

發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※申請案號：96146088

※申請日期：2007年12月4日

※IPC 分類：

G06F 13/16 (2006.01)

G06F 11/00 (2006.01)

一、發明名稱：(中文/英文)

以主動式預取寬度執行加強式停懸狀態之系統及方法

SYSTEM AND METHOD FOR IMPLEMENTING AN ENHANCED
HOVER STATE WITH ACTIVE PREFETCHES

二、申請人：(共1人)

姓名或名稱：(中文/英文)

美商·萬國商業機器公司

International Business Machines Corporation

代表人：(中文/英文)

安德森黎恩D

ANDERSON, LYNNE D.

住居所或營業所地址：(中文/英文)

美國紐約州阿蒙市

Armonk, New York 10504, U.S.A.

國籍：(中文/英文)

美國/USA

三、發明人：(共4人)

姓名：(中文/英文)

1. 費歐德斯二世詹姆士史帝芬/FIELDS JR., JAMES STEPHEN

2. 史達凱威廉約翰/STARKE, WILLIAM JOHN

3. 克拉愷李歐 J/CLARKE, LEO J.

4. 馬克派克金/MARK, PAK-KIN

國 籍：(中文/英文)

1.美國/USA

2.美國/USA

3.美國/USA

4.美國/USA

四、聲明事項：

☐ 主張專利法第二十二條第二項 ☐ 第一款或 ☐ 第二款規定之事實，其事實發生日期為： 年 月 日。

☒ 申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號 順序註記】

☒ 有主張專利法第二十七條第一項國際優先權：

美國；2006 年 12 月 19 日；11/612,568

☐ 無主張專利法第二十七條第一項國際優先權：

☐ 主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號 順序註記】

☐ 主張專利法第三十條生物材料：

☐ 須寄存生物材料者：

國內生物材料 【格式請依：寄存機構、日期、號碼 順序註記】

國外生物材料 【格式請依：寄存國家、機構、日期、號碼 順序註記】

☐ 不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

五、中文發明摘要：

一種以主動式預取執行加強式停懸狀態之系統及方法。根據本發明之一較佳具體實施例，處理單元中之一偵測器接收一全系統更新完成操作，其指示完成針對一特定位址之儲存體修改操作，其中該儲存體修改操作導致在一第一快取記憶體中產生一經修改之第一快取線。該偵測器在接收該全系統更新完成操作之前確定一第二快取記憶體是否保持一與該特定位址關聯之第二快取線。若如此，該偵測器發出一對於該經修改之第一快取線之一複本之預取請求，以替換該第二快取記憶體中之第二快取線。該偵測器使用該經修改之第一快取線之一複本更新該第二快取記憶體。

六、英文發明摘要：

A system and method for implementing an enhanced hover state with active prefetches. According to a preferred embodiment of the present invention, a snoopers in a processing unit receives a system-wide update complete operation indicating a completion of a storage-modifying operation targeting a particular address, where the storage-modifying operation results in a modified first cache line in a first cache memory. The snoopers determines if a second cache memory held a second cache line associated with the particular address prior to receiving the system-wide update complete operation. If so, the snoopers issues a prefetch request for a copy of the modified first cache line to replace the second cache line in the second cache memory. The snoopers updates the second cache memory with a copy of the modified first cache line.

七、指定代表圖：

(一)、本案指定代表圖為：第(1)圖。

(二)、本代表圖之元件代表符號簡單說明：

100 資料處理系統

102a 處理節點

104a PU

108b 系統記憶體

108a 系統記憶體

104b PU

110 系統互連

114 本機互連

104c PU

108d 系統記憶體

108c 系統記憶體

104d PU

102b 處理節點

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無

九、發明說明：

【發明所屬之技術領域】

本發明大體而言係關於資料處理系統之領域。特定而言，本發明係關於一種用於資料處理系統之快取同調協定。

【先前技術】

習知對稱多處理器（SMP）電腦系統（諸如一伺服器電腦系統）包括全部耦接至一系統互連之多個處理單元，該系統互連典型地包括一或多個位址、資料及控制匯流排。一系統記憶體耦接至該系統互連，該系統記憶體表示該多處理器電腦系統中最低級別之揮發性記憶體，且一般而言所有處理單元均可對其進行存取以用於讀取及寫入存取。為縮短對駐留於該系統記憶體中之指令及資料之存取潛時，每一處理單元典型地進一步受一相應多級快取階層支援，其該（等）下級可藉由一或多個處理器核心共享。

因為多個處理器核心可請求寫入存取同一資料快取線，及因為經修改之快取線不能立即與系統記憶體同步，多處理器電腦系統之快取階層典型地執行一快取記憶體同調協定，以在系統記憶體之內容之各種處理器核心的“檢視”之間確保至少有一最低級別之同調。特定而言，記憶體同調至少要求一處理單元存取一記憶體組塊之一複本及隨後存取該記憶體組塊之一更新之複本之後，該處理單元不能再次存取該記憶體組塊之舊的複本。

一快取記憶體同調協定典型地界定一組快取狀態以及一組同調訊息，其中該組快取狀態係與每一快取階層之快

取線關聯儲存，該組同調訊息用於在快取階層之間通信該快取狀態資訊。在常見具體實施例中，該快取狀態資訊採取吾人熟習之 MESI（修改、互斥、共享、無效）協定或其一變體之形式，且該同調訊息於一記憶體存取請求之該請求者及/或該等接收者之快取階層中指示一協定界定之同調狀態轉換。

在一多處理器電腦系統中，處理單元通常更新儲存於其本機快取中之快取線。當前，一旦第一處理單元更新在該本機快取中之一快取線，帶有該快取線複本之其他處理單元將其快取線複本標記為一需要更新之快取線。已經提出且執行某些協定，其中一快取控制器監視系統資料交易，並且，如果該快取線正好經過互連上之該快取控制器，則抓取該快取線之一更新之複本。

在當今之階層系統結構中，無任何保證該快取線將及時對所有感興趣的快取可見以避免潛時懲罰。其他協定涉及更新之後之該快取線之一全系統廣播，以便保證任何要求該快取線之新複本的快取可查看該更新之快取線。然而，此等協定效率很低，因為該更新之快取線之全系統廣播在互連上消耗不必要之頻寬。

因此，由於該先前技術之上述限制，需要一種有效地提供更新快取線至快取之系統及方法，其要求能不消耗在一資料處理系統中之不需要的頻寬的一快取線更新。

【發明內容】

本發明包括一種以主動式預取執行加強式停懸狀態之

系統及方法。根據本發明之一較佳具體實施例，在一處理單元中之一偵測器接收一系統級更新完成操作，指示完成針對一特定位址之儲存體修改操作，其中該儲存體修改操作導致在一第一快取記憶體中之經修改之第一快取線。該偵測器在接收該全系統更新完成操作之前確定一第二快取記憶體是否保持一與該特定位址關聯之第二快取線。若如此，該偵測器發出一對於該經修改之第一快取線之一複本之預取請求，以替換該第二快取記憶體中之第二快取線。該偵測器使用該經修改之第一快取線之一複本更新該第二快取記憶體。

