

申請日期： 92. 5. 12	IPC分類 6069 15/16
申請案號： 92112810	

(以上各欄由本局填註)

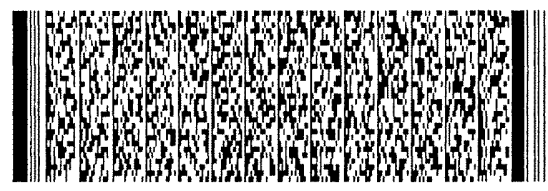
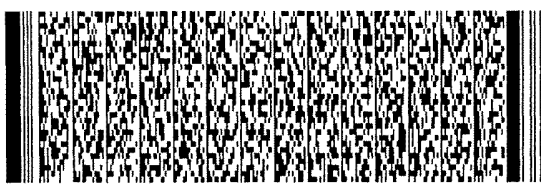
發明專利說明書

200401202

一、 發明名稱	中文	直接定址分享壓縮記憶體系統
	英文	DIRECT ADDRESSED SHARED COMPRESSED MEMORY SYSTEM

二、 發明人 (共4人)	姓名 (中文)	1. 卡斯特立, 維多里歐 2. 福蘭那奇, 彼得 A. 3. 海德伯格, 飛利浦
	姓名 (英文)	1. CASTELLI, Vittorio 2. FRANASZEK, Peter A. 3. HEIDELBERGER, Philip
	國籍 (中英文)	1. 美國 US 2. 美國 US 3. 美國 US
	住居所 (中文)	1. 美國紐約州10520克拉頓哈德森市克里夫蘭路72號 2. 美國紐約州10549奇斯寇山春峽路12號 3. 美國紐約州10567寇特蘭莊湖景大道西24號
	住居所 (英文)	1. 72 Cleveland Drive, Croton-on-Hudson, New York 10520 USA 2. 12 Spring Glen Drive, Mount Kisko, New York 10549 USA 3. 24 Lakeview Avenue West, Cortlandt Manor, New York 10567 USA

三、 申請人 (共1人)	名稱或姓名 (中文)	1. 國際商業機器股份有限公司
	名稱或姓名 (英文)	1. INTERNATIONAL BUSINESS MACHINES CORPORATION
	國籍 (中英文)	1. 美國 US
	住居所 (營業所) (中文)	1. 美國紐約州10504亞芒克市新奧爾察德路 (本地址與前向貴局申請者相同)
	住居所 (營業所) (英文)	1. New Orchard Road, Armonk, NY 10504, USA
	代表人 (中文)	1. 傑羅 羅森梭
	代表人 (英文)	1. Gerald Rosenthal

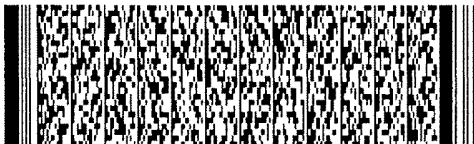


申請日期：	IPC分類
申請案號：	

(以上各欄由本局填註)

發明專利說明書

一、 發明名稱	中文	
	英文	
二、 發明人 (共4人)	姓名 (中文)	4. 羅賓森，約翰 提摩西
	姓名 (英文)	4. ROBINSON, John Timothy
	國籍 (中英文)	4. 美國 US
	住居所 (中文)	4. 美國紐約州10598約克鎮高地北鹿園大道3314號
	住居所 (英文)	4. 3314 North Deerfield Avenue, Yorktown Heights, New York 10598 USA
三、 申請人 (共1人)	名稱或 姓名 (中文)	
	名稱或 姓名 (英文)	
	國籍 (中英文)	
	住居所 (營業所) (中文)	
	住居所 (營業所) (英文)	
	代表人 (中文)	
	代表人 (英文)	



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	2002/05/30	10/158, 534	有

二、主張專利法第二十五條之一第一項優先權：

申請案號：

無

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：

無

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：

無

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)

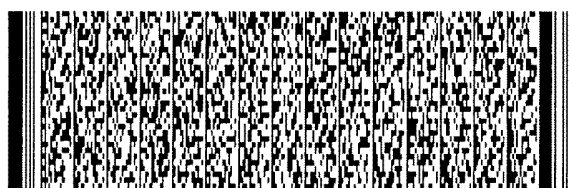
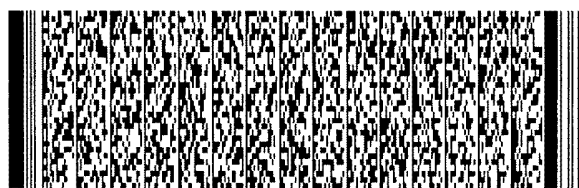
一、【發明所屬之技術領域】

本發明係關於一種電腦架構，並且特別關於一種網路連接之裝置，使得多個主機分享一記憶體段(sector)集合，其中記憶體段係用來儲存壓縮資料。此裝置以下被稱為一「直接定址分享壓縮記憶體系統(Direct Addressed Shared Compressed Memory System)」，並簡寫為SCMS。

本發明包括轉譯(translate)由主機產生之實位址(real address)為被SCMS管理之實位址，接著SCMS的實位址再轉成壓縮記憶體管理系統之實體位址(physical address)。本發明包括供確保記憶體內容保護之方法；本發明包括在複數個主機間分享記憶體內容之機制；本發明包括將一主機之實位址空間之一連續部份，分佈在多個SCMS間；本發明亦包括一方法，供確保每一主機能夠提供一特定數目之記憶體段。

二、【先前技術】

在電腦系統主記憶體使用資料壓縮，是一種新興的發展。實記憶體(real memory)，即對應到儲存在記憶體之資料的處理器位址(processor address)之集合，一般會切割為複數個成對不相交(pairwise disjoint)的區段(segment)，以對應一定數目的連續處理器位址。藉由成對不相交，即意味著每個實位址係對應至一旦唯一之此類區段。這些區段在此稱為記憶體線(memory lines)。記憶體



五、發明說明 (2)

線為壓縮的基本單位。存在壓縮記憶體中的記憶體線係經過壓縮，並且儲存於一變動數目的記憶體位置中，至於數目的大小則由其內容壓縮的程度決定。國際商業機器公司 (IBM) 具有若干有關電腦系統之主記憶體壓縮的專利。這類系統的例子揭露於美國專利US5,729,228，標題為「使用合作目錄所進行的平行壓縮/解壓縮(Parallel compression and decompression using a cooperative dictionary)」；美國專利US5,761,536，標題為「基於一最佳合適基礎，藉由指定餘數給分享記憶體區塊以降低記憶體零散破裂配置的系統與方法(System and method for reducing memory fragmentation by assigning remainders to share memory blocks on a best fit basis)」；以及美國專利US5,864,859，標題為「使用儲存定址以進行壓縮與解壓縮的系統與方法(System and method of compression and decompression using store addressing)」。

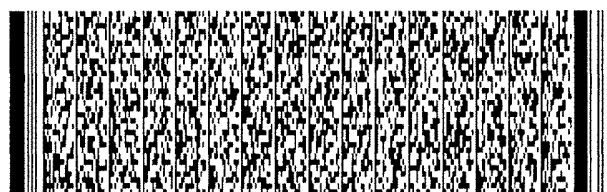
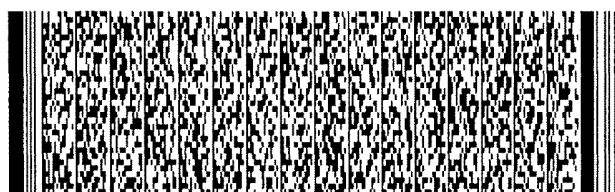
第1圖例示一範例系統，此系統具有此資料壓縮特徵。在第1圖中，中央處理單元(CPU)102從一或多個快取(cache)104讀取資料以及將資料寫入到這些快取104。當快取未中(cache miss)或欲將快取資料進行儲存時，將導致從壓縮主記憶體110讀取資料，或將資料寫入壓縮主記憶體110，而這些讀取與寫入係透過壓縮控制器106所提供的機制以達成。壓縮主記憶體110分割成兩個部分：資料部分



五、發明說明 (3)

108 及目錄107(也可視為壓縮變換表(compression translation table)，簡寫為CTT)。資料部分108則分割為成對不相交之段，亦即，實體記憶體位置的固定尺寸區間。舉例來說，一段可能由256個實體位元組(byte)構成，而此256個實體位元組具有連續實體位址。壓縮記憶體線之內容則盡可能存於最小數目的實體段中。包含壓縮線的實體段不需要具有連續實體位址，且可被安置在壓縮主記憶體110之資料部分108的任意位置。對於一位元組的實位址與包含此位元組的實體段位址之間的轉譯，係經由目錄或壓縮變換表107以達成。

第2圖包含進一步細節，以供更深入了解壓縮記憶體210的操作。處理器快取240包含未壓縮快取線241及快取目錄242，其儲存每一快取線的實位址。為了說明，以下將假設一快取線具有與一記憶體線(壓縮的單元)相同之大小。在快取未中資料時，快取提供造成快取未中的實位址270，以向記憶體的相對線要求資料。實位址切分為兩個部分：第一部份係 \log_2 (線長)個最低有效(least significant)位元，亦即自線中的位址之偏移值，其中 \log_2 是以2為基底的對數，第二部分，亦即剩餘的位元，則作為目錄220的索引(index)，其包含對於支援實位址範圍之每一線之線項目。在第2圖中，位址A1(271)對應線項目1(221)，位址A2(272)對應線項目2(222)，位址A3(273)對應線項目3(513)，並且位址A4(274)則對應線項目4(514)。在此例子中使用不同的

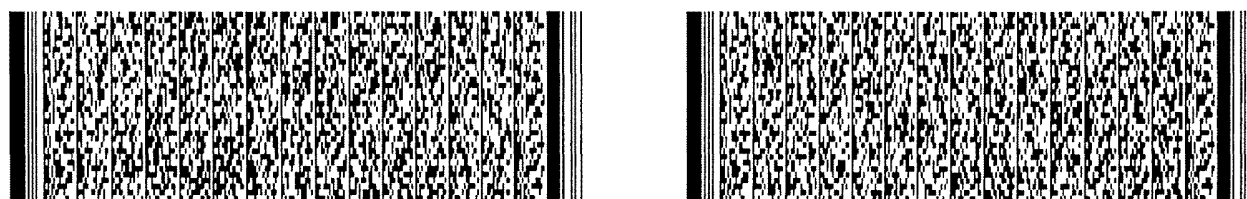


五、發明說明 (4)

位址，以顯示在壓縮主記憶體中儲存壓縮資料的不同方法。在此例中，具有位址A1的線壓縮地非常好(舉例來說，此線包含所有的0)。此種線係整個存於CCT項目221中，且並不需要記憶體段。在位址A2的線則壓縮較差，並且需要兩個記憶體段231及232，其存於資料區(data section) 230。線項目222包含指標指向記憶體段231及232。注意記憶體段232的最後一部份未被使用到。具有位址A3的線需要3個記憶體段233、234、及235。在段235剩餘的空間足夠大，以儲存具有實位址A4的壓縮線之一部份，此實位址A4輪流使用段236及段235的一部份。在位址A4及A3的線稱作同室者(roommate)。

當快取中的髒線(dirty lines)寫回到記憶體時，使用壓縮器261。在快取寫回時，髒線被壓縮。假如髒線可符合寫回之前所使用的記憶體數量時，則其直接存放在原處。否則，其被寫入適當數目的段中。假如要求的段數目減少時，未使用的段就加入自由段列表(free-sector list)。假如要求段數目增加時，它們則從自由區段列表中讀取。

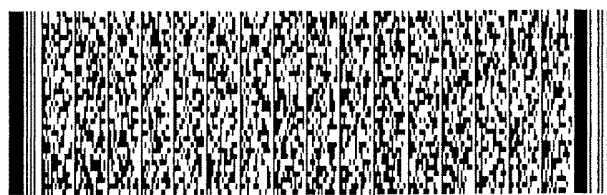
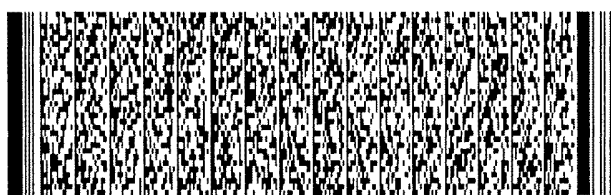
第3圖顯示目錄或壓縮變換表220中項目可能的組織方式。在此例示三種線組織。項目1(306)包含一旗標集合(a set of flags)(301)，以及四個段的位址。假如線的大小為1024位元組，並且記憶體段大小為256，則線需要最多4個段。項目2(307)包含一旗標集合、線使用之第一段位



五、發明說明 (5)

址、壓縮線的開始，與線使用的最後一段的位址。假如線需要超過兩個記憶體段，則此些段藉由一連結的指標表列而連結(亦即，每一記憶體段包含下一記憶體段之位址)。項目3包含一旗標集合，以及一高密度壓縮線，且此壓縮線壓縮至120位元或更少。在此例子中的旗標可為旗標302，辨識是否線以壓縮格式存放或未壓縮；旗標303辨識是否線為可高密度壓縮與是否整個存於目錄項目；旗標304(2位元)辨識線使用了多少的段；以及項目305(4位元)包含破碎離散資訊(fragment information)，亦即線所佔用了最後一個被使用段之多少部分(此資訊用來供同室者使用)。系統中依賴上述之壓縮記憶體結構，可達成的最大壓縮比係有賴於目錄的大小決定，亦即，實位址的最大數目等於目錄中的目錄項目之數目。

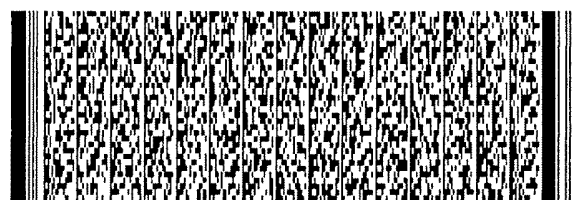
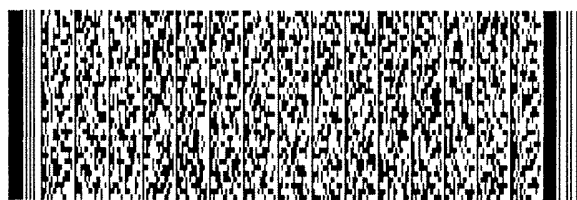
在相關領域中，已提出用來分割(partitioning)一未壓縮記憶體的方法，舉例來說，美國專利US4,564,903，標題「分割多處理器程式系統(Partitioned multiprocessor programming system)」。此專利提出一控制方法，以提供具有多個中央處理單元之多處理器(multiprocessor, MP)系統分享一主要儲存裝置(main storage, MS)及輸出/輸入(I/O)處理機制，以將複數個輸出/輸入裝置連接到主要儲存裝置，此控制方法使得多處理器系統可以使其上的多個中央處理單元，同時執行一單一處理器程式系統(Uniprocessor programming system, UPS)，即使此單一



五、發明說明 (6)

處理器程式系統僅針對單一處理器系統設計，且也只能在具有與多處理器系統相同或不同架構之單一處理器系統中執行。此專利所提出的裝置與方法，供適用於包含在多處理器系統中的非壓縮記憶體。然而，其未提出針對於一壓縮記憶體系統的裝置與方法，此壓縮記憶體系統並非多處理器系統的一部份，並且也非供多個不同電腦系統分享，並且此壓縮記憶體也非屬於任何電腦系統的一部份。

在美國專利US4,843,541，標題「資料處理系統的邏輯資源分割(Logical resource partitioning of a data processing system)」中，此專利提出一種方法，供在資料處理系統中，將一訪客操作(guest operation)可利用的資源限制為此系統配給一訪客(guest)的資源。系統的資源包括一或多個實中央處理單元，一系統主儲存裝置，一選擇性系統擴充儲存裝置(extended storage, ES)，以及複數個輸出/輸入通道，其使用複數個輸出/輸入處理器以將輸出/輸入裝置，透過複數個輸出/輸入裝置控制單元連接到系統，以及一包含軟體的主機超級使用者(host hypervisor)，主機監視管理(supervise)複數個軟體控制程式(訪客)，此複數個軟體同時及獨立地在系統中運作，並且可為同一或不同型態之控制程式，副通道(subchannel, SCHs)，供向輸出/輸入裝置指出此主機及此訪客，每一訪客則限制只能存取給定之系統資源的子集合(分區(partition))。然而，此專利未提出如何在不同主機



