

## 發明專利說明書

※ 申請案號：92117713

※ 申請日期：92年6月27日

※IPC分類：G06F13/00 G06F13/38

### 壹、發明名稱：(中文/英文)

多方共用儲存系統中快速複製之高效能鎖定管理

HIGH-PERFORMANCE LOCK MANAGEMENT FOR FLASH COPY  
IN N-WAY SHARED STORAGE SYSTEMS

### 貳、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)

傑拉德羅森瑟爾 / Gerald Rosenthal

住居所或營業所地址：(中文/英文)

美國紐約州阿蒙市新果園路

New Orchard Road, Armonk, New York 10504, USA

國籍：(中文/英文)

美國 / USA

### 參、發明人：(共2人)

姓名：(中文/英文)

1. 卡羅斯法蘭西斯可福恩特 / Carlos Francisco Fuente

2. 威廉詹姆士史蓋爾 / William James Scales

住居所地址：(中文/英文)

1. 英國漢普夏郡波茲懋斯市白鹿路43號

43 White Hart Road, Portsmouth, Hampshire PO1 2TY, United Kingdom

I226549

2.英國漢普夏郡斐爾漢市波特徹斯特鎮波特徹斯特路 5 號

5 Portchester Road, Prochester, Fareham, Hampshire PO16 8AD, United  
Kingdom

國 籍：(中文/英文)

1. 英國 / British

2. 英國 / British

**肆、聲明事項：**

◎本案申請前已向下列國家(地區)申請專利  主張國際優先權：

【格式請依：受理國家(地區)；申請日；申請案號數 順序註記】

英國；2002 年 11 月 29 日；0227825.7

## 玖、發明說明：

### 【發明所屬之技術領域】

本發明係有關於電腦儲存控制之領域，特別有關於可提供快速複製功能的多方共用儲存系統之高階功能儲存控制器。

### 【先前技術】

在電腦的儲存系統之領域內，對於“高階功能”(advanced functions)的需求愈來愈大。這種功能超越簡單的便利性儲存控制系統的輸入/輸出功能。高階功能在其專業領域內廣為人知且依賴中介資料(metadata)之控制，此中介資料用以保留有關於儲存在系統中的真實或是使用者資料之狀態資料(state data)。此種使用高階功能的運用提供多樣化的動作，此多樣化動作被資料的虛擬影像快速應用，使得真正的資料可以留給使用者應用程式使用。其中最廣為人知的高階功能就是“快速複製”(Flash Copy)。

在最高的階層中，快速複製含有提供某些資料的“第二影像(second image)”的功能。此功能有時被稱為“及時點”(Point-In-Time)複製，或是 T0 複製。第二影像的內容最初幾乎與第一影像一樣。且第二影像可“即時的”被取得。在實際的執行內，這表示第二影像可花費比所需求的還要更短的時間被創造出且可創造真實的，分開的，實體的複製，這表示第二影像可以在不干涉應用程式運作的情況下建立。

一旦建立後，第二複製可以運用在很多目的上，例如執行資料的備分，系統測試和資料開採(data mining)。原本就在使用中的應用程式仍將持續的使用第一複製在原本的目的上。比照沒有快速複製的備份方法，應用程式必須停止，然後在應用程式可以再度啟動之前，備份會整個被佔據住。要確實的估計應用程式停住的時間，是一件越來越困難的事。而且要維持住備份的資料也需要越來越多的資源。因此，快速複製所具有的，可以允許備份資料被系統運用卻不影響到應用程式的能力，將會有越來越高的商業價值。

快速複製藉由將讀取輸入/輸出(read I/O)對第二影像(今後稱為目標)的位址指向原始影像的步驟(今後稱為來源)，造成第二影像存在的假象，除非此區域已是寫入的主題。若有某區域是寫入的主題(subject)，不管是目標或來源，那麼便維持一個目標或來源都有資料備份的假象，且某段程序會被呼叫出在不影響其他程序的情況下停止寫入的命令。然後系統讀取來源的資料將其寫入目標，然後再釋放出寫入的指令。之後再有寫入的命令時，便不需要再停止該命令，因為目標已有複製的資料。這種”複製-寫入”(copy-on-write)的技術廣為所知且運用在很多方面。

所有的快速複製的實現都倚靠管理以上所述之判斷的資料結構，所謂的判斷就是，把從目標接收到的資料發送到目標或來源之判斷，以及寫的指令必須停止住等到複製-寫入的過程完成的判斷。此資料結構本質上是用以追蹤從

來源被複製到目標的區域或方向，資料被複製後，其區域和方向與沒有被複製的該區域會有很大的不同。

往後的文章中，將此資料結構稱為中介資料(metadata)，此中介資料之維持，是執行快速複製之演算法中最重要的部分。

快速複製時常會伴隨著一個單一中央處理器複合體(CPU complex)，一起在現今的儲存控制器運作，有時也會使用 SMP 處理器來代替中央處理器複合體。只要作一點小小的努力，使中介資料的複製動作能運用兩個中央處理器複合體，則可以製造出一種可容忍錯誤的快速複製。若第一個中央處理器複合體發生錯誤的時候，第二個中央處理器複合體便可繼續動作，而不會遺失目標影像的存取。

