

公告本

申請日期	P. P. 21
案 號	P. 120536
類 別	G06F 12/08

A4

C4

542958

(以上各欄由本局填註)

發明專利說明書

一、發明 名稱 <small>新型</small>	中 文	用以管線作業有序之輸入/輸出事務處理至分佈式記憶體，快取相關，多處理器系統中之相關記憶體之方法及裝置
	英 文	A METHOD AND APPARATUS FOR PIPELINING ORDERED INPUT/OUTPUT TRANSACTIONS TO COHERENT MEMORY IN A DISTRIBUTED MEMORY, CACHE COHERENT, MULTI-PROCESSOR SYSTEM
二、發明人 <small>創作人</small>	姓 名	1. 曼諾 凱爾 MANOJ KHARE 2. 艾奇利 庫瑪 AKHILESH KUMAR 3. 莉莉 P. 羅伊 LILY P. LOOI 4. 肯尼士 C. 克利塔 KENNETH C. CRETA
	國 籍	1.2. 印度 3.4. 美國
三、申請人	住、居所	1. 美國加州薩拉多加市奧瑞拉廣場12650號 2. 美國加州桑尼維爾市蘭特大道1256號 3. 美國俄勒岡州波特蘭市西北密查爾街15134號 4. 美國華盛頓州吉格港市西北第37街3010號
	姓 (名稱)	美商英特爾公司 INTEL CORPORATION
	國 籍	美國
	住、居所 (事務所)	美國加州聖塔卡拉瓦市米遜大學路2200號
	代表人 姓	F. 湯姆士. 當烈二世 F. THOMAS DUNLAP, JR.

由本局填寫	承辦人代碼：
	大類：
	I P C 分類：

A6

B6

本案已向：

國（地區）申請專利，申請日期：案號：，有 無主張優先權美國 2000年08月21日 09/643,380 有 無主張優先權

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

有關微生物已寄存於：，寄存日期：，寄存號碼：

五、發明說明(1)

發明背景

1. 發明範圍

本發明之一具體實施例與電腦系統之範圍有關，包含分散式記憶體，而且，更特別的是，本發明與在此類系統中，以管線作業之輸入/輸出事務處理之方法有關。

2. 相關先前技藝討論

某些輸入/輸出(I/O)匯流排，例如週邊元件相互連接(PCI)匯流排，具有嚴謹的事務處理次序之要求。在一電腦系統中，此一I/O匯流排與記憶體之間只有一條路徑，因此I/O匯流排對於管線作業有序之要求相對的較容易。該方法對於具有非分散式記憶體架構之系統作業較佳，因為一單一路徑存在於I/O匯流排與記憶體之間。

然而，對於具有一分散式記憶體架構之系統而言，一I/O匯流排與記憶體之間具有多重路徑，由於該複雜性包含透過多重路徑維持處理作業，因此常見的作業並非來自於I/O匯流排之管線作業有序之請求。使用一非管線作業之方法，在對記憶體發出下一個I/O匯流排請求之前，一I/O橋接器將完成第一個I/O匯流排請求，此一方法用以導引事務處理朝向分散式記憶體，其可限制I/O處理能力之完成。此方法可能造成所有系統效能下降，或是要求重要的軟體進行修正，以避免限制I/O處理能力之完成。

圖示簡述

本發明係藉由舉例之方式加以說明，而且並非對隨之圖示加以限制，其中類似的涵義表示相似的元件，而且其

裝
訂

線

五、發明說明 (2)

中：

圖1係一分散式記憶體、快取相關、多處理器系統之範例之方塊圖，其中根據一具體實施例所述，有序之I/O作業處理可有效率的進行管線作業。

圖2係一輸入/輸出橋接器之具體實施例之方塊圖，其對於有序之I/O事務處理提供有效率之管線作業，並且其可用於圖1之系統。

圖3係說明圖2之預先擷取引擎(engine)之作業流程圖。

圖4係說明圖2之取回引擎之作業流程圖。

圖5係說明一種用於管線作業有序之I/O事務處理具體實施例之方法之流程圖。

詳細描述

下面將描述一種用於管線作業有序之輸入/輸出事務處理至一分散式記憶體、快取相關系統之方法及裝置，在下面的描述中，將對特別的積體電路、系統及電路型態描述其目的，然而，吾人能了解，其他的具體實施例以另一種方式組態，應用於其他的積體電路及電路或系統之型態。

以一具體實施例來說，一預先擷取引擎預先擷取來自一分散式、相關記憶體之資料，且由一輸入/輸出匯流排回應一第一事務處理，並導引至分散式、相關記憶體。一相關快取緩衝器接收預先擷取資料，並且與在系統中之分散式、相關記憶體及其他相關快取記憶體維持相關性。

圖1代表一電路系統100之範例之方塊圖，其中並可以有效的使用本發明之一具體實施例。電腦系統100係一分散式

裝
訂
線

五、發明說明(3)

記憶體、快取相關、多處理器系統，該範例中系統100包含處理節點105及110，其間並包含多數的處理節點，此外在系統100包含一相互連接網路115及一輸入/輸出節點120。

每一個處理節點105及110分別包含一或多個處理指令之處理器125，一或多晶片組元件130及135，及一或多本地記憶體140及145，晶片組元件130及135具有許多功能，例如執行記憶體控制、多工處理支援、及/或對於個別之節點維持快取相關。一或多處理器125，每一個均包含及耦合至一快取記憶體147，對此一具體實施例來說，快取記憶體147係與彼此及與藉由記憶體140及145於分散式、相關記憶體之應用具有相關性。

此一具體實施例中之記憶體140及145作為一分散式、相關主記憶體共同作業，每一個記憶體140及145可視為一較大記憶體之部分或一區域，此外也包含非相關性區域，對此一具體實施例來說，使用眾所皆知的MESI(修改，互斥，共享或無效)狀態協定，在快取記憶體147與記憶體140及145之間維持相關性，並且與相關系統相互連接匯流排160、165及170產生關聯，為了在不同的記憶體140、145與在I/O節點120之一相關快取緩衝器之間維持相關性，相關系統相互連接匯流排160、165及170將藉由相互連接網路115，在處理節點105及110及I/O節點120之間傳遞資訊，也就是如下面所描述之細節。對此一具體實施例來說，相關系統相互連接匯流排160、165及170係在處理節點與I/O節點之間以點對點方式相互連接，通訊事務處理方

裝
訂

線

五、發明說明 (4)

面而言並沒有次序之限制。

相互連接網路115可於I/O節點120與處理節點(例如處理節點105及110)之間傳遞訊息，對此一具體實施例來說，相互連接網路115並未在這些訊息之間(I/O節點120與一或更多之處理節點105及110間之通訊)維持有序之作業。

