201917580

(19)中華民國智慧財產局

(12)發明說明書公開本 (11)公開編號：TW 201917580 A

(43)公開日： 中華民國 108 (2019) 年 05 月 01 日

(21)申請案號：107127263 (22)申請日： 中華民國 107 (2018) 年 08 月 06 日

(51)Int. Cl.： *G06F12/02 (2006.01)*

(30)優先權：2017/10/30 日本 2017-209344

(71)申請人：日商東芝記憶體股份有限公司 (日本) TOSHIBA MEMORY CORPORATION (JP)
日本

(72)發明人：管野伸一 KANNO, SHINICHI (JP)

(74)代理人：陳長文

申請實體審查：有 申請專利範圍項數：16 項 圖式數：49 共 109 頁

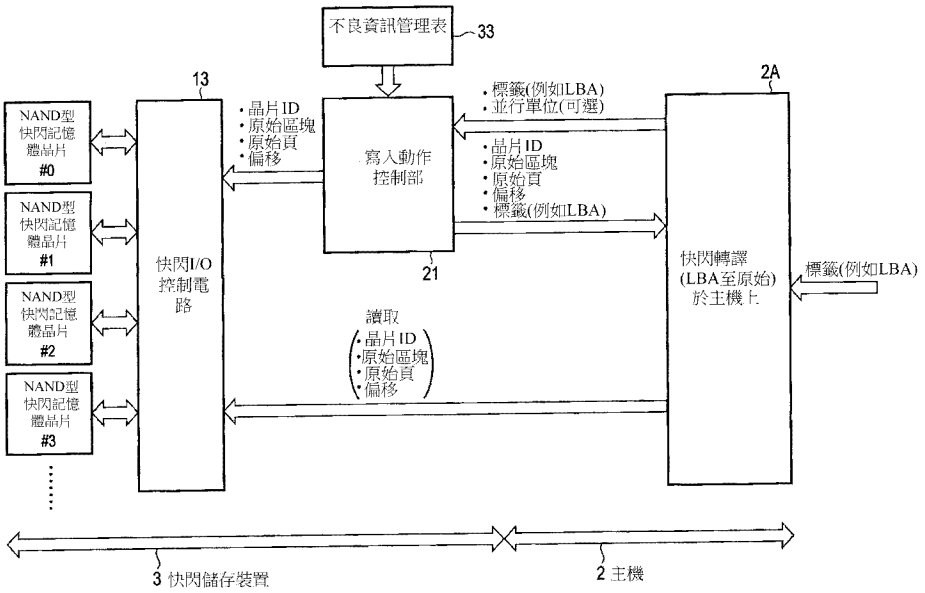
(54)名稱

計算機系統及儲存裝置之控制方法

(57)摘要

實施形態實現一種能夠改善 I/O 性能之計算機系統及儲存裝置之控制方法。實施形態之計算機系統將指定用以識別應寫入之第 1 資料之第 1 邏輯位址與上述第 1 資料之長度的寫入請求發送至儲存裝置。計算機系統係自上述儲存裝置接收第 1 實體位址、及上述第 1 邏輯位址，上述第 1 實體位址表示由上述儲存裝置自除不良區塊以外之區塊中選擇為用於上述第 1 資料之寫入目標區塊之第 1 區塊、及寫入了上述第 1 資料之上述第 1 區塊內之第 1 實體記憶位置之兩者。計算機系統於管理各邏輯位址與上述儲存裝置之各實體位址之間之映射之第 1 表中將上述第 1 實體位址映射至上述第 1 邏輯位址。

指定代表圖：



【圖20】

符號簡單說明：

2・・・主機

2A・・・快閃轉譯部

3・・・快閃儲存裝置

13・・・快閃 I/O 控制電路

21・・・寫入動作控制部

33・・・不良資訊管理表

【發明說明書】

【中文發明名稱】

計算機系統及儲存裝置之控制方法

【技術領域】

【0001】本發明之實施形態係關於一種計算機系統及儲存裝置(storage device)之控制方法。

【先前技術】

【0002】近年來，具備非揮發性記憶體之儲存裝置正廣泛地普及。作為此種儲存裝置之一，已知有基於NAND(Not And，反及)快閃技術之固態硬碟(SSD)。於資料中心之伺服器之類之計算機系統中利用之儲存裝置中，要求較高之I/O(Input/Output，輸入/輸出)性能。因此，最近，開始提出主機側與儲存裝置側之間之新介面。

【0003】但是，若非揮發性記憶體中包含之不良區塊之數量增加，則存在如下情形，即，於儲存裝置側，引起用以將不良區塊置換為其他區塊之置換資訊之量之增大、因該置換處理所致之讀取延遲時間之增加。該情況有時會成為使系統整體之I/O性能降低之因素。

【發明內容】

【0004】本發明之實施形態提供一種能夠改善I/O性能之計算機系統及控制方法。

【0005】根據實施形態，對包括分別包含複數個區塊之複數個非揮發性記憶體晶片(nonvolatile memory die)與控制器之儲存裝置進行控制的計算機系統具備：記憶體；及處理器，其構成為與上述記憶體電性連接，且執行上述記憶體中儲存之電腦程式。上述處理器將指定用以識別應

寫入之第1資料之第1邏輯位址與上述第1資料之長度的寫入請求發送至上述儲存裝置。上述處理器係自上述儲存裝置接收第1實體位址、及上述第1邏輯位址，上述第1實體位址表示由上述儲存裝置自除不良區塊以外之區塊中選擇為用於上述第1資料之寫入目標區塊之第1區塊及寫入了上述第1資料之上述第1區塊內之第1實體記憶位置之兩者。上述處理器係於管理各邏輯位址與上述儲存裝置之各實體位址之間之映射之第1表中將上述第1實體位址映射至上述第1邏輯位址。

【圖式簡單說明】

【0006】

圖1係表示主機與記憶體系統(快閃儲存裝置)之關係之方塊圖。

圖2係用以說明快閃儲存裝置與主機之間之功能分擔之圖。

圖3係表示快閃儲存裝置之構成例之方塊圖。

圖4係表示In-Drive-GC API用之指令之圖。

圖5係表示超級區塊API用之指令之圖。

圖6係表示設置於快閃儲存裝置之快閃I/O控制電路與複數個NAND型快閃記憶體晶片之關係的方塊圖。

圖7係表示由複數個區塊之集合構築之超級區塊(並行單位)之構成例之圖。

圖8係用以說明超級區塊之區塊位址與構成該超級區塊之複數個區塊各自之區塊位址之關係的圖。

圖9係用以說明將屬於某非揮發性記憶體晶片之各個不良區塊分別利用該非揮發性記憶體晶片之其他區塊置換之動作的圖。

圖10係用以說明不置換不良區塊而自超級區塊內之非不良區塊中選

擇寫入目標區塊之動作的圖。

圖11係用以說明對快閃儲存裝置應用之寫入指令之圖。

圖12係用以說明對圖11之寫入指令之應答之圖。

圖13係用以說明對快閃儲存裝置應用之Trim指令之圖。

圖14係表示對具有不良區塊之超級區塊寫入資料之動作之圖。

圖15係表示圖12之應答中包含之實體位址之構成例之圖。

圖16係表示超級區塊之區塊位址與該超級區塊中包含之各區塊之區塊位址之關係的圖。

圖17係用以說明將邏輯位址與資料之對寫入至區塊內之頁之動作的圖。

圖18係用以說明將資料寫入至區塊內之頁之用戶資料區域並將該資料之邏輯位址寫入至該頁之冗餘區域之動作的圖。

圖19係用以說明將複數個資料部與根據該等資料部算出之消失校正碼寫入至具有不良區塊之超級區塊之動作的圖。

圖20係表示主機內之快閃轉譯部與快閃儲存裝置內之寫入動作控制部之關係的方塊圖。

圖21係用以說明由主機與快閃儲存裝置執行之寫入動作及讀取動作的方塊圖。

圖22係表示由主機與快閃儲存裝置執行之寫入動作處理之序列之序列圖。

圖23係表示寫入對於已寫入之資料之更新資料之資料更新動作之圖。

圖24係用以說明更新由快閃儲存裝置管理之區塊管理表之動作之

圖。

圖25係用以說明更新由主機管理之查找表(邏輯實體位址轉換表)之動作之圖。

圖26係用以說明根據表示與應無效化之資料對應之實體位址之來自主機之通知更新區塊管理表之動作的圖。

圖27係用以說明對快閃儲存裝置應用之讀取指令之圖。

圖28係用以說明由快閃儲存裝置執行之讀取動作之圖。

圖29係表示由主機與快閃儲存裝置執行之讀取處理之序列之序列圖。

圖30係用以說明對快閃儲存裝置應用之垃圾回收(GC)控制指令之圖。

圖31係用以說明對快閃儲存裝置應用之強制垃圾回收(GC)指令之圖。

圖32係用以說明自快閃儲存裝置發送至主機之位址更新通知之圖。

圖33係表示由快閃儲存裝置執行之垃圾回收(GC)動作之順序之序列圖。

圖34係用以說明為了垃圾回收(GC)所執行之資料複製動作之例的圖。

圖35係用以說明基於圖34之資料複製動作之結果更新之主機之查找表之內容的圖。

圖36係用以說明對寫入指令之應答與GC用回調處理(位址更新通知)之關係之圖。

圖37係表示由主機執行之查找表更新處理之順序之流程圖。

圖38係表示用以管理參照計數之區塊管理表之構成例之圖。

圖39係用以說明對快閃儲存裝置應用之重複指令之圖。

圖40係表示由主機與快閃儲存裝置執行之參照計數遞增/遞減處理之序列圖。

圖41係表示由快閃儲存裝置執行之超級區塊分配處理之順序之流程圖。

圖42係用以說明以超級區塊內之所有非不良區塊自該超級區塊之開頭連續地邏輯性地配置之方式轉換存取對象區塊之位址的位址轉換動作之圖。

圖43係用以說明用於位址轉換動作之不良資訊管理表與位址轉換之例的圖。

圖44係表示主機內之快閃轉譯部與快閃儲存裝置內之不良區塊轉譯部之關係的方塊圖。

圖45係表示主機(計算機系統)之構成例之方塊圖。

圖46係表示內置快閃儲存裝置之主機之構成例之圖。

圖47係表示由主機執行之寫入動作之順序之流程圖。

圖48係表示由主機執行之讀取動作之順序之流程圖。

圖49係表示由主機執行之參照計數遞增/遞減處理之順序之流程圖。

【實施方式】

【0007】 以下，參照圖式，對實施形態進行說明。

【0008】 首先，參照圖1，對主機與記憶體系統之關係進行說明。

【0009】 該記憶體系統係以對非揮發性記憶體寫入資料並自非揮發性記憶體讀出資料之方式構成的半導體儲存裝置。該記憶體系統係以基於

NAND快閃技術之快閃儲存裝置3之形式實現。

【0010】 主機(主機裝置)2構成為對複數個快閃儲存裝置3進行控制。主機2係藉由構成為將由複數個快閃儲存裝置3構成之快閃陣列用作儲存器之計算機系統而實現。該計算機系統亦可為伺服器。

【0011】 再者，快閃儲存裝置3亦可用作設置於儲存陣列內之複數個儲存裝置之一個。儲存陣列亦可經由纜線或網路而連接於伺服器之類之計算機系統，儲存陣列包含控制該儲存陣列內之複數個儲存器(例如複數個快閃儲存裝置3)之控制器。於將快閃儲存裝置3應用於儲存陣列之情形時，該儲存陣列之控制器亦可作為快閃儲存裝置3之主機發揮功能。

【0012】 以下，例示伺服器之類之計算機系統作為主機2發揮功能之情形進行說明。

【0013】 主機(伺服器)2與複數個快閃儲存裝置3係經由介面50而相互連接(內部相互連接)。作為該用於內部相互連接之介面50，並不限定於此，可使用PCI Express(Peripheral Component Interconnect Express，快速周邊元件互連)(PCIe)(註冊商標)、NVM Express(Non-Volatile Memory Express，非揮發性記憶體標準)(NVMe)(註冊商標)、乙太網路(Ethernet，註冊商標)、NVMe over Fabrics(NVMeOF，非揮發性記憶體標準織物)等。

【0014】 作為以主機2之形式發揮功能之伺服器之典型例，可列舉資料中心內之伺服器。

【0015】 於主機2以資料中心內之伺服器之形式實現之實例中，該主機(伺服器)2亦可經由網路51而連接於複數個終端用戶終端(用戶端)61。主機2可對該等終端用戶終端61提供各種服務。

【0016】 主機(伺服器)2所能提供之服務例如有：(1)向各用戶端(各終端用戶終端61)提供系統運轉平台之平台即服務(PaaS，platform as a service)；(2)向各用戶端(各終端用戶終端61)提供虛擬伺服器之類之基礎架構之基礎架構即服務(IaaS，Infrastructure as a Service)等。

【0017】 複數個虛擬機亦可於作為該主機(伺服器)2發揮功能之實體伺服器上執行。於主機(伺服器)2上運行之該等虛擬機之各者可作為構成為向對應之若干個用戶端(終端用戶終端61)提供各種服務之虛擬伺服器發揮功能。

【0018】 主機(伺服器)2包含管理構成快閃陣列之複數個快閃儲存裝置3之儲存管理功能、及對各終端用戶終端61提供包含儲存存取之各種服務之前端功能。

【0019】 於先前型SSD中，NAND型快閃記憶體之區塊/頁之層級結構被SSD內之快閃轉譯層(FTL)隱蔽。即，先前型SSD之FTL具有如下等功能：(1)使用作為邏輯實體位址轉換表發揮功能之查找表，管理各邏輯位址與NAND型快閃記憶體之各實體位址之間之映射；(2)用以將以頁為單位之讀取/寫入及以區塊為單位之刪除動作隱蔽；(3)執行NAND型快閃記憶體之垃圾回收(GC)。各邏輯位址與NAND型快閃記憶體之實體位址之間之映射無法自主機看到。NAND型快閃記憶體之區塊/頁結構亦無法自主機看到。

【0020】 另一方面，於主機中，有時亦執行一種位址轉換(應用階層位址轉換)。該位址轉換係使用應用階層位址轉換表，管理應用階層用之各邏輯位址與SSD用之各邏輯位址之間之映射。又，於主機中，亦存在如下情形，即，為了消除SSD用之邏輯位址空間上產生之碎片，而執行用以