該偵測器在接收該系統級更新完成操作之前，藉由存取一與該第二快取記憶體關聯之快取目錄確定該第二快取記憶體是否保持與該特定位址關聯之第二快取線，以確定該第二快取線是否處於一停懸同調狀態。

若該偵測器確定尚未接收該經修改之第一快取線之複本，則該偵測器重新發送該預取請求且遞增一計數器以確定發送之預取請求重試之次數。

在以下之詳細書面說明中，本發明之所有目標、功能，及優點將變得顯而易見。

【實施方式】

在該等所附申請專利範圍中闡述了作為本發明特性之該等新穎功能。然而，當結合讀取該所附圖式時，藉由參照以下對一描述性具體實施例之詳細說明，將最佳地瞭解本發明以及一較佳使用模式。

現參照該等圖式，特定而言，參照第 1 圖，其中圖解說明根據本發明之一快取同調對稱多處理器（SMP）資料處理系統之一示例性具體實施例之一高階方塊圖。如圖所示，資料處理系統 100 包括多個用於處理資料及指令之處理節點 102a、102b。處理節點 102a、102b 耦接至一用於傳輸位址、資料及控制資訊之系統互連 110。舉例而言，系統互連 110 可作為一匯流排式互連、一交換器式互連或一混合互連來執行。

在所描述之具體實施例中，每一處理節點 102 作為一包含四個處理單元 104a-104d 之多晶片模組（MCM）實現，每一處理單元較佳地作為一相應積體電路實現。在每一處理節點 102 內之該等處理單元 104a-104d 被耦接以用於一本機互連 114 之通信，與系統互連 110 類似，該本機互連 114 可使用一或多個匯流排及/或交換器執行。

耦接至每一本機互連 114 之該等裝置不僅包括處理單元 104，還包括一或多個系統記憶體 108a-108d。駐留於系統記憶體 108 中之資料及指令一般地可藉由在資料處理系統 100 之任何處理節點 102 中之任何處理單元 104 中之一處理器核心存取及修改。在本發明之替代性具體實施例中，一或多個系統記憶體 108 可耦接至系統互連 110 而非一本機互連 114。

熟習此項技術者將瞭解 SMP 資料處理系統 100 可包括許多額外的未說明的組件，諸如互連橋接器、非揮發性儲存、用於連接至網路或附加裝置之埠，等等。因為此等額

外組件對於瞭解本發明並不必要，在第 1 圖中未對其進行圖解說明或在此未對其進一步討論。然而，亦應瞭解，本發明提供之該等加強適用於不同架構之快取同調資料處理系統，且決不由在第 1 圖中圖解說明之該一般化資料處理系統架構所限制。

現參照第 2 圖，其中描述根據本發明之一示例性處理單元 104 之一更詳細方塊圖。在所描述之該具體實施例中，每一處理單元 104 包括兩個處理器核心 200a、200b 以用於獨立處理指令及資料。每一處理器核心 200 包括至少一用於為執行擷取及排序指令之指令定序單元 (ISU) 208，及一或多個用於執行指令之執行單元 224。由執行單元 224 執行之該等指令包括請求存取一記憶體組塊或致使產生一用於存取一記憶體組塊之請求之指令。

每一處理器核心 200 之操作由一多級揮發性記憶體階層支援，在多級揮發性記憶體階層之最低級別具有共享系統記憶體 108a-108d，且在其較高級別具有一或多個快取記憶體級。在所描述之具體實施例中，每一處理單元 104 包括一積體記憶體控制器 (IMC) 206，該積體記憶體控制器控制對在其處理節點 102 之範圍內該等系統記憶體 108a-108d 之一相應者之讀取及寫入存取，以回應自處理器核心 200a-200b 接收之請求及藉由一或多個偵測器 236 在該本機互連 114 上偵測之操作。

在該描述性具體實施例中，處理單元 104 之快取記憶體階層包括一在每一處理器核心 200 內之快取儲存級別一

(L1) 快取 226 及一由該處理單元 104 之所有處理器核心 200a、200b 共享之級別二 (L2) 快取 230。L2 快取 230 包括一 L2 陣列及目錄 234、一主機 232 及一偵測器 236。主機 232 在本機互連 114 及系統互連 110 上啟動交易並存取 L2 陣列及目錄 234，以回應自該等關聯處理器核心 200a-200b 接收之記憶體存取 (及其他) 請求。偵測器 236 偵測本機互連 114 上之操作，提供適當回應，及執行任何該等操作所要求之對 L2 陣列及目錄 234 之存取。

儘管所圖解說明之快取階層僅包括兩個級別之快取，熟習此項技術者將瞭解替代性具體實施例可包括額外級別 (L3、L4，等) 之晶片上或晶片外線內或前瞻 (lookaside) 快取，其可完整包含、部分地包含，或不包含該等上層快取之內容。

每一處理單元 104 更包括回應邏輯 210 之一實施例，其實施執行在資料處理系統 100 內維護快取同調之分散式同調訊號機制之一部分。此外，每一處理單元 104 包括轉發邏輯 212 之一實施例，用於在其本機互連 114 及系統互連 110 之間選擇性地轉發通信。最後，每一處理單元 104 包括一積體 I/O (輸入/輸出) 控制器 214，其支援附接諸如 I/O 裝置 216 之一或多個 I/O 裝置。如以下進一步之詳述，I/O 控制器 214 可在本機互連 114 及/或系統互連 110 上發出操作，以回應 I/O 裝置 216 之請求。

現參照第 3 圖，其中圖解說明 L2 陣列及目錄 234 之一示例性具體實施例之一更詳細方塊圖。如圖所示，L2 陣

列及目錄 234 包括 L2 快取陣列 300 之該等內容之一組結合 L2 快取陣列 300 及一 L2 快取目錄 302。如在習知組結合快取中，利用在該系統記憶體（真實）位址內之預定索引位元，將系統記憶體 108 中之記憶體位置映射至快取陣列 300 內之特定同餘類。儲存在快取陣列 300 內之該等特定快取線記錄於快取目錄 302 中，其中包含快取陣列 300 中之每一快取線之目錄項。如熟習此項技術者所瞭解的，快取目錄 302 中之每一目錄項包括至少一標籤欄 304、一狀態欄 306，及一 LRU（最近最少使用）欄 308，該標籤欄 304 利用該對應真實位址之一標籤部分，指定儲存於快取陣列 300 中之特定快取線，該狀態欄 306 指示該快取線之同調狀態，該 LRU 欄 308 指示該快取線相對在相同同餘類中其他快取線之一替換排序。

本發明較佳地執行一快取同調協定，該快取同調協定管理儲存於多個快取中之快取線之多個副本之間之衝突，並維護該等快取及系統記憶體之間之資料的一致性。如在本項技術中所熟知的，利用一 MESI 協定之系統使用以下四個同調狀態之一標記每一快取線：