I/O節點120包含一連接一或多個I/O匯流排155之I/O橋接器150，例如一週邊元件相互連接匯流排(PCI)與處理節點105及110，一I/O橋接器120之具體實施例之進一步的細節將於下面，並參考圖2進行描述。

對其他的具體實施例來說，系統100可以另一種方式進行組態，例如，晶片組及/或記憶體可包含於一或多個處理器125之內，且/或相互連接網路115可不包含在內。而且，記憶體之相關性可依據不同的協定及/或使用不同型態之相互連接方法來維持，具有比一I/O匯流排及記憶體之間更多之路徑之其他的系統型態，也同樣的包含於不同的具體實施例之範圍內。

圖2係一I/O橋接器250之方塊圖，對此一具體實施例來說，可使用於圖1之I/O橋接器150之應用中。I/O橋接器250包含一或多個I/O快取緩衝器205，一或多個I/O事務處理請求緩衝器210，一預先擷取引擎215及一取回(retire)引擎220，其他的電路系統，例如其他的I/O事務處理方法，緩衝法(buffering)及控制電路系統(並未表示出來)也同樣包含在內。

下面將作更細部的描述，快取緩衝器205包含位址、資料

裝
訂
線

五、發明說明 (5)

及狀態欄位，且用來儲存預先擷取資料，並且回應一或更多I/O事務處理請求，以幫助I/O事務處理請求之管線作業，該位址欄位表示儲存於記憶體中所對應之資料欄位之資料位址，該快取緩衝器205之一具體實施例係與系統中其他的相關記憶體維持相關性，其中I/O橋接器也包含在內，例如，I/O橋接器係包含於圖1中之系統100，快取緩衝器205與分散式、相關記憶體140、145及相關快取記憶體147維持相關性，I/O快取緩衝器之狀態欄位係用於表示所對應之資料欄位之資料相關狀態(例如，此一具體實施例所使用之MESI協定，M(修改)、E(互斥)、S(共享)、I(無效))。

I/O事務處理請求緩衝器210(請參考I/O請求緩衝器)產生儲存事務處理請求，即由一或更多I/O匯流排255導引至圖1所描述之分散式、相關記憶體，此請求可改變成為對內的相關請求及對內的相關I/O請求。

對此一具體實施例而言，I/O事務處理請求緩衝器210可用於儲存所有的I/O事務處理請求，不論是否有導引他們至相關記憶體之範圍。I/O請求並未被導引至相關記憶體，因此可參考作為非相關事務處理請求，例如，可包含I/O埠(port)存取、架構存取、中斷及/或特殊之事務處理(例如鎖定)，因此，即使某些導引至相關記憶體範圍之I/O事務處理，亦可被分類成為非相關事務處理，例如，一鎖定序列期間，至相關記憶體進行存取之記憶體可被分類成為非相關事務處理。

所預期之相關快取緩衝器之尺寸及/或輸入/輸出請求緩衝

裝訂

線

五、發明說明 (6)

器可藉由平衡記憶體及所有I/O處理能力與緩衝器中可獲得之空間來決定。

同樣的，對於某些具體實施例而言，一I/O事務處理請求寫入或讀取具有多線條之記憶體，並儲存於I/O事務處理請求緩衝器210，以作為一單一事務處理，而非藉由預先擷取及取回引擎之方式作為記憶體線條範圍中之多重對單一之線條事務處理，對此一具體實施例而言，這些單一線條事務處理之每一個方法係根據下面所描述之方法，用於單一對單一之線條事務處理。

一具體實施例之I/O事務處理請求緩衝器210包含位址、資料及控制欄位，位址欄位表示與請求相結合之位址，例如存取記憶體位址。資料欄位產生對應之資料，並寫入有關於寫入請求之記憶體，或是恢復一讀取請求之資料位置。而且控制欄位可用於表示相對的指令有序之作業資訊，請求類型及/或與I/O事務處理請求有關之額外資訊。

I/O請求緩衝器210儲存來自I/O匯流排255之相關及非相關I/O事務處理請求，事務處理之類型可藉由位址與事務處理、請求類型及/或位於控制欄位之控制資訊相結合來決定。

對此一具體實施例而言，每一個耦合至I/O橋接器250之不同的I/O匯流排255，I/O橋接器250包含一類似I/O請求緩衝器210之分離之I/O請求緩衝器。對另一具體實施例而言，一單一之I/O請求緩衝器210提供暫時的儲存I/O事務處理請求，並由耦合至I/O橋接器250之所有I/O匯流排255接

裝
訂
線

五、發明說明 (7)

收。對這個具體實施例而言，I/O事務處理請求緩衝器210之控制欄位同樣表示接收事務處理之匯流排，對另一具體實施例而言，I/O橋接器250可包含一些類似於I/O事務處理請求緩衝器210之不同之I/O事務處理請求緩衝器，其中一或更多緩衝器係由並聯之I/O匯流排255所共享。

同樣的，快取緩衝器205可包含並聯之快取緩衝器或一單一之快取緩衝器。

預先擷取引擎215及收回引擎220可使用兩個分離之電路區塊執行，或將它們組合以提供下面所描述之功能。

在作業上而言，由一或更多I/O匯流排255所產生之對內的事務處理請求係暫時的儲存於I/O事務處理請求緩衝器210。以一具體實施例而言，I/O事務處理請求緩衝器210當作一先進先出(FIFO)緩衝器來應用，因此進來之事務處理依順序儲存於緩衝器，其中它們由I/O匯流排255接收，而且有序之事務處理並不需作其他的表示。以另一具體實施例而言，相關的事務處理有序作業可於控制欄位中表示。

預先擷取引擎215使至相關記憶體之事務處理請求之管線作業能夠作業，該預先擷取引擎藉由執行一非束縛(non-binding)預先擷取之資料，並與I/O事務處理請求緩衝器210之事務處理相結合，然後儲存該預先擷取之資料於I/O快取緩衝器205。當處理另一個事務處理或是等待一先前之預先擷取作業之回應時，預先擷取之資料將會被執行，另外，預先擷取引擎215可預先擷取資料，並與I/O事務處理

裝
訂
線

五、發明說明 (8)

請求緩衝器之任一個相關事務處理請求相結合，不論其是否依序被接收。以此種方式來說，管線作業將能執行，預先擷取引擎215之一具體實施例之作業請參考圖2及圖3，以進行更細部的描述。

如圖3之區塊305所表示，首先 預先擷取引擎215由I/O事務處理請求緩衝器210選擇一相關請求，一些方法中之任一個均可用於選擇懸置之相關請求，並由預先擷取引擎215進行作業，例如，預先擷取引擎215可單純的選擇緩衝器210之下一個懸置請求，如果具有並聯之I/O事務處理請求緩衝器，該預先擷取引擎215將可使用一劃分時段之方法，例如，選擇一緩衝器之下一個懸置相關請求。另外，由下一個懸置相關請求之預先擷取引擎之作業將可任意改變，然而，不管使用什麼方法，預先擷取引擎215在選擇懸置請求或是執行預先擷取作業時，將不需要遵守有序作業之規定。

