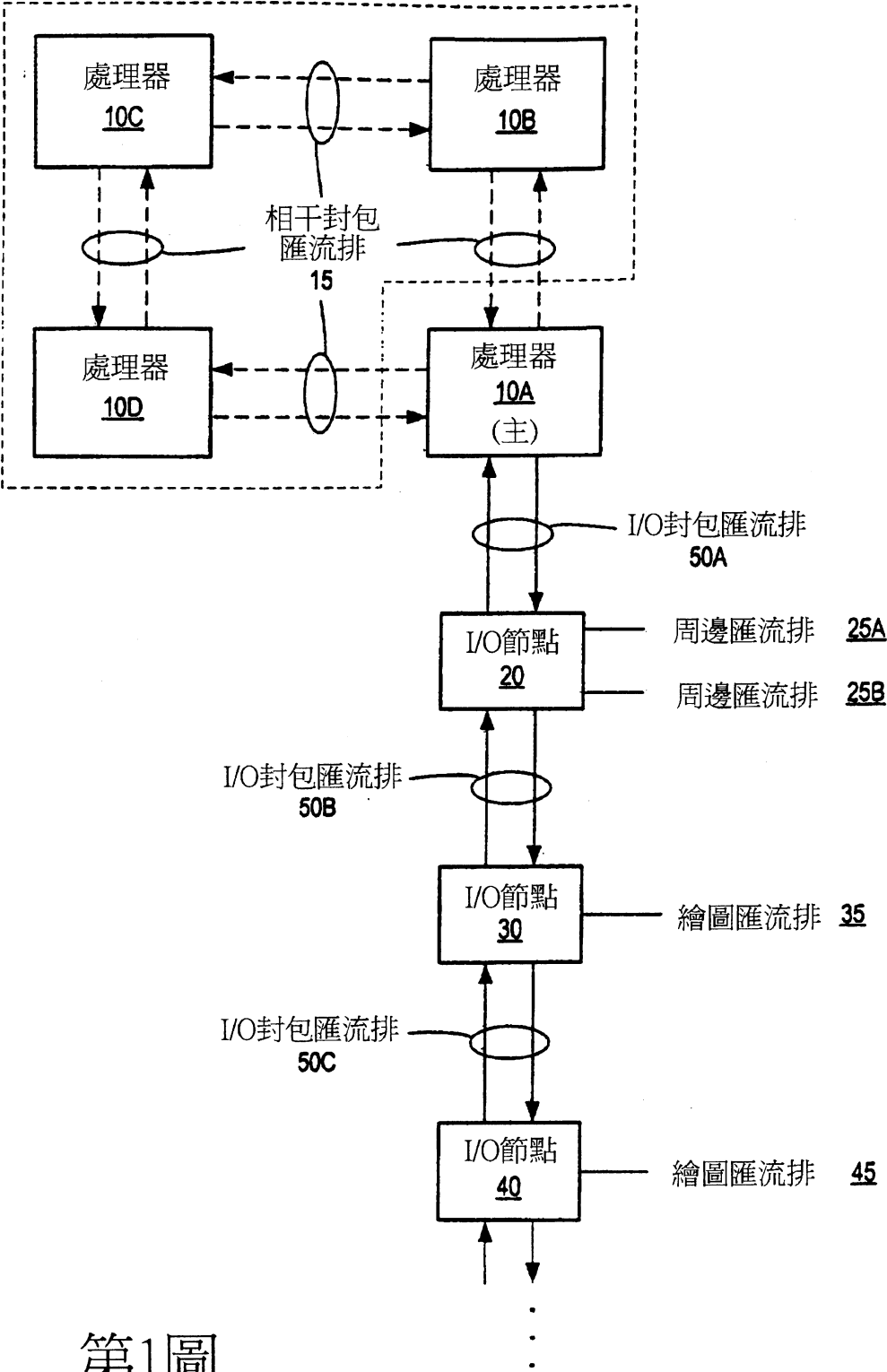


四、中文發明摘要 (發明名稱：電腦系統之輸入/輸出節點)