M（修改）：該快取線標籤及資料有效且僅存在於當前快取中，且被視為“髒（dirty）”，或修改自儲存於主記憶體中之快取線之值。

E（互斥）：該快取線標籤及資料有效且僅存在於當前快取中，且係“乾淨（clean）”，或與儲存於主記憶體中之對應線之值相同。

S (共享)：該快取線標籤及資料有效且可被儲存於該系統之一或多其他快取中。

I (無效)：該快取線標籤及資料無效。

一快取可讀取一處於除“無效”外之任何狀態的快取線。一處於一“無效”狀態之快取線必須擷取自另一快取或主記憶體以滿足該讀取。如果該快取線處於一“修改”或“互斥”狀態，則可執行一寫入。在該先前技術中，如果該快取線處於該共享狀態，則所有其他快取複本必須首先無效化，這典型地經由一廣播操作執行。

在某些修改之 MESI 同調協定中，可存在一個以上類型之“M”、“S”及“I”狀態。舉例而言，一“M_e”狀態指示相對記憶體，該快取線被修改，且獨佔式地保持在當前快取中。其他修改之 MESI 協定包括一“標誌”或“T”狀態，其指示相對記憶體該快取線被修改，但亦由其他快取共享。

本發明使用該 MESI 協定之一變體，其包括一指示該快取線之位址標籤有效及該快取線資料無效之“停懸”(H)狀態。如果一快取線被標記為一“H”狀態，則保持該“H”狀態中之快取線之快取將嘗試重新獲取該快取線資料之一更新之複本。舉例而言，假定一資料處理系統包括三個與快取(例如，快取 0、1 及 2)關聯之處理器(例如，處理器 0、1 及 2)，每一處理器保持一該“S”狀態中之快取線之一複本。如果處理器 0 修改儲存於快取 0 中之快取線之本機複本，則儲存於快取 0 中之快取線之複本被

標記為“M”。儲存於該等快取 1 及 2 中之快取線之複本將被標記為“H”。根據本發明，一旦處理器 0 已完成修改儲存於快取 0 中之快取線之複本，且已通知該修改之處理器 1 及 2（例如，經由一“更新完成”全系統、僅位址廣播），處理器 1 及 2 將發出資料預取請求以重新獲取該快取線。一旦該等預取請求被滿足，儲存於快取 1 及 2 中之快取線將返回一共享（例如，“S”）狀態。當然，熟習此項技術者將瞭解本發明必須發送該更新之快取線之位址以通知所有其他快取，其具有該更新之快取線之一先前複本作為一有效性協定之一部分。包含該更新之快取線之實際資料不被廣播至該資料處理系統中之所有快取，以節省經由系統互連之頻寬。僅儲存於該更新快取線之一先前複本中之快取將預取該快取線之一更新複本，以回應接收該“更新完成”全系統、僅位址廣播。

第 4 圖係根據本發明之一具體實施例之一示例性方法之一高階邏輯流程圖，快取記憶體藉由該方法服務一自其關聯處理單元接收之儲存體修改之請求。該處理開始於步驟 400 並繼續到步驟 402，其說明 L2 快取 230 之主機 232 自任一其關聯處理器核心 200a-200b 接收諸如一儲存請求或無效請求之一儲存體修改操作請求。該處理繼續至步驟 403，其描繪主機 232，該主機 232 確定作為該儲存體修改請求之目標的該特定快取線之本機同調狀態。如果 L2 陣列 300 中該快取線之狀態欄 306（第 3 圖）中該同調狀態處於一“修改”或“M”狀態，則該處理繼續至步驟 404，

其展示在該快取線上執行該儲存體修改操作之主機 232。
其後該處理結束於步驟 416。

返回步驟 403，如果 L2 陣列 300 中該快取線之狀態欄 306 中該同調狀態處於一同調狀態，其中主機 232 在沒有啟動系統互連 110 上之操作下，不能服務該處理器請求。（例如，“共享”或“無效”），該處理繼續至步驟 408，其描繪主機 232 在系統互連 110 上發出一適當儲存體修改操作。該處理繼續至步驟 410，其展示確定該儲存體修改操作是否成功之主機 232（舉例而言，藉由參照一全系統組合回應）。若該操作不成功，該處理返回步驟 408。

若該操作成功，該處理繼續至步驟 412，其說明主機 232 無效或更新 L2 陣列 300 中其目標快取線之本機複本。該快取線無效或更新之後，該處理繼續至步驟 414，其描繪主機 232 廣播一全系統、僅位址“更新完成”操作，以便指示該儲存體修改操作完成及該快取線可供在資料處理系統 100 內之任何其他處理單元預取，該處理單元在其 L2 快取 230 中保持緊接該接收該儲存體修改操作之前該快取線之一有效複本。然後該處理結束於步驟 416。

第 5A 圖係根據本發明之一具體實施例說明一示例性方法之一高階邏輯流程圖，一快取記憶體藉由該方法處理在一系統互連上偵測之儲存體修改操作。該處理開始於步驟 500，以回應藉由一偵測器 236 接收的系統互連 110 上偵測之操作。然後該處理繼續至步驟 501，其描繪偵測器 236 確定在系統互連 110 上偵測之操作是否針對一儲存於

關聯的 L2 陣列 300 (例如，快取命中) 中之有效快取線。若該操作不針對一儲存於 L2 陣列 300 中之快取線，則該處理於步驟 514 結束。

若該操作針對一儲存於 L2 陣列 300 (例如，快取命中) 中之有效快取線，則該處理繼續至步驟 502，其說明偵測器 236 確定該操作是否係一儲存體修改操作 (例如，一儲存請求或無效請求)。若偵測器 236 確定該操作不是一儲存體修改操作，則該處理繼續到步驟 503，其說明偵測器 236 確定該偵測操作是否係一“更新完成”操作之一全系統、僅位址廣播。若該偵測操作係一“更新完成”操作之一全系統、僅位址廣播，則該處理通過頁連接器“A”至如以下之討論之第 5B 圖。若該偵測操作不係一“更新完成”操作之全系統、僅位址廣播，該處理繼續到步驟 504，其描繪偵測器 236 執行該偵測操作之其他處理。然後該處理結束於步驟 514。

返回步驟 502，若偵測器 502 確定該偵測操作係一儲存體修改操作，則該處理繼續至任選步驟 506，其展示確定該快取線之該本機同調狀態之偵測器 236。步驟 506 係一任選步驟，因為每一偵測之儲存體修改操作自動導致一資料預取請求，可能不甚理想。該預取請求消耗所需系統互連頻寬，且可能因此被限制於具有特定同調狀態之快取線。舉例而言，在不久之將來可能更有可能存取一具有一本機同調狀態“M”之快取線，因為與該本機快取關聯之處理單元最近已修改該快取線。一具有一本機同調狀態

“S”之快取線被再次存取之可能較小。

若該快取線之該本機複本之本機同調狀態為“共享”，則該處理繼續至步驟 512，其展示將該快取線之同調狀態更新為“無效”之偵測器 236。然後該處理結束於步驟 514。