此外，以一具體實施例而言，如果(例如) 預先擷取引擎215在其他方面閒置(例如，在I/O事務處理請求緩衝器210中，並沒有懸置事務處理請求)，預先擷取引擎215將可預先擷取資料，以該具體實施例而言，預先擷取引擎215可使用一些方法中任何一種，以決定資料在預期之I/O事務處理上可預先擷取，例如，預先擷取引擎215可預先擷取下一個記憶體線條或是擷取接在預先擷取資料後之線條，在這個方法中，如果一即將來臨的對內之I/O事務處理請求在先前之請求後表示出連續的資料，該資料將會立即的被獲得，

裝
訂
線

五、發明說明 (9)

其他決定資料是否預先擷取之方法也包含在不同的具體實施例之範圍之內。

在決策方塊310中，其決定是否所選擇之相關請求不是讀取請求就是寫入請求，如果(例如)預先擷取引擎215作業在一懸置之讀取請求，那麼在區塊315中，預先擷取引擎215將決定是否一有效的請求記憶體線條之複製可於I/O快取緩衝器205獲得，如果如此的話，在區塊320之預先擷取引擎仍舊決定是否在I/O事務處理請求緩衝器210有更多之懸置請求，且在區塊305仍舊預先擷取資料，並與在I/O事務處理請求緩衝器210之其他懸置請求相結合。

在決策區塊315中，如果一有效之請求記憶體線條之複製並未在本地的I/O快取緩衝器205中獲得，那麼在區塊325中，預先擷取引擎215將透過相關系統相互連接270發出一讀取請求，以預先擷取所請求之記憶體線條，並進行讀取，在這裡I/O橋接器包含於類似圖1中之系統100之一系統，透過相關系統相互連接270發出一讀取請求，藉由相互連接網路傳遞資訊，並透過相關系統相互連接匯流排耦合至不同種類的系統處理節點，再連接至系統之分散式、相關記憶體，在這裡該系統包含以不同方法組態而成之I/O橋接器250，透過相關系統相互連接發出一讀取請求可藉由不同的路徑到達分散式、相關記憶體。

該預先擷取作業，如同一具體實施例之預先擷取引擎215所執行之其他的預先擷取作業，是一個非束縛之預先擷取作業，在本文中非束縛的意思就是，如果由系統中某些其

裝

訂

線

五、發明說明 (10)

他之記憶體/快取代理程式透過相關系統相互連接270接收一進入之窺探(snoop)請求，而且這個窺探請求撞擊I/O快取緩衝器205之一線條，對於分散式、相關記憶體(以一具體實施例之MESI協定而言)之記憶體/快取相關協定來說是被跟隨的(followed)，因此，所對應之事務處理請求之預先擷取作業及收回作業之間，儲存於I/O快取緩衝器205之預先擷取記憶體線條將可能無效，而且/或I/O快取緩衝器205可能失去線條本身，處理發出之方法請參考收回引擎220之作業，並於下面做更細部的討論。

繼續參考圖2及圖3，由相關系統相互連接中等待一回應時，在區塊320中，預先擷取引擎215可開始決定是否在I/O事務處理請求緩衝器210有更多之懸置請求，或是否需預先擷取額外的資料，預先擷取引擎能持續預先擷取資料，並對在I/O事務處理請求緩衝器210其他的懸置請求或是描述在上面之預先擷取資料之方式做出回應。

在區塊330中，透過相關系統相互連接170所接收之請求資料，預先擷取引擎215在I/O快取緩衝器205以一共享狀態配置資料，並藉由記憶體位址之表示與位於I/O快取緩衝器之位址欄位資料、位於對應之資料欄位之請求資料及位於對應之狀態欄位之快取相關狀態(在此範例中為共享)相結合來共同完成。

回到決策區塊310，如果選擇請求以寫入請求替代，那麼在區塊335中，該預先擷取引擎215將決定是否I/O快取緩衝器205擁有一對應至寫入請求之互斥的記憶體線條之複製

裝訂線

五、發明說明 (11)

(例如，使用 MESI 協定，不是以 E 就是 M 狀態來說明)，如果像這樣的話，在區塊 320 中，如先前所描述，該預先擷取引擎 215 可持續預先擷取對應至其他的懸置請求之資料或是如上面所描述之推測性的預先擷取資料。

如果在決策區塊 335 中，I/O 快取緩衝器 335 並未擁有與寫入請求相結合之一互斥的且特別之記憶體線條之複製，那麼在區塊 340 中，預先擷取引擎 215 發出一請求，以透過相關系統相互連接 270 預先擷取記憶體線條本身，對一具體實施例而言，請求之方式將依寫入請求是否為一部份的線條寫入請求或是全部的線條寫入請求而不同，例如，對全部的線條寫入請求來說，預先擷取引擎沒有對於記憶體線條之內容進行讀取，即可發出一使其他的記憶體複製無效之無效請求，然而，對部份的線條寫入請求來說，預先擷取引擎可發出對於使其他的記憶體複製無效之讀取請求，並且獲得記憶體線條之內容。

類似的讀取請求，當透過相互連接 270 等待窺探回應時，預先擷取引擎 215 可持續預先擷取與其他的懸置 I/O 事務處理請求相結合之資料。

在區塊 345 中，由上端接收一窺探請求時，預先擷取引擎在 I/O 快取緩衝器 205 之互斥狀態(例如，根據 MESI 協定之"E"狀態)中配置請求記憶體線條，那麼一具體實施例之方法將在區塊 320 中持續決定是否有更多之懸置 I/O 請求來進行作業。

與上面預先擷取引擎 215 平行之取回引擎 220，其依序的

裝訂線

五、發明說明 (12)

由I/O事務處理請求緩衝器210完成作業，以取回I/O事務處理請求，取回引擎215之一具體實施例之作業請參考圖2及圖4進行描述。

在區塊405中，取回引擎220依序的選擇下一個事務處理，以由I/O事務處理請求緩衝器210取回資料，對一具體實施例而言，取回引擎220由一特別之I/O匯流排接收時，可僅依序的取回事務處理請求，以另一具體實施例而言，取回引擎220可依據特定之有序規則取回事務處理請求，其可因所接收之事務處理順序而產生差異，對一具體實施例而言，這些有序之需求可在事務處理中詳細說明，然後根據狀態機(例如取回引擎220)進行處理。

