



(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 I494762 B

(45)公告日：中華民國 104 (2015) 年 08 月 01 日

(21)申請案號：098129955 (22)申請日：中華民國 98 (2009) 年 09 月 04 日

(51)Int. Cl. : **G06F12/08 (2006.01)** **G06F17/30 (2006.01)**

(30)優先權：2008/09/15 美國 61/097,079
2008/11/14 美國 12/271,472

(71)申請人：微軟公司(美國) MICROSOFT CORPORATION (US)
美國

(72)發明人：理剛麥密特 LYIGUN, MEHMET (US)；貝克依孚堅利 BAK, YEVGENLY (US)；
福汀麥克 FORTIN, MICHAEL (US)；費歐德大衛 FIELDS, DAVID (US)；愛爾根
森克 ERGAN, CENK (US)；柯薛恩巴姆亞歷山大 KIRSHENBAUM, ALEXANDER
(US)

(74)代理人：蔡坤財；李世章

(56)參考文獻：

TW	190585	TW	347515
US	7036040B2	US	2007/0294474A1

審查人員：林巧宜

申請專利範圍項數：19 項 圖式數：8 共 59 頁

(54)名稱

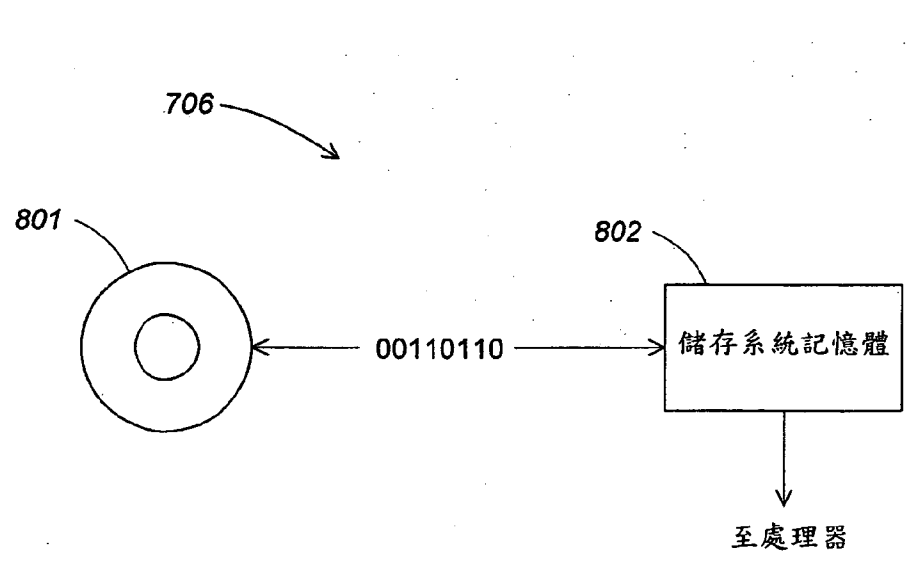
用於管理快取資料及元資料的方法、電腦儲存媒體與電腦

METHOD, COMPUTER STORAGE MEDIUM, AND COMPUTER FOR MANAGING CACHE DATA AND METADATA

(57)摘要

本發明之實施例等提供一種確定非揮發性記憶體裝置的內容可信賴，在通過電源轉換，像是再啟動時確實反映出磁碟儲存裝置上儲存的資料之技術。例如：本發明某些具體實施例提供決定該快取內容及/或磁碟內容在電源轉換期間是否遭修改，導致快取內容不再精確反映磁碟儲存裝置內所儲存資料之技術。進一步，某些具體實施例提供管理正常(「穩定狀態」)操作期間以及通過電源轉換時的快取元資料，確定可有效存取並可信賴地儲存以及通過電源轉換時復原快取元資料之技術。

Embodiments of the invention provide techniques for ensuring that the contents of a non-volatile memory device may be relied upon as accurately reflecting data stored on disk storage across a power transition such as a reboot. For example, some embodiments of the invention provide techniques for determining whether the cache contents and/or or disk contents are modified during a power transition, causing cache contents to no longer accurately reflect data stored in disk storage. Further, some embodiments provide techniques for managing cache metadata during normal (“steady state”) operations and across power transitions, ensuring that cache metadata may be efficiently accessed and reliably saved and restored across power transitions.



- 706 . . . 儲存裝置/
儲存系統
- 801 . . . 非揮發性記
錄媒體
- 802 . . . 揮發性記憶
體

第 8 圖

104年4月28日修正頁

發明專利說明書

公告本

(本說明書格式、順序，請勿任意更動，※記號部分請勿填寫；惟已有申請案號者請填寫)

※申請案號：98129955

G06F 12/08 (2006.01)

※申請日期：2009 年 9 月 4 日

※IPC 分類：G06F 17/30 (2006.01)

一、發明名稱：(中文/英文)

用於管理快取資料及元資料的方法、電腦儲存媒體與電腦/METHOD, COMPUTER STORAGE MEDIUM, AND COMPUTER FOR MANAGING CACHE DATA AND METADATA

二、中文發明摘要：

本發明之實施例等提供一種確定非揮發性記憶體裝置的內容可信賴，在通過電源轉換，像是再啟動時確實反映出磁碟儲存裝置上儲存的資料之技術。例如：本發明某些具體實施例提供決定該快取內容及/或磁碟內容在電源轉換期間是否遭修改，導致快取內容不再精確反映磁碟儲存裝置內所儲存資料之技術。進一步，某些具體實施例提供管理正常(「穩定狀態」)操作期間以及通過電源轉換時的快取元資料，確定可有效存取並可信賴地儲存以及通過電源轉換時復原快取元資料之技術。

三、英文發明摘要：

Embodiments of the invention provide techniques for ensuring that the contents of a non-volatile memory device may be relied upon as accurately reflecting data stored on disk storage across a power transition such as a reboot. For example, some embodiments of the invention provide techniques for determining whether the cache contents and/or or disk contents are modified during a power transition, causing cache contents to no longer accurately

reflect data stored in disk storage. Further, some embodiments provide techniques for managing cache metadata during normal ("steady state") operations and across power transitions, ensuring that cache metadata may be efficiently accessed and reliably saved and restored across power transitions.

四、指定代表圖：

(一)本案指定代表圖為：第(8)圖。

(二)本代表圖之元件符號簡單說明：

706 儲存裝置/儲存系統

802 揮發性記憶體

801 非揮發性記錄媒體

五、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無

六、發明說明：

【發明所屬之技術領域】

本發明係關於運用非揮發性記憶體裝置，像是可移除式和不可移除式非揮發性隨機存取記憶體(non-volatile random access memory, NVRAM)裝置的技術。

【先前技術】

某些傳統作業系統提供能力來運用非揮發性記憶體裝置(即是可操作來提供輔助儲存以及/或記憶體給電腦的周邊裝置，像是快閃記憶體 USB 隨身碟)當成區塊或檔案階層快取來用於慢速儲存裝置(例如磁碟儲存媒體或者一或多個可透過網路存取的儲存裝置)，來改善作業系統及/或應用程式的效能。在此範圍內，因為讀取自寫入至非揮發性記憶體裝置(此後簡稱為「快取裝置」)所執行的操作，顯著快過讀取自寫入至慢速儲存裝置，所以使用快取裝置快取儲存在這種慢速裝置內的資料，就可提供機會以顯著改善作業系統及/或應用程式的輸入/輸出(input/output, I/O)操作之速度。為此，由美國華盛頓州 Redmond 的微軟公司生產的 Microsoft Windows Vista 作業系統包含一種 ReadyBoost 功能，其允許使用者運用快取裝置快取慢速儲存裝置(為了方便起見也稱為「磁碟

儲存裝置」或「磁碟」，不過吾人應該瞭解，這些術語一般代表其 I/O 通常執行起來比快取裝置慢的任何儲存機構及/或裝置，包含可透過網路存取的儲存裝置)內的資料。

使用快取管理員組件可達成運用快取裝置來快取儲存在磁碟上的資料，在某些實施當中該組件為在作業系統的 I/O 堆疊中實施的驅動器。第 1A 圖至第 1B 圖說明範例高階處理 10A-10B，藉此快取管理員組件 100 管理至快取裝置 110 的資料快取。快取裝置 110 可使用有線及/或無線通訊基礎設施與協定，耦合至快取管理員 100 所在的電腦(未顯示)。例如：快取裝置 110 可從電腦(例如包含快閃記憶體 USB 隨身碟)當中移除，透過一或多個有線及/或無線網路不可移除及/或可存取至電腦。

在處理 10A(第 1A 圖)開始時，快取管理員 100 接收到寫入要求，指出應該將資料寫入至磁碟儲存裝置(即是快取磁碟 120 (Cached Volume))上的位址 X。快取管理員 100 利用在操作 101 上將資料寫入至快取磁碟 120 上的位址 X，並且也在操作 102 內寫入至快取裝置 110 上的位址 Y 來處理要求。10B(第 1B 圖)的處理包含在當已經接收指示應該讀取快取磁碟 120 上位址 X 之上儲存的資料之讀取要求之後執行的操作。快取管理員 100 決定將資料快取在快取裝置 110 的位址 Y 上，並且導致在操

作 103 內讀取位址 Y 上的資料。然後從快取裝置提供資料來滿足操作 104 內的讀取要求。

快取管理員維持元資料內磁碟位址(例如位址 X)映射至對應的快取位址(例如位址 Y)，並且此「快取元資料」通常運用在讀取自或寫入至快取裝置。一般而言，快取元資料維持在記憶體內，並且當接收到 I/O 要求時由快取管理員存取。如此，當快取管理員接收到指向磁碟偏移 X 的讀取要求時，快取管理員使用快取元資料判斷資料也儲存在快取偏移 Y 內，並且從快取偏移 Y 而非磁碟偏移 X 當中讀取資料來滿足要求。當快取管理員接收到指向磁碟偏移 X 的寫入要求，快取管理員運用快取元資料來判斷在該磁碟位址上的資料是否也儲存在快取內。若是(例如若資料儲存在快取位址 Y 上)，快取管理員會讓資料寫入快取內的適當位址，或驅逐(evict)該位址上的快取內容。若否，則快取管理員讓資料寫入快取內，並且可更新快取元資料，如此未來對於磁碟偏移 X 的讀取會由快取上儲存的資料所取代。

傳統作業系統可支援相對受限儲存磁碟的快取裝置，例如：Windows Vista 作業系統的 ReadyBoost 功能支援最多 4 GB (gigabyte)儲存容量的快取裝置(在 Windows Vista 上市之時，快取裝置的最大儲存容量大約是 2 GB)。近年來快取裝置的儲存容量迅速成長，某些快取裝

置甚至提供高達 16 GB 的儲存容量，在經過壓縮之後可儲存大約 32 GB 的資料。

【發明內容】

申請人已經瞭解，具有相當大儲存容量的快取裝置提供顯著機會，來改善作業系統以及應用程式所執行 I/O 操作的速度。申請人也瞭解，傳統作業系統支援只具有相當有限儲存容量的快取裝置之一個因素為，當發生特定種類電源轉移時(例如待機、休眠(或非 Microsoft Windows 作業系統所使用的同等模式)或再啟動)，必須重新載入快取內容。對於具有相當大儲存容量的快取裝置而言，重新載入快取內容會耗費可觀的時間，並且顯著消耗處理資源。就範例來說，使用背景 I/O 從磁碟以每秒大約 10 MB (megabyte) 的速度重新載入 8 GB 快閃記憶體裝置(最多容納 16 GB 壓縮資料)要耗費三十分鐘。這不僅抹煞了運用快取裝置所要獲得的任何效能優點，也顯著拖慢系統操作。

通過特定電源轉換必須重新載入快取內容的一項因素為，並無方式可信賴地確定在電腦重新啟動時快取內容準確代表磁碟儲存裝置的內容，因為可能在電源轉換期間已經修改過快取裝置及/或磁碟的內容。例如：當第一電腦關機，駭客可中斷與可移除式快取裝置的連線、將

其連線至其他電腦並且修改快取內容，如此若裝置重新連接至該第一電腦，則從快取裝置供應不正確的資料(此後稱為「偽造」資料)來滿足 I/O 要求。除了駭客的惡意動作以外，在電源轉換期間由於電腦或快取裝置硬體故障，所以快取內容也可變成毀損。