返回步驟 506，若該本機同調狀態為“修改”，則該處理繼續至步驟 520，其描繪偵測器 236 確定是否需要偵測器 236 將該快取線提供至一發出該偵測操作之請求處理器。若需要偵測器 236 將該快取線提供至該請求處理器，則該處理繼續至步驟 518，其說明偵測器 236 藉由快取至快取介入將該快取線發送至該請求處理器。然後該處理繼續至步驟 522，其展示偵測器 236 將該快取線之同調狀態更新為“停懸”。然而，若無需偵測器 236 將該快取線提供至該請求處理單元，則該處理直接繼續到步驟 522。然後該處理結束於步驟 514。熟習此項技術者將瞭解步驟 506 係一任選步驟，因為本發明包括偵測器 236 將該快取線之同調狀態更新為“停懸”，無論該快取線之先前同調狀態如何（例如，“M”、“S”、“I”，等）。

第 5B 圖係根據本發明之一具體實施例之一示例性方法之一高階邏輯流程圖，快取記憶體藉由該方法處理一在該系統互連上偵測之“更新完成”操作。該處理開始於頁連接器“A”，以回應在第 5A 圖步驟 503 確定藉由偵測器 236 偵測之操作，該操作係一“更新完成”操作之一僅位址、全系統廣播。該“更新完成”操作指示一作為一儲存

體修改操作之目標之快取線。

該處理繼續至步驟 524，其中若該關聯快取目錄 302 將該“更新完成”操作之目標位址與該“停懸”狀態關聯，則該處理繼續至步驟 526 及 528，其描繪偵測器 236 產生及發送一資料預取請求，以利用自該“更新完成”操作接收之位址擷取該快取線之一新近更新之複本，如在步驟 526 中所示。預取該快取線之一新近更新之複本減少了在處理單元 104a-104n 內之一管線延遲之機會。

事實上，儲存於 L2 快取目錄 302 中之快取線指示該關聯處理單元最近已參考該快取線且將可能再次參考該快取線。因為該處理單元將可能在不久之將來參考該快取線，在預期未來參考之情況下，預取該快取線之一新近更新之複本將減少一管線延遲之機會，因為當該處理單元要求該快取線用於處理時，該處理單元不得不等待擷取該新近更新之快取線（因此延遲該管線）之可能性將更小。

在步驟 528 之後，該處理繼續至步驟 530，其說明偵測器 236 確定是否已接收該請求的新近更新之快取線，以回應在步驟 528 中討論之預取請求。在本發明之一具體實施例中，偵測器 236 重試該預取請求，若需要，可多達 N 次，如步驟 534 所示。

返回步驟 530，若接收到該請求的新近更新之快取線，則偵測器 236 將該快取線更新為該正確的同調狀態（例如，“共享”）。然後該處理結束於步驟 538。

正如所討論的，本發明包括以主動式預取執行加強式

停懸狀態之系統及方法。根據本發明之一較佳具體實施例，一處理單元中之偵測器接收一全系統更新完成操作，指示完成針對一特定位址之儲存體修改操作，其中該儲存體修改操作導致在一第一快取記憶體中之經修改之第一快取線。該偵測器確定在接收該全系統更新完成操作之前，一第二快取記憶體是否保持一與該特定位址關聯之第二快取線。若如此，該偵測器發出一對於該經修改之第一快取線之一複本之預取請求，以替換該第二快取記憶體中之第二快取線。該偵測器使用該經修改之第一快取線之一複本更新該第二快取記憶體。

該偵測器藉由存取一與該第二快取記憶體關聯之快取目錄確定在接收該全系統更新完成操作之前，該第二快取記憶體是否保持與該特定位址關聯之該第二快取線，以便確定該第二快取線是否處於一停懸同調狀態。

熟習此項技術者將瞭解是否產生一預取請求以回應完成一儲存體修改操作，可依賴於在該系統互連上偵測之操作類型或交易類型。

儘管已參照一較佳具體實施例特別展示及說明本發明，熟習此項技術者將瞭解在不偏離本發明之精神及範疇內可對其形式及細節進行各種變更。

【圖式簡單說明】

第1圖係根據本發明之一示例性資料處理系統之一高階方塊圖；

第 2 圖係根據本發明之一處理單元之一更詳細方塊圖；

第 3 圖係在第 2 圖中說明之該 L2 快取陣列及目錄之一更詳細方塊圖；及

第 4 圖係根據本發明之一具體實施例之一示例性方法之一高階邏輯流程圖，快取記憶體藉由該方法服務一自其關聯之處理器核心接收之請求；

第 5A 圖係根據本發明之一具體實施例之一示例性方法之一高階邏輯流程圖，快取記憶體藉由該方法處理在該系統互連上偵測之一儲存體修改操作；及

第 5B 圖係根據本發明之一具體實施例之一方法之一高階邏輯流程圖，快取記憶體藉由該方法處理在該系統互連上偵測之一“更新完成”操作。

【主要元件符號說明】

100	資料處理系統
102a	處理節點
104a	PU
108b	系統記憶體
108a	系統記憶體
104b	PU
110	系統互連
114	本機互連
104c	PU
108d	系統記憶體

108c	系統記憶體
104d	PU
102b	處理節點
104	處理單元
200a	處理器核心
208	指令定序單元
224	執行單元
226	L1 快取
232	主控
234	L2 陣列及目錄
236	偵測器
230	L2 快取
210	回應邏輯
212	轉發邏輯
206	IMC
222	S
214	i/o 控制器
216	i/o 裝置
302	L2 快取目錄
304	標籤欄
306	狀態欄
308	LRU 欄
300	L2 快取陣列

十、申請專利範圍：

1. 一種用於為一資料處理系統實施一快取同調協定之方法，該資料處理系統具有經由一位址及資料互連進行通信之多個處理單元，該方法包括以下步驟：

經由該互連接收一全系統更新完成訊號，該訊號指示完成針對一特定位址之一儲存體修改操作，其中該儲存體修改操作在該多個處理單元中之一第一處理單元之一第一快取記憶體中導致一經修改之第一快取線；

在接收該全系統更新完成訊號之前，確定在該多個處理單元中之一第二處理單元之一第二快取記憶體是否保持一與該特定位址關聯之第二快取線；

其中，該確定步驟更包含存取一與該第二快取記憶體相關之快取目錄，以確定該第二快取線是否處於一停懸同調(H)狀態，其中該(H)狀態指示用於該第二快取線之該位址標籤係為有效，且該(H)狀態指示在該第二快取線中之該快取線資料目前係為無效；

其中，當一快取線經標記為一(H)狀態，則保持在該(H)狀態中之該快取線之該快取嘗試重新獲取來自另一快取之另一快取線之該快取線資料之一已更新複本，該另一快取保持在一修改(M)狀態中之該快取線；回應在接收該全系統更新完成操作之前確定該第二快取記憶體保持一與該特定位址關聯之第二快取線，在與該第二快取線之該特定位址關聯的資料被與該第二快取記憶體關聯之該處理單元請求之前，發送一對於該經修改之第一快取線之一