五、發明說明 (7)

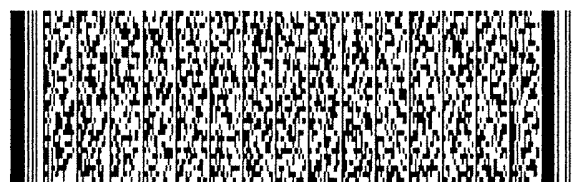
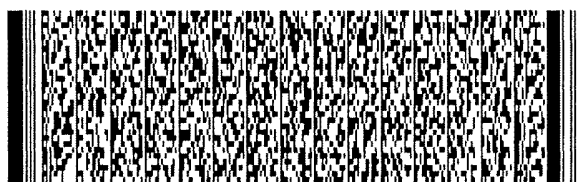
間分享一網路連接之壓縮記憶體。

在習知技術中的分區包含，透過在電腦上執行的軟體元件將所見到之實位址轉譯為被電腦所管理的實位址。這些軟體元件可在同一或多個作業系統中有多個翻版 (image)，因此它們可以執行邏輯-到-真實(logical-to-real)的位址轉譯。分區將這些軟體元件所產生的實位址，視為邏輯位址，並且執行進一步的邏輯-到-真實的位址轉譯。假如硬體不支援記憶體壓縮，實位址與實體位址是相等的。然而，假如記憶體經過壓縮，實位址則不會等於實體位址，並且需要進一步的轉譯。傳統的分區方法也包括保護機制，使得在一特定分區內所執行之軟體元件避免存取在其他分區內所執行之軟體元件的記憶體內容。

然而，習知技術所教示的分區未提及，在與電腦系統實體分離的裝置中，如何分享記憶體資源以及安全機制，其中使用資料的軟體元件就是在此電腦系統中執行，因此分享與安全機制並未被電腦系統控制，所以對於這類的電腦系統來說，上述問題是顯而易見的。

三、【發明內容】

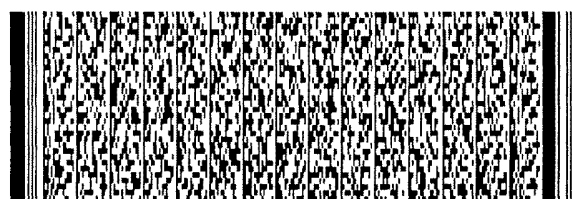
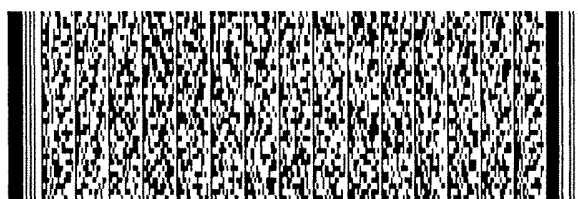
在電腦系統中，複數個主機經由一網路連接，而本發明係提供網路連結裝置，此網路連結裝置連接到此網路，並提供擴充主機的邏輯實記憶體的功能，以及在壓縮格式



五、發明說明 (8)

中維持它們記憶體的內容。本發明之此網路連接裝置係一直接定址分享壓縮記憶體系統(Shared Compressed Memory System, 以下簡稱SCMS)。SCMS將其內部實記憶體空間分割為相鄰的實位址區, 稱為區段(segment), 並且它們的實體記憶體(physical memory)被分割成由相鄰的實體位址區的共用區(common pool), 稱為段(sector)。主機能夠從這些裝置分配區段, 並且如同對它們內部記憶體定址的方式, 對這些區段之內容加以定址(address), 亦即, 透過一實位址為之。SCMS將主機所提供的記憶體位址轉譯成內部實位址。此係藉由上述壓縮記憶體目錄(CTT), 轉譯到段的實體位址而達成, 而壓縮資料就是存在這些段的實體位址。區段則動態關聯對照於實體段數目之變化值。當在區段內的資料之可壓縮性降低時, 關聯區段的實體段之數目將增加, 反之亦然。本發明特別提供裝置與方法, 供管理從共用區(common pool)的實體段, 完成對記憶體區段的指定, 而這是在不支援記憶體壓縮的裝置中無法達到的, 也是在一般支援記憶體壓縮的電腦無法完成的。

依據本發明之一實施例, 在一電腦系統中, 複數個主機經由一交互連接網路連接, 一裝置則連接到此交互連接網路, 以提供此複數個主機分享一記憶體段集合, 此記憶體段集合由複屬個記憶體段所組成, 而此複數個記憶體段供儲存壓縮資料。此裝置包括一網路卡, 供連接此裝置到交互連接網路; 亦包括一記憶體供存放此記憶體段集合;



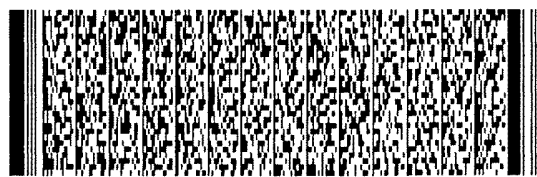
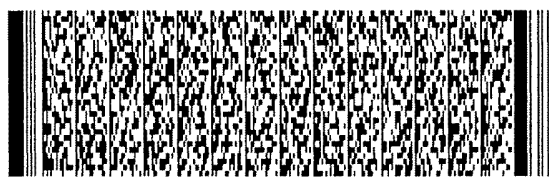
五、發明說明 (9)

亦包括控制邏輯，供管理此記憶體，此控制邏輯包括一記憶體壓縮器/解壓縮器。此記憶體更包括一目錄，以轉譯至少一主機之實位址為此裝置之一位址。

在另一實施例中，此裝置的控制邏輯更包括一暫存器矩陣，此矩陣的每一列對應到一不同的主機識別碼，並且每一行則對應到一不同的區段號碼，藉此，經由至少一主機之識別碼所決定的特定暫存器之內容與所需之區段連接到偏移值(offset)，以產生此裝置的實位址。

在另一實施例中，此裝置的控制邏輯包括一暫存器陣列，此陣列包括複數個暫存器，對應於裝置之複數個區段，其中每一暫存器對於每一支援主機包含一位元，藉此，假如此位元設為1，一需求區段則連接到偏移值，以產生裝置之實位址，假如此位元等於0，則此至少一主機將被拒絕存取此區段。

在本發明之另一實施例中，此裝置的控制邏輯包括一相聯記憶體。此相聯記憶體對於每一裝置的管理區段具有一對應列，且此相聯記憶體具有第一行與第二行，此第一行含有複數個鍵值(key)，此第二行含有複數個數值(value)，其中此複數個鍵值為此複數個主機對於需求區段所提供的起始實位址，且此複數個數值為此裝置的起始實位址，藉由上述之至少一主機之一搜尋鍵(search key)所



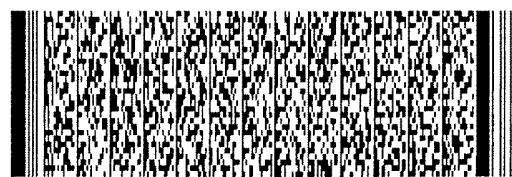
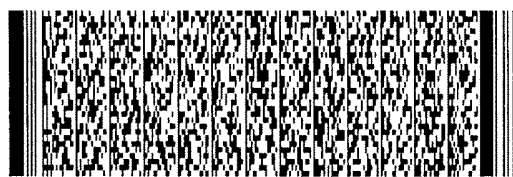
五、發明說明 (10)

決定之一數值連接到此偏移值，以產生此裝置的實位址。

依據另一實施例，提供一電腦系統，此電腦系統包括一經由一交互連接網路連接的主機；以及複數個裝置，每一裝置包括一記憶體段集合，此記憶體段集合由複數個記憶體段所組成，此複數個裝置連接於此交互連接網路，以容許此複數個主機在此複數個裝置間，分享此記憶體段集合，其中此複數個記憶體段供儲存壓縮資料。主機實位址的一連續子集合分佈在此複數個裝置上。

在本發明之另一實施例中，在一電腦系統中，複數個主機經由一交互連接網路連接，包括一連接到交互連接網路之裝置，以容許複數個主機分享一記憶體段集合，此記憶體段集合由複數個記憶體段所組成，而此複數個記憶體段供存放壓縮資料。並且，亦提供一種方法，供管理每一主機所使用之複數個記憶體段。此方法包括決定分配給每一主機之最大數目之段；分配給至少一主機一定限暫存器，供存放分配給此至少一主機之最大數目之段；分配給至少一主機一計數暫存器，供儲存此至少一主機使用之複數個段；並且比較此至少一主機的定限暫存器之值與計數暫存器之值，以決定是否此計數暫存器之值已經超過此定限暫存器之值。

此方法更包括，假如此計數暫存器之值已經超過此定



五、發明說明 (11)

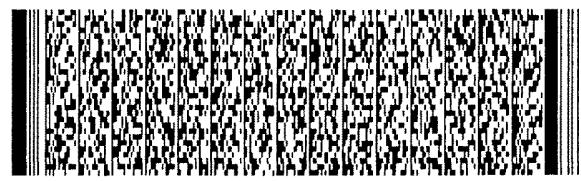
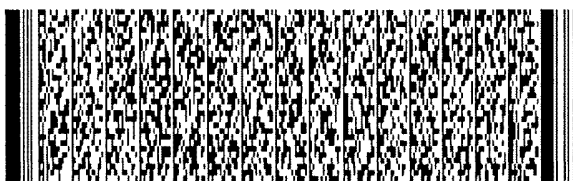
限暫存器之值，則避免至少一主機之寫入操作，並通知每一此複數個主機，此寫入操作為無效。

依據本發明之另一實施例，在一電腦系統中，複數個主機經由一交互連接網路連結，此交互連接網路包含一裝置，此裝置與此交互連結網路耦接以供此複數個主機分享一記憶體段集合，此記憶體段集合由複數個記憶體段所組成，此複數個記憶體段供存放壓縮資料，一種方法被提出以供將每一主機指定之實位址轉譯為此裝置之實位置。此方法包括由此至少一主機使用一主機指定實位址，以要求此裝置之一需求實位址；使用此主機指定實位址，以辨識此裝置的資料持有位置；使用此資料持有位置的內容，以決定一第一數值；以及將此第一數值接續到一偏移值，以產生此裝置之實位址。

四、【實施方式】

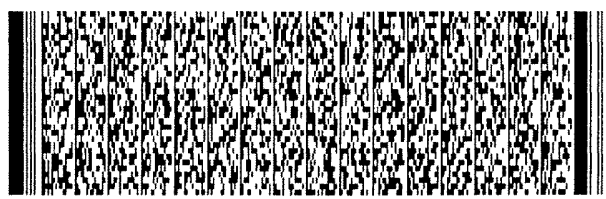
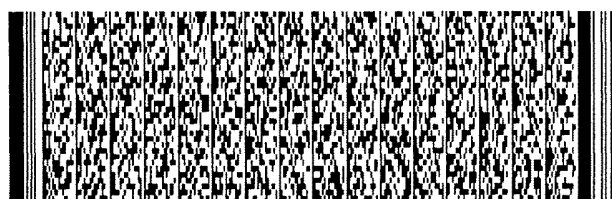
本發明之較佳實施例的描述，係假設網路系統為一分頁(paged)作業系統，例如Windows95、Windows98、WindowsNT、Windows2000、Linux、AIX及所有其他版本的UNIX、MacOS、IBM OS/400等等。習知技藝者當能清楚理解，本發明亦可實作於非分頁型作業系統。

在分頁作業系統中，「虛擬位址空間(virtual address space)」，亦即由一程式可定址的位址集合，被



五、發明說明 (12)

切分為頁，而頁為連續虛擬位址的集合且具有固定長度。一般來說，一頁包含4KB。程式的虛擬位址空間可能比可獲得的實體記憶體來得大。作業系統提供功能集合以支援此種技術特徵，而這些功能集合通稱為「虛擬記憶體管理員(virtual memory manager)」。為了支援比實體記憶體更大的虛擬記憶體空間，虛擬記憶體管理員在記憶體及在硬碟等第三級儲存裝置(tertiary store)中，存放虛擬頁。當一虛擬頁被存取，而此虛擬頁不在主記憶體中時，則從磁碟(頁內操作(page-in operation))中讀取此虛擬頁。假如沒有實體空間(physical space)供從磁碟讀取此頁時，則另一虛擬頁就被寫入磁碟(頁外操作(page-out operation))，並且釋放其原先佔據的空間。當一虛擬頁被從磁碟讀入，此虛擬頁被指定一起始實位址(亦即，從處理器所見到的位址)。實記憶體(此處理器的位址空間)則被分割為連續及成對不相交實位址區之集合，其具有與邏輯頁相同的大小。這些稱為「頁框(page frame)」。因此，當一邏輯頁被從記憶體讀取時，此邏輯頁被儲存在頁框中。在邏輯頁及真實頁間的轉譯有賴於分割頁的目錄結構，稱為「頁表(page table)」。每一邏輯頁在一頁表中具有單一項，稱為「頁表項(page table entry)」，其可能包含該頁框的起始實位址，其中該頁框包含該頁，而假如邏輯頁係存放在第三級儲存裝置時，頁表項亦可能包含磁碟之位置。自由頁框(free page frame)透過一個別的资料結構將以管理，此資料結構(在Windows NT及2000)稱為頁框號



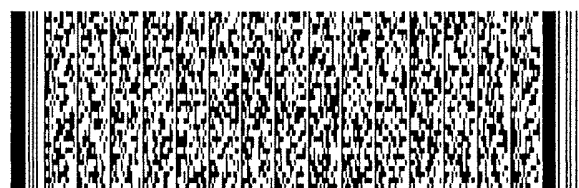
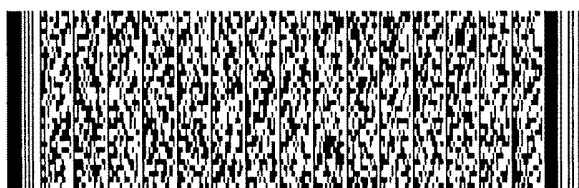
五、發明說明 (13)

碼資料庫(page frame number database)。

第4圖顯示具有本發明特徵之網路架構。複數個電腦主機401連接到一SCMS 403，而此SCMS 403透過一交互連接網路402以支援記憶體壓縮。其中，電腦主機401可為個人電腦、工作站、對稱多處理器機器(Symmetric Multiprocessor, SMP)、非單一記憶體存取機器(non-uniform memory access machines, NUMA)、純快取機器(Cache-only machine, COMA)、儲存伺服器，或SP節點(SP nodes)等等。

SCMS 403支援記憶體壓縮，並且提供電腦主機401分享一自由記憶體段集合，其中包含記憶體壓縮/解壓縮，在實位址與實體位址間進行轉譯，提供保護(亦即，避免一主機存取分配給其他主機的記憶體段)，負載平衡，效能分離(performance isolation)(亦即，將主機401對於其他主機的影響降到最低)等等。

第5圖顯示SCMS 403的基本架構。SCMS 403包括裝置集合501及記憶體502。裝置集合501包括網路卡503、管理記憶體502的控制邏輯504、及第三級儲存裝置505。管理記憶體的control邏輯504更包含記憶體壓縮器/解壓縮器、暫存器、包含目錄或壓縮變換表(CTT)項目的快取、定址電路、從記憶體的讀取及寫入、緩衝區等等。第三級儲存裝置505