快取內容也可因為在電源轉換期間磁碟上的資料已經更新而在轉換期間變成「過期」，如此當電腦重新啟動，該快取內容不再準確代表磁碟內容。例如：關機之後，使用者可將磁碟開機在未辨識快取裝置的其他作業系統內，並且修改磁碟上已經儲存並且快取至快取裝置上的資料，如此當電腦重新啟動，則快取內容不再反應磁碟上所儲存的資料。在其他範例中，於快取裝置無法存取作業系統之後(例如快取裝置關機之後)，關機期間電腦上會發生特定操作，如此快取內容無法正確反映出作業系統在此點之後對磁碟執行的任何寫入動作。通過電源轉移的許多事件任一者都可導致快取內容變成過期。

本發明的具體實施例提供管理這些與其他議題的技術，如此快取內容可信賴，在通過電源轉換時確實反映出磁碟上儲存的資料。例如：本發明的某些具體實施例提供確認快取內容通過電源轉換時仍舊為驗證的技術。此外，某些具體實施例提供可信賴地確定快取內容在通過電源轉換時不會變成過期的技術。進一步，某些具體

實施例提供管理快取元資料通過電源轉換，以及在正常（「穩定狀態」）操作期間，確定可有效存取並可信賴地儲存以及當發生電源轉換時復原快取元資料之技術。

利用提供確定快取內容可信賴，在通過電源轉換時確實反映磁碟上儲存的資料之技術，本發明的某些具體實施例可運用具有可觀儲存容量的快取裝置，來顯著加速作業系統及/或應用程式所執行的 I/O 操作。增加的 I/O 操作速度不僅加速電腦的正常「穩定狀態」操作，也顯著加速開機期間執行的操作，如此電腦更迅速備妥可供使用。

在某些具體實施例內，提供一種操作耦合至儲存媒體以及快取裝置的電腦之方法，該電腦包含作業系統。該方法包含下列步驟：(A) 在回應一要求以將一資料項目寫入該儲存媒體上一位址時，導致該資料項目寫入至該儲存媒體上該位址以及寫入至該快取裝置上一對應位址；(B) 在再啟動該作業系統之後，決定儲存在該快取裝置上該位址之上的該資料項目是否可信賴地用來滿足從該儲存媒體上該位址讀取該資料項目的一要求；以及(C) 若在(B)內決定儲存在該快取裝置上該位址之上的該資料項目可信賴地用來滿足從該儲存媒體上該位址讀取該資料項目的一要求，則導致從該快取裝置上該位址讀取該資料項目。

其他具體實施例提供至少一電腦儲存媒體其上儲存指令，當這些指令由包含一作業系統並且耦合至一快取裝置的電腦所執行時，執行包含下列步驟的方法：(A) 在回應一要求以將一資料項目寫入該電腦儲存媒體上一位址時，導致該資料項目寫入至該電腦儲存媒體上該位址以及寫入至該快取裝置上一對應位址；(B) 在再啟動該作業系統之後且回應從該電腦儲存媒體上該位址讀取該資料項目的一要求，決定儲存在該快取裝置上該位址之上的該資料項目是否在(A)之後修改過；以及(C) 若在(B)內決定儲存在該快取裝置上該位址之上的該資料項目並未在(A)之後修改過，則導致從該快取裝置上該位址讀取該資料項目。

仍舊是其他具體實施例提供一電腦，包含：一儲存媒體，其具有至少一其上儲存資料的位址；一快取裝置，其具有至少一位址其上儲存資料；至少一處理器，其透過一作業系統執行來：在回應一要求以將一資料項目寫入該儲存媒體上一第一位址時，導致該資料項目寫入至該儲存媒體上該第一位址以及寫入至該快取裝置上一對應位址；識別一寫入操作之發生，該寫入操作係對於該儲存媒體上的該位址執行，但是不對該快取裝置上的一位址執行；以及在再啟動該作業系統之後，利用決定一寫入操做已經針對該儲存媒體上該第一位址執行，但卻

沒有對該快取裝置上該對應位址執行，來決定儲存在該快取裝置上該對應位址之上的該資料項目是否可信賴，以用來滿足從該儲存媒體上該第一位址讀取該資料項目的一要求。

【實施方式】

本發明的某些具體實施例提供用於確認快取內容在通過電源轉換時確實反映磁碟儲存裝置內容之技術。例如：某些具體實施例提供用於確定快取內容在通過電源轉換時仍舊具有驗證及/或不會變成過期之能力。進一步，某些具體實施例提供用於管理快取元資料的技術，來確定元資料在電源轉換期間不會遭竄改。此外，某些具體實施例提供用於儲存快取元資料的能力，此資料可改善效率並且在電源轉換與正常操作之下執行。下列章節詳細說明這些具體實施例。

I. 確認快取資料驗證性

當電腦經歷電源轉換(例如進入待機或休眠模式，或再啟動)，快取裝置會與電腦中斷連線，並且改變其內容(例如由心懷不軌的駭客)。例如：當電腦進入待命或休眠模式，像是快閃記憶體驅動器這類可移除式快取裝置會與電腦中斷連線並且其內容可修改。即使像是內部 NVRAM

裝置這類非移除式裝置都可在作業系統再啟動時(即重新載入，藉此重新啟動電腦)中斷連線並且改變其內容。結果，當快取裝置重新連線至電腦，已經儲存與使用者所相信不同的資訊(即是快取內儲存的資料未通過「驗證」)。若從快取當中取出未驗證資料來滿足輸入/輸出(I/O)要求，則電腦的操作會受不良影響。

本發明的某些具體實施例提供用於偵測在電源轉換期間發生在快取內容的修改，以確定並未使用來自快取的未驗證資料來滿足 I/O 要求之技術。在某些具體實施例內，提供偵測在電源轉換期間以未驗證方式對快取內容所進行的任何「離線修改」之能力。

某些具體實施例運用至少部分資料的代表來提供此能力，例如：可用預定方式從至少部分資料及/或其他資訊當中計算或取得代表。當資料寫入快取時可第一次產生該代表，並且儲存在一或多個位置上。例如：代表可搭配資料寫入快取內，或寫入某些其他位置內。當儲存至快取時，該代表可用將其與資料相關聯的方式來儲存(例如將代表寫入與資料要寫入的位址相鄰之快取位址、寫入至與資料相關聯的快取元資料以及/或關聯於某些其他方式)。從快取讀取資料之後，也可取得該代表。該代表可用預定方式再生，並且再生的代表可與取得的代表互相比較，若代表吻合，則判斷從快取當中取得的資料

經過驗證，並用來滿足讀取要求，若未吻合，則發出要求從磁碟儲存裝置當中讀取資料來滿足讀取要求，並且驅逐快取上儲存的未驗證資料(例如刪除)。

第 2A 圖至第 2B 圖更詳細說明此處理。尤其是，第 2A 圖內顯示的處理 20A 包含將資料寫入快取時所執行的操作，並且第 2B 圖內顯示的處理 20B 包含從快取當中讀取資料之後所執行的操作。處理 20A-20B 每一者都包含由快取磁碟 120 (每一圖的左邊)、快取管理員 100 (中間) 以及快取裝置 110 (右邊) 所執行的操作。

在處理 20A (第 2A 圖) 開始時，快取管理員 100 在步驟 205 內接收到指引至磁碟儲存裝置上位址 X 的寫入要求。在步驟 210 內，快取管理員 100 運用快取元資料(未顯示)，來決定快取裝置 110 上也要寫入資料的位址 Y。快取管理員也產生至少部分資料的代表。

本發明的具體實施例可使用任何合適的技術產生此代表。在某些具體實施例內，可運用一或多個加密驗證技術來產生代表。例如：在某些具體實施例內，該代表可包含產自資料的訊息認證碼(message authentication code, MAC) 以及一組機密金鑰與每一資料項目序號。不過，本發明並不受限於這種實施，任何合適的代表產生技術都可運用。例如：不需要運用加密驗證技術。根據範例，可交替使用強散列及 / 或循環冗餘碼(cyclic

redundancy code, CRC)來呈現資料，並且可從儲存至快取的個別資料項目或用於一或多組資料項目當中產生。在此範圍內，申請人已經瞭解若變化資料驗證性的目的僅為偵測硬體毀損的證據(即是不關心資料遭駭)，則使用 CRC 就已足夠，並且可比起產生 MAC 給每一資料項目消耗更少處理資源。不過，若目的在於避免駭客或惡意行動者修改快取內容，則使用加密解決方案較佳，如此可使用像是 MAC 這類代表。本發明並不受限於任何特定實施，任何合適的技術都可運用。

在步驟 215 內，快取管理員 100 發出要求要快取裝置 110 將資料寫入快取位址 Y。快取管理員也發出要求要快取裝置 110 寫入該代表，例如：每一管理員 100 都可指定該代表應該運用任何其他技術，寫入一或多個與快取位址 Y 相鄰的位置，或儲存在資料的快取元資料內。如上面所討論，本發明並不受限於將該代表寫入任何特定位置(例如不需要寫入快取裝置 110)。若寫入快取，則該代表可用任何所要方式聯結於資料。

在步驟 220 內，快取裝置 110 接收要求，並且在步驟 225 內利用寫入資料和該代表來處理。

在步驟 230 內，快取管理員 100 發出對應要求要快取磁碟 120 將資料寫入磁碟位址 X。在步驟 235 內，快取磁碟 120 接收此要求，並且在步驟 240 內利用寫入資料

至位址 X 來處理。雖然第 2A 圖內顯示接著在步驟 215-225 內寫入至快取，步驟 230-240 可與步驟 215-225 同時執行，或在任何其他合適時間上執行，本發明並不受限於任何特定實施。然後完成處理 20A。

處理 20B (第 2B 圖) 包含執行來讀取儲存至快取的資料之操作。在步驟 245 內，快取管理員 100 接收讀取儲存在快取磁碟 120 內位址 X 上資料的要求。運用快取元資料(未顯示)，在步驟 250 內，快取管理員 100 決定資料儲存在快取裝置 110 內的位址 Y 上。在步驟 255 內，快取管理員 100 發出讀取要求至快取裝置 110，來取得儲存在位址 Y 上的資料以及相關代表。在步驟 260 內，快取裝置 110 接收要求並在步驟 265 內處理，然後快取裝置 110 將結果回傳給快取管理員 100。

在步驟 270 內，快取管理員 100 決定是否可確認從快取當中取得的資料。在某些具體實施例內，這由再生資料的代表，並且將再生的代表與原始在步驟 215 內產生的代表作比較所達成。例如：步驟 270 可包含再生 MAC 或 CRC 給資料，並且與在步驟 265 內從快取取得的代表作比較。

若在步驟 270 內決定可確認該代表，然後處理前往步驟 275，其中取得自快取裝置 110 的資料用來滿足讀取要求，然後完成處理 20B。若在步驟 270 內判斷無法確

認該代表，則處理前往步驟 280，其中快取管理員 100 發出要求要快取裝置 110 驅逐(例如抹除或使其不可存取)位址 Y 上儲存的資料。然後在步驟 285 上，快取管理員 100 發出要求至快取磁碟 120 來從磁碟上位址 X 讀取資料。在步驟 290 內接收此要求並且在步驟 295 內處理，之後資料回傳給快取管理員 100。然後從位址 X 讀取的資料用來滿足步驟 299 內的讀取要求。然後完成處理 20B。

若運用確認資料驗證性的加密解決方案，用於產生代表的任何金鑰可在電源轉換期間寫入至快取裝置以外的位置，避免駭客取得金鑰來再生改變資料項目的代表。例如：在某些具體實施例內，金鑰可儲存在磁碟儲存裝置內(例如當電腦關機時)，以避免未授權的存取。不過，本發明並不限於這種實施，如同金鑰並不需要儲存，並且若已儲存，可位於任何合適的位置內。例如：儲存的金鑰可放置在作業系統所提供於系統開機期間(例如系統在 Windows 內登錄時)可用的任何組態儲存裝置內，或根據某些使用者輸入(例如密碼)來再生，如此不需要儲存金鑰。

吾人應該瞭解，上述確認快取裝置上所儲存資料項目的驗證性之具體實施例僅為範例，並且可使用任何合適的技術來確認驗證性。例如：不需要利用當資料項目寫

入時產生至少部分資料項目的代表來確認資料項目驗證性，其稍後在資料項目讀取時會再生。可運用可信賴地確定從快取讀取的資料項目經過驗證並與之前寫入快取的資料項目吻合之任何合適技術。本發明並不受限於任何特定實施。

II. 避免快取資料過期

如上面所討論，傳統作業系統無法偵測在電源轉換期間何時對磁碟上所儲存資料項目執行寫入操作造成快取內容過期。本發明的某些具體實施例提供用於偵測何時發生這些「離線寫入」的機構，藉此確定快取內容確實反映出電源轉換之後磁碟上所儲存的資料。

