



(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 I706247 B

(45)公告日：中華民國 109 (2020) 年 10 月 01 日

(21)申請案號：104114781

(22)申請日：中華民國 104 (2015) 年 05 月 08 日

(51)Int. Cl. : G06F12/02 (2006.01)

G11C5/02 (2006.01)

G11C8/12 (2006.01)

(30)優先權：2014/05/08 美國

61/990,199

(71)申請人：美商美光科技公司 (美國) MICRON TECHNOLOGY, INC. (US)

美國

(72)發明人：墨菲 理查 C MURPHY, RICHARD C. (US)

(74)代理人：陳長文

(56)參考文獻：

TW 201411642A

TW 201414392A

CN 102157187A

US 2013/0138892A1

Gwangsun Kim, John Kim, <International Conference on Parallel Architectures and Compilation Techniques>, "Memory-centric System Interconnect Design with Hybrid Memory Cubes", 2013, <http://bt.nitk.ac.in/c/16a/co471/notes/Memory-centric%20System%20Interconnect%20Design%20with%20Hybrid%20Memory%20Cubes.pdf>

審查人員：吳鴻鎮

申請專利範圍項數：19 項 圖式數：7 共 28 頁

(54)名稱

記憶體中輕量一致性

(57)摘要

一種系統包含複數個主機處理器及經組態為用於該等主機處理器之一分散式共用記憶體之複數個混合記憶體立方體(HMC)裝置。一 HMC 裝置包含：複數個積體電路記憶體晶粒，其包含配置於一第二記憶體晶粒之頂部上之至少一第一記憶體晶粒，且該記憶體晶粒之記憶體之至少一部分經映射以包含一記憶體一致性目錄之至少一部分；及一邏輯基底晶粒，其包含：至少一個記憶體控制器，其經組態以管理至少一個第二裝置對該複數個記憶體晶粒之記憶體之三維(3D)存取；及邏輯電路，其經組態以判定儲存於該複數個記憶體晶粒之該記憶體中之資料之記憶體一致性狀態資訊、傳遞關於對記憶體之該存取之資訊且將該記憶體一致性資訊包含於該所傳遞資訊中。

A system includes a plurality of host processors and a plurality of HMC devices configured as a distributed shared memory for the host processors. An HMC device includes a plurality of integrated circuit memory die including at least a first memory die arranged on top of a second memory die and at least a portion of the memory of the memory die is mapped to include at least a portion of a memory coherence directory; and a logic base die including at least one memory controller configured to manage three-dimensional (3D) access to memory of the plurality of memory die by at least one second device, and logic circuitry configured to determine memory coherence state information for data stored in the memory of the plurality of memory die, communicate information regarding the access to memory, and include the memory coherence information in the communicated information.

指定代表圖：

符號簡單說明：

- 205 . . . 混合記憶體立方體
- 215 . . . 邏輯基底層
- 220 . . . 儲存庫/儲存庫儲存單元
- 225 . . . 儲存庫邏輯區塊

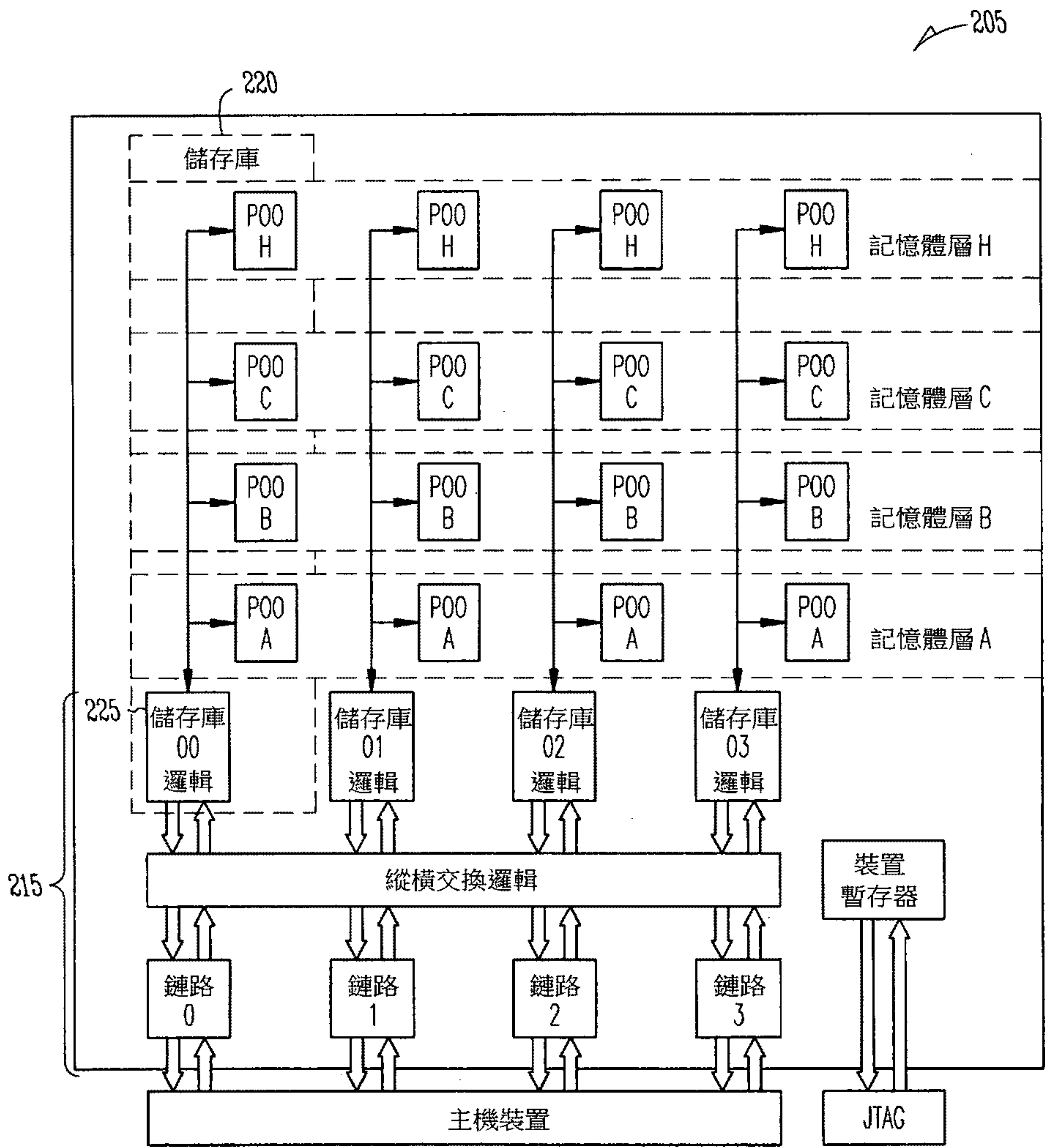


圖 2

發明摘要

※ 申請案號：104114781

※ 申請日：104年5月8日

※IPC 分類：G06F 12/02 (2006.01)
G11C 5/02 (2006.01)
G11C 8/12 (2006.01)

【發明名稱】

記憶體中輕量一致性

IN-MEMORY LIGHTWEIGHT COHERENCY

【中文】

一種系統包含複數個主機處理器及經組態為用於該等主機處理器之一分散式共用記憶體之複數個混合記憶體立方體(HMC)裝置。一HMC裝置包含：複數個積體電路記憶體晶粒，其包含配置於一第二記憶體晶粒之頂部上之至少一第一記憶體晶粒，且該記憶體晶粒之記憶體之至少一部分經映射以包含一記憶體一致性目錄之至少一部分；及一邏輯基底晶粒，其包含：至少一個記憶體控制器，其經組態以管理至少一個第二裝置對該複數個記憶體晶粒之記憶體之三維(3D)存取；及邏輯電路，其經組態以判定儲存於該複數個記憶體晶粒之該記憶體中之資料之記憶體一致性狀態資訊、傳遞關於對記憶體之該存取之資訊且將該記憶體一致性資訊包含於該所傳遞資訊中。

【英文】

A system includes a plurality of host processors and a plurality of HMC devices configured as a distributed shared memory for the host processors. An HMC device includes a plurality of integrated circuit memory die including at least a first memory die arranged on top of a second memory die and at least a portion of the memory of the memory die is mapped to include at least a portion of a memory coherence directory; and a logic base die including at least one memory controller configured to manage three-dimensional (3D) access to memory of the plurality of memory die by at least one second device, and logic circuitry configured to determine memory coherence state information for data stored in the memory of the plurality of memory die, communicate information regarding the access to memory, and include the memory coherence information in the communicated information.