變更該邏輯位址空間上之資料配置之一種GC(應用階層GC)。

【0021】 但是，於主機及SSD分別具有位址轉換表之冗餘構成(SSD具有作為邏輯實體位址轉換表發揮功能之查找表，主機具有應用階層位址轉換表)中，為了保存該等位址轉換表而消耗大量之記憶體資源。進而，包含主機側之位址轉換與SSD側之位址轉換之雙重之位址轉換亦成為使系統整體之I/O性能降低之因素。

【0022】 進而，主機側之應用階層GC成為使對SSD之資料寫入量增加至實際之用戶資料量之數倍(例如2倍)左右之因素。此種資料寫入量之增加雖不會增加SSD之寫入放大，但會使系統整體之儲存性能降低，且SSD之壽命亦變短。

【0023】 因此，於本實施形態中，使FTL之功能於主機2與快閃儲存裝置3之間被分擔。主機2管理作為邏輯實體位址轉換表發揮功能之查找表，但應使用於寫入之區塊及頁可由快閃儲存裝置3決定而並非由主機2決定。又，GC亦可由快閃儲存裝置3執行而並非由主機2執行。以下，將轉移至主機2之FTL功能稱為全域FTL。

【0024】 又，快閃儲存裝置3管理分別包括複數個區塊(複數個實體區塊)之複數個並行單位(複數個超級區塊)，以提高寫入/讀出速度。快閃儲存裝置3可並行地執行對於某並行單位內之複數個區塊之寫入動作及讀出動作。

【0025】 但是，由於NAND型快閃記憶體晶片中包含若干個不良區塊，故存在將並行單位內之各不良區塊以其他區塊置換，以確保並行度之情形。但是，若每一個NAND型快閃記憶體晶片中包含之不良區塊之數量增加，則存在用於置換管理之資訊量增大之可能性。

【0026】 若以一個並行單位中包含之區塊之數量為64之情形為例，則於根據數學規則自表示並行單位之區塊位址產生各區塊之區塊位址之實例中，將64個區塊中之第15個區塊以區塊位址2049之區塊置換之情形時，為了表達第15個需要至少6位元，為了表達2049需要12位元。於應置換之區塊數較多之情形時，需要與其成比例之資訊量。於將應置換之區塊之數量設為至多16個之情形時，針對每一並行單位而需要18位元×16＝294位元之資訊。

【0027】 又，於資料讀取動作中，必須利用置換資訊，執行用以將表示不良區塊之位址轉換為表示置換目標區塊之位址之位址轉換。因此，若不良區塊之數量增加，則伴隨置換資訊增大而該位址轉換處理所需之時間增加，因此讀取潛伏時間(read latency)增大。

【0028】 因此，於用以寫入來自主機2之資料之寫入動作中，快閃儲存裝置3避開寫入對象之並行單位內之不良區塊而自該並行單位內之非不良區塊中選擇寫入目標區塊，並決定該寫入目標區塊內之寫入目標位置。來自主機2之資料被寫入至該寫入目標位置。而且，快閃儲存裝置3將表示該寫入目標區塊及該寫入目標區塊內之寫入位置之兩者之實體位址通知給主機2。藉此，主機2可辨識實際寫入了資料之區塊與該區塊內之寫入目標位置，因此，於必須讀取該資料之情形時，可將指定該實體位址之讀取請求發送至快閃儲存裝置3。快閃儲存裝置3可基於該讀取請求所指定之實體位址，自上述所選擇之區塊內之寫入目標位置讀取資料。因此，於快閃儲存裝置3，無須進行位址轉換處理，可減少讀取潛伏時間。

【0029】 主機2之全域FTL亦可具有執行儲存服務之功能、管理查找表(LUT)之功能、軟體硬體控制功能、用以實現高可用性之功能、防止具

有相同內容之複數個重複資料部儲存於儲存器之重複資料刪除(De-duplication)功能等。

【0030】 另一方面，快閃儲存裝置3可執行底層抽象(LLA，low-level abstraction)。LLA係用於NAND型快閃記憶體之抽象之功能。LLA包含輔助資料配置之功能等。輔助資料配置之功能中包含決定用戶資料之寫入目標位置(區塊位址、該區塊內之位置)之功能、將表示寫入了用戶資料之該寫入目標位置之實體位址通知給上位層級(主機2)之功能、決定用於垃圾回收之複製源區塊與複製目標區塊之功能、將有效資料之複製目標位置通知給上位層級(主機2)之功能等。又，LLA亦具有針對每一個域(QoS(Quality of Service，服務質量)域)執行快閃儲存裝置3之資源管理之QoS控制功能。

【0031】 QoS控制功能中包含針對每一個QoS域決定存取單位(資料粒度)之功能。存取單位表示主機2可進行寫入/讀取之最小資料尺寸(資料粒度)。快閃儲存裝置3支持單個或複數個存取單位(資料粒度)，主機2係於快閃儲存裝置3支持複數個存取單位之情形時，可針對每一個QoS域對快閃儲存裝置3指示應使用之存取單位(資料粒度)。

【0032】 又，QoS控制功能中包含用以儘可能防止QoS域間之性能干涉之功能。該功能係用以基本上維持穩定之潛伏時間之功能。

【0033】 為了實現該目的，快閃儲存裝置3亦可將NAND型快閃記憶體內邏輯性地分割為複數個區域(複數個QoS域)。一個區域(一個QoS域)包含一個以上之並行單位(超級區塊)。各並行單位(超級區塊)僅屬於某一個區域(QoS域)。

【0034】 圖2表示包含快閃儲存裝置3與主機2之系統整體之層級結

構。

【0035】 於主機(伺服器)2中，執行用以對複數個終端用戶提供複數個虛擬機之虛擬機服務401。於虛擬機服務401上之各虛擬機中，執行對應之終端用戶所使用之作業系統及用戶應用402。

【0036】 又，於主機(伺服器)2中，執行與複數個用戶應用402對應之複數個I/O服務403。該等I/O服務403中亦可包含基於LBA(logical block address，邏輯區塊位址)之區塊I/O服務、鍵值儲存服務等。各I/O服務403包含管理各邏輯位址與快閃儲存裝置3之各實體位址之間之映射之查找表(LUT)。此處，所謂邏輯位址係指可識別存取對象之資料之識別碼(標籤)。該邏輯位址可為指定邏輯位址空間上之位置之邏輯區塊位址(LBA)，亦可為鍵值儲存之鍵，還可為檔案名之類之檔案識別碼。

【0037】 於基於LBA之區塊I/O服務中，亦可使用管理各邏輯位址(LBA)與快閃儲存裝置3之各實體位址之間之映射之LUT。

【0038】 於鍵值儲存服務中，亦可使用對各邏輯位址(鍵之類之標籤)與儲存有該等邏輯位址(即，鍵之類之標籤)所對應之資料之快閃儲存裝置3之各實體位址之間之映射進行管理的LUT。於該LUT中，亦可管理標籤、儲存有由該標籤識別之資料之實體位址、及該資料之資料長度之對應關係。

【0039】 各終端用戶可選擇應使用之定址方法(LBA、鍵值儲存之鍵、檔案識別碼等)。

【0040】 該等各LUT並非將來自用戶應用402之各邏輯位址轉換為快閃儲存裝置3用之各邏輯位址，而是將來自用戶應用402之各邏輯位址轉換為快閃儲存裝置3之各實體位址。即，該等各LUT係將快閃儲存裝置3

用之邏輯位址轉換為實體位址之表與應用階層位址轉換表整合(合併)所得之表。

【0041】 於主機(伺服器)2中，針對上述QoS域中之每一個而存在I/O服務403。屬於某QoS域之I/O服務403管理對應之QoS域內之用戶應用402所使用之各邏輯位址與分配給對應之QoS域之區域之各實體位址之間之映射。

【0042】 自主機(伺服器)2向快閃儲存裝置3之指令之發送及自快閃儲存裝置3向主機(伺服器)2之指令完成之應答等回傳係經由主機(伺服器)2及快閃儲存裝置3之各者中所存在之I/O佇列500而執行。該等I/O佇列500亦可分類為與複數個QoS域對應之複數個佇列組。

【0043】 快閃儲存裝置3包含與複數個QoS域對應之複數個寫入緩衝器(WB)601、與複數個QoS域對應之複數個垃圾回收(GC)功能602、NAND型快閃記憶體(NAND快閃陣列)603。

【0044】 圖3表示快閃儲存裝置3之構成例。

【0045】 快閃儲存裝置3具備控制器4及非揮發性記憶體(NAND型快閃記憶體)5。快閃儲存裝置3亦可同時具備隨機存取記憶體、例如DRAM(Dynamic Random Access Memory，動態隨機存取記憶體)6。

【0046】 NAND型快閃記憶體5包含記憶胞陣列，該記憶胞陣列包含呈矩陣狀配置之複數個記憶胞。NAND型快閃記憶體5可為二維結構之NAND型快閃記憶體，亦可為三維結構之NAND型快閃記憶體。

【0047】 NAND型快閃記憶體5之記憶胞陣列包含複數個區塊BLK0～BLKm－1。區塊BLK0～BLKm－1分別由多個頁(此處為頁P0～Pn－1)組成。區塊BLK0～BLKm－1作為刪除單位發揮功能。區塊有時亦被稱為

「刪除區塊」、「實體區塊」或「實體刪除區塊」。頁P0～Pn－1分別包含連接於同一字元線之複數個記憶胞。頁P0～Pn－1係資料寫入動作及資料讀入動作之單位。

【0048】 控制器4係經由Toggle、開放式NAND快閃介面(ONFI，Open NAND Flash Interface)之類之快閃I/O控制電路13而電性連接於作為非揮發性記憶體之NAND型快閃記憶體5。控制器4係構成為控制NAND型快閃記憶體5之記憶體控制器(控制電路)。

【0049】 NAND型快閃記憶體5包含複數個NAND型快閃記憶體晶片。控制器4管理上述複數個並行單位(並行單元)。各並行單位係以包含分別屬於不同之NAND型快閃記憶體晶片之複數個區塊(複數個實體區塊)之區塊組即超級區塊的形式實現。控制器4可並行地執行對於各並行單位(超級區塊)中包含之複數個區塊之寫入動作、讀出動作。複數個並行單位(超級區塊)分別具有固有之超級區塊位址(第1區塊位址)。各非揮發性記憶體晶片內之複數個區塊分別具有固有之區塊位址(第2區塊位址)。應包含於各並行單位(超級區塊)之各區塊之區塊位址係自各並行單位之超級區塊基於數學規則確定。

【0050】 控制器4包含主機介面11、CPU(Central Processing Unit，中央處理單元)12、快閃I/O控制電路13、及DRAM介面14等。該等CPU12、快閃I/O控制電路13、DRAM介面14係經由匯流排10而相互連接。

【0051】 該主機介面11係以執行與主機2之通信之方式構成之主機介面電路。該主機介面11例如可為PCIe控制器(NVMe控制器)。主機介面11自主機2接收各種請求(指令)。該等請求(指令)中包含寫入請求(寫入指

令)、讀取請求(讀取指令)、其他各種請求(指令)。

【0052】 CPU12係以控制主機介面11、快閃I/O控制電路13、DRAM介面14之方式構成之處理器。CPU12係響應快閃儲存裝置3之電源接通，自NAND型快閃記憶體5或未圖示之ROM(Read Only Memory，唯讀記憶體)將控制程式(韌體)載入至DRAM6，然後執行該韌體，藉此進行各種處理。再者，韌體亦可載入至控制器4內之未圖示之SRAM(Static Random-Access Memory，靜態隨機存取記憶體)上。該CPU12可執行用以處理來自主機2之各種指令之指令處理等。CPU12之動作係由CPU12所執行之上述韌體控制。再者，指令處理之一部分或全部亦可由控制器4內之專用硬體執行。

【0053】 CPU12可作為寫入動作控制部21、讀取動作控制部22、及GC動作控制部23發揮功能。於該等寫入動作控制部21、讀取動作控制部22、及GC動作控制部23中，安裝有用以實現圖2所示之系統構成之應用程式介面(API)。

【0054】 寫入動作控制部21自主機2接收指定用以識別應寫入之資料(用戶資料)之邏輯位址(標籤)的寫入請求(寫入指令)。於接收到寫入指令之情形時，寫入動作控制部21首先參照不良資訊管理表33，自寫入對象之並行單位(超級區塊)中包含之並非不良之區塊(非不良區塊)中選擇一個區塊作為寫入目標區塊。不良資訊管理表33保存不良資訊，該不良資訊保存針對各並行單位(超級區塊)中包含之每一區塊表示可使用或不可使用之至少1位元之資訊。與各超級區塊對應之不良資訊亦可為包含與各超級區塊中包含之區塊之數量為相同數量之位元的點陣圖。寫入動作控制部21藉由參照與寫入對象之並行單位(超級區塊)對應之不良資訊(不良資訊管理

表33)，而可辨識寫入對象之並行單位內之各區塊為不良區塊或非不良區塊中之哪一個。

【0055】 再者，不良資訊並不限定於上述點陣圖，例如亦可代替使用表示各區塊之刪除次數(編程/擦除週期數)之資訊作為不良資訊。於該情形時，控制器4亦可將具有某閾值以上之刪除次數之區塊判定為不良區塊。

【0056】 不良區塊係不可使用之區塊，亦被作為「壞區塊」加以參照。由不良資訊表示之不良區塊(壞區塊)中可包含在NAND型快閃記憶體之製造過程中產生之不良區塊(初級(primary)不良區塊)，亦可包含在快閃儲存裝置3之使用開始後產生之不良區塊(次級(grown)不良區塊)，還可包含該等初級不良區塊與次級不良區塊之兩者。

【0057】 寫入動作控制部21決定所選擇之寫入目標區塊內之寫入目標位置(頁、及該頁內之位置)。繼而，寫入動作控制部21將來自主機2之資料(寫入資料)寫入至該寫入目標區塊之寫入目標位置。於該情形時，寫入動作控制部21不僅可將來自主機2之資料寫入至寫入目標區塊，而且可將該資料與該資料之邏輯位址(標籤)之兩者寫入至寫入目標區塊。

【0058】 然後，寫入動作控制部21將所指定之邏輯位址(標籤)與表示寫入目標區塊及寫入目標位置之兩者之實體位址返回給主機2。

【0059】 於該情形時，該實體位址亦可由晶片識別碼、實體區塊位址(第2區塊位址)、及偏移表示。晶片識別碼係各非揮發性記憶體晶片所固有之識別碼。該實體位址中包含之晶片識別碼表示寫入目標區塊所屬之非揮發性記憶體晶片之晶片識別碼。實體區塊位址係用以識別各非揮發性記憶體晶片內之複數個區塊各自之區塊位址(區塊編號)。該實體位址中包