在某些作業系統(例如微軟公司提供的 Windows 作業系統家族)之下，特定電源轉換的涵義(例如待機與休眠模式)為在電源轉換期間無法修改非移除式儲存裝置(例如磁碟儲存裝置)上的資料。如此，對應至這種非移除式媒體上資料的快取內容一般並不會過期。不過，當電腦關機，會發生一些可能修改磁碟上資料的事情。例如：使用者可將磁碟開機到該電腦上的其他作業系統，或將磁碟連線至其他電腦，並且修改磁碟上儲存的資料。此外，如上面所討論，許多傳統作業系統的關機機構會在關機期間的某些點上，將快取裝置關閉並且不讓作業系

統存取，但是作業系統可繼續存取磁碟。如此，作業系統可更新磁碟上已經快取至快取裝置的資料項目。因為快取裝置已經關閉，作業系統就沒有辦法一併更新這些快取內容，如此快取內容變成過期。

若要管理這些與其他事情，本發明的某些具體實施例提供開始關機之後偵測磁碟上所儲存資料修改之技術，如此可更新、從快取裝置驅逐或處置因為這種修改造成的快取內容過期。

若要偵測關機操作期間發生在快取裝置關閉之後對磁碟儲存裝置執行的寫入，本發明某些具體實施例運用寫入記錄器組件。舉例而言，寫入記錄器組件可當成作業系統 I/O 路徑內的驅動器來實施，不過本發明並不限於這種實施。例如：寫入記錄器組件可為硬體式。針對範例，磁碟儲存硬體可提供一或多個介面，這些介面提供識別在特定時間週期內發生的修改集合，或識別在特定時間週期內是否發生修改之能力。例如：磁碟儲存硬體可提供自旋加快/電源開啟/開機計數器，這可用來推論至少某些儲存的資料項目已經更新，在此驅逐對應至磁碟上所儲存資料的快取內容(這不常發生，所以運用快取裝置仍舊有其優點)。本發明並不限於任何特定實施。

在某具體實施例內，寫入記錄器組件組態成當開始關機時啟用，並且追蹤對於磁碟儲存裝置執行的所有寫

入，直到完全關機。結果，當稍後重新啟動電腦，這些寫入會套用至快取內容。例如：當電腦重新啟動並且磁碟變成上線，然後快取管理員啟動，並且可追蹤寫入至磁碟。快取管理員可查詢寫入記錄器組件，來決定發生在快取裝置關閉之後的離線寫入，將這些寫入與快取管理員在開機時追蹤的部分合併，並且將合併的寫入集合套用至快取內容。將寫入套用至快取內容可包含例如：將對應至磁碟上資料的快取內容更新為指引寫入(例如對這些快取內容執行相同寫入操作)、驅逐這些快取內容、將兩者結合(例如將寫入操作套用至特定快取內容並驅逐其他部分)或執行某些其他操作。在將離線寫入套用至快取內容之後，可關閉寫入記錄器組件，並且快取裝置開始應付 I/O 要求。

第 3 圖說明用於追蹤離線寫入並且將這些寫入套用至快取內容的示範處理。尤其是，處理包含在電腦關機與後續再啟動期間由快取管理員 100、寫入記錄器 300、快取裝置 110 以及快取磁碟 120 所執行的操作。

在步驟 305 內，關於電腦關機期間所發生事項，快取管理員 100 啟動寫入記錄器 300，並且將「持續識別碼」供應給記錄器，這將識別將由寫入記錄器所追蹤的寫入操作之集合(例如產生)(底下詳細說明使用持續識別碼的範例)。在步驟 310 內，快取管理員 100 將持續識別碼，

以及記憶體內儲存的快取元資料寫入快取裝置 110。在關機處理的此點上，快取裝置 110 已經關閉並且變成快取管理員 100 無法存取。

在步驟 315 內，寫入記錄器 300 將在步驟 305 內傳遞給它的持續識別碼寫入快取磁碟 120，並且開始追蹤關機期間對快取磁碟 120 執行的任何寫入操作。例如：寫入記錄器 300 可在快取磁碟 120 上或在某些其他位置上建立記錄檔或一或多個其他資料結構，來指示磁碟上要執行寫入操作的位址，以及/或要寫入這些位址的資料。在完成步驟 315 時，電腦的關機操作已經完成。

此後，電腦重新啟動。針對部分開機處理，快取磁碟 120 上線，寫入記錄器 300 和快取管理員 100 重新起動。然後快取管理員 100 開始追蹤對快取磁碟 120 執行的寫入操作。例如：快取管理員 100 可建立記錄檔，並且儲存在快取裝置 110、快取磁碟 120 以及/或電腦記憶體(第 3 圖內未顯示)之上。在步驟 320 內，寫入記錄器 300 讀取在步驟 315 內記錄的磁碟變更，以及在步驟 315 內寫入快取磁碟 120 的持續識別碼。然後在步驟 325 內將磁碟變更與持續識別碼傳遞給快取管理員 100。

吾人應該瞭解，寫入記錄器 300 在快取裝置 110 關閉之後就無法追蹤對磁碟的所有寫入。例如：硬體資料毀損、不適時電源故障及/或寫入記錄檔發生問題都可讓寫

入記錄器 300 無法追蹤對磁碟所執行的所有離線寫入。在此案例中，寫入記錄器 300 可在步驟 325 內指示快取管理員 100，其無法可信賴地決定記錄已經完成並且確定記錄所執行的所有離線寫入。若發生此狀況，快取管理員 100 可隨著潛在不信賴，而驅逐整個快取內容或一部分(例如對應至寫入記錄器無法追蹤所有寫入操作的特定磁碟)。第 3 圖說明的剩餘部分假設寫入記錄器 300 可追蹤所有離線寫入。

在步驟 330 內，快取管理員 100 從快取裝置 110 內讀取快取元資料以及持續識別碼進入記憶體內。快取管理員 100 決定是否可確認持續識別碼(底下會進一步說明)。若否，快取管理員 100 可驅逐快取裝置 110 的整個內容或一部分(例如對應至無法確認持續識別碼的特定磁碟)。若可確認持續識別碼，則快取管理員 100 可將從電腦重新啟動以來對磁碟儲存裝置執行的寫入操作與寫入記錄器 300 所追蹤的任何寫入操作合併。例如：若一或多個記錄指出資料寫入至磁碟上每一位址，則快取管理員 100 可選擇對每一位址執行的最後更新並寫入記憶體內。

在某些具體實施例內，寫入記錄器 300 可組態成在電腦重新啟動之後繼續記錄寫入，如此快取管理員 100 不需要記錄該點之後執行的寫入並與寫入記錄器 300 追蹤

的寫入合併。取而代之，寫入記錄器 300 可簡單提供所有寫入記錄給快取管理員 100。

運用步驟 330 內讀取的快取元資料，在步驟 335 內快取管理員 100 將寫入集合套用至快取裝置 110 的內容。如上面所討論，套用寫入可包含驅逐快取內容、更新快取內容、兩者同時或執行某些其他操作。例如：利用驅逐對應快取內容可套用步驟 315 內寫入記錄器 300 所追蹤的離線寫入，同時利用更新對應快取內容可套用自電腦重新啟動以來快取管理員 100 所追蹤的寫入來反應寫入。將寫入操作套用至快取內容可用任何合適方式來執行，本發明並不受限於任何特定實施。

在步驟 335 完成時，第 3 圖的處理已完成。

吾人應該瞭解，本發明並不受限於運用組態成在開始關機時就變成啟用的寫入記錄器組件，也可/或另外追蹤關機期間未發生的寫入操作。例如；在某些實施當中，快取裝置很容易變成一段期間不可存取。例如：若透過一或多個網路存取快取裝置，則連線可能中斷，或若快取裝置可從電腦移除，則可能發生突然(例如意外)移除。結果，某些具體實施例可運用寫入記錄器來追蹤對磁碟執行的全部(或部分)寫入，不僅只是發生在關機期間，並且組態成定期擷取快取「快照」的快取裝置仍舊上線。如此，若快取在某些時間變成不可存取並且稍後

重新連線，則使用寫入記錄器所追蹤的寫入操作來更新最後的快取快照，而非完全重建。

吾人也應該瞭解，雖然第 3 圖的示範處理可偵測作業系統在關機期間執行的離線寫入，不過可能需要其他措施來偵測關機完成之後對磁碟執行的離線寫入。例如當使用者在關機之後將磁碟開機到其他作業系統，或在關機之後從電腦上移除磁碟並連線至其他電腦，以及修改磁碟上儲存的資料，如此會發生這種寫入。

在瞭解嘗試追蹤發生在關機之後的離線寫入(例如由其他作業系統)之相關困難度，本發明某些具體實施例嘗試避免發生這種情況。例如：某些具體實施例嘗試讓作業系統(未提供關機後的寫入記錄器組件)無法存取特定磁碟。這用任何許多方式都可達成。

在某些具體實施例內，寫入記錄器 300 可標示磁碟，使其變成不可由未提供寫入記錄器組件來追蹤離線寫入的作業系統安裝。例如：寫入記錄器 300 可修改磁碟識別碼，指示磁碟上所使用的檔案系統類型。在此範圍內，精通此技術的人士將瞭解，磁碟識別碼可讓作業系統識別磁碟上用來儲存資料的檔案系統類型，藉此讓作業系統瞭解磁碟上儲存的資料結構、何處找尋檔案等等。例如：若磁碟識別碼指示磁碟上使用 NT File System (NTFS) 檔案系統來儲存資料，則嘗試安裝該磁碟的其他作業系

統瞭解需要解析 NTFS 檔案系統並存取其上的資料。若提供的磁碟識別碼未指示磁碟上用來儲存資料的檔案系統類型，則大多數作業系統無法安裝該磁碟，因為並無可信賴的方式來瞭解其上所儲存資料的結構。如此，本發明某些具體實施例修改磁碟的磁碟識別碼使其不可存取，藉此避免使用者將磁碟開機至其他作業系統並且對磁碟上儲存的資料進行離線變更。

吾人知道，即使磁碟識別碼已修改，某些作業系統還是可識別磁碟上用來儲存資料的檔案系統類型，本發明某些具體實施例提供機構來偵測作業系統何時安裝磁碟。在此範圍內，若要安裝磁碟，任何作業系統都必須更新磁碟識別碼(例如指示已經運用 NTFS 檔案系統來將資料儲存在磁碟上)，以允許存取其上的資料。開機時應該可輕易偵測到任何這種更新。若偵測到這種更新，本發明某些具體實施例假設磁碟內容從上次關機之後已遭修改，並且驅逐對應至磁碟上所儲存資料的快取內容。

本發明的某些具體實施例提供一種磁碟可開機到也運用寫入記錄器組件的其他作業系統內之能力。例如：若磁碟從運行提供寫入記錄器組件的作業系統之電腦中移除，並且將該磁碟開機至提供寫入記錄器組件的其他作業系統內，則該其他作業系統可組態成知道已變更的磁碟識別碼指示該磁碟已經快取。結果，該其他作業系統

可新增至由第一作業系統所建立的離線寫入記錄(例如儲存在磁碟上)。

上述具體實施例設計成讓特定作業系統無法安裝磁碟對於依賴磁碟識別碼來執行特定功能的特定應用程式(例如備份應用程式)而言會產生問題。對於這些應用程式而言，若磁碟識別碼已經變更，則無法辨識該磁碟，因此無法備份。因此，本發明某些具體實施例提供一種機制來決定關機之後是否安裝檔案系統。若是，則假設已經對檔案系統內的資料進行變更，並且驅逐對應至檔案系統內資料的所有快取內容。

某些具體實施例可利用，將關機上的檔案系統紀錄變成嘗試安裝檔案系統的作業系統需要以某些方式修改記錄(例如變更其位置、新增新輸入等等)之狀態，來偵測關機之後檔案系統的安裝。例如：注意寫入記錄器 300 可在檔案系統未安裝時(例如在記錄本身內)，成為記錄離線寫入檔案系統記錄位置及/或內容的作業一部分。因為嘗試安裝檔案系統的任何作業系統都應變更該記錄(例如若檔案系統為 NTFS 檔案系統，則嘗試安裝該檔案系統的作業系統應新增輸入至該記錄)，若記錄在再啟動時未變更，則假設在電源轉換期間其他作業系統未安裝此檔案系統，如此對應至檔案系統內所儲存資料的快取內容未過期。相反地，若記錄已經用某些方式變更(例如

位置已經變更、已經新增輸入等等)，則假設檔案系統已經由其他作業系統安裝，儲存其中的資料已經變更，則對應至檔案系統內所儲存資料的快取內容已經過期。如此，則驅逐這些快取內容。