【代表圖】

【本案指定代表圖】：第（2）圖。

【本代表圖之符號簡單說明】：

- 205 混合記憶體立方體
- 215 邏輯基底層
- 220 儲存庫/儲存庫儲存單元
- 225 儲存庫邏輯區塊

【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】

記憶體中輕量一致性

IN-MEMORY LIGHTWEIGHT COHERENCY

【先前技術】

多處理器系統可包含允許一個以上處理器對相同資料執行操作之共用記憶體。對共用記憶體之存取可為緩慢的且可導致系統延時。為改良記憶體存取時間，多處理器共用記憶體系統可包含一相對較小但較快之存取快取記憶體。快取記憶體儲存可由處理器頻繁存取之資料之一複本。為維持經快取資料之正確性，後續接著一快取一致性協定。一快取一致性協定對藉由處理器之對經快取資料之存取施加規則以確保資料跨越所有處理器係正確的，儘管存在資料之多個複本。然而，快取一致性協定可消耗先前系統頻寬且產生資料存取之不均勻延時。本發明人已認識到需要多處理器共用記憶體系統中之經改良資料管理。

【圖式簡單說明】

圖1圖解說明可用以實踐本發明之某些實施例之一混合記憶體立方體之一表示。

圖2係圖解說明可用以實踐本發明之某些實施例之一混合記憶體立方體之一四分之一之一方塊圖。

圖3展示可用以實踐本發明之某些實施例之混合記憶體立方體之四個可能裝置拓撲之方塊圖。

圖4展示可用以實踐本發明之某些實施例之一種一致性目錄記憶體格式之一實例。

圖5展示可用以實踐本發明之某些實施例之用HMC裝置實施之一電腦系統之一實例之一方塊圖。

圖6展示操作一電腦系統之一方法之一實例之一流程圖。

圖7展示操作一電腦系統之一方法之另一實例之一流程圖。

【實施方式】

在本發明之實例性實施例之以下詳細說明中，藉由圖式及圖解說明方式而參考本發明之特定實例性實施例。充分詳細地闡述此等實例以使得熟習此項技術者能夠實踐本發明，且該等實例用於圖解說明本發明可如何應用於各種目的或實施例。存在本發明之其他實施例且該等實施例在本發明之範疇內，且可在不背離本發明之標的或範疇之情況下做出邏輯、機械、電及其他改變。本文中所闡述之本發明之各種實施例之特徵或限制雖然對其中併入該等特徵或限制之實例性實施例係必要的，但其並不限制本發明之其他實施例或作為一整體之本發明，且對本發明之任何參考(其元件、操作及應用)並不限制作為一整體之本發明，而是僅用於定義此等實例性實施例。因此，以下詳細說明並不限制僅由隨附申請專利範圍界定之本發明之各種實施例之範疇。

在習用電腦系統中，記憶體以二維方式組織成列及行。為改良記憶體存取延時做出之努力已包含開發具有一高頻寬介面之雙資料速率類型三(DDR3)或DDR4。DDR3及DDR4使用傳統列/行方法。

圖1圖解說明一混合記憶體立方體105 (HMC)之一表示。一HMC藉由堆疊積體電路(IC)記憶體晶粒110且使用達成在IC晶粒之頂部及底部兩者上之存取之穿矽導通體(TSV)結構(未展示)來使晶粒互連而包含一第三維度。TSV製造技術達成多個晶粒層之互連，以便構造三維晶粒。使多個晶粒層互連之此能力准許用記憶體儲存層與一或多個邏輯層之一組合來建造一記憶體裝置。以此方式，該裝置在一單個封

裝中提供實體記憶體儲存及邏輯記憶體異動處理。圖1中所展示之配置係用以圖解說明一概念且未必表示一實際晶粒佈局。

一HMC (例如，如混合記憶體立方體規範版本1.0中所規定)可提供具有多達每裝置320 GB/s之可用頻寬容量之一極精巧且功率高效封裝。HMC裝置經由一階層式及平行設計方法而能夠具有高頻寬。裝置階層跨越邏輯層而垂直地發生且硬體平行跨越一給定晶粒層而發生。在一HMC內，記憶體組織成儲存庫。每一儲存庫120使用穿矽導通體在晶粒內垂直地橫跨記憶體層中之每一者。在圖1之經簡化實例中，HMC組織成16個儲存庫，其中每一儲存庫具有一邏輯基底晶粒層115或邏輯基底。

圖2係圖解說明一HMC 205之一四分之一之一方塊圖。展示HMC之四個儲存庫，其中每一儲存庫220包含多個記憶體層。邏輯基底層215包含多個組件。舉例而言，每一儲存庫在邏輯基底中具有管理用於儲存庫之所有記憶體參考操作之一記憶體控制器。記憶體控制器提供對3D儲存層之靈活存取。3D分層方法准許記憶體異動不僅跨越一目標儲存陣列內之記憶體庫，且亦跨越平行儲存陣列平行地存在。邏輯基底層215可包含用以實施所闡述功能之一或多個處理器，且一HMC可為一記憶體中處理器(PIM)裝置。

邏輯基底層215提供接達至HMC裝置之外部鏈路以及內部路由及異動邏輯兩者。外部I/O鏈路由四個或八個邏輯鏈路提供。每一鏈路可包含十六個或八個串列I/O或SERDES雙向鏈路之一群組。四個鏈路裝置具有在10 Gbps、12.5 Gbps及15 Gbps下操作之能力。八個鏈路裝置具有在10 Gbps下操作之能力。在內部，鏈路附接至路由邏輯以便引導邏輯裝置處之控制每一垂直記憶體儲存單元之異動。在圖2之實例中，路由邏輯包含縱橫交換邏輯。

HMC鏈路結構支援將裝置附接至主機(處理器)或其他HMC裝置。

兩者之能力。鏈結(*chaining*)之此概念准許構造需要比一單個裝置大之容量而不擾動鏈路結構及經封包化異動協定之記憶體子系統。鏈路可以眾多拓撲經組態為主機裝置鏈路或直通鏈路。

圖3展示基於基底四鏈路HMC組態之四個可能裝置拓撲之方塊圖。HMC記憶體裝置305可以一傳統網路拓撲(諸如一網狀、環面(*torus*)或縱橫)組態。將多個HMC裝置鏈結在一起增加可用於一主機之總記憶體容量。

返回至圖2，針對每一獨立儲存庫，一儲存庫之邏輯基底層類似於一DIMM控制器單元。在圖2之實例中，儲存庫邏輯區塊225及其各別儲存庫儲存單元220可組織成四重單元。每一四重單元表示四個儲存庫單元。每一四重單元與最接近實體鏈路區塊鬆散地相關聯。以此方式，主機裝置具有以下能力：藉由邏輯上將請求封包發送至其相關聯四重單元實體上最接近於所需儲存庫之鏈路而使一HMC裝置之整個邏輯基底層之延時最小化。

一旦在一目標記憶體儲存庫內，記憶體儲存區便可分解成傳統概念之記憶體庫及動態隨機存取記憶體(DRAM)。穿過經堆疊記憶體層之垂直存取類似於選擇適當記憶體庫。下部記憶體庫可經組態於下部晶粒層中，而垂直上升(*vertical ascension*)選擇後續記憶體庫。一旦在一記憶體庫層內，便可使用列及行以傳統方式組織DRAM。儲存庫控制器將DRAM分解成各自定址16個位元組之一百萬位元(1Mb)區塊。針對每一行提取，可以32個位元組來執行至一目標記憶體庫之讀取或寫入請求。

針對一多主機多HMC系統，使用一經封包化格式來執行主機裝置與HMC裝置之間的頻帶內通信。此經封包化格式可包含三個主要封包分類：請求封包、回應封包及流量控制封包。封包可經組態為一單個16位元組流量單元或FLIT之一倍數。最大封包大小含有9個FLIT