然而，對單一的中央處理器複合體而言，輸入/輸出的能力有其限制。其輸入/輸出能力有其明顯的限制，輸入/輸出能力通常以每秒之輸入/輸出，或是頻寬(bandwidth)來表示，這也影響到了使用應用程式時的表現。此限制發生在許多快速複製的履行上，一個很好的例子就是儲存控制器(storage controller)。傳統的儲存控制器只有一個中央處理器複合體(或者另外有多餘的一對)，所以控制器在表現上常會被限制住。

當然可以安裝多個儲存控制器，但問題在於分別的儲存控制器並沒有共享中介資料的存取，所以並沒有辦法共同合作以管理快速複製的影像。儲存空間被分開，而且快速複製被單一儲存控制器的範圍限定住。來源和目標磁碟

兩者都必須被單一的儲存控制器管理。如此的情況下，當有一個控制器滿載的時候，其他的控制器可能還有一些多餘的空間，但不可能分割來源和目標磁碟，將目標磁碟置於新的控制器之管理下(很不幸的在新的快速複製下，當目標沒有與它相關的實體資料的時候，要移動目標是很簡單的步驟)。

如同限制了目標/來源對的整體體表現一樣，單一控制器之儲存功能的限制使得儲存環境的運作執行更為複雜。

然而，今天的儲存控制系統並沒有嘗試著去解決這個問題。仍以單一控制器來執行快速複製，因此也受到單一控制器的性能限制。

有一個簡單的方法可以讓多個控制器一起參與一快速複製的共享關係，就是指派其中一個控制器為解釋性資料的擁有者(owner)，而其他控制器遵循著此控制器所提出的讀和寫的需求。擁有者控制器處理其連接的主機伺服器傳來的需求，並根據上述的演算法處理之，完成每個 I/O 需求後送回給產生此需求的控制器(originating controller)。

然而此系統的主要缺點和沒有被廣泛使用的原因，是在於要傳送每個輸入/輸出需求，對系統來說是很大的負擔，其負擔可能是其他被廣為使用的系統成本的兩倍，而且也會造成系統效能的減半。

如大家所熟知的例子，在平行分佈的資料結構系統之領域內，為了維持任何資料在相連性關係(coherency relation)內的複製，必須要有一個可以提供二相鎖定協定

的分佈鎖定管理結構以持有資料的鎖定，以連貫關係維持該資料之任何複製本。但是，二相鎖定不但費時，而且會對程序執行造成可觀的訊息負擔。就其本身而言，依據先行技術中的未修改過的二相鎖定協定對於低階的軟體和硬體堆疊而言並沒有好處，例如對於擁有分佈的儲存控制器之儲存區域網路而言，鎖定控制訊息的通過之表現影響會比資料庫之控制層級的影響來得明顯。

因此，存在著使用快速複製以取得分散式鎖定管理優勢的需求，而且使用快速複製更包含擁有最少的訊息額外負擔 (message overhead) 的優點。

#### 【發明內容】

本發明之實施例的第一態樣係提供一種可在儲存控制器節點網路中運作的儲存控制設備，此儲存控制器節點包含一擁有者儲存控制器節點以在輸入/輸出動作中控制資料之一區域的鎖定，以及一訊息組件 (messaging component)，用以傳遞至少一訊息以要求鎖定，同意鎖定，要求鎖定釋放，及傳遞鎖定已解開的訊息；該設備至少包含一輸入/輸出執行組件，用以執行任何擁有者儲存控制器節點所擁有資料的輸入/輸出，其執行方式係符合該擁有者儲存控制器節點所控制的鎖定協定；其中資料之該區域的任何快速複製影像將會被維持在與資料之該區域有連貫性的關係；且其中該輸入/輸出執行組件可快取一先前的肯定回應 (positive confirmation)，此回應表示資料之該區域與

任何該快速複製影像連貫關係被留下，且輸入/輸出執行組件會根據先前的肯定回應執行輸入/輸出的動作。

較佳的是，此輸入/輸出組件可捨棄一已快取的肯定回應，然後要求重新鎖定的。

較佳的是，藉由選擇性的拋棄一已快取的肯定回應，以管理縮小的快取儲存區域。

較佳的是，資料之該區域的一肯定回應更包含與資料之該區域有連結一貫性之資料另外區域的任何肯定回應。

本發明之較佳實施例的第二態樣：提供一種在儲存控制器節點網路中的儲存控制方法，此儲存控制器節點包含一擁所有者儲存控制器節點以在輸入/輸出動作中控制資料之一區域的鎖定，以及一訊息組件，用以傳遞至少一訊息以要求鎖定，同意鎖定，要求鎖定釋放，及傳遞鎖定已解開的訊息；此方法至少包含下列步驟：執行輸入/輸出於任何該擁所有者儲存控制器節點所擁有的資料上，其係符合該擁所有者儲存控制器節點所控制的鎖定協定；維持資料之該區域的快速複製影像在與資料之該區域有連貫的關係；以及快取一先前肯定回應，該肯定回應表示資料之該區域與任何該快速複製影像連貫關係被留下，並根據該先前肯定回應執行輸入/輸出的動作。