除了提供機制避免離線寫入以外，本發明某些具體實施例提供管理快取內容不一致產生的能力。快取內容不一致產生的建立原因有許多，有一範例發生在當具有第一和第二快取裝置連線其上的第一和第二電腦運用此處所述技術來維持快取內容通過電源轉換時，若第二快取裝置連線至第一電腦(或第一快取裝置連線至第二電腦)並且第一電腦已經重新啟動，則從第二快取裝置會提供不正確的資料來滿足 I/O 要求。這是因為第一電腦的作業系統認為第二快取裝置的內容已驗證(因為從快取回傳的資料再生代表與原始產生的代表吻合)，並且未過期(因為離線寫入套用至快取內容)。其他範例為若第一快取裝置已經連線至電腦，該電腦已經關機(藉此持續快取內容)，然後該電腦重新啟動，已經連線第二快取裝置，並且該電腦再次關機(藉此再次持續快取內容)。然後若電腦在此重新啟動並且連線第一快取裝置，因為並無可信賴的方式來決定該第一快取裝置未儲存最後產生的快取內容，所以會提供不正確資料來滿足 I/O 要求。

某些具體實施例提供識別快取內容不一致產生的能

力，如此並不會誤用之前持續至最後一次關機的快取內容來滿足 I/O 要求。在某些具體實施例內，於任何許多方式中，透過唯一持續識別碼提供此能力，該識別碼在開始關機時產生(屬於範例)。例如：為此可運用 GUID 及/或加密隨機編號產生器。如上面參考第 3 圖的說明，持續識別碼可儲存在快取裝置上(例如在快取元資料內或一起)以及在電腦上(例如在磁碟及/或記憶體上)，並且在電腦啟動時確認(例如利用比較兩版本)。若驗證失敗，則可驅逐快取內容，並呈現之前的持續快取產生。

針對上面討論的驗證金鑰，用來產生持續識別碼的任何金鑰都可在電源轉換期間寫入至快取裝置以外的位置。例如：在某些具體實施例內，寫入記錄器組件可寫入金鑰以及持續識別碼到磁碟儲存裝置內(例如在關機時)。不過，本發明並不受限於這種實施，精通此技術的人士可想像其中可儲存金鑰的許多替代位置。例如：金鑰可保留在作業系統所提供的任何組態儲存裝置內，使其可在系統開機期間使用(例如 Windows 內的登錄)。

III. 快取元資料

如上面所說明，快取元資料可提供儲存資料項目的磁碟位址與快取裝置上快取這些資料項目的對應位址間之映射。本發明某些具體實施例提供儲存快取元資料的能

力，其顯著減少在系統運轉時間操作期間需要儲存快取元資料的記憶體數量。

此外，某些具體實施例提供讓快取元資料可信賴通過電源轉換或將快取裝置離線的任何其他事件(例如從電腦當中移除快取裝置、網路無法使用而讓網路快取裝置無法存取等等)之技術，如此當電腦重新啟動及/或快取裝置上線時可信賴地存取快取內容。在此範圍內，吾人應該了解，在特定電源轉換類型之下(例如待機與休眠模式)，因為在待機和休眠轉換期間保留了記憶體的內容，所以可接受僅將快取元資料儲存在記憶體內(即是RAM)。不過在再啟動期間，並不保留系統記憶體的內容。如此，本發明某些具體實施例提供在關機期間將快取元資料儲存在某些非揮發性媒體上，然後在再啟動時復原。例如：快取元資料可儲存在快取裝置上以及/或一個或多個分開的非揮發性媒體上。進一步，某些具體實施例可從其他當中取得快取元資料的某一部分，如此不需要儲存所有快取元資料。

某些具體實施例運用上面第 I 節內說明的技術來確認快取元資料的驗證性，如此偵測並避免當快取裝置離線時(例如電腦關機期間、從電腦中移除快取裝置、網路無法使用讓網路快取裝置無法存取等等)意外或惡意修改元資料。例如：當快取裝置上線，快取管理員可運用上

面參考第 2A 圖至第 2B 圖所說明的技術，確認其載入記憶體內的元資料之驗證性。若快取元資料的驗證性無法確認，則根據磁碟上所儲存資料更新對應的快取內容、驅逐或如上述來處理。

在某些具體實施例內，快取元資料可壓縮來減少關機與再啟動載入期間要儲存的元資料量。因為元資料壓縮需要儲存內涵有關元資料資訊的個別資訊片段(例如快取內的標題)，可運用上述技術來在再啟動時確認此資訊的驗證性。

本發明某些具體實施例提供以大幅減少任何時間上記憶體內所儲存快取元資料量之方式，來儲存快取元資料之技術，藉此減少將快取元資料載入記憶體以及從中載出所需的時間(例如運轉時間期間以及開機/關機操作期間)，並且大幅減少快取元資料的記憶體「佔用量」。在此範圍內，吾人應該瞭解，在使用具有相當大儲存容量的快取裝置時，需要顯著的元資料量來管理快取內容。例如：具有 16 GB 儲存容量的快取裝置最多可儲存 32 GB 的壓縮資料。在某些實施當中，在快取元資料中以代表 4 kB(kilobyte)磁碟儲存空間的「資料單元」來反映磁碟位址。如此，若要追蹤 32 GB 資料的位置，則需要八百萬個別資料單元。若在快取元資料內使用十六位元組映射(即是從磁碟位址到快取位址)來呈現每一個八百萬資

料單元，則這些映射需要 128 MB 的儲存空間。申請人已經瞭解，在記憶體內儲存 128 MB 的快取元資料會佔用許多電腦內非必需的大部分記憶體。此外，在關機時從記憶體將 128 MB 快取元資料寫入非揮發性媒體以及在再啟動時將 128 MB 快取元資料從非揮發性媒體復原到記憶體所需之時間相當耗時，並且消耗大量處理資源。

吾人明白無法輕易減少快取元資料量，本發明某些具體實施例提供設計來減少儲存快取元資料所需儲存資源，以及在關機和開機時儲存與復原快取元資料所需的時間與處理資源之技術。

在某些具體實施例內，利用將快取元資料儲存在一或多個階層資料結構內(例如樹狀、多層陣列等等)就可達成此目的。運用階層資料結構可允許低階層儲存在非揮發性媒體上(例如快取裝置)，而只有較高階層儲存在記憶體內。例如：在某些具體實施例內，只有較高階層儲存在記憶體內，如此快取元資料在記憶體內的「佔用量」可大幅減少，即使當需要支援具有顯著儲存容量的快取裝置之快取元資料量已經完全儲存。當然，只儲存快取元資料內的較高階層，如某些具體實施例可提供將較低階層上保存的某些資訊儲存在記憶體內，如此減少與重覆存取此資訊相關的 I/O 負擔。本發明並不受限於以任何特定方式來實施。

在系統操作期間，隨著處理讀取要求，從非揮發性媒體中讀取(即是從較低階層)來執行讀取操作的快取元資料可「分頁」到(即是從儲存媒體讀取到)記憶體內，如此可讓後續讀取要求更快速存取至相同磁碟/快取位址。當稍後電腦關機及/或快取裝置離線，則只有儲存在較高階層內的快取元資料，以及要儲存在較低階層內分頁到記憶體內的快取元資料，需要儲存到非揮發性媒體內。如此，在關機時將快取元資料從記憶體移動至非揮發性儲存裝置，以及在再啟動時將快取元資料從非揮發性儲存裝置復原至記憶體所需的時間顯著減少。

本發明某些具體實施例運用 B+樹狀來儲存至少部分快取元資料。精通此技術的人士將瞭解，B+樹狀可運用大量分支，因此減少所運用階層內的層數。運用上面的範例，若要在快取元資料內呈現八百萬資料單元並且要運用具有兩百個分支的 B+樹狀(如此階層內每一節點都有兩百個「子代」節點)，則只具有三層的資料結構就足以儲存元資料：最高層上單一個「根」節點、第二層上兩百個節點以及在第三層上四萬個節點，這四萬個節點每一都包含至兩百個資料單元(或總數八百萬個資料單元)的指引器。

第 4 圖說明此範例 B+樹狀，其包含根節點 400、第二層節點 410₁₋₂₀₀ 以及第三層節點 420₁₋₂₀₀。每一節點包含

兩百個元件，每一都由指引器分開至較低階層上的節點。例如：根節點 400 內的元件 402 由指引器 401 和 403 界定範圍。與已知金鑰(例如磁碟位址)相關的值(例如磁碟位址)可遵循節點內元件左邊或右邊的指引器所決定，若金鑰小於該元件，則遵循元件左邊的指引器，若金鑰大於該元件，則遵循元件右邊的指引器。例如：若要決定小於元件 402 的金鑰值，則遵循指引器 401 到達第二層節點 410_1 ，若要決定大於元件 402 但是小於元件 404 的金鑰值，則遵循指引器 403 到達第二層節點 410_2 (未顯示)，以此類推。類似地，在第二層節點上，遵循元件左邊或右邊的指引器(根據金鑰小於或大於節點內的元件)來到達第三層節點。在第三層上，遵循最終指引器(再次根據金鑰小於或大於節點內元件)到達該值，第三層上每一指引器都參照快取元資料內八百萬資料單元的其中之一。

吾人應該瞭解，具有大量分支的 B+ 樹狀提供相當「平坦」的階層，幾乎所有節點都位於階層的底層。也就是，樹狀內總共有 40,201 個節點，其中 40,000 個位於最底層。本發明某些具體實施例利用在開機時只將頂端兩層復原至記憶體，而階層最底層內的快取元資料則儲存在快取裝置內直到需要(例如可依照讀取要求的處理需求載入記憶體、緩慢載入等等)，來獲取其優點。因為記憶

體內只儲存部分階層資料結構，所以比起全部或大多數快取元資料都儲存在記憶體內，快取元資料只佔用非常少的記憶體空間。此外，當電腦關機，則只有頂端兩層上的資料以及在操作時載入記憶體的資料需要儲存在快取裝置上。結果，可迅速並有效率執行開機與關機操作。

如此，本發明某些具體實施例提供儲存在記憶體內一個階層上(在上面範例中為第二層)節點內的指引器，其參照儲存在快取裝置上其他階層上(在上面範例中為第三層)的節點。例如：當接收快取資料項目的讀取要求，本發明的具體實施例遵循指引器通過記憶體內儲存的一或多階層，然後到快取內所儲存較低層上的元資料，來決定資料項目儲存在快取內的位址。在某些具體實施例內，一旦決定資料項目的快取位址，則可儲存在記憶體內，如此可執行後續讀取該項目的要求，而不用從快取裝置當中讀取快取元資料。

第 5 圖說明根據本發明某些具體實施例用於管理快取元資料的示範系統。第 5 圖說明記憶體 500 和快取裝置 110，這兩者都可存取至電腦(未顯示)。當電腦已啟動，在操作 505 內將像是 B+ 樹狀這類包含一或多層階層資料結構的快取元資料載入記憶體 500 內。運用上述範例來說明，若快取元資料內呈現八百萬資料單元，則可使用三層階層資料結構來儲存快取元資料，然後階層的上兩

層可載入記憶體 500 內。當然，若元資料內呈現多於或少於八百萬資料單元，並且要使用多於或少於三層的階層資料結構，則可將不同的階層數載入記憶體 500 內。

此後，當指引讀取要求至快取內保留的資料項目，則利用存取快取裝置 110 內所儲存階層內儲存的快取元資料，來決定其上儲存資料項目的快取位址。然後將此快取元資料儲存在記憶體 510 內，如此可執行對資料項目的後續讀取或寫入，而不用讀取快取裝置上儲存的快取元資料，來決定其上儲存資料項目的快取位址。取而代之，可從記憶體讀取快取位址，這要比讀取快取來得迅速。

稍後，當電腦已關機，則步驟 515 將儲存在記憶體內的快取元資料(即是儲存在操作 505 內載入記憶體中階層的元資料，以及用於在操作 510 內滿足讀取要求寫入記憶體內的任何元資料)載入快取裝置 110。在記憶體內儲存相當少量快取元資料的結果之下，可迅速關機，而不需要後續處理資源。

吾人應該瞭解，B+樹狀為許多種可用來儲存快取元資料的資料結構的其中之一，因此可運用其他種資料結構(例如像是 AVL 樹狀階層結構、紅黑樹狀、二進位搜尋樹狀、B 樹狀及/或其他階層和非階層資料結構)。本發明並不限於運用任一資料結構或資料結構的組合來儲存