或144個位元組。最小16位元組(一個FLIT)封包含有一封包標頭及封包尾。用於所有記憶體有效負載大小之記憶體讀取請求封包可僅需要封包標頭、封包尾及各別實體記憶體位址。因此，讀取請求可使用一單個FLIT組態。然而，寫入請求及不可部分完成(atomic)請求封包亦含有分別用於寫入及讀取-修改-寫入操作之所需輸入資料。因此，此等請求類型具有2個FLIT至9個FLIT之封包寬度。

一HMC系統實施方案可具有多個封包重新排序點。經指定用於輔助裝置之到達封包可通過等待本端儲存庫存取之彼等輔助裝置。本端儲存庫亦可重新排序已佇列封包以便最高效地利用去往及來自各別儲存庫記憶體庫之頻寬。然而，一給定HMC實施方案中呈現之重新排序點維持一封包串流在一儲存庫內自一特定鏈路至一特定記憶體庫之次序。此確保後續接著記憶體讀取請求之記憶體寫入請求遞送正確及確定性記憶體行為。

如先前所解釋，一HMC提供一高頻寬裝置。然而，本發明系統互連規範通常缺少維持系統(其主記憶體係基於高頻寬裝置，諸如一HMC)中之適當平衡必需之充分功能性、靈活性及效能。舉例而言，一系統架構可包含用以維持一系統單晶片(SoC)與本端記憶體之間的一每秒數百十億位元組頻寬之一或多個HMC，但該系統架構可提供僅具有此可用頻寬之一小部分(例如，十分之一)之系統互連。該結果可為一高度不平衡系統。

此外，若附接至系統互連之處理器提供快取一致性機制，則單獨邏輯區塊可必須新增至該系統以提供充分頻寬及功能性，以維持共用記憶體及快取記憶體子系統兩者中之並行性及正確性。此產生與一組複雜周邊邏輯區塊之處理器互連、記憶體互連及系統互連及具有不同程度之延時之複雜協定規範。

低效快取一致性尤其係具有分散於多個SoC裝置或通訊端當中的

執行之多個執行緒(或任務)之應用中之一問題。若核心SoC或處理器支援功能資料快取，則SoC通訊端之間的快取一致性機制必須支援可為比可用於系統互連上之頻寬高之一數量級之一本端記憶體頻寬。再次，該結果係一高度不平衡系統。

用以改良系統層級延時之一方法係提供具有實質上高頻寬記憶體裝置之一頻寬之一記憶體一致性子系統。用以提供此之一種方式係使用一或多個HMC裝置實施系統記憶體一致性。由於HMC具有執行原位不可部分完成記憶體操作之能力，因此HMC可經延伸以提供一高頻寬及低延時記憶體一致性機制。

在一習用系統中，用於快取一致性之一種方法係使用一基於目錄之快取一致性協定。一致性目錄充當一篩選器，一處理器透過該篩選器尋求許可以自主記憶體將一項目載入至其快取。當一項目改變時，該目錄更新具有彼項目之其他快取或使該等其他快取失效。基於目錄之快取一致性包含對將狀態指派至快取線之一「目錄」之維持。每一快取線可被指派一「家節點(home node)」作為快取線之狀態之最終仲裁器。家節點擁有位址空間且准許修改快取線。目錄指派狀態且基於目錄之協定然後使用主機當中的點對點通信以處理後設資料。

處理器通常儲存目錄之顯著量之快取線狀態資訊且依賴於快速快取記憶體存取以讀取快取後設資料且判定快取線之狀態。一快取記憶體控制器保持追蹤哪一主機簽出快取線。在此組態中，快取目錄本身必須經快取(通常完全相關聯地)，從而形成維持於最後層級快取或記憶體控制器處之複雜且高耗電資料結構。針對大的對稱多處理器(SMP)組態，此等資料結構通常不適合於晶片上且必須由DRAM支持。

由於其實體記憶體儲存之階層式性質，因此一HMC裝置相較於傳統庫存之DRAM裝置(例如，DDR3或DDR4裝置)可具有一不同實體

定址及交錯模型。HMC裝置之實體位址編碼至含有儲存庫、記憶體庫及位址位元之一34位元欄位中。具有四個鏈路裝置之一實施方案利用該欄位之下部32個位元，且具有八個鏈路裝置之一實施方案利用該欄位之下部33個位元。

與DDR3或DDR4裝置相比，一HMC裝置可廉價地實施一抽象記憶體儲存介面。舉例而言針對一DDR類型裝置，當一處理器期望存取一位址時，真實或實際實體位址放置於一匯流排上且被存取。在一HMC協定中，實體位址可在其到達HMC處時重新映射至一不同記憶體位址。因此，對於一HMC，不存在對所寫入之實際記憶體位置之保證。另外，不需要HMC在一設定時間處產生一回應。

可利用一HMC之此位址重新映射能力以實施基於目錄之記憶體一致性。為提供對目錄記憶體一致性之支援，HMC之基底邏輯層可經組態以將一給定記憶體區塊之位置抽象化成一相關聯記憶體字組或「快取線」。因此，可將記憶體一致性目錄儲存於HMC儲存儲存庫內。此提供使用相同HMC協定管理傳統記憶體訊務及快取一致性訊務兩者之一方法；藉此提供本端記憶體資料存取及系統記憶體資料存取之統一性。

圖4展示一種一致性目錄記憶體格式之一實例。可在包含於基底實體位址欄位中之實體位址方面闡述一HMC之直接記憶體空間。在所展示之實例中，多達34個位元之實體位址儲存區係可用的。基底實體位址用作一記憶體操作之目標且可包含儲存庫、記憶體庫及DRAM位址位元。位址延伸立方體ID (CUB)欄位指示其中支持記憶體字組之記憶體駐存之HMC裝置。目錄位元欄位提供記憶體一致性位元以如同記憶體字組係一快取線一般地維持記憶體字組之狀態資訊。若記憶體一致性目錄使用MESI一致性協定，則目錄位元包含一經修改位元、一專屬位元、一共用位元及一无效位元。藉由存取與基底實體位址

址相關聯之記憶體字組加上CUB位移而執行將資料字組作為一快取線而存取。若一資料項目包含八個位元組，則一單個8GB HMC裝置可支援超過十億個並行目錄項目。

在記憶體一致性以此方式組織之情況下，可使用HMC不可部分完成記憶體操作對記憶體執行查找及修改操作。在使用HMC本機之異動協定與目錄互動同時維持目錄可變動性。此使得記憶體存取之類型較統一。

可使用一不可部分完成請求執行對目錄項目之存取。不可部分完成請求涉及：自DRAM讀取16個位元組之資料(如依請求位址欄位存取)；透過使用一16位元組運算元(亦包含於請求封包中)而對資料執行一操作；及然後將結果寫回至DRAM中之相同位置。讀取-更新-寫入序列以不可部分完成方式發生，此意指在完成不可部分完成請求之寫入之前無其他請求可存取相同記憶體庫。不可部分完成請求類似於一16位元組寫入請求，在於該等不可部分完成請求在請求封包中具有一16位元組資料有效負載，且可或可不傳回一寫入回應(取決於所發出之請求是投遞(posted)請求還是非投遞請求)。由不可部分完成操作產生之資料在一回應命令中不傳回。可使用不可部分完成記憶體操作(諸如(舉例而言)一HMC協定之一ADD立即請求命令)執行對目錄項目之修改。亦可使用不可部分完成記憶體操作(諸如(舉例而言)其中目標立即值為零之一ADD立即請求命令)執行對目錄項目之查找。

圖5展示用HMC裝置實施以包含一基於HMC之記憶體一致性目錄之一電腦系統500之一實例之一方塊圖。該系統包含複數個主機處理器。該等處理器可實施為SoC，其中每一處理器由一SoC通訊端502表示。該實例展示四個處理器，但該系統靈活地包含其他數目個處理器。