較佳的是，第二態樣的方法更包含捨棄一已快取的肯定回應並隨後要求重新鎖定的步驟。

較佳的是，第二態樣的方法藉由選擇性的拋棄一已快取過的肯定回應，以管理一縮小的快取儲存區域。

較佳的是，資料之該區域之一肯定回應更包含與資料之該區域有連結一貫性之資料另外區域的任何肯定回應。本發明之第三態樣提供一種嵌入一電腦可讀取媒體的電腦程式產品，其包含電腦程式碼裝置，當其被載入一電腦系統並於其上執行時，在儲存控制器節點網路中的儲存控制設備具有一擁有着儲存控制器節點以控制輸入/輸出動作過程中資料之一區域的鎖定，一訊息組件用以傳遞至少一訊息以要求一鎖定、同意鎖定、要求鎖定的釋放、和鎖定已被釋放的通知，使得該設備執行下列的步驟：執行輸入/輸出於任何該擁有着儲存控制器節點所擁有的資料上，其係符合該擁有着儲存控制器節點所控制的鎖定協定；維持資料之該區域的快速複製影像在與資料之該區域有連貫的關係；以及快取一先前肯定回應，該肯定回應表示資料之該區域與任何該快速複製影像連貫關係被留下，並根據該先前肯定回應執行輸入/輸出的動作。

本發明之較佳實施例運用一更改過的雙相訊息鎖定系統，用以共同處理多方系統中複數存取控制組件的動作。訊息在系統中的節點之間共同動作，但各節點依然負責實現各自的輸入/輸出。各節點擁有快取功能用以快取之前鎖定需求執行後的結果，如果該區域的狀況已被快取的先前結果明確的指出的話，可再次使用某些結果而不用詢問該區域的狀況。

#### 【實施方式】

若要比較清楚的了解本發明之較佳實施例，便必需詳盡的解釋雙相鎖定訊息控制在多方系統中複數存取控制組件的共同動作方式。

舉個例子來說，考慮到一個運用快速複製的多方系統。假設每個節點都可存取由共同運作的多節點所管理的儲存區域。其中一個節點被設定為與輸入/輸出儲存區域有關的中介資料的擁有者 102。其他節點則被設定為客戶端 (client)。在本發明一較佳實施例中，客戶端節點的其中一個更會被指定為備份的擁有者並維持解釋資訊的複製本，以便在擁有者節點發生錯誤的時候整個系統仍可按照正常的方式持續運作。

假設一個主機的輸入/輸出要求到達(104)某個客戶端節點('C')。假設主機的輸入/輸出要求如果不是目標磁碟的讀或寫，就是來源磁碟的寫入。客戶端 C 藉由暫停(106)輸入/輸出開始執行程序，然後 C 發出(108)一個 REQ 訊息給擁有者節點 O，詢問區塊 (grain) 是否已被複製。

接收到 REQ 後，O 會偵測本身的中介資料結構。如果 O 發覺區域已被複製，O 會以 NACK 訊息回應(110)。如果 O 發覺該區域尚未被複製，O 會放置一鎖定紀錄以鎖住與擁有中介資料結構之該區域相對應的中介資料，並且以 GNT 訊息回應(112)。鎖定紀錄必須確定與剛接收到並執行的要求有一致性 (compatibility)，更必須在 C 的程序持續執行的時候擁有影響相同中介資料的能力。關於鎖定紀錄如何維持，還有一致性的限制，都和輸入/輸出如何被 O

接收的原理一樣，廣為習知人士所熟知，在此便不贅述。

接收到 NACK 訊息後，C 便釋放原本的輸入/輸出要求 (114)。

接收到 GNT 訊息時，C 透過執行資料轉換或者快速複製演算法所需要的轉換 (116) 繼續執行程序。在讀取某特定目標的狀況下，這表示執行對來源磁碟的讀取動作。一段時間後，C 會完成讀取的動作 (118) 並發送 UNL 訊息給 O，同一時間亦完成原始的輸入/輸出要求並傳送給分配此輸入/輸出要求的主機。

接收到 UNL 訊息後，O 會將鎖定紀錄從中介資料列表移除 (122)，更可能解開了因為此鎖定被暫停住的輸入/輸出請求。在本發明一較佳實施例中，O 傳送一 UNLD 訊息給 C (124)，允許 C 再度使用與原始需求 (original request) 有關的資源。但快速複製演算法本身卻不需要這些資源。