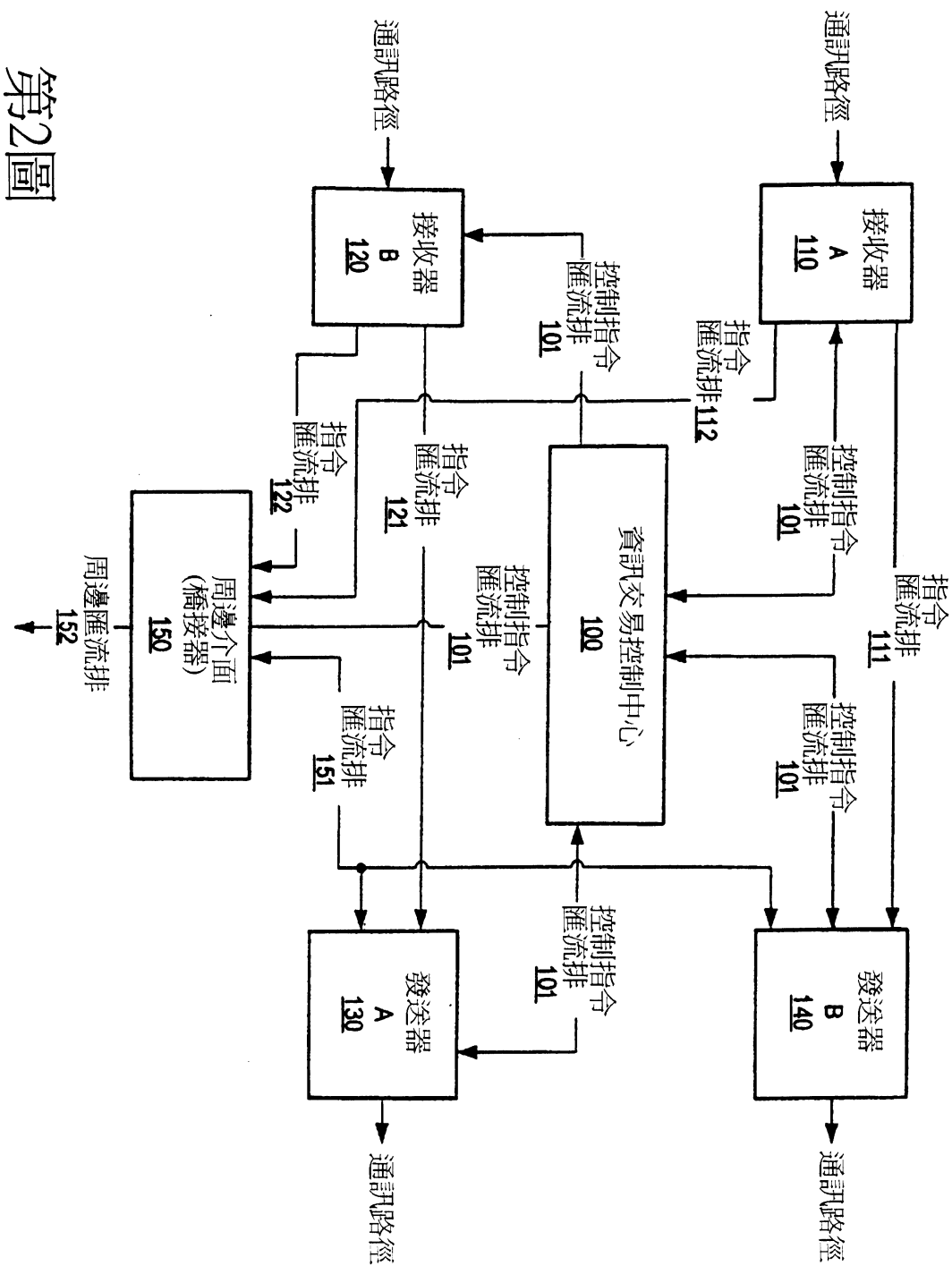
陸、英文發明摘要 (發明名稱：COMPUTER SYSTEM I/O NODE)

transmit commands corresponding to the selected
commands upon a peripheral bus.

