

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6536614号
(P6536614)

(45) 発行日 令和1年7月3日 (2019.7.3)

(24) 登録日 令和1年6月14日 (2019.6.14)

(51) Int.Cl.

F I

G O 6 F 12/00 (2006.01)

G O 6 F 3/06 (2006.01)

G O 6 F 3/08 (2006.01)

G O 6 F 12/00 5 5 0 Z

G O 6 F 3/06 3 0 5 Z

G O 6 F 3/06 3 0 4 N

G O 6 F 3/06 3 0 5 C

G O 6 F 3/08 H

請求項の数 10 (全 15 頁) 最終頁に続く

(21) 出願番号	特願2017-69250 (P2017-69250)	(73) 特許権者	000004237
(22) 出願日	平成29年3月30日 (2017.3.30)		日本電気株式会社
(65) 公開番号	特開2018-169980 (P2018-169980A)		東京都港区芝五丁目7番1号
(43) 公開日	平成30年11月1日 (2018.11.1)	(74) 代理人	100080816
審査請求日	平成30年7月5日 (2018.7.5)		弁理士 加藤 朝道
		(74) 代理人	100098648
			弁理士 内田 深人
		(74) 代理人	100119415
			弁理士 青木 充
		(74) 代理人	100168310
			弁理士 ▲高▼橋 幹夫
		(72) 発明者	菅沼 滋
			東京都港区芝五丁目7番1号 日本電気株式会社社内

最終頁に続く

(54) 【発明の名称】 ストレージ装置、ストレージ装置の制御方法及びプログラム

(57) 【特許請求の範囲】

【請求項1】

複数のSSD (Solid State Drive) それぞれにおける代替セクタに関する情報と、前記複数のSSDそれぞれに対するデータの書き込み回数と、に基づき各SSDの寿命値を推定する、寿命推定部と、

前記複数のSSDのうち前記推定された寿命値が所定の値よりも小さい第1のSSDに格納されたデータと、前記複数のSSDのうち前記第1のSSDとは異なる第2のSSDに格納されたデータと、を交換する、データ交換部と、

を備え、

前記データ交換部は、前記第1のSSDに格納されたデータを、前記複数のSSDのうち前記第1のSSDを除くSSDのそれぞれに格納した場合の各SSDにおける寿命短縮量を算出すると共に、前記寿命短縮量に基づき前記第2のSSDを選択する、ストレージ装置。

【請求項2】

前記寿命推定部は、

前記代替セクタに関する情報として、代替セクタの総容量と代替セクタの使用量を用いる、請求項1のストレージ装置。

【請求項3】

前記データ交換部は、

前記複数のSSDのうちいずれか1つのSSDに所定のサイズよりも大きいデータが書

き込まれる際に、前記第 1 の S S D と前記第 2 の S S D の間でデータを交換するか否かを判定する、請求項 1 又は 2 のストレージ装置。

【請求項 4】

前記データ交換部は、

前記寿命短縮量を算出した S S D のうち、前記寿命短縮量が最も小さい S S D を前記第 2 の S S D に選択する、請求項 1 乃至 3 のいずれか一項に記載のストレージ装置。

【請求項 5】

前記複数の S S D を用いて、R A I D (Redundant Arrays of Inexpensive Disks) を構成すべく、データ及び冗長符号を分散格納する機能を備える、請求項 1 乃至 4 のいずれか一項に記載のストレージ装置。

10

【請求項 6】

前記データ交換部は、前記第 1 及び第 2 の S S D 間でデータ交換する際のデータの入れ替え単位を、R A I D を構成するストライピンググループのサイズを基準とする、請求項 5 のストレージ装置。

【請求項 7】

前記データ交換部は、

前記第 1 及び第 2 の S S D 間でデータ交換する際、R A I D を構成する情報を書き換える、請求項 6 のストレージ装置。

【請求項 8】

前記複数の S S D それぞれに対応し、対応する S S D に対するデータの累計書き込み回数を計測する、複数の書き込み回数カウンタをさらに備える、請求項 1 乃至 7 のいずれか一項に記載のストレージ装置。

20

【請求項 9】

複数の S S D (Solid State Drive) それぞれにおける代替セクタに関する情報と、前記複数の S S D それぞれに対するデータの書き込み回数と、に基づき各 S S D の寿命値を推定するステップと、

前記複数の S S D のうち前記推定された寿命値が所定の値よりも小さい第 1 の S S D に格納されたデータと、前記複数の S S D のうち前記第 1 の S S D とは異なる第 2 の S S D に格納されたデータと、を交換するステップと、

を含み、

30

前記データを交換するステップは、

前記第 1 の S S D に格納されたデータを、前記複数の S S D のうち前記第 1 の S S D を除く S S D のそれぞれに格納した場合の各 S S D における寿命短縮量を算出すると共に、前記寿命短縮量に基づき前記第 2 の S S D を選択する、ストレージ装置の制御方法。

【請求項 10】

ストレージ装置に搭載されたコンピュータに実行させるプログラムであって、

複数の S S D (Solid State Drive) それぞれにおける代替セクタに関する情報と、前記複数の S S D それぞれに対するデータの書き込み回数と、に基づき各 S S D の寿命値を推定する処理と、

前記複数の S S D のうち前記推定された寿命値が所定の値よりも小さい第 1 の S S D に格納されたデータと、前記複数の S S D のうち前記第 1 の S S D とは異なる第 2 の S S D に格納されたデータと、を交換する処理と、

40

を前記コンピュータに実行させ、

前記データを交換する処理は、

前記第 1 の S S D に格納されたデータを、前記複数の S S D のうち前記第 1 の S S D を除く S S D のそれぞれに格納した場合の各 S S D における寿命短縮量を算出すると共に、前記寿命短縮量に基づき前記第 2 の S S D を選択する、プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