例如，對一具體實施例而言，一讀取事務處理請求與獲得於I/O快取緩衝器205之對應資料可於早先接收、讀取事務處理請求之前取回，如此一後續有效之讀取請求將傳達一未完成之先前的讀取請求，然而，對相同的具體實施例來說，在早先接收一寫入請求之前可不將一讀取請求取回。其他的有序規則之型態可利用於取回引擎之上，以確定錯誤之資料不會由系統中之記憶體寫入或讀出，該指令並將依照順序取回，以防止錯誤的讀取或參考寫入事務處理作為程式之順序。

而且，如上面所描述，對此一具體實施例而言，I/O事務處理請求緩衝器210亦可儲存非相關I/O事務處理請求，對類似一具體實施例而言，由取回引擎維持之取回順序也包含有序的撤回該非相關I/O事務處理請求。

裝訂線

五、發明說明 (13)

在區塊410中，一旦選取撤回之I/O事務處理請求，取回引擎220將決定在區塊415中是否該事務處理為一讀取事務處理或是寫入事務處理，如果該事務處理為一讀取之事務處理，那麼在區塊420中，取回引擎220將決定是否與該事務處理相結合之記憶體線條在I/O快取緩衝器205中為存在且有效的，如果如此的話，那麼在區塊425中，由I/O快取緩衝器205所對應之資料將透過I/O匯流排之對應資料送回至請求I/O代理程式255。

假如在決策區塊420中，其決定所對應之記憶體線條不存在於I/O快取緩衝器205就是無效的話(也許如上面所描述由系統中另一個快取代理程式產生窺探請求)，那麼在區塊440中，取回引擎將透過相關系統相互連接270發出一讀取請求，以擷取對應至事務處理請求之記憶體線條，一旦請求資料透過相關系統相互連接270送回至I/O橋接器250，該資料將提供至區塊425中之請求I/O代理程式。

在區塊430中，一旦I/O事務處理請求完成服務，其將由I/O事務處理請求緩衝器210中移除，這可以一些方法中之任何一種來完成，對一具體實施例而言(例如)，I/O事務處理請求緩衝器210之資料將被清除，以另一具體實施例而言，在緩衝器210之線條所獲得之後續的資料，可透過使用一旗標或是另一種方式來覆寫。

回到決策區塊415，如果將I/O事務處理請求取回以寫入請求取代，那麼在決策區塊445中，其將決定是否I/O快取緩衝器205在互斥狀態(例如，根據MESI協定，不是M就是

裝
訂
線

五、發明說明 (14)

E 狀態) 中擁有對應之記憶體線條，如上面所描述，即使記憶體線條已於先前預先擷取，I/O 快取緩衝器 205 可能在預先擷取記憶體線條時及取回事務處理請求時之間失去線條本身，這可能由於(例如)來自系統中另一個快取代理程式之一中間的窺探作業撞擊在 I/O 快取緩衝器 205 之請求快取線條。

如果 I/O 快取緩衝器 205 在互斥狀態中擁有記憶體線條，那麼在區塊 450 中，與記憶體線條相結合之資料將根據寫入請求及線條紀錄之狀態(如修改)於 I/O 快取緩衝器 205 中進行更新(該資料將根據暫時儲存於寫入緩衝器(並未表示出來)之寫入請求寫入記憶體線條，例如，預先寫入 I/O 快取緩衝器 205 之記憶體線條)。

如果寫入請求係一完全的線條寫入請求，全部的資料線條將根據寫入請求簡單的寫入 I/O 快取緩衝器 205 中所對應之資料欄位。假如寫入請求為部分的線條寫入請求，該資料將以熟悉此技藝之人士所熟知之方式寫入，並與目前在 I/O 快取緩衝器 205 之對應線條之資料合併。

以一具體實施例而言，假如在 I/O 快取緩衝器 205 之一被修改之線條以一新的線條取代，並對任何擷取、預先擷取或是獲得之作業做出回應時，最好容許對於 I/O 快取緩衝器 205 狀態進行追蹤，例如，I/O 橋接器 250 透過相關系統相互連接 270 送出一請求，以在分散式記憶體中適當的更新線條。如果在 I/O 快取緩衝器 205 之一未使用過之線條(例如目前無資料儲存於快取線條)對一預先擷取作業回應時被取

裝

訂

線

五、發明說明 (15)

代，例如，I/O橋接器250透過相關系統相互連接270送出一窺探請求，以表示對系統中快取緩衝器之狀態進行追蹤之動作。這些窺探請求可藉由包含於I/O橋接器(並未表示出來)內之其他的相關控制邏輯傳送，其他的相關控制邏輯也執行某些功能，例如對其他型態之動作做出回應之窺探，及/或解釋與I/O快取緩衝器205及/或系統中其他快取代理程式相結合之相關狀態。

請繼續參考圖2及圖4，在決策區塊445中，假如I/O快取緩衝器205並未擁有對應至互斥狀態之寫入請求之記憶體線條，那麼在區塊455中，收回引擎220將以上面所描述對應至預先擷取作業之一相似的方式，透過相關系統相互連接請求且獲得記憶體線條，那麼在區塊450之收回引擎將依照上面所描述之方式收回寫入請求，並持續更新在I/O快取緩衝器205之資料。

一旦完成在區塊430之請求後，該請求將由I/O事務處理請求緩衝器210中移除。

在決策區塊435中，其將決定在I/O事務處理請求緩衝器210中是否有更多之懸置請求，如果如此的話，它們將以類似的方式進行處理及收回。

使用上面所描述之方式，系統中管線作業有序之I/O事務處理請求，可能在一I/O匯流排及記憶體之間具有多重路徑(如同在系統中具有分散式、相關記憶體)，這個管線作業將部分的藉由一預先擷取引擎，執行促進非束縛、無次序之預先擷取作業，並回應至相關記憶體，以存取一或多個I/O

裝
訂

線

五、發明說明 (16)

匯流排之請求。此外，上面所描述之本地的I/O快取緩衝器雖然與分散式記憶體子系統及快取記憶體之剩餘數維持相關性，其將藉由儲存預先擷取作業之結果促進管線作業之運作。為使I/O請求之管線作業能夠改善I/O處理能力，因此，全面的系統效能將與先前多重處理器、分散式記憶體、快取相關系統之處理I/O事務處理之方法做比較。

圖5為說明一管線作業有序之I/O事務處理具體實施例之方法之一流程圖，在區塊505中，欲緩衝之I/O事務處理係由一I/O匯流排導引至分散式、相關記憶體，在區塊510中，資料係由分散式、相關記憶體預先擷取，並對第一欲緩衝之輸入/輸出事務處理做出回應，且暫時儲存於區塊515。在區塊520中，預先擷取資料及儲存於分散式、相關記憶體與其他快取記憶體之資料之間維持相關性，並且在區塊525中，該欲緩衝之I/O事務處理將依序的被收回，吾人將了解，對於其他之具體實施例而言，該方法可以只包含上面所描述之某些動作，或是可包含非上面所描述之額外的動作，此外，一或更多之動作將比上面所描述或是與上面所描述之另一個同時發生之動作，以不同之順序執行。