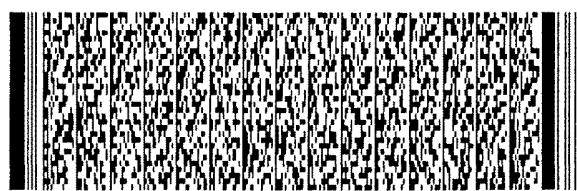
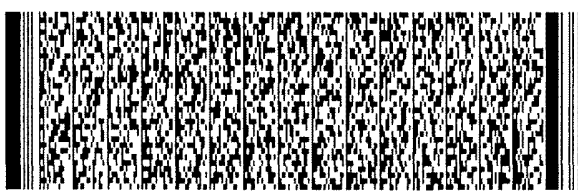


五、發明說明 (14)

包括唯讀記憶體(ROM)、電程式化唯讀記憶體(EPROM)、快閃記憶卡，及硬碟。在較佳實施例中，SCMS 403也與一個或多個主機401分享一個或多個硬碟機。

上述之記憶體502邏輯上分為記憶體目錄區域506(也被稱為壓縮變換表，CTT)、一直接映射區域507(舉例來說，位址以相對一暫存器的偏差值予以計算)、一未壓縮區508，以及一記憶體段池(a pool of memory sectors) 509。依據本發明之精神，每一個這些區域不需要展開成位址範圍(address range)，但也能展開成多個位址範圍，在後者之例子中，每一此類區域也可認為由複數個區域組成。

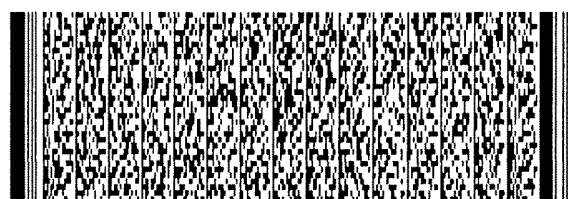
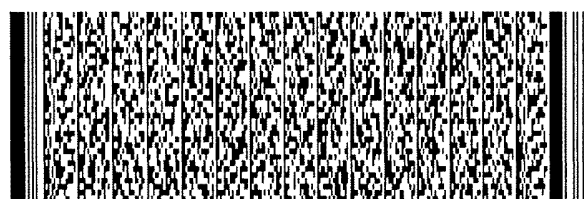
在較佳實施例中，SCMS 403以主機401的實記憶體空間的延伸方式進行操作。更精確地來說，主機401可指示SCMS 403提供實位址之範圍。當主機401需要存取(讀或寫)在實位址的範圍內之記憶體位置時，主機401要求SCMS 403進行所需運作。在較佳實施例中，為了使SCMS的硬體複雜性保持在合理的範圍，記憶體502可服務到H個不同的主機。對於想要使用SCMS的主機，如果SCMS目前服務的主機數目少於H台，則將自己登錄到SCMS，且SCMS分配給該主機一唯一識別碼，舉例來說，介於0到H-1，並且登錄程序因此而完成。



五、發明說明 (15)

在較佳實施例中，主機401可在SCMS中，以一預定的基本單位(granularity)分配實記憶體範圍，而此處的基本單位即包含連續實位址之固定數目，在此將簡稱為區段。區段大小用L代表，並且SCMS所支援的區段之數目則用S代表。舉例來說，在支援32Gb實記憶體的SCMS中， $H=16$ 而 $L=1\text{Gb}$ 。在此例中，SCMS所提供的實記憶體空間切分為32區段，且每一主機401可以分配1到17區段(此時，剩下的 $H-1$ 個主機僅僅使用一區段)。在一實施例中，SCMS限制區段數目，使得單一主機可分配到最大為N的區段，以避免一主機因為使用過多資源而影響到其他的主機。舉例來說，SCMS可限制主機使用的區段數目為8。假如尚有區段的話，並且主機尚未用到區段之限制，則主機可向SCMS要求區段，而SCMS也會配給此主機所要求之區段。主機也可以釋放區段，而釋放之區段可加入一自由區段池。

在較佳實施例中，當一主機分配一記憶體區段時，SCMS通知主機對應的實位址範圍。在此實施例中，假設使用64位元定址，位址的 $\log_2(H)$ 個最大位元為主機識別碼；在此例中， $H=16$ ，則前4個最大位元為主機識別碼。接下來的 $\log_2(N)$ 個位元(在此例中為3)辨識包含此位址的區段，而 $\log_2(L)$ 個最小位元則為區段內的偏移值。此種方法無須辨識定址位元(此例為27)的龐大數目。上述方法也可用在其他場合，舉例來說，假如主機401向多個SCMS登錄，並且獲得相同的識別碼。在這樣的情形中，非特定之位元也可

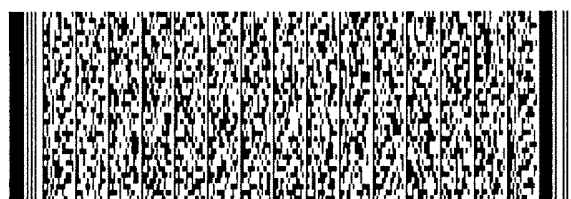


五、發明說明 (16)

用來區別不同SCMS所分配的區段，以下將對此提供說明。

如上所述，在支援記憶體壓縮的系統中，實位址與實體位址(physical address)之間的轉譯須依賴稱作CTT的目錄。依據本發明之精神，SCMS CTT 506分割成連續部分，每一部份用來進行一個區段的實-到-實體(real to physical)的轉譯。連續部分的數目因此等於SCMS所管理的最大數目之區段，並且每一連續部分的大小足以使其包含能夠定址一區段的CTT項之數目。在較佳實施例中，CTT的大小係由初始載入程式(Initial Program Load, IPL)所決定(亦即當機器開機時)。舉例來說，假如SCMS包含16Gb的實體記憶體，且支援32Gb的實記憶體，並且壓縮的單位為1K，則CTT將包含 32×2^{20} 個項目，並分為32個連續的部分，其中每一部份包含 2^{20} 項，以對應到每一個不同的區段。

本發明之一特徵為需要可供控制每一主機能使用多少的實體記憶體的硬體及/或軟體元件，此特徵無法見於支援記憶體壓縮之一般電腦、提供記憶體伺服器功能的裝置、或是不支援記憶體壓縮的裝置。此特徵係描述於第6圖。當一主機要求一區段，此主機也可要求複數個實體段，而此複數個實體段可視為分配或保留給區段之保證數量之實體段。假如保留段的總數目未超過段的總數目，則要求就被接受。

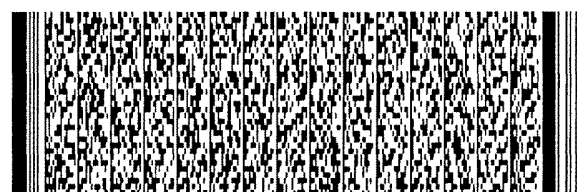
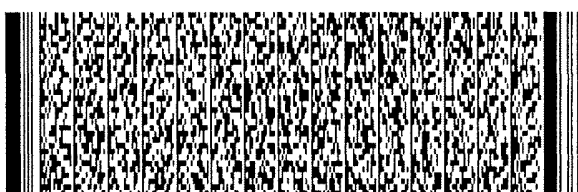


五、發明說明 (17)

SCMS 包含暫存器組，以存放定限值601及計數器602。分配給每一主機之實體段的最大數目儲存在專屬的定限暫存器，並且每一主機使用的實體段數目則存在專屬的計數暫存器。當主機進行寫入或釋放的記憶體操作時，它的主機識別碼603用來選擇關聯的定限值與計數器。假如操作為一釋放(release)時，計數器將減去主機所釋放的實體段數目，此實體段數目被主機所釋放之實位址所使用。假如操作為一寫入時，壓縮器/解壓縮器(604)將計數器減去一實體段數目，此數值即在寫入操作之前，對於每一記憶體線經由該寫入操作所造成之影響而決定之數值。壓縮器接著壓縮每一線，並且增加計數器之數值，增加之數值即該線所使用之實體段的數目。比較器605比較定限值與計數器之值。當計數器之值超過定限值時，比較器605產生一比較結果606，以顯示該主機已經使用超過被配給實體段之配額。否則，比較器產生一比較結果，顯示尚未超過定限值。

經由一(軟體或硬體)機制，當使用之區段數目超過一定限值時，比較結果606被用以通知一個或多個主機；此定限值可透過軟體設定為接近、但小於保留段計數器。此機制容許主機在接近極限值時，可以減少其使用之段數目。當主機未能成功減少段數目，則主機寫到SCMS之儲存操作(寫入)將會導致其使用段超過保留段之數目。在此種情形，有若干方案可供選取：

(1) SCMS可避免寫入操作之完成，並且發出一錯誤信



五、發明說明 (18)

號給主機。此可避免從主機，如主機0，進行之接續寫入操作，直到使用段數目降到此定限值以下才繼續進行。此外，不管主機0的壓縮比降級到多差的程度，仍容許其他主機繼續運作。

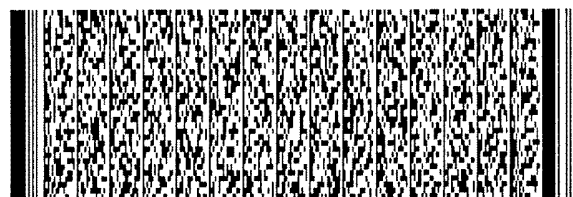
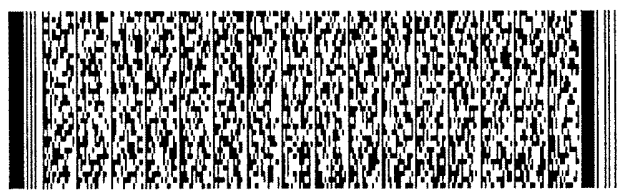
(2) 假如SCMS仍有足夠多的未使用段時，SCMS可同意此寫入操作。SCMS也可以發出另一種錯誤信號到主機，此可視為高優先權之訊息，以降低其段的使用。在此種情形，其他主機可繼續運作，然而，假如主機0的壓縮比持續衰減，則段最後仍有可能會用完。

第7圖顯示一較佳實施例，說明如何將一主機所指定之實位址轉譯成SCMS的實位址。主機指定實位址701分割為主機識別碼702、區段號碼703及偏移值704，以及部分可忽略的位元。如第5圖所示，SCMS的控制邏輯504包含一暫存器矩陣705，亦即資料持有位置。矩陣705的每一列對應到一不同的主機識別碼，因此矩陣705包含H列。矩陣的每一行對應一不同的區段號碼，因此矩陣包含L行。主機識別碼702用來選擇矩陣的列706，並且區段號碼703用來選擇行707。行707與列706一起辨識暫存器709。暫存器709的內容，亦即CTT用來定址需求區段的部分之開始點，則與偏移值704結合在矩陣705外的暫存器710，此係將包含在暫存器709的位元接續到偏移位元704。暫存器710現在包含SCMS的實位址，換言之，為被記憶體502所管理的需求資料的實位址。



五、發明說明 (19)

在另一實施例中，主機識別碼並非位址之一部份，而係從網路卡等處獲得。在此種實施例中，如第8圖所示情況，主機401所提供的預定部分802，亦即區段號碼，用來指定需求區段(因此，此部份長度為 $\log_2(S)$ 位元)，此時位址801的最低有效位元803為偏移值。習知技藝者當知，包含區段號碼的位址801之部分802的預先指定位置，不需要由鄰接位元構成，並且除了偏移值803使用的位置外，這些位元可以放在位址801的任意位置。主機識別碼805係從需求位址分別提供。如第5圖所示，SCMS的控制邏輯504包含暫存器804陣列(資料持有位置)，其中每一區段對應一暫存器，每一暫存器對於每一支援主機包含一位元。當一主機分配一區段，其指定何主機能夠讀取及寫入該區段，此係藉由對暫存器的相關位元設定為1而做到，此時其他的位元則設為0。當讀取/寫入操作進行，區段號碼802用來選擇暫存器陣列804的暫存器，並且關於主機識別碼805的項目之位元被檢測。假如位元設定為1，則產生致能信號806，其使得位址807可以被使用。位址807係由區段號碼802及偏移值803所組成。假如位元設為0，產生一錯誤信號，並且將之傳送給要求之主機。此種方法衍生出一種保護機制，以避免主機讀取或寫入其他主機所使用之記憶體區段。其亦提供一機制，以在多部主機401分享記憶體空間之部分下，對於此分享記憶體環境中提供存取控制。



五、發明說明 (20)

藉由第6圖所示之方法，控制邏輯504可對於每一主機包含一計數器，其包含此每一主機使用之區段數目，並且避免一主機被配給超過其被容許之區段數目。習知技藝者當知，如何將第8圖所示的保護機制整合到第6圖的架構中。

本發明也提供一種方法，無須使用暫存器陣列，其透過第7圖所示之方法建構一保護機制。如第7圖所示，位址中的主機識別碼部分個別由第8圖所提供(舉例來說，使用一變換表供轉譯從網路卡取得的主機網路位址，成為主機識別碼)。接著，對於一區段排他性的存取，由於識別碼702為位址801的一部份，而主機識別碼805獨立於位址之外，因此能夠藉由比較第7圖中主機識別碼702與第8圖的主機識別碼805，而輕易實作此功能。假如找不到比對結果，則表示該位址為一無效位址，並會發出一錯誤訊息。為了支援區段的分享，當分配一區段時，也指定將分享此區段的主機識別碼，SCMS找出對應指定之主機的轉譯矩陣705之列中未使用的項目，並且清空此項目，及拷貝分配區段的起始點位址到所有的項目中。此種方法可以對每個分享相同段的主機，衍生出不同的實位址。對於轉譯矩陣705之項目，SCMS控制邏輯不容許使用中區段之起始位置被寫入，因此可謂此種方法提供了保護機制。

假如主機識別碼非位址的一部份，但卻另外提供，則

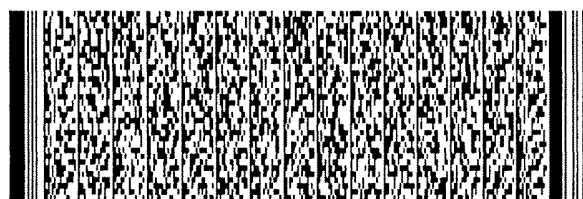


五、發明說明 (21)

第7圖所描述之方法可進一步修正，使得轉譯矩陣的列由另外提供的主機識別碼中選擇(也就是其不再為位址701的一部份)，而行係使用區段識別碼703選出。此方法也自動提供一保護機制，並且可以用來控制上述段落中，對於段的分享存取控制。

在另一實施例中，要求一區段之主機401也指定對於此區段之起始實位址，其與區段之邊界對齊。在此實施例中，SCMS包含邏輯器，供確認主機無法分配兩個不同的區段但指向相同之起始位址。第9圖顯示在主機所提供之實位址與SCMS實位址之間的轉譯。在此情況中，第5圖中的控制邏輯504包含一相聯記憶體904(資料持有位置)。此相聯記憶體904對於每一管理區段包含一列(這些管理區段包括那些目前尚未配置者)，以及兩行。第一行905包含複數個鍵值，其為主機401對於配置區段所提供之起始實位址。第二行906包含複數個數值，其為SCMS實記憶體空間的區段起始實位址。

當一主機401提供一位址901時，此位址被分為兩部分，即在區段內的偏移值903，其由最低有效位元構成，以及最高有效位元(the most significant bit)902，其被相聯記憶體904當作一搜尋鍵值使用。假如其中一列包含鍵值欄位之值與此搜尋鍵相同時，例如列907，則此相聯記憶體904傳回值欄位908之內容。值908接著結合此偏移值903到



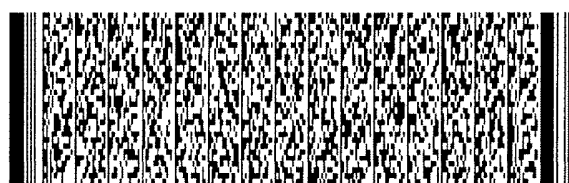
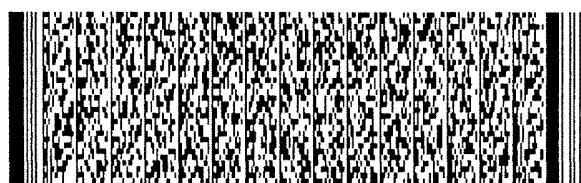
五、發明說明 (22)