該系統亦包含經組態為用於主機處理器之一分散式共用記憶體

之複數個HMC裝置。該等HMC裝置中之某些HMC裝置用作端點裝置且某些HMC裝置用作集線器裝置。集線器裝置可提供直通通信能力以提供系統內通信。

假定記憶體用具有以三維方式配置之記憶體晶粒之HMC裝置實施，共用記憶體經組態以用於3D記憶體存取。一HMC裝置505之記憶體晶粒之記憶體之至少一部分經映射以包含記憶體一致性目錄之至少一部分。此暗示習用快取一致性資料或後設資料可儲存於其中記憶體字組駐存之記憶體中。一HMC裝置之一邏輯基底晶粒包含管理至少一個第二裝置(諸如(舉例而言)一處理器或另一HMC裝置)對HMC裝置記憶體之三維(3D)存取之至少一個記憶體控制器。邏輯基底晶粒亦包含針對儲存於複數個記憶體晶粒之記憶體中之資料實施一記憶體一致性協定之邏輯電路。

由於一致性目錄儲存於HMC裝置中，因此一HMC裝置本身可作用於其自身之資料之家節點。記憶體一致性狀態資訊可與該資料儲存在一起。邏輯基底層在不將該資料儲存於一單獨快取中之情況下允許記憶體一致性狀態資訊在處理器當中傳遞。儲存於共用記憶體中之資料基本上係一快取線。另外，假定記憶體一致性由HMC裝置實施，與維持記憶體一致性狀態資訊相關聯之處理額外負擔自處理器移除。

如先前所解釋，針對一多主機多HMC系統，使用一經封包化協定來執行主機裝置與HMC裝置之間的I/O通信。記憶體一致性協定可包含於此I/O協定中。經封包化資訊可包含記憶體一致性狀態資訊。在某些實例中，記憶體一致性狀態資訊可如圖4之實例中一般與位址資訊包含在一起。將記憶體一致性協定與I/O協定包含在一起將兩種功能組合至一統一高效通信系統中。此使本端記憶體請求及系統記憶體請求兩者中之延時最小化且消除本端介面與系統介面之間的協定轉

譯。

狀態資訊判定一處理器是否可安全地將資料寫入至記憶體字組、記憶體字組當前是否保持有效資料且判定可如何將記憶體字組轉換至其中可能進行對記憶體之讀取或寫入之一狀態中。在某些實例中，一HMC裝置之邏輯電路實施一MSI快取一致性協定且判定記憶體晶粒之一記憶體字組之「經修改」狀態資訊、「共用」狀態資訊及「無效」狀態資訊且與記憶體晶粒之記憶體字組相關聯地儲存狀態資訊。在特定變化中，經修改狀態資訊、共用狀態資訊及無效狀態資訊分別包含一經修改位元、一共用位元及一無效位元。在特定實例中，一HMC裝置之邏輯電路實施判定專屬權狀態資訊之一MESI快取。在特定實例中，一HMC裝置之邏輯電路實施判定記憶體晶粒之一記憶體字組之所有權狀態資訊之一MOSI快取。邏輯電路與共用記憶體資料字組相關聯地儲存專屬權狀態資訊及所有權狀態中之一者或兩者。在某些實例中，集線器HMC裝置將路由資訊新增至一經封包化訊息。路由資訊可用以(例如，藉由一接收HMC裝置)判定記憶體一致性狀態資訊。

根據某些實例，一「簽出/簽入」機制由希望更新一特定記憶體字組之處理器使用。在此情形中，處理器請求對記憶體位置 X 之專屬存取。一旦彼存取由其中記憶體字組駐存之HMC裝置授予，記憶體字組便可標記為「經簽出」(在裝置上之一小相關聯表中或作為記憶體中之狀態之一額外位元)。當資料由處理器往回簽入時，HMC裝置推定該資料被更新。因此，記憶體字組本身將由處理器寫回(或「清空」)且記憶體中之空間可經重新使用以儲存額外狀態資訊(例如，指示簽出記憶體字組之裝置之狀態資訊)。對記憶體字組 X 之進一步請求可自組合兩條資訊(例如，快取線由處理器 P 簽出之狀態資訊)之HMC裝置接收一回應。基於目錄之一致性協定然後藉由使用記憶體一致性

狀態更新(例如，可要求處理器 P 清空記憶體內容等)而繼續進行。當處理器 P 將新資料清空回至家節點(記憶體字組之HMC裝置)時，可簽入記憶體字組且清除狀態資訊。

在某些實例中，HMC裝置回應於對一目標記憶體字組之一特定狀態請求。舉例而言，目標HMC裝置可自一處理器接收對對一記憶體字組之專屬存取之一請求(例如，對一專屬狀態之一請求)。HMC裝置之邏輯電路可起始對記憶體字組之狀態資訊之一讀取。若狀態資訊指示一第二裝置已具有對記憶體字組之專屬存取及所有權，則邏輯電路可使該記憶體字組無效且等待來自擁有之裝置之無效確認。當接收到確認時，HMC裝置可將專屬存取賦予原始請求裝置。

圖6展示操作一電腦系統(諸如圖5中所展示之實例性系統)之一方法600之一實例之一流程圖。在605處，諸如藉由一HMC協定而管理對該電腦系統之一分散式共用記憶體之存取。該共用記憶體經組態以用於三維存取，且包含一或多個HMC裝置。

在610處，將記憶體一致性狀態資訊儲存於分散式共用記憶體中作為一記憶體一致性目錄。在某些實例中，記憶體一致性目錄分散於電腦系統之多個HMC裝置上。

在615處，在運算系統之複數個處理器與共用記憶體之間傳遞經封包化資訊，且在620處，將快取一致性狀態資訊與所傳遞經封包化資訊包含在一起。在某些實例中，在620處，在電腦系統之處理器之間點對點地傳遞快取一致性狀態資訊。

本文中所闡述之實例性系統及方法可提供優於一習用多處理器分散式共用記憶體系統的數個優點。由HMC裝置提供之高頻寬產生一高頻寬記憶體及高頻寬系統互連。HMC裝置在系統架構中之分佈使本端記憶體之間以及與HMC裝置互連之其他系統層級記憶體之間的頻寬及系統延時特性平衡。將記憶體一致性協定併入至系統I/O協定

定中提供維持本端記憶體與系統層級記憶體之間的記憶體請求之不可部分完成性之一個統一協定。該統一協定藉由減少或消除本端記憶體存取與系統層級記憶體存取之間的協定轉譯且藉由使用用於本端記憶體存取及系統級記憶體存取兩者之一高效HMC封包規範而使延時最小化。藉由卸載記憶體一致性目錄之維持且藉由由SoC簡化一致性目錄查找而簡化SoC設計。

用於快取一致性之另一方法係使用一窺探協定。在一窺探協定中，一電腦系統之處理器監視使用可提供處理器當中的點對點通信之一記憶體一致性介面傳輸之記憶體訊務。與一基於目錄之方法相比，處理器自經由一致性介面之通信追蹤狀態資訊而非藉由存取一目錄來判定狀態資訊。針對一MSI窺探協定，記憶體一致性介面廣播一特定快取線是否處於經修改、共用或無效狀態中，從而允許快取記憶體及處理器因此更新共用複本之狀態。舉例而言，當修改一快取線時，必須使系統中之快取線之其他複本無效，使得可維持一致資料。

一電腦系統之HMC裝置可經組態以自處理器將窺探功能卸載(諸如藉由將記憶體一致性廣播協定併入至該系統之I/O協定中)。保持追蹤記憶體字組之狀態(或窺探)之一HMC裝置可回應於對共用記憶體字組之請求，猶如任何習用快取記憶體回應於對共用快取線之請求一樣。

舉例而言，一HMC裝置之一儲存庫之一邏輯基底層可包含判定駐存於其裝置中之一記憶體字組之狀態資訊之邏輯電路。當對記憶體字組進行讀取存取時，一邏輯基底層之邏輯電路可將彼記憶體字組之狀態資訊設定為「專屬的」以將專屬存取賦予請求者。可藉助對讀取存取之經封包化回應而傳回狀態資訊。其他HMC裝置可追蹤包含於經封包化通信中之狀態資訊(例如，藉由使用一經延伸HMC協定)。電腦系統中之某些HMC裝置(例如，集線器HMC裝置)包含直通鏈路以促