含之區塊位址表示該寫入目標區塊之區塊位址。偏移係區塊內偏移。該實體位址中包含之偏移表示自寫入目標區塊之開頭至寫入目標位置之偏移(偏移值)。該偏移亦可由寫入目標位置所屬之頁之頁位址、及與寫入目標位置對應之頁內偏移表示。

【0060】 或者，該實體位址亦可由超級區塊位址(第1區塊位址)與偏移表示。超級區塊位址係各超級區塊所固有之區塊位址。該實體位址中包含之超級區塊位址表示寫入對象之超級區塊之超級區塊位址。偏移表示自寫入對象之超級區塊之開頭至寫入目標位置之偏移(偏移值)。該偏移亦可由寫入目標區塊所屬之非揮發性記憶體晶片之晶片識別碼、寫入目標位置所屬之頁之頁位址、及與寫入目標位置對應之頁內偏移表示。

【0061】 實體位址亦被作為「快閃位址」加以參照。

【0062】 再者，寫入指令亦可不僅指定邏輯位址(標籤)，而且指定超級區塊位址。於接收到指定超級區塊位址之寫入指令之情形時，寫入動作控制部21自複數個超級區塊中，選擇具有寫入指令所指定之超級區塊位址之超級區塊作為寫入對象之並行單位(寫入目標超級區塊)。

【0063】 讀取動作控制部22係於自主機2接收到指定實體位址(表示讀取對象之區塊及該區塊內之讀取對象之位置)之讀取請求(讀取指令)的情形時，基於該實體位址，自讀取對象之區塊內之讀取對象之位置讀取資料。

【0064】 GC動作控制部23係於執行NAND型快閃記憶體5之垃圾回收之情形時，自NAND型快閃記憶體5內之多個區塊中選擇用於該垃圾回收之複製源區塊(GC源區塊)及複製目標區塊(GC目的地區塊)。該等GC源區塊及GC目的地區塊之各者可為超級區塊，亦可為實體區塊。

【0065】 GC動作控制部23通常選擇複數個複製源區塊(GC源區塊)、及一個以上之複製目標區塊(GC目的地區塊)。用以選擇複製源區塊(GC源區塊)之條件(GC策略)亦可由主機2指定。例如，可使用優先選擇有效資料量最少之區塊作為複製源區塊(GC源區塊)之GC策略，亦可使用其他GC策略。如此，複製源區塊(GC源區塊)及複製目標區塊(GC目的地區塊)之選擇並非由主機2執行，而是由快閃儲存裝置3之控制器4(GC動作控制部23)執行。控制器4亦可使用各區塊管理表來管理各區塊之有效資料量。

【0066】 有效資料/無效資料之管理亦可使用區塊管理表32而執行。該區塊管理表32例如亦可針對每一超級區塊而存在。於與某超級區塊對應之區塊管理表32中，儲存有表示該超級區塊內之各區塊之各個資料之有效/無效之點陣圖旗標。此處，所謂有效資料係指被LUT參照之資料(即，被邏輯位址作為最新之資料關聯起來之資料)，且之後有可能被主機2讀取之資料。所謂無效資料係指已無可能被主機2讀取之資料。例如，與某邏輯位址建立關聯之資料係有效資料，未與任何邏輯位址建立關聯之資料係無效資料。

【0067】 GC動作控制部23決定應寫入複製源區塊(GC源區塊)內所儲存之有效資料的複製目標區塊(GC目的地區塊)內之位置(複製目標位置)，將有效資料複製至複製目標區塊(GC目的地區塊)之該決定之位置(複製目標位置)。於該情形時，GC動作控制部23亦可將有效資料與該有效資料之邏輯位址之兩者複製至複製目標區塊(GC目的地區塊)。GC動作控制部23亦可藉由參照與複製源區塊(GC源區塊)對應之區塊管理表32而特定出GC源區塊內之有效資料。或者，於其他實施形態中，主機2亦可指定

GC源區塊與GC目的地區塊。該等GC源區塊與GC目的地區塊可為超級區塊，亦可為實體區塊。

【0068】 而且，GC動作控制部23將所複製之有效資料之邏輯位址(標籤)、表示所複製之有效資料之舊實體記憶位置之實體位址、及表示所複製之有效資料之新實體記憶位置之實體位址通知給主機2。

【0069】 於本實施形態中，如上所述，寫入動作控制部21可將來自主機2之資料(寫入資料)與來自主機2之邏輯位址(標籤)之兩者寫入至寫入目標區塊。因此，GC動作控制部23可容易地自複製源區塊(GC源區塊)中獲取該複製源區塊(GC源區塊)內之各資料之邏輯位址，故而可容易地將所複製之有效資料之邏輯位址通知給主機2。

【0070】 快閃I/O控制電路13係以於CPU12之控制下控制NAND型快閃記憶體5之方式構成的記憶體控制電路。DRAM介面14係以於CPU12之控制下控制DRAM6之方式構成之DRAM控制電路。DRAM6之記憶區域之一部分用於儲存讀取緩衝器(RB)30、寫入緩衝器(WB)31、區塊管理表32、不良資訊管理表33。再者，該等讀取緩衝器(RB)30、寫入緩衝器(WB)31、區塊管理表32、不良資訊管理表33亦可儲存於控制器4內之未圖示之SRAM。

【0071】 其次，對用作快閃儲存裝置3與主機2之間之軟體介面之API進行說明。於本實施形態中，該API大致區分為2種API(內置驅動GC(In-Drive-GC)API、超級區塊API。

【0072】 In-Drive-GC-API包含以快閃儲存裝置3單獨執行垃圾回收(GC)為前提之指令群。該等指令群中亦可包含寫入指令(Write without Physical Address，不包含實體位址之寫入)、讀取指令(Read with

Physical Address，包含實體位址之讀取)、禁用刪除指令(Trim)、重複指令(Duplicate)、位址更新通知(Address Update Notification(Device Initiated，由裝置發起))、強制GC指令(Forced Garbage-Collection，強制垃圾回收)、GC控制指令(Garbage Collection Control)等作為基本指令。

【0073】 寫入指令(Write without Physical Address)係指定識別應寫入之用戶資料之邏輯位址(標籤)但不指定寫入目標之實體位址的寫入指令。

【0074】 讀取指令(Read with Physical Address)係指定表示讀取對象之實體記憶位置(實體區塊、實體區塊內之讀取對象位置)之實體位址的讀取指令。

【0075】 禁用刪除指令(Trim)係指定應設為無效之資料之實體位址，並對儲存裝置3指示應使與該實體位址對應之資料無效化的指令。再者，於主機2支持防止具有相同內容之複數個重複資料部儲存於儲存器之重複資料刪除(De-duplication)功能之情形時，禁用刪除指令(Trim)係用作對儲存裝置3指示應使表示參照某資料之邏輯位址之數量之參照計數減小的指令。重複指令(Duplicate)係用作對儲存裝置3指示應使表示參照某資料之邏輯位址之數量之參照計數增加的指令。

【0076】 位址更新通知(Address Update Notification(Device Initiated))用以在快閃儲存裝置3執行用於GC之資料複製動作之後，快閃儲存裝置3將所複製之資料(有效資料)之邏輯位址、該有效資料之舊實體記憶位置、該有效資料之新實體記憶位置通知給主機2。

【0077】 強制GC指令(Forced Garbage-Collection，強制垃圾回收)係用以使快閃儲存裝置3強制性地執行GC之指令。

【0078】 GC控制指令(Garbage Collection Control：垃圾回收控制)係用以對快閃儲存裝置3指示用於開始GC之條件等之指令。

【0079】 圖4表示內置驅動-GC-API(In-Drive-GC-API)用之各指令之參數與返回值之例。

【0080】 於圖4中，接續於「主機：」之標記而記載之內容係由對應之指令指定之參數，接續於「裝置：」之標記而記載之內容係對該指令之應答中包含之參數(返回值)。

【0081】 寫入指令(Write without Physical Address，不包含實體位址之寫入)亦可包含用戶位址、長度、資料、QoS域識別碼。用戶位址係用以識別應寫入之資料之邏輯位址(標籤)。用戶位址例如包含LBA、鍵值儲存之鍵、檔案識別碼等。

【0082】 對寫入指令之應答亦可包含狀態(成功/失敗)、用戶位址、快閃位址(實體位址)、長度、剩餘之可寫入資料量(distance-to-block-boundary，至區塊邊界之距離)。剩餘之可寫入資料量(distance-to-block-boundary)係可選之返回值，表示可寫入至寫入了資料之超級區塊之剩餘資料量。剩餘之可寫入資料量(distance-to-block-boundary)亦可以上述資料粒度之倍數表示。資料存在跨及不良區塊之前後之2個非不良區塊寫入之情形。因此，對寫入指令之應答亦可包含各集包含用戶位址、快閃位址(實體位址)、長度之複數個集。

【0083】 讀取指令(Read with Physical Address，包含實體位址之寫入)亦可包含快閃位址、長度。讀取指令(Read with Physical Address)亦可包含各集包含快閃位址、長度之複數個集。對讀取指令之應答亦可包含狀態、用戶位址、長度、資料。對讀取指令之應答亦可包含各集包含用戶

位址、長度之複數個集。

【0084】 禁用刪除指令(Trim)/重複指令(Duplicate)亦可包含快閃位址、長度、參照計數之增減量(reference-count-to-add-or-subtract)。禁用刪除指令(Trim)/重複指令(Duplicate)亦可包含各集包含快閃位址、長度、參照計數之增減量之複數個集。

【0085】 位址更新通知(Address Update Notification(Device Initiated))亦可包含用戶位址、舊快閃位址、新快閃位址、參照計數、長度作為由快閃儲存裝置3通知給主機2之輸出參數。例如，快閃儲存裝置3自舊實體記憶位置將資料複製至新實體記憶位置之後，將位址更新通知(Address Update Notification(Device Initiated))發送至主機2。該位址更新通知(Address Update Notification(Device Initiated))中亦可包含該資料之用戶位址、表示該資料之舊實體記憶位置之舊快閃位址、表示該資料之新實體記憶位置之新快閃位址、表示參照該資料之邏輯位址之數量之參照計數、該資料之長度。位址更新通知(Address Update Notification (Device Initiated))亦可包含各集包含用戶位址、舊快閃位址、新快閃位址、參照計數、長度之複數個集。

【0086】 強制GC指令(Forced Garbage-Collection)亦可包含QoS域識別碼、源超級區塊位址(可選)。

【0087】 GC控制指令(Garbage Collection Control)亦可包含最大資料數(maximum-number-of-data)、QoS域識別碼、GC之方法(策略)。

【0088】 超級區塊API包含以主機2指定邏輯位址(標籤)與超級區塊且快閃儲存裝置決定該超級區塊內之寫入目標區塊與該寫入目標區塊內之寫入目標位置為前提的指令群。該等指令群中包含寫入指令(Write

without Page Address，不包含頁位址之寫入)、讀取指令(Read with Physical Address)、超級區塊釋放指令(Release Super Block to Unused Super Block Pool，釋放超級區塊至未用之超級區塊池)、超級區塊分配指令(Allocate Super Block and Open Write Buffer with Block Healing，分配超級區塊以及以區塊修復之方式打開寫入緩衝器)、關閉超級區塊指令(Close Super Block and Write Buffer，關閉超級區塊與寫入緩衝器)、超級區塊資訊指令(Super Block Information)、非複製資料設定指令(Set Data not to be Copied)、資料複製指令(Data Copy without Page Address，不包含頁位址之資料複製)等作為基本指令。

【0089】 寫入指令(Write without Page Address)係指定邏輯位址(標籤)與超級區塊位址之寫入指令。讀取指令與In-Drive-GC-API用之讀取指令相同。超級區塊釋放指令(Release Super Block to Unused Super Block Pool)係用以將已分配之超級區塊釋放之指令。超級區塊分配指令(Allocate Super Block and Open Write Buffer with Block Healing)係請求超級區塊之分配之指令。超級區塊分配指令(Allocate Super Block and Open Write Buffer with Block Healing)亦可包含指定表示可並行存取之區塊之數量之並行數之參數。超級區塊資訊指令(Super Block Information)係用以獲取與特定之超級區塊相關之資訊之指令。非複製資料設定指令(Set Data not to be Copied)係用以指定超級區塊內之不應複製之資料之指令。資料複製指令(Data Copy without Page Address)係用以為了GC而複製資料之指令。該等指令之參數與返回值之例示於圖5中。於圖5中，接續於「主機：」之標記而記載之內容表示由對應之指令指定之參數，接續於「裝置：」之標記而記載之內容表示對該指令之應答中包含

之參數(返回值)。

【0090】 圖6表示快閃I/O控制電路13與複數個NAND型快閃記憶體晶片之關係。

【0091】 如圖6所示，NAND型快閃記憶體5包含複數個NAND型快閃記憶體晶片。各NAND型快閃記憶體晶片係包括包含複數個區塊(實體區塊)BLK之記憶胞陣列及控制該記憶胞陣列之周邊電路的非揮發性記憶體晶片。各NAND型快閃記憶體晶片可獨立地動作。因此，NAND型快閃記憶體晶片作為最小之並行動作單位發揮功能。NAND型快閃記憶體晶片亦稱為「NAND型快閃記憶體晶片」或「非揮發性記憶體晶片」。於圖6中，例示出如下情形，即，於快閃I/O控制電路13連接有16個通道Ch0、Ch1、…Ch15，且於該等通道Ch0、Ch1、…Ch15之各者分別連接有相同數量(例如每個通道為1個晶片)之NAND型快閃記憶體晶片。各通道包含用以與對應之NAND型快閃記憶體晶片通信之通信線(記憶體匯流排)。

【0092】 控制器4係經由通道Ch0、Ch1、…Ch15而控制NAND型快閃記憶體晶片#0～#15。控制器4可同時驅動通道Ch0、Ch1、…Ch15。

【0093】 於圖6之構成例中，可使最多16個NAND型快閃記憶體晶片並行動作。

【0094】 於本實施形態中，控制器4對分別包括複數個區塊BLK之複數個並行單位(超級區塊)進行管理。超級區塊並不限定於此，亦可包含自連接於不同通道之NAND型快閃記憶體晶片#0～#15中逐一選擇出之共計16個區塊BLK。再者，NAND型快閃記憶體晶片#0～#15之各者亦可具有多平面構成。例如，於NAND型快閃記憶體晶片#0～#15之各者具有包含2個平面之多平面構成之情形時，一個超級區塊亦可包含自與NAND型快