快取元資料。

某些具體實施例可提供在任何時間上讓快取元資料的「目標量」維持在記憶體內。目標量可用任何合適的方式決定。例如：目標量可為電腦可用實際記憶體數量的百分比。例如：若電腦具備 1 GB 記憶體，則任何時間上在記憶體內可儲存 2 MB 快取元資料(範例)。如此，當電腦已關機，則只需要將 2 MB 快取元資料載入快取裝置。

在某些具體實施例內，快取元資料可循環進出記憶體。例如：若快取元資料的目標量已經儲存在記憶體內，並且已經執行需要從快取裝置當中讀取快取元資料的讀取，則元資料「分頁」至記憶體，並且抹除其他快取元資料(例如最不常存取的元資料)。例如：快取元資料可在寫入快取裝置之後抹除。另外，系統可決定自從快取元資料上次寫出之後是否變更，並且若為否，則只要抹除，如此免除寫入快取元資料所需的時間與處理資源。運用上述技術，可維持快取元資料佔用記憶體內少量的「佔用量」。

第 6 圖說明一個範例。尤其是，第 6 圖內顯示的處理包含由快取管理員 100 使用上述技術所執行，來讀取快取元資料的操作。

在處理開始時，在步驟 605 內接收要求來讀取磁碟位址 X 上儲存的資料。在步驟 610 上，決定是否可從記憶

體內儲存的快取元資料識別其上儲存資料的快取位址。若可，處理前往步驟 615，其中已決定已識別的快取位址，然後用來在步驟 620 內發出讀取要求至快取裝置 110。然後完成處理。若使用記憶體內儲存的快取元資料無法識別快取位址，則處理前往步驟 625，其中從快取裝置 110 讀取快取元資料，來決定其上儲存資料的快取位址。運用在步驟 625 內識別的快取偏移，則發出讀取要求給在步驟 620 內識別的快取偏移，然後完成處理。

吾人應該瞭解，將快取元資料儲存在快取裝置上並不會加快開機與關機期間載入與復原快取元資料的處理，但是亦可加快開機與關機期間執行的系統操作。在此範圍內，關機與開機通常牽涉到多次存取特定資料項目，並且對快取裝置執行兩次讀取操作通常要快過對磁碟儲存裝置執行一次讀取操作。結果，若在關機及/或開機期間存取的資料項目以及指定其位置的元資料都儲存在快取內，則比起若資料項目儲存在磁碟上可更迅速存取資料項目，因為讀取快取兩次(即是一次存取快取元資料來決定項目位置，另一次存取項目本身)通常比起讀取磁碟一次更迅速。如此，可促進關機與開機期間執行的個別操作。進一步，若在從快取第一次讀取快取元資料期間，其上儲存該項目的位址分頁進入記憶體內，則因為讀取記憶體執行起來通常比讀取快取更迅速，所以可更迅速

執行資料項目的後續讀取。

在一或多個電腦系統上，像是第 7 圖內所示的示例性電腦系統 700，可實施實現本發明特徵的許多系統與方法態樣。電腦系統 700 包含輸入裝置 702、輸出裝置 701、處理器 703、記憶體系統 704 和儲存裝置 706，這些都透過互連機構 705 來直接或間接耦合，其中該機構包含一或多個匯流排、開關、網路及/或任何其他合適的互連。輸入裝置 702 接收來自使用者或機器(例如操作員)的輸入，並且輸出裝置 701 顯示或傳輸資訊給使用者或機器(例如液晶顯示器)。處理器 703 通常執行稱為作業系統的電腦程式(例如 Microsoft Windows 家族作業系統，或任何其他合適的作業系統)，其控制其他電腦程式的執行，並且提供排程、輸入/輸出與其他裝置控制、帳號管理、編輯、儲存指派、資料管理、記憶體管理、通訊以及資料流控制。共同來說，處理器與作業系統定義應用程式所程式編輯以及其他電腦程式語言所撰寫之電腦平台。

處理器 703 也可執行一或多個電腦程式來實施許多功能，這些電腦程式可用任何電腦程式語言類型撰寫，包含程序程式編輯語言、物件導向程式編輯語言、巨集語言或這些的組合。這些電腦程式可儲存在儲存系統 706 內。儲存系統 706 可將資訊保留在揮發性或非揮發性媒

體上，並且可固定或可移除。第 8 圖內會更詳細顯示儲存系統 706。

儲存系統 706 通常包含電腦可讀取和可寫入非揮發性記錄媒體 801，其上儲存定義電腦程式或程式所使用資訊的信號。例如：媒體可為磁碟或快閃記憶體。通常，處理器 703 導致資料從非揮發性記錄媒體 801 讀取進入揮發性記憶體 802 (例如隨機存取記憶體或 RAM) 的操作，允許處理器 703 比媒體 801 更快速存取該資訊。記憶體 802 可位於儲存系統 706 內，如第 8 圖內所示，或在記憶體系統 704 內，如第 7 圖內所示。處理器 703 一般操縱積體電路記憶體 704、802 內的資料，然後在處理完成之後將資料複製到媒體 801。已知許多機構用於管理媒體 801 與積體電路記憶體元件 704、802 之間的資料移動，並且本發明並不受限於此。本發明也不受限於特定記憶體系統 704 或儲存系統 706。

進一步，本發明具體實施例也不受限於運用快取管理員組件，其實施為作業系統的 I/O 堆疊內之驅動器。另外或額外可使用任何合適的組件或組件組合，每一都可由作業系統或一或多個單機組件所實施。本發明並不受限於任何特定實施。

上述本發明的具體實施例可用任何許多方式實施，例如：可使用硬體、軟體或這兩者的組合來實施上面討論

的功能性。當在軟體內實施時，軟體程式碼可在單一電腦或分散在多部電腦之間的任何合適之處理器或處理器集合上執行。在此範圍內，吾人應該瞭解，執行此處所述功能的任何組件或組件集合一般可考慮為控制上面討論功能的一或多個控制器。一或多個控制器可用許多方式實施，像是使用專屬硬體，或運用一或多個使用微程式碼或軟體來程式編輯以執行上述功能的處理器。在此控制器儲存或提供資料給系統作業，這種資料可儲存在中央倉儲內、許多倉儲內或這些的組合內。

進一步，吾人應該瞭解，(用戶端或伺服器)電腦可用任意形式具體實施，像是框架安裝型電腦、桌上型電腦、膝上型電腦、平板電腦或其他種電腦。此外，(用戶端或伺服器)電腦可內嵌在一般看起來不像電腦但是具有合適處理能力的裝置內，包含個人數位助理(Personal Digital Assistant, PDA)、智慧型手機或任何其他合適的可攜式或固定式電子裝置。

另外，(用戶端或伺服器)電腦可具有一或多個輸入和輸出裝置。這些裝置除此以外可用來呈現使用者介面。可用來提供使用者介面的輸出裝置範例包含視覺呈現輸出的印表機或顯示器螢幕，以及聲音呈現輸出的喇叭或其他發聲裝置。可用來當成使用者介面的輸入裝置範例包含鍵盤以及指標裝置，像是滑鼠、觸控板以及數位板。

針對其他範例，電腦可透過語音辨識或其他聲音格式來接收輸入資訊。

這種電腦可用任何合適形式的一或多個網路互連，包含區域網路或廣域網路，像是企業網路或網際網路。這種網路可根據任何合適的技術，並且可根據任何合適的通訊協定來運作，並可包含無線網路、有線網路或光纖網路。另外，此處描述的許多方法或處理都可編碼為可在佈署許多作業系統或平台任一的一或多個處理器上執行之軟體。

此外，軟體可用任何許多合適的程式編輯語言及/或傳統程式編輯或描述工具撰寫，並且也可編譯為可在框架或虛擬機器上執行的可執行機器語言碼或中介碼。

在此範圍內，本發明可具體實施為用一或多種程式編碼的儲存媒體(或多重儲存媒體)(例如電腦記憶體、一或多個軟碟片、硬碟、光碟、磁帶、快閃記憶體、場可程式編輯閘極陣列的電路組態或其他半導體裝置，或者其他電腦儲存媒體)，當在一或多部電腦或其他處理器上執行時，會執行實施如上討論本發明許多具體實施例的方法。該儲存媒體可運輸，如此上面儲存的程式可載入一或多部不同的電腦或其他處理器，以實施上面討論的本發明許多態樣。

此處所用的「程式」或「軟體」等術語一般就是任何

種類的電腦程式碼或電腦可執行指令集，可用來程式編輯電腦或其他處理器以實施上面討論的本發明許多態樣。此外，吾人應該瞭解，根據本具體實施例的一個態樣，執行一或多個電腦程式時執行本發明方法並不需要位於單一電腦或處理器上，可用模組方式分散在許多不同電腦或處理器上，以實施本發明的許多態樣。

電腦可執行指令可為任何形式；像是由一或多部電腦或其他裝置所執行的程式模組。一般而言，程式模組包含例式、程式、物件、組件、資料結構等以及執行特定工作或實施特定摘要資料類型。一般而言，程式模組的功能性可在許多具體實施例內結合或分散。

本發明的許多態樣都可單獨使用、結合使用或以前面說明的具體實施例當中未特定討論的許多配置來使用，因此並不受限於應用在前述說明中公佈或圖式內說明的組件細節與配置內。例如：一個具體實施例內說明的態樣可用任何方式與其他具體實施例內說明的態樣結合。

申請專利範圍內使用像是「第一」、「第二」、「第三」等等序詞來修改申請專利範圍項目並未暗示一個申請專利範圍項目的任何優先順序、優先權或順序高於其他項目，或暗示所執行方法中步驟的暫時順序，而是只用來當成標籤以區別具有特定名稱的一個申請專利範圍項目與具有相同名稱的另一個項目(但是使用序詞)，以區別

申請專利範圍項目。

另外，此處所使用的語法與術語僅供說明，不應看成限制。此處所使用的「包含」、「包括」或「具有」、「內含」、「牽涉」以及其變化都用於囊括之後所列的項目和等同項以及附加項。

在說明本發明至少一個具體實施例的許多態樣之下，吾人將瞭解，精通此技術的人士將迅速明瞭許多改變、修改和改善。這些改變、改善與改善都屬於本發明的一部分，並且屬於本發明的精神與範疇，因此前述說明與圖式都僅為範例。

【圖式簡單說明】

第 1A 圖至第 1B 圖為說明根據先前技術寫入至與讀取自一快取裝置之技術的方塊圖；

第 2A 圖至第 2B 圖為說明根據本發明某些具體實施例寫入至與讀取自一快取裝置之示例性技術；

第 3 圖為說明根據本發明某些具體實施例，用於確定快取資料確實反映出電源轉換之後磁碟上所儲存資料之示例性技術的方塊圖；

第 4 圖為說明根據本發明某些具體實施例用於儲存快取元資料的示例性技術之方塊圖；

第 5 圖為說明根據本發明某些具體實施例用於快取元

資料的示例性儲存操作之方塊圖；

第 6 圖為說明根據本發明某些具體實施例用於使用快取裝置應付讀取要求的示例性技術之流程圖；

第 7 圖為說明可用於實施本發明態樣的示範電腦之方塊圖；以及

第 8 圖為說明其上可記錄實施本發明態樣的指令之示範電腦記憶體方塊圖。

【主要元件符號說明】

10A 處理	315 步驟
10B 處理	320 步驟
20A 處理	325 步驟
20B 處理	330 步驟
100 快取管理員	335 步驟
101 寫入位址 X	400 根節點
102 寫入位址 Y	401 指引器
103 操作	402 元件
104 操作	403 指引器
110 快取裝置	404 元件
120 快取磁碟	410 ₁₋₂₀₀ 第二層節點
300 寫入記錄器	420 ₁₋₂₀₀ 第三層節點
305 步驟	500 記憶體
310 步驟	505 操作

510 操作

515 步驟

700 電腦系統

701 輸出裝置

702 輸入裝置

703 處理器

704 記憶體系統

705 互連機構

706 儲存裝置/儲存系統

801 非揮發性記錄媒體

802 揮發性記憶體

104年4月28日修正 (本)

七、申請專利範圍：

1. 一種用於操作耦合於一儲存媒體及一快取裝置的一電腦之方法，該電腦包含一作業系統，該方法包含以下步驟：

(A) 啟動一寫入記錄器組件，該寫入記錄器組件係經配置以用於識別對該儲存媒體所實施而不對該快取裝置所實施的寫入操作；

(B) 為回應將一資料項目寫入於該儲存媒體的一位址處的一要求，將導致該資料項目被寫入於該儲存媒體的該位址處及該快取裝置的一對應位址處；

(C) 擷取寫入於該快取裝置的該對應位址處的該資料項目的一快照(snapshot)；

(D) 在重啟動(reboot)該作業系統之後，決定是否寫入於該快取裝置的該對應位址處的該資料項目能被可靠地用於滿足將該資料項目從該儲存媒體的該位址處讀取出的一要求，該決定之步驟包含透過該寫入記錄器組件來決定是否在完成步驟(B)之後，修改過寫入於該儲存媒體的該位址處的該資料項目；以及