進封包路由。直通鏈路可用以追蹤在回應封包中傳遞之狀態資訊。一端點(諸如一HMC主機鏈路或處理器)可藉由將一非修改請求傳輸至記憶體字組而自一記憶體字組之擁有者查詢狀態資訊。可以對非修改請求之一經封包化回應通信來傳回狀態資訊且其他HMC裝置可自回應封包偵測狀態資訊。

圖7展示操作一電腦系統之一方法700之另一實例之一流程圖。該電腦系統包含經組態為一共用記憶體之複數個HMC裝置。在705處，管理對該電腦系統之一分散式共用記憶體之存取(例如，使用一經延伸HMC協定)。在710處，分散式共用記憶體判定由記憶體存取產生之記憶體一致性狀態資訊。在某些實例中，共用記憶體分散於多個HMC裝置上且使用HMC裝置來判定記憶體一致性狀態資訊。

在715處，在運算系統之複數個處理器及共用記憶體端點當中傳遞經封包化資訊。將記憶體一致性狀態資訊與所傳遞經封包化資訊包含在一起。在某些實例中，使用一經延伸HMC協定來傳遞經封包化資訊。在720處，使用共用記憶體端點來追蹤或窺探廣播記憶體一致性狀態資訊。

經窺探無效及寫回將使資料之一致複本能夠維持於記憶體中且可產生系統協定之總體簡化。將協定併入至記憶體介面之經封包化通信中可增加回應代理程式(或回應端點)之數目，但HMC裝置之高頻寬亦可能達成較快且較具可擴縮性回應。

雖然本文中已圖解說明且闡述了特定實施例，但熟習此項技術者將瞭解，達成相同目的、結構或功能之任何配置可替代所展示之特定實施例。本申請案意欲涵蓋對本文中所闡述之本發明之實例性實施例之任何改編或變化形式。意欲本發明僅由申請專利範圍及其等效物之完全範疇限制。

【符號說明】

105	混合記憶體立方體
110	積體電路記憶體晶粒
115	邏輯基底晶粒層
120	儲存庫
205	混合記憶體立方體
215	邏輯基底層
220	儲存庫/儲存庫儲存單元
225	儲存庫邏輯區塊
305	混合記憶體立方體記憶體裝置
502	系統單晶片通訊端
505	混合記憶體立方體裝置

申請專利範圍

1. 一種運算系統，其包括：

複數個主機處理器；及

複數個經堆疊記憶體裝置，其經組態為用於該等主機處理器之一分散式共用記憶體，各經堆疊記憶體裝置包含：

複數個積體電路記憶體晶粒，其包含配置於一第二記憶體晶粒之頂部上之至少一第一記憶體晶粒，且該記憶體晶粒之記憶體之至少一部分經映射以包含一記憶體一致性目錄之至少一部分；及

一邏輯基底晶粒，其包含：至少一個記憶體控制器，其經組態以管理至少一個第二裝置對該複數個記憶體晶粒之記憶體之三維(3D)存取；及邏輯電路，其經組態以判定儲存於該複數個記憶體晶粒之該記憶體中之資料之記憶體一致性狀態資訊、將該記憶體一致性狀態資訊與該記憶體晶粒中之該資料儲存在一起、傳遞關於對記憶體之該存取之資訊且將該記憶體狀態一致性資訊包含於該所傳遞資訊中；

其中該記憶體一致性狀態資訊在該複數個主機處理器當中傳遞。

2. 如請求項1之運算系統，其中一經堆疊記憶體裝置之該邏輯電路經組態以使用用以存取該記憶體晶粒之一輸入/輸出(I/O)協定來傳遞經封包化資訊且作為該I/O協定之部分實施一記憶體一致性協定。

3. 如請求項2之運算系統，其中一經堆疊記憶體裝置之該邏輯電路經組態以作為該I/O協定之部分路由經封包化資訊且偵測與該經封包化資訊包含在一起之記憶體一致性狀態資訊。

4. 如請求項1之運算系統，其中一經堆疊記憶體裝置之該邏輯電路

經組態以判定該記憶體晶粒之一記憶體字組之經修改狀態資訊、共用狀態資訊及無效狀態資訊且與該記憶體晶粒之該記憶體字組相關聯地儲存該狀態資訊。

5. 如請求項1之運算系統，其中一經堆疊記憶體裝置之該邏輯電路經組態以判定該記憶體晶粒之一記憶體字組之專屬權狀態資訊及所有權狀態資訊中之至少一者且與共用記憶體資料字組相關聯地儲存該狀態資訊。
6. 如請求項1之運算系統，其包含經組態為一集線器經堆疊記憶體裝置之至少一個經堆疊記憶體裝置，其中該集線器經堆疊記憶體裝置之邏輯電路將路由資訊新增至一經封包化訊息且偵測該經封包化訊息中之記憶體一致性狀態資訊。
7. 如請求項1之運算系統，其中一經堆疊記憶體裝置之該邏輯基底晶粒之該記憶體控制器經組態以管理至少一個處理器及另一經堆疊記憶體裝置對該記憶體晶粒之記憶體之3D存取。
8. 如請求項1之運算系統，其中一經堆疊記憶體裝置之該記憶體晶粒包含配置為複數個儲存庫之記憶體，且該邏輯基底晶粒包含用於該記憶體晶粒之每一儲存庫之一記憶體控制器。
9. 如請求項1之運算系統，其中該複數個主機處理器之該等處理器係系統單晶片(SoC)處理器。
10. 一種操作一電腦系統之方法，該方法包括：
 - 管理對該電腦系統之一分散式共用記憶體之存取，其中該共用記憶體經組態以用於三維存取；
 - 由該分散式共用記憶體判定由記憶體存取產生之記憶體一致性狀態資訊；
 - 將該記憶體一致性狀態資訊與該記憶體晶粒中之該資料儲存在一起；

在運算系統之複數個處理器及共用記憶體端點當中傳遞經封包化資訊且將該記憶體一致性狀態資訊與該所傳遞經封包化資訊包含在一起；及

使用該等共用記憶體端點來追蹤該記憶體一致性狀態資訊。

11. 如請求項10之方法，其中使用該等共用記憶體端點來追蹤該記憶體一致性狀態資訊包含：使用包含該共用記憶體之經堆疊記憶體裝置來追蹤該記憶體一致性狀態資訊。
12. 如請求項10之方法，其包含將該記憶體一致性狀態資訊與一共用記憶體資料字組儲存在一起。
13. 如請求項10之方法，其包含使用包含用於該共用記憶體之至少一個記憶體控制器之一邏輯基底層來修改該記憶體一致性狀態資訊。
14. 如請求項10之方法，其中將記憶體一致性狀態資訊儲存於該共用記憶體中包含：將一經修改位元、一共用位元及一無效位元與該共用記憶體資料字組儲存在一起。
15. 一種經堆疊記憶體裝置，其包含：

複數個積體電路記憶體晶粒，其包含配置於一第二記憶體晶粒之頂部上之至少一第一記憶體晶粒，且該記憶體晶粒之記憶體之至少一部分經映射以包含一記憶體一致性目錄之至少一部分；及

一邏輯基底晶粒，其包含：至少一個記憶體控制器，經組態以管理至少一個第二裝置對該複數個記憶體晶粒之記憶體之三維(3D)存取；及邏輯電路，其經組態以判定儲存於該複數個記憶體晶粒之該記憶體中之資料之記憶體一致性狀態資訊、將該記憶體一致性狀態資訊與該記憶體晶粒中之該資料儲存在一起、傳遞關於對記憶體之該存取之資訊且將該記憶體狀態一致性資