50

本発明は、ストレージ装置、ストレージ装置の制御方法及びプログラムに関する。特に、複数の記憶媒体を有するストレージ装置、ストレージ装置の制御方法及びプログラムに関する。

【背景技術】

【0002】

近年、SSD (Solid State Drive) と称される記憶媒体が広く用いられている (特許文献1~4)。

【0003】

特許文献1は、複数のSSDそれぞれにおける書き込み回数を監視し、書き込み回数がSSD間で極端に偏る場合、データスワップを行い、書き込み回数の平準化を図る技術を開示している。

10

【0004】

特許文献2は、SSDへの書き込み回数が予め定めた回数よりも多い場合、データの一部をスペアディスク (スペアの記憶デバイス) に退避させる技術を開示している。

【0005】

特許文献3は、SSDに書き込み領域を割り当てる際、SSDの特性がSLC (Single-Level-Cell) かMLC (Multi-Level-Cell) であるかを考慮し、SSDに対する書き込み頻度の高い場合はSLCを割り当て、書き込み頻度が低い場合にはMLCを割り当てる技術を開示している。

【0006】

20

特許文献4は、複数の記憶媒体間の劣化を平準化するための技術を開示する。特許文献4には、第1論理記憶領域に割り当てられた記憶デバイスと異なる記憶デバイスを割り当てられ第1論理記憶領域の推定データより小さい推定データ量を有する第2論理記憶領域を選択し、第1論理記憶領域に格納された第1データを第2論理記憶領域に移動させる、と記載されている。

【0007】

特許文献5は、コンパクトフラッシュ (登録商標) に関する技術を開示している。特許文献5は、代替セクタ数が閾値を超える場合は商品販売データをメモリに記憶させ、商品販売データの登録終了に回答して、メモリに記憶した複数の商品販売データを一括してコンパクトフラッシュ (登録商標) カードに書込む技術を開示する。

30

【先行技術文献】

【特許文献】

【0008】

【特許文献1】特許第5242264号公報

【特許文献2】特開2010-55247号公報

【特許文献3】特許第5192352号公報

【特許文献4】国際公開第2014/141411号

【特許文献5】特開2006-155532号公報

【発明の概要】

【発明が解決しようとする課題】

40

【0009】

なお、上記先行技術文献の各開示を、本書に引用をもって繰り込むものとする。以下の分析は、本発明者らによってなされたものである。

【0010】

SSDへの書き込み回数は、SSDの寿命に影響を与える。そのため、上述のように、SSDの書き込みに関する種々の技術が存在する。しかし、SSDの寿命は書き込み回数にだけ依存して定まるものではなく、SSDへのデータ書き込みを抑制したとしてもSSDの寿命を保証することはできない。

【0011】

その理由は、実際のSSDが寿命を迎えるのは、代替セクタの枯渇が主要因として考え

50

られるからである。つまり、SSDへの書き込み回数はSSDの寿命を推測する方法の1つであって、代替セクタの存在を考慮していなければ不十分である。例えば、特許文献1には、代替セクタの使用量を観測する言及がない。

【0012】

また、SSDの内部制御回路は、SSDの読み書きが失敗した場合、代替セクタへの切り替えによって状態回復を図る。SSDの読み書き失敗は、書き込み回数によって確率的に発生し得るが、その発生頻度は、周辺環境やSSD素体の特性によって変化する。例えば、周辺温度が高い場合、書き込みエラーが発生する可能性が高くなる。また、個々のNANDフラッシュの特性がSLCかMLCかによって、エラー耐性が異なることも考慮すべきである。

10

【0013】

特許文献3には、記憶素子の特性に応じて（SLCかMLCかに応じて）、書き換え頻度の高い領域を割り当てるか、低い領域を割り当てるかを選択する方法が開示されている。しかし、当該文献に開示された技術においても、実際の代替セクタの使用量の多寡は検証されていない。そのため、書き込み回数が少なくても、代替セクタの使用量が多いSSDが存在する場合、書き込み時に代替セクタを消費する可能性が高くなる。その結果、SSD間のデータスワップ時の書き込みで、かえってSSD寿命が短くなる可能性もある。

【0014】

本発明は、複数のSSDからなるストレージシステム全体の寿命を長くする、ストレージ装置、ストレージ装置の制御方法及びプログラムを提供することを目的とする。

20

【課題を解決するための手段】

【0015】

本発明の第1の視点によれば、複数のSSD（Solid State Drive）それぞれにおける代替セクタに関する情報と、前記複数のSSDそれぞれに対するデータの書き込み回数と、に基づき各SSDの寿命値を推定する、寿命推定部と、前記複数のSSDのうち前記推定された寿命値が所定の値よりも小さい第1のSSDに格納されたデータと、前記複数のSSDのうち前記第1のSSDとは異なる第2のSSDに格納されたデータと、を交換する、データ交換部と、を備え、前記データ交換部は、前記第1のSSDに格納されたデータを、前記複数のSSDのうち前記第1のSSDを除くSSDのそれぞれに格納した場合の各SSDにおける寿命短縮量を算出すると共に、前記寿命短縮量に基づき前記第2のSSDを選択する、ストレージ装置が提供される。