(E) 若是在步驟(D)中決定出寫入於該快取裝置的該對應位址處的該資料項目是不能被可靠地用於滿足將該資料項目從該儲存媒體的該位址處讀取出

的該要求，則將導致使用來自該寫入記錄器組件的資料及該快照來更新該快取裝置。

2. 如申請專利範圍第 1 項所述之方法，其中該等步驟 (A)-(E) 是由該作業系統所實施。
3. 如申請專利範圍第 1 項所述之方法，其中該儲存媒體包含一磁碟儲存裝置。
4. 如申請專利範圍第 1 項所述之方法，其中該快取裝置是可從該電腦當中移除的。
5. 如申請專利範圍第 1 項所述之方法，其中為回應將該資料項目從該儲存媒體的該位址處讀取出的該請求，來實施該步驟 (D)，並且該方法另包含以下步驟：
驗證出在完成該步驟 (B) 之後，未修改過寫入於該快取裝置的該對應位址處的該資料項目。
6. 如申請專利範圍第 5 項所述之方法，其中步驟 (B) 另包含以下步驟：產生寫入於該儲存媒體的該位址處的至少該資料項目的一代表 (representation)，並且將該代表寫入於該快取裝置，其中在步驟 (D) 中的該驗證步驟包含以下步驟：

(D1) 取得曾寫入於該快取裝置的該代表；

(D2) 重新產生該代表；以及

(D3) 將該取得的代表與該重新產生的代表做比較，以決定是否寫入於該快取裝置的該對應位址處

的該資料項目能被可靠地用於滿足將寫入於該儲存媒體的該位址處的該資料項目讀取出的該要求。

7. 如申請專利範圍第 1 項所述之方法，其中當決定出該作業系統是正在關機(shut down)時，該作業系統係經配置以用於啟動該寫入記錄器組件。
8. 一種電腦儲存媒體，該電腦儲存媒體是儲存指令的一記憶體裝置，當該等指令由包含一作業系統並且耦合於一快取裝置的一電腦所執行時，將導致該電腦來實施包含下列步驟的一方法：

(A) 啟動一寫入記錄器組件，該寫入記錄器組件係經配置以用於識別對該電腦儲存媒體所實施而不對該快取裝置所實施的寫入操作；

(B) 為回應將一資料項目寫入於該電腦儲存媒體的一位址處的一要求，將導致該資料項目被寫入於該電腦儲存媒體的該位址處及該快取裝置的一對應位址處；

(C) 擷取寫入於該快取裝置的該對應位址處的該資料項目的一快照；

(D) 在重啟動該作業系統之後，決定是否寫入於該快取裝置的該對應位址處的該資料項目能被可靠地用於滿足將該資料項目從該電腦儲存媒體的該位址處讀取出的一要求，該決定之步驟包含透過該寫

入記錄器組件來決定是否在完成步驟(B)之後，修改過寫入於該電腦儲存媒體的該位址處的該資料項目；以及

(E) 若是在步驟(D)中決定出寫入於該快取裝置的該對應位址處的該資料項目是不能被可靠地用於滿足將該資料項目從該電腦儲存媒體的該位址處讀取出的該要求，則將導致使用來自該寫入記錄器組件的資料及該快照來更新該快取裝置。

9. 如申請專利範圍第 8 項所述之電腦儲存媒體，其中該步驟(E)另包含以下步驟：當在更新之後決定出寫入於該快取裝置的該對應位址處的該資料項目是不能被可靠地用於滿足將該資料項目從該電腦儲存媒體的該位址處讀取出的該要求時，將導致該資料項目是從該電腦儲存媒體的該位址處所讀取出。
10. 如申請專利範圍第 8 項所述之電腦儲存媒體，其中該步驟(E)另包含以下步驟：當在更新之後決定出寫入於該快取裝置的該對應位址處的該資料項目是不能被可靠地用於滿足將該資料項目從該電腦儲存媒體的該位址處讀取出的該要求時，將導致該資料項目是從該快取裝置的該對應位址處所驅逐出(evict)。
11. 如申請專利範圍第 8 項所述之電腦儲存媒體，其中該步驟(B)另包含以下步驟：產生寫入於該電腦儲存媒

體的該位址處的至少該資料項目的一代表，並且將該代表寫入於該快取裝置，其中在步驟(D)中的該決定步驟包含以下步驟：

(D1) 取得曾寫入於該快取裝置的該代表；

(D2) 重新產生該代表；以及

(D3) 將該取得的代表與該重新產生的代表做比較，以決定是否寫入於該快取裝置的該對應位址處的該資料項目能被可靠地用於滿足將寫入於該電腦儲存媒體的該位址處的該資料項目讀取出的該要求。

12. 如申請專利範圍第 8 項所述之電腦儲存媒體，其中該步驟(B)另包含以下步驟：產生寫入於該電腦儲存媒體的該位址處的至少該資料項目的一代表，並且將該代表寫入於該快取裝置，其中產生一代表的該步驟包含運用一加密技術。
13. 如申請專利範圍第 8 項所述之電腦儲存媒體，其中該步驟(B)另包含以下步驟：運用快取元資料來決定將該資料項目寫入於該快取裝置上的該對應位址，並且其中該步驟(D)另包含以下步驟：決定是否在步驟(B)之後，修改過該快取元資料。
14. 一種電腦，包含：
 - 一儲存媒體，其具有至少一個位址，將資料儲存於該位址處；

一 快取裝置，其具有至少一個位址，將資料儲存於該位址處；

一 寫入記錄器組件，其係經配置以用於識別對該儲存媒體所實施而不對該快取裝置所實施的寫入操作；

至少一個處理器，其透過一作業系統經程式編輯以下步驟：

為回應將一資料項目寫入於該儲存媒體的一第一位址處的一要求，將導致該資料項目被寫入於該儲存媒體的該第一位址處及該快取裝置的一對應位址處；

擷取寫入於該快取裝置的該對應位址處的該資料項目的一快照；

透過該寫入記錄器組件來識別對該儲存媒體的該第一位址處的一寫入操作的一發生，該寫入操作亦不對該快取裝置的該對應位址處來實施；

在該作業系統的一重啟動之後，至少部分地藉由透過該寫入記錄器組件來決定是否在該資料項目被寫入該儲存媒體的該第一位址處及該快取裝置的該對應位址處之後進而修改過寫入於該儲存媒體的該第一位址處的該資料項目之方式，來決定是否寫入於該快取裝置的該對應位址處的該資料項目能被

可靠地用於滿足將該資料項目從該儲存媒體的該第一位址處讀取出的一要求；以及

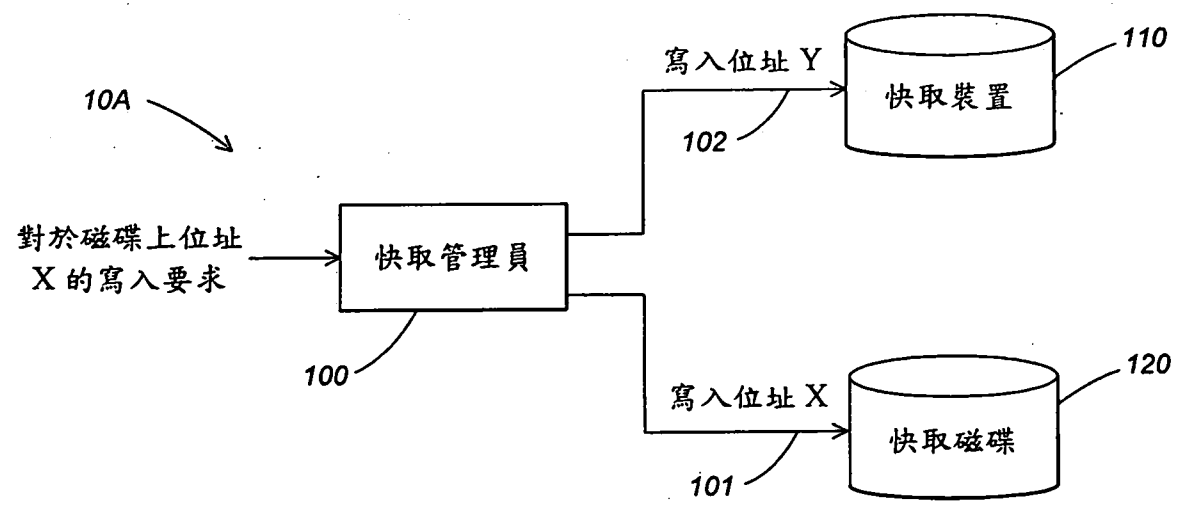
就在決定出在該快取裝置上的該對應位址處的該資料項目是不能被可靠地用於滿足將該資料項目從該儲存媒體的該第一位址處讀取出的該要求時，將導致使用來自該寫入記錄器組件的資料及該快照來更新該快取裝置。

15. 如申請專利範圍第 14 項所述之電腦，其中該至少一個處理器經程式編輯以下步驟：若決定出儲存於該快取裝置上的該對應位址處的該資料項目是不能被可靠地用於滿足將該資料項目從該儲存媒體的該第一位址處讀取出的該讀取要求，則 (i) 從該快取裝置的該對應位址處驅逐該資料項目，且 (ii) 藉由從該儲存媒體的該第一位址處讀取該資料項目來滿足該要求。
16. 如申請專利範圍第 14 項所述之電腦，其中當決定出該作業系統是正在關機時，該作業系統係經配置以啟動該寫入記錄器組件。
17. 如申請專利範圍第 14 項所述之電腦，其中該至少一個處理器經程式編輯以下步驟：避免對該儲存媒體上的該第一位址處的一寫入操作的一發生，該寫入操作並不對該快取裝置的該對應位址處來實施。
18. 如申請專利範圍第 17 項所述之電腦，其中該至少一

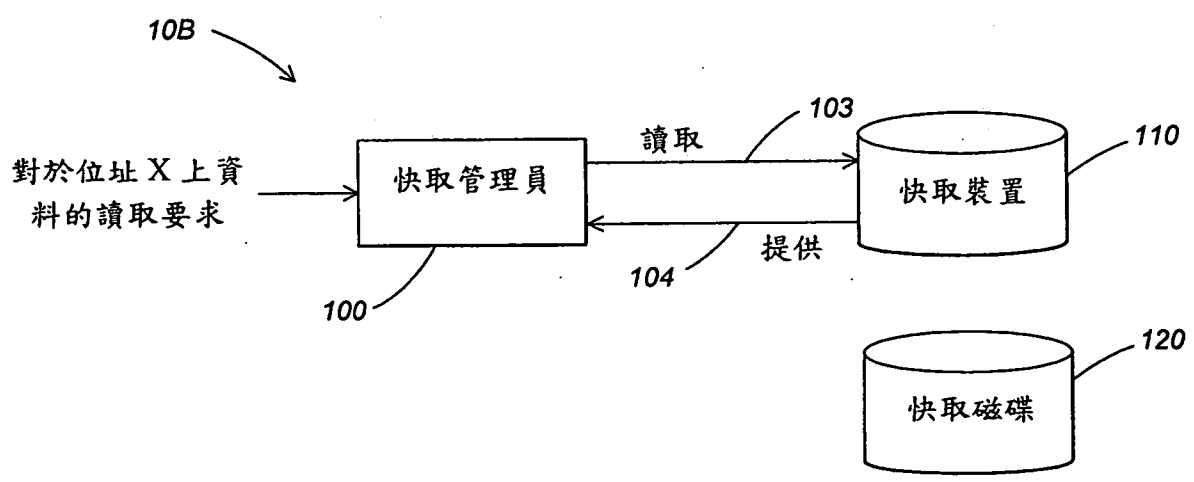
個處理器經程式編輯來避免的該寫入操作是針對在該儲存媒體上的一磁碟區(volume)，並且其中該至少一個處理器經程式編輯係藉由使該磁碟區成為不可存取的，來避免對該儲存媒體的該第一位址處的該寫入操作的該發生。

19. 如申請專利範圍第 14 項所述之電腦，其中該至少一個處理器程式經編輯來識別的該寫入操作是針對在該儲存媒體上的一檔案系統，且其中該至少一個處理器經程式編輯係藉由(i)在該作業系統的該重啟動之前，讓該檔案系統所使用的一記錄進入一特定狀態，且(ii)在該作業系統的該重啟動之後，決定是否該記錄是在該特定狀態內，來識別對該儲存媒體的一位址處的一寫入操作的一發生。