在之前的詳細說明中，本發明已參考特殊範例之具體實施例來進行描述，然而，吾人將了解不同的修改或改變而沒有偏離本發明申請專利範圍之廣泛的精神及範圍亦能採用，因此，該詳細說明及圖示將被當作一例證，而不是一限制性之意義。

裝
訂
線

四、中文發明摘要（發明之名稱： 用以管線作業有序之輸入/輸出事務處理至分佈式記憶體，快取相關，多處理器系統中之相關記憶體之方法及裝置)

一種以管線作業有序之輸入/輸出事務處理至分佈式記憶體、快取相關、多處理器系統中之相關記憶體之方法。一預先擷取引擎(prefetch engine)由分散式、相關記憶體預先擷取資料，其由一輸入/輸出匯流排回應此一處理，並導引至分散式、相關記憶體。一輸入/輸出相關快取緩衝器接收預先擷取之資料，並且與分散式、相關記憶體及系統中其他快取代理程式(agents)維持相關性。

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

英文發明摘要（發明之名稱： A METHOD AND APPARATUS FOR PIPELINING ORDERED INPUT/OUTPUT TRANSACTIONS TO COHERENT MEMORY IN A DISTRIBUTED MEMORY, CACHE COHERENT, MULTI-PROCESSOR SYSTEM)

An approach for pipelining ordered input/output transactions to coherent memory in a distributed memory, cache coherent, multi-processor system. A prefetch engine prefetches data from the distributed, coherent memory in response to a transaction from an input/output bus directed to the distributed, coherent memory. An input/output coherent cache buffer receives the prefetched data and is kept coherent with the distributed, coherent memory and with other caching agents in the system.