閃記憶體晶片#0～#15對應之32個平面中逐一選擇出之共計32個區塊BLK。

【0095】 於圖7中例示出一個超級區塊SB包括自NAND型快閃記憶體晶片#0～#14中逐一選擇出之共計16個區塊BLK的情形。於對於超級區塊SB之寫入動作中，資料係按照屬於NAND型快閃記憶體晶片#0之區塊BLK之頁P0、屬於NAND型快閃記憶體晶片#1之區塊BLK之頁P0、…屬於NAND型快閃記憶體晶片#15之區塊BLK之頁P0、屬於NAND型快閃記憶體晶片#0之區塊BLK之頁P1、屬於NAND型快閃記憶體晶片#2之區塊BLK之頁P1、…屬於NAND型快閃記憶體晶片#15之區塊BLK之頁P1、…之順序寫入。

【0096】 圖8表示超級區塊SB之區塊位址(超級區塊位址)與構成該超級區塊SB之複數個區塊(實體區塊)各自之區塊位址之關係。

【0097】 應包含於超級區塊SB之各區塊(實體區塊)之區塊位址係自該超級區塊SB之區塊位址(超級區塊位址)基於某數學規則確定。

【0098】 例如，亦可將藉由對超級區塊SB之超級區塊位址實施特定之四則運算而獲得之值決定為應包含於該超級區塊SB之各區塊之區塊位址。或者，亦可將藉由對表示超級區塊SB之超級區塊位址之複數位元實施特定之位元反轉或特定之位元移位而獲得之值決定為應包含於該超級區塊SB之各區塊之區塊位址。

【0099】 於圖8中，為了簡化圖示，例示出由具有與超級區塊SB之超級區塊位址相同之區塊位址之各區塊構成該超級區塊SB的情形。

【0100】 即，具有超級區塊位址0之超級區塊SB0係由NAND型快閃記憶體晶片#0(晶片#0)中包含之區塊位址0之區塊BLK0、NAND型快閃記

憶體晶片#1(晶片#1)中包含之區塊位址0之區塊BLK0、NAND型快閃記憶體晶片#2(晶片#2)中包含之區塊位址0之區塊BLK0、…、NAND型快閃記憶體晶片#15(晶片#15)中包含之區塊位址0之區塊BLK0所構成。

【0101】 同樣地，具有超級區塊位址1000之超級區塊SB1000係由NAND型快閃記憶體晶片#0(晶片#0)中包含之區塊位址1000之區塊BLK1000、NAND型快閃記憶體晶片#1(晶片#1)中包含之區塊位址1000之區塊BLK1000、NAND型快閃記憶體晶片#2(晶片#2)中包含之區塊位址1000之區塊BLK1000、…、NAND型快閃記憶體晶片#15(晶片#15)中包含之區塊位址1000之區塊BLK1000所構成。

【0102】 存在各晶片中包含若干個不良區塊之情形。通常，不良區塊之數量針對每一晶片而不同。

【0103】 於圖9中，示出將各晶片中包含之各不良區塊以屬於同一晶片之其他區塊置換之處理。

【0104】 於圖9中，設想如下情形，即，各晶片包含2048個區塊BLK，晶片#0包含100個不良區塊，晶片#1不包含不良區塊，晶片#2包含20個不良區塊，而且，晶片#15包含30個不良區塊。

【0105】 於晶片#0中，例如，不良區塊BLK1由晶片#0之非不良區塊BLK1948置換，不良區塊BLK5由晶片#0之非不良區塊BLK1949置換。如此一來，僅自晶片#0之開頭起合計1948(＝2048－100)個區塊成為可利用狀態，晶片#0之其餘之區塊BLK1948～BLK2047不可利用。

【0106】 因此，即便於晶片#1不包含不良區塊之情形時，晶片#1之區塊BLK1948～BLK2047亦無法用於構築超級區塊。因此，可構築之超級區塊SB之數量被限制為不良區塊最多之晶片#0內之非不良區塊之數

量。

【0107】 圖10表示不置換各不良區塊而自某超級區塊SB內之非不良區塊中選擇寫入目標區塊之動作。

【0108】 於圖10中，為了簡化圖示，例示出如下情形，即，某超級區塊(此處為超級區塊SB5)係由8個區塊、即晶片#0內之區塊BLK5、晶片#1內之區塊BLK5、晶片#2內之區塊BLK5、晶片#3內之區塊BLK5、晶片#4內之區塊BLK5、晶片#5內之區塊BLK5、晶片#6內之區塊BLK5、及晶片#7內之區塊BLK5所構成。

【0109】 於本實施形態中，設置有與各超級區塊對應之不良資訊管理表33。於超級區塊SB5用之不良資訊管理表33中，保存有包含針對每一區塊表示可使用或不可使用之1位元之資訊的不良資訊(點陣圖)。於不良資訊(點陣圖)中，“0”表示非不良區塊，“1”表示不良區塊。

【0110】 於圖10中，設想晶片#1內之區塊BLK5、晶片#4內之區塊BLK5、晶片#5內之區塊BLK5為不良區塊之情形。

【0111】 控制器4不執行將晶片#1內之區塊BLK5以晶片#1內之其他區塊置換之處理、將晶片#4內之區塊BLK5以晶片#4內之其他區塊置換之處理、將晶片#5內之區塊BLK5以晶片#5內之其他區塊置換之處理。但，控制器4參照超級區塊SB5用之不良資訊管理表33，自超級區塊SB5中包含之非不良區塊(此處為晶片#0內之區塊BLK5、晶片#2內之區塊BLK5、晶片#3內之區塊BLK5、晶片#6內之區塊BLK5、晶片#7內之區塊BLK5)中選擇一個區塊作為寫入目標區塊。控制器4決定該寫入目標區塊內之寫入目標位置，將來自主機2之寫入資料寫入至該寫入目標區塊內之寫入目標位置。而且，控制器4將表示該寫入目標區塊及該寫入目標位置之兩者

之實體位址通知給主機2。

【0112】 藉此，主機2可辨識實際寫入了寫入資料之區塊(寫入目標區塊)與該區塊內之寫入目標位置，因此，於必須讀取該資料之情形時，可將指定該通知之實體位址之讀取請求(讀取指令)發送至快閃儲存裝置3。即，首先，主機2將包含識別應寫入之資料之邏輯位址(標籤)之寫入請求(寫入指令)發送至快閃儲存裝置3。主機2自快閃儲存裝置3接收表示自除不良區塊以外之區塊中選擇為寫入目標區塊之區塊及該區塊內之寫入目標位置(實體記憶位置)之兩者的實體位址、及該資料之邏輯位址(標籤)。而且，主機2更新管理各邏輯位址(標籤)與快閃儲存裝置3之各實體位址之間之映射之主機2上之查找表(LUT)，將所接收到之實體位址映射至識別該資料之邏輯位址(標籤)。於必須讀取該資料之情形時，主機2參照主機2上之查找表(LUT)，獲取映射至該資料之邏輯位址(標籤)之實體位址，將指定所獲取之實體位址之讀取請求(讀取指令)發送至快閃儲存裝置3。

【0113】 如此，於本實施形態中，不將寫入對象之超級區塊內之不良區塊以該不良區塊所屬之晶片內之其他區塊置換，即可正常地執行對於該超級區塊之資料寫入及讀出動作。因此，即便不良區塊之數量增加，亦無須管理大量之置換資訊。又，亦不需要用於置換處理之位址轉換處理，因此可減少讀取潛伏時間。進而，基本上可構築與屬於各晶片之區塊之數量為相同數量之超級區塊，因此，即便不良區塊之數量增加，亦可利用幾乎所有之非不良區塊。

【0114】 圖11表示對快閃儲存裝置3應用之寫入指令(Write without Physical Address)。

【0115】 該寫入指令係對快閃儲存裝置3請求資料之寫入之指令。如

上所述，該寫入指令亦可包含指令ID(Identifier，識別碼)、QoS域ID、用戶位址、長度等。

【0116】 指令ID係表示該指令為寫入指令之ID(指令碼)，寫入指令中包含寫入指令用之指令ID。

【0117】 QoS域ID係可唯一地識別應寫入資料之QoS域之識別碼。根據來自某終端用戶之寫入請求自主機2發送之寫入指令亦可包含指定與該終端用戶對應之QoS域之QoS域ID。命名空間ID亦可作為QoS域ID進行處理。

【0118】 用戶位址係識別資料之邏輯位址(標籤)，例如LBA、鍵、檔案識別碼相當於用戶位址。

【0119】 長度表示應寫入之寫入資料之長度。長度可由LBA之數量指定，或者亦可以位元組指定其尺寸。

【0120】 圖12表示對圖11之寫入指令之應答。

【0121】 該應答包含用戶位址、快閃位址、長度、剩餘之可寫入資料量(distance-to-block-boundary)。

【0122】 用戶位址係圖11之寫入指令中包含之用戶位址。

【0123】 快閃位址表示根據圖11之寫入指令寫入了資料之NAND型快閃記憶體5內之實體記憶位置之實體位址。

【0124】 於本實施形態中，該實體位址例如藉由晶片識別碼、區塊位址、及偏移(區塊內偏移)之組合、或超級區塊位址與偏移(超級區塊內之偏移)之組合而指定。

【0125】 長度表示所寫入之寫入資料之長度。長度可由LBA之數量指定，或者亦可以位元組指定其尺寸。

【0126】 剩餘之可寫入資料量(distance-to-block-boundary)表示可對進行了資料寫入之超級區塊進一步寫入之資料量。

【0127】 圖13表示對快閃儲存裝置3應用之Trim指令。

【0128】 該Trim指令包含指令ID、快閃位址、長度、參照計數之減少量(reference-count-to-subtract)。

【0129】 指令ID係表示該指令為Trim指令之ID(指令碼)，Trim指令中包含Trim指令用之指令ID。

【0130】 快閃位址表示儲存有應無效化之資料(應使參照計數減小之資料)之最初之實體記憶位置。於本實施形態中，該快閃位址藉由晶片識別碼、區塊位址、及偏移(區塊內偏移)之組合、或超級區塊位址與偏移(超級區塊內之偏移)之組合而指定。

【0131】 長度表示應無效化之資料(應使參照計數減小之資料)之長度。該長度(資料長度)亦可由位元組指定。

【0132】 控制器4係使用區塊管理表32對表示複數個超級區塊之各者中包含之各個資料之有效/無效之旗標(點陣圖旗標)進行管理。於自主機2接收到包含表示儲存有應設為無效之資料之實體記憶位置之快閃位址的Trim指令之情形時，控制器4更新區塊管理表32，將與Trim指令中包含之快閃位址對應之實體記憶位置之資料所對應的旗標(點陣圖旗標)變更為表示無效之值。

【0133】 於支持重複資料刪除(De-duplication)功能之實例中，於區塊管理表32中，管理複數個超級區塊之各者中包含之資料對應之參照計數。參照計數之減少量表示應使參照計數減小之量。

【0134】 其次，參照圖14，就對具有不良區塊之超級區塊寫入資料

之動作進行說明。

【0135】 此處，為了簡化圖示，設想某一個超級區塊SB0由4個區塊BLK0(晶片#0)、BLK0(晶片#1)、BLK0(晶片#2)、BLK0(晶片#3)所構成，其中，BLK0(晶片#2)為不良區塊。

【0136】 控制器4係以避開不良區塊之方式，按照區塊BLK0(晶片#0)之頁0、區塊BLK0(晶片#1)之頁0、區塊BLK0(晶片#3)之頁0、區塊BLK0(晶片#0)之頁1、區塊BLK0(晶片#1)之頁1、區塊BLK0(晶片#3)之頁1、…之順序寫入資料。

【0137】 於頁尺寸為16 K位元組(16 KB)，且資料粒度為4 K位元組(4 KB)之情形時，最初之16 K位元組資料(D1～D4)被寫入至區塊BLK0(晶片#0)之頁0。隨後之16 K位元組資料(D5～D8)被寫入至區塊BLK0(晶片#1)之頁0。對於BLK0(晶片#2)之寫入被跳過，隨後之16 K位元組資料(D9～D12)被寫入至區塊BLK0(晶片#3)之頁0。

【0138】 圖15表示圖12之應答中包含之實體位址之構成例。

【0139】 實體位址如圖15之上部所示，係由選擇為寫入目標區塊之區塊所屬之晶片之晶片識別碼、與該選擇之區塊對應之區塊位址、及自該選擇之區塊之開頭至寫入目標位置之偏移所構成。自所選擇之區塊之開頭至寫入目標位置之偏移包含頁位址與頁內偏移。

【0140】 或者，如圖15之下部所示，實體位址係由寫入目標區塊所屬之超級區塊對應之區塊位址(超級區塊位址)、及自超級區塊之開頭至寫入目標位置之偏移所構成。自超級區塊之開頭至寫入目標位置之偏移包含晶片識別碼、頁位址、頁內偏移。

【0141】 圖16表示超級區塊之超級區塊位址(第1區塊位址)與該超級

區塊中包含之各區塊之區塊位址(第2區塊位址)之關係。

【0142】 為了使說明簡單，設想超級區塊SB0、SB1、SB2分別由4個區塊構成之情形。

【0143】 超級區塊SB0包含區塊80、區塊81、區塊82、及區塊83。該等區塊80、區塊81、區塊82、及區塊83之各者包含自超級區塊SB0之超級區塊位址(第1區塊位址)基於數學規則確定之區塊位址(第2區塊位址)。於區塊81為不良區塊之情形時，自主機2將資料寫入至自區塊80、區塊82、區塊83選擇之寫入目標區塊。因此，不良區塊81之第2區塊位址(不良區塊81所屬之晶片之晶片識別碼)不發回給主機2。

【0144】 超級區塊SB1包含區塊84、區塊85、區塊86及區塊87。該等區塊84、區塊85、區塊86、及區塊87之各者包含自超級區塊SB1之超級區塊位址(第1區塊位址)基於數學規則確定之區塊位址(第2區塊位址)。於區塊86及區塊87為不良區塊之情形時，自主機2將資料寫入至自區塊84、區塊85選擇之寫入目標區塊。因此，不良區塊86之第2區塊位址(不良區塊86所屬之晶片之晶片識別碼)、及不良區塊87之第2區塊位址(不良區塊87所屬之晶片之晶片識別碼)不發回給主機2。

【0145】 超級區塊SB2包含區塊88、區塊89、區塊90及區塊91。該等區塊88、區塊89、區塊90、及區塊91之各者包含自超級區塊SB2之超級區塊位址(第1區塊位址)基於數學規則確定之區塊位址(第2區塊位址)。於區塊88為不良區塊之情形時，自主機2將資料寫入至自區塊89、區塊90、及區塊91選擇之寫入目標區塊。因此，不良區塊88之第2區塊位址(不良區塊88所屬之晶片之晶片識別碼)不發回給主機2。