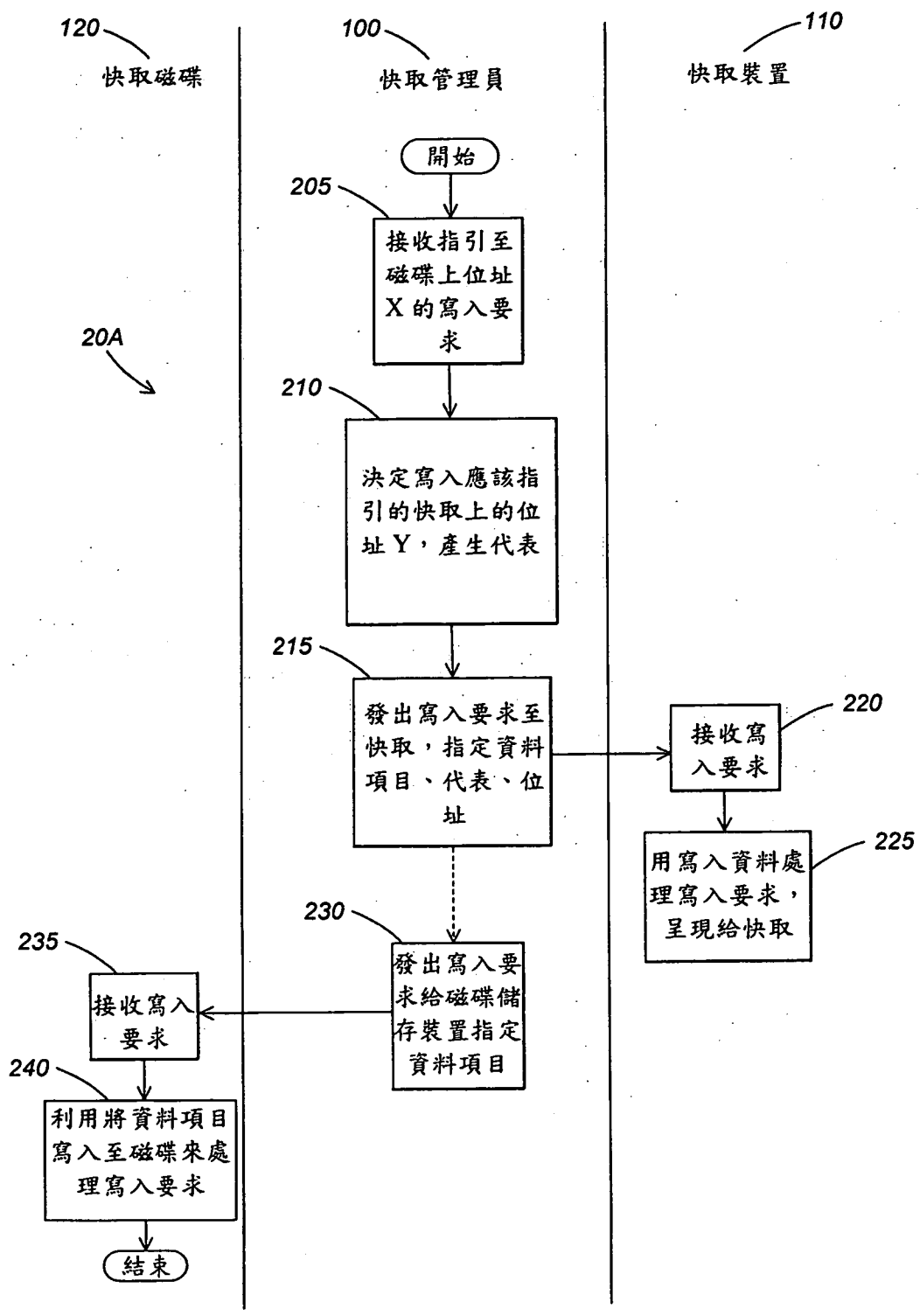
八、圖式：



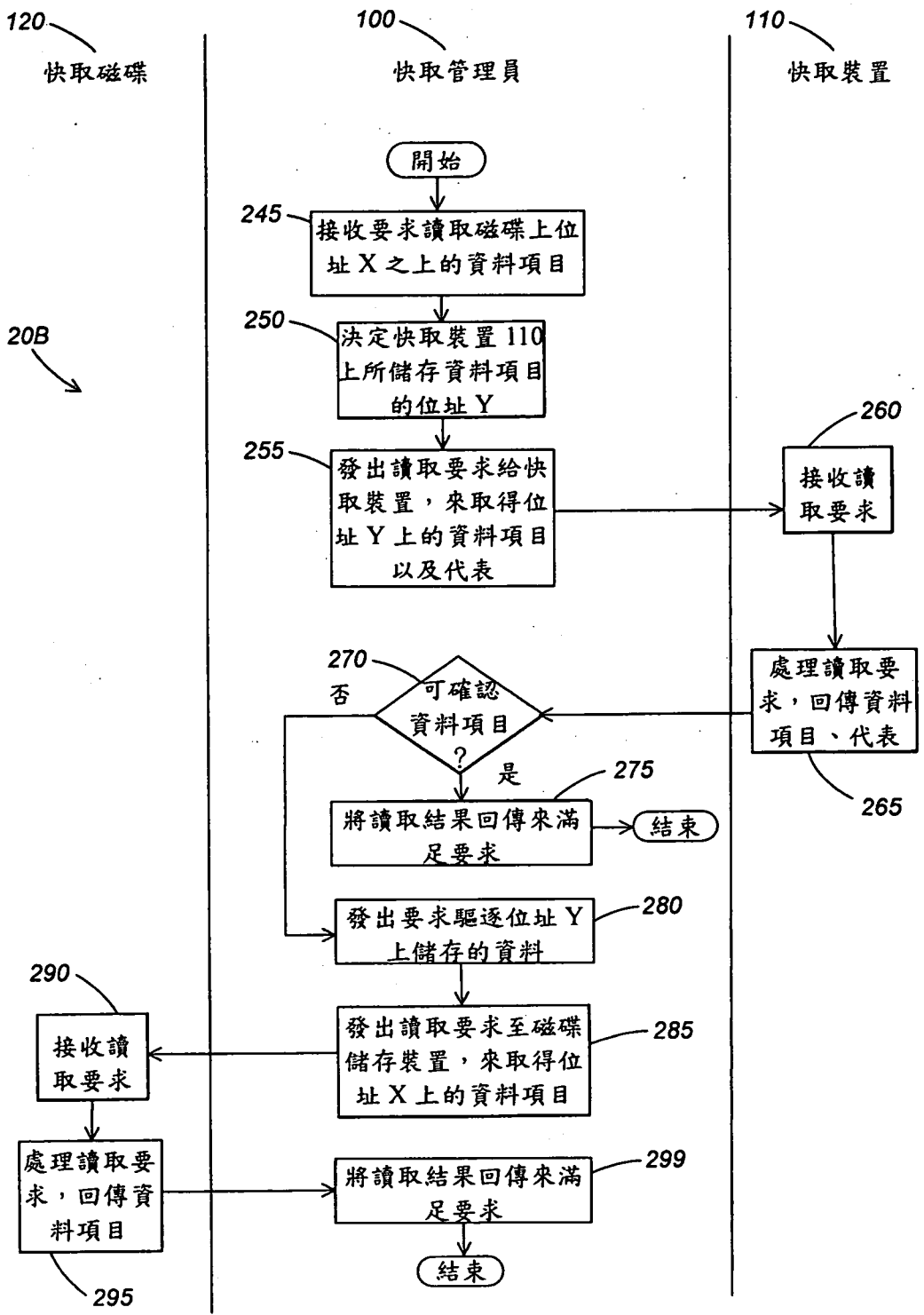
第 1A 圖



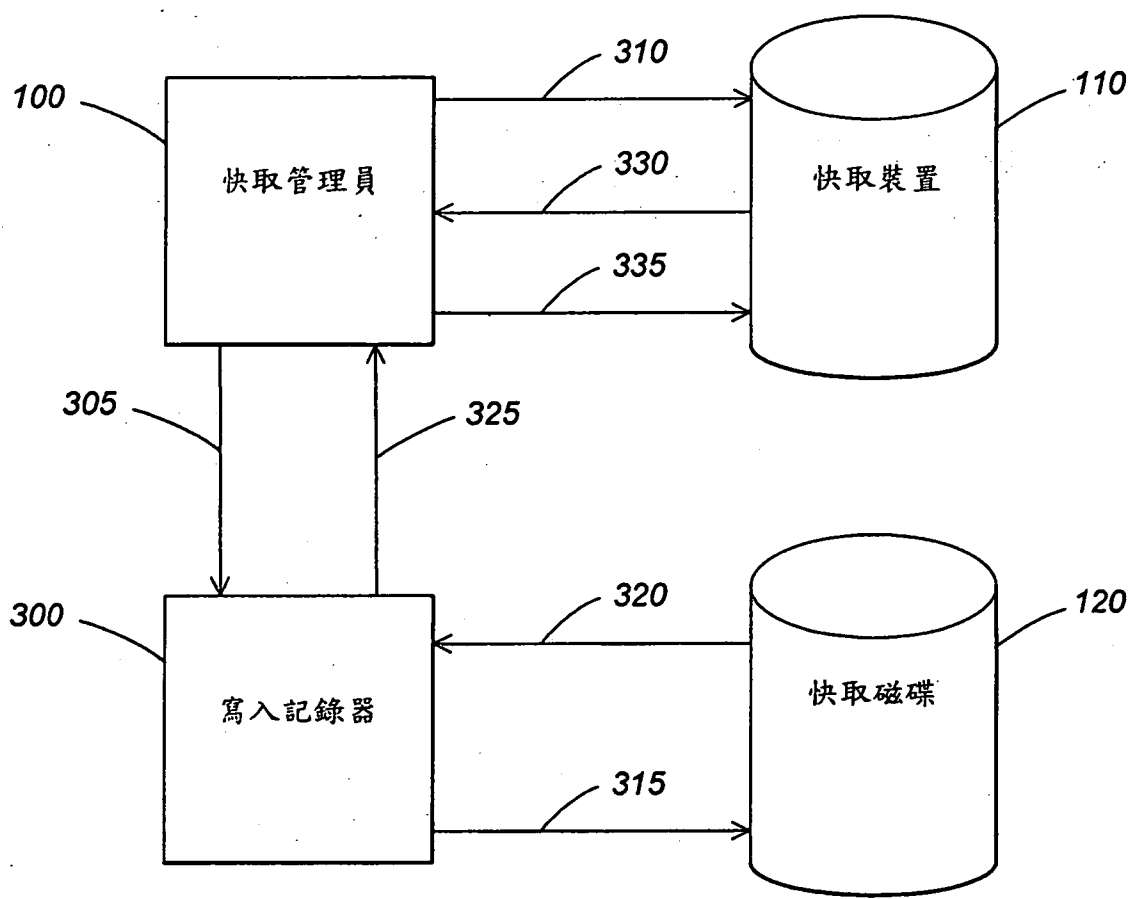
第 1B 圖



第 2A 圖

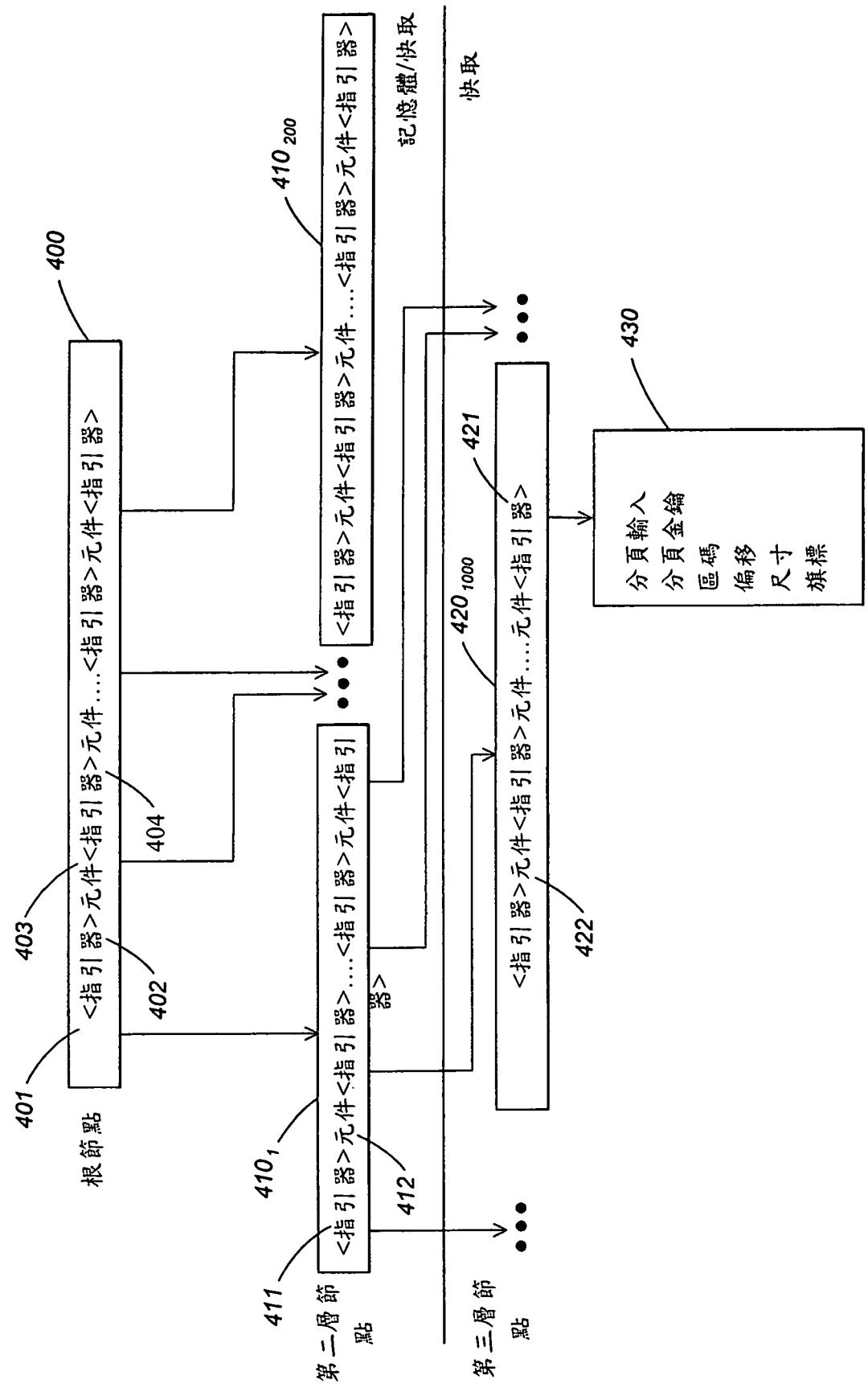


第 2B 圖