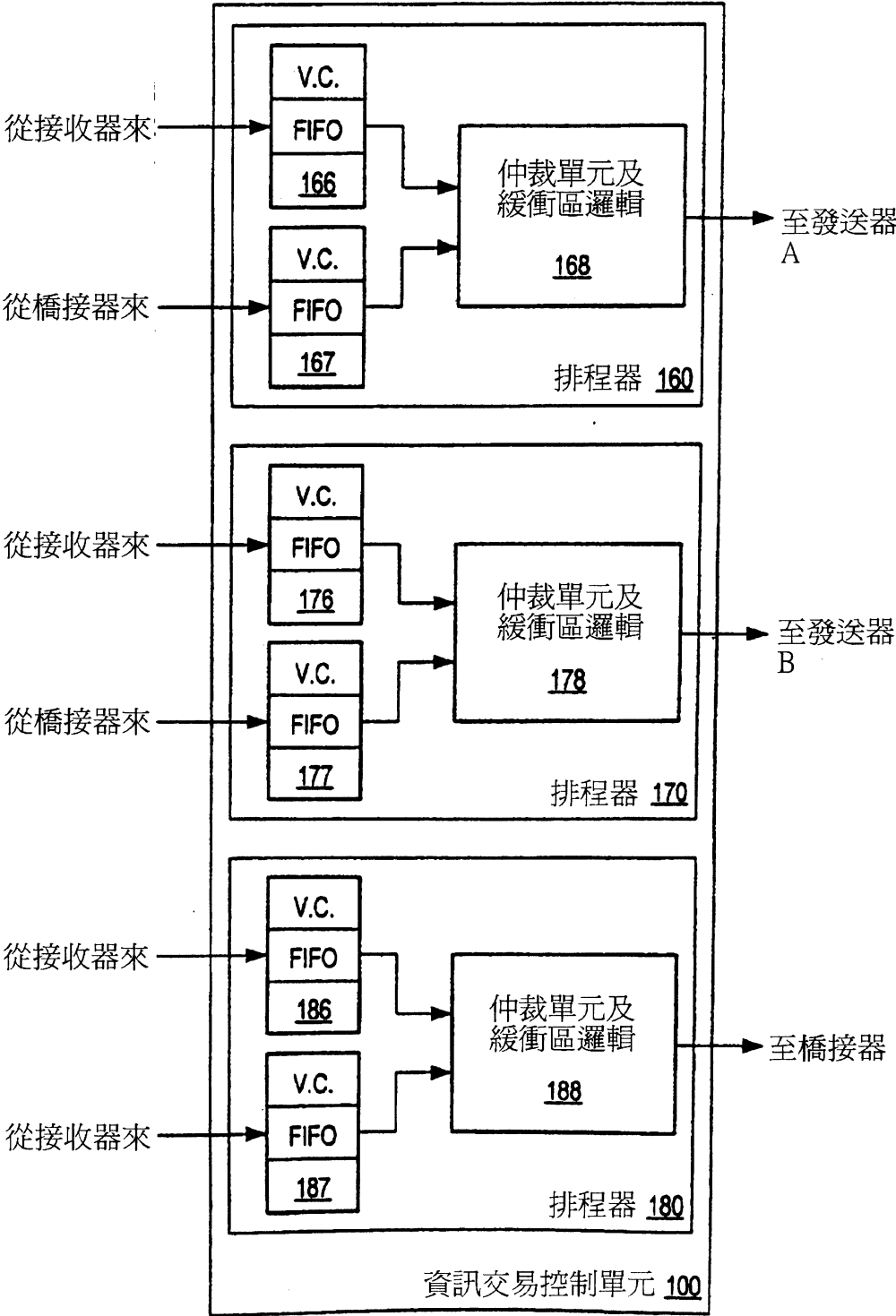




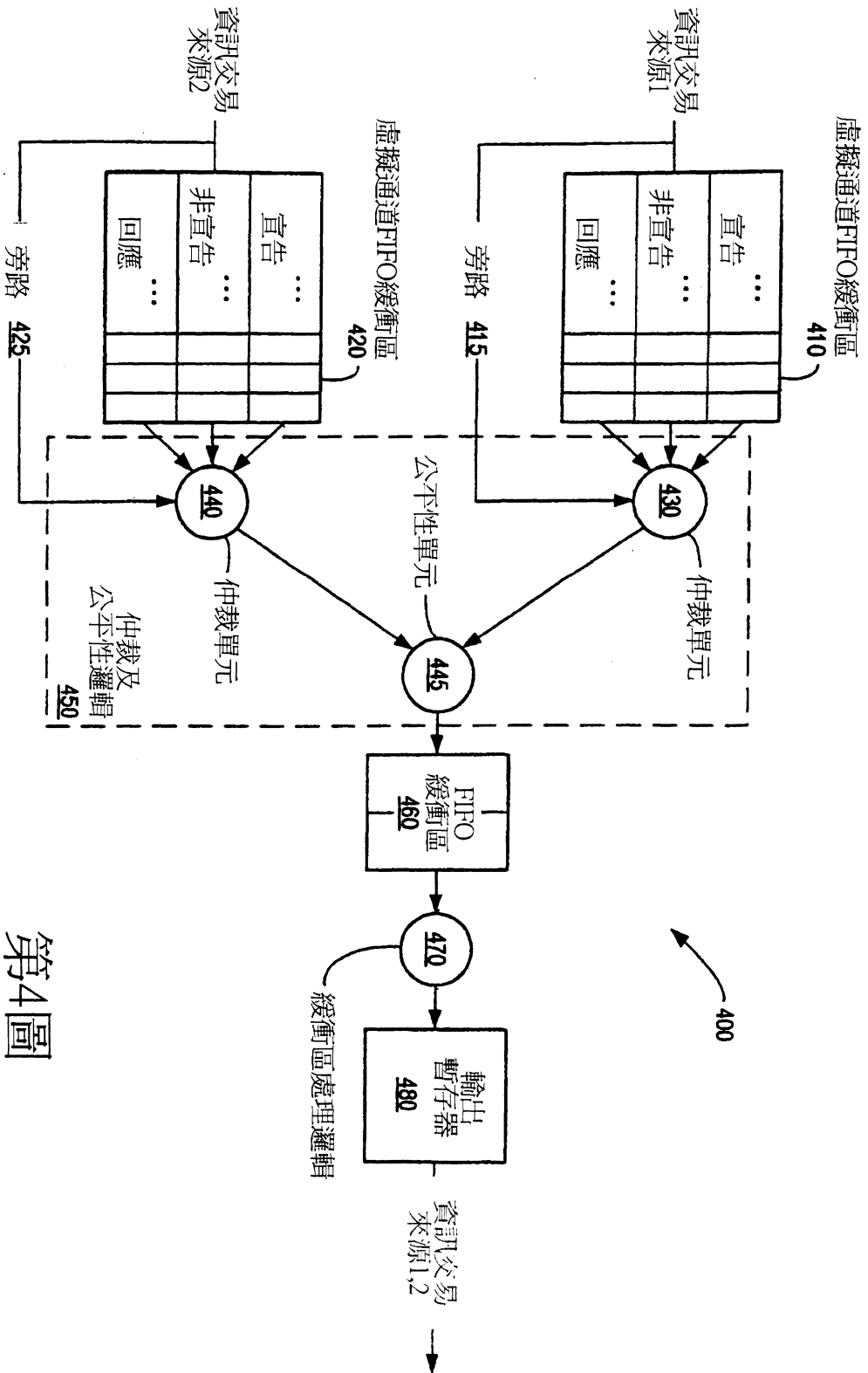
第1圖



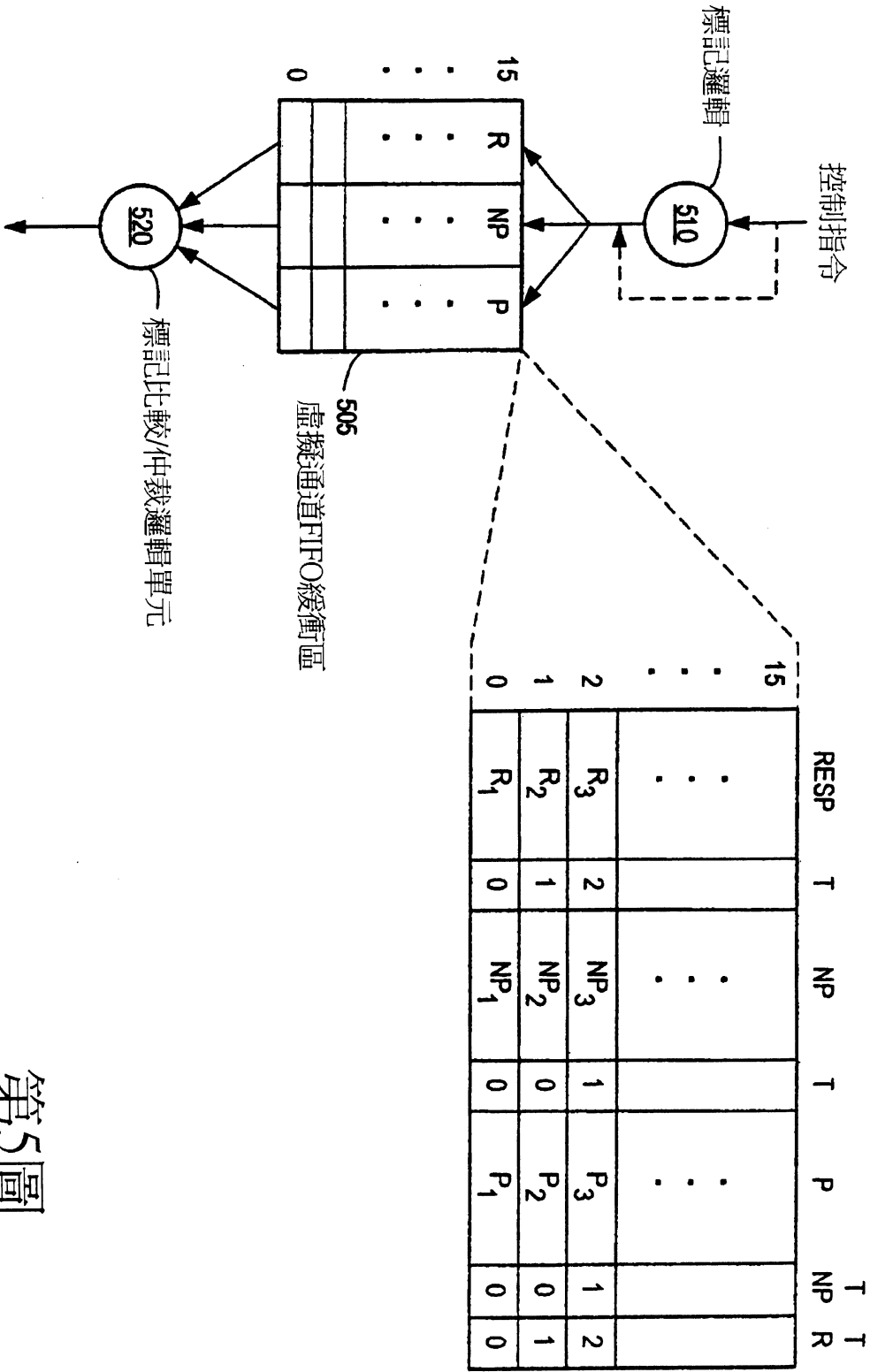
第2圖



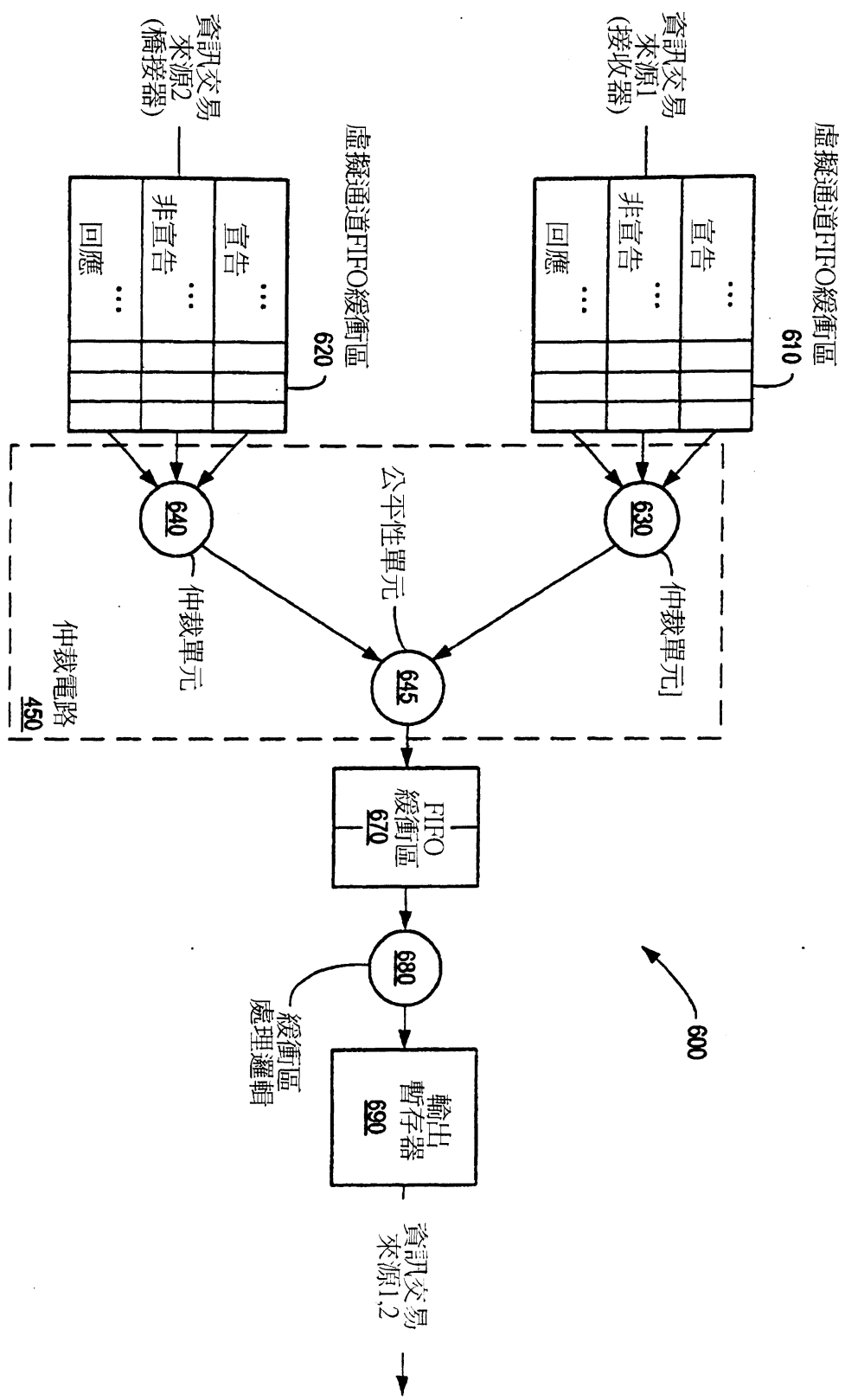