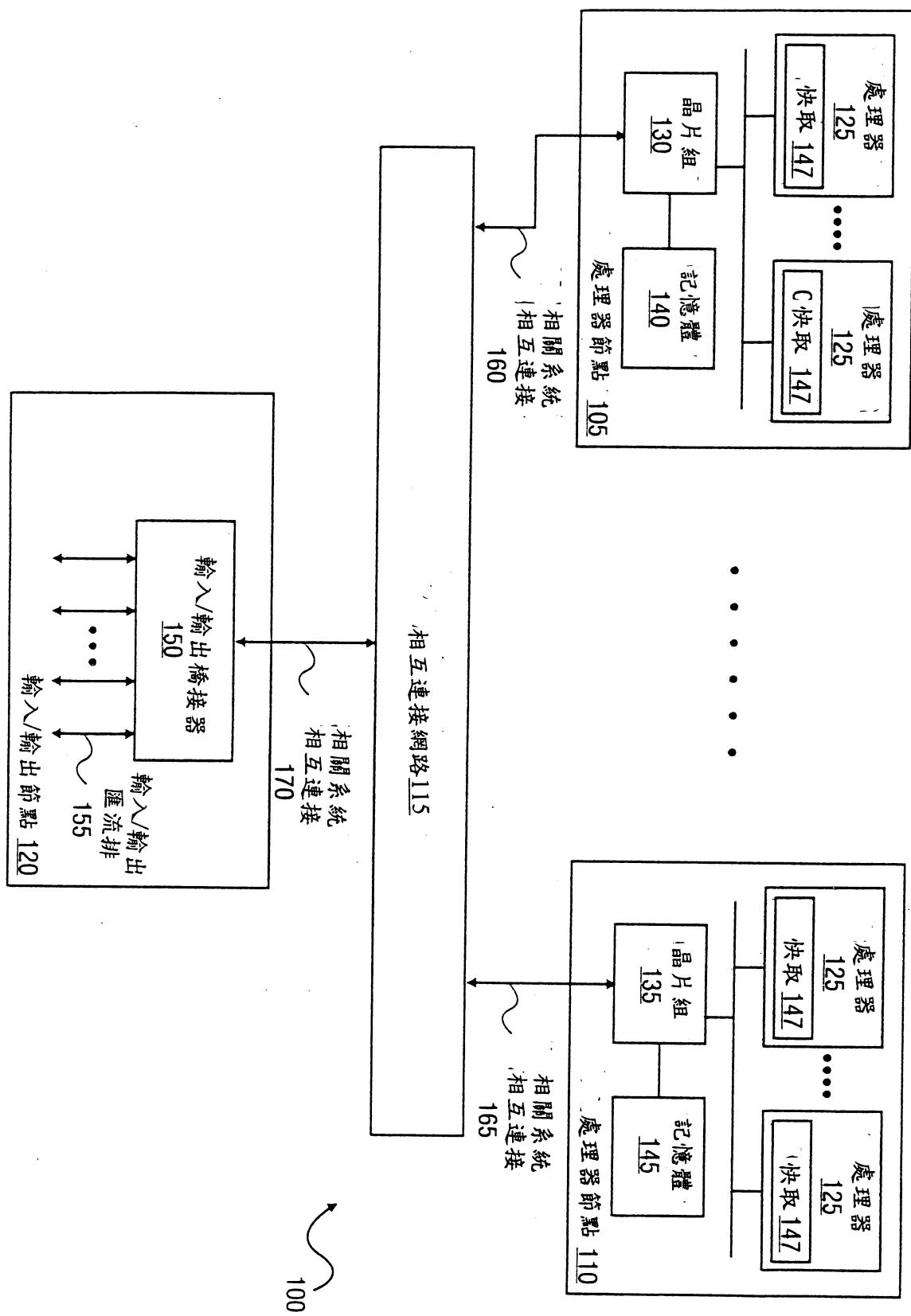


圖 1

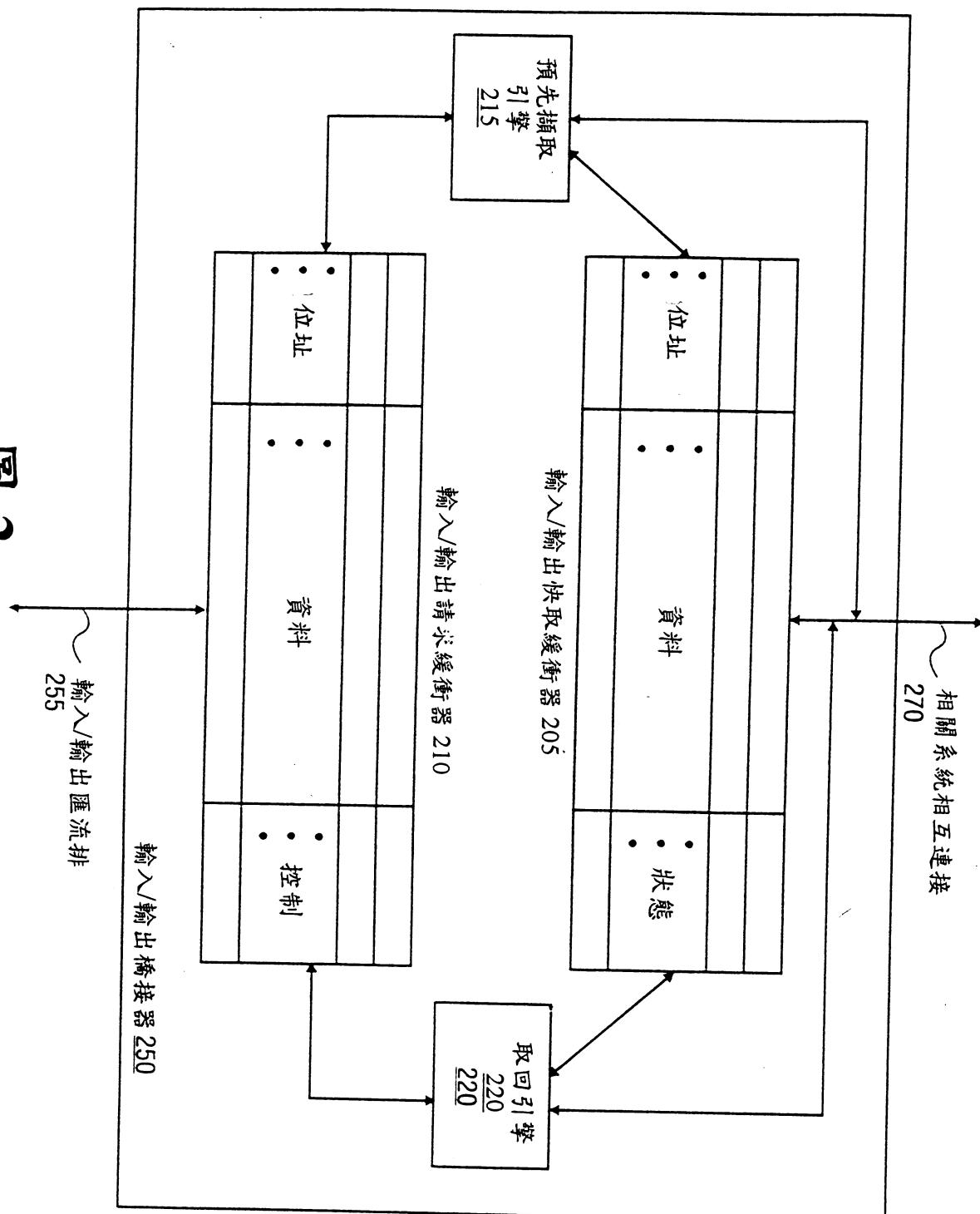


圖 2

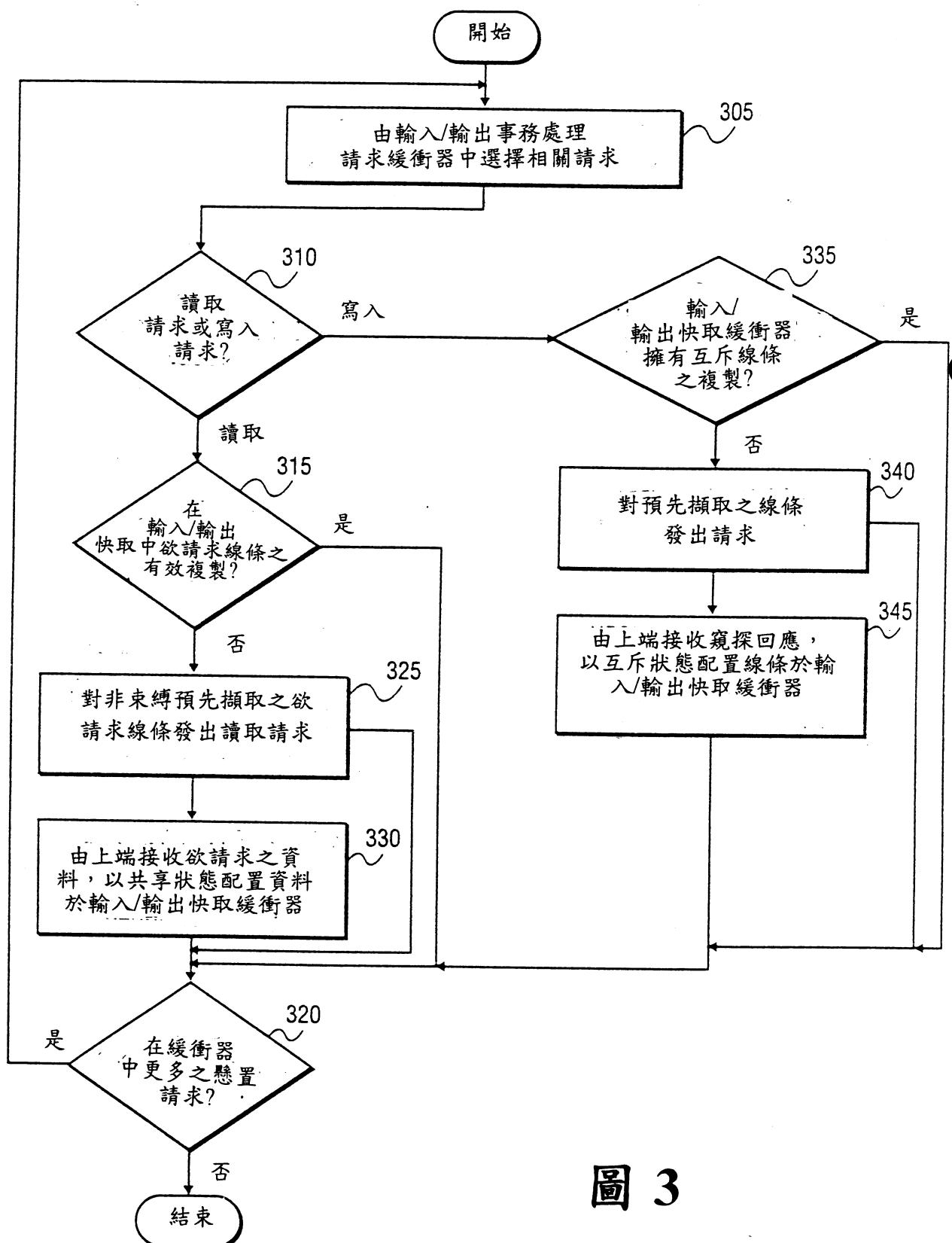


圖 3

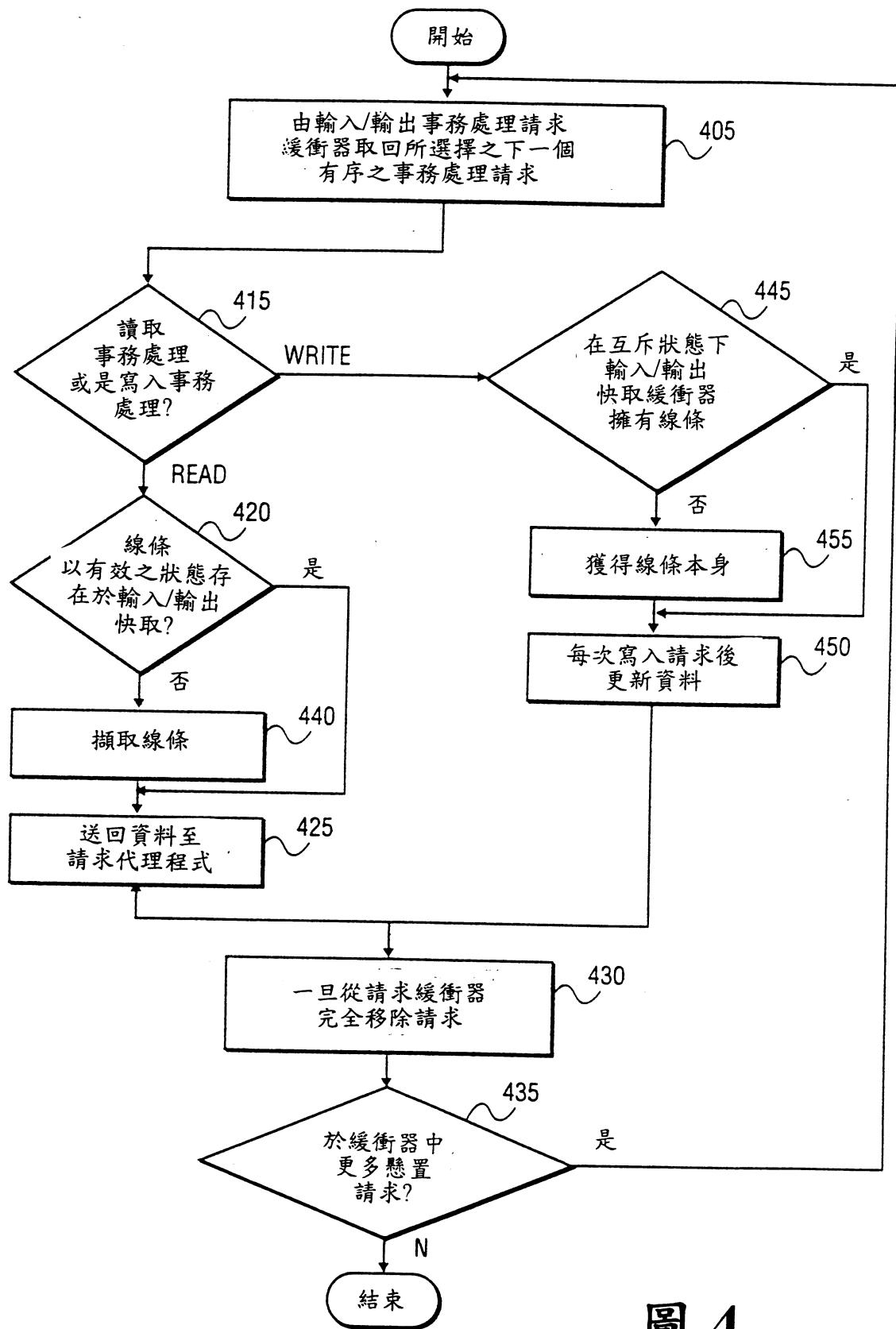


圖 4

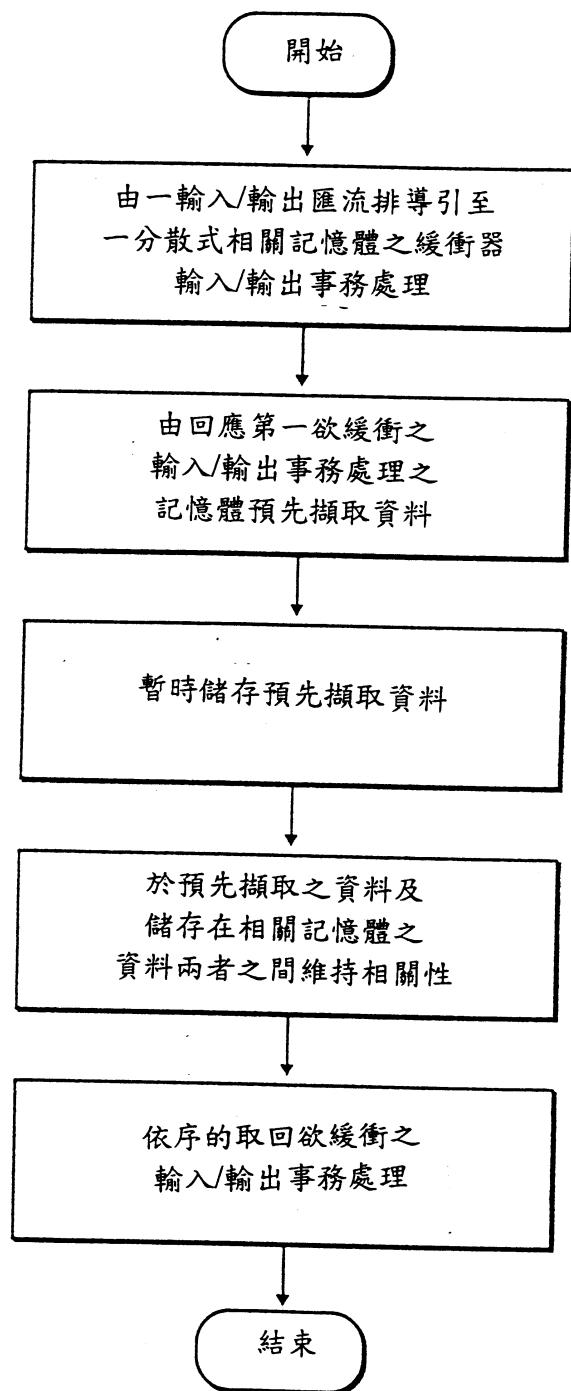


圖 5

92年5月1日 修正
補充

六、申請專利範圍

1. 一種用以管線作業輸入/輸出事務處理至相關記憶體之裝置，包括：

一由分散式、相關記憶體預先擷取資料之預先擷取引擎，其由一輸入/輸出匯流排回應第一事務處理，並導引至分散式、相關記憶體；及

一接收預先擷取資料之輸入/輸出相關快取緩衝器，該相關快取緩衝器與分散式、相關記憶體及與系統中其他的快取記憶體，包含輸入/輸出相關快取緩衝器，維持相關性。

2. 如申請專利範圍第1項之裝置，其中由預先擷取引擎執行之預先擷取作業係一非束縛之預先擷取作業，由相關快取緩衝器接收之預先擷取資料可藉由位於一分散式、相關記憶體之記憶體而改變。