訊包含於該所傳遞資訊中；

其中該記憶體一致性狀態資訊在複數個主機處理器當中傳遞。

16. 如請求項15之經堆疊記憶體裝置，其中該邏輯基底晶粒之該邏輯電路經組態以使用用以存取該記憶體晶粒之一輸入/輸出(I/O)協定來傳遞經封包化資訊且作為該I/O協定之部分實施一記憶體一致性協定。
17. 如請求項16之經堆疊記憶體裝置，其中該邏輯基底晶粒之該邏輯電路經組態以作為該I/O協定之部分路由經封包化資訊且偵測與該經封包化資訊包含在一起之記憶體一致性狀態資訊。
18. 如請求項15之經堆疊記憶體裝置，其中該邏輯基底晶粒之該邏輯電路經組態以判定該記憶體晶粒之一記憶體字組之經修改狀態資訊、共用狀態資訊及無效狀態資訊且與該記憶體晶粒之該記憶體字組相關聯地儲存該狀態資訊。
19. 如請求項15之經堆疊記憶體裝置，其中該邏輯基底晶粒之該邏輯電路經組態以判定該記憶體晶粒之一記憶體字組之專屬權狀態資訊及所有權狀態資訊中之至少一者且與共用記憶體資料字組相關聯地儲存該狀態資訊。