第3圖



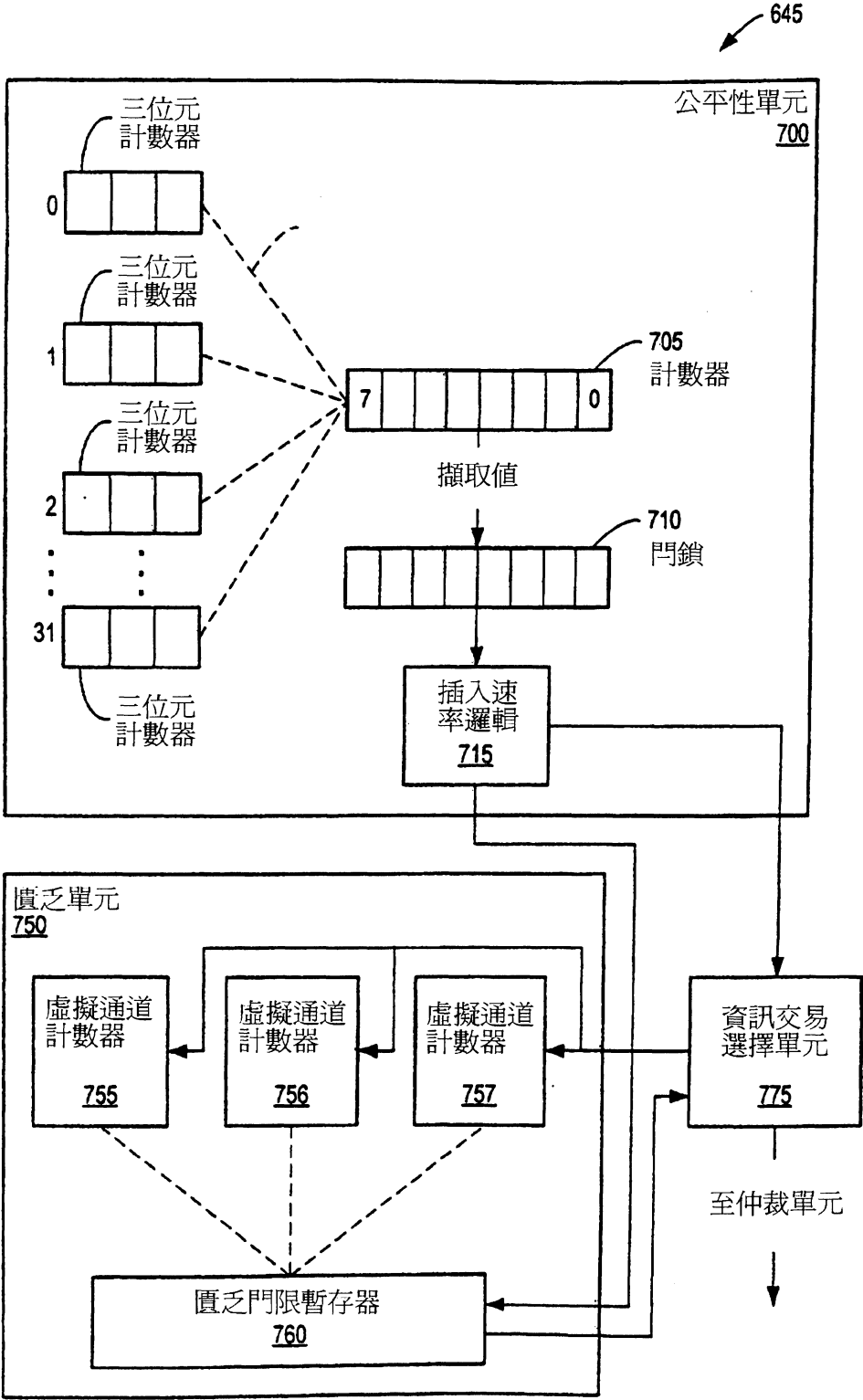
第4圖



第5圖



第6圖



第7圖

六、申請專利範圍

1. 一種用於電腦系統之輸入/輸出節點，包括：

經由配置以接收在第一通訊路徑上之第一指令的第一接收器單元；