30

【0016】

本発明の第2の視点によれば、複数のSSD（Solid State Drive）それぞれにおける代替セクタに関する情報と、前記複数のSSDそれぞれに対するデータの書き込み回数と、に基づき各SSDの寿命値を推定するステップと、前記複数のSSDのうち前記推定された寿命値が所定の値よりも小さい第1のSSDに格納されたデータと、前記複数のSSDのうち前記第1のSSDとは異なる第2のSSDに格納されたデータと、を交換するステップと、を含み、前記データを交換するステップは、前記第1のSSDに格納されたデータを、前記複数のSSDのうち前記第1のSSDを除くSSDのそれぞれに格納した場合の各SSDにおける寿命短縮量を算出すると共に、前記寿命短縮量に基づき前記第2のSSDを選択する、ストレージ装置の制御方法が提供される。

40

【0017】

本発明の第3の視点によれば、ストレージ装置に搭載されたコンピュータに実行させるプログラムであって、複数のSSD（Solid State Drive）それぞれにおける代替セクタに関する情報と、前記複数のSSDそれぞれに対するデータの書き込み回数と、に基づき各SSDの寿命値を推定する処理と、前記複数のSSDのうち前記推定された寿命値が所定の値よりも小さい第1のSSDに格納されたデータと、前記複数のSSDのうち前記第1のSSDとは異なる第2のSSDに格納されたデータと、を交換する処理と、を前記コンピュータに実行させ、前記データを交換する処理は、前記第1のSSDに格納されたデータを、前記複数のSSDのうち前記第1のSSDを除くSSDのそれぞれに格納した場

50

合の各ＳＳＤにおける寿命短縮量を算出すると共に、前記寿命短縮量に基づき前記第２のＳＳＤを選択する、プログラムが提供される。

なお、このプログラムは、コンピュータが読み取り可能な記憶媒体に記録することができる。記憶媒体は、半導体メモリ、ハードディスク、磁気記録媒体、光記録媒体等の非ランジェント（non-transient）なものとすることができる。本発明は、コンピュータプログラム製品として具現することも可能である。

【発明の効果】

【００１８】

本発明の各視点によれば、複数のＳＳＤからなるストレージシステム全体の寿命を長くすることに寄与する、ストレージ装置、ストレージ装置の制御方法及びプログラム、が提供される。

10

【図面の簡単な説明】

【００１９】

【図１】一実施形態の概要を説明するための図である。

【図２】第１の実施形態に係るストレージ装置の概略構成の一例を示す図である。

【図３】第１の実施形態に係るデバイスコントローラの処理構成の一例を示す図である。

【図４】第１の実施形態に係るストレージ装置の動作を説明するための図である。

【図５】第１の実施形態に係るストレージ装置の概略構成の他の一例を示す図である。

【発明を実施するための形態】

【００２０】

20

初めに、一実施形態の概要について説明する。なお、この概要に付記した図面参照符号は、理解を助けるための一例として各要素に便宜上付記したものであり、この概要の記載はなんらの限定を意図するものではない。また、各図におけるブロック間の接続線は、双方向及び単方向の双方を含む。一方向矢印については、主たる信号（データ）の流れを模式的に示すものであり、双方向性を排除するものではない。

【００２１】

一実施形態に係るストレージ装置１０は、寿命推定部２０１と、データ交換部２０２と、を備える（図１参照）。寿命推定部２０１は、複数のＳＳＤ（Solid State Drive）それぞれにおける代替セクタに関する情報と、複数のＳＳＤそれぞれに対するデータの書き込み回数と、に基づき各ＳＳＤの寿命値を推定する。データ交換部２０２は、複数のＳＳＤのうち推定された寿命値が所定の値よりも小さい第１のＳＳＤに格納されたデータと、複数のＳＳＤのうち第１のＳＳＤとは異なる第２のＳＳＤに格納されたデータと、を交換する。さらに、データ交換部２０２は、第１のＳＳＤに格納されたデータを、複数のＳＳＤのうち第１のＳＳＤを除くＳＳＤのそれぞれに格納した場合の各ＳＳＤにおける寿命短縮量を算出すると共に、寿命短縮量に基づき前記第２のＳＳＤを選択する。

30

【００２２】

上述のように、ＳＳＤの寿命は代替セクタの枯渇によるものと考えられるが、ＳＳＤの寿命を延ばす（ＳＳＤを延命）ためには書き込み回数の少ない領域への割り当ても必要になる。そこで、ストレージ装置１０は、ＳＳＤの代替セクタに関する情報（例えば、代替セクタの使用量）と書き込み回数を共に観測し、これらの情報を用いた演算結果から、ＳＳＤの寿命低下を予測する。ストレージ装置１０は、特定のＳＳＤの寿命が短くなっていると判断した場合には、他の寿命が長いＳＳＤとの間でデータスワップを行う。その結果、ＳＳＤを搭載したストレージシステム全体の寿命が長くなる。また、ストレージシステム全体の寿命が長くなることで、システム全体の信頼性を向上させることができる。

40

【００２３】

以下に具体的な実施の形態について、図面を参照してさらに詳しく説明する。なお、各実施形態において同一構成要素には同一の符号を付し、その説明を省略する。

【００２４】

〔第１の実施形態〕

第１の実施形態について、図面を用いてより詳細に説明する。

50

【 0 0 2 5 】

図 2 は、第 1 の実施形態に係るストレージ装置 1 の概略構成の一例を示す図である。図 2 を参照すると、ストレージ装置 1 は、ストレージコントローラ 1 1 と、複数の S S D 1 2 ~ 1 5 と、を含んで構成される。

【 0 0 2 6 】

ストレージコントローラ 1 1 は、ストレージ装置 1 の全体を制御する手段である。S S D 1 2 ~ 1 5 のそれぞれは、データを保持する記憶デバイスである。ストレージコントローラ 1 1 は、S S D 1 2 ~ 1 5 を制御するデバイスコントローラ 1 1 1 を備えている（デバイスコントローラ 1 1 1 が実装されている）。