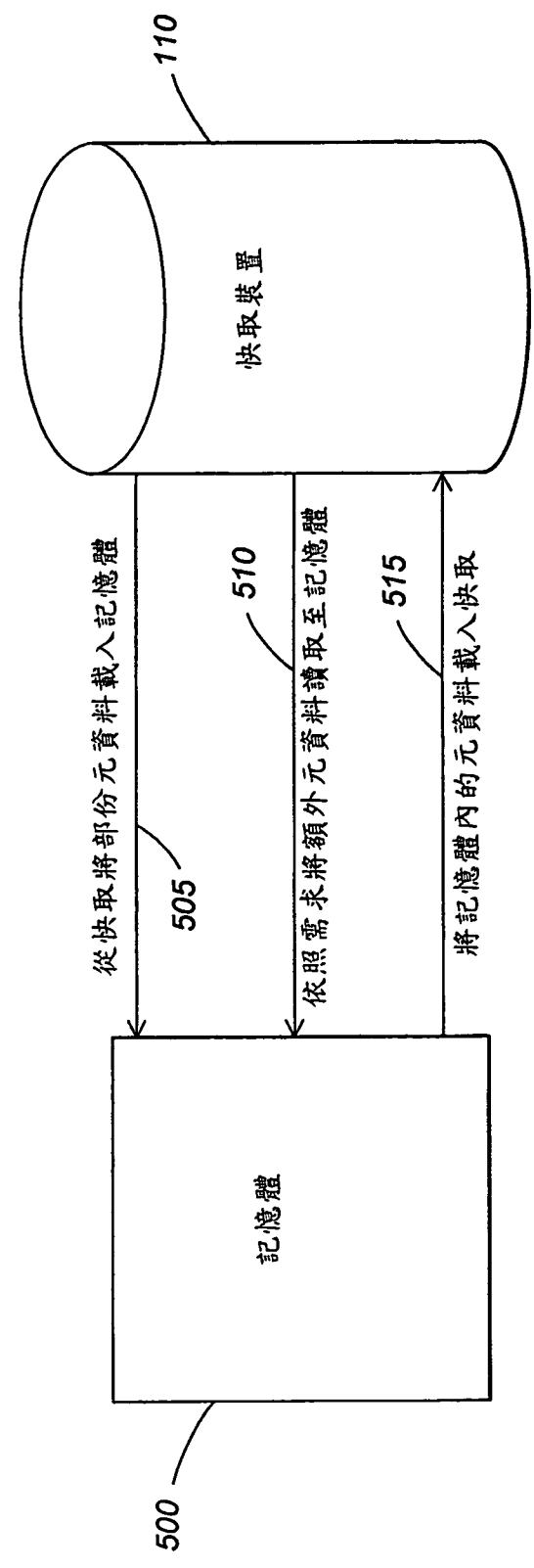


第 3 圖

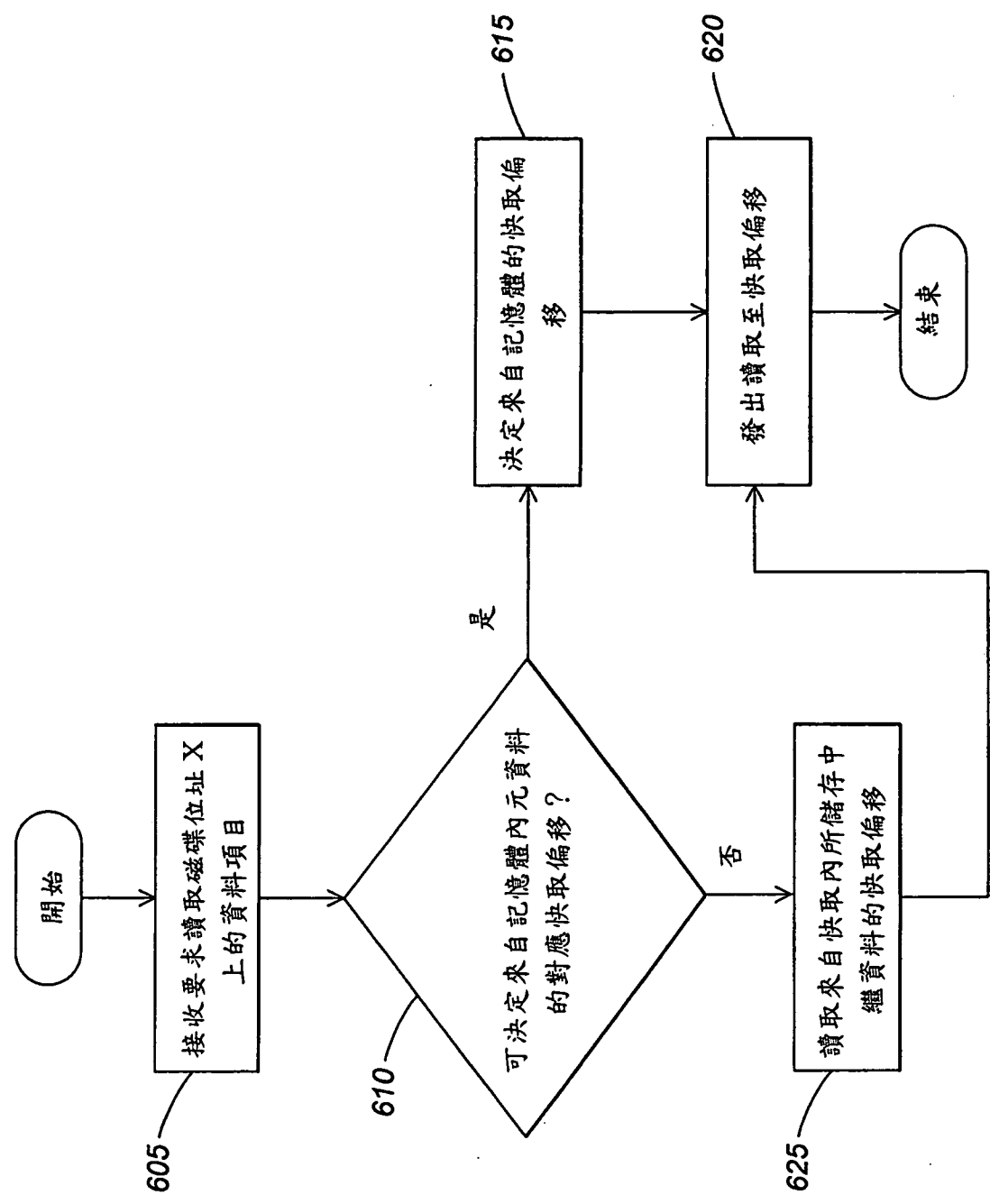
104年4月28日修正頁



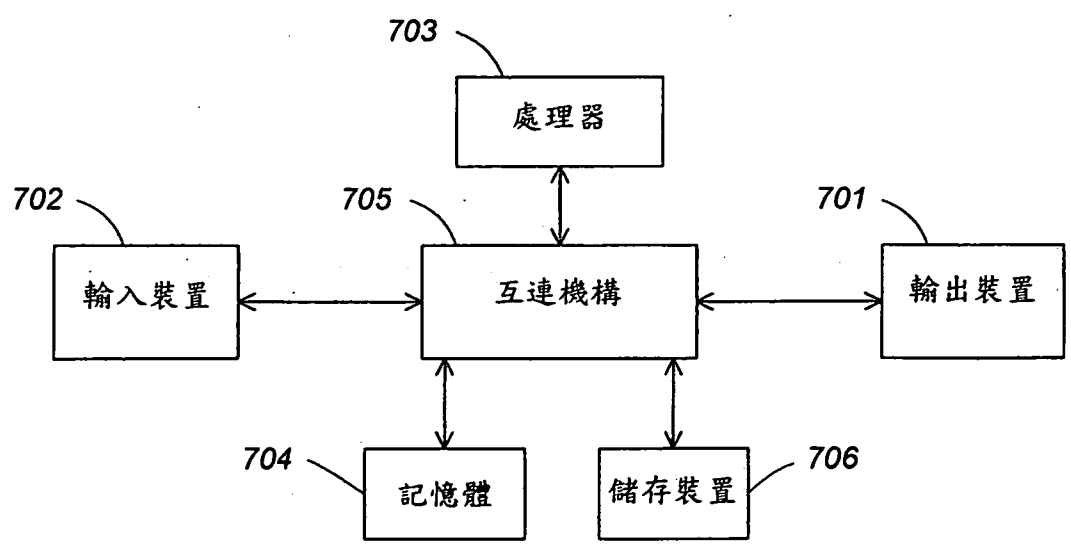
第 4 圖



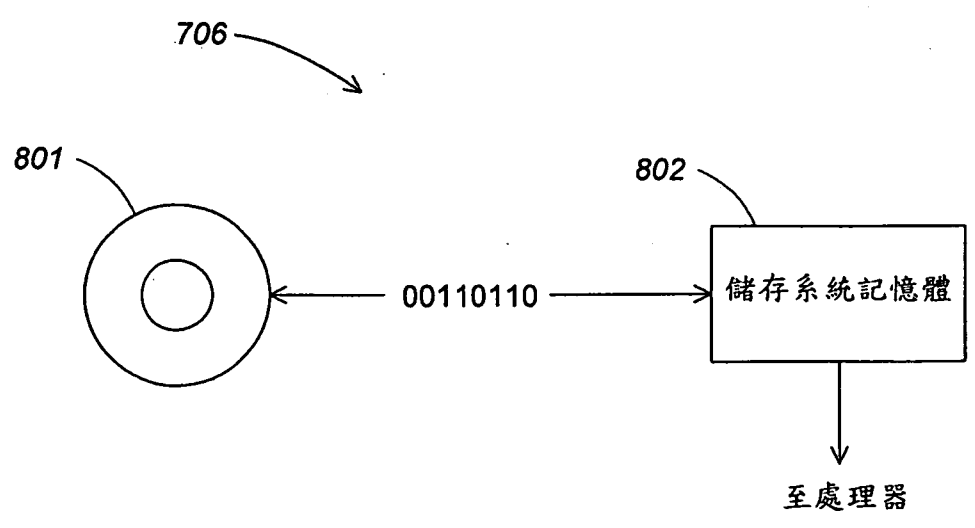
第 5 圖



第 6 圖



第 7 圖



第 8 圖