【0146】 圖17、圖18表示將邏輯位址與資料之對寫入至區塊內之頁

之動作。

【0147】 於各區塊中，各頁亦可包含用以儲存用戶資料之用戶資料區域與用以儲存管理資料之冗餘區域。頁尺寸為16 KB＋α。

【0148】 控制器4將4 KB用戶資料及該4 KB用戶資料對應之邏輯位址(例如LBA)之兩者寫入至寫入目標區塊BLK。於該情形時，如圖17所示，可將分別包含LBA與4 KB用戶資料之4個資料集寫入至同一頁中。區塊內偏移亦可表示集邊界。

【0149】 或者，如圖18所示，亦可將4個4 KB用戶資料寫入至頁內之用戶資料區域，將該等4個4 KB用戶資料對應之4個LBA寫入至該頁內之冗餘區域。

【0150】 利用圖19，對將複數個資料部與根據該等資料部算出之消失校正碼寫入至具有不良區塊之超級區塊的動作進行說明。

【0151】 為了使用超級區塊內之複數個區塊來實現RAID(Redundant Arrays of Independent Disks，獨立磁碟冗餘陣列)之機制，控制器4如圖19所示，跨及屬於一個超級區塊中包含之複數個區塊且分別具有相同頁位址的複數個頁，將複數個資料部與根據該等複數個資料部算出之1個以上之消失校正碼寫入。屬於複數個區塊且分別具有相同頁位址之複數個頁被作為超級頁加以參照。

【0152】 於圖19之上部，示出於超級區塊SB0中對超級頁寫入資料及消失校正碼之情況。超級頁包括BLK0(晶片#0)之頁0、BLK0(晶片#1)之頁0、BLK0(晶片#2)之頁0、BLK0(晶片#3)之頁0、BLK0(晶片#4)之頁0、BLK0(晶片#5)之頁0、BLK0(晶片#6)之頁0、BLK0(晶片#7)之頁0。

【0153】 資料被寫入至作為非不良區塊之BLK0(晶片#0)、BLK0(晶

片#2)、BLK0(晶片#3)、BLK0(晶片#4)之各頁0。

【0154】 消失校正碼例如包含里德-所羅門碼、同位(parity)等。消失校正碼被寫入至非不良區塊內之頁。又，消失校正碼係視為於各不良區塊內之頁儲存有預先規定之值(例如，全為“0”之位元串、或全為“1”之位元串)而算出。

【0155】 於圖19之上部所示之例中，例示出將2個消失校正碼寫入至超級區塊SB0之超級頁之情形。2個消失校正碼被寫入至最後之2個非不良區塊內之頁。於圖19之上部所示之例中，最後之區塊即BLK0(晶片#7)為非不良區塊，自最後數起為第2個之區塊即BLK0(晶片#6)為不良區塊，自最後數起為第3個之區塊即BLK0(晶片#5)為非不良區塊，因此，2個消失校正碼被寫入至BLK0(晶片#5)之頁0與BLK0(晶片#7)之頁0。

【0156】 於編碼處理中，控制器4係基於寫入至BLK0(晶片#0)之資料部、假定為寫入至作為不良區塊之BLK0(晶片#0)之預先規定之值、寫入至BLK0(晶片#2)之資料部、寫入至BLK0(晶片#3)之資料部、寫入至BLK0(晶片#4)之資料部、及假定為寫入至BLK0(晶片#6)之預先規定之值，算出2個消失校正碼。

【0157】 藉此，與超級區塊內之不良區塊/非不良區塊之模式無關，藉由用於編碼之相同運算便可容易地算出消失校正碼。

【0158】 又，控制器4係視為於各不良區塊內之頁儲存有預先規定之值而進行使用消失校正碼之解碼處理。

【0159】 於圖19之下部所示之例中，構成超級區塊SB1之最後之區塊即BLK1(晶片#7)為不良區塊，自最後數起為第2個之區塊即BLK1(晶片#6)為不良區塊，自最後數起為第3個之區塊即BLK1(晶片#5)為非不良區

塊，因此，2個消失校正碼被寫入至BLK1(晶片#5)之頁0與BLK1(晶片#6)之頁0。

【0160】 其次，利用圖20，對主機2內之快閃轉譯部2A與快閃儲存裝置3內之寫入動作控制部21之關係進行說明。

【0161】 於主機2側，快閃轉譯部2A係於進行資料之寫入之情形時，將包含識別該資料之標籤(例如LBA)之寫入指令發送至快閃儲存裝置3。再者，於使用超級區塊用之API之情形時，快閃轉譯部2A將指定識別資料之標籤(例如LBA)與並行單位(Parallel Unit)之區塊位址之寫入指令(Write without Page Address)發送至快閃儲存裝置3。由於並行單位(Parallel Unit)藉由一個超級區塊而實現，故並行單位(Parallel Unit)之區塊位址係該超級區塊之超級區塊位址。

【0162】 快閃儲存裝置3側包括寫入動作控制部21、快閃I/O控制電路13、不良資訊管理表33、NAND型快閃記憶體晶片。又，於快閃儲存裝置3內，以不良資訊管理表33管理保存針對每一超級區塊表示可使用或不可使用之至少1位元之資訊的不良資訊。

【0163】 於自主機2側接收到包含作為用以識別應寫入之資料之資訊之標籤(例如LBA)之寫入請求的情形時，寫入動作控制部21係參照不良資訊管理表33之不良資訊，自一個超級區塊中包含之非不良區塊中選擇一個區塊作為寫入目標區塊，並決定應寫入資料之所選擇之區塊內之寫入目標位置(寫入目標頁、該頁內之偏移)。

【0164】 再者，於自主機2接收到指定並行單位之區塊位址(超級區塊位址)之寫入指令的情形時，寫入動作控制部21係自具有該指定之區塊位址(超級區塊位址)之超級區塊中包含之非不良區塊中選擇寫入目標區

塊，並決定應寫入資料之所選擇之區塊內之寫入目標位置(寫入目標頁、該頁內之偏移)。

【0165】 而且，寫入動作控制部21將指定寫入目標區塊所屬之晶片之晶片識別碼(晶片ID)、寫入目標區塊之區塊位址(原始區塊(Raw Block))、寫入目標頁(原始頁(Raw Page))、寫入目標頁內之偏移(Offset)的寫入指示發送至快閃I/O控制電路13。

【0166】 又，寫入動作控制部21將表示寫入目標區塊及寫入目標位置之兩者之快閃位址(晶片ID、原始區塊、原始頁、偏移)、及標籤(例如LBA)通知給主機2。

【0167】 接收到寫入指示之快閃I/O控制電路13基於寫入指示，將資料寫入至寫入目標位置。

【0168】 於主機2側，快閃轉譯部2A若自快閃儲存裝置3接收快閃位址(晶片ID、原始區塊、原始頁、偏移)、及標籤(例如LBA)，則更新主機2中管理之LUT。此時，將快閃位址(晶片ID、原始區塊、原始頁、偏移)與標籤(例如LBA)建立對應。

【0169】 快閃轉譯部2A係於進行讀取請求之情形時，將指定快閃位址(晶片ID、原始區塊、原始頁、偏移)之讀取請求發送至快閃儲存裝置3。

【0170】 於快閃儲存裝置3側，自主機2接收到指定快閃位址(晶片ID、原始區塊、原始頁、偏移)之讀取請求的情形時，快閃I/O控制電路13基於快閃位址，讀取資料。讀取對象之區塊根據晶片ID與原始區塊而特定出。讀取對象之頁根據原始頁而特定出。讀取對象之頁內之讀取對象位置根據偏移而特定出。

【0171】 其次，利用圖21，對藉由主機2與快閃儲存裝置3執行之寫入動作及讀取動作進行說明。

【0172】

＜寫入動作＞

(1)寫入指令之接收

於快閃儲存裝置3中，自主機2接收之包含LBA之寫入指令、及資料被暫時儲存於快閃儲存裝置3內之寫入緩衝器31。

【0173】

(2)不良資訊之參照

寫入動作控制部21係參照以不良資訊管理表33管理之不良資訊，自寫入對象之超級區塊中包含之非不良區塊中選擇一個區塊作為寫入目標區塊，並決定寫入目標區塊內之寫入目標位置。

【0174】

(3)寫入指示

寫入動作控制部21當決定寫入目標區塊與寫入目標區塊內之寫入目標位置時，將指定表示寫入目標區塊與寫入目標區塊之兩者之快閃位址(原始位址)的寫入指示經由寫入緩衝器31發送至快閃I/O控制電路13。此處，原始位址中包含晶片ID、原始區塊、原始頁、偏移。接收到原始位址之快閃I/O控制電路13係基於原始位址，將寫入資料寫入至選擇為寫入目標區塊之區塊內之寫入目標位置。

【0175】

(4)向主機之寫入目標之通知

寫入動作控制部21將原始位址與LBA通知給主機2。藉此，主機2可

更新LUT並將該原始位址映射至該LBA。

【0176】

＜讀取動作＞

(11)LBA之通知

讀取剖析器2B若接收包含LBA之讀取指令，則將LBA通知給快閃轉譯部2A。

【0177】

(12)原始位址之獲取

快閃轉譯部2A若自讀取剖析器2B接收LBA，則參照LUT，獲取與所接收到之LBA對應之原始位址，並將所獲取之原始位址返回給讀取剖析器2B。藉此，讀取剖析器2B可獲取原始位址，從而可將包含LBA之讀取指令轉換為包含原始位址之讀取指令。

【0178】

(13)讀取指示

讀取剖析器2B將包含原始位址之讀取指令發送至快閃儲存裝置3。於快閃儲存裝置3中，接收到包含原始位址之讀取指令之快閃I/O控制電路13基於原始位址，讀取資料，並將讀取資料發送至讀取緩衝器30。於讀取緩衝器30，暫時記憶讀取資料。

【0179】

(14)讀取資料向主機之發送

讀取緩衝器30中暫時記憶之讀取資料被發送至主機2。

圖22之序列圖表示藉由主機2與快閃儲存裝置3執行之寫入動作處理之順序。

【0180】 主機2將包含QoS域ID、用戶位址(邏輯位址)、寫入資料、長度之寫入指令(Write without Physical Address)發送至快閃儲存裝置3。當快閃儲存裝置3之控制器4接收到該寫入指令時，控制器4自一個超級區塊(寫入對象之超級區塊)中包含之非不良區塊中選擇一個區塊作為寫入目標區塊，並決定所選擇之區塊內之寫入目標位置(步驟S11)。於步驟S11中，該寫入對象之超級區塊亦可為屬於由QoS域ID特定出之QoS域之超級區塊。於複數個超級區塊屬於該QoS域之情形時，將該等複數個超級區塊之一個選擇為寫入對象之超級區塊。

【0181】 控制器4將自主機2接收之寫入資料寫入至寫入目標位置(步驟S12)。於步驟S12中，控制器4將用戶位址(例如LBA)與寫入資料之兩者寫入至寫入目標區塊之寫入目標位置。

【0182】 控制器4更新區塊管理表32，將與所寫入之資料對應之點陣圖旗標(即，寫入了該資料之實體記憶位置之實體位址對應之點陣圖旗標)自0變更為1(步驟S13)。例如，設想如圖23所示般將起始LBA為LBAx之16 K位元組更新資料寫入至區塊BLK11之頁1之與偏移＋0～＋3對應之實體記憶位置的情形。於該情形時，如圖24所示，於區塊BLK11用之區塊管理表中，頁1之與偏移＋0～＋3對應之點陣圖旗標分別自0變更為1。

【0183】 控制器4將對該寫入指令之應答返回給主機2(步驟S14)。應答中包含用戶位址、實體位址(快閃位址)、長度。例如，若如圖23所示般，起始LBA為LBAx之16 K位元組更新資料已被寫入至區塊BLK11之頁1之與偏移＋0～＋3對應之實體記憶位置，則將包含LBAx、快閃位址(區塊BLK11所屬之晶片之晶片識別碼、區塊BLK11之區塊位址、頁位址(＝1)、頁內偏移(＝＋0))、長度(＝4)的應答自控制器4發送至主機2。快閃位

址藉由區塊位址、頁位址、頁內偏移之組合而表達，但於以下之說明中，為了簡化快閃位址之說明，而省略快閃位址內之晶片識別碼之說明。

【0184】當主機2接收到該應答時，主機2更新主機2所管理之LUT，將快閃位址(實體位址)映射至所寫入之寫入資料對應之各用戶位址(步驟S21)。如圖25所示，LUT包含與複數個用戶位址(邏輯位址)分別對應之複數個條目。於與某用戶位址(例如某LBA)對應之條目，儲存表示儲存有與該LBA對應之資料之NAND型快閃記憶體5內之位置(實體記憶位置)的實體位址。若如圖23所示般，起始LBA為LBAx之16 K位元組更新資料已被寫入至區塊BLK11之頁1之與偏移＋0～＋3對應之實體記憶位置，則如圖25所示，LUT更新，於與LBAx對應之條目儲存BLK11、頁1、偏移＋0，於與LBAx＋1對應之條目儲存BLK11、頁1、偏移＋1，於與LBAx＋2對應之條目儲存BLK11、頁1、偏移＋2，於與LBAx＋3對應之條目儲存BLK11、頁1、偏移＋3。

【0185】然後，主機2將用以使因上述更新資料之寫入而變得無用之以前之資料無效化的Trim指令發送至快閃儲存裝置3。如圖23所示，以前之資料儲存於區塊BLK0之頁0之與偏移＋0、偏移＋1、偏移＋2、偏移＋3對應之位置的情形時，如圖26所示，將指定快閃位址(區塊位址(＝BLK0)、頁位址(＝頁0)、頁內偏移(＝＋0))、長度(＝4)的Trim指令自主機2發送至快閃儲存裝置3。快閃儲存裝置3之控制器4根據該Trim指令，更新區塊管理表32(步驟S15)。於步驟S15中，如圖26所示，於區塊BLK0用之區塊管理表中，頁0之與偏移＋0～＋3對應之點陣圖旗標分別自1變更為0。

【0186】圖27表示對快閃儲存裝置3應用之讀取指令(Read with

Physical Address)。

【0187】 讀取指令係對快閃儲存裝置3請求資料之讀出之指令。該讀取指令包含指令ID、快閃位址、長度、傳送目的地指標。

【0188】 指令ID係表示該指令為讀取指令之ID(指令碼)，讀取指令中包含讀取指令用之指令ID。

【0189】 快閃位址表示應讀出資料之最初之實體記憶位置之快閃位址(實體位址)。長度表示應讀取之資料之長度。

【0190】 傳送目的地指標表示應被傳送所讀出之資料之主機2內之記憶體上之位置。

【0191】 一個讀取指令可指定複數個快閃位址(實體位址)與長度之組。

【0192】 此處，讀取指令中亦可包含2組以上之快閃位址及長度之組。作為包含2組以上之快閃位址及長度之組之情形，設想如下情形，即，由於進行了資料寫入之超級區塊中存在不良區塊，故寫入資料被寫入至隔著不良區塊之2個區塊。