在寫入的例子中 (不管是對目標或來源)，C 必須執行複製寫入 (copy-on-write) 的動作 (127)。完成複製寫入全部的步驟後，在原始寫入輸入/輸出的要求仍被暫停的情況下，C 送出 UNLC 的請求給 O。

在接收到 UNLC 後，O 會將該區域的中介資料做標記 (128) 為已複製，移除鎖定紀錄 (130)，並通知所有等候執行的請求該區域已被複製 (132)，然後傳送 UNLD 訊息給 C (134)。

在接收到 UNLD 訊息之後，C 會解開原本被暫停的寫入步驟 (136)，其將於稍後完成，然後 C 完成 (138) 將資料

寫入主機的動作。

當有磁碟輸入/輸出失敗、訊息系統失敗、或一個節點失常的時候，便需要修復的路徑，但以上的需求和執行方法都廣為習知人士所熟知，在此便不贅述。

以上敘述的是單一客戶端與單一輸入/輸出的情況，但亦清楚的表示將以相同的體制處理來自多數個客戶端節點的多數輸入/輸出，O將會使用相同的演算法去處理所有客戶端的要求。

接著看到第2圖，此圖所呈現的是本發明一較佳實施例所擁有的各個設備，在儲存控制組件網路中所實施的設備包括一擁有着(202)、一客戶端輸入/輸出執行組件(204)、一部分被儲存網路所控制的資料(208)的中介資料(206)、資料(208)的複製物件(209)，和連結溝通的工具。此設備亦含有所有權分配組件(ownership assignment component, 210)以及鎖定管理組件(lock management component, 212)。所有權管理組件用以分配中介資料的所有權給擁有着(202)，而鎖定管理組件的功用在於輸入/輸出動作中控制中介資料(206)的鎖定以讓其和複製的資料同步。此外擁有着(202)含有訊息組件(214)，此組件可以傳遞一個或以上訊息於客戶端(204)和擁有着(202)之間的，此訊息係要求一回應有關於中介資料的狀況、同意一鎖定、要求一鎖定的釋放與一鎖定已被解開的訊息。客戶端(204)可執行輸入/輸出於資料，此資料的中介資料被任何擁有着(202)所擁有，執行輸入/輸出資料的步驟被客戶

端與鎖定協定的限制所管理，而此鎖定協定隸屬於被擁  
有者(202)所控制的中介資料階層。

此系統和方法可以處理多方共用儲存控制組件網路中  
分散的鎖定管理，但它的缺點是會有可觀的訊息額外負擔  
(messaging overhead)。如果系統中只有少量的控制組件，  
或是控制組件中只有很少的互動，那麼這問題便不存在，  
但在現代的儲存系統中，像廣範圍的儲存區域網路，通常  
會有大量的控制組件和頻繁的儲存動作。在這樣的情況  
下，若能將訊息額外負擔的數量減少，將會對網路的速度  
非常有助益。

如此，為了改善系統的程序執行表現，在本發明之一  
較佳實施例中，每個客戶端節點具有維持從擁  
有者收到最後回應的紀錄之資訊的能力。特別的是(以根據第 3 圖和第  
1 圖之附加方式來敘述)，客戶端節點 C 被允許快取(308)  
含有本身在第 1 圖之步驟 114 後接收到之 NACK 訊息的資  
訊，或是本身指示及通知的，在第 1 圖之步驟 126 時和在  
步驟 134 之後的 UNLC/UNLD 訊息對。

在接收到一個(302)第 1 圖的步驟 104 中所產生的主機  
要求之時，客戶端 C 便提供一個定義過的鎖定控制演算  
法，如下所示。

C 首先偵測本身的快取資料(303)，檢查是否有欲複製  
之該區域已複製的肯定回應。如果有，則不傳送任何協定  
訊息給 O，然後繼續執行輸入/輸出的工作(304)。

如果快取的訊息中不含肯定指示，將使用如上所述的

未定義之協定。客戶端 C 將繼續執行第 1 圖中的步驟 106 和底下的步驟。NACK 或 UNLC/UNLD 的接收(306)將會導致快取訊息的更新(308)，此更新讓 C 不必傳送協定訊息而直接執行步驟 304。

此術語”悲觀快取”(pessimistic cache)常用以解釋本發明之一較佳實施例所需要的方法，這表示客戶端不用完全的更新擁有者的中介資料；客戶端預先認定某個資料區域需要被複製，除非收到擁有者的 NACK 訊息。然而，當擁有者知道此區域未被複製的時候，此客戶端也絕不能相信此區域已被複製。

在本發明之一較佳實施例中的鎖定快取，需含有一些特定的改變以供客戶端作為修正的操作。第一，在每個快速存取開始的時候(300a)，須重新設定快取訊息(301)(表示所有的區域都必須被複製)。這個動作可以用許多方法驅動，但一個從擁有者到客戶端的訊息是最直接的作用方法。第二，在許多時候客戶端可能會錯過某些從擁有者傳來的快取訊息已被重設的通知(300b)(有時只是因為網路的不穩定)，客戶端必須假設最糟糕的狀況並定時的重新設定(301)它的快取資訊。