暫存器909，藉以產生SCMS的實位址。對於習知技藝者應清楚，如何將第9圖的機制與第8圖的保護機制結合在一起。

假如SCMS管理大量數目的區段，第7圖的轉譯矩陣705、第8圖的暫存器804、第9圖的相聯記憶體904，以及那些從上述所衍生者，如果用第5圖所示之控制邏輯504之暫存器方法予以實做，將會造成相當高的製作成本。在較佳實施例中，包含在元件705、804、與904的資訊儲存於記憶體502中。控制邏輯504包含快取，而此快取只存有前述儲存資訊的一部份，以備使用。這些快取可以透過已知之快取管理方案予以管理，例如在快取中維持最近被使用的區段資訊。

本發明亦顯示如何將實記憶體空間，在多個SCMS間展開。第10圖顯示主機1001與多個SCMS 1003透過交互連接網路1002連接，並且部分主機的實記憶體空間散佈在SCMS間，使得每一記憶體1004包含複數個實位址。第9圖所示之方法可以用來確保主機1001的實記憶體空間之一連續子集合分布在SCMS 1003間。

在較佳實施例中，SCMS也包含連接到相同網路的其他SCMS之資訊，例如包括它們的號碼、每一個別SCMS的位址，以及，對於每一SCMS，分配區段的數目與壓縮資訊。

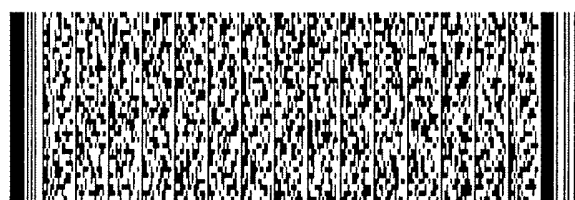
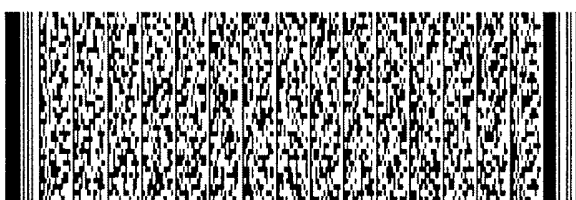


五、發明說明 (23)

可以理解的是，本發明可以實作於各式硬體、軟體、韌體、特殊用途處理器，或其各種組合。在一實施例中，本發明可實作於軟體成為一應用程式，其實際存於一程式儲存裝置中。此應用程式可以被上傳到一包含適當架構的機器，或由一包含適當架構的機器以執行。在較佳的情況中，此機器係以電腦平台實作，而具有中央處理單元、隨機存取記憶體，及輸出/輸入介面。此電腦平台也包括作業系統及微指令碼。在此描述的不同程序及功能可為微指令碼的一部份或應用程式的一部份(或其結合)，其可經由作業系統執行。此外，各種不同的周邊裝置可連接到此電腦平台，例如額外的資料儲存裝置與列印裝置。

也可以理解的是，由於附屬圖示中所描述的一些可行的系統元件與步驟可用軟體實作，因此，在系統元件間的實際連接(或程序步驟)可依據本發明程式撰寫的不同而發生差異。由於此處本發明的教示，習知技藝者將能夠斟酌這些及相似的實作，或是本發明的各種配置。

第11圖顯示主機系統如何將記憶體參照(memory reference)轉譯成對於SCMS的要求。在步驟1101，主機系統辨識何時實位址存於一網路連接裝置。在較佳實施例中，邏輯-到-實體(logical to physical)位址轉譯包含辨識存於SCMS的位址之部分。在此實施例中，頁表的適當旗標(flag)用來區別其係存於主機的主儲存裝置之頁或是存

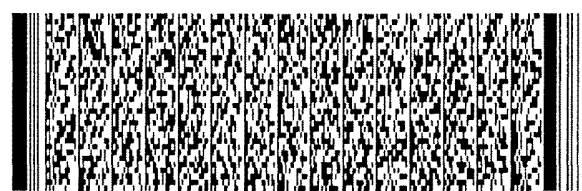
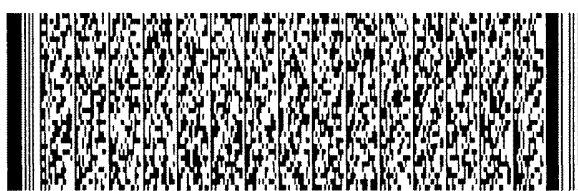


五、發明說明 (24)

於外部之頁。存於外部之頁接著經由一驅動程式或一硬體裝置存取，其係將此記憶體參照轉譯成一對網路連接裝置的要求。此轉譯將包括步驟1102，其中包含資料的SCMS網路位址將在位址目錄1103中，此位址目錄1103被存入複數個SCMS，以及被存在步驟1104中被要求的SCMS，此被要求的SCMS在步驟1102中被辨識。

在另一實施例中，存於SCMS的頁具有的實位址不屬於主機記憶體所支援的實位址。舉例來說，假如主機不支援記憶體壓縮，則主機記憶體所支援的實位址將從零到等於實體記憶體大小之最大值。假如主機支援壓縮，則實位址將由實-到-實體轉譯機制決定其範圍；假如該轉譯機制依賴一靜態目錄(CTT)，實位址的範圍由該目錄決定。在現今的作業系統中，當一實位址從主機所支援的實位置的範圍外所發出，則將會產生例外錯誤(exception)，並且也會產生中斷。在目前的實施例中，當作業系統模組因為此類中斷而被呼叫時，作業系統模組包含程式碼以執行1102及1104之作業。

在另一直接定址分享壓縮記憶體系統(SCMS)的實施例中，SCMS包括一CTT的集合，其中每個CTT對應一不同的實位址空間。在此實施例中，未被主機分享的位址空間可以連接到最多一個主機。當一個新的主機要求一個非分享之位址空間，並且實的位址空間仍可被獲得時，SCMS的控制

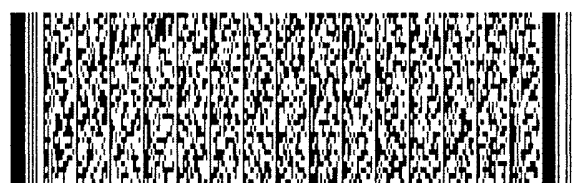
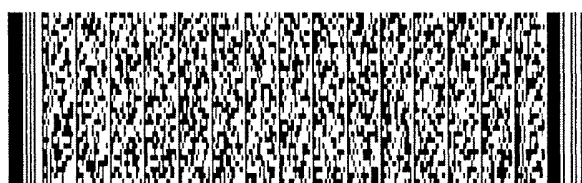


五、發明說明 (25)

邏輯指定一可用實位址空間與對應的CTT，並且將此位址空間與此主機識別碼關聯起來。當主機記憶體操作被導向SCMS時，SCMS將主機識別碼轉換成對應CTT的位址，並且將主機所提供之實位址轉譯成使用CTT的實體位址。

第12圖說明此實施例之一範例。SCMS 1202包括，但不限於，CCT集合1207以及記憶體1211。此記憶體1211更包括分享記憶體段1209、1210的集合。在運作時，主機1201對SCMS 1202發出一記憶體操作，並且送出一位址1203。SCMS從位址1203取出主機識別碼1204及由主機產生的實位址1206。接著，主機識別碼1204配合一主機識別碼-到-CTT變換表1205，從CCT的集合1207中選擇一個CTT。使用主機識別碼到-CTT變換表1205，SCMS選擇一CTT1208，此CTT1208係關聯於主機1201。選出之CTT1208用來將主機1206所產生的實位址轉譯到包含相對資料的實體位址。在此範例中，包含位址1206的記憶體線，使用兩個記憶體段1209及1210來儲存。

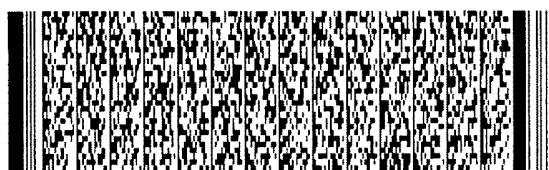
習知技藝者將能夠理解如何將本發明之前述其他特徵結合到此特徵。舉例來說，SCMS可容許多個主機分享關聯於CTT集合1207中之一個CTT的相同區段，此係藉由容許主機識別碼與CTT之間的對應為多對一之對應所達成。在此情況中，多個主機識別碼使用主機識別碼-到-CTT變換表，以將多個主機識別碼對應到單一之CTT。習知技藝者當知，



五、發明說明 (26)

由單一區段使用之實體記憶體數目之各種控制方案，在主機產生之位址與區段之實位址間之轉換等等，也可以應用在本發明之此特徵中。

雖然本發明已以一較佳實施例揭露如上，然其並非用以限定本發明，任何熟習此技藝者，在不脫離本發明之精神和範圍內，當可作各種之更動與潤飾，因此本發明之保護範圍當視後附之申請專利範圍所界定者為準。



圖式簡單說明

本發明上述及其他目的，特徵或優點可透過下列的詳細說明獲得更清楚的解釋，而這些解釋將參照下列圖示為之。

第1圖為一傳統電腦系統之結構，此傳統電腦系統支援壓縮主記憶體；

第2圖為一傳統電腦系統之記憶體系統之詳細結構，此傳統電腦系統支援第1圖所示之壓縮主記憶體；

第3圖為一傳統電腦系統之記憶體目錄之結構，此傳統電腦系統支援第1圖所示之壓縮主記憶體；

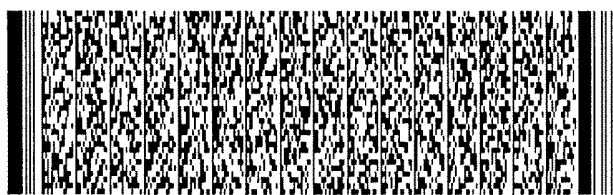
第4圖為一網路架構，此網路架構使用本發明之直接定址分享壓縮記憶體系統(SCMS)；

第5圖為根據本發明之SCMS的方塊圖；

第6圖為根據本發明之控制邏輯的結構，此控制邏輯用以追蹤SCMS中每一主機所使用的段數目；

第7圖為一方法，此方法利用一轉譯矩陣，供將一主機之位址轉譯為被SCMS所管理之位址；

第8圖為一方法，此方法利用具有保護位元之陣列，以提供



圖式簡單說明

對SCMS之區段的保護與存取控制；

第9圖為一方法，此方法利用一相聯記憶體，供將一主機之位址轉譯成被SCMS所管理之實位址；

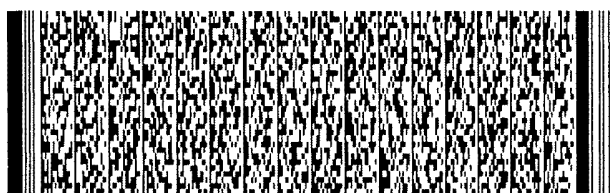
第10圖為一網路架構，其中一主機之實位址空間的一部份分散在多個SCMS間；

第11圖為一方法，供將一記憶體參照轉譯成一向SCMS的要求；以及

第12圖為根據本發明之SCMS之實施例，包括多個記憶體目錄(CTT)。

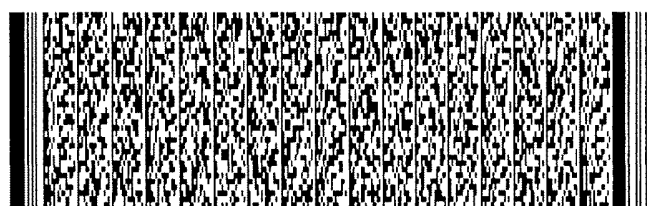
圖式元件符號說明

102	中央處理單元	104	快取
106	壓縮控制器	108	資料
107	目錄	110	壓縮主記憶體
210	壓縮記憶體	220	目錄
221	線項目 1	222	線項目 2
230	資料區	231	記憶體段
232	記憶體段	233	記憶體段
234	記憶體段	235	記憶體段
236	記憶體段	240	處理器快取



圖式簡單說明

241	未壓縮快取線	242	快取目錄
261	壓縮器	262	解壓縮器
270	實位址	271	位址 A 1
272	位址 A 2	273	位址 A 3
274	位址 A 4	301	旗標集合
302	壓縮/未壓縮	303	存於目錄
304	使用區塊數	305	破碎資訊
306	項目 1	307	項目 2
308	項目 3	401	主機
402	交互連接網路		
403	直接定址分享壓縮記憶體系統		
501	裝置集合	502	記憶體
503	網路卡	504	控制邏輯
505	第三級儲存裝置	506	目錄(CTT)
507	直接映射區域	508	未壓縮區
509	記憶體段池	513	線項目 3
514	線項目 4	601	定限值
602	計數器	603	主機識別碼
604	壓縮器/解壓縮器	605	比較器
606	比較結果	701	主機位址
702	主機識別碼	703	區段號碼
704	偏移值	705	矩陣
706	矩陣之列	707	矩陣之行
709	暫存器	710	暫存器



圖式簡單說明

801	主機位址	802	區段號碼
803	偏移值	804	暫存器陣列
805	主機識別碼	806	致能信號
807	S C M S 實位址	901	主機位址
902	最高有效位元	903	偏移值
904	相聯記憶體	905	相聯記憶體之第一行
906	相聯記憶體之第二行	907	相聯記憶體之一列
908	值	909	暫存器
1001	主機	1002	交互連接網路
1003	S C M S	1004	記憶體
1201	主機	1202	S C M S
1203	位址	1204	主機識別碼
1205	主機識別碼-到-CTT變換表		
1206	實位址	1207	CCT集合
1208	選出之CTT	1209	記憶體段
1210	記憶體段	1211	記憶體



四、中文發明摘要 (發明名稱：直接定址分享壓縮記憶體系統)

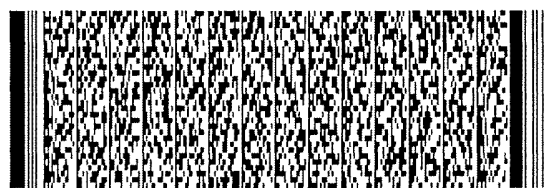
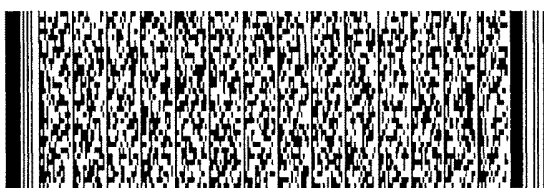
在一電腦系統中，複數個主機經由一交互連接網路連接，且一裝置亦連接到此交互連接網路，經由此裝置以提供此複數個主機分享一記憶體段集合，此記憶體段集合由複數個記憶體段所組成，其中，此複數個記憶體段儲存壓縮資料。此裝置包括一網路卡，供連接此裝置到此交互連接網路，另此裝置亦包括一記憶體，供儲存此記憶體段集合，且此裝置更包括控制邏輯，供管理此記憶體。此控制邏輯包括一記憶體壓縮器/解壓縮器。此記憶體更包括一目錄，供轉譯至少一主機之複數個實位址到此裝置之一實位址。此外，提供一種方法，供管理複數個記憶體段，此複數個記憶體段供每一主機使用，亦提供一種方法，供轉譯至少一主機指定之一實位址到此裝置之一實位址。

五、(一)、本案代表圖為：第 4 圖

(二)、本案代表圖之元件代表符號簡單說明：

六、英文發明摘要 (發明名稱：DIRECT ADDRESSED SHARED COMPRESSED MEMORY SYSTEM)

In a computer system in which a plurality of host is connected through an interconnection network, an apparatus coupled to the interconnection network for allowing the plurality of hosts to share a collection of memory sectors, the memory sectors storing compressed data, is provided. The apparatus includes a network adapter for coupling the apparatus to the interconnection



四、中文發明摘要 (發明名稱：直接定址分享壓縮記憶體系統)

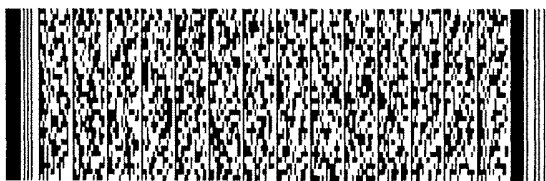
401 主機

402 交互連接網路

403 直接定址分享壓縮記憶體系統(SCMS)

六、英文發明摘要 (發明名稱：DIRECT ADDRESSED SHARED COMPRESSED MEMORY SYSTEM)

network; a memory for storing the collection of memory sectors; and control logic for managing the memory, the control logic including a memory compressor/decompressor. The memory further includes a directory for translating real addresses of at least one host to an address in the apparatus. A method for managing a number of memory sectors used by each host and a method for



四、中文發明摘要 (發明名稱：直接定址分享壓縮記憶體系統)

六、英文發明摘要 (發明名稱：DIRECT ADDRESSED SHARED COMPRESSED MEMORY SYSTEM)

translating a real address specified by at least one host into a real address of the apparatus is also provided.