複本之預取請求，以替換在該第二快取記憶體中之該第二快取線；及

當該預取請求歸還(return)該經修改之第一快取線之該複本時，利用該經修改之第一快取線之該複本來更新在該第二快取記憶體內之該第二快取線。

2. 如申請專利範圍第1項所述之方法，更包括以下步驟：

確定該第二快取線之一本機同調狀態；

回應確定該第二快取線之該本機同調狀態為共享，將該第二快取線之該本機同調狀態更新為無效。

3. 如申請專利範圍第1項所述之方法，更包括以下步驟：

回應確定尚未接收該經修改之第一快取線之該複本，重新發送該預取請求。

4. 如申請專利範圍第3項所述之方法，更包括以下步驟：

回應一預取請求重試之次數達到一預定數字，停止產生預取請求，並將該第二快取線之該本機同調狀態改變為無效。

5. 一種用於為一資料處理系統實施一快取同調協定之方法，該資料處理系統具有經由一位址及資料互連進行通信之多個處理單元，該方法包括以下步驟：

經由該互連接收一全系統更新完成訊號，該全系統更

新完成訊號指示完成針對一特定位址之一儲存體修改操作，其中該儲存體修改操作在該多個處理單元內之一第一處理單元之一第一快取記憶體中導致一經修改之第一快取線；

在接收該全系統更新完成訊號之前，確定在該多個處理單元內之一第二處理單元之一第二快取記憶體是否保持一與該特定位址關聯之第二快取線；

回應在接收該全系統更新完成操作之前確定該第二快取記憶體保持一與該特定位址關聯之第二快取線，發送一針對該經修改之第一快取線之一複本之預取請求，以替換在該第二快取記憶體中之該第二快取線；

當該預取請求歸還該經修改之第一快取線之該複本時，使用該經修改之第一快取線之該複本來更新在該第二快取記憶體內之該第二快取線；

確定第二快取線之一本機同調狀態；

回應確定該第二快取線之該本機同調狀態經修改，確定該第二快取線是否需要被調用(source)；

回應確定該第二快取線需要被調用，傳送該第二快取線至一請求處理單元，並將該第二快取線之該本機同調狀態更新為停懸；及

回應確定該第二快取線不需要被調用，將該第二快取線之該本機同調狀態更新為停懸。

6. 一種用於在一資料處理系統中實施一快取同調協定之

處理單元，該資料處理系統具有經由一位址及資料互連通信之多個處理單元，該處理單元包括：

至少一處理器核心；

一本機快取陣列，該本機快取陣列係耦接至該至少一處理器核心，以用於快取資料以供該至少一處理器核心存取；

至少一偵測器，該偵測器係用於：

接收一全系統更新完成訊號，該訊號指示完成一針對一特定位址之儲存體修改操作，其中該儲存體修改操作在該多個處理單元中導致與一第一處理單元關聯之一第一快取陣列中之經修改之第一快取線；

在接收該全系統更新完成訊號之前，確定該本機快取陣列是否保持一與該特定位址關聯之第二快取線；

其中，該偵測器存取該快取目錄以確定該第二快取線是否位於一停懸同調(H)狀態，其中該(H)狀態指示該第二快取線之該位址標籤係為有效及在該第二快取線中之該快取線資料目前係為無效；

其中，當一快取線被標記為一(H)狀態時，則保持在該(H)狀態中之該快取線的該快取嘗試重新獲取來自另一快取之另一快取線之該快取線資料的一已更新複本，該另一快取保持在一修改(M)狀態中之該快取線；回應在接收該全系統更新完成訊號之前確定該本機快取陣列保持一與該特定位址關聯之第二快取線，並在與該第二快取線之該特定位址關聯的資料被與該本機

快取陣列關聯之該處理單元請求之前，發送一對於該經修改之第一快取線之一複本之預取請求，以替換在該本機快取陣列中之該第二快取線；及

當該預取請求歸還該經修改之第一快取線之該複本時，利用該經修改之第一快取線之該複本，以更新該本機快取陣列之該第二快取線。

7. 一資料處理系統，該資料處理系統包括：

一系統互連；

一系統記憶體，該系統記憶體係耦接至該系統互連；

及

至少一如申請專利範圍第6項所述之處理單元，其中該至少一處理單元被耦接至該系統互連。

8. 如申請專利範圍第6項所述之處理單元，更包括：

一與該本機快取陣列關聯之快取目錄，以用於將一同調狀態與儲存於該本機快取陣列中之快取線關聯。

9. 如申請專利範圍第6項所述之處理單元，其中該偵測器將該第二快取線之一本機同調狀態更新為無效，以回應確定該第二快取線之該本機同調狀態為共享。

10. 如申請專利範圍第6項所述之處理單元，其中該偵測器確定是否需要調用該第二快取線，以回應確定該第二快取

線之一本機同調狀態被修改。

11. 一種用於在一資料處理系統中實施一快取同調協定之處理單元，該資料處理系統具有經由一位址及資料互連通信之多個處理單元，該處理單元包括：

至少一處理器核心；

一本機快取陣列，該本機快取陣列係耦接至該至少一處理器核心，以用於快取資料以供該至少一處理器核心存取；

至少一偵測器，該偵測器係用於：

接收一全系統更新完成訊號，該全系統更新完成訊號指示完成一針對一特定位址之儲存體修改操作，其中該儲存體修改操作在該多個處理單元中導致與一第一處理單元關聯之一第一快取陣列中之一經修改之第一快取線；

在接收該全系統更新完成訊號之前，確定該本機快取陣列是否保持與該特定位址相關聯之一第二快取線；

回應在接收該全系統更新完成訊號之前確定該本機快取陣列保持與該特定位址相關聯之一第二快取線，發送一針對該經修改之第一快取線之一複本的預取請求，以替換在該本機快取陣列中之該第二快取線；及