圖式

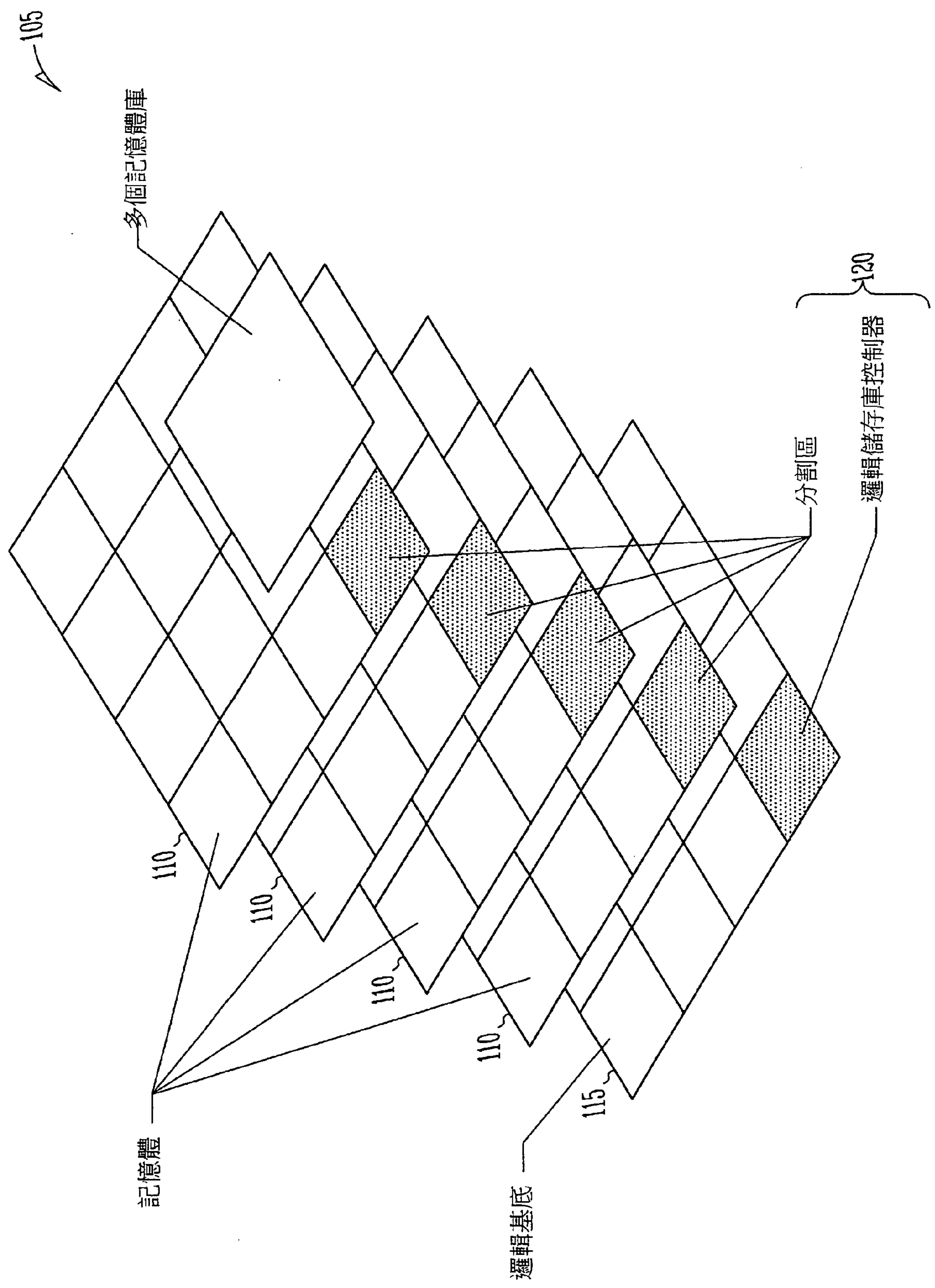


圖 1

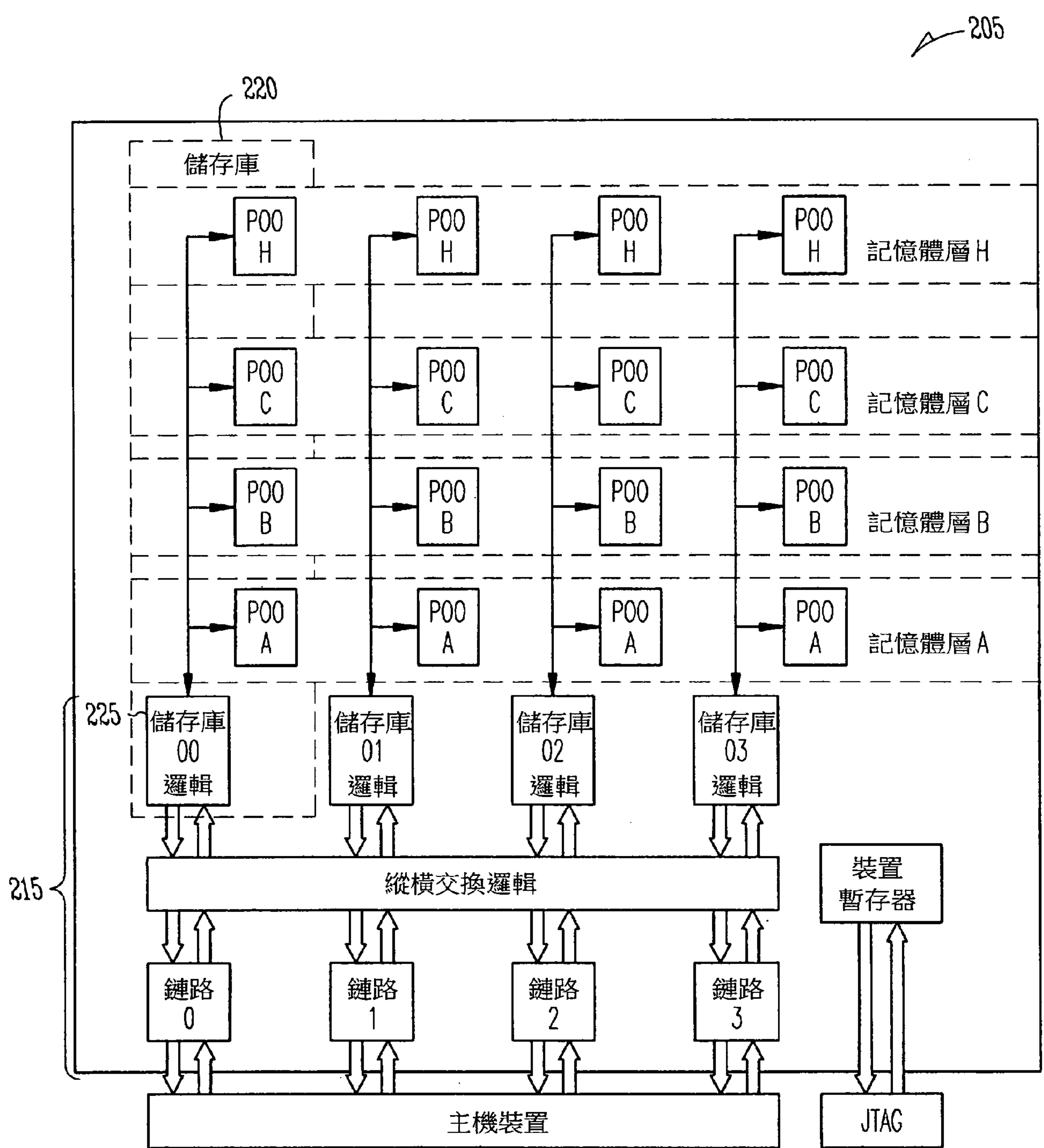
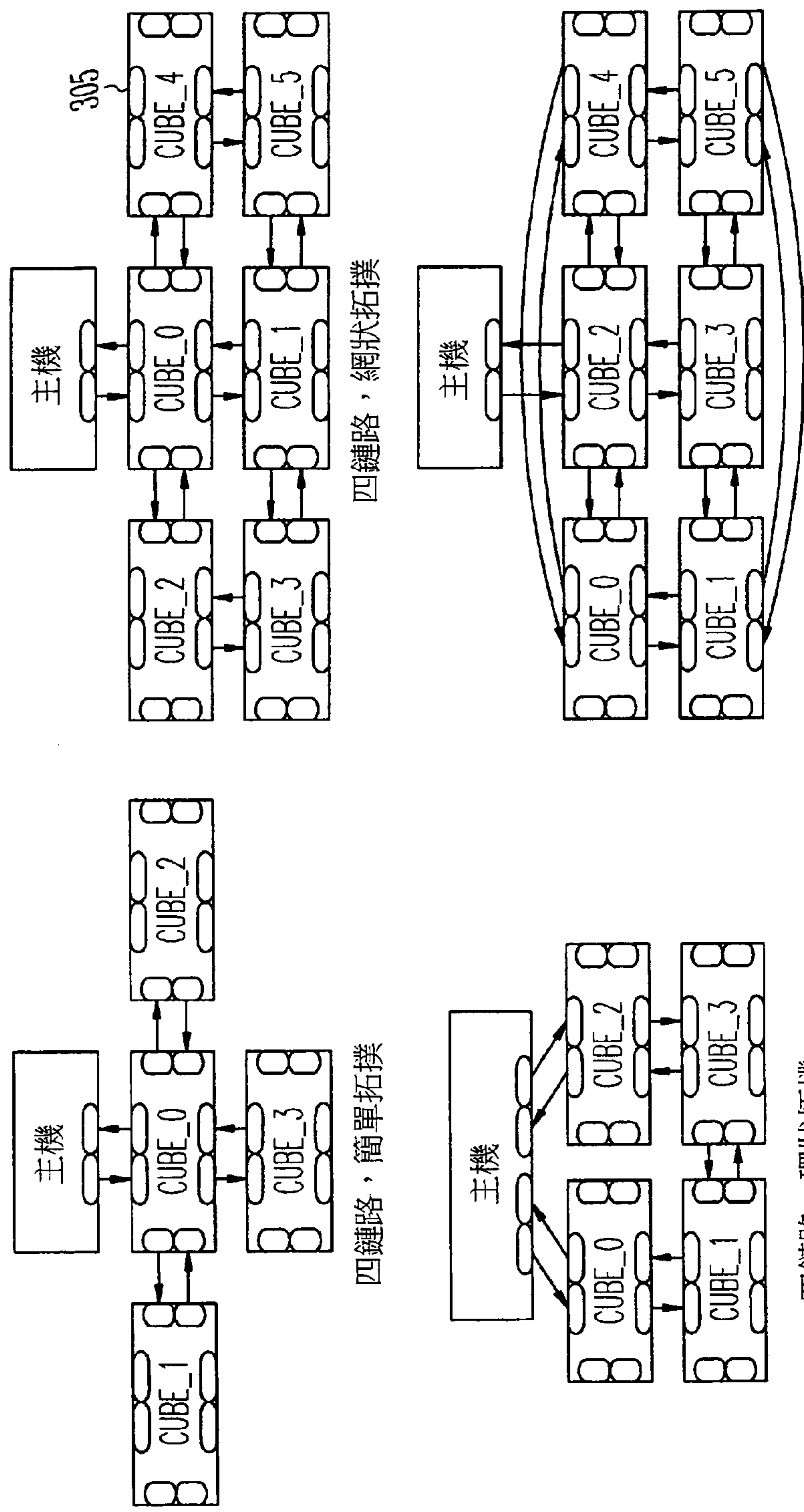


圖 2



四鏈路，網狀拓撲

四鏈路，簡單拓撲

四鏈路，環狀拓撲

四鏈路，2D 環面拓撲

圖 3

未使用 [63:46]	目錄 位元 [45:42]	CUB [41:34]	基底實體 位址 [33:0]
-------------	------------------	-------------	-------------------