六、申請專利範圍

1. 一種供複數個主機(host)分享一記憶體段(memory sector)集合的裝置，該記憶體段集合由複數個記憶體段所組成，該複數個主機係在一電腦系統中藉由一交互連結網路連接，每一主機具有一關聯實位址空間(real address space)，該複數個記憶體段儲存壓縮資料，該裝置包含：

一網路卡(network adapter)，供連接該裝置至該交互連結網路；

一記憶體，供存放該複數個記憶體段，該記憶體包含一目錄，供將至少一主機的實位址轉譯為該裝置之一位址，而在該裝置之該位址係供存放壓縮資料；以及

一控制邏輯(control logic)，供管理該記憶體，且該控制邏輯包含一記憶體壓縮器/解壓縮器。

2. 如申請專利範圍第1項所述之裝置，更包含一第三級儲存(tertiary store)。

3. 如申請專利範圍第1項所述之裝置，其中該記憶體更包含一直接映射區域(mapped area)及一未壓縮區域。

4. 如申請專利範圍第1項所述之裝置，更包含一暫存器組，該暫存器組包含複數個定限暫存器供存放每一主機所配給之一最大數目之記憶體段，且該暫存器組亦包含複數個計數暫存器供存放每一主機使用的複數個記憶體段。



六、申請專利範圍

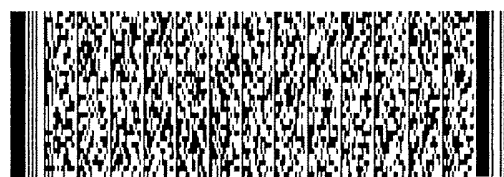
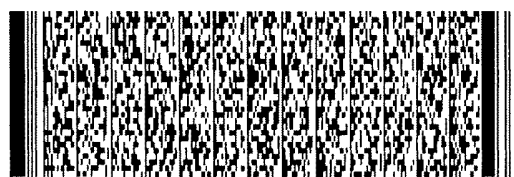
5. 如申請專利範圍第4項所述之裝置，更包含一比較器，該比較器供比較至少一主機之一定限暫存器之值與該至少一主機之一計數暫存器之值，藉此以判斷是否該計數暫存器之值已超過該定限暫存器之值。

6. 如申請專利範圍第1項所述之裝置，其中該至少一主機之該實位址包含該至少一主機識別碼、該裝置之一所需記憶體區段(memory segment)、以及該所需記憶體區段之一偏移值(offset)。

7. 如申請專利範圍第6項所述之裝置，其中該實位址更包含在不同裝置間之分配區段。

8. 如申請專利範圍第6項所述之裝置，其中該控制邏輯更包含一暫存器矩陣，該矩陣之每一列對應一不同的主機識別碼，且每一行則對應一不同的區段號碼(segment number)，藉此，至少一主機之該識別碼及所需記憶體區段決定一特定暫存器的內容，該特定暫存器的內容連接(concatenate)該偏移值，使該實位址被該裝置所管理。

9. 如申請專利範圍第1項所述之裝置，其中該控制邏輯更包含一暫存器陣列，該陣列包含複數個暫存器，對應至該裝置之複數個記憶體區段，其中每一暫存器相對於每一支援主機包含一位元，藉此當該位元為1時，一所需記憶體區段



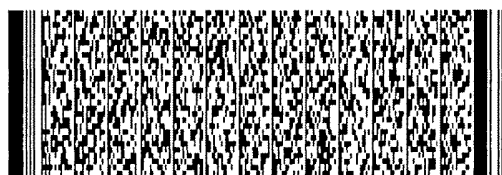
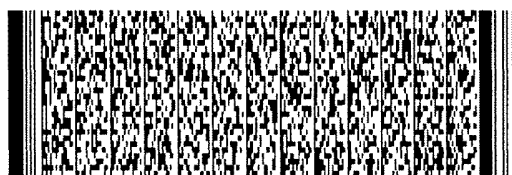
六、申請專利範圍

在連接該偏移值後使該實位址被該裝置所管理，並且當該位元為0時，該至少一主機被拒絕存取該區段。

10. 如申請專利範圍第1項所述之裝置，其中該控制邏輯更包含一相聯記憶體(associative memory)，該相聯記憶體含有一列對應於該裝置所管理之每一記憶體區段，且該相聯記憶體含有一第一行，該第一行含有複數個鍵值(key)以及一第二行含有複數個數值(value)，其中該複數個鍵值係由該複數主機所提供，以對應複數個需求記憶體之起始實位址，且該複數個數值係在該裝置中複數個區段的起始實位址，藉此，該至少一主機之一搜尋鍵(search key)所決定之一數值被接續到一偏移值，該偏移值使該實位址被該裝置所管理。

11. 如申請專利範圍第1項所述之裝置，其中該電腦系統更包含複數個裝置，該裝置包含其他裝置之資訊，例如該複數個裝置之數目、每一該複數個裝置之位址、每一該複數個裝置所配置的複數個分配區段，以及儲存資料的可壓縮性。

12. 如申請專利範圍第1項所述之裝置，其中該記憶體更包含複數個目錄、每一目錄對應每一該複數主機之一不同實位址空間。



六、申請專利範圍

13. 如申請專利範圍第12項所述之裝置，其中該記憶體更包含一主機識別碼對目錄變換表(id-to-directory translation table)，其中至少一主機產生一主機識別碼，用於選擇該複數個目錄中的一目錄，該複數個目錄中的一目錄對應該至少一主機，該目錄轉譯該至少一主機之一實位址為該裝置之一實體位址(physical address)。

14. 一種電腦系統包含一主機以及複數個裝置，該主機經由一交互連接網路連接，每一裝置包括一記憶體段集合，該記憶體段集合由複數個記憶體段所組成，且該每一裝置耦接到該交互連接網路，以容許該主機在該複數個裝置間分享該記憶體段集合，且該複數個記憶體段儲存壓縮資料。

15. 如申請專利範圍第14項所述之電腦系統，其中該主機之實位址之一連續子集合(contiguous subset)分佈在該複數個裝置間。

16. 如申請專利範圍第14項所述之電腦系統，其中該複數個裝置之至少一裝置包含其他裝置之資訊，例如該複數個裝置之數目、每一該複數個裝置之位址、對於每一該複數裝置所配置之一區段數目，以及儲存資料的可壓縮性。

17. 在一電腦系統中，複數個主機經由一交互連接網路連接，該電腦系統包括一裝置耦接於該交互連接網路，供容



六、申請專利範圍

許該複數個主機分享一記憶體段集合，該記憶體段集合由複數個記憶體段所組成，該複數個記憶體段儲存壓縮資料，其中一種方法供管理每一主機使用之複數個記憶體段，該方法包含下列步驟：

決定每一主機所配給之一最大數目之記憶體段；

分配給至少一主機一定限暫存器，供儲存該至少一主機被配給之最大數目之記憶體段；

分配該至少一主機一計數暫存器，供儲存該至少一主機所使用之複數個記憶體段；

比較該至少一主機之一定限暫存器之值與該至少一主機之一計數暫存器之值，以判斷是否該計數暫存器之值超過該定限暫存器之值。

18. 如申請專利範圍第17項所述之方法，更包含：

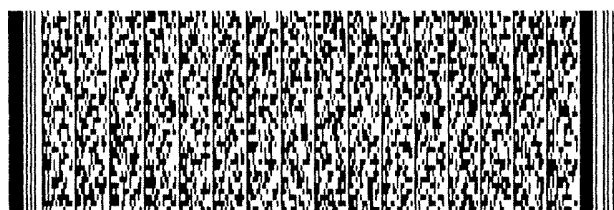
藉由該至少一主機之一識別碼，以選擇關聯該至少一主機之該定限暫存器及計數暫存器。

19. 如申請專利範圍第17項所述之方法，更包含：

當該至少一主機執行一釋放操作(release operation)，從該計數暫存器遞減複數個記憶體段，其中該複數個記憶體段係由該至少一主機所釋放之一實位址集合所使用。

20. 如申請專利範圍第17項所述之方法，更包含：

假使該計數暫存器之值超過該定限暫存器之值，則通知該



六、申請專利範圍

至少一主機其使用之記憶體段數目已經超過其所配給之數目。

21. 如申請專利範圍第17項所述之方法，更包含：

假如該計數暫存器之值超過該定限暫存器之值時，減少該至少一主機其使用之該複數個記憶體段數目。

22. 如申請專利範圍第17項所述之方法，更包含：

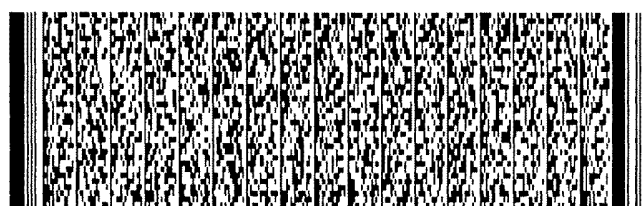
當該至少一主機執行一寫入操作(write operation)時，壓縮待寫入之複數個記憶體線(memory line)，並且增加該複數個記憶體線所使用之複數個記憶體段至該計數暫存器。

23. 如申請專利範圍第22項所述之方法，更包含：

假如該計數暫存器之值超過該定限暫存器之值時，拒絕該至少一主機進行該寫入操作。

24. 如申請專利範圍第23項所述之方法，更包含通知該至少一主機該寫入操作為無效。

25. 在一電腦系統中，複數個主機經由一交互連接網路連接，具有連接到該交互連接網路之一裝置，供該複數個主機分享一記憶體段集合，該記憶體段集合由複數個記憶體段組成，該複數個記憶體段儲存壓縮資料，一種方法將至少一主機所指定之一實位址轉譯到該裝置之一實位址，該



六、申請專利範圍

方法包含下列步驟：

利用一主機指定實位址，由該至少一主機要求該裝置之一需求實位址；

利用該主機指定實位址，辨識該裝置之一資料持有位置(data holding location)；

利用該資料持有位置的內容，以決定一第一數值；以及

連結該第一數值到一偏移值，以獲得該裝置之該實位址。

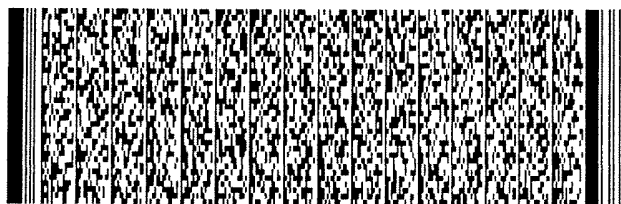
26. 如申請專利範圍第25項所述之方法，更包含從該主機指定實位址分解出一主機識別碼、一區段號碼及該偏移值。

27. 如申請專利範圍第26項所述之方法，其中該控制邏輯為一暫存器矩陣，該矩陣之每一列對應該至少一主機之一不同的主機識別碼，且每一行則對應一不同的區段號碼。

28. 如申請專利範圍第27項所述之方法，其中該辨識之步驟更包含：

利用該至少一主機之該主機識別碼，選擇該矩陣之一列，以及利用該區段號碼，選擇該矩陣之一行，以辨識在該矩陣中之一暫存器。

29. 如申請專利範圍第28項所述之方法，其中決定該第一數



六、申請專利範圍

值之步驟更包含，提供該被辨識暫存器之內容以作為該第一數值。

30. 如申請專利範圍第25項所述之方法，其中該資料持有位置為一暫存器陣列，該陣列含有複數個暫存器以對應該裝置之複數個區段。

31. 如申請專利範圍第30項所述之方法，其中每一暫存器相對於每一支援主機包含一位元，當該支援主機被容許存取該區段時，該位元被設為1，而當該支援主機不被容許存取該區段時，該位元被設為0。

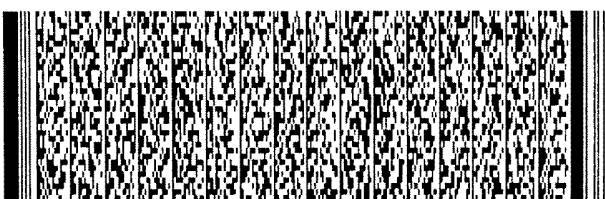
32. 如申請專利範圍第31項所述之方法，其中該辨識步驟更包含，對應從該主機指定實位址擷取出之一區段號碼，辨識該矩陣之一暫存器。

33. 如申請專利範圍第32項所述之方法，其中該決定步驟更包含：

對應該至少一主機所提供之一主機識別碼，檢查該被辨識暫存器中之該位元；以及

假如該位元等於1，提供該區段號碼以作為該第一數值。

34. 如申請專利範圍第32項所述之方法，其中該決定步驟更



六、申請專利範圍

包含，

對應該至少一主機所提供之一主機識別碼，檢查該被辨識暫存器中之該位元；以及

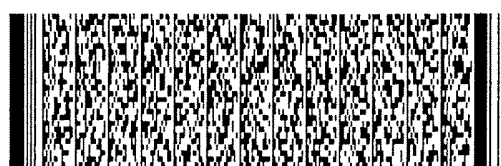
假如該位元為0，則產生一錯誤信號，供送到該至少一主機以辨識不被存取之該區段。

35. 如申請專利範圍第25項所述之方法，其中該資料持有位置係一相聯記憶體，該相聯記憶體含有一列，對應於該裝置所管理之每一記憶體區段，且該相聯記憶體含有一第一行，該第一行含有複數個鍵值以及一第二行含有複數個第二數值，其中該複數個鍵值係由該複數個主機所提供，以對應複數個需求記憶體之起始實位址，且該複數個第二數值係在該裝置中複數個區段的起始實位址。

36. 如申請專利範圍第35項所述之方法，更包含，從該主機指定位址分解出一搜尋鍵值及該偏移值。

37. 如申請專利範圍第36項所述之方法，其中該辨識步驟更包含，利用該至少一主機之該搜尋鍵值，以辨識該相聯記憶體之一列。

38. 如申請專利範圍第37項所述之方法，其中該決定步驟更包含，對應該被辨識列，提供該第二數值為該第一數值。



六、申請專利範圍

39. 在一電腦系統中，包括複數個主機經由一交互連接網路連接，使該複數個主機分享一記憶體段集合，該記憶體段集合由複數個記憶體段所組成，該複數個記憶體段儲存壓縮資料，機器可讀取之一程式儲存裝置(program storage device)，以具體的方式實施由指令組成之一程式，該程式則可由該機器執行以達成管理複數個每一主機所用之記憶體段之方法，該方法包含：