3. 如申請專利範圍第2項之裝置，其中第一事務處理請求係一記憶體讀取請求，而且預先擷取引擎發出一讀取請求，以由回應第一事務處理之分散式、相關記憶體中預先擷取資料並進行讀取。

4. 如申請專利範圍第2項之裝置，其中第一事務處理請求係一記憶體寫入請求，且位於分散式、相關記憶體之預先擷取引擎發出一預先擷取記憶體線條之請求，該記憶體線條並藉由第一事務處理請求來表示。

5. 如申請專利範圍第1項之裝置，進一步包含：

一輸入/輸出事務處理請求緩衝器，其暫時儲存由輸入/輸出匯流排接收之事務處理請求，並導引至分散式、相關記憶體。

裝
訂

92年5月5日
修正
補充

A8
B8
C8
D8

六、申請專利範圍

6. 如申請專利範圍第5項之裝置，其中

預先擷取引擎預先擷取對事務處理請求回應之資料，並儲存於輸入/輸出事務處理請求緩衝器。

7. 如申請專利範圍第6項之裝置，其中

預先擷取引擎預先擷取對事務處理請求回應之資料，並無依照順序的儲存於輸入/輸出事務處理請求緩衝器，其中該事務處理請求係由輸入/輸出匯流排所接收。

8. 如申請專利範圍第5項之裝置，進一步包含：

一取回引擎取回輸出/輸入事務處理請求，並於完成事務處理請求之後依程式順序儲存於事務處理請求緩衝器。

9. 如申請專利範圍第8項之裝置，其中

取回引擎進一步的檢查輸入/輸出相關快取緩衝器，以決定是否取回與輸入/輸出事務處理請求相結合之資料，且在輸入/輸出相關快取緩衝器以一有效之狀態出現。

10. 如申請專利範圍第1項之裝置，其中

輸入/輸出相關快取緩衝器及使用MESI協定之分散式、相關記憶體之間係維持相關性。

11. 一種用以管線作業輸入/輸出事務處理至相關記憶體之方法，包含：

由一輸入/輸出匯流排接收回應一第一輸入/輸出事務處理請求之預先擷取資料，並且導引至分散式、相關記憶體；

暫時儲存預先擷取資料；及

該預先擷取資料與儲存於分散式、相關記憶體之資料及儲存於其他的快取記憶體之資料間維持相關性。

92年5月5日
修正
補充

A8
B8
C8
D8

六、申請專利範圍

12. 如申請專利範圍第11項之方法，進一步包含：

緩衝由輸入/輸出匯流排所接收之輸入/輸出事務處理請求，並導引至分散式、相關記憶體。

13. 如申請專利範圍第12項之方法，進一步包含：

預先擷取回應第二及第三緩衝輸入/輸出事務處理之資料，其中

預先擷取回應可以任何順序執行之第一、第二及第三緩衝輸入/輸出事務處理之資料。

14. 如申請專利範圍第12項之方法，進一步包含：

依序的收回該欲緩衝之輸入/輸出事務處理，其中該欲緩衝之輸入/輸出事務處理係由輸入/輸出匯流排發出。

15. 如申請專利範圍第14項之方法，其中收回包含：

檢查對應至暫時儲存收回之事務處理請求，以決定該暫時儲存、預先擷取之資料是否為有效的資料。

16. 如申請專利範圍第11項之方法，其中

維持相關性，包含維持使用一MESI協定之相關性。

17. 如申請專利範圍第11項之方法，其中預先擷取包含：

對回應至該第一事務處理請求之資料發出一請求，及接收被請求之資料。

18. 如申請專利範圍第17項之方法，其中

預先擷取回應一第二輸入/輸出事務處理之資料，其由輸入/輸出匯流排接收，且導引至發生於發出該請求及接收被請求資料間之分散式、相關記憶體。

19. 一種電腦系統，包含：

裝
訂

97年5月5日
修正
補充

A8
B8
C8
D8

六、申請專利範圍

第一及第二處理節點，每一個節點至少包含一處理器及至少一快取代理程式；

一分散相關記憶體，其中部分之分散式相關記憶體係包含於每一個第一及第二處理節點之內；及

一耦合至第一及第二處理節點之輸入/輸出節點，該輸入/輸出節點包含

一由分散式、相關記憶體預先擷取資料之預先擷取引擎，回應來自第一輸入/輸出匯流排之第一事務處理，並導引至分散式、相關記憶體；及

一輸入/輸出相關快取緩衝器接收預先擷取資料，該相關快取緩衝器與分散式、相關記憶體及快取代理程式具有相關性。

20. 如申請專利範圍第19項之電腦系統，進一步包含：

一耦合每一個第一及第二處理節點至輸入/輸出節點之相關系統相互連接，該相關系統相互連接傳遞資訊以維持分散式、相關記憶體之相關性，且於輸入/輸出相關快取緩衝器與分散式、相關記憶體之間維持相關性。

21. 如申請專利範圍第20項之電腦系統，其中根據一MESI協定以維持相關性。

22. 如申請專利範圍第19項之電腦系統，進一步包含

一於第一及第二處理節點與輸入/輸出節點之間傳遞資訊之相互連接網路。

23. 如申請專利範圍第19項之電腦系統，進一步包含

一於第一及第二處理節點與複數個輸入/輸出匯流排間耦

裝
訂

92年5月5日
修正
補充

A8
B8
C8
D8

六、申請專利範圍

合之輸入/輸出橋接器，複數個輸入/輸出匯流排包含第一輸入/輸出匯流排，輸入/輸出橋接器包含預先擷取引擎及輸入/輸出相關快取緩衝器。

24. 如申請專利範圍第22項之電腦系統，其中輸入/輸出橋接器進一步包含：

至少一輸入/輸出事務處理請求緩衝器暫時的儲存由複數個輸入/輸出匯流排接收之輸入/輸出事務處理請求，並導引至分散式、相關記憶體。

25. 如申請專利範圍第24項之電腦系統，其中

預先擷取引擎預先擷取資料，且不管儲存事務處理請求之順序的對儲存於輸入/輸出事務處理請求緩衝器之事務處理請求做出回應。

26. 如申請專利範圍第24項之電腦系統，其中輸入/輸出橋接器進一步包含

一收回引擎，其進一步的檢查輸入/輸出相關快取緩衝器對應於收回之事務處理請求之有效的資料，

該收回引擎依照程式的順序收回儲存於輸入/輸出事務處理請求緩衝器之事務處理請求。

裝
訂