當該預取請求歸還該經修改之第一快取線之該複本，則利用該經修改之第一快取線之該複本來更新該本機快取陣列之該第二快取線；

其中，該偵測器確定該第二快取線是否需要被調用，以回應確定該第二快取線之一本機同調狀態係經修改；

其中，該偵測器傳送該第二快取線至一請求處理單元，並將該第二快取線之該本機同調狀態更新為停懸，以回應確定該第二快取線需要被調用。

12. 如申請專利範圍第6項所述之處理單元，其中該偵測器重新發送該預取請求，以回應確定尚未接收該經修改之第一快取線之該複本。

13. 如申請專利範圍第12項所述之處理單元，其中該偵測器停止產生預取請求，以回應一預取請求重試之次數達到一預定數字。

14. 一種用於在一資料處理系統中實施一快取同調協定之處理單元，該資料處理系統具有經由一位址及資料互連通信之多個處理單元，該處理單元包括：

至少一處理器核心；

一本機快取陣列，該本機快取陣列係耦接至該至少一處理器核心，以用於快取資料以供該至少一處理器核心存取；

至少一偵測器，該至少一偵測器係用於：

接收一全系統更新完成訊號，該全系統更新完成訊號指示完成一針對一特定位址之儲存體修改操作，其中

該儲存體修改操作在該多個處理單元中導致與一第一處理單元相關聯之一第一快取陣列中的一經修改之第一快取線；

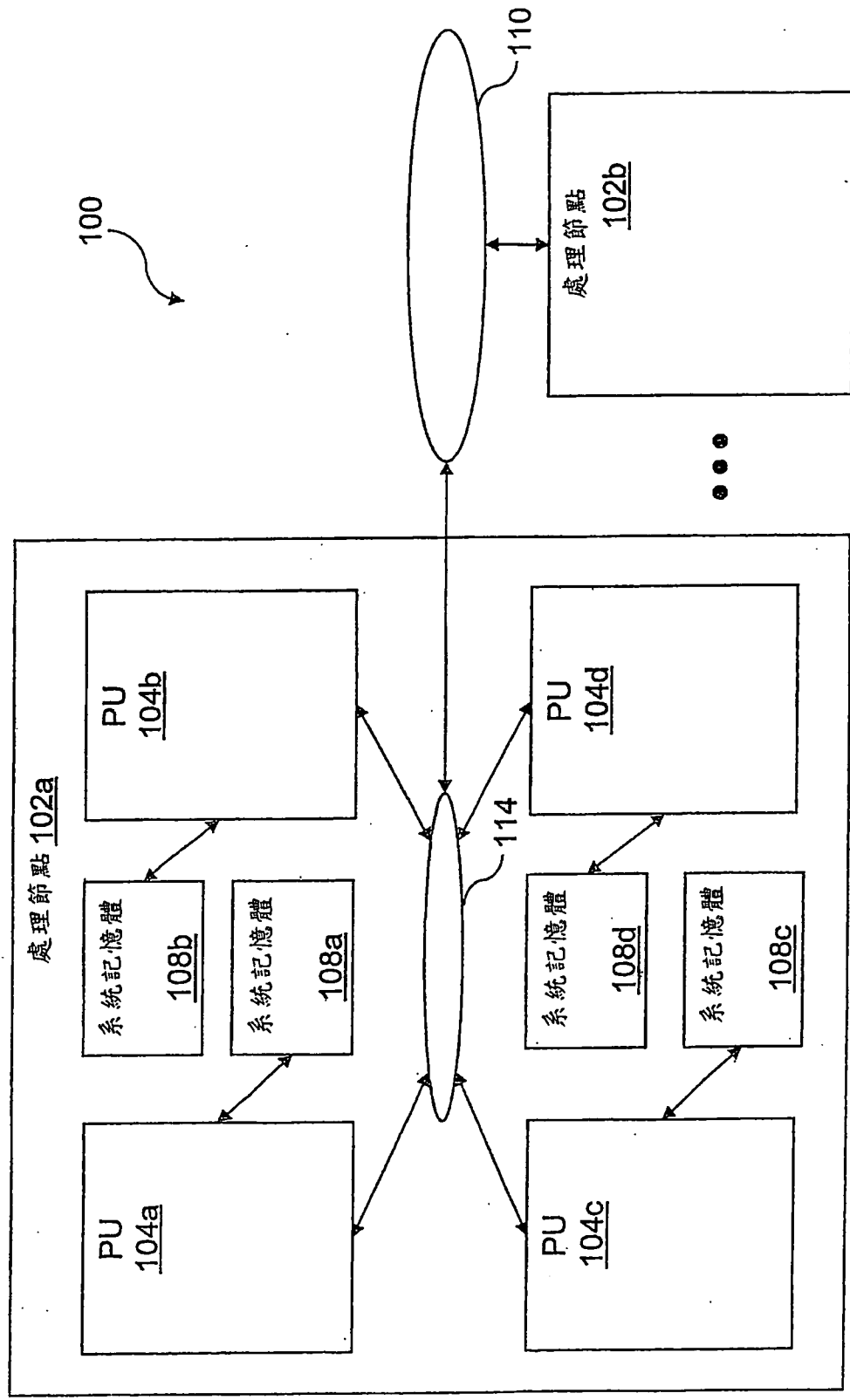
在接收該全系統更新完成訊號之前，確定該本機快取陣列是否保持與該特定位址相關聯之一第二快取線；

回應在接收該全系統更新完成訊號之前確定該本機快取陣列保持與該特定位址相關之一第二快取線，發送一針對該經修改之第一快取線之一複本的預取請求，以替換在該本機快取陣列中之該第二快取線；及

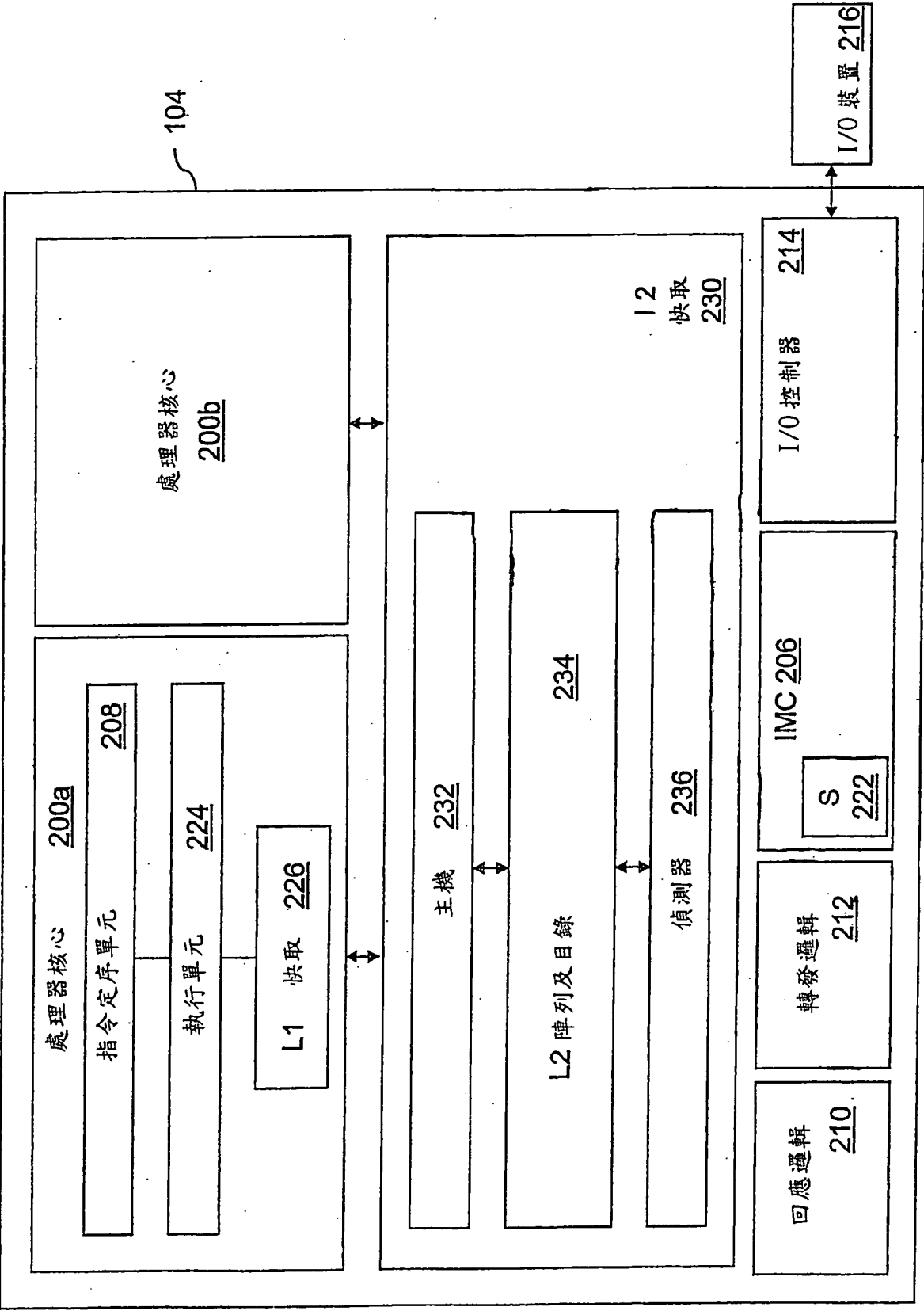
當該預取請求歸還該經修改之第一快取線之該複本，則利用該經修改之第一快取線之該複本來更新該本機快取陣列之該第二快取線；