經由連接以發送第一對應的指令之第一發送器單元，該指令對應於在第二通訊路徑上之該第一指令；

經由配置以接收在第三通訊路徑上之第二指令的第二接收器單元；

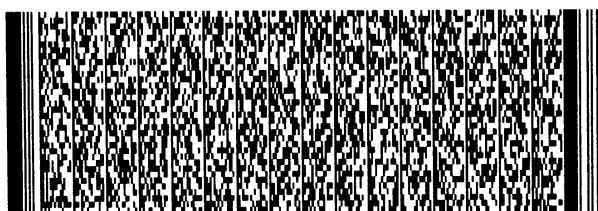
經由連接以發送第二對應的指令之第二發送器單元，該指令對應於在第四通訊路徑上之該第二指令；以及

經由連接以接收來自該第一接收器與該第二接收器之選擇的指令及經由配置以發送對應於該選擇的指令之指令於周邊匯流排上的橋接器單元。

2. 如申請專利範圍第1項之輸入/輸出節點，更包括經由連接以控制來自該第一通訊路徑至該第二通訊路徑及至該周邊匯流排，以及來自該第三通訊路徑至該第四通訊路徑及至該周邊匯流排之指令之傳送的控制單元。

3. 如申請專利範圍第2項之輸入/輸出節點，其中該控制單元更經由配置以控制來自該周邊匯流排至該第二通訊路徑及該第四通訊路徑之指令之傳送。

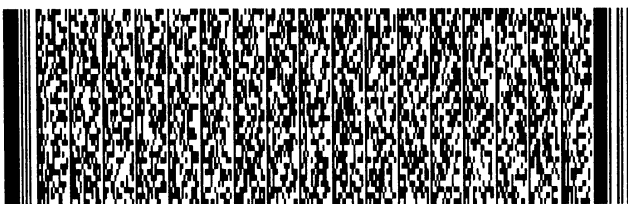
4. 如申請專利範圍第3項之輸入/輸出節點，其中該橋接器單元更經由配置以選擇性地提供對應於來自該周邊匯流排至該第一發送器及該第二發送器所接收之指令