【 0 0 2 7 】

デバイスコントローラ 1 1 1 は、各 S S D に対応するデバイスインターフェイス 1 1 2 ~ 1 1 5 を介して、S S D 1 2 ~ 1 5 に接続されている。また、ストレージコントローラ 1 1 は、デバイスインターフェイス 1 1 2 ~ 1 1 5 及び S S D 1 2 ~ 1 5 に対応する書き込み回数カウンタ 1 1 2 1 ~ 1 1 5 1 を備えている。

【 0 0 2 8 】

書き込み回数カウンタ 1 1 2 1 ~ 1 1 5 1 は、対応する S S D 1 2 ~ 1 5 に対する書き込み回数を監視する手段である。より具体的には、書き込み回数カウンタ 1 1 2 1 ~ 1 1 5 1 は、複数の S S D 1 2 ~ 1 5 それぞれに対応して設置され、対応する S S D へのデータの累計書き込み回数を計測する。

【 0 0 2 9 】

また、S S D 1 2 ~ 1 5 のそれぞれには、代替セクタ 1 2 1 ~ 1 5 1 が実装されている。代替セクタ 1 2 1 ~ 1 5 1 は、必要に応じて消費される。図 2 では、灰色で塗りつぶされた領域が代替セクタ使用領域 1 2 1 1 ~ 1 5 1 1 であって、代替セクタの消費分である。

【 0 0 3 0 】

各 S S D 1 2 ~ 1 5 の記憶領域は、S S D 1 2 ~ 1 5 の間で R A I D (Redundant Arrays of Inexpensive Disks) 構成されている。つまり、ストレージ装置 1 は、複数の S S D 1 2 ~ 1 5 を用いて、R A I D を構成すべく（R A I D を構成するための）、データおよび冗長符号を分散格納する機能を備える。また、ストライピンググループ 2 1 は、データ領域 2 1 2 ~ 2 1 5 で構成される。

【 0 0 3 1 】

なお、図 2 では、ストレージ装置 1 と接続されるホストや当該ホストとのインターフェイスの図示は省略している。

【 0 0 3 2 】

図 3 は、デバイスコントローラ 1 1 1 の処理構成の一例を示す図である。図 3 を参照すると、デバイスコントローラ 1 1 1 は、寿命推定部 1 0 1 と、データ交換部 1 0 2 と、を含んで構成される。なお、デバイスコントローラ 1 1 1 は、ストレージ装置 1 に搭載されたコンピュータに、上記寿命推定部 1 0 1、データ交換部 1 0 2 を含む処理モジュールを実行させるプログラムにより実現することができる。

【 0 0 3 3 】

寿命推定部 1 0 1 は、複数の S S D 1 2 ~ 1 5 それぞれにおける代替セクタに関する情報と、複数の S S D 1 2 ~ 1 5 それぞれに対するデータの書き込み回数と、に基づき各 S S D の寿命値を推定する手段である。寿命推定部 1 0 1 は、各 S S D における代替セクタの消費量と書き込み回数の組み合わせに基づき、各 S S D に関する寿命パラメータを作成する。

【 0 0 3 4 】

データ交換部 1 0 2 は、複数の S S D 1 2 ~ 1 5 のうち推定された寿命値（寿命パラメータ）が所定の値よりも小さい第 1 の S S D に格納されたデータと、複数の S S D 1 2 ~ 1 5 のうち第 1 の S S D とは異なる第 2 の S S D に格納されたデータと、を交換する手段である。データ交換の際、データ交換部 1 0 2 は、第 1 の S S D に格納されたデータを、

10

20

30

40

50

複数のSSD 12～15のうち第1のSSDを除くSSDのそれぞれに格納した場合の各SSDにおける寿命短縮量を算出すると共に、当該算出された寿命短縮量に基づき第2のSSDを選択する。

【0035】

つまり、データ交換部102は、寿命悪化が見込まれるSSDから余命が長いと推定されるSSDに書き込み処理をスワップする(SSD間でデータを移動する)。このようなデバイスコントローラ111の処理により、SSD間の見込み寿命が平準化され、ストレージシステム全体の寿命を長くすることができる。

【0036】

なお、データ交換部102は、データスワップをRAIDのストライピングごとに行う。より具体的には、データ交換部102は、上述の第1及び第2のSSD間でデータ交換する際のデータの入れ替え単位を、RAIDを構成するストライピンググループのサイズを基準とする。例えば、データ交換部102は、ブロック単位でSSD間でデータを交換し、1つのブロックよりも小さいデータ同士をSSD間で交換することはない。例えば、図2の例では、同じストライピンググループ21に属するデータ領域212～215の間でデータスワップ(SSD間のデータの移動)が行われる。

【0037】

また、データ交換部102は、SSD間でデータ交換する際、RAIDを構成する情報を書き換える。具体的には、データ交換部102は、複数のSSD 12～15を連携する情報のうち、データとその格納先のSSDを示す情報をSSD間のデータ移動の際に更新する。その結果、ホストに対してはSSD間のデータの移動は隠蔽される。換言するならば、SSD間のデータの移動はストレージコントローラ11(デバイスコントローラ111)側により隠ぺいされ、ホストが意識することはない。

【0038】

寿命推定部101は、以下の式(1)により、各SSDの寿命パラメータを作成する。

$$Lt_SSD_x = \alpha - \beta * Wcount_x - \gamma * Ssector_x * CapAll / (SsecCap * Wcount_x) \quad \cdots (1)$$