【0193】 圖28表示讀取動作。

【0194】 此處，設想自主機2接收到指定區塊位址(＝BLK2)、頁位址(＝頁1)、頁內偏移(＝＋1)、長度(＝3)之讀取指令的情形。快閃儲存裝置3之控制器4基於區塊位址(＝BLK2)、頁位址(＝頁1)、頁內偏移(＝＋1)、長度(＝3)，自BLK2讀取資料d1～d3。於該情形時，控制器4自BLK2之頁1讀取相當於1頁尺寸之資料，自該讀取資料抽取資料d1～資料d3。繼而，控制器4將資料d1～資料d3傳送至傳送目的地指標所指定之主機記憶體上。

【0195】 圖29之序列圖表示藉由主機2與快閃儲存裝置3執行之讀取處理之順序。

【0196】 主機2參照主機2所管理之LUT，將來自用戶應用之讀取請求中包含之用戶位址(邏輯位址)轉換為快閃位址。然後，主機2將指定該快閃位址、長度之讀取指令發送至快閃儲存裝置3。

【0197】 當快閃儲存裝置3之控制器4自主機2接收到讀取指令時，控制器4基於由該讀取指令指定之快閃位址，決定讀取對象之區塊、讀取對象之頁、讀取對象之頁內位置(步驟S31)。控制器4自NAND型快閃記憶體5讀取由快閃位址、長度規定之資料(步驟S32)，並將該讀取資料發送至主機2。

【0198】 圖30表示對快閃儲存裝置3應用之垃圾回收(GC)控制指令。

【0199】 GC控制指令亦可包含指令ID、策略(GC之方法)、QoS域ID、最大資料數(maximum-number-of-data)等。

【0200】 指令ID係表示該指令為GC控制指令之ID(指令碼)，GC控制指令中包含GC控制指令用之指令ID。

【0201】 策略(GC之方法)係指定用以選擇GC候補區塊(GC源區塊)之條件(GC策略)的參數。快閃儲存裝置3之控制器4支持複數個GC策略。

【0202】 控制器4所支持之GC策略中亦可包含優先選擇有效資料量較少之區塊作為GC候補區塊(GC源區塊)之策略(貪婪式)。

【0203】 又，控制器4所支持之GC策略中亦可包含如下策略，即，將集中有具有較低更新頻率之資料(冷資料)之區塊優先於集中有具有較高更新頻率之資料(熱資料)之區塊選擇為GC候補區塊(GC源區塊)。

【0204】 進而，GC策略亦可指定GC開始條件。GC開始條件例如亦可表示剩餘空閒區塊之個數。

【0205】 控制器4亦可利用活動區塊清單管理包含有效資料之超級區塊群，於執行GC之情形時，基於由GC控制指令指定之GC策略，自活動區塊清單所管理之超級區塊群中選擇一個以上之GC候補超級區塊(GC源區塊)。

【0206】 QoS域ID係指定應實施GC之QoS域之參數。控制器4係自屬於由QoS域ID指定之QoS域之超級區塊群、即該QoS域對應之活動區塊清單中，選擇一個以上之GC候補超級區塊(GC源區塊)。

【0207】 最大資料數表示執行GC時複製之資料量之上限值。即，GC動作不可超過最大資料數而執行。

【0208】 控制器4亦可於與QoS域對應之剩餘空閒區塊之數量成為策略所指定之閾值以下之情形時，開始GC。

【0209】 圖31表示對快閃儲存裝置3應用之強制GC指令(Forced Garbage-Collection)。

【0210】 強制GC指令亦可包含指令ID、QoS域ID、超級區塊位址等。當接收到該強制GC指令時，控制器立即執行GC。

【0211】 圖32表示自快閃儲存裝置3發送至主機2之位址更新通知(Address Update Notification(Device Initiated))。

【0212】 位址更新通知用以將快閃儲存裝置3中藉由執行GC動作而發生變更之資料之儲存位置通知給主機2。亦可包含用戶位址、舊快閃位址、新快閃位址、參照計數、長度等。

【0213】 用戶位址係識別所複製之資料之位址。

【0214】 舊快閃位址表示儲存有所複製之資料之舊實體記憶位置之實體位址(舊實體位址)。

【0215】 新快閃位址表示儲存有所複製之資料之新實體記憶位置之實體位址(新實體位址)。

【0216】 參照計數表示參照所複製之資料之用戶位址之數量。

【0217】 長度表示所複製之資料之長度。

【0218】 圖33之序列圖表示藉由快閃儲存裝置3執行之GC動作之順序。

【0219】 快閃儲存裝置3之控制器4自屬於主機2所指定之QoS域之超級區塊群中，選擇有效資料與無效資料混合存在之一個以上之GC源區塊(GC源超級區塊)(步驟S41)。繼而，控制器4自屬於該QoS域之空閒區塊群(空閒超級區塊群)中選擇一個以上之空閒區塊(空閒超級區塊)，並將所選擇之空閒區塊分配為GC目的地區塊(GC目的地超級區塊)(步驟S42)。

【0220】 控制器4將GC源區塊(GC源超級區塊)內之所有有效資料複製至GC目的地區塊(GC目的地超級區塊)(步驟S43)。於步驟S43中，控制器4不僅將GC源區塊(GC源超級區塊)內之有效資料，而且將該有效資料及該有效資料對應之用戶位址(邏輯位址)之兩者自GC源區塊(GC源超級區塊)複製至GC目的地區塊(GC目的地超級區塊)。藉此，可於GC目的地區塊(GC目的地超級區塊)內保存資料與用戶位址(邏輯位址)之對。

【0221】 然後，控制器4針對所複製之每一有效資料，使用位址更新通知將用戶位址(邏輯位址)、舊快閃位址、新快閃位址、長度通知給主機2(步驟S44)。

【0222】 當主機2接收到該位址更新通知時，主機2更新主機2所管理

之LUT，將新快閃位址映射至所複製之各有效資料對應之用戶位址(邏輯位址)(步驟S51)。

【0223】 圖34表示為了GC而執行之資料複製動作之例。

【0224】 於圖34中，設想如下情形，即，將GC源區塊(此處為區塊BLK50)之頁1之與偏移＋0對應之位置上儲存之有效資料(LBA＝10)複製至GC目的地區塊(此處為區塊BLK100)之頁0之與偏移＋0對應之位置，將GC源區塊(此處為區塊BLK50)之頁2之與偏移＋2對應之位置上儲存之有效資料(LBA＝20)複製至GC目的地區塊(此處為區塊BLK100)之頁0之與偏移＋1對應之位置。於該情形時，控制器4將｛LBA10、舊快閃位址(BLK50、頁1、偏移(＝＋0))、新快閃位址(LBA100、頁0、偏移(＝＋0))、長度(＝1)｝、｛LBA20、舊快閃位址(BLK50、頁2、偏移(＝＋2))、新快閃位址(LBA100、頁0、偏移(＝＋1))、長度(＝1)｝通知給主機2(位址更新通知)。

【0225】 圖35表示基於圖34之資料複製動作之結果更新之主機2之LUT之內容。

【0226】 於該LUT中，與LBA10對應之快閃位址(區塊位址、頁位址、及偏移(頁內偏移))自BLK50、頁1、偏移(＝＋0)更新為BLK100、頁0、偏移(＝＋0)。同樣地，與LBA20對應之快閃位址(區塊位址、頁位址、及偏移(頁內偏移))自BLK50、頁2、偏移(＝＋2)更新為BLK100、頁0、偏移(＝＋1)。

【0227】 LUT更新後，主機2亦可將指定BLK50、頁1、偏移(＝＋0)之Trim指令發送至快閃儲存裝置3，使BLK50之頁1之與偏移(＝＋0)對應之位置上儲存之資料無效化。進而，主機2亦可將指定BLK50、頁2、偏

移(＝＋2)之Trim指令發送至快閃儲存裝置3，使BLK50之頁2之與偏移(＝＋2)對應之位置上儲存之資料無效化。

【0228】 圖36表示對寫入指令之應答與GC用回調處理(位址更新通知)之關係。

【0229】 存在發生如下實例之情形，即，控制器4於複製與某用戶位址(邏輯位址)對應之有效資料之期間，自主機2接收指定該用戶位址之寫入指令。

【0230】 於圖36中，設想如下情形，即，於圖34之資料複製動作(與LBA10對應之資料複製動作)之執行中，自主機2接收到指定LBA10之寫入指令。

【0231】 控制器4將自主機2接收之寫入資料寫入至寫入目標區塊(此處係寫入至BLK3之頁0之與偏移＋0對應之位置)。然後，控制器4將｛LBA10、BLK3、頁0、偏移(＝＋0)｝作為對寫入指令之應答返回給主機2。

【0232】 主機2更新LUT，將與LBA10對應之區塊位址、頁位址、偏移(頁內偏移)自BLK50、頁1、偏移(＋0)變更為BLK3、頁0、偏移(＋0)。

【0233】 若之後將LBA10之新快閃位址自控制器4通知給主機2，則存在表示儲存有與LBA10對應之最新資料之位置之區塊位址、頁位址及偏移(BLK3、頁0、偏移(＋0))被錯誤地變更為與LBA10對應之新快閃位址(此處為BLK100、頁0、偏移(＝＋0))的可能性。

【0234】 於本實施形態中，控制器4不僅可將LBA10、新快閃位址(BLK100、頁0、偏移(＝＋0))、長度＝1，亦可將舊快閃位址(BLK50、

頁1、偏移(＝＋0))通知給主機2(位址更新通知)。主機2係於舊快閃位址(BLK50、頁1、偏移(＝＋0))與藉由LUT當前映射至LBA10之區塊位址、頁位址、偏移不一致之情形時，不更新LUT。藉此，可防止表示儲存有與LBA10對應之最新資料之位置之區塊位址、頁位址及偏移(BLK3、頁0、偏移(＋0))被錯誤地變更為與LBA10對應之新快閃位址(此處為BLK100、頁0、偏移(＝＋0))。

【0235】 圖37之流程圖表示主機2所執行之LUT更新處理之順序。

【0236】 主機2若接收到位址更新通知(步驟S101之是(YES))，則將舊快閃位址與LUT上之當前之實體位址加以比較(步驟S102)。比較之結果，若舊快閃位址與當前之實體位址一致(步驟S103之是)，則主機2將與用戶位址對應之當前之實體位址更新為新快閃位址(步驟S104)。繼而，主機2將Trim指令發送至快閃儲存裝置，使與舊快閃位址對應之位置上儲存之資料無效化(步驟S105)。

【0237】 另一方面，若舊快閃位址與當前之實體位址不一致(步驟S106之否(NO))，則主機2維持與用戶位址對應之當前之實體位址(步驟S106)。藉此，可防止位址更新通知所通知的表示儲存有最新資料之位置之區塊位址、頁位址及偏移被錯誤地變更為新快閃位址。

【0238】 於圖38中，例示出區塊BLK1用之區塊管理表。

【0239】 區塊BLK1用之區塊管理表包含與區塊BLK1之頁位址及頁內偏移之組之各者對應之複數個條目。

【0240】 例如，於頁0、偏移＋0對應之條目，儲存區塊BLK1之頁0、偏移＋0對應之位置上儲存之4 KB資料所對應的參照計數。同樣地，於頁0、偏移＋1對應之條目，儲存區塊BLK1之頁0、偏移＋1對應之位置

上儲存之4 KB資料所對應的參照計數。

【0241】 參照計數為1以上之資料係有效資料，參照計數為0之資料係無效資料。

【0242】 快閃儲存裝置3基於自主機2接收之重複指令/Ttim指令，使參照計數遞增/遞減。

【0243】 圖39表示為了參照計數之管理而對快閃儲存裝置3應用之重複指令。

【0244】 重複指令係請求快閃儲存裝置3使儲存於某實體位址(區塊位址、頁位址、頁內偏移)之資料之參照計數增加的指令。

【0245】 該重複指令亦可包含指令ID、快閃位址、長度等。

【0246】 指令ID係表示該指令為重複指令之ID(指令碼)，重複指令中包含重複指令用之指令ID。

【0247】 快閃位址表示儲存有應使參照計數增加之資料之最初之實體記憶位置。

【0248】 長度表示應使參照計數增加之資料之長度。

【0249】 控制器4係於自主機2接收到包含表示儲存有應使參照計數增加之資料之實體記憶位置之區塊位址、頁位址、頁內偏移之重複指令的情形時，更新區塊管理表32，使重複指令中包含之區塊編號、頁編號、頁內偏移對應之實體記憶位置之資料所對應的參照計數增加。

【0250】 圖40之序列圖表示參照計數遞增/遞減處理。

【0251】 當快閃儲存裝置3之控制器4自主機2接收到重複指令時，控制器4使由重複指令指定之快閃位址(區塊位址、頁位址、偏移(頁內偏移))對應之參照計數、即由該區塊位址、頁位址、偏移指定之NAND型快閃記

憶體5內之實體記憶位置上儲存之資料對應的參照計數例如增加1(步驟S61)。於該情形時，控制器4更新具有由重複指令指定之區塊位址之區塊所對應的區塊管理表32。藉由該區塊管理表32之更新，由重複指令指定之實體記憶位置對應之區塊管理表32內之條目中儲存之參照計數例如增加1。於由重複指令指定之長度為2以上之情形時，不僅由重複指令指定之頁位址及偏移對應之參照計數，而且接續於該頁位址及偏移之若干個頁位址及偏移對應之參照計數例如亦增加1。

【0252】 當快閃儲存裝置3之控制器4自主機2接收到Trim指令時，控制器4使由Trim指令指定之快閃位址(區塊位址、頁位址、偏移(頁內偏移))對應之參照計數、即由該區塊位址、頁位址、偏移指定之NAND型快閃記憶體5內之實體記憶位置上儲存之資料對應的參照計數例如減小1(步驟S62)。於該情形時，控制器4更新具有由Trim指令指定之區塊位址之區塊對應之區塊管理表32。藉由該區塊管理表32之更新，由Trim指令指定之頁位址及偏移對應之區塊管理表32內之條目中儲存之參照計數例如減小1。於由Trim指令指定之長度為2以上之情形時，不僅由Trim指令指定之頁位址及偏移對應之參照計數，而且接續於該偏移之若干個頁位址及偏移對應之參照計數例如亦減小1。