圖 4

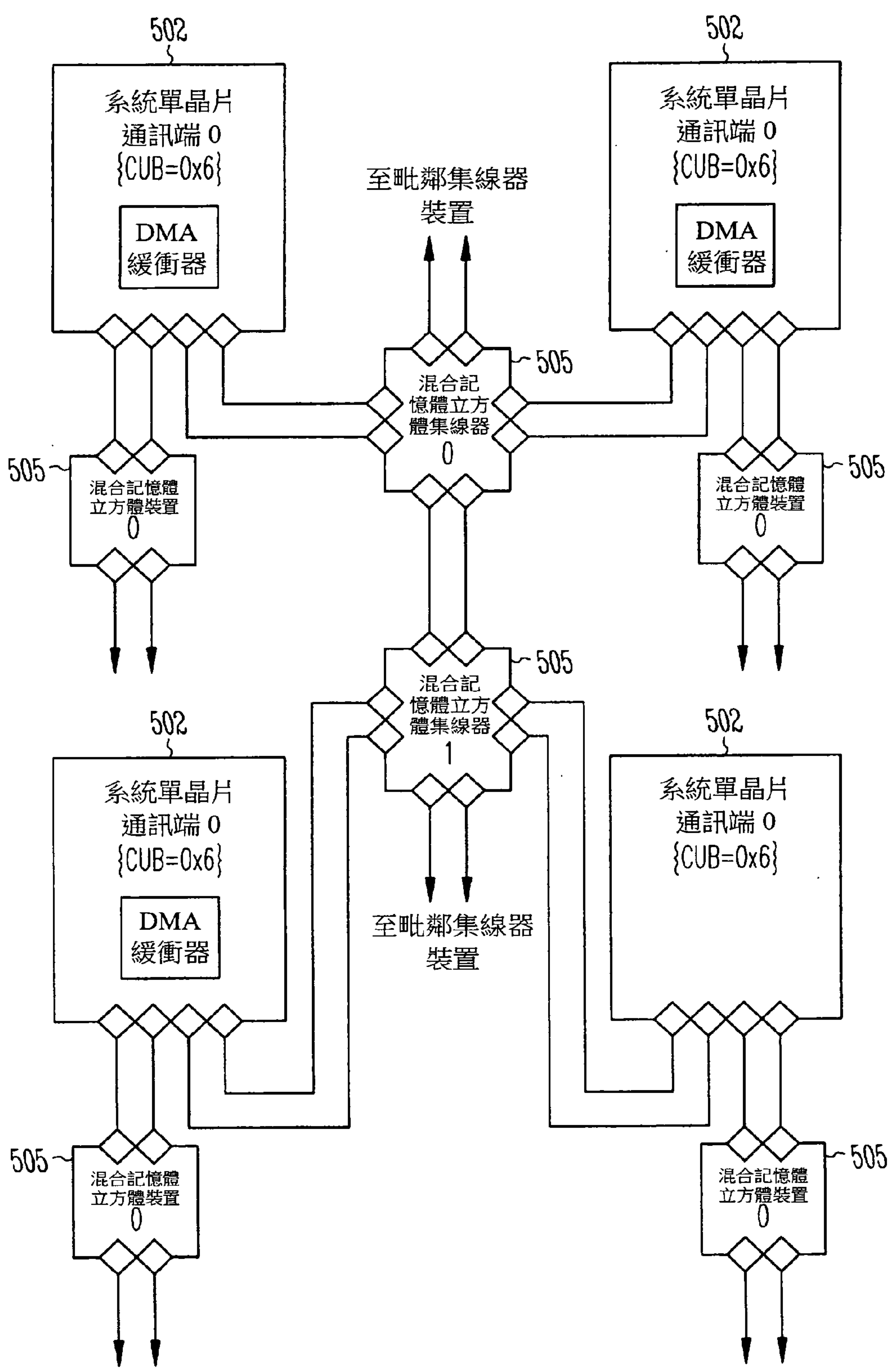


圖 5

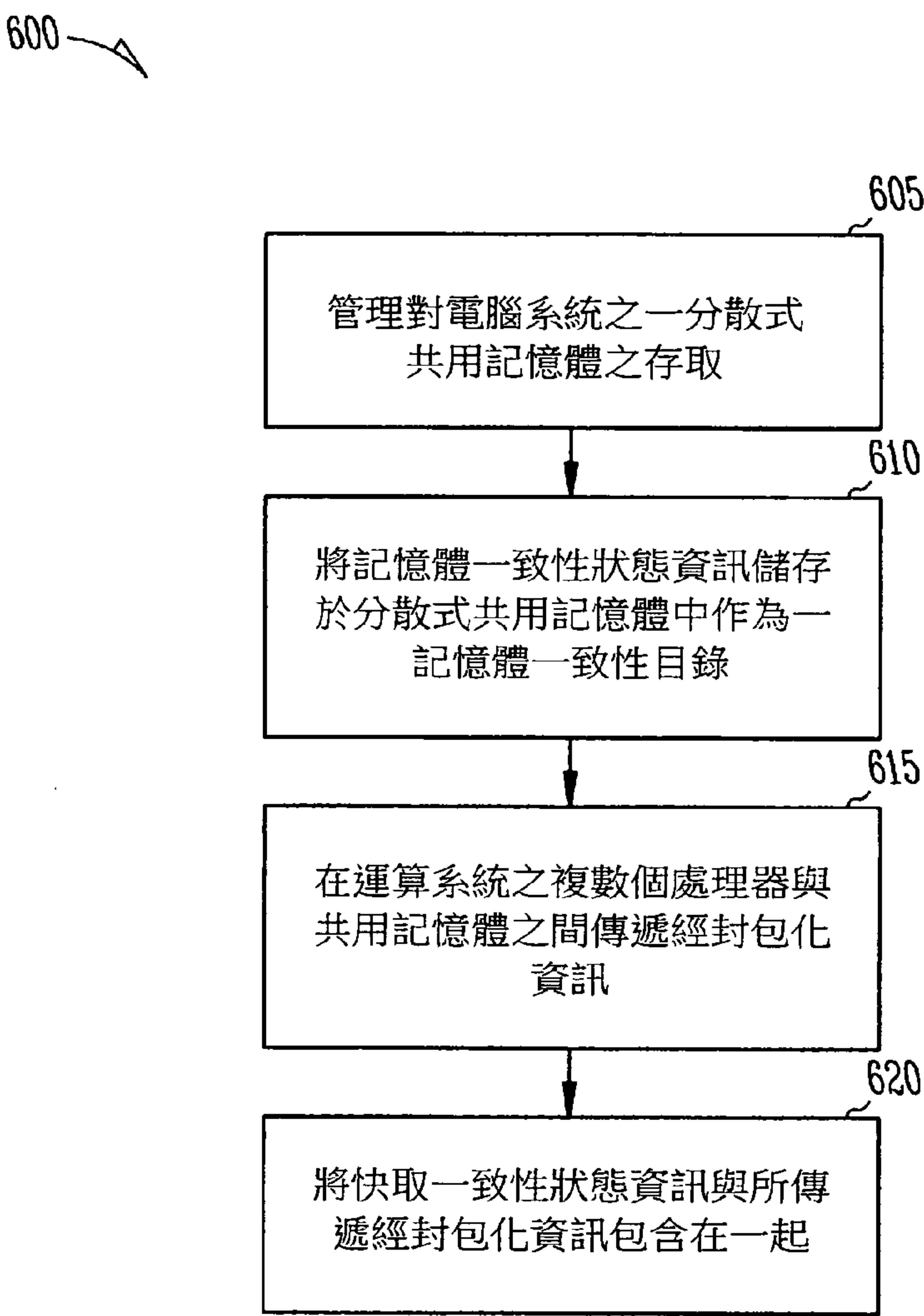


圖 6

700 ↗

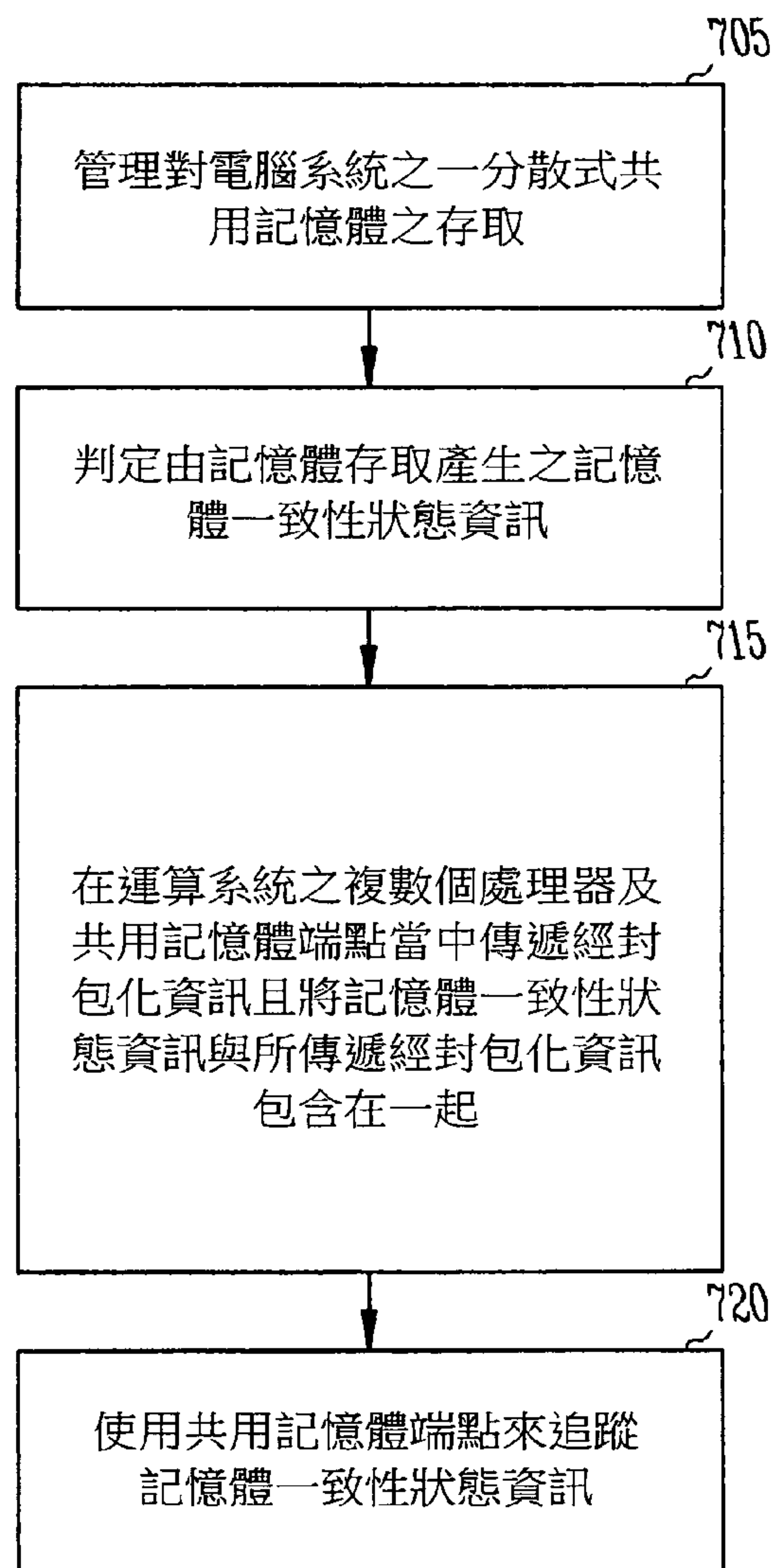


圖 7