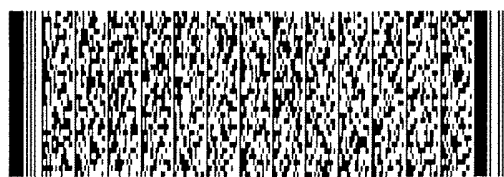
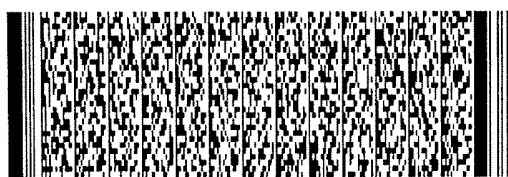
決定每一主機所配給之一最大數目之記憶體段；

分配給至少一主機一定限暫存器，供儲存該至少一主機所被配給之記憶體段之一最大值；

分配該至少一主機一計數暫存器，供儲存該至少一主機所使用之記憶體段之一數目；

比較該至少一主機之該定限暫存器與該至少一主機之該計數暫存器之一數目，以判斷是否該計數暫存器之值超過該定限暫存器之值，其中假如該計數暫存器之值超過該定限暫存器之值時，拒絕由該至少一主機進行該寫入操作。

40. 在一電腦系統中，複數個主機經由一交互連接網路連接，包括連接於該交互連接網路之一裝置，以提供該複數主機分享一記憶體段集合，該記憶體段集合由複數個記憶體段所組成，而該複數個記憶體段儲存壓縮資料，亦包括一機器可讀取之一程式儲存裝置，以具體的方式實施由指令組成之一程式，該機器執行該程式，以達成將該至少一



六、申請專利範圍

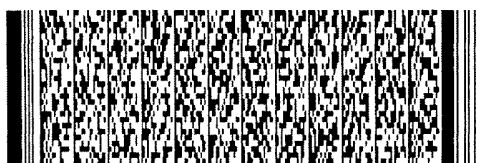
主機指定之一實位址，轉譯成該裝置之一實位置之方法，該方法包含：

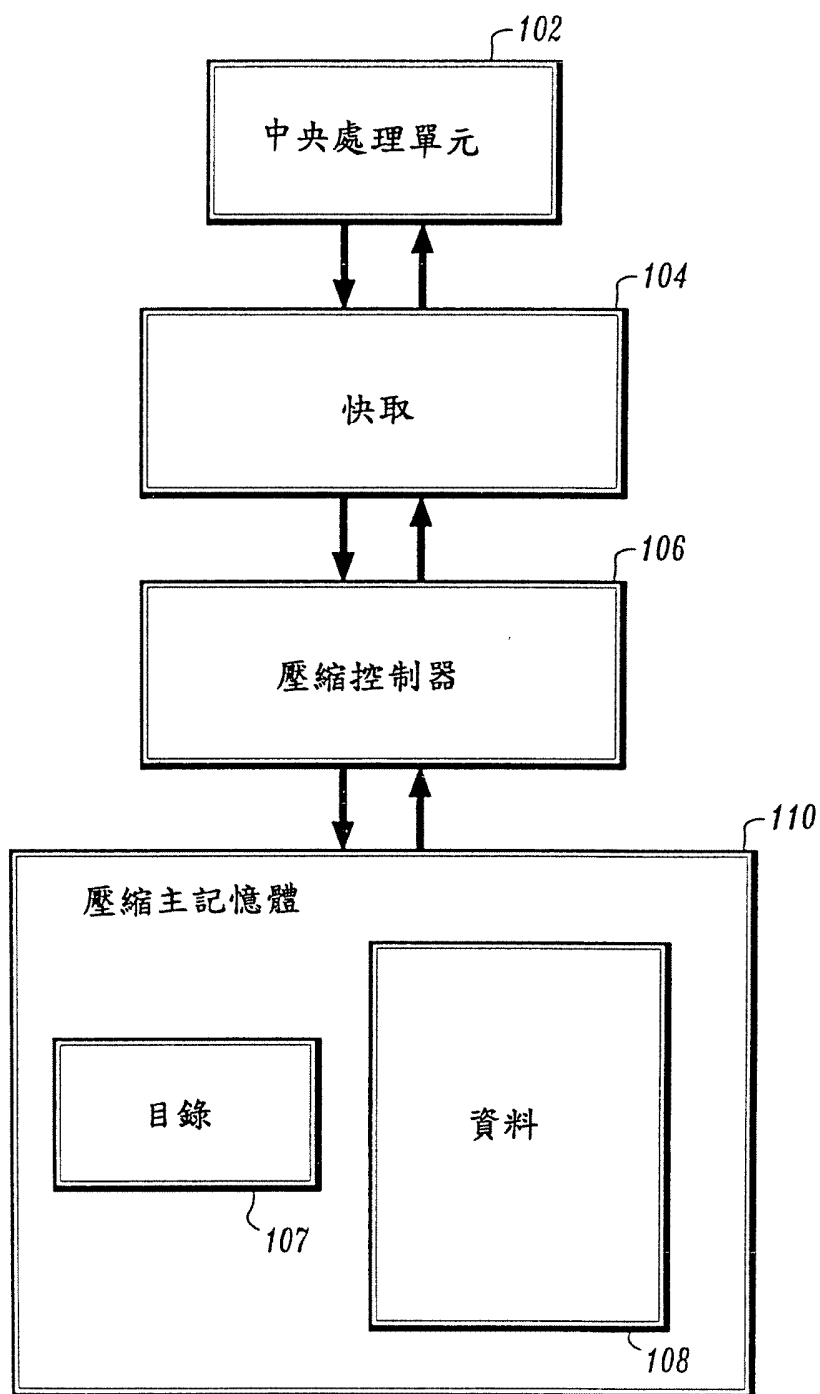
利用一主機指定實位址，由該至少一主機要求該裝置之一需求實位址；

利用該主機指定實位址，辨識該裝置之一資料持有位置；

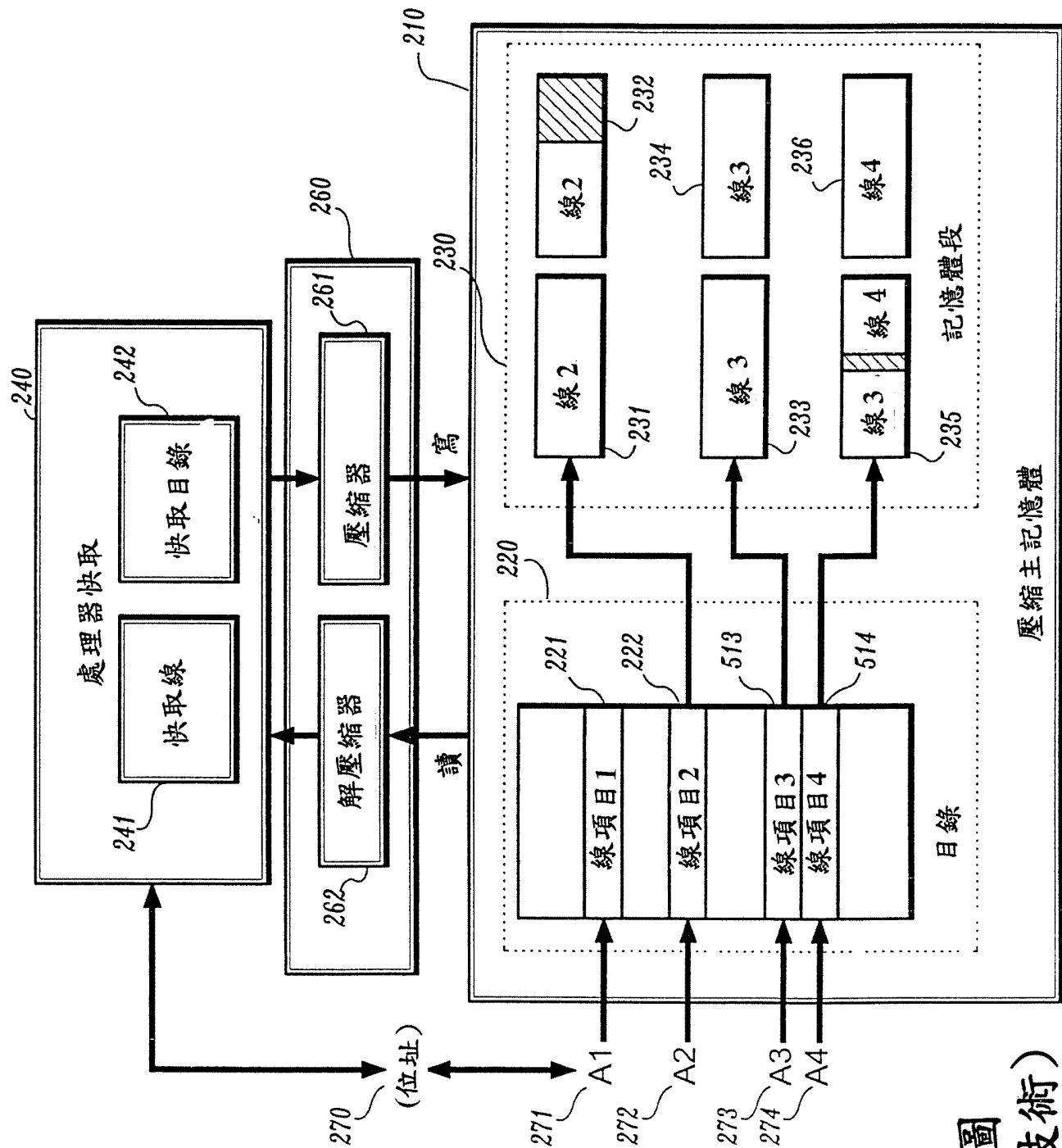
利用該資料持有位置的內容，以決定一第一數值；以及

連結該第一數值到一偏移值，以獲得該裝置之該實位址。

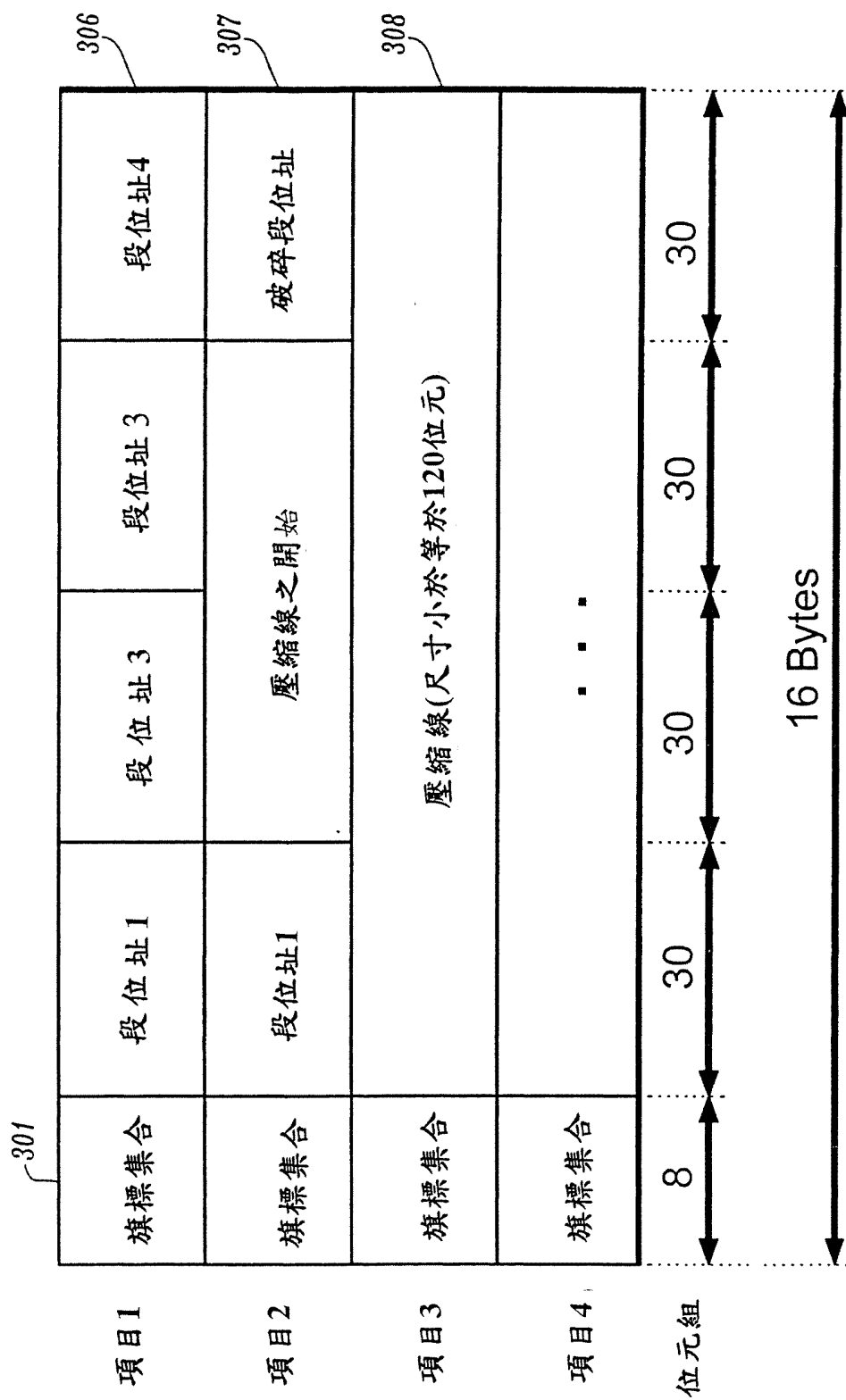




第1圖(習知技術)

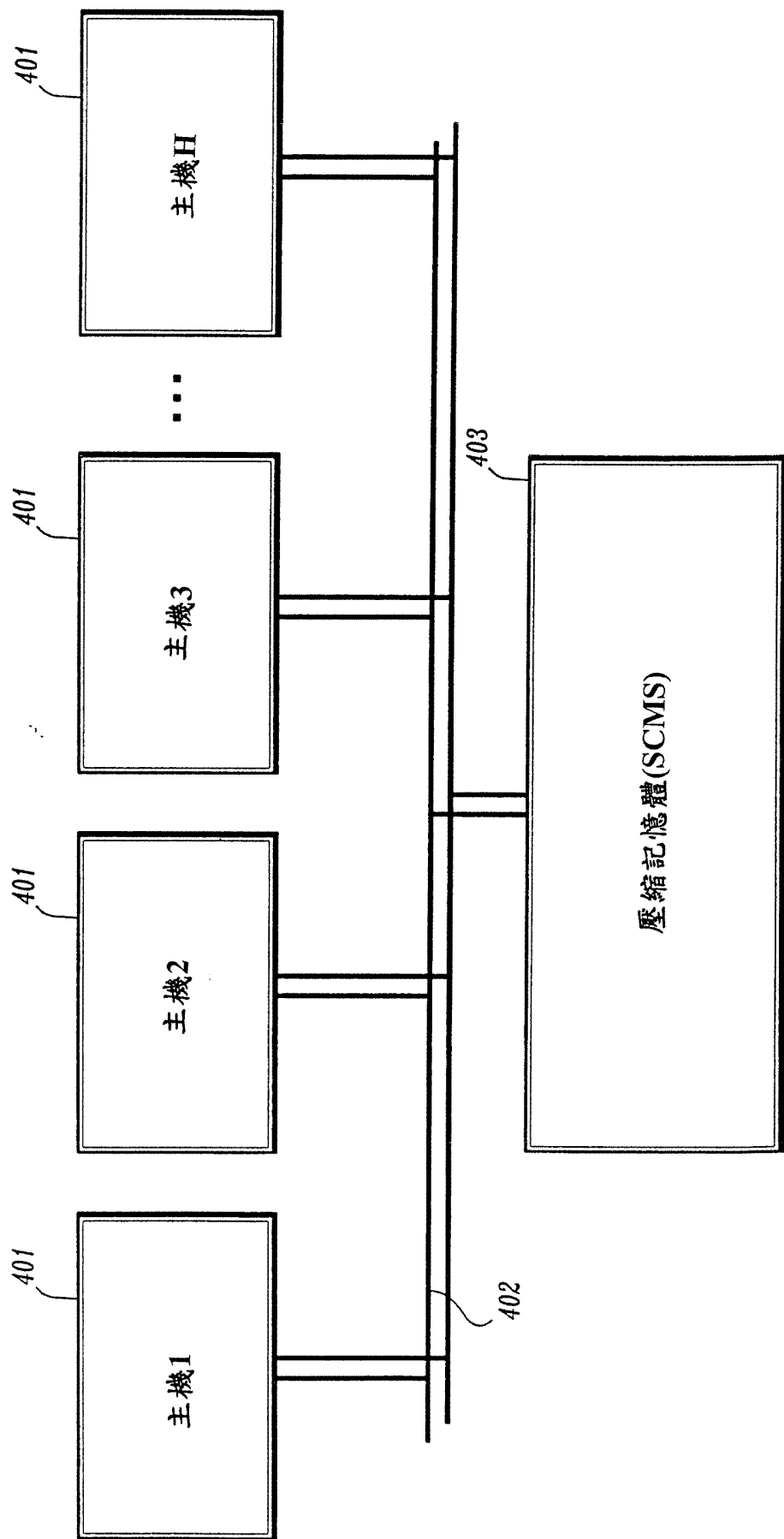


第2圖 (習知技術)