【0039】

なお、式(1)において、サフィックスxはSSDの番号(識別子)を示す。Lt_SSD_xは、各SSDの寿命の推測値を示す。Wcount_xは、各SSDへの書き込み回数を示す。Ssector_xは、代替セクタの使用量を示す。つまり、Ssector_xは、SSDでのエラー(代替)が発生したセクタ数と等価である。CapAllは、代替セクタを除くSSDの総記憶容量を示す。SsecCapは、各SSDの代替セクタ総容量を示す。α、β及びγはSSDの種類ごとのパラメータを示す。より具体的には、αはSSD寿命の基準値、βは書き込み回数による劣化係数、γは代替セクタ使用量での寿命の短縮係数である。

【0040】

上述の式(1)において、右辺の第二項は、書き込み回数による寿命の減少量を示す。また、右辺の第三項は、代替セクタの使用量が代替セクタの総容量に比べて大きい場合、及び、代替セクタの総容量が総記憶容量に比べて小さい場合、各SSDの寿命が短くなっていることを示す。

【0041】

寿命推定部101は、上記の式(1)を用いて各SSDに関する寿命の推定値(寿命パラメータ)を算出する。なお、上記の式(1)の計算に必要なパラメータのうち、Wcount_xは書き込み回数カウンタ1121～1151により得られる。また、Ssector_xは、代替セクタの使用量であるから、デバイスコントローラ111自身(寿命推定部101自身)が把握している情報である。他の情報、例えば、CapAll、SsecCap、α、β及びγ、は各SSDの仕様により定まる値である。従って、事前にこれらの情報をデバイスコントローラ111に入力しておけばよい。

【0042】

データ交換部102は、寿命推定部101により算出された寿命の推定値Lt_SSD

x に対して閾値処理を行い、当該推定値が所定の閾値よりも小さい場合には、データスワップを行う。あるいは、データ交換部 102 は、推定した寿命値が他の SSD と比較して小さい場合には余命が短いと判断し、データスワップを行うことで寿命悪化を回避してもよい。

【0043】

なお、データ交換部 102 によるデータスワップを行うか否かを判断は、複数の SSD のうちいずれか 1 つの SSD に所定のサイズよりも大きいデータが書き込まれるタイミングとする。例えば、1 ブロックよりも大きいデータが SSD に書き込まれる際に、上記データスワップの実行要否が判定される。あるいは、1 ブロックよりも大きいデータが SSD に書き込まれる際に、寿命推定部 101 により各 SSD の寿命が推定され、その後、データ交換部 102 によりデータスワップの実行要否が判定されてもよい。

10

【0044】

データ交換部 102 は、スワップする領域サイズから、寿命の変更差分を求める。つまり、データ交換部 102 は、データスワップした後の各 SSD の寿命を予測する。スワップ後の予測寿命を Lt_SSD_x' と表記すると、当該予測寿命 Lt_SSD_x' は下記の式 (2) により算出される。

【0045】

$$Lt_SSD_x' = \alpha - \beta * (Wcount_x + SSize_x / Block) - \gamma * Ssector_x * CapAll * (Wcount_x + SSize_x / Block) / (SsecCap * Wcount_x^2) \quad \dots (2)$$

【0046】

20

上記式 (2) において、 $SSize_x$ は、SSD の領域のうち更新頻度の高い領域でのサイズを示す。 $SSize_x$ は、データスワップで置き換えられるデータのサイズである。また、 $Block$ は SSD の種類ごとの書き込み更新単位を示す。 $Block$ は、SSD の最小書き換え単位とする。つまり、データ交換部 102 は、SSD 間でデータ交換する際のデータの入れ替え単位を、RAID を構成するストライピンググループのサイズ ($Block$) を基準とする。換言すれば、ブロック単位でデータが入れ替えられ、1 つのブロックよりも小さいサイズのデータ同士を入れ替えることはない。なお、 $SSize_x / Block$ を計算することで、データスワップ時の書き換え回数が求まる。

【0047】

ここで、代替セクタの消費量 $Ssector_x$ も、書き換え処理によって増加すると予測される。当該増加量は、書き換え回数 $Wcount_x$ の増加量の形で示される。データスワップ後の書き換え回数を $Wcount_x'$ とすると、当該書き換え回数は下記の式 (3) により求まる。

30

$$Wcount_x' = Wcount_x + SSize_x / Block \quad \dots (3)$$

【0048】

代替セクタの増加量は、元の書き換え回数との比率であらわされるから、増加比率 δ は以下の式 (4) により求まる。

$$\delta = Wcount_x' / Wcount_x = (Wcount_x + SSize_x / Block) / Wcount_x \quad \dots (4)$$

【0049】

式 (1) における右辺の第二項の書き換え回数を式 (3) における $Wcount_x'$ で置き換え、式 (1) における右辺の第三項に式 (4) の増加比率 δ を乗じて整理すると、式 (2) が得られる。但し、厳密には、右辺の第三項の分母を書き換え回数 $Wcount_x'$ により置き換えるのが好ましいが、代替セクタの消費量が少なく見積もられるため、式 (1) における右辺の第三項の分母は不変としている。

40

【0050】

データスワップの実行が決定されると、データ交換部 102 は、式 (1) と式 (2) を用いて、データ移動先の候補となる各 SSD の寿命の短縮量を計算する。

【0051】

ここで、SSD の寿命の短縮量が小さいということは、元の SSD の寿命が長く、寿命の変化量が小さいということと等価である。そこで、データ交換部 102 は、下記の式 (

50

5) に示すように、元の寿命の推定値 Lt_SSD_x よりデータスワップ後の寿命 Lt_SSD_x' を減算することで、寿命の短縮量を算出する。

【0052】

$Lt_SSD_x - Lt_SSD_x'$

$= (\beta + \gamma * Ssector_x * CapAll / (SsecCap * Wcount_x^2)) * SSize_x / Block) \dots (5)$