如習知人士所知，以上所述還可有很多的變化和延伸。舉例來說，快取到的通知是可丟棄的，但也可隨時從擁有者節點處復原，比起儲存所有的中介資料，客戶端將擁有更多可自由運用的空間。客戶端將根據它們執行輸入/輸出的區域來判定它們是否有從快取訊息的動作得到更大

的益處。

在本發明之另一實施例中，一個 NACK 訊息(也可以是 GNT 或 UNLD 訊息)可以帶回比來自 REQ/GNT/UNLC/UNLD 直接運作的區域的 NACK 訊息更多的資訊。有關鄰近區域的訊息被清除後，便可從擁有者被傳送到客戶端。

較佳的是，以上所闡述的方法可被架構在軟體上並且被使用於一或多個處理器，而且該軟體可以被視為一種附帶的電腦程式元件，附帶在任何適當的資料媒體上(如光碟、磁碟等)。而資料的傳輸管道可包含各種不同的訊號攜帶媒介，如有線的或無線的訊號媒介。

本發明可適當地體現於電腦系統使用的電腦程式產品。此類的實施方式包含一電腦可讀取指令之序列，無論是固定在有形的媒體如電腦可讀取媒體(如軟碟、CD-ROM、唯讀記憶體、或硬碟)上、或是可經由一數據機或其他介面元件傳輸至一電腦系統，或者透過一有形的媒體包含但不限定如光學或其他類似的傳輸線，或一無形地使用無線技術包含但不限定微波、紅外線或其他的傳輸技術。該電腦可讀取指令之序列體現全部或部分前述的功能性。

熟習該項技藝人士皆瞭解的是，本發明之電腦可讀取指令可用多種程式語言，以在各種不同的電腦架構和作業系統上執行。此外，該些指令可用各種記憶技術儲存，現在或未來包含但不限定如半導體、磁性或光學的，也可用各種傳輸技術加以傳輸，現在或未來包含但不限定如光

學、紅外線或微波。可預期的是此一電腦可讀取產品可以一種可攜式媒體伴隨印刷的或電子的使用說明散佈，例如包裝好的軟體、以一電腦系統預先載入至如系統唯讀記憶體或固定的磁碟，或是藉由伺服器或網路電子佈告板散佈，如網際網路或全球資訊網。

值得注意的是，以上所提的實施例之各種變化，對於習知技藝者而言，都是容易瞭解的。

#### 【圖式簡單說明】

以下藉由隨付之圖示來例示性地敘述本發明一較佳實施例，其中：

第 1 圖繪示了一雙相鎖定機制的流程圖，此機制使用鎖定訊息來控制一資料區域和此資料的快速複製影像的連貫；

第 2 圖繪示了本發明一較佳實施例中系統的各個系統組件；以及

第 3 圖繪示了本發明一較佳實施例中額外的步驟。

#### 【元件代表符號簡單說明】

102-308 流程步驟

200 設備

204 客戶端

208 資料

210 所有權控制

202 擁有者

206 中介資料

209 資料影像

212 鎖定管理者

214 訊息傳送組件

### 伍、中文發明摘要：

一種在儲存控制器節點之網路中運作的儲存控制設備，此儲存控制器節點包含擁有者節點，用以在 I/O 動作中控制資料區域鎖定，以及一訊息組件，用以要求鎖定、同意鎖定、要求釋放鎖定及鎖定釋放訊號。此設備至少包含一輸入/輸出執行組件，用以執行任何擁有者節點所擁有資料的輸入/輸出，用以符合該擁有者節點所控制的鎖定協定。資料區域的任何快速複製影像將會被維持在與此資料區域連貫的關係；且輸入/輸出執行組件可快取一先前的肯定回應，此回應表示資料區域已伴隨著任何快速複製被留下，且輸入/輸出執行組件會根據先前肯定回應執行輸入/輸出的動作。

### 陸、英文發明摘要：

Storage control apparatus is operable in a network of storage controller nodes having an owner node operable to control locking of a region of data during I/O activity, a messaging component operable to request locks, grant locks, request release of locks and signal lock release. The apparatus comprise an I/O performing component to perform I/O on data owned by any owner node, subject to compliance with lock protocols controlled by the owner node. Any Flash Copy image of the region of data is maintained in a coherency relation with the region of data, and the I/O performing component is operable to cache a previous positive confirmation that the region of data has been left in a coherency relation with any said Flash Copy image, and to perform I/O activity on the basis of said previous positive confirmation.