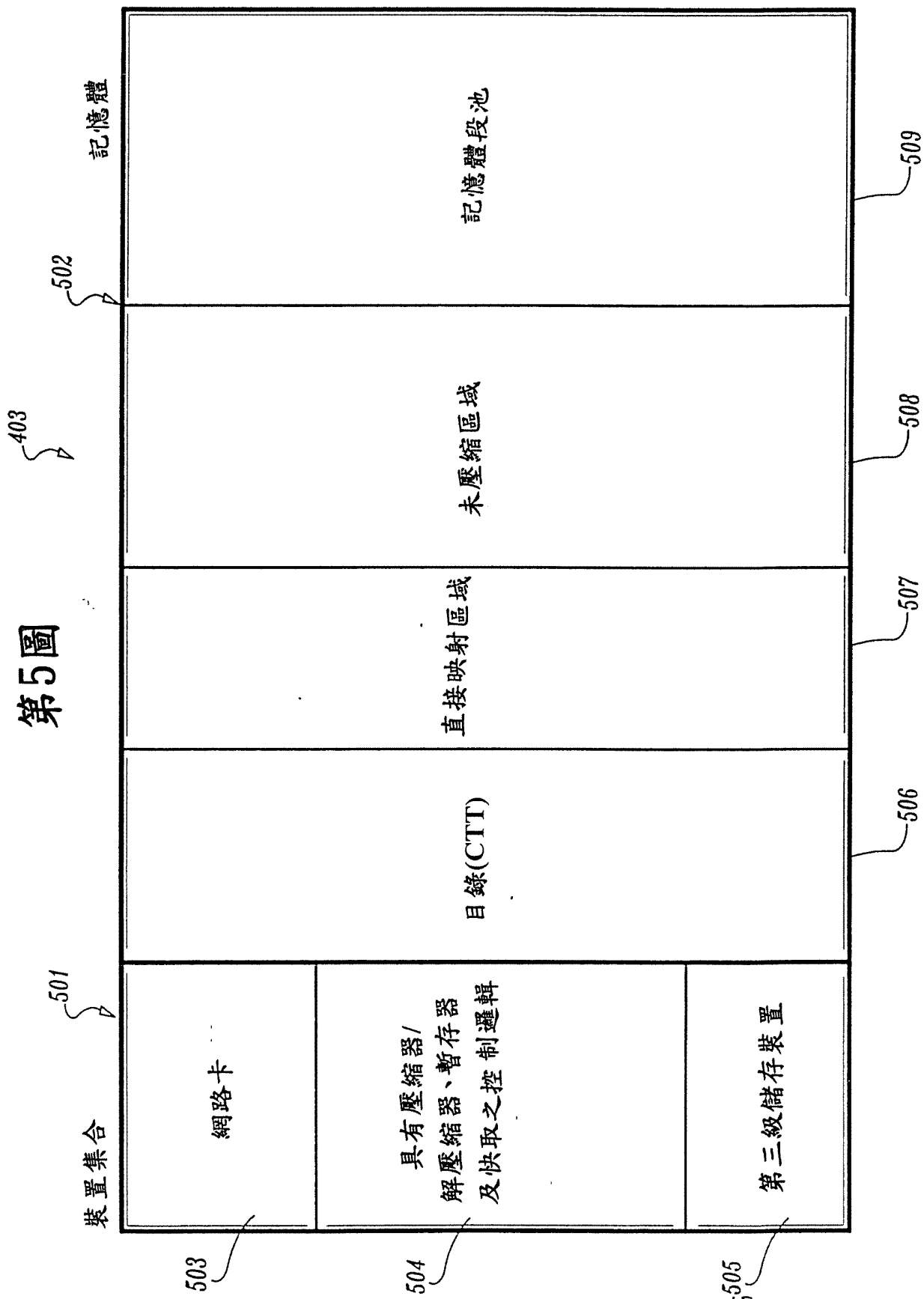
- Flags
- 302 壓縮/未壓縮 (1位元)
 - 303 存於目錄 (1位元)
 - 304 使用區塊數 (2位元)
 - 305 破碎資訊 (4位元)

第3圖
(習知技術)