【0253】 於GC中，控制器4參照與GC源區塊對應之區塊管理表，以尺寸為4 KB之資料為單位，判定GC源區塊內之資料為有效資料還是無效資料。控制器4將參照計數為0之資料判定為無效資料，將參照計數為1以上之資料判定為有效資料。而且，控制器4將有效資料(參照計數為1以上之資料)及該有效資料對應之邏輯位址自GC源區塊(GC源超級區塊)複製至GC目的地區塊(GC目的地超級區塊)。

【0254】 更詳細而言，控制器4係於執行NAND型快閃記憶體5之垃圾回收之情形時，選擇用於垃圾回收之GC源區塊(GC源超級區塊)及GC目的地區塊(GC目的地超級區塊)。控制器4將儲存於GC源區塊(GC源超級區塊)內的參照計數為1以上之第1資料(有效資料)與第1資料之邏輯位址之兩者複製至GC目的地區塊(GC目的地超級區塊)。而且，控制器4將第1資料之用戶位址(邏輯位址)、第1資料之複製目標實體記憶位置(新實體記憶位置)之實體位址、及第1資料之複製源實體記憶位置(舊實體記憶位置)之實體位址通知給主機2。

【0255】 圖41之流程圖表示快閃儲存裝置3所執行之超級區塊分配處理之順序。

【0256】 若接收來自主機2之超級區塊分配請求(並行單位分配請求)(步驟S71之是)，則控制器4選擇包含由該分配請求指定之並行數以上之非不良區塊之超級區塊(步驟S72)。

【0257】 其次，控制器4對主機2分配所選擇之超級區塊(步驟S73)。

【0258】 繼而，若接收來自主機2之寫入請求(步驟S74之是)，則控制器4將寫入資料寫入至該分配之超級區塊內之非不良區塊(步驟S75)。然後，控制器4向主機返回應答(步驟S76)。

【0259】 其次，利用圖42，對以超級區塊內之所有非不良區塊自該超級區塊之開頭連續地邏輯性地配置之方式轉換存取對象區塊之位址的位址轉換動作進行說明。

【0260】 於圖42中，示出如下情況，即，超級區塊SB5係由BLK5(晶片#0)、BLK5(晶片#1)、BLK5(晶片#2)、BLK5(晶片#3)、BLK5(晶片#4)、BLK5(晶片#5)、BLK5(晶片#6)、BLK5(晶片#7)所構

成，其中，BLK5(晶片#2)、BLK5(晶片#5)為不良區塊。又，BLK5(晶片#0)、BLK5(晶片#1)、BLK5(晶片#2)、BLK5(晶片#3)、BLK5(晶片#4)、BLK5(晶片#5)、BLK5(晶片#6)、BLK5(晶片#7)分別以區塊編號(區塊位址)0、1、2、3、4、5、6、7自外部指定。

【0261】 位址轉換係以超級區塊SB5內之6個非不良區塊(BLK5(晶片#0)、BLK5(晶片#1)、BLK5(晶片#3)、BLK5(晶片#4)、BLK5(晶片#6)、BLK5(晶片#7))作為區塊編號0～6之區塊進行處理之方式執行。

【0262】 當執行位址轉換動作時，BLK5(晶片#0)、BLK5(晶片#1)、BLK5(晶片#2)、BLK5(晶片#3)、BLK5(晶片#4)各自之區塊編號分別以1→1、2→3、3→4、4→6、5→7之方式轉換。藉由該位址轉換動作，可使得在自超級區塊SB5之開頭開始之連續之偏移位址內不包含不良區塊之區塊編號。由此，藉由該位址轉換動作，而超級區塊SB5內之所有非不良區塊自該超級區塊SB5之開頭連續地邏輯性地配置，因此，可將超級區塊SB5視為宛如不包含不良區塊之小尺寸之超級區塊(僅由6個非不良區塊構成之超級區塊)。其結果，即便為較長之資料，亦可藉由表示該資料之起始實體記憶位置之一個實體位址與一個長度之組合而表達該資料之實體記憶位置。

【0263】 其次，利用圖43，對用於位址轉換動作之不良資訊管理表33與位址轉換之例進行說明。再者，於圖43中，區塊編號以16進制表示。

【0264】 不良資訊管理表33表示自開頭起(區塊編號由小至大)發現之不良區塊數之累計。於圖43中，示出不良資訊管理表33之區塊編號為4、9之區塊係不良區塊。於該情形時，於不良資訊管理表33中，於與區

塊編號0對應之位置、與區塊編號1對應之位置、與區塊編號2對應之位置、與區塊編號3對應之位置儲存“0”，於與區塊編號4對應之位置、與區塊編號5對應之位置、與區塊編號6對應之位置、與區塊編號7對應之位置、與區塊編號8對應之位置儲存“1”，於與區塊編號9對應之位置、與區塊編號A對應之位置、與區塊編號B對應之位置、與區塊編號C對應之位置、與區塊編號D對應之位置儲存“1”。即，表示儲存之數值發生變化之部分對應之區塊編號之區塊為不良區塊。

【0265】再者，於不良資訊管理表33之區塊編號E、及F對應之位置儲存不可能有之數值。若每個區塊之不良資訊之位元數為2位元，則於區塊編號E、及F對應之位置儲存“3”。

【0266】又，當執行位址轉換動作時，存取對象之區塊之區塊編號被轉換為該區塊之區塊編號與不良區塊數之累計之和。例如，關於區塊編號0～3，由於不良區塊之累計為“0”，故區塊編號不變。

【0267】另一方面，例如，於存取對象之區塊之區塊編號為4之情形時，將該區塊編號＋1而轉換為區塊編號“5”。因此，於寫入對象之區塊之區塊編號為4之情形時，資料實際上被寫入至區塊編號“5”之區塊。控制器4將區塊編號4作為資料寫入目標區塊之實體位址通知給主機2。於接收到將區塊編號4指定為讀取對象區塊之實體位址之讀取請求的情形時，由於該區塊編號4被轉換為區塊編號“5”，故控制器4自區塊編號“5”之區塊讀取資料。

【0268】同樣地，於存取對象之區塊之區塊編號為5之情形時，將該區塊編號＋1而轉換為區塊編號“6”。因此，於寫入對象之區塊之區塊編號為5之情形時，資料實際上被寫入至區塊編號“6”之區塊。控制器4將

區塊編號5作為資料寫入目標區塊之實體位址通知給主機2。於接收到將區塊編號5指定為讀取對象區塊之實體位址之讀取請求的情形時，由於該區塊編號5被轉換為區塊編號“6”，故控制器4自區塊編號“6”之區塊讀取資料。

【0269】 圖44表示主機2內之快閃轉譯部2A與快閃儲存裝置3內之不良區塊轉譯部24之關係。

【0270】 於主機2側，快閃轉譯部2A係於進行資料之寫入之情形時，將包含識別該資料之標籤(例如LBA)之寫入指令發送至快閃儲存裝置3。再者，於使用超級區塊用之API之情形時，快閃轉譯部2A將指定識別資料之標籤(例如LBA)與並行單位(Parallel Unit)之區塊位址之寫入指令(Write without Page Address)發送至快閃儲存裝置3。由於並行單位(Parallel Unit)藉由一個超級區塊而實現，故並行單位(Parallel Unit)之區塊位址係該超級區塊之超級區塊位址。

【0271】 快閃儲存裝置3側包括不良區塊轉譯部24、快閃I/O控制電路13、不良資訊管理表33、NAND型快閃記憶體晶片。又，於快閃儲存裝置3內，以不良資訊管理表33管理保存針對每一超級區塊表示可使用或不可使用之至少2位元之資訊的不良資訊。

【0272】 於自主機2側接收到包含作為用以識別應寫入之資料之資訊之標籤(例如LBA)之寫入請求的情形時，不良區塊轉譯部24係參照不良資訊，執行以寫入對象之超級區塊中包含之所有非不良區塊自寫入對象之超級區塊之開頭連續地邏輯性地配置之方式轉換存取對象區塊之位址的位址轉換，藉此，選擇寫入對象之超級區塊內之非不良區塊作為寫入目標區塊，並決定所選擇之區塊內之寫入目標位置(寫入目標頁、該頁內之偏

移)。

【0273】 再者，於自主機2接收到指定並行單位之區塊位址(超級區塊位址)之寫入指令的情形時，不良區塊轉譯部24參照不良資訊，選擇具有該指定之區塊位址(超級區塊位址)之超級區塊作為寫入對象之超級區塊。

【0274】 而且，不良區塊轉譯部24將指定寫入目標區塊所屬之晶片之晶片識別碼(晶片ID)、寫入目標區塊之區塊位址(原始區塊)、寫入目標頁(原始頁)、寫入目標頁內之偏移(Offset)的寫入指示發送至快閃I/O控制電路13。

【0275】 又，不良區塊轉譯部24將表示位址轉換前之存取對象區塊及寫入目標位置之兩者的實體位址、及標籤(例如LBA)通知給主機2。該實體位址亦可由表示寫入目標之超級區塊之並行單位(Parallel Unit)、及該超級區塊內之偏移(Offset)表示。表示寫入目標之超級區塊之並行單位(Parallel Unit)係寫入目標之超級區塊之超級區塊位址。

【0276】 接收到寫入指示之快閃I/O控制電路13基於寫入指示，將資料寫入至寫入目標位置。

【0277】 於主機2側，快閃轉譯部2A若接收實體位址(並行單位、偏移)、及標籤(例如LBA)，則更新主機2中管理之LUT，將所接收到之實體位址映射至所接收到之標籤(例如LBA)。

【0278】 快閃轉譯部2A係於進行讀取請求之情形時，將指定實體位址(並行單位、偏移)之讀取請求發送至快閃儲存裝置3。

【0279】 於快閃儲存裝置3側，自主機2接收到指定實體位址(並行單位、偏移)之讀取請求之情形時，不良區塊轉譯部24執行上述位址轉換，

將實體位址(並行單位、偏移)轉換為晶片識別碼(晶片ID)、區塊位址(原始區塊)、頁位址(原始頁)、頁內之偏移(Offset)。而且，不良區塊轉譯部24將指定晶片識別碼(晶片ID)、區塊位址(原始區塊)、頁位址(原始頁)、頁內之偏移(Offset)的寫入指示發送至快閃I/O控制電路13。快閃I/O控制電路13基於晶片識別碼(晶片ID)、區塊位址(原始區塊)、頁位址(原始頁)、頁內之偏移(Offset)，讀取資料。

【0280】 圖45表示主機2(計算機系統)之構成例。

【0281】 該主機2(計算機系統)包含處理器(CPU)101、主記憶體102、BIOS(Basic Input/Output System，基本輸入輸出系統)-ROM103、網路控制器105、周邊介面控制器106、控制器107、及嵌入式控制器(EC)108等。

【0282】 處理器101係以控制該計算機系統之各組件之動作之方式構成之CPU。該處理器101執行自複數個快閃儲存裝置3之任一個載入至主記憶體102之各種程式。主記憶體102包含DRAM之類之隨機存取記憶體。由處理器101執行之程式包含應用軟體層41、作業系統(OS)42、檔案系統43、驅動器44等。此處，檔案系統43中包含快閃儲存管理器45。再者，該快閃儲存管理器45亦可包含於驅動器44，而並非包含於檔案系統43。

【0283】 又，處理器101亦執行作為非揮發性記憶體之BIOS-ROM103中儲存之基本入輸出系統(BIOS)。BIOS係用於硬體控制之系統程式。

【0284】 網路控制器105係有線LAN(Local Area Network，區域網路)控制器、無線LAN控制器之類之通信裝置。周邊介面控制器106係以執

行與USB(Universal Serial Bus，通用串列匯流排)裝置之類之周邊裝置之通信之方式構成。

【0285】 控制器107係以執行與分別連接於複數個連接器107A之裝置之通信之方式構成。複數個快閃儲存裝置3亦可分別連接於複數個連接器107A。控制器107為SAS(Serial Attached SCSI(Small Computer System Interface，小型電腦系統介面)，串列連接SCSI)擴展器、PCIe交換器、PCIe擴展器、或RAID控制器等。

【0286】 EC108作為構成為執行計算機系統之電力管理之系統控制器發揮功能。EC108根據用戶對電源開關之操作而將計算機系統接通電源及斷開電源。EC108係以單晶片微控制器之類之處理電路之形式實現。

【0287】 快閃儲存管理器45係作為上述快閃轉譯部2A發揮功能之程式模組。快閃儲存管理器45具有管理各用戶位址(邏輯位址)與快閃儲存裝置3之各實體位址之間之映射之上述LUT。於使用LBA作為用戶位址(邏輯位址)之情形時，快閃儲存管理器45亦可設置於驅動器44內。

【0288】 快閃儲存管理器45將指定用以識別應寫入之資料(寫入資料)之用戶位址(邏輯位址)與該寫入資料之長度的寫入指令發送至快閃儲存裝置3。快閃儲存管理器45係自快閃儲存裝置3接收表示由快閃儲存裝置3自除不良區塊以外之區塊中選擇之寫入目標區塊及寫入目標區塊內之寫入目標位置之兩者的實體位址、及上述用戶位址(邏輯位址)，更新LUT而將該實體位址映射至該用戶位址(邏輯位址)。詳細而言，接收之實體位址表示由快閃儲存裝置3自寫入對象之一個並行單位中包含之不良區塊以外之區塊中選擇為用於寫入資料之寫入目標區塊的寫入目標區塊、及寫入了該寫入資料之寫入目標區塊內之實體記憶位置之兩者。

【0289】 又，快閃儲存管理器45參照LUT，獲取映射至讀取對象之資料對應之用戶位址(邏輯位址)之實體位址，並將指定該實體位址之讀取指令發送至快閃儲存裝置3。

【0290】 圖46表示內置快閃儲存裝置3之主機(計算機系統)2之構成例。

【0291】 該計算機系統具備能夠收容於機架之薄箱形之殼體201。亦可於殼體201內配置多個快閃儲存裝置3。於該情形時，各快閃儲存裝置3亦可能夠卸除地插入設置於殼體201之前表面201A之插槽中。

【0292】 系統板(母板)202配置於殼體201內。於系統板(母板)202上，安裝有包括CPU101、記憶體102、網路控制器105、控制器107之各種電子零件。該等電子零件作為主機2發揮功能。