【0053】

上記式(5)の値が寿命の短縮量であるから、データ交換部は、上記値が最も小さくなるSSDをスワップ対象として選択する。

【0054】

次に、図4を参照しつつ、第1の実施形態に係るストレージ装置1の動作を説明する。図4において、各SSD12～15のそれぞれの代替セクタ121～151は、灰色領域の容量が消費されているものとする(代替セクタ121～151は、代替セクタ使用領域1211～1511の大きさが使用されている)。また、各SSD12～15に対する書き込み回数(書き込み回数カウンタ1121～1151が保持するカウント値)は、N1、N2、N3、N4とする。

【0055】

上記書き込み回数の大小関係は、下記の式(6)のとおりとする。

【0056】

$N4 > N3 > N2 > N1 \dots (6)$

【0057】

デバイスコントローラ111は、各SSDの代替セクタ使用領域1211～1511のサイズと各書き込み回数カウンタ1121～1151の値により、各SSDの予測寿命を算出する。具体的には、デバイスコントローラ111は、SSD12～15のそれぞれについて式(1)を適用し、下記の式(7)～(10)により各SSDの予測寿命を算出する。

【0058】

$Lt_SSD_{12} = \alpha - \beta * N1 - \gamma * Ssector_{12} * CapAll / (SsecCap * N1) \dots (7)$

$Lt_SSD_{13} = \alpha - \beta * N2 - \gamma * Ssector_{13} * CapAll / (SsecCap * N2) \dots (8)$

$Lt_SSD_{14} = \alpha - \beta * N3 - \gamma * Ssector_{14} * CapAll / (SsecCap * N3) \dots (9)$

$Lt_SSD_{15} = \alpha - \beta * N4 - \gamma * Ssector_{15} * CapAll / (SsecCap * N4) \dots (10)$

【0059】

なお、SSDがすべて同一種類であるとする、 α 、 β 及び γ は固定値となる。また、上記の式(7)～(10)では、各SSDを示す符号をサフィックスxとしている。

【0060】

上記(7)～(10)による計算の結果、SSD12の予測寿命値 Lt_SSD_{12} が最も小さく、かつ、当該予測寿命値が所定の閾値よりも小さい場合、デバイスコントローラ111は、SSD12に格納されたデータのスワップを決定する。

【0061】

デバイスコントローラ111は、SSD12のデータをスワップする際の相手先となるSSDを下記の式(11)～(13)により求める。

【0062】

$Lt_SSD_{13} - Lt_SSD_{13}'$

$= (\beta + \gamma * Ssector_{13} * CapAll / (SsecCap * N2^2)) * SSize_x / Block) \dots (11)$

$Lt_SSD_{14} - Lt_SSD_{14}'$

$= (\beta + \gamma * Ssector_{14} * CapAll / (SsecCap * N3^2)) * SSize_x / Block) \dots (12)$

$Lt_SSD_{15} - Lt_SSD_{15}'$

$= (\beta + \gamma * Ssector_{15} * CapAll / (SsecCap * N4^2)) * SSize_x / Block) \dots (13)$

【0063】

式(11)～(13)を確認すると、データスワップ後のSSDの寿命の短縮量は、代替セクタの消費量が多い物ほど大きくなる。具体的には、式(11)により計算されるS

10

20

30

40

50

SSD13の寿命短縮量が最も小さく、式(12)により計算されるSSD14の寿命短縮量が最も大きくなる。

【0064】

デバイスコントローラ111は、SSDの寿命短縮量が最も小さいSSDをデータスワップの相手先として選出する。上記の例ではSSD13がデータスワップの相手となる。

【0065】

このように、デバイスコントローラ111は、データスワップの相手(データの移動先)とするSSDは、書き込み回数に加え、代替セクタの消費量を少ないものを選出する。図4の例では、デバイスコントローラ111は、SSD12の寿命が短くなっていると判断した場合には、SSD12とSSD13との間でデータスワップを実施する。

10

【0066】

上記実施形態にて説明したストレージ装置1の構成や動作は例示であって、システムの構成を限定する趣旨ではない。例えば、各SSDがRAIDを構成しておらず、図4のデータ領域212~215が互いに独立していてもよい(図5参照)。このような場合であっても、データ領域の書き込み単位が定まっていれば、データスワップは可能である。

【0067】

以上のように、第1の実施形態に係るストレージ装置1では、SSDの代替セクタの使用量を監視することで、実際のSSDの寿命をより正確に予測する。代替セクタの枯渇がSSDの寿命そのものに影響するため、書き込み回数の観測より直接的にSSDの状態を確認できるためである。

20

【0068】

上記の説明により、本発明の産業上の利用可能性は明らかであるが、本発明は、複数の記憶媒体を含むディスクアレイ装置等に好適に適用可能である。ディスクアレイ装置に適用することで、当該装置の長寿命化が実現できる。即ち、SSD単体の寿命は、SSD内部でのウェアレベリングや、代替セクタの大容量化などで増加傾向にあり、これに装置全体での延命効果を付加すれば、SSDの寿命を装置寿命程度まで伸ばすことが可能となり、SSDの交換を減らすことが可能になる。

【0069】

上記の実施形態の一部又は全部は、以下の付記のようにも記載され得るが、以下には限られない。

30

[付記1]

上述の第1の視点に係るストレージ装置のとおりである。

[付記2]

前記寿命推定部は、

前記代替セクタに関する情報として、代替セクタの総容量と代替セクタの使用量を用いる、付記1のストレージ装置。

[付記3]