柒、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：

102-138 流程步驟

捌、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無

## 拾、申請專利範圍：

1. 一種儲存控制設備，可在儲存控制器節點的網路中執行，該儲存控制器節點具有一擁有着儲存控制器節點以在輸入/輸出動作中控制資料之一區域的鎖定，以及一訊息組件，用以傳遞至少一訊息以要求鎖定、同意鎖定、要求鎖定釋放、及通知鎖定已釋放的訊息，該儲存控制設備至少包含：

一輸入/輸出執行組件，用以執行任何該擁有着儲存控制器節點所擁有資料的輸入/輸出，該輸入/輸出執行組件係符合該擁有着儲存控制組件節點所控制的鎖定協定；

其中，資料之該區域的任何快速複製影像係與資料之該區域維持在一連貫關係；以及

其中，該輸入/輸出執行組件係可用以快取一先前肯定回應，該先前肯定回應表示資料之該區域與任何該快速複製影像連貫關係被留下，以及可用以根據該先前肯定回應執行輸入/輸出的動作。

2. 如申請專利範圍第 1 項所述之儲存控制設備，其中該輸入/輸出執行組件可用以捨棄一已快取的肯定回應，然後要求一重新的鎖定。

3. 如申請專利範圍第 2 項所述之儲存控制設備，其中一降低快取儲存區域係由選擇性的捨棄一已快取的肯定回應所管理。

4. 如申請專利範圍第 1 項所述之儲存控制設備，其中一資料之該區域的肯定回應更包含任何與資料之該區域有連結一貫性的資料另外區域的肯定回應。

5. 一種在儲存控制器節點網路上的儲存控制方法，該儲存控制器節點包含一擁有着儲存控制器節點用以在輸入/輸出動作中控制資料之一區域的鎖定，以及一訊息組件用以傳遞至少一訊息以要求鎖定、同意鎖定、要求鎖定釋放、及通知鎖定已釋放的訊息，該儲存控制方法至少包含：

執行輸入/輸出於任何該擁有着儲存控制器節點所擁有的資料上，其係符合該擁有着儲存控制器節點所控制的鎖定協定；

維持資料之該區域的快速複製影像在與資料之該區域有連貫的關係；以及

快取一先前肯定回應，該肯定回應表示資料之該區域與任何該快速複製影像連貫關係被留下，並根據該先前肯定回應執行輸入/輸出的動作。

6. 如申請專利範圍第 5 項所述之方法，更包含捨棄一已快取的肯定回應，隨後要求一重新鎖定之步驟。

7. 如申請專利範圍第 6 項所述之方法，更包含藉由選擇性的捨棄一已快取的肯定回應以管理一縮小的快取儲存區

域。

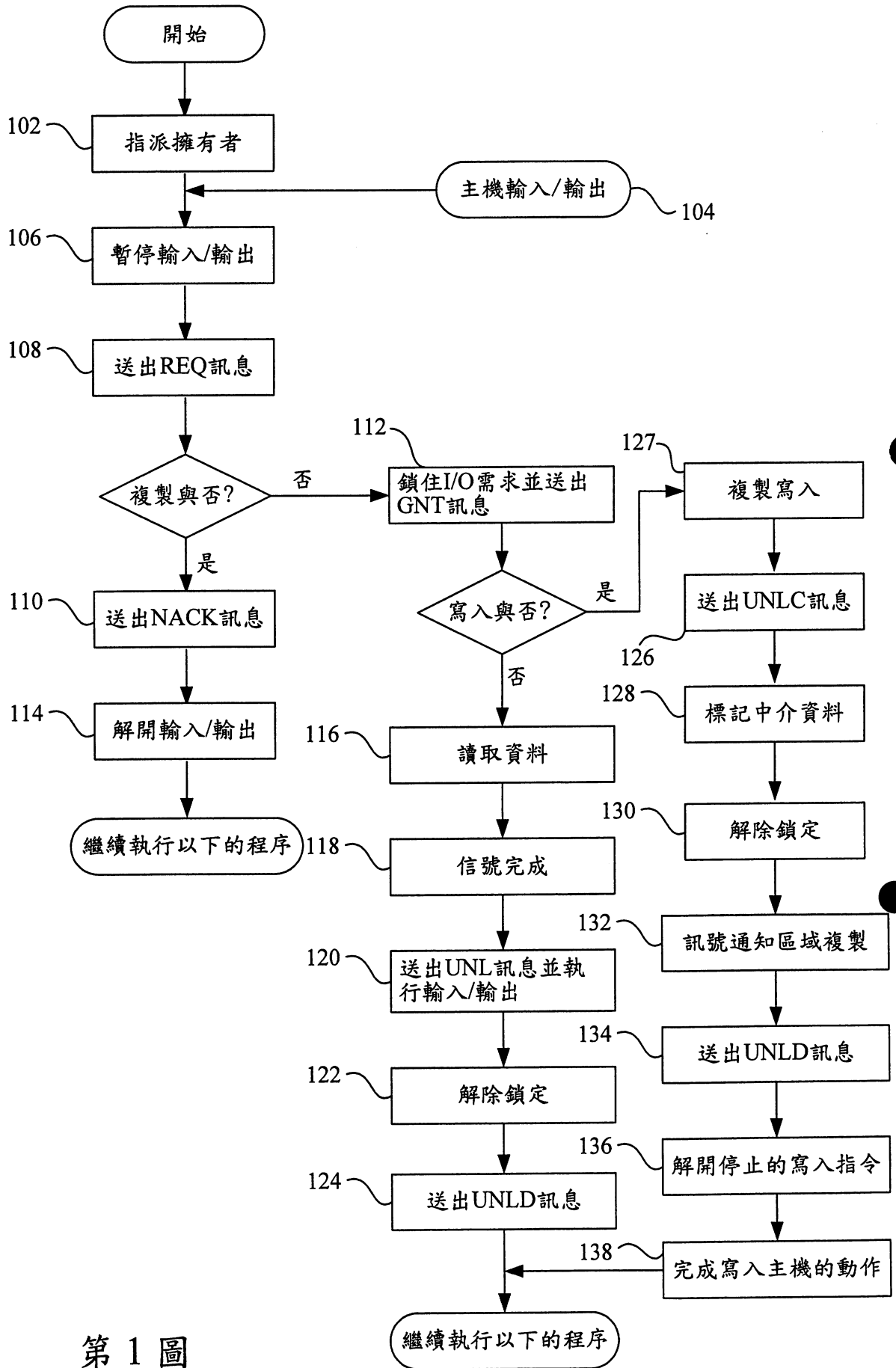
8. 如申請專利範圍第 5 項所述之方法，其中資料之該區域的肯定回應更包含任何與資料之該區域有連結一貫性之資料另外區域的肯定回應。