其中，該偵測器確定該第二快取線是否需要被調用，以回應確定該第二快取線之一本機同調狀態係經修改；

其中，該偵測器將該第二快取線之該本機同調狀態更新為停懸，以回應確定該第二快取線不需要被調用。



第 1 圖

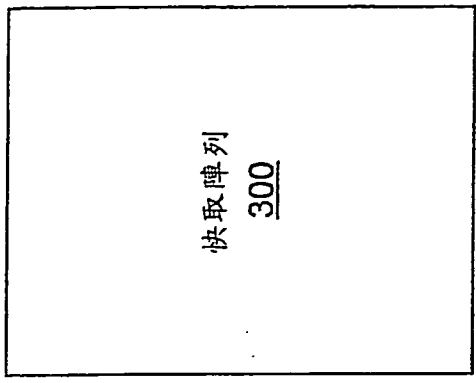


第 2 圖

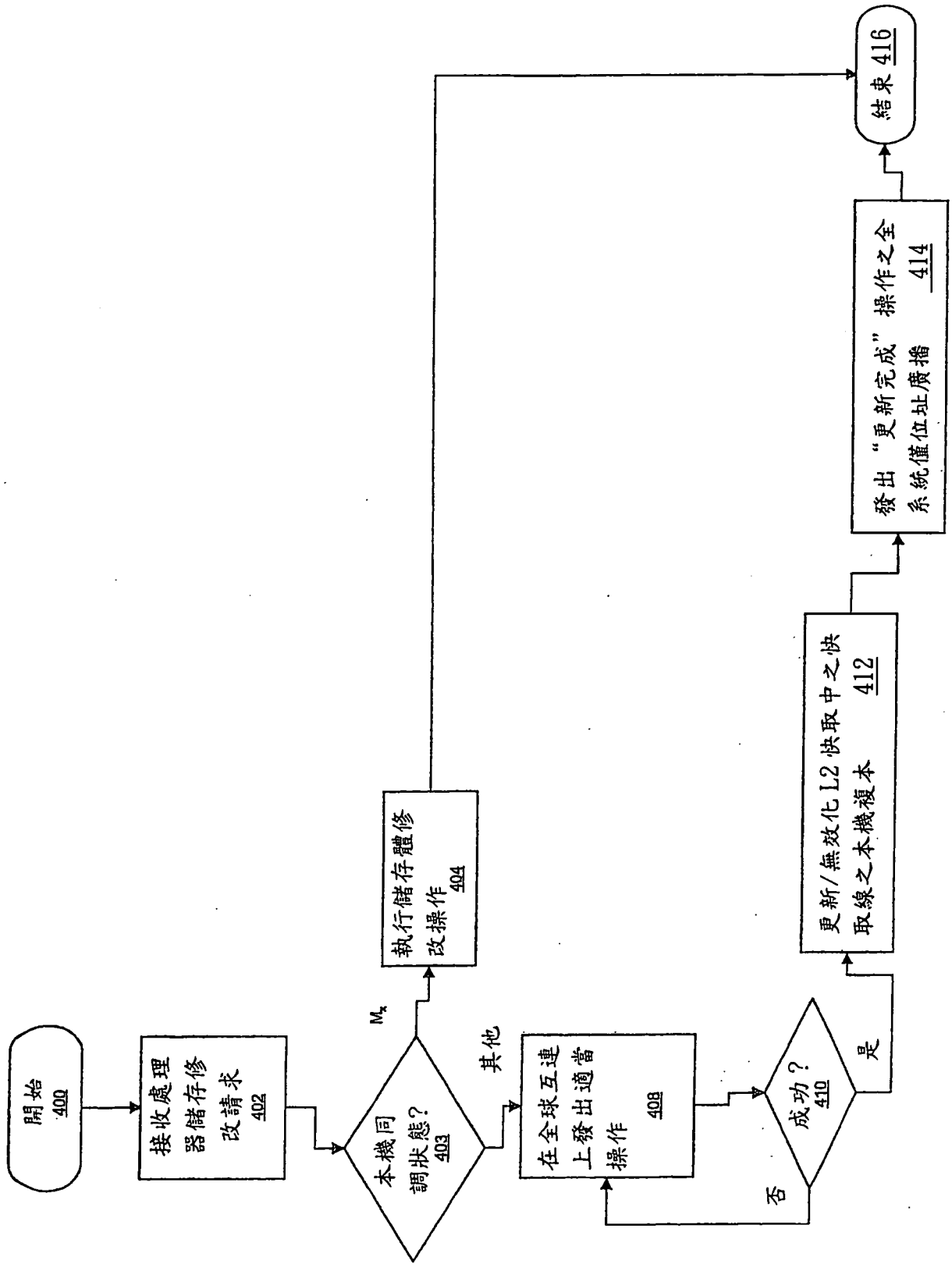