前記データ交換部は、

前記複数のSSDのうちいずれか1つのSSDに所定のサイズよりも大きいデータが書き込まれる際に、前記第1のSSDと前記第2のSSDの間でデータを交換するか否かを判定する、付記1又は2のストレージ装置。

40

[付記4]

前記データ交換部は、

前記寿命短縮量を算出したSSDのうち、前記寿命短縮量が最も小さいSSDを前記第2のSSDに選択する、付記1乃至3のいずれかーに記載のストレージ装置。

[付記5]

前記複数のSSDを用いて、RAID(Redundant Arrays of Inexpensive Disks)を構成すべく、データ及び冗長符号を分散格納する機能を備える、付記1乃至4のいずれかーに記載のストレージ装置。

[付記6]

50

前記データ交換部は、前記第 1 及び第 2 の S S D 間でデータ交換する際のデータの入れ替え単位を、R A I D を構成するストライピンググループのサイズを基準とする、付記 5 のストレージ装置。

[付記 7]

前記データ交換部は、

前記第 1 及び第 2 の S S D 間でデータ交換する際、R A I D を構成する情報を書き換える、付記 6 のストレージ装置。

[付記 8]

前記複数の S S D それぞれに対応し、対応する S S D に対するデータの累計書き込み回数を計測する、複数の書き込み回数カウンタをさらに備える、付記 1 乃至 7 のいずれかに記載のストレージ装置。

10

[付記 9]

上述の第 2 の視点に係るストレージ装置の制御方法のとおりである。

[付記 1 0]

上述の第 3 の視点に係るプログラムのとおりである。

[付記 1 1]

複数の S S D を内蔵するストレージ装置であって、

当該 S S D 内部の代替セクタの総容量と、使用済み代替セクタの容量を確認する手段を有し、また、当該 S S D 毎の累計の書き込み回数を計測する機能を有し、

当該 S S D 毎の前記代替セクタ総容量と使用済み代替セクタ容量、及び累計書き込み回数を元に当該 S S D 毎の寿命を予測し、

20

かつ、

当該 S S D 毎に、ストレージ装置内で定められた、ある一定容量のデータ単位を書き込んだ場合、各々の S S D の寿命の短縮量を予測し、

前記書き込み後の予測寿命の短縮量が最も小さい S S D を選択する機能を有し、

最も予測寿命の短い S S D が、ストレージ装置内で定められた閾値より予測寿命が短くなった場合、前記最も予測寿命の短い S S D のデータと、前記予測寿命の短縮量が最も小さい S S D のデータとを入れ替える機能を有することを特徴とする、ストレージ装置。

[付記 1 2]

複数の S S D を内蔵するストレージ装置であって、

30

当該 S S D 内部の代替セクタの総容量と、使用済み代替セクタの容量を確認する手段を有し、

また、当該 S S D 毎の累計の書き込み回数を計測する機能を有し、

当該 S S D 毎の前記代替セクタ総容量と使用済み代替セクタ容量、及び累計書き込み回数を元に当該 S S D 毎の寿命を予測し、

かつ、

当該 S S D 毎に、ストレージ装置内で定められた、ある一定容量のデータ単位を書き込んだ場合、各々の S S D の寿命の変化を予測し、

前記書き込み後の予測寿命の短縮量が最も小さい S S D を選択する機能を有し、

最も予測寿命の短い S S D が、ストレージ装置内で定められた閾値より予測寿命が短くなった場合、前記最も予測寿命の短い S S D のデータと、前記予測寿命の短縮量が最も小さい S S D のデータとを入れ替える機能を有することを特徴とする、ストレージ制御方式。

40

[付記 1 3]

前記付記 1 1 に記載されたストレージ装置であって、

R A I D と呼ばれる冗長記録を行うディスクアレイ装置の機能を有し、

前記付記 1 1 のデータの入れ替えを行う際、データの入れ替え単位を、R A I D を構成するストライピンググループのサイズを基準にすることを特徴とした、ディスクアレイ装置。

[付記 1 4]

50

前記付記 13 に記載されたディスクアレイ装置であって、
データを入れ替える際に、RAID を構成する情報を同時に書き換えることにより、
データ入れ替えの前後で、RAID の再構築手順に変化がないように制御することを特徴とする、ディスクアレイ装置。

なお、付記 9 の形態及び付記 10 の形態は、付記 1 の形態と同様に、付記 2 の形態～付記 8 の形態に展開することが可能である。

【0070】

なお、引用した上記の特許文献等の各開示は、本書に引用をもって繰り込むものとする。本発明の全開示（請求の範囲を含む）の枠内において、さらにその基本的技術思想に基づいて、実施形態ないし実施例の変更・調整が可能である。また、本発明の全開示の枠内において種々の開示要素（各請求項の各要素、各実施形態ないし実施例の各要素、各図面の各要素等を含む）の多様な組み合わせ、ないし、選択が可能である。すなわち、本発明は、請求の範囲を含む全開示、技術的思想にしたがって当業者であればなし得るであろう各種変形、修正を含むことは勿論である。特に、本書に記載した数値範囲については、当該範囲内に含まれる任意の数値ないし小範囲が、別段の記載のない場合でも具体的に記載されているものと解釈されるべきである。