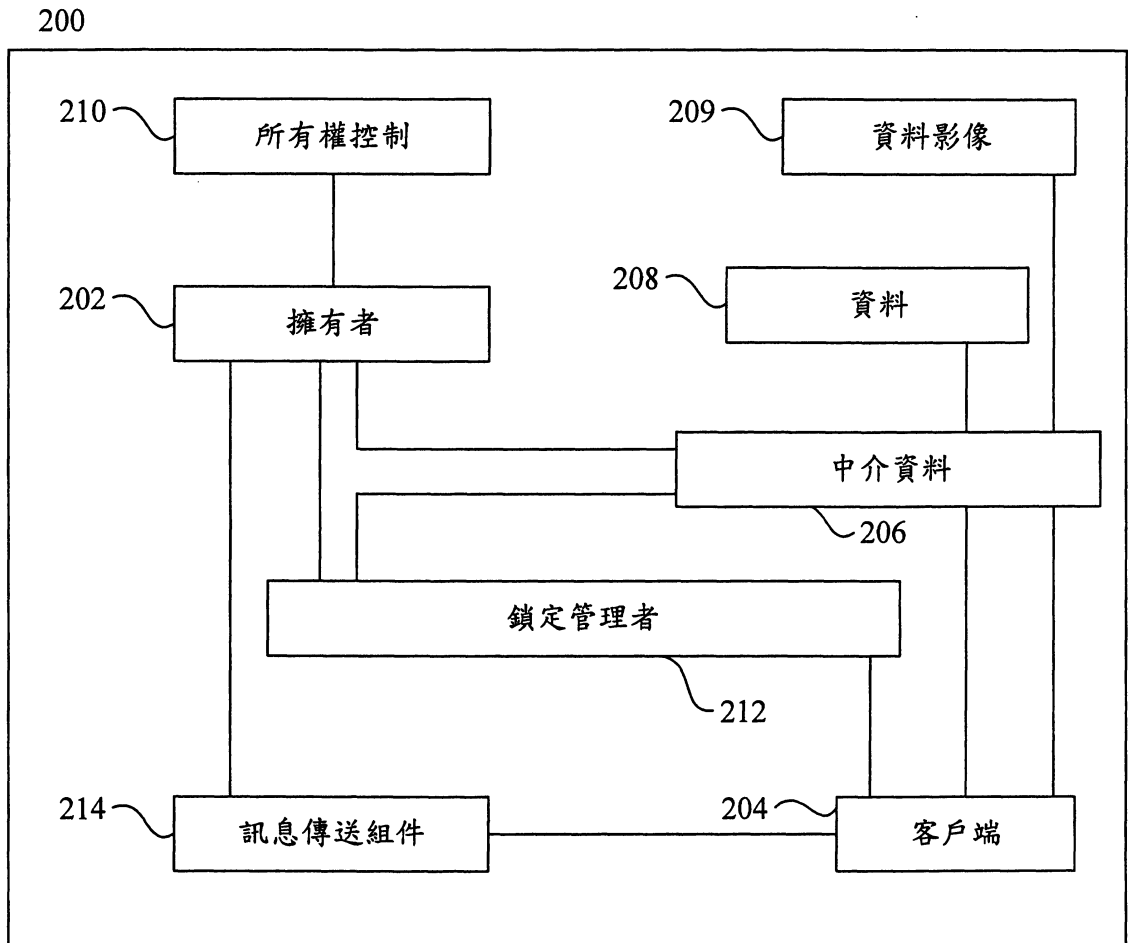
9. 一種嵌入一電腦可讀取媒體的電腦程式產品，至少包含電腦程式碼裝置，當其被載入一電腦系統並於其上執行時，在儲存控制器節點網路中的儲存控制設備具有一擁有者儲存控制器節點以控制輸入/輸出動作過程中資料之一區域的鎖定，一訊息組件用以傳遞至少一訊息以要求一鎖定、同意鎖定、要求鎖定的釋放、和鎖定已被釋放的通知，使得該設備執行下列的步驟：