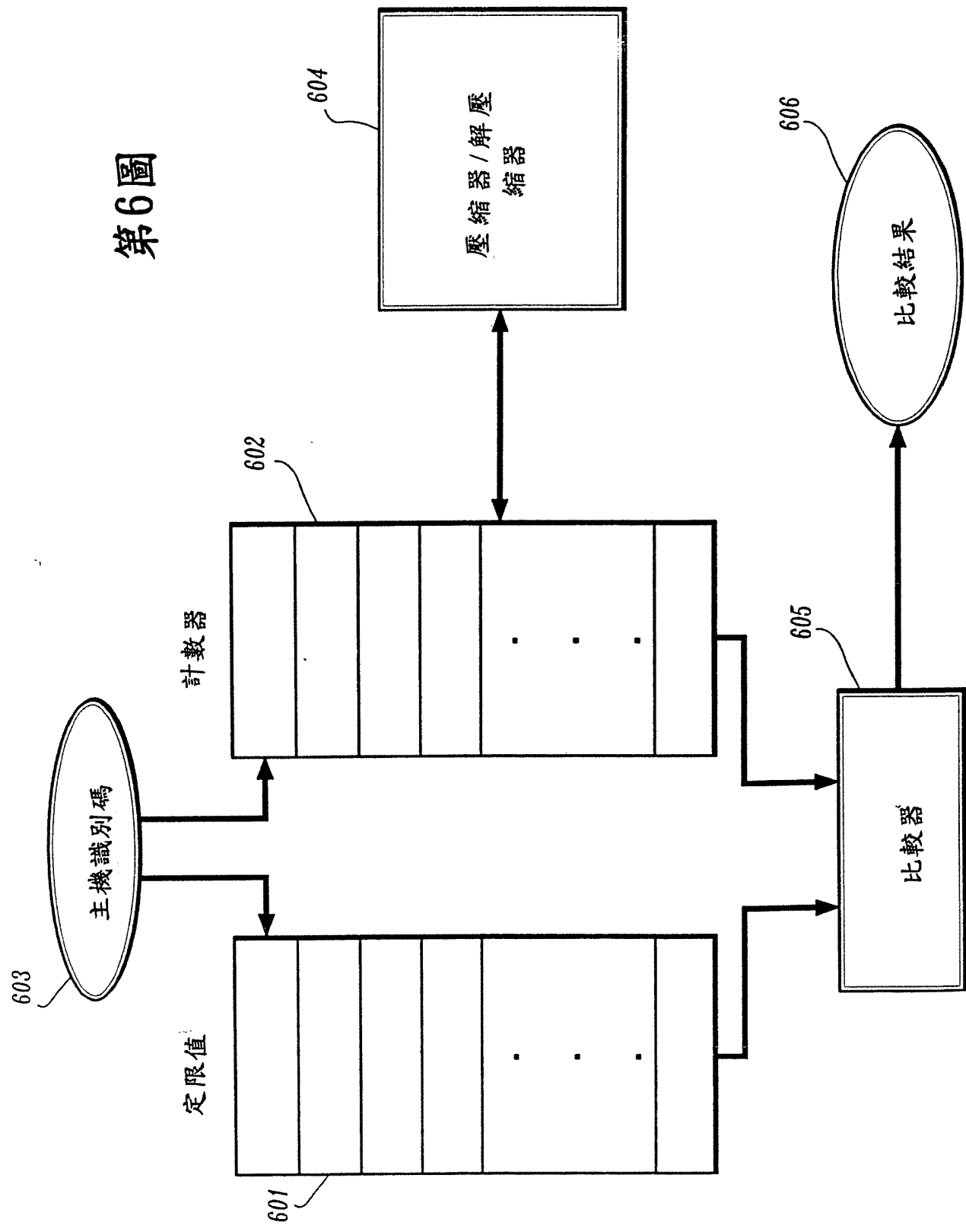


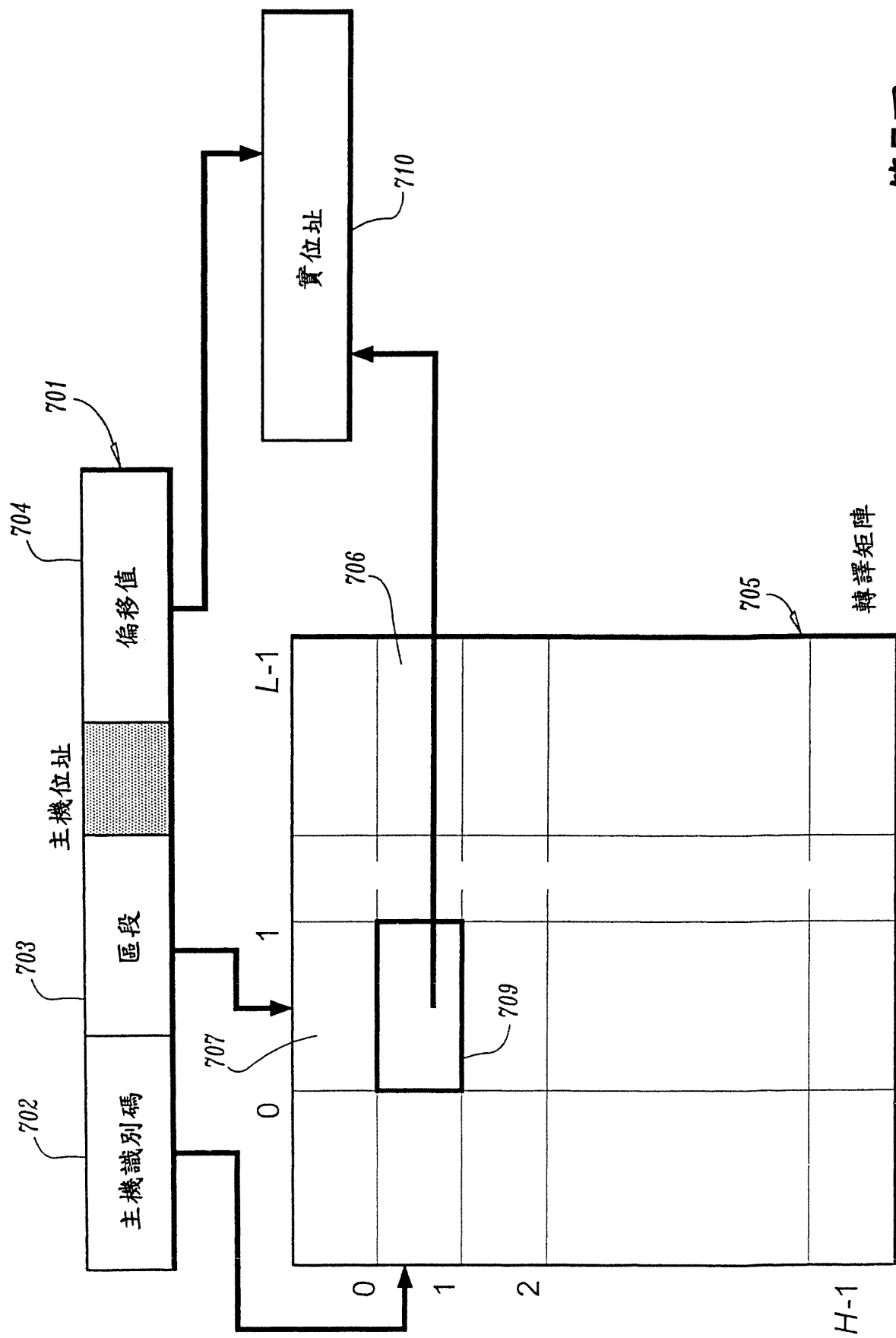
第4圖

第5圖

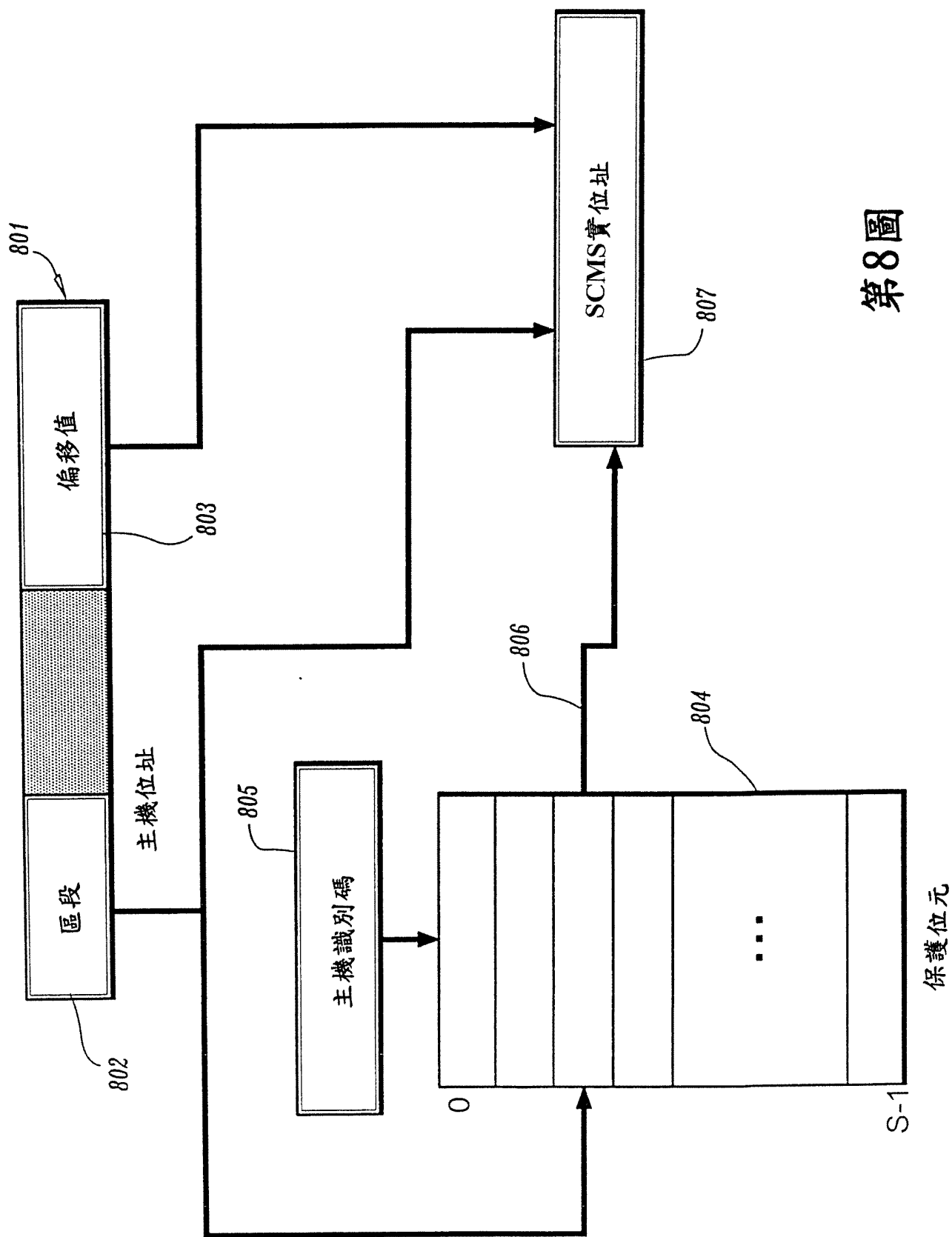


第6圖

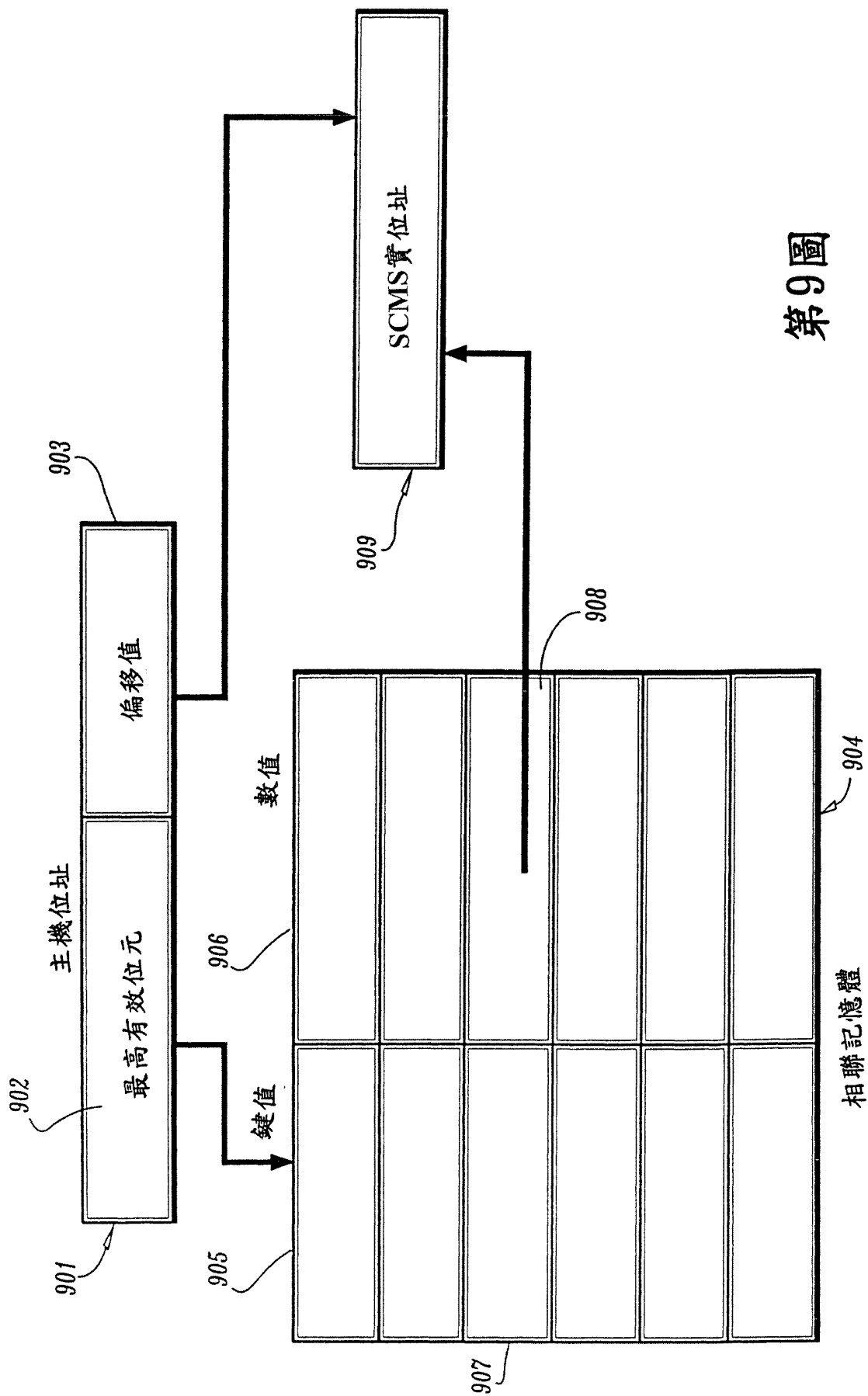




第7圖

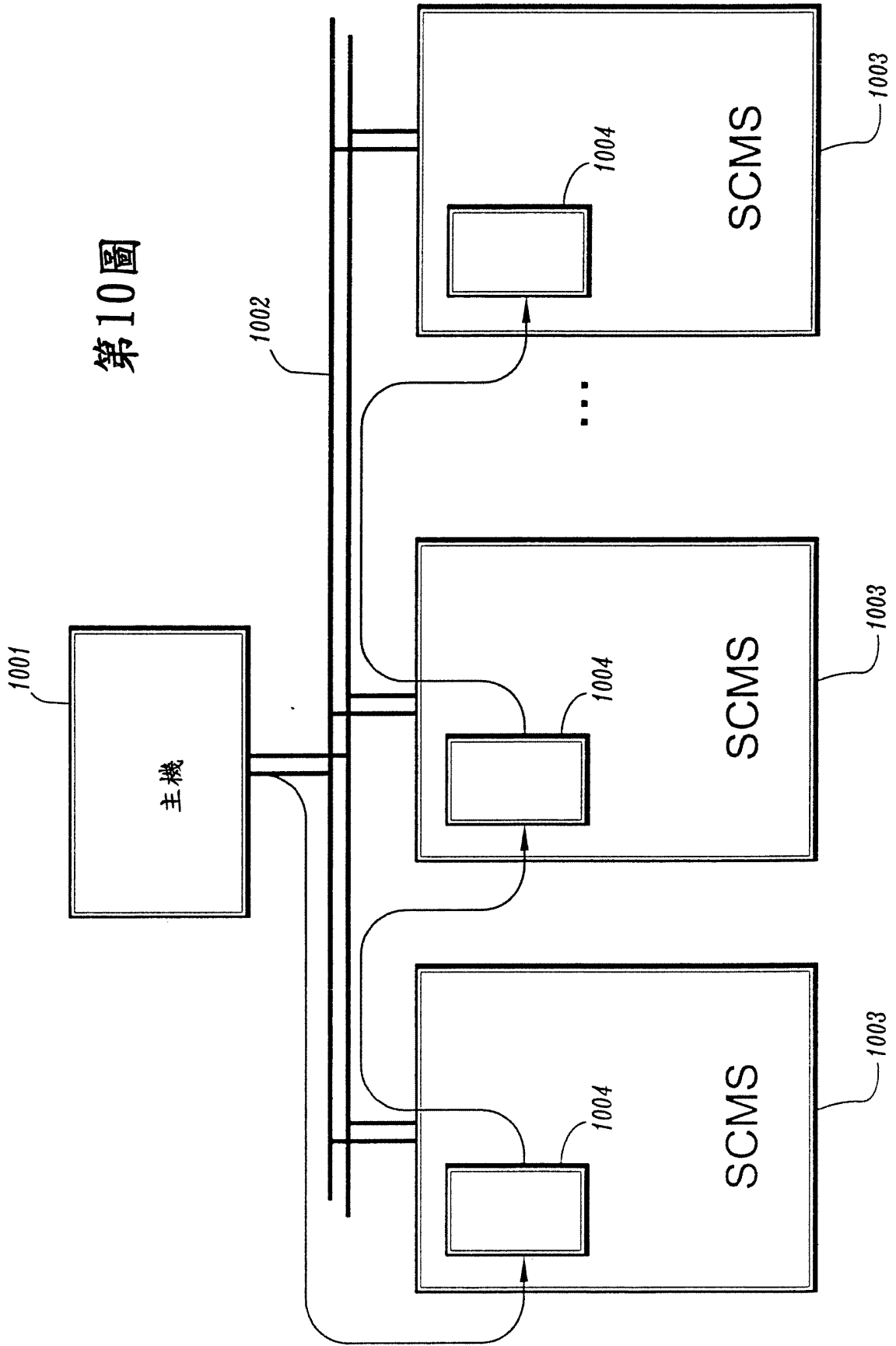


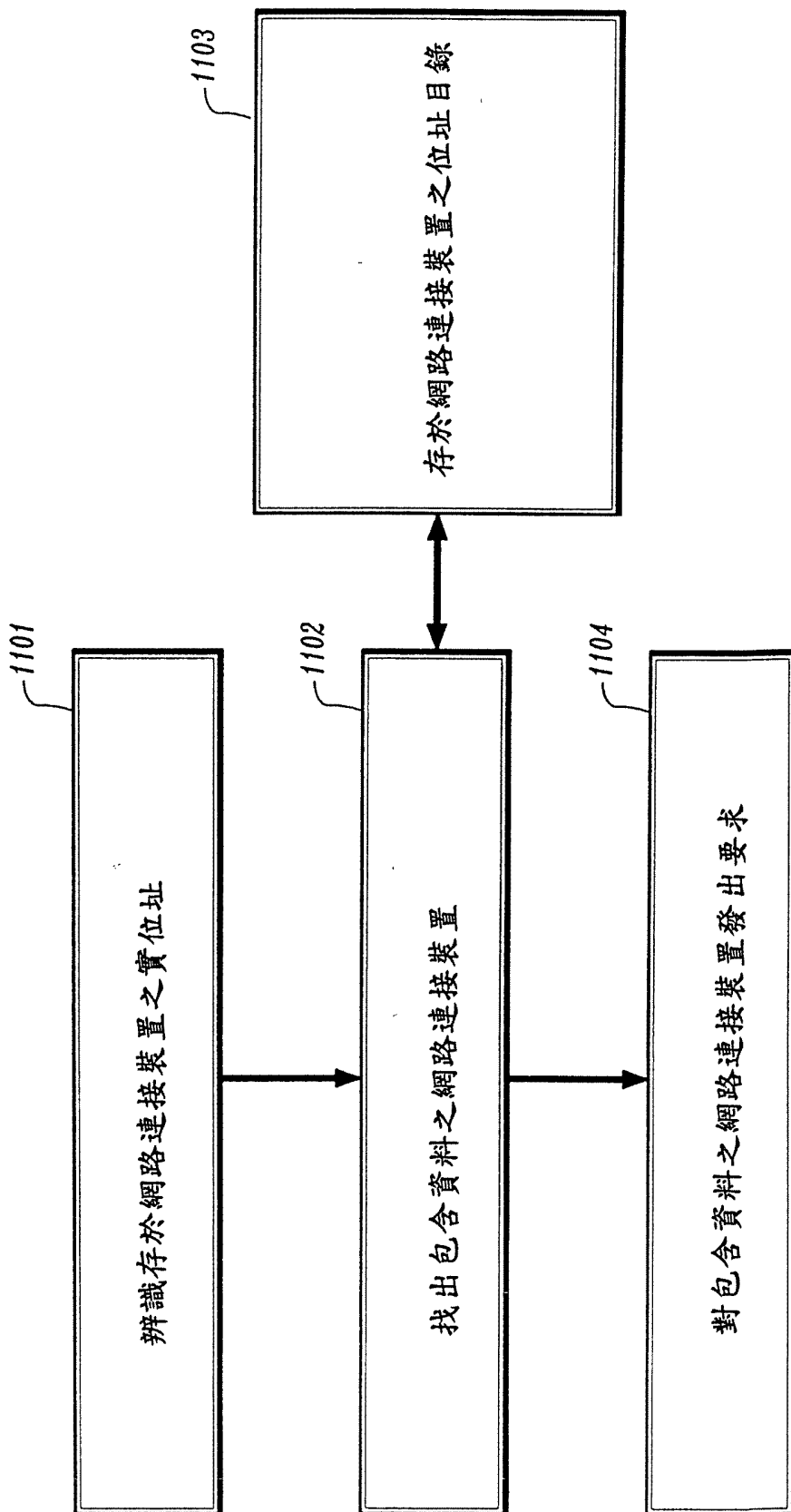
第8圖



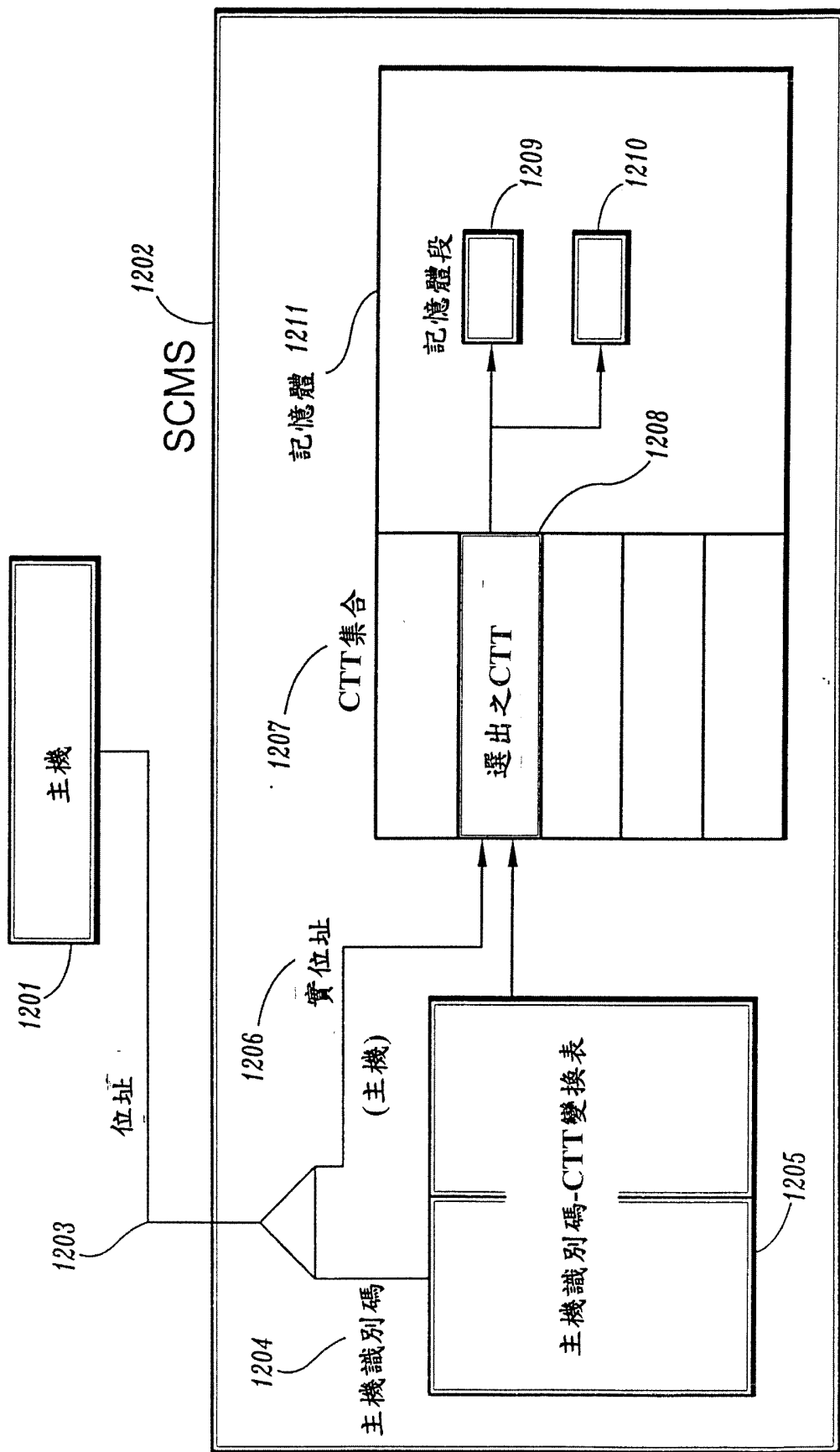
第9圖

第10圖





第11圖



第12圖