【符号の説明】

【0071】

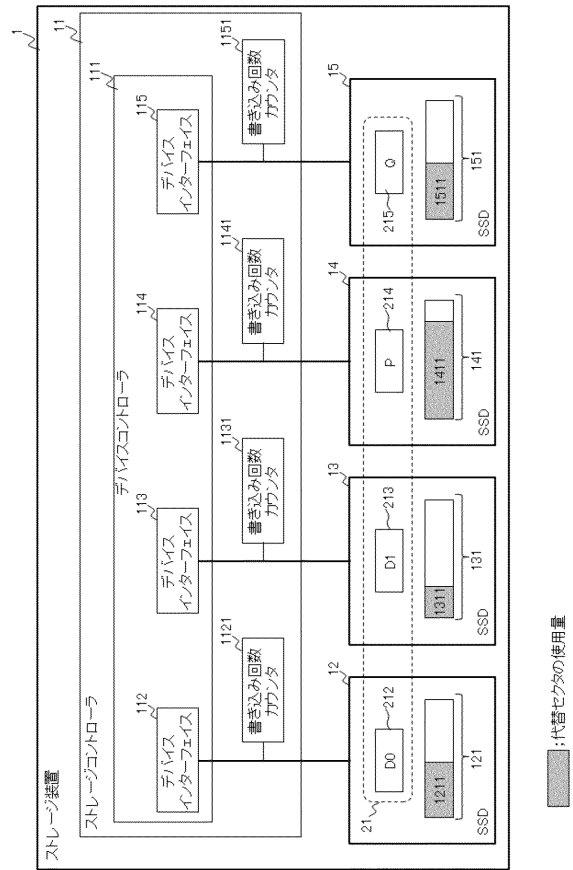
- 1、10 ストレージ装置
- 11 ストレージコントローラ
- 12～15 SSD
- 21 ストライピンググループ
- 101、201 寿命推定部
- 102、202 データ交換部
- 111 デバイスコントローラ
- 112～115 デバイスインターフェイス
- 121～151 代替セクタ
- 212～215 データ領域
- 1121～1151 書き込み回数カウンタ
- 1211～1511 代替セクタ使用済み領域

10

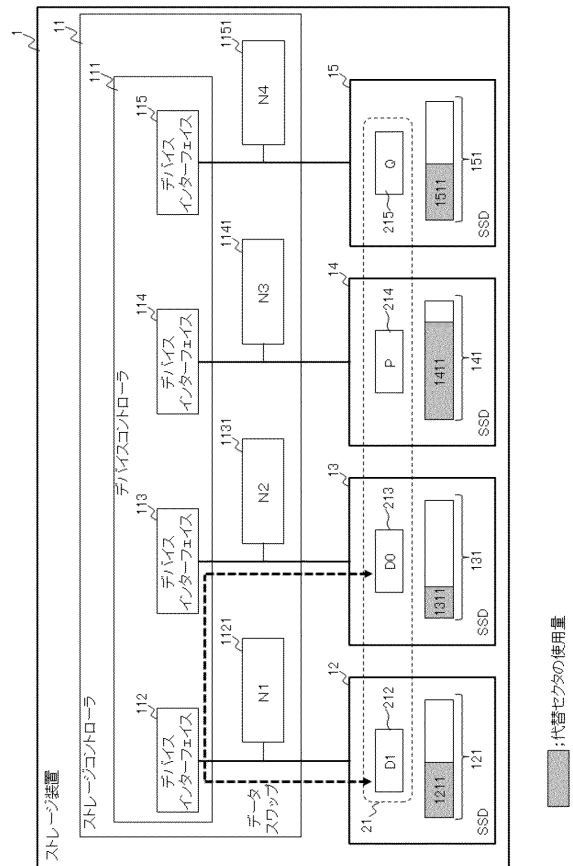
20

30

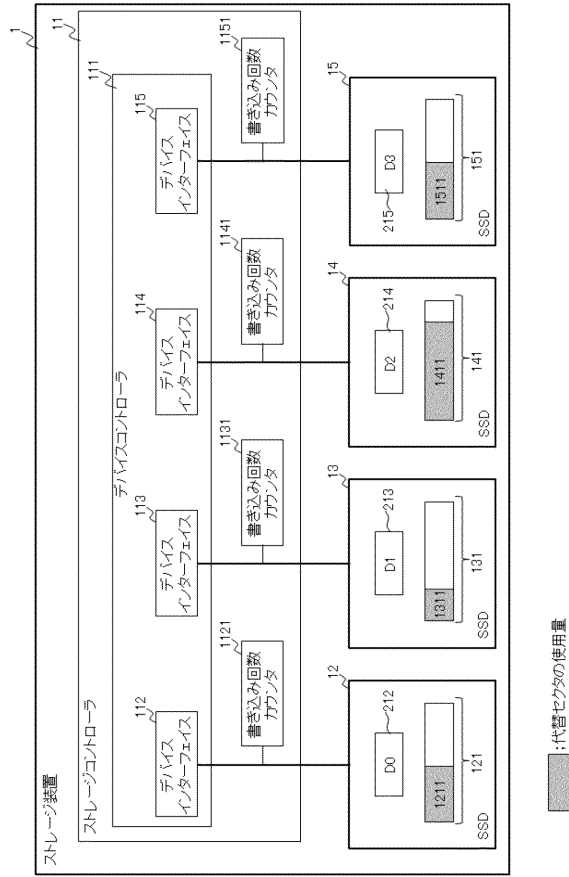
【 図 2 】



【 図 4 】



【図 5】



フロントページの続き

(51)Int.Cl. F I
G 0 6 F 3/06 5 4 0

審査官 篠塚 隆

(56)参考文献 特表2015-505078(JP,A)
国際公開第2015/181923(WO,A1)
特開2009-230660(JP,A)
特開2000-181805(JP,A)
特開2008-046923(JP,A)

(58)調査した分野(Int.Cl., DB名)
G 0 6 F 1 2 / 1 6
G 0 6 F 1 2 / 0 0
G 0 6 F 3 / 0 6
G 0 6 F 3 / 0 8
G 0 6 F 1 1 / 0 8 - 1 1 / 1 0
G 1 1 C 2 9 / 0 0