執行輸入/輸出於任何該擁有者儲存控制器節點所擁有的資料上，其係符合該擁有者儲存控制器節點所控制的鎖定協定；

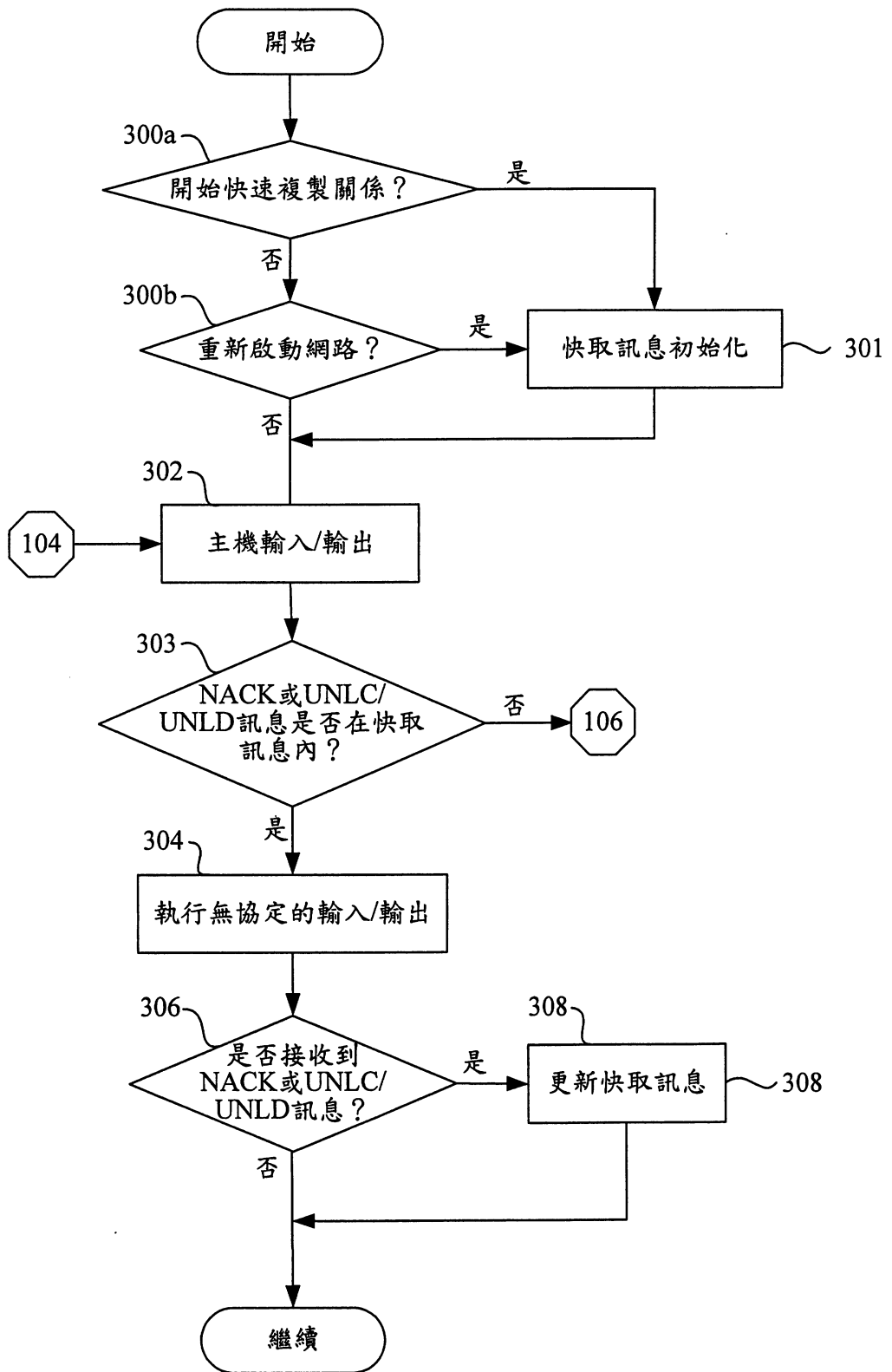
維持資料之該區域的快速複製影像在與資料之該區域有連貫的關係；以及

快取一先前肯定回應，該肯定回應表示資料之該區域與任何該快速複製影像連貫關係被留下，並根據該先前肯定回應執行輸入/輸出的動作。





第 2 圖



第 3 圖