234

L2 快取
目錄 302

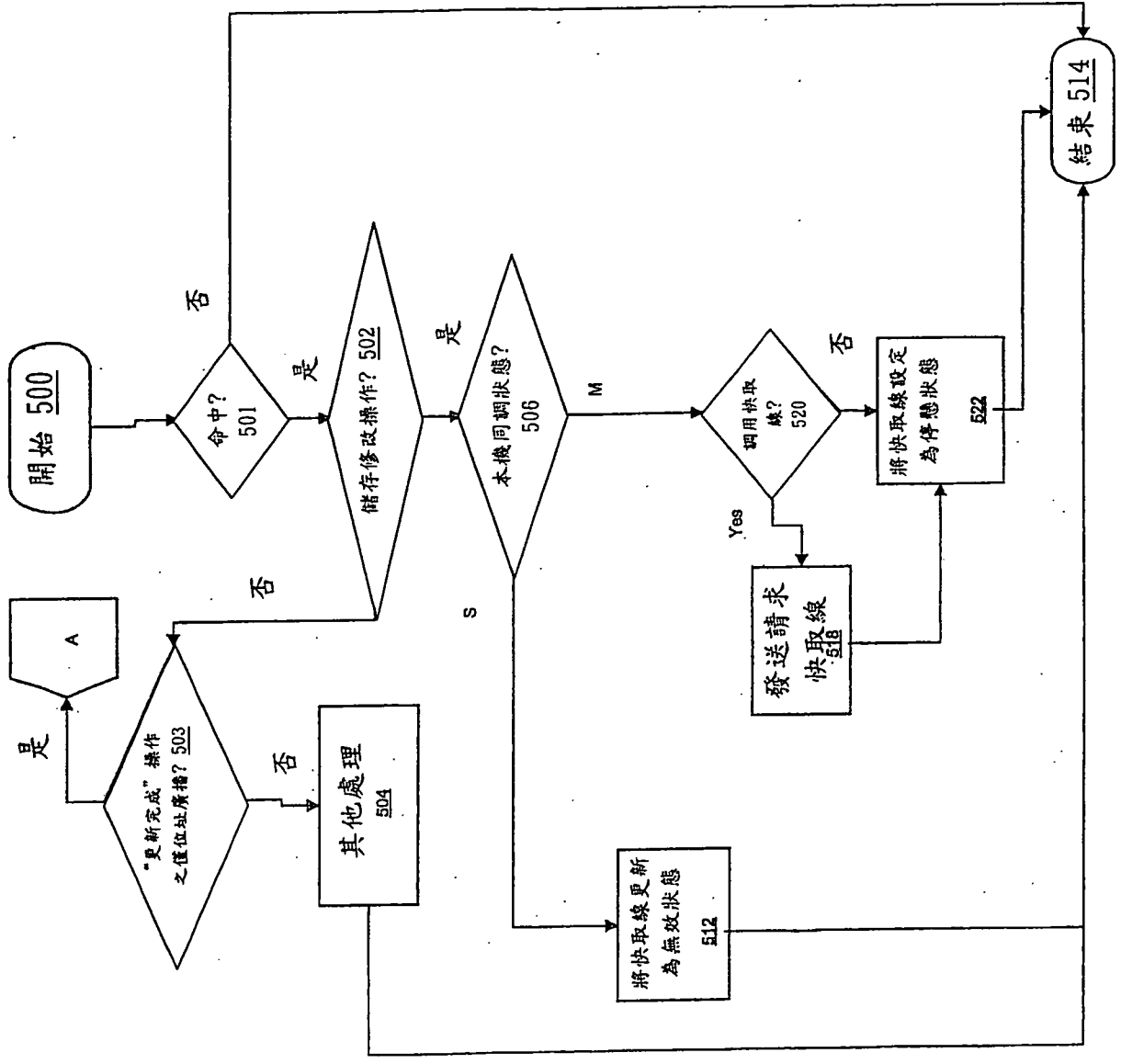
標籤欄 304	狀態欄 306	LRU 欄 308



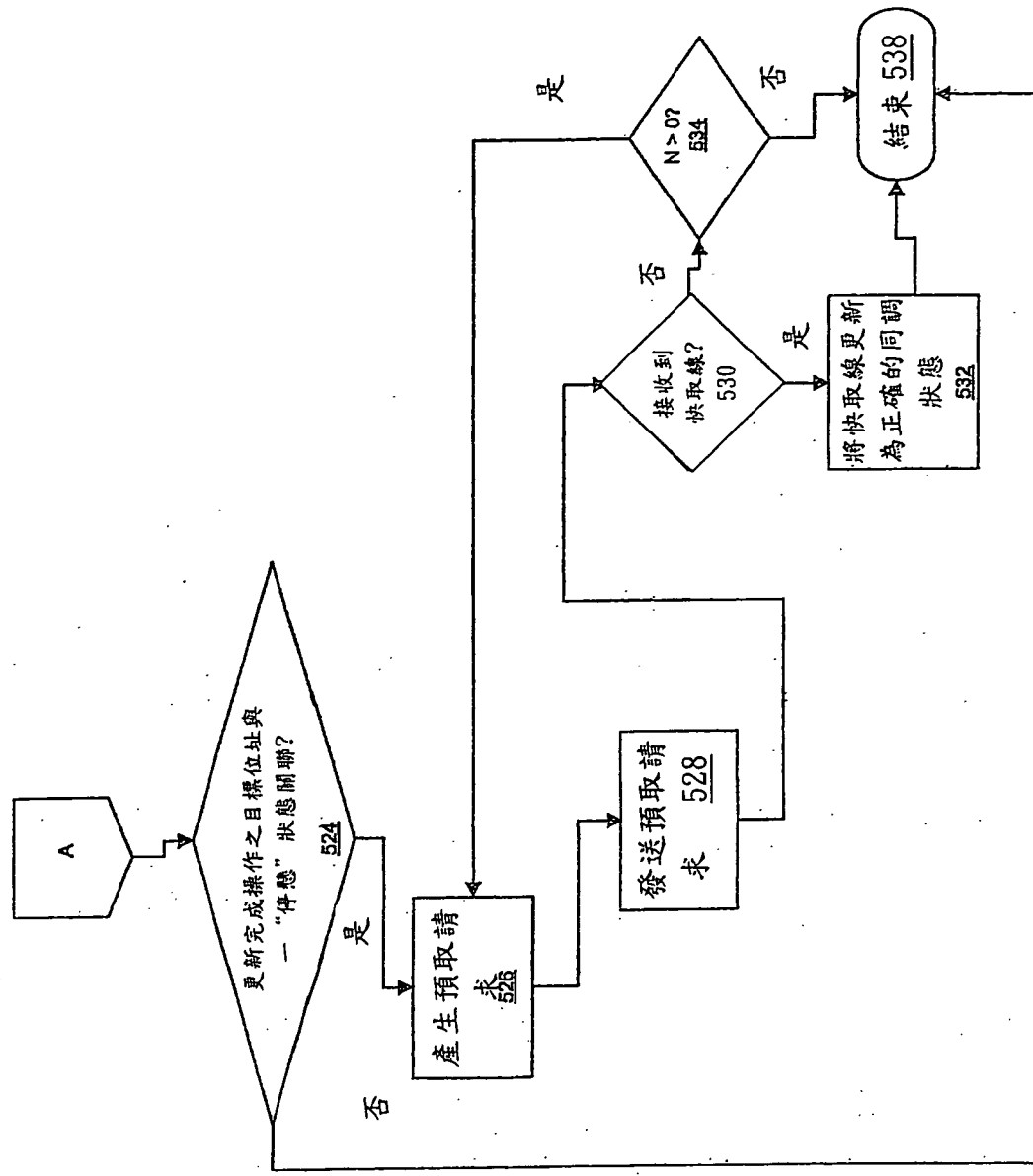
第 3 圖



第 4 圖



第 5A 圖



第 5B 圖