【0293】 圖47之流程圖表示主機(計算機系統)2執行之寫入動作之順序。

【0294】 主機2之處理器101係藉由執行主記憶體102中儲存之電腦程式(此處為快閃儲存管理器45等)而執行以下順序。

【0295】 即，處理器101判定是否需要發送寫入指令(步驟S201)，若需要發送寫入指令(步驟S201之是)，則將指定用戶位址、快閃位址、長度之寫入指令發送至快閃儲存裝置3(步驟S202)。

【0296】 然後，處理器101若接收到來自快閃儲存裝置3之寫入動作完成之應答(包含用戶位址、快閃位址、長度)(步驟S203之是)，則更新LUT(步驟S204)。繼而，處理器101將指定與舊資料對應之實體位址之Trim指令發送至快閃儲存裝置3，對快閃儲存裝置3指示應使上述舊資料無效化(或應使表示參照舊資料之邏輯位址之數量之參照計數減小)(步驟

S205)。

【0297】 圖48之流程圖表示主機2執行之讀取動作之順序。

【0298】 處理器101判定是否需要發送讀取指令(步驟S301)，若需要發送讀取指令(步驟S301之是)，則將指定快閃位址、長度之讀取指令發送至快閃儲存裝置3(步驟S302)。

【0299】 其後，處理器101判定是否已接收到讀取資料(步驟S303)，若已接收到讀取資料(步驟S303之是)，則動作完成。

【0300】 圖49之流程圖表示主機2執行之參照計數遞增/遞減處理之順序。

【0301】 處理器101判定是否需要發送用以使舊資料之參照數增加之重複指令(步驟S401)，若需要發送重複指令(步驟S401之是)，則將重複指令發送至快閃儲存裝置3(步驟S402)。

【0302】 又，處理器101判定是否需要發送用以使舊資料之參照數減小之Trim指令(步驟S403)，若需要發送Trim指令(步驟S403之是)，則將Trim指令發送至快閃儲存裝置3(步驟S404)。

【0303】 如以上所說明般，根據本實施形態，主機(計算機系統)2將指定用以識別寫入資料之邏輯位址之寫入請求發送至儲存裝置3，並自儲存裝置3接收實體位址、及該邏輯位址，上述實體位址表示由儲存裝置3自除不良區塊以外之區塊中選擇為用於寫入資料之寫入目標區塊的寫入目標區塊、及寫入了寫入資料之寫入目標區塊內之實體記憶位置之兩者。而且，主機(計算機系統)2更新LUT，將所接收到之實體位址映射至所接收到之邏輯位址。因此，可利用主機(計算機系統)2之LUT管理直接表示不良區塊以外之實體記憶位置之實體位址，故而主機(計算機系統)2可視需

要將指定直接表示該實體記憶位置之實體位址的讀取請求發送至儲存裝置3。其結果，於儲存裝置3側，不將不良區塊以該不良區塊所屬之晶片內之其他區塊置換，便可正常地執行對於不良區塊以外之區塊之資料寫入及讀出動作。因此，即便不良區塊之數量增加，亦無須管理大量之置換資訊。又，亦不需要用於置換處理之位址轉換處理，因此可減少讀取潛伏時間。因此，可提昇系統整體之I/O性能。

【0304】 再者，於本實施形態中，例示出NAND型快閃記憶體作為非揮發性記憶體。但是，本實施形態之功能例如亦可應用於MRAM(Magnetoresistive Random Access Memory，磁阻式隨機存取記憶體)、PRAM(Phase change Random Access Memory，相變隨機存取記憶體)、ReRAM(Resistive Random Access Memory，可變電阻式隨機存取記憶體)、或FeRAM(Ferroelectric Random Access Memory，鐵電式隨機存取記憶體)之類之其他各種非揮發性記憶體。

【0305】 已對本發明之若干實施形態進行了說明，但該等實施形態係作為示例而提出，並不意圖限定發明之範圍。該等新穎之實施形態能夠以其他多種形態實施，可於不脫離發明主旨之範圍內進行各種省略、置換、變更。該等實施形態或其變化包含於發明之範圍或主旨中，並且包含於申請專利範圍所記載之發明及其均等之範圍內。

【0306】

[相關申請案]

本申請案享有以日本專利申請案2017-209344號(申請日：2017年10月30日)為基礎申請案之優先權。本申請案藉由參照該基礎申請案而包含基礎申請案之全部內容。

【符號說明】

【0307】

2	主機
2A	快閃轉譯部
2B	讀取剖析器
3	快閃儲存裝置
4	控制器
5	NAND型快閃記憶體
6	DRAM
10	匯流排
11	主機介面
12	CPU
13	快閃I/O控制電路
14	DRAM介面
21	寫入動作控制部
22	讀取動作控制部
23	GC動作控制部
24	不良區塊轉譯部
30	讀取緩衝器
31	寫入緩衝器
32	區塊管理表
33	不良資訊管理表
41	應用軟體層

42	作業系統
43	檔案系統
44	驅動器
45	快閃儲存管理器
50	介面
51	網路
61	終端用戶終端
80	區塊
81	區塊
82	區塊
83	區塊
84	區塊
85	區塊
86	區塊
87	區塊
88	區塊
89	區塊
90	區塊
91	區塊
101	處理器
102	主記憶體
103	BIOS-ROM
105	網路控制器

106	周邊介面控制器
107	控制器
107A	連接器
108	嵌入式控制器
201	殼體
201A	前表面
202	系統板
401	虛擬機服務
402	用戶應用
403	I/O服務
500	I/O佇列
601	寫入緩衝器
602	垃圾回收功能
603	NAND型快閃記憶體
BLK0～BLKm－1	區塊
BLK50	區塊
BLK100	區塊
BLK1000	區塊
BLK1948～BLK2047	區塊
Ch0～Ch15	通道
d1～d3	資料
P0～Pn－1	頁
SB	超級區塊

SB0	超級區塊
SB1	超級區塊
SB2	超級區塊
SB5	超級區塊
SB1000	超級區塊

201917580

【發明摘要】

【中文發明名稱】

計算機系統及儲存裝置之控制方法

【中文】

實施形態實現一種能夠改善I/O性能之計算機系統及儲存裝置之控制方法。

實施形態之計算機系統將指定用以識別應寫入之第1資料之第1邏輯位址與上述第1資料之長度的寫入請求發送至儲存裝置。計算機系統係自上述儲存裝置接收第1實體位址、及上述第1邏輯位址，上述第1實體位址表示由上述儲存裝置自除不良區塊以外之區塊中選擇為用於上述第1資料之寫入目標區塊之第1區塊、及寫入了上述第1資料之上述第1區塊內之第1實體記憶位置之兩者。計算機系統於管理各邏輯位址與上述儲存裝置之各實體位址之間之映射之第1表中將上述第1實體位址映射至上述第1邏輯位址。

【指定代表圖】

圖20

【代表圖之符號簡單說明】

2	主機
2A	快閃轉譯部
3	快閃儲存裝置
13	快閃I/O控制電路
21	寫入動作控制部
33	不良資訊管理表

【發明申請專利範圍】

【第1項】

一種計算機系統，其係對包括分別包含複數個區塊之複數個非揮發性記憶體晶片與控制器之儲存裝置進行控制者，且包括：

記憶體；及

處理器，其構成為與上述記憶體電性連接，且執行上述記憶體中儲存之電腦程式；

上述處理器係構成為：

將指定用以識別應寫入之第1資料之第1邏輯位址與上述第1資料之長度的寫入請求發送至上述儲存裝置，

自上述儲存裝置接收第1實體位址、及上述第1邏輯位址，上述第1實體位址係表示以下兩者：由上述儲存裝置自除不良區塊以外之區塊中選擇作為用於上述第1資料之寫入目標區塊之第1區塊及寫入了上述第1資料之上述第1區塊內之第1實體記憶位置，

於管理各邏輯位址與上述儲存裝置之各實體位址之間之映射之第1表中將上述第1實體位址映射至上述第1邏輯位址。

【第2項】

如請求項1之計算機系統，其中上述處理器係構成為：

參照上述第1表，獲取映射至上述第1邏輯位址之上述第1實體位址，

將指定上述第1實體位址之讀取請求發送至上述儲存裝置。

【第3項】

如請求項1之計算機系統，其中上述處理器係構成為：

於藉由上述儲存裝置將上述第1資料自舊實體記憶位置複製至新實體記憶位置之後，自上述儲存裝置接收上述第1邏輯位址、表示上述舊實體記憶位置之第2實體位址、及表示上述新實體記憶位置之第3實體位址，

參照上述第1表，獲取映射至上述第1邏輯位址之當前之實體位址，

於上述第2實體位址與上述獲取之當前之實體位址一致之情形時，於上述第1表中，將上述第3實體位址映射至上述第1邏輯位址，

於上述第2實體位址與上述獲取之當前之實體位址不一致之情形時，於上述第1表中維持上述當前之實體位址。

【第4項】

如請求項3之計算機系統，其中上述處理器係構成為：於上述第1表將上述第3實體位址映射至上述第1邏輯位址之後，將指定上述第2實體位址之第1指令發送至上述儲存裝置，藉此，對上述儲存裝置指示應使與上述第2實體位址對應之資料無效化。

【第5項】

如請求項1之計算機系統，其中上述處理器係構成為：於上述第1資料為已經寫入至上述儲存裝置之舊資料之更新資料之情形時，於上述第1表中將上述第1實體位址映射至上述第1邏輯位址之後，將指定與上述舊資料對應之實體位址之第1指令發送至上述儲存裝置，藉此，對上述儲存裝置指示應使上述舊資料無效化。

【第6項】

如請求項1之計算機系統，其中上述處理器係構成為：於上述第1資料為已經寫入至上述儲存裝置之舊資料之更新資料，且上述舊資料為被複數個邏輯位址參照之重複資料的情形時，於上述第1表中將上述第1實體位

址映射至上述第1邏輯位址之後，將指定與上述舊資料對應之實體位址之第1指令發送至上述儲存裝置，藉此，對上述儲存裝置指示應使表示參照上述舊資料之邏輯位址之數量之參照計數減小。

【第7項】

如請求項6之計算機系統，其中上述計算機系統不保存上述參照計數。

【第8項】

一種計算機系統，其係對儲存裝置進行控制者，上述儲存裝置包括複數個非揮發性記憶體晶片與控制器且構成為對分別包含屬於不同非揮發性記憶體晶片之各個區塊之複數個並行單位進行管理，該計算機系統包含：

記憶體；及

處理器，其構成為與上述記憶體電性連接，且執行上述記憶體中儲存之電腦程式；

上述處理器係構成為：

將指定用以識別應寫入之第1資料之第1邏輯位址與上述第1資料之長度的寫入請求發送至上述儲存裝置，

自上述儲存裝置接收第1實體位址、及上述第1邏輯位址，上述第1實體位址係表示以下兩者：由上述儲存裝置自寫入對象之一個並行單位中包含之不良區塊以外之區塊中選擇作為用於上述第1資料之寫入目標區塊的第1區塊、及寫入了上述第1資料之上述第1區塊內之第1實體記憶位置，

於管理各邏輯位址與上述儲存裝置之各實體位址之間之映射之第1表中將上述第1實體位址映射至上述第1邏輯位址。

【第9項】

如請求項8之計算機系統，其中上述寫入請求指定上述第1邏輯位址、及表示上述複數個並行單位內之一個並行單位之第1區塊位址，且

上述寫入對象之上述一個並行單位係具有上述指定之第1區塊位址之並行單位。

【第10項】

如請求項8之計算機系統，其中上述處理器係構成為：

參照上述第1表，獲取映射至上述第1邏輯位址之上述第1實體位址，

將指定上述第1實體位址之讀取請求發送至上述儲存裝置。

【第11項】

如請求項8之計算機系統，其中上述處理器係構成為：

於藉由上述儲存裝置將上述第1資料自舊實體記憶位置複製至新實體記憶位置之後，自上述儲存裝置接收上述第1邏輯位址、表示上述舊實體記憶位置之第2實體位址、及表示上述新實體記憶位置之第3實體位址，

參照上述第1表，獲取映射至上述第1邏輯位址之當前之實體位址，

於上述第2實體位址與上述獲取之當前之實體位址一致之情形時，於上述第1表中將上述第3實體位址映射至上述第1邏輯位址，

於上述第2實體位址與上述獲取之當前之實體位址不一致之情形時，於上述第1表中維持上述當前之實體位址。

【第12項】

如請求項11之計算機系統，其中上述處理器係構成為：於上述第1表中將上述第3實體位址映射至上述第1邏輯位址之後，將指定上述第2實體位址之第1指令發送至上述儲存裝置，藉此，對上述儲存裝置指示應使與

上述第2實體位址對應之資料無效化。

【第13項】

一種儲存裝置之控制方法，其係由對包括分別包含複數個區塊之複數個非揮發性記憶體晶片與控制器之儲存裝置進行控制的計算機系統執行，且包含以下步驟：

將指定用以識別應寫入之第1資料之第1邏輯位址與上述第1資料之長度的寫入請求發送至上述儲存裝置；

自上述儲存裝置接收第1實體位址、及上述第1邏輯位址，上述第1實體位址係表示以下兩者：由上述儲存裝置自除不良區塊以外之區塊中選擇以作為用於上述第1資料之寫入目標區塊之第1區塊、及寫入了上述第1資料之上述第1區塊內之第1實體記憶位置；及

於管理各邏輯位址與上述儲存裝置之各實體位址之間之映射之第1表中將上述第1實體位址映射至上述第1邏輯位址。

【第14項】

如請求項13之儲存裝置之控制方法，其進而包含以下步驟：

參照上述第1表，獲取映射至上述第1邏輯位址之上述第1實體位址；及

將指定上述第1實體位址之讀取請求發送至上述儲存裝置。

【第15項】

如請求項13之儲存裝置之控制方法，其進而包含以下步驟：

於藉由上述儲存裝置將上述第1資料自舊實體記憶位置複製至新實體記憶位置之後，自上述儲存裝置接收上述第1邏輯位址、表示上述舊實體記憶位置之第2實體位址、及表示上述新實體記憶位置之第3實體位址；

參照上述第1表，獲取映射至上述第1邏輯位址之當前之實體位址；

於上述第2實體位址與上述獲取之當前之實體位址一致之情形時，於上述第1表中將上述第3實體位址映射至上述第1邏輯位址；及

於上述第2實體位址與上述獲取之當前之實體位址不一致之情形時，於上述第1表中維持上述當前之實體位址。

【第16項】

如請求項15之儲存裝置之控制方法，其進而包含以下步驟：於上述第1表中將上述第3實體位址映射至上述第1邏輯位址之後，將指定上述第2實體位址之第1指令發送至上述儲存裝置，藉此，對上述儲存裝置指示應使與上述第2實體位址對應之資料無效化。