六、申請專利範圍

的指令。

5. 如申請專利範圍第4項之輸入/輸出節點，其中該控制單元更經由配置而基於來自該第一接收器、該第二接收器及該橋接器單元所接收之複數個控制指令以選擇性地控制該指令之傳送。
6. 如申請專利範圍第5項之輸入/輸出節點，其中每一個該控制指令含有由該第一與第二接收器及該橋接器單元所接收之子集合之對應的指令。
7. 如申請專利範圍第6項之輸入/輸出節點，其中該控制單元更經由配置以透過控制指令匯流排接收該控制指令。
8. 如申請專利範圍第1項之輸入/輸出節點，其中該周邊匯流排為周邊組件互連接(PCI)匯流排。
9. 如申請專利範圍第1項之輸入/輸出節點，其中該周邊匯流排為繪圖匯流排。
10. 如申請專利範圍第1項之輸入/輸出節點，其中該第一與第三通訊路徑及該第二與第四通訊路徑為 HyperTransport™ 連結。
11. 一種電腦系統，包括：
 - 一個或一個以上之處理器；
 - 連接在一起並且連接至該一個或一個以上之處理器的一個或一個以上之輸入/輸出節點，每個該輸入/輸出節點包含：
 - 經由配置以接收在第一通訊路徑上之第一指令的



六、申請專利範圍

第一接收器單元；

經由連接以發送第一對應的指令之第一發送器單元，該指令對應於在第二通訊路徑上之該第一指令；

經由配置以接收在第三通訊路徑上之第二指令的第二接收器單元；

經由連接以發送第二對應的指令之第二發送器單元，該指令對應於在第四通訊路徑上之該第二指令；
以及

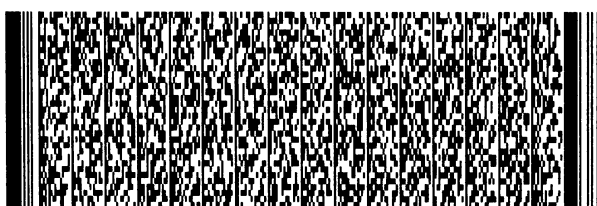
經由連接以接收來自該第一接收器與該第二接收器之選擇的指令及經由配置以發送對應於該選擇的指令之指令於周邊匯流排上的橋接器單元。

12.如申請專利範圍第11項之電腦系統，其中該輸入/輸出節點更包括經由連接以控制來自該第一通訊路徑至該第二通訊路徑及至該周邊匯流排，以及來自該第三通訊路徑至該第四通訊路徑及至該周邊匯流排之指令之傳送的控制單元。

13.如申請專利範圍第12項之電腦系統，其中該控制單元更經由配置以控制來自該周邊匯流排至該第二通訊路徑及該第四通訊路徑之指令之傳送。

14.如申請專利範圍第13項之電腦系統，其中該橋接器單元更經由配置以選擇性地提供對應於來自該周邊匯流排至該第一發送器及該第二發送器所接收之指令的指令。

15.如申請專利範圍第14項之電腦系統，其中該控制單元



六、申請專利範圍

更經由配置而基於來自該第一接收器、該第二接收器及該橋接器單元所接收之複數個控制指令以選擇性地控制該指令之傳送。

16. 如申請專利範圍第 15 項之電腦系統，其中每一個該控制指令含有由該第一與第二接收器及該橋接器單元所接收之子集合之對應的指令。
17. 如申請專利範圍第 16 項之電腦系統，其中該控制單元更經由配置以透過控制指令匯流排接收該控制指令。
18. 如申請專利範圍第 11 項之電腦系統，其中該周邊匯流排為周邊組件互連接 (PCI) 匯流排。
19. 如申請專利範圍第 11 項之電腦系統，其中該周邊匯流排為繪圖匯流排。
20. 如申請專利範圍第 11 項之電腦系統，其中該第一與第三通訊路徑及該第二與第四通訊路徑為 HyperTransport™ 連結。

